



**Universidad de Santiago de Compostela**  
Facultad de Medicina y Odontología  
Departamento de Medicina

**BÚSQUEDA DE FACTORES  
GENÉTICOS DE SUSCEPTIBILIDAD A LA  
ARTROSIS:  
fenotipos extremos, variantes raras,  
microsatélites y metaanálisis de genes  
candidatos.**

**Tesis doctoral**  
**María Cristina Rodríguez Fontenla**  
**2013**



El Dr. Antonio González Martínez-Pedrayo, investigador principal del Laboratorio de Investigación 10 del Instituto de Investigación Sanitaria del Complejo Hospitalario Universitario de Santiago de Compostela, y el Prof. Dr. Juan Jesús Gómez-Reino Carnota, Jefe del Servicio de Reumatología del Hospital Clínico Universitario de Santiago de Compostela y Catedrático de Medicina del Departamento de Medicina de la Universidad de Santiago de Compostela,

CERTIFICAN QUE:

El presente trabajo que lleva por título: **Búsqueda de factores genéticos de susceptibilidad a la artrosis: fenotipos extremos, variantes raras, microsatélites y metaanálisis de genes candidatos**, realizado por María Cristina Rodríguez Fontenla en el Departamento de Medicina de la Universidad de Santiago de Compostela bajo nuestra dirección, ha sido revisado y está en disposición de ser presentado para optar al grado de Doctora en Biología.

Fdo: Dr. Antonio González Martínez-Pedrayo Fdo: Prof. Dr. Juan J. Gómez-Reino Carnota





*El amor por todas las criaturas vivientes*

*es el más noble atributo del hombre.*

*Charles Darwin (1809-1882)*





## AGRADECIMIENTOS

Me puedo imaginar que todas las personas que han comenzado una tesis y alcanzado el final se les hace difícil creer que este momento ha llegado, el de encontrarse escribiendo estas líneas, las últimas de un largo y duro camino, con mejores y peores momentos.

En primer lugar, quiero darle las gracias a mi director de tesis, el Dr Antonio González por haberme dado la oportunidad de formar parte de su laboratorio y ayudarme a realizar este trabajo. Gracias también al Dr Juan Gómez-Reino, mi codirector de tesis, por su apoyo y por permitirme realizarla en colaboración con el Servicio de Reumatología del Hospital Clínico Universitario de Santiago de Compostela.

También agradezco la cooperación de los pacientes artrosis y de los controles así como la colaboración de los médicos e investigadores, tanto nacionales como extranjeros, que nos han brindado la oportunidad de trabajar con sus muestras de ADN y nos han permitido establecer colaboraciones con sus grupos. En especial, al TreatOA por proporcionarnos gran cantidad de datos y ayudarme a conocer la dinámica de trabajo a gran escala en un consorcio integrado por numerosos investigadores.

Es inevitable ponerse algo melancólica y sensible después de casi seis años junto a magníficas personas que no sólo han sido los mejores compañeros de trabajo que me podía encontrar si no también estupendos amigos con los que he compartido lloros, alegrías, viajes y congresos, protocolos, momentos frikis biológicos y matemáticos, inquietudes, chistes... Aún recuerdo el primer día que llegué al laboratorio y si no me podía imaginar lo que iba a aprender ni los proyectos de investigación que iba a comenzar tampoco se me pasaban por la cabeza las valiosas personas que iba a conocer. En el fondo, creo que ha sido un viaje bastante comparable al que se emprende al poner rumbo a un país desconocido. Sabes que vas a ir a un lugar cercano o lejano pero no conoces lo que te encontrarás por el camino, intuyes que vivirás experiencias positivas y negativas y seguro que sufrirás algún problema aunque pienses que lo tienes todo planificado, pero no sabes con que personas o paisajes tropezarán contigo cuando lo atraveses. A medida que pasan los días en este nuevo país es fácil encariñarse con todo lo que te rodea de tal manera que cuando llega el momento de marchar, de alguna manera te encuentras diferente y comienzas a echar de menos sus calles, gentes y ambiente aunque el avión no haya despegado. Ese es justo el momento exacto de escribir estos agradecimientos y pensar en todos los compañeros que me encontré en este largo trayecto.

A Julio, por enseñarme tanto de artrosis y genética, ser mi profesor durante mis primeros meses en el laboratorio y por su maravillosa filosofía. A Chicha por ser la primera técnico del laboratorio que conocí, tan simpática, bromista y extremadamente profesional. A Rebe, porque te echo muchísimo de menos, por todo lo que me has enseñado, por tus lecciones de vida, tus piropos y por todos nuestros momentos micromamífero que no fueron pocos. A Pombo, por ser un compañero tan genial y divertido siempre lleno de ocurrencias con el que todos nos hemos reído mucho. A Marián, por ser tan buena amiga, haberme escuchado siempre que lo necesité, transmitirme tranquilidad y equilibrio y porque el viaje a New York nunca lo olvidaremos. A Carmen por ser siempre de gran ayuda, ser una técnico muy competente y por soportarnos a todos durante muchos años, lo cuál no debe ser fácil. A Elisa, por compartir tantísimas cosas, ser amiga, compañera de conservatorio, carrera y de tesis, por tener “viajitis” como yo y por no dudar en unirse a los destinos más locos aunque hicieran falta bastantes barcos para llegar. A Ana, por ser genial, valiente y única, por tu capacidad de salir adelante y por mantener tu amistad con nosotros a pesar de no estar ya en el laboratorio. No hay nadie como tú y por eso todos te tenemos un cariño especial. A Ariana, porque conectamos desde el primer momento, por escucharme siempre, enseñarme tanto, ser tan humilde a pesar de todos tus conocimientos y porque me encanta tu carácter. A May, porque desde aquel congreso OARSI en Bruselas ya formaste parte de nuestro indestructible y fantástico equipo OA (¡menudo equipazo!), por mostrarme que no siempre todo es fácil pero quien la sigue la consigue y por todas las risas que nos echamos juntas. A Aida, porque te conocí en la carrera en aquellas míticas clases de antropología molecular que difícilmente se borran de la mente y porque compartimos preocupaciones y aficiones. A Marta y a Mari porque a pesar de estar tan poco tiempo en el laboratorio siempre se os recuerda con cariño. A Nacho, por ser tan profesional y buen compañero, dar tantos ánimos y ayudar a que nada parezca imposible. A Romi, porque eres increíble, ordenada, muy capaz e inteligente y porque tu fantástica y única risa nunca se olvida. A Laura, por permitirme enseñarte algo de lo que aprendí estos años, por ser la última en entrar en el laboratorio y por tanto la “nueva”, lo cuál no siempre es fácil.

He dejado para el final, de todas las personas que han pasado por el laboratorio a un compañero y amigo al que tengo que dar un especial agradecimiento ya que sin su ayuda difícilmente podría haber llegado hasta aquí. Manolo, no puedo describir con palabras todo lo que me has enseñado y ayudado durante estos años, eres una persona muy valiosa y profesional. Te doy las gracias por enfrentarte a nuevos retos con determinación e iniciativa, tener tantísimos conocimientos, no rendirte nunca ni perder la ilusión y trabajar conmigo siempre que lo necesité. Tengo la profunda convicción de que llegarás muy lejos.

Tengo que mencionar a muchos de los compañeros que he conocido en otros laboratorios. A los “cardíacos”: Ana, Diego, Vane y Sandra por hacerme más fácil el día a día con vuestras bromas y alegrías, porque siempre sentimos que formáis parte del laboratorio 10 a pesar de estar separados por un muro. Hacéis

un gran equipo y demostráis que se puede trabajar muy duro sin perder la sonrisa. Al laboratorio 8; a Bea por saber tanto, ser tan tenaz y “nice” y darme tanto apoyo y a Paula por su simpatía y alegría. A Vane (laboratorio 9) por ser una excelente y dulce persona. A Isa porque a pesar de conocernos hace poco eres una excelente persona que siempre sabe escuchar y ser la mejor compañía en las comidas.

A todos los compañeros de los laboratorios del IDIS a los que he conocido durante estos años, muchos ya se han ido pero otros siguen ahí: Rudy, Bruno, Rocío, Toño que ya estáis en el extranjero. A Javi, Morena, Anna, Giuseppe (laboratorio 9); Miguel, Marina, María (laboratorio 11); a Marta; a Turo, Lucía (laboratorio de endocrinología); a Marga y María (laboratorio de digestivo); a Sonia y Ezequiel. A todos vosotros gracias por estar siempre allí y escucharme. También quiero agradecer a Oreste, Fran y Carmen por ayudarme siempre que lo necesité.

A mi familia por apoyarme y empujarme hacia delante todos estos años. Si a veces este camino ha sido arduo y difícil para mi, sé que mucho más para vosotros. Gracias por ayudarme a luchar por lo que quiero conseguir, ofrecerme consuelo, lograr que sea feliz y que nunca me rinda para poder trabajar en lo que me gusta de verdad y por animarme a perseguir mis sueños por muy difíciles que éstos puedan parecer. Siempre he encontrado en vosotros la misma sensación que el príncipe Andrei en Guerra y Paz cuando buscaba y encontraba de nuevo aquel roble perdido en el bosque y al verlo brotar “le inundaba un sentimiento de alegría y renovación”, valorando su vida positivamente a pesar de los baches sufridos y sintiendo un reverdecer de esperanzas para el futuro.

Si estos agradecimientos los hubiera escrito con pluma y papel como se hacía hace cientos de años, estas líneas estarían bastante borrosas ya que se me han escapado unas cuántas lágrimas.

*A todos vosotros gracias.*





***ÍNDICE***

---





# ÍNDICE

<b>ABREVIATURAS</b> .....	3
<b>INTRODUCCIÓN</b> .....	9
1. <i>La artrosis(OA)</i> .....	9
1.1 Diagnóstico y clasificación .....	10
1.2 Prevalencia e incidencia .....	12
1.3 Factores de riesgo .....	13
2. <i>El cartilago articular</i> .....	16
2.1 Función y estructura del cartilago articular.....	16
2.2 El condrocito .....	19
2.2.1 El proceso de condrogénesis y la formación de la articulación .....	19
2.2.2 El papel del condrocito en la OA .....	22
2.3 La matriz extracelular (ECM) .....	22
2.3.1 El colágeno .....	23
2.3.2 Proteoglicanos .....	24
2.3.3 COMP .....	26
2.3.4 Mecanismos de formación y degradación de la ECM.....	26
2.3.4.1 Biosíntesis de la ECM .....	27
2.3.4.1.1 <i>Superfamilia TGF-<math>\beta</math> (Transforming Growth Factor <math>\beta</math>)</i> .....	27
2.3.4.1.2 <i>IGFs (Insuline like Growth Factors)</i> .....	29
2.3.4.1.3 <i>FGFs (Fibroblast Growth Factors)</i> .....	29
2.3.4.2 Destrucción de la ECM .....	30
2.3.4.2.1 <i>Citoquinas (IL-1, TNF-<math>\alpha</math>)</i> .....	30
2.3.4.2.2 <i>Óxido nítrico (NO) y ROS (especies reactivas de oxígeno)</i> .....	31
2.3.4.2.3 <i>Proteasas de matriz extracelular</i> .....	32
2.3.4.2.3.1 <i>Metaloproteasas (MMPs)</i> .....	33
Metaloproteasas de matriz extracelular.....	33
2.3.4.2.3.2 <i>ADAMTS, ADAMS y TIMPs</i> .....	35
2.3.4.2.3.2 <i>Serín proteasas</i> .....	37
2.3.5 Otros mecanismos biológicos que participan en el desarrollo de la OA: inflamación y angiogénesis.....	38
3. <i>Genética de la OA</i> .....	41
3.1 Estudios de ligamiento .....	45
3.2 Estudios de genes candidato .....	46
3.3 Estudios de asociación de genoma completo (GWAS) .....	48
3.4 Factores genéticos de susceptibilidad a la OA .....	52

3.4.1 VNTR en asporina que afecta a la actividad de TGF- $\beta$ .....	53
3.4.2 Un polimorfismo en el promotor de <i>CALM-1</i> con asociación no confirmada.....	54
3.4.3 Evidencia inconcluyente de asociación de un polimorfismo en <i>SMAD3</i> .....	55
3.4.4 <i>MATN3</i> e <i>IL1</i> cluster ligados a OA de mano pero sin confirmación definitiva .....	56
3.4.5 Un microsatélite funcional de <i>BMP-5</i> asociado con OA de cadera en mujeres del Reino Unido.....	58
3.4.6 <i>IL4-R</i> : otro locus ligado a OA pero no confirmado .....	59
3.4.7 <i>FRZB</i> : asociación de dos SNPs no sinónimos no confirmados en estudios más recientes .....	60
3.4.7 Variantes en <i>DIO2</i> posiblemente asociadas con OA de cadera a través de modulación de la tiroxina .....	61
3.4.8 <i>DVWA</i> : el mejor candidato del primer GWAS en OA realizado en población japonesa.....	62
3.4.9 <i>LRCHI</i> : uno de los resultados de un GWAS realizado con DNA pools.....	63
3.4.10 <i>PTGS2</i> : un magnífico candidato funcional cuya asociación no se ha confirmado .....	64
3.4.11 Posible asociación de <i>A2BPI</i> con OA de mano.....	65
3.4.12 <i>HLA</i> : locus asociado con nivel de GWAS en asiáticos pero no en europeos.....	66
3.4.13 <i>GDF-5</i> : primer locus asociado con nivel de GWAS en europeos y asiáticos.....	67
3.4.14 Locus 7q22: segundo locus asociado con nivel GWAS en europeos. ....	69
3.4.15 Resultados de la primera fase del proyecto arcOGEN.....	71
3.4.16 <i>COL11A1</i> : el mejor candidato entre las señales del proyecto arcOGEN fase 1 .....	72
3.4.17 Resultados de la segunda fase del proyecto arcOGEN .....	73
3.4.18 <i>DOTIL</i> .....	76
3.4.19 <i>MCF2L</i> .....	77
3.5 Visión general y futuro de los estudios genéticos de OA.....	78
<b>OBJETIVOS</b> .....	83
<b>PROCEDIMIENTO EXPERIMENTAL Y RESULTADOS</b> .....	87
<b>Publicación 1:</b> Estudio de la relación entre la edad de inicio de los síntomas de OA y la carga genética de la enfermedad .....	91
<b>Publicación 2:</b> Resequenciado del gen <i>GDF-5</i> buscando variantes raras asociadas .....	101
<b>Publicación 3:</b> Estudio de asociación caso-control en OA de rodilla del microsatélite <i>D6S1276</i> de <i>BMP-5</i> .....	115
<b>Análisis de genes candidatos en un metaanálisis de 9 estudios de asociación de genoma completo</b> .....	129
MATERIAL Y MÉTODOS.....	129

Colecciones de muestras.....	129
Selección de genes candidatos y SNPs.....	130
Genotipado e imputación.....	132
Análisis estadísticos.....	132
RESULTADOS.....	137
Identificación de genes candidatos.....	137
Metaanálisis de asociación de los SNPs de los genes candidatos.....	137
Estudio de extensión y combinado.....	140
<b>DISCUSIÓN</b> .....	147
Relación entre la edad de inicio de los síntomas de OA y la carga genética de la enfermedad.....	149
Resequenciado del gen <i>GDF-5</i> buscando variantes raras asociadas.....	154
Estudio de asociación caso-control del microsatélite <i>D6S1276</i> de <i>BMP-5</i> .....	157
Análisis de genes candidatos en un metaanálisis de 9 GWAS.....	161
<b>CONCLUSIONES</b> .....	169
<b>BIBLIOGRAFÍA</b> .....	171
<i>Anexo</i> .....	201
<b>Publicación 4:</b> <i>Nuevas variantes en la secuencia de la región de HLA clase II/III están asociadas con susceptibilidad a OA de rodilla mediante GWAS.</i> ....	203
<b>Publicación 5:</b> <i>Variantes comunes en genes relacionados con los estrógenos están asociadas con artrosis severa: un estudio multicentro y funcional</i> .....	211





***ABREVIATURAS***

---



## ABREVIATURAS

**3'UTR:** Región 3' no traducida (del inglés, *3' untranslated region*)

**5'UTR:** Región 5' no traducida (del inglés, *5' untranslated region*)

**A2BP1:** Proteína 1 de unión a la ataxina 2 (del inglés, *ataxin 2-binding protein 1*)

**ACR:** Colegio Americano de Reumatología (del inglés, *American College of Rheumatology*)

**ADAMs:** Metaloproteasas y desintegrinas (del inglés, *a disintegrin and metalloproteinase*)

**ADAMTS:** Metaloproteasas y desintegrinas con motivos trombospondina (del inglés, *a disintegrin and metalloproteinase with thrombospondin motifs*)

**arcOGEN:** *Arthritis Research UK Osteoarthritis Genetics*

**ASTN2:** Astrotactina 2

**BMD:** Densidad mineral ósea (del inglés, *bone mineral density*)

**BMI:** Índice de masa corporal (del inglés, *body mass index*)

**BMP:** Proteína morfogenética ósea (del inglés *bone mineral protein*)

**BTNL-2:** *butyrophilin-like 2*

**CALM1:** Calmodulina 1

**CDMP-1:** Proteína morfogenética 1 derivada del cartilago (del inglés, *cartilage-derived morphogenetic protein-1*)

**CHST11:** Condroitin. 4-sulfotransferasa 11 (del inglés, *carbohydrate (chondroitin 4) sulfotransferase 11*)

**eNOS:** Sintasa de óxido nítrico constitutiva (del inglés, *constitutive nitric oxide synthase*)

**DAE:** Expresión alélica diferencial (del inglés, *different allelic expression*)

**DIO2:** Iodotironina deiodinasa de tipo 2 (del inglés, *type II iodothyronine deiodinase*)

**DNA:** Ácido desoxirribonucleico (del inglés, *deoxyribonucleic acid*)

**DUS4L:** Dihidrouridina sintasa 4 similar (del inglés, *dihydrouridine synthase 4-like*)

**DVWA:** *Double von Willebrand factor A domains*

**COG5:** Subunidad 5 del complejo oligomérico conservado de Golgi (del inglés, *conserved oligomeric Golgi complex subunit 5*)

**COMP:** Proteína oligomérica de la matriz del cartilago (del inglés, *cartilage oligomeric matrix protein*)

**COX-2:** Ciclooxygenasa de tipo 2 (del inglés, *cyclooxygenase 2*)

**DOT1L:** Histona H3 metiltransferasa ó similar a DOT1 (del inglés, *DOT1-like*)

**ECM:** Matriz extracelular (del inglés, *extracellular matrix*)

**EMSA:** Ensayo de cambio de movilidad electroforética (del inglés, *electrophoretic mobility shift assay*)

**ENCODE:** ENCyclopedia Of DNA Elements

**ER:** Retículo endoplasmático (del inglés, *endoplasmic reticulum*)

**FGF:** Factor de crecimiento de fibroblastos (del inglés, *fibroblast growth factor*)

**FILIP1:** Proteína 1 de interacción con filamina A (del inglés, *filamin-A-interacting protein 1*)

**GAG:** Glucosaminoglicanos

**GARP:** *Genetics, Arthrosis and Progression*

**GDF-5:** Factor de crecimiento y diferenciación 5 (del inglés, *growth differentiation factor-5*)

**GLT8D1:** *Glycosyltransferase 8 domain containing 1*

**GNL3:** Proteína 3 de unión al nucleótido guanina (del inglés, *guanine nucleotide-binding protein-like 3*)

**GPR22:** *G protein-coupled receptor 22*

**GWAS:** Estudio de asociación de genoma completo (del inglés, *genome wide association study*)

**HBP1:** *HMG-box transcription factor 1*

**HLA:** Antígeno leucocitario humano (del inglés, *human leukocyte antigen*)

**IGFs:** Factores de crecimiento similares a la insulina (del inglés, *insulin like growth factors*)

**IGFBPs:** Proteínas de unión al factor de crecimiento similar a insulina (del inglés, *insulin-like growth factor-binding protein*)

**IGFRs:** Receptores del factor de crecimiento parecido a la insulina (del inglés, *insulin like growth factor receptors*)

**IFN-I:** Interferón de tipo I

**IL1:** Interleucina 1 (del inglés, *interleukin 1*)

**IL1R:** Receptor de interleucina 1 (del inglés, *IL1 receptor*)

**IL-1RA:** Receptor antagonista de interleucina 1 (del inglés, *IL1 receptor antagonist*)

**iNOS:** Sintasa de óxido nítrico inducible (del inglés, *inducible nitric oxide synthase*)

**JSW:** Anchura del espacio mínimo articular (del inglés, *joint space width*)

**KL:** Escala de Kellgren y Lawrence

**LD:** Desequilibrio de ligamiento (del inglés, *linkage disequilibrium*)

**LDH:** Hernia de disco lumbar (del inglés, *lumbar disc herniation*)

**MATN3:** Matrilina 3

**MHC:** Complejo mayor de histocompatibilidad (del inglés, *major histocompatibility*)



*complex*)

**MICAL3:** Monooxigenasa asociada a microtúbulos, calponina y dominio LIM de tipo 3 ( del inglés, *microtubule associated monoxygenase, calponin and LIM domain containing 3*)

**MMPs:** Metaloproteasas (del inglés, *metalloproteinases*)

**NFκB:** Factor nuclear κB (del inglés, *nuclear factor κB*)

**NGF:** Factor de crecimiento nervioso (del inglés, *neurotrophin growth factor*)

**NO:** Óxido nítrico (del inglés, *nitric oxide*)

**nsSNP:** SNP no sinónimo

**OR:** Odds ratio

**PCR:** Reacción en cadena de la polimerasa (del inglés, *polimerase chain reaction*)

**PTGS2:** Prostaglandina endoperoxidasa sintasa 2 (del inglés, *prostaglandin-endoperoxide synthase 2*)

**PTH1H:** Hormona similar a la paratiroidea (del inglés, *parathyroid hormone-like hormone*)

**PRG4:** Proteoglicano de tipo 4 ó lubricina

**PRKAR2B:** Proteína quinasa reguladora 2B dependiente de AMPc (del inglés, *cAMP-dependent protein kinase type II-beta regulatory subunit*)

**qPCR:** PCR cuantitativa (del inglés, *quantitative PCR*)

**RNA:** Ácido ribonucleico (del inglés, *ribonucleic acid*)

**ROS:** Especies reactivas de oxígeno (del inglés, *reactive oxygen species*)

**RUNX2:** *Runt-related transcription factor 2*

**SLE:** Lupus eritematoso sistémico (del inglés, *systemic lupus erythematosus*)

**SLRPs:** Proteoglicanos de pequeño tamaño ricos en leucina (del inglés, *small leucin rich proteoglicans*)

**SMAD:** *Small Mothers Against Decapentaplegic*

**SOX-9:** Sry-related HMG box-9

**SUPT3H:** *Suppressor of Ty3 homolog*

**tag SNP:** SNP marcador

**TACE:** Enzima convertidora de TNF ( del inglés, *TNF converting enzyme*)

**TGF-β:** Factor de crecimiento transformante beta 1 (del inglés, *transforming growth factor beta 1*)

**THR:** Reemplazamiento articular de cadera (del inglés, *total hip replacement*)

**TIMPs** Inhibidores de metaloproteasas de tejido (del inglés, *tissue inhibitors of metalloproteinases*)

**TKR:** Reemplazamiento articular de rodilla (del inglés, *total knee replacement*)

**TLR:** Receptor tipo Toll (del inglés, *Toll-like receptor*)

**TNF- $\alpha$** : Factor de necrosis tumoral- $\alpha$  (del inglés, *tumor necrosis factor- $\alpha$* )  
**TreatOA**: *Translation Research in Europe Applied Technologies for Osteoarthritis*  
**TRIB1**: Homólogo 1 de *tribbles* (del inglés, *tribbles homolog 1*)  
**VEGF**: Factor de crecimiento endotelial vascular (del inglés, vascular *endothelial growth factor*)  
**VNTR**: Repetición en tándem variable en número (del inglés, *variable number tandem repeat*)  
**As**: Riesgo relativo entre hermanos





## **INTRODUCCIÓN**

---



# INTRODUCCIÓN

## 1. *La artrosis(OA)*

La artrosis (OA: osteoarthritis) (OMIM 165720) es una enfermedad osteoarticular caracterizada por una degeneración progresiva del cartilago y por diversas alteraciones en los tejidos de la articulación. Los cambios que se producen en la articulación afectada tienen como consecuencia una pérdida del cartilago articular hialino junto con modificaciones del hueso subcondral como esclerosis, desarrollo de osteófitos y crecimiento marginal. Además, otras estructuras blandas de la articulación como los ligamentos y los músculos articulares también sufren cambios volviéndose más laxos (Figura 1).

La destrucción del cartilago se origina por la ruptura del equilibrio existente entre las señales anabólicas y catabólicas de la ECM (*extracellular matrix*). Esta pérdida del equilibrio lleva finalmente al fallo de la articulación, al perderse el papel protector y amortiguador que el cartilago proporciona. Así, en la OA avanzada, la superficie del cartilago articular que normalmente es suave y lisa muestra una serie de cambios hasta que se pierde, quedando expuesto el hueso subcondral de la articulación.

Acompañando a la degradación del cartilago articular también se puede observar inflamación del tejido sinovial así como presencia de infiltrado inflamatorio

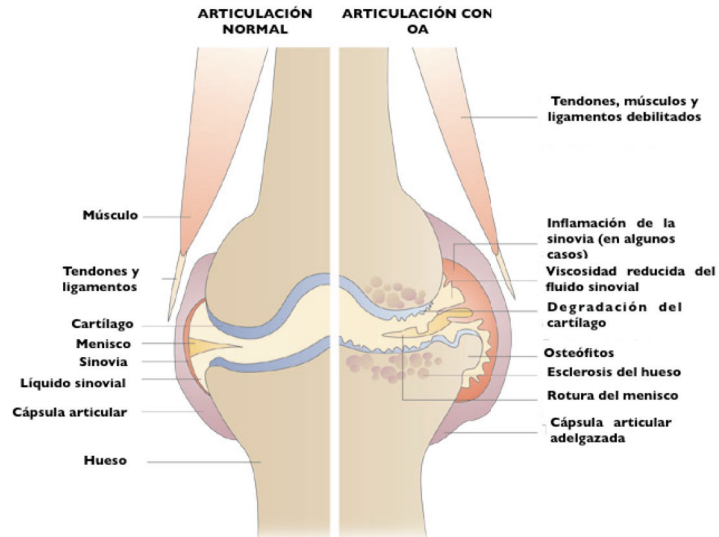


Figura 1. Esquema de una articulación sana vs una artrósica. En la articulación normal el cartilago articular no está dañado y permanece lubricado por el líquido sinovial lo que permite el movimiento de la articulación. La OA provoca la degradación del cartilago primero adelgazándolo y finalmente rompiendo su superficie. Cuando el daño de la articulación es severo se producen cambios en el hueso como osteófitos y esclerosis del hueso subcondral e inflamación de la sinovia. También puede producirse ensanchamiento o deformidad de las articulaciones. Tomado de Wieland H.A et al. *Nature Reviews Drug Discovery* 4, 331-334 (2005).

### 1.1 Diagnóstico y clasificación

Clínicamente, la OA se caracteriza por dolor en la articulación, limitación del movimiento y posible inflamación. Junto con la presencia de estos síntomas, el examen radiológico de la articulación es el principal método para determinar la existencia de OA y el más ampliamente utilizado sobre todo en estudios epidemiológicos. El sistema de clasificación utilizado es

el de Kellgren y Lawrence (KL) que puntúa la gravedad de las lesiones en una escala del 0-4. Se define la OA radiográficamente con un  $KL \geq 2$ .

Se suele hablar de OA sintomática y radiográfica. La OA sintomática se define con la presencia de características radiográficas en combinación con presencia de síntomas como dolor o rigidez articular. Sin embargo, la OA radiográfica puede estar definida por examen radiográfico y no presentar síntomas <sup>2</sup>.

Debido a la heterogeneidad de la OA se han establecido una serie de criterios de clasificación para la enfermedad por la ACR (American College of Rheumatology), que varían según la articulación y que ayudan a identificar a los pacientes con OA clínica, basándose principalmente en la presencia de dolor articular en combinación con cambios estructurales u otras anomalías radiológicas<sup>3</sup>.

La OA también se puede clasificar en primaria y secundaria de acuerdo a la causa o factor principal desencadenante. La OA primaria o idiopática, es la forma más común de OA y no tiene una etiología identificable. Por el contrario la OA secundaria tiene un factor causante conocido aunque patológicamente las lesiones pueden ser indistinguibles de la artrosis primaria. Entre los factores que se han reconocido como causantes de artrosis secundaria se encuentran los traumatismos y los cambios metabólicos, anatómicos e inflamatorios <sup>4</sup>. Otra categorización posible se establece en función de la articulación afectada: artrosis de cadera, de mano, de rodilla, etc... También es posible otra subclasificación según el número de articulaciones afectadas, pudiendo ser monoarticular, la que implica a una sola articulación o poliarticular en la que un mismo individuo presenta la enfermedad en varias articulaciones. También se

utilizan como criterios de clasificación la presencia de inflamación o erosiones, asociadas normalmente a OA de mano.

### **1.2 Prevalencia e incidencia**

La prevalencia se define como la proporción de individuos dentro de una población que presentan una enfermedad. Existen numerosas estimaciones de la prevalencia de la OA que son muy variables dependiendo de la articulación, del sexo, de la edad y de la etnia. Las estimaciones del riesgo de desarrollar OA durante la vida de un individuo rondan el 24% para OA de rodilla, 11% para OA de cadera y un 43% para OA de mano<sup>5</sup>.

Una de las cohortes más utilizadas en los estudios epidemiológicos de OA es la de Framingham. En esta cohorte se estimó que un 19% de los adultos mayores de 45 años padecen OA de rodilla. Este dato es fuertemente dependiente del sexo comprobándose que existe un mayor número de mujeres (1.7 veces más) con OA sintomática que hombres. También es dependiente de la edad pues al incluir sujetos de más de 80 años la prevalencia de la OA radiográfica aumenta situándose entorno al 40%<sup>6</sup>.

La incidencia es el número de casos nuevos en un período determinado. La incidencia de OA de rodilla radiográfica en la cohorte de Framingham es del 1% y varía dependiendo del sexo, siendo 1,7 veces mayor en mujeres que en hombres<sup>7</sup>. La edad también afecta a la incidencia que se ve incrementada, concretamente aumentando de 2 a 10 veces, entre los 30 y los 65 años de edad<sup>8</sup>.



### 1.3 Factores de riesgo

La OA es una enfermedad compleja, con una naturaleza multifactorial en la que participan factores sistémicos, biomecánicos y genéticos (Figura 2). Entre los factores sistémicos más relevantes se incluyen la edad y el sexo.

La edad es uno de los factores de riesgo más importantes en la OA debido a los cambios que provoca en el cartílago articular y el condrocito. Así, la edad produce desgaste y adelgazamiento de la superficie articular y el condrocito pierde la capacidad de mantener y reparar la ECM <sup>9</sup>.

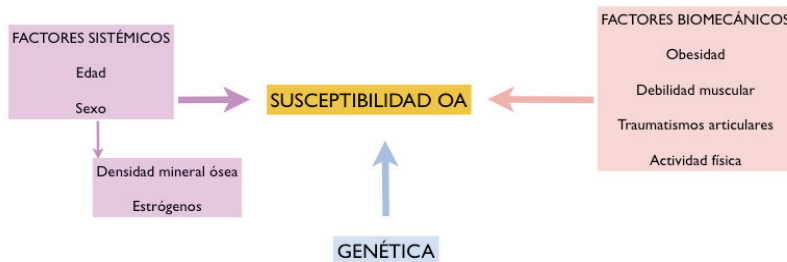


Figura 2. Esquema que resume los factores de riesgo para OA. Tomado de Felson D.T et al. *Annals of internal Medicine* 133(8): 635-46 (2000).

En relación al sexo, en las mujeres se observa una mayor prevalencia y severidad de la OA que en los hombres. Esto se debe en parte a que las características físicas del cartílago son diferentes entre ambos sexos. Las mujeres poseen un cartílago más delgado y con menor volumen que los hombres y además también existe una relación entre los niveles de estrógenos y el desarrollo de OA <sup>10</sup>. En estudios de cohortes se observó que la incidencia de la OA aumenta después de la menopausia lo que

## INTRODUCCIÓN

---

pareció indicar un papel protector de los estrógenos frente a la enfermedad<sup>11</sup>. También se sugirió una relación inversa entre OA y osteoporosis. Mientras que en la OA se observa un aumento de la *BMD* (*bone mineral density*), en la osteoporosis se encuentra un descenso de la misma que se asocia con la reducción en los niveles de estrógenos propia de la menopausia. En estudios con pacientes tratadas con terapia de sustitución hormonal para suplir la carencia de estrógenos, se comprobó como el riesgo de OA es mayor, al tiempo que se produce un aumento de *BMD*<sup>12</sup>. Los estrógenos no sólo afectan a los huesos de la articulación sino que también pueden influenciar la función del condrocito a diversos niveles, interactuando con citoquinas o factores de transcripción<sup>13</sup>.

Entre los factores biomecánicos que aumentan el riesgo de OA se incluyen la obesidad, los traumatismos articulares, la actividad física y la debilidad muscular. En relación a la obesidad se ha observado la existencia de una correlación positiva entre el *BMI* (*body mass index*) y el riesgo de desarrollar OA<sup>14</sup>. La obesidad supone una sobrecarga en las articulaciones tanto en la cadera como en la rodilla lo que provoca estrés mecánico en el cartílago. También se sabe que no sólo existe una relación de tipo mecánico entre este factor y la enfermedad. La existencia de una relación entre OA de mano y obesidad sugiere que existe otro vínculo<sup>15</sup>. El tejido adiposo no sólo almacena lípidos también secreta hormonas por lo que se puede considerar un órgano endocrino que puede desempeñar un papel inflamatorio. La secreción de quimoquinas y adipoquinas que funcionan como mediadores activos metabólicamente pueden tener un papel relevante en el inicio y progresión de la OA<sup>16</sup>.

Los traumatismos articulares son otro de los factores a tener en cuenta. Se conoce que el daño o rotura del menisco es causante de aparición de OA de rodilla de inicio temprano<sup>17,18</sup>. Otra lesión importante es la rotura del ligamento cruzado anterior que también se asocia con un riesgo mayor de OA<sup>19</sup>. En relación con la actividad física, cuando se produce una carga excesiva de la articulación puede favorecer el desarrollo de OA<sup>20</sup>. Por el contrario, la actividad física moderada es beneficiosa al favorecer un incremento de la fuerza muscular periarticular que ayuda a estabilizar la articulación<sup>21</sup>.

Los diferentes factores sistémicos y biomecánicos comentados se pueden incluir dentro del componente ambiental que afecta a la progresión y desarrollo de OA. Al tratarse de una enfermedad compleja, los factores ambientales interactúan con múltiples factores genéticos<sup>22</sup>. El componente genético de la OA es el centro de este estudio por lo que será tratado en detalle en apartados posteriores.

### ***2. El cartílago articular***

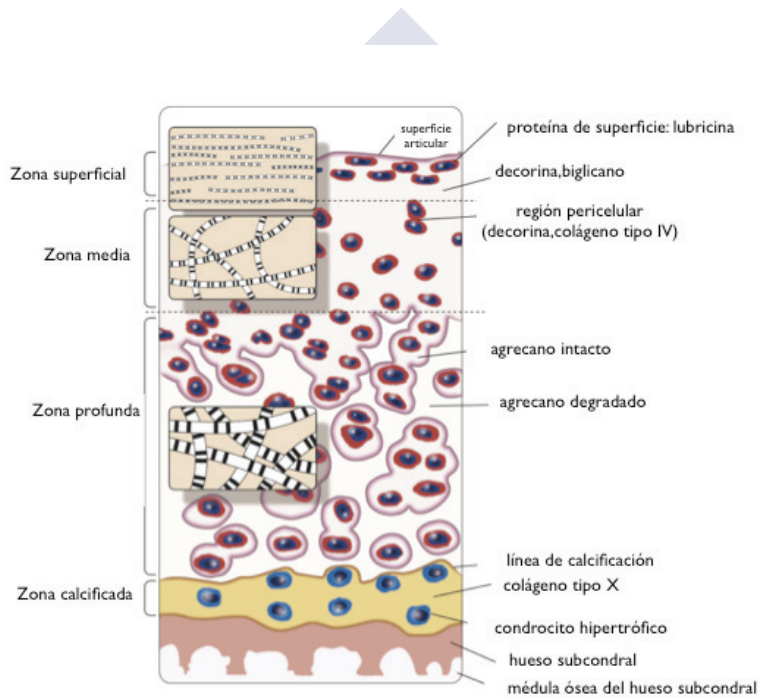
El cartílago articular es un cartílago hialino que carece de vascularización e inervación, permanece anclado al hueso subcondral y consta de un único tipo celular, los condrocitos que están embebidos en la matriz extracelular (ECM). Los condrocitos constituyen una parte minoritaria del cartílago articular; aproximadamente un 1-2%, siendo el resto la ECM. La ECM está constituida fundamentalmente por agua (60-80%) y componentes fibrilares: colágeno y proteoglicanos (20-40%). Considerando el peso seco del cartílago, el colágeno es el componente mayoritario (50%), fundamentalmente colágeno de tipo II ; seguido por proteoglicanos 30-35%, glicoproteínas y proteoglicanos minoritarios (15-20%)<sup>23</sup>.

#### **2.1 Función y estructura del cartílago articular**

El cartílago articular posee una serie de características físicas que le hacen idóneo para el correcto funcionamiento articular. Es un tejido altamente compresible, lo que le proporciona la capacidad de amortiguar la presión ejercida en las articulaciones. Además, reduce la fricción al proporcionar una superficie lisa en la que las articulaciones se deslizan al moverse <sup>24</sup>. Por ello, la degradación del cartílago articular es el evento central en la patogénesis de la OA.

En un corte transversal del cartílago articular maduro se observa una estructura claramente diferenciada (Figura 3). La zona superficial está en contacto con el líquido sinovial y representa aproximadamente un 10-20% del grosor del cartílago. Posee un contenido en condrocitos mayor que las

otras zonas que tienen morfología discoide y se disponen paralelamente a la superficie del cartilago. Los condrocitos están incluidos en la ECM, formada mayoritariamente por colágeno de tipo II y decorina <sup>25</sup>. Estos condrocitos de la zona superficial se caracterizan por la producción de lubricina o PRG4 que actúa como lubricante de la articulación y juega un papel en la homeostasis sinovial<sup>26</sup>. La desorganización del colágeno y la pérdida de proteoglicanos en esta zona son los primeros cambios observados al inicio de la OA<sup>27</sup>.



**Figura 3. Estructura del cartilago articular humano mostrando sus cuatro zonas principales y el contenido de cada una. Además se muestra la orientación de las fibras de colágeno en las diferentes regiones. Tomado de Firestein: Kelley's Textbook of Rheumatology, 8th ed.**

La siguiente zona del cartilago representa la zona media también llamada de transición. La morfología de los condrocitos se vuelve más redondeada y pierden la orientación de la región anterior siguiendo un patrón más

## INTRODUCCIÓN

---

desorganizado. Estos condrocitos se caracterizan por una gran producción de proteoglicanos mientras que las fibras de colágeno se muestran entrecruzadas y pierden la orientación paralela de la superficie<sup>28,29</sup>.

La siguiente zona es la profunda o radial que puede ocupar un 80% del grosor del cartílago y que se extiende hasta la zona calcificada, separada de ella por la denominada línea de marea (*tidemark*)<sup>30</sup>. Los condrocitos en la zona profunda tienen una morfología más alargada y están entre las fibras de colágeno dispuestas radialmente. En un corte histológico transversal las células se disponen en columnas con las fibras de colágeno entre ellas. Los condrocitos en esta capa, se encuentran en agrupaciones de 3 o más células con una elevada deposición de matriz extracelular a su alrededor<sup>31</sup>.

En esta zona, se producen cambios característicos de la OA como la presencia de calcificación anormal y desorganización de los componentes fibrilares. En la zona limitante con la línea de calcificación, aparecen condrocitos hipertróficos que producen colágeno de tipo X. Estos condrocitos se vuelven mucho más activos metabólicamente. Comienzan a producir diversas proteínas como MMP-13 (*metalloproteinase-13*), osteocalcina, osteopontina, fosfatasa alcalina y el mencionado colágeno de tipo X, que son marcadores del cartílago hipertrófico<sup>32</sup>. Estos cambios son similares a los que se producen en la placa de crecimiento de los huesos.

La zona calcificada es una región mineralizada del cartílago que lo ancla al hueso subcondral. En ella los condrocitos son hipertróficos o han sufrido apoptosis y sólo queda el lugar vacío que habían ocupado en el cartílago (*lacuna*). Esta capa ancla el cartílago sobre el hueso subcondral.

## 2.2 El condrocito

El condrocito es el único tipo celular del cartílago articular. Su forma varía dependiendo de la zona del cartílago articular donde se localiza, siendo típicamente redondeada, poligonal o discoide. Son células muy activas metabólicamente, cuya función principal es sintetizar y regular la degradación de los componentes de la ECM como colágenos, glicoproteínas, proteoglicanos y ácido hialurónico. Al estar incluidos en un tejido no vascularizado soportan un ambiente con una baja tensión de  $O_2$ <sup>33</sup>.

También hay condrocitos en el cartílago de las epífisis óseas donde participan en el crecimiento durante el desarrollo del esqueleto. En la formación del hueso estos cambios acaban en la osificación endocondral. Los cambios que producen los condrocitos desde células mesenquimales hasta condrocitos hipertróficos se parecen a los que tienen lugar en la OA y se han usado como modelo de los mismos. De esta forma, moléculas centrales en la diferenciación de los condrocitos como SOX, BMPs ó TGF- $\beta$  son también moléculas relevantes en el desarrollo de la OA<sup>34</sup>

### 2.2.1 El proceso de condrogénesis y la formación de la articulación

Los condrocitos tienen tres posibles orígenes embrionarios; pueden proceder de las células de la cresta neural, del mesodermo paraxial o de las células mesenquimales de las placas laterales. Durante el proceso de embriogénesis, moléculas de las rutas de señalización WNT (*wingless* +

## INTRODUCCIÓN

---

*Int-1*), BMPs (*bone morphogenetic proteins*) y TGF- $\beta$  (*transforming growth factor beta*) están implicadas en el inicio del proceso de diferenciación desde células del mesodermo a condrocitos<sup>35</sup>.

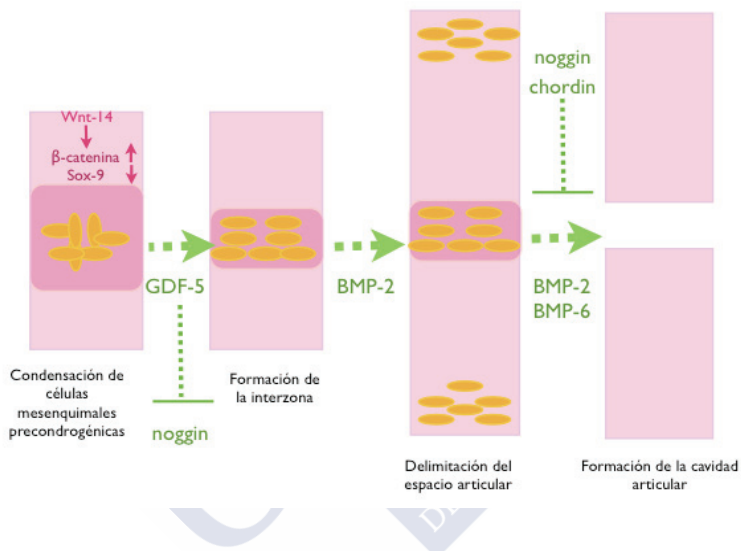
El primer evento que tiene lugar es la condensación de las células precursoras que formarán el primordio de los huesos. En su agrupación intervienen las cadherinas y las BMPs<sup>36</sup>. Los condrocitos comienzan a sintetizar ECM, incluyendo colágeno de tipo II. Además, comienza el proceso de diferenciación, regulado principalmente por el factor de transcripción SOX-9 (*Sry-related HMG box-9*) y de proliferación para formar el molde del futuro hueso.

Después de esta fase de diferenciación y proliferación, los condrocitos localizados en el núcleo de las agrupaciones condrogénicas inician el proceso de diferenciación hipertrófica para finalmente entrar en apoptosis. La matriz de este cartílago servirá de molde para la deposición de la matriz mineral ósea. El cartílago localizado en el centro del hueso se degrada y es invadido por vasos sanguíneos que forman la médula ósea. En los extremos de los huesos largos se localizan las denominadas placas de crecimiento del cartílago que continúan el proceso de diferenciación condrogénica.

Los condrocitos que ocupan los extremos del hueso en formación no siguen el proceso de diferenciación hasta condrocitos hipertróficos, y una vez que el patrón global del esqueleto está establecido, derivan su destino hacia la formación de la futura articulación<sup>37</sup>. Este proceso (Figura 4), se inicia por la condensación de las células mesenquimales entre los dos huesos en formación. Estas células de la futura interzona poseen una morfología aplanada y comienzan a expresar fundamentalmente, WNT9a



ó WNT14 que se consideran la señales iniciadoras de formación de la articulación<sup>38</sup>. La expresión de estas moléculas provoca una disminución de la expresión de SOX-9 y de colágeno de tipo II, interrumpiendo de esta forma la maduración hacia cartílago<sup>39</sup>.



**Figura 4. Etapas de la formación de la articulación con las principales moléculas encargadas de iniciar y dirigir cada proceso. Modificado de Jorgenesen C et al. *Annals of Rheumatic Diseases* 61(1): 6-9 (2002).**

Además, como consecuencia de la expresión de WNT-14 comienza la sobreexpresión de GDF-5 que provoca la formación de la interzona que marca la zona de la futura cavidad articular. GDF-5 desencadena la expresión de moléculas de la familia de las BMPs que dan paso a las siguientes etapas del proceso<sup>37</sup>. Concretamente, BMP-2 es de especial relevancia ya que participa en el proceso de cavitación que produce la separación de los elementos del esqueleto y la formación del espacio articular<sup>40</sup>. La cavitación está ligada al movimiento embrionario, el

crecimiento longitudinal de las extremidades y la síntesis de ácido hialurónico para la formación del líquido sinovial<sup>41,42</sup>. A lo largo de todo el proceso de formación de la articulación la regulación de las moléculas antagonistas de las BMPs, *noggin* y *chordin*, es fundamental para la sucesión coordinada de las distintas etapas<sup>43</sup>.

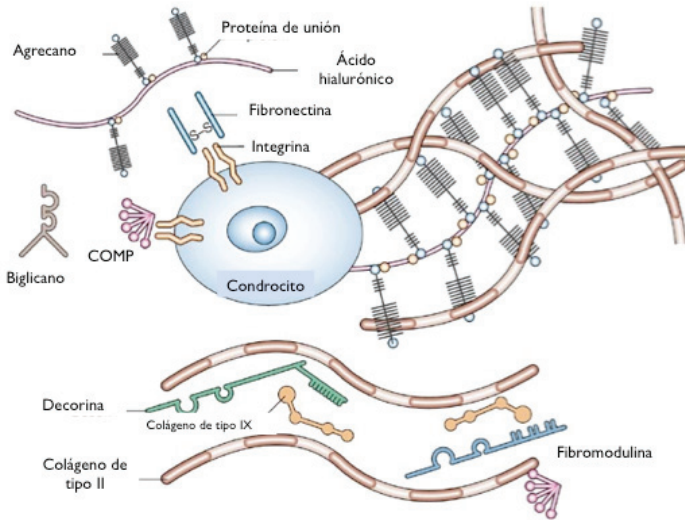
### 2.2.2 El papel del condrocito en la OA

El condrocito en el cartílago con OA, pierde el fenotipo proliferativo y adquiere el hipertrófico. Existen múltiples marcadores de hipertrofia del condrocito como el colágeno de tipo X, MMP-13, Runx-2, fosfatasa alcalina ó VEGF. En una etapa inicial del proceso el condrocito comienza la producción de MMP-13. Durante la progresión de la OA se degrada la ECM que lo rodea y comienza a disminuir la síntesis de colágeno de tipo II y agreganos. Se ha demostrado que existe una correlación entre el cambio de fenotipo hacia el hipertrófico y la degradación del cartílago, pues cuando la degradación es inhibida, también se inhibe la hipertrofia<sup>44</sup>.

### 2.3 La matriz extracelular (ECM)

En el cartílago articular existe un equilibrio entre las señales anabólicas y catabólicas. En el cartílago con OA, este equilibrio se rompe aumentando la degradación de la ECM e inhibiéndose la formación de nueva ECM.

Las dos moléculas mayoritarias que forman la ECM son el colágeno II (proteína estructural) y el agregano. Además también existe un porcentaje pequeño otras proteínas como COMP y proteoglicanos minoritarios (Figura 5).



**Figura 5.** Figura que muestra el condrocito y la ECM con sus tres clases de proteínas mayoritarias: colágenos (de tipo II mayoritariamente), proteoglicanos (agregano) y otras proteínas como fibronectina y COMP, así como los SLRPs (biglicano, decorina y fibromodulina). Tomado de Faye H Chen. *Nature Clinical Practice Rheumatology*. 2, 373-382 (2006).

### 2.3.1 El colágeno

La familia del colágeno se compone de 12 genes que codifican para 20 tipos de proteínas fibrilares diferentes<sup>45</sup>. Todos los componentes tienen en común poseer una triple hélice proteica formada por dos cadenas  $\alpha 1$  y una cadena  $\alpha 2$ . Cada una de estas cadenas incluye un péptido de reconocimiento de señal para su transporte hacia el ER (retículo endoplasmático). La primera transformación de este pre-propéptido es la pérdida del péptido señal dando lugar al pro-péptido del colágeno. A continuación, tiene lugar la hidroxilación de prolina y algunas glicosilaciones, así como el enrollamiento de los tres pro-péptidos

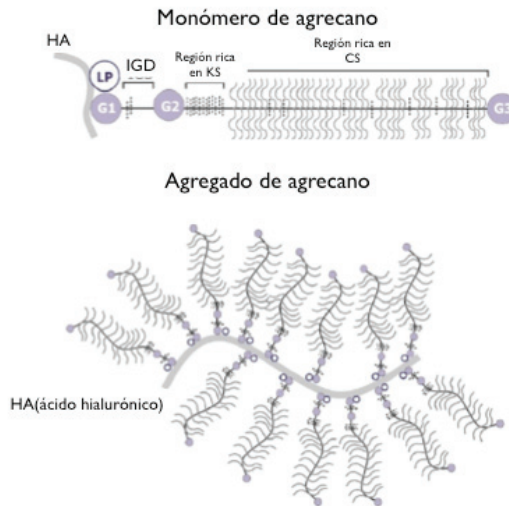
formando la triple hélice. El procolágeno es empaquetado en vesículas para su transporte al aparato de Golgi, dónde tras modificaciones postraduccionales es secretado fuera de la célula. Una vez en el exterior, pierde los extremos N y C terminal por la acción de peptidasas de colágeno, formándose la molécula de tropocolágeno. La unión de moléculas de tropocolágeno entre sí da lugar a las fibrillas de colágeno. El colágeno de tipo II (COL2A1) es el colágeno mayoritario del cartílago articular, seguido por los colágenos de tipo IX (COL9A1, COL9A2, COL9A3) y XI (COL11A1, COL11A2). El colágeno X (COL10A1) no es un componente del cartílago articular sano. Sólo se encuentra en la capa más profunda del cartílago calcificado adyacente al hueso subcondral<sup>46</sup>. También se pueden encontrar algunos colágenos muy minoritarios como el colágeno de tipo III que se sintetiza en respuesta al daño articular y que se une al colágeno de tipo II<sup>47</sup>.

Los colágenos II-IX-XI forman una matriz heteropolimérica que constituye el entramado fundamental de la ECM. El colágeno de tipo II constituye la mayor parte de la estructura, el colágeno de tipo IX facilita la interacción entre las fibras con moléculas de proteoglicanos mientras que el colágeno de tipo XI regula el espaciado entre las fibras de colágeno de tipo II<sup>48</sup>.

### **2.3.2 Proteoglicanos**

Los proteoglicanos son una familia de glicoproteínas muy ricas en polisacáridos. Están formados por una proteína central a la se unen covalentemente cadenas de GAG (glucosaminoglicanos). Los proteoglicanos se clasifican en función de los GAG que contengan.

El agregano es un proteoglicano de gran tamaño que se une a moléculas de agua confiriendo a la ECM sus propiedades de comprensibilidad y al cartílago sus particulares características biomecánicas. Tiene tres dominios globulares; el G1 y el G2 en el extremo N-terminal y el G3 en el extremo C-terminal.



**Figura 6.** Estructura del monómero de agregano con sus diferentes dominios: IGD(dominio interglobular), G1,G2 y G3 (dominios globulares), KS (región rica en keratán sulfato) y CS (región rica en condroitín sulfato) . En el agregado de agregano hasta 100 monómeros se unen a una sola molécula de ácido hialurónico (HA) a través de proteínas de unión (LP). Tomado de Fosang A. *Proteinases and OA*. Primer OARSI 2012.

El dominio G1 tiene el papel principal de unir las moléculas individuales de agregano a lo largo de una molécula de ácido hialurónico para formar agregados. A cada molécula de ácido hialurónico se le pueden unir hasta 100 monómeros de agregano a través de una proteína de unión<sup>49</sup>(Figura 6). Las regiones interglobulares desde los dominios globulares G1 a G2 y

desde los dominios G2 a G3, se unen a múltiples moléculas de condroitín sulfato y keratán sulfato que son los principales glucosaminoglicanos.

En el cartílago, además del agrecano existen una serie de proteoglicanos minoritarios. Estos proteoglicanos incluyen los SLRPs (*small leucin rich proteoglicans*): decorina, biglicano, fibromodulina, lumicano y asporina. En una concentración todavía más baja se encuentran, el versicano y el perlecano que pertenecen a la familia del agrecano y se caracterizan por ser proteoglicanos de gran peso molecular<sup>50</sup>.

### 2.3.3 COMP

La proteína COMP (proteína oligomérica de matriz del cartílago) es una glicoproteína pentamérica que se encuentra en la ECM rodeando a los condrocitos. La función de COMP es permanecer unida a los colágenos de tipo II y III ayudando a su estabilización<sup>51</sup>. Es una proteína de la ECM de especial relevancia ya que mutaciones en el gen COMP provocan diferentes tipos de osteocondrodisplasias y los niveles en suero de esta proteína se encuentran elevados en los pacientes con OA de rodilla<sup>52,53</sup>.

### 2.3.4 Mecanismos de formación y degradación de la ECM

El mantenimiento de la integridad de la ECM se logra mediante la acción de una serie de moléculas anabólicas y catabólicas. Precisamente, cuando este equilibrio se altera a favor de rutas catabólicas comienza la degradación de la ECM, proceso clave en la OA<sup>54</sup>. A continuación, se detalla el papel de las principales moléculas que intervienen en los procesos de biosíntesis y degradación de la ECM.

#### **2.3.4.1 Biosíntesis de la ECM**

Entre las moléculas que favorecen la formación de ECM se encuentran diferentes factores de crecimiento como TGF- $\beta$ , IGFs y FGFs.

##### **2.3.4.1.1 Superfamilia TGF- $\beta$ (Transforming Growth Factor $\beta$ )**

Esta superfamilia de proteínas consta de 35 miembros que incluyen a TGF- $\beta$  y las BMPs. TGF- $\beta$ , es uno de los principales factores implicados en la biosíntesis de la ECM. Entre sus funciones se encuentran la de promover la diferenciación de los condrocitos y favorecer la síntesis de ECM (proteoglicanos y colágeno de tipo II) <sup>55</sup>.

TGF- $\beta$  tiene tres isoformas (TGF- $\beta$ 1, 2 y 3) que se expresan en el cartílago articular y cuya acción se ejerce a través de receptores situados en la superficie del condrocito. TGF- $\beta$ , se une al receptor TGF- $\beta$ RII el cuál recluta a TGF $\beta$ R1 que se activa mediante fosforilación. Después de la dimerización de los receptores, se desencadena la fosforilación intracelular de las proteínas SMAD-2 y 3 (*Small Mothers Against Decapentaplegic 2 y 3*) que se unen a SMAD-4 formando un heterocomplejo que se transloca al núcleo dónde regula la expresión de los genes diana. Esta ruta es inhibida por SMAD-6 y -7 <sup>56</sup>. Aunque la acción de TGF- $\beta$  es clave en el anabolismo del cartílago se ha constatado un papel diferente en el cartílago joven y en el cartílago envejecido. Con la edad se produce una alteración en la señalización de TGF- $\beta$  lo que induce diferenciación hipertrófica y aumenta la expresión de MMP-13 <sup>32</sup>. Además, estimula la formación de osteófitos y la fibrosis sinovial en el cartílago articular adulto <sup>57,58</sup>. Es posible que en la OA haya una alteración o disminución en la señalización por esta ruta ya que la expresión de sus tres isoformas se encontró aumentada en el cartílago articular con OA <sup>59</sup>.

## INTRODUCCIÓN

---

Las BMPs (*bone morphogenetic proteins*), se caracterizan por su efecto pleiotrópico. Tal como se ha tratado anteriormente, juegan un papel importante en el desarrollo de la articulación y del cartílago<sup>60</sup>. Sus funciones anabólicas en el cartílago son compartidas con TGF- $\beta$ ; estimular la síntesis de moléculas de la ECM y la diferenciación celular. Se ha descrito el papel de BMP-2, BMP-7 y BMP-13 en la diferenciación de las células madre mesenquimales de condrocitos y osteoblastos<sup>61</sup>. BMP-7 es una de las proteínas con más capacidad anabólica al poseer la capacidad de estimular la producción *in vitro* de proteoglicanos en condrocitos artrósicos<sup>62</sup>. BMP-5 es sin duda una de las BMPs más estudiada en OA tanto en estudios genéticos (3.2.2) como funcionales. Su papel es relevante en el crecimiento y reparación del cartílago así como en la regulación de la proliferación y diferenciación de los condrocitos<sup>63</sup>. El gen de *BMP-5* es regulado por elementos en *cis* y esta modulación de la expresión es clave para la formación del patrón del esqueleto en el embrión<sup>64</sup>. La regulación de las BMPs se ejerce mediante múltiples moléculas y a diferentes niveles. La proteína *noggin* es la principal antagonista de las BMP-2, -4 y -7. SMAD-6 y -7 también inhiben la acción de las BMPs al igual que otras moléculas como TOB y SMURF-1<sup>65</sup>.

Tanto TGF- $\beta$  como BMPs ejercen una función anabólica pero también pueden favorecer algún aspecto de la OA. Así por ejemplo, BMP-2 se encuentra expresada en condrocitos artrósicos concretamente en la capa superficial del cartílago<sup>66</sup>.



#### **2.3.4.1.2 IGFs (*Insuline like Growth Factors*)**

Esta familia consta de 6 proteínas de unión IGFBP (1-6) (*IGF binding proteins*), dos receptores IGFR (1 y 2) y dos ligandos IGF (1 y 2). IGF-1, es el miembro de la familia más conocido por su papel en el mantenimiento de la homeostasis en el cartílago estimulando a los condrocitos a producir proteínas de la ECM <sup>67</sup>. De hecho, en áreas de cartílago articular con elevada densidad de fibras en la ECM se encuentra un aumento de expresión del mRNA de IGF-1<sup>68</sup>. También se ha comprobado que IGF-1 aumenta la síntesis de colágeno de tipo II al controlar transcripcionalmente, conjuntamente con SOX-6 y SOX-9, un enhancer localizado en el primer intrón de COL2A1 <sup>69</sup>.

Las IGFBPs, secretadas por los condrocitos, controlan la accesibilidad al receptor de IGF-1 (IGFR1) situado en la superficie celular. La regulación del sistema es compleja e implica también al hueso subcondral donde la activación de IGF-1 conllevaría a su vez a una producción excesiva de IGFBP. La competición de estas proteínas de unión por IGF-1 lleva a la pérdida de ECM en el cartílago artrósico al no poder ejercer su función anabólica en el tejido<sup>70</sup>. En el cartílago artrósico se ha observado sobreexpresión de dos de las proteínas de unión en comparación con condrocitos de cartílago normal, IGFBP-3 (24x) e IGFBP-5 (16x) <sup>71</sup>.

#### **2.3.4.1.3 FGFs (*Fibroblast Growth Factors*)**

Los FGFs son polipéptidos de efecto pleiotrópico secretados por los condrocitos. Constituyen una familia proteica de 22 miembros y ejercen su acción a través de los receptores FGFR <sup>72,73</sup>.

FGF es un factor estimulador y mitogénico de la placa de crecimiento del cartílago y de la síntesis de proteoglicano en la misma <sup>74</sup>. Sin embargo, otros estudios han revelado su posible efecto catabólico al aumentar la síntesis de metaloproteasas (MMPs), agreganasas y óxido nítrico(NO) <sup>75</sup>.

El estudio de ratones *knockout* (KO) para diferentes FGFs ha demostrado su papel en la regulación de la formación de nueva ECM. Así, los ratones KO para FGF-2 son más propensos a desarrollar OA <sup>76</sup>. Por el contrario, deleciones de alguno de los receptores de la familia como FGFR-1, han mostrado un efecto condroprotector al provocar una disminución de la producción de MMP-13 y un aumento de FGFR-3<sup>77</sup>.

### 2.3.4.2 Destrucción de la ECM

Entre las moléculas que propician el catabolismo en el cartílago articular se encuentran diversas citoquinas proinflamatorias, el óxido nítrico (NO) y sus derivados así como las proteasas de matriz extracelular.

#### 2.3.4.2.1 Citoquinas (IL-1, TNF- $\alpha$ )

La IL-1 (*interleukin-1*), es la molécula pro-catabólica más potente que ejerce su acción en la ECM del cartílago. Esta citoquina junto con TNF- $\alpha$  (*tumor necrosis factor  $\alpha$* ) muestra la importancia de la inflamación en el desarrollo de la OA y su relación con la destrucción de la ECM. Existen dos proteínas solubles IL-1; IL-1 $\alpha$  e IL-1 $\beta$  que ejercen su acción al unirse a los receptores (IL1R1) situados en condrocitos y sinoviocitos. Además, también existe un proteína inhibidora, IL-1RA (*interleukin 1 receptor antagonist*), codificada por el gen IL1RN. El papel de IL-1RA es actuar

como inhibidor de la acción de IL-1 al interceptar su unión con los receptores IL1R1<sup>78-80</sup>. El papel de IL1-RA se ha estudiado como posible diana terapéutica contra la progresión de la OA. En concreto, mediante terapia génica se comprobó que un vector insertado en condrocitos portando IL-1RA protege de la degradación del cartílago articular que les rodea<sup>81</sup>.

TNF- $\alpha$  es una molécula que comparte propiedades catabólicas con IL-1<sup>82</sup>. Al igual que IL-1, TNF- $\alpha$  actúa aumentando la expresión de MMP-3. También hay que destacar el papel de TACE (*tumor necrosis factor- $\alpha$ -converting enzyme*) que es la enzima que procesa TNF- $\alpha$  a su forma activa. En el cartílago artrósico se ha detectado un aumento de TACE y también de TNF- $\alpha$ <sup>83</sup>.

#### **2.3.4.2.2 Óxido nítrico (NO) y ROS (especies reactivas de oxígeno)**

El NO se sintetiza a partir de la L-arginina y moléculas de O<sub>2</sub><sup>84</sup>. Es una molécula pleiotrópica con acciones a nivel de diferentes órganos y tejidos como en vasos sanguíneos y el cerebro. El papel del NO en el cartílago es de especial relevancia dado que los condrocitos son las células capaces de producir NO en mayores cantidades<sup>85</sup>. Los condrocitos poseen una cNOS (*constitutive nitric oxide synthase*) que es la encargada de mantener los niveles de NO en condiciones normales. Sin embargo, cuando existe inflamación, se activa iNOS (*inducible nitric oxide synthase*), generadora de elevadas cantidades de NO. Al aumentar la producción de NO se favorece la destrucción del cartílago articular y se inhibe la síntesis de proteoglicanos de la ECM<sup>86-90</sup>. El papel del NO en la OA está íntimamente ligado al de diferentes moléculas como IL-1 $\beta$ , TNF- $\alpha$  e

IFN- $\gamma$ . Estas moléculas actúan de manera autocrina/paracrina encargándose de activar su propia producción y la del NO <sup>91</sup>.

A pesar del papel del NO en el catabolismo de la ECM, cada vez existen más evidencias de que puede tener una posible función protectora de la destrucción de la ECM. Se ha comprobado en cultivos de condrocitos que el NO exógeno disminuye la activación inflamatoria mientras que los derivados reducidos del NO tienen el efecto contrario siendo potencialmente más dañinos<sup>92</sup>. Los derivados del NO, como el peroxinitrito, se forman a partir de la reacción de ROS (*reactive oxygen species*) con el NO, y tienen un efecto muy destructivo en el cartílago <sup>93</sup>. Así, el peroxinitrito es capaz de inducir apoptosis de los condrocitos mientras que el NO no es capaz de provocar el mismo efecto dañino <sup>94</sup>.

### **2.3.4.2.3 Proteasas de matriz extracelular**

Las proteasas de matriz extracelular son sintetizadas por el condrocito y su regulación es clave en el desplazamiento del equilibrio existente entre síntesis y degradación de los componentes de la ECM. La clasificación de las proteasas de matriz extracelular se presenta en la siguiente tabla (Tabla 1).

**Tabla 1. Clasificación de las proteasas de matriz extracelular basada en su mecanismo catalítico.**

**I-Metaloproteasas (metizincinas):**

**I.1- Metaloproteasas de ECM (MMPs)**

I.1.1- Colagenasas: MMP-1, -8, -13 y -18

I.1.2- Estromelisininas: MMP-3, -10 y -11

I.1.3-Gelatinasas: MMP-2 y -9

I.1.4-Matrilisininas: MMP-7 y -26

I.1.5-De membrana: MTP-MMP-14, -15, -16, -17, -23, -24 y -25

I.1.6-Otras: MMP-12, -19, -20, -21, -27 y -28

I.2-ADAMTS: ADAMTS-5, -4, -1, -15, -9

I.3-ADAMs: ADAM-12, -10, -15 y -13

**II-Serín proteasas**

II.1-Miembros del sistema plasminógeno-plasmina: PLAT, u-PA, PAI

II.2-Otras: catepsina G, elastasa de PMN (polimorfonucleares)

II.3-Matriptasa

**III-Cisteín proteasas**

III.1-Catepsinas -B, -H, -K, -L y -S

III.1-Calpaínas

**IV-Glutamyl proteasas**

**V-Treonín proteasas**

**VI-Aspartil proteasas**

**2.3.4.2.3.1 Metaloproteasas (MMPs)**

Las metaloproteasas son una familia de proteasas muy amplia y con múltiples funciones. Algunas tienen la capacidad de romper y degradar colágeno de tipo II y agregano, dos de los componentes mayoritarios de la ECM<sup>95</sup>. Entre las metaloproteasas se deben destacar dos grupos: las adamolisinas (ADAMs y ADAMTS) y las metaloproteasas de matriz extracelular. Los dos grupos tienen un átomo de zinc en su dominio catalítico.

*Metaloproteasas de matriz extracelular*

Las metaloproteasas de matriz extracelular se sintetizan como proenzimas que se secretan de forma inactiva y que deben ser

## INTRODUCCIÓN

---

procesadas para poder llevar a cabo su actividad. Su estructura consta de varios motivos: péptido señal, prodominio, dominio catalítico que incluye un átomo de Zn junto con una metionina conservada y un dominio C-terminal específico<sup>96</sup>. El prodominio se caracteriza por poseer la secuencia denominada *cysteine switch*, PRCGXPD. Esta región mantiene a las MMPs en su forma inactiva de pro-MMPs<sup>97</sup>. El dominio catalítico consta de unos 170 aminoácidos (aa) y en el caso de MMP-2 y MMP-9 también incluye 3 dominios fibronectina que permiten interactuar con el colágeno<sup>98</sup>. El dominio C-terminal (hasta 210 aa) es variable y específico de cada MMP dependiendo de su función.

Las MMPs se pueden clasificar por su acción catalítica. Las colagenasas (MMP-1, -8 y -13) degradan principalmente el colágeno, siendo MMP-13 específica del cartílago artrósico y la proteasa más importante por su acción degradativa. La acción de MMP-13 se estudió en ratones en un modelo inducido de OA. Los ratones deficientes en esta metaloproteasa muestran protección frente a la erosión del cartílago comparado con los ratones *wild-type*<sup>99</sup>. En la degradación proteolítica del colágeno también pueden intervenir la gelatinasa A (MMP-2) y MMP-14 que es una proteína transmembrana<sup>100</sup>. El agregano aunque es procesado mayoritariamente por ADAMTS (se trata en el próximo apartado) también puede ser degradado por MMPs en una región diferente del dominio interglobular<sup>101,102</sup>.

La expresión de las MMPs se regula a nivel transcripcional mientras que diferentes mecanismos post-transcripcionales están implicados en el control de su función. Los promotores de las MMPs se regulan en *cis* por una serie de trans-activadores siendo el sitio de regulación más

importante, el de unión a AP-1, localizado unas 70 pb *upstream* del TSS (sitio de inicio de la transcripción)<sup>103,104</sup>. Otros sitios de unión son los denominados PEA3 que unen factores de transcripción Ets, los cuáles junto con otras proteínas como c-fos ó c-jun influyen en la expresión de las MMPs <sup>105</sup>. A nivel post-transcripcional, las MMPs pueden ver regulada su función a través de moléculas como TGF- $\beta$  y de hormonas como el cortisol. Los mi-RNA, también intervienen en la regulación de las MMPs. Son moléculas de unos 22 nucleótidos que se unen a secuencias complementarias del mRNA reprimiendo su traducción. Uno de los miRNAs importantes en la patogénesis de la OA es el miRNA-140 que actúa como regulador negativo de MMP-13 <sup>106</sup>.

#### 2.3.4.2.3.2 ADAMTS, ADAMS y TIMPs

La familia de las ADAMTS (*A Disintegrin And Metalloproteinase with Thrombospondin Motifs*) consta de 19 proteínas con una estructura en múltiples dominios. Algunos de ellos son dominios típicos de proteasas y otros característicos de esta familia: el de similitud con la desintegrina, una serie de repeticiones trombospondina (TSR), un dominio rico en cisteína (CRD), un dominio espaciador de longitud variable y una serie de módulos C-terminales que pueden ser dominios PLAC o dominios CUB dependiendo de los miembros de la familia <sup>107</sup>

La principal función de las ADAMTS en el cartilago articular es participar en la degradación del agregano <sup>108</sup>. Las ADAMTS-1, -4, -5 y -9 son las denominadas agrecanasas de las cuáles, la ADAMTS-4 (agrecanasa 1) y la ADAMTS-5 (agrecanasa 2) son las principales agrecanasas *in vivo*. Se caracterizan por romper las moléculas de agregano en la unión E<sub>373</sub>↓<sub>374</sub>A en el dominio interglobular<sup>109</sup>. Su papel en la

## INTRODUCCIÓN

---

degradación del agrecano se comprobó mediante siRNAs (RNA de silenciamiento) específicos para estas agrecanasas en explantes de cartílago OA en los cuáles se redujo la degradación del agrecano <sup>110</sup>.

Estudios en ratones *KO* (*knockout*) resaltan la importancia de las agrecanasas y ADAMTS-5 en la patogénesis de OA. Los ratones deficientes para ADAMTS-5 muestran protección frente a la degradación del cartílago y la pérdida de agrecano. Esto no se observa en los ratones deficientes para ADAMTS-4. ADAMTS-4 es una agrecanasa que se expresa sólo en cartílago artrósico mientras que ADAMTS-5 se expresa tanto en cartílago normal como artrósico <sup>111-113</sup>. El ratón doble knockout para las dos agrecanasas muestra una protección similar contra la degradación del agrecano que el knockout de ADAMTS-5, lo cuál parece evidenciar un papel más relevante de esta agrecanasa en la patogénesis de la OA aunque en el humano el papel de ambas es similar <sup>114</sup>.

La regulación de las agrecanasas se ejerce mediante moléculas como NFkB y en el caso de ADAMTS-4 también por una regulación epigenética que provoca la desmetilación de sitios específicos del promotor <sup>115,116</sup>. El proceso de agrecanolisis que llevan a cabo las ADAMTS se considera el primer evento en el desarrollo de la OA siendo previo a la colágenolisis llevada a cabo por MMPs <sup>117</sup>. IL-1 y TNF- $\alpha$  estimulan la pérdida de agrecano al activar a las agrecanasas pero no son suficientes para activar la ruptura del colágeno por MMPs <sup>118,119</sup>. Sin embargo, otros autores afirman que la ruptura del colágeno sería el primer proceso degradativo en la OA temprana <sup>120</sup>.

Las ADAM se diferencian de las ADAMTS en que carecen del dominio trombospondina en la zona C-terminal y comparten funciones con las



ADAMTS en la degradación de la ECM <sup>121</sup>. La expresión de diferentes ADAMs, ADAM-12, ADAM-10 y ADAM-15, se encuentra aumentada en cartílago artrósico respecto al normal <sup>122-124</sup>.

En la regulación de la acción de MMPs, ADAMTS y ADAM interviene los TIMPS (inhibidores tisulares de las metaloproteasas). Existen cuatro TIMPs: TIMP-1, TIMP-2, TIMP-3 y TIMP-4. La producción de TIMPs y su acción aparece ligada a la producción de moléculas de la ECM <sup>95</sup>. Los TIMPs también son los inhibidores más efectivos de las ADAMTS. ADAMTS-4 y-5 son inhibidas por TIMP-3 aunque son insensibles a los demás miembros de TIMPs <sup>125</sup>.

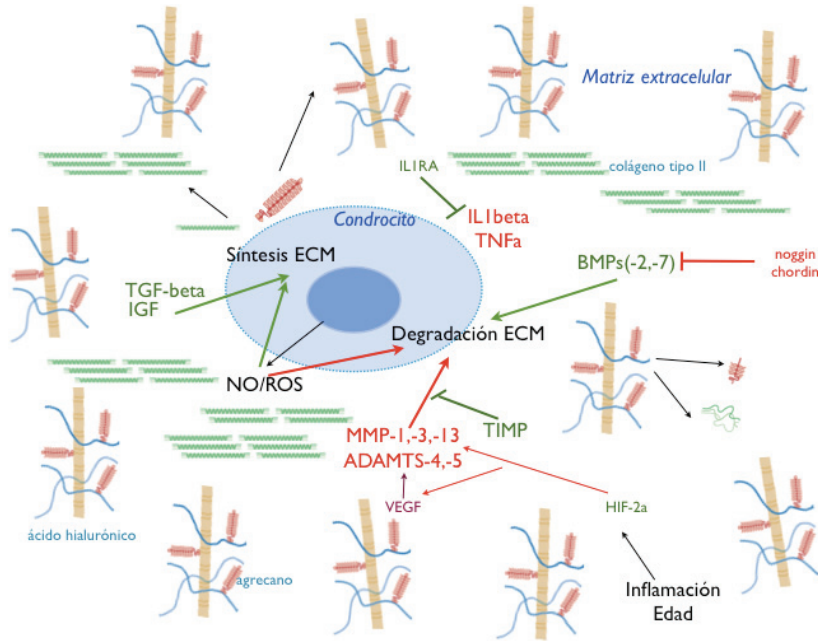
#### *2.3.4.2.3.2 Serín proteasas*

Las serín proteasas se caracterizan por poseer una serina en su centro activo. Conforman la familia un amplio grupo de proteínas de las cuáles algunas forman parte del sistema plasminógeno-plasmina. La plasmina es una serín proteasa de dos cadenas cuyo zimógeno es el plasminógeno y su función es degradar la fibrina y componentes de la ECM. El activador del plasminógeno puede ser de tipo urokinasa (u-PA) o el denominado tisular (t-PA), siendo ambas proteasas las que se encargan de la conversión de plasminógeno a plasmina <sup>126</sup>. Las concentraciones de uPA, tPA son más elevadas en el fluido sinovial de OA de rodilla que en la articulación no afectada lo que revela su importancia en el proceso catabólico <sup>127</sup>. Además de los componentes citados anteriormente también se encuentran los receptores activadores de plasminógeno (PAR) encargados de regular la proteólisis mediada por serín proteasas. Su expresión también se encontró aumentada en condrocitos con OA y regulan diferentes metaloproteasas <sup>128</sup>. Además de las moléculas citadas anteriormente, recientemente, se ha

descubierto como otra serín proteasa, la matripsina que tiene especial relevancia en la destrucción del cartílago. Se encontró aumentada su expresión en el cartílago artrósico y está implicada en el procesamiento de la pro-MMP3 a su forma activa <sup>129</sup>.

### **2.3.5 Otros mecanismos biológicos que participan en el desarrollo de la OA: inflamación y angiogénesis.**

La integración de diferentes señales inflamatorias o relacionadas con la edad, es crucial para la regulación de diferentes factores anabólicos (TGF- $\beta$ , factores de crecimiento) y catabólicos (IL-1, proteasas de matriz extracelular). Estos factores desempeñan un papel central en la patogénesis de la OA al ser las moléculas efectoras de la síntesis o degradación de los distintos componentes de la ECM (Figura 7). La OA se considera clásicamente como una enfermedad degenerativa cuyo evento central es la destrucción de la ECM del cartílago. Sin embargo, también ha cobrado especial relevancia el papel que desempeña el proceso inflamatorio en la iniciación y progresión de la enfermedad. Moléculas inflamatorias como TNF- $\alpha$  e IL-1 estimulan la síntesis por parte de los condrocitos de metaloproteasas. Además de estas moléculas, se ha descrito recientemente en estudios en ratones, la implicación del complemento en el desarrollo de la OA. El sistema del complemento potencia la respuesta inflamatoria induciendo la expresión de metaloproteasas, mediadores inflamatorios y efectores del complemento <sup>130</sup>.



**Figura 7. Esquema que resume las relaciones entre las principales moléculas anabólicas (verde) y catabólicas (rojo) en el mantenimiento de la integridad de la ECM**

Junto con la inflamación, la angiogénesis también influye en la progresión de la OA estando ambos procesos íntimamente relacionados. Por una parte, la angiogénesis puede facilitar la aparición de inflamación y a su vez la inflamación puede estimular el proceso angiogénico<sup>131</sup>. Se debe destacar que el cartílago normal adulto es avascular mientras que el cartílago artrósico puede presentar invasión de algunos vasos sanguíneos de las capas profundas a través de la capa calcificada. El principal factor angiogénico es VEGF (*vascular endothelial growth factor*) cuya producción se ve inducida por la acción de diferentes moléculas mediadoras como HIFs (*hypoxia inducible factors*) y RUNX-2 (*Runt-related transcription factor*). VEGF además de provocar la

## *INTRODUCCIÓN*

---

vascularización del cartílago induce la síntesis de metaloproteasas y ADAMTSs<sup>132,133</sup>.

En conclusión, los procesos de inflamación y angiogénesis interactúan y se regulan mutuamente, rompiendo el equilibrio existente en la ECM a favor de moléculas catabólicas que promueven su degradación.



### *3. Genética de la OA*

Los primeros estudios genéticos que han permitido mostrar el componente hereditario de la OA y su naturaleza poligénica han sido los estudios de segregación en familias, estudios de gemelos y de riesgo relativo entre hermanos <sup>134</sup>.

Los estudios de gemelos se basan en comparar la similaridad entre gemelos monozigotos (MZ) y dizigotos (DZ). A partir del porcentaje de concordancia fenotípica en gemelos MZ (100% de identidad genética) se puede calcular la heredabilidad. La concordancia entre gemelos DZ se utiliza para corregir la estimación anterior. La heredabilidad, cuantifica la importancia de los factores genéticos de la enfermedad y es diferente en función de la articulación afectada. Para OA de rodilla y OA de mano se sitúa en torno al 40-65%, mientras que para OA de cadera sería aproximadamente de un 60%. La OA de columna es la que presenta un heredabilidad más elevada de un 70% <sup>22</sup>.

Otro tipo de estudios de heredabilidad son los estudios de riesgo relativo de los hermanos ( $\lambda_s$ ) que cuantifican la heredabilidad basándose en el riesgo de padecer la enfermedad si se tiene un hermano enfermo respecto al riesgo de la población general. Se han realizado diferentes estudios de este tipo en OA para las diferentes articulaciones. En OA de cadera se estima un riesgo relativo entre 2 y 5 veces mayor entre hermanos respecto a la población general dependiendo del estudio y de la población considerada <sup>135-137</sup>. Para OA de mano se ha estimado un  $\lambda_s$  de 5 <sup>138</sup> mientras que para OA de rodilla y de cadera es de 2,32 <sup>139</sup>.

## INTRODUCCIÓN

---

Estos trabajos han sido los pioneros en el campo genético y han dado paso al inicio de otras estrategias de estudio para aclarar los factores genéticos de susceptibilidad a la OA: estudios de genes candidato, estudios de ligamiento de genoma completo, estudios de asociación caso-control, GWAS (estudios de asociación de genoma completo) y los meta-análisis de los mismos.

Si hay que destacar alguna característica del estudio de la genética de la OA es su evolución en etapas marcadas por el desarrollo de nuevas tecnologías. En los estudios de asociación se descubrieron múltiples polimorfismos asociados con un efecto pequeño repartidos por todo el genoma. Con el desarrollo de la tecnología de los GWAS se esperaba confirmar parte de estas señales y/o descubrir nuevos SNPs asociados. Sin embargo, los resultados de los GWAS señalaron un único gen ya descubierto en los estudios de asociación (*GDF-5*) y descubrieron un número relativamente bajo de nuevas señales que alcanzan umbral de asociación requerido en estos estudios ( $p=5 \times 10^{-8}$ )<sup>140,141</sup>. Esta situación es particular en comparación con otras enfermedades complejas en las que se encuentra un número más elevado de señales asociadas<sup>142</sup>. Parte de los problemas para encontrar nuevos loci asociados viene dada en parte por la propia naturaleza poligénica de la enfermedad. Así, múltiples polimorfismos con distintas frecuencias alélicas contribuyen con efectos pequeños o modestos. Sin embargo, también existen otras dificultades específicas de la OA que se deben fundamentalmente a la heterogeneidad de la enfermedad y la definición de controles.

El estudio de la primera fase de arcOGEN reveló que la arquitectura genética de la OA es poligénica. Este análisis vuelve a dejar claro que en

la etiología de la OA, no se encuentran SNPs asociados de gran efecto como en otras enfermedades complejas<sup>143</sup>.

En relación con la heterogeneidad de la enfermedad, se ha comprobado que un mismo factor genético no tiene igual efecto en la susceptibilidad a la enfermedad en distintas poblaciones. Esto se refleja claramente en los estudios genéticos realizados en población europea y asiática<sup>144</sup>. También hay que tener en cuenta que los factores genéticos de susceptibilidad son específicos de articulación (OA de rodilla, cadera y mano) y de sexo. Esta compartimentalización dificulta el estudio de la genética de la OA sobre todo a la hora de intentar reunir un amplio tamaño muestral ya que éste se ve a la fuerza reducido por las múltiples estratificaciones requeridas en los análisis.

Otro tipo de heterogeneidad viene introducida por los fenotipos de los pacientes (OA radiográfica vs sintomática). Recientemente se ha recomendado una estandarización en los futuros estudios genéticos para los pacientes introducidos según presenten uno u otro fenotipo. Para ello se recomienda que en los futuros GWAS se estratifiquen los análisis por edad, sexo, BMI y presencia de OA radiográfica o sintomática. Dentro de la OA radiográfica (ROA), se deben definir los fenotipos en función de otra serie de características como el número de osteófitos (ROA de rodilla y mano) y la presencia de estrechamiento articular (ROA de cadera)<sup>145</sup>. La importancia de seguir esta serie de estandarizaciones se debe a que introducir fenotipos definidos de forma diferente reduce el poder estadístico de encontrar asociaciones genéticas consistentes en OA<sup>146</sup>.

En relación con la definición de controles, la principal dificultad radica en conseguir controles libres de OA. La OA es una enfermedad de inicio

## *INTRODUCCIÓN*

---

tardío y de gran prevalencia en la población. Si los casos incluidos en el estudio genético se han seleccionado según criterio radiográfico y tienen una edad avanzada parece lógico pensar que también lo deberían estar los controles. El problema es que es realmente complicado justificar una exploración radiográfica en un individuo sano y muchas veces la OA es asintomática. Además, si los controles que se seleccionan no tienen una edad avanzada es imposible excluir que desarrollarán OA en el futuro cercano y por lo tanto pueden tener una posible carga genética similar a la de los pacientes. Se ha evaluado cómo influye utilizar controles poblacionales o controles libres de OA en la primera fase del estudio de arcOGEN. Para ello se han utilizado los resultados de asociación para tres SNPs en loci relevantes en OA (*GDF-5*, *MCF2L* y el locus 7q22), así como las nuevas señales que se descubren en este trabajo. Sólo para el SNP de *GDF-5* se observó un efecto mayor y un p-valor más significativo al utilizar controles libres de OA. En el resto de los SNPs analizados se obtuvieron idénticos resultados o en algunos casos mejores, al utilizar los controles poblacionales. Por lo que se concluyó que utilizar controles libres de OA no produce una mejoría detectable en los resultados de asociación<sup>141</sup>. Sin embargo, este resultado no puede considerarse definitivo ya que los controles poblacionales utilizados son los mismos utilizados para descubrir muchas de estas asociaciones.

A continuación se describen las principales estrategias de estudio que se han llevado a cabo en el campo de la genética de la OA.



### 3.1 Estudios de ligamiento.

Un estudio de ligamiento evalúa la cosegregación de un conjunto de marcadores genéticos con la enfermedad en familias. Los marcadores genéticos más utilizados para realizar estos estudios son los microsatélites, secuencias nucleotídicas de 2-6 p.b que se repiten de manera consecutiva y que son muy polimórficas. La existencia de ligamiento de un determinado locus con la enfermedad se comprueba si varios familiares afectados comparten la misma variante alélica para el microsatélite. Una vez localizada la región genética que segrega con la enfermedad, ésta se satura con más microsatélites hasta que la región de ligamiento se acorte y los genes se puedan priorizar por su función (identificación de genes candidato en loci de ligamiento).

Los estudios de ligamiento en OA se llevaron a cabo en diferentes cohortes con distinto fenotipo:

1-OA de mano (Framingham (EEUU), Islandia y Finlandia)

2-OA de rodilla y/o cadera (Reino Unido, GARP (Holanda))

Los estudios de ligamiento han llevado a la identificación de diversas regiones de ligamiento en OA de cadera, rodilla y mano, aunque la mayoría de los hallazgos no han tenido una repercusión importante ni han sido replicados posteriormente. Esto es debido a que las regiones de ligamiento son normalmente muy extensas por lo que identificar la variante de susceptibilidad es complicado. Además, los *loci* detectados tienen un efecto débil y el tamaño de las colecciones muestrales utilizado

fue pequeño.

Dentro de las regiones genómicas delimitadas en los estudios de ligamiento se seleccionan a posteriori los genes candidatos más probables incluidos en ellas para realizar estudios de asociación caso-control (estudios de asociación caso-control de genes candidato). Este tipo de estudios utiliza colecciones caso-control reuniendo un mayor tamaño muestral, por lo tanto, confiriendo un mayor poder estadístico.

### **3.2 Estudios de genes candidato**

Los estudios de genes candidato intentan buscar genes diana basándose en los mecanismos moleculares implicados en el inicio y la progresión de la OA. Tomando como base la etiología de la OA, se seleccionan genes en los que la modulación de su función, podría ser clave en la susceptibilidad a la enfermedad.

En los estudios de genes candidato se realizan estudios de asociación caso-control. En ellos se comparan las frecuencias alélicas de los polimorfismos de interés localizados en los genes candidatos seleccionados, en individuos con OA no emparentados y controles libres de la enfermedad. Si la frecuencia alélica de un determinado polimorfismo se encuentra significativamente elevada en los casos, se considera asociado con OA. Los estudios de asociación en OA también pueden ser estudios de cohortes. Este diseño no afecta al tipo de análisis estadístico de los datos pero sí a la selección de casos y controles. Son poblaciones más homogéneas y seguidas prospectivamente que permiten evaluar la aparición de nuevos casos de OA. El problema de estos trabajos

es que no son viables para el estudio de OA avanzadas que requerirían un tiempo de seguimiento muy largo.

Para el análisis estadístico se construyen tablas de contingencia 2x2 con las frecuencias de los alelos y se realiza un test  $\chi^2$ . Una de las medidas que tiene especial relevancia en los estudios de asociación caso-control es la OR (*odds ratio*) y su intervalo de confianza (95% IC). La OR utiliza la distribución de alelos para proporcionar una estimación de la magnitud de la asociación. Cuando la OR tiene un valor menor de 1 el alelo es de protección, si tiene un valor mayor de 1 el alelo es de susceptibilidad.

Los genes candidatos estudiados en OA se pueden dividir de una forma genérica en tres grandes grupos:

- 1-Genes que codifican componentes de la ECM.
- 2-Genes que codifican reguladores del anabolismo o catabolismo de la ECM.
- 3-Genes que codifican moléculas inflamatorias.

Los estudios de asociación caso-control tienen capacidad de detectar efectos pequeños con facilidad y las colecciones muestrales pueden alcanzar un tamaño relativamente grande. Una de sus aportaciones más importantes es a la vez una desventaja. Este tipo de estudios genéticos se centran *a priori* en una selección de polimorfismos localizados en genes de interés por su función pero a la vez se está excluyendo del estudio toda la variabilidad genética restante existente en el genoma. Esta limitación es solventada por otro tipo de estudios genéticos, los GWAS que relegaron a un segundo plano a los estudios de asociación caso-control.

También se debe comentar que sólo uno de los hallazgos de estos estudios (*GDF-5*) ha sido replicado posteriormente alcanzando el umbral de significación requerido en los GWAS. Esto indica que los resultados de los estudios de genes candidatos se interpretaron con optimismo excesivo.

### 3.3 Estudios de asociación de genoma completo (GWAS)

Los GWAS permiten evaluar la variabilidad genética a lo largo de todo el genoma sin hipótesis previa. Se intenta estudiar y capturar toda la variabilidad del genoma humano basándose en la existencia de LD (*linkage disequilibrium*) a lo largo de éste. El genoma humano se encuentra organizado en bloques de LD que muestran una baja frecuencia de recombinación separados por puntos de elevada recombinación. En los bloques hay información redundante entre los SNPs. Si se seleccionan sólo los SNPs informativos (tagSNPs) se obtiene una representación de la región sin necesidad de obtener los alelos de todas las variantes. Se pueden conocer cuáles son los tagSNPs característicos de una región genómica gracias al proyecto HapMap que estudió un gran número de SNPs en individuos de varias poblaciones<sup>147</sup>.

Basándose en esta información, los *arrays* de genotipado de alta densidad utilizados en la mayoría de los GWAS contienen entre 300000 y 500000 SNPs representativos (*Illumina Human Hap300* y *Affymetrix 500K*, respectivamente). Este número es suficiente para proporcionar una cobertura lo más completa posible de la variabilidad presente en el genoma. Recientemente se ha comenzado a utilizar otros arrays que pueden cubrir mayor número de SNPs, hasta 900000 y además genotipar CNVs (variaciones en el número de copias). El trabajo a gran escala que

suponen los GWAS desarrolló en paralelo técnicas informáticas que permiten el manejo y análisis de la gran cantidad de datos generados.

En los GWAS se comparan la variación de las frecuencias alélicas de miles de SNPs genotipados, entre dos grupos de individuos, no afectos (controles) y afectos (casos). Si la frecuencia alélica de un SNP es mayor en los casos que en los controles, la variante genética se considera asociada con la enfermedad. El umbral de significación en los GWAS al corregir por múltiples tests, se considera  $p < 5 \times 10^{-8}$ . Este umbral viene dado por la estimación de que existen aproximadamente  $10^6$  SNPs independientes en el genoma humano ( $0,05/10^6 = 5 \times 10^{-8}$ ). Este umbral tiene la ventaja de reducir la tasa de posibles falsos positivos. Esta exigencia en el umbral de significación combinada con colecciones muestrales de elevado tamaño permite obtener resultados sólidos con suficiente potencia estadística para detectar efectos modestos. En general, se puede decir que la potencia de un GWAS depende de diferentes factores: el tamaño muestral de las colecciones utilizadas, el tamaño del efecto de los SNPs y su frecuencia del alelo menor.

El GWAS que demostró la importancia y utilidad de esta estrategia ha sido el llevado a cabo por el Wellcome Trust en el año 2007, en el que se incluyeron 14000 casos de siete enfermedades complejas y 3000 controles compartidos. El genotipado incluyó hasta 500568 SNPs y se consideró un umbral de significación de  $p < 5 \times 10^{-7}$ <sup>148</sup>. Después de este trabajo se realizaron más de 600 GWAS en unas 150 enfermedades diferentes, encontrándose alrededor de 800 SNPs con p-valor  $< 5 \times 10^{-8}$  (umbral de significación)<sup>149</sup>.

## INTRODUCCIÓN

---

En OA, al igual que en el resto de las enfermedades complejas, el estudio de las variantes genéticas se focalizó en los GWAS. El número de loci asociados encontrados es relativamente bajo y sólo tres de ellos se consideraban hallazgos sólidos hasta hace relativamente poco tiempo, el gen *GDF-5*, el locus 7q22 y *MCF2L*<sup>150-152</sup>. El número de loci continúa aumentando paulatinamente. Recientemente, se han incorporado los descubiertos en el estudio de arcOGEN (*GNL3*, *GLT8D1*, *ASTN2*, *FILIP1-SEN6*, *KLHDC-5-PTHLH*, *CHST11*, *TP63* y *SUPT3H-CDC5L*). Aún así, el número resultante de loci asociados es relativamente bajo lo que se atribuye comúnmente a dos factores: el tamaño muestral de las colecciones disponibles y la heterogeneidad de fenotipos en la OA<sup>153</sup>.

En relación al efecto del tamaño muestral se puede comentar un ejemplo concreto con el meta-análisis que reveló asociación en el locus 7q22 con OA de rodilla. Las colecciones de muestras incluidas reunían un total de 6709 casos y 35909 controles. Aunque pueda parecer un tamaño muestral elevado, el poder estadístico de este estudio para descubrir variantes de efectos pequeños ( $OR=1-1.15$ ) no es suficiente. En este caso particular se necesitarían 7000 muestras más de OA de rodilla, para detectar el efecto de SNPs con  $MAF=20\%$  y una  $OR$  de 1.15. En resumen, los GWAS de OA detectan bien variantes comunes de efectos fuertes ( $OR>1.15$ ) pero las colecciones muestrales disponibles no son suficientes para identificar efectos más modestos<sup>151</sup>.

Uno de los problemas que se plantea en el estudio genético de la OA es que no se trata de un fenotipo definido y existen grandes diferencias en la clasificación de los pacientes. La heterogeneidad en la definición de los fenotipos incluidos en los GWAS es otro factor que reduce el poder

estadístico de encontrar asociaciones consistentes. Concretamente se ha comprobado que las asociaciones entre OA sintomática y edad, sexo ó BMI son diferentes de las de la OA radiográfica. Por ello se propone realizar una estandarización en los GWAS de OA y realizar una estratificación de los diferentes fenotipos incluidos. Esta estrategia reduciría el tamaño muestral pero a la vez podría ganar en poder estadístico al reducir la heterogeneidad<sup>154</sup>.

Centrar los GWAS en endofenotipos particulares también puede ser una forma de aumentar el poder de detección de nuevas variantes. Los endofenotipos son fenotipos internos que están más próximos a la etiología biológica de la enfermedad que a sus signos o síntomas, por lo que se considera que tienen una influencia genética más fuerte que ambiental. Los endofenotipos en OA son características del cartílago o la forma de la articulación. Se ha llevado a cabo un GWAS con un endofenotipo particular que mide la anchura del espacio articular en las radiografías, JSW (*joint space width*) y que refleja la existencia de adelgazamiento del cartílago. En este trabajo se descubre un nuevo locus asociado, *DOTIL* (3.4.18), que alcanza el nivel de significación requerido en los GWAS<sup>155</sup>

Otro de los fenotipos que se propone estudiar por separado es el dolor de la articulación que es una de las principales características de la OA sintomática. Así, se descubrió la asociación entre *PACE-4*, seleccionando casos de OA de rodilla que presentan dolor y OA radiográfica y dos grupos de controles; los que no presentan dolor ni OA radiográfica y los que no presentan dolor y sí OA radiográfica<sup>156</sup>.

Por otra parte, se debe ser consciente de que los GWAS tienen algunas limitaciones. Así, los arrays de genotipado utilizados no cubren de una manera tan completa la variación genética determinada por los SNPs con frecuencia  $<5\%$ , debido a que presentan un LD muy débil con otros SNPs. Además, ni las variantes genéticas raras con MAF  $<0.5\%$ , ni las variaciones en número de copias (VNTR) son incluidas en los arrays lo que resulta en un estudio incompleto de la variabilidad genética<sup>157</sup>. Otra de las limitaciones de los GWAS, es que los SNPs que pueden encontrarse asociados no tienen porque ser las variantes causales de la enfermedad. Para poder obtener toda esta información se debe realizar un mapeado fino del gen, mediante resecuenciación a la par que se coteja esta información con la contenida en el proyecto ENCODE (ENCyclopedia Of DNA elements) recientemente finalizado. La información proporcionada es esencial para conocer la posible funcionalidad de los SNPs asociados y poder validarla en posteriores experimentos<sup>158</sup>.

### **3.4 Factores genéticos de susceptibilidad a la OA**

En este apartado se reúne la información de los diferentes loci que han sido objeto de estudio en la genética de la OA en los últimos años. Muchos de ellos se consideraron relevantes en el momento de su descubrimiento. Sin embargo, se puede decir que en la actualidad la mayoría de las señales de genes candidatos o estudios de ligamiento no se han replicado en los GWAS. Por otro lado, muchos de estos factores genéticos no se pueden excluir de forma definitiva por falta de estudios en todos los fenotipos de OA.



### 3.4.1 VNTR en asporina que afecta a la actividad de TGF- $\beta$

La asporina es una proteína que forma parte de la familia de los SLRPs al igual que la decorina y el biglicano. La asporina comparte homología de secuencia aminoacídica con la decorina. Sin embargo, en el extremo N-terminal de la asporina se localiza una cadena de ácido aspártico (10-19) variable en longitud. Las diferencias en el número de repeticiones del triplete nucleotídico vienen determinadas por un VNTR (*variable number tandem repeat*) y originan una cadena variable en número de ácido aspártico en la proteína. El primer estudio que identificó este VNTR de asporina como un factor genético de susceptibilidad a la OA se realizó en población japonesa. La frecuencia del alelo D14 (14 repeticiones) se encontró elevada en los pacientes con OA de rodilla respecto a los controles ( $p=0.000066$ ). La repetición D13 mostró un efecto menor pero protector, encontrándose su frecuencia elevada en controles<sup>159</sup>. Un estudio en población de Reino Unido, mostró una menor asociación ( $p=0.016$ ) aunque con la misma tendencia que en población asiática<sup>160</sup>. Un metaanálisis de seis estudios de asociación en cohortes europeas y asiáticas mostró una elevada heterogeneidad entre ambas. Así, se demuestra la existencia de un componente étnico diferencial que potencia el efecto de susceptibilidad del alelo D14 en población asiática respecto a la europea<sup>161</sup>. Esta diferencia de efecto podría deberse en parte a una combinación del efecto del alelo D14 con otros polimorfismos de susceptibilidad que tienen mayor frecuencia en asiáticos que en europeos. El principal papel funcional de la asporina es modular la vía de señalización de TGF- $\beta$  uniéndose a él en la ECM y evitando así su acoplamiento al receptor TGF- $\beta$ RI<sup>162,163</sup>. En el mismo trabajo en el que se

descubrió su asociación se comprobó *in vitro* que la asporina D14 inhibe la expresión de TGF- $\beta$  más fuertemente que la D13. El papel funcional de esta unión diferencial de la asporina D13 y D14 a TGF- $\beta$ , concuerda perfectamente con la asociación genética descubierta para los alelos D13 (protección) y D14 (susceptibilidad)<sup>159</sup>.

La asporina se considera un factor genético de susceptibilidad más relevante en población asiática que en europea aunque en ésta última no se ha estudiado a nivel de GWAS al ser un microsatélite y no estar incluido en los chips de genotipado. Su importancia en europeos no se puede descartar aunque la actualidad aunque su relevancia ha quedado relegada a un segundo plano tras el descubrimiento de los SNPs asociados en los GWAS.

### **3.4.2 Un polimorfismo en el promotor de *CALM-1* con asociación no confirmada**

La calmodulina es una proteína mediadora en la señalización por calcio que regula multitud de procesos celulares. La primera asociación genética con *CALMI* se encontró en población japonesa con un SNP localizado en el promotor del gen, rs12665715 C/T ( $p= 9.8 \times 10^{-7}$ ). El alelo T de susceptibilidad de este SNP medió una reducción de la expresión del gen en los experimentos funcionales<sup>164</sup>. Sin embargo, pese a este prometedor hallazgo en japoneses la asociación de este polimorfismo no se confirmó en estudios de Reino Unido ni de Grecia<sup>165,166</sup>. Además, un estudio de ligamiento de genoma completo destacó la región en la que se encuentra *CALM-1* junto con *DIO-2*. Esta zona se cubrió con tag-SNPs para un

estudio más detallado. Ningún SNP de *CALM-1* mostró asociación aunque SNPs localizados en *DIO-2* resultaron significativos<sup>167</sup>.

La calmodulina se consideró en el momento de su descubrimiento un buen candidato de estudio. Se expresa en condrocitos y modula su adherencia a la ECM. Además, participa en el metabolismo del cartilago en las vías inducidas mediante mecanotransducción<sup>168</sup>. Sin embargo, al no haberse confirmado la asociación de *CALMI* con OA en ningún estudio posterior, ni en población china ni en los GWAS, esta asociación debe considerarse como muy dudosa<sup>169</sup>.

### **3.4.3 Evidencia inconcluyente de asociación de un polimorfismo en *SMAD3***

*SMAD3* es un mensajero intracelular de la ruta del TGF- $\beta$  que media su acción reguladora de la transcripción<sup>170</sup>. En un estudio realizado con pacientes y controles de Reino Unido, se encontró asociación del rs12901499 (intrón 1) con OA de rodilla ( $p=7.5 \times 10^{-6}$ ) y de cadera ( $p=4 \times 10^{-4}$ )<sup>170</sup>. En relación al papel funcional de *SMAD3*, numerosos estudios han relacionado un descenso de la actividad de la proteína con la aparición de un fenotipo artrósico en modelos animales<sup>171</sup>. Recientemente, se encontraron mutaciones en *SMAD3* relacionadas con AOS, un síndrome de herencia autosómica recesiva que se caracteriza por anomalías esqueléticas, artrosis de inicio temprano y la presencia de aneurismas<sup>172,173</sup>. Pese a la importancia de *SMAD-3* en la ruta de TGF- $\beta$  y su papel funcional, ningún SNP en el gen se ha replicado en los GWAS. Por ello, no se ha confirmado su papel como factor genético relevante en OA.

### 3.4.4 *MATN3* e *IL1* cluster ligados a OA de mano pero sin confirmación definitiva

Estos loci se comenzaron a estudiar en OA de mano en los estudios de ligamiento de Islandia (*MATN3*) y Finlandia (*IL1 cluster*).

El estudio de Islandia incluyó una amplia colección de muestras de un fenotipo particular, OA de la articulación del carpo-metacarpiano del primer dedo (CMC1). En este trabajo se identificó ligamiento en tres regiones cromosómicas: 2p, 3p y 4q. En la región de ligamiento del cromosoma 2, se localiza el gen *MATN3* que codifica para una proteína de la ECM del cartílago, la matrilina-3. Un *screening* más detallado de la región llevó a la identificación de una mutación asociada que codifica para un cambio aminoacídico en la proteína. La frecuencia de esta mutación es de un 2% en pacientes<sup>174</sup>. Esta mutación, T303M junto con otros seis polimorfismos del gen se seleccionaron para un estudio de asociación en dos colecciones de muestras: Rotterdam y GARP. La cohorte de Rotterdam incluyó sujetos con degeneración de disco espinal y OA de mano en múltiples articulaciones y el estudio de GARP incluyó sujetos con OA de mano (DIP y CMC1). No se confirmó la asociación de la mutación T303M en ninguna de las dos colecciones de muestras. Sin embargo, otro SNP en el gen mostró asociación y también se detectó asociación en relación con degeneración de disco. Debido a ello, el resultado no fue considerado totalmente negativo<sup>175</sup>.

*MATN3* es un componente de la ECM del cartílago con un papel funcional estudiado ampliamente. Por un lado, participa en la formación de la ECM siendo capaz de mediar la producción de colágeno de tipo II y agregano y bloquear de la acción de ADAMTS-5<sup>176</sup>. *MATN3* también

regula la homeostasis del cartílago articular participando en la degeneración del cartílago articular al regular a la baja la expresión de SOX9<sup>177</sup>.

En conclusión, *MATN3* se considera como uno de los posibles factores genéticos de susceptibilidad en OA de mano. Pese a ello, se debe señalar que en el GWAS posterior llevado a cabo en OA de mano no se volvió a encontrar asociación de ningún SNP en este gen <sup>178</sup>.

Otro estudio de ligamiento en OA de mano, realizado en Finlandia, identificó cuatro regiones cromosómicas con evidencia de ligamiento (2q12-q13, 4q26-q27, 7p15-p21 y Xcen). En la primera de estas regiones se encuentra el cluster de la *IL-1* <sup>179</sup>. En un estudio detallado pero pequeño que cubrió la variabilidad de la región se encontró asociación de dos SNPs en *IL1B* (rs1143634 y rs1143633)<sup>180</sup>. Recientemente, para intentar aclarar la asociación del cluster (*IL1R1-IL1A-IL1B-IL1RN*), se llevó a cabo un metaanálisis incluyendo SNPs de *IL1B* e *IL1RN* y también las combinaciones haplotípicas de estos polimorfismos. A pesar de incluir 9132 controles, 5013 sujetos con OA de rodilla y 3595 con OA de cadera, no se encontró ninguna asociación <sup>181</sup>.

Como se trató anteriormente, *IL-1* participa en la destrucción de la ECM por lo que el estudio de este cluster resultó muy prometedor en un comienzo. Sin embargo, la falta de replicación de los hallazgos en estudios posteriores, hace que en la actualidad no se considere un locus relevante en la susceptibilidad a OA.

### 3.4.5 Un microsatélite funcional de *BMP-5* asociado con OA de cadera en mujeres del Reino Unido

El grupo de John Loughlin en el Reino Unido, realizó un estudio de ligamiento en OA de cadera y de rodilla identificando varias regiones con un LOD score indicativo de ligamiento<sup>182,183</sup>. Una de ellas, 6q12-q13, presentó un LOD score de 2.33 con OA de cadera en mujeres. Esta región es de particular relevancia al incluir dos genes con una posible repercusión en la susceptibilidad a la OA, *COL9A1* y *BMP-5*.

Se realizó un estudio más refinado de *COL9A1* y ninguno de los 20 SNPs genotipados mostró asociación<sup>184</sup>. También se encontraron resultados negativos para *BMP-5* en un trabajo en el que se identificaron los SNPs tanto en la secuencia exónica como en las regiones 5' y 3'UTR y se estudió su asociación con OA de cadera<sup>185</sup>. Sin embargo, tras genotipar numerosos microsatélites del gen, se encontró asociación de uno de ellos, localizado en el intrón 1, con OA de cadera. Este hallazgo motivó estudios posteriores de *BMP-5*<sup>186</sup>. Así, la variabilidad genética del intrón 1 se cubrió en detalle al genotiparse 9 SNPs, 8 microsatélites y 1 INDEL (inserción/delección), en mujeres del Reino Unido. El microsatélite D6S1276 mostró asociación ( $p=0.018$ ) con OA de cadera siendo una variante con efecto en la actividad transcripcional del gen *in vitro*<sup>187</sup>. Posteriormente, un estudio realizado con DAE (*different allelic expression*) comprobó que la expresión de *BMP-5* está regulada por polimorfismos en *cis* pero de una forma muy variable, dependiendo del tejido específico articular estudiado. Parte de esta regulación en *cis* se asoció con un SNP situado en la región 3' aunque esta asociación no fue compartida por otros tejidos analizados<sup>188</sup>. Esta regulación tan compleja del gen ya había sido observada en estudios en ratones, en los que se

había encontrado regulación independiente de BMP-5 incluso para su expresión en distintas porciones de las costillas. Esta característica, al igual que la de otros genes implicados en el desarrollo está influenciada por los denominados *long range enhancers*. Son regiones muy extensas adyacentes al gen, no codificantes y que participan en su regulación<sup>189</sup>. En el caso particular de *BMP-5* se localizan en la región 3' aunque también existen reguladores en el propio gen concretamente en los intrones<sup>64</sup>. Se sabe que BMP-5 ejerce su acción en condrocitos a través de sus receptores de membrana y activa la expresión de marcadores hipertróficos, como el colágeno de tipo X, que se encuentran aumentados en el cartílago artrósico<sup>190</sup>.

El microsatélite localizado en el intrón 1 de *BMP-5* se considera un posible factor de susceptibilidad a OA de cadera. No se ha comprobado su asociación en otros fenotipos como OA de rodilla o de mano. Además como los microsatélites no están incluidos en los chips de genotipado de los GWAS tampoco se puede descartar su asociación.

#### **3.4.6 *IL4-R*: otro locus ligado a OA pero no confirmado**

Otra de las regiones con evidencias de ligamiento con OA de cadera en mujeres en el estudio del Reino Unido es la región 16p12-p13 en la que se localiza *IL4R* (*interleukin 4 receptor*). Dos SNPs no sinónimos del gen mostraron asociación con OA de cadera en mujeres aunque no han sido estudiados desde el punto de vista funcional<sup>191</sup>. Estos resultados de asociación ni se replicaron en GWAS posteriores ni en un amplio estudio de replicación también realizado en el Reino Unido<sup>192</sup>. El papel funcional de *IL4R* le hace un buen candidato de estudio al estar implicado en la

regulación de las MMPs y participar en la señalización inducida por mecanotransducción. Sin embargo, la falta de confirmación de las asociaciones encontradas en un primer momento hacen que no se considere una señal de asociación importante en OA de cadera.

### **3.4.7 *FRZB*: asociación de dos SNPs no sinónimos no confirmados en estudios más recientes**

Un estudio británico encontró ligamiento con OA de rodilla en una región (2q24-q31) en la que se encuentran ocho genes que son buenos candidatos de estudio. Entre estos genes, un estudio de asociación caso-control destacó a *FRZB* (*frizzled-related protein gene*) al encontrarse dos SNPs no sinónimos (Arg324Gly y Arg200Trp) asociados con OA de rodilla. Estos SNPs se localizan en la secuencia codificante para el extremo C-terminal de la proteína. También, se realizaron múltiples estudios tratando de replicar esta asociación hasta que finalmente un metaanálisis de 14 estudios mostró resultados totalmente negativos<sup>193,194</sup>. Los dos SNPs se han asociado con diferentes formas de la cabeza del fémur que se relacionan con una mayor o menor probabilidad de desarrollar OA de cadera<sup>195</sup>.

En relación con el papel funcional de *FRZB*, actúa como antagonista de la ruta WNT implicada en la patogénesis de la OA. Sin embargo, ni los GWAS recientes ni el metaanálisis encontraron asociación de SNPs en *FRZB*, por lo que no se considera un factor genético de susceptibilidad a OA.



### 3.4.7 Variantes en *DIO2* posiblemente asociadas con OA de cadera a través de modulación de la tiroxina

En el estudio GARP (Holanda) se llevó a cabo un análisis de ligamiento de genoma completo que destacó regiones de ligamiento en los cromosomas 6, 10, 13 y 14. En la región del cromosoma 14q32.1 (LOD score=3.03) se encuentran tres genes, *CALMI*, *FLRT2* y *DIO2*, que se seleccionaron para un estudio más detallado mediante tagSNPs. Finalmente, este estudio sólo encontró asociación de dos SNPs en *DIO2*, el rs225014 (nsSNP) y el rs12885300 (5'UTR). El alelo C del rs225014 mostró un efecto protector mientras que el alelo T del rs12885300 mostró un efecto de susceptibilidad. En un metaanálisis de tres colecciones independientes que incluyó sujetos caucásicos y asiáticos, el haplotipo de susceptibilidad formado por estos dos SNPs mostró asociación en mujeres con OA de cadera ( $p= 2.02 \times 10^{-5}$ )<sup>196</sup>. En la misma colección GARP, este SNP se asoció con la morfología de la cadera. Como consecuencia, se ha propuesto que *DIO2* podría estar implicado en el proceso de morfogénesis del cartílago articular y en la geometría acetabular. Así, una forma poco óptima de la articulación llevaría a la aparición de OA de cadera con el paso del tiempo<sup>197</sup>.

En cuanto a su papel funcional, *DIO2* (*type II iodothyronine deiodinase*) es la enzima encargada de la conversión de la triyodotironina (T3) a tiroxina (T4). Se conoce su papel en la diferenciación de los condrocitos y su capacidad de modificar la respuesta inflamatoria al regular los niveles de COX-2 (*cyclooxygenase II*) e IL1 $\beta$ <sup>198,199</sup>. El mismo grupo que descubrió la asociación, realizó un estudio de DAE, observando una mayor expresión del alelo T de susceptibilidad del rs225014 en el cartílago de pacientes con OA. Esto indica que además de este nsSNP

debe haber otros polimorfismos reguladores en *cis*<sup>200</sup>.

La asociación de *DIO2* con OA de cadera no se ha observado ni en los GWAS ni en otros estudios posteriores. Por lo tanto, aunque se considera un posible factor de susceptibilidad apoyado por más de un estudio, no se puede considerar como probado o establecido.

### **3.4.8 DVWA : el mejor candidato del primer GWAS en OA realizado en población japonesa**

El primer GWAS en OA se realizó en japoneses con OA de rodilla. En una primera fase se encontraron unos 2000 SNPs con  $p < 0.01$  que se seleccionaron para genotipar en una colección más amplia e independiente de muestras. En esta segunda fase se identificó un SNP (rs3773472) asociado ( $p = 0.000017$ ). Un análisis detallado seleccionando cuarenta SNPs del mismo bloque de LD, llevó a identificar dos SNPs no sinónimos asociados: el rs7639618 y el rs11718863 ( $p = 7.3 \times 10^{-8}$ ,  $p = 2.1 \times 10^{-4}$ , respectivamente). Estos polimorfismos se localizaban en una región (*LOC344875*) que se transcribía en una proteína desconocida. Esta proteína posee dos dominios VWA por lo que se denominó DVWA (*double von Willebrand factor A domains*)<sup>201</sup>. Con posterioridad se han realizado dos estudios tratando de replicar la asociación de DVWA. El primero realizado en el Reino Unido no la encontró. El segundo incluyó otras muestras británicas, así como holandesas, españolas y griegas. En estas muestras tampoco se reveló asociación aunque sí una tendencia en el mismo sentido a la encontrada previamente en japoneses y chinos (rs7639618, OR=1.16,  $p = 0.06$  en europeos). La diferencia entre grupos étnicos se observó en la intensidad de la asociación y en las frecuencias alélicas<sup>202,203</sup>.

La función de *DVWA* ha sido muy poco estudiada. En el trabajo que descubrió su asociación se comprobó que *DVWA* interactúa con la  $\beta$ -tubulina de manera diferencial dependiendo de los alelos de los dos SNPs no sinónimos. Posteriormente, se han planteado dudas sobre su identidad ya que según algunos autores se trata de la porción 5' de otro gen, el *COL6A4*<sup>204</sup>. Sin embargo, la reciente clonación del gen, reveló dos isoformas de la proteína, una larga y otra corta. La corta no tiene homología con el *COL6A4* y tiene una expresión específica en cartílago. Se hipotetiza con un posible papel de *DVWA* en la diferenciación de los condrocitos y el tráfico intracelular<sup>205</sup>.

*DVWA* se consideró uno de los factores genéticos más sólidos para OA de rodilla pero al mostrar heterogeneidad entre poblaciones y no replicarse la asociación en ningún GWAS ha ido perdiendo relevancia e interés.

#### **3.4.9 *LRCHI*: uno de los resultados de un GWAS realizado con DNA pools**

En uno de los primeros estudios de tipo GWAS de OA en población europea se estudiaron unos 25000 SNPs y su posible asociación con OA de rodilla en tres cohortes independientes: una de descubrimiento y dos de replicación. Hay que destacar que este estudio se realizó con mezclas de DNA (*DNA pools*). En la primera colección de muestras se encontró asociación del rs912428 C/T en el intrón 1 del gen *LRCHI* (*leucine-rich repeats and calponin homology (CH) domain containing 1*). Esta asociación se confirmó en las dos cohortes de replicación y en el análisis combinado de las tres colecciones ( $OR=1.45$ ;  $p=5 \times 10^{-4}$ )<sup>206</sup>. Sin embargo, un estudio de asociación caso-control posterior en población británica no

replicó este hallazgo<sup>207</sup>. También mostró resultados negativos el metanálisis que incluyó población asiática y los estudios realizados en europeos<sup>208</sup>.

No está muy claro cuál es el papel de *LRCHI* en OA ya que a nivel funcional sólo se conoce que la familia de proteínas LRHCs podrían estar implicadas en la regulación del citoesqueleto y en el control de la morfología celular<sup>209</sup>.

Al igual que en el caso de otros genes comentados con anterioridad, las asociaciones encontradas en *LRHCI* no se replicaron en ningún GWAS. Por ello, *LRHCI* es uno de los genes cuyo estudio perdió relevancia en la genética de la OA.

### **3.4.10 *PTGS2*: un magnífico candidato funcional cuya asociación no se ha confirmado**

La asociación de *PTGS2* se descubrió en un GWAS en el que se genotiparon diferentes colecciones muestrales. En este trabajo, se realizó una primera fase de *screening*, utilizando mezclas de DNA, en 255 controles y 387 casos. A partir de ella, se seleccionaron 28 SNPs para estudiar en dos colecciones del Reino Unido y de EEUU en las que se obtuvieron resultados significativos para siete de ellos. Estos SNPs fueron genotipados en otra colección independiente, la cohorte de Rotterdam. Finalmente, el rs4140564, situado en la región 5'UTR del gen *PTGS2* (*prostaglandin-endoperoxide synthase 2*), mostró asociación tanto en la cohorte final de replicación como en el metanálisis global de todas las colecciones (OR=1.55;  $p=7 \times 10^{-7}$ ). En el mismo trabajo, utilizando

expresión alélica diferencial, se comprobó que podría haber posibles polimorfismos reguladores en *cis*<sup>210</sup>.

*PTGS-2* codifica para *COX-2* (*cyclooxygenase-2*) que tiene un papel relevante en el proceso inflamatorio, en la proliferación y diferenciación de condrocitos y en el proceso de degradación del cartílago<sup>80,90,211</sup>.

Al igual que en el caso de otros loci, *PTGS2* es una de las asociaciones más sólidas en el campo de la genética de la OA que no ha sido confirmada ni en los GWAS ni en los metaanálisis de los mismos. Por ello, su calificación como factor de riesgo de OA es incierta.

#### **3.4.11 Posible asociación de *A2BP1* con OA de mano**

La asociación de *A2BP1* se identificó en un GWAS realizado en dos etapas: descubrimiento y replicación. En la fase de descubrimiento se escogieron los 100 SNPs más asociados en la cohorte de Twins UK ( $p < 3.6 \times 10^{-4}$ ) y se genotiparon en una selección de las muestras de la cohorte de Rotterdam. Cuatro SNPs asociados se seleccionaron para una primera etapa de replicación en las muestras de Chingford. De entre ellos sólo el rs716508, se genotipó en las tres colecciones muestrales de replicación (GARP, Chuvasha y las muestras restantes de la cohorte de Rotterdam). Finalmente, este SNP mostró asociación ( $p = 4.75 \times 10^{-5}$ ) en el metanálisis de todas las colecciones, con heterogeneidad significativa y en el metanálisis de la fase de replicación ( $p = 1.81 \times 10^{-5}$ ) sin heterogeneidad. Este SNP está localizado en el intrón 1 del gen *A2BP1* (*ataxin 2-binding protein 1*).

## INTRODUCCIÓN

---

Se sabe que *A2BP1* es un regulador transcripcional del *splicing* de *CGRP* (*calcitonin gene related peptide*), cuyos receptores se expresan en osteoblastos y osteoclastos lo que parece sugerir alguna ruta de regulación conjunta del hueso y del cartílago. Por ello, en el mismo trabajo se analizó cuál es la relación entre el rs716508 y BMD (*bone mineral density*), demostrándose que el alelo C (protector) se correlaciona con una disminución de la BMD <sup>178</sup>.

*A2BP1* se considera el loci genético más prometedor en OA de mano ya que hasta el momento, el rs716508 es el SNP más fuertemente asociado con este fenotipo.

### **3.4.12 HLA: locus asociado con nivel de GWAS en asiáticos pero no en europeos**

Un GWAS realizado en población japonesa con OA de rodilla descubrió dos SNPs asociados, rs7775228 en *HLA-DQ1B* (*human leukocyte antigen DQB1*) y rs10947262 en *BTNL-2* (*butyrophilin-like 2*) ( $p=2.4 \times 10^{-8}$  y  $p=6.7 \times 10^{-8}$ ; respectivamente). Ambos SNPs están localizados en la región del HLA de clase II/III (*human leukocyte antigen*) y muestran una elevada correlación entre ellos ( $r^2=0.86$ ). Debido a estos resultados significativos, ambos SNPs se seleccionaron para genotipar en muestras europeas de Grecia y España. En el metanálisis de población asiática y caucásica sólo se encontró asociación a nivel de GWAS de uno de ellos, el rs10947262 ( $p=5 \times 10^{-9}$ ). Un trabajo posterior genotipó ambos polimorfismos en dos cohortes independientes de China y Australia y no confirmó la asociación <sup>212</sup>. Tampoco mostró asociación un estudio con

gran tamaño muestral en población europea con OA de rodilla (36000 controles y 5749 casos)<sup>213</sup>.

La importancia del estudio de polimorfismos en estos genes reside en que podrían explicar parte del componente inflamatorio de la OA. Las moléculas de HLA de clase II se expresan en las células presentadoras de antígenos (macrófagos, células dendríticas y linfocitos B). Por otro lado, *BTNL-2* interviene en la regulación de la activación de células T a través de un receptor desconocido.

El rs7775228 y el rs10947262 se incluyen entre los pocos polimorfismos asociados con  $p < 5 \times 10^{-8}$  en los estudios genéticos de OA. Sin embargo hay que destacar que esta asociación está restringida a población japonesa y no se confirmó en europeos.

#### **3.4.13 *GDF-5*: primer locus asociado con nivel de GWAS en europeos y asiáticos.**

El primer estudio genético que descubrió la asociación de un SNP (rs143383, C/T) en la región 5'UTR de *GDF-5* se realizó en población japonesa. La asociación se observó en OA de cadera ( $p = 1,8 \times 10^{-13}$ ) replicándose con menor magnitud en OA de rodilla<sup>214</sup>. Estudios posteriores confirmaron la asociación del rs143383 tanto en asiáticos como en europeos<sup>194,215</sup>. Un metaanálisis de quince cohortes europeas, y dos cohortes no europeas, japonesa y china, confirmó la asociación del rs143383 a nivel de GWAS con OA de rodilla ( $p = 6.2 \times 10^{-11}$ )<sup>194</sup>. Este mismo SNP se encontró asociado con displasia de cadera y con la heredabilidad de la altura<sup>216,217</sup>. Este resultado también se confirmó en el estudio de Rotterdam en el que el rs143383 se encontró asociado con OA,

## INTRODUCCIÓN

---

altura, tamaño óseo y riesgo de fractura vertebral en mujeres<sup>218</sup>. También se ha demostrado la existencia de asociación entre el genotipo del rs143383 y la severidad radiográfica de OA de rodilla medida por la escala de Kellgren-Lawrence<sup>219</sup>.

*GDF-5*, también denominado *CDMP-1* (*cartilage-derived morphogenetic protein-1*) es un miembro de la superfamilia del TGF- $\beta$ . Participa en el desarrollo del esqueleto y en la formación de la articulación, en el proceso de condrogénesis y en la reparación del hueso y del cartílago<sup>220,221</sup>. Mutaciones en *GDF-5* causan braquidactilia de tipo C y otras condrodisplasias producidas por anomalías del hueso y del cartílago<sup>222-224</sup>. Además de los fenotipos humanos producidos por mutaciones en el gen, diferentes estudios con modelos animales demostraron el papel de *GDF-5* en la esqueletogénesis y mantenimiento de la articulación. Ratones deficientes en *GDF-5* mostraron una reducción en la densidad del hueso subcondral y daño articular<sup>225</sup>. La pérdida de expresión de este gen causa defectos en la formación del cartílago y articulación, observándose fenotipos más severos en los ratones dobles mutantes de ratón para *GDF-5* y otros miembros de la familia como *GDF-6* y *GDF-7*<sup>226</sup>.

En el artículo del grupo de Shiro Ikegawa en el que se descubrió la asociación de *GDF-5* se realizaron estudios funcionales *in vitro* con SNPs comprendidos entre las posiciones +34 y +128, incluyendo el rs143383 y el rs143384. Este trabajo descubrió que diferencias alélicas en estos dos SNPs influyen en la actividad transcripcional de *GDF-5* en líneas condrogénicas. El grupo de John Loughlin realizó estudios *ex vivo* de expresión alélica diferencial (*DAE: different allelic expression*). Mediante esta técnica se comprobó que el rs143383 mostraba una



expresión del alelo T disminuida respecto al alelo C en el cartílago de pacientes con OA. En este mismo trabajo se observó que otro SNP, rs56366915, localizado en la región 3'UTR influencia la expresión alélica independientemente del SNP del promotor<sup>227</sup>. Debido a que los alelos del SNP rs143383 son C/T y mostraron una expresión diferencial, se planteó la existencia de una posible regulación epigenética. Para aclarar el papel de la metilación del DNA en *GDF-5*, se realizaron distintos experimentos tanto en líneas celulares como en tejido articular humano. Se observó que la desmetilación del alelo C del rs143383 y del rs143384 también C/T, está asociada con el aumento de regulación transcripcional de *GDF-5*, comprobándose la regulación epigenética del locus<sup>228</sup>.

*GDF-5* es el factor genético más consistentemente demostrado en los estudios genéticos de OA.

#### **3.4.14 Locus 7q22: segundo locus asociado con nivel GWAS en europeos.**

La asociación de este locus con OA de rodilla se descubrió en un GWAS que genotipó 500510 SNPs. Para esta primera fase de descubrimiento se utilizó la cohorte de Rotterdam que incluyó casos de OA de rodilla, cadera y mano. Presentaron asociación 18 SNPs en 12 loci diferentes, entre ellos tres SNPs en *GDF-5* y tres en el locus 7q22 (rs3815148, rs1548524, rs997311), concretamente en el gen *COG5*. Estos SNPs asociados con OA de rodilla se seleccionaron para el genotipado en 12 colecciones muestrales adicionales que incluyeron los tres fenotipos iniciales. Finalmente, sólo un SNP, rs3815148, presentó asociación a nivel de GWAS (OR = 1.14,  $p = 8 \times 10^{-8}$ ).

## INTRODUCCIÓN

---

Este locus ocupa unas 500 kb, y contiene aproximadamente unos 400 SNPs y seis genes diferentes: *PRKAR2B* (*cAMP-dependent protein kinase type II-beta regulatory subunit*), *HBPI* (*HMG-box transcription factor 1*), *COG5* (*conserved oligomeric Golgi complex subunit 5*), *GPR22* (*G protein-coupled receptor 22*), *DUS4L* (*dihydrouridine synthase 4-like*) y *BCAP29* (*B-cell receptor-associated protein 29*)<sup>229</sup>. Posteriormente se realizó un metanálisis con los datos de 4 GWAS (deCODE, Rotterdam Study, Framingham, Twins UK) en el que las variantes que alcanzaron un valor de  $p < 2 \times 10^{-5}$  se seleccionaron para replicar en nueve colecciones europeas adicionales (arcOGEN, Grecia, España Finlandia, Nottingham, Chingford, GARP, Estonia y Suecia). El metaanálisis global incluyó 6709 casos y unos 40.000 controles. El rs4730250 (*DUS4L*) fue el SNP más asociado tanto en la fase inicial como en el análisis global de todas las colecciones (OR = 1.17,  $p = 9.2 \times 10^{-9}$ )<sup>151</sup>.

De entre todos los genes del locus, *GPR22* resultó ser el mejor candidato para realizar los primeros estudios funcionales. Esto es debido a que un SNP localizado *upstream* del gen y en elevado ligamiento con el rs3815148, se asoció previamente con variaciones en la expresión del gen en linfoblastos. En condrocitos en cultivo y mediante qPCR se mostró una baja expresión de *GPR22* respecto a los otros cinco genes del locus. Sin embargo, utilizando inmunohistoquímica se encontró expresión de *GPR22* en los condrocitos de la capa superficial de ratones con OA<sup>229</sup>. Por lo tanto, existen discrepancias en los resultados de expresión de *GPR22* en cartilago en diferentes estudios. Recientemente, se ha comprobado la posible existencia de polimorfismos reguladores en *cis* en el locus 7q22. Para ello se realizaron experimentos de DAE en cartilago de pacientes y se analizaron los niveles de expresión de cada gen presente en el locus. A

continuación se estratificaron los datos obtenidos en función del genotipo presente para los dos SNPs previamente asociados; rs381548 y rs4730250. Curiosamente, no se encontró expresión de *GPR22* en el cartílago mientras que la expresión de *HBPI* se reducía significativamente en presencia del alelo G del rs3815148. Este resultado permite resaltar la importancia dentro del locus de *HBPI* (factor de transcripción HMG box) importante en la ruta de señalización WNT<sup>230</sup>.

En conclusión, el locus 7q22 es uno de los loci más consistentemente asociados con OA de rodilla junto con GDF-5.

#### **3.4.15 Resultados de la primera fase del proyecto arcOGEN**

Las asociaciones de estos genes se descubrieron en el GWAS realizado en las muestras de la primera fase de arcOGEN (*Arthritis Research UK Osteoarthritis Genetics*). Este estudio incluyó 3177 casos y 4894 controles del Reino Unido y constituyó un gran esfuerzo en el esclarecimiento de las bases genéticas de susceptibilidad a la OA. En la etapa de descubrimiento se genotiparon 514898 SNPs en pacientes con OA de rodilla y cadera. En la segunda etapa, los SNPs independientes con una  $r^2 < 0.4$  y con  $p < 0.0001$ , se seleccionaron para realizar una replicación *in silico* con los datos ya disponibles de las colecciones de deCODE, Rotterdam y Framingham. Además, unos 52 SNPs se seleccionaron para replicar *de novo* en la colección de Twins UK. Finalmente, se metaanalizaron los resultados obtenidos tanto *in silico* como en el genotipado de la colección de arcOGEN y se escogieron 36 SNPs para replicar en otras cohortes europeas. En la primera fase se obtuvieron dos hallazgos principales: el rs4512391 (OR = 1.17;  $p = 1.8 \times 10^{-6}$ ) con OA

de rodilla, localizado *upstream* del gen *TRIB1* (*tribbles homolog 1*) y el rs4977469 (OR = 1.3;  $p = 1.2 \times 10^{-6}$ ) con OA de cadera, en el intrón 3 del gen *FAM154A* (*family with sequence similarity 154, member*). Los resultados obtenidos en el metanálisis de todos los estudios de replicación encontraron tres señales principales: el rs2277831 (OR = 1.07;  $p = 2.3 \times 10^{-3}$ ) asociado con OA rodilla y/o cadera localizado en el intrón 32 de *MICAL3* (*microtubule associated monooxygenase, calponin and LIM domain containing 3*); el rs11280 (OR = 1.1;  $p = 3.2 \times 10^{-5}$ ) con OA de rodilla, en *c6orf130* y el rs261577 en el *COL11A1* que se trata con detalle en el apartado siguiente.

La asociación de estos SNPs no se confirmó en la siguiente fase de arcOGEN en la que sin embargo, se descubrieron señales con  $p < 5 \times 10^{-8}$  que se tratan en el apartado **3.4.17**.

### **3.4.16 COL11A1: el mejor candidato entre las señales del proyecto arcOGEN fase 1**

En la primera fase del estudio arcOGEN se descubrió la asociación de rs2615977 ( $p = 1.1 \times 10^{-5}$ ), en *COL11A1*. Anteriormente, se había identificado la asociación de otro SNP (rs1676486) en el mismo gen, con hernia de disco lumbar (LDH) en japoneses. En este mismo estudio, se observó como la expresión de *COL11A1* disminuye en disco intervertebral de pacientes con LDH lo que se correlacionó con la presencia del alelo de susceptibilidad (C) de rs1676486<sup>231</sup>. Curiosamente, en un estudio de expresión alélica en cartílago artroscópico también se correlacionó la expresión de *COL11A1* con este SNP aunque no esté asociado con OA. Sin embargo, no se encontró ninguna correlación entre

los niveles de expresión del gen y los alelos del rs2615977, asociado en el estudio de arcOGEN<sup>232</sup>.

El *COL11A1*, desempeña un papel funcional importante en la estructura de la ECM. Codifica para una de las tres fibrillas que forman la molécula de colágeno de tipo XI, colágeno minoritario de la ECM. Las otras dos fibrillas son codificadas por otros dos genes, *COL11A2* y *COL2A1*. La función específica del *COL11A1* es unirse a los agregados de proteoglicanos y anclarlos a la red de colágeno de la ECM. Mutaciones en el gen provocan distintas condrodisplasias en ratones, el síndrome de Marshall y el síndrome de Stickler en humanos y fibrocondrogénesis<sup>233-235</sup>. También se encontraron mutaciones de baja frecuencia en el gen asociadas con formas no sindrómicas de OA temprana<sup>236</sup>.

Pese a los resultados positivos de asociación y a su importante papel funcional, *COL11A1* no se incluye entre los loci definitivamente asociados de susceptibilidad a OA. Esto se debe a que la asociación con OA de cadera no alcanzó el nivel requerido en los GWAS ni se confirmó en la segunda fase del estudio de arcOGEN con un mayor tamaño muestral.

#### **3.4.17 Resultados de la segunda fase del proyecto arcOGEN.**

Recientemente, se ha publicado el GWAS que incluye el genotipado de todas las muestras de arcOGEN (7410 casos con OA de rodilla o de cadera, 11009 controles) genotipadas con un array de *Illumina* que cubre unos 500.000 SNPs. En este trabajo también se realizó la replicación *in silico* en un número similar de casos y controles en seis colecciones europeas y de *novo* en otra cohorte de Reino Unido.

## INTRODUCCIÓN

---

En el análisis global de la fase de descubrimiento y replicación se encontraron cinco señales con  $p < 5 \times 10^{-8}$ . Los dos *top* SNPs, el rs6976 y el rs11177, presentan una  $r^2=1$  entre ellos por lo que representan una única señal asociada con reemplazamiento articular tanto de rodilla como de cadera. El rs6976 (OR = 1.12;  $p = 7.2 \times 10^{-11}$ ) está situado en la región 3'UTR de *GLT8D1* (*glycosyltransferase 8 domain containing 1*) y el rs11177 (OR = 1.12;  $p = 1.25 \times 10^{-10}$ ) es un SNP no sinónimo localizado en el exón 3 de *GNL3* (*guanine nucleotide binding protein-like 3/nucleostemine*). Las cuatro señales restantes están asociadas con OA de cadera y son: el rs4836732 (OR = 1.2;  $p = 6.11 \times 10^{-10}$ ) en el gen *ASTN2* (*astrotactin 2*) (asociación específica de mujeres), el rs9350591 (OR = 1.18;  $p = 2.4 \times 10^{-9}$ ) localizado entre los genes *FILIP1* (*filamin A interacting protein 1*) y *SENP6* (*SUMO1/sentrin specific peptidase 6*), el rs10492367 (OR = 1.11;  $p = 1.48 \times 10^{-8}$ ) localizado entre los genes *KLHDC-5* (*Kelch domain containing 5*) y *PTHLH* (*parathyroid hormone-like hormone*) y el rs835487 (OR = 1.13;  $p = 1.64 \times 10^{-8}$ ) en el gen *CHST11* (*carbohydrate (chondroitin 4) sulfotransferase 11*).

Otras tres señales mostraron asociación que estuvo cerca del umbral de  $p < 5 \times 10^{-8}$ . El rs12107036 (OR = 1.21;  $p = 6.71 \times 10^{-8}$ ) en *TP63* (*transformation related-protein 63*) mostró asociación con OA de rodilla en mujeres. Además se encontraron dos asociaciones específicas de sexo e independientes de articulación: el rs8044769 (OR = 1.11;  $p = 6.85 \times 10^{-8}$ ) en el gen *FTO* (*fat mass and obesity*) en mujeres y el rs10948172 (OR = 1.14;  $p = 7,92 \times 10^{-8}$ ) localizado entre los genes *SUPT3H* (*suppressor of Ty3 homologue*) y *CDC5L* (*CDC5 cell division cycle 5-like*) y específico de hombres.

En el mismo trabajo se discute el posible papel funcional de cada uno de estos genes ó de genes vecinos con SNPs relacionados por  $r^2$ . *GNL3* se expresa en células mesenquimales que son las precursoras de los condrocitos y además su expresión se encuentra aumentada en condrocitos artrósicos. *ASTN2* se expresa en cerebro dónde regula la migración neuronal y de momento no se le conoce su función en el cartílago articular. *Upstream* a los genes *FILIP1-SEN6* se encuentra el *COL12A1*, con un papel importante en la ECM. En relación a los genes, *KLHDC-5* y *PTHLH*, se conoce la repercusión funcional del último de ellos. Se comprobó que la proteína relacionada con la hormona paratiroidea (PTHrP) mostró una expresión elevada en el cartílago de pacientes con OA y participa en el desarrollo del hueso subcondral<sup>237</sup>. *CHST11* está implicado en regulación de la síntesis de componentes de proteoglicanos concretamente del condroitín sulfato. Además, el condroitín sulfato se da como tratamiento contra la OA aunque existen opiniones controvertidas acerca de su efectividad. El siguiente gen, *TP63* está relacionado con el desarrollo de la articulación y se ha descrito como uno de los loci que determinan la morfología facial en humanos<sup>238</sup>. Cerca del locus *CDC5L-SUPT3H* se encuentra *RUNX-2* que es esencial tanto en la diferenciación de los osteoblastos como en la morfogénesis del esqueleto. El gen *FTO* se relaciona con obesidad por mecanismos no totalmente aclarados pero que podrían incluir regulación de la ingesta o producción de energía. Al corregir por BMI, la asociación encontrada disminuye por lo que se considera que *FTO* ejerce su acción en la OA a través de la obesidad<sup>141</sup>.

### 3.4.18 DOTIL

Un GWAS con un enfoque original utilizó como variable a explicar la anchura del espacio articular (*JSW: joint space width*). Esta característica aporta varias ventajas desde un punto de vista teórico: es una variable cuantitativa, lo que aumenta la potencia estadística; es un endofenotipo de OA, lo que permite una mayor proximidad a los procesos patogénicos y es una variable precisa, lo que podría disminuir la heterogeneidad entre pacientes. En cualquier caso, este enfoque permitió descubrir la asociación del rs1298744 en *DOTIL* (*DOT1-like, histone H3 methyltransferase*) utilizando la anchura del espacio articular de la cohorte de Rotterdam (6523 sujetos). El metaanálisis global de estos datos con los de las cohortes de replicación (tres colecciones de Reino Unido) mostró asociación del rs1298744 ( $p=1.1 \times 10^{-11}$ ). Recientemente, se publicó asociación del mismo SNP con OA de cadera en hombres ( $p=7.8 \times 10^{-9}$ ) en un que reunió un amplio tamaño muestral. Además, se debe comentar que el rs1298744 se había encontrado asociado previamente con altura<sup>239</sup>.

Los estudios funcionales realizados en ratones probaron que *DOTIL* se expresa en el cartílago de pacientes con OA e interviene en el proceso de condrogénesis. *DOT1L* interactúa con TCF-4 regulando la transcripción de los genes de la ruta Wnt e influenciando la diferenciación condrogénica<sup>155</sup>.

*DOTIL* es uno de los loci asociados en OA al nivel requerido en los GWAS y por lo tanto es una de las señales asociadas más firmemente con la enfermedad.



### 3.4.19 *MCF2L*

Una forma de detectar posibles variantes genéticas asociadas con OA y no identificadas directamente en los GWAS, es combinarlos con la información del Proyecto 1000 Genomas (1KGP). Este proyecto nace de una colaboración entre grupos de todo el mundo y pretende proporcionar un catálogo completo de toda la variación genética posible que alberga el genoma humano. Mediante secuenciación de última generación seguida de distintas etapas de genotipado y validación, se catalogan además de variantes con MAF >5%, variantes de baja frecuencia (0.5-5%) y variantes raras (<0.5%)<sup>240</sup>.

La información de 1KGP se ha utilizado para realizar imputación de SNPs y variantes raras no estudiadas directamente en el GWAS de arcOGEN. Esta imputación ayudó a confirmar la asociación del rs11841874 en *MCF2L* (*MCF.2 cell-line-derived transforming sequence-like*) con OA de cadera. Para ello, se testaron más de siete millones de variantes con MAF>1% de entre las cuáles se seleccionaron ocho SNPs en seis loci para genotipar directamente en las muestras del GWAS original. Finalmente, el rs11841874 fue el único SNP que permaneció asociado y se confirmó su asociación en otras cohortes de muestras (OR=1.17; p metanálisis global=2 x 10<sup>-8</sup>). Este SNP estaba genotipado en las muestras del GWAS pero no se consideró relevante porque era el único asociado en la región con un nivel de asociación modesto (p=1 x 10<sup>-5</sup>). Sin embargo, la imputación utilizando 1KGP reveló otros SNPs asociados y correlacionados con el rs11841874, lo que estableció la importancia del locus. Además, este trabajo revela la importancia de utilizar la

información contenida en 1KGP como herramienta útil para descubrir y/o confirmar variantes asociadas en OA <sup>152</sup>.

En relación al papel funcional de *MCF2L* se sabe que es un intercambiador de nucleótidos. Además, participa en la regulación de la neurotrofina-3 que pertenece a la familia NGF (*neurotrophin growth factors*). NGF actúa como factor pro-angiogénico y su expresión se encontró aumentada en condrocitos artrósicos <sup>241</sup>.

En la actualidad *MCF2L* es uno de los loci más firmemente asociados con OA de cadera.

### 3.5 Visión general y futuro de los estudios genéticos de OA.

El estudio de los factores genéticos implicados en la susceptibilidad a la OA es un campo complejo que a pesar de las dificultades ha progresado notablemente en los últimos años. En la actualidad hay 9 loci asociados con OA en población europea al nivel requerido en los GWAS y otros 3 cerca de este nivel. La mayoría fueron descubiertos en el estudio arcOGEN. Otros son *MCF2L*, *DOTIL* y dos loci que se descubrieron con anterioridad a estos trabajos: 7q22 y *GDF-5*. El número de genes de susceptibilidad encontrados es relativamente bajo si lo comparamos con el de otras enfermedades complejas. En este trabajo experimental se han abordado varias aproximaciones de estudio que pueden facilitar el hallazgo de nuevas variantes asociadas con OA:

1. El estudio de fenotipos particulares, de endofenotipos o de fenotipos extremos puede ayudar a un mejor conocimiento de la arquitectura

genética de la enfermedad y a un diseño más efectivo de los estudios genéticos. Un ejemplo de fenotipos extremos lo constituyen la aparición de la enfermedad a una edad temprana. Para explorar esta campo se plantea la posibilidad de que la carga genética de loci previamente asociados aumenta de forma inversa a la edad. Una relación de este tipo ha sido demostrada en otras enfermedades complejas como SLE<sup>242</sup>. Además, formas sindrómicas de OA se desarrollan en edades tempranas debido a mutaciones en genes estructurales de la ECM que son distintas a las causantes de las formas tardías. Este patrón dicotómico entre formas de inicio temprano y tardío llevan a hipotetizar que las formas comunes de OA se podrían desarrollar antes en sujetos con un mayor número de alelos de riesgo<sup>134,243,244</sup>. Testar esta hipótesis, es de interés para el diseño de estudios clínicos ya que se facilitaría la selección de pacientes con progresión rápida de la enfermedad. Además, también se podría incrementar la posibilidad de identificar loci de susceptibilidad al seleccionar pacientes con edad temprana de inicio de los síntomas y controles de edad avanzada en los estudios genéticos.

2. Otra de las aproximaciones que pueden ser de utilidad es el resecuenciado de genes con variantes comunes previamente asociadas. En otras enfermedades se ha observado que en estos genes se pueden encontrar otras variantes raras que también están asociadas. Por lo tanto, los genes asociados son un buen principio para empezar a buscar variantes raras asociadas con OA. Esto es importante porque las variantes de baja frecuencia son una fuente potencial de la heredabilidad no explorada. Los estudios de secuenciación se puede realizar de una manera tradicional seleccionando los genes de interés y resecuenciándolos mediante el método de Sanger en controles y pacientes. También se pueden utilizar las

## *INTRODUCCIÓN*

---

tecnologías de secuenciación de última generación (NGS) que permiten secuenciar con rapidez el genoma completo o el exoma.

3. Otra potencial fuente de variabilidad no explicada reside en variantes estructurales como CNVs, microsatélites o inserciones y deleciones. Estas variantes no están recogidas en los GWAS o lo están de forma parcial. Además, el efecto genético de estas variantes puede no ser detectado a través de los SNPs incluidos en los chips de genotipado ya que sus numerosos alelos pueden no tener suficiente LD con ellos. En OA no se ha realizado ningún estudio específico de CNVs aunque dos microsatélites. Sin embargo dos variantes de este tipo, un VNTR en asporina y un microsatélite en *BMP5* mostraron asociación con la enfermedad. Dado que ninguna de estas asociaciones alcanzó el nivel de GWAS requieren una confirmación posterior. A este campo no se le ha prestado demasiada atención pero puede tener resultados prometedores.

4. Otra cuestión que se plantea es cómo extraer más información de los GWAS publicados hasta el momento en OA. Todos estos trabajos han realizado metaanálisis con los datos de múltiples colecciones de muestras lo que constituye una manera muy eficiente de aumentar el poder estadístico del estudio. También es muy útil para detectar nuevas asociaciones ya que cada estudio proporciona información para la mayor parte de los SNPs del genoma tanto genotipados directamente como mediante imputación. Sin embargo, aún es posible descubrir nuevas asociaciones a partir de estos datos centrándose en conjuntos particulares de genes de los cuáles existía una evidencia previa de asociación (genes candidatos). Una aproximación de este tipo no ha sido realizada todavía en OA. Además, serviría para evaluar el estado actual de los genes

candidatos incluidos en los numerosos estudios de asociación previos a los GWAS.

Todas estas aproximaciones serán exploradas en la parte experimental de esta tesis.





---

## OBJETIVOS

El objetivo de esta tesis es encontrar nuevos factores genéticos de susceptibilidad a OA empleando múltiples aproximaciones que incluyen:

1. Estudio de fenotipos extremos en relación con la edad de inicio de los síntomas mediante análisis de la relación entre esa edad y el número de alelos de riesgo de SNPs asociados con OA de rodilla.
2. Estudio de variantes raras de alta penetrancia mediante la búsqueda de variantes de baja frecuencia en el gen *GDF-5*. Esta búsqueda es de especial importancia ya que los GWAS no cubren toda la variabilidad genética y parte de la heredabilidad inexplorada podría ser explicada por variantes raras.
3. Estudio de microsatélites funcionales, que son otro tipo de variantes no estudiadas en los GWAS, mediante genotipado del microsatélite en el intrón 1 del gen *BMP-5 (D6S1276)* asociado previamente con OA de cadera y que es funcional.
4. Estudio de variantes con efecto pequeño que pueden estar ya entre los resultados de los GWAS pero sin alcanzar el nivel de significación requerido y que se podrían priorizar gracias a un metaanálisis de genes candidatos con los resultados de asociación de 9 GWAS y un estudio de replicación con los nuevos SNPs asociados.







**PROCEDIMIENTO EXPERIMENTAL**

**Y RESULTADOS**

---



# PROCEDIMIENTO EXPERIMENTAL Y RESULTADOS

## Esquema general

*1.1 Estudio de la relación entre la edad de inicio de los síntomas de OA y la carga genética de la enfermedad*

**Publicación 1: Rodríguez-Fontenla C, López-Golán Y, Calaza M, Pombo-Suarez M, Gómez-Reino JJ, González A. Genetic risk load and age at symptom onset of knee osteoarthritis. J Orthop Res. 2012 Jun;30(6):905-9.**

*1.2 Estudio de resecuenciado del gen GDF-5 buscando variantes raras asociadas.*

**Publicación 2: Dodd AW\*, Rodríguez-Fontenla C\*, Calaza M, Carr A, Gomez-Reino JJ, Tsezou A, Reynard LN, Gonzalez A, Loughlin J. Deep sequencing of GDF5 reveals the absence of rare variants at this important osteoarthritis susceptibility locus. Osteoarthritis Cartilage. 2011 Apr;19(4):430-4**

\*Contribución equivalente

*1.3 Estudio de asociación caso-control en OA de rodilla del microsatélite D6S1276 de BMP-5*

**Publicación 3: Rodríguez-Fontenla C, Carr A, Gomez-Reino JJ, Tsezou A, Loughlin J, Gonzalez A. Association of a BMP5 microsatellite with knee osteoarthritis: case-control study. Arthritis Res Ther. 2012 Nov 27;14(6):R257**

*1.4 Análisis de genes candidatos de artrosis en un metaanálisis de 9 GWAS*





**PUBLICACIÓN 1**

---



## Genetic Risk Load and Age at Symptom Onset of Knee Osteoarthritis

Cristina Rodríguez-Fontenla,<sup>1</sup> Yolanda López-Golán,<sup>1</sup> Manuel Calaza,<sup>1</sup> Manuel Pombo-Suarez,<sup>1</sup> Juan J. Gómez-Reino,<sup>1,2</sup> Antonio González<sup>1</sup>

<sup>1</sup>Laboratorio Investigación 10, Rheumatology Unit, Instituto Investigación Sanitaria, Hospital Clínico Universitario de Santiago, Travesía Choupana sn., 15706 Santiago de Compostela, Spain, <sup>2</sup>Medicine Department, Universidad de Santiago de Compostela, Santiago de Compostela, Spain

Received 22 July 2011; accepted 31 October 2011

Published online 18 November 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jor.22018

**ABSTRACT:** To test whether a higher genetic risk load for knee osteoarthritis (OA) is associated with an earlier age at symptom onset. Six polymorphisms in *GDF5*, *PTGS2*, 7q22 locus, *DVWA*, *DIO3*, and *ASPN* that have been associated with knee OA were analyzed in 255 patients that had undergone total knee replacement (TKR) because of primary OA and in 457 healthy controls. We looked for association between the number of risk alleles in each patient and his age at symptom onset with linear regression and *t*-tests between the upper and lower quartiles. There was not even a weak trend in the direction of a younger age at symptom onset in the patients carrying more risk alleles. Patients in the upper quartile of age at symptom onset ( $67.0 \pm 2.8$  years) carried the same number of OA risk alleles ( $5.4 \pm 1.4$  vs.  $5.3 \pm 1.0$ ) than patients in the lower quartile ( $44.6 \pm 5.5$  years). We did not find any evidence in support of the hypothesis of an earlier knee OA symptom onset associated with higher genetic risk load as determined by the six loci. This result suggests that old age and genetic risk act as independent factors in the pathogenesis of OA. It also indicates that designing OA genetic studies with patients selected for early symptom onset will not provide any substantial power gain. © 2011 Orthopaedic Research Society. Published by Wiley Periodicals, Inc. *J Orthop Res* 30:905–909, 2012

**Keywords:** knee osteoarthritis; genetics; age of onset; susceptibility; senescence

Osteoarthritis (OA) is a complex disease affecting all tissues of the joints and whose etiology, progression, and relationship between anatomical anomalies and clinical manifestations are poorly understood.<sup>1,2</sup> Old age is a major contributing factor in OA.<sup>3</sup> The peak incidence of radiographic knee OA is reached at 70–79 years, although the first cases appear before 40 years of age.<sup>4</sup> This has been known for long time, but the causes of this relationship are still unclear.<sup>5,6</sup> Possible mechanisms include accumulated mechanical damage to the joint or, alternatively, chondrocyte senescence that makes these cells less efficient in response to insults or aging of cartilage matrix with changes that make the tissue more brittle.

One of the many aspects of interest in these questions is the consideration of the possible role of genetic factors.<sup>7</sup> It is well known that some syndromic forms of OA develop at early ages due to deleterious mutations in structural cartilage genes.<sup>7–9</sup> These syndromic forms of OA are different from the common form of OA that appears at older ages and that is under the influence of numerous genetic factors of much weaker individual effect. A similar dichotomy is observed in other diseases like Alzheimer's, Parkinson's, and breast cancer. This dichotomous pattern of early- and late-onset forms makes it tempting to hypothesize that among the patients with the common form those with a higher genetic load will have an earlier disease onset than those with a lower load. This hypothetical relationship implies that OA develops earlier in subjects with a larger number of OA risk alleles because each

of them contributes to lower the threshold for overt disease. Damage could be either because risk alleles determine a faulty structure of joint tissues, incompetent repair mechanisms, or dysregulated inflammatory responses, as has been already suggested for some OA risk genes.<sup>10–14</sup> Therefore, patients with higher genetic load will require less contribution from mechanical or other environmental factors.

In spite of the plausibility of the above-mentioned hypothesis there are very few pieces of evidence supporting a relationship between genetic risk load and age at disease onset in common complex diseases. We are not aware of any direct proof of it in OA, and we only know of the positive result from a very recent study in other complex disease, SLE.<sup>15</sup> Testing this hypothesis in OA is of interest for the understanding of the disease, but also of practical interest for the design of clinical trials, because it could facilitate selecting patients with rapid progression to overt disease,<sup>16</sup> and genetic studies, because selecting patients with early onset and controls of old age could increase the chances of identifying susceptibility loci.<sup>17</sup>

In the current study, we have done multiple analyses directed to find whether there is any relationship between the number of OA risk alleles and age at symptom onset or not. We have used for these analyses the best supported susceptibility loci for knee OA and a collection of patients with advanced primary OA that required total knee replacement (TKR). None of the analyses performed showed any hint of a higher genetic risk load in the patients with younger age at symptom onset; nor of a lower load in older controls free of overt OA.

### METHODS

#### Patients and Controls

Cases were selected from consecutive patients, aged 55–80 years at the time of surgery, undergoing total knee

Additional Supporting Information may be found in the online version of this article.

The authors declare no conflict of interest.

Correspondence to: Antonio González (T: 34 981 950 903; F: 34 981 951 068;

E-mail: antonio.gonzalez.martinez-pedrayo@sergas.es)

© 2011 Orthopaedic Research Society. Published by Wiley Periodicals, Inc.

replacement. They were included only if a rheumatologist considered them to suffer from severe primary OA. Exclusion criteria included inflammatory, infectious or traumatic joint pathology and lesions due to crystal deposition or osteonecrosis. Time of symptom onset was provided by patients in response to a questionnaire. Controls were recruited among subjects older than 55 years of age undergoing preoperative work-up for elective surgeries other than joint surgery and who did not show clinical manifestations or complaints compatible with OA when specifically explored and questioned. This study was approved by the Ethical Committee for Clinical Research of Galicia and all cases and controls gave their written informed consent to participate according to the Declaration of Helsinki.

**Genotyping of Knee OA Associated Loci**

We have analyzed five SNPs and a microsatellite that are in knee OA susceptibility loci. They were rs143383 in *GDF-5*<sup>10,18,19</sup>, rs4140564 in *PTGS2*<sup>14</sup>, rs12535761 (which is perfect proxy of rs4730250) in the 7q22 locus<sup>20,21</sup>; rs7639618 in *DVWA*<sup>22,23</sup>, rs945006 in *DIO3*<sup>13</sup> and the aspartic acid repeat microsatellite in *ASPN*.<sup>11,24</sup> The SNPs were genotyped by single-base extension with the SNaPshot Multiplex Kit (Applied Biosystems, Foster City, CA) after amplification of 30 ng of genomic DNA in a multiplex PCR reaction (KAPA2G Fast HotStart, Kapa Biosystems, Woburn, MA). Samples were analyzed in an ABI Prism 3130xl Genetic Analyzer (Applied Biosystems). Several samples with different genotypes were sequenced to assess the accuracy of results with the Big Dye v3.1 Ready Reaction Kit (Applied Biosystems). The number of repeats of the *ASPN* microsatellite was determined by capillary electrophoresis on an ABI 3130xl Genetic Analyzer (Applied Biosystems) after amplification with primers labeled with FAM. Oligonucleotide sequences are available from the authors upon request. Genotype calling percentages were uniform across patients, controls and polymorphisms (mean calling rate was 0.985). Genotype frequencies were in Hardy-Weinberg equilibrium (HWE) for all SNPs.

**Statistical Analysis**

The Statistica (Statsoft, Tulsa, OK) software was used thoroughly. Analysis of the effect of risk loci on the age at symptom onset was performed with different approaches and separately in cases and controls. Age of the controls at the time of recruitment was considered as the minimum age free of OA symptoms. A combined analysis of the six susceptibility loci was possible because each locus was in a different chromosome and, therefore, they were inherited in an independent way. This analysis was carried out with lineal regression between age at symptom onset for patients, or minimum age free of symptoms for controls, and number of OA risk alleles. It was repeated after adjustment for gender and BMI. Finally, age distributions between carrier and non-carrier strata were compared for each of the polymorphisms separately with lineal regression and with Q-Q plots. Power analyses were done with GPower 3.0.10.<sup>25</sup>

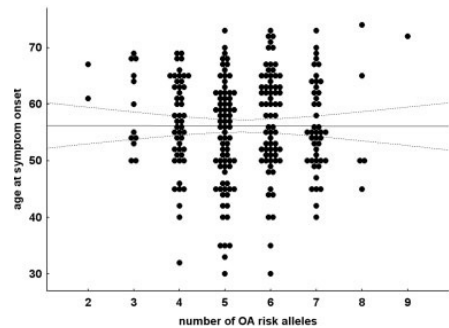
**RESULTS**

Six polymorphisms corresponding to six well-validated knee OA susceptibility loci were genotyped in 255 patients with knee OA (206 women and 49 men) and in 456 controls without articular pain or deformity (160 women and 296 men), all of them of Spanish

origin and older than 55 years of age. Mean age at symptom onset was 56.3 years. There were not significant differences between women (55.8 ± 8.9 years) and men (58.0 ± 8.8 years; *p* = 0.11). The range was very wide extending from 30 to 74 years. It is also, interesting to consider that our knee OA patients were 68.0 ± 5.7 years old at the time of recruitment. The controls were of similar age at the time of recruitment (68.9 ± 8.9 years; *p* = 0.14).

Genotypes were obtained with a 98.5% call rate and fulfilling other quality control criteria including confirmation by sequencing of some results and lack of deviation of genotype frequencies from the expected under Hardy-Weinberg equilibrium. Percentages of the risk alleles were nominally higher in the knee OA patients for each of the six polymorphisms, but none of these differences was significant (not shown), as was expected given the known effect sizes of the polymorphisms.<sup>13,14,19-21,23,24</sup>

The number of risk alleles ranged from 2 to 9 in patients with knee OA, but not any trend was appreciable relating them with the age at symptom onset in linear regression analysis (Fig. 1). This result was not modified after adjusting by BMI and gender. This analysis had enough power to detect an effect of allele risk load explaining 2.5% of the variance in age at symptom onset with 80% sensitivity. As a complementary analysis we compared the two extremes of the distribution both in the number of risk alleles and in the age at symptom onset. The 62 patients with 2-4 risk alleles showed a similar age at symptom onset (57.3 ± 8.1 years) than the 50 patients with 7-9 risk alleles (56.4 ± 8.6 years; *p* = 0.6). This analysis was powered to detect a difference of 4.0 years in age at symptom onset with 80% sensitivity. In a similar way, the lowest quartile in age at symptom onset showed a similar number of risk alleles (5.3 ± 1.0) than the highest quartile (5.4 ± 1.4; *p* = 0.6) in spite of the



**Figure 1.** Lack of relationship between age at symptom onset and number of risk alleles. Number of OA risk alleles at the six examined OA susceptibility loci in each of the 255 knee OA patients were represented against their age at symptom onset. The linear regression fitting line with its two 95% confidence boundaries was drawn.



large difference in age at symptom onset between the two patient quartiles ( $44.6 \pm 5.5$  vs.  $67.0 \pm 2.8$  years). In this instance, the minimal detectable difference with 80% power was of 0.56 risk alleles.

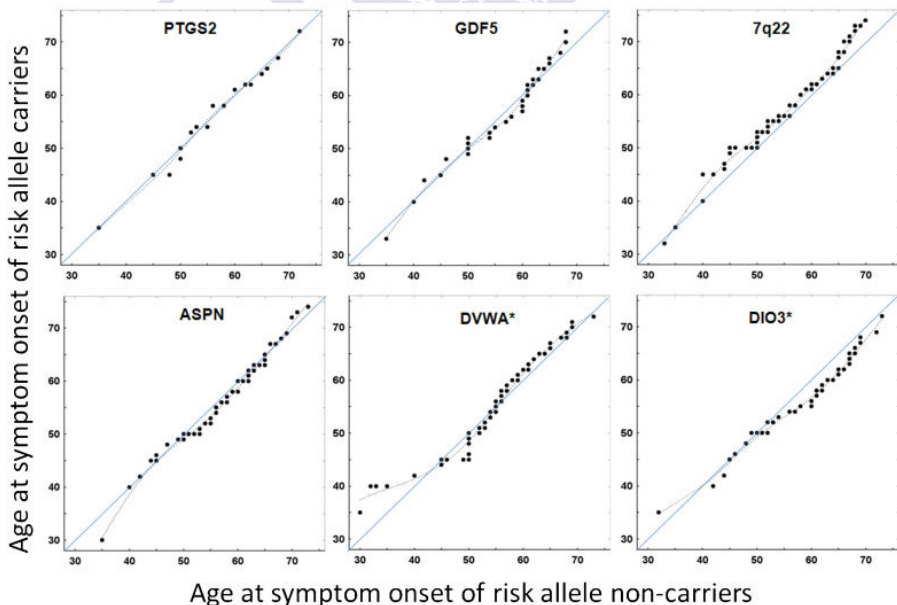
As a further test of the validity of our results we repeated the linear regression analysis considering only the two loci that have shown association with knee OA according with the  $5 \times 10^{-8}$  *p*-value threshold required in GWAS, *GDF5* and *7q22*. This analysis did not disclose any association with age at symptom onset ( $R^2 = 0.0028$ ;  $p = 0.4$ ; Supplementary Fig. 1).

In addition, there was not any significant difference in the minimum age at which controls were free of symptoms (that is, age at recruitment) in function of the number of risk alleles at the six polymorphisms in any of the analyses done. These analyses included linear regression analysis ( $R^2 = 0.0006$ ;  $p = 0.6$ ) with 80% power to detect a difference accounting for 1.35% of the variance; comparison of age free of symptoms between controls with 2–4 risk alleles and controls with 7–9 risk alleles ( $70.3 \pm 9.4$  vs.  $68.6 \pm 8.6$  years, respectively); and comparison of the number of risk alleles between the upper and the lower quartiles in age at symptom onset ( $5.3 \pm 1.3$  alleles vs.  $5.2 \pm 1.2$  alleles, respectively). Detectable differences for these two latter analyses were 3.23 years and 0.4 alleles, respectively.

Individual linear regression analysis of age at symptom onset in function of each of the polymorphisms did not show any significant association either unadjusted or after adjusting by BMI and gender (not shown). As it was possible that a relationship between genotypes and age at symptom onset could be more evident in a limited range of ages, we have analyzed the quantile–quantile plots of age at symptom onset between carriers and non-carriers of the risk allele at each polymorphism (Fig. 2). However, none of the plots showed any convincing difference towards a younger age at disease onset of the carriers or risk alleles than in the non-carriers. In a similar way, minimum age free of symptoms in controls was not associated with the genotypes at any of the six polymorphisms (not shown).

**DISCUSSION**

We have not obtained any evidence in support of the hypothesis of an increased genetic risk load in patients with an earlier onset of symptoms of knee OA. This result indicates that the genetic component of knee OA acts with independence of age to increase disease liability. Also, it suggests that there will not be a significant gain in selecting patients with earlier disease onset for the identification of new susceptibility loci, or



**Figure 2.** Q–Q plots comparing age at symptom onset in risk allele carriers with non-carriers. Quantiles are connected with least squares fitting lines. A bias towards a younger age at onset will be shown by a fitting line departing from the bisector to the lower half of the plot. Patients homozygous for the risk allele against non-homozygous given the low number of non-carriers.

in selecting patients with higher genetic load for clinical trials. Finally, it also shows the difficulties of the proposed classification of primary OA patients in separated subsets of different etiology based on the genetic or the old age contribution.<sup>5</sup>

One of the limitations of our study is that it has been done when we know relatively little about OA genetic loci. The six loci we have analyzed are arguably the ones with soundest support,<sup>10,11,13,14,18-24</sup> but only two of them have reached the GWAS significant status (with  $p < 5 \times 10^{-8}$ ), *GDF5* and the *7q22* locus.<sup>19,21</sup> However, we judged it opportune to analyze this question at this time given that it could help design more powerful genetic studies, and because it seemed that there are not many more OA loci near to be discovered. In fact, only a new sound locus associated with knee OA has been reported since we started this study, *SMAD3*,<sup>26</sup> in spite of the recent publication of the largest OA GWAS ever done<sup>27</sup> and of a large meta-analysis of OA GWAS.<sup>21</sup> We also acknowledge that a small effect of genetic risk load in the age at symptom onset could have escaped our detection. This has been quantified as 2.5% of the variance in age at symptom onset for 80% statistical power. In addition, analysis of controls free of OA symptoms support the results obtained with the knee OA patients.

The small percentage of variance detectable in our study reflects the high statistical power of regression analysis when both the dependent and the independent variables are continuous and approach a normal distribution, as happens in our study with age at symptom onset and genetic risk load, respectively. This high power contrasts with the lower power of genetic studies where the dependent variable is dichotomous, case or control, and the independent variable has only the 3 levels 0, 1, and 2 for the aa, Aa, and AA genotypes of a SNP. Availability of a larger number of OA susceptibility loci will improve further the sensitivity of the analysis by increasing the favorable properties of the genetic load score. For example, an analysis identical to the reported here but including ten OA loci in place of six will increase power from 80% to 99% for the same effect. Therefore, we cannot exclude that future discovery of new OA loci will lead to identification of a relationship between genetic risk load and age at symptom onset. On the contrary, modifying the range or distribution of age at symptom onset will not provide significant gains in power according with our estimations.

It is interesting to try to define the implications of our results regarding OA mechanisms. The hypothesis we have tested implicates an acceleration of OA determined by the level of genetic risk. As our results do not support this hypothesis, age, and genetic risk load seem to be independent contributors to OA. This could happen in disease models where there is a dynamic interplay between damage and repair that could be dominated by one or the other process in different phases. Therefore, disease progression will be phasic

and with the possibility of regression or of periods of stability. This type of scenario is consistent with some recent findings in the follow-up of cartilage damage biomarkers<sup>28,29</sup> and MRI lesions.<sup>30-33</sup> Therefore, it is possible that old age contribution to overt OA will be due to a bias in the balance towards damage beyond a no-return point because of age-related factors like cell senescence or matrix fragility.<sup>5,6</sup>

Our results have direct implications for the design of genetic studies aiming to identify new OA associated loci. These studies have been particularly difficult.<sup>19-21,27</sup> They have obtained very few consistently demonstrated associations. It is unclear if the poor performance of the studies done until now is due to defects in their design, especially regarding selection of patients and controls,<sup>34</sup> or to the intrinsic genetic architecture of OA that could be more polygenic than earlier suspected,<sup>27</sup> or more dependent of genetic variants or interactions that are beyond the reach of current studies. A way to improve efficiency of genetic studies is to compare extreme phenotypes.<sup>17</sup> One of the tactics implementing this approach will be to compare knee OA patients of young onset with controls free of OA at old ages. Our results do not support the use of this tactic in knee OA. They were negative for the two sides of the proposed tactic, the side of the patient and the side of the control.

Lack of relationship between genetic risk load and age at symptom onset does not exclude that genetic variation at a particular locus could affect the age at symptom onset. This is particularly likely for gene variants that could influence tissue aging or pain sensitivity. In any case, none of the six loci we have studied had any effect on the age at symptom onset considered individually.

#### ACKNOWLEDGMENTS

We thank Carmen Pena-Pena for her excellent technical help. We also thank all participants for generously contributing their samples and time. Author contribution: CR-F genotyped all samples, analyzed the data and wrote the first draft of the manuscript; YL-G recruited patients and controls; MC and JGG-R have participated in the analysis and interpretation of results; AG designed the study and has supervised all aspects of its development. All authors have participated in writing the manuscript and all read and approved the final manuscript. This study was supported by Fondo de Investigación Sanitaria of the Instituto de Salud Carlos III (Spain), grants 08/0744 and 09/01431 that are partially financed by the European Regional Development Fund of the European Union, and by grants from the Xunta de Galicia. CR-F is the recipient of a predoctoral contract from Instituto de Salud Carlos III (Spain).

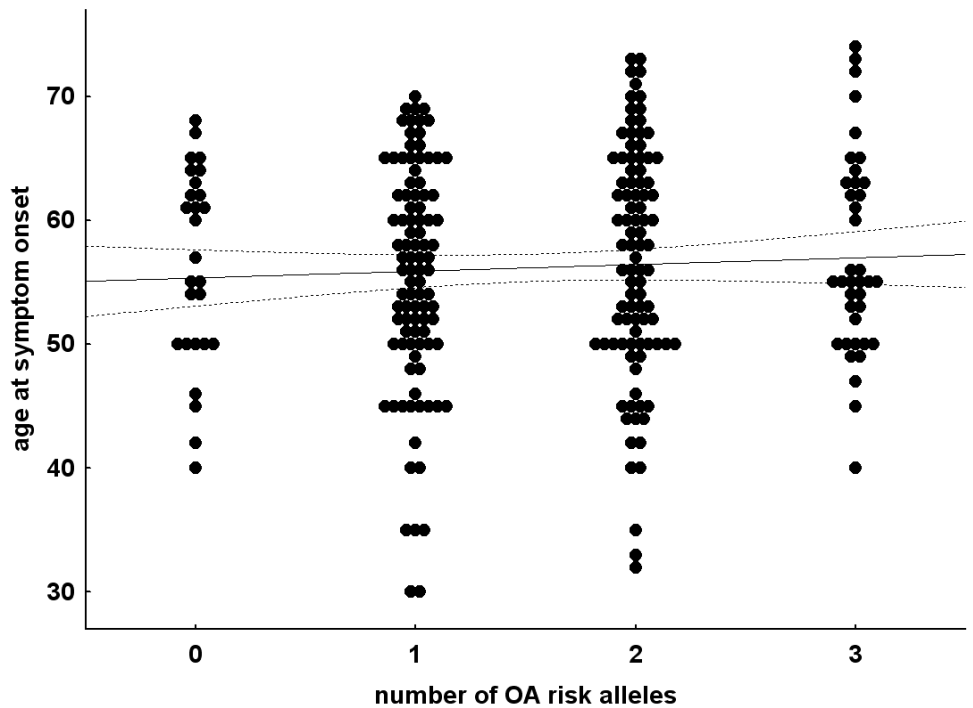
#### REFERENCES

1. Goldring MB, Goldring SR. 2007. Osteoarthritis. *J Cell Physiol* 213:626-634.
2. Bedson J, Croft PR. 2008. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC Musculoskelet Disord* 9:116.

3. Blagojevic M, Jinks C, Jeffery A, et al. 2010. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis Cartilage* 18: 24–33.
4. Oliveria SA, Felson DT, Reed JI, et al. 1995. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. *Arthritis Rheum* 38:1134–1141.
5. Loeser RF. 2009. Aging and osteoarthritis: the role of chondrocyte senescence and aging changes in the cartilage matrix. *Osteoarthritis Cartilage* 17:971–979.
6. Martin JA, Buckwalter JA. 2002. Aging, articular cartilage chondrocyte senescence and osteoarthritis. *Biogerontology* 3:257–264.
7. Valdes AM, Spector TD. 2009. The contribution of genes to osteoarthritis. *Med Clin North Am* 93:45–66, x.
8. Herrero-Beaumont G, Roman-Blas JA, Castaneda S, et al. 2009. Primary osteoarthritis no longer primary: three subsets with distinct etiological, clinical, and therapeutic characteristics. *Semin Arthritis Rheum* 39:71–80.
9. Kannu P, Bateman JF, Belluoccio D, et al. 2009. Employing molecular genetics of chondrodysplasias to inform the study of osteoarthritis. *Arthritis Rheum* 60:325–334.
10. Southam L, Rodriguez-Lopez J, Wilkins JM, et al. 2007. An snp in the 5'-utr of *gdf5* is associated with osteoarthritis susceptibility in europeans and with in vivo differences in allelic expression in articular cartilage. *Hum Mol Genet* 16:2226–2232.
11. Kizawa H, Kou I, Iida A, et al. 2005. An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. *Nat Genet* 37:138–144.
12. Meulenbelt I, Min JL, Bos S, et al. 2008. Identification of *dio2* as a new susceptibility locus for symptomatic osteoarthritis. *Hum Mol Genet* 17:1867–1875.
13. Meulenbelt I, Bos SD, Chapman K, et al. 2011. Meta-analyses of genes modulating intracellular T3 bio-availability reveal a possible role for the *DIO3* gene in osteoarthritis susceptibility. *Ann Rheum Dis* 70:164–167.
14. Valdes AM, Loughlin J, Timms KM, et al. 2008. Genome-wide association scan identifies a prostaglandin-endoperoxide synthase 2 variant involved in risk of knee osteoarthritis. *Am J Hum Genet* 82:1231–1240.
15. Webb R, Kelly JA, Somers EC, et al. 2011. Early disease onset is predicted by a higher genetic risk for lupus and is associated with a more severe phenotype in lupus patients. *Ann Rheum Dis* 70:151–156.
16. Lohmander LS, Felson D. 2004. Can we identify a 'high risk' patient profile to determine who will experience rapid progression of osteoarthritis? *Osteoarthritis Cartilage* 12:S49–S52.
17. Morton NE, Collins A. 1998. Tests and estimates of allelic association in complex inheritance. *Proc Natl Acad Sci USA* 95:11389–11393.
18. Miyamoto Y, Mabuchi A, Shi D, et al. 2007. A functional polymorphism in the 5' UTR of *GDF5* is associated with susceptibility to osteoarthritis. *Nat Genet* 39:529–533.
19. Valdes AM, Evangelou E, Kerkhof HJ, et al. 2010. The *GDF5* rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. *Ann Rheum Dis* 70:873–875.
20. Kerkhof HJ, Lories RJ, Meulenbelt I, et al. 2010. A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. *Arthritis Rheum* 62:499–510.
21. Evangelou E, Valdes AM, Kerkhof HJ, et al. 2011. Meta-analysis of genome-wide association studies confirms a susceptibility locus for knee osteoarthritis on chromosome 7q22. *Ann Rheum Dis* 70:349–355.
22. Miyamoto Y, Shi D, Nakajima M, et al. 2008. Common variants in *dvwa* on chromosome 3p24.3 are associated with susceptibility to knee osteoarthritis. *Nat Genet* 40:994–998.
23. Meulenbelt I, Chapman K, Dieguez-Gonzalez R, et al. 2009. Large replication study and meta-analyses of *dvwa* as an osteoarthritis susceptibility locus in european and asian populations. *Hum Mol Genet* 18:1518–1523.
24. Nakamura T, Shi D, Tzetis M, et al. 2007. Meta-analysis of association between the *aspn* d-repeat and osteoarthritis. *Hum Mol Genet* 16:1676–1681.
25. Faul F, Erdfelder E, Lang AG, et al. 2007. G\*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 39:175–191.
26. Valdes AM, Spector TD, Tamm A, et al. 2010. Genetic variation in the *SMAD3* gene is associated with hip and knee osteoarthritis. *Arthritis Rheum* 62:2347–2352.
27. Panoutsopoulou K, Southam L, Elliott KS, et al. 2010. Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study. *Ann Rheum Dis* 70:864–867.
28. Sharif M, Kirwan JR, Elson CJ, et al. 2004. Suggestion of nonlinear or phasic progression of knee osteoarthritis based on measurements of serum cartilage oligomeric matrix protein levels over five years. *Arthritis Rheum* 50:2479–2488.
29. Chaganti RK, Kelman A, Lui L, et al. 2008. Change in serum measurements of cartilage oligomeric matrix protein and association with the development and worsening of radiographic hip osteoarthritis. *Osteoarthritis Cartilage* 16: 566–571.
30. Dore D, Quinn S, Ding C, et al. 2010. Natural history and clinical significance of mri-detected bone marrow lesions at the knee: a prospective study in community dwelling older adults. *Arthritis Res Ther* 12:R223.
31. Wirth W, Larroque S, Davies RY, et al. 2011. Comparison of 1-year vs 2-year change in regional cartilage thickness in osteoarthritis results from 346 participants from the osteoarthritis initiative. *Osteoarthritis Cartilage* 19:74–83.
32. Kornaat PR, Kloppenburg M, Sharma R, et al. 2007. Bone marrow edema-like lesions change in volume in the majority of patients with osteoarthritis; associations with clinical features. *Eur Radiol* 17:3073–3078.
33. Zhang Y, Nevitt M, Niu J, et al. 2011. Fluctuation of knee pain and changes in bone marrow lesions, effusions, and synovitis on magnetic resonance imaging. *Arthritis Rheum* 63:691–699.
34. Kerkhof HJ, Meulenbelt I, Akune T, et al. 2011. Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. *Osteoarthritis Cartilage* 19:254–264.



**Sup Fig 1:** Lack of relationship between age at symptom onset and number of risk alleles when only the 7q22 SNP rs12535761 and the GDF5 SNP rs143383 were considered. Numbers of risk alleles at these two SNPs in each of the 255 knee OA patients were represented against their age at symptom onset. The linear regression fitting line with its two 95 % confidence boundaries was drawn.







**PUBLICACIÓN 2**

---





# Osteoarthritis and Cartilage



## Deep sequencing of GDF5 reveals the absence of rare variants at this important osteoarthritis susceptibility locus

A.W. Dodd<sup>†,a</sup>, C. Rodriguez-Fontenla<sup>‡,a</sup>, M. Calaza<sup>‡</sup>, A. Carr<sup>§</sup>, J.J. Gomez-Reino<sup>‡</sup>, A. Tsezou<sup>||</sup>, L.N. Reynard<sup>†</sup>, A. Gonzalez<sup>‡</sup>, J. Loughlin<sup>†\*</sup>

<sup>†</sup>Newcastle University, Institute of Cellular Medicine, Newcastle, UK

<sup>‡</sup>Instituto Investigacion Sanitaria, Hospital Clinico Universitario de Santiago, Santiago de Compostela, Spain

<sup>§</sup>University of Oxford, NDORMS, Oxford, UK

<sup>||</sup>University of Thessaly, Medical School, Department of Biology, Larissa, Greece

### ARTICLE INFO

#### Article history:

Received 1 November 2010

Accepted 22 January 2011

#### Keywords:

GDF5

DNA variants

Sequencing

Genetic susceptibility

### SUMMARY

**Objective:** The common single nucleotide polymorphism (SNP) rs143383 in the 5' untranslated region (5'UTR) of growth and differentiation factor 5 (GDF5) is strongly associated with osteoarthritis (OA) and influences GDF5 allelic expression *in vitro* and in the joint tissues of OA patients. This effect is modulated *in cis* by another common SNP, also located within the 5'UTR, whilst a common SNP in the 3'UTR influences allelic expression independent of rs143383. DNA variants can be common, rare or extremely rare/unique. To therefore enhance our understanding of the allelic architecture of this very important OA susceptibility locus we sequenced the gene for potentially functional and novel rare variants.

**Method:** Using the Sanger method we sequenced GDF5 in 992 OA patients and 944 controls, with DNA changes identified by sequencing software. We encompassed the protein-coding region of the two GDF5 exons, both untranslated regions and approximately 100 bp of the proximal promoter of the gene.

**Results:** We detected 13 variants. Six were extremely rare with minor allele frequencies (MAFs) of  $\leq 0.0006$ . One is in a predicted transcription factor binding site in the GDF5 promoter whilst two substitute conserved amino acids. The remaining seven variants were common and are previously known variants, with MAFs ranging from 0.025 to 0.39. There was a complete absence of variants with frequencies in-between the extremely rare ( $n = 6$ ) and the common ( $n = 7$ ).

**Conclusions:** This is the first report of the deep sequencing of an OA susceptibility locus. The absence of rare variants informs us that within the regions of the gene that we have sequenced GDF5 does not harbour any novel variants that are able to contribute, at a population level, to the OA association signal mediated by rs143383 nor does it harbour, at a population level, any novel variants that can influence OA susceptibility independent of rs143383.

© 2011 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

### Introduction

Growth and differentiation factor 5 (GDF5), also known as cartilage derived morphogenetic protein 1 (CDMP1), is an extracellular signalling molecule that participates in the development, maintenance and repair of bone, cartilage and other tissues of the synovial joint<sup>1</sup>, with penetrant and rare deleterious mutations of the GDF5 gene resulting in dominant skeletal defects<sup>2</sup>. Based on this

functional knowledge the GDF5 gene was chosen by a Japanese group for examination as a potential osteoarthritis (OA) susceptibility locus<sup>3</sup>. A number of common GDF5 polymorphisms were genotyped in cases and in controls and association to the common single nucleotide polymorphism (SNP) rs143383, a T to C transition located in the 5' untranslated region (5'UTR) of the gene, was observed. In subsequent studies association of this SNP with OA was observed in Europeans, demonstrating that the association has global relevance<sup>4–6</sup>. Functional studies have suggested that rs143383 is itself the polymorphism influencing OA susceptibility with the OA-associated T-allele mediating reduced GDF5 transcription relative to the C-allele<sup>7,8</sup>. Furthermore, association of rs143383 to other musculoskeletal phenotypes, including variation in normal height, Achilles tendon pathology, fracture risk and congenital dysplasia of the hip, has been reported. This highlights

\* Address correspondence and reprint requests to: John Loughlin, Newcastle University, Institute of Cellular Medicine, Musculoskeletal Research Group, 4th Floor, Catherine Cookson Building, The Medical School, Framlington Place, Newcastle upon Tyne NE2 4HH, UK.

E-mail address: john.loughlin@ncl.ac.uk (J. Loughlin).

<sup>a</sup> These authors contributed equally to this work.

the tendency of a common genotype to influence multiple phenotypes (pleiotropy) and of the potential role that developmental factors can have on conditions that manifest in the mature individual<sup>9–13</sup>.

The negative effect of the OA-associated T-allele of rs143383 on *GDF5* transcription is modulated by a second common polymorphism located in the 5'UTR of the gene, rs143384; in a luciferase assay using chondrogenic cells it was demonstrated that the T-allele of rs143383 mediates a reduction in luciferase activity when it is on the background of a T-allele of rs143384 but not when it is on the background of the C-allele of this SNP<sup>8</sup>. Furthermore, using RNA extracted from the joint tissues of OA patients, this study demonstrated that a third SNP in *GDF5*, rs56366915 (referred to as 2250ct in ref<sup>7</sup>), which is located in the 3'UTR, affects allelic expression of *GDF5* independent of rs143383.

Mouse models have further developed our understanding of the role of *GDF5* in skeletogenesis and joint maintenance. The brachypodism mouse is homozygote for a frame-shift mutation of *Gdf5* that causes a premature termination codon resulting in a null allele<sup>14</sup>. These mice have a number of abnormalities including a shortened appendicular skeleton, a decrease in the length of the limb long bones, and soft tissue deformities that clearly demonstrate the important role of *GDF5* in normal development<sup>14–16</sup>. A haplo-insufficient mouse carrying one copy of the brachypodism allele does not show such developmental abnormalities but does show an increased propensity of developing an OA phenotype when challenged<sup>17</sup>. This model, which from a molecular genetics perspective mimics the effect of SNP rs143383 in human OA, confirms the importance of maintaining adequate levels of *GDF5* for normal joint function and emphasizes again the detrimental effects of quantitative alterations in its expression on joint tissue maintenance.

Overall, the rs143383 polymorphism is a very significant contributor to OA susceptibility, albeit with modest effect sizes (odds ratios < 1.2). This polymorphism is functionally complex, with the allelic expression mediated by rs143383 being modulated by another common SNP and with a third SNP influencing allelic expression in an independent manner. Cataloguing the variation at this gene is therefore critical to our full understanding of this important locus and we set out to assess the allelic architecture of *GDF5* in OA populations from northern and southern Europe using a deep-sequencing approach.

## Patients and methods

### Cases and controls

OA cases and controls were studied from the UK, Spain and Greece. All were of Caucasian ethnicity. The UK cohort were north Europeans of white British origin, the Spanish cohort were Europeans of Spanish origin and the Greek cohort were Europeans of central Greek origin. There were 992 cases that had each undergone a total knee or a total hip joint replacement due to severe primary OA and 944 controls that lacked clinical signs of the disease. Comprehensive details about the cases and the controls have been published<sup>7,18</sup>.

In the UK cohort there were 502 cases (383 females and 119 males; 220 knee cases and 282 hip) and 460 controls (184 females and 276 males). The cases were ascertained using the criteria of signs and symptoms of OA sufficiently severe to require joint replacement surgery. The radiological stage of the disease was a Kellgren and Lawrence grade of 2 or more in all cases with over 90% being grade 3 or 4. Inflammatory arthritis (rheumatoid, polyarthritic or autoimmune disease) was excluded, as was post-traumatic or post-septic arthritis. The cases had an age range of

56–85 years. The controls had no signs or symptoms of arthritis or joint disease (pain, swelling, tenderness or restriction of movement) and had an age range of 55–89 years.

In the Spanish cohort there were 264 cases (214 females and 50 males, all knee cases) and 294 controls (207 females and 87 males). Cases were selected as consecutive patients, aged 55–80 years of age at the time of surgery, undergoing total knee replacement if a rheumatologist considered them to suffer from severe primary OA. Exclusion criteria were inflammatory, infectious, traumatic or congenital joint pathology and lesions due to crystal deposition or osteonecrosis. Controls were recruited among subjects older than 55 years of age undergoing preoperative work-up for elective surgeries other than joint surgery and who did not show clinical manifestations of OA.

In the Greek cohort there were 226 cases (200 females and 26 males, all knee cases) and 190 controls (137 females and 53 males). The cases had a Kellgren and Lawrence grade of greater than 2 prior to total knee replacement and were randomly selected. Individuals with rheumatoid arthritis and other autoimmune diseases as well as chondrodysplasias, infection-induced OA, and post-traumatic OA were excluded. The cases had an age range of 40–92 years. The controls had undergone treatment for injuries and fractures and had an age range of 46–88 years. All controls had a Kellgren and Lawrence grade of 0.

Ethical approvals for the use of the DNAs in OA genetic studies were obtained from local ethical committees: Oxford for the UK cohort, Santiago for the Spanish cohort and Larissa for the Greek cohort.

### DNA sequencing

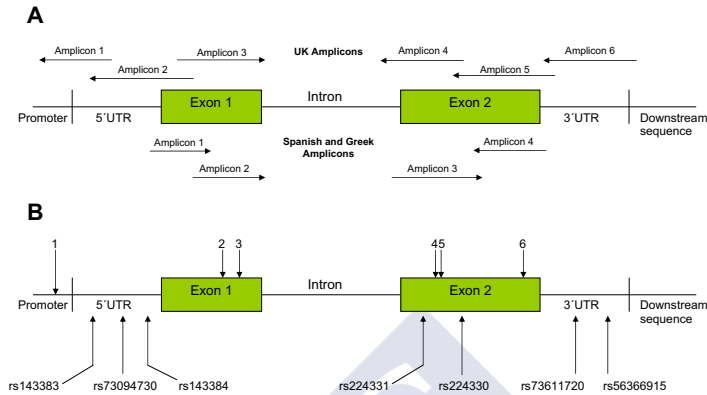
Using 50 ng of DNA extracted from peripheral blood samples by phenol–chloroform or guanidine hydrochloride *GDF5* was polymerase chain reaction (PCR) amplified with overlapping primers and these were then subjected to dideoxy Sanger sequencing using the BigDye v3.1 kit (Applied Biosystems) and 3130 Genetic Analyzer capillary sequencers (Applied Biosystems). In the Spanish and Greek cohorts the search for variants focused on the protein-coding regions of the two exons of *GDF5* and the intron/exon flanking sequences, whilst the UK search encompassed the protein-coding regions, the intron/exon flanking sequences, both UTRs and 100 bp of the proximal promoter of *GDF5* (Fig. 1A). The sequences of the primers used are listed in Supplementary Table 1. The Polyphred and SeqScape software's (Applied Biosystems) were used to align sequences and to identify DNA changes. Novel variants were confirmed by designing a new PCR amplicon and repeating the sequencing on a fresh aliquot of DNA.

Variant 1 was genotyped in the Spanish and Greek cohorts using the SNaPshot minisequencing protocol (Applied Biosystems) and the PCR forward primer 5'TTCCGTTTCCAATCTGAG3', the PCR reverse primer 5'TGACTGAGGGCTTGAAGGAG3' and the SNaPshot primer 5'AAC TAGGGGAAAAAAAACCTGGAGCACACAAGCAGCAT TACGCCATT3'.

## Results

### Known variants

*GDF5* is 4.9 kb in length and contains two exons coding for a 150 amino acid protein. We identified seven common variants that were previously known to exist in the gene: rs143383, rs73094730, rs143384, rs224331, rs224330, rs73611720 and rs56366915 (Fig. 1B, Supplementary Table II). rs224331 and rs224330 were observed in the UK, Spanish and Greek cohorts, since the region of *GDF5* harbouring these two variants (exon 2) was encompassed in the



**Fig. 1.** The sections of *GDF5* sequenced and the location of the newly discovered variants. (A) The six amplicons used for the sequencing in the UK cohort and the four amplicons used for the Spanish and Greek cohorts. (B) At the top the arrows point to the approximate location of each of the six novel variants. At the bottom the arrows point to the approximate location of the seven known variants that are present in dbSNP and which we confirmed.

sequencing of all three of these cohorts (Fig. 1A). The remaining five variants were observed only in the UK cohort, since the regions harbouring these five variants (the 5' and 3' UTRs) were only sequenced in this cohort (Fig. 1A).

None of the known variants identified by us demonstrated association to OA ( $P < 0.05$ ), with association analysis performed unstratified and stratified by cohort, by sex and by joint (data not shown). We were not particularly surprised by this result since it is known that large sample sizes are required to generate robust association to rs143383<sup>4–6</sup> and by comparison our sample size is small and underpowered. Furthermore, and as noted above, rs224331 and rs224330 were the only SNPs with data available in all three cohorts, thus reducing further any power for the five variants with genotype data available only from the UK cohort.

There are several other *GDF5* SNPs listed in dbSNP but these either lack frequency data or have so far been shown to be polymorphic only in non-European samples. These are rs34534075, rs61754581, rs28936397, rs76603468, rs61754580, rs28936683, rs17853055, rs5841206, rs6120942 and rs79051206 (<http://www.ncbi.nlm.nih.gov/projects/SNP/>). We did not detect these variants despite their coverage in our sequence analysis and we conclude therefore that they are not polymorphic in Europeans, at least in the large sample that we have studied.

**Novel variants**

We identified six novel and extremely rare variants, all in the UK cohort (Fig. 1, Table I). Each variant was detected in a single individual, two in cases and four in controls. The six UK individuals carrying a variant allele were each heterozygote for the variant

allele that they carried. Variant 1 was discovered in a female control following the successful sequencing of 458 cases and 418 controls and therefore has a MAF in the UK sample of 0.0006 (1 mutant allele in 1752 chromosomes). Variants 2 (female case with knee OA) and 3 (male case with knee OA) were discovered following the successful sequencing of 491 cases and 425 controls and therefore have minor allele frequencies (MAFs) in the UK sample of 0.0005 (1 mutant allele of each variant in 1832 chromosomes). Variants 4 (male control) and 5 (female control) were discovered following the successful sequencing of 474 cases and 389 controls and therefore have MAFs in the UK sample of 0.0006 (1 mutant allele of each variant in 1726 chromosomes). Variant 6 (female control) was discovered following the successful sequencing of 495 cases and 436 controls and therefore has a MAF in the UK sample of 0.0005 (1 mutant allele in 1862 chromosomes). For variants 2–6 these MAFs are upper-bound estimates, as the variants were not observed in the Spanish or Greek samples. The region of *GDF5* harbouring variant 1 was not included in the sequence analysis of the Spanish and Greek cohorts. We therefore directly genotyped this variant in these cohorts, using as a positive the UK female carrying the variant. The variant was not detected in the Spanish or Greek cohorts. Variant 1, like variants 2–6, was therefore observed only in the UK cohort.

We then used bioinformatics tools to see if any of these novel variants had the potential to alter the *GDF5* protein or affect transcription factor binding to the *GDF5* gene. Variant 1 is located in the promoter of *GDF5*, 41 bp upstream of the transcription initiation site. Using the transcription factor binding profile databases JaspAr ([http://jaspAr.genereg.net/cgi-bin/jaspAr\\_db.pl](http://jaspAr.genereg.net/cgi-bin/jaspAr_db.pl)), TFSearch (<http://www.cbrc.jp/research/db/TFSEARCH.html>) and Algen

**Table I**

The six novel variants detected in *GDF5*. Each variant was detected in single individuals from the UK. The variants were not observed in the Spanish or in the Greek cohorts

Variant	Individual	Wild-type allele	Mutant allele	Location	Predicted effect
1	Control	C	A	Promoter	Alteration of a SOX9 binding site
2	Case	G	A	Exon 1	Synonymous, Gly67Gly
3	Case	G	A	Exon 1	Non-synonymous, Gly81Arg
4	Control	C	T	Exon 2	Synonymous, Gly285Gly
5	Control	C	T	Exon 2	Synonymous, Asp287Asp
6	Control	C	G	Exon 2	Non-synonymous, Thr469Arg

Homo sapiens Gly81Arg	PPLARN-VFRPGGHSYGGGATNANARAKGGT	OTGGLTQPKKDEPKKLLP	99
Homo sapiens	PPLARN-VFRPGGHSYGGGATNANARAKGGT	OTGGLTQPKKDEPKKLLP	99
Pan troglodytes	PPLARN-VFRPGGHSYGGGATNANARAKGGT	OTGGLTQPKKDEPKKLLP	99
Equus caballus	PPLARN-VFRPGGHSYGGGAT--SARAKGGT	OTGGPTQPKKDEPKKLLP	97
Mus musculus	PPLARN-VFRPGGHIYGVGAT--NARAKGSSG	OTGGPTQPKKDEPKKMP	93
Gallus gallus	TASARAGAPRAANHGFSTGTS--KARAKSNACAG	ALLAKNDESQRVLS	97
Xenopus tropicalis	NPLPKVNATRTGILGHGVGLQ--KGRSKVPL	OSR-IFLSKNEDIKKQAA	95
Homo sapiens Thr469Arg	HAVIQTLMNSMDPESTPPTCCVFR	RLSPSILFIDSANNVVYKQYEDMVV	495
Homo sapiens	HAVIQTLMNSMDPESTPPTCCVFR	RLSPSILFIDSANNVVYKQYEDMVV	495
Pan troglodytes	HAVIQTLMNSMDPESTPPTCCVFR	RLSPSILFIDSANNVVYKQYEDMVV	495
Equus caballus	HAVIQTLMNSMDPESTPPTCCVFR	RLSPSILFIDSANNVVYKQYEDMVV	493
Mus musculus	HAVIQTLMNSMDPESTPPTCCVFR	RLSPSILFIDSANNVVYKQYEDMVV	489
Gallus gallus	HAVIQTLMNSMDPESTPPTCCVFR	RLSPSILFIDSANNVVYKQYEDMVV	494
Xenopus tropicalis	HAVIQTLMNSMDPESTPPTCCVFR	RLSPSILYTDSANVVYKQYEDMVV	490

Fig. 2. The amino acid sequence alignments for variants 3 and 6. The amino acid sequences of variant 3, Gly81Arg (top, highlighted in blue), and of variant 6, Thr469Arg (bottom, highlighted in red), were aligned with human wild-type *GDF5* and with *GDF5* of other species.

(<http://algsen.lsi.upc.es/>) we ascertained that the variant was predicted to abolish HSF and CAP binding sites, to create CdxA and Ubx binding sites and to alter SOX9, SOX10 and NFATC2 binding sites.

The glycine to arginine substitution coded for by variant 3 is at a residue conserved in mammals but not in vertebrates (Fig. 2). The substitution was predicted to be benign by the PolyPhen database, which predicts the functional effect of non-synonymous human SNPs (<http://genetics.bwh.harvard.edu/pph/>). The threonine to arginine substitution coded for by variant 6 is in the mature part of the protein at a highly conserved residue, conserved in mammals and other vertebrates. PolyPhen predicted this substitution to be potentially damaging.

#### Absence of rare variants

Our sequencing of *GDF5* did not detect any rare variants, that is variants with a MAF in the region of 0.001–0.025 (i.e., 0.1–2.5%). Instead, variants were either extremely rare (MAFs  $\leq 0.0006$ , i.e.,  $\leq 0.06\%$ ,  $n = 6$ ) or common (MAFs  $> 0.025$ , i.e., 2.5%,  $n = 7$ ), as represented in Fig. 3.

#### Discussion

It is becoming increasingly apparent that loci harbouring major susceptibility alleles for a particular trait are also likely to be

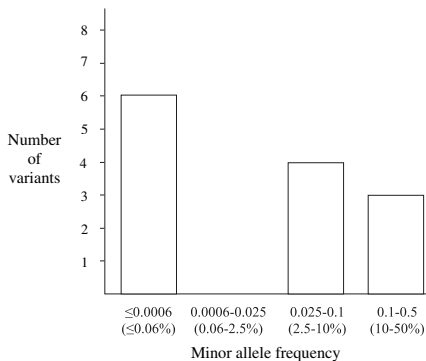


Fig. 3. The number of variants at *GDF5* displayed according to their MAFs in the UK cohort.

repositories for other alleles influencing that trait<sup>19</sup>. The growth and differentiation factor 5 gene *GDF5* contains a common functional SNP, rs143383, the T-allele of which is significantly associated with the development of OA and of other musculoskeletal phenotypes. From a previous study we had established that the effect of rs143383 on allelic expression imbalance of *GDF5* is influenced by genotype at a second 5'UTR SNP, rs143384, whilst the 3'UTR SNP rs56366915 correlates with allelic expression imbalance independent of rs143383<sup>8</sup>. rs143384 and rs56366915 are both common SNPs that are present in public databases. In this study we therefore set out to assess the allelic architecture of *GDF5* to assess whether this gene harboured other potentially functional variants that may be of relevance to OA aetiology. As far as we are aware this is the first example of a deep-sequencing analysis of an OA susceptibility locus.

We sequenced *GDF5* in cohorts of OA cases and non-OA controls from northern and southern Europe. This focus on cases and controls, rather than on individuals drawn from the general population, increased our likelihood of detecting penetrant risk or protective alleles. We demonstrated the accuracy of our sequencing by detecting all of the known variants from dbSNP that reside within the sections of the gene that we covered. We also demonstrated its sensitivity by identifying six extremely rare variants, present in single individuals from amongst the large cohort sequenced. Three of these extremely rare variants are potentially functional; variant 1, which is located in predicted transcription factor binding sites, variant 3, which codes for the substitution of a conserved glycine residue and variant 6, which codes for the substitution of a highly conserved threonine residue. Since these variants are extremely rare they cannot have an impact, at a population level, on OA susceptibility, on allelic expression imbalance mediated by rs143383 or on allelic expression imbalance independent of rs143383. If they are functional then they may have an effect at the individual and family level. Highly penetrant, deleterious mutations of *GDF5* can result in autosomal dominant brachydactyly, characterized by malformation of the phalanges<sup>2</sup>. Ethical policies with regards to genetic studies do not allow us to re-contact the UK individuals used in this study to discuss the impact of genetics on their health, so it is not possible for us to assess whether the individuals that harbour variant 1 (a control), variant 3 (a case) or variant 6 (a control) have a skeletal phenotype that was missed at their recruitment. Of the two non-synonymous variants that we detected, Gly81Arg was predicted to be benign by the PolyPhen database whilst Thr469Arg was predicted to be potentially damaging. Neither of these residues have been reported to be sites for brachydactyly mutations (<http://www.hgmd.cf.ac.uk/ac/index.php>).

In conclusion, our deep sequencing of the transcript sequence and promoter of *GDF5* has revealed that this gene harbours

extremely rare variants and that it harbours common variants but that there is an absence of variants of intermediate frequency. In Europeans there are therefore unlikely to be additional proximal variants of this locus that, at a population level, influence allelic expression imbalance of the gene and contribute to OA susceptibility. The currently known variants should be the target for future functional studies.

#### Role of the funding source

The funding bodies had no role in the design of the study, data collection, analysis and interpretation of the data, the writing of the manuscript or in the decision to submit the manuscript.

#### Author contributions

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Prof. Loughlin had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design: Dodd, Rodriguez-Fontenla, Gonzalez, Loughlin.

Acquisition of data: Dodd, Rodriguez-Fontenla, Calaza, Carr, Gomez-Reino, Tsezou, Reynard, Gonzalez, Loughlin.

Analysis and interpretation of data: Dodd, Rodriguez-Fontenla, Reynard, Gonzalez, Loughlin.

#### Conflicts of interest

The authors declare that there are no conflicts of interest.

#### Acknowledgements

We thank the individuals who contributed DNA samples and who therefore made our study possible.

Source of funding: The UK work was funded by Arthritis Research UK and by the UK NIHR Biomedical Research Centre for Ageing and Age-related disease award to the Newcastle upon Tyne Hospitals NHS Foundation Trust. The Spanish work was funded by Fondo de Investigacion Sanitaria of the Instituto de Salud Carlos III (Spain), grant 09/10431 that is partially financed by the Fondo Europeo de Desarrollo Regional program of the European Union, and by grants from the Xunta de Galicia. AWD is the recipient of a Dr William Harker Foundation studentship. CR-F is the recipient of an ISCI pre-doctoral fellowship.

#### Supplementary material

Supplementary data associated with this article can be found in online version at doi:10.1016/j.joca.2011.01.014.

#### References

- Francis-West PH, Abdelfattah A, Chen P, Allen C, Parish J, Ladhur R, et al. Mechanisms of GDF-5 action during skeletal development. *Development* 1999;126:1305–15.
- Byrnes AM, Racacho L, Nikkel SM, Xiao F, MacDonald H, Underhill TM, et al. Mutations in GDF5 presenting as semidominant brachydactyly A1. *Hum Mutat* 2010;31:1155–62.
- 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. *Nat Genet* 2007;39:529–33.
- Chapman K, Takahashi A, Meulenbelt I, Watson C, Rodriguez-Lopez J, Egli R, et al. A meta-analysis of European and Asian cohorts reveals a global role of a functional SNP in the 5' UTR of GDF5 with osteoarthritis susceptibility. *Hum Mol Genet* 2008;17:1497–504.
- Evangelou E, Chapman K, Meulenbelt I, Karassa FB, Loughlin J, Carr A, et al. Large-scale analysis of association between GDF5 and FRZB variants and osteoarthritis of the hip, knee and hand. *Arthritis Rheum* 2009;60:1710–21.
- 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* (in press). doi:10.1136/ard.2010.134155
- Southam L, Rodriguez-Lopez J, Wilkins JM, Pombo-Suarez M, Snelling S, Gomez-Reino JJ, et al. An SNP in the 5'-UTR of GDF5 is associated with osteoarthritis susceptibility in Europeans and with in vivo differences in allelic expression in articular cartilage. *Hum Mol Genet* 2007;16:2226–32.
- Egli R, Southam L, Wilkins JM, Lorenzen I, Pombo-Suarez M, Gonzalez A, et al. Functional analysis of the osteoarthritis susceptibility-associated GDF5 regulatory polymorphism. *Arthritis Rheum* 2009;60:2055–64.
- Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen WM, Bonnycastle LL, et al. Common variants in the GDF5-UGCC region are associated with variation in human height. *Nat Genet* 2008;40:198–203.
- Posthumus M, Collins M, Cook J, Handley CJ, Ribbens WJ, Smith RK, et al. Components of the transforming growth factor- $\beta$  family and the pathogenesis of human Achilles tendon pathology – a genetic association study. *Rheumatology (Oxford)* 2010;49:2090–7.
- Vaes RB, Rivadeneira F, Kerkhof JM, Hofman A, Pols HA, Uitterlinden AG, et al. Genetic variation in the GDF5 region is associated with osteoarthritis, height, hip axis length and fracture risk: the Rotterdam study. *Ann Rheum Dis* 2009;68:1754–60.
- Dai J, Shi D, Zhu P, Qin J, Ni H, Xu Y, et al. Association of a single nucleotide polymorphism in growth differentiating factor 5 with congenital dysplasia of the hip: a case-control study. *Arthritis Res Ther* 2008;10: R126.
- Rouault K, Scotet V, Autret S, Gaucher F, Dubrana F, Tanguy D, et al. Evidence of association between GDF5 polymorphisms and congenital dislocation of the hip in a Caucasian population. *Osteoarthritis Cartilage* 2010;18:1144–9.
- Storm EE, Huynh TV, Copeland NG, Jenkins NA, Kingsley DM, Lee SJ. Limb alterations in brachypodism mice due to mutations in a new member of the TGF $\beta$ -superfamily. *Nature* 1994;368:639–43.
- Settle SH, Rountree RB, Sinha A, Thacker A, Higgins K, Kingsley DM. Multiple joint and skeletal patterning defects caused by single and double mutations in the Gdf6 and Gdf5 genes. *Dev Biol* 2003;254:116–30.
- Harada M, Takahara M, Zhe P, Otsuji M, Iuchi Y, Takagi M, et al. Developmental failure of the intra-articular ligaments in mice with absence of growth differentiation factor 5. *Osteoarthritis Cartilage* 2007;15:468–74.
- Daans M, Luyten FP, Lories RJ. GDF5 deficiency in mice is associated with instability-driven joint damage, gait and subchondral bone changes. *Ann Rheum Dis* 2011;70:208–13.
- Kaliakatsos M, Tzetis M, Kanavakis E, Fytily P, Chouliaras G, Karachalios T, et al. Aspirin and knee osteoarthritis in patients of Greek origin. *Osteoarthritis Cartilage* 2006;14:609–11.
- Allen HL, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010;467:832–8.



**Supplementary Table I.** Primers used for the sequencing of *GDF5*

Sequencing of the Spanish and Greek cohorts	
Amplicon	Primers used, forward / reverse (5'-3')
1	TTTTTCCTGCCAAACCAGAG / GTCCTGGCTTGGGTTTCAG
2	GCCCGGAACGTCTTCAGGCCAG / TGCCAGGGCTTTGAAAGCCCTC
3	CTCTCAGTGGTTGGCTCTTG / AGAGCGGGCCCTTAAATCTCAT
4	CCCTGTTCCCTGGTGTGG / ATTCCAGGAGTGCAGGAAGG
Sequencing of the UK cohort	
Amplicon	Primers used, forward / reverse (5'-3')
1	ATTTTCACTGGAAAAGGATTC / AGGGGCACCCAACACAGTGC
2	ACACCCGGCAGACTTCAAGAG / GACAGATCCTGCTTTTGGGGG
3	ATTGGCCAAAGCAGAGGCCAAG / GTGAGATATGTGGTCAGTGTC
4	ATTCGAAGTGACTGGCTCCCTTG / TGGAAAGCCCTCGTACTCAAGG
5	GCCAGGACGATAAGACC / GAGTCTGTCTCCCTGGACCTG
6	TTCTGCACTCCTGGAAATCAC / GATAGGGCTCTGAAAACCTGAG





**Supplementary Table II**

The frequency of the seven common known SNPs in the UK cohort of hip and knee cases and controls, and of rs224331 and rs224330 in the Spanish and Greek cohorts of knee cases and controls

rs143383 - UK		Genotype			Allele	
		CC	CT	TT	C	T
All cases	Count	64	199	195	327	589
	%	14.0	43.4	42.6	35.7	64.3
Hip cases	Count	43	116	123	202	362
	%	15.3	41.1	43.6	35.8	64.2
Knee cases	Count	22	90	78	134	246
	%	11.6	47.4	41.0	35.3	64.7
Controls	Count	59	196	163	314	522
	%	15.3	45.3	39.4	38.0	62.0

rs73094730 - UK		Genotype			Allele	
		AA	GA	GG	A	G
All cases	Count	0	20	486	20	992
	%	0	4.0	96.0	2.0	98.0
Hip cases	Count	0	13	291	13	595
	%	0	4.3	95.7	2.1	97.9
Knee cases	Count	0	11	209	11	429
	%	0	5.0	95.0	2.5	97.5
Controls	Count	0	21	396	21	813
	%	0	5.0	95.0	3.0	97.0

rs143384 - UK		Genotype			Allele	
		CC	CT	TT	C	T
All cases	Count	78	224	189	380	602
	%	15.9	45.6	38.5	38.7	61.3
Hip cases	Count	47	130	111	224	352
	%	16.3	45.1	38.6	38.9	61.1
Knee cases	Count	32	104	85	168	274
	%	14.5	47.1	38.4	38.0	62.0
Controls	Count	67	207	151	341	509
	%	14.8	45.8	33.4	40.1	59.9

**PUBLICACIÓN 2**

rs224331 - UK		Genotype			Allele	
		TT	TG	GG	T	G
All cases	Count	218	191	61	627	313
	%	46.4	40.6	13.0	66.7	33.3
Hip cases	Count	134	121	41	389	203
	%	45.2	40.9	13.9	65.7	34.3
Knee cases	Count	91	77	21	259	119
	%	48.2	40.7	11.1	68.5	31.5
Controls	Count	171	169	47	511	263
	%	44.2	43.7	12.1	66.0	34.0

rs224330 - UK		Genotype			Allele	
		AA	AG	GG	A	G
All cases	Count	2	39	433	43	905
	%	0.4	8.2	91.4	4.5	95.5
Hip cases	Count	0	24	274	24	572
	%	0	8.1	91.9	4.0	96.0
Knee cases	Count	2	17	172	21	361
	%	1.1	8.9	90.0	5.5	94.5
Controls	Count	3	27	359	33	745
	%	0.8	6.9	92.3	4.2	95.8

rs73611720 - UK		Genotype			Allele	
		CC	CA	AA	C	A
All cases	Count	2	44	443	48	930
	%	0.4	9.0	90.6	4.9	95.1
Hip cases	Count	0	24	257	24	538
	%	0	8.5	91.5	4.3	95.7
Knee cases	Count	2	21	201	25	423
	%	0.9	9.4	89.7	5.6	94.4
Controls	Count	2	38	370	42	778
	%	0.5	9.3	90.2	5.1	94.9

rs56366915 - UK		Genotype			Allele	
		CC	CT	TT	C	T
All cases	Count	1	42	442	44	926
	%	0.2	8.7	91.1	4.5	95.5
Hip cases	Count	0	23	257	23	537
	%	0	8.2	91.8	4.1	95.9
Knee cases	Count	1	20	200	22	420
	%	0.5	9.0	90.5	5.0	95.0
Controls	Count	3	36	370	42	776
	%	0.7	8.8	90.5	5.1	94.9

rs224331 - Spain		Genotype			Allele	
		TT	TG	GG	T	G
Cases	Count	114	82	60	310	202
	%	44.5	32.0	23.4	60.5	39.5
Controls	Count	120	121	48	361	217
	%	41.5	41.9	16.6	62.5	37.5

rs224330 - Spain		Genotype			Allele	
		AA	AG	GG	A	G
Cases	Count	4	28	210	36	448
	%	1.7	11.6	86.8	7.4	92.6
Controls	Count	3	40	243	46	526
	%	1.0	14.0	85.0	8.0	92.0

rs224331 - Greece		Genotype			Allele	
		TT	TG	GG	T	G
Cases	Count	70	110	21	250	152
	%	34.8	54.7	10.4	62.2	37.8
Controls	Count	62	93	35	217	163
	%	32.6	48.9	18.4	57.1	42.9

rs224330 - Greece		Genotype			Allele	
		AA	AG	GG	A	G
Cases	Count	2	22	181	26	384
	%	1.0	10.7	88.3	6.3	93.7
Controls	Count	0	15	164	15	343
	%	0	8.4	91.6	4.2	95.8





**PUBLICACIÓN 3**

---



## RESEARCH ARTICLE

## Open Access

# Association of a *BMP5* microsatellite with knee osteoarthritis: case-control study

Cristina Rodríguez-Fontenla<sup>1</sup>, Andrew Carr<sup>2</sup>, Juan J Gomez-Reino<sup>1,3</sup>, Aspasia Tsezou<sup>4</sup>, John Loughlin<sup>5</sup> and Antonio Gonzalez<sup>1\*</sup>

## Abstract

**Introduction:** We aimed to explore the involvement of a multiallelic functional polymorphism in knee osteoarthritis (OA) susceptibility as a prototype of possible genetic factors escaping GWAS detection.

**Methods:** OA patients and controls from three European populations (Greece, Spain and the UK) adding up to 1003 patients (716 women, 287 men) that had undergone total knee joint replacement (TKR) due to severe primary OA and 1543 controls (758 women, 785 men) lacking clinical signs or symptoms of OA were genotyped for the *D6S1276* microsatellite in intron 1 of *BMP5*. Genotype and multiallelic trend tests were used to compare cases and controls.

**Results:** Significant association was found between the microsatellite and knee OA in women ( $P$  from  $3.1 \times 10^{-4}$  to  $4.1 \times 10^{-4}$  depending on the test), but not in men. Three of the alleles showed significant differences between patients and controls, one of them of increased risk and two of protection. The gender association and the allele direction of change were very concordant with those previously reported for hip OA.

**Conclusions:** We have found association of knee OA in women with the *D6S1276* functional microsatellite that modifies in *cis* the expression of *BMP5* making this a sounder OA genetic factor and extending its involvement to other joints. This result also shows the interest of analysing other multiallelic polymorphisms.

## Introduction

Investigation of the genetic component of osteoarthritis (OA) susceptibility has yielded the identification of several loci achieving genome-wide significance or consistent replication [1-3]. These loci are insufficient to account for the heritability of OA, which has been estimated as 49 to 78% for knee OA in twin and family studies [4-6]. Multiple possible causes of this discrepancy, which has become known as the missing heritability problem of complex diseases, have been discussed [7]. Some have a particular importance in OA, like the difficulties in discriminating between cases and controls, the variability in phenotype definitions and the relatively small size of genome wide association studies (GWAS) in which OA has been studied in comparison with other diseases [3,8]. Other causes are shared by most complex

diseases. They include genetic variants that are not well ascertained with the current GWAS and interactions between genetic variants, and between genetic and environmental factors that are beyond the analysis and design of current studies. Among the variants escaping GWAS detection are rare-frequency variants, even if they have a large effect, and variants with multiple alleles. The latter group is made of large structural variants and small repetitive sequences. They could escape GWAS detection because their multiple alleles cannot have sufficient linkage disequilibrium (LD) with bi-allelic markers as the single nucleotide polymorphisms (SNPs) analyzed in GWAS. One of these multi-allelic polymorphisms, a variable number tandem repeat (VNTR) polymorphism in the gene for asporin, has shown consistent association with OA in Asians and less so in Europeans [9,10]. Other multi-allelic polymorphisms have also been found associated with OA, but none of these studies has been followed by an attempt of replication [11-14].

One of the OA-associated multi-allelic polymorphisms is an intronic microsatellite in *BMP5* (bone morphogenetic

\* Correspondence: antonio.gonzalez.martinez-pedrayo@sergas.es

<sup>1</sup>Laboratorio Investigación 10 & Rheumatology, Instituto de Investigación Sanitaria - Hospital Clínico Universitario de Santiago, Choupana s/n 15706, Santiago de Compostela, Spain

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

protein 5). It has been associated with hip OA in UK women and its allelic variants showed a regulatory effect on the *BMP5* promoter *in vitro* [11]. *BMP5* is a good candidate gene for OA because BMPs are members of the transforming growth factor (TGF)-beta superfamily that were identified by their involvement in cartilage and bone development. Now, they are known to give morphogenetic signals directing tissue organization throughout the body. In particular, *BMP5* is implicated in bone morphogenesis and in the formation of the skeletal pattern, in addition to having roles in other organs [15]. It is expressed in proliferating zone chondrocytes of the growth plate and is very markedly increased by hypertrophic differentiation. This is in contrast with other BMPs [16]. In chondrocytes, upon binding to the BMP membrane receptors, *BMP5* leads to phosphorylation of p38 MAP kinase, ERK and SMAD-1,-5, and -8. Nuclear accumulation of phospho-SMADs leads to over-expression of genes involved in cartilage homeostasis as proteoglycan and of markers of hypertrophic differentiation [17]. At a cellular level, this translates into an increase of chondrocyte proliferation and in the synthesis of cartilage matrix [16]. In addition, *BMP5* expression is decreased in OA synovia [18] and bone [19]. These changes in expression together with the *BMP5* chondrogenic role and involvement in chondrocyte hypertrophy indicate *BMP5* relevance for OA pathogenesis.

We have genotyped the *D6S1276 BMP5* microsatellite in a large set of knee OA cases and disease-free controls and found an association with knee OA in women. The associated alleles and the gender specificity are similar to those previously found in UK women with hip OA. This finding reinforces the need to explore the involvement of *BMP5* genetic variation in OA and extends it to OA in other joints. It also shows that other multi-allelic variants are worthy of examination for their possible involvement in OA.

## Materials and methods

### Patients and controls

OA patients and controls were recruited at Thessaly in Greece, Santiago de Compostela in Spain and Oxford in the UK. All were European Caucasians. There were 1,003 patients (716 women, 287 men) who had undergone total knee joint replacement (TKR) due to severe primary OA. The 1,543 controls (758 women, 785 men) lacked clinical signs of OA. Full details about the patients and the controls have been reported [20]. In the Greek collection, there were 369 TKR patients (298 women and 71 men) and 383 controls (241 women and 142 men). TKR patients had a Kellgren-Lawrence (K/L) grade > 2 prior to surgery and an age range of 40 to 90 years. Rheumatoid arthritis, other autoimmune diseases, chondrodysplasias, infection-induced OA, and post-traumatic OA were excluded. The controls had a K/L grade of 0 and had undergone treatment for

injuries and fractures. They were from 46 to 88 years old. In the Spanish collection, there were 274 TKR patients (222 women and 52 men) and 462 controls (161 women and 301 men). Patients were selected as consecutive patients, of 55-80 years of age at the time of surgery, and undergoing TKR if a rheumatologist considered them to suffer from primary OA. Exclusion criteria were inflammatory, infectious, or traumatic joint pathology and lesions due to crystal deposition or osteonecrosis. Controls were older than 55 years of age and were selected during preoperative work-up for elective surgery other than joint surgery and without symptoms or signs of OA. The UK collection included 360 TKR patients (196 women and 164 men) of 56 to 85 years of age and 698 controls (356 women and 342 men) of 55 to 89 years of age. Patients showed symptoms and signs of OA of sufficient severity to require surgery, including a K/L grade  $\geq 2$  (> 90% of them had grade 3 or 4). Inflammatory arthritis (rheumatoid, polyarthritic or autoimmune disease) was excluded, as was post-traumatic or post-septic arthritis. Controls had no signs or symptoms of arthritis or joint disease (pain, swelling, tenderness or restriction of movement). All patients and controls gave their written informed consent and the use of their DNA for OA genetic studies was approved by the respective ethics committees.

### Genotyping

The *BMP5* microsatellite *D6S1276* was genotyped by length analysis of the PCR products obtained with primers 5'-FAM-atgctctggcaaatgtcaagt-3' and 5'-gccagatccctgataag-3'. Size of the fluorescence-labeled products was determined by capillary electrophoresis on an ABI 3130xl Avant Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and the corresponding microsatellite genotypes were determined using GeneMapper 3.5 software (Applied Biosystems). Several samples with different genotypes were sequenced to assess the accuracy of results with the Big Dye v3.1 Ready Reaction Kit (Applied Biosystems).

### Statistical analysis

Analyses were based on Power Marker software [21] and on the R software implementation of the Cochran-Mantel-Haenszel stratified test [22]. Exact tests for Hardy Weinberg equilibrium (HWE) were run with the Markov Chain Monte Carlo method. Deviation from HWE motivated analyses based in genotype  $2 \times n$  contingency tables and the multi-allelic trend test. These two tests are totally insensitive to HWE deviations [23]. Significance of the genotype test was obtained with the Markov Chain Monte Carlo method to avoid the limitations that sparse contingency tables pose to the asymptotic chi squared distribution method. Rare genotypes, with frequency below 1.0% were excluded from analysis. In this way, the degrees of freedom were reduced and the tests became more robust.



These analyses were conducted collection by collection and with the three collections together. Heterogeneity between collections was assessed with the population differentiation test applied to the controls and with the inconsistency ( $I^2$ ) statistic [24] applied to the Cramer's V effect size, which is the effect size most commonly used for  $2 \times n$  contingency tables. Combined analysis was done considering all data together and stratified by sample collection with the Cochran-Mantel-Haenszel method for  $2 \times n$  tables.

**Results**

**D6S1276 genotypes**

A total of 2,545 DNA samples were available for study. Valid genotypes for the *D6S1276* microsatellite were obtained from 93.32% of them with a similar distribution of the genotype call rate between the three sample collections and between cases and controls. A total of ten different alleles were observed, ranging from five to fourteen tetranucleotide repeats (Table 1). The most common was the ten-repeats allele (41.2%), followed by the eight-, nine- and eleven-repeats alleles (20.1, 20.2% and 13.4%, respectively). All the remaining alleles showed a frequency below 1.0%. These 10 alleles appeared in 28 different genotypes (not shown). The most frequent was the 10/10 genotype (17.7%) followed by the 10/9 and 10/8 genotypes (16.8 and 15.9%, respectively). Up to 14 genotypes showed a frequency below 1.0% and they were excluded from further study (for remaining genotype frequencies, see Tables S1- S3 in Additional file 1). In this way, 62 samples (or 2.4% of all the genotyped samples) were left out, with 2,313 remaining in subsequent analyses. They showed 14 different genotypes made of five alleles (from seven to eleven tetranucleotide repeats).

There was a significant deviation from HWE in the UK ( $P = 0.008$ ) controls, but not in the Spanish ( $P = 0.6$ ) or

Greek controls ( $P = 0.3$ ). The deviation motivated a careful revision of the genotypes, and their confirmation by sequencing. Also we compared the reported *BMP5* allele frequency distribution [11] with the one obtained by us in UK women and no difference was found ( $P = 0.7$ ). In addition, no difference was detected between the three collections in our study ( $P = 0.9$ ). These results showed the accuracy of the genotypes. However, we used tests insensitive to deviations from HWE as more appropriate for our data: genotype test and the multi-allelic trend tests [23].

**Association analyses by sample collection**

Previously reported association of *BMP5* with hip OA was only observed in women [11], therefore a gender-stratified analysis was done (Tables 2 and 3). Significant differences between knee OA patients and controls were only observed in women from the UK and from Greece (Table 4). In the Greek women, the multi-allelic trend test and the genotype test provided very similar results, with a significant difference between knee OA patients and controls in both of them ( $P = 0.021$  and  $P = 0.032$ ). In the UK women the two tests were discordant. The multi-allelic trend test showed a significant difference ( $P = 0.028$ ), whereas the genotype test was not significant ( $P = 0.3$ ). This discordance could be partly related to the greater power of the first test that has four degrees of freedom, in contrast with the thirteen corresponding to the genotype test. On comparison, no significant differences were observed in the Spanish women or in any of the men (Table 4).

**Combined association across sample collections**

In a first analysis, we simply combined all samples together without considering their origin (Table 5). This approach was justified in the lack of heterogeneity of *D6S1276* effects between the three sample collections

**Table 1 Allele counts and frequency of the *D6S1276* *BMP5* microsatellite in patients with knee osteoarthritis (OA) and in controls from the three sample collections**

Allele	Controls								Patients with knee OA							
	Spain		Greece		UK		Total		Spain		Greece		UK		Total	
	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%
5	1	0.1					1	0.0							3	0.2
6	1	0.1	2	0.3	2	0.2	5	0.2	1	0.2	4	0.6			5	0.3
7	39	4.4	23	3.4	58	4.4	120	4.2	18	3.6	12	1.7	35	4.9	65	3.4
8	160	18.2	126	18.9	256	19.6	542	19.0	116	23.2	139	20.1	158	22.3	413	21.7
9	188	21.4	147	22.0	282	21.6	617	21.6	97	19.4	139	20.1	104	14.7	340	17.9
10	362	41.2	264	39.5	525	40.3	1151	40.4	210	42.0	292	42.2	303	42.8	805	42.4
11	115	13.1	98	14.7	168	12.9	381	13.4	53	10.6	102	14.7	101	14.3	256	13.5
12	11	1.3	6	0.9	12	0.9	29	1.0	5	1.0	4	0.6	4	0.6	13	0.7
13			2	0.3	1	0.1	3	0.1								
14	1	0.1					1	0.0								
Total	878		668		1,304		2,850		500		692		708		1,900	

**Table 2 Allele counts and frequency of the D6S1276 BMP5 microsatellite in women with knee osteoarthritis (OA) and control women from the three sample collections**

Alleles	Controls								Patients with knee OA							
	Spain		Greece		UK		Total		Spain		Greece		UK		Total	
	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%
5					1	0.2	1	0.1	1	0.2	4	0.7	1	0.3	1	0.1
6									16	3.9	11	2	17	4.4	44	3.2
7	16	5.2	20	4.8	28	4.4	64	4.7	91	22.2	114	20.2	95	24.6	300	22.1
8	52	17	70	16.8	117	18.2	239	17.5	82	20	113	20	56	14.5	251	18.5
9	66	21.6	99	23.8	137	21.3	302	22.1	174	42.4	240	42.6	164	42.5	578	42.5
10	126	41.2	159	38.2	263	41	548	40.2	43	10.5	79	14	50	13	172	12.6
11	42	13.7	65	15.6	87	13.6	194	14.2	3	0.7	3	0.5	3	0.8	9	0.7
12	4	1.3	3	0.7	9	1.4	16	1.2								
Total	306		416		642		1,364		410		564		386		1,360	

( $P^2 = 0.0\%$ ) and lack of differences in the microsatellite genotypes between the three populations (above). Both the multi-allelic trend test and the genotype test were clearly significant in women ( $P = 3.3 \times 10^{-4}$ , and  $P = 4.1 \times 10^{-4}$ , respectively). We also performed a combined analysis accounting for origin of the samples with the Cochran-Mantel-Haenszel test applied to the  $2 \times 14 \times 3$  genotype contingency table (Table 5). The result was very similar to the simple pooling showing association in the female patients ( $P = 3.8 \times 10^{-4}$ ). No association was observed in men (Table 3).

Inspection of the genotypes showed that the difference between female patients with knee OA and female controls was not due to a single genotype but to several (Figure 1), including an increase in patients of the 10/8 genotype (20.4% vs. 12.6%, odds ratio (OR) 1.77, 95% CI 1.32-2.39,  $P = 0.0002$ ) and a decrease of the 10/9 genotype (14.4% vs. 21.0%, OR 0.63, 95% CI 0.47-0.84,  $P = 0.002$ ). No differences were detected when the same comparisons

were made in men although variation was in the same direction

**Discussion**

Our study has identified the D6S1276 microsatellite in BMP5 as a new genetic factor for knee OA in women. The D6S1276 microsatellite has previously been associated with hip OA [11]. Concordance with this study extends beyond the association because the two studies found association in women and with a similar pattern of alleles. The concordance of our findings was very clear with the first set of samples in the hip OA study, where the risk allele was also the eight-repeats allele ( $P = 0.006$ ) and protection was mediated by the seven-repeats allele ( $P = 0.03$ ), and other alleles showed changes in the same direction [11]. The second set of samples was not associated with D6S1276 in Wilkins et al., but in the combined analysis of the two sets the association remained as well as the protective effect of the seven-repeats allele ( $P = 0.03$ ) and the

**Table 3 Allele counts and frequency of the D6S1276 BMP5 microsatellite in men with knee osteoarthritis (OA) and in control men from the three sample collections**

Alleles	Controls								Knee OA							
	Spain		Greece		UK		Total		Spain		Greece		UK		Total	
	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%
5	1	0.2					1	0.1					2	0.6	2	0.4
6	1	0.2	2	0.8	1	0.2	4	0.3								
7	23	4	3	1.2	30	4.5	56	3.8	2	2.2	1	0.8	18	5.6	21	3.9
8	108	18.9	56	22.2	139	21	303	20.4	25	27.8	25	19.5	63	19.5	113	20.9
9	122	21.3	48	19	145	21.9	315	21.2	15	16.7	26	20.3	48	14.9	89	16.5
10	236	41.3	105	41.7	262	39.6	603	40.6	36	40	52	40.6	139	43.2	227	42.0
11	73	12.8	33	13.1	81	12.2	187	12.6	10	11.1	23	18	51	15.8	84	15.6
12	7	1.2	3	1.2	3	0.5	13	0.9	2	2.2	1	0.8	1	0.3	4	0.7
13			2	0.8	1	0.2	3	0.2								
14	1	0.2					1	0.1								
Total	572		252		662		1,486		90		128		322		540	

**Table 4 Association of the D6S1276 BMP5 microsatellite with knee osteoarthritis (OA) by sample collection and by gender**

Collection	Gender	Number		P-values	
		Controls	Knee OA	Multi-allelic trend test	Genotype test
Spain	women	149	200	0.2	0.5
	men	277	43	0.4	0.4
	all	426	243	0.1	0.5
Greece	women	205	275	0.021	0.032
	men	120	63	0.8	0.6
	all	325	338	0.2	0.09
UK	women	312	187	0.028	0.2
	men	325	157	0.09	0.3
	all	637	344	0.0059	0.1

direction of change of the eight-repeats allele. Therefore, our study does not amount to a formal replication because of the difference in the involved joint, but strongly reinforces the evidence in support of the implication of this *BMP5* polymorphism in OA susceptibility. In addition, our finding indicates that this microsatellite could be associated with independence of the joint.

It is very unlikely that the association of *D6S1276* with knee OA could be due to confounding by population stratification. The association was found in two independent sample collections with origin in Greece and the UK, and reinforced in the analyses with samples from Spain without any detectable heterogeneity as shown by the lack of inconsistency ( $I^2 = 0\%$ ) and by the very similar results obtained pooling data from the three collections or using the Cochran-Mantel-Haenszel approach. Lack of differences in *D6S1276* between collections was also shown by the absence of population differentiation between the controls ( $P = 0.91$ ), and by lack of significant differences in any of the pairwise comparisons between controls, and differences only in patients from Greece (Table S2 in Additional file 1). In addition, there are arguments to exclude significant population stratification in each of the three sample collections taken individually. The UK samples were genotyped in the recently reported arcOGEN GWAS [1]. In this study no significant population stratification was detected (genomic control inflation factor  $\lambda_{1000} = 1.009$ ). The Spanish

samples are from subjects reporting a Spanish origin extending for two or more generations. In fact, 94.7% of the patients and 95.4% of the controls are of an even more homogeneous ancestry, as all their known ancestors were from Galicia, a region in the North-West of Spain. Excluding subjects of a wider Spanish ancestry did not alter the results (not shown). Finally, the Greek collection was made of individuals of Greek descent who were inhabitants of Central Greece and ethnic or racial differences between patients and controls were not detected in multiple previous studies [13,14,20,25].

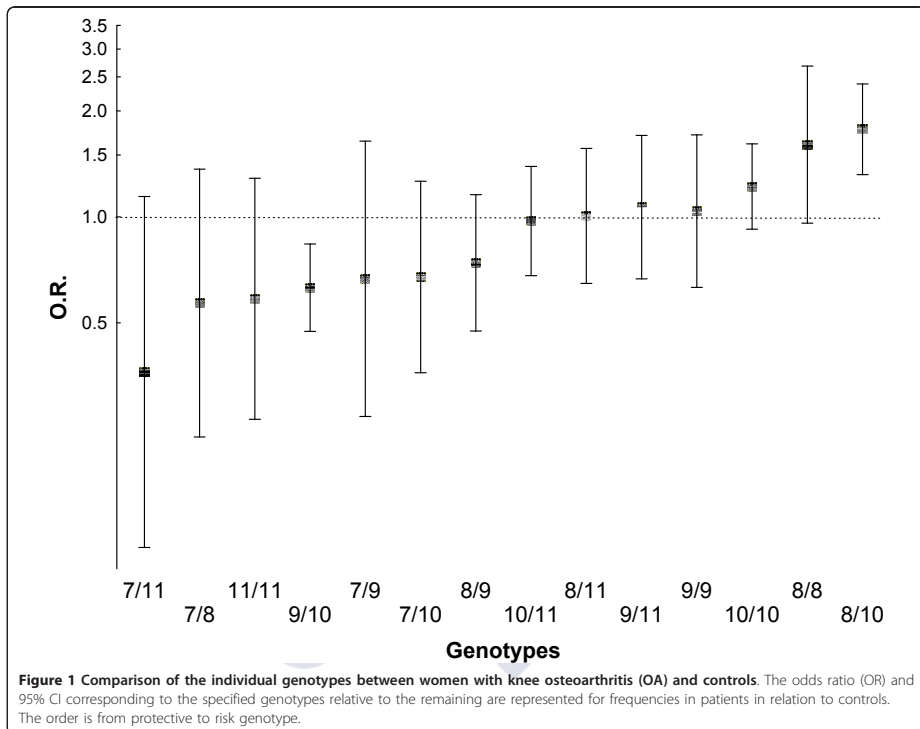
The identification of the *D6S1276* microsatellite as an important polymorphism in hip OA has been the fruit of successive studies by the group of John Loughlin [11,26-29]. Initially, linkage with hip OA was detected in chromosome 6 after stratification by affected joint of a genome-wide linkage study performed in sibling-pairs who had undergone total hip replacement (THR) for primary OA [26]. The linked region was very broad ( $> 50$  cM). Fine mapping with additional microsatellites and THR families reduce the locus to 11.4 cM at 6p12.3-q13 and showed that it was specific to women [27]. This study also increased the statistical evidence for linkage. Two candidate genes in the interval were excluded by association analysis with SNPs mapping to their coding sequence and promoter, *COL9A1* [27] and *BMP5* [28]. However, a subsequent study including additional microsatellites confirmed the same peak of linkage and found association in intron 1 (*D6S1276*) and 3' to *BMP5*, bringing forward the possibility that the causal polymorphism could be in *cis*-regulatory sequences of this gene [29]. This hypothesis was tested in the last study to date, where more samples and polymorphisms within intron 1 were analyzed [11]. Here, association was confirmed with *D6S1276* and found only with one of the new polymorphisms. Subsequent functional analysis with reporter gene assays showed that only *D6S1276* was able to modify the expression of the luciferase gene under the *BMP5* promoter [11]. This analysis also demonstrated a significant amount of variability of modifier effect depending on the cell line where the study was done, the *D6S1276* allele and the alleles at three nearby SNPs.

The variability of effects of *D6S1276* is congruent with what is known of the regulation of *BMP5*. In effect, regulation of this gene is very modular, precise and complex involving multiple *cis*-acting enhancers with tissue- and location-specific effects [17,30,31]. This has been clearly demonstrated in the mouse, where mutations in *BMP-5* showed cartilage and bone defects affecting seemingly unrelated skeletal elements in addition to several soft-tissue organs [15]. This indicated a complex pattern of transcription regulation that was confirmed with the identification of multiple *cis*-regulatory modular enhancers each of them acting in particular tissues or even in

**Table 5 Association of the D6S1276 BMP5 microsatellite with knee osteoarthritis (OA) across sample collections**

Gender	Multi-allelic trend test		Genotype test	
	Simple pooling	CMH <sup>a</sup>	Simple pooling	CMH <sup>a</sup>
Women	0.00033	0.00041	0.00038	0.00038
Men	0.11	0.2	0.4	0.4
All	0.00073	0.0076	0.0092	0.0092

<sup>a</sup>CMH, Cochran-Mantel-Haenszel stratified analysis



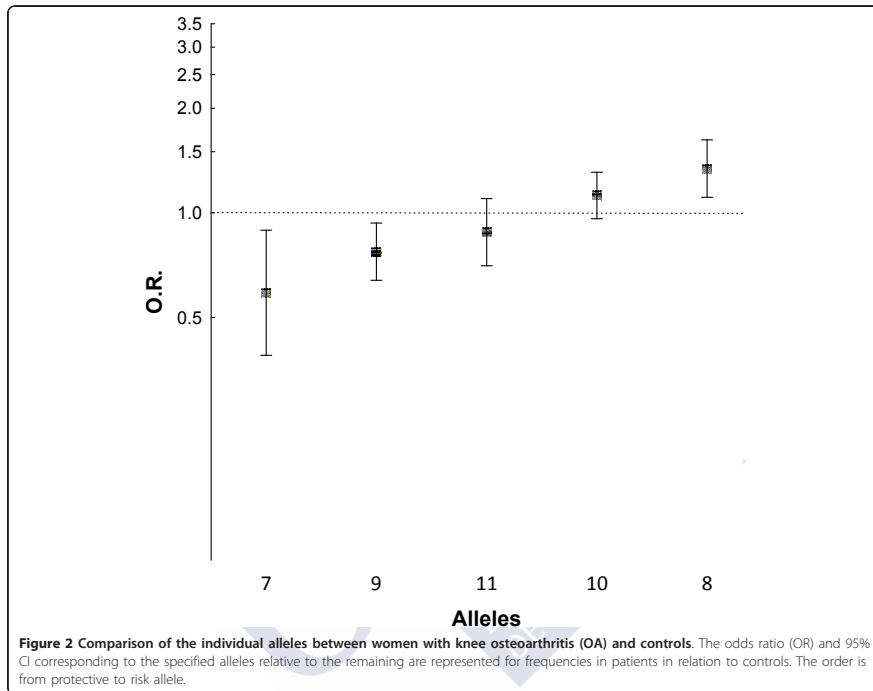
different regions of the same organ [30,31]. Most of these enhancers are 3' to the coding sequence but also an enhancer directing expression in different areas of the mouse ribs has been identified in one of the *Bmp5* introns [31]. A similar complex and precise regulation seems to be present in humans because analysis of differential allelic expression showed extreme variability between the joint tissues with differences even between regions of the same cartilage surface in function of their proximity to the OA lesion [32].

Association of *D6S1276* is not contradicted by the lack of replication in the Spanish collection because it was most likely due to lack of power (below 0.46 for the most significant differences found in the current study) and not to significant differences between the Spanish and the other two collections (Table S4 in Additional file 1). In addition, the association we have found was not questioned by the lack of an association signal in the GWAS because the multiplicity of the microsatellite-associated alleles makes for very poor correlation with any bi-allelic

SNP. This has been already shown in the previous studies of hip OA, where none of the *BMP5* SNPs analyzed accounted for the association of *D6S1276* [11,28]. Therefore, our initial hypothesis positing that some of the unidentified genetic susceptibility to OA could be in multi-allelic polymorphisms is supported by our results in *BMP5* and should motivate analysis of other similar loci. In addition, the lack of linkage with knee OA in the original genome-wide linkage study that brought attention to the *BMP5* locus is not against the association we have found, because this study included fewer families with knee OA (34 families, including affected women and men) than with hip OA in women (85 families plus 44 families with hip OA in men), and therefore was less sensitive to detect any difference specific to the women of this subgroup [26].

**Conclusions**

We have found association of knee OA in women with the *D6S1276* functional microsatellite that modifies in *cis*



the expression of *BMP5*. The results are very concordant with those observed previously for hip OA, making this association sounder, and extending its range of involvement to other joints. These results also show the interest of analyzing other multi-allelic polymorphisms for their possible role in OA susceptibility.

**Additional material**

**Additional file 1: Genotype frequencies and comparison across populations.** Genotype counts and frequencies in patients with knee OA and in controls from the three sample collections are provided as total, in women and in men. The genotype frequencies were also compared pairwise between the sample collections.

**Abbreviations**

BMP5: bone morphogenetic protein 5; HWE: Hardy Weinberg equilibrium; K/L: Kellgren-Lawrence; LD: linkage disequilibrium; OA: osteoarthritis; PCR: polymerase chain reaction; SNP: single nucleotide polymorphism; THR: total hip replacement; TKR: total knee replacement; VNTR: variable number tandem repeat.

**Acknowledgements**

We thank Carmen Pena-Pena for technical assistance and Manuel Calaza for help with data management. C R-F has a pre-doctoral bursary from the Instituto de Salud Carlos III (Spain). The study was supported by grant PI09/01431 of the Instituto de Salud Carlos III that is partially financed by the European Regional Development Fund of the European Union, and by grants from the Xunta de Galicia. None of the funding sources was involved in any decision about this manuscript.

**Author details**

<sup>1</sup>Laboratorio Investigación 10 & Rheumatology, Instituto de Investigación Sanitaria - Hospital Clínico Universitario de Santiago, Choupana s/n 15706, Santiago de Compostela, Spain. <sup>2</sup>Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Windmill Road, OX3 7HE, Oxford, UK. <sup>3</sup>Department of Medicine, University of Santiago de Compostela, Santiago de Compostela, San Francisco s/n, 15701, Spain. <sup>4</sup>Department of Biology, University of Thessaly, Medical School, 22 Papakyriazi Street, 41222, Larissa, Greece. <sup>5</sup>Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

**Authors' contributions**

C R-F contributed to the acquisition and analysis of data, drafting the article and final approval of the submitted version. AC, JIG-R, AT and JL contributed to the acquisition and interpretation of data, critical revision of the article for important intellectual content and final approval of the submitted version. AG contributed to the conception and design of the

study, analysis and interpretation of data, drafting the article and final approval of the submitted version. AG takes responsibility for the integrity of the work as a whole, from inception to finished article. All authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

Received: 6 September 2012 Revised: 19 November 2012  
 Accepted: 22 November 2012 Published: 27 November 2012

#### References

- Zeggini E, Panoutsopoulou K, Southam L, Day-Williams A, Lopes M, Boraska V, Beazley C, Blackburn H, Bumpstead S, Hunt S, Potter S, Shin S, Yadav V, Deloukas P, Rayner N, Elliott K, Esko T, Metspalu A, Evangelou E, Ioannidis J, Hofman A, Kerkhof H, Rivadeneira F, Utterlinden A, van Meurs J, Houwing-Duistermaat J, Kloppenburg M, Bos S, Slagboom P, Meulenbelt I, et al: **Identification of new susceptibility loci for osteoarthritis (arCOGEN): a genome-wide association study.** *Lancet* 2012, **380**:815-823.
- Day-Williams AG, Southam L, Panoutsopoulou K, Rayner NW, Esko T, Estrada K, Helgadottir HT, Hofman A, Ingvarsson T, Jonsson H, Keis A, Kerkhof HJ, Thorleifsson G, Arden NK, Carr A, Chapman K, Deloukas P, Loughlin J, McCaskie A, Ollier WE, Ralston SH, Spector TD, Wallis GA, Wilkin JM, Aslam N, Birell F, Carluke I, Joseph J, Rai A, Reed M, et al: **A variant in MCF2L1 is associated with osteoarthritis.** *Am J Hum Genet* 2011, **89**:446-450.
- Loughlin J: **Osteoarthritis year 2010 in review: genetics.** *Osteoarthritis Cartilage* 2011, **19**:342-345.
- Neame RL, Muir K, Doherty S, Doherty M: **Genetic risk of knee osteoarthritis: a sibling study.** *Ann Rheum Dis* 2004, **63**:1022-1027.
- MacGregor AJ, Li Q, Spector TD, Williams FM: **The genetic influence on radiographic osteoarthritis is site specific at the hand, hip and knee.** *Rheumatology (Oxford)* 2009, **48**:277-280.
- Page WF, Hoaglund FT, Steinbach LS, Heath AC: **Primary osteoarthritis of the hip in monozygotic and dizygotic male twins.** *Twin Res* 2003, **6**:147-151.
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy ML, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarron SA, Visscher PM: **Finding the missing heritability of complex diseases.** *Nature* 2009, **461**:747-753.
- Kerkhof HJ, Meulenbelt I, Akune T, Arden NK, Aromaa A, Bierma-Zeinstra SM, Carr A, Cooper C, Dai J, Doherty M, Doherty SA, Felson D, Gonzalez A, Gordon A, Harilainen A, Hart DJ, Hauksson VB, Heliövaara M, Hofman A, Ikegawa S, Ingvarsson T, Jiang Q, Jonsson H, Jonsdottir I, Kawaguchi H, Kloppenburg M, Kujala UM, Lane NE, Leino-Arjas P, Lohmander LS, et al: **Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium.** *Osteoarthritis Cartilage* 2011, **19**:254-264.
- Kizawa H, Kou I, Iida A, Sudo A, Miyamoto Y, Fukuda A, Mabuchi A, Kotani A, Kawakami A, Yamamoto S, Uchida A, Nakamura K, Notoya K, Nakamura Y, Ikegawa S: **An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis.** *Nat Genet* 2005, **37**:138-144.
- Nakamura T, Shi D, Tzatis M, Rodriguez-Lopez J, Miyamoto Y, Tsezou A, Gonzalez A, Jiang Q, Kamatani N, Loughlin J, Ikegawa S: **Meta-analysis of association between the ASPN D-repeat and osteoarthritis.** *Hum Mol Genet* 2007, **16**:1676-1681.
- Wilkins JM, Southam L, Mustafa Z, Chapman K, Loughlin J: **Association of a functional microsatellite within intron 1 of the BMP5 gene with susceptibility to osteoarthritis.** *BMC Med Genet* 2009, **10**:141.
- Magana JJ, Galvez-Rosas A, Gonzalez-Huerta C, Duarte-Salazar C, Lara-Alvarado L, Soria-Bastida MA, Cortes-Gonzalez S, Miranda-Duarte A: **Association of the calcitonin gene (CA) polymorphism with osteoarthritis of the knee in a Mexican mestizo population.** *Knee* 2010, **17**:157-160.
- Fytilli P, Giannatou E, Papanikolaou V, Stripeli F, Karachalios T, Malizos K, Tsezou A: **Association of repeat polymorphisms in the estrogen receptors alpha, beta, and androgen receptor genes with knee osteoarthritis.** *Clin Genet* 2005, **68**:268-277.
- Fytilli P, Giannatou E, Karachalios T, Malizos K, Tsezou A: **Interleukin-10G and interleukin-10R microsatellite polymorphisms and osteoarthritis of the knee.** *Clin Exp Rheumatol* 2005, **23**:621-627.
- King JA, Marker PC, Seung KJ, Kingsley DM: **BMP5 and the molecular, skeletal, and soft-tissue alterations in short ear mice.** *Dev Biol* 1994, **166**:112-122.
- Mallhot G, Yang M, Mason-Savas A, Mackay CA, Leav I, Odgren PR: **BMP-5 expression increases during chondrocyte differentiation in vivo and in vitro and promotes proliferation and cartilage matrix synthesis in primary chondrocyte cultures.** *J Cell Physiol* 2008, **214**:56-64.
- Snelling SJ, Hulley PA, Loughlin J: **BMP5 activates multiple signaling pathways and promotes chondrogenic differentiation in the ATDC5 growth plate model.** *Growth Factors* 2010, **28**:268-279.
- Bramlage CP, Haupt T, Kaps C, Ungethüm U, Krenn V, Pruss A, Müller GA, Strutz F, Burmester GR: **Decrease in expression of bone morphogenetic proteins 4 and 5 in synovial tissue of patients with osteoarthritis and rheumatoid arthritis.** *Arthritis Res Ther* 2006, **8**:R58.
- Hogwood B, Tsykin A, Findlay DM, Fazalari NL: **Microarray gene expression profiling of osteoarthritic bone suggests altered bone remodelling, WNT and transforming growth factor-beta/bone morphogenic protein signalling.** *Arthritis Res Ther* 2007, **9**:R100.
- Rodriguez-Lopez J, Mustafa Z, Pombo-Suarez M, Malizos KN, Rego I, Blanco FJ, Tsezou A, Loughlin J, Gomez-Reino JJ, Gonzalez A: **Genetic variation including nonsynonymous polymorphisms of a major aggrecanase, ADAMTS-5, in susceptibility to osteoarthritis.** *Arthritis Rheum* 2008, **58**:435-441.
- Liu K, Muse SV: **PowerMarker: an integrated analysis environment for genetic marker analysis.** *Bioinformatics* 2005, **21**:2128-2129.
- Team RDC: **R: A Language and Environment for Statistical Computing.** Vienna, Austria: R Foundation for Statistical Computing; 2011.
- Czika W, Weir BS: **Properties of the multiallelic trend test.** *Biometrics* 2004, **60**:69-74.
- Higgins JP, Thompson SG, Deeks JJ, Altman DG: **Measuring inconsistency in meta-analyses.** *BMJ* 2003, **327**:557-560.
- Dieguez-Gonzalez R, Galaza M, Shi D, Meulenbelt I, Loughlin J, Tsezou A, Dai J, Malizos KN, Slagboom EP, Kloppenburg M, Chapman K, Jiang Q, Kremer D, Gomez-Reino JJ, Nakajima N, Ikegawa S, Gonzalez A: **Testing the druggable endothelial differentiation gene 2 knee osteoarthritis genetic factor for replication in a wide range of sample collections.** *Ann Rheum Dis* 2009, **68**:1017-1021.
- Loughlin J, Mustafa Z, Iven C, Smith A, Carr AJ, Sykes B, Chapman K: **Stratification analysis of an osteoarthritis genome screen-suggestive linkage to chromosomes 4, 6, and 16.** *Am J Hum Genet* 1999, **65**:1795-1798.
- Loughlin J, Mustafa Z, Dowling B, Southam L, Marcelline L, Raina SS, Ala-Kokko L, Chapman K: **Finer linkage mapping of a primary hip osteoarthritis susceptibility locus on chromosome 6.** *Eur J Hum Genet* 2002, **10**:562-568.
- Southam L, Chapman K, Loughlin J: **Genetic association analysis of BMP5 as a potential osteoarthritis susceptibility gene.** *Rheumatology (Oxford)* 2003, **42**:911-912.
- Southam L, Dowling B, Ferreira A, Marcelline L, Mustafa Z, Chapman K, Bentham G, Carr A, Loughlin J: **Microsatellite association mapping of a primary osteoarthritis susceptibility locus on chromosome 6p123-q13.** *Arthritis Rheum* 2004, **50**:3910-3914.
- DiLeone RJ, Marcus GA, Johnson MD, Kingsley DM: **Efficient studies of long-distance Bmp5 gene regulation using bacterial artificial chromosomes.** *Proc Natl Acad Sci USA* 2000, **97**:1612-1617.
- Guenther C, Pantalena-Filho L, Kingsley DM: **Shaping skeletal growth by modular regulatory elements in the Bmp5 gene.** *PLoS Genet* 2008, **4**: e1000308.
- Wilkins JM, Southam L, Price AJ, Mustafa Z, Carr A, Loughlin J: **Extreme context specificity in differential allelic expression.** *Hum Mol Genet* 2007, **16**:537-546.

doi:10.1186/ar4102  
**Cite this article as:** Rodriguez-Fontenla et al: Association of a BMP5 microsatellite with knee osteoarthritis: case-control study. *Arthritis Research & Therapy* 2012 **14**:R257.

**Table S1:** Genotype counts and frequencies in patients with knee OA and in controls from the three sample collections after excluding genotypes with frequency  $< 0.01$  overall

Genotype	UK				GR				ES			
	Controls		Knee OA		Controls		Knee OA		Controls		Knee OA	
	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency
10-10	110	0,1735	68	0,1977	52	0,1600	60	0,1775	80	0,1878	48	0,1975
10-11	60	0,0946	43	0,1250	35	0,1077	40	0,1183	44	0,1033	22	0,0905
10-7	23	0,0363	12	0,0349	10	0,0308	3	0,0089	16	0,0376	9	0,0370
10-8	84	0,1325	57	0,1657	48	0,1477	77	0,2278	64	0,1502	47	0,1934
10-9	126	0,1987	51	0,1483	65	0,2000	51	0,1509	71	0,1667	34	0,1399
11-11	16	0,0252	10	0,0291	7	0,0215	4	0,0118	7	0,0164	5	0,0206
11-7	13	0,0205	5	0,0145	6	0,0185	2	0,0059	5	0,0117		
11-8	33	0,0521	20	0,0581	27	0,0831	26	0,0769	24	0,0563	12	0,0494
11-9	27	0,0426	11	0,0320	16	0,0492	26	0,0769	27	0,0634	9	0,0370
7-8	8	0,0126	6	0,0174	4	0,0123	4	0,0118	8	0,0188	4	0,0165
7-9	9	0,0142	4	0,0116	3	0,0092	3	0,0089	9	0,0211	3	0,0123
8-8	38	0,0599	29	0,0843	11	0,0338	6	0,0178	13	0,0305	16	0,0658
8-9	56	0,0883	17	0,0494	22	0,0677	20	0,0592	36	0,0845	19	0,0782
9-9	31	0,0489	11	0,0320	19	0,0585	16	0,0473	22	0,0516	15	0,0617
Total	634	1	344	1	325	1	338	1	426	1	243	1

**Table S2:** Genotype counts and frequencies in women with knee OA and in control women from the three sample collections after excluding genotypes with frequency < 0.01 overall

Genotype	UK				GR				ES			
	Controls		Knee OA		Controls		Knee OA		Controls		Knee OA	
	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency
10-10	56	0,1812	39	0,2086	28	0,1366	51	0,1855	26	0,1745	41	0,2050
10-11	25	0,0809	18	0,0963	22	0,1073	31	0,1127	21	0,1409	17	0,0850
10-7	12	0,0388	7	0,0374	8	0,0390	3	0,0109	5	0,0336	7	0,0350
10-8	37	0,1197	35	0,1872	25	0,1220	63	0,2291	20	0,1342	37	0,1850
10-9	67	0,2168	24	0,1283	46	0,2244	40	0,1455	27	0,1812	30	0,1500
11-11	11	0,0356	5	0,0267	4	0,0195	2	0,0073	2	0,0134	3	0,0150
11-7	4	0,0129	2	0,0107	6	0,0293	2	0,0073	1	0,0067		
11-8	18	0,0583	10	0,0535	16	0,0780	21	0,0764	8	0,0537	11	0,0550
11-9	16	0,0518	8	0,0428	13	0,0634	21	0,0764	7	0,0470	9	0,0450
7-8	5	0,0162	1	0,0053	3	0,0146	3	0,0109	6	0,0403	4	0,0200
7-9	5	0,0162	2	0,0107	3	0,0146	3	0,0109	4	0,0268	3	0,0150
8-8	17	0,0550	20	0,1070	6	0,0293	6	0,0218	3	0,0201	13	0,0650
8-9	23	0,0744	9	0,0481	14	0,0683	15	0,0545	11	0,0738	12	0,0600
9-9	13	0,0421	7	0,0374	11	0,0537	14	0,0509	8	0,0537	13	0,0650
Total	309		187		205		275		149		200	



**Table S3:** Genotype counts and frequencies in men with knee OA and in control men from the three sample collections after excluding genotypes with frequency < 0.01 overall

Genotype	UK				GR				ES			
	Controls		Knee OA		Controls		Knee OA		Controls		Knee OA	
	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency
10-10	54	0,1662	29	0,1847	24	0,2000	9	0,1429	54	0,1949	7	0,1628
10-11	35	0,1077	25	0,1592	13	0,1083	9	0,1429	23	0,0830	5	0,1163
10-7	11	0,0338	5	0,0318	2	0,0167			11	0,0397	2	0,0465
10-8	47	0,1446	22	0,1401	23	0,1917	14	0,2222	44	0,1588	10	0,2326
10-9	59	0,1815	27	0,1720	19	0,1583	11	0,1746	44	0,1588	4	0,0930
11-11	5	0,0154	5	0,0318	3	0,0250	2	0,0317	5	0,0181	2	0,0465
11-7	9	0,0277	3	0,0191					4	0,0144		
11-8	15	0,0462	10	0,0637	11	0,0917	5	0,0794	16	0,0578	1	0,0233
11-9	11	0,0338	3	0,0191	3	0,0250	5	0,0794	20	0,0722		
7-8	3	0,0092	5	0,0318	1	0,0083	1	0,0159	2	0,0072		
7-9	4	0,0123	2	0,0127					5	0,0181		
8-8	21	0,0646	9	0,0573	5	0,0417			10	0,0361	3	0,0698
8-9	33	0,1015	8	0,0510	8	0,0667	5	0,0794	25	0,0903	7	0,1628
9-9	18	0,0554	4	0,0255	8	0,0667	2	0,0317	14	0,0505	2	0,0465
Total	325		157		120		63		277		43	

**Table S4:** Comparisons of D6S1276 genotype frequencies between the sample collections.

	UK vs. GR	UK vs. ES	GR vs. ES
All			
Controls	> 0.49	> 0.34	> 0.76
Knee OA	0.0006	> 0.54	0.01
Women			
Controls	> 0.80	> 0.33	> 0.57
Knee OA	0.0006	> 0.76	> 0.16

Comparisons were done with Clump<sup>1</sup> and the reported P value is the lowest obtained with any of the four test included in this application  
<sup>1</sup>Sham PC, Curtis D: **Monte Carlo tests for associations between disease and alleles at highly polymorphic loci.**  
*Ann Hum Genet* 1995, **59**:97-105.



**ANÁLISIS DE GENES CANDIDATOS EN UN  
METAANÁLISIS DE 9 ESTUDIOS DE ASOCIACIÓN  
DE GENOMA COMPLETO**

---



## MATERIAL Y MÉTODOS

La realización de este trabajo se llevó a cabo gracias a la colaboración establecida con el Treat OA (*Translational Research in Europe Applied Technologies for Osteoarthritis*). Por ello, debemos diferenciar dos apartados en el material y métodos. Uno de ellos es el metanálisis de los 9 GWAS a partir del cuál elaboramos el estudio de los genes candidatos. Los datos de este metanálisis fueron proporcionados por Evangelis Evangelou que fue quién lo realizó. La selección de SNPs, la definición de los genes candidatos, los análisis estadísticos para el cálculo de SNPs independientes así como la replicación de los SNPs seleccionados corresponde a la contribución de esta tesis.

### **Colecciones de muestras.**

El metaanálisis se realizó a partir de los resultados de asociación de 9 GWAS de ascendencia europea incluyendo pacientes con OA de rodilla, OA de cadera y controles (Tabla 1): deCODE (Islandia)<sup>245</sup>, tres colecciones del estudio de Rotterdam (Rotterdam I, II y III)<sup>246</sup>, USA (Framingham Osteoarthritis study)<sup>247</sup>, UK (Twins UK), GARP<sup>248</sup>, arcOGEN<sup>143</sup> y EGCUT(Estonia)<sup>249</sup>. Otras colecciones de muestras no incluidas en los GWAS se utilizaron para completar el análisis de los resultados significativos (Tabla 1). Incluyeron colecciones del Norte de España<sup>250-252</sup>, de Grecia Central<sup>253</sup>, de Nottingham y del estudio de GOAL<sup>254,255</sup>. Las características de estas colecciones muestrales se describen en detalle en artículos previos. El estudio de deCODE incluye pacientes con reemplazamiento articular tanto de rodilla como de cadera y

controles de la población, excluyendo aquellos individuos con susceptibilidad a OA (rodilla, mano, cadera) incluidos en las listas de hospitales y centros de salud de Islandia. Los estudios de Rotterdam, Framingham y Twins UK incluyen pacientes con OA radiográfica y controles de la misma cohorte sin OA radiográfica, definidos según la estandarización de fenotipos propuesta recientemente con casos con  $KL > 2$  y controles con  $KL < 2$ <sup>154</sup>. El estudio GARP está formado por pacientes con OA confirmada clínica y radiográficamente en dos o más articulaciones<sup>248</sup>. La fase I de arcOGEN incluyó OA de cadera y OA de rodilla determinada mediante evidencia radiográfica y clínica que requiere reemplazamiento articular en la mayoría de los pacientes<sup>143</sup>. El estudio de EGCUT incluyó OA confirmada radiográficamente ( $KL > 2$ ) y los controles no mostraron ningún síntoma de OA. Todas las colecciones de muestras empleadas en el estudio de extensión provienen de estudios caso-control. Los pacientes fueron sujetos con reemplazamiento articular debido a OA de cadera confirmada radiográficamente. Los pacientes del estudio de GOAL se seleccionaron por OA sintomática severa de rodilla. Todas estas colecciones muestrales y estudios genéticos han sido aprobados por los comités éticos correspondientes y las muestras se obtuvieron con el consentimiento informado de los pacientes.

### **Selección de genes candidatos y SNPs.**

Para identificar los genes candidatos que han sido estudiados hasta el momento hemos utilizado la herramienta Phenopedia del HuGE (*Human Genome Epidemiology*) Navigator<sup>256</sup>. Esta base de datos recopila los estudios genéticos publicados desde el año 2000. Los términos de

búsqueda a fecha del 20 de abril del 2013, fueron “artrosis” y otros dos fenotipos relacionados: “osteofitosis espinal” y “desplazamiento del disco intervertebral”. Los genes candidatos de estos dos últimos fenotipos mostraron gran solapamiento con los de OA (90% y 72% respectivamente). Los 17 genes no solapados se consideraron de interés para el descubrimiento de nuevos loci de OA. Los genes del cromosoma X se excluyeron porque no es posible realizar la imputación de los genotipos. También se excluyó el gen *FRA1H* por corresponder a una referencia bibliográfica errónea. Además, todos los genes de GWAS de Europeos que mostraron asociación previa se excluyeron. Las posiciones de los loci que comprenden a los genes candidatos se definieron tomando un criterio arbitrario de selección de 50 kb *downstream* del codón de parada y 50 kb *upstream* del codón de inicio. Estas posiciones se obtuvieron de acuerdo con el ensamblaje GhCR37 (Febrero 2007) y la base de datos de Ensembl ([http://www.ensembl.org/Homo\\_sapiens/Info/Index](http://www.ensembl.org/Homo_sapiens/Info/Index)). Los loci solapantes se fusionaron en un único locus. La herramienta del UCSC Genome browser, Batch Coordinate Conversion (*liftOver*) (<http://genome.ucsc.edu/cgi-bin/hgLiftOver>) se utilizó para convertir las coordenadas genómicas obtenidas en Ensembl a las de HapMap CEU release 27 (fases I+II+III). Los SNPs con MAF >5% en población europea CEU de los correspondientes genes candidatos, se descargaron de HapMap mediante un programa realizado en Perl que interactúa con el servidor HapMart (<http://hapmap.ncbi.nlm.nih.gov/biomart/martview/>). Todos los SNPs se alinearon respecto a la cadena positiva del DNA para evitar ambigüedades.

### Genotipado e imputación.

Las tecnologías de genotipado utilizadas para los GWAS incluídas en el metaanálisis fueron diferentes y han sido descritas con detalle en estudios previos<sup>143,210,229</sup>. Para el genotipado e imputación de los SNPs no genotipados se utilizaron los programas MACH<sup>257</sup> e IMPUTE<sup>258</sup> con los datos de HapMap CEU release 22 (*phases* I+II). El genotipado de las colecciones de replicación (España y Grecia) se realizó mediante extensión de una única base utilizando el kit SNaPshot Multiplex (Applied Biosystems). Las colecciones de Nottingham y GOAL se genotiparon en UK por Kbioscience Ltd, mediante química KASPar.

### Análisis estadísticos.

Cada equipo responsable de cada uno de los GWAS realizó un análisis por alelos para testar la asociación tanto en OA de rodilla como de cadera. El factor de inflación lambda ( $\lambda$ ) se calculó para el efecto específico de cada sexo utilizando el método de control genómico y los errores estándar se corrigieron por la raíz cuadrada del factor de inflación lambda ( $SE_{corrected} = SE_{observed} \times \sqrt{\lambda}$ ). En los estudios GARP y deCODE los errores estándar se estimaron teniendo en cuenta las relaciones de parentesco. Para el meta-análisis, el *effect size* (tamaño del efecto) de cada SNP se calculó utilizando el modelo de efectos fijos ponderado por la inversa de la varianza que sintetiza los tamaños de efecto y los errores estándar corregidos por el factor  $\lambda$ . La heterogeneidad se calculó con el estadístico  $I^2$  y se realizó un metaanálisis con el software METAL para combinar los efectos obtenidos. Se consideraron seis análisis incluyendo



dos niveles por tipo de articulación (rodilla y cadera) y tres niveles por género (mujeres, hombres y total). Para el cálculo del número efectivo de tests (número de SNPs independientes) se utilizó una adaptación del algoritmo de simple M desarrollado por Gao<sup>259</sup>. La modificación de este algoritmo consiste en sustituir la matriz de correlaciones observada por la matriz semidefinida positiva más cercana, implementada en el paquete del software R, *corpcor*. Esta modificación es necesaria para corregir las desviaciones introducidas por los genotipos incompletos. Los resultados del estudio de extensión se combinaron con el test de Mantel-Haenszel. El análisis conjunto del estudio de extensión con los datos de los GWAS se realizó utilizando un metaanálisis de efectos fijos con el software R. El poder estadístico del estudio se analizó con el software *Power and Sample Size*<sup>260</sup>. Un análisis del power se proporciona en la Figura S1 como un ejemplo ilustrativo. El poder estadístico del estudio fue de un 80% para detectar la asociación de un SNP con MAF=20% y OR=1.15 en OA de rodilla y 1.16 en OA de cadera.



Metaanálisis	OA de rodilla		OA de cadera		diagnóstico de OA
	Casos	Controles	Casos	Controles	
arcOGEN	1643	4896	1728	4896	TJR and ROA <sup>a</sup>
deCODE*	1312	2318	1423	2318	TJR <sup>a</sup> and clinical ROA <sup>a</sup>
Framingham Study	417	1667	NA	1671	ROA <sup>a</sup> and clínica
GARP	154	1671	106	3233	ROA <sup>a</sup>
Rotterdam Study I	1476	3233	760	1472	ROA <sup>a</sup>
Rotterdam Study II	369	1472	159	1487	ROA <sup>a</sup>
Rotterdam Study III	152	1487	41	228	ROA <sup>a</sup>
TwinsUK	113	228	68	2531	ROA <sup>a</sup>
EGCUT			64		ROA <sup>a</sup>
Total GWAS	5636	16972	4349	17836	
Spain <sup>c</sup>	-	-	695	784	TJR <sup>b</sup>
Greece	-	-	92	358	TJR <sup>b</sup>
Nothingham <sup>d</sup>	-	-	1246	713	TJR <sup>b</sup>
GOAL <sup>d</sup>	-	-	1270	763	TJR and ROA <sup>a</sup>
Total extension			3303	2618	

**Tabla 1. Características de las colecciones de muestras incluídas en el metaanálisis de GWAS y en el estudio de extensión.**

<sup>a</sup> ROA = OA radiográfica; <sup>b</sup> TJR = reemplazamiento articular de cadera

\* tamaño efectivo deCODE. <sup>c</sup> Muestras de las tres colecciones del Norte de España: Santiago, Santander y A Coruña <sup>d</sup> Estas muestras sólo se genotiparon para el rs833058(VEGF)



## RESULTADOS

### Identificación de genes candidatos.

Se seleccionaron 199 genes candidatos por su estudio previo en la genética de la OA mediante una búsqueda en el navegador HuGe Navigator (20 abril del 2013). Algunos de los genes son de loci que mostraron asociación con OA por primera vez en los GWAS (*HLA de clase II/III*, *A2BP1* y *LRCH1*) y otros en estudios de ligamiento de genoma completo (*MATN3*, *DIO2*, *FRZB* y *BMP5*). Estos loci no se excluyeron porque la identificación de algunos de ellos como genes de susceptibilidad viene en parte de análisis post hoc de su posible función biológica y porque ninguno de ellos alcanzó el nivel de significación requerido en los GWAS  $p < 5 \times 10^{-8}$ . Todos los genes se muestran en la Tabla 2. Los 199 genes se agruparon en 158 segmentos genómicos no solapantes que contienen 27501 SNPs autosómicos (MAF>5%) con genotipos conocidos en la población CEU de HapMap.

### Metaanálisis de asociación de los SNPs de los genes candidatos.

Se combinaron los tamaños de efecto (log OR) de los SNPs de cada uno de los 9 GWAS en un metanálisis. 25839 SNPs de los 27501 incluidos en los genes candidatos pasaron los controles de calidad de genotipado o de imputación. El tamaño de efecto de cada uno de ellos en los 9 GWAS fue calculado mediante metaanálisis de efectos fijos. Los valores de p resultantes fueron interpretados considerando como umbral de significación  $p=1.58 \times 10^{-5}$ . Este umbral fue obtenido al aplicar la

**ANÁLISIS DE GENES CANDIDATOS EN UN METAANÁLISIS DE 9 ESTUDIOS DE ASOCIACIÓN DE GENOMA COMPLETO**

corrección de Bonferroni con los 3156 tests independientes calculados teniendo en cuenta que los SNPs muestran desequilibrio de ligamiento. (Tabla 3). Esta estimación se realizó mediante el algoritmo de simple M.

**Tabla 2. Genes candidatos (199) seleccionados para OA mediante la base de datos HuGE Navigator.**

<i>A2BP1</i>	<i>CNTF</i>	<i>HAPLN1</i>	<i>KIR2DL1</i>	<i>NOS2</i>	<i>TLR4</i>
<i>ACAN</i>	<i>COL10A1</i>	<i>HFE</i>	<i>KIR2DL2</i>	<i>NOS3</i>	<i>TLR9</i>
<i>ACE</i>	<i>COL11A1</i>	<i>HIF1A</i>	<i>KIR2DL3</i>	<i>OLIG3</i>	<i>TNF</i>
<i>ADAM12</i>	<i>COL11A2</i>	<i>HLA-B</i>	<i>KIR2DL4</i>	<i>OPRM1</i>	<i>TNFAIP3</i>
<i>ADAMTS14</i>	<i>COL1A1</i>	<i>HLA-DQA1</i>	<i>KIR2DL5A</i>	<i>PAPPA2</i>	<i>TNFAIP6</i>
<i>ADAMTS3</i>	<i>COL1A2</i>	<i>HLA-DQB1</i>	<i>KIR2DP1</i>	<i>PAPSS2</i>	<i>TNFRSF11B</i>
<i>ADAMTS5</i>	<i>COL2A1</i>	<i>HLA-DRB1</i>	<i>KIR2DS1</i>	<i>PCSK6</i>	<i>TNFRSF1A</i>
<i>AIF1</i>	<i>COL3A1</i>	<i>HLA-DRB5</i>	<i>KIR2DS2</i>	<i>PITX1</i>	<i>TNFRSF1B</i>
<i>ANP32A</i>	<i>COL5A1</i>	<i>HMGA2</i>	<i>KIR2DS3</i>	<i>PLA2G4A</i>	<i>TNFSF11</i>
<i>APOE</i>	<i>COL5A2</i>	<i>HOXB9</i>	<i>KIR2DS4</i>	<i>PPARG</i>	<i>TP53</i>
<i>APOM</i>	<i>COL6A1</i>	<i>HTRA1</i>	<i>KIR3DL1</i>	<i>PTGS2</i>	<i>TRAF1</i>
<i>ASPEN</i>	<i>COL9A1</i>	<i>ICAM1</i>	<i>KIR3DL2</i>	<i>PTH</i>	<i>TRPV1</i>
<i>B2M</i>	<i>COL9A2</i>	<i>IBSP</i>	<i>KIR3DL3</i>	<i>PTPN22</i>	<i>TXNDC3</i>
<i>BAT2</i>	<i>COL9A3</i>	<i>IFNG</i>	<i>KIR3DS1</i>	<i>RAMP2</i>	<i>TXNRD2</i>
<i>BAT3</i>	<i>COMP</i>	<i>IGF1</i>	<i>KL</i>	<i>RHOB</i>	<i>USP33</i>
<i>BAT4</i>	<i>COMT</i>	<i>IGFBP7</i>	<i>LECT1</i>	<i>ROR2</i>	<i>VDR</i>
<i>BAT5</i>	<i>CRP</i>	<i>IHH</i>	<i>LEP</i>	<i>SCN9A</i>	<i>VEGF</i>
<i>BDNF</i>	<i>CSNK1D</i>	<i>IL10</i>	<i>LEPR</i>	<i>SELS</i>	<i>WISP1</i>
<i>BMP2</i>	<i>CSNK2B</i>	<i>IL12A</i>	<i>LOC344875</i>	<i>SEPP1</i>	<i>WRN</i>
<i>BMP5</i>	<i>CYP19A1</i>	<i>IL13</i>	<i>LPAR1/EDG2</i>	<i>SERPINE1</i>	
<i>BTNL2</i>	<i>CYP2D6</i>	<i>IL18</i>	<i>LRCH1</i>	<i>SERPINA3</i>	
<i>C5</i>	<i>DCT</i>	<i>IL18R1</i>	<i>LRP5</i>	<i>SLC26A2</i>	
<i>C6orf10</i>	<i>DIO2</i>	<i>IL18RAP</i>	<i>LRP6</i>	<i>SLC6A4</i>	
<i>C6orf47</i>	<i>DIO3</i>	<i>IL1A</i>	<i>LTA</i>	<i>SMAD3</i>	
<i>CALCA</i>	<i>ENG</i>	<i>IL1B</i>	<i>LY6G5C</i>	<i>SOC1</i>	
<i>CALCRL</i>	<i>ENPP1</i>	<i>IL1R1</i>	<i>MATN1</i>	<i>SOST</i>	
<i>CALM1</i>	<i>EPAS1</i>	<i>IL1R2</i>	<i>MATN3</i>	<i>SPP1</i>	
<i>CALM2</i>	<i>ESR1</i>	<i>IL1RL1</i>	<i>MMP1</i>	<i>STAT4</i>	
<i>CAT</i>	<i>ESR2</i>	<i>IL1RL2</i>	<i>MMP13</i>	<i>TBX4</i>	
<i>CASP9</i>	<i>FRZB</i>	<i>IL1RN</i>	<i>MMP2</i>	<i>TGFB1</i>	
<i>CCL2</i>	<i>GHR</i>	<i>IL2</i>	<i>MMP3</i>	<i>THBS2</i>	
<i>CCR5</i>	<i>GNB3</i>	<i>IL4</i>	<i>MMP8</i>	<i>THRA</i>	
<i>CD36</i>	<i>GPX1</i>	<i>IL4R</i>	<i>MMP9</i>	<i>TIMP2</i>	
<i>CILP</i>	<i>GSTM1</i>	<i>IL6</i>	<i>MTHFR</i>	<i>TIMP4</i>	
<i>CLEC3B</i>	<i>GSTP1</i>	<i>IL8</i>	<i>NCOR2</i>	<i>TLR2</i>	
<i>CNRI</i>	<i>GSTT1</i>	<i>KIAA1217</i>	<i>NFKB1</i>	<i>TLR3</i>	

**Tabla 3. Estimación del número de tests independientes que corresponden a los SNPs en los genes candidatos de cada cromosoma. Sólo se consideraron SNPs con MAF>5% .**

Cromosoma	# SNPs <sup>a</sup>	# nr SNPs <sup>b</sup>	# test independientes
1	2295	1547	350
2	3136	2315	259
3	743	568	82
4	1108	810	169
5	875	636	64
6	4596	3393	569
7	1184	1040	83
8	598	466	57
9	1619	1090	221
10	2210	1741	164
11	1262	961	119
12	1311	1182	123
13	936	726	106
14	636	545	63
15	1315	1104	180
16	1331	1183	126
17	1097	896	190
19	370	327	90
20	302	283	44
21	276	242	48
22	301	235	49
<b>Total</b>	<b>27501</b>	<b>21290</b>	<b>3156</b>

<sup>a</sup>Estimación con una modificación del algoritmo simple M descrito por Gao X, Starmer J, Martin ER. Genet Epidemiol. 2008; 32:361-9

<sup>b</sup> nr = no redundante, se selecciona un solo SNP de entre un grupo con  $r^2 = 1$  entre ellos.

Aplicando este umbral de significación, dos genes candidatos, *COL11A1* y *VEGF*, mostraron asociación con OA de cadera y ningún gen mostró asociación con OA de rodilla.

Dos SNPs, rs4907986 y rs1241164, en los extremos 5' y 3' del *COL11A1* respectivamente, se encontraron asociados en el análisis sin estratificar (Tabla 4) (Figura 1A). Son dos asociaciones independientes con una débil  $r^2$  entre ellas (0.09). Un SNP del *COL11A1*, rs2615977, ya había mostrado asociación en la fase 1 del estudio de arcOGEN ( $p=1.1 \times 10^{-5}$ ) que se solapa con este metaanálisis. Sin embargo, ninguno de los dos SNPs independientes asociados en el metaanálisis mostró una elevada  $r^2$  con el rs2615977 ( $r^2 < 0.4$ ). En el análisis estratificado por sexo se encontró asociación de otro SNP del *COL11A1* en el subgrupo de mujeres. Este SNP mostró una  $r^2 = 0.68$  con el rs4907986 pero no está en LD con el rs1241164.

El segundo gen candidato asociado con OA de cadera fue *VEGF*. Sólo se encontró un SNP asociado, rs833058 que alcanzó el umbral requerido en el subgrupo de varones (Tabla 4). No se encontró ningún otro SNP con  $r^2 > 0.5$  (moderada o fuerte) con el rs833058 (Figura 1B).

Ninguno de los otros 197 genes candidatos mostró asociación con OA de cadera o rodilla al nivel requerido

### **Estudio de extensión y combinado.**

Tratando de obtener más evidencia sobre los tres SNPs asociados con OA de cadera, se analizaron en nuevas muestras. Esta parte del estudio no se planeó como una replicación de los resultados debido al pequeño tamaño



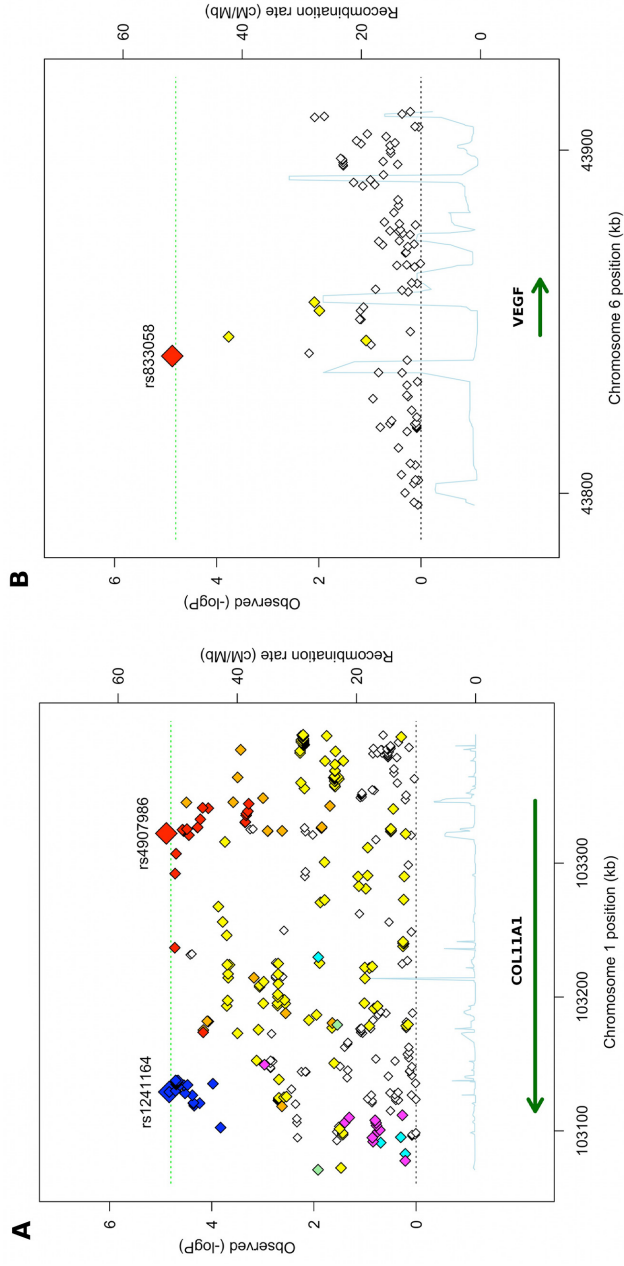
de las colecciones muestrales independientes disponibles. Se han genotipado los dos top SNPs del *COL11A1* en 1929 muestras de Grecia y España (784 pacientes de OA de cadera y 1142 controles). No se encontró asociación significativa en este análisis pero cuando estos datos se combinan con el metaanálisis de GWAS, la asociación del rs1241164 se hace más clara ( $p=5.3 \times 10^{-6}$ ) (Tabla 5). La segunda señal que corresponde al rs4907986 se mostró asociada más débilmente en el análisis combinado.

El SNP de *VEGF*, rs833058, se genotipó en 5921 muestras adicionales (3303 pacientes con OA de cadera y 2618 controles) procedentes de Reino Unido, España y Grecia. No se encontró asociación significativa en hombres (1466 pacientes y 1263 controles) (Tabla 5) aunque la dirección del cambio fue la misma observada en el metaanálisis de GWAS. Los resultados combinados mostraron un nivel de p menor en hombres ( $p=2.6 \times 10^{-5}$ ) que el encontrado en el metaanálisis. Sin embargo, este SNP mostró una asociación débil en el estudio de extensión ( $p=0.03$ ) al considerarse tanto hombres como mujeres.



**Tabla 4. SNPs asociados (  $p < 1.58 \times 10^{-5}$  ) independientes con OA de cadera en todas las muestras o en los análisis estratificados en mujeres y hombres. Se muestran las OR y sus intervalos de confianza (95 %) relativos al primer alelo.**

Locus	SNP	Chr	Posición	Alelos	OR (95 % CI)	P-valor	$I^2$	Valor de $P_{het}$
Todo	<i>COL11A1</i>	1	103322221	T/C	1.12 (1.06-1.17)	$1.29 \times 10^{-5}$	0	0.63
	<i>COL11A1</i>	1	103129065	T/C	0.82 (0.74-0.89)	$1.47 \times 10^{-5}$	0	0.92
Mujeres	<i>COL11A1</i>	1	103345324	A/T	0.87 (0.82-0.92)	$1.29 \times 10^{-5}$	33.2	0.16
Hombres	<i>VEGF</i>	6	43839832	T/C	0.85 (0.79-0.91)	$1.35 \times 10^{-5}$	18.6	0.29

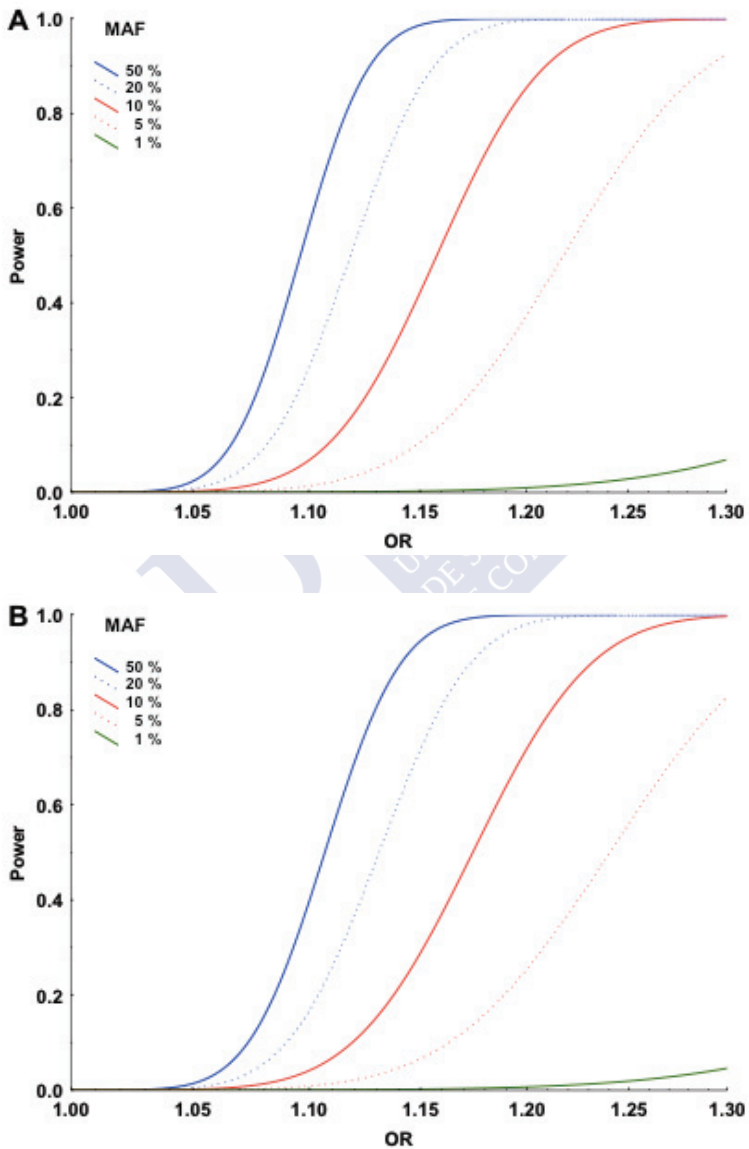


**Figura 1. Gráfico de los niveles de asociación de A) COL11A1 and B) VEGF con OA de cadera.** Cada SNP es uno de los rombos localizados en el mapa genético (eje X) y la escala de  $-\log 10$  representa su p-valor de asociación (eje Y). La línea horizontal punteada muestra el umbral de significación,  $1.58 \times 10^{-5}$ . La línea continua azul representa la tasa de recombinación. En A, la correlación de cada SNP con el rs4907986 se muestra desde un mínimo representado en color blanco hasta un máximo ( $r^2 \geq 0.8$ ) en color rojo, a través de color naranja para la  $r^2 \geq 0.5 < 0.8$ . Sin embargo, la correlación de cada SNP con el rs1241164 se muestra desde un mínimo en color blanco hasta un máximo ( $r^2 \geq 0.8$ ) en color azul a través de color fucsia para la  $r^2 \geq 0.5 < 0.8$ . En B, la correlación de cada SNP con el top asociado (rs833058) se muestra desde un mínimo en color blanco hasta un máximo ( $r^2 \geq 0.8$ ) en color rojo.

**Tabla 5: Resultados del metaanálisis de GWAS, estudio de extensión y análisis combinado de ambos para los dos SNPs independientes del COL11A1 y VEGF asociados con OA de cadera. No se detectó heterogeneidad significativa ( $p < 0.05$ ) en ninguno de los análisis combinados.**

Gen	SNP	Sexo	GWAS metaanálisis		Estudio de extensión		Combinado: GWAS meta + extensión	
			O.R. <sub>M-H</sub> (95%CI)	p-valor	O.R. <sub>M-H</sub> (95%CI)	p-valor	O.R. (95%CI)	p-valor
<i>COL11A1</i> <sup>a</sup>	rs1241164	All	0.82 (0.74, 0.89)	$1.47 \times 10^{-3}$	0.85 (0.68, 1.06)	0.15	0.82 (0.75, 0.89)	$5.3 \times 10^{-6}$
		Female	0.79 (0.7, 0.89)	$1.27 \times 10^{-4}$	0.96 (0.70, 1.32)	0.81	0.81 (0.72, 0.91)	$2.4 \times 10^{-4}$
		Male	0.84 (0.74, 0.97)	0.012	0.77 (0.56, 1.06)	0.11	0.83 (0.73, 0.94)	$3.7 \times 10^{-3}$
<i>COL11A1</i> <sup>b</sup>	rs4907986	All	1.12 (1.06, 1.17)	$1.29 \times 10^{-5}$	0.98 (0.85, 1.12)	0.77	1.09 (1.05, 1.15)	$5.8 \times 10^{-5}$
		Female	1.13 (1.07, 1.21)	$1.05 \times 10^{-4}$	1.15 (0.95, 1.39)	0.13	1.13 (1.07, 1.21)	$3.2 \times 10^{-5}$
		Male	1.1 (1.03, 1.19)	0.0069	0.95 (0.77, 1.18)	0.67	1.08 (1.01, 1.16)	0.015
<i>VEGF</i>	rs833058	All	0.92 (0.88, 0.97)	0.0022	0.92 (0.85, 0.99)	0.03	0.92 (0.88, 0.96)	$1.9 \times 10^{-4}$
		Female	0.99 (0.93, 1.06)	0.75	0.91 (0.81, 1.01)	0.07	0.97 (0.91, 1.02)	0.21
		Male	0.85 (0.79, 0.91)	$1.3 \times 10^{-5}$	0.94 (0.84, 1.06)	0.31	0.87 (0.82, 0.93)	$2.6 \times 10^{-5}$

Figura S1: Potencia estadística del metaanálisis de GWAS para una asociación con  $p < 1.58 \times 10^{-5}$  en OA de rodilla (A) o de cadera (B), correspondiente a SNPs con distintas MAF y OR. Estas estimaciones asumen ausencia de heterogeneidad entre GWAS.





**DISCUSIÓN**

---





### **Relación entre la edad de inicio de los síntomas de OA y la carga genética de la enfermedad.**

En este trabajo hemos querido comprobar si los individuos portadores de un mayor número de alelos de riesgo de los polimorfismos estudiados, podrían tener una edad más temprana de inicio de los síntomas de OA de rodilla. Ningún estudio previo analizó la existencia de esta relación en OA, aunque un trabajo en SLE (lupus eritematoso sistémico) comprobó que los pacientes con un mayor número de alelos de riesgo mostraban fenotipos más severos de la enfermedad<sup>242</sup>. Sin embargo, los resultados obtenidos en OA indican que no existe ninguna evidencia que sostenga la hipótesis planteada y por lo tanto el componente genético es independiente de la edad de inicio.

Una de las consecuencias directas de este estudio sería que seleccionar pacientes con una edad temprana de inicio de los síntomas no conllevaría ninguna ventaja en el proceso de identificación de nuevos loci de susceptibilidad. Lo mismo ocurriría si se seleccionan pacientes con una elevada carga genética para ensayos clínicos. También se pone de relieve la dificultad de llevar a cabo una clasificación de los pacientes de OA primaria en subgrupos de diferente etiología basados en la genética o la edad avanzada como se ha propuesto en algún estudio<sup>243</sup>.

La importancia de realizar este trabajo viene dada por el contexto en el que se encontraba hace un par de años el estudio de la genética de la OA. Los estudios que buscaban encontrar nuevos loci asociados eran negativos y/o obtenían muy pocas asociaciones demostradas consistentemente. Los

## *DISCUSIÓN*

---

dos únicos SNPs asociados con  $p > 5 \times 10^{-8}$  en aquel momento pertenecían a *GDF-5* y el locus 7q22.

Esta situación, llevó a cuestionar si el diseño de los estudios genéticos en OA era el correcto. ¿Podría deberse este número bajo de loci de susceptibilidad asociados a un problema en la selección de los fenotipos incluidos? En otras enfermedades como la esquizofrenia o la obesidad se presta especial atención a la heterogeneidad de los fenotipos estudiados<sup>261,262</sup>. La heterogeneidad en la definición de los fenotipos en los diferentes estudios genéticos reduce el poder estadístico de encontrar señales asociadas consistentemente<sup>263</sup>. En el caso de la OA, la OA sintomática o la OA radiográfica se consideran conjuntamente, a pesar de que se conoce que las definiciones de fenotipos influyen la asociación con determinadas variables clínicas, como la edad, sexo y BMI<sup>145</sup>. Otro factor a considerar es que se pueden realizar diseños de estudios genéticos más potentes seleccionando fenotipos más extremos, es decir controles libres de OA de edad avanzada y pacientes con edad temprana de inicio de los síntomas. Es en este punto concreto dónde se encuadraría este trabajo. Aunque los resultados son negativos y seleccionar pacientes con edad temprana no supondría una ventaja hay que considerar que no se deben extrapolar estos resultados a todos los loci del genoma<sup>264,265</sup>.

El pequeño porcentaje de varianza que es detectable en este estudio refleja el elevado poder estadístico del análisis de regresión cuando las variables dependientes (edad de inicio de los síntomas) e independientes (carga genética) son continuas y se aproximan a una distribución normal. Éste no es el caso de los análisis en los que la variable dependiente es dicotómica (caso-control) y la independiente tiene tres niveles que son los genotipos

del SNP (AA, Aa, aa). Desde que se realizó este estudio se han descubierto nuevos loci asociados. Si se incrementara el número de loci de susceptibilidad en el análisis mejoraría su sensibilidad. Así, incluyendo diez loci de susceptibilidad en lugar de los seis analizados, el poder estadístico pasaría a ser de un 99%, en lugar de un 80%. Además como cada locus puede mostrar características propias diferenciales, no se puede descartar que en el futuro se descubran loci que tengan una relación con la edad de inicio de los síntomas.

Es interesante tratar de definir las implicaciones de estos resultados en relación con los mecanismos de la OA. La hipótesis testada implica una aceleración de la OA determinada por el nivel de carga genética. Los resultados no la sostienen, lo que implica que edad y riesgo genético contribuyen independientemente a la OA. Esta separación podría ocurrir en modelos de enfermedad en los que existe una relación dinámica entre daño y reparación. Además, la progresión de la enfermedad sería en fases con posibilidad de remisión o períodos de estabilidad.

Este tipo de escenario sería consistente con hallazgos recientes en el seguimiento de biomarcadores del cartílago. En algunos pacientes con OA de rodilla se produce progresión de la enfermedad. Sin embargo, en otros la enfermedad se estabiliza y no progresa con el paso del tiempo. Estos patrones diferentes se relacionan con los niveles de la proteína COMP (*serum cartilage oligomeric matrix protein*) y son consistentes con la hipótesis de una progresión en distintas fases de la OA<sup>266</sup>. Otra evidencia que apoya la evolución en etapas, también se ha obtenido al seguir las lesiones detectadas mediante resonancia magnética en la médula ósea y cartílago de la rodilla. Estas lesiones se asocian con pérdida del cartílago

## *DISCUSIÓN*

---

y cambios estructurales que conllevan dolor en la articulación. El descubrimiento de estas lesiones permite predecir la progresión de las siguientes fases de la enfermedad<sup>267-269</sup>.

Esta evolución en fases no contradice la asociación entre edad avanzada y OA. La edad avanzada contribuye a la pérdida de equilibrio entre síntesis y degradación del cartílago, favoreciendo éste último proceso. El tejido pierde la capacidad de mantener la homeostasis y llega un momento en que no es posible su reparación. Esto es debido al proceso senescente de las condrocitos acompañado de un incremento de la fragilidad en la ECM. Con la edad, los condrocitos articulares pierden la capacidad no sólo proliferativa sino también de síntesis de componentes de la ECM. Sin embargo continúan produciendo mediadores inflamatorios y enzimas degradativas. Todos estos cambios contribuyen al desarrollo de OA<sup>270</sup>.

Sólo los SNPs de *GDF-5* y *7q22* han alcanzado el umbral de significación requerido en los GWAS y por lo tanto sólo ellos pueden considerarse como confirmados. Esto es una limitación del estudio ya que cuando se inició no había más loci asociados a ese nivel. El resto de SNPs estudiados mostraron asociación pero muy por debajo del umbral de significación de los GWAS. Este es el caso de *DVWA*, *DIO3*, *PTGS2* y *ASPN*<sup>161,202,210,271</sup>. En la actualidad hay otros loci asociados con OA de rodilla a este nivel como *MCF2L*<sup>152</sup> y *GNL3/GLT8D1* descubierto en el estudio de arcOGEN. No podemos asegurar que su inclusión en el estudio no pudiera proporcionar resultados diferentes.

En conclusión, se ha encontrado que no existe relación entre la carga genética de riesgo para OA de rodilla y la edad de inicio de los síntomas. Esto no excluye que la variación genética en un locus particular pueda

influenciar la edad de inicio de los síntomas. Esto sería probable en loci que influyen en el envejecimiento del tejido o en el dolor. Sin embargo, ninguno de los seis loci estudiados tuvieron este efecto.



### **Resecuenciado del gen *GDF-5* buscando variantes raras asociadas**

Tras resecuenciar los dos exones y las regiones 5' y 3' UTR de *GDF-5* en un amplio tamaño muestral no se han encontrado variantes raras en el gen que puedan estar asociadas con OA. Este resultado indica que aparte de los SNPs descritos con anterioridad<sup>140</sup>, no existe otro tipo de variantes en el gen que influyeran la susceptibilidad a OA.

Se ha demostrado en numerosos estudios que aquellos loci que muestran algún tipo de variantes genéticas asociadas con una enfermedad con frecuencia pueden portar otros tipos de variantes de susceptibilidad. Por ejemplo, un estudio genético de altura demostró múltiples SNPs comunes asociados de forma independiente en los mismos loci en los que se encontraban variantes raras<sup>239</sup>. Otro ejemplo es el mismo gen *GDF-5* en el que se encuentran mutaciones que afectan a la forma y función de las articulaciones y variantes comunes asociadas con OA de rodilla y altura<sup>194,216,217,220</sup>. Por ello, con este estudio se ha querido completar el mapa de la arquitectura genética de *GDF-5* debido a la relevancia de este locus en la etiología de la OA y utilizarlo como un punto de partida para el estudio de variantes raras en la OA.

Se identificaron seis variantes extremadamente raras, cuatro en controles y dos en casos en la cohorte de Reino Unido. Su frecuencia es muy baja,  $\leq 0,06$ , menor que la requerida para ser consideradas variantes raras (0,1-2,5%). Se confirma además, la presencia de siete variantes comunes con  $MAF > 1\%$  ya descritas en bases de datos como dbSNP.

Tres de las nuevas variantes extremadamente raras podrían tener un posible papel funcional. La variante 1 se encontró en el promotor del gen y se localiza en un sitio de unión predicho para SOX-9, factor clave en el crecimiento del cartílago. La variante 3 es exónica y provoca una sustitución nucleotídica en un residuo conservado de glicina. La variante 6 también es exónica y codifica un cambio de treonina por arginina en el exón 2. Ninguna de estas variantes coinciden con las mutaciones descritas previamente en *GDF-5* para el fenotipo de braquidactilia<sup>223</sup>. Debido a la muy baja frecuencia de estas variantes se descartó que puedan tener efecto en la susceptibilidad a OA. En todo caso podrían tener un papel funcional específico a nivel individual o familiar aunque no fue posible validar esta hipótesis. Otro problema de su baja frecuencia es la imposibilidad de llevar a cabo estudios de expresión alélica diferencial como en el caso del rs143383<sup>188</sup>.

Una de estas seis variantes, la 1, que codifica un cambio de C por A y situada en la posición -41, fue estudiada con detalle en una publicación posterior. Se encuentra en una posición conservada y predicha como zona de unión de reguladores en *trans*<sup>272</sup>. Mediante ensayos luciferasa, se comprobó que el alelo A provoca un aumento de la expresión de *GDF-5*, capaz de neutralizar la reducción de la expresión provocada por el alelo T del rs143383. Además de SOX-9 se pueden unir potencialmente otros reguladores en el sitio -41 pb como YY1, VDR ó SEF-4. Mediante ensayo EMSA (*electrophoretic mobility shift assay*) con extractos nucleares de líneas celulares de osteosarcoma y condrosarcoma, se comprobó que YY1 (*transcriptional repressor YY1*) se une con más intensidad al alelo raro A en el sitio -41 de *GDF-5*. Se debe comentar que recientemente, se han

## *DISCUSIÓN*

---

descrito otra serie de reguladores en *trans* (Sp1, Sp3, p15 y DEAF-1) que interactúan con el rs143383 en la región 5'UTR, reduciendo la expresión del gen. Estos resultados resaltan la importancia de este tipo de regulación de *GDF-5*<sup>273</sup>.

Es posible que la ausencia de variantes raras con MAF entre 0,1-2,5 %, en *GDF-5* se deba a que es un locus que ha estado sujeto a selección positiva durante la evolución. De hecho, recientemente se ha secuenciado el gen en individuos de distintas poblaciones: europeos, asiáticos (del norte y este) y africanos, todos ellos no afectados por la enfermedad. Se encontraron evidencias de posible selección positiva en población europea y asiática que mantiene la frecuencia elevada de los SNPs asociadas con OA, como el rs143383 y el rs143384, en detrimento del resto de las variantes del locus. Los alelos comunes de estos dos SNPs conllevan una reducción de la expresión del gen lo cual se relaciona con un aumento del riesgo de OA y con una disminución de la altura en la población. Este patrón de selección apoya la idea de que la reducción del tamaño corporal durante el Pleistoceno y el Holoceno puede ser un proceso adaptativo influenciado por factores genéticos<sup>274</sup>.

En conclusión, se encontró ausencia de variantes raras en *GDF-5* que sólo porta variantes comunes o muy raras pero no de frecuencia intermedia. Por lo tanto, en población europea es poco probable que existan variantes adicionales en *GDF-5* que influyan en su expresión y que contribuyan a la susceptibilidad a OA.



---

## Estudio de asociación caso-control del microsatélite D6S1276 de BMP-5

Este trabajo identificó asociación entre el microsatélite *D6S1276*, localizado en el intrón 1 de *BMP-5* y OA de rodilla (TKR: *total knee replacement*) en mujeres. Este microsatélite había sido previamente asociado sólo con OA de cadera (THR: *total hip replacement*) y es funcional<sup>187</sup>.

La concordancia de estos resultados con los del estudio previo en cadera, no sólo se limitan a la asociación ya que se observa concordancia en otros aspectos. Por un lado, se identifica la asociación en el subgrupo de mujeres y además se percibe una distribución similar en el número de repeticiones encontradas (5-14 repeticiones tetranucleotídicas del motivo TCAT) y sus correspondientes frecuencias alélicas. Además, el alelo de riesgo es el que porta 8 repeticiones y el alelo de protección tiene 7 repeticiones. Esta concordancia con los resultados en OA de cadera es más clara con el primer grupo de muestras estudiadas por Wilkins y colaboradores. Los alelos restantes mostraron cambios en la misma dirección que la previamente observada. El segundo grupo de muestras estudiadas por Wilkins y colaboradores no mostró asociación. El análisis combinado de los dos grupos muestrales mostró asociación del alelo de 7 repeticiones con OA de cadera y efecto protector.

Nuestro estudio no es una replicación formal de este trabajo anterior, ya que la articulación estudiada es distinta. Sin embargo, proporciona mayor solidez a la implicación del microsatélite de *BMP-5* en la susceptibilidad a OA con independencia de la articulación afectada.

Otro punto que se debe comentar, ya que se realizó un análisis conjunto de tres poblaciones con diferente origen, es que los resultados de asociación obtenidos no se deben a un efecto de estratificación poblacional. En este sentido es notable que no se observaron diferencias en las frecuencias alélicas entre controles o entre los pacientes de las tres poblaciones (excepto en los pacientes de Grecia). Tampoco se observó heterogeneidad al combinar los resultados ( $I^2 = 0$ ). Además, hay razones para excluir estratificación en cada una de las tres poblaciones. Las muestras de la colección de Reino Unido están incluidas en el GWAS de arcOGEN en el que no se detectaron evidencias de estratificación poblacional de acuerdo con el valor de control genómico obtenido en el estudio<sup>141</sup>. Las muestras de Grecia tanto casos como controles son de individuos de población de Grecia Central que no han mostrado evidencias de heterogeneidad<sup>275</sup>. Lo mismo ocurre con las muestras de la colección de Santiago que tienen un ancestro homogéneo localizado en el norte de Galicia para un 94,7% de los pacientes y un 95,4% de los controles. Excluir los sujetos que tienen un ancestro español pero no gallego no alteró los resultados.

La ausencia de asociación de *BMP-5* en los GWAS de OA no contradice la asociación observada ya que las variantes multialélicas no tienen suficiente LD con los SNPs, bialélicos como para que estos sean marcadores válidos y además, las variantes de este tipo no están incluidas en los GWAS<sup>276</sup>. Algunos VNTRs, se estudiaron con anterioridad en OA como el de asporina que mostró una asociación más consistente en asiáticos que en europeos<sup>161</sup> y también otros microsatélites en genes como *CA* ó *ESR-2* aunque ninguno de ellos se replicó posteriormente<sup>275,277</sup>. Sin embargo, el *D6S1276* es muy probablemente la variante causal ya que se

ha demostrado que afecta a la regulación de la expresión de *BMP-5* y se ha analizado su repercusión en la actividad transcripcional del gen.

En efecto, se ha analizado cómo afectan a la actividad transcripcional de *BMP-5*, combinaciones de las repeticiones del microsatélite junto alelos de SNPs del gen<sup>186</sup>. Los ensayos *gen reporter* determinaron que sólo *D6S1276* modifica la expresión del gen de luciferasa bajo el promotor de *BMP-5*<sup>187</sup>. Este análisis también mostró una gran variabilidad de efecto de los alelos de *D6S1276* dependiendo de múltiples factores. Estos factores incluyen la línea celular donde se realizan los estudios y los alelos de otros tres SNPs del mismo locus. Esta complejidad es característica de *BMP-5* como se ha mostrado en estudios realizados en ratones. La regulación es modular, precisa y compleja. Incluye múltiples *enhancers* en *cis* y efectos específicos de tejido e incluso de localización<sup>64</sup>. Una regulación similar se encuentra en la regulación de *GDF-5* en el hombre ya que estudios de DAE muestran variabilidad extrema entre los tejidos de la articulación e incluso diferencias entre regiones de la misma superficie del cartílago dependiendo de la proximidad con la lesión de artrosis<sup>188</sup>.

En conclusión, se ha demostrado asociación de OA de rodilla en mujeres con el microsatélite *D6S1276* que tiene un probado efecto funcional en la expresión de *BMP-5*. La asociación encontrada demuestra la importancia de estudiar otros marcadores genéticos multialélicos además de los SNPs. Aunque pueden existir SNPs en LD elevado con un microsatélite también puede ocurrir, como en el caso del *D6S1276*, que no exista ningún SNP con elevada correlación con él<sup>187</sup>. Por ello el estudio de otros microsatélites en OA es especialmente importante ya que quizás pueda

## *DISCUSIÓN*

---

informar de parte de la heredabilidad no explicada por los SNPs recogidos en los GWAS.



### **Análisis de genes candidatos en un metaanálisis de 9 GWAS.**

En este metaanálisis de genes candidatos de OA de 9 GWAS, se analizaron 199 genes en 158 regiones genómicas detectándose asociaciones en dos genes (*COL11A1* y *VEGF*). Éste era uno de los objetivos importantes de este estudio, descubrir nuevos SNPs asociados en loci de susceptibilidad no especialmente relevantes hasta el momento. Otro objetivo era evaluar la validez de la aproximación de genes candidatos en el estudio de la genética de la OA. En relación a este punto, ninguno de los 197 genes restantes mostraron asociación con OA.

Los estudios genéticos han avanzado rápidamente gracias a la tecnología de los GWAS que permitió aumentar el tamaño muestral, incrementar la cobertura de los SNPs del genoma y mejorar los controles de calidad. Estas características positivas de los GWAS se potencian al combinarse en un metaanálisis disminuyendo la probabilidad de encontrar falsos negativos en el estudio<sup>278</sup>. Este metaanálisis tiene un gran poder estadístico para analizar la mayor parte de los genes candidatos, al menos para  $OR > 1.15$  y frecuencias alélicas  $> 0.2$ . Además, cubre gran parte del genoma al haberse realizado una imputación de los SNPs no directamente genotipados. El análisis centrado en genes candidatos aumenta las probabilidades de descubrir nuevas asociaciones debido a dos factores. Por un lado existe un aumento de probabilidad previa de detectar asociaciones y por otro se establece un umbral de asociación menos exigente<sup>279,280</sup>. Una serie de estudios de otras enfermedades se han centrado en el primer aspecto considerando los genes candidatos de previos estudios de asociación (al igual que en este trabajo) o seleccionando de la bibliografía genes de repercusión en la enfermedad.

## *DISCUSIÓN*

---

El segundo aspecto, un nivel de asociación menos exigente, se relaciona con un problema que afecta a todas las enfermedades complejas. Los tamaños de efecto de la mayoría de los factores genéticos tienen una  $OR < 1.2$  por lo que se necesitan tamaños muestrales muy amplios para detectar asociaciones a nivel de genoma completo. En este contexto, realizar un análisis focalizado en un conjunto de genes permite utilizar un umbral de significación menos exigente para una validación posterior.

Otro punto que se debe comentar es que pequeñas variaciones de tamaño entre las colecciones muestrales provocan grandes diferencias en el poder estadístico. Esto se debe a que la mayor parte de los factores genéticos tienen un tamaño de efecto pequeño y el poder estadístico disminuye de forma muy marcada cuando la OR se aproxima a 1.0. Es muy probable que esta variación de sensibilidad de las ORs entre 1.0-1.15 explique por qué algunos factores genéticos confirmados no muestran asociación en estudios de gran tamaño. Esto se debería a que no son lo suficientemente grandes para compensar las pequeñas fluctuaciones del tamaño de efecto (OR). Por lo tanto ni siquiera con este metaanálisis es posible excluir completamente asociación de alguno de los genes candidatos analizados. Si la OR o la MAF son pequeños, también hay una caída drástica de la potencia estadística con  $MAF > 10\%$  (Figura Suplementaria 1).

Un aspecto que requiere comentario es el uso de la estimación del número de test independientes y de la corrección de Bonferroni. La utilización de este tipo de estrategias viene dada por la necesidad de establecer un nivel de significación restrictivo para evitar los falsos positivos asociados al gran número de tests. La existencia de LD entre los SNPs del estudio hace que no sea apropiado utilizar la corrección de Bonferroni por el número

de tests independientes que esos SNPs representan. El mejor método es el proporcionado por el algoritmo de simple M de Gao que proporciona una estimación segura y eficiente basada en la matriz de correlaciones de  $r^2$  de los SNPs. Una vez obtenido este número se utiliza la corrección de Bonferroni para establecer el umbral de significación a partir del cuál un SNP se considera asociado en el estudio ( $0,05/\text{número de tests independientes}$ ).

*COL11A1*, uno de los genes asociados en el trabajo, es un gen candidato relevante en OA ya que pertenece a la familia de los colágenos que contribuyen a la estabilización y organización estructural de la ECM. Concretamente, *COL11A1* es un componente minoritario de la ECM del cartílago, responsable de la interacción entre los agregados de proteoglicanos y otros colágenos. Mutaciones en este gen son responsables de diferentes condrodisplasias en ratones que desarrollaban OA más tempranamente<sup>233</sup>. En humanos, mutaciones en *COL11A1* son las causantes de los síndromes de Stickler y de Marshall<sup>231,234,281</sup>. Centrándonos en los estudios genéticos, en la primera fase de arcOGEN previa a la realización de este metaanálisis se encontró asociación del rs2615977 en el intrón 31 del *COL11A1*, con OA de cadera<sup>143</sup>. Además, se mostró asociación de SNPs de *COL11A1* con un fenotipo relacionado, hernia de disco lumbar, en población japonesa<sup>231</sup>.

Los resultados del metaanálisis muestran asociación de dos SNPs que marcan distintos efectos genéticos en el gen que cubre unas 154 kb. Por un lado, estaría el SNP más asociado, el rs4907986 en el intrón 1 y todos los SNPs relacionados con él por  $r^2$ . Otro efecto independiente debido a la elevada recombinación que los separa, vendría marcado por el rs1241164

## DISCUSIÓN

---

en el intrón 55 y sus SNPs correlacionados. Además, se podría señalar una tercera señal independiente (rs12139120) aunque no alcanza el umbral de significación requerido. De esta manera se puede decir que a lo largo del gen se localizan diferentes señales que contribuyen con efectos genéticos independientes. Así, se puede decir que con este metaanálisis de 9 GWAS se aclara la contribución de *COL11A1* en la susceptibilidad a OA de cadera.

En este metaanálisis se encontró asociación de *VEGF* con OA de cadera solamente en hombres. Este gen codifica para un factor angiogénico implicado en la osificación endocondral y en la formación del cartílago articular<sup>282</sup>. Es uno de los marcadores que se sobreexpresa en el cartílago hipertrófico. Contribuye a cambios en modelos animales de OA al estimular la proliferación de condrocitos, la apoptosis y la producción de mediadores catabólicos<sup>283</sup>. No existen estudios previos que muestren asociación de SNPs en este gen con OA. Sin embargo, un SNP en HIF-2a que regula la activación transcripcional de *VEGF*, mostró asociación con OA de rodilla en población japonesa<sup>284</sup>. Además, el rs833058 también se encontró asociado con asma infantil<sup>285</sup>. *VEGF* es un gen de pequeño tamaño (76 kb) y sólo otros cuatro SNPs se relacionan con el rs833058 aunque con una  $r^2$  comprendida entre 0,2-0,5. Esto indica la existencia de una sola señal de asociación en el gen.

En el resto de los genes candidatos, no se encontró asociación a pesar del gran número de muestras utilizadas y el umbral de asociación menos exigente. Esto no excluye que existan SNPs con efecto débil o que muestren heterogeneidad entre las colecciones de muestras incluidas. Otra limitación inherente a cualquier estudio basado en chips de genotipado es



la no inclusión de variabilidad genética de VNTRs, CNVs y variantes raras. En este estudio no se cubren las variantes con una frecuencia < 5% ni polimorfismos multialélicos como los que se encuentran en los genes de *ASPN* y *BMP-5*<sup>161,187</sup>. Otras asociaciones que son específicas de algún subfenotipo concreto de OA no se incluyen en el análisis, como *MATN-3* que se encontró asociada con OA de mano del primer metacarpo<sup>174</sup>. A pesar de estas limitaciones, no podemos negar la falta de reproducibilidad de asociación en 197 genes candidatos de OA.

Así pues, nuestros hallazgos indican que los estudios de genes candidatos realizados hasta el momento siguiendo los métodos tradicionales no han sido muy fructíferos en la genética de la OA. Este es un hallazgo común en las enfermedades complejas y parece debido a los bajos estándares con los que muchos estudios de genes candidatos se han realizado. Se han propuesto una serie de recomendaciones para aumentar la calidad y el valor de este tipo de estudios<sup>286,287</sup>. Así, se propone utilizar colecciones muestrales más amplias, realizar una replicación en muestras adicionales y uniformizar los fenotipos incluidos en el estudio. Los estudios de genes candidatos siguiendo estas recomendaciones se consideran todavía útiles para validar o refinar loci encontrados en los GWAS.

En resumen, este metaanálisis de genes candidatos en 9 GWAS de OA reveló la asociación de *COL11A1* y *VEGF* con OA de cadera y no encontró asociación en los restantes 197 genes candidatos estudiados.

## *DISCUSIÓN*

---

En resumen, en esta tesis se han abordado distintos aspectos de la variabilidad genética de la OA. Por un lado se analizó la relación entre edad de inicio y carga genética. Análisis que podía haber señalado una vía para modificar el diseño de los estudios genéticos de OA. Los resultados indicaron que no se podía mejorar seleccionando pacientes con una edad temprana de inicio de los síntomas ya que esto no proporcionaría ninguna ventaja en el descubrimiento de nuevos loci asociados con OA. También se ha explorado el campo no estudiado hasta el momento de secuenciación en busca de variantes raras. Se comprobó que *GDF-5*, uno de los loci más sólidamente asociados en OA, no porta variantes de susceptibilidad de baja frecuencia. Además, se ha estudiado un ejemplo de variabilidad no recogida en los GWAS, concretamente el microsatélite *D6S1276* en el gen *BMP-5*. Al demostrarse su asociación con OA de rodilla se ha puesto de manifiesto la necesidad de explorar los polimorfismos multialélicos como contribuyentes a la genética de la OA. Por último, se extrajo más información de los GWAS publicados, descubriendo nuevas asociaciones en dos genes, *COL11A1* y *VEGF* que deben ser confirmadas en más estudios pero que son las más prometedoras entre todos los genes candidatos considerados.

Estas cuatro aproximaciones diferentes han servido para descartar la relación entre la edad de inicio de los síntomas y la carga genética de OA, así como la ausencia de variantes raras en *GDF-5*. Además, también se han obtenido resultados positivos como la asociación de *D6S1276* con OA de rodilla y el descubrimiento de asociaciones en SNPs de *COL11A1* y *VEGF*.



## **CONCLUSIONES**

---



## CONCLUSIONES

Las conclusiones de los trabajos que forman parte de esta tesis son las siguientes:

1. No existe relación entre el número de alelos de riesgo de los polimorfismos estudiados y la edad de inicio de los síntomas de OA de rodilla.
2. No hay variantes raras en *GDF-5* asociadas con OA de rodilla.
3. El microsatélite funcional *D6S1276* de *BMP-5* se asocia con OA de rodilla en mujeres.
4. Sólo dos de los 199 genes candidatos analizados, *COL11A1* y *VEGF*, están asociados con OA en un metaanálisis de 9 GWAS.
5. Por lo tanto, no se ha encontrado apoyo para la idea de que seleccionando fenotipos extremos en la edad de inicio de la OA se pueda mejorar la eficiencia de los estudios genéticos; no se ha encontrado ninguna variante rara asociada a OA en el gen mejor estudiado, *GDF-5*; se confirma la posibilidad de que polimorfismos multialélicos contribuyan a la susceptibilidad a OA. Se muestra el bajo rendimiento de los estudios de genes candidatos aunque se identifican dos genes para priorizar en futuros estudios.

## *CONCLUSIONES*

---

Estos resultados indican que el futuro de los estudios genéticos de OA estará en mayores colecciones de muestras buscando tanto variantes comunes como raras y posiblemente multialélicas.





## **BIBLIOGRAFÍA**

---





## BIBLIOGRAFÍA

1. Scanzello, C. R. & Goldring, S. R. The role of synovitis in osteoarthritis pathogenesis. *Bone* (2012).
2. Corti, M. C. & Rigon, C. Epidemiology of osteoarthritis: prevalence, risk factors and functional impact. *Aging Clinical and Experimental Research* **15**, 359–63 (2003).
3. Altman, R. D. The classification of osteoarthritis. *The Journal of Rheumatology. Supplement* **43**, 42–3 (1995).
4. Arden, N. & Nevitt, M. C. Osteoarthritis: epidemiology. *Best practice & research. Clinical Rheumatology* **20**, 3–25 (2006).
5. Pereira, D. *et al.* The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **19**, 1270–85 (2011).
6. Felson, D. T. *et al.* The prevalence of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis and Rheumatism* **30**, 914–8 (1987).
7. Felson, D. T. *et al.* The incidence and natural history of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. *Arthritis and Rheumatism* **38**, 1500–5 (1995).
8. Oliveria, S. A., Felson, D. T., Reed, J. I., Cirillo, P. A. & Walker, A. M. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. *Arthritis and Rheumatism* **38**, 1134–41 (1995).
9. Leong, D. J. & Sun, H. B. Events in articular chondrocytes with aging. *Current Osteoporosis Reports* **9**, 196–201 (2011).
10. Maleki-Fischbach, M. & Jordan, J. M. New developments in osteoarthritis. Sex differences in magnetic resonance imaging-based biomarkers and in those of joint metabolism. *Arthritis Research & Therapy* **12**, 212 (2010).
11. Zhang, Y. *et al.* Estrogen replacement therapy and worsening of radiographic knee osteoarthritis: the Framingham Study. *Arthritis and Rheumatism* **41**, 1867–73 (1998).
12. Sandmark, H., Hogstedt, C., Lewold, S. & Vingård, E. Osteoarthrosis of the knee in men and women in association with overweight, smoking, and hormone therapy. *Annals of the Rheumatic Diseases* **58**, 151–5 (1999).

13. Richette, P., Corvol, M. & Bardin, T. Estrogens, cartilage, and osteoarthritis. *Joint, bone, spine: revue du rhumatisme* **70**, 257–62 (2003).
14. Blagojevic, M., Jinks, C., Jeffery, A. & Jordan, K. P. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **18**, 24–33 (2010).
15. Carman, W. J., Sowers, M., Hawthorne, V. M. & Weissfeld, L. A. Obesity as a risk factor for osteoarthritis of the hand and wrist: a prospective study. *American Journal of Epidemiology* **139**, 119–29 (1994).
16. Rai, M. F. & Sandell, L. J. Inflammatory mediators: tracing links between obesity and osteoarthritis. *Critical reviews in eukaryotic gene expression* **21**, 131–42 (2011).
17. Englund, M. *et al.* Effect of meniscal damage on the development of frequent knee pain, aching, or stiffness. *Arthritis and Rheumatism* **56**, 4048–54 (2007).
18. Englund, M., Guermazi, A. & Lohmander, L. S. The meniscus in knee osteoarthritis. *Rheumatic Diseases Clinics of North America* **35**, 579–90 (2009).
19. Lohmander, L. S., 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 and Rheumatism* **50**, 3145–52 (2004).
20. McDaniel, G., Renner, J. B., Sloane, R. & Kraus, V. B. Association of knee and ankle osteoarthritis with physical performance. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **19**, 634–8 (2011).
21. Dunlop, D. D., Song, J., Semanik, P. A., Sharma, L. & Chang, R. W. Physical activity levels and functional performance in the osteoarthritis initiative: a graded relationship. *Arthritis and Rheumatism* **63**, 127–36 (2011).
22. Spector, T. D. & MacGregor, A. J. Risk factors for osteoarthritis: genetics. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **12 Suppl A**, S39–44 (2004).
23. Salter, D. M. ( II ) Cartilage. *Rheumatic Disease Clinics Of North America* **31**, 251–257 (1998).
24. Sophia Fox, A. J., Bedi, A. & Rodeo, S. A. The Basic Science of Articular Cartilage: Structure, Composition, and Function. *Sports Health: A Multidisciplinary Approach* **1**, 461–468 (2009).

25. Youn, I., Choi, J. B., Cao, L., Setton, L. A. & Guilak, F. Zonal variations in the three-dimensional morphology of the chondron measured in situ using confocal microscopy. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **14**, 889–97 (2006).
26. Schumacher, B. L., Hughes, C. E., Kuettner, K. E., Caterson, B. & Aydelotte, M. B. Immunodetection and partial cDNA sequence of the proteoglycan, superficial zone protein, synthesized by cells lining synovial joints. *Journal of Orthopaedic Research: official publication of the Orthopaedic Research Society* **17**, 110–20 (1999).
27. Desrochers, J., Amrein, M. W. & Matyas, J. R. Viscoelasticity of the articular cartilage surface in early osteoarthritis. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **20**, 413–21 (2012).
28. Korver, G. H., Van de Stadt, R. J., Van Kampen, G. P. & Van der Korst, J. K. Composition of proteoglycans synthesized in different layers of cultured anatomically intact articular cartilage. *Matrix (Stuttgart, Germany)* **10**, 394–401 (1990).
29. Aydelotte, M. B. & Kuettner, K. E. Differences between subpopulations of cultured bovine articular chondrocytes. I. Morphology and cartilage matrix production. *Connective tissue research* **18**, 205–22 (1988).
30. Lyons, T. J., Stoddart, R. W., McClure, S. F. & McClure, J. The tidemark of the chondro-osseous junction of the normal human knee joint. *Journal of Molecular Histology* **36**, 207–15 (2005).
31. Quinn, T. M., Hunziker, E. B. & Häuselmann, H.-J. Variation of cell and matrix morphologies in articular cartilage among locations in the adult human knee. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **13**, 672–8 (2005).
32. Van der Kraan, P. M. & Van den Berg, W. B. Chondrocyte hypertrophy and osteoarthritis: role in initiation and progression of cartilage degeneration? *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **20**, 223–32 (2012).
33. Archer, C. W. & Francis-West, P. The chondrocyte. *The International Journal of Biochemistry & Cell Biology* **35**, 401–404 (2003).
34. Kawaguchi, H. Regulation of osteoarthritis development by Wnt-beta-catenin signaling through the endochondral ossification process. *Journal of bone and mineral research: the official journal*

- of the American Society for Bone and Mineral Research **24**, 8–11 (2009).
35. Quintana, L., Zur Nieden, N. I. & Semino, C. E. Morphogenetic and regulatory mechanisms during developmental chondrogenesis: new paradigms for cartilage tissue engineering. *Tissue engineering. Part B, Reviews* **15**, 29–41 (2009).
  36. Ghone, N. V & Grayson, W. L. Recapitulation of mesenchymal condensation enhances in vitro chondrogenesis of human mesenchymal stem cells. *Journal of Cellular Physiology* (2012).doi:10.1002/jcp.24078
  37. Jorgensen, C. Could inflammatory arthritis be triggered by progenitor cells in the joints? *Annals of the Rheumatic Diseases* **61**, 6–9 (2002).
  38. Hartmann, C. & Tabin, C. J. Wnt-14 plays a pivotal role in inducing synovial joint formation in the developing appendicular skeleton. *Cell* **104**, 341–51 (2001).
  39. Tsumaki, N. *et al.* Bone morphogenetic protein signals are required for cartilage formation and differently regulate joint development during skeletogenesis. *Journal of bone and mineral research: the official journal of the American Society for Bone and Mineral Research* **17**, 898–906 (2002).
  40. Khan, I. M. *et al.* The development of synovial joints. *Current topics in developmental biology* **79**, 1–36 (2007).
  41. Pitsillides, A. A., Skerry, T. M. & Edwards, J. C. Joint immobilization reduces synovial fluid hyaluronan concentration and is accompanied by changes in the synovial intimal cell populations. *Rheumatology (Oxford, England)* **38**, 1108–12 (1999).
  42. Pitsillides, A. A. Early effects of embryonic movement: “a shot out of the dark”. *Journal of Anatomy* **208**, 417–31 (2006).
  43. Reddi, A. H. Interplay between bone morphogenetic proteins and cognate binding proteins in bone and cartilage development: noggin, chordin and DAN. *Arthritis Research* **3**, 1–5 (2001).
  44. Borzì, R. M. *et al.* Matrix metalloproteinase 13 loss associated with impaired extracellular matrix remodeling disrupts chondrocyte differentiation by concerted effects on multiple regulatory factors. *Arthritis and Rheumatism* **62**, 2370–81 (2010).
  45. Hulmes, D. J. The collagen superfamily--diverse structures and assemblies. *Essays in Biochemistry* **27**, 49–67 (1992).
  46. Eyre, D. Collagen of articular cartilage. *Arthritis Research* **4**, 30–5 (2002).

47. Wu, J.-J., Weis, M. A., Kim, L. S. & Eyre, D. R. Type III collagen, a fibril network modifier in articular cartilage. *The Journal of Biological Chemistry* **285**, 18537–44 (2010).
48. Cremer, M. A., Rosloniec, E. F. & Kang, A. H. The cartilage collagens: a review of their structure, organization, and role in the pathogenesis of experimental arthritis in animals and in human rheumatic disease. *Journal of Molecular Medicine* **76**, 275–88 (1998).
49. Roughley, P. J. & Lee, E. R. Cartilage proteoglycans: structure and potential functions. *Microscopy Research and Technique* **28**, 385–97 (1994).
50. Knudson, C. B. & Knudson, W. Cartilage proteoglycans. *Seminars in Cell & Developmental Biology* **12**, 69–78 (2001).
51. Posey, K. L. & Hecht, J. T. The role of cartilage oligomeric matrix protein (COMP) in skeletal disease. *Current Drug Targets* **9**, 869–77 (2008).
52. Posey, K. L., Hayes, E., Haynes, R. & Hecht, J. T. Role of TSP-5/COMP in pseudoachondroplasia. *The International Journal of Biochemistry & Cell Biology* **36**, 1005–12 (2004).
53. Hoch, J. M., Mattacola, C. G., Medina McKeon, J. M., Howard, J. S. & Lattermann, C. Serum cartilage oligomeric matrix protein (sCOMP) is elevated in patients with knee osteoarthritis: a systematic review and meta-analysis. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **19**, 1396–404 (2011).
54. Finsson, K. W., Chi, Y., Bou-Gharios, G., Leask, A. & Philip, A. TGF- $\beta$  signaling in cartilage homeostasis and osteoarthritis. *Frontiers in Bioscience* **4**, 251–68 (2012).
55. Verrecchia, F. & Mauviel, A. Transforming growth factor-beta signaling through the Smad pathway: role in extracellular matrix gene expression and regulation. *The Journal of Investigative Dermatology* **118**, 211–5 (2002).
56. Massagué, J. & Chen, Y. G. Controlling TGF-beta signaling. *Genes & Development* **14**, 627–44 (2000).
57. Blaney Davidson, E. N., Vitters, E. L., Van der Kraan, P. M. & Van den Berg, W. B. Expression of transforming growth factor-beta (TGFbeta) and the TGFbeta signalling molecule SMAD-2P in spontaneous and instability-induced osteoarthritis: role in cartilage degradation, chondrogenesis and osteophyte formation. *Annals of the Rheumatic Diseases* **65**, 1414–21 (2006).

58. Blaney Davidson, E. N. *et al.* Resemblance of osteophytes in experimental osteoarthritis to transforming growth factor beta-induced osteophytes: limited role of bone morphogenetic protein in early osteoarthritic osteophyte formation. *Arthritis and Rheumatism* **56**, 4065–73 (2007).
59. Pombo-Suarez, M., Castaño-Oreja, M. T., Calaza, M., Gomez-Reino, J. & Gonzalez, A. Differential upregulation of the three transforming growth factor beta isoforms in human osteoarthritic cartilage. *Annals of the Rheumatic Diseases* **68**, 568–71 (2009).
60. Bragdon, B. *et al.* Bone morphogenetic proteins: a critical review. *Cellular signalling* **23**, 609–20 (2011).
61. Dorman, L. J., Tucci, M. & Benghuzzi, H. In vitro effects of BMP-2, BMP-7, and BMP-13 on proliferation and differentiation of mouse mesenchymal stem cells. *Biomedical Sciences Instrumentation* **48**, 81–7 (2012).
62. Stöve, J., Schneider-Wald, B., Scharf, H.-P. & Schwarz, M. L. Bone morphogenetic protein 7 (BMP-7) stimulates proteoglycan synthesis in human osteoarthritic chondrocytes in vitro. *Biomedicine & Pharmacotherapy* **60**, 639–43 (2006).
63. Mailhot, G. *et al.* BMP-5 expression increases during chondrocyte differentiation in vivo and in vitro and promotes proliferation and cartilage matrix synthesis in primary chondrocyte cultures. *Journal of Cellular Physiology* **214**, 56–64 (2008).
64. Guenther, C., Pantalena-Filho, L. & Kingsley, D. M. Shaping skeletal growth by modular regulatory elements in the BMP-5 gene. *PLoS Genetics* **4**, (2008).
65. Chen, D., Zhao, M. & Mundy, G. R. Bone morphogenetic proteins. *Growth Factors* **22**, 233–41 (2004).
66. Nakase, T. *et al.* Localization of bone morphogenetic protein-2 in human osteoarthritic cartilage and osteophyte. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **11**, 278–84 (2003).
67. McQuillan, D. J. *et al.* Stimulation of proteoglycan biosynthesis by serum and insulin-like growth factor-I in cultured bovine articular cartilage. *The Biochemical Journal* **240**, 423–30 (1986).
68. Middleton, J. F. & Tyler, J. A. Upregulation of insulin-like growth factor I gene expression in the lesions of osteoarthritic human articular cartilage. *Annals of the Rheumatic Diseases* **51**, 440–7 (1992).
69. Renard, E. *et al.* Sox9/Sox6 and Sp1 are involved in the insulin-like growth factor-I-mediated upregulation of human type II



- collagen gene expression in articular chondrocytes. *Journal of Molecular Medicine* **90**, 649–66 (2012).
70. Martel-Pelletier, J., Di Battista, J. A., Lajeunesse, D. & Pelletier, J. P. IGF/IGFBP axis in cartilage and bone in osteoarthritis pathogenesis. *Inflammation Research* **47**, 90–100 (1998).
  71. Olney, R. C. *et al.* Chondrocytes from osteoarthritic cartilage have increased expression of insulin-like growth factor I (IGF-I) and IGF-binding protein-3 (IGFBP-3) and -5, but not IGF-II or IGFBP-4. *The Journal of Clinical Endocrinology and Metabolism* **81**, 1096–103 (1996).
  72. Itoh, N. & Ornitz, D. M. Fibroblast growth factors: from molecular evolution to roles in development, metabolism and disease. *Journal of Biochemistry* **149**, 121–30 (2011).
  73. Ornitz, D. M. & Itoh, N. Fibroblast growth factors. *Genome Biology* **2**, review S3005 (2001).
  74. Rosselot, G., Vasilatos-Younken, R. & Leach, R. M. Effect of growth hormone, insulin-like growth factor I, basic fibroblast growth factor, and transforming growth factor beta on cell proliferation and proteoglycan synthesis by avian postembryonic growth plate chondrocytes. *Journal of Bone and Mineral Research: the official journal of the American Society for Bone and Mineral Research* **9**, 431–9 (1994).
  75. Ellman, M. B., An, H. S., Muddasani, P. & Im, H.-J. Biological impact of the fibroblast growth factor family on articular cartilage and intervertebral disc homeostasis. *Gene* **420**, 82–9 (2008).
  76. Vincent, T. Explaining the FGF paradox in osteoarthritis: Lessons from conditional knockout mice. *Arthritis and Rheumatism* **64**:3835-8 (2012).
  77. Weng, T. *et al.* Genetic inhibition of FGFR1 in cartilage attenuates articular cartilage degeneration in adult mice. *Arthritis and Rheumatism*, **64**:3982-92 (2012).
  78. Martel-Pelletier, J. & Pelletier, J. P. Importance of interleukin-1 receptors in osteoarthritis. *Revue du Rhumatisme (Ed. française: 1993)* **61**, 109S–113S (1994).
  79. Weber, A., Wasiliew, P. & Kracht, M. Interleukin-1 (IL-1) pathway. *Science Signaling* **3**, (2010).
  80. Goldring, M. B. & Berenbaum, F. The regulation of chondrocyte function by proinflammatory mediators: prostaglandins and nitric oxide. *Clinical Orthopaedics and Related Research* S37–46 (2004).

81. Baragi, V. M. *et al.* Transplantation of transduced chondrocytes protects articular cartilage from interleukin 1-induced extracellular matrix degradation. *The Journal of Clinical Investigation* **96**, 2454–60 (1995).
82. Kapoor, M., Martel-Pelletier, J., Lajeunesse, D., Pelletier, J.-P. & Fahmi, H. Role of proinflammatory cytokines in the pathophysiology of osteoarthritis. *Nature Reviews. Rheumatology* **7**, 33–42 (2011).
83. Amin, A. R. Regulation of tumor necrosis factor-alpha and tumor necrosis factor converting enzyme in human osteoarthritis. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **7**, 392–4 (1999).
84. Palmer, R. M., Ashton, D. S. & Moncada, S. Vascular endothelial cells synthesize nitric oxide from L-arginine. *Nature* **333**, 664–6 (1988).
85. Stadler, J. *et al.* Articular chondrocytes synthesize nitric oxide in response to cytokines and lipopolysaccharide. *Journal of Immunology* **147**, 3915–20 (1991).
86. De Isla, N. G., Mainard, D., Muller, S. & Stoltz, J. F. In vitro effects of diacerein on NO production by chondrocytes in response to proinflammatory mediators. *Bio-medical materials and engineering* **18**, S99–104 (2008).
87. Studer, R., Jaffurs, D., Stefanovic-Racic, M., Robbins, P. D. & Evans, C. H. Nitric oxide in osteoarthritis. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **7**, 377–9 (1999).
88. Scher, J. U., Pillinger, M. H. & Abramson, S. B. Nitric oxide synthases and osteoarthritis. *Current Rheumatology Reports* **9**, 9–15 (2007).
89. Scher, J. U., Pillinger, M. H. & Abramson, S. B. Nitric oxide synthases and osteoarthritis. *Current Rheumatology Reports* **9**, 9–15 (2007).
90. Amin, A. R. *et al.* The expression and regulation of nitric oxide synthase in human osteoarthritis-affected chondrocytes: evidence for up-regulated neuronal nitric oxide synthase. *The Journal of Experimental Medicine* **182**, 2097–102 (1995).
91. Weinberg, J. B., Fermor, B. & Guilak, F. Nitric oxide synthase and cyclooxygenase interactions in cartilage and meniscus: relationships to joint physiology, arthritis, and tissue repair. *Subcellular Biochemistry* **42**, 31–62 (2007).



92. Abramson, S. B. Nitric oxide in inflammation and pain associated with osteoarthritis. *Arthritis Research & Therapy* **10 Suppl 2**, S2 (2008).
93. Pacher, P., Beckman, J. S. & Liaudet, L. Nitric oxide and peroxynitrite in health and disease. *Physiological Reviews* **87**, 315–424 (2007).
94. Del Carlo, M. & Loeser, R. F. Nitric oxide-mediated chondrocyte cell death requires the generation of additional reactive oxygen species. *Arthritis and Rheumatism* **46**, 394–403 (2002).
95. Ishiguro, N. *et al.* Relationship of matrix metalloproteinases and their inhibitors to cartilage proteoglycan and collagen turnover: analyses of synovial fluid from patients with osteoarthritis. *Arthritis and Rheumatism* **42**, 129–36 (1999).
96. Nagase, H. Matrix Metalloproteinases. *Journal of Biological Chemistry* **274**, 21491–21494 (1999).
97. Visse, R. & Nagase, H. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circulation Research* **92**, 827–39 (2003).
98. Steffensen, B., Wallon, U. & Overall, C. Extracellular matrix binding properties of recombinant fibronectin type II-like modules of human 72-kDa gelatinase/type IV collagenase. High affinity binding to native type I collagen but not native type IV collagen. *J. Biol. Chem.* **270**, 11555–11566 (1995).
99. Little, C. B. *et al.* Matrix metalloproteinase 13-deficient mice are resistant to osteoarthritic cartilage erosion but not chondrocyte hypertrophy or osteophyte development. *Arthritis and Rheumatism* **60**, 3723–33 (2009).
100. Tam, E. M., Moore, T. R., Butler, G. S. & Overall, C. M. Characterization of the distinct collagen binding, helicase and cleavage mechanisms of matrix metalloproteinase 2 and 14 (gelatinase A and MT1-MMP): the differential roles of the MMP hemopexin c domains and the MMP-2 fibronectin type II modules in collagen. *The Journal of Biological Chemistry* **279**, 43336–44 (2004).
101. Little, C. B. *et al.* Blocking aggrecanase cleavage in the aggrecan interglobular domain abrogates cartilage erosion and promotes cartilage repair. *The Journal of Clinical Investigation* **117**, 1627–36 (2007).
102. Struglics, A. *et al.* Human osteoarthritis synovial fluid and joint cartilage contain both aggrecanase- and matrix metalloproteinase-generated aggrecan fragments. *Osteoarthritis and Cartilage /*

- OARSI, Osteoarthritis Research Society International* **14**, 101–13 (2006).
103. Yan, C. & Boyd, D. D. Regulation of matrix metalloproteinase gene expression. *Journal of Cellular Physiology* **211**, 19–26 (2007).
104. Benbow, U. & Brinckerhoff, C. E. The AP-1 site and MMP gene regulation: what is all the fuss about? *Matrix biology: journal of the International Society for Matrix Biology* **15**, 519–26 (1997).
105. Crawford, H. C. & Matrisian, L. M. Mechanisms controlling the transcription of matrix metalloproteinase genes in normal and neoplastic cells. *Enzyme & Protein* **49**, 20–37 (1996).
106. Liang, Z.-J. *et al.* MiRNA-140 is a negative feedback regulator of MMP-13 in IL-1 $\beta$ -stimulated human articular chondrocyte C28/I2 cells. *Inflammation Research* **61**, 503–9 (2012).
107. Apte, S. S. A disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motifs: the ADAMTS family. *The International Journal of Biochemistry & Cell Biology* **36**, 981–5 (2004).
108. Tang, B. L. ADAMTS: a novel family of extracellular matrix proteases. *The International Journal of Biochemistry & Cell Biology* **33**, 33–44 (2001).
109. Tortorella, M. D. *et al.* Purification and cloning of aggrecanase-1: a member of the ADAMTS family of proteins. *Science* **284**, 1664–6 (1999).
110. Song, R.-H. *et al.* Aggrecan degradation in human articular cartilage explants is mediated by both ADAMTS-4 and ADAMTS-5. *Arthritis and Rheumatism* **56**, 575–85 (2007).
111. Naito, S. *et al.* Expression of ADAMTS4 (aggrecanase-1) in human osteoarthritic cartilage. *Pathology International* **57**, 703–11 (2007).
112. Glasson, S. S. *et al.* Characterization of and osteoarthritis susceptibility in ADAMTS-4-knockout mice. *Arthritis and Rheumatism* **50**, 2547–58 (2004).
113. Rogerson, F. M. *et al.* Evidence of a novel aggrecan-degrading activity in cartilage: Studies of mice deficient in both ADAMTS-4 and ADAMTS-5. *Arthritis and Rheumatism* **58**, 1664–73 (2008).
114. Majumdar, M. K. *et al.* Double-knockout of ADAMTS-4 and ADAMTS-5 in mice results in physiologically normal animals and prevents the progression of osteoarthritis. *Arthritis and Rheumatism* **56**, 3670–4 (2007).

115. Bondeson, J., Wainwright, S., Hughes, C. & Caterson, B. The regulation of the ADAMTS4 and ADAMTS5 aggrecanases in osteoarthritis: a review. *Clinical and Experimental Rheumatology* **26**, 139–45
116. Cheung, K. S. C., Hashimoto, K., Yamada, N. & Roach, H. I. Expression of ADAMTS-4 by chondrocytes in the surface zone of human osteoarthritic cartilage is regulated by epigenetic DNA demethylation. *Rheumatology International* **29**, 525–34 (2009).
117. Kozaci, L. D., Buttle, D. J. & Hollander, A. P. Degradation of type II collagen, but not proteoglycan, correlates with matrix metalloproteinase activity in cartilage explant cultures. *Arthritis and Rheumatism* **40**, 164–74 (1997).
118. Pratta, M. A., Scherle, P. A., Yang, G., Liu, R.-Q. & Newton, R. C. Induction of aggrecanase 1 (ADAM-TS4) by interleukin-1 occurs through activation of constitutively produced protein. *Arthritis and Rheumatism* **48**, 119–33 (2003).
119. Beekman, B., Verzijl, N., De Roos, J. A. & TeKoppele, J. M. Matrix degradation by chondrocytes cultured in alginate: IL-1 beta induces proteoglycan degradation and proMMP synthesis but does not result in collagen degradation. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **6**, 330–40 (1998).
120. Bau, B. *et al.* Relative messenger RNA expression profiling of collagenases and aggrecanases in human articular chondrocytes in vivo and in vitro. *Arthritis and Rheumatism* **46**, 2648–57 (2002).
121. Van Goor, H., Melenhorst, W. B. W. H., Turner, A. J. & Holgate, S. T. Adamalysins in biology and disease. *The Journal of Pathology* **219**, 277–86 (2009).
122. Kerna, I. *et al.* Association of ADAM12-S protein with radiographic features of knee osteoarthritis and bone and cartilage markers. *Rheumatology International* **32**, 519–23 (2012).
123. Chubinskaya, S., Cs-Szabo, G. & Kuettner, K. E. ADAM-10 message is expressed in human articular cartilage. *The Journal of Histochemistry and Cytochemistry: official journal of the Histochemistry Society* **46**, 723–9 (1998).
124. Böhm, B. B. *et al.* Up-regulation of MDC15 (metargidin) messenger RNA in human osteoarthritic cartilage. *Arthritis and Rheumatism* **42**, 1946–50 (1999).
125. Kashiwagi, M., Tortorella, M., Nagase, H. & Brew, K. TIMP-3 is a potent inhibitor of aggrecanase 1 (ADAM-TS4) and aggrecanase

- 2 (ADAM-TS5). *The Journal of Biological Chemistry* **276**, 12501–4 (2001).
126. Flemmig, M. & Melzig, M. F. Serine-proteases as plasminogen activators in terms of fibrinolysis. *The Journal of Pharmacy and Pharmacology* **64**, 1025–39 (2012).
127. Belcher, C., Fawthrop, F., Bunning, R. & Doherty, M. Plasminogen activators and their inhibitors in synovial fluids from normal, osteoarthritis, and rheumatoid arthritis knees. *Annals of the Rheumatic Diseases* **55**, 230–6 (1996).
128. Boileau, C. *et al.* Activation of proteinase-activated receptor 2 in human osteoarthritic cartilage upregulates catabolic and proinflammatory pathways capable of inducing cartilage degradation: a basic science study. *Arthritis Research & Therapy* **9**, R121 (2007).
129. Milner, J. M. *et al.* Matriptase is a novel initiator of cartilage matrix degradation in osteoarthritis. *Arthritis and Rheumatism* **62**, 1955–66 (2010).
130. Wang, Q. *et al.* Identification of a central role for complement in osteoarthritis. *Nature Medicine* **17**, 1674–9 (2011).
131. Bonnet, C. S. & Walsh, D. A. Osteoarthritis, angiogenesis and inflammation. *Rheumatology* **44**, 7–16 (2005).
132. Husa, M., Liu-Bryan, R. & Terkeltaub, R. Shifting HIFs in osteoarthritis. *Nature Medicine* **16**, 641–4 (2010).
133. Pufe, T. *et al.* Vascular endothelial growth factor (VEGF) induces matrix metalloproteinase expression in immortalized chondrocytes. *The Journal of Pathology* **202**, 367–74 (2004).
134. Valdes, A. M. & Spector, T. D. The contribution of genes to osteoarthritis. *Rheumatic diseases clinics of North America* **34**, 581–603 (2008).
135. Lindberg, H. Prevalence of primary coxarthrosis in siblings of patients with primary coxarthrosis. *Clinical Orthopaedics and Related Research* 273–5 (1986).
136. Lanyon, P., Muir, K., Doherty, S. & Doherty, M. Assessment of a genetic contribution to osteoarthritis of the hip: sibling study. *BMJ* **321**, 1179–83 (2000).
137. Ingvarsson, T. Prevalence and inheritance of hip osteoarthritis in Iceland. *Acta orthopaedica Scandinavica. Supplementum* **298**, 1–46 (2000).
138. Jonsson, H. *et al.* The inheritance of hand osteoarthritis in Iceland. *Arthritis and Rheumatism* **48**, 391–5 (2003).

139. Chitnavis, J. *et al.* Genetic influences in end-stage osteoarthritis. Sibling risks of hip and knee replacement for idiopathic osteoarthritis. *The Journal of Bone and Joint Surgery. British volume* **79**, 660–4 (1997).
140. Valdes, A. M. *et al.* The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. *Annals of the Rheumatic Diseases* **70**, 873–5 (2011).
141. Zeggini E, Panoutsopoulou K, Southam L, Rayner NW, Day-Williams AG, Lopes MC, Boraska V, Esko T, Evangelou E, Hofman A, Houwing-Duistermaat JJ, Ingvarsson T, Jonsdottir I, Jonsson H, Kerkhof HJ, Kloppenburg M, Bos SD, Mangino M, Metrustry S, Slagboom PE, , L. J. Identification of new susceptibility loci for osteoarthritis (arcOGEN ): a genome-wide association study. *Lancet* **380**, 815-23 (2012).
142. Visscher, P. M., Brown, M. A., McCarthy, M. I. & Yang, J. Five years of GWAS discovery. *American Journal of Human Genetics* **90**, 7–24 (2012).
143. Panoutsopoulou, K. *et al.* Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study. *Annals of the Rheumatic Diseases* **70**, 864–7 (2011).
144. Valdes, A. M. *et al.* Sex and ethnic differences in the association of ASPN, CALM1, COL2A1, COMP, and FRZB with genetic susceptibility to osteoarthritis of the knee. *Arthritis and Rheumatism* **56**, 137–46 (2007).
145. Kerkhof, H. J. M. *et al.* Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **19**, 254–64 (2011).
146. Ioannidis, J. P. A. *et al.* Assessment of cumulative evidence on genetic associations: interim guidelines. *International Journal of Epidemiology* **37**, 120–32 (2008).
147. International, T. & Consortium, H. A haplotype map of the human genome. *Nature* **437**, 1299–320 (2005).
148. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–78 (2007).
149. Manolio, T. A. Genomewide Association Studies and Assessment of the Risk of Disease. *N Engl J Med* 2010;363:166-76.
150. Valdes, A. M. *et al.* The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide

- statistical significance. *Annals of the Rheumatic Diseases* **70**, 873–5 (2011).
151. Evangelou, E. *et al.* Meta-analysis of genome-wide associations studies confirms a susceptibility locus for knee osteoarthritis on chromosome 7q22. *Annals of the Rheumatic Diseases* **70**, 349–355 (2011).
152. Day-Williams, A. G. *et al.* A variant in MCF2L is associated with osteoarthritis. *American Journal of Human Genetics* **89**, 446–50 (2011).
153. Van Meurs, J. B. J. & Uitterlinden, A. G. Osteoarthritis year 2012 in review: genetics and genomics. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* (2012).doi:10.1016/j.joca.2012.08.007
154. Kerkhof, H. J. M. *et al.* Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **19**, 254–64 (2011).
155. Castaño Betancourt, M. C. *et al.* Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8218–23 (2012).
156. Malfait, A.-M. *et al.* A role for PACE4 in osteoarthritis pain: evidence from human genetic association and null mutant phenotype. *Annals of the Rheumatic Diseases* **71**, 1042–8 (2012).
157. Gibson, G. Rare and common variants: twenty arguments. *Nature Reviews. Genetics* **13**, 135–45 (2011).
158. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
159. Kizawa, H. *et al.* An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. *Nature Genetics* **37**, 138–44 (2005).
160. Mustafa, Z. *et al.* Investigating the aspartic acid (D) repeat of asporin as a risk factor for osteoarthritis in a UK Caucasian population. *Arthritis and Rheumatism* **52**, 3502–6 (2005).
161. Nakamura, T. *et al.* Meta-analysis of association between the ASPN D-repeat and osteoarthritis. *Human Molecular Genetics* **16**, 1676–81 (2007).
162. Kou, I., Nakajima, M. & Ikegawa, S. Binding characteristics of the osteoarthritis-associated protein asporin. *Journal of Bone and Mineral Metabolism* **28**, 395–402 (2010).



163. Sakao, K. *et al.* Asporin and transforming growth factor-beta gene expression in osteoblasts from subchondral bone and osteophytes in osteoarthritis. *Journal of Orthopaedic Science: official journal of the Japanese Orthopaedic Association* **14**, 738–47 (2009).
164. Mototani, H. *et al.* A functional single nucleotide polymorphism in the core promoter region of CALM1 is associated with hip osteoarthritis in Japanese. *Human Molecular Genetics* **14**, 1009–17 (2005).
165. Loughlin, J., Sinsheimer, J. S., Carr, A. & Chapman, K. The CALM1 core promoter polymorphism is not associated with hip osteoarthritis in a United Kingdom Caucasian population. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **14**, 295–8 (2006).
166. Poulou, M. *et al.* Association of the CALM1 core promoter polymorphism with knee osteoarthritis in patients of Greek origin. *Genetic testing* **12**, 263–5 (2008).
167. Shi, D. *et al.* Lack of association between the CALM1 core promoter polymorphism (-16C/T) and susceptibility to knee osteoarthritis in a Chinese Han population. *BMC Medical Genetics* **9**, 91 (2008).
168. Valhmu, W. B. & Raia, F. J. myo-Inositol 1,4,5-trisphosphate and Ca(2+)/calmodulin-dependent factors mediate transduction of compression-induced signals in bovine articular chondrocytes. *The Biochemical Journal* **361**, 689–96 (2002).
169. Husa, M., Liu-Bryan, R. & Terkeltaub, R. Shifting HIFs in osteoarthritis. *Nature Medicine* **16**, 641–4 (2010).
170. Valdes, A. M. *et al.* Genetic variation in the SMAD3 gene is associated with hip and knee osteoarthritis. *Arthritis and Rheumatism* **62**, 2347–52 (2010).
171. Wu, Q. *et al.* Induction of an osteoarthritis-like phenotype and degradation of phosphorylated Smad3 by Smurf2 in transgenic mice. *Arthritis and Rheumatism* **58**, 3132–44 (2008).
172. Van de Laar, I. M. B. H. *et al.* Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. *Nature Genetics* **43**, 121–6 (2011).
173. Van de Laar, I. M. B. H. *et al.* Phenotypic spectrum of the SMAD3-related aneurysms-osteoarthritis syndrome. *Journal of Medical Genetics* **49**, 47–57 (2012).
174. Stefánsson, S. E. *et al.* Genomewide scan for hand osteoarthritis: a novel mutation in matrilin-3. *American Journal of Human Genetics* **72**, 1448–59 (2003).

175. Min, J. L. *et al.* Association of matrilin-3 polymorphisms with spinal disc degeneration and osteoarthritis of the first carpometacarpal joint of the hand. *Annals of the Rheumatic Diseases* **65**, 1060–6 (2006).
176. Jayasuriya, C. T., Goldring, M. B., Terek, R. & Chen, Q. Matrilin-3 induction of IL-1 receptor antagonist is required for up-regulating collagen II and aggrecan as well as down-regulating ADAMTS-5 gene expression. *Arthritis Research & Therapy* **14**, R197 (2012).
177. Vincourt, J.-B. *et al.* Increased expression of matrilin-3 not only in osteoarthritic articular cartilage but also in cartilage-forming tumors, and down-regulation of SOX9 via epidermal growth factor domain 1-dependent signaling. *Arthritis and Rheumatism* **58**, 2798–808 (2008).
178. Zhai, G. *et al.* A genome-wide association study suggests that a locus within the ataxin 2 binding protein 1 gene is associated with hand osteoarthritis: the Treat-OA consortium. *Journal of Medical Genetics* **46**, 614–6 (2009).
179. Leppävuori, J. *et al.* Genome scan for predisposing loci for distal interphalangeal joint osteoarthritis: evidence for a locus on 2q. *American Journal of Human genetics* **65**, 1060–7 (1999).
180. Solovieva, S. *et al.* Association between interleukin 1 gene cluster polymorphisms and bilateral distal interphalangeal osteoarthritis. *The Journal of Rheumatology* **36**, 1977–86 (2009).
181. Kerkhof, H. J. M. *et al.* Large-scale meta-analysis of interleukin-1 beta and interleukin-1 receptor antagonist polymorphisms on risk of radiographic hip and knee osteoarthritis and severity of knee osteoarthritis. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **19**, 265–71 (2011).
182. Loughlin, J. *et al.* Stratification analysis of an osteoarthritis genome screen-suggestive linkage to chromosomes 4, 6, and 16. *American Journal of Human genetics* **65**, 1795–8 (1999).
183. Loughlin, J. *et al.* Linkage analysis of chromosome 2q in osteoarthritis. *Rheumatology* **39**, 377–81 (2000).
184. Loughlin, J. *et al.* Finer linkage mapping of a primary hip osteoarthritis susceptibility locus on chromosome 6. *European Journal of Human Genetics : EJHG* **10**, 562–8 (2002).
185. Southam, L., Chapman, K. & Loughlin, J. Genetic association analysis of BMP5 as a potential osteoarthritis susceptibility gene. *Rheumatology* **42**, 911–2 (2003).



186. Southam, L. *et al.* Microsatellite association mapping of a primary osteoarthritis susceptibility locus on chromosome 6p12.3-q13. *Arthritis and Rheumatism* **50**, 3910–4 (2004).
187. Wilkins, J. M., Southam, L., Mustafa, Z., Chapman, K. & Loughlin, J. Association of a functional microsatellite within intron 1 of the BMP5 gene with susceptibility to osteoarthritis. *BMC Medical Genetics* **10**, 141 (2009).
188. Wilkins, J. M. *et al.* Extreme context specificity in differential allelic expression. *Human molecular genetics* **16**, 537–46 (2007).
189. Pregizer, S. & Mortlock, D. P. Control of BMP gene expression by long-range regulatory elements. *Cytokine & Growth Factor reviews* **20**, 509–15
190. Snelling, S. J. B., Hulley, P. A. & Loughlin, J. BMP5 activates multiple signaling pathways and promotes chondrogenic differentiation in the ATDC5 growth plate model. *Growth Factors* **28**, 268–79 (2010).
191. Forster, T., Chapman, K. & Loughlin, J. Common variants within the interleukin 4 receptor alpha gene (IL4R) are associated with susceptibility to osteoarthritis. *Human Genetics* **114**, 391–5 (2004).
192. Limer, K. L. *et al.* Attempt to replicate published genetic associations in a large, well-defined osteoarthritis case-control population (the GOAL study). *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **17**, 782–9 (2009).
193. Snelling, S., Ferreira, A. & Loughlin, J. Allelic expression analysis suggests that cis-acting polymorphism of FRZB expression does not contribute to osteoarthritis susceptibility. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **15**, 90–2 (2007).
194. Evangelou, E. *et al.* Large-scale analysis of association between GDF5 and FRZB variants and osteoarthritis of the hip, knee, and hand. *Arthritis and Rheumatism* **60**, 1710–21 (2009).
195. Baker-Lepain, J. C. *et al.* Variant alleles of the Wnt antagonist FRZB are determinants of hip shape and modify the relationship between hip shape and osteoarthritis. *Arthritis and Rheumatism* **64**, 1457–65 (2012).
196. Meulenberg, I. *et al.* Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. *Human Molecular Genetics* **17**, 1867–75 (2008).

197. Waarsing, J. H. *et al.* Osteoarthritis susceptibility genes influence the association between hip morphology and osteoarthritis. *Arthritis and Rheumatism* **63**, 1349–54 (2011).
198. Miura, M. *et al.* Thyroid hormones promote chondrocyte differentiation in mouse ATDC5 cells and stimulate endochondral ossification in fetal mouse tibias through iodothyronine deiodinases in the growth plate. *Journal of Bone and Mineral Research: the official journal of the American Society for Bone and Mineral Research* **17**, 443–54 (2002).
199. Cheng, A. W. M., Bolognesi, M. & Kraus, V. B. DIO2 modifies inflammatory responses in chondrocytes. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **20**, 440–5 (2012).
200. Bos, S. D. *et al.* Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. *Annals of the Rheumatic Diseases* **71**, 1254–8 (2012).
201. Miyamoto, Y. *et al.* Common variants in DVWA on chromosome 3p24.3 are associated with susceptibility to knee osteoarthritis. *Nature Genetics* **40**, 994–8 (2008).
202. Valdes, A. M. *et al.* Association of the DVWA and GDF5 polymorphisms with osteoarthritis in UK populations. *Annals of the Rheumatic Diseases* **68**, 1916–20 (2009).
203. Meulenbelt, I. *et al.* Large replication study and meta-analyses of DVWA as an osteoarthritis susceptibility locus in European and Asian populations. *Human Molecular Genetics* **18**, 1518–23 (2009).
204. Wagener, R., Gara, S. K., Kobbe, B., Paulsson, M. & Zaucke, F. The knee osteoarthritis susceptibility locus DVWA on chromosome 3p24.3 is the 5' part of the split COL6A4 gene. *Matrix Biology: journal of the International Society for Matrix Biology* **28**, 307–10 (2009).
205. Nakajima, M., Miyamoto, Y. & Ikegawa, S. Cloning and characterization of the osteoarthritis-associated gene DVWA. *Journal of Bone and Mineral Metabolism* **29**, 300–8 (2011).
206. Spector, T. D. *et al.* Association between a variation in LRCH1 and knee osteoarthritis: a genome-wide single-nucleotide polymorphism association study using DNA pooling. *Arthritis and Rheumatism* **54**, 524–32 (2006).

207. Snelling, S., Sinsheimer, J. S., Carr, A. & Loughlin, J. Genetic association analysis of LRCH1 as an osteoarthritis susceptibility locus. *Rheumatology (Oxford, England)* **46**, 250–2 (2007).
208. Jiang, Q. *et al.* Lack of association of single nucleotide polymorphism in LRCH1 with knee osteoarthritis susceptibility. *Journal of Human Genetics* **53**, 42–7 (2008).
209. Foussard, H. *et al.* LRCH proteins: a novel family of cytoskeletal regulators. *PLoS One* **5**, e12257 (2010).
210. Valdes, A. M. *et al.* Genome-wide association scan identifies a prostaglandin-endoperoxide synthase 2 variant involved in risk of knee osteoarthritis. *American Journal of Human Genetics* **82**, 1231–40 (2008).
211. Amin, A. R., Dave, M., Attur, M. & Abramson, S. B. COX-2, NO, and cartilage damage and repair. *Current Rheumatology reports* **2**, 447–53 (2000).
212. Shi, D. *et al.* Association of single-nucleotide polymorphisms in HLA class II/III region with knee osteoarthritis. *Osteoarthritis and Cartilage* **18**, 1454–1457 (2010).
213. Valdes, A. M. *et al.* Large scale replication study of the association between HLA II/BTNL2 variants and osteoarthritis of the knee in European-descent populations. *PLoS One* **6**, e23371 (2011).
214. Miyamoto, Y. *et al.* A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. *Nature Genetics* **39**, 529–33 (2007).
215. Chapman, K. *et al.* A meta-analysis of European and Asian cohorts reveals a global role of a functional SNP in the 5' UTR of GDF5 with osteoarthritis susceptibility. *Human Molecular Genetics* **17**, 1497–504 (2008).
216. Williams, F. M. K. *et al.* GDF5 single-nucleotide polymorphism rs143383 is associated with lumbar disc degeneration in Northern European women. *Arthritis and Rheumatism* **63**, 708–12 (2011).
217. Sanna, S. *et al.* Common variants in the GDF5-UQCC region are associated with variation in human height. *Nature Genetics* **40**, 198–203 (2008).
218. Vaes, R. B. A. *et al.* Genetic variation in the GDF5 region is associated with osteoarthritis, height, hip axis length and fracture risk: the Rotterdam study. *Annals of the Rheumatic Diseases* **68**, 1754–60 (2009).
219. Valdes, A. M. *et al.* Genetic contribution to radiographic severity in osteoarthritis of the knee. *Annals of the Rheumatic Diseases* **71**, 1537–40 (2012).

220. Edwards, C. J. & Francis-West, P. H. Bone morphogenetic proteins in the development and healing of synovial joints. *Seminars in Arthritis and Rheumatism* **31**, 33–42 (2001).
221. Francis-West, P. H. *et al.* Mechanisms of GDF-5 action during skeletal development. *Development* **126**, 1305–15 (1999).
222. Everman, D. B. *et al.* The mutational spectrum of brachydactyly type C. *American Journal of Medical Genetics* **112**, 291–6 (2002).
223. Byrnes, A. M. *et al.* Mutations in GDF5 presenting as semidominant brachydactyly A1. *Human Mutation* **31**, 1155–62 (2010).
224. Stelzer, C., Winterpacht, A., Spranger, J. & Zabel, B. Grebe dysplasia and the spectrum of CDMP1 mutations. *Pediatric Pathology & Molecular Medicine* **22**, 77–85
225. Daans, M., Luyten, F. P. & Lories, R. J. U. GDF5 deficiency in mice is associated with instability-driven joint damage, gait and subchondral bone changes. *Annals of the Rheumatic Diseases* **70**, 208–13 (2011).
226. Settle, S. H. *et al.* Multiple joint and skeletal patterning defects caused by single and double mutations in the mouse *Gdf6* and *Gdf5* genes. *Developmental Biology* **254**, 116–30 (2003).
227. Egli, R. J. *et al.* Functional analysis of the osteoarthritis susceptibility-associated GDF5 regulatory polymorphism. *Arthritis and Rheumatism* **60**, 2055–64 (2009).
228. Reynard, L. N., Bui, C., Canty-Laird, E. G., Young, D. A. & Loughlin, J. Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. *Human Molecular Genetics* **20**, 3450–60 (2011).
229. Kerkhof, H. J. M. *et al.* A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. *Arthritis and Rheumatism* **62**, 499–510 (2010).
230. Raine, E. V. A., Wreglesworth, N., Dodd, A. W., Reynard, L. N. & Loughlin, J. Gene expression analysis reveals HBP1 as a key target for the osteoarthritis susceptibility locus that maps to chromosome 7q22. *Annals of the Rheumatic Diseases* **71**, 2020–7 (2012).
231. Mio, F. *et al.* A functional polymorphism in COL11A1, which encodes the alpha 1 chain of type XI collagen, is associated with susceptibility to lumbar disc herniation. *American Journal of Human Genetics* **81**, 1271–7 (2007).
232. Raine, E. V. A., Dodd, A. W., Reynard, L. N. & Loughlin, J. Allelic expression analysis of the osteoarthritis susceptibility gene

- COL11A1 in human joint tissues. *BMC musculoskeletal disorders* **14**, 85 (2013).
233. Xu, L. *et al.* Osteoarthritis-like changes and decreased mechanical function of articular cartilage in the joints of mice with the chondrodysplasia gene (cho). *Arthritis and Rheumatism* **48**, 2509–18 (2003).
234. Annunen, S. *et al.* Splicing mutations of 54-bp exons in the COL11A1 gene cause Marshall syndrome, but other mutations cause overlapping Marshall/Stickler phenotypes. *American Journal of Human Genetics* **65**, 974–83 (1999).
235. Tompson, S. W. *et al.* Fibrochondrogenesis results from mutations in the COL11A1 type XI collagen gene. *American Journal of Human Genetics* **87**, 708–12 (2010).
236. Jakkula, E. *et al.* The role of sequence variations within the genes encoding collagen II, IX and XI in non-syndromic, early-onset osteoarthritis. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society International* **13**, 497–507 (2005).
237. Okano, K. *et al.* Expression of parathyroid hormone-related peptide in human osteoarthritis. *Journal of Orthopaedic Research: official publication of the Orthopaedic Research Society* **15**, 175–80 (1997).
238. Liu, F. *et al.* A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS Genetics* **8**, (2012).
239. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–8 (2010).
240. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–73 (2010).
241. Iannone, F. *et al.* Increased expression of nerve growth factor (NGF) and high affinity NGF receptor (p140 TrkA) in human osteoarthritic chondrocytes. *Rheumatology* **41**, 1413–8 (2002).
242. Webb, R. *et al.* Early disease onset is predicted by a higher genetic risk for lupus and is associated with a more severe phenotype in lupus patients. *Annals of the Rheumatic Diseases* **70**, 151–6 (2011).
243. Herrero-Beaumont, G., Roman-Blas, J. A., Castañeda, S. & Jimenez, S. A. Primary osteoarthritis no longer primary: three subsets with distinct etiological, clinical, and therapeutic characteristics. *Seminars in Arthritis and Rheumatism* **39**, 71–80 (2009).

## BIBLIOGRAFÍA

---

244. Kannu, P., Bateman, J. F., Belluoccio, D., Fosang, A. J. & Savarirayan, R. Employing molecular genetics of chondrodysplasias to inform the study of osteoarthritis. *Arthritis and Rheumatism* **60**, 325–34 (2009).
245. Styrkarsdottir, U. *et al.* New sequence variants associated with bone mineral density. *Nature genetics* **41**, 15–7 (2009).
246. Hofman, A. *et al.* The Rotterdam Study: objectives and design update. *European Journal of Epidemiology* **22**, 819–29 (2007).
247. Kiel, D. P. *et al.* Genome-wide association with bone mass and geometry in the Framingham Heart Study. *BMC medical genetics* **8 Suppl 1**, S14 (2007).
248. Riyazi, N. *et al.* Evidence for familial aggregation of hand, hip, and spine but not knee osteoarthritis in siblings with multiple joint involvement: the GARP study. *Annals of the Rheumatic Diseases* **64**, 438–43 (2005).
249. Metspalu, A., Köhler, F., Laschinski, G., Ganten, D. & Roots, I. [The Estonian Genome Project in the context of European genome research]. *Deutsche medizinische Wochenschrift* **129 Suppl**, S25–8 (2004).
250. Rodriguez-Lopez, J., Pombo-Suarez, M., Liz, M., Gomez-Reino, J. J. & Gonzalez, A. Further evidence of the role of frizzled-related protein gene polymorphisms in osteoarthritis. *Annals of the Rheumatic Diseases* **66**, 1052–5 (2007).
251. Rego-Pérez, I., Fernández-Moreno, M., Fernández-López, C., Arenas, J. & Blanco, F. J. Mitochondrial DNA haplogroups: role in the prevalence and severity of knee osteoarthritis. *Arthritis and Rheumatism* **58**, 2387–96 (2008).
252. Riancho, J. A. *et al.* Wnt receptors, bone mass, and fractures: gene-wide association analysis of LRP5 and LRP6 polymorphisms with replication. *European Journal of Endocrinology* **164**, 123–31 (2011).
253. Kostopoulou, F. *et al.* Central role of SREBP-2 in the pathogenesis of osteoarthritis. *PloS One* **7**, e35753 (2012).
254. Valdes, A. M. *et al.* Variation at the ANP32A gene is associated with risk of hip osteoarthritis in women. *Arthritis and Rheumatism* **60**, 2046–54 (2009).
255. Holliday, K. L. *et al.* Lifetime body mass index, other anthropometric measures of obesity and risk of knee or hip osteoarthritis in the GOAL case-control study. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society* **19**, 37–43 (2011).



256. Yu, W., Gwinn, M., Clyne, M., Yesupriya, A. & Khoury, M. J. A navigator for human genome epidemiology. *Nature Genetics* **40**, 124–5 (2008).
257. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype imputation. *Annual Review of Genomics and Human Genetics* **10**, 387–406 (2009).
258. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature Genetics* **39**, 906–13 (2007).
259. Gao, X., Starmer, J. & Martin, E. R. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genetic Epidemiology* **32**, 361–9 (2008).
260. Dupont, W. D. & Plummer, W. D. Power and sample size calculations. A review and computer program. *Controlled clinical trials* **11**, 116–28 (1990).
261. Sabb, F. W. *et al.* Challenges in phenotype definition in the whole-genome era: multivariate models of memory and intelligence. *Neuroscience* **164**, 88–107 (2009).
262. Kring, S. I. I. *et al.* Genotype-phenotype associations in obesity dependent on definition of the obesity phenotype. *Obesity Facts* **1**, 138–45 (2008).
263. Moonesinghe, R., Khoury, M. J., Liu, T. & Ioannidis, J. P. A. Required sample size and nonreplicability thresholds for heterogeneous genetic associations. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 617–22 (2008).
264. Morton, N. E. & Collins, A. Tests and estimates of allelic association in complex inheritance. *Proceedings of the National Academy of Sciences* **95**, 11389–11393 (1998).
265. Hochberg, M. C., Yerges-Armstrong, L., Yau, M. & Mitchell, B. D. Genetic epidemiology of osteoarthritis: recent developments and future directions. *Current Opinion in Rheumatology* **25**, 192–7 (2013).
266. Sharif, M., Kirwan, J. R., Elson, C. J., Granell, R. & Clarke, S. Suggestion of nonlinear or phasic progression of knee osteoarthritis based on measurements of serum cartilage oligomeric matrix protein levels over five years. *Arthritis and Rheumatism* **50**, 2479–88 (2004).
267. Dore, D. *et al.* Natural history and clinical significance of MRI-detected bone marrow lesions at the knee: a prospective study in

- community dwelling older adults. *Arthritis Research & Therapy* **12**, R223 (2010).
268. Wirth, W. *et al.* MRI-based extended ordered values more efficiently differentiate cartilage loss in knees with and without joint space narrowing than region-specific approaches using MRI or radiography--data from the OA initiative. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society* **19**, 689–99 (2011).
269. Kornaat, P. R. *et al.* Osteoarthritis of the knee: association between clinical features and MR imaging findings. *Radiology* **239**, 811–7 (2006).
270. Loeser, R. F. Aging processes and the development of osteoarthritis. *Current opinion in Rheumatology* **25**, 108-13 (2012).
271. Meulenbelt, I. *et al.* Meta-analyses of genes modulating intracellular T3 bio-availability reveal a possible role for the DIO3 gene in osteoarthritis susceptibility. *Annals of the Rheumatic Diseases* **70**, 164–7 (2011).
272. Dodd, A. W., Syddall, C. M. & Loughlin, J. A rare variant in the osteoarthritis-associated locus GDF5 is functional and reveals a site that can be manipulated to modulate GDF5 expression. *European journal of human genetics : EJHG* **21**, 517-21 (2012).
273. Syddall, C. M., Reynard, L. N., Young, D. A. & Loughlin, J. The identification of trans-acting Factors that regulate the expression of GDF5 via the osteoarthritis susceptibility SNP, rs143383. *PLoS Genetics* **9**, 6 (2013).
274. Wu, D.-D., Li, G.-M., Jin, W., Li, Y. & Zhang, Y.-P. Positive selection on the osteoarthritis-risk and decreased-height associated variants at the GDF5 gene in East Asians. *PloS One* **7**, 8 (2012).
275. Fytili, P. *et al.* Association of repeat polymorphisms in the estrogen receptors alpha, beta, and androgen receptor genes with knee osteoarthritis. *Clinical Genetics* **68**, 268–77 (2005).
276. Girirajan, S. & Eichler, E. E. Phenotypic variability and genetic susceptibility to genomic disorders. *Human Molecular Genetics* **19**, R176–87 (2010).
277. Magaña, J. J. *et al.* Association of the calcitonin gene (CA) polymorphism with osteoarthritis of the knee in a Mexican mestizo population. *The Knee* **17**, 157–60 (2010).
278. Zeggini, E. & Ioannidis, J. P. A. Meta-analysis in genome-wide association studies. *Pharmacogenomics* **10**, 191–201 (2009).



279. Raychaudhuri, S. *et al.* Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genetics* **5**, 6 (2009).
280. Khoury, M. J. *et al.* Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases. *American Journal of Epidemiology* **170**, 269–79 (2009).
281. Parentin, F., Sangalli, A., Mottes, M. & Perissutti, P. Stickler syndrome and vitreoretinal degeneration: correlation between locus mutation and vitreous phenotype. Apropos of a case. *Graefe's archive for clinical and experimental ophthalmology* **239**, 316–9 (2001).
282. Stempel, J., Fritsch, H., Pfaller, K. & Blumer, M. J. F. Development of articular cartilage and the metaphyseal growth plate: the localization of TRAP cells, VEGF, and endostatin. *Journal of Anatomy* **218**, 608–18 (2011).
283. Ludin, A. *et al.* Injection of vascular endothelial growth factor into knee joints induces osteoarthritis in mice. *Osteoarthritis and Cartilage / OARSI, Osteoarthritis Research Society* **21**, 491–7 (2013).
284. Saito, T. *et al.* Transcriptional regulation of endochondral ossification by HIF-2alpha during skeletal growth and osteoarthritis development. *Nature Medicine* **16**, 678–86 (2010).
285. Sharma, S. *et al.* Association of VEGF polymorphisms with childhood asthma, lung function and airway responsiveness. *The European Respiratory Journal : official journal of the European Society for Clinical Respiratory Physiology* **33**, 1287–94 (2009).
286. Ioannidis, J. P. A., Trikalinos, T. A., Ntzani, E. E. & Contopoulos-Ioannidis, D. G. Genetic associations in large versus small studies: an empirical assessment. *Lancet* **361**, 567–71 (2003).
287. Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S. & Hirschhorn, J. N. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genetics* **33**, 177–82 (2003).





***ANEXO: Publicaciones 4 y 5***

---



## *Anexo*

### ***Estudios de asociación caso-control de SNPs de la región de HLA clase II/III y de los genes CYP19A1 y ESR1.***

***Publicación 4:*** Nakajima M, Takahashi A, Kou I, **Rodríguez-Fontenla C**, Gomez-Reino JJ, Furuichi T, Dai J, Sudo A, Uchida A, Fukui N, Kubo M, Kamatani N, Tsunoda T, Malizos KN, Tsezou A, Gonzalez A, Nakamura Y, Ikegawa S. **New sequence variants in HLA class II/III region associated with susceptibility to knee osteoarthritis identified by genome-wide association study.** PLoS One. 2010 Mar 18;5(3):e9723.

***Publicación 5:*** Riancho JA, García-Ibarbia C, Gravani A, Raine EV, **Rodríguez-Fontenla C**, Soto-Hermida A, Rego-Perez I, Dodd AW, Gómez-Reino JJ, Zarrabeitia MT, Garcés CM, Carr A, Blanco F, González A, Loughlin J. **Common variations in estrogen-related genes are associated with severe large-joint osteoarthritis: a multicenter genetic and functional study.** Osteoarthritis Cartilage. 2010 Jul;18(7):927-33.



# New Sequence Variants in HLA Class II/III Region Associated with Susceptibility to Knee Osteoarthritis Identified by Genome-Wide Association Study

Masahiro Nakajima<sup>1</sup>, Atsushi Takahashi<sup>2</sup>, Ikuyo Kou<sup>1</sup>, Cristina Rodríguez-Fontenla<sup>3</sup>, Juan J. Gomez-Reino<sup>3,4</sup>, Tatsuya Furuichi<sup>1</sup>, Jin Dai<sup>1,5</sup>, Akihiro Sudo<sup>6</sup>, Atsumasa Uchida<sup>6</sup>, Naoshi Fukui<sup>7</sup>, Michiaki Kubo<sup>8</sup>, Naoyuki Kamatani<sup>2</sup>, Tatsuhiko Tsunoda<sup>9</sup>, Konstantinos N. Malizos<sup>10,11</sup>, Aspasia Tsezou<sup>12</sup>, Antonio Gonzalez<sup>3</sup>, Yusuke Nakamura<sup>13,14</sup>, Shiro Ikegawa<sup>1\*</sup>

**1** Laboratory for Bone and Joint Diseases, Center for Genomic Medicine, RIKEN, Tokyo, Japan, **2** Laboratory for Statistical Analysis, Center for Genomic Medicine, RIKEN, Tokyo, Japan, **3** Laboratorio Investigación 10 and Rheumatology Unit, Hospital Clínico Universitario de Santiago, Santiago de Compostela, Spain, **4** Department of Medicine, University of Santiago de Compostela, Santiago de Compostela, Spain, **5** The Center of Diagnosis and Treatment for Joint Disease, Drum Tower Hospital Affiliated to Medical School of Nanjing University, Nanjing, China, **6** Department of Orthopaedic Surgery, Mie University Faculty of Medicine, Mie, Japan, **7** Department of Pathomechanisms, Clinical Research Center for Rheumatology and Allergy, National Hospital Organization Sagami National Hospital, Kanagawa, Japan, **8** Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, Kanagawa, Japan, **9** Laboratory for Medical Informatics, Center for Genomic Medicine, RIKEN, Kanagawa, Japan, **10** Department of Orthopaedics University of Thessaly, Larissa, Greece, **11** Institute for Biomedical Research and Technology, Larissa, Greece, **12** Department of Biology, University of Thessaly Medical School, Larissa, Greece, **13** Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, **14** Center for Genomic Medicine, RIKEN, Kanagawa, Japan

## Abstract

Osteoarthritis (OA) is a common disease that has a definite genetic component. Only a few OA susceptibility genes that have definite functional evidence and replication of association have been reported, however. Through a genome-wide association study and a replication using a total of ~4,800 Japanese subjects, we identified two single nucleotide polymorphisms (SNPs) (rs7775228 and rs10947262) associated with susceptibility to knee OA. The two SNPs were in a region containing HLA class II/III genes and their association reached genome-wide significance (combined  $P = 2.43 \times 10^{-8}$  for rs7775228 and  $6.73 \times 10^{-9}$  for rs10947262). Our results suggest that immunologic mechanism is implicated in the etiology of OA.

**Citation:** Nakajima M, Takahashi A, Kou I, Rodríguez-Fontenla C, Gomez-Reino JJ, et al. (2010) New Sequence Variants in HLA Class II/III Region Associated with Susceptibility to Knee Osteoarthritis Identified by Genome-Wide Association Study. PLoS ONE 5(3): e9723. doi:10.1371/journal.pone.0009723

**Editor:** Sudha Agarwal, Ohio State University, United States of America

**Received:** December 28, 2009; **Accepted:** February 26, 2010; **Published:** March 18, 2010

**Copyright:** © 2010 Nakajima et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The Japanese study is supported by the Leading Project for Personalized Medicine in the Ministry of Education, Culture, Sports, Science and Technology. The European study is supported by Instituto de Salud Carlos III from the Spanish Health Ministry with participation of FEDER (Fondo Europeo para el Desarrollo Regional), the Foundation Mutua Madrile (Madrid). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

\* E-mail: sikegawa@ims.u-tokyo.ac.jp

## Introduction

We are living in the “Bone and Joint Decade” (<http://www.boneandjointdecade.org/>). As the WHO initiative shows, bone and joint diseases are serious problems all over the world, putting us under severe medical, economical and social burden. Osteoarthritis (OA; MIM 165720) is one of the most common diseases among them. OA affects synovial joints of all over the body, mainly knee, hip, hand and spine. OA is characterized by progressive loss of articular cartilage and, often, proliferation of synovium and bone, which lead to pain, loss of joint function and disability. More than tens of millions patients in the world are suffering from this non-lethal, but intractable disease, and the number is relentlessly increasing; however, its etiological picture remains unclear and we have no fundamental treatment for it.

OA is a polygenic disease. Both environmental and genetic factors contribute to its etiology and pathogenesis [1]. To understand its genetic factor, identification of its susceptibility

gene(s) must be the first step. Many OA susceptibility genes identified by candidate-gene association studies have been reported, but only a few have supporting functional evidence and replication of the results in different populations [1,2]. Large-scale association studies including the genome-wide association study (GWAS) using high-density single nucleotide polymorphisms (SNPs) have been reported by a few groups in Asia and Europe [3–6], but only a gene fulfilled genome-wide significance level [2]. The genetic basis of OA susceptibility remains largely uncharacterized. To identify OA susceptibility gene(s), we conducted a GWAS for knee OA and identified two SNPs with genome-wide significance level.

## Methods

### Samples

Characteristics of each cohort group are shown in Table 1. Case samples of GWAS for the Japanese population were obtained from

**Table 1.** Basal characteristics of the subjects.

Cohort	Source	Platform	Number of samples	Nationality	Female (%)	Age (mean $\pm$ sd)	BMI (mean $\pm$ sd)	Severity <sup>a</sup> (% severe OA)
<b>GWAS</b>								
knee OA	RIKEN	Illumina HumanHap550	899	Japanese	759 (84.4)	71.6 $\pm$ 7.6	24.9 $\pm$ 3.6	76.5
control	ORC+BioBank Japan	Illumina HumanHap550	3,396	Japanese	1,491 (43.9)	52.5 $\pm$ 15.2	22.5 $\pm$ 3.7	–
<b>Replication</b>								
Japanese								
knee OA	RIKEN	Invader assay	167	Japanese	124 (74.3)	73.8 $\pm$ 6.1	24.5 $\pm$ 3.3	48.5
control	RIKEN	Invader assay	347	Japanese	223 (64.3)	65.9 $\pm$ 8.7	22.3 $\pm$ 2.7	–
European Caucasian								
knee OA	Santiago de Compostela	SNaPshot	243	Spanish	197 (81.1)	68.0 $\pm$ 5.7	32.8 $\pm$ 4.8	ND <sup>b</sup>
control	Santiago de Compostela	SNaPshot	426	Spanish	165 (38.7)	68.4 $\pm$ 9.1	28.3 $\pm$ 3.8	–
knee OA	University of Thessaly	SNaPshot	570	Greek	468 (82.1)	65.8 $\pm$ 8.7	29.1 $\pm$ 3.3	77.1
control	University of Thessaly	SNaPshot	645	Greek	417 (64.6)	59.5 $\pm$ 11.6	25.4 $\pm$ 3.7	–

OA: osteoarthritis, ORC: Osaka-Midosuji Rotary Club.

<sup>a</sup>Kellegren-Laurence grade  $\geq 3$  was considered as severe OA.

<sup>b</sup>All cases underwent TKR (total knee replacement) surgery.

doi:10.1371/journal.pone.0009723.t001

several medical institutes in Japan, as previously described [5,7]. Knee OA was diagnosed on the basis of clinical and radiographic findings using previously described criteria [5,7]. Rheumatoid arthritis (RA) and polyarthritis associated with autoimmune diseases were excluded, as were secondary OA due to crystal deposition (gout and pseudogout), posttraumatic OA and infection-induced OA. Patients who had clinical and radiographic findings suggestive of skeletal dysplasias, including overt short stature, multiple symmetric involvements of epiphyses and a definitely positive Mendelian family history were also excluded from the study. The control groups consisted of 3,396 individuals that were registered in the Leading Project for Personalized Medicine in the Ministry of Education, Culture, Sports, Science and Technology, Japan as the subjects with diseases unrelated to OA and the volunteers in the Osaka-Midosuji Rotary Club, Osaka, Japan [8]. For replication study, we recruited population-based cohorts from inhabitants of Odai and Minami-ise town (previously Miyagawa village and Nansai town, respectively in the Mie prefecture in Japan) [9]. The Spanish and Greek knee OA and control populations were recruited as described previously from the Hospital Clinico de Santiago, the Departments of Biology and Genetics and of Orthopaedics, University of Thessaly and the Institute of Musculoskeletal Sciences [10]. All the participants provided written informed consent. This research project was approved by the ethical committees at Center for Genomic Medicine (formerly, SNP Research Center), RIKEN and the participating institutions.

#### SNP genotyping

For the GWAS, we genotyped 906 patients with OA and 3,396 controls using Illumina HumanHap550v3 Genotyping BeadChip. After excluding seven cases with call rate of  $<0.98$ , we applied SNP QC (call rate of  $\geq 0.99$  in both cases and controls and  $P$  value of Hardy-Weinberg equilibrium test of  $\geq 1.0 \times 10^{-6}$  in controls). Finally, 459,393 SNPs on autosomal chromosomes passed the QC filters and were further analyzed. Among the SNPs analyzed in the GWAS, we selected top 15 SNPs showing the smallest  $P$  values ( $P < 1 \times 10^{-5}$ ) for the replication study using an independent 514 Japanese subjects

from a resident cohort. SNPs with minor allele frequency of  $\leq 0.1$  in both case and control samples were excluded from the further analysis. In the replication analysis, we genotyped SNPs using the multiplex PCR-based invader assay (Third Wave Technologies) or by direct sequencing of PCR products using ABI 3700 DNA analyzers (Applied Biosystems), or by SNaPshot Multiplex System (Applied Biosystems) according to manufacturers' protocols.

#### Statistical analysis

In the GWAS and replication analyses, we applied Fisher's exact test to two-by-two contingency table in three genetic models: an allele frequency model, a dominant-effect model, and a recessive-effect model. We conducted the meta-analysis using the Mantel-Haenszel method. We examined heterogeneity among studies by using the Breslow-Day test. Significance levels after the Bonferroni correction for multiple testing were  $P = 1.09 \times 10^{-7}$  (0.05/459,393). Age, gender- and BMI-adjusted odds ratios were obtained by logistic regression analysis [11]. Odds ratios and confidence intervals were calculated using the risk allele as a reference. We analyzed the haplotype association using Haploview software [12]. We conducted a principal component analysis to detect population stratification [13].

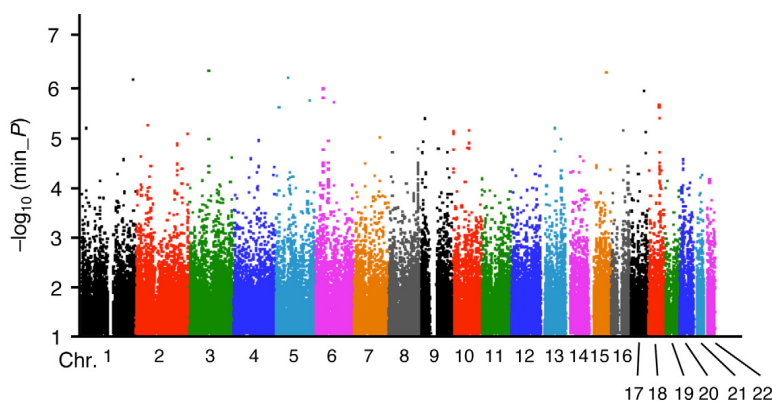
#### Software

For general statistical analysis, we used R statistical environment version 2.6.1 or Microsoft Excel. Drawing the LD map, estimation of haplotype frequencies and analysis of haplotype association were performed by Haploview software.

#### Results

To identify genetic variants that determine OA susceptibility, we conducted a GWAS in Japanese knee OA. We examined 906 individuals with knee OA and 3,396 control individuals (Table 1) using Illumina HumanHap550v3 Genotyping BeadChip. After confirming the data quality, we compared the results of 459,393 SNPs between cases and controls by Fisher's exact test for three genetic models: allelic, dominant or recessive (Figure 1). Fifteen





**Figure 1. Results of a genome-wide association study ( $-\log_{10} P$  value plot).** Each  $P$  value is the minimum of Fisher's exact tests for three models: dominant, recessive and allele frequency model. doi:10.1371/journal.pone.0009723.g001

SNPs selected for the replication study that had the smallest  $P$  values (minimum  $P < 1 \times 10^{-5}$ ) were next genotyped in an independent set of 167 Japanese knee OA individuals and 347

Japanese controls from a resident cohort study. Through these studies, only two SNPs, rs7775228 (combined  $P = 2.43 \times 10^{-8}$ ; OR = 1.34; 95% CI = 1.21–1.49) and rs10947262 (combined

**Table 2. Association of rs7775228 and rs10947262 with knee osteoarthritis.**

SNP (Nearest gene)	Allele	Population	Minor allele Frequency		OR (95% CI) <sup>a</sup>	$P^b$	$P_{het}^c$
			Case	Control			
rs7775228 (HLA-DQB1)	C/T	Japanese					
		GWAS	0.318	0.379	1.31 (1.18–1.47)	1.38 E-06	
		Replication	0.290	0.385	1.53 (1.15–2.02)	3.07 E-03	
		Combined <sup>d</sup>			1.34 (1.21–1.49)	2.43 E-08	0.33
		European Caucasian					
		Spanish	0.194	0.209	1.10 (0.83–1.45)	0.521	
		Greek	0.094	0.075	0.78 (0.58–1.03)	0.084	
European combined <sup>e</sup>			0.93 (0.76–1.13)	0.178	0.09		
All combined <sup>f</sup>			–	–	0.003		
rs10947262 (BTNL2)	C/T	Japanese					
		GWAS	0.358	0.419	1.30 (1.16–1.44)	2.45 E-06	
		Replication	0.332	0.422	1.47 (1.12–1.93)	5.74 E-03	
		Combined <sup>d</sup>			1.32 (1.19–1.46)	6.73 E-08	0.40
		European Caucasian					
		Spanish	0.122	0.136	1.13 (0.81–1.59)	0.465	
		Greek	0.068	0.094	1.43 (1.06–1.92)	0.019	
European combined <sup>e</sup>			1.29 (1.03–1.61)	0.025	0.32		
All combined <sup>f</sup>			1.31 (1.20–1.44)	5.10 E-09	0.63		

Odds ratios (ORs) and  $P$  values for the independence test were calculated by the Mantel-Haenszel method.

<sup>a</sup>OR of the risk allele from the two-by-two allele frequency table.

<sup>b</sup> $P$  values of the Pearson's  $\chi^2$  test for the allele model.

<sup>c</sup>Results of the Breslow-Day test.

<sup>d</sup>Meta-analysis of Japanese studies.

<sup>e</sup>Meta-analysis of European studies.

<sup>f</sup>Meta-analysis of all studies.

doi:10.1371/journal.pone.0009723.t002

$P=6.73 \times 10^{-8}$ ; OR = 1.32; 95% CI = 1.19–1.46) were significant even after the Bonferroni correction for multiple testing ( $P=1.09 \times 10^{-7}$ ) (Table 2). The two SNPs showing significant associations are located within a 340-kb region within the HLA locus, including *BTNL2*, *HLA-DRA*, *HLA-DRB5*, *HLA-DRB1*, *HLA-DQA1* and *HLA-DQB1* (Figure 2). Although the HLA region is known to show extensive linkage disequilibrium (LD) spanning over 7 Mb, only SNPs in the 340-kb region showed strong associations with OA (Figure 2), and SNPs outside of this region did not have significant association.

Application of the Cochran-Armitage test to all the tested SNPs indicated that the genetic inflation factor lambda was 1.08 for GWAS (Figure 3), implying a low possibility of false positive associations due to population stratification. We also carried out age, gender- and BMI-adjusted analysis using a logistic regression model, and confirmed similar association after adjustment (data not shown). The principal component analysis [13] revealed that there was no evidence for population stratification between the two control groups used for the GWAS (Figure S1).

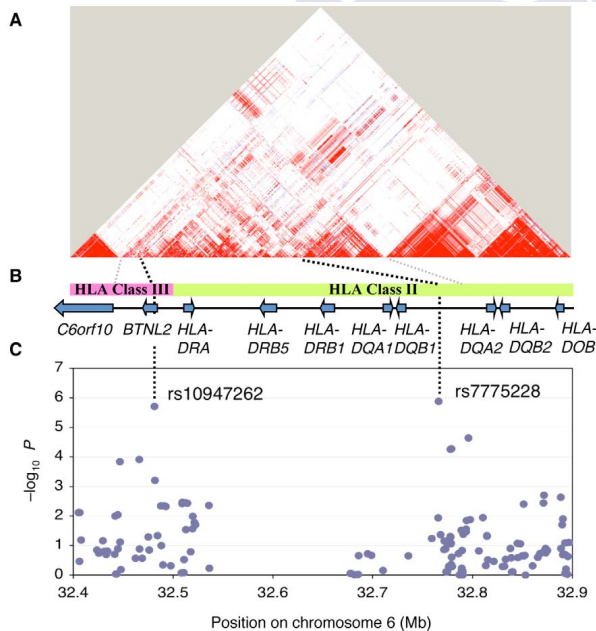
To check the association of rs7775228 and rs10947262 in different ethnic populations, we examined the association of the SNPs with knee OA in two European Caucasian populations from Greece and Spain. We genotyped a total of 813 OA and 1,071 control subjects (Table 1). We conducted the meta-analysis using the Mantel-Haenszel method. The combined European results for

rs7775228 were not significant with OR (95%CI) of 0.93 (0.76–1.13) (Table 2), while those for rs10947262 were supportive with OR (95%CI) of 1.29 (1.03–1.61). rs10947262 showed replication in the Greek population and the same trend in the Spanish population (Table 2). A meta-analysis of the Japanese and two European studies gave more significant association (combined  $P=5.10 \times 10^{-9}$ ).

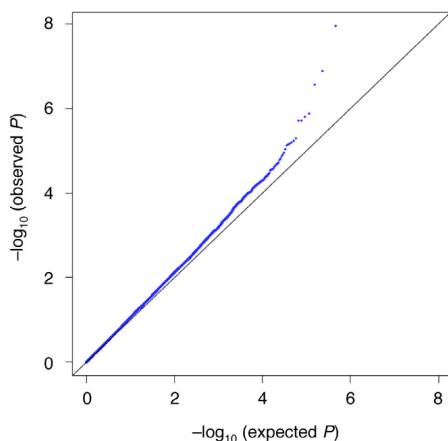
We estimated the pairwise LD indexes ( $D'$  and  $r^2$ ) between rs7775228 and rs10947262 using the genotype data of Japanese populations (GWAS and the replication study), and found that they were in strong LD with each other ( $D'=0.82$ ,  $r^2=0.56$ ). They formed two frequent haplotypes (Haplotype I and II; Table 3) accounting for about 90% of all observed chromosomes. The haplotypes were also significantly associated with knee OA; Haplotype I, the most frequent haplotype was a risk haplotype ( $P=1.48 \times 10^{-8}$ ; OR = 1.33; 95% CI = 1.20–1.46).

## Discussion

We performed a GWAS followed by a replication in an independent population using a total of ~4,800 Japanese subjects, and identified two SNPs (rs7775228 and rs10947262) in the HLA class II/III locus associated with susceptibility to knee OA. To our knowledge, this study represents the first GWAS of OA with extensive coverage (~550,000 markers) and definite genome-wide significance even after Bonferroni's correction, which is very



**Figure 2. Case-control association analysis and linkage disequilibrium (LD) map of the HLA class II/III region of chromosome 6.** (A) The LD map based on  $D'$  was drawn using HapMap data release 24 for the JPT population. (B) Genomic structure within the extended HLA-II/III region. (C) Results of GWAS for osteoarthritis in Japanese population. The  $\log_{10}$ -transformed  $P$  values are plotted on the  $y$  axis. doi:10.1371/journal.pone.0009723.g002



**Figure 3. GWAS of knee osteoarthritis.** Q-Q plot with Cochran-Armitage trend  $P$  in the GWAS. Horizontal and vertical lines represent expected  $P$  values under a null distribution and observed  $P$  values, respectively. The genetic inflation factor lambda is 1.08. doi:10.1371/journal.pone.0009723.g003

conservative. There were no effects of population stratification and confounding factors. Since two groups of controls were used in the GWAS, we evaluated the possibility of genetic heterogeneity between the two groups by a principal component analysis and found it unlikely (Figure S1). Although there was large age difference between the case and control groups of GWAS (Table 1), significant association was observed after the age adjustment.

In the NCBI genome database, rs7775228 and rs10947262 located between upstream region of *HLA-DQA2* and *HLA-DQB1*, and within the intron 1 of *BTNL2*, respectively (Figure 2). *HLA-DQA2* and *HLA-DQB1* encode HLA-DQ  $\alpha$  and  $\beta$  chains, which belong to the HLA class II molecules. HLA class II molecules are expressed in antigen presenting cells (B lymphocytes, dendritic cells and macrophages) and play a central role in the immune system by presenting peptides derived from extracellular proteins [14]. The HLA-DQA2 protein is expressed, but at a very low level in comparison with the HLA-DQA1 protein [15,16]. Moreover,

the HLA-DQA2  $\alpha$  chain does not dimerize with class II  $\beta$  chains [16]. *BTNL2* encodes butyrophilin-like 2, a member of butyrophilin family that shares sequence homology with the B7 co-stimulatory molecules. *BTNL2* regulates T-cell activation through unknown receptor, distinct from CD28 and CTLA-4 [17]. In Japanese population, the haplotype association was more significant than those of respective SNPs (Tables 2 and 3). Therefore, there may be hidden SNP(s) with a lower  $P$  value than rs7775228 and rs10947262, or the haplotype may be implicated in the OA susceptibility. An association of sarcoidosis with rs2076530, a coding SNP on exon 5 of the *BTNL2* gene has been reported [18], but the SNP was not in LD with rs10947262 ( $D' = 0.11$ ,  $r^2 = 0$ ).

The 340-kb region of HLA locus, where the two SNPs are located also includes *HLA-DRA*, *HLA-DRB1*, *HLA-DRB3*, *HLA-DRB4*, *HLA-DRB5* and *HLA-DQA1*. *HLA-DRA*, *HLA-DRB1/3/4/5* and *HLA-DQA1* encode HLA-DR  $\alpha$ ,  $\beta$  and HLA-DQ  $\alpha$  chains, which could also belong to the HLA class II molecules. *HLA-DRB1* is present in all individuals. Allelic variants of *HLA-DRB1* are linked with either none or one of the genes *HLA-DRB3*, *HLA-DRB4* and *HLA-DRB5* [19]. Among these genes, *HLA-DRB1* is strongly associated with RA. Some subtypes of HLA-DRB1 alleles, such as \*0101, \*0401, and \*0405, is associated with RA [20], but not with generalized OA [21].

Although OA has generally been considered a non-inflammatory disease, accumulating evidences suggest that this is not the case. Inflammation involving activated T cells in the synovial membrane of OA patients is well documented [22]. Recently, we identified a genetic variant of *EDG2* gene encoding lysophosphatidic acid receptor associated with knee OA [23]. A GWAS has identified a genetic variant of the *PTGS2* gene encoding cyclooxygenase-2 involved in risk for knee OA [6]. These genetic associations of genes such as *EDG2* and *PTGS2* underscore the potential role of inflammatory pathways in the pathogenesis of knee OA.

Several studies have suggested associations of OA with HLA class I and class II alleles. Study on generalized OA revealed association with HLA A1-B8 in Caucasian [24] and with HLA-Cw4 in Japanese [21]. An association of the HLA-DRB1\*02 alleles with knee and hip OA was identified in a cohort of 106 patients [25]. Interestingly, chondrocyte, which are normally HLA-DP, DQ and DR-negative, become positive for them in OA [26,27], suggesting their function as antigen-presenting cells. Cartilage fragments are mechanically shaved from the joint surface and frequently found in the synovial membrane of OA patients [28]. So, physical interaction between chondrocytes and T cells is conceivable. Peripheral blood T cells from OA patients show significantly higher proliferative responses to autologous chondrocytes [27]. Our results further support the concept that OA is an immunologic disorder.

**Table 3. Haplotype association analysis for knee osteoarthritis susceptibility SNPs, rs7775228 and rs10947262.**

Haplotype <sup>a</sup>	Case	Frequency		Haplotype effect <sup>b</sup>	
		Control	OR (95% CI)	P value	
I TC	0.610	0.541	1.33 (1.20–1.46)	1.48 E-8	
II CT	0.277	0.340	0.74 (0.67–0.83)	5.07 E-8	
III CC	0.077	0.080	0.96 (0.80–1.15)	0.653	
IV TT	0.037	0.040	0.91 (0.71–1.18)	0.475	

OR: odds ratio, CI: confidence interval.

<sup>a</sup>Alleles of rs7775228 (C/T) and rs10947262 (C/T) with this order.

<sup>b</sup>A haplotype vs. all others.

doi:10.1371/journal.pone.0009723.t003

## Supporting Information

**Figure S1** Principal component analysis of GWAS samples. Samples in the GWAS and in HapMap database are analyzed by a program of Smartpca [12], and plotted for the first (X axis) and the second (Y axis) principal component (PC), respectively.

Found at: doi:10.1371/journal.pone.0009723.s001 (0.16 MB TIF)

## Acknowledgments

We thank all participating patients, doctors and staffs of collaborating institutes. We also thank Ms. Yoshie Takamashi, Carmen Pena-Pena and technical staffs of Laboratories for Genotyping Development and for Bone and Joint Diseases of CGM, RIKEN, and Laboratorio Investigación 10 and Rheumatology Unit, Hospital Clínico Universitario de Santiago for genotyping.

### Author Contributions

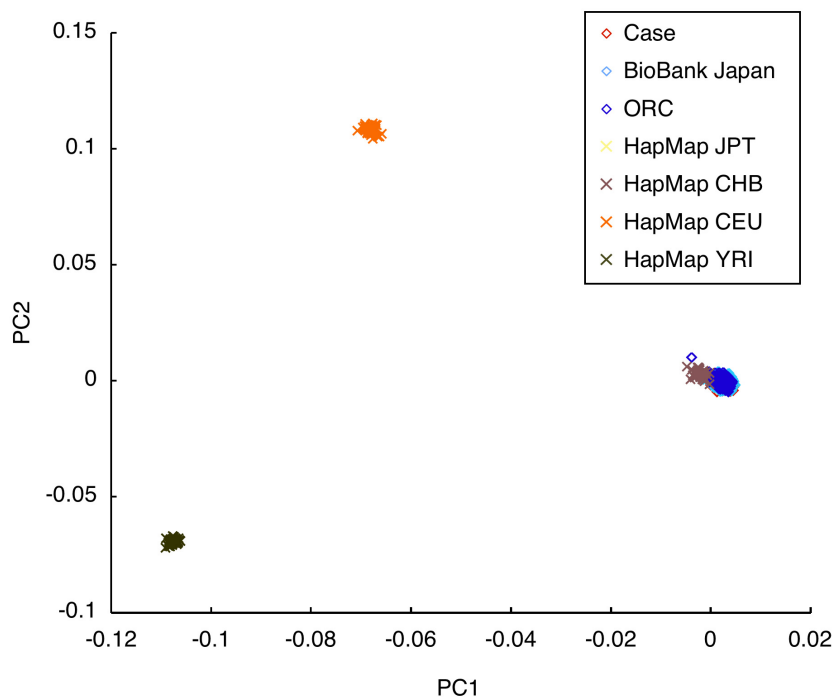
Wrote the paper: MN SI. Planned and supervised the whole project: SI. Performed the Japanese association study: MN IK TF JD. Managed the European association study: A. Tsezou AG. Helped with statistic analysis:

A. Takahashi NK TT. Managed DNA sample and/or clinical information: CR-F JGG-R AS AU NF KNM A. Tsezou AG YN. Oversaw a genotyping of GWAS: MK YN.

### References

- Ikegawa S (2007) New gene associations in osteoarthritis: what do they provide, and where are we going? *Curr Opin Rheumatol* 19: 429–434.
- Dai J, Ikegawa S (2010) Recent advances in association studies of osteoarthritis susceptibility genes. *J Hum Genet* 55: 77–80.
- Mototani H, Mabuchi A, Saito S, Fujioka M, Iida A, et al. (2005) A functional single nucleotide polymorphism in the core promoter region of CALM1 is associated with hip osteoarthritis in Japanese. *Hum Mol Genet* 14: 1009–1017.
- Spector TD, Reneland RH, Mah S, Valdes AM, Hart DJ, et al. (2006) Association between a variation in LRRCH1 and knee osteoarthritis: a genome-wide single-nucleotide polymorphism association study using DNA pooling. *Arthritis Rheum* 54: 524–532.
- Miyamoto Y, Shi D, Nakajima M, Ozaki K, Sudo A, et al. (2008) Common variants in DVWA on chromosome 3p24.3 are associated with susceptibility to knee osteoarthritis. *Nat Genet* 40: 994–998.
- Valdes AM, Loughlin J, Timms KM, van Meurs JJ, Southam L, et al. (2008) Genome-wide association scan identifies a prostaglandin-endoperoxide synthase 2 variant involved in risk of knee osteoarthritis. *Am J Hum Genet* 82: 1231–1240.
- Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, et al. (2007) A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. *Nat Genet* 39: 529–533.
- Nakamura Y (2007) The BioBank Japan Project. *Clin Adv Hematol Oncol* 5: 696–697.
- Kizawa H, Kou I, Iida A, Sudo A, Miyamoto Y, et al. (2005) An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. *Nat Genet* 37: 138–144.
- Rodriguez-Lopez J, Pombo-Suarez M, Loughlin J, Tsezou A, Blanco FJ, et al. (2009) Association of a rsSNP in ADAMTS14 to some osteoarthritis phenotypes. *Osteoarthritis Cartilage* 17: 321–327.
- Andrieu N, Goldstein AM (1998) Epidemiologic and genetic approaches in the study of gene-environment interaction: an overview of available methods. *Epidemiol Rev* 20: 137–147.
- Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263–265.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909.
- Alberts B, Johnson A, Lewis J, Raff M, Roberts K, et al. (2007) *Molecular biology of the cell* - Fifth edition. New York: Garland Science, pp 1577–1579.
- Yu LP, Sheehy MJ (1991) The cryptic HLA-DQA2 ("DX alpha") gene is expressed in human B cell lines. *J Immunol* 147: 4393–4397.
- Rudy GB, Lew AM (1997) The nonpolymorphic MHC class II isotype, HLA-DQA2, is expressed on the surface of B lymphoblastoid cells. *J Immunol* 158: 2116–2125.
- Arnett HA, Escobar SS, Gonzalez-Suarez E, Budelsky AL, Steffen LA, et al. (2007) BTNL2, a butyrophilin/B7-like molecule, is a negative costimulatory molecule modulated in intestinal inflammation. *J Immunol* 178: 1523–1533.
- Valentonyte R, Hampe J, Huse K, Rosenstiel P, Albrecht M, et al. (2005) Sarcoidosis is associated with a truncating splice site mutation in BTNL2. *Nat Genet* 37: 357–364.
- Clayton J, Lonjou C, Whittle D (1997) Allele and haplotype frequencies for HLA loci in various ethnic groups. In: Charron D, Fauchet R, Albert E, Bodmer J, Cambon-Thomsen A, et al. (1997) Genetic diversity of HLA, functional and medical implication. EDK, Paris, pp 665–820.
- Evans TI, Han J, Singh R, Moxley G (1995) The genotypic distribution of shared-epitope DRB1 alleles suggests a recessive mode of inheritance of the rheumatoid arthritis disease-susceptibility gene. *Arthritis Rheum* 38: 1754–1761.
- Wakitani S, Imoto K, Mazuka T, Kim S, Murata N, et al. (2001) Japanese generalised osteoarthritis was associated with HLA class I-A study of HLA-A, B, Cw, DQ, DR in 72 patients. *Clin Rheumatol* 20: 417–419.
- Sakkas LI, Platsoucas CD (2007) The role of T cells in the pathogenesis of osteoarthritis. *Arthritis Rheum* 56: 409–424.
- Mototani H, Iida A, Nakajima M, Furuchi T, Miyamoto Y, et al. (2008) A functional SNP in EDG2 increases susceptibility to knee osteoarthritis in Japanese. *Hum Mol Genet* 17: 1790–1797.
- Patrick M, Manhire A, Ward AM, Doherty M (1989) HLA-A, B antigens and alpha 1-antitrypsin phenotypes in nodal generalised osteoarthritis and erosive osteoarthritis. *Ann Rheum Dis* 48: 470–475.
- Moos V, Menard J, Sieper J, Sparmann M, Muller B (2002) Association of HLA-DRB1\*02 with osteoarthritis in a cohort of 106 patients. *Rheumatology (Oxford)* 41: 666–669.
- Lance EM, Kimura LH, Manibog CN (1993) The expression of major histocompatibility antigens on human articular chondrocytes. *Clin Orthop Relat Res* pp 266–282.
- Sakata M, Masuko-Hongo K, Nakamura H, Onuma H, Tsuruha JJ, et al. (2003) Osteoarthritic articular chondrocytes stimulate autologous T cell responses in vitro. *Clin Exp Rheumatol* 21: 704–710.
- Revell PA, Mayston V, Lalor P, Mapp P (1988) 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* 47: 300–307.

Figure S1





# Osteoarthritis and Cartilage



## Common variations in estrogen-related genes are associated with severe large-joint osteoarthritis: a multicenter genetic and functional study

J.A. Riancho †\*, C. García-Ibarbia †, A. Gravani ‡, E.V.A. Raine ‡, C. Rodríguez-Fontenla §, A. Soto-Hermida ||, I. Rego-Perez ||, A.W. Dodd ‡, J.J. Gómez-Reino §, M.T. Zarrabeitia †, C.M. Garcés †, A. Carr ¶, F. Blanco ||, A. González §, J. Loughlin ‡

† Servicios de Medicina Interna, Medicina Legal and Cirugía Ortopédica, Hospital U.M. Valdecilla-IFIMAV, Universidad de Cantabria, RETICEF, Santander, Spain

‡ Newcastle University, Institute of Cellular Medicine, Musculoskeletal Research Group, Newcastle upon Tyne, UK

§ Laboratorio Investigación 10 and Unidad de Reumatología, Instituto de Investigación Sanitaria-Hospital Clínico Universitario de Santiago, Santiago de Compostela, Spain

|| INIBIC-Complejo Hospitalario Universitario A Coruña, División de Reumatología, A Coruña, Spain

¶ University of Oxford, NDORMS, Oxford, UK

### ARTICLE INFO

#### Article history:

Received 26 January 2010

Accepted 14 April 2010

#### Keywords:

Osteoarthritis

Estrogens

Aromatase

Gene expression

Genetic association study

Multicenter study

### SUMMARY

**Objective:** Several lines of evidence suggest that estrogens influence the development of osteoarthritis (OA). The aim of this study was to explore the association of two common polymorphisms within the aromatase (*CYP19A1*) and estrogen receptor (ER) alpha (*ESR1*) genes with severe OA of the lower limbs. **Methods:** The rs1062033 (*CYP19A1*) and rs2234693 (*ESR1*) single nucleotide polymorphisms were genotyped in 5528 individuals (3147 patients with severe hip or knee OA, and 2381 controls) from four centres in Spain and the United Kingdom. Gene expression was measured in femoral bone samples from a group of patients.

**Results:** In the global analysis, both polymorphisms were associated with OA, but there was a significant sex interaction. The GG genotype at rs1062033 was associated with an increased risk of knee OA in women [odds ratio (OR) 1.23;  $P = 0.04$ ]. The CC genotype at rs2234693 tended to be associated with reduced OA risk in women (OR 0.76,  $P = 0.028$ , for knee OA; OR = 0.84,  $P = 0.076$  for hip OA), but with increased risk of hip OA in men (OR 1.28;  $P = 0.029$ ). Women with unfavourable genotypes at both loci had an OR of 1.61 for knee OA ( $P = 0.006$ ). The rs1062033 genotype associated with higher OA risk was also associated with reduced expression of the aromatase gene in bone.

**Conclusions:** Common genetic variations of the aromatase and ER genes are associated with the risk of severe OA of the large joints of the lower limb in a sex-specific manner. These results are consistent with the hypothesis that estrogen activity may influence the development of large-joint OA.

© 2010 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

### Introduction

Osteoarthritis (OA) is a prevalent disorder that frequently impairs the quality of life in middle aged and elderly individuals and represents a considerable economical burden to health-care systems. Among individuals over 45 years of age, several studies have reported a prevalence of hip OA of between 7 and 27%, and of approximately 16% for knee OA<sup>1</sup>. Joint replacement surgery, mostly at the hip and the knee, represents a major component of the activity of orthopaedic surgery departments, with the number of

operations expected to increase six-fold by the year 2030<sup>2</sup>. Therefore, a better understanding of the pathophysiology of OA is important to establish effective preventive strategies. OA is currently seen as a complex disorder resulting from the interaction between acquired and genetic factors. The heritability of OA at different sites has been estimated to be approximately 75% at the spine, 65% at the hand, 60% at the hip, and 40% at the knee, with some suggestions for a more important contribution of heredity in females than in males<sup>3,4</sup>. Although the damage to the cartilage is frequently emphasized in OA, there is evidence suggesting that changes to the synovium and the subchondral bone are also involved in the OA disease process<sup>5</sup>.

Estrogens are critical modulators of bone homeostasis, in females and in males<sup>6</sup>. They have also been shown to modulate chondrocyte activity and the synthesis of a variety of factors, including metalloproteinases, nitric oxide and reactive oxygen

\* Address correspondence and reprint requests to: J.A. Riancho, Departamento de Medicina Interna, Hospital U.M. Valdecilla, Universidad de Cantabria, Av Valdecilla s/n, 39008 Santander, Spain. Fax: 34-942201695.

E-mail address: rianchoj@unican.es (J.A. Riancho).

species, involved in the anabolism and catabolism of the cartilage matrix<sup>7–10</sup>. Estrogens act through the binding to two types of specific estrogen receptors (ERs), encoded by different genes:  $\alpha$  (or *ESR1*) and  $\beta$  (or *ESR2*). *ESR1* seems to have a more consistent role in both sexes<sup>6</sup>. ERs are expressed by a variety of cells in the skeleton, including stromal cells, osteoblasts and chondrocytes<sup>11</sup>. ER genes have been considered as appealing candidates potentially contributing to the genetic component of OA, but conflicting results have so been reported [recently reviewed by Ryder *et al.*<sup>12</sup>]. However, various studies suggest that estrogens may indeed influence the development of OA<sup>13</sup>; low levels of estrogens have been associated with an increased risk of OA in some studies<sup>14</sup>, whilst estrogen replacement therapy may have a protective role on the development of OA, both in experimental animals and in humans<sup>15–18</sup>. In fact, increased osteophytosis has been reported in ER knock-out mice<sup>19</sup>. However, the protective effect of estrogens has not been demonstrated in all studies, and conflicting results have been reported in primate models of OA<sup>17,20,21</sup>.

In postmenopausal women and in men, the aromatization of androgenic precursors in peripheral tissues is the main source of estrogens. The reaction is catalyzed by aromatase, the product of the *CYP19A1* gene, located on chromosome 15 (15q21.1). The synthesis of estrogens by aromatase-expressing bone cells may have important effects at the skeletal tissue level, independent of the circulating levels of estrogens. In fact, we have recently shown that aromatase expression is reduced in the bone tissue of patients with severe hip OA, in comparison to patients with hip fractures<sup>22</sup>. Therefore, we hypothesized that allelic variants of genes influencing estrogen synthesis and/or response could influence the risk of OA. The objective of this study was to explore the association of two common polymorphisms of the aromatase and *ER $\alpha$*  genes with primary OA of the lower limb.

## Methods

### Subjects

We studied 5528 individuals aged 55 years and older (2176 patients with hip OA, 971 patients with knee OA and 2381 controls) recruited at three centres in Spain (Santander, Santiago and Coruña) and one centre in the United Kingdom, Oxford (Table 1). The Spanish cohorts included Caucasian individuals of Spanish ancestry living in two regions in Northern Spain (Cantabria, Santander cohort; or Galicia, Santiago and Coruña cohorts). The Oxford cohort comprised individuals of Caucasian ethnicity from Oxford and from other parts of the UK. Patients had severe, radiographically confirmed, primary OA requiring hip or knee joint replacement surgery. Exclusion criteria included individuals with secondary OA (due to systemic diseases, rheumatoid arthritis, infections, trauma, ischemic necrosis, neurological diseases, etc.), who were excluded by clinical, laboratory and radiographic studies. The control group included volunteers recruited by voice and

written announcements, patient's spouses, and individuals taking part in a cohort study of osteoporosis risk factors. They had no clinically evident OA at the lower extremities, but X-rays were not obtained routinely in the absence of a clinical justification. The study was approved by local ethics committees and all individuals gave informed written consent.

### Genotyping

DNA was extracted from the peripheral blood or buccal swabs using standard procedures. Two single nucleotide polymorphisms (SNPs) were genotyped; the rs1062033 SNP located in the 5' untranslated region (5'UTR) of the aromatase gene *CYP19A1* and the rs2234693 SNP located in intron 1 of the *ER $\alpha$*  gene *ESR1*. rs1062033 has been shown to be associated with differences in gene transcription at *CYP19A1* and with BMD in postmenopausal women<sup>23,24</sup>, whilst rs2234693 has been suggested to influence OA risk in some studies<sup>25–27</sup>. In the Santander and Oxford cohorts the SNPs were genotyped by TaqMan allelic discrimination assays (Applied Biosystems, Foster City, CA, USA), as previously described<sup>23</sup>. In the Santiago and Coruña cohorts they were genotyped by a single-base extension procedure, involving a multiplex polymerase chain reaction (PCR) (Qiagen Multiplex PCR, Valencia, CA, USA; oligonucleotide primer sequences and PCR conditions are available upon request) followed by single-base extension reactions performed with the SNaShot Multiplex Kit (Applied Biosystems). In each laboratory, 5% samples were analyzed at least twice in different days to confirm the genotyping reproducibility. Twelve samples were genotyped in different laboratories to check the consistency of results with both genotyping methods. Alleles were designated according to the human reference sequences (minus strand for the *CYP19A1* gene, plus strand for the *ESR1* gene; Genome build 36.3).

### Gene expression

*CYP19A1* and *ESR1* expression was studied by real-time quantitative PCR, as previously described<sup>22</sup>. Biopsy samples were obtained from the femoral head removed during joint replacement for hip OA. Small fragments of trabecular bone were extensively washed with PBS, snap-frozen in liquid nitrogen and stored at  $-70^{\circ}\text{C}$ . Unthawed fragments were ground to a powder using a tissue homogenizer into Trizol (Invitrogen) to extract RNA, whose integrity was then monitored by gel electrophoresis and whose concentration was measured by absorbometry. Aliquots of RNA (approximately 1  $\mu\text{g}$ ) were reverse transcribed with the Superscript III kit (Invitrogen), using random hexamers as primers. In negative control reactions reverse transcriptase was omitted. After reverse transcription, the expression of *CYP19A1* and *ESR1* was determined by real-time PCR in an ABI7300 apparatus (Applied Biosystems). The reactions were performed in triplicate in 96-well plates containing aliquots of cDNA obtained by reverse transcription, 5  $\mu\text{l}$  of universal PCR master mix, and specific primers and probes for *CYP19A1* and *ESR1* (TaqMan Gene Expression Assays, Applied Biosystems).

The cycle threshold (Ct) was determined. This represents the cycle at which a significant increase in fluorescence is first detected and is inversely related to the amount of target cDNA in the starting material. The results were normalized to *TBP* (TATA box binding protein) expression analyzed in the same reaction plate. Control samples of reverse transcribed reference human RNA (Stratagene, La Jolla, CA) were also included. The relative RNA levels were calculated using the formula  $10 \times 2^{\Delta\text{Ct}1 - \Delta\text{Ct}2}$ , where  $\Delta\text{Ct}1$  is the difference between the control cDNA Ct and the sample cDNA Ct when the target gene (*CYP19A1*, *ESR1*) is amplified, and  $\Delta\text{Ct}2$  is the

**Table 1**  
Characteristics of study subjects

	Santander	Coruña	Santiago	Oxford
Individuals (n)	1433	496	1014	2585
Controls (n)	802	244	473	862
Hip OA (n)	359	252	287	1278
Knee OA (n)	272	–	254	445
Age (mean, SD)				
Controls	71 $\pm$ 10	65 $\pm$ 13	68 $\pm$ 9	69 $\pm$ 7
Hip OA	71 $\pm$ 7	67 $\pm$ 14	68 $\pm$ 5	65 $\pm$ 6
Knee OA	72 $\pm$ 7	–	68 $\pm$ 6	64 $\pm$ 5
Sex (% women)	61	62	55	56



**Table II**  
Genotype and allele frequencies in males and females (percentages in parentheses)

Cohort	Group	rs1062033				rs2234693			
		CC/CG	GG	C	G	TT/CT	CC	T	C
<i>Males</i>									
Santander	Control	182 (64)	103 (36)	0.36	0.64	229 (80)	56 (20)	0.55	0.45
	Knee OA	58 (61)	37 (39)	0.36	0.64	79 (83)	16 (17)	0.58	0.42
	Hip OA	109 (61)	71 (39)	0.35	0.65	133 (74)	47 (26)	0.50	0.50
Santiago	Control	190 (64)	105 (36)	0.34	0.66	250 (85)	45 (15)	0.62	0.38
	Knee OA	22 (47)	25 (53)	0.37	0.63	34 (72)	13 (28)	0.48	0.52
	Hip OA	76 (69)	34 (31)	0.31	0.69	91 (83)	19 (17)	0.59	0.41
Coruña	Control	67 (69)	30 (31)	0.46	0.54	79 (81)	18 (19)	0.56	0.44
	Knee OA	—	—	—	—	—	—	—	—
	Hip OA	63 (70)	27 (30)	0.46	0.54	72 (80)	18 (20)	0.59	0.41
Oxford	Control	330 (70)	141 (30)	0.46	0.54	385 (82)	86 (18)	0.55	0.45
	Knee OA	125 (71)	51 (29)	0.49	0.51	144 (82)	32 (18)	0.54	0.46
	Hip OA	363 (72)	140 (28)	0.47	0.53	395 (79)	108 (21)	0.56	0.44
<i>Females</i>									
Santander	Control	358 (69)	159 (31)	0.43	0.57	394 (76)	123 (24)	0.52	0.48
	Knee OA	107 (61)	70 (39)	0.40	0.60	140 (79)	37 (21)	0.53	0.47
	Hip OA	114 (64)	65 (36)	0.40	0.60	146 (82)	33 (14)	0.57	0.43
Santiago	Control	129 (73)	49 (27)	0.44	0.56	143 (80)	35 (20)	0.57	0.43
	Knee OA	132 (64)	75 (36)	0.43	0.57	174 (84)	33 (16)	0.55	0.45
	Hip OA	108 (61)	69 (39)	0.40	0.60	145 (82)	32 (18)	0.58	0.42
Coruña	Control	95 (65)	52 (35)	0.40	0.60	116 (79)	31 (21)	0.55	0.45
	Knee OA	—	—	—	—	—	—	—	—
	Hip OA	116 (72)	46 (28)	0.48	0.52	130 (80)	32 (20)	0.57	0.43
Oxford	Control	281 (72)	110 (28)	0.47	0.53	310 (79)	81 (21)	0.55	0.45
	Knee OA	188 (70)	81 (30)	0.46	0.54	224 (83)	45 (17)	0.56	0.44
	Hip OA	535 (69)	240 (31)	0.44	0.56	626 (81)	149 (19)	0.57	0.43

difference between control cDNA Ct and the sample cDNA Ct when the control gene (*TBP*) is amplified.

#### Data analysis

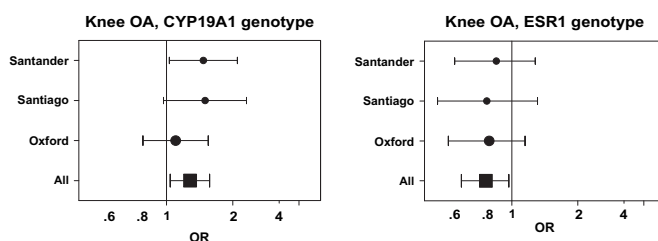
Hardy–Weinberg equilibrium in controls was tested with HWSIM software (<http://krunch.med.yale.edu/hwsim/hwsim.doc>). Differences in genotype distribution in each cohort and in the combined study population were analyzed with the  $\chi^2$  test and by unconditional logistic regression analysis after genotype grouping. As measures of association, the odds ratios (ORs) and their 95% confidence intervals (CIs) were calculated for each cohort. The global OR combining all cohorts and the adjusted Mantel–Haenszel OR (weighted by each stratum variance) were calculated as summary estimates of effect. Assuming a minor allele frequency of 0.4, the study power to detect an effect size associated with an OR of at least 1.4 was higher than 90%, in the global and the sex-

stratified analysis, under a log-additive model. The non-parametric Jonckheere–Terpstra test for ordered categories was used to analyze the differences in gene expression between subjects with different genotypes. SPSS (SPSS Inc, Chicago, IL, USA) and EPIDAT software (<http://dxsp.sergas.es/ApliEdatos/Epidat/cas/default.asp>) were used in the statistical analyses. Study power calculations were done with QUANTO (<http://hydra.usc.edu/gxe>).

#### Results

##### Genetic variations and OA

In the whole study group, the genotype frequencies were similar to those reported in other Caucasian populations. The *CYP19A1* SNP rs1062033 genotypes were: CC 20%, CG 48%, and GG 32%. The *ESR1* SNP rs2234693 genotype frequencies were: TT 30%, TC 50%, CC 20%. There was no evidence for departure from the Hardy–Weinberg



**Fig. 1.** Association of *CYP19A1* SNP rs1062033 GG genotype and *ESR1* SNP rs2234693 CC genotype with knee OA in women. The individual ORs, the global combined OR and the 95% CIs are shown.

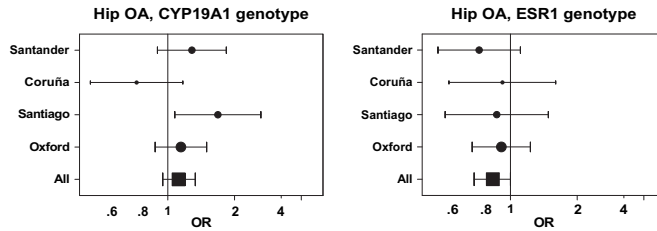


Fig. 2. Association of *CYP19A1* SNP rs1062033 GG genotype and *ESR1* SNP rs2234693 CC genotype with hip OA in women. The individual ORs, the global combined OR and the 95% CIs are shown.

equilibrium in the control groups ( $P > 0.1$ ). The analyses of the unstratified data did not show significant differences in genotype distributions between cases and controls for either SNP. However, in sex-adjusted logistic regression models the genotypes at both SNPs were associated with OA (rs1062033,  $P = 0.023$ ; rs2234693,  $P = 0.03$ ), and there was a significant interaction between sex and the genotypes ( $P = 0.04$  for *CYP19A1* and  $0.006$  for *ESR1*). Therefore, all subsequent analyses were stratified by sex. A dominant model for rs1062033 and a recessive model for rs2234693 best fitted the data. The genotype frequency distributions in the study cohorts are shown in Table II.

In women there was a consistent trend towards an increased risk of knee OA in those bearing the rs1062033 GG genotype. The global OR was 1.28 (95% CI 1.04–1.57;  $P = 0.020$ ) and the combined Mantel–Haenszel OR was 1.32 (95% CI 1.06–1.63;  $P = 0.011$ ), without evidence of heterogeneity between the three study populations ( $P = 0.41$ ). On the other hand, women with the CC genotype at rs2234693 had a lower risk of knee OA, with a global OR of 0.76 (0.59–0.97;  $P = 0.028$ ) and combined Mantel–Haenszel OR of 0.80 (0.62–1.02;  $P = 0.08$ ), without evidence of heterogeneity ( $P = 0.9$ ) (Fig. 1). Both SNPs were independently associated with the risk of knee OA, and there was no significant interaction between them. In the male subgroup, there were no significant associations between knee OA and either SNP (not shown).

Similarly, the GG genotype at rs1062033 was associated with a nonsignificant trend for increased risk of hip OA in women (OR 1.12, 95% CI 0.95–1.33,  $P = 0.18$ ; Mantel–Haenszel OR 1.18, 95% CI 0.98–1.41,  $P = 0.075$ ), with relatively wide between-study differences (heterogeneity  $P$ -value = 0.08; proportion of between-study variance to total variance  $I^2 = 0.5$ ) (Fig. 2). As for knee OA, there was a consistent trend for lower risk of hip OA in women bearing the CC genotype at rs2234693, with a global OR of 0.84, almost reaching the conventional threshold for statistical significance (95% CI 0.69–1.02,  $P = 0.07$ ; Mantel–Haenszel OR 0.86, 95% CI 0.70–1.06,

$P = 0.16$ ; heterogeneity  $P$ -value 0.8). In the male subgroup rs1062033 was not associated with hip OA. However, unlike women, men with the CC genotype at rs2234693 had a higher risk of hip OA, with a global OR of 1.28 (1.02–1.59;  $P = 0.029$ ; Mantel–Haenszel OR 1.25; 95% CI 1.00–1.57,  $P = 0.048$ ; heterogeneity  $P$ -value 0.9) (Fig. 3).

As shown in Table III, when the genotypes at both loci were combined, in women the gradient risk was somewhat higher than with any individual genotype, particularly for knee OA, with ORs between 1.24 and 1.61 for knee OA. In men the combined analysis did not generate more significant data than when the SNPs were analyzed individually (not shown).

#### *CYP19A1* and *ESR1* expression

Gene transcription of *CYP19A1* and *ESR1* was studied in bone samples of 49 patients undergoing hip arthroplasty (29 men, 20 women) and the results were analyzed according to the donor's genotype. rs1062033 was associated with differences in the abundance of *CYP19A1* transcripts, which were more abundant in samples from individuals bearing C alleles than in those bearing G alleles (upper panel, Table IV). Intriguingly, significant differences in *CYP19A1* transcript abundance between individuals with different genotypes of rs2234693 were also observed (lower panel, Table IV). However, neither the rs2234693 genotypes nor the rs1062033 genotypes were associated with differences in the amount of *ESR1* transcripts in the bone samples (Table V).

#### Discussion

Estrogens are critical in the acquisition and maintenance of bone mass, in women as well as in men<sup>28,29</sup>, and estrogen deficiency plays a major role in involuntional osteoporosis. The possible role of estrogens in OA is less well established. However, several

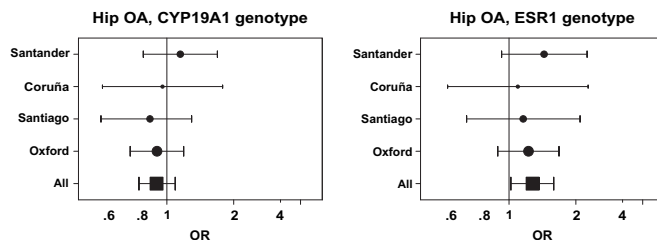


Fig. 3. Association of *CYP19A1* SNP rs1062033 GG genotype and *ESR1* SNP rs2234693 CC genotype with hip OA in men. The individual ORs, the global combined OR and the 95% CIs are shown.

**Table III**  
Combined analysis of rs1062033 and rs2234693 genotypes in women

rs1062033 rs2234693	No-GG CC	No-GG No-CC	GG CC	GG No-CC	P-value	P-value for trend
Controls (n)	193	670	77	293		
Knee OA (n)	77	350	38	188		
OR	1	1.31 (0.97–1.75)	1.24 (0.77–1.98)	1.61 (1.16–2.22)	0.031	0.006
Hip OA (n)	156	717	90	330		
OR	1	1.32 (1.05–1.68)	1.45 (1.00–2.09)	1.39 (1.07–1.81)	0.063	0.051

experimental and epidemiological studies suggest that estrogen deficiency may favour the development of this disease<sup>30</sup>. For instance, Sowers et al found an inverse relationship between serum levels of estradiol and the incidence of knee OA, even after adjustment for age and BMI<sup>14</sup>, but others did not confirm these results<sup>31</sup>. Likewise, conflicting results have been published regarding the possible preventive effect of hormone replacement therapy on OA<sup>32,15</sup>.

Such conflicting results are not totally unexpected, as the local synthesis of estrogens in bone may have important effects in nearby tissue that are not captured by measuring circulating serum estradiol, whilst a single point measurement may not reflect adequately the overall exposure of skeletal cells to estrogens during the lifetime period. In this regard, it has been suggested that studies of genetic variants modulating the synthesis or activity of biological factors may be more informative than single measurements or even short interventional studies, as they originate the so-called “Mendelian randomization”<sup>33,34</sup>. In this multicenter study we found that two common polymorphisms related to the estrogen pathway are associated with the risk of OA. Thus, women bearing a particular genotype at the aromatase-coding gene *CYP19A1* are at an increased risk for knee OA. The effect size was relatively small and not significant in individual cohorts, but it was consistent and statistically significant in the global analysis, with an OR greater than 1.2. The risk genotype was common, being present in about one third of the individuals genotyped. Therefore, it may have important consequences not only at the individual level, but also at the population level. This risk genotype was associated with reduced *CYP19A1* expression in bone samples from patients with OA, thus confirming previous results obtained in patients with fractures<sup>35,24</sup>. We do not have data on estrogen levels, but we speculate that individuals with genotypes associated with reduced *CYP19A1* expression may have lower levels of aromatase and have therefore had a lower cumulative exposure to estrogens during their life. This would be consistent with the hypothesis that a relative estrogen deficiency, either local or systemic, may favour the development of OA<sup>13,36</sup>, and with a recent report of reduced aromatase expression in OA<sup>22</sup>.

Aromatase is responsible for the synthesis of estrogens in postmenopausal women and in men<sup>37</sup>. Its important role is emphasized by studies showing a marked decline in bone mass following therapy with aromatase inhibitors<sup>38</sup>. On the other hand, type 1 estrogen receptors (ER $\alpha$ ) are regarded as the main receptors transducing the

effect of estrogens on target cells. Based on previous results<sup>23,35,39</sup>, we hypothesized that genetic variations of ER could also modulate the risk of OA. In fact, although to our knowledge the association of aromatase gene polymorphisms with OA has not been reported previously, several investigators have explored the relationship between *ESR1* genetic variants and OA<sup>25–27,40,41</sup>. In common with several of these studies, we found that the *ESR1* SNP rs2234693 was also associated with the risk of OA in women, particularly at the knee (results for hip OA did not reach the conventional 0.05 statistical significance threshold) and with hip OA in men. The polymorphism is located in intron 1 and we speculated that it might influence gene transcription. However, we found no significant differences in the abundance of *ESR1* transcripts across the three *ESR1* genotypes. Interestingly, *ESR1* genotypes were associated with statistically significant differences in *CYP19A1* transcript abundance. There is no clear explanation for this result and it obviously merits further genetic and functional investigation; a relevant observation from other studies is that ER may have ligand-independent and ligand-dependent effects on the expression of genes including *CYP19A1*<sup>42</sup>, suggesting a trans-interaction between these two genes located in different chromosomes. Whatever the explanation might be, our results suggest that the alleles associated with lower aromatase levels are associated with an increased risk of OA. They do not prove a direct causal link between the genotypes and OA, but suggest that genotype-dependent differences in estrogen levels or action influence OA risk.

Although aromatase is critical for estrogen synthesis in both postmenopausal women and men, we observed an interaction between sex and the genotypes. Thus, whereas both rs1062033 and rs2234693 genotypes were associated with knee OA in women, in men rs1062033 genotypes were not associated with either hip or knee OA, and rs2234693 genotypes were associated with hip OA, but in an opposite direction to that observed in women. The exact mechanisms explaining the sex-related differences are unclear. They may be the consequence of differences in the relative importance of acquired factors (such as physical activity) in men and women, or reflect some sex-specific responses to changes in estrogen levels, similar to what is observed in aromatase knock-out mice and in experimental models of OA<sup>43,44</sup>.

This multicenter study included a large number of individuals, which allowed us to increase the precision of the estimates of the genetic effects. However, it had several limitations. First, we did not systematically obtain X-rays of the control subjects. Therefore,

**Table IV**  
*CYP19A1* gene expression according to *CYP19A1* and *ESR1* genotypes. Values are means  $\pm$  SEM (arbitrary units) and 95% CIs in parentheses

<i>CYP19A1</i> genotype	CG (n = 21)	GG (n = 21)	P-value
CC (n = 8)	15.5 $\pm$ 3.4 (8.7–22.3)	11.4 $\pm$ 2.9 (5.8–17.0)	0.036
<i>ESR1</i> genotype			
TT (n = 14)	16.7 $\pm$ 3.1 (10.6–22.8)	23.6 $\pm$ 9.2 (5.6–41.6)	0.011
TC (n = 28)			

**Table V**  
*ESR1* gene expression according to *CYP19A1* and *ESR1* genotypes (data on 37 subjects). Values are means  $\pm$  SEM (arbitrary units) and 95% CIs in parentheses

<i>CYP19A1</i> genotype	CG (n = 16)	GG (n = 17)	P-value
CC (n = 4)	13.2 $\pm$ 6.2 (1.1–25.3)	5.1 $\pm$ 1.4 (2.3–7.9)	0.22
<i>ESR1</i> genotype			
TT (n = 13)	11.8 $\pm$ 5.6 (0.9–22.7)	4.4 $\pm$ 2.1 (2.7–6.1)	0.90
TC (n = 18)			

some of them may have mild asymptomatic OA, which could bias the results towards the null effect. On the other hand, although we excluded cases with secondary OA, we had limited information about anthropometric, nutritional and other environmental factors that may influence the risk of OA. Therefore, we could not explore possible interactions between genetic and acquired factors that can be important in OA.

In conclusion, in this multicenter study we have shown that common genetic variation of the aromatase gene *CYP19A1* and the ER gene *ESR1*, which are associated with gene expression, are also associated with the risk of severe OA of the large joints of the lower limb. The sex-stratified analysis suggested that the influence of the *CYP19A1* SNP is more important in women than in men, and in knee OA than in hip OA. On the other hand, the *ESR1* SNP may have a sex-specific influence on OA risk. Overall, these results are consistent with the hypothesis that estrogen activity may influence the development of large-joint OA.

#### Author contributions

Conception and design: JAR, JL.  
 Provision of patients and samples: CMG, JJGR, AC, AP, IRP.  
 Clinical and laboratory data acquisition: CGI, AG, EAR, CRF, FJB, IRP AD, JJGR, CMG, AC.  
 Genotyping supervision: MTZ, AG, FB, JL, IRP.  
 Data analysis and interpretation: JAR, JL, AG, FB.  
 Drafting of the article: JAR, JL, AG, FB.  
 Final approval of the article: all coauthors.  
 Obtaining of funding: JAR, JL, AG, FB.  
 Study coordination and responsibility for the integrity of the work: JAR.

#### Conflicts of interest

Authors do not have conflicts of interest relevant to this paper.

#### Acknowledgements

We acknowledge the technical assistance of Verónica Mijares, Carolina Sañudo and Carmen Pena-Pena. We thank Bridget Watkins and Kim Cliphsham for their assistance in the collection of the Oxford cohort.

**Funding sources:** this study was supported in part by grants from the Fondo de Investigaciones Sanitarias-ISCIII (PI 06/0034 and PI08/2028), by the Arthritis Research UK, by the UK NIHR Biomedical Research Centre for Ageing and Age-related disease award to the Newcastle upon Tyne Hospitals NHS Foundation Trust, by the Dr William Harker Foundation, by the UK NIHR Musculoskeletal Biomedical Research Unit in Oxford, by the Collisson and Jean Shanks Foundations and by grants from the Xunta de Galicia. C. R-F is the recipient of a predoctoral contract of the ISCIII.

#### References

- Zhang Y, Jordan JM. Epidemiology of osteoarthritis. *Rheum Dis Clin North Am* 2008;34:515–29.
- Bitton R. The economic burden of osteoarthritis. *Am J Manag Care* 2009;15:S230–5.
- Loughlin J. Genetic epidemiology of primary osteoarthritis. *Curr Opin Rheumatol* 2001;13:111–6.
- Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. *Osteoarthritis Cartilage* 2004;12(Suppl A):S39–44.
- Goldring MB, Goldring SR. Osteoarthritis. *J Cell Physiol* 2007;213:626–34.
- Syed FS, Khosla S. Mechanisms of sex steroid effects on bone. *Biochem Biophys Res Commun* 2005;328:688–96.
- Sniekers YH, van Osch GJ, Jahr H, Weinans H, van Leeuwen JP. Estrogen modulates iodoacetate-induced gene expression in bovine cartilage explants. *J Orthop Res* 2009.
- Richette P, Dumontier MF, Tahiri K, Widerak M, Torre A, Benallaoua M, et al. Estrogens inhibit interleukin-1{beta}-mediated nitric oxide synthase expression in articular chondrocytes through NF-κB impairment. *Ann Rheum Dis* 2007;66:345–50.
- Claassen H, Schluter M, Schunke M, Kurz B. Influence of 17beta-estradiol and insulin on type II collagen and protein synthesis of articular chondrocytes. *Bone* 2006;39:310–7.
- Claassen H, Schunke M, Kurz B. Estradiol protects cultured articular chondrocytes from oxygen-radical-induced damage. *Cell Tissue Res* 2005;319:439–45.
- Oshima Y, Matsuda K, Yoshida A, Watanabe N, Kawata M, Kubo T. Localization of estrogen receptors alpha and beta in the articular surface of the rat femur. *Acta Histochem Cytochem* 2007;40:27–34.
- Ryder JJ, Garrison K, Song F, Hooper L, Skinner J, Loke Y, et al. Genetic associations in peripheral joint osteoarthritis and spinal degenerative disease: a systematic review. *Ann Rheum Dis* 2008;67:584–91.
- Richette P, Corvol M, Bardin T. Estrogens, cartilage, and osteoarthritis. *Joint Bone Spine* 2007;70:257–62.
- Sowers MR, McConnell D, Jannausch M, Buyuktur AG, Hochberg M, Jamadar DA. Estradiol and its metabolites and their association with knee osteoarthritis. *Arthritis Rheum* 2006;54:2481–7.
- 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.
- Ravn P, Warming L, Christgau S, Christiansen C. The effect on cartilage of different forms of application of postmenopausal estrogen therapy: comparison of oral and transdermal therapy. *Bone* 2004;35:1216–21.
- Ham KD, Loeser RF, Lindgren BR, Carlson CS. Effects of long-term estrogen replacement therapy on osteoarthritis severity in cynomolgus monkeys. *Arthritis Rheum* 2002;46:1956–64.
- Ma HL, Blanchet TJ, Peluso D, Hopkins B, Morris EA, Glasson SS. Osteoarthritis severity is sex dependent in a surgical mouse model. *Osteoarthritis Cartilage* 2007;15:695–700.
- Sniekers YH, van Osch GJ, Ederveen AG, Inzunza J, Gustafsson JA, van Leeuwen JP, et al. Development of osteoarthritic features in estrogen receptor knockout mice. *Osteoarthritis Cartilage* 2009;17:1356–61.
- Olson EJ, Lindgren BR, Carlson CS. Effects of long-term estrogen replacement therapy on the prevalence and area of periarticular tibial osteophytes in surgically postmenopausal cynomolgus monkeys. *Bone* 2007;41:282–9.
- Sniekers YH, Weinans H, Bierma-Zeinstra SM, van Leeuwen JP, van Osch GJ. Animal models for osteoarthritis: the effect of ovariectomy and estrogen treatment – a systematic approach. *Osteoarthritis Cartilage* 2008;16:533–41.
- Hernandez JL, Garcés CM, Sumillera M, Fernandez-Aldasoro EV, Garcia-Ibarbia C, Ortiz JA, et al. Aromatase expression in osteoarthritic and osteoporotic bone. *Arthritis Rheum* 2008;58:1696–700.
- Riancho JA, Zarrabeitia MT, Valero C, Sanudo C, Mijares V, Gonzalez-Macias J. A gene-to-gene interaction between aromatase and estrogen receptors influences bone mineral density. *Eur J Endocrinol* 2006;155:53–9.
- Riancho JA, Sanudo C, Valero C, Pipaon C, Olmos JM, Mijares V, et al. Association of the aromatase gene alleles with BMD: epidemiological and functional evidence. *J Bone Miner Res* 2009;24:1709–18.

25. Ushiyama T, Ueyama H, Inoue K, Nishioka J, Ohkubo I, Hukuda S. Estrogen receptor gene polymorphism and generalized osteoarthritis. *J Rheumatol* 1998;25:134–7.
26. Lian K, Lui L, Zmuda JM, Nevitt MC, Hochberg MC, Lee JM, *et al.* Estrogen receptor alpha genotype is associated with a reduced prevalence of radiographic hip osteoarthritis in elderly Caucasian women. *Osteoarthritis Cartilage* 2007;15:972–8.
27. Bergink AP, van Meurs JB, Loughlin J, Arp PP, Fang Y, Hofman A, *et al.* Estrogen receptor alpha gene haplotype is associated with radiographic osteoarthritis of the knee in elderly men and women. *Arthritis Rheum* 2003;48:1913–22.
28. Khosla S, Melton LJ, Riggs BL. Estrogen and the male skeleton. *J Clin Endocrinol Metab* 2002;87:1443–50.
29. Riggs BL, Khosla S, Melton III LJ. Sex steroids and the construction and conservation of adult skeleton. *Endocr Rev* 2002;23:279–302.
30. Roman-Blas JA, Castaneda S, Largo R, Herrero-Beaumont G. Osteoarthritis associated with estrogen deficiency. *Arthritis Res Ther* 2009;11:241.
31. de Klerk BM, Schiphof D, Groeneveld FP, Koes BW, van Osch GJ, van Meurs JB, *et al.* No clear association between female hormonal aspects and osteoarthritis of the hand, hip and knee: a systematic review. *Rheumatology (Oxford)* 2009;48:1160–5.
32. de Klerk BM, Schiphof D, Groeneveld FP, Koes BW, van Osch GJ, van Meurs JB, *et al.* Limited evidence for a protective effect of unopposed oestrogen therapy for osteoarthritis of the hip: a systematic review. *Rheumatology (Oxford)* 2009;48:104–12.
33. Casas JP, Bautista LE, Smeeth L, Sharma P, Hingorani AD. Homocysteine and stroke: evidence on a causal link from Mendelian randomisation. *Lancet* 2005;365:224–32.
34. Smith GD, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol* 2004;33:30–42.
35. Valero C, Perez-Castrillon JL, Zarrabeitia MT, Hernandez JL, Alonso MA, Pino-Montes J, *et al.* Association of aromatase and estrogen receptor gene polymorphisms with hip fractures. *Osteoporos Int* 2008;19:787–92.
36. Li F, Yu S, Feng X, Wang H, Fu S, Ding Y. Locally synthesized estrogen: a novel target in the development of OA. *Med Hypotheses* 2009;73:144–6.
37. Simpson ER. Role of aromatase in sex steroid action. *J Mol Endocrinol* 2000;25:1549–56.
38. McCloskey E. Effects of third-generation aromatase inhibitors on bone. *Eur J Cancer* 2006;42:1044–51.
39. Leskela HV, Oikku A, Lehtonen S, Mahonen A, Koivunen J, Turpeinen M, *et al.* Estrogen receptor alpha genotype confers interindividual variability of response to estrogen and testosterone in mesenchymal-stem-cell-derived osteoblasts. *Bone* 2006;39:1026–34.
40. Valdes AM, Van Oene M, Hart DJ, Surdulescu GL, Loughlin J, Doherty M, *et al.* Reproducible genetic associations between candidate genes and clinical knee osteoarthritis in men and women. *Arthritis Rheum* 2006;54:533–9.
41. Jin SY, Hong SJ, Yang HI, Park SD, Yoo MC, Lee HJ, *et al.* Estrogen receptor-alpha gene haplotype is associated with primary knee osteoarthritis in Korean population. *Arthritis Res Ther* 2004;6:R415–21.
42. Chen S, Ye J, Kijima I, Kinoshita Y, Zhou D. Positive and negative transcriptional regulation of aromatase expression in human breast cancer tissue. *J Steroid Biochem Mol Biol* 2005;95:17–23.
43. Kinney RC, Schwartz Z, Week K, Lotz MK, Boyan BD. Human articular chondrocytes exhibit sexual dimorphism in their responses to 17beta-estradiol. *Osteoarthritis Cartilage* 2005;13:330–7.
44. Oz OK, Zerwekh JE, Graves K, Nanu L, Millsaps R, Simpson ER. Bone has a sexually dimorphic response to aromatase deficiency. *J Bone Miner Res* 2000;15:507–14.