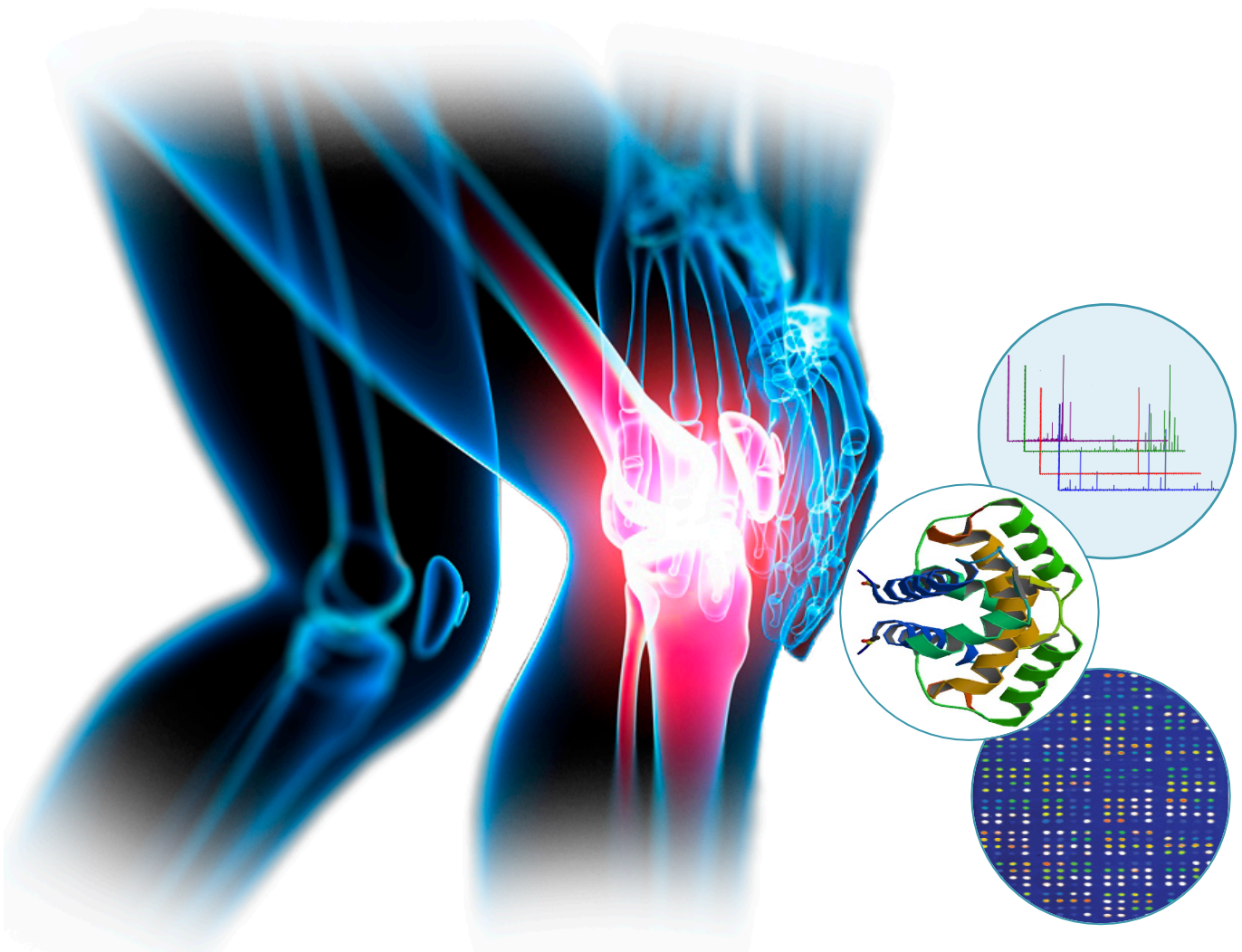


Departamento de Medicina
Facultad de Ciencias de la Salud
Universidad de A Coruña

Búsqueda de biomarcadores de artrosis mediante técnicas proteómicas



TESIS DOCTORAL

Lucía María Lourido Salas
A Coruña, 2015

Búsqueda de biomarcadores de artrosis mediante técnicas proteómicas

MEMORIA PRESENTADA PARA OPTAR AL GRADO DE DOCTOR INTERNACIONAL

Lucía María Lourido Salas

**TESIS DOCTORAL
2015**

Directores: Dr. Francisco Javier Blanco García y Dra. Cristina Ruiz Romero

Tutor: Dr. Francisco Javier de Toro Santos

Grupo de Proteómica-Servicio de Reumatología-Instituto de Investigación Biomédica

de A Coruña



Los directores de esta tesis doctoral, Dr. Francisco Javier Blanco García y Dra. Cristina Ruiz Romero,

CERTIFICAN:

Que el presente trabajo que lleva por título “Búsqueda de biomarcadores de artrosis mediante técnicas proteómicas” fue realizado por Lucía María Lourido Salas, bajo nuestra dirección en el Instituto de Investigación Biomédica de A Coruña, y reúne las condiciones de originalidad y rigor científico para optar al grado de doctor internacional en Ciencias de la Salud.

A Coruña, a 29 de Julio de 2015

Fdo. Dr. Francisco J. Blanco García

Fdo. Dra. Cristina Ruiz Romero

Dr. Francisco Javier de Toro Santos, tutor de esta tesis

CERTIFICA:

Que el presente trabajo que lleva por título “Búsqueda de biomarcadores de artrosis mediante técnicas proteómicas” fue realizado por Lucía María Lourido Salas, bajo mi tutela desde el Departamento de Medicina de la Universidad de A Coruña, y reúne las condiciones de originalidad y rigor científico para optar al grado de doctor internacional en Ciencias de la Salud.

A Coruña, a 29 de Julio de 2015

Fdo. Dr. Francisco de Toro Santos

SciLifeLab

Parte de la investigación presentada en esta tesis ha sido realizada bajo la supervisión del Dr. Peter Nilsson, Catedrático en Proteómica del *KTH-Royal Institute of Technology*, director de la plataforma de proteómica de afinidad en el Scilifelab y codirector del *Human Protein Atlas* en el SciLifeLab de Estocolmo, durante una estancia de investigación predoctoral en el año 2014.

Part of the research included in this thesis was been carried out under the supervision of Dr. Peter Nilsson, Professor in Proteomics at KTH-Royal Institute of Technology, Director of Affinity Proteomics at SciLifeLab and the Site Director of the Human Protein Atlas at SciLifeLab Stockholm, during a pre-doctoral fellowship in 2014.

A mis padres

AGRADECIMIENTOS

Llegado este momento, no puedo ni acercarme a describir lo afortunada que me siento por la gran cantidad de gente maravillosa que he conocido y con la que he colaborado durante estos cinco años. Todos habéis sido una enorme motivación para continuar trabajando y aprendiendo.

En primer lugar me gustaría agradecer a mis directores de tesis, el Dr. Francisco Blanco (Paco) y la Dra. Cristina Ruiz Romero (Cris) por la confianza que han depositado en mi para realizar este trabajo. Gracias a Paco por enseñarme siempre el lado práctico de la investigación que realizamos y la importancia que tiene ésta para los pacientes; y a Cris, gracias por tu continuo apoyo así como por tu contagiosa fuerza, vitalidad y positivismo ante todo.

A mis compañeras, amigas y familia postiza Bea, Carol y Vale. Gracias por los maravillosos momentos que hemos vivido juntas, por ser mi apoyo incondicional en el trabajo y por ayudarme con vuestra experiencia durante todos estos años. A Valentina, gracias por compartir conmigo las pequeñas pero muy satisfactorias y emocionantes recompensas que nos ha dado este trabajo. Para mí eres una crack y espero haber aprendido al menos un poquito de todo lo que me has enseñado. También me gustaría darle las gracias a Patri, porque siempre has estado ahí, sacándome una sonrisa y has sido un gran apoyo estos últimos meses. Y a Jesús, porque nunca olvidaré la ilusión que ponías por enseñarme cuando empecé en el lab. Ha sido un verdadero placer trabajar con todos vosotros. Gracias por aguantarme también en mis malos momentos... Os debo mucho.

A todo el grupo de genómica, en especial a Angelito y a Mary, es una suerte inmensa poder teneros como compañeros y amigos. Gracias por estar siempre ahí. A Sara por su paciencia a la hora de pedirle muestras y a Nacho, Estefa y Merce por sus ánimos en la recta final. A la Dra. Oreiro por las tardes que ha dedicado a revisar conmigo los datos clínicos de los pacientes. Y por supuesto a Juan, mi compi y amigo, por su infinita paciencia conmigo y por su disponibilidad incondicional para mis dudas, preguntas y peticiones fuera cual fuera el día y la hora. No habrá croissantitos de chocolate suficientes para pagar todo tu esfuerzo... Mil gracias por todo.

Me gustaría darle las gracias también a Romi por animarme un día a mandar el curriculum al INIBIC. Ese día cambió mi vida. Gracias.

También a Tami, MariaJo, Noa y Puri, por su paciencia con este pequeño desastre con las guardias y por su voluntad de ayudarme siempre.

No entiendo este trabajo si no es en grupo. Creo que lo hace más productivo y por supuesto, más divertido, y yo he encontrado aquí sin duda el mejor. Gracias a TODOS por hacer del INIBIC un perfecto ambiente de trabajo.

Me gustaría también agradecer a todo el lab11 del CIC de Salamanca por acogerme y aguantarme en las innumerables estancias allí, en especial a Manuel por confiar siempre en mí y darme todas las oportunidades posibles para aprender, y a María, por su inestimable ayuda. Gracias de todo corazón.

A Peter Nilsson del Scilifelab en Estocolmo por hacerme sentir parte de su laboratorio desde el primer momento. Gracias por tu cercanía, generosidad y confianza en mí. Me gustaría agradecer en especial a Burcu y a Frauke, mis compañeras de fatigas. Gracias por todo vuestro apoyo e implicación en el trabajo. Y en general, gracias a todo el inmenso PAPP group (Jochen, Ellin, Arash, David, Cecilia, Eni, Maria, Anna, Kimi....) por vuestra hospitalidad y amabilidad durante mi estancia. Ha sido genial trabajar con todos vosotros. Nunca hubiera imaginado que me pudiera sentir como en casa estando tan lejos de la mía. Mil gracias a todos. Y por supuesto, a mi familia de Chusiños, y a Tony. Non sei que faría sen vos...Siempre estaré eternamente agradecida por vuestro apoyo y cariño.

A toda mi familia, lo más importante y valioso de mi vida. A mis padres, por enseñarme siempre el poder del trabajo y que con eso llegaría a donde me propusiera. Gracias por apoyarme siempre y por animarme a seguir adelante. Gracias a vosotros ahora entiendo todo lo que unos padres son capaces de hacer por unos hijos. Nunca os podré agradecer todo lo que me habéis dado. Todo lo que he conseguido es gracias a vosotros y esta tesis es también un triunfo vuestro. A la mejor hermana del mundo, Paula, por su apoyo incondicional. Y a mis abuelas, las mujeres más luchadoras que conozco.

Y por supuesto, a mi Ángel, mi compañero de viaje. Después de vivir contigo toda esta etapa, sólo puedo decir que te admiro profundamente y que para mí eres el mejor. Formamos un gran equipo y contigo a mi lado, mi vida ha cobrado todo el sentido. Gracias por ser como eres, por tu apoyo, por entenderme y por cuidarme.

GRACIAS

RESUMEN

Resumen

La Artrosis (OA) es la enfermedad reumática más frecuente e invalidante a nivel mundial. Se caracteriza principalmente por una degradación progresiva del cartílago articular pero también por cambios en el hueso subcondral y la inflamación de la membrana sinovial. Pese a su elevada prevalencia, los métodos de diagnóstico actuales de la OA son poco sensibles y no existe tratamiento eficaz frente a la enfermedad. Estas limitaciones han originado un considerable interés por encontrar marcadores biológicos específicos que reflejen las variaciones cuantitativas y dinámicas de la remodelación de la articulación durante la artrosis.

Teniendo en cuenta que el proceso de OA implica diferentes tejidos y complejos procesos biológicos, la estrategia diagnóstica más prometedora podría ser la combinación del estudio de diferentes biomarcadores con el fin de facilitar el diagnóstico precoz, evaluar la progresión de la enfermedad y posibilitar la monitorización de terapias alternativas.

Así, en esta tesis se han empleado diferentes técnicas proteómicas de expresión diferencial basadas en espectrometría de masas y en *arrays* de proteínas para abordar el estudio de la búsqueda de potenciales biomarcadores de artrosis.

En el primer trabajo presentado, se analizó el perfil de degradación del cartílago artrósico con la finalidad de identificar potenciales marcadores tempranos y de progresión de la enfermedad. Para ello, se utilizaron muestras de diferentes zonas del cartílago artrósico humano y también cartílago sano. En este estudio, se identificó un panel de 76 proteínas que estaban diferencialmente liberadas por el cartílago artrósico al compararlo con el cartílago sano. A continuación, se validó por inmunodetección el potencial de dos proteínas como marcadores de progresión de la degradación del cartílago artrósico. En el segundo trabajo, se utilizaron muestras de líquido sinovial de pacientes con OA y pacientes con artritis reumatoide (AR), con el objetivo de caracterizar el perfil proteico del líquido sinovial en la OA. Posteriormente, se validaron por inmunodetección los resultados obtenidos para 4 de las proteínas cuantificadas por espectrometría de masas.

A continuación, utilizamos *arrays* de proteínas para buscar potenciales biomarcadores séricos de la OA.

Así, en el tercer trabajo, se utilizaron muestras de suero de pacientes con artrosis, artritis reumatoide y controles sanos con el objetivo de identificar autoanticuerpos con potencial biomarcador de la artrosis empleando *arrays* de

antígenos. En este trabajo, detectamos que los niveles de inmunorreactividad frente a 7 antígenos permitían distinguir los pacientes artrósicos de los individuos sanos mientras que los niveles de inmunorreactividad frente a otros 7 antígenos, permitían distinguir entre los pacientes con OA y con AR.

Por último, se construyeron *arrays* de anticuerpos en suspensión con el fin de identificar, en un gran número de muestras, un panel de proteínas en el suero con potencial biomarcador para el diagnóstico de la artrosis. Esta metodología se utilizó para analizar muestras de suero de pacientes con diferentes enfermedades reumáticas (artrosis, artritis reumatoide y artritis psoriásica) e individuos sanos. Los resultados obtenidos proporcionaron un perfil proteico con potencial biomarcador para el diagnóstico de la artrosis de rodilla.

Resumo

A artrose (OA) é a enfermidade reumática mais frecuente e invalidante a nivel mundial. Caracterízase principalmente por unha degradación progresiva da cartilaxe articular pero tamén por cambios no óso subcondral e a inflamación da membrana sinovial. Pese a súa prevalencia, os métodos actuais de diagnose da OA son pouco sensibles e non existe tratamento eficaz fronte a enfermidade. Estas limitacións orixinaron un considerable interés por encontrar marcadores biolóxicos específicos que reflexen variacións cuantitativas e dinámicas da remodelación da articulación durante a OA.

Tendo en conta que o proceso artrósico implica a diferentes texidos e complexos procesos biolóxicos, a estratexia diagnóstica mais prometedora podería ser a combinación do estudio de diferentes biomarcadores coa finalidade de facilitar a diagnose temperá, avaliar a progresión da enfermidade e posibilitar a monitorización de terapias alternativas.

Así, nesta tese se empregaron diferentes técnicas proteómicas de expresión diferencial basadas en espectrometría de masas e en *arrays* de proteínas para abordar o estudo da búsqueda de biomarcadores de artrose.

No primeiro traballo presentado, se analizou o perfil de degradación proteica da cartilaxe artrósica co obxectivo de indentificar potenciais marcadores temperáns e de progresión da enfermidade. Para isto, se utilizaron mostras de diferentes zonas da cartilaxe artrósica humana e tamén cartilaxe sana. En este estudio, se identificou un panel de 76 proteínas que estaban diferencialmente liberadas pola cartilaxe artrósica en comparación coa cartilaxe sana. A continuación, validamos por inmunodetección o potencial de dúas proteínas como marcadores da degradación progresiva da cartilaxe artrósica. No segundo traballo, se empregaron mostras de líquido sinovial de pacientes con artrose e pacientes con artritis reumatoide (AR), co obxectivo de caracterizar o perfil proteico do líquido sinovial na OA. Posteriormente, se validaron por inmunodetección os resultados obtidos para dúas das proteínas cuantificadas por espectrometría de masas.

A continuación, utilizamos *arrays* de proteínas para buscar potenciais biomarcadores séricos da OA.

Así, no terceiro traballo, se utilizaron mostras de soro de pacientes con OA, pacientes con AR e individuos sanos para buscar autoanticorpos con potencial biomarcador da artrose empregando *arrays* de antíxenos. Neste traballo, detectamos que os niveis de inmureactividade fronte a 7 antíxenos permitían distinguir entre

pacientes con OA e individuos sanos, mentres que os niveis de inmureactividade fronte a outros 7 antíxenos permitían distinguir entre pacientes con OA e AR.

Por último, se construíron *arrays* de anticorpos en suspensión coa finalidade de identificar, nun gran número de mostras, un panel de proteínas no soro con potencial biomarcador para a diagnose da artrose. Así, se analizaron mostras de soro de doentes con diferentes enfermidades reumáticas (artrose, artritis reumatoide e artritis psoriásica) e individuos sanos. Os resultados obtidos proporcionaron un perfil proteico con potencial biomarcador para a diagnose da artrose de xeonllo.

Abstract

Osteoarthritis (OA) is the most common and disabling rheumatic disease worldwide. It is characterized by the progressive loss of cartilage, subchondral bone remodelling and synovial inflammation. Despite its high prevalence, current diagnosis methods are poor sensitive and there is no efficient treatment for the disease. These limitations have prompted a considerable interest in finding specific biological markers to reflect quantitative and dynamic joint changes that occur in OA.

Taking in account that OA process involves different tissues and complex biological process, the most promising strategy for early diagnosis and monitoring of the disease, could be the combination of the study of different biomarkers.

Therefore, in this thesis, different proteomic technologies based on mass spectrometry and protein arrays were employed to address the research of potential OA biomarkers.

Firstly, we aimed to study the protein profile of OA cartilage degradation to find potential early and progression OA biomarkers. For this, we used OA and healthy human cartilage explants. A panel of 76 differentially secreted proteins from OA cartilage compared to normal cartilage was found. Then, two proteins with potential value as progression biomarkers were validated. Next, we used synovial fluid from OA patients and rheumatoid arthritis (RA) patients to find a characteristic protein profile of OA process using mass spectrometry. Four modulated proteins were validated.

Then we used protein arrays to find potential serum OA biomarkers. Therefore, antigen arrays were used to search for serum autoantibodies in OA comparing the autoantibody profile from OA patients, healthy donors as well as RA patients. We observed immunoreactivity levels towards 7 antigens allowed to distinguish between OA patients and healthy controls, whereas the immunoreactivity towards other 7 antigens enables to discriminate between OA and RA patients.

Finally, we used antibody suspension bead arrays to identify, in a large set of samples, a panel of serum proteins with potential value for OA diagnosis. This methodology was used to analyse serum samples from patients with different rheumatic diseases (OA, RA and psoriatic arthritis) and healthy donors. The results obtained provide a serum protein profile with a great potential for knee OA diagnosis.

INDICE

| | |
|--|-----------|
| I. INTRODUCCIÓN..... | 1 |
| 1. La articulación humana | 3 |
| 1.1 Estructura y composición del cartílago articular..... | 4 |
| 2. La artrosis..... | 5 |
| 2.1. Definición..... | 5 |
| 2.2. Fisiopatología de la artrosis..... | 6 |
| 2.3. Factores de riesgo..... | 9 |
| 2.4. Epidemiología e impacto socioeconómico..... | 12 |
| 2.5. Diagnóstico y tratamiento..... | 14 |
| 3. La proteómica..... | 17 |
| 3.1. Definición y aplicaciones | 17 |
| 3.2. Biomarcadores proteicos en la artrosis..... | 19 |
| 4. Estrategias proteómicas para la búsqueda de biomarcadores..... | 22 |
| 4.1. Técnicas basadas en espectrometría de masas (MS)..... | 23 |
| 4.1.1. Digestión de proteínas..... | 24 |
| 4.1.2. Separación de proteínas..... | 24 |
| 4.1.2.1. Métodos de separación basados en gel..... | 25 |
| 4.1.2.2. Métodos de separación basados en cromatografía..... | 27 |
| 4.1.3. Espectrometría de masas..... | 29 |
| 4.1.3.1. Fuentes de ionización..... | 30 |
| 4.1.3.2. Analizadores de masas..... | 31 |
| 4.1.3.3. Identificación de proteínas mediante MS..... | 35 |
| 4.1.4. Cuantificación de proteínas mediante MS..... | 35 |
| 4.1.4.1. Cuantificación basada en gel | 35 |
| 4.1.4.2. Cuantificación basada en MS | 37 |
| 4.2. Técnicas basadas en <i>arrays</i> de proteínas..... | 43 |
| 4.2.1. Formatos..... | 44 |
| 4.2.1.1. <i>Arrays</i> planos..... | 44 |
| 4.2.1.2. <i>Arrays</i> en esfera..... | 47 |
| 4.2.2. Clasificación..... | 48 |
| 4.2.2.1. <i>Arrays</i> de anticuerpos..... | 48 |
| 4.2.2.2. <i>Arrays</i> de antígenos..... | 54 |

| | |
|---|------------|
| 4.2.2.3. Arrays de fase reversa..... | 62 |
| II. OBJETIVOS..... | 63 |
| III. CAPÍTULOS..... | 67 |
| <u>Capítulo I.</u> “Quantitative proteomic profiling of human articular cartilage degradation in osteoarthritis”..... | 69 |
| <u>Capítulo II.</u> “Differential protein profiling of synovial fluid from rheumatoid arthritis and osteoarthritis patients using LC–MALDI TOF/TOF”..... | 85 |
| <u>Capítulo III.</u> “Analysis of autoantibody profiles in osteoarthritis using comprehensive protein array concepts”..... | 97 |
| <u>Capítulo IV.</u> “Discovery of circulating proteins associated to knee radiographic osteoarthritis”..... | 113 |
| IV. DISCUSIÓN..... | 129 |
| Discusión capítulo I..... | 131 |
| Discusión capítulo II..... | 135 |
| Discusión capítulo III..... | 139 |
| Discusión capítulo IV..... | 143 |
| Discusión general..... | 147 |
| V. CONCLUSIONES..... | 153 |
| VI. CONCLUSIONS..... | 157 |
| VII. BIBLIOGRAFÍA..... | 161 |
| VIII. ANEXOS..... | 183 |
| Anexo I: Material suplementario capítulo I..... | 185 |
| Anexo II: Material suplementario capítulo II..... | 188 |
| Anexo III: Material suplementario capítulo III..... | 189 |
| Anexo IV: Material suplementario capítulo IV..... | 193 |
| Anexo V: Relación de artículos no directamente relacionados con los objetivos de esta tesis en los que he participado durante mi etapa como estudiante predoctoral..... | 197 |

Índice de abreviaturas

| | |
|---------------|---|
| 2-DE | Electroforesis bidimensional (2-Dimensional Electrophoresis) |
| ACN | Acetonitrilo (Acetonitrile) |
| ACR | Colegio americano de reumatología (American Colleague of Rheumatology) |
| ADAMT | Adamalisina proteasa con motivos de trombospondina (A Disintegrin And Metalloproteinase with Thrombospondin Motifs) |
| BMPs | Proteínas morfogenéticas del hueso (Bone Morphogenetic Proteins) |
| BSA | Suero bovino fetal (Bovine Serum Albumin) |
| CID | Disociación inducida por colisión (Collision Induced Dissociation) |
| DAMPs | Patrones moleculares asociados al daño (Damage Associated Molecular Patters) |
| DAPA | Micromatriz de proteínas a partir de DNA (DNA to Protein Array) |
| DIGE | Electroforesis diferencial en gel (Difference In Gel Electrophoresis) |
| DNA | Ácido deoxirribonucleico (DeoxyriboNucleic Acid) |
| DTT | Ditiotreitol (Dithiothreitol) |
| ECM | Matriz extracelular (Extracellular Matrix) |
| ELISA | Ensayo por inmunoabsorción ligado a enzimas (Enzyme-linked Immunosorbent Assay) |
| ESI | Ionización por electrospray (Electrospray Ionization) |
| Fab | Fracción de unión al antígeno (Fragment antigen-binding) |
| Fc | Fragmento constante (constant Fragment) |
| FT-ICR | Espectrómetro de masas de resonancia ciclotrónica de iones con transformada de Fourier (Fourier Transform Ion Cyclotron Resonance mass spectrometer) |
| GST | Glutathion-S-transferasa (Glutathion-S-Transferase) |
| HPA | Atlas del proteoma humano (Human Protein Atlas) |
| HPLC | Cromatografía líquida de alta resolución (High Performance Liquid Chromatography) |
| HUPO | Organización del proteoma humano (Human Proteome Organization) |
| IAA | Iodoacetamida (Iodoacetamide) |
| IEF | Isoelectroenfoque (Isoelectric Focusing) |
| IgG | Inmunoglobulina G (Inmunoglobulin G) |
| IGF | Factor de crecimiento insulínico (Insuline Growth Factor) |
| IL | Interleuquina (Interleukin) |
| IT | Trampa iónica (Ion Trap) |
| iTRAQ | Marcaje isotópico para cuantificación relativa y absoluta (Isotope Tagged for Relative and Absolute Quantification) |
| K/L | Kellgren-Lawrence (Kellgren-Lawrence) |
| KDa | Kilodalton (KiloDalton) |
| KV | Kilovoltio (Kilovolt) |

| | |
|--------------|---|
| LC | Cromatografía líquida (<i>Liquid Chromatography</i>) |
| LIT | Trampa de iones lineal (<i>Linear Ion Trap</i>) |
| LTQ | Trampa lineal cuadrupolar (<i>Linear Trap Quadrupole</i>) |
| m/z | Masa/carga (<i>Mass/charge</i>) |
| MALDI | Desorción/ionización mediante láser asistida por matriz (<i>Matrix-Assisted Laser Desorption Ionization</i>) |
| MFI | Intensidad de Fluorescencia media (<i>Median Fluorescence Intensity</i>) |
| mL | Mililitro (<i>Milliliter</i>) |
| MMP | Metaloproteasa de la matriz (<i>Matrix Metaloproteases</i>) |
| MRM | Monitorización de reacciones múltiple (<i>Multiple Reaction Monitoring</i>) |
| mRNA | Ácido ribonucleico mensajero (<i>messenger Ribonucleic Acid</i>) |
| MS | Espectrometría de masas (<i>Mass Spectrometry</i>) |
| MS/MS | Espectrometría de masas en tandem (<i>Tandem Mass Spectrometry</i>) |
| NAPPA | Micromatriz de proteínas programada por ácidos nucleicos (<i>Nucleic Acid Programmable Protein Array</i>) |
| ng | Nanogramo (<i>nanogram</i>) |
| NIH | Instituto nacional de salud (<i>National Institute of Health</i>) |
| nm | Nanometros (<i>nanometers</i>) |
| NO | Óxido nítrico (<i>Nitric Oxide</i>) |
| OA | Artrosis (<i>Osteoarthritis</i>) |
| OARSI | Sociedad Internacional de la Investigación en Artrosis (<i>Osteoarthritis Research Society International</i>) |
| ORF | Región abierta de lectura (<i>Open Reading Frame</i>) |
| PAGE | Electroforesis en gel de poliacrilamida (<i>Polyacrilamide Gel Electrophoresis</i>) |
| PCR | Reacción en cadena de la polimerasa (<i>Polymerase Chain Reaction</i>) |
| PE | Ficoeritrina (<i>Phycoerythrin</i>) |
| pg | Picogramo (<i>picogram</i>) |
| PGs | Prostaglandinas (<i>Prostaglandins</i>) |
| pI | Punto isoeléctrico (<i>Isoelectric Point</i>) |
| PISA | Micromatriz de proteínas <i>in situ</i> (<i>Protein in situ Array</i>) |
| pL | picolitro (<i>Picoliter</i>) |
| PrEST | Epítipo proteico con una etiqueta característica (<i>Protein Epitope Signature Tag</i>) |
| PUCA | Micromatriz de proteínas por captura de puomicina (<i>Puromycin Capture Protein Array</i>) |
| Q | Cuadrupolo (<i>Quadrupole</i>) |
| RF | Radiofrecuencia (<i>Radiofrequency</i>) |
| RP | Fase reversa (<i>Reverse Phase</i>) |
| ScFv | Cadena sencilla de la fracción variable (<i>single chain of variable Fraction</i>) |
| SDS | Dodecil sulfato sódico (<i>Sodium Dodecil Sulfate</i>) |

| | |
|-------------------------------|---|
| SER | Sociedad Española de Reumatología |
| SILAC | Marcaje isotópico estable con aminoácidos en cultivo celular (<i>Stable Isotope Labelling with Amino acids in Cell culture</i>) |
| ssDNA | Ácido deoxiribonucleico monocatenario (s ingle- s tranded D eoxyribonucleic A cid) |
| TNFα | Factor de necrosis tumoral-alfa (<i>Tumor Necrosis Factor alpha</i>) |
| TOF | Tiempo de vuelo (<i>Time of Flight</i>) |
| μg | Microgramo (M icrogramo) |
| μm | Micrometro (M icrometer) |

Índice de figuras

| | | |
|-------------------|--|----|
| Figura 1. | Estructura de la articulación humana. | 3 |
| Figura 2. | Estructura del cartílago articular hialino. | 4 |
| Figura 3. | Representación de los procesos degenerativos más importantes que tienen lugar durante la artrosis. | 6 |
| Figura 4. | Moléculas que intervienen en el desequilibrio hacia los procesos catabólicos que se producen en la artrosis. | 7 |
| Figura 5. | Representación de las principales implicaciones del sistema inmune en la artrosis. | 8 |
| Figura 6. | Incidencia de la artrosis de mano, rodilla y cadera por sexo y edad de los pacientes. | 13 |
| Figura 7. | Representación de los costes económicos que supone la artrosis por paciente en España. | 14 |
| Figura 8. | Imagen que representa la complejidad del proteoma humano en comparación con el genoma. | 18 |
| Figura 9. | Principales aplicaciones de la proteómica. | 19 |
| Figura 10. | Flujo de trabajo habitual en la búsqueda de biomarcadores proteicos. | 22 |
| Figura 11. | Estrategias proteómicas basadas en espectrometría de masas empleadas en la búsqueda de biomarcadores. | 23 |
| Figura 12. | Proceso de separación de proteínas mediante la técnica de electroforesis bidimensional. | 26 |
| Figura 13. | Esquema representativo de la técnica de electroforesis diferencial en gel (DIGE) para la separación y cuantificación de proteínas. | 27 |
| Figura 14. | Componentes básicos de un espectrómetro de masas. | 29 |
| Figura 15. | Principales fuentes de ionización empleadas en espectrometría de masas. | 31 |
| Figura 16. | Tipos de analizadores de masas más utilizados en proteómica. | 32 |
| Figura 17. | Esquema de las principales estrategias para el marcaje de péptidos y proteínas mediante isótopos estables. | 38 |
| Figura 18. | Representación de los diferentes componentes del reactivo TRAQ. | 40 |
| Figura 19. | Flujo de trabajo en un experimento iTRAQ. | 41 |

| | | |
|-------------------|---|----|
| Figura 20. | Representación de la cuantificación sin marcaje basada en conteo espectral. | 42 |
| Figura 21. | Estructura básica de un anticuerpo de tipo inmunoglobulina G. | 49 |
| Figura 22. | Tipos de epítomos que puede reconocer un anticuerpo. | 50 |
| Figura 23. | Formatos de <i>arrays</i> de anticuerpos. | 51 |
| Figura 24. | Flujo de trabajo del HPA. | 54 |
| Figura 25. | Representación esquemática de los métodos para producir <i>arrays</i> de proteínas <i>in situ</i> . | 61 |

Índice de tablas

| | | |
|-----------------|--|----|
| Tabla 1. | Criterios radiológicos y clínicos para la artrosis de mano, rodilla y cadera. | 15 |
| Tabla 2. | Escala radiográfica Kellgren-Lawrence para la clasificación de pacientes artrósicos. | 16 |
| Tabla 3. | Biomarcadores de artrosis en fases de investigación. | 21 |
| Tabla 4. | Principales características de los métodos de tinción de proteínas más frecuentes. | 36 |

I. INTRODUCCIÓN

1. La articulación humana

Una articulación se define como el punto de unión de dos o más huesos y es el órgano encargado de que el movimiento se produzca sin dolor. En función de la naturaleza del material que separa los huesos y de su movimiento podemos distinguir: articulaciones fibrosas (escaso movimiento), articulaciones cartilagosas (semimóviles) y articulaciones sinoviales o diartroidales (gran movilidad).

Las articulaciones diartroidales (Figura 1) son las más comunes en los humanos. Estas articulaciones se encuentran rodeadas de una cápsula articular. La cápsula articular es una prolongación del hueso que cubre la cavidad articular cuya función es proteger la articulación. Esta cápsula está constituida por una capa externa de tejido fibroso y una capa interna que es la membrana sinovial, formada por un tejido conjuntivo vascularizado que secreta el líquido sinovial. El líquido sinovial es un líquido transparente formado por un filtrado del plasma y ácido hialurónico. Su función es nutrir al cartílago y mantener la lubricación de la articulación.

Uno de los componentes más importantes de la articulación es el cartílago hialino. El cartílago articular hialino es un tejido que cubre el hueso y que amortigua las fuerzas que ejercen presión sobre las articulaciones. Por su parte, los ligamentos y tendones permiten el deslizamiento de las superficies articulares y ayudan a estabilizar el movimiento [1, 2].

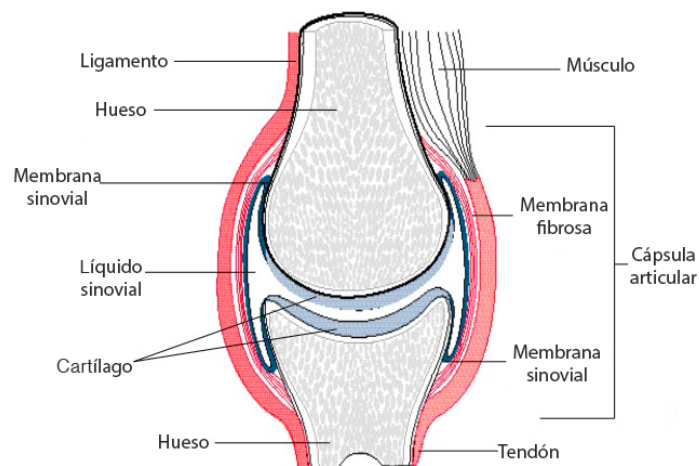


Figura 1. Estructura de la articulación diartroïdal humana.

1.1. Estructura y composición del cartílago articular

El cartílago articular es un tejido avascular, aneuronal y alifático que cubre las articulaciones, favorece el movimiento de los huesos y amortigua la carga. Este tejido está constituido principalmente por una matriz extracelular (ECM) (98%) que es sintetizada por los condrocitos (2%), las únicas células residentes en el cartílago maduro. La ECM está compuesta por agua (70-80%), colágenos VI, IX, XI, XII y XIV (10-20%), así como por otras proteínas como los proteoglicanos y el ácido hialurónico (5-10%), que son macromoléculas altamente hidrofílicas que se encuentran absorbidas en el interior de la red de colágenos y son las responsables de aportar elasticidad al tejido [2].

La composición y organización del cartílago varía desde la superficie a la profundidad del tejido (Figura 2).

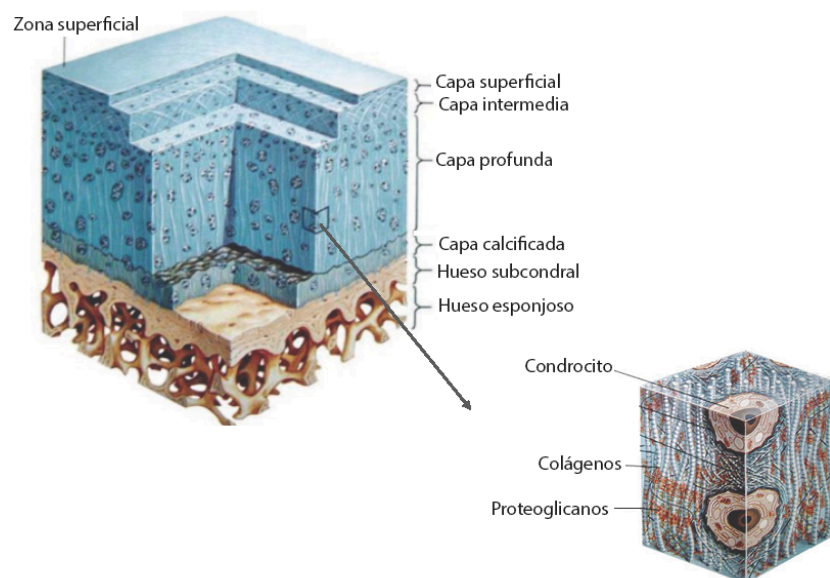


Figura 2. Estructura del cartílago articular hialino mostrando las cuatro capas que lo componen: la capa superficial, la capa intermedia, la capa profunda y la capa calcificada. Imagen modificada de Elsevier.Inc.Netterimagenes.com

En la capa superficial del cartílago articular, los condrocitos son pequeños y aplanados. Esta capa tiene una mayor cantidad de agua, las fibras de colágeno tienen un menor diámetro y se disponen paralelas a la superficie articular atrapando a las grandes moléculas de proteoglicanos y protegiendo al cartílago de la abrasión.

En las capas profundas las fibras están más separadas, su diámetro es mayor y se encuentran alineadas de forma vertical [3].

El cartílago calcificado es una capa delgada que separa el hueso subcondral del cartílago articular y está constituida por condrocitos pequeños rodeados de cartílago calcificado [4].

Una estructura íntimamente ligada al cartílago es el hueso subcondral. El hueso subcondral comprende el tejido subarticular mineralizado que se extiende desde el frente de calcificación hasta el inicio de la médula ósea. Sus funciones principales consisten en dar soporte al cartílago articular suprayacente, distribuir la carga mecánica así como absorber la tensión de los impactos mecánicos continuos y nutrir las capas profundas del cartílago hialino, especialmente en el período de crecimiento [5].

2. La artrosis

2.1. Definición

La artrosis (OA) es la patología reumática más frecuente en los países desarrollados. La OA se ha definido clásicamente como una enfermedad articular degenerativa caracterizada únicamente por la degradación progresiva del cartílago. Sin embargo, hoy se sabe que la OA es una enfermedad compleja que se origina como consecuencia de un grupo heterogéneo de procesos que provocan una afectación global de todos los componentes de la articulación [6].

La *Osteoarthritis Research Society International* (OARSI) ha publicado recientemente una definición actualizada de la artrosis de manera que ésta se define como una enfermedad que puede afectar a cualquier articulación y que está caracterizada por el estrés celular y la degradación de la matriz extracelular que se inicia por micro o macro lesiones en la articulación que el tejido no es capaz de revertir activando las vías de reparación y la respuesta del sistema inmune. La enfermedad se manifiesta primero como un desequilibrio metabólico del tejido articular, seguido por cambios fisiológicos y anatómicos como son la degradación del cartílago, la remodelación del hueso, la formación de osteofitos, la inflamación de la articulación y la pérdida de la función articular normal [7].

En la Figura 3 se muestran los cambios que tienen lugar en la articulación de la rodilla durante el proceso artrósico.

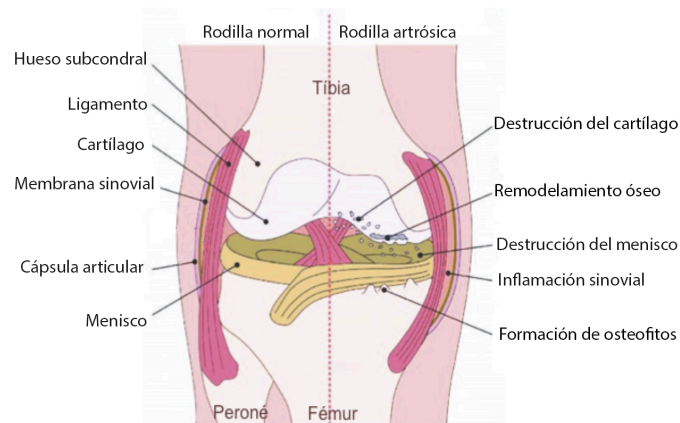


Figura 3. Representación de los principales componentes de la articulación de rodilla y los cambios más importantes que se producen durante la artrosis.

2.2. Fisiopatología de la artrosis

La artrosis se entiende como la insuficiencia global de toda la articulación, donde participan todas las estructuras articulares, teniendo especial importancia la interacción entre el cartílago articular, el hueso subcondral y la membrana sinovial [8].

Aunque el cartílago es un tejido con baja capacidad de renovación, en un cartílago fisiológicamente normal, los condrocitos son las células encargadas de producir todas las moléculas que intervienen en los procesos de síntesis de la matriz (anabólicos) y en los procesos de reparación de aquellas moléculas que están dañadas (catabólicos).

Sin embargo, en la artrosis se produce una disfunción en la homeostasis del metabolismo normal del cartílago (Figura 4). Los procesos catabólicos de degradación del cartílago, que incluyen la activación de factores como el factor de necrosis tumoral alpha ($TNF\alpha$) o la interleucina- 1β ($IL1\beta$), aumentan la expresión de proteasas como las colagenasas (MMP1, MMP13, MMP8), las agrecanasas (MMP3, MMP14, ADAMTS-1, 4,5) o las gelatinasas (MMP2, MMP9) que degradan los componentes de la matriz extracelular del cartílago. En la artrosis, este catabolismo no puede ser compensado por los factores anabólicos del tejido como el factor de crecimiento insulínico (IGF) o las proteínas morfogenéticas del hueso (BMPs), produciéndose así un desbalance en el metabolismo del cartílago [9].

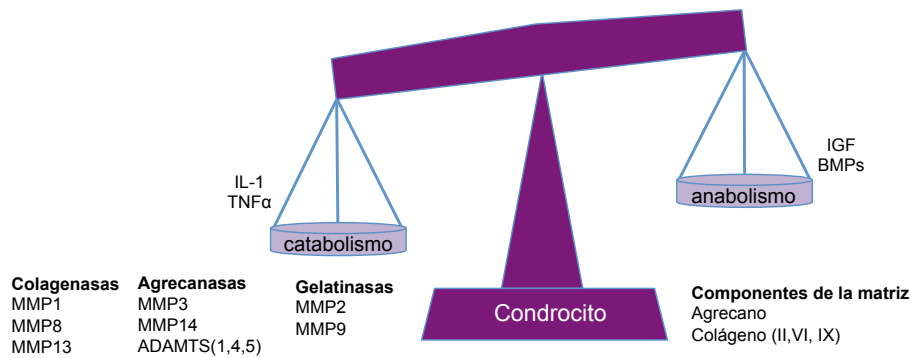


Figura 4. Moléculas que regulan el metabolismo del cartílago articular. Los procesos catabólicos promueven la degradación de la matriz extracelular, y los anabólicos, la síntesis de la matriz extracelular.

Una de las causas más simples para explicar este desequilibrio es la apoptosis o pérdida de viabilidad de los condrocitos. Según esta teoría, la reducción en el número de condrocitos limitaría la capacidad del tejido para mantener la homeostasis del cartílago y reparar el daño durante el proceso artrósico [9]. La otra hipótesis es la de la senescencia celular. Esta teoría postula que los condrocitos sufren un envejecimiento prematuro como consecuencia del estrés oxidativo al que están sometidos limitando su capacidad para realizar sus funciones normales y llevando en último término a la destrucción del tejido [9].

Una tercera explicación sería el fallo en la capacidad de las células madre mesenquimales presentes en el cartílago para diferenciarse a condrocitos con el fin de restaurar la pérdida de morfología y funciones del cartílago [10].

A nivel histológico, los cambios degenerativos del cartílago provocan el emblandecimiento del tejido, fibrilación de las capas superficiales del cartílago y disminución del espesor del cartílago [11].

Otra de las estructuras implicadas en la artrosis es el hueso subcondral. Este hueso en los pacientes artrósicos es esclerótico, denso y presenta un elevado remodelado óseo [12]. Su papel en la artrosis continúa siendo enigmático puesto que aún no se ha definido si los cambios que suceden en este tejido son posteriores o preceden a la degradación del cartílago [13-15]. Esta última teoría se basa en que las células del hueso (osteoblastos) de los pacientes artrósicos tienen un fenotipo más catabólico que los de los individuos normales. Los factores catabólicos generados por estos osteoblastos actuarían sobre los condrocitos a través de una invasión vascular en el cartílago de forma que esta lesión no podría ser reparada por los condrocitos a

través de sus factores antiangiogénicos y se perpetuaría la progresión de la enfermedad [16, 17].

Por otro lado, en los últimos años varios estudios han revelado el papel de la respuesta inmune en la fisiopatología de la artrosis [18-23] (Figura 5). Aunque clásicamente la artrosis se ha considerado una enfermedad no inflamatoria, varios estudios epidemiológicos han demostrado que la inflamación de la membrana sinovial conocida como sinovitis, síntoma originalmente relegado únicamente a los pacientes con artritis reumatoide, es un fenómeno que también ocurre en el 50% de los pacientes con artrosis y frecuentemente en las fases más tempranas de la enfermedad [24].

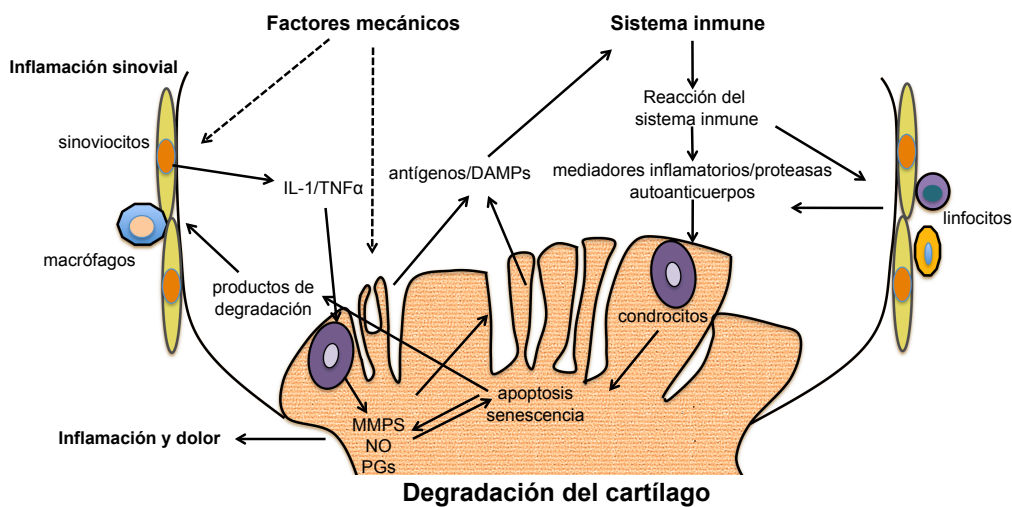


Figura 5. Figura esquemática de las implicaciones del sistema inmune en la OA.

Por un lado, se produce una respuesta inflamatoria humoral en la que las citoquinas pro-inflamatorias liberadas por la membrana sinovial inflamada (IL1, IL6, IL17, IL18, TNF α , prostaglandinas y óxido nítrico) provocan un aumento de los procesos catabólicos en el cartílago, produciendo un aumento en la degradación de la matriz y una progresión en los síntomas de la artrosis [25].

Otro efector crucial de la respuesta inmune es la activación de una cascada de proteínas fuertemente reguladas que constituyen el sistema del complemento. La activación del complemento y su modulación han sido también descritos a nivel del cartílago, líquido sinovial y suero de pacientes artrósicos [26].

Por otro lado, varios estudios también han observado infiltrados de células del sistema inmune (linfocitos T, linfocitos B, células dendríticas y macrófagos) en la membrana sinovial inflamada de los pacientes con OA. La presencia de estas células sugiere un papel de la respuesta inmune adaptativa en el proceso artrósico [27-29].

Además, el análisis de los receptores de las poblaciones oligoclonales de células T y células B en la membrana sinovial de pacientes con OA sugiere que estas células se activan en respuesta a antígenos específicos dentro del compartimento sinovial [30, 31]. En este sentido, diversos autores han propuesto la hipótesis de que los fragmentos liberados de la matriz extracelular como consecuencia de la degradación del cartílago y de las alteraciones en el hueso subcondral se comportan como patrones moleculares asociados al daño (DAMPs) en la articulación [32] y que podrían actuar como autoantígenos que desencadenen la respuesta inmune en la artrosis, activando al sistema del complemento y perpetuando los procesos catabólicos y la inflamación en la articulación [33].

Basándose en esta teoría, a lo largo de los últimos años varios componentes del cartílago se han propuesto como responsables de la expansión oligoclonal de las células del sistema inmune que se han observado en pacientes artrósicos. En estos estudios se han encontrado anticuerpos en el cartílago, líquido sinovial y el suero de pacientes con artrosis dirigidos frente a proteínas derivadas del cartílago como *triosephosphate isomerase* (TPIS) [34], *cartilage intermediate layer protein-1* (CILP-1) [35], *chitinase-3-like protein 2* (YKL-39) [36], *osteopontin* [37] y *fibulin-4* [38]. Los autores de estos trabajos sugieren que estos autoanticuerpos podrían tener relevancia para el diagnóstico de la artrosis y, a su vez, que la supresión de estos antígenos podrían tener un papel como diana terapéutica en la enfermedad.

Por lo tanto, todas estas evidencias sugieren un papel de todos los aspectos de la respuesta inmune en la patología de la artrosis, que junto con el desequilibrio en el metabolismo del condrocito y otros factores que afectan al desarrollo de la artrosis (genética, sexo, edad, obesidad, efectos mecánicos, etc) hacen de esta enfermedad una de las patologías más complejas.

2.3. Factores de riesgo

La artrosis es una enfermedad multifactorial, con aspectos comunes y diferenciales entre los individuos afectados que conducen a manifestaciones clínicas similares y a pronósticos variables en cada paciente.

Se han descrito diversos factores que afectan a la prevalencia y progresión de la artrosis. Estos factores se pueden dividir en sistémicos y mecánicos [39, 40].

Factores de riesgo sistémicos:

Edad.- Aunque la artrosis no es una consecuencia del envejecimiento normal del individuo, desde hace tiempo se sabe que existe una fuerte asociación entre la artrosis y la edad [41]. Probablemente esto es debido a la relación que existe entre los cambios que se producen en los compartimentos de la articulación (anatómicos, estructurales y bioquímicos) y la pérdida de la capacidad biomecánica del individuo para soportar las cargas. Por ejemplo, con la edad se pierde masa muscular, por lo que toda la carga se deposita en la articulación y el hueso, sufriendo éstos un mayor desgaste [42-45].

Género.- Existen varios estudios en los que se asocia la incidencia y prevalencia de la artrosis con el sexo [46, 47]. Concretamente, se sabe que las mujeres tienen un riesgo mayor que los hombres de padecer artrosis de rodilla, sobre todo después de la menopausia [48]. Esta observación ha llevado a la realización de varios estudios para entender qué relación podría tener la pérdida de los niveles de estrógenos con el desarrollo y severidad de la artrosis. Sin embargo, los resultados hasta el momento son controvertidos y se deben de realizar más trabajos para clarificar la relación entre la carga hormonal y la incidencia y severidad de la enfermedad [49-52].

Factores genéticos.- Se estima que el componente hereditario contribuye en un 40-60% a la susceptibilidad de padecer artrosis de manos, cadera y rodilla [53]. Hasta la fecha se han identificado polimorfismos en 11 localizaciones dentro del genoma que se asocian con una mayor susceptibilidad a desarrollar artrosis: en GDF5, que codifica para una proteína que regula la expresión de diferentes proteínas estructurales y fundamentales del cartílago como son colágeno II y el agregano [54, 55]; 6 genes (*PRKAR2B*, *HPB1*, *COG5*, *GPR22*, *DUS4L* y *BCAP29*) en el cromosoma 7q22 [56], polimorfismos en el gen MCF2L [57], en el cromosoma 3 donde se encuentra un polimorfismo para en GNL3 y GLT8D1; ASTN2 en el cromosoma 9; FILIP1 y SENP6 en el cromosoma 6; KLHDC5 y PTHLH en el cromosoma 9; y en el cromosoma 12 donde está CHST11 [58].

También se han identificado polimorfismos en genes asociados con la sensibilidad al dolor como COMT o TRPV1[59].

Por otra parte, los polimorfismos en el DNA mitocondrial los cuales determinan los denominados “haplogrupos” también han demostrado su papel en la susceptibilidad al desarrollo de la artrosis [60-63].

Nutrición.- Existe la hipótesis de que la presencia de agentes oxidantes aumenta el daño en el cartílago y favorece la susceptibilidad de la artrosis. Diversos estudios sugieren que la ingesta de vitamina C y D protege frente a la progresión de la enfermedad [64-66].

Factores de riesgo mecánicos:

Obesidad.- La obesidad es uno de los factores que más fuertemente afecta a la artrosis [39]. Existe una relación entre el índice de masa corporal con el grado de incidencia radiográfica y sintomática de la artrosis rodilla, y se ha visto que índices de masa corporal alto aceleran su progresión [67-72]. La obesidad también es un factor de riesgo en la artrosis de cadera aunque en este caso la asociación no es tan fuerte [42]. Parece obvio pensar que la obesidad incrementa la probabilidad de padecer artrosis puesto que la carga sobre la articulación es mayor en una persona obesa. Sin embargo, el papel de la grasa corporal en la articulación no está aún claro. El hecho de que en la artrosis de mano la relación entre obesidad y artrosis no sea tan evidente hace pensar que la grasa corporal está afectando de alguna otra forma al desarrollo de la artrosis [73]. Algunos estudios relacionan la obesidad con los efectos sistémicos (inflamación y efectos metabólicos) que se producen en la articulación durante la artrosis, relacionando el aumento de adipocinas inflamatorias con el desarrollo y progresión de la artrosis [74-76].

Densidad ósea.- Una característica importante de la artrosis es la alteración a nivel óseo que desencadena la formación de osteofitos y esclerosis del hueso subcondral. Por este motivo, se ha estudiado el papel de la densidad ósea en el desarrollo y progresión de la enfermedad. Según los estudios realizados hasta el momento parece que una elevada densidad ósea favorece el desarrollo de la artrosis mientras que la progresión de la enfermedad puede estar asociada con la pérdida ósea a nivel sistémico o local [77-79].

Actividad física y traumatismos.- Estudios poblacionales han demostrado que la actividad física puede tener beneficios para fortalecer la musculatura periarticular y que los niveles normales de actividad física no están asociados con la artrosis sintomática o radiográfica [80]. Sin embargo, si esta actividad es más prolongada o no se realiza correctamente, el ejercicio se convierte en un factor de riesgo para la artrosis [81]. Por otra parte, otros estudios muestran que los traumatismos en la rodilla pueden aumentar hasta 6 veces el riesgo de padecer artrosis [82, 83].

Morfología y alineamiento de la articulación- Tanto la morfología como el alineamiento de las articulaciones determinarán la carga que sufre la articulación y, por lo tanto, son factores que se deben tener en cuenta para entender la epidemiología de la enfermedad. Se ha comprobado que la displasia en la cadera es un factor de riesgo para la artrosis de cadera [84, 85] y que un mal alineamiento de las rodillas también es un factor predictivo de la progresión de la artrosis de rodilla [86].

2.4. Epidemiología e impacto socioeconómico de la artrosis

La artrosis puede afectar a cualquier articulación, pero las más afectadas son las rodillas, las caderas, las manos y la columna, todos ellos componentes fundamentales para la movilidad y la autonomía de las personas. Por lo tanto, la disfunción de estas articulaciones, puede provocar dependencia del individuo y en algunos casos, la incapacidad laboral permanente del paciente.

Los estudios de prevalencia de la artrosis varían dependiendo de la articulación objeto del análisis y de otros factores como el sexo o la edad. La mayor discrepancia entre los distintos estudios disponibles recae en función del método de diagnóstico utilizado, es decir, si el estudio se ha hecho teniendo en cuenta las manifestaciones clínicas del paciente, las evidencias radiológicas o ambas variables (artrosis sintomática).

Aunque por estos motivos, es difícil establecer cifras globales, existen datos que reflejan la importancia y el gasto que supone esta enfermedad para la sociedad. Así, se estima que el 5% de la población general mundial sufre artrosis de manos y que la prevalencia de la artrosis de rodilla es del 20%, mientras que la artrosis de cadera afecta alrededor del 5% en personas mayores de 40 años [2]. Por otro lado, según el estudio más reciente realizado en el 2010 por la Organización Mundial de la Salud para evaluar el daño de las enfermedades, la artrosis de rodilla y cadera se sitúa en el puesto número 11 en la lista de enfermedades que más contribuyen a la incapacidad del individuo. Además, esta incapacidad supone una pérdida media de 2.2 años en la vida laboral de las personas afectadas por la enfermedad [87, 88]. En España, el estudio epidemiológico EPISER realizado por la Sociedad Española de Reumatología (SER) en 2001 indica que el 6,2% de la población mayor de 65 años sufre artrosis sintomática de manos y un 10,2% artrosis sintomática de rodilla [89].

Por otra parte, existen muy pocos trabajos en los que se analice la incidencia de la artrosis debido a que es difícil identificar el momento en el que comienza la enfermedad y porque requieren estudios longitudinales y prospectivos. Los datos más

recientes para caracterizar la incidencia de la artrosis sintomática de mano, rodilla y cadera provienen del estudio que realizó Oliveria *et al* [90] (Figura 6).

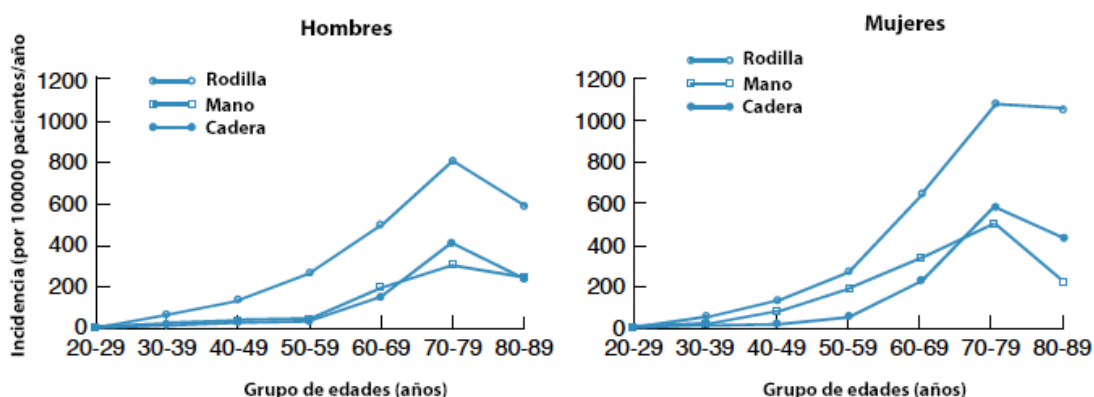


Figura 6. Incidencia de la artrosis de mano, rodilla y cadera determinados en el estudio de Oliveria *et al*. Figura adaptada de Arden N. *et al* [39]

En este estudio los rangos de incidencia se ajustaron por sexo y edad, de modo que observaron que la incidencia de la OA de manos era de 100 por 100 000 personas/año, la OA de rodilla de 240 por 100 000 personas /año y la OA de cadera 88 por 100 000 personas /año. Para las tres articulaciones analizadas, la incidencia aumentaba con la edad y las mujeres tenían ratios de incidencia más altos que los hombres, especialmente después de los 50 años.

La artrosis también tiene importantes costes económicos relacionados con el diagnóstico, seguimiento y tratamiento de los pacientes. Junto al cáncer, es la segunda patología que genera más gastos sanitarios, sólo por detrás de las enfermedades cardiovasculares. Según datos del estudio ArtRoCad, elaborado por la SER y Semergen, el tratamiento de la artrosis de rodilla y cadera, las más invalidantes, supone un coste directo (costes de tipo social) e indirecto (costes de tipo sanitario) de casi 5000 millones anuales, lo que supone un gasto medio por paciente de 1.500 euros al año [91] (Figura 7).

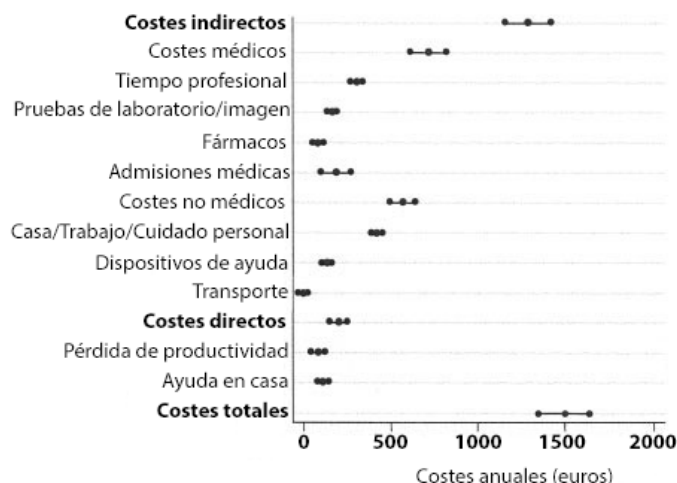


Figura 7. Representación de los costes económicos que supone la artrosis por cada paciente en España. El círculo en el medio de cada línea representa la media del coste analizado en cada caso y los círculos a los extremos de la línea representan el intervalo de confianza al 95% de los costes que supone el parámetro medido. Figura adaptada de Loza et al. 2009 [91].

Además, como se ha comentado anteriormente, la artrosis es una patología asociada a la vejez. Este factor junto con la ausencia de medidas preventivas y terapéuticas para la enfermedad, hace preveer que, debido al aumento de la esperanza de vida en los países desarrollados, la prevalencia y el impacto económico de la artrosis aumente con el envejecimiento de la población.

2.5. Diagnóstico y tratamiento de la artrosis

Actualmente el diagnóstico de la artrosis de mano, cadera y rodilla, se basa en los criterios establecidos por el *American Colleague of Rheumatology* (ACR). Estos criterios incluyen manifestaciones clínicas, signos radiológicos y varios parámetros biológicos [92] (Tabla 1).

Los síntomas clínicos evaluados son: dolor, rigidez articular, limitación funcional progresiva, crepitación ósea, hinchazón y derrame articular ocasional.

Los parámetros biológicos analizados son la velocidad de sedimentación globular y los niveles de proteína C-reactiva. Estos datos proporcionan información útil acerca del proceso de inflamación general en algunos pacientes con artrosis pero no son específicos de la articulación afectada y sus niveles no correlacionan con los daños del cartílago que realmente tiene el paciente [93].

Tabla 1. Criterios del *American Colleague of Rheumatology* para la clasificación de la artrosis de mano, cadera y rodilla.

| |
|---|
| Mano |
| Dolor en la mano junto al menos 3 de los siguientes criterios: |
| <ul style="list-style-type: none"> • Engrosamiento de estructuras óseas de más de 2 de 10 articulaciones interfalángicas seleccionadas de ambas manos • Engrosamiento de 2 o más articulaciones interfalángicas distales • Inflamación en más de 2 articulaciones interfalángicas • Deformación de al menos 1 de 10 articulaciones seleccionadas de ambas manos |
| Cadera |
| Dolor en la cadera junto con al menos 2 de los siguientes criterios: |
| <ul style="list-style-type: none"> • Velocidad de sedimentación globular <20mm/h • Radiografía con osteofitos • Radiografía con estrechamiento del espacio articular |
| Rodilla |
| Dolor en la rodilla junto con alguno de los siguientes criterios: |
| <ul style="list-style-type: none"> • Edad >50 años • Rigidez matutina inferior a 30 minutos • Crepitación en la movilización activa de la rodilla • Engrosamiento de estructuras óseas en la exploración • No aumento de temperatura cutánea en la rodilla |

Las técnicas de imagen utilizadas para el diagnóstico son la radiografía y la resonancia magnética.

Aunque la radiografía es la técnica más comúnmente empleada en el diagnóstico de la artrosis presenta varias limitaciones: su uso frecuente puede ser dañino para el paciente, es poco sensible para detectar cambios incipientes y es difícil evaluar la progresión de la enfermedad puesto que no es posible visualizar tejidos blandos como el cartílago y el parámetro utilizado para evaluar la destrucción del cartílago es la disminución del espacio articular [94-96].

La escala radiológica más utilizada en la valoración del daño estructural de la articulación es la de Kellgren y Lawrence (K/L) [97] que incluye 5 categorías (Tabla 2) que miden varios parámetros: disminución del espacio articular debido a una pérdida de volumen del cartílago, esclerosis del hueso subcondral y la presencia de osteofitos en los márgenes articulares.

Tabla 2. Escala radiográfica Kellgren-Lawrence para la clasificación de la artrosis.

| Grado | Diagnóstico | Criterios de evaluación |
|-------|-------------|---|
| 0 | Sano | <ul style="list-style-type: none">• Ausencia de signos radiológicos |
| I | Dudoso | <ul style="list-style-type: none">• Posible presencia de osteofitos• Reducción dudosa del espacio articular |
| II | Leve | <ul style="list-style-type: none">• Presencia de osteofitos definidos• Reducción leve del espacio articular |
| III | Moderada | <ul style="list-style-type: none">• Osteofitos definidos de tamaño medio• Reducción del espacio articular en más de un 50%• Esclerosis leve |
| IV | Severa | <ul style="list-style-type: none">• Reducción severa del espacio articular• Numerosos osteofitos• Esclerosis grave |

La otra técnica de imagen utilizada para evaluar el grado de enfermedad, aunque menos frecuente, es la resonancia magnética. Esta técnica de imagen es más sensible y de mayor especificidad que la radiografía, llegando a detectarse cambios mínimos en la estructura del cartílago. Su principal inconveniente es que es una tecnología muy cara y de difícil aplicación en la clínica [94, 98].

El tratamiento de la artrosis se centra en reducir el dolor del paciente y lo podemos dividir en farmacológico y no farmacológico [99, 100].

El tratamiento no farmacológico se basa en indicarle al paciente ciertos hábitos de su vida diaria que debe cambiar como por ejemplo evitar la sobrecargas en la articulación o modificar sus hábitos nutricionales en el caso de las personas con sobrepeso.

La terapia farmacológica es complementaria a la anterior y se basa en la administración de una serie de fármacos cuya finalidad es paliar el dolor del paciente.

Los fármacos más frecuentemente recetados son los analgésicos como el paracetamol, antiinflamatorios no esteroideos (AINES), los opiáceos como el tramadol, ácido hialurónico, los fármacos de acción lenta modificadores de síntomas como el condroitin sulfato o la glucosamina y en los casos más extremos se emplean los corticoides intrarticulares.

Sin embargo, cuando la artrosis ya está avanzada, el dolor es intenso y el daño en la articulación afecta fuertemente a la movilidad del paciente, se procede al tratamiento quirúrgico que consiste en el reemplazamiento de la articulación por una

prótesis. Aunque esta opción sólo es posible en el caso de que la articulación afectada sea la cadera o la rodilla.

En términos generales, se puede decir que la principal limitación en el diagnóstico es que la artrosis tiene una primera fase de duración desconocida en la que la enfermedad se muestra clínicamente silente pero donde ya se han producido alteraciones metabólicas al nivel del tejido. La consecuencia es que, cuando el paciente llega a la consulta, la enfermedad ya está avanzada y es muy difícil recuperar la capacidad funcional articular del individuo afectado.

Con todas las limitaciones que existen a nivel del diagnóstico y la falta de tratamiento eficaz frente a la artrosis, en los últimos años ha aumentado el interés por buscar marcadores biológicos que reflejen cambios en la remodelación de la articulación y que nos permitan realizar un diagnóstico precoz de la enfermedad, evaluar el pronóstico de la misma y monitorizar terapias alternativas y personalizadas para cada paciente [101].

3. Proteómica

3.1. Definición y aplicaciones

La proteómica es el término empleado para describir la disciplina encargada de estudiar los proteomas.

El término “proteoma” se usó por primera vez en 1995 [102, 103], para describir el conjunto de proteínas que se expresan a partir de un genoma. Por su parte, las proteínas son biomoléculas formadas por secuencias concretas de aminoácidos y son las encargadas de ejecutar las funciones encodificadas por el código genético.

El proteoma, a diferencia del genoma, es un elemento altamente dinámico, cuyos componentes varían en cada célula del organismo como consecuencia de la activación o supresión de genes, del procesamiento de los transcritos que codifica el genoma así como de las modificaciones postraduccionales de las diferentes variantes proteicas que a su vez se modulan dependiendo de las condiciones fisiológicas y el entorno (Figura 8).

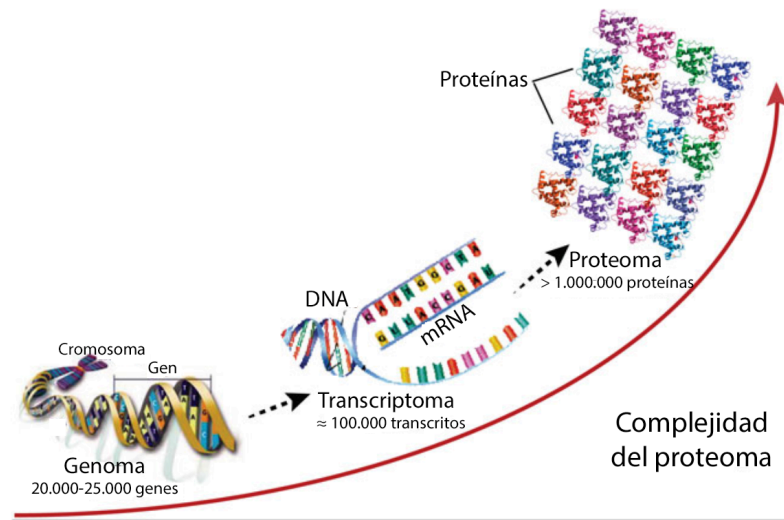


Figura 8. Figura mostrando la complejidad y variabilidad de las proteínas

La proteómica abarca un conjunto de tecnologías de alto rendimiento que nos permiten analizar cientos de proteínas en el mismo ensayo. Es una disciplina que comprende varias áreas de estudio [104]: la *proteómica descriptiva o estructural* cuyo objetivo es la identificación y caracterización (ej. modificaciones postraduccionales) a gran escala de las proteínas en una condición biológica determinada, la *proteómica comparativa o de expresión diferencial* que se encarga de identificar los cambios en el nivel de expresión de proteínas asociados a diferentes condiciones biológicas, y la *proteómica funcional* que se encarga de estudiar la localización y distribución subcelular de las proteínas así como sus interacciones (Figura 9).

Una de las aplicaciones más interesantes de la proteómica es la proteómica comparativa, que posibilita identificar y cuantificar las proteínas cuya presencia y/o abundancia resulten característicos de un determinado proceso biológico, y es la aplicación proteómica que se ha utilizado para la realización de esta tesis.

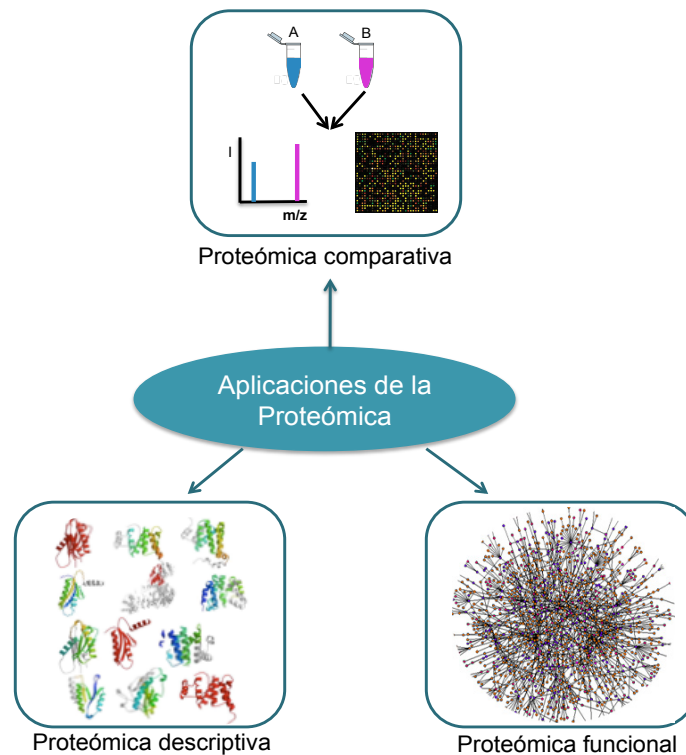


Figura 9. Aplicaciones de la proteómica

La proteómica está adquiriendo una gran relevancia en el ámbito científico. A raíz de la publicación del Proyecto Genoma Humano en el año 2001, la comunidad científica internacional se ha propuesto un nuevo reto, el Proyecto del Proteoma Humano (2010) [105]. El objetivo de este proyecto es identificar, localizar y caracterizar funcionalmente todas las proteínas humanas así como sus interacciones y alteraciones bajo diferentes condiciones fisiológicas, como puede ser una enfermedad o su respuesta a un determinado tratamiento. El uso y la combinación de diferentes técnicas proteómicas, concretamente la espectrometría de masas, el uso de anticuerpos como moléculas de afinidad y diferentes herramientas informáticas, forman en su conjunto un pilar básico e imprescindible para abordar el Proyecto del Proteoma Humano.

3.2. Biomarcadores proteicos en la artrosis

La irrupción de la proteómica en la última década ha abierto también nuevas perspectivas en la investigación clínica, principalmente en la búsqueda de nuevos biomarcadores.

El término biomarcador fue propuesto por primera vez a finales de los 90 por el *National Institute of Health* (NIH) americano para describir una característica que es

objetivamente medible y evaluada y que es un indicador del proceso biológico normal, de un proceso patológico o de una respuesta a una intervención terapéutica.

Un biomarcador ideal debería presentar una alta especificidad y sensibilidad para detectar estos procesos y cambios. La sensibilidad de un marcador se define como la proporción de verdaderos positivos (ej. presencia de enfermedad) identificados correctamente por el biomarcador y la especificidad se define como la proporción de verdaderos negativos (ej. no presencia de enfermedad) que son correctamente identificados por el biomarcador [106].

Los biomarcadores pueden ser moléculas biológicas como proteínas, metabolitos, carbohidratos y ácidos nucleicos. Los biomarcadores proteicos, sobre los que se centra esta tesis, se pueden clasificar en diferentes tipos de acuerdo a su función: biomarcadores de *diagnóstico*, los cuales permiten detectar una determinada patología y establecer el grado de la enfermedad, biomarcadores de *pronóstico* para detectar y predecir el curso de la enfermedad y biomarcadores de *respuesta* que refleja la eficacia de un determinado tratamiento farmacológico [107]. Estas moléculas pueden ser medidas en células, tejidos y en fluidos biológicos como el suero o la orina.

En el caso de la artrosis, los mejores biomarcadores son proteínas estructurales o fragmentos del cartílago, hueso o del compartimento sinovial que reflejen cambios en el metabolismo o degradación de la articulación. Estos biomarcadores pueden ser específicos de una articulación o comunes a todas ellas. Los compartimentos utilizados normalmente para buscar biomarcadores en artrosis son el cartílago, las proteínas liberadas por el cartílago (secretoma), el líquido sinovial, el suero y la orina.

En la última reunión de la Sociedad Europea sobre los Aspectos Clínicos y Económicos de la Osteoporosis y la Artrosis se definieron todos los biomarcadores hasta ahora descritos en la artrosis y su potencial valor en la práctica clínica (Tabla 3). [108-110]. La mayoría de estos biomarcadores son proteínas asociadas con el metabolismo del colágeno en el cartílago y en el hueso. Otras corresponden a proteínas de inflamación y/o fibrosis.

Tabla 3. Biomarcadores de artrosis descritos hasta el momento que se encuentran en fase de investigación, y compartimento donde han sido descritos (S:suero; O:orina; LS: Líquido sinovial; C:cartílago)

| Tipos | Biomarcadores | Compartimento |
|---|---|----------------------|
| Biomarcadores asociados al metabolismo del colágeno | Telopéptido carboxiterminal del colágeno tipo II (CTX-II) | S/O |
| | Fragmentos helicoidales del colágeno tipo II (α -CTX-II) | S/O |
| | Propéptidos del colágeno tipo II (PIINP, PIIANP, PIIBNP, PIICP, CPII) | S |
| | Pyridinolina | O |
| | Glucosil- galacteosil- pyridolina (Glc-Gal-PYD) | S |
| | Producto de fragmentación del colágeno tipo II (C2C) | S |
| | Neoepítipo del colágeno tipo II (C2M) | O |
| | Telopéptido carboxiterminal del colágeno tipo I (CTX-I, α -CTX-I) | S/O |
| | Telopéptido aminoterminal del colágeno tipo I (NTX-I) | S/O |
| | Propéptido aminoterminal del colágeno tipo I (PINP) | S |
| Neoepítipos de la fragmentación del colágeno I y II (C1,C2) | S | |
| Biomarcadores asociados al metabolismo del agregano | Fragmentos del agregano (ARGS, FFGV y neoepítipos) | C/LS |
| | Epítipo 846 del condroitin sulfato | LS |
| | Keratan sulfato | S/LS |
| Biomarcadores asociados a otras proteínas de la matriz | <i>Cartilage oligomeric matrix protein</i> (COMP y su forma deaminada D-COMP) | S |
| | Péptidos de la <i>fibulin 3</i> (Fib3-1y Fib3-2) | O |
| | <i>Follistatin-like protein 1</i> (FSTL-1) | S/LS |
| | Ácido hialurónico | S |
| | Metaloproteasas (MMP-1, MMP-3, MMP-9, MMP-13) | S |
| | Inhibidores de metaloproteasas (TIMPs) | S |
| | <i>Cartilage glycoprotein 39</i> (YKL-40) | S |
| Receptor soluble para productos glicosilados (sRAGE) | S/LS | |
| Inflamación/ Fibrosis | Proteína C reactiva, IL-1 β ,IL-6 y COX-2 | S |
| | Proteínas del complemento | S |
| | Adipoquinas (adiponectina, leptina, visfatina) | S |
| | Receptor soluble de la leptina (sOB-Rb) | S |
| | Periostin | C/LS |
| | Inhibidores de la ruta <i>Wnt</i> (DKKs y SOST) | C |
| Ácido úrico | S | |

Aunque todos ellos aportan valiosa información sobre la enfermedad, ninguno ha llegado a la fase de aplicación en la clínica.

4. Estrategias proteómicas para la búsqueda de biomarcadores.

El proceso para generar un biomarcador es largo y tiene varios pasos que van desde el descubrimiento de los potenciales biomarcadores, su validación por otras técnicas hasta su aplicación final en la práctica clínica (Figura 10).

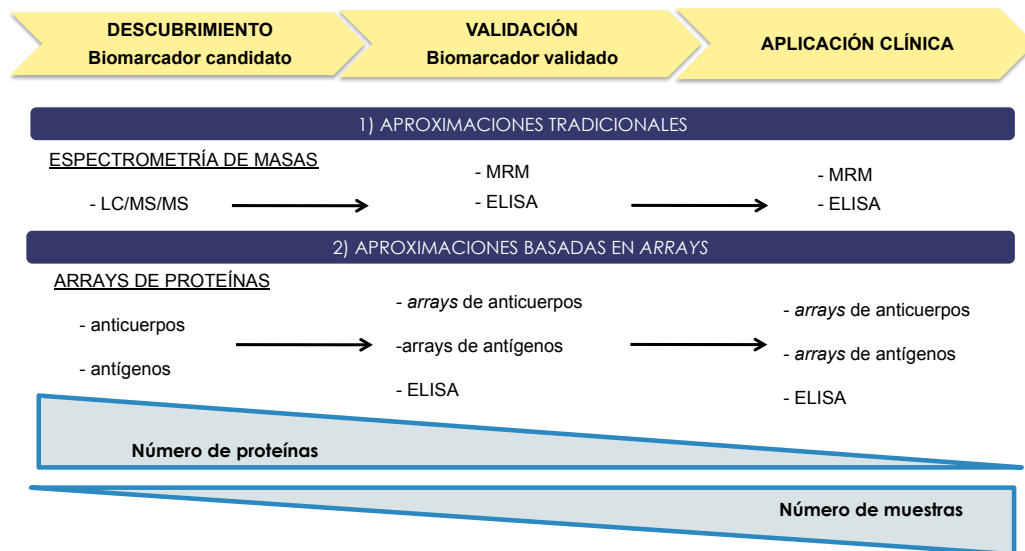


Figura 10. Flujo de trabajo habitual en la búsqueda de biomarcadores proteicos. La proteómica es especialmente útil en las fases de descubrimiento y verificación, empleando técnicas basadas en espectrometría de masas o en *arrays* de proteínas.

En la fase de descubrimiento y búsqueda de potenciales biomarcadores, las técnicas proteómicas clásicas se basan en el uso de espectrometría de masas. Normalmente son técnicas proteómicas no dirigidas conocidas con el término en inglés de *shotgun proteomics*, cuya finalidad es identificar y/o cuantificar todas las proteínas de la muestra en la condición biológica de interés. La espectrometría de masas también se puede utilizar en la validación de los biomarcadores candidatos. En este caso se utilizan estrategias dirigidas como son, la *Multiple Reaction Monitoring* (MRM) [111].

Sin embargo, en los últimos años, los *arrays* de proteínas han emergido como una potente tecnología para identificar moléculas candidatas en las primeras etapas del proceso de generación de biomarcadores. Por otra parte, aunque el inmunoensayo por ELISA sigue siendo la técnica estándar más utilizada en todos los laboratorios para la fase de validación, los *arrays* de proteínas también se están empleando como una potente herramienta para la validación, en un mayor número de muestras, de las proteínas candidatas con potencial biomarcador.

4.1. Técnicas basadas en espectrometría de masas

Las técnicas basadas en espectrometría de masas para la búsqueda de biomarcadores son las estrategias no dirigidas cuyo objetivo es realizar una identificación global de todas las proteínas de la muestra. Aunque la espectrometría de masas tiene una elevada especificidad para la identificación de proteínas, una de las principales limitaciones es su sensibilidad analítica para analizar muestras biológicas complejas. Esta tecnología ofrece una sensibilidad del rango de $\mu\text{g/mL}$ para identificar las proteínas. Por esta razón, estas estrategias conllevan normalmente la separación de las proteínas de la muestra biológica objeto del estudio, la digestión de las proteínas y finalmente, la identificación de los péptidos correspondientes por espectrometría de masas (MS) (Figura 11).

La digestión de las proteínas se puede realizar antes o después de la separación. Las técnicas de separación de proteínas pueden ser de dos tipos: la electroforesis en gel y la cromatografía líquida (LC).

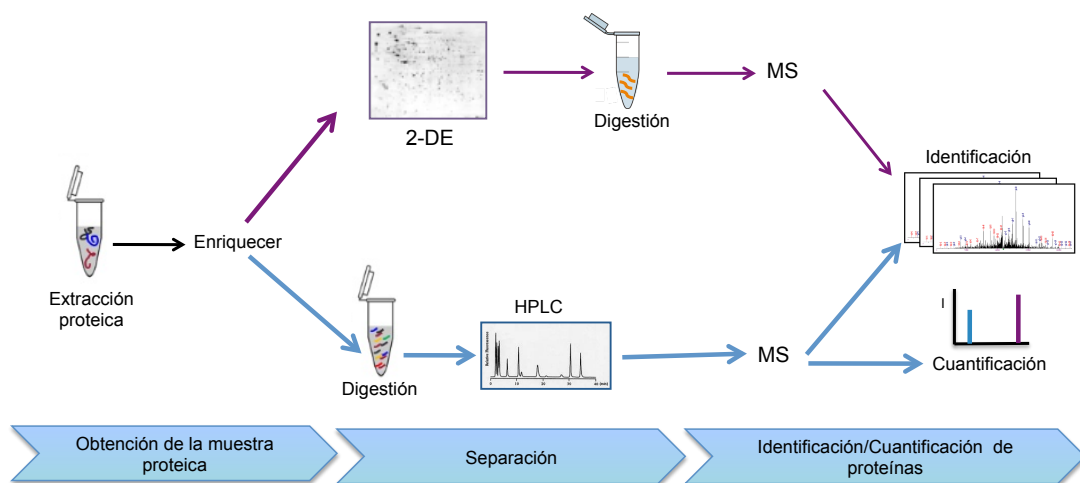


Figura 11. Estrategias proteómicas basadas en espectrometría de masas empleadas en la búsqueda de biomarcadores.

4.1.1. Digestión de las proteínas

La digestión proteica consiste en romper las proteínas por sus enlaces peptídicos dando lugar a sus fragmentos, los péptidos.

El proceso de digestión se suele realizar después de la separación en gel de las proteínas o directamente sobre una solución proteica, y los péptidos resultantes son posteriormente separados por cromatografía líquida.

En general, los dos protocolos de digestión necesitan un paso previo en el que se desnaturalizan las proteínas, bien por calor o bien a través de un agente químico como el dodecil sulfato sódico (SDS) o la urea [112]. Una vez desnaturalizadas, las proteínas se someten a un siguiente paso de rotura de los puentes disulfuro, denominado reducción, seguido de la alquilación de los grupos sulfhidrilo para evitar la formación de los puentes disulfuro de nuevo.

Los agentes reductores más empleados son el beta-mercaptoetanol y el ditioneitol (DTT), mientras que el agente alquilante más empleado tanto en la digestión en gel como en la digestión en solución es la iodoacetamida (IAA) [113].

Finalmente, las proteínas se digieren utilizando unas enzimas denominadas peptidasas que catalizan la rotura hidrolítica de los enlaces peptídicos. Algunas peptidasas cortan específicamente en el enlace peptídico entre dos aminoácidos concretos, dando lugar a una digestión proteica predecible y de forma reproducible. Este es el caso de una de las enzimas más empleadas en proteómica, la tripsina. Esta enzima hidroliza un enlace peptídico después de los residuos de lisina (Lys) y arginina (Arg). La digestión con tripsina se realiza a 37°C durante aproximadamente 16h, sin superar nunca este tiempo con el fin de evitar la autólisis de la enzima.

Otra de las enzimas más comúnmente utilizadas son la quimiotripsina, la Asp-N-proteasa, la pepsina o la endopeptidasa. Todas ellas difieren en la especificidad que tienen por el punto de corte entre aminoácidos.

4.1.2. Separación de las proteínas

Uno de los principales obstáculos para buscar biomarcadores es que se trabaja normalmente con muestras biológicas complejas como puede ser el suero humano, en el que hay una gran variedad de proteínas distintas con un rango dinámico muy amplio.

Con el fin de identificar el mayor número de proteínas posibles, se pueden combinar varias técnicas de separación de proteínas previo al análisis de MS, lo que permite ganar resolución y sensibilidad en la identificación del proteoma. Los dos métodos más utilizados para el fraccionamiento de proteínas son los métodos de

separación basados en gel y los que se basan en la separación por cromatografía líquida.

4.1.2.1. Métodos de separación basados en gel

La técnica en gel más ampliamente utilizada es la electroforesis desnaturizante en gel de poliacrilamida (SDS-PAGE).

En esta técnica, las proteínas se tratan con SDS, que aporta cargas negativas a la proteína y hace que las proteínas se separen en el gel de acuerdo únicamente a su peso molecular después de aplicar un campo eléctrico sobre el gel.

La porosidad del gel variará dependiendo de la concentración a la que se utilice la poliacrilamida, y determina finalmente el rango de separación de las proteínas en el gel.

A partir del SDS-PAGE han surgido otros métodos como la electroforesis bidimensional (2-DE) y la electroforesis diferencial en gel (DIGE) que mejoran el rango de separación de proteínas.

La 2-DE es una técnica de fraccionamiento proteico que fue desarrollada en los años 70 por Kenrick & Margolis y por O' Farrell [114]. Esta técnica ha sido muy utilizada en proteómica para buscar biomarcadores, puesto que tiene una gran resolución al combinar dos pasos de separación. En primer lugar se realiza un isoelectroenfoque (IEF) donde las proteínas se separan en un gradiente continuo de pH en función de su punto isoeléctrico (pI), que se define como el pH en el que la carga neta de la proteína o el péptido es cero.

A continuación, se lleva a cabo una separación SDS-PAGE, donde las proteínas se separan en función de su masa molecular. La técnica 2-DE, al contrario que el SDS-PAGE, tiene la ventaja de que permite la identificación de isoformas proteicas y de modificaciones post-traduccionales. Sin embargo, es una técnica poco reproducible entre los geles, y tiene poca sensibilidad para detectar proteínas poco abundantes y para resolver proteínas que tienen puntos isoeléctricos extremos ($3 > pI$ o $pI > 11$) [115] (Figura 12).

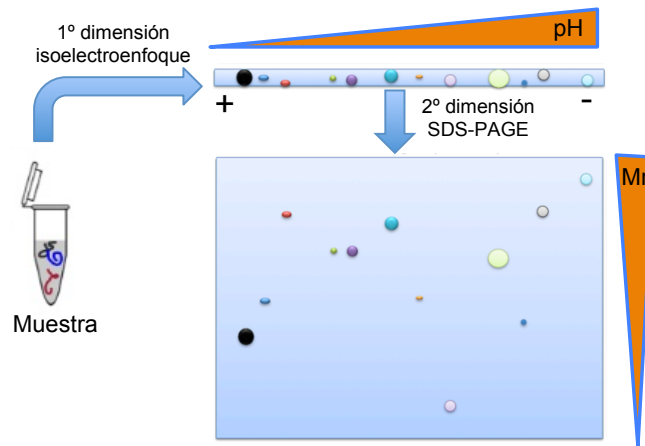


Figura 12. Separación de proteínas mediante 2-DE.

Por su parte, la técnica DIGE combina el marcaje fluorescente de las proteínas con la electroforesis 2-DE. En este caso, se pueden comparar condiciones biológicas diferentes marcando las proteínas de cada condición con fluoróforos derivados de la cianina CyDye (Cy2, Cy3 y Cy5) que se unen de forma covalente a las lisinas de las proteínas [116]. En esta técnica se utiliza un estándar interno marcado con Cy2 y compuesto por una mezcla proteica formada por las condiciones biológicas que se quieren comparar. Después del marcaje, se mezclan las 3 condiciones marcadas y se separan las proteínas en un único gel por 2-DE. Tras la electroforesis, se captura la imagen del patrón proteico descrito por la excitación de cada uno de los fluoróforos y se procede con el análisis de imagen para detectar las manchas proteicas diferenciales en las distintas condiciones (Figura 13). El marcaje fluorescente, la separación proteica simultánea de las dos condiciones biológicas y la presencia del estándar interno en el diseño experimental del DIGE hacen que ésta sea una técnica más sensible, más reproducible y que permite una cuantificación más precisa y robusta que la 2-DE.

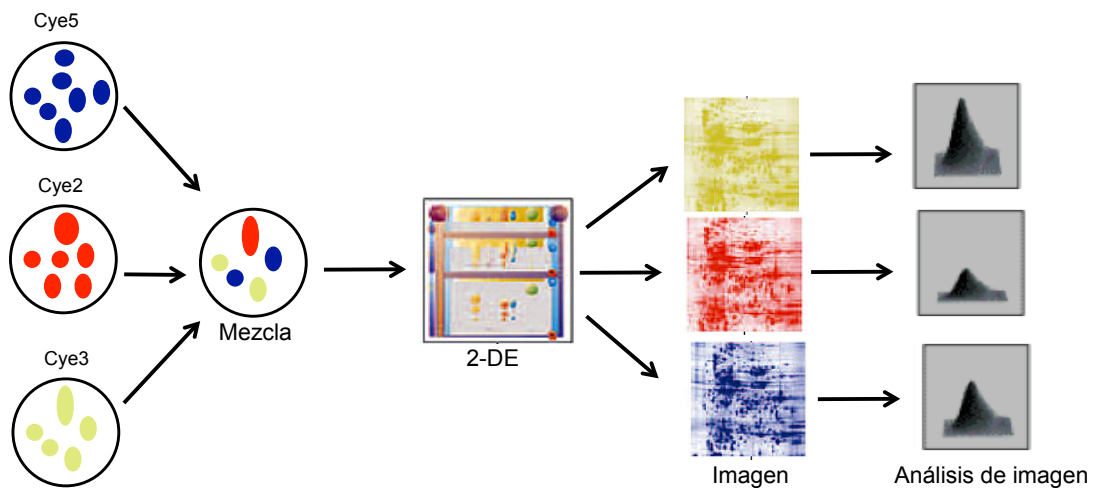


Figura 13. Flujo de trabajo del DIGE para la separación y cuantificación de proteínas. Las dos muestras se marcan con los fluoróforos Cy3 y Cy5 respectivamente, y el estándar interno formado por la mezcla de cantidades iguales de las dos condiciones se marca con el Cy2. Se mezclan las tres muestras marcadas y se separan mediante electroforesis bidimensional. Por último, los geles se escanean a diferentes longitudes de onda, obteniendo las imágenes de los tres fluoróforos.

4.1.2.2 Métodos de separación cromatográficos

Con el nombre de cromatografía se agrupan varias técnicas que separan los componentes de una mezcla por su interacción con una fase estacionaria y una fase móvil. En proteómica, la cromatografía más utilizada es la cromatografía líquida (LC), donde la fase estacionaria es un sólido contenido en una columna y la fase móvil es un líquido. Las fases estacionarias más empleadas en la separación de muestras proteicas son las columnas de fase reversa (RP), de exclusión molecular, de intercambio iónico y de afinidad.

Cromatografía de fase reversa: La cromatografía de fase reversa es una cromatografía de reparto en la que la fase estacionaria es apolar y la fase móvil polar. Frecuentemente, la fase estacionaria está compuesta de partículas de sílica químicamente modificadas con hidrocarburos, y la fase móvil es un compuesto relativamente polar como el propanol o el acetonitrilo (ACN). En este tipo de cromatografía, los componentes más polares de la muestra eluyen primero de la columna y al disminuir la polaridad de la fase móvil, por ejemplo aumentando gradualmente la concentración de ACN, se van separando los componentes más hidrófobos de la muestra.

Cromatografía de exclusión molecular: En la cromatografía de exclusión molecular, la fase estacionaria es un gel o una resina formados por poros de diferentes tamaños. Las moléculas de mayor tamaño que los poros no pueden entrar en la resina, y eluyen de la columna con mayor rapidez. Las moléculas de menor tamaño pueden entrar en los poros y eluyen más tarde. En este tipo de cromatografía las proteínas se separan por orden decreciente de tamaño molecular.

Cromatografía de intercambio iónico: En este tipo de cromatografía, las proteínas se separan por diferencia de carga. En este caso, la fase estacionaria presenta en la superficie grupos funcionales cargados que interactúan con iones de carga opuesta presentes en la muestra. La elución de las proteínas retenidas en la columna se realiza aumentando la fuerza iónica de la fase móvil, de forma que primero eluyen las proteínas menos cargadas y por último las de mayor carga. Este tipo de cromatografía puede ser de dos tipos: de intercambio catiónico (fase estacionaria cargada negativamente) e intercambio aniónico (fase estacionaria cargada positivamente).

Cromatografía de afinidad: En la cromatografía de afinidad las proteínas se separan por su afinidad al ligando que tapiza la fase estacionaria. Este ligando puede ser un anticuerpo, un metal, una enzima, un carbohidrato o cualquier otra molécula que interaccione específicamente con la proteína que queremos separar en la muestra. Este tipo de cromatografía es la que se suele utilizar para deplecionar las proteínas abundantes de muestras biológicas complejas como el líquido sinovial o el suero puesto que evita la depleción de proteínas minoritarias por tratarse de una unión específica de las proteínas abundantes con sus anticuerpos inmovilizados en la columna cromatográfica.

Si la cromatografía se realiza después de la digestión enzimática de las proteínas, la mezcla peptídica que se obtiene es muy compleja por lo que se hace necesario realizar una cromatografía multidimensional que consiste en acoplar dos o más cromatografías que se basen en métodos de separación diferentes. Para la primera dimensión se usan normalmente dos métodos: primero se realiza una cromatografía de intercambio catiónico o RP a pH básico. Una segunda separación por RP, debido a su elevada eficacia y a los solventes usados en la elución, es siempre el método final que se acopla directamente al espectrómetro de masas.

Un tipo especial de cromatografía es la Cromatografía Líquida de Alta Eficacia (HPLC). HPLC es el resultado de la optimización de los parámetros que determinan la eficiencia de la separación cromatográfica líquida que son fundamentalmente el tamaño de la columna y el tamaño de las partículas que forman su matriz.

Estas mejoras han hecho de este tipo cromatografía la más utilizada en proteómica para separar proteínas y péptidos de forma rápida y eficaz.

Desde el punto de vista de la separación, en HPLC también podemos tener cromatografía de intercambio iónico, de separación en gel, de fase reversa o de afinidad.

4.1.3. Espectrometría de masas

La espectrometría de masas (MS) es una técnica analítica que permite identificar y cuantificar moléculas después de transformarlas en iones cargados. Los analitos se identifican posteriormente por la relación masa/carga (m/z) de sus iones.

Los espectrómetros de masas consisten en 3 elementos básicos: una fuente de iones que convierten los analitos en iones en fase gaseosa, un analizador de masas que separa los analitos ionizados basándose en su relación m/z , y un detector que registra el número de iones de cada una de los valores m/z [117]. Los espectrómetros de masas fueron utilizados por primera vez para caracterizar proteínas y péptidos a finales de los años 50 [118]. La evolución más grande de esta tecnología, se produjo en la década de los 70 y 80, con el desarrollo de los espectrómetros de masas en tándem [119] y las fuentes de ionización.

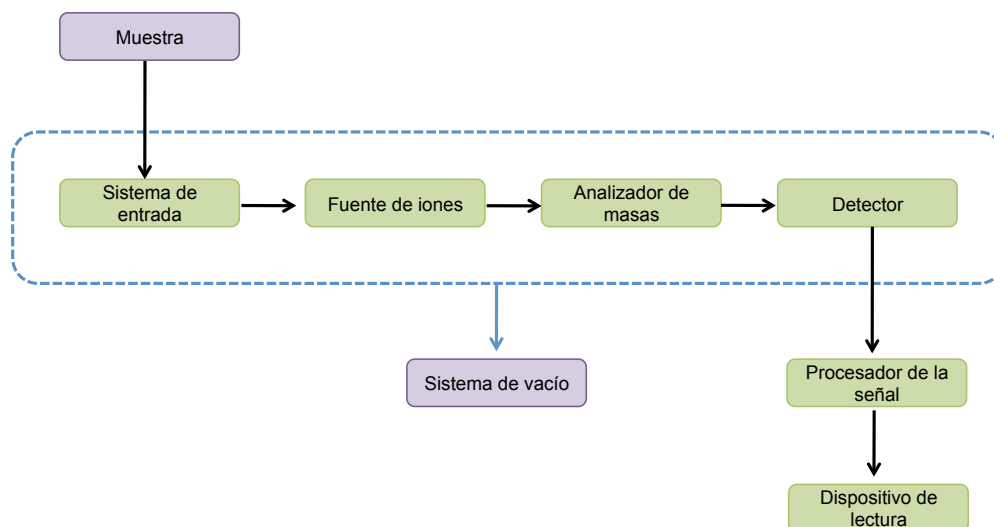


Figura 14. Componentes de un espectrómetro de masas

4.1.3.1 Fuente de ionización

La fuente de ionización se encarga de convertir los analitos en iones gaseosos. Si la muestra a analizar está en fase gas o es un líquido volátil se utilizan fuentes de impacto de electrones, ionización química o desorción por campo. Si se trata de sólidos o disoluciones en fase líquida se pueden utilizar las modernas fuentes de electro spray (ESI), ionización química a presión atmosférica, ionización por láser asistida por matriz (MALDI), ionización por iones secundarios o incluso las más recientes técnicas de ionización directa [120].

Las fuentes más empleadas en proteómica son la fuente de desorción/ionización por láser asistida por una matriz (MALDI) y la ionización por electro spray (ESI) (Figura 15).

MALDI: En esta fuente de ionización, la muestra se mezcla con una matriz orgánica y se deja cristalizar sobre una superficie metálica. En este caso, la ionización se produce por sublimación de la muestra sólida al incidir un láser sobre la matriz dando lugar a moléculas protonadas en estado gas que son dirigidas hacia el analizador de masas a través de un campo eléctrico [121]. Las matrices más empleadas en MALDI son el ácido α -ciano-4 hidroxicinámico (CHCA), el ácido 2,5 dihidroxibenzoico (DHB), y el ácido 3,5-dimetoxi-4-hidroxicinámico, que absorben energía en la longitud de onda del ultravioleta. La preparación de la matriz y la cristalización influyen de forma directa en la ionización y en la calidad de los espectros, de modo que cuando una cristalización no es homogénea disminuye la resolución y la correlación entre la señal y la concentración del analito en la muestra.

ESI: En este tipo de ionización, la muestra que está en estado líquido pasa a través de una aguja de acero a la que se le aplica un voltaje (2.5-6kV) de modo que las partículas de la muestra se dispersan a modo de pequeñas gotas cargadas del orden de micrómetros (*electrospray*). Estas microgotas se evaporan rápidamente, dando lugar a gotas de menor tamaño y a un incremento de la densidad de carga que supera la tensión superficial y hace que las moléculas se separen en varios analitos con diferentes cargas que, por acción del campo eléctrico, se transfieren finalmente al analizador de masas [122].

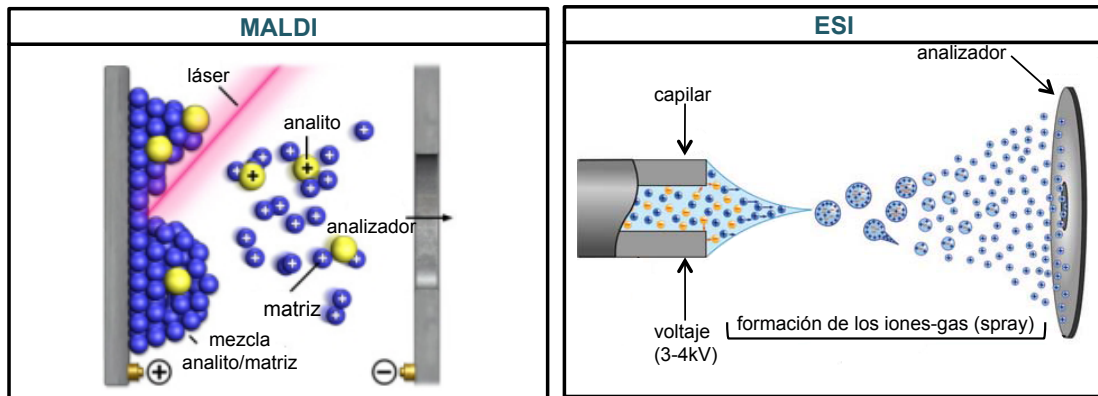


Figura 15. Fuentes de ionización más empleadas en espectrometría de masas.

4.1.3.2. Analizadores de masas

El analizador de masas es el elemento más importante de la espectrometría de masas y de sus características depende la sensibilidad, resolución y la precisión para generar información acerca del espectro m/z (MS) y del espectro de fragmentación (MS/MS) de los analitos.

Básicamente hay 4 tipos de analizadores de masas que se utilizan habitualmente en la investigación proteómica: tiempo de vuelo (TOF), el analizador de cuadrupolo, la trampa de iones (IT) y el analizador de masas de transformada de Fourier y resonancia iónica ciclotrónica (FT-ICR) [123]. (Figura 16).

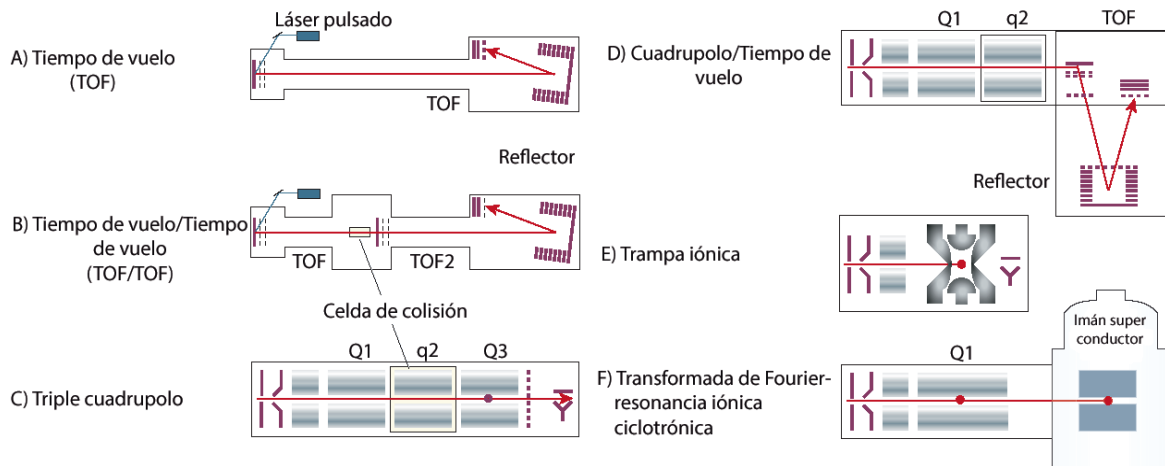


Figura 16. Tipo de analizadores de masas más empleados en proteómica. A) Analizador de tiempo de vuelo (TOF), que separa los iones en base a sus diferentes velocidades de vuelo. B) Analizador TOF-TOF que incorpora una celda de colisión donde se fragmentan los iones seleccionados en el primer TOF y después los fragmentos, se separan en el segundo TOF. C) Analizador de triple cuadrupolo que separa los iones por variaciones en el tiempo de los campos eléctricos que se generan entre los cuadrupolos. Los iones se seleccionan en el primer cuadrupolo (Q1), se fragmentan en el segundo (q2) y se separan y se escanean en el tercer cuadrupolo (Q3) por resonancia del campo magnético aplicado. En la trampa lineal, los iones se seleccionan y se escanean en Q3. D) Analizador cuadrupolo/tiempo de vuelo (Q/TOF) en el que el tercer cuadrupolo se sustituye por un TOF que mide la masa de los iones. E) El analizador de trampa iónica, que captura los iones en la trampa donde los fragmenta y los escanea para dar lugar a su espectro de masas. F) El analizador de masas transformada de Fourier y resonancia ciclotrónica (FT-ICR) que atrapa los iones a través de campos magnéticos. Figura adaptada de Aebersold y Mann [123].

TOF

Los analizadores tipo TOF miden la masa de los péptidos intactos y se suelen acoplar a las fuentes de ionización MALDI.

En el TOF los iones son acelerados en el interior de un tubo hasta llegar a un reflector donde los iones son dirigidos hacia un detector. Los iones se separan según diferencias en las cinéticas de los iones y en el detector se mide y cuantifican los iones que llegan de cada una de las m/z .

En el analizador TOF-TOF, que es un espectrómetro de masas en tándem, se incorpora una celda de colisión (CID) entre los dos TOF de modo que iones con una

determinada m/z son seleccionados en el primer TOF, fragmentados en la CID, y las masas de los fragmentos son separadas y detectadas en el siguiente TOF.

Cuadrupolo

Los cuadrupolos están formados por cuatro barras equidistantes a una distancia determinada de un eje central y en su conjunto forman un cilindro por el que pasan los iones.

A cada par de barras se le aplica una combinación de corriente continua y corriente alterna en el rango de las radiofrecuencias (RF), ambas de amplitud variable que crean en su interior un campo eléctrico denominado cuadripolar.

Los iones que entran en el analizador son sometidos al efecto de este campo eléctrico que los hace oscilar y los desvía en función de su valor m/z de forma que para una combinación de potenciales y RF sólo los iones en un estrecho rango de valores m/z llegan a incidir en el detector.

El triple cuadrupolo es el más utilizado en proteómica. Este espectrómetro está formado por tres cuadrupolos y trabaja en dos modos, primero los iones atraviesan los tres cuadrupolos obteniendo las relaciones m/z de todos los iones, y en el segundo modo, el primer cuadrupolo filtra los iones que posteriormente se fragmentan en el segundo cuadrupolo y se escanean en el tercer cuadrupolo.

El cuadrupolo también se puede combinar con un TOF, de forma que tenemos un triple cuadrupolo con un reflector TOF para medir las masas de los iones.

Trampa de iones

En este tipo de analizadores, los iones son atrapados durante un intervalo de tiempo prolongado en unas orbitas concéntricas tridimensionales creadas por tres cuadrupolos (trampa).

La muestra se ioniza y los iones se aceleran hacia el interior de la trampa. Todos los iones se acumulan en la trampa y posteriormente, mediante la variación de la RF, se seleccionan los iones para fragmentar. Para la fragmentación se introduce gas He en la trampa. Una característica adicional de la trampa es que los fragmentos generados pueden ser a su vez seleccionados y fragmentados en la misma trampa como si fuera un sistema de espectrometría de masas en tándem. Sin embargo, las trampas tienen baja capacidad de resolución y una limitada capacidad para atrapar los iones. Con el fin de mejorar sus cualidades se creó la trampa de iones lineal (LIT) y la trampa linear cuadrupolar (LTQ), que tienen mayor resolución de masa.

Un tipo especial de trampa de iones es el orbitrap o trampa orbital, en donde los iones son atrapados bajo un campo eléctrico estático en el interior de una trampa que se compone de un electrodo exterior en forma de barril y un electrodo coaxial interior de huso alrededor de un electrodo central donde los iones oscilan en dirección axial formando órbitas. Gracias a su resolución y velocidad de análisis es la trampa que se ha introducido más rápidamente en el mercado y se utiliza principalmente para la identificación de péptidos modificados. Las trampas de iones se suelen utilizar con un modo de ionización ESI.

Transformada de Fourier y resonancia iónica ciclotrónica

El espectrómetro de masas FT-ICR funciona de forma parecida a una trampa iónica, pero en este caso los iones gaseosos adquieren un movimiento circular (movimiento ciclotrónico) causado por un poderoso campo magnético que actúa dentro de la celda. La frecuencia de este movimiento (frecuencia ciclotrónica) es característica de los iones con una determinada m/z .

Para detectar los iones se aplica sobre la trampa un campo eléctrico y aquellos iones cuya frecuencia ciclotrónica coincida con el voltaje transmitido, alargan sus órbitas de movimiento de modo que la frecuencia es registrada en un detector en función del tiempo. La energía aplicada también se puede ajustar para producir fragmentación o para expulsar a los iones de la trampa.

Posteriormente, utilizando la transformación matemática de Fourier, las frecuencias detectadas se transforman en valores m/z con una gran resolución, sensibilidad y exactitud.

En los últimos años, los analizadores de masas se han ido modernizando con el objetivo de incrementar su sensibilidad, su exactitud en las masas y su capacidad de tratar una mayor número de muestras simultáneamente. Con esta finalidad han surgido los espectrómetros de masas híbridos, que están formados por la combinación de más de un espectrómetro de masas.

Todos ellos pueden funcionar solos o en combinación con otros analizadores de masas de modo que se acumulan las ventajas y desventajas de ambos.

4.1.3.3. Identificación de proteínas mediante MS

Las proteínas se identifican normalmente mediante la estrategia conocida como *bottom up*. En ella las proteínas se desnaturalizan, se separan, se digieren y los fragmentos peptídicos se identifican por espectrometría de masas.

El método *bottom up* más habitual para la identificación es utilizar la huella peptídica de la proteína que se corresponde con el espectro de masas (MS) correspondiente a los péptidos resultado de la digestión enzimática de la proteína. La huella peptídica se utiliza normalmente cuando la muestra proteica corresponde a un *spot* o a una banda definida, teñida y visualizada en un gel de poliacrilamida o cuando se trata de una proteína purificada mediante algún método cromatográfico.

Si el espectro de MS no es suficiente para la identificación de la proteína, se procede a la fragmentación de los péptidos, obteniéndose las masas de los fragmentos en un espectro de masas en tándem (MS/MS) de la proteína.

Posteriormente, el espectro de MS o el de MS/MS es volcado mediante un programa informático en una base de datos donde podemos comparar el espectro obtenido con el de espectros de digestión *in silico* [124].

4.1.4. Cuantificación de proteínas mediante MS

La cuantificación implica la medida de cantidad de proteína existente en una muestra dentro de un rango de magnitudes. En proteómica muchas veces no es suficiente identificar qué proteínas hay en la muestra si no también saber en qué proporción está esa proteína en una condición biológica determinada.

La cuantificación puede ser absoluta o relativa, siendo esta última mucho más utilizada en los estudios de proteómica cuantitativa. En las cuantificaciones relativas, la abundancia de las proteínas se mide por comparación entre distintas condiciones biológicas y ha sido la cuantificación utilizada en esta tesis. Las cuantificaciones se pueden realizar en gel o mediante MS.

4.1.4.1. Cuantificación de proteínas basada en gel

Las cuantificaciones que se realizan en las técnicas proteómicas basadas en gel son obtenidas a partir de las cantidades de proteína detectadas en los geles de electroforesis usados para la separación de las mismas. Después de la electroforesis, tanto en el SDS-PAGE como en el 2-DE, las proteínas se pueden visualizar utilizando tinciones colorimétricas como el azul de Coomassie y la tinción de plata o tinciones fluorescentes como el Sypro o el Cy5. La aplicación de un método u otro depende

de factores como la sensibilidad, el rango lineal, el equipo de adquisición de imagen disponible o la compatibilidad con las técnicas de identificación por MS posteriores (Tabla 4).

Tabla 4. Características de los principales métodos de tinción utilizados para cuantificar proteínas. Adaptada de Miller *et al* 2006 [125].

| Método | Detección | Sensibilidad | Rango lineal | Compatibilidad con MS |
|-------------------|---------------|--------------|--------------|-----------------------|
| Azul de Coomassie | Colorimétrica | 10ng | ++ | + |
| Tinción de plata | Colorimétrica | 1ng | + | + |
| Sypro Ruby | Fluorescente | 1ng | +++ | +++ |
| CyDye | Fluorescente | <1ng | +++ | ++ |

El colorante azul de Coomassie (R-250 y G-250) se une a las proteínas a través de interacciones iónicas con los grupos amino de las proteínas. La principal diferencia entre los dos tipos de Coomassie es que el G-250 presenta un menor ruido de fondo en la tinción y es más estable que la tinción R-250. La principal ventaja de la tinción con Coomassie es que presenta un bajo coste y una elevada compatibilidad con MS. Sin embargo, no es muy sensible, puesto que sólo puede detectar hasta 10 ng de proteína [126].

La tinción de plata puede ser o no compatible con MS dependiendo del reactivo utilizado (nitrato de plata o amoniaco de plata). La tinción que es compatible con MS se basa en impregnar el gel con una mezcla que contiene nitrato de plata y formaldehído a pH básico y posteriormente, revelar la señal por reducción de la proteína con una solución de carbonato cálcico. Este tipo de tinción es muy sensible permitiendo detectar hasta 1ng de proteína [125]. Las desventajas de la tinción con plata son el rango lineal limitado que posee (menos de un orden de magnitud) y la dificultad de automatización de los protocolos puesto que es un proceso relativamente largo. El otro inconveniente es que, en el revelado, podemos llegar a saturar la señal. Esto produce una disminución en la reproducibilidad entre los geles y por lo tanto, una reducción en la precisión de la cuantificación entre diferentes muestras proteicas.

La tinción fluorescente con Sypro Ruby es compatible con MS, es tan sensible como lo es la tinción de plata y ofrece un rango dinámico lineal más amplio. Además el proceso de tinción es sencillo y no implica muchos pasos como en la plata, lo cual hace

posible su aplicación automatizada en estudios de alto rendimiento donde se trabaja con varios geles y muchas muestras [127].

Otro protocolo de tinción fluorescente es el que deriva de los fluorocromos cianina CyDye que se aplica antes de la electroforesis en gel. Esta tinción tiene una sensibilidad mayor que las tinciones comentadas anteriormente y además, permite una mayor reproducibilidad [116]. El principal inconveniente de los métodos de tinción fluorescentes es que después de analizar las imágenes resultantes, el gel tiene que ser teñido a *posteriori* con técnicas colorimétricas para visualizar las proteínas, poder escindir las del gel y analizarlas por MS.

Para analizar las imágenes resultantes de las tinciones, se han desarrollado varios programas informáticos que incluyen la visualización de los geles, la detección de las proteínas en función de la densidad óptica que presentan tras su tinción y la correlación entre las distintas muestras y geles analizados para realizar la cuantificación relativa de sus proteínas en función de la intensidad de su tinción.

4.1.4.2 Cuantificación de proteínas basada en MS

En las estrategias proteómicas basadas en espectrometría de masas, la cuantificación puede ser absoluta o relativa pero normalmente se utiliza esta última debido a la falta de estándares de proteínas para realizar una cuantificación absoluta. En muchos casos la cuantificación relativa se lleva a cabo mediante un marcaje metabólico o químico utilizando distintos tipos de isótopos estables que se unen a las proteínas o a sus péptidos. Los péptidos marcados producen, en el análisis de MS o de MS/MS, dependiendo del método de marcaje empleado, parejas de picos con una diferencia de masa característica a la de los isótopos usados. La cuantificación en todos los casos se basa en la comparación de las intensidades de las señales correspondientes a los isótopos utilizados (Figura 17).

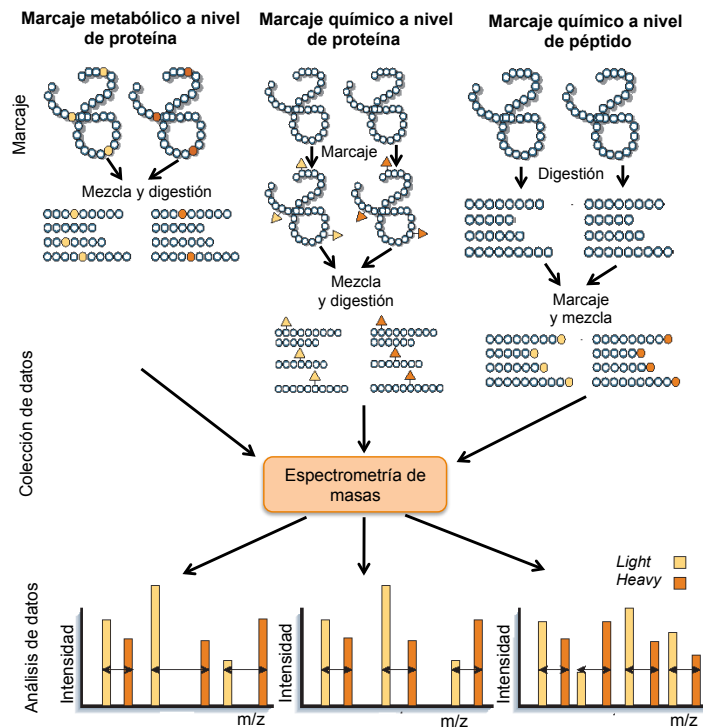


Figura 17. Estrategias para el marcaje de péptidos y proteínas mediante isótopos estables. Figura adaptada de Aebersold et al [123].

- **Marcaje metabólico mediante isótopos estables- SILAC**

El marcaje metabólico más utilizado en proteómica es el *Stable Isotope Labelling by Amino acids in Cell Culture* (SILAC) [128]. En esta técnica, los isótopos estables se incorporan a las proteínas durante el crecimiento y producción de proteínas por parte de células en cultivo. Los experimentos típicos de SILAC solo pueden comparar dos condiciones de cultivo celular de modo que las células de una determinada condición experimental se cultivan en un medio enriquecido con aminoácidos no marcados (N^{14} y C^{12}) y las células con otra condición se cultivan con aminoácidos marcados con isótopos estables de nitrógeno y carbono (N^{15} y C^{13}).

Los aminoácidos más ampliamente utilizados en SILAC son la arginina y la lisina, puesto que es entre estos aminoácidos donde corta la tripsina que es la enzima que se suele utilizar en los experimentos SILAC y de este modo se asegura la presencia de un aminoácido marcado en todos los péptidos.

Al final del tiempo de cultivo, se extraen las proteínas, se mezclan las proteínas de las dos condiciones en proporción 1:1, y se separan por SDS-PAGE. A continuación el gel se tiñe mediante una tinción colorimétrica, las bandas se someten a digestión

tríptica y los péptidos resultantes se separan por cromatografía líquida acoplada a un espectrómetro de masas.

En el SILAC la incorporación de los isótopos estables no interfiere en el crecimiento normal de las células ni a su síntesis proteica, pero sí que produce una variación de masa en las proteínas sintetizadas en la condición marcada con respecto a las proteínas de la condición sin marcar. Esta diferencia de masa nos permite identificar los espectros de masas de los péptidos de cada condición y cuantificar los niveles relativos de esa proteína en la muestras comparando las intensidades de todos los péptidos marcados y sin marcar correspondientes a la proteína.

La principal ventaja del SILAC es que se mezclan las dos condiciones proteicas desde el principio del estudio, por lo que cualquier paso o alteración en la parte experimental se producirá de forma homogénea en las dos muestras. La principal limitación es que solo es aplicable a estudios que impliquen cultivo celular, y que no permite comparar muchas condiciones ya que aumenta demasiado la complejidad del espectro de masas.

- **Marcaje químico mediante isótopos estables- iTRAQ**

En el año 2004, *Applied Biosystems* desarrolló una nueva generación de reactivos de marcaje isotópico llamada *Isotope Tagged Relative and Absolute Quantification* (iTRAQ) [129]. Recientemente, se han desarrollado ocho formas isotópicamente diferentes de los reactivos iTRAQ, pero, en sus orígenes, salió al mercado con sólo cuatro reactivos. Estos marcadores se componen de tres regiones bien diferenciadas 1) el grupo reactivo, 2) el grupo compensador y 3) el grupo reportero (Figura 18).

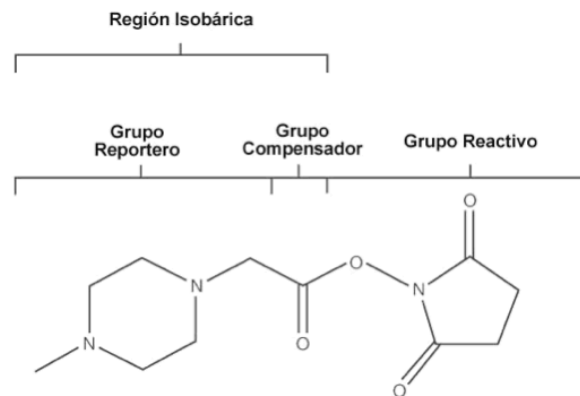


Figura 18. El reactivo iTRAQ es un marcador isobárico que presenta un grupo reportero, un grupo reactivo y una porción compensadora. El grupo reportero presenta una relación m/z que va desde 113,1 a 121,1 en la versión de ocho reactivos y de 114,1 a 117,1 en la de cuatro. Así, el grupo compensador compensa la diferencia que presentan entre sí los diferentes marcadores.

El grupo reactivo es derivado de la N-hidroxisuccimida, y reacciona con el grupo amino de los residuos de lisina y el extremo N-terminal de todos los péptidos. Durante la fragmentación de los péptidos marcados, los grupos reporteros se liberan de los mismos y pueden ser detectados en el espectro de MS/MS como picos con una relación m/z de 113,1; 114,1; 115,1; 116,1; 117,1; 118,1; 119,1 y 121,1 en la versión de ocho reactivos o de 114,1 a 117,1 en la de cuatro. El grupo compensador ecualiza las masas de los distintos reactivos, de manera que la masa total de cada reactivo sea la misma. De esta manera, dado que la suma de los pesos moleculares de las tres partes de cada reactivo es constante, un péptido marcado es detectado como un pico único en el espectro MS, pero, como cada grupo reportero difiere en su peso molecular, los fragmentos originados tras la ruptura de cada marcador se observan como picos distintos en el espectro de MS/MS. Así, la concentración relativa de los péptidos se obtiene por comparación de las intensidades de las señales MS/MS de estos grupos reporteros. En el protocolo estandarizado para el marcaje de iTRAQ (Figura 19), las proteínas son primero digeridas con tripsina y posteriormente, los péptidos resultantes se marcan con los distintos reactivos iTRAQ. Después del marcaje peptídico, las muestras se combinan, se fraccionan por ejemplo por cromatografía HPLC y las fracciones resultantes se analizan por nano cromatografía líquida de fase reversa

acoplada a espectrometría de masas (LC-MS/MS). En este tipo de marcaje, el espectro de MS/MS se usa tanto en la identificación como en la cuantificación de las proteínas.

Los reactivos de iTRAQ suponen un gran avance en los métodos de marcaje al permitir el análisis simultáneo de hasta ocho muestras en la versión actual de los mismos. Además, el uso del iTRAQ asegura un completo marcaje de todos los péptidos presentes en la muestra, lo que aumenta la cobertura de las secuencias polipeptídicas identificadas y por lo tanto, la confianza de los resultados obtenidos. El mayor inconveniente que presenta esta técnica es que incorpora una mayor complejidad a la muestra original por lo que el correcto fraccionamiento de la muestra antes del análisis por espectrometría de masas es de gran importancia y afectará en gran medida a la resolución y calidad de los resultados obtenidos.

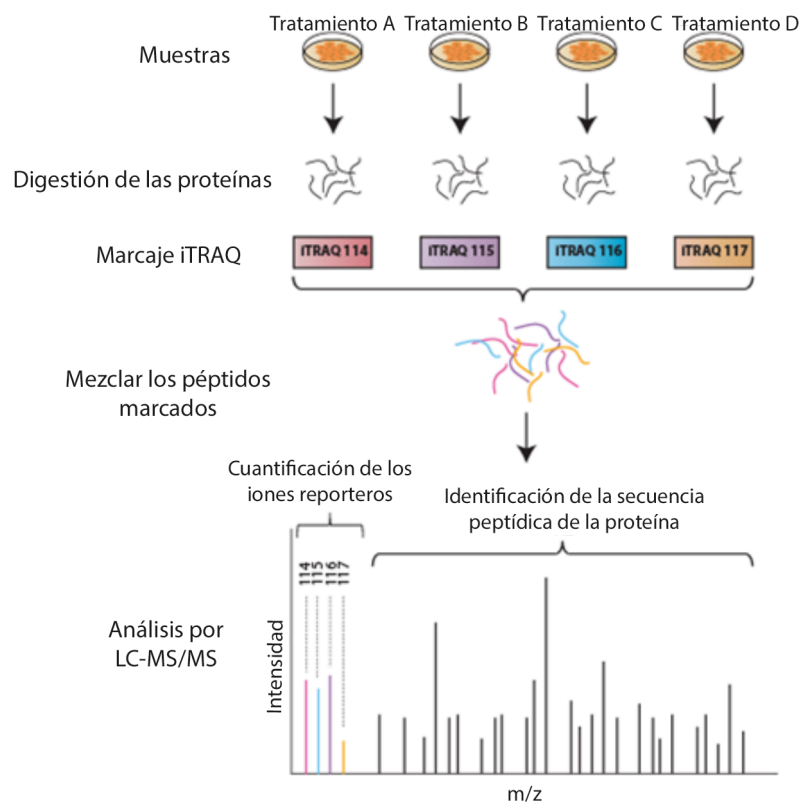


Figura 19. Flujo de trabajo en un experimento iTRAQ.

- **Técnicas de cuantificación sin marcaje (*label-free*)**

Sin embargo, las proteínas también se pueden cuantificar sin ningún tipo de marcaje, es lo que se conoce como técnicas "*label free*". Una de las aproximaciones más utilizadas en *label free* es el conteo espectral [130] (Figura 20). Esta aproximación se basa en la observación de que cuanto más proteína haya en la muestra analizada, más péptidos de esa proteína encontraremos en el análisis de MS. De esta forma, en el conteo espectral, la cuantificación de la abundancia relativa de la proteína se produce contando el número de espectros de MS/MS identificado y acumulado para cada proteína en cada uno de los experimentos de LC-MS/MS. Esta tecnología presenta una menor sensibilidad y precisión en la cuantificación que el marcaje puesto que el espectro depende de la ionización de los analitos y ésta puede ser variable entre todos los experimentos realizados. Sin embargo, es un tipo de cuantificación mucho más económica que el marcaje isotópico y en la actualidad el aumento en la reproducibilidad de los nuevos espectrómetros de masas ha incrementado la fiabilidad de estos experimentos *label-free* [131].

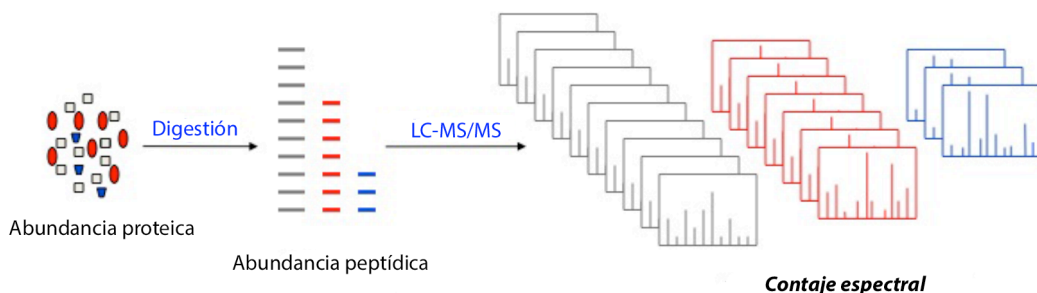


Figura 20. Representación mostrando la aproximación para la cuantificación sin marcaje basada en conteo espectral.

4.2. Técnicas proteómicas basadas en *arrays* de proteínas

Un array de proteínas es una plataforma sólida en la que se inmovilizan de forma ordenada proteínas para ser posteriormente interrogadas con otras moléculas con el fin de detectar interacciones específicas. Un array de proteínas es en su esencia un inmunoensayo en miniatura. El concepto inicial de esta tecnología fue desarrollado por Roger Ekins en 1989 [132], quien propuso que pequeñas cantidades de moléculas de captura inmovilizadas como *microspots* sobre una superficie sólida, capturarían sólo una pequeña fracción del analito específico y que esta fracción reflejaría la abundancia del analito en la muestra independientemente de la concentración de la molécula de captura y el volumen de muestra analizado. Más tarde Ekins y Chu [133, 134] propusieron que esta miniaturización favorecería el análisis de múltiples moléculas simultáneamente.

En la actualidad, los *arrays* de proteínas constituyen una tecnología de alto rendimiento que permite el estudio de forma simultánea y masiva de miles de proteínas. Al contrario que la espectrometría de masas, es una técnica que no requiere la separación previa de la muestra. Por lo tanto, mezclas proteicas complejas y no fraccionadas pueden ser analizadas utilizando los *arrays* de proteínas.

La naturaleza de las proteínas inmovilizadas puede ser muy variada: pueden ser proteínas purificadas como enzimas, péptidos, antígenos o anticuerpos así como también lisados celulares o fluidos biológicos.

Las aplicaciones de los *arrays* de proteínas son múltiples, como el cribado de muestras biológicas para buscar biomarcadores, el estudio de interacciones proteína-proteína, la identificación de sustratos para enzimas o la identificación de proteínas o péptidos diana para fármacos. Todas estas aplicaciones hacen que esta tecnología tenga un enorme potencial en diferentes áreas científicas: en la investigación básica para elucidar la red de interacciones proteicas, en el área clínica (diagnóstico, monitorización del paciente durante el tratamiento, etc.), y en el área farmacéutica para la búsqueda de nuevas drogas.

Por otra parte, el desarrollo actual de los *arrays* de proteínas ofrece un importante avance en la miniaturización y automatización del análisis del proteoma a gran escala. Esta miniaturización favorece la reducción de costes asociado al bajo consumo de reactivos y sobre todo al bajo consumo de muestra biológica sin reducir la sensibilidad del ensayo y permitiendo la detección de proteínas poco abundantes del rango de pg/mL en proteomas no fraccionados.

El principal problema de esta tecnología recae en mantener la estabilidad y funcionalidad de las proteínas inmovilizadas en el array. Las proteínas tienen unas propiedades bioquímicas muy diferentes dependientes de su composición aminoacídica y de sus modificaciones postraduccionales, las cuales van a determinar su estructura tridimensional, su afinidad, su especificidad por otras moléculas así como su estabilidad en el array.

Para solventar este posible inconveniente, en los últimos años han surgido los *arrays* de proteínas sintetizados *in situ* que permiten sintetizar las proteínas en el momento de realización de los ensayos.

Los *arrays* se pueden construir en un formato plano o en formato esfera. El uso de uno u otro formato dependerá del diseño experimental del ensayo concreto. Por otra parte, según el contenido proteico inmovilizado en el array, los *arrays* se pueden clasificar en *arrays* de anticuerpos, *arrays* de antígenos y *arrays* de fase reversa.

4.2.1. Formatos

Las proteínas pueden estar inmovilizadas en *arrays* de formato plano o en formato de esfera.

4.2.1.1. Arrays planos

Los *arrays* en formato plano se generan por inmovilización de una gran cantidad de proteínas en una superficie sólida y plana pudiendo llegar a una densidad espacial de 2000 proteínas/cm² [135].

En general, para los *arrays* planos, sólo son necesarios pequeños volúmenes de muestra de entre 50-500 pL. Este volumen da lugar a *spots* con un tamaño de entre 100-300µm dependiendo de las propiedades del soporte sólido y de la viscosidad de la muestra [136].

Los primeros estudios con este tipo de *arrays* se realizaron en el año 2000 por MacBeath y Schreider [137] y pocos años después, los *arrays* planos de proteínas purificadas ya habían demostrado su utilidad y robustez para analizar proteomas completos [138-140].

Los sistemas de impresión o "*printing*" de las proteínas en el array se pueden dividir en dos categorías: el *printing* de contacto, que usa agujas que tocan la superficie del array para la inmovilización de las proteínas y el *printing* de inyección, que genera gotas que contiene la muestra proteica y que se propulsan y se depositan en la superficie del array sin ningún tipo de contacto. Los dos tipos de tecnologías tienen sus

ventajas e inconvenientes, pero ambas proporcionan una elevada consistencia y rapidez en la creación del array.

El *printing* de contacto es un sistema relativamente sencillo que consiste en agujas que captan la muestra por capilaridad y la depositan en el array. El sistema de lavado para eliminar los restos de la muestra depositada previamente es relativamente simple. Los *printing* de inyección, que utilizan técnicas para la generación y propulsión de las gotas en la superficie del array, tales como la tecnología piezoeléctrica, son técnicamente más complejos y requieren sistemas de lavado más laboriosos pero sin embargo, cuando funcionan correctamente, ofrecen una alta velocidad de *printing* y la formación de gotas es muy consistente y reproducible entre lotes de *arrays*. Sin embargo, comparado con los *printing* de contacto, los de inyección están bastante limitados por el rango de viscosidad de la muestra con el que pueden trabajar; de hecho, pequeñas variaciones en las propiedades fluídicas de las muestras, pueden hacer que se tengan que optimizar y ajustar los parámetros que controlan la formación y deposición de la gota (voltaje, presión, frecuencia, etc) entre las muestras, ralentizando finalmente el tiempo del *printing*.

Aparte del sistema de *printing*, los parámetros más importantes que hay que tener en cuenta para la construcción de *arrays* planos es la selección de la superficie y la estrategia utilizada para inmovilizar las proteínas.

Existe una gran número de superficies sólidas disponibles para construir *arrays* planos, que van desde la nitrocelulosa a las plataformas de oro o vidrio [141]. Entre todas estas, las de vidrio son las más utilizadas actualmente en la construcción de *arrays* planos.

Las estrategias de inmovilización pueden ser clasificadas en físicas, covalentes o de afinidad [142-144]. Las superficies de nitrocelulosa y las superficies de vidrio activadas con compuestos con cargas positivas como el aminosilano permiten una unión física por absorción de la proteína a la membrana. Ésta es una unión débil que puede provocar que las proteínas no se orienten correctamente y que pierdan su funcionalidad para el ensayo.

Por otra parte, las superficies tratadas con epoxi permiten una unión covalente y, por lo tanto, fuerte con las proteínas a través de su extremo amino (-NH₂) o hidroxilo (-OH).

Las superficies tratadas por ejemplo con la avidina permiten una unión no covalente pero específica con aquellas proteínas que lleven una etiqueta con biotina en alguno de sus extremos.

En general, se puede decir que no hay un soporte o una química universal que sea adecuada para todos los ensayos que requieran el uso de los *arrays* de proteínas y es necesario dedicarle un tiempo considerable a la optimización de las condiciones para la realización del ensayo concreto.

Otro factor importante a tener en cuenta con el fin de minimizar las uniones inespecíficas es la elección del sistema de bloqueo del array. En este sentido, se suelen utilizar soluciones de bloqueo clásicas como la albúmina de suero bovino (BSA) o la leche en polvo a diferentes concentraciones.

Otros factores que también debemos de tener presentes a la hora de fabricar *arrays* planos es la solución en la que resuspendemos las proteínas que vamos a imprimir en el array de modo que ésta mantenga la estabilidad de la proteína en el soporte; y las condiciones de humedad durante el *printing* de forma que se evite, en la medida de lo posible, la evaporación de la muestra [145].

Finalmente, debemos evaluar la reproducibilidad y calidad del array. Para esto es necesario incluir determinados controles en el diseño del array como controles positivos y negativos del *printing*, controles de hibridación con la muestra, etc.

La generación y la detección de la señal se puede realizar utilizando diferentes técnicas: fluorescencia, quimioluminiscencia o métodos sin marcaje como el quantum dots o el plasmón de resonancia [146, 147]. Sin embargo, los métodos basados en fluorescencia son los más utilizados puesto que la señal ofrece una elevada sensibilidad y amplio rango dinámico. La otra razón principal es que existen escáneres compatibles con fluorescencia que también se usan en los *arrays* de DNA y se han heredado e implementado también en el campo de los *arrays* de proteínas. Los fluoróforos más utilizados son las cianinas Cy3 y Cy5, Alexa Fluor y el complejo R-ficoeritina (R-PE)

La gran ventaja de los *arrays* en formato plano es que permiten analizar una gran número de analitos en la misma superficie. Sin embargo, esta ventaja a su vez limita el número de muestras que pueden ser analizadas en el mismo ensayo. Por esta razón, los *arrays* en formato plano se suelen utilizar como herramienta para el descubrimiento e identificación de nuevos biomarcadores evaluados en un pequeño número de muestras. Sin embargo, los resultados deben de ser verificados con otras tecnologías con una mayor capacidad para analizar un gran número de muestras simultáneamente, tales como los *arrays* de proteínas basados en esferas.

4.2.1.2. Arrays en esferas

El principio de los *arrays* en esfera, o también llamados *arrays* en suspensión, consiste en inmovilizar las proteínas en conjuntos de superficies circulares individuales y diferenciales que actúan como soportes sólidos y son separadas e identificadas usando un sistema tipo citometría de flujo.

Las esferas son soportes sólidos que llevan grupos funcionales como carboxilos (-COOH) o tioles (-SH) que favorecen la unión de proteínas como pueden ser antígenos o anticuerpos.

El uso de este tipo de plataformas ya fue descrito hace 40 años para inmovilizar antígenos con el fin de identificar anticuerpos específicos frente a esos antígenos en muestras de suero [148]. Posteriormente, en la década de los 90, se incluyeron diversos fluoróforos dentro de las esferas para crear hasta 64 poblaciones de esferas diferentes [149].

El último avance en esta tecnología fue creado por la compañía Luminex que comercializa las microesferas fluorescentes para realizar ensayos múltiples bajo el nombre XMap® technology.

Actualmente, la tecnología XMap es la más empleada para producir *arrays* de proteínas y se basa en construir el *array* utilizando unos soportes sólidos de poliestireno en formato microesfera, los cuales llevan internamente un color fluorescente que es el resultado de la mezcla de dos o tres fluoróforos diferentes con distintos perfiles de emisión. Este sistema hace que cada población de esferas pueda ser distinguida por su color, acoplada con una proteína (antígeno o anticuerpo) concreta y combinada posteriormente con otras poblaciones de esferas fluorescentes para formar un array de esferas en suspensión y poder así analizar varios analitos simultáneamente.

Actualmente, la tecnología Luminex posee una combinación de fluoróforos que permite analizar hasta 500 analitos en el mismo array y pueden analizar hasta 384 muestras simultáneamente.

La tecnología XMap® contiene esferas estándar, no magnéticas, con un diámetro de 5,6µm que son las esferas Microplex, pero también puede usar la MagPlex, que son esferas magnéticas con un diámetro de 6.5µm. Las esferas magnéticas, al contrario que las estándar, incrementan la automatización del proceso de creación del array, lo que es muy importante en términos de reproducibilidad del ensayo.

En los dos casos, las microesferas se activan mediante el uso de compuestos carboxilados de modo que las proteínas se van a unir de forma covalente a las esferas a través de sus grupos amino.

El método de detección preferiblemente utilizado en los *arrays* en esfera es también la fluorescencia. Para la lectura del ensayo se utiliza un sistema basado en la citometría de flujo en la que hay dos láseres, uno de los láser con una longitud de onda de 650nm excita el fluoróforo interno de la esfera, identificando de esta forma el analito que estamos detectando, y con el otro láser a 532 nm identificamos la intensidad de fluorescencia media (MFI) del fluorocromo reportero, normalmente se usa la R-PE, que indica la cantidad de analito unido a la esfera.

Con el fin de evaluar la fiabilidad y sensibilidad de esta tecnología, en los últimos años, varios estudios han comparado la tecnología de las esferas con los ELISA, la tecnología *estándar* en inmunoensayos para validar biomarcadores. En estos ensayos se midieron diferentes citoquinas y se ha visto que las dos tecnologías tienen una alta correlación de señal y que incluso para la tecnología de detección múltiple con las esferas, la sensibilidad es mayor que en el ELISA [150-152]. Estas dos tecnologías también se han comparado para detectar autoanticuerpos en el suero y se ha visto que no hay diferencias de sensibilidad entre una y otra, pero que los *arrays* en esfera consumían menos muestra y era una tecnología más rápida que el ELISA [153, 154].

Al contrario que los *arrays* en formato plano, los *arrays* en esfera no requieren escáneres para la lectura del ensayo, *softwares* de análisis de imagen o sofisticados equipos de *printing*. Además, es una tecnología muy versátil y flexible que permite diseñar de forma rápida y sencilla un array para su uso en investigación y la práctica clínica.

4.2.2. Clasificación

4.2.2.1. Arrays de anticuerpos

Los *arrays* de anticuerpos son *arrays* analíticos cuya principal aplicación es la detección de proteínas diferencialmente expresadas y su abundancia en muestras biológicas complejas con el fin de buscar biomarcadores proteicos.

Estos *arrays* se generan por inmovilización de anticuerpos en la superficie del array (plano o en esfera), los cuales son utilizados como moléculas de captura de una diana concreta.

- **Anticuerpos: definición y clasificación**

Los anticuerpos se pueden definir como proteínas especializadas para funcionar como reactivos de afinidad por otras moléculas (Figura 21).

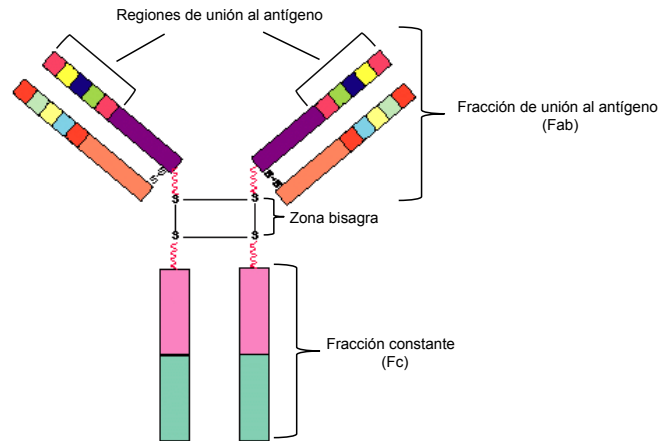


Figura 21. Estructura de un anticuerpo del tipo inmunoglobulina G (IgG). Esta inmunoglobulina contiene el lugar que reconoce elementos extraños específicos (antígeno). Esta región del anticuerpo se llama fracción de unión al antígeno o región Fab. Está compuesta de un dominio constante y otro variable de cada una de las cadenas del anticuerpo. Contiene también una fracción constante (Fc) que participa en la modulación de la respuesta inmune. Entre estas dos regiones se encuentra la zona bisagra formada por puentes disulfuro.

Los anticuerpos que se utilizan en investigación son normalmente generados por el sistema inmunitario de un animal hospedador al que se le inmuniza con una proteína desconocida por ese organismo que es lo que se conoce como inmunógeno o antígeno [155]. Los antígenos pueden ser proteínas completas o fragmentos proteicos.

Las células que responden a esos antígenos son los linfocitos B o células B. Cada célula B produce un anticuerpo específico frente a una parte del antígeno que es lo que se conoce como epítipo. Puesto que hay diferentes células B en el organismo, el suero del animal inmunizado contendrá una mezcla de anticuerpos frente al antígeno. Estos anticuerpos reciben el nombre de policlonales puesto que reconocen diferentes epítipos del antígeno. Sin embargo, es posible crear una línea celular con clones de células B que reconozcan un único epítipo del antígeno.

Para ello, se seleccionan clones concretos de células B y se fusionan con una línea celular de mieloma dando lugar a lo que se conoce como hibridoma [156]. Este hibridoma tiene una capacidad ilimitada de producir anticuerpos idénticos frente a un

epítipo concreto. Los anticuerpos generados por la tecnología del hibridoma reciben el nombre de anticuerpos monoclonales.

La producción de los anticuerpos monoclonales es más largo y costoso que la producción de anticuerpos policlonales pero sin embargo, son más específicos y no hay variabilidad entre los lotes de anticuerpo producido, al contrario de lo que ocurre cuando se trabaja con anticuerpos policlonales [157].

Para los *arrays* de anticuerpos se pueden utilizar tanto anticuerpos monoclonales como policlonales. La elección de uno u otro dependerá de si estamos interesados o no en determinar un epítipo concreto de las proteínas diana.

Los epítipos que reconocen los anticuerpos, tanto monoclonales como policlonales, pueden ser divididos en dos clases: epítipos lineales (continuos), formados por una secuencia aminoacídica consecutiva o epítipos conformacionales (discontinuos), los cuales se forman cuando residuos aminoacídicos distantes se aproximan debido al plegamiento de la proteína [158] (Figura 22).

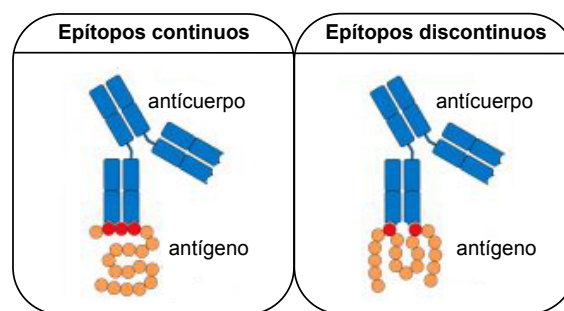


Figura 22. Tipos de epítipos que pueden reconocer los anticuerpos.

Por lo tanto, es importante tener en cuenta el diseño del anticuerpo y adaptar las condiciones experimentales al ensayo en base a ello, puesto que si el anticuerpo reconoce epítipos lineales, se requerirá la desnaturalización previa de las proteínas para linearizar los epítipos y que el anticuerpo tenga acceso al fragmento antigénico frente al cual se ha generado [159].

En los últimos años, también se han empezado a utilizar los anticuerpos recombinantes en la producción de *arrays* de anticuerpos. Los anticuerpos recombinantes son anticuerpos producidos *in vitro*. Se generan clonando y expresando en bacterias los fragmentos variables de cadena sencilla (scFv) o los fragmentos de unión al antígeno (Fab) del anticuerpo mediante el método conocido como *phage display* [160, 161]. Es un método que no requiere el uso de animales para la

generación de los anticuerpos y es un método más económico para producir moléculas de afinidad para inmovilizar en los *arrays*. Sin embargo, el proceso de selección de los fragmentos con ScFv o Fab con afinidad específica para un antígeno es un proceso muy largo y laborioso.

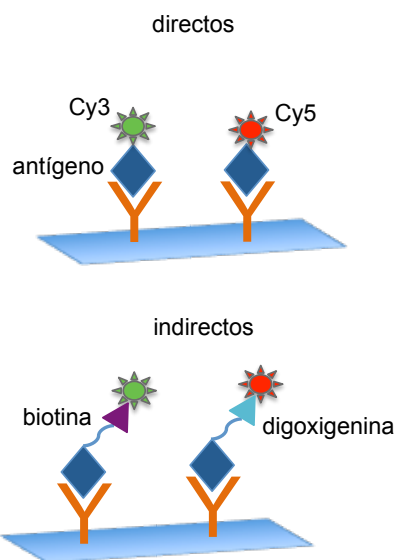
- **Tipos de *arrays* de anticuerpos**

Los *arrays* de anticuerpos se pueden presentar en varios formatos dependiendo del número de anticuerpos utilizados para detectar la proteína diana y de los procesos de marcaje e hibridación de la muestra a analizar.

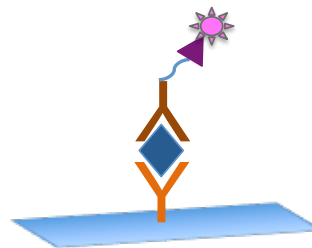
Según esto, los *arrays* de anticuerpos se pueden clasificar en directos, si usan un único anticuerpo para detectar la proteína concreta, o *arrays* tipo “sándwich”, si usan dos anticuerpos para capturar la proteína diana [162] (Figura 23).

En los *arrays* directos, las proteínas son capturadas por un único anticuerpo inmovilizado en el soporte y la señal de la proteínas se detecta a través del marcaje de la muestras. El marcaje puede ser de dos tipos, directo o indirecto.

A. Planos competitivos



B. Planos tipo “sándwich”



C. Esferas en suspensión

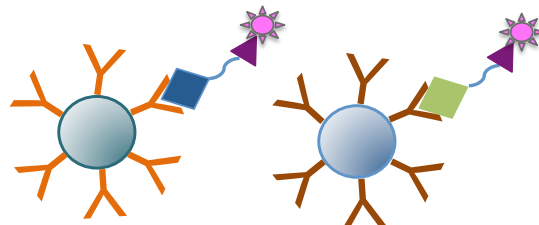


Figura 23. Formatos de *arrays* de anticuerpos

En el marcaje directo, las proteínas se marcan normalmente con un fluorocromo tipo cianina (Cy3 o Cy5). 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 que tienen afinidad por la molécula intermedia. Con esta última aproximación se consigue amplificar la señal de la proteína detectada.

Los *arrays* que usan un sólo anticuerpo permiten la incubación de dos muestras diferentes simultáneamente, cada una marcada con un fluorocromo distinto. Son por lo tanto ensayos competitivos, es decir, las proteínas de las dos muestras compiten por su unión con el anticuerpo. La ventaja de estos ensayos es que conducen a la mejora en la linealidad y el rango de la señal comparado con los ensayos no competitivos.

El número de proteínas potencialmente determinables en este tipo de *arrays* va a depender de la disponibilidad de anticuerpos y del espacio disponible en el soporte donde se inmovilicen los anticuerpos.

La principal desventaja de los *arrays* directos está relacionada con la estabilidad del marcaje, lo cual podría limitar la detección de la señal, así como la sensibilidad y especificidad del ensayo.

En el formato “sándwich”, los anticuerpos inmovilizados capturan proteínas no marcadas. En estos *arrays*, la proteína diana es detectada usando dos anticuerpos, un anticuerpo inmovilizado en el soporte y un anticuerpo secundario marcado que dará lugar a una señal (fluorescente, quimioluminiscente, etc) que se corresponderá con la abundancia de la proteína en la muestra.

El uso de dos anticuerpos dirigidos frente a dos epítopos diferentes de cada proteína incrementa la especificidad del ensayo comparado con los *arrays* que usan sólo un anticuerpo. Esta especificidad se refleja en un reducido ruido de fondo, lo que provoca que se incremente la sensibilidad del ensayo [162].

Otra ventaja es que en los ensayos tipo “sándwich” se puede hacer una cuantificación absoluta de la proteína en la muestra. Para ello, es necesario tener proteínas o antígenos diana purificados con los que realizar una curva estándar de concentraciones conocidas. Sin embargo, muchas veces estos estándares no están disponibles comercialmente y no es posible realizar una cuantificación absoluta de las proteínas. El array en formato “sándwich” no permite ensayos competitivos, es decir, solamente una muestra puede ser incubada sobre cada array.

Las dos principales desventajas de estos *arrays* es que requieren pares de anticuerpos específicos para cada proteína, lo que muchas veces se hace muy difícil

de encontrar, y que existe un límite práctico de 40 proteínas que pueden ser medidas simultáneamente debido a la interferencia entre los diferentes anticuerpos utilizados [162]. Debido a estos inconvenientes, los *arrays* tipo sándwich son más difíciles de desarrollar de forma múltiple.

La precisión y fiabilidad de los *arrays* de anticuerpos depende de la afinidad y la especificidad de los anticuerpos por sus epítomos correspondientes. Un posible problema que puede surgir cuando se trabaja con anticuerpos es la reactividad cruzada, es decir, que el anticuerpo esté reconociendo la proteína equivocada [163]. Muchas proteínas son similares en sus secuencias aminoacídicas primarias y hay muchas proteínas con funciones similares que comparten dominios de unión muy parecidos. Si un anticuerpo está dirigido frente a estas regiones, pueden unirse a la proteína errónea, generando resultados no fiables y conclusiones erróneas. Normalmente, la unión inespecífica es mucho más débil que la unión con la proteína específica cuando las dos proteínas se encuentran en concentraciones equimolares. Sin embargo, si la proteína diana se encuentra en menores proporciones que la errónea, el resultado será que el anticuerpo se va a unir a la proteína equivocada.

El factor más limitante para el uso de *arrays* de anticuerpos es que, desafortunadamente, sólo un pequeño porcentaje de anticuerpos frente a proteínas humanas conocidas están bien caracterizados. Esto limita el uso y aplicación de los *arrays* de anticuerpos en investigación para la búsqueda de biomarcadores.

- ***Human Protein Atlas (HPA)***

Con el fin de producir anticuerpos bien caracterizados para su uso en investigación, la *Human Proteome Organization* (HUPO) creó la iniciativa *Human Protein Atlas (HPA)* [164]. Esta iniciativa tiene el objetivo de generar al menos dos anticuerpos contra cada producto proteico codificado por el genoma humano con el fin de crear un mapa de la localización de todas las proteínas humanas a nivel celular y tisular.

Los anticuerpos del HPA son anticuerpos policlonales que reconocen epítomos lineales y que se generan por inmunización de conejos con fragmentos de proteínas recombinantes de entre 50-150 aminoácidos que tiene una homología menor al 20% con otras proteínas humanas.

El flujo de trabajo del HPA (Figura 24) comienza con seleccionar *in silico* los fragmentos proteicos [165-167]. Estos fragmentos reciben el nombre de *Protein Epitope Signature Tag* (PrEST) [168]. Los PrEST poseen un fragmento de unión a la

albúmina, que se utiliza para aumentar la inmunogenicidad del fragmento, y una cola de seis histidinas en el extremo aminoterminal de la secuencia aminoacídica para favorecer la posterior purificación de los fragmentos.

Para generar los anticuerpos, los PrEST se expresan en bacterias *E.Coli* y posteriormente, éstos se purifican y se utilizan para inmunizar a los conejos. Finalmente, se extrae el suero de los conejos, donde se encuentran los anticuerpos policlonales frente a los distintos fragmentos de la proteína recombinante. Los anticuerpos se purifican en columnas cromatográficas de inmunoafinidad y posteriormente se validan por diferentes metodologías (inmunohistoquímica, *western blot*, inmunofluorescencia y *arrays* de antígenos tipo PrEST) [168].

En la purificación se utilizan los mismos PrEST frente a los que fue generado el anticuerpo. Por esta razón a estos anticuerpos también se les conoce con el nombre de anticuerpos mono-específicos.

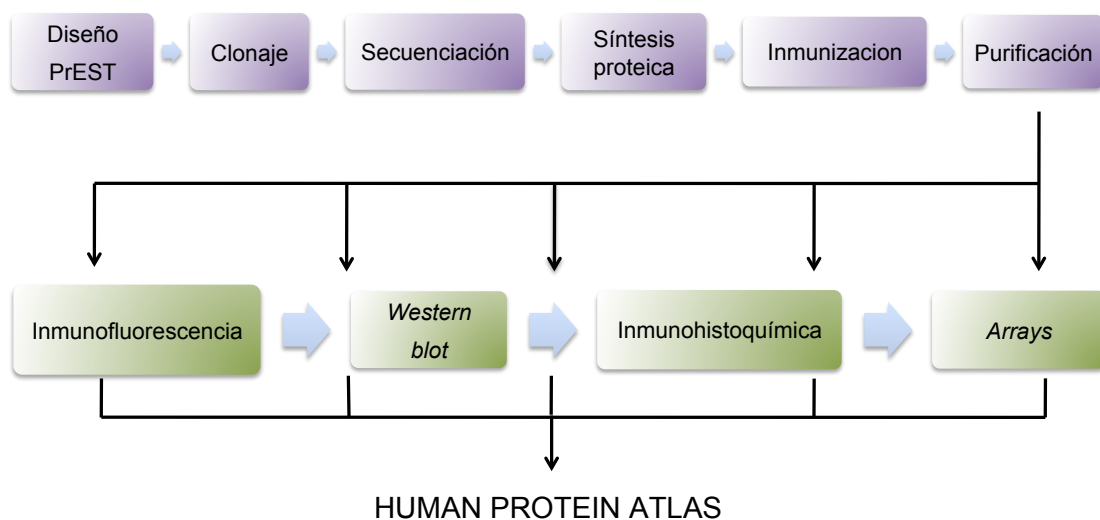


Figura 24. Flujo de trabajo del *Human Protein Atlas* (HPA).

4.2.2.2 Arrays de antígenos

Los *arrays* de antígenos, tanto planos como en esfera, se construyen para identificar interacciones con otras moléculas. Los antígenos que se inmovilizan en el array suelen ser proteínas completas o fragmentos proteicos.

Usando este tipo de *arrays* se pueden realizar tanto ensayos analíticos, para determinar la abundancia de una proteína en la muestra biológica de interés, como funcionales, donde los antígenos son inmovilizados en el array para ver uniones del

antígeno con otras moléculas como interacciones proteína-proteína, proteína- fármaco, proteína-péptido o proteína-ácido nucleico [143, 169].

Los *arrays* de antígenos analíticos que son los que se han utilizado en esta tesis, se emplean fundamentalmente para detectar anticuerpos frente al propio antígeno en la muestra biológica de interés.

- **Estrategias para construir *arrays* de antígenos**

Las proteínas son moléculas muy variadas debido a su composición aminoacídica, plegamiento, conformación, estado de oligomerización y modificaciones postraduccionales presentes en la proteína. Todas estas características hacen que las proteínas tengan propiedades bioquímicas exclusivas que determinarán su características funcionales.

Por todo esto, el primer paso crítico para construir *arrays* de proteínas es disponer las proteínas en el *array* manteniendo la estabilidad y las propiedades funcionales de la proteínas inmovilizadas.

Para fabricar los *arrays* de antígenos se pueden utilizar fundamentalmente dos estrategias: las proteínas se pueden producir *in vivo*, es decir, se pueden inmovilizar en el *array* después de expresarlas en bacterias como proteínas recombinantes y purificarlas, o bien pueden ser expresadas *in situ* sobre el *array*, usando sistemas de expresión libres de células.

La forma clásica de construir *arrays* de proteínas utiliza sistemas de expresión basados en células de forma que la región codificante (ORF, *Open Reading Frame*) de la proteína de interés se clona en bacterias, se expresa *in vivo* como una proteína recombinante y finalmente se purifica y se inmoviliza en el *array*. La fiabilidad de producir grandes cantidades de proteínas expresadas *in vivo* e inmovilizarlas posteriormente en el *array* fue demostrada por Zhu *et al* en el año 2001 en un trabajo en el que construyeron un *array* con 6000 proteínas de levadura expresadas como proteínas de fusión con la etiqueta GTS-His₆.

En la actualidad, ya existen varias compañías que distribuyen *arrays* de proteínas recombinantes humanas en formato plano. Uno de los *arrays* comerciales más conocidos es el *ProtoArrayHuman@Protein Microarray*, que contiene en su versión actual 5.0, 9000 proteínas humanas completas en cada *array* expresadas como proteínas recombinantes con la etiqueta GST. Recientemente, también la compañía *CDI Laboratories* ha comercializado el *HuProt@ Human Proteome Microarray*, que en su última versión contiene 19,832 proteínas humanas completas, lo que se

corresponde ya con el 75% de los genes humanos codificantes para proteínas. Sin embargo, la gran desventaja de estas plataformas comerciales es su elevado coste.

Por otra parte, los fragmentos proteicos humanos generados en el HPA, han demostrado también su fiabilidad para producir antígenos humanos a gran escala para utilizarlos en investigación [170]. El HPA tiene aproximadamente 46.000 fragmentos proteicos humanos verificados, que representan el 80% de los genes que codifican para proteínas humanas. Como se comentó previamente en la sección anterior (4.2.2.2), el flujo de trabajo del HPA conlleva el clonaje en *E.Coli* de fragmentos proteicos humanos de 50-150 aminoácidos que se expresan como fragmentos proteicos de fusión (PrEST). Estos PrEST se utilizan para producir anticuerpos pero después también son utilizados para generar *arrays* de antígenos en formato plano. Estos *arrays* están formados por 384 antígenos seleccionados de forma aleatoria y son utilizados de rutina en el flujo de trabajo del HPA para comprobar la correcta unión del anticuerpo generado frente a su PrEST.

Sin embargo, pese a todas estas iniciativas y plataformas comerciales, la producción de grandes cantidades de proteínas mediante expresión *in vivo* es un proceso muy laborioso, costoso y su rendimiento es muy variable dependiendo de la proteína objeto del estudio.

Por otra parte, los *arrays* de antígenos también pueden ser sintetizados utilizando péptidos sintéticos. Los *arrays* de péptidos comparten muchas características con los *arrays* de proteínas pero en lugar de estudiar la unión de proteínas completas o fragmentos grandes, en estos *arrays* se generan péptidos de entre 6-25 aminoácidos.

Tradicionalmente, los *microarrays* de péptidos se producen usando la tecnología SPOT™ [171] donde gotas que contienen un aminoácido de los 20 existentes, son depositadas en posiciones específicas de membranas de celulosa funcionales seguido por el lavado de los aminoácidos no unidos y la eliminación de grupos de protección para permitir la unión del siguiente aminoácido al péptido en formación. De esta forma, se generan ciclos repetidos para generar un *array* de péptidos de secuencia conocida en posiciones específicas del soporte [172]. Los *arrays* de péptidos son ampliamente utilizados para el descubrimiento de nuevas drogas. La ventaja de este tipo de *arrays* es que son químicamente más estables que los *arrays* de proteínas o fragmentos proteicos y que su fabricación es más económica. La principal limitación de los *arrays* de péptidos es que sólo permiten detectar epítomos continuos.

En la actualidad es posible evitar los inconvenientes de la expresión de proteínas *in vivo* y las limitaciones de los *arrays* de péptidos, construyendo *arrays* de proteínas *in situ*, también conocidos como *arrays* auto- ensamblados.

- ***Arrays de proteínas in situ***

En este caso, para construir el *array*, se utilizan los DNAs que codifican para cada proteína. Los DNAs pueden estar insertados dentro de un plásmido o pueden ser productos de PCR.

Para ello, el DNA se imprime en el *array* y se sintetiza el producto proteico a partir de un sistema que contiene toda la maquinaria para la transcripción del DNA a RNAm (factores de iniciación, elongación y terminación) y ribosomas para la traducción de la proteína. Esta mezcla de compuestos es externamente suplementada con otros componentes como aminoácidos, sales y cofactores que son necesarios para la producción de las proteínas.

Los sistemas de expresión *in situ* fueron descubiertos por Nirenberg y Matthaei en 1961 [173] y desde entonces, han sido ampliamente utilizados en la comunidad científica para solventar los problemas de la expresión *in vivo* de las proteínas. Los primeros sistemas de expresión *in situ* se generaron a partir de las bacterias *E.Coli*, seguido más tarde por sistemas procedentes de células eucariotas [174-176]. Los sistemas de bacterias pueden producir hasta miligramos de proteína pero, al contrario que los sistemas eucariotas, están limitados a la producción de proteínas con un tamaño menor a 65 KDa y no producen modificaciones postraduccionales de la proteínas expresadas [177].

Los sistemas de expresión eucariotas pueden ser de germen de trigo, de insecto, reticulocitos de conejo o lisados de células humanas [177]. Aunque el reticulocito de conejo produce menos proteína que otros sistemas (0.2ug/10uL de reacción) y es bastante caro, es el sistema *in vitro* más utilizado. En gran parte, esto es debido a que la producción de la proteína con este sistema es muy rápida, produce proteínas con un amplio rango de tamaños distintos, y crea modificaciones postraduccionales en las proteínas expresadas. Además, el reticulocito de conejo contiene también chaperonas de modo que la proteína se puede expresar en su conformación tridimensional. La principal desventaja de este sistema es que cada vial de reticulocitos proviene de un conejo distinto, de modo que puede haber variabilidad de un lote a otro de reticulocitos y por lo tanto, una variabilidad de eficiencia en la producción de las proteínas [177]. Por otro lado, el desarrollo de sistemas de expresión

basados en líneas celulares humanas ha permitido solventar el inconveniente del sistema basado en reticulocitos minimizando las diferencias en la eficiencia de producción entre ensayos.

Desde el punto de vista de la investigación en la búsqueda de biomarcadores, los *arrays* de proteínas *in situ* ofrecen una alternativa rápida y eficaz para la detección múltiple de proteínas con un bajo consumo de reactivos.

Hay diferentes tipos de *arrays* de proteínas *in situ* (Figura 25):

PISA

PISA (del inglés, *Protein In Situ Array*) fue la primera estrategia desarrollada como un *array* de proteínas *in situ* y fue diseñado por primera vez por He y Taussing en el año 2001[178].

Este método utiliza un DNA amplificado por PCR que contiene un promotor transcripcional fuerte y una secuencia corta que codifica para una etiqueta en el extremo amino o carboxi-terminal de la proteína expresada. En PISA, la superficie del *array* se trata con una molécula para capturar la etiqueta de la proteína expresada y después de la expresión utilizando el sistema de *E.Coli* o el de reticulocitos de conejo, las proteínas sintetizadas son directamente capturadas en el *array*. El método PISA original era un sistema que incluía 24 pocillos y que requería el uso de 25 μ L de la mezcla DNA-lisado por pocillo. Sin embargo, más tarde, se miniaturizó la tecnología utilizando el sistema de deposición múltiple [179], de modo que podían imprimir en pequeños *spots* primero el DNA y después el sistema de expresión, reduciendo el volumen requerido a nanolitros y aumentando la densidad teórica del *array* a 13000 *spots*.

Sin embargo, hasta la fecha y con esta estrategia, sólo han podido demostrar la expresión simultánea de 384 proteínas funcionales.

DAPA

Esta técnica fue desarrollada por He *et al* en el año 2008 [180]. DAPA (del inglés, *DNA Array to Protein Array*) es una técnica derivada del PISA pero permite reutilizar el DNA molde para imprimir hasta 20 copias del mismo *array* de proteínas. El DAPA requiere un sistema formado por dos plataformas y una membrana permeable en el medio. Una de las plataformas está inmovilizada con los DNAs producto de la PCR que van a codificar para la proteína recombinante, la otra plataforma tiene el agente de captura para la proteína recombinante expresada y entre las dos hay una

membrana en el que se deposita el sistema de expresión *in vitro*. De esta forma, la proteína de fusión producida difunde a través de la membrana y es capturada en la otra plataforma por el agente de captura. A parte de que es un proceso largo, la principal desventaja de este sistema es que la difusión puede limitar la síntesis de proteínas multiméricas. Debido también a los efectos de la difusión, los *spots* ocupan mucho espacio de en *array* y los *arrays* más extensos que se han publicado hasta el momento utilizando DAPA cuentan solamente con un centenar de proteínas inmovilizadas [181].

PuCa

PuCa (del inglés, *Puromycin Capture protein array*) es un *array* de proteínas autoensamblado desarrollado por Tao y Zhu en el año 2006 [182] que se basa en la afinidad de la puromicina por las proteínas o los péptidos que se expresan directamente en el *array*.

Para este método, en primer lugar, el DNA se transcribe a RNAm *in vitro*. El extremo 3' del RNAm se inmoviliza en la superficie del *array* con una secuencia de oligonucleótidos de cadena sencilla (ssDNA) y con otra ssDNA complementaria que contiene biotina y puromicina. La biotina sirve para inmovilizar el RNAm en la superficie del *array* tapizada con avidina mientras que la puromicina sirve de anclaje para la proteína recién expresada por el sistema de expresión *in situ* que se añade al *array*.

Sin embargo, esta estrategia no ha tenido mucha utilidad puesto que ofrece dificultades para expresar proteínas con un peso molecular mayor de 60 KDa y generalmente el rendimiento de producción proteica es muy bajo [143].

NAPPA

Las estrategias mencionadas hasta el momento, solo han sido testadas con un bajo número de proteínas comparado con los *arrays* de proteínas recombinantes sintetizadas *in vivo*. Por lo tanto, no han demostrado la robustez necesaria para producir *arrays* de alto contenido proteico.

En el año 2004, Labaer y colaboradores desarrollaron un *array* de proteínas autoensamblado de alta densidad conocido como *Nucleic Acid Programmable Protein Array* (NAPPA) [183]. Esta tecnología se basa en imprimir sobre el *array* el DNA codificante para la proteína que está insertado dentro de un vector de expresión mediante la tecnología Gateway®. Este vector contiene un promotor transcripcional, una secuencia codificante para un péptido/proteína que será la etiqueta que se exprese

en el extremo carboxiterminal de la proteína y un vector de resistencia a ampicilina para amplificar el vector en bacterias.

El clonar un inserto de DNA en un vector especializado requiere una inversión de tiempo mayor que el realizar una simple PCR, como es el caso de la tecnología PISA. Sin embargo, tiene varias ventajas: (1) una vez el clon es generado y almacenado adecuadamente, se convierte en una fuente renovable e inagotable de plásmido que puede ser compartida con otros laboratorios, (2) puesto que el clon es verificado por secuenciación, la secuencia será estable y válida durante largos períodos de tiempo, (3) en el NAPPA, el epítipo etiqueta se expresa de forma constitutiva en el vector lo que hace que el epítipo sea estable y funcional en el momento de la expresión [184].

En la mayoría de los NAPPA, las proteínas expresadas llevan la etiqueta *Glutathione S-Transferase* (GST) pero también se pueden utilizar otras etiquetas como Flag, HA, c-myc y Halo.

En el NAPPA, el DNA codificante se clona en un vector de expresión (<http://DNASU.org>) y se transforma en bacterias. A continuación, las bacterias transformadas se cultivan a gran escala y son seleccionadas por su resistencia al antibiótico ampicilina. Posteriormente, las bacterias se lisan y el DNA plasmídico se purifica utilizando una cromatografía de intercambio iónico. Esta cromatografía permite obtener plásmidos de la elevada pureza requerida para los ensayos NAPPA. Finalmente, los plásmidos se imprimen sobre una superficie sólida activada con aminosilano junto con un intercalante homobifuncional (BS³), BSA y un anticuerpo anti-GST. En esta mezcla, cada componente tiene un papel definido: el BSA aumenta la eficiencia de la unión del DNA a la plataforma y reduce las uniones inespecíficas, el anti-GST es el anticuerpo de captura que se va unir a la proteína recién expresada y el BS³ ayuda a la orientación del anticuerpo de captura en el *array*. Cuando el sistema de expresión *in situ* es añadido al *array*, se produce una reacción en la que se acopla la transcripción y la traducción proteica. Después de la expresión, la proteína se inmoviliza en el *array* mediante el anticuerpo anti-GST que captura la etiqueta GST expresada en el extremo carboxiterminal de la proteína, de forma que se asegura la completa traducción de la proteína sobre el *array*.

El método creado por Labaer y colaboradores ha permitido construir *arrays* de proteínas de hasta 1000 genes humanos clonados en plásmidos de expresión [184]. En este ensayo realizado en el año 2008, demostraron que el 96% de los genes impresos en el *array* codificaban para proteínas y que eran capaces de expresar una amplia

variedad de proteínas diferentes, desde proteínas solubles hasta proteínas de membrana [184]. Comprobaron también que el tamaño de la proteína no era un factor limitante para esta tecnología puesto que con el NAPPA conseguían expresar el 98% de las proteínas menores de 50KDa y el 88% de las proteínas mayores de 100KDa [184]. Por lo tanto, el NAPPA es una tecnología fiable y versátil que puede ser aplicada a diferentes ensayos proteómicos tanto analíticos como funcionales [185].

Hasta la fecha, los NAPPA se ha utilizado principalmente para la búsqueda de autoanticuerpos biomarcadores de enfermedades como el cáncer de mama [186, 187] o enfermedades autoinmunes como diabetes tipo I [188], espondilitis anquilosante [189] y artritis reumatoide juvenil [190].

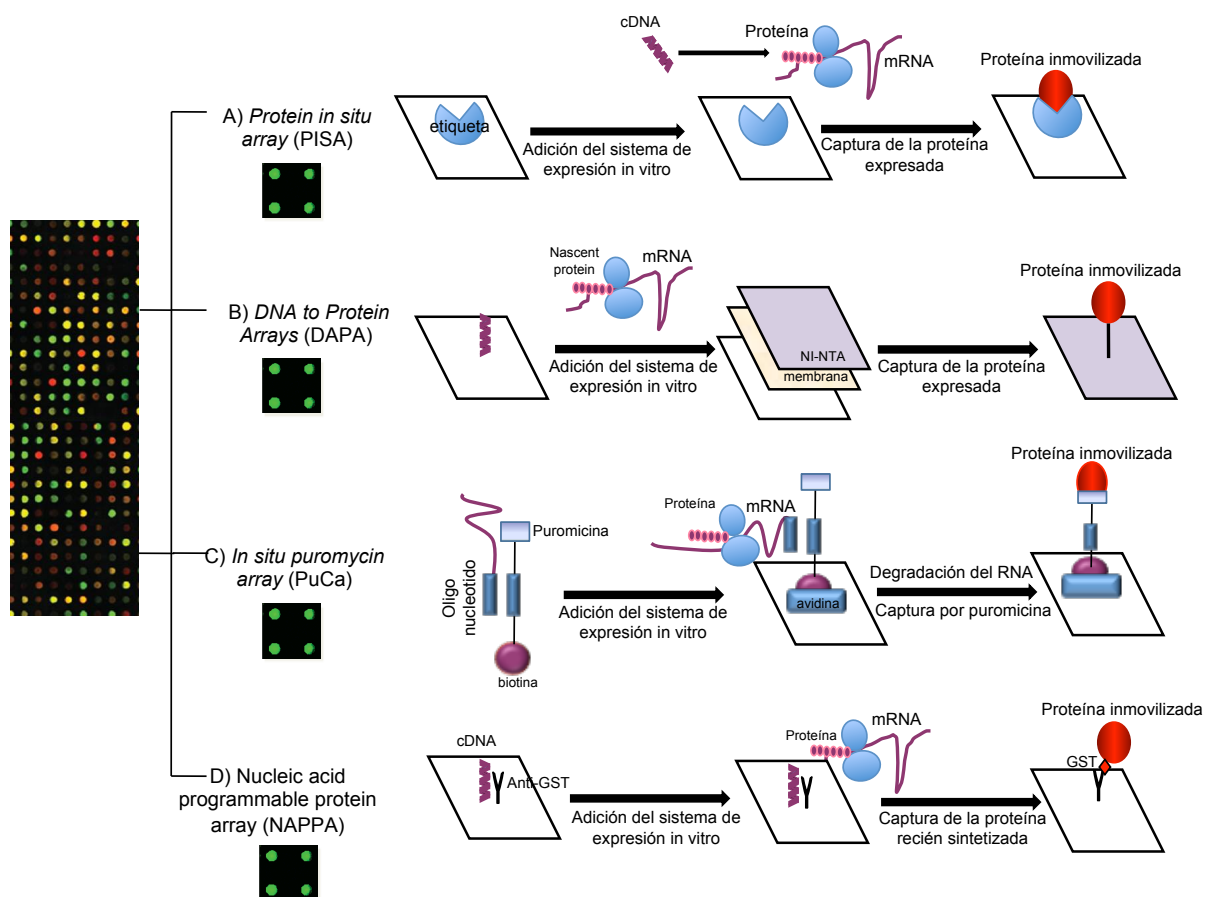


Figura 25. Figura esquemática de los cuatro métodos que se utilizan para producir *arrays* de proteínas *in situ*. A) *Protein In Situ Array* (PISA); B) *DNA to Protein Array* (DAPA); C) *In situ puromycin array* (PuCa); D) *Nucleic Acid Programmable protein Array* (NAPPA). Figura adaptada de Casado-Vela *et al* [177]

4.2.2.3. Arrays de fase reversa

En los *arrays* de fase reversa se pueden definir como plataformas miniaturizadas de las clásicas inmunohistoquímicas o *dot blots*. En los *arrays* de fase reversa, en lugar de un reactivo de captura, en la superficie del *array* se depositan cantidades mínimas de las muestras biológicas objeto de estudio. Normalmente las muestras se inmovilizan en plataformas de vidrio tratadas con nitrocelulosa [191]. Posteriormente, las muestras son interrogadas utilizando anticuerpo específico frente a la diana de interés. Los primeros *arrays* de fase reversa fueron descritos por Paweletz en 2001 [192]. Posteriormente, otros estudios demostraron también la utilidad de los *arrays* de fase reversa para estudiar modificaciones postraduccionales como las fosforilaciones con el fin de elucidar las vías de señalización celular en lisados celulares procedentes de diferentes tipos de cáncer [193-195].

Aunque no es lo más frecuente, en los *arrays* de fase reversa, también se pueden inmovilizar muestras de suero. El principal inconveniente de esta tecnología es que solamente permite detectar proteínas relativamente abundantes en la muestra biológica del orden de 500-700ng/mL de proteína [196, 197].

II. OBJETIVOS

El objetivo general de esta tesis es buscar perfiles proteicos asociados a la artrosis con potencial biomarcador de la enfermedad, aplicando diferentes técnicas proteómicas basadas en espectrometría de masas y *arrays* de proteínas.

Para la consecución de este objetivo general, nos propusimos los siguientes objetivos:

1. Analizar el perfil proteico característico de la degradación del cartílago artrósico humano en un modelo *in vitro* de explantes de tejido.
2. Estudiar el perfil proteico característico del líquido sinovial de pacientes con artrosis.
3. Buscar autoanticuerpos en suero de pacientes con artrosis con potencial valor diagnóstico de la enfermedad.
4. Identificar un panel de proteínas en suero que permitan diagnosticar a los pacientes artrósicos.

III. CAPÍTULOS

Capítulo I: *“Quantitative proteomic profiling of human articular cartilage degradation in osteoarthritis”*

La artrosis (OA) es una enfermedad caracterizada principalmente por la degradación del cartílago. Por esta razón, el conjunto de proteínas secretado por el cartílago, que es lo que se conoce como secretoma, es un compartimento interesante para buscar biomarcadores moleculares de la enfermedad, puesto que las proteínas derivan del lugar donde se está produciendo el daño. Además, las proteínas identificadas en este compartimento pueden ser un reflejo de las proteínas que podemos encontrar en otros fluidos biológicos como el líquido sinovial o la sangre.

El objetivo de este estudio fue caracterizar el patrón de degradación del cartílago en la artrosis con el fin de buscar potenciales marcadores moleculares tempranos y de progresión de la enfermedad. Para ello, nos propusimos identificar el perfil de las proteínas fisiológicamente secretadas por explantes de cartílago articular humano y cuantificar, mediante la técnica iTRAQ, las diferencias entre el perfil proteico secretado desde las zonas dañadas del cartílago artrósico (zonas lesionadas, WOA), las zonas adyacentes a la lesión, macroscópicamente intactas del mismo tejido artrósico (zonas no lesionadas, UWOA) y el perfil de proteínas secretadas por cartílagos procedentes de individuos sanos (N). Para este trabajo, se utilizaron un total de 2 caderas y 2 rodillas procedentes de pacientes artrósicos sometidos a reemplazamiento articular y 2 caderas y 2 rodillas de individuos sin enfermedad articular.

A continuación, se extrajeron 5 explantes desde las zonas WOA y UWOA de cada cartílago artrósico así como 5 explantes de cada uno de los cartílagos normales. Los explantes se cultivaron durante 6 días con medio sin suero bovino fetal y el medio se renovó a día 1 y 3 de cultivo. A día 6, se recogió el secretoma y el tejido de la placa de cultivo. A continuación, se juntaron los secretomas correspondientes a día 3 y día 6 de cultivo.

Para el análisis proteómico, se cogió el mismo volumen del secretoma de cada una de las tres condiciones de estudio (N, UWOA, WOA), el contenido proteico se precipitó con acetona, se dirigió con tripsina y se marcó con iTRAQ. Después del marcaje, se juntaron las 3 condiciones y se eliminó el exceso de sales utilizando una columna en fase reversa.

Los péptidos se fraccionaron mediante una primera separación por cromatografía líquida en fase reversa a pH básico. Las fracciones resultantes se agruparon y se sometieron a una segunda separación por nanocromatografía líquida

en fase reversa a pH ácido acoplada a espectrometría de masas de tipo MALDI-TOF/TOF.

En total se realizaron 4 experimentos iTRAQ, cada uno con explantes de cartílago procedentes de donantes normales y artrósicos. Después de analizar los 4 experimentos, identificamos un total de 232 proteínas con al menos dos péptidos. Para determinar la cuantificación relativa entre las distintas condiciones, se consideraron moduladas aquellas proteínas que presentaran un ratio >1.8 o <0.55 . Siguiendo este criterio, identificamos un total de 76 proteínas diferencialmente representadas en los secretomas de las zonas WOA, UWOA del cartílago artrósico y de los cartílagos normales. La mayoría de estas proteínas estaban relacionadas con la estructura y organización de la matriz extracelular del cartílago.

Posteriormente, las proteínas se clasificaron en grupos según la abundancia que presentaban en las 3 condiciones comparadas. Este análisis permitió la caracterización de diferentes paneles de proteínas de acuerdo a su perfil de liberación desde el cartílago. Así, encontramos 42 proteínas cuya abundancia estaba significativamente modulada en zonas macroscópicamente normales del cartílago OA (UWOA) con respecto a los cartílagos normales y estas diferencias se mantenían cuando comparábamos el secretoma de zonas lesionadas (WOA) y el de cartílagos normales. Otras 31 proteínas, sólo se modulaban cuando la degradación en el cartílago OA era más evidente (WOA). También encontramos un pequeño grupo de 3 proteínas que mostraban una secreción diferencial en la zona macroscópicamente normal que desaparecía cuando el tejido ya estaba visiblemente dañado.

De estas 76 proteínas, validamos por inmunodetección la modulación de dos proteínas: la *osteoprotegerin* (OPG) y la *periostin* (POSTN).

Para los ensayos de validación, se utilizaron 12 muestras de cartílagos procedentes de pacientes artrósicos e individuos sanos diferentes a los utilizados en los análisis proteómicos mediante iTRAQ.

Los niveles de OPG disminuían en el secretoma de UWOA del cartílago OA con respecto a los cartílagos normales y sus niveles seguían bajando en WOA. La POSTN aumentaba en WOA del cartílago OA con respecto a cartílagos normales. Además comprobamos por PCR e inmunohistoquímica que los niveles de POSTN estaban más aumentados en WOA del cartílago OA que en UWOA y en estas últimas, más que en cartílagos normales. Para esta parte del trabajo, también utilizamos 5 cartílagos normales y las zonas WOA y UWOA de 5 cartílagos OA diferentes a los que se había

utilizado para los análisis anteriores. Estos resultados mostraban el potencial de OPG y POSTN como marcadores de progresión de la degradación del cartílago en la artrosis.

Finalmente, comprobamos que los niveles de POSTN estaban significativamente incrementados en el líquido sinovial de pacientes OA con respecto a los niveles observados en pacientes con artritis reumatoide (AR) y artritis psoriásica (APS).

Globalmente, este trabajo describe un panel de proteínas relacionado con las etapas de degradación del cartílago que podrían ser útiles para el desarrollo de nuevos diagnósticos de la OA.

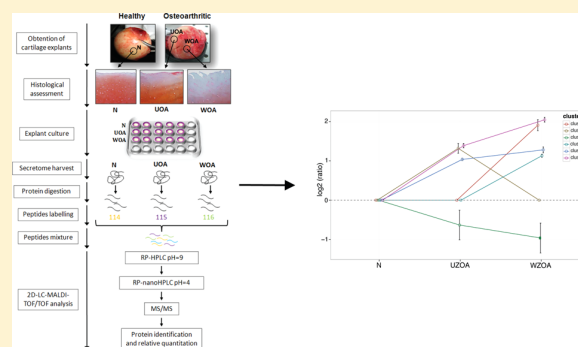
Quantitative Proteomic Profiling of Human Articular Cartilage Degradation in Osteoarthritis

Lucía Lourido,[†] Valentina Calamia,[†] Jesús Mateos,[†] Patricia Fernández-Puente,[†] Juan Fernández-Tajes,^{†,‡} Francisco J Blanco,^{*,†,§} and Cristina Ruiz-Romero^{*,†,||}[†]Proteomics Group-PBR2-ProteoRed/ISCIII, Rheumatology Division, [§]RIER-RED de Inflamación y Enfermedades Reumáticas,^{||}CIBER-BBN Instituto de Salud Carlos III, Instituto de Investigación Biomédica de A Coruña (INIBIC), Complejo Hospitalario Universitario de A Coruña (CHUAC), Sergas, Universidade da Coruña (UDC), As Xubias, 84, 15006-A Coruña, Spain[‡]Wellcome Trust Centre For Human Genetics, McCarthy's Group, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom

S Supporting Information

ABSTRACT: Osteoarthritis (OA) is the most common rheumatic pathology and is characterized primarily by articular cartilage degradation. Despite its high prevalence, there is no effective therapy to slow disease progression or regenerate the damaged tissue. Therefore, new diagnostic and monitoring tests for OA are urgently needed, which would also promote the development of alternative therapeutic strategies. In the present study, we have performed an iTRAQ-based quantitative proteomic analysis of secretomes from healthy human articular cartilage explants, comparing their protein profile to those from unwounded (early disease) and wounded (advanced disease) zones of osteoarthritic tissue. This strategy allowed us to identify a panel of 76 proteins that are distinctively released by the diseased tissue. Clustering analysis allowed the classification of proteins according to their different profile of release from cartilage.

KEYWORDS: osteoarthritis, cartilage, secretome, proteomics, biomarkers



Among these proteins, the altered release of osteoprotegerin (decreased in OA) and periostin (increased in OA), both involved in bone remodelling processes, was verified in further analyses. Moreover, periostin was also increased in the synovial fluid of OA patients. Altogether, the present work provides a novel insight into the mechanisms of human cartilage degradation and a number of new cartilage-characteristic proteins with possible biomarker value for early diagnosis and prognosis of OA.

INTRODUCTION

Osteoarthritis (OA) is the most common global arthritic disease and is among the most frequent health problems for middle aged and older people. Worldwide estimates are that 9.6% of men and 18.0% of women aged over 60 years have symptomatic osteoarthritis. Almost 80% of patients with OA have some degree of limitation of movement, and 25% cannot perform their major daily activities of life. Therefore, OA is already one of the 10 most disabling diseases in developed countries, and it is becoming even more prevalent as the population ages and obesity rates rise.¹

OA is a degenerative joint disease that is characterized by progressive articular cartilage destruction and bone changes, occasionally accompanied by synovial inflammation.² One of the major research goals in OA is the development of strategies for early diagnosis and accurate monitoring of the disease progression. This disease is clinically silent in most patients in their early stages; thus the deterioration of cartilage is already extensive at the time of diagnosis. Diagnostic procedures

carried out at present are mainly symptomatic (based on the description of pain or stiffness) and based on X-ray, a technique that provides only indirect information on the state of cartilage and lacks sensitivity to detect small changes. This deficit of accurate diagnostic methods is especially important in the case of OA, as there are no disease-modifying treatments available to date, with the only existing therapies primarily comprising analgesics. Indeed, at advanced stages the most successful intervention in OA is joint replacement, with the high social and economic costs this entails.

This background highlights the great need to find new biomarkers that permit the molecular characterization of the disease, thus enabling the early diagnosis and evaluation of both disease progression and efficacy of novel therapies.³ In this area, proteomics has emerged in the last few years as a powerful tool for biomarker discovery.⁴ Given the specific characteristics of

Received: October 1, 2014

Published: November 10, 2014

OA, the subset of cartilage secreted proteins (secretome) is an interesting compartment for the search of such biomarkers, as it is derived directly from the site where the disease occurs. The cartilage explant model, maintaining and supporting cells (chondrocytes) within their natural physicochemical micro-environment, overcomes the typical limitations of an *in vitro* cellular model, mainly associated with the dedifferentiation of chondrocytes in monolayer cultures. Furthermore, identifying proteins released from explants may give an indication of the cartilage proteins most likely to be present in the biofluids (synovial fluid or blood) of OA patients.⁵

To date, the large-scale quantitative analysis of secretomes from cartilage explants has only background in samples from animals, such as calf,^{6,7} mouse,⁸ or horse.^{5,9} Furthermore, in all these studies an *in vitro* model of tissue degradation was evaluated, either by the use of mechanic compression or by the addition of catabolic agents such as proinflammatory cytokines. In contrast to these works, in the present study we aimed to identify the profiles of proteins physiologically secreted by human articular cartilage explants and to quantitatively analyze the differences of these profiles between the damaged areas of osteoarthritic cartilage (wounded zones in OA tissue, or WZOA), the macroscopically intact areas of the same diseased tissue (unwounded zones, or UZOA), and normal tissue (from healthy controls, or N). This would allow characterizing cartilage degradation in osteoarthritis, and ultimately identifying putative protein biomarker candidates for this disease, which can be useful for its molecular diagnosis and monitoring.

MATERIALS AND METHODS

Patients

In this study, OA patients were diagnosed following the criteria determined by the American College of Rheumatology.¹⁰ Articular cartilage for the proteomic analysis was obtained either from femoral heads or condyles of eight donors (mean age 74 ± 14 years), a half undergoing joint replacement (OA) and a half with no history of joint disease (N). For the validation studies (immunohistochemistry, real-time PCR and Western blot), cartilage samples from 12 additional sex- and age-matched individuals were included (mean age 78 ± 10 years, 50% women). All tissue samples were provided by the Tissue Bank and the Autopsy Service at Hospital Universitario de A Coruña. Synovial fluid samples were collected after informed consent from knee joints by arthrocentesis in the Rheumatology Department at Hospital Universitario de A Coruña, following our institutional regulations and procedures for sample collection. Patients with osteoarthritis ($n = 14$), rheumatoid arthritis (RA, $n = 7$), and psoriatic arthritis (PSA, $n = 7$) selected for this study were diagnosed following the criteria determined by the American College of Rheumatology/European League Against Rheumatism.¹¹ The study was approved by the local Ethics Committee (Galicia, Spain).

Histological–Histochemical Grading of Cartilage Samples

The Mankin score¹² was employed for the histopathological classification of the severity of osteoarthritic lesions of the cartilage samples employed in this work. Briefly, tissue sections ($4 \mu\text{m}$) were stained with hematoxylin and eosin (cellular architecture), toluidine blue and safranin O/fast green (proteoglycan contents of matrix). Histological scoring of the cartilage was performed by two blinded investigators and averaged by using a modified Mankin scoring system¹³ with the following three categories: cartilage structure (0–7 points),

cellular abnormalities (0–2 points), and matrix staining (0–4 points). These three different aspects of the score were determined and summed up leading to a scale that ranges between 0 and 13. The Mankin score 0–2 represents normal cartilage, 3–5 superficial fibrillation, 6–7 moderate cartilage destruction, 8–10 severe damage of cartilage, and over 10 complete loss of cartilage.

Explants Culture

Tissue explants were obtained from the dissection of N and OA cartilages. Among the OA samples, we differentiated the wounded zones (WZOA) from those corresponding to the area adjacent to the lesion, or unwounded zones (UZOA). Cartilage shavings from each donor were cut into 6 mm discs using a sterile biopsy punch. After extensive washes with PBS, five discs per donor were placed into 96-well plates (one disc/well), containing $200 \mu\text{L}$ of serum-free DMEM supplemented with 100 units/mL penicillin and 100 $\mu\text{g}/\text{mL}$ streptomycin. Plates were incubated overnight at 37°C , 5% CO_2 .

Processing of Conditioned Media

The conditioned media (CM) from each well were collected at days 1 and 3, and replaced with fresh medium. At day 6, both CM and the correspondent cartilage disc were collected. Subsequently, phenylmethylsulfonyl fluoride (PMSF) and protease inhibitor cocktail (1:1000) were added to the CM. Protein concentrations were determined by the Bradford assay (Sigma). Finally, samples from the same donor, zone, and day were mixed and stored at -80°C until use. For the proteomic analysis, 1 mL of CM from each condition (N, UZOA, WZOA) was concentrated using Amicon ultrafiltration units (10 kDa MWCO, Millipore, Billerica, MA, USA), cleaned up by acetone precipitation, and stored at -80°C .

One-Dimensional Gel Electrophoresis and In-Gel Digestion of Proteins

Two micrograms of extracellular proteins from days 1, 3, and 6 of culture were resolved onto a 10% SDS-PAGE gel. The gel was stained with silver nitrate, and the most intense bands were cut and subjected to tryptic digestion according to reported procedures to allow subsequent MS identification.^{14,15}

Protein Digestion and iTRAQ Labeling

Trypsin digestion and labeling of peptides with iTRAQ reagents were performed according to the supplier's instructions (ABSciex, Foster City, CA, USA). The protein pellets from each condition were reconstituted in $25 \mu\text{L}$ of dissolution buffer. Then, samples were denatured with 2% sodium dodecyl sulfate (SDS) in 0.5 M tryethylammonium bicarbonate (TEAB) (ABSciex, Foster City, CA, USA), reduced for 1 h at 60°C using 50 mM tris(2-carboxyethyl) phosphine (TCEP) (ABSciex), and cysteine-blocked with 200 mM methylmethanethiosulfonate (MMTS) (ABSciex) at room temperature in the dark for 30 min. Then, the proteins were digested with spectrometry grade trypsin (Gold Mass, Promega, Madison, WI, USA) at a concentration of 1:20 trypsin/protein for 14 h at 37°C . Each peptide solution was labeled for 90 min at room temperature using the iTRAQ reagents previously reconstituted in $70 \mu\text{L}$ of ethanol, following the iTRAQ Reagents Multiplex kit-Plasma Protocol (ABSciex). The samples were labeled with iTRAQ reagents as follows: N, 114; UZOA, 115; WZOA, 116. The reaction was stopped by adding deionized water, and the labeled samples (114, 115, and 116) were combined. The mixture was desalted using reversed-phase columns (Pierce C18 Spin Columns) prior to LC-MS analysis.

Liquid Chromatography (LC)

The peptides mixture was first separated by off-line reversed-phase chromatography at basic pH (pH 10) to lower its complexity. Dried labeled peptides were resuspended in 140 μL of buffer A (10 mM ammonium hydroxide, 5% acetonitrile) and 130 μL was injected into an HP 1200 system (Agilent). The separation was performed on a C18 reversed-phase column (Zorbax extend C18, 100 \times 2.1 mm id, 3.5 μm , 300 \AA ; Agilent), at a flow rate of 0.2 mL/min. The chromatogram was obtained using a UV detector at 214 nm. Sixty fractions were collected every 90 s using a Gilson FC203B fraction collector (Gilson, Middleton, WI, USA). Several fractions were pooled postcollection based on the peak intensity of the UV trace, yielding to 12–15 LC runs per experiment. Each fraction was dried in a vacuum concentrator and stored at $-20\text{ }^\circ\text{C}$ for the next step of analysis. For a second dimension separation, the dried peptide fractions were dissolved in 0.1% TFA and 2% ACN. Then, 5 μL of this sample were injected onto a capillary trap column (0.5 \times 2 mm, Michrom Bioresources, Auburn, CA, USA) at a flow rate of 15 $\mu\text{L}/\text{min}$. Peptides were desalted for 10 min and loaded onto a C18 column (Integrat C18, Proteopep II, 75 μm id, 10.2 cm, 5 μm , 300 \AA ; New Objective) to perform the separation using linearly increasing concentration of ACN in buffer B. The reversed-phase fractions were collected and mixed with matrix α -cyano-4-hydroxycinnamic acid (Sigma-Aldrich) at 3 mg/mL in 0.1% TFA in 70% ACN at a flow rate of 1.2 $\mu\text{L}/\text{min}$, using the Sun Collect MALDI Spotter/Micro Collector (SunChrom Wissenschaftliche Geräte GmbH, Germany). Fractions collected every 15 s were then spotted onto a matrix-assisted laser desorption/ionization (MALDI) plate for MS analysis.

MS/MS Analysis

Data acquisition was accomplished in a positive ion mode using a 4800 analyzer equipped with time-of-flight (TOF)/TOF ion optics (ABSciex) and 4000 Series Explorer software version 3.5.1 (ABSciex). MS spectra from m/z 800–4000 were acquired for each fraction using 1500 laser shots processed with internal calibration. We used 3 fmol/spot of angiotensin, diluted in the matrix, as an internal standard ($m/z = 1046.50$), and a laser intensity of 3800. After screening of all LC-MALDI sample positions in MS positive reflector mode, the fragmentation of automatically selected precursors was performed at a collision energy of 1 kV with CID gas (air). Up to 12 of the most intense ion signals per spot position with signal/noise ratio (S/N) above 80 were selected as precursors for MS/MS acquisition, excluding common trypsin autolysis peaks and matrix ion signals. The number of shots was 1800 for MS/MS, and the laser intensity was set to 4700. A second MS/MS was acquired excluding the precursors selected in the previous MS/MS run. Precursors with $S/N > 30$ were selected to identify proteins that were not identified in the previous run.

Mass Spectrometry Data Analysis

Protein identification was carried out using ProteinPilot software v.3.0 (ABSciex). The Paragon algorithm in ProteinPilot software served as the default search program with trypsin as the digestion agent and MMTS as a fixed modification of cysteine. Biological modifications were programmed in the algorithm. Each MS/MS spectrum was searched in the Uniprot/Swissprot database (downloaded in May 2012) for *Homo sapiens*. Only proteins identified with at least 95% confidence, or a Prot Score (protein confidence measure) of at least 1.3 were reported. The ProteinPilot software also

calculated a confidence percentage, the Unused score, which reflects the probability of a hit being a “false positive,” at a 95% confidence. While this software automatically accepts all peptides with an identification level $>1\%$, only proteins having at least one peptide above the 95% confidence level were initially recorded. Searches against a concatenated database containing both forward and reversed sequences allowed the false discovery rate to be kept at 1%. The PSPEP program was used independently to calculate false discovery rates. The results obtained from ProteinPilot were exported to Microsoft Excel for further analyses.

Western Blot Analysis

One-dimensional Western blot analyses were performed utilizing standard procedures. Briefly, 15 μg of secreted proteins were loaded and resolved using standard 8% SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Equivalent loadings of each lane were ensured by running in parallel the same samples on another gel and staining it with silver nitrate. The separated proteins were then transferred to polyvinylidene fluoride (PVDF) membranes (Immobilon P, Millipore Co., Bedford, MA, USA) by electroblotting and probed with specific antibodies against POSTN (1:100) and OPG (1:500). Immunoreactive bands were detected by chemiluminescence using corresponding horseradish peroxidase (HRP)-conjugated secondary antibodies and enhanced chemiluminescence (ECL) detection reagents (GE Healthcare), and then digitized using the LAS 3000 image analyzer. Quantitative changes in band intensities were evaluated using ImageQuant 5.2 software (GE Healthcare).

Real-time PCR Assays

Total RNA was isolated from cartilage using Trizol Reagent (Invitrogen, Carlsbad, CA, USA), following the manufacturer's instructions. cDNA was synthesized from 1 μg of total RNA, using the Transcriptor First Strand cDNA Synthesis kit (Roche Applied Science) in accordance with the manufacturer's instructions, and analyzed by quantitative real-time PCR. Quantitative real-time PCR assay was performed in the LightCycler 480 instrument (Roche Applied Science) using 96-well plates. Primers for POSTN and the housekeeping gene RPLP13A were designed using the Universal Probe Library tool from the Roche Web site (<http://www.roche-applied-science.com>). Primer sequences were as follows: POSTN forward, 5'-TGTTGCCCTGGTTATATGAGAA-3'; POSTN reverse, 5'-ACATGGTCAATGGGCAAAAC-3'; RPLP13A forward, 5'-CAAGCGGATGAACACCAAC-3'; RPLP13A reverse 5'-TGTGGGGCAGCATACTC-3'. The results were analyzed using the LightCycler 480 software release 1.5.0 (Roche), which automatically recorded the threshold cycle (Ct). Target gene Ct values were normalized against RPLP13A. Data were analyzed using the $2^{-\Delta\Delta\text{Ct}}$ method and expressed as fold change of the test sample (WZOA, UZOA) compared to the normal condition.¹⁶

Immunohistochemical Analysis

Frozen cartilage samples from five healthy and five OA donors different from those selected for the proteomics strategy were cut at 4 μm on a cryostat (Leica Microsystems, Barcelona, Spain). Sections were incubated with primary antibodies to detect the presence of Periostin (POSTN, 1:50, Santa Cruz Biotechnology, Heidelberg, Germany). The peroxidase/DAB ChemMate™ DAKO EnVision™ detection kit (Dako, Barcelona, Spain) was used to determine antigen-antibody

interactions. Negative staining controls were achieved by omitting the primary monoclonal antibody. Samples were visualized using an optical microscope.

A densitometric analysis of the images was performed to quantify the relative abundance of those proteins detected in the immunohistochemical analyses. With this aim, at least three images from the same assay were quantified with AnalySIS Image Processing software 5.0 (Olympus Soft Imaging Solutions GmbH, Münster, Germany), using the staining intensity assigned by the program for the immunohistochemistry values. Data were expressed as the percentage of positive staining intensity of those sections examined under the same magnification and light microscopy conditions. The total area of each cartilage section was used to normalize all values of intensities.

Suspension Bead Array Analysis

A quantitative protein analysis by magnetic suspension bead array technology was performed according manufacturer's recommendation (MILLIPLEX MAP-2Plex kit; HCMBMAG-22K) for the multiplex detection of periostin (POSTN) and osteoproteogelin (OPG) in synovial fluid (SF) samples obtained from 14 osteoarthritis (OA), 7 psoriatic arthritis (PSA), and 7 rheumatoid arthritis (RA) patients. Lyophilized standards of the recombinant proteins tested, included in the kit, were reconstituted and diluted at seven serial concentrations (standard curves). All SF samples were diluted 1:10 in assay buffer, and 25 μL of each sample was bound for overnight incubation at 4 $^{\circ}\text{C}$ in a 96-well plate. Then, the wells were washed three times with wash buffer and incubated with 25 μL of detection secondary antibodies mix for 1 h at room temperature (RT). Twenty-five microliters of streptavidin-phycoerythrin were subsequently added to each well and incubated 30 min at RT to label the target analyte in the samples. Finally, wells were washed again three times and 100 μL of Sheat Fluid was added per well. The MAGPIX System (Luminex Europe, Oosterhout, The Netherlands) was used to detect the fluorescence signal. Concentration of each analyte was obtained by interpolating fluorescence intensity to the standard curve and calculated by the MILLIPLEX Analyst 5.1 Software.

Bioinformatic Analyses

In the proteomic analysis, normalization tools and statistical package from ProteinPilot software were employed. To determine the cutoff value for significant fold changes, coefficient of variation and cumulative frequency were calculated for the common proteins employing the R statistical package.¹⁷ Gene ontology (GO) enrichment analysis on the identified proteins was performed using DAVID v6.7.¹⁸

For the clustering analysis, the log-transformed mean signature from the final list of modulated proteins was clustered using the k-means method implemented in R. The clusters were calculated allowing a maximum of 1 000 000 iterations; 250 000 was the number of random sets chosen for initializing the algorithm. Between 3 and 10 clusters were examined to determine both the best fit (Silhouette plots) and especially the biologically useful number of clusters with which to interpret the system. A several samples location test using spatial ranks (similar to Hotellings T square) was performed to determine whether the clusters were significantly different from one another and thus likely to describe unique biological phenomena. The enrichment of GO terms was also performed on these clusters, using the R library "topGO" from

Bioconductor. We applied the three available algorithms of this R package, the *classic*, the *elim*, and the *weight* methods. The first of them computes the significance of a GO node independent of the significance of the neighboring nodes, whereas the others take into account the significance of neighbor nodes integrating the dependency structure of GO terms between nodes to enhance the statistical inference.

Each validation experiment was repeated at least four times. The statistical significance of the differences between mean values was determined using a two-tailed *t* test, considering significant *p* values ≤ 0.05 . Where appropriate, results are expressed as the mean \pm standard error.

RESULTS

Selection and Processing of Human Normal and Osteoarthritic Cartilage Specimens

The experimental workflow followed in this work for the proteomic profiling of cartilage degradation in osteoarthritis is illustrated in Figure 1. First, human articular cartilage explants were carefully selected from OA patients and healthy donors. An histological-histochemical assessment of the tissue charac-

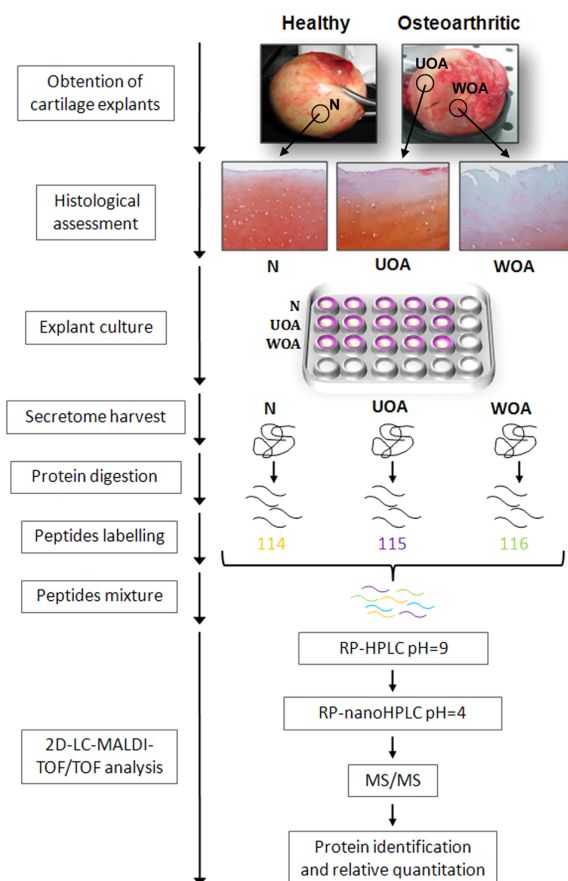


Figure 1. Schematic workflow followed in this work. Histological-histochemical characterization of the selected cartilage explants was performed according to the Mankin score system. Samples from four N, four UZOA, and four WZOA (the last two conditions obtained from the same donor) were included in the study. Then, proteins present in the conditioned media of the explants cultures were processed for iTRAQ labeling and subsequent 2D-LC-MS/MS analysis.

Table 1. List of Proteins Differentially Released from Unwounded (UZO) and Wounded (WZO) Zones of OA Cartilage Compared to Normal Tissue (N)

| N | accession number ^a | symbol | protein name | pept (95%) ^b | UZO/N | | | WZO/N | | | cluster ^f |
|----|-------------------------------|--------|--|-------------------------|--------------------|-------------------|-----------------|--------------------|-------------------|-----------------|----------------------|
| | | | | | ratio ^c | PVal ^d | EF ^e | ratio ^c | PVal ^d | EF ^e | |
| 1 | P02458 | CO2A1 | collagen alpha-1(II) chain | 100/8 | 2.13 | 0.0023 | 1.50 | 1.80 | 0.067 | 1.90 | 1 |
| 2 | P24821 | TENA | tenascin | 79/43 | 2.53 | 0 | 1.40 | 5.68 | 0 | 1.38 | 4 |
| 3 | P10915 | HPLN1 | hyaluronan and proteoglycan link protein 1 | 71/17 | 2.29 | 0 | 1.25 | 2.13 | 0 | 1.32 | 1 |
| 4 | P12111 | CO6A3 | collagen alpha-3(VI) chain | 69/52 | 1.60 | 0 | 1.14 | 3.37 | 0 | 1.18 | 2 |
| 5 | P22105 | TENX | tenascin-X | 58/45 | 1.81 | 0 | 1.26 | 3.33 | 0 | 1.32 | 1 |
| 6 | Q15782 | CH3L2 | chitinase-3-like protein 2 | 53/13 | 2.03 | 0.0052 | 1.59 | 1.94 | 0.0053 | 1.55 | 1 |
| 7 | P02461 | CO3A1 | collagen alpha-1(III) chain | 52/4 | 2.52 | 0.013 | 1.82 | 3.22 | 0.0102 | EF > 2 | 4 |
| 8 | P21810 | PGS1 | biglycan | 49/13 | 1.93 | 0.0453 | 1.90 | 2.58 | 0.0031 | 1.76 | 1 |
| 9 | P02452 | CO1A1 | collagen alpha-1(I) chain | 48/2 | 2.35 | 0.0401 | EF > 2 | 5.30 | 0.0382 | EF > 2 | 4 |
| 10 | P01857 | IGHG1 | Ig gamma-1 chain C region | 46/4 | 2.28 | 0.0281 | 1.93 | 3.62 | 0.0115 | EF > 2 | 4 |
| 11 | Q92954 | PRG4 | proteoglycan 4 | 45/13 | 1.36 | 0.0055 | 1.23 | 2.08 | 0 | 1.21 | 3 |
| 12 | P08123 | CO1A2 | collagen alpha-2(I) chain | 42/9 | 1.71 | 0.0779 | 1.84 | 3.69 | 0.0011 | 1.87 | 2 |
| 13 | P36955 | PEDF | pigment epithelium-derived factor | 40/14 | 2.38 | 0.0004 | 1.52 | 3.77 | 0 | 1.44 | 4 |
| 14 | O00300 | TR11B | tumor necrosis factor receptor superfamily member 11B | 38/16 | 0.81 | 0.0131 | 1.18 | 0.55 | 0 | 1.17 | 6 |
| 15 | P98160 | PGBM | basement membrane-specific heparan sulfate proteoglycan core protein | 37/23 | 1.44 | 0 | 1.17 | 2.45 | 0 | 1.25 | 3 |
| 16 | P20774 | MIME | mimectan | 34/11 | 1.69 | 0.0018 | 1.34 | 2.01 | 0.0015 | 1.47 | 3 |
| 17 | P08670 | VIME | vimentin | 33/22 | 2.29 | 0 | 1.24 | 5.13 | 0 | 1.43 | 4 |
| 18 | Q9NQ79 | CRAC1 | cartilage acidic protein 1 | 32/14 | 1.59 | 0.0001 | 1.21 | 3.28 | 0 | 1.29 | 2 |
| 19 | Q15582 | BGH3 | transforming growth factor-beta-induced protein ig-h3 | 31/16 | 3.58 | 0 | 1.33 | 5.02 | 0 | 1.25 | 4 |
| 20 | Q99715 | COCA1 | collagen alpha-1(XII) chain | 31/20 | 1.59 | 0 | 1.22 | 2.39 | 0 | 1.20 | 3 |
| 21 | P01834 | IGKC | Ig kappa chain C region | 28/5 | 2.50 | 0.0691 | EF > 2 | 3.42 | 0.019 | EF > 2 | 4 |
| 22 | P12109 | CO6A1 | collagen alpha-1(VI) chain | 28/14 | 1.62 | 0.0001 | 1.21 | 2.26 | 0 | 1.30 | 3 |
| 23 | P02787 | TRFE | serotransferrin | 24/19 | 1.79 | 0 | 1.26 | 2.44 | 0 | 1.27 | 3 |
| 24 | P05090 | APOD | apolipoprotein D | 22/6 | 0.36 | 0.0001 | 1.29 | 0.28 | 0 | 1.28 | 6 |
| 25 | P08253 | MMP2 | 72 kDa type IV collagenase | 21/16 | 3.11 | 0.0003 | 1.72 | 3.93 | 0.0001 | 1.79 | 4 |
| 26 | P05452 | TETN | tetranectin | 21/7 | 2.32 | 0.0016 | 1.51 | 3.61 | 0.0001 | 1.47 | 4 |
| 27 | P02743 | SAMP | serum amyloid P-component | 19/8 | 2.22 | 0.0017 | 1.51 | 1.49 | 0.0046 | 1.27 | 5 |
| 28 | Q7Z7G0 | TARSH | target of Nesh-SH3 | 18/11 | 0.48 | 0 | 1.31 | 0.82 | 0.0843 | 1.25 | 6 |
| 29 | P00751 | CFAB | complement factor B | 17/14 | 1.44 | 0.1694 | 1.70 | 1.84 | 0.0196 | 1.64 | 3 |
| 30 | P07998 | RNAS1 | ribonuclease pancreatic | 15/4 | 0.60 | 0.0137 | 1.39 | 0.45 | 0.0127 | 1.68 | 6 |
| 31 | Q8IUX7 | AEBP1 | adipocyte enhancer-binding protein 1 | 14/8 | 1.98 | 0 | 1.23 | 2.50 | 0 | 1.30 | 1 |
| 32 | P12110 | CO6A2 | collagen alpha-2(VI) chain | 14/6 | 1.24 | 0.2867 | 1.53 | 2.11 | 0 | 1.24 | 3 |
| 33 | P02647 | APOA1 | apolipoprotein A-I | 13/9 | 1.99 | 0.0014 | 1.43 | 4.07 | 0 | 1.45 | 4 |
| 34 | P13942 | COBA2 | collagen alpha-2(XI) chain | 12/8 | 1.74 | 0.0008 | 1.27 | 2.24 | 0.0003 | 1.36 | 3 |
| 35 | P23142 | FBLN1 | fibulin-1 | 10/8 | 1.82 | 0.0228 | 1.64 | 3.70 | 0.0005 | 1.71 | 4 |
| 36 | P39059 | COFA1 | collagen alpha-1(XV) chain | 10/9 | 1.50 | 0.0945 | 1.62 | 2.67 | 0.0016 | 1.68 | 3 |
| 37 | P14555 | PA2GA | phospholipase A2, membrane associated | 10/4 | 2.96 | 0.058 | 1.75 | 1.36 | 0.0105 | 1.21 | 5 |
| 38 | P01876 | IGHA1 | Ig alpha-1 chain C region | 9/6 | 1.75 | 0.0532 | 1.77 | 3.54 | 0 | 1.33 | 2 |
| 39 | P19827 | ITIH1 | interalpha-trypsin inhibitor heavy H1 | 9/7 | 1.79 | 0.2358 | EF > 2 | 3.03 | 0.0251 | EF > 2 | 2 |
| 40 | P00338 | LDHA | L-lactate dehydrogenase A chain | 9/7 | 1.73 | 0.0081 | 1.45 | 2.02 | 0.0006 | 1.39 | 3 |
| 41 | P02763 | AIAG1 | alpha-1-acid glycoprotein 1 | 8/3 | 2.57 | 0.0061 | 1.54 | 3.30 | 0.0096 | 1.90 | 4 |
| 42 | P13611 | CSPG2 | versican core protein | 8/3 | 1.46 | 0.2148 | EF > 2 | 2.67 | 0.0159 | 1.88 | 3 |
| 43 | P08476 | INHBA | inhibin beta A chain | 8/6 | 2.16 | 0.0016 | 1.48 | 2.43 | 0.0043 | 1.70 | 1 |
| 44 | O75596 | CLC3A | C-type lectin domain family 3 member A | 8/4 | 1.49 | 0.0028 | 1.21 | 1.84 | 0.0678 | 1.96 | 3 |
| 45 | P00738 | HPT | haptoglobin | 7/5 | 2.89 | 0.0116 | EF > 2 | 5.22 | 0.0008 | 1.92 | 4 |
| 46 | P02545 | LMNA | prelamin-A/C | 7/7 | 2.16 | 0.005 | 1.60 | 4.52 | 0 | 1.46 | 4 |
| 47 | Q15063 | POSTN | periostin | 7/5 | 2.54 | 0.0558 | EF > 2 | 4.10 | 0.0054 | EF > 2 | 4 |
| 48 | P06733 | ENOA | alpha-enolase | 7/6 | 1.98 | 0.0643 | EF > 2 | 2.35 | 0.0215 | 1.99 | 1 |
| 49 | P49746 | TSP3 | thrombospondin-3 | 6/2 | 2.93 | 0.0655 | EF > 2 | 6.78 | 0.0332 | EF > 2 | 4 |
| 50 | P55058 | PLTP | phospholipid transfer protein | 6/5 | 3.03 | 0.0262 | EF > 2 | 3.77 | 0.0086 | EF > 2 | 4 |
| 51 | Q99880 | H2B1L | histone H2B type 1-L | 6/3 | 1.73 | 0.1628 | EF > 2 | 2.96 | 0.0178 | EF > 2 | 2 |
| 52 | P30101 | PDIA3 | protein disulfide-isomerase A3 | 6/4 | 2.16 | 0.0175 | 1.73 | 2.83 | 0.0002 | 1.23 | 1 |
| 53 | P02766 | TTHY | transthyretin | 6/4 | 1.09 | 0.8554 | EF > 2 | 2.72 | 0.0394 | EF > 2 | 3 |
| 54 | P04406 | G3P | glyceraldehyde-3-phosphate dehydrogenase | 6/5 | 1.45 | 0.0864 | 1.57 | 2.06 | 0.0288 | 1.84 | 3 |
| 55 | P02760 | AMBP | protein AMBP | 6/5 | 1.59 | 0.0405 | 1.54 | 1.98 | 0.0491 | 1.97 | 3 |

Table 1. continued

| N | accession number ^a | symbol | protein name | pept (95%) ^b | UZOA/N | | | WZOA/N | | | cluster ^f |
|----|-------------------------------|--------|--|-------------------------|--------------------|-------------------|-----------------|--------------------|-------------------|-----------------|----------------------|
| | | | | | ratio ^c | PVal ^d | EF ^e | ratio ^c | PVal ^d | EF ^e | |
| 56 | P05109 | S10A8 | protein S100-A8 | 5/4 | 1.71 | 0.0154 | 1.45 | 6.72 | 0.0006 | 1.71 | 2 |
| 57 | P0CF74 | LAC6 | Ig lambda-6 chain C region | 5/3 | 1.88 | 0.0631 | EF > 2 | 3.81 | 0.0165 | EF > 2 | 4 |
| 58 | Q16610 | ECM1 | extracellular matrix protein 1 | 5/4 | 1.49 | 0.2078 | EF > 2 | 2.69 | 0.0203 | EF > 2 | 3 |
| 59 | P00736 | C1R | complement C1r subcomponent | 5/4 | 2.02 | 0.0207 | 1.69 | 2.52 | 0.0648 | EF > 2 | 1 |
| 60 | P14618 | KPYM | pyruvate kinase isozymes M1/M2 | 5/4 | 1.14 | 0.4735 | 1.55 | 1.96 | 0.0035 | 1.39 | 3 |
| 61 | P62805 | H4 | histone H4 | 4/3 | 3.36 | 0.0239 | EF > 2 | 4.03 | 0.0044 | 1.77 | 4 |
| 62 | P02741 | CRP | C-reactive protein | 4/3 | 3.54 | 0.0163 | EF > 2 | 3.33 | 0.0293 | EF > 2 | 4 |
| 63 | P07451 | CAH3 | carbonic anhydrase 3 | 4/3 | 3.31 | 0.0147 | EF > 2 | 3.09 | 0.005 | 1.62 | 4 |
| 64 | Q15084 | PDIA6 | protein disulfide-isomerase A6 | 4/4 | 2.15 | 0.0194 | 1.70 | 2.78 | 0.0048 | 1.54 | 1 |
| 65 | P18669 | PGAM1 | phosphoglycerate mutase 1 | 4/4 | 1.58 | 0.0559 | 1.61 | 2.33 | 0.0289 | EF > 2 | 3 |
| 66 | P62937 | PPIA | peptidyl-prolyl cis-trans isomerase A | 4/3 | 1.22 | 0.5028 | EF > 2 | 1.99 | 0.0329 | 1.74 | 3 |
| 67 | P26038 | MOES | moesin | 4/4 | 0.86 | 0.0527 | 1.16 | 1.83 | 0.0007 | 1.23 | 3 |
| 68 | P10451 | OSTP | osteopontin | 4/3 | 2.37 | 0.0088 | 1.42 | 1.59 | 0.00752 | 1.79 | 5 |
| 69 | Q14112 | NID2 | nidogen-2 | 3/1 | 2.88 | 0.0608 | EF > 2 | 5.30 | 0.0456 | EF > 2 | 4 |
| 70 | P60903 | S10AA | protein S100-A10 | 3/2 | 2.43 | 0.1002 | EF > 2 | 4.45 | 0.0035 | 1.47 | 4 |
| 71 | P02144 | MYG | myoglobin | 3/2 | 3.47 | 0.0528 | EF > 2 | 3.38 | 0.1674 | EF > 2 | 4 |
| 72 | Q12841 | FSTL1 | follistatin-related protein 1 | 3/3 | 1.66 | 0.0431 | 1.62 | 2.28 | 0.0092 | 1.63 | 3 |
| 73 | P17936 | IBP3 | insulin-like growth factor-binding protein 3 | 3/1 | 2.32 | 0.0168 | 1.81 | 1.94 | 0.1204 | EF > 2 | 1 |
| 74 | P02652 | APOA2 | apolipoprotein A-II | 2/2 | 1.07 | 0.7619 | EF > 2 | 3.33 | 0.0559 | EF > 2 | 2 |
| 75 | P32119 | PRDX2 | peroxiredoxin-2 | 2/2 | 1.81 | 0.0736 | EF > 2 | 2.95 | 0.0526 | EF > 2 | 1 |
| 76 | P62987 | RL40 | ubiquitin-60S ribosomal protein L40 | 2/2 | 1.18 | 0.304 | 1.55 | 1.97 | 0.0114 | 1.47 | 3 |

^aAccession number according to SwissProt and TrEMBL databases. ^bNumber of unique peptides (Pept) used for protein identification and iTRAQ ratio calculations, both at 95% confidence. ^cAverage iTRAQ ratios ($n = 4$) that represent the relative protein abundance in UZOA and WZOA secretome versus secretome from normal cartilage calculated by Protein Pilot 3.0 software. ^d p value (measure of the statistical results), calculated by Protein Pilot 3.0 software. ^eError factor (a measure of the error in the average ratio), calculated by Protein Pilot 3.0 software for quantification accuracy. ^fProtein classification according to the clustering analysis.

teristics was performed on the explants by Mankin scoring.¹⁹ Referring to the classification of Mankin and colleagues, the human specimens were classified into four stages based on their scores: nearly normal ($0 \leq$ Mankin score < 2), early OA ($2 \leq$ Mankin score < 6), moderate OA ($6 \leq$ Mankin score < 10), and late OA ($10 \leq$ Mankin score ≤ 13). Samples from four healthy controls (Mankin score 0–1) and four OA joints were selected for the proteomic analysis. In the OA joints, explants were obtained both from the macroscopically normal zone of the tissue (UZOA, Mankin score 2) and the lesion (WZOA, Mankin score 7–9). Representative images of the Safranin O/ Fast Green staining of cartilage proteoglycans for Mankin score assessment are shown in Figure 1.

The cartilage explants were then cultured as described in Materials and Methods. Culture conditions (washings, media replacement, and culture length) were optimized by evaluating the different secretome profiles that were obtained on silver stained SDS-PAGE gels. Clear variations in the protein profiles were found between N and OA cartilages and between UZ and WZ among the OA samples (Supplementary Figure S1, Supporting Information). Since no significant differences were observed between days 3 and 6 of culture of the same sample (data not shown), these were mixed for the subsequent phases of the work. Proteins recovered at day 1 were always discarded due to their high amount of serum contaminants, mainly detected in the WZOA samples. We obtained an average of $0.24 \pm 0.10 \mu\text{g}/\mu\text{L}$ proteins from the healthy tissues, $0.32 \pm 0.14 \mu\text{g}/\mu\text{L}$ from the UZOA and $0.46 \pm 0.25 \mu\text{g}/\mu\text{L}$ from the WZOA. The global quantitative differences between secretomes from these three conditions were statistically significant (Supplementary Figure S2).

Identification and Relative Quantitation of Proteins Released from Healthy and Osteoarthritic Cartilage Tissue

The iTRAQ-based proteomic analysis led to the total identification of 401 proteins (Supplementary Table S1). These data were the result of four distinct iTRAQ experiments, each one with cartilage explants from different normal and OA donors. The identification record was then filtered by a stringent threshold of two distinct peptides ($>95\%$) and a protein pilot total score ≥ 2 , leading to a final list of 234 high-confidence unique proteins identified in the secretome of cartilage explants (Supplementary Table S2). A gene ontology (GO) enrichment analysis was performed on this set of proteins with DAVID bioinformatics resources,¹⁸ and demonstrated the significant presence of extracellular proteins in our samples, since the highest statistical significances were found for GO terms and cellular compartments relative to Extracellular Region and keyword “secreted” (Supplementary Table S3).

iTRAQ ratios were calculated for 232 of these proteins using the Protein Pilot software, and subsequently the cutoff value for significant fold changes was established at ≥ 1.8 (or ≤ 0.55) using bioinformatics tools (see Materials and Methods and Supplementary Figure S3). Following these considerations, 76 proteins were finally identified as differentially released from OA tissue (either UZ or WZ) when compared to healthy cartilage. These proteins are listed in Table 1, together with their abundance ratios in the OA conditions compared to Normal.

For the classification of the different profiles of protein release from the cartilage explants, a clustering analysis was applied on this set of altered proteins leading to their

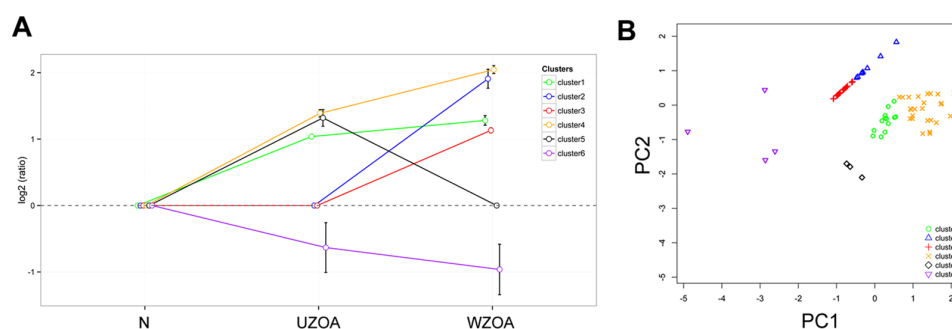


Figure 2. Multivariate analysis of the differentially released proteins. (A) k-means analysis clustered the proteins into six different groups, according to their profile of release from cartilage. Data are presented as \log_2 ratios of proteins from unwounded zones and wounded zones of OA tissue (UZOA and WZOA, respectively) compared to healthy cartilage (N). The graphical representation of each independent cluster is shown in Supplementary Figure S4. (B) Principal component analysis differentiated the proteins by two principal components that distinguish the variance. The colors of the proteins in the PCA graphic correlate with the six clusters.

classification into six different clusters composed with completely independent protein subsets. A graphical representation of this grouping is shown in Figure 2A, whereas the graphics obtained from each independent cluster are supplied in Supplementary Figure S4. The clustering analysis allowed the characterization of different panels of proteins according to their profile of release by osteoarthritic cartilage. As shown in Figure 2A, the release of proteins belonging to the clusters 1, 4, 5, and 6 is significantly increased (or decreased, in the case of cluster 6) already in UZOA. However, whereas this release keeps significantly increasing in WZOA for the case of proteins from cluster 4, the secretion of proteins slows its tendency (cluster 1) or even decreases (cluster 5) when the tissue is macroscopically damaged. Additionally to these different release profiles, proteins from clusters 2 and 3 do not show quantitative differences in the secretion from UZOA cartilage, but only when the degradation is evident. A principal component analysis was carried out to visualize the differential behavior of the proteins grouped in the six independent clusters (Figure 2B).

Functional Proteomic Analysis of Cartilage Degradation Processes

To further explore the biological functions involved in the process of cartilage degradation and gain understanding on the entailed molecular mechanisms, we carried out a pathway analysis using String software on the set of 76 quantitatively altered proteins from the OA tissue included in the clustering analysis. As shown in Supplementary Figure S5, the most represented functional group corresponds to proteins involved in ECM organization and turnover, such as collagens or small proteoglycans.

We also performed enrichment analyses on the six clusters of proteins identified as altered. As shown in the Supplementary Table S4, the GO terms with a higher enrichment were ECM organization for proteins in clusters 1 and 4, collagen catabolic process for cluster 2, ECM disassembly for cluster 3, inflammatory response for cluster 5, and regulation of bone resorption for cluster 6.

Verification of the Altered Release of Two Bone Remodeling Proteins, Periostin and Osteoprotegerin, from Osteoarthritic Cartilage

Two proteins were selected for further verification analyses of their differential release: one belonging to cluster 4 (periostin, POSTN, whose release increases linearly with the tissue

degradation) and one to cluster 6 (osteoprotegerin, OPG, whose release decreases) based on their different behavior, but mainly on their involvement in bone remodeling pathways, which are known to be important processes for OA pathogenesis²⁰ and a putative good source of OA biomarkers.²¹

Western blot analyses were carried out on independent samples of cartilage secretomes, different from those employed in the proteomic screening ($n = 12$, 4 normal tissues, 4 UZOA, and 4 WZOA), in order to evaluate the levels of POSTN and OPG secretion from the tissue. As shown in Figure 3A, a significant increase of POSTN was found in the secretome of WZOA, which was semiquantified by densitometric analysis of the blot images, obtaining a mean ratio vs N of 2.54, and $p = 0.03$. On the contrary, as shown in Figure 3B, secretion of OPG was significantly diminished in wounded OA cartilage explants. On this occasion, the mean ratio WZOA/N was 0.399 (2.5-fold decreased), with a p value = 0.018. In both cases (POSTN and OPG), slight differences were already found in macroscopically normal OA cartilage, although they were not significant. Altogether, these results confirm the data previously obtained by proteomics and suggest a potential value of POSTN and OPG as markers of disease.

Periostin Expression and Protein Abundance Is Increased in Osteoarthritic Cartilage

Additional experiments were performed to evaluate the levels of POSTN in the human cartilage explants, in order to investigate if the cause for the differential secretion relies on an increased amount of this protein in the pathological tissue. First, real-time PCR analyses were performed on the tissue explants to explore POSTN gene expression. As shown in Figure 4A, a significant increase of POSTN expression was detected in both UZ and WZ explants from OA cartilage compared to normal: WZOA/N = 6.46, $p = 0.01$; UZOA/N = 1.93, $p = 0.047$.

The increase of POSTN in OA cartilage was also confirmed at the protein level by immunohistochemistry analyses on the tissue explants ($n = 15$). As illustrated by representative images of the assays and a densitometry graph of mean percentages of POSTN positivity (Figure 4B), clearly higher amounts of POSTN were detected in the pathological tissue already in macroscopically normal zones.

Periostin is Increased in Synovial Fluid from OA patients

To extend the investigations on the putative usefulness of POSTN as OA biomarker, quantitative measurements of this protein were carried out on human articular synovial fluid from

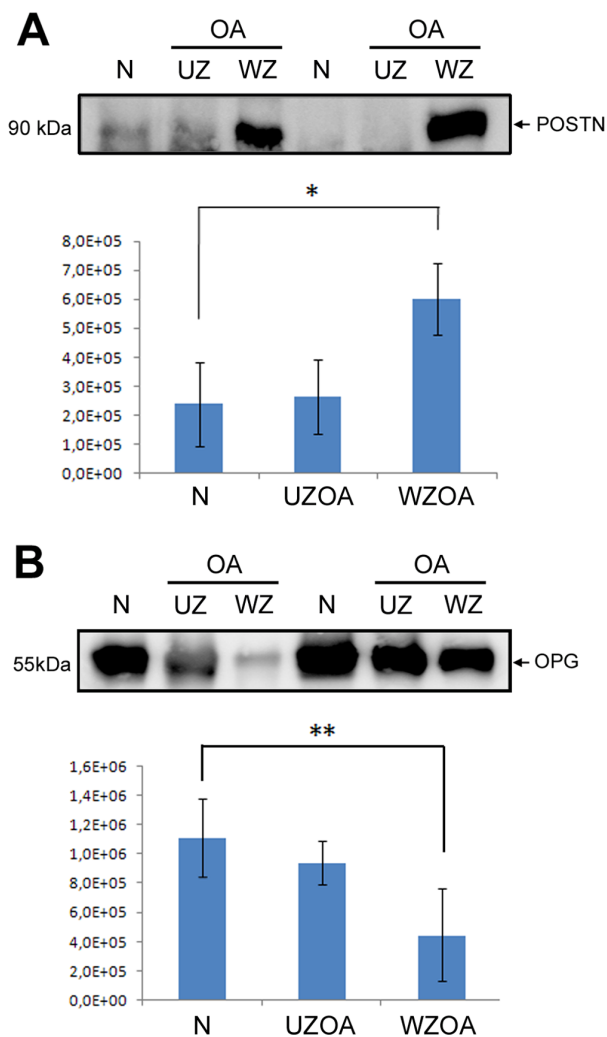


Figure 3. Verification experiments of POSTN (A) and OPG (B) altered secretion from OA cartilage. Western blot analyses were performed using conditioned media from explants of N, UZOA, and WZOA cartilages ($n = 12$, four biological replicates of each condition). Representative images of the blots are shown in the upper panels, whereas the results of the semiquantitative analysis performed by their densitometric analysis are shown at the bottom. *, $p = 0.03$; **, $p = 0.018$.

osteoarthritis patients. Because of the unavailability of this type of specimen from healthy donors, we decided to employ synovial fluid from patients suffering other rheumatic diseases (rheumatoid arthritis and psoriatic arthritis) as a comparator group.

The study was performed using a beads-based assay ($n = 28$, synovial fluids from 14 OA and 14 non-OA individuals). As shown in Figure 5, a significant ($p = 0.027$) increase of POSTN was found in the samples from OA patients (mean of 147.62 ng/mL POSTN) compared to non-OA (mean of 91.52 ng/mL).

DISCUSSION

Altogether, the data presented in this paper provide several tools for gaining knowledge about the mechanisms involved in the cartilage degradation processes that take place during

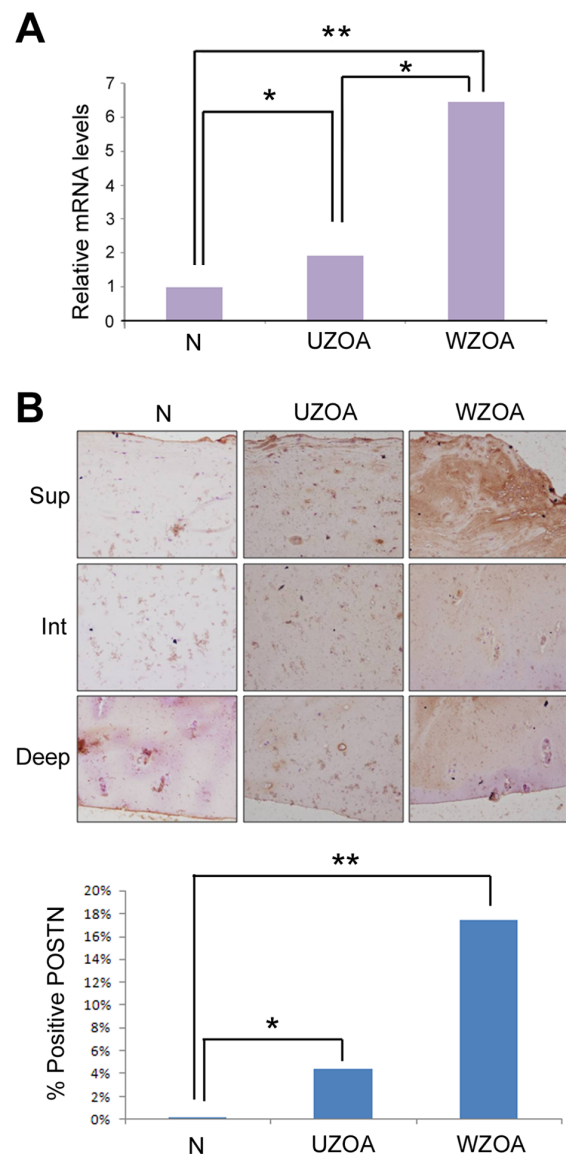


Figure 4. Periostin expression and protein abundance are increased in osteoarthritic cartilage. (A) Expression of periostin (POSTN) was evaluated by real-time PCR. Gene expression data were normalized against RPLP13A expression and are presented as the mean of 15 independent samples (5 N, 5 UZOA, and 5 WZOA), analyzed in duplicate and relativized to N cartilage (* $p < 0.05$; ** $p < 0.01$). (B) Representative images of the immunohistochemistry assays performed on human cartilage (5 N, 5 UZOA, and 5 WZOA), showing the different abundance of POSTN in the three cartilage layers. In the bottom panel, graphical representation of the densitometric analysis of the immunohistochemical images, showing the percentage of positivity of POSTN that was detected (*, $p < 0.05$; **, $p < 0.01$).

osteoarthritis and also for advancing in the search for novel and specific osteoarthritis biomarkers.

First, we describe a number of proteins that are differentially released to the media from the diseased cartilage. To allow finding differences between early and late stages of tissue degeneration, samples from unwounded (UZ) and wounded zones (WZ) from OA cartilage were separated for comparison. The quantitative proteomic analysis subsequently performed on

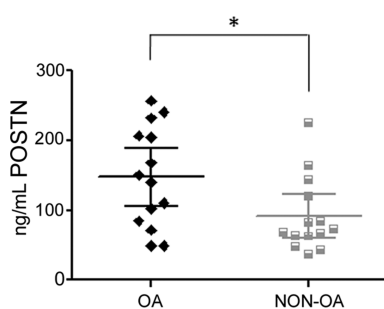


Figure 5. Periostin protein levels are increased in synovial fluid from OA patients when compared to other rheumatic diseases. Graphic shows the quantitative values obtained for POSTN in synovial fluids from patients suffering rheumatic diseases ($n = 28$, 14 OA and 14 non-OA). *, $p = 0.027$.

four biological replicates led to the identification of 76 proteins that can be assembled into three different release profiles: A, B, and C. Proteins with an A release profile are those whose secretion is already differential in macroscopically normal OA cartilage (UZOA) compared to healthy tissue, and then their differential release continues in WZOA. 42 proteins showed this profile and were grouped into clusters 1, 4, and 6 using the k-means algorithm, being putative early disease and progression markers. According to the enrichment analyses performed, these proteins are mainly involved in ECM organization processes (such as type II collagen, HPLN1, biglycan or tenascins). The profile B is formed by 32 proteins whose secretion is only detected as differential when the tissue is macroscopically damaged (WZOA). These were grouped into clusters 2 and 3 by the k-means algorithm and are putative useful predictors of progression or therapeutic response, but not of incidence. Members of this profile are proteins involved in collagen catabolic processes and ECM disassembly, such as the proteoglycan core protein (PGBM) and collagens type VI, XI, XII, or XV. Finally, a small group of three proteins showed a release profile C, with evident differential secretion in UZOA that disappears when the cartilage is wounded. These proteins were put together into cluster 5 and are related to inflammatory response.

The data summarized above provide highly valuable information about protein markers of cartilage degradation. It is essential to highlight the stringent cutoff values that we employed in the proteomic screening, both for protein identification and quantitation. The latter was set by bioinformatics analysis on the variability found between the biological replicates, which in our experiments were rather high. The most likely explanation for this is the fact that in our iTRAQ analysis we did not compared equal protein amounts from each condition, but equal secretome volumes obtained from cartilage explants of equal size: As an advantage of this, artifactual differences are diminished, by example those found on serum contaminants that were quite frequently identified in the cartilage secretomes. As drawbacks, probably due to these strict guidelines (cut off ratio >1.8) some proteins with a previously suggested biomarker value for OA were not included in the differential panel shown in Table 1. Among these, we can underscore cartilage oligomeric matrix protein (COMP),²² proteoglycan 4/lubricin (PRG4),²³ cartilage intermediate layer protein 1 (CILP1),²⁴ and Chitinase 3-like protein 1 (CH3L1/YKL40).²⁵ In all these cases, a UZOA/N increased ratio with highly significant p values was quantified (1.47 for COMP, 1.36

for PRG4, 1.74 for CILP1, and 1.63 for YKL40; see Supplementary Table S2). This demonstrates the usefulness of our approach for finding consistent and significant differences in protein release from the pathological tissue, but also invites a thorough review of all proteins identified in the present work (Supplementary Table S2) for uncovering other molecules with putative biomarker utility.

Interestingly, the largest group of proteins differentially released from cartilage was detected already in unwounded zones (profile A). We next focused on these proteins due to their potential value in the development of tests for early diagnosis. The altered secretion of two of these proteins, POSTN and OPG, was validated by Western blot and demonstrates the participation of cartilage in signaling processes of bone remodelling. Osteoprotegerin (OPG or TR11B/tumor necrosis factor receptor superfamily member 11B) is a widely expressed protein that is known to regulate bone density²⁶ by acting as decoy receptor for RANKL and thereby neutralizing its function in osteoclastogenesis.²⁷ The clear decreased presence of OPG that we found in secretomes from wounded OA cartilage might be an indicator of increased bone resorption in the diseased joint, and is consistent with previous findings of a reduced OPG:RANKL ratio observed in chondrocytes from OA patients.²⁸ However, we did not find any significant modulation of OPG in OA cartilage neither at the RNA level (by real-time PCR assays) nor at the protein level (by immunohistochemistry). Furthermore, quantification of OPG levels in synovial fluid from OA and non-OA patients also showed no significant differences (data not shown). These results are in contrast with previous reports that provided evidence of a significant association between OPG and OA. Valdes et al. pointed to an increased gene expression of OPG in arthritic cartilage compared to healthy tissue,²⁹ while other authors found higher levels of OPG in synovial fluid from OA patients compared to RA.³⁰ Moreover, a strong correlation between the severity of osteoarthritis (according to the Kellgren and Lawrence classification) and synovial OPG values was also reported by Pilichou et al.³¹ Because of this discrepancy, further investigations in a greater cohort of clinical samples are required to establish the marker value of OPG for OA.

On the other hand, the higher release of periostin from OA cartilage that we detected in this work went along with a significant increased protein expression in the diseased tissue, as detected by IHC. POSTN is a vitamin K-dependent protein produced by osteoblasts and also found in the periosteum,³² which may play a role in ECM mineralization by enhancing incorporation of bone morphogenetic protein 1 in the fibronectin matrix of connective tissues.³³ Interestingly, this protein was lately detected also in the cartilage matrix of a mouse model of OA obtained by surgical destabilization of the medial meniscus (DMM).³⁴ Using this model and by microarray analysis, these authors found an upregulation of POSTN gene expression in DMM vs sham joints. Furthermore, most recently a very significant upregulation of POSTN was also reported in human subchondral bone from osteoarthritis patients, when comparing OA medial with lateral tibial plateaus.³⁵ Additionally, the data presented in this work demonstrate for the first time the increase of periostin gene expression and protein abundance in OA human cartilage, which is accompanied by an augmented release of this protein from the diseased tissue to the media. This ultimately leads to the detection of increased levels in OA synovial fluid compared

to other rheumatic diseases, thus clearly suggesting a putative usefulness of POSTN as osteoarthritis biomarker.

Altogether, this study describes a panel of cartilage-related proteins that would not only be useful for the development of new diagnostic and monitoring tests for OA, but also for the design of new therapeutic possibilities by providing clinically relevant end points.³⁶ In this area, the potential biomarkers described herein could be employed in drug development or mesenchymal stem cell-based strategies for cartilage repair. Major efforts in the conception and implementation of efficient clinical trials (probably with the use of novel biochemical markers such as those reported in this work) are needed to solve those unmet clinical needs in osteoarthritis.

■ ASSOCIATED CONTENT

📄 Supporting Information

Supplementary Figure S1. Alterations in the secretomes of N, UZOA, and WZOA cartilage explants. Silver-stained SDS-PAGE gel showing the differences in the secretome profiles. Supplementary Figure S2. Differential secretion from N, UZOA and WZOA cartilage explants. Supplementary Figure S3. Determination of cutoff value for significant fold changes in the iTRAQ experiments. Supplementary Figure S4. K-means clustering analysis of the differentially released proteins. Supplementary Figure S5. Functional profiling of proteins differentially released from osteoarthritic cartilage. Supplementary Table S1. Crude identification and quantitation data obtained in the iTRAQ-based proteomic analysis. Supplementary Table S2. List of high-confidence identified proteins. Supplementary Table S3. GO enrichment analysis of the proteins identified in the secretomes of human articular cartilage explants. Supplementary Table S4. GO enrichment of the proteins belonging to each of the six clusters identified by the k-means algorithm. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Authors

*(F.J.B.) Phone: 34-981-176399; Fax: 34-981-176398 E-mail: francisco.blanco.garcia@sergas.es.

*(C.R.-R.) E-mail: cristina.ruiz.romero@sergas.es.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge the Pathology Service from the Orthopaedics Department of the CHU A Coruña for providing cartilage samples, and to Purificación Filgueira and Noa Goyanes for their help in the histological assessment of Mankin score and the histochemistry assays. This study was funded by grants from Fondo Investigación Sanitaria-Spain (PI11/02397, PI12/00329) and Secretaría I+D+I Xunta de Galicia (10CSA916058PR). J.M. (CA11/00050) and P.F.-P. (ProteoRed/ISCI) are supported by Fondo Investigación Sanitaria-Spain. C.R.-R. is supported by the Miguel Servet program from Fondo Investigación Sanitaria-Spain (CP09/00114).

■ REFERENCES

(1) *Global Burden of Disease Study*; WHO: Geneva, Switzerland, 2010.

(2) Wieland, H. A.; Michaelis, M.; Kirschbaum, B. J.; Rudolphi, K. A. Osteoarthritis - an untreatable disease? *Nat. Rev. Drug Discov* **2005**, *4* (4), 331–44.

(3) DeGroot, J.; Bank, R. A.; Tchetverikov, I.; Verzijl, N.; TeKoppele, J. M. Molecular markers for osteoarthritis: the road ahead. *Curr. Opin. Rheumatol.* **2002**, *14* (5), 585–9.

(4) Ruiz-Romero, C.; Blanco, F. J. Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis Cartilage* **2010**, *18* (4), 500–9.

(5) Clutterbuck, A. L.; Smith, J. R.; Allaway, D.; Harris, P.; Liddell, S.; Mobasher, A. High throughput proteomic analysis of the secretome in an explant model of articular cartilage inflammation. *J. Proteomics* **2011**, *74* (5), 704–15.

(6) Stevens, A. L.; Wishnok, J. S.; Chai, D. H.; Grodzinsky, A. J.; Tannenbaum, S. R. A sodium dodecyl sulfate-polyacrylamide gel electrophoresis-liquid chromatography tandem mass spectrometry analysis of bovine cartilage tissue response to mechanical compression injury and the inflammatory cytokines tumor necrosis factor alpha and interleukin-1 beta. *Arthritis Rheum.* **2008**, *58*, 489–500.

(7) Stevens, A. L.; Wishnok, J. S.; White, F. M.; Grodzinsky, A. J.; Tannenbaum, S. R. Mechanical injury and cytokines cause loss of cartilage integrity and upregulate proteins associated with catabolism, immunity, inflammation, and repair. *Mol. Cell Proteomics* **2009**, *8* (7), 1475–89.

(8) Wilson, R.; Belluoccio, D.; Little, C. B.; Fosang, A. J.; Bateman, J. F. Proteomic characterization of mouse cartilage degradation in vitro. *Arthritis Rheum.* **2008**, *58* (10), 3120–31.

(9) Swan, A. L.; Hillier, K. L.; Smith, J. R.; Allaway, D.; Liddell, S.; Bacardit, J.; Mobasher, A. Analysis of mass spectrometry data from the secretome of an explant model of articular cartilage exposed to pro-inflammatory and anti-inflammatory stimuli using machine learning. *BMC Musculoskelet. Disord.* **2013**, *14*, 349.

(10) Altman, R.; Asch, E.; Bloch, D.; Bole, G.; Borenstein, D.; Brandt, K.; Christy, W.; Cooke, T. D.; Greenwald, R.; Hochberg, M.; et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum.* **1986**, *29* (8), 1039–49.

(11) Aletaha, D.; Neogi, T.; Silman, A. J.; Funovits, J.; Felson, D. T.; Bingham, C. O., 3rd; Birnbaum, N. S.; Burmester, G. R.; Bykerk, V. P.; Cohen, M. D.; Combe, B.; Costenbader, K. H.; Dougados, M.; Emery, P.; Ferraccioli, G.; Hazes, J. M.; Hobbs, K.; Huizinga, T. W.; Kavanaugh, A.; Kay, J.; Kvien, T. K.; Laing, T.; Mease, P.; Menard, H. A.; Moreland, L. W.; Naden, R. L.; Pincus, T.; Smolen, J. S.; Stanislawski-Biernat, E.; Symmons, D.; Tak, P. P.; Upchurch, K. S.; Vencovsky, J.; Wolfe, F.; Hawker, G. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann. Rheum. Dis.* **2010**, *69* (9), 1580–8.

(12) Mankin, H. J.; Dorfman, H.; Lippiello, L.; Zarins, A. Biochemical and metabolic abnormalities in articular cartilage from osteoarthritic human hips. II. Correlation of morphology with biochemical and metabolic data. *J. Bone Joint Surg. Am.* **1971**, *53* (3), 523–37.

(13) Pascual Garrido, C.; Hakimiyan, A. A.; Rappoport, L.; Oegema, T. R.; Wimmer, M. A.; Chubinskaya, S. Anti-apoptotic treatments prevent cartilage degradation after acute trauma to human ankle cartilage. *Osteoarthritis Cartilage* **2009**, *17* (9), 1244–51.

(14) Calamia, V.; Ruiz-Romero, C.; Rocha, B.; Fernández-Puente, P.; Mateos, J.; Montell, E.; Vergés, J.; Blanco, F. J. Pharmacoproteomic study of the effects of chondroitin and glucosamine sulfate on human articular chondrocytes. *Arthritis Res. Ther.* **2010**, *12* (4), R138.

(15) Calamia, V.; Rocha, B.; Mateos, J.; Fernández-Puente, P.; Ruiz-Romero, C.; Blanco, F. J. Metabolic labeling of chondrocytes for the quantitative analysis of the interleukin-1-beta-mediated modulation of their intracellular and extracellular proteomes. *J. Proteome Res.* **2011**, *10* (8), 3701–11.

(16) Livak, K. J.; Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **2001**, *25* (4), 402–8.

- (17) Jin, J.; Ku, Y. H.; Kim, Y.; Kim, K.; Lee, J. Y.; Cho, Y. M.; Lee, H. K.; Park, K. S. Differential proteome profiling using iTRAQ in microalbuminuric and normoalbuminuric type 2 diabetic patients. *Exp Diabetes Res.* **2012**, *2012*, 168602.
- (18) Huang, d. W.; Sherman, B. T.; Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4* (1), 44–57.
- (19) Ehrlich, M. G.; Houle, P. A.; Vigliani, G.; Mankin, H. J. Correlation between articular cartilage collagenase activity and osteoarthritis. *Arthritis Rheum.* **1978**, *21* (7), 761–6.
- (20) Burr, D. B.; Gallant, M. A. Bone remodelling in osteoarthritis. *Nat. Rev. Rheumatol.* **2012**, *8* (11), 665–73.
- (21) Lotz, M.; Martel-Pelletier, J.; Christiansen, C.; Brandi, M. L.; Bruyere, O.; Chapurlat, R.; Collette, J.; Cooper, C.; Giacovelli, G.; Kanis, J. A.; Karsdal, M. A.; Kraus, V.; Lems, W. F.; Meulenbelt, I.; Pelletier, J. P.; Raynaud, J. P.; Reiter-Niesert, S.; Rizzoli, R.; Sandell, L. J.; Van Spil, W. E.; Reginster, J. Y. Value of biomarkers in osteoarthritis: current status and perspectives. *Ann. Rheum. Dis.* **2013**, *72* (11), 1756–63.
- (22) Verma, P.; Dalal, K. Serum cartilage oligomeric matrix protein (COMP) in knee osteoarthritis: a novel diagnostic and prognostic biomarker. *J. Orthop. Res.* **2013**, *31* (7), 999–1006.
- (23) Ruan, M. Z.; Erez, A.; Guse, K.; Dawson, B.; Bertin, T.; Chen, Y.; Jiang, M. M.; Yustein, J.; Gannon, F.; Lee, B. H. Proteoglycan 4 expression protects against the development of osteoarthritis. *Sci. Transl. Med.* **2013**, *5* (176), 176ra34.
- (24) Lorenzo, P.; Bayliss, M. T.; Heinegard, D. Altered patterns and synthesis of extracellular matrix macromolecules in early osteoarthritis. *Matrix Biol.* **2004**, *23* (6), 381–91.
- (25) Conrozier, T.; Carlier, M. C.; Mathieu, P.; Colson, F.; Debard, A. L.; Richard, S.; Favret, H.; Bienvenu, J.; Vignon, E. Serum levels of YKL-40 and C reactive protein in patients with hip osteoarthritis and healthy subjects: a cross sectional study. *Ann. Rheum. Dis.* **2000**, *59* (10), 828–31.
- (26) Simonet, W. S.; Lacey, D. L.; Dunstan, C. R.; Kelley, M.; Chang, M. S.; Luthy, R.; Nguyen, H. Q.; Wooden, S.; Bennett, L.; Boone, T.; Shimamoto, G.; DeRose, M.; Elliott, R.; Colombero, A.; Tan, H. L.; Trail, G.; Sullivan, J.; Davy, E.; Bucay, N.; Renshaw-Gegg, L.; Hughes, T. M.; Hill, D.; Pattison, W.; Campbell, P.; Sander, S.; Van, G.; Tarpley, J.; Derby, P.; Lee, R.; Boyle, W. J. Osteoprotegerin: a novel secreted protein involved in the regulation of bone density. *Cell* **1997**, *89* (2), 309–19.
- (27) Luan, X.; Lu, Q.; Jiang, Y.; Zhang, S.; Wang, Q.; Yuan, H.; Zhao, W.; Wang, J.; Wang, X. Crystal structure of human RANKL complexed with its decoy receptor osteoprotegerin. *J. Immunol.* **2012**, *189* (1), 245–52.
- (28) Kwan Tat, S.; Amiable, N.; Pelletier, J. P.; Boileau, C.; Lajeunesse, D.; Duval, N.; Martel-Pelletier, J. Modulation of OPG, RANK and RANKL by human chondrocytes and their implication during osteoarthritis. *Rheumatology (Oxford)* **2009**, *48* (12), 1482–90.
- (29) Valdes, A. M.; Hart, D. J.; Jones, K. A.; Surdulescu, G.; Swarbrick, P.; Doyle, D. V.; Schafer, A. J.; Spector, T. D. Association study of candidate genes for the prevalence and progression of knee osteoarthritis. *Arthritis Rheum.* **2004**, *50* (8), 2497–507.
- (30) Skoumal, M.; Kolarz, G.; Haberhauer, G.; Woloszczuk, W.; Hawa, G.; Klingler, A. Osteoprotegerin and the receptor activator of NF-kappa B ligand in the serum and synovial fluid. A comparison of patients with longstanding rheumatoid arthritis and osteoarthritis. *Rheumatol Int.* **2005**, *26* (1), 63–9.
- (31) Pilichou, A.; Papassotiropoulos, I.; Michalakakou, K.; Fessatou, S.; Fandridis, E.; Papachristou, G.; Terpos, E. High levels of synovial fluid osteoprotegerin (OPG) and increased serum ratio of receptor activator of nuclear factor-kappa B ligand (RANKL) to OPG correlate with disease severity in patients with primary knee osteoarthritis. *Clin. Biochem.* **2008**, *41* (9), 746–9.
- (32) Couto, D. L.; Wu, J. H.; Monette, A.; Rivard, G. E.; Blostein, M. D.; Galipeau, J. Periostin, a member of a novel family of vitamin K-dependent proteins, is expressed by mesenchymal stromal cells. *J. Biol. Chem.* **2008**, *283* (26), 17991–8001.
- (33) Gillan, L.; Matei, D.; Fishman, D. A.; Gerbin, C. S.; Karlan, B. Y.; Chang, D. D. Periostin secreted by epithelial ovarian carcinoma is a ligand for alpha(V)beta(3) and alpha(V)beta(5) integrins and promotes cell motility. *Cancer Res.* **2002**, *62* (18), 5358–64.
- (34) Loeser, R. F.; Olex, A. L.; McNulty, M. A.; Carlson, C. S.; Callahan, M. F.; Ferguson, C. M.; Chou, J.; Leng, X.; Fetrow, J. S. Microarray analysis reveals age-related differences in gene expression during the development of osteoarthritis in mice. *Arthritis Rheum.* **2012**, *64* (3), 705–17.
- (35) Chou, C. H.; Wu, C. C.; Song, I. W.; Chuang, H. P.; Lu, L. S.; Chang, J. H.; Kuo, S. Y.; Lee, C. H.; Wu, J. Y.; Chen, Y. T.; Kraus, V. B.; Lee, M. T. Genome-wide expression profiles of subchondral bone in osteoarthritis. *Arthritis Res. Ther.* **2013**, *15* (6), R190.
- (36) Kraus, V. B.; Burnett, B.; Coindreau, J.; Cottrell, S.; Eyre, D.; Gendreau, M.; Gardiner, J.; Garner, P.; Hardin, J.; Henrotin, Y.; Heinegard, D.; Ko, A.; Lohmander, L. S.; Matthews, G.; Menetski, J.; Moskowitz, R.; Persiani, S.; Poole, A. R.; Rousseau, J. C.; Todman, M. Application of biomarkers in the development of drugs intended for the treatment of osteoarthritis. *Osteoarthritis Cartilage* **2011**, *19* (5), 515–42.

Capítulo II: “Differential protein profiling of synovial fluid from rheumatoid arthritis and osteoarthritis patients using LC–MALDI TOF/TOF”

El líquido sinovial es una muestra que presenta una gran ventaja como fuente para la búsqueda de biomarcadores de artrosis debido a su proximidad a los compartimentos más afectados por la enfermedad como son el cartílago, el hueso y la membrana sinovial. El uso, por tanto, de muestras de líquido sinovial posibilita identificar un mayor número de proteínas relacionadas con los procesos que están teniendo lugar en la articulación durante el proceso artrósico, ya que se detectan antes de su dilución en el torrente sanguíneo.

La hipótesis que nos planteamos para este trabajo es que el perfil de las proteínas presentes en líquidos sinoviales de pacientes diagnosticados con OA es diferente del perfil proteico de donantes con otras enfermedades reumáticas como la artritis reumatoide (AR). Las proteínas que muestran alteración entre estas dos enfermedades reumáticas podrían ser potenciales marcadores específicos de artrosis.

Con esta premisa, el objetivo que nos propusimos en este trabajo fue el de identificar las proteínas en el líquido sinovial de pacientes con artrosis y comparar el perfil proteico de estos pacientes con los pacientes con artritis reumatoide utilizando cromatografía líquida de alta resolución acoplada a espectrometría de masas.

Para ello, utilizamos 20 muestras de líquido sinovial procedente de pacientes con OA y 20 de pacientes con AR. Cada grupo de estudio, se agrupó en 2 *pools* de 10 muestras cada uno. Posteriormente, utilizamos una columna de inmovilización para deplecionar las 20 proteínas más abundantes del líquido sinovial y de esta forma poder aumentar el rango de proteínas minoritarias que puedan ser identificadas con potencial biomarcador. A continuación, las proteínas de cada *pool* de muestras deplecionadas se separaron por electroforesis en gel de una dimensión (SDS-PAGE).

Para el análisis por espectrometría de masas, los *pools* se dividieron en dos réplicas técnicas. Las bandas se escindieron del gel y posteriormente, se aplicó el protocolo de digestión en gel. Los péptidos se separaron por nanocromatografía líquida de fase reversa a pH ácido y las fracciones resultantes se analizaron utilizando la espectrometría de masas del tipo MALDI-TOF/TOF.

Finalmente, se realizó la identificación y cuantificación relativa de las proteínas en las dos condiciones de estudio mediante la técnica *label-free* basada en contaje espectral. En total, se identificaron 136 proteínas, 93 proteínas se encontraron presentes en las dos enfermedades, 29 proteínas sólo se identificaron en pacientes AR

y 14 eran específicas para OA. La mayoría de las proteínas aumentadas en el líquido sinovial OA eran proteínas estructurales de la matriz extracelular del cartílago como *fibronectin* (FINC), *kininogen-1* (KNG-1), *cartilage acidic protein 1* (CRAC1) y *cartilage oligomeric matrix protein* (COMP). Nuestros resultados mostraron también una menor abundancia relativa en los líquidos sinoviales de los pacientes con OA de proteínas relacionadas con la inflamación y la respuesta inmune, tales como las metaloproteasas de la matriz y varias proteínas relacionadas con la activación de neutrófilos y el complemento. Posteriormente, algunos de los resultados obtenidos fueron validados por inmunodetección. Concretamente, utilizando 8 pacientes OA y 8 pacientes AR validamos que los niveles de las proteínas FINC y *gelsolin* (GELS) estaban incrementados en pacientes OA con respecto a pacientes AR. Utilizando 12 pacientes OA y 12 pacientes AR, validamos también por inmunodetección que los niveles de las proteínas *metalloprotease-1* (MMP-1) y *transforming-growth factor-beta* (BIGH3) estaban disminuidos en pacientes OA con respecto a pacientes AR.

El perfil proteico identificado en este trabajo refleja diferencias entre las dos enfermedades estudiadas: la AR es una enfermedad con un fenotipo más inflamatorio mientras que la OA muestra un incremento en la abundancia de las proteínas relacionadas con la formación y remodelado de la matriz extracelular, como consecuencia de la degradación del cartílago.

La conclusión principal de este trabajo es que alguna de las proteínas identificadas en este trabajo pueden ser relevantes para aumentar el conocimiento en la etiología y patogénesis de la OA y también pueden ser útiles como potenciales biomarcadores de la enfermedad debido a su presencia en el líquido sinovial antes de su dilución en el suero.

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot

Differential protein profiling of synovial fluid from rheumatoid arthritis and osteoarthritis patients using LC–MALDI TOF/TOF [☆]

Jesús Mateos^a, Lucía Lourido^a, Patricia Fernández-Puente^a,
Valentina Calamia^a, Carlos Fernández-López^b, Natividad Oreiro^b,
Cristina Ruiz-Romero^a, Francisco J. Blanco^{a, b, c, *}

^aRheumatology Division, ProteoRed/ISCIII Proteomics Group, INIBIC-Hospital Universitario de A Coruña, 15006-A Coruña, Spain

^bRheumatology Division, Research Clinic Unit, INIBIC-Hospital Universitario de A Coruña, 15006-A Coruña, Spain

^cCIBER-BBN Instituto de Salud Carlos III, A Coruña, Spain

ARTICLE INFO

Available online 8 January 2012

Keywords:

Synovial fluid
Osteoarthritis
Rheumatoid arthritis
Biomarkers
Proteomics

ABSTRACT

The purpose of this study was to identify those proteins relatively more abundant in the synovial fluid (SF) of patients suffering from rheumatoid arthritis (RA) and osteoarthritis (OA) using high performance liquid chromatography coupled to mass spectrometry. 20 individual SF samples from each disease were pooled into two groups (RA and OA) to reduce the contribution of extreme individual values. Prior to the proteomic analysis, samples were immunodepleted from the top 20 most abundant plasma proteins, to enrich the lower-abundance protein fractions. Then, they were subjected to protein size fractionation and in-gel digestion, followed by reversed-phase peptide separation in a nano-LC system and subsequent peptide identification by MALDI-TOF/TOF. This strategy led to the identification of 136 different proteins in SF, which is the largest number of SF proteins described up to date by proteomics.

A relative quantification of the proteins between RA and OA was carried out by spectral counting analysis. In RA, our results show a greater relative abundance of proteins related to complement activation, inflammation and the immune response, such as the major matrix metalloproteinases and several neutrophil-related proteins. In OA, we detected an increase in proteins involved in the formation and remodeling of the extracellular matrix (ECM), such as fibronectin, kininogen-1, cartilage acidic protein 1 and cartilage oligomeric matrix protein. The results obtained for MMP-1, BGH3, fibronectin and gelsolin were verified by immunoblotting analyses. Some of the novel proteins identified in this work might be relevant not only for increasing knowledge on the etiopathogenesis of RA and OA processes, but also as putative disease biomarkers, as their presence in SF is a prior step to their dilution in serum. This article is part of a Special Issue entitled: Proteomics: The clinical link.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Synovial fluid (SF) is a serum filtrate located in the joint, where it also receives protein contributions from the surrounding

tissues, articular cartilage, synovial membrane and bone [1]. In a healthy joint, articular cartilage and SF cushion the bones and allow them to move easily. Disorders causing destruction of the joint, such as osteoarthritis (OA) and rheumatoid arthritis

[☆] This article is part of a Special Issue entitled: Proteomics: The clinical link.

* Corresponding author at: Rheumatology Division, INIBIC-Hospital Universitario A Coruña, C/Xubias, 84, 15006-A Coruña, Spain. Tel.: +34 981 176399; fax: +34 981 176398.

E-mail address: fblagar@sergas.es (F.J. Blanco).

1874-3919/\$ – see front matter © 2012 Elsevier B.V. All rights reserved.

doi:10.1016/j.jprot.2011.12.042

(RA), are the most common rheumatic pathologies. Although both diseases result in the destruction of articular cartilage, they have very different etiopathogeneses [2,3]. RA is an autoimmune disease that correlates with a strong inflammation that is responsible for cartilage destruction, mediated by the activation of macrophages and B and T lymphocytes. OA is characterized by alterations in metabolism and cell signaling, and by an imbalance of redox mechanisms leading to degradation of cartilage and surrounding tissues through the action of activated matrix metalloproteinases (MMPs) [4,5]. The identification of a biomarker or a panel of biomarkers related to the OA process would be of great clinical importance not only for the early diagnosis of the disease, but also for its treatment and the development of experimental and clinical therapeutic trials and promoting a deeper understanding of the disease. Because SF is in direct contact with the affected tissues, cartilage, synovial membrane and bone, it could be an excellent source for discovery of biomarkers [6]. Furthermore, the identification of a differential proteomic signature of these processes will increase our understanding of their underlying molecular basis.

Liquid chromatography coupled to tandem mass spectrometry (LC-MS) has been shown to be a powerful tool for protein profiling in biological fluids and, in recent years, has been widely used for large-scale identification of complex protein mixtures [7]. The high dynamic range in the protein content of body fluids, such as serum, plasma, urine, or synovial fluid, creates the need for the preliminary steps of depleting highly abundant proteins and multi-fractionating the sample [8]. Furthermore, recent improvements in the robustness and mass accuracy of LC-MS platforms make label-free quantitative methods an attractive alternative to such classical labeling methods as isobaric labeling such as iTRAQ, 18O, and others [9]. We have designed a label-free relative quantitative approach using spectral counting, a well-documented label free quantitative technique [10] to identify SF proteins, and we have compared the differential protein profiles between RA and OA. Among the different label-free quantitative proteomics strategies, which essentially use the relationship between sampling statistics and protein abundance [11], spectral counting has proven its utility for profiling complex proteomes [12]. The information about the proteins identified in this work would help us to understand the etiopathogenesis of RA and OA at the molecular level. Moreover, these proteins might also hold an important value as joint-specific disease biomarkers.

2. Materials and methods

2.1. Patients

Patients with RA and OA selected for this study were diagnosed following the criteria determined by the American College of Rheumatology (ACR, 1988). The mean age of the RA subjects was 56 ± 10 years old ($n=20$), while that of the OA subjects was 73 ± 8 years old ($n=20$). All SF samples were collected after informed consent from knee joints by arthrocentesis in the Rheumatology Department at Hospital Universitario de A Coruña, following our institutional regulations and procedures for sample collection. The study was approved by the local Ethics Committee (Galicia, Spain). The SF samples were visually inspected for blood contamination and stored in liquid nitrogen.

2.2. Sample processing

Four separate pools, two for RA and two for OA, were created each containing 200 μ L of SF from ten individual SF samples with similar total protein concentrations, ranging from 25 to 30 mg/mL as determined by direct measurement at 280 nm in a nano-drop instrument. A 1 mL aliquot from each pool was treated with 1 μ g/mL of hyaluronidase (Sigma-Aldrich, St. Louis, (MO), USA) at 37 °C for 1 h to reduce viscosity and then centrifuged for 5 min at 1000 \times g. The supernatant was re-centrifuged through a 0.22 μ m cellulose acetate membrane (CLS810, Costar Spin-X, Corning, Tokio, Japan) tube filter at 5000 \times g for 15 min to remove insoluble material and debris. The filtered flow-through was collected, aliquoted and stored at -80 °C pending immuno-depletion.

2.3. Immunodepletion

Filtered SF pooled samples were processed to deplete the twenty most abundant proteins in serum and SF using a ProteoPrep20 column (Sigma-Aldrich, St. Louis, (MO), USA) in a 1200 HPLC system (Agilent, Wilmington, DE, USA). Injections of 100 μ L were done following the manufacturer's instructions. ProteoPrep Equilibration Buffer (Phosphate Buffered Saline (PBS), Buffer A; Sigma-Aldrich) was used for equilibration and washing steps. The flow-through fractions containing the lower-abundance proteins were collected. ProteoPrep Elution buffer (0.1 M glycine-HCl, pH 2.5, and Octyl β -D-glucopyranoside, Buffer B; Sigma-Aldrich) was used to elute the bound highly abundant proteins from the column. Lower-abundance protein fractions were combined and concentrated by Amicon filters (MWCO 3000 kDa, Millipore, Billerica, (MA), USA) to a final volume of approximately 150 μ L. The total protein concentration was determined by direct measurement at 280 nm and ranged from 1.3 to 1.8 μ g/ μ L.

2.4. Two-dimensional gel electrophoresis analysis

2D-PAGE analysis of SF samples was performed as described previously [1]. Briefly, 100 μ g of proteins was applied to 24 cm, pH 3–10 NL, IPG strips (GE Healthcare) using passive overnight rehydration. Isoelectric focusing was performed at 20 °C for a total of 64,000 Vhr (IPGphor, GE Healthcare). Second dimension was carried out on 10% T, 2.6% C polyacrylamide gels according to the procedure of Laemmli, but using Tris-glycine electrophoresis buffer for the lower (anode) buffer and 2 \times Tris-glycine for the upper (cathode) buffer. Samples were run at 2.5 W/gel for the first 30 min and then at 17 W/gel until the bromophenol blue dye reached the bottom of the gel. Finally, gels were stained with silver nitrate as described in our previous reports [13].

2.5. In-gel fractioning

For each pool, 30 μ g of protein was loaded and separated by SDS-PAGE in 10% acrylamide in-house gels in a Protean@ mini-gel system (Bio-Rad, Hercules, (CA), USA). The gels were lightly stained with coomassie blue for 10 min to check protein integrity. Each entire lane was size-fractionated into 15 sections that were subsequently processed independently. Each section was de-stained with methanol, divided into two

identical parts to create two experimental replicates, diced into small pieces and in-gel digested following a standard procedure [14]. Briefly, the samples were desiccated with acetonitrile, reduced with DTT, alkylated with iodoacetamide and trypsin-digested (6 ng/ μ L, Promega MSgold, Madison, (WI), USA) for 16 h at 37 °C. Peptides were then extracted, dried in a speed-vac, reconstituted in 0.1% trifluoroacetic acid and de-salted using nu-tipC18 (Glygen, Columbia, (MA), USA).

2.6. NanoLC–MALDI-TOF/TOF

The peptide fractions were separated using reversed phase chromatography in a nanoLC system (Tempo, Eksigent, Dublin, (CA), USA) by loading through a trapping column into a C18 silica-based column (New Objective, Woburn, MA, USA) with an internal diameter of 300 Å. Peptides were eluted at a flow rate of 0.35 μ L/min during a 1 hour linear gradient from 2 to 40%B (mobile phase A: 0.1% trifluoroacetic acid 2% acetonitrile, mobile phase B: 0.1% trifluoroacetic acid 95% acetonitrile), mixed with α -cyano matrix (4 mg/mL at a flow rate of 1.2 μ L/min) and deposited onto a MALDI plate using an automatic spotter (SunCollect, Sunchrome, Friedrichsdorf, Germany). Chromatograms corresponding to each gel section were composed by 240 spots, each one comprising a 15 s deposition.

The MS runs for each chromatogram were acquired and analyzed in a MALDI-TOF/TOF instrument (4800 ABSciex, Framingham, (MA), USA) using a fixed laser intensity of 3800 kV and 1500 shots/spectrum. Automated precursor selection was done using a Job-Wide interpretation method (up to 10 precursors/fraction, signal-to-noise lower threshold 50) with a laser voltage of 4800 and 1500 shots/spectrum at medium CID collision energy range. A second Job-Wide precursor selection was done excluding those precursors previously fragmented and using a lower signal-to-noise threshold of 30, to identify peptides coming from low-abundance proteins. Data from both MS/MS acquisitions were used for data processing and subsequent protein identification.

2.7. Database search and protein identification

The complete MS and MS/MS raw data for both diseases were processed separately using the Phenyx 2.6 software platform (Geneva Informatics, Geneva, Switzerland). Peptide identification was performed against the last uniprot_sprot-release (2010_09 of 10-Aug-2010, 519348 sequence entries). Search parameters included carbamidomethylation of cysteines (fixed), oxidation of methionine (variable, ≤ 4), 1 missed trypsin normal cleavage, precursor ion mass range 800–4000 Da, ± 50 ppm tolerance on precursor and ± 0.3 Da on fragmentation of ions. The scoring model was defined by the algorithm and peptide acceptance parameters were peptide length ≥ 5 , minimum peptide z-score of 5.0, maximum peptide p-value of $1.0E-7$ and an AC score of 8.0. False Discovery Rate (FDR) was estimated by searching against both uniprot_sprot and uniprot_sprot decoy databases.

2.8. Relative protein quantitation

Phenyx 2.6 Spectral Counting module (Geneva Informatics, Geneva, Switzerland) was used for relative quantification comparing OA and RA samples as previously described [15,16]. Briefly,

the four RA and OA replicates were classified as the “RA state” and “OA state,” respectively. The software calculates individual ratios between MS/MS runs belonging to different states. In this case, we obtained up to sixteen individual ratios for RA/OA and used them to calculate an average ratio that indicates the relative abundance of each protein between RA and OA samples with their standard deviation (SD) to assess reproducibility of the individual ratios for each protein when comparing the replicates.

2.9. Gene ontology (GO) classification and pathway studio analysis

The Panther Classification System (www.pantherdb.org) was used to categorize the SF proteins according to their gene ontology (GO) biological process definition, as previously described [17]. Pathway studio software (Ariadne Genomics, Rockville, (MD), USA) was used to visualize molecular pathways of those proteins with different presence in the two diseases as previously described [18].

2.10. Immunoblotting procedures

Western-blot analysis was performed on 20 μ g of total protein from individual SF samples as previously described [19]. The primary antibodies used were mouse monoclonal anti-human MMP-1 (Millipore, Billerica, (MA), USA), mouse polyclonal anti-human BGH3, mouse monoclonal anti-gelsolin and rabbit polyclonal anti-fibronectin (all from Santa Cruz Biotech., Santa Cruz, (CA), USA). Secondary anti-mouse and anti-rabbit antibodies (GE Healthcare, Amersham Biotechnology, UK) were used to visualize proteins by the ECL™ Western Blotting Analysis System (GE Healthcare, Amersham Biotechnology, UK). Optimal concentrations for each antibody were determined empirically. Blots were digitized using the LAS 3000 image analyzer (GE Healthcare). Densitometric analysis of the band intensities was performed using ImageQuant 5.2 software (GE Healthcare). The statistical significance of these results was determined using the Microsoft Excel statistical package.

3. Results and discussion

3.1. Handling and immunodepletion of SF samples for proteomic analysis

Synovial fluid is a potential compartment for joint disease biomarkers, as it is derived directly from the site where the pathology occurs. Consequently, many proteomic strategies in the rheumatology field have been designed to identify putative biomarkers in SF, and then try to validate them in serum [6].

SF can be quite easily obtained from knee joints by aseptical aspiration, avoiding blood contamination. Due to its high viscosity, hyaluronidase treatment is a recommended procedure to digest the hyaluronic acid network of the samples prior to the proteomic analysis [1]. As a serum filtrate, a high quantitative intersample variability has been reported for a number of proteins by a proteomic approach [20], thus indicating the need for analyzing a high number of samples in order to determine disease-dependent protein changes. To minimize this

variability, we performed the analysis on SF pools, each representing ten individual samples. This sample pooling approach has been recently followed by our group for the discovery of novel OA biomarkers in serum [21], demonstrating its usefulness.

SF fluid has a similar high dynamic range as serum [22], with a reduced number of high abundant proteins (representing almost 95% of the total protein volume) that mask the low abundant ones [23]. Immunodepletion of these high abundant proteins by affinity chromatography has been commonly accepted as a practical and effective strategy to improve detection and determination of low-abundant proteins [24], which usually have a higher biomarker value. Therefore, we have employed for the first time on SF samples a multi-affinity removal system that removes the top 20 proteins in plasma [25]. To investigate the efficiency of the immunodepletion on the SF proteome, we resolved non-depleted and depleted samples by 2D-PAGE, and compared their protein profiles. As shown in Fig. 1, the immunodepletion step markedly removes most of the albumin, IgG and other high abundant proteins present in the sample, as well as enhances the detection of a remarkably higher number of proteins. These data demonstrate the usefulness of the affinity removal system for proteomic studies on SF.

3.2. Identification of SF proteins from immunodepleted RA and OA samples

A proteomic analysis was then carried out on the depleted SF samples obtained from RA and OA patients. We analyzed two sample pools from each condition, each representing ten individual samples. The pools were further divided into two experimental replicates, thus creating four replicates from each disease. The analysis led to the identification of 122 different proteins in RA SF and 107 in OA SF, using Phenyx 2.0 as the search engine and combining all the peak lists for a given condition for the identifications.

In RA SF samples, 2498 valid peptides were used to identify a total of 122 proteins, 120 of them with at least two valid peptides. A global FDR of less than 2% was calculated by running a parallel search in a decoy database (data not shown). A complete list of the identified proteins is found in Supplementary Table 1. In OA SF samples, 2582 valid peptides were used to identify a total of 107 proteins, 105 of them with at least two valid peptides. A global FDR of 3% was calculated by running a parallel search in a decoy database (data not shown). A complete list of the identified proteins is shown in Supplementary Table 2. Fig. 2 shows the complementarities of RA and OA SF samples at both peptide and protein levels: From all identified proteins, a total of 93 were common in SF from the two pathologies, 29 specific for RA and 14 for OA, for a total of 136 different proteins identified. This is, to our knowledge, the largest proteomic analysis performed up to date on human synovial fluid. A previous published work [26] listed 135 unique GenInfo accession-identified proteins in OA SF, but this number is reduced to 99 different proteins. We now report a deeper insight on SF proteome, mainly due to the immunodepletion step that we have performed. Gobezie et al. [26] identified in their work all those abundant proteins that are targeted by the immunodepletion column. In our approach,

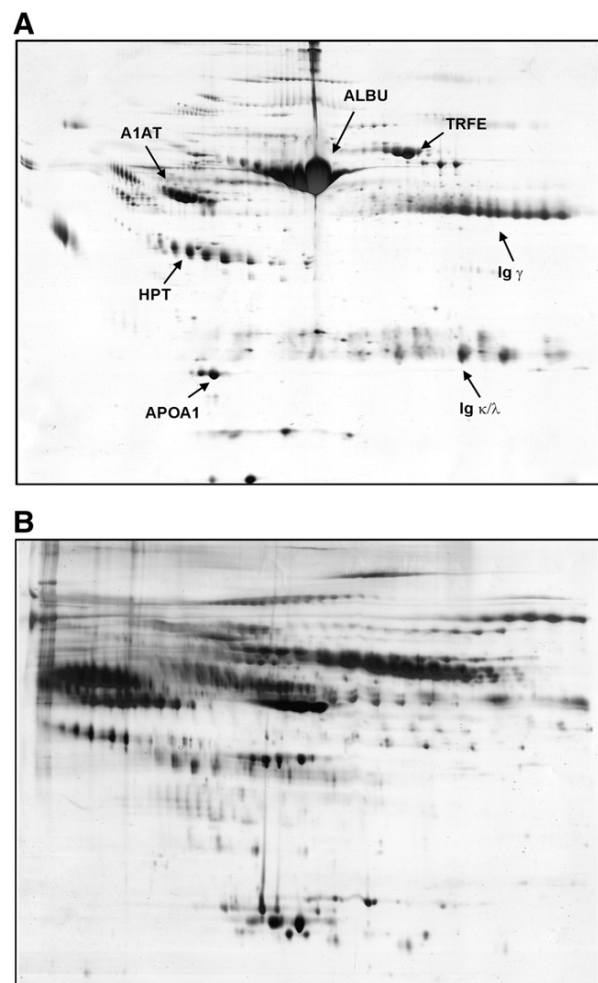


Fig. 1 – 2D-PAGE analysis of non-depleted (A) and immunodepleted (B) SF samples. Proteins were resolved on pH 3–10 NL (24 cm) IPG strips in the first dimension, and on 10% T polyacrylamide gels in the second dimension. Gels were stained with silver nitrate. Some of the typical most abundant plasma proteins are pointed in the non-depleted map. A1AT: α_1 -antitrypsin; ALBU: serum albumin; APOA1: apolipoprotein A-I; HPT: haptoglobin; Ig γ : immunoglobulin heavy chains; Ig κ/λ : immunoglobulin light chains; TRFE: transferrin.

only four of these “Top 20” proteins were identified, and always below the 10 highest scored identifications: serum albumin (rank 18 in RA and 17 in OA), apolipoprotein A-I, Ig κ and fibrinogen.

3.3. Spectral counting analysis shows differentially abundant proteins in RA compared to OA

Once the differential protein profiles of OA and RA SF have been characterized, we decided to use MS data to relatively quantify those differentially accumulated proteins. Previous works have been focused on the same topic, but employing

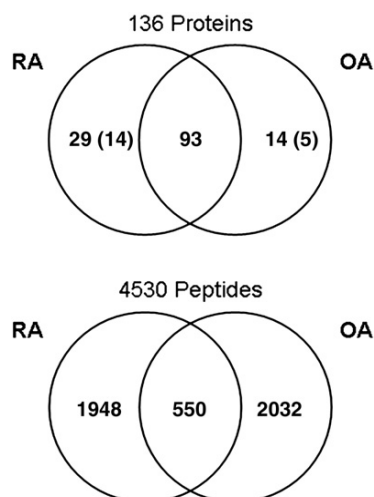


Fig. 2 – Summary of the LC-MS/MS data of the SF samples. Venn diagram representing the overlap of proteins (top) and unique peptides (bottom) identified by MS/MS among the two pathological processes analyzed, RA and OA. Numbers in parenthesis represent the number of proteins identified only in one condition with at least 4 unique peptides. Phenyx 2.6 was used as search engine.

substantially different strategies, such as two-dimensional gel electrophoresis [27,28], or surface-enhanced laser desorption/ionization (SELDI) [29,30]. Other recent studies, based on LC-MS/MS analysis, have been focused on analyzing specific SF sub-proteomes for the search of OA biomarkers, such as the group of SF high abundant proteins [26], SF endogenous

peptides [31], or peptides bound to weak cation exchange chromatography magnetic beads [32].

The method that we have employed, based on spectral counting for relative quantification, is a well-documented technique suitable for this purpose, that avoids the time- and cost-consuming in vitro labeling of the samples [10]. In most cases, label-free quantification approaches have been performed using online LC-ESI-MS/MS platforms, since they are the most reproducible and accurate MS techniques at present. Nevertheless, several previous studies have demonstrated the utility of nanoLC coupled offline to MALDI-TOF/TOF-MS for comparative analyses [33–35]. Therefore, we have employed the latter strategy to identify those proteins that are relatively more abundant in RA SF (Table 1) or in OA SF (Table 2). The quantification has been performed using Unique peptides for each protein, since this kind of peptides provides better correlation with protein abundance [36]. Those proteins that were only detected in RA or OA are classified as detected/not detected or not detected/detected, respectively. It is important to clarify that extreme ratios do not necessarily indicate that a given protein is absent in one of the conditions, only that the concentration in SF of such protein is below the detection limit of the followed technique. An example of this issue occurs for MMP-1: we identified this protein only in RA samples, although its presence in synovial fluid from OA patients has been demonstrated previously by a targeted (and therefore more sensitive) assay [37].

A number of other proteins that were identified in both conditions were found to be relatively more abundant in RA or in OA, as calculated by spectral counting analysis. 1.2 and 0.8 were the cut-off fold change of the mean ratio necessary for a protein to be considered as relatively more abundant in RA or OA, respectively. In this case, a wide range of SD values was detected, indicating that the reproducibility of the

Table 1 – Proteins identified as relatively more abundant in rheumatoid arthritis (RA) synovial fluid.

| SwissProt ID | Protein name | Valid unique peptides/score RA ^{a)} | Valid unique peptides/score OA ^{a)} | Mean Ratio RA/OA ^{b)} | SD ^{c)} | Frequency RA/OA ^{d)} |
|--------------|---|--|--|--------------------------------|------------------|-------------------------------|
| CAP7_HUMAN | Azurocidin | 8/31.8 | – | d/n.d.* | – | 3/0 |
| CO4A_HUMAN | Complement C4gamma chain | 8/27.4 | – | d/n.d.* | – | 4/0 |
| APOB_HUMAN | Apolipoprotein B-48 | 6/26.4 | – | d/n.d.* | – | 2/0 |
| MMP1_HUMAN | 27 kDa interstitial collagenase | 7/25.8 | – | d/n.d.* | – | 3/0 |
| ILEU_HUMAN | Leukocyte elastase inhibitor | 6/23.1 | – | d/n.d.* | – | 2/0 |
| PZP_HUMAN | Pregnancy zone protein | 7/22.0 | – | d/n.d.* | – | 3/0 |
| BGH3_HUMAN | Transforming growth factor-beta-induced protein ig-h3 | 6/19.7 | – | d/n.d.* | – | 3/0 |
| H4_HUMAN | Histone H4 | 5/19.5 | – | d/n.d.* | – | 4/0 |
| NGAL_HUMAN | Neutrophil gelatinase-associated lipocalin | 4/18.1 | – | d/n.d.* | – | 2/0 |
| DEF1_HUMAN | Neutrophil defensin 2 | 4/17.9 | – | d/n.d.* | – | 2/0 |
| S10A8_HUMAN | Protein S100-A8 | 18/62.3 | – | d/n.d.* | – | 4/0 |
| PLSL_HUMAN | Plastin-2 | 18/91.5 | – | d/n.d.* | – | 4/0 |
| CO8G_HUMAN | Complement component C8 gamma chain | 11/52.1 | – | d/n.d.* | – | 4/0 |
| PRG4_HUMAN | Proteoglycan 4 C-terminal part | 12/56.7 | 2/9.1 | 3.86 | 2.47 | 3/1 |
| APOE_HUMAN | Apolipoprotein E | 17/71.2 | 10/47.8 | 2.97 | 0.8 | 3/4 |
| MMP3_HUMAN | Stromelysin-1 | 47/160.9 | 18/69.0 | 2.73 | 1.76 | 4/4 |
| FCN3_HUMAN | Ficolin-3 | 10/40.7 | 2/13.1 | 1.70 | 1.07 | 3/1 |
| CO8A_HUMAN | Complement component C8 alpha chain | 14/61.5 | 7/28.6 | 1.67 | 0.99 | 4/4 |

a) Valid unique peptides used by Phenyx 2.6 and AC score for each identified protein.

b) Relative abundance calculated by Spectral Counting using Phenyx 2.6 software. *Detected/not detected denotes proteins only detected in RA samples.

c) Reproducibility of individual ratios.

d) Number of replicates for each condition in which the protein was identified.

method should be improved. For these reason, data should be assessed cautiously, particularly for those proteins identified using a small number of peptides. Therefore, in this work we have considered as differently abundant only those proteins identified by at least four unique peptides in one of the diseases. The reproducibility of our method was also determined by analyzing the number of replicates from each condition in which a specific protein was detected (Frequency, Tables 1 and 2).

The analysis reveals quantitative differences in the SF protein profile between RA and OA (Tables 1 and 2). In RA, it is clear that those proteins occurring at higher frequencies are involved in three primary processes: immune response, complement activation and inflammation (Table 1). In OA samples, we found an increased accumulation of proteins that are either structural components of the cartilage extracellular matrix (ECM), such as aggrecan, cartilage oligomeric matrix protein (COMP) and cartilage acidic protein 1 (CRAC1), or are involved in tissue remodeling, such as thrombospondin-4 and fibronectin (FINC) (Table 2). The relative abundance of these proteins in SF samples from OA patients strongly supports the hypothesis that enhanced ECM production and turnover by metalloproteases is a core constituent of the OA pathogenic process, leading to progressive erosion of structural integrity of the joint and eventual loss of functional performance. These differences between the RA and OA SF protein profiles are represented in Fig. 3.

3.4. RA synovial fluid proteomic signature

On the set of RA-specific proteins that we report (Table 1), we found S100A8 (MRP-8 or calgranulin-A). This protein is a known marker of chronic inflammation that was identified in most previous proteomic studies of RA markers in SF [28–30], thus being a “positive control” of the workflow that

we propose. Nevertheless, most of the proteins that we describe in this work have been identified for the first time in SF by proteomics. Given the fact that neutrophils are the predominant cell type in RA SF, being recruited in very large numbers to the rheumatoid cavity [38], many of the novel RA SF proteins that we identified are neutrophil-related proteins, such as azurocidin, defensin 2, neutrophil gelatinase-associated lipocalin or the leukocyte elastase inhibitor. Their identification will help to increase the understanding of RA pathophysiology, as neutrophil enzymatic activities are closely related with the persistence of inflammation, cartilage degradation and progression of joint damage [39]. We also detected for the first time in SF the presence of transforming growth factor-beta-induced protein ig-h3 (BGH3), an adhesion protein that is involved in cell–collagen binding (it binds to type VI collagen) and endochondral bone formation [40]. In a previous proteomics-based study performed by our group, we found an increased secretion of this protein from cartilage cells stimulated with the proinflammatory cytokine IL-1 β [41]. The present identification of BGH3 in SF supports its value as joint inflammatory biomarker. Finally, although the complement activation is not a prominent feature of RA [42], we detected two components of the complement pathway specifically in RA SF: C4 and C8 gamma chains, which highlight the contribution of this system to the RA pathological process. Whereas C4 gamma chain (CO4A) was previously characterized accumulated in synovial fluid of RA patients [43], this is the first time C8 is found.

3.5. OA synovial fluid proteomic signature

Regarding those proteins identified only in OA SF samples, we identified two proteins that have been previously linked to the osteoarthritic process, such as aggrecan core protein 2 and

Table 2 – Proteins identified as relatively more abundant in osteoarthritis (OA) synovial fluid.

| SwissProt ID | Protein name | Valid unique peptides/score RA ^a | Valid unique peptides/score OA ^a | Mean Ratio RA/OA ^b | SD ^c | Frequency RA/OA ^d |
|--------------|---|---|---|-------------------------------|-----------------|------------------------------|
| PGCA_HUMAN | Aggrecan core protein 2 | – | 18/29.7 | n.d./d* | – | 0/2 |
| COMP_HUMAN | Cartilage oligomeric matrix protein | – | 14/74.6 | n.d./d* | – | 0/3 |
| CFAD_HUMAN | Complement factor D | – | 6/35.9 | n.d./d* | – | 0/2 |
| TETN_HUMAN | Tetranectin | – | 5/30.3 | n.d./d* | – | 0/2 |
| ITIH1_HUMA | Inter-alpha-trypsin inhibitor heavy chain H1 | 19/77.0 | 24/129.0 | 0.8 | 0.52 | 4/4 |
| GELS_HUMAN | Gelsolin | 12/28.01 | 24/94.3 | 0.75 | 0.79 | 4/4 |
| IC1_HUMAN | Plasma protease C1 inhibitor | 54/85.8 | 76/172.8 | 0.74 | 0.71 | 4/4 |
| ITIH4_HUMAN | 35 kDa inter-alpha-trypsin inhibitor heavy chain H4 | 46/135.6 | 71/165.6 | 0.73 | 0.52 | 4/4 |
| A2AP_HUMAN | Alpha-2-antiplasmin | 21/59.8 | 32/90.7 | 0.65 | 0.35 | 4/4 |
| KNG1_HUMAN | Low molecular weight growth-promoting factor | 67/172.6 | 98/252.3 | 0.60 | 0.3 | 4/4 |
| CRAC1_HUMAN | Cartilage acidic protein 1 | 21/75.9 | 40/130.3 | 0.60 | 0.37 | 4/4 |
| FINC_HUMAN | Fibronectin | 22/114.2 | 42/222.9 | 0.5 | 0.16 | 4/4 |
| PEDF_HUMAN | Pigment epithelium-derived factor | 14/91.9 | 32/142.8 | 0.49 | 0.34 | 4/4 |
| HRG_HUMAN | Histidine-rich glycoprotein | 21/55.1 | 61/154.1 | 0.42 | 0.42 | 4/4 |
| THRB_HUMAN | Thrombin heavy chain | 34/100.2 | 57/161.9 | 0.31 | 0.07 | 4/4 |
| CH3L1_HUMAN | Chitinase-3-like protein 1 | 2/8.7 | 4/26.5 | 0.29 | 0 | 2/1 |
| A1BG_HUMAN | Alpha-1B-glycoprotein | 19/89.1 | 46/130.2 | 0.28 | 0.17 | 4/4 |

b) Relative abundance calculated by Spectral Counting using Phenyx 2.6 software. *Not detected/detected denotes proteins only detected in OA samples. Other headings, as in Table 1.

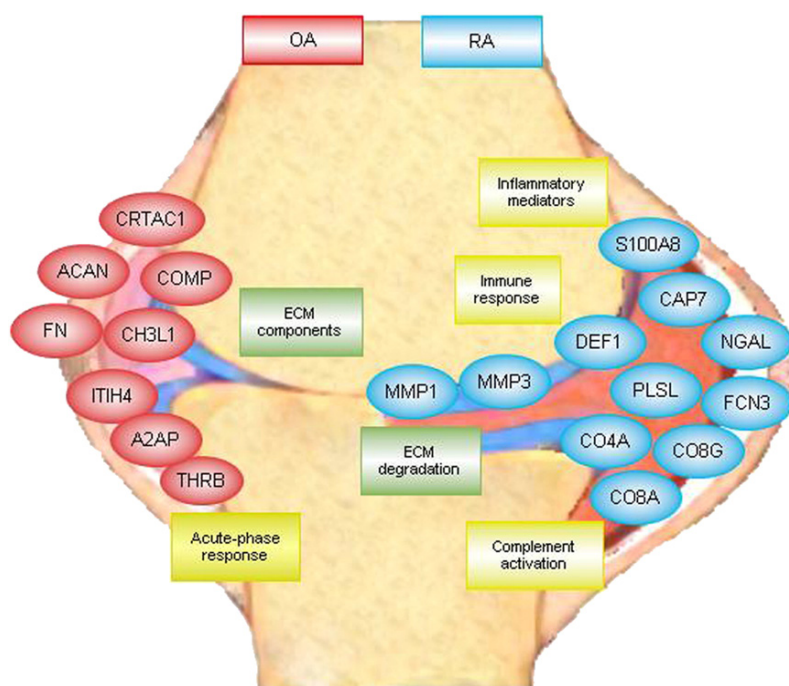


Fig. 3 – Differential protein profiling of SF from RA vs. OA patients. Representation of the proteins found in synovial fluids from knees with rheumatoid arthritis (RA) and osteoarthritis (OA). Protein abbreviations correspond to their SwissProt ID (see Tables 1 and 2). ECM: extracellular matrix.

cartilage oligomeric matrix protein (COMP). Aggrecan core protein 2 is a major structural component of the ECM, whose defects have been linked to spondylodysplasia and degenerative arthropathy [44], whereas COMP has been proposed to be

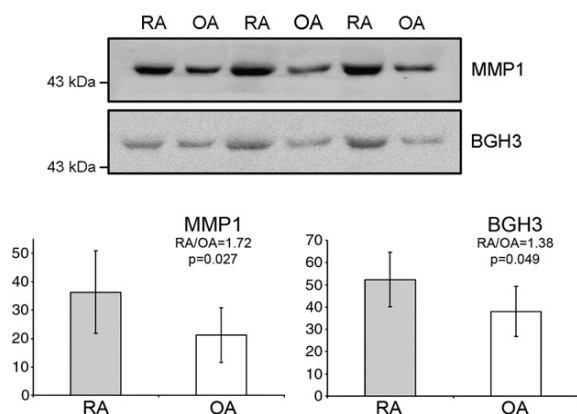


Fig. 4 – Accumulation of MMP-1 and BGH3 in SF from RA patients. Representative images of the Western-blot analyses performed against matrix metalloproteinase-1 (MMP-1) and transforming growth factor-beta-induced protein ig-h3 (BGH3) in 24 individual synovial fluid (SF) samples from knees of patients with rheumatoid arthritis (RA, n=12) and osteoarthritis (OA, n=12). Graphics represent the semi-quantitative values (in arbitrary units) obtained by densitometric analysis of the blots, together with the statistic significance of the results.

important for maintaining the structural integrity of cartilage, and its levels in serum have shown correlation with cartilage loss [45]. Interestingly, two other OA-specific SF proteins identified in this work (complement factor D and tetranectin) have been very recently proposed as novel serum OA biomarkers in a proteomic analysis carried out by our group on sera from OA and non-symptomatic controls [21]. This is the first time complement factor D is found in SF, which points to a role of the alternative pathway in OA pathogenesis. Finally, we report several other novel proteins as increased in OA (Table 2) that have been identified for the first time in SF. These include two inhibitors of angiogenesis (TSP4 and PEDF) and the chitinase-like protein 1 (CH3L1), a protein secreted at high amounts by cartilage cells that has been proposed as OA biomarker [46].

3.6. Verification of the increased presence of MMP1 and BGH3 in RA SF by immunoblotting on individual samples

We pursued to verify some of our proteomic data by an orthogonal strategy on the SF individual samples. Therefore, immunoblot analyses were performed using antibodies against two proteins belonging to the RA-characteristic SF protein group: matrix metalloprotease 1 (MMP-1). For the validation, we randomly chose 24 individual samples from each of the two pools that were employed in the proteomic analysis (n=12 for each condition).

Matrix metalloprotease 1 (MMP-1), also known as interstitial collagenase, is responsible for cleavage of collagen types I, II and III in ECM remodeling [47,48]. Thus, it plays a key role in both RA and OA pathogenic processes. This protein

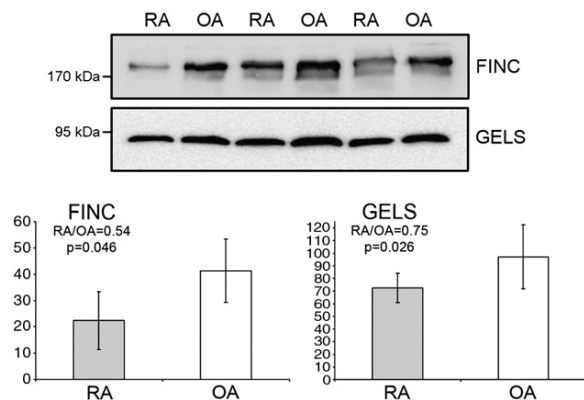


Fig. 5 – Increased abundance of fibronectin and gelsolin in SF from OA patients when compared to RA. Representative images of the Western-blot analyses of fibronectin and gelsolin in 16 individual synovial fluid (SF) samples from knees of patients with rheumatoid arthritis (RA, n=8) and osteoarthritis (OA, n=8). Graphics represent the semi-quantitative values (in arbitrary units) obtained by densitometric analysis of the blots, together with the statistic significance of the results.

was identified in the proteomic analysis only in the RA samples, $SD=0$ and $Frequency=3/0$ (Table 1). As shown in Fig. 4, Western-blot analysis showed the increased abundance of MMP-1 in RA SF than in OA SF, with a RA/OA ratio of 1.72 ($p=0.027$) calculated by densitometry analysis of the blots. With this antibody-based strategy we were able to detect MMP-1 in OA samples, as described previously [37]. Therefore, the value of the ratio RA/OA obtained by immunoblot would not be comparable to the proteomic data, in which the lack of MMP-1 identification in OA samples led to an arbitrary ratio of “9999” which is shown in Table 1 as “detected/non-detected”.

The other protein from the RA-specific group that we verified was transforming growth factor-beta ($TGF-\beta$)-induced protein ig-h3 (BGH3), an adhesion protein that belongs to the $TGF-\beta$ pathway that may play an important role in cell–collagen interactions through modulation of COL6 microfibril function [49,50]. Interestingly, we demonstrated in a previous study how the proinflammatory cytokine IL-1 β increased the secretion of this protein by cartilage cells [41]. As happened with MMP1, we could detect by immunoblotting BGH3 in both RA and OA SF samples, and we found an RA-specific increase of BGH3 in SF, with a ratio RA/OA=1.38 ($p=0.049$), that confirms the higher levels of this protein in RA that we noticed by the proteomic approach.

3.7. Verification of the increase of fibronectin and gelsolin in OA SF samples

We also aimed to validate the results obtained on two proteins from the OA-specific SF proteomic signature. Among the candidate proteins identified as increased in OA SF, we chose two proteins with a putative biomarker value reported previously: fibronectin (FINC), and gelsolin (GELS). FINC is an ECM protein essential for cell adhesion and motility [51] whose fragments appear in osteoarthritic cartilage [52], whereas gelsolin (GELS)

is an actin-modulating protein whose plasma isoform was recently reported as decreased in RA patients [53].

FINC relative quantification exhibited a low SD in the spectral counting analysis, with a ratio RA/OA=0.5, $SD=0.16$ and $Frequency=4/4$ (see Table 2). FINC forms a polymeric form, superfibronectin, and is responsible for cell–collagen interactions [54]. As shown in the representative immunoblotting pictures illustrated in Fig. 5, FINC (MW~250 kDa) is more abundant in OA SF samples compared to RA SF samples ($n=16$, 8 RA and 8 N). Densitometric analysis of the blots allowed the quantification of an RA/OA ratio of 0.54 ($p=0.046$, see Fig. 5). These data are in agreement with previous studies indicating the increase of FINC in OA, which suggest that its excess (or excess of its proteolytic fragments) promotes cartilage destruction through stimulation of matrix metalloproteases and aggrecanases [55].

On the other hand, GELS was found in the proteomic study with a ratio RA/OA=0.75, although with a very high SD (Table 2). Nevertheless, based on the previously reported putative value of GELS as RA biomarker, we tried to investigate its abundance in the synovial fluids by Western blotting. As shown in Fig. 5, GELS was decreased in RA samples compared to OA, with a ratio RA/OA of 0.75 and in this case exhibiting a reduced SD and high consistency of the values ($p=0.026$, $n=16$).

4. Conclusions

Protein profiling of complex biological samples has improved in the recent years, largely due to the development of the liquid chromatography coupled to mass spectrometry technique (LC–MS). In this study, we semi-quantitatively analyzed the protein content of knee SF, comparing samples from two highly prevalent pathological conditions affecting articular joints, RA and OA. We demonstrated that the LC–MALDI-TOF/TOF technique is suitable for the ongoing efforts to discover early indicators for RA and OA in SF as the source. Efforts to further enhance the sensitivity and reproducibility of the technique are necessary to improve the accuracy for detecting and quantifying the very low-abundance protein fractions, including growth factors, signaling molecules and cytokines.

For the first time, we have employed an immunodepletion column to remove in SF those top 20 most abundant plasma proteins, thus narrowing the high dynamic range of SF proteins. This strategy led to the largest number of proteins identified in SF up to date. Moreover, it allowed the characterization of two RA- and OA-proteomic signatures. The protein profiles identified in this work reflect the differences between the diseases: RA is a pathology characterized by an accumulation in SF of proteins related to the immune response, complement activation and inflammation, whereas OA SF exhibits an increased abundance of proteins involved in the formation and remodeling of the extracellular matrix, as a consequence of cartilage degradation.

Finally, most of the novel proteins identified in this work might have an essential value not only for increasing knowledge on the etiopathogenesis of both RA and OA processes. They also have a possible usefulness as disease biomarkers, as their

presence in SF is a prior step to their dilution in serum. A targeted approach, using either MRM-based or antibody-based strategies, would then be useful to evaluate their differential occurrence in plasma samples from RA patients, OA patients and controls.

Acknowledgments

We thank the ISCIII Networked Proteomics Platform (ProteoRed) and its members for support and helpful discussions, Nieves Domenech for antibody donations and Sara Relaño for technical help. This study was supported by grants from the Instituto de Salud Carlos III (CIBER BBN CB06-01-0040 and FIS-PI 08/2028), with participation of funds from FEDER (European Community); and from Xunta de Galicia (PGIDIT06PXIB916358PR, 10CSA916058PR). P. F.-P. and C.R.-R. are supported by ISCIII.

Appendix A. Supplementary data

Supplementary data to this article can be found online at [doi:10.1016/j.jprot.2011.12.042](https://doi.org/10.1016/j.jprot.2011.12.042).

REFERENCES

- Ruiz-Romero C, Calamia V, Carreira V, Mateos J, Fernandez P, Blanco FJ. Strategies to optimize two-dimensional gel electrophoresis analysis of the human joint proteome. *Talanta*. England: 2009Elsevier B.V; 2010. p. 1552–60.
- Kim HR, Kim KW, Jung HG, et al. Macrophage migration inhibitory factor enhances osteoclastogenesis through upregulation of RANKL expression from fibroblast-like synoviocytes in patients with rheumatoid arthritis. *Arthritis Res Ther* 2011;13:R43.
- Waarsing J, Kloppenburg M, Slagboom P, et al. Osteoarthritis susceptibility genes influence the association between hip morphology and osteoarthritis. *Arthritis Rheum* 2011;63(5): 1349–54.
- Dreier R. Hypertrophic differentiation of chondrocytes in osteoarthritis: the developmental aspect of degenerative joint disorders. *Arthritis Res Ther* 2010;12:216.
- Heinegård D, Saxne T. The role of the cartilage matrix in osteoarthritis. *Nat Rev Rheumatol* 2011;7:50–6.
- Ruiz-Romero C, Blanco FJ. Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis Cartilage* 2010;18:500–9.
- Schiess R, Wollscheid B, Aebersold R. Targeted proteomic strategy for clinical biomarker discovery. *Mol Oncol* 2009;3:33–44.
- Brewis IA, Brennan P. Proteomics technologies for the global identification and quantification of proteins. *Adv Protein Chem Struct Biol*, Vol. 80. Netherlands: Elsevier Inc; 2010. p. 1–44.
- Neilson KA, Ali NA, Muralidharan S, et al. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 2011;11:535–53.
- Hoehenwarter W, Wienkoop S. Spectral counting robust on high mass accuracy mass spectrometers. *Rapid Commun Mass Spectrom* 2010;24:3609–14.
- Zhang B, VerBerkmoes NC, Langston MA, Uberbacher E, Hettich RL, Samatova NF. Detecting differential and correlated protein expression in label-free shotgun proteomics. *J Proteome Res* 2006;5:2909–18.
- Ghaemmaghami S, Huh WK, Bower K, et al. Global analysis of protein expression in yeast. *Nature* 2003;425:737–41.
- Ruiz-Romero C, Lopez-Armada MJ, Blanco FJ. Proteomic characterization of human normal articular chondrocytes: a novel tool for the study of osteoarthritis and other rheumatic diseases. *Proteomics* 2005;5:3048–59.
- Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 2006;1:2856–60.
- Ahrne E, Masselot A, Binz PA, Muller M, Lisacek F. A simple workflow to increase MS2 identification rate by subsequent spectral library search. *Proteomics* 2009;9:1731–6.
- Colinge J, Chiappe D, Lagache S, Moniatte M, Bougueleret L. Differential proteomics via probabilistic peptide identification scores. *Anal Chem* 2005;77:596–606.
- Mi H, Dong Q, Muruganujan A, Gaudet P, Lewis S, Thomas PD. PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. *Nucleic Acids Res* 2010;38:D204–10.
- Nikitin A, Egorov S, Daraselina N, Mazo I. Pathway studio—the analysis and navigation of molecular networks. *Bioinformatics* 2003;19:2155–7.
- Chang X, Cui Y, Zong M, et al. Identification of proteins with increased expression in rheumatoid arthritis synovial tissues. *J Rheumatol* 2009;36:872–80 Canada.
- Yamagiwa H, Sarkar G, Charlesworth MC, McCormick DJ, Bolander ME. Two-dimensional gel electrophoresis of synovial fluid: method for detecting candidate protein markers for osteoarthritis. *J Orthop Sci* 2003;8:482–90.
- Fernandez-Puente P, Mateos J, Fernandez-Costa C, et al. Identification of a panel of novel serum osteoarthritis biomarkers. *J Proteome Res* 2011;10(11):5095–101.
- Fam H, Bryant JT, Kontopoulou M. Rheological properties of synovial fluids. *Biorheology* 2007;44:59–74.
- Ray S, Reddy PJ, Jain R, Gollapalli K, Moiyadi A, Srivastava S. Proteomic technologies for the identification of disease biomarkers in serum: advances and challenges ahead. *Proteomics Clin Appl* 2011;5:559.
- Echan LA, Tang HY, Ali-Khan N, Lee K, Speicher DW. Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* 2005;5:3292–303.
- Faca V, Pitteri SJ, Newcomb L, et al. Contribution of protein fractionation to depth of analysis of the serum and plasma proteomes. *J Proteome Res* 2007;6:3558–65.
- Gobezie R, Kho A, Krastins B, et al. High abundance synovial fluid proteome: distinct profiles in health and osteoarthritis. *Arthritis Res Ther* 2007;9:R36.
- Sinz A, Bantscheff M, Mikkat S, et al. Mass spectrometric proteome analyses of synovial fluids and plasmas from patients suffering from rheumatoid arthritis and comparison to reactive arthritis or osteoarthritis. *Electrophoresis* 2002;23: 3445–56.
- Drynda S, Ringel B, Kekow M, et al. Proteome analysis reveals disease-associated marker proteins to differentiate RA patients from other inflammatory joint diseases with the potential to monitor anti-TNFalpha therapy. *Pathol Res Pract* 2004;200:165–71.
- Uchida T, Fukawa A, Uchida M, Fujita K, Saito K. Application of a novel protein biochip technology for detection and identification of rheumatoid arthritis biomarkers in synovial fluid. *J Proteome Res* 2002;1:495–9.
- Baillet A, Trocme C, Berthier S, et al. Synovial fluid proteomic fingerprint: S100A8, S100A9 and S100A12 proteins discriminate rheumatoid arthritis from other inflammatory joint diseases. *Rheumatology (Oxford)* 2010;49:671–82 England.
- Kamphorst J, van der Heijden R, DeGroot J, et al. Profiling of endogenous peptides in human synovial fluid by

- NanoLC-MS: method validation and peptide identification. *J Proteome Res* 2007;6:4388–96.
- [32] Pan X, Huang L, Chen J, Dai Y, Chen X. Analysis of synovial fluid in knee joint of osteoarthritis: 5 proteome patterns of joint inflammation based on matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Int Orthop* 2012;36(1):57–64 Jan.
- [33] Zegels G, Van Raemdonck GA, Coen EP, Tjalma WA, Van Ostade XW. Comprehensive proteomic analysis of human cervical-vaginal fluid using colposcopy samples. *Proteome Sci* 2009;7:17.
- [34] Andrews G, Lewis D, Notey J, Kelly R, Muddiman D. Part II: defining and quantifying individual and co-cultured intracellular proteomes of two thermophilic microorganisms by GeLC-MS2 and spectral counting. *Anal Bioanal Chem* 2010;398:391–404.
- [35] Huang Y, Houston NL, Tovar-Mendez A, et al. A quantitative mass spectrometry-based approach for identifying protein kinase clients and quantifying kinase activity. *Anal Biochem* 2010;402:69–76.
- [36] Ishihama Y, Oda Y, Tabata T, et al. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 2005;4:1265–72.
- [37] Tchetcherikov I, Lohmander LS, Verzijl N, et al. MMP protein and activity levels in synovial fluid from patients with joint injury, inflammatory arthritis, and osteoarthritis. *Ann Rheum Dis* 2005;64:694–8 England.
- [38] Pillinger MH, Abramson SB. The neutrophil in rheumatoid arthritis. *Rheum Dis Clin North Am* 1995;21:691–714.
- [39] Wright HL, Moots RJ, Bucknall RC, Edwards SW. Neutrophil function in inflammation and inflammatory diseases. *Rheumatology (Oxford)* 2010;49:1618–31 England.
- [40] Munier FL, Frueh BE, Othenin-Girard P, et al. BIGH3 mutation spectrum in corneal dystrophies. *Invest Ophthalmol Vis Sci* 2002;43:949–54.
- [41] Calamia V, Rocha B, Mateos J, Fernandez-Puente P, Ruiz-Romero C, Blanco FJ. Metabolic labeling of chondrocytes for the quantitative analysis of the interleukin-1-beta-mediated modulation of their intracellular and extracellular proteomes. *J Proteome Res* 2011;10:3701–11.
- [42] Okroj M, Heinegård D, Holmdahl R, Blom AM. Rheumatoid arthritis and the complement system. *Ann Med* 2007;39: 517–30.
- [43] Swaak AJ, Van Rooyen A, Planten O, Han H, Hattink O, Hack E. An analysis of the levels of complement components in the synovial fluid in rheumatic diseases. *Clin Rheumatol* 1987;6: 350–7.
- [44] Gleghorn L, Ramesar R, Beighton P, Wallis G. A mutation in the variable repeat region of the aggrecan gene (AGC1) causes a form of spondyloepiphyseal dysplasia associated with severe, premature osteoarthritis. *Am J Hum Genet* 2005;77: 484–90.
- [45] Vilim V, Olejárová M, Macháček S, Gatterová J, Kraus VB, Pavelka K. Serum levels of cartilage oligomeric matrix protein (COMP) correlate with radiographic progression of knee osteoarthritis. *Osteoarthritis Cartilage* 2002;10:707–13.
- [46] Huang K, Wu LD. YKL-40: a potential biomarker for osteoarthritis. *J Int Med Res* 2009;37:18–24.
- [47] Desrochers PE, Jeffrey JJ, Weiss SJ. Interstitial collagenase (matrix metalloproteinase-1) expresses serpinase activity. *J Clin Invest* 1991;87:2258–65.
- [48] Rumbaugh J, Turchan-Cholewo J, Galey D, et al. Interaction of HIV Tat and matrix metalloproteinase in HIV neuropathogenesis: a new host defense mechanism. *FASEB J* 2006;20:1736–8.
- [49] Hanssen E, Reinboth B, Gibson MA. Covalent and non-covalent interactions of betaig-h3 with collagen VI. Beta ig-h3 is covalently attached to the amino-terminal region of collagen VI in tissue microfibrils. *J Biol Chem* 2003;278:24334–41.
- [50] Reinboth B, Thomas J, Hanssen E, Gibson MA. Beta ig-h3 interacts directly with biglycan and decorin, promotes collagen VI aggregation, and participates in ternary complexing with these macromolecules. *J Biol Chem* 2006;281:7816–24.
- [51] Di Cesare PE, Chen FS, Moergelin M, et al. Matrix–matrix interaction of cartilage oligomeric matrix protein and fibronectin. *Matrix Biol* 2002;21:461–70.
- [52] Zack MD, Amer EC, Anglin CP, Alston JT, Malfait AM, Tortorella MD. Identification of fibronectin neoepitopes present in human osteoarthritic cartilage. *Arthritis Rheum* 2006;54:2912–22.
- [53] Osborn TM, Verdrengh M, Stossel TP, Tarkowski A, Bokarewa M. Decreased levels of the gelsolin plasma isoform in patients with rheumatoid arthritis. *Arthritis Res Ther* 2008;10:R117.
- [54] Sasaki T, Brakebusch C, Engel J, Timpl R. Mac-2 binding protein is a cell-adhesive protein of the extracellular matrix which self-assembles into ring-like structures and binds beta1 integrins, collagens and fibronectin. *EMBO J* 1998;17:1606–13.
- [55] Sofat N. Analysing the role of endogenous matrix molecules in the development of osteoarthritis. *Int J Exp Pathol* 2009;90: 463–79.

Capítulo III: “Analysis of autoantibody profiles in osteoarthritis using comprehensive protein array concepts”

Aunque la artrosis (OA) no es una enfermedad autoinmune, los cambios que se producen durante la enfermedad pueden ser detectados específicamente por el sistema inmune, produciendo autoanticuerpos frente a antígenos específicos que se generan como consecuencia de la destrucción de la articulación. Estos autoanticuerpos podrían tener un potencial valor biomarcador en la OA.

En este trabajo, nos propusimos utilizar dos estrategias proteómicas diferentes basadas en *arrays* de proteínas con el fin de comparar el perfil de autoanticuerpos séricos de pacientes con artrosis (OA), pacientes con artritis reumatoide (AR) y controles sanos, y buscar así potenciales autoanticuerpos biomarcadores de la OA. Por un lado, se utilizaron los *Nucleic Acid Programmable Protein Array* (NAPPA) y por otra parte, se construyeron *arrays* de fragmentos proteicos en formato plano y en formato de esferas.

Los NAPPA son *arrays* de proteínas que nos permiten sintetizar proteínas completas funcionales a partir de un cDNA impreso en una plataforma químicamente modificada empleando un sistema de transcripción y traducción *in vitro*. En este trabajo se construyeron NAPPA expresando 80 proteínas distintas seleccionadas por su posible relevancia en la patología artrósica. Los *arrays* de fragmentos proteicos en formato plano se construyeron inmovilizando 3840 fragmentos antigénicos diseñados dentro del *Human Protein Atlas* (HPA) y seleccionados aleatoriamente. Estos antígenos representaban fragmentos correspondientes a 3207 proteínas diferentes.

En las dos plataformas, se hibridaron un total de 21 sueros de pacientes OA, 20 sueros de pacientes AR y 21 sueros de individuos sanos. En los dos casos, los autoanticuerpos fueron detectados usando un anticuerpo frente a la inmunoglobulina G humana marcado con fluorescencia.

Con los NAPPA se identificaron 4 autoanticuerpos frente a *Carbohydrate sulfotransferase 14* (CHST14), CD44, *Leptin* (LEP) y *Procollagen C-endopeptidase enhancer 1* (PCOLCE) cuyos niveles en suero distinguían significativamente ($p < 0.05$) entre pacientes OA y controles sanos. También identificamos 3 proteínas *Insulin-like growth factor-binding protein 6* (IGFBP6), *Insulin-like growth factor-binding protein 4* (IGFBP4) y LEP, cuyos niveles de reactividad en suero permitieron distinguir entre los pacientes OA y AR. También hemos observado que los niveles de reactividad frente a la *interleukin-6* (IL6) son mayores en pacientes AR que en controles.

Una de estas proteínas, la CHST14, fue seleccionada para una siguiente fase de validación por el método de ELISA utilizando las mismas muestras que en el NAPPA. Pudimos comprobar también por esta técnica que el nivel de reactividad frente a CHST14 en suero de pacientes OA es mayor que en controles sanos.

Tras un primer análisis con los *arrays* de fragmentos proteicos en formato plano, se seleccionaron 232 antígenos reactivos para una siguiente fase de verificación utilizando *arrays* de fragmentos proteicos en formato de esferas. A este panel de 232 antígenos, se añadieron 12 antígenos más correspondientes a fragmentos de las 232 proteínas seleccionadas para esta fase, 129 antígenos candidatos identificados en otras enfermedades y 88 antígenos que representaban fragmentos de las proteínas expresadas en el NAPPA. En total 373 antígenos se acoplaron a esferas magnéticas fluorescentes para verificar los resultados obtenidos con los *arrays* planos y determinar las diferencias entre los perfiles de autoanticuerpos de individuos OA, AR y controles sanos.

Con los *arrays* de fragmentos proteicos en formato plano y en formato esfera se identificaron 3 antígenos, *Homeobox protein Hox-D3* (HOXD3), *Torsin-1A-interacting protein 2* (IFRG15) y *Prolyl 3-hydroxylase 3* (LEPREL2) cuyos niveles de reactividad significativamente ($P.value < 0.05$) distinguían pacientes OA de controles sanos. HOXD3 y LEPREL2 sólo presentaron reactividad en pacientes OA y IFRG15 sólo en controles.

Un antígeno, *G protein-activated inward rectifier potassium channel 2* (KCNJ6) presentaba más reactividad en pacientes OA que en AR. Mientras que tres antígenos, *Coiled-coil domain-containing protein 85C* (CCDC85C), *Potassium voltage-gated channel subfamily B member 2* (KCNB2) y *Zinc finger protein 784* (ZNF787) presentaba menor reactividad en pacientes OA. Por otra parte, la comparación entre pacientes AR y controles reveló 6 antígenos, HOXD3, KCNB2, KCNJ6, IFRG15, *WW domain binding protein 1* (WBP1) y *Kuppler like-factor protein* (KLF2) cuya reactividad estaba diferencialmente modulada en pacientes AR y en controles. De estos 6 antígenos, 4 presentaban una mayor reactividad en AR (HOXD3, WBP1, KCNB2 y KLF2) y 2 antígenos (KCNJ6 y IFRG15) mostraban mayor reactividad en controles.

Los niveles de los autoanticuerpos identificados en los NAPPA y en los *arrays* de antígenos respectivamente, han permitido distinguir entre pacientes OA y controles, así como entre las distintas enfermedades reumáticas estudiadas (OA y AR), por lo que pueden tener potencial valor en la práctica clínica como biomarcadores séricos de la artrosis.

En conjunto, este trabajo demuestra la combinación de estos dos conceptos de *arrays* de proteínas para buscar perfiles de autoanticuerpos específicos de la artrosis.

Analysis of Autoantibody Profiles in Osteoarthritis Using Comprehensive Protein Array Concepts

Frauke Henjes,^{†,▽} Lucía Lourido,^{‡,▽} Cristina Ruiz-Romero,^{‡,§} Juan Fernández-Tajes,^{‡,#} Jochen M. Schwenk,[†] María Gonzalez-Gonzalez,[‡] Francisco J. Blanco,^{‡,||} Peter Nilsson,^{*,†} and Manuel Fuentes^{*,‡}

[†]Affinity Proteomics, SciLifeLab, School of Biotechnology, KTH - Royal Institute of Technology, Stockholm SE 171- 21, Sweden

[‡]Rheumatology Division, ProteoRed/ISCIII Proteomics Group, INIBIC – Hospital Universitario de A Coruña, A Coruña 15006, Spain

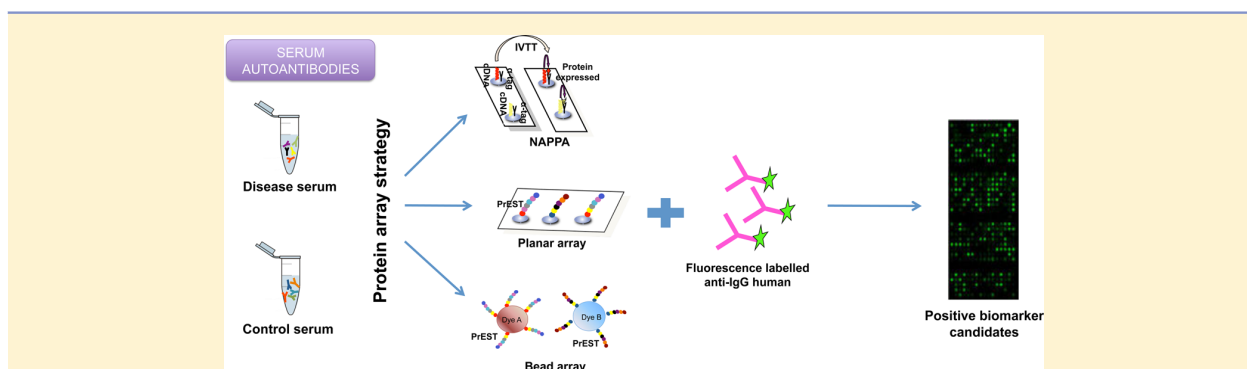
[§]CIBER-BBN Instituto de Salud Carlos III, INIBIC-CHUAC, A Coruña 15006, Spain

^{||}RIER-RED de Inflamación y Enfermedades Reumáticas, INIBIC-CHUAC, A Coruña 15006, Spain

[‡]Cancer Research Center (CIC), Department of Medicine, Proteomics Unit ProteoRed/ISCIII, IBSAL. University of Salamanca, Salamanca 37007, Spain

[#]Wellcome Trust Centre For Human Genetics, McCarthy's Group, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom

S Supporting Information



ABSTRACT: Osteoarthritis (OA) is the most common rheumatic disease and one of the most disabling pathologies worldwide. To date, the diagnostic methods of OA are very limited, and there are no available medications capable of halting its characteristic cartilage degeneration. Therefore, there is a significant interest in new biomarkers useful for the early diagnosis, prognosis, and therapeutic monitoring. In the recent years, protein microarrays have emerged as a powerful proteomic tool to search for new biomarkers. In this study, we have used two concepts for generating protein arrays, antigen microarrays, and NAPPAs (nucleic acid programmable protein arrays), to characterize differential autoantibody profiles in a set of 62 samples from OA, rheumatoid arthritis (RA), and healthy controls. An untargeted screen was performed on 3840 protein fragments spotted on planar antigen arrays, and 373 antigens were selected for validation on bead-based arrays. In the NAPPAs approach, a targeted screening was performed on 80 preselected proteins. The autoantibody targeting CHST14 was validated by ELISA in the same set of patients. Altogether, nine and seven disease related autoantibody target candidates were identified, and this work demonstrates a combination of these two array concepts for biomarker discovery and their usefulness for characterizing disease-specific autoantibody profiles.

KEYWORDS: Protein microarrays, autoantibody profiles, osteoarthritis, NAPPAs

INTRODUCTION

OA as Multifactor Disease

Osteoarthritis (OA) is the most common arthritic disease, and it is thereby among the most frequent symptomatic diseases for middle aged and elderly people. Worldwide estimates are that 10% of men and 18% of women over 60 years of age have symptomatic osteoarthritis. About 80% of those with OA have

significant limitations of movement, and 25% cannot perform their major activities of daily living.¹ Therefore, OA is already

Special Issue: Proteomics of Human Diseases: Pathogenesis, Diagnosis, Prognosis, and Treatment

Received: July 28, 2014

Published: September 16, 2014

one of the 10 most disabling diseases in developed countries, and it is becoming even more prevalent as the population ages and obesity rates rise. Between 2002 and 2007, OA moved from the 12th to the 6th leading cause of years lost to disability or morbidity.¹

Currently, OA diagnosis is mainly symptomatic, resting on the description of pain symptoms, stiffness of the affected joints, and radiography, which has been the reference technique to determine the grade of joint destruction. This is of key importance because radiography provides only indirect information about the joint tissue, and it has poor sensitivity to detect changes. Strategies followed at present are only symptomatic, focused on reducing pain, but are not able to arrest or reverse disease progression. In fact, the most successful therapy in advanced stages of OA is joint replacement with prosthesis. This disease manifests not only by cartilage degradation but also as an alteration of the whole joint structure, with progressive synovial inflammation, deterioration of menisci and ligaments, and changes on the subchondral bone. These are the reasons for the substantial interest in finding new specific biological markers of osteoarthritis that could facilitate not only early diagnosis of joint destruction and early interventions but also for disease prognosis or progression monitoring. Over the years, a series of biological markers have been proposed that may reflect the synthesis or degradation of the three main joint tissues (cartilage, synovial membrane and bone), and this list is continuously expanding.² However, despite the active research in this field, no single biomarker is sufficiently validated and qualified for its systematic use in OA diagnosis. Thus, no blood tests are yet available to make an early diagnosis of OA, predict how it will develop, or respond to therapy.

Although OA is not an autoimmune disease in terms of clinical manifestations, the changes involved in joint degeneration can be specifically detected by the immune system, leading to a humoral immune response producing immunoglobulins against these antigens (autoantibodies).³ Autoantibodies are stable circulating proteins, which are easily measurable in serum and may be detectable before clinical manifestations of the disease.⁴ Thus, serum autoantibody profiling can facilitate the discovery of OA diagnostic, prognostic, or progression biomarkers.

Until now, only very few proteins related with osteoarthritis have been purified and used in enzyme-linked immunosorbent assays (ELISA) and Western blot to detect immune response in serum from osteoarthritis patients.³ However, these conventional approaches are not applicable for multiplex- and high-throughput screens. In the last years, protein microarrays have emerged as a high-throughput strategy to provide a suitable tool for the identification of disease autoantibody signatures.

Array Platforms for Biomarker Discovery by Differential Serum Profiles

One application of affinity proteomics is to discover protein or autoantibody profiles in body fluids of patients. For the identification of new autoantibody target candidates, the use of high-throughput antigen microarrays offers an approach to screen simultaneously thousands of antigens in parallel in an unbiased manner and using a minimal amount of sample, and they have already been shown to be a suitable tool for autoantibody profiling in various disease entities.⁵ Those antigens can be either produced, purified, and subsequently immobilized to solid surfaces⁶ or directly expressed on-site.^{7,8} In this study, we applied two different concepts to complementarily analyze autoantibody signatures in patients in the context of OA.

The Human Protein Atlas Project (HPA) aims at generating antibodies against all human proteins and its representative immunogenic protein fragments.⁹ Within that project, fragments from protein coding genes covering the human proteome are selected based on sequences with low similarity to other proteins coding genes. Those fragments are routinely cloned, expressed, purified, and immobilized on microarrays to determine selectivity of antibodies, which are produced and affinity-purified with these protein fragments. So far more than 43 000 protein fragments (covering >16 000 human proteins) have been produced and arrayed. Each printed set of arrays is built of 384 different antigens from the routine antibody production workflow and are thereby unbiased with regard to protein selection. Therefore, these antigen arrays allow for untargeted autoantibody screens of patients samples.

The protein fragments contain mostly nonhomologous sequences to other proteins and are therefore specific for the target protein. However, autoantibodies can also be generated against conformational epitopes that might be lost in those short fragments due to missing sequences or incorrect folding. A complementing method is presented by the NAPPA technology, a new generation of self-assembled protein microarray where cDNAs, encoding human proteins with a tag, are spotted onto chemically modified surfaces. In the NAPPA, full-length proteins are produced in situ by using mammalian in vitro expression systems.¹⁰ The nascent protein is captured by immobilized antibodies specific for the tag encoded at the carboxy-terminus of the amino acid sequence, which ensures the full-length translation of the captured protein. Full-length proteins have a higher possibility to fold properly and thereby cover conformational epitopes. It has previously been shown that proteins expressed on the NAPPA arrays are biologically active when this technology was applied for the study of protein–protein interaction.¹⁰

A challenge in the discovery of new biomarkers is the high diversity and variety of patients, meaning numerous autoantibodies are found in single individuals only. Hence, there is a need to validate findings in larger sample cohorts, for example, by utilizing arrays that allow for the analysis of several hundreds of samples within a single assay. Therefore, we pursued a combinatory array strategy to screen for an OA-specific immune response in serum by combining two platforms, planar and bead arrays, as well as two concepts applying printed protein fragments and in situ expressed full-length proteins.

■ MATERIALS AND METHODS

Patients

The sera used in this work were obtained from peripheral blood of osteoarthritis (OA) and rheumatoid arthritis (RA) patients and from controls with no history of joint disease at the Rheumatology Division of the Hospital Universitario de A Coruña (CHUAC). Inclusion criteria for OA are fully described in a previous publication from our group,¹¹ whereas RA patients were diagnosed according to the American College of Rheumatology (ACR, 1988). The patients were of both genders and ages ranged from 41 to 99 years. The study was approved by the local Ethics Committee (Galicia, Spain). The distribution of gender, age, and diagnosis is depicted in Table 1. A total of 62 sera samples were screened in planar and bead arrays and 58 out of these in the NAPPA approach.

Table 1. Overview of Samples Used in the Multiarray-Screening Platform

| patient diagnosis | age range (mean \pm sd) | gender | |
|-----------------------------------|---------------------------|--------|------|
| | | F(%) | M(%) |
| rheumatoid arthritis ($n = 20$) | 46–75 (59 \pm 10) | 80% | 20% |
| osteoarthritis ($n = 21$) | 63–93 (74 \pm 9) | 62% | 38% |
| controls ($n = 21$) | 41–99 (67 \pm 16) | 48% | 52% |

Planar Antigen Arrays

In the discovery phase of the study, planar antigen arrays produced in the routine workflow of the Human Protein Atlas (HPA) were applied. Each slide consisted of 21 subarrays each containing 384 individual antigens. Immobilization of the antigens was performed through noncontact printing (Marathon, ArrayJet) onto epoxy slides (CapitalBio). A detailed protocol of the procedure was recently published by Ayoglu and colleagues.⁵ Ten print batches were from the HPA-workflow were randomly chosen. Thereby, a total of 3840 (3799 unique) antigens, denoted as protein epitope signature tags (PrESTs), covering 3307 Ensembl Gene IDs were analyzed. These PrESTs are 80–100 amino acids long and contain a N-terminal dual affinity tag (His₆-ABP) consisting of a hexahistidyl tag and albumin binding protein.

The screening procedure was performed as described previously with minor changes.⁵ Briefly, samples were diluted 1:250 (v/v) in assay buffer (5% nonfat milk powder, 3% BSA, 0.1% Tween-20 in PBS) containing 160 ng/mL His₆-ABP and preincubated for 1 h at RT on an orbital shaker. Samples (60 μ L/well) were transferred to the subarrays and incubated for 75 min. Slides were then washed with PBS-T (0.1% Tween-20 in PBS) and subsequently incubated with goat antihuman DyeLight 649 (Jackson) diluted 1:25 000 (v/v) in assay buffer for 75 min. Prior to scanning, slides were washed in deionized H₂O and dried by centrifugation. Scanning was performed using an array scanner (Agilent G2565BA) at a wavelength of 635 and 532 nm, and obtained images were analyzed using GenePix 5.0, extracting MFIs for data analysis.

Suspension Bead Arrays

For validation, a subset of antigens was coupled to color-coded magnetic beads (MagPlex, Luminex). Antigens were coupled as reported previously⁵ and described in more detail in the supplements. Briefly, 6 μ g antigens as well as 2.5 μ g antihuman IgG, 6 μ g His₆-ABP, and 4.5 μ g EBNA1 in protein-free buffer (2[N-Morpholino]-ethanesulfonic acid, pH 5.0) and blocking reagent for ELISA (BRE, Roche) as controls were coupled per 7.5×10^5 beads per identity. After coupling, all beads were combined into one suspension bead array, and coupling efficiency was evaluated using anti-His₆-ABP antibody detection with Alexa 555.

For analysis, samples were diluted 1:200 in assay buffer containing His₆-ABP and preincubated as described for planar arrays. Subsequently, 45 μ L was transferred to 5 μ L of beads and incubated for 2 h at RT on an orbital shaker. Beads were washed and incubated using a R-PE labeled antihuman IgG F(ab')₂ fragment for 30 min. After washing, fluorescence intensities were collected using the FlexMAP3D instrument (Luminex), and for data analysis, the median fluorescent intensities (MFIs) were used.

Self-Assembled (NAPPA) Protein Arrays

The list of OA biomarker candidates selected for targeted screening in the NAPPA platform was built on the basis of

previous works by our group^{12–15} and other unpublished results as well as database searches. The 80 cDNAs (encoding corresponding full-length human proteins) employed in the array are grouped by biological function in Supplementary Table S1 in the Supporting Information.

cDNA Preparation. NAPPA protein arrays were performed following the protocols of LaBaer's lab.¹⁰ *Escherichia coli* bearing a total of 80 sequence-verified full-length human genes in pANT7_cGST or pLDNT7_nFLAG were obtained from the Center for Personalized Diagnostics at the Arizona State University and are publicly available (www.dnasu.org). Bacteria were grown for 24 h at 37 °C in 100 mL of Luria–Bertani medium supplemented with ampicillin. The cells were pelleted and their plasmid DNA was purified using NucleoBond Xtra Midi EF/Max (Macherey-Nagel, Bethlehem, PA) according to the manufacturer's instruction.

Array Preparation. Samples of the previously purified cDNAs were precipitated by the addition of 0.8 \times volumes of isopropanol and centrifugation at 4000g for 30 min at RT. Precipitated cDNAs were then washed with 80% ethanol and allowed to air-dry. Fifteen μ g of each purified DNA was supplemented with 30 μ L of a Mastermix containing capture antibody (50 μ g/mL antiglutathione S-transferase (anti-GST) antibody (GE Healthcare Biosciences, Piscataway, NJ)) or anti-FLAG antibody (Sigma-Aldrich, St. Louis, MO), BSA (Sigma-Aldrich), and cross-linker (BS3; Pierce, Rockford, IL) and transferred to a 384-plate prior to printing onto the array surface. Aminosilane-functionalized microscope glass slides (VWR, Germany) were prepared according to the protocol described by González-González et al.¹⁶ The master mix plates and aminosilane functionalized slides were loaded to a MicroGrid II robotic microarray spotter (Isogen, The Netherlands) configured to use 48 pins that produced 150 μ m features.

In Situ Protein Expression. The in vitro transcription and translation step was performed as previously described by Manzano-Román et al. (2012).¹⁷ The transcription/translation lysate mix was added onto the slide. The arrays were incubated for 1.5 h at 30 °C for protein expression and captured by anti-GST polyclonal, respectively.

Analysis of Protein Display on NAPPA Arrays. The quality of DNA deposition was assessed by PicoGreen dye (Invitrogen, USA) and protein expression using monoclonal anti-GST antibody (Cell Signaling Technologies, Beverly, MA) diluted 1:100 (v/v).

Detection of Serum Autoantibodies. For autoantibody profiling, printed arrays were after expression incubated with serum diluted 1:600 (v/v) in the dilution buffer (5% milk with 0.2% Tween 20 in PBS) at 4 °C overnight with gentle mixing. Human antibodies reacting with expressed proteins on array were detected by HRP-linked antihuman IgG (Jackson ImmunoResearch Laboratories, West Grove, PA) diluted 1:1000 in Superblock Buffer. Then, slides were washed and incubated with tyramide signal amplification (TSA, Perkin Elmer) diluted 1:50 (v/v) in the dilution buffer supplied for the kit. Finally, the slides were rinsed with deionized water, dried using compressed air, and scanned using a ProScanArray HT scanner (PerkinElmer). The array images were analyzed and quantified using the GenePix Software version 6.0 (GenePix, USA).

ELISA Assay

A GST protein fused to the full-length CHST14 was expressed in vitro using the rabbit reticulocyte lysates, as previously described.

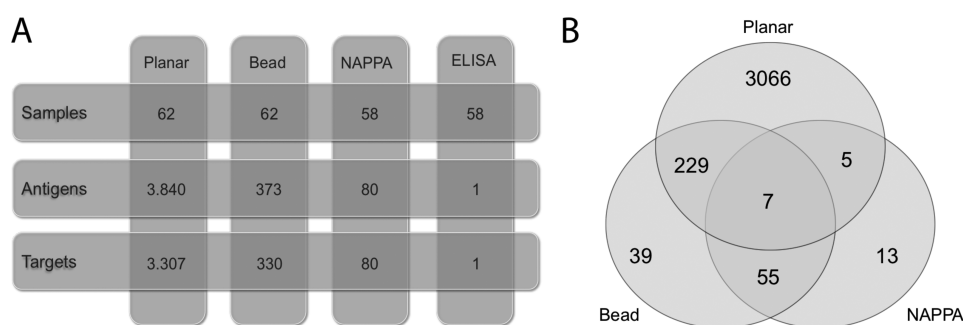


Figure 1. (A) Overview of the number of samples, antigens, and targets analyzed on the different platforms within this study. (B) Overlap between targets analyzed on both platforms.

The CHST14 protein was then applied to an anti-GST-coated 96-well plate. Plates were washed in PBS-0.02% Tween and blocked with PBS-Tween with 5% milk, overnight at 4 °C. After washing, the plate was incubated with the 1:300 (v/v) diluted sera used in the NAPPA screening.¹⁸

The presence of specific protein autoantibodies was detected by incubation with HRP-linked antihuman IgG secondary antibody (Jackson ImmunoResearch Laboratories, West Grove, PA) diluted 1:1000 (v/v) in assay buffer. Tetramethyl-benzidine substrate was then added and the reaction stopped. The 450 nm signals were read and the absorbances were subjected to *t*-test analyses to determine the significance of any variation in antibody levels between samples.

Statistical Analysis

All three applied methods differ with regard to material, production, binding chemistry, and incubation procedures resulting in different background and signal variation. Therefore, different criteria were applied to define reactivity.

The planar antigens arrays were only utilized as a first screening tool to identify reactive antigens, which were then transferred to bead arrays for further analysis. Therefore, more unbiased criteria were applied to identify antigens that show either reactivity in many samples or high signal intensities. First, fluorescence intensities were converted to a binary score defining reactivity according to a sample-specific threshold. For each print batch containing 384 antigens, signal intensities higher than the median plus either 20 or 250 times the median absolute deviation (MAD) were scored as reactivity. As additional selection criterion, antigens with a low cutoff of 20× MAD had to reveal reactivity in ≥ 10 samples to select only antigens that were reactive in at least 15% of the samples. Likewise, antigens showing signal intensities higher than 250× MAD above the sample median had to be reactive in ≥ 2 individual samples to select antigens with high signal intensities but to exclude all antigens that were reactive in individual samples only. As a third selection criterion, results from different MAD multiple values between 10 and 250 were combined, and Fisher's exact test was applied to identify significant ($p < 0.05$) differences in two group comparisons between the disease groups.

For the data generated on bead arrays containing the selected antigens, the sample median plus 10× MAD was applied as cutoff criterion to convert the data into binary information, and Fisher's exact test was applied to compare the count of reactive samples per disease group. For cluster analysis, all values below 0 meaning the signal intensities were below the median were set to 0.

For the data obtained from NAPPA, the total intensity (532 nm) for each spot generated on NAPPA arrays was converted to a normalized volume calculated by first removing the background signal estimated by the first quartile of the non-DNA spots (Mastermix with no cDNA). For interarray normalization, the previous normalized intensity was divided by the median of the excess intensity of the non-DNA spots in all arrays screened. As a first selection criterion, the median of the 12 replicates per antigen was selected as a reference of the reactivity of this antigen. As a second selection criterion, a cutoff >1 was used to determine reactivity in all samples. Then, the log 2 of the spot-normalized volume was used to compare the autoantibody abundance, and differential spot analysis was performed across the two samples by Wilcoxon signed-rank test (p value < 0.05).¹⁹

RESULTS

In this study, immunoreactivity profiles measuring the specificities of the IgG repertoire in serum of patients suffering from OA ($n = 21$), RA ($n = 20$), as well as healthy controls (HC; $n = 21$) were generated utilizing two antigen array concepts (Figure 1). The samples were applied in an untargeted screen on planar microarrays, followed by a technical validation applying a selected set of antigens using suspension beads arrays. In parallel, a targeted screening on 80 preselected full-length human recombinant proteins was carried out on the NAPPA platform.

Antigen Microarrays - Screening Phase

The set of samples was screened on 3840 antigens immobilized on planar microarrays. Out of these, 1738 and 301 were found reactive for low and high cutoff (median $+20\times/250\times$ MAD), respectively. The reactivity profiles observed in the initial screening showed a high variability between the individual samples, as depicted for two examples in Figure 2A. As a result of this heterogeneous profiles, most antigens, 51 and 80% for the two selected cutoff values, showed reactivity only in individual samples (Figure 2C). The total number of immune reactivities per sample was similar between the disease groups (Figure 2B), with a median number of reactivities per sample of 61, 60, and 60 (MAD: 19, 10.5, 9) for the low and 8, 5.5, and 6 (MAD: 4, 2.5, 2) for the high cutoff for OA, RA, and HC, respectively. However, one healthy control sample showed more than 300 immune reactivities (low cutoff) and was therefore excluded as an outlier from further analysis. Over all samples, no significant correlation ($r^2 = 0.08$; $p = 0.55/r^2 = 0.23$; $p = 0.07$ (low/high)) of the number of immune reactivities with age was observed (Supplementary Figure S2 in the Supporting Information).

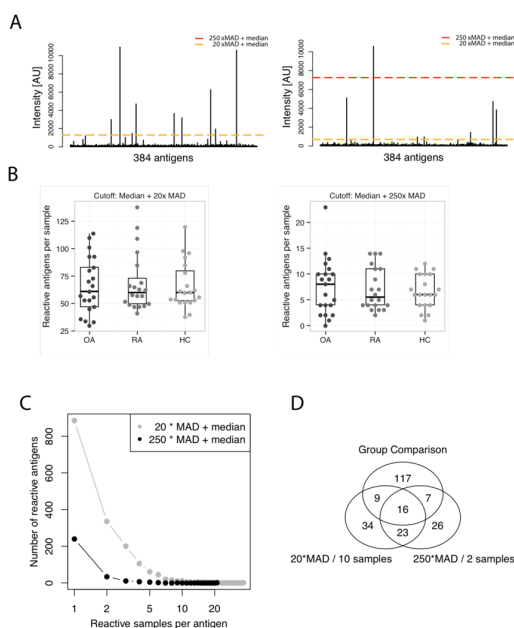


Figure 2. Summary of planar antigen array results. (A) Example of two different sample profiles on the same set of 384 antigens. The dashed lines represent the cutoff values of 20 \times MAD (orange) and 250 \times MAD (red). (B) Group comparison of the number of reactive antigens per sample showed a similar distribution of reactivities across all disease groups for the low (top) and the high (bottom) cutoff value. (C) Plot depicts the number of samples that show reactivity toward a certain antigen. Most antigens are reactive in individual samples only. (D) Summary of the 232 antigens selected from three different selection criteria.

For the selection of interesting antigens for further analysis, the high and low cutoff antigens had to be reactive in at least 2 or 10 samples, respectively. For the identification of disease-specific targets, different cutoff values between 10 and 250 \times MAD were applied and significant targets were identified applying Fisher's exact test. Combining all, low and high stringency as well as group comparisons, 232 antigens out of 3480 antigens were selected (Figure 2D).

Antigen Microarrays - Validation Phase

For the validation of the findings on another platform, the selected antigens were applied in bead-based assays. In addition to the 232 selected antigens, 12 antigens representing different regions of the selected target proteins were added. Furthermore, 129 antigens covering 92 targets from a list of other in-house identified candidates were included, from which 88 antigens covered 63 targets that were also analyzed using the NAPPA technology (Figure 1B).

After the screening on planar arrays, the 373 selected antigens were coupled to magnetic beads, and the 62 samples were reanalyzed using suspension bead arrays. Technical replication of the bead-based experiment showed a Pearson correlation coefficient of $r^2 = 0.99$, indicating high reproducibility of the data. The sample-specific background was more even on bead arrays compared with planar arrays (Supplementary Figure 3A in the Supporting Information). Therefore, a lower threshold compared to the planar arrays (median + 10 \times MAD) was applied to identify reactive antigens (Supplementary Figure 2B,C in the Supporting Information). While the primary aim in the screening phase was the selection of reactive

antigens for further validation, the goal in this phase was to determine disease-specific differences. Two group comparisons showed that most antigens had similar distributed reactivity ratios among all disease groups, as shown in Figure 3A. In all disease groups, most of the antigens were reactive in up to 40% of the samples per disease group. Comparison of the three disease groups revealed that the number of reactivities per patient was slightly higher in OA patients compared with both other groups, with a median number of 31 compared with 29 (Figure 3C).

To identify potential candidate antigens, group-wise comparison between OA, RA, and health controls was performed. This comparison revealed significant differences in nine antigens, namely, coiled-coil domain-containing protein 85C (CCDC85C), homeobox D3 (HOXD3), Torsin-1A-interacting protein 2 (IFRG15), potassium voltage-gated channel, Shab-related subfamily, member 2 (KCNB2), potassium inwardly rectifying channel, subfamily J, member 6 (KCNJ6), Kruppel-like factor 2 (KLF2), Leprecan-like protein 2 (LEPREL2), WW domain binding protein 1 (WBP1), and zinc finger protein 784 (ZNF784) (Table 2). The reactivity profile of these nine antigens with the percentage of reactive samples is summarized in Figure 4A.

Three antigens, namely, HOXD3, IFRG15, and LEPREL2, were significantly different between OA and healthy controls, whereby HOXD3 and LEPREL2 showed reactivity only in OA patients, while IFRG15 was reactive in controls (Figure 4B). One antigen, KCNJ6, was significantly higher in OA patients compared with RA patients, while three antigens, CCDC85C, KCNB2, and ZNF784, showed higher reactivity in RA compared with OA patients (Figure 4C). The comparison between RA and controls revealed six differential antigens. From those, four were higher in RA samples, namely, HOXD3, WBP1, KCNB2, and KLF2, and three antigens, namely, KCNJ6 and IFRG15, showed reactivity on more controls samples (Figure 4D). Out of these nine antigens, HOXD3, LEPREL2, and WBP1 showed no reactivity in the healthy controls (Supplementary Figure 4A in the Supporting Information).

Serum Screening by NAPPA Arrays

In the NAPPA approach, a total of 80 proteins were probed with sera from 20 OA, 20 RA, and 18 HC samples. The percentage of positive hits detected is represented in Figure 3D for all samples. The plot displays different profiles and a high variability between samples as well as a higher reactivity in the pathological (OA, RA) compared with HC samples. Additionally, different immune reactivities of the antigens according to the sample groups were observed. The percentage of positive hits for each antigen from the NAPPA array data was plotted and correlated to illustrate the differences of immune reactivity between the sample groups (Figure 3B). These results indicate a generally high immune reactivity for all antigens and different profiles between groups. The comparison between groups shows that the percentage of reactive samples for particular antigens is higher in OA as well as RA samples compared with HC.

Identification of Differential Serum Profiles

The statistical analysis of the autoantibody levels based on Wilcoxon rank-sum test revealed that immunoreactivity against seven full-length human recombinant proteins, CD44 antigen (CD44), carbohydrate sulfotransferase 14 (CHST14), leptin (LEP), procollagen C-endopeptidase enhancer 1 (PCOLCE), insulin-like growth factor-binding protein 4 (IGFBP4), insulin-like growth factor-binding protein 6 (IGFBP6), and interleukin-6

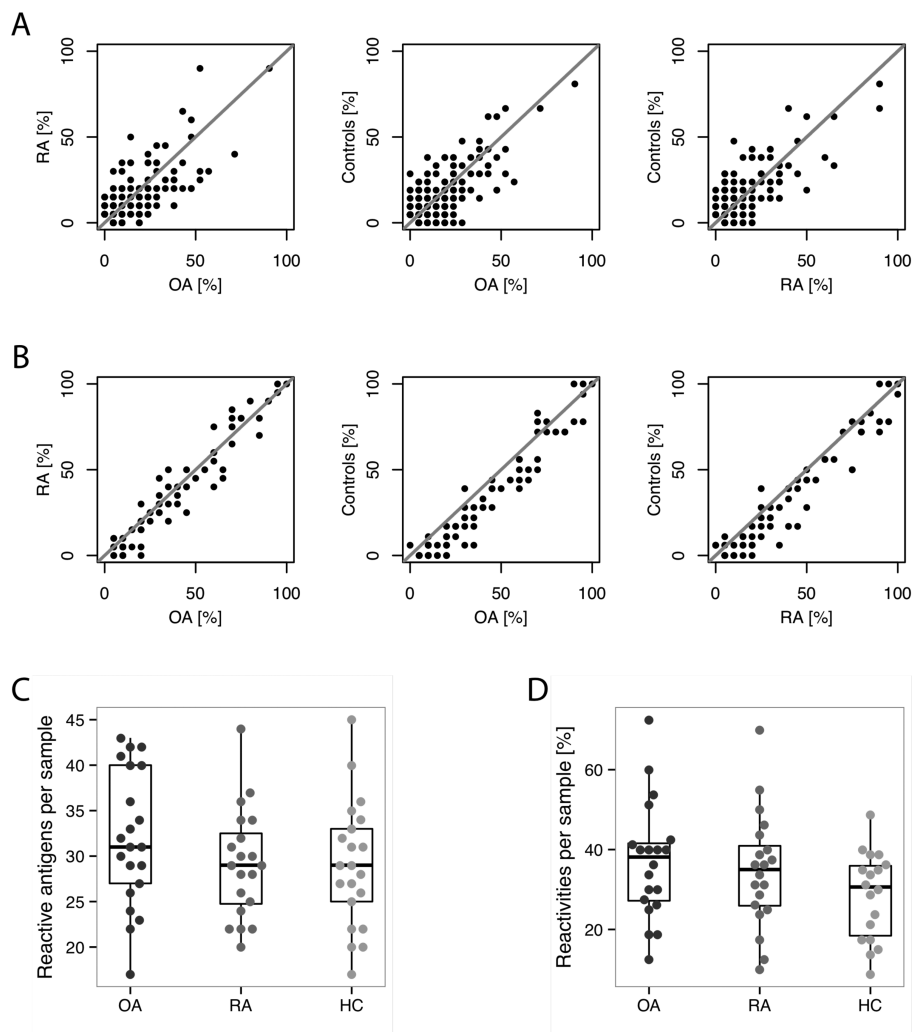


Figure 3. Summary of antigen bead arrays and NAPPA results. Group-wise comparison of the ratio of reactivities per antigen for antigen arrays (A) and NAPPA (B). Comparison of the number of reactive antigens per sample shown per disease group shown in absolute numbers for antigen arrays (C) and as percentage of all positive autoantibodies (D).

(IL6) showed significant differences ($p < 0.05$) between sample groups. An overview of the presence or absence for these autoantibodies in all patient samples is represented in Figure 5A. Hierarchically clustered heatmaps based on Euclidian distance for each pairwise comparison are shown in Figure 5B–D. The cluster analysis comparing OA and healthy control samples indicated that autoantibody levels against CHST14, CD44, LEP, and PCOLCE were significantly higher in OA compared with control samples (Figure 5B). The hierarchically clustered heatmap comparing RA and healthy control sera indicated that anti-IL6 autoantibody levels significantly distinguish these two groups of samples showing higher levels in RA compared with controls (Figure 5C). Finally, the comparison between the two groups of pathological samples (OA and RA) revealed that LEP, IGFBP6, and IGFBP4 antigens showed a higher immunoreactivity in OA compared with RA samples (Figure 5D).

The results observed for the CHST14 antigen in OA and HC samples used in the screening were validated using an ELISA assay. As shown in Figure 6A, the levels of this specific autoantibody were confirmed to be significantly higher in OA than in HC sera.

From the 88 protein fragments on the bead arrays that represented the 80 human recombinant proteins on NAPPA arrays, 34 showed no reactivity, while 3, namely, SPOCK2, SDC1, and CSAR1, were reactive in more than 40% of the samples. However, none of the antigens revealed significant differences in the reactivity between the groups on either of the platforms.

From the seven targets identified as candidates on NAPPA arrays, six were included on the antigen bead-array. Out of these six, four targets revealed reactivity, namely, IGFBP4, IGFBP6, CD44, and IL6, whereby the latter showed more reactivities in RA, while the others had more in OA. None of those antigens revealed reactivities in HC. Figure 7 shows the profiles of IL6 and CHST14 for the antigen arrays (upper panel) and NAPPA (lower panel). IL6 was represented by two different protein fragments with adjacent sequences. One of those showed reactivities in RA as well as OA samples but none in HC, while the other antigen was not reactive. The protein fragment representing CHST14 showed no reactivity on the antigen arrays.

Table 2. Antigen Information of Targets Screened in Both Platforms

| antigen | antigen sequence | gene | gene description | ENSG | Uniprot |
|--------------------------|--|---------|--|-----------------|---------|
| HPRR2510215 | PEAQRSSPTPESRSQEPDLVLVPPDDCRPGTPPS-DLIEIQVVKVTD'TLVPPEPPEPGSFHCALC-PAAFRLVSELLFHEHGHLA | ZNF784 | zinc finger protein 784 | ENSG00000179922 | Q8NCA9 |
| HPRR3700258 | KAAYENPGLFGGYGYSKTTDTYGYSTPHQ-PYPPPAASSLTDYDYPGSACSIQSSAPLRAPAHK-GADLN | HOXD3 | homeobox D3 | ENSG00000128652 | P31249 |
| HPRR3760489 | SFHETYETSTPPLSAKELAEAS-RAELPLSWSVSSKLNQHALETEEEEEKNLEEQ-TERNGDVANLENES | KCNJ6 | potassium inwardly rectifying channel, subfamily J, member 6 | ENSG00000157542 | P48051 |
| HPRR3920097 | GPGLHYAPPAPFAFLFDDA | KLF2 | Kruppel-like factor 2 (lung) | ENSG00000127528 | Q9Y5W3 |
| HPRR3790921 | DSGIELCPCPASGEGEPVKEVRSATLPDLE-DYSPCALPESVPQIFPMGLSSSEG | WBP1 | WW domain binding protein 1 | ENSG00000239779 | Q96G27 |
| HPRR3410375 | SDNSHCPCDGGQWFPSELGHWLYQTELVEHCYQVFLDRINRADYCEPCYPDNPANRSLV | IFRG15 | Torsin-1A-interacting protein 2, isoform IFRG15 | ENSG00000258664 | Q9H496 |
| HPRR4050663 | RLLGDKLLAQAGSGEFRTLKRGFSPYHSESQ-LASLPPSYQDSLQNGPACPAPPEL | CCDC85C | coiled-coil domain containing 85C | ENSG00000205476 | A6NKD9 |
| HPRR4030464 | SFENKYQEVSKQDSHEQLNNTSSSSPQHLSAQ-KLEMLYNEITKTQPHSHPNPDCQEKPERP-SAYEEEEIEEMEEVVCPEQ | KCNB2 | Potassium voltage-gated channel subfamily B member 2 | ENSG00000182674 | Q92953 |
| HPRR3460075 ^a | MEHLGTSFKDPDPWTPAALPEALREKLRED-QEKRPWDHEPVKPKPLTYWKDVLLEGVTLTQDSRQLNGSERAVLDGLLT-PAECGVLLQLAKDA | LEPREL2 | Leprecan-like protein 2 | ENSG00000110811 | Q8IVL6 |
| HPRR330267 | DDVSSGSSSERSSTSGGYIFYTFSTVHPIDEDSP-WITDSTDRIPATTLMSTSATATETATKR-QETWDFWFSWLFLPSESKNHLHTTTQ-MAGTSSNTISAGWEPNEENEDERDRHLSFSGS-GIDDEDFISSITSTTPR | CD44 | CD44 molecule (Indian blood group) | ENSG0000026508 | P16070 |
| HPRR4120127 | GILAEKPLPLHPPGREGTAWRGKAPKPGGLSL-RAGDADLQVRQDVRNRTL-RAVCGQPGMPRPDPWDLPV | CHST14 | carbohydrate sulfotransferase 14 | ENSG00000169105 | Q8NCHO |
| HPRR3140768 | EDARVPVQGSQSELHRLERLAASQSRTHED-LYIIPNCDNRNGNFHPKQCH-PALDGGQRGKCWCVDKRTGVKLPGLGLEPK-GELDCHQLADSFRE | IGFBP4 | insulin-like growth factor binding protein 4 | ENSG00000141753 | P22692 |
| HPRR620032 | PGLQCHPPKDDDEAPLRALLGRGRCLPARAPA-VAEENPKESKPQAG-TARPQDVNRRDQQRNPGTSTTPSQPN-SAGVQDTEMGPCRRHLDVSLQQLQTEVYR-GAQTLYVPNCDFHRG-FYRKRQCRRSQGQRRGPCWC | IGFBP6 | insulin-like growth factor binding protein 6 | ENSG00000167779 | P24592 |
| HPRR2760215 | KLQAQNWQLQDMTTHLLRSFKEFLQSSL-RALRQM | IL6 | interleukin 6 (interferon, beta 2) | ENSG00000136244 | P05231 |
| HPRR330007 | KDVAAPHRQPLTSSERIDKQIRYILDGIALR-KETCNKSNMCESSKEALAENNLNLPK-MAEKDGCFSQGFNEETCLVKIITGLLEFEVY-LEYLQNRFESEEQARAVQMSTKVLIQFLQK-KAKNLDAITTPDPTTNASLLTK | IL6 (2) | interleukin 6 (interferon, beta 2) | ENSG00000136244 | P05231 |
| HPRR2760330 | TLIKTIVTRINDISHTQSVSSKQKVTGLDFIPGLH-PILTLSKMDQTLAVYQQILTSRPSRVN | LEP | leptin | ENSG00000174697 | P41159 |

^aRemoved from last ensemble version.

In summary, nine and seven targets were identified as potential biomarker candidates on antigen arrays and NAPPA, respectively.

DISCUSSION

In this study, we aimed for the unbiased identification of autoantibody profiles in sera of OA patients to perform further studies for the progression or early detection of the disease before clinical onset. Therefore, two comprehensive protein array concepts were applied in an untargeted screening as well as a targeted approach to cover as many immunoreactive moieties as possible. The applied antigen arrays for untargeted discovery contained randomly selected protein fragments. Thereby, 62 patient sera were screened on a total of 3840 antigens representing 3307 proteins. Antigens showing reactivity were then applied on a bead-based array, which avoids the process of antigen printing which can lead to high variation between the subarrays.

The same sera samples were in parallel also screened using the NAPPA platform. The NAPPA approach avoids the delicate task and concerns associated with the expression of recombinant proteins in heterologous systems including protein purification and stability because proteins are expressed directly on-chip prior to sample incubation. Moreover, using the in vitro transcription/translation (IVTT) expression systems based on rabbit reticulocytes lysates, a broad variety of protein classes and sizes can be displayed in a similar range of protein levels in all spots along the array.¹⁰ This IVTT expression system includes chaperons and cofactors to allow the folding of mammalian proteins. The immunogenicity observed with NAPPA is most likely related to the epitope of folded proteins accessible for the autoantibody. Therefore, different folding in the expression may lead to the detection of different reactivities and includes the possibility to identify conformational epitopes.

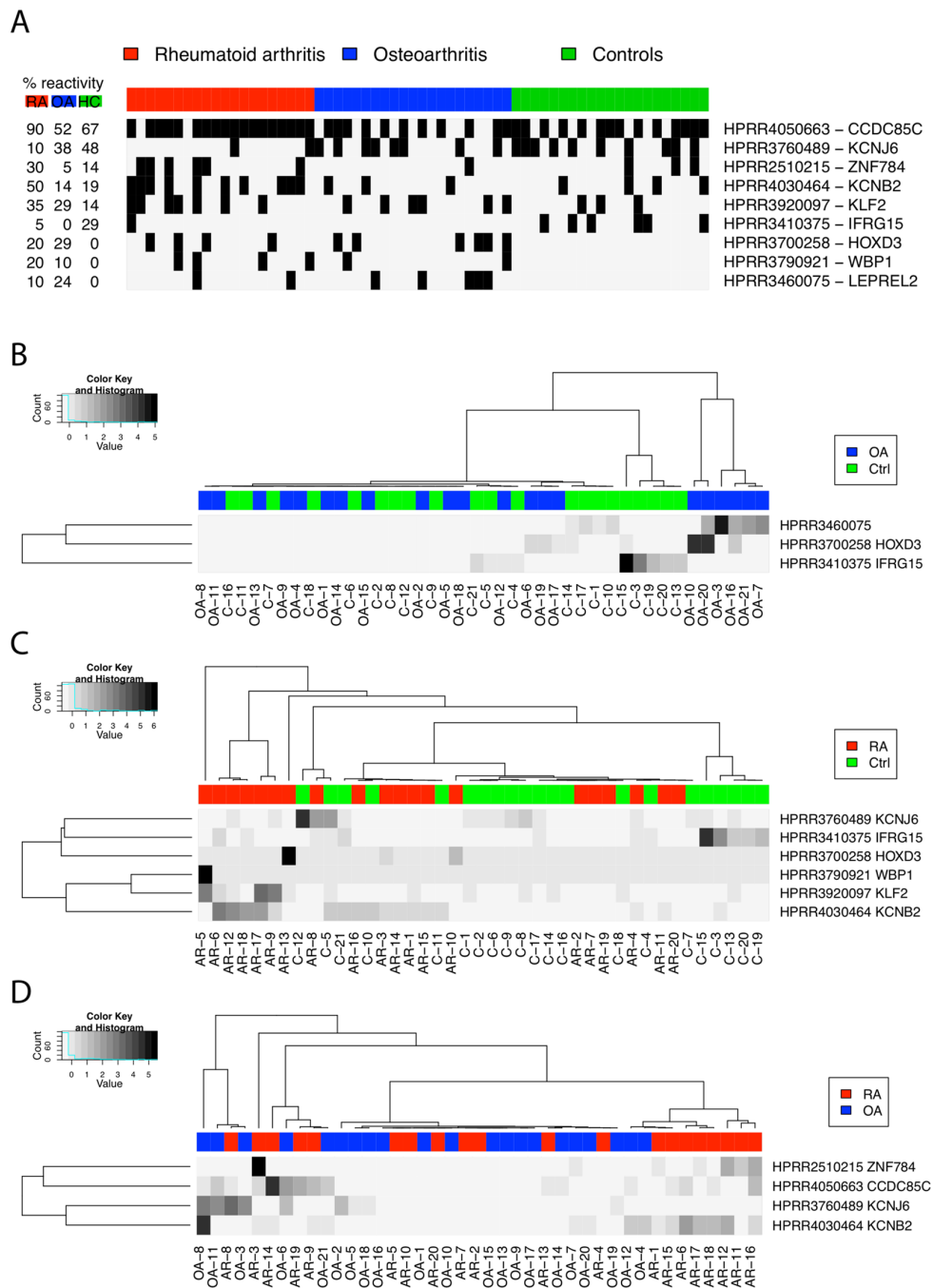


Figure 4. Antigen bead results. (A) Overview of the reactivity of each sample for the nine selected antigens, which showed differences between the three disease groups. Black squares show reactivity. Cluster analysis applying the MAD multiple above the median calculated for the differential antigens of OA samples versus Ctrl (B), RA versus Ctrl (C), and RA versus OA (D).

For both concepts, antigen bead arrays and NAPPA, IgG immunoreactivity was analyzed. 3 out of 373 and 4 out of 80 targets revealed differences between OA and controls on the antigen bead arrays and NAPPA, respectively. In contrast with RA, which is an autoimmune disease resulting in chronic systemic inflammation, OA is usually not regarded as an autoimmune disease but is rather seen as a disease being caused mainly by the aging process or joint overload, which leads a musculoskeletal disorders. However, the infiltration of lymphocytes,

macrophages, and monocytes in OA synovial fluid has been already described and well-characterized.^{20–22} Furthermore, recent reports suggest that extracellular matrix (ECM) proteins progressive release from articular cartilage may become endogenous catabolic factors during joint damage. Activation of pro-inflammatory pathways by ECM proteins has led to their description as damage-associated molecular patterns (DAMPs).²³ Thus, taken together, these data support the role of immunologic involvement in the pathogenesis of OA.

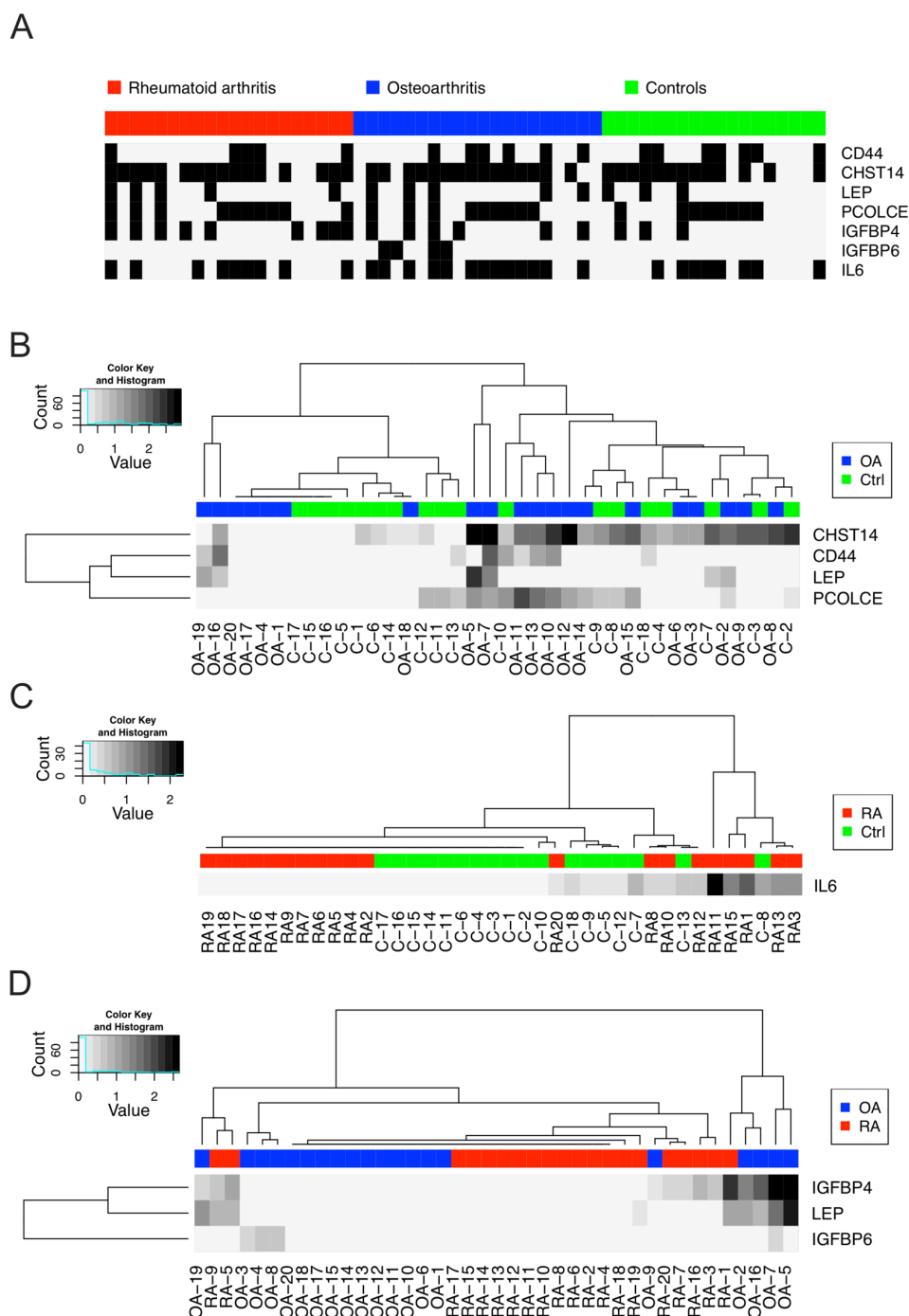


Figure 5. NAPPA results. (A) Overview of the reactivity of each sample for the selected antigens, which showed significant differences between groups. Black squares show their reactivity and white squares is no reactivity. (B–D) Cluster analysis applying Wilcoxon signed-rank tests showing significant (p value < 0.05) differences between osteoarthritis (OA) compared with control (Ctrl) samples (B); Rheumatoid arthritis (RA) compared with OA (C); and RA compared with Ctrl samples (D).

The overlap of reactivities observed on both antigen and NAPPA arrays was low, which can be explained by the intrinsic differences between these two concepts. From the analyzed proteins, 63 were represented on both concepts. While all targets except TIMP2 showed reactivity on NAPPA, no reactivities were observed for 41 out of 88 protein fragments representing those proteins. Examples of these divergences are

illustrated in Figure 7, which shows the reactivity observed in the antigen screening for two hits of the NAPPA screening (Figure 5): CHST14 (differential between OA patients and controls) and IL6 (differential in RA vs controls). As shown in Figure 7, no reactivity above the threshold $10 \times \text{MAD} + \text{median}$ was found for the antigen representing CHST14, while the presence of autoantibodies against this protein was detected

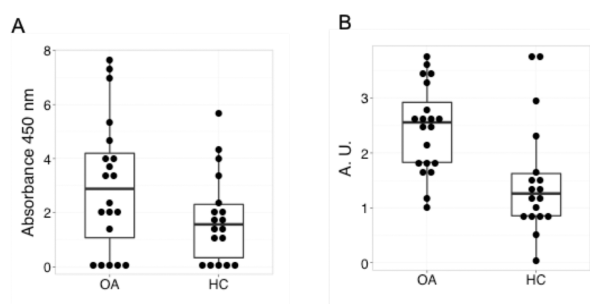


Figure 6. Reactivity levels observed for CHST14 antigen by NAPPA and ELISA. (A) The level of this specific autoantibody was confirmed by ELISA to be significantly higher (paired *t* test *p* value <0.05) in osteoarthritis (OA) than in healthy controls (Ctrl). (B) CHST14 reactivity detected for OA and Ctrl samples by the NAPPA approach. Black horizontal lines represent the mean value of CHST14 reactivity for each group of samples in both ELISA and NAPPA assays.

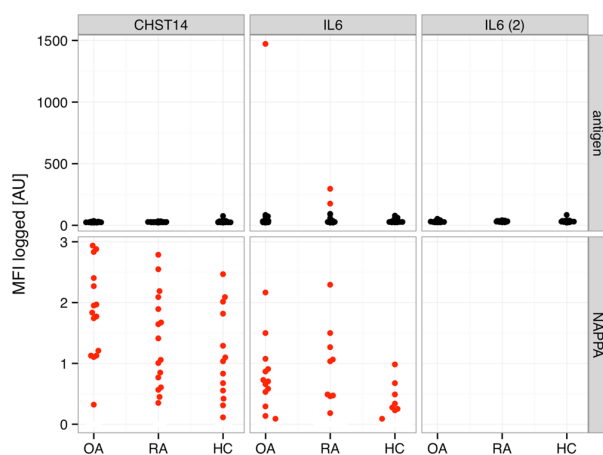


Figure 7. Antigen array – NAPPA comparison. Results for two selected targets identified on NAPPA (CHST14 and IL6) are shown for antigen arrays (upper panel) and NAPPA (lower panel). The reactive samples from NAPPA (15 out of 20 OA, 16 out of 20 RA, and 13 HC out of 18 for CHST14; 13 out of 20 OA, 9 out of 20 RA, and 8 out of 18 HC for IL6) are depicted in red; nonreactive samples are depicted in black. Two different antigens for IL6 were present on the antigen arrays (Table 2). One of them shows reactivity, and the reactive samples are depicted in red (1 OA and 2 RA samples).

by NAPPA and confirmed by ELISA (Figure 6). CHST14 is a carbohydrate sulfotransferase essential for the synthesis of dermatan sulfate,²⁴ a glycosaminoglycan that participates in the maintenance of the collagen fiber network in cartilage. More investigations will be highly valuable to validate these findings and gain knowledge on the mechanisms, underlying the presence of significantly increased amounts of CHST14 antibodies in OA patient's sera. The cytokine IL6 is key in the pathogenesis of RA²⁵ and a highly relevant therapeutic target for this disease.²⁶ Two different IL6 antigens with adjacent sequences were screened on the bead arrays, and one promoted immunoreactivity (Figure 7), thus showing different antigenic properties of the two epitopes. This example illustrates the advantages of the protein fragment screening to detect reactivities to specific and particular antigen sequences for the identification of putative disease-dependent antigenic specificities.

Although we have only screened a small set of samples, our results suggest the need to apply a combinatory screening strategy

to obtain a higher and better output for the identification of autoantibody targets.

Interestingly, from the antigens, which were identified in the unbiased screening approach, two out of nine were potassium-channels, KCNB2 and KCNJ6. Both were significantly different between OA and RA, whereby the latter was more abundant in OA. Recently, the potassium-channel KCNJ10 has been found as immune target in multiple sclerosis,²⁷ indicating a potential role of autoantibodies targeting potassium-channels in disease pathology. Potassium channels are involved in inflammatory diseases. For example, the voltage-gated potassium-channel KCNA3 has been reported to be upregulated in T-cells in the synovial fluid of RA patients,²⁸ as well as in inflammatory infiltrates of MS brains.²⁹ Other voltage-gated potassium-channels like KCNB2 or autoantibodies targeting those proteins might therefore also play a role in inflammatory processes and in the pathology of RA.

The 80–150 amino acid long protein fragments immobilized on the antigen arrays most likely present the whole spectrum from almost linear epitopes to folded secondary structures and thereby cover different possible recognition sites. However, because these fragments in most cases do not represent the full-length protein, only a subset of possible epitopes is represented, and physiological folding might not occur. These protein fragments are constructed for unique sequences of the target proteins, and thus no domain-specific but rather protein-specific autoantibodies ought to be detected. However, this may lead to false-negatives in case the fragment does not represent the autoantibody-epitope of interest. Unlike antigen analysis, full-length proteins are expressed and presumably properly folded in the NAPPA approach.^{10,17} Therefore, it is likely to obtain conformational epitopes, which are bound by the autoantibodies. However, conformational epitopes are not solely sequence-dependent, and thus domain-specific autoantibodies can be detected as well. Furthermore, the IVTT is an open system that allows studying the implications of possible modifications produced during joint damage. In future studies, one additional focus could be the implication of the proteins' post-transcriptional modifications in the generation of autoantibodies in OA.

CONCLUSIONS

Currently in early stages of biomarker discovery, the first studies are often based on a small cohort of samples, which are analyzed in one single array concept. However, bearing in mind our results, a comprehensive array approach (antigen microarrays and NAPPA arrays) could be considered as a strategy for early stages in biomarker discovery to identify specific serum fingerprints. In addition, it might identify the array concept suitable for further biomarker validation. In summary, this methodological approach presents a new strategy in biomarker discovery based on autoantibody detection. The combination of results obtained from different screening methods increases the number of identified potential targets and thereby the probability to validate a target as potential biomarker. The benefit of this combination has yet to be confirmed in further validation studies applying more samples, covering more pathologies, and possible additional array concepts.

ASSOCIATED CONTENT

Supporting Information

Supplementary Figure S1. NAPPA workflow. Supplementary Figure S2. Distribution of reactivities per sample generated on

planar antigen arrays plotted by age and gender. Supplementary Figure S3. Sample profiles of the same samples shown in Figure 1 on the set of 373 antigens on bead-arrays. Number of samples that show reactivity towards a certain antigen. Group comparison of the number of reactive antigens per sample, depicted for the cutoff median + 10×/50× MAD. Supplementary Figure S4. Summary of the percentage of reactivities per disease group: osteoarthritis (OA), rheumatoid arthritis (RA), and healthy controls (HC) of the selected antigens. Plots are shown for the selected targets from antigen array and NAPPA. Supplementary Table S1. List of the 80 Clones included in the NAPPA screening. Experimental details. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*P.N.: E-mail: peter.nilsson@scilifelab.se.

*M.F.: E-mail: mfuentes@usal.es.

Author Contributions

▽ F.H. and L.L. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support from Euro-TEAM (EU FP7), Knut and Alice Wallenberg Foundation, SciLifeLab Stockholm, the Carlos III Health Institute of Spain (FIS PI11/02114, PI11/02397, PI12/00329; CIBER-CB06/01/0040; RETIC-RIER-RD12/0009/0018), FEDER (European Community), Xunta de Galicia (Red Gallega REDICENT), and Junta Castilla-León SA198A12-2. The proteomics unit at CIC belongs to ProteoRed, PRB2-ISCIII, supported by grant PT13/0001.

REFERENCES

- Cross, M.; Smith, E.; Hoy, D.; Nolte, S.; Ackerman, I.; Fransen, M.; Bridgett, L.; Williams, S.; Guillemin, F.; Hill, C. L.; Laslett, L. L.; Jones, G.; Cicuttini, F.; Osborne, R.; Vos, T.; Buchbinder, R.; Woolf, A.; March, L. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann. Rheum. Dis.* **2014**, *73* (7), 1323–1330.
- Rousseau, J. C.; Delmas, P. D. Biological markers in osteoarthritis. *Nat. Clin. Pract. Rheumatol.* **2007**, *3* (6), 346–356.
- Du, H.; Masuko-Hongo, K.; Nakamura, H.; Xiang, Y.; Bao, C. D.; Wang, X. D.; Chen, S. L.; Nishioka, K.; Kato, T. The prevalence of autoantibodies against cartilage intermediate layer protein, YKL-39, osteopontin, and cyclic citrullinated peptide in patients with early-stage knee osteoarthritis: evidence of a variety of autoimmune processes. *Rheumatol. Int.* **2005**, *26* (1), 35–41.
- Dudas, S. P.; Chatterjee, M.; Tainsky, M. A. Usage of cancer associated autoantibodies in the detection of disease. *Cancer Biomarkers* **2010**, *6* (5–6), 257–270.
- Ayoglu, B.; Haggmark, A.; Khademi, M.; Olsson, T.; Uhlen, M.; Schwenk, J. M.; Nilsson, P. Autoantibody profiling in multiple sclerosis using arrays of human protein fragments. *Mol. Cell. Proteomics* **2013**, *12* (9), 2657–2672.
- MacBeath, G.; Schreiber, S. L. Printing proteins as microarrays for high-throughput function determination. *Science* **2000**, *289* (5485), 1760–1763.
- He, M.; Taussig, M. J. Single step generation of protein arrays from DNA by cell-free expression and in situ immobilisation (PISA method). *Nucleic Acids Res.* **2001**, *29* (15), E73–E73.
- Ramachandran, N.; Hainsworth, E.; Bhullar, B.; Eisenstein, S.; Rosen, B.; Lau, A. Y.; Walter, J. C.; LaBaer, J. Self-assembling protein microarrays. *Science* **2004**, *305* (5680), 86–90.

- Uhlen, M.; Oksvold, P.; Fagerberg, L.; Lundberg, E.; Jonasson, K.; Forsberg, M.; Zwahlen, M.; Kampf, C.; Wester, K.; Hober, S.; Wernerus, H.; Bjorling, L.; Ponten, F. Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* **2010**, *28* (12), 1248–1250.

- Ramachandran, N.; Raphael, J. V.; Hainsworth, E.; Demirkan, G.; Fuentes, M. G.; Rolfs, A.; Hu, Y.; LaBaer, J. Next-generation high-density self-assembling functional protein arrays. *Nat. Methods* **2008**, *5* (6), 535–538.

- Rego, I.; Fernandez-Moreno, M.; Fernandez-Lopez, C.; Gomez-Reino, J. J.; Gonzalez, A.; Arenas, J.; Blanco, F. J. Role of European mitochondrial DNA haplogroups in the prevalence of hip osteoarthritis in Galicia, Northern Spain. *Ann. Rheum. Dis.* **2010**, *69* (1), 210–213.

- Ruiz-Romero, C.; Carreira, V.; Rego, I.; Remeseiro, S.; López-Armada, M. J.; Blanco, F. J. Proteomic analysis of human osteoarthritic chondrocytes reveals protein changes in stress and glycolysis. *Proteomics* **2008**, *8* (3), 495–507.

- Ruiz-Romero, C.; Calamia, V.; Mateos, J.; Carreira, V.; Martinez-Gomariz, M.; Fernandez, M.; Blanco, F. J. Mitochondrial dysregulation of osteoarthritic human articular chondrocytes analyzed by proteomics: a decrease in mitochondrial superoxide dismutase points to a redox imbalance. *Mol. Cell. Proteomics* **2009**, *8* (1), 172–189.

- Fernández-Puente, P.; Mateos, J.; Fernández-Costa, C.; Oreiro, N.; Fernández-López, C.; Ruiz-Romero, C.; Blanco, F. J. Identification of a panel of novel serum osteoarthritis biomarkers. *J. Proteome Res.* **2011**, *10* (11), 5095–5101.

- Mateos, J.; Lourido, L.; Fernandez-Puente, P.; Calamia, V.; Fernandez-Lopez, C.; Oreiro, N.; Ruiz-Romero, C.; Blanco, F. J. Differential protein profiling of synovial fluid from rheumatoid arthritis and osteoarthritis patients using LC-MALDI TOF/TOF. *J. Proteomics* **2012**, *75* (10), 2869–2878.

- González-González, M.; Bartolome, R.; Jara-Acevedo, R.; Casado-Vela, J.; Dasilva, N.; Matarraz, S.; García, J.; Alcazar, J. A.; Sayagues, J. M.; Orfao, A.; Fuentes, M. Evaluation of homo- and hetero-functionally activated glass surfaces for optimized antibody arrays. *Anal. Biochem.* **2014**, *450*, 37–45.

- Manzano-Román, R.; Díaz-Martín, V.; González-González, M.; Matarraz, S.; Álvarez-Prado, A. F.; LaBaer, J.; Orfao, A.; Pérez-Sánchez, R.; Fuentes, M. Self-assembled protein arrays from an *Ornithodoros moubata* salivary gland expression library. *J. Proteome Res.* **2012**, *11* (12), 5972–5982.

- Wong, J.; Sibani, S.; Lokko, N. N.; LaBaer, J.; Anderson, K. S. Rapid detection of antibodies in sera using multiplexed self-assembling bead arrays. *J. Immunol. Methods* **2009**, *350* (1–2), 171–182.

- Anderson, K. S.; Ramachandran, N.; Wong, J.; Raphael, J. V.; Hainsworth, E.; Demirkan, G.; Cramer, D.; Aronzon, D.; Hodi, F. S.; Harris, L.; Logvinenko, T.; LaBaer, J. Application of protein microarrays for multiplexed detection of antibodies to tumor antigens in breast cancer. *J. Proteome Res.* **2008**, *7* (4), 1490–1499.

- Sakkas, L. I.; Johanson, N. A.; Scanzello, C. R.; Platsoucas, C. D. Interleukin-12 is expressed by infiltrating macrophages and synovial lining cells in rheumatoid arthritis and osteoarthritis. *Cell Immunol.* **1998**, *188* (2), 105–110.

- Jasin, H. E. Autoantibody specificities of immune complexes sequestered in articular cartilage of patients with rheumatoid arthritis and osteoarthritis. *Arthritis Rheum.* **1985**, *28* (3), 241–248.

- Revell, P. A.; Mayston, V.; Lalor, P.; Mapp, P. The synovial membrane in osteoarthritis: a histological study including the characterisation of the cellular infiltrate present in inflammatory osteoarthritis using monoclonal antibodies. *Ann. Rheum. Dis.* **1988**, *47* (4), 300–307.

- Foell, D.; Wittkowski, H.; Roth, J. Mechanisms of disease: a 'DAMP' view of inflammatory arthritis. *Nat. Clin. Pract. Rheumatol.* **2007**, *3* (7), 382–390.

- Evers, M. R.; Xia, G.; Kang, H. G.; Schachner, M.; Baenziger, J. U. Molecular cloning and characterization of a dermatan-specific N-acetylgalactosamine 4-O-sulfotransferase. *J. Biol. Chem.* **2001**, *276* (39), 36344–36353.

(25) Boissier, M. C.; Semerano, L.; Challal, S.; Saidenberg-Kermanac'h, N.; Falgarone, G. Rheumatoid arthritis: from autoimmunity to synovitis and joint destruction. *J. Autoimmun.* **2012**, *39* (3), 222–228.

(26) Semerano, L.; Thiolat, A.; Minichiello, E.; Clavel, G.; Bessis, N.; Boissier, M. C. Targeting IL-6 for the treatment of rheumatoid arthritis: Phase II investigational drugs. *Expert Opin. Invest. Drugs* **2014**, *23* (7), 979–999.

(27) Srivastava, R.; Aslam, M.; Kalluri, S. R.; Schirmer, L.; Buck, D.; Tackenberg, B.; Rothhammer, V.; Chan, A.; Gold, R.; Berthele, A.; Bennett, J. L.; Korn, T.; Hemmer, B. Potassium channel KIR4.1 as an immune target in multiple sclerosis. *New Engl. J. Med.* **2012**, *367* (2), 115–123.

(28) Beeton, C.; Wulff, H.; Standifer, N. E.; Azam, P.; Mullen, K. M.; Pennington, M. W.; Kolski-Andreaco, A.; Wei, E.; Grino, A.; Counts, D. R.; Wang, P. H.; LeeHealey, C. J.; B, S. A.; Sankaranarayanan, A.; Homerick, D.; Roeck, W. W.; Tehranzadeh, J.; Stanhope, K. L.; Zimin, P.; Havel, P. J.; Griffey, S.; Knaus, H. G.; Nepom, G. T.; Gutman, G. A.; Calabresi, P. A.; Chandy, K. G. Kv1.3 channels are a therapeutic target for T cell-mediated autoimmune diseases. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (46), 17414–17419.

(29) Rus, H.; Pardo, C. A.; Hu, L.; Darrach, E.; Cudrici, C.; Niculescu, T.; Niculescu, F.; Mullen, K. M.; Allie, R.; Guo, L.; Wulff, H.; Beeton, C.; Judge, S. I.; Kerr, D. A.; Knaus, H. G.; Chandy, K. G.; Calabresi, P. A. The voltage-gated potassium channel Kv1.3 is highly expressed on inflammatory infiltrates in multiple sclerosis brain. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (31), 11094–11099.

Capítulo IV: *“Discovery of circulating proteins associated to knee radiographic osteoarthritis”*

Los *arrays* de anticuerpos son una potente herramienta para analizar proteínas concretas en muestras complejas como el suero. Por otra parte, los *arrays* de anticuerpos en formato de esferas o *arrays* en suspensión permiten analizar una gran cantidad de proteínas simultáneamente en un gran número de individuos con un bajo consumo de muestra.

Así, en este trabajo nos propusimos por primera vez utilizar *arrays* de anticuerpos en formato de esfera con el objetivo de encontrar un perfil proteico en el suero, que permita discriminar entre pacientes con artrosis radiográfica de rodilla, controles sanos y pacientes con AR y artritis psoriásica (APs).

Se analizaron un total de 78 proteínas propuestas en base a trabajos previos de nuestro grupo relacionados con la OA y se incluyeron también otras proteínas recogidas en la bibliografía relacionada con las enfermedades reumáticas. Los anticuerpos se seleccionaron según su disponibilidad dentro del *Human Protein Atlas* (HPA) y se intentó, en todos los casos posibles, seleccionar al menos dos anticuerpos por proteína diana analizada. Para crear el *array* se acoplaron cantidades iguales de anticuerpo a cada población de esferas fluorescentes. Cuando se comprobó la correcta unión de todos los anticuerpos a las esferas, éstas se mezclaron para crear el *array* de esferas en suspensión.

Para este estudio, se utilizaron dos cohortes de muestras de suero independientes, una para el cribado inicial y otra para la verificación de los resultados. La cohorte de cribado estaba formada por 288 pacientes OA con grado Kellgren-Lawrence (K/L) entre 2 y 4, 96 controles sanos así como 288 pacientes AR y 288 APS. La cohorte de verificación estaba formada por 196 pacientes OA con grado K/L entre 2 y 4, 92 controles sanos así como 192 pacientes AR y 192 APS. En todas las muestras, el contenido proteico de todos los sueros se marcó directamente con biotina. A continuación, se desnaturalizaron las proteínas con calor y las muestras se hibridaron directamente en el *array*.

En el cribado se analizaron un total de 174 anticuerpos dirigidos frente a 78 proteínas diferentes. Después de este primer análisis, se detectaron 12 proteínas significativamente moduladas ($p < 0.05$) entre pacientes OA con respecto a controles; 40 proteínas diferencialmente expresadas entre pacientes OA y AR; y 40 proteínas cuya abundancia en el suero distinguía a los pacientes OA y APS. En total, estas proteínas

representaban 46 proteínas diferentes que se seleccionaron para una siguiente fase de verificación.

Para la fase de verificación se creó un *array* formado por 106 anticuerpos dirigidos frente a estas 46 proteínas. Los resultados indicaron que, de las 12 proteínas que distinguían entre pacientes OA y controles, 6 se confirmaban en la cohorte de verificación. De las 40 proteínas que distinguían pacientes OA y AR, 32 se modulaban también en la cohorte de verificación; y de las 40 proteínas que diferenciaban pacientes OA y APS, 19 fueron también confirmadas en la segunda cohorte de muestras analizadas.

Así, el análisis de cribado y verificación mostraba un panel de 6 proteínas formado por S100A6, Leptin (LEP), *Complement 3 (C3)*, *Inter-alpha-trypsin inhibitor heavy chain H1 (ITIH1)*, *Vitamin D-binding protein (GC)* y *Apolipoprotein A-I (APOA1)* que permitían distinguir los sueros de los pacientes OA de los sueros procedentes de individuos sanos. Este panel mostró un *Area under the curve (AUC)* superior a 70 para distinguir los pacientes OA de los controles sanos en las dos cohortes analizadas, demostrando el potencial del panel de proteínas encontrado para el diagnóstico de la artrosis. En concreto, los niveles de S100A6, LEP, C3 e ITIH1 estaban aumentados en pacientes OA con respecto a controles sanos mientras que los niveles de GC y APOA1, estaban disminuidos. De estas 6 proteínas, sólo S100A6 permitía significativamente ($p < 0.05$) distinguir todos los grados K/L de los controles en las dos cohortes de muestras analizadas, sugiriendo el potencial de esta proteína para el diagnóstico y pronóstico de la artrosis de rodilla.

Además, encontramos 16 proteínas cuyos niveles en el suero permitían distinguir las 3 enfermedades reumáticas. De estas 16 proteínas, 8 estaban aumentadas en pacientes OA: *Apolipoprotein A-IV (APOA4)*, *Plasma protease C1 inhibitor (SERPING1)*, C3, *Extracellular matrix protein 1 (ECM1)*, *Insulin-like growth factor-binding protein 6 (IGFBP6)*, *SPARC-like protein 1 (SPARCL1)*, *Matrix Gla protein (MGP)*. De estas 8 proteínas, solo la C3 permitía distinguir también los pacientes OA y los controles sanos. Estos resultados sugieren la utilidad y el potencial de C3 para el diagnóstico específico de la artrosis.

Capítulo IV: *“Discovery of circulating proteins associated to knee radiographic osteoarthritis”*

INTRODUCTION

Osteoarthritis (OA) is the most common rheumatic disease of the developed world and it is increasingly important in current ageing populations, leading to patient chronic disability [198]. This disease manifests not only by cartilage degradation but also as an alteration of the whole joint structure, with progressive synovial inflammation and changes on the subchondral bone [7].

Currently, OA diagnosis is mainly symptomatic, resting on the description of pain symptoms, stiffness of the affected joints and evaluation of cartilage radiography or magnetic resonance imaging (MRI) [92]. However, the sensitivity of radiography is not adequate for detecting small changes, thus when radiographic diagnosis is established, significant joint damage has often already occurred [94-96]. In contrast, MRI is a quite sensible technique and it has been developed for the evaluation of cartilage damage in OA, but it is very expensive and requires a large instrumentation time, which reduces its applicability [94, 98]. Moreover, OA has little efficient therapeutics, probably as a consequence of the lack of early diagnosis strategies and techniques for its precise monitoring.

In the last years, biochemical biomarkers have emerged as promising tools, with more sensitivity and reliability than plain radiography, to detect joint changes that occur in OA [199]. Such markers of osteoarthritis could facilitate early diagnosis of joint destruction, disease prognosis and progression monitoring, which could be detectable with an early biochemical test [101]. Over the years, a series of biological markers have been proposed that may reflect the synthesis or degradation of the joint tissues. Despite the active research in this field, currently no single marker is sufficiently validated for its use in OA diagnosis [108-110]. This is mainly due to the lack of validation studies in large populations, which makes the findings questionable to be considered as robust biomarkers for OA [200].

In the present study, 1632 serum samples from groups of OA patients, healthy control subjects and disease control samples from patients with rheumatoid arthritis (RA) and psoriatic arthritis (PsA) were analysed using a high-throughput affinity proteomic approach based on antibody suspension bead arrays, with the potential to screen hundreds of proteins in hundreds of body fluid samples in parallel [201]. Here,

we aimed to identify a panel of serum proteins able to discriminate knee radiographic OA patients from controls.

MATERIALS AND METHODS

Patients and controls

All the individuals analysed in this study were attended in Hospital Universitario de A Coruña. The 484 OA participants were classified according to the American College of Rheumatology (ACR) criteria, and knee radiographies were classified using the Kellgren-Lawrence (K/L) score. 180 individuals were classified as controls considering the following inclusion criteria: not autoimmune disease and non-radiographic knee OA. Additionally, a total of 480 RA and 480 PsA fulfilling the ACR and CASPAR criteria, respectively, were included in the study.

The clinical data of the patients is summarized in Table 1.

Table 1. Clinical characteristics of the patients and control subjects used in the study.

| Screening set (N=960) | Ctrl | OA=288 | | | | |
|-------------------------------------|-------|--------|------|-------|-------|-------|
| | | K/L2 | K/L3 | K/L4 | RA | PsA |
| Number of patients | 96 | 118 | 121 | 58 | 288 | 288 |
| Female(%) | 50 | 75 | 78 | 86 | 74 | 40 |
| Body mass index(kg/m ²) | 29±6 | 31±5 | 32±6 | 32±5 | - | - |
| Age | 63±10 | 67±8 | 69±9 | 70±10 | 55±15 | 55±13 |
| Verification set (N=672) | Ctrl | OA=196 | | | | |
| | | K/L2 | K/L3 | K/L4 | RA | PsA |
| Number of patients | 92 | 43 | 31 | 112 | 192 | 192 |
| Female(%) | 57 | 72 | 77 | 74 | 73 | 42 |
| Body mass index(kg/m ²) | 24±8 | 31±5 | 33±1 | 32±6 | - | - |
| Age | 56±21 | 69±10 | 71±8 | 73±7 | 59±12 | 55±14 |

Samples

Blood samples from all patients and controls were collected after overnight fast in plain tubes containing a separation gel. They were allowed to stand for 20 min and then centrifuged at 2800 rpm for 10 min. The serum was aliquoted and stored at -80°C until use.

Antibody selection and bead array generation

Protein targets proposed to generate the bead array were selected based on thorough mining of experimental evidence in the literature of rheumatic diseases and in previous in-house efforts in the field of osteoarthritis using mass spectrometry analyses [202-204]. The antibody set employed in the experiments was finalized according to their availability within the Human Protein Atlas (HPA) [164]. A total of 174 protein microarray validated antibodies [170] targeting 78 unique proteins were included, selecting at least one antibody for each target protein (Supplementary Table 1).

The bead arrays were created as previously described [205] by diluting 1.6 μg of each antibody into 100 μL of dilution antibody buffer. All antibodies were immobilized onto color-coded magnetic beads (MagPlex-C, Luminex Corp.) with one bead identity corresponding to a certain antibody. The coupling of each antibody on the beads was confirmed via R-phycoerythrin-conjugated donkey antirabbit-IgG antibody (Jackson ImmunoResearch).

Serum profiling

The procedure for serum profiling was performed as described previously [206]. Briefly, 3 μL of each sample were diluted 1:10 in PBS, and randomized in 96-well plates. Then, the protein content was directly labelled with biotin (Life Technologies). Samples were further diluted 1:50 in an assay buffer, heat treated at 56°C for 30 min, combined into a 384-well microtiter plate, and incubated with the bead array at RT on a shaker overnight. Unbound proteins were removed by washing and proteins on the beads were detected through a streptavidin-conjugated fluorophore (Invitrogen.Com). Results from the FlexMap3D instrument (Luminex Corp.Com) were reported per bead identity as median fluorescence intensities (MFI).

Serum study setup

An overview of the study design is illustrated in Figure 1. A set of 960 samples (denoted as screening set) containing 288 OA patients, 96 control subjects, 288 RA and 288 PsA patients were first analyzed (screening phase) using the panel of 174 antibodies targeting 78 proteins immobilized on bead arrays (Supplementary Table 1). In total, 46 different proteins showed altered levels in comparisons across the sample groups (Supplementary Figure 1). Then, a smaller panel comprising 107 antibodies targeting these 46 proteins (Supplementary Table 2) was used to verify the protein profiles identified in the screening phase in a new set of 672 samples (denoted as

verification set), which was composed of 196 OA, 92 controls as well as 192 RA and 192 PsA individuals.

Statistical analysis

Data were processed and visualized in R. MFI values were normalized in each 384-plate by probabilistic quotient normalization (PQN) as accounting for any potential sample dilution effects. In addition, potential batch effects were adjusted using the ComBat function included in the 'sva' R Package. The outliers that were identified by robust principal component analysis (rPCA, R package "rrcov") were excluded from further analysis.

The technical variation was assessed by calculating the coefficient of variation (C.V.), which was lower than 20 % based on replicates of pooled samples distributed across all the plates (Supplementary Figure 2).

For biological interpretation, a linear regression analysis adjusting for sex, age and body mass index (BMI) was applied to normalized data in order to identify differences in protein profiles between the compared groups. The nonparametric Wilcoxon rank-sum test was applied to normalized data to discriminate between controls and the different OA grades. Differences were denoted statistically significant if they revealed unadjusted $p < 0.05$ in screening and verification phases.

We employed logistic regression to compare the classification power of the significant proteins concordantly distinguishing OA and controls in the two independent sample cohorts analyzed, where leave one-out method was selected as cross-validation option. The ROC curves were generated using the R package "pROC" for screening and verification sets of samples.

RESULTS

Initial screening phase

In the screening phase we analysed a cohort composed of 288 OA, 96 Controls, 288 RA and 288 PsA subjects using an array with 174 different antibodies. We found 12 proteins significantly ($p < 0.05$) differentiating OA patients and controls (Figure 1 and Table 2), 40 proteins whose levels were modulated between OA and RA patients (Figure 1 and Supplementary Table 3) and 40 proteins distinguishing OA and PsA subjects (Figure 1 and Supplementary Table 4). Altogether, the results of this screening phase narrowed the list of candidates to 107 antibodies against 46 proteins. A total of

106 antibodies targeting these 46 proteins were used to evaluate them in a further verification analysis.

Verification phase

Following the screening, a verification study was performed to profile the panel of 46 proteins in an independent cohort of serum samples from 196 OA patients and 92 control subjects, as well as 192 RA and 192 PsA patients. In accordance with the findings observed in the screening, 6 proteins out of 14 (Figure 1, Table 2), were found to display levels allowing to significantly ($p < 0.05$) distinguish OA and control individuals. Additionally, 32 proteins from the 40 were verified to discriminate between OA and RA patients (Figure 1, Supplementary Table 3), whereas 19 from the 40 protein candidates could discriminate between OA and PsA individuals (Figure 1, Supplementary Table 4).

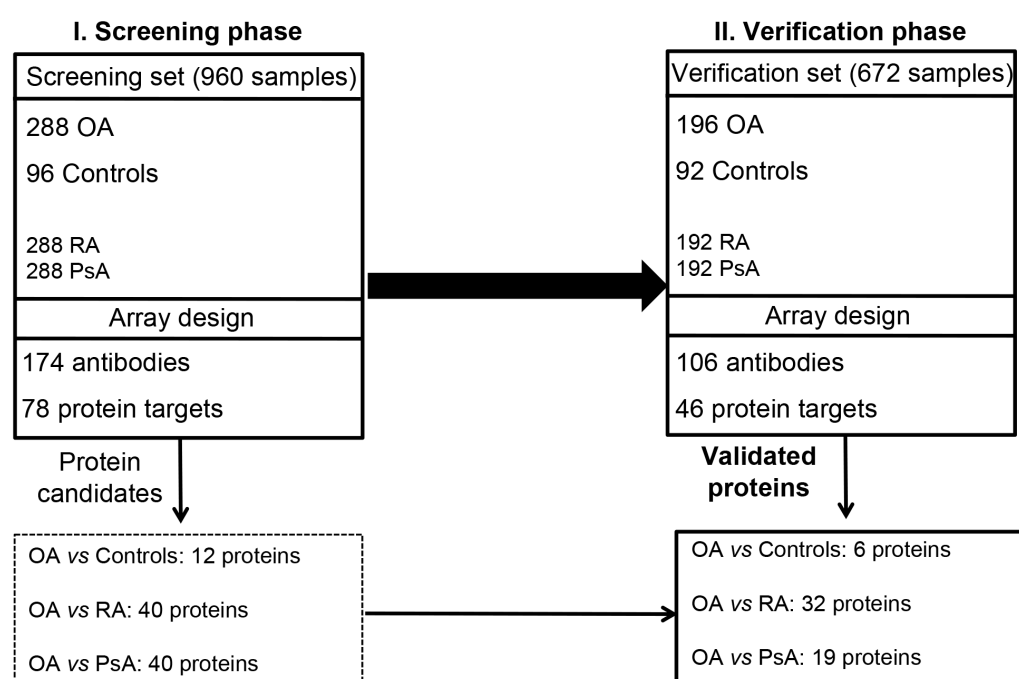


Figure 1. Schematic overview of the study. A screening phase was performed using a set of 960 serum samples. After this first step, we found a total of 46 different biomarker candidates significantly ($p < 0.05$) separating osteoarthritis (OA) patients and controls individuals as well as OA and rheumatoid arthritis (RA) and psoriatic arthritis (PsA) patients. Then we built another array targeting these 46 protein candidates and a Verification phase was finally performed using the 46 multiplex antibody array to validate 6 out of the 12 biomarker candidates separating OA and Controls, 32 from the 40 separating OA and RA patients and 19 from the 40 protein candidates to discriminate OA and PsA individuals.

Table 2. List of antibodies for which the protein profiles revealed differences with $p < 0.0$ (light grey) or $p < 0.01$ (grey) for the comparison OA versus control serum samples. The modulation of these six common proteins found both in screening and the verification is indicated with increased (\uparrow) or decreased levels (\downarrow) in serum OA individuals for the certain protein.

| Protein name | Antibody | OA versus Control | | |
|--------------|-----------|-------------------|-----------------|-----------------|
| | | Modulation | Screening | Verification |
| C3 | HPA020432 | \uparrow OA | 1.45E-04 | 5.36E-05 |
| S100A6 | HPA007575 | \uparrow OA | 5.95E-04 | 3.36E-05 |
| ITIH1 | HPA041639 | \uparrow OA | 2.66E-02 | 8.90E-05 |
| KLF2 | HPA055964 | \uparrow OA | 3.80E-02 | |
| LEP | HPA068565 | \uparrow OA | 2.12E-02 | 2.71E-02 |
| | HPA057322 | \uparrow OA | 3.03E-02 | |
| | HPA030721 | \uparrow OA | 4.13E-02 | |
| APOA1 | HPA046715 | \downarrow OA | 2.91E-02 | 3.15E-08 |
| GC | HPA019855 | \downarrow OA | 3.31E-02 | 8.32E-04 |
| LMNA | HPA006660 | \downarrow OA | 4.70E-04 | |
| SERPINC1 | HPA001816 | \downarrow OA | 2.90E-03 | |
| BGLAP | HPA046415 | \downarrow OA | 1.18E-02 | |
| OGN | HPA067445 | \downarrow OA | 1.50E-02 | |
| PLTP | HPA047207 | \downarrow OA | 1.63E-02 | |

Protein profiles for OA diagnosis

The screening and verification phases concordantly allowed the detection of six antibodies targeting six different proteins revealing significant differences ($p < 0.05$) between radiographic knee OA patients and Control individuals (Table 2). These six antibodies were generated towards S100 calcium binding protein A6 regulator (S100A6), vitamin D-binding protein (GC), inter-alpha trypsin inhibitor heavy chain (ITIH1), apolipoprotein A-1 (APOA1), leptin (LEP) and complement 3 (C3).

Four of the six mentioned proteins, S100A6, ITIH1, LEP and C3 showed higher serum levels in OA patients when compared to the control group. In contrast, the protein levels of GC and APOA1 were decreased in OA compared to control individuals (Figure 2).

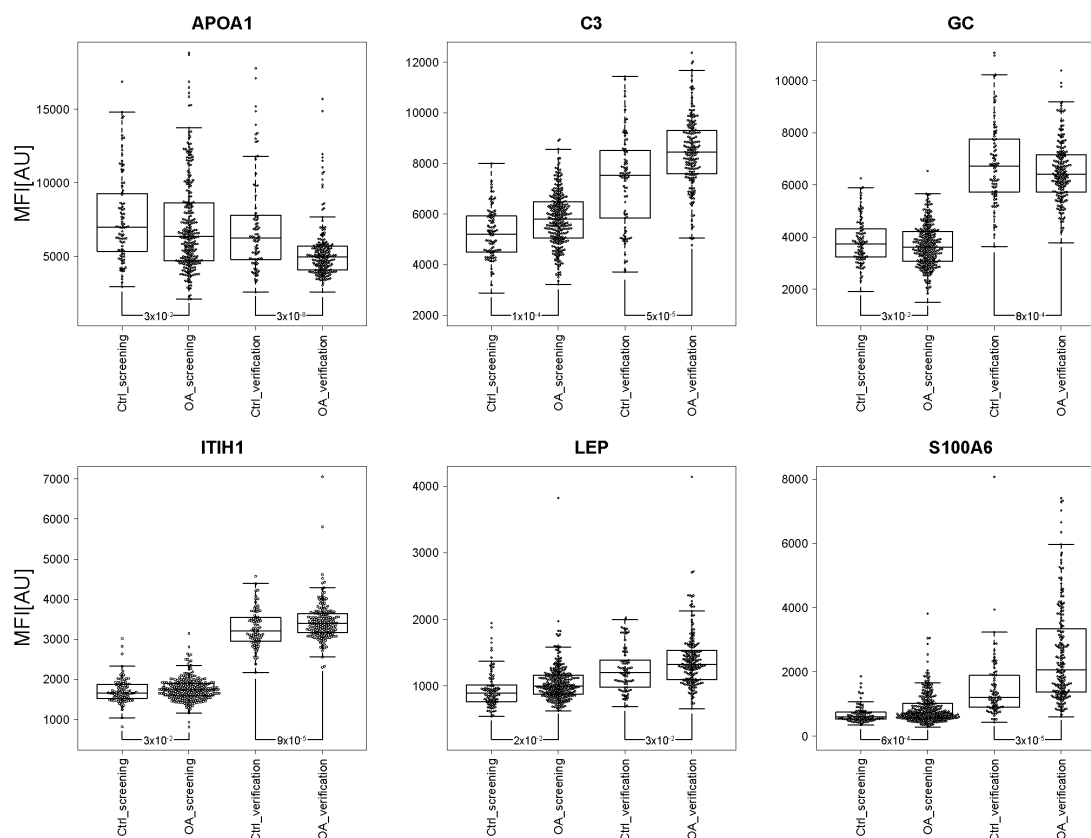


Figure 2. Identification of concordant protein profiles separating osteoarthritis patients from control groups in screening and verification sample set. Box-plots illustrate the proteins revealing significant differences ($p < 0.05$) in different sample cohorts for group comparisons between osteoarthritis (OA) patients and controls (Ctrl).

The classification performance of the identified and concordant protein profiles was visualized by receiver operator characteristic (ROC) curve. This protein panel had an area under the curve (AUC) > 0.70 for classification between OA and controls in the screening sample set. This panel gave a better classification of OA and controls in the verification cohort (AUC=0.81) (Figure 3, Supplementary Table 5).

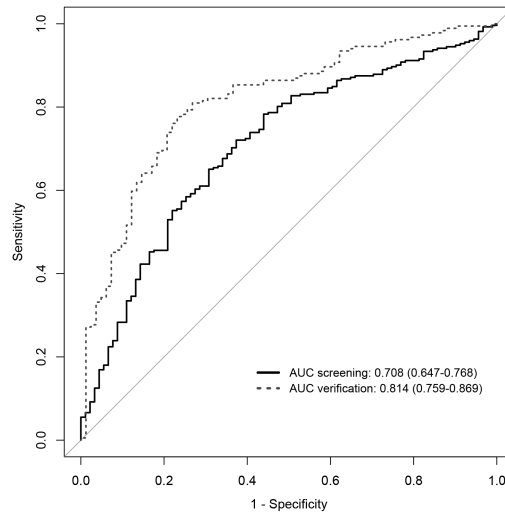


Figure 3. ROC curves represent the classification power of a panel composed of S100A6, GC, ITIH1, APOA1, LEP and C3 profiles between osteoarthritis patients and Control individuals in the screening and the verification cohorts.

Proteins associated with radiographic severity

A non-parametrical analysis was used to compare the profiles of these six proteins between K/L scores and healthy controls in the screening and verification sets of samples.

Patients with K/L=2 showed significantly higher levels of C3, LEP and S100A6 in the screening phase (Figure 4A). Profiles for S100A6 and LEP were confirmed in the verification phase (Figure 4B). However, APOA1 levels were significantly lower in OA than controls in the screening and verification phases (Figure 4A), while GC levels were significantly lower only in the verification phase (Figure 4B).

Patients with K/L=3 showed higher levels of C3, LEP and S100A6 than control individuals in the screening sample set (Figure 4A). The profiles were confirmed for S100A6 and C3 in the verification set, whereas the results for LEP failed to be confirmed (Figure 4B). Interestingly, APOA1 serum protein levels were significantly lower in these patients compared to controls in the verification set (Figure 4B). Finally, patients with K/L=4 showed higher levels of C3, LEP and S100A6 than controls in both screening and verification sets (Figures 4A-B). In this case, ITIH1 and APOA1 levels were significantly higher or lower than in controls, respectively, only in the verification set of samples (Figure 4B).

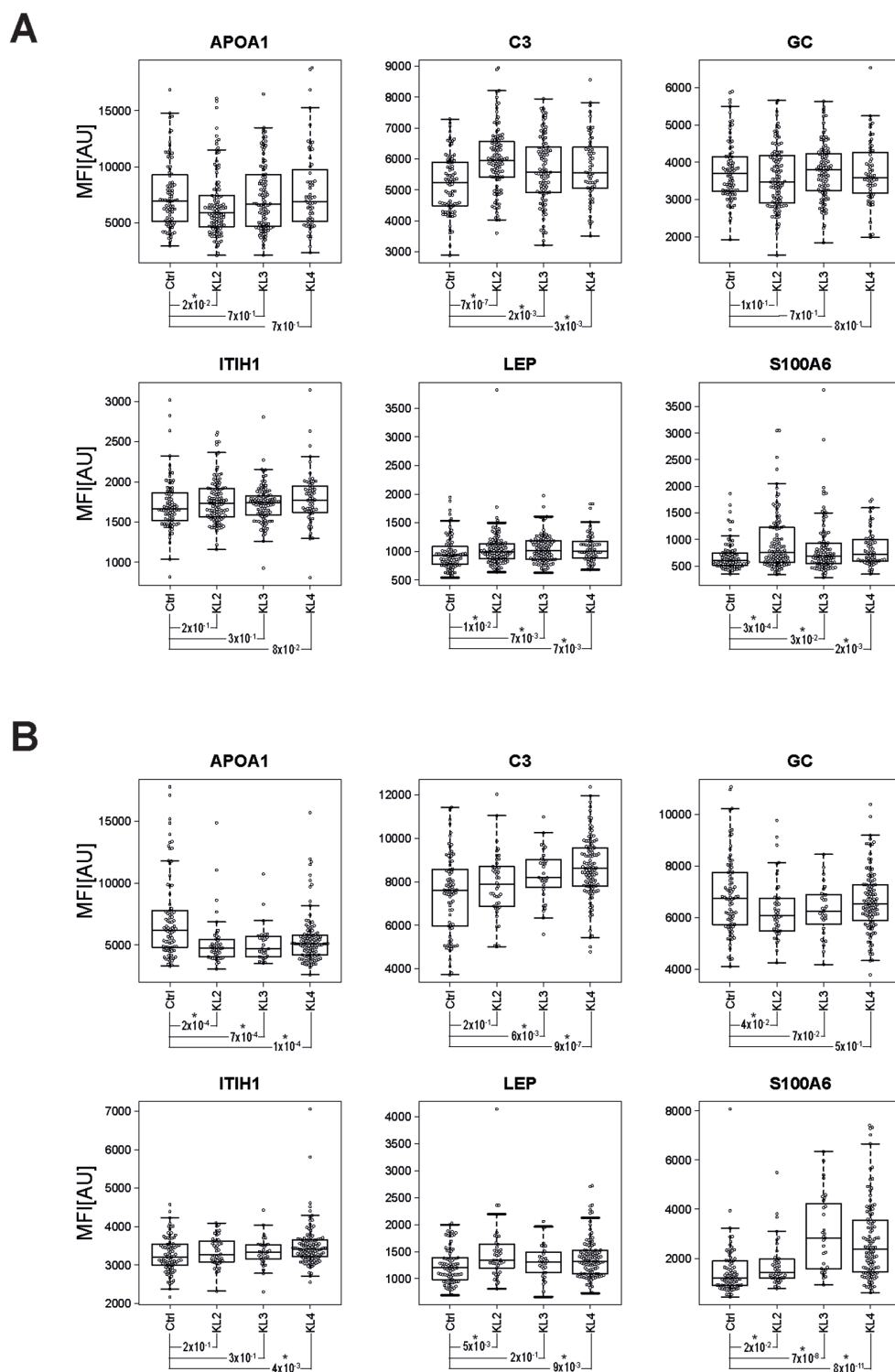


Figure 4. A) Identification of protein revealing differences between different K/L scores in osteoarthritis (OA) patients and control group in screening sample set. B). Identification of protein revealing differences between different K/L scores in osteoarthritis (OA) patients and control group in verification sample set. Significant differences ($p < 0.05$) are indicated with a *

Proteins distinguishing OA and RA and PsA patients

The analysis of screening and verification phases concordantly allowed the detection of 32 proteins to significantly ($p < 0.05$) discriminate OA from RA patients (Supplementary Table 6A). Of these, 18 were increased in OA compared to RA, and 14 were decreased. 60% of these 32 proteins were related to inflammatory processes, lipid metabolism and bone formation. We found that ITIH1, a protein also distinguishing between OA and controls, is also increased in OA patients compared to RA individuals. However, LEP and APOA1 levels are increased in RA compared to OA patients. Comparing OA and PsA patients, 19 proteins were significantly ($p < 0.05$) modulated in screening and verification phases (Supplementary Table 6B). Ten of these proteins were increased in OA compared to PsA, and nine were decreased. Almost 45% of these 19 proteins are also involved in inflammatory response and lipid metabolism.

Among the 16 concordantly identified proteins separating OA and RA as well as OA and PsA patients (Supplementary Figure 3), eight were increased in OA individuals in both comparisons. These proteins are apolipoprotein 4 (APOA4), a protein involved in lipid metabolism; plasma protease C1 inhibitor (SERPING1) and complement 3 (C3), two different proteins related to immune response, extracellular matrix protein 1 (ECM1), a secreted protein related to angiogenesis processes; insuline growth factor binding protein 6 (IGFBP6), involved in negative cell proliferation; pre-lamin A/C (LMNA) which participates in apoptosis process; SPARC-like protein 1 (SPARCL1), a protein involved in bone formation and matrix Gla Protein (MGP) which plays an important role in cartilage formation (Supplementary Table 6A-B).

DISCUSSION

We have performed an extensive profiling of serum samples using two different suspension antibody bead arrays with the aim of identifying protein profiles associated to OA. To our knowledge, this high-throughput technology has been applied for the first time for the analysis of sample collections of this size within osteoarthritis, where serum samples from a total of 1632 individuals were analysed to evaluate the levels of up to 78 different proteins. Using this affinity proteomic approach, we identified six proteins (S100A6, GC, ITIH1, APOA1, LEP and C3) whose levels in serum were able to discriminate between radiographic OA and control individuals, and are therefore likely to be potential biochemical markers for this disease.

To evaluate the specificity of the protein panel, we analyzed samples from two other rheumatic diseases (RA and PsA) along with OA and controls. Interestingly,

among all significantly modulated proteins between the three rheumatic diseases analyzed, only C3 was significantly increased in OA compared to the three other groups (Control, RA and PsA). Complement activation is described to happen in the first stages of OA, and its role has emerged as a crucial factor in experimental OA progression [26]. Complement components are expressed by normal chondrocytes. However, their production is increased in the presence of fragments of extracellular matrix components (ECM). It is known that in OA the catabolic processes compromise the integrity of the cartilage, and fragments of proteins released from ECM as fibromodulin, COMP and osteoadherin lead to the activation of C1q component to further activate the classical and alternative pathways of complement factors [207].

Increased levels of complement C3 in serum from late OA patients compared to healthy donors were already identified in previous proteomics screenings [202]. Furthermore, one of its fragments (C3f) was also described as increased in early OA compared to normal individuals and RA patients [208]. In the present work, we have found that protein C3 levels are higher in serum from radiographic OA compared to all control subjects (Control, RA and PsA), being significant when more severe OA stages (K/L=3 and K/L=4) were compared to controls. Using an antibody-based approach, our result confirms those previous data obtained by mass spectrometry and the potential of C3 evaluation as marker of OA.

Besides to the additional evidence of the role of complement activation in the development of OA, our work also provides new insights into other pathogenic mechanisms of this disease. The identification of significantly increased levels of S100A6 (or calcyclin) in all K/L=2 groups of OA samples compared to controls, and further increased with severe OA (K/L=3 and K/L=4) in screening and verification sample sets, points to a role of this protein in the OA process. Although this calcium-binding protein is known to be expressed in OA cartilage [209], and there is an evidence of its expression in chondrocytes (data shown in HPA portal, www.proteinatlas.org), its role in OA cartilage has not been yet described. It has been suggested that S100A6 could be involved in cell survival by interaction with advanced glycation end products (RAGE), consequent formation of reactive oxygen species (ROS), activation of ERK pathway and changes in NFκ-beta transcriptional activity as well as promoting catabolic process in the cartilage [210, 211]. Our results suggest a potential value of this protein for diagnosis, prognosis and monitoring studies of OA, and underline the need of further functional studies to elucidate the specific role of S100A6 in OA.

Analogous to calcyclin, the levels of inter-alpha-trypsin inhibitor heavy chain (ITIH1) higher in all OA K/L grades compared to controls, although this was only statistically significant in K/L=4 scores in the verification set of samples and thus suggests a value of this protein as marker of advanced OA. It is known that ITIH1 is synthesized by chondrocytes and binds to hyaluronic acid and other extracellular matrix components [212, 213]. This protein was found at higher levels in synovial fluids from OA patients compared to RA [203], and this has been also detected in serum in the present work (Supplementary Table 6A).

In contrast to S100A6 and ITIH1, which are increased in OA, we found decreased levels of GC protein in OA sera compared to controls in the first stages of the disease, although only reaching statistical significance when K/L=2 OA are compared to control from the verification sample set. It has been previously reported that this vitamin-D binding protein is secreted from OA human cartilage [214] and mediates bone resorption by activating osteoclasts [215], thus our findings suggest a novel pathogenic pathway of OA in which vitamin D might play an important role.

Apart from these three proteins that might be directly related with cartilage and bone turnover, our work has also identified in the panel of putative markers for OA two proteins involved in lipid metabolism: APOA1 and LEP. Altered lipid metabolism has been implicated as a critical player in the development of OA. Many studies have demonstrated that people suffering overweight are more likely to develop OA compared to healthy individuals [216]. Other studies suggest that excess of fat mass triggers symptoms as insulin resistance, hyperinsulinemia, glucose intolerance, hypertension, dyslipidemia and cardiovascular disorders that may be related to OA [217, 218]. Our finding of the altered presence of APOA1 and LEP highlights the importance of lipid metabolism in OA pathogenesis.

APOA1 is the major structural protein of high-density lipids (HDL). Its role in OA is still controversial: Although this protein is usually considered as an anti-inflammatory agent [219], other studies showed their role in proinflammatory effects in OA [220]. On the other hand, it has been reported that ApoA1 mRNA and protein expression levels were significantly reduced in OA compared to normal cartilage, and this was accompanied by intracellular lipid accumulation at the early stages of OA [221]. Finally, other studies enhance the hypothesis of a local APOA1 dysregulation inside the OA joint cavity due to the weak correlations between APOA1 levels in blood when compared to synovial fluid levels [220, 222]. In our study, we found that APOA1 serum protein levels were significantly lower in K/L=2 scores compared to Controls in both

screening and verification assays. APOA1 was also decreased in all K/L scores compared to controls although only reaching statistical significance in the verification sample cohort. These results suggest that low levels of APOA1 in serum can be an indicator of OA. Moreover, APOA1 levels are higher in RA compared to OA (Supplementary Table 6A), suggesting that an increased APOA1 could be associated to a more inflammatory arthritis phenotype.

Here, we found the levels of the adipokine leptin to be significantly higher in all K/L groups of OA samples compared to controls, which further increased with increasing K/L (except in the comparison K/L=3, where significantly increased levels were observed only for the screening set). In this sense, OA has been suggested as a systemic disorder in which homeostatic dysregulation within the joint structure might be due to the activity of adipokines [223]. According to this, significant levels of LEP were observed in human OA cartilage compared to normal tissue [224]. Furthermore, LEP was also found in synovial fluids of OA joints, and its expression was related to the grade of cartilage destruction [225]. Altogether, the studies available indicate a pro-inflammatory and catabolic role of leptin on cartilage metabolism [226]. Therefore, our results confirm the association between LEP and OA, and show that increased levels of this protein in serum might be indicative of radiographic OA even at early stages.

CONCLUSION

We present for the first time an affinity proteomics approach comparing serum protein levels of OA and control individuals in a total of 1632 samples. We detected six proteins: S100A6, GC, ITIH1, APOA1, LEP and C3, which were significantly modulated between OA patients and healthy individuals, with a potential value to facilitate OA diagnosis. Interestingly, one of these proteins, C3, discriminated among OA, RA and PsA patients, which suggests C3 as a specific protein for OA. Taken together, upon further verification in independent sample collections, these findings help to a better understanding of OA pathology and provide a great insight in OA biomarker field.

IV. DISCUSIÓN

Discusión capítulo I

En este trabajo nos propusimos caracterizar el patrón de degradación del cartílago en la artrosis (OA) con el fin de buscar potenciales biomarcadores moleculares tempranos y de progresión de la enfermedad.

Para ello, utilizamos un modelo experimental *in vitro* basado en explantes de cartílago articular humano procedente de las zonas lesionadas del cartílago artrósico (WOA), de zonas no dañadas del mismo cartílago artrósico (UOA) así como de cartílagos de personas sin artrosis (N). La principal ventaja de usar explantes de cartílago es que los condrocitos se mantienen en su ambiente natural dentro del tejido, y se evita, por lo tanto, las limitaciones del cultivo celular *in vitro*, asociadas fundamentalmente con la desdiferenciación de los condrocitos.

Después de realizar 4 experimentos iTRAQ, describimos 76 proteínas que están diferencialmente liberadas al secretoma desde diferentes zonas del cartílago artrósico en degradación (Tabla 1). Según el perfil que presentaron en todas las comparaciones, estas proteínas se pudieron clasificar en 6 grupos o *clusters* (Figura 2). Estas proteínas podrían ser agrupadas a su vez en 3 perfiles proteicos (A, B y C).

Así, identificamos 42 proteínas pertenecientes a los *clusters* 1, 4 y 6 cuya abundancia está diferencialmente modulada en las zonas macroscópicamente no dañadas del tejido OA con respecto a los niveles de estas proteínas encontradas en los secretomas de cartílagos normales y esta modulación también se observa en las zonas lesionadas (perfil A). Estos resultados indican que estas proteínas podrían ser indicadores tempranos y de progresión de la enfermedad. De acuerdo con el análisis de enriquecimiento (Tabla Suplementaria 4), estas proteínas están relacionadas con procesos de organización de la matriz extracelular como el colágeno tipo II (CO2A1), *biglycan* (PGS1) o las *tenascins* (TENA, TENX). Identificamos también 31 proteínas pertenecientes a los *clusters* 2 y 3 que se detectan solamente diferencialmente moduladas cuando el tejido está macroscópicamente dañado (perfil B). Varias de estas proteínas están relacionadas con procesos catabólicos de desestructuración de la matriz como el proteoglicano (PGBM) y los colágenos VI, XI, XII, y XV. Finalmente, un pequeño grupo de 3 proteínas pertenecientes al *cluster* 5 presentaban una secreción diferencial en la zona macroscópicamente normal del tejido artrósico que desaparecía cuando el tejido ya estaba visiblemente dañado (perfil C). Estas proteínas corresponden a proteínas que participan en procesos inflamatorios.

Debido a la variabilidad entre los 4 experimentos realizados, utilizamos un ratio de modulación bastante estricto (ratio >1.8 o <0.55). La variabilidad observada puede ser debida a que no comparamos cantidades iguales de proteína entre las 3 condiciones de estudio, sino que comparamos volúmenes iguales procedentes del secretoma de explantes del mismo tamaño. La ventaja de esta estrategia es que se evitan artefactos experimentales debido a la notable presencia de proteínas abundantes en el secretoma. Los inconvenientes de haber utilizado este ratio estricto es que, algunas proteínas que ya habían sido sugeridas como biomarcadores de la artrosis como *cartilage oligomeric matrix protein* (COMP) [227], *proteoglycan4/lubricin* (PRG4) [228], *cartilage intermediate layer protein 1* (CILP1) [229] y *Chitinase 3-like protein 1* (CH3L1/ YKL40) [230] no fueron incluidas en el panel de las 76 proteínas moduladas. En todos estos casos, encontramos un ratio significativo pero menor de 1.8 (1.47 para COMP, 1.36 para PRG4, 1.74 para CILP1, y 1.63 para YKL40). Aunque en nuestro caso, estas proteínas no se han incluido en el panel de proteínas moduladas, los resultados encontrados para estas proteínas son consistentes con los datos publicados en la bibliografía e invita a revisar todas las proteínas identificadas en este trabajo, independientemente del ratio de cuantificación, por su posible utilidad como potencial biomarcador de la artrosis.

Por otro lado, debido a su potencial uso en el diagnóstico temprano y evaluación de la enfermedad, para los ensayos de validación, nos centramos en grupo de proteínas del perfil A. Dentro de este grupo decidimos validar dos proteínas que participan en la remodelación del hueso: *osteoprotegerin* (OPG) y *periostin* (POSTN). La validación de estas proteínas por inmunodetección (Figura 3) demuestra la participación del cartílago en los procesos de señalización del remodelado óseo.

OPG es un inhibidor de la resorción ósea [231]. La disminución de los niveles de esta proteína en las zonas lesionadas del cartílago OA puede estar indicando un aumento en la destrucción del hueso y es consistente con los resultados encontrados en condrocitos OA [232]. Sin embargo, no hemos observado ninguna modulación de su expresión a nivel de RNAm ni a nivel del cartílago con las inmunohistoquímicas realizadas. Además la cuantificación de los niveles de OPG en el líquido sinovial de pacientes con artrosis y pacientes con artritis reumatoide tampoco mostraron diferencias significativas (datos no mostrados). Los resultados encontrados en nuestro trabajo, contradicen los resultados preliminares de otros grupos en los que veían un aumento en la expresión del gen que codifica para OPG en cartílagos OA en comparación con cartílagos sanos [233] y que los niveles de OPG están aumentados

en líquidos sinoviales de pacientes OA en relación con los niveles encontrados en pacientes AR [234]. Además, otros autores han encontrado una relación entre el grado de severidad de la enfermedad y los niveles de OPG en el líquido sinovial de pacientes OA [235]. Debido a las discrepancias de nuestros resultados con la bibliografía descrita hasta la fecha, será necesario realizar más estudios para demostrar el valor y el papel de la OPG en la OA.

Por otra parte, el aumento de POSTN en el secretoma del cartílago OA que se ha detectado en este trabajo concuerda con un aumento significativo de su expresión en el cartílago dañado (Figura 4A, Figura 4B).

POSTN es una proteína dependiente de la vitamina K producida por los osteoblastos que participa en los procesos de mineralización de la matriz extracelular [236]. En un modelo de OA en ratón, se encontró que la expresión de esta proteína estaba incrementada en las articulaciones dañadas con respecto a las no dañadas [237]. Recientemente, también se ha descrito la expresión de POSTN en el hueso subcondral de pacientes con artrosis [238].

Los datos presentados en este trabajo, demuestran por primera vez el aumento en la expresión y en la abundancia proteica de POSTN en el cartílago OA humano, el cual se acompaña con un aumento en la liberación de esta proteína desde el tejido dañado al medio. Además, hemos observado que los niveles de POSTN se encuentran aumentados en el líquido sinovial de pacientes artrósicos con respecto a los niveles de esta proteína en los líquidos sinoviales de pacientes con artritis reumatoide y con artritis psoriásica. Todos estos datos sugieren que la POSTN puede tener valor como marcador específico del proceso artrósico.

Hasta la fecha, este es el primer estudio en el que se realiza un análisis cuantitativo de los secretomas procedentes de explantes de cartílago humano. En global, los resultados mostrados en este trabajo proporcionan una valiosa información acerca de los marcadores proteicos de degradación del cartílago humano en la artrosis y su utilidad debería ser investigada en otros fluidos biológicos para el desarrollo de nuevos métodos de diagnóstico de la OA y la monitorización de terapias frente a la enfermedad.

Discusión capítulo II

En este estudio, realizamos un análisis semicuantitativo del contenido proteico del líquido sinovial procedente de las rodillas de pacientes con artrosis (OA) así como de pacientes con artritis reumatoide (AR). Con la finalidad de minimizar la variabilidad entre los pacientes [239], agrupamos los líquidos sinoviales en 2 *pools* de 10 muestras cada uno.

Este trabajo fue el primero en el que se utilizó la inmudepleción del líquido sinovial para estudios proteómicos. El líquido sinovial posee un elevado rango dinámico, con proteínas muy abundantes que representan el 95% del volumen proteico total, y que enmascaran las proteínas menos abundantes, las cuales tienen un mayor valor biomarcador [240]. Por este motivo, decidimos utilizar un método de inmunodeplección para eliminar las 20 proteínas más abundantes de la muestra y aumentar el número de péptidos detectados y proteínas identificadas. Esta estrategia llevó a la identificación de un total de 136 proteínas en las muestras de líquido sinovial analizadas, lo que corresponde al mayor número de proteínas descrito hasta ese momento en los estudios proteómicos utilizando este tipo de muestras. Además, la mayoría de las proteínas identificadas no habían sido descritas previamente. Por otro lado, de las 20 proteínas más abundantes en el líquido sinovial [241], sólo 4 fueron identificadas en este estudio, demostrando la eficacia del método de depleción utilizado.

En las muestras de pacientes AR se detectaron un total de 2498 péptidos correspondientes a un total de 122 proteínas identificadas con al menos dos péptidos únicos (Tabla suplementaria 1). En muestras de pacientes OA se identificaron 2582 péptidos correspondientes a un total de 107 proteínas con dos péptidos específicos (Tabla suplementaria 2). De las 136 proteínas totales identificadas, 29 fueron identificadas sólo en pacientes AR, 14 sólo estaban presentes en el líquido sinovial de pacientes OA y 93 fueron identificadas en las dos condiciones de estudio (Figura 2). Una vez caracterizados los perfiles proteicos de cada grupo de muestras, decidimos utilizar los espectros de MS para cuantificar de forma relativa las proteínas diferencialmente moduladas entre pacientes AR y OA mediante contaje espectral. El contaje espectral es un método que ha sido ampliamente utilizado en proteómica para la cuantificación relativa de proteínas con el fin de evitar los costes y el tiempo que conlleva el marcaje químico de las muestras. Aunque normalmente este método se acopla a la ionización por electrospray, las técnicas donde se acopla la

nanocromatografía con la espectrometría MALDI-TOF/TOF también ha demostrado su potencial para realizar análisis comparativos [242-244]. Para la cuantificación, se utilizaron los péptidos únicos de cada proteína puesto que éstos han demostrado presentar una mayor correlación con la abundancia de la proteína en la muestra [245].

Con el conteo espectral se calculó una media del ratio entre los péptidos identificados en cada condición de estudio y se consideró que la proteína estaba modulada entre las diferentes condiciones si presentaba un ratio menor 0.8 o mayor de 1.2. El rango de desviaciones estándar fue bastante amplio indicando que la reproducibilidad entre las réplicas podría ser mejorada. Por esta razón, en este trabajo consideramos como diferencialmente abundantes sólo aquellas proteínas identificadas con al menos 4 péptidos en una de las dos enfermedades analizadas. La reproducibilidad del método se describe también con el número de réplicas en las que la proteína ha sido detectada, indicado como frecuencia en las Tablas 1 y 2.

El análisis reveló diferencias cuantitativas entre los líquidos sinoviales de pacientes con OA y AR. Las proteínas más abundantes en pacientes con AR correspondían a proteínas relacionadas con la respuesta inmune, a la activación del complemento y a la inflamación. Una de estas proteínas es la S100A8 que es una proteína con un papel importante en la respuesta de inflamación aguda, ya identificada previamente en pacientes con AR [246-248]. Los neutrófilos son células predominantes en pacientes con AR [249]. Muchas de las proteínas identificadas en estos pacientes correspondían a proteínas relacionadas con neutrófilos que nunca antes se habían descrito en el compartimento sinovial como son las proteínas azurocidin, defensin 2, lipocalin asociada a gelatinasa, o el inhibidor de elastasa leucocitaria.

En muestras OA, las proteínas más abundantes estaban relacionadas con componentes estructurales de la matriz como el *agrecan core protein*, *cartilage oligomeric matrix protein* (COMP), *cartilage acidic protein 1* (CRAC1), o proteínas relacionadas con la remodelación del cartílago como *trombospondin-4* (TSP4) y *fibronectin* (FINC). También encontramos proteínas que nunca se habían descrito antes como dos inhibidores de la angiogénesis (TSP1 y PEDF) y la *chitinase-like protein 1* (CH3L1), una proteína secretada por las células del cartílago que ha sido previamente propuesta como marcador de artrosis.

Después de este análisis, nos propusimos validar por inmunodetección algunos de los resultados obtenidos por espectrometría de masas. Así, realizamos la inmunodetección frente a dos proteínas que pertenecían a las proteínas características del grupo de pacientes AR como eran la MMP1 y la BIGH3. Para ello, utilizamos

muestras individuales de 12 pacientes OA y 12 pacientes AR incluidas previamente para los *pools*.

MMP1 es una colagenasa conocida por romper los colágenos tipo I, II y III de la matriz extracelular del cartílago por lo tanto juega un papel esencial en las dos enfermedades reumáticas estudiadas donde se produce una degradación del cartílago [250, 251]. Mediante esta técnica, vimos que la MMP1 estaba significativamente disminuida en pacientes OA frente a los AR (Figura 4). El ratio obtenido no se corresponde con el observado por el análisis de masas, en el que no conseguimos detectar esta proteína en OA. Como se ha comprobado por inmunodetección, esto no significa que la proteína no esté presente, sino que su concentración en el líquido sinovial está por debajo del límite de detección de la técnica de espectrometría de masas que se ha utilizado en este trabajo.

La BGH3 es una proteína que juega un papel importante en la adhesión célula-colágeno [252] y es la primera vez que se detecta en líquido sinovial. La disminución de los niveles de esta proteína en pacientes OA con respecto a pacientes AR fue confirmado también por inmunodetección. (Figura 4). En un trabajo previo de nuestro grupo encontramos que esta proteína está incrementada después de tratar los condrocitos con la citoquina proinflamatoria interleucina 1 β [253]. Por lo tanto, el presente trabajo validaría estudios previos de nuestro grupo y soporta el valor de esta proteína como marcador de inflamación en la articulación.

Por otra parte, validamos algunas de las proteínas características de pacientes OA como *fibronectin* (FINC) y *gelsolin* (GELS). Para ello, utilizamos 8 muestras individuales de pacientes OA y 8 AR utilizados para preparar los *pools* del análisis proteómico previo.

FINC es una proteína de la matriz extracelular esencial para la adhesión celular [254, 255] y hemos verificado por inmunodetección que es una proteína que está en mayores niveles en OA que en AR (Figura 5). Estos resultados coinciden con otros resultados publicados en los que presentan que los fragmentos de esta proteína estimulan a las agregasas y metaloproteasas de la matriz que promueven la degradación del cartílago [256].

Por otra parte, la GELS ha sido descrita previamente como biomarcador de AR [257]. Sin embargo, nosotros hemos encontrado que tanto en el análisis de masas como en la inmunodetección (Tabla 2, Figura 5), el ratio de modulación coincide en los dos aproximaciones, indicando que esta proteína está más aumentada en pacientes

OA que en AR. Por lo tanto, estos resultados indican que es necesario realizar más estudios para determinar el papel de esta proteína en la OA.

Las proteínas identificadas en este trabajo reflejan las diferencias entre las dos enfermedades reumáticas: la AR es una enfermedad caracterizada por la acumulación en el líquido sinovial de proteínas relacionadas con la respuesta inmune, mientras que la OA muestra un incremento en la abundancia de las proteínas relacionadas con la remodelación de la matriz extracelular, como consecuencia de la degradación del cartílago. Además, la mayoría de las proteínas identificadas en este trabajo no habían sido descritas previamente y aportan valiosa información acerca de la etiopatogénesis de las dos enfermedades reumáticas a la vez que proporcionan una serie de proteínas con potencial biomarcador tanto en la OA como en la AR, puesto que su presencia en el líquido sinovial es anterior a su dilución en el suero del paciente. Su valor biomarcador podría ser analizado y evaluado en el suero o plasma de pacientes con OA, AR y también controles sanos sin enfermedad reumática.

Discusión capítulo III

Aunque la artrosis no es una enfermedad con origen autoinmune, varios trabajos reflejan el papel de la respuesta inmune en la enfermedad [27, 32, 258] y muchos consideran que esta respuesta se produce en las primeras fases del desarrollo de la enfermedad [24].

En este trabajo, nos hemos propuesto utilizar dos tipos de *arrays* de antígenos con el objetivo de explorar el repertorio de posibles autoanticuerpos en sueros de pacientes con artrosis (OA), pacientes con artritis reumatoide (AR) y controles sanos. La finalidad de utilizar dos plataformas diferentes de *arrays* de antígenos es cubrir el mayor número de inmureactividades posibles utilizando tanto fragmentos de proteínas como proteínas completas.

Los *arrays* de fragmentos proteicos en formato plano se emplearon para un análisis no dirigido de la inmunorreactividad donde un total de 62 muestras de suero se interrogaron frente a 3840 fragmentos proteicos que representaban 3307 proteínas seleccionadas aleatoriamente desde el *Human Protein Atlas* (HPA). Los antígenos que mostraron reactividad, se seleccionaron para una siguiente fase de validación utilizando *arrays* de fragmentos proteicos en formato de esferas.

Las mismas muestras se analizaron paralelamente con los *arrays* tipo NAPPA. Esta plataforma se utilizó para hacer un análisis dirigido de la inmunorreactividad puesto que se analizaron 80 proteínas completas seleccionadas específicamente por su posible relevancia en la patología artrósica.

Dos de los 9 antígenos reactivos observados en el *array* de fragmentos proteicos, KCNB2 y KCNJ6, son proteínas constituyentes de canales de potasio. KCNB2 presentó una mayor reactividad en AR y KCNJ6 presentó una mayor reactividad en OA. Hasta la fecha, hay varios trabajos que relacionan las proteínas de canales de potasio con enfermedades inflamatorias como la esclerosis múltiple [259] y la artritis [260]. Por lo tanto, las proteínas de canales de potasio, KCNB2 y KCNJ6 encontrados en este trabajo como posibles antígenos, pueden tener algún papel interesante en el proceso inflamatorio de la AR y de la OA respectivamente.

Por su parte, los NAPPA son *arrays* que evitan los problemas asociados con la expresión de proteínas recombinantes en sistemas heterólogos, incluyendo los inconvenientes asociados con la purificación de las proteínas expresadas y el mantenimiento de su estabilidad. El sistema de expresión *in vitro* utilizado en este trabajo, permite además el plegamiento de la proteína en su estado conformacional.

Las proteínas inmunorreactivas encontradas en NAPPA, pertenecen a proteínas de proliferación celular (IGBP4 e IGFBP6), metabolismo de los lípidos (LEP), proteínas de adhesión celular (CD44) y proteínas de la matriz extracelular (CHST14 y PCOLCE). Interesantemente, una proteína, LEP, presentaba mayor reactividad en OA en comparación con controles sanos y con pacientes AR. Esta proteína es una adipoquina involucrada en la degradación de la matriz extracelular y en la inflamación [223, 226] y los niveles de esta proteína en suero han sido también descritos para estar aumentados en pacientes con artrosis en comparación con pacientes sanos [225]. Los resultados mostrados en este trabajo, podrían sugerir por tanto una respuesta inmune frente a los elevados niveles de LEP en pacientes con OA.

Por otro lado, la superposición entre los resultados observados en los *arrays* de fragmentos proteicos y en el NAPPA fue baja. Sólo 41 de los 81 fragmentos proteicos que representaban las proteínas expresadas en el NAPPA, presentaron reactividad en el *array* de fragmentos proteicos. Esta divergencia se ejemplifica en la Figura 7, la cual muestra la reactividad observada para CHST14 e IL6, dos proteínas que significativamente presentaban una reactividad diferente en OA y AR respectivamente cuando los comparábamos con pacientes sanos en la plataforma NAPPA. Como se muestra en esta figura, en el *array* de fragmentos proteicos, no se observaron reactividades para el fragmento antigénico que representaba a la proteína CHST14 en el *array*, mientras que la presencia de anticuerpos para esta proteína fue detectada en el NAPPA y confirmada por ELISA.

La inmunogenicidad observada con los NAPPA y no en el *array* de fragmentos proteicos puede estar relacionada con los epítomos conformacionales accesibles para los autoanticuerpos en el NAPPA pero no así en el *array* de fragmentos proteicos.

La CHST14 es una sulfotransferasa implicada en la síntesis de dermatán sulfato, un glicosaminoglicano que participa en la estabilidad de la red de fibras de colágeno embebidas en la matriz extracelular del cartílago. Aunque es necesario realizar más investigaciones para validar estos resultados, los anticuerpos frente a la CHST14 en suero de pacientes con OA puede ser relevantes como indicadores de la enfermedad y su diagnóstico.

Por otra parte, la citoquina IL6 es clave en la patogénesis de la AR [261] y una diana terapéutica muy interesante para la enfermedad [262]. Esta proteína fue representada en el *array* de fragmentos proteicos con dos epítomos lineales, y sólo uno de ellos presentaba reactividad en pacientes con AR. Este ejemplo refleja las ventajas de este tipo de *array*, puesto que éstos utilizan secuencias concretas de los antígenos

y con esta plataforma, se puede detectar una inmunorreactividad más específica. Sin embargo, la mayoría de los epítomos representados en el *array* de fragmentos proteicos son epítomos lineales y por lo tanto, los epítomos conformacionales no van a ser detectados con esta plataforma, liderando a falsos negativos en los resultados, en el caso de que los fragmentos no representen el epítomo específico frente a los que se generan los autoanticuerpos.

La aproximación metodológica utilizada en este trabajo representa una nueva estrategia en el descubrimiento de nuevos biomarcadores en artrosis basada en la detección de autoanticuerpos. La combinación de los resultados obtenidos desde los dos enfoques técnicos utilizados, aumenta el número de potenciales dianas inmurreactivas en la OA y presenta a estos autoanticuerpos encontrados en sueros de pacientes con OA como una fuente de potenciales biomarcadores de la enfermedad.

Discusión capítulo IV

En este trabajo hemos analizado una gran cantidad de muestras usando dos *arrays* de anticuerpos en suspensión con el objetivo de buscar perfiles proteicos asociados a la artrosis. Este estudio supone la primera vez en la que esta tecnología se aplica en la artrosis para analizar hasta 78 proteínas diferentes en un total de 1632 individuos.

Usando esta aproximación, se identificó un panel de 6 proteínas (S100A6, GC, ITIH1, APOA1, LEP and C3) que permitían distinguir entre pacientes con artrosis radiográfica de rodilla e individuos sin daño radiográfico de rodilla (Figura 2). Para evaluar la especificidad del perfil de proteínas encontrado, se analizaron también dos cohortes de pacientes con otras enfermedades reumáticas (AR y APs) junto con los pacientes con artrosis y los controles.

Así, hemos identificado que los niveles de C3 están aumentados en los pacientes OA con respecto a los pacientes AR, APs e individuos control y que sus niveles están significativamente aumentados en grados radiológicos avanzados cuando los comparamos con los controles.

A nivel experimental, varios autores han descrito que la activación del complemento ocurre en las primeras fases de la artrosis y que podría tener un papel importante en la progresión de la enfermedad [26]. Por otro lado, esta proteína ya había sido descrita previamente por nuestro grupo para identificar la artrosis avanzada en el suero de los pacientes artrósicos cuando estos se comparaban con individuos sanos [202]. Además, un fragmento de esta proteína, C3f, ha sido descrita también por otro grupo para distinguir los sueros de los pacientes OA, AR e individuos sanos [208]. Por lo tanto, los resultados presentados aquí confirman los resultados obtenidos por espectrometría de masas y la utilidad de C3 como marcador de artrosis.

A parte de reforzar la evidencia de la implicación de C3 en la artrosis, nuestro trabajo también aporta información acerca de otros posibles mecanismos implicados en la enfermedad.

Así, hemos encontrado que la S100A6 está aumentada en todos los grados radiográficos K/L comparado con los controles sanos en las dos cohortes de muestras analizadas y que sus niveles aumentan con relación a la severidad del daño en la articulación (Figura 4A,4B). S100A6 es una proteína de unión al calcio que se expresa en los condrocitos y en el cartílago artrósico [209]. Se desconoce todavía la función de esta proteína en la artrosis pero podría estar implicada en la regulación de la

supervivencia celular [210, 211]. Los resultados presentados en este trabajo sugieren el potencial valor de esta proteína como marcador de diagnóstico y pronóstico de la artrosis y abre nuevas perspectivas para la realización de estudios funcionales para determinar el papel específico de esta proteína en la artrosis.

Por otra parte, hemos visto que ITIH1 está aumentada en pacientes OA con respecto a controles, pero que sus niveles en suero sólo están significativamente aumentados en grados K/L=4 comparado con los controles en la cohorte de verificación indicando el valor de esta proteína como marcador de una artrosis avanzada. ITIH1 es una proteína que se expresa en los condrocitos y que se une a las proteínas de la matriz extracelular. En trabajos previos de nuestro grupo también hemos observado que los niveles de esta proteína están aumentados en los líquidos sinoviales de pacientes OA con respecto a pacientes con AR [203]. Estos resultados también se han confirmado en el presente trabajo en el suero (Tabla suplementaria 5A).

Interesantemente, también encontramos que los niveles de la proteína GC están significativamente disminuidos en las primeras fases de la enfermedad (K/L=2) con respecto a los controles. Esta proteína es secretada por los condrocitos [214] y produce la adsorción del hueso y la activación de los osteoclastos [215]. Así, nuestros resultados sugieren una nueva vía patogénica de la artrosis en la que la vitamina D puede tener un papel importante.

A parte de estas tres proteínas relacionadas directamente con el cartílago y el metabolismo del hueso. Nuestro trabajo también pone de relieve la implicación del metabolismo lipídico con la identificación de APOA1 y LEP en el panel de las 6 proteínas con potencial marcador de diagnóstico de la artrosis.

APOA1 es el principal componente estructural de los lípidos de alta densidad (HDL). Su papel en la artrosis no está del todo claro puesto que hay estudios que le adjudican un papel antiinflamatorio [219], mientras que otros han comprobado experimentalmente su posible papel proinflamatorio en la enfermedad [220]. Por otra parte, hay estudios que han comprobado que los niveles de expresión de APOA1 están disminuidos en los cartílagos artrósicos [221]. En nuestro estudio, hemos observado que los niveles de APOA1 están disminuidos en el suero de individuos control con respecto a pacientes con artrosis en todos los grados K/L, aunque estos resultados sólo fueron significativos en la cohorte de verificación. Estos resultados sugieren que los bajos niveles de APOA1 en el suero pueden ser un indicador de artrosis. Además, hemos observado que los niveles de APOA1 están más aumentados en el suero de pacientes AR con respecto a los pacientes OA, lo que sugiere que los niveles

aumentados de APOA1 pueden estar asociados con un fenotipo de artritis más inflamatoria.

Finalmente, hemos observado que los niveles de la adipocina leptina están significativamente aumentados en todos los grados K/L en comparación con los controles. En concordancia con esto, la artrosis se ha definido como un desorden metabólico en el cual se produce un desequilibrio en la articulación debido a la influencia de las adipocinas [223]. Apoyando esta teoría, se han observado niveles más altos de leptina en los cartílagos artrósicos con respecto a los normales [224] y también que los niveles de LEP en el líquido sinovial de pacientes artrósicos aumentaban con el grado de destrucción del cartílago [225]. Así, nuestros resultados confirman la asociación entre la LEP y la artrosis y muestran que los niveles altos de LEP en el suero pueden ser indicadores de la artrosis incluso en fases tempranas de la enfermedad.

En este trabajo, presentamos un estudio utilizando *arrays* de anticuerpos en formato de esfera que permitió identificar un panel de 6 proteínas que distinguían a los pacientes artrósicos de los controles sanos analizados. Aunque el valor de estas proteínas debe ser explorado en un mayor número de pacientes, estos resultados reflejan el potencial de este panel para el diagnóstico de la artrosis de rodilla.

Discusión general

El interés por la búsqueda de nuevos biomarcadores de diagnóstico, pronóstico, monitorización y predicción de respuesta a tratamiento está aumentando en los últimos años. Estos biomarcadores podrían ayudar al diagnóstico precoz de las enfermedades, a la estratificación de los pacientes y al desarrollo de terapias personalizadas para cada individuo. Dentro del campo de los biomarcadores, la proteómica es una disciplina que ha demostrado su capacidad para identificar proteínas con potencial marcador de las enfermedades.

Esta tesis se ha centrado en la búsqueda de potenciales biomarcadores de artrosis combinando diferentes técnicas proteómicas basadas en espectrometría de masas y en *arrays* de proteínas.

La artrosis es la enfermedad reumática con mayor prevalencia en la sociedad. Es una enfermedad que, en sus fases más iniciales, se muestra clínicamente silente, de modo que, cuando el paciente llega a la consulta con algún síntoma, la enfermedad ya está avanzada y no es posible recuperar la funcionalidad normal del individuo. Además no existe terapia eficaz frente a la enfermedad. Por este motivo, existe un elevado interés por buscar biomarcadores de la artrosis que permitan realizar un diagnóstico precoz y una monitorización de los posibles tratamientos alternativos frente a la enfermedad.

La artrosis se caracteriza principalmente por una degradación del cartílago articular que va acompañado por alteraciones en el hueso y la membrana sinovial. El encontrar marcadores específicos de estos procesos, ayudaría a desarrollar métodos de diagnóstico y seguimiento de la enfermedad.

Así, el primer trabajo presentado en esta tesis, se centró en identificar el perfil proteico de la degradación del cartílago en la artrosis mediante marcaje iTRAQ e identificación por MALDI-TOF/TOF utilizando un modelo *in vitro* de explantes de cartílago humano. En este trabajo, se identificaron 76 proteínas que indicaban un perfil de degradación diferencial del cartílago artrósico. De estas 76 proteínas, se identificaron 31 proteínas con potencial indicador de la degradación tardía del cartílago OA, 42 proteínas con potencial marcador del daño temprano en el cartílago y de la progresión de la degradación en el cartílago OA; y 3 proteínas que indicaban únicamente una degradación temprana del tejido. De todas ellas, se validaron dos, la OPG y la POSTN, como candidatos para indicar una degradación progresiva del cartílago y por lo tanto, con potencial biomarcador en la artrosis. En este mismo

trabajo, también se utilizaron *arrays* de anticuerpos en formato de esfera para detectar que los niveles de POSTN estaban aumentados en líquidos sinoviales de pacientes con artrosis en comparación con los niveles encontrados en los líquidos sinoviales procedentes de pacientes con otras enfermedades reumáticas (AR y APs). Estos resultados sugieren la especificidad de esta proteína como marcadora de la degradación del cartílago artrósico.

En el segundo trabajo, se comparó el perfil proteico de los líquidos sinoviales procedentes de pacientes con artrosis con el encontrado en los pacientes con artritis reumatoide con el objetivo de identificar el perfil proteico característico del compartimento sinovial en la artrosis. El 66% de las proteínas totales identificadas en este trabajo habían sido también detectadas en el secretoma de los cartílagos artrósicos analizados en el primer trabajo. Estos datos indican que las proteínas encontradas en el líquido sinovial de los pacientes artrósicos proceden del cartílago dañado y confirman la validez del modelo *in vitro* utilizado en esta tesis para analizar el perfil de degradación del cartílago en la artrosis.

Algunas de las proteínas identificadas en el primer trabajo como marcadoras de la degradación diferencial del cartílago artrósico son PEDF, CRAC1, MMP2 e ITIH1. En el segundo estudio presentado en esta tesis, estas proteínas se encontraron aumentadas en los líquidos sinoviales de pacientes artrósicos en comparación con pacientes con artritis reumatoide. Estos resultados podrían sugerir que estas 4 proteínas son marcadores específicos de la degradación del cartílago en la artrosis.

Por otra parte, la proteína POSTN, identificada y validada en el primer trabajo como posible marcador de la degradación progresiva del cartílago en artrosis y aumentada en líquidos sinoviales de pacientes con artrosis en comparación con pacientes AR y APS, no se identificó en ninguno de los líquidos sinoviales analizados en el segundo trabajo. Las discrepancias entre estos resultados pueden ser debidas a la sensibilidad analítica de las diferentes técnicas proteómicas utilizadas. En el primer trabajo, se detectaron diferencias en los niveles de POSTN en el líquido sinovial empleando una técnica dirigida basada en anticuerpos, mientras que en el segundo trabajo se utilizó una técnica proteómica no dirigida utilizando LC-MALDI-TOF/TOF. El MALDI-TOF/TOF es un analizador de masas que tiene un rango dinámico reducido en estudios de proteómica. Por lo tanto, en muestras con un amplio rango dinámico como el líquido sinovial, las proteínas que están en muy bajas concentraciones, pueden no ser detectadas.

Por otra parte, el suero es una muestra biológica muy utilizada para buscar biomarcadores debido a su rápida disponibilidad y fácil aplicabilidad en la práctica clínica. Con esta premisa, en esta tesis decidimos utilizar *arrays* de proteínas con el objetivo de buscar proteínas con potencial biomarcador en el suero de pacientes con artrosis.

Hay trabajos que indican que la respuesta inmune en la artrosis se produce en las primeras etapas de la enfermedad y que se generan autoanticuerpos frente a los fragmentos liberados desde los compartimentos articulares, siendo éstos autoanticuerpos potenciales indicadores del daño temprano en la articulación. Siguiendo esa hipótesis, en el tercer trabajo presentado en esta tesis, utilizamos dos tipos de *arrays* de antígenos para buscar autoanticuerpos en el suero de pacientes con artrosis. Los resultados obtenidos en este trabajo, indican que los niveles de inmunorreactividad frente a HOXD3, LEPREL2, CHST14, CD44, LEP y PCOLCE están incrementados en pacientes artrósicos con respecto a controles sanos. Por otra parte, observamos que la inmunorreactividad frente KCNJ6, IGFBP4, IGFBP6 y LEP está incrementada en pacientes artrósicos con respecto a pacientes con artritis reumatoide mientras que la reactividad frente a CCDC85C, KCNB2 y ZNF784 está disminuida.

Finalmente, en el último trabajo presentado en esta tesis, se utilizaron *arrays* de anticuerpos en formato esfera para determinar potenciales biomarcadores de la artrosis en el suero. Empleando esta aproximación, se observó que los niveles de las proteínas S100A6, C3, ITIH1 y LEP están aumentados en los sueros de pacientes con artrosis de rodilla con respecto a controles sanos. Mientras que los niveles de GC y APOA1, están disminuidos. Este panel de 6 proteínas permitió clasificar con una buena sensibilidad y especificidad a los pacientes artrósicos de dos poblaciones independientes.

Además, la proteína S100A6 permitió distinguir todos los grados K/L radiológicos de los controles sanos, indicando el potencial valor de esta proteína como biomarcador de la severidad en la artrosis de rodilla. Por otra parte, los niveles de C3 estaban aumentados en pacientes artrósicos en relación con los niveles encontrados en pacientes con AR y APS. Estos resultados sugieren el potencial de esta proteína para diagnosticar específicamente a los pacientes con artrosis de rodilla.

Por otro lado, ITIH1 es una proteína expresada en el cartílago articular humano. En esta tesis, ITIH1 fue identificada en el secretoma de los cartílagos artrósicos. Sus niveles se encontraron aumentados en el líquido sinovial de pacientes OA con respecto a pacientes AR y también en el suero de pacientes OA con respecto a controles sanos

y pacientes AR. Aunque los niveles de esta proteína sólo están significativamente aumentados en los sueros de pacientes con artrosis de rodilla K/L=4 con respecto a los controles analizados, a la vista de los resultados recogidos en esta tesis, sería interesante realizar más estudios para explorar el papel de esta proteína en la enfermedad así como para estudiar su utilidad como posible marcador de la artrosis.

En este cuarto trabajo también se ha detectado que la LEP está incrementada en el suero de pacientes OA en comparación con controles sanos. Por otra parte, en el tercer trabajo presentado en esta tesis, también se ha visto que los niveles de autoanticuerpos frente a LEP están aumentados en pacientes OA con respecto a controles y pacientes con artritis reumatoide. Sería interesante, por lo tanto, en un futuro, estudiar la relación de estas observaciones y ver, si los niveles de LEP en el suero se correlacionan con los niveles de autoanticuerpos en esos mismos pacientes.

En la artrosis no existen biomarcadores validados en la clínica. Uno de los motivos es porque en la gran mayoría de los casos, los resultados tienen que ser validados en muestras de sangre o suero y en un gran número de pacientes. Esto de por sí supone un reto importante para la proteómica puesto que, como se ha mencionado a lo largo de la tesis, este compartimento biológico, posee un rango dinámico muy amplio y puede que no se encuentre el biomarcador investigado.

Por otro lado, un inconveniente inherente a la búsqueda de biomarcadores es la gran variabilidad que existe entre los pacientes. Esta variabilidad en la artrosis se puede deber en gran medida a la gran heterogeneidad que existe entre los fenotipos de la artrosis y a que es muy difícil diagnosticar el inicio de la enfermedad. Así, en los estudios de verificación de los biomarcadores es necesario incluir muchas variables clínicas de seguimiento de los pacientes acompañado de un componente de análisis de imagen para encontrar biomarcadores que nos permitan estratificar a los pacientes y definir el valor del biomarcador encontrado en la artrosis.

Teniendo en cuenta que la artrosis afecta a diferentes tejidos, en esta tesis, se han empleado diferentes tipos de muestras, desde el secretoma del cartílago y líquido sinovial hasta el suero de pacientes con la finalidad de buscar potenciales biomarcadores específicos de la artrosis.

El trabajo presentado aquí, pone de manifiesto la utilidad de combinar la espectrometría de masas con la proteómica de afinidad en todas las fases de la búsqueda de biomarcadores, puesto que ambas pueden aportar datos complementarios y relevantes para el desarrollo de biomarcadores. Por otra parte, en el tercer trabajo presentado en esta tesis, los *arrays* de proteínas también han

demostrado su capacidad para utilizarlos en las fases más tempranas de la búsqueda de biomarcadores.

Este trabajo indica además, la importancia de incluir en la búsqueda de biomarcadores, a parte de casos y controles sanos, también pacientes con otras enfermedades para determinar la especificidad del potencial biomarcador encontrado.

Globalmente, en esta tesis se han identificado perfiles proteicos asociados a la artrosis con potencial valor para el diagnóstico de la enfermedad. Sin embargo, el descubrimiento de estos candidatos con potencial biomarcador no los hace un biomarcador. Así, los hallazgos presentados aquí deben de ser verificados por otras técnicas y validados en largas colecciones de muestras procedentes de diferentes localizaciones con el fin de definir el valor clínico de las proteínas encontradas y seguir avanzando en la búsqueda, verificación y validación de biomarcadores de la artrosis.

V. CONCLUSIONES

Basándonos en el trabajo presentado en esta tesis, podemos establecer una serie de conclusiones que se enumeran a continuación:

1. La combinación del marcaje iTRAQ y el análisis LC-MS/MS para el estudio de la degradación del cartílago artrósico humano ha permitido identificar 76 proteínas diferencialmente liberadas a partir del cartílago artrósico.
2. De estas 76 proteínas, 3 proteínas reflejan un daño temprano en el cartílago artrósico, 31 indican un daño tardío, y 42 proteínas señalan tanto un daño temprano como tardío en el tejido, lo que sugiere que este último grupo de proteínas podrían ser potenciales marcadores tempranos y de progresión de la degradación del cartílago artrósico.
3. La OPG y la POSTN han sido validadas como posibles marcadores de la progresión de la degradación del cartílago en la artrosis.
4. Los niveles de POSTN son mayores en líquido sinovial de pacientes con artrosis que en pacientes con artritis reumatoide y artritis psoriásica, lo que sugiere el potencial valor de esta proteína como marcador específico de la artrosis.
5. La inmunodepleción seguida de análisis por LC-MALDI-MS/MS ha permitido realizar una cuantificación relativa de las proteínas presentes en el líquido sinovial de pacientes con artrosis. Se ha validado que FINC y GELS se encuentran aumentadas en el líquido sinovial de pacientes con artrosis con respecto a pacientes con artritis reumatoide, mientras que MMP1 y BIGH3 están disminuidas.
6. Las proteínas aumentadas en líquido sinovial de pacientes con artrosis con respecto a pacientes con artritis reumatoide se relacionan principalmente con el remodelado de la matriz extracelular del cartílago, mientras que la mayoría de las proteínas aumentadas en pacientes con artritis reumatoide están involucradas en procesos inflamatorios.

7. Utilizando dos estrategias diferentes de *arrays* de antígenos, se han encontrado autoanticuerpos en el suero de pacientes con artrosis. Los niveles de inmunorreactividad frente a HOXD3, LEPREL2, CHST14, CD44, LEP y PCOLCE están incrementados en pacientes artrósicos con respecto a controles sanos. La inmunorreactividad frente KCNJ6, IGFBP4, IGFBP6 y LEP está incrementada en pacientes artrósicos con respecto a pacientes con artritis reumatoide, mientras que la inmunorreactividad para los antígenos CCDC85C, KCNB2 y ZNF784 está disminuida.
8. Se ha identificado un panel de 6 proteínas formado por APOA1, GC, ITIH1, S100A6 y C3, cuyos niveles en suero permiten diferenciar pacientes con artrosis radiográfica de rodilla de controles sanos.
9. De todas las proteínas del panel, sólo la S100A6 está aumentada en todos los grados K/L radiográficos con respecto a los controles sanos, por lo que esta proteína podría ser indicadora del daño radiográfico de la artrosis de rodilla.
10. De todas las proteínas del panel, sólo la proteína C3 permite diferenciar los pacientes artrósicos de los controles sanos y de las dos enfermedades reumáticas (artritis reumatoide y artritis psoriásica) analizadas, indicando el potencial de esta proteína como biomarcador específico para el diagnóstico de la artrosis de rodilla.

VI. CONCLUSIONS

According to the chapters presented and discussed in this thesis we can establish the following conclusions:

1. The combination of iTRAQ labelling and LC-MS/MS analysis to study the osteoarthritic cartilage degradation enabled the identification of 76 proteins that are differentially released by osteoarthritic cartilage.
2. Among these 76 proteins, 3 proteins reflect an early damage in the osteoarthritic cartilage, 31 indicate only a late damage, and 42 proteins show both an early and late damage in the tissue. This last group of proteins could be putative early and progression of osteoarthritic cartilage degradation.
3. We validated OPG and POSTN as potential progression biomarkers of cartilage degradation in osteoarthritis.
4. POSTN protein levels are higher in synovial fluids from patients with osteoarthritis than those observed in rheumatoid arthritis and psoriatic arthritis individuals, which suggests the potential value of this protein as specific biomarker of osteoarthritis.
5. Immunodepletion followed by LC-MALDI-MS/MS analysis enabled to carry out a relative quantitation of proteins in synovial fluids from osteoarthritis patients. We validated that FINC and GELS protein levels are increased in synovial fluids from osteoarthritis patients compared to rheumatoid arthritis patients, whereas MMP1 and BIGH3 are decreased.
6. Proteins increased in synovial fluids from osteoarthritis patients compared to rheumatoid arthritis patients are mainly related to the extracellular matrix remodelling, whereas most of the proteins increased in rheumatoid arthritis are involved in inflammatory processes.

7. The application of two different approaches of antigen arrays enabled the identification of autoantibodies in serum from osteoarthritis patients. The immunoreactivity towards HOXD3, LEPREL2, CHST14, CD44, LEP and PCOLCE is increased in osteoarthritis patients compared to healthy individuals. The immunoreactivity against KCNJ6, IGFBP4, IGFBP6 and LEP is elevated in osteoarthritis patients compared to rheumatoid arthritis patients, whereas the immunoreactivity towards CCDC85C, KCNB2 and ZNF784 is decreased.
8. The serum protein levels of a panel composed of APOA1, GC, ITIH1, S100A6 and C3, enable to distinguish patients with radiographic knee osteoarthritis and healthy individuals.
9. Among all the proteins from the panel, only S100A6 is increased in all radiographic K/L categories compared to healthy individuals, which suggests the potential value of this protein to indicate radiographic knee damage in osteoarthritis.
10. Among all the proteins from the panel, only C3 enables to distinguish osteoarthritis patients, healthy controls and the other two rheumatic diseases analysed (rheumatoid arthritis and psoriatic arthritis), indicating the potential of this protein as a specific biomarker for the diagnosis of knee osteoarthritis.

VII. BIBLIOGRAFÍA

- [1] Benito-Ruiz P. Estructura y función del cartílago. Manual SER de la artrosis. 2002.
- [2] Benito-Ruiz P. Artrosis. Fisiopatología, Diagnóstico y tratamiento. 2010.
- [3] H. M. The water of articular cartilage. The human joint in health and disease. 1978.
- [4] Thonar EJ MK, Manicourt DH, Kuettner KE. Structure and function of normal human articular cartilage. Springer. 1999:1-19.
- [5] Brandt KD, Dieppe P, Radin EL. Etiopathogenesis of osteoarthritis. Rheumatic diseases clinics of North America. 2008;34:531-59.
- [6] Loeser RF, Goldring SR, Scanzello CR, Goldring MB. Osteoarthritis: a disease of the joint as an organ. Arthritis Rheum. 2012;64:1697-707.
- [7] Kraus VB, Blanco FJ, Englund M, Henrotin Y, Lohmander LS, Losina E, et al. OARSI Clinical Trials Recommendations: Soluble biomarker assessments in clinical trials in osteoarthritis. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society. 2015;23:686-97.
- [8] Man GS, Mologhianu G. Osteoarthritis pathogenesis - a complex process that involves the entire joint. Journal of medicine and life. 2014;7:37-41.
- [9] Aigner T, Soder S, Gebhard PM, McAlinden A, Haag J. Mechanisms of disease: role of chondrocytes in the pathogenesis of osteoarthritis--structure, chaos and senescence. Nat Clin Pract Rheumatol. 2007;3:391-9.
- [10] Michigami T. Regulatory mechanisms for the development of growth plate cartilage. Cellular and molecular life sciences : CMLS. 2013;70:4213-21.
- [11] Buckwalter JA, Mankin HJ. Articular cartilage: tissue design and chondrocyte-matrix interactions. Instructional course lectures. 1998;47:477-86.
- [12] Bailey AJ, Mansell JP, Sims TJ, Barse X. Biochemical and mechanical properties of subchondral bone in osteoarthritis. Biorheology. 2004;41:349-58.
- [13] Burr DB. Anatomy and physiology of the mineralized tissues: role in the pathogenesis of osteoarthritis. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society. 2004;12 Suppl A:S20-30.
- [14] Burr DB, Gallant MA. Bone remodelling in osteoarthritis. Nature reviews Rheumatology. 2012;8:665-73.
- [15] Neogi T. Clinical significance of bone changes in osteoarthritis. Therapeutic advances in musculoskeletal disease. 2012;4:259-67.
- [16] Lajeunesse D, Reboul P. Subchondral bone in osteoarthritis: a biologic link with articular cartilage leading to abnormal remodeling. Current opinion in rheumatology. 2003;15:628-33.
- [17] Lories RJ, Luyten FP. The bone-cartilage unit in osteoarthritis. Nature reviews Rheumatology. 2011;7:43-9.
- [18] Goldring MB, Otero M. Inflammation in osteoarthritis. Current opinion in rheumatology. 2011;23:471-8.
- [19] Sellam J, Berenbaum F. The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. Nature reviews Rheumatology. 2010;6:625-35.

- [20] Haywood L, McWilliams DF, Pearson CI, Gill SE, Ganesan A, Wilson D, et al. Inflammation and angiogenesis in osteoarthritis. *Arthritis Rheum.* 2003;48:2173-7.
- [21] Haseeb A, Haqqi TM. Immunopathogenesis of osteoarthritis. *Clinical immunology.* 2013;146:185-96.
- [22] Hill CL, Hunter DJ, Niu J, Clancy M, Guermazi A, Genant H, et al. Synovitis detected on magnetic resonance imaging and its relation to pain and cartilage loss in knee osteoarthritis. *Ann Rheum Dis.* 2007;66:1599-603.
- [23] de Lange-Brokaar BJ, Ioan-Facsinay A, van Osch GJ, Zuurmond AM, Schoones J, Toes RE, et al. Synovial inflammation, immune cells and their cytokines in osteoarthritis: a review. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society.* 2012;20:1484-99.
- [24] Ene R, Sinescu RD, Ene P, Cirstoiu MM, Cirstoiu FC. Synovial inflammation in patients with different stages of knee osteoarthritis. *Romanian journal of morphology and embryology = Revue roumaine de morphologie et embryologie.* 2015;56:169-73.
- [25] Yuan GH, Masuko-Hongo K, Kato T, Nishioka K. Immunologic intervention in the pathogenesis of osteoarthritis. *Arthritis Rheum.* 2003;48:602-11.
- [26] Liu-Bryan R, Terkeltaub R. Emerging regulators of the inflammatory process in osteoarthritis. *Nature reviews Rheumatology.* 2015;11:35-44.
- [27] Revell PA, Mayston V, Lalor P, Mapp P. The synovial membrane in osteoarthritis: a histological study including the characterisation of the cellular infiltrate present in inflammatory osteoarthritis using monoclonal antibodies. *Ann Rheum Dis.* 1988;47:300-7.
- [28] Ishii H, Tanaka H, Katoh K, Nakamura H, Nagashima M, Yoshino S. Characterization of infiltrating T cells and Th1/Th2-type cytokines in the synovium of patients with osteoarthritis. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society.* 2002;10:277-81.
- [29] Nakamura H, Yoshino S, Kato T, Tsuruha J, Nishioka K. T-cell mediated inflammatory pathway in osteoarthritis. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society.* 1999;7:401-2.
- [30] Krenn V, Hensel F, Kim HJ, Souto Carneiro MM, Starostik P, Ristow G, et al. Molecular IgV(H) analysis demonstrates highly somatic mutated B cells in synovialitis of osteoarthritis: a degenerative disease is associated with a specific, not locally generated immune response. *Laboratory investigation; a journal of technical methods and pathology.* 1999;79:1377-84.
- [31] Shiokawa S, Matsumoto N, Nishimura J. Clonal analysis of B cells in the osteoarthritis synovium. *Ann Rheum Dis.* 2001;60:802-5.
- [32] Foell D, Wittkowski H, Roth J. Mechanisms of disease: a 'DAMP' view of inflammatory arthritis. *Nat Clin Pract Rheumatol.* 2007;3:382-90.
- [33] Alsalameh S, Mollenhauer J, Hain N, Stock KP, Kalden JR, Burmester GR. Cellular immune response toward human articular chondrocytes. T cell reactivities against chondrocyte and fibroblast membranes in destructive joint diseases. *Arthritis Rheum.* 1990;33:1477-86.

- [34] Xiang Y, Sekine T, Nakamura H, Imajoh-Ohmi S, Fukuda H, Nishioka K, et al. Proteomic surveillance of autoimmunity in osteoarthritis: identification of triosephosphate isomerase as an autoantigen in patients with osteoarthritis. *Arthritis Rheum.* 2004;50:1511-21.
- [35] Tsuruha J, Masuko-Hongo K, Kato T, Sakata M, Nakamura H, Nishioka K. Implication of cartilage intermediate layer protein in cartilage destruction in subsets of patients with osteoarthritis and rheumatoid arthritis. *Arthritis Rheum.* 2001;44:838-45.
- [36] Tsuruha J, Masuko-Hongo K, Kato T, Sakata M, Nakamura H, Sekine T, et al. Autoimmunity against YKL-39, a human cartilage derived protein, in patients with osteoarthritis. *The Journal of rheumatology.* 2002;29:1459-66.
- [37] Sakata M, Tsuruha JI, Masuko-Hongo K, Nakamura H, Matsui T, Sudo A, et al. Autoantibodies to osteopontin in patients with osteoarthritis and rheumatoid arthritis. *The Journal of rheumatology.* 2001;28:1492-5.
- [38] Xiang Y, Sekine T, Nakamura H, Imajoh-Ohmi S, Fukuda H, Yudoh K, et al. Fibulin-4 is a target of autoimmunity predominantly in patients with osteoarthritis. *Journal of immunology.* 2006;176:3196-204.
- [39] Arden N, Nevitt MC. Osteoarthritis: epidemiology. *Best practice & research Clinical rheumatology.* 2006;20:3-25.
- [40] Litwic A, Edwards MH, Dennison EM, Cooper C. Epidemiology and burden of osteoarthritis. *British medical bulletin.* 2013;105:185-99.
- [41] Felson DT, Naimark A, Anderson J, Kazis L, Castelli W, Meenan RF. The prevalence of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis Rheum.* 1987;30:914-8.
- [42] Karlson EW, Mandl LA, Aweh GN, Sangha O, Liang MH, Grodstein F. Total hip replacement due to osteoarthritis: the importance of age, obesity, and other modifiable risk factors. *The American journal of medicine.* 2003;114:93-8.
- [43] Felson DT, Zhang Y, Hannan MT, Naimark A, Weissman BN, Aliabadi P, et al. The incidence and natural history of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis Rheum.* 1995;38:1500-5.
- [44] Hurley MV. The role of muscle weakness in the pathogenesis of osteoarthritis. *Rheumatic diseases clinics of North America.* 1999;25:283-98, vi.
- [45] Dekker J, Tola P, Aufdemkampe G, Winckers M. Negative affect, pain and disability in osteoarthritis patients: the mediating role of muscle weakness. *Behaviour research and therapy.* 1993;31:203-6.
- [46] Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society.* 2005;13:769-81.

- [47] Maleki-Fischbach M, Jordan JM. New developments in osteoarthritis. Sex differences in magnetic resonance imaging-based biomarkers and in those of joint metabolism. *Arthritis research & therapy*. 2010;12:212.
- [48] Zhang Y, McAlindon TE, Hannan MT, Chaisson CE, Klein R, Wilson PW, et al. Estrogen replacement therapy and worsening of radiographic knee osteoarthritis: the Framingham Study. *Arthritis Rheum*. 1998;41:1867-73.
- [49] Hanna FS, Wluka AE, Bell RJ, Davis SR, Cicuttini FM. Osteoarthritis and the postmenopausal woman: Epidemiological, magnetic resonance imaging, and radiological findings. *Seminars in arthritis and rheumatism*. 2004;34:631-6.
- [50] Wluka AE, Davis SR, Bailey M, Stuckey SL, Cicuttini FM. Users of oestrogen replacement therapy have more knee cartilage than non-users. *Ann Rheum Dis*. 2001;60:332-6.
- [51] Cirillo DJ, Wallace RB, Wu L, Yood RA. Effect of hormone therapy on risk of hip and knee joint replacement in the Women's Health Initiative. *Arthritis Rheum*. 2006;54:3194-204.
- [52] Nevitt MC, Felson DT, Williams EN, Grady D. The effect of estrogen plus progestin on knee symptoms and related disability in postmenopausal women: The Heart and Estrogen/Progestin Replacement Study, a randomized, double-blind, placebo-controlled trial. *Arthritis Rheum*. 2001;44:811-8.
- [53] Spector TD, Cicuttini F, Baker J, Loughlin J, Hart D. Genetic influences on osteoarthritis in women: a twin study. *Bmj*. 1996;312:940-3.
- [54] Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, Saito S, et al. A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. *Nature genetics*. 2007;39:529-33.
- [55] Valdes AM, Evangelou E, Kerkhof HJ, Tamm A, Doherty SA, Kisand K, et al. The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. *Ann Rheum Dis*. 2011;70:873-5.
- [56] Evangelou E, Valdes AM, Kerkhof HJ, Styrkarsdottir U, Zhu Y, Meulenbelt I, et al. Meta-analysis of genome-wide association studies confirms a susceptibility locus for knee osteoarthritis on chromosome 7q22. *Ann Rheum Dis*. 2011;70:349-55.
- [57] Day-Williams AG, Southam L, Panoutsopoulou K, Rayner NW, Esko T, Estrada K, et al. A variant in MCF2L is associated with osteoarthritis. *American journal of human genetics*. 2011;89:446-50.
- [58] arc OC, arc OC, Zeggini E, Panoutsopoulou K, Southam L, Rayner NW, et al. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet*. 2012;380:815-23.
- [59] Belfer I, Dai F, Kehlet H, Finelli P, Qin L, Bittner R, et al. Association of functional variations in COMT and GCH1 genes with postherniotomy pain and related impairment. *Pain*. 2015;156:273-9.

- [60] Loughlin J. The genetic epidemiology of human primary osteoarthritis: current status. *Expert reviews in molecular medicine*. 2005;7:1-12.
- [61] Rego-Perez I, Fernandez-Moreno M, Fernandez-Lopez C, Arenas J, Blanco FJ. Mitochondrial DNA haplogroups: role in the prevalence and severity of knee osteoarthritis. *Arthritis Rheum*. 2008;58:2387-96.
- [62] Fernandez-Moreno M, Tamayo M, Soto-Hermida A, Mosquera A, Oreiro N, Fernandez-Lopez C, et al. mtDNA haplogroup J modulates telomere length and nitric oxide production. *BMC musculoskeletal disorders*. 2011;12:283.
- [63] Rego-Perez I, Fernandez-Moreno M, Soto-Hermida A, Fernandez-Lopez C, Oreiro N, Blanco FJ. Mitochondrial genetics and osteoarthritis. *Frontiers in bioscience*. 2013;5:360-8.
- [64] Sowers M, Lachance L. Vitamins and arthritis. The roles of vitamins A, C, D, and E. *Rheumatic diseases clinics of North America*. 1999;25:315-32.
- [65] McAlindon TE, Jacques P, Zhang Y, Hannan MT, Aliabadi P, Weissman B, et al. Do antioxidant micronutrients protect against the development and progression of knee osteoarthritis? *Arthritis Rheum*. 1996;39:648-56.
- [66] Lane NE, Gore LR, Cummings SR, Hochberg MC, Scott JC, Williams EN, et al. Serum vitamin D levels and incident changes of radiographic hip osteoarthritis: a longitudinal study. Study of Osteoporotic Fractures Research Group. *Arthritis Rheum*. 1999;42:854-60.
- [67] Niu J, Zhang YQ, Torner J, Nevitt M, Lewis CE, Aliabadi P, et al. Is obesity a risk factor for progressive radiographic knee osteoarthritis? *Arthritis Rheum*. 2009;61:329-35.
- [68] Spector TD, Hart DJ, Doyle DV. Incidence and progression of osteoarthritis in women with unilateral knee disease in the general population: the effect of obesity. *Ann Rheum Dis*. 1994;53:565-8.
- [69] Felson DT, Anderson JJ, Naimark A, Walker AM, Meenan RF. Obesity and knee osteoarthritis. The Framingham Study. *Annals of internal medicine*. 1988;109:18-24.
- [70] Felson DT, Zhang Y, Anthony JM, Naimark A, Anderson JJ. Weight loss reduces the risk for symptomatic knee osteoarthritis in women. The Framingham Study. *Annals of internal medicine*. 1992;116:535-9.
- [71] Christensen R, Bartels EM, Astrup A, Bliddal H. Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: a systematic review and meta-analysis. *Ann Rheum Dis*. 2007;66:433-9.
- [72] Wills AK, Black S, Cooper R, Coppack RJ, Hardy R, Martin KR, et al. Life course body mass index and risk of knee osteoarthritis at the age of 53 years: evidence from the 1946 British birth cohort study. *Ann Rheum Dis*. 2012;71:655-60.
- [73] Yusuf E, Nelissen RG, Ioan-Facsinay A, Stojanovic-Susulic V, DeGroot J, van Osch G, et al. Association between weight or body mass index and hand osteoarthritis: a systematic review. *Ann Rheum Dis*. 2010;69:761-5.
- [74] Sandell LJ. Obesity and osteoarthritis: is leptin the link? *Arthritis Rheum*. 2009;60:2858-60.

- [75] Terlain B, Presle N, Pottie P, Mainard D, Netter P. [Leptin: a link between obesity and osteoarthritis?]. *Bulletin de l'Academie nationale de medecine*. 2006;190:1421-35; discussion 35-7, 75-7.
- [76] Berenbaum F, Eymard F, Houard X. Osteoarthritis, inflammation and obesity. *Current opinion in rheumatology*. 2013;25:114-8.
- [77] Nevitt MC, Zhang Y, Javaid MK, Neogi T, Curtis JR, Niu J, et al. High systemic bone mineral density increases the risk of incident knee OA and joint space narrowing, but not radiographic progression of existing knee OA: the MOST study. *Ann Rheum Dis*. 2010;69:163-8.
- [78] Jacobsen S, Jensen TW, Bach-Mortensen P, Hyldstrup L, Sonne-Holm S. Low bone mineral density is associated with reduced hip joint space width in women: results from the Copenhagen Osteoarthritis Study. *Menopause*. 2007;14:1025-30.
- [79] Li B, Aspden RM. Composition and mechanical properties of cancellous bone from the femoral head of patients with osteoporosis or osteoarthritis. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*. 1997;12:641-51.
- [80] Messier SP, Loeser RF, Miller GD, Morgan TM, Rejeski WJ, Sevick MA, et al. Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: the Arthritis, Diet, and Activity Promotion Trial. *Arthritis Rheum*. 2004;50:1501-10.
- [81] Spector TD, Harris PA, Hart DJ, Cicuttini FM, Nandra D, Etherington J, et al. Risk of osteoarthritis associated with long-term weight-bearing sports: a radiologic survey of the hips and knees in female ex-athletes and population controls. *Arthritis Rheum*. 1996;39:988-95.
- [82] Gelber AC, Hochberg MC, Mead LA, Wang NY, Wigley FM, Klag MJ. Joint injury in young adults and risk for subsequent knee and hip osteoarthritis. *Annals of internal medicine*. 2000;133:321-8.
- [83] Lohmander LS, Ostenberg A, Englund M, Roos H. High prevalence of knee osteoarthritis, pain, and functional limitations in female soccer players twelve years after anterior cruciate ligament injury. *Arthritis Rheum*. 2004;50:3145-52.
- [84] Lynch JA, Parimi N, Chaganti RK, Nevitt MC, Lane NE, Study of Osteoporotic Fractures Research G. The association of proximal femoral shape and incident radiographic hip OA in elderly women. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society*. 2009;17:1313-8.
- [85] Doherty M, Courtney P, Doherty S, Jenkins W, Maciewicz RA, Muir K, et al. Nonspherical femoral head shape (pistol grip deformity), neck shaft angle, and risk of hip osteoarthritis: a case-control study. *Arthritis Rheum*. 2008;58:3172-82.
- [86] Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. *Jama*. 2001;286:188-95.

- [87] Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis.* 2014;73:1323-30.
- [88] Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet.* 2012;380:2197-223.
- [89] Carmona L, Ballina J, Gabriel R, Laffon A, Group ES. The burden of musculoskeletal diseases in the general population of Spain: results from a national survey. *Ann Rheum Dis.* 2001;60:1040-5.
- [90] Oliveria SA, Felson DT, Reed JI, Cirillo PA, Walker AM. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. *Arthritis Rheum.* 1995;38:1134-41.
- [91] Loza E, Lopez-Gomez JM, Abasolo L, Maese J, Carmona L, Battle-Gualda E, et al. Economic burden of knee and hip osteoarthritis in Spain. *Arthritis Rheum.* 2009;61:158-65.
- [92] Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society.* 2007;15 Suppl A:A1-56.
- [93] Rousseau JC, Delmas PD. Biological markers in osteoarthritis. *Nat Clin Pract Rheumatol.* 2007;3:346-56.
- [94] Guermazi A, Hayashi D, Roemer FW, Felson DT. Osteoarthritis: a review of strengths and weaknesses of different imaging options. *Rheumatic diseases clinics of North America.* 2013;39:567-91.
- [95] Finan PH, Buenaver LF, Bounds SC, Hussain S, Park RJ, Haque UJ, et al. Discordance between pain and radiographic severity in knee osteoarthritis: findings from quantitative sensory testing of central sensitization. *Arthritis Rheum.* 2013;65:363-72.
- [96] Felson DT, Niu J, Guermazi A, Sack B, Aliabadi P. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. *Ann Rheum Dis.* 2011;70:1884-6.
- [97] Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis.* 1957;16:494-502.
- [98] Ding C, Zhang Y, Hunter D. Use of imaging techniques to predict progression in osteoarthritis. *Current opinion in rheumatology.* 2013;25:127-35.
- [99] Altman RD, Lozada CJ. Practice guidelines in the management of osteoarthritis. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society.* 1998;6 Suppl A:22-4.
- [100] Hochberg MC, Altman RD, Brandt KD, Clark BM, Dieppe PA, Griffin MR, et al. Guidelines for the medical management of osteoarthritis. Part I. Osteoarthritis of the hip. *American College of Rheumatology. Arthritis Rheum.* 1995;38:1535-40.
- [101] Rousseau J, Garnero P. Biological markers in osteoarthritis. *Bone.* 2012;51:265-77.

- [102] Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, et al. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis*. 1995;16:1090-4.
- [103] Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, et al. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnology & genetic engineering reviews*. 1996;13:19-50.
- [104] Lau AT, He QY, Chiu JF. Proteomic technology and its biomedical applications. *Sheng wu hua xue yu sheng wu wu li xue bao Acta biochimica et biophysica Sinica*. 2003;35:965-75.
- [105] HUPO. A gene-centric human proteome project: HUPO--the Human Proteome organization. *Molecular & cellular proteomics : MCP*. 2010;9:427-9.
- [106] Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. *Bmj*. 1994;308:1552.
- [107] Chatterjee SK, Zetter BR. Cancer biomarkers: knowing the present and predicting the future. *Future oncology*. 2005;1:37-50.
- [108] Lotz M, Martel-Pelletier J, Christiansen C, Brandi ML, Bruyere O, Chapurlat R, et al. Republished: Value of biomarkers in osteoarthritis: current status and perspectives. *Postgraduate medical journal*. 2014;90:171-8.
- [109] Kraus VB, Burnett B, Coindreau J, Cottrell S, Eyre D, Gendreau M, et al. Application of biomarkers in the development of drugs intended for the treatment of osteoarthritis. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society*. 2011;19:515-42.
- [110] van Spil WE, DeGroot J, Lems WF, Oostveen JC, Lafeber FP. Serum and urinary biochemical markers for knee and hip-osteoarthritis: a systematic review applying the consensus BIPED criteria. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society*. 2010;18:605-12.
- [111] Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecular & cellular proteomics : MCP*. 2012;11:O111 016717.
- [112] Gorg A, Weiss W, Dunn MJ. Current two-dimensional electrophoresis technology for proteomics. *Proteomics*. 2004;4:3665-85.
- [113] Herbert B, Galvani M, Hamdan M, Olivieri E, MacCarthy J, Pedersen S, et al. Reduction and alkylation of proteins in preparation of two-dimensional map analysis: why, when, and how? *Electrophoresis*. 2001;22:2046-57.
- [114] O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*. 1975;250:4007-21.
- [115] Valcu CM, Valcu M. Reproducibility of two-dimensional gel electrophoresis at different replication levels. *J Proteome Res*. 2007;6:4677-83.
- [116] Unlu M, Morgan ME, Minden JS. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis*. 1997;18:2071-7.

- [117] Hoffmann D, Iinuma Y, Herrmann H. Development of a method for fast analysis of phenolic molecular markers in biomass burning particles using high performance liquid chromatography/atmospheric pressure chemical ionisation mass spectrometry. *Journal of chromatography A*. 2007;1143:168-75.
- [118] Biemann K, Gapp G., and Seibl, J. Applications of Mass Spectrometry to Structure Proteins.I.Amino Acids Sequence and Peptides. *Journal of the American Chemical Society*. 1959;81(9):2274:5.
- [119] Hunt DF, Buko AM, Ballard JM, Shabanowitz J, Giordani AB. Sequence analysis of polypeptides by collision activated dissociation on a triple quadrupole mass spectrometer. *Biomedical mass spectrometry*. 1981;8:397-408.
- [120] Proteómica SEd. Manual de Proteómica. 2015.
- [121] Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry*. 1988;60:2299-301.
- [122] Loo JA, Udseth HR, Smith RD. Peptide and protein analysis by electrospray ionization-mass spectrometry and capillary electrophoresis-mass spectrometry. *Anal Biochem*. 1989;179:404-12.
- [123] Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003;422:198-207.
- [124] Di Girolamo F, Lante I, Muraca M, Putignani L. The Role of Mass Spectrometry in the "Omics" Era. *Current organic chemistry*. 2013;17:2891-905.
- [125] Miller I, Crawford J, Gianazza E. Protein stains for proteomic applications: which, when, why? *Proteomics*. 2006;6:5385-408.
- [126] Diezel W, Kopperschlager G, Hofmann E. An improved procedure for protein staining in polyacrylamide gels with a new type of Coomassie Brilliant Blue. *Anal Biochem*. 1972;48:617-20.
- [127] Lopez MF, Berggren K, Chernokalskaya E, Lazarev A, Robinson M, Patton WF. A comparison of silver stain and SYPRO Ruby Protein Gel Stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling. *Electrophoresis*. 2000;21:3673-83.
- [128] Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics : MCP*. 2002;1:376-86.
- [129] Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & cellular proteomics : MCP*. 2004;3:1154-69.
- [130] Neilson KA, Keighley T, Pascovici D, Cooke B, Haynes PA. Label-free quantitative shotgun proteomics using normalized spectral abundance factors. *Methods in molecular biology*. 2013;1002:205-22.

- [131] Nahnsen S, Bielow C, Reinert K, Kohlbacher O. Tools for label-free peptide quantification. *Molecular & cellular proteomics* : MCP. 2013;12:549-56.
- [132] Ekins RP. Multi-analyte immunoassay. *Journal of pharmaceutical and biomedical analysis*. 1989;7:155-68.
- [133] Ekins R, Chu F. Multianalyte microspot immunoassay. The microanalytical 'compact disk' of the future. *Annales de biologie clinique*. 1992;50:337-53.
- [134] Ekins RP, Chu F. Developing multianalyte assays. *Trends in biotechnology*. 1994;12:89-94.
- [135] Wingren C, Borrebaeck CA. Progress in miniaturization of protein arrays--a step closer to high-density nanoarrays. *Drug discovery today*. 2007;12:813-9.
- [136] Espina V, Mehta AI, Winters ME, Calvert V, Wulfschlegel J, Petricoin EF, 3rd, et al. Protein microarrays: molecular profiling technologies for clinical specimens. *Proteomics*. 2003;3:2091-100.
- [137] MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. *Science*. 2000;289:1760-3.
- [138] Phizicky E, Bastiaens PI, Zhu H, Snyder M, Fields S. Protein analysis on a proteomic scale. *Nature*. 2003;422:208-15.
- [139] Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, et al. Analysis of yeast protein kinases using protein chips. *Nature genetics*. 2000;26:283-9.
- [140] Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, et al. Global analysis of protein activities using proteome chips. *Science*. 2001;293:2101-5.
- [141] Kusnezow W, Hoheisel JD. Solid supports for microarray immunoassays. *Journal of molecular recognition* : JMR. 2003;16:165-76.
- [142] Rusmini F, Zhong Z, Feijen J. Protein immobilization strategies for protein biochips. *Biomacromolecules*. 2007;8:1775-89.
- [143] Sutandy FX, Qian J, Chen CS, Zhu H. Overview of protein microarrays. *Current protocols in protein science / editorial board, John E Coligan [et al]*. 2013;Chapter 27:Unit 27 1.
- [144] Tomizaki KY, Usui K, Mihara H. Protein-detecting microarrays: current accomplishments and requirements. *Chembiochem : a European journal of chemical biology*. 2005;6:782-99.
- [145] Kusnezow W, Jacob A, Walijew A, Diehl F, Hoheisel JD. Antibody microarrays: an evaluation of production parameters. *Proteomics*. 2003;3:254-64.
- [146] Chandra H, Reddy PJ, Srivastava S. Protein microarrays and novel detection platforms. *Expert review of proteomics*. 2011;8:61-79.
- [147] Ray S, Mehta G, Srivastava S. Label-free detection techniques for protein microarrays: prospects, merits and challenges. *Proteomics*. 2010;10:731-48.
- [148] Horan PK, Wheelless LL, Jr. Quantitative single cell analysis and sorting. *Science*. 1977;198:149-57.

- [149] Kettman JR, Davies T, Chandler D, Oliver KG, Fulton RJ. Classification and properties of 64 multiplexed microsphere sets. *Cytometry*. 1998;33:234-43.
- [150] de Jager W, te Velthuis H, Prakken BJ, Kuis W, Rijkers GT. Simultaneous detection of 15 human cytokines in a single sample of stimulated peripheral blood mononuclear cells. *Clinical and diagnostic laboratory immunology*. 2003;10:133-9.
- [151] dupont NC, Wang K, Wadhwa PD, Culhane JF, Nelson EL. Validation and comparison of luminex multiplex cytokine analysis kits with ELISA: determinations of a panel of nine cytokines in clinical sample culture supernatants. *Journal of reproductive immunology*. 2005;66:175-91.
- [152] Elshal MF, McCoy JP. Multiplex bead array assays: performance evaluation and comparison of sensitivity to ELISA. *Methods*. 2006;38:317-23.
- [153] Martins TB, Burlingame R, von Muhlen CA, Jaskowski TD, Litwin CM, Hill HR. Evaluation of multiplexed fluorescent microsphere immunoassay for detection of autoantibodies to nuclear antigens. *Clinical and diagnostic laboratory immunology*. 2004;11:1054-9.
- [154] Martins TB, Litwin CM, Hill HR. Evaluation of a multiplex fluorescent microsphere immunoassay for the determination of epstein-barr virus serologic status. *American journal of clinical pathology*. 2008;129:34-41.
- [155] Hanly WC, Artwohl JE, Bennett BT. Review of Polyclonal Antibody Production Procedures in Mammals and Poultry. *ILAR journal / National Research Council, Institute of Laboratory Animal Resources*. 1995;37:93-118.
- [156] Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*. 1975;256:495-7.
- [157] Leenaars M, Hendriksen CF. Critical steps in the production of polyclonal and monoclonal antibodies: evaluation and recommendations. *ILAR journal / National Research Council, Institute of Laboratory Animal Resources*. 2005;46:269-79.
- [158] Anderer FA, Strobel G. Recognition of conjugated and native peptide determinants. I. Conformational and sequential specificities of rabbit antibodies versus tobacco mosaic virus. *European journal of immunology*. 1972;2:274-7.
- [159] Schwenk JM, Igel U, Neiman M, Langen H, Becker C, Bjartell A, et al. Toward next generation plasma profiling via heat-induced epitope retrieval and array-based assays. *Molecular & cellular proteomics : MCP*. 2010;9:2497-507.
- [160] Clackson T, Hoogenboom HR, Griffiths AD, Winter G. Making antibody fragments using phage display libraries. *Nature*. 1991;352:624-8.
- [161] McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*. 1990;348:552-4.
- [162] Haab BB. Antibody arrays in cancer research. *Molecular & cellular proteomics : MCP*. 2005;4:377-83.
- [163] Van Regenmortel MH. From absolute to exquisite specificity. Reflections on the fuzzy nature of species, specificity and antigenic sites. *J Immunol Methods*. 1998;216:37-48.

- [164] Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, et al. Towards a knowledge-based Human Protein Atlas. *Nature biotechnology*. 2010;28:1248-50.
- [165] Lindskog M, Rockberg J, Uhlen M, Sterky F. Selection of protein epitopes for antibody production. *BioTechniques*. 2005;38:723-7.
- [166] Berglund L, Andrade J, Odeberg J, Uhlen M. The epitope space of the human proteome. *Protein science : a publication of the Protein Society*. 2008;17:606-13.
- [167] Berglund L, Bjorling E, Jonasson K, Rockberg J, Fagerberg L, Al-Khalili Szigyarto C, et al. A whole-genome bioinformatics approach to selection of antigens for systematic antibody generation. *Proteomics*. 2008;8:2832-9.
- [168] Nilsson P, Paavilainen L, Larsson K, Odling J, Sundberg M, Andersson AC, et al. Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling. *Proteomics*. 2005;5:4327-37.
- [169] Zhu H, Qian J. Applications of functional protein microarrays in basic and clinical research. *Advances in genetics*. 2012;79:123-55.
- [170] Sjoberg R, Sundberg M, Gundberg A, Sivertsson A, Schwenk JM, Uhlen M, et al. Validation of affinity reagents using antigen microarrays. *New biotechnology*. 2012;29:555-63.
- [171] Frank R. The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *J Immunol Methods*. 2002;267:13-26.
- [172] Volkmer R. Synthesis and application of peptide arrays: quo vadis SPOT technology. *Chembiochem : a European journal of chemical biology*. 2009;10:1431-42.
- [173] Nirenberg MW, Matthaei JH. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proceedings of the National Academy of Sciences of the United States of America*. 1961;47:1588-602.
- [174] Carlson ED, Gan R, Hodgman CE, Jewett MC. Cell-free protein synthesis: applications come of age. *Biotechnology advances*. 2012;30:1185-94.
- [175] Endo Y, Sawasaki T. Cell-free expression systems for eukaryotic protein production. *Current opinion in biotechnology*. 2006;17:373-80.
- [176] Spirin ASaS, J.R Cell-Free Protein Synthesis Systems: Historical Landmarks, Classification and general methods. Wiley. 2008:1-34.
- [177] Casado-Vela JG-G, M.; Matarraz, S.; Martinez-Esteso, M.J; Vilella, M.; Sayagues, J.M; Fuentes, M.; Carlos Lacal, J. Protein arrays: recent achievements and their application to study the human proteome. *Curren Proteomics*. 2013;10:83-97.
- [178] He M, Taussig MJ. Single step generation of protein arrays from DNA by cell-free expression and in situ immobilisation (PISA method). *Nucleic acids research*. 2001;29:E73-3.
- [179] Angenendt P, Kreutzberger J, Glokler J, Hoheisel JD. Generation of high density protein microarrays by cell-free in situ expression of unpurified PCR products. *Molecular & cellular proteomics : MCP*. 2006;5:1658-66.

- [180] He M, Stoevesandt O, Palmer EA, Khan F, Ericsson O, Taussig MJ. Printing protein arrays from DNA arrays. *Nature methods*. 2008;5:175-7.
- [181] Stoevesandt O, Vetter M, Kastelic D, Palmer EA, He M, Taussig MJ. Cell free expression put on the spot: advances in repeatable protein arraying from DNA (DAPA). *New biotechnology*. 2011;28:282-90.
- [182] Tao SC, Zhu H. Protein chip fabrication by capture of nascent polypeptides. *Nature biotechnology*. 2006;24:1253-4.
- [183] Ramachandran N, Hainsworth E, Bhullar B, Eisenstein S, Rosen B, Lau AY, et al. Self-assembling protein microarrays. *Science*. 2004;305:86-90.
- [184] Ramachandran N, Raphael JV, Hainsworth E, Demirkan G, Fuentes MG, Rolfs A, et al. Next-generation high-density self-assembling functional protein arrays. *Nature methods*. 2008;5:535-8.
- [185] Ramachandran N, Srivastava S, Labaer J. Applications of protein microarrays for biomarker discovery. *Proteomics Clinical applications*. 2008;2:1444-59.
- [186] Anderson KS, Ramachandran N, Wong J, Raphael JV, Hainsworth E, Demirkan G, et al. Application of protein microarrays for multiplexed detection of antibodies to tumor antigens in breast cancer. *J Proteome Res*. 2008;7:1490-9.
- [187] Anderson KS, Sibani S, Wallstrom G, Qiu J, Mendoza EA, Raphael J, et al. Protein microarray signature of autoantibody biomarkers for the early detection of breast cancer. *J Proteome Res*. 2011;10:85-96.
- [188] Miersch S, Bian X, Wallstrom G, Sibani S, Logvinenko T, Wasserfall CH, et al. Serological autoantibody profiling of type 1 diabetes by protein arrays. *J Proteomics*. 2013;94:486-96.
- [189] Wright C, Sibani S, Trudgian D, Fischer R, Kessler B, LaBaer J, et al. Detection of multiple autoantibodies in patients with ankylosing spondylitis using nucleic acid programmable protein arrays. *Molecular & cellular proteomics : MCP*. 2012;11:M9 00384.
- [190] Gibson DS, Qiu J, Mendoza EA, Barker K, Rooney ME, LaBaer J. Circulating and synovial antibody profiling of juvenile arthritis patients by nucleic acid programmable protein arrays. *Arthritis research & therapy*. 2012;14:R77.
- [191] Speer R, Wulfkühle J, Espina V, Aurajo R, Edmiston KH, Liotta LA, et al. Development of reverse phase protein microarrays for clinical applications and patient-tailored therapy. *Cancer genomics & proteomics*. 2007;4:157-64.
- [192] Paweletz CP, Charboneau L, Bichsel VE, Simone NL, Chen T, Gillespie JW, et al. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene*. 2001;20:1981-9.
- [193] Boyd ZS, Wu QJ, O'Brien C, Spoerke J, Savage H, Fielder PJ, et al. Proteomic analysis of breast cancer molecular subtypes and biomarkers of response to targeted kinase inhibitors using reverse-phase protein microarrays. *Molecular cancer therapeutics*. 2008;7:3695-706.

- [194] Chruscinski AJ, Singh H, Chan SM, Utz PJ. Broad-scale phosphoprotein profiling of beta adrenergic receptor (beta-AR) signaling reveals novel phosphorylation and dephosphorylation events. *PloS one*. 2013;8:e82164.
- [195] Luckert K, Gujral TS, Chan M, Sevecka M, Joos TO, Sorger PK, et al. A dual array-based approach to assess the abundance and posttranslational modification state of signaling proteins. *Science signaling*. 2012;5:pl1.
- [196] Aguilar-Mahecha A, Cantin C, O'Connor-McCourt M, Nantel A, Basik M. Development of reverse phase protein microarrays for the validation of clusterin, a mid-abundant blood biomarker. *Proteome science*. 2009;7:15.
- [197] Janzi M, Odling J, Pan-Hammarstrom Q, Sundberg M, Lundeberg J, Uhlen M, et al. Serum microarrays for large scale screening of protein levels. *Molecular & cellular proteomics : MCP*. 2005;4:1942-7.
- [198] Grazio S, Balen D. [Obesity: risk factor and predictor of osteoarthritis]. *Lijecnicki vjesnik*. 2009;131:22-6.
- [199] Golightly YM, Marshall SW, Kraus VB, Renner JB, Villaveces A, Casteel C, et al. Biomarkers of incident radiographic knee osteoarthritis: do they vary by chronic knee symptoms? *Arthritis Rheum*. 2011;63:2276-83.
- [200] Hunter DJ, Losina E, Guermazi A, Burstein D, Lassere MN, Kraus V. A pathway and approach to biomarker validation and qualification for osteoarthritis clinical trials. *Current drug targets*. 2010;11:536-45.
- [201] Ayoglu B, Haggmark A, Neiman M, Igel U, Uhlen M, Schwenk JM, et al. Systematic antibody and antigen-based proteomic profiling with microarrays. *Expert review of molecular diagnostics*. 2011;11:219-34.
- [202] Fernandez-Puente P, Mateos J, Fernandez-Costa C, Oreiro N, Fernandez-Lopez C, Ruiz-Romero C, et al. Identification of a panel of novel serum osteoarthritis biomarkers. *J Proteome Res*. 2011;10:5095-101.
- [203] Mateos J, Lourido L, Fernandez-Puente P, Calamia V, Fernandez-Lopez C, Oreiro N, et al. Differential protein profiling of synovial fluid from rheumatoid arthritis and osteoarthritis patients using LC-MALDI TOF/TOF. *J Proteomics*. 2012;75:2869-78.
- [204] Lourido L, Calamia V, Mateos J, Fernandez-Puente P, Fernandez-Tajes J, Blanco FJ, et al. Quantitative proteomic profiling of human articular cartilage degradation in osteoarthritis. *J Proteome Res*. 2014;13:6096-106.
- [205] Drobin K, Nilsson P, Schwenk JM. Highly multiplexed antibody suspension bead arrays for plasma protein profiling. *Methods in molecular biology*. 2013;1023:137-45.
- [206] Schwenk JM, Gry M, Rimini R, Uhlen M, Nilsson P. Antibody suspension bead arrays within serum proteomics. *J Proteome Res*. 2008;7:3168-79.

- [207] Sjoberg A, Onnerfjord P, Morgelin M, Heinegard D, Blom AM. The extracellular matrix and inflammation: fibromodulin activates the classical pathway of complement by directly binding C1q. *J Biol Chem*. 2005;280:32301-8.
- [208] de Seny D, Sharif M, Fillet M, Cobraiville G, Meuwis MA, Maree R, et al. Discovery and biochemical characterisation of four novel biomarkers for osteoarthritis. *Ann Rheum Dis*. 2011;70:1144-52.
- [209] Zreiqat H, Belluoccio D, Smith MM, Wilson R, Rowley LA, Jones K, et al. S100A8 and S100A9 in experimental osteoarthritis. *Arthritis research & therapy*. 2010;12:R16.
- [210] Donato R, Cannon BR, Sorci G, Riuzzi F, Hsu K, Weber DJ, et al. Functions of S100 proteins. *Current molecular medicine*. 2013;13:24-57.
- [211] Leclerc E, Fritz G, Weibel M, Heizmann CW, Galichet A. S100B and S100A6 differentially modulate cell survival by interacting with distinct RAGE (receptor for advanced glycation end products) immunoglobulin domains. *J Biol Chem*. 2007;282:31317-31.
- [212] Yoshihara Y, Plaas A, Osborn B, Margulis A, Nelson F, Stewart M, et al. Superficial zone chondrocytes in normal and osteoarthritic human articular cartilages synthesize novel truncated forms of inter-alpha-trypsin inhibitor heavy chains which are attached to a chondroitin sulfate proteoglycan other than bikunin. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society*. 2008;16:1343-55.
- [213] Zhao M, Yoneda M, Ohashi Y, Kurono S, Iwata H, Ohnuki Y, et al. Evidence for the covalent binding of SHAP, heavy chains of inter-alpha-trypsin inhibitor, to hyaluronan. *J Biol Chem*. 1995;270:26657-63.
- [214] De Ceuninck F, Marcheteau E, Berger S, Caliez A, Dumont V, Raes M, et al. Assessment of some tools for the characterization of the human osteoarthritic cartilage proteome. *Journal of biomolecular techniques : JBT*. 2005;16:256-65.
- [215] Yan X, Zhao Y, Pan J, Fang K, Wang Y, Li Z, et al. Vitamin D-binding protein (group-specific component) has decreased expression in rheumatoid arthritis. *Clinical and experimental rheumatology*. 2012;30:525-33.
- [216] Sowers MR, Karvonen-Gutierrez CA. The evolving role of obesity in knee osteoarthritis. *Current opinion in rheumatology*. 2010;22:533-7.
- [217] Erb N, Pace AV, Douglas KM, Banks MJ, Kitas GD. Risk assessment for coronary heart disease in rheumatoid arthritis and osteoarthritis. *Scandinavian journal of rheumatology*. 2004;33:293-9.
- [218] Singh G, Miller JD, Lee FH, Pettitt D, Russell MW. Prevalence of cardiovascular disease risk factors among US adults with self-reported osteoarthritis: data from the Third National Health and Nutrition Examination Survey. *The American journal of managed care*. 2002;8:S383-91.

- [219] Zamanian-Daryoush M, Lindner D, Tallant TC, Wang Z, Buffa J, Klipfell E, et al. The cardioprotective protein apolipoprotein A1 promotes potent anti-tumorigenic effects. *J Biol Chem*. 2013;288:21237-52.
- [220] de Seny D, Cobraiville G, Charlier E, Neuville S, Lutteri L, Le Goff C, et al. Apolipoprotein-A1 as a damage-associated molecular patterns protein in osteoarthritis: ex vivo and in vitro pro-inflammatory properties. *PLoS one*. 2015;10:e0122904.
- [221] Tsezou A, Iliopoulos D, Malizos KN, Simopoulou T. Impaired expression of genes regulating cholesterol efflux in human osteoarthritic chondrocytes. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*. 2010;28:1033-9.
- [222] Oliviero F, Sfriso P, Baldo G, Dayer JM, Giunco S, Scanu A, et al. Apolipoprotein A-I and cholesterol in synovial fluid of patients with rheumatoid arthritis, psoriatic arthritis and osteoarthritis. *Clinical and experimental rheumatology*. 2009;27:79-83.
- [223] Conde J, Gomez R, Bianco G, Scotece M, Lear P, Dieguez C, et al. Expanding the adipokine network in cartilage: identification and regulation of novel factors in human and murine chondrocytes. *Ann Rheum Dis*. 2011;70:551-9.
- [224] Dumond H, Presle N, Terlain B, Mainard D, Loeuille D, Netter P, et al. Evidence for a key role of leptin in osteoarthritis. *Arthritis Rheum*. 2003;48:3118-29.
- [225] Ku JH, Lee CK, Joo BS, An BM, Choi SH, Wang TH, et al. Correlation of synovial fluid leptin concentrations with the severity of osteoarthritis. *Clinical rheumatology*. 2009;28:1431-5.
- [226] Simopoulou T, Malizos KN, Iliopoulos D, Stefanou N, Papatheodorou L, Ioannou M, et al. Differential expression of leptin and leptin's receptor isoform (Ob-Rb) mRNA between advanced and minimally affected osteoarthritic cartilage; effect on cartilage metabolism. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society*. 2007;15:872-83.
- [227] Verma P, Dalal K. Serum cartilage oligomeric matrix protein (COMP) in knee osteoarthritis: a novel diagnostic and prognostic biomarker. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*. 2013;31:999-1006.
- [228] Ruan MZ, Erez A, Guse K, Dawson B, Bertin T, Chen Y, et al. Proteoglycan 4 expression protects against the development of osteoarthritis. *Science translational medicine*. 2013;5:176ra34.
- [229] Lorenzo P, Bayliss MT, Heinegard D. Altered patterns and synthesis of extracellular matrix macromolecules in early osteoarthritis. *Matrix biology : journal of the International Society for Matrix Biology*. 2004;23:381-91.
- [230] Conrozier T, Carlier MC, Mathieu P, Colson F, Debard AL, Richard S, et al. Serum levels of YKL-40 and C reactive protein in patients with hip osteoarthritis and healthy subjects: a cross sectional study. *Ann Rheum Dis*. 2000;59:828-31.
- [231] Luan X, Lu Q, Jiang Y, Zhang S, Wang Q, Yuan H, et al. Crystal structure of human RANKL complexed with its decoy receptor osteoprotegerin. *Journal of immunology*. 2012;189:245-52.

- [232] Kwan Tat S, Amiabile N, Pelletier JP, Boileau C, Lajeunesse D, Duval N, et al. Modulation of OPG, RANK and RANKL by human chondrocytes and their implication during osteoarthritis. *Rheumatology*. 2009;48:1482-90.
- [233] Valdes AM, Hart DJ, Jones KA, Surdulescu G, Swarbrick P, Doyle DV, et al. Association study of candidate genes for the prevalence and progression of knee osteoarthritis. *Arthritis Rheum*. 2004;50:2497-507.
- [234] Skoumal M, Kolarz G, Haberhauer G, Woloszczuk W, Hawa G, Klingler A. Osteoprotegerin and the receptor activator of NF-kappa B ligand in the serum and synovial fluid. A comparison of patients with longstanding rheumatoid arthritis and osteoarthritis. *Rheumatol Int*. 2005;26:63-9.
- [235] Pilichou A, Papassotiriou I, Michalakakou K, Fessatou S, Fandridis E, Papachristou G, et al. High levels of synovial fluid osteoprotegerin (OPG) and increased serum ratio of receptor activator of nuclear factor-kappa B ligand (RANKL) to OPG correlate with disease severity in patients with primary knee osteoarthritis. *Clinical biochemistry*. 2008;41:746-9.
- [236] Gillan L, Matei D, Fishman DA, Gerbin CS, Karlan BY, Chang DD. Periostin secreted by epithelial ovarian carcinoma is a ligand for alpha(V)beta(3) and alpha(V)beta(5) integrins and promotes cell motility. *Cancer research*. 2002;62:5358-64.
- [237] Loeser RF, Olex AL, McNulty MA, Carlson CS, Callahan MF, Ferguson CM, et al. Microarray analysis reveals age-related differences in gene expression during the development of osteoarthritis in mice. *Arthritis Rheum*. 2012;64:705-17.
- [238] Chou CH, Wu CC, Song IW, Chuang HP, Lu LS, Chang JH, et al. Genome-wide expression profiles of subchondral bone in osteoarthritis. *Arthritis research & therapy*. 2013;15:R190.
- [239] Yamagiwa H, Sarkar G, Charlesworth MC, McCormick DJ, Bolander ME. Two-dimensional gel electrophoresis of synovial fluid: method for detecting candidate protein markers for osteoarthritis. *Journal of orthopaedic science : official journal of the Japanese Orthopaedic Association*. 2003;8:482-90.
- [240] Echan LA, Tang HY, Ali-Khan N, Lee K, Speicher DW. Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics*. 2005;5:3292-303.
- [241] Gobezie R, Kho A, Krastins B, Sarracino DA, Thornhill TS, Chase M, et al. High abundance synovial fluid proteome: distinct profiles in health and osteoarthritis. *Arthritis research & therapy*. 2007;9:R36.
- [242] Zegels G, Van Raemdonck GA, Coen EP, Tjalma WA, Van Ostade XW. Comprehensive proteomic analysis of human cervical-vaginal fluid using colposcopy samples. *Proteome science*. 2009;7:17.

- [243] Andrews G, Lewis D, Notey J, Kelly R, Muddiman D. Part II: defining and quantifying individual and co-cultured intracellular proteomes of two thermophilic microorganisms by GeLC-MS2 and spectral counting. *Analytical and bioanalytical chemistry*. 2010;398:391-404.
- [244] Huang Y, Houston NL, Tovar-Mendez A, Stevenson SE, Miernyk JA, Randall DD, et al. A quantitative mass spectrometry-based approach for identifying protein kinase clients and quantifying kinase activity. *Anal Biochem*. 2010;402:69-76.
- [245] Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, et al. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular & cellular proteomics : MCP*. 2005;4:1265-72.
- [246] Drynda S, Ringel B, Kekow M, Kuhne C, Drynda A, Glocker MO, et al. Proteome analysis reveals disease-associated marker proteins to differentiate RA patients from other inflammatory joint diseases with the potential to monitor anti-TNFalpha therapy. *Pathology, research and practice*. 2004;200:165-71.
- [247] Uchida T, Fukawa A, Uchida M, Fujita K, Saito K. Application of a novel protein biochip technology for detection and identification of rheumatoid arthritis biomarkers in synovial fluid. *J Proteome Res*. 2002;1:495-9.
- [248] Baillet A, Trocme C, Berthier S, Arlotto M, Grange L, Chenau J, et al. Synovial fluid proteomic fingerprint: S100A8, S100A9 and S100A12 proteins discriminate rheumatoid arthritis from other inflammatory joint diseases. *Rheumatology*. 2010;49:671-82.
- [249] Pillinger MH, Abramson SB. The neutrophil in rheumatoid arthritis. *Rheumatic diseases clinics of North America*. 1995;21:691-714.
- [250] Desrochers PE, Jeffrey JJ, Weiss SJ. Interstitial collagenase (matrix metalloproteinase-1) expresses serpinase activity. *The Journal of clinical investigation*. 1991;87:2258-65.
- [251] Rumbaugh J, Turchan-Cholewo J, Galey D, St Hillaire C, Anderson C, Conant K, et al. Interaction of HIV Tat and matrix metalloproteinase in HIV neuropathogenesis: a new host defense mechanism. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*. 2006;20:1736-8.
- [252] Munier FL, Frueh BE, Othenin-Girard P, Uffer S, Cousin P, Wang MX, et al. BIGH3 mutation spectrum in corneal dystrophies. *Investigative ophthalmology & visual science*. 2002;43:949-54.
- [253] Calamia V, Rocha B, Mateos J, Fernandez-Puente P, Ruiz-Romero C, Blanco FJ. Metabolic labeling of chondrocytes for the quantitative analysis of the interleukin-1-beta-mediated modulation of their intracellular and extracellular proteomes. *J Proteome Res*. 2011;10:3701-11.
- [254] Di Cesare PE, Chen FS, Moergelin M, Carlson CS, Leslie MP, Perris R, et al. Matrix-matrix interaction of cartilage oligomeric matrix protein and fibronectin. *Matrix biology : journal of the International Society for Matrix Biology*. 2002;21:461-70.

- [255] Zack MD, Arner EC, Anglin CP, Alston JT, Malfait AM, Tortorella MD. Identification of fibronectin neoepitopes present in human osteoarthritic cartilage. *Arthritis Rheum.* 2006;54:2912-22.
- [256] Sofat N. Analysing the role of endogenous matrix molecules in the development of osteoarthritis. *International journal of experimental pathology.* 2009;90:463-79.
- [257] Osborn TM, Verdrengh M, Stossel TP, Tarkowski A, Bokarewa M. Decreased levels of the gelsolin plasma isoform in patients with rheumatoid arthritis. *Arthritis research & therapy.* 2008;10:R117.
- [258] Sakkas LI, Johanson NA, Scanzello CR, Platsoucas CD. Interleukin-12 is expressed by infiltrating macrophages and synovial lining cells in rheumatoid arthritis and osteoarthritis. *Cell Immunol.* 1998;188:105-10.
- [259] Srivastava R, Aslam M, Kalluri SR, Schirmer L, Buck D, Tackenberg B, et al. Potassium channel KIR4.1 as an immune target in multiple sclerosis. *The New England journal of medicine.* 2012;367:115-23.
- [260] Beeton C, Wulff H, Standifer NE, Azam P, Mullen KM, Pennington MW, et al. Kv1.3 channels are a therapeutic target for T cell-mediated autoimmune diseases. *Proceedings of the National Academy of Sciences of the United States of America.* 2006;103:17414-9.
- [261] Boissier MC, Semerano L, Challal S, Saidenberg-Kermanac'h N, Falgarone G. Rheumatoid arthritis: from autoimmunity to synovitis and joint destruction. *J Autoimmun.* 2012;39:222-8.
- [262] Semerano L, Thiolat A, Minichiello E, Clavel G, Bessis N, Boissier MC. Targeting IL-6 for the treatment of rheumatoid arthritis: Phase II investigational drugs. *Expert Opin Investig Drugs.* 2014;23:979-99.

VIII. ANEXOS

Anexo I. Material suplementario del capítulo I

Figura Suplementaria 1. Alteraciones en el secretoma de los explantes de cartílago N, UZOA y WZOA. Tinción de plata de un gel SDS-PAGE mostrando las diferencias en los perfiles de los secretomas. Las proteínas representativas del secretoma fueron identificadas por espectrometría de masas.

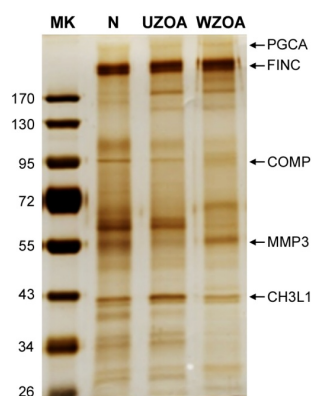


Figura Suplementaria 2. Secreción diferencial desde los explantes de cartílago N, UZOA y WZOA. La concentración proteica fue medida en el secretoma de 10 cartílagos N, 10 UZOA y 10 WZOA empleados en este trabajo. Los resultados están expresados como la media \pm SD (*, $p < 0.05$; **, $p < 0.005$).

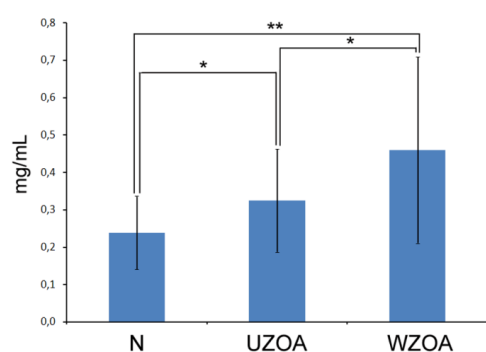


Figura Suplementaria 3. Cálculo del *cutoff* para determinar los ratios significativos en los experimentos iTRAQ. Para determinar el *cutoff* se utilizó el % de variación de las proteínas comunes entre las 4 réplicas biológicas. El eje vertical representa la frecuencia acumulada, mientras que el horizontal representa el porcentaje de variación de los experimentos iTRAQ. El 90% de las proteínas se encuentran dentro del 80% de la variación experimental. Así, consideramos un ratio > 1.8 o < 0.55 para definir las diferencias significativas entre replicas biológicas.

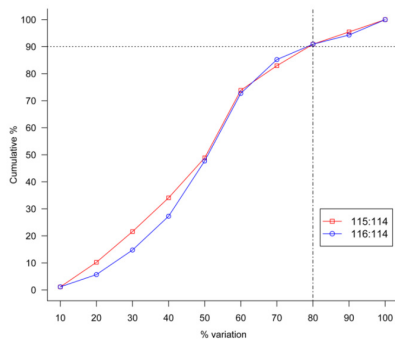


Figura Suplementaria 4. Análisis de *clustering K-means* de las proteínas diferencialmente secretadas. Los datos muestran el \log_2 de los ratios entre las proteínas liberadas por las zonas macroscópicamente no lesionadas y de las zonas lesionadas del cartílago OA (UZOA y WZOA, respectivamente) en comparación con el cartílago normal (N).

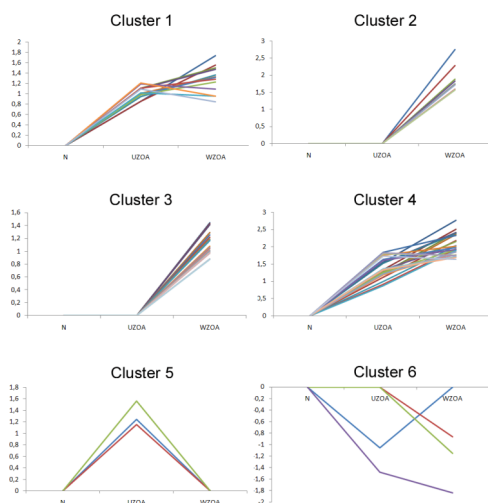


Figura Suplementaria 5. Perfil funcional de las proteínas diferencialmente secretadas por el cartílago artrósico. El análisis de las vías de señalización usando el programa String muestra que los tres grupos funcionales en los cuales se agrupan las proteínas diferencialmente secretadas son: la organización y el recambio de la ECM, la respuesta inmune y la inflamación.

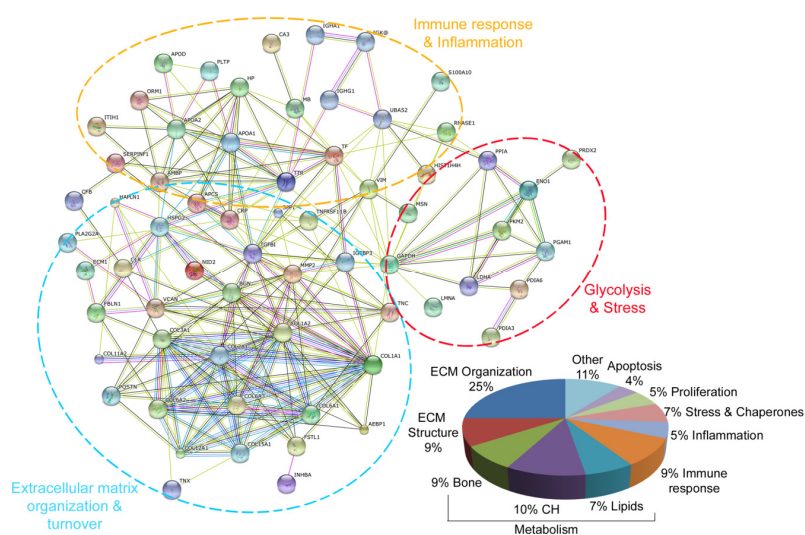


Tabla Suplementaria 1. Lista de las proteínas identificadas y cuantificadas en los experimentos iTRAQ. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 2. Lista de las proteínas identificadas con alta confianza. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 3. Análisis de enriquecimiento basado en GO de las proteínas identificadas en este trabajo. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 4. Análisis de enriquecimiento de las proteínas que pertenecen a cada uno de los 6 *clusters* identificados. Incluida en el DVD adjunto a esta tesis.

Anexo II. Material suplementario del capítulo II

Tabla Suplementaria 1. Lista de las 122 proteínas identificadas en el líquido sinovial de pacientes con artritis reumatoide. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 2. Lista de las 107 proteínas identificadas en el líquido sinovial de pacientes con artrosis. Incluida en el DVD adjunto a esta tesis.

Anexo III. Material suplementario capítulo III

Figura Suplementaria 1. Flujo de trabajo del NAPPA.

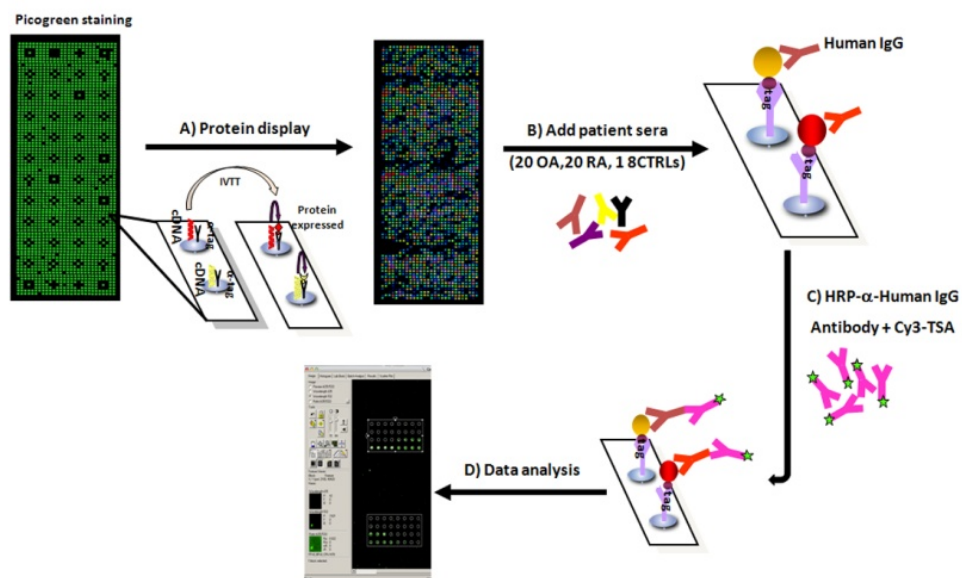


Figura Suplementaria 2. Distribución de las reactividades en los *arrays* y sexo (B).

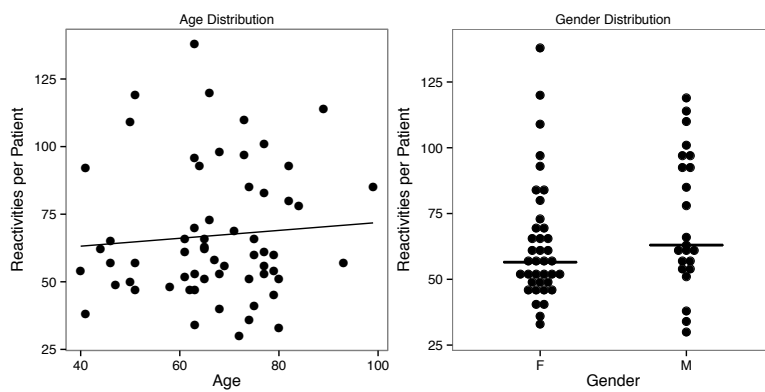


Figura Suplementaria 3. A: Perfiles de la reactividad de las muestras en los *arrays* en formato esfera con de 373 fragmentos antigénicos. La línea de puntos representa los valores de diferentes *cutoff* utilizados para definir la reactividad, 10x MAD (naranja) y 50x MAD (rojo). **B:** El gráfico representa el número de muestras que presentan reactividad hacia cierto antígeno. **C:** Comparaciones entre grupos según el número de antígenos reactivos por muestra, dependiendo del *cutoff* medio + 10x/50x MAD.

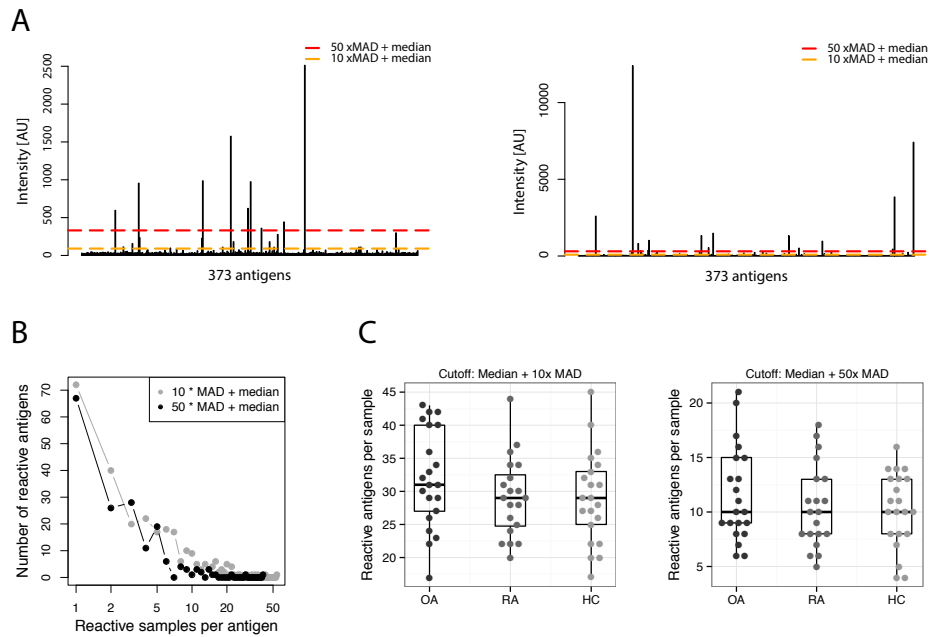


Figura Suplementaria 4. Resumen del porcentaje de reactividades por grupo de estudio; artrosis (OA), artritis reumatoide (RA) y controles sanos (HC) de los antígenos seleccionados en el array de fragmentos antigénicos (A) y el NAPPA (B). Los números que están encima de las barras representan el número de muestras reactivas para ese antígeno.

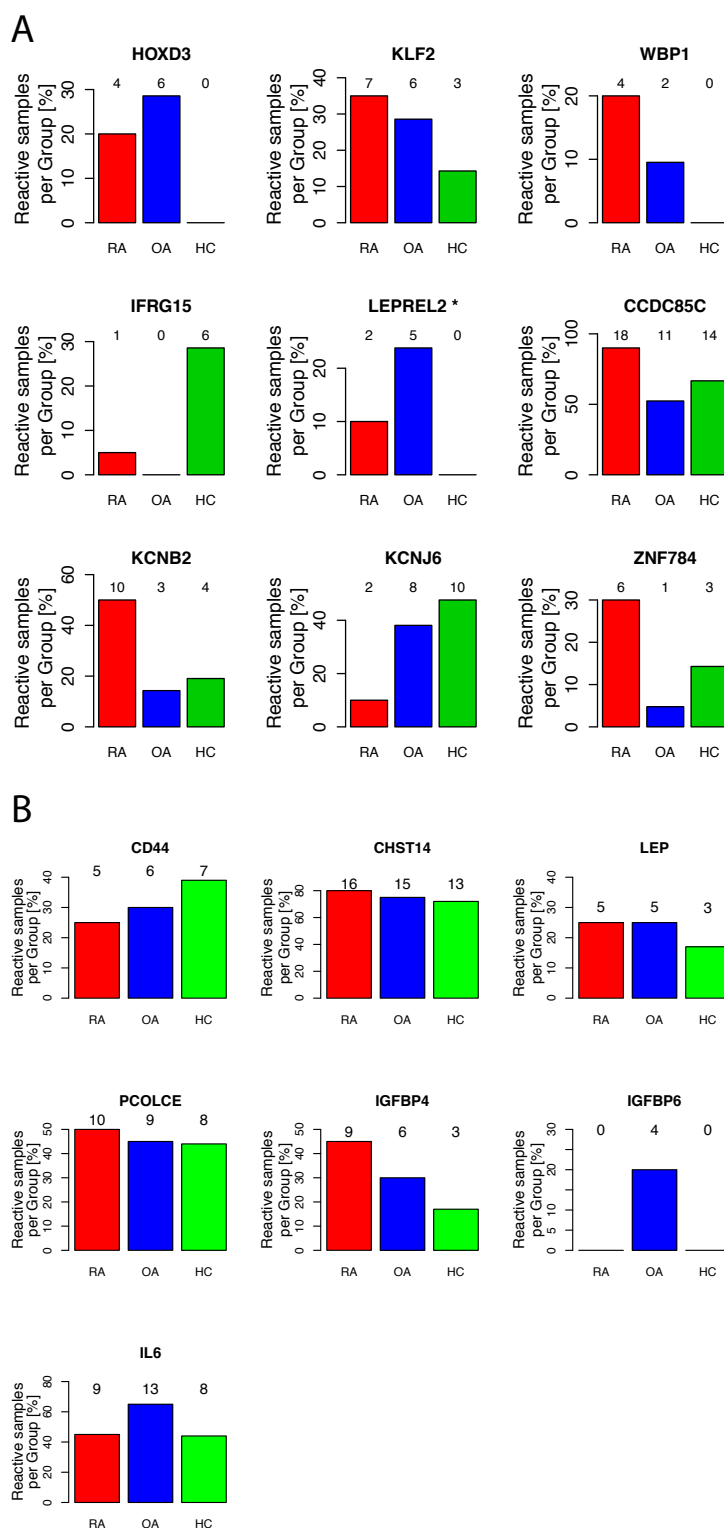


Tabla Suplementaria 1. Lista de los 80 clones incluidos en el NAPPA. Incluida en el DVD adjunto a esta tesis.

Anexo IV. Material suplementario del capítulo IV

Figura Suplementaria 1. Diagrama de Venn para mostrar el número de proteínas significativamente diferentes ($p < 0.05$) entre las diferentes comparaciones, OA y Controles (Ctrl), OA comparado a pacientes con artritis psoriásica (PsA) y OA comparado a pacientes con artritis reumatoide (RA) en la fase de cribado. En total, se identificaron 46 proteínas que distinguían estas 3 comparaciones.

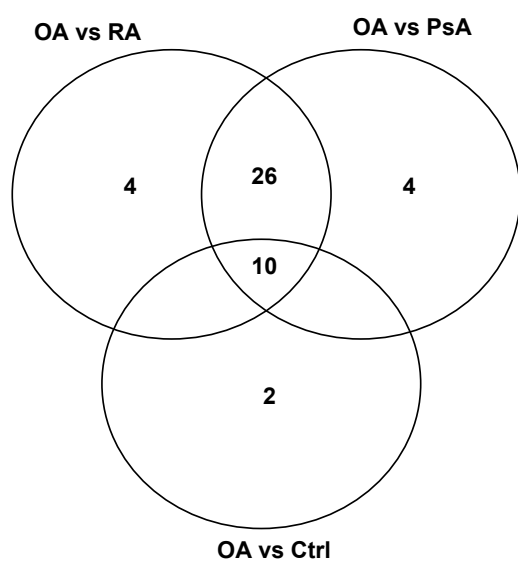


Figura Suplementaria 2. Distribución de la variación intra-ensayo y los ensayos de cribado y verificación. El gráfico muestra el coeficiente de variación (C.V) en las réplicas técnicas. La calidad técnica de los ensayos fue testada por el CV intraensayo usando los valores MFI de las muestras de suero replicadas a lo largo de todos los ensayos. El C.V de todos los anticuerpos entre las réplicas en el cribado con las 960 muestras y en la verificación con las 672 muestras fue menor del 20%.

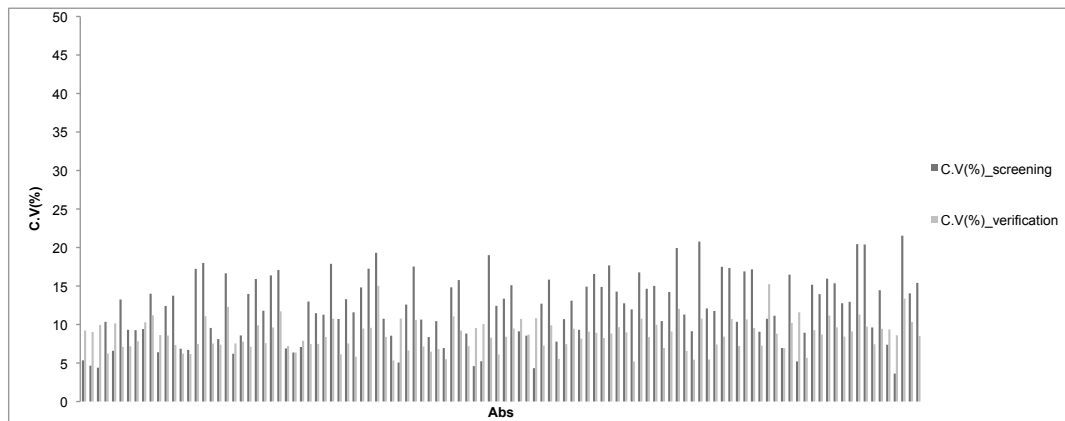


Figura Suplementaria 3. Diagrama de Venn para mostrar el número de proteínas significativamente diferentes ($p < 0.05$) entre las diferentes comparaciones, OA y Controles (Ctrl), OA comparado a pacientes con artritis psoriásica (PsA) y OA comparado a pacientes con artritis reumatoide (RA) en la fase de verificación. Los niveles de 16 proteínas permitieron diferenciar a los pacientes OA de los RA y de los PsA. Una de estas proteínas, también permitía distinguir entre los pacientes OA y los Controles.

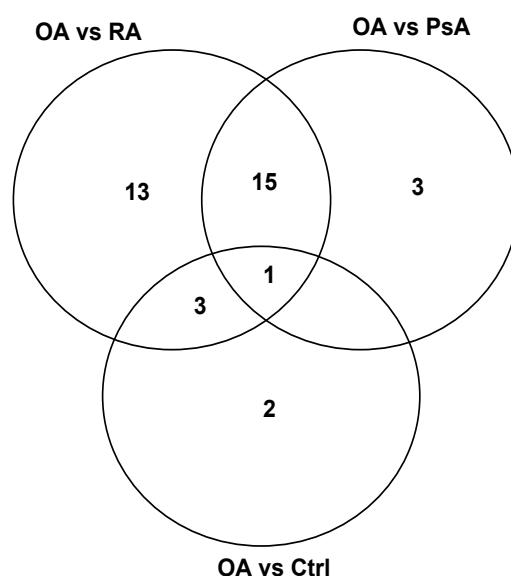


Tabla Suplementaria 1. Lista de los anticuerpos del HPA usados para generar el array en suspensión utilizado en la fase de cribado. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 2. Lista de los anticuerpos del HPA usados para generar el array en suspensión utilizado en la fase de cribado. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 3. Esta tabla resume el total de proteínas cuyos niveles en suero diferencian significativamente ($p < 0.05$) a los pacientes artrósicos de los pacientes con artritis reumatoide (RA) en la fase de cribado, mostrando los pvalues obtenidos para cada anticuerpo utilizado para detectar la proteína concreta. Las proteínas marcadas en negro están también moduladas en la fase de verificación. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 4. Esta tabla resume el total de proteínas cuyos niveles en suero diferencian significativamente ($p < 0.05$) a los pacientes artrósicos de los pacientes con artritis reumatoide (RA) en la fase de cribado, mostrando los p-values obtenidos para cada anticuerpo utilizado para detectar la proteína concreta. Las proteínas marcadas en negro están también moduladas en la fase de verificación. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 5. Tabla mostrando el área bajo la curva ROC (AUC) y el intervalo de confianza al 95% de cada proteína del panel de seis proteínas que permite distinguir pacientes OA y controles sanos en las muestras analizadas en la fase de cribado y en la fase de verificación. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 6A. Proteínas que significativamente ($p < 0.05$) distinguen a los pacientes OA de los RA en la fase de cribado y verificación. Incluida en el DVD adjunto a esta tesis.

Tabla Suplementaria 6B. Proteínas que significativamente ($p < 0.05$) distinguen a los pacientes OA de los PsA en la fase de cribado y verificación. Incluida en el DVD adjunto a esta tesis.

Anexo V: Relación de artículos no relacionados directamente con esta tesis en los que he participado durante mi etapa de estudiante doctoral:

1. **Lourido L.**, Diez P., Dasilva N., Gonzalez-Gonzalez M., Ruiz-Romero C., Blanco F.J, Orfao A., LaBaer J., and Fuentes M. *Protein Microarrays: Overview, Applications and Challenges Genomics and Proteomics for Clinical Discovery and Development*, Translational Bioinformatics 6, DOI 10.1007/978-94-017-9202-8_8
2. **Lourido L.**, Calamia V., Fernández-Puente P., Mateos J., Oreiro N, Blanco FJ., Ruiz-Romero C. *Secretome analysis of human articular chondrocytes unravels catabolic effects of nicotine on the joint*. Proteomics Clin Appl. 2015 Apr 25. doi: 10.1002/prca.201400186.
3. Calamia V., Mateos J., Fernández-Puente P., **Lourido L.**, Rocha B., Fernández-Costa C., Montell E., Vergés J., Ruiz-Romero C., Blanco FJ. *A pharmacoproteomic study confirms the synergistic effect of chondroitin sulfate and glucosamine*. Sci Rep. 2014 Jun 10;4:5069. doi: 10.1038/srep05069.
4. Calamia V., **Lourido L.**, Fernández-Puente P., Mateos J., Rocha B., Montell E., Vergés J., Ruiz-Romero C., Blanco FJ. *Secretome analysis of chondroitin sulfate-treated chondrocytes reveals anti-angiogenic, anti-inflammatory and anti-catabolic properties*. Arthritis Res Ther. 2012 Oct 2;14(5):R202. doi: 10.1186/ar4040.
5. Calamia V., Fernández-Puente P., Mateos J., **Lourido L.**, Rocha B., Montell E., Vergés J., Ruiz-Romero C., Blanco FJ. *Pharmacoproteomic study of three different chondroitin sulfate compounds on intracellular and extracellular human chondrocyte proteomes*. Mol Cell Proteomics. 2012 Jun;11(6):M111.013417. doi: 10.1074/mcp.M111.013417.

