

# Caracterización genética del fenotipo de progresión rápida en la patología artrósica

Autor/a: Alejandro Durán-Sotuela

---

Tesis doctoral UDC / 2023

Directores: Prof. Francisco Javier Blanco García

Dr. José Ignacio Rego Pérez

Tutor/a: Dr. José Ignacio Rego Pérez

Programa de doctorado en Ciencias de la Salud (RD99/2011)

*Grupo de Investigación en Reumatología (GIR).*

*Instituto de Investigación Biomédica da Coruña (INIBIC).*



El Prof. Francisco Javier Blanco García y el Dr. José Ignacio Rego Pérez, investigadores del Grupo de Investigación en Reumatología (GIR) del Instituto de Investigación Biomédica de la Coruña (INIBIC) y directores del presente trabajo de tesis doctoral,

CERTIFICAN QUE:

La presente memoria de tesis titulada “Caracterización genética del fenotipo de progresión rápida en la patología artrósica” presentada por Don Alejandro Durán Sotuela, ha sido realizada bajo nuestra dirección y supervisión en el Grupo de Investigación en Reumatología en el INIBIC y reúne las condiciones necesarias para ser defendida públicamente optando al Grado de Doctor con mención Internacional.

Para que así conste, firman el siguiente certificado en A Coruña, con fecha

Prof. Francisco J Blanco García

Dr. José Ignacio Rego Pérez





A continuación, se adjunta en este texto el justificante de estancia predoctoral realizada en Portugal, con el fin de optar al grado de Doctor con mención Internacional.



Porto, 11 September 2019

## DECLARATION

To Whom It May Concern,

I hereby confirm that the PhD student Alejandro Durán stayed for a period of 3 months (from 1<sup>st</sup> of April to the 30<sup>th</sup> of June 2019) in the Diagnostic Unit of the Institute of Molecular Pathology and Immunology of the University of Porto - IPATIMUP.

Do not hesitate to contact me if you need further information.

Yours faithfully,

---

Prof. Doutor José Luis Costa  
Professor Afiliado, Faculdade de Medicina da Universidade do Porto  
Investigador Principal, IPATIMUP/i3S  
Rua Alfredo Allen 208  
Tel. 220408808  
Email: [jcosta@ipatimup.pt](mailto:jcosta@ipatimup.pt)



*A mis padres,*

*A Catuxa*

*A Máximo*

*y como no...*

*a Arthas.*







# *Agradecimientos*







Parece mentira que ya hayan pasado cinco años desde que comencé esta etapa de mi vida y, como todo, llegado el momento, se acaba. Cabe destacar que la tesis es un proceso de aprendizaje constante, con sus altibajos, frustrante, pero a la vez enriquecedora, y en ella, no solo intervienes tú, sino tu entorno, ya que en ellos te vas a basar para desestresarte, apoyarte, solicitar ayuda y aprender. Es por esto último por lo que, en este punto, echas la vista atrás, y te das cuenta de la enorme cantidad de personas que sientes que en algún momento han hecho este proceso más ameno, y por lo que hay que darles las gracias.

En primer lugar, me gustaría agradecer a mis directores, el Dr. Francisco Blanco García y el Dr. Ignacio Rego Pérez. Al Dr. Blanco por tenerme bajo su tutela y darme la oportunidad de desarrollar este trabajo en su grupo y, al doctor Rego, por el apoyo, la confianza y la enseñanza constante. “Nacho” ha intentado siempre mantener la formalidad, pero, en un gran número de ocasiones acaba cayendo en mis tocaditas de narices y desesperándose (lo cual me llena de orgullo y satisfacción). También me gustaría agradecer a las instituciones que me han permitido realizar mi formación académica: la Fundación Profesor Novoa Santos, la Sociedad Española de Reumatología, la *Axencia Galega de Innovación* y el Instituto de Salud Carlos III.

En segundo lugar, debo de agradecer también al resto de compañeros de genómica, tanto a los que han estado como a los que están actualmente, primero a Sara, ya que me enseñó todo lo necesario para poder desenvolverme adecuadamente en el laboratorio. A María Eugenia por todas esas luchas territoriales en poyata y por el bueno rollo. A Morena, buena amiga a la que es fácil tomarle el pelo, *ecco* se te echa de menos. A Merce, ojalá se me contagie tu metodismo y solvente todo mi caos. También obviamente a mis compañeros *predocs* de genómica, con los que he pasado mis años, Andrea Dalmao, Paula Ramos y más recientemente Jorge, los cuales comparten tus agobios laborales día a día contigo, lo que nos permite empatizar muy bien entre nosotros y apoyarnos.

Pero el grupo en el que he realizado mi tesis doctoral es muy grande, no sólo genómica, hay mucha más gente que me han aportado cosas estos años. Por tanto, en tercer lugar, es necesario agradecerle a cada uno de los miembros del grupo de investigación en reumatología su papel en esta tesis, ya que siempre han aportado algo. Al grupo de proteómica con el que he compartido tantas reuniones, gracias a todas: Valentina, Patri, Rocío, Lucía, Cris, Bea y María Camacho. También agradecer al grupo de epigenética, liderado por May, con la que me he fascinado por su exacerbado entusiasmo por la investigación y también a Marga y Carmen, a las que he ayudado y me han ayudado en diversas técnicas. Al grupo de medicina regenerativa, destacando a mi querida Tamara, siempre está disponible para echar una mano y largarme una colleja, gracias por estar ahí, también a Noa, Elena y Charlie. Con las compañeras de biología del cartílago también he tenido mis peripecias en congresos, gracias a Bea y a Irene por estar ahí, especialmente a esta última, con sus miniagobios y también por echarnos siempre una mano mutuamente. También me gustaría dar las gracias a los miembros clínicos del grupo, especialmente a Natividad, con la que también he compartido viajes a congresos, nervios y también por la excelente gestión en la cohorte PROCOAC.

Además de miembros del servicio de reumatología, he tenido más compañeras, aunque no hayan pertenecido a dicho grupo. En primer lugar, a mis compis de poyata, las chicas de inflamación, Oly por siempre estar sacándome esa sonrisilla a pesar de que la situación pudiera ser dura, *that's what she said*, y a Jenny, que siempre deja alguna frase perfecta para soltarle un buen zasca, pero en el buen sentido. También gracias al resto de *predocs* con los que he compartido espacio, Rocío, Andrés, Adrián, Marta, Amanda, Marina, Alba, Dani, Andrea Díaz, Maca, Gloria, Andrea Rodríguez, Miriam y Christian. Y además de a los *predocs*, tengo que

## Agradecimientos

agradecer en gran medida a María Moreno, tanto por su capacidad resolutive en pequeños trámites como en su capacidad informativa, a Estefanía Cives, la cual siempre me ha solventado todos los temas burocráticos que le he presentado, a Paula Juiz, por conseguir ayudarme a controlar mi sonrojez y, a Mariajo por enseñarme lo poco que sé de cultivos.

También debo de agradecer en cuarto lugar a mis compañeros portugueses del Instituto de patología molecular e inmunología de la universidad de Oporto (IPATIMUP), gracias a ellos aprendí mucho con respecto a las técnicas fundamentales de esta tesis, y sin ellos no habría sido posible. Gracias a Jose Luis Costa, por haberme permitido realizar la estancia, y a Luis Cirnes, un gran profesional de la patología del cáncer, que me ayudó durante la estancia y que siempre consiguió alegrarme el día con su peculiar forma de ser. También agradecer a mis compañeras de estancia que más tiempo pasaron conmigo, Maria João, Joana Pereira, Márcia Baixia, Ana Mafalda y Ron.

Los que me conocéis sabéis que siempre he sido una persona muy activa, necesito rutinas de entreno para poder mantener mi cabeza amueblada. En este ámbito, lo que especialmente más me llena y me libera de mis agobios son los deportes de contacto. Muchos años antes de comenzar la tesis, concretamente con catorce años, había estado practicado Jiu-jitsu tradicional japonés. Por tanto, en quinto lugar, me gustaría agradecer a mi primer maestro y buen amigo David Pardavila, con el cuál echo mucho de menos ir a practicar suelo durante horas, entre risas, paranoias mentales y leñazos, con él comencé en este mundillo el cuál es a día de hoy mi forma rápida de deshincharme. A pesar de que con el tiempo dejara de lado este deporte, a lo largo de esos años otros muchos compañeros me han aportado mucho, tanto en los entrenos como en las competiciones, por lo que he de hacer mención de todos ellos en este documento, gracias a todos vosotros por haber estado ahí, Álex Álvarez, Domínguez, Dani Marín, Sariña, Jaco y sin olvidarme de Canosa. Actualmente me he centrado en el Jiu-jitsu brasileño, que es la parte que más me gustaba del anterior arte marcial, y los nuevos compañeros y el maestro, Adriano, también me tratan muy bien, gracias a todos.

En sexto lugar, les tengo que dar las gracias a mis amigos. Con quince años se comenzó a formar mi grupo de amigos más cercanos, estoy hablando de Miguel, Lidia, Jugo y Adrián. Estas cuatro personas se han mantenido constantes, aunque el grupo de colegas aumentase o disminuyese, y siguen estando a día de hoy ahí, y además introdujeron a buenos fichajes al grupo que siempre aportan buen rollo, hablo de Jusé, Raquel y Alba. Me da mucha pena hacernos mayores, que no tengamos todas las tardes y noches libres para hacer el imbécil, que las responsabilidades no me dejen estar con vosotros tanto como desearía y sentir que cada día que no os veo me alejo más. Pero por suerte, cuando por fin consigo estar un pequeño rato con vosotros, todo eso se disipa, me apoyáis y solo está presente un fuerte sentimiento de cariño... os quiero. Obviamente también ha habido nuevas incorporaciones a mi círculo de amigos desde entonces. Gracias a David, por haberme ayudado durante la carrera de Biología, no solo como profesor adjunto sino como amigo, y por meterme en el mundillo de juegos de mesa, ya que gracias a eso conocí también a Alex, el cuál es sorprendentemente fácil de picar y divertirse con él, gracias a ambos por los buenos ratos. También con respecto a la carrera, quiero destacar a Manuel Pimentel, con él hice mi trabajo de fin de grado, y consiguió que algo que no me gusta, la botánica, fuera apasionante, y además de profesor, también continúa siendo un buen amigo a día de hoy. También agradecer a una *predoc* que no he mencionado hasta ahora, mi amiga María Piñeiro, con la cual el doctorado era más sencillo y una vez que se fue más cuesta arriba, por suerte no he perdido contacto con ella ni con su reciente marido, Alejandro Pan, un friki de los buenos y con los que me lo paso como un enano hablando de lo que sea, gracias a los dos



por estar ahí. Por último, Iván y Alba, con las que las quedadas de tertulia y juegos de mesa nunca son aburridas.

Y en séptimo lugar, y no por ello menos importante, a mi familia. Primero agradecer a mis padres Ana y Joaquín, siempre se han desvivido por mí, intentando tener acceso a todo aquello a lo que ellos no pudieron por causas del destino. Mi madre siempre ha estado ahí, con reflejos de leona para apoyarme, escucharme, cuidarme y también para discutir, sé que desde que me independicé no nos vemos todo lo que quieres y sufres mucho por ello, pero eso no hace que te quiera menos o que no te eche de menos, boba. Mi padre desde muy niño me ha llevado a explorar el mundo que nos rodea y tocarlo todo, haciendo que mi interés por la biología creciera exponencialmente, ha sido mi compañero de travesuras y también mi verdugo porque si no, la cosa se iba de las manos. Si bien mi madre me ha apoyado con toda su fuerza, mi padre ha sabido canalizar dicho impulso hacia donde estoy hoy, y desde mi punto de vista, han hecho un trabajo mejor valorado que cualquier tesis doctoral, os quiero mucho. A mi abuelo Domiciano, el cuál siempre ha sabido guiarme desde donde esté, aunque ya hayan pasado más de quince años, en realidad nunca te has ido de mi lado, muchas gracias abuelo, te echo de menos. Y también a Paqui, que, a pesar de todo siempre me ha querido y ayudado. Por suerte o por desgracia, el resto de mi familia siempre ha estado muy lejos de mí y solo podía verlos brevemente una o dos veces al año, pero eso no quita que en algún momento me hayan marcado, por desgracia no puedo decirles a cada uno lo que pienso o estos agradecimientos serán eternos, gracias a todos mis tíos y tías: Pili, Martín, Conchi, Geli, Jaime, Javi, Carmen, Eva, Montse, Jose, Manolo, Mari Ángeles, Pepi y Paco, y claramente gracias a todos mis primos y primas: Víctor, Miky, Linsey, Lidia, Adrián, Gerard, Marc, Roger y Edgar (especialmente estos cuatro últimos son unos capullos) y, también, muchas gracias a mi abuela Marisol, ella siempre busca la gresca con todo el mundo, pero en el fondo, nos ha cuidado y querido a todos con todo su corazón. Independientemente de lo que haya pasado, o pueda pasar, siempre seremos familia.

También tengo que agradecer a mi familia más reciente, hablo de Luis y Pilar, los cuales me recibieron con los brazos abiertos. Siempre me han tratado bien, Luis, que siempre es un “fuguillas” nos ha apoyado y nos ha dado buenos momentos entre pavos, peluches y tute. Pili a pesar de que se haga la dura es un cacho de pan, y siempre cocina algo que me gusta, aunque no fuera su plan inicial (y además me enseñó a hacer filloas, cosa que agradeceré eternamente). Mis cuñados no se quedan atrás. Inés es súper fácil de vacilar, y eso hace que me ría mucho con ella, también sabe escuchar, aconsejar, y aunque siempre me haga el duro y no se lo diga a menudo, los dos sabemos que, en el fondo, no podría tener una hermana política mejor. Con Jorge, creo que inicialmente nos costó mucho soltarnos, pero no cambiaría nada de todo ese proceso, ya que sin ello la relación que tenemos actualmente no sería la misma. Gracias a los cuatro por estar ahí.

Es imposible pasar por alto a Paula Catarina, alias Catuxa. ¿Qué te puedo decir nena que no sepas? Desde que comenzó nuestro viaje juntos nunca me has defraudado, incluso ni cuando discutimos, y mira que lo hacemos con lo cabezones que somos los dos, pero ya sabes que, a mí, me encanta discutirlo todo. Tú conoces todos mis puntos fuertes y también algunos que incluso yo desconozco, siempre has sabido apoyarme y cubrir todas mis debilidades, haciendo que yo me venga arriba. Día a día crecemos juntos, como pareja y viejóvenes, y ponemos un pequeño adoquín en el camino que vamos formando, a un ritmo constante, y eso me encanta. Dentro de poco se nos viene el Máximo desafío encima, pero dudo que no podamos con él. Gracias por estar ahí día tras día, en nuestro humilde hogar. Te amo nena.

## *Agradecimientos*

Y finalmente a ti, Máximo. Estoy escribiendo esto cuando todavía falta algo de tiempo para que llegues a este mundo, pero espero que algún día puedas leerlo. Gracias por haberme ilusionado tantísimo, cada día que pasa mi mente piensa más en como interaccionar y jugar contigo y menos en las cosas que me puedan agobiar. Estoy deseando que todo vaya bien y poder verte, hijo mío.

# *Resumo*

Esta tese de doutoramento trata de identificar variantes xenéticas capaces de alterar o risco de desenvolver o fenotipo de artrose de xeonllo con progresión rápida. Para iso utilizáronse técnicas de secuenciación para atopar polimorfismos presentes no ADN nuclear e mitocondrial que poidan favorecer o desenvolvemento do fenotipo de interese.

Detectouse unha variante mitocondrial, presente na posición m.16519C, que estivo significativamente asociada coa rápida progresión. Despois de realizar un pequeno estudo funcional, detectáronse niveis máis altos de interleucina 6 e cicloxixenasa 2 nos cíbridos portadores da variante, indicando que, o aumento da resposta inflamatoria podería ser un mecanismo desencadeante para unha progresión rápida. Por outra banda, detectouse por primeira vez unha interacción entre o ADN nuclear e o mitocondrial no campo da artrose. A presenza do haplogrupo mitocondrial UK e do polimorfismo nuclear rs12107036 tamén aumenta o risco de progresión rápida na artrose do xeonllo.

Por último, xerouse un modelo predictivo de gran relevancia clínica e farmacolóxica, baseado en variables xenéticas e clínicas capaces de predecir o risco dun individuo de desenvolver o fenotipo de artrose de xeonllo con progresión rápida.

# *Resumen*

Esta tesis doctoral trata de identificar variantes genéticas capaces de alterar el riesgo de desarrollar el fenotipo de artrosis de rodilla con progresión rápida. Para ello, se han empleado técnicas de secuenciación para buscar polimorfismos presentes en el ADN nuclear y en el mitocondrial que puedan favorecer el desarrollo del fenotipo de interés.

Se detectó una variante mitocondrial, presente en la posición m.16519C, que se asoció significativamente con la progresión rápida. Tras realizar un pequeño estudio funcional, se detectaron mayores niveles de interleucina 6 y ciclooxigenasa 2 en los cíbridos portadores de la variante, pudiendo ser el aumento de la respuesta inflamatoria un mecanismo desencadenante en la progresión rápida. Por otro lado, se ha detectado por primera vez en el campo de la artrosis una interacción entre el ADN nuclear y el mitocondrial. La presencia del haplogrupo mitocondrial UK y el polimorfismo nuclear rs12107036 también aumenta el riesgo de progresión rápida en la artrosis de rodilla.

Finalmente, se generó un modelo predictivo de gran relevancia clínica y farmacológica, basado en variables genéticas y clínicas capaz de predecir el riesgo que tiene un individuo de desarrollar el fenotipo de artrosis de rodilla con progresión rápida.

## *Abstract*

This thesis project tries to identify genetic variants capable of altering the risk of developing the rapidly progressive phenotype in knee osteoarthritis. To achieve this, sequencing techniques have been used to search for polymorphisms present in nuclear and mitochondrial DNA that may favor the development of the phenotype of interest.

A mitochondrial variant, present at m.16519C, was detected and was significantly associated with rapid progression. After carrying out a small functional study, higher levels of interleukin 6 and cyclooxygenase 2 were detected in cybrids carrying this variant, and the increased inflammatory response could be a triggering mechanism for rapid progression. On the other hand, an interaction between nuclear and mitochondrial DNA has been detected for the first time in the field of osteoarthritis. The presence of the mitochondrial haplogroup UK and the rs12107036 nuclear polymorphism also increases the risk of rapid progression in knee osteoarthritis.

Finally, a predictive model of great clinical and pharmacological relevance was generated, based on genetic and clinical variables, which is capable of predicting the risk of developing the rapidly progressive phenotype of knee osteoarthritis.

# *Índice de contenidos*

<b>INTRODUCCIÓN</b>	1
• ARTROSIS	2
• POLIMORFISMOS GENÉTICOS NUCLEARES	25
• LA MITOCONDRIA	42
• TÉCNICAS DE SECUENCIACIÓN PARA ANALIZAR EL GENOMA	77
<b>JUSTIFICACIÓN</b>	88
<b>HIPÓTESIS Y OBJETIVOS</b>	92
<b>MATERIAL Y MÉTODOS</b>	96
• MATERIAL Y MÉTODOS OBJETIVO ESPECÍFICO NÚMERO 1	98
• MATERIAL Y MÉTODOS OBJETIVO ESPECÍFICO NÚMERO 2	104
• MATERIAL Y MÉTODOS OBJETIVO ESPECÍFICO NÚMERO 3	118
• MATERIAL Y MÉTODOS OBJETIVO ESPECÍFICO NÚMERO 4	124
<b>RESULTADOS</b>	136
• RESULTADOS DEL OBJETIVO ESPECÍFICO NÚMERO 1	138
• RESULTADOS DEL OBJETIVO ESPECÍFICO NÚMERO 2	144
• RESULTADOS DEL OBJETIVO ESPECÍFICO NÚMERO 3	150
• RESULTADOS DEL OBJETIVO ESPECÍFICO NÚMERO 4	154
<b>DISCUSIÓN</b>	162
<b>CONCLUSIONES</b>	192
<b>BIBLIOGRAFÍA</b>	196









# *Índice de figuras*

FIGURA 1. CLASIFICACIÓN DE LOS TIPOS DE ARTICULACIONES SINOVIALES.	0
FIGURA 2. LOS ELEMENTOS ARTICULARES QUE COMPONEN LA RODILLA.	5
FIGURA 3. LA ESTRUCTURA EN CAPAS DEL CARTÍLAGO ARTICULAR.	7
FIGURA 4. TIPOS DE MAL ALINEAMIENTO DE RODILLA.	12
FIGURA 5. EFECTO DE LA DIETA EN EL ESTRÉS Y DESARROLLO DE PATOLOGÍAS.	14
FIGURA 6. CARACTERÍSTICAS DE LA ESCALA KL.	16
FIGURA 7. LA COMPARTIMENTACIÓN DE LA RODILLA.	19
FIGURA 8. LA ESTRUCTURA DEL ADN DE DOBLE CADENA.	26
FIGURA 9. LOS DISTINTOS NIVELES DE EMPAQUETAMIENTO DE LA MOLÉCULA DE ADN.	27
FIGURA 10. LA ORGANIZACIÓN DE UN GEN.	28
FIGURA 11. LA FASE DE ELONGACIÓN DE LA REPLICACIÓN.	29
FIGURA 12. LA TRANSCRIPCIÓN.	30
FIGURA 13. LA TRADUCCIÓN.	31
FIGURA 14. REPRESENTACIÓN DEL DESEQUILIBRIO DE LIGAMIENTO.	36
FIGURA 15. RESUMEN DE LOS PASOS PARA LLEVAR A CABO UN ESTUDIO GWAS.	37
FIGURA 16. EJEMPLO DE UN GRÁFICO MANHATTAN.	38
FIGURA 17. RESUMEN DE POLIMORFISMOS QUE SE ASOCIARON CON LA ARTROSIS.	41
FIGURA 18. LA ESTRUCTURA DE LA MITOCONDRIA.	42
FIGURA 19. LA CADENA DE TRANSPORTE DE ELECTRONES Y LA FOSFORILACIÓN OXIDATIVA.	43
FIGURA 20. LA VÍA INTRÍNSECA DE LA APOPTOSIS CELULAR.	45
FIGURA 21. ADN MITOCONDRIAL.	48
FIGURA 22. SISTEMA DE EMPAQUETAMIENTO DEL ADNMT BASADO EN NUCLEOIDES.	49
FIGURA 23. REPLICACIÓN DEL ADNMT.	50
FIGURA 24. EL FENÓMENO DE SEGREGACIÓN REPLICATIVA.	52
FIGURA 25. MAPA MIGRATORIO DE LA ESPECIE HUMANA DESDE LA EVA MITOCONDRIAL.	56
FIGURA 26. MODELOS FUNCIONALES CON CÍBRIDOS.	69
FIGURA 27. CREACIÓN DE RATONES CONPLÁSTICOS.	70
FIGURA 28. COMUNICACIÓN NÚCLEO-MITOCONDRIA.	72
FIGURA 29. EL PROCESO DE SECUENCIACIÓN QUÍMICA DE MAXAM Y GILBERT.	78
FIGURA 30. LOS PASOS NECESARIOS PARA REALIZAR UN PROCESO DE SECUENCIACIÓN SANGER.	79
FIGURA 31. PROCESO DE BRIDGE AMPLIFICATION DE ILLUMINA.	81
FIGURA 32. PROCESO DE SECUENCIACIÓN POR SÍNTESIS DE ILLUMINA.	81
FIGURA 33. PROCESO DE SECUENCIACIÓN DE ION TORRENT.	83
FIGURA 34. PROCESO DE SECUENCIACIÓN MEDIANTE TECNOLOGÍA SOLID.	84
FIGURA 35. PROCESO DE SECUENCIACIÓN EN PACBIO.	86
FIGURA 36. PROCESO DE SECUENCIACIÓN EN LA METODOLOGÍA DE NANOPORE.	87
FIGURA 37. POSIBLES PERÍODOS DE SEGUIMIENTO EN LAS TRES COHORTES.	101
FIGURA 38. PROCESO DE CUANTIFICACIÓN DE LAS MUESTRAS DE ADNMT CON QUBIT.	105
FIGURA 39. COBERTURA DEL PRECISION ID MTDNA WHOLE GENOME PANEL.	106
FIGURA 40. LOS ELEMENTOS QUE COMPONEN UN FRAGMENTO DE LIBRERÍA DE ADN.	107
FIGURA 41. PROTOCOLO DE GENERACIÓN DE LIBRERÍAS CON EL ION CHEF.	108
FIGURA 42. EL CARTUCHO DE REACTIVOS DEL KIT PRECISION ID DL8 KIT.	108
FIGURA 43. PROCESO DE AMPLIFICACIÓN CLONAL EN ION TORRENT.	110
FIGURA 44. PROCESO DE CARGA DEL CHIP.	111
FIGURA 45. PROCESO DE SECUENCIACIÓN DE ION TORRENT.	111
FIGURA 46. REPRESENTACIÓN GRÁFICA DE LOS VALORES DETECTADOS POR EL SEMICONDUCTOR.	112
FIGURA 47. PROCESO DE SECUENCIACIÓN Y EL PROCESAMIENTO DE LOS DATOS CON ION TORRENT.	113
FIGURA 48. ELECTROFORESIS EN GEL DE AGAROSA DE LA POSICIÓN M.16519.	115

## *Índice de figuras*

FIGURA 49. PURIFICACIÓN DEL PRODUCTO DE PCR CON EXOSAP-IT.	115
FIGURA 50. RESULTADOS DE SECUENCIACIÓN CAPILAR PARA LA VARIANTE MITOCONDRIAL.	117
FIGURA 51. OBTENCIÓN DE LOS CÍBRIDOS TRANSMITOCONDRIALES EMPLEADOS.	118
FIGURA 52. ESQUEMA REPRESENTATIVO DEL PROCESO DE SBE.	125
FIGURA 53. PCR MÚLTIPLE DE LOS SIETE FRAGMENTOS DE LOS SNPS NUCLEARES.	128
FIGURA 54. SBE Y FAST-AP.	129
FIGURA 55. REPRESENTACIÓN DEL PROCESO DE MINI-SECUENCIACIÓN CAPILAR.	130
FIGURA 56. EJEMPLO DE DATOS DE SBE PARA LOS SIETE SNPS NUCLEARES INICIALES.	130
FIGURA 57. EJEMPLO DE PCR MÚLTIPLEX DE HAPLOGRUPOS CAUCÁSICOS.	132
FIGURA 58. EJEMPLO DE DATOS DE SBE PARA LOS HAPLOGRUPOS MITOCONDRIALES.	133
FIGURA 59. FILTRADO DE LOS DATOS BRUTOS OBTENIDOS DE LA NGS.	144
FIGURA 60. FOREST PLOT RESULTANTE DEL META-ANÁLISIS DE LA VARIANTE M.16519C.	149
FIGURA 61. ANÁLISIS DE COMPONENTES PRINCIPALES DE LOS DATOS DE RNA-SEQ.	151
FIGURA 62. PROCESOS BIOLÓGICOS AFECTADOS POR LOS DEGS SOBREEXPRESADOS.	151
FIGURA 63. ANÁLISIS MEDIANTE QPCR DE LOS NIVELES DE IL6 Y COX2.	152
FIGURA 64. COMPARATIVA DE LAS ÁREAS BAJO LA CURVA.	160
FIGURA 65. NOMOGRAMA PARA EL FENOTIPO DE ARTROSIS DE RODILLA CON PROGRESIÓN RÁPIDA	161

# Índice de tablas

TABLA 1. COMPARATIVA DE LOS VALORES DE PREVALENCIA ENTRE LOS ESTUDIOS EPISER 2000 Y 2016.	3
TABLA 2. RESUMEN DE LOS CRITERIOS PROPUESTOS PARA DEFINIR UN PROGRESOR RÁPIDO.	20
TABLA 3. RESUMEN DE LAS PRINCIPALES ASOCIACIONES ENTRE POLIMORFISMOS MITOCONDRIALES Y DIVERSAS PATOLOGÍAS EN DIFERENTES POBLACIONES HUMANAS.	62
TABLA 4. TABLA RESUMEN DE ALGUNOS EJEMPLOS DE INTERACCIONES GENÉTICAS NUCLEO-MITOCONDRIA QUE SE ASOCIAN CON DIVERSAS PATOLOGÍAS HUMANAS.	74
TABLA 5. OLIGONUCLEÓTIDOS EMPLEADOS EN LA AMPLIFICACIÓN Y SECUENCIACIÓN DE LA POSICIÓN MITOCONDRIAL DE INTERÉS.	115
TABLA 6. INFORMACIÓN REFERENTE A LOS POLIMORFISMOS NUCLEARES SELECCIONADOS.	124
TABLA 7. PAREJAS DE OLIGONUCLEÓTIDOS DE PCR DISEÑADAS PARA CADA UNO DE LOS SNPS NUCLEARES ASOCIADOS CON SUSCEPTIBILIDAD A ARTROSIS DE RODILLA.	126
TABLA 8. OLIGONUCLEÓTIDOS DE SBE SELECCIONADOS PARA CADA UNO DE LOS SNPS NUCLEARES ASOCIADOS CON SUSCEPTIBILIDAD A ARTROSIS DE RODILLA.	127
TABLA 9. OLIGONUCLEÓTIDOS DE PCR Y SBE EMPLEADOS EN LA DETERMINACIÓN DE HAPLOGRUPOS CAUCÁSICOS.	131
TABLA 10. CLASIFICACIÓN DE LOS HAPLOGRUPOS CAUCÁSICOS.	132
TABLA 11. RECUENTO DE INDIVIDUOS SELECCIONADOS DENTRO DE LOS CRITERIOS DE PROGRESIÓN DRIBAN Y GIR-BLANCO EN LA COHORTE OAI.	138
TABLA 12. VARIABLES CLÍNICAS Y DEMOGRÁFICAS DE LOS INDIVIDUOS SELECCIONADOS EN BASE AL CRITERIO DE PROGRESIÓN DRIBAN DE TRES GRUPOS DE PROGRESIÓN.	139
TABLA 13. VARIABLES CLÍNICAS Y DEMOGRÁFICAS DE LOS INDIVIDUOS SELECCIONADOS EN BASE AL CRITERIO DE PROGRESIÓN DRIBAN DE DOS GRUPOS DE PROGRESIÓN.	139
TABLA 14. VARIABLES CLÍNICAS Y DEMOGRÁFICAS DE LOS INDIVIDUOS SELECCIONADOS EN BASE AL CRITERIO DE PROGRESIÓN GIR-BLANCO DE TRES GRUPOS DE PROGRESIÓN.	140
TABLA 15. VARIABLES CLÍNICAS Y DEMOGRÁFICAS DE LOS INDIVIDUOS SELECCIONADOS EN BASE AL CRITERIO DE PROGRESIÓN GIR-BLANCO DE DOS GRUPOS DE PROGRESIÓN.	140
TABLA 16. ANÁLISIS DE LA CORRELACIÓN DE LA CLASIFICACIÓN DE LA PROGRESIÓN RÁPIDA EN LOS CRITERIOS DE TRES CATEGORÍAS.	141
TABLA 17. ANÁLISIS DE LA CORRELACIÓN DE LA CLASIFICACIÓN DE LA PROGRESIÓN RÁPIDA EN LOS CRITERIOS DE DOS CATEGORÍAS.	141
TABLA 18. RECUENTO DE INDIVIDUOS SELECCIONADOS DENTRO DE LOS CRITERIOS DE PROGRESIÓN DRIBAN Y GIR-BLANCO EN LA COHORTE DE VALIDACIÓN EUROPEA.	142
TABLA 19. VARIABLES CLÍNICAS Y DEMOGRÁFICAS DE LOS INDIVIDUOS SELECCIONADOS EN BASE AL CRITERIO DE PROGRESIÓN GIR-BLANCO DE TRES GRUPOS DE PROGRESIÓN DE LA COHORTE DE VALIDACIÓN EUROPEA.	143
TABLA 20. VARIABLES CLÍNICAS Y DEMOGRÁFICAS DE LOS INDIVIDUOS SELECCIONADOS EN BASE AL CRITERIO DE PROGRESIÓN GIR-BLANCO DE DOS GRUPOS DE PROGRESIÓN EN LAS COHORTE EUROPEA DE VALIDACIÓN.	143
TABLA 21. LISTA DE LAS 76 VARIANTES MITOCONDRIALES QUE SUPERARON LOS FILTROS INICIALES DE MAF Y FRECUENCIA DE HETEROPLASMA EN EL CRITERIO GIR-BLANCO DE DOS GRUPOS DE PROGRESIÓN.	145
TABLA 22. LISTA DE VARIANTES CON SOBRRERREPRESENTACIÓN DENTRO DEL GRUPO DE PROGRESORES RÁPIDOS EN LAS 450 MUESTRAS INICIALES USADAS EN LA NGS EN EL CRITERIO GIR-BLANCO DE DOS GRUPOS DE PROGRESIÓN.	146
TABLA 23. ANÁLISIS UNIVARIANTE DE LA VARIANTE M.16519C EN LOS INTEGRANTES DE LA COHORTE OAI DE LOS CRITERIOS DE PROGRESIÓN GIR-BLANCO.	147
TABLA 24. ANÁLISIS UNIVARIANTE DE LA VARIANTE M.16519C EN LOS INTEGRANTES DE LAS COHORTES EXTERNAS DE LOS CRITERIOS DE PROGRESIÓN GIR-BLANCO.	148

## Índice de tablas

TABLA 25. ANÁLISIS DEL EFECTO DE LA VARIANTE M.16519C EN LA PROGRESIÓN RÁPIDA EN LA FASE DE EXPLORACIÓN Y VALIDACIÓN.	148
TABLA 26. LISTADO DE VARIANTES MITOCONDRIALES DIFERENCIALES ENTRE LAS DOS LÍNEAS DE CÍBRIDOS TRANSMITOCONDRIALES CREADAS.	150
TABLA 27. LISTADO DE LOS DIEZ DEGS MÁS SOBREENPRESADOS O REPRIMIDOS EN LOS CÍBRIDOS M.16519C.	152
TABLA 28. ANÁLISIS MULTIVARIANTE DE LOS NIVELES DE IL6 SÉRICOS DE LOS 229 PARTICIPANTES ANALIZADOS DE LA COHORTE OAI.	153
TABLA 29. ANÁLISIS UNIVARIANTES PARA LAS FRECUENCIAS ALÉLICAS DE LOS OCHO SNPS NUCLEARES EN LOS CRITERIOS GIR-BLANCO DE TRES GRUPOS DE PROGRESIÓN.	154
TABLA 30. ANÁLISIS UNIVARIANTES PARA LAS FRECUENCIAS ALÉLICAS DE LOS OCHO SNPS NUCLEARES EN LOS CRITERIOS GIR-BLANCO DE DOS GRUPOS DE PROGRESIÓN.	155
TABLA 31. ANÁLISIS UNIVARIANTES PARA LAS FRECUENCIAS DE LOS CLUSTERS DE HAPLOGRUPOS MITOCONDRIALES EN LOS CRITERIOS GIR-BLANCO DE TRES GRUPOS DE PROGRESIÓN.	155
TABLA 32. ANÁLISIS UNIVARIANTES PARA LAS FRECUENCIAS DE LOS CLUSTERS DE HAPLOGRUPOS MITOCONDRIALES EN LOS CRITERIOS GIR-BLANCO DE DOS GRUPOS DE PROGRESIÓN.	155
TABLA 33. ANÁLISIS UNIVARIANTES PARA LAS FRECUENCIAS ALÉLICAS DE LOS OCHO SNPS NUCLEARES EN LOS CRITERIOS GIR-BLANCO DE TRES GRUPOS DE PROGRESIÓN ESTRATIFICANDO LA POBLACIÓN EN BASE A SU CLUSTER MITOCONDRIAL.	156
TABLA 34. ANÁLISIS UNIVARIANTES PARA LAS FRECUENCIAS ALÉLICAS DE LOS OCHO SNPS NUCLEARES EN LOS CRITERIOS GIR-BLANCO DE DOS GRUPOS DE PROGRESIÓN ESTRATIFICANDO LA POBLACIÓN EN BASE A SU CLUSTER MITOCONDRIAL.	157
TABLA 35. MODELO PREDICTIVO PARA EL FENOTIPO DE ARTROSIS DE RODILLA CON PROGRESIÓN RÁPIDA DESARROLLADO EN LA COHORTE OAI.	158
TABLA 36. COMPARATIVA DE LOS EFECTOS DE LAS VARIABLES INCLUIDAS EN EL MODELO PREDICTIVO PARA FENOTIPO DE ARTROSIS DE RODILLA CON PROGRESIÓN RÁPIDA ENTRE LOS TRES GRUPOS DE PROGRESIÓN DEL CRITERIO GIR-BLANCO.	159
TABLA 37. CÁLCULO DE LA PROPORCIÓN DE RIESGO ATRIBUIBLE A LA INTERACCIÓN ENTRE EL ALELO G DEL SNP R12107036 DE TP63 Y EL CLUSTER MITOCONDRIAL UK.	160
TABLA 38. ÁREAS BAJO LA CURVA DEL MODELO DESARROLLADO Y DE SU VALIDACIÓN INTERNA CRUZADA.	160



# *Índice de abreviaturas*

%GC	Porcentaje de guanina-citosina
.BAM	<i>Binary alignment map</i>
.VCF	<i>Variant call format</i>
µm	Micrómetro(s)
8-OHdG	8-hidroxiguanosina
A	Adenina
Å	Angstrom
ABHD2	<i>Abhydrolase domain containing 2, acylglycerol lipase</i>
ACR	<i>American College of Rheumatology</i>
ACS	Suero autólogo condicionado
ADN	Ácido desoxirribonucleico
ADNc	ADN complementario
ADNmt	ADN mitocondrial
ADULT	<i>Acro-dermo-ungual-lacrima-tooth</i>
AINES	Antiinflamatorios no esteroideos
APPROACH	<i>Applied Public-Private Research enabling Osteoarthritis Clinical Headway</i>
ARN	Ácido ribonucleico
ARNm	ARN mensajero
ARNmt	ARN mitocondrial
ARNr	ARN ribosómico
ARNt	ARN de transferencia
ASPN	<i>Asporin</i>
ATP	<i>Adenosine triphosphate</i>

## *Índice de abreviaturas*

AUC	<i>Area under the ROC curve</i>
Bak	<i>Bcl2 antagonist/killer</i>
Bax	<i>Bcl2 associated x, apoptosis regulator</i>
Bcl2	<i>B-cell lymphoma 2</i>
BER	Sistema de reparación por escisión de base
BH	Benjamin-hochberg
BH3	<i>Bcl2 homology region 3</i>
Bid	<i>BH3 interacting-domain death agonist</i>
C	Citosina
Ca <sup>2+</sup>	Calcio
CCL5	<i>C-c motif chemokine ligand 5</i>
CH <sub>3</sub>	Grupo metilo
CHECK	<i>Cohort Hip and Cohort Knee</i>
CNV	Variante(s) de número de copias
COI	<i>Cytochrome oxidase c subunit 1</i>
COII	<i>Cytochrome oxidase c subunit 2</i>
COIII	<i>Cytochrome oxidase c subunit 3</i>
COMP	<i>Encoding cartilage oligomeric Matrix protein</i>
COX2	Ciclooxigenasa 2 o PTSG2 ( <i>Prostaglandin-Endoperoxide Synthase 2</i> )
CRISPR	<i>Clustered regularly interspaced short palindromic repeats</i>
CTE	Cadena de transporte de electrones
ddNTPs	Didesoxinucleótidos terminadores
DEG	Gen(es) diferencialmente expresados
DIGICOD	<i>Digital cohort Design</i>
DISC	<i>Death-inducing signaling complex</i>
D-Loop	<i>Displacement loop</i>

## Índice de abreviaturas

DMOAD	<i>Disease modifying osteoarthritis drug</i>
dNTPs	Desoxinucleótidos
DSBR	Sistema de reparación de roturas de doble cadena
DUS4L	<i>Dihydrouridine synthase 4-like</i>
EDTA	Ácido etilendiaminotetraacético
EEG	Ecuacion(es) de Estimación generalizadas
EMA	<i>European medicines agency</i>
ePCR	Amplificación en emulsión
eQTL	<i>Expression quantitative trait loci</i>
EULAR	Liga europea contra los reumatismos
FADD	<i>Fas-associated protein with death domain</i>
FADH <sub>2</sub>	Dinucleótido de flavina adenina
Fc	<i>Fold change</i>
FDA	<i>Food and Drug Administration</i>
FDR	<i>False Discovery Rate</i>
FGF	Factor de crecimiento fibroblástico
FIS1	<i>Mitochondrial fission protein 1</i>
fMet	N-formilmetionina
FTO	<i>Fat mass and obesity-associated protein</i>
G	Guanina
GAG	Glucosaminoglucano(s)
GDF5	<i>Growth differentiation factor 5</i>
GNL3	<i>Guanine nucleotide binding protein-like 3</i>
GWAS	Metodologías de asociación de genoma completo
H <sup>+</sup>	Protón(es)
H <sub>2</sub> O <sub>2</sub>	Peróxido de hidrógeno



## *Índice de abreviaturas*

HA	Ácido hialurónico
HapMap	<i>Haplotype map</i>
HIF1 $\alpha$	<i>Hypoxia-inducible factor 1-alpha</i>
HO <sub>2</sub> <sup>-</sup>	Hidroperoxilo
HSP	<i>Heavy strand promoter</i>
HtrA2	<i>High temperature requirement protein A2</i>
I2	Índice de heterogeneidad entre estudios
IC95%	Intervalo de confianza del 95%
IGF	Factor de crecimiento insulínico
IGFBP1	<i>Insulin like growth factor binding protein 1</i>
IGV	<i>Integrative genomics viewer</i>
IL1 $\beta$	Interleuquina-1 $\beta$
IL24	Interleuquina 24
IL4	<i>Interleuquina 4</i>
IL6	Interleuquina 6
IL8	Interleuquina 8
IMC	Índice de masa corporal
JSW	Espacio intra-articular
KL	Escala Kellgren y Lawrence
KOOS	<i>Knee injury and Osteoarthritis Outcomes Score</i>
LC3-II	<i>Microtubule-associated protein 1A/1B-light chain 3</i>
LHON	Neuropatía óptica de Lheber
LSP	<i>Light strand promoter</i>
MAF	<i>Minor allele frequency</i>
MCF2L	<i>Mcf2 transforming sequence-like protein</i>
MEC	Matriz extracelular del cartílago

## Índice de abreviaturas

MELAS	<i>Mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes syndrome</i>
meQTLs	<i>Methylation quantitative trait loci</i>
mJSW	Espacio mínimo intra-articular
mL	Mililitro(s)
MMPs	Metaloproteasas
MMR	Sistema de reparación de errores de apareamiento
MOST	<i>Multicentre osteoarthritis</i>
MT-ATP6	<i>Mitochondrially encoded ATP synthase membrane subunit 6</i>
MT-ND6	<i>Mitochondrially encoded nadh:ubiquinone oxidoreductase core subunit 6</i>
mtSSB	Proteína de unión monocatenaria mitocondrial
MT-TL1	<i>Mitochondrially encoded trna-Leu</i>
NADH	Dinucleótido de nicotinamida adenina
NARP	<i>Neurogenic muscle weakness, ataxia and retinitis pigmentosa</i>
NCBI	<i>National Center for Biotechnology Information</i>
NER	Sistema de reparación por escisión de nucleótido
NFκB	Factor nuclear potenciador de las cadenas ligeras kappa de las células B activadas
NGF	Factor de crecimiento nervioso
NGS	<i>Next-generation sequencing</i>
NIBAN3	<i>Niban apoptosis regulator 3</i>
nm	Nanómetro(s)
NO	Óxido nítrico
NRF2	<i>Nuclear factor erythroid 2-related factor 2</i>
NT5DC2	<i>5'-nucleotidase domain containing 2</i>
NUMTs	<i>Nuclear mitochondrial DNA segment</i>
O <sub>2</sub> <sup>-</sup>	Anión superóxido
OAI	<i>Osteoarthritis initiative</i>

## Índice de abreviaturas

OARSI	<i>Osteoarthritis research society international</i>
OH	Grupo hidroxilo
O <sub>h</sub>	Origen de replicación de la cadena pesada
OH <sup>-</sup>	Ion hidroxilo
O <sub>l</sub>	Origen de replicación de la cadena ligera
ONOO <sup>-</sup>	Peroxinitrito
OR	<i>Odds ratio</i>
ORC	<i>Origin recognition complex</i>
p62	<i>Sequestosome 1</i>
PAF	<i>Population attributable risk proportion</i>
pb	Par(es) de bases
PCA	Análisis de componentes principales
PCR	<i>Polymerase chain reaction</i>
PDGF	Factor de crecimiento derivado de plaquetas
PF	Oligonucleótido <i>forward</i>
PG	Proteoglicano(s)
PGC-1s	Coactivador del receptor gamma 1-alfa activado por el proliferador de peroxisomas
PGC1 $\alpha$	<i>Peroxisome proliferator-activated receptor-<math>\gamma</math> coactivator 1<math>\alpha</math></i>
Phb	Prohibitina
pM	Picomol(es)
PolG	Polimerasa- $\gamma$
POLRMT	ARN polimerasa mitocondrial
<i>pools</i>	Conjuntos
PPAR	Receptor(es) activados por proliferadores peroxisomales
ppi	Grupo pirofosfato
PR	Oligonucleótido <i>reverse</i>

## Índice de abreviaturas

PROCOAC	<i>Prospective cohort of A Coruña</i>
PRP	Plasma rico en plaquetas
p-rpS6	<i>Phospho-ribosomal protein S6</i>
PTGS2	Prostaglandina E2
PVT1	<i>Plasmacytoma variant translocation 1</i>
qPCR	PCR cuantitativa
rCRS	Secuencia de referencia de Cambridge revisada
RLBP1	<i>Retinaldehyde binding protein 1</i>
RNA-seq	<i>Secuenciación de ARN</i>
RNS	Radicales libres de nitrógeno
ROS	Radicales libres de oxígeno
<i>Run Plan</i>	Plan de carrera
RUNX2	<i>Runx family transcription factor 2</i>
S	Unidad de coeficiente de sedimentación Svedberg
SBE	<i>Single base extension</i>
SMAC	<i>Second mitochondria-derived activator of caspase</i>
SNPs	Polimorfismos de un solo nucleótido
SOD2	Superóxido dismutasa mitocondrial
SOLiD	<i>Sequencing by Oligonucleotide ligation and Detection</i>
SSB	<i>Single-strand-binding proteins</i>
SUPTH3	<i>Spt3 homolog</i>
T	Timina
TAE	Tris-acético-edta
TASOAC	<i>Tasmanian older adult cohort</i>
TBCA	<i>Tubulin folding cofactor a</i>
TFAM	Factor de transcripción mitocondrial A

## *Índice de abreviaturas*

TGFA	<i>Transforming growth factor alpha</i>
TGFβ	Factor de crecimiento transformante
TIMPs	Inhibidores de metaloproteasas
T <sub>m</sub>	Temperatura de <i>melting</i>
TNFR	<i>Tumor necrosis factor receptor</i>
TNFSF14	<i>Tnf superfamily member 14</i>
TNFα	Factor de necrosis tumoral α
TOARP	<i>Tool for Osteoarthritis Risk Prediction</i>
TP63	<i>Tumor protein 63</i>
TRPM1	<i>Transient receptor potential cation channel subfamily m member 1</i>
TVC	<i>Torrent variant caller</i>
UCPs	Proteínas desacopladoras
VEGF	Factor de crecimiento endotelial vascular
WOMAC	<i>Western Ontario and mcmaster Universities Osteoarthritis Index</i>
WSCD2	<i>WSC domain containing 2</i>
ZMW	<i>Zero mode waveguide</i>
ρ0	Célula rho-0

# *Introducción*







## Artrosis

### *Origen del término y definición actual de la artrosis.*

La artrosis como dolencia ha estado presente desde la antigüedad, pero el origen de su término es relativamente reciente. Hasta hace aproximadamente 250 años todas las formas de artritis crónica eran designadas como manifestaciones de la gota, la cual es la forma más común de artritis (Neogi, 2016). En el año 1816, Heberden fue capaz de separar la artritis reumatoide de la gota gracias a los denominados *digitorum nodi*, conocidos actualmente como nódulos de Heberden en la artrosis de mano. Aun así, por aquel entonces todavía se consideraban tanto a la artritis reumatoide como a la artrosis como diferentes grados de la misma entidad, la “artritis deformante”. Existe cierta discrepancia sobre en qué punto de la historia se diferenciaron claramente la artritis reumatoide de la artrosis, algunos autores atribuyen este mérito a Archibald E. Garrod en 1890 (Dequeker y Luyten, 2008), sin embargo, bibliografía más reciente afirma que este término surgió antes, en la década de 1850, donde apareció por primera vez el término artrosis de la mano gracias a Richar Von Volkman, diferenciando así ambas patologías (Dobson *et al.*, 2018).

Dicho concepto de artrosis fue evolucionando con el tiempo; inicialmente se consideraba como una enfermedad que afectaba únicamente al cartílago articular, produciendo su degeneración paulatina debido al desgaste natural. Actualmente esto se considera erróneo, y se entiende que el proceso de la artrosis consiste en una remodelación de los tejidos de la articulación, todo esto dirigido por mediadores inflamatorios dentro de la articulación afectada. Por tanto, los cambios patológicos no afectan sólo al cartílago, sino que incluyen el engrosamiento del hueso subcondral, la formación de osteofitos, la presencia de diversos niveles de inflamación de la membrana sinovial, la degeneración de los ligamentos y/o meniscos en la rodilla, la hipertrofia de la cápsula articular, así como la presencia de alteraciones en los músculos periarticulares, de los nervios y de la bolsa sinovial o “bursa”. Es decir, el concepto a día de hoy es que la artrosis afecta a todos los tejidos de la articulación como si fuera un órgano, y como resultado, se produce el fallo total de dicho órgano (Loeser *et al.*, 2012; Blanco, 2018).

Actualmente, la artrosis se define como un “desorden que afecta a las articulaciones móviles y que se caracteriza por un estrés celular y degradación de la matriz extracelular, iniciado por micro y macro lesiones que activan respuestas de reparación maladaptativas, incluyendo vías de señalización pro-inflamatorias de la respuesta inmune. La enfermedad comienza manifestándose como un desorden molecular (metabolismo anormal del tejido articular) seguido por trastornos anatómicos, y/o fisiológicos (caracterizados por la degradación del cartílago, remodelación ósea, formación de osteofitos, inflamación de la articulación y pérdida de la función normal de la articulación), que culminan en dolencia” (Kraus *et al.*, 2015).

### *Prevalencia e impacto socio-económico de la artrosis.*

Las enfermedades reumáticas son patologías crónicas que representan una de las principales causas de discapacidad en países desarrollados. Además, teniendo en cuenta el incremento en la esperanza de vida de la población y el incremento de la tasa de obesidad a nivel mundial, se espera que su prevalencia en la población siga aumentando (Lawrence *et al.*, 2008). Concretamente, en nuestro país se han llevado a cabo dos estudios sobre la prevalencia de enfermedades reumáticas en la población española, uno en el año 2000 (Carmona *et al.*, 2001) y otro en el 2016 (Blanco *et al.*, 2020). Es importante resaltar un par de aspectos a la hora de comparar los datos de dichos estudios:

## Introducción

- La edad de la población de estudio: en el estudio del año 2000 se analizaba la prevalencia de artrosis en la población de 20 años o más, mientras que en el de 2016, se analizaba la prevalencia de artrosis en la población de 40 años o más.
- La definición de la artrosis: en el estudio del año 2000 se aplicaron únicamente criterios de artrosis sintomática, mientras que en el del 2016 se incluyeron, además, criterios clínico-radiológicos.

Teniendo en cuenta estos elementos podemos observar que, en general, las cifras de prevalencia del año 2016 son más altas que las obtenidas en el año 2000 (Tabla 1). Otro dato relevante obtenido de este reciente estudio es que la prevalencia de artrosis en la población española en una o más articulaciones es del 29,35%.

Tabla 1. Comparativa de los valores de prevalencia de artrosis en distintas articulaciones entre los estudios EPISER 2000 y 2016.

Estudio	Artrosis rodilla		Artrosis manos		Artrosis cadera		Artrosis cervical		Artrosis lumbar	
	Frecuencia	IC 95%	Frecuencia	IC 95%	Frecuencia	IC 95%	Frecuencia	IC 95%	Frecuencia	IC 95%
Episer 2000	223/2192 (10,2%)	8,5- 11,9%	136/2192 (6,2%)	5,9- 6,5%	No		No		No	
Episer 2016	452/3336 (13,83%)	12,66- 15,11%	276/3336 (7,73%)	6,89- 8,67%	166/3336 (5,13%)	4,40- 5,99%	345/3336 (10,10%)	9,07- 11,24%	540/3336 (15,52%)	14,30- 16,83%

IC 95%: intervalo de confianza al 95% de significación estadística

Con los datos que se acaban de mencionar, no es de extrañar que esta patología suponga un gasto socio-económico elevado en España. En el artículo de Loza y colaboradores (Loza *et al.*, 2009) se estimaron los costes directos (gasto médico) e indirectos (como indemnizaciones o bajas laborables) de la artrosis de rodilla y cadera en España. Los autores incluyeron un total de 1.071 sujetos de 50 años o más con artrosis sintomática y radiológica en dichas articulaciones, y calcularon el coste medio anual por paciente, el cuál fue de 1.502 euros, siendo el 86% de dicho valor debido a costes directos. A partir de estos datos infirieron el coste anual nacional, el cual ascendía a 4.738 millones de euros, una cantidad que equivale al 0,5% del Producto Interior Bruto de España.

### La articulación.

Una articulación se define como la unión entre dos o más huesos por sus superficies de contacto (de Toro, 2010). La forma y el grado de movilidad varía ampliamente en relación a su función, y esto permite distinguir entre articulaciones con gran libertad de movimiento y articulaciones prácticamente inmóviles. Dependiendo de la naturaleza del material que separa los huesos, las articulaciones se pueden clasificar de la siguiente forma:

1. Articulaciones fibrosas: presentan escasa o nula movilidad, y por ello reciben el nombre de articulaciones sinartrodiales. En ellas los huesos se mantienen unidos por tejido conjuntivo fibroso y, además, carecen de cavidad articular. Un ejemplo de articulación fibrosa son las suturas entre los huesos del cráneo.
2. Articulaciones cartilaginosas: se caracterizan por su movilidad limitada, y reciben el nombre de articulaciones anfiartrodiales. En este tipo de articulaciones los huesos se mantienen unidos mediante cartílago y, al igual que las articulaciones fibrosas, también carecen de cavidad articular. Un ejemplo de articulación cartilaginosa serían la sínfisis púbica, la cual mantiene unida los huesos de la pelvis.

3. Articulaciones sinoviales: son las más abundantes y se caracterizan por su elevada movilidad, por ello reciben el nombre de articulaciones diartrodiales. A diferencia de los otros dos tipos de articulaciones, éstas constan de cavidad articular, cápsula articular y ligamentos. Los huesos que unen dicho tipo de articulación constan de superficies de contacto con morfología variable y están recubiertas de cartílago articular. El grado de movilidad de este tipo de articulaciones depende de la forma de las superficies de contacto, de esta forma podemos subclasificar las articulaciones sinoviales en uniaxiales, si se mueven alrededor de un único eje, biaxiales si pueden moverse en dos, o multiaxiales si permiten el movimiento en tres ejes (Fig. 1).

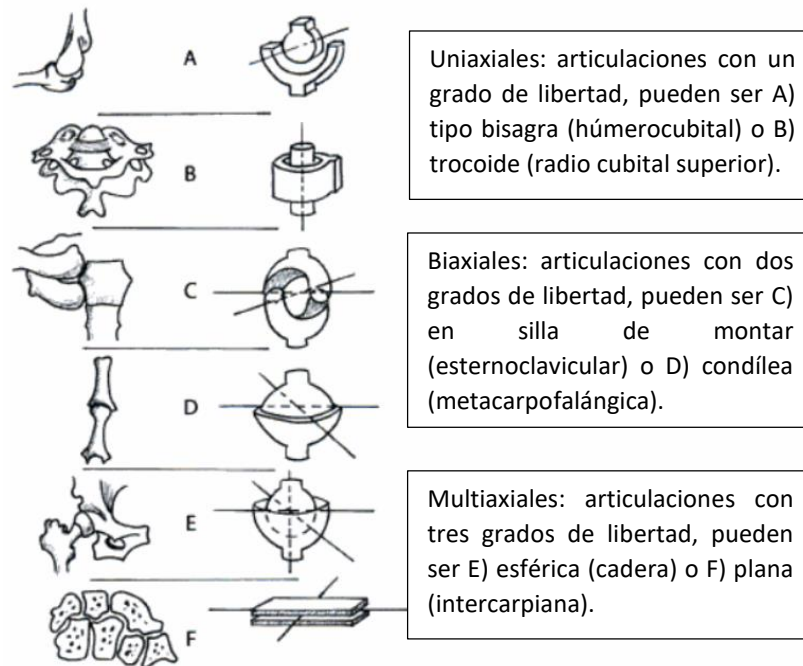


Figura 1. Clasificación de los tipos de articulaciones sinoviales según el grado de movilidad. Imagen modificada de de Toro, 2010.

Articulación sinovial.

La estructura de la articulación sinovial está compuesta por diversos tipos de elementos, algunos de ellos están presentes en la figura 2, donde se observa un esquema de la estructura de la rodilla humana. Los elementos articulares se pueden clasificar en cinco grupos, los de protección, los de amortiguación de presiones, contacto y deslizamiento, los elementos de adaptación de las superficies articulares, los elementos de mantenimiento y, por último, la cavidad articular.

1. Elementos de protección: éstos constituyen la cápsula articular o bursa, cuya morfología es similar a un manguito que se extiende entre los elementos óseos y se continúa con el periostio de éstos. La capa más externa está formada por una membrana fibrosa que es muy resistente y poco elástica perforada por vasos y nervios articulares, es rica en fibras de colágeno y pobre en células. La capa interna es la membrana sinovial, la cual es delgada, transparente y es la encargada de secretar el líquido sinovial. La membrana sinovial también presenta dos láminas, una subintimal, formada por tejido conjuntivo y que está en contacto con la membrana fibrosa externa, y una lámina celular íntima que contiene los sinoviocitos.
2. Elementos de amortiguación de presiones, contacto y deslizamiento:

## Introducción

- a. Cartílago articular. En el próximo apartado se profundizará con mayor detalle en la descripción de este elemento, el cual se caracteriza por su elevada capacidad viscoelástica, lo que permite que este tejido posea una gran capacidad de amortiguación de presiones y, además, favorezca el deslizamiento entre superficies articulares.
  - b. Líquido sinovial. Está formado a partir de filtrado de plasma sanguíneo y contiene unas 300 células por  $\text{mm}^3$  (entre ellas monocitos, linfocitos y sinoviocitos) y ácido hialurónico. Es el elemento articular encargado de la lubricación y nutrición del cartílago articular y, además de estar presente en la cavidad articular, también lo está en las bolsas serosas y en las vainas de los tendones.
  - c. Rodete, menisco y disco. Son estructuras clave en el proceso de deslizamiento articular y la correcta amortiguación y distribución de las cargas que sufren las distintas articulaciones del cuerpo. El rodete es una estructura particular del hombro, el disco articular es propio de la articulación temporomandibular (Matamala *et al*, 2006) y finalmente los meniscos son fundamentales en la rodilla.
3. Elementos de adaptación de las superficies articulares: facilitan la congruencia de las superficies articulares. Entre ellos vuelven a aparecer elementos de deslizamiento articular como el rodete articular, el cual es un fibrocartilago que aumenta la estabilidad articular. Otro ejemplo sería el disco articular, que es un tabique intraarticular fibrocartilaginoso adherido a la cápsula articular, esta estructura aumenta la estabilidad articular y contribuye en la amortiguación de las presiones.
  4. Elementos de mantenimiento o estabilidad articular: entre ellos está presente la membrana fibrosa de la cápsula articular, la cual otorga elevada robustez a la articulación. También están presentes en este grupo los ligamentos, que son láminas fibrosas que afianzan las articulaciones, y los tendones de los músculos.
  5. La cavidad articular: este último elemento está limitado por los cartílagos articulares y la membrana sinovial. Está relleno de una moderada cantidad de líquido sinovial.

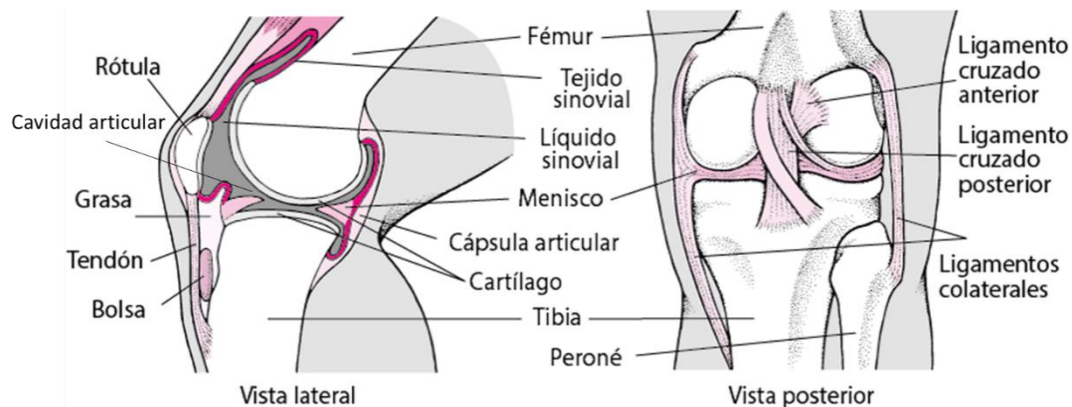


Figura 2. Esquema de los elementos articulares que componen la rodilla. Imagen modificada de: <https://www.msmanuals.com/es/professional/trastornos-de-los-tejidos-musculares-y-conectivo/evaluación-del-paciente-con-síntomas-articulares>.

### *Cartílago articular, estructura y función.*

El cartílago articular es un tejido de tipo hialino presente en las articulaciones diartrodiales. Está desprovisto de vasos sanguíneos, vasos linfáticos y nervios. Su superficie tiene una transparencia variable y un color blanquecino. Su grosor suele ser de entre 3 o 4 mm,

aunque éste puede variar en las articulaciones que soportan mayores cargas, por ejemplo el cartílago patelar presente en la rodilla cuenta con mayor grosor, rondando los 6 o 7 mm (Espanha, 2010; Brody, 2015). La nutrición de este tejido depende fundamentalmente de la difusión del líquido de la cavidad articular, proceso favorecido gracias al propio movimiento de la articulación (Muñíos, 2010).

Los componentes y la estructura del cartílago son fundamentales para que realice sus dos funciones principales: deformarse al sufrir una carga mecánica y actuar como una suave superficie de deslizamiento en las articulaciones. Esto permite a la articulación absorber el estrés mecánico y realizar su movimiento con bajos niveles de fricción (Creamer y Hochberg, 1997).

El cartílago articular está formado por un reducido número de células, los condrocitos (2-3% del cartílago), y una matriz extracelular del cartílago (MEC) que los rodea (97- 98% restante) (Benito, 2010). Los condrocitos representan el único elemento celular presente en el cartílago y son los encargados de la síntesis del colágeno y otras macromoléculas, así como de la reparación y renovación de la MEC. Estas células no están repartidas de forma equitativa por todo el cartílago, sino que su abundancia disminuye con la profundidad. La unidad funcional básica de este tejido es el condrón, formado por el condrocito, el lugar que ocupa en la matriz (laguna), la matriz pericelular sintetizada por el propio condrocito, y una capa de fibras colágenas que otorgan protección mecánica. La MEC consta de dos fases:

1. Fase líquida: constituye el 80% del peso total del cartílago y está compuesta de agua. La distribución de este componente varía según la profundidad, presentando mayor concentración de agua en la superficie. Su interacción con las macromoléculas de la fase sólida de la matriz proporciona al cartílago elasticidad, y además permite la nutrición de los condrocitos y una mayor lubricación de la articulación.
2. Fase sólida: representa aproximadamente el 20% del peso total del cartílago. Está formada por un armazón de fibras de colágeno y un gel constituido principalmente por agregados de proteoglicanos (PG), otras glucoproteínas y proteínas no colágenas. El colágeno más abundante es el tipo II (representando el 90-95% de todas las fibras de colágeno del cartílago), que es sintetizado por los condrocitos, y forma un entramado en forma de malla que proporciona a la matriz resistencia a la tensión. El PG más importante es el agregano, una macromolécula con un eje central proteico al que se conectan glucosaminoglucanos (GAG) como el ácido hialurónico (HA del inglés *hyaluronic acid*), cuya función principal es conferir elasticidad al cartílago. El agregano es sintetizado también por los condrocitos.

La estructura del cartílago articular está representada en la figura 3, y está formado por las siguientes capas (Benito 2010, Brody, 2015):

1. Zona superficial o tangencial: en esta zona los condrocitos presentan forma alargada y, además, posee el mayor contenido de colágeno ya que está especializada en el deslizamiento. En esta capa, las fibras de colágeno se orientan paralelamente con respecto a la superficie, y esta distribución aumenta la resistencia a fuerzas de rozamiento. Es la zona más rica en células y fibras, pero los condrocitos están casi metabólicamente inactivos, lo que implica que la reparación de esta capa sea difícil; por tanto, una lesión en esta zona puede exponer a las capas inferiores a fuerzas de desgaste.
2. Zona media o de transición: representa entre el 40% y el 60% del grosor total del cartílago. Las fibras de colágeno tienen un diámetro más ancho que en la capa superficial

## Introducción

y se orientan de forma oblicua a la superficie del tejido. Las fibras están distribuidas aleatoriamente junto con los condrocitos, los cuales presentan una morfología esférica en esta capa, son de mayor tamaño que en la capa superficial y, además, metabólicamente hablando, son más activos, lo que sugiere una mayor capacidad de regeneración.

3. Zona profunda o radial: posee unas fibras de colágeno más alargadas en comparación con las capas superiores. Los condrocitos se distribuyen verticalmente con respecto al hueso subcondral, continúan presentando morfología esférica y se organizan en columnas de dos a seis células. Esta capa presenta un elevado contenido en PG y un menor contenido en agua. Esta peculiaridad hace que esta capa interna tenga una baja capacidad para combatir fuerzas de rozamiento, sin embargo, es muy eficaz dispersando las fuerzas de compresión gracias a la disposición vertical de las fibras de colágeno y al rico contenido en PG.
4. *Tide mark*: Zona basófila que marca el límite entre zona profunda y zona calcificada.
5. Zona calcificada: aquí la matriz se encuentra mineralizada y el contenido en PG está muy reducido. Los condrocitos son muy escasos, grandes e hipertróficos. Esta zona marca el límite entre el cartílago y el hueso subcondral: por encima de ella los nutrientes son distribuidos por difusión del líquido sinovial y, por debajo, los nutrientes son transportados desde los vasos sanguíneos del hueso subcondral. Durante la adultez, el cartílago articular no es capaz de obtener nutrientes de los vasos del hueso, dejando al cartílago dependiente solo del fluido sinovial y, comprometiendo así la capacidad de reparación del tejido.

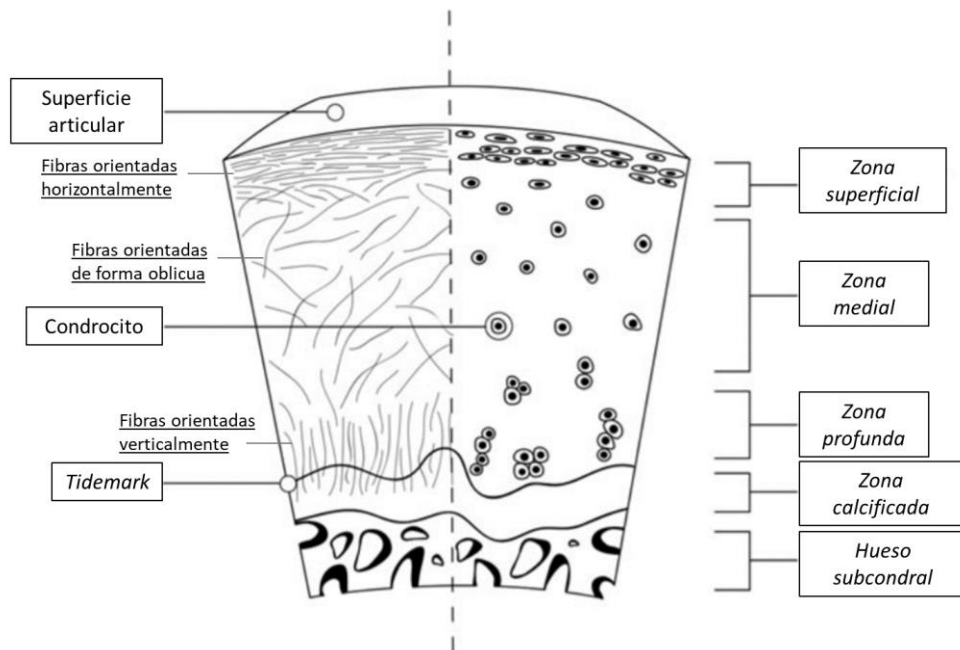


Figura 3. Esquema representativo de la estructura en capas del cartílago articular. Imagen modificada de [https://link.springer.com/chapter/10.1007/978-3-319-44785-8\\_1](https://link.springer.com/chapter/10.1007/978-3-319-44785-8_1).

### Mantenimiento del cartílago articular.

Los condrocitos son los principales encargados del mantenimiento de la MEC. En condiciones normales estas células mantienen un equilibrio entre procesos anabólicos (síntesis) y catabólicos (degradación) de la MEC, promoviendo así la constante renovación de esta estructura. Para poder mantener su función, el condrocito necesita glucosa, aminoácidos,

vitaminas y trazas de minerales esenciales, y estos nutrientes deben difundir desde la membrana sinovial y atravesar la MEC para poder llegar hasta el condrocito.

Los condrocitos son células altamente glicolíticas, lo que quiere decir que obtienen una gran parte de su energía a partir del metabolismo de la glucosa. Para mantener una adecuada homeostasis celular es indispensable un suministro regular de estos nutrientes, con lo que cambios en las reservas de glucosa o en el microambiente extracelular del condrocito pueden llevar a la inactivación de las actividades del condrocito, promoviendo así patologías articulares (Mobasher *et al.*, 2002). Diversos factores ambientales/mecánicos y/o bioquímicos pueden producir alteraciones en el equilibrio metabólico de los condrocitos, favoreciendo los procesos catabólicos frente a los anabólicos y degradando así el cartílago (Umlauf *et al.*, 2010).

1. Los factores mecánicos son debidos a la sobrecarga que sufren las articulaciones. Dicho proceso puede generar desgaste por interacción de las superficies de la articulación, apareciendo abrasión y arañazos, o también puede generar desgaste por fatiga, produciendo la deformación de las superficies en contacto y acumulación de daños microscópicos (Trujillo, 2002). A pesar de que la carga mecánica en la articulación puede producir daños en ésta, es un proceso fundamental para la generación de estímulos biomecánicos en los condrocitos y la nutrición de estos, por tanto, una inmovilización de la articulación supondría cambios en la composición del cartílago articular (Musumeci, 2016).
2. Los factores bioquímicos hacen referencia a elementos que alteran el equilibrio entre los factores que son liberados por los condrocitos. Para realizar tanto su función catabólica como la anabólica estas células pueden sintetizar y secretar una gran variedad de compuestos. Para favorecer el anabolismo secretan distintos factores de crecimiento, entre ellos el factor de crecimiento insulínico (IGF, del inglés *insulin growth factor*), el factor de crecimiento transformante  $\beta$  (TGF $\beta$  del inglés *Transforming growth factor*  $\beta$ ), el factor de crecimiento derivado de plaquetas (PDGF del inglés *platelet derived growth factor*) y el factor de crecimiento fibroblástico (FGF, del inglés *fibroblast growth factor*). Estos elementos aumentan la síntesis de colágeno II, la síntesis de PG y la síntesis de inhibidores de metaloproteasas (TIMPs del inglés *tissue inhibitor of matrix metalloproteinase*), y aumentan también la proliferación celular. En cambio, bajo determinados estímulos, como por ejemplo procesos inflamatorios, los condrocitos son capaces de decantarse hacia el catabolismo del cartílago, pudiendo sintetizar distintas citoquinas, prostaglandinas, óxido nítrico (NO), proteasas y glicosidasas (Trujillo, 2002).

En la artrosis tiene lugar un proceso inflamatorio que favorece la expresión de factores bioquímicos pro-inflamatorios, los cuales desequilibran el metabolismo de los condrocitos hacia la vía catabólica. Entre estos elementos, los más predominantes son la interleuquina-1 $\beta$  (IL1 $\beta$ ) y el factor de necrosis tumoral  $\alpha$  (TNF $\alpha$  del inglés *tumor necrosis factor*  $\alpha$ ). Estas moléculas proinflamatorias son citoquinas sintetizadas mayoritariamente por los sinoviocitos presentes en la membrana sinovial, aunque también, en menor medida por los condrocitos, y estimulan cambios fenotípicos en los condrocitos que fomentan la producción de estas moléculas de forma autocrina o paracrina (Chou *et al.*, 2020). Ambas moléculas favorecen la vía catabólica ya que inducen la síntesis de metaloproteasas (MMPs), agregasas, NO, prostaglandinas, así como otras citoquinas como las interleuquinas 8 y 6 (IL8 e IL6), y, además, también son capaces de reducir la síntesis de los TIMP (Umlauf *et al.*, 2010; Vela, 2015).

Como resultado de este proceso inflamatorio, la degradación del cartílago articular se ve favorecida frente a la síntesis, dando lugar a daños en el cartílago. El condrocito articular

## Introducción

juega un papel clave durante el proceso artrósico. Otro problema añadido es que el número de condrocitos en el tejido es limitado, ya que tras la madurez esquelética no suelen dividirse. Cualquier reducción en el número de condrocitos o en su función puede afectar significativamente a la salud del cartílago articular; es normal con el proceso de envejecimiento que el volumen de condrocitos disminuya, además, en edades avanzadas los condrocitos producen agregados más pequeños e irregulares. Estas consecuencias del proceso de envejecimiento comprometen la función de los condrocitos, afectando a los procesos de reparación del tejido, de esta forma las células están más expuestas al daño mecánico, generando así un ciclo vicioso. El envejecimiento no es la única amenaza que puede comprometer la función del condrocito, sino que procesos tales como la disfunción mitocondrial, la senescencia celular o la producción de radicales libres de oxígeno (ROS, del inglés *reactive oxygen species*) son elementos que pueden producir apoptosis celular, inflamación, daños en diversas estructuras del condrocito o incluso alterar su capacidad de generar componentes de la matriz en buen estado (Brody, 2015).

Es importante destacar que durante el avance de la artrosis, el condrocito también está expuesto a múltiples factores bioquímicos, como se ha comentado, e inicialmente modificará su fenotipo a uno activo o maduro para aumentar su actividad anabólica, en un intento por reparar el cartílago articular (Fukui *et al.*, 2008). Conforme avanza el proceso artrósico, el condrocito comenzará a desdiferenciarse y a expresar distintos genes de matriz que no se expresan de forma natural en el cartílago articular, como por ejemplo el colágeno I y el colágeno III. El fenotipo celular más común durante la artrosis es el condrocito hipertrófico, que se caracteriza por producir principalmente colágeno X. Los condrocitos hipertróficos expresan genes relacionados con la diferenciación osteogénica y producen proteínas de la MEC mineralizadas, produciendo la calcificación del propio condrocito (Charlier *et al.*, 2019; Rim *et al.*, 2020). Además del fenotipo hipertrófico, existen otros fenotipos presentes en el cartílago, como el desdiferenciado o fibroblástico, que se caracteriza por secretar principalmente colágeno tipo I y II.

### *Factores de riesgo.*

El desarrollo y la progresión de la artrosis depende de múltiples factores que pueden actuar de forma simultánea; dichos factores de riesgo se pueden clasificar en factores personales y factores propios de la articulación. Los factores personales pueden subdividirse en factores modificables y no modificables.

#### Factores modificables

- **Obesidad:** un índice de masa corporal (IMC) mayor a 30kg/m<sup>2</sup> se ha asociado de forma robusta con la artrosis de rodilla en diversos estudios como el de Silverwood y colaboradores (Silverwood *et al.*, 2015). Otro estudio reciente también ha mostrado el posible papel sinérgico que pueden ejercer la obesidad junto con la depresión en la progresión de la artrosis de rodilla (Jacobs *et al.*, 2020). En dicho estudio observaron que los pacientes obesos con artrosis y depresión desarrollan mayores niveles de dolor y de pérdida de funcionalidad articular, de marcadores de deterioro articular y de remodelación ósea con respecto a pacientes con artrosis y obesidad, pero sin depresión. A pesar de su fuerte asociación con la artrosis de rodilla, también se asocia con la artrosis de manos y en menor medida con la artrosis de cadera, pudiendo además favorecer la progresión de la enfermedad. La obesidad promueve la artrosis de rodilla y cadera mediante dos mecanismos, por un lado hay un potente efecto biomecánico sobre ambas



articulaciones, donde el exceso de peso genera en ellas una importante sobrecarga, produciendo la rotura del cartílago y un fallo de las estructuras de soporte (Felson *et al.*, 2000). Por otro lado, la obesidad también puede favorecer el desarrollo de la artrosis promoviendo cambios metabólicos que favorezcan el proceso inflamatorio de la articulación, provocando así daños en el cartílago articular (Misra *et al.*, 2019). Esta última faceta de la obesidad es especialmente relevante en otras articulaciones como por ejemplo la mano, la cual no tiene una sobrecarga mecánica tan importante, sin embargo el modo de acción de la obesidad es principalmente a través de efectos metabólicos e inflamatorios (Carman *et al.*, 1940).

- Dieta: la dieta típica occidental consiste en el consumo elevado de carne roja, productos lácteos ricos en grasas y granos refinados, los cuales se asocian con mayores niveles de proteína C reactiva e IL6, lo que promueve la inflamación sistémica. Por otro lado, la dieta mediterránea consiste en un elevado consumo de frutas, vegetales, legumbres, nueces, semillas, cereales, pescado y marisco, junto con un consumo moderado de productos lácteos, carne de ave de corral, huevos, vino tinto y aceite de oliva, lo que conlleva menores niveles de inflamación. En 2019 Veronese y colaboradores (Veronese *et al.*, 2020) realizaron un estudio longitudinal de 4 años de seguimiento en una cohorte de pacientes de 4.330 sujetos, y pudieron observar que aquellos individuos con mayor adherencia a la dieta mediterránea mostraban menores niveles de dolor, menor empeoramiento y sintomatología de la artrosis de rodilla. En la revisión sistemática de 2018 de Ivorra y colaboradores (Morales-Ivorra *et al.*, 2018), estos autores también concluyeron que la adherencia a esta dieta reducía la prevalencia de la artrosis, y que también se reducía tanto la presencia de biomarcadores de inflamación como de la IL1 $\beta$ . Los beneficios de la dieta mediterránea están en parte motivados por la presencia de polifenoles, que previenen la inflamación y la destrucción del cartílago, así como de los ácidos grasos omega 3, que son capaces de reducir los niveles de inflamación sistémicos (Schmitz y Ecker, 2008; Morales-Ivorra *et al.*, 2018).
- Ocupación laboral: el uso repetitivo de la articulación puede aumentar el riesgo de desarrollar artrosis. Aquellos individuos con puestos de trabajo que requieran agacharse o arrodillarse tienen el doble de riesgo de desarrollar artrosis de rodilla que aquellos que no necesitan actividad física, sobre todo en aquellas personas que presentan sobrepeso o cuyos trabajos impliquen el levantamiento de cargas elevadas (Messier *et al.*, 2009). También, aquellos trabajos que requieran estar durante elevados períodos de tiempo de pie estáticos o levantando peso se asocian con artrosis de cadera, y aquellos trabajos que requieren destreza manual se asocian con artrosis de manos (Johnson y Hunter, 2014).

### Factores no modificables

- Edad: la relación entre la artrosis y la edad es clara pero no está totalmente caracterizada, esto es debido a que es una relación multifactorial, donde intervienen el daño oxidativo, el estrechamiento del cartílago articular, una reducción en la fuerza muscular y la reducción de la propiocepción del individuo. A esto hay que añadirle que los mecanismos celulares que mantienen la homeostasis del tejido empeoran con el paso del tiempo, produciendo respuestas inadecuadas al estrés o una lesión articular resultando en una destrucción y pérdida del tejido articular (Litwic *et al.*, 2013).
- Sexo: el sexo femenino se asocia con mayor prevalencia y severidad de la artrosis, presentando mayor afectación de artrosis en las manos, pies y rodillas con respecto a los hombres. Además, las mujeres tienden a sufrir artrosis de rodilla radiográficamente

## Introducción

más severa que los hombres, especialmente después de la menopausia (Srikanth *et al.*, 2005). Una posible explicación es que los estrógenos pueden tener un papel clave en el desarrollo de la enfermedad, pudiendo tener un efecto directo en el metabolismo del cartílago mediante la regulación de citoquinas proinflamatorias y de los factores de crecimiento y, también, promoviendo cambios en el hueso subcondral (Nevitt *et al.*, 2001). Otro motivo causante de esta disparidad entre sexos pueden ser las diferencias en la fuerza ósea, el alineamiento articular, el embarazo o la fuerza neuromuscular (Johnson y Hunter, 2014).

- Genética: este elemento juega un papel fundamental en el desarrollo de la artrosis; se estima que la genética condiciona en un 60% el desarrollo de la artrosis de cadera y en un 40% la artrosis de rodilla (Palazzo *et al.*, 2016). En 2021, con el uso de metodologías de asociación de genoma completo (GWAS, del inglés *genome-wide association studies*), se habían identificado un total de 124 polimorfismos de un solo nucleótido (SNPs, del inglés *Single Nucleotide Polymorphisms*) asociados con el riesgo de desarrollar artrosis, siendo uno de los más robustos en la artrosis de rodilla el SNP nuclear rs143383 del gen GDF5 (del inglés *Growth Differentiation Factor 5*). En general, el riesgo de artrosis estimado para cada uno de estos 124 SNPs es muy pequeño, por lo que se piensa que el desarrollo de la artrosis se ve condicionado por el efecto conjunto de múltiples polimorfismos que actúan de forma simultánea (Aubourg *et al.*, 2022). Un estudio muy reciente incrementó este número (Boer *et al.*, 2021). En dicho estudio se trabajó con un total de 13 cohortes internacionales pertenecientes a 9 poblaciones distintas y, sumó un total de 826.690 individuos, de los cuales 177.517 eran pacientes de artrosis. Como resultado, además de volver a encontrar las variantes ya mencionadas, encontraron un total de 52 nuevos polimorfismos que no se habían asociado previamente con la artrosis.

En cuanto a los factores propios de la articulación, destacan los siguientes elementos:

- Lesión previa: las más comunes en deportistas de alto nivel son las que afectan a estructuras de soporte, como la rotura del ligamento cruzado anterior o las lesiones en las estructuras de distribución de cargas compresivas, los meniscos. La rotura del ligamento cruzado anterior suele ir acompañada de daño en el cartílago articular, el hueso subcondral, los ligamentos colaterales y daño en los meniscos (Von Porat *et al.*, 2004). El daño en los meniscos produce un mal reparto de las cargas articulares, y a la larga promueve la progresión de la artrosis de rodilla (Walker *et al.*, 2011). Como resultado de estos procesos tiende a favorecerse el desarrollo de la artrosis y, aproximadamente el 50% de los individuos que sufren este tipo de lesiones suelen desarrollar la enfermedad mostrando dolor e incapacidad funcional pasados de 10 a 20 años tras la lesión (Lohmander *et al.*, 2007).
- Fuerza muscular: el debilitamiento del cuádriceps puede tener un papel clave en el desarrollo de la artrosis de rodilla. El déficit de fuerza muscular y de la activación muscular junto con la pérdida de propiocepción suelen ser comunes en pacientes con artrosis de rodilla, y pueden ocurrir simultáneamente como consecuencia de la artrosis. Además, la pérdida de fuerza muscular no solo se asocia con el desarrollo de la artrosis, sino que también se relaciona con la progresión de la artrosis de rodilla (Bennell *et al.*, 2013).

- Alineamiento: el alineamiento incorrecto de la sección inferior de la pierna tanto en dirección valgo como varo (Fig. 4) influye en la distribución de las cargas en la superficie articular de la rodilla. Posibles causas de un mal alineamiento de la rodilla pueden ser tanto genéticas como defectos durante el desarrollo o incluso eventos traumáticos, como por ejemplo la rotura del ligamento cruzado anterior o daños en los meniscos. El alineamiento incorrecto se ha asociado de forma robusta con la progresión en la artrosis de rodilla, pudiendo ser un factor clave en el desarrollo de esta enfermedad si se combina con elevados niveles de IMC. Concretamente el alineamiento varo se asocia con riesgo de progresión en la zona medial, mientras que el valgo, se asocia con riesgo en la zona lateral de la articulación (Tanamas *et al.*, 2009).

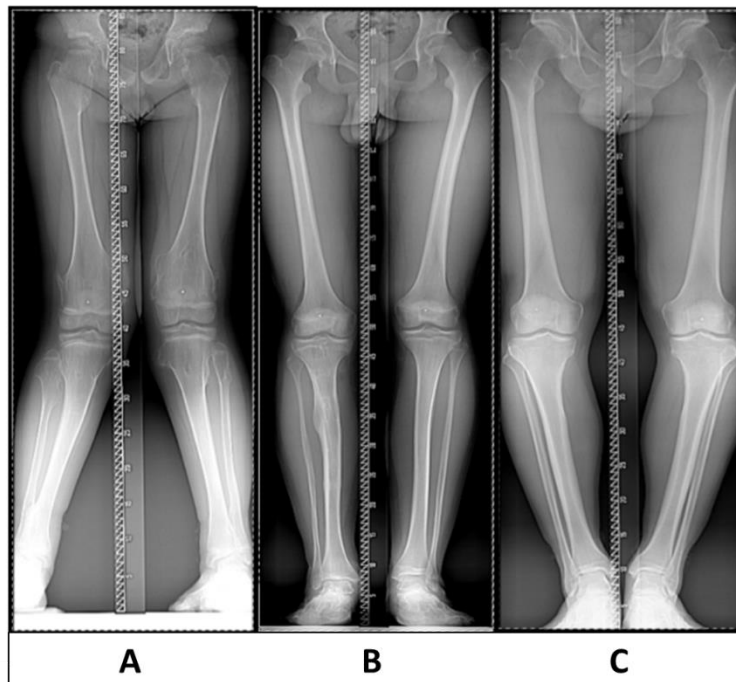


Figura 4. Tipos de mal alineamiento de rodilla. A: desviación valga; B: alineamiento normal; C: desviación vara. Imagen modificada de <https://epos.myesr.org/poster/esr/ecr2013/C-1044/background>.

#### Fenotipos de artrosis.

La búsqueda de un tratamiento para la artrosis tiene como finalidad obtener una mejoría funcional articular, reducir el dolor y frenar la progresión de la enfermedad. El objetivo es incluir nuevos fármacos dentro del grupo denominado DMOADs (del inglés *disease modifying osteoarthritis drug*), los cuales, como su nombre indica, tendrían la capacidad de interferir en la progresión de los cambios articulares (Ruiz, 2005). Diversos potenciales DMOADs han sido propuestos como efectivos en estudios tempranos y preclínicos, pero, por desgracia, han fallado a la hora de alcanzar criterios de valoración estructurales y clínicos en la fase 3 de los ensayos clínicos. Una de las posibles causas es que la artrosis es una enfermedad muy heterogénea, donde existen muchos factores que pueden afectar al desarrollo de la enfermedad, y por lo general las intervenciones terapéuticas se centran en uno o pocos de estos mecanismos, por lo que puede no ser efectivo si la causa subyacente no concuerda con el objetivo del tratamiento (Van Spil *et al.*, 2019).

En los últimos años, diversos grupos de investigación de esta área de la medicina han centrado sus esfuerzos en tratar de clasificar a la artrosis en una serie de fenotipos con el fin de

## Introducción

poder enfrentar esta heterogeneidad. Un fenotipo se define como una serie de características visibles en un individuo que son el resultado de la interacción entre su genotipo (información genética) y el ambiente en el que vive. En la pasada década diversos autores (Bijlsma *et al.*, 2011; Sellam y Berenbaum, 2013) establecieron una clasificación fenotípica muy simple basada en factores de riesgo, creando un total de cinco fenotipos de artrosis:

1. Artrosis asociada a lesión previa: típica en pacientes jóvenes, generalmente menores de 45 años. La principal causa es el estrés mecánico debido a la lesión, y afecta principalmente a la rodilla, el pulgar, el hombro y el tobillo.
2. Artrosis metabólica: afecta a pacientes de mediana edad de entre 45 y 65 años. La presencia de desórdenes metabólicos asociados al paciente, como por ejemplo la diabetes mellitus de tipo 2, incrementan los niveles de elementos pro-inflamatorios como las ROS, favoreciendo así el proceso artrósico (Schett *et al.*, 2013). Este fenotipo de artrosis afecta especialmente a la rodilla y a la mano.
3. Artrosis asociada al envejecimiento: presente en individuos con edades superiores a los 65 años. La causa principal son los daños asociados al envejecimiento y la senescencia del condrocito articular. Afecta a las tres principales articulaciones sinoviales, rodilla, cadera y mano.
4. Artrosis de origen genético: a pesar de no estar considerada como una enfermedad hereditaria, tanto la genética como las modificaciones epigenéticas juegan un papel clave en el desarrollo de esta patología. Este tipo de artrosis no afecta a un grupo de edad específico y ocurre principalmente en la mano, la rodilla, la cadera y en la espina dorsal.
5. Artrosis asociada a dolor: este fenotipo tampoco afecta a un grupo de edad específico. La causa son procesos inflamatorios, remodelaciones óseas y percepción aberrante del dolor. Afecta a la rodilla, la cadera y la mano.

Sin embargo, los mecanismos funcionales que se encuentran detrás de estos fenotipos no son exclusivos, sino que entre ellos pueden compartir rutas de inflamación, como es el caso de la inflamación leve crónica (Berenbaum, 2013). Hay que tener en cuenta que, de forma general, el desarrollo de artrosis en un individuo no suele depender de un único factor, sino de la combinación de varios, lo cual provoca que la clasificación basada en factores de riesgo comience a estar obsoleta. Para una correcta asignación del fenotipo es necesario incluir en las bases de datos clínicas toda una serie de variables muy diversas y variables de respuesta bien definidas. Algunos de los ejemplos de datos que se pueden incluir son las variables socio-demográficas, las variables clínicas, datos de imagen, marcadores biomecánicos, medidas mecánicas e incluso datos genéticos u ómicos. Para poder definir fenotipos basándonos en esta gran cantidad de información, tanto clínica como biológica, es necesario aplicar aproximaciones metodológicas de gran computación, como “*Big Data*” o “*Machine learning*” (Deveza *et al.*, 2019). En esta línea, de la mano de Dell’Isola y colaboradores (Dell’Isola *et al.*, 2016) surgió un nuevo tipo de clasificación de los fenotipos de artrosis de rodilla. En este trabajo analizaron en la literatura, desde el año 1984 hasta 2016, todas aquellas variables y *sets* de información que se asociaron con diversos fenotipos de artrosis, y tuvieron en cuenta un total de 841 trabajos de investigación de los cuales incluyeron 25 en el estudio. Como resultado de los análisis de este estudio fueron capaces de establecer seis *sets* de variables, representando cada una un fenotipo de artrosis:

1. Fenotipo de dolor crónico: caracterizado por alteraciones del sistema nervioso central y del dolor como papel clave en la fisiopatología de la enfermedad.

2. Fenotipo inflamatorio: caracterizado por una elevada inflamación crónica.
3. Fenotipo asociado al síndrome metabólico: típico de pacientes con desórdenes metabólicos que incluyen sobrepeso, diabetes, hipertensión o dislipemia. Como resultado se produce, por un lado, un estrés mecánico que se asocia directamente con la artrosis de rodilla y, por otro lado, un estrés metabólico, que podría ser una de las causas de la artrosis de manos (Courties *et al.*, 2019)(Fig. 5)
4. Fenotipo de metabolismo óseo y del cartílago: caracterizado por alteraciones en el metabolismo óseo y del cartílago articular.
5. Fenotipo de sobrecarga mecánica: típico en pacientes con sobrepeso, lesión previa o mal alineamiento de la articulación.
6. Fenotipo de enfermedad articular mínima: presente en pacientes con bajos niveles de deterioro articular, síntomas leves y baja progresión en el tiempo.

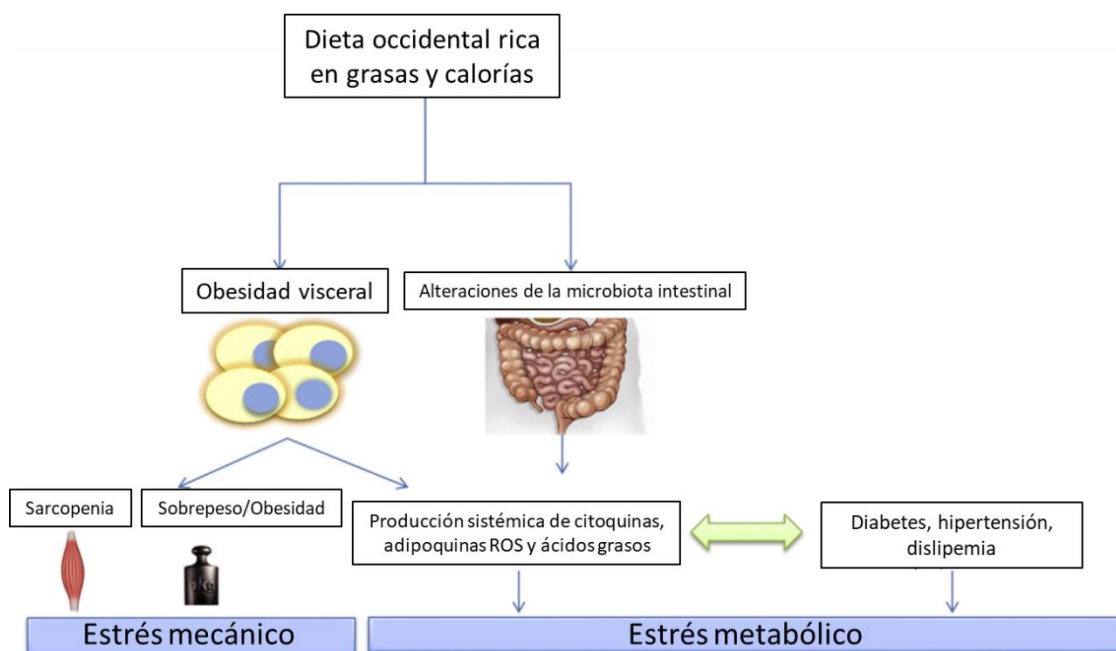


Figura 5. Efecto de la dieta en el estrés y desarrollo de patologías. Una dieta hipercalórica rica en grasas puede producir desórdenes sistémicos. Por un lado, se fomenta la obesidad visceral, el exceso de IMC incrementa las cargas que sufren las articulaciones y también puede producir alteraciones musculares, favoreciendo el debilitamiento de la musculatura, promoviendo así el estrés mecánico. Por otro lado, se pueden producir alteraciones en la microbiota intestinal, favoreciendo la presencia de bacterias proinflamatorias que, junto con el exceso de tejido adiposo, pueden promover la producción sistémica de factores inflamatorios, fomentando el estrés metabólico. Finalmente, la presencia de comorbilidades asociadas con el estrés metabólico como la diabetes, la hipertensión y la dislipemia también pueden incrementar todavía más este estrés metabólico, aumentando notoriamente el riesgo de desarrollar la artrosis. Imagen modificada de (Courties, Berenbaum and Sellam, 2019).

A día de hoy no existe un consenso en la clasificación de fenotipos de artrosis, pero el nuevo enfoque que se ha tomado es bastante prometedor. Si finalmente se consigue hacer una clara distinción de estos fenotipos, se podría desarrollar y aplicar terapias efectivas. Esta ambiciosa aproximación ya se ha completado en patologías como el asma (Borroni y Benussi, 2019), en la cual ya se están desarrollando ensayos con biomarcadores capaces de identificar casos severos de asma, y otras como la enfermedad pulmonaria obstructiva crónica, fallo cardíaco e incluso en el cáncer de mama (Deveza *et al.*, 2019).

## Introducción

### *Clasificación radiográfica de la artrosis.*

La artrosis es una enfermedad que progresa con el paso del tiempo, alterando poco a poco la articulación afectada. Debido a los cambios que esta enfermedad produce, fue necesario el desarrollo de un sistema capaz de identificar el estado patológico y la severidad que presentaba la articulación afectada en un momento dado. Debido a esto, se desarrollaron diversas clasificaciones de la artrosis y, en los próximos subapartados, se comentarán las principales escalas de clasificación empleadas actualmente.

➤ Escala Kellgren-Lawrence.

Ésta es la escala que ha sido usada de forma más tradicional para clasificar la artrosis en diversos niveles basados en el grado de deterioro de la estructura articular. Fue definida en 1957 por Kellgren y Lawrence (Kellgren y Lawrence, 1957). En dicha publicación establecieron un sistema de grados para la artrosis de mano, espina cervical, espina lumbar, cadera, rodilla y pie. Las características radiológicas que tuvieron en cuenta para llevar a cabo este sistema fueron las siguientes:

1. Formación de osteofitos en los márgenes de la articulación o, en el caso de artrosis de rodilla, en las espinas tibiales.
2. Formación de pequeños huesos (osículos) alrededor de las articulaciones interfalángicas distales y proximales.
3. Estrechamiento del cartílago articular, asociado con la esclerosis del hueso subcondral.
4. La aparición de pequeñas áreas pseudocísticas con paredes esclerotizadas en el hueso subcondral.
5. Deformación de la zona final de los huesos, especialmente en la cabeza del fémur.

En base a estas características radiológicas fueron capaces de clasificar la severidad de la artrosis en cinco grados KL (Fig. 6):

1. Grado 0 (sin evidencias de artrosis): ausencia clara de cambios radiológicos típicos de la artrosis.
2. Grado 1 (artrosis dudosa): cambios dudosos en el estrechamiento del espacio intra-articular y posibles osteofitos de tamaño muy reducido.
3. Grado 2 (artrosis mínima): osteofitos de pequeño tamaño claramente definidos y posible reducción el espacio intra-articular. A partir de este punto se considera que la articulación sufre de artrosis radiográfica.
4. Grado 3 (artrosis moderada): aparición de múltiples osteofitos de tamaño pequeño, clara reducción del espacio intra-articular, leve aparición de esclerosis del hueso subcondral y posible deformación de las zonas finales de los huesos.
5. Grado 4 (artrosis severa): osteofitos de gran tamaño, una remarcada reducción del espacio intra-articular, esclerosis severa y clara deformación de la zona final de los huesos.

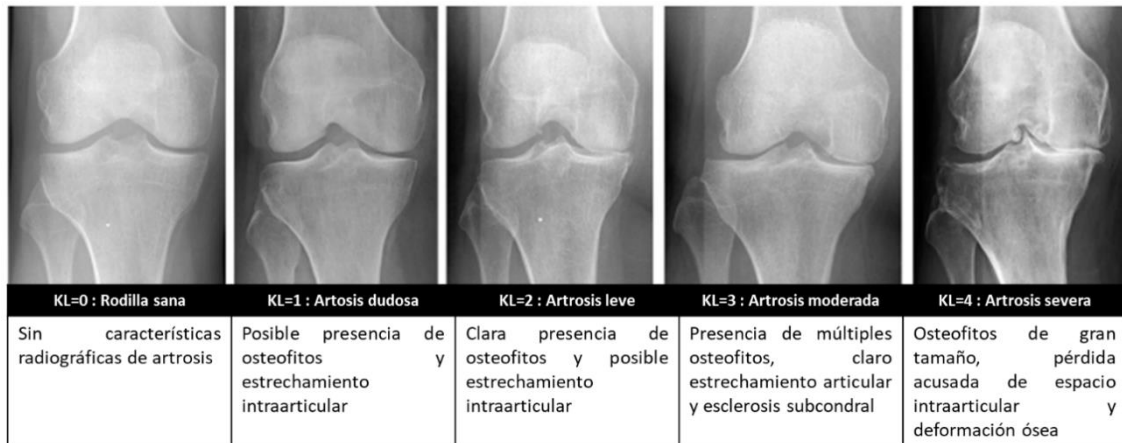


Figura 6. Esquema de características y radiografías representativas de los distintos grados de artrosis de la escala KL. Imagen modificada de (Bayramoglu, Nieminen and Saarakkala, 2020).

➤ Escala Osteoarthritis Research Society International (OARSI)

Otra escala empleada habitualmente para la clasificación del grado de artrosis según la severidad es la de la OARSI. Fue creada en 1996 (Altman *et al.*, 1995) y revisada en 2007 (Altman y Gold, 2007). Como ya se comentó anteriormente, en la escala KL se definen cinco grados de severidad usando una combinación de la presencia de osteofitos, el estrechamiento del espacio intra-articular y la esclerosis subcondral. La escala OARSI usa mediciones semicuantitativas separadas para el estrechamiento intra-articular y la formación de osteofitos y, de esta forma, es capaz de graduar ambas características de forma independiente en una escala del 0 a 3. Además, la escala OARSI también gradúa la artrosis de forma independiente en los compartimentos tibiofemorales lateral y medial. De esta forma, en esta escala existen tres umbrales independientes de presencia de artrosis en la articulación: mostrar un grado 2 de estrechamiento intra-articular o; llegar a un grado 2 de presencia de osteofitos o; tener de forma simultánea un grado 1 de estrechamiento intra-articular y grado 1 de presencia de osteofitos.

Un aspecto a favor de la escala OARSI es que estudia de forma individual el estrechamiento intra-articular o la formación de osteofitos; ambos elementos surgen por mecanismos etiológicos distintos y su progresión no es constante o proporcional, lo cual podría afectar a la estimación real de la incidencia en las poblaciones (Oka *et al.*, 2010). En la literatura se ha sugerido que el umbral de incidencia del criterio KL y del OARSI es similar en ambos sistemas, pero esto no se había comparado directamente hasta 2015 gracias a Culvenor y colaboradores (Culvenor *et al.*, 2015). En este estudio compararon las tasas de artrosis tibiofemoral usando la escala KL y la escala OARSI y observaron dos aspectos: primero, que la incidencia de artrosis tibiofemoral casi se duplica al usar la escala OARSI; y segundo, que el umbral de incidencia de la escala KL se corresponde muy bien con el criterio de estrechamiento intra-articular de grado 2 de la escala OARSI. En vista de los datos concluyeron que realmente estas dos escalas no son comparables y que la escala OARSI puede ser muy útil para identificar individuos que están en un estado muy temprano de la enfermedad donde hay pequeños cambios radiológicos en la articulación. Ambas escalas están aceptadas a día de hoy, y a pesar de este inconveniente, la escala KL sigue siendo la más utilizada de forma rutinaria para clasificar las articulaciones según la severidad.

## Introducción

### *Fenotipo de artrosis con progresión rápida.*

Como ya se ha comentado anteriormente, la artrosis es una enfermedad altamente heterogénea donde intervienen distintos factores. Esto produce que sea muy difícil definir la incidencia de esta patología, produciendo que hayan surgido diversas definiciones. Las más empleadas son las definiciones de incidencia de artrosis radiográfica y artrosis sintomática (Pereira *et al.*, 2011). En la primera, la definición de incidencia de artrosis radiográfica, se puede llevar a cabo mediante diversos sistemas, siendo uno de los más empleados la escala de Kellgren y Lawrence (KL) descrita en el apartado anterior (Kellgren y Lawrence, 1957), y se centra solamente en cambios radiográficos, en concreto se define incidencia como la evolución de un grado KL 0-1 a un grado mayor o igual a 2 en el tiempo (Wesseling *et al.*, 2015; Felson *et al.*, 2011). En la segunda, la definición de artrosis sintomática, se considera que la incidencia ocurre cuando coexisten síntomas radiográficos y articulares, es decir, no solo alteraciones radiográficas de la articulación sino también rigidez o pérdida de función entre otras características.

Por otro lado, la progresión radiográfica en esta patología se define clásicamente como un avance en la escala KL a partir de un grado 2 (Wesseling *et al.*, 2015; Felson *et al.*, 2011), pero esto puede ser muy impreciso sobre todo en el grado KL 3, donde el rango de estrechamiento articular es tan amplio que dificulta la detección de cambios en la progresión. Autores como Felson (Felson *et al.*, 2011) recomiendan centrarse únicamente en el estrechamiento articular mediante el uso de medidas semicuantitativas. Cabe destacar que la artrosis también es heterogénea en cuanto a su progresión, pudiendo alternar con períodos de inactividad (Felson *et al.*, 2013). Esto implica que el paso de un grado KL 0-1 a un grado KL 3-4 pueda ocurrir a lo largo de un período de tiempo amplio. En este sentido, en 1995 se estimó que la tasa de progresión promedio habitual es del 4% por año (Felson *et al.*, 1995). Sin embargo, dicha velocidad de progresión en realidad es muy variable (Riddle *et al.*, 2016; Walker *et al.*, 2011), existiendo una fracción de individuos que sufren este proceso de avance de la enfermedad en períodos de tiempo mucho más cortos, de pocos años o incluso meses. Esto es lo que se conoce como fenotipo de artrosis de rodilla con progresión rápida.

La primera mención de este fenotipo en la literatura fue realizada por Forestier en el año 1957 en la cadera (Forestier, 1957), y en 1959 por Coste (Coste *et al.*, 1959). Los primeros en acuñar el nombre de artrosis de cadera de progresión rápida fueron Postel y Kerboull (Postel y Kerboull, 1970). Estos autores describieron que, al inicio de la enfermedad, la cadera parece seguir una progresión normal con un dolor moderado, pero que en un período de 6 a 12 meses la destrucción puede ser extrema y con un dolor constante. En su trabajo describieron tres características principales: i) una destrucción severa de la cabeza femoral y del acetábulo, pero sin formación de osteofitos, ii) la desaparición total del cartílago articular y la aparición de una esclerosis del hueso subcondral, así como iii) la ausencia de inflamación. Como resultado, el paciente muestra un alto grado de dolor sin reducciones significativas de la movilidad articular y, recalcan la dificultad de identificar este fenotipo, siendo fácil confundirlo con casos de artritis reumatoide y necrosis. El mismo año en el que se describió esta patología, Lequesne definió la artrosis de cadera con progresión rápida como una tasa de pérdida anual del espacio intra-articular de 2 mm o bien una pérdida anual del 50% del espacio intra-articular (Lequesne y Amouroux, 1970).

Las causas que puedan provocar este fenotipo son desconocidas a día de hoy, esto es debido a que la identificación de factores de riesgo que puedan promover el desarrollo de este



subtipo de artrosis es compleja, ya que la mayoría de factores propuestos también están descritos como factores de riesgo de incidencia o de progresión normal (D. T. Felson *et al.*, 2013). Actualmente se consideran como potenciales factores de riesgo para el fenotipo de artrosis con progresión rápida el incorrecto alineamiento de la rodilla, los daños en los meniscos, el sexo femenino, la edad avanzada, la lesión previa, los defectos en el hueso subcondral, la osteoporosis y la osteonecrosis (Mitrovic y Riera, 1992; Yamamoto y Bullough, 2000; Blair-Levy *et al.*, 2008; Walker, Davis y Mosher, 2011; Driban *et al.*, 2014, 2020; Mavrogenis *et al.*, 2015; Hart y Fehring, 2016).

Con el paso de los años aumentó el conocimiento sobre este fenotipo. Se ha descrito en nuevas articulaciones como el hombro (Flemming y Gustas-French, 2017) o la rodilla (Emrani *et al.*, 2009). En relación a la artrosis de rodilla con progresión rápida, su incidencia en la población adulta se estima en un 3,4%. Entre los pacientes con artrosis de rodilla, uno de cada siete pueden sufrir un proceso de progresión rápida, y además, dos de cada tres casos con progresión rápida pueden alcanzar un elevado grado de severidad en un período de 12 meses (Driban *et al.*, 2020). A lo largo de los años, se han propuesto diferentes criterios para definir la progresión rápida en la artrosis, sin embargo, actualmente no existe ninguno que haya sido aceptado como el *gold standard*. Los distintos criterios que se han descrito hasta la fecha pueden clasificarse en los siguientes grupos (Tabla 2):

1. Criterios basados en la pérdida del espacio intra-articular o JSW (del inglés *joint space width*): normalmente, la pérdida media del JSW es de  $0,13 \pm 0,15$  mm/año (Emrani *et al.*, 2009). El criterio más empleado es el de progresión rápida de tipo 1 (Pivec *et al.*, 2013; Roemer *et al.*, 2015); este criterio, el cual está basado en la definición de Lequesne de 1970 (Lequesne y Amouroux, 1970), caracteriza a este fenotipo como una reducción de 2 mm en el espacio mínimo intra-articular (mJSW, del inglés *minimal joint space width*) de la zona medial de la rodilla en un período de un año, pero sin destrucción ósea. Por su parte, Bartlett y colaboradores (Bartlett *et al.*, 2011) definen como progresor rápido a aquel individuo que sufre una pérdida de JSW de 2,1 mm durante dos años de seguimiento. En el proyecto *Applied Public-Private Research enabling Osteoarthritis Clinical Headway* (APPROACH) establecen como un progresor rápido a todos aquellos individuos que tienen una pérdida de mJSW mayor o igual a 0,3 mm/año mantenida en un período de 2 o 3 años (van Helvoort *et al.*, 2021). Otros criterios se centran en la reducción del JSW en un compartimento específico de la articulación, por ejemplo Neogi y colaboradores (Neogi *et al.*, 2012) en su estudio se centran la reducción de JSW específicamente en el compartimento medial en la posición  $x=0,250$  (Fig. 7), y definen como progresión rápida a aquellos individuos que sufren en esta posición una reducción anual de JSW de 0,25 mm durante el período de seguimiento.
2. Criterios basados en la evolución temporal de la escala Kellgren-Lawrence (KL): en esta categoría están aquellos criterios que, en vez de fijarse en un único aspecto de la articulación como es la pérdida de espacio intra-articular o la formación de osteofitos, estudian diversas características de este órgano usando una escala radiográfica, siendo la más utilizada la escala KL. Dentro de este grupo destaca el criterio que define a la artrosis de rodilla acelerada, propuesto por Driban y colaboradores (Driban, Stout, *et al.*, 2016). La artrosis acelerada se define como una progresión desde un grado KL 0-1 (rodilla sana) a un estado avanzado de grado KL 3-4 en un período menor o igual a 48

## Introducción

meses. También se puede incluir en esta categoría el criterio descrito por Wesseling y colaboradores (Wesseling *et al.*, 2015) donde definen a un progresor rápido como a aquel individuo que partiendo de un grado KL 0, durante un período de seguimiento de 4 o 5 años avanza al menos dos grados KL. Finalmente, en este subgrupo también está presente la denominada artrosis incidente rápidamente progresiva; este término fue acuñado por Riddle y colaboradores (Riddle *et al.*, 2016), y se define como un avance desde un grado KL 0-1 hasta un grado KL 3-4 durante un período de seguimiento de cinco años, es decir, es muy similar a la artrosis acelerada, solo que cambia el período de seguimiento establecido.

3. Criterios basados en la destrucción articular: dentro de este grupo se encuentra la artrosis de progresión rápida de tipo 2, la cual está caracterizada por una pérdida acelerada del cartílago articular junto con destrucción excesiva del hueso subcondral (Hochberg, 2015). Durante la fase III de los ensayos con un prometededor DMOAD denominado tanezumab en pacientes con artrosis, se observó que algunos pacientes presentaban como efecto adverso la progresión rápida de la artrosis de tipo 2 (Hochberg *et al.*, 2016). Este hecho aumentó notoriamente el interés por el fenotipo rápido de artrosis de rodilla.

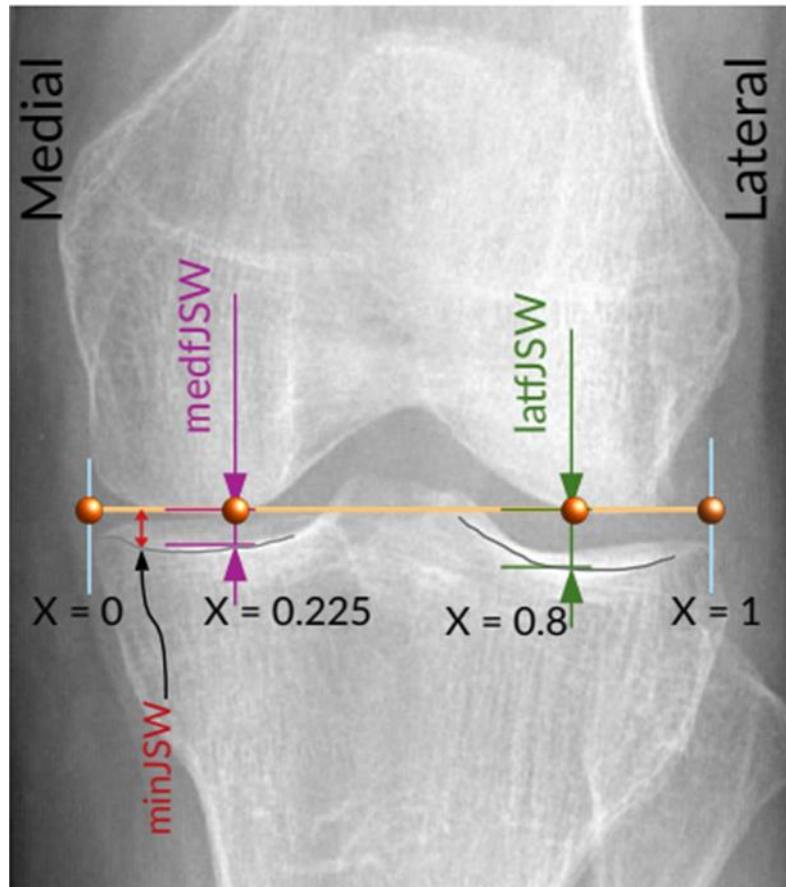


Figura 7. Compartimentación de la rodilla. Esta imagen frontal sirve como ejemplo de cómo se pueden realizar las mediciones de mJSW o JSW fijo en una posición. Para medir el JSW fijo desde la zona medial (medfJSW) es necesario trazar un eje horizontal imaginario desde la zona medial hasta la lateral, que va desde  $x=0,0$  hasta  $x=1,0$ . Dentro de este eje se puede medir el JSW en posiciones concretas como la  $x=0,25$  o, el que se suele usar más comúnmente  $x=0,225$ . En este ejemplo también se puede observar el JSW fijo de la zona lateral (latfJSW) y el mJSW (minJSW). Imagen obtenida de (Bayramoglu, Nieminen and Saarakkala, 2020).

Debido a que actualmente no existe un criterio consenso para definir la artrosis de rodilla con progresión rápida, según la astringencia de los criterios utilizados, la tasa de prevalencia de este fenotipo de la enfermedad puede variar (Driban, Stout, *et al.*, 2016). Además, ya en el año 1970 Postel y Kerboul resaltaron la importancia de la correcta identificación de estos pacientes, ya que ciertas condiciones similares pueden ser vistas en casos de artritis reumatoide y osteonecrosis, y que el mal diagnóstico puede llevar a complicaciones desastrosas (Postel y Kerboul, 1970). Pivec y colaboradores también resaltan la dificultad a la hora de distinguir entre la artrosis, la osteonecrosis y el fenotipo de progresión rápida (Pivec *et al.*, 2013).

Tabla 2. Resumen de los distintos criterios propuestos para definir un progresor rápido.

Grupo de criterios	Autores	Año	Definición de progresor rápido
Pérdida de JSW	Lequesne y Amoroux	1970	Reducción de mJSW medial $\geq 2$ mm /año sin destrucción ósea
	Bartlett y colaboradores	2011	Reducción de mJSW $\geq 2,1$ mm durante 2 años de seguimiento
	Neogi y colaboradores	2012	Reducción de JSW en $x=0,250 \geq 0,25$ mm /año durante 2 años de seguimiento
	van Helvoort y colaboradores	2021	Reducción de mJSW $\geq 0,3$ mm /año sostenida durante 2-3 años de seguimiento
Evolución temporal del grado KL	Wesseling y colaboradores	2015	Incremento KL $\geq 2$ desde un grado KL basal = 0-1 en 4-5 años de seguimiento
	Driban y colaboradores	2016	Avance de un KL 0 o 1 hasta KL 3 o 4 en 4 años de seguimiento o menos
	Riddle y colaboradores	2016	Avance de un KL 0 o 1 hasta KL 3 o 4 en 5 años de seguimiento o menos
Destrucción articular	Hochberg	2015	Pérdida acelerada de cartilago articular y destrucción del hueso subcondral

JSW: joint space width; mJSW: minimal joint space width; KL: grado en escala Kellgren y Lawrence

En 2014 la liga europea contra los reumatismos (EULAR, el inglés *European Alliance of Associations for Rheumatology*) estableció como una de sus líneas de investigación prioritarias el desarrollo de una herramienta capaz de identificar el fenotipo de artrosis con progresión rápida (Conaghan *et al.*, 2014). Teniendo en cuenta la importancia de este fenotipo, distintos grupos de investigación han centrado sus esfuerzos en identificar biomarcadores clínicos y/o moleculares que permitan estimar el riesgo de desarrollar este fenotipo. Ya en la década de los 2000 se descubrió que los pacientes con artrosis de cadera rápidamente destructiva presentaban en orina mayores niveles de los péptidos de colágeno II helical (Helix II) y C-telopéptido (Garnero *et al.*, 2003, 2006). Riddle y colaboradores (Riddle *et al.*, 2016) desarrollaron un modelo que incorporaba las variables de contralateralidad, IMC, dolor WOMAC (del inglés *Western Ontario and McMaster Universities Osteoarthritis Index*) y el grado KL 1. En su estudio, Jacobs y colaboradores (Jacobs *et al.*, 2020) vieron como la combinación de obesidad y depresión se asociaba con un mayor riesgo de progresión rápida de la artrosis. Driban y colaboradores encontraron una fuerte asociación entre la lesión previa en la rodilla diana con la artrosis acelerada (Driban *et al.*, 2014). Todo este esfuerzo por parte de la comunidad científica tiene como objetivo el desarrollo de un modelo basado en biomarcadores, que nos permita clasificar a los pacientes según su riesgo de sufrir una progresión rápida de esta enfermedad para:

1. Realizar un seguimiento clínico más exhaustivo y aplicar terapias personalizadas.
2. Incluir a estos pacientes en ensayos clínicos de nuevos fármacos para combatir la artrosis, ya que al sufrir el avance en un período tan corto de tiempo, servirían como un modelo a pequeña escala de todo el proceso, pudiendo así testar la eficacia de estas nuevas moléculas (Conaghan *et al.*, 2014)

## Introducción

Teniendo en cuenta estos dos aspectos, los estudios en este campo son por tanto una ventana de oportunidades para la medicina personalizada y tratamiento de la artrosis, ya que como se verá más adelante, en esta enfermedad no existe actualmente ningún fármaco o terapia que permita prevenir, frenar o detener el proceso artrósico. Si bien es cierto que la artrosis no es una enfermedad que produzca la muerte del paciente, y por ello no se le da la importancia que merece, sí que es una enfermedad crónica que merma de manera notable la calidad de vida del paciente que la sufre. De modo que, con el objetivo de permitir que la población pueda envejecer sin grandes limitaciones en su calidad de vida, es de vital importancia desarrollar líneas de investigación como los que representa esta tesis doctoral.

### *Tratamiento de la Artrosis.*

Los tratamientos para la artrosis son muy variados, desde el ejercicio y ritmo de vida saludable a operaciones quirúrgicas como las artroplastias pasando por tratamientos farmacológicos. Actualmente los regímenes de tratamiento más comunes se centran principalmente en controlar el dolor y la inflamación articular ya que, a día de hoy, no existe un tratamiento que sea capaz de prevenir, frenar o detener el proceso artrósico (Kloppenborg y Berenbaum, 2020). Sin embargo, estudios recientes están tratando de implementar novedosas terapias biológicas que tienen como objetivo remodelar la articulación a nivel biomecánico, fomentando procesos anabólicos en el cartílago y el hueso subcondral (Richards *et al.*, 2016). Se pueden clasificar los tratamientos para la artrosis en cuatro grupos: los no farmacológicos, los farmacológicos, la terapia biológica y los tratamientos quirúrgicos.

1. Tratamientos no farmacológicos. La primera línea de tratamiento comprende la práctica de ejercicio, la pérdida de peso (en individuos con un IMC de 25 o superior) y la re-educación del paciente.

- Terapia física: los programas de ejercicio físico sobre suelo o en el agua han mostrado mejorías en la funcionalidad y reducción del dolor de la artrosis de rodilla y cadera, sin embargo, hay menos evidencia de que puedan reducir el dolor de la artrosis de manos. Este tipo de terapias deben centrarse primero en mejorar la fuerza muscular circundante a las articulaciones afectadas y, en segundo lugar, en enfatizar el ejercicio aeróbico general (Martel-Pelletier *et al.*, 2016). Además, el ejercicio físico no es importante solo para el alivio sintomático de la artrosis, sino que puede prevenir hasta 35 condiciones crónicas y mejorar la sintomatología de otras 26. Dos de cada tres pacientes con artrosis de rodilla y cadera tiene una o más comorbilidades como diabetes tipo 2, hipertensión o depresión y, por tanto, el ejercicio y la actividad física son cruciales para mantener una buena salud y bienestar en pacientes con artrosis (Skou y Roos, 2020).

El ejercicio físico puede producir dolor muscular temporal, sin embargo, los tratamientos farmacológicos pueden tener asociados efectos secundarios más graves en el estómago, hígado o sistema cardiovascular. Además, estos brotes de dolor suelen ser el resultado del proceso de adaptación del cuerpo a la rutina de ejercicio, por lo que suelen desaparecer tras cinco o seis semanas.

Dentro del ejercicio físico, el Tai chi es un arte marcial que ha sido recomendada para pacientes con artrosis de rodilla o cadera por Kolasinski y colaboradores, ya que su práctica mejora la fuerza muscular, el equilibrio, ayuda a prevenir caídas y puede ayudar a combatir la depresión (Kolasinski *et al.*, 2020). Del mismo modo, otra disciplina útil para combatir la artrosis de rodilla sería el

- Yoga, la cual también favorece la obtención de fuerza muscular, flexibilidad y movilidad funcional (Deepeshwar *et al.*, 2018).
- Pérdida de peso: una corrección de la alimentación y de hábitos de vida que suponga una reducción de peso de entre 5 y 10% se ha asociado de forma significativa con una mejoría del dolor, de la función articular y de la reducción de la progresión del daño estructural. Reducciones de peso mayores también se han asociado con mayores mejoras en estos parámetros (Kolasinski *et al.*, 2020). Además, la combinación de la reducción de peso con la terapia física ha mostrado un efecto aditivo en la reducción del dolor y en el aumento de función articular (Martel-Pelletier *et al.*, 2016; Skou y Roos, 2020).
  - Educación del paciente: el efecto de un programa de terapia física es beneficioso para combatir la artrosis, pero su eficacia disminuye con el tiempo debido a que los pacientes tienden a tener baja adherencia al régimen de ejercicio y a los cambios de estilo de vida (Martel-Pelletier *et al.*, 2016). La educación del paciente, junto con sesiones de seguimiento, mejora la eficacia y adherencia a la terapia física. Estos elementos deben educar al paciente sobre las causas, factores de riesgo, mecanismos de la enfermedad, las consecuencias de la inactividad física y todas las herramientas necesarias para ayudar a que el paciente pueda administrar su enfermedad de forma exitosa. Una vez más, el efecto combinado de la terapia física y la educación del paciente tiene un efecto aditivo en la reducción del dolor y el aumento de la función articular (Skou y Roos, 2020).
  - Tratamiento no farmacológico suplementario: cuando ninguno de los elementos citados previamente ha dado resultado, el terapeuta puede ofrecer tratamientos suplementarios (Skou y Roos, 2020). Entre ellos se encuentra la acupuntura o los ejercicios de equilibrio; sin embargo, es necesario realizar más estudios para demostrar su eficacia (Kolasinski *et al.*, 2020).
2. Tratamientos farmacológicos. En el caso de que los tratamientos no farmacológicos no hayan dado resultado, la siguiente vía de actuación consiste en el uso de tratamientos farmacológicos tópicos, orales o inyectables. Este tipo de terapias se centran principalmente en combatir el dolor y la inflamación.
- Analgésicos: uno de los primeros medicamentos recomendados por diversas guías de artrosis es el uso del paracetamol, y a pesar de tener pocos efectos adversos como la hepatotoxicidad, solo proporciona una mejoría mínima en el dolor y en la funcionalidad (Kloppenburger y Berenbaum, 2020). También se puede prescribir el uso inicial del analgésico capsaicina de forma tópica.
  - Antiinflamatorios no esteroideos (AINES): la guía 2019 de la OARSI para el tratamiento no quirúrgico de la artrosis (Bannuru *et al.*, 2019) recomienda el uso tópico de los AINES en la artrosis de rodilla, ya que presentan una mejor relación entre eficacia y efectos adversos que los AINES de uso oral. Esto es debido a que el tratamiento tópico representa una dosis de medicamento substancialmente menor que la dosis del tratamiento oral. Sin embargo, la guía 2019 de la *American College of Rheumatology* para el tratamiento de la artrosis (Kolasinski *et al.*, 2020) no recomienda el uso tópico de AINES en la artrosis de cadera, debido a la profundidad bajo la piel a la que está la articulación, lo cual dificulta que la aplicación tópica muestre algún tipo de mejoría. Y tampoco la recomienda para la artrosis de mano, ya que su frecuente lavado dificulta el efecto de los AINES tópicos. Los AINES tradicionales más empleados han sido el

diclofenaco y el naproxeno. Con el paso del tiempo han aparecido nuevos AINES que inhiben de forma selectiva a la ciclooxigenasa 2 (COX2 o PTSG2 del inglés *Prostaglandin-Endoperoxide Synthase 2*). La expresión de esta enzima está aumentada durante los procesos inflamatorios debido a la presencia de la IL6 y de otros factores proinflamatorios. COX2 produce la formación de prostaglandinas, las cuales son moléculas necesarias para la homeostasis de diversos órganos, pero también están relacionadas con el dolor y la inflamación (García Meijide y Gómez-Reino Carnota, 2000). En esta nueva generación de AINES específicos de COX2 se encuentran: celecoxib, etoricoxib, lumiracoxib, parecoxib y rofecoxib. Sin embargo, el gran problema de los AINES son los efectos adversos, tanto gastrointestinales como cardiovasculares, motivo por el cual no está recomendado su uso en pacientes con comorbilidades relacionadas (Bannuru *et al.*, 2019).

- Anticuerpos monoclonales: este tipo de terapia busca diseñar elementos capaces de interferir con moléculas clave en la patogénesis de la artrosis, tratando así de aliviar la sintomatología durante más tiempo que los tratamientos tradicionales citados previamente, intentando frenar así la progresión de la artrosis. En la revisión de Zheng y colaboradores (Zheng *et al.*, 2016) se describen algunas de las moléculas diana más usadas en el desarrollo de anticuerpos monoclonales, como son el TNF, la IL1 $\beta$ , el factor de crecimiento endotelial vascular (VEGF del inglés *Vascular endothelial growth factor*) o el factor de crecimiento nervioso (NGF del inglés *Nerve growth factor*), entre otros. Algunos de estos tratamientos se encuentran todavía en distintas fases de ensayos clínicos. Y por desgracia algunos medicamentos como el tanezumab, que tiene como diana terapéutica el NGF, producen efectos adversos al combinarse con AINES, aumentando la tasa de progresión de la artrosis en un reducido número de pacientes (Hochberg, 2015).
- Nutracéuticos: en este grupo se encuentran tratamientos como la glucosamina y el condroitín sulfato, los cuales son compuestos naturales ricos en GAGs. Durante el proceso artrósico y el deterioro del cartílago articular se produce una pérdida de GAGs en dicho tejido; los nutracéuticos, gracias a su composición, permiten reponer dichos elementos en el cartílago. Sin embargo, distintas guías de tratamiento no recomiendan su uso debido a que no hay suficientes hallazgos que muestren un efecto analgésico significativo (Martel-Pelletier *et al.*, 2016).
- Inyecciones intra-articulares: el HA es un GAG que se encuentra en las capas superficiales del cartílago y que es además el componente principal del líquido sinovial. Las inyecciones de soluciones de HA (también conocido como viscosuplementación) pueden actuar como lubricantes y también ayudar a amortiguar las cargas de las articulaciones, protegiendo así la superficie articular durante la función articular, mejorando la funcionalidad y reduciendo el dolor (Cheng *et al.*, 2012; Clegg, Caborn y Mauffrey, 2013). Otro compuesto que suele inyectarse intra-articularmente son los corticosteroides, cuya función principal es la reducción del dolor y la inflamación local. La guía OARSI 2019 sobre el tratamiento de la artrosis recomienda el uso de ambas inyecciones intra-articulares, sin embargo, la inyección de corticosteroides produce una reducción a corto plazo del dolor, mientras que la

- viscosuplementación puede tener efectos en el dolor hasta 12 semanas tras el tratamiento, dando mayores resultados a largo plazo (Bannuru *et al.*, 2019).
3. Terapia biológica: a diferencia de los tratamientos farmacológicos, la terapia biológica tiene como objetivo estimular el anabolismo del hueso y del cartílago con el fin de reparar o regenerar la zona dañada. Sin embargo, a pesar de los hallazgos *in-vitro* y en estudios preclínicos, sigue habiendo una falta de evidencia que apoye la eficacia clínica de estos tratamientos (de Girolamo *et al.*, 2016). De hecho, la guía OARSI 2019 no recomienda su uso. Los tratamientos más destacables son los siguientes:
- Factores de crecimiento: juegan un papel clave en la homeostasis y reparación de tejidos ya que son capaces de estimular las respuestas pro-anabólicas de los condrocitos. Uno de los más conocidos es el TGF-  $\beta$ .
  - Derivados de la sangre: estos tratamientos tratan de influenciar el medio que rodea a la articulación produciendo la regeneración de los tejidos dañados. Los más conocidos son el suero autólogo condicionado (ACS del inglés *Autologous conditioned serum*) y el plasma rico en plaquetas (PRP del inglés *Platelet rich plasma*).
  - Células madre: Estas células son capaces de diferenciarse hacia distintos tipos celulares como osteocitos, condrocitos, adipocitos, miocitos, cardiomiocitos, hepatocitos y neuronas (Roberts *et al.*, 2011). A parte de su capacidad de diferenciación, las células madre son capaces de migrar a un tejido que sufre un proceso de inflamación del mismo modo que haría un leucocito, poseen propiedades inmunomoduladoras y antiinflamatorias y son hipoinmunogénicas, por lo que no activan al sistema inmune, haciéndolas idóneas para la realización de trasplantes a otros pacientes (trasplante heterólogo) (Chen y Tuan, 2008).
4. Tratamientos quirúrgicos. En el momento en que esta patología llega a un estado muy avanzado de destrucción articular y de dolor, la opción más recomendada en la rodilla y en la cadera es la artroplastia o reemplazo articular por una prótesis. En la cadera, la artroplastia es total, reemplazando tanto la cabeza femoral como el tronco femoral. En el caso de la rodilla, la artroplastia puede ser de un solo compartimento (unicondilar) o total. La principal complicación en este proceso es la infección, siendo en el 99% de los casos una infección superficial; otras complicaciones que pueden aparecer son la rigidez, el desarrollo de trombos o inflamación (Robinson *et al.*, 2018). En esta línea, cabe destacar que esta no es una solución definitiva y tiende a fallar tras un período de tiempo, por lo que la revisión de la prótesis es necesaria. El tiempo de vida de una prótesis es aproximadamente de 20 años (Kloppenborg y Berenbaum, 2020).

## Polimorfismos genéticos nucleares.

### *ADN nuclear: estructura y función.*

El ácido desoxirribonucleico (ADN) es un componente celular presente en el núcleo de las células eucariotas, cuya función es el almacenamiento de la información genética necesaria para codificar las secuencias de proteínas y moléculas de ácido ribonucleico (ARN) para el correcto funcionamiento celular (Travers y Muskhelishvili, 2015).

El ADN está formado por nucleótidos, los cuales están compuestos de tres elementos:

1. Una molécula de desoxirribosa: esta molécula de azúcar sirve como estructura de soporte para los otros elementos estructurales.
2. Un grupo fosfato unido al carbono 5' de la desoxirribosa: necesario para unirse con el grupo hidroxilo (OH) del carbono 3' del siguiente nucleótido de la cadena de ADN mediante un enlace fosfodiéster de tipo covalente.
3. Bases nitrogenadas: puede ser de dos tipos, dependiendo del compuesto del que provengan, si derivan de la purina se conocen como bases púricas y serían la A (adenina) y la G (guanina); por el contrario, si son derivados de la pirimidina se conocen como bases pirimidínicas y serían la T (timina) y la C (citosina).

La estructura de esta molécula fue descrita por primera vez por Watson y Crick en 1953 gracias a los hallazgos de Rosalind Franklin. El hallazgo fue tan importante que les permitió obtener el premio Nobel de medicina en el año 1962. Según el modelo de estos investigadores, el ADN es una doble hélice de nucleótidos, y cada una de las hebras de la hélice está unida a la otra mediante enlaces por puentes de hidrógeno entre las parejas de bases nitrogenadas de los nucleótidos, donde la adenina se une con la timina mediante dos puentes de hidrógeno, y la guanina se une con la citosina mediante tres puentes de hidrógeno. Los nucleótidos de una hélice están unidos entre sí mediante enlaces fosfodiéster y perpendicularmente están presentes las bases nitrogenadas que se unen a la otra hélice mediante los puentes de hidrógeno. Esto le confiere una estructura similar a una escalera, donde las barras laterales serían las cadenas de desoxirribosas y fosfatos, y los peldaños las parejas de bases nitrogenadas (Fig. 8 A). Las dos hélices del ADN son antiparalelas, de tal forma que presentan orientaciones contrarias, una de ellas está orientada en sentido 5'→3' mientras que la otra lo está en 3'→5'. Además, la doble hélice del ADN presenta un enrollamiento plectonémico, donde las dos hebras no pueden separarse sin desenrollarla previamente, y dextrógiro, gira en el sentido de las agujas del reloj (Fig. 8 B) (Burgos, 2012).

En total, el genoma humano tiene aproximadamente  $3 \cdot 10^9$  pares de bases (pb). Teniendo en cuenta que la longitud de un nucleótido es de 3,4 Angstrom (Å) (Un Å equivale a  $10^{-10}$  m), la longitud total del ADN por célula se calcularía en base a la siguiente fórmula:

$$(3,4 \text{ \AA} \times (3 \times 10^9) \text{ pb}) \times \frac{1 \text{ \AA}}{10000000000 \text{ m}} = 1,02 \text{ m}$$

Al tratarse de organismos diploides, es decir, con dos copias de ADN por célula somática, cada célula en total tiene 2,04 m de ADN en su interior. El cuerpo humano contiene aproximadamente  $5 \times 10^{19}$  células, por lo que en total un cuerpo humano medio contiene  $10^{20}$  metros de ADN, lo que equivale a 100.000.000.000.000.000 km. Teniendo en cuenta que la



distancia media entre la Tierra y el sol es de 149.600.000 km, nuestro contenido total de ADN cubre esta distancia un total 668.449.197,9 veces. A pesar de ser muy variable entre tejidos, el tamaño de las células humanas está en la escala de los micrómetros ( $\mu\text{m}$ ) ( $1 \mu\text{m}$  equivale a  $10^{-6}$  m), por lo que para contener esta abrumadora cantidad de ADN las células han tenido que desarrollar un mecanismo complejo de organización y compactación del ADN nuclear.

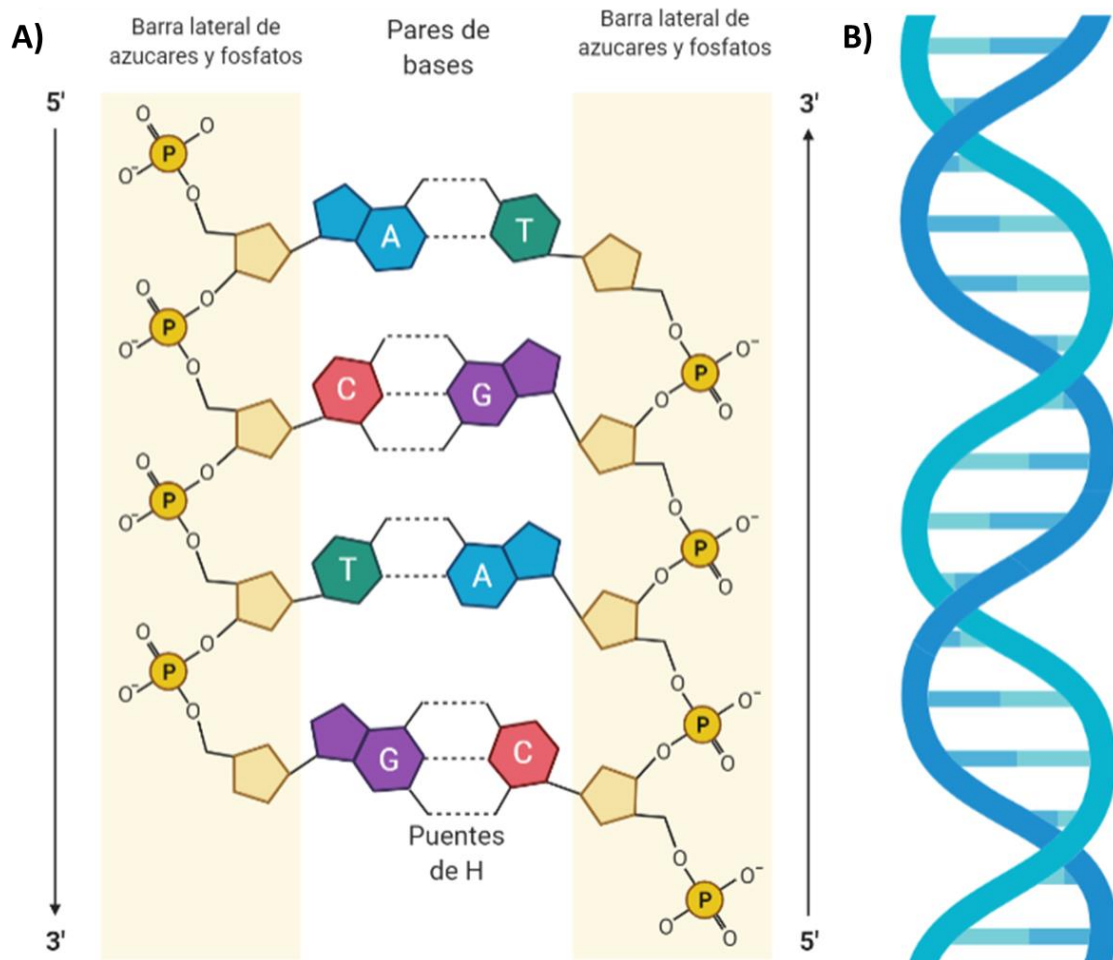


Figura 8. Esquema representativo de la estructura de la molécula de ADN de doble cadena. A: estructura molecular del ADN, en la parte central se pueden observar los elementos que conforman los cuatro tipos de desoxinucleótidos (dNTPs) y como se unen entre ellos mediante puentes de hidrógeno, en los laterales quedan situados tanto los azúcares como los grupos fosfato (P); B: morfología de la estructura de la molécula de ADN bicatenaria con menos aumentos para observar el enrollamiento que sufre esta molécula. Imagen creada en app.biorender.com.

Debido a que el ADN contiene la información necesaria para el correcto funcionamiento celular es importante que el empaquetamiento siga permitiendo a la maquinaria celular acceder a dicha información genética (Travers y Muskhelishvili, 2015). El empaquetamiento del ADN depende de las histonas, y al conjunto de ADN e histonas se le conoce como cromatina. La estructura primaria de la cromatina es una fibra de 10 nanómetros (nm) de diámetro donde la unidad fundamental es el nucleosoma. El nucleosoma se compone de 150 pb de ADN enrolladas alrededor de un núcleo de histonas; es un octámero formado por dos copias de cada una de las siguientes histonas: H2A, H2B, H3 y H4. Además, también se une a este complejo la histona de enlace H1 (o H5), la cual se une al nucleosoma e impide que se desenrolle el ADN. Como resultado de este empaquetamiento se obtiene una fibra de aproximadamente 10 nm de diámetro.

## Introducción

La estructura secundaria de la cromatina se forma gracias a las interacciones entre los nucleosomas. En esta estructura, las histonas de enlace forman un eje central alrededor del cual se disponen los nucleosomas, generándose la fibra de 30 nm de diámetro. Finalmente, la estructura terciaria se forma gracias a diversas proteínas ancladas en la matriz nuclear que producen el pliegue de la fibra de 30 nm en dominios en forma de *loop* de aproximadamente 300 nm. Cuando la célula va a dividirse, estas proteínas de matriz nuclear se asocian y forman un eje central, alrededor del cual se disponen los dominios de 300 nm, produciendo una fibra de 700 nm, la cual se usa para formar las cromátidas (o brazos) de los cromosomas humanos (Szerlong y Hansen, 2011; Burgos, 2012)(Fig. 9).

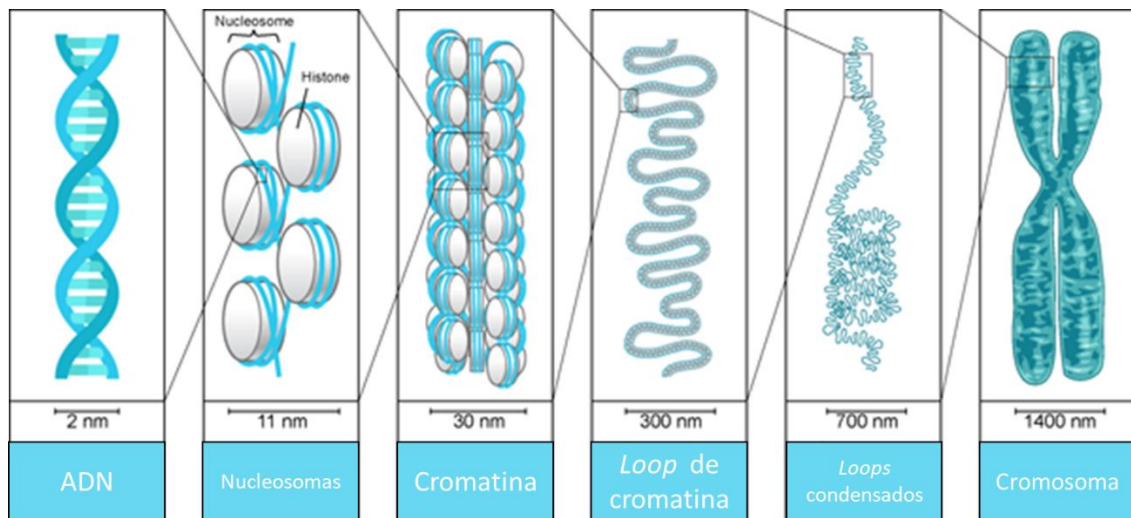


Figura 9. Esquema de los distintos niveles de empaquetamiento de la molécula de ADN, desde la molécula sin empaquetar hasta los cromosomas humanos. Imagen modificada de <https://www.shmoop.com/study-guides/biology/dna/dna-packaging>.

En las regiones en las que la cromatina se encuentra más compactada, el ADN no es accesible, por lo que se encuentra en un estado inactivo o silenciado. Estas zonas de elevado nivel de compactación reciben el nombre de *heterocromatina*, mientras que las regiones menos compactadas donde el ADN es accesible reciben el nombre de *euromatina*. La cromatina es dinámica y puede pasar de un estado a otro modificando así la expresión génica. Es en el estado de euromatina donde las ARN-polimerasas producen el proceso de transcripción para producir la expresión de dichas regiones (o genes) para obtener el producto proteico de interés (Burgos, 2012).

En el ADN nuclear no toda la secuencia es codificante, las regiones codificantes se denominan exones, y se encuentran intercaladas por grandes regiones no codificantes denominadas intrones. Los intrones no generan ningún producto proteico, pero sí contienen elementos reguladores que son necesarios para la transcripción como los elementos *cis* y *trans*. Los elementos *cis* son los promotores, los *enhancers* y los silenciadores, los cuales son capaces de regular la transcripción de genes cercanos, por otro lado los elementos *trans* pueden alterar la expresión de genes distantes al combinarse con sus secuencias diana (Wang *et al.*, 2019). A pesar de su gran tamaño, el ADN nuclear codifica para un total de aproximadamente 25.000 genes y, por lo general, las regiones codificantes de los genes contienen un alto porcentaje en guaninas y citosinas, esto es debido a que los codones de los aminoácidos más abundantes

suelen contener muchas guaninas y citosinas. En cambio las regiones no codificantes suelen ser ricas en adenina y timina (Travers y Muskhelishvili, 2015)(Fig. 10).

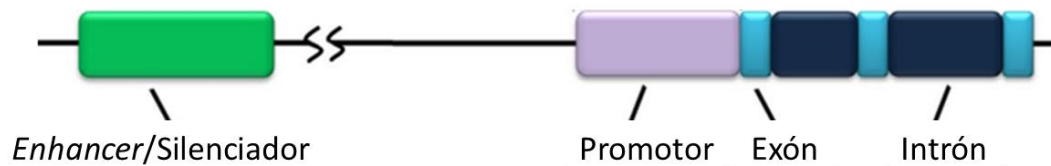


Figura 10. Representación de la organización de un gen. En caso de que los elementos cis distantes como los enhancers o los silenciadores estén presentes, algunos elementos trans, como proteínas pueden unirse a éstos y en conjunción alterar la expresión del gen. El promotor es la región desde donde comenzará la transcripción del gen. Los exones son la región codificante de los genes, mientras que los intrones, son regiones no codificantes que suelen ser eliminadas durante el proceso de splicing. Imagen modificada de (Bitas et al., 2016).

#### ADN nuclear: replicación, transcripción y traducción.

La información genética contenida en el ADN, que es necesaria para generar proteínas, necesita sufrir tres importantes procesos para llevar a cabo este fin: la replicación del ADN, la transcripción y la traducción. En el primero tiene lugar la replicación de la molécula de ADN, en el segundo se transforma el ADN en ARN mensajero (ARNm), y en el tercer proceso, el ARNm se traduce a proteína.

En primer lugar, se comentará brevemente el proceso de replicación del ADN nuclear. La replicación de esta molécula se realiza de forma semiconservativa, es decir, la doble hélice del ADN se separa y cada una de sus hebras sirve de molde para sintetizar una nueva hebra, y también es bidireccional, ocurre tanto en una cadena en sentido 5'→3' como en la otra en dirección 3'→5'. Este proceso ocurre en el núcleo celular y, en los organismos eucariotas, el proceso se inicia de forma simultánea en diversos puntos denominados "orígenes de replicación", que se activan una única vez por ciclo de replicación y son secuencias cortas de nucleótidos ricas en pares de AT. Estos puntos son identificados por un complejo proteico conocido como ORC (del inglés *origin recognition complex*), que se une a la secuencia del origen de replicación y lo activa, separando las dos hebras de ADN y generando la estructura denominada horquilla de replicación, tras lo cual da inicio la replicación.

Para que se produzca este proceso es necesario ir separando las cadenas de ADN desde el origen de replicación, esto es llevado a cabo por tres tipos de proteínas: la ADN-girasa, que es una topoisomerasa que desenrolla el ADN, las helicasas, que se unen al ADN y rompen los enlaces de hidrógeno entre las bases nitrogenadas, y finalmente para que no se vuelva a producir la unión de las dos hebras, intervienen las proteínas SSB (del inglés *single-strand-binding proteins*). Una vez separadas las hebras, comienza la fase de la elongación donde intervienen una serie de proteínas denominadas ADN-polimerasas. Estas enzimas tienen actividad de síntesis de ADN en dirección 5'→3' y necesitan, para poder sintetizar ADN, que esté presente un extremo 3' con OH libre, para ello, una enzima ADN primasa suministra una secuencia corta de ARN denominada cebador. La polimerasa añade el primer nucleótido al extremo libre del cebador suministrado, y desde ahí comienza a añadir uno a uno los nucleótidos siguientes por complementariedad de bases, produciéndose la extensión o elongación del ADN en dirección 5'→3'. Debido a esta última peculiaridad de las polimerasas y a que cada hebra del ADN discurre en un sentido distinto, la cadena que va en sentido 3'→5' se sintetizará de forma continua y se denominará cadena líder, mientras que la otra cadena, la que discurre en sentido 5'→3' se sintetizará de forma discontinua y recibe el nombre de cadena retardada. La cadena

## Introducción

retardada generará con la ayuda de múltiples cebadores, creados por las ADN primasas, pequeños segmentos en dirección 5'→3' denominados fragmentos de Okazaki, los cuales son unidos posteriormente mediante una ADN-ligasa. El proceso de elongación continuará hasta haber replicado la totalidad de la molécula, momento en el que la replicación se detiene (Benito, 2012)(Fig. 11).

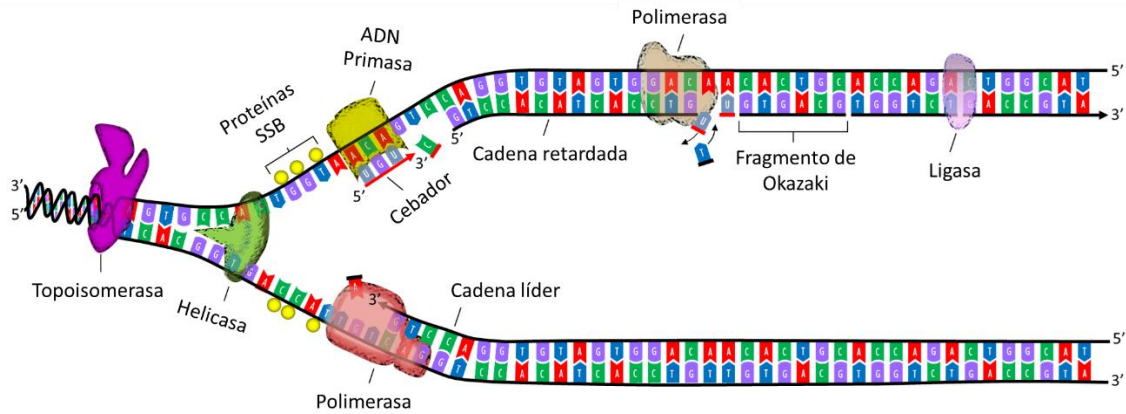


Figura 11. Esquema representativo de la fase de elongación de la replicación. La topoisomerasa, la helicasa y las proteínas SSB son las encargadas de producir el desenrollamiento y la separación de las dos hebras del ADN molde. La cadena líder se sintetiza de forma continua en sentido 5'→3' con la ayuda de un único cebador. Sin embargo, la cadena retardada se sintetiza de forma discontinua en la misma dirección 5'→3' pero generando pequeños fragmentos de Okazaki que, finalmente son unidos mediante una ADN ligasa. Imagen modificada de [https://commons.wikimedia.org/wiki/File:DNA\\_Replication\\_in\\_prokaryotes.png](https://commons.wikimedia.org/wiki/File:DNA_Replication_in_prokaryotes.png).

Cabe destacar que las polimerasas que intervienen en este proceso tienen una elevada fiabilidad a la hora de replicar la secuencia, pero igualmente pueden producir errores, por lo que suelen producir errores cada 100.000 o 1.000.000 nucleótidos. Pero, hay que tener en cuenta que algunas de las polimerasas que intervienen en este proceso tienen capacidad de corrección de errores en sentido 3'→5', y junto con los sistemas de reparación del ADN nuclear (explicados en apartados posteriores), se estima que la tasa de error en la replicación es de entorno a un error por cada mil millones de nucleótidos y por ciclo de replicación (Preston *et al.*, 2010). Esto también está presente en la *Taq polimerasa 1*, ampliamente utilizada en la reacción en cadena de la polimerasa (PCR, del inglés *Polymerase Chain Reaction*). Esta enzima tiene una tasa de error estimada de entre  $1 \cdot 10^{-5}$  a  $2 \cdot 10^{-4}$  (Potapov y Ong, 2017), es decir entre 1 de cada 5.000 y 1 de cada 100.000 pb.

El segundo proceso mencionado, la transcripción del ADN (Fig. 12), ocurre en el núcleo y, para que comience, es necesario que la cromatina sea accesible, pudiéndose así reconocer las secuencias reguladoras de la transcripción en el ADN, que son el promotor, la zona que se va a transcribir y una señal de terminación. Las enzimas principales encargadas de este proceso en células eucariotas son la ARN-polimerasa II y diversos factores de transcripción. Una vez se reconocen las secuencias reguladoras en la cadena de ADN molde, la ARN-polimerasa y los factores de transcripción forman un complejo proteico, se produce un cambio conformacional en la ARN-polimerasa y se liberan todos los factores de transcripción, lo que permite que esta enzima se desplace desde el promotor por el resto de la cadena incorporando los ribonucleótidos necesarios. La zona de terminación de la transcripción presenta repeticiones invertidas en su secuencia y, cuando la polimerasa llega a este punto, la complementariedad de bases obliga a formar una estructura secundaria en forma de horquilla seguida de uracilos. Esta estructura promueve que la ARN-polimerasa detenga su actividad, liberándose así el pre-ARNm y finalizando la transcripción. Durante este proceso el pre-ARNm es procesado, modificando el

extremo 5' y generándose así una estructura denominada caperuza, que protege al ARN de la degradación y que es necesaria para el proceso de traducción. En segundo lugar, se eliminan los intrones mediante un proceso de corte y empalme (*splicing*). Finalmente, se modifica el extremo 3', añadiendo en él un número variable de adeninas, generándose una estructura terminal llamada la cola de poli-adeninas. Tras este proceso se produce la maduración del pre-ARNm a ARNm. (Linacero, 2012).

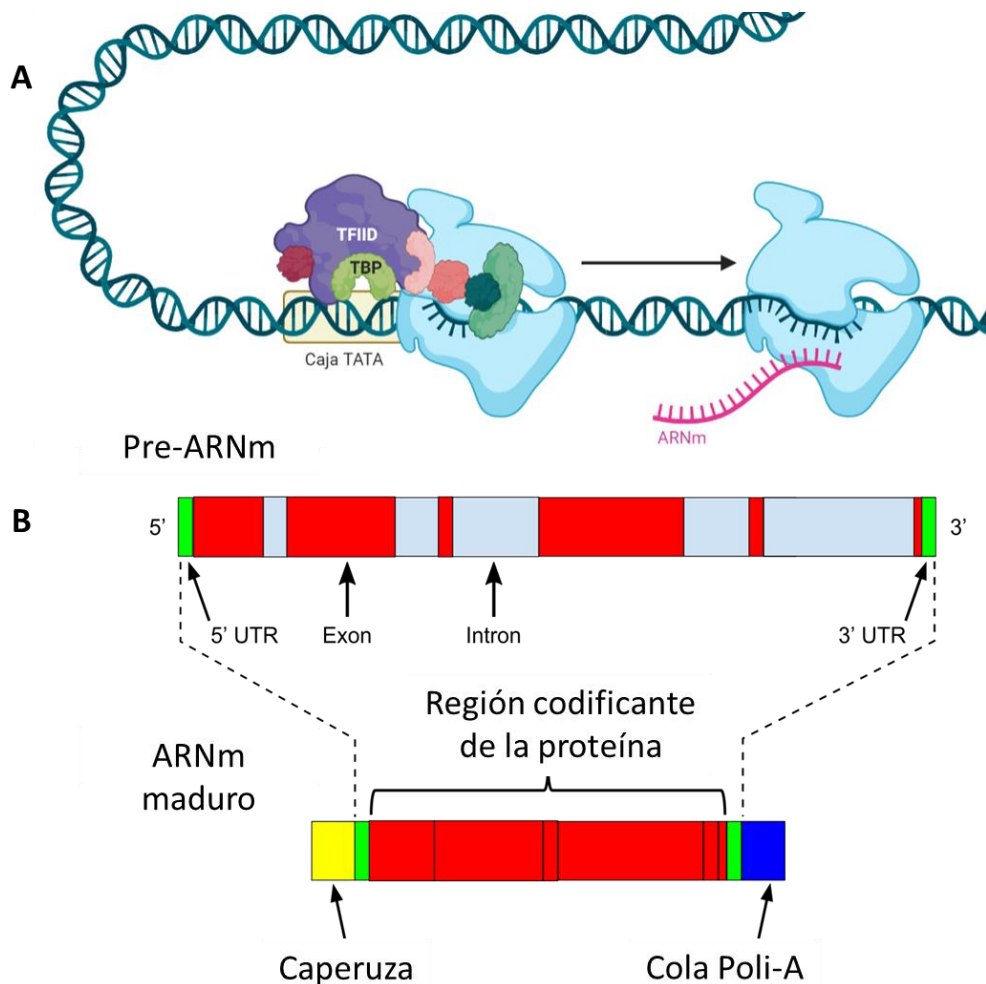


Figura 12. Esquema representativo del proceso de transcripción. A: durante el proceso de iniciación de la transcripción la ARN polimerasa II necesita que el polipéptido de unión a caja TATA (TBP del inglés TATA binding protein) del factor de transcripción II D (TFIID) reconozca a la secuencia "caja TATA" del promotor, tras esto el resto de factores de transcripción se unen a la polimerasa formando el complejo proteico. Acto seguido se produce un cambio de conformación de la polimerasa, se liberan los factores de transcripción y comienza a desplazarse por la cadena generando ARNm usando ribonucleótidos. Este proceso continuará hasta la región de terminación donde se liberará la molécula de pre-ARNm.; B: durante la transcripción el pre-ARNm sufre un proceso de maduración donde se añade una caperuza protectora en el extremo 5' de la cadena y una cola de poli-adeninas en el extremo 3' y, además, se realiza un proceso de *splicing* donde se eliminan las regiones no codificantes, denominadas intrones. Imagen A creada con [app.biorender.com](https://www.biorender.com), imagen B modificada de [https://es.wikipedia.org/wiki/ARN\\_mensajero#/media/Archivo:Pre-mRNA.svg](https://es.wikipedia.org/wiki/ARN_mensajero#/media/Archivo:Pre-mRNA.svg).

El ARNm maduro es fundamental para el último proceso, la traducción, donde la información que posee esta molécula es necesaria para el correcto ensamblaje y orden de los aminoácidos de la proteína diana (Fig. 13). A diferencia de la transcripción, este proceso ocurre en el citoplasma celular, concretamente en los ribosomas, los cuales son unos orgánulos formados por dos subunidades compuestas por proteínas y distintos tipos de ARN ribosómico

## Introducción

(ARNr), la mayor, ARNr 60S (unidad de coeficiente de sedimentación Svedberg) y 49 proteínas, y la menor por ARNr 40S y otras 33 proteínas. El proceso de traducción puede ocurrir de forma simultánea mediante la unión de un mismo ARNm y varios ribosomas, o también denominado polirribosoma.

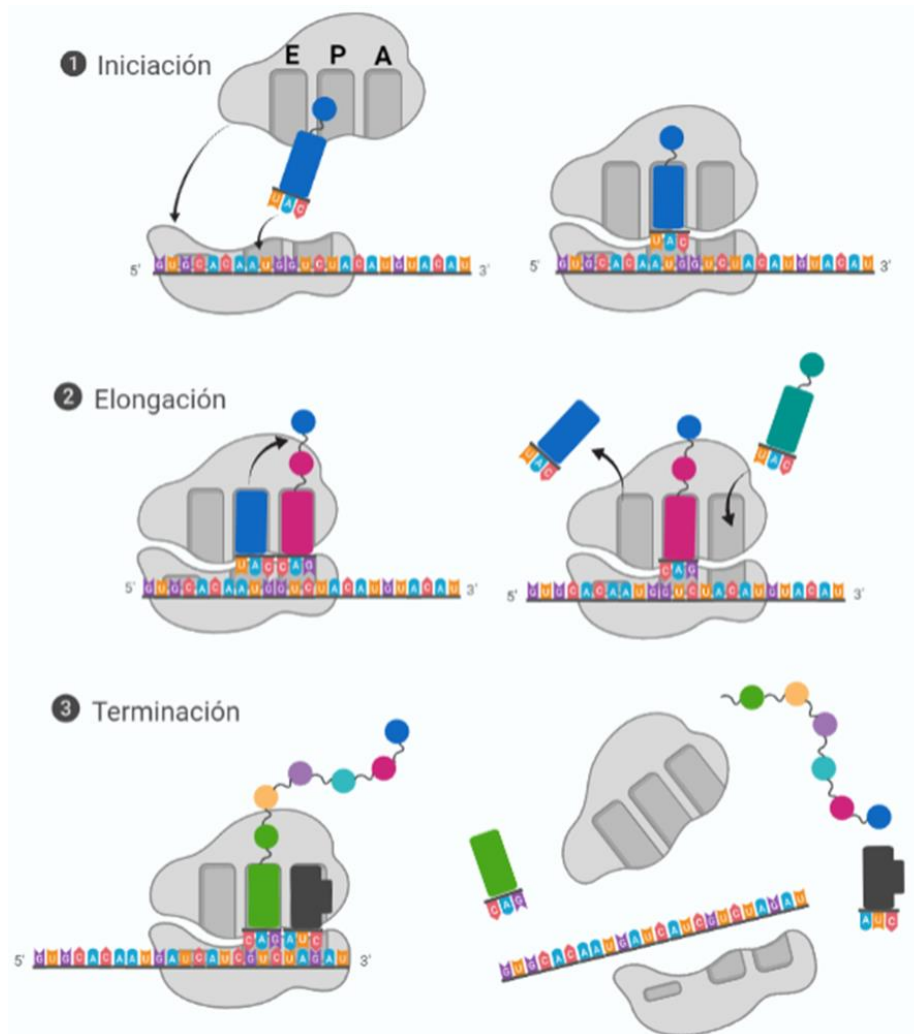


Figura 13. Esquema representativo de la traducción. 1. **Iniciación:** en eucariotas este proceso se inicia gracias a la caperuza presente en el ARNm, donde se produce la unión entre el ARNm y la subunidad menor del ribosoma. En este punto la subunidad pequeña del ribosoma se desplaza sobre el ARNm hasta que encuentra un triplete o codón de iniciación, en eucariotas este triplete es el AUG que codifica para el aminoácido metionina. Una vez identificado un ARNt iniciador que porta el aminoácido de la metionina se empareja con el codón AUG presente en el ARNm gracias al anticodón presente en el ARNt, y en este punto se crea el complejo de iniciación de la traducción. Tras esto, se incorpora también la subunidad ribosómica mayor. En este paso el ARNt iniciador está situado en el sitio P (peptidil) del ribosoma, y el sitio A (aminoacil) está vacío. 2. **Elongación:** durante este proceso los siguientes ARNt con los aminoácidos que se vayan a incorporar se introducirán en el sitio A de forma repetitiva. Las peptidiltransferasas son las encargadas de formar el enlace peptídico entre los aminoácidos presentes en el sitio P y A. Mientras ocurre este proceso de unión, se separan el aminoácido y el ARNt presentes en el sitio P, y el ARNt sin aminoácido se desplaza al sitio E (del inglés exit), y el ARNt unido al dipéptido se desplaza al sitio P. Tras esto, el complejo formado por ARNm y ARNt unido al péptido se desplaza al siguiente trinucleótido del ARNm y se incorpora al sitio A el nuevo ARNt unido al aminoácido correspondiente. Este proceso continuará de forma secuencial. 3. **Terminación:** este proceso ocurre cuando el complejo se desplaza hasta una zona con uno de los tres posibles codones terminadores, UAA, UAG o UGA, cuyos ARNt carecen de aminoácido. Cuando uno de estos ARNt se introduce en el sitio A, se activan los factores de terminación, que son unas proteínas que cortan la cadena polipeptídica del ARNt, liberándose así la proteína sintetizada.

También intervienen en este proceso los ARN de transferencia (ARNt), los cuales tienen secuencias invertidas que les permiten formar estructuras secundarias dándoles una apariencia de trébol de tres hojas. Los ARNt presentan en su extremo 3' una secuencia que termina en CCA, donde las aminoacil-ARNt-sintetasas son capaces de unir a esta última adenina un aminoácido mediante un enlace covalente. El aminoácido que se une específicamente a este ARNt depende de una secuencia de tres nucleótidos denominada anticodón. Gracias al descubrimiento del código genético, a día de hoy conocemos para que aminoácido codifica cada uno de los 61 tripletes de ADN que existen (Espino, 2012).

### *Mutaciones del ADN.*

La secuencia del ADN no es inalterable, en ella pueden aparecer cambios y éstos pueden ser heredados por la descendencia. Cuando la secuencia de un individuo sufre una alteración, ésta se denomina mutación. Dependiendo de en qué célula del organismo esté presente la mutación, ésta puede ser somática (células no reproductivas), las cuales no se transmiten a la descendencia, o germinal (en células reproductivas), las cuales sí podrían transmitirse a su descendencia y estar presentes en todas las células de los nuevos organismos. Las mutaciones somáticas tienden a acumularse con el paso del tiempo en nuestras células y la gran mayoría de ellas son inofensivas, sin embargo, algunas pueden afectar a genes o elementos reguladores de la transcripción y producir consecuencias biológicas que sean perjudiciales para el organismo. Un ejemplo de la consecuencia de evolución somática es el envejecimiento e incluso el cáncer (Martincorena y Campbell, 2015). También puede ocurrir que una mutación sea neutra y no tenga ningún efecto sobre la supervivencia del individuo, o incluso pueda llegar a ser ventajosa. Como se ha mencionado previamente, las mutaciones pueden transmitirse a la descendencia, y esto puede llegar a producir que con el paso del tiempo la frecuencia de ciertas mutaciones pueda ser mayor del 1% y fijarse en las poblaciones. Cuando esto ocurre y la alteración de la secuencia pasa a ser una característica poblacional entonces se denomina polimorfismo (Karki *et al.*, 2015)

Existen diferentes tipos de mutaciones génicas, que son aquellas que afectan a nucleótidos aislados y provocan cambios en la estructura del ADN, y que se describen a continuación (Figueras, 2012):

1. Sustitución: en ellas una única base se ve afectada. Pueden ser transiciones si el cambio es entre bases púricas, o entre bases pirimidínicas, o tranversiones si es una purina por una pirimidina o viceversa.
2. Inserción y/o deleción: adición o eliminación respectivamente de uno o más nucleótidos a la secuencia del ADN. Si el número de bases ganadas o perdidas no es un múltiplo de tres esto puede desencadenar un cambio de la pauta de lectura del ADN.
3. Duplicaciones: repeticiones de un segmento de ADN.
4. Inversión: un segmento de ADN invierte en 180 grados su orientación.
5. Transposición: un transposón es un elemento de ADN que se pueden mover dentro o entre genomas. Al cambiar de posición pueden afectar a secuencias codificantes de genes.

Las mutaciones también se pueden clasificar en función del efecto que provocan:

1. Mutaciones de pérdida o ganancia de función génica: en el primer caso se altera algún codón del producto proteico, afectando a la funcionalidad de la proteína. En el segundo, la mutación puede afectar a la zona de regulación de transcripción del gen aumentando la síntesis o expresión, y por tanto la cantidad de producto proteico.

## Introducción

2. Mutaciones nulas o sin sentido: produce la formación de un codón de finalización de la traducción y por tanto el producto proteico se interrumpe antes de su finalización.
3. Mutación supresora: se genera una nueva mutación que suprime el efecto de una mutación previa. Si ocurre exactamente en el mismo punto de la mutación original se denomina retromutación.
4. Cambio de pauta de lectura: las inserciones o deleciones de un número de bases que no sea múltiplo de 3 ocasionan un cambio en el marco de lectura, afectando así a la transcripción de todo el producto genético.
5. Mutación silenciosa: el cambio codifica para el mismo codón que la secuencia original y, por tanto, no se ve afectado el producto proteico.

Los mecanismos por los cuales puede surgir una mutación son variados. Por un lado, las mutaciones espontáneas son aquellas que aparecen debido a los procesos biológicos o químicos del organismo y que alteran la estructura del ADN. Estas ocurren durante el proceso de replicación; las enzimas responsables, las polimerasas, no son 100% eficaces y pueden introducir una base errónea durante este proceso, tal y como se mencionó anteriormente. Por otro lado, las mutaciones inducidas son debidas al daño causado por agentes mutágenos como:

1. Agentes físicos: como por ejemplo la radiación, que puede generar moléculas conocidas como radicales libres. Los radicales libres son moléculas con electrones desapareados, lo que les permite interactuar con moléculas como el ADN y producir cambios en su secuencia. Entre los radicales libres destacan el  $H^+$ , los radicales libres de nitrógeno (RNS del inglés *radical nitrogen species*) como por ejemplo el óxido nítrico, y las ROS, como el ion superóxido ( $O_2^-$ ), el peróxido de hidrógeno ( $H_2O_2$ ) o el ion hidroxilo ( $OH^-$ ), siendo este último el más reactivo y destructivo de los que se forman durante el metabolismo celular. Se estima que las ROS generan hasta 50.000 lesiones en el ADN por célula y día, produciendo modificaciones de base, roturas de las cadenas de ADN y enlaces intercatenarios, los cuales pueden afectar a la replicación del ADN (Maynard *et al.*, 2015).
2. Agentes químicos: son aquellos que pueden entrar al organismo mediante ingestión, inhalación o absorción a través de la piel y, una vez dentro pueden producir lesiones en el ADN. Un ejemplo son los análogos de base, que tienen estructuras similares a las bases nitrogenadas y se pueden incorporar al ADN durante la replicación fomentando en futuras replications el mal apareamiento de bases. Otro ejemplo muy conocido es el bromuro de etidio, el cual se puede intercalar con la secuencia de ADN modificando la pauta de lectura.

Para combatir esta ingente cantidad de modificaciones, la célula cuenta con diversos mecanismos de reparación. En el caso de los mamíferos existen cuatro mecanismos básicos de reparación del ADN nuclear (Maynard *et al.*, 2015; Figueras, 2012):

1. El sistema de reparación por escisión de base (BER, del inglés *Base excision repair*): elimina principalmente el daño en el ADN debido a procesos oxidativos, agentes alquilantes que producen modificaciones en las bases nitrogenadas de los nucleótidos, rayos X e incluso roturas de una de las cadenas del ADN. La ADN-glicosilasa escinde la base modificada, generándose así un sitioapurínico, si se elimina una purina, o apirimidínico, si se elimina una pirimidina. La endonucleasa AP reconoce el azúcar que no está unido y lo elimina. Finalmente, una polimerasa rellena el espacio vacío con el nucleótido correcto y la ADN ligasa lo sella.



2. El sistema de reparación por escisión de nucleótido (NER, del inglés *Nucleotide excision repair*): elimina principalmente lesiones debidas a radiación ultravioleta como los fotodímeros, y también se encarga de lesiones que afecten a la estructura de la hélice del ADN como los abultamientos. Es muy similar al BER, solo que en este caso un complejo enzimático busca las distorsiones en la doble hélice, una vez detectado otro complejo separa las dos hebras y una endonucleasa elimina un segmento de ADN con varios nucleótidos. Finalmente, una polimerasa rellena el espacio vacío con el nucleótido correcto y la ADN ligasa lo sella.
3. El sistema de reparación de errores de apareamiento (MMR, del inglés *Mismatch repair*): elimina errores debidos a fallos de apareamiento entre bases durante la replicación. Este sistema de proteínas detecta la distorsión producida en la estructura del ADN por el apareamiento incorrecto, elimina la base insertada erróneamente e introduce la correcta.
4. El sistema de reparación de roturas de doble cadena (DSBR, del inglés *Double-strand break repair*): corrige roturas de doble cadena producidas por rayos X y también elimina las uniones intercatenarias. Este sistema es el más desconocido y se sabe que puede funcionar de dos maneras, de forma no homóloga, uniendo de esta forma los extremos rotos, o de forma homóloga, donde se usa la información de la otra cromátida para reparar el daño.

La acumulación de daño en el ADN puede producir inestabilidad genómica e inducir cascadas de señalización celular que desembocan en procesos de senescencia celular o incluso muerte celular. Los sistemas de reparación tienden a declinar con el envejecimiento celular, llevando a la aparición de mutaciones que pueden ser oncogénicas. Tanto la senescencia como la muerte celular son procesos que juegan un papel fundamental en el envejecimiento.

Como se ha mencionado en este apartado, la presencia de mutaciones en el genoma puede alterar el riesgo de desarrollar diversas enfermedades. Esto ha provocado que tanto la ciencia como la medicina moderna se hayan interesado enormemente en la genética de los individuos con el fin de predecir, prevenir o incluso desarrollar curas para distintas enfermedades. Para ello ha sido necesario profundizar en el estudio de la secuencia del ADN humano, desarrollándose con el tiempo diversas técnicas y nuevos enfoques de estudio.

### *Estudios GWAS y artrosis*

Dos de las metodologías más empleadas en la década de los 90 para analizar las asociaciones entre mutaciones y enfermedades fueron el enfoque de mapeo por ligamiento y la estrategia de genes candidatos (Kwon y Goate, 2000). Aunque ambas metodologías tienen como objetivo encontrar polimorfismos que pueden asociarse con el riesgo de sufrir diversas patologías, ambas funcionan de forma antagónica.

El enfoque de mapeos por ligamiento consiste en escanear de forma sistemática fragmentos del genoma de numerosos miembros de familias afectadas por un desorden, usando *oligonucleótidos* que anclan en posiciones conocidas. Con este enfoque los investigadores pueden comparar la frecuencia de ciertas mutaciones en individuos pertenecientes a familias afectadas, y en caso de ser muy elevada se podría asociar esta variante con la enfermedad. La gran ventaja que tienen este tipo de estudios es que no necesitan conocimiento previo de la enfermedad, sin embargo, necesitan un gran número de individuos pertenecientes a familias afectadas y no afectadas.

## Introducción

La estrategia de genes candidatos, sí necesita del conocimiento previo de la patología, pero sin embargo no es necesario que los individuos del estudio estén emparentados. En esta segunda aproximación lo que se hace es seleccionar genes candidatos de estudio en función de la patología; por ejemplo, en el caso del alcoholismo, se podrían estudiar genes asociados al metabolismo del alcohol como la alcohol-deshidrogenasa. Después sería necesario seleccionar los polimorfismos del gen a estudiar dependiendo de los posibles efectos que causarían en la proteína resultante, y finalmente, se estudiaría la frecuencia de los polimorfismos de interés en un grupo formado por individuos afectados por la enfermedad e individuos sanos.

Estas dos estrategias, a pesar de ser efectivas, se centraban de forma específica en unas pocas regiones del genoma, resultando en un número limitado de asociaciones. Con la obtención de la primera secuencia completa del genoma humano en el año 2001, el mundo de la genética sufrió una revolución tecnológica, y como resultado se comenzaron a cimentar las bases de lo que a día de hoy se conocen como estudios GWAS. Estos estudios consisten en análisis estadísticos con un gran tamaño muestral, compuesto por un grupo de individuos sanos (grupo de control), un grupo de individuos enfermos (grupo de casos), y una ingente cantidad de SNPs. Los primeros intentos de estudios GWAS aparecieron al principio de la década de los 2000, y se basaban en métodos de PCR multiplex de alto rendimiento en los que se analizaban numerosos SNPs presentes en genes preseleccionados. Fue en el año 2005 cuando se publicó el primer trabajo que sirvió como referente de los estudios GWAS. El trabajo consistió en la identificación de SNPs asociados con la degeneración macular asociada a la edad (Narayanan *et al.*, 2007). Este estudio se diferenció de otros que había en ese momento en dos aspectos, por un lado, en vez de buscar SNPs en genes preseleccionados buscaron polimorfismos en zonas aleatorias de genes y regiones intergénicas y, por otro lado, utilizaron una metodología de análisis basada en *arrays* de SNPs que se explicará más adelante en este apartado. Inicialmente, en los primeros GWAS se analizaban cientos de miles de SNPs en cohortes de individuos de aproximadamente 100 casos y 100 controles.

Otro hito que continuó la revolución del campo de la genética sucedió entre los años 2005 y 2007, donde se finalizaron respectivamente la fase I y fase II del proyecto internacional HapMap (del inglés, *haplotype map*). Este ambicioso proyecto tenía como objetivo la validación y caracterización de millones de SNPs, y estudiar sus patrones de desequilibrio de ligamiento en poblaciones con ascendencia europeas, asiáticas y africanas. El concepto de desequilibrio de ligamiento quiere decir que, aquellos SNPs que estén muy próximos en los cromosomas, durante el proceso de recombinación de las cromátidas hermanas en la gametogénesis van a migrar juntos, y que, por tanto, si está presente el SNP diana en la descendencia, los que están en desequilibrio de ligamiento con éste también estarán presentes, aunque no estemos analizándolos directamente, lo que se conoce como un enfoque de asociación indirecta (Fig. 14). Los resultados de este estudio promovieron la generación de los bloques de desequilibrio de ligamiento de SNPs, y ahora en un único *array* de secuenciación se podían analizar muchísimos más SNPs entre el análisis directo e indirecto, lo que supuso que en el año 2007 sucediera el “boom” de los estudios GWAS, conociéndose este año como “El año de los GWAS” (Ku *et al.*, 2010; Uitterlinden, 2016).

Con el paso del tiempo, con el incremento en el número de SNPs estudiados y la mejora de los *arrays* de secuenciación, también se incrementó el número de individuos que se analizaban en los estudios GWAS, pasando de los aproximadamente 100 casos y 100 controles, a estudios con cientos de miles de controles y casos como el estudio de Hyde y colaboradores de 2016 donde trataban de buscar SNPs asociados con la depresión en individuos con

ascendencia europea (Hyde *et al.*, 2016). A pesar de que la proporción entre casos y controles *gold standard* es la 1:1, en términos de poder estadístico, la más óptima es la relación 1:4 (Uitterlinden, 2016).

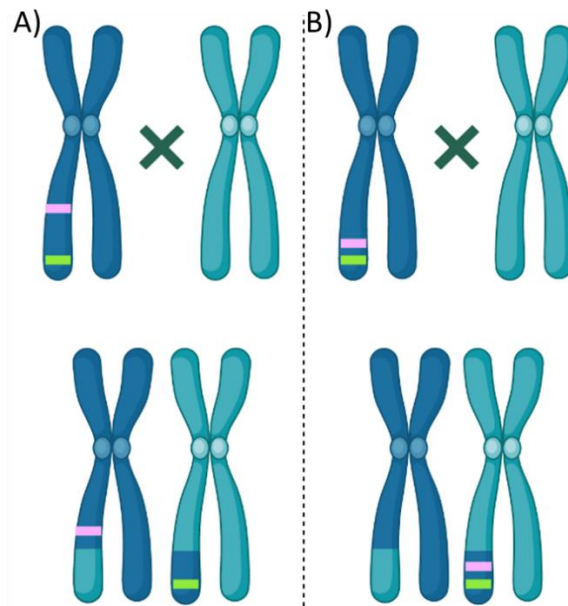


Figura 14. Representación del desequilibrio de ligamiento durante la recombinación cromosómica. A) Cuando dos SNPs están muy alejados físicamente en la cromátida, en caso de que se produzca una recombinación con la cromátida de otro cromosoma, estos SNPs serán segregados de forma independiente. B) Sin embargo, si estos dos SNPs están muy próximos físicamente, sucederá un proceso de desequilibrio de ligamiento, donde debido a la proximidad, durante el proceso de replicación ambos SNPs se segregan de forma conjunta, de tal manera que siempre que esté presente uno, también estará el otro. Imagen creada con [app.biorender.com](http://app.biorender.com).

Como se comentó anteriormente, para analizar esta gran cantidad de polimorfismos se usan los denominados *arrays* de genotipado de SNPs. Éstos consisten en *oligonucleótidos* que están unidos a un soporte sólido. Dichos *oligonucleótidos* son específicos de una secuencia pequeña de aproximadamente 50 pb que está próxima al SNP de interés. El proceso de genotipado consiste en varios pasos: primero, es necesario realizar una amplificación por PCR de los distintos fragmentos de interés de la muestra; tras esto, la muestra es introducida en el *array* para que los fragmentos de ADN de interés se unan a los oligonucleótidos del *array* dependiendo de si poseen alguno de los SNPs de interés o no, además durante esta hibridación se usan dideoxinucleótidos marcados con fluorocromos para identificar los fragmentos que se han unido a algún oligonucleótido del *array*. A continuación se realiza el lavado del chip para eliminar aquellos fragmentos que no se han unido y, finalmente, se usan láseres de alta definición para detectar los productos marcados con el fluorocromo y así identificar qué SNPs están presentes en la muestra (Uitterlinden, 2016)(Fig. 15).

En estos *arrays*, muchos de los oligonucleótidos incorporados identifican SNPs que están en desequilibrio de ligamiento con otros SNPs cercanos, permitiendo así la identificación de una gran cantidad de SNPs. Por ejemplo, el chip *Genome-Wide Human SNP Array 6.0* de Affymetrix (EEUU) es capaz de detectar 906.000 SNPs y más de 946.000 variantes de número de copias (CNV, del inglés *copy number variation*). Con los datos obtenidos del *array* se realiza un test de

## Introducción

asociación entre los casos y los controles, normalmente se emplean gráficos QQ o Manhattan (Fig. 16), donde se otorga un p-valor para identificar si un SNP se asocia de forma significativa o no con el grupo de casos. Debido a que se realizan múltiples comparaciones sobre la misma muestra es necesario realizar una corrección del p-valor, normalmente es necesario un p-valor igual o inferior a  $5 \cdot 10^{-8}$  para considerar como significativa una asociación de genoma completo. Este valor viene derivado de dividir el p-valor más ampliamente utilizado, de 0,05, entre 1 millón de test independientes debidos a los bloques de SNPs en desequilibrio de ligamiento. Cuando se obtienen datos de múltiples GWAS que de forma independiente tratan de buscar asociaciones para el mismo evento, es posible realizar metodologías de meta-análisis, las cuales permiten combinar todos estos datos independientes para conseguir aumentar la robustez estadística. Además, con el fin de sacar conclusiones que afecten a una población en general es importante replicar los datos obtenidos en otras cohortes externas, con el fin de cerciorarse que lo observado se reproduce en otras poblaciones (Uitterlinden, 2016).

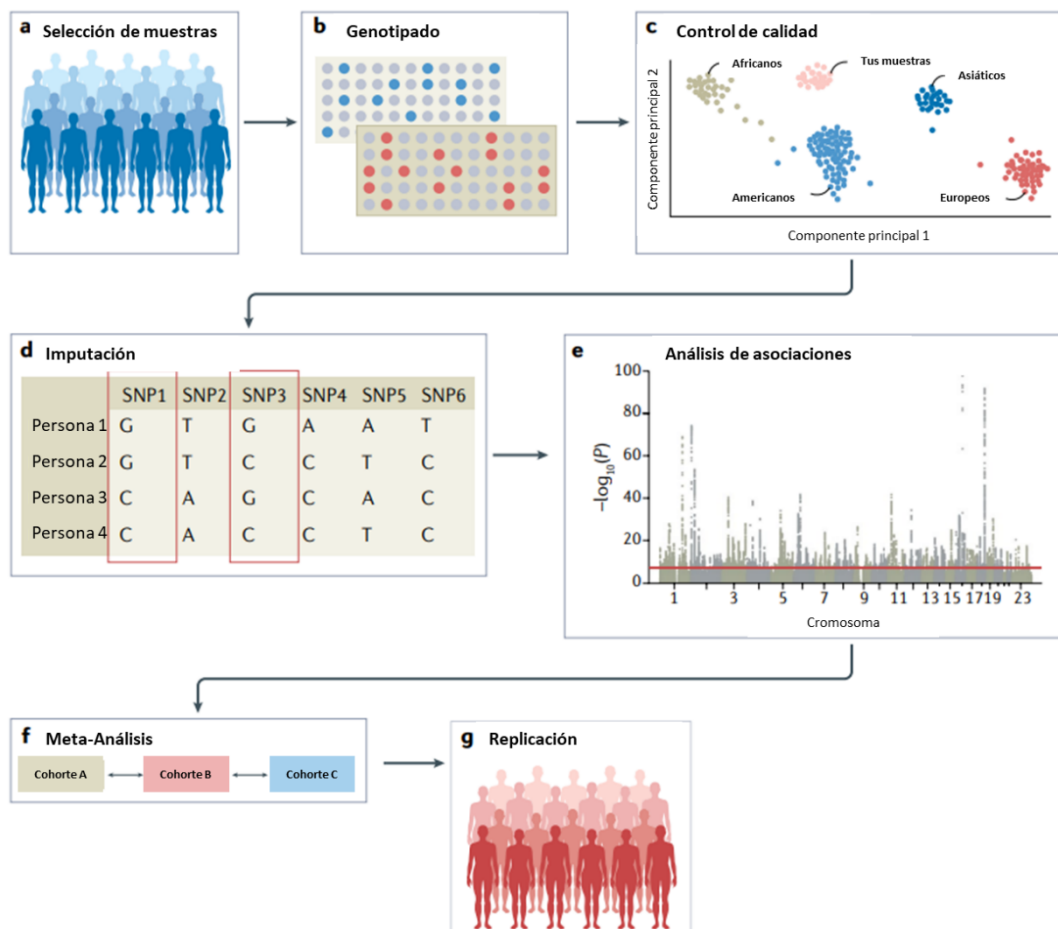


Figura 15. Resumen de los pasos para llevar a cabo un estudio GWAS. A) Se pueden obtener muestras de cohortes de casos y controles o, directamente acceder a los datos genéticos y fenotípicos disponibles en los biobancos. B) Los datos genéticos se pueden obtener mediante arrays o métodos de secuenciación masiva. C) Se realiza un control de calidad para comprobar las señales genéticas obtenidas, también se realizan pruebas de componentes principales con el objetivo de comprobar la agrupación poblacional de las muestras de estudio. D) Se analizan los datos genéticos obtenidos y se comparan con la secuencia de referencia de repositorios como el proyecto de 1000 genomas. E) Se realizan tests de asociación genética para cada variante. F) Los resultados de múltiples cohortes se pueden combinar para ganar robustez estadística. G) Se realiza una replicación de los datos, ya sea de forma interna o externa. Imagen modificada de (Uffelmann et al., 2021).

Gracias a este tipo de estudios se han identificado muchas variantes asociadas a rasgos fenotípicos. Con fecha del 11 de enero de 2019 se habían realizado un total de 3.730 GWAS, pero desde entonces el número de estudios ha continuado aumentando y el 26 de agosto de

2021 se publicó una revisión de estudios GWAS donde el número de GWAS realizados ascendía a más de 5.700. A 16 de junio de 2022, las estadísticas recogidas en la base de datos *GWAS Catalog* (<https://www.ebi.ac.uk/gwas/home>) muestran que, con todos los estudios GWAS realizados, se han generado un total de 30.315 publicaciones. Todo este esfuerzo ha permitido que decenas de miles de SNPs se hayan asociado con múltiples enfermedades como la anorexia nerviosa, el desorden depresivo, cánceres y subtipos de cánceres, diabetes mellitus tipo 2, enfermedades de las arterias coronarias, esquizofrenia, enfermedad intestinal inflamatoria, insomnio, índice de masa corporal y otras muchas (Bosse y Amos, 2018; Uffelmann *et al.*, 2021). La identificación de dichos SNPs ha permitido identificar nuevos mecanismos biológicos detrás de diversas patologías y, además, muchos de estos SNPs tienen aplicabilidad clínica ya que pueden ser usados para identificar individuos con elevado riesgo de desarrollar patologías. Cabe destacar que, hasta la fecha, la mayoría de los GWAS se han centrado en el estudio de SNPs, pero también se pueden estudiar otro tipo de alteraciones como los CNV, mosaicismo clónico, haplotipos, variaciones en el número de repeticiones tándem, transposones, inserciones, deleciones e incluso inversiones.

Sin embargo, este tipo de estudios también tiene sus desventajas, siendo la primera la gran corrección que se tiene que hacer al p-valor, lo cual hace necesario tener cohortes de individuos muy grandes. Además, estos tests detectan muy pocas señales, lo cual representa una heredabilidad estimada muy baja para aquellos rasgos que son complejos. Otro problema añadido es que los GWAS son incapaces de detectar los fenómenos de epistasia en humanos, es decir, que la expresión de uno o más genes dependen de la expresión de otro gen. Por último, la mayoría de las señales detectadas tienen un efecto muy pequeño en el desarrollo de una patología, y por tanto su valor clínico predictivo, en determinados casos, puede ser limitado (Tam *et al.*, 2019).

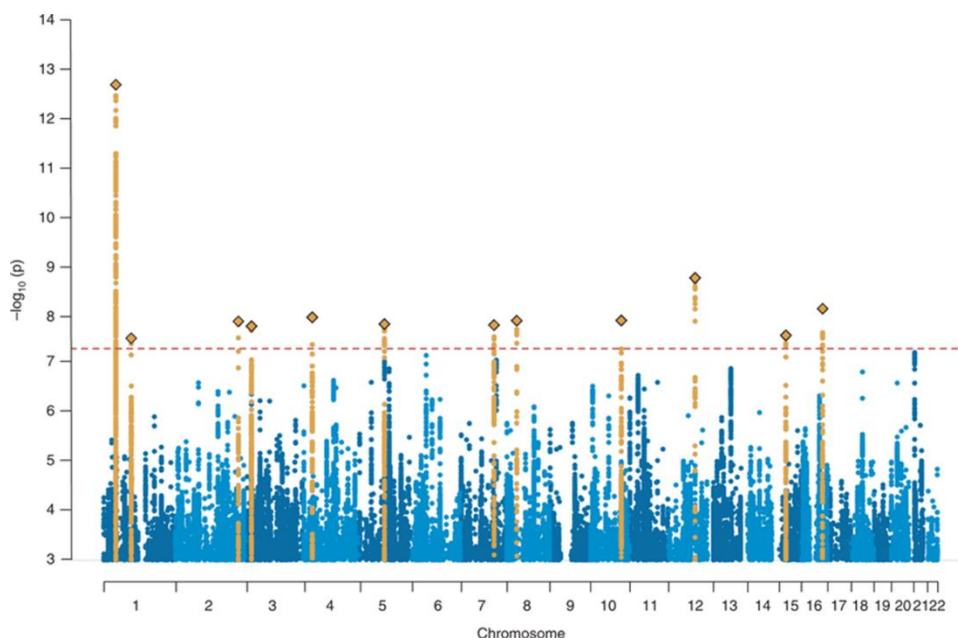


Figura 16. Ejemplo de un gráfico Manhattan. Cada uno de los puntos que se observa en la imagen representa una señal obtenida de un SNP, su posición en el eje X permite identificar en qué cromosoma humano está presente dicho polimorfismo, su altura en el eje Y representa el p-valor de su asociación en escala logarítmica. Aquellas señales que sean capaces de sobrepasar la línea discontinua serán las que alcancen unos niveles significativos tras la corrección del p-valor. Imagen obtenida de Demontis y colaboradores 2019 (Demontis *et al.*, 2019).

## Introducción

Los estudios GWAS también han sido empleados para identificar SNPs que puedan ser la causa de la artrosis en distintas articulaciones ya que los polimorfismos pueden ser específicos de mano, cadera o rodilla. Hasta el año 2021, se han asociado un total de 176 SNPs con la artrosis (Boer *et al.*, 2021; Aubourg *et al.*, 2022), los cuales afectan a loci distribuidos por todo el genoma, siendo el que más veces se ha asociado de forma robusta con la artrosis de rodilla el SNP rs143383 de GDF5 (Valdes *et al.*, 2011; Pan *et al.*, 2014; Jiang *et al.*, 2016). El efecto que confieren dichos alelos tiende a ser pequeño, siendo la mayoría de los estimadores de riesgo (OR, del inglés *Odds ratios*) menores de 1,5. A pesar de estos bajos efectos, se considera a la artrosis como una enfermedad poligénica, en la que el efecto conjunto de múltiples SNPs puede acabar desencadenando la patología.

Uno de los GWAS pioneros, y por ello de los más importantes en el campo de la artrosis, es el del estudio arcOGEN de 2012, que tenía como objetivo buscar SNPs que se asociaran con la susceptibilidad a desarrollar la enfermedad (Zeggini *et al.*, 2012). En el año 2017, Warner y Valdés (Warner y Valdés, 2017) realizaron una revisión de distintos estudios GWAS centrados en la susceptibilidad de la artrosis. La figura 17 pertenece a este artículo y muestra algunos de los genes que presentan SNPs asociados de forma robusta, en distintos estudios GWAS, con la susceptibilidad a desarrollar artrosis en distintas articulaciones. Pero estos estudios focalizados en artrosis no se han centrado solamente en la susceptibilidad. El estudio Arthrotest de 2015 (Blanco *et al.*, 2015) se centró en la búsqueda de SNPs que se asociasen con el riesgo de mal pronóstico, esto es, la probabilidad de incluir una prótesis articular en el paciente artrósico. Como resultado de dicho estudio se encontraron un total de 23 SNPs que se asociaron con este tipo de progresión. El estudio GWAS más reciente se publicó en septiembre de 2021 y en él se meta-analizaron los genotipos de un total de 826.690 individuos (de los cuales 117.517 tenían artrosis) e identificaron un total de 100 variantes asociadas de forma independiente con 11 fenotipos de artrosis, 52 de estas variantes no se habían asociado previamente con la enfermedad. La señal más robusta que encontraron fue el SNP rs747952496, que es una variante intrónica que se asoció de forma significativa con el riesgo de desarrollar artrosis de cadera (p-valor=4,91 x 10<sup>-11</sup>; OR=7,02, IC95%=3,93–12,55), seguida de otra variante intrónica, rs571734653, que se asoció de forma significativa con el riesgo de desarrollar el artrosis de rodilla (p-valor=5,56 x 10<sup>-9</sup>; OR=6,03 IC95%=3,30–11,03)(Boer *et al.*, 2021).

En el año en que comenzó este proyecto de tesis doctoral, Warner y Valdés publicaron una revisión sobre los SNPs que de forma más robusta se habían asociado hasta la fecha con la susceptibilidad a desarrollar artrosis de rodilla, entre otras articulaciones (Warner y Valdes, 2017). Entre los SNPs que los autores destacaron en este estudio, para la realización de esta tesis doctoral se seleccionaron siete con el fin de analizar su posible asociación con el fenotipo de artrosis de rodilla con progresión rápida. Más adelante, y debido a la importancia que tiene el endotipo asociado a la reducción del mJSW en la progresión rápida, se decidió añadir otro polimorfismo asociado de forma robusta con esta característica, el SNP rs3771501 del gen TGFA (del inglés, *Transferring Growth Factor Alpha*). A continuación, se detallan los datos de los ocho polimorfismos seleccionados:

1. rs12107036 del gen TP63 (del inglés *tumor protein 63*): este polimorfismo consiste en una transición de purinas A→G. El alelo G de este SNP se asoció robustamente con el riesgo de desarrollar cirugía de reemplazo articular de rodilla en mujeres (Zeggini *et al.*, 2012)
2. rs8044769 del gen FTO (del inglés *fat mass and obesity-associated protein*). consiste en una transición de pirimidinas C→T. En el estudio GWAS arcOGEN (Zeggini *et al.*, 2012),

- el alelo C de este SNP se asoció de forma muy robusta con el riesgo de desarrollar artrosis en mujeres, y se observó una asociación fuerte entre el IMC y este SNP.
3. rs11177 del gen GNL3 (del inglés *guanine nucleotide binding protein-like 3*): el SNP rs11177 de este gen consiste en un cambio entre bases pirimidínicas C→T. En el estudio arcOGEN (Zeggini *et al.*, 2012) el alelo T de este polimorfismo se asoció con un aumento del riesgo de desarrollar artrosis de rodilla y de cadera en una población europea. En el estudio de Liu y colaboradores de 2018 (Liu *et al.*, 2018) también se asoció este SNP con el riesgo aumentado de artrosis de rodilla en una población china.
  4. rs4730250 del gen DUS4L (del inglés *dihydrouridine synthase 4-like*): Este polimorfismo consiste en una transición entre purinas A→G, y es el alelo G el que se asocia con susceptibilidad a artrosis de rodilla (Evangelou *et al.*, 2011).
  5. rs143383 del gen GDF5: como se mencionó previamente, este polimorfismo es el que se ha asociado, de manera reiterada y robusta, con la artrosis de rodilla, pero también se ha asociado con la artrosis de cadera. El polimorfismo rs143383 consiste en una transición entre bases pirimidínicas C→T, y está presente en la región promotora del gen.
  6. rs10948172 del gen SUPT3H (del inglés *SPT3 Homolog*): consiste en una transición entre bases púricas A→G. La primera asociación entre el alelo G de este polimorfismo y la artrosis de rodilla se llevó a cabo en una población caucásica utilizada en el estudio arcOGEN en 2012 (Zeggini *et al.*, 2012).
  7. rs11842874 del gen MCF2L (del inglés *MCF2 Transforming Sequence-Like Protein*): este polimorfismo consiste en una transición entre purinas A→G, siendo el alelo A el que se ha asociado con el riesgo de desarrollar artrosis de rodilla y cadera (Shepherd *et al.*, 2015).
  8. rs3771501 del gen TGFA: el polimorfismo rs3771501 de este gen consiste en una transición entre pirimidinas C→T. Concretamente el alelo T de este SNP se ha asociado de forma muy robusta con la reducción del mJSW y con el riesgo de desarrollar artrosis en cualquier articulación (Zeggini *et al.*, 2018).

# Introducción

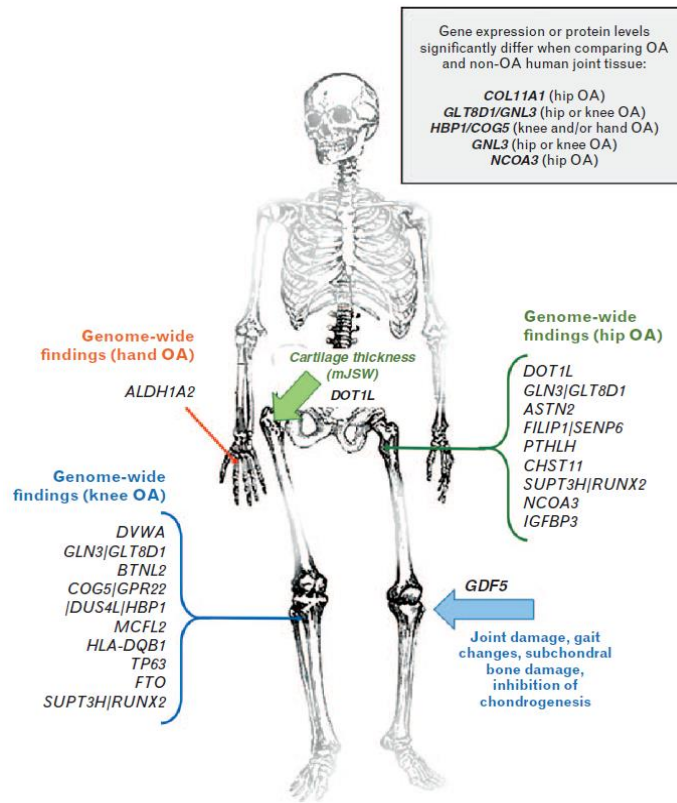


Figura 17. Resumen de genes en los que se han encontrado polimorfismos que se asociaron con la susceptibilidad a la artrosis en distintas articulaciones. Imagen obtenida de Warner y Valdés 2017 (Warner y Valdes, 2017).



## La mitocondria.

### *Origen, estructura y función de las mitocondrias.*

El origen de este orgánulo celular se explica mediante la teoría endosimbionte, que fue propuesta por Lynn Margulis en 1967 (Sagan, 1967). En ella se teoriza sobre el origen de la célula eucariota hace aproximadamente 1.200 millones de años. Esto ocurrió en un ambiente muy primitivo rico en oxígeno, donde unas células heterótrofas ingirieron microbios aerobios y establecieron entre ellos una relación simbiótica. Con el paso de millones de años, esta unión evolucionó y se produjeron una gran cantidad de cambios evolutivos graduales a medida que el endosimbionte y el huésped se integraron. Este proceso sigue siendo estudiado a día de hoy, y diversos estudios apuntan a que las células eucariotas ya estaban presentes previamente a estas condiciones ricas en oxígeno (Brocks *et al.*, 1999; Roger, Muñoz-Gómez y Kamikawa, 2017). Como resultado, las células heterótrofas obtuvieron las mitocondrias y cloroplastos, estos últimos presentes sólo en las células de las plantas y, se formó la célula eucariota.

Centrándonos en las mitocondrias, su tamaño oscila entre 0,5 y 1  $\mu\text{m}$  de diámetro y su abundancia dentro de la célula es muy variada, pudiendo depender del organismo, tipo de tejido y el tipo celular. Las únicas células que carecen de mitocondrias son los glóbulos rojos. La forma y el tamaño de estos orgánulos es muy variable, debido a la gran plasticidad que les otorgan los procesos de fusión y fisión a los que están sometidas constantemente para poder responder de forma eficiente a los requerimientos energéticos de la célula (Sauvanet *et al.*, 2010). Este orgánulo está envuelto por una doble membrana, una membrana externa, la cual está en contacto con el citosol y regula el intercambio de metabolitos entre la segunda membrana, o membrana interna, y el citosol. Ambas membranas están separadas por el espacio intermembrana. La membrana interna mitocondrial presenta un gran número de invaginaciones denominadas “crestas mitocondriales”, y es en estas estructuras donde están localizadas las proteínas de la cadena de transporte de electrones (CTE), necesaria para la producción de energía. Ambas membranas delimitan en su interior la matriz mitocondrial, que contiene el ADN mitocondrial (ADNmt), los mitorribosomas y el ARN mitocondrial (ARNmt), y es en esta zona donde tiene lugar el ciclo de Krebs y la beta-oxidación de los ácidos grasos (Fig. 18) (Rodríguez-Violante *et al.*, 2010; Bartolák-Suki *et al.*, 2017).

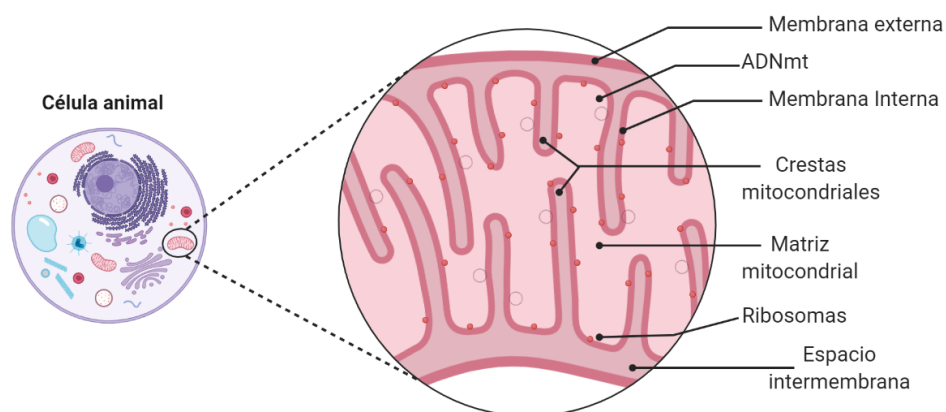


Figura 18. Esquema representativo de la estructura de la mitocondria. Imagen creada con [app.biorender.com](https://www.biorender.com).

La función principal de la mitocondria es la generación de energía en forma de adenosín-trifosfato (ATP, del inglés *Adenosine triphosphate*), motivo por el cual reciben el título de la planta energética de la célula, pero además es importante destacar su implicación en otras funciones (Murphy *et al.*, 2016):

## Introducción

1. Generación de ATP (Feduchi *et al.*, 2011): gracias a los procesos de glucólisis, la desaminación oxidativa de los aminoácidos y la  $\beta$ -oxidación de los ácidos grasos, se obtienen moléculas de acetil-CoA. Esta molécula es empleada durante el ciclo de Krebs en el interior de la mitocondria de modo que, por cada acetil-CoA se obtienen tres moléculas de dinucleótido de nicotinamida adenina (NADH del inglés *Nicotinamide adenine dinucleotide*) o coenzima 1 y una molécula de dinucleótido de flavina adenina (FADH<sub>2</sub> del inglés *Flavin adenine dinucleotide*). Éstas son moléculas reducidas, es decir, tienen ganancia de electrones, y son empleadas en el sistema de transporte electrónico de las crestas mitocondriales para generar ATP. Dicho sistema está compuesto por cuatro complejos transportadores proteicos, que poseen diversos grupos prostéticos capaces de aceptar y donar electrones, una coenzima Q y los citocromos. En este proceso se van a transportar electrones desde las moléculas reducidas hasta el aceptor final de electrones, el oxígeno. Durante este proceso los complejos I, III y IV transportan protones (H<sup>+</sup>) en contra de gradiente al espacio intermembrana. Aprovechando este gradiente de H<sup>+</sup>, se realiza el paso final, denominado fosforilación oxidativa, donde la enzima ATP sintasa, también conocida como complejo V, usa esta fuerza protón-motriz para generar ATP a partir de ADP y fosfato inorgánico. (Fig. 19)
2. Homeostasis del calcio (Ca<sup>2+</sup>): el Ca<sup>2+</sup> funciona como una señal intracelular, regulando

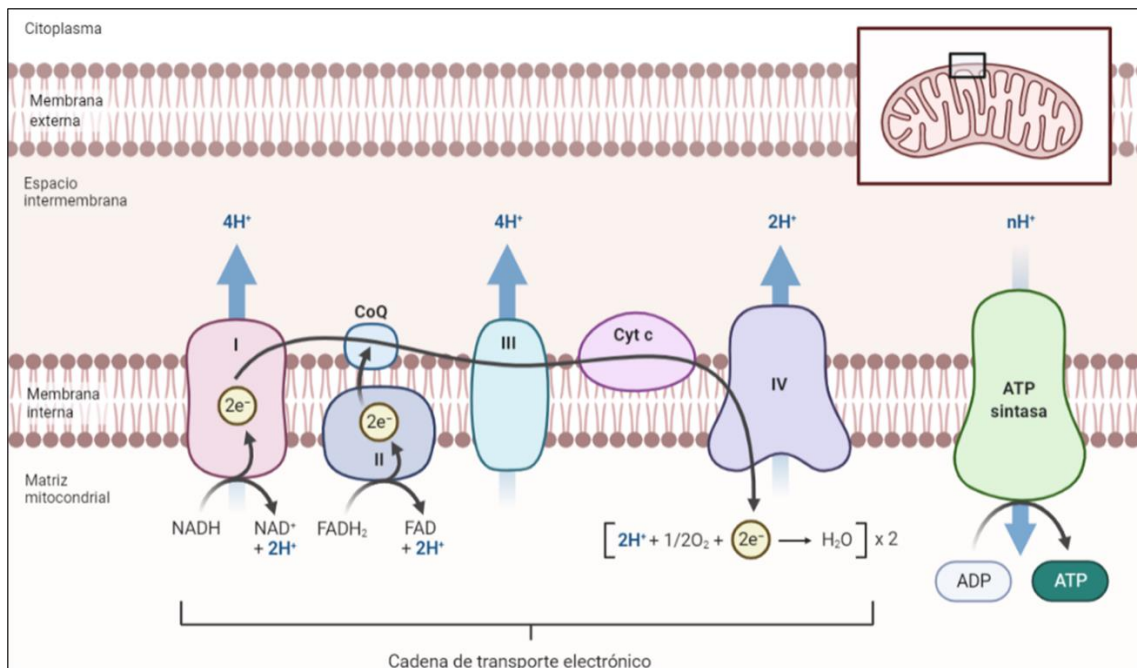


Figura 19. Esquema representativo de la cadena de transporte de electrones y la fosforilación oxidativa. Complejo I, NADH deshidrogenasa: transporta electrones de una molécula de NADH a la coenzima Q; Complejo II, Succinato deshidrogenasa: al igual que el complejo I también transporta electrones a la coenzima Q, pero en vez de ser desde el NADH lo hace desde una molécula de FADH<sub>2</sub>. Una vez reducida por los dos complejos, la coenzima Q difunde libremente por la membrana transportando los electrones hasta el complejo III; Complejo III, citocromo c óxidorreductasa: acopla la transferencia de electrones desde la coenzima Q al citocromo C; Complejo IV, citocromo c oxidasa: una vez que el citocromo C entra en un estado reducido, éste transfiere los electrones cedidos por el complejo III hasta el complejo IV. Finalmente, éste mueve los electrones cedidos por el citocromo C hasta el oxígeno, el cual se reduce a agua; ATP sintasa: aquí se realiza la fosforilación oxidativa, esta enzima usa el gradiente de protones generado para obtener una molécula de ATP a partir de ADP y fosfato inorgánico. Imagen creada con [app.biorender.com](http://app.biorender.com).

diversos procesos celulares como la síntesis de proteínas, la transcripción genética, la remodelación genética, la remodelación celular y la proliferación celular (Berridge, 2012). Además de estas funciones, el Ca<sup>2+</sup> actúa indirectamente en la generación de ATP,

ya que la actividad de la enzima piruvato deshidrogenasa fosfatasa depende de la presencia de este ion, siendo esta enzima la encargada de desfosforilar la piruvato deshidrogenasa, la cual cataliza la transformación de piruvato a acetil-CoA, el sustrato inicial del ciclo de Krebs (Denton *et al.*, 1972).

Es de vital importancia mantener en un correcto equilibrio la concentración citosólica del  $\text{Ca}^{2+}$ , para ello las mitocondrias son capaces de “secuestrar” el  $\text{Ca}^{2+}$  citoplasmático, funcionando como un reservorio natural de este ion y, manteniendo estable su concentración, permiten así el correcto funcionamiento celular. La alteración de la actividad mitocondrial condiciona la homeostasis del  $\text{Ca}^{2+}$ , conduce a un incremento de la concentración de  $\text{Ca}^{2+}$  en el citosol, pudiendo generar consecuencias patológicas (Hernández-Beltrán *et al.*, 2013).

3. Apoptosis celular: el proceso de muerte celular programada tiene como objetivo la eliminación de células dañadas o no deseadas por el organismo. Este mecanismo cuenta con un alto grado de regulación, pudiendo ser activado por señales intracelulares (vía intrínseca) o extracelulares (vía extrínseca). Ambas rutas culminan en la activación de las caspasas, unas proteínas que tienen la capacidad de adherirse a diversos componentes celulares y dismantelarlos.

**La vía intrínseca**, explicada en la figura 20, responde a estímulos como el daño en el ADN, agentes quimioterapéuticos, inanición sérica, radiación UV y daño celular (Kim y Blanco, 2007; Wang y Youle, 2009). El principal orgánulo regulador de la vía intrínseca es la mitocondria, y en ella participan los siguientes elementos:

- Proteínas de la familia Bcl2 (del inglés *B-cell lymphoma 2*): estas proteínas se dividen en tres grandes grupos, las antiapoptóticas, las proapoptóticas como Bax (del inglés *BCL2 Associated X, Apoptosis Regulator*) y Bak (del inglés *BCL2 Antagonist/Killer*), y las BH3 (del inglés *Bcl2 homology region 3*), estas últimas son capaces de inhibir a las Bcl2 antiapoptóticas.
- Caspasas: son proteínas sintetizadas como proenzimas, y que se activan mediante proteólisis irreversible, por tanto, su activación está muy regulada por los inhibidores de apoptosis. En ausencia de señales apoptóticas, los inhibidores de apoptosis suprimen dicho proceso manteniéndose unidas a las caspasas (Hwang y Kim, 2015). Estas proteínas están presentes en distintos compartimentos celulares, por ejemplo las caspasas 7 y 8 están presentes únicamente en el citosol, mientras que la 2, la 3 y la 9 están presentes en la mitocondria y en el citosol (Zhivotovsky *et al.*, 1999). Estas proteínas se pueden clasificar en tres grupos:
  - Iniciadoras: destacan las caspasas 2, 8, 9 y 10. Son importantes en la vía apoptótica extrínseca.
  - Inflamatorias: destacan las caspasas 1, 4, 5, 11, 12 y 13. Las caspasas inflamatorias no están relacionadas con la vía clásica de la apoptosis, sino con la muerte celular inducida por patógeno.
  - Efectoras: destacan las caspasas 3, 6 y 7. Las caspasas efectoras son las que se unen a diversas estructuras celulares para dismantelarlas.
- Citocromo C: además de su rol en la CTE mitocondrial, tiene un papel fundamental en la activación del complejo citosólico de factores activadores de proteasas apoptóticas. Este complejo es capital en la proteólisis y activación de las caspasas.
- SMAC (del inglés *second mitochondria-derived activator of caspase*, también conocida como DIABLO) y HtrA2 (del inglés *High temperature requirement*

## Introducción

*protein A2*, también conocida como Omi): estas proteínas están encerradas dentro de la mitocondria y son capaces de unirse a los inhibidores de apoptosis de las caspasas 3, 7 y 9 para liberarlas y activar el proceso apoptótico.

- Endonucleasa G y el factor de inducción de apoptosis: también están encerradas dentro de la mitocondria, y su función es translocarse al núcleo y fragmentar el ADN nuclear.

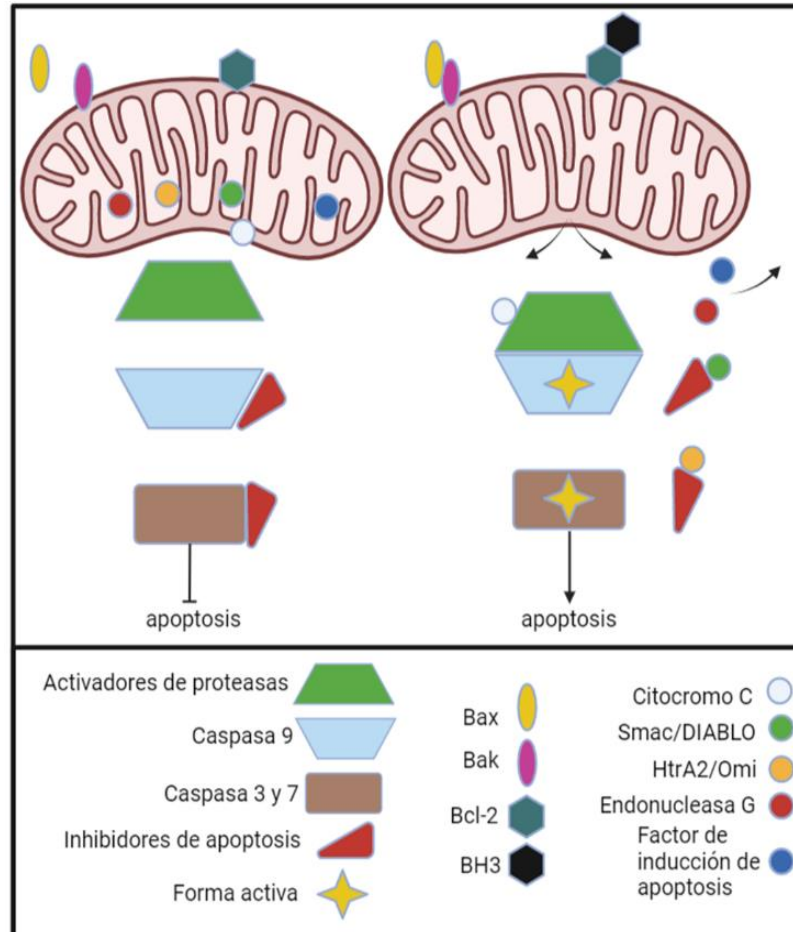


Figura 20. Representación de la vía intrínseca de la apoptosis celular. Existen dos posibles modelos para explicar el inicio de la señal apoptótica en la vía intrínseca, o bien las proteínas BH3 activan a Bax y a Bak, translocándose Bax del citosol celular a la superficie mitocondrial donde sufre cambios conformacional y Bak, que reside en la mitocondria, sufre también cambios conformacionales; o bien BH3 inhibe las proteínas Bcl2 antiapoptóticas, permitiendo así la activación de Bax y Bak. En ambos casos la activación de estas proteínas proapoptóticas promueve la apertura de los poros mitocondriales que conectan la matriz mitocondrial con el exterior, de esta forma aumenta la permeabilidad de la membrana mitocondrial. Con el aumento de la permeabilidad mitocondrial se liberan al citosol el citocromo C, Smac/Diablo, Omi/HtrA2, la endonucleasa G y el factor de inducción de apoptosis. Smac/Diablo y Omi/HtrA2 eliminan los inhibidores de apoptosis que están unidos a la caspasa iniciadora 9 y a las caspasas efectoras 3 y 7. La endonucleasa G y el factor de inducción de apoptosis se translocan al núcleo para iniciar la fragmentación del ADN nuclear. El citocromo C se une y activa el complejo de factores activadores de proteasas apoptóticas, el cual va proteolizar a la caspasa iniciadora 9, activándola, y esta va su vez hace lo mismo con las caspasas efectoras 3 y 7, las cuales comienzan a unirse a otras proteínas efectoras que finalmente desmantelaran el contenido celular (Wang and Youle, 2009; Hwang and Kim, 2015). Imagen creada con app.biorender.com.

En la **vía extrínseca** cobran especial importancia los receptores de muerte, como por ejemplo TNFR (del inglés *tumor necrosis factor receptor*), los cuales se caracterizan por la presencia de un dominio de muerte, un dominio citosólico y un dominio extracelular rico en cisteínas. Cuando el respectivo ligando se une a su receptor se forma un complejo denominado DISC (del inglés *death-inducing signaling complex*), que recluta a una proteína adaptadora denominada FADD (del inglés *Fas-associated protein*

*with death domain*) y también a la caspasa iniciadora 8. Una vez finalizado el reclutamiento, la caspasa 8 es proteolizada y se activa, iniciando una cascada de activación de caspasas efectoras como la 3. De esta forma, al igual que ocurría en la vía intrínseca, se inicia el proceso que dará lugar a la apoptosis.

Algunos autores destacan que puede haber una interconexión entre la vía extrínseca y la vía intrínseca, donde la caspasa 8 se une a la proteína Bid (del inglés *BH3 interacting-domain death agonist*), produciendo un cambio conformacional a tBid y, permitiendo que este complejo se transloque a la mitocondria. De esta forma puede activar a Bax/Bak e iniciar también la vía intrínseca de forma simultánea (Hwang y Kim, 2015).

4. Mantenimiento del estado redox: las moléculas de ROS son un producto natural de la actividad celular. Existen diversas fuentes de ROS, como por ejemplo los iones de metales de transición, la actividad de los peroxisomas, el retículo endoplasmático, el catabolismo de las timidinas y las poliaminas, entre otros. Sin embargo, la fuente más importante de ROS endógeno es la mitocondria (Snezhkina *et al.*, 2020). Esto es debido a que, durante el proceso de transporte electrónico, en diversos puntos de este mecanismo, como el complejo I o el complejo III, puede ocurrir una pérdida de electrones produciendo la reducción parcial del oxígeno, generándose así el anión superóxido ( $O_2^-$ ), el cuál es liberado de forma natural a la matriz mitocondrial. Además el  $O_2^-$  puede reaccionar con el NO y generar peroxinitrito (ONOO<sup>-</sup>), cuyas propiedades químicas lo convierten en un oxidante biológico importante y en un mediador patógeno central en una variedad de estados de enfermedad, incluidos los cardiovasculares, trastornos neurodegenerativos e inflamatorios (Ferrer-Sueta y Radi, 2009). La mitocondria cuenta con diversos mecanismos para protegerse de la sobrecarga de ROS, siendo uno de los más importantes la superóxido dismutasa mitocondrial (MnSOD o SOD2, del inglés *superoxide dismutase 2*), una enzima capaz de transformar el  $O_2^-$  en  $H_2O_2$ . Por su parte, la catalasa descompone en los peroxisomas el  $H_2O_2$  en agua y oxígeno.

El mantenimiento de la concentración de ROS celular es vital para la supervivencia de la célula, ya que tener unos niveles de ROS bajos ayuda a mantener la homeostasis celular y es importante en diversos procesos de señalización celular (Shadel y Horvath, 2015). Sin embargo, la acumulación excesiva de ROS produce daños en las proteínas, lípidos y también en el ADN mitocondrial y nuclear, generando mutaciones somáticas y provocando inestabilidad genómica, lo que puede dar lugar a alteraciones metabólicas. Las consecuencias de estos daños son, por un lado, la despolarización de la membrana mitocondrial y, por otro, la activación del factor nuclear potenciador de las cadenas ligeras kappa de las células B activadas (NFkB del inglés *nuclear factor kappa-light-chain-enhancer of activated B cells*), un importante regulador de la inflamación. Estos procesos pueden culminar en la activación de la apoptosis celular (Mao *et al.*, 2020).

## Introducción

### *El ADN mitocondrial (ADNmt).*

Debido a su origen endosimbionte, la mitocondria posee su propio ADN, que difiere bastante estructuralmente del ADN nuclear. El ADNmt es una molécula circular que está constituida por dos cadenas con diferente composición nucleotídica, la cadena ligera y la cadena pesada. La cadena ligera debe su nombre a su elevada composición de guaninas, mientras que la cadena pesada se lo debe a la elevada composición en citosinas. La molécula de ADNmt (Fig. 21) consta de 16.569 pb y codifica para 37 genes: 13 polipéptidos esenciales de la cadena respiratoria mitocondrial, 22 ARNt y dos ARNr (12S y 16S). Entre los genes mitocondriales del ARNt fenilalanina (m.577 – m.647) y prolina (m.15956 – m.16023) se encuentra una región de 1,1 kb conocida como la región no codificante o región control. Esta parte del ADNmt no codificante se encuentran subdividida en tres zonas que se conocen como segmentos hipervariables; el segmento hipervariable 1, que va desde la posición m.16024 a la m.16383, el segmento hipervariable 2, de la m.257 a la m.372, y el segmento hipervariable 3, de la m.438 a la m.574. Esta región de control contiene los promotores reguladores de la transcripción de la cadena pesada y ligera y, además, también está presente en ella el origen de replicación de la cadena pesada ( $O_h$ ), que es el lugar donde se inicia la replicación. Existe un segundo origen de replicación, el de la cadena ligera ( $O_l$ ), pero no se encuentra en esta zona (Wallace *et al.*, 1999). En la región no codificante puede estar presente una tercera cadena de ADNmt lineal de 650 pb denominada 7S, que produce que se forme una estructura estable conocida como el D-Loop (del inglés “*displacement loop*”). Se estima que aproximadamente el 10% de las moléculas de ADNmt presentes en el ser humano presentan el D-loop en un momento dado. La región no codificante del ADNmt contiene los elementos necesarios para la replicación y la transcripción de este genoma, por tanto la presencia de alteraciones en esta región puede condicionar el correcto desarrollo de las funciones de la mitocondria (Nicholls y Minczuk, 2014). De hecho, variaciones en la región no codificante se han asociado con patologías como el cáncer, la obesidad, la diabetes, enfermedades cardiovasculares y neurodegenerativas. Esta región del ADNmt es en la que se han encontrado mayores niveles de modificaciones epigenéticas, como metilación e hidroximetilación, las cuales modifican los niveles de replicación y transcripción del ADNmt. Este podría ser otro motivo por el que alteraciones en esta región se asocian con diversas patologías humanas (Stoccoro y Coppedè, 2021).

El empaquetamiento del ADNmt no es tan elevado como el del ADN nuclear. El ADNmt carece de histonas, de modo que se encuentra empaquetado en unas estructuras llamadas nucleoides (Fig. 22), junto con las proteínas necesarias para la replicación y la transcripción. Estas estructuras presentan un tamaño aproximado de 100 nm de diámetro, y se localizan en la cara interior de la membrana mitocondrial, muy próximos a la CTE. En la mayoría de los casos, cada nucleoide presenta una única molécula de ADNmt acompañado de las siguientes proteínas (Chocron *et al.*, 2019; Falkenberg 2018):

1. Polimerasa- $\gamma$  (PolG): es la única polimerasa replicativa presente en la mitocondria. Además de esta función también realiza reparaciones en el ADNmt. Está compuesta por una subunidad catalítica POL $\gamma$ A y dos subunidades POL $\gamma$ B. POL $\gamma$ A contiene un dominio 3'  $\rightarrow$  5' exonucleasa correctora de errores, POL $\gamma$ B incrementa la interacción con el ADN molde e incrementa la actividad catalítica y procesividad de POL $\gamma$ A. No puede usar ADN de doble cadena como molde por lo que depende de las helicasas para funcionar.
2. *Twinkle*: helicasa del ADNmt necesaria durante la replicación, proceso durante el cual viaja justo por delante de PolG separando las dos hebras de la molécula de ADNmt.

3. Proteína de unión monocatenaria mitocondrial (mtSSB, del inglés *mitochondrial single stranded binding protein*): necesaria durante el proceso de replicación del ADNmt. Se une a los ADNmt monocatenarios generados por *Twinkle* para protegerlos de nucleasas y evitar la formación de estructuras secundarias. Además, estimula la actividad de *Twinkle* y la procesividad de PolG.
4. Prohibitina (Phb): proteína reguladora de la transcripción mitocondrial.
5. Factor de transcripción mitocondrial A (TFAM del inglés *Mitochondrial transcription factor A*): tiene un papel clave en la transcripción del ADNmt, en la regulación del número de copias ADNmt en células de mamíferos y en el empaquetamiento del ADNmt en nucleoides. En el nucleoide están presentes múltiples proteínas TFAM, en una proporción de 1 por cada 16-17 pb de ADNmt.

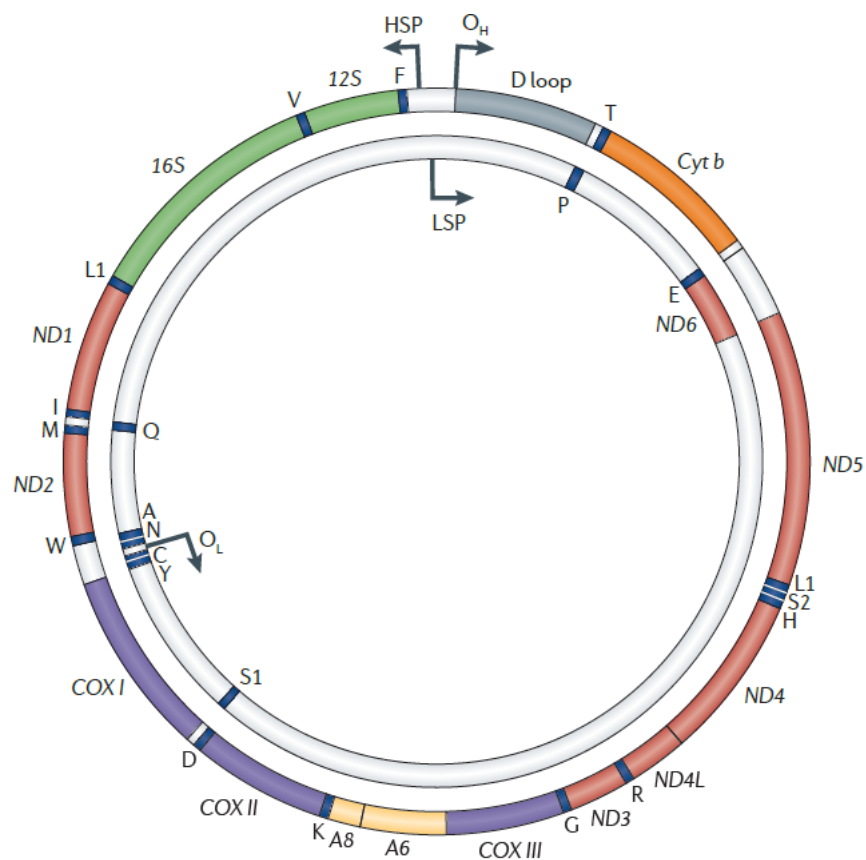


Figura 21. ADN mitocondrial. El ADNmt codifica para 37 genes: 22 ARNt (representados por una sola letra correspondiente a cada aminoácido), 2 ARNr (16S y 12S) y 13 ARNm que se corresponden con subunidades de los componentes de la cadena de transporte electrónico: ND1, ND2, ND3, ND4, ND4L, ND5 y ND6 codifican para componentes del complejo I; *cyt b* codifica para una subunidad del complejo III; COX I, COX II y COX III codifican para subunidades del complejo IV; A6 y A8 codifican para el complejo V. O<sub>H</sub> y O<sub>L</sub> orígenes de replicación de las cadenas pesada y ligera, respectivamente. HSP y LSP promotores de transcripción de la cadena pesada y ligera, respectivamente. Imagen obtenida de Schon et al. 2012 (Schon et al., 2012).

El ADNmt contiene la información necesaria para la síntesis de los 37 elementos para los que codifica y, al igual que sucede con el ADN nuclear, las etapas necesarias para la síntesis de proteínas a partir de este material genético son la replicación, la transcripción y la traducción. Los genes codificantes de polipéptidos mitocondriales que están presentes en la molécula de ADNmt contienen la información necesaria para generar distintas subunidades de los complejos I, III, IV y V de la CTE. El complejo I consta de 40 proteínas, de las cuales 7 son codificadas por el ADNmt, el complejo III está formado por 11 proteínas siendo sólo una, el citocromo b, codificada por el ADNmt, el complejo IV lo forman 14 proteínas donde COI (del inglés *cytochrome oxidase*

## Introducción

*c subunit 1*), COII (del inglés *cytochrome oxidase c subunit 2*) y COIII (del inglés *cytochrome oxidase c subunit 3*) son codificadas por el ADNmt; por último, el complejo V o ATPasa lo conforman 16 proteínas, siendo las subunidades ATP6 y ATP8 codificadas por el ADNmt. El resto de elementos de estos complejos de la CTE son codificados por el ADN nuclear (García *et al.*, 2017).

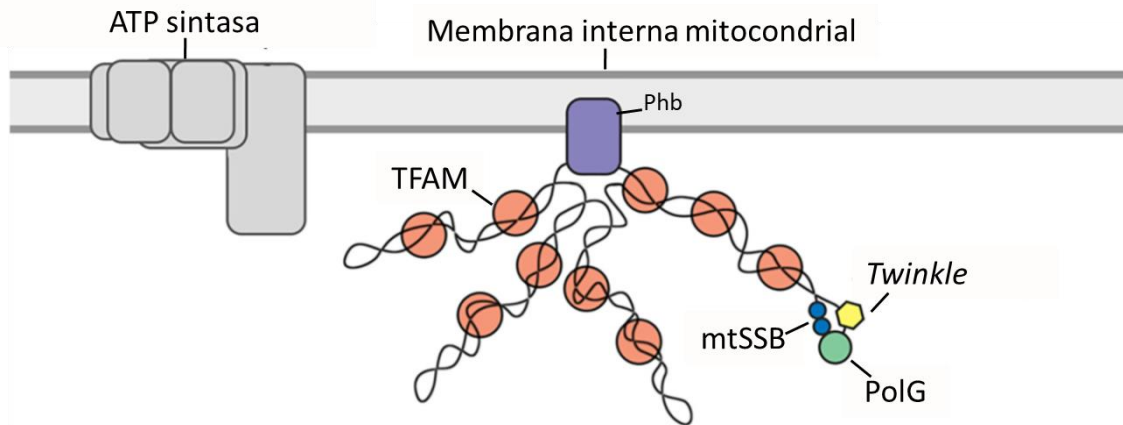


Figura 22. Esquema representativo del sistema de empaquetamiento del ADNmt basado en nucleoides. Imagen modificada de Falkenberg *et al.* 2018.

El ADNmt no dispone de toda la información genética necesaria para que la mitocondria desarrolle sus funciones y se estima que, para llevarlas a cabo, la mitocondria necesita alrededor de 1.000 proteínas codificadas por el núcleo, traducidas en el citosol e importadas directamente a la mitocondria. Estas proteínas nucleares no solo constituyen parte de la CTE, sino que son necesarias en otros procesos como el ciclo de Krebs, la  $\beta$ -oxidación, la síntesis de pirimidinas y otras funciones como la replicación y la transcripción mitocondrial (Pagliarini *et al.*, 2008). La regulación de estos procesos es sumamente importante, ya que determinarán los niveles de metabolitos y proteínas mitocondriales, afectando directamente al metabolismo energético de la mitocondria (D'Souza y Minczuk, 2018).

El modelo de **replicación** de ADNmt más aceptado fue propuesto en 1972, y es el denominado modelo de desplazamiento de hebra (Falkenberg, 2018). El proceso es similar a la replicación del ADN nuclear, sin embargo, en él intervienen proteínas distintas y ocurre de forma continua en las dos hebras de la molécula, comenzando en el  $O_H$  y, cuando aproximadamente se ha replicado 2/3 del genoma mitocondrial, se inicia la síntesis desde el  $O_L$  (Fig 23).

El proceso de **transcripción** ocurre de forma simultánea en los dos promotores presentes en la región no codificante; el promotor LSP (del inglés *Light strand promoter*) transcribe la cadena ligera y el HSP (del inglés *Heavy strand promoter*) la pesada. A diferencia del ADN nuclear, en este proceso se generan transcritos policistrónicos, es decir, que contienen la información de más de un gen. La mayoría de las regiones codificantes están separadas por secuencias de ARNt mitocondriales y se piensa que la escisión endonucleolítica de estos ARNt produce la liberación de los ARNm y ARNr, proceso conocido como modelo de puntuación del ARNt (Ojala *et al.*, 1981). Sin embargo, este modelo no explica como maduran aquellas regiones que no tienen ningún gen de ARNt cercano. Una vez liberados del transcrito, todos los ARNm, a



excepción del de MT-ND6 (del inglés *Mitochondrially Encoded NADH: Ubiquinone Oxidoreductase Core Subunit 6*), sufren un proceso de poliadenilación de su extremo 3'.

La **traducción** mitocondrial ocurre en los mitorribosomas 55S (Greber y Ban, 2016). Al igual que los citoplasmáticos, están formados por dos subunidades, la mayor (39S) y la menor (28S). El mitorribosoma está formado por un total de 82 proteínas, de las cuales 36 son proteínas reclutadas en la mitocondria, 22 son necesarias en la subunidad mayor y 14 en la menor. El proceso de traducción depende de numerosas proteínas reguladoras que son codificadas por el núcleo celular.

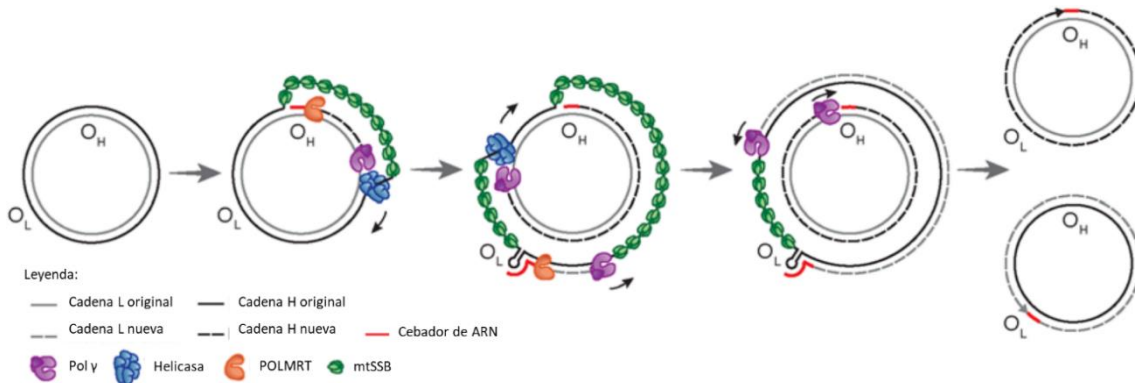


Figura 23. Replicación del ADNmt. Se inicia en O<sub>H</sub> y procede unidireccionalmente para producir la hebra H naciente de longitud completa. Las proteínas mtSSB se unen y evitan que la hebra H parental expuesta hibride. Cuando la horquilla de replicación pasa el O<sub>L</sub>, se forma una estructura de bucle de tallo que bloquea la unión de mtSSB, presentando una región de bucle monocatenario a partir de la cual POLRMT (ARN polimerasa mitocondrial) puede iniciar la síntesis de cebadores. La transición a la síntesis de ADN de cadena L tiene lugar después de aproximadamente 25 nt, cuando la polimerasa y reemplaza a POLRMT en el extremo 3' del cebador. La síntesis de las dos hebras procede de manera continua hasta que se han formado dos moléculas completas de ADN de doble hebra, momento en el cual se produce la separación de las dos moléculas de ADNmt. Imagen modificada de Falkenberg 2018.

### Inestabilidad mitocondrial.

La molécula de ADNmt fue totalmente secuenciada por primera vez en el año 1981 (Anderson *et al.*, 1981). El ADNmt puede sufrir defectos que condicionan las funciones de la mitocondria, lo cual es especialmente relevante en tejidos post-mitóticos con elevada demanda energética. En este sentido, diversos grupos de investigación han puesto de manifiesto que la presencia de algunos polimorfismos específicos puede modificar el riesgo de sufrir enfermedades neurológicas degenerativas como la neuropatía óptica de Lheber (LHON del inglés *Lheber optical neuropathy*), el párkinson, el alzheimer, la esclerosis múltiple o la esclerosis lateral amiotrófica (Chinnery *et al.*, 2018); enfermedades cardiovasculares, como el fallo cardíaco, la hipertrofia cardíaca o la cardiomiopatía diabética (Martín-Fernández y Gredilla, 2016); la pérdida de función y masa muscular asociada a la edad (también denominada sarcopenia) (Bellanti *et al.*, 2021); o también pueden alterar el riesgo de desarrollar distintos tipos de cánceres (Zong *et al.*, 2016). Además, tiene un papel clave en el envejecimiento así como en las enfermedades asociadas a este proceso (Chocron *et al.*, 2019).

El principal mecanismo de reparación del ADNmt es el sistema BER mitocondrial, aunque también están presentes otros mecanismos como los sistemas mitocondriales DSBR y el MMR, e incluso la reparación directa del daño en el ADN (Gredilla *et al.*, 2010). Estos mecanismos de reparación dependen en gran medida de la maquinaria nuclear, ya que muchas de las proteínas que intervienen son codificadas por el núcleo, pero contienen señales de localización mitocondrial que promueven su translocación a la mitocondria (Kazak *et al.*, 2012). Dichos

## Introducción

mecanismos de reparación del ADNmt tienden a ser menos eficientes que los del ADN nuclear. Además, el ADNmt tiene un empaquetamiento más pobre que el ADN nuclear, y una exposición continua a las ROS generadas durante el proceso de transporte electrónico, ya que los nucleoides están muy próximos a los puntos donde se forman estas moléculas.

La tasa de mutación del ADNmt es entre 10 y 20 veces más elevada que la del ADN nuclear. Esto es debido, principalmente, a una serie de factores que incluyen su bajo grado de compactación con respecto al ADN nuclear, haciéndolo más susceptible al daño; su deficiente sistema de reparación comparado con el ADN nuclear, lo que dificulta la eliminación de daños en esta molécula; y en tercer lugar, a su proximidad a la principal fuente de producción de ROS (Gredilla *et al.*, 2010). Cuando aparece una nueva mutación en una copia de ADNmt, esta molécula portadora de la mutación convive con el resto de moléculas “salvajes” no portadoras de esta mutación. Esta condición se conoce como heteroplasmia (Wallace, 2015), y es debido a la presencia de múltiples mitocondrias por célula y a que cada mitocondria porta varias copias de ADNmt. La heteroplasmia puede considerarse como una herramienta defensiva, ya que la proporción de moléculas salvajes con respecto a las mutantes tiende a ser mucho mayor, y es necesario que el porcentaje de moléculas mutantes llegue unos porcentajes de heteroplasmia elevados para que produzcan un efecto negativo, esto es lo que se conoce como efecto umbral o *threshold effect*. El porcentaje de moléculas de ADNmt necesario para producir el efecto negativo suele ser superior al 60%, sin embargo, dependiendo del tipo de mutación y del tejido afectado en el que esta mutación esté presente puede ocurrir que el porcentaje de heteroplasmia necesario para desarrollar una consecuencia fenotípica sea mayor, alcanzando valores del 90% (Rossignol *et al.*, 2003).

Es necesaria una gran acumulación de mutaciones durante la vida del individuo para que ocurra un fallo mitocondrial. La mayoría de ellas ocurren de forma somática en tejidos no germinativos y tienden a acumularse con el paso de los años las cuales, junto con modificaciones epigenéticas como la metilación o la hidroximetilación, así como cambios en el número de copias de ADNmt, pueden acabar promoviendo la disfunción mitocondrial y fallos a nivel de tejidos y órganos (Chocron *et al.*, 2019). Como se comentó en el párrafo anterior, el ADNmt tiene una elevada tasa de mutación debido a varios factores. Uno de los principales causantes de estas mutaciones es la exposición a las ROS mitocondriales, que modifica las guaninas a 8-hidroxiguanosina (8-OHdG); la 8-OHdG puede unirse de forma errónea con las adeninas en vez de con las citosinas, lo que genera una mutación puntual de tipo transversión de GC a TA. Este tipo de mutaciones debidas a la prolongada exposición a ROS generan un círculo vicioso donde la acumulación de mutaciones provoca un aumento significativo de la disfunción mitocondrial lo que, a su vez, aumenta la producción de ROS (Harman, 1956). Finalmente, la elevada producción de ROS acaba generando una importante cantidad de daños a nivel celular que terminan promoviendo la senescencia y la muerte de la célula (Balaban *et al.*, 2005). Este fenómeno es especialmente importante en tejidos con elevadas demandas energéticas, pudiendo ser la causa de algunas enfermedades comentadas al principio de este apartado.

También puede ocurrir que estas mutaciones afecten a las células de la línea germinativa. En este caso, los niveles de heteroplasmia de una mutación mitocondrial germinativa pueden ser drásticamente alterados debido al proceso de segregación replicativa. Este concepto se refiere a que, durante la replicación celular, las mitocondrias son aleatoriamente segregadas entre células hijas y, en el caso de que sean las células germinativas femeninas las que tienen el defecto heredado, pueden ocurrir tres situaciones. La primera es que el oocito resultante tenga una secuencia totalmente normal, también denominada como situación de homoplasmia salvaje; el segundo caso es que tenga una secuencia totalmente anormal, situación de homoplasmia mutante; o finalmente, que tenga una proporción de moléculas mutantes aleatoria, situación de heteroplasmia (Fig. 24) (Christodoulou y Jansen, 2000).

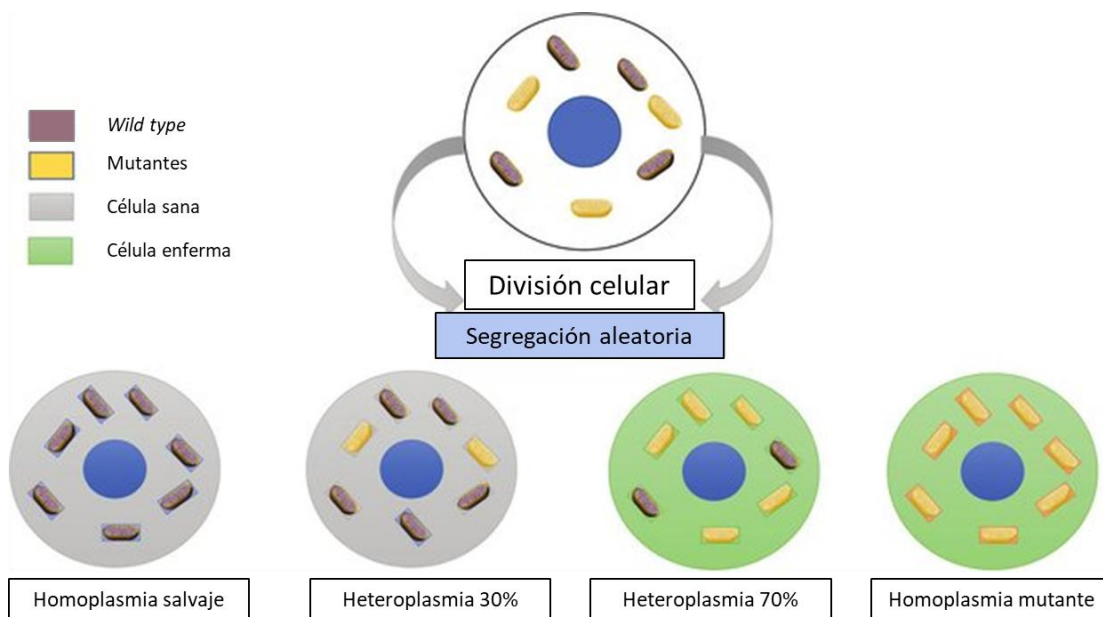


Figura 24. Esquema representativo del fenómeno de segregación replicativa. A partir de una célula precursora de gametos y debido al proceso de gametogénesis, se produce una segregación aleatoria de su contenido celular. De esta forma, si la célula precursora poseía un fenómeno de heteroplasmia en sus mitocondrias, este puede ser transmitido a la descendencia de formas muy diversas, pudiéndose generar gametos sin mitocondrias mutantes (homoplasmia salvaje), gametos con solo mitocondrias mutantes (homoplasmia mutante), y por otro lado también se pueden generar gametos con un porcentaje de heteroplasmia muy variable, pudiendo sobrepasar inmediatamente el efecto umbral dando lugar a consecuencias fenotípicas. Imagen modificada de <https://www.fernandogalangalan.com/blog?id=jfkgmoarstlx&start=140>.

Dependiendo del resultado de la segregación replicativa, el umbral puede ser sobrepasado de forma inmediata en la descendencia y, según el efecto que pueda tener esta mutación germinativa, la podremos clasificar (Wallace *et al.*, 1999) como:

1. Mutación neutra: parecen tener poca importancia para la supervivencia del individuo y no suelen afectar a la longevidad.
2. Mutación ligeramente deletérea: no afectan en gran medida a la supervivencia del individuo, sin embargo, es posible que alguna variante pueda causar reducciones en la capacidad de la CTE, lo que, unido a la acumulación continua de mutaciones somáticas, pueden a largo plazo dar lugar a un fallo orgánico en edades avanzadas. Por ejemplo, la variante m.1555A>G, que afecta al ARNr 12S se ha asociado en linajes árabes con el desarrollo de sordera.
3. Mutación moderadamente deletérea: según aumenta la severidad de la mutación, la frecuencia y severidad de sus manifestaciones clínicas aumenta. Estas mutaciones

## Introducción

pueden afectar en gran medida a la supervivencia del individuo y, generalmente, con la muerte del individuo estas mutaciones desaparecen de la población. Si las moléculas mutantes son muy predominantes, pueden aumentar en gran medida la incidencia y la progresión de alguna enfermedad. Un ejemplo de este tipo de variante es la m.11778G>A, la cual es la causa más común de LHON; esta mutación afecta al gen que codifica para la subunidad 4 del complejo I de la CTE, comprometiendo la generación de energía por parte de la mitocondria.

4. Mutación severamente deletérea: los individuos con mutaciones moderadamente deletéreas suelen permanecer razonablemente sanos durante la niñez. Sin embargo, las mutaciones más severas suelen dar lugar a enfermedades infantiles e incluso letales. Estas mutaciones reducen enormemente la supervivencia del individuo. Un ejemplo de esto es la variante m.8993T>G. Este polimorfismo causa un amplio espectro de presentaciones clínicas que van desde la retinosis pigmentaria periférica leve, pasando por la atrofia olivopontocerebelosa y la degeneración macular, hasta el síndrome de Leigh infantil letal acompañado de lesiones en los ganglios basales. Esto se debe a que esta mutación bloquea el canal de protones de la ATP sintasa, inhibiendo así la generación de ATP a partir de la fosforilación oxidativa.

La zona del ADNmt con mayor tasa de mutación es la región de 1,1 kb no codificante, o región control. Como se comentó previamente, esta zona contiene los elementos necesarios para la transcripción y la replicación (L. Wang *et al.*, 2021). Debido a esto, diferentes autores proponen que la presencia de polimorfismos en esta región puede afectar a los niveles de transcripción mitocondriales, también que algunos polimorfismos pueden producir sitios CpGs susceptibles de sufrir procesos de metilación condicionando así el correcto funcionamiento de la mitocondria, o incluso afectar al número de copias de ADNmt (Navaglia *et al.*, 2006; Akouchejian *et al.*, 2009). Diversos estudios han puesto de manifiesto que, conforme aumenta la edad del individuo, en esta región se acumulan elevados niveles de mutaciones puntuales en el ADNmt de tejidos como el cerebro, la sangre, la piel, el músculo y el corazón (Michikawa *et al.*, 1999; Wang *et al.*, 2001). La presencia y acumulación de mutaciones somáticas en esta zona también puede incrementar el riesgo de sufrir diversas patologías como enfermedades crónicas del riñón, endometriosis, cáncer colorectal, cáncer de mama, migrañas o desórdenes gastrointestinales, entre otros (Covarrubias *et al.*, 2008; Camilleri *et al.*, 2009; Zaki *et al.*, 2009; Bai *et al.*, 2014; Guo *et al.*, 2016; Li *et al.*, 2020). Un ejemplo de variante mitocondrial en la región control, que ha sido muy estudiada, es la de la posición 189, (m.189A>G). Entre los efectos propuestos, destaca el que puede afectar al mantenimiento y a la replicación del ADNmt debido a su proximidad al origen de replicación (Nicholls y Minczuk, 2014). Otro ejemplo conocido es la variante m.150C>T pero, en este caso, en vez de promover un efecto deletéreo, la presencia del alelo T se asoció con un incremento de la longevidad en poblaciones irlandesas, turcas, finlandesas y japonesas; sin embargo, este efecto no se replicó en algunas poblaciones chinas (Chocron *et al.*, 2019). Estudios funcionales mostraron que la presencia de esta mutación reduce los niveles de ROS, aunque los autores insisten en tomar con cautela los datos obtenidos ya que no se descarta que la presencia de otras mutaciones adicionales pudiera estar detrás de este fenotipo asociado a longevidad.

Además de las mutaciones de un solo nucleótido, las grandes deleciones que pueden aparecer en la molécula de ADNmt también juegan un papel importante en la disfunción mitocondrial. Estas mutaciones pueden eliminar zonas del genoma mitocondrial de entre 1,8 y 8 kb (Nissanka *et al.*, 2019) y, por tanto, pueden eliminar genes vitales para el funcionamiento de la mitocondria. Este tipo de mutaciones se han asociado con numerosas enfermedades,

siendo una de las más comunes la delección de 4.977 pb entre las posiciones m.8482 y m.13640. Esta alteración implica la eliminación de 4 genes que codifican para subunidades del complejo I, un gen del complejo IV, dos genes del complejo V, y cinco genes de ARNts. Asimismo, esta delección se ha asociado con diferentes patologías humanas como el Síndrome de Kearns-Sayre espontáneo (Shoffner *et al.*, 1989), la presbiacusia (Han *et al.*, 2019), el desarrollo de neoplasmas benignos uterinos (Luo *et al.*, 2019) o el Alzheimer (Phillips *et al.*, 2014) e incluso con el desarrollo de la artrosis (Chang *et al.*, 2005), entre otras. Además de esta gran delección, existen otras que pueden modular el riesgo a desarrollar ciertas enfermedades; por ejemplo, delecciones de aproximadamente 3,1 a 6 kb se han asociado con el síndrome de Pearson (Lee *et al.*, 2007), u otra de casi 7 kb se ha asociado con el síndrome de Fanconi (Zhu *et al.*, 2019). Este tipo de mutaciones suelen generarse por errores durante el proceso de replicación del ADNmt (Nissanka *et al.*, 2019), y su frecuencia tiende a aumentar con la edad, sobre todo en tejidos como el músculo esquelético (Herbst *et al.*, 2018).

### *Origen de los haplogrupos mitocondriales.*

La especie humana moderna, *Homo sapiens sapiens*, es relativamente joven, originándose en el continente africano hace aproximadamente entre 150.000 y 200.000 años (Stringer, 2016). Fue en ese momento cuando surgió lo que se conoce como la “Eva mitocondrial”, la primera mujer con las moléculas de ADNmt a partir de las cuales descienden todas las moléculas de ADNmt de todos los individuos del planeta. Como se habló en el apartado anterior, cuando una mutación tiene un efecto neutral o incluso una ventaja adaptativa, ésta se fija en la población como un polimorfismo. Este tipo de polimorfismos en el ADNmt se han acumulado secuencialmente a lo largo del tiempo, heredándose a través de la vía materna, aunque excepcionalmente puede ocurrir un fenómeno de herencia biparental del ADNmt (Luo *et al.*, 2018). Se hipotetiza que durante el proceso de migración desde África hacia el resto del mundo, las poblaciones comenzaron a dispersarse y a aislarse y, durante este proceso, determinados polimorfismos favorecieron la adaptación geoclimática de poblaciones específicas, fijándose en éstas. Como resultado de este proceso, los linajes de ADNmt fueron divergiendo debido a la acumulación de distintos polimorfismos que permitían a poblaciones concretas adaptarse mejor al entorno que les rodeaba. Debido a este fenómeno se generaron combinaciones de polimorfismos específicos en poblaciones humanas concretas que en su conjunto se denominan haplogrupos mitocondriales (Wallace, 2015).

La hipótesis más robusta a día de hoy sobre el origen los haplogrupos mitocondriales es precisamente aquella que propone que éstos han sido el producto de un proceso de adaptación climática. Esto es debido a que las variantes genéticas que componen los haplogrupos pueden afectar a dos procesos fundamentales relacionados con el funcionamiento de la ETC, que son la producción de energía en forma de ATP y la generación de calor. Así, por un lado, pueden potenciar la eficiencia de la fosforilación oxidativa, cuya función principal es la producción de energía en forma de ATP o, por otro lado, las variantes de los haplogrupos pueden favorecer el proceso de desacoplamiento mitocondrial, por el que se reducirá la producción de energía en detrimento de una mayor producción de calor. El proceso de desacoplamiento mitocondrial puede ocurrir mediante tres mecanismos (Demine *et al.*, 2019):

1. Los protones pueden directamente escapar a través de la membrana interna mitocondrial por difusión, proceso denominado como filtración de protones (*proton leak* en inglés). Este proceso depende de la composición de la membrana interna mitocondrial.

## Introducción

2. Los desacopladores son unas proteínas hidrofóbicas denominadas UCPs (del inglés *uncoupling proteins*), las cuales poseen un protón disociable y son capaces de traspasar la membrana mitocondrial interna transportando otros protones. En presencia de estas moléculas, el transporte de electrones se realiza de forma normal pero no se genera ATP, sino que estas moléculas disipan la fuerza protonmotriz generándose calor. Una de las más estudiadas es UCP-1 o termogenina, la cual está presente en mitocondrias del tejido adiposo y se activa en presencia de ácidos grasos generados a partir de triacilglicéridos.
3. Los electrones pueden filtrarse desde la ETC y unirse directamente al oxígeno, generando así  $O_2^-$  o hidroperoxilo ( $HO_2^-$ ).

Si los polimorfismos de un haplogrupo mitocondrial favorecen que la fosforilación oxidativa sea muy eficiente, se generará mucha más energía que calor, permitiendo una mejor aclimatación a ambientes tropicales. Por el contrario, si se favorecen los procesos de desacoplamiento, se generará más calor, favoreciendo la aclimatación a ambientes más fríos. Esta hipótesis descrita por Ruiz-Pesini y colaboradores en 2004, y publicada en la revista *Science* (Ruiz-Pesini *et al.*, 2004), serviría para explicar que, debido a la influencia del clima más frío, se habría producido un proceso de adaptación por el cual una selección de variantes mitocondriales beneficiosas se irían fijando en las poblaciones, dando lugar a los haplogrupos mitocondriales. Debido al impacto que tuvo este trabajo, algunos grupos de investigación trataron de validar esta hipótesis y, trabajos como el de Amo y Brand en 2007 (Amo y Brand, 2007) no fueron capaces de hacerlo, con lo que la ponen en duda. En este trabajo los autores diseñaron cíbridos transmitocondriales, un modelo celular donde se generan líneas de células con el mismo ADN nuclear pero distinto ADNmt, para estudiar el efecto que el genoma mitocondrial tiene sobre la célula. Los autores generaron líneas celulares portadoras de haplogrupos árticos y tropicales y compararon la eficiencia de fosforilación oxidativa que presentaban. Como resultado, no observaron diferencias en los niveles de desacoplamiento energético entre ambos grupos de cíbridos. Sin embargo, una de las grandes autocríticas de este estudio es que la línea que se usó para crear los cíbridos era portadora del haplogrupo H, que se hipotetiza que está relativamente adaptado al frío (Mishmar *et al.*, 2003; Ruiz-Pesini *et al.*, 2004), y quizá el haplogrupo tropical no se complementa con el nuclear de forma correcta, pudiendo favorecer el bajo desacoplamiento, enmascarando la eficiencia real.

Teniendo en cuenta esta teoría de adaptación climática se han desarrollado distintos mapas de los movimientos migratorios humanos junto con la aparición de diversos haplogrupos mitocondriales. El haplogrupo original de la Eva mitocondrial se denomina haplogrupo L, que con el paso del tiempo se subdividió en 7 grandes familias, L0, L1, L2, L3, L4, L5 y L6, siendo el más abundante el haplogrupo L2. Hace aproximadamente 85.000 años, y debido a que durante ese tiempo la especie humana estaba bajo un período interglacial conocido como *Riss-Wurm*, ésta comenzó a expandirse hacia otros continentes. Debido a esta migración, la familia L3 dio lugar a los haplogrupos M y N. Hace 80.000 años comenzó un período glacial conocido como la glaciación de *Würm*, que produjo nuevos eventos migratorios, promoviendo que el haplogrupo M emigrara y colonizara Asia hace alrededor de 60.000 años. Con el paso del tiempo, a partir de este haplogrupo se originaron los haplogrupos A, B, C, D, G y F. Además, algunos individuos del continente asiático también migraron hacia América en dos eventos, el primero ocurrió hace 26.000-34.000 años, que llevó los haplogrupos A, C y D; y el segundo, ocurrió hace 12.000-15.000 años, que llevó el haplogrupo B. De forma paralela a lo que ocurrió con el haplogrupo M, el haplogrupo N emigró hacia Europa y, durante este proceso se originó el haplogrupo R, el cual,

hace 45.000 años, se ramificó en los 9 haplogrupos mitocondriales europeos actuales: H, I, J, K, T, U, V, W y X (Mancuso *et al.*, 2011; Wallace, 2015; Chinnery y Gomez-Duran, 2018)(Fig. 25).

A día de hoy este proceso continúa siendo estudiado, ya que las metodologías de secuenciación presentes permiten identificar nuevos haplogrupos minoritarios y sub-haplogrupos dentro de los principales haplogrupos continentales, favoreciendo así la precisión en el análisis de los procesos migratorios humanos. Un buen ejemplo de esto son los haplogrupos caucásicos, ya que, inicialmente solo se conocían cuatro haplogrupos, H, I, J y K (Wallace, 2015), pero al año siguiente Torroni y colaboradores fueron capaces de identificar 5 nuevos haplogrupos mayoritarios caucásicos T, U, V, W y X, aumentando así la variabilidad de los haplogrupos europeos.

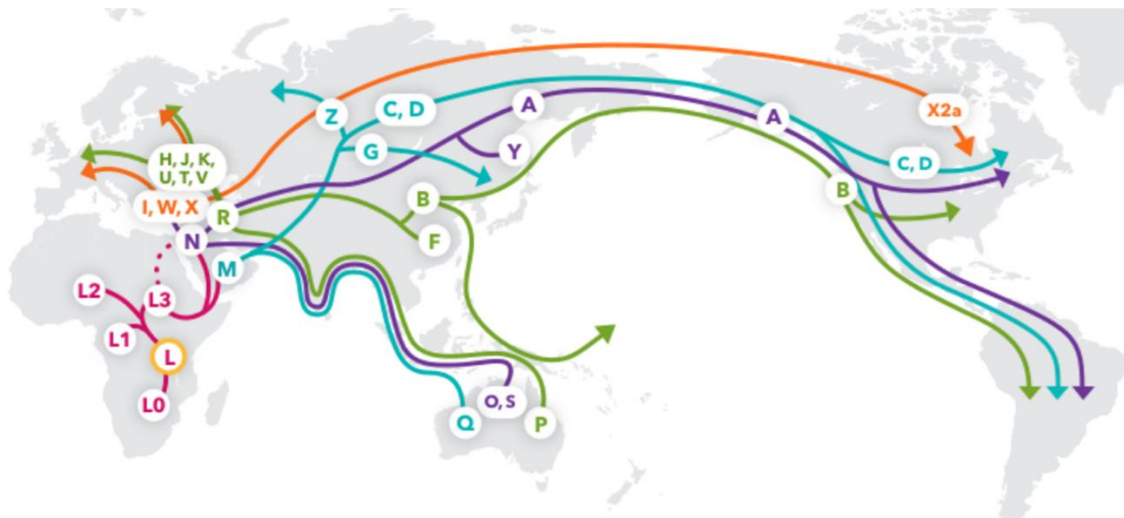


Figura 25. Mapa migratorio de la especie humana desde la Eva mitocondrial, con el haplogrupo mitocondrial L, originado hace aproximadamente 200.000 años. Imagen obtenida de <https://blog.23andme.com>.

*Polimorfismos mitocondriales y enfermedades humanas.*

La presencia de determinados polimorfismos mitocondriales de forma individual puede modificar el riesgo de desarrollar diversas patologías, debido a que alteran en gran medida la función mitocondrial. Por ejemplo, en 2021 Na Cai y colaboradores describieron como la presencia de diversos polimorfismos mitocondriales puede afectar a los niveles de N-formilmetionina, la cual es fundamental en la regulación de la proteoestasis mitocondrial y citosólica, pudiendo llegar a alterar el riesgo de desarrollar multitud de patologías (Cai *et al.*, 2021). Pero los polimorfismos mitocondriales pueden actuar también de forma conjunta como en el caso de los haplogrupos, ya que no sólo tienen la capacidad de modificar la función mitocondrial, sino que tienen una repercusión importante en el comportamiento del núcleo celular, tal y como se describirá más adelante. De hecho, muchos trabajos de investigación han puesto de manifiesto que la presencia de determinados haplogrupos mitocondriales se ha asociado con multitud de patologías y fenotipos no clínicos. Entre estas asociaciones se pueden destacar las siguientes:

1. Enfermedad del Alzheimer: en 2009 se llevó a cabo un estudio para evaluar la influencia de los haplogrupos caucásicos en el riesgo de desarrollar esta patología. Como resultado, encontraron que los haplogrupos con origen filogenético común dentro del clúster de HV se asociaron de forma significativa con el riesgo de desarrollar esta patología en una cohorte de individuos polacos (Maruszak *et al.*, 2009). Más adelante, otro grupo de investigadores analizaron si alguno de los sub-

haplogrupos del H se asociaba de forma específica con esta patología, y encontraron una asociación entre el haplogrupo H5 y el riesgo de Alzheimer en poblaciones italianas (Santoro *et al.*, 2010). Una posible explicación radicaría en que éstos haplogrupos pertenecientes al haplogrupo H poseen una fosforilación oxidativa con un alto grado de acoplamiento, esto implicaría una elevada eficiencia energética a costa de generar mayores niveles de ROS. Cuando los niveles de ROS superan a la capacidad de la célula para combatirlos se produce el denominado estrés oxidativo, el cuál comienza a producir daños en diversos elementos celulares y en los tejidos. El estrés oxidativo es una de las causas por las cuales se desarrollan diversas enfermedades neurodegenerativas, siendo una de éstas el Alzheimer.

En el estudio de Bi y colaboradores (Bi *et al.*, 2015) se realizó un estudio similar a los mencionados en el párrafo anterior, pero en una población asiática. En este caso fue el sub-haplogrupo B5 el que se asoció de forma significativa con el riesgo de desarrollar esta patología neurodegenerativa. Para conocer las causas de esta asociación, los autores realizaron estudios funcionales para analizar el efecto de la una de las variantes que definen el sub-haplogrupo H5, la m.8584G>A. Este polimorfismo afecta directamente al gen MT-ATP6 (del inglés *mitochondrially encoded ATP synthase membrane subunit 6*), que codifica para una subunidad de la ATPasa mitocondrial. Como resultado observaron que la presencia de esta variante reduce la función mitocondrial promoviendo mayor generación de ROS mitocondriales, una reducción de la producción de ATP, una pérdida de masa mitocondrial y menor tasa de consumo de oxígeno.

2. Enfermedad de Parkinson: en 2003 se realizó un estudio para analizar el efecto de los haplogrupos caucásicos en el desarrollo de la enfermedad de Parkinson (Van Der Walt *et al.*, 2003). En los resultados con respecto al haplogrupo H, el haplogrupo J mostro un efecto protector frente al Parkinson. Los autores describen que este posible efecto protector puede ser debido a que el haplogrupo J, al igual que el K, posee la variante mitocondrial m.10398G>A, la cual afecta a un gen que codifica para una subunidad del complejo I de la CTE. Esta variante podría alterar la efectividad de este complejo I, reduciendo así la generación de ROS en individuos portadores de estos haplogrupos. Otros estudios también apoyan estos resultados y, del mismo modo que ocurría con el Alzheimer, el cluster mitocondrial HV se asoció de nuevo con el riesgo de desarrollar la enfermedad de Parkinson debido al aumento del estrés oxidativo (Hudson, Nalls, *et al.*, 2013). En este mismo estudio también se observó que el clúster de haplogrupos JT se asociaba de forma significativa con un efecto protector en el desarrollo de la enfermedad del Parkinson. Más concretamente, la variante m.2158T>C, la cual define específicamente el sub-haplogrupo J1b se asoció con este efecto. La explicación radica en que este polimorfismo afecta al gen que codifica para la humanina, un gen que tiene un efecto protector para enfermedades neurodegenerativas, de modo que esta mutación podría alterar los niveles de expresión de esta proteína reduciendo el riesgo de desarrollar este tipo de patologías.
3. Diabetes mellitus tipo 2: la variante mitocondrial m.16189T>C, presente en la región no codificante del ADNmt, tiene una alta representación en las poblaciones asiáticas y, de hecho, se asoció en poblaciones asiáticas con un riesgo aumentado de desarrollar diabetes mellitus tipo 2 (Liao *et al.*, 2008). Una posible explicación para esta asociación fue propuesta en el año 1998 (Poulton *et al.*, 1998), donde los autores describen que la presencia del alelo C reduce la afinidad de unión de la



proteína mtSSB. Esta proteína juega un papel determinante en el empaquetamiento del nucleoide y en el mantenimiento del ADNmt. Por tanto, la presencia de esta variante podría afectar a la estabilidad de esta molécula. Esta variante no tiene tanta representación en haplogrupos caucásicos, estando presente solamente en sub-haplogrupos del cluster HV. Debido a esta baja presencia de m.16189C, no se ha podido confirmar su asociación en dichas poblaciones. (Kwak y Park, 2016). En poblaciones marroquíes tampoco se ha asociado esta variante con el riesgo de esta patología (Charoute *et al.*, 2018).

A nivel de haplogrupos, en la revisión bibliográfica de Nishigaki y colaboradores (Nishigaki *et al.*, 2010), los autores analizan la asociación entre los haplogrupos y diferentes enfermedades en una población japonesa y coreana. En dicha revisión diversos estudios apuntan a que los haplogrupos F y A se asocian con un riesgo aumentado de desarrollar diabetes mellitus tipo 2 en poblaciones japonesas, mientras que, en poblaciones coreanas, tanto el haplogrupo D5 como el sub-haplogrupo D4b se asociaron con dicho riesgo. Por otro lado, en ambas poblaciones el haplogrupo N9a se asocia con un efecto protector frente a esta patología. Para explicar el posible efecto protector de N9a, piensan que se puede deber, por un lado, a la baja eficiencia de la fosforilación oxidativa, ya que al generar menos energía sería necesario aumentar la respiración mitocondrial con el fin de compensar este déficit; otra posibilidad podría ser el hecho de que los individuos con N9a generan menores niveles de estrés oxidativo, protegiendo frente a la diabetes mellitus tipo 2.

4. Cáncer: existen dos grandes tipos de mutaciones del ADNmt que pueden promover el cáncer, o bien variantes severas que pueden contribuir a la inducción del cáncer, o variantes menos deletéreas que son importantes en la adaptación de las células cancerígenas (Kopinski *et al.*, 2021). Algunos ejemplos de mutaciones severas asociadas con este tipo de patologías están presentes en el cáncer de próstata. Estudios funcionales han puesto de manifiesto que las variantes m.8993T>G (Petros *et al.*, 2005) o m.6124T>C (Arnold *et al.*, 2013) promueven que la célula genere mayores niveles de ROS mitocondriales. Las ROS, además de producir daño, pueden promover cambios epigenéticos alterando la expresión de diversos genes, como por ejemplo HIF1 $\alpha$  (del inglés, *Hypoxia-inducible factor 1-alpha*), el cual podría ser un mecanismo que favorece la replicación de las células cancerígenas. En contraste con este tipo de mutaciones severas, los haplogrupos tienen efectos menos notorios, pero que pueden llegar a favorecer el desarrollo de este tipo de patologías. En la revisión de Kopinski y colaboradores (Kopinski *et al.*, 2021) se muestran diferentes asociaciones de haplogrupos de distintos continentes con el riesgo de desarrollar múltiples tipos de cáncer. En el cáncer de mama, en poblaciones asiáticas, se ha asociado el haplogrupo D5 con un aumento del riesgo (Fang *et al.*, 2010), mientras en caucásicos, poseer el sub-haplogrupo T1a1 se asoció con un menor riesgo de este tipo de cáncer (Blein *et al.*, 2015). El haplogrupo U se ha asociado con cáncer renal y también con cáncer de próstata (Booker *et al.*, 2006). El haplogrupo asiático D4a se asoció con mayor riesgo de cáncer de tiroides (Fang *et al.*, 2010). En el cáncer colorrectal, la presencia del haplogrupo caucásico T se ha asociado con un aumento del riesgo de sufrir dicha patología (Li *et al.*, 2015).

Como se ha mencionado previamente, estas mutaciones mitocondriales pueden promover el desarrollo de este tipo de patologías a través de diversos mecanismos: pueden alterar la producción de ROS, modificando el estado redox de la célula y

favoreciendo así los procesos de crecimiento de las células cancerosas; asimismo, también pueden inducir cambios en genes mitocondriales que pueden alterar los metabolitos producidos por este orgánulo. A su vez, estos metabolitos pueden alterar la epigenética del ADN nuclear alterando los niveles de expresión de genes importantes en el desarrollo de esta patología (Wallace *et al.*, 2010).

5. Esclerosis múltiple: además de la influencia de este orgánulo en el desarrollo del Parkinson y el Alzheimer, la disfunción mitocondrial también puede modificar el riesgo de progresión en la esclerosis múltiple. Esta enfermedad se caracteriza porque el sistema inmune ataca a la vaina protectora de mielina presente en los axones de las neuronas, pudiendo producir a largo plazo lesiones en los nervios y generando finalmente fallos de comunicación entre el cerebro y el resto del cuerpo. La disfunción mitocondrial da lugar a una serie alteraciones funcionales, como por ejemplo la reducción de producción de energía en forma de ATP, el aumento de la generación de ROS y la desregulación de los niveles de  $Ca^{2+}$  citoplasmáticos, entre otros. Estos cambios pueden promover la degeneración de los axones y de las vainas de mielina, favoreciendo así el avance de esta patología (Mahad *et al.*, 2008).

Con respecto a la incidencia de la esclerosis múltiple, en 2015 un estudio trató de analizar si la presencia de variantes mitocondriales alteraba el riesgo de desarrollar esta patología (Tranah *et al.*, 2015). En dicho estudio observaron que los individuos portadores del haplogrupos J o T (cluster mitocondrial JT) exhibían un mayor riesgo de desarrollarla.

6. Enfermedades cardiovasculares: en el estudio de Nishigaki y colaboradores (Nishigaki *et al.*, 2010), además de analizar la diabetes mellitus de tipo 2, también analizaron como los haplogrupos modifican el riesgo de infarto de miocardio en hombres. Observaron que el haplogrupo N9b tiene un efecto protector frente a esta patología, mientras que el G1 aumentaba el riesgo. Otras conclusiones que sacaron de este estudio fueron, por un lado, que el haplogrupo A puede modificar el riesgo de infarto cerebral aterotrombótico en mujeres japonesas y, por otro lado, que tanto el haplogrupo N9a como el G1 y el D5, tienen un efecto protector en mujeres japonesas frente al síndrome metabólico.

Otra patología destacable dentro de las enfermedades cardiovasculares es la enfermedad de la arteria coronaria. Esta patología se caracteriza por un endurecimiento y estrechamiento de las arterias coronarias, debido a la acumulación de materiales en la capa interna de las paredes de estos vasos, proceso conocido como arterioesclerosis, siendo uno de los más destacados el colesterol. Como consecuencia, el flujo de sangre hacia el corazón se ve alterado, pudiendo derivar en complicaciones como por ejemplo la angina de pecho o el infarto de miocardio. El exceso en la producción de ROS producido por la disfunción de las mitocondrias podría ser uno de los detonantes del proceso de arterioesclerosis, ya que comprometería la función endotelial, la proliferación de las células de los vasos sanguíneos y alteraría los niveles de apoptosis de estos tejidos, por tanto afectando estructuralmente a las arterias (Madamanchi y Runge, 2007).

En esta línea, Kofler y colaboradores (Kofler *et al.*, 2009) trataron de analizar si existía alguna influencia de los haplogrupos mitocondriales caucásicos sobre esta patología. Como resultado encontraron una asociación entre el haplogrupo T y el riesgo de desarrollar la enfermedad de las arterias coronarias. Una de las posibles hipótesis que barajan los autores es que los portadores del haplogrupo T podrían ser más vulnerables al daño producido por el estrés oxidativo, lo cual favorecería los

procesos de daño en los tejidos arteriales, aumentando así el riesgo de desarrollar la patología descrita.

Otros estudios destacables se centraron en otra patología cardiovascular, la cardiomiopatía isquémica. En 2012 se encontraron asociaciones entre esta patología y los haplogrupos caucásicos H y J, los cuales mostraron un aumento del riesgo y un efecto protector respectivamente (Fernández-Caggiano *et al.*, 2012). En 2013, estos mismos autores analizaron si los haplogrupos mitocondriales caucásicos condicionaban el riesgo de desarrollar miocardiopatía dilatada idiopática y, observaron que, una vez más, el haplogrupo H se asociaba con el riesgo de desarrollar dicha patología (Fernández-Caggiano *et al.*, 2013). Por otra parte, en un estudio de 2006, los autores mostraron una nueva asociación entre el haplogrupo T, definido por la variante m.13368G>A, y el riesgo de desarrollar cardiomiopatía hipertrófica (Castro *et al.*, 2006).

7. Enfermedades mitocondriales: éstas constituyen un conjunto de patologías que comparten una característica, la presencia de mitocondrias disfuncionales. Son enfermedades muy heterogéneas y pueden ocurrir en diversos estados de desarrollo y afectar a múltiples órganos y tejidos, especialmente a aquellos con alta dependencia del metabolismo aerobio. Por lo general, normalmente son producidas por mutaciones patogénicas en el ADNmt que afectan a subunidades estructurales de los complejos de la CTE, aunque también pueden producir las mutaciones en genes que afectan a la síntesis de proteínas o incluso genes nucleares necesarios para el correcto funcionamiento de la mitocondria (Gorman *et al.*, 2016).

Algunas enfermedades mitocondriales destacadas son por ejemplo el LHON, que se caracteriza por una pérdida de visión repentina. Es causada por mutaciones puntuales en las posiciones m.11778G>A, m.14484T>C y m.3460G>A, las cuales afectan todas al complejo I de la CTE (Farrar *et al.*, 2013). Otra patología mitocondrial conocida es el MELAS (del inglés *Mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes síndrome*), caracterizada por la aparición repentina de episodios similares a un accidente cerebrovascular, también puede incluir sordera, diabetes mellitus, retinopatía pigmentada, cardiomiopatía, ataxia cerebelosa, convulsiones, encefalopatía, acidosis láctica y miopatía mitocondrial. Esta patología es debida a mutaciones que afectan al gen MT-TL1 (del inglés *mitochondrially encoded tRNA-Leu*), siendo una de las más conocidas la m.3243A>G (Pavlakis *et al.*, 1984). También destaca el NARP (del inglés *Neurogenic muscle weakness, ataxia and retinitis pigmentosa*), caracterizada por una neuropatía sensitivomotora, ataxia y retinopatía pigmentaria junto con otras complicaciones. Es producida por dos mutaciones distintas en el gen MT-ATP6, ambas en la misma posición m.8993T>G y m.8993T>C (siendo esta última menos severa) (Ortiz *et al.*, 1993). Finalmente se podría destacar también el síndrome de Leigh, aunque existen otras muchas patologías mitocondriales. Dicho síndrome aparece durante la niñez y es causado por mutaciones en más de 75 genes nucleares y mitocondriales. Se caracteriza por períodos agudos de regresión del neurodesarrollo seguidos de parcial recuperación, hipotonía, distonía, hipopnea, disfagia, epilepsia, retraso del crecimiento, encefalopatía y lesiones de los ganglios basales y del tronco encefálico (Lake *et al.*, 2016).

8. Fenotipos no clínicos: Alexe y colaboradores (Alexe *et al.*, 2007) mostraron que en una población japonesa los haplogrupos D4a, D5 y D4b2b estaban sobrerrepresentados en individuos centenarios o semicentenarios. El haplogrupo

D4a también contribuye a la longevidad en otras poblaciones asiáticas; así, en el artículo de Cai y colaboradores (Cai *et al.*, 2009) los individuos centenarios de una cohorte china mostraban una mayor frecuencia del haplogrupo D4a, mientras que los haplogrupos M9, N9 y B4a eran menos abundantes en individuos de elevada edad. En Europa también se han realizado diversos estudios para analizar la posible asociación de los haplogrupos caucásicos con la longevidad, y se ha observado que la frecuencia del haplogrupo J aumenta conforme mayor sea la edad de la población en italianos (De Benedictis *et al.*, 1999), irlandeses (Ross *et al.*, 2001) y finlandeses (Niemi *et al.*, 2003). La causa de esta asociación podría residir, tal y como explicaba la hipótesis de la adaptación climática de los haplogrupos caucásicos (Ruiz-Pesini *et al.*, 2004), en que éste haplogrupo tiene una menor eficiencia en la fosforilación oxidativa, generando más calor y menos energía, y por tanto menos ROS, lo cual favorece la longevidad debido a la menor acumulación de daño debido a estrés oxidativo (Dato *et al.*, 2004). Sin embargo, otros estudios han rebatido estos datos en otras poblaciones; concretamente, en el estudio de Pinós y colaboradores (Pinós *et al.*, 2012) no observaron un aumento significativo de la frecuencia del haplogrupo J en una población española, y esto mismo ocurrió en una población del sur de Italia (Dato *et al.*, 2004). Debido a estas discrepancias entre estudios, resulta complicado concluir que la longevidad está afectada de forma inequívoca por variaciones en el ADNmt. Lo más probable es que el incremento de la longevidad sea debido no solo al ADNmt, sino a la interacción entre este elemento junto con las posibles alteraciones presentes en el ADN nuclear, específicas de las poblaciones de los individuos, y los eventos estocásticos a los que estén sometidos los portadores de estas moléculas (Rose *et al.*, 2001). En el caso específico del haplogrupo caucásico J, se piensa que éste no tiene un efecto directo sobre la longevidad, sino que más bien aumenta la penetrancia de mutaciones que son beneficiosas para el envejecimiento (De Benedictis *et al.*, 2000). En la población caucásica, el haplogrupo J además de aumentar la penetrancia de variantes que se asocian con la longevidad, la presencia de este haplogrupo también aumenta la penetrancia de mutaciones con efecto negativo como la m.11778G>A y m.14484T>C, las cuales, en conjunción con las variantes propias del haplogrupo J, se han asociado con un riesgo elevado de desarrollar LHON (Torroni *et al.*, 1997; Brown *et al.*, 2002) *al.*, 2002).

Con todo lo mencionado, el efecto de la variación del ADNmt en el desarrollo de diversas patologías es notorio (Tabla 3). Es importante destacar que suele ser frecuente ver que el mismo haplogrupo se pueda asociar con numerosos fenotipos patológicos; esto es debido a que muchas de las variantes que componen estos conjuntos de SNPs tienen un efecto pleiotrópico, es decir, la misma variante puede producir más de un fenotipo: por ejemplo en el estudio de Marom y colaboradores (Marom *et al.*, 2017) los autores realizaron un meta-análisis de una gran colección de estudios donde se asociaban haplogrupos con diversas patologías y fenotipos no clínicos. Los resultados de este estudio enfatizaron la importancia de la heterogeneidad alélica presente en el genoma mitocondrial, ya que, por un lado los resultados mostraban que algunos de los haplogrupos se asociaban de forma específica con ciertos fenotipos concretos, como la asociación del haplogrupo T con un menor riesgo de desarrollar Párkinson, o el haplogrupo asiático A, que se asoció con un mayor riesgo de desarrollar la diabetes mellitus de tipo 2; mientras que, por otro lado, contrastando con las asociaciones individuales, en este estudio los haplogrupos H, K y J se asociaron con susceptibilidad con múltiples fenotipos patológicos, entre

ellos Alzheimer, Parkinson y distintos tipos de cáncer, denotando claramente un efecto pleiotrópico.

Tabla 3. Resumen de las principales asociaciones entre polimorfismos mitocondriales y diversas patologías en diferentes poblaciones humanas.

Patología	Variante	Riesgo	Referencia
Cáncer de próstata	m.8993G	Aumentado	Petrós <i>et al.</i> , 2005
	m.6124C	Aumentado	Arnold <i>et al.</i> , 2013
Diabetes mellitus tipo 2	m.16189C	Aumentado	Poulton <i>et al.</i> , 1998
LHON	m.3460A	Aumentado	Farrar <i>et al.</i> , 2013
	m.14484C	Aumentado	Farrar <i>et al.</i> , 2013
	m.11778A	Aumentado	Wallace 1994
Longevidad	m.150T	Aumentado	Chocron <i>et al.</i> , 2019
MELAS	m.3243G	Aumentado	Pavakis <i>et al.</i> , 1984
NARP	m.8993G	Aumentado	Ortiz <i>et al.</i> , 1993
	m.8993C	Aumentado	Ortiz <i>et al.</i> , 1993
Síndrome de Leigh	m.8993G	aumentado	Wallace 1994
Sordera	m.1555G	Aumentado	Wallace 1994
Patología	Haplogrupo	Riesgo	Referencia
Alzheimer	B5	Aumentado	Bi <i>et al.</i> , 2015
	Cluster HV	Aumentado	Maruszak <i>et al.</i> , 2009
	H5	Aumentado	Santoro <i>et al.</i> , 2010
Cáncer colorectal	T	Aumentado	Li <i>et al.</i> , 2015
Cáncer de mama	D5	Aumentado	Fang <i>et al.</i> , 2010
	T1a1	Reducido	Blein <i>et al.</i> , 2015
Cáncer de próstata	U	Aumentado	Booker <i>et al.</i> , 2006
Cáncer de tiroides	D4a	Aumentado	Fang <i>et al.</i> , 2010
Cáncer renal	U	Aumentado	Booker <i>et al.</i> , 2006
Cardiomiopatía hipertrófica	T	Aumentado	Castro <i>et al.</i> , 2006
Cardiomiopatía isquémica	H	Aumentado	Fernández-Caggiano <i>et al.</i> , 2012
	J	Reducido	Fernández-Caggiano <i>et al.</i> , 2012
Diabetes mellitus tipo 2	A	Aumentado	Nishigaki <i>et al.</i> , 2010
	D4b	Aumentado	Nishigaki <i>et al.</i> , 2010
	D5	Aumentado	Nishigaki <i>et al.</i> , 2010
	F	Aumentado	Nishigaki <i>et al.</i> , 2010
	N9a	Reducido	Nishigaki <i>et al.</i> , 2010
Enfermedad de las arterias coronarias	T	Aumentado	Kofler <i>et al.</i> , 2009
Esclerosis múltiple	J	Aumentado	Tranah <i>et al.</i> , 2015
	T	Aumentado	Tranah <i>et al.</i> , 2015
Infarto cerebral aterotrombótico	A	Reducido	Nishigaki <i>et al.</i> , 2010
Infarto de miocardio	G1	Aumentado	Nishigaki <i>et al.</i> , 2010
	N9b	Reducido	Nishigaki <i>et al.</i> , 2010
Longevidad	B4a	Reducido	Cai <i>et al.</i> , 2009
	D4a	Aumentado	Alexe <i>et al.</i> , 2007; Cai <i>et al.</i> , 2009
	D4b2b	Aumentado	Alexe <i>et al.</i> , 2007
	D5	Aumentado	Alexe <i>et al.</i> , 2007
	J	Aumentado	De benedictis <i>et al.</i> , 1999; Ross <i>et al.</i> , 2001; Niemi <i>et al.</i> , 2003
	M9	Reducido	Cai <i>et al.</i> , 2009
	N9	Reducido	Cai <i>et al.</i> , 2009
Parkinson	Cluster HV	Aumentado	Hudson <i>et al.</i> , 2013
	Cluster JT	Reducido	Hudson <i>et al.</i> , 2013
	J	Reducido	Van Der Walt <i>et al.</i> , 2003
Síndrome metabólico	D5	Reducido	Nishigaki <i>et al.</i> , 2010
	G1	Reducido	Nishigaki <i>et al.</i> , 2010
	N9a	Reducido	Nishigaki <i>et al.</i> , 2010

LHON: Lebers hereditary optical neuropathy; MELAS: Mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes syndrome; NARP: Neurogenic muscle weakness, ataxia and retinitis pigmentosa

## Introducción

Finalmente cabe destacar que, a pesar de todo lo comentado, no existe un haplogrupo ideal que evite el desarrollo de cualquier tipo de patología. Un mismo haplogrupo puede ser beneficioso o perjudicial dependiendo del *background* nuclear y las condiciones ambientales, ya que, por un lado una fosforilación oxidativa muy eficiente conlleva mayores niveles de ROS mitocondriales, aumentando el riesgo de desarrollar patologías sobre todo neurodegenerativas y, por otro lado, una fosforilación oxidativa con mayor grado de desacoplamiento producirá más calor a costa de generar menores niveles de energía, aumentando así el riesgo de desarrollar enfermedades asociadas con deficiencias energéticas como el LHON (Ruiz-Pesini *et al.*, 2004).

### *Mitocondria y artrosis.*

Durante los últimos años, la disfunción mitocondrial y la variación del ADNmt se han revelado como factores clave en el desarrollo del proceso artrósico (Blanco *et al.*, 2011; Blanco *et al.*, 2018). La disfunción mitocondrial se manifiesta principalmente como una disminución de la producción de ATP, un aumento del estrés oxidativo, la desregulación del calcio, un aumento de la permeabilidad de la membrana mitocondrial y alteraciones del ADNmt, que dan como resultado la degradación del cartílago articular (Mao *et al.*, 2020). La disfunción mitocondrial altera el correcto funcionamiento de las células, favoreciendo el desarrollo de diferentes patologías, además de la artrosis. Asimismo, la presencia de determinados polimorfismos genéticos mitocondriales puede modular la función mitocondrial, condicionando la predisposición a desarrollar estas patologías.

Los condrocitos presentes en el cartílago articular son células altamente glicolíticas que, debido a que están embebidos en la matriz del cartílago y que este tejido presenta un gradiente de concentración de nutrientes y oxígeno en profundidad, obtienen la mayoría de la energía que necesitan a través del metabolismo anaerobio de la glucosa. Sin embargo, se estima que aproximadamente el 25% de la energía que necesitan la obtienen a través de la fosforilación oxidativa mitocondrial (Terkeltaub *et al.*, 2002), por lo que este orgánulo es muy necesario para cubrir la alta demanda energética del condrocito. Se ha observado que durante el proceso artrósico, los condrocitos muestran una reducción de la actividad mitocondrial en los complejos II y III de la CTE y que, además, para intentar compensar esta reducción de la actividad, presentan una mayor masa mitocondrial si se compara con condrocitos sanos (Maneiro *et al.*, 2003). Por otro lado, se ha demostrado que la inhibición de los complejos III y V de la CTE mediante inhibidores específicos como antamicina y oligomicina, promueve la generación de citoquinas pro-inflamatorias, metaloproteinasas y un aumento de ROS (Cillero-Pastor *et al.*, 2008; Vaamonde-García *et al.*, 2012). Tal y como se explicó previamente, el metabolismo mitocondrial es una importante fuente de ROS y, debido a la disfunción mitocondrial se pueden generar excesivas cantidades de éstas moléculas. Además, en la disfunción mitocondrial se ha observado una menor expresión de la SOD2 en las capas superficiales de cartílagos con artrosis, reduciendo la capacidad de estos tejidos de combatir el estrés oxidativo (Ruiz-Romero *et al.*, 2009). Todo esto reduce la capacidad de reparación del condrocito y compromete su función celular, pudiendo desarrollar artrosis asociada al envejecimiento (Blanco *et al.*, 2004; Grishko *et al.*, 2009; Collins *et al.*, 2016).

La disfunción mitocondrial también puede inducir cambios anormales en el condrocito del cartílago articular humano que acaben promoviendo la degradación de este tejido durante la artrosis, por ejemplo los condrocitos pueden tener peor respuesta a factores de crecimiento, un aumento de la apoptosis celular a través de la vía mitocondrial, defectos en la autofagia (proceso por el cual la célula es capaz de eliminar proteínas, orgánulos o fracciones citoplasmáticas), una reducción de la longitud telomérica, mayor senescencia celular y un

aumento de la respuesta inflamatoria (Kim y Blanco, 2007; Tamayo *et al.*, 2011; Lotz y Loeser, 2012). Asimismo, se ven afectadas por todo este proceso diversas dinámicas mitocondriales que se han relacionado recientemente con la artrosis. Entre estas dinámicas podemos encontrar los fenómenos mitocondriales de biogénesis, término que incluye a los procesos fusión y fisión que permiten que estos orgánulos se unan o se dividan generando grandes redes o nuevas mitocondrias y, la mitofagia, que consiste en la eliminación de mitocondrias dañadas. Estos procesos de dinámica mitocondrial, que tienen como objetivo adaptar a estos orgánulos a las necesidades energéticas de las células y mejorar la supervivencia celular, se ven altamente afectados por la disfunción mitocondrial, pudiendo comprometer la función del condrocito y la supervivencia celular (Naik *et al.*, 2019; Blanco y Fernández-Moreno, 2020). En definitiva, como consecuencia de la disfunción mitocondrial, la función y supervivencia del condrocito se ven enormemente afectadas, empeorando el correcto mantenimiento de la matriz del cartílago articular y favoreciendo los procesos catabólicos, lo cual acaba derivando a lo largo del tiempo en el proceso de la artrosis (Blanco *et al.*, 2018).

La artrosis es una enfermedad multifactorial y, tal y como se ha reflejado a lo largo de esta tesis, la prevalencia, la progresión y la incidencia de esta enfermedad pueden verse influenciadas por la presencia de diversos factores de riesgo. El papel de la mitocondria en el desarrollo de la artrosis se vio definitivamente reforzado a raíz del descubrimiento de la asociación entre la genética mitocondrial, a través de los haplogrupos, y diferentes aspectos relacionados con esta enfermedad:

- Prevalencia: en el año 2008, Rego-Pérez y colaboradores publicaron el primer estudio que relacionó los haplogrupos mitocondriales con la susceptibilidad a padecer artrosis de rodilla. Dicho estudio se realizó en la cohorte de pacientes española PROspective COhort of A Coruña (PROCOAC) y en él se observó que los individuos portadores del haplogrupo mitocondrial J tenían un menor riesgo de desarrollar artrosis de rodilla. En este mismo trabajo los investigadores también encontraron que, en caso de desarrollar esta patología, ser portador de dicho haplogrupo reducía la severidad de la artrosis de rodilla, a nivel de su grado KL. Sin embargo, ser portador del haplogrupo U la incrementaba (Rego-Pérez *et al.*, 2008). Este mismo grupo, en el año 2010, replicó esta misma asociación en una cohorte española de pacientes con artrosis de cadera (Rego *et al.*, 2010). El siguiente paso de este grupo fue analizar si esto ocurría en otras poblaciones, y realizaron un estudio similar en el año 2014 (A. Soto-Hermida *et al.*, 2014), donde compararon las frecuencias de los haplogrupos en una cohorte de pacientes del Reino Unido con los datos que tenían de las poblaciones españolas. En la población española, el haplogrupo J mostró una menor frecuencia en los individuos con artrosis, mientras que, en la cohorte de Reino Unido, fue el haplogrupo T el que mostró el efecto protector. Entre ambas cohortes los investigadores observaron diferencias en la distribución del haplogrupo J, presentando en los individuos sanos de España una mayor frecuencia que en los individuos sanos de Reino Unido, seguramente debido al proceso de adaptación climática previamente mencionado. Además, cabe destacar que la composición de los sujetos con haplogrupo J del Reino Unido es diferente a la de los sujetos españoles con haplogrupo J, presentando estos últimos mayor frecuencia de la variante m.3394T>C. Esto puede promover que un mismo haplogrupo, en dos poblaciones diferentes, presente distintos comportamientos (A. Soto-Hermida *et al.*, 2014).

A pesar de que algunos estudios no consiguieron replicar estos resultados en otras cohortes europeas (Hudson, Panoutsopoulou, *et al.*, 2013), en un meta-análisis

## Introducción

realizado en el año 2014 en el que se englobaron 6 estudios de casos y controles sumando un total aproximado de 17.000 participantes (Shen *et al.*, 2014), sí se replicó la asociación del cluster mitocondrial JT como factor protector para el desarrollo de la artrosis.

También cabe destacar que en poblaciones del sur de China, se encontraron asociaciones significativas entre la presencia del haplogrupo G con un mayor riesgo de desarrollar artrosis, mientras que los haplogrupos B y B4 se asociaron con un menor riesgo (Fang *et al.*, 2014)

- Progresión: Trabajos posteriores del grupo de Rego y colaboradores en la misma cohorte de pacientes española PROCOAC mostraron que los individuos pertenecientes al clúster mitocondrial JT presentaban menores tasas de progresión radiográfica que los individuos pertenecientes al cluster UK, y que los pacientes portadores del haplogrupo H presentaban una mayor tendencia a necesitar una cirugía de reemplazo articular de rodilla (Soto-Hermida *et al.*, 2015). Además, empleando muestras de ADN de la cohorte Americana de la *Osteoarthritis Initiative* (OAI), estos mismos autores mostraron que el haplogrupo T también se asociaba, no sólo con una menor tasa de progresión radiológica de la artrosis de rodilla, sino también con una menor formación de osteofitos, un menor estrechamiento intraarticular, una menor esclerosis subcondral y una reducción de la pérdida de integridad del cartílago articular durante 48 meses (Soto-Hermida *et al.*, 2014). Finalmente, en el año 2017 se llevó a cabo un estudio de replicación seguido de un meta-análisis incluyendo datos de la cohorte OAI, de la cohorte de los países bajos *Cohort Hip and Cohort Knee* (CHECK) y de la la cohorte PROCOAC (Fernández-Moreno *et al.*, 2017). En dicho estudio se confirmó la asociación del haplogrupo T y del clúster JT con una menor tasa de progresión de artrosis de rodilla.
- Incidencia: El uso de cohortes longitudinales, como la OAI, la CHECK o la PROCOAC, no sólo son ideales para llevar a cabo estudios de asociación genética relacionados con la progresión de la enfermedad, sino también para realizar aquellos estudios relacionados con la tasa de incidencia. En esta línea, Fernández-Moreno y colaboradores (Fernández-Moreno *et al.*, 2017) realizaron un estudio en estas cohortes para analizar la tasa de incidencia de la artrosis de rodilla en base a los haplogrupos mitocondriales. Los criterios de incidencia que utilizaron en la cohorte OAI en este estudio fueron los propuestos por Felson y colaboradores (Felson *et al.*, 2011), de tal manera que se consideraron pacientes con incidencia de artrosis a todos aquellos sujetos que, teniendo un grado KL basal  $\leq 1$ , alcanzaban un grado  $\geq 2$  durante el seguimiento. Por otro lado, en la cohorte CHECK emplearon los criterios definidos por Thorstensson y colaboradores (Thorstensson *et al.*, 2009), considerando pacientes con incidencia de artrosis a aquellos sujetos que, teniendo un grado basal KL  $< 1$ , alcanzaban un grado  $\geq 1$  durante el seguimiento. Tras definir los grupos apropiadamente y realizar los pertinentes análisis estadísticos, Fernández-Moreno y colaboradores pudieron observar que, en ambas cohortes, los sujetos portadores del haplogrupo J presentaban una menor tasa de incidencia de la enfermedad a 96 meses comparado con los sujetos portadores del haplogrupo H. El estudio concluye con un meta-análisis de 3.214 casos, en el que se ratifica de forma significativa esta asociación.

Además de estudios de asociación en poblaciones caucásicas, también se ha analizado el impacto de los haplogrupos mitocondriales en poblaciones asiáticas. En esta línea, existe un trabajo que analiza el impacto de estas variantes genéticas mitocondriales en la incidencia de la artrosis de rodilla en poblaciones coreanas. Los



autores encontraron que el haplogrupo B se asociaba de forma significativa con una mayor tasa de incidencia de artrosis de rodilla en esta población (Koo *et al.*, 2019).

Debido a la importancia cada vez mayor que tienen las alteraciones mitocondriales en el desarrollo de la artrosis, algunos autores han planteado la posibilidad de establecer la mitocondria como una diana terapéutica, para así poder desarrollar fármacos que puedan mejorar la actividad mitocondrial. Algunos estudios preclínicos indican que la regulación de la función mitocondrial podría ayudar a reducir la apoptosis del condrocito (Kim y Blanco, 2007), siendo uno de los enfoques más prometedores en este campo el desarrollo de inhibidores sintéticos de la caspasa 9. Un ejemplo de ello es un estudio en perros donde se les inducía artrosis de rodilla mediante un corte en el ligamento cruzado anterior; los cartílagos de estos perros, al ser tratados en condiciones *in vitro* con un inhibidor de caspasa 9, Z-LEHD-FMK, mostraron menores niveles de apoptosis en los condrocitos (Pelletier *et al.*, 2001). Otros estudios se centraron en un enfoque diferente que consistía en la búsqueda de compuestos que permiten reducir la síntesis de ROS o bloquear sus efectos. Por ejemplo, el resveratrol, el cuál es un compuesto natural que parece tener un efecto antiinflamatorio *in vitro*, incluso presenta capacidades condroprotectoras ya que suprime la producción de ROS, de prostaglandinas, de IL-1 y aumenta los niveles de la proteína supresora de tumores p53 (Csaki *et al.*, 2008).

Algunos de los enfoques terapéuticos más novedosos a día de hoy, se centran directamente en la mitocondria *per se*. Uno de ellos consiste en editar directamente el genoma mitocondrial con el fin de corregir aquellas variantes presentes en heteroplasmia que puedan llevar a la disfunción mitocondrial. Entre estas herramientas se encuentran las terapias antígenómicas, el uso de endonucleasas de restricción, las nucleasas con dedos de zinc, nucleasas efectoras similares a activadores de transcripción y la tecnología CRISPR (del inglés *clustered regularly interspaced short palindromic repeats*)(Zhong *et al.*, 2022). Otra aproximación novedosa consiste en el trasplante de mitocondrias sanas desde células madre estromales. Por ejemplo, un estudio en 2020 obtuvo mitocondrias sanas de células madre estromales obtenidas a partir de médula ósea de ratas sanas y, éstas fueron trasplantadas a condrocitos artrósicos de ratas de avanzada edad mediante un co-cultivo. Realizaron diversos análisis y pudieron observar que, en comparación con los condrocitos artrósicos sin mitocondrias trasplantadas, las células tratadas presentaban un mayor potencial de membrana, mayor actividad de los complejos de CTE, mayor producción de ATP, menores niveles de apoptosis celular, y mayor secreción de colágeno tipo II y PGs (R. Wang *et al.*, 2021). Otro estudio similar fue desarrollado en 2022 y en él analizaron los efectos *in vitro* e *in vivo* del trasplante de mitocondrias sanas. Se obtuvieron mitocondrias sanas de células humanas musculares y de la línea celular de ratas L6. Las mitocondrias humanas se co-cultivaron junto con condrocitos artrósicos humanos, mientras que las de rata se trasplantaron a condrocitos de rata y, también se inyectaron en rodillas de ratas con artrosis inducida. Como resultado, en ambos modelos *in vitro* el trasplante de mitocondrias a condrocitos de rata y el co-cultivo de condrocitos humanos con mitocondrias sanas mostraban una mejora de la función mitocondrial. Finalmente en el modelo *in vivo* las ratas artrósicas con mitocondrias trasplantadas mostraron mejorías en el dolor, una menor destrucción del cartílago y pérdida de hueso, así como una menor producción de factores inflamatorios (Lee *et al.*, 2022). Todas estas metodologías permitirían restaurar el estado normal de las mitocondrias pudiendo generar efectos protectores frente a la artrosis, siendo muy prometedoras; sin embargo, cabe destacar que todavía están realizando sus primeros pasos y existen numerosos desafíos todavía por resolver.

## Introducción

### *Los haplogrupos mitocondriales como marcadores genéticos complementarios en la artrosis.*

Entre los años 2009 y 2012 se realizaron diferentes estudios que trataron de valorar el impacto de la presencia de determinados haplogrupos en los niveles séricos de diferentes proteínas relacionadas con el proceso artrósico. El primer estudio consistió en analizar los niveles de biomarcadores moleculares séricos resultantes del proceso del metabolismo del colágeno II. En este trabajo, los individuos portadores de haplogrupo H mostraban mayores niveles de péptidos del colágeno II resultantes del catabolismo del cartílago articular, mientras que los portadores de haplogrupo J presentaban niveles más bajos (Rego-Pérez *et al.*, 2010). Al año siguiente analizaron si existían diferencias en los niveles séricos de determinadas enzimas proteolíticas que participan en el proceso artrósico y, en su estudio descubrieron que los individuos portadores del haplogrupos H presentaban mayores niveles de enzimas catabólicas, como las metaloproteinasas 3 y 13, que los individuos portadores del haplogrupo J (Rego-Pérez *et al.*, 2011). En ese mismo año también analizaron la producción de NO, una molécula de RNS (del inglés, *reactive nitrogen species*), y la longitud telomérica en leucocitos periféricos de individuos con distintos haplogrupos. Tras analizar los datos concluyeron que los condrocitos aislados a partir de individuos portadores del haplogrupo J generaban menores niveles de NO que los condrocitos portadores de otros haplogrupos. Asimismo, también detectaron que los leucocitos de individuos del haplogrupo J tenían una mayor longitud telomérica (Fernández-Moreno *et al.*, 2011). En línea con todos los hallazgos que fueron obteniendo, en el año 2012 se publicó una revisión que incluía el análisis de 12 marcadores séricos de artrosis entre individuos portadores de los haplogrupos H y J. La principal conclusión de este trabajo fue la recomendación de utilizar los haplogrupos mitocondriales como marcadores complementarios de otros biomarcadores proteicos relacionados con la artrosis (Fernández-Moreno *et al.*, 2012).

Además, los haplogrupos pueden servir como marcadores indirectos de artrosis, ya que pueden condicionar a otros factores de riesgo de la enfermedad que favorecen su desarrollo. Por ejemplo, en 2018 se analizó el efecto de los haplogrupos mitocondriales en el desarrollo de grandes lesiones de médula ósea en el compartimento medial de la rodilla. Este tipo de lesiones se han asociado de forma robusta con la progresión sintomática y el empeoramiento radiológico de la articulación en períodos de 24 meses, posicionando a estas lesiones como un marcador de actividad de la enfermedad. En dicho estudio llevado a cabo en sujetos de la cohorte del OAI se observó que, en comparación con los individuos portadores de haplogrupo mitocondrial H, los portadores del haplogrupo J mostraban un menor riesgo de desarrollar grandes lesiones de médula ósea en el compartimento medial de la rodilla (Rego-Pérez *et al.*, 2018). Otro factor de riesgo especialmente importante en la artrosis es la obesidad, ya que como se mencionó previamente el aumento de peso aumenta la carga de las articulaciones y también contribuye al desarrollo de moléculas inflamatorias. También en 2018 se realizó un estudio en la cohorte OAI con el objetivo de ver si la presencia de determinados haplogrupos se podían asociar con la obesidad. Como resultado en comparación con individuos portadores de los clusters mitocondriales HV, los portadores de haplogrupos minoritarios I, W y X mostraron un riesgo aumentado del 48% de sufrir obesidad (Veronese *et al.*, 2018). En esta línea, un estudio reciente llevado a cabo en una cohorte de pacientes mejicana analizó como la interacción entre los haplogrupos mitocondriales y la obesidad condicionaban el riesgo de desarrollar artrosis de rodilla. En este estudio clasificaron a los participantes en base a la presencia de artrosis y la obesidad. Analizaron el haplogrupo de los participantes y, en los análisis realizados, detectaron una interacción significativa entre el haplogrupo A y la obesidad que se asociaba de forma significativa con el riesgo de desarrollar artrosis de rodilla (Ramos-Louro *et al.*, 2022).

### *Metodologías para analizar el significado funcional de variantes en el genoma mitocondrial.*

Con el fin de identificar los efectos que determinadas variantes mitocondriales causan en el metabolismo y en el comportamiento celular, se han desarrollado diferentes estudios funcionales utilizando tanto modelos celulares como modelos animales. Los modelos celulares se basan en el uso de los cíbridos transmitocondriales. Los cíbridos son células que se crean al fusionar plaquetas (células sanguíneas enucleadas con mitocondrias) portadoras de un haplogrupo conocido, con células a las que se les ha eliminado previamente su ADNmt, denominadas Rho-0 ( $\rho 0$ ). Existen multitud de líneas celulares que se pueden emplear como célula  $\rho 0$ , algunos ejemplos serían la línea celular de osteosarcoma denominada 143B.TK<sup>-</sup> Rho-0 o la de células pigmentarias epiteliales de retina ARPE-19. Como resultado de este proceso de fusión, se obtienen células con el mismo núcleo, pero con distintas mitocondrias, siendo el modelo ideal para analizar y comparar el efecto de los haplogrupos mitocondriales.

Usando esta metodología, en el trabajo de Fernández-Moreno y colaboradores (Fernández-Moreno *et al.*, 2017), los autores también desarrollaron un modelo funcional con cíbridos, usando una línea celular 143B.TK<sup>-</sup> Rho-0 y plaquetas de individuos portadores del haplogrupo H y del haplogrupo J. Los cíbridos portadores del haplogrupo J mostraron menores niveles de ROS, mayores niveles de supervivencia bajo condiciones de estrés oxidativo y menores niveles de apoptosis, aunque también mostraron menores niveles de respiración y glucólisis, lo cual provoca una menor producción de ATP. Estos datos, al igual que muchos de los mencionados a lo largo de esta tesis doctoral, continúan mostrando la misma tendencia que hasta ahora, y van de la mano con la hipótesis ya mencionada de Ruiz-Pesini y colaboradores (Ruiz-Pesini *et al.*, 2004)(Fig. 26), por la cual una menor producción de ATP lleva aparejada una menor producción de ROS. Otro trabajo que emplea líneas celulares de cíbridos en la artrosis es el de Dalmao-Fernández y colaboradores de 2020. En él generaron cíbridos a partir de la línea celular 143B.TK<sup>-</sup> Rho-0 y plaquetas de donantes sanos y artrósicos. Compararon entre los distintos cíbridos los metabolismos de la glucosa y ácidos grasos, y detectaron que los cíbridos creados a partir de mitocondrias artrósicas presentaban un metabolismo de ácidos grasos incompleto en comparación con los cíbridos sanos. En este estudio también se analizó la flexibilidad metabólica de los distintos cíbridos, es decir la capacidad de cambiar de un metabolismo de glucosa a uno de ácidos grasos dependiendo de las condiciones del medio. Los resultados indicaron que los cíbridos sanos tenían una mayor flexibilidad metabólica que los cíbridos artrósicos ya que los donantes sanos mantuvieron el fenotipo glucolítico, mientras que los donantes artrósicos mostraron preferencia por el metabolismo del ácido oleico. Los resultados obtenidos en este trabajo con cíbridos podrían ser extrapolados a los condrocitos que, a pesar de ser células altamente glucolíticas, también necesitan ácidos grasos para generar energía, usarlos como componentes estructurales o moléculas de señalización, lo cual podría justificar la relación entre la disfunción mitocondrial y el procesos artrósicos (Dalmao-Fernández *et al.*, 2020). Con el descubrimiento del haplogrupo G como factor de riesgo y el haplogrupo B4 como factor protector frente a la artrosis en poblaciones asiáticas, Fang y colaboradores en 2016 llevaron a cabo un estudio con cíbridos a partir de la línea celular 143B.TK<sup>-</sup> Rho-0 y plaquetas de donantes con haplogrupo G y B4. En dicho estudio demostraron que los niveles de actividad de los complejos I y III de la CTE son más altos en los cíbridos G que en los portadores del haplogrupo B4. También comprobaron que los cíbridos G generan más ATP a través de la fosforilación oxidativa que a través de la glucólisis, disminuyendo así la viabilidad de estos cíbridos en condiciones de hipoxia en comparación con los cíbridos B4. El perfil de expresión génica de los cíbridos G y B4 fue analizado utilizando secuenciación masiva y mostró que 404 de 575 genes se

## Introducción

expresaban diferencialmente entre los híbridos G y B4, afectando a 17 vías, de las cuales 11 participan en la artrosis. Por tanto, concluyeron que el haplogrupo asiático G parece aumentar el riesgo de desarrollar artrosis debido a que cambia el perfil metabólico de la célula de glucólisis a fosforilación oxidativa, y también debido a que se activan más vías de señalización relacionadas con la artrosis (Fang *et al.*, 2016).

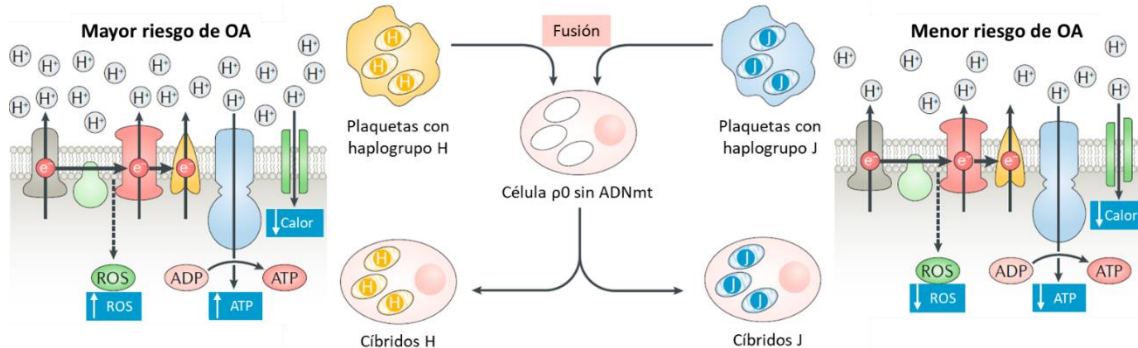


Figura 26. Modelos funcionales con híbridos. Los híbridos son células artificiales creadas a partir de la fusión de plaquetas con ADNmt, pero sin ADN nuclear, con una línea celular portadora de ADN nuclear, pero sin ADNmt. Como resultado se obtienen células que solamente se diferencian en base al ADNmt, siendo idóneas para el estudio de este elemento celular en el desarrollo de patologías. Estos híbridos han permitido observar que los individuos portadores del haplogrupo J generan menores niveles de ROS y energía, a cambio de generar más calor, esto les proporciona un efecto protector frente a patologías como la artrosis. Por otro lado, los individuos portadores del haplogrupo H generan mucho más ATP y ROS, a costa de generar menos calor, lo cual les proporciona un efecto protector frente a enfermedades con elevada dependencia energética, pero representa un factor de riesgo para enfermedades neurodegenerativas e incluso la artrosis. Imagen modificada de Blanco *et al.* 2018.

En cuanto a los modelos animales, estos están basados en el uso de ratones conplásticos. Los ratones conplásticos son el resultado de realizar retrocruzamientos entre dos cepas de ratones específicas, y como resultado se obtienen ratones con núcleos idénticos, pero con distintos ADNmt, permitiendo de nuevo comparar variantes mitocondriales (Fig 27). Estudios como el de Latorre-Pellicer y colaboradores (Latorre-Pellicer *et al.*, 2016) o el de Brinker y colaboradores en el campo de la oncología (Brinker *et al.*, 2017) han demostrado el impacto que tienen las variantes mitocondriales en la salud en general, así como en diferentes rasgos relacionados como son la longevidad, la proteostasis mitocondrial, la generación de ROS, la señalización mediada por insulina, la obesidad, el acortamiento telomérico, la disfunción mitocondrial, y otros caracteres complejos relacionados con diferentes enfermedades humanas.

En el campo de la artrosis, el primer trabajo que se publicó utilizando estos modelos animales data del año 2021. En este estudio, los autores generan un modelo de ratón con artrosis inducida quirúrgicamente y demuestran que la composición nucleotídica del ADNmt puede alterar drásticamente el curso de la enfermedad. Así, en este estudio, los autores utilizan dos cepas de ratones C57BL/6, una de las cuales contenía tanto el ADN nuclear como el ADNmt de la cepa original (C57BL/6<sup>C57</sup>), mientras que la segunda tenía el ADN nuclear original pero el ADNmt de la cepa NZB (C57BL/6<sup>NZB</sup>). Tras inducir la artrosis se evaluaron en las rodillas de estos animales diversas características histológicas, y también se analizaron distintos procesos clave como la autofagia, la apoptosis, el estrés oxidativo y la senescencia celular. Los ratones con el genoma mitocondrial NZB mostraron, a nivel del cartílago articular, una mayor celularidad y una menor puntuación de la escala histopatológica OARSI; además, también presentaban mayores niveles de autofagia y menores niveles de apoptosis. Como resultado, los ratones C57BL/6<sup>NZB</sup> desarrollaban artrosis de rodilla menos severa comparada con la cepa original. (Scotece *et al.*, 2021). En un estudio posterior en el año 2022, estos mismos autores trataron de analizar la influencia de las variantes mitocondriales en el daño articular en un modelo de envejecimiento basado en ratones conplásticos. Para ello trabajaron exactamente con las mismas dos cepas,

C57BL/6<sup>C57</sup> C57BL/6<sup>NZB</sup> y se sacrificaron ratones a 25, 75 y 90 semanas de edad. Tras el sacrificio se graduó la patología articular, se analizaron marcadores de autofagia, de senescencia y proliferación, además también se midieron los niveles de 8-OHdG y de caspasa 3 proteolizada. Los resultados, al igual que en artículo anterior, mostraron que los ratones C57BL/6<sup>NZB</sup> presentaban menores niveles de daño articular a las 25, 75 y 90 semanas de edad, también mostraron mayores niveles de marcadores de autofagia y proliferación y menores niveles de marcadores de senescencia. Además, la cepa original C57BL/6<sup>C57</sup>, mostró mayores niveles de daño por moléculas de ROS y mayores niveles de caspasa 3 proteolizada, indicando un mayor nivel de apoptosis celular (Scotece *et al.*, 2022). Por tanto, estos dos trabajos demostraron, por primera vez en la artrosis, que el ADNmt altera drásticamente el curso de la enfermedad.

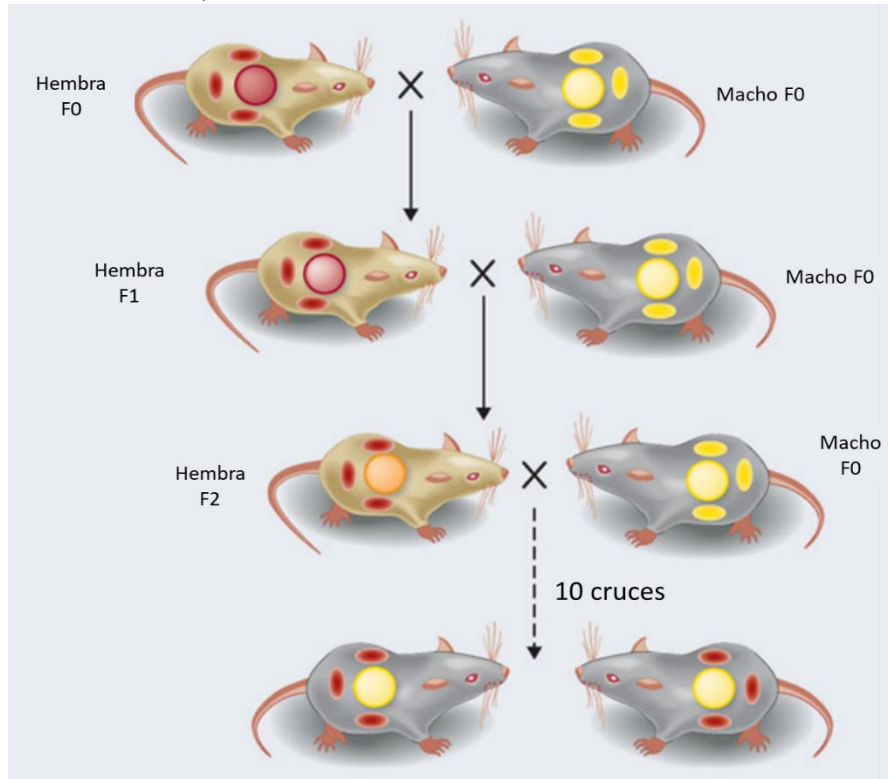


Figura 27. Creación de ratones conplásticos. Este tipo de ratones se establecen mediante una serie de retrocruzamientos que involucran dos cepas endogámicas. Este esquema de cría consiste en seleccionar solo hembras para las generaciones posteriores. Las hembras F1 portan mitocondrias (óvalos rojos) heredadas únicamente de su madre, mientras que el 50 % de su genoma nuclear se hereda de su padre y el 50 % de su genoma nuclear se hereda de su madre. Luego, las hembras F1 se cruzan con machos de la cepa progenitora original, generando hembras F2 que todavía tienen todas sus mitocondrias de la cepa progenitora materna, mientras que el 75% de su genoma nuclear proviene de su padre y solo el 25% de su genoma nuclear es de su madre. Después de 10 generaciones de retrocruzamiento, el genoma nuclear (círculo amarillo) es esencialmente idéntico al genoma nuclear de la cepa progenitora masculina original, mientras que las mitocondrias lo son únicamente de la cepa progenitora femenina original. Imagen modificada de Bussard and Siracusa, 2017.

#### Interacciones núcleo-mitocondria.

El ADN nuclear y el ADNmt no funcionan de forma estanca, de hecho, como ya se ha explicado en esta tesis, para realizar diversos procesos mitocondriales necesitan interactuar entre ellos, pero, también uno u otro material genético puede alterar la expresión de genes de la otra molécula. La expresión de los genes nucleares puede verse afectada por la presencia de variantes en el ADNmt. Esto es debido a que diversos productos resultantes del metabolismo mitocondrial, como las moléculas de ATP, el acetil-CoA, el  $\alpha$ -cetoglutarato o incluso la

## Introducción

producción de ROS, pueden producir modificaciones en proteínas señalizadoras citoplasmáticas o añadir alteraciones a proteínas nucleares que, junto con las modificaciones epigenéticas ya presentes en el ADN nuclear, constituyen lo que se denomina el epigenoma (Pakula, 2019). Las modificaciones epigenéticas cambian los niveles de expresión génica de diversos genes sin alterar la secuencia del ADN. Estas modificaciones pueden clasificarse en tres tipos (Shen *et al.*, 2017):

1. Las modificaciones postranscripcionales de las histonas: las cuales modifican la conformación de la cromatina.
2. Los ARN no codificantes: estos son capaces de alterar durante y tras la transcripción la expresión de los ARNm, impidiendo realizar su traducción o incluso degradándolo.
3. La metilación del ADN: la adición de grupos metilo en las zonas promotoras de los genes pueden alterar los niveles de expresión de éstos.

Teniendo en cuenta lo explicado en el párrafo anterior y la teoría del origen de los haplogrupos del ADNmt, la cual explica que la variación de haplogrupos es el producto de una adaptación al clima, se puede interpretar que la mitocondria actúa como un sensor clave para adaptarse a cambios medioambientales. Los cambios en el medioambiente afectan a la bioenergética de este orgánulo, alterando la producción de moléculas mitocondriales. Las concentraciones alteradas de estas moléculas llevan a modificaciones del epigenoma, reprogramando el patrón de expresión del ADN nuclear. En este punto, los cambios de expresión en el ADN nuclear producen, a su vez, modificaciones en la expresión génica mitocondrial con el fin de reestablecer el balance energético, estableciéndose lo que se denomina una comunicación bidireccional. Sin embargo, si los cambios locales son demasiado fuertes como para ser controlados por el ADNmt, este sistema de balance puede fallar, produciéndose un desequilibrio en las proteínas mito-nucleares, que ocasiona una reducción de la respiración mitocondrial, por tanto se genera menos ATP y tiene lugar finalmente la muerte celular (Quirós *et al.*, 2016; Pakula, 2019).

La adecuada coordinación entre las funciones nucleares y mitocondriales es esencial para mantener tanto la funcionalidad como los requerimientos energéticos celulares. La comunicación entre estos dos sistemas puede ocurrir en dos sentidos (Fig. 28):

1. Comunicación anterógrada: el núcleo manda señales a la mitocondria para que ésta modifique tanto su actividad como su cantidad. Como resultado, se modulan los niveles de genes nucleares que codifican para proteínas mitocondriales, genes de regulación de la traducción mitocondrial, genes de biogénesis mitocondrial, autofagia y dinámica mitocondrial. Los receptores nucleares y los factores de transcripción que intervienen en este proceso dependen de señales intra y extracelulares, como los niveles de nutrientes u hormonas que desencadenan esta respuesta nuclear. En este proceso de comunicación anterógrada destaca tanto la superfamilia de receptores activados por proliferadores peroxisomales (PPAR, del inglés *Perosyxomel proliferator activated receptors*) o coactivador del receptor gamma 1-alfa activado por el proliferador de peroxisomas (PGC-1s, del inglés *Peroxisome proliferator-activated receptor-gamma coactivator*), los cuales se encargan de regular diversos genes relacionados con el metabolismo de la glucosa y los lípidos (Mello *et al.*, 2019)
2. Comunicación retrógrada: la mitocondria informa al núcleo acerca del estrés oxidativo, la cantidad de ATP y metabolitos, fallos en la fosforilación oxidativa, cambios en el potencial de membrana o plegamientos proteicos incorrectos. Como

resultado, se activa la transcripción de diversos elementos nucleares para compensar estos problemas. En este tipo de señalización, la mitocondria puede estimular la transcripción de genes nucleares a través de fluctuaciones en los niveles de metabolitos, estrés oxidativo, estrés energético o niveles de  $Ca^{2+}$ . Esta respuesta nuclear puede favorecer la adaptación a hipoxia (muy importante en el desarrollo de tumores) o también pueden modular la función de NRF2 (del inglés *nuclear factor erythroid 2-related factor 2*), produciendo respuestas nucleares frente al estrés oxidativo. Como resultado, se inician procesos de regulación epigenética sobre el ADN nuclear, promoviendo así respuestas al estrés oxidativo o energético, o incluso favorecer procesos de autofagia, con el fin de adaptar la función mitocondrial a las condiciones metabólicas necesarias, promoviendo así la supervivencia celular (Mello *et al.*, 2019)

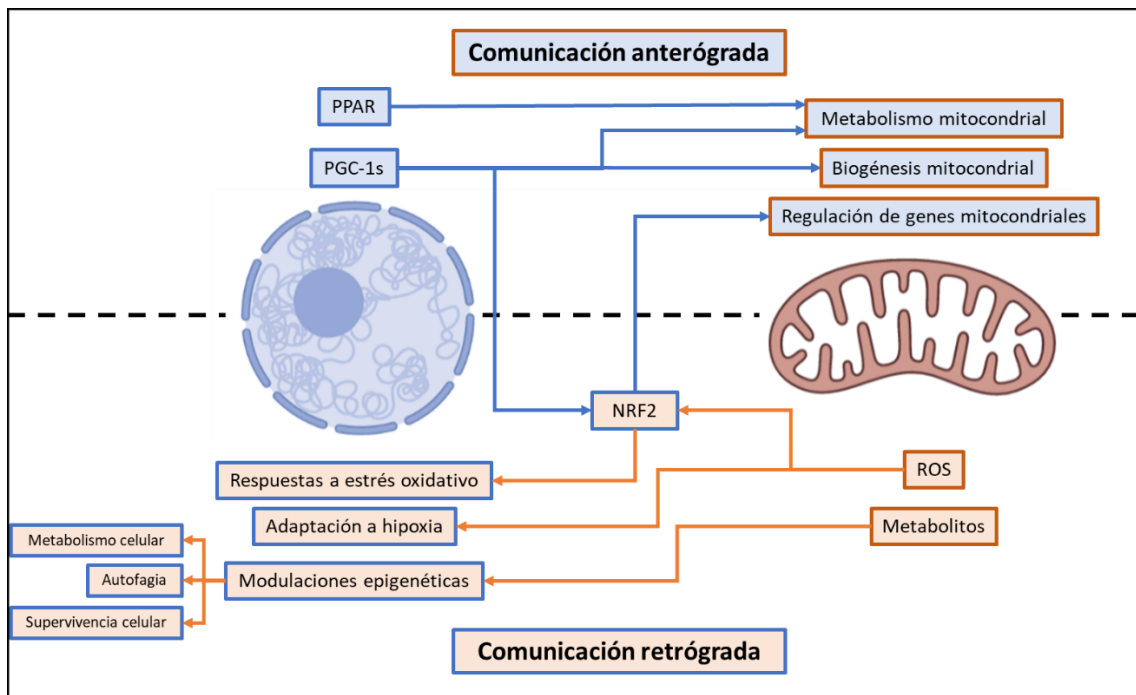


Figura 28. Comunicación núcleo-mitocondria. En la comunicación anterógrada, miembros de la familia PPAR regulan el metabolismo de la glucosa y el metabolismo lipídico, mientras que los miembros de la familia PGC-1s son los principales encargados de regular la integridad, función y biogénesis mitocondrial, además los PGC-1s también pueden modular la expresión de genes nucleares y mitocondriales relacionados con la CTE a través de la expresión de NRF2. En la comunicación retrógrada, subproductos de las funciones mitocondriales pueden producir respuestas nucleares, las moléculas de ROS pueden generar respuestas nucleares que favorezcan la adaptación a la hipoxia (muy importante en el desarrollo de tumores) o también pueden modular la función de NRF2, produciendo respuestas nucleares frente al estrés oxidativo. Los metabolitos resultantes del metabolismo mitocondrial pueden producir modulaciones epigenéticas sobre el ADN nuclear, como cambios en los patrones de metilación de diversos genes, promoviéndose distintos tipos de respuestas celulares que favorezcan procesos de autofagia e incluso, se pueden desregular oncogenes y genes supresores de tumores, que promoverían el metabolismo celular y la supervivencia celular (siendo respuestas especialmente importantes en la generación de tumores).

Como se ha comentado anteriormente, uno de los principales mecanismos por los que se cree que actúan las interacciones entre el núcleo y la mitocondria es a través de modificaciones en el epigenoma (modificaciones epigenéticas). En la artrosis se hipotetiza que uno de los principales cambios epigenéticos es la metilación (Shen *et al.*, 2017). La metilación es un proceso de silenciamiento génico en el que tiene lugar la transferencia de grupos metilo ( $-CH_3$ ) a determinadas citosinas localizadas previamente a una guanina, principalmente en las regiones promotoras con una determinada densidad de bases CG, o islas CpG. Así como la

## Introducción

presencia de determinados polimorfismos nucleares puede inducir ligeras modificaciones en el perfil de metilación a través de los denominados meQTLs (del inglés, *methylation quantitative trait loci*), la presencia de polimorfismos genéticos mitocondriales también. Por ejemplo, en un estudio de expresión génica nuclear con líneas celulares portadoras de diversos niveles de heteroplasma de la mutación m.3243A>G se trató de analizar si existían cambios a nivel del perfil de expresión de genes nucleares. Como resultado, los investigadores pudieron observar que cada nivel de heteroplasma se asociaba con un patrón de expresión génica nuclear distinto, mostrando una clara regulación del ADN nuclear por parte del mitocondrial (Picard *et al.*, 2014). Centrándonos en los haplogrupos mitocondriales, existen dos estudios independientes con híbridos transmitocondriales que muestran que los niveles globales de metilación del ADN nuclear están alterados por la presencia de distintos haplogrupos; así, en el año 2012 Bellizzi y colaboradores analizaron los patrones de metilación globales de ADN de sangre periférica de individuos con distintos haplogrupos, y pudieron observar que los individuos portadores de haplogrupos J tenían unos niveles de metilación globales mayores que los portadores de otros haplogrupos. También realizaron estudios *in vitro* con híbridos transmitocondriales, y observaron que los híbridos portadores de haplogrupos J mostraban mayor expresión de genes asociados a metilación y menor producción de ATP y ROS (Bellizzi *et al.*, 2012). Ya en el año 2015 Atilano y colaboradores desarrollaron híbridos portadores de haplogrupos H y J a partir de células de la retina humana y, en sus resultados, detectaron diferencias en la expresión de distintos genes nucleares relacionados con la metilación y la acetilación de histonas. Cuando trataron los híbridos con un inhibidor de la metilación, las diferencias de expresión de estos genes desaparecieron, demostrando que estos híbridos presentaban distintos perfiles de metilación en función de su haplogrupo mitocondrial (Atilano *et al.*, 2015).

Con respecto a la artrosis, un estudio del año 2019 analizó, por primera vez, el efecto de la variación genética del ADNmt en el perfil de metilación del cartílago articular. Dicho estudio mostró que los cartílagos portadores de los haplogrupos H y J mostraban patrones de metilación diferentes independientemente del diagnóstico de artrosis. En este proyecto los autores detectaron un total de 538 loci diferencialmente metilados entre células portadores del haplogrupo H y células portadores del haplogrupo J, así como 2.384 genes con una expresión diferencial entre ambos haplogrupos. De todos estos genes, 17 mostraron una correlación inversa entre los niveles de metilación y expresión. Concretamente, los autores observaron que en los cartílagos portadores del haplogrupo H la apoptosis era más frecuente, mientras que en los cartílagos de individuos J, la apoptosis estaba más reprimida. Además, en comparación con los cartílagos H, las muestras que presentaban el haplogrupo J mostraron mayores niveles de hipometilación en genes relacionados con el proceso de desarrollo, mientras que los cartílagos H mostraban menores niveles de metilación en genes asociados con procesos metabólicos y de regulación de la expresión génica y de procesos celulares. En definitiva, los autores, en su conclusión, barajan la hipótesis de que las modificaciones epigenéticas que tienen lugar en la artrosis y que afectan a factores clave como la apoptosis o las alteraciones metabólicas, son diferentes en función del fondo mitocondrial. Lo cual podría determinar, hasta cierto punto, la evolución de la enfermedad (Cortés-Pereira *et al.*, 2019)

Queda claro que la composición genética nuclear y mitocondrial juega un papel crucial en la regulación de la comunicación núcleo-mitocondria. Hasta el punto de que, tal y como se describió anteriormente, la presencia de determinados polimorfismos mitocondriales condiciona el papel que algunos polimorfismos nucleares tienen en la susceptibilidad a desarrollar diferentes patologías, y viceversa (Andrews *et al.*, 2019). A continuación, se describirán algunos ejemplos (Tabla 4):



Tabla 4. Tabla resumen de algunos ejemplos de interacciones genéticas nucleo-mitocondria que se asocian con diversas patologías humanas.

SNP nuclear	Alelo	SNP mitocondrial	Alelo	Haplogrupo	Patología	Riesgo	Referencia
PRS		-		K	Alzheimer	Aumentado	Andrews <i>et al.</i> 2019
				T		Aumentado	
				T		Reducido	
				V		Aumentado	
rs2107538	A	-		J	Esclerosis múltiple	Aumentado	Kozin <i>et al.</i> 2018
rs2114358	G						
rs1077667	C						
rs2243250	C						
rs1640923	A						
rs17320521	A	rs869096886	G	-	Cáncer de mama	Aumentado	Bushel <i>et al.</i> 2022
rs11666267	A	rs2854128	G	-	Parkinson	Aumentado	Pickett <i>et al.</i> 2022
rs254412	C	rs2853493	G				
rs384109	G	rs2853493	G				
rs441492	A	rs2853493	G				
rs1606610	G	rs2853493	G				
rs6493454	C	rs28357980	A	-	Degeneración macular	Reducido	Persad <i>et al.</i> 2017
rs11459118	GC	rs11459118	A			Aumentado	
rs11459118	GC	rs11459118	G			Aumentado	
rs144871045	A	rs11459118	A			Aumentado	
rs144871045	A	rs11459118	G			Aumentado	

PRS: polygenic risk score; SNP: single nucleotide polymorphism

En la patología del Alzheimer se realizó un estudio donde se analizaba la interacción entre los haplogrupos mitocondriales y una serie de polimorfismos nucleares asociados tradicionalmente con esta enfermedad, con el fin de analizar si dichas interacciones podían condicionar tanto el desarrollo de esta patología, como la edad de inicio (Andrews *et al.*, 2020). En dicho trabajo se calculó una puntuación en base a la presencia de hasta 19.630 SNPs nucleares asociados con el Alzheimer, y se analizó su interacción con la presencia de distintos haplogrupos mitocondriales. Los resultados de este estudio permitieron ver que, dependiendo del haplogrupo mitocondrial del sujeto, los efectos de los polimorfismos nucleares se alteraban. Las asociaciones más interesantes que encontraron fueron las siguientes: en primer lugar, observaron que la presencia del haplogrupo K o de los polimorfismos nucleares incrementaba, de forma independiente, el riesgo de desarrollar la enfermedad. Sin embargo, la presencia simultánea del haplogrupo K, o del haplogrupo T, y de los polimorfismos nucleares que forman parte de la puntuación nuclear reducía significativamente el riesgo de desarrollar la enfermedad. En segundo lugar, también vieron que la presencia simultánea del haplogrupo T y de determinados polimorfismos nucleares disminuía el riesgo de un inicio temprano de la enfermedad, mientras que la interacción de estos mismos polimorfismos con el haplogrupo V aumentaba dicho riesgo. Los autores recalcan que estos resultados sugieren que la epistasis entre los genomas nuclear y mitocondrial, en la que el efecto de un gen depende de la presencia de otro gen o conjunto de genes, influye en el Alzheimer. A pesar de no estar seguros de por qué puede ocurrir esto, una de las hipótesis que barajan está relacionada con las asociaciones entre la variación genética mitocondrial y la apolipoproteína E.

En el año 2018, Kozin y colaboradores (Kozin *et al.*, 2018) descubrieron una interacción entre una serie de variantes genéticas nucleares y mitocondriales capaz de modificar el riesgo de desarrollar esclerosis múltiple en una cohorte de individuos rusa. En su trabajo muestran que entre los individuos portadores del haplogrupo J, la incidencia de la esclerosis múltiple casi se duplicaba con respecto a otros haplogrupos, asociándose de forma significativa con el riesgo de desarrollar esta enfermedad. También analizaron de forma individual el efecto de diversos

## Introducción

polimorfismos nucleares con el riesgo de desarrollar esta patología, y tan solo un SNP, la variante intergénica rs1640923, se reveló de forma significativa como un factor de riesgo. Adicionalmente, realizaron un análisis donde combinaban la presencia de diversos polimorfismos nucleares (rs2107538 del gen CCL5 (del inglés *C-C Motif Chemokine Ligand 5*), rs2114358 del gen PVT1 (del inglés *Plasmacytoma Variant Translocation 1*), rs1077667 del gen TNFSF14 (del inglés *TNF Superfamily Member 14*), rs2243250 del gen IL4 (interleuquina 4) y la variante intergénica rs1640923) con el haplogrupo J. Todas las interacciones mostraron una asociación significativa con el riesgo de desarrollar esclerosis múltiple, y concretamente, la combinación del polimorfismo rs1640923 con el haplogrupo J generó un efecto aditivo por el cual el riesgo de desarrollar la enfermedad era mayor que el que otorgaba la presencia del SNP de forma individual. Una posible explicación sería que de por sí el haplogrupo J tiende a producir menores niveles de ATP, lo cual es preocupante en tejidos con alta demanda energética como es el sistema nervioso. Si a ello se le añaden los defectos en algunos de estos genes, el fenómeno sinérgico podría ser el desencadenante de la esclerosis múltiple. Por ejemplo, la interacción sinérgica entre el haplogrupo J y el alelo rs2107538\*A del gen CCL5. CCL5 desempeña un papel esencial en el metabolismo del ácido glutámico, el cuál es un neurotransmisor excitatorio. Los polimorfismos en este gen pueden alterar su función y afectar a los niveles de glutamato, favoreciendo el daño y la destrucción de neuronas, que se puede combinar con la menor capacidad energética del haplogrupo J.

En la patología del cáncer, se publicó en el año 2022 cómo la presencia simultánea de determinados polimorfismos genéticos en el ADN nuclear y en el ADNmt podía condicionar el fenotipo morfológico del cáncer de mama (Bushel *et al.*, 2022). Inicialmente, en este estudio trabajaron con muestras de tejidos teñidos con hematoxilina y eosina de mujeres con cáncer de mama (286) y de mujeres sanas (259). Encontraron un total de 115 interacciones entre mutaciones nucleares y mitocondriales en la cohorte de mujeres enfermas, mientras que no detectaron ninguna de estas interacciones entre las mujeres sanas. Con el objetivo de confirmar las interacciones, realizaron el mismo análisis en otra cohorte distinta compuesta por 523 muestras con cáncer de mama y, de las 115 interacciones inicialmente detectadas en la cohorte inicial, solamente una se mantuvo significativa. La interacción de interés ocurría entre el SNP nuclear rs17320521 del gen WSCD2 (del inglés *WSC domain containing 2*) y el SNP mitocondrial rs869096886 (m.11251A>G). Los resultados del estudio concluían que tener ambos polimorfismos incrementaba de forma significativa el tamaño medio de los núcleos celulares de la muestra en un 12%. Esta característica, junto con el tamaño y la cantidad de ADN dentro del núcleo, son algunas de los aspectos que analizan los patólogos para clasificar el nivel de severidad del cáncer de mama.

También, en el año 2021 se publicó un artículo en el que se analizaba cómo determinados factores, entre los cuales estaban presentes determinados polimorfismos nucleares, podían condicionar los niveles de heteroplasmia del ADNmt. Concretamente, en dicho estudio se genotiparon un total de 982.072 individuos y entre las asociaciones que encontraron, destacó una fuerte asociación entre elevados niveles de heteroplasmia en el ADNmt y el SNP nuclear rs1049342, que está situado muy próximo al gen TFAM, muy importante en la replicación y empaquetamiento del ADNmt. Esta interacción puede ser debida a la importancia que tiene este gen nuclear en la replicación del ADNmt y los altos niveles de heteroplasmia pueden condicionar el correcto funcionamiento de la mitocondria, alterando el riesgo de desarrollar diversas patologías (Nandakumar *et al.*, 2021).

Finalmente, existen casos de interacciones genéticas entre núcleo y mitocondria en el Parkinson. Un artículo reciente publicado en el año 2022 encontró cinco interacciones que se asociaban de forma significativa con el riesgo de desarrollar dicha patología. De estas cinco interacciones, una estaba presente entre el SNP mitocondrial rs2854128 (m.2706A>G) y el nuclear rs11666267 del gen NIBAN3 (del inglés *Niban Apoptosis Regulator 3*), mientras que el resto de interacciones estaban presentes entre el SNP mitocondrial rs2853493 (m.11467A>G) y los SNPs nucleares rs254412, rs384109 y rs441492 presentes en el gen TBCA (del inglés *Tubulin Folding Cofactor A*), y también con la variante nuclear intergénica rs1606610 (Pickett *et al.*, 2022). Los dos polimorfismos mitocondriales de interés m.2706 y m.11467 definen los haplogrupos H y U respectivamente, los cuales fueron previamente asociados con el riesgo de desarrollar Parkinson (Hudson *et al.*, 2013; Marom *et al.*, 2017); por su parte, los polimorfismos nucleares condicionan el riesgo de desarrollar diversas patologías neurológicas, por lo que probablemente exista un efecto aditivo entre las variantes nucleares y mitocondriales que promueve el desarrollo de la patología del Parkinson (Pickett *et al.*, 2022).

Sin embargo, un gran problema que tienen este tipo de estudios es el tamaño muestral de la población de estudio. Esto es debido a la estratificación de la población analizada, ya que, al analizar el efecto de un determinado polimorfismo nuclear en presencia de un haplogrupo o variante mitocondrial específica, el tamaño muestral de cada grupo se ve reducido. Como resultado, se reduce también el poder estadístico de los análisis, pudiendo obtenerse interacciones que en realidad son falsos positivos. Por lo tanto, para realizar estos análisis es necesario disponer de un elevado tamaño muestral. Un ejemplo es el trabajo de Persad y colaboradores de 2017 (Persad *et al.*, 2017), donde tratan de analizar si la presencia simultánea de variantes genéticas nucleares y mitocondriales puede modificar el riesgo de desarrollar la degeneración macular asociada a la edad. En este estudio, los autores analizaron un total de 17.832 controles y 16.144 casos. Como resultado, identificaron tres interacciones de interés: la primera ocurría entre la posición mitocondrial m.4917A>G (rs28357980) y el SNP nuclear rs6493454 de TRPM1 (del inglés *Transient Receptor Potential Cation Channel Subfamily M Member 1*). La presencia simultánea del alelo A de rs28357980 y el alelo C del SNP rs6493454 se asociaba de forma significativa con un menor riesgo de desarrollar la patología. La segunda y la tercera interacción ocurría entre los SNPs rs11459118 y rs144871045, de ABHD2 (del inglés *Abhydrolase Domain Containing 2, Acylglycerol Lipase*) y RLBP1 (del inglés *Retinaldehyde Binding Protein 1*) respectivamente, con la posición mitocondrial m.12771A>G (rs11459118). La presencia de ambos SNPs nucleares se asociaba significativamente con un riesgo aumentado de desarrollar la patología en presencia del alelo m.12771A, mientras que en presencia del alelo m.12771G las interacciones seguían mostrándose como factores de riesgo pero con un efecto mucho menor (Persad *et al.*, 2017).

En lo que respecta a la artrosis, a día de hoy no se ha publicado ningún trabajo que estudie las interacciones entre el genoma nuclear y el mitocondrial que puedan condicionar la susceptibilidad o progresión de la patología, ni de sus fenotipos clínicos. Como se ha comentado a lo largo de esta introducción, múltiples estudios han puesto de manifiesto que, de forma independiente, tanto polimorfismos nucleares como mitocondriales se asocian con distintos aspectos de esta enfermedad, por lo que es altamente probable que la coexistencia de estos polimorfismos presentes en los dos genomas pueda también tener un efecto crucial en esta patología.

## Técnicas de secuenciación para analizar el genoma.

Además de los arrays de SNPs mencionados en el apartado de estudios GWAS, existen otras formas ampliamente utilizadas para el análisis del genoma, que son las técnicas de secuenciación. Estas técnicas permiten conocer la secuencia, u orden de nucleótidos de uno o múltiples segmentos de ADN.

Todo el conocimiento que se tiene a día de hoy sobre el genoma humano ha sido gracias a estudios internacionales cimentados sobre uno de los grandes hallazgos en este campo, la obtención de la secuencia completa de las  $3 \cdot 10^9$  pb del genoma humano. Este proyecto, denominado como *Proyecto Genoma Humano*, fue en el año 1990 un proyecto muy ambicioso. Debido a las limitaciones de su tiempo, costó un total de 3.000 millones de dólares y participaron en él un total de 20 centros de investigación. Durante los inicios de este proyecto se empleó una tecnología de secuenciación bastante rudimentaria y costosa, la cual permitía secuenciar fragmentos de una longitud aproximada de entre 600 y 800 pb. Con el paso de los años esta metodología fue actualizándose, permitiendo la automatización del proceso, el aumento de la velocidad, una mejora de la calidad de los datos, junto con una reducción de los costes asociados (Hood y Rowen, 2013). Gracias a todo este esfuerzo colaborativo, en el año 2001 se obtuvo la primera secuencia de referencia del genoma humano que, posteriormente en el año 2004, fue revisada y mejorada (Abdellah *et al.*, 2004). Este ambicioso proyecto ha revolucionado los campos de la biología y la medicina, permitiendo el desarrollo de la medicina personalizada gracias al diagnóstico molecular. Asimismo, tal y como se comentó anteriormente, también sirvió para impulsar el avance de las tecnologías de secuenciación con el paso del tiempo. En líneas generales, las diferentes tecnologías de secuenciación se pueden clasificar en las siguientes generaciones:

### *Secuenciación de primera generación.*

Existieron diversos métodos de secuenciación inicialmente, uno bastante conocido fue el propuesto por Maxam y Gilbert en 1977 (Maxam y Gilbert, 1977). Este método consistía en producir roturas de ADN químicamente al tratar la muestra de ADN con diversos productos químicos. De esta forma se empleaba dimetilsulfuro para eliminar solamente las guaninas, dimetilsulfuro y ácido fórmico para eliminar las adeninas y las guaninas, hidracina para eliminar las citosinas y las timinas, e hidracina y cloruro sódico para eliminar solo las citosinas. Después, al correr el producto resultante en un gel de electroforesis se podía leer la secuencia obtenida (Fig. 29). El problema que tenía este método de secuenciación es que empleaba compuestos químicos peligrosos, era técnicamente complejo y difícil de escalar.

En ese mismo año también surgió otro método de secuenciación, el cual fue el primer método ampliamente utilizado, y es la denominada secuenciación Sanger, la cual está basada en los terminadores de cadena (Sanger *et al.*, 1977). Para realizar esta técnica es necesario tener un ADN molde, un oligonucleótido que hibride en la región de interés del ADN molde, una enzima polimerasa, dNTPs y didesoxinucleótidos (ddNTPs) terminadores. Este último elemento se diferencia de los dNTPs en que carecen del OH en el carbono 3', de modo que en el momento en el que este ddNTP es incorporado por la *Taq polimerasa* a la cadena molde, se interrumpe la elongación. Inicialmente esta técnica se realizaba de la siguiente forma:

1. Se crean cuatro mezclas de secuenciación distintas para cada uno de los cuatro ddNTPs (adenina, timina, citosina o guanina) marcados radiactivamente; además,

- cada una de las mezclas incluía también el oligonucleótido que hibrida en la zona que se desea secuenciar, la enzima polimerasa y los dNTPs.
- Se introduce el mismo ADN molde en cada uno de los cuatro tubos y se aplican las condiciones de temperatura necesarias para que la polimerasa pueda funcionar. Durante este proceso, los dNTPs se irán uniendo a la secuencia creada, y de forma aleatoria, se incorporarán los ddNTPs que, al carecer del grupo hidroxilo, no pueden formar el enlace con el siguiente nucleótido, terminando así la secuencia y quedando ésta marcada radiactivamente en su extremo 3'. Como resultado de este proceso, en cada uno de los cuatro tubos, se producirán todos los posibles tamaños de fragmentos de ADN interrumpidos por los ddNTPs correspondientes al tubo.
  - Finalmente, los cuatro productos resultantes de este proceso son sometidos a un proceso de electroforesis en un gel de poliacrilamida. Se introducen de forma separada cada uno de los cuatro productos en sus respectivos pocillos del gel. Debido a que el ADN está cargado negativamente por la presencia de los grupos fosfato, tras someter a la muestra a un campo eléctrico, los fragmentos de cada una de las reacciones migrarán hacia el polo positivo. Una vez finalizada, se revela el resultado del gel mediante una radiografía, pudiéndose observar las bandas gracias al marcaje radioactivo de los ddNTPs presentes en los extremos 3' de los fragmentos. Se podrá saber la secuencia leyendo las bandas desde el polo positivo, donde están las secuencias más pequeñas, hacia el polo negativo, donde estarán las secuencias de mayor tamaño.

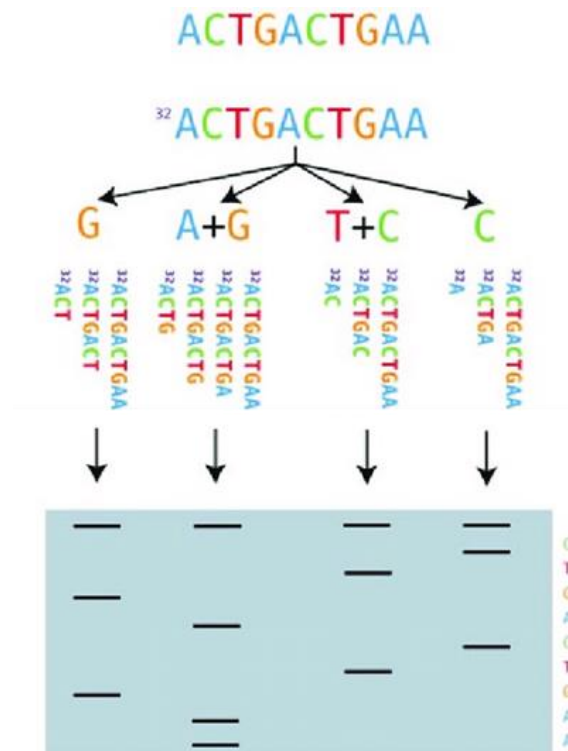


Figura 29. Esquema representativo del proceso de secuenciación química de Maxam y Gilbert. Dependiendo del tratamiento químico realizado se producirán roturas aleatorias de distintos nucleótidos. Al someter los cuatro productos resultantes en una electroforesis capilar podremos conocer la secuencia leyendo las bandas desde abajo hacia arriba. Imagen modificada de [https://www.researchgate.net/figure/An-example-of-Maxam-Gilbert-Sequencing-Technique-showing-specific-cleavage-of-DNA\\_fig3\\_335867452](https://www.researchgate.net/figure/An-example-of-Maxam-Gilbert-Sequencing-Technique-showing-specific-cleavage-of-DNA_fig3_335867452).

## Introducción

Esta técnica demostró una gran robustez, precisión y sencillez, lo cual hizo que se popularizara en su época. El amplio uso de esta metodología trajo consigo una serie de actualizaciones: en primer lugar, el cambio del marcaje radioactivo de los ddNTPs por un marcaje con fluorocromos y, a partir de este punto, cada uno de los cuatro posibles ddNTPs tenía un marcaje específico según la base nitrogenada presente, haciendo posible llevar a cabo la reacción de secuenciación en un único tubo en vez de cuatro. Sin embargo, para detectar este marcaje fue necesario substituir los geles de acrilamida por sistemas de electroforesis capilar (Fig. 30). La segunda gran mejora en esta técnica fue la automatización del proceso mediante secuenciadores automáticos conectados a ordenadores personales. Otros descubrimientos en el campo de la biología molecular, llevados a cabo por esta época, contribuyeron a mejorar la calidad de las secuencias obtenidas por este proceso; es el caso de la PCR (Mullis *et al.*, 1986; Saiki *et al.*, 1988), cuya aportación a las técnicas de secuenciación permitió obtener muestras de ADN mucho más concentradas y de mayor calidad, mejorando sustancialmente el resultado final.

La secuenciación Sanger permite obtener secuencias de un tamaño máximo aproximado de 1.000 pb, por lo que, para analizar fragmentos de mayor tamaño, se desarrollaron técnicas como la secuenciación *shotgun*. Este procedimiento consiste en la secuenciación de fragmentos superpuestos de forma separada, los cuales son posteriormente ensamblados por las regiones compartidas, permitiendo obtener secuencias más largas (Heather y Chain, 2016).

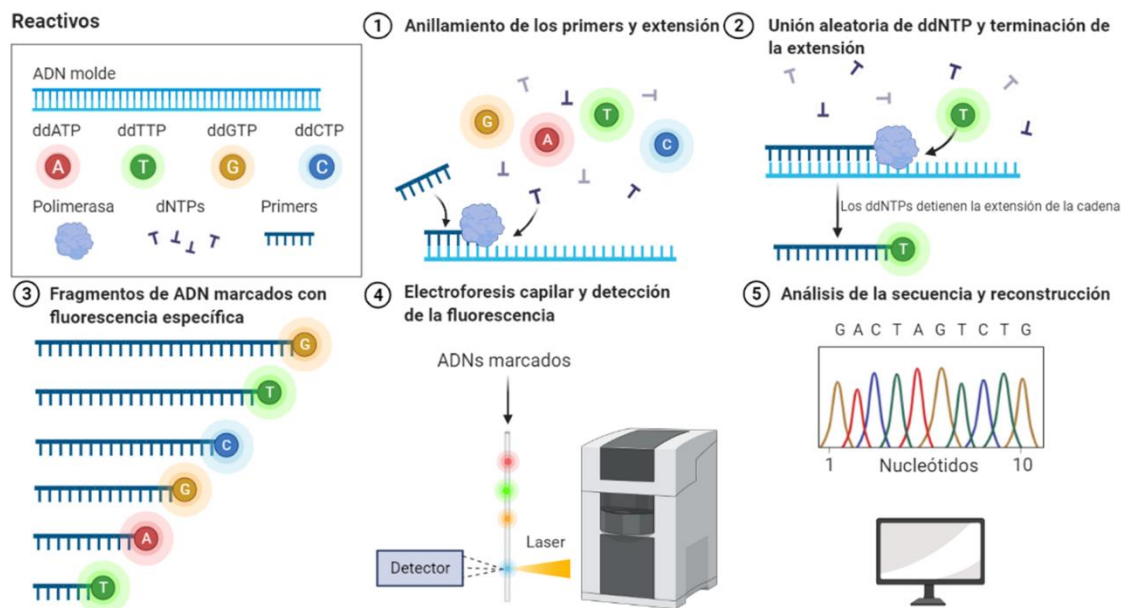


Figura 30. Esquema de los pasos necesarios para realizar un proceso de secuenciación Sanger mediante electroforesis capilar. Imagen creada en [app.biorender.com](https://www.biorender.com).

### Secuenciación de segunda generación.

En la época de los 90 se desarrolló la pirosecuenciación (Nyrén, 2015). Esta metodología cambió el paradigma de la secuenciación ya que, en vez de observar el marcaje de los fragmentos una vez finalizada la reacción de secuenciación, se podía seguir la actividad de la polimerasa durante la incorporación de los nucleótidos a la cadena de ADN, concepto que se conoce como secuenciación por síntesis. Durante la pirosecuenciación, cada vez que un dNTP se une a la cadena de ADN, este liberará un grupo pirofosfato (ppi) que desencadena una serie de reacciones enzimáticas produciendo finalmente una señal lumínica (Harrington *et al.*, 2013). Al

mismo tiempo surgió la PCR de emulsión, donde todo el proceso de PCR ocurre dentro de unos microcompartimentos formados por emulsiones de agua en aceite; de la unión de estas dos tecnologías, la pirosecuenciación y la PCR de emulsión, surgió en el año 2000 la empresa 454 Life Sciences, la cual creó la primera plataforma de *next-generation sequencing* (NGS) del mercado, conocido como el GS 20. Esta máquina era capaz de realizar una secuenciación *shotgun* de pequeños genomas como el de *Mycoplasma genitalium* (580.069 pb) con un 96% de cobertura y un 99,96% de precisión en un único experimento.

Con el paso de los años, las plataformas de NGS continuaron evolucionando y se centraron en la secuenciación paralela y cíclica de múltiples amplicones clonales separados espacialmente, lo cual permitía generar sobre 500.000.000 pb de secuencia en bruto o billones de bases en un único experimento. Todas estas tecnologías tienen en común cuatro pasos fundamentales en su metodología (Head *et al.*, 2014):

1. Preparación de las librerías: una librería es el conjunto de todos los fragmentos de ADN de una muestra, estando cada uno de estos fragmentos delimitado en sus extremos por unas secuencias conocidas denominadas adaptadores. Para realizar este paso lo primero es extraer los ácidos nucleicos de la muestra de interés, y realizar un análisis de su concentración y calidad antes de proceder con los siguientes pasos. Este es un paso sencillo pero crítico, ya que de la precisa cuantificación y calidad del material de partida depende en gran medida el éxito de la secuenciación. En segundo lugar, para generar las librerías, es necesario realizar una fragmentación inicial de la muestra de ADN mediante sonicación o tratamiento enzimático. A estos fragmentos es necesario ligarles a continuación los adaptadores, de modo que cada secuencia de ADN flanqueada por adaptadores se denomina amplicón, y el conjunto de amplicones de una muestra constituye la librería. Otra forma de realizar este paso consiste en utilizar conjuntos de oligonucleótidos capaces de amplificar de forma solapada los fragmentos de interés de una secuencia diana (ya sea un panel de genes, un exoma o un genoma completo) en vez de realizar la fragmentación de la muestra. Esta tecnología se conoce como *ampliseq*.
2. Enriquecimiento: una vez obtenidas las librerías mediante una metodología u otra, se realizará un proceso de selección y enriquecimiento de las mismas para sobrerrepresentar el producto deseado frente a los artefactos. Tras el enriquecimiento se realiza un proceso de lavado para eliminar dichos artefactos y quedarnos sólo con los amplicones con buena calidad. Finalmente, antes de la secuenciación se suele llevar a cabo una amplificación clonal de los fragmentos de interés mediante técnicas de PCR de emulsión.
3. Secuenciación: el fundamento de la secuenciación es específico de la tecnología empleada, pero todas ellas comparten una elevada capacidad de trabajo de múltiples secuencias en paralelo.
4. Análisis de datos: finalmente, la cantidad de datos generados mediante estas plataformas es ingente, por lo que es necesario un extenso tratamiento de los mismos con el objetivo de eliminar el ruido, filtrar las secuencias en base a unos criterios determinados de calidad y poder analizar las secuencias correctamente.

En esta generación surgieron distintas plataformas de secuenciación, cada una con su propio fundamento (Heather y Chain, 2016):

1. *Illumina*: una de las peculiaridades de esta tecnología es su proceso de amplificación, denominado *bridge amplification* (Fig. 31). En este proceso se va a emplear un soporte

## Introducción

físico denominado célula de flujo, el cual posee en su superficie secuencias complementarias a las presentes en los adaptadores de los amplicones. Como resultado de este proceso de amplificación clonal se generarán clústers de amplicones.

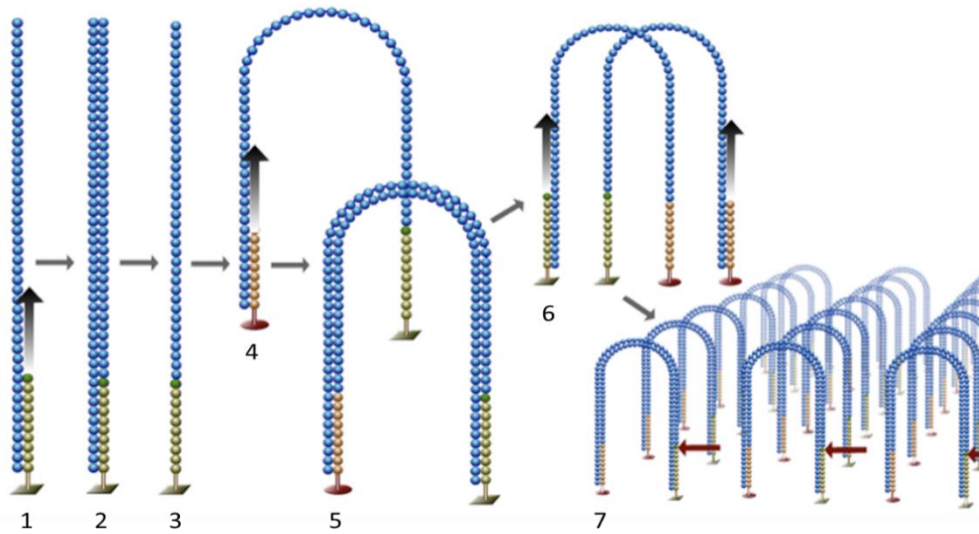


Figura 321. Proceso de bridge amplification de Illumina. 1: inicialmente se unen fragmentos monocatenarios a los adaptadores complementarios presentes en la célula de flujo; 2: tras la unión se produce la elongación de la otra cadena del fragmento, siendo esta cadena la que quedará anclada de forma definitiva a la célula de flujo; 3: se elimina la cadena inicial que no está unida a la célula de flujo; 4: el otro extremo del fragmento monocatenario se dobla y se une a otros adaptadores cercanos complementarios, formando la típica estructura de puente; 5: se produce la elongación de la otra cadena, cada una de las cadenas está unida a un punto distinto de la célula de flujo; 6: se rompe la unión complementaria de las dos cadenas y éstas se vuelven a unir a otros adaptadores libres cercanos; 7: este proceso se repetirá cíclicamente generando zonas ricas en fragmentos denominados clústers de amplificación. Imagen obtenida de <http://zon.trilinkbiotech.com/2019/12/10/polymerase-chain-reaction-pcr-day-december-20th/bridge-2>.

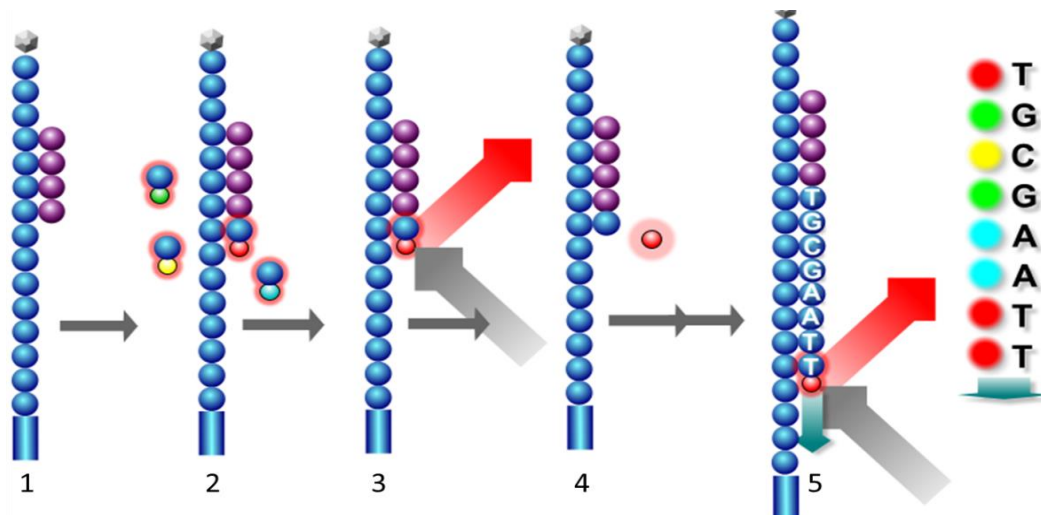


Figura 312 . Proceso de secuenciación por síntesis de Illumina. 1: el oligonucleótido de secuenciación se une a la región diana de los amplicones; 2: solamente un ddNTP marcado se va a unir al nucleótido siguiente al oligonucleótido y la elongación se detiene temporalmente debido a que estos ddNTPs poseen un elemento terminador que puede ser eliminado y que impide el enlace fosfodiéster; 3: en este momento se excita el ddNTP unido con un láser; 4: el láser produce el desbloqueo del ddNTP y la liberación de la molécula fluorescente, la cual es detectada por una cámara y se registra el nucleótido que se haya unido; 5: el desbloqueo del ddNTP permite la unión de un siguiente ddNTP bloqueado, repitiéndose el ciclo unas 300 veces y obteniéndose finalmente la secuencia del fragmento. Imagen obtenida de <https://pediaa.com/how-does-illumina-sequencing-work>.



El proceso de secuenciación está resumido en la figura 32 y en él intervienen ddNTPs terminadores de cadena reversibles marcados con fluorescencia, es decir, poseen el grupo hidroxilo en el carbono 3' pero tienen un elemento terminador que puede ser eliminado que impide el enlace fosfodiéster. En esta metodología de secuenciación se va a leer la secuencia tanto en un sentido como en el otro, característica que se conoce como *paired-end* (Ravi *et al.*, 2018) y, permite evitar la pérdida de calidad de las secuencias debido al *phasing* (Tewhey *et al.*, 2011), fenómeno por el cual conforme avanza la secuenciación se pierde calidad de la secuencia debido a que las moléculas del mismo clúster pierden sincronía entre ellas. Además, el hecho de que los nucleótidos se unan de uno en uno reduce la presencia de errores en zonas de homopolímeros, problema muy habitual en la tecnología de pirosecuenciación.

2. *Ion Torrent*: esta tecnología consiste en un proceso de secuenciación por síntesis, sin embargo, no se utilizan ddNTPs marcados con fluorescencia. En el caso de esta metodología el fundamento es muy similar a la pirosecuenciación, solo que, en vez de observar una reacción lumínica, se observará un cambio de pH que produce un cambio de voltaje, de ahí que a este tipo de secuenciadores se los denomine “semiconductores”. La preparación de librerías se puede llevar a cabo por fragmentación o bien mediante tecnología *ampliseq*, generándose así directamente amplicones de un tamaño conocido de las secuencias de interés. Una vez obtenidos los amplicones, es necesario realizar la limpieza, enriquecimiento y amplificación de los fragmentos de interés mediante una PCR de emulsión, quedando miles de copias del amplicón unidos a una misma esfera. El siguiente paso será ya la secuenciación. Al igual que ocurre en la tecnología *Illumina*, es necesario usar un soporte físico, en este caso un Chip formado por millones de micropocillos donde solo cabe una esfera con amplicones unidos a su superficie. Cada uno de estos micropocillos tiene en su base un medidor capaz de detectar los cambios de voltaje debidos a las modificaciones en el pH del medio. El proceso de secuenciación es cíclico, incorporándose un único dNTP a la muestra en un orden específico. En cada ciclo se unirá a la secuencia molde el número de nucleótidos correspondiente, y cada nucleótido que se une liberará un H<sup>+</sup>, acidificando el medio y produciendo un cambio de voltaje. El cambio de voltaje será proporcional al número de bases incorporadas, y esto será detectado por el medidor en la base del pocillo. Una vez finalizado un ciclo, se produce un lavado para retirar el exceso de dNTPs, y se comienza el siguiente, introduciendo un nuevo tipo de dNTP. El proceso se repetirá un número determinado de ciclos (Fig. 33). A diferencia de *Illumina*, esta tecnología carece de metodología *paired-end* y, además, no se incorporan los nucleótidos de uno en uno, ya que, dependiendo del número de repeticiones de la secuencia, pueden incorporarse un elevado número de bases idénticas a la vez. Una de las grandes ventajas de esta tecnología es que no necesita ningún tipo de marcaje, lo cual abarata mucho el coste y reduce los tiempos de secuenciación, haciéndola más accesible para todo el mundo (Schadt *et al.*, 2010). Debido a que esta es la metodología de secuenciación que se empleó en el desarrollo de esta tesis doctoral, se explicará con más detalle en la sección de material y métodos.

3.

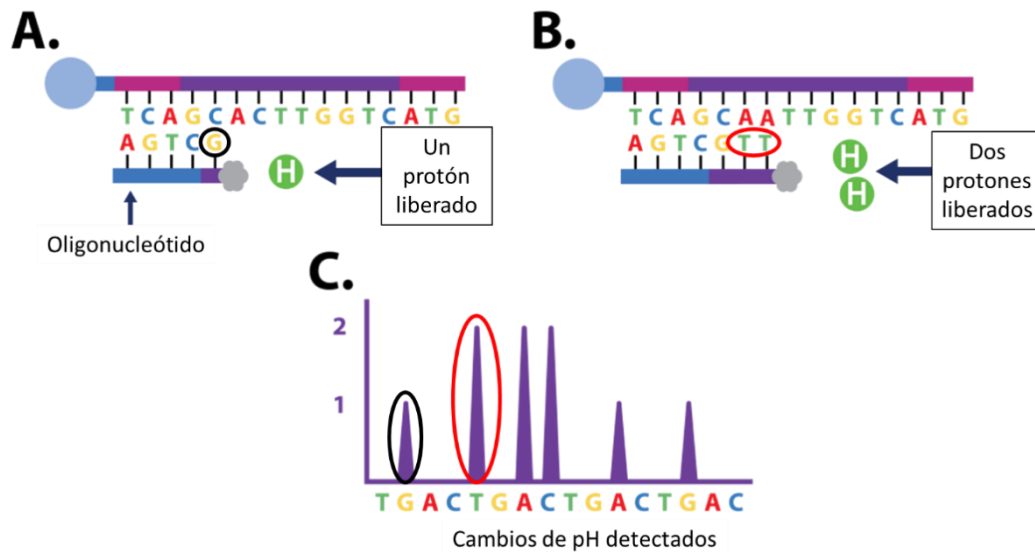


Figura 33. . Esquema representativo del proceso de secuenciación de Ion Torrent. A. Tras la unión del oligonucleótido al fragmento de interés unido a la esfera comienzan a introducirse dNTPs de forma secuencial (T, G, A y C). Cuando por complementariedad de bases se produzca la unión de un dNTP, se liberará un protón que acidificará el medio y será detectado. B. Puede ocurrir que se una más de un dNTP a la secuencia durante un ciclo, en este caso se liberan tantos protones como número de dNTPs unidos, en este ejemplo debido a la unión de dos timinas se liberan dos protones al medio. C. Durante todo este proceso se anotan los cambios de pH que sufre el medio, la intensidad del cambio será proporcional al número de dNTPs. Con este gráfico podemos conocer la secuencia al leer el eje X de izquierda a derecha, y el eje Y nos permitirá saber cuántos dNTPs se han unido en cada ciclo. Imagen modificada de <https://apollo-institute.org/2021/11/24/ion-torrent-sequencing/>.

4. SOLiD (del inglés *Sequencing by Oligonucleotide Ligation and Detection*): a diferencia de las otras tecnologías, las cuales están basadas en el principio de secuenciación por síntesis, es decir, realizan el proceso de secuenciación observando moléculas liberadas durante la unión de nucleótidos a la secuencia (señales lumínicas, colorimétricas o cambios de pH), la plataforma SOLiD se basan en el principio de secuenciación por ligación (Thudi *et al.*, 2012), donde se observa la señal previamente a su liberación. Una vez generadas las librerías y llevado a cabo el enriquecimiento, comienza la fase de secuenciación, cuyo fundamento consiste en el uso de 16 sondas distintas de ocho nucleótidos, seis de los cuales son inespecíficos, mientras que los otros dos son específicos de la secuencia. Las sondas presentan un marcaje de color con un patrón específico de la casa comercial. Cuando se inicia el proceso de secuenciación se usan unos *oligonucleótidos* que se unen a los adaptadores presentes en el fragmento de interés. Dependiendo de las dos bases siguientes al *oligonucleótido*, se unirá específicamente una sonda marcada, y en este momento se capta la información lumínica. Tras esto se produce el corte del fluoróforo, un lavado y se une la siguiente sonda mediante una ligasa. Este proceso se repetirá un total de 7 veces, y se conoce como "ronda de secuenciación". Una vez finalizada una ronda de secuenciación, se usará un *oligonucleótido* de adaptador distinto, que se unirá específicamente a una posición con una diferencia de un único pb con respecto a la primera ronda, y se llevará a cabo la ronda de secuenciación. En total se realizan cinco rondas de secuenciación completas (Fig. 34). Con esta aproximación, en total cada nucleótido de la secuencia es analizado dos veces, mejorando así la calidad de los datos. Con este proceso se puede

conseguir más de 100 gigabases de información por carrera de secuenciación, pero es un proceso lento que puede llevar hasta 6 días.

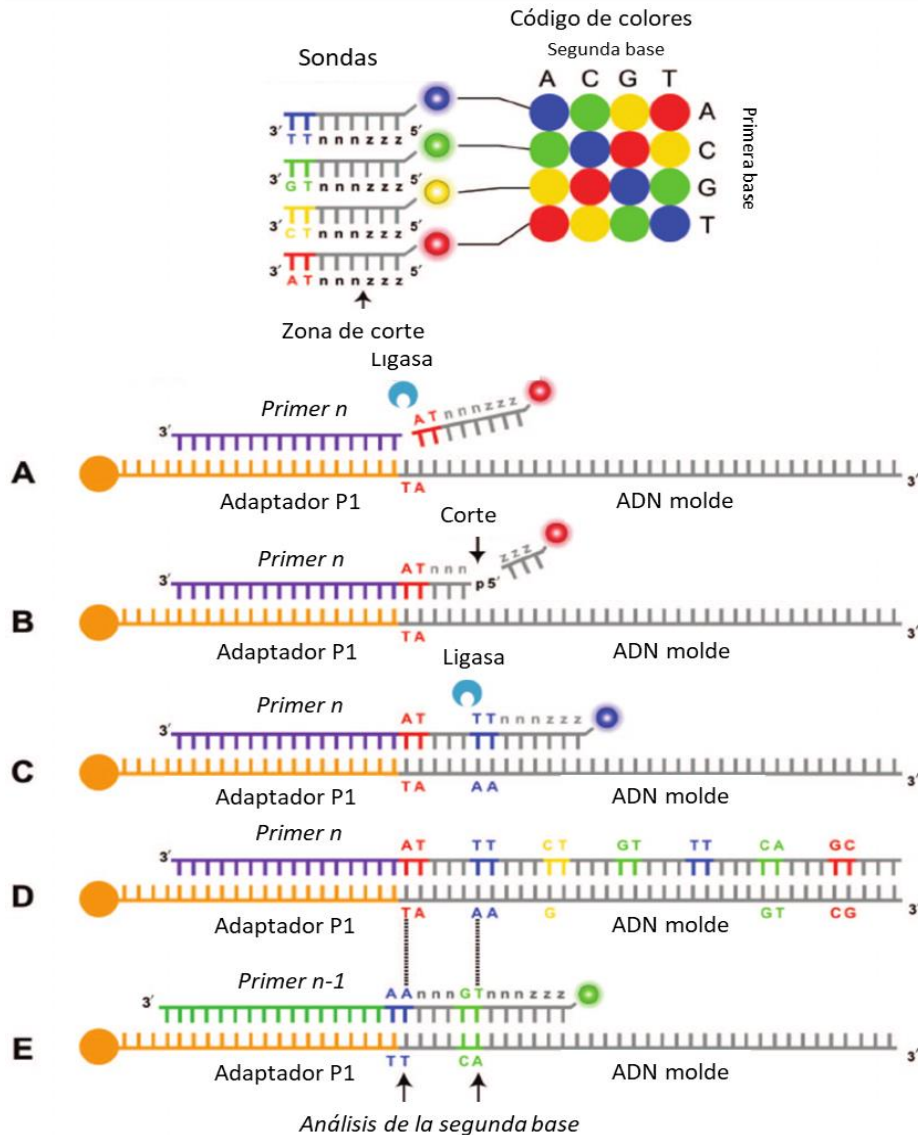


Figura 34. Proceso de secuenciación mediante tecnología SOLiD. El código de colores presente en la sonda dependerá de las dos primeras bases que porte ésta. A: el primer oligonucleótido se une a la región del Adaptador P1, y con la ayuda de una ligasa se unirá la sonda correspondiente a las dos bases siguientes al oligonucleótido, es en este momento cuando se detecta la señal lumínica; B: tras la unión de las 3 bases inespecíficas se produce el corte de la sonda, liberándose así la molécula fluorescente; C: una nueva sonda se une con la ayuda de una ligasa; D: el proceso se repite hasta un total de 7 veces completando una ronda de secuenciación; E: se inicia una nueva ronda de secuenciación empleando un oligonucleótido que ancla colindante al oligonucleótido empleado en la ronda anterior y se repite todo el proceso. Imagen modificada de [https://www.researchgate.net/figure/Applied-Biosystems-SOLiD-sequencing-by-ligation-Top-SOLiD-color-space-coding-Each\\_fig3\\_24043867](https://www.researchgate.net/figure/Applied-Biosystems-SOLiD-sequencing-by-ligation-Top-SOLiD-color-space-coding-Each_fig3_24043867).

### Secuenciación de tercera generación.

Uno de los grandes problemas de la segunda generación es que generan lecturas de tamaños relativamente cortos, lo cual puede llevar a fallos en el ensamblaje o incluso vacíos entre secuencias. Y a pesar de ser muy útiles para detectar variantes de pequeño tamaño como SNPs o pequeñas inserciones y deleciones, las variantes de mayor tamaño son más difíciles de detectar (Slatko *et al.*, 2018; van Dijk *et al.*, 2018). Además, todos los métodos de secuenciación

## Introducción

de segunda generación necesitan un paso de amplificación previo a la secuenciación para facilitar la detección de las señales por los detectores, y este proceso puede promover errores de secuencia o favorecer determinadas secuencias frente a otras, cambiando la frecuencia relativa y la cantidad de unos fragmentos de ADN frente a otros (Pareek *et al.*, 2011).

En la tercera generación, las compañías que han diseñado estas metodologías se han centrado, por un lado, en poder realizar la secuenciación desde la muestra de ADN original sin necesidad de PCR previa y, por otro lado, en generar lecturas de gran tamaño. Actualmente existen dos plataformas de secuenciación de tercera generación (van Dijk *et al.*, 2018):

1. PacBio: estos secuenciadores fueron diseñados por la compañía *Pacific Biosciences* (EEUU). Se basan en el uso de la tecnología *single molecule real time sequencing* o SMRT, y de media generan lecturas de aproximadamente 20.000 pb llegando a máximos de 50.000 pb. La generación de librerías es distinta a la del resto de tecnologías mencionadas y necesita un ADN de partida de elevada calidad. A diferencia de las otras tecnologías, los amplicones generados son bicatenarios y los adaptadores de sus extremos tienen forma de *loop*, lo que le da a los amplicones una morfología circular característica. El proceso tiene lugar sobre un soporte físico, que es una célula de flujo con miles de pocillos denominados *zero mode waveguide* (ZMW). En el fondo de cada uno de los pocillos está anclada una única molécula de enzima polimerasa mediante biotina y estreptavidina. Además, para el proceso de secuenciación se emplean dNTPs fosfoligados, es decir, en vez de tener la fluorescencia anclada a la base nitrogenada presente en el carbono 1 del nucleótido, la molécula fluorescente está unida a los grupos fosfatos presentes en el carbono 5 del nucleótido.

Una vez que se carga la célula de flujo con la muestra comienza el proceso de secuenciación, de modo que solamente un amplicón va a entrar en un pocillo ZMW, e interactuará con la polimerasa en el fondo del pocillo. Esta enzima va a comenzar a secuenciar de forma cíclica y repetida la molécula circular, secuenciando ambas cadenas en cada uno de los procesos. Cada vuelta completa de la secuencia se denomina pase. Cada vez que se une un dNTP fosfoligado, se liberará un grupo pirofosfato marcado, que al ser incidido por un haz de luz emitirá una fluorescencia específica según la base que se haya unido. La captación de la fluorescencia ocurre *in vivo* en un proceso que denominan película. En dicha película se registra qué nucleótido se une y también el tiempo entre uniones, lo cual se conoce como duración interpulsos. Una particularidad de esta tecnología es que puede detectar modificaciones epigenéticas, ya que éstas son capaces de alterar la duración interpulsos (Fig. 35).

Los múltiples pases producidos durante el proceso generan una secuencia consenso circular, cuya precisión y fidelidad será mayor cuantos más pases se hayan dado. Se estima que, en 25 pases, la precisión alcanza valores de 99,999%, y en 50 pases 99,9999%. Además, como la polimerasa incorpora los nucleótidos de uno en uno, no tienen problemas de homopolímeros. Sin embargo, esta metodología genera muy pocas lecturas por carrera y tiene un elevado coste.

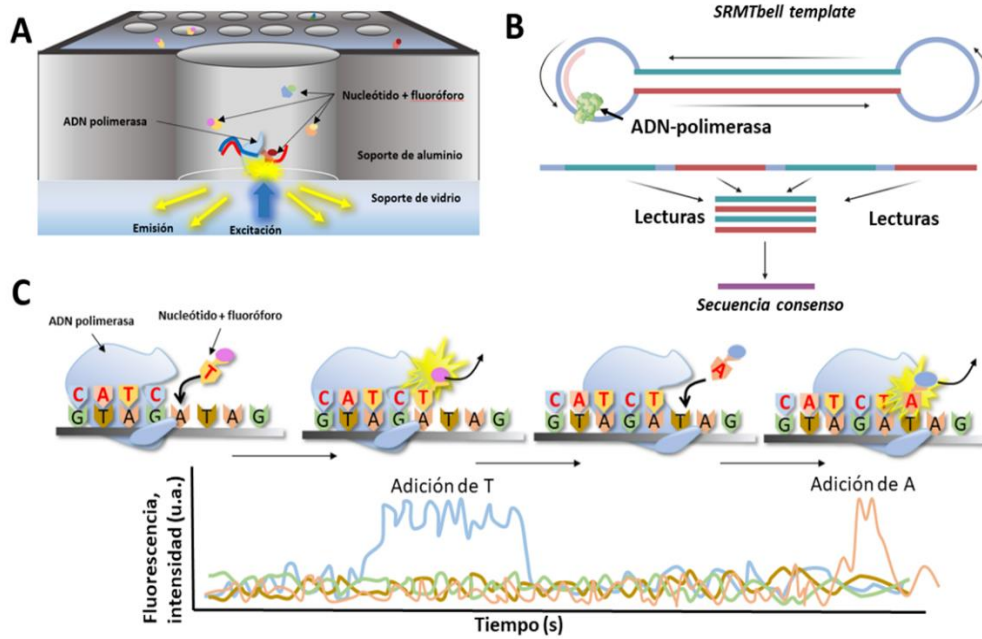


Figura 35. Proceso de secuenciación en PacBio. A: esquema representativo de los pocillos ZMW de la tecnología SMRT; en la base de dicho pocillo hay una molécula de polimerasa anclada, tras incorporar la muestra, solamente un amplicón entrará a cada uno de los pocillos y quedará unido a la polimerasa. Durante el proceso de secuenciación cada vez que se incorpore un nucleótido se liberará una molécula fluorescente que, tras ser excitada con un haz de luz, emitirá una fluorescencia específica según qué base nucleotídica se haya unido; B: representación de la estructura típica de un amplicón utilizado en esta metodología, la polimerasa presente en el pocillo ZMW se unirá a uno de los adaptadores y comenzará el proceso de secuenciación produciendo lecturas cíclicas del amplicón. El conjunto de las lecturas que se obtiene dará como resultado la secuencia consenso; C: película de secuenciación, durante la incorporación de los nucleótidos en el proceso de secuenciación se liberan los fluoróforos específicos, esto emite un pulso de fluorescencia que es detectado y se registra tanto el nucleótido como el momento en el que pasó. Imagen obtenida de (Valderrama, Ortigosa and Cañas, 2020)

2. Nanopore: esta metodología ha sido desarrollada por la compañía Oxford Nanopore (Reino Unido). Tiene un principio similar a la de IonTorrent, ya que se basa en mediciones de cambios de voltaje; sin embargo, la secuenciación se hace a través de un poro de membrana y no se tienen en cuenta los cambios de pH. Del mismo modo que ocurre en PacBio, se generan fragmentos de gran tamaño, pero el tamaño de los fragmentos no está limitado por la química que emplean, sino por el tamaño de la molécula a secuenciar, llegando a conseguir amplicones de hasta 1 megabase (Jain *et al.*, 2018).

Para generar las librerías en esta metodología es necesario un ADN de partida de elevada calidad; asimismo, cabe destacar que los fragmentos de éstas no son romos, y el inserto está rodeado por adaptadores con distintos tamaños, poseyendo uno de ellos una proteína motora al final de su secuencia. El soporte físico consiste en una célula de flujo con dos compartimentos rellenos de soluciones iónicas separados por una membrana con un gran número de nanoporos incorporados. Tras cargar la célula de flujo con las librerías, la proteína motora desplaza a los amplicones hasta los nanoporos, y al aplicar corriente eléctrica estos migran a través del poro junto con dicha corriente. Cada nucleótido que pasa a través del poro modifica la corriente eléctrica de forma específica, y esta información es recolectada por un detector (Fig. 36). Esta tecnología presenta una serie de ventajas e inconvenientes. Entre las ventajas destaca su bajo coste, el escaso grado de requerimiento de reactivos, su ausencia de *phasing* y que no

## Introducción

sufre la pérdida de la actividad enzimática, que es el principal factor limitante de los tamaños de los amplicones de PacBio. Además, tiene un tamaño muy pequeño y consta de kits comerciales que permiten la preparación de librerías en tiempos muy cortos, de entre 10 y 15 minutos, lo cual lo hace ideal para usarlo directamente en muestreos en campo abierto. Por otro lado, presenta un gran inconveniente, y es que cada molécula de ADN solamente se lee una vez, lo que genera una tasa de error muy elevada. Actualmente se está trabajando en tratar de reducir la probabilidad de artefactos.

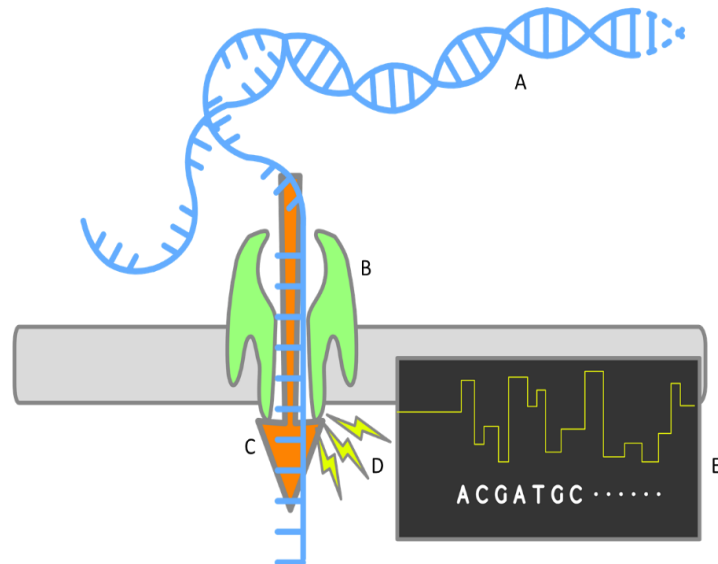


Figura 36. Esquema representativo del proceso de secuenciación en la metodología de Nanopore A: ADN molde, B: célula de flujo con el nanoporo; C: proteína motora: está unida a una de las cadenas del ADN y la transporta a través del poro de membrana; D: corriente eléctrica: atraviesa constantemente al nanoporo, de modo que al pasar el ADN, cada uno de los nucleótidos altera la corriente eléctrica de forma específica; E: representación de la toma de datos, cada nucleótido es identificado gracias al patrón específico con el que modifica la corriente eléctrica. Imagen modificada de <https://togotv.dbcls.jp/togopic.2020.01.html>.

# *Justificación*





1. La progresión de la artrosis de rodilla es muy heterogénea; la mayoría de pacientes sufren un empeoramiento gradual a lo largo del tiempo, sin embargo, existe un pequeño grupo de pacientes que avanzan desde una articulación sana, hasta una articulación con un alto grado de severidad en un período muy corto de tiempo, de pocos años e incluso meses. Este fenotipo rápido de artrosis de rodilla se identificó por primera vez en 1957, y desde entonces distintos grupos de investigación han establecido sus propias definiciones, pero a día de hoy, no existe un *gold standard* para definirlo.
2. Además de la heterogeneidad en su progresión, el desarrollo de la artrosis puede depender de múltiples factores, entre los cuales está presente la genética. Si bien esta patología no es hereditaria, diversos estudios han puesto de manifiesto que la presencia de determinados polimorfismos, tanto nucleares como mitocondriales, pueden afectar a su desarrollo. Concretamente, en el ámbito del ADN mitocondrial, muchos trabajos están incrementando sus esfuerzos en la búsqueda de posibles variantes genéticas mitocondriales que puedan estar asociadas con la incidencia y la progresión de la patología, debido a la importancia de la mitocondria en la supervivencia del condrocito, el único tipo celular presente en el cartílago articular. Sin embargo, cabe destacar en primer lugar la ausencia de estudios que analicen el efecto del ADN mitocondrial en el desarrollo del fenotipo rápido de artrosis de rodilla, y, por otro lado, en el campo de la artrosis, nunca se han evaluado interacciones entre polimorfismos nucleares y mitocondriales que puedan condicionar el desarrollo de la patología, ni tampoco como éstas pueden afectar al fenotipo de progresión rápida.
3. Múltiples estudios han tratado de establecer herramientas predictivas basadas en distintas variables predictoras, con el fin de tratar de predecir la incidencia y la progresión de esta patología, pero pocos estudios se han centrado en desarrollar herramientas predictivas para el fenotipo rápido de artrosis de rodilla. Una herramienta en este campo es de suma importancia, ya que permitiría identificar pacientes con un elevado riesgo de progresar rápidamente. Por un lado, esto permitiría al reumatólogo clínico aplicar terapias preventivas antes de que se desencadene el proceso y, por otro lado, estos pacientes podrían incluirse en futuros ensayos clínicos para tratar de identificar posibles tratamientos para la artrosis, la cual a día de hoy continúa sin tener un tratamiento eficaz que sea capaz de revertir o frenar totalmente la progresión de la enfermedad.



## *Hipótesis y objetivos*



*Hipótesis*

La interacción entre la genética mitocondrial y nuclear define biomarcadores moleculares del fenotipo de rodilla con progresión rápida.

*Objetivo principal:*

Búsqueda de polimorfismos genéticos que se asocien con el riesgo de desarrollar el fenotipo de artrosis de rodilla con progresión rápida.

*Objetivos específicos:*

1. Establecimiento de los criterios de progresión y selección de los pacientes.
2. Analizar el genoma mitocondrial mediante metodologías de NGS con el fin de identificar posibles variantes asociadas con el fenotipo de artrosis de rodilla con progresión rápida en una cohorte exploratoria.
3. Exploración del significado funcional de las variantes mitocondriales asociadas con el fenotipo de progresión rápida mediante análisis del transcriptoma por RNA-seq.
4. Analizar el impacto de las interacciones entre el genoma nuclear y mitocondrial en el desarrollo del fenotipo de artrosis de rodilla con progresión rápida, y desarrollo de un modelo predictivo del fenotipo con aplicación clínica.



# *Material y métodos*





## Material y métodos objetivo específico número 1: establecimiento de los criterios de progresión y selección de los pacientes.

### *Cohortes de estudio.*

En el desarrollo de los siguientes objetivos de esta tesis doctoral se emplearon tres cohortes de pacientes: la cohorte americana del OAI, y un conjunto de dos cohortes europeas de validación, la cohorte española PROCOAC y la holandesa CHECK. Todas las cohortes utilizadas en este estudio consisten en estudios prospectivos en los cuales los pacientes realizan visitas clínicas repetidas a lo largo del tiempo. A continuación, se detallan las características de cada una de las tres cohortes utilizadas en este trabajo.

### *Cohorte Osteoarthritis Initiative (OAI).*

Esta cohorte pertenece al Instituto Nacional de Salud de los Estados Unidos de América (<https://nda.nih.gov/oai/>). Contiene datos clínicos, demográficos y radiológicos de artrosis de rodilla de 4.796 pacientes. Los criterios de inclusión que la institución americana tuvo en cuenta para reclutar individuos fueron los siguientes:

- Sujetos de edades comprendidas entre 45 y 79 años.
- Individuos con (o con riesgo de desarrollar) artrosis de rodilla tibial-femoral sintomática.
- Los individuos seleccionados fueron agrupados en tres subcohortes en función de una serie de características clínicas:
  - 1) Subcohorte de progresión: compuesta por individuos con artrosis de rodilla tibial-femoral sintomática en la visita de reclutamiento o *screening* (grado KL  $\geq 2$ ). En esta visita los pacientes presentan osteofitos en al menos una rodilla en la zona tibial-femoral y síntomas como dolor y rigidez a diario. Este subgrupo lo componen un total de 1.389 pacientes en los que se quiere estudiar la progresión de la enfermedad.
  - 2) Subcohorte de incidencia: compuesta por individuos que no presentan artrosis de rodilla tibial-femoral sintomática, tal y como se define para la subcohorte de progresión, pero presentan dos o más factores de riesgo para desarrollarla en una o ambas rodillas como por ejemplo, sobrepeso, tener 70 o más años de edad, registro clínico de lesión previa, historial de cirugía de rodilla, reemplazamiento total de rodilla debido a artrosis en un hermano o padre biológico y actividad física diaria elevada, entre otros. Este es el subgrupo más numeroso, con un total de 3.285 pacientes que se espera que desarrollen artrosis de rodilla en un futuro.
  - 3) Subcohorte de control: compuesta por individuos que carecen de artrosis de rodilla tibial-femoral o patelofemoral sintomática y factores de riesgo. Este grupo es el más pequeño, con un total de 122 individuos.

El estudio de la cohorte OAI comenzó en el año 2004 y cuenta con visitas anuales hasta los 108 meses (9 años de seguimiento), si bien en esta tesis se ha manejado la información hasta los 96 meses de seguimiento. En la visita basal se recogen muestras de sangre y de orina, se toman datos clínicos del individuo, se analizan mediante técnicas de imagen ambas rodillas y se estudian parámetros de actividad física. Además, los pacientes responden a una batería de *tests* relacionados con su sintomatología, estado de salud y medicación. El protocolo realizado en la

## Material y métodos

visita basal se repite durante las cuatro primeras visitas hasta alcanzar los 48 meses de seguimiento, tras este período de tiempo se alternaron visitas donde solo se realizaba la batería de *tests*, en las visitas de 60 y 84 meses, y visitas donde se repetía el mismo procedimiento aplicado en la visita basal, concretamente en las visitas de 72 y 96 meses.

La edad media de los individuos que conforman esta cohorte es de  $61,71 \pm 9,31$  años, con un IMC promedio de  $28,09 \pm 4,65$  kg/m<sup>2</sup>, y el 58% son mujeres. A lo largo del estudio, el 37% de los participantes desarrollaron artrosis radiográfica de rodilla, el 7% artrosis radiográfica de cadera y el 48% desarrollaron artrosis en ambas articulaciones.

### *Cohorte Prospective Cohort of Osteoarthritis from A Coruña (PROCOAC).*

La cohorte PROCOAC (Oreiro-Villar *et al.*, 2022) la integran pacientes del servicio de Reumatología del Hospital Universitario de A Coruña. Al igual que en la cohorte OAI, se dispone de información de seguimiento clínica, demográfica y radiológica. En cuanto a los criterios de inclusión, se describen a continuación:

- Pacientes mayores de 55 años atendidos en el Servicio de Urgencias por un cuadro de dolor abdominal y a los que se les realizó una radiografía simple de abdomen (incluidas ambas caderas).
- Pacientes con dolor a nivel de manos y con un diagnóstico de artrosis según los criterios de la *American College of Rheumatology* (ACR).
- Pacientes con dolor de rodilla y con un diagnóstico de artrosis radiográfica de rodilla siguiendo los criterios de la ACR.
- Pacientes con dolor en la cadera y diagnosticados de artrosis radiográfica de cadera siguiendo los criterios de la ACR.

El proceso de inclusión comenzó en el año 2006 y se mantiene activo actualmente. Tras firmar el consentimiento informado, se revisa la información clínica del paciente y se realiza una entrevista para completar el historial médico, actualizando la información clínica e incluyendo la información demográfica. También se recogen muestras biológicas y se realizan radiografías de ambas manos, rodillas y caderas.

Entre las variables que se recogen está parte del historial médico personal y familiar del individuo; asimismo, se anotan diferentes comorbilidades que pueda padecer el paciente, se hace una exploración física de las articulaciones y una batería de *tests*. Este procedimiento se hace en la primera visita, y se repite en posteriores visitas bianuales. La valoración de las articulaciones se basa en la escala KL. Para clasificar la severidad de la enfermedad se codifican a los pacientes dependiendo del número de articulaciones afectadas.

En el año 2018, la cohorte PROCOAC constaba de 937 pacientes, de los cuales el 93,1% (873 pacientes) presentaba artrosis radiográfica de rodilla, 83,5% (783) artrosis radiográfica de cadera, y el 72,4% (679) artrosis radiográfica de manos. El 53,3% de los participantes presentaban artrosis en más de una articulación. La edad media de la población era de  $63,9 \pm 8,9$  años y una media de IMC de  $29,6 \pm 5,1$  kg/m<sup>2</sup>. La gran mayoría de los participantes eran mujeres (75,1%). Actualmente, en el año 2022 la cohorte cuenta ya con 1.537 pacientes reclutados.

### *Cohorte Cohort Hip and Cohort Knee (CHECK).*

La cohorte CHECK, de la Fundación de Artritis Holandesa (<https://www.check-onderzoek.nl/>), es una cohorte de artrosis de rodilla y de cadera temprana, compuesta por 1.002

participantes y que combina información biológica, psicológica y social. Los criterios de inclusión de los pacientes son los siguientes:

- Pacientes de edades comprendidas entre 45 y 65 años.
- Todos los pacientes presentan dolor y/o rigidez en la rodilla y/o cadera al inicio del estudio.
- Fueron incluidos antes de consultar estos síntomas con el especialista, o hasta 6 meses tras la consulta.
- Al inicio del estudio no presentan ninguna condición patológica que pueda explicar el dolor como por ejemplo otra enfermedad reumática, reemplazo articular, displasia congénita, fracturas o roturas meniscal entre otras.

Los pacientes de la cohorte CHECK realizan visitas hasta los 10 años de seguimiento. En la visita basal se estudia el grado KL de las articulaciones de los pacientes con técnicas de imagen, se recogen muestras de sangre y orina y, además, los sujetos cumplimentan el *test* WOMAC para estudiar el dolor, la rigidez y la funcionalidad de la articulación, y también el *test* SF-36 (*Short Form 36-item health status survey questionnaire*) para estudiar la fatiga y diversos factores ambientales y personales. En base a los síntomas que presenten en la visita basal, los pacientes se subdividieron en dos grupos (Wesseling *et al.*, 2016), aquellos con síntomas leves siguieron un itinerario de visitas de 2, 5, 8 y 10 años tras la visita basal, en los cuales se registraron datos clínicos de seguimiento, mientras que los pacientes con síntomas más severos realizaron visitas anuales. Los análisis realizados en la visita basal se repitieron en las visitas posteriores en ambos grupos de pacientes.

La edad promedio de los individuos de la cohorte CHECK es de  $56 \pm 5$  años, siendo el 79% de los individuos mujeres y con un IMC promedio de  $26 \pm 4$  kg/m<sup>2</sup>. El 41% de los individuos presentaban artrosis de rodilla, el 17% artrosis de cadera y el 42% restante mostraban artrosis en ambas articulaciones. Con respecto al grado KL, la gran mayoría de los participantes iniciaron el estudio con un grado KL 0 o 1 (68 y 25% respectivamente), el resto de individuos comenzaron en grado KL 2 (6%) o 3 (1%).

### *Establecimiento de los criterios de progresión rápida.*

Con el objetivo de seleccionar pacientes con distintos grados de progresión radiológica dentro de las cohortes, se establecieron unos criterios de selección específicos. Dichos criterios están basados en un empeoramiento de la severidad de la artrosis en base a la escala KL en un período de tiempo de 48 meses (o 60 meses en los pacientes de la cohorte CHECK). Para esta tesis se definieron dos tipos de criterios, dependiendo del estado radiológico de la articulación del sujeto en su visita basal:

1. *Progresión rápida en sujetos sin daño radiológico en la visita basal (criterio Driban).*

Este criterio fue descrito por Driban y colaboradores (Driban *et al.*, 2016), los cuales acuñaron el término de *Accelerated Knee Osteoarthritis*. En base a este criterio, todos los individuos parten de una situación de grado KL basal de 0 o 1. Por lo tanto, los grupos de progresión establecidos en base a este criterio fueron los siguientes:

- Progresor rápido: individuo que partiendo de un grado KL basal de 0 o 1, llega a un grado KL igual o superior a 3 en un período de seguimiento no superior a 48 meses.
- Progresor no-rápido: individuo que partiendo de un grado KL basal de 0 o 1, llega a un grado KL igual 2 en un período de seguimiento no superior a 48 meses; o lo que es lo mismo, el paciente tarda más de 48 meses en alcanzar un grado KL mayor o igual a 3.

## Material y métodos

- No progresor: individuo que partiendo de un grado KL basal de 0 o 1 se mantiene estable bilateralmente durante el período de seguimiento de 48 meses.

Tal y como se mostrará posteriormente en el apartado de *Resultados*, esta clasificación se ha colapsado en dos grupos de progresión, progresores rápidos y progresores no-rápidos, incluyendo en estos últimos a todos aquellos individuos que no cumplen los criterios de progresión rápida.

### 2. Progresión rápida aplicada a sujetos sin daño radiológico basal y sujetos con daño radiológico basal moderado-bajo (criterio GIR-blanco).

Este segundo criterio fue creado en base a la experiencia de los reumatólogos del servicio de Reumatología del Complejo Hospitalario Universitario A Coruña. Mediante este criterio, se añaden, a los pacientes clasificados como progresores rápidos en base al criterio Driban, a todos aquellos pacientes que presentan un grado KL basal de 2, que son los que mayoritariamente frecuentan las consultas de reumatología. De esta manera se consigue incrementar también el tamaño muestral del estudio. Los grupos de progresión establecidos en base a este criterio serían los siguientes:

- Progresor rápido: individuo que partiendo de un grado KL basal de 0 o 1, llega a un grado KL igual o superior a 3 en un período de seguimiento no superior a 48 meses; o individuo que, partiendo de un grado KL basal de 2, llega a un grado KL de 4 o sufre una cirugía de reemplazo articular durante el período de seguimiento indicado.
- Progresor no-rápido: individuo que partiendo de un grado KL basal de 0 o 1, llega a un grado KL igual 2 en un período de seguimiento no superior a 48 meses; o individuo que, partiendo de un grado KL basal de 2, llega a un grado KL máximo de 3 durante el período de seguimiento.
- No progresor: individuo que partiendo de un grado KL basal de 0, 1 o 2 se mantiene estable bilateralmente durante el período de seguimiento de 48 meses.

Del mismo modo que se realizó en el criterio Driban, esta clasificación también se ha colapsado en progresores rápidos y no-rápidos en el criterio GIR-blanco. Los dos criterios descritos en este trabajo estudian la evolución radiográfica del individuo durante un período de seguimiento de 48 meses. Teniendo en cuenta los regímenes de visitas dentro de las cohortes, se establecieron los períodos de seguimiento que se muestran en la figura 37. Tanto en la cohorte OAI como en la PROCOAC se pudieron establecer períodos de seguimiento de 48 meses. Sin embargo, debido al régimen de visitas de la cohorte CHECK, se estableció un único período de seguimiento de 60 meses, tal y como se describe en otros trabajos publicados (Wesseling *et al.*, 2015).

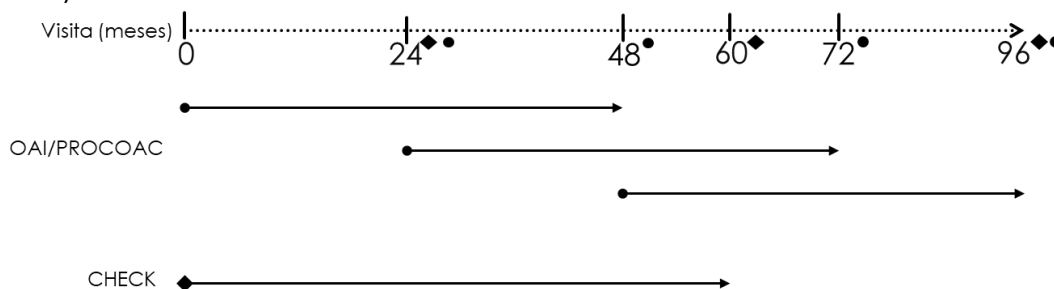


Figura 37. Posibles períodos de seguimiento en las tres cohortes. Los períodos de seguimiento están representados con flechas, los círculos marcan los períodos de seguimientos y las visitas disponibles en la cohorte OAI y PROCOAC, los rombos identifican los períodos de seguimiento y las visitas disponibles en la cohorte CHECK.

*Análisis estadístico del objetivo número 1.*

Tanto en la cohorte OAI, como en las cohortes CHECK y PROCOAC en conjunto, se realizó un análisis univariante de las variables demográficas de interés. Las variables demográficas cuantitativas se analizaron mediante un análisis Kruskal Wallis en los criterios de tres grupos de progresión, realizando comparaciones entre parejas, mientras que, para los criterios colapsados con dos grupos de progresión, se analizaron las variables cuantitativas mediante regresión logística. Por otro lado, las variables cualitativas se analizaron en todos los criterios de progresión mediante análisis de  $\chi^2$ .

Con el fin de determinar qué criterio era el más adecuado para la realización de esta tesis doctoral, se realizó un análisis univariante comparativo entre los individuos de los dos criterios de progresión (N=852), con el fin de estudiar si había correlación en la clasificación mediante el estadístico Kappa en el *software IBM-SPSS, release 25* (IBM, Armonk, NY, USA).



## Material y métodos objetivo específico número 2: analizar el genoma mitocondrial mediante metodologías de NGS con el fin de identificar posibles variantes asociadas con el fenotipo de artrosis de rodilla con progresión rápida.

Para la realización de este objetivo se trabajó en dos fases, la primera consistió en una fase de exploración en la cohorte OAI, donde se aplicaron técnicas de secuenciación masiva de segunda generación para identificar variantes mitocondriales que se pudieran asociar con el fenotipo de artrosis de rodilla con progresión rápida. En la segunda fase de este objetivo específico, se trabajó con un conjunto de dos cohortes externas (CHECK y PROCOAC), donde se llevó a cabo una validación de los datos obtenidos en la fase de exploración mediante técnicas de secuenciación capilar.

### *Secuenciación masiva del ADNmt en la cohorte OAI.*

La fase de exploración en la cohorte OAI tiene como objetivo identificar variantes genéticas mitocondriales que se puedan asociar con el fenotipo de artrosis de rodilla con progresión rápida y, para ello, se empleó una metodología de secuenciación masiva de segunda generación. En este trabajo se empleó la metodología *Ion Torrent* mediante una plataforma Ion S5 XL (ThermoFisher) y el robot de preparación de librerías Ion Chef (ThermoFisher). A continuación, se hace un resumen de los procesos generales que componen esta metodología, y más adelante se explicarán en detalle en sus correspondientes apartados:

- Preparación de librerías: una librería se define como una representación fragmentada de la muestra inicial, de modo que cada uno de estos fragmentos estará flanqueado por adaptadores de secuencia conocida (universal) y específicos de la tecnología empleada. En este proceso se utilizó un panel de oligonucleótidos específico para amplificar las regiones de interés del ADNmt con el fin de generar las librerías. Este proceso se realizó íntegramente en el Ion Chef.
- Carga en soporte sólido o *Templating*: también se realiza con el Ion Chef. Tras la generación de las librerías es necesario introducirlas en un soporte sólido para su posterior secuenciación. Con esta plataforma de secuenciación, el soporte físico donde se cargan las muestras es un chip formado por millones de micropocillos, cada uno con su propio medidor semiconductor en la base. Cada fragmento que compone la librería se unirá individualmente a una esfera magnética y será amplificado de forma clonal sobre la superficie de esta. Tras la amplificación sobre las esferas magnéticas, estas serán introducidas dentro del chip semiconductor, entrando una única esfera en un único micropocillo.
- Secuenciación: los chips cargados se introducirán en el secuenciador Ion S5 XL. Durante el proceso de secuenciación, el chip será rellenado con uno de los cuatro dNTPs en un orden específico durante distintos ciclos. Cada vez que un dNTP se une a un fragmento presente en un micropocillo se liberará un H<sup>+</sup>, acidificando el medio. Este leve cambio es detectado por el medidor semiconductor situado en la base del micropocillo, transformando la señal de pH a señal eléctrica, la cual será proporcional al número de nucleótidos incorporados. Este proceso se repetirá un número determinado de veces y el secuenciador recogerá datos de cada uno de los pocillos del chip, obteniendo de

## Material y métodos

forma simultánea información nucleotídica de múltiples muestras y fragmentos distintos. Una vez finalizado los datos serán transportados al servidor del secuenciador para su análisis.

### *Selección y cuantificación de las muestras de la cohorte OAI para su secuenciación.*

El estudio de exploración de las variantes genéticas mitocondriales mediante técnicas de secuenciación masiva se llevó a cabo en un subgrupo inicial de muestras de la cohorte del OAI. En total, para la secuenciación masiva del ADNmt se seleccionaron 450 individuos en este estudio inicial de exploración, de los cuales 184 pertenecían al grupo de progresores rápidos y 266 a las otras dos categorías, que incluyeron 128 progresores no-rápidos y 138 no progresores.

Debido a la alta sensibilidad de la técnica de secuenciación masiva, es necesario introducir una cantidad específica de ADN con el fin de conseguir el resultado deseado. Por ello es recomendable realizar una cuantificación mediante técnicas muy precisas; en este caso se utilizó un fluorómetro *Qubit 4* (Thermofisher) y el kit *dsDNA HS Kit* (Thermofisher). Con esta combinación se pueden cuantificar muestras que estén en una concentración comprendida entre 10 pg/ $\mu\text{L}$  y 100 ng/ $\mu\text{L}$ . El kit *dsDNA HS* contiene dos elementos, el reactivo de cuantificación y los estándares de ADN 1 y 2 para realizar la correcta calibración del instrumento. Es necesario realizar la calibración del *Qubit 4* previamente a la cuantificación, para ello se mezclaron 10  $\mu\text{L}$  del estándar 1 y 190  $\mu\text{L}$  del reactivo de cuantificación en un tubo *Qubit Assay* (Thermofisher) y 10  $\mu\text{L}$  del estándar 2 y 190  $\mu\text{L}$  del reactivo de cuantificación en otro tubo *Qubit Assay*. Se incubaron durante 2 minutos en oscuridad y se calibró siguiendo las instrucciones del fabricante. Debido a que el panel de oligonucleótidos necesario para generar las librerías necesita únicamente 0,1 ng en un volumen final de 15  $\mu\text{L}$ , fue necesario realizar una dilución inicial 1:100 de cada muestra, de modo que 10  $\mu\text{L}$  de muestra se diluyeron en un volumen final de 1000  $\mu\text{L}$  en un tubo *DNA LoBind* (Eppendorf) de 1,5 mililitros (mL). A continuación, se pipetearon 2  $\mu\text{L}$  de muestra y 198  $\mu\text{L}$  del reactivo *dsDNA HS Kit* en un tubo *Qubit Assay* (Thermofisher), y se incubó 2 minutos en oscuridad a temperatura ambiente. Acto seguido se analizó la concentración de cada dilución con el fluorómetro *Qubit 4*. Por último, siguiendo las indicaciones del panel de secuenciación de ADNmt, se realizó una dilución final de la muestra a 0,1 ng/ $\mu\text{L}$  en un tubo *DNA LoBind* de 1,5 mL. La figura 38 muestra un esquema de este proceso.

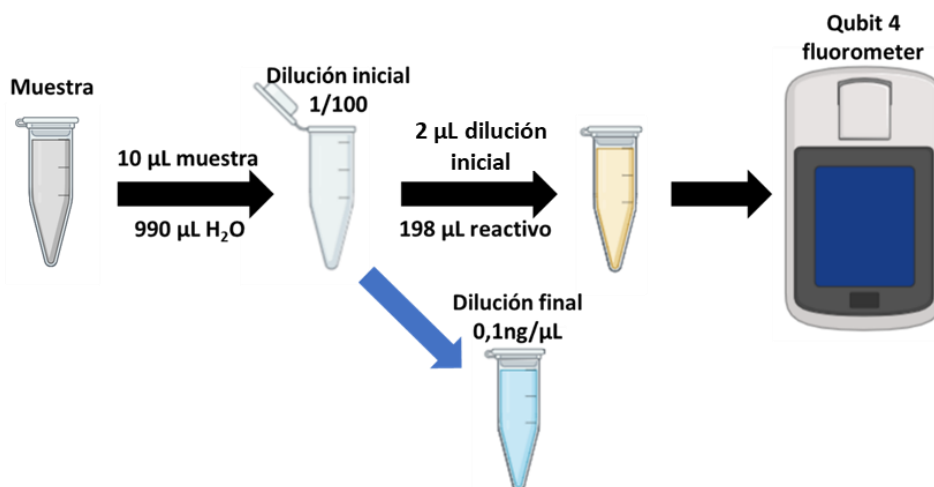


Figura 38. Proceso de cuantificación de las muestras de ADNmt con Qubit. Primero se diluye inicialmente la muestra y se cuantifica (flechas negras). Tras la cuantificación (flecha azul), se realiza la dilución final de la muestra a una concentración de 0,1 ng/ $\mu\text{L}$ .



### Preparación de las librerías de las muestras con la plataforma Ion Chef.

Para generar las librerías, es necesario amplificar todo el genoma mitocondrial mediante la tecnología *Ampliseq* (Thermofisher). Esta amplificación se realiza mediante paneles de oligonucleótidos previamente diseñados y validados por la compañía; en el caso del genoma mitocondrial, el panel de oligonucleótidos empleado fue el *Precision ID mtDNA Whole Genome Panel* (Thermofisher), que es capaz de generar todos los fragmentos de ADN mitocondrial necesarios para obtener la secuencia completa de 16.569 pares de bases de esta molécula. Este panel contiene dos conjuntos (*pools*) de 81 parejas de oligonucleótidos cada uno, con un grado mínimo de solapamiento entre los *pools*, generando un total de 162 amplicones. Además, de los 324 oligonucleótidos totales que se compone el panel, 283 son degenerados para evitar que la presencia de diferentes polimorfismos o SNPs impida la amplificación de ciertos fragmentos. De media, cada amplicón mide 163 pb, y en conjunto cubren todo el genoma mitocondrial (Fig. 39).

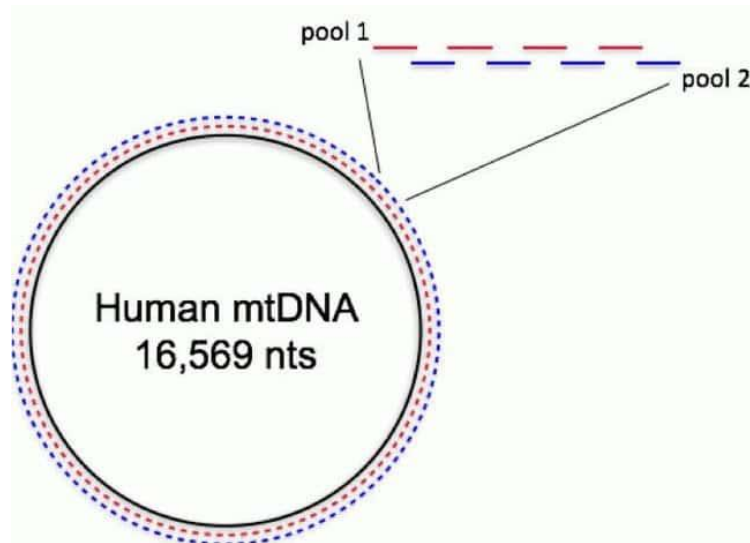


Figura 39. Cobertura del Precision ID mtDNA Whole Genome Panel. Imagen obtenida de <https://www.thermofisher.com>.

Además del panel de oligonucleótidos fue necesario emplear el kit de preparación de librerías *Precision ID DL8 Kit* (Thermofisher), que contiene los reactivos necesarios para generar de forma automática hasta 32 muestras en cuatro *pools* de 8 muestras cada uno. Este kit incluye:

- Cuatro conjuntos de plásticos desechables.
- Cuatro cartuchos de reactivos y cuatro cartuchos de soluciones. Con los elementos y enzimas necesarios para generar las librerías.
- Cuatro placas de PCR. Cada una de estas placas lleva una mezcla de adaptadores y *barcodes*. Los *barcodes* están presentes en la sexta columna de la placa, y este elemento es fundamental en el proceso de secuenciación masiva ya que permite añadir a todas las secuencias del mismo individuo un identificador específico. Para ello, las cuatro placas contienen una combinación de 8 *barcodes* únicas por placa, lo cual nos va a permitir identificar a cada muestra dentro del *pool* final de 32 librerías. Los adaptadores son necesarios para llevar a cabo los procesos de preparación de librerías y *templating*.

Los fragmentos que componen la librería creada con el kit *Precision ID DL8* constan de los siguientes elementos (Fig. 40):

## Material y métodos

- Adaptador A. En esta región se unirá el oligonucleótido de secuenciación. Contiene la *Key sequence*, que es una secuencia de 4 nucleótidos (AGCT) que sirve para calibrar la señal y también de control de calidad interno del proceso.
- *Barcode*. Secuencia nucleotídica que identifica a la muestra; todos los fragmentos pertenecientes a un individuo tienen el mismo código. Este kit contiene 32 *barcodes* distintos, lo que permite secuenciar simultáneamente hasta 32 muestras.
- Secuencia de interés. Región nucleotídica resultante de la amplificación con el panel *Precision ID mtDNA Whole Genome*. Esta es la parte que se quiere estudiar por secuenciación masiva del fragmento de librerías.
- Adaptador P1B. Necesario para la unión a las esferas magnéticas durante los procesos de generación de librerías y carga del chip.

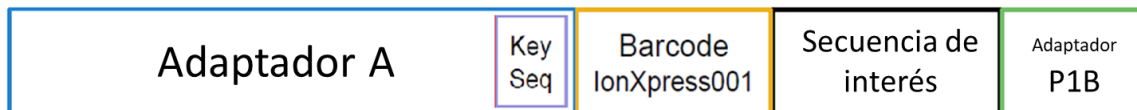


Figura 40. Representación de los elementos que componen un fragmento de librería de ADN elaborado con el kit *Precision ID DL8 Kit*. Imagen modificada de <https://www.thermofisher.com>.

El proceso general para elaborar las librerías con esta tecnología es el siguiente:

1. Amplificación con el panel de oligonucleótidos de ADNmt.
2. Eliminación de los oligonucleótidos de amplificación: usando el reactivo FuPa<sup>®</sup> se eliminan los oligonucleótidos, dejando los extremos romos.
3. Ligación de adaptadores al fragmento de interés usando una ligasa.
4. Enriquecimiento de las librerías: debido a que la unión de los adaptadores no es perfecta pueden quedar fragmentos sin adaptadores. El enriquecimiento consiste en aumentar la cantidad de fragmentos con adaptadores frente a aquellos que no tienen, mediante una amplificación de 5 ciclos de los fragmentos con los adaptadores.
5. *Equalizer*: en este proceso se iguala la concentración de todas las librerías que irán finalmente juntas en un único *pool* al final del proceso. Para ello se usan unas esferas con un número específico de puntos de anclaje, a las cuales se unirá una cantidad determinada de fragmentos de cada librería hasta saturarse. Sólo seleccionando estas esferas se puede conseguir la misma cantidad de fragmentos de cada una de las librerías, trabajando por lo tanto con librerías que tienen la misma concentración.
6. Generación del *pool* de librerías: gracias a la presencia de los *barcodes*, todos los fragmentos pertenecientes a la librería de un individuo presentan una misma secuencia identificadora que es distinta del resto de librerías. Esto permite unir las 8 librerías obtenidas en un mismo tubo o *pool* de librerías, y que el secuenciador identifique a qué muestra pertenece cada fragmento.

El protocolo de preparación de un *pool* de 8 librerías está representado en la figura 41 y consta de los siguientes pasos:

1. Atemperar, al menos 20 minutos antes de la carga en el sistema Ion Chef, el reactivo *ID DL8 reagents* y una alícuota de cada *pool* de oligonucleótidos del *Precision ID mtDNA Whole Genome Panel*.
2. Introducir 14  $\mu\text{L}$  de agua a cada pocillo de la columna 1 de la placa con los *barcodes*. A continuación, se añade 1  $\mu\text{L}$  de la dilución de cada muestra a 0,1  $\text{ng}/\mu\text{L}$ . Así, en cada uno de los 8 pocillos de la columna 1 tendremos una muestra distinta a una concentración de 0,1  $\text{ng}$  en 15  $\mu\text{L}$ .

3. Añadir 150  $\mu\text{L}$  del *pool* de oligonucleótidos 1 a la posición A (Fig. 42) del cartucho de reactivos, y 150  $\mu\text{L}$  del *pool* de oligonucleótidos 2 a la posición B. Abrir y quitar las tapas de los tubos.

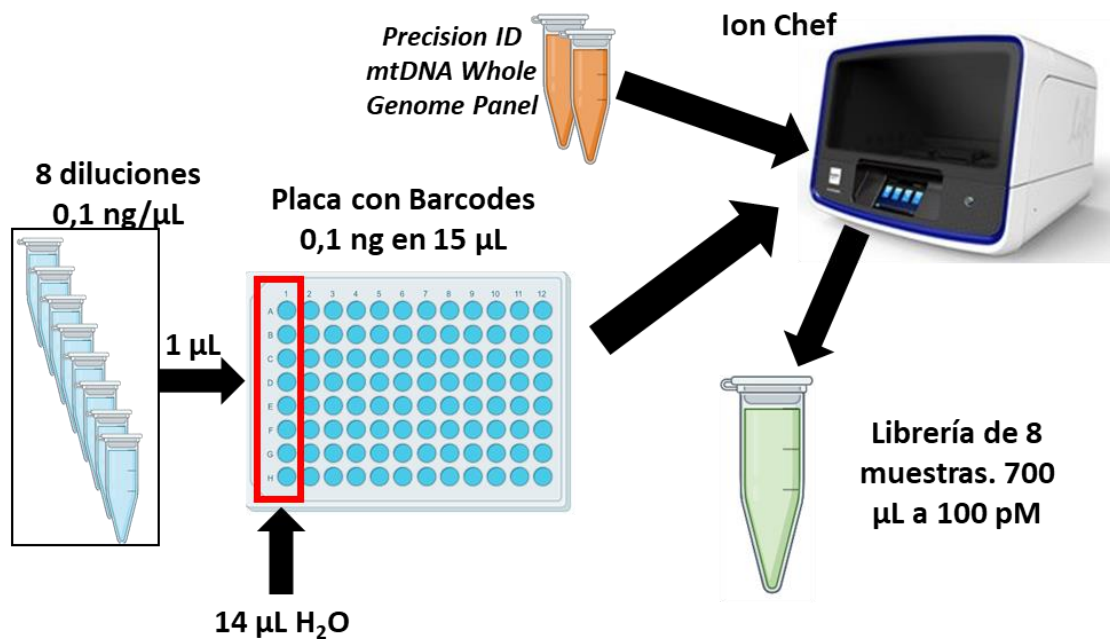


Figura 41. Protocolo de generación de librerías con el Ion Chef, Precision ID mtDNA Whole Genome Panel y el kit Precision ID DL8. La primera columna de la placa de barcodes marcada con el rectángulo rojo es donde se introducen las muestras

4. Introducir los reactivos en el Ion Chef. En las condiciones del programa es necesario especificar el número de *pools* de oligonucleótidos, el número de ciclos y el tiempo de anillamiento y de extensión. Para el kit *Precision ID mtDNA Whole Genome Panel*, el número de oligonucleótidos es 2, el número de ciclos recomendado es 22, y el tiempo de anillamiento y extensión es de 4 minutos.
5. Tras el proceso, el *pool* de 8 librerías estará en la posición D del cartucho de reactivos a una concentración de 100 picomoles (pM) en un volumen final de 700  $\mu\text{L}$ .

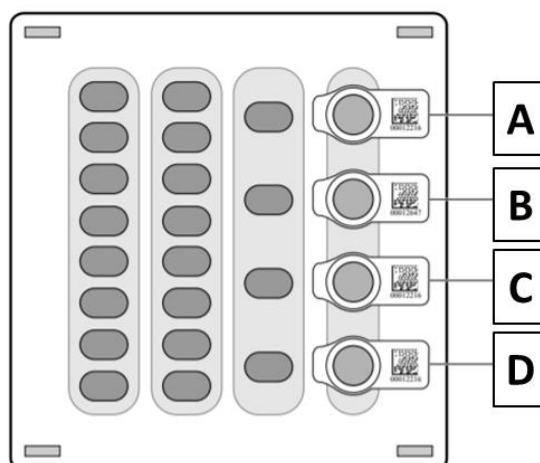


Figura 42. Representación del cartucho de reactivos del kit Precision ID DL8 Kit. En la posición A se introduce el *pool* de oligonucleótidos 1, en la posición B el 2, la posición C es un tubo vacío y la posición D es donde finalmente estará el *pool* de 8 librerías. Imagen modificada de <https://www.thermofisher.com>

## Material y métodos

6. Almacenar el *pool* de librerías a 2-8°C durante una semana, o a -30°C -10°C para períodos de tiempo más largos.

Antes de proceder con el siguiente apartado del protocolo, es necesario llevar a cabo la cuantificación de cada *pool* de librerías.

### *Carga del chip con Ion Chef.*

Este proceso es intermedio a la preparación de librerías y secuenciación, y consiste en introducir las librerías de secuenciación en un soporte sólido para su posterior análisis en el Ion S5 XL. Los kits empleados para realizar la carga del chip fueron:

- *Ion S5 Chef Supplies* (Thermofisher). Cuatro cajas que contienen todo el material fungible necesario para el sistema Ion Chef.
- *Ion S5 Chef Solutions* (Thermofisher) e *Ion 510 & Ion 520 & Ion 530 Chef reagents* (Thermofisher). Ambos están compuestos por 4 cartuchos con soluciones de trabajo del Ion Chef.
- *Ion 530 Chip Kit* (Thermofisher). Contiene 8 chips, que son el soporte físico usado en esta metodología.

El proceso tiene tres etapas principales: amplificación clonal, enriquecimiento y carga del chip.

1. Amplificación clonal (Fig. 43A): se unen las librerías a las esferas que tiene secuencias coincidentes con el adaptador P1, con una mezcla de PCR y aceites de emulsión. Sobre las esferas ocurrirá un proceso de amplificación en emulsión (ePCR), de tal manera que cada una de las esferas quedará cubierta por múltiples copias del mismo fragmento. Debido a que el *pool* de librerías tiene una concentración muy baja, de aproximadamente a 30 pM en 25  $\mu$ L, teóricamente solamente un fragmento se va a unir a una esfera y, tras la ePCR, se generará una esfera monoclonal con múltiples repeticiones del mismo fragmento. Esto es el escenario ideal, sin embargo, no siempre ocurre de esta manera, por lo que es necesario realizar un proceso de enriquecimiento.
2. Enriquecimiento: el proceso de amplificación clonal no es 100% eficaz. En contraste con el escenario ideal de que un fragmento se una a una esfera (monoclonal, figura 43B a), puede ocurrir que sean varios los fragmentos que se unen a la misma esfera (policlonales, figura 43B b) o puede haber esferas a las que no se une ningún fragmento (esferas vacías, figura 43B c). El proceso de enriquecimiento tiene como objetivo la eliminación de las esferas vacías y, para llevarlo a cabo, son fundamentales los adaptadores A presentes en los amplicones. Los adaptadores A están biotinizados y pueden interactuar con esferas magnéticas marcadas con estreptovidina. La biotina y la estreptovidina se unen, de tal manera que usando un imán podremos seleccionar aquellas esferas que tienen fragmentos en su superficie gracias al adaptador A, y eliminar las que están vacías, ya que no se unen a ningún fragmento. Por otro lado, las esferas policlonales no pueden ser eliminadas durante el *templating*, sin embargo, los datos de éstas esferas son detectados y eliminados posteriormente por el software del secuenciador.
3. Carga del chip: Previo a este paso hay que llevar a cabo un proceso de desnaturalización con sosa, que permite que los fragmentos bicatenarios que están en la superficie de las esferas se desnaturalicen y sean monocatenarios. Gracias a esto, durante la fase de secuenciación el oligonucleótido de secuenciación se unirá a los fragmentos y podrán

incorporarse los nucleótidos de forma secuencial. Tras este último paso el chip estará listo para su secuenciación.

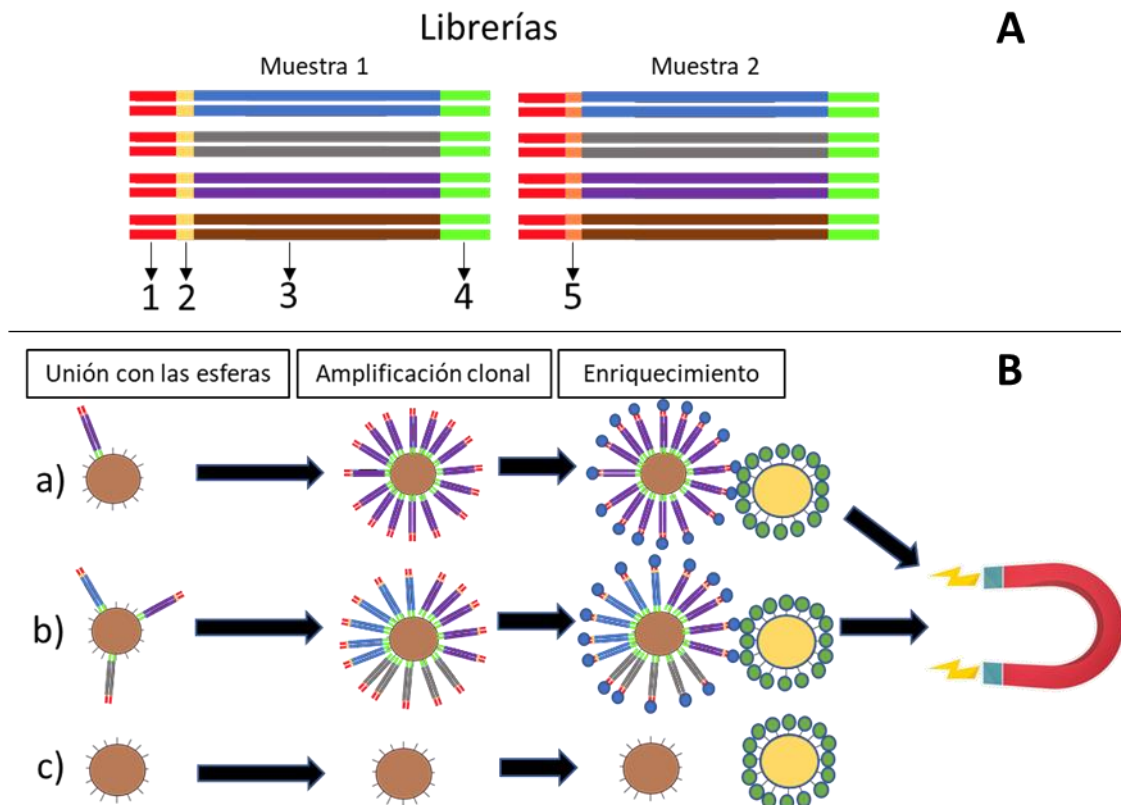


Figura 43. Proceso de amplificación clonal en Ion Torrent. A: Representación de los elementos de las librerías: 1. Adaptador A, 2. Barcode que identifica la muestra nº1, 3. Secuencia de interés del fragmento, 4. Adaptador P1b, 5. Barcode que identifica la muestra nº2. B: Representación de los procesos de amplificación clonal y enriquecimiento: a) esfera monoclonal, b) esfera policlonal y c) esfera vacía.

Debido a las especificaciones del aparato, el proceso de carga se puede hacer de forma simultánea en dos chips (Fig. 44). El protocolo a seguir es el siguiente:

1. Atemperar 45 minutos antes de la carga en el sistema Ion Chef el cartucho *Ion 510 & Ion 520 & Ion 530 Chef reagents*.
2. Crear un plan de carrera (*Run Plan*) para cada uno de los chips con el *Torrent Suite Software*; esto permitirá asignar una identidad de muestra a cada uno de los 32 *barcodes* que están presentes en el chip.
3. Preparar los dos *pools* de librerías: uno de cuatro librerías de 8 muestras cada una para el primer chip y otro *pool* de otras cuatro librerías distintas para el segundo chip. Cada uno de estos dos *pools* deben de estar a una concentración final de 30 pM en un volumen final de 25 µL.
4. Introducir los dos *pools* de librerías en el cartucho *Ion 510 & Ion 520 & Ion 530 Chef reagents*.
5. Iniciar el Ion Chef, introducir los fungibles del kit *Ion S5 Chef Solutions, Ion S5 Chef Supplies, Ion 530 Chip Kit Chef Supplies* y los dos chips *Ion 530*. Comprobar si cada Chip tiene asociado en el Ion Chef el plan de carrera que se diseñó en el segundo punto de este protocolo.
6. Iniciar el proceso automático de carga del chip. Una vez finalizado, la vida útil de los chips es de entre 6 y 8 horas.

## Material y métodos

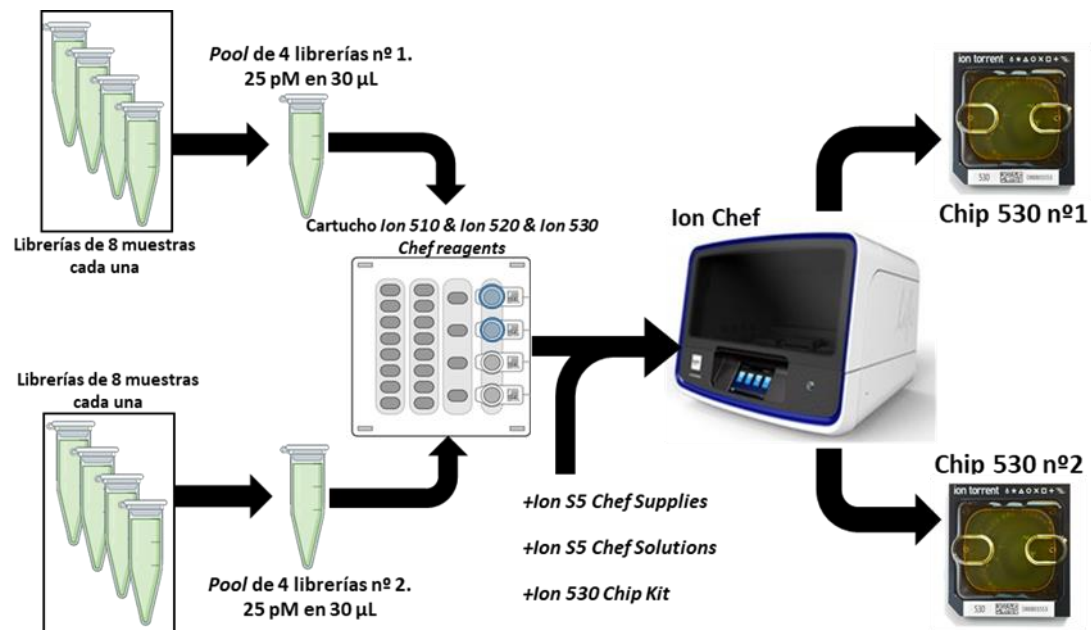


Figura 44. Proceso de carga del chip. A partir de dos grupos de 4 pools de 8 librerías cada uno, se generan dos pools finales de 32 librerías en un volumen final de 30 µL a 25 pM. El pool de librerías 1 se carga en la posición A, mientras que el pool de librerías 2 se carga en la posición B del cartucho Ion 510 & Ion 520 & Ion 530 Chefreagents, y se introduce posteriormente en el Ion Chef junto con los kits Ion S5 Chef Solutions, Ion S5 Chef Supplies e Ion 530 Chip Kit. Tras el proceso de templating obtendremos dos chips de 32 muestras cada uno, uno por cada pool final.

### Secuenciación masiva con el Ion S5 XL.

Tras realizar la amplificación, la generación de librerías y la carga del chip, queda el último paso, la secuenciación. Con esta tecnología vamos a trabajar sobre un soporte sólido, el chip, el cual tiene millones de pocillos, cada uno de los cuales tiene en su base un medidor semiconductor. Es en estos pocillos donde se introducirán las esferas mono o policlonales. En cada pocillo se realiza individualmente un proceso de secuenciación. Para ello se rellena el chip durante 10 segundos con un nucleótido, a continuación, durante 5 segundos se realiza un lavado del chip. Este proceso de 15 segundos de carga de un nucleótido y lavado se denomina *flow*, y ocurre de forma repetitiva siguiendo un orden específico a nivel de nucleótidos: T, A, C y G.

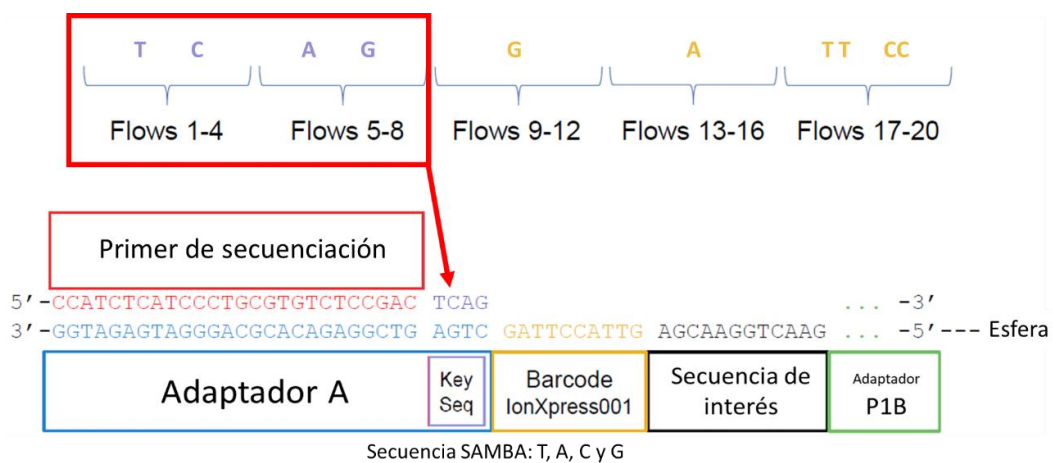


Figura 45. Representación del inicio del proceso de secuenciación de Ion Torrent. Tras la unión del oligonucleótido de secuenciación, los nucleótidos se irán introduciendo de forma separada (*flow*) siguiendo el orden establecido (T A C G). Los 8 primeros nucleótidos completarán la key sequence, necesaria para la correcta autocalibración del aparato. A continuación, el secuenciador comenzará a secuenciar el barcode identificador de cada muestra, y continuará secuenciando todo el fragmento presente en el micropocillo hasta completar el programa de secuenciación. Imagen modificada de <https://www.thermofisher.com>.



## Material y métodos

- *Ion S5 Sequencing solutions*: contiene las soluciones *Ion S5 Wash solution* e *Ion S5 Cleaning solution*. La primera sirve para resuspender los nucleótidos del cartucho de *reagents* y para realizar los *flows* de la secuenciación. La segunda sirve para hacer los lavados entre *flows*.

Antes de proceder con el protocolo de secuenciación es necesario inicializar el *Ion S5 XL*. En este proceso se introducirán los kits mencionados en el secuenciador y se resuspenderán los nucleótidos del cartucho de *reagents* para su uso. Tras la inicialización, el secuenciador podrá realizar un programa de secuenciación en un rango de 24 horas, este proceso está representado en la figura 47 y los pasos a seguir son los siguientes:

1. Introducir el primer chip en el secuenciador y seleccionar el *run plan* asociado a éste. Guardar el segundo chip en oscuridad a 4°C, atemperarlo veinte minutos antes de su uso.
2. Seleccionar el protocolo de secuenciación.
3. Iniciar la secuenciación del primer chip, el proceso dura unas 3 horas aproximadamente. Cuando termine se tiene que repetir el proceso para el segundo chip.

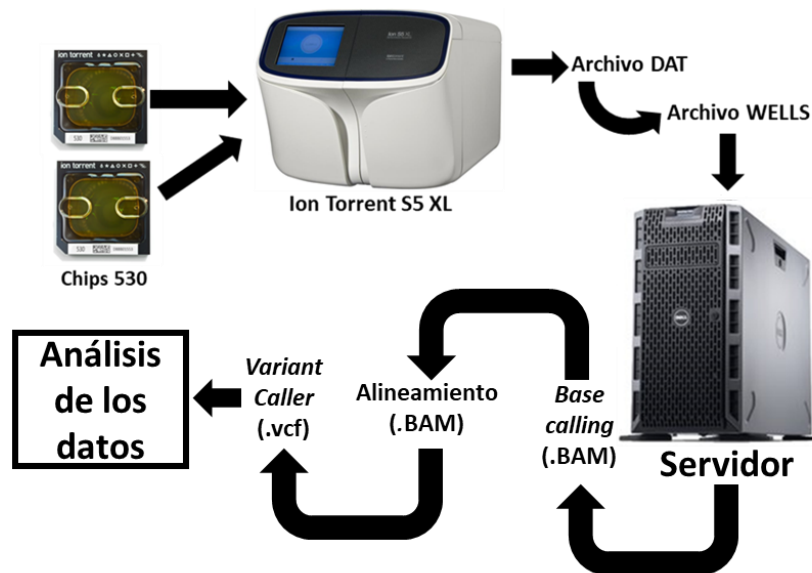


Figura 47. Representación del proceso de secuenciación en el *Ion Torrent S5 XL* y el procesamiento de los datos con el servidor *Torrent Suite*.

### *Análisis bioinformático de los resultados de secuenciación masiva.*

Las lecturas de *Ion S5 XL* fueron procesadas con el software *Ion Torrent Suite v5.12* (ThermoFisher Scientific). Para realizar el análisis de las variantes se usó una canalización adaptada para bibliotecas *Ampliseq* usando un modo de baja rigurosidad en el complemento *Torrent Variant Caller (TVC) (OCPv1i3 Panel-530-Chef-Somatic- Low stringency)*.

Con el paso del tiempo, en diferentes especies se han acumulado en el genoma nuclear diversos fragmentos de ADNmt denominados NUMTs (del inglés *Nuclear mitochondrial DNA segment*). Al realizar la amplificación y la secuenciación, además de obtener lecturas de los ADNmt presentes en las mitocondrias, también se pueden generar lecturas pertenecientes a estos NUMTs, los cuales generarían ruido en el análisis final. Para confirmar que los resultados no están “contaminados” con la presencia de estas secuencias, se llevó a cabo una canalización paralela, utilizando los mismos pasos, pero agregando un proceso previo de limpieza de lecturas



potencialmente derivadas de NUMT. Para ello, se realinearón las lecturas contra el genoma nuclear de referencia (GRCh28, cromosomas 1-23) y el genoma mitocondrial (secuencia de referencia de Cambridge revisada (rCRS)(Andrews *et al.*, 1999)) usando la optimización de alineación TMAP (Li y Homer, 2010). Tras el realineamiento, se obtienen unas puntuaciones de alineación para cada lectura y genoma mitocondrial; aquellas lecturas con una puntuación de alineación superior en el genoma nuclear que la obtenida contra la referencia rCRS, se etiquetaron como potencialmente derivadas de NUMTs y posteriormente se eliminaron del archivo .BAM. Dada la circularidad del genoma mitocondrial, y para evitar errores de alineamiento en los extremos de la referencia rCRS, el estándar de *Ion Torrent Suite* utiliza la referencia *PrecisionID\_mtDNA\_rCRS* que repite los primeros 80 nucleótidos de rCRS al final de la secuencia. La alineación se realiza nuevamente utilizando una versión de rCRS con un punto de corte alternativo. En esta versión alternativa de rCRS, se unen los extremos de la secuencia de referencia original rCRS y se desplaza el punto de corte entre los nucleótidos m.8284 y m.8285, lo que permite confirmar variantes identificadas en posiciones a menos de 500 pb de los extremos de la secuencia de referencia original.

Las variantes mitocondriales resultantes obtenidas después de aplicar el flujo de trabajo descrito anteriormente se filtraron inicialmente por su frecuencia de población. De este modo, se descartaron aquellas variantes genéticas mitocondriales con una frecuencia mínima alélica (MAF, del inglés *minor allele frequency*) inferior a 0,05. El segundo filtro consistió en fijar un umbral de detección de heteroplasma conservadora del 20 %. Aun así, después de la asignación automática de las variantes mitocondriales, se realizó una verificación manual utilizando el software *Integrative Genomics Viewer* (IGV) v2.8.0. El tercer y último filtro consistió en la selección de aquellas variantes genéticas de mtDNA que presentaban una distribución de frecuencia diferencial entre progresores rápidos y progresores no rápidos. Para ello, se aplicó un análisis univariante de  $\chi^2$  para obtener la relación de probabilidades (u OR, del inglés *Odds ratio*) y el intervalo de confianza del 95% (IC 95%) con el fin de seleccionar aquellas variantes representadas diferencialmente entre los grupos de progresión con un p-valor  $\leq 0,15$  (Harrell, 2001). A partir de estas variantes, se aplicó la corrección de Benjamin-Hochberg (BH) para controlar el *False Discovery rate* (FDR). Finalmente, se seleccionaron como posibles candidatas a aquellas variantes con un p-valor corregido menor o igual a 0,05.

*Asignación de la variante de interés por secuenciación capilar en el resto de la cohorte OAI y en las cohortes externas de validación.*

Tras el análisis bioinformático, la variante de interés obtenida en las 450 muestras analizadas por secuenciación masiva se analizó en el resto de la cohorte OAI (645 muestras) mediante un ensayo de secuenciación capilar, aplicando el método Sanger, diseñando los oligonucleótidos de PCR oportunos (Tabla 5). Estos oligonucleótidos consisten en parejas de pequeños fragmentos de nucleótidos que por complementariedad de bases hibridan en una región diana del ADNmt, y, permiten primero amplificar el fragmento que contiene al polimorfismo de interés mediante PCR, que es un paso crucial que permite aumentar la señal de fluorescencia obtenida durante el segundo proceso, la secuenciación Sanger. Las secuencias del ADNmt fueron obtenidas a partir de la base de datos del NCBI (*National Center for Biotechnology Information*) (<https://www.ncbi.nlm.nih.gov/>).

Como se ha comentado, el primer paso consistió en una amplificación mediante PCR del fragmento de ADNmt que contiene la posición de interés. Para ello, se emplearon 80 ng por muestra. Las concentraciones finales de los reactivos de PCR fueron: 1X de tampón de reacción, 2 mM de  $MgCl_2$ , 0,2 mM de cada dNTP, 0,3  $\mu M$  de PF, 0,3  $\mu M$  de PR y 0,05 U/ $\mu L$  de *Taq*

## Material y métodos

polimerasa en un volumen final de 10  $\mu$ L. Las condiciones de PCR consistieron en una desnaturalización inicial de 94°C durante 5 minutos, 35 ciclos de 94°C un minuto, 60°C 1 minuto y 72°C 1 minuto, y una elongación final de 72°C durante 10 minutos. A continuación, para comprobar el resultado de la PCR se realizó una electroforesis en gel de agarosa al 1,5% en un tampón TAE (Tris-Acético-EDTA). Al gel se le añadió una cantidad recomendada por el fabricante del reactivo SYBR Safe (Thermofisher), y un marcador de peso molecular de 100 pb DNA ladder, 100 bp plus, peqGOLD (VWR Peqlab). Los geles se analizaron en la cámara de revelado Amershan Imager 600 (GE Healthcare Life Sciences). El fragmento de la variante de interés mostró el tamaño esperado de 596 pb (Fig. 48).

Tabla 5. Oligonucleótidos empleados en la amplificación y secuenciación de la posición mitocondrial de interés.

Sentido	Secuencia 5' -> 3'	Tm	GC%	Longitud
PF	CAACTGCAACTCCAAAGCCA	59,25	50,00	596
PR	AGTGGCTGTGCAGACATTCAA	60,48	47,62	

PF: oligonucleótido forward; PR: oligonucleótido reverse Tm: temperatura de melting; GC%: porcentaje de guanina-citosina

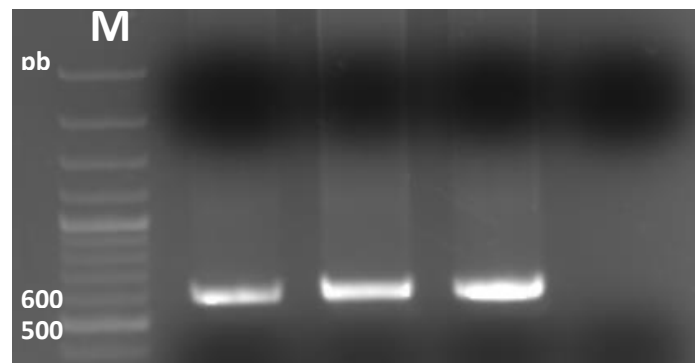


Figura 48. Electroforesis en gel de agarosa del proceso de PCR para el fragmento que contiene la posición mitocondrial de interés. M: marcador de peso molecular; pb: pares de bases.

Tras confirmar los tamaños, es necesario purificar el producto de PCR, ya que, en la muestra, además de los fragmentos de interés, también están presentes dNTPs y oligonucleótidos que pueden generar ruido durante el proceso de secuenciación. Para ello se trataron los productos de PCR con ExoSap-IT (Thermofisher). Este reactivo consta de dos enzimas, la Exonucleasa I, que desnaturaliza fragmentos monocatenarios correspondientes a oligonucleótidos que no han anillado, y una fosfatasa alcalina de camarón que elimina los dNTPs no incorporados (Fig. 49). Para realizar este paso, se añadieron 0,8  $\mu$ L de ExoSap-IT por cada 2  $\mu$ L de muestra de PCR. A continuación, se sometió la mezcla a un programa de temperatura de 37°C durante 15 minutos, para activar las enzimas, seguido de otro ciclo de temperatura a 80°C durante otros 15 minutos para inactivarlas.

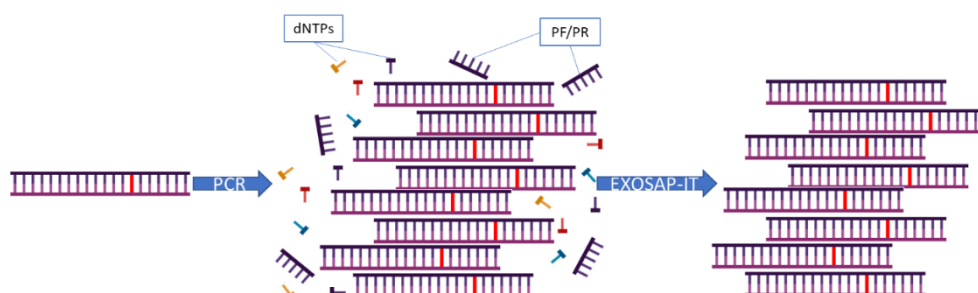


Figura 49. Representación del proceso de purificación del producto de PCR con EXOSAP-IT. dNTPs: desoxinucleótidos; PF: oligonucleótido en sentido forward; PR: oligonucleótido en sentido reverse.

Acto seguido, para llevar a cabo la reacción de secuenciación, se prepararon dos mezclas individuales de secuenciación, una para cada sentido (*Forward 5'-3'* y *Reverse 3'-5'*). Cada mezcla estaba formada por 2  $\mu\text{L}$  de tampón de reacción, 0,5  $\mu\text{L}$  del reactivo *Big dye terminator kit* (Thermofisher), 0,6  $\mu\text{M}$  (0,3  $\mu\text{L}$ ) de PF o PR del fragmento a analizar, completando con  $\text{H}_2\text{O}$  destilada hasta un volumen final de 11  $\mu\text{L}$ . La mezcla de secuenciación fue añadida a las muestras, sumando un volumen de 13,8  $\mu\text{L}$ , y éstas fueron sometidas a un programa de temperaturas con una desnaturalización inicial de 96°C durante un minuto, 25 ciclos de 96°C un minuto y cuatro minutos, y cinco segundos a 60°C. Una vez finalizada la reacción de secuenciación en el termociclador, fue necesario realizar otra purificación adicional del producto de la reacción de secuenciación, ya que pueden quedar sales y ddNTPs que pueden modificar el proceso de secuenciación Sanger. Para llevar a cabo este proceso, se empleó el kit *Montage SEQ96 sequencing reaction cleanup* (Millipore). Brevemente, las muestras se introdujeron en los correspondientes pocillos de las placas de purificación en membrana de sílice, se añadió a la muestra 25  $\mu\text{L}$  de *Wash Buffer* y se aplicó una presión negativa de 25-30 mm de Hg hasta vaciar el contenido del pocillo. Este proceso se repitió de nuevo para generar un segundo lavado y finalmente se resuspendieron los pocillos con 25  $\mu\text{L}$  de *Sequencing Buffer* y se incubó la placa en agitación durante 10 minutos.

Finalmente, las muestras se introdujeron en el secuenciador capilar ABI 3130XL, *Genetic analyzer* (Applied Biosystems). Este equipo consta de 16 capilares de 36 cm de longitud, rellenos de un soporte polimérico especial (*Performance Optimized Polymer 4, POP4*, Thermofisher) que permite la separación de fragmentos con una resolución de un nucleótido. En la secuenciación se empleó un tiempo de inyección de 18 segundos con un voltaje de 1,2 kilovoltios y 2.500 segundos de tiempo de carrera con un voltaje de 15 kilovoltios. Una vez finalizado el proceso de secuenciación, los datos fueron analizados con el software *Sequencing Analysis 7* (Applied Biosystems). La figura 50 muestra un ejemplo del resultado obtenido del proceso de secuenciación Sanger. Como se comentó al inicio del apartado, todo este proceso empleado en la cohorte de exploración OAI, también fue aplicado en las cohortes de validación externas CHECK y PROCOAC.

### *Análisis estadístico de los datos del objetivo específico número 2.*

Tras genotipar la variante mitocondrial de interés en el resto de la cohorte OAI, se analizó mediante un análisis univariante de  $\chi^2$  la distribución de los alelos entre los grupos de progresión en los criterios de dos y tres grupos de participantes. También se aplicó el análisis univariante de  $\chi^2$  en las cohortes de validación externas. Se analizó mediante regresión logística el efecto de esta variante en la progresión rápida, para ello se empleó dicha variante como variable independiente y, como factores de confusión la edad, el sexo, el IMC, la artrosis contralateral y el dolor WOMAC. Se empleó como variable dependiente el estado de progresión de artrosis (progresores rápidos frente al resto de progresores). Estos análisis se llevaron a cabo en el software *IBM-SPSS, release 25* (IBM, Armonk, NY, USA).

Finalmente, se meta-analizaron los resultados de la fase de exploración y de la fase de validación usando el programa *R v3.6.3 (The R Foundation for Statistical Computing)*. Se consideraron modelos de efectos fijos y aleatorios para el meta-análisis con datos de resultados binarios (progresores rápidos frente al resto de individuos). Se utilizó el método de la varianza inversa para calcular las estimaciones de efectos fijos y aleatorios. Se utilizó la estimación de los OR ajustados y su IC del 95 % como medidas del tamaño del efecto para la asociación entre la variante mitocondrial y la progresión rápida de artrosis de rodilla. Se calculó el índice  $I^2$  para explorar la heterogeneidad. Los resultados del meta-análisis se presentaron en un gráfico de

## Material y métodos

diagrama de bosque. Todos los cálculos se realizaron con el software *R* v3.6.3 (*The R Foundation for Statistical Computing*), utilizando el paquete *meta*.

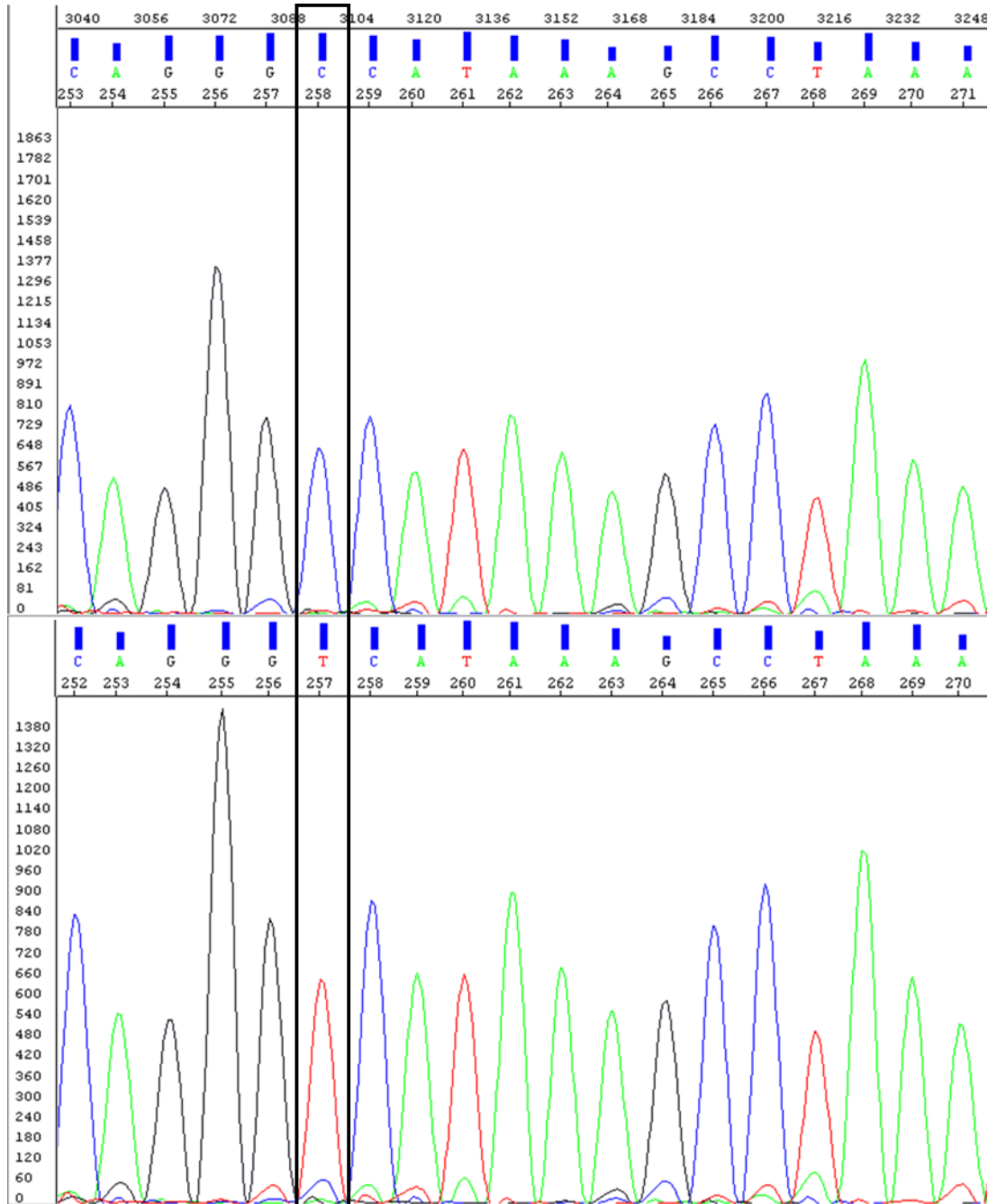


Figura 50. Ejemplo de resultados de secuenciación capilar para la posición de la variante mitocondrial de interés. En el electroferograma superior se muestra un ejemplo de la secuencia para un portador de la variante con el alelo C y, en la imagen inferior, está presente un electroferograma de un portador del alelo salvaje T.

### Material y métodos objetivo específico número 3: exploración del significado funcional de las variantes mitocondriales asociadas con el fenotipo de progresión rápida mediante análisis del transcriptoma por RNA-seq.

Debido a que en la literatura ningún trabajo ha profundizado en el efecto funcional de la variante detectada en el objetivo específico número 2, se decidió desarrollar un breve estudio funcional usando cíbridos transmitocondriales para poder entender cómo esta variante mitocondrial podía alterar el perfil de expresión del genoma. Para ello se volvió a emplear la metodología de secuenciación de segunda generación para realizar un proceso de RNA-seq (secuenciación de ARN) del transcriptoma de los cíbridos transmitocondriales. En líneas generales, el proceso es muy similar al que se realizó durante la secuenciación del ADNmt, sin embargo, son necesarios algunos pasos previos como el tratamiento con DNAsa I y la retrotranscripción. Además, los reactivos de cuantificación y de preparación de librerías son distintos.

#### *Desarrollo de las líneas de cíbridos transmitocondriales.*

Gracias a trabajos previos del grupo GIR, cuando se desarrolló esta tesis doctoral ya existía un modelo celular que se pudo emplear para evaluar el impacto de la variante mitocondrial. Se disponía de dos líneas de cíbridos, ambas portadoras del haplogrupo H, una de las líneas era portadora de la variante mitocondrial de interés, mientras que la otra línea poseía el alelo salvaje. Ambas líneas se generaron siguiendo el protocolo descrito por Chomyn (Chomyn *et al.*, 1994; Hofhaus *et al.*, 1996).

Para generar estas líneas se fusionaron células Rho-0 ( $\rho 0$ ), desprovistas de ADNmt, de la línea de osteosarcoma 143B.TK, con plaquetas de donantes sanos portadores del haplogrupo H y la variante mitocondrial o el alelo salvaje, en presencia de polietilenglicol (40%). Como resultado de este proceso, una vez constituidos los cíbridos, se pudo disponer de células con el mismo genoma nuclear, pero ligeramente diferentes en su ADNmt (Fig. 51). Las dos líneas celulares se cultivaron en el medio Dulbecco's Modified Eagle Medium suplementado con un 10% de suero bovino fetal y penicilina y estreptomycin (100 U/ml y 100  $\mu$ g/ml respectivamente, Gibco).

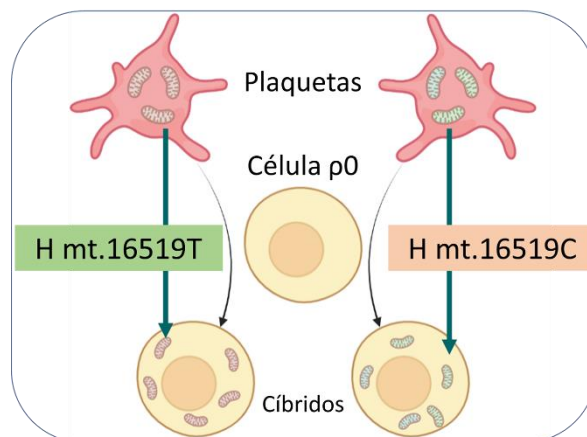


Figura 51. Esquema representativo del proceso de obtención de los cíbridos transmitocondriales empleados.

## Material y métodos

Además, con el fin de verificar que ambos cíbridos poseían los ADNmt deseados, y para comprobar las posibles variantes adicionales que diferenciaban ambas líneas, se realizó una secuenciación completa del ADNmt de ambas líneas celulares siguiendo exactamente el protocolo descrito en el objetivo específico número 2 de esta tesis doctoral. Previamente, fue necesario realizar la extracción de ADN de los cíbridos utilizando el kit Qlamp DNA Mini Kit (Qiagen®). Finalmente se seleccionaron un total de cinco clones de cada una de las líneas de cíbridos con el fin de realizar el RNA-seq.

### *Extracción de ARN de las dos líneas de cíbridos mitocondriales.*

Se empleó el siguiente protocolo de extracción de ARN para extraer el ARN de los diez clones empleados en este objetivo específico:

1. Inicialmente se eliminó el medio de cultivo y se lavaron las células con suero salino.
2. Se añadieron 500 mL de Trizol y se traspasaron las células en un tubo *Eppendorf* de 1,5 mL.
3. A las muestras se les añadieron 200  $\mu$ L de cloroformo (Sigma-Aldrich, Francia) por cada mililitro de Trizol y se agitaron manualmente durante quince segundos.
4. Las muestras se incubaron durante 3 minutos a temperatura ambiente.
5. Las mezclas resultantes se centrifugaron durante 15 minutos a 12.000 xg y a una temperatura de 4°C.
6. Se recuperó la fase superior de las muestras y se transfirieron a nuevos tubos *Eppendorf*, donde se les añadieron, a cada una de ellas, 1  $\mu$ L de glicógeno (Sigma-Aldrich, EEUU) y 500  $\mu$ L de isopropanol (Sigma-Aldrich, EEUU) por cada mililitro de Trizol.
7. Las muestras se incubaron durante 48 horas a una temperatura de -20°C para precipitar el ARN.
8. Pasado este tiempo se centrifugaron durante 10 minutos a 12.000 xg y a una temperatura de 4°C. Tras esto, se desecharon los sobrenadantes y se lavó cada pellet de ARN con 1 mL de etanol 70° por cada mililitro de Trizol.
9. Se centrifugaron los ARNs durante 5 minutos a 7.500 xg y una temperatura de 4°C, se desecharon los sobrenadantes y se lavaron los pellets con 1 mL de etanol 100° por cada mililitro de Trizol.
10. Finalmente se centrifugaron las muestras durante 5 minutos a 7.500 xg y una temperatura de 4°C, se desecharon los sobrenadantes, se dejaron secar los pellets de ARN y se resuspendieron cada uno de ellos en un tubo *Eppendorf* con 10  $\mu$ L de agua de biología molecular (Millipore, EEUU).

### *Cuantificación del ARN de los cíbridos mitocondriales.*

Tras finalizar la extracción del ARN de los cíbridos mitocondriales y, debido a la sensibilidad de la técnica de RNA-seq, fue necesario determinar la concentración de estos ARNs para ajustarse a las recomendaciones del protocolo. Del mismo modo que se hizo en la secuenciación del ADNmt, para realizar la cuantificación de ARN se utilizó un fluorómetro *Qubit 4* (ThermoFisher) y el kit *RNA HS Assay Kit* (ThermoFisher). Con esta combinación se pueden cuantificar muestras que estén en una concentración comprendida entre 4 y 200 ng/ $\mu$ L de ARN. Del mismo modo que se realizó con el kit de cuantificación de ADN *dsDNA HS kit* (ThermoFisher), se realizó la calibración del *Qubit 4* previamente a la cuantificación. A continuación, se pipetearon 2  $\mu$ L de muestra y 198  $\mu$ L del reactivo *RNA HS Assay Kit* en un tubo *Qubit Assay*

## Material y métodos

(Thermofisher), y se incubó 2 minutos en oscuridad a temperatura ambiente. Acto seguido se analizó la concentración de cada dilución con el fluorómetro *Qubit 4*.

### *Purificación del ARN con DNAsa I.*

La presencia de ADN en la muestra de ARN puede generar ruido y un contaje erróneo de genes monoexónicos. Por tanto, con el fin de eliminar cualquier contaminación con ADN de las muestras, éstas se trataron con DNAsa I (Thermofisher). El protocolo fue el siguiente:

- Se introdujo 1 µg de la muestra a un tubo de 200 µL.
- A la muestra se le añadió 1 µL de *buffer* de reacción con MgCl<sub>2</sub> 10X, 1 µL de DNAsa I (1 U/µL), y se añadió agua de biología molecular (Millipore, EEUU) hasta un volumen final de 10 µL.
- La mezcla se incubó a 37°C durante 30 minutos.
- Pasado este tiempo se incorporó a la mezcla 1 µL de ácido etilendiaminotetraacético (EDTA) 50 mM, y se incubó de nuevo a 65°C durante 10 minutos.

### *Retrotranscripción y Preparación de las librerías de RNA-seq.*

Para generar las librerías de RNA-seq se empleó el kit *Ion AmpliSeq™ Transcriptome Human Gene Expression Panel, Chef-Ready Kit* (Thermofisher). Este kit es capaz de medir de forma simultánea la expresión de más de 20.000 genes en un único ensayo a partir de 10 ng de ARN. Este kit que contiene los reactivos necesarios para generar hasta 32 muestras en cuatro *pools* de 8 muestras e incluye:

- Cuatro conjuntos de plásticos desechables.
- Cuatro cartuchos de reactivos y cuatro cartuchos de soluciones. Con los elementos y enzimas necesarios para generar las librerías.
- Cuatro placas de PCR. Cada una de estas placas lleva una mezcla de adaptadores y *barcodes* en los 8 pocillos de la primera columna.
- El panel de oligonucleótidos *Ion AmpliSeq™ Transcriptome Human Gene Expression Panel* (Thermofisher). Que está formado por un único *pool* con numerosos oligonucleótidos capaces de amplificar los fragmentos de cADN de cada muestra.

El proceso general para elaborar las librerías es muy similar al descrito durante el proceso de secuenciación de ADNmt del objetivo específico 2. Sin embargo, existen algunas diferencias que cabe puntualizar:

1. Fue necesario realizar la retrotranscripción del ARN. La retrotranscripción consistió en transformar el ARN en ADN complementario (ADNc) usando enzimas retrotranscriptasas. En este proceso se empleó el kit *SuperScript™ VILO™ cDNA Synthesis Kit* (Thermofisher); para ello se introdujo un total de 10,5 µl de ARN con una concentración de 0,95 ng/µL en la columna A de la placa de librerías del kit *Ion AmpliSeq™ Transcriptome Human Gene Expression Panel, Chef-Ready* (Fig. 41), y se añadieron a cada pocillo 3 µl del reactivo 5X *VILO RT reaction Mix* y 1,5 µl del reactivo *10X superscript III enzyme mix*. Finalmente, para obtener los cADN se introdujo la placa en el termociclador siguiendo el programa de temperaturas de 30 minutos a 42°C y 5 minutos a 85°C.
2. En este caso se emplearon 10 ng de ADNc en la preparación de las librerías.

## Material y métodos

3. Se cargaron 150  $\mu\text{L}$  del panel de oligonucleótidos a las posiciones A y B del cartucho de reactivos (fig. 42).
4. En las condiciones del programa fue necesario especificar que el número de *pools* de oligonucleótidos, en este caso es 1, el número de ciclos recomendado es 13, y el tiempo de anillamiento y extensión es de 16 minutos.

Tras el proceso, el *pool* de 8 librerías estará en la posición D del cartucho de reactivos a una concentración de 100 pM en un volumen final de 700  $\mu\text{L}$ . El *pool* de librerías se almacenó a 2-8°C durante una semana, o a -30°C -10°C para períodos de tiempo más largos.

### *Carga del chip de RNA-seq.*

Cuando se realizó la secuenciación de ADNmt, debido al tamaño del chip 530 y al proceso de generación de librerías de ADNmt, se podían introducir hasta 32 muestras en el mismo chip. Sin embargo, en el caso del RNA-seq, se generan muchas más amplicones, por lo que se necesitan chips de mayor capacidad de lectura. Los chips empleados durante la secuenciación del ADNmt permitían generar entre 15 y 20 millones de lecturas, lo cuál es más que suficiente para conseguir una buena cobertura y profundidad de lectura, por otro lado, chips de mayor tamaño como el Ion 550 Chip (Thermofisher), son capaces de generar entre 100 y 130 millones de lecturas, lo que permite alcanzar buenas coberturas, aunque se empleen paneles de oligonucleótidos que abarquen más regiones y de mayor tamaño. Los kits empleados para realizar la carga del chip fueron:

- *Ion S5 Chef Supplies* (Thermofisher). Cuatro cajas que contienen todo el material fungible necesario para el sistema Ion Chef.
- *Ion S5 Chef 550 Solutions* (Thermofisher) e *Ion 550 Chef reagents* (Thermofisher). Ambos están compuestos por 4 cartuchos con soluciones de trabajo del Ion Chef.
- *Ion 550 Chip Kit* (Thermofisher). Contiene cuatro chips, que son el soporte físico usado en esta metodología. En estos chips se pueden introducir dos librerías de 8 muestras cada una, sumando un total de 16 muestras de cADN.

A pesar de que algunos reactivos cambian, el proceso de carga del chip es prácticamente igual al que se realizó durante la secuenciación del ADNmt en el objetivo específico número 2. Las únicas diferencias radican en que:

1. Al crear el *run plan* se emplearán únicamente 16 *barcodes* ya que el protocolo de secuenciación del transcriptoma humano recomienda secuenciar, como máximo, 16 muestras en un chip 550.
2. Los *pools* de librerías que se introducen en el cartucho están formados únicamente por dos librerías de ADNc de 8 muestras cada una, y deben de estar a una concentración final de 70 pM en un volumen final de 25  $\mu\text{L}$ .

### *Secuenciación del chip de RNA-seq.*

Finalmente, los chips de cDNA se introducen en el *Ion Torrent S5 XL* (Thermofisher) para realizar la secuenciación de las muestras. Este proceso es totalmente idéntico al descrito en la secuenciación masiva del ADNmt empleada en el objetivo específico número 2.

### *Análisis bioinformático de los datos de RNA-seq.*

Tras el análisis de RNA-seq en el *Torrent Suite Software* se obtuvo un archivo .CHP para cada una de las muestras analizadas, dicho archivo contiene las secuencias de cADN



normalizadas en lecturas por millón. Los diez archivos .CHP se introdujeron en el *software Transcriptome Analysis Console software v.4* (Thermofisher) para su análisis. Inicialmente se realizó un análisis de componentes principales (PCA) sin supervisión, con el objetivo de ver cómo agrupaba a las muestras. A continuación, se realizó un análisis de los genes diferencialmente expresados (DEG, del inglés *differentially expressed genes*) entre los grupos de muestras obtenidos del PCA. Solamente se consideraron como diferencialmente expresados aquellos genes con una diferencia de expresión (Fc del inglés *Fold Change*) entre grupos  $>\pm 2$  y con un p-valor corregido por FDR  $<0,05$ . Con los resultados obtenidos se construyó un *volcano plot* y, además, el listado de DEGs considerados se cargó en la herramienta web Enrichr (<https://maayanlab.cloud/Enrichr/>) para poder detectar qué procesos biológicos podían estar afectados por estas diferencias de expresión.

### *Validación por PCR cuantitativa (qPCR).*

Con el objetivo de cerciorarnos de que los datos de RNA-seq eran correctos, se decidió validar estos resultados mediante una técnica *gold standard*, en este caso la qPCR. Para ello se seleccionaron dos de los genes con mayor diferencia de expresión entre ambos híbridos.

Para llevar a cabo esta validación por qPCR, inicialmente se emplearon los mismos ARNs de los diez clones que se secuenciaron por RNA-seq, se purificaron de nuevo con DNAsa I y se retrotranscribieron. Para hacer la qPCR se usaron dos ensayos *dual labelled Taqman* HS00174131\_m1 de IL6 (Thermofisher) y HS00153133-m1 de COX2/PTSG2 (Thermofisher) usando en ambos casos como gen de referencia GAPDH (del inglés *glyceraldehyde 3-phosphate dehydrogenase*). Para ello se introdujo en una placa de 96 pocillos de Lightcycler (Roche) una mezcla compuesta por 5  $\mu\text{L}$  de *Master Mix*, 0,5  $\mu\text{L}$  de la sonda de interés y 0,2  $\mu\text{L}$  de los PF y PR, 0,02  $\mu\text{L}$  de la sonda de GAPDH, 3,28  $\mu\text{L}$  de agua de biología molecular (Millipore, EEUU) y 1  $\mu\text{L}$  de cDNA. Tras cargar la placa, ésta se introdujo en el termociclador LightCycler® 480 II (Roche). Inicialmente se activó la polimerasa sometiendo la placa a una temperatura inicial de 95°C durante 20 segundos, a continuación, se repitieron 50 ciclos de amplificación que consistieron en 3 segundos a 95°C, seguidos de 30 segundos a 60°C. Finalmente, los datos de la qPCR se introdujeron en el software qBase plus (Biogazelle, Bélgica) para comparar los niveles de expresión de IL6 y COX2/PTSG2 entre los cinco clones portadores de la mutación mitocondrial m.16519C y los cinco clones portadores del alelo salvaje m.16519T.

### *Análisis de los niveles de IL6 en la cohorte OAI.*

Gracias a una colaboración interna entre los miembros del grupo GIR, a algunas de las muestras de los 1.095 participantes incluidos en este estudio se les midieron los niveles de IL6 en el suero de la visita basal. En total, 25 progresores rápidos y 204 progresores no rápidos tenían datos de niveles de IL6 detectados con el kit *MILLIPLEX MAP High Sensitivity Human Cytokine multiplex* (Millipore).

Con este dato, se analizaron los niveles de IL6 entre los progresores rápidos y no rápidos mediante un análisis univariado con un test de Mann-Whitney. Posteriormente, se realizó un modelo de regresión logística ajustando por variables de confusión para comparar así los niveles de IL6 entre progresores rápidos y no rápidos. Esto permitiría trasladar los resultados obtenidos en las líneas celulares de híbridos a la progresión rápida. El análisis se llevó a cabo en el *software IBM-SPSS, release 25* (IBM, Armonk, NY, USA).



Material y métodos objetivo específico número 4: analizar el impacto de las interacciones entre el genoma nuclear y mitocondrial en el desarrollo del fenotipo de artrosis de rodilla con progresión rápida, y desarrollo de un modelo predictivo del fenotipo con aplicación clínica.

En la realización de este objetivo se trabajó única y exclusivamente con la cohorte OAI, ya que era la única que contaba con un tamaño muestral relativamente grande (N=1.095) como para poder realizar el análisis de interacciones entre el ADN nuclear y el ADNmt. Para poder llevar a término este análisis se empleó un ensayo de reacción de extensión de base simple o *Single Base Extension* (SBE), el cual es un proceso de mini-secuenciación que permite analizar de forma simultánea múltiples polimorfismos genéticos, tanto nucleares como mitocondriales.

*Selección de polimorfismos nucleares.*

Se seleccionaron un total de ocho SNPs (Tabla 6), siete de los cuales se habían asociado previamente de forma robusta con la susceptibilidad a la artrosis de rodilla en estudios GWAS (Warner y Valdes 2017); y un octavo SNP relacionado de manera robusta con la reducción del espacio articular (mJSW, del inglés *minimal Joint Space Width*) (Zengini *et al.*, 2018).

Tabla 6. Información referente a los polimorfismos nucleares seleccionados.

GEN	Id del SNP	Cambio	Alelo de riesgo
TP63	rs12107036	A>G	G
FTO	rs8044769	C>T	C
GNL3	rs11177	C>T	T
DUS4L	rs4730250	A>G	G
GDF5	rs143383	C>T	T
SUPT3H	rs10948172	A>G	G
MCF2L	rs11842874	A>G	A
TGFA	rs3771501	C>T	T

SNP: single nucleotide polymorphism

*Single Base Extension (SBE).*

Esta técnica de mini-secuenciación tiene como fundamento el marcaje con fluorescencia de la posición nucleotídica de interés. Para ello, es necesario el diseño de un oligonucleótido complementario que anille en la región contigua al SNP que queremos interrogar. El siguiente paso es realizar una elongación con una enzima polimerasa y ddNTPs terminadores marcados, los cuales se diferencian de los dNTPs en que carecen del grupo OH en el carbono 3' del azúcar, y, por lo tanto, no pueden continuar la elongación. De esta forma, cuando se inicie el proceso de elongación desde el oligonucleótido contiguo al SNP, la primera y única base que se incorpora en sentido 3' es la complementaria a la posición de interés, quedando así marcada con fluorescencia y finalizando la elongación (Fig. 52).

## Material y métodos

Previamente a la realización de esta técnica es necesario el diseño de dos tipos de oligonucleótidos:

- 1- Oligonucleótidos de PCR: son parejas de oligonucleótidos que nos permiten amplificar el fragmento que contiene al SNP de interés. Este paso es fundamental para aumentar la señal de fluorescencia obtenida durante el proceso de SBE.
- 2- Oligonucleótidos de SBE: oligonucleótidos individuales que anillan de manera contigua al extremo 3' del SNP de interés, de tal manera que la primera base que incorpore la *Taq polimerasa* durante la elongación será la correspondiente a la posición que se quiere identificar.

Con esta técnica pueden analizarse de forma simultánea hasta diez posiciones en la misma reacción, reduciendo el número de ensayos necesarios por experimento y, por tanto, abaratando costes. Además, debido a su fundamento, es una técnica idónea para analizar SNPs tanto del ADN nuclear como del ADN mitocondrial.

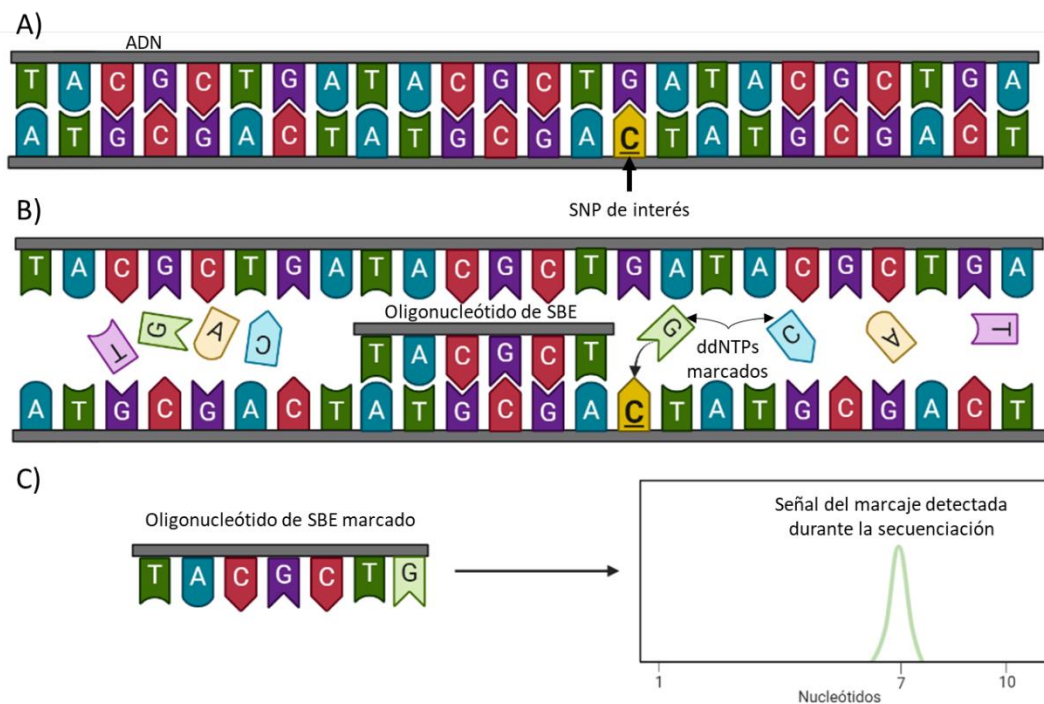


Figura 52. Esquema representativo del proceso de SBE: A) este proceso va a permitir identificar la base presente en el ADN en la posición de interés. B) en el proceso de SBE se van a emplear ddNTPs marcados con fluorescencia y un oligonucleótido específico que va a hibridar de forma contigua al extremo 3' del SNP de interés. C) como resultado de este proceso se producirá la elongación de una única base desde el oligonucleótido, permitiendo obtener un fragmento de un tamaño determinado (en este caso siete nucleótidos) con un marcaje en su extremo 3' que viene determinado por el propio ddNTP. Cuando el producto resultante del SBE se secuencie, se va a detectar una señal de fluorescencia cuya longitud vendrá determinada por el número de nucleótidos del oligonucleótido, y dependiendo de la fluorescencia emitida, sabremos cual es la base presente en dicha posición (en este caso concreto, hemos detectado una guanina, por lo que sabremos que en la otra cadena hay una citosina).

### Diseño de oligonucleótidos de PCR para analizar los SNPs nucleares.

Para evitar realizar una PCR individual para cada fragmento, se decidió poner a punto un ensayo de PCR múltiple, donde se amplifican en cada muestra y de forma simultánea los 8 SNPs de interés. Esto es fundamental para ahorrar tiempo y gastos, pero supone un riguroso diseño de los oligonucleótidos necesarios para la PCR y el SBE. Las secuencias de los fragmentos que contienen los SNPs de interés fueron obtenidas en formato FASTA de la base de datos de

SNPs del NCBI (<https://www.ncbi.nlm.nih.gov/projects/SNP/>). Para su diseño se tuvieron en cuenta los siguientes aspectos:

- Tanto el PF como el PR deben de anillar a un mínimo 70 pb de distancia de la posición de interés, con el fin de que sea posible el diseño posterior del oligonucleótido de SBE que facilitará el marcaje con fluorescencia.
- La Tm de todos los oligonucleótidos que formen parte del mismo ensayo múltiplex debe de ser la misma, o muy similar. En este caso, la Tm empleada fue de 60°C.
- El %GC no debe de ser superior a 60 ni inferior a 30 y tampoco muy dispar entre las parejas de oligonucleótidos.
- Los fragmentos diseñados deben de diferenciarse en al menos 50 pb para poder así comprobar el resultado de la PCR en una electroforesis en gel de agarosa.

Siguiendo los criterios descritos en el párrafo anterior, se usó el software *Primer-Blast* ([https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK\\_LOC=BlastHome](https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome)) y se seleccionaron los parámetros adecuados para el diseño. El listado de oligonucleótidos seleccionados para la amplificación de cada fragmento que contiene a los diferentes SNPs nucleares se muestra en la tabla 7.

Tabla 7. Parejas de oligonucleótidos de PCR diseñadas para cada uno de los SNPs nucleares asociados con la susceptibilidad a la artrosis de rodilla.

GEN	SNP	Sentido	Secuencia 5' -> 3'	Tm	GC%	Longitud
GNL3	rs11177	Forward	AGGGAAATAGGTGAACCAAGTGAC	59,99	47,83	539
		Reverse	ACCATGGGCACAAGCTCTTA	59,30	50,00	
GDF5	rs143383	Forward	CCTAAGCCTAGAGTCCCGC	58,96	63,16	716
		Reverse	GGAGTTTGGGGAGTCTCATCC	59,79	57,14	
DUS4L	rs4730250	Forward	CTGCATCTTGTTTCCACAAC	59,93	43,48	614
		Reverse	AGGAGTTACAACCACCACCAT	58,94	47,62	
FTO	rs8044769	Forward	CACCCTGCTTAAGGACTAGGT	58,81	52,38	426
		Reverse	CCCTAAGCCTCTGTTTACTG	58,09	52,38	
TP63	rs12107036	Forward	AGCCCTTAAAGGATTGGTGAGTA	59,15	43,48	343
		Reverse	TCCTTGTTATGAGAAAGGTCAGG	57,83	43,48	
SUPT3H	rs10948172	Forward	AGGAATCCAAGACCCAAAGTT	59,55	45,45	228
		Reverse	TACTCGTTTTGGACCCACAGG	61,72	57,14	
MCF2L	rs11842874	Forward	GTGCATTCTGGCAGGTGTA	59,72	50,00	793
		Reverse	GGGCAGCAGACCTTAGGAAT	60,60	55,00	
TGFA	rs3771501	Forward	TCCTGCCTAGGTGGCTAGAA	59,97	55,00	386
		Reverse	GGCAAATGGGAACATTTTG	60,17	40,00	

SNP: single nucleotide polymorphism; Tm: temperatura de melting; GC%: porcentaje de guanina y citosina.

#### Diseño de oligonucleótidos de SBE de los SNPs nucleares.

Del mismo modo que en el diseño de los oligonucleótidos de PCR, fue necesario establecer unos criterios de diseño para los oligonucleótidos de SBE con el fin de poder realizar un ensayo múltiple. Los criterios utilizados fueron los descritos por Rego y colaboradores (Rego *et al.*, 2008):

- Los oligonucleótidos de SBE deben de anillar de forma contigua al extremo 3' del SNP que se vaya a interrogar.
- El tamaño mínimo de los oligonucleótidos debe de ser de al menos 20 nucleótidos.
- La Tm de todos los oligonucleótidos que formen parte del mismo ensayo múltiplex debe de ser la misma, en este caso, debido a las condiciones del ensayo lo ideal es que sea aproximadamente de 60°C.
- El %GC no debe de ser superior a 60 ni inferior a 30.

## Material y métodos

- Para que durante el proceso de electroforesis capilar los fragmentos no pasen de forma simultánea por el detector, es necesario que dichos fragmentos presenten una diferencia de tamaño mínima entre ellos de entre 4 y 6 nucleótidos. Si por restricciones de la secuencia contigua al SNP de interés esta diferencia es inalcanzable, pueden añadirse, en el extremo 5' del oligonucleótido, polinucleótidos inespecíficos de T, A, C o repeticiones de la secuencia GATC.

Dentro de las secuencias FASTA de cada polimorfismo, se identificaron las regiones colindantes al SNP de interés para generar los oligonucleótidos de SBE. Para comprobar si los posibles oligonucleótidos cumplían las condiciones de  $T_m$  y GC% se empleó la herramienta <http://insilico.ehu.eus/tm.php>. Finalmente se seleccionaron aquellos oligonucleótidos que cumplían los criterios para su uso en el ensayo de SBE (Tabla 8).

Tabla 8. Oligonucleótidos de SBE seleccionados para cada uno de los SNPs nucleares asociados con la susceptibilidad a la artrosis de rodilla.

GEN	SNP	Sentido	Secuencia 5' -> 3'	$T_m$	GC%	Longitud
TP63	rs12107036	Reverse	5' TGACCTTTCTCTCTAGCCGGGTTATC 3'	59,5	50,0	26
FTO	rs8044769	Reverse	GGTGTATCTCACCTACAAATGAAGTGATTTTTTC	59,6	35,3	34
GNL3	rs111177	Reverse	ttAACATCATCGAAAATTAAGAAAGGAGGCTAAAAAGC	59,9	33,3	38
DUS4L	rs4730250	Forward	ccGAACTAAGATGAAATTTCCCTTTTATGCTCTGACTTTTTA	60,4	30,0	42
GDF5	rs143383	Reverse	(gact) <sub>3</sub> aaTGAAAATAACTCGTCTTGAAAGGAGAAAGCC	59,3	37,5	46
LOC	rs10948172	Forward	(gact) <sub>4</sub> GTTACTCATCAGGCAAGAGTTACGAATAGATATA	59,6	35,3	50
MCF2L	rs11842874	Forward	(gact) <sub>6</sub> AACGGCAGGCAAACTAATGTATGGTGACA	60,3	43,3	54
TGFA	rs3771501	Forward	GGAGTGTGTGAGTAAAAGGAACTCTGTATTCTA	60,8	38,24	34

Las letras en minúscula se refieren a la secuencia 5' inespecífica utilizada para ajustar los tamaños de los fragmentos. SNP: single nucleotide polymorphism;  $T_m$ : temperatura de melting; GC%: porcentaje de guanina y citosina.

### Puesta a punto de las condiciones del ensayo.

Todos los oligonucleótidos de los SNPs asociados con la susceptibilidad a la artrosis de rodilla se diseñaron para analizarse en un ensayo múltiple. Debido a que más tarde se incluyó el SNP asociado a la reducción del mJSW, rs3771501 del gen TGFA, este polimorfismo se analizó de forma individual.

Inicialmente, se realizó el análisis individual de cada pareja de oligonucleótidos. En todas las pruebas individuales se utilizaron 80 ng de 3 muestras distintas. En cada reacción de PCR individual las concentraciones finales de los reactivos de PCR fueron: 1X de tampón de reacción, 2 mM de  $MgCl_2$ , 0,2 mM de cada dNTP, 0,3  $\mu M$  de PF, 0,3  $\mu M$  de PR y 0,05 U/ $\mu L$  de *Taq polimerasa* en un volumen final de 10  $\mu L$ . Las condiciones de PCR consistieron en una desnaturalización inicial de 94°C durante 5 minutos, 35 ciclos de 94°C 1 minuto, 60°C 1 minuto y 72°C 1 minuto, y una elongación final de 72°C durante 10 minutos. A continuación se realizó una electroforesis en gel de agarosa al 1,5% en un tampón TAE. Al gel se le añadió una cantidad recomendada por el fabricante del reactivo *SYBR Safe* (ThermoFisher), y un marcador de peso molecular DNA ladder, 100 bp plus, peqGOLD (VWR Peqlab). Los geles se analizaron en la cámara de revelado *Amersham Imager 600* (GE Healthcare Life Sciences). Todos los fragmentos mostraron en los geles de agarosa el tamaño esperado.

Para los siete SNPs nucleares asociados con la susceptibilidad a la artrosis de rodilla, el siguiente paso fue la puesta a punto del ensayo de PCR múltiplex. En este proceso se emplearon 80 ng de ADN y una concentración final de 1X de tampón de reacción, 2 mM de  $MgCl_2$ , 0,2 mM

## Material y métodos

de cada dNTP, 0,3  $\mu\text{M}$  de cada uno de los siete PF, 0,3  $\mu\text{M}$  de cada uno de los siete PR y 0,05 U/ $\mu\text{L}$  de *Taq polimerasa* en un volumen final de 10  $\mu\text{L}$ . Tras realizar la mezcla, las muestras se sometieron a un programa de temperaturas en un termociclador de gradiente con el objetivo de obtener la temperatura idónea para llevar a cabo la amplificación simultánea de los siete fragmentos en una sola reacción. El programa consistió en una desnaturalización inicial de 94°C durante 5 minutos, 35 ciclos de 94°C 1 minuto, un gradiente de 50 a 60°C 1 minuto y 72°C 1 minuto, y una elongación final de 72°C durante 10 minutos. La figura 53 muestra el resultado de la electroforesis en gel, en la que se puede observar las amplificaciones de los siete SNPs iniciales a distintas temperaturas de anillamiento. Teniendo en cuenta la intensidad de las bandas, se estableció que la temperatura de anillamiento adecuada para realizar el ensayo multiplex era de 60°C.

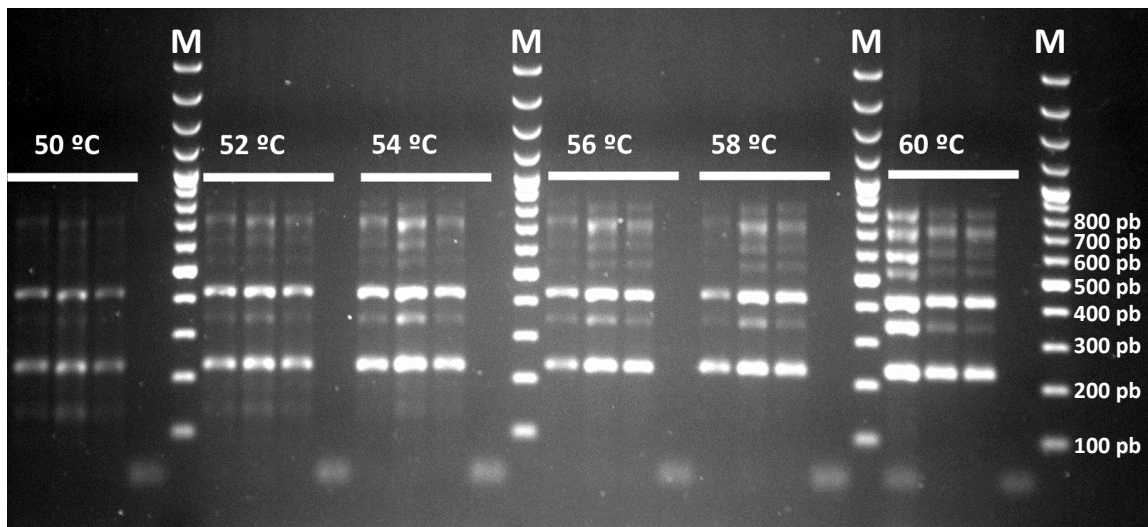


Figura 53. Ensayo de PCR múltiple de los siete fragmentos de los SNPs nucleares asociados con la susceptibilidad a la artrosis de rodilla. Las columnas con M contienen el marcador de peso molecular. El orden de los fragmentos desde arriba a abajo es el siguiente: 793 pb de rs11842874 de MCF2L, 716 pb de rs143383 de GDF5, 614 pb de rs4730250 de DUS4L, 539 pb de rs11177 de GNL3, 426 pb de rs8044769 de FTO, 343 pb de rs12107036 de TP63 y 228 pb de rs10948172 de SUPT3H.

Una vez obtenido el producto de PCR, el siguiente paso para realizar el SBE consistió en la purificación del mismo con ExoSap-IT (ThermoFisher). Para ello, se añadieron 0,8  $\mu\text{L}$  de ExoSap-IT a cada 2  $\mu\text{L}$  de muestra. A continuación se sometió la mezcla a un programa de temperatura de 37°C durante 15 minutos, para activar las enzimas, seguido de otro ciclo de temperatura a 80°C durante 15 minutos para inactivarlas.

Una vez que los productos de PCR han sido purificados, el siguiente paso en el que se realizará el marcaje específico de las bases de interés es la clave del SBE (Fig. 54). Para llevar a cabo este paso se generó una mezcla de reacción que incluía, 2,8  $\mu\text{L}$  del producto de PCR purificado, los siete oligonucleótidos de SBE a una concentración de 0,3  $\mu\text{M}$  (0,15  $\mu\text{L}$ ) y 1,5  $\mu\text{L}$  del reactivo *Snapshot Multiplex Kit* (ThermoFisher), el cual incluye un tampón, ddNTPs marcados y la *Taq polimerasa*, y agua hasta un volumen final de 7,2  $\mu\text{L}$ . En el caso del SNP 3771501 de TGFA se usó una mezcla con 1  $\mu\text{L}$  de *Snapshot Multiplex Kit*, con el correspondiente oligonucleótido SBE a una concentración de 0,3  $\mu\text{M}$  (0,15  $\mu\text{L}$ ) y agua hasta un volumen final de 7,2  $\mu\text{L}$ . En ambos ensayos las muestras se sometieron a un programa de temperaturas en un termociclador con una desnaturalización inicial de 94°C durante 1 minuto, seguido de 25 ciclos de 94°C durante 10 segundos, 60°C durante 5 segundos y 60°C durante 30 segundos.

## Material y métodos

Del mismo modo que ocurre durante el proceso de PCR, este paso no es 100% eficaz y pueden quedar ddNTPs marcados no incorporados durante la reacción de SBE, los cuales pueden finalmente generar artefactos durante la electroforesis capilar. Para eliminar dichos productos no deseados se sometió al producto de la reacción de SBE a un proceso de purificación. En este caso la enzima utilizada fue la *Fast-AP* (Thermofisher) (Fig. 54). Para ello, a los 10  $\mu\text{L}$  de la mezcla de reacción SBE se le añadieron 2  $\mu\text{L}$  de tampón, 1  $\mu\text{L}$  de enzima y 7  $\mu\text{L}$  de agua. Esta mezcla se sometió a un programa de temperaturas de 37°C durante 5 minutos para activar la enzima, seguido de un paso a 75°C durante 10 minutos para inactivarla.

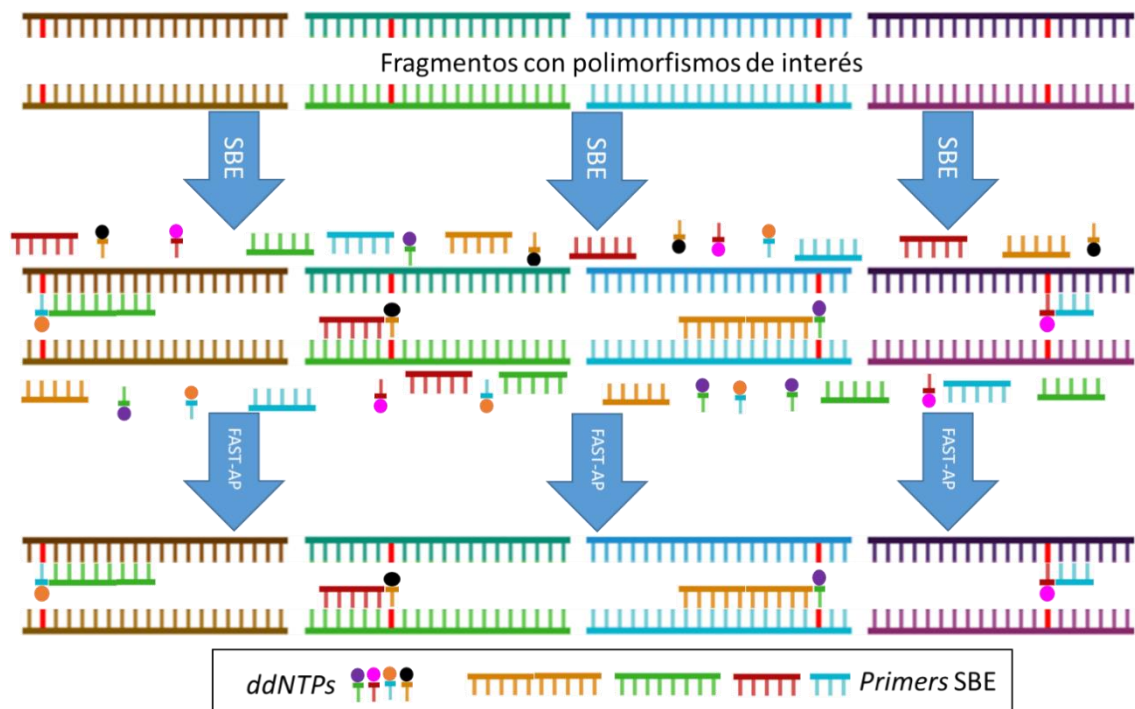


Figura 54. SBE y FAST-AP. En primer lugar, a los productos de PCR purificados se les añade la mezcla de SBE que contiene el reactivo snapshot y los oligonucleótidos de SBE. Al someter la mezcla al programa de temperaturas, estos oligonucleótidos van a anclar específicamente a regiones colindantes a las posiciones de interés de cada fragmento (marcadas en rojo), se producirá la unión de un único ddNTP marcado y se detendrá la elongación. Como resultado, todos los fragmentos van a poseer un marcaje en la posición deseada, pero también queda exceso de ddNTPs marcados con fluorescencia. Usando la enzima FAST-AP se purifica la muestra, obteniendo un producto marcado limpio.

Por último, se llevó a cabo el proceso de desnaturalización, que consiste en la preparación de la muestra para su circulación a lo largo de los capilares del secuenciador. Se mezclaron 0,5  $\mu\text{L}$  de muestra con 9,35  $\mu\text{L}$  de *Hi-Di formamide* (Thermofisher) y 0,15  $\mu\text{L}$  del marcador de peso molecular *Genescan 120 LIZ dye size standard* (Thermofisher). Se sometió a la muestra a una temperatura de 95°C durante 5 minutos para desnaturalizar el marcador de peso molecular y producir la desnaturalización de las dos hebras de ADN, y acto seguido se bajó la temperatura rápidamente hasta 4°C, evitando la rehibridación de las hebras de ADN. Tras esto, se introdujeron las muestras en el secuenciador capilar ABI 3130XL, *Genetic analyzer* (Applied Biosystems). Este equipo consta de capilares de 36 cm de longitud, rellenos de un soporte polimérico especial (*Performance Optimized Polymer 4, POP4*, Thermofisher) que permite la separación de fragmentos con una resolución de un nucleótido. En la secuenciación se empleó un tiempo de inyección de 22 segundos con un voltaje de 2 kilovoltios y, 1.000 segundos de tiempo de carrera con un voltaje de 15 kilovoltios. Una vez finalizado el proceso de secuenciación, los datos fueron analizados con el software *Genemapper v4.0* (Thermofisher) anotando la información de los SNPs tras comparar los resultados con una secuencia estándar



previamente cargada. La figura 55 muestra una representación esquemática del proceso de mini-secuenciación capilar y la figura 56 muestra un ejemplo de los datos.

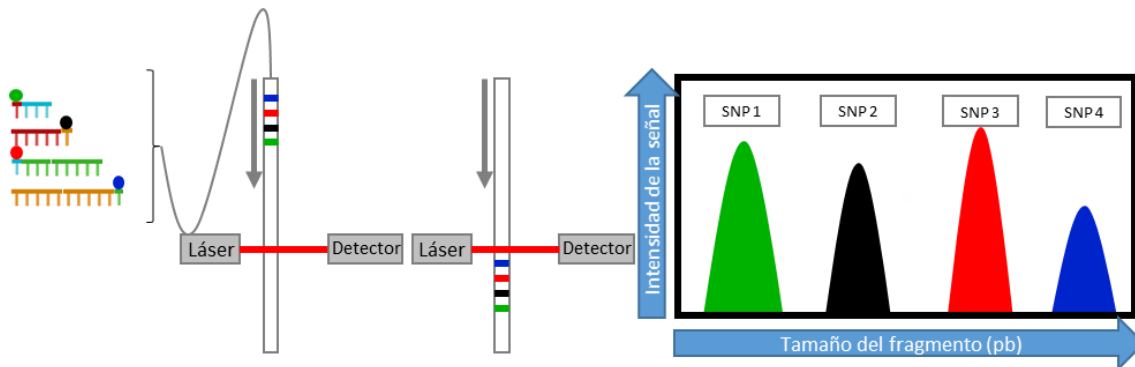


Figura 55. Representación del proceso de mini-secuenciación capilar. A la izquierda podemos ver los fragmentos con los SNPs de interés marcados con ddNTPs. Estos fragmentos son introducidos en un capilar donde migrarán del polo negativo al positivo al someterse a un campo eléctrico. Cuando los fragmentos atraviesan el láser, los ddNTPs marcados emiten una fluorescencia específica en función de la base nucleotídica que sea y esta señal es captada por un detector. Los fragmentos con menor tamaño (pb) atravesarán el láser antes que aquellas con un tamaño mayor. Finalmente se obtiene un electroferograma donde el eje Y refleja la intensidad de la señal detectada y el eje X el tamaño del fragmento detectado.

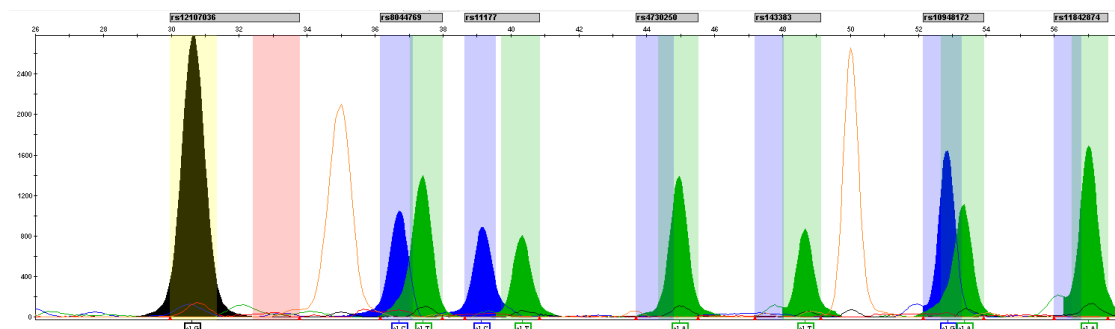


Figura 56. Ejemplo de datos de SBE para los siete SNPs nucleares iniciales. En el eje Y están los valores de intensidad de fluorescencia de cada fragmento detectado. En el eje X se muestran los tamaños de los fragmentos. En el eje X también se muestra la posición en la que aparece cada fragmento de SBE y a qué SNP nuclear pertenece cada uno.

#### Validación de la técnica de SBE por secuenciación.

Actualmente, el *gold standard* en secuenciación de ADN es la secuenciación capilar Sanger, por tanto, el siguiente paso fue el empleo de esta metodología para llevar a cabo la verificación de los resultados obtenidos con la técnica de análisis de fragmentos de SBE en el 15% del total de las muestras analizadas. Para ello, primero se purificaron con Exosap-IT los fragmentos de PCR que contenían a cada uno de los ocho SNPs nucleares incluidos en el estudio. Después, se prepararon un total de dieciséis mezclas individuales de secuenciación, ocho para cada sentido (*Forward* 5'-3' y *Reverse* 3'-5'). Cada mezcla estaba formada por 2  $\mu$ L de tampón de reacción, 0,5  $\mu$ L del reactivo *Big dye terminator kit* (Thermofisher), 0,6  $\mu$ M (0,3  $\mu$ L) de PF o PR del fragmento a analizar, completando con H<sub>2</sub>O destilada hasta un volumen final de 11  $\mu$ L. La mezcla de secuenciación fue añadida a las muestras, sumando un volumen de 13,8  $\mu$ L. Finalmente fueron sometidas a un programa de temperaturas con una desnaturalización inicial de 96°C durante 1 minuto, 25 ciclos de 96°C 1 minuto y 4 minutos y 5 segundos a 60°C. Una vez finalizado el programa, la muestra fue purificada con el kit *Montage SEQ96 sequencing reaction cleanup* (Millipore); brevemente, se introdujeron las muestras en los correspondientes pocillos de las placas de purificación en membrana de sílice, se añadió a la muestra 25  $\mu$ L de *Wash Buffer*

## Material y métodos

y se aplicó una presión negativa de 25-30 mm de Hg hasta vaciar el contenido del pocillo. Este proceso se repitió de nuevo para generar un segundo lavado y finalmente se resuspendieron los pocillos con 25 µL de *Sequencing Buffer* y se incubó la placa en agitación durante 10 minutos. A continuación, las muestras fueron traspasadas a nuevas placas para realizar su secuenciación en el ABI 3130XL, *Genetic analyzer* (Applied Biosystems), con un tiempo de inyección de 18 segundos y un voltaje de 1,2 kilovoltios y, un tiempo de carrera de 2.500 segundos a 15 kilovoltios.

### Asignación de los haplogrupos mitocondriales.

En este trabajo se asignaron los haplogrupos mitocondriales caucásicos mediante la técnica de SBE descrita en los apartados previos de este objetivo específico. La determinación de los haplogrupos se llevó a cabo utilizando el protocolo descrito por Rego y colaboradores (Rego *et al.*, 2008). La tabla 9 muestra los oligonucleótidos de PCR y de SBE utilizados en este procedimiento. En la fase de amplificación se emplearon 70 ng de ADN total, y una concentración final de tampón de 1X, 2 mM de MgCl<sub>2</sub>, 0,2 mM de cada dNTP, 0,05 U/µL de enzima *Taq polimerasa* y 0,5 µM de los oligonucleótidos de la posición m.12308, 0,36 µM de los oligonucleótidos de las posiciones m.4580 y m.4216, y 0,3 µM del resto de oligonucleótidos en un volumen final de 10 µL. Las muestras fueron sometidas a un programa de temperaturas con una desnaturalización inicial de 94°C durante 5 minutos, 45 ciclos de 94°C 1 minuto, 60°C 1 minuto y 72°C 1 minuto, y una elongación final de 72°C durante 10 minutos.

Tabla 9. Oligonucleótidos de PCR y SBE empleados en la determinación de los haplogrupos caucásicos.

#### Oligonucleótidos PCR

Posición mtDNA	Sentido	Secuencia 5' -> 3'	Tm	GC%	Longitud
m.7028	Forward	CTGACTGGCATTGTATTAGCA	53,0	42,9	473
	Reverse	GTATACGGGTTCTTGAATG	50,8	45,0	
m.14766	Forward	GAGAAGGCTTAGAAGAAAACCCAC	57,9	48,0	349
	Reverse	GTGGGCGATTGATGAAAAGGC	57,4	52,4	
m.10398	Forward	GGCCTATGAGTGACTACAAAAA	52,8	40,9	163
	Reverse	TATTCCTAGAAGTGAGATGGT	49,6	38,1	
m.4580	Forward	CCTACCACTCACCTAGCATTAC	56,5	52,2	935
	Reverse	TAGGAATGCGGTAGTAGTTAG	51,0	42,9	
m.12308	Forward	CAACCCGACATCATTACCGGGT	61,3	56,5	307
	Reverse	GGGTTAACGAGGGTGGTAAGG	56,9	57,1	
m.4216	Forward	CCTACCACTCACCTAGCATTAC	56,5	52,2	357
	Reverse	GCGAGCTTAGCGCTGTGATGAG	60,9	59,1	

#### Oligonucleótidos SBE

Posición mtDNA	Sentido	Secuencia 5' -> 3'	Tm	GC%	Longitud
m.7028	Forward	ACACGACACGTACTACGTTGTAGC	59,4	50,0	24
m.14766	Reverse	cgatcATGAGTGGTTAATTAATTTTATTAGGGGGTTA	59,1	32,4	37
m.10398	Forward	ataTATGAGTGACTACAAAAAGGATTAGACTGA	56,2	30,3	33
m.4580	Forward	(at) <sub>7</sub> TTTTTACCTGAGTAGGCCTAGAAATAAACAT	56,9	21,7	46
m.12308	Forward	(tacg)5aCCATTGGTCTTAGGCCCCAA	69,6	51,2	41
m.4216	Forward	cgCCA CTCA CCCTAGCATTACTTATATGA	60,1	44,8	29

Tm: temperatura de melting; GC%: porcentaje de guanina y citosina.

## Material y métodos

Al finalizar la PCR, las muestras siguieron el mismo procedimiento que en el análisis de los SNPs nucleares; así, en primer lugar, se analizó el resultado mediante un gel de agarosa (Fig. 57). A continuación, se purificaron los productos de PCR con ExoSap-IT. Seguidamente, para el ensayo de SBE se tomaron 1,5  $\mu$ L de reactivo de *snapshot* y una concentración final de oligonucleótidos de SBE de 0,4  $\mu$ M para las posiciones m.12308 y m.4216, y de 0,3  $\mu$ M para el resto de oligonucleótidos, todo en un volumen final de 7,2  $\mu$ L. A esta mezcla se añadió un volumen de muestra de 2,8  $\mu$ L tratada con ExoSap-IT. Tras este proceso se continuó con la purificación mediante *Fast-AP* y la consiguiente desnaturalización. Por último, las muestras se introdujeron en el secuenciador para su análisis. Una vez realizado el procedimiento se procedió al análisis de los datos usando el software *Genemapper v4.0*. La tabla 10 muestra las combinaciones de SNPs que dan lugar a los haplogrupos caucásicos, y la figura 58 muestra ejemplos de los espectros obtenidos por SBE para cada haplogrupo.

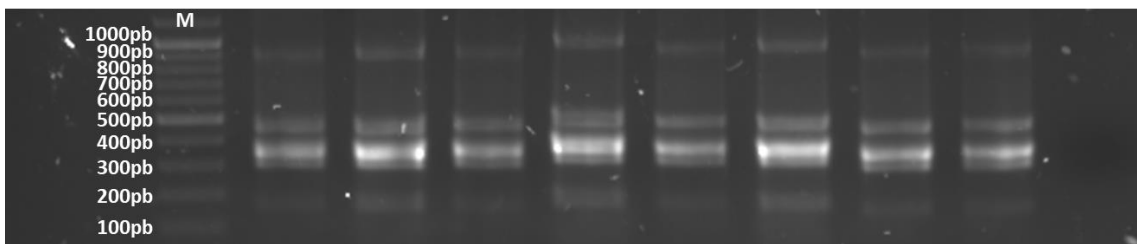


Figura 57. Ejemplo de PCR múltiplex de haplogrupos caucásicos. El orden de los fragmentos desde arriba a abajo es el siguiente: 935 pb del fragmento de la posición m.4580, 473 pb del fragmento de la posición m.7028, 357 pb del fragmento de la posición m.4216, 349 pb del fragmento de la posición m.14766, 307 pb del fragmento de la posición m.12308 y 163pb del fragmento de la posición m.10398.

Tabla 10. Clasificación de los haplogrupos Caucásicos. Cualquier combinación de SNPs que no esté representada en la tabla se designó como grupo "Otros".

Haplogrupo	Posición mtDNA					
	m.7028	m.4216	m.10398	m.12308	m.14766	m.4580
H	C	T	A	A	C	G
V	T	T	A	A	C	A
K	T	T	G	G	T	G
U	T	T	A	G	T	G
T	T	C	A	A	T	G
J	T	C	G	A	T	G

### Análisis estadístico de los datos del objetivo específico número 4.

Tras genotipar las variantes genéticas nucleares seleccionadas y los haplogrupos mitocondriales en la cohorte OAI, primero se llevó a cabo un análisis univariante de  $\chi^2$  individual para analizar la distribución de los alelos entre los grupos de progresión para cada uno de los ocho SNPs nucleares seleccionados, y otro análisis para analizar la distribución de los clusters mitocondriales en los grupos de progresión. El segundo análisis se centró en la búsqueda de interacciones entre SNPs nucleares y clusters mitocondriales, para ello se estratificó la población de estudio de la cohorte OAI en base a su cluster mitocondrial y, se reanalizó por  $\chi^2$  la distribución de los alelos entre los grupos de progresión para cada uno de los ocho SNPs nucleares seleccionados. Los p-valores obtenidos se ajustaron mediante una corrección de Bonferroni debido a las múltiples comparaciones. Se analizó mediante regresión logística el efecto de las posibles interacciones en la progresión rápida de rodilla ajustando el modelo por las variables clásicas de confusión: edad, sexo, IMC, artrosis contralateral, lesión previa y el dolor

## Material y métodos

WOMAC. Estos análisis se llevaron a cabo en el software IBM-SPSS, release 25 (IBM, Armonk, NY, USA).

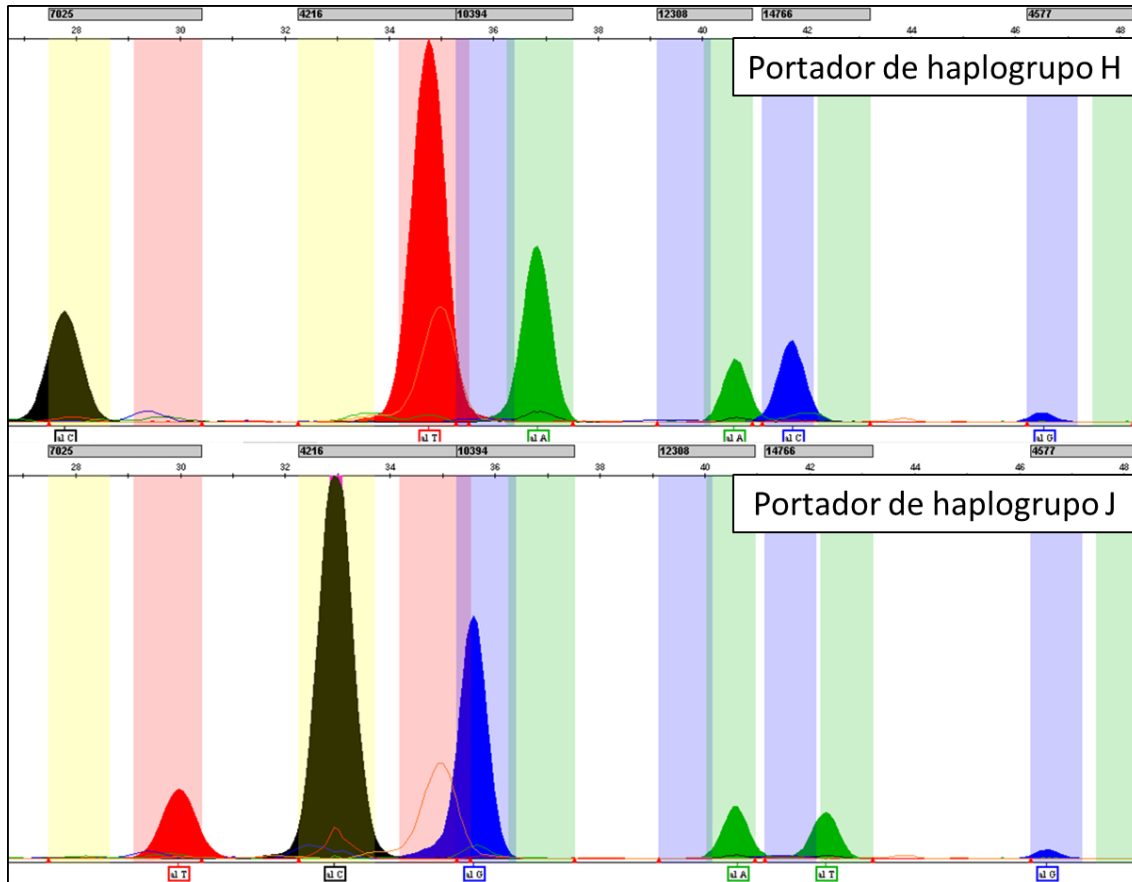


Figura 58. Ejemplo de datos de SBE para los haplogrupos mitocondriales. En la imagen superior se observa un electroferograma de un individuo portador de haplogrupo H, mientras que en la imagen inferior se observa un electroferograma de un portador de haplogrupo J.

*Desarrollo del modelo predictivo del fenotipo de artrosis de rodilla con progresión rápida usando variables clínicas y genéticas.*

Finalmente, se diseñó una herramienta de aplicación clínica capaz de identificar, de forma temprana, a aquellos individuos con un elevado riesgo de desarrollar el fenotipo de artrosis de rodilla con progresión rápida. Para llevarlo a cabo se generó un modelo predictivo que incluye tanto variables clínicas como las variantes genéticas que obtuvieron valores significativos en esta tesis doctoral.

Para desarrollar dicha herramienta predictiva, se empleó una regresión logística basada en Ecuaciones de Estimación Generalizadas (EEG), aplicando el modelo genético aditivo. Se incluyeron las siguientes variables en el modelo: edad, sexo, IMC, artrosis contralateral, lesión previa, el dolor WOMAC, el alelo de riesgo del polimorfismo nuclear de interés, el cluster mitocondrial UK, la interacción genética entre estos dos elementos y la variante mitocondrial detectada por secuenciación masiva. Todos estos análisis se realizaron con el software R software v3.6.3 (The R Foundation for Statistical Computing).

Debido al reducido tamaño muestral de las cohortes CHECK y PROCOAC, y a la ausencia de la variable de lesión previa, fue imposible aplicar los modelos de EEG en estas cohortes, por lo que se optó por validación mediante técnicas de validación cruzada dentro de la propia

## *Material y métodos*

cohorte OAI. En la validación cruzada se tomó un valor  $k=10$ , generándose así diez subgrupos dentro de la cohorte OAI. La precisión del modelo se midió a través del AUC (del inglés *Area under the ROC curve*), tomando como umbral un punto de corte estándar (0,5).

Finalmente, se diseñó un nomograma que sirve como una herramienta visual de aplicación clínica que permite estimar el riesgo que tiene un individuo de desarrollar el fenotipo de progresión rápida en base a sus variables clínicas y genéticas. Dicho nomograma se calibró mediante *bootstrapping* con 1.000 remuestreos. Tanto la validación cruzada, como el nomograma, se desarrollaron con el R software v3.6.3 (The R Foundation for Statistical Computing).



# *Resultados*





## Resultados del objetivo específico número 1: establecimiento de los criterios de progresión y selección de los pacientes.

Esta tesis doctoral se inició con una selección inicial de participantes dentro de la cohorte de exploración OAI siguiendo los criterios definidos por Driban y colaboradores (Driban *et al.*, 2016), donde los individuos parten de un grado KL basal de 0 o 1 y alcanzan un grado KL  $\geq 3$  en 48 meses o menos. Con la aplicación de estos criterios se seleccionaron un total de 184 progresores rápidos. A continuación, utilizando el *software* Epidat v4.0 (Xunta de Galicia) se seleccionaron, de manera aleatoria, 138 y 128 sujetos no-progresores y progresores no-rápidos respectivamente. Así, la cohorte inicial de exploración que sería posteriormente sometida a los procesos de secuenciación masiva del ADNmt quedó constituida por un total de 450 muestras.

Posteriormente, con ayuda de la supervisión de los facultativos del servicio de Reumatología del Hospital, se decidió incluir para este estudio a aquellos individuos que, mostrando basalmente un grado KL 2 desarrollaban, al cabo de 48 meses o menos, un grado KL 4 o eran sometidos a una cirugía de reemplazo articular. De esta forma, tal y como se describe en la metodología de esta tesis, con la inclusión de estos nuevos pacientes se constituirían los criterios GIR-blanco de progresión rápida. En base a estos criterios, y con el fin de alcanzar un adecuado tamaño muestral, se seleccionaron un total de 268 progresores rápidos y, siguiendo una proporción aproximada de 1:3 entre casos y controles, 827 progresores no-rápidos (repartidos entre 471 sujetos no-progresores y 356 progresores no-rápidos). De esta forma, el número total de sujetos adscritos a los criterios GIR-blanco ascendió a 1.095. Por su parte, el incremento del tamaño muestral para conseguir una proporción de 1:3 en base a los criterios Driban, hizo que el número de sujetos ascendiese a 852, de los que 189 eran progresores rápidos y 663 no-rápidos (repartidos entre 274 sujetos no-progresores y 379 progresores no-rápidos).

En ambos criterios, y con el fin de realizar diversos análisis, las categorías de progresores no-rápidos y no-progresores se colapsaron en una única categoría de individuos no rápidos. Esto permitió realizar la comparación de individuos progresores rápidos frente al resto, pudiendo analizar de forma más precisa el efecto de las distintas variables estudiadas en el desarrollo del fenotipo de artrosis de rodilla con progresión rápida. Los números finales correspondientes al tamaño muestral en base a ambos criterios se encuentran descritos en la tabla 11.

Tabla 11. Recuento de individuos seleccionados dentro de los criterios de progresión Driban y GIR-blanco en la cohorte OAI. En la parte superior se muestra el recuento de los criterios con tres grupos de progresión y, en la parte inferior, los criterios con dos grupos de progresión.

Criterio	Tres grupos de progresión			Total
	No rápidos	Rápidos	No progresores	
Driban	274	189	389	852
GIR-blanco	356	268	471	1095
Criterio	Dos grupos de progresión			Total
	No rápidos	Rápidos		
Driban	663	189		852
GIR-blanco	827	268		1095

Las tablas 12, 13, 14 y 15 muestran, respectivamente, los análisis univariantes de las variables clínicas y demográficas de las poblaciones de individuos seleccionados en base a los

## Resultados

criterios Driban y GIR-blanco dentro de la cohorte del OAI, y que formarán parte posteriormente de los modelos de predicción.

Tabla 12. Variables clínicas y demográficas de los individuos seleccionados en base al criterio de progresión Driban de tres grupos de progresión.

Criterio Driban		Tres grupos de progresión				Comparación	P-valor	P-valor ajustado
Variables cuantitativas		No rápidos	Rápidos	No progresores				
Edad		60,340 ± 8,620	64,870 ± 8,310	61,960 ± 9,399		P.N.R. vs N.P.	0,039*	0,116*
						P.N.R. vs P.R.	<0,001*	<0,001*
						N.P. vs P.R.	<0,001*	<0,001*
IMC (kg/m <sup>2</sup> )		28,515 ± 4,466	29,398 ± 4,519	26,809 ± 4,160		P.N.R. vs N.P.	<0,001*	<0,001*
						P.N.R. vs P.R.	0,056	0,169
						N.P. vs P.R.	<0,001*	<0,001*
WOMAC		1,756 ± 2,506	2,086 ± 2,986	1,251 ± 2,068		P.N.R. vs N.P.	0,005*	0,016*
						P.N.R. vs P.R.	0,406	1
						N.P. vs P.R.	0,001*	0,002*
Variables cualitativas		No rápidos	Rápidos	No progresores	Total	P-valor		
Sexo	Hombre	94 (34,3%)	70 (37%)	172 (44,2%)	336	0,027*		
	Mujer	180 (65,7%)	119 (63%)	217 (55,8%)	516			
Contralateralidad	No	169 (61,7%)	84 (44,4%)	301 (77,4%)	554	<0,001*		
	Sí	105 (38,3%)	105 (55,6%)	88 (22,6%)	298			
Lesión previa	No	172 (62,8%)	102 (54%)	291 (74,8%)	565	<0,001*		
	Sí	102 (37,2%)	87 (46%)	98 (25,2%)	287			

Las variables cuantitativas fueron analizadas mediante el test de Kruskal-Wallis, y los p-valores corregidos mediante Bonferroni; las variables cualitativas fueron analizadas mediante un análisis de  $\chi^2$ . Valores de las variables cuantitativas: valor promedio  $\pm$  desviación estándar; valores de las variables cualitativas: número de casos con la frecuencia de cada grupo entre paréntesis; P.R: progresores rápidos; P.N.R: progresores no rápidos; N.P: no progresores; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; \*: significación estadística <0,05.

Tabla 13. Variables clínicas y demográficas de los individuos seleccionados en base al criterio de progresión Driban de dos grupos de progresión.

Criterio Driban		Dos grupos de progresión			P-valor	OR	IC 95%	
Variables cuantitativas		No rápidos	Rápidos				Inf	SUP
Edad		61,290 ± 9,113	64,870 ± 8,310		<0,001*	1,045	1,026	1,064
IMC (kg/m <sup>2</sup> )		27,514 ± 4,367	29,398 ± 4,519		<0,001*	1,097	1,058	1,138
WOMAC		1,460 ± 2,271	2,086 ± 2,986		0,002*	1,098	1,033	1,166
Variables cualitativas		No rápidos	Rápidos	Total	P-valor	OR	IC 95%	
							Inf	SUP
Sexo	Hombre	266 (40,1%)	70 (37%)	336	0,444	1,139	0,816	1,590
	Mujer	397 (59,9%)	119 (63%)	516				
Artrosis Contralateral	No	470 (70,9%)	84 (44,4%)	554	<0,001*	3,044	2,184	4,244
	Sí	193 (29,1%)	105 (55,6%)	298				
Lesión previa	No	463 (69,8%)	102 (54%)	565	<0,001*	1,975	1,419	2,748
	Sí	200 (30,2%)	87 (46%)	287				

Se analizó el efecto de las variables cuantitativas mediante un análisis univariado para la obtención del correspondiente OR e IC95%, las variables cualitativas fueron analizadas mediante un análisis de  $\chi^2$ ; Valores de las variables cuantitativas: valor promedio  $\pm$  desviación estándar; valores de las variables cualitativas: número de casos con la frecuencia de cada grupo entre paréntesis; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; \*: significación estadística <0,05; OR: odds ratio; IC 95%: intervalo de confianza al 95% con los límites inferior y superior.

Tabla 14. Variables clínicas y demográficas de los individuos seleccionados en base al criterio de progresión GIR-blanco de tres grupos de progresión.

Criterio GIR-blanco	Tres grupos de progresión				Comparación	P-valor	P-valor ajustado	
	Variables cuantitativas		No rápidos	Rápidos				No progresores
Edad	60,880 ± 8,653		65,090 ± 8,240	61,960 ± 9,224		P.N.R. vs N.P.	0,108	0,323
						P.N.R. vs P.R.	<0,001*	<0,001*
						N.P. vs P.R.	<0,001*	<0,001*
IMC (kg/m <sup>2</sup> )	28,660 ± 4,434		29,352 ± 4,489	27,169 ± 4,303		P.N.R. vs N.P.	<0,001*	<0,001*
						P.N.R. vs P.R.	0,095	0,285
						N.P. vs P.R.	<0,001*	<0,001*
WOMAC	1,928 ± 2,678		2,493 ± 3,117	1,319 ± 1,805		P.N.R. vs N.P.	0,006*	0,018*
						P.N.R. vs P.R.	0,371	1
						N.P. vs P.R.	0,001*	0,002*
Variables cualitativas s		No rápidos	Rápidos	No progresores	Total	P-valor		
Sexo	Hombre	125 (35,5%)	96 (35,8%)	208 (44,2%)	429	0,013*		
	Mujer	231 (64,9%)	172 (64,2%)	263 (55,8%)	666			
Artrosis contralateral	No	202 (56,7%)	94 (35,5%)	306 (65%)	602	<0,001*		
	Sí	154 (43,3%)	171 (64,5%)	165 (35%)	490			
Lesión previa	No	224 (62,9%)	145 (54,3%)	343 (72,8%)	712	<0,001*		
	Sí	132 (37,1%)	122 (45,7%)	128 (27,2%)	382			

Las variables cuantitativas fueron analizadas mediante el test de Kruskal-Wallis, y los p-valores corregidos mediante Bonferroni; las variables cualitativas fueron analizadas mediante un análisis de  $\chi^2$ . Valores de las variables cuantitativas: valor promedio  $\pm$  desviación estándar; valores de las variables cualitativas: número de casos con la frecuencia de cada grupo entre paréntesis; P.R: progresores rápidos; P.N.R: progresores no rápidos; N.P: no progresores; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; \*: significación estadística <0,05.

Tabla 15. Variables clínicas y demográficas de los individuos seleccionados en base al criterio de progresión GIR-blanco de dos grupos de progresión.

Criterio GIR-blanco	Dos grupos de progresión				P-valor	OR	IC 95%	
	Variables cuantitativas		No rápidos	Rápidos			Inf	SUP
Edad	61,500 ± 8,993		65,090 ± 8,240	<0,001*	1,047	1,030	1,064	
IMC (kg/m <sup>2</sup> )	27,811 ± 4,419		29,352 ± 4,489	<0,001*	1,079	1,046	1,112	
WOMAC	1,582 ± 2,396		2,493 ± 3,117	<0,001*	1,127	1,073	1,184	
Variables cualitativas		No rápidos	Rápidos	Total	P-valor	OR	IC 95%	
							Inf	SUP
Sexo	Hombre	333 (40,3%)	96 (35,8%)	429	0,195	1,208	0,908	1,607
	Mujer	494 (59,7%)	172 (64,2%)	666				
Artrosis Contralateral	No	508 (61,4%)	94 (35,5%)	602	<0,001*	2,897	2,172	3,864
	Sí	319 (38,6%)	171 (64,5%)	490				
Lesión previa	No	567 (68,6%)	145 (54,3%)	712	<0,001*	1,835	1,384	2,433
	Sí	260 (31,4%)	122 (45,7%)	382				

Se analizó el efecto de las variables cuantitativas mediante un análisis univariado para la obtención del correspondiente OR e IC95%, las variables cualitativas fueron analizadas mediante un análisis de  $\chi^2$ ; Valores de las variables cuantitativas: valor promedio  $\pm$  desviación estándar; valores de las variables cualitativas: número de casos con la frecuencia de cada grupo entre paréntesis; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; \*: significación estadística <0,05; OR: odds ratio; IC 95%: intervalo de confianza al 95% con los límites inferior y superior.

En los sujetos seleccionados en base a los criterios Driban de tres grupos de progresión, se encontraron diferencias significativas a nivel de todas las variables analizadas, siendo el grupo de progresores rápidos los pacientes con más edad (64,870  $\pm$  8,310), mayor IMC (29,398  $\pm$  4,519) y el dolor WOMAC (2,086  $\pm$  2,986); asimismo, presentar artrosis contralateral (55,6%) y haber sufrido una lesión previa en la rodilla diana (46%) son factores de riesgo para desarrollar el

## Resultados

fenotipo de progresión rápida. En el análisis de los criterios Driban de dos grupos de progresión, de nuevo se encontraron diferencias significativas a nivel de todas las variables analizadas, siendo el grupo de progresores rápidos los pacientes con más edad ( $64,870 \pm 8,310$ ), mayor IMC ( $29,398 \pm 4,519$ ) y dolor WOMAC ( $2,086 \pm 2,986$ ); también la artrosis contralateral (55,6%) y la lesión previa en la rodilla diana (46%) volvieron a mostrar una mayor frecuencia en los progresores rápidos.

En los sujetos seleccionados en base a los criterios GIR-blanco de tres grupos de progresión, se encontraron diferencias significativas a nivel de todas las variables analizadas, siendo el grupo de progresores rápidos los pacientes con más edad ( $65,090 \pm 8,240$ ), mayor IMC ( $29,352 \pm 4,489$ ) y dolor WOMAC ( $2,493 \pm 3,117$ ); asimismo, presentar artrosis contralateral (64,5%) y haber sufrido una lesión previa en la rodilla diana (45,7%) son factores de riesgo para desarrollar el fenotipo de progresión rápida. En el análisis en base a los criterios GIR-blanco de dos grupos de progresión, otra vez se encontraron diferencias significativas a nivel de todas las variables analizadas, siendo el grupo de progresores rápidos los pacientes con más edad ( $65,090 \pm 8,240$ ), mayor IMC ( $29,352 \pm 4,489$ ) y dolor WOMAC ( $2,493 \pm 3,117$ ); la frecuencia del sexo femenino (63%), de la artrosis contralateral (64,5%) y de la lesión previa en la rodilla diana (45,7%) continuaban estando sobrerrepresentadas en los individuos progresores rápidos, si bien el sexo no alcanzó la significación estadística.

Teniendo en cuenta que las variables demográficas analizadas en el apartado anterior muestran unos valores y una tendencia muy similar tanto en los criterios Driban como en los GIR-blanco y, que ambos criterios comparten una población común ( $N=852$ ), se decidió analizar la correlación en la clasificación de los individuos entre los criterios con el fin de decidir aplicar únicamente el criterio GIR-blanco en los futuros análisis, ya que, además de tener los mismos participantes que el criterio Driban, se incrementaba el tamaño muestral a 1.095. Para llevar a cabo este análisis se usó el estadístico de correlación Kappa en un análisis univariado tanto para los criterios de tres (Tabla 16) como los de dos grupos de progresión (Tabla 17).

Tabla 16. Análisis de la correlación de la clasificación de la progresión rápida en los criterios de tres grupos de progresión.

Tres grupos de progresión		Driban			Medida de acuerdo Kappa
		No rápidos	Rápidos	No progresores	
GIR-blanco	No rápidos	272	0	0	0,996
	Rápidos	2	189	0	
	No progresores	0	0	389	

Tabla 17. Análisis de la correlación de la clasificación de la progresión rápida en los criterios de dos grupos de progresión.

Dos grupos de progresión		Driban		Medida de acuerdo Kappa
		No rápidos	Rápidos	
GIR-blanco	No rápidos	661	0	0,993
	Rápidos	2	189	

En ambos análisis, el coeficiente de correlación Kappa es muy elevado, rozando el valor de 1, es decir, existe una correlación casi absoluta entre ambos criterios. En el caso de los criterios de tres grupos de progresión el valor del coeficiente de correlación es de 0,996, mientras que en los de dos grupos, éste alcanzó un valor de 0,993. Los dos análisis mostraron

que, de las 852 muestras comunes entre los criterios, 850 se clasifican exactamente igual, mientras que existen dos individuos que en el criterio Driban se clasificaron como individuos no rápidos, pero que con el criterio GIR-blanco se clasifican como progresores rápidos.

Teniendo en cuenta la elevada correlación en la clasificación entre ambos criterios, la similitud de las variables demográficas entre éstos y el notable aumento del tamaño muestral que aporta el criterio GIR-blanco, se decidió utilizar el criterio GIR-blanco para los análisis descritos en los siguientes apartados de la tesis.

Con el objetivo de validar los hallazgos de la cohorte de exploración, se trabajó con un subgrupo de muestras pertenecientes a las cohortes CHECK y PROCOAC. Aplicando el criterio GIR-blanco, en la cohorte CHECK se seleccionaron un total de 66 individuos progresores rápidos, 153 progresores no-rápidos y 158 no progresores, sumando un total de 377 participantes. Por otro lado, en la cohorte PROCOAC, aplicando el criterio GIR-blanco, el tamaño muestral ascendió hasta los 329 participantes, donde 41 eran progresores rápidos, 83 no-rápidos y 205 no progresores. Debido al escaso tamaño muestral de las dos cohortes, y a las similitudes en las frecuencias de progresión rápida, se decidió agrupar estas dos cohortes en una única cohorte de validación, denominada cohorte europea de validación, permitiendo así mejorar el poder estadístico del estudio (Tabla 18). El resumen de los análisis univariantes de las variables demográficas en las cohortes de validación para el criterio GIR-blanco de tres y dos grupos de progresión está presente en las tablas 19 y 20 respectivamente.

En la cohorte de validación europea, en los sujetos seleccionados en base a los criterios GIR-blanco de tres grupos de progresión, se encontraron menos diferencias significativas a nivel de todas las variables analizadas, siendo los progresores rápidos los que presentaron un porcentaje de artrosis contralateral significativamente mayor (30,5%). En los criterios de dos categorías, también mostraron, además de un mayor porcentaje de contralateralidad (30,8%), un mayor porcentaje de mujeres (86,9%) y mayores valores de dolor WOMAC ( $5,933 \pm 3,659$ ).

Tabla 18. Recuento de individuos seleccionados dentro de los criterios de progresión Driban y GIR-blanco en la cohorte de europea de validación. En la parte superior se muestra el recuento de los criterios con tres grupos de progresión y, en la parte inferior, los criterios con dos grupos de progresión.

Criterio	Tres grupos de progresión			
	No rápidos	Rápidos	No progresores	Total
GIR-blanco	236	107	363	706
Criterio	Dos grupos de progresión			
	No rápidos	Rápidos	Total	
GIR-blanco	599	107	706	

## Resultados

Tabla 19. Variables clínicas y demográficas de los individuos seleccionados en base al criterio de progresión GIR-blanco de tres grupos de progresión de la cohorte europea de validación.

Variables		Tres grupos de progresión			P-valor	OR	P-valor ajustado
Cuantitativas		No rápidos	Rápidos	No progresores			
Edad		60,152 ± 7,745	60,787 ± 6,995	61,925 ± 7,304	P.N.R. vs N.P.	0,003*	0,008*
					P.N.R. vs P.R.	0,462	1
					N.P. vs P.R.	0,117	0,351
IMC (kg/m <sup>2</sup> )		28,390 ± 5,006	28,642 ± 4,883	27,049 ± 4,341	P.N.R. vs N.P.	0,002*	0,007*
					P.N.R. vs P.R.	0,662	1
					N.P. vs P.R.	0,006*	0,019*
WOMAC		5,228 ± 3,626	5,934 ± 3,634	4,868 ± 4,151	P.N.R. vs N.P.	0,111	0,333
					P.N.R. vs P.R.	0,172	0,517
					N.P. vs P.R.	0,010*	0,030*
Cualitativas		No rápidos	Rápidos	No progresores	Total	P-valor	
Sexo	Hombre	47 (21%)	15 (13,9%)	81 (24,3%)	143	0,073	
	Mujer	177 (79%)	93 (86,1%)	253 (75,7%)	523		
Contralateralidad	No	178 (81,3%)	73 (69,5%)	302 (90,7%)	553	<0,001*	
	Sí	41 (18,7%)	32 (30,5%)	31 (9,3%)	104		

Las variables cuantitativas fueron analizadas mediante el test de Kruskal-Wallis, y los p-valores corregidos mediante Bonferroni; las variables cualitativas fueron analizadas mediante un análisis de  $\chi^2$ . Valores de las variables cuantitativas: valor promedio  $\pm$  desviación estándar; valores de las variables cualitativas: número de casos con la frecuencia de cada grupo entre paréntesis; P.R: progresores rápidos; P.N.R: progresores no rápidos; N.P: no progresores; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; \*: significación estadística <0,05.

Tabla 20. Variables clínicas y demográficas de los individuos seleccionados en base al criterio de progresión GIR-blanco de dos grupos de progresión en la cohorte europea de validación.

Criterio GIR-blanco		Dos grupos de progresión			P-valor	OR	IC 95%	
Variables cuantitativas		No rápidos	Rápidos				Inf	SUP
Edad		60,730 ± 7,790	60,820 ± 7,019		0,910	1,002	0,975	1,029
IMC (kg/m <sup>2</sup> )		27,526 ± 4,733	28,580 ± 4,876		0,054	1,045	0,999	1,092
WOMAC		4,896 ± 3,992	5,933 ± 3,659		0,035*	1,065	1,005	1,128
Variables cualitativas		No rápidos	Rápidos	Total	P-valor	OR	IC 95%	
							Inf	SUP
Sexo	Hombre	136 (22,7%)	14 (13,1%)	150	0,025*	1,951	1,078	3,532
	Mujer	463 (77,3%)	93 (86,9%)	556				
Contralateralidad	No	517 (87,2%)	72 (69,2%)	589	<0,001*	3,023	1,869	4,891
	Sí	76 (12,8%)	32 (30,8%)	108				

Se analizó el efecto de las variables cuantitativas mediante un análisis univariado para la obtención del correspondiente OR e IC95%, las variables cualitativas fueron analizadas mediante un análisis de  $\chi^2$ ; Valores de las variables cuantitativas: valor promedio  $\pm$  desviación estándar; valores de las variables cualitativas: número de casos con la frecuencia de cada grupo entre paréntesis; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; \*: significación estadística <0,05; OR: odds ratio; IC 95%: intervalo de confianza al 95% con los límites inferior y superior.

## Resultados del objetivo específico número 2: analizar el genoma mitocondrial mediante metodologías de NGS con el fin de identificar posibles variantes asociadas con el fenotipo de artrosis de rodilla con progresión rápida.

En este proyecto de tesis se realizó la secuenciación completa del ADNmt mediante metodologías de secuenciación masiva en plataformas *Ion Torrent* (Thermofisher). Como ya se explicó en el apartado de metodologías, es necesario hacer un importante análisis bioinformático para filtrar las variantes más robustas (Fig. 59). Una vez finalizada la secuenciación de las 450 muestras iniciales (184 progresores rápidos y 266 individuos no rápidos y no progresores) se encontraron un total de 1.335 variantes mitocondriales con respecto a la referencia empleada. El primer paso consistió en un filtrado donde sólo se mantuvieron en el estudio aquellas variantes con una frecuencia de heteroplasma mayor o igual a 20%, y una MAF mayor o igual al 5%. Aplicando estos parámetros de filtrado, de estas 1.335 variantes mitocondriales iniciales, solo 76 se seleccionaron para el estudio (Tabla 21).

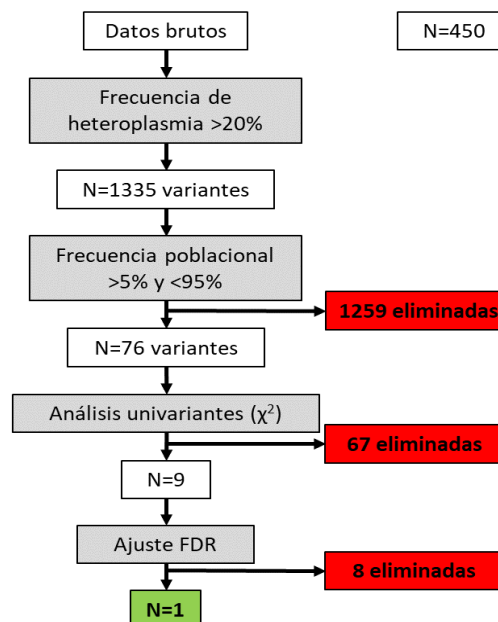


Figura 59. Proceso de filtrado de los datos brutos obtenidos de la NGS del ADNmt de las 450 muestras iniciales de la cohorte OAI. FDR: False Discovery Rate.

Tras este primer filtrado, el siguiente paso consistió en comparar la frecuencia de estas variantes entre los grupos de progresión del criterio GIR-blanco, para tratar de ver si alguna de ellas presentaba una frecuencia superior dentro del grupo de progresores rápidos. Aunque los resultados se van a centrar en el criterio GIR-blanco de dos grupos de progresión, de forma complementaria, también se realizaron estos análisis con el criterio GIR-blanco de tres grupos de progresión, aunque estos datos no se han mostrado en esta tesis doctoral. En estos análisis se utilizó un filtro adicional consistente en la aplicación de un p-valor nominal  $\leq 0,15$ . De las 76 variantes, solamente nueve se mantuvieron en el estudio (Tabla 22): mt.146T>C (p-valor nominal=0,15), mt.930G>A (p-valor nominal=0,11), mt.5147G>A (p-valor nominal=0,044), mt.10463T>C (p-valor nominal=0,106), mt.11812A>G (p-valor nominal=0,042), mt.14233A>G (p-valor nominal=0,042), mt.16294C>T (p-valor nominal=0,07), mt.16296C>T (p-valor

## Resultados

nominal=0,045) y mt.16519T>C (p-valor nominal=0,005). Debido a que se realizan múltiples comparaciones sobre la misma muestral de estudio, es necesario realizar un ajuste de este p-valor mediante una corrección de FDR. La única variante que mantuvo la significación estadística, tras la corrección, en el criterio de dos grupos de progresión fue la correspondiente a la posición mt.16519C (p-valor ajustado=0,045). En el criterio GIR-blanco de tres categorías de progresión, la variante mostraba un p-valor nominal significativo (p-valor nominal = 0,01), pero tras el ajuste del FDR, este rozaba la significación estadística (p-valor ajustado = 0,065), probablemente debido a la pérdida de poder estadístico al subdividir los 450 participantes en tres grupos de progresión.

Tabla 21. Lista de las 76 variantes mitocondriales que superaron los filtros iniciales de MAF y frecuencia de heteroplasmia en el criterio GIR-blanco de dos grupos de progresión. MAF: frecuencia mínima alélica, & el alelo menos común es el alelo de referencia.

Variante	MAF	Referencia	Cambio	No rápidos (N=266)	Rápidos (N=184)	Región ADNmt
m.73	45,625&	A	G	139 (52,3)	104 (56,5)	MT-HV2
m.146	9,583	T	C	21 (7,9)	22 (12,0)	MT-HV2
m.150	7,708	C	T	17 (6,4)	18 (9,8)	MT-HV2
m.152	22,292	T	C	58 (21,8)	39 (21,2)	MT-HV2
m.185	6,875	G	A	18 (6,8)	11 (6,0)	MT-HV2
m.195	17,083	T	C	52 (19,5)	27 (14,7)	MT-HV2
m.204	6,875	T	C	19 (7,1)	11 (6,0)	MT-HV2
m.207	5,000	G	A	16 (6,0)	7 (3,8)	MT-HV2
m.228	5,625	G	A	15 (5,6)	10 (5,4)	MT-HV2
m.295	6,458	C	T	20 (7,5)	8 (4,3)	MT-HV2
m.462	6,458	C	T	17 (6,4)	10 (5,4)	MT-HV3
m.489	7,083	T	C	27 (10,2)	13 (7,1)	MT-HV3
m.497	5,833	C	T	18 (6,8)	9 (4,9)	MT-HV3
m.709	16,875	G	A	43 (16,2)	36 (19,6)	MT-RNR1
m.930	9,583	G	A	21 (7,9)	23 (12,5)	MT-RNR1
m.1189	6,875	T	C	20 (7,5)	12 (6,5)	MT-RNR1
m.1719	6,667	G	A	17 (6,4)	12 (6,5)	MT-RNR2
m.1811	15,000	A	G	38 (14,3)	30 (16,3)	MT-RNR2
m.1888	9,375	G	A	22 (8,3)	21 (11,4)	MT-RNR2
m.2706	43,333&	A	G	152 (57,1)	100 (54,3)	MT-RNR2
m.3010	22,917	G	A	58 (21,8)	46 (25,0)	MT-RNR2
m.3197	8,333	T	C	21 (7,9)	15 (8,2)	MT-RNR2
m.3480	9,583	A	G	26 (9,8)	18 (9,8)	MT-ND1
m.4216	18,333	T	C	48 (18,0)	34 (18,5)	MT-ND1
m.4917	9,583	A	G	23 (8,6)	21 (11,4)	MT-ND2
m.5147	6,667	G	A	13 (4,9)	18 (9,8)	MT-ND2
m.6776	5,417	T	C	12 (4,5)	12 (6,5)	MT-CO1
m.7028	43,125&	C	T	152 (57,1)	101 (54,9)	MT-CO1
m.8251	6,250	G	A	19 (7,1)	8 (4,3)	MT-CO2
m.8697	10,208	G	A	24 (9,0)	22 (12,0)	MT-ATP6
m.9055	10,208	G	A	27 (10,2)	19 (10,3)	MT-ATP6
m.9477	8,125	G	A	21 (7,9)	14 (7,6)	MT-CO3
m.9698	10,000	T	C	27 (10,2)	19 (10,3)	MT-CO3
m.10398	18,125	A	G	47 (17,7)	32 (17,4)	MT-ND3
m.10463	9,583	T	C	21 (7,9)	23 (12,5)	MT-TR (tRNA-Arg)
m.10550	9,583	A	G	26 (9,8)	18 (9,8)	MT-ND4L
m.11251	17,292	A	G	44 (16,5)	33 (17,9)	MT-ND4
m.11299	9,583	T	C	26 (9,8)	18 (9,8)	MT-ND4
m.11467	23,542	A	G	62 (23,3)	43 (23,4)	MT-ND4
m.11719	48,750&	G	A	135 (50,8)	94 (51,1)	MT-ND4
m.11812	7,500	A	G	15 (5,6)	20 (10,9)	MT-ND4
m.11914	6,667	G	A	15 (5,6)	16 (8,7)	MT-ND4



m.12308	23,958	A	G	62 (23,3)	45 (24,5)	MT-TL2 (tRNA-Leu)
m.12372	24,167	G	A	63 (23,7)	45 (24,5)	MT-ND5
m.12612	8,125	A	G	23 (8,6)	12 (6,5)	MT-ND5
m.12705	9,792	C	T	28 (10,5)	16 (8,7)	MT-ND5
m.13368	9,792	G	A	23 (8,6)	22 (12,0)	MT-ND5
m.13617	8,125	T	C	21 (7,9)	14 (7,6)	MT-ND5
m.13708	11,250	G	A	33 (12,4)	17 (9,2)	MT-ND5
m.14167	9,583	C	T	26 (9,8)	18 (9,8)	MT-ND6
m.14233	7,500	A	G	15 (5,6)	20 (10,9)	MT-ND6
m.14766	47,917	C	T	125 (47,0)	89 (48,4)	MT-CYB
m.14798	16,667	T	C	45 (16,9)	29 (15,8)	MT-CYB
m.14905	9,167	G	A	21 (7,9)	21 (11,4)	MT-CYB
m.15452	17,292	C	A	44 (16,5)	33 (17,9)	MT-CYB
m.15607	9,167	A	G	21 (7,9)	21 (11,4)	MT-CYB
m.15924	5,833	A	G	14 (5,3)	11 (6,0)	MT-TT (tRNA-Thr)
m.15928	9,375	G	A	22 (8,3)	21 (11,4)	MT-TT (tRNA-Thr)
m.16069	8,125	C	T	23 (8,6)	12 (6,5)	MT-HV1
m.16093	7,083	T	C	18 (6,8)	14 (7,6)	MT-HV1
m.16126	18,750	T	C	50 (18,8)	34 (18,5)	MT-HV1
m.16129	5,833	G	A	13 (4,9)	12 (6,5)	MT-HV1
m.16172	5,000	T	C	14 (5,3)	9 (4,9)	MT-HV1
m.16183	5,625	A	C	12 (4,5)	14 (7,6)	MT-HV1
m.16189	9,583	T	C	25 (9,4)	19 (10,3)	MT-HV1
m.16192	5,625	C	T	14 (5,3)	9 (4,9)	MT-HV1
m.16223	10,000	C	T	27 (10,2)	18 (9,8)	MT-HV1
m.16224	9,583	T	C	26 (9,8)	18 (9,8)	MT-HV1
m.16256	5,208	C	T	16 (6,0)	6 (3,3)	MT-HV1
m.16270	7,500	C	T	18 (6,8)	14 (7,6)	MT-HV1
m.16294	10,208	C	T	22 (8,3)	25 (13,6)	MT-HV1
m.16296	5,000	C	T	9 (3,4)	14 (7,6)	MT-HV1
m.16304	9,792	T	C	24 (9,0)	20 (10,9)	MT-HV1
m.16311	19,167	T	C	51 (19,2)	33 (17,9)	MT-HV1
m.16362	7,917	T	C	25 (9,4)	12 (6,5)	MT-HV1
m.16519	31,458 <sup>&amp;</sup>	T	C	169 (63,5)	140 (76,1)	Control region

Los valores de las columnas de progresores rápidos y no rápidos son número de pacientes con su respectivo porcentaje entre paréntesis; MAF: frecuencia mínima alélica, & el alelo menos común es el alelo de referencia.

Tabla 22. Lista de variantes con sobrerrepresentación dentro del grupo de progresores rápidos en las 450 muestras iniciales usadas en la NGS en el criterio GIR-blanco de dos grupos de progresión.

Variante	Referencia	Cambio	No rápidos (N=266)	Rápidos (N=184)	OR	IC 95%	p-valor		mtDNA region
							nominal	ajustado	
m.146	T	C	21 (7,9)	22 (12,0)	1,584	0,844 - 2,975	0,150	0,150	MT-HV2
m.930	G	A	21 (7,9)	23 (12,5)	1,667	0,893 - 3,111	0,106	0,120	MT-RNR1
m.5147	G	A	13 (4,9)	18 (9,8)	2,11	1,007 - 4,422	0,044*	0,080	MT-ND2
m.10463	T	C	21 (7,9)	23 (12,5)	1,667	0,893 - 3,111	0,106	0,120	MT-TR (tRNA-Arg)
m.11812	A	G	15 (5,6)	20 (10,9)	2,041	1,016 - 4,100	0,042*	0,080	MT-ND4
m.14233	A	G	15 (5,6)	20 (10,9)	2,041	1,016 - 4,100	0,042*	0,080	MT-ND6
m.16294	C	T	22 (8,3)	25 (13,6)	1,744	0,951 - 3,199	0,070	0,100	MT-HV1
m.16296	C	T	9 (3,4)	14 (7,6)	2,352	0,996 - 5,555	0,045*	0,080	MT-HV1
m.16519	T	C	169 (63,5)	140 (76,1)	1,826	1,199 - 2,782	0,005*	0,045**	Control region

Se realizó la comparación en la distribución de las variantes entre los grupos mediante un análisis de  $\chi^2$ . Los valores de las columnas de progresores rápidos y no rápidos son número de pacientes con su respectivo porcentaje entre paréntesis; \*: significación estadística nominal <0,05; \*\*: significación estadística ajustada <0,05; IC95%: intervalo de confianza al 95; OR: Odds ratio

En este estudio no se incluyeron variantes mitocondriales con una MAF inferior al 5% ni las posiciones multiplásmicas m.310, m.514 y m.16189, aunque sí fueron identificadas en las

## Resultados

muestras secuenciadas. Tras realizar esta secuenciación en profundidad, la posición m.16519 se analizó en el resto de muestras de la cohorte OAI (N=1.095) mediante técnicas de secuenciación capilar. Tras asignar los genotipos de todas las muestras se procedió a comparar la frecuencia de los alelos entre los grupos de progresión. Como se observa en la tabla 23, la variante m.16519C mostró una frecuencia significativamente mayor dentro del grupo de progresores rápidos en el criterio de tres grupos de progresión (p-valor=0,016). Con el criterio GIR-blanco de dos grupos de progresión pudimos observar que, en los análisis univariantes, el alelo C de esta variante se asocia de forma significativa con el riesgo de progresar rápidamente (p-valor=0,004; OR=1,558; IC95% 1,148-2,113).

Con el objetivo de validar esta asociación, se realizó la misma aproximación de secuenciación capilar utilizada en la cohorte OAI en las cohortes externas CHECK y PROCOAC (N=699). Tras asignar los genotipos de todas las muestras se procedió a comparar la frecuencia de los alelos entre los grupos de progresión. Como se observa en la tabla 24, de nuevo, la variante m.16519C mostró una frecuencia significativamente mayor dentro del grupo de progresores rápidos en el criterio de tres grupos de progresión (p-valor=0,013). Con el criterio GIR-blanco de dos grupos de progresión pudimos observar que, en los análisis univariantes, el alelo C de esta variante se asocia de forma significativa con el riesgo de progresar rápidamente (p-valor=0,043; OR=1,647; IC95% 1,013-2,678).

Tabla 23. Análisis univariante de la variante m.16519C en los integrantes de la cohorte OAI de los criterios de progresión GIR-blanco.

Criterio GIR-blanco con tres grupos de progresión										
m.16519	No rápidos		Rápidos		No progresores		Total		P-valor	
	N	Frec.	N	Frec.	N	Frec.	N	Frec.		
T	128	0,360	72	0,269	173	0,367	373	0,341	0,016*	
C	228	0,640	196	0,731	298	0,633	722	0,659		
Total	356	1	268	1	471	1	1095	1		
Criterio GIR-blanco con dos grupos de progresión										
m.16519	No rápidos		Rápidos		Total		P-valor	OR	IC 95%	
	N	Frec.	N	Frec.	N	Frec.			Inf	SUP
T	301	0,364	72	0,269	373	0,341	0,004*	1,558	1,148	2,113
C	526	0,636	196	0,731	722	0,659				
Total	827	1	268	1	1095	1				

Se realizó la comparación en la distribución de la variante entre los grupos mediante un análisis univariado de  $\chi^2$ . \*: significación estadística nominal <0,05; Frec: frecuencia alélica; IC95%: intervalo de confianza al 95%; OR: Odds ratio.

Finalmente, se analizó el efecto de esta variante en la progresión rápida mediante una regresión logística binaria (criterio GIR-blanco de dos grupos de progresión), ajustando por las variables clínicas de confusión. Debido a que en las cohortes externas no está presente la variable lesión previa, ésta fue excluida de estos análisis (Tabla 25).

En la fase de exploración con el criterio GIR-blanco de dos categorías, la variante m.16519C se presentó como un factor de riesgo para el fenotipo de artrosis de rodilla con progresión rápida (p-valor=0,008; OR=1,559; IC95% 1,125-2,161) independientemente de la edad (p-valor<0,001; OR=1,054; IC95% 1,036-1,072), el IMC (p-valor<0,001; OR=1,073; IC95% 1,037-1,110), la artrosis contralateral (p-valor<0,001; OR=2,475; IC95% 1,830-3,348) y el dolor WOMAC (p-valor<0,001; OR=1,122; IC95% 1,065-1,183). Por otro lado, tras el ajuste en la fase de validación, la variante m.16519C bordeó, pero no alcanzó la significación estadística (p-valor=0,162). Sin embargo, se pudo observar que la tendencia de la variante era casi idéntica a la detectada en la fase de exploración (OR=1,506; IC95% 0,848-2,675) (Tabla 25).

Tabla 24. Análisis univariante de la variante m.16519C en los integrantes de las cohortes externas de los criterios de progresión GIR-blanco.

Criterio GIR-blanco con tres grupos de progresión									
m.16519	No rápidos		Rápidos		No progresores		Total		P-valor
	N	Frec.	N	Frec.	N	Frec.	N	Frec.	
T	87	0,373	24	0,224	104	0,290	215	0,308	0,013*
C	146	0,627	83	0,776	255	0,710	484	0,692	
Total	223	1	107	1	359	1	699	1	

Criterio GIR-blanco con dos grupos de progresión										
m.16519	No rápidos		Rápidos		Total		P-valor	OR	IC 95%	
	N	Frec.	N	Frec.	N	Frec.			inf	SUP
T	191	0,323	24	0,224	215	0,308	0,043*	1,647	1,013	2,678
C	401	0,677	83	0,776	484	0,692				
Total	827	1	107	1	699	1				

Se realizó la comparación en la distribución de la variante entre los grupos mediante un análisis de  $\chi^2$ . \*: significación estadística nominal  $<0,05$ ; Frec: frecuencia alélica; IC95%: intervalo de confianza al 95%; OR: Odds ratio.

Tabla 25. Análisis del efecto de la variante m.16519C en la progresión rápida en la fase de exploración y validación.

Variable	Exploración (OAI)				Validación (CHECK/PROCOAC)			
	P-valor	OR	IC 95%		P-valor	OR	IC 95%	
			inf	SUP			inf	SUP
Edad	<0,001*	1,054	1,036	1,072	0,138	0,972	0,937	1,009
IMC	<0,001*	1,073	1,037	1,110	0,727	1,010	0,957	1,066
Sexo (Femenino)	0,142	1,258	0,926	1,710	0,190	1,582	0,796	3,144
Contralateralidad (Sf)	<0,001*	2,475	1,830	3,348	0,048*	2,056	1,005	4,208
WOMAC	<0,001*	1,122	1,065	1,183	0,060	1,062	0,997	1,130
m.16519C	0,008*	1,559	1,125	2,161	0,162	1,506	0,848	2,675

Se analizó el efecto de la variante m.16519C en la progresión rápida mediante una regresión logística binaria, usando como variable dependiente el tipo de progresión, y como variables predictoras la edad, el IMC, el sexo, la contralateralidad, el WOMAC y la variante m.16519C. \*: significación estadística nominal  $<0,05$ ; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; IC95%: intervalo de confianza al 95%; OR: Odds ratio.

Para llevar a cabo el abordaje del meta-análisis, sólo se incluyeron los resultados obtenidos a partir de las cohortes de este trabajo; no se consideró ningún otro resultado porque no existe en la literatura científica ningún estudio de asociación similar que incluya a este fenotipo. Con todo, el tamaño muestral final incluyó a un total de 1.794 sujetos. Se realizaron dos meta-análisis, uno de ellos incluía los resultados individuales de la cohorte del OAI y de la cohorte europea de validación (Fig. 60 A), y otro meta-análisis que concentró los resultados del análisis univariado de las tres cohortes por separado, OAI, CHECK y PROCOAC (Fig. 60 B).

En el meta-análisis correspondientes a los resultados de las dos cohortes no se detectó heterogeneidad entre estudios ( $I^2=0,0\%$ ,  $p\text{-valor}=0,9173$ ), y se observó que la cohorte OAI contribuyó mucho más al peso relativo de los modelos que la cohorte europea de validación (75,5% vs 24,5% tanto para el modelo de efectos fijos como para el modelo de efectos aleatorios). Tanto el modelo de efectos fijo como el modelo de efectos aleatorios ( $p\text{-valor}=0,0027$ ;  $OR=1,546$ ;  $IC95\% 1,163\text{-}2,054$ ) mostraron una clara asociación estadísticamente

## Resultados

significativa de la variante mitocondrial m.16519C con el riesgo de desarrollar el fenotipo de artrosis de rodilla con progresión rápida.

Por su parte, en el meta-análisis correspondientes a los resultados de las tres cohortes individuales tampoco se detectó heterogeneidad entre estudios ( $I^2=0,0\%$ ,  $p$ -valor= $0,8756$ ). De nuevo, la cohorte OAI contribuyó mucho más al peso relativo de los modelos que las otras dos cohortes de validación (71,6% vs 16% y 12,4% para el modelo de efectos fijos, y 71,8% vs 16,5% y 11,7% para el modelo de efectos aleatorios). Tanto el modelo de efectos fijo ( $p$ -valor= $0,0005$ ; OR= $1,584$ ; IC95%  $1,224-2,050$ ) como el modelo de efectos aleatorios ( $p$ -valor= $0,0005$ ; OR= $1,583$ ; IC95%  $1,222-2,049$ ) mostraron una clara asociación estadísticamente significativa de la variante mitocondrial m.16519C con el riesgo de desarrollar el fenotipo de artrosis de rodilla con progresión rápida.

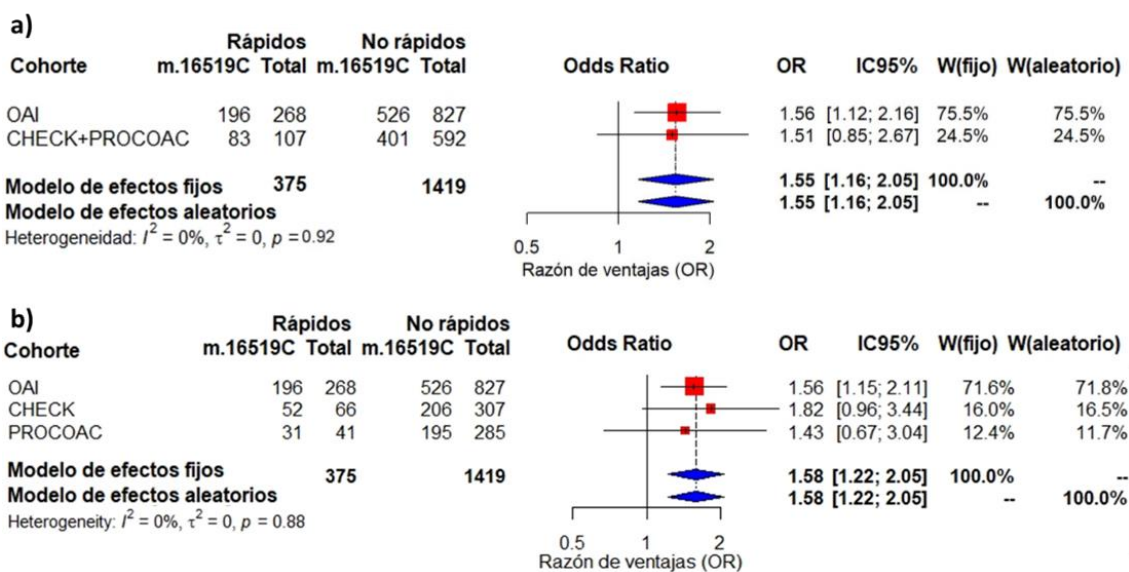


Figura 60. A: Forest plot resultante del meta-análisis de la variante m.16519C. Los datos de la cohorte de exploración y replicación fueron meta-analizados conjuntamente. Los datos de las columnas de rápidos y no rápidos muestran el número de individuos portadores del alelo m.16519C y del número de individuos totales para cada grupo. OAI: Osteoarthritis Initiative; CHECK: Cohort hip and cohort Knee; PROCOAC: Prospective cohort of A Coruña; IC95%: intervalo de confianza al 95%; OR: Odds ratio;  $I^2$ : índice de heterogeneidad entre estudios;  $p$ :  $p$ -valor. B: Forest plot resultante del meta-análisis de la variante m.16519C. Los datos de la cohorte de exploración y las dos de replicación fueron meta-analizados conjuntamente. Los datos de las columnas de rápidos y no rápidos muestran el número de individuos portadores del alelo m.16519C y del número de individuos totales para cada grupo. OAI: Osteoarthritis Initiative; CHECK: Cohort hip and cohort Knee; PROCOAC: Prospective cohort of A Coruña; IC95%: intervalo de confianza al 95%; OR: Odds ratio;  $I^2$ : índice de heterogeneidad entre estudios;  $p$ :  $p$ -valor

## Resultados del objetivo específico número 3: exploración del significado funcional de las variantes mitocondriales asociadas con el fenotipo de progresión rápida mediante análisis del transcriptoma por RNA-seq.

Tras detectar una asociación significativa entre la variante m.16519C y la progresión rápida de artrosis de rodilla en nuestras cohortes de estudio, tratamos de estudiar el significado funcional de esta variante, en términos de expresión génica, usando cíbridos transmitocondriales. Tras generar las dos líneas de cíbridos fusionando células p0 de la línea de osteosarcoma 143B.TK, con plaquetas de donantes sanos portadores del haplogrupo H y la variante m.16519C o la variante m.16519T, se llevó a cabo la secuenciación total de la molécula de ADNmt de estos cíbridos para analizar el *background* genético que poseían.

Como resultado, se detectaron un total de 17 variantes mitocondriales diferenciales entre las dos líneas de cíbridos (Tabla 26). Los cíbridos portadores del alelo salvaje m.16519T presentan un ADNmt representativo del haplogrupo H6b2, mientras que los portadores de la variante m.16519C pertenecen al haplogrupo H76.

Tabla 26. Listado de variantes mitocondriales diferenciales entre las dos líneas de cíbridos transmitocondriales creadas.

Cíbrido 509.H6b2				Cíbrido 513.H76			
posición	Referencia	Variante#	mtDNA región ADNmt	posición	Referencia	Variante#	Región ADNmt
m.150	C	T	MT-OHR	m.152	T	C	MT-OHR
m.239	T	C	MT-OHR	m.4117	T	C	MT-ND1
m.4225	A	G	MT-ND1	m.6040	A	G	MT-CO1
m.6467	C	T	MT-CO1	m.8572	G	A	MT-ATP6/8
m.14040	G	A	MT-ND5	m.8943	C	T	MT-ATP6
m.16300	A	G	MT-HV1	m.12017	A	T	MT-ND4
m.16325	A	C	MT-HV1	m.16291	C	T	MT-HV1
m.16362	T	C	MT-HV1	m.16519	T	C	Región control
m.16482	A	G	Región control				

#: con respecto a la secuencia de referencia rCRS: ADNmt: ADN mitocondrial

Tras verificar el ADNmt de ambos cíbridos, se llevó a cabo la RNA-seq del transcriptoma de cinco clones de cada cíbrido para observar los cambios que producía la presencia de la variante m.16519C en el perfil de expresión celular. Los datos de RNA-seq de cada cíbrido fueron inicialmente analizados en un PCA sin supervisión (Fig. 61). En dicho análisis se establecieron un total de tres componentes principales que resumen el 83,9% de la variabilidad de las muestras. El primer componente fue capaz de clasificar en dos grupos a las diez muestras, agrupándolas correctamente en base a la base nucleotídica presente en la posición m.16519.

El siguiente paso consistió en el análisis de los DEG que mostraran un *Fc* significativo de  $\pm 2$ . En total se detectaron 460 DEGs, de los cuales 167 estaban sobreexpresados y 293 reprimidos en los cíbridos con el alelo de riesgo m.16519C (Tabla 27 y tabla suplementaria 1), pudiendo destacar por su rol en el proceso inflamatorio (importantísimo en la patología de la artrosis) la IL6, IGFBP1 (del inglés *insulin like growth factor binding protein 1*), COX2 (PTGS2) y la interleuquina 24 (IL24). Además, se analizaron qué procesos biológicos podían estar afectados por la presencia de estos DEGs. Los 10 principales procesos significativamente afectados son los presentes en la figura 62. Sin embargo, sólo tres de estos procesos aparecieron alterados con

## Resultados

un nivel de significación corregido por FDR. De estos tres, los dos primeros están relacionados con la inflamación, siendo uno la regulación positiva de la respuesta inflamatoria aguda (GO:0002675) y el otro la regulación positiva de la migración de leucocitos (GO:0002687); mientras que el tercer proceso consistió en la regulación negativa de procesos celulares (GO:0048523).

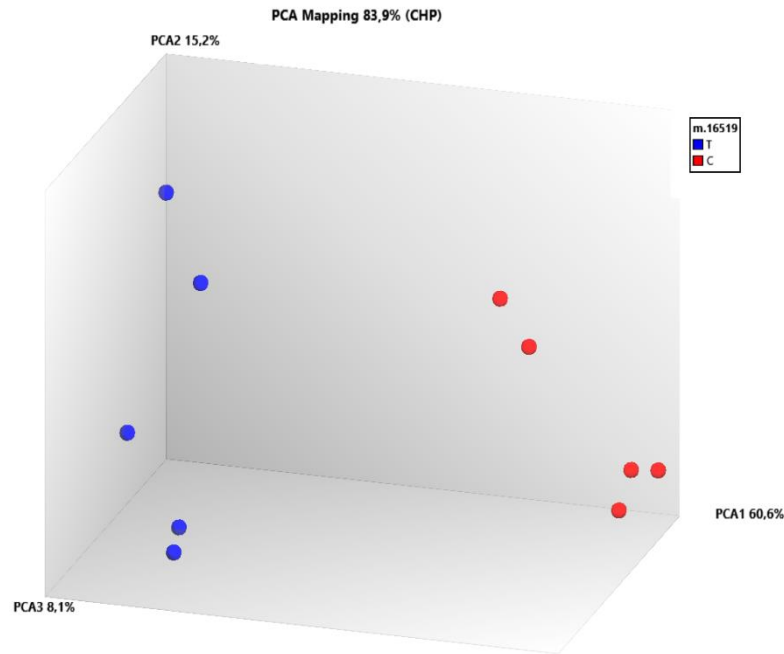


Figura 61. Análisis de componentes principales de los datos de RNA-seq de los híbridos transmitocondriales estudiados.

**GO Biological Process 2021** Bar Graph **Table** Clustergram Appyter

Hover each row to see the overlapping genes.

10 entries per page Search:

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	positive regulation of acute inflammatory response (GO:0002675)	0.000005957	0.004906 *	44.22	532.02
2	positive regulation of leukocyte migration (GO:0002687)	0.000003860	0.004906 *	25.47	317.54
3	negative regulation of cellular process (GO:0048523)	0.00008355	0.04587 *	3.45	32.43
4	cytokine-mediated signaling pathway (GO:0019221)	0.0002298	0.07615	3.13	26.23
5	regulation of cell population proliferation (GO:0042127)	0.0002312	0.07615	2.90	24.24
6	positive regulation of epithelial to mesenchymal transition (GO:0010718)	0.0004098	0.09146	12.78	99.71
7	positive regulation of cell migration (GO:0030335)	0.0004442	0.09146	4.29	33.10
8	negative regulation of cell population proliferation (GO:0008285)	0.0003540	0.09146	3.73	29.64
9	positive regulation of glial cell proliferation (GO:0060252)	0.0006817	0.09356	80.12	584.16
10	regulation of acute inflammatory response (GO:0002673)	0.0005872	0.09356	21.32	158.65

Figura 62. Principales 10 procesos biológicos afectados por los DEGs sobreexpresados en los híbridos m.16519C. \*: significación estadística ajustada < 0,05.

Tabla 27. Listado de los diez DEGs más sobreexpresados (diez primeros) o reprimidos (diez últimos) en los híbridos m.16519C.

GEN	Fc	p-val	q-val
IGFBP1	23,37	2,26E-05	0,0087
IL6	<b>14,36</b>	<b>1,61E-05</b>	<b>0,0075</b>
TINAGL1	14,05	4,53E-07	0,0013
IL24	11,35	9,27E-05	0,0141
SIRPB1	9,57	0,0002	0,0176
PTGS2	<b>9,16</b>	<b>0,0004</b>	<b>0,0249</b>
ERF1	7,67	0,0019	0,0475
SERPINB4	7,27	8,51E-06	0,006
SERPINB3	7,07	1,34E-05	0,0071
CCL20	6,48	0,0002	0,0186
HAS2	-5,78	0,0013	0,0405
MANSC1	-5,84	1,21E-05	0,007
CHFR	-5,89	2,48E-08	0,0003
MAPRE2	-6,09	2,60E-06	0,0034
H2AFX	-7,11	0,0006	0,0284
ZNF256	-7,63	3,26E-06	0,0034
MMP7	-12,66	0,0021	0,0499
FABP4	-14,36	5,41E-05	0,011
TM4SF18	-15,36	0,0006	0,0289
GAP43	-30,54	0,0002	0,018

Fc: fold change; p-val: p-valor nominal; q-val: p-valor ajustado por FDR

Como validación de los datos de RNA-seq, se decidió comprobar los niveles de expresión de IL6 y COX2 entre los diez clones de híbridos mediante un ensayo de qPCR (Fig. 63 A y B respectivamente). Los cinco clones de la línea m.16519C mostraron una expresión de IL6 (Fc=14,06; p-valor=0,0285) y COX2 (Fc=3,029; p-valor=0,02857) significativamente mayor que los cinco clones de la línea m.16519T.

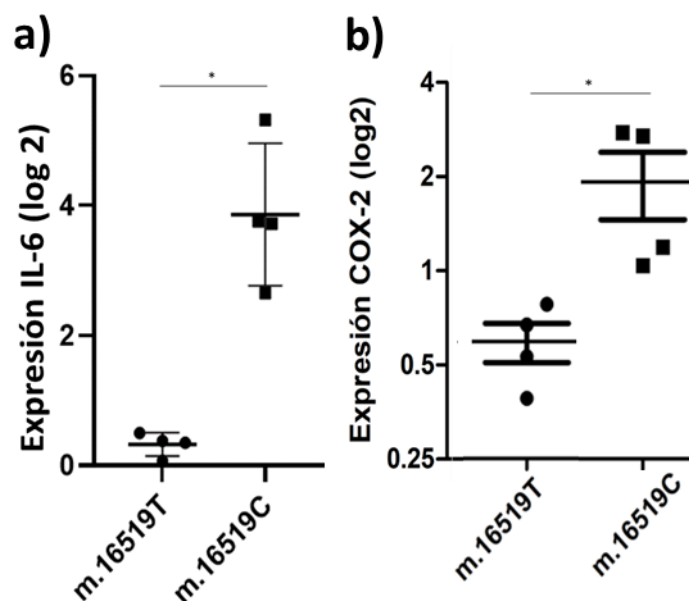


Figura 63. Análisis mediante qPCR de los niveles de IL6 (A) y COX2 (B) entre los clones de las dos líneas de híbridos m.16519. \*: significación estadística <math>< 0,05</math>

## Resultados

Finalmente, debido a que algunas muestras basales de la cohorte OAI cuentan con niveles de IL6 detectados en suero, se analizaron los niveles de esta citoquina en un total de 25 progresores rápidos y 204 progresores no rápidos mediante un análisis univariante U de Mann-Whitney. Dicho análisis reveló que los progresores rápidos presentaban unos niveles basales de IL6 en suero significativamente mayores que los progresores no rápidos ( $p$ -valor=0,044). Al realizar el modelo de regresión ajustando por las variables clínicas de confusión, de nuevo los niveles basales de suero de la IL6 aparecían significativamente elevados en el grupo de progresores rápidos (Tabla 28). Estos resultados parecen indicar que la variante m.16519C podría favorecer el desarrollo del fenotipo de progresión rápida aumentando los niveles de moléculas inflamatorias como la IL6.

Tabla 28. Análisis multivariante de los niveles basales de IL6 séricos de los 229 participantes analizados de la cohorte OAI.

Variable	P-valor	OR	IC 95%	
			inf	SUP
Edad	0,017*	1,063	1,011	1,118
IMC	0,027*	1,121	1,013	1,240
Sexo (Femenino)	0,264	1,699	0,670	4,306
Contralateralidad (Sí)	0,031*	2,715	1,094	6,738
WOMAC	0,136	1,117	0,966	1,291
Mediana IL6	0,045*	2,627	1,021	6,760

Se analizó el efecto de los niveles de IL6 en la progresión rápida mediante una regresión logística binaria, usando como variable dependiente el tipo de progresión, y como variables predictoras la edad, el IMC, el sexo, la contralateralidad, el WOMAC y la mediana de IL6; \*: significación estadística nominal  $<0,05$ ; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; IC95%: intervalo de confianza al 95%; OR: Odds ratio.



## Resultados del objetivo específico número 4: analizar el impacto de las interacciones entre el genoma nuclear y mitocondrial en el desarrollo del fenotipo de artrosis de rodilla con progresión rápida, y desarrollo de un modelo predictivo del fenotipo con aplicación clínica.

Esta tesis doctoral cuenta con otro aspecto pionero en el campo de la artrosis, la búsqueda de interacciones entre el genoma nuclear y el mitocondrial que puedan condicionar el desarrollo del fenotipo de artrosis de rodilla con progresión rápida. Debido a la aproximación estadística que se llevó a cabo, el escaso tamaño muestral de las cohortes externas hizo imposible la validación por esta vía, por lo que todo este objetivo específico se llevó a cabo en la cohorte OAI, utilizándose en este caso un método de validación interna cruzada por *bootstrap*.

Inicialmente, tras el análisis de los SNPs nucleares y los haplogrupos mitocondriales mediante técnicas de minisequenciación de SBE, se llevó a cabo un análisis de distribución de frecuencias de estos elementos entre los grupos de progresión. Por un lado, en las tablas 29 y 30 están presentes los análisis univariantes de  $\chi^2$  de los SNPs nucleares en los criterios GIR-blanco de tres y dos grupos de progresión respectivamente. Por otro lado, en las tablas 31 y 32 están presentes los análisis univariantes de  $\chi^2$  de los clusters de haplogrupos mitocondriales en los criterios GIR-blanco de tres y dos grupos de progresión respectivamente.

Tabla 29. Análisis univariantes para las frecuencias alélicas de los ocho SNPs nucleares en los criterios GIR-blanco de tres grupos de progresión.

GIR-blanco tres grupos de progresión								
SNPs nucleares	Alelo de riesgo	No rápidos		Rápidos		No progresores		P-valor
		N	Frec.	N	Frec.	N	Frec.	
rs12107036 (A/G)	G	374	0,525	303	0,567	485	0,515	0,141
rs3771501 (C/T)	T	346	0,486	250	0,473	451	0,479	0,906
rs143383 (C/T)	T	432	0,607	344	0,652	589	0,625	0,273
rs4730250 (A/G)	G	135	0,190	90	0,170	157	0,167	0,454
rs8044769 (C/T)	C	385	0,541	282	0,534	473	0,502	0,247
rs11177 (C/T)	T	283	0,397	200	0,379	388	0,412	0,459
rs10948172 (A/G)	G	211	0,296	160	0,303	268	0,285	0,732
rs11842874 (A/G)	A	652	0,916	494	0,936	876	0,930	0,364

Se realizó la comparación en la distribución de las variantes nucleares entre los grupos mediante un análisis de  $\chi^2$ . Los datos de las columnas representan el conteo de alelos de riesgo dentro de cada grupo de progresión; Frec.: frecuencia alélica dentro de cada grupo de progresión.

Como se observa en las tablas mencionadas, cuando se analizan los ocho SNPs nucleares o los clusters de haplogrupos mitocondriales por separado, no se observa ningún elemento que se asocie de forma significativa con el riesgo de desarrollar el fenotipo rápido de artrosis de rodilla. Si bien, el alelo G del SNP rs12107036 del gen TP63 bordeó la significación estadística (p-valor nominal = 0,053) con una tendencia hacia factor de riesgo de desarrollar este fenotipo (Tabla 30).

La última parte de los análisis univariantes es la que se centró en la búsqueda de posibles interacciones entre el genoma nuclear y el mitocondrial. Para ello se estratificó la población en base a su cluster mitocondrial y, para cada SNP nuclear, se analizó la distribución de frecuencias

## Resultados

de sus alelos entre los grupos de progresión. Las tablas 33 y 34 muestran los resultados de estos análisis en los criterios GIR-blanco de tres y dos grupos de progresión respectivamente.

Tabla 30. Análisis univariantes para las frecuencias alélicas de los ocho SNPs nucleares en los criterios GIR-blanco de dos grupos de progresión.

GIR-blanco dos grupos de progresión									
SNPs nucleares	Alelo de riesgo	No rápidos		Rápidos		P-valor	OR	IC 95%	
		N	Frec.	N	Frec.			inf	SUP
rs12107036 (A/G)	G	859	0,519	303	0,567	0,053	1,214	0,997	1,478
rs3771501 (C/T)	T	797	0,482	250	0,473	0,737	0,967	0,795	1,177
rs143383 (C/T)	T	1021	0,617	344	0,652	0,157	1,159	0,945	1,422
rs4730250 (A/G)	G	292	0,177	90	0,17	0,749	0,958	0,739	1,243
rs8044769 (C/T)	C	858	0,519	282	0,534	0,539	1,064	0,874	1,294
rs11177 (C/T)	T	671	0,406	200	0,379	0,272	0,893	0,73	1,093
rs10948172 (A/G)	G	479	0,29	160	0,303	0,555	1,067	0,861	1,321
rs11842874 (A/G)	A	1528	0,924	494	0,936	0,366	1,198	0,809	1,773

Se realizó la comparación en la distribución de las variantes nucleares entre los grupos mediante un análisis de  $\chi^2$ . Los datos de las columnas representan el conteo de alelos de riesgo dentro de cada grupo de progresión; Frec.: frecuencia alélica dentro de cada grupo de progresión; IC95%: intervalo de confianza inferior y superior al 95%; OR: Odds ratio.

Tabla 31. Análisis univariantes para las frecuencias de los clusters de haplogrupos mitocondriales en los criterios GIR-blanco de tres grupos de progresión.

GIR-blanco tres grupos de progresión							
Cluster ADNmt	No rápidos		Rápidos		No progresores		P-valor
	N	%		%	N	%	
HV	175	49,16%	132	49,44%	223	47,35%	0,816
UK	85	23,88%	63	23,60%	108	22,93%	0,947
TJ	61	17,13%	50	18,73%	89	18,90%	0,791
Otros	35	9,83%	22	8,24%	51	10,83%	0,526
Total	356	100,00%	267	100,00%	471	100,00%	0,934

Se realizó la comparación en la distribución de los clusters de haplogrupos mitocondriales entre los grupos mediante un análisis de  $\chi^2$ . Los datos de las columnas representan el conteo de individuos pertenecientes a dicho cluster dentro de cada grupo de progresión; %: porcentaje de individuos con dicho cluster de haplogrupos dentro del grupo de progresión correspondiente.

Tabla 32. Análisis univariantes para las frecuencias de los clusters de haplogrupos mitocondriales en los criterios GIR-blanco de dos grupos de progresión.

GIR-blanco dos grupos de progresión									
Cluster ADNmt	No rápidos		Rápidos		P-valor	OR	IC 95%		
	N	%	N	%			inf	SUP	
HV	398	48,13%	132	49,44%	0,709	0,949	0,72	1,25	
UK	193	23,34%	63	23,60%	0,931	0,986	0,712	1,365	
TJ	150	18,14%	50	18,73%	0,829	0,962	0,674	1,371	
Otros	86	10,40%	22	8,24%	0,304	1,292	0,792	2,11	
Total	827	100,00%	267	100,00%	0,785	-	-	-	

Se realizó la comparación en la distribución de los clusters de haplogrupos mitocondriales entre los grupos mediante un análisis de  $\chi^2$ . Los datos de las columnas representan el conteo de individuos pertenecientes a dicho cluster dentro

de cada grupo de progresión; %: porcentaje de individuos con dicho cluster de haplogrupos dentro del grupo de progresión correspondiente; IC95%: intervalo de confianza inferior y superior al 95%; OR: Odds ratio.

Tabla 33. Análisis univariantes para las frecuencias alélicas de los ocho SNPs nucleares en los criterios GIR-blanco de tres grupos de progresión estratificando la población en base a su cluster mitocondrial.

GIR-blanco tres grupos de progresión										
Cluster ADNmt	SNPs	Alelo de riesgo	No rápidos		Rápidos		No progresores		P-valor	Q-valor
			N	Frec.	N	Frec.	N	Frec.		
HV	rs12107036 (A/G)	G	188	0,537	150	0,568	226	0,507	0,277	1,000
	rs3771501 (C/T)	T	170	0,486	121	0,465	215	0,482	0,872	1,000
	rs143383 (C/T)	T	220	0,629	176	0,677	272	0,610	0,201	1,000
	rs4730250 (A/G)	G	67	0,191	42	0,162	85	0,191	0,568	1,000
	rs8044769 (C/T)	C	184	0,526	136	0,523	225	0,504	0,810	1,000
	rs11177 (C/T)	T	209	0,597	165	0,635	268	0,601	0,594	1,000
	rs10948172 (A/G)	G	101	0,289	73	0,281	136	0,305	0,769	1,000
	rs11842874 (A/G)	A	314	0,897	243	0,935	413	0,926	0,185	1,000
UK	rs12107036 (A/G)	G	85	0,500	79	0,627	100	0,463	0,012*	0,096
	rs3771501 (C/T)	T	79	0,465	64	0,508	103	0,477	0,755	1,000
	rs143383 (C/T)	T	94	0,553	79	0,627	132	0,611	0,365	1,000
	rs4730250 (A/G)	G	32	0,188	19	0,151	31	0,144	0,467	1,000
	rs8044769 (C/T)	C	92	0,541	62	0,492	106	0,491	0,567	1,000
	rs11177 (C/T)	T	110	0,647	80	0,635	112	0,519	0,019*	0,152
	rs10948172 (A/G)	G	50	0,294	40	0,317	57	0,264	0,555	1,000
	rs11842874 (A/G)	A	160	0,941	116	0,921	202	0,935	0,776	1,000
JT	rs12107036 (A/G)	G	60	0,492	48	0,490	101	0,567	0,315	1,000
	rs3771501 (C/T)	T	57	0,467	47	0,480	79	0,444	0,833	1,000
	rs143383 (C/T)	T	75	0,615	58	0,592	114	0,640	0,719	1,000
	rs4730250 (A/G)	G	23	0,189	21	0,214	25	0,140	0,261	1,000
	rs8044769 (C/T)	C	71	0,582	64	0,653	91	0,511	0,070	0,560
	rs11177 (C/T)	T	73	0,598	58	0,592	109	0,612	0,938	1,000
	rs10948172 (A/G)	G	39	0,320	36	0,367	50	0,281	0,330	1,000
	rs11842874 (A/G)	A	114	0,934	94	0,959	163	0,916	0,386	1,000
Otros	rs12107036 (A/G)	G	41	0,586	26	0,591	58	0,569	0,959	1,000
	rs3771501 (C/T)	T	40	0,571	18	0,409	54	0,529	0,229	1,000
	rs143383 (C/T)	T	43	0,614	31	0,705	71	0,696	0,464	1,000
	rs4730250 (A/G)	G	13	0,186	8	0,182	16	0,157	0,867	1,000
	rs8044769 (C/T)	C	38	0,543	20	0,455	51	0,500	0,651	1,000
	rs11177 (C/T)	T	37	0,529	25	0,568	65	0,637	0,348	1,000
	rs10948172 (A/G)	G	21	0,300	11	0,250	25	0,245	0,705	1,000
	rs11842874 (A/G)	A	64	0,914	41	0,932	98	0,961	0,438	1,000

Se realizó la comparación en la distribución de las variantes nucleares entre los grupos mediante un análisis de  $\chi^2$ . Los datos de las columnas representan el conteaje de alelos de riesgo dentro de cada grupo de progresión; \*: p-valor nominal significativo <0,05; Frec: frecuencia alélica dentro de cada grupo de progresión.

## Resultados

Tabla 34. Análisis univariantes para las frecuencias alélicas de los ocho SNPs nucleares en los criterios GIR-blanco de dos grupos de progresión estratificando la población en base a su cluster mitocondrial.

GIR-blanco dos grupos de progresión											
Cluster ADNmt	SNPs	Alelo de riesgo	No rápidos		Rápidos		P-valor	Q-valor	O.R	IC 95%	
			N	Frec.	N	Frec.				min	MAX
HV	rs12107036 (A/G)	G	414	0,520	150	0,568	0,175	1,000	1,214	0,917	1,607
	rs3771501 (C/T)	T	385	0,484	121	0,465	0,608	1,000	0,929	0,702	1,230
	rs143383 (C/T)	T	492	0,618	176	0,677	0,088	0,704	1,295	0,962	1,742
	rs4730250 (A/G)	G	152	0,191	42	0,162	0,288	1,000	0,816	0,561	1,187
	rs8044769 (C/T)	C	409	0,514	136	0,523	0,795	1,000	1,038	0,784	1,373
	rs11177 (C/T)	T	477	0,599	165	0,635	0,310	1,000	0,861	0,645	1,150
	rs10948172 (A/G)	G	237	0,298	73	0,281	0,602	1,000	0,921	0,675	1,256
	rs11842874 (A/G)	A	727	0,913	243	0,935	0,276	1,000	1,357	0,783	2,352
UK	rs12107036 (A/G)	G	185	0,479	79	0,627	0,004*	0,032**	1,826	1,209	2,759
	rs3771501 (C/T)	T	182	0,472	64	0,508	0,477	1,000	1,157	0,774	1,730
	rs143383 (C/T)	T	226	0,585	79	0,627	0,410	1,000	1,190	0,787	1,800
	rs4730250 (A/G)	G	63	0,163	19	0,151	0,741	1,000	0,910	0,521	1,590
	rs8044769 (C/T)	C	198	0,513	62	0,492	0,684	1,000	0,920	0,615	1,375
	rs11177 (C/T)	T	222	0,575	80	0,635	0,236	1,000	0,778	0,514	1,179
	rs10948172 (A/G)	G	107	0,277	40	0,317	0,386	1,000	1,213	0,784	1,876
	rs11842874 (A/G)	A	362	0,938	116	0,921	0,501	1,000	0,769	0,357	1,656
JT	rs12107036 (A/G)	G	161	0,537	48	0,490	0,420	1,000	0,829	0,525	1,308
	rs3771501 (C/T)	T	136	0,453	47	0,480	0,651	1,000	1,111	0,704	1,755
	rs143383 (C/T)	T	189	0,630	58	0,592	0,499	1,000	0,852	0,534	1,357
	rs4730250 (A/G)	G	48	0,160	21	0,214	0,218	1,000	1,432	0,807	2,539
	rs8044769 (C/T)	C	162	0,540	64	0,653	0,050*	0,400	1,603	0,998	2,576
	rs11177 (C/T)	T	182	0,607	58	0,592	0,794	1,000	1,064	0,668	1,693
	rs10948172 (A/G)	G	89	0,297	36	0,367	0,191	1,000	1,377	0,852	2,224
	rs11842874 (A/G)	A	277	0,923	94	0,959	0,220	1,000	1,951	0,658	5,788
Otros	rs12107036 (A/G)	G	99	0,576	26	0,591	0,854	1,000	1,065	0,544	2,087
	rs3771501 (C/T)	T	94	0,547	18	0,409	0,104	0,832	0,574	0,293	1,125
	rs143383 (C/T)	T	114	0,663	31	0,705	0,599	1,000	1,213	0,590	2,494
	rs4730250 (A/G)	G	29	0,169	8	0,182	0,836	1,000	1,096	0,462	2,599
	rs8044769 (C/T)	C	89	0,517	20	0,455	0,456	1,000	0,777	0,400	1,510
	rs11177 (C/T)	T	102	0,593	25	0,568	0,765	1,000	1,107	0,567	2,163
	rs10948172 (A/G)	G	46	0,267	11	0,250	0,815	1,000	0,913	0,426	1,955
	rs11842874 (A/G)	A	162	0,942	41	0,932	0,803	1,000	0,844	0,222	3,206

Se realizó la comparación en la distribución de las variantes nucleares entre los grupos mediante un análisis de  $\chi^2$ . Los datos de las columnas representan el conteo de alelos de riesgo dentro de cada grupo de progresión; \*: p-valor nominal significativo <0,05; \*\*: p-valor ajustado por Bonferroni significativo <0,05; Frec: frecuencia alélica dentro de cada grupo de progresión; IC95%: intervalo de confianza al 95%; OR: Odds ratio.

Como se puede observar en el primer análisis de interacciones (Tabla 33), dos interacciones mitonucleares muestran p-valores nominales significativos, ambas dentro del cluster mitocondrial UK. Sin embargo, tras ajustar el p-valor, ambas asociaciones pierden la significación estadística. Concretamente, en el caso del rs12107036 de TP63 y el cluster

mitocondrial UK, se puede observar como la frecuencia del alelo de riesgo G es muy superior dentro del grupo de progresores rápidos, alcanzando un 62% de frecuencia, mientras que los otros dos grupos de progresión, no-rápidos y no-progresores respectivamente, están en el 50% y 46,3%. En el análisis posterior, con el criterio de dos grupos de progresión (Tabla 34), donde se comparan los individuos rápidos frente a individuos con progresiones más lentas, se pueden observar dos interacciones que alcanzan un p-valor nominal significativo, otra vez la presente entre el SNP rs12107036 y el cluster UK y una nueva dentro del cluster mitocondrial JT. Sin embargo, tras el ajuste, únicamente la interacción que afecta a TP63 es la que alcanza un p-valor significativo, asociándose con un mayor riesgo de desarrollar el fenotipo de progresión rápida (p-valor nominal: 0,004; p-valor ajustado: 0,032; OR: 1,826; IC 95%: 1,209-2,759).

Por último, se emplearon todos los datos genéticos obtenidos durante esta tesis para tratar de desarrollar una herramienta predictiva, la cual fuera de fácil aplicación clínica, y capaz de identificar precozmente a aquellos individuos con un elevado riesgo de desarrollar el fenotipo de artrosis de rodilla con progresión rápida. Para ello, trabajamos con la cohorte OAI, donde se aplicó estadística predictiva en un modelo de regresión binario aditivo donde la variable dependiente era la progresión rápida y, como variables predictoras clínicas, se introdujeron la edad, el IMC, el sexo, la artrosis contralateral, la lesión previa y el dolor WOMAC; por otro lado, como variables predictoras genéticas se añadieron aquellos polimorfismos e interacciones que mostraron alguna asociación con el fenotipo de progresión rápida en otros objetivos específicos de este trabajo, véase el alelo G de rs12107036 de TP63, el cluster de haplogrupos mitocondriales UK, la interacción entre estos dos elementos y la variante mitocondrial m.16519C (Tabla 35). Además, se analizó el efecto de las variables entre los grupos de progresión realizando comparaciones dos a dos (Tabla 36).

Tabla 35. Modelo predictivo para el fenotipo de artrosis de rodilla con progresión rápida desarrollado en la cohorte OAI.

Variables clínicas				
Parámetro	p-valor	OR	IC 95%	
			Inf	SUP
Edad	<0,001*	1,058	1,040	1,077
IMC	<0,001*	1,070	1,034	1,106
Sexo femenino	0,107	1,291	0,946	1,761
Lesión previa (Sí)	<0,001*	1,915	1,400	2,620
Contralateralidad (Sí)	<0,001*	2,471	1,821	3,352
WOMAC	0,001*	1,103	1,042	1,169
Variables genéticas				
Parámetro	p-valor	OR	IC 95%	
			Inf	SUP
rs12107036.G	0,529	1,080	0,849	1,374
Cluster mitocondrial UK	0,096	0,690	0,446	1,069
m.16519C	0,008*	1,571	1,128	2,189
rs12107036.G + Cluster UK	0,037*	1,670	1,031	2,706

\*: significación estadística nominal <0,05; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; IC95%: intervalo de confianza al 95%; OR: Odds ratio.

En el análisis binario de individuos rápidos frente al resto de grupos de progresión (Tabla 35) se puede observar que todas las variables clínicas a excepción del sexo se asociaron de forma significativa con el riesgo de progresar rápidamente, véase edad (p-valor: <0,001; OR: 1,058; IC95% 1,040-1,077), IMC (p-valor: <0,001; OR: 1,070; IC 95% 1,034-1,106), lesión previa (p-valor: <0,001; OR: 1,915; IC 95%: 1,400-2,620), contralateralidad (p-valor: <0,001; OR: 2,471; IC 95%

## Resultados

1,821-3,352) y dolor WOMAC (p-valor: 0,001; OR: 1,103; IC95% 1,042-1,169). Con respecto a las variables genéticas, por separado, ni el alelo rs12107036G de TP63 ni el cluster mitocondrial UK se asociaron de forma significativa con el riesgo de progresar rápidamente. Sin embargo, cuando se analizaba su interacción, esta sí mostró un riesgo significativamente mayor de desarrollar este fenotipo (p-valor: 0,037; OR: 1,670; IC95% 1,031-2,706); por su parte, la variante mitocondrial m.16519C también mostró una asociación significativa con el riesgo de progresar rápidamente (p-valor: 0,008; OR: 1,571; IC95% 1,128-2,189).

También se realizó el análisis de regresión logística binaria comparando los tres grupos de progresión por parejas (Tabla 36). En la comparación entre progresores rápidos frente a no rápidos (sin incluir a los que no progresan) se obtuvieron unos datos muy similares al primer modelo de regresión (Tabla 35), de modo que todas las variables clínicas a excepción del sexo mostraron una asociación significativa con el riesgo de progresar rápidamente y, de nuevo, tanto la interacción entre el haplogrupo UK y el SNP rs12107036 (p-valor: 0,048; OR: 1,759; IC95% 1,006-3,077) como la variante m.16519C (p-valor: 0,017; OR: 1,574; IC95% 1,084-2,287) se asociaron significativamente con el riesgo de progresar rápidamente. Cuando se analizan progresores rápidos frente a no progresores (excluyendo a los progresores no rápidos) todas las variables clínicas se asocian de forma significativa con el riesgo de progresar rápidamente, y nuevamente, tanto la interacción entre el haplogrupo UK y el alelo G del SNP rs12107036 (p-valor: 0,048; OR: 1,686; IC95% 1,004-2,832) como la variante m.16519C (p-valor: 0,016; OR: 1,580; IC95% 1,091-2,289) se asociaron significativamente con el riesgo de progresar rápidamente. Por último, al comparar aquellos individuos que no progresan de forma rápida, véase progresores no rápidos y no progresores, solamente las variables clínicas, a excepción de la edad y la contralateralidad, se asociaron de forma significativa con el riesgo de progresar de forma no rápida.

Tabla 36. Comparativa de los efectos de las variables incluidas en el modelo predictivo para el fenotipo de artrosis de rodilla con progresión rápida entre los tres grupos de progresión del criterio GIR-blanco.

Comparativa	Rápidos v No rápidos (R)				Rápidos v No progresores (R)				No rápidos v No progresores (R)			
Parámetro	p-valor	OR	IC 95%		p-valor	OR	IC 95%		p-valor	OR	IC 95%	
			Inf	SUP			Inf	SUP			Inf	SUP
			<i>Edad</i>	<0,001*			1,068	1,046			1,090	<0,001*
<i>IMC</i>	0,046*	1,041	1,001	1,083	<0,001*	1,103	1,058	1,149	<0,001*	1,078	1,042	1,115
<i>Sexo femenino</i>	0,998	1,000	0,699	1,432	0,004*	1,717	1,194	2,470	0,001*	1,683	1,247	2,273
<i>Lesión previa (Sí)</i>	0,006*	1,640	1,151	2,336	<0,001*	2,170	1,516	3,107	0,010*	1,505	1,103	2,055
<i>Contralateralidad (Sí)</i>	<0,001*	2,183	1,545	3,085	<0,001*	2,752	1,953	3,879	0,187	1,222	0,907	1,647
<i>WOMAC</i>	0,067	1,061	0,996	1,131	<0,001*	1,162	1,088	1,242	0,008*	1,086	1,022	1,155
Variables genéticas												
Parámetro	p-valor	OR	IC 95%		p-valor	OR	IC 95%		p-valor	OR	IC 95%	
			Inf	SUP			Inf	SUP			min	Inf
			<i>rs12107036.G</i>	0,957			1,008	0,762			1,332	0,333
<i>Cluster mitocondrial UK</i>	0,066	0,626	0,379	1,032	0,166	0,713	0,442	1,151	0,727	1,076	0,714	1,622
<i>m.16519C</i>	0,017*	1,574	1,084	2,287	0,016*	1,580	1,091	2,289	0,601	1,083	0,804	1,458
<i>rs12107036.G + Cluster UK</i>	0,048*	1,759	1,006	3,077	0,048*	1,686	1,004	2,832	0,785	1,067	0,669	1,703

(R): grupo de progresión de referencia; \*: significación estadística nominal <0,05; IMC: índice de masa corporal; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; IC95%: intervalo de confianza al 95%; OR: Odds ratio.

Como análisis adicional característico de este tipo de abordajes, se realizó el cálculo de la proporción de riesgo atribuible a la interacción (PAF, del inglés *population attributable risk proportion*). En este caso se expresa sin ajustar por factores demográficos ni clínicos. En la tabla 37 se muestran los PAFs para cada variable de forma independiente. Se observa que la

proporción de casos de progresión rápida atribuibles a presentar el alelo G del SNP rs12107036 entre los sujetos con el cluster mitocondrial UK es del 23%, mucho más elevada que entre los sujetos pertenecientes a un cluster mitocondrial distinto de UK (3%), resultados que siguen la tendencia observada en los modelos propuestos hasta el momento.

Tabla 37. Cálculo de la proporción de riesgo atribuible a la interacción entre el alelo G del SNP rs12107036 de TP63 y el cluster mitocondrial UK.

Population Attributable Risk Proportion (PAF) sin ajustar por covariables	Estimación	IC95%	
		inf	SUP
Proporción de casos atribuibles a rs12107036G	8,05	0,157	15,9
Proporción de casos atribuibles a Cluster = UK	0,275	0	4,41
Proporción de casos atribuibles a rs12107036G en individuos Uk	23,4	7,89	38,9
Proporción de casos atribuibles a rs12107036G en individuos no Uk	3,1	0	12,3

IC95%: intervalo de confianza inferior y superior al 95%.

Finalmente, se validó este modelo predictivo mediante una validación cruzada interna (*crossvalidation*) dentro de la propia cohorte del OAI (Tabla 38). Inicialmente se valoró la precisión del modelo desarrollado a través del cálculo de la especificidad y sensibilidad midiendo el AUC, tomando como umbral un punto de corte estándar (0,5). A continuación, se corrigió esta AUC recalculando 500 veces este valor mediante técnicas de *bootstrap*. Finalmente, se generaron un total de diez subgrupos aleatorios dentro de la cohorte OAI (*k-fold crossvalidation* =10) para calcular la media de AUC de la validación (Tabla 38 y fig. 64).

Tabla 38. Áreas bajo la curva (AUC) del modelo desarrollado y de su validación interna cruzada.

Modelo	AUC	IC 95%	
		Inf	SUP
Desarrollo	0,735	0,711	0,759
Bootstrap corregido (B=500)	0,728	0,703	0,753
	<b>Media (AUC)</b>	<b>inf (AUC)</b>	<b>SUP(AUC)</b>
k-fold CV (k=10)	0,7296	0,6909	0,7696

AUC: area bajo la curva; IC95%: intervalo de confianza inferior y superior al 95%.

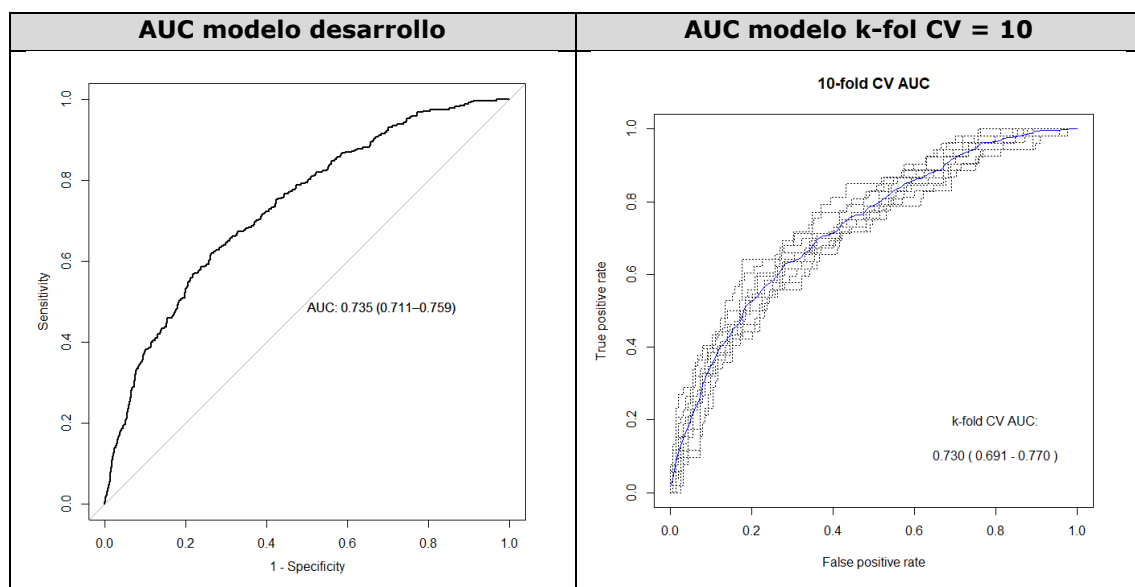


Figura 64. Comparativa de las áreas bajo la curva (AUC) entre el modelo desarrollado originalmente (izquierda) y de los modelos obtenidos de las diez validaciones cruzadas (derecha). El eje X representa la sensibilidad mientras que el eje Y la especificidad.

## Resultados

En el modelo desarrollado originalmente se obtuvo un AUC de 0,735 (IC95%: 0,711-0,759) y, en la validación cruzada se obtiene un AUC de 0,730 (IC95%: 0,691-0,770), lo que permite validar el modelo predictivo para el fenotipo de artrosis de rodilla con progresión rápida. Además, también se calculó la tasa de predicción del modelo validado, la cual alcanzó un valor de 77,1%, es decir con las variables presentes, el modelo es capaz de clasificar correctamente el 77,1% de los individuos.

Gracias a toda esta información, se desarrolló una herramienta visual para poder aplicar directamente en consulta el modelo predictivo, en este caso fue un nomograma (Fig. 65). El nomograma permite estimar un riesgo para desarrollar el fenotipo de artrosis de rodilla con progresión rápida según las variables clínicas y genéticas del individuo. Si por ejemplo quisiéramos calcular el riesgo de una paciente de 65 años, con sobrepeso (IMC=26), 14 puntos en la escala de dolor WOMAC, con historia de lesión previa en la rodilla y artrosis contralateral, que además es portadora de las variantes genéticas m.16519C, rs12107036.G y el haplogrupo UK; se sumarían un total de 272 puntos, los cuales se traducirían en una probabilidad de desarrollar el fenotipo de artrosis de rodilla con progresión rápida cercana al 80%.

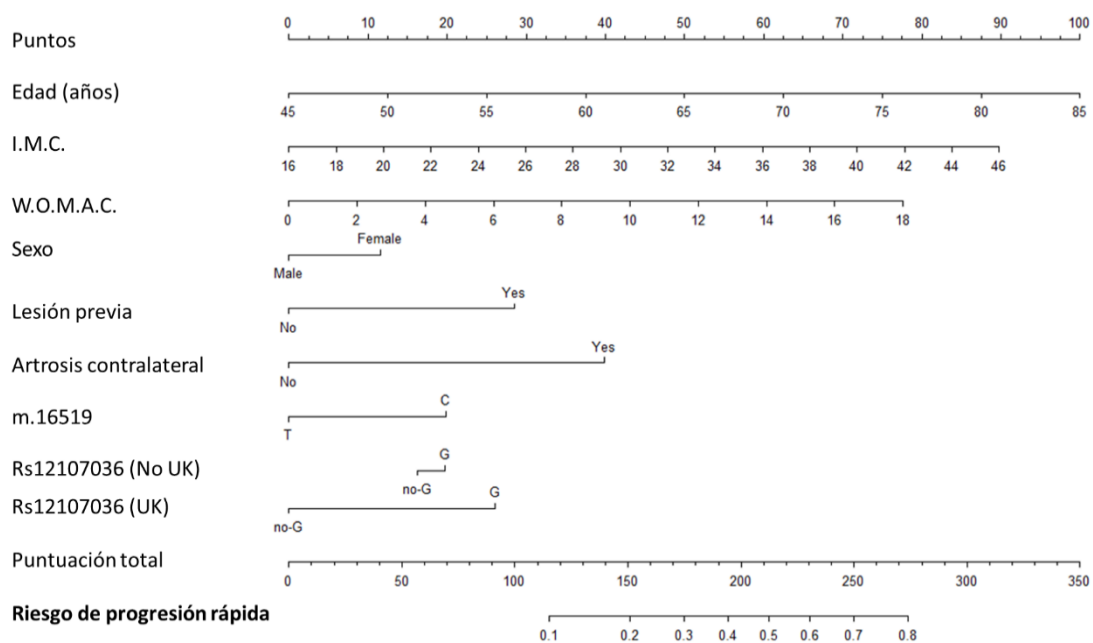


Figura 65. Nomograma de aplicación clínica directa resultante del modelo predictivo para el fenotipo de artrosis de rodilla con progresión rápida.



## *Discusión*



*La progresión rápida en la artrosis de rodilla.*

La progresión en la artrosis presenta una característica muy destacable, su heterogeneidad. Normalmente, esta patología se puede considerar un trastorno lentamente progresivo, en el que la mayoría de los pacientes empeoran gradualmente durante años o décadas (Emrani *et al.*, 2008); sin embargo, un pequeño porcentaje de pacientes muestra una rápida evolución o progresión estructural de la enfermedad (Riddle *et al.*, 2016). Se estima que uno de cada siete pacientes de artrosis de rodilla desarrollará este fenotipo, y su prevalencia en la población general es aproximadamente del 3,4% (Driban *et al.*, 2020).

El concepto de artrosis con progresión rápida no es nuevo, ya que en 1957 fue descrito por Forestier en la articulación de la cadera (Forestier, 1957). Desde entonces hasta ahora se ha encontrado este fenotipo de artrosis en diversas articulaciones además de la cadera, y también se han realizado numerosas definiciones para este fenotipo, destacando algunas como la artrosis acelerada, la progresión rápida, la artrosis rápidamente progresiva, la artrosis rápidamente destructiva o el cambio radiográfico rápido (Driban *et al.*, 2020). Debido a ello, no se ha alcanzado un consenso para describir este fenotipo, lo cual produce que su tasa de prevalencia pueda variar entre estudios (Driban, Stout, *et al.*, 2016). Sin embargo, sí se han llegado a crear definiciones bastante aceptadas; entre estas podemos destacar la artrosis rápidamente progresiva de tipo 1 (Lequesne y Amouroux, 1970), caracterizada por una reducción del JSW mayor o igual a 2 mm en un año sin destrucción ósea, la artrosis rápidamente progresiva de tipo 2 (Hochberg, 2015), caracterizada por la drástica destrucción ósea, y también podrían citarse la definición de la artrosis incidente rápidamente progresiva (Riddle *et al.*, 2016) y la artrosis acelerada (Driban *et al.*, 2014).

Concretamente esta última definición, la artrosis acelerada, se centra principalmente en la artrosis de rodilla, y se define tradicionalmente como un establecimiento y progresión desde un grado KL de 0 o 1, hasta un grado de severidad muy elevado (grado KL 3 o 4) en un período de 4 años o menos (Driban *et al.*, 2014). El desarrollo de la artrosis acelerada comienza al menos dos años antes de la aparición de evidencias radiográficas, y las etapas previas a su desarrollo se pueden subdividir en tres fases (Driban *et al.*, 2020):

1. Dos o más años previos al desarrollo radiográfico: los adultos con artrosis de rodilla acelerada pueden experimentar sinovitis y efusión de líquido sinovial, presentar un cartílago articular más grueso, una patología meniscal difusa y cambios degenerativos en los ligamentos cruzados y el mecanismo extensor.
2. Uno o dos años antes del desarrollo radiográfico: se observan mayores tasas de sinovitis efusional junto con lesiones de la médula ósea y el cartílago articular.
3. Un año o menos antes del desarrollo radiográfico: aparecen más síntomas articulares, los pacientes llevan a cabo un uso frecuente de medicamentos para el dolor, se observa una hinchazón frecuente de la rodilla y dolor de rodilla diario.

Ésta definición de artrosis acelerada se diferencia de la descrita por Wesseling como progresión rápida (Wesseling *et al.*, 2015) en que la primera, partiendo de un grado KL de 0 o 1 se llega a un grado KL 3 o 4 en 48 meses, mientras que en la segunda, partiendo de un grado KL 0 se avanzan dos o más grados en 48 o 60 meses. Esto puede provocar que se definan como progresores rápidos aquellos individuos que partiendo de un grado KL 0 llegan, en cuatro o cinco años, a incidir con un grado KL 2; de esta manera no todos los individuos considerados como progresores rápidos mostrarán un grado de severidad de artrosis tan elevado como en la definición de artrosis acelerada. La artrosis incidente rápidamente progresiva fue descrita por

## Discusión

Riddle (Riddle *et al.*, 2016) y es muy similar a la artrosis acelerada, la principal diferencia es el tiempo de seguimiento, siendo de cuatro años en la artrosis acelerada y de cinco en la artrosis incidente rápidamente progresiva.

Las causas que puedan provocar este fenotipo son desconocidas a día de hoy, y actualmente se han propuesto diferentes factores de riesgo potenciales, como por ejemplo el incorrecto alineamiento de la rodilla (Walker *et al.*, 2011; Driban *et al.*, 2014), los daños en los meniscos (Walker *et al.*, 2011; Driban *et al.*, 2014), los defectos en el hueso subcondral (Mitrovic y Riera, 1992; Blair-Levy *et al.*, 2008) o la osteonecrosis (Mavrogenis *et al.*, 2015; Hart y Fehring, 2016). Gracias a la extensa información disponible en las bases de datos de las cohortes utilizadas en esta tesis, para desarrollar los modelos predictivos del fenotipo de artrosis de rodilla con progresión rápida hemos tenido en cuenta la edad, el sexo, el IMC, la contralateralidad, la lesión previa, el dolor WOMAC y distintas variables genéticas. Con respecto a éstas últimas, a pesar de la más que demostrada implicación de la mitocondria y de la variación del ADNmt en diferentes aspectos relacionados con la artrosis, su impacto en el desarrollo de este fenotipo no ha sido explorado hasta la fecha, lo cual es un enfoque pionero en el campo de esta patología.

El interés por estudiar este fenotipo de artrosis resurgió en los últimos años debido a su aparición, como efecto adverso, en ensayos clínicos del fármaco tanezumab (Hochberg, 2015; Hochberg *et al.*, 2016). El tanezumab es un anticuerpo monoclonal contra el NGF, el cuál demostró una gran eficacia en ensayos clínicos contra el dolor crónico producido por la artrosis de rodilla (Brown *et al.*, 2012) y de cadera (Brown *et al.*, 2013). Durante la fase III de los ensayos clínicos, se observó que 87 pacientes mostraban como efecto adverso la osteonecrosis, lo cual produjo que la *Food and Drug Administration* (FDA) de Estados Unidos llevase a cabo una retención clínica parcial para todas las indicaciones con este medicamento, a excepción del dolor por cáncer. En 2016 se llevó a cabo un estudio con el fin de revisar los efectos adversos del tanezumab, y lo que se observó es que el tanezumab no se asociaba con osteonecrosis, sino con la artrosis con progresión rápida caracterizada por la pérdida acelerada del cartílago articular y la destrucción del hueso subcondral, también definida como progresión rápida de tipo 2 (Hochberg, 2015). Concretamente, altas dosis de tanezumab o el uso de este anticuerpo monoclonal en combinación con otros AINES podían aumentar el riesgo de desarrollar este tipo de progresión (Hochberg *et al.*, 2016). Actualmente, tras la última actualización del 27 de marzo de 2020, el tanezumab continúa sin estar aprobado por la FDA (<https://www.drugs.com/history/tanezumab.html#>). En Europa, la *European Medicines Agency* (EMA) tampoco ha concedido la autorización pertinente para el uso de este compuesto en la unión Europea (<https://www.ema.europa.eu/en/medicines/human/EPAR/raylumis>).

Debido a este repunte en el interés por el fenotipo, organizaciones destacadas como la EULAR, establecieron como una de las líneas de investigación prioritaria el desarrollo de herramientas capaces de identificar en la población a aquellos individuos con elevado riesgo de desarrollar el fenotipo (Conaghan *et al.*, 2014). Gracias a estas herramientas, se podría realizar un seguimiento más exhaustivo de estos sujetos y, a mayores, incluirlos en ensayos clínicos para desarrollar fármacos contra la artrosis, ya que sirven como un modelo a pequeña escala de la enfermedad. Esta es la necesidad que trata de solventar este proyecto de tesis doctoral, la generación de un modelo predictivo capaz de identificar a sujetos con elevado riesgo de desarrollar el fenotipo rápido de artrosis de rodilla. Para llevarlo a cabo, en el desarrollo de esta tesis se han empleado dos tipos de criterios para definir la progresión rápida, ambos basados en la artrosis acelerada, denominados como Driban y GIR-blanco. El primero aplica exactamente la

definición de artrosis acelerada descrita en 2014 (Driban *et al.*, 2014), incluyendo única y exclusivamente sujetos con artrosis de rodilla pre-radiográfica al inicio del seguimiento (KL grado 0-1). Sin embargo, dado que en este estudio se contaba con cohortes prospectivas bien caracterizadas y con un tiempo de seguimiento bastante largo, se propuso desarrollar una herramienta predictiva que permitiera identificar sujetos propensos a sufrir una evolución rápida de la enfermedad en términos del empeoramiento radiográfico de la articulación. Por lo tanto, se generó un segundo criterio que abarca tanto a individuos con artrosis acelerada como individuos con artrosis establecida (KL grado 2), denominado criterio GIR-blanco. Bajo este criterio se incluyeron sujetos con artrosis de rodilla pre-radiográfica al inicio (KL grado 0-1) y sujetos con artrosis de rodilla consolidada al inicio (KL grado 2), pero que en ambos casos su evolución, a cuatro años, consistiría en la consecución de un grado KL 3-4 o 4 respectivamente.

Con el fin de determinar qué criterio era el más adecuado, se decidió comparar ambos tipos de criterios mediante un análisis estadístico de correlación y, el coeficiente de correlación *kappa* fue muy elevado. Este dato demostró que ambos criterios clasificaban de forma casi idéntica a los 852 participantes que compartían. Por ello, se decidió emplear para el desarrollo de este proyecto de tesis únicamente los criterios GIR-blanco ya que presentan un tamaño muestral superior, permitiendo obtener un mayor poder estadístico y, por otro lado, este nuevo criterio tiene en cuenta a pacientes que basalmente ya han incidido, es decir, presentan un KL 2 al inicio del seguimiento. Esto es especialmente importante a nivel clínico ya que, normalmente cuando un paciente llega a consulta con dolor, incapacidad o pérdida de función en la rodilla, suelen presentar artrosis de rodilla, siendo poco frecuente que lleguen al reumatólogo con un grado KL 0-1.

#### *Cohortes prospectivas para el estudio de la artrosis.*

En las dos fases que se llevaron a cabo para el desarrollo de esta tesis, exploración y validación, se trabajó con cohortes prospectivas. Los estudios con cohortes son estudios observacionales compuestos por un conjunto de participantes que cumplen una serie de características basales comunes, a los que se les hace un seguimiento durante un período de tiempo determinado. Esto permite generar subgrupos de participantes en función de características concretas, permitiendo así comparar dichos subgrupos en distintos períodos de tiempo, con el fin de tratar de buscar asociaciones entre un resultado (por ejemplo: incidir en determinada patología) y alguna característica particular (por ejemplo: ser fumador) (Merrill, R.M, 2015). Los estudios con cohortes pueden clasificarse como prospectivos o retrospectivos, esto dependerá de cuándo ocurre el resultado que se quiera detectar; si este no ha ocurrido al inicio del estudio entonces será una cohorte prospectiva, por el contrario, si este ya ha ocurrido al inicio del estudio, será una cohorte retrospectiva (Euser *et al.*, 2009).

Una gran ventaja del estudio de cohortes es la capacidad de estudiar múltiples resultados que pueden asociarse con una o varias variables. Incluso se puede determinar el efecto combinado de múltiples características en el resultado. Los estudios de cohortes también permiten el estudio de exposiciones raras y los investigadores pueden seleccionar específicamente sujetos expuestos a un determinado factor. La naturaleza longitudinal de los estudios de cohortes significa que los cambios en los niveles de exposición a diversos factores a lo largo del tiempo y los cambios en el resultado pueden medirse para proporcionar información sobre la relación dinámica entre la variable de interés y el resultado. Los estudios prospectivos y retrospectivos tienen diferentes fortalezas y debilidades. En primer lugar, los estudios de cohortes prospectivos se llevan a cabo desde el presente hasta el futuro y, por lo tanto, tienen la ventaja de ser precisos en cuanto a la información recopilada sobre variables, puntos finales

## Discusión

y factores de confusión. La desventaja podría ser el largo período de seguimiento mientras se espera que ocurran los eventos, lo que lleva una alta tasa de pérdida de seguimiento de los participantes. En segundo lugar, los estudios retrospectivos se basan en datos recopilados en el pasado para identificar tanto las variables como los resultados. Estos estudios utilizan datos que ya han sido recopilados, haciéndolos eficientes en términos de tiempo y coste. Sin embargo, muchos estudios de cohortes retrospectivos utilizan datos recopilados en el pasado con otro objetivo diferente al que los investigadores quieren estudiar en sus trabajos, por lo que ellos carecen de control sobre la recopilación de datos, provocando que la medición de las variables puede ser inexacta o inconsistente, lo que resulta en una fuente de sesgo de información (Wang y Kattan, 2020).

Los estudios de enfermedades humanas usando cohortes prospectivas han permitido asociar la presencia de diversas variables con un gran número de patologías, entre las cuales también está presente la artrosis. Concretamente, la artrosis es el diagnóstico más común realizado en pacientes mayores con dolor de rodilla o cadera. El diagnóstico puede basarse en síntomas, signos y hallazgos radiográficos y, como tal, puede definirse mediante varios conjuntos y combinaciones de criterios. Esto provoca que el pronóstico de la artrosis para el paciente individual sea incierto; el curso de los síntomas, signos clínicos, discapacidad y cambios radiográficos son difíciles de predecir (Wesseling *et al.*, 2009). Además, se ha demostrado que existe una inconsistencia manifiesta entre el cambio radiográfico y la severidad del dolor articular con discapacidad acompañante (Creamer *et al.*, 1999). Por tanto, debido a estas dificultades, son necesarias grandes cohortes prospectivas que comiencen temprano en la etapa de la enfermedad con el fin de saber más de ella. Aquí es donde radica la importancia de las cohortes prospectivas en la artrosis. Entre las cohortes más empleadas en estudios prospectivos de artrosis podemos citar la OAI (Eckstein *et al.*, 2007), la CHECK (Wesseling *et al.*, 2016), la PROCOAC (Oreiro-Villar *et al.*, 2022), la *DIGital COhort Design* (DIGICOD) (Sellam *et al.*, 2021), la *Multicentre Osteoarthritis* (MOST) (Perry *et al.*, 2020) o la TASOAC (*Tasmanian Older Adult Cohort*) (Doré *et al.*, 2012).

De las cohortes mencionadas previamente, en esta tesis doctoral se emplearon la OAI, la CHECK y la PROCOAC. En la fase de exploración se comenzó trabajando con la cohorte OAI, ya que es la que tiene un mayor número de participantes (N=4.796). Además, esta cohorte comenzó en 2004, y realiza seguimientos clínicos y radiográficos anuales (los cuatro primeros años) y bianuales (sexto, octavo año y décimo año) (<https://www.niams.nih.gov/grants-funding/funded-research/osteoarthritis-initiative>). Estas características tan rigurosas hacen que sea una cohorte muy robusta para realizar estudios prospectivos de diversas patologías, de hecho, al realizar una búsqueda rápida en *pubmed* en septiembre de 2022 con el término “*data from the osteoarthritis initiative*” se encuentran un total de 794 publicaciones que emplean esta cohorte en sus trabajos. En la fase de validación, se emplearon participantes de las cohortes europeas CHECK y PROCOAC. La primera está compuesta por participantes con artrosis sintomática temprana de cadera o de rodilla y consta de un total de 1002 participantes que fueron reclutados entre los años 2002 y 2005, a los que se les realizó un seguimiento hasta los 10 años. Al igual que en la cohorte OAI, existen unos períodos de seguimiento bien definidos donde a los dos, cinco, ocho y diez años se realizan visitas clínicas en las que se obtienen imágenes radiológicas (Wesseling *et al.*, 2016). Tanto la cohorte OAI como la cohorte CHECK presentan unos criterios de inclusión y de exclusión que las sitúan como cohortes tempranas de artrosis. Sin embargo, cabe destacar que, al comparar ambas cohortes, se observó que, basalmente, la artrosis radiográfica de rodilla estaba presente en un menor número de participantes en la cohorte CHECK en comparación con la OAI y, por el contrario, la artrosis

sintomática de rodilla estaba presente en un mayor porcentaje de individuos pertenecientes en la cohorte CHECK que en la OAI. Esto sitúa a la cohorte CHECK como una cohorte todavía más temprana de la enfermedad (Wesseling *et al.*, 2009).

La última cohorte empleada, la PROCOAC, lleva a cabo el seguimiento de pacientes diagnosticados con artrosis, permitiendo realizar estudios de incidencia, progresión y fenotipos de la enfermedad. En 2018, la cohorte estaba compuesta por un total de 937 participantes, de los cuales 873 presentan artrosis radiográfica de rodilla, 783 artrosis radiográfica de cadera y 679 artrosis de manos. Actualmente forman parte de esta cohorte 1.537 pacientes. A diferencia de la OAI y la CHECK, las cuales se centran respectivamente en artrosis de rodilla y artrosis de rodilla o cadera, la cohorte PROCOAC incluye la afectación en las manos. La artrosis de rodilla es más notable en la PROCOAC y en la OAI que en la CHECK, y en el caso de la artrosis de cadera, esta también es más notable en la PROCOAC que en la CHECK. Finalmente cabe destacar que, las cohortes CHECK y OAI tienen un objetivo distinto a la PROCOAC, ya que las dos primeras se centran en la detección temprana de la artrosis. Es debido a esto que, en comparación con las otras dos cohortes, la PROCOAC cuenta con un número de individuos radiológicamente sanos muy escaso, ya que, se centra principalmente en el avance de la enfermedad y el estudio de los fenotipos y actitudes terapéuticas frente a estos fenotipos (Oreiro-Villar *et al.*, 2022).

#### *Discusión de los resultados*

En esta tesis doctoral, además de incorporar y analizar las variables clínicas que se han empleado de forma rutinaria en el desarrollo de modelos predictivos, también se incluyeron variables genéticas. Aunque tradicionalmente la artrosis no está considerada como una enfermedad hereditaria (Brody, 2015), muchos estudios han puesto de manifiesto que la presencia de determinados polimorfismos nucleares y mitocondriales pueden condicionar su desarrollo (Soto-Hermida *et al.*, 2015; Palazzo *et al.*, 2016; Fernández-Moreno *et al.*, 2017; Bonakdari *et al.*, 2022). La importancia de la inclusión de este tipo de biomarcadores radica en que éstos podrían mejorar la precisión de modelos predictivos de diagnóstico y pronóstico y, por tanto, mejorar enormemente medicina personalizada de esta enfermedad y el desarrollo de posibles tratamientos.

Teniendo en cuenta esto, este estudio ha tratado de analizar cómo determinadas variables genéticas pueden condicionar el riesgo de desarrollar un fenotipo específico de artrosis de rodilla, la progresión rápida. Para ello se han llevado a cabo tres objetivos pioneros en el campo de la artrosis. En primer lugar, se ha estudiado mediante técnicas de secuenciación masiva la totalidad del ADNmt, con el objetivo de identificar potenciales variantes mitocondriales que condicionen el riesgo de progresar rápidamente. En segundo lugar, se han analizado por primera vez posibles interacciones entre el ADNmt y polimorfismos genéticos nucleares que puedan condicionar el riesgo de desarrollar el fenotipo rápido. Y en tercer y último lugar, se ha tratado de desarrollar una herramienta predictiva de aplicación clínica, que incluye tanto variables clínicas como variables genéticas nucleares, mitocondriales y la interacción entre éstas.

## Discusión

*Influencia de la variante m.16519 en el fenotipo de artrosis de rodilla con progresión rápida, posibles efectos funcionales en la mitocondria y en la célula e hipótesis acerca de su impacto en la artrosis.*

Además de los haplogrupos, existen otras variantes genéticas mitocondriales que, a título individual, tienen la capacidad de modificar el riesgo de desarrollar diversas patologías o fenotipos no clínicos, como por ejemplo el cáncer de próstata (Petros *et al.*, 2005; Arnold *et al.*, 2013), la diabetes mellitus de tipo 2 (Poulton *et al.*, 1998), el LHON (Ortiz *et al.*, 1993; Farrar *et al.*, 2013), la longevidad (Chocron *et al.*, 2019), el MELAS (Pavlakakis *et al.*, 1984), el NARP (Ortiz *et al.*, 1993), síndrome de Leigh (Ortiz *et al.*, 1993) o la sordera (Ortiz *et al.*, 1993). En esta tesis se llevó a cabo un enfoque pionero en el campo de la artrosis, el uso de tecnologías de NGS para buscar posibles variantes genéticas mitocondriales que condicionen el fenotipo de artrosis de rodilla con progresión rápida. Para llevarlo a cabo se trabajó con plataformas de secuenciación de segunda generación de *Ion Torrent* (ThermoFisher), las cuales cuentan con paneles de oligonucleótidos validados y testados por la compañía que permiten amplificar regiones precisas del genoma humano, en este caso concreto se trabajó con un panel que permite secuenciar la totalidad de la molécula de ADNmt.

Como resultado del análisis exploratorio del ADNmt en la cohorte OAI, se detectó una variante que estaba sobrerrepresentada de forma significativa dentro de los progresores rápidos del criterio GIR-blanco, asociándose de forma significativa con el riesgo de desarrollar el fenotipo de progresión rápida en la cohorte OAI. Esta variante estaba presente en la posición 16519 y consistía en un cambio de timina por una citosina (m.16519C). En la fase de validación de los datos obtenidos en la cohorte OAI, se analizó mediante técnicas de secuenciación capilar en el *pool* de cohortes externas CHECK y PROCOAC y, se observó que, en el análisis univariado, la variante se asociaba nuevamente, de forma significativa, con la progresión rápida, sin embargo, al realizar el ajuste de este p-valor mediante la regresión logística, la variante m.16519 mostraba una tendencia similar a la observada durante la fase de exploración, que no llegaba a alcanzar valores significativos, por lo que se decidió finalmente tratar de validar los datos mediante un meta-análisis de las cohortes. El meta-análisis mostró resultados satisfactorios, revelando que la variante en cuestión aparecía sobre-representada significativamente dentro de los progresores rápidos, incrementando el riesgo de desarrollar el fenotipo de artrosis de rodilla con progresión rápida. Además, de forma complementaria se llevó a cabo un meta-análisis con las tres cohortes por separado (sin agrupar CHECK y PROCOAC), y de nuevo se observó una asociación significativa entre la variante m.16519C y el riesgo de desarrollar fenotipo de artrosis de rodilla con progresión rápida.

La variante m.16519C se encuentra en la región *D-loop* del ADNmt, la cual es la zona de esta molécula con una mayor tasa de mutación, debido a que en ella se encuentran los elementos reguladores de la transcripción y replicación del ADNmt (L. Wang *et al.*, 2021). Diversos estudios han puesto de manifiesto que la presencia de mutaciones en esta región se asocia con el riesgo de desarrollar determinadas patologías como por ejemplo la enfermedad de Huntington (Mousavizadeh *et al.*, 2015) o la enfermedad renal crónica (Bai *et al.*, 2014). Y no solo la presencia de variantes en esta región se asocia con patologías, sino que en los últimos años ha surgido el concepto de mitoepigenética. Este nombre hace referencia a los cambios epigenéticos que ocurren sobre la molécula de ADNmt y que afectan a su replicación y expresión génica (van der Wijst y Rots, 2015). Este concepto ha sido, y continúa siendo, objeto de debate desde hace años, principalmente debido a limitaciones metodológicas y a las diferencias presentes con respecto al ADN nuclear, ya que el ADNmt carece de islas CpG, está organizado



en nucleoides y no presenta histonas (Coppedè y Stocco, 2019). Se hipotetiza que el principal método de regulación epigenética ocurre mediante metilación del ADNmt e hidroximetilación (Stimpfel *et al.*, 2018). Sin embargo, cabe destacar que los niveles de metilación e hidroximetilación presentes en el ADNmt son mucho menores que los presentes en el ADN nuclear, pero a pesar de ser pequeños, juegan un papel fundamental en la replicación y transcripción de esta molécula. En esta línea se han encontrado correlaciones inversas con los niveles de metilación en la región *D-loop* del ADNmt con el número de copias de ADNmt en células sanguíneas periféricas (Sanyal *et al.*, 2018), en muestras de tejido de cáncer colorrectal (Gao *et al.*, 2015) y en la placenta humana (Janssen *et al.*, 2015). También se han encontrado correlaciones entre los niveles de metilación del ADNmt y la expresión génica en células mesenquimales estromales (Yu *et al.*, 2017), en muestras de tejidos de cáncer colorrectal (Feng *et al.*, 2012) y en muestras de sangre periférica (Sanyal *et al.*, 2018). Finalmente, los patrones de metilación del ADNmt se han asociado con diversas enfermedades humanas, tales como la arterioesclerosis y/o eventos trombóticos agudos (Baccarelli y Byun, 2015), el cáncer colorrectal (Tong *et al.*, 2017), la obesidad (Zheng *et al.*, 2015), la disfunción mitocondrial en pacientes con síndrome de Down (Infantino *et al.*, 2011) y las enfermedades neurodegenerativas como el Parkinson (Blanch *et al.*, 2016; Stocco y Coppedè, 2021), el Alzheimer (Blanch *et al.*, 2016; Stocco *et al.*, 2017; Stocco y Coppedè, 2018) o la esclerosis lateral amiotrófica (Stocco *et al.*, 2018).

Distintos trabajos han relacionado la presencia de ambos alelos de la posición m.16519 con un amplio abanico de enfermedades; por ejemplo, la presencia del alelo T se ha asociado con un mayor riesgo de desarrollar diabetes mellitus de tipo 2 asociada al cáncer de páncreas (Navaglia *et al.*, 2006), también con migrañas y el síndrome de vómitos cíclicos (Zaki *et al.*, 2009), el síndrome del intestino irritable con herencia materna (Van Tilburg *et al.*, 2014) y el edema pulmonar en altura (Yan Wang *et al.*, 2022). Por otro lado, la presencia del alelo C se ha asociado con el síndrome de muerte súbita del lactante en mujeres (Läer *et al.*, 2014), la disminución de la tasa de cambio del VO<sub>2</sub>max y de la actividad de la citrato sintasa como resultado del entrenamiento de resistencia (Murakami *et al.*, 2002), el riesgo de desarrollar diabetes mellitus de tipo 2 (Liao *et al.*, 2008), la esquizofrenia (Sequeira *et al.*, 2015) o la enfermedad de Huntington (Mousavizadeh *et al.*, 2015). También se ha descrito que este alelo empeora el pronóstico del cáncer colorrectal y de mama, ya que se piensa que puede producir un bloqueo metabólico y resistencia a la apoptosis (Bai *et al.*, 2007; Akouchekian *et al.*, 2009).

En el caso de la artrosis, hasta el desarrollo de esta tesis doctoral no existía ningún trabajo que hubiera realizado un estudio funcional para analizar cómo esta variante puede afectar a alguna patología en concreto. Fue por ello que se planteó como un objetivo específico de esta tesis el inicio del estudio funcional con el fin de identificar procesos que vinculen la variante y la patología de interés. Para ello se generaron dos líneas de híbridos a partir de la línea de osteosarcoma 143B.TK Rho-0 y plaquetas de donantes sanos portadores del haplogrupo H, poseyendo una línea la variante m.16519C y la otra el alelo salvaje m.16519T. Con estos híbridos se llevaron a cabo técnicas de RNAseq que permitieron detectar un total de 460 DEGs, de los cuales 167 estaban sobre-expresados en los híbridos portadores de la variante m.16519C, siendo dos de los más destacados, especialmente por su papel en la artrosis, la IL6 y COX2. Además, empleando estos DEGs, se pudieron identificar diversos procesos biológicos que podían estar potencialmente afectados debidos a la presencia de la variante, entre ellos la respuesta inflamatoria aguda. Finalmente, aprovechando la disponibilidad de datos relativos a los niveles de IL6 séricos en algunos participantes de la cohorte OAI, se pudo observar que los progresores

## Discusión

rápidos mostraban unos niveles basales de IL6 significativamente mayores que los progresores no rápidos.

Como se explicó al inicio de esta tesis, el correcto balance entre factores anabólicos y catabólicos es clave para el correcto mantenimiento del cartílago articular. Inicialmente, la artrosis era considerada una enfermedad de rozamiento y desgaste, pero este paradigma cambió con el tiempo, destacando la importancia del proceso inflamatorio en el desarrollo de la patología. Los factores inflamatorios que intervienen en esta patología pueden provenir del hueso subcondral, del propio cartílago o de la membrana sinovial (Berenbaum, 2013). Especialmente esta último tejido, la membrana sinovial, puede sufrir un proceso de inflamación denominado como sinovitis, el cuál se ha asociado con la incidencia y progresión de artrosis radiográfica (Ayrál *et al.*, 2005; Roemer *et al.*, 2011). Las principales moléculas que se considera que participan en este proceso son IL1 $\beta$ , TNF $\alpha$ , IL8 e IL6, entre otras muchas. Estas citoquinas son capaces de estimular su propia expresión y activar a los condrocitos, promoviendo que sintetizen MMPs, NO, proteasas, prostaglandinas y otras moléculas, las cuales favorecen la degradación del cartílago (Rahmati *et al.*, 2016). Esta inflamación de la membrana sinovial suele co-ocurrir en lesiones traumáticas de menisco, las cuales son relevantes para la incidencia y progresión de la patología artrósica, pero la causa exacta por la que se inflama la membrana sinovial es desconocida, siendo una de las hipótesis más aceptadas que, una vez degradado el cartílago articular, los fragmentos entran en contacto con la membrana sinovial y, debido a que los propios sinoviocitos que la componen detectan a estos fragmentos como cuerpos extraños, éstos producen los mediadores inflamatorios, activando a los condrocitos presentes en la capa superficial del cartílago, los cuales generan metaloproteinasas para degradar estos fragmentos de cartílago degradado, pero incrementando también la degradación del propio cartílago articular en un círculo vicioso (Sellam y Berenbaum, 2010).

Concretamente en el caso de la IL6, ésta es una glicoproteína que consta de 184 residuos de aminoácidos, que tras el procesamiento postraduccional, se convierte en una estructura interconectada de cuatro hélices  $\alpha$  (Hammacher *et al.*, 1994). En pacientes con artrosis, se ha observado un incremento en los niveles de IL6 en el líquido sinovial y suero de pacientes con artrosis (Kaneko *et al.*, 2000). La IL6 puede actuar sinérgicamente con otras citoquinas importantes en el proceso artrósico, como por ejemplo la IL1 $\beta$  y el TNF $\alpha$ , produciendo una reducción en los niveles de producción de colágeno II e incrementando la producción de MMPs (Rowan *et al.*, 2001; Porée *et al.*, 2008). De hecho, estos efectos catabólicos de la IL6 y la IL1 $\beta$  se pueden ver potenciados en presencia de una lesión previa, ya que este proceso libera en el tejido adyacentes citoquinas proinflamatorias, incrementando todavía más el proceso inflamatorio (Sui *et al.*, 2009). A este último dato, hay que sumarle que en el caso de la artrosis de rodilla con progresión rápida, la lesión previa se ha propuesto como un importante factor de riesgo (Driban *et al.*, 2014). También la IL6 puede interaccionar con otras citoquinas y producir conjuntamente alteraciones en el hueso subcondral, promoviendo la formación de osteoclastos y favoreciendo así la pérdida de hueso (Steeve *et al.*, 2004). En línea con lo citado, en 2022 usando modelos animales de ratón, se puso de manifiesto que la IL6 es capaz de modular la degradación del cartílago y el dolor de la artrosis, postulando a la IL6 como una interesante diana terapéutica para el posible desarrollo de DMOADs contra esta patología (Liao *et al.*, 2022). Por tanto, teniendo en cuenta todo lo comentado en este párrafo, una posible explicación del efecto de la variante m.16519C en la progresión rápida podría ser que la presencia de ésta favorezca el desarrollo del fenotipo de artrosis de rodilla con progresión rápida alterando los niveles de citoquinas proinflamatorias como la IL6, ya que éstas moléculas tienen un papel clave en el mantenimiento del cartílago y el hueso subcondral, fomentando así la pérdida de cartílago

articular y, por tanto, la reducción de mJSW, un endotipo muy relevante en la progresión de la artrosis de rodilla.

Por otro lado, la proteína resultante del gen COX2 también es una glicoproteína perteneciente a la familia de las ciclooxigenasas (Park *et al.*, 2006). Existen tres isoformas de COX, todas ellas capaces de sintetizar a partir de moléculas de ácido araquidónico unas moléculas lipídicas denominadas prostaglandinas; la primera, COX-1, se expresa de forma constitutiva y se encarga de la regulación homeostática del organismo generando diversas prostaglandinas, como por ejemplo la regulación de la mucosa gástrica e influyendo el flujo de sangre renal entre otras funciones. La segunda, COX2, es una enzima inducible que, en respuesta a un traumatismo, estímulo o molécula señalizadora genera estas prostaglandinas, las cuales son mediadores capaces de inducir efectos inflamatorios y el dolor, motivo por el cual muchos AINES son inhibidores específicos de COX2 (Simon, 1999). Finalmente, COX3, la cuál deriva de COX1, pero es estructural y funcionalmente distinta, teniendo un rol fundamental en la producción de fiebre y dolor a nivel central pero no inflamatorio (Pradilla-Vesga, 2004). En lo referente a COX2, como se ha mencionado, esta puede ser activada mediante moléculas señalizadoras, y cabe destacar, que entre estas moléculas señalizadoras se han identificado a NF- $\kappa$ B e IL6 entre otras (Hashemi-Goradel *et al.*, 2019). Teniendo en cuenta este último dato y los resultados de esta tesis doctoral en lo referente a la variante m.16519C e IL6, es probable que, la presencia de esta variante mitocondrial pueda también incrementar los niveles de COX2 y por tanto la síntesis de prostaglandinas, favoreciendo el proceso artrósico, e incrementando enormemente el dolor, algo muy frecuente en la progresión rápida. Además, estos individuos portadores de la m.16519C parece que también tienen mayores niveles de IL6, que podría actuar de forma sinérgica con COX2, induciendo esta enzima y generando todavía más prostaglandinas, acelerando mucho más el proceso artrósico.

Otros posibles efectos de la variante m.16519C en la progresión rápida han surgido recientemente gracias a que el grupo GIR-blanco ha profundizado todavía más en el ensayo funcional de la variante m.16519C (Durán-Sotuela *et al.*, 2022). Los autores emplearon las mismas líneas de híbridos usadas en esta tesis doctoral para realizar otros ensayos y, como resultado, observaron que los híbridos portadores de la variante m.16519C presentan mayor número de copias de ADNmt que los híbridos sin la variante. También, en comparación con los híbridos con el alelo salvaje, en los híbridos m.16519C detectaron unos niveles menores de expresión e PGC1 $\alpha$  (del inglés *peroxisome proliferator-activated receptor- $\gamma$  coactivator 1 $\alpha$* ), el cuál cumple un papel importante dentro de la biogénesis mitocondrial, y también menores niveles de FIS1 (del inglés *mitochondrial fission protein 1*) importante en los procesos de fusión y fisión mitocondrial; por otro lado, se detectaron en los híbridos m.16519C mayores niveles de la proteína p62 (*sequestosome 1*), LC3-II (del inglés *microtubule-associated protein 1A/1B-light chain 3*) y p-rpS6 (del inglés *phospho-ribosomal protein S6*), los cuales cumplen un papel fundamental en la autofagia. Finalmente, también realizaron ensayos de producción de ROS mitocondrial y respuestas a estrés oxidativo, y de nuevo, los híbridos portadores de la variante m.16519C generaban mayores niveles de O $_2^-$ , y presentaban una menor supervivencia celular en presencia de H $_2$ O $_2$ .

Todos los procesos descritos en el párrafo anterior pueden relacionarse de una u otra forma con alteraciones que podrían condicionar la patología artrósica. En primer lugar, el número de copias de ADNmt está regulado para asegurar que la mitocondria sea capaz de mantener sus funciones, entre ellas la señalización celular y la generación de energía (Zhang *et al.*, 2017). Se ha observado en patologías como el autismo (Gu *et al.*, 2013), el carcinoma

## Discusión

escamoso laríngeo (Guo *et al.*, 2013), en la lesión renal crónica en pacientes hipertensos (Eirin *et al.*, 2016) e, incluso la artrosis (Fang *et al.*, 2014), que el número de copias está aumentado, y se sospecha que esto consiste en un mecanismo que trata de compensar los bajos niveles de energía producidos por la CTE (Eirin *et al.*, 2016). En segundo lugar, la biogénesis mitocondrial también se vio afectada en los cíbridos m.16519C y, del mismo modo que ocurre con éstos cíbridos, en estudios llevados a cabo en condrocitos humanos también se han observado menores niveles de biogénesis mitocondrial (Wang *et al.*, 2015). En línea con esto, los menores niveles de FIS1 podrían afectar a los procesos de fusión y fisión mitocondrial, y por tanto a la biogénesis. Además, FIS1 se ha asociado con la autofagia en estudios *in vitro*, los cuales revelaron que los niveles de esta proteína están reducidos en los condrocitos artrósicos, y que la depleción de FIS1 lleva a una acumulación de agregados de LC3, inhibiendo la autofagia, acumulando daños a nivel celular y, favoreciendo así la apoptosis del condrocito (Kim *et al.*, 2016). Finalmente, en tercer lugar, la variante m.16519C también parece aumentar los niveles de ROS mitocondriales, pudiendo superar la capacidad antioxidante del condrocito articular y, por otro lado, reduce la supervivencia de la célula en presencia de estrés oxidativo. Esta acumulación de daño oxidativo puede producir lesiones celulares, una activación de las metaloproteasas, sobreproducción de citoquinas pro-inflamatorias, modificar la señalización celular y degradar el cartílago, contribuyendo con todos estos procesos a la progresión de la artrosis (Henrotin, Bruckner y Pujol, 2003; Reed *et al.*, 2014; Bolduc, Collins y Loeser, 2019).

Por tanto, además de la vía inflamatoria que se ha estudiado en esta tesis doctoral, existen otros posibles mecanismos que podrían vincular esta variante mitocondrial con la progresión de la artrosis de rodilla. Si bien es cierto que muchos de los procesos mencionados están regulados parcialmente por una comunicación anterógrada del ADN nuclear sobre la mitocondria, es posible que la ocurrencia simultánea de todos estos procesos, junto con otros factores desconocidos, puedan incrementar dramáticamente la progresión de la enfermedad. Sin embargo, cabe destacar que este es el primer estudio funcional que analiza el efecto de esta variante, por lo que cuenta con ciertas limitaciones y serían necesarios estudios continuistas que pudieran profundizar más en la importancia funcional de la variante m.16519C.

### *Interacciones núcleo-mitocondria en el fenotipo de artrosis de rodilla con progresión rápida.*

En lo referente al ADN nuclear, se han identificado hasta la fecha un total de 176 SNPs que se asocian con la susceptibilidad de artrosis en distintas articulaciones (Boer *et al.*, 2021; Aubourg *et al.*, 2022). La mayoría de estos SNPs tienen un efecto pequeño, con unos valores generales de OR por debajo de 1,5. De hecho, en muchos estudios centrados en artrosis, el OR de los polimorfismos estudiados tiende a ser menor que el de variables clínicas como la edad y el IMC (Kulm *et al.*, 2022). Esto es debido a que en los análisis se suele incluir un número bajo de polimorfismos, es por ello que se clasifica a la artrosis como una enfermedad poligénica, donde la enfermedad ocurre debido a la herencia de múltiples alelos de riesgo con un efecto individual modesto (Aubourg *et al.*, 2022). Actualmente, el SNP que se ha asociado con mayor OR ha sido de 16,70, y fue descubierto en la artrosis de cadera. Se trata de una variante rara sin sentido que afecta al gen COMP (del inglés *encoding cartilage oligomeric Matrix protein*) y que consiste en una transversión que genera un cambio aminoacídico de asparagina a histidina en la posición 369 de la proteína COMP, que es un componente funcional importante en la MEC del cartílago (Styrkarsdottir *et al.*, 2017). Pero, cabe destacar que esta variante está presente en un *pedigree* de Islandia, estando ausente en el resto de poblaciones europeas, con lo que su impacto en la población en general quedaría en entredicho.

Por otro lado, con respecto a la progresión de la enfermedad, es importante destacar que el efecto de los SNPs nucleares, en comparación con la susceptibilidad, ha sido poco estudiada, y los pocos estudios que han sido publicados carecen del poder estadístico de los estudios GWAS. En 2013 se desarrolló un estudio para buscar genes candidatos para la progresión de artrosis de mano, y encontraron una asociación significativa entre el SNP rs13301537 del gen ASPN (del inglés *asporin*) y la progresión de artrosis de manos (Bijsterbosch *et al.*, 2013). Respecto a la progresión de la artrosis de rodilla, en 2015 el estudio Arthrotest generó una herramienta capaz de predecir la progresión severa de la artrosis de rodilla. Los autores definieron la progresión severa como el avance desde un grado KL 2-3 hasta 4 o cirugía de reemplazo articular en un período de ocho años. Dicha herramienta está basada en biomarcadores genéticos, y consta de un total de ocho SNPs, como variables genéticas, y la edad como variable clínica (Blanco *et al.*, 2015). En 2020, se generó un modelo predictivo de progresión de artrosis de rodilla en un período de dos años, que combinaba por primera vez la expresión de genes inflamatorios en sangre periférica (COX2, IL1 $\beta$  y TNF $\alpha$ ) y marcadores de imagen (lesiones de médula ósea); el modelo resultante además incluía también las variables clínicas tradicionales, la edad, el sexo y el IMC (Attur *et al.*, 2020). Más adelante, debido a la importancia de la estadística predictiva en esta tesis doctoral, en el apartado de la discusión centrado en modelos predictivos en la artrosis se profundizará más en el desarrollo de herramientas pronóstico para esta patología.

Para el desarrollo de esta tesis doctoral se decidió analizar el efecto, en la progresión rápida, de siete SNPs nucleares que se han asociado de forma muy robusta con la susceptibilidad a la artrosis de rodilla en distintos estudios GWAS (Warner y Valdes, 2017). Asimismo, debido a que la drástica reducción del mJSW es una característica de la artrosis con progresión rápida, también se analizó un octavo SNP relacionado con la reducción del espacio intraarticular (Zengini *et al.*, 2018). Si bien es cierto que estos SNPs se han asociado de forma muy robusta con la susceptibilidad y la disminución del mJSW, esto no descarta que se pudieran asociar también con la progresión rápida. Por otro lado, hay que tener en cuenta que, con el paso del tiempo, nuevos estudios han puesto de manifiesto nuevos SNPs que podrían ser útiles para predecir este fenotipo, y sería interesante su inclusión para proyectos futuros (Boer *et al.*, 2021; Aubourg *et al.*, 2022). Por ejemplo, en el año 2022 se llevó a cabo un estudio que incluyó 16.032 pacientes con artrosis con un elevado grado de severidad y que se habían sometido a una cirugía de reemplazo articular, y 373.516 controles. Con estos participantes se realizó un estudio GWAS con millones de variantes, ajustando los datos por edad, IMC, sexo y ser fumador. Como resultado, los investigadores encontraron un total de siete variantes asociadas significativamente con artrosis en estados muy avanzados de la enfermedad, que suele ser algo común dentro de la artrosis acelerada. Pero entre estas variantes no estaba presente ninguna de las empleadas en este estudio (Kulm *et al.*, 2022).

En base a lo descrito en la bibliografía, las consecuencias funcionales de cada uno de los SNPs incluidos en este estudio se detallan a continuación:

1. rs12107036 de TP63: el gen TP63 pertenece a la familia de factores de transcripción de p53, y la proteína que codifica, p63, cumple importantes funciones en la proliferación, la supervivencia y la apoptosis en diversas células y tejidos. Este gen también tiene un rol importante en la formación de la epidermis, las extremidades y en el desarrollo craneoencefálico (Kawata *et al.*, 2017). La presencia de mutaciones en este gen en seres

humanos se ha asociado de forma robusta con diversos desórdenes, destacando el síndrome ADULT (del inglés *Acro-dermo-ungual-lacrima-tooth*) y el síndrome *limb-mammary* (Sutton y van Bokhoven, 1993). Estudios previos en la artrosis han puesto de manifiesto que este gen puede afectar a la supervivencia del condrocito articular y por tanto condicionar el desarrollo y la progresión de la enfermedad. Con la ayuda de modelos animales en ratones se ha observado que este gen está sobrerrepresentado en condrocitos artrósicos, y la eliminación de este gen mediante estrategias de *knock out* produce la disminución de los niveles de apoptosis de los condrocitos en estos modelos (Taniguchi *et al.*, 2017).

Por otro lado, el gen podría modificar los niveles de autofagia de los condrocitos. En el estudio de Wangyang y colaboradores (Wangyang *et al.*, 2017), los autores analizaron los niveles de expresión de p63 en condrocitos de pacientes artrósicos y controles sanos, y de nuevo observaron que p63 presentaba mayores niveles de expresión en pacientes artrósicos. También realizaron diversas técnicas para analizar si la sobreexpresión de TP63 modificaba la viabilidad del condrocito y los niveles de autofagia de este tipo de célula y, como resultado, observaron que la sobreexpresión de esta proteína reducía la viabilidad de los condrocitos a 24, 36 y 48 horas de tratamiento. Además, también se reducían los niveles de autofagia de los condrocitos, pudiendo comprometer la supervivencia celular por una nueva vía a parte de la apoptótica ya mencionada. El papel del gen TP63 tanto en el proceso de autofagia, así como en la apoptosis, se ha demostrado en un trabajo reciente de Wang y colaboradores (Yichao Wang *et al.*, 2022), en el que demostraron que el factor TGF- $\beta$ 1 promueve la autofagia e inhibe la apoptosis de células de cáncer de mama al inhibir la expresión de TP63.

El SNP rs12107036 (A/G) está presente en el intrón 12 del gen TP63. Este polimorfismo ya fue empleado en un modelo predictivo para identificar progresores estructurales de artrosis de rodilla usando metodologías computacionales de *machine learning* (Bonakdari *et al.*, 2022). El alelo G se ha asociado con cirugía de reemplazo articular en mujeres (Zeggini *et al.*, 2012) y también se asocia con la enfermedad de Kashin-Beck, la cual en individuos adultos presenta unas manifestaciones clínicas muy similares a la artrosis (Cheng *et al.*, 2021). En este último estudio, los autores también realizaron un análisis de expresión en el cartílago articular, observando, al igual que en el estudio de Taniguchi y colaboradores, que la proteína p63 presentaba una gran abundancia en el cartílago artrósico. Los autores de este trabajo también realizaron estudios de supresión del gen TP63 y, pudieron ver que, al reprimir dicho gen, algunos marcadores de apoptosis como la caspasa 3 presentaban una menor abundancia, lo cual parece indicar que TP63 puede tener un rol en la apoptosis del condrocito, y que la sobreexpresión de éste gen puede reducir la supervivencia de este tipo celular.

2. rs8044769 del gen FTO: diversos estudios han puesto de manifiesto que la presencia de polimorfismos en este gen se asocia con mayores niveles de IMC, y aunque el efecto tiende a ser pequeño, algunos de estos polimorfismos, como por ejemplo tener dos copias del alelo de riesgo A del SNP rs9939609, se asocian con un incremento medio de 3 kg comparado con individuos heterocitogicos, y se estimó que cada copia del alelo A incrementa de media el IMC en aproximadamente 0,4 kg/m<sup>2</sup> (Frayling *et al.*, 2007; Fawcett y Barroso, 2010). En el año 2017, mediante estrategias de meta-análisis, también se encontró otro SNP en este gen, el rs1558902, que se asoció de forma robusta con la artrosis de rodilla (Yau *et al.*, 2017). Sin embargo, estas asociaciones no siempre se replican, y por ejemplo en el estudio de Dai y colaboradores (Dai *et al.*, 2018) las

asociaciones entre rs8044769 y el IMC, así como entre el mismo polimorfismo y la artrosis, no fueron replicadas en una población China. Esto produjo cierta controversia y se empezó a especular sobre que quizá este SNP no incrementaba el riesgo de artrosis a través de la obesidad. Para testar esta hipótesis, en 2021 Takaoka y colaboradores (Takaoka *et al.*, 2022) analizaron en una población japonesa el efecto del polimorfismo rs8044769 en el riesgo de desarrollar artrosis en la articulación temporomandibular, la cual está muy poco afectada por el IMC. Como conclusión de este estudio, el tener dos copias del alelo de riesgo C del polimorfismo rs8044769 se asociaba de forma significativa con el riesgo de desarrollar artrosis en esta articulación, independientemente de la obesidad. Sería necesario en el futuro continuar con esta línea de investigación con el fin de esclarecer si realmente existe una interacción entre este polimorfismo y el IMC pero, independientemente de esta asociación, lo que está claro es que la presencia del polimorfismo es capaz de modificar el riesgo de desarrollar artrosis.

3. rs11177 del gen GNL3: la proteína para la que codifica este gen se expresa en células madre mesenquimales, de las cuales derivan los condrocitos del cartílago articular. Esta proteína juega un rol importante en la regulación del ciclo celular y en el mantenimiento de la actividad de las enzimas telomerasas, que tienen a su vez un papel destacado en la proliferación de las células madre y en la regulación del ciclo celular (Lee *et al.*, 2015). El alelo T del SNP intrónico rs11177 (C/T) presente en este gen, se ha asociado en poblaciones caucásicas y chinas con el riesgo de desarrollar artrosis de rodilla (Zeggini *et al.*, 2012; Liu *et al.*, 2018), artrosis de cadera en poblaciones caucásicas (Zeggini *et al.*, 2012) y artrosis de mano en poblaciones asiáticas (X. Wang *et al.*, 2022). Una posible explicación del efecto de este alelo puede ser que, si bien es una variante intrónica y esta no modifica la estructura de la proteína, este polimorfismo puede funcionar como un eQTL (del inglés *expression quantitative trait loci*), modificando la expresión de otros genes cercanos relevantes para el desarrollo de la artrosis, como por ejemplo NT5DC2 (del inglés *5'-Nucleotidase Domain Containing 2*), el cual se vinculó previamente en otros estudios GWAS con el riesgo de desarrollar artrosis de rodilla (X. Wang *et al.*, 2022).
4. rs4730250 del gen DUS4L: El gen DUS4L codifica para una proteína que cataliza la síntesis de dihidrouridinas, unas bases modificadas que se pueden encontrar en moléculas de ARNt y ARNr. Este polimorfismo está presente en un bloque de desequilibrio de ligamiento junto con otros cinco. La primera vez que esta región se asoció con la artrosis fue en el año 2010, en el estudio de Kerkhof y colaboradores (Kerkhof *et al.*, 2010), donde encontraron una fuerte asociación entre un polimorfismo presente en la región 7q22 con la artrosis. Fue en el año 2011 cuando, por primera vez, Evangelou y colaboradores (Evangelou *et al.*, 2011) mostraron la asociación del SNP rs4730250 de DUS4L con la artrosis de rodilla. Una posible explicación de cómo este gen puede afectar al desarrollo de la artrosis de rodilla viene de la mano de Raine y colaboradores (Raine *et al.*, 2012), los cuales en su estudio observaron que los portadores del polimorfismo rs4730250 de DUS4L presentaron una reducción de la expresión de este gen en la grasa intrapatelar de la rodilla. La grasa intrapatelar juega un papel importante en la patogénesis de la artrosis de rodilla, ya que puede secretar mediadores inflamatorios como TNF $\alpha$  e IL6, incrementando así el proceso artrósico. Es posible que la reducción de la expresión de este gen pueda tener efectos en la grasa intrapatelar y por tanto en el proceso artrósico, pero serían necesarios más estudios con el fin de aclarar de manera más definitiva, el efecto de este polimorfismo.

5. rs143383 del gen GDF5: este gen codifica para una proteína con funciones de morfogénesis ósea, y tiene un papel fundamental en el desarrollo, el mantenimiento y la reparación del cartílago articular y del hueso de las articulaciones. El alelo C se ha asociado con un efecto protector en poblaciones caucásicas, asiáticas, y africanas (Jia *et al.*, 2021); mientras que el alelo T del polimorfismo se ha asociado con el riesgo de artrosis en poblaciones europeas (Yau *et al.*, 2017), americanas (García-Alvarado *et al.*, 2018) y asiáticas (Tawonsawatruk *et al.*, 2011). El motivo por el que el alelo T se asocia con la artrosis es debido a que reduce los niveles de expresión de este gen (Valdes *et al.*, 2011) y, por tanto, las funciones de esta proteína pueden comprometerse, provocando que se favorezca el proceso artrósico. Cabe destacar también que los patrones de metilación que sufra este gen en su región promotora pueden alterar el desequilibrio de expresión producido por el SNP. Este proceso ocurre en una isla CpG situada 4 pb antes del SNP rs143383; así, cuando esta isla está desmetilada, aumenta la expresión del gen, siendo mayor con el alelo C en comparación con el T, favoreciendo el desequilibrio de expresión alélico, sin embargo, cuando esta isla está metilada, se reprime mucho más la expresión del gen, produciendo que los niveles de expresión con el alelo C y el T sean más similares, reduciéndose así el desequilibrio (Reynard *et al.*, 2014).
6. rs10948172 del gen SUPT3H: este gen está en una región con un alto desequilibrio de ligamiento, concretamente en el brazo corto del cromosoma 6, región 2 banda 1 subbanda 1 (6p21.1). En este bloque de ligamiento, junto con SUPT3H también está presente el gen RUNX2 (del inglés *RUNX Family Transcription Factor 2*), que codifica para un factor de transcripción esencial para la diferenciación de los osteoblastos y el desarrollo esquelético. El motivo por el que se vincula al alelo G de este polimorfismo con un aumento en el riesgo de desarrollar artrosis es debido a que la presencia de este SNP puede alterar los niveles de metilación de los elementos reguladores de RUNX2, pudiendo comprometer sus funciones en el correcto desarrollo y mantenimiento de la articulación (Rice *et al.*, 2018)
7. rs11842874 del gen MCF2L: en células humanas, el gen MCF2L regula la migración celular de las células de Schawnn inducida por neurotrofina-3, la cual es un miembro de la familia del NGF y un factor asociado al dolor en la artrosis, siendo este el motivo por el que está asociado a esta enfermedad (Day-Williams *et al.*, 2011; Panoutsopoulou y Zeggini, 2013). El producto resultante de este gen actúa como una proteína señalizadora importante en el proceso de desarrollo e hipertrofia de los condrocitos, y también puede activar proteínas relacionadas con las propiedades invasivas de los sinoviocitos fibroblásticos (Wang y Beier, 2005).  
Este SNP rs11842874 (A/G) está presente en el intrón número cuatro del gen MCF2L, y es el alelo A el que se ha asociado con el riesgo de desarrollar artrosis. Esto es debido a que este alelo aumenta la expresión de MCF2L, pudiendo alterar el correcto desarrollo de los condrocitos articulares. Además, este polimorfismo también puede actuar como un eQTL, modificando la expresión de otros genes cercanos (Shepherd *et al.*, 2015)
8. rs3771501 del gen TGFA: este gen codifica para el receptor del factor de crecimiento epidermal, que tiene una función importante en la señalización y en la función celular. Estudios funcionales han puesto de manifiesto que este gen regula la conversión de cartílago hacia hueso durante el proceso de crecimiento del hueso endocondral (Usmani *et al.*, 2012) y también juega un rol importante en el desarrollo craneofacial y cierre palatino (Miettinen *et al.*, 1999). El alelo T del SNP rs3771501 (C/T) se ha asociado con



una reducción del mJSW (Zengini *et al.*, 2018), pero actualmente no existen estudios funcionales que evalúen como produce este efecto dicho polimorfismo.

En esta tesis se realizó el genotipado de las ocho posiciones nucleares de interés mediante tecnologías de SBE (realizando un control de calidad en el 10% de las muestras) en las 1.095 muestras que abarca el criterio criterio GIR-blanco. En los análisis univariantes con el criterio GIR de dos y tres categorías, ninguno de los alelos de riesgo de los SNPs de interés mostró resultados significativos de forma individual, siendo el alelo de riesgo más próximo a la significación estadística el G del SNP rs12107036 de TP63. Los progresores rápidos mostraban, tanto en el criterio de dos como en el de tres grupos de progresión, una mayor frecuencia de este alelo frente a los otros grupos, sin embargo, en ambos casos esta diferencia rozaba la significación estadística, alcanzando un p-valor de 0,141 en el criterio GIR-blanco de tres grupos de progresión, y un p-valor de 0,053 en el de dos.

Con respecto al ADNmt, esta molécula es fundamental para que la mitocondria pueda llevar a cabo las funciones imprescindibles para la supervivencia del organismo (Murphy *et al.*, 2016). El ADNmt presenta una tasa de mutación mayor que el ADN nuclear debido a la elevada producción de ROS mitocondriales, también a que se replica constantemente y a que tiene unos mecanismos de reparación más pobres que el ADN nuclear (Lechuga-Vieco *et al.*, 2021). Además, debido a esta elevada tasa de mutación y al proceso de adaptación al clima que llevó a cabo la especie humana tras la migración desde África al resto del mundo, diversas mutaciones se han fijado en distintas poblaciones estableciendo conjuntos de polimorfismos estables denominados haplogrupos mitocondriales, los cuales, además de las mutaciones individuales, también condicionan el comportamiento de la mitocondria y modifican el riesgo de desarrollar patologías (Wallace, 2015).

Previamente a este estudio se había observado que en la patología artrósica la presencia de distintos haplogrupos condiciona la prevalencia, incidencia y progresión de esta patología (Fang *et al.*, 2014; Shen *et al.*, 2014; Blanco *et al.*, 2018). Con respecto a la progresión de la artrosis, tal y como se comentó en la introducción de esta tesis doctoral, tanto el haplogrupo T como el cluster JT se asocian con una menor tasa de progresión en la artrosis de rodilla (Fernández-Moreno *et al.*, 2017; Zhao *et al.*, 2020). Además, los individuos portadores del haplogrupo H, presentan mayor riesgo de necesitar cirugía de reemplazo articular de rodilla (Soto-Hermida *et al.*, 2015). Sin embargo, en nuestro estudio, el análisis de la distribución de los clusters de haplogrupos mitocondriales entre los grupos de progresión del criterio GIR-blanco, no mostró ninguna asociación significativa individual entre estos clusters y el fenotipo de artrosis de rodilla con progresión rápida, con unos valores de frecuencia de los clusters de haplogrupos mitocondriales muy similar entre los grupos de progresión.

Hasta este punto del trabajo, ningún SNP ni cluster de haplogrupos mitocondriales ha mostrado ninguna asociación individual con el fenotipo de artrosis de rodilla con progresión rápida, pudiendo indicar que el efecto individual de estos polimorfismos es incapaz de generar cambios lo suficientemente potentes como para inducir una progresión rápida. Por ello, se decidió llevar a cabo un estudio de interacciones entre SNPs nucleares y clusters de haplogrupos mitocondriales. Debido a la gran dependencia y comunicación que se lleva a cabo entre estos dos elementos (ya sea de forma retrógrada o anterógrada), la coexistencia de polimorfismos en ambos genomas puede conllevar cambios en el perfil de expresión celular, promover la disfunción mitocondrial y, como consecuencia, desarrollarse distintos fenotipos patológicos,

## Discusión

como por ejemplo el Alzheimer (Andrews *et al.*, 2020), la esclerosis múltiple (Kozin *et al.*, 2018), el cáncer de mama (Bushel *et al.*, 2022), el Parkinson (Pickett *et al.*, 2022), la degeneración macular asociada al envejecimiento (Persad *et al.*, 2017), o incluso elevados niveles de heteroplasmia en el ADNmt (Nandakumar *et al.*, 2021).

Los análisis de interacciones entre cluster mitocondriales y los ocho SNPs nucleares revelaron una interacción significativa entre el alelo de riesgo G de rs12107036 de TP63 y el cluster de haplogrupos UK. Dicha interacción se mantuvo significativa incluso tras el ajuste debido a las comparaciones múltiples, y ser portador de ambas variables genéticas incrementaba el riesgo de progresar rápidamente. Además de esta asociación, también se pudo observar otra interacción, entre el alelo de riesgo T de rs11177 de GNL3 con el cluster de haplogrupos UK, sin embargo, tras ajustar el p-valor, esta perdía la significación estadística, y tal vez, para estudios futuros sería deseable explorar su posible impacto en el fenotipo estudiándola en cohortes prospectivas con un mayor tamaño muestral.

El efecto de la interacción entre el cluster UK y el alelo G del SNP rs12107036 de TP63 se detectó en la fase exploratoria en la cohorte OAI. A la hora de validar dicha interacción se descartó la utilización de las cohortes externas PROCOAC y CHECK debido al tamaño muestral, ya que algunas combinaciones presentaban unos tamaños muestrales muy bajos. Por tanto, se decidió llevar a cabo, de manera exitosa, un enfoque de validación cruzada interna usando técnicas de *bootstrap*. Con esta técnica, los 1.095 individuos de la cohorte OAI se subdividieron de forma aleatoria y equitativa en diez subgrupos. Se realizaron diez análisis usando estos subgrupos donde un grupo se usaba como *test* (validación) mientras que los otros nueve grupos se empleaban para realizar el *training* del modelo. Tras realizar el *training*, el grupo *test* se emplea para evaluar el modelo. Este paso se repitió un total de diez veces donde en cada análisis se empleó un subgrupo de *test* distinto.

Ninguna de estas dos variantes genéticas (rs12107036 y cluster UK) había mostrado ningún tipo de asociación individual con el fenotipo de artrosis de rodilla con progresión rápida. Sin embargo, la coexistencia de ambos genotipos incrementa el riesgo de desarrollar este fenotipo. De hecho, el análisis de PAFs mostraba una mayor proporción de casos de progresión rápida atribuibles a presentar la interacción (23,4%) en comparación con individuos con el SNP nuclear pero otros clusters de haplogrupos (8,05%). Como se mencionó previamente, el gen TP63 cumple importantes funciones en la proliferación, la supervivencia y la apoptosis en diversas células y tejidos (Taniguchi *et al.*, 2017), y la presencia del alelo G puede afectar a los niveles de expresión del gen, pudiendo llegar a modificar los niveles de autofagia de los condrocitos (Wangyang *et al.*, 2017). Teniendo en cuenta esto, es posible que la presencia de este polimorfismo pueda producir defectos en el cartílago que, a largo plazo, puedan derivar en el desarrollo y/o avance de la patología artrósica.

Por otro lado, el cluster de haplogrupos UK se han vinculado con unos niveles de severidad de artrosis mayores, en términos de grado KL (Rego-Pérez *et al.*, 2008), lo cual es el estado final que suelen alcanzar los progresores rápidos. En un estudio reciente, un grupo de investigadores trataron de analizar cómo las variantes mitocondriales más frecuentes en la población eran capaces de condicionar el desarrollo de patologías asociadas al envejecimiento (Cai *et al.*, 2021). Para ello combinaron datos genéticos con datos metabólicos, proteómicos y rasgos hematológicos de 16.620 individuos. Encontraron un total de quince variantes en el ADNmt que condicionaban los niveles de N-formilmetionina (fMet), catorce de las cuales eran

variantes que definían el cluster mitocondrial UK. La fMet es un aminoácido fundamental en la síntesis de proteínas en las bacterias y, debido a su origen endosimbionte, es necesario para iniciar la traducción de proteínas mitocondriales. Los autores también realizaron un estudio funcional con híbridos H y UK, que tenía como objetivo ver cómo estas variantes afectaban a los niveles de fMet. Pudieron ver que los híbridos UK generaban mayores niveles de fMet, lo cual, a pesar de parecer contradictorio, podía acarrear una reducción en la síntesis de proteínas mitocondriales (entre ellas las pertenecientes a los complejos I y IV de la CTE); asimismo, también condicionaba la síntesis y degradación de proteínas citosólicas, afectando por tanto a la funcionalidad de la célula. Finalmente, analizaron cómo los niveles de fMet podían condicionar el desarrollo de patologías en una cohorte de 11.966 participantes durante 20 años de seguimiento, y vincularon estos altos niveles del metabolito con enfermedad renal, fallo cardíaco, la enfermedad de las arterias coronarias, aneurismas aórticas abdominales, enfermedad de las arterias periféricas y mortalidad.

Si además tenemos en cuenta que la mitocondria y el núcleo se comunican constantemente para mantener la funcionalidad y los requerimientos energéticos celulares (De Benedictis *et al.*, 2000; Quirós *et al.*, 2016; Mello *et al.*, 2019), se abre una nueva posibilidad, donde el efecto sinérgico de ambos polimorfismos, los cuales de forma individual se ha comentado que pueden tener un efecto moderado en la progresión rápida, se dispare, produciendo que los individuos portadores de ambas variantes genéticas experimenten un rápido y acusado deterioro de la articulación de la rodilla debido a la artrosis. El por qué la presencia de ambas condiciones genéticas puede disparar la progresión de la artrosis puede residir en que una misma mutación, presente en ambientes genéticos distintos (ya sean mitocondriales y/o nucleares), puede producir diferentes efectos fenotípicos (Bénil *et al.*, 2010). Recientemente se demostró en un estudio que analizaba el *pedigree* de una familia, que nuevas variantes, que de forma espontánea aparecen en el ADNmt, podían mostrar incompatibilidades con las variantes presentes previamente en la propia molécula de ADNmt. Concretamente, la incompatibilidad detectada estaba presente entre la variante m.13708G>A y el haplogrupo H7, que resultaba en una disfunción mitocondrial debida a una regulación anormal de los niveles de calcio, produciendo un gran abanico de fenotipos clínicos en la descendencia, incluyendo mortalidad en el útero. Además, en este mismo estudio, observaron que la presencia de variantes en el ADN nuclear (concretamente la presencia de la variante W317X en el gen histidina descarboxilasa), podía compensar el efecto de estas variantes deletéreas, protegiendo a los portadores de las incompatibilidades genéticas mitocondriales del riesgo de muerte (Schaefer *et al.*, 2022).

Por tanto, puede que la coexistencia del alelo G del SNP nuclear rs12107036 de TP63 junto con el haplogrupo UK amplifique de algún modo el efecto deletéreo de estos polimorfismos, lo que no se observaría al combinar el SNP nuclear con cualquier otro haplogrupo mitocondrial o viceversa. Sin embargo, son necesarios más estudios para tratar de buscar un significado funcional a esta prometedora interacción núcleo-mitocondria, que es la primera que se ha detectado hasta la fecha en el campo de la artrosis.

#### *Modelos predictivos en la artrosis.*

Con todos los datos obtenidos a lo largo de esta tesis doctoral se generó un modelo predictivo para el fenotipo de artrosis de rodilla con progresión rápida. Teniendo en cuenta esto, el estudio aquí descrito podría englobarse dentro de la investigación predictiva. Este tipo de estudios pueden centrarse en el diagnóstico de un evento o patología (probabilidad de que un evento esté presente pero no haya sido detectado) o en su pronóstico (probabilidad de un

## Discusión

evento de ocurrir en el futuro). Los estudios en este campo pueden clasificarse del siguiente modo (Wolff *et al.*, 2019):

- Estudios de predictores: buscan identificar variables que permitan predecir un resultado. Estas variables predictoras pueden servir tanto como factores de riesgo o factores de pronóstico para el desarrollo de modelos predictivos.
- Estudios de modelos predictivos: buscan desarrollar, validar y/o actualizar un modelo predictivo multivariable. En estos estudios se desarrollan herramientas que cuentan con múltiples variables predictoras que son capaces de alterar un evento o patología. Con las herramientas obtenidas en este tipo de trabajos se puede predecir la probabilidad de un individuo de tener una patología o evento (en caso de ser diagnóstico) o la probabilidad de desarrollarla en el futuro (en caso de ser pronóstico).
- Estudios del impacto del modelo predictivo: comparan el efecto de usar un modelo predictivo existente para guiar el cuidado del paciente con respecto a no usar dicha herramienta.

Esta tesis, por tanto, podría clasificarse como un estudio de modelo predictivo. Las herramientas resultantes de estos trabajos, los modelos predictivos, representan matemáticamente procesos biológicos, los cuales evalúan distintos escenarios con diversas características predictoras con el fin de determinar la probabilidad de que ocurra un suceso, lo que permite tener tiempo de respuesta y gestionar el resultado. Para generar los modelos predictivos de variables binarias y categóricas existen numerosas metodologías, siendo las más comunes la regresión logística, los árboles de clasificación y regresión, las redes neuronales o las máquinas vectoriales, entre otras (Fenlon *et al.*, 2018). El modelo predictivo desarrollado en esta tesis doctoral se ha generado mediante regresión logística, y funciona como una herramienta de pronóstico del fenotipo de artrosis de rodilla con progresión rápida.

Los modelos predictivos se han usado en multitud de patologías. Entre ellas están presentes el cáncer de mama (Altman, 2009), el cáncer de próstata (Shariat *et al.*, 2008), el derrame cerebral (Counsell y Dennis, 2001), la muerte por traumatismo hemorrágico (Perel *et al.*, 2012) o el desarrollo de enfermedades cardiovasculares, entre otras (Hippisley-Cox *et al.*, 2008; Damen *et al.*, 2016). En el campo de la artrosis también se han desarrollado trabajos que se centran en la predicción. En las últimas décadas se han identificado y confirmado numerosos factores de riesgo para el desarrollo de la artrosis de rodilla, siendo los empleados de forma más habitual la edad, el sexo y el IMC, aunque se podrían destacar muchos otros como la educación, la ocupación laboral, el estilo de vida, el entorno, la nutrición, la actividad física, la pérdida de peso, la lesión previa, el historial familiar, la circunferencia de la cintura, las enfermedades articulares concomitantes, el dolor en la rodilla, la rigidez en la rodilla, el tratamiento farmacológico, distintos biomarcadores en suero, plasma y orina, datos radiográficos, datos cuantitativos de resonancia magnética, biomarcadores genéticos (Warner y Valdes, 2017), marcadores basados en sitios CpG metilados (Papathanasiou *et al.*, 2019) e incluso microARNs circulantes (como él miR-320 (Ali *et al.*, 2022)). En esta tesis doctoral concretamente, para desarrollar el modelo predictivo, se incluyeron como factores de riesgo la edad, el sexo, el IMC, la lesión previa, la contralateralidad, el dolor WOMAC y distintas variables genéticas tanto nucleares como mitocondriales. Por tanto, el modelo predictivo generado en esta tesis es el primero que incluye variables genéticas para predecir el desarrollo del fenotipo de artrosis de rodilla con progresión rápida.

Dentro de las variables cuantitativas introducidas en el modelo predictivo para la artrosis de rodilla con progresión rápida, la comparación de las variables demográficas entre los

grupos de progresión del criterio GIR-blanco dentro de la cohorte OAI indicaba que los individuos progresores rápidos presentan edades significativamente mayores con respecto a los otros dos grupos de progresión. La edad es un factor de riesgo para numerosas enfermedades asociadas al envejecimiento, entre ellas la artrosis. El envejecimiento es un proceso complejo donde se acumulan daños moleculares, celulares y orgánicos, que llevan, en el caso de la artrosis, a la pérdida de función de la articulación, incrementando la vulnerabilidad de esta estructura frente a la progresión de esta enfermedad (Valdes y Stocks, 2018). Inicialmente se pensaba que la relación entre estos dos elementos se debía única y exclusivamente a un proceso de desgaste entre las superficies articulares debido al estrés mecánico y al paso del tiempo, pero actualmente se sabe que otros elementos pueden jugar un papel importante en esta asociación; por ejemplo, con el envejecimiento, el sistema musculoesquelético se debilita, además también hay una pérdida de propiocepción, lo cual puede promover que se sufran lesiones en las articulaciones que pueden condicionar el desarrollo de la patología artrósica (Shane Anderson y Loeser, 2010). Otra posible explicación de por qué puede estar fomentada esta relación entre artrosis y envejecimiento es a través de la mitocondria. Este orgánulo tiene numerosas funciones y, con el paso del tiempo, tiende a acumular diversas mutaciones somáticas que puede producir que se vuelvan mitocondrias disfuncionales, afectando a múltiples procesos como una disminución de la producción de ATP, un aumento del estrés oxidativo, la desregulación del calcio, un aumento de la permeabilidad de la membrana mitocondrial y alteraciones del ADNmt, que comprometen la supervivencia del condrocito y pueden culminar en la degradación del cartílago articular (Mao *et al.*, 2020).

En lo referente al IMC y al índice de dolor WOMAC, los individuos progresores rápidos y progresores no rápidos mostraban en la cohorte OAI unos valores significativamente mayores que los individuos no progresores, sin embargo, no se encuentran diferencias significativas entre estos dos grupos de progresores, y es que ambos factores son causa de riesgo de prevalencia de artrosis. En primer lugar, un IMC mayor de 30kg/m<sup>2</sup> se ha asociado de forma robusta con la incidencia de artrosis (Silverwood *et al.*, 2015). La obesidad promueve la artrosis de rodilla y cadera mediante un potente efecto biomecánico debido al exceso de carga, produciendo daños en la articulación (Felson *et al.*, 2000) y también genera cambios metabólicos que favorezcan el proceso inflamatorio de la articulación (Kulkarni *et al.*, 2016; Misra *et al.*, 2019). Además, la artrosis debida a obesidad puede ir acompañada de sarcopenia producida por un dolor incapacitante. Esto produce un ciclo vicioso de inflamación, inactividad, pérdida de musculatura y ganancia de tejido adiposo, lo cual favorece el fenotipo obeso y la prevalencia y la progresión de la patología (Godziuk *et al.*, 2018). También existe una relación sinérgica entre la obesidad y la depresión a efectos de la progresión de la artrosis de rodilla, los pacientes obesos con artrosis y depresión desarrollan mayores niveles de dolor, de pérdida de funcionalidad articular, de marcadores de deterioro articular y de remodelación ósea con respecto a pacientes con artrosis y obesidad pero sin depresión (Jacobs *et al.*, 2020). Tanto la obesidad como la edad son dos de los factores más importantes en el desarrollo de la artrosis acelerada, destacando particularmente dos grupos de individuos, aquellos con una edad superior a 65 años y un IMC con sobrepeso u obesidad (menor de 35 kg/m<sup>2</sup>) y, aquellos con edades menores de 65 pero con un IMC mayor de 32,5 kg/m<sup>2</sup> (Driban, Eaton, *et al.*, 2016). En segundo lugar, el WOMAC consiste en un índice que permite evaluar la actividad de la enfermedad de la artrosis. Se centra en tres aspectos fundamentalmente, dolor, rigidez y capacidad funcional (Bellamy *et al.*, 1988). Este indicador también ha sido empleado en el desarrollo de modelos predictivos para el fenotipo de artrosis de rodilla acelerada (Riddle *et al.*, 2016). El principal motivo es debido a que, doce meses

## Discusión

antes de desarrollar la artrosis de rodilla acelerada, tienden a sufrir un intenso dolor en la rodilla, llegando incluso a necesitar medicación; con lo que este parámetro podría servir como un posible indicador del fenotipo acelerado (Driban *et al.*, 2020).

Al analizar las variables cualitativas en la cohorte OAI, se puede observar que tanto el sexo, como la artrosis contralateral y la lesión previa muestran diferencias significativas entre los tres grupos de progresión. Los individuos no progresores tienen un mayor porcentaje de individuos varones con respecto a los otros dos grupos (el 44,2% son varones), en los cuales es más predominante el sexo femenino (55,8%). El sexo femenino se ha asociado con mayores niveles de prevalencia y severidad en la artrosis especialmente después de la menopausia (Srikanth *et al.*, 2005); se piensa que esto puede ser debido al papel de los estrógenos en el metabolismo del cartílago. Menores niveles de estrógenos contribuyen a la progresión de la enfermedad debido a que se incrementa la pérdida de masa muscular, afectando a las funciones musculares que pueden resultar en inestabilidad articular, desequilibrios en el reparto de carga mecánica en la articulación y finalmente a daños en el cartílago. Además, la deficiencia en los niveles de estrógenos también afecta a la pérdida de hueso (Peshkova *et al.*, 2022). Aunque hay evidencia de que la progresión de esta patología se suele asociar con una mayor densidad ósea (Hardcastle *et al.*, 2015), una menor densidad ósea, contrariamente, suele favorecer su desarrollo (Burr y Gallant, 2012). Además, en línea con esto último, una menor densidad ósea puede favorecer el desarrollo de fracturas por insuficiencia subcondral, las cuales son un tipo de lesiones atraumáticas del hueso subcondral que se han asociado con la artrosis de rodilla con progresión rápida (Yamamoto y Bullough, 2000; Gourlay *et al.*, 2015). Por otro lado, a diferencia de lo que ocurre en el sexo masculino, en el caso de las mujeres puede condicionar al desarrollo de esta patología la existencia de diferencias en la fuerza ósea, el alineamiento articular, la fuerza neuromuscular o las consecuencias derivadas del embarazo (Johnson y Hunter, 2014).

La contralateralidad en la cohorte OAI es mucho más frecuente en los individuos progresores rápidos. La contralateralidad se define como la presencia de artrosis en una articulación en un lado del cuerpo, que puede condicionar el desarrollo de artrosis en la articulación sana opuesta en el otro lado del cuerpo. La contralateralidad está considerada como un factor de riesgo que puede afectar al desarrollo de la artrosis de rodilla acelerada en un período de dos a cuatro años previo al establecimiento de la artrosis (Driban *et al.*, 2020). En consecuencia, este parámetro fue incluido en modelos predictivos para la identificación de este fenotipo (Riddle *et al.*, 2016). Estudios recientes han puesto de manifiesto el papel que tiene la contralateralidad en la progresión de la artrosis de rodilla y en el aumento de la probabilidad de aplicar cirugías de reemplazo articular. Así, en el año 2022 se llevó a cabo un estudio con pacientes que presentaban una rodilla artrósica y una rodilla con artrosis temprana. El objetivo consistió en ver si la progresión en la rodilla menos dañada se podía ver afectada por la presencia de distintos niveles de severidad de la rodilla artrósica. Para ello trabajaron con dos grupos de pacientes, uno con rodillas con grados KL 3 y 4 (N=449), y otro con grados KL 0, 1 y 2 (N=1.276). Como resultado, pudieron observar unas tasas de progresión significativamente mayores dentro del grupo con mayor severidad en la rodilla contralateral en períodos de uno y dos años. Además el riesgo de necesitar una cirugía articular también era significativamente mayor en el grupo con mayor severidad (Wu *et al.*, 2022).

Del mismo modo que ocurre con la contralateralidad, en la cohorte OAI los progresores rápidos presentan una mayor frecuencia de lesión previa en su rodilla diana. Las lesiones previas que afectan a estructuras de soporte tienden a producir daños en el cartílago articular y el hueso subcondral (Von Porat *et al.*, 2004). Por ejemplo, el daño en los meniscos genera un mal reparto de las cargas que acaba favoreciendo la progresión de la artrosis de rodilla (Walker *et al.*, 2011). En el año 2014 se realizó un estudio para evaluar cómo la lesión previa podía condicionar el desarrollo de la artrosis de rodilla acelerada (Driban *et al.*, 2014). Clasificaron a individuos con un grado KL basal 0 o 1 en progresores con artrosis acelerada, progresores normales y no progresores. Como resultado encontraron una fuerte asociación entre la lesión previa y el desarrollo del fenotipo de artrosis acelerado, además la lesión reciente (hace menos de un año) se asociaba de forma más robusta con la progresión acelerada e incluso con la progresión normal de la artrosis de rodilla. Además, como se ha comentado, la lesión previa en la articulación puede ser debido a un proceso de envejecimiento u obesidad que desembocan en una sarcopenia e incluso en una pérdida severa de la propiocepción, lo cual favorece las caídas y el desarrollo de lesiones. La lesión previa reciente (uno o dos años antes de desarrollar la artrosis) facilita el desarrollo de artrosis acelerada, especialmente en adultos con edades avanzadas (65 o más años) e individuos con sobrepeso (Driban, Eaton, *et al.*, 2016).

Finalmente se introdujeron distintas variables genéticas nucleares y mitocondriales debido a la importancia que tienen estos elementos en la susceptibilidad, incidencia y progresión de la artrosis de rodilla (Blanco *et al.*, 2015; Warner y Valdes, 2017; Blanco *et al.*, 2018). Si bien es cierto que el uso de estos biomarcadores en el campo de la artrosis no es algo nuevo, esta tesis doctoral presenta dos aspectos pioneros; en primer lugar, el uso de metodologías de secuenciación masiva con el fin de identificar variantes mitocondriales que se asocien con el fenotipo rápido de artrosis de rodilla y, en segundo lugar, la búsqueda de interacciones genéticas núcleo-mitocondria que puedan condicionar dicho fenotipo.

#### *Desarrollo de modelos predictivos convencionales para la artrosis.*

El desarrollo de modelos predictivos para la artrosis ha cobrado gran importancia en la actualidad, ya que el diagnóstico de esta patología ocurre principalmente durante la etapa moderada a grave de la enfermedad, momento en el cual el tejido articular a menudo ya se ha dañado de manera irreversible. Además, el progreso en el desarrollo de medicamentos para la artrosis o terapias que detienen o reducen la progresión del deterioro del tejido articular es lento y, cabe destacar que, actualmente, a pesar de los numerosos ensayos clínicos que están en marcha, las agencias reguladoras aún no han aprobado ningún tratamiento del grupo de los DMOADs, capaces de alterar el curso de la enfermedad modificando su prevenir y/o su progresión. Los medicamentos actuales son efectivos solo para aliviar los síntomas y/o el dolor, y su uso podría inducir eventos adversos importantes e incluso la mortalidad (Solomon *et al.*, 2010). Además de la ausencia de un tratamiento eficaz, existen otros obstáculos importantes que se deben superar en el manejo de la artrosis; en primer lugar estaría la identificación de los pacientes que podrían beneficiarse más de ciertos medicamentos o medidas preventivas, incluidos los que se encuentran en una etapa temprana de la enfermedad o aquellos para los que la enfermedad no es demasiado grave, pero puede progresar rápidamente; y en segundo lugar, debido a la heterogeneidad y los numerosos fenotipos de la artrosis, la administración de una terapia efectiva para el paciente es una tarea muy compleja. Es en este segundo aspecto donde actualmente la medicina personalizada se ha enfocado enormemente, tratando de buscar las peculiaridades únicas de cada individuo con el fin de identificar la terapia que mejor se adapte a esa persona, evitando tratamientos generalizados que pueden generar efectos

## Discusión

secundarios y no mejorar la sintomatología del sujeto. Pero, a día de hoy, los procedimientos de diagnóstico clínico actuales continúan sin satisfacer adecuadamente la necesidad de los médicos de ayudar a los pacientes a reducir el riesgo de progresión de la enfermedad o la necesidad de la industria de la salud de desarrollar nuevos fármacos efectivos.

He aquí donde radica la importancia de los modelos de predicción, ya que permitirían clasificar a la población, identificando fenotipos de la enfermedad o detectando pacientes con la enfermedad en etapas tempranas e incluso pronosticando su desarrollo, permitiendo un mejor seguimiento del paciente y, también, mejorando la selección de individuos que participen en ensayos clínicos para desarrollar nuevos fármacos para combatir la artrosis. Los modelos predictivos para el fenotipo de artrosis de rodilla con progresión rápida obtenidos durante esta tesis doctoral podrían servir como una herramienta de aplicación clínica directa. En el año 2014 se destacó la importancia de diseñar una herramienta que fuera capaz de realizar esta tarea (Conaghan *et al.*, 2014), ya que permitiría aplicar el enfoque de la medicina personalizada en los pacientes que sufren esta enfermedad; además, facilitaría la identificación de individuos con alto riesgo de desarrollar este fenotipo para incluirlos en ensayos clínicos dada la rápida evolución en el deterioro estructural de la articulación.

Inicialmente, los primeros trabajos predictivos en el campo de la artrosis se centraban en llevar un enfoque dirigido similar al nuestro, es decir, los autores realizaron una búsqueda bibliográfica para identificar las variables predictoras de interés, que más adelante fueron introducidas en el modelo. El primer estudio que desarrolló modelos predictivos en artrosis de rodilla usando factores de riesgo convencionales apareció en 2011 (Zhang *et al.*, 2011). En este trabajo se trabajó con tres cohortes de pacientes, entre ellas la OAI, y se desarrollaron tres modelos predictivos de artrosis de rodilla, los dos primeros consistieron en un modelo de incidencia radiográfica (alcanzar durante el seguimiento un grado KL mayor o igual a dos) y otro de incidencia sintomática (alcanzar durante el seguimiento un grado KL mayor o igual a dos y tener dolor en esa misma rodilla). En ambos modelos las variables predictoras fueron la edad, el sexo femenino, el IMC, la ocupación laboral, el historial familiar y la lesión previa. El tercer modelo se centraba en la progresión de la artrosis de rodilla, definiendo esta como el avance de un grado KL o más en una rodilla que ya presenta artrosis establecida; en este tercer modelo las variables que incluyeron fueron la edad, el sexo femenino, la lesión previa, la actividad física regular y el número de compartimentos de rodilla con artrosis radiográfica. Estos modelos fueron creados en una cohorte retrospectiva de 12 años de seguimiento, y validados en tres cohortes, aunque con un bajo poder de discriminación, probablemente debido al bajo tamaño muestral (N=424). A pesar de estas debilidades, este primer estudio estableció un pilar en el campo de la artrosis, a partir del cual otros grupos de investigación generaron sus propios modelos predictivos.

En el año 2014 se llevó a cabo el primer estudio que, además de las variables clínicas clásicas, incorporó a los modelos predictivos una serie de variables genéticas (Kerkhof *et al.*, 2014). Este estudio tenía como objetivo desarrollar y validar un modelo predictivo para la artrosis de rodilla incipiente en una población envejecida. El desarrollo del modelo se llevó a cabo en la subcohorte I del estudio Rotterdam y, se validó en la subcohorte II del estudio Rotterdam y en la cohorte del estudio Chingford. El modelo inicial contenía solamente las variables edad, sexo e IMC, pero generaron nuevos modelos incrementando poco a poco las variables. El segundo modelo añadía las variables obtenidas a partir de diferentes cuestionarios relacionados con la funcionalidad. El tercer modelo incorporaba variables genéticas, que consistían en nueve SNPs previamente asociados con la artrosis de rodilla en meta-análisis de



poblaciones caucásicas. El cuarto modelo, denominado modelo completo, incorporaba al modelo tres las variables radiográficas (como el grado KL basal y el dolor en la rodilla). Finalmente crearon un quinto modelo donde incorporaban la edad, el sexo, el IMC, las variables de los cuestionarios y las variables radiográficas, pero no las genéticas. Con estos cinco modelos pudieron hacer estudios comparativos que permitían observar la precisión que aportaban cada una de las variables en la fase de validación. Como resultado observaron que la incorporación al primer modelo de las variables genéticas y de los cuestionarios no aumentaba la especificidad y la sensibilidad del modelo inicial; sin embargo, la incorporación de las variables radiográficas sí aumentaba estos valores. Por tanto, una conclusión que obtuvieron de este estudio era que los factores genéticos no parecían ser un buen predictor de la artrosis de rodilla radiográfica en poblaciones envejecidas. Sin embargo, esta conclusión es algo precipitada ya que en primer lugar hay que tener en cuenta que la heredabilidad de la artrosis es moderada (Brody, 2015), es decir, las variables genéticas sí juegan un papel importante en esta patología, pero no son el único factor determinante, sino que predisponen basalmente a cierto grado de riesgo. Además en segundo lugar, a pesar de que se han realizado numerosos estudios GWAS en la artrosis, en la gran mayoría de ellos los SNPs detectados presentan un OR moderado y se piensa que es el efecto poligénico el que puede desencadenar esta patología (Boer *et al.*, 2021), por lo que puede que los SNPs que se emplearon en el análisis no sean los más adecuados para analizar el efecto genético en la población de estudio. Recientemente en 2022, Kulm y colaboradores (Kulm *et al.*, 2022) también mostraron una tendencia similar en un estudio en el que trataron de desarrollar un modelo predictivo de prótesis de rodilla incluyendo variables clínicas (edad, sexo e IMC) y genéticas. De nuevo ellos pudieron observar que las variables genéticas, en comparación con las clínicas, no aportaban ningún valor predictivo. Sin embargo, un dato muy importante es que al realizar el análisis seleccionando tan solo individuos con edades menores de 60 años, el impacto de las variables genéticas era mucho mayor, por tanto, esto abre la posibilidad de que las variables genéticas jueguen un papel clave en la predicción temprana de una artrosis severa.

Otro estudio interesante se llevó a cabo en el año 2017 y fue el que generó la herramienta denominada TOARP (del inglés *Tool for Osteoarthritis Risk Prediction*) (Joseph *et al.*, 2018). En este estudio, los autores diseñaron una herramienta capaz de predecir la artrosis moderada o severa de rodilla en un período de 8 años. En este estudio se utilizaron datos de la cohorte OAI y se realizó una validación cruzada interna mediante técnicas de *bootstrap*. Se generaron tres modelos, todos ellos tenían en común las variables clásicas de edad, sexo e IMC. El primer modelo se centraba en factores de riesgo como la lesión previa en algún momento, la lesión previa en los últimos 12 meses, el historial familiar de cirugía de reemplazo articular, el grado KL y el ángulo de alineamiento. El segundo modelo incorporaba, adicionalmente, defectos en el cartílago y desgarros en los meniscos. El último modelo añadía medidas compartimentarizadas automáticamente del cartílago articular del fémur, tibia y rotula. Al realizar los análisis de sensibilidad y especificidad, el modelo basal (edad, sexo e IMC) alcanzaba unos valores de AUC de 0,60, la adición a este modelo del grado KL y de la lesión previa aumentaba este valor hasta 0,67. El segundo y el tercer modelo alcanzaban unos valores muy similares de 0,71 y 0,72, existiendo diferencias significativas entre el modelo 1 y el modelo 3. Una vez más, la adición de las variables radiográficas era lo que más potenciaba la precisión del modelo, aunque en este estudio los defectos en el cartílago articular también mostraron su efecto claro en la artrosis de rodilla (Joseph *et al.*, 2018). Recientemente, en 2022 se publicó un estudio que generaba una herramienta capaz de predecir el riesgo de un individuo de sufrir una cirugía de reemplazo articular en un período de cinco años (Liu *et al.*, 2022). En este estudio trabajaron con la cohorte MOST (N=2.658) y validaron los datos en la cohorte OAI (N=4.060) y

## Discusión

generaron distintos modelos empleando variables clínicas. Tras realizar la validación obtuvieron un modelo con gran capacidad predictiva; dicha herramienta incluía la edad, el sexo, la raza, el dolor frecuente en la rodilla, el uso común de analgésicos, el dolor WOMAC y el grado KL y, con este modelo, se obtuvo un AUC de 0,87, mostrando por tanto una elevada precisión.

También es importante destacar el papel de los biomarcadores proteómicos en el desarrollo de modelos predictivos de esta enfermedad. En 2019, Camacho-Encina y colaboradores descubrieron que se podían emplear autoanticuerpos como predictores de artrosis de rodilla temprana (Camacho-Encina *et al.*, 2019). En este estudio se empleó la subcohorte de incidencia de la OAI y, concretamente, detectaron que los niveles altos del autoanticuerpo contra la metionina adenosiltransferasa humana 2 $\beta$  se correlacionaban con la incidencia de artrosis de rodilla y con el desarrollo temprano de esta patología. Al incorporar a los modelos predictivos este biomarcador, se mejoró significativamente la identificación y clasificación de los pacientes que podían desarrollar el desorden. Además, los autores fueron capaces de replicar los resultados en una subcohorte de la OAI independiente. En 2021, otros miembros de este grupo identificaron nuevos biomarcadores proteicos capaces de predecir la artrosis de rodilla radiográfica (Lourido *et al.*, 2021). Detectaron que los altos niveles del inhibidor Inter-Alfa-tripsina de la cadena pesada H1 se asociaban con el desarrollo de la artrosis de rodilla radiográfica. Al incluir este biomarcador proteico en el modelo predictivo compuesto por la edad, el sexo, el IMC, la lesión previa y el WOMAC, la capacidad predictiva del modelo mejoró, pasando de un AUC de 0,754, a 0,786.

Pero no solamente se han publicado modelos predictivos en la rodilla. En el año 2018, Hosnijeh y colaboradores (Hosnijeh *et al.*, 2018), desarrollaron y validaron un modelo predictivo para la incidencia de la artrosis de cadera. El estudio se llevó a cabo en la subcohorte Rotterdam I y se validó en la subcohorte Rotterdam II y en la cohorte CHECK, e incorporó muchas variables previamente descritas en otros trabajos aquí mencionados. Generaron un modelo basal que además de la edad, el sexo y el IMC, incluía el hábito tabáquico, los niveles de proteína C reactiva, el historial familiar de la artrosis, la diabetes, el dolor en la cadera y el nivel de educación. A partir de este modelo generaron otros cuatro, en el primero solo añadieron al modelo inicial datos de imagen, en el segundo incorporaron únicamente al modelo inicial marcadores bioquímicos del colágeno II, en el tercero incluyeron exclusivamente al modelo inicial marcadores genéticos y, en el cuarto, añadieron al modelo inicial tanto los datos de imagen como los marcadores bioquímicos de colágeno II. Tras la validación, pudieron observar que el modelo básico presentaba un AUC de 0,60 en Rotterdam II y 0,54 en CHECK. Si al modelo se añadían solamente datos genéticos o bioquímicos, este dato de sensibilidad mejoraba levemente a 0,62 y 0,57 para la genética, y 0,65 y 0,61 para el marcador bioquímico de colágeno II. El mejor predictor resultó ser el modelo que incorporaba solamente los datos de imagen que incrementaba el AUC hasta 0,75 y 0,71, aunque la adición a este último modelo de las variables bioquímicas no modificaba significativamente este dato. Los resultados obtenidos en este estudio centrado en cadera van en consonancia con los obtenidos en los estudios mencionados en rodilla, donde los datos de imagen son los que parecen aportar más sensibilidad y especificidad al modelo.

Centrándonos en el fenotipo de progresión rápida, en el año 2016 se llevó a cabo un estudio que tenía como objetivo el diseño de una herramienta predictiva de artrosis incidente acelerada a 4 ó 5 años (Riddle *et al.*, 2016) . En este trabajo realizaron la fase de exploración en la cohorte MOST (N=1.690) y la fase de validación de los datos en la cohorte OAI (N=2.422). Las variables que introdujeron en el modelo fueron, artrosis de rodilla contralateral, el grado KL

basal, el IMC y el WOMAC. Con este modelo fueron capaces de alcanzar un AUC de 0,79 en la exploración y 0,81 en la validación.

Un aspecto negativo de todos estos estudios dirigidos puede ser que se centran en variables que precisan de procedimientos que requieren tiempo, como los ensayos de laboratorio o técnicas de imagen, los cuales son variables que no se recogen de forma rutinaria en atención primaria. Otro aspecto negativo es que la mayoría de estos estudios dirigidos cuentan con un tamaño humilde, rondando un valor pequeño de aproximadamente 400 participantes, aunque existen trabajos que alcanzan tamaños moderados de 3.000 participantes, pero generalmente, el poder estadístico del modelo puede verse comprometido. Para tratar de solventar estas dificultades, en 2020 se generó un modelo de predicción del desarrollo de artrosis en un período de 5 en cualquier articulación basado en variables rutinarias presentes en las bases de datos de los centros médicos (Black *et al.*, 2020). Este modelo ambicioso permitiría eliminar la necesidad de obtener datos adicionales y realizar una estimación en tiempo real una vez el paciente llega a la consulta por primera vez. Este estudio utilizó datos de una cohorte formada por 383.117 participantes donde se realizó una validación interna cruzada. El modelo resultante incorporaba la edad, el IMC, el sexo femenino, la lesión previa y la presencia de un diagnóstico previo de osteoporosis.

#### *Desarrollo de modelos predictivos para la artrosis basados en metodologías computacionales.*

Todos los estudios mencionados hasta ahora realizan modelos de forma dirigida, es decir, se basan en datos previos en la literatura. Actualmente existen bases de datos que cuentan con numerosos casos y variables, como es el caso de la OAI. Con esta ingente cantidad de datos, se podrían buscar nuevas variables menos convencionales que pudieran aumentar el poder predictivo de los modelos relacionados con la artrosis. Pero, el problema que tiene esta metodología convencional es que es incapaz de soportar cantidades ingentes de información. Para solventar esto, a día de hoy existe una nueva tendencia en la investigación predictiva que trata de aplicar potentes métodos computacionales, entre los que destacan los basados en las metodologías de *big data*, *data mining* y *machine learning* (Jamshidi *et al.*, 2019).

En el año 2020, el grupo de Leung y colaboradores empleó la tecnología de *Deep learning* en la cohorte OAI para desarrollar una herramienta capaz de predecir el riesgo de progresar hasta la cirugía de reemplazo articular en la artrosis de rodilla en un período de 9 años. Para ello usaron una red neuronal y analizaron radiografías de 324 pacientes que se sometieron a cirugías de reemplazo articular, y radiografías de 324 controles a los que no se les sometió a esta cirugía (Leung *et al.*, 2020). El modelo generado mediante este procedimiento alcanzó un AUC de 0,87 (IC 0,85-0,90). Posteriormente, en el año 2021, el grupo de Jamshidi y colaboradores emplearon metodologías de *machine learning* para generar un modelo capaz de predecir en una rodilla artrósica, el riesgo y el tiempo hasta el reemplazo articular en pacientes de la cohorte del OAI (Jamshidi *et al.*, 2021). Este trabajo muestra la particularidad que, de las 1.107 características documentadas en la cohorte del OAI, los autores fueron capaces de identificar, con esta metodología computacional, las 10 que aportaban un mayor riesgo de reemplazo articular. Generaron modelos con distintas herramientas y, entre los modelos seleccionados, obtuvieron uno que se centraba solamente en tres características, las lesiones de médula ósea, el grado KL y los síntomas en la rodilla, con el cual obtenían un AUC de 0,86. Como se puede ver, estos trabajos no incluyen variables genéticas, sino que se centran en datos obtenidos a través de imagen. Sin embargo, en el año 2022, y utilizando metodologías de *machine learning*, Bonakdari y colaboradores analizaron si alguno de los SNPs nucleares descritos en esta tesis y los

## Discusión

haplogrupos mitocondriales, de forma independiente o combinada, eran capaces de identificar individuos que sufren una progresión estructural de artrosis muy temprana (Bonakdari *et al.*, 2022). Para ello emplearon datos procedentes de 901 participantes de la cohorte OAI para entrenar y testar distintos modelos y metodologías. Los datos obtenidos fueron sometidos posteriormente a un proceso de validación interna cruzada mediante *bootstrap*, y finalmente, se analizó la reproducibilidad de los datos externamente en la cohorte TASSOAC. En total fueron capaces de generar 277 modelos, todos ellos incluían como variables clínicas la edad y el IMC. Los dos modelos más robustos fueron, en primer lugar, el que incorporaba los SNPs en los genes TP63, DUS4L, GDF5 y FTO, alcanzando una precisión del 85%; mientras que el segundo, incorporaba los haplogrupos mitocondriales y los SNPs en FTO y SUPT3H, con una precisión del 82,5%.

También en 2022, se evaluó el uso de la metodología de *machine learning* en la cohorte del estudio IMI-APPROACH, para predecir la progresión del dolor o la progresión estructural de la enfermedad (van Helvoort *et al.*, 2022). Usando distintos algoritmos, se calcularon unas puntuaciones de predicción de progresión de dolor y progresión de daño estructural y se compararon estas puntuaciones con los valores reales detectados en dicha cohorte, los cuales se basaban en mediciones de mJSW y los valores de dolor del cuestionario *Knee injury and Osteoarthritis Outcomes Score* (KOOS). Ambos predictores mostraron buenos valores y correlación con los datos reales, especialmente en el caso de los predictores de dolor, mostrando capacidad de distinción entre progresores y no progresores.

Todo este esfuerzo colaborativo de los distintos grupos de investigación permite poco a poco avanzar hacia la medicina personalizada en la artrosis. Esto, en primer lugar, permitirá mejorar la efectividad de los ensayos clínicos, mejorando la selección de participantes con el fin de desarrollar un tratamiento eficaz contra la artrosis. También favorecería la identificación de aquellos pacientes que podrían responder mejor a determinados tratamientos, facilitando así el asesoramiento de los médicos clínicos. En lo referente al modelo de predicción de artrosis de rodilla con progresión rápida desarrollado en esta tesis, este podría ser empleado directamente en la clínica, donde tras extraer una pequeña muestra de sangre del paciente, se podría llevar a cabo el análisis de sus variables genéticas en menos de 48 horas. De esta forma en un plazo corto de tiempo, se podrían evaluar las variables clínicas y genéticas, pudiendo predecir en poco tiempo la probabilidad de dicho individuo de progresar rápidamente en la artrosis de rodilla, otorgando cierto tiempo de respuesta al reumatólogo para prevenir el avance de la enfermedad. Por otro lado, este modelo también podría ser empleado en ensayos clínicos para identificar pacientes progresores rápidos y poder así generar cohortes de participantes que funcionan como un modelo a pequeña escala de la enfermedad, permitiendo así poder mejorar la evaluación de distintos DMOADs, tan necesarios para poder combatir esta incapacitante patología crónica.

### Limitaciones del estudio

Del mismo modo que ocurre en la gran mayoría de los proyectos de investigación, siempre existen limitaciones en los estudios desarrollados y, esta tesis doctoral no es una excepción.

1. Las cohortes de estudio empleadas: la cohorte OAI se ha empleado en numerosos estudios predictivos con diversas patologías, y gracias a que consta de un generoso tamaño muestral, períodos de seguimiento bien definidos y un amplio arsenal de variables, esta cohorte permite obtener resultados muy robustos, reforzando así la

fiabilidad de los datos obtenidos en este estudio. Si bien es cierto que además se ha podido realizar la validación de los datos mediante validación interna cruzada, siempre es más idóneo poder realizar la validación de los datos en cohortes externas, ya que permite aumentar la variabilidad poblacional, ajustándose la predicción del modelo más a la sociedad real en la que vivimos. Las cohortes externas empleadas en la fase de validación de esta tesis doctoral constan de un tamaño muestral pequeño. Aunque este tamaño fue suficiente para poder validar la asociación de la variante mitocondrial m.16519C, el número de muestras se tornó especialmente crítico cuando se realizó el estudio de interacciones ya que algunas combinaciones de SNPs y haplogrupos presentaban unos números bajos. Esto produjo que tanto la interacción mito-nuclear como el modelo predictivo tuvieran que ser validados en la cohorte OAI mediante metodologías de validación cruzada.

Además, si nos centramos en el número total de progresores rápidos con el criterio GIR-blanco, debido a la frecuencia de este fenotipo en la población, el número de individuos con este fenotipo era relativamente bajo, siendo de 107 en las cohortes externas, lo cual era menos de la mitad de lo que se obtenía en la cohorte OAI (N=268). Quizá en estudios futuros de este fenotipo sería ideal desarrollar una cohorte enfocada en este grupo de individuos, permitiendo así alcanzar un poder estadístico mayor, pudiendo detectar nuevas asociaciones que probablemente en este estudio no fueron detectadas, favoreciendo así la caracterización del fenotipo.

2. La selección de variables predictoras: en este trabajo, con el objetivo de desarrollar una herramienta capaz de predecir el fenotipo rápido de artrosis de rodilla, se emplearon variables predictoras previamente asociadas de forma robusta con este fenotipo en la literatura, lo cual refuerza los datos obtenidos en esta tesis. Sin embargo, la robustez del modelo predictivo podría ser mejorada, ya que la cohorte OAI contiene numerosas variables clínicas y, por tanto, en estudios futuros se podría tratar de usar un método computacional basado en técnicas de *machine learning* o *big data* que permita identificar más variables que se asocien específicamente con el fenotipo de progresión rápida. Por otro lado, en lo referente a las variables genéticas nucleares, las que se seleccionaron en este trabajo se asociaron previamente con la susceptibilidad a la enfermedad en diferentes estudios GWAS, habiendo hecho su selección a partir de una revisión publicada en el año 2017 (Warner y Valdes, 2017). Desde entonces hasta ahora se han detectado nuevos polimorfismos asociados con la enfermedad (Boer *et al.*, 2022) que quizás puedan tener un mayor efecto en el fenotipo que los seleccionados para esta tesis.
3. Estudio funcional: en este estudio se ha analizado, por primera vez, el posible efecto funcional de la variante mitocondrial m16519C usando cíbridos transmitocondriales. Sin embargo, los datos del estudio funcional se han centrado en cíbridos obtenidos a partir de la línea de osteosarcoma 143B.TK, por lo que no son completamente extrapolables a los condrocitos, que es el tipo celular presente en el cartílago articular. Cabe destacar que actualmente existen pocas líneas de condrocitos inmortalizadas, y cuentan con numerosas limitaciones por lo que el desarrollo de cíbridos con estas líneas todavía es un desafío. Por otro lado, las dos líneas de cíbridos empleadas en esta tesis son portadores del haplogrupo H, que es el haplogrupo mitocondrial con una menor divergencia a nivel de secuencia (Zaki *et al.*, 2009), sin embargo, una es un H76 y la otra H6b2, por lo que además de presentar genotipos distintos para la variante m.16519C, también están presentes otras 16 variantes mitocondriales que no son haplogrupo-específicas. Por lo que no se pueden confirmar rotundamente que los datos obtenidos

## *Discusión*

son única y exclusivamente debidos a la variante m.16519C. Por desgracia, a día de hoy no existe ninguna metodología de mutagénesis dirigida perfectamente establecida que permita editar y corregir una variante mitocondrial en homoplasmia.

# *Conclusiones*





1. *Se ha definido un nuevo criterio para artrosis de rodilla con progresión rápida (GIR-blanco) que permite clasificar un mayor número de pacientes con este fenotipo y con aplicación en la práctica clínica.*
2. *La variante m.16519C es un factor de riesgo para padecer el fenotipo de artrosis de rodilla con progresión rápida.*
3. *La presencia de la variante m.16519C condiciona los niveles de citoquinas proinflamatorias como la IL6, e incrementa los niveles de la COX2.*
4. *La interacción entre el alelo G del polimorfismo nuclear rs12107036 de TP63 con el cluster de haplogrupos UK se asocia de forma significativa con el riesgo de desarrollar el fenotipo de artrosis de rodilla con progresión rápida.*
5. *Se ha desarrollado una herramienta de aplicación clínica que permite seleccionar individuos para incluirlos en ensayos clínicos en los que se estudie fármacos contra esta patología.*



# *Bibliografía*



## Artículos:

Abdellah, Z. *et al.* (2004) 'Finishing the euchromatic sequence of the human genome', *Nature*. *Nature*, 431(7011), pp. 931–945. doi: 10.1038/nature03001.

Akouchekian, M. *et al.* (2009) 'High rate of mutation in mitochondrial DNA displacement loop region in human colorectal cancer', *Diseases of the Colon and Rectum*, 52(3), pp. 526–530. doi: 10.1007/DCR.0b013e31819acb99.

Alexe, G. *et al.* (2007) 'Enrichment of longevity phenotype in mtDNA haplogroups D4b2b, D4a, and D5 in the Japanese population', *Human Genetics*, 121(3–4), pp. 347–356. doi: 10.1007/s00439-007-0330-6.

Ali, S. A. *et al.* (2022) 'Circulating microRNAs differentiate fast-progressing from slow-progressing and non-progressing knee osteoarthritis in the Osteoarthritis Initiative cohort', *Therapeutic Advances in Musculoskeletal Disease*. *Ther Adv Musculoskelet Dis*, 14. doi: 10.1177/1759720X221082917.

Altman, D. G. (2009) 'Prognostic models: A methodological framework and review of models for breast cancer', *Cancer Investigation*. *Cancer Invest*, pp. 235–243. doi: 10.1080/07357900802572110.

Altman, R. D. *et al.* (1995) 'Atlas of individual radiographic features in osteoarthritis', *Osteoarthritis and Cartilage*. England, 3(SUPPL. A), pp. 3–70.

Altman, R. D. and Gold, G. E. (2007) 'Atlas of individual radiographic features in osteoarthritis, revised', *Osteoarthritis and Cartilage*. Elsevier, 15(SUPPL. 1), pp. 1–56. doi: 10.1016/j.joca.2006.06.017.

Amo, T. and Brand, M. D. (2007) 'Were inefficient mitochondrial haplogroups selected during migrations of modern humans? A test using modular kinetic analysis of coupling in mitochondria from cybrid cell lines', *Biochemical Journal*, 404(2), pp. 345–351. doi: 10.1042/BJ20061609.

Anderson, S. *et al.* (1981) 'Sequence and organization of the human mitochondrial genome', *Nature*, 290(5806), pp. 457–465. doi: 10.1038/290457a0.

Andrews, R. M. *et al.* (1999) 'Reanalysis and revision of the cambridge reference sequence for human mitochondrial DNA [5]', *Nature Genetics*. Nature Publishing Group, p. 147. doi: 10.1038/13779.

Andrews, S. J. *et al.* (2020) 'Mitonuclear interactions influence Alzheimer's disease risk', *Neurobiology of Aging*. Elsevier, 87, pp. 138.e7-138.e14. doi: 10.1016/j.neurobiolaging.2019.09.007.

Arnold, R. S. *et al.* (2013) 'An inherited heteroplasmic mutation in mitochondrial gene COI in a patient with prostate cancer alters reactive oxygen, reactive nitrogen and proliferation', *BioMed Research International*. Hindawi Publishing Corporation, 2013. doi: 10.1155/2013/239257.

Atilano, S. R. *et al.* (2015) 'Mitochondrial DNA variants can mediate methylation status of inflammation, angiogenesis and signaling genes', *Human Molecular Genetics*. Hum Mol Genet, 24(16), pp. 4491–4503. doi: 10.1093/hmg/ddv173.

Attur, M. *et al.* (2020) 'The combination of an inflammatory peripheral blood gene expression and imaging biomarkers enhance prediction of radiographic progression in knee osteoarthritis', *Arthritis Research and Therapy*. *Arthritis Res Ther*, 22(1). doi: 10.1186/s13075-

## Bibliografia

020-02298-6.

Aubourg, G. *et al.* (2022) 'Genetics of osteoarthritis', *Osteoarthritis and Cartilage*. Elsevier Ltd, 30(5), pp. 636–649. doi: 10.1016/j.joca.2021.03.002.

Ayral, X. *et al.* (2005) 'Synovitis: A potential predictive factor of structural progression of medial tibiofemoral knee osteoarthritis - Results of a 1 year longitudinal arthroscopic study in 422 patients', *Osteoarthritis and Cartilage*. Osteoarthritis Cartilage, 13(5), pp. 361–367. doi: 10.1016/j.joca.2005.01.005.

Baccarelli, A. A. and Byun, H. M. (2015) 'Platelet mitochondrial DNA methylation: A potential new marker of cardiovascular disease', *Clinical Epigenetics*. Clin Epigenetics, 7(1). doi: 10.1186/s13148-015-0078-0.

Bai, R. K. *et al.* (2007) 'Mitochondrial genetic background modifies breast cancer risk', *Cancer Research*, 67(10), pp. 4687–4694. doi: 10.1158/0008-5472.CAN-06-3554.

Bai, Y. *et al.* (2014) 'Association of sequence polymorphism in the mitochondrial D-loop with chronic kidney disease', *Renal Failure*. Informa Healthcare USA, Inc, 36(5), pp. 781–784. doi: 10.3109/0886022X.2014.890842.

Balaban, R. S., Nemoto, S. and Finkel, T. (2005) 'Mitochondria, oxidants, and aging', *Cell*, 120(4), pp. 483–495. doi: 10.1016/j.cell.2005.02.001.

Bannuru, R. R. *et al.* (2019) 'OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis', *Osteoarthritis and Cartilage*. Elsevier Ltd, 27(11), pp. 1578–1589. doi: 10.1016/j.joca.2019.06.011.

Bartlett, S. J. *et al.* (2011) 'Identifying common trajectories of joint space narrowing over two years in knee osteoarthritis', *Arthritis Care and Research*, 63(12), pp. 1722–1728. doi: 10.1002/acr.20614.

Bartolák-Suki, E. *et al.* (2017) 'Regulation of mitochondrial structure and dynamics by the cytoskeleton and mechanical factors', *International Journal of Molecular Sciences*, 18(8), pp. 7–11. doi: 10.3390/ijms18081812.

Bayramoglu, N., Nieminen, M. T. and Saarakkala, S. (2020) 'A Lightweight CNN and Joint Shape-Joint Space (JS<sup>2</sup>) Descriptor for Radiological Osteoarthritis Detection', in *Communications in Computer and Information Science*, pp. 331–345. doi: 10.1007/978-3-030-52791-4\_26.

Bellamy, N. *et al.* (1988) 'Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee.', *The Journal of rheumatology*, 15(12), pp. 1833–40. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3068365>.

Bellanti, F., Buglio, A. Lo and Vendemiale, G. (2021) 'Mitochondrial impairment in sarcopenia', *Biology*, 10(1), pp. 1–16. doi: 10.3390/biology10010031.

Bellizzi, D. *et al.* (2012) 'Global DNA methylation levels are modulated by mitochondrial DNA variants', *Epigenomics*, 4(1), pp. 17–27. doi: 10.2217/epi.11.109.

De Benedictis, G. *et al.* (1999) 'Mitochondrial DNA inherited variants are associated with successful aging and longevity in humans', *The FASEB Journal*. John Wiley & Sons, Ltd, 13(12), pp. 1532–1536. doi: 10.1096/fasebj.13.12.1532.

De Benedictis, G. *et al.* (2000) 'Inherited variability of the mitochondrial genome and successful aging in humans', in *Annals of the New York Academy of Sciences*. Ann N Y Acad Sci, pp. 208–218. doi: 10.1111/j.1749-6632.2000.tb06648.x.

- Bénil, P. *et al.* (2010) 'Genetic background influences mitochondrial function: modeling mitochondrial disease for therapeutic development', *Trends in Molecular Medicine*. Elsevier Ltd, 16(5), pp. 210–217. doi: 10.1016/j.molmed.2010.03.001.
- Bennell, K. L. *et al.* (2013) 'Update on the Role of Muscle in the Genesis and Management of Knee Osteoarthritis', *Rheumatic Disease Clinics of North America*, 39(1), pp. 145–176. doi: 10.1016/j.rdc.2012.11.003.
- Berenbaum, F. (2013) 'Osteoarthritis as an inflammatory disease (osteoarthritis is not osteoarthrosis!)', *Osteoarthritis and Cartilage*, 21(1), pp. 16–21. doi: 10.1016/j.joca.2012.11.012.
- Berridge, M. J. (2012) 'Calcium signalling remodelling and disease', *Biochemical Society Transactions*, 40(2), pp. 297–309. doi: 10.1042/BST20110766.
- Bi, R. *et al.* (2015) 'Mitochondrial DNA haplogroup B5 confers genetic susceptibility to Alzheimer's disease in Han Chinese', *Neurobiology of Aging*. Neurobiol Aging, 36(3), pp. 1604.e7-1604.e16. doi: 10.1016/j.neurobiolaging.2014.10.009.
- Bijlsma, J. W. J., Berenbaum, F. and Lafeber, F. P. J. G. (2011) 'Osteoarthritis: An update with relevance for clinical practice', *The Lancet*, 377(9783), pp. 2115–2126. doi: 10.1016/S0140-6736(11)60243-2.
- Bijsterbosch, J. *et al.* (2013) 'Association study of candidate genes for the progression of hand osteoarthritis', *Osteoarthritis and Cartilage*. Osteoarthritis Cartilage, 21(4), pp. 565–569. doi: 10.1016/j.joca.2013.01.011.
- Biłas, R. *et al.* (2016) 'Cis-regulatory elements used to control gene expression in plants', *Plant Cell, Tissue and Organ Culture*, pp. 269–287. doi: 10.1007/s11240-016-1057-7.
- Black, J. E., Terry, A. L. and Lizotte, D. J. (2020) 'Development and evaluation of an osteoarthritis risk model for integration into primary care health information technology.', *International journal of medical informatics*. Int J Med Inform, 141, p. 104160. doi: 10.1016/j.ijmedinf.2020.104160.
- Blair-Levy, J. M. *et al.* (2008) 'A type I collagen defect leads to rapidly progressive osteoarthritis in a mouse model', *Arthritis and Rheumatism*, 58(4), pp. 1096–1106. doi: 10.1002/art.23277.
- Blanch, M. *et al.* (2016) 'Altered mitochondrial DNA methylation pattern in Alzheimer disease-related pathology and in Parkinson disease', *American Journal of Pathology*. Am J Pathol, 186(2), pp. 385–397. doi: 10.1016/j.ajpath.2015.10.004.
- Blanco, F. J. *et al.* (2015) 'Improved prediction of knee osteoarthritis progression by genetic polymorphisms: The Arthrotest Study', *Rheumatology (United Kingdom)*. Rheumatology (Oxford), 54(7), pp. 1236–1243. doi: 10.1093/rheumatology/keu478.
- Blanco, F. J. (2018) 'Osteoarthritis and Atherosclerosis in Joint Disease', *Reumatología Clínica (English Edition)*. Elsevier España, S.L.U. and Sociedad Española de Reumatología y Colegio Mexicano de Reumatología, 14(5), pp. 251–253. doi: 10.1016/j.reumae.2018.08.001.
- Blanco, F. J. and Fernández-Moreno, M. (2020) 'Mitochondrial biogenesis: a potential therapeutic target for osteoarthritis', *Osteoarthritis and Cartilage*. Osteoarthritis Cartilage, pp. 1003–1006. doi: 10.1016/j.joca.2020.03.018.
- Blanco, F. J., López-Armada, M. J. and Maneiro, E. (2004) 'Mitochondrial dysfunction in osteoarthritis', *Mitochondrion*. Mitochondrion, 4(5-6 SPEC. ISS.), pp. 715–728. doi: 10.1016/j.mito.2004.07.022.

## Bibliografía

- Blanco, F. J., Rego, I. and Ruiz-Romero, C. (2011) 'The role of mitochondria in osteoarthritis', *Nature Reviews Rheumatology*. Nature Publishing Group, 7(3), pp. 161–169. doi: 10.1038/nrrheum.2010.213.
- Blanco, F. J., Valdes, A. M. and Rego-Pérez, I. (2018) 'Mitochondrial DNA variation and the pathogenesis of osteoarthritis phenotypes', *Nature Reviews Rheumatology*, 14(6), pp. 327–340. doi: 10.1038/s41584-018-0001-0.
- Blein, S. *et al.* (2015) 'An original phylogenetic approach identified mitochondrial haplogroup T1a1 as inversely associated with breast cancer risk in BRCA2 mutation carriers', *Breast Cancer Research*, 17(1), pp. 1–15. doi: 10.1186/s13058-015-0567-2.
- Blount, S. G. *et al.* (1954) 'Isolated valvular pulmonic stenosis clinical and physiologic response to open valvuloplasty.', *Circulation*, 10(2), pp. 161–172. doi: 10.1161/01.CIR.10.2.161.
- Boer, C. G. *et al.* (2021) 'Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations', *Cell*, 184(18), pp. 4784–4818.e17. doi: 10.1016/j.cell.2021.07.038.
- Bolduc, J. A., Collins, J. A. and Loeser, R. F. (2019) 'Reactive oxygen species, aging and articular cartilage homeostasis', *Free Radical Biology and Medicine*. Free Radic Biol Med, pp. 73–82. doi: 10.1016/j.freeradbiomed.2018.08.038.
- Bonakdari, H. *et al.* (2022) 'Single nucleotide polymorphism genes and mitochondrial DNA haplogroups as biomarkers for early prediction of knee osteoarthritis structural progressors: use of supervised machine learning classifiers', *BMC Medicine*. BMC Med, 20(1), p. 316. doi: 10.1186/s12916-022-02491-1.
- Booker, L. M. *et al.* (2006) 'North American white mitochondrial haplogroups in prostate and renal cancer', *Journal of Urology*. J Urol, 175(2), pp. 468–473. doi: 10.1016/S0022-5347(05)00163-1.
- Borroni, B. and Benussi, A. (2019) 'Recent advances in understanding frontotemporal degeneration', *F1000Research*, 8, pp. 1–11. doi: 10.12688/f1000research.20330.1.
- Bosse, Y. and Amos, C. I. (2018) 'A decade of GWAS results in lung cancer', *Cancer Epidemiology Biomarkers and Prevention*, 27(4), pp. 363–379. doi: 10.1158/1055-9965.EPI-16-0794.
- Brinker, A. E. *et al.* (2017) 'Mitochondrial haplotype alters mammary cancer tumorigenicity and metastasis in an oncogenic driver-dependent manner', *Cancer Research*, 77(24), pp. 6941–6949. doi: 10.1158/0008-5472.CAN-17-2194.
- Brocks, J. J. *et al.* (1999) 'Archean molecular fossils and the early rise of eukaryotes', *Science*. Science, 285(5430), pp. 1033–1036. doi: 10.1126/science.285.5430.1033.
- Brody, L. T. (2015) 'Knee osteoarthritis: Clinical connections to articular cartilage structure and function', *Physical Therapy in Sport*. Elsevier Ltd, 16(4), pp. 301–316. doi: 10.1016/j.ptsp.2014.12.001.
- Brown, M. D. *et al.* (2002) 'The role of mtDNA background in disease expression: A new primary LHON mutation associated with Western Eurasian haplogroup J', *Human Genetics*, 110(2), pp. 130–138. doi: 10.1007/s00439-001-0660-8.
- Brown, M. T. *et al.* (2012) 'Tanezumab reduces osteoarthritic knee pain: Results of a randomized, double-blind, placebo-controlled phase III trial', *Journal of Pain*. J Pain, 13(8), pp. 790–798. doi: 10.1016/j.jpain.2012.05.006.
- Brown, M. T. *et al.* (2013) 'Tanezumab reduces osteoarthritic hip pain: Results of a



- randomized, double-blind, placebo-controlled phase III trial', *Arthritis and Rheumatism*. *Arthritis Rheum*, 65(7), pp. 1795–1803. doi: 10.1002/art.37950.
- Burr, D. B. and Gallant, M. A. (2012) 'Bone remodelling in osteoarthritis', *Nature reviews. Rheumatology*. *Nat Rev Rheumatol*, 8(11), pp. 665–673. doi: 10.1038/NRRHEUM.2012.130.
- Bushel, P. R. *et al.* (2022) 'Mitochondrial-nuclear epistasis underlying phenotypic variation in breast cancer pathology', *Scientific Reports*. Nature Publishing Group UK, 12(1), pp. 1–10. doi: 10.1038/s41598-022-05148-4.
- Bussard, K. M. and Siracusa, L. D. (2017) 'Understanding mitochondrial polymorphisms in cancer', *Cancer Research*. American Association for Cancer Research Inc., pp. 6051–6059. doi: 10.1158/0008-5472.CAN-17-1939.
- Cai, N. *et al.* (2021) 'Mitochondrial DNA variants modulate N-formylmethionine, proteostasis and risk of late-onset human diseases', *Nature Medicine*. *Nat Med*, 27(9), pp. 1564–1575. doi: 10.1038/s41591-021-01441-3.
- Cai, X. Y. *et al.* (2009) 'Association of mitochondrial DNA haplogroups with exceptional longevity in a Chinese population', *PLoS ONE*, 4(7), pp. 3–8. doi: 10.1371/journal.pone.0006423.
- Camacho-Encina, M. *et al.* (2019) 'Discovery of an autoantibody signature for the early diagnosis of knee osteoarthritis: Data from the Osteoarthritis Initiative', *Annals of the Rheumatic Diseases*. *Ann Rheum Dis*, 78(12), pp. 1699–1705. doi: 10.1136/annrheumdis-2019-215325.
- Camilleri, M. *et al.* (2009) 'Mitochondrial DNA and gastrointestinal motor and sensory functions in health and functional gastrointestinal disorders', *American Journal of Physiology - Gastrointestinal and Liver Physiology*. *Am J Physiol Gastrointest Liver Physiol*, 296(3), pp. 510–516. doi: 10.1152/ajpgi.90650.2008.
- Carman, W. J. *et al.* (1940) 'American journal of physics', *Journal of the Franklin Institute*, 229(2), p. 280. doi: 10.1016/s0016-0032(40)90500-2.
- Carmona, L. *et al.* (2001) 'The burden of musculoskeletal diseases in the general population of Spain: Results from a national survey', *Annals of the Rheumatic Diseases*, 60(11), pp. 1040–1045. doi: 10.1136/ard.60.11.1040.
- Castro, M. G. *et al.* (2006) 'Mitochondrial DNA haplogroups in Spanish patients with hypertrophic cardiomyopathy', *International Journal of Cardiology*. *Int J Cardiol*, 112(2), pp. 202–206. doi: 10.1016/j.ijcard.2005.09.008.
- Chang, M. C. *et al.* (2005) 'Accumulation of mitochondrial DNA with 4977-bp deletion in knee cartilage - An association with idiopathic osteoarthritis', *Osteoarthritis and Cartilage*, 13(11), pp. 1004–1011. doi: 10.1016/j.joca.2005.06.011.
- Charlier, E. *et al.* (2019) 'Chondrocyte dedifferentiation and osteoarthritis (OA)', *Biochemical Pharmacology*. Elsevier, 165(January), pp. 49–65. doi: 10.1016/j.bcp.2019.02.036.
- Charoute, H. *et al.* (2018) 'Novel variants of mitochondrial DNA associated with Type 2 diabetes mellitus in Moroccan population', *Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis*. Taylor & Francis, 29(1), pp. 9–13. doi: 10.1080/24701394.2016.1233530.
- Chen, F. H. and Tuan, R. S. (2008) 'Mesenchymal stem cells in arthritic diseases', *Arthritis Research and Therapy*, 10(5), p. 223. doi: 10.1186/ar2514.

## Bibliografia

Cheng, B. *et al.* (2021) 'Genetic association scan of 32 osteoarthritis susceptibility genes identified TP63 associated with an endemic osteoarthritis, Kashin-Beck disease', *Bone*. Elsevier Inc., 150(May), p. 115997. doi: 10.1016/j.bone.2021.115997.

Cheng, O. T. *et al.* (2012) 'Evidence-Based Knee Injections for the Management of Arthritis', *Pain Medicine (United States)*, 13(6), pp. 740–753. doi: 10.1111/j.1526-4637.2012.01394.x.

Chinnery, P. F. and Gomez-Duran, A. (2018) 'Oldies but goldies mtDNA population variants and neurodegenerative diseases', *Frontiers in Neuroscience*, 12(OCT), pp. 1–11. doi: 10.3389/fnins.2018.00682.

Chocron, E. S., Munkácsy, E. and Pickering, A. M. (2019) 'Cause or casualty: The role of mitochondrial DNA in aging and age-associated disease', *Biochimica et Biophysica Acta - Molecular Basis of Disease*. Elsevier, 1865(2), pp. 285–297. doi: 10.1016/j.bbadis.2018.09.035.

Chomyn, A. *et al.* (1994) 'Platelet-mediated transformation of mtDNA-less human cells: Analysis of phenotypic variability among clones from normal individuals-and complementation behavior of the tRNA(Lys) mutation causing myoclonic epilepsy and ragged red fibers', *American Journal of Human Genetics*. Elsevier, 54(6), pp. 966–974. Available at: /pmc/articles/PMC1918182/?report=abstract (Accessed: 11 October 2022).

Chou, C. H. *et al.* (2020) 'Synovial cell cross-talk with cartilage plays a major role in the pathogenesis of osteoarthritis', *Scientific Reports*. Nature Publishing Group UK, 10(1), pp. 1–14. doi: 10.1038/s41598-020-67730-y.

Christodoulou, J. and Jansen, R. (2000) 'Genetic defects causing mitochondrial respiratory chain disorders and disease', *Human Reproduction*, 15(SUPPL. 2), pp. 28–43. doi: 10.1093/humrep/15.suppl\_2.28.

Cillero-Pastor, B. *et al.* (2008) 'Mitochondrial dysfunction activates cyclooxygenase 2 expression in cultured normal human chondrocytes', *Arthritis and Rheumatism*, 58(8), pp. 2409–2419. doi: 10.1002/art.23644.

Clegg, T. E., Caborn, D. and Mauffrey, C. (2013) 'Viscosupplementation with hyaluronic acid in the treatment for cartilage lesions: A review of current evidence and future directions', *European Journal of Orthopaedic Surgery and Traumatology*, 23(2), pp. 119–124. doi: 10.1007/s00590-012-0940-0.

Collins, J. A. *et al.* (2016) 'Oxidative stress promotes peroxiredoxin hyperoxidation and attenuates pro-survival signaling in aging chondrocytes', *Journal of Biological Chemistry*. J Biol Chem, 291(13), pp. 6641–6654. doi: 10.1074/jbc.M115.693523.

Conaghan, P. G. *et al.* (2014) 'Osteoarthritis research priorities: A report from a EULAR ad hoc expert committee', *Annals of the Rheumatic Diseases*. BMJ Publishing Group Ltd, 73(8), pp. 1442–1445. doi: 10.1136/annrheumdis-2013-204660.

Coppedè, F. and Stoccoro, A. (2019) 'Mitoepigenetics and neurodegenerative diseases', *Frontiers in Endocrinology*. Frontiers Media SA, p. 86. doi: 10.3389/fendo.2019.00086.

Coste, F., Laurent, F. and Benichou, C. (1959) 'Coxarthrose avec lyse de la tete fémorale Coxarthrose usante', *Rev. du Rhumat*, 26, pp. 305–308.

Counsell, C. and Dennis, M. (2001) 'Systematic review of prognostic models in patients with acute stroke', *Cerebrovascular diseases (Basel, Switzerland)*. Cerebrovasc Dis, 12(3), pp. 159–170. doi: 10.1159/000047699.

Courties, A., Berenbaum, F. and Sellam, J. (2019) 'The Phenotypic Approach to Osteoarthritis: A

Look at Metabolic Syndrome-Associated Osteoarthritis', *Joint Bone Spine*. Elsevier Masson SAS, 86(6), pp. 725–730. doi: 10.1016/j.jbspin.2018.12.005.

Covarrubias, D. *et al.* (2008) 'Mitochondrial DNA variant interactions modify breast cancer risk', *Journal of Human Genetics*. *J Hum Genet*, 53(10), pp. 924–928. doi: 10.1007/s10038-008-0331-x.

Creamer, P. and Hochberg, M. C. (1997) 'Osteoarthritis', *Lancet*, 350(9076), pp. 503–509. doi: 10.1016/S0140-6736(97)07226-7.

Creamer, P., Lethbridge-Cejku, M. and Hochberg, M. C. (1999) 'Determinants of pain severity in knee osteoarthritis: Effect of demographic and psychosocial variables using 3 pain measures', *Journal of Rheumatology*, 26(8), pp. 1785–1792. Available at: <https://pubmed.ncbi.nlm.nih.gov/10451078/>.

Csaki, C. *et al.* (2008) 'Regulation of inflammation signalling by resveratrol in human chondrocytes in vitro', *Biochemical Pharmacology*. *Biochem Pharmacol*, 75(3), pp. 677–687. doi: 10.1016/j.bcp.2007.09.014.

Culvenor, A. G. *et al.* (2015) 'Defining the presence of radiographic knee osteoarthritis: a comparison between the Kellgren and Lawrence system and OARSI atlas criteria', *Knee Surgery, Sports Traumatology, Arthroscopy*, 23(12), pp. 3532–3539. doi: 10.1007/s00167-014-3205-0.

D'Souza, A. R. and Minczuk, M. (2018) 'Mitochondrial transcription and translation: Overview', *Essays in Biochemistry*, 62(3), pp. 309–320. doi: 10.1042/EBC20170102.

Dai, J. *et al.* (2018) 'FTO variant is not associated with osteoarthritis in the Chinese Han population: Replication study for a genome-wide association study identified risk loci', *Journal of Orthopaedic Surgery and Research*. *Journal of Orthopaedic Surgery and Research*, 13(1), pp. 4–9. doi: 10.1186/s13018-018-0769-2.

Dalmao-Fernández, A. *et al.* (2020) 'Impaired Metabolic Flexibility in the Osteoarthritis Process: A Study on Transmitochondrial Cybrids', *Cells*, 9(4). doi: 10.3390/cells9040809.

Damen, J. A. A. G. *et al.* (2016) 'Prediction models for cardiovascular disease risk in the general population: Systematic review', *BMJ (Online)*. *BMJ*. doi: 10.1136/bmj.i2416.

Dato, S. *et al.* (2004) 'Association of the mitochondrial DNA haplogroup J with longevity is population specific', *European Journal of Human Genetics*. Nature Publishing Group, 12(12), pp. 1080–1082. doi: 10.1038/sj.ejhg.5201278.

Day-Williams, A. G. *et al.* (2011) 'A variant in MCF2L is associated with osteoarthritis', *American Journal of Human Genetics*. The American Society of Human Genetics, 89(3), pp. 446–450. doi: 10.1016/j.ajhg.2011.08.001.

Deepeshwar, S. *et al.* (2018) 'Effect of yoga based lifestyle intervention on patients with knee osteoarthritis: A randomized controlled trial', *Frontiers in Psychiatry*. Frontiers Media SA, 9(MAY), p. 180. doi: 10.3389/fpsy.2018.00180.

Dell'Isola, A. *et al.* (2016) 'Identification of clinical phenotypes in knee osteoarthritis: a systematic review of the literature', *BMC Musculoskeletal Disorders*. *BMC Musculoskeletal Disorders*, 17(1), pp. 1–12. doi: 10.1186/s12891-016-1286-2.

Demine, S., Renard, P. and Arnould, T. (2019) 'Mitochondrial uncoupling: A key controller of biological processes in physiology and diseases', *Cells*. Multidisciplinary Digital Publishing Institute (MDPI). doi: 10.3390/cells8080795.

## Bibliografia

Demontis, D. *et al.* (2019) 'Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder', *Nature Genetics*, 51(1), pp. 63–75. doi: 10.1038/s41588-018-0269-7.

Denton, R. M., Randle, P. J. and Martin, B. R. (1972) 'Stimulation by calcium ions of pyruvate dehydrogenase phosphate phosphatase.', *The Biochemical journal*, 128(1), pp. 161–163. doi: 10.1042/bj1280161.

Dequeker, J. and Luyten, F. P. (2008) 'The history of osteoarthritis-osteoarthrosis', *Annals of the Rheumatic Diseases*, 67(1), pp. 5–10. doi: 10.1136/ard.2007.079764.

Deveza, L. A., Nelson, A. E. and Loeser, R. F. (2019) 'Phenotypes of osteoarthritis: current state and future implications', *Clinical and experimental rheumatology*, 37(5), pp. 64–72.

van Dijk, E. L. *et al.* (2018) 'The Third Revolution in Sequencing Technology', *Trends in Genetics*. *Trends Genet*, 34(9), pp. 666–681. doi: 10.1016/j.tig.2018.05.008.

Dobson, G. P. *et al.* (2018) 'Defining the osteoarthritis patient: back to the future', *Osteoarthritis and Cartilage*, 26(8), pp. 1003–1007. doi: 10.1016/j.joca.2018.04.018.

Doré, D. *et al.* (2012) 'A longitudinal study of the association between dietary factors, serum lipids, and bone marrow lesions of the knee', *Arthritis Research and Therapy*. BioMed Central, 14(1), pp. 1–10. doi: 10.1186/ar3689.

Driban, J. B. *et al.* (2014) 'Association of Knee Injuries With Accelerated Knee Osteoarthritis Progression: Data From the Osteoarthritis Initiative', *Arthritis Care and Research*. NIH Public Access, 66(11), pp. 1673–1679. doi: 10.1002/acr.22359.

Driban, J. B., Stout, A. C., *et al.* (2016) 'Best performing definition of accelerated knee osteoarthritis: data from the Osteoarthritis Initiative', *Therapeutic Advances in Musculoskeletal Disease*, 8(5), pp. 165–171. doi: 10.1177/1759720X16658032.

Driban, J. B., Eaton, C. B., *et al.* (2016) 'Overweight older adults, particularly after an injury, are at high risk for accelerated knee osteoarthritis: Data from the Osteoarthritis Initiative', *Clinical Rheumatology*. *Clin Rheumatol*, 35(4), pp. 1071–1076. doi: 10.1007/s10067-015-3152-2.

Driban, J. B. *et al.* (2020) 'Risk factors and the natural history of accelerated knee osteoarthritis: A narrative review', *BMC Musculoskeletal Disorders*. *BMC Musculoskeletal Disorders*, 21(1), pp. 1–11. doi: 10.1186/s12891-020-03367-2.

Eckstein, F. *et al.* (2007) 'Two year longitudinal change and test-retest-precision of knee cartilage morphology in a pilot study for the osteoarthritis initiative', *Osteoarthritis and Cartilage*. *Osteoarthritis Cartilage*, 15(11), pp. 1326–1332. doi: 10.1016/j.joca.2007.04.007.

Eirin, A. *et al.* (2016a) 'Urinary Mitochondrial DNA Copy Number Identifies Chronic Renal Injury in Hypertensive Patients', *Hypertension*. *Hypertension*, 68(2), pp. 401–410. doi: 10.1161/HYPERTENSIONAHA.116.07849.

Eirin, A. *et al.* (2016b) 'Urinary Mitochondrial DNA Copy Number Identifies Chronic Renal Injury in Hypertensive Patients', *Hypertension*. *Hypertension*, 68(2), pp. 401–410. doi: 10.1161/HYPERTENSIONAHA.116.07849.

Emrani, P. S. *et al.* (2008) 'Joint space narrowing and Kellgren-Lawrence progression in knee osteoarthritis: an analytic literature synthesis', *Osteoarthritis and Cartilage*, 16(8), pp. 873–882. doi: 10.1016/j.joca.2007.12.004.

Espanha, M. M. (2010) 'Articular cartilage: structure and histochemical composition', *Acta Reumatologica Portuguesa*, 35(4), pp. 424–433. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/21245810> (Accessed: 30 June 2017).

Euser, A. M. *et al.* (2009) 'Cohort studies: Prospective versus retrospective', *Nephron - Clinical Practice*. Karger Publishers, pp. c214–c217. doi: 10.1159/000235241.

Evangelou, E. *et al.* (2011) 'Meta-analysis of genome-wide association studies confirms a susceptibility locus for knee osteoarthritis on chromosome 7q22', *Annals of the Rheumatic Diseases*, 70(2), pp. 349–355. doi: 10.1136/ard.2010.132787.

Falkenberg, M. (2018) 'Mitochondrial DNA replication in mammalian cells: Overview of the pathway', *Essays in Biochemistry*, 62(3), pp. 287–296. doi: 10.1042/EBC20170100.

Fang, H. *et al.* (2010) 'Cancer type-specific modulation of mitochondrial haplogroups in breast, colorectal and thyroid cancer', *BMC Cancer*. BioMed Central, 10(1), pp. 1–10. doi: 10.1186/1471-2407-10-421.

Fang, H. *et al.* (2014) 'Role of mtDNA haplogroups in the prevalence of knee osteoarthritis in a southern Chinese population', *International Journal of Molecular Sciences*, 15(2), pp. 2646–2659. doi: 10.3390/ijms15022646.

Fang, H. *et al.* (2016) 'Mitochondrial DNA haplogroups modify the risk of osteoarthritis by altering mitochondrial function and intracellular mitochondrial signals', *Biochimica et Biophysica Acta - Molecular Basis of Disease*. Biochim Biophys Acta, 1862(4), pp. 829–836. doi: 10.1016/j.bbdis.2015.12.017.

Farrar, G. J. *et al.* (2013) 'Mitochondrial disorders: Aetiologies, models systems, and candidate therapies', *Trends in Genetics*. Elsevier Ltd, 29(8), pp. 488–497. doi: 10.1016/j.tig.2013.05.005.

Fawcett, K. A. and Barroso, I. (2010) 'The genetics of obesity: FTO leads the way', *Trends in Genetics*, 26(6), pp. 266–274. doi: 10.1016/j.tig.2010.02.006.

Felson, D. *et al.* (2013) 'Progression of osteoarthritis as a state of inertia', *Annals of the Rheumatic Diseases*, 72(6), pp. 924–929. doi: 10.1136/annrheumdis-2012-201575.

Felson, D. T. *et al.* (1995) 'The incidence and natural history of knee osteoarthritis in the elderly, the framingham osteoarthritis study', *Arthritis & Rheumatism*. John Wiley & Sons, Ltd, 38(10), pp. 1500–1505. doi: 10.1002/art.1780381017.

Felson, D. T. *et al.* (2000) 'Osteoarthritis: New insights - Part 1: The disease and its risk factors', *Annals of Internal Medicine*, 133(8), pp. 635–646. doi: 10.7326/0003-4819-133-8-200010170-00016.

Felson, D. T. *et al.* (2011) 'Defining radiographic incidence and progression of knee osteoarthritis: Suggested modifications of the Kellgren and Lawrence scale', *Annals of the Rheumatic Diseases*, 70(11), pp. 1884–1886. doi: 10.1136/ard.2011.155119.

Felson, D. T. *et al.* (2013) 'Valgus malalignment is a risk factor for lateral knee osteoarthritis incidence and progression: Findings from the multicenter osteoarthritis study and the osteoarthritis initiative', *Arthritis and Rheumatism*. Arthritis Rheum, 65(2), pp. 355–362. doi: 10.1002/art.37726.

Feng, S. *et al.* (2012) 'Correlation between increased ND2 expression and demethylated displacement loop of mtDNA in colorectal cancer', *Molecular Medicine Reports*. Spandidos Publications, 6(1), pp. 125–130. doi: 10.3892/mmr.2012.870.

Fenlon, C. *et al.* (2018) 'A discussion of calibration techniques for evaluating binary and categorical predictive models', *Preventive Veterinary Medicine*. Prev Vet Med, pp. 107–114. doi: 10.1016/j.prevetmed.2017.11.018.

## Bibliografía

- Fernández-Caggiano, M. *et al.* (2012) 'Mitochondrial Haplogroups H and J: Risk and Protective Factors for Ischemic Cardiomyopathy', *PLoS ONE*. PLoS One, 7(8). doi: 10.1371/journal.pone.0044128.
- Fernández-Caggiano, M. *et al.* (2013) 'Mitochondrial DNA haplogroup H as a risk factor for idiopathic dilated cardiomyopathy in Spanish population', *Mitochondrion*. Mitochondrion, 13(4), pp. 263–268. doi: 10.1016/j.mito.2013.03.005.
- Fernández-Moreno, M. *et al.* (2011) 'MtDNA haplogroup J Modulates telomere length and Nitric Oxide production', *BMC Musculoskeletal Disorders*, 12(1), p. 283. doi: 10.1186/1471-2474-12-283.
- Fernández-Moreno, M. *et al.* (2012) 'Mitochondrial haplogroups define two phenotypes of osteoarthritis', *Frontiers in Physiology*, 3 MAY, p. 129. doi: 10.3389/fphys.2012.00129.
- Fernández-Moreno, M., Soto-Hermida, A., Vázquez-Mosquera, M. E., Cortés-Pereira, E., Pértega, S., *et al.* (2017) 'A replication study and meta-analysis of mitochondrial DNA variants in the radiographic progression of knee osteoarthritis', *Rheumatology (United Kingdom)*. Rheumatology (Oxford), 56(2), pp. 263–270. doi: 10.1093/rheumatology/kew394.
- Fernández-Moreno, M., Soto-Hermida, A., Vázquez-Mosquera, M. E., Cortés-Pereira, E., Relaño, S., *et al.* (2017) 'Mitochondrial DNA haplogroups influence the risk of incident knee osteoarthritis in OAI and CHECK cohorts. A meta-Analysis and functional study', *Annals of the Rheumatic Diseases*, 76(6), pp. 1114–1122. doi: 10.1136/annrheumdis-2016-210131.
- Ferrer-Sueta, G. and Radi, R. (2009) 'Chemical biology of peroxynitrite: Kinetics, diffusion, and radicals', *ACS Chemical Biology*, pp. 161–177. doi: 10.1021/cb800279q.
- Flemming, D. J. and Gustas-French, C. N. (2017) 'Rapidly Progressive Osteoarthritis: a Review of the Clinical and Radiologic Presentation', *Current Rheumatology Reports*. Current Rheumatology Reports, 19(7). doi: 10.1007/s11926-017-0665-5.
- Forestier, F. (1957) *Coxites rhumatismales subaigues et chroniques*. Paris. Available at: <https://data.bnf.fr/de/temp-work/d4b57dd9af4b1a0c9d8cbcc92dd60639/>.
- Frayling, T. M. *et al.* (2007) 'A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity', *Science*. American Association for the Advancement of Science, 316(5826), pp. 889–894. doi: 10.1126/science.1141634.
- Fukui, N. *et al.* (2008) 'Regional differences in chondrocyte metabolism in osteoarthritis: A detailed analysis by laser capture microdissection', *Arthritis and Rheumatism*, 58(1), pp. 154–163. doi: 10.1002/art.23175.
- Gao, J. *et al.* (2015) 'De-methylation of displacement loop of mitochondrial DNA is associated with increased mitochondrial copy number and nicotinamide adenine dinucleotide subunit 2 expression in colorectal cancer', *Molecular Medicine Reports*. Mol Med Rep, 12(5), pp. 7033–7038. doi: 10.3892/mmr.2015.4256.
- García-Alvarado, F. *et al.* (2018) 'Association between the SNP rs143383 + 104T/C in the GDF5 Gene and the Risk of Knee Osteoarthritis in a Population from Northern Mexico-A Case-Control Study', *Genetic Testing and Molecular Biomarkers*. Genet Test Mol Biomarkers, 22(8), pp. 503–506. doi: 10.1089/gtmb.2018.0142.
- García, I. *et al.* (2017) 'The little big genome: The organization of mitochondrial DNA', *Frontiers in Bioscience - Landmark*, 22(4), pp. 710–721. doi: 10.2741/4511.
- García Meijide, J. A. and Gómez-Reino Carnota, J. J. (2000) 'Fisiopatología de la ciclooxigenasa-

- 1 y ciclooxigenasa-2', *Revista Española de Reumatología*, 27(1), pp. 33–35.
- Garnero, P. *et al.* (2003) 'Urinary type II collagen C-telopeptide levels are increased in patients with rapidly destructive hip osteoarthritis', *Annals of the Rheumatic Diseases*. *Ann Rheum Dis*, 62(10), pp. 939–943. doi: 10.1136/ard.62.10.939.
- Garnero, P. *et al.* (2006) 'Increased urinary type II collagen helical and C telopeptide levels are independently associated with a rapidly destructive hip osteoarthritis', *Annals of the Rheumatic Diseases*. *Ann Rheum Dis*, 65(12), pp. 1639–1644. doi: 10.1136/ard.2006.052621.
- de Girolamo, L. *et al.* (2016) 'Regenerative approaches for the treatment of early OA', *Knee Surgery, Sports Traumatology, Arthroscopy*. Springer Berlin Heidelberg, 24(6), pp. 1826–1835. doi: 10.1007/s00167-016-4125-y.
- Godziuk, K. *et al.* (2018) 'The impact of sarcopenic obesity on knee and hip osteoarthritis: A scoping review', *BMC Musculoskeletal Disorders*, 19(1). doi: 10.1186/s12891-018-2175-7.
- Gorman, G. S. *et al.* (2016) 'Mitochondrial diseases', *Nature Reviews Disease Primers*. Macmillan Publishers Limited, 2, pp. 1–23. doi: 10.1038/nrdp.2016.80.
- Gourlay, M. L. *et al.* (2015) 'Subchondral insufficiency fracture of the knee: A non-traumatic injury with prolonged recovery time', *BMJ Case Reports*. *BMJ Case Rep*, 2015. doi: 10.1136/bcr-2015-209399.
- Greber, B. J. and Ban, N. (2016) 'Structure and Function of the Mitochondrial Ribosome', *Annual Review of Biochemistry*, 85(1), pp. 103–132. doi: 10.1146/annurev-biochem-060815-014343.
- Gredilla, R., Bohr, V. A. and Stevnsner, T. (2010) 'Mitochondrial DNA repair and association with aging - An update', *Experimental Gerontology*. Elsevier Inc., 45(7–8), pp. 478–488. doi: 10.1016/j.exger.2010.01.017.
- Grishko, V. I. *et al.* (2009) 'Diminished mitochondrial DNA integrity and repair capacity in OA chondrocytes', *Osteoarthritis and Cartilage*. *Osteoarthritis Cartilage*, 17(1), pp. 107–113. doi: 10.1016/j.joca.2008.05.009.
- Gu, F. *et al.* (2013) 'Alterations in mitochondrial DNA copy number and the activities of electron transport chain complexes and pyruvate dehydrogenase in the frontal cortex from subjects with autism.', *Translational psychiatry*. *Transl Psychiatry*, 3(9). doi: 10.1038/tp.2013.68.
- Guo, W. *et al.* (2013) 'Mutations in the D-loop region and increased copy number of mitochondrial DNA in human laryngeal squamous cell carcinoma', *Molecular Biology Reports*. *Mol Biol Rep*, 40(1), pp. 13–20. doi: 10.1007/s11033-012-1939-7.
- Guo, Z. *et al.* (2016) 'Identification of sequence polymorphisms in the D-Loop region of mitochondrial DNA as a risk factor for colon cancer', *Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis*. *Mitochondrial DNA A DNA Mapp Seq Anal*, 27(6), pp. 4244–4245. doi: 10.3109/19401736.2014.1003920.
- Hammacher, A. *et al.* (1994) 'Structure-Function analysis of human IL-6: Identification of two distinct regions that are important for receptor binding', *Protein Science*. Wiley-Blackwell, 3(12), pp. 2280–2293. doi: 10.1002/pro.5560031213.
- Han, B. *et al.* (2019) 'Correlation between mitochondrial DNA 4977bp deletion and presbycusis: A system review and meta-analysis', *Medicine (United States)*. *Medicine (Baltimore)*. doi: 10.1097/MD.00000000000016302.

## Bibliografía

- Hardcastle, S. A. *et al.* (2015) 'Osteoarthritis and bone mineral density: are strong bones bad for joints?', *BoneKEy Reports*, 4, p. 624. doi: 10.1038/bonekey.2014.119.
- Harman, D. (1956) 'Aging: a theory based on free radical and radiation chemistry.', *Journal of gerontology*. *J Gerontol*, 11(3), pp. 298–300. doi: 10.1093/geronj/11.3.298.
- Harrell, F. E. (2001) 'General Aspects of Fitting Regression Models', in: Springer, New York, NY, pp. 11–40. doi: 10.1007/978-1-4757-3462-1\_2.
- Harrington, C. T. *et al.* (2013) 'Fundamentals of pyrosequencing', *Archives of Pathology and Laboratory Medicine*, 137(9), pp. 1296–1303. doi: 10.5858/arpa.2012-0463-RA.
- Hart, G. and Fehring, T. (2016) 'Rapidly destructive osteoarthritis can mimic infection', *Arthroplasty Today*. Elsevier, 2(1), pp. 15–18. doi: 10.1016/j.artd.2015.11.003.
- Hashemi Goradel, N. *et al.* (2019) 'Cyclooxygenase-2 in cancer: A review', *Journal of Cellular Physiology*, 234(5), pp. 5683–5699. doi: 10.1002/jcp.27411.
- Head, S. R. *et al.* (2014) 'Library construction for next-generation sequencing: Overviews and challenges', *BioTechniques*, 56(2), pp. 61–77. doi: 10.2144/000114133.
- Heather, J. M. and Chain, B. (2016) 'The sequence of sequencers: The history of sequencing DNA', *Genomics*. The Authors, 107(1), pp. 1–8. doi: 10.1016/j.ygeno.2015.11.003.
- van Helvoort, E. M. *et al.* (2021) 'Baseline clinical characteristics of predicted structural and pain progressors in the IMI-APPROACH knee OA cohort', *RMD Open*, 7(3). doi: 10.1136/rmdopen-2021-001759.
- van Helvoort, E. M. *et al.* (2022) 'Predicted and actual 2-year structural and pain progression in the IMI-APPROACH knee osteoarthritis cohort', *Rheumatology*. Rheumatology (Oxford). doi: 10.1093/rheumatology/keac292.
- Henrotin, Y. E., Bruckner, P. and Pujol, J. P. L. (2003) 'The role of reactive oxygen species in homeostasis and degradation of cartilage', *Osteoarthritis and Cartilage*. W.B. Saunders Ltd, 11(10), pp. 747–755. doi: 10.1016/S1063-4584(03)00150-X.
- Herbst, R. S., Morgensztern, D. and Boshoff, C. (2018) 'The biology and management of non-small cell lung cancer', *Nature*. Nature Publishing Group, 553(7689), pp. 446–454. doi: 10.1038/nature25183.
- Hernández-Beltrán, N., Moreno, C. B. and Gutiérrez-álvarez, Á. M. (2013) 'El papel de la mitocondria en el dolor de la neuropatía diabética', *Endocrinología y Nutrición*, pp. 25–32. doi: 10.1016/j.endonu.2012.03.005.
- Hippisley-Cox, J. *et al.* (2008) 'Predicting cardiovascular risk in England and Wales: Prospective derivation and validation of QRISK2', *BMJ*. *BMJ*, 336(7659), pp. 1475–1482. doi: 10.1136/bmj.39609.449676.25.
- Hochberg, M. C. (2015) 'Serious joint-related adverse events in randomized controlled trials of anti-nerve growth factor monoclonal antibodies', *Osteoarthritis and Cartilage*. Elsevier Ltd, 23, pp. S18–S21. doi: 10.1016/j.joca.2014.10.005.
- Hochberg, M. C. *et al.* (2016) 'When is osteonecrosis not osteonecrosis?: adjudication of reported serious adverse joint events in the tanezumab clinical development program', *Arthritis and Rheumatology*. *Arthritis Rheumatol*, 68(2), pp. 382–391. doi: 10.1002/art.39492.
- Hofhaus, G. *et al.* (1996) 'Respiration and growth defects in transmitochondrial cell lines carrying the 11778 mutation associated with Leber's hereditary optic neuropathy', *Journal of*



*Biological Chemistry*. Elsevier, 271(22), pp. 13155–13161. doi: 10.1074/jbc.271.22.13155.

Hood, L. and Rowen, L. (2013) 'The human genome project: Big science transforms biology and medicine', *Genome Medicine*, 5(9), p. 1. doi: 10.1186/gm483.

Hudson, G., Panoutsopoulou, K., *et al.* (2013) 'No evidence of an association between mitochondrial DNA variants and osteoarthritis in 7393 cases and 5122 controls', *Annals of the Rheumatic Diseases*. *Ann Rheum Dis*, 72(1), pp. 136–139. doi: 10.1136/annrheumdis-2012-201932.

Hudson, G., Nalls, M., *et al.* (2013) 'Two-stage association study and meta-analysis of mitochondrial DNA variants in Parkinson disease', *Neurology*, 80(22), pp. 2042–2048. doi: 10.1212/WNL.0b013e318294b434.

Hwang, H. S. and Kim, H. A. (2015) 'Chondrocyte apoptosis in the pathogenesis of osteoarthritis', *International Journal of Molecular Sciences*, 16(11), pp. 26035–26054. doi: 10.3390/ijms161125943.

Hyde, C. L. *et al.* (2016) 'Identification of 15 genetic loci associated with risk of major depression in individuals of European descent', *Nature Genetics*. Nature Publishing Group, 48(9), pp. 1031–1036. doi: 10.1038/ng.3623.

Infantino, V. *et al.* (2011) 'Impairment of methyl cycle affects mitochondrial methyl availability and glutathione level in Down's syndrome', *Molecular Genetics and Metabolism*. *Mol Genet Metab*, 102(3), pp. 378–382. doi: 10.1016/j.ymgme.2010.11.166.

Jacobs, C. A. *et al.* (2020) 'Rapid Progression of Knee Pain and Osteoarthritis Biomarkers Greatest for Patients with Combined Obesity and Depression: Data from the Osteoarthritis Initiative', *Cartilage*, 11(1), pp. 38–46. doi: 10.1177/1947603518777577.

Jain, M. *et al.* (2018) 'Nanopore sequencing and assembly of a human genome with ultra-long reads', *Nature Biotechnology*, 36(4), pp. 338–345. doi: 10.1038/nbt.4060.

Jamshidi, A. *et al.* (2021) 'Machine Learning–Based Individualized Survival Prediction Model for Total Knee Replacement in Osteoarthritis: Data From the Osteoarthritis Initiative', *Arthritis Care and Research*. *Arthritis Care Res (Hoboken)*, 73(10), pp. 1518–1527. doi: 10.1002/acr.24601.

Jamshidi, A., Pelletier, J. P. and Martel-Pelletier, J. (2019) 'Machine-learning-based patient-specific prediction models for knee osteoarthritis', *Nature Reviews Rheumatology*. *Nat Rev Rheumatol*, pp. 49–60. doi: 10.1038/s41584-018-0130-5.

Janssen, B. G. *et al.* (2015) 'Placental mitochondrial methylation and exposure to airborne particulate matter in the early life environment: An ENVIRONAGE birth cohort study', *Epigenetics*. *Epigenetics*, 10(6), pp. 536–544. doi: 10.1080/15592294.2015.1048412.

Jia, B. *et al.* (2021) 'Correlation between growth differentiation factor 5 (rs143383) gene polymorphism and knee osteoarthritis: an updated systematic review and meta-analysis', *Journal of Orthopaedic Surgery and Research*. *Journal of Orthopaedic Surgery and Research*, 16(1), pp. 1–12. doi: 10.1186/s13018-021-02269-w.

Jiang, D. *et al.* (2016) 'Association between GDF5 +104T/C polymorphism and knee osteoarthritis in Caucasian and Asian populations: A meta-analysis based on case-control studies', *Journal of Orthopaedic Surgery and Research*. *Journal of Orthopaedic Surgery and Research*, 11(1), pp. 1–7. doi: 10.1186/s13018-016-0436-4.

Johnson, V. L. and Hunter, D. J. (2014) 'The epidemiology of osteoarthritis', *Best Practice and*

## Bibliografía

Research: *Clinical Rheumatology*. Elsevier Ltd, 28(1), pp. 5–15. doi: 10.1016/j.berh.2014.01.004.

Joseph, G. B. *et al.* (2018) 'Tool for osteoarthritis risk prediction (TOARP) over 8 years using baseline clinical data, X-ray, and MRI: Data from the osteoarthritis initiative.', *Journal of magnetic resonance imaging : JMRI*. John Wiley & Sons, Ltd, 47(6), pp. 1517–1526. doi: 10.1002/jmri.25892.

Kaneko, S. *et al.* (2000) 'Interleukin-6 and interleukin-8 levels in serum and synovial fluid of patients with osteoarthritis', *Cytokines, Cellular and Molecular Therapy*. Cytokines Cell Mol Ther, 6(2), pp. 71–79. doi: 10.1080/13684730050515796.

Karki, R. *et al.* (2015) 'Defining "mutation" and "polymorphism" in the era of personal genomics', *BMC Medical Genomics*. doi: 10.1186/s12920-015-0115-z.

Kawata, M. *et al.* (2017) 'Different regulation of limb development by p63 transcript variants', *PLoS ONE*, 12(3), pp. 1–19. doi: 10.1371/journal.pone.0174122.

Kazak, L., Reyes, A. and Holt, I. J. (2012) 'Minimizing the damage: Repair pathways keep mitochondrial DNA intact', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 13(10), pp. 659–671. doi: 10.1038/nrm3439.

Kellgren, J. H. and Lawrence, J. S. (1957) 'Radiological assessment of osteo-arthritis.', *Annals of the rheumatic diseases*, 16(4), pp. 494–502. doi: 10.1136/ard.16.4.494.

Kerkhof, H. J. M. *et al.* (2010) 'A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22', *Arthritis and Rheumatism*, 62(2), pp. 499–510. doi: 10.1002/art.27184.

Kerkhof, H. J. M. *et al.* (2014) 'Prediction model for knee osteoarthritis incidence, including clinical, genetic and biochemical risk factors', *Annals of the Rheumatic Diseases*. BMJ Publishing Group Ltd, 73(12), pp. 2116–2121. doi: 10.1136/annrheumdis-2013-203620.

Kim, D. *et al.* (2016) 'Fis1 depletion in osteoarthritis impairs chondrocyte survival and peroxisomal and lysosomal function', *Journal of Molecular Medicine*. J Mol Med (Berl), 94(12), pp. 1373–1384. doi: 10.1007/s00109-016-1445-9.

Kim, H. and Blanco, F. (2007) 'Cell Death and Apoptosis in Osteoarthritic Cartilage', *Current Drug Targets*, 8(2), pp. 333–345. doi: 10.2174/138945007779940025.

Kloppenborg, M. and Berenbaum, F. (2020) 'Osteoarthritis year in review 2019: epidemiology and therapy', *Osteoarthritis and Cartilage*. Elsevier Ltd, 28(3), pp. 242–248. doi: 10.1016/j.joca.2020.01.002.

Kofler, B. *et al.* (2009) 'Mitochondrial DNA haplogroup T is associated with coronary artery disease and diabetic retinopathy: A case control study', *BMC Medical Genetics*, 10. doi: 10.1186/1471-2350-10-35.

Kolasinski, S. L. *et al.* (2020) '2019 American College of Rheumatology/Arthritis Foundation Guideline for the Management of Osteoarthritis of the Hand, Hip, and Knee', *Arthritis Care and Research*, 72(2), pp. 149–162. doi: 10.1002/acr.24131.

Koo, B. S. *et al.* (2019) 'Association of Asian mitochondrial DNA haplogroup B with new development of knee osteoarthritis in Koreans', *International Journal of Rheumatic Diseases*. Int J Rheum Dis, 22(3), pp. 411–416. doi: 10.1111/1756-185X.13453.

Kopinski, P. K. *et al.* (2021) 'Mitochondrial DNA variation and cancer', *Nature Reviews Cancer*. Springer US, 21(7), pp. 431–445. doi: 10.1038/s41568-021-00358-w.

- Kozin, M. S. *et al.* (2018) 'Variants of mitochondrial genome and risk of multiple sclerosis development in Russians', *Acta Naturae*, 10(4), pp. 79–86. doi: 10.32607/20758251-2018-10-4-79-86.
- Kraus, V. B. *et al.* (2015) 'Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use', *Osteoarthritis and Cartilage*, 23(8), pp. 1233–1241. doi: 10.1016/j.joca.2015.03.036.
- Ku, C. S. *et al.* (2010) 'The pursuit of genome-wide association studies: Where are we now', *Journal of Human Genetics*. Nature Publishing Group, 55(4), pp. 195–206. doi: 10.1038/jhg.2010.19.
- Kulkarni, K. *et al.* (2016) 'Obesity and osteoarthritis', *Maturitas*. Maturitas, pp. 22–28. doi: 10.1016/j.maturitas.2016.04.006.
- Kulm, S. *et al.* (2022) 'Characterization of Genetic Risk of End-Stage Knee Osteoarthritis Treated with Total Knee Arthroplasty: A Genome-Wide Association Study.', *The Journal of bone and joint surgery. American volume*, pp. 1–8. doi: 10.2106/JBJS.22.00364.
- Kwak, S. H. and Park, K. S. (2016) 'Role of mitochondrial DNA variation in the pathogenesis of diabetes mellitus', *Frontiers in Bioscience - Landmark*. Frontiers in Bioscience, 21(6), pp. 1151–1167. doi: 10.2741/4447.
- Kwon, J. M. and Goate, A. M. (2000) 'The candidate gene approach', *Alcohol Research and Health*, 24(3), pp. 164–168.
- Läer, K. *et al.* (2014) 'Mitochondrial deoxyribonucleic acid may play a role in a subset of sudden infant death syndrome cases', *Acta Paediatrica, International Journal of Paediatrics*, 103(7), pp. 775–779. doi: 10.1111/apa.12609.
- Lake, N. J. *et al.* (2016) 'Leigh syndrome: One disorder, more than 75 monogenic causes', *Annals of Neurology*. Ann Neurol, pp. 190–203. doi: 10.1002/ana.24551.
- Latorre-Pellicer, A. *et al.* (2016) 'Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing', *Nature*, 535(7613), pp. 561–565. doi: 10.1038/nature18618.
- Lawrence, R. C. *et al.* (2008) 'Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II', *Arthritis and Rheumatism*, 58(1), pp. 26–35. doi: 10.1002/art.23176.
- Lechuga-Vieco, A. V., Justo-Méndez, R. and Enríquez, J. A. (2021) 'Not all mitochondrial DNAs are made equal and the nucleus knows it', *IUBMB Life*, pp. 511–529. doi: 10.1002/iub.2434.
- Lee, A. R. *et al.* (2022) 'Mitochondrial Transplantation Ameliorates the Development and Progression of Osteoarthritis', *Immune Network*. Immune Netw, 22(2). doi: 10.4110/in.2022.22.e14.
- Lee, H. F. *et al.* (2007) 'The neurological evolution of Pearson syndrome: Case report and literature review', *European Journal of Paediatric Neurology*. Eur J Paediatr Neurol, 11(4), pp. 208–214. doi: 10.1016/j.ejpn.2006.12.008.
- Lee, M. *et al.* (2015) 'GNL3 and SKA3 are novel prostate cancer metastasis susceptibility genes', *Clinical and Experimental Metastasis*. Springer, 32(8), pp. 769–782. doi: 10.1007/s10585-015-9745-y.
- Lequesne, M. and Amouroux, J. (1970) 'La coxarthrose destructrice rapide.', *La Presse medicale*, 78(32), pp. 1435–1439.

## Bibliografía

- Leung, K. *et al.* (2020) 'Prediction of total knee replacement and diagnosis of osteoarthritis by using deep learning on knee radiographs: Data from the osteoarthritis initiative', *Radiology*. *Radiology*, 296(3), pp. 584–593. doi: 10.1148/radiol.2020192091.
- Li, H. and Homer, N. (2010) 'A survey of sequence alignment algorithms for next-generation sequencing', *Briefings in Bioinformatics*. *Brief Bioinform*, pp. 473–483. doi: 10.1093/bib/bbq015.
- Li, X. *et al.* (2020) 'Association between mitochondrial DNA D-loop region polymorphisms and endometriosis in a Chinese population', *Journal of Assisted Reproduction and Genetics*. *J Assist Reprod Genet*, 37(9), pp. 2171–2179. doi: 10.1007/s10815-020-01853-z.
- Li, Y. *et al.* (2015) 'Association of genes, pathways, and haplogroups of the mitochondrial genome with the risk of colorectal cancer: The multiethnic cohort', *PLoS ONE*. *Public Library of Science*, 10(9), p. e0136796. doi: 10.1371/journal.pone.0136796.
- Liao, W. Q. *et al.* (2008) 'Novel mutations of mitochondrial DNA associated with type 2 diabetes in Chinese Han population', *Tohoku Journal of Experimental Medicine*, 215(4), pp. 377–384. doi: 10.1620/tjem.215.377.
- Liao, Y. *et al.* (2022) 'Interleukin-6 signaling mediates cartilage degradation and pain in posttraumatic osteoarthritis in a sex-specific manner', *Science Signaling*. *Sci Signal*, 15(744). doi: 10.1126/scisignal.abn7082.
- Lillie, P. J. *et al.* (2020) 'Novel coronavirus disease (Covid-19): The first two patients in the UK with person to person transmission', *Journal of Infection*, 80(5), pp. 578–606. doi: 10.1016/j.jinf.2020.02.020.
- Litwic, A. *et al.* (2013) 'Epidemiology and burden of osteoarthritis', *British Medical Bulletin*, 105(1), pp. 185–199. doi: 10.1093/bmb/lds038.
- Liu, B. *et al.* (2018) 'Common variants in the GNL3 contribute to the increasing risk of knee osteoarthritis in Han Chinese population', *Scientific Reports*. Springer US, 8(1), p. 9610. doi: 10.1038/s41598-018-27971-4.
- Liu, Q. *et al.* (2022) 'Prediction models for the risk of total knee replacement: development and validation using data from multicentre cohort studies', *The Lancet Rheumatology*. *Lancet Rheumatol*, 4(2), pp. e125–e134. doi: 10.1016/S2665-9913(21)00324-6.
- Loeser, R. F. *et al.* (2012) 'Osteoarthritis: A disease of the joint as an organ', *Arthritis and Rheumatism*, 64(6), pp. 1697–1707. doi: 10.1002/art.34453.
- Lohmander, L. S. *et al.* (2007) 'The long-term consequence of anterior cruciate ligament and meniscus injuries: Osteoarthritis', *American Journal of Sports Medicine*, 35(10), pp. 1756–1769. doi: 10.1177/0363546507307396.
- Lotz, M. and Loeser, R. F. (2012) 'Effects of aging on articular cartilage homeostasis', *Bone*, 51(2), pp. 241–248. doi: 10.1016/j.bone.2012.03.023.
- Lourido, L. *et al.* (2021) 'A clinical model including protein biomarkers predicts radiographic knee osteoarthritis: a prospective study using data from the Osteoarthritis Initiative', *Osteoarthritis and Cartilage*. *Osteoarthritis Cartilage*, 29(8), pp. 1147–1154. doi: 10.1016/j.joca.2021.04.011.
- Loza, E. *et al.* (2009) 'Economic burden of knee and hip osteoarthritis in Spain', *Arthritis Care and Research*, 61(2), pp. 158–165. doi: 10.1002/art.24214.
- Luo, S. *et al.* (2018) 'Biparental inheritance of mitochondrial DNA in humans', *Proceedings of*

- the *National Academy of Sciences of the United States of America*, 115(51), pp. 13039–13044. doi: 10.1073/pnas.1810946115.
- Luo, Y. *et al.* (2019) 'The mitochondrial DNA 4977-bp deletion and copy number alteration in Han Chinese samples with uterine fibroids', *Annals of Human Genetics*. *Ann Hum Genet*, 83(4), pp. 220–230. doi: 10.1111/ahg.12303.
- Madamanchi, N. R. and Runge, M. S. (2007) 'Mitochondrial dysfunction in atherosclerosis', *Circulation Research*. *Circ Res*, 100(4), pp. 460–473. doi: 10.1161/01.RES.0000258450.44413.96.
- Mahad, D., Lassmann, H. and Turnbull, D. (2008) 'Review: Mitochondria and disease progression in multiple sclerosis', *Neuropathology and Applied Neurobiology*, 34(6), pp. 577–589. doi: 10.1111/j.1365-2990.2008.00987.x.
- Mancuso, M. *et al.* (2011) 'May mitochondrial eve and mitochondrial haplogroups play a role in neurodegeneration and alzheimer's disease?', *International Journal of Alzheimer's Disease*, 2011. doi: 10.4061/2011/709061.
- Maneiro, E. *et al.* (2003) 'Mitochondrial respiratory activity is altered in osteoarthritic human articular chondrocytes', *Arthritis and Rheumatism*, 48(3), pp. 700–708. doi: 10.1002/art.10837.
- Mao, X. *et al.* (2020) 'Mitochondria: Potential Targets for Osteoarthritis', *Frontiers in Medicine*. *Front Med (Lausanne)*, 7(November), pp. 1–17. doi: 10.3389/fmed.2020.581402.
- Marom, S., Friger, M. and Mishmar, D. (2017) 'MtDNA meta-analysis reveals both phenotype specificity and allele heterogeneity: A model for differential association', *Scientific Reports*. *Sci Rep*, 7. doi: 10.1038/srep43449.
- Martel-Pelletier, J. *et al.* (2016) 'Osteoarthritis', *Nature Reviews Disease Primers*, 2. doi: 10.1038/nrdp.2016.72.
- Martín-Fernández, B. and Gredilla, R. (2016) 'Mitochondria and oxidative stress in heart aging', *Age*. *AGE*, 38(4), pp. 225–238. doi: 10.1007/s11357-016-9933-y.
- Martincorena, I. and Campbell, P. J. (2015) 'Somatic mutation in cancer and normal cells', *Science*, 349(6255), pp. 1483–1489. doi: 10.1126/science.aab4082.
- Maruszak, A. *et al.* (2009) 'Mitochondrial haplogroup H and Alzheimer's disease-Is there a connection?', *Neurobiology of Aging*, 30(11), pp. 1749–1755. doi: 10.1016/j.neurobiolaging.2008.01.004.
- Mavrogenis, A. F. *et al.* (2015) 'Rapid destructive arthritis of the hip revisited', *European Journal of Orthopaedic Surgery and Traumatology*. *Eur J Orthop Surg Traumatol*, 25(7), pp. 1115–1120. doi: 10.1007/s00590-015-1676-4.
- Maxam, A. M. and Gilbert, W. (1977) 'A new method for sequencing DNA', *Proceedings of the National Academy of Sciences of the United States of America*. *National Academy of Sciences*, 74(2), pp. 560–564. doi: 10.1073/pnas.74.2.560.
- Maynard, S. *et al.* (2015) 'DNA damage, DNA repair, aging, and neurodegeneration', *Cold Spring Harbor Perspectives in Medicine*, 5(10), pp. 1–18. doi: 10.1101/cshperspect.a025130.
- Mello, T., Simeone, I. and Galli, A. (2019) 'Mito-nuclear communication in hepatocellular carcinoma metabolic rewiring', *Cells*, 8(5), p. 417. doi: 10.3390/cells8050417.
- Messier, S. P. *et al.* (2009) 'The Intensive Diet and Exercise for Arthritis (IDEA) trial: Design and rationale', *BMC Musculoskeletal Disorders*. *BioMed Central*, 10(1), p. 93. doi: 10.1186/1471-

## Bibliografía

2474-10-93.

Michikawa, Y. *et al.* (1999) 'Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication', *Science*, 286(5440), pp. 774–779. doi: 10.1126/science.286.5440.774.

Miettinen, P. J. *et al.* (1999) 'Epidermal growth factor receptor function is necessary for normal craniofacial development and palate closure', *Nature Genetics*. Nat Genet, 22(1), pp. 69–73. doi: 10.1038/8773.

Mishmar, D. *et al.* (2003) 'Natural selection shaped regional mtDNA variation in humans', *Proceedings of the National Academy of Sciences of the United States of America*, 100(1), pp. 171–176. doi: 10.1073/pnas.0136972100.

Misra, D. *et al.* (2019) 'Risk of Knee Osteoarthritis With Obesity, Sarcopenic Obesity, and Sarcopenia', *Arthritis and Rheumatology*. NIH Public Access, 71(2), pp. 232–237. doi: 10.1002/art.40692.

Mitrovic, D. R. and Riera, H. (1992) 'Synovial, articular cartilage and bone changes in rapidly destructive arthropathy (osteoarthritis) of the hip', *Rheumatology International*. Rheumatol Int, 12(1), pp. 17–22. doi: 10.1007/BF00246872.

Mobasheri, A. *et al.* (2002) 'Glucose transport and metabolism in chondrocytes: A key to understanding chondrogenesis, skeletal development and cartilage degradation in osteoarthritis', *Histology and Histopathology*, 17(4), pp. 1239–1267. doi: 10.14670/HH-17.1239.

Morales-Ivorra, I. *et al.* (2018) 'Osteoarthritis and the mediterranean diet: A systematic review', *Nutrients*, 10(8), pp. 1–11. doi: 10.3390/nu10081030.

Mousavizadeh, K. *et al.* (2015) 'Usage of mitochondrial D-loop variation to predict risk for Huntington disease', *Mitochondrial DNA*, 26(4), pp. 579–582. doi: 10.3109/19401736.2013.878902.

Mullis, K. *et al.* (1986) 'Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction', *Cold Spring Harbor Symposia on Quantitative Biology*. Cold Spring Harb Symp Quant Biol, 51(1), pp. 263–273. doi: 10.1101/sqb.1986.051.01.032.

Murakami, H. *et al.* (2002) 'Polymorphisms in control region of mtDNA relates to individual differences in endurance capacity or trainability', *Japanese Journal of Physiology*, 52(3), pp. 247–256. doi: 10.2170/jjphysiol.52.247.

Murphy, E. *et al.* (2016) *Mitochondrial Function, Biology, and Role in Disease: A Scientific Statement from the American Heart Association*, *Circulation Research*. doi: 10.1161/RES.000000000000104.

Musumeci, G. (2016) 'The effect of mechanical loading on articular cartilage', *Journal of Functional Morphology and Kinesiology*. Multidisciplinary Digital Publishing Institute, 1(2), pp. 154–161. doi: 10.3390/jfmk1020154.

Naik, P. P., Birbrair, A. and Bhutia, S. K. (2019) 'Mitophagy-driven metabolic switch reprograms stem cell fate', *Cellular and Molecular Life Sciences*. Cell Mol Life Sci, pp. 27–43. doi: 10.1007/s00018-018-2922-9.

Nandakumar, P. *et al.* (2021) 'Nuclear genome-wide associations with mitochondrial heteroplasmy', *Science Advances*, 7(12), pp. 1–10. doi: 10.1126/sciadv.abe7520.

Narayanan, R. *et al.* (2007) 'Complement Factor H Polymorphism in Age-Related Macular

- Degeneration', *Ophthalmology*, 114(7), pp. 1327–1331. doi: 10.1016/j.ophtha.2006.10.035.
- Navaglia, F. *et al.* (2006) 'Mitochondrial DNA D-loop in pancreatic cancer: Somatic mutations are epiphenomena while the germline 16519T variant worsens metabolism and outcome', *American Journal of Clinical Pathology*, 126(4), pp. 593–601. doi: 10.1309/GQFCCJMH5KHNWX73.
- Neogi, T. *et al.* (2012) 'Identifying trajectories of medial joint-space width loss and associated risk factors', *Osteoarthritis and Cartilage*. Elsevier Ltd, 20(2012), pp. S182–S183. doi: 10.1016/j.joca.2012.02.281.
- Neogi, T. (2016) 'Gout', *Annals of Internal Medicine*, 165(1), pp. ITC1–ITC15. doi: 10.7326/AITC201607050.
- Nevitt, M. C. *et al.* (2001) 'The effect of estrogen plus progestin on knee symptoms and related disability in postmenopausal women: The heart and estrogen/progestin replacement study, a randomized, double-blind, placebo-controlled trial', *Arthritis and Rheumatism*, 44(4), pp. 811–818. doi: 10.1002/1529-0131(200104)44:4<811::AID-ANR137>3.0.CO;2-F.
- Nicholls, T. J. and Minczuk, M. (2014) 'In D-loop: 40 years of mitochondrial 7S DNA', *Experimental Gerontology*. Elsevier Inc., 56, pp. 175–181. doi: 10.1016/j.exger.2014.03.027.
- Niemi, A. K. *et al.* (2003) 'Mitochondrial DNA polymorphisms associated with longevity in a Finnish population', *Human Genetics*, 112(1), pp. 29–33. doi: 10.1007/s00439-002-0843-y.
- Nishigaki, Y., Fuku, N. and Tanaka, M. (2010) 'Mitochondrial haplogroups associated with lifestyle-related diseases and longevity in the Japanese population', *Geriatrics and Gerontology International*, 10(SUPPL. 1). doi: 10.1111/j.1447-0594.2010.00599.x.
- Nissanka, N., Minczuk, M. and Moraes, C. T. (2019) 'Mechanisms of Mitochondrial DNA Deletion Formation', *Trends in Genetics*. Elsevier Ltd, 35(3), pp. 235–244. doi: 10.1016/j.tig.2019.01.001.
- Nyrén, P. (2015) 'The history of Pyrosequencing®', *Methods in Molecular Biology*, 1315, pp. 3–15. doi: 10.1007/978-1-4939-2715-9\_1.
- Ojala, D., Montoya, J. and Attardi, G. (1981) 'TRNA punctuation model of RNA processing in human mitochondria', *Nature*, 290(5806), pp. 470–474. doi: 10.1038/290470a0.
- Oka, H. *et al.* (2010) 'Normal and threshold values of radiographic parameters for knee osteoarthritis using a computer-assisted measuring system (KOACAD): The ROAD study', *Journal of Orthopaedic Science*, 15(6), pp. 781–789. doi: 10.1007/s00776-010-1545-2.
- Oreiro-Villar, N. *et al.* (2022a) 'PROCOAC (PROspective COhort of A Coruña) description: Spanish prospective cohort to study osteoarthritis', *Reumatología Clínica (English Edition)*. *Reumatol Clin (Engl Ed)*, 18(2), pp. 100–104. doi: 10.1016/j.reumae.2020.08.004.
- Oreiro-Villar, N. *et al.* (2022b) 'PROCOAC (PROspective COhort of A Coruña) description: Spanish prospective cohort to study osteoarthritis', *Reumatología Clínica (English Edition)*. *Reumatol Clin (Engl Ed)*, 18(2), pp. 100–104. doi: 10.1016/j.reumae.2020.08.004.
- Ortiz, R. G. *et al.* (1993) 'Variable Retinal and Neurologic Manifestations in Patients Harboring the Mitochondrial DNA 8993 Mutation', *Archives of Ophthalmology*. *Arch Ophthalmol*, 111(11), pp. 1525–1530. doi: 10.1001/archophth.1993.01090110091031.
- Pagliarini, D. J. *et al.* (2008) 'A Mitochondrial Protein Compendium Elucidates Complex I Disease Biology', *Cell*, 134(1), pp. 112–123. doi: 10.1016/j.cell.2008.06.016.

## Bibliografia

- Pakula, A. (2019) 'Mitochondrial DNA in evolution and disease', *Methods Molecular Biology*, 176(5), pp. 139–148. doi: 10.1038/nature18902.Mitochondrial.
- Palazzo, C. *et al.* (2016) 'Risk factors and burden of osteoarthritis', *Annals of Physical and Rehabilitation Medicine*, 59(3), pp. 134–138. doi: 10.1016/j.rehab.2016.01.006.
- Pan, F. *et al.* (2014) 'Association between GDF5 rs143383 polymorphism and knee osteoarthritis: An updated meta-analysis based on 23,995 subjects', *BMC Musculoskeletal Disorders*, 15(1), pp. 1–11. doi: 10.1186/1471-2474-15-404.
- Panoutsopoulou, K. and Zeggini, E. (2013) 'Advances in osteoarthritis genetics', *Journal of Medical Genetics*, 50(11), pp. 715–724. doi: 10.1136/jmedgenet-2013-101754.
- Papathanasiou, I. *et al.* (2019) 'DNA methylation regulates miR-140-5p and miR-146a expression in osteoarthritis', *Life Sciences*. Life Sci, 228, pp. 274–284. doi: 10.1016/j.lfs.2019.05.018.
- Pareek, C. S., Smoczynski, R. and Tretyn, A. (2011) 'Sequencing technologies and genome sequencing', *Journal of Applied Genetics*, 52(4), pp. 413–435. doi: 10.1007/s13353-011-0057-x.
- Park, J. Y., Pillinger, M. H. and Abramson, S. B. (2006) 'Prostaglandin E2 synthesis and secretion: The role of PGE2 synthases', *Clinical Immunology*. Academic Press, 119(3), pp. 229–240. doi: 10.1016/j.clim.2006.01.016.
- Pavlakakis, S. G. *et al.* (1984) 'Mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes: A distinctive clinical syndrome', *Annals of Neurology*. Ann Neurol, 16(4), pp. 481–488. doi: 10.1002/ana.410160409.
- Pelletier, J. P. *et al.* (2001) 'Chondrocyte death in experimental osteoarthritis is mediated by MEK 1/2 and p38 pathways: Role of cyclooxygenase-2 and inducible nitric oxide synthase', *Journal of Rheumatology*, 28(11), pp. 2509–2519. Available at: www.jrheum.org (Accessed: 24 June 2022).
- Pereira, D. *et al.* (2011) 'The effect of osteoarthritis definition on prevalence and incidence estimates: A systematic review', *Osteoarthritis and Cartilage*. Elsevier Ltd, 19(11), pp. 1270–1285. doi: 10.1016/j.joca.2011.08.009.
- Perel, P. *et al.* (2012) 'Predicting early death in patients with traumatic bleeding: development and validation of prognostic model.', *BMJ (Clinical research ed.)*. BMJ, 345. doi: 10.1136/bmj.e5166.
- Perry, T. A. *et al.* (2020) 'Occupation and risk of knee osteoarthritis and knee replacement: A longitudinal, multiple-cohort study', *Seminars in Arthritis and Rheumatism*. Semin Arthritis Rheum, 50(5), pp. 1006–1014. doi: 10.1016/j.semarthrit.2020.08.003.
- Persad, P. J. *et al.* (2017) 'Joint Analysis of Nuclear and Mitochondrial Variants in Age-Related Macular Degeneration Identifies Novel Loci TRPM1 and ABHD2/RLBP1', *Investigative Ophthalmology and Visual Science*, 58(10), pp. 4027–4038. doi: 10.1167/iovs.17-21734.
- Peshkova, M. *et al.* (2022) 'Gender-Related Aspects in Osteoarthritis Development and Progression: A Review', *International Journal of Molecular Sciences*. Multidisciplinary Digital Publishing Institute (MDPI). doi: 10.3390/ijms23052767.
- Petros, J. A. *et al.* (2005) 'MtDNA mutations increase tumorigenicity in prostate cancer', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 102(3), pp. 719–724. doi: 10.1073/pnas.0408894102.
- Phillips, N. R., Simpkins, J. W. and Roby, R. K. (2014) 'Mitochondrial DNA deletions in



- Alzheimer's brains: A review', *Alzheimer's and Dementia*. *Alzheimers Dement*, pp. 393–400. doi: 10.1016/j.jalz.2013.04.508.
- Picard, M. *et al.* (2014) 'Progressive increase in mtDNA 3243A>G heteroplasmy causes abrupt transcriptional reprogramming', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 111(38), pp. E4033–E4042. doi: 10.1073/pnas.1414028111.
- Pickett, S. J. *et al.* (2022) 'Interactions between nuclear and mitochondrial SNPs and Parkinson's disease risk', *Mitochondrion*. Elsevier B.V., 63, pp. 85–88. doi: 10.1016/j.mito.2022.02.002.
- Pinós, T. *et al.* (2012) 'Are mitochondrial haplogroups associated with extreme longevity? A study on a Spanish cohort', *Age*, 34(1), pp. 227–233. doi: 10.1007/s11357-011-9209-5.
- Pivec, R. *et al.* (2013) 'Differentiation, diagnosis, and treatment of osteoarthritis, osteonecrosis, and rapidly progressive osteoarthritis', *Orthopedics*, 36(2), pp. 118–125. doi: 10.3928/01477447-20130122-04.
- Von Porat, A., Roos, E. M. and Roos, H. (2004) 'High prevalence of osteoarthritis 14 years after an anterior cruciate ligament tear in male soccer players: A study of radiographic and patient relevant outcomes', *Annals of the Rheumatic Diseases*, 63(3), pp. 269–273. doi: 10.1136/ard.2003.008136.
- Porée, B. *et al.* (2008) 'Interleukin-6 (IL-6) and/or soluble IL-6 receptor down-regulation of human type II collagen gene expression in articular chondrocytes requires a decrease of Sp1-Sp3 ratio and of the binding activity of both factors to the COL2A1 promoter', *Journal of Biological Chemistry*. *J Biol Chem*, 283(8), pp. 4850–4865. doi: 10.1074/jbc.M706387200.
- Postel, M. and Kerboull, M. (1970) 'Total prosthetic replacement in rapidly destructive arthrosis of the hip joint.', *Clinical Orthopaedics and Related Research*, 72, pp. 138–144. doi: 10.1097/00003086-197009000-00016.
- Potapov, V. and Ong, J. L. (2017) 'Examining sources of error in PCR by single-molecule sequencing', *PLoS ONE*, 12(1), pp. 1–19. doi: 10.1371/journal.pone.0169774.
- Poulton, J. *et al.* (1998) 'A common mitochondrial DNA variant is associated with insulin resistance in adult life', *Diabetologia*. *Diabetologia*, 41(1), pp. 54–58. doi: 10.1007/s001250050866.
- Pradilla-Vesga, O. E. (2004) 'Ciclooxigenasa 3: La nueva iso-enzima en la familia', *Med UNAB*, 7(21), pp. 181–184.
- Preston, B. D., Albertson, T. M. and Herr, A. J. (2010) 'DNA replication fidelity and cancer', *Seminars in Cancer Biology*. NIH Public Access, pp. 281–293. doi: 10.1016/j.semcancer.2010.10.009.
- Quirós, P. M., Mottis, A. and Auwerx, J. (2016) 'Mitonuclear communication in homeostasis and stress', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 17(4), pp. 213–226. doi: 10.1038/nrm.2016.23.
- Rahmati, M., Mobasheri, A. and Mozafari, M. (2016) 'Inflammatory mediators in osteoarthritis: A critical review of the state-of-the-art, current prospects, and future challenges', *Bone*. *Bone*, pp. 81–90. doi: 10.1016/j.bone.2016.01.019.
- Raine, E. V. A. *et al.* (2012) 'Gene expression analysis reveals HBP1 as a key target for the osteoarthritis susceptibility locus that maps to chromosome 7q22', *Annals of the Rheumatic*

## Bibliografía

*Diseases*, 71(12), pp. 2020–2027. doi: 10.1136/annrheumdis-2012-201304.

Ramos-Louro, P. *et al.* (2022) 'mtDNA haplogroup A enhances the effect of obesity on the risk of knee OA in a Mexican population', *Scientific Reports*, 12(1), p. 5173. doi: 10.1038/s41598-022-09265-y.

Ravi, R. K., Walton, K. and Khosroheidari, M. (2018) 'Miseq: A next generation sequencing platform for genomic analysis', in *Methods in Molecular Biology*, pp. 223–232. doi: 10.1007/978-1-4939-7471-9\_12.

Reed, K. N. *et al.* (2014) 'The role of mitochondrial reactive oxygen species in cartilage matrix destruction', *Molecular and Cellular Biochemistry*. Mol Cell Biochem, 397(1–2), pp. 195–201. doi: 10.1007/s11010-014-2187-z.

Rego-Pérez, I. *et al.* (2008) 'Mitochondrial DNA haplogroups: Role in the prevalence and severity of knee osteoarthritis', *Arthritis and Rheumatism*, 58(8), pp. 2387–2396. doi: 10.1002/art.23659.

Rego-Pérez, I. *et al.* (2010) 'Mitochondrial DNA haplogroups modulate the serum levels of biomarkers in patients with osteoarthritis', *Annals of the Rheumatic Diseases*, 69(5), pp. 910–917. doi: 10.1136/ard.2009.117416.

Rego-Pérez, I. *et al.* (2011) 'Mitochondrial DNA haplogroups and serum levels of proteolytic enzymes in patients with osteoarthritis', *Annals of the Rheumatic Diseases*, 70(4), pp. 646–652. doi: 10.1136/ard.2010.133637.

Rego-Pérez, I. *et al.* (2018) 'Mitochondrial DNA haplogroups associated with MRI-detected structural damage in early knee osteoarthritis', *Osteoarthritis and Cartilage*. W.B. Saunders Ltd, 26(11), pp. 1562–1569. doi: 10.1016/j.joca.2018.06.016.

Rego, I. *et al.* (2008) 'The Use of Single Base Extension (SBE) for the Assessment of European Mitochondrial DNA Haplogroups in Galician Population', *The Open Cell Development & Biology Journal*, 1(1), pp. 10–16. doi: 10.2174/1874085500801010010.

Rego, I. *et al.* (2010) 'Role of European mitochondrial DNA haplogroups in the prevalence of hip osteoarthritis in Galicia, Northern Spain', *Annals of the Rheumatic Diseases*, 69(1), pp. 210–213. doi: 10.1136/ard.2008.105254.

Reynard, L. N. *et al.* (2014) 'CpG methylation regulates allelic expression of GDF5 by modulating binding of SP1 and SP3 repressor proteins to the osteoarthritis susceptibility SNP rs143383', *Human Genetics*, 133(8), pp. 1059–1073. doi: 10.1007/s00439-014-1447-z.

Rice, S. J. *et al.* (2018) 'Identification of a novel, methylation-dependent, RUNX2 regulatory region associated with osteoarthritis risk', *Human Molecular Genetics*, 27(19), pp. 3464–3474. doi: 10.1093/hmg/ddy257.

Richards, M. M. *et al.* (2016) 'Intra-articular treatment of knee osteoarthritis: from anti-inflammatories to products of regenerative medicine', *Physician and Sportsmedicine*. Taylor & Francis, 44(2), pp. 101–108. doi: 10.1080/00913847.2016.1168272.

Riddle, D. L., Stratford, P. W. and Perera, R. A. (2016) 'The incident tibiofemoral osteoarthritis with rapid progression phenotype: development and validation of a prognostic prediction rule', *Osteoarthritis and Cartilage*. W.B. Saunders, 24(12), pp. 2100–2107. doi: 10.1016/j.joca.2016.06.021.

Rim, Y. A., Nam, Y. and Ju, J. H. (2020) 'The role of chondrocyte hypertrophy and senescence in osteoarthritis initiation and progression', *International Journal of Molecular Sciences*, 21(7).

doi: 10.3390/ijms21072358.

- Roberts, S. *et al.* (2011) 'Prospects of stem cell therapy in osteoarthritis', *Regenerative Medicine*, 6(3), pp. 351–366. doi: 10.2217/rme.11.21.
- Robinson, P. D. *et al.* (2018) 'Osteoarthritis and arthroplasty of the hip and knee', *British Journal of Hospital Medicine*, 79(4), pp. C54–C59. doi: 10.12968/hmed.2018.79.4.C54.
- Rodríguez-Violante, M. *et al.* (2010) 'Papel de la función mitocondrial en las enfermedades neurodegenerativas', *Archivos de Neurociencias*, 15(1), pp. 39–46.
- Roemer, F. W. *et al.* (2011) 'Presence of MRI-detected joint effusion and synovitis increases the risk of cartilage loss in knees without osteoarthritis at 30-month follow-up: The MOST study', *Annals of the Rheumatic Diseases*. *Ann Rheum Dis*, 70(10), pp. 1804–1809. doi: 10.1136/ard.2011.150243.
- Roemer, F. W. *et al.* (2015) 'Imaging atlas for eligibility and on-study safety of potential knee adverse events in anti-NGF studies (Part 1)', *Osteoarthritis and Cartilage*. Elsevier Ltd, 23, pp. S22–S42. doi: 10.1016/j.joca.2014.09.015.
- Roger, A. J., Muñoz-Gómez, S. A. and Kamikawa, R. (2017) 'The Origin and Diversification of Mitochondria', *Current Biology*, 27(21), pp. R1177–R1192. doi: 10.1016/j.cub.2017.09.015.
- Rose, G. *et al.* (2001) 'Paradoxes in longevity: Sequence analysis of mtDNA haplogroup J in centenarians', *European Journal of Human Genetics*, 9(9), pp. 701–707. doi: 10.1038/sj.ejhg.5200703.
- Ross, O. A. *et al.* (2001) 'Mitochondrial DNA polymorphism: Its role in longevity of the Irish population', *Experimental Gerontology*. Pergamon, 36(7), pp. 1161–1178. doi: 10.1016/S0531-5565(01)00094-8.
- Rossignol, R. *et al.* (2003) 'Mitochondrial threshold effects', *Biochemical Journal*, 370(3), pp. 751–762. doi: 10.1042/BJ20021594.
- Rowan, A. D. *et al.* (2001) 'Synergistic effects of glycoprotein 130 binding cytokines in combination with interleukin-1 on cartilage collagen breakdown', *Arthritis and Rheumatism*, 44(7), pp. 1620–1632. doi: 10.1002/1529-0131(200107)44:7<1620::AID-ART285>3.0.CO;2-B.
- Ruiz-Pesini, E. *et al.* (2004) 'Effects of Purifying and Adaptive Selection on Regional Variation in Human mtDNA', *Science*, 303(5655), pp. 223–226. doi: 10.1126/science.1088434.
- Ruiz-Romero, C. *et al.* (2009) 'Mitochondrial dysregulation of osteoarthritic human articular chondrocytes analyzed by proteomics: A decrease in mitochondrial superoxide dismutase points to a redox imbalance', *Molecular and Cellular Proteomics*. American Society for Biochemistry and Molecular Biology, 8(1), pp. 172–189. doi: 10.1074/mcp.M800292-MCP200.
- Ruiz, A. A. (2005) 'DMOA : revisión crítica de la evidencia', *Revista española de Reumatología*, 32(1), pp. 2–7.
- Saberi Hosnijeh, F. *et al.* (2018) 'Development of a prediction model for future risk of radiographic hip osteoarthritis', *Osteoarthritis and Cartilage*. W.B. Saunders Ltd, 26(4), pp. 540–546. doi: 10.1016/j.joca.2018.01.015.
- Sagan, L. (1967) 'On the origin of mitosing cells', *Journal of Theoretical Biology*. Academic Press, 14(3), pp. 225–274. doi: 10.1016/0022-5193(67)90079-3.
- Saiki, R. *et al.* (1988) 'Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase', *Science*. American Association for the Advancement of Science, 239(4839),

## Bibliografia

pp. 487–491. doi: 10.1126/science.239.4839.487.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) 'DNA sequencing with chain-terminating inhibitors.', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 74(12), pp. 5463–5467. doi: 10.1073/pnas.74.12.5463.

Santoro, A. *et al.* (2010) 'Evidence for sub-haplogroup H5 of mitochondrial DNA as a risk factor for late onset alzheimer's disease', *PLoS ONE*, 5(8), pp. 1–11. doi: 10.1371/journal.pone.0012037.

Sanyal, T. *et al.* (2018) 'Hypomethylation of mitochondrial D-loop and ND6 with increased mitochondrial DNA copy number in the arsenic-exposed population', *Toxicology*. Toxicology, 408, pp. 54–61. doi: 10.1016/j.tox.2018.06.012.

Sauvanet, C. *et al.* (2010) 'Energetic requirements and bioenergetic modulation of mitochondrial morphology and dynamics', *Seminars in Cell and Developmental Biology*, 21(6), pp. 558–565. doi: 10.1016/j.semcdb.2009.12.006.

Schadt, E. E., Turner, S. and Kasarskis, A. (2010) 'A window into third-generation sequencing', *Human Molecular Genetics*, 19(R2), pp. 227–240. doi: 10.1093/hmg/ddq416.

Schaefer, P. M. *et al.* (2022) 'Combination of common mtDNA variants results in mitochondrial dysfunction and a connective tissue dysregulation', *Proceedings of the National Academy of Sciences of the United States of America*, 119(45), pp. 1–11. doi: 10.1073/pnas.2212417119.

Schett, G. *et al.* (2013) 'Diabetes is an independent predictor for severe osteoarthritis: Results from a longitudinal cohort study', *Diabetes Care*, 36(2), pp. 403–409. doi: 10.2337/dc12-0924.

Schmitz, G. and Ecker, J. (2008) 'The opposing effects of n-3 and n-6 fatty acids', *Progress in Lipid Research*, 47(2), pp. 147–155. doi: 10.1016/j.plipres.2007.12.004.

Scotece, M. *et al.* (2021) 'Mitochondrial DNA impact on joint damaged process in a conplastic mouse model after being surgically induced with osteoarthritis', *Scientific reports*. Nature Publishing Group, 11(1), p. 9112. doi: 10.1038/s41598-021-88083-0.

Scotece, M. *et al.* (2022) 'mtDNA variability determines spontaneous joint aging damage in a conplastic mouse model.', *Aging*. Aging (Albany NY), 14(15), pp. 5966–5983. doi: 10.18632/aging.204153.

Sellam, J. *et al.* (2021) 'The DIGICOD cohort: A hospital-based observational prospective cohort of patients with hand osteoarthritis – methodology and baseline characteristics of the population', *Joint Bone Spine*. Joint Bone Spine, 88(4). doi: 10.1016/j.jbspin.2021.105171.

Sellam, J. and Berenbaum, F. (2010) 'The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis', *Nature Reviews Rheumatology*. Nat Rev Rheumatol, pp. 625–635. doi: 10.1038/nrrheum.2010.159.

Sellam, J. and Berenbaum, F. (2013) 'Is osteoarthritis a metabolic disease?', *Joint Bone Spine*. Elsevier Masson SAS, 80(6), pp. 568–573. doi: 10.1016/j.jbspin.2013.09.007.

Sequeira, A. *et al.* (2015) 'Mitochondrial Mutations in Subjects with Psychiatric Disorders', *PLoS ONE*. Public Library of Science, 10(5), p. 127280. doi: 10.1371/journal.pone.0127280.

Shadel, G. S. and Horvath, T. L. (2015) 'Mitochondrial ROS Signaling in Organismal Homeostasis', *Cell*, 163(3), pp. 560–569. doi: 10.1016/j.cell.2015.10.001.

Shane Anderson, A. and Loeser, R. F. (2010) 'Why is osteoarthritis an age-related disease?', *Best Practice and Research: Clinical Rheumatology*. NIH Public Access, pp. 15–26. doi:

10.1016/j.berh.2009.08.006.

Shariat, S. F. *et al.* (2008) 'Comparison of nomograms with other methods for predicting outcomes in prostate cancer: A critical analysis of the literature', *Clinical Cancer Research*. Clin Cancer Res, pp. 4400–4407. doi: 10.1158/1078-0432.CCR-07-4713.

Shen, J. *et al.* (2017) 'Inflammation and epigenetic regulation in osteoarthritis', *Connective Tissue Research*. Connect Tissue Res, pp. 49–63. doi: 10.1080/03008207.2016.1208655.

Shen, J. M., Feng, L. and Feng, C. (2014) 'Role of mtDNA haplogroups in the prevalence of osteoarthritis in different geographic populations: A meta-analysis', *PLoS ONE*. Edited by R. Trullas, 9(10), p. e108896. doi: 10.1371/journal.pone.0108896.

Shepherd, C. *et al.* (2015) 'Expression analysis of the osteoarthritis genetic susceptibility locus mapping to an intron of the MCF2L gene and marked by the polymorphism rs11842874', *BMC Medical Genetics*. BMC Med Genet, 16(1), pp. 1–11. doi: 10.1186/s12881-015-0254-2.

Shoffner, J. M. *et al.* (1989) 'Spontaneous Kearns-Sayre/chronic external ophthalmoplegia plus syndrome associated with a mitochondrial DNA deletion: A slip-replication model and metabolic therapy', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 86(20), pp. 7952–7956. doi: 10.1073/pnas.86.20.7952.

Silverwood, V. *et al.* (2015) 'Current evidence on risk factors for knee osteoarthritis in older adults: A systematic review and meta-analysis', *Osteoarthritis and Cartilage*. Elsevier Ltd, 23(4), pp. 507–515. doi: 10.1016/j.joca.2014.11.019.

Simon, L. S. (1999) 'Role and regulation of cyclooxygenase-2 during inflammation', in *American Journal of Medicine*. Elsevier Inc., pp. 37S-42S. doi: 10.1016/S0002-9343(99)00115-1.

Skou, S. T. and Roos, E. M. (2020) 'Erratum: Physical therapy for patients with knee and hip osteoarthritis: Supervised, active treatment is current best practice study (Clinical and Experimental Rheumatology (2019) 37:Suppl.120 (S112-S117))', *Clinical and Experimental Rheumatology*, 38(5), p. 1036.

Slatko, B. E., Gardner, A. F. and Ausubel, F. M. (2018) 'Overview of Next-Generation Sequencing Technologies', *Current Protocols in Molecular Biology*, 122(1), pp. 1–11. doi: 10.1002/cpmb.59.

Snezhkina, A. V. *et al.* (2020) 'ROS generation and antioxidant defense systems in normal and malignant cells', *Oxidative Medicine and Cellular Longevity*. Hindawi, 2019. doi: 10.1155/2019/6175804.

Solomon, D. H. *et al.* (2010) 'The comparative safety of analgesics in older adults with arthritis', *Archives of Internal Medicine*. Arch Intern Med, 170(22), pp. 1968–1978. doi: 10.1001/archinternmed.2010.391.

Soto-Hermida, Angel *et al.* (2014) 'Mitochondrial DNA (mtDNA) haplogroups influence the progression of knee osteoarthritis. Data from the osteoarthritis initiative (OAI)', *PLoS ONE*, 9(11). doi: 10.1371/journal.pone.0112735.

Soto-Hermida, A. *et al.* (2014) 'MtDNA haplogroups and osteoarthritis in different geographic populations', *Mitochondrion*. Elsevier B.V., 15(1), pp. 18–23. doi: 10.1016/j.mito.2014.03.001.

Soto-Hermida, A. *et al.* (2015) 'Mitochondrial DNA haplogroups modulate the radiographic progression of Spanish patients with osteoarthritis', *Rheumatology International*, 35(2), pp. 337–344. doi: 10.1007/s00296-014-3104-1.

Van Spil, W. E. *et al.* (2019) 'Osteoarthritis phenotypes and novel therapeutic targets',

## Bibliografia

- Biochemical Pharmacology*. Elsevier, 165(January), pp. 41–48. doi: 10.1016/j.bcp.2019.02.037.
- Srikanth, V. K. *et al.* (2005) 'A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis', *Osteoarthritis and Cartilage*, 13(9), pp. 769–781. doi: 10.1016/j.joca.2005.04.014.
- Steeve, K. T. *et al.* (2004) 'IL-6, RANKL, TNF-alpha/IL-1: Interrelations in bone resorption pathophysiology', *Cytokine and Growth Factor Reviews*. Elsevier BV, pp. 49–60. doi: 10.1016/j.cytogfr.2003.10.005.
- Stimpfel, M., Jancar, N. and Virant-Klun, I. (2018) 'New Challenge: Mitochondrial Epigenetics?', *Stem Cell Reviews and Reports*. Stem Cell Rev Rep, pp. 13–26. doi: 10.1007/s12015-017-9771-z.
- Stoccoro, A. *et al.* (2017) 'Decreased Methylation of the Mitochondrial D-Loop Region in Late-Onset Alzheimer's Disease', *Journal of Alzheimer's Disease*. J Alzheimers Dis, 59(2), pp. 559–564. doi: 10.3233/JAD-170139.
- Stoccoro, A. *et al.* (2018) 'Mitochondrial DNA copy number and D-loop region methylation in carriers of amyotrophic lateral sclerosis gene mutations', *Epigenomics*. Epigenomics, 10(11), pp. 1431–1443. doi: 10.2217/epi-2018-0072.
- Stoccoro, A. and Coppedè, F. (2018) 'Role of epigenetics in Alzheimer's disease pathogenesis', *Neurodegenerative disease management*. Neurodegener Dis Manag, 8(3), pp. 181–193. doi: 10.2217/nmt-2018-0004.
- Stoccoro, A. and Coppedè, F. (2021) 'Mitochondrial dna methylation and human diseases', *International Journal of Molecular Sciences*, 22(9), pp. 1–27. doi: 10.3390/ijms22094594.
- Stringer, C. (2016) 'The origin and evolution of homo sapiens', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1698). doi: 10.1098/rstb.2015.0237.
- Styrkarsdottir, U. *et al.* (2017) 'Whole-genome sequencing identifies rare genotypes in COMP and CHADL associated with high risk of hip osteoarthritis', *Nature Genetics*. Nature Publishing Group, 49(5), pp. 801–805. doi: 10.1038/ng.3816.
- Sui, Y. *et al.* (2009) 'Mechanical injury potentiates proteoglycan catabolism induced by interleukin-6 with soluble interleukin-6 receptor and tumor necrosis factor  $\alpha$  in immature bovine and adult human articular cartilage', *Arthritis and Rheumatism*. Arthritis Rheum, 60(10), pp. 2985–2996. doi: 10.1002/art.24857.
- Sutton, V. R. and van Bokhoven, H. (1993) *TP63-Related Disorders*, *GeneReviews*<sup>®</sup>. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20556892>.
- Szerlong, H. J. and Hansen, J. C. (2011) 'Nucleosome distribution and linker DNA: Connecting nuclear function to dynamic chromatin structure', *Biochemistry and Cell Biology*, 89(1), pp. 24–34. doi: 10.1139/O10-139.
- Takaoka, R. *et al.* (2022) 'Involvement of an FTO gene polymorphism in the temporomandibular joint osteoarthritis', *Clinical Oral Investigations*, 26(3), pp. 2965–2973. doi: 10.1007/s00784-021-04278-9.
- Tam, V. *et al.* (2019) 'Benefits and limitations of genome-wide association studies', *Nature Reviews Genetics*. Springer US, 20(8), pp. 467–484. doi: 10.1038/s41576-019-0127-1.
- Tamayo, M. *et al.* (2011) 'Decreased length of telomeric DNA sequences and increased numerical chromosome aberrations in human osteoarthritic chondrocytes', *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*. Mutat Res, 708(1–2), pp. 50–58. doi: 10.1016/j.mrfmmm.2011.01.007.

- Tanamas, S. *et al.* (2009) 'Does knee malalignment increase the risk of development and progression of knee osteoarthritis? A systematic review', *Arthritis Care and Research*, 61(4), pp. 459–467. doi: 10.1002/art.24336.
- Taniguchi, Y. *et al.* (2017) 'Regulation of Chondrocyte Survival in Mouse Articular Cartilage by p63', *Arthritis and Rheumatology*, 69(3), pp. 598–609. doi: 10.1002/art.39976.
- Tawonsawatruk, T. *et al.* (2011) 'A genetic association study between growth differentiation factor 5 (GDF 5) polymorphism and knee osteoarthritis in Thai population', *Journal of Orthopaedic Surgery and Research*. J Orthop Surg Res, 6(1). doi: 10.1186/1749-799X-6-47.
- Terkeltaub, R. *et al.* (2002) 'Invited review: The mitochondrion in osteoarthritis', *Mitochondrion*, pp. 301–319. doi: 10.1016/S1567-7249(01)00037-X.
- Tewhey, R. *et al.* (2011) 'The importance of phase information for human genomics', *Nature Reviews Genetics*. Nature Publishing Group, 12(3), pp. 215–223. doi: 10.1038/nrg2950.
- Thorstensson, C. A. *et al.* (2009) 'Natural course of knee osteoarthritis in middle-aged subjects with knee pain: 12-Year follow-up using clinical and radiographic criteria', *Annals of the Rheumatic Diseases*, 68(12), pp. 1890–1893. doi: 10.1136/ard.2008.095158.
- Thudi, M. *et al.* (2012) 'Current state-of-art of sequencing technologies for plant genomics research', *Briefings in Functional Genomics*, 11(1), pp. 3–11. doi: 10.1093/bfpg/elr045.
- Van Tilburg, M. A. L. *et al.* (2014) 'Irritable bowel syndrome may be associated with maternal inheritance and mitochondrial DNA control region sequence variants', *Digestive Diseases and Sciences*, 59(7), pp. 1392–1397. doi: 10.1007/s10620-014-3045-2.
- Tong, H. *et al.* (2017) 'Methylation of mitochondrial DNA displacement loop region regulates mitochondrial copy number in colorectal cancer', *Molecular Medicine Reports*. Mol Med Rep, 16(4), pp. 5347–5353. doi: 10.3892/mmr.2017.7264.
- Torrioni, A. *et al.* (1997) 'Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484', *American Journal of Human Genetics*, 60(5), pp. 1107–1121.
- Tranah, G. J. *et al.* (2015) 'Mitochondrial DNA sequence variation in multiple sclerosis', *Neurology*, 85(4), pp. 325–330. doi: 10.1212/WNL.0000000000001744.
- Travers, A. and Muskhelishvili, G. (2015) 'DNA structure and function', *FEBS Journal*, 282(12), pp. 2279–2295. doi: 10.1111/febs.13307.
- Uffelmann, E. *et al.* (2021) 'Genome-wide association studies', *Nature Reviews Methods Primers*. Nature Publishing Group, 1(1), pp. 1–21. doi: 10.1038/s43586-021-00056-9.
- Uitterlinden, A. G. (2016) 'An Introduction to Genome-Wide Association Studies: GWAS for Dummies', *Seminars in Reproductive Medicine*, 34(4), pp. 196–204. doi: 10.1055/s-0036-1585406.
- Umlauf, D. *et al.* (2010) 'Cartilage biology, pathology, and repair', *Cellular and Molecular Life Sciences*, 67(24), pp. 4197–4211. doi: 10.1007/s00018-010-0498-0.
- Usmani, S. E. *et al.* (2012) 'Transforming growth factor alpha controls the transition from hypertrophic cartilage to bone during endochondral bone growth', *Bone*. Bone, 51(1), pp. 131–141. doi: 10.1016/j.bone.2012.04.012.
- Vaamonde-García, C. *et al.* (2012) 'Mitochondrial dysfunction increases inflammatory

## Bibliografía

responsiveness to cytokines in normal human chondrocytes', *Arthritis and Rheumatism*, 64(9), pp. 2927–2936. doi: 10.1002/art.34508.

Valderrama, M., Ortigosa, F. and Cañas, Y. R. A. (2020) 'Métodos de secuenciación : tercera generación Secuenciación a tiempo real de una sola molécula ( SMRT , PacBio )', *Encuentros de la biología*, XIII(175), pp. 15–21.

Valdes, A. M. *et al.* (2011) 'The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance', *Annals of the Rheumatic Diseases*. NIH Public Access, pp. 873–875. doi: 10.1136/ard.2010.134155.

Valdes, A. M. and Stocks, J. (2018) 'Osteoarthritis and Ageing', *European Medical Journal*, (March). Available at: <https://www.emjreviews.com/rheumatology/article/osteoarthritis-and-ageing/>.

Vargas, F. M., Fernández, R. F. and Casanova, M. C. (2006) 'Morfología y morfometría del disco de la articulación témporomandibular en fetos y adultos humanos', *International Journal of Morphology*, 24(2), pp. 245–250.

Veronese, N. *et al.* (2018) 'Mitochondrial genetic haplogroups and incident obesity: A longitudinal cohort study', *European Journal of Clinical Nutrition*. Eur J Clin Nutr, 72(4), pp. 587–592. doi: 10.1038/s41430-018-0097-y.

Veronese, N. *et al.* (2020) 'Mediterranean Diet and KNEE Osteoarthritis Outcomes: a Longitudinal Cohort Study The Cambridge Centre for Sport and Exercise Sciences, Anglia Ruskin Europe PMC Funders Group', *Clinical nutrition*, 38(6), pp. 2735–2739.

Walker, E. A., Davis, D. and Mosher, T. J. (2011) 'Rapidly progressive osteoarthritis: Biomechanical considerations', *Magnetic Resonance Imaging Clinics of North America*. Elsevier Ltd, 19(2), pp. 283–294. doi: 10.1016/j.mric.2011.02.008.

Wallace, D. C. (2015) 'Mitochondrial DNA Variation in Human Radiation and Disease', *Cell*, 163(1), pp. 33–38. doi: 10.1016/j.cell.2015.08.067.

Wallace, D. C., Brown, M. D. and Lott, M. T. (1999) 'Mitochondrial DNA variation in human evolution and disease.', *Gene*, 238(1), pp. 211–30. doi: 10.1016/s0378-1119(99)00295-4.

Wallace, D. C., Fan, W. and Procaccio, V. (2010) 'Mitochondrial energetics and therapeutics', *Annual Review of Pathology: Mechanisms of Disease*, 5, pp. 297–348. doi: 10.1146/annurev.pathol.4.110807.092314.

Van Der Walt, J. M. *et al.* (2003) 'Mitochondrial polymorphisms significantly reduce the risk of Parkinson disease', *American Journal of Human Genetics*. Am J Hum Genet, 72(4), pp. 804–811. doi: 10.1086/373937.

Wang, C. and Youle, R. J. (2009) 'The role of mitochondria in apoptosis', *Annual Review of Genetics*. NIH Public Access, 43, pp. 95–118. doi: 10.1146/annurev-genet-102108-134850.

Wang, G. and Beier, F. (2005) 'Rac1/Cdc42 and RhoA GTPases antagonistically regulate chondrocyte proliferation, hypertrophy, and apoptosis', *Journal of Bone and Mineral Research*. J Bone Miner Res, 20(6), pp. 1022–1031. doi: 10.1359/JBMR.050113.

Wang, L. *et al.* (2021) 'Clinical significance of the D-Loop gene mutation in mitochondrial DNA in laryngeal cancer', *OncoTargets and Therapy*, 14, pp. 3461–3466. doi: 10.2147/OTT.S304836.

Wang, Q. *et al.* (2019) 'Evolution of cis- And trans-regulatory divergence in the chicken genome between two contrasting breeds analyzed using three tissue types at one-day-old', *BMC Genomics*. BioMed Central Ltd., 20(1), pp. 1–10. doi: 10.1186/s12864-019-6342-5.



- Wang, R. *et al.* (2021) 'Mitochondrial transfer from bone-marrow-derived mesenchymal stromal cells to chondrocytes protects against cartilage degenerative mitochondrial dysfunction in rats chondrocytes', *Chinese Medical Journal*. Chin Med J (Engl), 134(2), pp. 212–218. doi: 10.1097/CM9.0000000000001057.
- Wang, X. *et al.* (2022) 'Common variants in GNL3 gene contributed the susceptibility of hand osteoarthritis in Han Chinese population', *Scientific Reports*. Sci Rep, 12(1). doi: 10.1038/s41598-022-20287-4.
- Wang, X. and Kattan, M. W. (2020) 'Cohort Studies: Design, Analysis, and Reporting', *Chest*. Elsevier Inc, pp. S72–S78. doi: 10.1016/j.chest.2020.03.014.
- Wang, Y. *et al.* (2001) 'Muscle-specific mutations accumulate with aging in critical human mtDNA control sites for replication', *Proceedings of the National Academy of Sciences of the United States of America*, 98(7), pp. 4022–4027. doi: 10.1073/pnas.061013598.
- Wang, Y. *et al.* (2015) 'Mitochondrial biogenesis is impaired in osteoarthritis chondrocytes but reversible via peroxisome proliferator-activated receptor  $\gamma$  coactivator 1 $\alpha$ ', *Arthritis and Rheumatology*. Arthritis Rheumatol, 67(8), pp. 2141–2153. doi: 10.1002/art.39182.
- Wang, Yan *et al.* (2022) 'Association of variants m.T16172C and m.T16519C in whole mtDNA sequences with high altitude pulmonary edema in Han Chinese lowlanders', *BMC Pulmonary Medicine*. BMC Pulm Med, 22(1). doi: 10.1186/s12890-021-01791-1.
- Wang, Yichao *et al.* (2022) 'TGF- $\beta$ 1 Promotes Autophagy and Inhibits Apoptosis in Breast Cancer by Targeting TP63', *Frontiers in Oncology*. Frontiers Media S.A., 12, p. 865067. doi: 10.3389/fonc.2022.865067.
- Wangyang, Y. W. *et al.* (2017) 'Upregulation of P63 inhibits chondrocyte autophagy thereby enhancing the malignant progression of osteoarthritis', *Pharmazie*, 72(6), pp. 361–364. doi: 10.1691/ph.2017.6933.
- Warner, S. C. and Valdes, A. M. (2017) 'Genetic association studies in osteoarthritis: Is it fairytale?', *Current Opinion in Rheumatology*, 29(1), pp. 103–109. doi: 10.1097/BOR.0000000000000352.
- Wesseling, J. *et al.* (2009) 'CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative.', *Annals of the rheumatic diseases*. BMJ Publishing Group, 68(9), pp. 1413–9. doi: 10.1136/ard.2008.096164.
- Wesseling, J. *et al.* (2015) 'Worsening of pain and function over 5 years in individuals with "early" OA is related to structural damage: Data from the osteoarthritis initiative and CHECK (Cohort Hip & Cohort Knee) study', *Annals of the Rheumatic Diseases*, 74(2), pp. 347–353. doi: 10.1136/annrheumdis-2013-203829.
- Wesseling, J. *et al.* (2016) 'Cohort profile: Cohort Hip and Cohort Knee (CHECK) study', *International Journal of Epidemiology*. Int J Epidemiol, 45(1), pp. 36–44. doi: 10.1093/ije/dyu177.
- van der Wijst, M. G. P. and Rots, M. G. (2015) 'Mitochondrial epigenetics: An overlooked layer of regulation?', *Trends in Genetics*. Trends Genet, pp. 353–356. doi: 10.1016/j.tig.2015.03.009.
- Wolff, R. F. *et al.* (2019) 'PROBAST: A tool to assess the risk of bias and applicability of prediction model studies', *Annals of Internal Medicine*. Ann Intern Med, 170(1), pp. 51–58. doi: 10.7326/M18-1376.
- Wu, R. *et al.* (2022) 'Contralateral advanced radiographic knee osteoarthritis predicts

- radiographic progression and future arthroplasty in ipsilateral knee with early-stage osteoarthritis', *Clinical Rheumatology*. Springer, pp. 1–7. doi: 10.1007/s10067-022-06235-7.
- Yamamoto, T. and Bullough, P. G. (2000) 'The role of subchondral insufficiency fracture in rapid destruction of the hip joint: A preliminary report', *Arthritis and Rheumatism*, 43(11), pp. 2423–2427. doi: 10.1002/1529-0131(200011)43:11<2423::AID-ANR8>3.0.CO;2-Z.
- Yau, M. S. *et al.* (2017) 'Genome-Wide Association Study of Radiographic Knee Osteoarthritis in North American Caucasians', *Arthritis and Rheumatology*. *Arthritis Rheumatol*, 69(2), pp. 343–351. doi: 10.1002/art.39932.
- Yu, D. *et al.* (2017) 'Mitochondrial DNA hypomethylation is a biomarker associated with induced senescence in human fetal heart mesenchymal stem cells', *Stem Cells International*. *Stem Cells Int*, 2017. doi: 10.1155/2017/1764549.
- Zaki, E. A. *et al.* (2009) 'Two common mitochondrial DNA polymorphisms are highly associated with migraine headache and cyclic vomiting syndrome', *Cephalalgia*. *Cephalalgia*, 29(7), pp. 719–728. doi: 10.1111/j.1468-2982.2008.01793.x.
- Zeggini, E. *et al.* (2012) 'Identification of new susceptibility loci for osteoarthritis (arcOGEN): A genome-wide association study', *The Lancet*. Elsevier Ltd, 380(9844), pp. 815–823. doi: 10.1016/S0140-6736(12)60681-3.
- Zengini, E. *et al.* (2018) 'Genome-wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis', *Nature Genetics*, 50(4), pp. 549–558. doi: 10.1038/s41588-018-0079-y.
- Zhang, R. *et al.* (2017) 'Independent impacts of aging on mitochondrial DNA quantity and quality in humans', *BMC Genomics*. *BMC Genomics*, 18(1), pp. 1–14. doi: 10.1186/s12864-017-4287-0.
- Zhang, W. *et al.* (2011) 'Nottingham knee osteoarthritis risk prediction models', *Annals of the Rheumatic Diseases*. BMJ Publishing Group Ltd, 70(9), pp. 1599–1604. doi: 10.1136/ard.2011.149807.
- Zhao, Zhenxing *et al.* (2020) 'Mitochondrial DNA haplogroups participate in osteoarthritis: current evidence based on a meta-analysis', *Clinical Rheumatology*. *Clin Rheumatol*, pp. 1027–1037. doi: 10.1007/s10067-019-04890-x.
- Zheng, L. D. *et al.* (2015) 'Insulin resistance is associated with epigenetic and genetic regulation of mitochondrial DNA in obese humans', *Clinical Epigenetics*. Springer Verlag, 7(1), pp. 1–9. doi: 10.1186/s13148-015-0093-1.
- Zheng, S. *et al.* (2016) 'Monoclonal antibodies for the treatment of osteoarthritis', *Expert Opinion on Biological Therapy*, 16(12), pp. 1529–1540. doi: 10.1080/14712598.2016.1229774.
- Zhivotovsky, B. *et al.* (1999) 'Caspases: Their intracellular localization and translocation during apoptosis', *Cell Death and Differentiation*, 6(7), pp. 644–651. doi: 10.1038/sj.cdd.4400536.
- Zhong, G., Madry, H. and Cucchiari, M. (2022) 'Mitochondrial Genome Editing to Treat Human Osteoarthritis—A Narrative Review', *International Journal of Molecular Sciences*, 23(3). doi: 10.3390/ijms23031467.
- Zhu, Y. H., Zhang, Q. J. and Wang, Q. J. (2019) '[Mitochondrial DNA deletion syndrome: a case report and literature review].', *Lin chung er bi yan hou tou jing wai ke za zhi = Journal of clinical otorhinolaryngology, head, and neck surgery*. *Lin Chung Er Bi Yan Hou Tou Jing Wai Ke Za Zhi*, 33(9), pp. 808–813. doi: 10.13201/j.issn.1001-1781.2019.09.004.

Zong, W. X., Rabinowitz, J. D. and White, E. (2016) 'Mitochondria and Cancer', *Molecular Cell*. Elsevier Inc., 61(5), pp. 667–676. doi: 10.1016/j.molcel.2016.02.011.

*Capítulos de libros y tesis doctorales:*

Benito, P. (2010). Estructura y función de la articulación. En: Monfort Faure J., Benito-Ruiz, P., Blanco García, F.J., Tornero Molina, J., Möller Parera I., Monfort Faure, J y Batlle-Gualda, E. (Eds). Artrosis: fisiopatología, diagnóstico y tratamiento. Madrid: Panamericana, pp. 81-91.

Benito Jiménez, C. (2012). Replicación y recombinación. En: Benito Jiménez, C., Espino Nuño, F.J. (Eds). Genética conceptos esenciales. Madrid: Panamericana, pp. 189-215.

Burgos, M. (2012). ADN y ARN. Estructura y propiedades de los ácidos nucleicos. En: Benito Jiménez, C., Espino Nuño, F.J. (Eds). Genética conceptos esenciales. Madrid: Panamericana, pp. 155-166.

Burgos, M. (2012). Organización del material hereditario. En: Benito Jiménez, C., Espino Nuño, F.J. (Eds). Genética conceptos esenciales. Madrid: Panamericana, pp. 155-166.

De Toro Santos, F.J. (2010). Anatomía y biomecánica de la articulación humana. En: Monfort Faure J., Benito-Ruiz, P., Blanco García, F.J., Tornero Molina, J., Möller Parera I., Monfort Faure, J y Batlle-Gualda, E. (Eds). Artrosis: fisiopatología, diagnóstico y tratamiento. Madrid: Panamericana, pp. 73-81.

Espino Nuño, F.J. (2012). Genes y proteínas. En: Benito Jiménez, C., Espino Nuño, F.J. (Eds). Genética conceptos esenciales. Madrid: Panamericana, pp. 257-274.

Feduchi Canosa, E., Blasco Castiñeyra, I., Romero Magdalena, C.S. y Yáñez Conde, E. (2011). Bioenergética. En: Feduchi Canosa, E., Blasco Castiñeyra, I., Romero Magdalena, C.S. y Yáñez Conde, E.(Eds). Bioquímica: conceptos esenciales. Madrid: Panamericana, pp. 113-131

Figueiras Merino, A.M. (2012). Mutación génica y reparación del ADN. En: Benito Jiménez, C., Espino Nuño, F.J. (Eds). Genética conceptos esenciales. Madrid: Panamericana, pp. 215-234.

Linacero de la Fuente, R. (2012). Transcripción. En: Benito Jiménez, C., Espino Nuño, F.J. (Eds). Genética conceptos esenciales. Madrid: Panamericana, pp. 237-256.

Merrill, R.M. (2015). Design strategies and statistical methods in analitic epidemiology. En: Merrill, R.M (Eds). Introduction to Epidemiology, 7th ed. Utah (USA): Jones & Bartlett learning, pp. 141-159.

Muñíos, E., Blanco, F.J. y Díaz, S.M. (2010). La membrana amniótica humana: caracterización de las células madre y su aplicación en terapia celular para reparar lesiones de cartílago articular humano (Tesis doctoral). Universidade da Coruña, La Coruña, España.

Trujillo, E. (2002). Mecanismos de destrucción y reparación del cartílago. En: Batlle-Gualda, E., Benito Ruiz, P., Blanco García, F.J., Martín Mola, E. (Eds). Manual S.E.R. de la Artrosis. Madrid: IMyC, pp. 43-61.

Vela, A., Meijide, R. y Fernández, E. (2015). El sulfuro de hidrógeno como diana terapéutica en la artrosis (Tesis doctoral). Universidade da Coruña, La Coruña, España.