

Importancia de la bioinformática en la implementación de la medicina personalizada de precisión en los sistemas sanitarios: el caso del cáncer de pulmón



Lorena Aguilera Cobos
Tesis Doctoral - 2023

Director: Manuel Gonzalo Claros Díaz
Tutor: Juan Antonio García Ranea

Programa de Doctorado:
Biotecnología Avanzada





UNIVERSIDAD
DE MÁLAGA

AUTORA: Lorena Aguilera Cobos

 <https://orcid.org/0000-0001-7549-7099>

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional:

<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Cualquier parte de esta obra se puede reproducir sin autorización pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de Málaga (RIUMA): riuma.uma.es



Universidad de Málaga

Facultad de Ciencias

TESIS DOCTORAL

Importancia de la bioinformática en la
implementación de la medicina personalizada
de precisión en los sistemas sanitarios: el caso
del cáncer de pulmón

Programa de Doctorado: Biotecnología Avanzada

Lorena Aguilera Cobos

Málaga, mayo 2023



UNIVERSIDAD
DE MÁLAGA



Universidad de Málaga

Facultad de Ciencias

TESIS DOCTORAL

Importancia de la bioinformática en la
implementación de la medicina personalizada
de precisión en los sistemas sanitarios: el caso
del cáncer de pulmón

Programa de Doctorado: Biotecnología Avanzada

Lorena Aguilera Cobos

Málaga, mayo 2023



UNIVERSIDAD
DE MÁLAGA

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D./Dña LORENA AGUILERA COBOS

Estudiante del programa de doctorado BIOTECNOLOGÍA AVANZADA de la Universidad de Málaga, autor/a de la tesis, presentada para la obtención del título de doctor por la Universidad de Málaga, titulada: IMPORTANCIA DE LA BIOINFORMÁTICA EN LA IMPLEMENTACIÓN DE LA MEDICINA PERSONALIZADA DE PRECISIÓN EN LOS SISTEMAS SANITARIOS: EL CASO DEL CÁNCER DE PULMÓN.

Realizada bajo la tutorización de JUAN ANTONIO GARCÍA RANEA y dirección de MANUEL GONZALO CLAROS DÍAZ (si tuviera varios directores deberá hacer constar el nombre de todos)

DECLARO QUE:

La tesis presentada es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, conforme al ordenamiento jurídico vigente (Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), modificado por la Ley 2/2019, de 1 de marzo.

Igualmente asumo, ante a la Universidad de Málaga y ante cualquier otra instancia, la responsabilidad que pudiera derivarse en caso de plagio de contenidos en la tesis presentada, conforme al ordenamiento jurídico vigente.

En Málaga, a 26 de MAYO de 2023

Fdo.: LORENA AGUILERA COBOS Doctorando/a	Fdo.: JUAN ANTONIO GARCÍA RANEA Tutor/a
Fdo.: MANUEL GONZALO CLAROS DÍAZ Director/es de tesis	



UNIVERSIDAD
DE MÁLAGA

D. M. Gonzalo Claros Díaz, catedrático del departamento de Biología Molecular y Bioquímica de la Universidad de Málaga.

INFORMA:

Que Doña Lorena Aguilera Cobos, ha realizado bajo mi dirección, dentro del programa de doctorado Biotecnología Avanzada de la Facultad de Ciencias de la Universidad de Málaga, el trabajo de investigación recogido en la presente memoria de tesis doctoral que lleva por título: *“Importancia de la bioinformática en la implementación de la medicina personalizada de precisión en los sistemas sanitarios: el caso del cáncer de pulmón”*.

Tras la revisión de la presente memoria, he estimado oportuna su presentación ante la Comisión de Evaluación correspondiente, por lo que autorizo su exposición y defensa para optar al grado de doctora.

Y para que así conste, en cumplimiento de las disposiciones legales vigentes, firmo el presente informe.

Director de la presente tesis,

Dr. M. Gonzalo Claros Díaz



UNIVERSIDAD
DE MÁLAGA

D. Juan Antonio García Ranea, catedrático del departamento de Biología Molecular y Bioquímica de la Universidad de Málaga.

INFORMA:

Que Doña Lorena Aguilera Cobos, ha realizado bajo mi tutorización, dentro del programa de doctorado Biotecnología Avanzada de la Facultad de Ciencias de la Universidad de Málaga, el trabajo de investigación recogido en la presente memoria de tesis doctoral que lleva por título: *“Importancia de la bioinformática en la implementación de la medicina personalizada de precisión en los sistemas sanitarios: el caso del cáncer de pulmón”*.

Tras la revisión de la presente memoria, he estimado oportuna su presentación ante la Comisión de Evaluación correspondiente, por lo que autorizo su exposición y defensa para optar al grado de doctora.

Y para que así conste, en cumplimiento de las disposiciones legales vigentes, firmo el presente informe.

Tutor de la presente tesis,

Dr. Juan Antonio García Ranea



UNIVERSIDAD
DE MÁLAGA

*Las leyendas en su boca huelen a café
Y de sus palabras brotan todo lo que hoy sé
De la tierra, del abono, del cultivo
Del amor que se cimenta y el hermano que sigue el camino
De las espinas de una flor, de los amigos
De la sensación de estar unidos
Adaptado de “Raíces”. Safree, Las Migas*



UNIVERSIDAD
DE MÁLAGA

Agradecimientos

Afronto la redacción de estos agradecimientos consciente de los múltiples riesgos que esconden. El riesgo de convertirse en las líneas más leídas. El riesgo del hastío tras más de 4 años de tesis doctoral. El riesgo del equilibrio entre la ñoñería y la estricta formalidad. El riesgo del olvido de más de un nombre. Asumiendo estos retos me tiro de cabeza a por las que son para mí las últimas líneas de mi doctorado.

“A veces hay que llamar a las puertas y otras veces hay que abrirlas a cabezazos.” –

Lorena Aguilera Cobos

No recuerdo si hubo cabezazos en noviembre de 2017, pero Gonzalo Claros Díaz y Juan Antonio García Ranea, director y tutor respectivamente de esta tesis, me aceptaron entre sus filas para mi beca de Garantía Juvenil. Desde entonces su tutela ha sido providencial para que comenzara, y posteriormente no abandonara, este proyecto vital en el que se convierte una tesis doctoral.

Durante 15 meses, mantuve un encarnizado combate con la bioinformática. Por suerte no estuve sola en el ring; los desayunos a pie de lago, los debates políticos en el almuerzo y la sobremesa de desahogo vital ayudaban a superar cada uno de los asaltos. Gracias a todos los compañeros del Centro de Supercomputación y Bioinformática, y especialmente a Rocío, por colocarme los guantes, secarme las lágrimas, curar mis heridas, y sobre todo no tirar la toalla hasta salir con el puño en alto.

“No tengo mucha suerte, por eso la gestiono para los momentos cruciales.” –

Lorena Aguilera Cobos

Hay momentos cruciales que no sabes que lo son hasta que puedes mirarlos con la perspectiva que da el tiempo. Uno de esos momentos fue el 31 de enero de 2019, cuando realicé las azarosas gestiones para continuar mi deambular científico en la Agencia de Evaluación de Tecnologías Sanitarias de Andalucía. Gracias a AETSA y FPS por confiar más en mis capacidades que yo misma. Desde entonces han llovido muchas tecnologías, muchos informes, muchos fichajes, muchas búsquedas, muchos contratos, muchas PICO, mucho comisario y mucho R. Pero también ha llovido mucho café, mucha

tortilla de papas, mucho feminismo, mucho humor negro y mucha caseta. Y ya se sabe, la lluvia en Sevilla es una maravilla. Gracias AETSianos.

“La comedia es tragedia más tiempo.” –

Ignatius Farray

La representación de la tragicomedia que es una tesis doctoral sería imposible sin los entremeses que me regala mi reparto vital porque, aunque esto pueda parecer un monólogo, ellos son los directores, maquilladores, guionistas, tramoyistas, figurantes, apuntadores y secundarios de lujo de este doctorado. Cuenta este particular elenco con mujeres que me acompañan desde que éramos niñas hasta convertirnos en señoras amiguillas, señoras que trabajan en las oficinas y colegios por las que tanto han luchado, señoras capaces de mantener la compostura ante todo menos ante un intercambio de arcilla, señoras que intercalan el debate político más profundo con el cotilleo del momento, señoras que ya acunan bebés y pasean perretes sin dejar de perseguir sueños. Se cuelan entre bastidores numerosas batas blancas; unas se encuentran ya colgadas y cambiadas por el uniforme y la placa que tanto costaron conseguir, otras han sido reemplazadas por botas que pisan islas en cualquier confín del mundo, una comparte origen y sigue endulzando los gases y la vida de todos los que le rodean, un par se quedaron en mi hogar y dispensan y ensayan toda medicación que esté a su alcance, una me espera para alumbrar juntas nuestros mejores proyectos, otra voló tan lejos que si no fuera por su brillo sería indistinguible en el paisaje polar, y la última tiene bordadas las palabras Crick y libertad. Y no es casualidad que este párrafo resume teatralidad, porque aunque desde el 5ºD no encontráramos las clases, entre las tablas, junto a mi tripulación, navegué mis Martirios, y parece que nunca me querré despegar la etiqueta de Lorena Teatro.

“Me quedaré contigo. Eres verde, fuerte, da gusto verte. Juntos no habrá quien se atreva a toserme.” –

Asno, Shrek

La RAE define los cimientos como: “principio y raíz de algo”. Debe ser causalidad, que nuestra raíz y principio fueran para placas fotovoltaicas a prueba de terremotos en el otro lado del mundo, capaces de soportarlo todo sin importar fuerzas ni distancias. El

algoritmo prometía una trompa azul y a Marshal de acompañante, en su lugar me encontré la entrada de una final y al Doctor Cateto. Gracias por hacerme sentir que eres mi Champions, por sentir que soy tu Europa League, porque pase lo que pase nadie las quiere como nosotros. Que parezca para siempre, creo que somos muy conscientes de lo que tenemos. Cuídamela.

“Míos son los pasos, pero gracias por tan buenos zapatos.” –

Lorena Aguilera Cobos

No es banal que la dedicatoria de esta tesis surja de un tema titulado Raíces. Por no considerarme profeta de mi tierra y dado mi errante vagar, durante mucho tiempo creí carecer de ellas. Nada más lejos de la realidad, sin ser consciente de ello creo haber desarrollado una red radicular de alimento y sostén fuera de la cual sería imposible entenderme. En los párrafos anteriores, de estos (quizás ya demasiado extensos) agradecimientos, se vislumbran numerosas de las bifurcaciones de esta red, pero surgen todas ellas de la raíz central de una flor entre espinas. De ellos fui semilla, ellos mi abono y de ellos mi no siempre fácil cultivo. Gracias por seguir sulfatándome incluso sin entender mis decisiones. Este es uno de los frutos de vuestra constante, cariñosa y atenta crianza. La previsible próxima cosecha ya ha emprendido su camino y amenaza con ser una de las mejores añadas que se recuerden. Gracias nano porque dos manzanas podridas recogidas del mismo contenedor sobreviven mejor juntas en el mismo cesto.

Y, por último, y por ello más importante, gracias a la persona en la que me he convertido capaz de llegar hasta aquí aprendiendo, soltando, respirando y confiando ante todo en mí. El síndrome de la impostora amenaza con aparecer detrás de cada esquina, pero mi confianza es más rápida.

Y a partir de aquí, disfruten de la que considero una de mis mejores obras hasta el momento. Buena lectura.



UNIVERSIDAD
DE MÁLAGA

Contenido

Listado de abreviaturas	23
Resumen	25
1. Introducción	29
1.1 ¿Qué es la medicina personalizada de precisión?	29
1.2 Origen y desarrollo de las ciencias ómicas	30
1.3 Las tecnologías de secuenciación.	32
1.3.1 Tecnologías de secuenciación de lecturas cortas.....	35
1.3.2 Tecnologías de secuenciación de una sola molécula	38
1.3.3 Tecnologías de secuenciación de ARN	40
1.4 La bioinformática	42
1.4.1 Desarrollo de la computación y su aplicación en las metodologías de secuenciación	43
1.4.2 Explosión de la bioinformática en las últimas décadas.....	49
1.5 Evaluación e implementación de la medicina personalizada de precisión.	51
1.6 Evaluación en el manejo del cáncer de pulmón en el contexto de la MPP.....	53
1.6.1 Alteraciones genéticas en el cáncer de pulmón	54
1.6.2 Aplicación de la MPP en el diagnóstico de cáncer de pulmón	57
1.6.3 Aplicación de la MPP en el tratamiento de cáncer de pulmón.....	61
2. Objetivos	65
3. Metodología	69
3.1 Revisión sistemática.....	69
3.1.1 Formulación de la pregunta de investigación	70
3.1.2 Desarrollo de un protocolo de la revisión sistemática y registro.....	71
3.1.3 Estrategia de búsqueda.	71
3.1.4 Selección de estudios.	73

3.1.5	Valoración del riesgo de sesgo de los estudios seleccionados.	74
3.1.6	Síntesis de la evidencia de forma cualitativa o cuantitativa.	80
3.2	Obtención de datos ARNseq para su análisis bioinformático	87
3.3	Análisis bioinformático de expresión diferencial de datos de ARNseq.....	88
3.4	Análisis de correlación de los datos de expresión	90
3.5	Interpretación biológica de datos del análisis de datos de ARNseq	91
4.	Resultados	95
4.1	Elaboración de un marco de evaluación cualitativo y cuantitativo adaptable a distintos sistemas sanitarios para evaluar el nivel de evaluación de la implementación de la MPP.	95
4.1.1	Generación del marco de evaluación de la implementación de la MPP a nivel nacional: revisión sistemática	96
4.1.2	Generación del marco de evaluación de la implementación de la MPP a nivel nacional: sistema de puntuación.....	118
4.2	Evaluación de la implementación de la MPP a nivel nacional en distinto sistemas sanitarios mediante el marco generado.	119
4.3	Metanálisis de la efectividad de la biopsia líquida para el diagnóstico y la adecuación del tratamiento en cáncer de pulmón	121
4.3.1	Revisión sistemática de la bibliografía publicada.....	122
4.3.2	Metanálisis de la efectividad para la adecuación del tratamiento del análisis de las mutaciones en <i>EGFR</i> en muestras de biopsia líquida en pacientes de NSCLC.....	128
4.4	Expresión diferencial de ARNlnc entre tejido tumoral y tejido sano de pacientes con diferentes tipos de cáncer de pulmón y diferentes etnias	133
4.5	Correlación de la expresión de ARNlnc y ARNm para identificar posibles biomarcadores nuevos para biopsia líquida en los pacientes con cáncer de pulmón.....	136
4.5.1	Correlaciones de expresión diferencial entre ARNlnc y ARNm.....	136

4.5.2 Interpretación funcional de las correlaciones ARNlnc-ARNm en los pacientes caucásicos	138
5. Discusión	151
5.1 Evaluación de la implementación de la MPP a nivel nacional.....	151
5.2 Biopsia líquida para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón.....	153
5.3 Análisis bioinformático de expresión diferencial de ARNlnc y búsqueda de nuevos biomarcadores para cáncer de pulmón.....	155
6. Conclusiones.....	159
Referencias	163
Anexos	187
Anexo 1: Estrategias de búsqueda de la revisión sistemática para la generación de un marco de evaluación de la implementación de la MPP a nivel nacional.	187
Anexo 2: Resultados de la evaluación de la calidad de los estudios incluidos en la revisión sistemática para la generación de un marco de evaluación de la implementación de la MPP a nivel nacional.....	199
Anexo 3: Tabla detallada de los estudios incluidos de los que se ha extraído cada uno de los elementos evaluables que constituyen los dominios.	200
Anexo 4: Tabla con las puntuaciones para cada uno de los elementos evaluables y los resultados obtenidos en la evaluación de los países seleccionados	205
Anexo 5: Estrategias de búsqueda de la revisión sistemática para el análisis de la efectividad de la biopsia líquida en el cáncer de pulmón	211
Anexo 6: Tablas detalladas de la información extraída de las revisiones sistemáticas incluidas para el análisis de la efectividad de la biopsia líquida en el cáncer de pulmón.....	221
Anexo 7: Datos extraídos de los estudios clínicos incluidos para el MA de la eficacia de la detección de mutaciones de <i>EGFR</i> en BL para la adecuación del tratamiento	230

Anexo 8: Resultados de la herramienta QUADAS-2 de la evaluación de la calidad de los estudios incluidos en el MA de la eficacia de la detección de mutaciones de <i>EGFR</i> en BL para la adecuación del tratamiento.....	233
Anexo 9: <i>Forest plots</i> de la sensibilidad y especificidad obtenida en el metanálisis todos los estudios clínicos incluidos para evaluar la efectividad de la adecuación del tratamiento del análisis de las mutaciones en <i>EGFR</i> en muestras de biopsia líquida en pacientes de NSCLC al compararlas con biopsias tisulares.	238
Anexo 10: Listado de los ARNlnc con expresión diferencial comunes en los 4 tipos de muestras.	241

Listado de abreviaturas

ADNc:	ADN complementario
ARNcirc:	ARN circulares
ARNInc:	ARN largo no-codificante
ARNm:	ARN mensajero
ARNnc:	ARN no codificante
ARNpi:	ARN pequeño de interferencia
ARNseq:	<i>ARN sequencing</i>
BAG:	Biopsia con aguja gruesa
BL:	Biopsia líquida
cM:	centiMorgan
COSMIC:	<i>Catalogue of somatic mutations in cancer</i>
CRT:	<i>Cyclic reversible termination</i>
DAVID:	<i>Database for annotation, visualization and integrated discovery</i>
ddNTP:	2',3'-didesoxinucleótidos
dNTP:	desoxinucleótidos
EGA:	<i>European genome-phenome archive</i>
Embase:	<i>Excerpta medica database</i>
ENIAC:	<i>Electronic numerical integrator and computer</i>
EPOC:	Enfermedad pulmonar obstructiva crónica
FPKM:	<i>Fragments per kilobase per million</i>
GO:	Gene ontology
INAHTA:	<i>International network of agencies for health technology assessment</i>
ELSI:	<i>Ethical, legal and social implications</i>
HTA:	<i>Health technologies assessment</i>
kb:	kilobases
KEGG:	<i>Kyoto encyclopedia of genes and genomes</i>
LAC:	<i>Lung adenocarcinoma</i>
LAC50:	100 muestras de 50 pacientes de LAC
LAC7:	14 muestras pareadas de 7 pacientes de LAC

LCC: *Large cancer cells*

logFC: *Logarithmic fold change*

MA: metanálisis

miARN: microARN

MLE: *Maximum likelihood estimation*

MPP: medicina personalizada de precisión

MSA: *Multiple sequence alignment*

NCBI: *National center of biotechnology information*

NGS: *Next-Generation Sequencing*

NSCLC: *Non-small cell lung cancer*

PAAF: Punción aspirativa con aguja fina

PAM: *Point accepted mutations*

pb: pares de bases

RES: red española de supercomputación

RISE: *RNA interactome from sequencing experiments*

RNAInter: *RNA interactome database*

ROC: *receiver operating characteristic*

ROSE: *Rapid on-site evaluation*

RPKM: *Read per kilobase per million*

RS: revisión sistemática

SBL: *Sequencing by ligation*

SBS: *Sequencing by synthesis*

SCBI: Centro de Supercomputación y Bioinformática

SCLC: *Small cell lung cancer*

SCLC17: 34 muestras de 17 pacientes de SCLC

SLR: *Synthetic long-reads*

SMRT: *Single-molecule real-time sequencing*

SNA: *Single-nucleotide addition*

SNP: *Single nucleotide polymorphism*

SqCC: *Squamous cell carcinoma*

SqCC8: 16 muestras pareadas de 8 pacientes de SqCC

TKI: *tyrosine kinase inhibitors*

Resumen

La medicina personalizada de precisión (MPP) es la identificación y aplicación del abordaje preventivo, diagnóstico y terapéutico más efectivo para cada paciente, utilizando como herramienta la medicina de precisión. Constituye una nueva frontera en la atención sanitaria que combina las ciencias ómicas, el análisis de grandes datos y la salud de la población.

La implementación de la MPP en los sistemas sanitarios frente a la medicina clásica tiene como objetivo eliminar o mitigar los posibles resultados perjudiciales para los pacientes de la aplicación de la medicina clásica. Para conseguir un alto grado de implementación de la MPP debe realizarse un análisis del sistema sanitario en su conjunto. Este enfoque holístico podría permitir la detección de las necesidades que surgirían en el sistema con esta implementación para que fueran tenidas en cuenta al elaborar las políticas sanitarias.

Para conseguir la implementación de la MPP han surgido numerosas iniciativas que contemplan diversas finalidades terapéuticas (diagnóstica, predictiva presintomática, análisis de portadores, diagnóstica prenatal, diagnóstica preimplantacional o farmacogenética) y su aplicación para diferentes enfermedades. En todas ellas se indica que es fundamental su uso en el cáncer. Este aspecto común de todas las iniciativas es debido a la profunda caracterización molecular y genética del cáncer y a su alta prevalencia en la población mundial (en 2020 se contabilizaron aproximadamente 19,3 millones de casos nuevos a nivel mundial, según últimos datos disponibles en el proyecto GLOBOCAN). Entre los más de 100 tipos de cáncer, el cáncer de pulmón es el segundo tipo de cáncer más prevalente tras el de mama, con 2.206.771 de pacientes diagnosticados en 2020 a nivel mundial y 29.549 pacientes diagnosticados en España en 2021.

Esta tesis doctoral aborda la MPP en distintas fases de su desarrollo, desde la investigación hasta la implementación en la rutina clínica, y desde distintos enfoques. Un primer enfoque holístico de la implementación y un segundo enfoque centrado en

una enfermedad y aplicación concreta: la eficacia de la biopsia líquida y los análisis de expresión de ARN largos no codificantes (ARNlnc) y ARN mensajeros (ARNm) para la mejora del diagnóstico y del tratamiento en el cáncer de pulmón.

Se propone en este trabajo un marco de evaluación cualitativo y cuantitativo de los sistemas sanitarios, resultado de una revisión sistemática, que demuestra la correlación directa entre el contexto económico y el nivel de desarrollo de los países respecto al nivel de implementación de la MPP a nivel nacional.

Siguiendo la misma estrategia de trabajo que la desarrollada para el marco de evaluación, determinamos mediante metanálisis que la detección de mutaciones en el gen *EGFR* en ADN circulante libre en biopsia líquida para el tipo de cáncer de pulmón de células no pequeñas con la finalidad de adecuación del tratamiento es una alternativa muy ventajosa a la biopsia tisular, tanto si el análisis se realiza en suero como en plasma del paciente. Estos resultados son más fiables, en términos de sensibilidad y especificidad, si el paciente no ha recibido previamente tratamiento de quimioterapia ni inhibidores de la tirosina quinasa.

A la par, realizamos estudios bioinformáticos de 14 muestras pareadas de 7 pacientes caucásicos con adenocarcinoma de pulmón, 16 muestras pareadas de 8 pacientes caucásicos con cáncer de pulmón epidermoide, 34 muestras de 17 pacientes asiáticos con cáncer de pulmón de células pequeñas y 100 muestras de 50 pacientes asiáticos con adenocarcinoma de pulmón, que confirman la expresión diferencial de ARNlnc entre tejido sano y tejido tumoral en los distintos tipos de cáncer de pulmón, aunque la etnia también afecta mucho al resultado.

Posteriormente, realizamos un análisis de las correlaciones de expresión entre los ARNlnc y los ARNm que presentaban expresión diferencial entre tejido sano y tejido tumoral en las poblaciones caucásicas incluidas. Los resultados obtenidos nos permiten proponer posibles ARNlnc biomarcadores en biopsia líquida para esta población: para el adenocarcinoma de pulmón *CRNDE*, *FENDRR*, *DRAIC* y *UCA1*, y para el cáncer de pulmón epidermoide, *FENDRR*, *FOXD2-AS1* y *SENCR*.

Introducción



UNIVERSIDAD
DE MÁLAGA

1. Introducción

1.1 ¿Qué es la medicina personalizada de precisión?

Las últimas décadas han traído consigo grandes avances en materia de salud, con diversas intervenciones que aumentan la longevidad, reducen el sufrimiento y promueven la eficacia y la eficiencia de la asistencia sanitaria. Con la combinación de los continuos desarrollos (bio)tecnológicos y la necesidad de una toma de decisiones centrada en la persona, la medicina ha llegado a una era en la que una mayor personalización es posible y necesaria.

En este contexto surgen los términos medicina personalizada y medicina de precisión. Términos que están muy relacionados entre sí, ya que ambos se refieren al uso de características únicas de los pacientes para seleccionar el mejor tratamiento. Sin embargo, no existe consenso universal en su definición, utilizándose ambos términos indistintamente en muchos casos.

La **primera aparición** como término científico del concepto medicina personalizada o medicina de precisión aparece en 1991 (Visvikis-Siest, Gorenjak y Stathopoulou 2018). Cabe destacar que la idea ya estaba en la medicina muchos años atrás, aunque no existiera un término que la definiera. Este concepto se aplicaba cuando un tratamiento se usaba de forma selectiva con unos pacientes y no con otros sobre los que podría ser perjudicial. Un claro ejemplo se puede encontrar en las donaciones de sangre dependientes del grupo sanguíneo realizadas desde 1900 (Yamamoto et al. 1990).

Formalmente las definiciones para ambos conceptos, según la bibliografía encontrada son:

El International *Consortium for Personalised Medicine* (ICPerMed) define la **medicina personalizada** como una estrategia dirigida a la prevención, diagnóstico y tratamiento de la enfermedad en función del perfil específico de un individuo, es decir, teniendo en cuenta la heterogeneidad genética entre los individuos, el medio ambiente y el estilo de vida, incluida la dieta. Sin embargo, muchas veces se supone erróneamente que, debido

Introducción

a la naturaleza del término "personalizada", la medicina personalizada desarrolla tratamientos únicos para cada individuo, y esto no siempre es así (Vicente, Ballensiefen y Jönsson 2020).

Por otro lado, según el *National Research Council* (NRC) de EE. UU. y la *Precision Medicine Initiative* (PMI), la **medicina de precisión** se define como una estrategia emergente para el tratamiento y prevención de las enfermedades que conlleva una adaptación a las características individuales de cada paciente (Sankar y Parker 2017). En esta definición se incluye además la capacidad de clasificar a los pacientes en subpoblaciones basada en perfiles complejos y multimodales que utilizan datos biológicos, clínicos y ambientales (Orfao et al. 2019).

Debido a este conflicto en la terminología surge el término **medicina personalizada de precisión** (Rodríguez Weber y Ramírez Arias 2022) como la identificación y aplicación del abordaje preventivo, diagnóstico y terapéutico más efectivo para cada paciente, utilizando como herramienta la medicina de precisión. Constituye una nueva frontera en la atención sanitaria que combina las ciencias ómicas, el análisis de grandes datos y la salud de la población.

Tras establecer estas definiciones, el concepto que más se ajusta al propósito de esta tesis es el de medicina personalizada de precisión (MPP). Por ello, se usará este término de forma preferencial en este trabajo.

1.2 Origen y desarrollo de las ciencias ómicas

Como ya se ha comentado, desde hace al menos un siglo se aplican intervenciones médicas de forma diferencial entre pacientes atendiendo a características individuales. Sin embargo, el actual concepto de MPP previamente expuesto va mucho más allá y no se puede entender sin el desarrollo de las ciencias ómicas.

Existen tantas **ciencias ómicas** como elementos biológicos o moleculares susceptibles de ser estudiados y, en general, se nombran añadiendo el sufijo “-ómica” al conjunto de moléculas y elementos estudiados. El análisis de la información derivada de las ciencias ómicas permite una mejor comprensión de la fisiopatología y la contextualización de las

enfermedades, para una mejor prevención, diagnóstico y adaptación del tratamiento, basado en las diferencias interindividuales (Orfao et al. 2019). En la tabla 1 se definen las principales ciencias ómicas.

Nombre	Definición
Genómica	Estudio del conjunto del material genético (genoma) presente en un organismo.
Farmacogenómica	Estudio de los genes que afectan a la respuesta de una persona a determinados fármacos.
Nutrigenómica	Estudio de la interacción entre genes y nutrientes.
Transcriptómica	Estudio de los perfiles de expresión de los ARN mensajeros (ARNm) y los ARN no codificantes (ARNnc).
Epigenómica	Estudio de los elementos que controlan la expresión génica sin modificar la secuencia de nucleótidos del genoma.
Proteómica	Estudio del conjunto de proteínas expresadas en un organismo en un momento determinado y particular de cada tipo celular o tisular.
Metabolómica	Identificación y cuantificación de productos metabólicos de pequeño tamaño (metabolitos) de un sistema biológico (célula, tejido, fluido biológico u órgano).
Lipidómica	Estudio y caracterización del conjunto de los lípidos celulares, las moléculas con las que interactúa y sus funciones en el organismo.
Secretómica	Estudio de las moléculas orgánicas y compuestos inorgánicos secretados por las células, tejidos u órganos.
Interactómica	Estudio de las interacciones moleculares en un sistema biológico.
Citómica	Estudio de la colección de procesos celulares complejos y dinámicos subyacentes a los procesos fisiológicos, así como de la diversidad funcional y estructural de conjunto de células de un organismo.
Fenómica	Estudio del conjunto de los fenotipos mostrados por una célula, resultado de la interacción del genotipo y el ambiente.
Exposómica	Estudio del conjunto de factores ambientales a los que se expone una persona a lo largo de su vida.
Metagenómica	Estudio del conjunto de microorganismo de una muestra ambiental para proporcionar información de la diversidad ecológica de un ambiente determinado, mediante la obtención secuencias del genoma de los diferentes microorganismos que componen la comunidad.
Microbiómica	Estudio del material genético de los microorganismos en un nicho específico, como el tubo digestivo, la cavidad bucal o la piel.

Tabla 1: Descripción de las principales ciencias ómicas (Orfao et al. 2019).

Introducción

El origen de las ciencias ómicas ha sido posible gracias, entre otros factores, al desarrollo de las tecnologías de secuenciación para analizar los elementos biológicos y de la bioinformática para analizar y dar fundamento biológico y clínico a la información obtenida.

1.3 Las tecnologías de secuenciación.

La publicación de la estructura del ADN por parte de Watson y Crick (Watson y Crick 1953), junto con la acumulación de datos y conocimientos de la bioquímica y la estructura de los ácidos nucleicos y las proteínas por parte de los estudios de Rosalind Franklin (Cobb y Comfort 2023), Linus Pauling, Robert Corey y Gopalasamudram Ramachandran (Canduri 2017) supusieron un punto de inflexión en el desarrollo de la biología molecular y las tecnologías desarrolladas basadas en sus avances.

Gracias a ello surge la primera metodología de secuenciación de ADN desarrollada por **Maxam-Gilbert** en 1976. Era una técnica muy compleja basada en la reactividad química de los ácidos nucleicos marcados con radioactividad (Maxam y Gilbert 1977). Esta técnica fue rápidamente reemplazada por la metodología de secuenciación de terminación de cadena Sanger, desarrollada por Frederick Sanger en 1977, y que se terminó imponiendo por ser más sencilla y precisa (Sanger, Nicklen y Coulson 1977). En la primera metodología se secuenciaba directamente el ADN celular, mientras que en la secuenciación Sanger se secuenciaban copias de ADN purificado.

El método de **secuenciación de terminación de cadena Sanger** se basa en la síntesis de una cadena de ADN complementaria por parte de la ADN polimerasa en presencia de 20 desoxinucleótidos (dNTP) y 2',3'-didesoxinucleótidos (ddNTP). Cada vez que se incorpora un ddNTP durante la síntesis de ADN, la reacción se termina. Mediante electroforesis en un gel de poliacrilamida se separan por tamaño las hebras resultantes, se identifican los ddNTP terminales marcados con radiactividad (posteriormente este marcaje pasaría a realizarse con fluoróforos), revelándose así la secuencia de ADN de la cadena molde. Las secuencias obtenidas con este método tienen una longitud de 200 bases (Sanger, Nicklen y Coulson 1977). El primer genoma que se secuenció con este método fue el del bacteriófago Φ -X-174 de 5386 pares de bases (pb), posteriormente

también se secuenció el primer genoma viral del virus *Vaccina*. El primer programa informático dedicado al análisis de los resultados de las lecturas de la metodología Sanger fue publicado en 1977 por Roger Staden (Staden 1977). Las limitaciones de este método de secuenciación eran la necesidad de amplificación de los fragmentos de ADN a secuenciar, su elevado coste, la baja capacidad de automatización de la técnica para la secuenciación de genomas completos, y el gran coste de tiempo de interpretación de los resultados. Ante estas limitaciones, surgió la necesidad de un nuevo proceso robusto, automatizado y barato.

Entre 1977 y 1996 aparecieron avances y mejoras en las metodologías de secuenciación. En concreto, en 1986 se publica el primer reporte de método de **automatización de la secuenciación de ADN**, con el uso de terminadores fluorescentes como variante del método Sanger (Gauthier et al. 2019). Esta modificación utiliza una molécula fluorescente diferente unida a cada ddNTP y las secuencias obtenidas, de 500 a 1000 bases de largo, se leen por ordenador.

En 1987 aparece el **primer secuenciador automático**, ABI370A, generado por la empresa Applied Biosystems. Este aparato sigue la metodología de secuenciación de Sanger y empleando a continuación electroforesis capilar separa cada fragmento por tamaño, al estar los fragmentos marcados fluorescentemente se obtiene una secuencia ordenada de los fragmentos en cromatogramas. Con esta tecnología se determinó el primer genoma procariota (*Hemophylus influenzae*) y el primero eucariota (*Saccharomyces cerevisiae*).

Para continuar con el eje temporal hay que mencionar el **Proyecto del Genoma Humano** (*Human Genome Project*), que ha marcado enormemente el desarrollo actual de la secuenciación y su análisis bioinformático. Se inició en el año 1990 y finalizó en 2003. En 1994 se publicó un mapeo detallado del genoma humano que incluía 5840 locus con una media de espaciado de 0,7 cM (centiMorgan). En 1999 se publica por primera vez la secuencia completa de un cromosoma humano, el 22. Para ese mismo año y basándose en la tecnología de secuenciación por capilares de Applied Biosystems (ABI3700), la empresa Celera comienza la secuenciación del genoma completo. El 25 de junio de 2000, se anunció públicamente el borrador del genoma humano y finalmente en el año 2004

Introducción

se publicó la versión definitiva (Bentley 2000). Además de por este hito, este proyecto debe ser destacado por la amplia colaboración internacional, la gran inversión de la que fue dotado (3000 millones de dólares), y la gran relevancia que este conocimiento ha supuesto en los estudios de medicina y genética clínica (Venter et al. 2001). La secuenciación del genoma humano se completa con el pangenoma en 2023 (Liao et al. 2023).

La finalización del Proyecto del Genoma Humano puso de relieve la necesidad de mejora de las tecnologías de secuenciación, para superar los desafíos biológicos y computacionales que surgieron durante este proyecto. El primer método de secuenciación masiva (NGS, *Next-Generation Sequencing*) se basó en la **pirosecuenciación**, fue publicado en el año 2005 en la revista *Nature*, y llevado al mercado por la empresa 454, adquirida posteriormente por Roche. La plataforma permitía secuenciar 25 millones de bases con una precisión del 99% en 4 horas. El lanzamiento de esta plataforma supuso una reducción de costes de 50000 veces respecto a lo que costó el primer genoma humano secuenciado (Wetterstrand 2013).

En la figura 1 se muestra una línea cronológica en la que se destacan los principales hitos de las tecnologías de secuenciación.

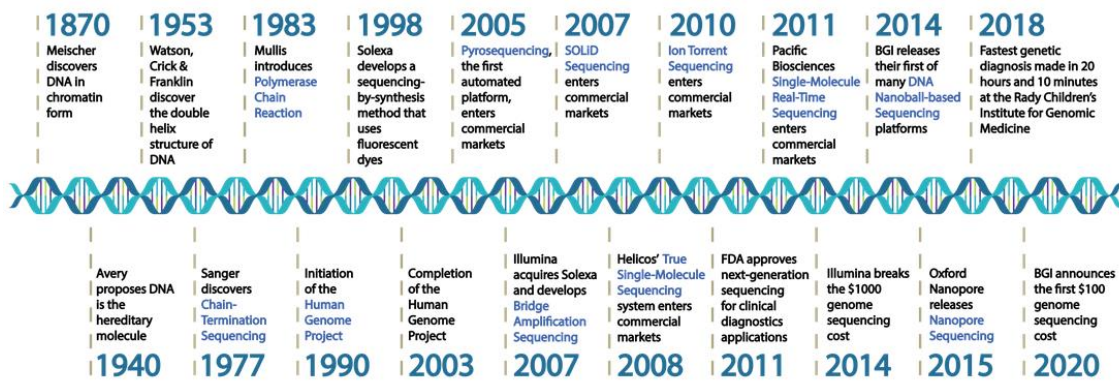


Figura 1: Evolución cronológica de las metodologías de secuenciación. Ref: Apollo Institute

Tras estas evoluciones tecnológicas, se produce por parte de diferentes casas comerciales el desarrollo de una serie de plataformas de NGS genómicas basadas en diferentes metodologías, lo cual ha permitido un fuerte crecimiento en la calidad de los resultados y al mismo tiempo el decrecimiento en el precio y tiempo necesarios. Para su

descripción las vamos a dividir entre las que generan lecturas cortas y las que leen una sola molécula.

1.3.1 Tecnologías de secuenciación de lecturas cortas

Las primeras tecnologías de NGS generaban lecturas muy cortas, originalmente de 25-30 pb, y después se llegó incluso a 250 pb. El ADN se amplifica clonalmente en una superficie sólida para obtener muchos miles de copias idénticas de cada fragmento de ADN, lo que garantiza que la señal pueda distinguirse del ruido de fondo. Se secuencian simultáneamente muchos miles de moléculas diferentes (Goodwin, Mcpherson y McCombie 2016). En este tipo de **secuenciación de lecturas cortas** se pueden distinguir 2 tipos de tecnologías:

- En la **secuenciación por ligación** (SBL, *Sequencing by ligation*) una secuencia de sonda unida a un fluoróforo se hibrida con un fragmento de ADN y se liga a un oligonucleótido adyacente. El espectro de emisión del fluoróforo indica la identidad de la base o bases complementarias a posiciones específicas dentro de la sonda. La plataforma que utilizó esta estrategia fue *SOLiD* y logra lecturas de una longitud de hasta 75 bp (Valouev et al. 2008). La última versión apareció en 2011 (*SOLiD 4*) y desapareció tras la compra de *Applied Biosystems* por *ThermoFischer Scientific*.
- En la **secuenciación por síntesis** (SBS, *Sequencing by synthesis*) se utiliza una polimerasa y se usa una señal luminosa o un cambio en la concentración iónica para identificar la incorporación de un nucleótido en una cadena de elongación (Metzker 2010). A su vez la SBS se puede dividir en 2 subtipos de tecnologías:
 - La **terminación cíclica reversible** (CRT, *cyclic reversible termination*) se caracteriza por el uso de moléculas terminadoras similares a las utilizadas en la secuenciación de Sanger, en las que el grupo 3'-OH de la ribosa está modificado para bloquear la elongación. Para comenzar el proceso, se cebe una plantilla de ADN con una secuencia complementaria a una región adaptadora, que iniciará la unión de la polimerasa a esta región de ADN de doble cadena. En cada ciclo, se añade una mezcla de los cuatro dNTP etiquetados individualmente y con la posición 3' bloqueada. Después de la incorporación de un solo dNTP a cada hebra

complementaria en elongación, se eliminan los dNTP no unidos y se obtiene una imagen de la superficie para identificar cuál se incorporó en cada grupo. El fluoróforo y el grupo de bloqueo se retiran y se inicia un nuevo ciclo (Ju et al. 2006). Dos plataformas utilizaron esta tecnología de secuenciación para sus dispositivos *Illumina* y *Qiagen*, consiguiendo lecturas de una longitud de 150 bp, que han logrado incrementar en algunos casos a 300 bp.

- La **adición de nucleótidos simples** (SNA, *single-nucleotide addition*) se basa en una única señal para marcar la incorporación de un dNTP en una cadena en elongación. En consecuencia, cada uno de los cuatro nucleótidos debe añadirse de forma iterativa a una reacción de secuenciación para garantizar que sólo un dNTP es responsable de la señal. Además, esto no requiere que los dNTP estén bloqueados, ya que la ausencia del siguiente nucleótido en la reacción de secuenciación impide la elongación (Margulies et al. 2005)(Rothberg et al. 2011). Dos plataformas utilizaron esta tecnología de secuenciación para sus dispositivos *454 pirosecuenciación* e *Ion Torrent*, logrando lecturas de hasta 1000 y 400 pares de bases. La gran diferencia entre ambas plataformas que usan este tipo de secuenciación es que la primera detecta la señal por emisión de luz al acoplarle una reacción de luciferasa (de ahí que se denominara pirosecuenciación), mientras que en el segundo se mide el cambio de pH porque en la incorporación de un nucleótido se libera un H^+ . Aunque *ThermoFischer* sigue vendiendo nuevas versiones de *Ion Torrent*, *Roche* dejó de comercializar este aparato y kits de pirosecuenciación en 2016.

La figura 2 resume los tipos de tecnologías de secuenciación de ADN de lecturas cortas.

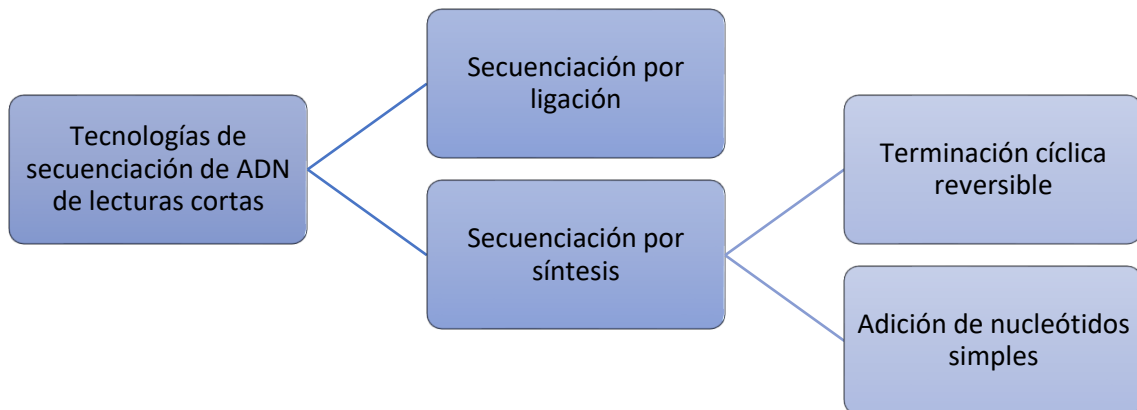


Figura 2: Tipos de tecnologías de secuenciación de ADN de lecturas cortas.

En la secuenciación de ADN de lecturas cortas la casa comercial dominante es *Illumina*. Este dominio se debe a la gran madurez de su tecnología, el alto grado de compatibilidad entre sus plataformas, la existencia de modelos de distinta escala y la constante renovación y mejora de sus plataformas. Esta amplia gama le permite adaptarse a las necesidades de cada centro de investigación, disponiendo de numerosas opciones en cuanto a tiempo de ejecución, estructura de la lectura y longitud de la misma (Goodwin, Mcpherson y Mccombie 2016). Esta casa comercial divide los instrumentos actualmente disponibles en secuenciadores de sobremesa y secuenciadores a escala de producción. En la gama de secuenciadores de sobremesa se encuentran las plataformas *iSeq 100*, *MiniSeq*, *MiSeq Series*, plataformas de la serie *NextSeq 550* y plataformas de la serie *NextSeq 1000 & 2000*, con estos equipos se obtienen una longitud de lectura de 2 x 150 pb y un número de lecturas máxima por carrera que oscila entre los 4 y los 1,1 mil millones. En la gama de secuenciadores a escalas de producción cuentan con las series *NextSeq 550*, *NextSeq 1000*, *NextSeq 2000* y *NovaSeq 6000*. Con los equipos de esta gama se obtiene una longitud de lectura de hasta 2 x 250 pb (en el caso de *NovaSeq 6000*) y un número de lecturas máxima por carrera que oscila entre los 400 millones de la serie *NextSeq 550* y los 20 mil millones en el caso del secuenciador *NovaSeq 6000*.

Introducción

1.3.2 Tecnologías de secuenciación de una sola molécula

Los genomas tienen una estructura compleja, con muchos elementos repetitivos, y posibles alteraciones del número de copias y variaciones estructurales, que pueden ser relevantes para el estudio de la evolución o de alguna enfermedad. Sin embargo, la complejidad de estos elementos hace muy difícil resolverlos con la secuenciación de lecturas cortas. La **secuenciación de una sola molécula** proporciona lecturas de varias kilobases (kb) de una única molécula (sin amplificar), lo que permite la resolución de estas grandes características estructurales. Estas lecturas largas pueden abarcar regiones complejas o repetitivas con una sola lectura continua, eliminando así la ambigüedad en las posiciones o el tamaño de los elementos genómicos (Goodwin, Mcpherson y Mccombie 2016). En este tipo de secuenciación se puede distinguir 3 tipos de tecnologías:

- La tecnología de **secuenciación de una sola molécula en tiempo real** (SMRT, *single-molecule real-time sequencing*) se diferencia de la secuenciación de lecturas cortas en que no depende de una población clonal de fragmentos de ADN amplificados para generar una señal detectable, ni requiere ciclos químicos para cada dNTP añadido (van Dijk et al. 2018). La casa comercial *PacBio* ha desarrollado plataformas basadas en esta tecnología, disminuyendo las tasas de error y aumentando el tamaño de las lecturas secuenciadas en los nuevos modelos (Quail et al. 2012).
- La tecnología de **secuenciación sintética** (SLR, *synthetic long-reads*) se basa en un sistema de código de barras para asociar fragmentos que se secuencian con técnicas de secuenciación de lecturas cortas. Tras la secuenciación de lecturas cortas, los datos se dividen según su código y se reensamblan con la premisa de que las lecturas que comparten código se derivan de un mismo fragmento original (McCoy et al. 2014). Hay 2 plataformas de secuenciación que usan esta tecnología *Illumina SLR platform* y *10X Genomics emulsion-based system*.
- La tecnología de **secuenciación por nanoporos** se diferencia de los otros en que detecta directamente los nucleótidos sin síntesis activa de ADN, ya que un tramo largo de ADN monocatenario atraviesa un nanoporo proteico estabilizado en una membrana polimérica resistente a la electricidad. Al establecer un voltaje a

través de esta membrana, los sensores detectan los cambios de corriente iónica desplazados por nucleótidos que ocupan el poro en tiempo real a medida que la molécula de ADN pasa a través de él (Kono y Arakawa 2019). Esta metodología ha sido utilizada en las plataformas de *Oxford Nanopore Technologies*.

La figura 3 resume los tipos de tecnologías de secuenciación de ADN de lecturas largas.

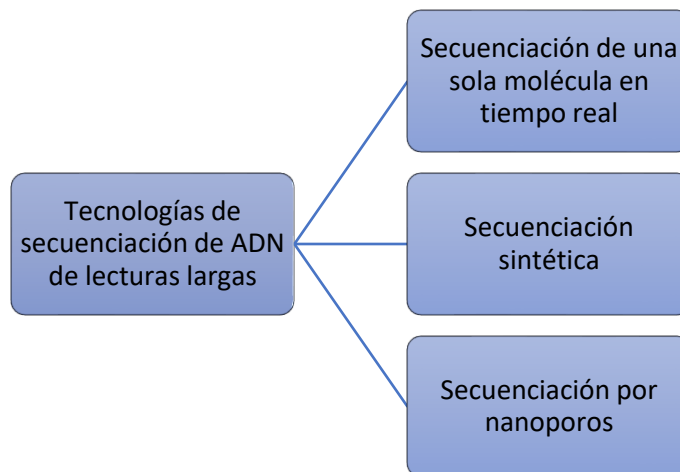


Figura 3: Tipos de tecnologías de secuenciación de ADN de lecturas largas.

Las tecnologías de secuenciación de una única molécula presentan un enorme potencial, como ha quedado patente en el **Consorcio T2T**, cuyo objetivo es obtener el ensamblado de cromosomas de punta a punta resolviendo las regiones no resueltas (genes de ARN ribosómico y satélites). Este consorcio ha logrado numerosos hitos como presentar en 2018 el centrómetro del cromosoma Y (Jain et al. 2018), en 2020 la secuencia completa del cromosoma X (Miga et al. 2020), y en 2021 la secuencia completa del cromosoma 8 (Logsdon et al. 2021), hasta en 2022 conseguir su objetivo con la publicación del genoma humano (Nurk et al. 2022). Esta publicación impulsa la tecnología para aumentar drásticamente el rendimiento de los ensamblajes completos y de calidad de telómero a telómero.

Introducción

1.3.3 Tecnologías de secuenciación de ARN

Las tecnologías de secuenciación de ARN (ARNseq, *ARN sequencing*) comenzaron a desarrollarse hace más de una década, desde entonces se han convertido en una herramienta omnipresente en la biología molecular. Anteriormente, el análisis del ARN se realizaba mediante *microarrays*, herramienta muy útil, pero que está perdiendo importancia con el desarrollo del ARNseq. El flujo de trabajo estándar de esta tecnología se ha mantenido en estos más de 10 años de evolución, comienza con la extracción del ARN en laboratorio, el enriquecimiento de ARNm, ARN pequeños o ARN total (dependiendo del objetivo de nuestro análisis), la síntesis de ADN complementario al ARN extraído (ADNc) y la preparación de una librería de secuencias ligadas a un adaptador (Wang, Gerstein y Snyder 2009).

La ARNseq es más sencilla que la secuenciación de genomas completos, ya que el transcriptoma es más pequeño representando menos del 2% del genoma, y no se necesita conocer el genoma del organismo. Aun así, sus finalidades son múltiples, al ser el objetivo más común de estas investigaciones el análisis de la expresión diferencial de genes, también ha permitido revelar la gran importancia de los mecanismos de corte y empalme en el ARN y de los ARNnc (Marguerat y Bähler 2010).

Dentro de la gran homogeneización de los protocolos de ARNseq, destacada anteriormente, se pueden distinguir tres tipos de plataformas (Mailler et al. 2018):

- **Plataformas ARNseq de lecturas cortas:** la tecnología ARNseq de lecturas cortas se ha convertido gracias a los secuenciadores de la casa comercial *Illumina* en el principal método de ARNseq desde la desaparición de la pirosecuenciación, representando el 95% de los datos de ARN almacenados en el repositorio SRA (*Short Read Archive*). Constituye el método más barato para obtener datos completos y de alta calidad del transcriptoma. El protocolo incluye los pasos principales anteriormente enumerados (extracción de ARN, enriquecimiento de ARN, síntesis de ADNc, ligación de adaptadores y amplificación por PCR), tras ello se realiza el proceso de secuenciación y análisis. Este protocolo genera fragmentos de ADNc de unas 200 pb de longitud debido a la fragmentación del ARNm. En cada librería se obtiene una media de 20 a 30 millones de lecturas. La

principal limitación de este tipo de secuenciación es la capacidad de identificar y cuantificar correctamente las isoformas.

- **Plataformas de ARNseq de una única molécula:** aunque la secuenciación de *Illumina* sea la dominante, *PacBio* ofrece una tecnología alternativa de lectura larga que permite la secuenciación de moléculas completas de ARN tras su conversión en ADNc. Al eliminar la necesidad de ensamblar las lecturas cortas de ARN, estos métodos superan algunos de los problemas asociados a las lecturas cortas. Por ejemplo, se reduce la ambigüedad en el mapeo de las lecturas de secuencia, y se pueden identificar transcritos más largos, lo que conduce a una captura más completa de la diversidad de isoformas.
- **Plataformas de secuenciación directa de ARN:** Tanto los métodos de lectura corta como los de única molécula se basan en la conversión del ARNm en ADNc. La plataforma de secuenciación *Oxford Nanopore* puede secuenciar directamente el ARN sin síntesis de ADNc ni amplificación por PCR durante la preparación de la biblioteca, gracias a su tecnología de secuenciación por nanoporos (Leonardi y Leger 2021). Esta estrategia elimina los sesgos generados por la amplificación y síntesis de ADNc, permite conservar la información epigenética y estudiar las modificaciones postranscripcionales de los transcritos.

La figura 4 resume los tipos de tecnologías ARNseq.

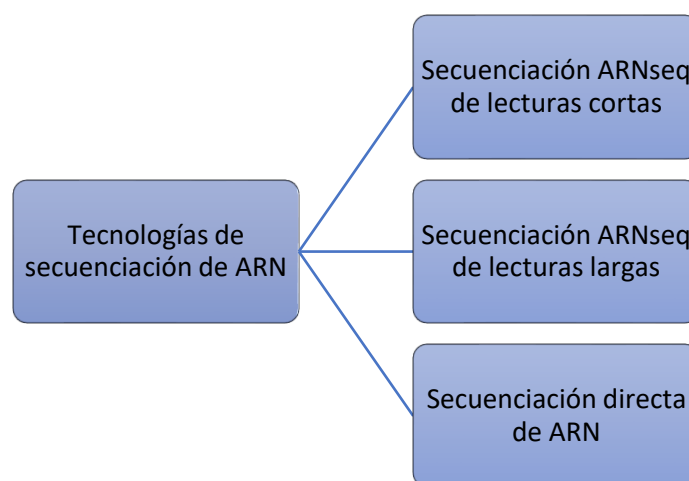


Figura 4: Tipos de tecnologías de secuenciación de ARN.

Introducción

1.4 La bioinformática

Numerosos hechos históricos precedieron a la aparición de la bioinformática. El trabajo pionero de Margaret Oakley Dayhoff para la determinación de la estructura 3D de proteínas, lleva a considerarla la **madre de la bioinformática** (Gauthier et al. 2019). Dayhoff junto a otros autores publicó el libro *Atlas of Protein Sequence and Structure*, considerado uno de los pilares fundamentales de la bioinformática, y desarrolló el programa **COMPROTEIN**, un ensamblador de novo de secuencias proteicas (Dayhoff y Ledley 1962), que era el único tipo de secuenciación realizada hasta ese momento. Además, Dayhoff generó a finales de los años 1960 la primera base de datos bioinformática al reunir todos los datos de secuenciación proteica generados hasta ese momento.

En 1970, Needleman y Wunsch desarrollaron el primer algoritmo de programación para la alineación de secuencias proteicas por pares. A pesar de este avance, no fue hasta principios de la década de 1980 cuando surgieron los primeros **algoritmos de alineamiento de secuencias** múltiples (MSA, *multiple sequence alignment*). El primer algoritmo de MSA consistía en una generalización del algoritmo propuesto por Needleman y Wunsch y resultó poco práctico en términos de rendimiento. El primer enfoque verdaderamente práctico del MSA fue desarrollado por Da-Fei Feng y Russell F. Doolite en 1987. Con una simplificación del algoritmo, Feng-Doolite desarrolló en 1988 el programa informático **CLUSTAL**, utilizado hasta hoy en día para el alineamiento de secuencias (Gauthier et al. 2019).

En 1978, Dayhoff, Schwartz y Orcutt contribuyeron a otro hito de la bioinformática al desarrollar el primer modelo probabilístico de sustituciones de aminoácidos. Basado en la observación de 1572 mutaciones puntuales aceptadas (PAM, *point accepted mutations*) de 71 familias de proteínas que comparten más del 85% de identidad. Como resultado obtuvieron una matriz que contenía los valores de probabilidad basados en las mutaciones observadas para cada aminoácido, es decir, la probabilidad para cada aminoácido de cambiar tras un intervalo evolutivo determinado. Esta matriz se denominó **matriz PAM** e introdujo el número de sustituciones de aminoácidos como medida de cambio evolutivo (Hogeweg 2011).

1.4.1 Desarrollo de la computación y su aplicación en las metodologías de secuenciación

Junto con el desarrollo de las metodologías de secuenciación, se destaca como elemento primordial de la evolución de la bioinformática el desarrollo de la computación. Sin duda, en primer lugar, cabe destacar a Ada Lovelace bautizada como la **primera programadora** por ser la primera persona en describir un lenguaje de programación. En 1946 se desarrolló la primera computadora llamada **ENIAC** (*Electronic Numerical Integrator And Computer*). Desde entonces el crecimiento en la capacidad de los ordenadores ha sido exponencial, hoy España cuenta con 16 supercomputadores situados en diferentes centros de investigación y universidades, pertenecientes a la **Red Española de Supercomputación** (RES). Dentro de esta red, liderada por el computador Mare Nostrum del Barcelona Supercomputing Center, se localiza en la Universidad de Málaga el **superordenador Picasso**, instalado en 1997, que contaba con 10136 cores de cómputo y 77440 GB de RAM, de HDR200/HDR100 cuando se realizó esta tesis, y que recientemente ha cuádruplicado su capacidad de cómputo y almacenamiento.

Es imprescindible el desarrollo de la computación a la par que el de las metodologías de secuenciación, ya que sin este desarrollo no sería posible el análisis y la comprensión de los datos obtenidos. El tipo de procesamiento de los datos variará según la finalidad del análisis. Sin embargo, se puede establecer un flujo de procesamiento inicial común que dependerá del tipo de molécula secuenciada y cuyos ficheros finales serán los que se utilicen en sucesivos y distintos programas dependiendo del objetivo del análisis bioinformático.

En el procesamiento bioinformático de los datos obtenidos por **secuenciación de ADN** se deben de destacar las siguientes fases (Bao et al. 2014):

- **Preprocesamiento de las lecturas:** se eliminan los adaptadores (para ello se deberá de tener en cuenta la plataforma de secuenciación utilizada), los posibles contaminantes y las lecturas que no superen un umbral de calidad preestablecido. Se realiza sobre un archivo en formato .fastq (Cock et al. 2010) que se genera por los propios secuenciadores, y contiene 4 líneas por cada una de las secuencias que corresponden al identificador, la secuencia, el identificador

y la calidad. La calidad es un valor determinado por la probabilidad de que la base sea incorrecta, y se calcula con el coeficiente de calidad Phred. Muchas herramientas se han desarrollado para realizar este preprocesamiento de las lecturas, por ejemplo, SeqTrimNext, FastQC, FastQ Screen, FASTX-Toolkit, PRINSEQ, Cutadapt, Trimmomatic (Garber et al. 2011) y la recientemente publicada ultrarrápida fastp (Chen et al. 2018).

- **Ensamblaje con una secuencia de referencia o ensamblaje *de novo*:** en el ensamblaje se distinguen dos etapas (ensamblado de las lecturas en *contigs* y ensamblado de los *contigs* en *scaffolds*). Esta fase es de una gran complejidad computacional debido a la gran cantidad de fragmentos, solamente mantener en memoria los fragmentos y el genoma humano ocupa 12 GB. Además, un genoma contiene gran cantidad de secuencias repetidas de las cuales es difícil determinar la posición. En el caso de contar con un genoma de referencia y optar por el **ensamblaje con una secuencia de referencia**, debe realizarse un alineamiento y un mapeo. El algoritmo más ampliamente utilizado para el mapeo es la transformación de Burrows-Wheeler, una técnica para optimizar la memoria y detectar los alineamientos. Este algoritmo es utilizado por varios programas de mapeo con genomas de referencia como *Burrows-Wheeler Aligner* (BWA) y Bowtie. El formato de alineamiento resultante es .sam que se comprime a .bam (Chen, Shi y Shi 2017; Costa-silva, Domingues y Lopes 2017; Martin y Wang 2011; Carruthers et al. 2018; Garber et al. 2011). Para el **ensamblaje *de novo*** (sin genoma de referencia) se ha mejorado el algoritmo de grafos de De Bruijn, implementando numerosos programas como SPADes, SOAP, Velvet y ABySS.
- **Anotación de genomas:** es el proceso de búsqueda de elementos funcionales en el genoma ensamblado. Normalmente esta anotación se realiza para genes codificadores de proteínas, pero también puede incluir transcritos no codificantes, islas CpG (regiones donde existe una gran concentración de pares de citosina y guanina) o regiones de hipersensibilidad a las DNAsas. Esta tarea suele dividirse en dos categorías: la **predicción *ab initio*** (predicción computacional de la estructura exón-intrón mediante modelos estadísticos) y los **modelos basados en el alineamiento de secuencias**. Algunos métodos de

anotación combinan ambas fuentes de predicción para generar un conjunto final de anotaciones. Los programas de anotación más populares son: ROSETTA, SGP-1/-2, TWINSKAN, SLAM, EvoGene, ExoniPhy, DOGFISH, N-SCAN, CONTRAST y DOUBLESCAN (Armstrong et al. 2019).

En el procesamiento bioinformático de los datos obtenidos por **ARNseq** se deben destacar las siguientes fases:

- **Preprocesamiento de las lecturas:** este procesamiento en el caso de las secuencias obtenidas por secuenciación de ARN es igual al expuesto para la secuenciación de ADN. No obstante, deben de establecerse los parámetros adecuados para la eliminación de adaptadores y el umbral de calidad de las muestras.
- **Ensamblaje de transcriptomas:** las herramientas mencionadas anteriormente para el ensamblaje de los genomas no pueden usarse en transcriptomas por tres motivos: la profundidad de los transcritos varía en varios órdenes de magnitud, hay solapamiento de transcritos en sentido y antisentido, y hay exones que no estarán en todos los transcritos de un mismo gen. Las estrategias para el ensamblaje de transcriptomas pueden estar basadas en un transcriptoma de referencia o seguir una estrategia de ensamblaje *de novo* (Martin y Wang 2011; Raghavan et al. 2022).
 - **Ensamblaje de transcriptomas con transcriptoma de referencia:** En primer lugar, las lecturas de ARNseq se mapean con un genoma de referencia utilizando un mapeador que tenga en cuenta el empalme, como STAR o HISAT2. En segundo lugar, las lecturas solapadas de cada locus se agrupan para construir un grafo que represente todas las isoformas posibles. El último paso consiste en recorrer el grafo para resolver las isoformas individuales. Algunos ejemplos de programas que utilizan esta estrategia son Cufflinks y Scripture.
 - **Ensamblaje de transcriptomas *de novo*:** no utiliza un genoma de referencia, aprovecha la redundancia de la secuenciación de lectura corta para encontrar solapamientos entre las lecturas y ensamblarlas en transcritos. Los ensambladores *de novo* Rnnotator, Multiple-k y Trans-

ABYSS siguen la misma estrategia: ensamblan el conjunto de datos varias veces utilizando una estrategia basada en el grafo de De Bruijn para reconstruir transcritos a partir de una amplia gama de niveles de expresión y, a continuación, postprocesan el ensamblaje para fusionar *contigs* y eliminar la redundancia. Por el contrario, otros ensambladores, como Trinity y Oases, recorren directamente el grafo de De Bruijn para ensamblar cada isoforma. Las lecturas más largas de tecnologías de secuenciación más desarrolladas también pueden integrarse en los ensamblajes *de novo* del transcriptoma, lo que mejora la capacidad de resolver isoformas alternativas.

- **Cuantificación de la expresión** (Garber et al. 2011): La expresión de un gen se define como la suma de la expresión de todas sus isoformas. Cuando se utiliza la secuenciación de ARN para estimar la expresión génica, los recuentos de lecturas deben **normalizarse** adecuadamente para extraer las estimaciones de expresión significativas. Existen dos **fuentes de variabilidad sistemática** principales que requieren normalización. En primer lugar, la fragmentación del ARN durante la construcción de la librería hace que los transcritos más largos generen más lecturas en comparación con los transcritos más cortos. En segundo lugar, la variabilidad en el número de lecturas producidas para cada ejecución causa fluctuaciones en el número de fragmentos mapeados a través de las muestras. Para tener en cuenta esta variabilidad, se utilizó la fórmula de lecturas por kilobase de transcrito por millón de lecturas mapeadas (RPKM, *read per kilobase per million*) que normaliza el recuento de lecturas de un transcrito tanto por su longitud como por el número total de lecturas mapeadas en la muestra. Cuando los datos proceden de secuenciación por pares, la métrica análoga a RPKM son los fragmentos por kilobase de transcrito por millón de lecturas mapeadas (FPKM, *fragments per kilobase per million*).

Como muchos genes tienen múltiples isoformas, muchas de las cuales comparten exones, algunas lecturas no pueden asignarse inequívocamente a un transcrito. Esta incertidumbre en la asignación de lecturas afecta a la precisión de la cuantificación de la expresión. Una estrategia consiste en estimar los valores de expresión a nivel de isoforma contando sólo las lecturas que se

asignan unívocamente a una única isoforma. Aunque esto funciona para algunos genes empalmados alternativamente, falla para genes que no contienen exones únicos a partir de los cuales estimar la expresión de isoformas. Los métodos alternativos denominados **métodos de expresión de isoformas**, como Cufflinks, manejan la incertidumbre construyendo una función de verosimilitud que modela el proceso de secuenciación e identifica las estimaciones de abundancia de isoformas que mejor explican las lecturas obtenidas. Esta estimación se denomina estimación de máxima verosimilitud (MLE, *maximum likelihood estimation*). Para genes poco expresados, la MLE no es una estimación precisa de la expresión; la inferencia Bayesiana mejora la robustez de la cuantificación de la expresión mediante el muestreo de estimaciones alternativas de abundancia alrededor de la MLE, proporcionando al mismo tiempo una medida de confianza en la estimación.

- **Análisis de expresión diferencial** (Garber et al. 2011): los métodos iniciales modelaban las cuentas de lecturas observadas utilizando distribuciones basadas en el recuento, como la distribución de Poisson. Sin embargo, estas distribuciones no tienen en cuenta la variabilidad biológica entre muestras. Para superar este hándicap y proporcionar una medida de significación en ausencia de un gran número de réplicas biológicas, se utilizan programas que modelan la varianza de recuento entre réplicas como una función no lineal de los recuentos medios distribuidos como una binomial negativa y utilizando diferentes estrategias paramétricas. Algunos de estos programas son DESeq, EdgeR y Cuffdiff.

Tanto si el análisis bioinformático se ha desarrollado a partir de datos de secuenciación de ADN o de ARNseq, el último paso es su **interpretación funcional**. Las posibles alteraciones detectadas en los pasos anteriormente detallados pueden tener consecuencias a nivel biológico, pudiendo llegar a ser patológicas. Para interpretar estas repercusiones biológicas, hay que usar un vocabulario unificado fácilmente analizable por un ordenador. Los vocabularios más extendidos y útiles son los GO (*Gene Ontology*) y KEGG (*Kyoto Encyclopedia of Genes and Genomes*).

Introducción

- GO: vocabularios y clasificaciones estructurados que abarcan varios ámbitos de la biología molecular y celular para la anotación de genes, productos génicos y secuencias. La base de datos GO integra los vocabularios y las anotaciones dando acceso a esta información en varios formatos. Esta base de datos ofrece dos recursos principales: la ontología GO (la estructura lógica que describe toda la complejidad de la biología, compuesta por las clases de las diferentes funciones biológicas, las rutas por las que se llevan a cabo y las localizaciones celulares en las que se producen), y el corpus de anotaciones GO (las afirmaciones rastreables que relacionan un producto génico o un gen con términos ontológicos para describir su función biológica normal). En conjunto, la ontología y las anotaciones proporcionan un modelo completo de los sistemas biológicos (Blake y Harris 2002).
- KEGG: base de datos desarrollada para representar la información biológica en términos de interacciones moleculares y redes de interacciones. Los datos recogidos son tomados de artículos publicados, sin contar con modelos de inteligencia artificial o aprendizaje automático. Usado para analizar posibles funciones a nivel celular o de organismo ocultas en secuencias de genomas y metagenomas (Kanehisa et al. 2023).

Para analizar esta información funcional existen numerosas bases de datos, entre las que destacan:

- DAVID (*Database for Annotation, Visualization and Integrated Discovery*): proporciona un conjunto de seis herramientas de anotación funcional que se basan en múltiples fuentes de anotaciones funcionales y cuya finalidad es comprender el significado biológico de grandes listas de genes (Dennis et al. 2003). Estas herramientas son: agrupación de anotaciones funcionales, gráfico de anotaciones funcionales, tabla de anotaciones funcionales, clasificación funcional de genes, conversor de identificadores de genes y visor de la lista de genes (Sherman et al. 2022).
- String: base de datos que integra todas las asociaciones conocidas y previstas entre las proteínas codificadas, incluidas tanto las interacciones físicas como las asociaciones funcionales (Szklarczyk et al. 2021).

- COSMIC (*Catalogue Of Somatic Mutations In Cancer*): base de datos que incluye casi 6 millones de mutaciones codificantes en 1,4 millones de muestras tumorales, seleccionadas a partir de más de 26.000 publicaciones. Además de las mutaciones codificantes, COSMIC abarca todos los mecanismos genéticos por los que las mutaciones somáticas promueven el cáncer, incluidas las mutaciones no codificantes, las fusiones de genes, las variantes del número de copias y las mutaciones de resistencia a fármacos (Tate et al. 2019).
- ExoRbase: repositorio de ARNlnc (ARN largo no codificante), ARNm y ARNcirc (ARN circulares) detectados en vesículas extracelulares en análisis de diferentes fluidos humanos (Lai et al. 2022).
- Lncbook: base de datos que incluye anotaciones de ARNlnc a diferentes niveles ómicos, permitiendo establecer posibles firmas funcionales de ARNlnc en enfermedades humanas y cualquier contexto biológico (Li et al. 2023).
- RNAInter (*RNA Interactome Database*): base de datos que integra datos de interactomas de ARN validados experimentalmente y predichos computacionalmente a partir de la bibliografía y otras bases de datos (Kang et al. 2022)
- RISE (*RNA Interactome from Sequencing Experiments*): base de datos de interactomas de ARN obtenidos de experimentos de secuenciación. Contiene interacciones de ARNm, miARN (microARN)y ARNnc de humanos, ratones y levaduras (Gong et al. 2018).

1.4.2 Explosión de la bioinformática en las últimas décadas

Con el gran desarrollo de las tecnologías de secuenciación y la computación, descritas en los apartados anteriores de esta introducción, no es de extrañar que la expansión de la bioinformática haya ido al unísono.

La secuenciación del primer genoma humano que se publicó en 2001 tuvo un coste entre 500 millones y 1000 millones de dólares. Un año después de la publicación tras la puesta a punto de la secuenciación y el procesamiento, el costo cayó hasta los 150 millones de dólares por genoma humano. Actualmente, la secuenciación de un genoma humano

Introducción

completo ronda los 1000 dólares, según el *National Institutes of Health*. Esta disminución de los costes se plasma en la figura 5. Con la creciente escala de los estudios de genética humana y el número cada vez mayor de aplicaciones clínicas de la secuenciación genómica, se está prestando aún más atención a la comprensión de los costes subyacentes de la generación de una secuencia del genoma humano. Determinar con precisión el coste de secuenciar un genoma no es sencillo. Hay muchos parámetros que definir y matices que considerar. De hecho, es difícil citar cifras precisas sobre el coste de la secuenciación del genoma que signifiquen lo mismo para todo el mundo porque, en realidad, los distintos investigadores, instituciones de investigación y empresas suelen registrar y contabilizar esos costes de diferentes maneras. En los costes no se suele incluir actividades bioinformáticas, como el control de calidad, la alineación de la secuencia generada con un genoma de referencia, el ensamblaje de secuencias, la determinación de variantes genómicas o la anotación.

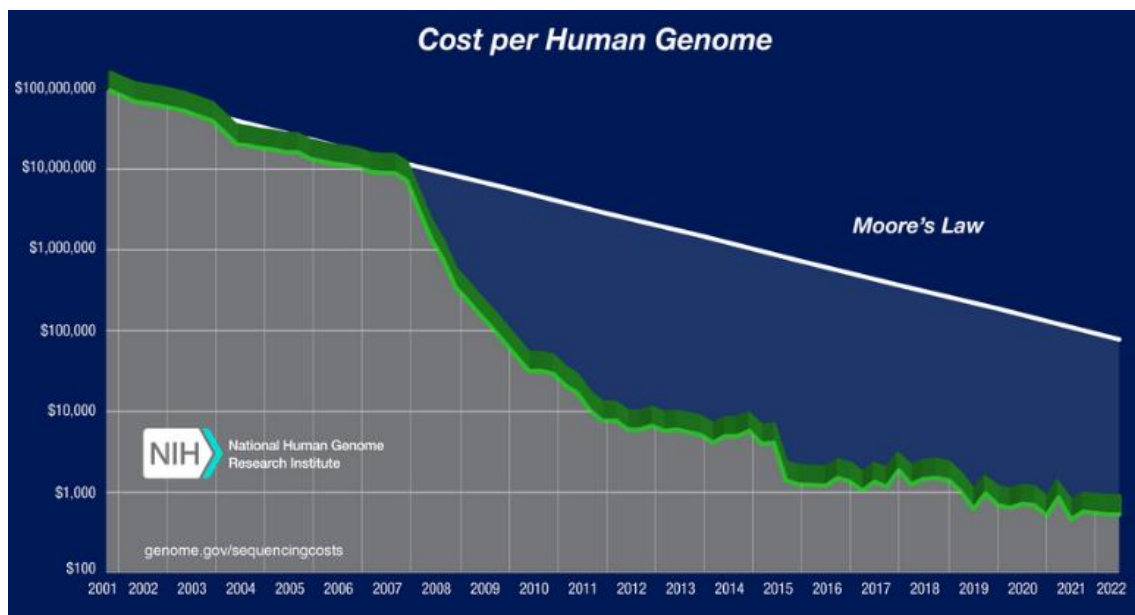


Figura 5: Coste de la secuenciación del genoma humano. Ref: National Institutes of Health.

Esta **disminución de costes** junto a los avances de las tecnologías ómicas, ha generado un aumento exponencial de generación de datos y de las secuencias publicadas en los repositorios bioinformáticos. La comunidad científica ha generado datos que superan el nivel de los exabytes (1.018 bytes) (Gauthier et al. 2019). Se necesitan grandes recursos informáticos para almacenar, organizar y facilitar el acceso y el uso de toda la información publicada.

Esta expansión se puede corroborar también a nivel de publicaciones e interés científico en la materia. En el repositorio de publicaciones científicas PubMed el número de publicaciones indexadas con el término “bioinformatics” pasó de 13.974 en 2009 a 47.197 en 2020 según la base de datos PubMed.

1.5 Evaluación e implementación de la medicina personalizada de precisión.

Tras este periplo de desarrollo de las tecnologías ómicas, fundamentado en los avances de las tecnologías de secuenciación y en la bioinformática, se llega al concepto actual de MPP. Sin embargo, tras este desarrollo y conceptualización es necesario que exista un proceso de implementación en los sistemas sanitarios para que estas innovaciones se trasladen de forma real al paciente.

La **MPP** supone un cambio total en el paradigma de la medicina clásica basada principalmente en la sintomatología fenotípica y en tratamientos cuya efectividad desigual entre pacientes se basaba en el ensayo-error (Visvikis-Siest, Gorenjak y Stathopoulou 2018). Estas prácticas podían tener como resultado la falta de diagnóstico para pacientes con enfermedades poco prevalentes o enfermedades con expresividad fenotípica desigual entre los afectados, la falta de respuesta al tratamiento de un importante porcentaje de los pacientes tratados, la posibilidad de reacciones adversas no esperadas y la incapacidad para establecer planes de predicción y prevención de enfermedades con un componente genético.

La implementación de la MPP en los sistemas sanitarios frente a la medicina clásica tiene como objetivo eliminar o mitigar los resultados perjudiciales para los pacientes anteriormente detallados, consiguiendo una medicina P4. El concepto **medicina P4** nace de la necesidad de que la medicina discurra a ser calificada con estas cuatro palabras que empiezan por p: predictiva, preventiva, personalizada y participativa (Sagner et al. 2017). Según los autores que acuñaron este concepto, la medicina P4 es “un marco para promover y facilitar la colaboración de múltiples partes interesadas con un lenguaje común orquestado y un modelo de atención integrada para aumentar la duración de la salud” (Sagner et al. 2017). Este desarrollo hasta la medicina P4 se plasma en la figura 6.

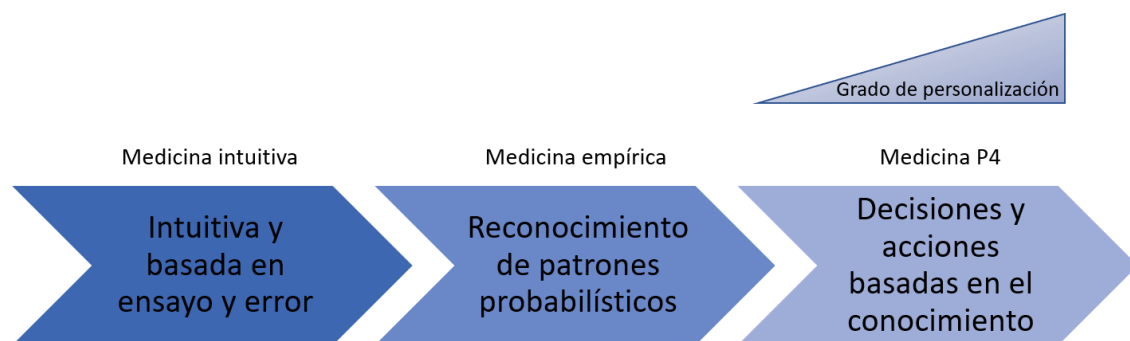


Figura 6: Transición a la medicina P4.

Para conseguir un alto grado de implementación de la MPP y lograr los objetivos de la medicina P4 debe realizarse un **análisis del sistema sanitario** en su conjunto. Este enfoque holístico podría permitir la detección de las necesidades que surgirían en el sistema con esta implementación para que fueran tenidas en cuenta al elaborar las políticas sanitarias (BigMed y Lighthouse 2021).

Dada la gran envergadura que suponen estos cambios en los sistemas sanitarios, han surgido numerosos **proyectos y redes nacionales e internacionales** para analizar y posibilitar esta implementación. Destacan por ejemplo las iniciativas: *European Alliance for Personalised Medicine*, *ICPerMed*, *2025 France Genomic Medicine Initiative* y *All of Us*. Todas ellas comparten algunos rasgos comunes, a pesar de la diversidad entre los sistemas sanitarios. Entre esas similitudes cabe destacar la necesidad de creación de comités interdisciplinares, el sustento de una financiación suficiente y estable, y la concienciación y participación de científicos, clínicos, administrativos, políticos y ciudadanía de los beneficios de esta implementación.

Las iniciativas de implementación de MPP contemplan diversas finalidades terapéuticas (diagnóstica, predictiva presintomática, análisis de portadores, diagnóstica prenatal, diagnóstica preimplantacional o farmacogenética) y su aplicación para diferentes enfermedades. En todas ellas se indica que es fundamental su uso en el **cáncer**. Este aspecto común de todas las iniciativas es debido a la profunda caracterización molecular y genética del cáncer y a su alta prevalencia en la población mundial (en 2020 se contabilizaron aproximadamente 19,3 millones de casos nuevos a nivel mundial, según últimos datos disponibles en el proyecto GLOBOCAN) (Sung et al. 2021).

1.6 Evaluación en el manejo del cáncer de pulmón en el contexto de la MPP

El cáncer es una enfermedad genética. Esta premisa que puede parecer simple ha llevado a los científicos décadas de investigación sobre los mecanismos biológicos que subyacen en ella. La acumulación de alteraciones moleculares en el genoma es la que produce la progresión del tumor. Estas alteraciones pueden ser heredadas o producirse de manera aleatoria como resultado de errores en la mitosis o por exposición a agentes carcinogénicos (Flores-Pérez et al. 2019).

Para poder afirmar que el cáncer, al igual que otras enfermedades, tiene origen genético fue necesario tener la secuencia de un **genoma de referencia** para establecer las diferencias con el genoma de las células tumorales, y detectar las alteraciones del genoma que convierten una célula sana en una tumoral. Debe tenerse en cuenta en los estudios genéticos que estas alteraciones pueden diferir entre las distintas poblaciones humanas convergiendo en un mismo proceso oncológico. Además, es importante distinguir en el ADN tumoral las mutaciones inductoras (*driver mutations*) de las mutaciones pasajeras (*passenger mutations*). Las mutaciones inductoras son las que provocan el comportamiento anómalo de las células tumorales que lleva a la formación del tumor, mientras que las mutaciones pasajeras se deben a la alteración de los mecanismos de regulación en las células tumorales, pero no contribuyen a la formación, crecimiento o expansión tumoral (Pon y Marra 2015).

La aplicación de estos conocimientos a la práctica clínica conlleva la **implementación de la MPP** en el campo oncológico para la prevención, diagnóstico, pronóstico y adecuación del tratamiento. Dentro del campo oncológico y las diversas aplicaciones de la MPP, este trabajo se centrará en el diagnóstico y adecuación del tratamiento del cáncer de pulmón.

El **cáncer de pulmón** es una enfermedad que se consideraba rara a principios del siglo pasado, pero que con la difusión del tabaco aumentó su incidencia hasta convertirse en un problema de salud pública. El número de pacientes diagnosticados con cáncer de pulmón en 2020 a nivel mundial fue de 2.206.771, convirtiéndose en el segundo tipo de cáncer más prevalente tras el de mama, ya que representa el 11,4% de los casos totales de cáncer diagnosticados. En España, el número de nuevo diagnosticados en 2021 fue

Introducción

de 29.549 pacientes. En cuanto a mortalidad, el cáncer de pulmón es el más letal tanto nivel mundial (1.796.144 fallecimientos en 2020) como de España (22.930 fallecimientos en 2020) (Sociedad Española de Oncología Médica 2021).

Las cifras expuestas a nivel de diagnósticos y fallecimientos muestran el problema de salud pública que supone la enfermedad, lo cual hace apremiante su investigación a nivel básico y traslacional. Además, resulta acuciante que estas mejoras sean trasladadas de forma eficiente a la práctica clínica, y es en este punto de traslado donde radica la importancia de la MPP para el cáncer de pulmón.

Cabe destacar que la aplicación de la MPP tanto en el cáncer de pulmón como en cualquier otra enfermedad no supone la sustitución total de las metodologías diagnósticas usadas hasta ese momento, sino una actualización en su aplicación o interpretación de resultados y la posibilidad de complementarse con nuevas pruebas.

1.6.1 Alteraciones genéticas en el cáncer de pulmón

Algunas de las alteraciones conocidas en el cáncer de pulmón hacen aumentar la susceptibilidad de padecerlo a lo largo de la vida, otras determinan el tipo de tratamiento a aplicar y otras establecen el pronóstico del paciente. Por tanto, conocer estas alteraciones y su significado puede tener una aplicación directa sobre la prevención, el diagnóstico, el pronóstico y la adecuación del tratamiento del paciente con cáncer de pulmón.

Las alteraciones genéticas presentes en las células tumorales, tanto de cáncer de pulmón como de cualquier otro proceso oncológico, pueden afectar directamente a la secuencia del genoma o a su conformación (Pecorino 2021):

- **Alteraciones a nivel de la secuencia:** estas alteraciones son las conocidas como mutaciones y suponen cambios en la secuencia de nucleótidos del genoma celular. Las mutaciones pueden ser desde grandes traslocaciones, inversiones o deleciones de regiones genómicas hasta cambios de un solo nucleótido (SNP, *single nucleotide polymorphism*). Las regiones del genoma se pueden dividir en dos subtipos principales, al ser posibles las mutaciones en ambas:

- Las **regiones transcripcionalmente activas**: regiones del genoma que generan transcritos de ARN que pueden ser a su vez:
 - ARN codificante: se le denomina ARNm porque se traduce a proteínas.
 - ARNnc: no generan proteínas. Durante mucho tiempo se desconoció su funcionamiento, pero ahora se sabe que tienen funciones reguladoras al interactuar directa o indirectamente con el ADN o con otros tipos de ARN. Existen muchos tipos de ARNnc y se clasifican en función del número de su longitud (largos > 200 pb y pequeños < 200 pb) y de su conformación, entre ellos destacan: miARN, ARNlnc, ARN circular o ARNpi (ARN pequeño de interferencia).
- Las **regiones transcripcionalmente inactivas**: no generan ningún tipo de transcritos, sino que tienen una función reguladora, principalmente sobre la transcripción de otras regiones del genoma cercanas, pero también pueden afectar a regiones alejadas.
- **Alteraciones de la conformación del ADN**: se denominan también alteraciones epigenómicas y no afectan a la secuencia del ADN, pero sí a su estructura y conformación por lo que pueden alterar la transcripción. Pueden afectar a las histonas o a la metilación del ADN.

Las tecnologías genómicas se utilizan para detectar cambios de secuencia del ADN, las transcriptómicas para detectar cambios en los perfiles de expresión de los ARN (codificantes o no codificantes) y las epigenómicas para los cambios de conformación del ADN. El grado de desarrollo y uso en la práctica clínica de estas tecnologías es desigual, estando algunas de ellas en fase de implementación y otras en fase de investigación. Esta tesis se centra en el uso de estas tres tecnologías para su aplicación en el cáncer de pulmón:

- **Tecnologías genómicas**: los pacientes de cáncer de pulmón pueden presentar una serie de alteraciones ómicas, desde reordenamientos a nivel cromosómico, detectables con un cariotipo, hasta mutaciones puntuales en genes. Entre los genes para los que existen más resultados de mutaciones destacan: *EGFR*, *BRAF*

Introducción

(codifica la proteína B-Raf), *KRAS* (codifica la proteína K-Ras perteneciente a la ruta de señalización RAS/MAPK) y *ALK* (codifica el receptor de la tirosina quinasa) (Lyu et al. 2019).

- **Tecnologías transcriptómicas:** los pacientes de cáncer de pulmón pueden presentar alteraciones en los ARNm o en los ARNnc, como los miARN, ARNlnc o ARN circulares (Yi et al. 2021; Wu et al. 2020).
- **Tecnologías epigenómicas:** las alteraciones epigenómicas podrían conducir a desregulaciones de la expresión. Suelen presentarse como regiones hipermetiladas o hipometiladas, principalmente relacionadas con la regulación de genes involucrados en la iniciación o desarrollo del tumor, como la metilación de regiones promotoras de genes supresores de tumores (Fujiwara et al. 2005).

A nivel transcriptómico, numerosos estudios han evaluado la correlación de diferentes tipos de ARN en función a sus niveles de expresión. Estos estudios han demostrado correlaciones en los niveles de expresión diferencial en ARNm-miARN-ARNlnc entre tejido tumoral de cáncer de pulmón y tejido sano de los mismos pacientes (Wang et al. 2020; 2021). La existencia de estas correlaciones pone de relieve la función reguladora de la transcripción que tienen los miARN y los ARNlnc (Li et al. 2018) (Mattick et al. 2023):

- Los **miARN** son una clase de ARN monocatenarios no codificantes de 19-25 nucleótidos que regulan transcripcionalmente o postranscripcionalmente la expresión génica mediante su unión a un ARNm diana e influye en su degradación y traducción. Cada vez hay más pruebas de que los niveles aberrantes de miARN están relacionados con la proliferación, la angiogénesis y la metástasis en varias enfermedades humanas. Además, los miARN han suscitado una gran atención como diana en terapias contra el cáncer o como marcadores de diagnóstico o pronóstico.
- Los **ARNlnc** son ARN no codificantes cuya longitud oscila entre 200 nucleótidos y las 100kb. Muchos de los ARNlnc son específicos del tipo de célula y regulan muchos aspectos de la diferenciación celular, el desarrollo y otros procesos fisiológicos. Los ARNlnc tienen funciones mecánicamente diversas en la célula, y en el núcleo celular se ha demostrado que los ARNlnc regulan la expresión génica

en *cis* o *trans* reclutando un complejo modificador de la cromatina al promotor de un gen diana. También tienen papeles importantes en el citoplasma, incluso en la regulación de la traducción, el metabolismo y la señalización.

1.6.2 Aplicación de la MPP en el diagnóstico de cáncer de pulmón

Las principales técnicas clásicas de diagnóstico para el cáncer de pulmón se han basado en pruebas de imagen. Entre estas **pruebas de imagen** destacan: la radiografía de tórax, la tomografía computerizada, la resonancia magnética, la ecografía endobronquial y la tomografía por emisión de positrones. Una de las principales desventajas que pueden tener estas técnicas es que en algunos casos son prescritas cuando el paciente ha presentado síntomas compatibles con el cáncer de pulmón, lo que en algunos casos puede suponer un estadio de desarrollo tumoral alto. Este posible retraso es principalmente relevante en el cáncer de pulmón cuyo primer síntoma suele ser la tos, pero muchos de los pacientes se encuentran afectados previamente por enfermedad pulmonar obstructiva crónica (EPOC) debido al hábito tabáquico, por lo que puede ser más difícil identificar la primera sintomatología por parte del paciente (Cooley 2000).

El siguiente paso diagnóstico, en caso de detectar una posible masa tumoral, consiste en la toma de una muestra de dicha masa (**biopsia tisular**), para lo que se dispone de diversas metodologías:

- Citología del esputo: muestra de esputo (mucosidad que al toser sale de los pulmones) se examina en un laboratorio para determinar si tiene células cancerosas. Esta prueba es más probable que ayude a encontrar cánceres que se originan en las vías respiratorias principales del pulmón, como los cánceres de pulmón de células escamosas. Aunque presenta una alta especificidad, su sensibilidad es baja, de ahí que no sea una prueba idónea para el cribado de cáncer de pulmón. Si el médico sospecha cáncer de pulmón, se realizarán más pruebas, incluso si no se encuentran células cancerosas en el esputo.
- Toracocentesis: si se ha acumulado líquido alrededor de los pulmones (un derrame pleural), los médicos pueden extraer líquido para su estudio citológico

Introducción

en el laboratorio con el fin de identificar células neoplásicas y determinar si dicho derrame se debe a la propagación del cáncer al revestimiento de los pulmones (pleura).

- Biopsia con aguja: son las técnicas más ampliamente utilizadas, ya que permiten una mejor obtención de células para su análisis. Se usa una aguja hueca para obtener una muestra pequeña de la posible masa tumoral. La principal ventaja de las biopsias con aguja consiste en que no requieren de una intervención quirúrgica. Sin embargo, en ocasiones la muestra obtenida puede no ser suficiente para el diagnóstico o para la evaluación de biomarcadores necesarios para el tratamiento del cáncer de pulmón (Aguilera-Cobos et al. 2022a):
 - Punción aspirativa con aguja fina (PAAF): la aspiración puede ser transtraqueal o transbronquial con aguja fina. Se realiza pasando la aguja a través de la pared de la tráquea o de los bronquios durante una broncoscopia o ecografía endobronquial, obteniéndose una muestra citológica para su estudio.
 - Biopsia con aguja gruesa (BAG): se usa una aguja de mayor calibre para extraer uno o más pequeños cilindros de tejido. Las muestras obtenidas mediante BAG incluyen una mayor cantidad de tejido, y a menudo son preferidas porque son más representativas de la lesión que las obtenidas mediante PAAF.
 - Biopsia transtorácica con aguja: utilizada cuando el tumor se encuentra en la parte exterior de los pulmones. La aguja se introduce a través de la piel y se dirige hacia la posible masa tumoral con monitorización de los pulmones mediante fluoroscopia o tomografía computarizada.

Una vez tomada la muestra se procede su **examen histopatológico**. Este análisis en laboratorio puede estar precedido por un análisis in situ en el lugar de biopsiado realizado por un anatomopatólogo. A esta evaluación previa se le conoce como **técnica ROSE** (*rapid on-site evaluation*), y se ha demostrado el aumento del rendimiento de las muestras biopsiadas principalmente para los casos de cáncer de pulmón. Según su histología el cáncer de pulmón se puede clasificar en (Travis et al. 2015):

- Cáncer de pulmón de células pequeñas o microcítico (SCLC, *small cell lung cancer*): procede de las células neuroendocrinas del epitelio bronquial. Supone un 20% de los cánceres de pulmón y tiende a diseminarse con rapidez, aunque responde a la quimioterapia mejor que otros subtipos.
- Cáncer de pulmón no microcítico (NSCLC, *non-small cell lung cancer*): incluye a la mayoría de los cánceres de pulmón (80% de los casos) y se disemina con más lentitud que el microcítico. Se subdivide en tres tipos:
 - Carcinoma de células escamosas o epidermoide (SqCC, *squamous cell carcinoma*): es el más relacionado con el tabaco, suele afectar a los bronquios lobulares o segmentarios. Este tipo tumoral presenta un crecimiento lento y supone un 30% de los casos de NSCLC.
 - Adenocarcinoma de pulmón (LAC, *lung adenocarcinoma*): es el más abundante, supone casi el 60% de los casos, y procede del epitelio con diferenciación glandular. Este tipo histológico es el más frecuente en personas que nunca han fumado.
 - Carcinoma de células grandes (LCC, *large cancer cells*): es el menos frecuente, supone un 10% de los casos. Es un tumor indiferenciado, de localización periférica y desarrollado principalmente en fumadores.

La figura 7 resume la clasificación histopatológica del cáncer de pulmón.

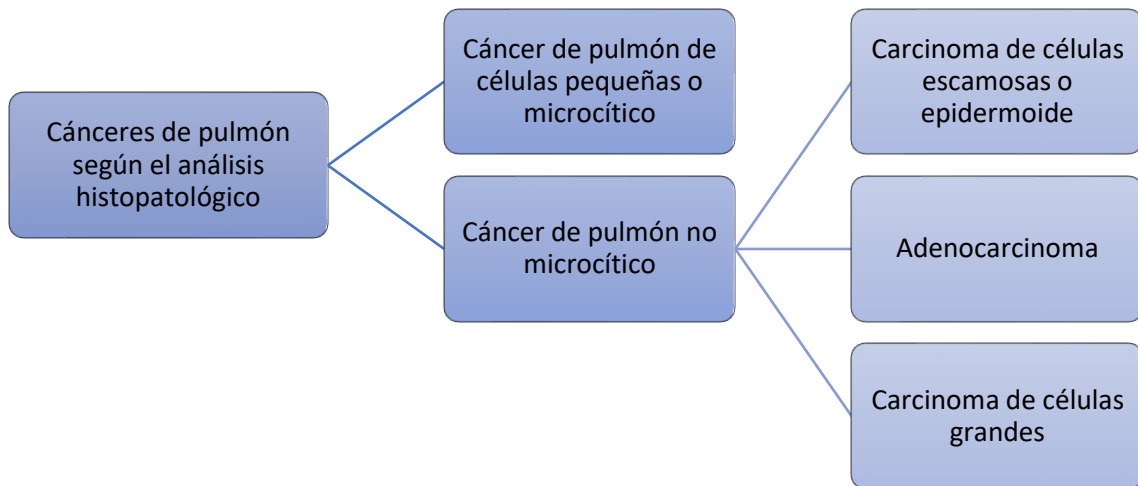


Figura 7: Clasificación histopatológica del cáncer de pulmón.

El análisis histopatológico de la muestra ya supone la aplicación de la MPP, ya que los diferentes resultados obtenidos de estos análisis pueden suponer una adecuación del tratamiento y un pronóstico para el paciente. Pero además de esta aplicación, las tecnologías ómicas descritas anteriormente junto a la bioinformática pueden complementar el diagnóstico, obtenerlo en un estadio precoz o detectar otras dianas terapéuticas determinantes para la adecuación del tratamiento.

Los **análisis ómicos** pueden ser realizados sobre biopsia del tejido o biopsia líquida. El análisis de la biopsia tisular ha sido el abordaje clásico para el diagnóstico y la caracterización tumoral, si bien la invasividad del biopsiado, la posible heterogeneidad celular que puede presentar el tumor y el desarrollo de las tecnologías ómicas, han convertido a la biopsia líquida en una opción plausible para el diagnóstico, pronóstico y adecuación del tratamiento del paciente. La **biopsia líquida (BL)** es el proceso por el cual la sangre u otros fluidos corporales son analizados para detectar posibles alteraciones, y puede ayudar a esclarecer las características del cáncer de pulmón identificando células u otros metabolitos tumorales liberados en la sangre por el crecimiento o la apoptosis de las células cancerosas. Por lo tanto, la detección en sangre u otros fluidos (saliva o fluido del lavado broncoalveolar) podría tanto diagnosticar el cáncer, como revelar características del tumor (tipo de cáncer, mutaciones genéticas...) (Li et al. 2022).

En la rutina clínica ya se está incorporando la detección en BL de algunas mutaciones claves en el desarrollo del cáncer de pulmón, principalmente las mutaciones en el gen *EGFR* para el diagnóstico y la adecuación del tratamiento. Otros biomarcadores en BL que aún no están incorporados en la rutina clínica, pero de los que existe numerosa evidencia científica, son mutaciones en los genes *ALK*, *BRAF* y *KRAS* (Bonanno et al. 2022).

1.6.3 Aplicación de la MPP en el tratamiento de cáncer de pulmón

Desde que en 2001 se realizaran los primeros ensayos clínicos dirigidos a una diana terapéutica, concretamente el EGFR (Herbst et al. 2002), las **terapias personalizadas** han cambiado el tratamiento del cáncer de pulmón. Hoy se aplican en la rutina clínica muchos tratamientos personalizados, principalmente inhibidores de *EGFR*, *ALK*, *RET*, *BRAF*, *ROS1*, *NTKR*, *MET* y *KRAS*. Entre los tratamientos más conocidos para estas dianas se encuentran (Wang, Herbst, Roy S y Boshoff 2021): Gefitinib, Erlotinib, Afatinib, Osimertinib o Amivantamab (en los pacientes con mutaciones en el gen *EGFR*), Capmatinib (en los pacientes con mutaciones en el gen *MET*), Sotorasib (en los pacientes con mutaciones en el gen *KRAS*), Lorlatinib (en los pacientes con mutaciones en los genes *ALK* y *ROS*), Brigatinib (en los pacientes con mutaciones en el gen *ALK*), Selpercetinib (en los pacientes con mutaciones en el gen *RET*) y Dabrafenib (en los pacientes con mutaciones en el gen *BRAF*) (Wang, Herbst, Roy S. y Boshoff 2021).

La existencia de estos tratamientos hace imprescindible conocer las alteraciones tumorales a nivel molecular que presenta el paciente, para poder adecuar su tratamiento. Desafortunadamente, menos de un 25% de los pacientes pueden beneficiarse de estos tratamientos personalizados y casi su totalidad desarrollan resistencias durante el tratamiento (Wang, Herbst, Roy S y Boshoff 2021). Por ello, se hace imprescindible la búsqueda de nuevas dianas terapéuticas que puedan ser el objetivo de nuevos tratamientos personalizados y revelar información sobre la evolución de la enfermedad.

Este uso de la MPP para la adecuación del tratamiento puede englobarse dentro del campo de la farmacogenética o farmacogenómica. La **farmacogenómica** es la

Introducción

adecuación del tratamiento del paciente en función a las características moleculares de la enfermedad que hayan revelado las pruebas ómicas.

Objetivos



UNIVERSIDAD
DE MÁLAGA

2. Objetivos

Esta tesis doctoral centrada en la aplicación de la MPP y la bioinformática para el cáncer de pulmón parte de la siguiente **hipótesis**:

Distintas poblaciones con distintas tipologías de cáncer de pulmón llevan a distintos abordajes de la MPP.

Para verificar la hipótesis mencionada se plantea como **objetivo general**:

Mejorar la aplicación de la MPP por coexpresión de biomarcadores para pacientes con cáncer de pulmón.

Para alcanzar este objetivo se plantean estos **objetivos específicos**:

- Elaboración de un marco de evaluación cualitativo y cuantitativo adaptable a distintos sistemas sanitarios para evaluar el nivel de evaluación de la implementación de la MPP.
- Evaluación de la implementación de la MPP a nivel nacional en distintos sistemas sanitarios mediante el marco generado.
- Análisis de la efectividad de la biopsia líquida para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón.
- Expresión diferencial de ARNlnc entre tejido tumoral y tejido sano de pacientes de diferentes etnias con distintos tipos de cáncer de pulmón.
- Correlación de la expresión de ARNlnc y ARNm para identificar posibles nuevos biomarcadores para biopsia líquida en los pacientes con cáncer de pulmón.



UNIVERSIDAD
DE MÁLAGA

Metodología



UNIVERSIDAD
DE MÁLAGA

3. Metodología

En esta sección se exponen las metodologías desarrolladas en este doctorado para la consecución de los objetivos expuestos, cuyos resultados permitirían aceptar o rechazar la hipótesis planteada.

3.1 Revisión sistemática

Una revisión sistemática (RS) es un estudio secundario para responder una pregunta de investigación a través de búsquedas exhaustivas de la evidencia disponible y la síntesis de los resultados encontrados (Fernandez-Chinguel et al. 2019). Estos estudios requieren una metodología reproducible, transparente y crítica, que permita reducir los sesgos en la investigación (Linares-Espinós et al. 2018). Para realizar una RS, deben seguirse los **pasos** mostrados en la figura 8 (Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ 2022):

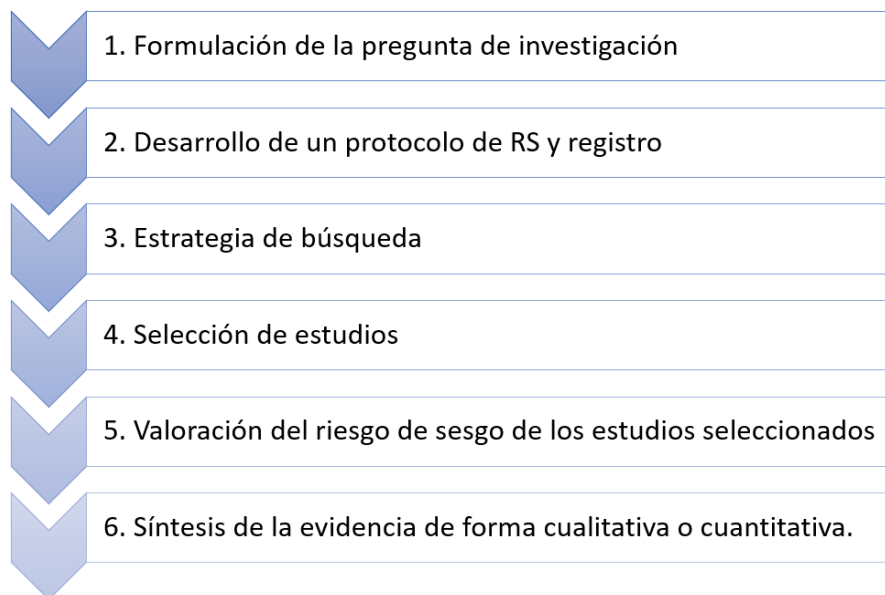


Figura 8: Pasos para realizar una revisión sistemática.

En esta tesis se desarrollan dos RS totalmente independientes entre ellas: la primera para generar un marco de evaluación de la implementación de la MPP y la segunda para evaluar la efectividad de la biopsia líquida en el diagnóstico y la adecuación del

tratamiento en el cáncer de pulmón. En los siguientes subapartados se explican los detalles de cada una de estas fases. Esta misma metodología ha sido utilizada por la autora de esta tesis en otros trabajos ya publicados (Aguilera-Cobos et al. 2022; Aguilera-Cobos et al. 2022a; Aguilera-Cobos et al. 2022b; Aguilera-Cobos et al. 2022c; Aguilera-Cobos et al. 2022d).

3.1.1 Formulación de la pregunta de investigación

La pregunta de investigación contará de unos elementos que dependerán del tipo de RS. Esto origina diferentes **tipos de preguntas** que serán denominadas mediante acrónimos de los elementos que constan en inglés. Según el tipo de RS las diferentes estructuras de pregunta de investigación son (Roqué et al. 2020):

- Evaluaciones de prevalencia de una enfermedad: estructura de la pregunta de investigación CoCoPop-S (*condition, context, population y study design*).
- Evaluación del pronóstico de una enfermedad: estructura de la pregunta de investigación CoCoPop-S (*condition, context, population y study design*).
- Evaluación de factores pronósticos de una enfermedad: estructura de la pregunta de investigación PICOT-S (*population, intervention factor, comparison, outcome, time y study design*) o PFO-S (*population, factor or model, outcome y study design*).
- Evaluación de modelos de predicción para el pronóstico de una enfermedad: estructura de la pregunta de investigación PICOT-S (*population, intervention or factor, comparison, outcome, time y study design*).
- Evaluación de la exactitud de pruebas diagnósticas: estructura de la pregunta de investigación PIRD-S (*population, index test, reference test, diagnosis of interest y study design*).
- Evaluación de los efectos de la aplicación de una intervención: estructura de la pregunta de investigación PICO-S (*population, intervention, outcome of interest y study design*).

Las dos RS planteadas en esta tesis doctoral responden a diferentes preguntas por lo que su estructura es diferente. La pregunta de la primera RS es: ¿Cuáles son los dominios para la determinación del nivel de implementación de la MPP a nivel nacional? La pregunta de la segunda RS es: ¿Es la biopsia líquida tan efectiva como la biopsia de tejido

tumoral en la detección de biomarcadores ómicos para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón?

3.1.2 Desarrollo de un protocolo de la revisión sistemática y registro.

Tras disponer la pregunta de investigación, según la estructura que le corresponda, se debe desarrollar un **protocolo** de la revisión sistemática. Este paso es necesario para que el autor establezca previamente la metodología de las siguientes fases de la RS. El objetivo es reducir los sesgos que se pudieran producir, por parte del autor de la RS, según la evidencia encontrada. Cualquier desviación respecto al protocolo debe de quedar registrada y explicar los motivos de las modificaciones (Linares-Espinós et al. 2018).

Existen diversos repositorios web en los que poder registrar el protocolo de una RS. Tras el registro, se asigna un identificador único a cada uno de los protocolos. A través del identificador los revisores y lectores de la RS podrán comprobar si se siguió el protocolo preestablecido por los autores, y en caso de existir algunas modificaciones, si estas fueron debidamente justificadas. Los repositorios web más conocidos y ampliamente utilizados para el registro de protocolos de revisiones sistemáticas son: PROSPERO (<https://www.crd.york.ac.uk/prospero/>) y Open Science Framework (<https://osf.io/>).

En el caso de las dos RS realizadas en esta tesis doctoral, solamente el protocolo de la primera de ellas (relativa a la implementación de la MPP) pudo ser registrado en un repositorio web. Se registró en el registro de protocolos PROSPERO donde se le asignó el identificador CRD42022338611 y puede ser consultado. El protocolo de la segunda RS no pudo ser registrada por falta de permisos.

3.1.3 Estrategia de búsqueda.

Las estrategias de búsqueda de las revisiones sistemáticas son procesos iterativos de búsqueda para la identificación de estudios que respondan a la pregunta de investigación, y que después serán filtradas. Estas estrategias deben ser completas, objetivas y reproducibles (Linares-Espinós et al. 2018).

En primer lugar, se deben determinar las **bases de datos** en las que se realizará la búsqueda. Para ello hay que tener en cuenta el tipo de estudios que se desea incluir y el objeto de evaluación de la RS.

Metodología

En segundo lugar, es importante establecer cuantos de los componentes de la pregunta de investigación se incluyen en la estrategia de búsqueda. Si se incluyen todos los elementos, pueden no ser detectados estudios que respondan a la pregunta, en cambio, si se incluyen muy pocos el resultado de la búsqueda sería un alto número de estudios, pero pocos responderían a la pregunta. Por lo tanto, es necesario un balance entre la precisión (número de estudios relevantes detectados sobre el total de estudios existentes) y la sensibilidad (número de estudios relevantes detectados sobre el total de estudios relevantes existentes) (Linares-Espinós et al. 2018). Para hacer este balance se realizan diversas pruebas de **estrategias de búsqueda**, antes de dar el aprobado a la búsqueda definitiva, cuyos términos y resultados son revisados por más de uno de los investigadores incluidos en el proyecto.

En tercer lugar, la estrategia de búsqueda debe adaptarse a cada una de las bases de datos, que incluyen los descriptores y la terminología libre adecuada. Los **descriptores** son palabras o términos incluidos en el tesoro de la base de datos para representar conceptos, y que aparecen asociados a los estudios según su área de estudio, temática o metodología. La **terminología libre** son otras palabras que deben incluirse porque pueden quedar fuera de los descriptores de esa base de datos o no encontrarse correctamente asociados a los estudios que se quieren identificar. Cabe destacar que las principales bases de datos se encuentran en inglés, por lo que será este el idioma utilizado en las estrategias de búsqueda. Por ejemplo, en una búsqueda para cáncer de pulmón se debería incluir el descriptor que existe para esta enfermedad (*Lung Neoplasms*), y terminología libre como *lung tumour* o *pulmonary sarcoma*.

En cuarto lugar, deben establecerse las correctas relaciones entre sinónimos y términos mediante las conjunciones “o” o “y”. Entre las diferentes partes de la pregunta de investigación, por ejemplo, tecnología en estudio (biopsia líquida) y enfermedad (cáncer de pulmón), el booleano que se debe emplear es “y”. En cambio, entre los diferentes descriptores o terminología libre usados durante la búsqueda para referirse a un mismo concepto, por ejemplo, *Lung neoplasms*, *lung tumour* o *pulmonary sarcoma* para referirse a cáncer de pulmón, debe usarse el booleano “o”.

Por último, junto con los elementos de la pregunta de investigación pueden incluirse **otras limitaciones** en la estrategia de búsqueda como el año de publicación o el idioma

de los estudios, que ayudarán en el balance de número de estudios identificados. Esto también debe quedar registrado y argumentado en el protocolo de la RS (Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ 2022).

Para las 2 RS presentadas en este trabajo, se construyeron 2 estrategias de búsqueda independientes adecuadas a las preguntas de investigación planteadas en cada una de ellas. Estas estrategias se mostrarán en el apartado de Resultados.

3.1.4 Selección de estudios.

Una vez obtenidos los estudios detectados en las diversas bases de datos y eliminados los duplicados por haber sido detectados en más de una base de datos, se procede a la selección de los estudios que responden a la pregunta de investigación.

El proceso de selección debe realizarse de acuerdo con los **criterios de inclusión y exclusión** establecidos previamente en el protocolo. Estos criterios deben de contener todos los elementos de la pregunta de investigación, pudiendo contener además otros criterios adicionales como el año de publicación o el idioma de los estudios.

La selección de estudios consta de dos fases, realizándose ambas por pares (dos investigadores deben de estar de acuerdo en su inclusión o exclusión de forma cegada) y en caso de desacuerdo la discrepancia se resuelve mediante consenso o con el criterio de un tercer investigador (Roqué et al. 2020):

- **Fase 1.** Selección de los estudios por título y resumen: los investigadores deben establecer si atendiendo al título y al resumen del estudio, éste parece responder a la pregunta de investigación. En caso de duda, se recomienda la inclusión del estudio para su consulta a texto completo.
- **Fase 2.** Selección de los estudios por texto completo: los estudios incluidos en primera fase por ambos investigadores son valorados a texto completo para determinar si responden a la pregunta de investigación. En caso negativo, debe especificarse cuál es el criterio de inclusión que no cumple. En caso afirmativo, el estudio es finalmente incluido para las siguientes fases de la revisión.

Metodología

Con la finalidad de que ambas fases sean cegadas, se han desarrollado herramientas para el filtrado de estudios. Entre ellas destaca la herramienta online **Covidence** (<https://www.covidence.org/>) que permite la gestión de las RS. En primer lugar, almacena de manera online los estudios identificados en la estrategia de búsquedas. En segundo lugar, permite que los investigadores tomen en las fases de selección la decisión de incluir o excluir cada uno de los estudios de manera independiente y cegada en función a unos criterios de inclusión y exclusión previamente establecidos por ellos. En tercer lugar, cuenta con la opción de añadir etiquetas personalizables y comentarios a todos los estudios, lo cual puede ser muy útil especialmente en la resolución de conflictos.

Tras su finalización, el proceso de selección de estudios se resume en un diagrama de flujo que estratifica el número de estudios seleccionados y descartados en cada una de las fases. Esta representación sigue los criterios establecidos por la Declaración PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*) (Liberati et al. 2009).

En ambas RS realizadas en este trabajo se usó la herramienta Covidence para la selección de estudios, y se mostrarán los resultados del número de estudios incluidos y excluidos en sus respectivos diagramas de flujo.

3.1.5 Valoración del riesgo de sesgo de los estudios seleccionados.

Los riesgos de sesgo de un estudio dependerán tanto del tipo de estudio como del objetivo evaluado por el mismo. Por ello, existen diferentes herramientas que contienen distintos parámetros de evaluación, que deben ser elegidas adecuadamente por parte de los autores de la RS para valorar el riesgo de sesgo de los estudios seleccionados. Algunas de las herramientas más ampliamente utilizadas son:

- QUADAS-2 (Whiting et al. 2011): evaluación del riesgo de sesgo de ensayos clínicos de pruebas diagnósticas.
- ROB-2 (*risk-of-bias tool for randomized trials*) (Sterne et al. 2019): evaluación del riesgo de sesgo de ensayos clínicos aleatorizados de intervenciones clínicas.

- ROBINS-I (*risk of bias in non-randomised studies of interventions*) (Sterne et al. 2016): evaluación del riesgo de sesgo de estudios observacionales de intervenciones clínicas.
- NOS (*Newcastle-Ottawa scale*) (Stang 2010): evaluación del riesgo de sesgo de estudios observacionales de intervenciones clínicas.
- QUIPS (*quality In prognosis studies*) (Grooten et al. 2019): evaluación del riesgo de sesgo de ensayos clínicos aleatorizados o no-aleatorizados pronósticos.
- PROBAST (*prediction model risk of bias assessment tool*) (Wolff et al. 2019): evaluación del riesgo de sesgo de modelos de predicción.

Si los estudios incluidos son revisiones o metanálisis, también existen herramientas específicas para su evaluación. El uso de estas herramientas dependerá del tipo de revisiones.

- AMSTAR-II (*a measurement tool to assess systematic reviews*) (Shea et al. 2017): evaluación de la calidad de revisiones sistemáticas y metanálisis.
- SANRA (*scale for the assessment of narrative review articles*) (Baethge, Goldbeck-Wood y Mertens 2019): evaluación de la calidad de revisiones narrativas.
- Lista de verificación publicada por Humphrey-Murto (Humphrey-Murto et al. 2017): evaluación de la calidad de entrevistas y paneles de expertos.

Independientemente del tipo de estudios incluidos, el objetivo de la RS y la herramienta utilizada, esta valoración debe realizarse también siguiendo una metodología por pares cegada. En caso de desacuerdo, se podrá resolver por consenso o ser resuelto por un tercer investigador.

La primera de las RS realizada en este trabajo, cuenta entre sus estudios incluidos con RS, revisiones narrativas, entrevistas y paneles de expertos. Por lo que se usaron 3 herramientas diferentes para la valoración de sesgos de los estudios incluidos: AMSTAR-II (para valorar RS), SANRA (para evaluar revisiones narrativas) y la Lista de verificación publicada por Humphrey-Murto (para evaluar entrevistas y paneles de expertos). Dado

Metodología

que son las herramientas usadas en esta RS, se muestran a continuación los criterios de evaluación de cada una ellas.

La herramienta **AMSTAR-II** consta de 16 preguntas para la evaluación de la calidad de las RS y con preguntas específicas para aquellas RS que contengan un MA. Todas las preguntas deben ser respondidas con sí, sí parcial o no. Estas preguntas son:

1. ¿Las preguntas de investigación y los criterios de inclusión para la revisión incluyen los componentes PICO?
2. ¿El reporte de la revisión contiene una declaración explícita de que los métodos de la revisión fueron establecidos con anterioridad a su realización y justifica cualquier desviación significativa del protocolo?
3. ¿Los autores de la revisión explicaron su decisión sobre los diseños de estudio a incluir en la revisión?
4. ¿Los autores de la revisión usaron una estrategia de búsqueda bibliográfica exhaustiva?
5. ¿Los autores de la revisión realizaron la selección de estudios por duplicado?
6. ¿Los autores de la revisión realizaron la extracción de datos por duplicado?
7. ¿Los autores de la revisión proporcionaron una lista de estudios excluidos y justificaron las exclusiones?
8. ¿Los autores de la revisión describieron los estudios incluidos con el detalle adecuado?
9. ¿Los autores de la revisión utilizaron una técnica satisfactoria para evaluar el riesgo de sesgo en los estudios individuales que se incluyeron en la revisión?
10. ¿Los autores de la revisión informaron sobre las fuentes de financiación para los estudios primarios incluidos en la revisión?
11. Si se realizó un metanálisis, ¿usaron los autores de la revisión los métodos apropiados para la combinación estadística de los resultados?
12. Si se realizó el metanálisis, ¿los autores evaluaron el impacto potencial de riesgo de sesgo en estudios individuales sobre los resultados del metanálisis u otra síntesis de evidencia?
13. ¿Los autores de la revisión dieron cuenta de riesgo de sesgo en los estudios primarios al interpretar/discutir los resultados de la revisión?

14. ¿Los autores de la revisión proporcionaron una explicación satisfactoria y una discusión sobre cualquier heterogeneidad observada en los resultados de la revisión?
15. Si realizó una síntesis cuantitativa, ¿ejecutaron los autores de la revisión una investigación adecuada del sesgo de publicación y discutieron su posible impacto en los resultados de la revisión?
16. ¿Los autores de la revisión informaron sobre posibles fuentes de conflicto de intereses, incluido el financiamiento que recibieron para realizar la revisión?

Según la respuesta a estas preguntas, la herramienta AMSTAR-II otorga una calidad a la RS analizada, de mejor a peor calidad: alta calidad, calidad moderada, baja calidad o calidad críticamente baja.

La herramienta **SANRA** (Baethge, Goldbeck-Wood y Mertens 2019) consta de 6 preguntas para evaluar la calidad de las revisiones narrativas. Todas las preguntas se valoran con 0, 1 o 2, pudiendo obtener un estudio la puntuación máxima de 12 puntos. Estas preguntas y las posibles respuestas son:

1. Justificación de la importancia de la revisión para los lectores:
Puntuación 0: Importancia no justificada
Puntuación 1: Se alude a su importancia, pero no está explícita.
Puntuación 2: Su importancia está justificada y explícita.
2. Objetivos de la revisión narrativa:
Puntuación 0: No se formulan los objetivos de la revisión.
Puntuación 1: Objetivos de la revisión formulados de forma general.
Puntuación 2: Objetivos y preguntas de investigación de la revisión claramente formulados.
3. Descripción de la búsqueda de bibliografía:
Puntuación 0: No se presenta ninguna estrategia de búsqueda.
Puntuación 1: La búsqueda en la bibliografía se describe brevemente.
Puntuación 2: La estrategia de búsqueda se presenta en detalle, incluidos los términos de búsqueda y los criterios de inclusión.

4. Referencias:

Puntuación 0: Las principales afirmaciones de la revisión no se sustentan en referencias.

Puntuación 1: Las principales afirmaciones de la revisión se sustentan de forma inconsistente en referencias.

Puntuación 2: Las principales afirmaciones de la revisión se sustentan correctamente en referencias.

5. Base científica:

Puntuación 0: La revisión no se basa en premisas científicas apropiadas.

Puntuación 1: La revisión se basa en premisas científicas de forma selectiva.

Puntuación 2: La revisión se sustenta en la adecuada evidencia científica.

6. Adecuada presentación de los datos:

Puntuación 0: Los datos se presentan de forma inadecuada.

Puntuación 1: Algunos de los datos no se presentan de la forma más adecuada.

Puntuación 2: Los datos relevantes se presentan adecuadamente.

La **lista de verificación publicada por Humphrey-Murto** (Humphrey-Murto et al. 2017) consta de 11 preguntas para evaluar la calidad de las entrevistas y los paneles de expertos. Las 11 preguntas deben ser respondidas con un sí o un no. En los casos de respuesta afirmativa, se le otorga a ese campo un valor de 1. En los casos de respuesta negativa, se le otorga a ese campo un valor de 0. Las preguntas de esta lista de verificación son:

1. ¿Define el propósito u objetivo del estudio?
2. ¿Resume cada paso del proceso y justifica las modificaciones del protocolo, en caso de haberse producido?
3. ¿Describe la selección y preparación de las pruebas para los participantes?
4. ¿Describe cómo se seleccionaron los elementos para incluirlos en el cuestionario inicial?

5. ¿Describe cómo se seleccionó a los participantes y sus cualificaciones?
6. ¿Describe la formación de los facilitadores?
7. ¿Describe el número de rondas previstas o los criterios para poner fin al proceso?
8. ¿Describe claramente cómo se definió el consenso?
9. ¿Reporta los índices de respuesta y los resultados después de cada ronda?
10. ¿Describe el tipo de retroalimentación proporcionada a los participantes después de cada ronda?
11. ¿Describe cómo se mantuvo el anonimato de los participantes?

En cuanto a la segunda RS presentada en este trabajo, se incluyeron ensayos clínicos de pruebas diagnósticas, por lo que para la evaluación del riesgo de sesgo de estos estudios se usó la herramienta **QUADAS-2** (Whiting et al. 2011). Esta herramienta consta de 10 preguntas divididas en 4 dominios de evaluación: selección de pacientes, prueba índice, prueba de referencia y flujo y tiempo. Todas las preguntas deben responderse con: sí, no o no especificado.

Dominio 1: Selección de pacientes. Consta de 3 preguntas que evalúan si la selección de pacientes podría haber introducido sesgos en el estudio.

1. ¿Se enroló una muestra consecutiva o aleatoria de pacientes?
2. ¿Se evitó un diseño de casos y controles?
3. ¿Se evitaron exclusiones inapropiadas?

Dominio 2: Prueba índice. Consta de 2 preguntas que evalúan si la realización o la interpretación de la prueba índice podría haber introducido sesgos en el estudio.

1. ¿Fueron interpretados los resultados de la prueba índice sin conocimiento de los resultados de la de referencia?
2. Si se utilizó un umbral para definir la positividad o la negatividad de la prueba índice, ¿fue especificado previamente?

Dominio 3: Prueba de referencia. Consta de 2 preguntas que evalúan si la realización o la interpretación de la prueba de referencia podría haber introducido sesgos en el estudio.

Metodología

1. ¿Es probable que la prueba de referencia valore correctamente la condición diana?
2. ¿Fueron interpretados los resultados de la prueba de referencia sin conocimiento de los resultados de la prueba índice?

Dominio 4: Flujo y tiempo. Consta de 3 preguntas que evalúan si el flujo de pacientes podría haber introducido sesgos en el estudio.

1. ¿Hubo un intervalo apropiado entre la prueba índice y la prueba de referencia?
2. ¿Fue aplicada en todos los individuos la misma prueba de referencia?
3. ¿Fueron incluidos todos los pacientes en el análisis?

3.1.6 Síntesis de la evidencia de forma cualitativa o cuantitativa.

La síntesis de la evidencia es la exposición de forma resumida y comprensible de los resultados de los estudios incluidos en la RS, siguiendo la misma metodología por pares expuesta en fases anteriores. La extracción de los datos y su síntesis se realizará en tablas diseñadas en la fase de diseño de protocolo y adaptadas para cada RS.

En una primera fase de síntesis de la evidencia deben incluirse características generales de los estudios incluidos como la referencia del estudio, la fecha de publicación y la metodología y objetivo de este.

En una segunda fase de síntesis deben extraerse los resultados de interés de los estudios incluidos. Esta síntesis puede ser cualitativa o cuantitativa dependiendo del tipo de revisión que se esté realizando y del tipo de resultados de interés (Linares-Espinós et al. 2018):

- Si los resultados extraídos son **cualitativos**, se debe realizar un **análisis temático** de los estudios incluidos. Esta metodología permite la identificación, el análisis y la interpretación de patrones de caracteres dentro de una serie de datos cualitativos (Thomas y Harden 2008).
- Si los resultados extraídos son **cuantitativos** y se incluyen dos o más estudios, se puede realizar un **metanálisis** de la evidencia. Un metanálisis es una técnica estadística para resumir en un único valor los resultados de dos o más estudios que hayan comparado dos grupos (Fernandez-Chinguel et al. 2019).

En los siguientes apartados de esta sección se desarrollan ambas técnicas de síntesis de la evidencia.

Análisis temático

El análisis temático es una metodología de síntesis de la evidencia en RS para resultados de tipo cualitativo. Este análisis es un enfoque flexible que permite a los investigadores generar nuevas percepciones y conceptos derivados de los datos de los estudios incluidos. Esta flexibilidad puede llegar a suponer un hándicap para esta técnica, ya que existen muchas formas diferentes de interpretar el significado de un conjunto de datos, lo que conduce a subjetividad. Otra de sus desventajas, es que se centra en la búsqueda de patrones de caracteres y significados pudiendo quedar ignorados aquellos datos que aparecen en pocos estudios respecto al total de estudios incluidos, pero que pudieran ser de gran relevancia (Thomas y Harden 2008).

Esta técnica de síntesis de la evidencia cualitativa se realiza en varias **fases** (Cruzes y Dybå 2011):

1. Lectura de los estudios completos y anotaciones de las ideas fundamentales y posibles códigos.
2. Creación de códigos que representen los significados y patrones de los datos que se consideran de interés para responder la pregunta de investigación. Los códigos pueden agruparse en un libro de códigos para llevar el registro. Todos los significantes que representen un mismo significado deben quedar asignados bajo el mismo código.
3. Agrupación de los significantes asociados a un mismo código.
4. Agrupación de los códigos en temas, pudiendo convertirse algunos temas en subtemas. Los temas pueden permitir estudiar posibles tendencias y patrones de los datos.
5. Revisión de los temas, y si es necesaria su modificación, considerando que cada uno de los temas debe contener suficientes datos que respalden su entidad como tema y que lo diferencien del resto. Se puede contemplar la posibilidad de eliminar aquellos temas que no cuenten con datos suficientes o fusionar temas similares.

Metodología

Este análisis debe realizarse siguiendo los pasos anteriormente mencionados mediante una metodología por pares (síntesis y análisis de los datos por parte de dos investigadores de manera independiente y cegada). Existen también herramientas en línea para automatizar todo este proceso o parte de él, como: ATLAS.ti (<https://atlasti.com/>) o Delve (<https://delvetool.com/>).

La primera de las RS desarrolladas en esta tesis tiene como finalidad determinar los dominios para evaluar la implementación de la MPP. Por ello, la información extraída de los estudios incluidos es de tipo cualitativo y su análisis se realiza mediante análisis temático siguiendo las fases anteriormente expuestas. No se utilizó en este caso ninguna herramienta de automatización, pero sí se mantuvo la síntesis y el análisis por pares.

Metanálisis de intervenciones diagnósticas o adecuación del tratamiento

Un metanálisis es una técnica estadística para la síntesis de la evidencia de datos cuantitativos de una RS. Dependiendo del tipo de datos analizados y el objetivo de la RS, existen distintos tipos de metanálisis, por ejemplo, el tipo de variables objeto de análisis no será el mismo para la evaluación de intervenciones diagnósticas que para intervenciones pronósticas. Dado que en esta tesis doctoral se utilizará esta técnica para la evidencia recabada sobre una intervención diagnóstica o de adecuación de tratamiento (evaluación de la efectividad de la biopsia de tejido tumoral en la detección de biomarcadores ómicos para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón), este apartado metodológico se focalizará especialmente en el análisis de los datos para este tipo de intervenciones.

Una prueba diagnóstica puede aportar información sobre la presencia o ausencia de una determinada enfermedad, la estadificación de una enfermedad o la existencia de una determinada condición clínica, no necesariamente patológica. En los ensayos clínicos, estas intervenciones se evalúan comparándolas frente a la **prueba de referencia o gold standard**. La prueba de referencia es la intervención ampliamente consensuada y usada en la rutina clínica para obtener el resultado diagnóstico en estudio (Bravo-Grau y Cruz 2015).

La finalidad de un metanálisis de una intervención diagnóstica es evaluar si la prueba en evaluación es igual, mejor o peor que la prueba de referencia para el diagnóstico de una

enfermedad determinada. Además de los resultados estadísticos obtenidos, debe ser tenido en cuenta por el evaluador si la prueba en evaluación será usada como una prueba de selección (*triage*), como una prueba complementaria a la prueba de referencia o pretende reemplazarla (Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ 2022).

Para evaluar los datos de los estudios debe considerarse si los desenlaces son **dicotómicos** o **continuos**. Un desenlace dicotómico es aquel que solamente admite dos respuestas: sí o no. Un desenlace continuo es aquel que se expresa en valores numéricos entre un rango de posibilidades. En las pruebas diagnósticas el desenlace evaluado debe ser dicotómico (presencia o ausencia de enfermedad), si bien puede evaluarse con un valor numérico (por ejemplo, niveles de colesterol en sangre) o con un resultado positivo/negativo (por ejemplo, tiene o no tiene una determinada mutación) (Fernandez-Chinguel et al. 2019). Para las pruebas evaluadas con valor numérico debe establecerse un **valor umbral**, a partir del cual el resultado se considere positivo. Dado que este apartado se centra en el metanálisis para pruebas diagnósticas, los valores de medidas de precisión que se muestran en este apartado son para desenlaces dicotómicos.

Tras estas consideraciones, el primer paso en los metanálisis de intervenciones diagnósticas es la generación de la **tabla de contingencia 2x2**. Esta tabla con dimensiones 2x2, se genera con los valores acumulados de verdaderos positivos (TP, *true positive*), falsos positivos (FP, *false positive*), falsos negativos (FN, *false negative*) y verdaderos negativos (TN, *true negative*) de los datos de los estudios incluidos de la intervención diagnóstica en evaluación frente a la prueba de referencia. Se considera:

- TP: resultado positivo tanto en la prueba en evaluación como en la prueba de referencia.
- FP: resultado positivo en la prueba en evaluación y negativo en la prueba de referencia.
- FN: resultado negativo en la prueba en evaluación y positivo en la prueba de referencia.
- TN: resultado negativo tanto en la prueba en evaluación como en la prueba de referencia.

		Prueba de referencia	
		Positivo	Negativo
Prueba en evaluación	Positivo	Verdaderos positivos (TP)	Falsos positivos (FN)
	Negativo	Falsos negativos (FN)	Verdaderos negativos (TN)

Figura 9: Ejemplo de tabla de contingencia 2 x 2

Con estos datos acumulados se pueden obtener fácilmente los valores de los siguientes parámetros que permitirán la evaluación de la prueba (Shreffler y Huecker 2022):

- Sensibilidad (sen, *sensitivity*): proporción de individuos correctamente diagnosticados con la prueba en evaluación. Se calcula dividiendo el número de TP entre el sumatorio de TP + FN.
- Especificidad (spe, *specificity*): proporción de individuos correctamente diagnosticados con ausencia de la condición o enfermedad por la prueba diagnóstica en evaluación. Se calcula dividiendo el número de TN entre el sumatorio de FP + TN.
- Valor predictivo positivo (VPP): probabilidad condicional de que el paciente tenga la enfermedad, tras obtener un resultado positivo en la prueba diagnóstica en evaluación. Se calcula dividiendo el número de TP entre el sumatorio de TP + FP.
- Valor predictivo negativo (VPN): probabilidad condicional de que el paciente no tenga la enfermedad, tras obtener un resultado negativo en la prueba diagnóstica en evaluación. Se calcula dividiendo el número de TN entre el sumatorio de FN + TN.
- Razón de verosimilitud positiva (LR+, *likelihood ratio +*): probabilidad de que la prueba en evaluación sea positiva en una paciente con la enfermedad dividida entre la probabilidad de que la prueba sea positiva en un paciente sin la enfermedad. Se calcula dividiendo la sensibilidad ente (1 – especificidad), y tiene un valor entre 1 e infinito.

- Razón de verosimilitud negativa (LR-, *likelihood ratio* -): probabilidad de que un paciente que tiene una enfermedad obtenga un resultado negativo en la prueba en evaluación dividido entre la probabilidad de que un paciente que no tiene una enfermedad dé negativo. Se calcula dividiendo $(1 - \text{sensibilidad})$ entre la especificidad, y tiene un valor entre 0 y 1.
- Exactitud: probabilidad de que el resultado del test prediga correctamente la presencia o ausencia de la enfermedad en el paciente. Se calcula dividiendo el sumatorio VP + VN entre el sumatorio del total de los pacientes incluidos.
- *Odds ratio* diagnóstico (DOR): medida de efecto para desenlaces dicotómicos. Representa la asociación entre el resultado de la prueba y la enfermedad. Los valores de DOR varían desde cero a infinito. Un valor DOR igual a 1 significa que la prueba no discrimina entre pacientes enfermos y no enfermos, por lo que el test es totalmente inútil. Los valores mayores de 1 significan que es más probable que la prueba sea positiva en pacientes enfermos que en sanos, cuantos más altos son los valores mejor discrimina la prueba pacientes enfermos y sanos. Las medidas de efecto se suelen presentar junto a su intervalo de confianza (IC), usualmente el IC 95%.

Además de los valores de medida de precisión diagnóstica, tiene que interpretarse estadísticamente la variabilidad entre los efectos de los estudios incluidos para permitir una mejor comprensión de los resultados del metanálisis. Esta interpretación estadística se denomina **heterogeneidad**. La heterogeneidad para los metanálisis de pruebas diagnósticas suele valorarse con (Fernandez-Chinguel et al. 2019):

- Valor de I^2 de Higgins. Este parámetro evalúa si el azar es el responsable de la variabilidad. Una I^2 igual a 0% sugiere que el azar es el responsable de la variabilidad, mientras que un 100% sugiere una variabilidad excesiva. No existe un punto de corte consenso a partir del cual la heterogeneidad sería excesiva para el MA, aunque algunos manuales sugieren que es admisible hasta un valor de 40%.
- Prueba de χ^2 . Esta prueba tiene como hipótesis nula que todos los estudios tengan el mismo valor de la medida del efecto. Si el valor de la p para esta prueba

Metodología

es menor o igual a 0,10 se podría interpretar que hay una heterogeneidad excesiva entre los estudios.

En los casos en los que exista una heterogeneidad excesiva pueden adoptarse distintas estrategias, como la valoración individualizada de cada uno de los estudios para detectar posibles estudios con datos atípicos o contemplar la posibilidad de analizar los datos de forma desagregada dividiendo los estudios incluidos en dos o más subgrupos.

Los resultados estadísticos obtenidos pueden ser representados para facilitar la interpretación visual de los resultados obtenidos en el análisis. Estas **representaciones** pueden mostrar valores agregados para todos los estudios incluidos o desagregados para cada uno de ellos o por subgrupos. Las representaciones más comunes de los resultados estadísticos de metanálisis de pruebas diagnósticas son (Hajian-Tilaki 2013):

- *Forest plot*: gráfico que muestra los resultados de la medida de efecto y su IC para cada uno de los estudios y el resumen estadístico de dichos estudios.
- *Curva ROC (receiver operating characteristic)*: esta representación se realiza en el espacio ROC que está delimitado por los ejes cartesianos sensibilidad frente 1 – especificidad. Esta curva responde a la capacidad de la prueba diagnóstica a discernir entre los pacientes enfermos y los pacientes sanos. Una curva ROC situada en la línea diagonal, refleja que la prueba no supera el nivel del azar por lo que sus resultados no están relacionados con la enfermedad. Cuanto mayor sea la capacidad discriminante de la prueba diagnóstica se situará más cerca de la esquina superior izquierda del gráfico.

Actualmente existen **herramientas**, como Meta-DiSc (Zamora et al. 2006), que permiten realizar de manera automatizada estos análisis, e incluso obtener sus representaciones. Además, en R existen paquetes específicos para el análisis estadístico de MA de pruebas diagnósticas, el más conocido es *mada (Meta-Analysis of Diagnostic Accuracy)*.

Tal y como se ha mencionado al comienzo de este apartado, en la segunda RS realizada (evaluación de la efectividad de la biopsia de tejido tumoral en la detección de biomarcadores ómicos para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón) la información extraída de los estudios es de tipo cuantitativo. Por ello, se

realiza un metanálisis de aquellos biomarcadores y finalidades terapéuticas en los que es posible. La herramienta Meta-DiSc fue usada para este análisis.

3.2 Obtención de datos ARNseq para su análisis bioinformático

El análisis bioinformático de datos ARNseq se realizó en 4 tipos de muestras de diferentes tipos tumorales de cáncer de pulmón y de diferente procedencia. Todas ellas habían sido obtenidas antes del comienzo de esta tesis doctoral por la unidad de bioinformática del Centro de Supercomputación y Bioinformática (SCBI) de la Universidad de Málaga. Igualmente, en todos los casos se trata de muestras pareadas, contándose con muestras tanto de tejido pulmonar sano como de tejido tumoral de todos los pacientes.

La procedencia de las primeras muestras fue el BioBanco de IBIMA, tras la necesaria aprobación por el Comité Ético del Hospital Regional de Málaga y la correspondiente solicitud, se obtuvieron **14 muestras pareadas de 7 pacientes de LAC (LAC7) y 16 muestras pareadas de 8 pacientes de SqCC (SqCC8)**. Los especímenes incluidos cumplían los siguientes criterios adicionales: (1) la muestra procedía de pacientes de cáncer de pulmón en estadio IIA-B o IIIA, (2) el paciente no había recibido quimioterapia ni radioterapia antes de la intervención, (3) no se habían detectado neoplasias en el paciente antes de la intervención y (4) el tejido biopsiado tanto del tumor como del tejido sano debía ser suficiente para extraer el ARN de ambos. Posteriormente, se extrajo el ARN total de todas las muestras y se secuenciaron en el *NextSeq 550* de *Illumina* que la Universidad de Málaga oferta en el Servicio de Ultrasecuenciación del SCBI, siguiendo el protocolo estándar marcado por el fabricante para ARNseq para obtener lecturas pareadas de 2 x 100nt.

Dado el bajo número de pacientes y muestras incluidos, la unidad realizó una búsqueda en datos públicos de secuenciación por ARNseq para incluir un mayor número de muestras ya secuenciadas. Los datos debían de cumplir una serie de criterios para considerar su inclusión: (1) el tejido secuenciado haya sido crionizado, y no parafinado, (2) el protocolo seguido haya permitido la correcta conservación del ARN, (3) la secuenciación haya sido de ARN total, (4) existan muestras pareadas del mismo paciente de tejido sano y tumoral. En la base de datos EGA (*European Genome-phenome Archive*) se encontró un estudio con lecturas pareadas 2 x 75nt de **34 muestras de 17 pacientes**

de **SCLCC (SCLCC17)** obtenidas en un HiSeq2000 de Illumina®. En la base de datos NCBI (*National Center of Biotechnology Information*) se detectó un experimento con lecturas pareadas 2 x 100 nt de **100 muestras de 50 pacientes de LAC (LAC50)** secuenciadas también en un HiSeq2000 de *Illumina*. Tras solicitar los permisos pertinentes, se descargaron los proyectos EGA y NCBI, cuyos identificadores son respectivamente EGAS00001000334 y ERP001058.

Es relevante destacar que, si bien las muestras proporcionadas por IBIMA son de pacientes malagueños, los datos de las muestras de pacientes de SCLCC y LAC descargadas de bases de datos proceden de pacientes de origen asiático. Esta puntualización puede ser importante porque algunos estudios han señalado las grandes diferencias a nivel genómico y transcriptómico que existen entre poblaciones orientales y occidentales en el análisis del cáncer de pulmón (Izumi et al. 2020; Adib et al. 2022).

3.3 Análisis bioinformático de expresión diferencial de datos de ARNseq

Con los datos de ARNseq de los 4 tipos de muestras obtenidos por la unidad de bioinformática del SCBI, se procedió al análisis bioinformático de la expresión diferencial. La expresión diferencial se evaluó para los 2 tipos de biomoléculas en estudio (ARNInc y ARNm) utilizando el mismo flujo de trabajo. La única adaptación realizada según el tipo de biomolécula en estudio fue el archivo con las posiciones genómicas que se debían enfrentar entre el genoma de referencia y las muestras en análisis. Las posiciones genómicas codificantes de ARNm y de ARNInc se obtuvieron respectivamente de los archivos de anotación en formato .gft del genoma humano filtrando por el tipo de transcrito que generan.

El flujo de trabajo utilizado en este análisis es una adaptación del conocido flujo Tuxedo (Trapnell et al. 2012; Ghosh y Chan 2016). Este flujo de trabajo cuenta con una **fase de limpieza y pre-procesamiento de las lecturas**, para lo cual se utilizó el programa SeqTrimNext (Falgueras et al. 2010) con la configuración de parámetros específica para secuenciación con *Illumina*. Se eliminaron lecturas de baja calidad, adaptadores, secuencias contaminadas, colas polyA/polyT y secuencias de menos de 20 pb. Tras la fase de limpieza, se procede al **mapeo** de las lecturas limpias frente al genoma humano

de referencia hg38 usando el software STAR v2.5 (Dobin et al. 2013). Una vez mapeadas las lecturas, se usa el paquete Cufflinks para **ensamblar el transcriptoma y compararlo** entre ambas condiciones en estudio. Cufflinks cuenta con 4 herramientas usadas de forma secuencial (Ghosh y Chan 2016):

- Cufflinks: ensamblaje de los transcritos para cada una de las condiciones.
- Cuffcompare: los transcritos obtenidos son comparados con el archivo de anotación. En este punto se utiliza el archivo de anotación correspondiente a la biomolécula en estudio en cada caso (ARNm o ARNlnc).
- Cuffmerge: fusión de los ensamblajes obtenidos para cada una de las condiciones.
- Cuffdiff: compara los ensamblajes de ambas condiciones mediante análisis estadístico para determinar aquellos que se expresan de forma diferencial entre ambas condiciones.

En el análisis de expresión diferencial realizado en esta tesis el análisis comparativo estadístico realizado por Cuffdiff se basó en probar la hipótesis nula de que el cambio logarítmico ($\log FC$, *logarithmic fold change*) entre los niveles de expresión de los ARNm o ARNlnc entre las muestras de tejido sano y tejido tumoral del mismo paciente es cero, es decir, su expresión no se ve afectada por la enfermedad. Sin embargo, los cambios pequeños, aunque sean estadísticamente muy significativos, no se consideraron candidatos para el análisis de su expresión diferencial, acotando en nuestro análisis la expresión diferencial a aquellos transcritos con valores $-1 > \log FC > 1$. En este caso se tomaron tanto valores positivos como negativos de $\log FC$ porque se quiere estudiar tanto los ARNlnc o ARNm que se encuentran infraexpresados como los sobreexpresados. Además, también se filtraron los resultados de expresión diferencial por P , que tiene en cuenta los niveles de expresión del resto de genes. Si el P está por debajo del nivel de significación (en este caso $P < 0,05$), podemos rechazar la hipótesis nula a favor de la hipótesis alternativa, es decir, concluimos que la expresión diferencial observada es el resultado de un efecto real y no producida por azar. Es reseñable también que al compararse en este estudio muestras del mismo paciente, se elimina el posible ruido que podrían originar las diferencias interpersonales no relacionadas con el proceso oncológico.

Metodología

El archivo final con los transcritos que presentan expresión diferencial es un archivo .csv separado por comas en el que se presentan los valores logFC y P para cada uno de ellos en los pacientes analizados.

La expresión diferencial de ARNm y miARN en las muestras de pacientes malagueños ya ha sido analizada y publicada por otros miembros del grupo (Arroyo et al. 2019; Espinosa Garcia et al. 2022), por lo que los resultados en esta tesis se centran en la expresión diferencial de los ARNlnc en los 4 tipos de muestras.

3.4 Análisis de correlación de los datos de expresión

En este estudio se pretende analizar la posible correlación de la expresión entre ARNlnc y ARNm con expresión diferencial entre tejido tumoral y tejido sano de los mismos pacientes con datos de ARNseq. Este análisis se pudo desarrollar gracias a los archivos obtenidos previamente de expresión diferencial en formato .csv de ARNlnc y ARNm. Para el análisis bioinformático de esta correlación se usó un *script* desarrollado en R y Python (<https://github.com/Unit-Bioinformatic-SCBI-UMA/miRNA-mRNA-correlation>) y que ha sido utilizado en publicaciones previas por la unidad de bioinformática del SCBI (Espinosa Garcia et al. 2022).

El *script* utilizado evalúa el coeficiente de correlación r de Pearson entre los valores logFC de la expresión diferencial de los transcritos en estudio para cada paciente, es decir, evalúa la medida de dependencia lineal de la variación de los valores de expresión de ambos transcritos.

Para evaluar esta correlación, en primer lugar, el *script* genera dos matrices (una para ARNm y otra para ARNlnc) con los valores de logFC de uno de los transcritos (filas) para cada uno de los pacientes (columnas). En segundo lugar, compara ambas matrices gracias a la función `rcorr` del paquete `Hmisc` de R, especificando que el análisis estadístico será la correlación de Pearson y se mantienen el resto de parámetros predeterminados. Al final se obtiene una matriz de correlación de ARNlnc y ARNm basada en el logFC, esta matriz contiene las siguientes columnas: ARNlnc, ARNm, r y P .

En nuestro estudio la matriz final fue filtrada por $-0,8 \geq r \geq 0,8$ y por $P < 0,05$. Al seleccionarse valores de correlación tanto positivos como negativos se obtienen tanto aquellos casos en los que las expresiones diferenciales de ARNm y ARNlnc sean en el

mismo sentido (infraexpresado/infraexpresado o sobreexpresado/sobreexpresado) como en los que sea en sentido opuesto (infraexpresado/sobreexpresado o viceversa), siempre que los niveles de expresión diferencial estén correlacionados.

3.5 Interpretación biológica de datos del análisis de datos de ARNseq

La expresión diferencial de ARNm y ARNlnc entre el tejido tumoral y el tejido normal de los diferentes tipos de muestras, puede ser relevante para el estudio del proceso oncológico, así como para la detección de posible nuevos biomarcadores para el diagnóstico o la adecuación del tratamiento en estas enfermedades.

En primer lugar, se consultan bases de datos para la predicción del análisis funcional de la correlaciones de expresión ARNlnc-ARNm obtenidas, como RNAInter (Kang et al. 2022), RISE (Gong et al. 2018) y Lncbase (Karagkouni et al. 2020). Las dos primeras bases de datos permiten su descarga en formato .txt, por lo que se descargaron y consultaron de forma recurrente para las correlaciones matemáticas obtenidas previamente. La base de datos Lncbase no permite su descarga, por lo que se consultó de manera manual a través de su web. Este análisis permite la validación experimental de las correlaciones de ARNlnc-ARNm con expresión diferencial obtenidas matemáticamente. En los siguientes pasos se trabaja únicamente con los ARNlnc y ARNm con expresión diferencial correlacionada validada en alguna de estas bases de datos.

En segundo lugar, se construye la red de interacciones proteína-proteína con la información referente a los ARNm, aquellos con expresión diferencial correlacionados con la expresión diferencial de algún ARNlnc, en el programa String (Szklarczyk et al. 2021). Además, se analizan las rutas biológicas en las que se encuentran involucrados con la información de las bases de datos ShinyGO (Ge, Jung y Yao 2020) y KEGG Pathways (Kanehisa et al. 2023; Chen et al. 2017).

En tercer lugar, se usará la base de datos ExoRbase (Lai et al. 2022) para consultar si aquellos ARNlnc con expresión diferencial correlacionados con la expresión diferencial de algún ARNm han sido detectados en biopsia líquida de individuos sanos o de pacientes oncológicos. Esta base de datos no permite su descarga, por lo que se consultó de manera manual a través de su web. Este análisis permite conocer si los ARNlnc con expresión diferencial en los casos de cáncer de pulmón en estudio pueden ser objeto de

Metodología

análisis en biopsia líquida para el diagnóstico o adecuación del tratamiento de esta enfermedad.

Resultados



UNIVERSIDAD
DE MÁLAGA

4. Resultados

En esta sección se exponen los resultados, obtenidos al emplear las metodologías expuestas, cuya interpretación permitirá aceptar o rechazar la hipótesis planteada en esta tesis. Los resultados se muestran en apartados que corresponden a los objetivos específicos programados.

4.1 Elaboración de un marco de evaluación cualitativo y cuantitativo adaptable a distintos sistemas sanitarios para evaluar el nivel de implementación de la MPP.

Con el objetivo de evaluar el grado de implementación de la MPP a nivel nacional, se identificaron marcos de evaluación para evaluar cualitativa y cuantitativamente esta implementación. Sin embargo, los marcos encontrados no eran adaptables a diferentes entornos sanitarios (Agarwal et al. 2021; Doyle et al. 2018; Lee y Kim 2022). El estudio de Agarwal y colaboradores (Agarwal et al. 2021) presenta un marco de evaluación específico para organizaciones sanitarias estadounidenses. El marco publicado en 2018 por Doyle y colaboradores (Doyle et al. 2018) recoge solamente los aspectos organizacionales de los centros que proporcionan MPP. Y por último, en el estudio de Lee y Kim (Lee y Kim 2022) se ha desarrollado un marco que contempla aspectos tanto técnicos como organizacionales de las estrategias de implementación de MPP, pero centrados únicamente en Corea del Sur.

Se justifica así la necesidad de creación de un marco de evaluación de implementación de la MPP que sea adaptable a distintos entornos sanitarios. Una vez creado este marco, se podría proceder al proceso de evaluación.

Para la creación del marco de evaluación del grado de implementación de MPP a nivel nacional, se realiza una RS de la bibliografía existente para la identificación de los elementos evaluables que debe de contener dicho marco y el sistema de puntuación para un análisis cuantitativo. Esta RS se realizó siguiendo la metodología PRISMA (*Preferred Reporting Items for Systematic Review and Meta-Analyses*)(Liberati et al.

Resultados

2009), y su protocolo se registró en el repositorio PROSPERO con el identificador: CRD42022338611. Este protocolo se puede consultar en la web del repositorio (<https://www.crd.york.ac.uk/prospero/>) a través del identificador asignado. Dado que los datos extraídos son de tipo cualitativo (elementos y dominios evaluables) la RS finaliza con un análisis temático de los estudios incluidos, siguiendo la metodología de Thomas y Harden (Thomas y Harden 2008).

En los siguientes apartados se plasman los resultados de cada una de las etapas de la RS y el marco de evaluación generado.

4.1.1 Generación del marco de evaluación de la implementación de la MPP a nivel nacional: revisión sistemática

Formulación de la pregunta de investigación

La pregunta de investigación para esta RS es: ¿Cuáles son los dominios para la determinación del nivel de implementación de la MPP a nivel nacional?

Descripción de los criterios de inclusión y exclusión de los estudios

Los estudios que se incluyen en esta RS deben responder a la pregunta de investigación planteada y cumplir los siguientes criterios de inclusión: (1) respecto a la intervención debe versar sobre la implementación de la medicina personalizada y los servicios ómicos en los sistemas sanitarios públicos, (2) respecto al diseño de los estudios se incluyeron revisiones sistemáticas, revisiones narrativas, informes técnicos, informes gubernamentales y textos legislativos vigentes, (3) por fecha se restringió la búsqueda a estudios posteriores a enero de 2016. Se excluyen aquellos estudios que no cumplan alguno de los criterios de inclusión o que no estén disponibles en español o inglés.

Estrategias de búsqueda

Para recopilar la evidencia científica se realiza una búsqueda bibliográfica (fecha de última actualización de la búsqueda: 14 de febrero de 2022) en las siguientes **bases de datos** de referencia: Medline (Ovid), Embase (*Excerpta Medica Database*), WoS (*Science Citation Index*) y PubMed (*Ahead of print/First online*). Para la búsqueda de estudios se utilizó tanto lenguaje controlado (*Genetic Testing, Precision Medicine, Health Plan Implementation*) como terminología libre (*genomic service, pharmacogenetics, personalized medicine, health policy*), habiéndose adaptado la

estrategia inicial a la sintaxis de cada base de datos. Estas búsquedas se limitaron por fecha a los estudios publicados a partir de 2016, y por idioma a los estudios publicados en inglés o español. En el anexo 1 se detallan las estrategias de búsqueda de la RS para cada una de las bases de datos interrogadas. Por otro lado, se realizó una **búsqueda secundaria** de recursos, que arrojó estudios adicionales. Esta búsqueda secundaria se realizó a través de la base de datos INAHTA (*International Network of Agencies for Health Technology Assessment*) para detectar informes de agencias internacionales de evaluación de tecnologías sanitarias y a través de recursos institucionales o gubernamentales de documentos, estrategias y normativas nacionales de política sanitaria.

El número total de estudios recuperados entre todas las bases de datos fue de 1.432 estudios. Tras eliminar los duplicados, quedaron 1.369 estudios potencialmente relevantes.

Filtrado de los estudios a incluir

Las referencias identificadas durante las búsquedas primarias y secundarias en las bases de datos mencionadas se importaron a la sección de gestión de referencias de la herramienta **Covidence** (<https://www.covidence.org/>). Se procedió a una metodología de filtrado por pares en dos fases. En la primera fase, se filtraron los estudios por sus títulos y resúmenes utilizando los criterios de inclusión y exclusión preestablecidos. Después, siguiendo la misma metodología, se filtraron los estudios a texto completo siguiendo los mismos criterios. Cualquier discrepancia en ambas rondas fue resuelta por ambos revisores mediante consenso.

Tras la primera fase del filtrado se obtuvieron 115 estudios potencialmente relevantes. Estos estudios pasaron a la segunda fase de filtrado, tras ella se obtuvieron 19 estudios que fueron los que se incluyeron para la extracción de datos (Kovanda, Zimani y Peterlin 2021; Pearce et al. 2019; Pitini, D'Andrea, De Vito, Rosso, Unim, Marzuillo, Federici, Maria, et al. 2019; Unim et al. 2019; Bilkey et al. 2019; Burns et al. 2019; Chong, Allotey y Chaiyakunapruk 2018; Doyle et al. 2018; Kurnat-Thoma 2020; Lee y Kim 2022; Pastorino et al. 2021a; Vrijenhoek et al. 2021; Whitsel, Laurie P. et al. 2019; Agarwal et al. 2021; Isaacson Barash 2016; Chanfreau-Coffinier et al. 2019; Delnord, M. et al. 2021; Pritchard et al. 2017; Ricciardi y Boccia 2017).

Resultados

El proceso de filtrado se muestra en el diagrama de flujo PRISMA de la figura 10.

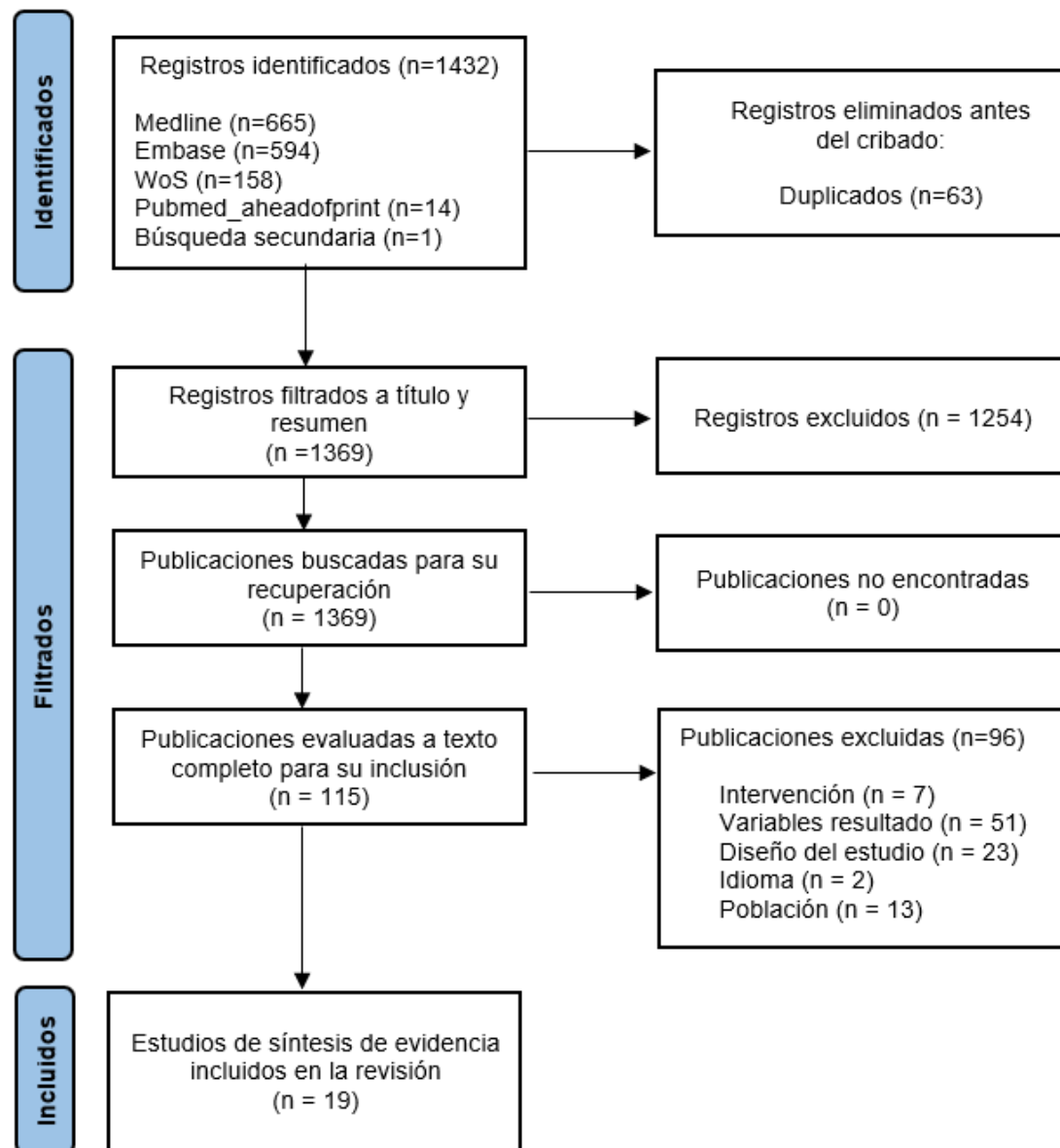


Figura 10: Diagrama PRISMA del filtrado de estudios a incluir.

Evaluación de la calidad de los estudios incluidos

La evaluación de la calidad de los 19 estudios incluidos se realizó también por pares. Se utilizaron las siguientes herramientas en función del tipo de estudio: **AMSTAR-II** (Shea et al. 2017) para evaluar la calidad de las revisiones sistemáticas, **SANRA** (Baethge, Goldbeck-Wood y Mertens 2019) para evaluar las revisiones narrativas, y la **lista de verificación publicada por Humphrey-Murto** (Humphrey-Murto et al. 2017) para las entrevistas y los paneles de expertos. Algunos de los estudios incluidos aparecen evaluados con más de una herramienta, esto se debe a que contienen partes con

distintas metodologías y la evaluación se hizo de forma desagregada para cada una de ellas.

La calidad de las revisiones sistemáticas fue, según los criterios de la herramienta AMSTAR-II (Shea et al. 2017), moderada, baja o críticamente baja sin alcanzar ninguna de las revisiones incluidas la calidad alta que es la máxima para esta herramienta. La calidad de las revisiones narrativas osciló, según los criterios de la herramienta SANRA (Baethge, Goldbeck-Wood y Mertens 2019), entre las puntuaciones 12 y 5, al ser la puntuación máxima posible un 12 para esta herramienta. Los autores de la herramienta SANRA indican que un valor 4 o inferior indica una calidad muy pobre de la revisión. La calidad de las entrevistas y paneles de expertos obtuvo unas puntuaciones, según la lista de evaluación publicada por Humphrey-Murto (Humphrey-Murto et al. 2017), entre 11 y 3, al ser la puntuación máxima posible para esta herramienta de 11. Finalmente, no se excluyeron estudios por motivos de baja calidad. En el anexo 2 se detallan los resultados de la evaluación de la calidad de cada uno de los estudios incluidos en la RS.

Extracción de la información general de los estudios incluidos

La información general de los estudios incluidos se extrajo por pares en tablas de Microsoft Excel adaptadas para esta RS. Algunos de los datos extraídos son independientes del tipo de estudio (autores y año de publicación), pero la mayoría no lo son. Por ello, estas tablas se dividieron según el tipo de estudios en tres (RS, RN y otro tipo de estudios) debido a que sus diferentes características requerían extraer distintos datos. Algunos de los estudios incluidos aparecen en más de una tabla de extracción de descripción, ya que contienen partes con distintas metodologías y la extracción de datos se hizo de forma desagregada para cada una de ellas.

- Para las **revisiones sistemáticas** incluidas, 4 estudios, la información extraída sobre sus características generales fue: el primer autor y año de publicación de la revisión, bases de datos interrogadas y fecha de búsqueda, criterios de inclusión de estudios, número de estudios incluidos y objetivo de la revisión. En la tabla 2 se muestran estos datos para las revisiones sistemáticas incluidas.
- Para las **revisiones narrativas** incluidas, 9 estudios, la información extraída sobre sus características generales fue: primer autor y año de publicación, principales

Resultados

características de la metodología y objetivo de la revisión. En la tabla 3 se muestran estos datos para las revisiones narrativas incluidas.

- La metodología del resto de estudios, aquellos que no son revisiones sistemáticas o narrativas o partes de estas revisiones que contenían otra metodología, incluyen: **encuestas, cuestionarios, entrevistas, debates y paneles de expertos**. La información extraída para este tipo de estudios, 7 estudios, fue: primer autor y año de publicación, principales características de la metodología y objetivo del estudio. En la tabla 4 se muestran estos datos para las encuestas, cuestionarios, entrevistas, debates y paneles de expertos incluidos.

Autor y año	Bases de datos y fecha de búsqueda	Criterios de selección	Número de estudios incluidos	Objetivo
Kovanda et al.(Kovanda, Zimani y Peterlin 2021) 2021	PubMed, Google, European Genome Phenome Archive-EGA Abril 2020	Incluye proyectos de implementación de la medicina genómica a nivel nacional. Se excluyeron proyectos ya concluidos en la fecha de búsqueda, proyectos internacionales, y proyectos que solamente tengan como objetivo la obtención de muestras de secuenciación.	41 estudios incluidos de 86 países	Ofrecer una visión general de la información disponible sobre los proyectos genómicos nacionales activos en todo el mundo para identificar las características comunes y las diferencias entre ellos.
Pearce et al.(Pearce et al. 2019) 2019	Web of Science y PubMed Abril 2018	Incluye estudios de investigación primaria y revisiones sistemáticas. La búsqueda no se limitó a los estudios del Reino Unido, ni a países con sistemas sanitarios similares. Se excluyeron los artículos que no estaban escritos en inglés.	55 estudios incluidos (EEUU = 29, Reino Unido = 10, Canadá = 4, Holanda = 2, Estonia = 1, Corea del Sur = 1 y Cuba = 1)	Identificación de los dominios clave para implantar la medicina genómica en la práctica clínica y la evaluación del estado actual del Servicio Nacional de Salud del Reino Unido para implementar un Servicio de Medicina Genómica.

Resultados

		Se excluyeron los editoriales, los comentarios, los resúmenes de conferencias y documentos metodológicos sin presentación de los resultados de la investigación. También se excluyeron los artículos con un enfoque molecular/biológico, trabajos sobre genética peri o prenatal, y los artículos sobre investigación genética.		
Pitini et al.(Pitini, D'Andrea, De Vito, Rosso, Unim, Marzuillo, Federici, Di	PubMed, SCOPUS, ISI Web of Knowledge, Google Scholar y Google, y los sitios web de las principales agencias gubernamentales y organizaciones de investigación que participan en la evaluación de las pruebas genéticas	Cualquier documento que describiera un marco de evaluación original para las pruebas genéticas.	29 estudios incluidos	Identificar y comparar los marcos existentes de evaluación de pruebas genómicas, centrándose especialmente en sus criterios de evaluación (la validez analítica y clínica, la utilidad clínica...).

Resultados

Maria, et al. 2019) 2019	Abril 2017			
Unim et al.(Unim et al. 2019) 2019	PubMed, Scopus, Web Of Science, Google y Google Scholar 2015	Incluye artículos de informes relevantes sobre estudios piloto, mejores prácticas y proyectos financiados inherentes a la prestación de servicios genéticos; provisión de todo tipo de pruebas genéticas por parte de equipos de especialistas en genética y profesionales sanitarios que ejercen en atención primaria o secundaria; intervenciones realizadas en países europeos y extraeuropeos (anglófonos) (Estados Unidos, Canadá, Australia y Nueva Zelanda).	117 estudios incluidos	Evaluar los servicios genéticos en países europeos y no-europeos (anglófonos) para identificar y clasificar los modelos de implementación de las pruebas genéticas.

Resultados

		Se incluyeron estudios publicados en inglés e italiano entre 2000 y 2015. Se excluyeron estudios que sólo informaran sobre servicios de asesoramiento genético; estudios descriptivos en los que las vías de atención no estaban bien definidas; y estudios que no especificaban el tipo de prueba genética considerada.		
--	--	--	--	--

Tabla 2: Descripción de las revisiones sistemáticas incluidas en la revisión.

Resultados

Autor y año	Características metodológicas del estudio	Objetivo
Bilkey et al.(Bilkey et al. 2019) 2019	-	Descripción de las aplicaciones de la atención sanitaria genómica, su alcance y los retos éticos, legales y sociales que plantean en las distintas fases del ciclo de desarrollo.
Burns et al.(Burns et al. 2019) 2019	-	Consideraciones clave para integrar con éxito las tecnologías genómicas en los sistemas sanitarios. Descripción de la implementación efectiva y sostenible de las pruebas genómicas dentro de las áreas estratégicas prioritarias del Marco de Política Nacional de Genómica Sanitaria de Australia.
Chong et al.(Chong, Allotey y Chaiyakunapruk 2018) 2018	47 recursos incluidos (38 artículos, 6 guías de organismos gubernamentales y 3 sitios webs). Bases de datos: PubMed, EMBASE, CINAHL, Cochrane Library y Web Of Science, y literatura gris. Fecha búsqueda: 31 enero 2017. Países incluidos: Indonesia, Malasia, Singapur y Tailandia.	Describir los enfoques actuales de las políticas y programas en el sudeste asiático para promover la adopción y la aplicación de la medicina personalizada. Destacar los desafíos y proponer estrategias para el desarrollo futuro.

Resultados

Doyle et al.(Doyle et al. 2018) 2018	Redacción en grupo de trabajo con proceso interactivo de 8 miembros de <i>Genomics and Population Health Action Collaborative</i> . Se identifican 38 medidas para la implementación de la medicina genómica.	Desarrollo de un marco organizativo para la implementación de los servicios genómicos, a través de la identificación de amplios dominios de los programas estatales de genómica de la salud pública (con énfasis para las enfermedades: síndrome de Lynch, cáncer de colon hereditario y cáncer de mama y ovario hereditario). Estas medidas de la implementación formarán parte de un conjunto de herramientas de salud pública genómica disponible en línea.
Kurnat-Thoma(Kurnat-Thoma 2020) 2020	80 recursos incluidos (51 artículos y 29 recursos online). Bases de datos: PubMed, CINAHL y GoogleScholar. MeSH terms: Clinical Laboratory Improvement Amendments certified laboratory genetic tests, health system implementation, genetic test implementation, precision medicine/precision health genomic medicine implementation, workforce development, and clinical decision support.	Descripción de las tendencias actuales (2014-2019) en Estados Unidos del uso de pruebas genéticas en el sistema sanitario, de los recursos de educación y formación del personal sanitario en genética/genómica y de las herramientas de implementación de la medicina genómica en la práctica clínica.
Lee & Kim(Lee y Kim 2022)	5647 proyectos de I+D relacionados con medicina personalizada incluidos.	Desarrollo de un marco de innovación global preciso con dimensiones regionales, técnicas y organizativas para establecer la colaboración entre las

2022	Bases de datos: Instituto Nacional de Ciencia y Tecnología Corea del Sur.	diferentes partes interesadas, creando así una base para una estrategia nacional global de medicina personalizada en Corea del Sur.
Pastorino et al.(Pastorino et al. 2021b) 2021	-	Descripción de las cuestiones éticas planteadas con la implementación de la medicina personalizada, cómo pueden estas cuestiones afectar a su eficacia o eficiencia y mecanismos para mitigar estos efectos.
Vrijenhoek et al.(Vrijenhoek et al. 2021) 2021	-	Proporcionar una visión general de los servicios genéticos ofertados en Estonia, Finlandia y Holanda, y considerar el impacto de hechos históricos y recientes en el desarrollo de la genética clínica. Describir las características claves de los servicios genómicos en estos 3 países e identificar semejanzas y diferencias entre los sistemas a nivel de cultura, estructura y práctica clínica, y cómo esto influye en su futuro desarrollo.
Whitsel et al. (Whitsel, Laurie P et al. 2019) 2019	-	Destacar el importante papel del gobierno estadounidense, en sus diferentes niveles organizativos, en las diferentes etapas de desarrollo de la medicina personalizada (desde la ciencia básica hasta la traslación).

Tabla 3: Descripción de las revisiones narrativas incluidas en la revisión.

Autor y año	Características metodológicas del estudio	Objetivo
Agarwal et al.(Agarwal et al. 2021) 2021	Encuesta estructurada a 4 organizaciones sanitarias de Estados Unidos líderes en MP: dos sistemas de salud comunitarios, una red de distribución y un sistema de salud académico. Las entrevistas se llevaron a cabo de forma doblemente ciega. No se revelan los nombres de las instituciones que fueron entrevistadas.	Conocer la forma en que los sistemas sanitarios han adoptado la gestión de proyectos, comprender los retos asociados a la integración clínica de la gestión de proyectos y desarrollar una perspectiva sobre el alcance de la adopción de la gestión de proyectos en estas diferentes instituciones. Estas entrevistas fueron la base para identificar los parámetros valorados por las instituciones que están a la vanguardia de la gestión de proyectos y los parámetros clave con los que estas organizaciones miden el nivel de adopción de la gestión de proyectos. A partir de estas entrevistas se definieron ocho criterios de evaluación para el marco: tres de estos criterios reflejan las pruebas realizadas y los datos recogidos para hacer posible la medicina personalizada, y las categorías restantes eran la orientación de las pruebas y la accesibilidad de los datos, el apoyo del liderazgo, la financiación interna, la utilización de los datos y los esfuerzos para compartirlos.
Barash(Isaacson Barash 2016) 2016	Encuesta a 60 asistentes a la 11ª Conferencia de Asia-Pacífico sobre Genética Humana, 2015, Hanoi (Vietnam).	Declaración regional de consenso sobre las necesidades para avanzar en la implementación de la medicina personalizada y la genómica en la práctica

Resultados

	Cuestionario de 5 preguntas, 2 de ellas abiertas.	clínica para ayudar a las organizaciones de financiación y a los responsables políticos a promover el intercambio de beneficios genómicos a nivel mundial.
Chanfrau-Coffinier et al.(Chanfreau-Coffinier, Peredo, Russell, Marcia M, et al. 2019) 2019	Entrevista a 57 participantes de la Conferencia de la <i>Veterans Health Administration</i> , organizada el 25 agosto 2016 en Arlington, Virginia, (11 administradores/gestores, 22 clínicos, 6 pacientes/familiares, 5 responsables políticos/pagadores, 13 investigadores).	Desarrollo de un modelo para facilitar la implementación de la medicina de precisión en <i>Veterans Health Administration</i> y otras organizaciones sanitarias.
Delnord et al.(Delnord, M et al. 2021) 2021	Encuesta a 14 expertos del <i>ComPerMed Board</i> consistente en un cuestionario con 28 preguntas agrupadas en 4 categorías (calidad, capacidad de respuesta, compromiso de las partes interesadas e integración del conocimiento). Sistema de puntuación con escala Likert, documentación aportada previamente <i>NGS Roadbook</i> .	Evaluar la implementación del <i>NGS Roadbook</i> en el sistema sanitario belga y el impacto de esta iniciativa en el uso de las NGS en oncohematología.
Pitini(Pitini, D'Andrea, De	Panel de expertos con metodología Delphi en 3 rondas; a 55 miembros de <i>Italian Network of Public</i>	Desarrollo de un framework para la evaluación integral de las pruebas genéticas que incluya la evaluación de la prestación de servicios.

Resultados

Vito, Rosso, Unim, Marzuillo, Federici, Di Maria, et al. 2019)	2019	<i>Health Genomics</i> (34 médicos, 20 biólogos y 1 economista).	
Pritchard et al.(Pritchard et al. 2017)	2017	Encuesta a grupo de Trabajo de Sanidad compuesto por representantes de 49 organizaciones de servicios de salud (19 centros de salud académicos, 12 sistemas de salud comunitarios, 16 organizaciones de apoyo a la prestación de salud, y 2 grupos de médicos). Debate a 54 participantes divididos en 3 grupos focales (17 proveedores, 23 expertos de la industria y 14 pacientes), a cada uno se les planteó una tanda de 5 preguntas diferentes.	Identificación de áreas sanitarias que requieren mejoras para la implementación de la medicina personalizada y las estrategias para abordar estas mejoras analizadas principalmente en las organizaciones sanitarias de Estados Unidos.

Ricciardi et al.(Ricciardi y Boccia 2017) 2017	Debate de un grupo de expertos en sanidad, política pública, salud pública, legislación y ética de Europa, Estados Unidos y Canadá sobre el futuro de la asistencia sanitaria, organizado por la <i>Foundation for Genomics and Population Health</i> .	Descripción de los retos, los beneficios y las consecuencias de la implantación de la medicina personalizada por parte de responsables políticos y personal sanitario.
---	---	--

Tabla 4: Descripción de otros tipos de estudios incluidos (encuestas, cuestionarios, entrevistas, debates y paneles de expertos) incluidos en la revisión.

Resultados

Análisis temático

La información extraída de los estudios incluidos es de tipo cualitativo, por lo que para su análisis se adopta una metodología que permite el análisis de este tipo de datos: el análisis temático. Siguiendo la metodología descrita por Thomas y Harden (Thomas y Harden 2008), se identificaron y extrajeron por pares los ítems considerados en cada uno de los estudios incluidos. La clasificación en dominios se basó en la lista de estos ítems recuperada del análisis temático, mediante un proceso de consenso. Tras esta clasificación, los ítems adquirieron el estatus de elementos evaluables del marco dentro de su dominio correspondiente. Se incluyeron un total de 11 dominios que contienen 37 elementos evaluables. La figura 11 plasma los dominios incluidos en este marco.



Figura 11: Representación de los dominios de evaluación de la implementación de la MPP a nivel nacional.

La tabla 5 resume estos dominios y los elementos evaluables que contienen. En el anexo 3 se detalla de cuál de los estudios incluidos se ha extraído cada uno de los elementos evaluables que constituyen los dominios

Políticas sanitarias
• Puesta en marcha de medidas legislativas (proyecto o iniciativa existente para alcanzar este objetivo) • Legislación/Regulación • Grupos de trabajo de colaboración entre las distintas partes interesadas
Integración de la MPP en el sistema sanitario
• Cartera de servicios de MPP • Nivel de accesibilidad de la MPP en el sistema sanitario • Cartera de finalidades asistenciales
Investigación básica y traslacional
• Plan de fomento de la investigación básica en MPP y su traslación a la investigación clínica
Infraestructuras y recursos
• Proyectos de recogida de información ómica a nivel poblacional • Biobancos (muestras biológicas a nivel poblacional o de enfermedades específicas) • Plataformas de almacenamiento de historias clínicas electrónicas • Plataformas de almacenamiento de datos ómicos • Plataformas de análisis de datos ómicos y big data
Gestión y disponibilidad de datos
• Armonización, calidad y protección de las historias clínicas electrónicas • Armonización, calidad y protección de los datos ómicos • Incorporación de datos ómicos a la historia clínica electrónica • Acceso a los datos ómicos e interoperabilidad entre profesionales y entidades • Uso de datos ómicos en la toma de decisiones clínicas
Áreas organizacionales
• Interoperabilidad entre organizaciones y recursos de investigación básica y traslacional • Introducción de áreas especializadas en MPP • Refuerzo de las áreas no especializadas en MPP

Resultados

• Desarrollo y adopción de directrices de procedimiento • Estructura organizativa de los ensayos ómicos • Equipamiento de ensayos ómicos
Implicaciones éticas, legales y sociales (ELSI)
• Nivel asistencial de información al paciente • Formularios normalizados de consentimiento informado del paciente y de aceptación y compromiso del paciente • Mecanismos de protección de datos
Necesidades educativas
• Educación y formación del personal sanitario especializado y no especializado en MPP • Educación de pacientes y familiares en MPP • Actividades de sensibilización y divulgación para la ciudadanía
Evaluación de tecnologías sanitarias (HTA)
• Plan de HTA específico de MPP • Órgano de HTA • Metodología de HTA de MPP • Grupo de toma de decisiones de HTA
Evaluación de implementación
• Plan de evaluación de la aplicación del MPP • Órgano de evaluación de la aplicación • Metodología de evaluación de la aplicación
Financiación
• Previsión de presupuesto de ejecución MPP

Tabla 5: Resumen de los dominios y elementos evaluables del marco para la evaluación de la implementación de la MPP a nivel nacional.

A continuación, se definen los dominios, los elementos evaluables y se referencian en cada uno de ellos los estudios incluidos de los que han sido extraídos:

Políticas de salud (Agarwal et al. 2021; Whitsel, Laurie P et al. 2019; Ricciardi y Boccia 2017; Pitini, D'Andrea, De Vito, Rosso, Unim, Marzuillo, Federici, Maria, et al. 2019; Bilkey et al. 2019): este dominio hace referencia a la puesta en marcha de medidas

legislativas que proporcionen un marco legal adecuado para el acceso de los pacientes a la MPP con todas las garantías. En este dominio se incluyen como elementos valorables la existencia de legislación o regulación específica para uno o más ámbitos de la MPP y el establecimiento de grupos de trabajo legislativos para la elaboración de estas normas, proyectos o iniciativas.

Integración de la MPP en el sistema sanitario (Agarwal et al. 2021; Unim et al. 2019; Burns et al. 2019; Chong, Allotey y Chaiyakunapruk 2018; Doyle et al. 2018; Kurnat-Thoma 2020): este dominio contempla la incorporación de la MPP y las ciencias ómicas en el sistema sanitario nacional. Los elementos cuantificables incluidos en este dominio son la existencia de una cartera de servicios de MPP, un nivel de accesibilidad alto a la MPP para los pacientes del sistema sanitario y la recogida de las finalidades asistenciales para las que se considera el uso de la MPP.

Investigación básica y traslación (Pritchard et al. 2017; Ricciardi y Boccia 2017; Doyle et al. 2018): este dominio contiene un único elemento puntuable, la existencia de un plan de fomento de la investigación básica y traslacional en MPP.

Infraestructuras y recursos (Kovanda, Zimani y Peterlin 2021; Delnord, M. et al. 2021; Chong, Allotey y Chaiyakunapruk 2018; Doyle et al. 2018; Kurnat-Thoma 2020): este dominio versa sobre la obtención de muestras y datos clínicos a nivel poblacional y de su correcto almacenamiento y análisis. En este dominio se incluyen cinco elementos puntuables: la existencia de proyectos de obtención de información ómica a nivel poblacional, la existencia o creación de biobancos para el depósito y custodia de muestras biológicas, y la existencia de plataformas para el almacenamiento de historiales clínicos electrónicos y datos ómicos y su análisis.

Gestión y disponibilidad de datos (Isaacson Barash 2016; Kovanda, Zimani y Peterlin 2021; Chanfreau-Coffinier et al. 2019; Delnord, M. et al. 2021; Pritchard et al. 2017; Unim et al. 2019; Chong, Allotey y Chaiyakunapruk 2018; Doyle et al. 2018): este dominio contempla las necesidades específicas a nivel de los datos que generan las pruebas ómicas y la MPP para que puedan ser usados de forma efectiva y segura en la toma de decisiones clínicas. Para determinar este dominio se han recogido cinco elementos puntuables: armonización, calidad y protección de historiales clínicos

Resultados

electrónicos y datos ómicos, incorporación de los datos ómicos en los historiales clínicos electrónicos, acceso e interoperabilidad entre facultativos y entidades a los datos, y utilización de los datos ómicos en la toma de decisiones clínicas.

Áreas organizacionales (Lee y Kim 2022; Agarwal et al. 2021; Kovanda, Zimani y Peterlin 2021; Pastorino et al. 2021a; Whitsel, Laurie P et al. 2019; Delnord, M. et al. 2021; Ricciardi y Boccia 2017; Unim et al. 2019; Chong, Allotey y Chaiyakunapruk 2018; Kurnat-Thoma 2020): se requieren cambios a nivel organizativo al contemplar la MPP como un cambio en el enfoque de la medicina. Estos cambios surgen de la necesidad de especialización del sistema y del personal sanitario, y se recogen en los seis elementos medibles de los que consta este dominio: interoperabilidad entre organizaciones y recursos de investigación, implementación de áreas sanitarias especializadas en MPP, refuerzo de áreas sanitarias no especializadas involucradas en MPP, desarrollo y adopción de procedimientos que incluyan la MP, creación de estructuras organizativas para la realización de pruebas ómicas, y existencia de equipos para la realización de las pruebas.

Implicaciones éticas, legales y sociales (ELSI, del inglés *ethical, legal and social implications*) (Isaacson Barash 2016; Pastorino et al. 2021a; Vrijenhoek et al. 2021; Delnord, M. et al. 2021; Doyle et al. 2018): la correcta información al paciente y la cobertura legal de sus datos deben ser tenidas en cuenta en la MPP dada la sensibilidad de los datos, las posibles repercusiones de los resultados en su salud y la posible detección en algunas de las pruebas de hallazgos accidentales. Por ello en este dominio se incluyen los elementos valorables: existencia de formularios estandarizados de consentimiento informado al paciente antes de la realización de la prueba, mecanismos de protección de los datos y los resultados, y estandarización del nivel asistencial encargado de la información al paciente.

Necesidades educativas (Vrijenhoek et al. 2021; Whitsel, Laurie P et al. 2019; Delnord, M. et al. 2021; Ricciardi y Boccia 2017; Pitini, D'Andrea, De Vito, Rosso, Unim, Marzuillo, Federici, Maria, et al. 2019): dada la gran innovación que supone la MPP para el personal sanitario, los pacientes, los familiares y la sociedad, es fundamental su correcta formación y concienciación de su uso. Para contemplar todas las necesidades educativas se recogen tres elementos: la existencia de planes de educación y formación para el

personal sanitario especializado y no especializado en MPP, la existencia de planes educativos para pacientes y familiares en MPP, y la existencia de actividades para la concienciación de la ciudadanía en la utilidad de la MPP.

Evaluación de tecnologías sanitarias (HTA, *health technologies assessment*) (Bilkey et al. 2019; Doyle et al. 2018; Chong, Allotey y Chaiyakunapruk 2018): las tecnologías sanitarias usadas en las pruebas ómicas tienen unas particularidades que hacen necesaria la adaptación de la metodología de evaluación de tecnologías sanitarias. Para que estas particularidades no supongan un obstáculo para su evaluación, y por lo tanto para su implantación, se han recogido cuatro elementos: existencia de un plan de evaluación específico para el HTA en MPP, creación de un organismo de referencia para el HTA de tecnologías de MPP, adaptación específica para la MPP de una metodología de HTA y creación de un grupo de expertos para la toma de decisiones en la HTA de estas tecnologías.

Evaluación de la implementación (Pastorino et al. 2021a; Ricciardi y Boccia 2017; Pitini, D'Andrea, De Vito, Rosso, Unim, Marzuillo, Federici, Maria, et al. 2019; Burns et al. 2019; Chong, Allotey y Chaiyakunapruk 2018): el actual auge de la implementación de las tecnologías relacionadas con la MPP hace necesaria la existencia de planes de evaluación de la implementación para su seguimiento y detección de áreas de mejora. Los elementos contemplados en este dominio son: la existencia de un plan de evaluación de la implementación a nivel nacional de la MPP, de un organismo encargado y de una metodología de evaluación adecuada.

Financiación (Pastorino et al. 2021a; Ricciardi y Boccia 2017): la creación y el desarrollo de todos los dominios destacados anteriormente deben contar con una financiación suficiente. Para evaluar este dominio se recoge un único elemento puntuable, la existencia de un plan de financiación para la implementación de la MPP.

Resultados

4.1.2 Generación del marco de evaluación de la implementación de la MPP a nivel nacional: sistema de puntuación

Una vez construido el marco de evaluación de la implementación de la MPP, se desarrolló un sistema para la determinación cualitativa y cuantitativa de esta implementación basada en los elementos recogidos en el marco.

Dentro de cada uno de los dominios para determinar el nivel de implementación de la MPP, deben evaluarse varios elementos. En este marco se recogen 11 dominios que constan de un total de 37 elementos cuantificables de forma individual. Estos elementos se pueden evaluar de forma cualitativa (Sí/No/No consta información) o de forma cuantitativa mediante este **sistema de puntuación**:

- Si el plan de implementación nacional no contempla este elemento o no se encuentra información sobre él se valorará con una puntuación de 0.
- Si el plan de implementación nacional contempla este elemento se valorará con una puntuación de 1.
- Si este elemento ya se encuentra en fase de implementación se valorará con una puntuación de 2.

Basándonos en la metodología aplicada en dos de los estudios incluidos (Agarwal et al. 2021; Kovanda, Zimani y Peterlin 2021), se aplica una fórmula matemática para la normalización del peso de los dominios dentro de la valoración final de la implementación para que aquellos dominios que cuentan con un mayor número de elementos evaluables no tengan un mayor peso en la valoración final de la implementación. Como resultado, la máxima puntuación total que se puede obtener es de 22 puntos.

Para usar este marco de evaluación, el usuario debe previamente recopilar la información relativa a cada uno de los elementos puntuables en las páginas web de los países en evaluación y de los proyectos de implementación propuestos, tanto los que se encuentren en fase de redacción, ejecución o finalizados, siempre que haya información disponible. Este marco pretende ser una guía de toda la información a recopilar y permitir su evaluación una vez recopilada.

Al aplicar este marco con el objeto de comparar o establecer una clasificación entre países se debe observar si ha sido posible obtener y recopilar la información suficiente de cada uno de ellos. De no ser así, la comparación podría no ser representativa de la realidad en cada uno de los sistemas sanitarios comparados.

Otras finalidades con las que podría usarse este marco de evaluación son: identificar países que hayan obtenido una elevada puntuación con este marco para que su sistema de implementación sirva como modelo a otros países, servir como guía para el desarrollo de planes de implementación de la MPP gracias a la identificación de todos los dominios involucrados y detectar debilidades en los planes de implementación en curso o finalizados para poder subsanarlos.

4.2 Evaluación de la implementación de la MPP a nivel nacional en distintos sistemas sanitarios mediante el marco generado.

Una vez generado el marco de evaluación cuantitativa y cualitativa de la MPP gracias a la metodología anteriormente expuesta se procedió a su aplicación. Para ello, se eligieron diversos países a nivel mundial con grandes diferencias entre sus sistemas sanitarios, lo cual validaría la adaptabilidad del marco a diferentes sistemas sanitarios. Se decidió excluir en esta etapa de evaluación a España por la posible parcialidad en su evaluación debido al conocimiento previo de su sistema sanitario por parte de la autora de esta tesis, lo cual sería una fuente de heterogeneidad en este estudio. Con la finalidad de obtener la información necesaria para la evaluación de cada uno de los países seleccionados se consultaron los sitios web de los Ministerios de Salud, los documentos políticos publicados, las directrices nacionales, los sitios web de las principales agencias gubernamentales y organizaciones de investigación y los proyectos de implementación de la MPP en redacción, ejecución o finalizados.

Con la intención de probar la adaptabilidad del marco creado a distintos entornos sanitarios, la selección de países se realizó de manera que incluyera a **países de todos los continentes** (Europa, Asia, Norteamérica, Sudamérica, África y Oceanía) y representara distintos niveles de desarrollo económico según la clasificación del *The World Bank*. La clasificación del *The World Bank* clasifica a los países en cuatro **niveles**

Resultados

de desarrollo económico, que es de mayor a menor: renta alta (*high income*), renta media-alta (*upper middle income*), renta media-baja (*lower middle income*) y renta baja (*low income*).

- Los **países europeos** seleccionados para evaluar la implementación de la MPP en ellos fueron: Finlandia (renta alta), Reino Unido (renta alta), Italia (renta alta) y Francia (renta alta).
- Los **países asiáticos** seleccionados para evaluar la implementación de la MPP en ellos fueron: China (renta media-alta), Tailandia (renta media-alta), Japón (renta alta) y Pakistán (renta media-baja).
- El **país sudamericano** seleccionado para evaluar la implementación de la MPP en él fue: Brasil (renta media-alta)
- Los **países norteamericanos** seleccionados para evaluar la implementación de la MPP en ellos fueron: Estados Unidos (renta alta) y Canadá (renta alta).
- El **país oceánico** seleccionado para evaluar la implementación de la MPP fue: Australia (renta alta).
- Los **países africanos** no se pudieron evaluar de forma desagregada por falta de datos, pero sí se encontraron evidencias que permitieron la evaluación de la implementación de la MPP en el conjunto del continente. Si bien esta información se incluye en este trabajo debe ser asumida con cautela por las particularidades que puede presentar cada país.

Las puntuaciones obtenidas para cada uno de los países son las que se muestran en la tabla 6. En el anexo 4 se detalla en una tabla las puntuaciones para cada uno de los elementos evaluables y los resultados obtenidos en la evaluación de los países seleccionados.

País evaluado	Puntuación obtenida
Finlandia	19,75
Reino Unido	19,5
Italia	15
Francia	16,8
China	13,53
Tailandia	5,27
Japón	15,53

Pakistán	0
Brasil	4,6
Estados Unidos	17,67
Canadá	19,77
Australia	20,16

Tabla 6: Resultados obtenidos por el marco generado de la evaluación de la implementación de la MPP en los países seleccionados.

Se obtuvo un rango de puntuación que cubría casi todo el espectro de puntuaciones posibles del marco generado, lo cual valida la selección de estos países y la utilidad del marco. Además, **las puntuaciones parecen estar más relacionadas con el contexto económico y de desarrollo de los países que con su localización geográfica**. Los países clasificados por *The World Bank* por ser países de renta alta presentan una mayor puntuación en la evaluación de la implementación de la MPP.

4.3 Metanálisis de la efectividad de la biopsia líquida para el diagnóstico y la adecuación del tratamiento en cáncer de pulmón

La implementación de la MPP incluye entre otras la incorporación de las tecnologías ómicas y la bioinformática en la rutina clínica. En la introducción de esta tesis se plantea que estas tecnologías pueden ser usadas en distintos puntos del abordaje médico y sobre distintos tipos de muestras. En el caso del cáncer de pulmón, los análisis ómicos, que pueden facilitar el diagnóstico, la adecuación del tratamiento y el pronóstico, pueden realizarse sobre muestras de tejido tumoral o en muestras de sangre, líquido pleural o esputo del paciente. A las muestras alternativas a las biopsias tisulares se las denomina BL. Así pues, nos vamos a centrar en la efectividad de la BL para el diagnóstico y la adecuación del tratamiento. Para **evaluar mediante metanálisis la efectividad de la BL frente a la biopsia tisular para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón** se realizó una nueva RS de la bibliografía publicada. Los resultados de esta segunda RS fueron usados para realizar el posterior MA en aquellos biomarcadores de los que existía evidencia suficiente y los estudios eran comparables.

Resultados

4.3.1 Revisión sistemática de la bibliografía publicada

Formulación de la pregunta de investigación

La pregunta de investigación para esta revisión sistemática es: **¿Es la biopsia líquida tan efectiva como la biopsia de tejido tumoral en la detección de biomarcadores ómicos para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón?**

Esta pregunta de investigación sigue una estructura PIRD-S (*population, index test, reference test, diagnosis of interest y study design*).

Descripción de los criterios de inclusión y exclusión de los estudios

Los estudios que se incluyeron en esta RS debían responder a la pregunta de investigación planteada y cumplir los siguientes criterios de inclusión en ambas fases de filtrado: (1) respecto a la población analizada debe tratarse de pacientes con cáncer de pulmón (de cualquier tipo); (2) respecto a la intervención debe analizar componentes ómicos obtenidos de una BL; (3) respecto al comparador debe ser biopsia tisular (punción por aspiración con aguja fina, biopsia con aguja gruesa o biopsia quirúrgica); (4) respecto a los resultados analizados debe incluir la tasa de concordancia de resultados en BL frente a biopsia tisular, o la sensibilidad y la especificidad de la prueba, o las tasas de falsos positivos y falsos negativos; (5) respecto al diseño de los estudios incluidos solamente se incluirán revisiones sistemática o metanálisis. Se excluyeron aquellos estudios que no cumplieran alguno de los criterios de inclusión o que no estén disponibles en español o inglés.

Estrategias de búsqueda

Para recopilar la evidencia científica se realizó una búsqueda bibliográfica (fecha de última actualización de la búsqueda: 25 de julio de 2022) en las siguientes bases de datos de referencia: Medline (Ovid), Embase (*Excerpta Medica Database*), WoS (*SCI Science Citation Index*) y PubMed (*Ahead of print/First online*).

Para la búsqueda de estudios se utilizó tanto lenguaje controlado (*lung neoplasm, liquid biopsy, biomarkers*) como terminología libre (*circulating tumour DNA, circulating tumour cells, ARNm, ARNInc, miARN*), habiéndose adaptado la estrategia inicial a la sintaxis de cada base de datos. Estas búsquedas no se limitaron por fecha, pero sí por tipo de estudio limitándose a revisiones sistemáticas y/o metanálisis, y por idioma a los

estudios publicados en inglés o español. En el anexo 5 se detallan las estrategias de búsqueda de la RS para cada una de las bases de datos interrogadas.

El número total de estudios recuperados entre todas las bases de datos fue de 1463 estudios. Tras eliminar los duplicados, quedaron **1285 estudios potencialmente relevantes**.

Filtrado de los estudios a incluir

Las referencias identificadas durante las búsquedas en las bases de datos mencionadas, tras la eliminación de duplicados, se importaron a la sección de gestión de referencias de la herramienta informática **Covidence** (<https://www.covidence.org/>). Se procedió a una metodología de filtrado por pares en dos fases. En la primera fase, se filtraron los estudios por sus títulos y resúmenes utilizando los criterios de inclusión y exclusión preestablecidos. Después, siguiendo la misma metodología y criterios, se filtraron los estudios a texto completo. Cualquier discrepancia entre ambos revisores en las dos rondas fue resuelta por consenso.

Tras la primera fase del filtrado se obtuvieron 198 estudios potencialmente relevantes. Estos estudios pasaron a la segunda fase de filtrado, tras ella se obtuvieron **59 estudios** que fueron los que se incluyeron para la extracción de datos. Este proceso de filtrado se muestra en el diagrama de flujo PRISMA de la figura 12.

Resultados

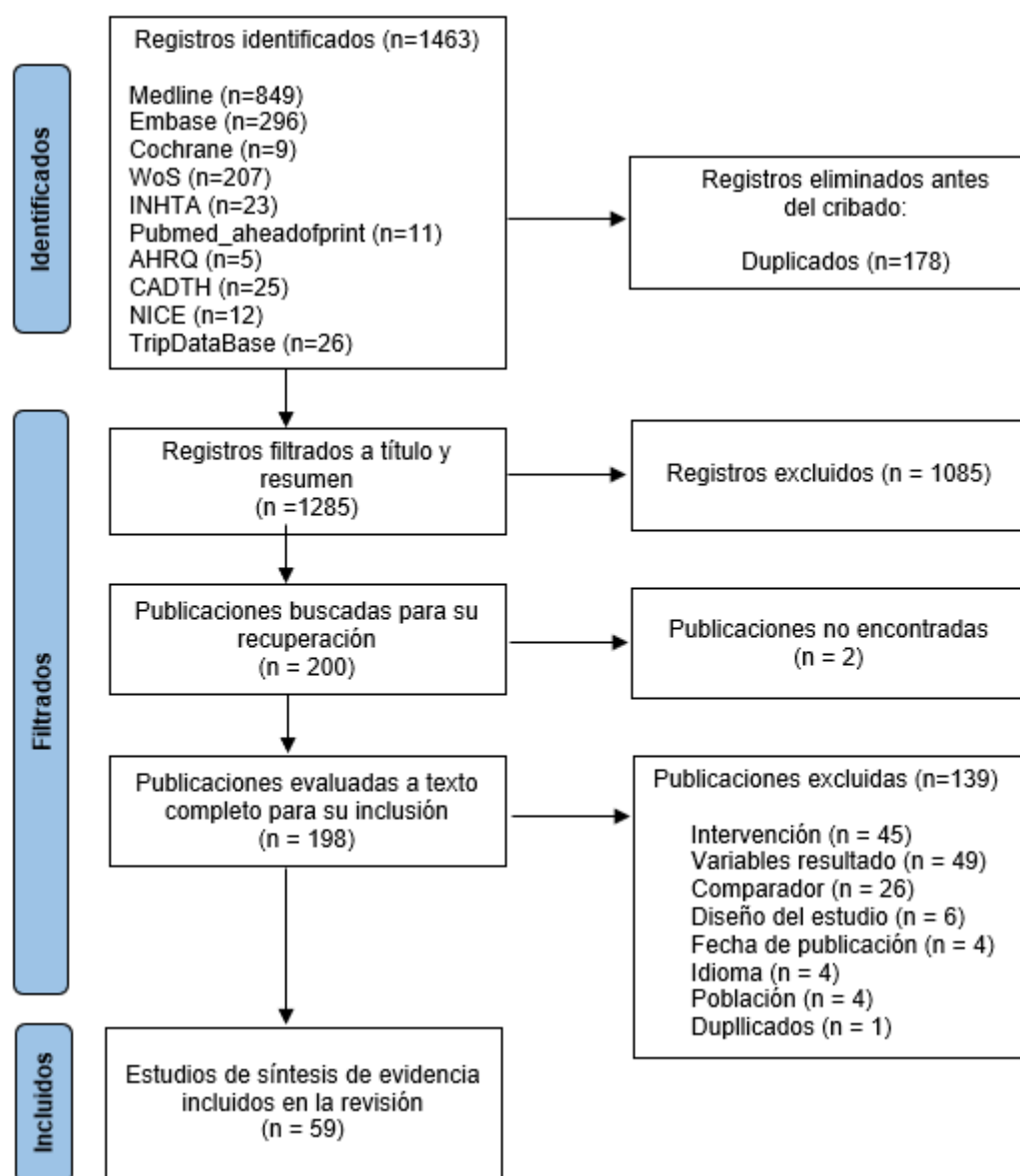


Figura 12: Diagrama PRISMA del filtrado de estudios a incluir.

Extracción de la información general de los estudios incluidos

La extracción de los datos de las revisiones sistemáticas se realizó por pares en tablas en Microsoft Excel diseñadas *ad hoc* para esta revisión. Los datos extraídos incluyen la referencia de la RS incluida, el número de pacientes incluidos en la RS, el tipo de cáncer de pulmón de los pacientes incluidos, las alteraciones ómicas analizadas, el material biológico analizado para la BL, la finalidad del estudio (diagnóstico o adecuación del tratamiento) y los resultados de sensibilidad y especificidad de cada estudio al comparar los resultados de la BL frente a la biopsia de tejido (*goldstandard*). En la tabla 7 se muestra un resumen del número de RS incluidas según las alteraciones ómicas analizadas, algunas de las RS analizan más de una alteración. Todas las RS incluidas

tenían como finalidad el diagnóstico, salvo las 12 RS que analizan las mutaciones puntuales en los genes *EGFR*, *KRAS* o *BRAF* cuya finalidad es la de adecuación del tratamiento. En el anexo 6 se detalla en tablas, separadas según la alteración ómica analizada, toda la información extraída de las revisiones sistemáticas incluidas y sus referencias.

Alteraciones ómicas analizadas	Número de RS incluidas
Mutaciones puntuales en los genes <i>EGFR</i> , <i>KRAS</i> o <i>BRAF</i> en ADN circulante tumoral o en ADN de células tumorales circulantes	12
Alteraciones en los patrones de metilación	5
Alteraciones en la expresión de ARNinc	4
Alteraciones en la expresión de ARNcirc	3
Detección de células circulantes tumorales	4
Cuantificación de ADN circulante	3
Mutaciones en la secuencia genómica	1
Alteraciones en la expresión de miARN	26
Análisis del contenido de exosomas	1

Tabla 7: Número de RS incluidas según las alteraciones ómicas analizadas.

Los criterios para establecer si un biomarcador era susceptible para realizar un MA fueron que al menos 5 de las RS incluidas debían de tener como objetivo principal la evaluación de este biomarcador y entre todas las revisiones incluidas sobre dicho biomarcador debían de aparecer un mínimo de 10 estudios clínicos diferentes. Además, los estudios clínicos incluidos en las revisiones sistemáticas debían estar correctamente referenciados para poder ser encontrados en las bases de datos y ser comparables: detección del mismo biomarcador, análisis del mismo material biológico y poblaciones comparables (mismo tipo de cáncer de pulmón). A pesar de la apabullante cantidad de RS incluidas y datos extraídos, el único biomarcador que contaba con suficiente evidencia para realizar un metanálisis, según los criterios preestablecidos, fue la **detección de mutaciones en *EGFR* en ADN circulante libre en BL para el tipo de cáncer NSCLC con la finalidad de adecuación del tratamiento.**

Para los biomarcadores de diagnóstico o de adecuación del tratamiento que no se consideraron susceptibles de metanálisis, se muestra en la tabla 8 el rango de resultados de las RS para la sensibilidad y especificidad de ese biomarcador en la intervención BL frente a biopsia tisular (*goldstandar*). Es destacable que la mayoría de las alteraciones en BL presentan mayor especificidad que sensibilidad y que presentan una gran

Resultados

heterogeneidad en sus valores. Las alteraciones ómicas que presentan una mayor sensibilidad en BL son las alteraciones en la expresión de miARN o ARNlnc, la cuantificación de ADN circulante y el análisis del contenido de los exosomas. Las alteraciones ómicas que presentan una mayor especificidad en BL son la detección de células tumorales circulantes, el análisis NGS de mutaciones puntuales, las mutaciones puntuales en el gen *BRAF* y las alteraciones en la expresión de miARN. En cualquier caso, estos valores deben de ser tenidos en cuenta con cautela por la gran heterogeneidad de las RS incluidas y de los estudios clínicos que ellas contienen.

Alteraciones ómicas analizadas	Rango de sensibilidad en las RS incluidas	Rango de especificidad en las RS incluidas
Mutaciones puntuales en <i>BRAF</i>	0,31	0,99
Mutaciones puntuales en <i>KRAS</i>	0,39-0,65	0,92-0,95
Alteraciones en los patrones de metilación	0,39-0,62	0,73-0,92
Alteraciones en la expresión de ARNlnc	0,70-0,82	0,67-0,92
Alteraciones en la expresión de ARNcirc	0,76-0,77	0,76-0,77
Detección de células circulantes tumorales	0,63-0,75	0,89-0,96
Cuantificación de ADN circulante	0,79-0,81	0,77-0,81
Análisis NGS de mutaciones puntuales	0,76	0,99
Alteraciones en la expresión de miARN	0,66-0,88	0,75-0,9
Análisis del contenido de exosomas	0,82	0,84

Tabla 8: Resultados de las RS incluidas para la sensibilidad y especificidad de ese biomarcador en la intervención biopsia líquida frente a biopsia tisular.

Al centrarse en el biomarcador seleccionado para su MA (alteraciones puntuales en *EGFR* en ADN circulante libre para la adecuación de tratamiento), se extrajeron además de los datos de las RS correspondientes incluidas, los datos de los estudios clínicos que contenían dichas RS y que fueron localizables con las referencias proporcionadas. Finalmente se incluyeron **86 estudios clínicos** procedentes de las RS incluidas.

La extracción de datos de los estudios clínicos se realizó, al igual que en las otras extracciones, en tablas de Microsoft Excel diseñadas *ad hoc*. Estas tablas incluyen la

siguiente información de cada uno de los estudios: referencia, año de publicación, número de TP, FP, FN y TN, tipo de material biológico analizado para la BL y si el análisis tanto de biopsia tisular como de BL se habían realizado sobre muestras del paciente antes de que este recibiera ningún tratamiento. No se tuvo en cuenta como tratamiento la resecabilidad del tumor, ya que la toma de la muestra de la biopsia tisular podía proceder de este proceso de resección. Sí se tuvieron en cuenta como tratamiento la quimioterapia y fármacos inhibidores de la tirosina quinasa (TKI, *tyrosine kinase inhibitors*). Para aquellos estudios que incluían pacientes sometidos a estos tratamientos y pacientes no sometidos a tratamiento, se consensuó que se incluirían en el subgrupo que representara más del 75% de la cohorte incluida. La tabla construida con esta información se adjunta en el anexo 7.

Evaluación de la calidad de los estudios incluidos

La evaluación de la calidad de los estudios incluidos se realizó por pares. Al incluirse solamente revisiones sistemáticas o metanálisis, la herramienta para la evaluación de su calidad fue AMSTAR-II (Shea et al. 2017). En concreto, se usó la versión *online* de esta herramienta: https://amstar.ca/Amstar_Checklist.php.

En las 59 RS incluidas en este estudio y evaluadas con **AMSTAR-II**, todos los estudios obtuvieron una valoración de calidad moderada. Las preguntas que fueron respondidas de forma negativa en muchos de los estudios incluidos fueron las preguntas 7 y 10, respecto al reporte de los motivos de exclusión de los estudios excluidos y de las fuentes de financiación de los estudios incluidos.

Adicionalmente, también se evaluó la calidad de los 86 estudios clínicos sobre la detección de mutaciones en *EGFR* en ADN libre circulante en BL extraídos de las RS incluidas. Al igual que la evaluación de la calidad para las RS incluidas, esta evaluación también se realizó por pares. La herramienta utilizada fue **QUADAS-2**, al ser una herramienta específica para la evaluación de estudios clínicos de pruebas diagnósticas (Whiting et al. 2011).

Para la evaluación de los 86 estudios clínicos incluidos en el MA de la eficacia de la detección de mutaciones de *EGFR* en BL para la adecuación del tratamiento con la herramienta QUADAS-2, se eliminó la pregunta 2 del dominio 2, al no utilizarse un

Resultados

umbral en esta prueba diagnóstica, sino la presencia/ausencia de mutación en *EGFR*. La mayoría de los estudios incluidos obtuvieron un riesgo de sesgo bajo. Las preguntas con un mayor porcentaje de respuestas negativas o no especificadas fueron la pregunta 1 del dominio 1 (en los estudios en los que no se explicitaba concretamente el método de selección de la cohorte de pacientes) y la pregunta 2 del dominio 2 (en los estudios en los que se conocían los resultados del análisis de biopsia tisular cuando se analizaban los resultados de la BL). En la figura 13 se muestran por porcentaje las respuestas a cada una de las preguntas planteadas por la herramienta y su implicación sobre el riesgo de sesgo y su aplicabilidad (representando en color rojo el porcentaje de estudios que presentan un riesgo de sesgo elevado en esa pregunta, en color amarillo el porcentaje de estudios que presentan un riesgo de sesgo incierto y en color verde el porcentaje de estudios que presentan un riesgo de sesgo bajo). Los resultados de la herramienta QUADAS-2 desagregados por preguntas para cada uno de los estudios clínicos incluidos, se adjuntan en el anexo 8.

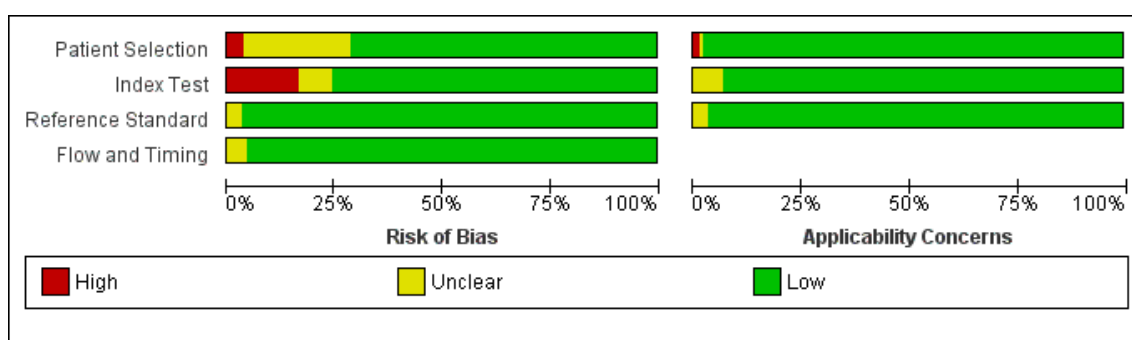


Figura 13: Resultados de la herramienta QUADAS-2 de los estudios clínicos para la evaluación de la eficacia de la detección de mutaciones de *EGFR* en BL

4.3.2 Metanálisis de la efectividad para la adecuación del tratamiento del análisis de las mutaciones en *EGFR* en muestras de biopsia líquida en pacientes de NSCLC.

Un metanálisis es una técnica estadística usada para resumir un único valor de los resultados de dos o más estudios que hayan comparado dos grupos (un grupo intervención y un grupo control) (Fernandez-Chinguel et al. 2019). El MA llevado a cabo tiene como finalidad **evaluar la efectividad para la adecuación del tratamiento del análisis de las mutaciones en el gen *EGFR* en ADN libre circulante en muestras de biopsia líquida en pacientes con NSCLC.**

En este MA se incluyeron los 86 estudios clínicos detectados a partir de las RS incluidas. La evaluación de su calidad con la herramienta QUADAS-2 (Whiting et al. 2011), se ha expuesto en el apartado anterior y en el anexo 7 se recopila la información extraída de ellos (referencia, año de publicación, número de TP, FP, FN y TN, tipo de material biológico analizado para la BL y si el análisis tanto de biopsia tisular como de BL se habían realizado sobre muestras del paciente antes de que este recibiera ningún tratamiento).

En primer lugar, se realizó un MA de los 86 estudios para comparar los resultados de mutaciones en EGFR obtenidos por BL frente a los obtenidos por biopsia tisular (*goldstandard*) con el programa **MetaDisC** (https://metadisc-irycis.shinyapps.io/oct_metadisc/). Se introdujeron los valores de TP, FP, FN y TP extraídos de cada uno de los estudios. Después de eliminar dos de los estudios que no ofrecían ningún resultado TN (He, Tan y Ma 2017; Taniguchi et al. 2011), las siguientes etapas del MA se realizaron con 84 estudios. Las variables resultantes fueron la sensibilidad, la especificidad, el valor DOR y los valores LR+ y LR-. Estos resultados junto a los 95% IC superior e inferior de cada uno de ellos se recogen en la tabla 9. Es reseñable que, al igual que ocurría con los resultados de las RS de los otros biomarcadores, el valor de la especificidad es mayor que el de la sensibilidad. En el anexo 10 se muestran dos representaciones *forest plot* correspondientes a la sensibilidad y especificidad obtenida en el metanálisis todos los estudios clínicos incluidos.

Variable	Valor	95% IC
Sensibilidad	0,69	0,64-0,73
Especificidad	0,97	0,95-0,98
DOR	77,07	44,43-113,68
LR+	24,55	14,75-40,84
LR-	0,32	0,27-0,36

Tabla 9: Resultados del metanálisis para todos los estudios incluidos para evaluar la efectividad de la adecuación del tratamiento del análisis de las mutaciones en EGFR en muestras de biopsia líquida en pacientes de NSCLC al compararlas con biopsias tisulares.

La relación entre la sensibilidad y la especificidad obtenidas en este metanálisis para el análisis de las mutaciones en EGFR en muestras de BL frente a su análisis en biopsia de tejido en pacientes de NSCLC, se muestra en un gráfico tipo curva ROC en la figura 14. En esta figura se puede observar que la prueba objeto de estudio (BL) posee una gran

Resultados

capacidad discriminante, es decir, es capaz de discernir correctamente entre pacientes con la mutación en estudio y pacientes sin la mutación, ya que la curva ROC se sitúa cerca de la esquina superior izquierda del gráfico y alejado de la diagonal. Es importante destacar que, aunque la mayoría de los estudios se sitúan por encima de o en la curva ROC, una minoría de ellos aparecen cerca de la diagonal del gráfico, lo que indica que en estos estudios la prueba no pudo discriminar pacientes con la mutación de pacientes sin la mutación.

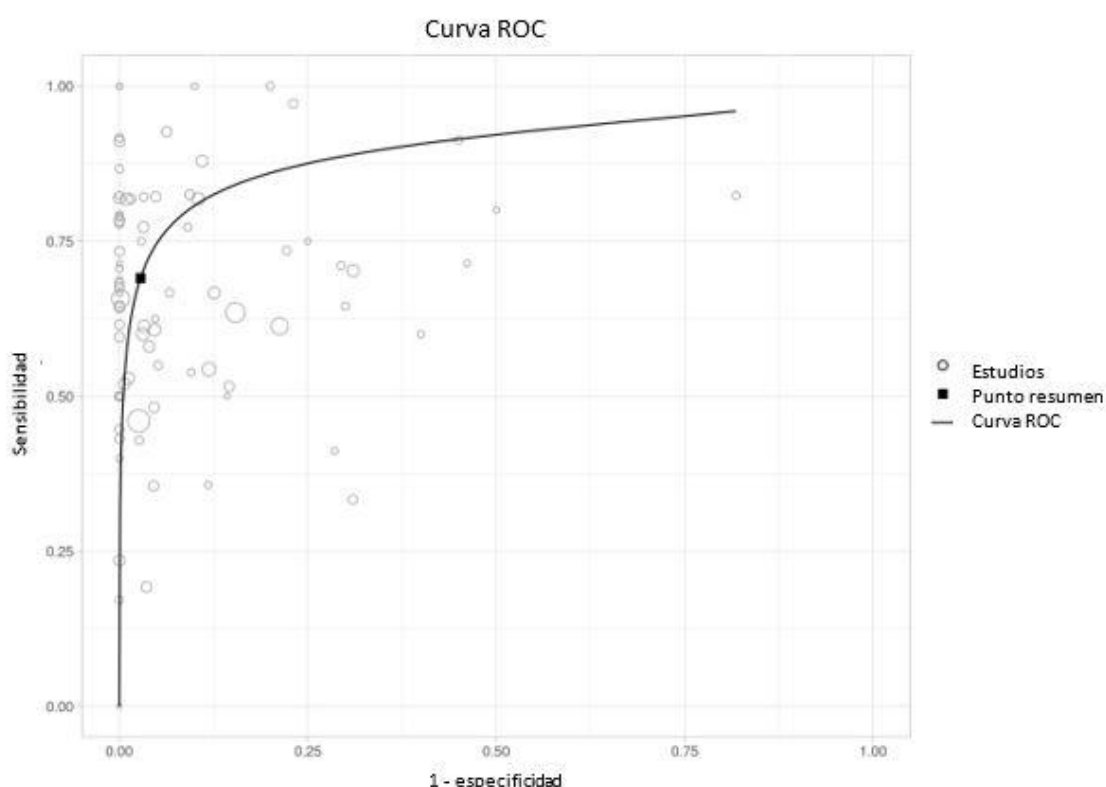


Figura 14: Curva ROC resultado del metanálisis de los estudios clínicos incluidos para evaluar la efectividad de la adecuación del tratamiento del análisis de las mutaciones en EGFR en muestras de biopsia líquida en pacientes de NSCLC al compararlos con biopsias tisulares.

En segundo lugar, se realizó un **metanálisis desagregado** por subgrupos según el **tipo de material biológico analizado** para la BL de los 83 estudios incluidos para el metanálisis. Se excluyó un estudio al ser el único que utilizaba muestras de efusión pleural para la BL, que obtuvo un valor de sensibilidad y especificidad de 0,88 y 0,90 respectivamente (Liu et al. 2018). En este MA se incluyeron 73 estudios que analizaban el plasma de los pacientes y 10 estudios que analizaban su suero. Igualmente se obtuvieron las variables resultado: sensibilidad, especificidad, valor DOR y valores LR+ y LR-. Estos resultados junto a los 95% IC superior e inferior de cada uno de ellos se recogen en la tabla 10. El

resultado de la metarregresión de ambos grupos proporciona un **P** de **0,42**, indicando este valor que no existe una diferencia estadísticamente significativa entre los dos grupos. Esta diferencia no significativa se puede observar de forma visual en la figura 15 en la que se ilustran las curvas ROC para las muestras de plasma (curva roja) y de suero (curva azul). En esta figura se puede observar que ambas curvas casi se solapan.

Variable	Muestras de plasma		Muestras de suero	
	Valor	95% IC	Valor	95% IC
Sensibilidad	0,69	(0,64 - 0,74)	0,66	(0,49 - 0,79)
Especificidad	0,97	(0,95 - 0,98)	0,99	(0,95 - 0,99)
DOR	70,69	(39,50 - 126,49)	166,86	(32,41 - 859,03)
LR+	22,61	(13,25 - 38,58)	58,09	(12,79 - 264,01)
LR-	0,32	(0,27 - 0,37)	0,35	(0,22 - 0,55)

Tabla 10: Resultados del metanálisis desagregado por subgrupos según el material biológico analizado para evaluar la efectividad de la adecuación del tratamiento del análisis de las mutaciones en EGFR en muestras de biopsia líquida en pacientes de NSCLC.

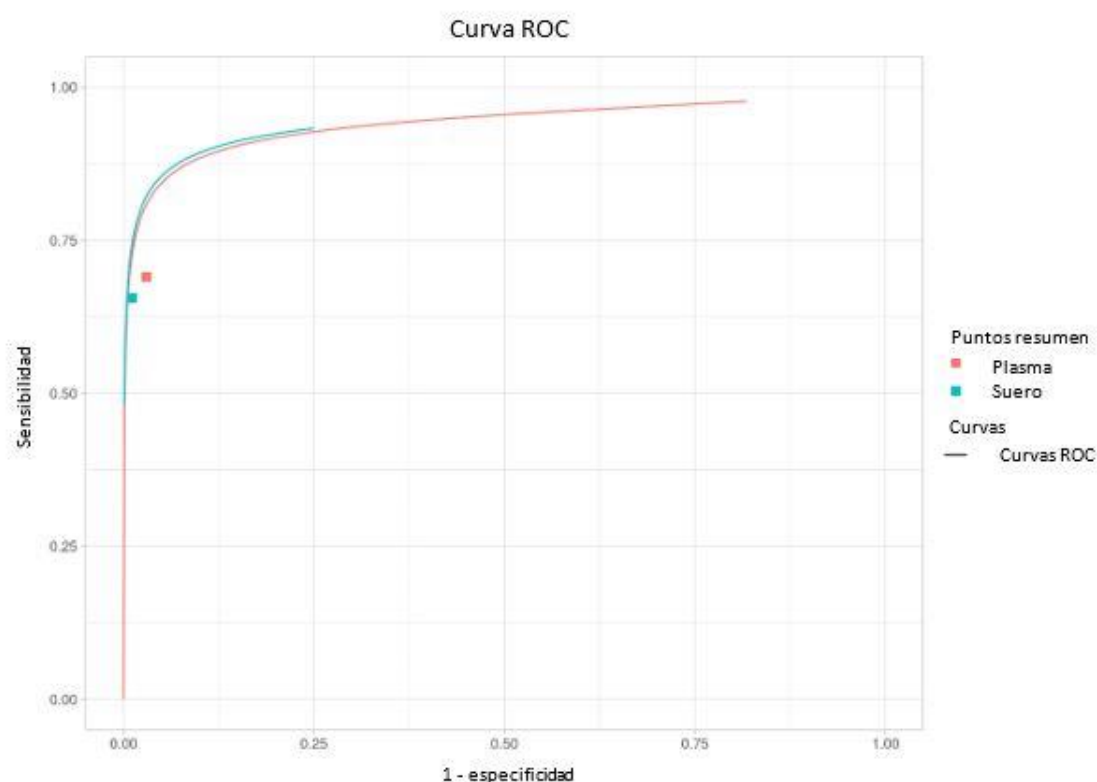


Figura 15: Curva ROC resultado del metanálisis desagregado por subgrupos según el material biológico analizado para evaluar la efectividad de la adecuación del tratamiento del análisis de las mutaciones en EGFR en muestras de biopsia líquida en pacientes de NSCLC.

En tercer lugar, se realizó un **metanálisis desagregado** por subgrupos dependiendo de si el análisis tanto de biopsia tisular como de BL se habían realizado sobre **muestras del**

Resultados

paciente antes de que este recibiera ningún tratamiento (pretratamiento) **o después** (postratamiento). Dos de los estudios clínicos no se incluyeron en este análisis por subgrupos al no presentar más del 75% de los pacientes de ninguno de los subgrupos (Han et al. 2016; Taniguchi et al. 2011). Por lo tanto, este MA se realizó con los resultados de 83 de los estudios incluidos: 53 de los estudios incluían más del 75% de su cohorte pretratamiento y 30 de los estudios incluían más del 75% de su cohorte postratamiento. Al igual que en los metanálisis anteriores las variables resultado analizadas fueron: sensibilidad, especificidad, valor DOR y valores LR+ y LR-. Estos resultados junto a los 95% IC superior e inferior de cada uno de ellos se recogen en la tabla 11. El resultado de la meta regresión de ambos grupos proporciona un **P** de **0,004**, indicando este valor una diferencia significativa entre los dos subgrupos. Esta diferencia significativa se puede observar en la figura 16 en la que se ilustran las curvas ROC para las muestras de pacientes pretratamiento (curva azul) y de pacientes postratamiento (curva roja). En esta figura se puede observar que ambas curvas se encuentran cercanas y en ambos casos la BL sería capaz de discriminar pacientes con la mutación de pacientes sin la mutación en estudio, sin embargo, ambas curvas no se solapan y en el caso de las muestras de pacientes pretratamiento la prueba posee una mayor capacidad discriminante.

Variable	Muestras pretratamiento		Muestras postratamiento	
	Valor	95% IC	Valor	95% IC
Sensibilidad	0,70	(0,64 - 0,75)	0,67	(0,58 - 0,74)
Especificidad	0,98	(0,97 - 0,99)	0,92	(0,85 - 0,96)
DOR	130,218	(69,4 - 244,36)	24,69	(11,28 - 54,07)
LR+	39,55	(22,02 - 71,02)	8,87	(4,52 - 17,44)
LR-	0,304	(0,25 - 0,37)	0,36	(0,28 - 0,46)

Tabla 11: Resultados del metanálisis desagregado por subgrupos según antes o después de tratamiento para evaluar la efectividad de la adecuación del tratamiento del análisis de las mutaciones en EGFR en muestras de biopsia líquida en pacientes de NSCLC.

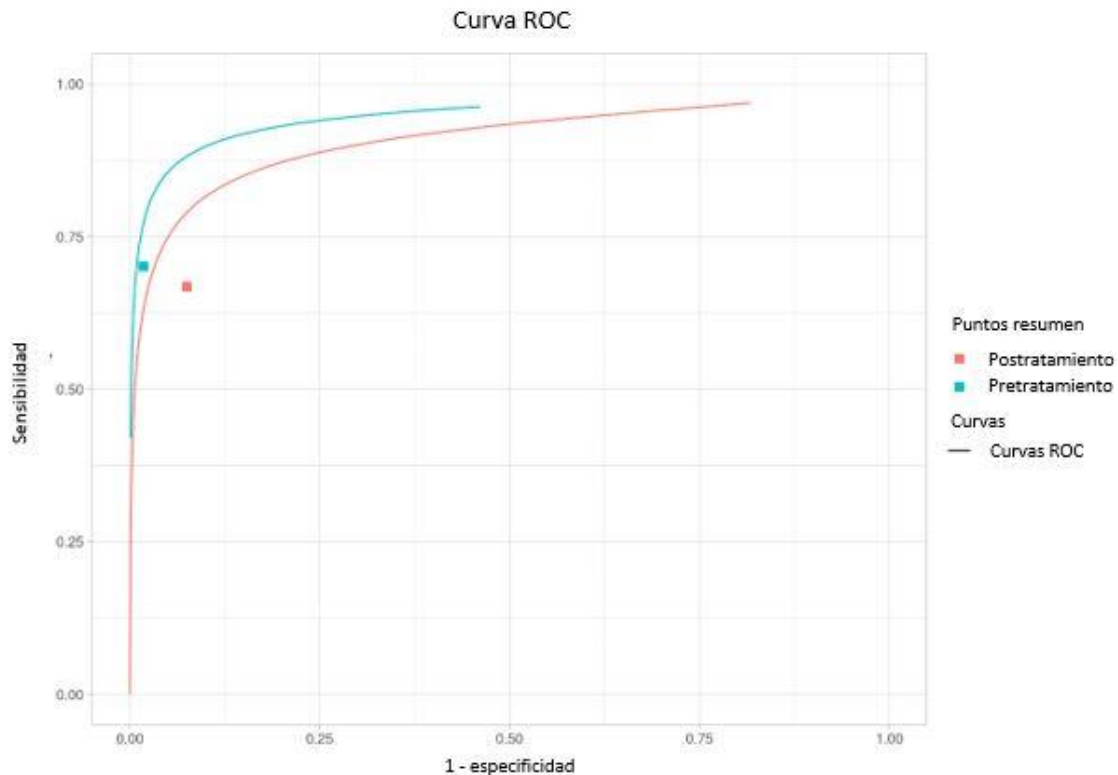


Figura 16: Curva ROC resultado del metanálisis desagregado por subgrupos antes o después de tratamiento para evaluar la efectividad de la adecuación del tratamiento del análisis de las mutaciones en EGFR en muestras de biopsia líquida en pacientes de NSCLC.

4.4 Expresión diferencial de ARNlnc entre tejido tumoral y tejido sano de pacientes con diferentes tipos de cáncer de pulmón y diferentes etnias

En la implementación de la MPP y la interpretación de los resultados de las tecnologías ómicas es imprescindible el correcto uso de la bioinformática. Esta disciplina y su importancia en la MPP ya ha sido recogida previamente en esta tesis dentro del marco de evaluación de la implementación de la MPP para el análisis de los datos ómicos y su interpretación. En esta sección esta importancia queda patente al usarse la bioinformática como herramienta para proponer nuevos biomarcadores en tejido pulmonar tumoral. Con esta finalidad, se realiza el análisis de la expresión diferencial de ARNlnc entre tejido tumoral y tejido sano de pacientes con diferentes tipos de cáncer de pulmón y de diferentes etnias.

El análisis de expresión diferencial de ARNlnc planteado, pretende analizar los diferentes patrones de expresión de los ARNlnc entre la muestra tumoral y la muestra de pulmón sano de un mismo paciente. Este análisis se realizó sobre **4 tipos de muestras**: 14 muestras pareadas de 7 pacientes malagueños de LAC (LAC7), 16 muestras pareadas de

Resultados

8 pacientes malagueños de SqCC (SqCC8), 34 muestras de 17 pacientes asiáticos de SCLC (SCLC17) y 100 muestras de 50 pacientes asiáticos de LAC (LAC50). En la sección 3.2 *Obtención de datos ARNseq para su análisis bioinformático* se explicitan los detalles de todos los tipos de muestras y la metodología seguida es el flujo de trabajo mostrado en el apartado 3.3 *Análisis bioinformático de expresión diferencial de datos de ARNseq*. La expresión se consideró diferencial de ARNlnc entre el tejido sano y el tejido tumoral del mismo paciente si $P < 0,05$ y $-1 > \log FC > 1$.

En la tabla 12 se muestra el número de ARNlnc con expresión diferencial para cada uno de los tipos de muestras. Además, este número de ARNlnc se divide entre los ARNlnc sobreexpresados en el tejido tumoral (ARNlnc sobreexpresados) y los que se encuentran infraexpresados (ARNlnc infraexpresados).

Muestras	ARNlnc con expresión diferencial	ARNlnc sobreexpresados	ARNlnc infraexpresados
LAC7	352	191	161
SqCC8	544	362	182
SCLC17	726	432	294
LAC50	843	592	251

Tabla 12: Número de ARNlnc con expresión diferencial para cada una de las muestras analizadas.

Tras estos resultados, se analizó si los ARNlnc con expresión diferencial eran compartidos por las diferentes muestras. Con esta finalidad se representó un diagrama de Venn, cuyos resultados se plasman en la figura 17a. En esta figura se puede observar que tanto el tipo de cáncer de pulmón como la etnia influyen en los ARNlnc que muestran expresión diferencial, ya que se observa un bajo número de ARNlnc comunes tanto en muestras que comparten etnia, pero no tipo de cáncer (LAC7-SqCC8: caucásica y LAC50-SCLC17: asiática), como en aquellas muestras que comparten tipo de cáncer, pero no etnia (LAC7-LAC50). Por ello se puede concluir que tanto el tipo de cáncer como la etnia son fuentes de variabilidad, y que el estudio de los ARNlnc que presentan expresión diferencial en los 4 tipos de muestras puede ser de interés por ser conservada. Otros estudios ya habían confirmado la variabilidad en los patrones de expresión entre las distintas etnias y tipos de cáncer de pulmón en otras biomoléculas (ARNm y miARN) (Izumi et al. 2020; Adib et al. 2022).

Además, en las figuras 17b y 17c se muestran dos diagramas de Venn de los ARNlnc con expresión diferencial comunes en las 4 muestras, representando la figura 17b los ARNlnc sobreexpresados y la figura 17c los ARNlnc infraexpresados. Se puede observar gracias a estos diagramas que la mayoría de los ARNlnc con expresión diferencial comunes entre las muestras comparten el sentido de desregulación y que el número de ARNlnc comunes sobreexpresados e infraexpresados es muy similar, a pesar de que en los no comunes existe una mayoría de sobreexpresados. Aquellos ARNlnc comunes entre los diferentes tipos de cáncer y etnias y comparten sentido de desregulación son de especial interés porque pueden ser mutaciones esenciales para el inicio o progreso del cáncer de pulmón y se convertirían en posibles dianas con las que tratar distintos tipos de cáncer. En el anexo 10 se detalla el listado de los 56 ARNlnc que muestran expresión diferencial en los 4 tipos de muestras.

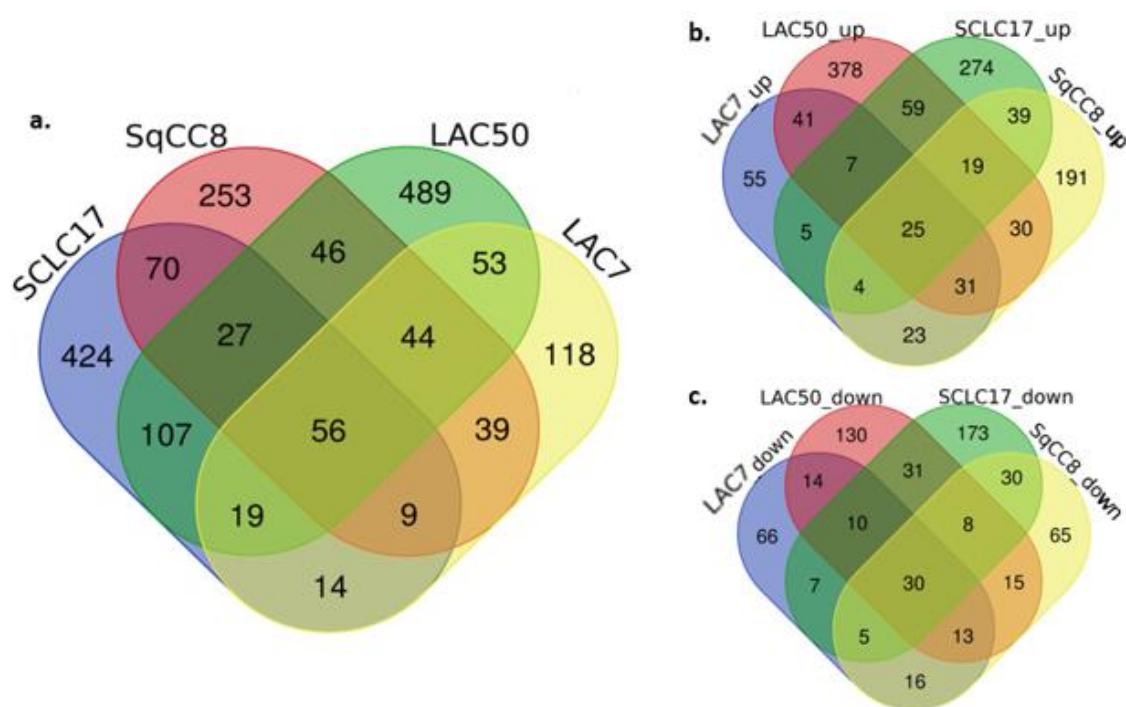


Figura 17: a. Diagrama de Venn representando el número de ARNlnc con expresión diferencial entre el tejido tumoral y el tejido sano en las muestras analizadas. b. Diagrama de Venn representando el número de ARNlnc sobreexpresados diferencialmente en el tejido tumoral de las muestras analizadas. c. Diagrama de Venn representando el número de ARNlnc infraexpresados diferencialmente en el tejido tumoral de las muestras analizadas.

Resultados

4.5 Correlación de la expresión de ARNlnc y ARNm para identificar posibles biomarcadores nuevos para biopsia líquida en los pacientes con cáncer de pulmón.

Tras obtener los datos de expresión diferencial de ARNlnc entre tejido tumoral y tejido sano de los cuatro tipos de pacientes de cáncer de pulmón analizados (pacientes caucásicos con adenocarcinoma, pacientes caucásicos con carcinoma epidermoide, pacientes asiáticos con adenocarcinoma y pacientes asiáticos con cáncer de células pequeñas o microcítico) se procedió a la correlación de los niveles de expresión de ARNlnc y ARNm. Los valores de expresión diferencial de ARNm de los cuatro tipos de muestras analizados en esta tesis han sido publicados en estudios previos (Espinosa Garcia et al. 2022) de nuestro laboratorio. Obtener la correlación ARNlnc-ARNm permite interpretar posibles repercusiones de la expresión diferencial de los ARNlnc y proponer posibles nuevos biomarcadores.

4.5.1 Correlaciones de expresión diferencial entre ARNlnc y ARNm

Gracias al *script* explicado en detalle en el apartado 3.6 *Análisis de correlación de expresión de datos de ARNseq*, se evaluó r , el coeficiente de correlación de Pearson, entre los perfiles de expresión de ARNlnc y ARNm. Se incluyeron en estos análisis los ARNlnc y los ARNm que mostraron expresión diferencial entre el tejido sano y el tejido tumoral. Se seleccionaron los pares ARNlnc-ARNm definidos por $|r| \geq 0,8$ y por $P < 0,05$. El número de correlaciones ARNlnc-ARNm obtenidas según estos criterios (tanto en el mismo sentido como en sentidos opuestos) en cada una de las muestras se recogen en la Tabla 13.

Muestras	Número de ARNlnc con expresión diferencial	Número de ARNm con expresión diferencial	Número de correlaciones ARNlnc-ARNm
LAC7	352	2595	18539
SqCC8	544	3098	14319
SCLC17	726	4705	331
LAC50	843	3344	149

Tabla 13: Número de correlaciones significativas ARNlnc-ARNm en los diferentes tipos de muestras.

Las correlaciones ARNlnc-ARNm obtenidas fueron comparadas entre los 4 tipos de muestras, para poder obtener posibles correlaciones comunes o dependientes del tipo de cáncer de pulmón o de la etnia de los pacientes (LAC7-SqCC8: caucásica y LAC50-

SCLC17: asiática). Esta comparación se muestra en el diagrama de Venn de la figura 18. Las correlaciones compartidas entre los diferentes tipos de muestras fueron escasas y ninguna de ellas era compartida entre los 4 tipos. Esto no resulta sorprendente porque numerosos estudios (incluidos TFG y TFM de nuestro laboratorio) han publicado ya las grandes diferencias en los patrones de expresión entre estos diferentes tipos de cáncer y entre las etnias caucásica y asiática (Izumi et al. 2020; Adib et al. 2022). Es muy destacable que, aunque no pertenezcan al mismo tipo de cáncer pulmonar, las muestras de pacientes caucásicos (LAC7 y SqCC8) son las que comparten un mayor número de correlaciones, 810 correlaciones, siendo 791 de ellas no compartidas con las muestras de etnia asiática (LAC50 y SCLC17).

Tras confirmar a nivel de correlaciones ARNlnc-ARNm la escasa similitud entre los diferentes tipos de cáncer de pulmón y las etnias y dado que los que presentaban un mayor número de correlaciones comunes eran los dos tipos de muestras de pacientes caucásicos, se decidió centrar el estudio en las muestras de los pacientes caucásicos, aunque esto significara una gran reducción del número de pacientes estudiados.

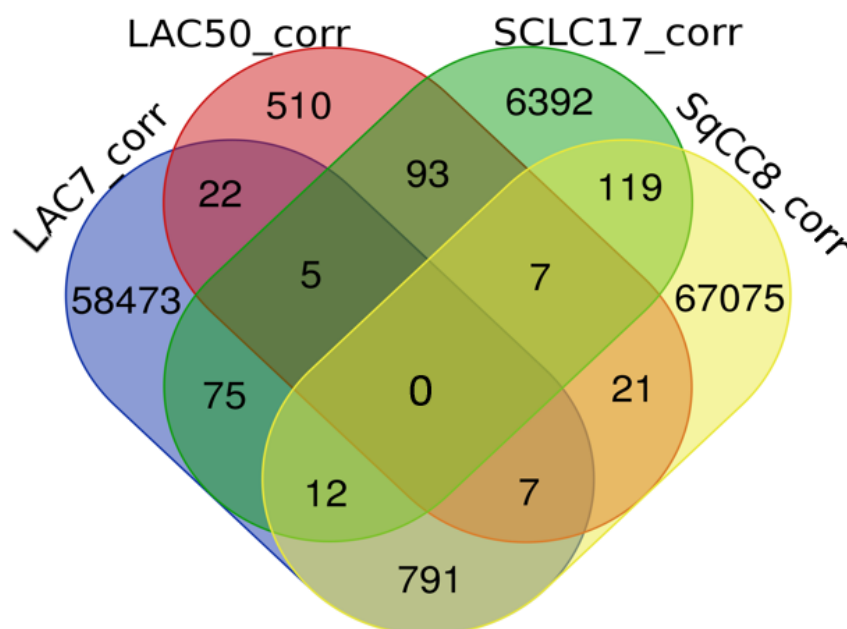


Figura 18: Diagrama de Venn representando las correlaciones ARNlnc-ARNm con expresión diferencial en las muestras analizadas.

Resultados

4.5.2 Interpretación funcional de las correlaciones ARNInc-ARNm en los pacientes caucásicos

Con la finalidad de proponer nuevos biomarcadores para el adenocarcinoma de pulmón y el carcinoma epidermoide pulmonar en los pacientes caucásicos, se interpretaron funcionalmente las correlaciones ARNInc-ARNm obtenidas en las muestras LAC7 y SqCC8 para extraer las parejas que no solo tengan buena correlación sino también una plausible motivación funcional.

Interpretación de las correlaciones ARNInc-ARNm en las muestras LAC7

En primer lugar, las 18.539 correlaciones ARNInc-ARNm obtenidas matemáticamente por correlación de expresión diferencial de ARNInc y ARNm con el *script* previamente explicado, fueron filtradas por su presencia en las bases de datos de correlación ARN-ARN: RNAInter (Kang et al. 2022) y RISE (Gong et al. 2018). Este paso permitió validar 13 de las correlaciones matemáticas obtenidas.

ARNInc	Signo del FC del ARNInc	ARNm	Signo del FC del ARNm	Correlación	<i>P</i>
CRNDE	++	KLF10	++	0,95	0,003
FENDRR	--	BDH2	--	0,95	0,002
FENDRR	--	RNF144B	--	0,97	0,001
FENDRR	--	GRASP	--	0,98	0,0007
FENDRR	--	SLC40A1	--	0,98	0,0003
TFAP2A-AS2	++	INPP1	++	0,96	0,002
CRNDE	++	TSKU	--	-0,99	0,0002
CRNDE	++	TFAP2A	--	-0,96	0,002
CRNDE	++	JAM2	--	-0,96	0,002
DRAIC	++	CASS4	--	-0,95	0,003
FENDRR	--	ABCB1	++	-0,95	0,003
TFAP2A-AS2	++	EMP2	--	-0,98	0,0004
UCA1	++	CDKN1A	--	-0,96	0,003

Tabla 14: Valor de las correlaciones matemáticas ARNInc-ARNm en las muestras LAC7 validadas por su presencia en las bases de datos.

Para estudiar la posible funcionalidad celular dependiente de los ARNm implicados en las correlaciones, en primer lugar, se extrajeron de la versión 11.5 de la base de datos String las interacciones de las proteínas que codifican estos ARNm en función de vecindad a nivel de genoma, por fusión de genes, por coexpresión o por aparición de la

interacción en bases de datos. Además, añade proteínas a la red en los casos en los que la adición de una nueva proteína permita representar la correlación de dos de las proteínas incluidas previamente. En este caso, se obtuvo la red de interacciones de la figura 19, y gracias a las anotaciones funcionales que se pueden consultar desde String, entre otras KEGG Pathways y WikiPathways, se pudo identificar que 2 de los ARNm en estudio (*CDKN1A* y *ABCB1*) codifican proteínas involucradas en rutas relacionadas con la enfermedad en estudio. En concreto, se marcaron en rojo las proteínas involucradas en rutas relacionadas con cáncer de pulmón microcítico (según la base de datos KEGG Pathways), en azul las proteínas involucradas en rutas relacionadas con cáncer de pulmón no microcítico (según la base de datos KEGG Pathways) y en verde las proteínas relacionadas con mecanismos de resistencia terapéutica mediados por ARNlnc (según la base de datos WikiPathways).

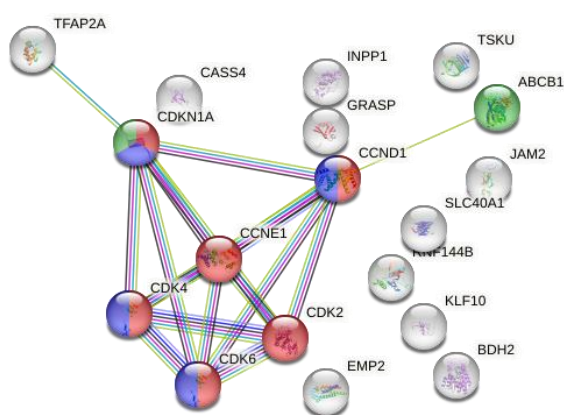


Figura 19: Representación de la red de interacciones proteína-proteína de los ARNm implicados en las correlaciones obtenida en el portal de String.

En segundo lugar, se estudiaron por separado las rutas y procesos en los que estaban implicados todos los ARNm y ARNlnc de las correlaciones, buscándolos en la bibliografía. Además, en el caso de los ARNlnc se buscó información de ellos en la base de datos exorbase para determinar si según esta base de datos estos ARNlnc son detectables en BL.

ARNm	Función en células no tumorales	Implicaciones en células cancerosas	Tipo de cáncer relacionado
KLF10	La proteína codificada es un represor transcripcional que actúa como efector de la	La actividad de la proteína puede inhibir el crecimiento de cánceres.	Cáncer de páncreas y pulmón

Resultados

	señalización del factor de crecimiento beta.		
BDH2	Codifica una proteína que interviene en la diferenciación de las células epiteliales y en la betaoxidación de los ácidos grasos.	En células tumorales esta proteína puede desencadenar la muerte celular inducida por ROS y la autofagia mediante la promoción de la ubiquitinación de <i>Nrf2</i> .	Cáncer gástrico
RNF144B	Induce la apoptosis a través de un mecanismo dependiente de p53/TP53 pero independiente de la caspasa.	Esta proteína se relaciona con mecanismos de proliferación de células cancerosas en ausencia de estrógeno.	Cáncer de endometrio y mama
GRASP	Codifica proteínas de apilamiento del complejo de Golgi.	Esta proteína se relaciona con la capacidad de migración e invasión de las células tumorales.	Carcinoma hepatocelular
SLC40A1	Codifica una proteína que transporta hierro desde las células intestinales, esplénicas y hepáticas a la sangre para proporcionárselo a otros tejidos.	Su infraexpresión podría aumentar la capacidad de proliferación, migración e invasión de las células tumorales	Carcinoma hepatocelular
INPP1	INPP1 codifica la enzima inositol polifosfato-1-fosfatasa, una de las enzimas implicadas en las vías de señalización del fosfatidilinositol.	Su sobreexpresión contribuye a la capacidad de crecimiento, migración e invasión de las células tumorales.	Melanoma y cáncer de ovario
TSKU	Codifica la proteína de matriz extracelular Tsukushi o tsukushin.	En estudio	NSCLC
TFAP2A	Codifica el factor de transcripción AP-2 α que interviene en el desarrollo antes del nacimiento.	Involucrada en vías de señalización para posibilitar la metástasis.	Cáncer de pulmón
JAM2	La proteína codificada por este gen es una proteína de membrana de tipo I que se localiza en las uniones estrechas de las células epiteliales y endoteliales.	Su sobreexpresión se relaciona con una reducción de la capacidad de invasión y migración de las células tumorales.	Cáncer de mama
CASS4	Codifica una proteína que permite la actividad de	La sobreexpresión de CASS4 promueve la	NSCLC

	unión de la proteína tirosina quinasa. Participa en la regulación positiva de varios procesos, como la señalización de la proteína quinasa B y la actividad de la proteína tirosina quinasa.	invasión en el NSCLC mediante la activación de la vía de señalización AKT y la inhibición de la expresión de E-cadherina.	
ABCB1	El gen ABCB1 codifica un transportador transmembrana que bombea una amplia gama de compuestos xenobióticos de las células.	Se relaciona con resistencia a fármacos en las neoplasias. El ARNInc <i>FENDRR</i> sensibiliza la resistencia a la doxorrubicina de las células de osteosarcoma mediante la regulación a la baja de <i>ABCB1</i> y <i>ABCC1</i> .	Osteosarcoma
EMP2	La proteína codificada regula la composición de la membrana celular.	<i>EMP2</i> se identificó como gen supresor de tumores en el carcinoma urotelial del tracto urinario.	Carcinoma urotelial
CDKN1A	Es un inhibidor de la quinasa dependiente de ciclina capaz de inhibir todos los complejos ciclina/CDK, aunque se asocia principalmente con la inhibición de CDK2.	La infraexpresión de <i>CDKN1A</i> se asocia con mal pronóstico de adenocarcinoma gástrico y NSCLC y SCLC, y con mecanismos de resistencia terapéutica mediados por ARNInc.	Adenocarcinoma gástrico, NSCLC y SCLC

Tabla 15: Información funcional de los ARNm implicados en las correlaciones ARNInc-ARNm en las muestras LAC7.

ARNInc	Implicaciones en células cancerosas	Tipo de cáncer relacionado	Evidencia detección en BL
CRNDE	Asociado al crecimiento y la invasión de células de glioma a través de la vía de señalización mTOR.	Cáncer colorrectal y carcinoma de células escamosas de lengua	Sí
FENDRR	Como supresor tumoral, <i>FENDRR</i> regula la expresión de genes diana y desempeña funciones cruciales tanto en la tumorigénesis como en la progresión del cáncer. Además, <i>FENDRR</i> ha sido identificado como un biomarcador potencial para terapias dirigidas. La sobreexpresión terapéutica de	Diversos cánceres gástricos, urológicos y en vías respiratorias	Sí

Resultados

	<i>FENDRR</i> puede frenar el crecimiento tumoral y la metástasis a través de vías moleculares específicas.		
TFAP2A-AS2	Implicado en la regulación de la expresión génica del desarrollo funcionando como represor en células no pluripotentes y como activador en células pluripotentes. Regula la expresión específica de muchos genes frecuentemente relacionados con la tumorigénesis.	LAC y carcinoma naseofaríngeo	NA
DRAIC	Inhibe la proliferación y metástasis de células de cancerosas a través de la interferencia con la desubiquitinación de <i>NFRKB</i> mediada por <i>UCHL5</i> .	Cáncer gástrico	Sí
UCA1	Desempeña un papel regulador en la proliferación celular. Este transcrito regula varios genes, entre ellos <i>CREB</i> . Además, induce resistencia adquirida a Gefitinib por silenciamiento epigenético de <i>CDKN1A</i> .	Cáncer de vejiga y NSCLC	Sí

Tabla 16: Información funcional de los ARNlnc implicados en las correlaciones ARNlnc-ARNm en las muestras LAC7.

En tercer lugar, se buscó información adicional de las interacciones y correlaciones registradas en las bases de datos de cada uno de los emparejamientos ARNlnc-ARNm, y si existía, su posible relación con las neoplasias. El único emparejamiento para el que existía información previa en los recursos consultados fue *CDKN1A* y *UCA1*, relacionado además con la adquisición de resistencias a uno de los tratamientos más habituales en el cáncer de pulmón: el gefitinib.

A la vista de los resultados expuestos, respecto a los ARNm con expresión diferencial correlacionada con la expresión de ARNlnc, se concluye que:

- 5 de los 13 ARNm en estudio se han relacionado con cáncer de pulmón (*KLF10*, *TSKU*, *TFAP2A*, *CASS4* y *CDKN1A*). 1 de ellos (*CDKN1A*) se relaciona con mecanismos de resistencia terapéutica mediados por ARNlnc, en concreto se relaciona con el ARNlnc *UCA1*, coincidiendo con uno de los resultados obtenidos en nuestro estudio.

- 6 de los 13 ARNm en estudio se han relacionado con otros tipos de cánceres (*BDH2*, *RNF144B*, *INPP1*, *JAM2*, *ABCB1* y *EMP2*). Se proponen estos ARNm como **posibles biomarcadores de adenocarcinoma pulmonar en poblaciones caucásicas**. 1 de ellos (*ABCB1*) se relaciona con mecanismos de resistencia terapéutica mediados por ARNlnc, en concreto se relaciona con el ARNlnc *FENDERR1*, coincidiendo con uno de los resultados obtenidos en nuestro estudio.

Respecto a los ARNlnc con expresión diferencial correlacionada con la expresión de ARNm:

- 3 de los 5 ARNlnc en estudio (*UCA1*, *FENDRR* y *TFAP2A-AS2*) se han relacionado previamente con cáncer de pulmón.
- 2 de los 5 ARNlnc en estudio se proponen como **posibles nuevos biomarcadores de adenocarcinoma pulmonar en poblaciones caucásicas** (*CRNDE* y *DRAIC*). Ya se había descrito en la bibliografía la relación de estos ARNlnc con otras neoplasias, pero no con cáncer de pulmón
- Especialmente valiosos son aquellos ARNlnc en los que es posible su análisis en biopsia líquida (*CRNDE*, *UCA1*, *FENDRR* y *DRAIC*).

Adicionalmente, se recomienda el estudio de las nuevas correlaciones ARNlnc-ARNm que involucran a ARNlnc ya recogidos en bases de datos como implicados en las neoplasias pulmonares, ya que el único emparejamiento para el que existía información previa en los recursos consultados ha sido *CDKN1A* y *UCA1*.

Interpretación de las correlaciones ARNlnc-ARNm en las muestras SqCC8

La metodología empleada para la interpretación de las correlaciones ARNlnc-ARNm en las muestras SqCC8 fue la misma que la empleada en las muestras LAC7. En primer lugar, las 14.319 correlaciones ARNlnc-ARNm obtenidas se filtraron por su presencia en las bases de datos de correlación ARN-ARN RNAInter (Kang et al. 2022) y RISE (Gong et al. 2018). Este paso permitió validar 9 de ellas.

Resultados

ARNInc	Signo FC ARNInc	ARNm	Signo FC ARNm	Correlación	P
FENDRR	++	FOXF1	++	0,90	0,005
FENDRR	++	ARHGAP29	++	0,91	0,004
FENDRR	++	PTPRM	++	0,94	0,001
FENDRR	++	RBMS3	++	0,94	0,001
FOXD2-AS1	--	OXTR	--	0,91	0,005
SENCR	++	H2AFX	++	0,90	0,005
SENCR	++	NOL6	++	0,95	0,0009
TFAP2A-AS2	--	ELMO1	++	-0,94	0,001
TFAP2A-AS2	--	OCIAD2	--	0,92	0,004

Tabla 17: Valor de las correlaciones ARNInc-ARNm en las muestras SqCC8 validadas por su presencia en bases de datos.

En primer lugar, la extracción de las interacciones de la base de datos String no arrojó posibles relaciones entre los ARNm objeto de estudio.

En segundo lugar, se estudiaron por separado las rutas y procesos en los que estaban implicados todos los ARNm y ARNInc de las correlaciones, buscándolos en la bibliografía. Además, en el caso de los ARNInc se buscó información de ellos en la base de datos exoRbase para determinar si según los datos de esta base de datos estos ARNInc son detectables en biopsia líquida.

ARNm	Función en células no tumorales	Implicaciones en células cancerosas	Tipo de cáncer relacionado
FOXF1	Este gen pertenece a la familia de factores de transcripción <i>forkhead</i> . La función específica de este gen aún no se ha determinado; sin embargo, puede desempeñar un papel en la regulación de los genes pulmonares, así como en el desarrollo embrionario.	La activación epigenética de <i>FOXF1</i> confiere propiedades de célula madre cancerosa en NSCLC resistente al cisplatino.	NSCLC
ARHGAP29	Codifica una proteína que media la regulación cíclica de pequeñas proteínas de unión a GTP como RhoA.	La expresión de <i>ARHGAP29</i> se asocia con la proliferación y la invasión celular.	Cánceres de mama y próstata
PTPRM	Codifica proteínas de señalización que regulan una gran variedad de procesos celulares, como el	Una expresión elevada de <i>PTPRM</i> predice un mal pronóstico, favorece el crecimiento tumoral y	Cáncer de cuello uterino

	crecimiento celular, la diferenciación, el ciclo mitótico y la transformación oncogénica.	la metástasis en los ganglios linfáticos.	
RBMS3	Este gen codifica una proteína de unión a ARN implicada en funciones tan diversas como la replicación del ADN, la transcripción génica, la progresión del ciclo celular y la apoptosis.	<i>RBMS3</i> inhibe la proliferación y metástasis.	Cáncer de mama
OXTR	La proteína codificada por este gen pertenece a la familia de los receptores acoplados a proteínas G y actúa como receptor de la oxitocina. El sistema oxitocina-receptor de oxitocina desempeña un papel importante en el útero durante el parto.	<i>OXTR</i> está directamente implicado en los mecanismos patológicos de múltiples cánceres.	Cánceres de mama, próstata y ovario
H2AFX	Este gen codifica una histona independiente de la replicación que pertenece a la familia de las histonas H2A y genera dos transcritos.	<i>H2AFX</i> se ha descrito como un posible biomarcador pronóstico.	Carcinoma hepatocelular
NOL6	Este gen codifica una proteína asociada al ARN nucleolar que está altamente conservada entre especies. Otros estudios sugieren que la proteína está asociada a la biogénesis de los ribosomas a través de su interacción con los transcritos primarios de pre-ARNr.	<i>NOL6</i> se ha descrito como un oncogén y diana de <i>miR-590-3p</i> .	Cáncer de próstata
ELMO1	Codifica un factor de intercambio de nucleótidos de guanina.	Su sobreexpresión se ve relacionada con una mayor capacidad invasiva.	Cáncer de esófago
OCIAD2	Participa en la endocitosis, la homeostasis de las células madre hematopoyéticas y la regulación positiva de la vía de señalización del receptor a través de JAK-STAT.	Su nivel de expresión aumenta durante el curso de la progresión tumoral.	Cánceres de pulmón, ovario, tiroides y hematológicos

Resultados

Tabla 18: Información funcional de los ARNm implicados en las correlaciones ARNInc-ARNm en las muestras SqCC8.

ARNInc	Implicaciones en células cancerosas	Tipo de cáncer relacionado	Evidencia detección en BL
FENDRR	Como supresor tumoral, <i>FENDRR</i> regula la expresión de genes diana y desempeña funciones cruciales tanto en la tumorigénesis como en la progresión del cáncer. Además, <i>FENDRR</i> ha sido identificado como un biomarcador potencial para terapias dirigidas. La sobreexpresión terapéutica de <i>FENDRR</i> puede frenar el crecimiento tumoral y la metástasis a través de vías moleculares específicas.	Diversos cánceres gástricos, urológicos y en vías respiratorias	Sí
FOXD2-AS1	Se localiza en el cromosoma 1p33 y contiene 2527 nucleótidos, es un nuevo ARN no codificante identificado como oncogén en varios tipos de cánceres a través de su interacción con <i>miR-204-3p</i> .	Cáncer de vejiga, NSCLC, cáncer gástrico y carcinoma esofágico de células escamosas	Sí
SENCR	Su sobreexpresión se relaciona con la resistencia al cisplatino y el crecimiento tumoral.	NSCLC	Sí
TFAP2A-AS2	Implicado en la regulación de la expresión génica del desarrollo funcionando como represor en células no pluripotentes y como activador en células pluripotentes. Regula la expresión específica de muchos genes frecuentemente relacionados con la tumorigénesis.	LAC y carcinoma naseofaríngeo	NA

Tabla 19: Información funcional de los ARNInc implicados en las correlaciones ARNInc-ARNm en las muestras SqCC8.

En tercer lugar, se buscó información adicional de las interacciones y correlaciones registradas en las bases de datos de cada uno de los emparejamientos ARNInc-ARNm, y si existía, su posible relación con el cáncer. Solamente se localizó información para la

correlación *FOXF1* y *FENDRR*, que además se asocia en los casos de cáncer de pulmón no microcítico con resistencia al tratamiento con cisplatino.

A la vista de los resultados expuestos, respecto a los ARNm con expresión diferencial correlacionada con la expresión de ARNlnc, se concluye que:

- 1 de los 9 ARNm en estudio se ha relacionado con cáncer de pulmón (*FOXF1*) y se asocia a casos de resistencia al tratamiento con cisplatino en los casos de cáncer de pulmón. Además, se relaciona con el ARNlnc *FENDRR*, coincidiendo con uno de los resultados obtenidos en nuestro estudio.
- 1 de los 9 ARNm en estudio se ha relacionado con cáncer de pulmón (*OCIAD2*) aunque no se ha podido confirmar la correlación de este ARNm con *TFAP2A-AS2* en los recursos consultados.
- 7 de los 9 ARNm en estudio se han relacionado con otros tipos de cánceres (*ARHGAP29*, *PTPRM*, *RBMS3*, *OXTR*, *H2AFX*, *NOL6* y *ELMO1*). Se proponen estos ARNm como **posibles biomarcadores de carcinoma epidermoide pulmonar en poblaciones caucásicas**.

Respecto a los ARNlnc con expresión diferencial correlacionada con la expresión de ARNm:

- Los 4 ARNlnc en estudio (*FENDRR*, *FOXD2-AS1*, *TFAP2A-AS* y *SENCR*) se asocian con diferentes tipos de cáncer de pulmón. Este estudio añade evidencia para proponer estos ARNlnc como **posibles nuevos biomarcadores de carcinoma epidermoide pulmonar en poblaciones caucásicas**

Adicionalmente, se recomienda el estudio de las nuevas correlaciones ARNlnc-ARNm que involucren a estos ARNlnc ya recogidos en bases de datos como implicados en las neoplasias pulmonares.



UNIVERSIDAD
DE MÁLAGA

Discusión



UNIVERSIDAD
DE MÁLAGA

5. Discusión

El trabajo presentado en esta tesis doctoral aborda la MPP en distintas fases de su desarrollo, desde la investigación hasta la implementación en la rutina clínica, y desde distintos enfoques. Un primer enfoque holístico de la implementación y un segundo enfoque centrado en una enfermedad y aplicación concreta: la eficacia de la biopsia líquida y los análisis de expresión de ARNlc y ARNm para la mejora del diagnóstico y del tratamiento en el cáncer de pulmón.

La MPP supone un reto frente a la medicina clásica, desde las fases de investigación hasta su desarrollo, evaluación e implementación (Fosse et al. 2023; 2022; Papanicolas, Rajan y Karanikolos 2022). En primer lugar, desde la investigación se requieren novedosas tecnologías y especialistas en ellas, como las tecnologías de secuenciación y la bioinformática. En segundo lugar, los estudios y ensayos clínicos necesarios para el desarrollo y validación de las aplicaciones investigadas puede suponer un reto para los investigadores por el propio diseño que deben cumplir los estudios. En tercer lugar, las metodologías de evaluación de tecnologías sanitarias deben adaptarse a las particularidades de las tecnologías usadas en la MPP suponiendo un reto metodológico para los evaluadores. En cuarto lugar, para la implementación de la MPP en la rutina clínica, lo que supondrá el acceso del paciente a sus beneficios, debe de estudiarse en el sistema sanitario en su conjunto dados los numerosos puntos del sistema en los que la aplicación de la MPP podría cambiar el abordaje del paciente (cribado, pronóstico, diagnóstico o adecuación del tratamiento) y las diferentes enfermedades en las que podría aplicarse. Estos aspectos son los que se han contemplado en el marco para la evaluación de la implementación de la MPP a nivel nacional resultado de una RS.

5.1 Evaluación de la implementación de la MPP a nivel nacional

Dados los numerosos aspectos diferenciales de la MPP frente a la medicina clásica expuestos anteriormente, es necesario un marco específico que contemple todas estas particularidades y englobe todos los puntos del sistema sanitario que se verían afectados por esta implementación. Existen otros marcos desarrollados para la evaluación específica de la implementación de la MPP en un país determinado (Agarwal

et al. 2021), sin que puedan ser extrapolables a otros, y otras que contemplan solamente algunos de los dominios, como los aspectos organizacionales (Doyle et al. 2018) o la investigación e innovación (Lee y Kim 2022). Sin embargo, el marco para la evaluación de la implementación de la MPP, presentado en la figura 11 y la tabla 5, de esta tesis doctoral supone el primer marco de este tipo especializado en MPP. Además, tanto la obtención de los dominios y elementos que la constituyen como el sistema de puntuación utilizado se basan en los resultados de una RS de la bibliografía.

Los 11 dominios del marco de evaluación, que contienen los elementos evaluables, engloban aspectos políticos, organizacionales, científicos, clínicos, educativos y financieros, poniendo de relieve la necesidad de colaboración entre estos actores y entre estos actores y los pacientes para la correcta implementación de la MPP.

Este marco, que permite la evaluación cualitativa y cuantitativa, posibilita: la detección de debilidades dentro de la implementación de cada país, la guía para el desarrollo de políticas y planes estratégicos de implementación de MPP, y la evaluación de múltiples países para detectar aquellos con un mayor grado de implementación cuyas estrategias puedan servir como modelo.

Junto al diseño del marco, se ha procedido a su validación evaluando la implementación de la MPP en una serie de países de diferentes continentes con diferentes niveles de desarrollo económico según la clasificación del *World Bank*. Los valores obtenidos, recogidos en la tabla 6, muestran que el marco cubre todo el rango de puntuaciones y que hay una relación directamente proporcional entre el desarrollo económico del país y la puntuación del grado de implementación de la MPP obtenido por el marco generado. Estos resultados son esperables, pues un mayor grado de desarrollo económico permite una mayor inversión y respaldo para las iniciativas y recursos que permiten sistemas sanitarios adaptables a las necesidades que requiere la implementación de la MPP (Vicente, Ballensiefen y Jönsson 2020; Papanicolas, Rajan y Karanikolos 2022).

5.2 Biopsia líquida para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón

Para analizar las aplicaciones específicas de la MPP, se ha escogido el cáncer de pulmón, aunque las 2 aplicaciones objeto de análisis se encuentran en diferentes estados de desarrollo. La primera aplicación es la biopsia líquida para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón que se encuentra ya en fase de evaluación e implementación. La segunda es el análisis de la expresión de ARNlnc y su correlación con ARNm en el tejido tumoral del cáncer de pulmón que se encuentra en fase de investigación. Al incluir 2 aplicaciones de la MPP en distintos estadios de desarrollo de esa tecnología, junto a un primer enfoque holístico del desarrollo e implementación de la MPP en su conjunto, se pretende cubrir el análisis de todas las etapas de desarrollo de las tecnologías de MPP y sus repercusiones para los sistemas sanitarios.

Centrándonos en la aplicación de la BL (biopsia líquida) para el diagnóstico y la adecuación del tratamiento del cáncer de pulmón, esta técnica es de gran interés porque puede permitir el diagnóstico precoz, la adecuación del tratamiento o el pronóstico del cáncer de pulmón, sin necesidad de tomar una biopsia de tejido del paciente. Esto puede ser especialmente beneficioso para los pacientes de los que no se pueda tomar una biopsia del tejido tumoral, para los casos en los que exista heterogeneidad celular a nivel tumoral o para la disminución de reintervenciones para los casos en los que se sospeche que es necesario una modificación del tratamiento (Bonanno et al. 2022).

En nuestro estudio, se pretendió a través de una RS de la bibliografía y un metanálisis conocer el estado de desarrollo y las aplicaciones de la BL para el cáncer de pulmón y las diferentes alteraciones ómicas objeto de análisis. Esta RS permitió conocer que actualmente están siendo objeto de estudios clínicos una gran cantidad de alteraciones ómicas a nivel de BL tanto para el diagnóstico como para la adecuación del tratamiento de cáncer de pulmón. Estas alteraciones ómicas comprenden desde el análisis de mutaciones puntuales de genes hasta alteraciones en patrones de expresión de diferentes tipos de ARN o la detección de células tumorales circulantes. El análisis de estas alteraciones en BL aporta valores interesantes de concordancia entre BL y biopsia tisular, especialmente en cuanto a su especificidad, recogidos en la tabla 8. Si bien, los

Discusión

estudios son difícilmente comparables debido a la diversidad de alteraciones analizadas. Este hecho es esperable al estar esta técnica aún en fase de evaluación y existir numerosos biomarcadores que pueden ser prometedores.

Dados unos criterios preestablecidos, el único biomarcador en BL que pudo ser objeto de metanálisis para evaluar la concordancia entre la biopsia tisular y la BL fue: mutaciones en el gen *EGFR* en ADN libre circulante en pacientes con NSCLC. Como cabía esperar este resultado ha seleccionado un gen en el que se sabe que recaen muchas mutaciones asociadas al NSCLC, como las deleciones en el exón 19 (en el codón 747, 748, 749 o 750) o las mutaciones puntuales en el exón 21 (especialmente en el codón 858), y tratamientos aplicables cuando aparecen estas mutaciones, como gefitinib, erlotinib, afatinib u osimertinib (Yoneda et al. 2019). Además, se ha estudiado que el estado mutacional de este gen puede variar según el estadio tumoral o el tratamiento aplicado y puede ser un factor pronóstico de la evolución de la enfermedad (Zhao, Wang y He 2021). Por tanto, parece una buena estrategia para seleccionar marcadores óptimos al menos en el cáncer de pulmón, por lo que probablemente se pueda extender con éxito a otras enfermedades con base genética.

El metanálisis para evaluar la concordancia entre la BL y la biopsia tisular en mutaciones en el gen *EGFR* en ADN libre circulante en pacientes con NSCLC, se ejecutó con los resultados de los estudios clínicos que incluían las RS y sus resultados se plasman en la tabla 9. Al evaluar estos estudios de manera conjunta se obtuvo una gran especificidad de la BL al compararla con la biopsia tisular, sin embargo, el valor de la sensibilidad fue de 0,69 por lo que se debería ser cauteloso con los resultados negativos por la posibilidad de obtener FN. Por otra parte, se desarrollaron dos metanálisis desagregados. El primero de ellos, cuyos resultados se recogen en la tabla 10, según el tipo de muestra biológica analizada para la BL (suero o plasma) para el cual se obtuvo que no había diferencias significativas entre ambos grupos, como se puede observar en la figura 14. En el segundo de los metanálisis desagregados, cuyos resultados se plasman en la tabla 11, los subgrupos se establecieron según si los pacientes habían recibido o no tratamiento (se incluyó como tratamiento la quimioterapia y fármacos inhibidores de la TKI), en este caso sí se obtuvo una diferencia significativa entre ambos subgrupos mostrando una mayor concordancia entre la biopsia de tejido y la BL aquellos pacientes

que no habían recibido ninguno de estos tratamientos, como se puede observar en la figura 15. Hasta el momento no se ha identificado ningún estudio previo que realizara este metanálisis desagregado y obtuviera estos resultados, por lo que puede ser de gran relevancia para saber en qué momento es más sensible y específica la BL, y puede ser una alternativa fiable frente a la biopsia de tejido.

5.3 Análisis bioinformático de expresión diferencial de ARNlnc y búsqueda de nuevos biomarcadores para cáncer de pulmón

El análisis bioinformático de la expresión de ARNlnc y su correlación con ARNm en el tejido tumoral del cáncer de pulmón ha sido también estudiado en esta tesis doctoral. Esta aplicación de la MPP, se encuentra aún en fase de investigación y se centra en uno de los dominios del marco de evaluación, el análisis de los datos ómicos a través de la bioinformática. Si bien este dominio de análisis, almacenamiento e integración de datos ómicos es uno de los más recientemente desarrollados, es imprescindible para la MPP, tanto en las fases de implementación como en las de desarrollo e investigación (como es el caso de su uso en esta tesis).

El análisis bioinformático ha confirmado las diferencias a nivel de patrones de expresión de ARNlnc entre los diferentes cánceres de pulmón y entre las diferentes etnias en estudio (asiática y caucásica). Estas diferencias se pueden observar en la figura 17 que recoge los ARNlnc con expresión diferencial en el tejido tumoral para las diferentes poblaciones en estudio, existiendo un bajo número de ARNlnc con expresión diferencial comunes a los 4 tipos (LAC7, SqCC8, LAC50 y SCLC17). Estas diferencias ya habían sido descritas en otras publicaciones para ARNm y miARN, y acentúan la necesidad de una MPP en estas enfermedades para su correcto diagnóstico, tratamiento y pronóstico integrando todas las particularidades del paciente y de su enfermedad.

Dadas estas diferencias entre las 4 poblaciones en estudio, se decidió centrar el análisis bioinformático posterior en las poblaciones caucásicas de la provincia de Málaga (LAC7 y SqCC8). Lo más notable de este análisis bioinformático a pequeña escala (muchos dirían con razón que 7 u 8 pacientes no son una población significativa) es que se han obtenido resultados validados por otras publicaciones y bases de datos, como las correlaciones de expresión entre los ARNm *CDKN1A* y *FOXF1* con los ARNlnc *UCA1* y

Discusión

FENDRR, respectivamente. Pero también se han obtenido resultados novedosos, como ARNlnc cuya expresión diferencial se correlaciona con la expresión de ARNm en las enfermedades en estudio. Estas correlaciones obtenidas matemáticamente han sido validadas en las bases de datos RISE y RNAInter, pero es la primera vez que han sido descritas para adenocarcinoma o carcinoma escamoso de pulmón en población caucásica, según la bibliografía encontrada. Por eso se proponen nuevos biomarcadores para estos cánceres de pulmón, tanto ARNm como ARNlnc, y son especialmente interesantes los ARNlnc cuya presencia en BL ha sido recogida en la base de datos ExoRbase. En concreto, como biomarcadores en BL el análisis de los niveles de expresión de ARNlnc para el LAC: *CRNDE*, *FENDRR*, *DRAIC* y *UCA1*, y para el SqCC: *FENDRR*, *FOXD2-AS1* y *SENCR*. Está claro que merece la pena hacer prospecciones de marcadores en poblaciones pequeñas (con el consiguiente ahorro de costes, tiempo y molestias a los pacientes) que se puedan validar bioinformáticamente a gran escala antes de proponerlas como posibles biomarcadores en futuros ensayos clínicos. Además, el hecho de que *FENDRR* aparezca como candidato para los dos tipos de cáncer y que también haya sido identificado en otros tipos de cáncer (tablas 16 y 18) sugiere que sería una diana preferente de tratamiento para varios cánceres diferentes, incluidos los dos de pulmón considerados en este trabajo.

Conclusiones



UNIVERSIDAD
DE MÁLAGA

6. Conclusiones

- Gracias al marco de evaluación cualitativo y cuantitativo de los sistemas sanitarios desarrollado en este trabajo hemos demostrado que existe una correlación directa entre el contexto económico y el nivel de desarrollo de los países respecto al nivel de implementación de la MPP a nivel nacional.
- Con la misma estrategia del marco de evaluación hemos determinado que la detección de mutaciones en *EGFR* en ADN circulante libre en BL para el tipo de cáncer NSCLC con la finalidad de adecuación del tratamiento es una alternativa muy ventajosa a la biopsia tisular, tanto si el análisis se realiza en suero como en plasma del paciente. Estos resultados son más fiables, en términos de sensibilidad y especificidad, si el paciente no ha recibido previamente tratamiento de quimioterapia ni inhibidores de la tirosina quinasa. Aunque los miARN son otra fuente esperanzadora para lo mismo, los resultados todavía no son suficientemente concluyente para decantarnos por alguno en concreto.
- Los estudios bioinformáticos realizados en grupos reducidos de pacientes confirman la expresión diferencial de ARNInc entre tejido sano y tejido tumoral en los distintos tipos de cáncer de pulmón, aunque la etnia también afecta mucho al resultado. Por lo tanto, confirmamos que la prospección de marcadores en grupos reducidos de pacientes puede ser una manera rápida y barata de obtener nuevos biomarcadores para BL al menos de las poblaciones locales.
- Proponemos los siguientes ARNInc como posibles biomarcadores en BL para población caucásica con sospecha o pacientes de cáncer de pulmón del subtipo LAC: *CRNDE*, *FENDRR*, *DRAIC* y *UCA1*. En los casos de sospecha o pacientes de cáncer de pulmón del subtipo SqCC, proponemos: *FENDRR*, *FOXD2-AS1* y *SENCR*.



UNIVERSIDAD
DE MÁLAGA

Referencias



UNIVERSIDAD
DE MÁLAGA

Referencias

- ADIB, E., NASSAR, A.H., ABOU ALAIWI, S., GROHA, S., AKL, E.W., SHOLL, L.M., MICHAEL, K.S., AWAD, M.M., JÄNNE, P.A., GUSEV, A. y KWIATKOWSKI, D.J., 2022. Variation in targetable genomic alterations in non-small cell lung cancer by genetic ancestry, sex, smoking history, and histology. *Genome Medicine* [en línea], vol. 14, no. 1, pp. 39. ISSN 1756-994X. DOI 10.1186/s13073-022-01041-x. Disponible en: <https://doi.org/10.1186/s13073-022-01041-x>.
- AGARWAL, A., PRITCHARD, D., GULLETT, L., AMANTI, K.G. y GUSTAVSEN, G., 2021a. A quantitative framework for measuring personalized medicine integration into us healthcare delivery organizations. *Journal of Personalized Medicine*, vol. 11, no. 3. ISSN 20754426. DOI 10.3390/jpm11030196.
- AGUILERA-COBOS, L., ROSARIO-LOZANO, M.P., PONCE-POLO, A., BLASCO-AMARO, J.A. y EPSTEIN, D., 2022. Barriers for the evaluation of advanced therapy medicines and their translation to clinical practice: Umbrella review. *Health Policy* [en línea], vol. 126, no. 12, pp. 1248-1255. ISSN 0168-8510. DOI <https://doi.org/10.1016/j.healthpol.2022.10.007>.
- AGUILERA-COBOS, L., BAÑOS ÁLVAREZ, E., ROSARIO-LOZANO, MP. y BLASCO-AMARO, JA., 2022a. Efectividad, eficacia y seguridad de la citología con evaluación in situ en punción aspiración con aguja fina y biopsia con aguja gruesa ecoguiadas. *Ministerio de Sanidad, AETSA Área de Evaluación de Tecnologías Sanitarias* [en línea]. DOI <http://doi.org/10.52766/BEUL1278>
- AGUILERA-COBOS L., ISABEL-GÓMEZ, R. y BLASCO-AMARO, JA., 2022b. Efectividad de la limitación de la movilidad en la evolución de la pandemia por Covid-19. *Ministerio de Sanidad, AETSA Área de Evaluación de Tecnologías Sanitarias* [en línea]. DOI <http://doi.org/10.52766/PYUI7071>
- AGUILERA-COBOS L., ISABEL-GÓMEZ, R. y BLASCO-AMARO, JA., 2022c. Limitación del deporte exterior en la evolución de la pandemia por Covid-19. *Ministerio de Sanidad, AETSA Área de Evaluación de Tecnologías Sanitarias* [en línea]. DOI <http://doi.org/10.52766/RINR3952>

Referencias

- AGUILERA-COBOS L, BAÑOS-ÁLVAREZ E, ISABEL-GÓMEZ R, BLASCO-AMARO JA., 2022d. Ultrasonido focalizado de alta intensidad (HIFU) extracorpóreo en tumores sólidos. *Ministerio de Sanidad, AETSA Área de Evaluación de Tecnologías Sanitarias* [en línea]. DOI <http://www.doi.org/10.52766/JGAJ5300>
- ARMSTRONG, J., FIDDES, I.T., DIEKHANS, M. y PATEN, B., 2019. Whole-Genome Alignment and Comparative Annotation. *Annual review of animal biosciences*, vol. 7, pp. 41-64. ISSN 2165-8110 (Electronic). DOI 10.1146/annurev-animal-020518-115005.
- ARROYO, M., LARROSA, R., CLAROS, M.G. y BAUTISTA, R., 2019. Expression Change Correlations Between Transposons and Their Adjacent Genes in Lung Cancers Reveal a Genomic Location Dependence and Highlights Cancer-Significant Genes BT - *Bioinformatics and Biomedical Engineering*. Springer International Publishing, pp. 84-92. ISBN 978-3-030-17938-0.
- BAETHGE, C., GOLDBECK-WOOD, S. y MERTENS, S., 2019. SANRA—a scale for the quality assessment of narrative review articles. *Research Integrity and Peer Review*, vol. 4, no. 1, pp. 1-8. DOI 10.1186/s41073-019-0064-8.
- BAO, R., HUANG, L., ANDRADE, J., TAN, W., KIBBE, W.A., JIANG, H. y FENG, G., 2014. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Informatics*, vol. 13, pp. 67-82. DOI 10.4137/CIN.S13779.
- BENTLEY, D.R., 2000. The human genome project—an overview. *Medicinal research reviews*, vol. 20, no. 3, pp. 189-196. ISSN 0198-6325.
- BIGMED y LIGHTHOUSE, N.I.C.T., 2021. *Reflections on the clinical implementation of precision medicine*. S.l.: s.n. ISBN 9788247300916.
- BILKEY, G.A., BURNS, B.L., COLES, E.P., BOWMAN, F.L., BEILBY, J.P., PACHTER, N.S., BAYNAM, G., J S DAWKINS, H., NOWAK, K.J. y WEERAMANTHRI, T.S., 2019. Genomic Testing for Human Health and Disease Across the Life Cycle: Applications and Ethical, Legal, and Social Challenges. *Frontiers in public health* [en línea], vol. 7, no. 101616579 PG-40, pp. 40. DOI <https://dx.doi.org/10.3389/fpubh.2019.00040>.

- BLAKE, J.A. y HARRIS, M.A., 2002. The Gene Ontology (GO) project: structured vocabularies for molecular biology and their application to genome and expression analysis. *Current protocols in bioinformatics*, vol. Chapter 7, pp. Unit 7.2. ISSN 1934-340X (Electronic). DOI 10.1002/0471250953.bi0702s00.
- BONANNO, L., DAL MASO, A., PAVAN, A., ZULATO, E., CALVETTI, L., PASELLO, G., GUARNERI, V., CONTE, P. y INDRACCOLO, S., 2022. Liquid biopsy and non-small cell lung cancer: are we looking at the tip of the iceberg? *British journal of cancer*, vol. 127, no. 3, pp. 383-393. ISSN 1532-1827 (Electronic). DOI 10.1038/s41416-022-01777-8.
- BRAVO-GRAU, S. y CRUZ, J.P., 2015. Estudios de exactitud diagnóstica: Herramientas para su Interpretación. *Revista chilena de radiología*, vol. 21, pp. 158-164. DOI 10.4067/S0717-93082015000400007.
- BURNS, B.L., BILKEY, G.A., COLES, E.P., BOWMAN, F.L., BEILBY, J.P., PACHTER, N.S., BAYNAM, G., DAWKINS, H.J.S., WEERAMANTHRI, T.S. y NOWAK, K.J., 2019. Healthcare System Priorities for Successful Integration of Genomics: An Australian Focus. *Frontiers in public health* [en línea], vol. 7, no. 101616579 PG-41, pp. 41. DOI <https://dx.doi.org/10.3389/fpubh.2019.00041>.
- CANDURI, F., 2017. Bioinformatics : an overview and its applications. *Genetics and molecular research*, vol. 16, no. 1, pp. 1-21. DOI 10.4238/gmr16019645.
- CARRUTHERS, M., YURCHENKO, A.A., AUGLEY, J.J., ADAMS, C.E., HERZYK, P. y ELMER, K.R., 2018. De novo transcriptome assembly , annotation and comparison of four ecological and evolutionary model salmonid fish species. *BMC genomics*, pp. 1-17. DOI 10.1186/s12864-017-4379-x.
- CHANFREAU-COFFINIER, C., PEREDO, J., RUSSELL, Marcia M, YANO, E.M., HAMILTON, A.B., LERNER, B., PROVENZALE, D., KNIGHT, S.J., VOILS, C.I. y SCHEUNER, M.T., 2019. A logic model for precision medicine implementation informed by stakeholder views and implementation science. *Genetics in medicine : official journal of the American College of Medical Genetics* [en línea], vol. 21, no. 5 PG-1139-1154, pp. 1139-1154. DOI <https://dx.doi.org/10.1038/s41436-018-0315-y>.

Referencias

- CHEN, G., SHI, T. y SHI, L., 2017. Characterizing and annotating the genome using RNA-seq data. *Science China. Life Science*, vol. 60, no. 2, pp. 116-125. DOI 10.1007/s11427-015-0349-4.
- CHEN, L., ZHANG, Y.-H., WANG, S., ZHANG, Y., HUANG, T. y CAI, Y.-D., 2017. Prediction and analysis of essential genes using the enrichments of gene ontology and KEGG pathways. *PloS one*, vol. 12, no. 9, pp. e0184129. ISSN 1932-6203 (Electronic). DOI 10.1371/journal.pone.0184129.
- CHEN, S., ZHOU, Y., CHEN, Y. y GU, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics (Oxford, England)*, vol. 34, no. 17, pp. i884-i890. ISSN 1367-4811 (Electronic). DOI 10.1093/bioinformatics/bty560.
- CHONG, H.Y., ALLOTEY, P.A. y CHAIYAKUNAPRUK, N., 2018. Current landscape of personalized medicine adoption and implementation in Southeast Asia. *BMC medical genomics* [en línea], vol. 11, no. 1 PG-94, pp. 94. DOI <https://dx.doi.org/10.1186/s12920-018-0420-4>.
- COBB, M., COMFORT., 2023. What Rosalind Franklin truly contributed to the discovery of DNA's structure. *Nature*, vol. 616, no. 7958, pp. 657-660. ISSN 0028-0836. DOI <https://dx.doi.org/10.1038/d41586-023-01313-5>.
- COCK, P.J.A., FIELDS, C.J., GOTO, N., HEUER, M.L. y RICE, P.M., 2010. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic acids research*, vol. 38, no. 6, pp. 1767-1771. ISSN 1362-4962. DOI 10.1093/nar/gkp1137.
- COOLEY, M.E., 2000. Symptoms in adults with lung cancer: a systematic research review. *Journal of pain and symptom management*, vol. 19, no. 2, pp. 137-153. DOI 10.1016/s0885-3924(99)00150-5
- COSTA-SILVA, J., DOMINGUES, D. y LOPES, F.M., 2017. RNA-Seq differential expression analysis: An extended review and a software tool. *PLoS One*, pp. 1-18. DOI 10.1371/journal.pone.0190152
- CRUZES, D. y DYBÅ, T., 2011. Recommended Steps for Thematic Synthesis in Software Engineering. *International Symposium on Empirical Software Engineering and*

- Measurement*, pp. 275-284. DOI 10.1109/ESEM.2011.36
- DAYHOFF, M.O. y LEDLEY, R.S., 1962. Comproteín: a computer program to aid primary protein structure determination. *Proceedings of the December 4-6, 1962, fall joint computer conference*.
- DELNORD, M, VAN VALCKENBORGH, E., HEBRANT, A., ANTONIOU, A., VAN HOOF, W., WAEYTENS, A. y VAN DEN BULCKE, M., 2021. Precision cancer medicine: What has translated into clinical use in Belgium?. *Seminars in cancer biology* [en línea] DOI <https://dx.doi.org/10.1016/j.semcancer.2021.06.010>.
- DENNIS, G.J., SHERMAN, B.T., HOSACK, D.A., YANG, J., GAO, W., LANE, H.C. y LEMPICKI, R.A., 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome biology*, vol. 4, no. 5, pp. P3. ISSN 1474-760X (Electronic).
- DOBIN, A., DAVIS, C.A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. y GINGERAS, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*, vol. 29, no. 1, pp. 15-21. ISSN 1367-4811 (Electronic). DOI 10.1093/bioinformatics/bts635.
- DOYLE, DL., CLYNE, M., RODRIGUEZ, JL., CRAGUN, DL., SENIER, L. y HURST, G., 2018. Proposed outcomes measures for state public health genomic programs. *Genetics in Medicine*. vol. 20, no. 9, pp. 995-1003. DOI 10.1038/gim.2017.229
- ESPINOSA GARCIA, E., ARROYO VARELA, M., LARROSA JIMENEZ, R., GOMEZ-MALDONADO, J., COBO DOLS, M.A., CLAROS, M.G. y BAUTISTA MORENO, R., 2022. Construction of miRNA–mRNA networks for the identification of lung cancer biomarkers in liquid biopsies. *Clinical and Translational Oncology* [en línea], ISSN 1699-3055. DOI 10.1007/s12094-022-02969-7.
- FALGUERAS, J., LARA, A.J., FERNÁNDEZ-POZO, N., CANTÓN, F.R., PÉREZ-TRABADO, G. y CLAROS, M.G., 2010. SeqTrim: a high-throughput pipeline for pre-processing any type of sequence read. *BMC bioinformatics*, vol. 11, pp. 38. ISSN 1471-2105 (Electronic). DOI 10.1186/1471-2105-11-38.
- FERNANDEZ-CHINGUEL, J.E., ZAFRA-TANAKA, J.H., GOICOCHEA-LUGO, S., PERALTA, C.I. y TAYPE-RONDAN, A., 2019. Aspectos básicos sobre la lectura de revisiones

Referencias

- sistemáticas y la interpretación de meta-análisis. *Acta Médica Peruana* [en línea], vol. 36, pp. 157-169. ISSN 1728-5917.
- FLORES-PÉREZ, J.A., DE LA, F., OLIVA, R., ARGENES, Y. y MENESES-GARCIA, A., 2019. Translational Research and Onco-Omics Applications in the Era of Cancer Personal Genomics. *Advances in Experimental Medicine and Biology*. Springer. ISBN 9783030240998.
- FOSSE, V., OLDONI, E., BIETRIX, F., BUDILLON, A., DASKALOPOULOS, E.P., FRATELLI, M., GERLACH, B., GROENEN, P.M.A., HÖLTER, S.M., MENON, J.M.L., MOBASHERI, A., OSBORNE, N., RITSKES-HOITINGA, M., RYLL, B., SCHMITT, E., USSI, A., ANDREU, A.L. y MCCORMACK, E., 2023. Recommendations for robust and reproducible preclinical research in personalised medicine. *BMC medicine*, vol. 21, no. 1, pp. 14. ISSN 1741-7015. DOI 10.1186/s12916-022-02719-0.
- FOSSE, V., OLDONI, E., GERARDI, C., BANZI, R., FRATELLI, M., BIETRIX, F., USSI, A., ANDREU, A.L., MCCORMACK, E. y GROUP, T.P., 2022. Evaluating Translational Methods for Personalized Medicine-A Scoping Review. *Journal of personalized medicine*, vol. 12, no. 7. ISSN 2075-4426. DOI 10.3390/jpm12071177.
- FUJIWARA, K., FUJIMOTO, N., TABATA, M., NISHII, K., MATSUO, K., HOTTA, K., KOZUKI, T., AOE, M., KIURA, K. y UEOKA, H., 2005. Identification of epigenetic aberrant promoter methylation in serum DNA is useful for early detection of lung cancer. *Clinical Cancer Research*, vol. 11, no. 3, pp. 1219-1225.
- GARBER, M., GRABHERR, M.G., GUTTMAN, M. y TRAPNELL, C., 2011. Computational methods for transcriptome annotation and quantification using RNA-seq. *Nature Publishing Group*, vol. 8, no. 6, pp. 469-477. ISSN 1548-7091. DOI 10.1038/nmeth.1613.
- GAUTHIER, J., VINCENT, A.T., CHARETTE, S.J. y DEROME, N., 2019. A brief history of bioinformatics. *Briefings in bioinformatics*, vol. 20, no. June 2018, pp. 1981-1996. DOI 10.1093/bib/bby063.
- GE, S.X., JUNG, D. y YAO, R., 2020. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, vol. 36, no. 8, pp. 2628-2629. ISSN 1367-4811.

- DOI 10.1093/bioinformatics/btz931.
- GHOSH, S. y CHAN, C.-K.K., 2016. Analysis of RNA-Seq Data Using TopHat and Cufflinks. *Methods in molecular biology*, vol. 1374, pp. 339-361. ISSN 1940-6029. DOI 10.1007/978-1-4939-3167-5_18.
- o. 5, pp. 139-148. ISSN 0000000000. DOI 10.1038/gim.2017.229.Proposed.
- GONG, J., SHAO, D., XU, K., LU, Z., LU, Z.J., YANG, Y.T. y ZHANG, Q.C., 2018. RISE: a database of RNA interactome from sequencing experiments. *Nucleic acids research*, vol. 46, no. D1, pp. D194-D201. ISSN 1362-4962 (Electronic). DOI 10.1093/nar/gkx864.
- GOODWIN, S., MCPHERSON, J.D. y MCCOMBIE, W.R., 2016. Coming of age : ten years of next- generation sequencing technologies. *Nature Publishing Group*, vol. 17, no. 6, pp. 333-351. ISSN 1471-0056. DOI 10.1038/nrg.2016.49.
- GROOTEN, W.J.A., TSELI, E., ÄNG, B.O., BOERSMA, K., STÅLNACKE, B.-M., GERDLE, B. y ENTHOVEN, P., 2019. Elaborating on the assessment of the risk of bias in prognostic studies in pain rehabilitation using QUIPS—aspects of interrater agreement. *Diagnostic and Prognostic Research* [en línea], vol. 3, no. 1, pp. 5. ISSN 2397-7523. DOI 10.1186/s41512-019-0050-0.
- HAIJIAN-TILAKI, K., 2013. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian journal of internal medicine*, vol. 4, no. 2, pp. 627-635. ISSN 2008-6164.
- HAN, J.-Y., CHOI, J.-J., KIM, J.Y., HAN, Y.L. y LEE, G.K., 2016. PNA clamping-assisted fluorescence melting curve analysis for detecting EGFR and KRAS mutations in the circulating tumor DNA of patients with advanced non-small cell lung cancer. *BMC cancer*, vol. 16, pp. 627. ISSN 1471-2407. DOI 10.1186/s12885-016-2678-2.
- HE, J., TAN, W. y MA, J., 2017. Circulating tumor cells and DNA for real-time EGFR detection and monitoring of non-small-cell lung cancer. *Future oncology*, vol. 13, no. 9, pp. 787-797. ISSN 1744-8301. DOI 10.2217/fon-2016-0427.
- HERBST, R.S., MADDOX, A.-M., ROTHENBERG, M.L., SMALL, E.J., RUBIN, E.H., BASELGA,

Referencias

- J., ROJO, F., HONG, W.K., SWAISLAND, H. y AVERBUCH, S.D., 2002. Selective oral epidermal growth factor receptor tyrosine kinase inhibitor ZD1839 is generally well-tolerated and has activity in non-small-cell lung cancer and other solid tumors: Results of a phase I trial. *Journal of Clinical Oncology*, vol. 20, no. 18, pp. 3815-3825.
- HIGGINS JPT, THOMAS J, CHANDLER J, CUMPSTON M, LI T, PAGE MJ, W.V., 2022. Cochrane Handbook for Systematic Reviews of Interventions version 6.3. *Cochrane*.
- HOGEWEG, P., 2011. The Roots of Bioinformatics in Theoretical Biology. *PLoS Computational Biology*, vol. 7, no. 3, pp. 1-5. DOI 10.1371/journal.pcbi.1002021.
- HUMPHREY-MURTO, S., VARPIO, L., GONSALVES, C. y WOOD, T.J., 2017. Using consensus group methods such as Delphi and Nominal Group in medical education research. *Medical Teacher*, vol. 39, no. 1, pp. 14-19. DOI 10.1080/0142159X.2017.1245856.
- ISAACSON BARASH, C., 2016. Translating translational medicine into global health equity: What is needed? *Applied and Translational Genomics*, vol. 9, pp. 37-39. DOI 10.1016/j.atg.2016.03.004.
- IZUMI, M., SUZUMURA, T., OGAWA, K., MATSUMOTO, Y., SAWA, K., YOSHIMOTO, N., TANI, Y., WATANABE, T., KANEDA, H., MITSUOKA, S., ASAI, K. y KAWAGUCHI, T., 2020. Differences in molecular epidemiology of lung cancer among ethnicities (Asian vs. Caucasian). *Journal of thoracic disease*, vol. 12, no. 7, pp. 3776-3784. ISSN 2072-1439. DOI 10.21037/jtd.2019.08.61.
- JAIN, M., OLSEN, H.E., TURNER, D.J., STODDART, D., BULAZEL, K. V, PATEN, B., HAUSSLER, D., WILLARD, H.F., AKESON, M. y MIGA, K.H., 2018. Linear assembly of a human centromere on the Y chromosome. *Nature Biotechnology*, vol. 36, no. 4. DOI 10.1038/nbt.4109.
- JU, J., KIM, D.H., BI, L., MENG, Q., BAI, X., LI, Z., LI, X., MARMA, M.S., SHI, S. y WU, J., 2006. Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. *Proceedings of the National Academy of Sciences*, vol. 103, no. 52, pp. 19635-19640. ISSN 0027-8424.
- KANEHISA, M., FURUMICHI, M., SATO, Y., KAWASHIMA, M. y ISHIGURO-WATANABE, M., 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids*

- Research*, vol. 51, no. D1, pp. D587-D592. ISSN 0305-1048. DOI 10.1093/nar/gkac963.
- KANG, J., TANG, Q., HE, J., LI, L., YANG, N., YU, S., WANG, M., ZHANG, Y., LIN, J., CUI, T., HU, Y., TAN, P., CHENG, J., ZHENG, H., WANG, D., SU, X., CHEN, W. y HUANG, Y., 2022. RNAInter v4.0: RNA interactome repository with redefined confidence scoring system and improved accessibility. *Nucleic Acids Research* [en línea], vol. 50, no. D1, pp. D326-D332. ISSN 0305-1048. DOI 10.1093/nar/gkab997.
- KARAGKOUNI, D., PARASKEVOPOULOU, M.D., TASTOGLOU, S., SKOUFOS, G., KARAVANGELI, A., PIERROS, V., ZACHAROPOULOU, E. y HATZIGEORGIOU, A.G., 2020. DIANA-LncBase v3: indexing experimentally supported miRNA targets on non-coding transcripts. *Nucleic Acids Research*, vol. 48, no. D1, pp. D101-D110. ISSN 0305-1048. DOI 10.1093/nar/gkz1036.
- KONO, N. y ARAKAWA, K., 2019. Nanopore sequencing: Review of potential applications in functional genomics. *Development Growth and Differentiation*, vol. 61, no. 5, pp. 316-326. ISSN 1440169X. DOI 10.1111/dgd.12608.
- KOVANDA, A., ZIMANI, A.N. y PETERLIN, B., 2021. How to design a national genomic project—a systematic review of active projects. *Human Genomics*, vol. 15, no. 1, pp. 1-13. ISSN 14797364. DOI 10.1186/s40246-021-00315-6.
- KURNAT-THOMA, E., 2020. Educational and Ethical Considerations for Genetic Test Implementation Within Health Care Systems. *Network and systems medicine*, vol. 3, no. 1 PG-58-66, pp. 58-66. DOI <https://dx.doi.org/10.1089/nsm.2019.0010>.
- LAI, H., LI, Yuchen, ZHANG, H., HU, J., LIAO, J., SU, Y., LI, Q., CHEN, B., LI, C., WANG, Z., LI, Yan, WANG, J., MENG, Z., HUANG, Z. y HUANG, S., 2022. exoRBase 2.0: an atlas of mRNA, lncRNA and circRNA in extracellular vesicles from human biofluids. *Nucleic Acids Research*, vol. 50, no. D1, pp. D118-D128. ISSN 0305-1048. DOI 10.1093/nar/gkab1085.
- LEE, D. y KIM, K., 2022. Public R&D Projects-Based Investment and Collaboration Framework for an Overarching South Korean National Strategy of Personalized Medicine. *International journal of environmental research and public health*, vol.

Referencias

- 19, no. 3. DOI 10.3390/ijerph19031291.
- LEONARDI, T. y LEGER, A., 2021. Nanopore RNA Sequencing Analysis. *Methods in molecular biology*, vol. 2284, pp. 569-578. ISSN 1940-6029. DOI 10.1007/978-1-0716-1307-8_31.
- LI, L., PENG, M., XUE, W., FAN, Z., WANG, T., LIAN, J., ZHAI, Y., LIAN, W., QIN, D. y ZHAO, J., 2018. Integrated analysis of dysregulated long non-coding RNAs/microRNAs/mRNAs in metastasis of lung adenocarcinoma. *Journal of Translational Medicine*, vol. 16, no. 1, pp. 372. ISSN 1479-5876. DOI 10.1186/s12967-018-1732-z.
- LI, W., LIU, J. Bin, HOU, L.K., YU, F., ZHANG, J., WU, W., TANG, X.M., SUN, F., LU, H.M., DENG, J., BAI, J., LI, J., WU, C.Y., LIN, Q.L., LV, Z.W., WANG, G.R., JIANG, G.X., MA, Y.S. y FU, D., 2022. Liquid biopsy in lung cancer: significance in diagnostics, prediction, and treatment monitoring. *Molecular Cancer*, vol. 21, no. 1, pp. 1-15. DOI 10.1186/s12943-022-01505-z.
- LI, Z., LIU, L., FENG, C., QIN, Y., XIAO, J., ZHANG, Z. y MA, L., 2023. LncBook 2.0: integrating human long non-coding RNAs with multi-omics annotations. *Nucleic Acids Research*, vol. 51, no. D1, pp. D186-D191. ISSN 0305-1048. DOI 10.1093/nar/gkac999.
- LIAO, W.W., ASRI, M., EBLER, J. y DOERR, D., 2023. A draft human pangenome reference. *Nature*, vol. 617, no. 7960 pp. 312-324. ISSN 0028-0836. DOI 10.1038/s41586-023-05896-x.
- LIBERATI, A., ALTMAN, D.G., TETZLAFF, J., MULROW, C., GØTZSCHE, P.C., IOANNIDIS, J.P.A., CLARKE, M., DEVEREAUX, P.J., KLEIJNEN, J. y MOHER, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of clinical epidemiology*, vol. 62, no. 10, pp. e1-e34. ISSN 0895-4356.
- LINARES-ESPINÓS, E., HERNÁNDEZ, V., DOMÍNGUEZ-ESCRIG, J.L., FERNÁNDEZ-PELLO, S., HEVIA, V., MAYOR, J., PADILLA-FERNÁNDEZ, B. y RIBAL, M.J., 2018. Metodología de una revisión sistemática. *Actas Urológicas Españolas*, vol. 42, no. 8, pp. 499-506.

- ISSN 0210-4806. DOI <https://doi.org/10.1016/j.acuro.2018.01.010>.
- LIU, N., SUN, R.-Z., DU, J., DONG, Q.-Z., FAN, C.-F., LI, Q.-C., WANG, E.-H. y LIU, Y., 2018. Comparison of Epidermal Growth Factor Receptor Gene Mutations Identified Using Pleural Effusion and Primary Tumor Tissue Samples in Non-Small Cell Lung Cancer. *Applied immunohistochemistry & molecular morphology: AIMM*, vol. 26, no. 4, pp. e44-e51. ISSN 1533-4058. DOI 10.1097/PAI.0000000000000543.
- LOGSDON, G.A., VOLLGER, M.R., HSIEH, P., MAO, Y., LISKOYKH, M.A., KOREN, S., NURK, S., MERCURI, L., DISHUCK, P.C., RHIE, A., LIMA, L.G. De, DVORKINA, T., PORUBSKY, D., HARVEY, W.T., MIKHEENKO, A., BZIKADZE, A. V, KREMITZKI, M., GRAVES-LINDSAY, T.A., JAIN, C., HOEKZEMA, K., MURALI, S.C., MUNSON, K.M., BAKER, C., SORENSEN, M., LEWIS, A.M., SURTI, U., GERTON, J.L., LARIONOV, V., VENTURA, M., MIGA, K.H., PHILLIPPY, A.M. y EICHLER, E.E., 2021. The structure , function and evolution of a complete human chromosome 8. *Nature*, vol. 593, no. May. ISSN 1476-4687. DOI 10.1038/s41586-021-03420-7.
- LYU, M., ZHOU, J., NING, K. y YING, B., 2019. The diagnostic value of circulating tumor cells and ctDNA for gene mutations in lung cancer. *OncoTargets and Therapy*, vol. 12, pp. 2539-2552. DOI 10.2147/OTT.S195342.
- MAILLER, E., PAILLART, J., MARQUET, R. y VIVET-BOUDOU, V., 2018. The evolution of RNA structural probing methods: From gels to next-generation sequencing. *Wiley Interdisciplinary Reviews RNA*, no. October, pp. 1-20. DOI 10.1002/wrna.1518.
- MARGUERAT, S. y BÄHLER, J., 2010. RNA-seq: from technology to biology. *Cellular and molecular life science*, vol.67, no. 4, pp. 569-79. DOI 10.1007/s00018-009-0180-6
- MARGULIES, M., EGHOLM, M., ALTMAN, W.E., ATTIYA, S., BADER, J.S., BEMBEN, L.A., BERKA, J., BRAVERMAN, M.S., CHEN, Y.J. y CHEN, Z., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, vol. 437, no. 7057, pp. 376-380. ISSN 1476-4687.
- MARTIN, J.A. y WANG, Z., 2011. Next-generation transcriptome assembly. *Nature reviews. Genetics*, vol. 12, no. 10, pp. 671-682. ISSN 1471-0064 (Electronic). DOI 10.1038/nrg3068.

Referencias

- MATTICK, J.S., AMARAL, P.P., CARNINCI, P., CARPENTER, S., CHANG, H.Y., CHEN, L.L., CHEN, R., DEAN, C., DINGER, M.E., FITZGERALD, K.A., GINGERAS, T.R., GUTTMAN, M., HIROSE, T., HUARTE, M., JOHNSON, R., KANDURI, C., KAPRANOV, P., LAWRENCE, B., LEE, J.T., MENDELL, J.T., MERECER, T.R., MOORE, K.J., NAKAGAWA, S., RINN, J.L., SPECTOR, D.L., ULITSKY, I., WAN, Y., WILUSZ, J.E. y WU, M., 2023. Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nature Reviews Molecular Cell Biology*, vol. 24, pp. 430–447. <https://doi.org/10.1038/s41580-022-00566-8>
- MAXAM, A.M. y GILBERT, W., 1977. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences*, vol. 74, no. 2, pp. 560-564. ISSN 0027-8424.
- MCCOY, R.C., TAYLOR, R.W., BLAUWKAMP, T.A., KELLEY, J.L., KERTESZ, M., PUSHKAREV, D., PETROV, D.A. y FISTON-LAVIER, A.-S., 2014. Illumina TruSeq synthetic long-reads empower de novo assembly and resolve complex, highly-repetitive transposable elements. *PloS One*, vol. 9, no. 9, pp. e106689. ISSN 1932-6203. DOI: 10.1371/journal.pone.0106689
- METZKER, M.L., 2010. Sequencing technologies—the next generation. *Nature reviews genetics*, vol. 11, no. 1, pp. 31-46. ISSN 1471-0064. DOI 10.1038/nrg2626
- MIGA, K.H., KOREN, S., RHIE, A., VOLLGER, M.R., GERSHMAN, A., BZIKADZE, A., BROOKS, S., HOWE, E., PORUBSKY, D., LOGSDON, G.A., SCHNEIDER, V.A., POTAPOVA, T., WOOD, J., CHOW, W., ARMSTRONG, J., MULLIKIN, J.C., PEVZNER, P.A., GERTON, J.L. y SULLIVAN, B.A., 2020. Telomere-to-telomere assembly of a complete human X chromosome. *Nature*, vol. 585, no. July 2019. ISSN 1476-4687. DOI 10.1038/s41586-020-2547-7.
- NURK, S., KOREN, S., RHIE, A., RAUTIAINEN, M., BZIKADZE, A. V, MIKHEENKO, A., VOLLGER, M.R., ALTEMOSE, N., URALSKY, L., GERSHMAN, A., AGANEZOV, S., HOYT, S.J., DIEKHANS, M., LOGSDON, G.A., ALONGE, M., ANTONARAKIS, S.E., BORCHERS, M., BOUFFARD, G.G., BROOKS, S.Y., CALDAS, G. V, CHEN, N.-C., CHENG, H., CHIN, C.-S., CHOW, W., DE LIMA, L.G., DISHUCK, P.C., DURBIN, R., DVORKINA, T., FIDDES, I.T., FORMENTI, G., FULTON, R.S., FUNGTAMMASAN, A., GARRISON, E., GRADY, P.G.S., GRAVES-LINDSAY, T.A., HALL, I.M., HANSEN, N.F., HARTLEY, G.A.,

- HAUKNESS, M., HOWE, K., HUNKAPILLER, M.W., JAIN, C., JAIN, M., JARVIS, E.D., KERPEDJIEV, P., KIRSCH, M., KOLMOGOROV, M., KORLACH, J., KREMITZKI, M., LI, H., MADURO, V. V, MARSCHALL, T., MCCARTNEY, A.M., MCDANIEL, J., MILLER, D.E., MULLIKIN, J.C., MYERS, E.W., OLSON, N.D., PATEN, B., PELUSO, P., PEVZNER, P.A., PORUBSKY, D., POTAPOVA, T., ROGAEV, E.I., ROSENFELD, J.A., SALZBERG, S.L., SCHNEIDER, V.A., SEDLAZECK, F.J., SHAFIN, K., SHEW, C.J., SHUMATE, A., SIMS, Y., SMIT, A.F.A., SOTO, D.C., SOVIĆ, I., STORER, J.M., STREETS, A., SULLIVAN, B.A., THIBAUD-NISSEN, F., TORRANCE, J., WAGNER, J., WALENZ, B.P., WENGER, A., WOOD, J.M.D., XIAO, C., YAN, S.M., YOUNG, A.C., ZARATE, S., SURTI, U., MCCOY, R.C., DENNIS, M.Y., ALEXANDROV, I.A., GERTON, J.L., O'NEILL, R.J., TIMP, W., ZOOK, J.M., SCHATZ, M.C., EICHLER, E.E., MIGA, K.H. y PHILLIPPY, A.M., 2022. The complete sequence of a human genome. *Science*, vol. 376, no. 6588, pp. 44-53. DOI 10.1126/science.abj6987.
- ORFAO, A., BENÍTEZ, J., CORRALES, F., MARTÍN-SUBERO, I., ORDOVÁS, J.M., CARRACEDO, Á. y LAPUNZINA, P., 2019. Ciencias ómicas. *Instituto Roche*, pp. 32.
- PAPANICOLAS, I., RAJAN, D. y KARANIKOLOS, M., 2022. Health system performance assessment: A framework for policy analysis. *European Observatory on Health Systems and Policies*. Health Policy Series, no. 57. ISBN-13: 9789240042476
- PASTORINO, R., LANG, M., ZAWATI, M.H, CALABRÒ, G.E., HOXHAJ, I., HOUWINK, E.J.F., SASSANO, M. y BOCCIA, S., 2021. Ethico-legal and Policy Issues Surrounding Personalised Medicine. *Personalised Health Care: Fostering Precision Medicine Advancements for Gaining Population Health Impact*, pp. 57–76. DOI 10.1007/978-3-030-52399-2_4
- PEARCE, C., GOETTKE, E., HALLOWELL, N., MCCORMACK, P., FLINTER, F. y MCKEVITT, C., 2019. Delivering genomic medicine in the United Kingdom National Health Service: a systematic review and narrative synthesis. *Genetics in medicine : official journal of the American College of Medical Genetics*, vol. 21, no. 12, pp. 2667-2675. DOI <https://dx.doi.org/10.1038/s41436-019-0579-x>.
- PECORINO, L., 2021. Molecular biology of cancer: mechanisms, targets, and therapeutics. *Oxford university press*. ISBN 0198833024.

Referencias

- PITINI, E., D'ANDREA, E., DE VITO, C., ROSSO, A., UNIM, B., MARZUILLO, C., FEDERICI, A., DI MARIA, E. y VILLARI, P., 2019. A proposal of a new evaluation framework towards implementation of genetic tests., *PloS one* [en línea], vol. 14, no. 8 PG-e0219755, pp. e0219755. DOI <https://dx.doi.org/10.1371/journal.pone.0219755>. Disponible en: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med16&NEWS=N&AN=31381569> NS -.
- PITINI, E., D'ANDREA, E., DE VITO, C., ROSSO, A., UNIM, B., MARZUILLO, C., FEDERICI, A., MARIA, E. Di y VILLARI, P., 2019. A proposal of a new evaluation framework towards implementation of genetic tests. *PLoS One*, vol. 14, no. 8, pp. 1-15. ISSN 19326203. DOI 10.1371/journal.pone.0219755.
- PON, J.R. y MARRA, M.A., 2015. Driver and passenger mutations in cancer. *Annual review of pathology*, vol. 10, pp. 25-50. ISSN 1553-4014. DOI 10.1146/annurev-pathol-012414-040312.
- PRITCHARD, D.E., MOECKEL, F., VILLA, M.S., HOUSMAN, L.T., MCCARTY, C.A. y MCLEOD, H.L., 2017. Strategies for integrating personalized medicine into healthcare practice. *Personalized medicine*, vol. 14, no. 2, pp. 141-152. DOI <https://dx.doi.org/10.2217/pme-2016-0064>.
- QUAIL, M.A., SMITH, M., COUPLAND, P., OTTO, T.D., HARRIS, S.R., CONNOR, T.R., BERTONI, A., SWERDLOW, H.P. y GU, Y., 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC genomics*, vol. 13, no. 1, pp. 1-13. ISSN 1471-2164. DOI 10.1186/1471-2164-13-341
- RAGHAVAN, V., KRAFT, L., MESNY, F. y RIGERTE, L., 2022. A simple guide to de novo transcriptome assembly and annotation. *Briefings in bioinformatics*, vol. 23, no. 2. ISSN 1477-4054. DOI 10.1093/bib/bbab563.
- RICCIARDI, W. y BOCCIA, S., 2017. New challenges of public health: bringing the future of personalised healthcare into focus. *European journal of public health*, vol. 27, no. suppl_4 PG-36-39, pp. 36-39. DOI <https://dx.doi.org/10.1093/eurpub/ckx164>.

- RODRÍGUEZ WEBER, F.L. y RAMÍREZ ARIAS, J.L., 2022. Medicina personalizada, individualizada, de precisión y centrada en el paciente; diferencias o sinónimos. Su importancia. *Acta Médica Grupo Ángeles*, vol. 20, no. 1, pp. 111-112. DOI 10.35366/103572.
- ROQUÉ, M., MARTÍNEZ-GARCÍA, L., SOLÀ, I., ALONSO-COELLO, P., BONFILL, X. y ZAMORA, J., 2020. Toolkit of methodological resources to conduct systematic reviews. *F1000Research*, vol. 9, pp. 82. ISSN 2046-1402. DOI 10.12688/f1000research.22032.3.
- ROTHBERG, J.M., HINZ, W., REARICK, T.M., SCHULTZ, J., MILESKE, W., DAVEY, M., LEAMON, J.H., JOHNSON, K., MILGREW, M.J. y EDWARDS, M., 2011. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, vol. 475, no. 7356, pp. 348-352. ISSN 1476-4687. DOI 10.1038/nature10242
- SAGNER, M., MCNEIL, A., PUSKA, P., AUFRAY, C., PRICE, N.D., HOOD, L., LAVIE, C.J., HAN, Z.-G., CHEN, Z., BRAHMACHARI, S.K., MCEWEN, B.S., SOARES, M.B., BALLING, R., EPEL, E. y ARENA, R., 2017. The P4 Health Spectrum – A Predictive, Preventive, Personalized and Participatory Continuum for Promoting Healthspan. *Progress in Preventive Medicine*, vol. 2, no. 1, pp. e0002-e0002. ISSN 00000000000000. DOI 10.1097/pp9.0000000000000002.
- SANGER, F., NICKLEN, S. y COULSON, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences*, vol. 74, no. 12, pp. 5463-5467. ISSN 0027-8424.
- SANKAR, P.L. y PARKER, L.S., 2017. The Precision Medicine Initiative's All of Us Research Program: an agenda for research on its ethical, legal, and social issues. *Genetics in medicine : official journal of the American College of Medical Genetics*, vol. 19, no. 7, pp. 743-750. DOI 10.1038/gim.2016.183.
- SHEA, B.J., REEVES, B.C., WELLS, G., THUKU, M., HAMEL, C., MORAN, J., MOHER, D., TUGWELL, P., WELCH, V., KRISTJANSSON, E. y HENRY, D.A., 2017. AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, vol. 358, pp. 1-9.

Referencias

- ISSN 17561833. DOI 10.1136/bmj.j4008.
- SHERMAN, B.T., HAO, M., QIU, J., JIAO, X., BASELER, M.W., LANE, H.C., IMAMICHI, T. y CHANG, W., 2022. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*, vol. 50, no. W1, pp. W216-W221. ISSN 0305-1048. DOI 10.1093/nar/gkac194.
- SHREFFLER, J. y HUECKER, M.R., 2022. Diagnostic Testing Accuracy: Sensitivity, Specificity, Predictive Values and Likelihood Ratios. *StatPearls. Treasure Island*. PMID 32491423.
- SOCIEDAD ESPAÑOLA DE ONCOLOGÍA MÉDICA, 2021. *Las cifras del cáncer en España 2021*. ISBN 9788409277049.
- STADEN, R., 1977. Sequence data handling by computer. *Nucleic acids research*, vol. 4, no. 11, pp. 4037-4051. ISSN 0305-1048. DOI 10.1093/nar/4.11.4037.
- STANG, A., 2010. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *European journal of epidemiology*, vol. 25, no. 9, pp. 603-605. ISSN 0393-2990. DOI 10.1007/s10654-010-9491-z.
- STERNE, J.A.C., HERNÁN, M.A., REEVES, B.C., SAVOVIĆ, J., BERKMAN, N.D., VISWANATHAN, M., HENRY, D., ALTMAN, D.G., ANSARI, M.T., BOUTRON, I., CARPENTER, J.R., CHAN, A.-W., CHURCHILL, R., DEEKS, J.J., HRÓBJARTSSON, A., KIRKHAM, J., JÜNI, P., LOKE, Y.K., PIGOTT, T.D., RAMSAY, C.R., REGIDOR, D., ROTHSTEIN, H.R., SANDHU, L., SANTAGUIDA, P.L., SCHÜNEMANN, H.J., SHEA, B., SHRIER, I., TUGWELL, P., TURNER, L., VALENTINE, J.C., WADDINGTON, H., WATERS, E., WELLS, G.A., WHITING, P.F. y HIGGINS, J.P.T., 2016. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, vol. 355, pp. i4919. DOI 10.1136/bmj.i4919.
- STERNE, J.A.C., SAVOVIĆ, J., PAGE, M.J., ELBERS, R.G., BLENCOWE, N.S., BOUTRON, I., CATES, C.J., CHENG, H.-Y., CORBETT, M.S., ELDRIDGE, S.M., EMBERSON, J.R., HERNÁN, M.A., HOPEWELL, S., HRÓBJARTSSON, A., JUNQUEIRA, D.R., JÜNI, P., KIRKHAM, J.J., LASSERSON, T., LI, T., MCALEENAN, A., REEVES, B.C., SHEPPERD, S.,

- SHRIER, I., STEWART, L.A., TILLING, K., WHITE, I.R., WHITING, P.F. y HIGGINS, J.P.T., 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* [en línea], vol. 366, pp. l4898. DOI 10.1136/bmj.l4898. Disponible en: <http://www.bmj.com/content/366/bmj.l4898.abstract>.
- SUNG, H., FERLAY, J., SIEGEL, R.L., LAVERSANNE, M., SOERJOMATARAM, I., JEMAL, A. y BRAY, F., 2021. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209-249. ISSN 0007-9235. DOI 10.3322/caac.21660.
- SZKLARCZYK, D., GABLE, A.L., NASTOU, K.C., LYON, D., KIRSCH, R., PYYSALO, S., DONCHEVA, N.T., LEGEAY, M., FANG, T., BORK, P., JENSEN, L.J. y VON MERING, C., 2021. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, vol. 49, no. D1, pp. D605-D612. ISSN 1362-4962. DOI 10.1093/nar/gkaa1074.
- TANIGUCHI, K., UCHIDA, J., NISHINO, K., KUMAGAI, T., OKUYAMA, T., OKAMI, J., HIGASHIYAMA, M., KODAMA, K., IMAMURA, F. y KATO, K., 2011. Quantitative detection of EGFR mutations in circulating tumor DNA derived from lung adenocarcinomas. *Clinical cancer research : an official journal of the American Association for Cancer Research*, vol. 17, no. 24, pp. 7808-7815. ISSN 1557-3265. DOI 10.1158/1078-0432.CCR-11-1712.
- TATE, J.G., BAMFORD, S., JUBB, H.C., SONDKA, Z., BEARE, D.M., BINDAL, N., BOUTSELAKIS, H., COLE, C.G., CREATORE, C., DAWSON, E., FISH, P., HARSHA, B., HATHAWAY, C., JUPE, S.C., KOK, C.Y., NOBLE, K., PONTING, L., RAMSHAW, C.C., RYE, C.E., SPEEDY, H.E., STEFANCSIK, R., THOMPSON, S.L., WANG, S., WARD, S., CAMPBELL, P.J. y FORBES, S.A., 2019. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic acids research*, vol. 47, no. D1, pp. D941-D947. ISSN 1362-4962. DOI 10.1093/nar/gky1015.
- THOMAS, J. y HARDEN, A., 2008. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, vol. 8, pp. 1-10. ISSN 14712288. DOI 10.1186/1471-2288-8-45.

Referencias

- TRAPNELL, C., ROBERTS, A., GOFF, L., PERTEA, G., KIM, D., KELLEY, D.R., PIMENTEL, H., SALZBERG, S.L., RINN, J.L. y PACHTER, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, vol. 7, no. 3, pp. 562-578. ISSN 17542189. DOI 10.1038/nprot.2012.016.
- TRAVIS, W.D., BRAMBILLA, E., BURKE, A.P., MARX, A. y NICHOLSON, A.G., 2015. Introduction to the 2015 World Health Organization classification of tumors of the lung, pleura, thymus, and heart. *Journal of Thoracic Oncology*, vol. 10, no. 9, pp. 1240-1242. DOI 10.1097/JTO.0000000000000663
- UNIM, B., PITINI, E., LAGERBERG, T., ADAMO, G., DE VITO, C., MARZUILLO, C. y VILLARI, P., 2019. Current Genetic Service Delivery Models for the Provision of Genetic Testing in Europe: A Systematic Review of the Literature. *Frontiers in genetics*, vol. 10, no. 101560621, pp. 552. DOI <https://dx.doi.org/10.3389/fgene.2019.00552>.
- VALOUEV, A., ICHIKAWA, J., TONTHAT, T., STUART, J., RANADE, S., PECKHAM, H., ZENG, K., MALEK, J.A., COSTA, G. y MCKERNAN, K., 2008. A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome research*, vol. 18, no. 7, pp. 1051-1063. ISSN 1088-9051.
- VAN DIJK, E.L., JASZCZYSZYN, Y., NAQUIN, D. y THERMES, C., 2018. The third revolution in sequencing technology. *Trends in Genetics*, vol. 34, no. 9, pp. 666-681. ISSN 0168-9525. DOI 10.1101/gr.076463.108
- VENTER, J.C., ADAMS, M.D., MYERS, E.W., LI, P.W., MURAL, R.J., SUTTON, G.G., SMITH, H.O., YANDELL, M., EVANS, C.A. y HOLT, R.A., 2001. The sequence of the human genome. *Science*, vol. 291, no. 5507, pp. 1304-1351. ISSN 0036-8075. DOI 10.1126/science.1058040
- VICENTE, A.M., BALLENSIEFEN, W. y JÖNSSON, J.I., 2020. How personalised medicine will transform healthcare by 2030: The ICPeMed vision. *Journal of Translational Medicine*, vol. 18, no. 1, pp. 1-4. ISSN 14795876. DOI 10.1186/s12967-020-02316-w.
- VISVIKIS-SIEST, S., GORENJAK, V. y STATHOPOULOU, M.G., 2018. Personalised medicine: The odyssey from hope to practice. *Journal of Personalized Medicine*, vol. 8, no. 4.

DOI 10.3390/jpm8040031.

VRIJENHOEK, T., TONISSON, N., KÄÄRIÄINEN, H., LEITSALU, L. y RIGTER, T., 2021. Clinical genetics in transition—a comparison of genetic services in Estonia, Finland, and the Netherlands. *Journal of community genetics*, vol. 12, no. 2 PG-277-290, pp. 277-290. DOI 10.1007/s12687-021-00514-7.

WANG, M., HERBST, Roy S y BOSHOF, C., 2021. Toward personalized treatment approaches for non-small-cell lung cancer. *Nature Medicine*, vol. 27, no. 8, pp. 1345-1356. DOI 10.1038/s41591-021-01450-2.

WANG, M., HERBST, Roy S. y BOSHOF, C., 2021. Toward personalized treatment approaches for non-small-cell lung cancer. *Nature Medicine*, vol. 27, no. 8, pp. 1345-1356. ISSN 1546170X. DOI 10.1038/s41591-021-01450-2.

WANG, M., ZHENG, S., LI, X., DING, Y., ZHANG, M., LIN, L., XU, Hao, CHENG, Y., ZHANG, X., XU, Hui y LI, S., 2020. Integrated Analysis of lncRNA-miRNA-mRNA ceRNA Network Identified lncRNA EPB41L4A-AS1 as a Potential Biomarker in Non-small Cell Lung Cancer. *Frontiers in Genetics*, vol. 11. ISSN 1664-8021. DOI 10.3389/fgene.2020.511676.

WANG, T., YANG, C., LI, B., XING, Y., HUANG, J., ZHANG, Y., BU, S. y GE, H., 2021. Identification of lncRNA-miRNA-mRNA Networks Linked to Non-small Lung Cancer Resistance to Inhibitors of Epidermal Growth Factor Receptor. *Frontiers in genetics*, vol. 12, pp. 758591. ISSN 1664-8021. DOI 10.3389/fgene.2021.758591.

WANG, Z., GERSTEIN, M. y SNYDER, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, vol. 10, no. 1, pp. 57-63. DOI 10.1038/nrg2484

WATSON, J.D. y CRICK, F.H.C., 1953. The structure of DNA. *Cold Spring Harbor symposia on quantitative biology*, pp. 123-131. ISBN 0091-7451.

WETTERSTRAND, K.A., 2013. DNA sequencing costs: data from the NHGRI Genome Sequencing Program (GSP). *NHGRI Genome Sequencing Program*.

WHITING, P.F., RUTJES, A.W.S., WESTWOOD, M.E., MALLETT, S., DEEKS, J.J., REITSMA,

Referencias

- J.B., LEEFLANG, M.M.G., STERNE, J.A.C. y BOSSUYT, P.M.M., 2011. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine*, vol. 155, no. 8, pp. 529-536. ISSN 0003-4819. DOI 10.7326/0003-4819-155-8-201110180-00009
- WHITSEL, Laurie P, WILBANKS, J., HUFFMAN, M.D. y HALL, J.L., 2019. The Role of Government in Precision Medicine, Precision Public Health and the Intersection With Healthy Living. *Progress in cardiovascular diseases*, vol. 62, no. 1, pp. 50-54. DOI <https://dx.doi.org/10.1016/j.pcad.2018.12.002>.
- WHITSEL, Laurie P., WILBANKS, J., HUFFMAN, M.D. y HALL, J.L., 2019. The Role of Government in Precision Medicine, Precision Public Health and the Intersection With Healthy Living. *Progress in Cardiovascular Diseases*, vol. 62, no. 1, pp. 50-54. ISSN 18731740. DOI 10.1016/j.pcad.2018.12.002.
- WOLFF, R.F., MOONS, K.G.M., RILEY, R.D., WHITING, P.F., WESTWOOD, M., COLLINS, G.S., REITSMA, J.B., KLEIJNEN, J. y MALLETT, S., 2019. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Annals of internal medicine*, vol. 170, no. 1, pp. 51-58. ISSN 1539-3704. DOI 10.7326/M18-1376.
- WU, Z., XU, Z., YU, Boyao, ZHANG, J. y YU, Bentong, 2020. The Potential Diagnostic Value of Exosomal Long Noncoding RNAs in Solid Tumors: A Meta-Analysis and Systematic Review. *BioMed Research International*, vol. 2020. DOI 10.1155/2020/6786875.
- YAMAMOTO, F., CLAUSEN, H., WHITE, T., MARKEN, J. y HAKOMORI, S., 1990. Molecular genetic basis of the histo-blood group ABO system. *Nature*, vol. 345, no. 6272, pp. 229-233. DOI 10.1038/345229a0.
- YI, M., LIAO, Z., DENG, L., XU, L., TAN, Y., LIU, K., CHEN, Z. y ZHANG, Y., 2021. High diagnostic value of miRNAs for NSCLC: quantitative analysis for both single and combined miRNAs in lung cancer. *Annals of Medicine*, vol. 53, no. 1, pp. 2178-2193. DOI 10.1080/07853890.2021.2000634.
- YONEDA, K., IMANISHI, N., ICHIKI, Y. y TANAKA, F., 2019. Treatment of Non-small Cell Lung Cancer with EGFR-mutations. *Journal of UOEH*, vol. 41, no. 2, pp. 153-163. ISSN 0387-821X. DOI 10.7888/juoeh.41.153.

- ZAMORA, J., ABRAIRA, V., MURIEL, A., KHAN, K. y COOMARASAMY, A., 2006. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC medical research methodology*, vol. 6, pp. 31. ISSN 1471-2288. DOI 10.1186/1471-2288-6-31.
- ZHAO, Y., WANG, H. y HE, C., 2021. Drug resistance of targeted therapy for advanced non-small cell lung cancer harbored EGFR mutation: from mechanism analysis to clinical strategy. *Journal of Cancer Research and Clinical Oncology*, vol. 147, no. 12, pp. 3653-3664. ISSN 1432-1335. DOI 10.1007/s00432-021-03828-8.



UNIVERSIDAD
DE MÁLAGA

Anexos



UNIVERSIDAD
DE MÁLAGA

Anexos

Anexo 1: Estrategias de búsqueda de la revisión sistemática para la generación de un marco de evaluación de la implementación de la MPP a nivel nacional.

- Estrategia de búsqueda en la base de datos Medline

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily <1946 to February 14, 2022>

Search Strategy:

-
- 1 exp Genomics/ (135052)
 - 2 exp Genetic Testing/ (51057)
 - 3 ((gene or genetic or genom* or multigene or multi-gene) adj3 (test* or panel*)).tw. (58499)
 - 4 Genetic Services/ or Genetic Counseling/ (15402)
 - 5 (genetic adj3 (service or counsel* or screening?)).tw. (29322)
 - 6 (genomic? adj3 (resource or service? or portfolio)).tw. (1316)
 - 7 (Prenatal adj3 (Noninvasive or Screening)).tw. (5527)
 - 8 ((mass or carrier) and screening).tw. (42895)
 - 9 *molecular diagnostic techniques/ (7905)
 - 10 ((molecular or genetic) and techni*).tw. (157471)
 - 11 Pharmacogenetics/ (12868)
 - 12 (pharmacogenomic\$ or pharmacogenetic\$).ti,ab. (17398)
 - 13 Precision Medicine/ (23193)

Anexos

- 14 (((individualized or personalized or precision or predictive or stratified) adj3 medicine) or 'p health' or p-health or theranostic or theranostics).tw. (35431)
- 15 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 (506540)
- 16 Health Plan Implementation/ or Translational Medical Research/ (18521)
- 17 (implement* or disseminat\$ or translation or transfer or incorporation or integration).tw. (1653569)
- 18 16 or 17 (1663276)
- 19 Health Resources/ (14527)
- 20 National Health Programs/ (33072)
- 21 ((health adj3 (national or program? or insurance)) or "GOVERNMENT PROGRAMS").ti. (39614)
- 22 Regional Health Planning/ (5454)
- 23 ((health adj3 planning) or (annual adj3 plan?) or (Appropriateness adj3 Review)).tw. (13725)
- 24 Health Systems Plans/ (194)
- 25 (Health adj3 Plan?).tw. (19681)
- 26 Program Development/ (30160)
- 27 (program adj3 (development or description?)).tw. (9740)
- 28 (strategy adj3 (national or plan)).tw. (3864)
- 29 Health Policy/ (70488)
- 30 (Polic* adj3 (Health or Healthcare)).ti,ab. (52908)
- 31 "Delivery of Health Care"/ (104818)
- 32 (Delivery adj3 (Health or Healthcare)).tw. (28823)
- 33 Health Services Accessibility/ (82457)
- 34 (Access* adj3 (care or Services or Therap* or treatment?)).tw. (59767)
- 35 Community Health Services/og, st [Organization & Administration, Standards] (10550)
- 36 (Community adj3 Service?).tw. (18435)

- 37 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 (483316)
- 38 15 and 18 and 37 (1241)
- 39 animals/ not (animals/ and humans/) (4924998)
- 40 38 not 39 (1235)
- 41 limit 40 to yr="2016 -Current" (665)

- Estrategia de búsqueda en la base de datos EMBASE

	594
#40	
#39 AND [embase]/lim NOT ([embase]/lim AND [medline]/lim) AND [2016-2022]/py AND [humans]/lim	2,981
#39	
#15 AND #19 AND #38	809,453
#38	
#20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37	16,622
#37	
(community NEAR/3 service?):ti,ab	10,637
#36	
'community care'/de AND 'organization and management'/de	80,490
#35	
(access* NEAR/3 (care OR services OR therap* OR treatment?):ti,ab	73,462
#34	
'health care access'/de	34,770
#33	
(delivery NEAR/3 (health OR healthcare)):ti,ab	193,477
#32	
'health care delivery'/de	61,732
#31	

Anexos

(polic* NEAR/3 (health OR healthcare)):ti,ab	207,582
#30	
'health care policy'/de	5,350
#29	
(strategy NEAR/3 (national OR plan)):ti,ab	12,346
#28	
(program NEAR/3 (development OR description?):ti,ab	25,279
#27	
'program development'/de	11,852
#26	
(health NEAR/3 plan?):ti,ab	12
#25	
'health systems plans':ti,ab	16,066
#24	
((health NEAR/3 planning):ti,ab) OR ((annual NEAR/3 plan?):ti,ab) OR ((appropriateness NEAR/3 review):ti,ab)	70
#23	
'regional health planning':ti,ab	38,517
#22	
((health NEAR/3 (national OR program? OR insurance)):ti) OR 'government programs':ti	216,521
#21	
'public health'/de	7,080
#20	
'health resources':ti,ab	2,106,575
#19	
#16 OR #17 OR #18	2,005,263
#18	
implement*:ti,ab OR disseminat*:ti,ab OR translation:ti,ab OR transfer:ti,ab OR incorporation:ti,ab OR integration:ti,ab	127,416
#17	
'health care planning'/exp OR 'translational research'/exp	55
#16	
('health plan' NEAR/3 implementation):ti,ab	

		645,628
#15		
#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14		
		47,430
#14		
(((individualized OR personalized OR precision OR predictive OR stratified) NEAR/3 medicine):ti,ab) OR 'p health':ti,ab OR theranostic:ti,ab OR theranostics:ti,ab		
		56,978
#13		
'personalized medicine'/de		
		25,311
#12		
pharmacogenomic*:ti,ab OR pharmacogenetic*:ti,ab		
		21,303
#11		
'pharmacogenetics'/de		
		192,357
#10		
(molecular:ti,ab OR genetic:ti,ab) AND techni*:ti,ab		
		6,874
#9		
'molecular diagnosis'/mj		
		60,752
#8		
(mass:ti,ab OR carrier:ti,ab) AND screening:ti,ab		
		9,439
#7		
(prenatal NEAR/3 (noninvasive OR screening)):ti,ab		
		411
#6		
(genomic? NEAR/3 (resource OR service? OR portfolio)):ti,ab		
		30,830
#5		
(genetic NEAR/3 (service OR counsel* OR screening?)):ti,ab		
		36,406
#4		
'genetic service'/de OR 'genetic counseling'/de		
		94,709
#3		
((gene OR genetic OR genom* OR multigene OR 'multi gene') NEAR/3 (test* OR panel*)):ti,ab		
		103,617
#2		
'genetic screening'/exp		
		120,990
#1		
'genomics'/exp		

Anexos

- Estrategia de búsqueda en la base de datos WOS

61

#60 AND #42 AND #39

[158](#)

60

**#59 OR #58 OR #57 OR #56 OR #55 OR #54 OR #53 OR #52 OR #51 OR #50 OR #49 OR
#48 OR #47 OR #46 OR #45 OR #44 OR #43**

[339,311](#)

59

TI=(Community NEAR/3 Service?) OR AB=(Community NEAR/3 Service?)

[20,156](#)

58

TI=("Health Services organization")

[22](#)

57

**TI= ("access care" or "access services" or "access therapy" or "access treatment") OR
AB= ("access care" or "access services" or "access therapy" or "access treatment")**

[2,147](#)

56

192

TI=("Health Services Accessibility") OR AB=("Health Services Accessibility")

[57](#)

55

TI=("Delivery Health" or "Delivery Healthcare") OR AB=("Delivery Health" or "Delivery Healthcare")

[224](#)

54

TI="Delivery of Health Care" OR AB="Delivery of Health Care"

[1,499](#)

53

TI=(Polic* Health" or "Polic* Healthcare") OR AB=(Polic* Health" or "Polic* Healthcare")

[4](#)

52

TI="Health Policy" OR AB="Health Policy"

[22,925](#)

51

TI=("national strategy " or "national plan") OR AB=("national strategy " or "national plan")

[4,184](#)

50

TI=(program NEAR/3 development or description?) OR AB=(program NEAR/3 development or description?)

[193,836](#)

49

TI="Program Development" OR AB="Program Development"

[5,238](#)

48

TI=(Health NEAR/3 Plan?) OR AB=(Health NEAR/3 Plan?)

[13,107](#)

47

TI="Health Systems Plans" OR AB="Health Systems Plans"

[4](#)

46

TI=(health NEAR/3 planning) or TI=(annual NEAR/3 plan?) or TI=(Appropriateness NEAR/3 Review) OR AB=(health NEAR/3 planning) or AB=(annual NEAR/3 plan?) or AB=(Appropriateness NEAR/3 Review)

[34,473](#)

45

194

TI=("Regional Health Planning") OR AB=("Regional Health Planning")

[45](#)

44

TI=("health national" or "health program?" or "health insurance" OR "GOVERNMENT PROGRAMS") OR AB=("health national" or "health program?" or "health insurance" OR "GOVERNMENT PROGRAMS")

[56,344](#)

43

TI=("Health Resources" OR "National Health Programs") OR AB=("Health Resources" OR "National Health Programs")

[5,285](#)

42

#41 OR #40

[1,208,620](#)

41

TI=(implementation or dissemination or translation or transfer or incorporation or integration) OR AK=(implementation or dissemination or translation or transfer or incorporation or integration)

[1,208,553](#)

40

Anexos

**TI=("Health Plan Implementation" or "Translational Medical Research") or
AB=("Health Plan Implementation" or "Translational Medical Research")**

[74](#)

39

#30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38

[1,090,964](#)

38

**TI=("individualized medicine" or "personalized medicine" or "precision medicine" or
"predictive medicine" or "stratified medicine")**

[10,328](#)

37

TI=("Precision Medicine") OR AB=("Precision Medicine")

[14,061](#)

36

**TI=(pharmacogenomic or pharmacogenetic) OR AB=(pharmacogenomic or
pharmacogenetic)**

[11,072](#)

35

TI=("molecular diagnostic techniques") OR AB=("molecular diagnostic techniques")

[474](#)

196

34

TI=("genomic resource" or "genomic service" or "genomic portfolio") OR
AB=("genomic resource" or "genomic service" or "genomic portfolio")

[622](#)

33

TI=(genetic NEAR/3 service or counsel* or screening?) OR AB=(genetic NEAR/3 service
or counsel* or screening?)

[161,271](#)

32

TI=(gene AND test* or panel*) OR TI=(genetic AND test* or panel*) OR TI=(genome
AND test* or panel*) OR TI=(multiline AND test* or panel*) OR TI=(multi-gene AND
test* or panel*) OR AB=(gene AND test* or panel*) OR AB=(genetic AND test* or
panel*) OR AB=(genome AND test* or panel*) OR AB=(multiline AND test* or panel*)
OR AB=(multi-gene AND test* or panel*)

[859,521](#)

31

TI=("Genetic Services" or Genetic Counseling") or AB=("Genetic Services" or Genetic
Counseling")

[364](#)

30

TI=(Genomics OR "Genetic Testing") OR AB=(Genomics OR "Genetic Testing")

Anexo 2: Resultados de la evaluación de la calidad de los estudios incluidos en la revisión sistemática para la generación de un marco de evaluación de la implementación de la MPP a nivel nacional.

Estudio incluido (autor y año)	Tipo de estudio	Herramienta de evaluación	Resultado de la calidad
Kovanda et al. 2021	RS	AMSTARII	Calidad baja
Pearce et al. 2019	RS	AMSTARII	Calidad moderada
Pitini, 2019	RS	AMSTARII	Calidad críticamente baja
Unim et al. 2019	RS	AMSTARII	Calidad baja
Bilkey et al. 2019	RN	SANRA	8/12
Burns et al. 2019	RN	SANRA	7/12
Chong et al. 2018	RN	SANRA	12/12
Doyle et al. 2018	RN	SANRA	10/12
Kurnat-Thoma, 2020	RN	SANRA	11/12
Lee & Kim, 2022	RN	SANRA	12/12
Pastorino et al. 2021	RN	SANRA	5/12
Vrijenhoek et al. 2021	RN	SANRA	6/12
Whitsel et al. 2019	RN	SANRA	5/12
Agarwal et al. 2021	Entrevista	Humphrey-Murto	8/11
Barash, 2016	Encuesta	Humphrey-Murto	4/11
Chanfrau-Coffinier et al. 2019	Entrevista	Humphrey-Murto	6/11
Delnord et al. 2021	Encuesta	Humphrey-Murto	8/11
Pitini, 2019	Panel de expertos	Humphrey-Murto	11/11
Pritchard et al. 2017	Debate en grupo y encuesta	Humphrey-Murto	6/11
Ricciardi et al. 2017	Debate en grupo	Humphrey-Murto	3/11

Anexo 3: Tabla detallada de los estudios incluidos de los que se ha extraído cada uno de los elementos evaluables que constituyen los dominios.

Políticas sanitarias	Estudios incluidos
Puesta en marcha de medidas legislativas (proyecto o iniciativa existente para alcanzar este objetivo)	Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Vrijenhoek et al.(Vrijenhoek et al. 2021)
Legislación/Regulación	Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Vrijenhoek et al.(Vrijenhoek et al. 2021), Agarwal et al.(Agarwal et al. 2021), Ricciardi et al.(Ricciardi y Boccia 2017)
Grupos de trabajo de colaboración entre las distintas partes interesadas	Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Agarwal et al.(Agarwal et al. 2021), Delnord et al.(Delnord et al. 2021)
Integración de la MPP en el sistema sanitario	Estudios incluidos
Cartera de servicios de MPP	Pearce et al.(Pearce et al. 2019), Unim et al.(Unim et al. 2019), Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Agarwal et al.(Agarwal et al. 2021)
Nivel de accesibilidad de la MPP en el sistema sanitario	Pearce et al.(Pearce et al. 2019), Burns et al.(Burns et al. 2019), Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Doyle et al.(Doyle et al. 2018)
Cartera de finalidades asistenciales	Pearce et al.(Pearce et al. 2019), Chong et al.(Chong, Allotey y Chaiyakunapruk 2018)
Investigación básica y traslacional	Estudios incluidos
Plan de fomento de la investigación básica en MPP y su traslación a la investigación clínica	Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Chanfreau-Coffinier et al.(Chanfreau-Coffinier et al. 2019), Delnord et al.(Delnord et al. 2021)
Infraestructuras y recursos	Estudios incluidos

Proyectos de recogida de información ómica a nivel poblacional	Doyle et al.(Doyle et al. 2018), Barash et al.(Isaacson Barash 2016)
Biobancos (muestras biológicas a nivel poblacional o de enfermedades específicas)	Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Doyle et al.(Doyle et al. 2018)
Plataformas de almacenamiento de historias clínicas electrónicas	Kovanda et al.(Kovanda, Zimani y Peterlin 2021)
Plataformas de almacenamiento de datos ómicos	Burns et al.(Burns et al. 2019), Barash et al.(Isaacson Barash 2016)
Plataformas de análisis de datos ómicos y big data	Doyle et al.(Doyle et al. 2018), Barash et al.(Isaacson Barash 2016)
Gestión y disponibilidad de datos	Estudios incluidos
Armonización, calidad y protección de las historias clínicas electrónicas	Pearce et al.(Pearce et al. 2019), Chanfreau-Coffinier et al.(Chanfreau-Coffinier et al. 2019)
Armonización, calidad y protección de los datos ómicos	Whitsel et al.(Whitsel et al. 2019), Barash et al.(Isaacson Barash 2016)
Incorporación de datos ómicos a la historia clínica electrónica	Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Chanfreau-Coffinier et al.(Chanfreau-Coffinier et al. 2019)
Acceso a los datos ómicos e interoperabilidad entre profesionales y entidades	Kovanda et al.(Kovanda, Zimani y Peterlin 2021), Bilkey et al.(Bilkey et al. 2019), Whitsel et al.(Whitsel et al. 2019)
Uso de datos ómicos en la toma de decisiones clínicas	Kovanda et al.(Kovanda, Zimani y Peterlin 2021), Bilkey et al.(Bilkey et al. 2019), Whitsel et al.(Whitsel et al. 2019)
Áreas organizacionales	Estudios incluidos
Interoperabilidad entre organizaciones y recursos de investigación básica y traslacional	Pearce et al.(Pearce et al. 2019), Agarwal et al.(Agarwal et al. 2021)
Introducción de áreas especializadas en MPP	Kovanda et al.(Kovanda, Zimani y Peterlin 2021), Kurnat-Thoma et al.(Kurnat-Thoma 2020)
Refuerzo de las áreas no especializadas en MPP	Burns et al.(Burns et al. 2019), Kurnat-Thoma et al.(Kurnat-Thoma 2020), Vrijenhoek et al.(Vrijenhoek et al. 2021), Barash et al.(Isaacson Barash 2016)
Desarrollo y adopción de directrices de procedimiento	Kovanda et al.(Kovanda, Zimani y Peterlin 2021),

	Vrijenhoek et al.(Vrijenhoek et al. 2021), Delnord et al.(Delnord et al. 2021)
• Estructura organizativa de los ensayos ómicos	Unim et al.(Unim et al. 2019), Pritchard et al.(Pritchard, Hulick y Wells 2021)
• Equipamiento de ensayos ómicos	Pritchard et al.(Pritchard, Hulick y Wells 2021)
Implicaciones éticas, legales y sociales (ELSI)	Estudios incluidos
• Nivel asistencial de información al paciente	Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Kurnat-Thoma et al.(Kurnat-Thoma 2020)
• Formularios normalizados de consentimiento informado del paciente y de aceptación y compromiso del paciente	Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Pastorino et al.(Pastorino et al. 2021), Barash et al.(Isaacson Barash 2016)
• Mecanismos de protección de datos	Bilkey et al.(Bilkey et al. 2019)
Necesidades educativas	Estudios incluidos
• Educación y formación del personal sanitario especializado y no especializado en MPP	Pastorino et al.(Pastorino et al. 2021), Barash et al.(Isaacson Barash 2016), Delnord et al.(Delnord et al. 2021)
• Educación de pacientes y familiares en MPP	Pastorino et al.(Pastorino et al. 2021), Barash et al.(Isaacson Barash 2016)
• Actividades de sensibilización y divulgación para la ciudadanía	Chong et al.(Chong, Allotey y Chaiyakunapruk 2018), Vrijenhoek et al.(Vrijenhoek et al. 2021)
Evaluación de tecnologías sanitarias (HTA)	Estudios incluidos
• Plan de HTA específico de MPP	Pitini et al.(Pitini et al. 2019), Chong et al.(Chong, Allotey y Chaiyakunapruk 2018)
• Órgano de HTA	Pitini et al.(Pitini et al. 2019), Chong et al.(Chong, Allotey y Chaiyakunapruk 2018)
• Metodología de HTA de MPP	Pitini et al.(Pitini et al. 2019), Burns et al.(Burns et al. 2019)
• Grupo de toma de decisiones de HTA	Pitini et al.(Pitini et al. 2019)
Evaluación de implementación	Estudios incluidos
• Plan de evaluación de la aplicación del MPP	Unim et al.(Unim et al. 2019), Burns et al.(Burns et al. 2019), Ricciardi et al.(Ricciardi y Boccia 2017)
• Órgano de evaluación de la aplicación	Unim et al.(Unim et al. 2019), Kurnat-Thoma et al.(Kurnat-Thoma 2020)

• Metodología de evaluación de la aplicación	Unim et al.(Unim et al. 2019), Delnord et al.(Delnord et al. 2021)
Financiación	Estudios incluidos
• Previsión de presupuesto de ejecución MPP	Burns et al.(Burns et al. 2019), Delnord et al.(Delnord et al. 2021)

Anexos

Anexo 4: Tabla con las puntuaciones para cada uno de los elementos evaluables y los resultados obtenidos en la evaluación de los países seleccionados

Dominios / Países evaluados	Finlandia	Reino Unido	Italia	Francia	China	Tailandia	Japón	Pakistán	Brasil	Estados Unidos	Canadá	Australia
Políticas sanitarias												
• Puesta en marcha de medidas legislativas (proyecto o iniciativa existente para alcanzar este objetivo)	2	2	1	1	1	0	2	0	0	2	2	2
• Legislación/Regulación	2	2	2	2	1	0	2	0	0	2	2	2
• Grupos de trabajo de colaboración entre las distintas partes interesadas	2	2	2	2	2	0	2	0	0	2	2	2
Integración de la MPP en el sistema sanitario												
• Cartera de servicios de MPP	2	2	1	2	1	0	2	0	1	0	2	2
• Nivel de accesibilidad de la MPP en el sistema sanitario	2	2	2	1	1	1	2	0	0	1	1	2
• Cartera de finalidades asistenciales	2	2	2	2	2	2	2	0	2	2	2	2

Investigación básica y traslacional													
• Plan de fomento de la investigación básica en MPP y su traslación a la investigación clínica	2	2	2	2	2	0	2	0	1	2	2	2	2
Infraestructuras y recursos													
• Proyectos de recogida de información ómica a nivel poblacional	2	2	2	2	1	2	2	0	0	2	2	2	2
• Biobancos (muestras biológicas a nivel poblacional o de enfermedades específicas)	2	2	2	1	1	1	1	0	1	2	2	2	2
• Plataformas de almacenamiento de historias clínicas electrónicas	2	2	1	2	2	1	1	0	0	2	2	2	2
• Plataformas de almacenamiento de datos ómicos	2	2	2	2	2	2	2	0	0	2	2	2	2
• Plataformas de análisis de datos ómicos y big data	2	2	2	2	2	2	2	0	0	2	2	2	2
Gestión y disponibilidad de datos													

Anexos

• Armonización, calidad y protección de las historias clínicas electrónicas	2	2	1	2	2	0	2	0	1	2	2	2
• Armonización, calidad y protección de los datos ómicos	2	2	2	2	2	0	2	0	1	2	2	2
• Incorporación de datos ómicos a la historia clínica electrónica	2	2	1	2	2	0	1	0	0	2	1	2
• Acceso a los datos ómicos e interoperabilidad entre profesionales y entidades	2	2	2	2	1	0	2	0	0	2	1	2
• Uso de datos ómicos en la toma de decisiones clínicas	2	2	0	2	1	0	1	0	0	2	2	2
Áreas organizacionales												
• Interoperabilidad entre organizaciones y recursos de investigación básica y traslacional	2	2	2	0	0	2	2	0	0	2	2	2
• Introducción de áreas especializadas en MPP	2	2	2	2	2	2	2	0	1	2	2	2

• Refuerzo de las áreas no especializadas en MPP	2	2	0	2	1	2	1	0	0	2	1	2
• Desarrollo y adopción de directrices de procedimiento	2	2	2	2	1	0	2	0	1	2	2	2
• Estructura organizativa de los ensayos ómicos	2	2	1	2	1	0	2	0	0	2	2	2
• Equipamiento de ensayos ómicos	2	2	2	2	2	0	2	0	0	2	2	2
Implicaciones éticas, legales y sociales (ELSI)												
• Nivel asistencial de información al paciente	2	2	2	0	1	0	2	0	0	2	2	2
• Formularios normalizados de consentimiento informado del paciente y de aceptación y compromiso del paciente	2	2	2	2	2	0	2	0	1	2	2	2
• Mecanismos de protección de datos	2	2	2	2	1	0	2	0	0	2	2	2
Necesidades educativas												
• Educación y formación del personal sanitario	2	2	2	2	2	2	2	0	2	2	2	2

Anexos

especializado y no especializado en MPP													
• Educación de pacientes y familiares en MPP	2	2	0	2	0	0	1	0	1	1	2	2	
• Actividades de sensibilización y divulgación para la ciudadanía	2	2	0	1	0	0	0	0	1	1	2	2	
Evaluación de tecnologías sanitarias (HTA)													
• Plan de HTA específico de MPP	2	0	0	2	0	0	1	0	0	0	2	0	
• Órgano de HTA	1	2	0	2	2	0	1	0	0	0	2	2	
• Metodología de HTA de MPP	2	2	1	0	0	0	0	0	0	0	2	2	
• Grupo de toma de decisiones de HTA	2	2	1	0	0	0	0	0	0	0	2	2	
Evaluación de implementación													
• Plan de evaluación de la aplicación del MPP	0	0	0	0	0	0	0	0	0	2	0	0	
• Órgano de evaluación de la aplicación	0	0	0	0	0	0	0	0	0	0	2	2	
• Metodología de evaluación de la aplicación	0	0	0	0	0	0	0	0	0	2	0	0	
Financiación													

Anexos

• Previsión de presupuesto de ejecución MPP	2	2	2	2	2	1	1	0	0	2	2	2
Puntuación total	19,75	19,5	15	16,8	13,53	5,27	15,53	0	4,6	17,67	19,77	20,16

Anexo 5: Estrategias de búsqueda de la revisión sistemática para el análisis de la efectividad de la biopsia líquida en el cáncer de pulmón

- Estrategia de búsqueda en la base de datos Medline

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily <1946 to August 24, 2022>

Search Strategy:

-
- 1 exp Lung Neoplasms/ (263995)
 - 2 ((lung* or pulmonary or bronch*) adj3 (cancer* or neoplasm* or carcinoma* or tumor* or lymphoma* or metast* or malignan* or blastoma* or carcinogen* or adenocarcinoma* or angiosarcoma* or chondrosarcoma* or sarcoma* or teratoma* or microcytic*)).ti,ab. (303730)
 - 3 ((pancoast* or superior sulcus or pulmonary sulcus) adj4 (tumor* or syndrome*)).tw. (883)
 - 4 ((lung* or pulmonary or bronch*) adj4 (oat or small or non-small) adj4 cell*).tw. (92285)
 - 5 (SCLC or NSCLC).tw. (61462)
 - 6 Carcinoma, Non-Small-Cell Lung/ (65950)
 - 7 (((Bronch* or Lung or Pulmon*) adj ("Non Small Cell?" or "Nonsmall Cell?") adj (Cancer* or Carcino*)) or ("Non Small Cell?" or "Nonsmall Cell?") adj (Bronchial or Lung or Pulmon*) adj (Cancer* or Carcino*)) or NSCLC*).ti,ab. (82502)
 - 8 1 or 2 or 3 or 4 or 5 or 6 or 7 (375207)
 - 9 Liquid Biopsy/ (2340)
 - 10 (Biops* adj3 (Liquid or fluid)).ti,ab. (6710)
 - 11 Neoplastic Cells, Circulating/ (11948)
 - 12 ("circulating tumor cell" or CTCs).ti,ab. (5815)
 - 13 exp Cell-Free Nucleic Acids/ (6209)
 - 14 ('cell-free tumor DNA' or cfDNA).ti,ab. (3246)
 - 15 Circulating Tumor DNA/ (2180)

Anexos

- 16 (circulating tumour DNA or ctDNA).ti,ab. (3407)
- 17 ('circulating tumour ribonucleic acid' or ctRNA).ti,ab. (34)
- 18 Exosomes/ or exosomes.ti,ab. (19132)
- 19 ('Archer Reveal ctDNA 28' or BDX-XL217 or 'BDX-XL2' or CancerSEEK or CellMax or 'EarlyCDT-Lung' or 'Epi proLung' or GRAIL or IvyGene or Lunar-1 or Lunar-2).ti,ab. (2165)
- 20 Biomarkers, Tumor/bl, ge [Blood, Genetics] (73746)
- 21 (('tumor marker' adj3 (biochemical or biologic*)) or "biolarker tum?r?").ti,ab. (72)
- 22 RNA, Messenger/ (388510)
- 23 (rna or mrna).ti,ab. (1009827)
- 24 MicroRNAs/bl, ge, me [Blood, Genetics, Metabolism] (101150)
- 25 (((micro or small) adj3 rna) or microrna? or mirna or strna).ti,ab. (164148)
- 26 RNA, Long Noncoding/bl, ge, me [Blood, Genetics, Metabolism] (25010)
- 27 (lnc-rna or lncrna or 'linc rna' or lincrna? or ('long non' adj3 'coding rna') or 'long ncrna?').ti,ab. (30003)
- 28 Blood Platelets/ (81304)
- 29 (platelet? or thrombocyt*).ti,ab. (293144)
- 30 "tumor-educated platelets".ti,ab. (74)
- 31 Endothelial Cells/ (56733)
- 32 (endothelial adj3 cell?).ti,ab. (194597)
- 33 ('circulating tumor endothelial cells' or CTECs).ti,ab. (108)
- 34 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 (1742437)
- 35 8 and 34 (45017)
- 36 (letter or "case report*" or "historical article*" or (comment or editorial or in vitro or news)).pt. (4684508)
- 37 35 not 36 (42766)
- 38 Systematic review/ or Meta-Analysis as Topic/ or meta analy*.tw. or metaanaly*.tw. or Meta-Analysis/ or (systematic* adj (review* or overview*)).tw. or exp Review Literature as Topic/ or cochrane.ab. or embase.ab. or psychlit.ab. or psyclit.ab. or psychinfo.ab. or psycinfo.ab. or cinahl.ab. or "science citation index".ab. or bids.ab. or cancerlit.ab. or "reference list".ab. or bibliography*.ab. or hand-search*.ab. or "relevant journals".ab. or (manual adj1 search*).ab. or "selection criteria".ab. or "study

selection".ab. or "data extraction".ab. or "data sources".ab. or (search adj1 strateg*).ab.
(535307)

39 37 and 38 (852)

40 animals/ not (animals/ and humans/) (5005636)

41 39 not 40 (849)

- Estrategia de búsqueda en la base de datos EMBASE

296

#46

#45 AND [embase]/lim NOT ([embase]/lim AND [medline]/lim) AND [humans]/lim

983

#45

#43 AND #44

675,618

#44

'systematic review'/exp OR 'meta analysis'/exp OR ((meta NEXT/3 analy*):ti,ab) OR
metaanalys*:ti,ab OR ((systematic NEXT/3 (review* OR overview*)):ti,ab) OR
cancerlit:ab OR cochrane:ab OR embase:ab OR psychlit:ab OR psyclit:ab OR psychinfo:ab
OR psycinfo:ab OR cinahl:ab OR cinhal:ab OR 'science citation index':ab OR bids:ab OR
'reference lists':ab OR bibliograph*:ab OR 'hand-search*':ab OR 'manual search*':ab OR
'relevant journals':ab OR (('data extraction':ab OR 'selection criteria':ab) AND review:pt)

44,182

#43

#41 NOT #42

7,935,942

213

Anexos

#42

'conference abstract'/it OR 'conference paper'/it OR 'note'/it OR 'short survey'/it OR
'abstract report'/exp OR 'abstract report' OR 'letter'/exp OR 'letter'

60,913

#41

#8 AND #40

2,033,155

#40

#9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #22
OR #23 OR #24 OR #25 OR #28 OR #29 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37
OR #38 OR #39

143

#39

'circulating tumor endothelial cells':ti,ab OR ctecs:ti,ab

215,657

#38

(endothelial NEAR/3 cell?):ti,ab

175,278

#37

'endothelium cell'/de

103

#36

'tumor-educated platelets':ti,ab

248,563

214

#35

platelet?:ti,ab OR thrombocyt*:ti,ab

124,649

#34

'thrombocyte'/de

33,845

#33

'lnc rna':ti,ab OR lncrna:ti,ab OR 'linc rna':ti,ab OR lincrna?:ti,ab OR (('long non' NEAR/3
'coding rna'):ti,ab) OR 'long ncrna?':ti,ab

2,507

#32

#30 AND #31

2,024,673

#31

'blood'/de OR 'genetic'/de OR 'metabolism'/de

45,047

#30

'long untranslated rna'/de

175,627

#29

((micro OR small) NEAR/3 rna):ti,ab) OR microrna?:ti,ab OR mirna:ti,ab OR strna:ti,ab

14,820

#28

#26 AND #27

215

Anexos

2,024,673

#27

'blood'/de OR 'genetic'/de OR 'metabolism'/de

175,051

#26

'microrna'/de

1,253,379

#25

rna:ti,ab OR mrna:ti,ab

632,994

#24

'messenger rna'/de

105

#23

((('tumor marker' NEAR/3 (biochemical OR biologic*)):ti,ab) OR 'biolarker tum?r?':ti,ab

6,939

#22

#20 AND #21

704,523

#21

'blood'/de OR 'genetic'/de

94,163

#20

216

'tumor marker'/de

2,851

#19

'archer reveal ctdna 28':ti,ab OR 'bdx-xl2':ti,ab OR cancerseek:ti,ab OR cellmax:ti,ab OR 'earlycdt-lung':ti,ab OR 'epi prolong':ti,ab OR grail:ti,ab OR ivygene:ti,ab OR 'lunar 1':ti,ab OR 'lunar 2':ti,ab

29

#18

'bdx xl217':ti,ab OR cancerseek:ti,ab OR 'epi prolong':ti,ab OR ivygene:ti,ab OR 'lunar 1':ti,ab

47,089

#17

'exosome'/de OR exosome?:ti,ab

9,319

#16

circulating:ti,ab AND tumour:ti,ab AND dna:ti,ab OR ctdna:ti,ab

7,236

#15

'circulating tumor dna'/de

7,430

#14

'cell-free tumor dna':ti,ab OR cfdna:ti,ab

1,634

#13

217

Anexos

'cell free nucleic acid'/de

10,107

#12

'circulating tumor cell':ti,ab OR ctcs:ti,ab

4,225

#11

'tumor embolism'/de

11,329

#10

(biops* NEAR/3 (liquid OR fluid)):ti,ab

8,564

#9

'liquid biopsy'/exp

588,688

#8

#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7

101

#7

((bronch* OR lung OR pulmon*) NEAR/1 ('non small cell?' OR 'nonsmall cell?') NEAR/1 (cancer* OR carcino*)):ti,ab) OR (('non small cell?' OR 'nonsmall cell?') NEAR/1 (bronchial OR lung OR pulmon*) NEAR/1 (cancer* OR carcino*)):ti,ab)

147,264

#6

'non small cell lung cancer'/de

218

109,563

#5

sclc:ti,ab OR nsclc:ti,ab

144,518

#4

((lung* OR pulmonary OR bronch*) NEAR/4 (oat OR small OR 'non small') NEAR/4 cell*):ti,ab

490

#3

((pancoast* OR 'superior sulcus' OR 'pulmonary sulcus') NEAR/4 (tumo?r* OR syndrome*)):ti,ab

426,547

#2

((lung* OR pulmonary OR bronch*) NEAR/3 (cancer* OR neoplasm* OR carcinoma* OR tumor* OR lymphoma* OR metast* OR malignan* OR blastoma* OR carcinogen* OR adenocarcinoma* OR angiosarcoma* OR chondrosarcoma* OR sarcoma* OR teratoma* OR microcytic*)):ti,ab

511,082

#1

'lung tumor'/exp

- Estrategia de búsqueda en la base de datos WOS

1	(TI="Lung Neoplasm*") OR (AB="Lung Neoplasm*") OR TI=("Carcinoma Non-Small-Cell Lung") OR AB=("Carcinoma Non-Small-Cell Lung") Editions: WOS.SCI Results: 715
---	--

2	TI=(lung* or pulmonary or bronch* NEAR/3 cancer* or neoplasm* or carcinoma* or tumor* or lymphoma* or metast* or malignan* or blastoma* or carcinogen* or adenocarcinoma* or angiosarcoma* or chondrosarcoma* or sarcoma* or teratoma* or microcytic*) OR AB=(lung* or pulmonary or bronch* NEAR/3 cancer* or neoplasm* or carcinoma* or tumor* or lymphoma* or metast* or malignan* or blastoma* or carcinogen* or adenocarcinoma* or angiosarcoma* or chondrosarcoma* or sarcoma* or teratoma* or microcytic*) Editions: WOS.SCI Results: 3826219
3	#2 OR #1 Editions: WOS.SCI Results: 3826219
4	TI=("liquid biopsy" OR "circulating tumor cell" OR "circulating tumor cluster" OR "exosome" OR "ctDNA" OR "circulating tumor DNA" OR "cell free DNA" OR "cfDNA" OR "liquid biopsies" OR "circulating tumor cells" OR "circulating tumor clusters" OR "exosomes" OR "exosomal" OR "microparticle*" OR "circulating endothelial cell*" OR "biomarker tumor*")OR AB=("liquid biopsy" OR "circulating tumor cell" OR "circulating tumor cluster" OR "exosome" OR "ctDNA" OR "circulating tumor DNA" OR "cell free DNA" OR "cfDNA" OR "liquid biopsies" OR "circulating tumor cells" OR "circulating tumor clusters" OR "exosomes" OR "exosomal" OR "microparticle*" OR "circulating endothelial cell*" OR "biomarker tumor*") Editions: WOS.SCI Results: 80983
5	#4 AND #3 Editions: WOS.SCI Results: 44214
6	TI=("Systematic review") OR AB=("Systematic review") Editions: WOS.SCI Results: 227058
7	#5 AND #6 Editions: WOS.SCI Results: 223
8	#5 AND #6 and Meeting Abstract or Early Access or Editorial Material or Correction (Exclude – Document Types) Editions: WOS.SCI Results: 207

Anexo 6: Tablas detalladas de la información extraída de las revisiones sistemáticas incluidas para el análisis de la efectividad de la biopsia líquida en el cáncer de pulmón

- Mutaciones puntuales en el gen *EGFR* en ADN circulante tumoral o en ADN de células tumorales circulantes para la adecuación del tratamiento

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Wang et al. 2021)	5995 pacientes NSCLC (40 estudios)	Mutaciones en <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 68% (95% CI=60-75%) Especificidad: 98% (95% CI=95-99%)
(Zhou, Huang y Cao 2020)	4527 pacientes NSCLC (32 estudios)	Mutaciones en <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 70% (95% CI=0,63-0,75) Especificidad: 98% (95% CI=0,96-0,99)
(Passiglia et al. 2018)	1639 pacientes NSCLC (19 estudios)	Mutación en <i>EGFR</i> -T790M en ADN circulante tumoral de suero o plasma	Sensibilidad: 67% (95% CI=0,64-0,70) Especificidad: 80% (95% CI=0,77-0,83)
(Lyu et al. 2019)	NA (47 estudios)	Mutaciones en <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 67,1% (95% CI=0,647-0,695) Especificidad: 96,1% (95% CI=0,954-0,968)
(Ma et al. 2017)	2094 pacientes NSCLC (15 estudios)	Mutaciones en <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 69% (95% CI=0,59-0,78) Especificidad: 97% (95% CI=0,94-0,99)
(Biaoxue y Shuanying 2018)	2922 pacientes NSCLC (19 estudios)	Mutaciones en <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 65% (95% CI=0,61-0,68) Especificidad: 91% (95% CI=0,89-0,92)
(Qian et al. 2016)	3938 pacientes NSCLC (33 estudios)	Deleción del exón 19 o mutación en L858R de <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 60% (95% CI=0,57-0,62) Especificidad: 94% (95% CI=0,93-0,95)

(Luo, Shen y Zheng 2014)	2012 pacientes NSCLC (20 estudios)	Mutaciones en <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 67,4% (95% CI=0,517-0,800) Especificidad: 93,5% (95% CI=0,888-0,963)
(Qiu et al. 2015)	3110 pacientes NSCLC (27 estudios)	Mutaciones en <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 62% (95% CI=0,513-0,716) Especificidad: 95,9% (95% CI=0,929-0,977)
(Li, Zhijun et al. 2014)	1591 pacientes NSCLC (13 estudios)	Mutaciones en <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 65% (95% CI=61-68%) Especificidad: 88% (95% CI=86-90%)
(Hu et al. 2018)	255 pacientes NSCLC (8 estudios)	Mutaciones en <i>EGFR</i> en células circulantes tumorales de suero o plasma	Sensibilidad: 82% (95% CI=0,50-0,95) Especificidad: 95% (95% CI=0,24-1,00)
(Liu et al. 2016)	170 pacientes NSCLC (8 estudios)	Mutaciones en <i>EGFR</i> en células circulantes tumorales de suero o plasma	Sensibilidad: 91% (95% CI=0,55-0,99) Especificidad: 99% (95% CI=0,59-1,00)

- Mutaciones puntuales en el gen *KRAS* en ADN circulante tumoral para la adecuación del tratamiento

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Lyu et al. 2019)	NA (47 estudios)	Mutaciones en <i>EGFR</i> en ADN circulante tumoral de suero o plasma	Sensibilidad: 65,1% (95% CI=0,558-0,736) Especificidad: 95,5% (95% CI=0,932-0,972)

- Mutaciones puntuales en el gen *BRAF* en ADN circulante tumoral para la adecuación del tratamiento

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Lyu et al. 2019)	NA (47 estudios)	Mutaciones en EGFR en ADN circulante tumoral de suero o plasma	Sensibilidad: 31,3% (95% CI=0,141-0,532) Especificidad: 99,5% (95% CI=0,978-1,000)

- Alteraciones en los patrones de metilación del ADN circulante tumoral para el diagnóstico

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Liu et al. 2021)	2006 incluidos (1014 pacientes/992 controles) (14 estudios)	Metilación del promotor del gen APC en suero, esputo o fluido del lavado bronquioalveolar	Sensibilidad: 0,43 (95% CI=0,40-0,45) Especificidad: 0,92 (95% CI=0,90-0,94)
(Chen, Ying y Bao 2021)	1192 incluidos (629 pacientes/563 controles) (12 estudios)	Metilación del promotor del gen <i>MGMT</i> en plasma o fluido del lavado bronquioalveolar	Sensibilidad: 0,39 (95% CI=0,31-0,49) Especificidad: 0,92 (95% CI=0,77-0,97)
(Zhang et al. 2019)	1424 incluidos (826 pacientes/598 controles) (12 estudios)	Hipermetilación del promotor del gen <i>RASSF1A</i> en suero o esputo	Sensibilidad: 0,45 (95% CI=0,41-0,48) Especificidad: 0,99 (95% CI=0,98-1,00)
(Tuo et al. 2018)	3095 incluidos (1772 pacientes/1323 controles) (26 estudios)	Metilación del promotor del gen <i>P16INK4a</i> en suero, esputo o fluido del lavado bronquioalveolar	Sensibilidad: 0,46 (95% CI=0,43-0,48) Especificidad: 0,90 (95% CI=0,88-0,91)
(Wang, Xu et al. 2014)	2048 incluidos (1208 pacientes/1200 controles) (22 estudios)	Patrones de metilación aberrantes en diversas regiones del ADN en esputo	Sensibilidad: 0,62 (95% CI=0,59-0,65) Especificidad: 0,73 (95% CI=0,70-0,75)

- Alteraciones en los patrones de expresión de ARNlnc para el diagnóstico

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Pan et al. 2020)	1281 incluidos (677 pacientes / 604 controles) (6 estudios)	Análisis cuantitativo de la expresión de <i>MALAT1</i> circulante en suero o plasma	Sensibilidad: 0,81 (95% CI=0,78-0,84) Especificidad: 0,67 (95% CI=0,63-0,71)
(Dai, Jin y Li 2018)	5642 incluidos (3044 pacientes / 2598 controles) (28 estudios)	Análisis cuantitativo de la expresión de 20 ARNlnc en suero o plasma	Sensibilidad: 0,82 (95% CI=0,79-0,84) Especificidad: 0,82 (95% CI=0,78-0,84)
(Chen et al. 2017)	824 incluidos (450 pacientes / 374 controles) (14 estudios)	Análisis cuantitativo de la expresión de <i>MALAT1</i> circulante en suero o plasma	Sensibilidad: 0,70 (95% CI=0,60-0,81) Especificidad: 0,92 (95% CI=0,85-0,96)
(Yang, Hongyun et al. 2017)	1405 incluidos (775 pacientes / 630 controles) (8 estudios)	Análisis cuantitativo de la expresión de más de 60 ARNlnc	Sensibilidad: 0,87 (95% CI=0,83-0,90) Especificidad: 0,74 (95% CI=0,64-0,98)

- Alteraciones en los patrones de expresión de ARNcirc para el diagnóstico

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Yang, Xianglin et al. 2021)	1525 incluidos (774 pacientes / 751 controles) (18 estudios)	Análisis cuantitativo de la expresión de 18 circrna en plasma o suero	Sensibilidad: 0,76 (95% CI=0,71-0,80) Especificidad: 0,77 (95% CI=0,73-0,81)
(Yang, Qian et al. 2021)	1968 incluidos (1045 pacientes / 923 controles) (17 estudios)	Análisis cuantitativo de la expresión de 15 circrna en plasma o suero	Sensibilidad: 0,77 (95% CI=0,73-0,81) Especificidad: 0,75 (95% CI=0,71-0,79)
(Xiao et al. 2019)	821 incluidos (668 pacientes / 153 controles) (10 estudios)	Análisis cuantitativo de la expresión de 9 circrna en plasma o suero	Sensibilidad: 0,77 (95% CI=0,73-0,80) Especificidad: 0,76 (95% CI=0,69-0,82)

- Detección de células tumorales circulantes para el diagnóstico

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Zhao et al. 2021)	3997 incluidos (2714 pacientes / 1283 controles) (21 estudios)	Análisis de la presencia de células circulantes tumorales en sangre	Sensibilidad: 0,72 (95% CI=0,65-0,79) Especificidad: 0,96 (95% CI=0,91-0,98)
(Ye et al. 2019)	1601 incluidos (1093 pacientes / 508 controles) (8 estudios)	Análisis de la presencia de células circulantes tumorales en sangre	Sensibilidad: 0,75 (95% CI=0,73-0,78) Especificidad: 0,89 (95% CI=0,86-0,92)
(Huang et al. 2018)	699 incluidos (460 pacientes / 239 controles) (5 estudios)	Análisis de la presencia de células circulantes tumorales en sangre	Sensibilidad: 0,75 (95% CI=0,54-0,88) Especificidad: 0,92 (95% CI=0,82-0,97)
(Zeng et al. 2017)	2298 incluidos (1402 pacientes / 896 controles) (15 estudios)	Análisis de la presencia de células circulantes tumorales en sangre	Sensibilidad: 0,63 (95% CI=0,60-0,65) Especificidad: 0,94 (95% CI=0,92-0,95)

- Cuantificación de ADN circulante para el diagnóstico

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(He et al. 2022)	1253 incluidos (618 pacientes / 635 controles) (8 estudios)	Cuantificación de ADN circulante en plasma o suero	Sensibilidad: 0,79 (95% CI=0,75-0,82) Especificidad: 0,81 (95% CI=0,78-0,84)
(Jiang et al. 2016)	2125 incluidos (1193 pacientes / 1059 controles) (15 estudios)	Cuantificación de ADN circulante en plasma o suero	Sensibilidad: 0,81 (95% CI=0,74-0,86) Especificidad: 0,85 (95% CI=0,77-0,91)
(Zhang et al. 2010)	1387 incluidos (752 pacientes / 635 controles) (10 estudios)	Cuantificación de ADN circulante en plasma o suero	Sensibilidad: 0,80 (95% CI=0,77-0,83) Especificidad: 0,77 (95% CI=0,74-0,80)

Anexos

- Mutaciones puntuales en la secuencia de ADN tumoral circulante para el diagnóstico

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Sebastião et al. 2020)	2116 pacientes (19 estudios)	Análisis NGS de mutaciones puntuales en los genes <i>EGFR</i> , <i>ALK</i> , <i>ROS-1</i> , <i>BRAF</i> , <i>RET</i> y <i>MET</i> en plasma	Sensibilidad: 0,77 (95% CI=0,678-0,835) Especificidad: 0,99 (95% CI=0,990-1,000)

- Alteraciones en los patrones de expresión de miARN para el diagnóstico

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Wang, Xu y Ding 2022)	2238 incluidos (1109 pacientes / 1129 controles) (17 estudios)	Nivel de expresión de <i>miRNA-21</i> en suero o esputo	Sensibilidad: 0,73 (95% CI=0,67-0,78) Especificidad: 0,81 (95% CI=0,75-0,85)
(Wang et al. 2022)	3707 incluidos (2124 pacientes / 1583 controles) (31 estudios)	Nivel de expresión de <i>miRNA-21</i> en plasma, suero o esputo	Sensibilidad: 0,77 (95% CI=0,72-0,81) Especificidad: 0,86 (95% CI=0,80-0,90)
(Wu et al. 2014)	1205 incluidos (676 pacientes / 529 controles) (11 estudios)	Nivel de expresión de <i>miRNA-21</i> en sangre o esputo	Sensibilidad: 0,66 (95% CI=0,57-0,74) Especificidad: 0,82 (95% CI=0,74-0,88)
(Yang et al. 2014)	868 incluidos (500 pacientes / 368 controles) (7 estudios)	Nivel de expresión de <i>miRNA-21</i> en sangre	Sensibilidad: 0,71 (95% CI=0,57-0,82) Especificidad: 0,84 (95% CI=0,76-0,89)
(Li et al. 2020)	931 incluidos (480 pacientes / 451 controles) (6 estudios)	Nivel de expresión de <i>miRNA-25</i> en suero	Sensibilidad: 0,75 (95% CI=0,69-0,80) Especificidad: 0,81 (95% CI=0,76-0,86)
(Sun et al. 2020)	956 incluidos (493 pacientes / 463 controles) (5 estudios)	Nivel de expresión de <i>miRNA-126</i> en suero, plasma y esputo	Sensibilidad: 0,83 (95% CI=0,59-0,94) Especificidad: 0,83 (95% CI=0,71-0,90)

(Tao et al. 2020)	1394 incluidos (9 estudios)	Nivel de expresión de <i>miRNA-145</i> en suero y plasma	Sensibilidad: 0,78 (95% CI=0,7-0,83) Especificidad: 0,75 (95% CI=0,67-0,82)
(Cheng et al. 2021)	2347 incluidos (831 pacientes / 1516 controles) (5 estudios)	Nivel de expresión de <i>miRNA-148/152</i> en suero y plasma	Sensibilidad: 0,79 (95% CI=0,74-0,83) Especificidad: 0,81 (95% CI=0,76-0,85)
(Shao et al. 2019)	799 incluidos (457 pacientes / 342 controles) (8 estudios)	Nivel de expresión de <i>miRNA-155</i> en suero o plasma	Sensibilidad: 0,82 (95% CI=0,78-0,88) Especificidad: 0,78 (95% CI=0,71-0,84)
(Li et al. 2017)	1231 incluidos (564 pacientes / 667 controles) (12 estudios)	Nivel de expresión de <i>miRNA-155</i> en sangre	Sensibilidad: 0,88 (95% CI=0,78-0,94) Especificidad: 0,78 (95% CI=0,66-0,86)
(Hu et al. 2020)	1604 incluidos (858 pacientes / 746 controles) (14 estudios)	Nivel de expresión de <i>miRNA-210</i> en suero, plasma o esputo	Sensibilidad: 0,66 (95% CI=0,59-0,73) Especificidad: 0,79 (95% CI=0,74-0,84)
(Yang et al. 2016)	993 incluidos (554 pacientes / 439 controles) (9 estudios)	Nivel de expresión de <i>miRNA-210</i> en suero, plasma o esputo	Sensibilidad: 0,66 (95% CI=0,57-0,75) Especificidad: 0,82 (95% CI=0,72-0,89)
(Shen et al. 2022)	2011 incluidos (1132 pacientes / 879 controles) (19 estudios)	Nivel de expresión de <i>miRNA-486</i> en suero, plasma o esputo	Sensibilidad: 0,8 (95% CI=0,8-0,9) Especificidad: 0,9 (95% CI=0,9-0,9)
(Li et al. 2021)	(119 estudios)	Nivel de expresión de diferentes <i>miRNA</i> en plasma, suero o esputo	Sensibilidad: 0,76 (95% CI=0,73-0,78) Especificidad: 0,80 (95% CI=0,77-0,81)
(Huang et al. 2021)	1956 incluidos (16 estudios)	Nivel de expresión de diferentes <i>miRNA</i> en plasma, suero o efusión pleural	Sensibilidad: 0,77 (95% CI=0,72-0,80) Especificidad: 0,83 (95% CI=0,78-0,86)
(Hu et al. 2019)	1927 incluidos (973 pacientes / 954 controles) (16 estudios)	Nivel de expresión en <i>miRNA-210</i> y otros <i>miRNA</i> en plasma o esputo	Sensibilidad: 0,76 (95% CI=0,72-0,79) Especificidad: 0,88 (95% CI=0,86-0,90)
(Wang, Guan y Wang 2019)	3954 incluidos (2487 pacientes / 1467 controles) (29 estudios)	Niveles de expresión de <i>miRNA</i> en sangre o fluido del lavado bronquioalveolar	Sensibilidad: 0,77 (95% CI=0,74-0,79) Especificidad: 0,83 (95% CI=0,81-0,85)
(Zhang, Wang y Zhang 2019)	2015 incluidos (1009 pacientes / 1006	Niveles de expresión de <i>miRNA</i> en esputo	Sensibilidad: 0,75 (95% CI=0,72-0,78)

	controles) (14 estudios)		Especificidad: 0,88 (95% CI=0,86-0,90)
(Yu et al. 2018)	NA (17 estudios)	Niveles de expresión de 35 miRNA en suero o plasma	Sensibilidad: 0,78 (95% CI=0,52-1) Especificidad: 0,79 (95% CI=0,43-0,93)
(Jiang et al. 2018)	9099 incluidos (5111 pacientes / 3988 controles) (71 estudios)	Niveles de expresión de múltiples miRNA en suero, plasma, esputo o fluido del lavado bronquioalveolar	Sensibilidad: 0,85 (95% CI=0,82-0,88) Especificidad: 0,88 (95% CI=0,85-0,9)
(Yang, Yanlong et al. 2017)	13983 incluidos (6919 pacientes / 7064 controles) (65 estudios)	Niveles de expresión de múltiples miRNA en suero o plasma	Sensibilidad: 0,83 (95% CI=0,80-0,85) Especificidad: 0,84 (95% CI=0,82-0,86)
(He et al. 2015)	975 incluidos (510 pacientes / 465 controles) (13 estudios)	Niveles de expresión de múltiples miRNA en suero, plasma o esputo	Sensibilidad: 0,797 (95% CI=0,756-0,832) Especificidad: NA
(Chen y Jin 2014)	2623 incluidos (1563 pacientes / 1060 controles) (20 estudios)	Niveles de expresión de múltiples miRNA en suero, plasma o esputo	Sensibilidad: 0,76 (95% CI=0,72-0,80) Especificidad: 0,80 (95% CI=0,77-0,84)
(Guo, Zhao y Wang 2014)	3802 incluidos (2142 pacientes / 1660 controles) (33 estudios)	Niveles de expresión de múltiples miRNA en suero, plasma o esputo	Sensibilidad: 0,72 (95% CI=0,68-0,75) Especificidad: 0,74 (95% CI=0,70-0,77)
(Liao et al. 2014)	1005 incluidos (514 pacientes / 491 controles) (8 estudios)	Niveles de expresión de múltiples miRNA en esputo	Sensibilidad: 0,70 (95% CI=0,66-0,70) Especificidad: 0,89 (95% CI=0,86-0,91)
(Shen et al. 2013)	2066 incluidos (1187 pacientes / 879 controles) (18 estudios)	Niveles de expresión de múltiples miRNA en suero o plasma	Sensibilidad: 0,85 (95% CI=0,83-0,87) Especificidad: 0,84 (95% CI=0,81-0,86)

- Análisis del contenido de exosomas para el diagnóstico

Referencia RS incluida	Población	Alteraciones ómicas analizadas y material biológico para BL	Resultados BL frente biopsia tejido
(Song, Wang y Liu 2019)	2413 incluidos (1338 pacientes / 1075 controles) (13 estudios)	Análisis del contenido de exosomas (lípidos,	Sensibilidad: 0,87 (95% CI=0,83-0,90)

		proteínas o miRNA) en suero o plasma	Especificidad: 0,86 (95% CI=0,81-0,90)
--	--	---	---

Anexos

Anexo 7: Datos extraídos de los estudios clínicos incluidos para el MA de la eficacia de la detección de mutaciones de *EGFR* en BL para la adecuación del tratamiento

Referencia estudio	TP	FP	FN	TN	Tipo de muestra	Pretratamiento o postratamiento
(Denis et al. 2019)	9	0	5	112	Plasma	Pretratamiento
(Ding et al. 2019)	11	0	5	10	Plasma	Pretratamiento
(Hong et al. 2019)	11	13	22	29	Plasma	Postratamiento
(Leighl et al. 2019)	18	2	4	201	Plasma	Pretratamiento
(Xu et al. 2019)	63	1	56	83	Plasma	Pretratamiento
(Li et al. 2019)	29	0	8	90	Plasma	Pretratamiento
(Ito et al. 2018)	8	0	26	117	Plasma	Pretratamiento
(Kobayashi et al. 2018)	7	0	0	3	Plasma	Postratamiento
(Liu et al. 2018)	102	8	14	65	Efusión pleural	Pretratamiento
(Veldore et al. 2018)	41	0	4	87	Plasma	Pretratamiento
(Shi et al. 2018)	27	5	11	12	Plasma	Pretratamiento
(Arriola et al. 2018)	17	4	5	121	Plasma	Pretratamiento
(Buder et al. 2018)	28	9	6	2	Plasma	Postratamiento
(Wang et al. 2017)	32	8	30	47	Plasma	Postratamiento
(Zhang et al. 2017)	57	4	36	118	Plasma	Postratamiento
(Suzawa et al. 2017)	9	1	12	37	Plasma	Postratamiento
(Jenkins et al. 2017)	255	27	161	100	Plasma	Postratamiento
(Mellert et al. 2017)	13	0	2	40	Plasma	Pretratamiento
(Kasahara et al. 2017)	5	6	2	7	Plasma	Pretratamiento
(Yoshida et al. 2017)	4	0	6	11	Plasma	Postratamiento
(Wu et al. 2017)	7	2	10	5	Plasma	Postratamiento
(Gu et al. 2017)	23	4	5	79	Plasma	Pretratamiento
(Guo et al. 2017)	8	0	4	8	Plasma	Pretratamiento
(He et al. 2017)	83	0	23	14	Plasma	Postratamiento
(Ma et al. 2017)	47	3	34	73	Plasma	Pretratamiento
(Schwaederlé et al. 2017)	7	2	6	19	Plasma	Pretratamiento
(Yao et al. 2017)	12	0	5	22	Plasma	Pretratamiento
(Zhou et al. 2017)	5	1	3	20	Plasma	Pretratamiento
(Chai et al. 2016)	25	6	9	21	Plasma	Pretratamiento
(Karlovich et al. 2016)	55	0	20	24	Plasma	Pretratamiento
(Ma et al. 2016)	54	4	36	125	Plasma	Pretratamiento
(Sundaresan et al. 2016)	9	4	6	6	Plasma	Postratamiento
(Vázquez et al. 2016)	13	1	12	148	Suero	Pretratamiento
(Sacher et al. 2016)	41	0	9	124	Plasma	Pretratamiento

(Que et al. 2016)	33	6	7	58	Plasma	Pretratamiento
(Oxnard et al. 2016)	111	18	47	40	Plasma	Postratamiento
(Reckamp et al. 2016)	38	4	3	60	Plasma	Postratamiento
(Takahama et al. 2016)	20	3	11	7	Plasma	Postratamiento
(Paweletz et al. 2016)	8	2	2	2	Plasma	Postratamiento
(Seki et al. 2016)	5	0	2	3	Plasma	Postratamiento
(Thompson et al. 2016)	19	0	5	26	Plasma	Postratamiento
(Han et al. 2016)	34	18	17	125	Plasma	NA
(Que et al. 2016)	187	0	13	0	Plasma	Pretratamiento
(Newman et al. 2016)	22	0	2	33	Plasma	Pretratamiento
(Pécuchet et al. 2016)	28	0	19	62	Plasma	Postratamiento
(Rachiglio et al. 2016)	17	2	5	20	Plasma	Pretratamiento
(Reck et al. 2016)	87	25	102	948	Plasma	Pretratamiento
(He, Tan y Ma 2017)	83	0	23	14	Plasma	Pretratamiento
(Duan et al. 2015)	19	0	19	56	Plasma	Pretratamiento
(Ishii et al. 2015)	10	0	5	2	Plasma	Postratamiento
(Lam et al. 2015)	34	9	1	30	Plasma	Pretratamiento
(Thress et al. 2015)	23	1	5	30	Plasma	Pretratamiento
(Zhu et al. 2015)	18	1	4	63	Plasma	Pretratamiento
(Pasquale et al. 2015)	16	0	10	70	Plasma	Pretratamiento
(Uchida et al. 2015)	56	22	47	163	Plasma	Pretratamiento
(Douillard et al. 2014)	69	1	36	546	Plasma	Pretratamiento
(Li, Xuefei et al. 2014)	27	3	29	62	Plasma	Pretratamiento
(Weber et al. 2014)	17	8	11	164	Plasma	Postratamiento
(Wang, Shuai et al. 2014)	15	2	63	54	Plasma	Postratamiento
(Couraud et al. 2014)	11	2	9	37	Plasma	Pretratamiento
(Akca et al. 2013)	14	0	3	21	Plasma	Pretratamiento
(Jing et al. 2014)	31	0	17	72	Plasma	Pretratamiento
(Kim, Hye-Ryoun et al. 2013)	6	0	29	5	Plasma	Postratamiento
(Kim, Seung Tae et al. 2013)	8	3	4	42	Suero	Pretratamiento
(Liu et al. 2013)	27	0	13	46	Plasma	Pretratamiento
(Zhang et al. 2013)	15	0	7	64	Plasma	Pretratamiento
(Zhao et al. 2013)	16	3	29	63	Plasma	Pretratamiento
(Lv et al. 2013)	0	0	3	3	Plasma	Postratamiento
(Goto et al. 2012)	22	0	29	35	Suero	Pretratamiento
(Huang et al. 2012)	188	81	108	445	Plasma	Pretratamiento
(Xu et al. 2012)	4	0	4	16	Suero	Postratamiento
(Yam et al. 2012)	32	1	0	4	Plasma	Postratamiento
(Punnoose et al. 2012)	4	0	0	14	Plasma	Postratamiento
(Hu et al. 2012)	22	0	2	23	Suero	Pretratamiento
(Nakamura et al. 2012)	21	0	26	23	Plasma	Pretratamiento
(Jiang et al. 2011)	14	0	4	40	Suero	Pretratamiento

Anexos

(Sriram et al. 2011)	3	0	3	58	Suero	Pretratamiento
(Brevet et al. 2011)	5	2	9	15	Plasma	Pretratamiento
(Taniguchi et al. 2011)	32	0	12	0	Plasma	NA
(Bai et al. 2009)	63	16	14	137	Plasma	Postratamiento
(He et al. 2009)	8	1	0	9	Plasma	Postratamiento
(Kuang et al. 2009)	21	9	2	11	Plasma	Postratamiento
(Yung et al. 2009)	15	0	4	16	Plasma	Pretratamiento
(Kimura et al. 2007)	6	1	2	33	Suero	Postratamiento
(Kimura, Kasahara, Kawaishi, et al. 2006)	3	2	1	6	Suero	Pretratamiento
(Kimura, Kasahara, Shibata, et al. 2006)	2	1	2	6	Suero	Pretratamiento

Anexo 8: Resultados de la herramienta QUADAS-2 de la evaluación de la calidad de los estudios incluidos en el MA de la eficacia de la detección de mutaciones de *EGFR* en BL para la adecuación del tratamiento

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Akca 2013	+	+	+	+	+	+	+
Arriola 2018	+	+	+	+	+	+	+
Bai 2009	+	+	+	+	+	+	+
Brevet 2011	+	+	+	+	+	+	+
Buder 2018	+	-	+	+	+	+	+
Chai 2016	+	+	+	+	+	+	+
Courad 2014	+	+	+	+	+	+	+
Denis 2019	+	+	+	+	+	+	+
Ding 2019	+	-	+	?	+	+	+
Douillard 2014	+	+	+	+	+	+	+
Duan 2015	+	+	+	+	+	+	+
Goto 2012	+	+	+	+	+	+	+
Gu 2017	?	+	+	+	+	+	+
Guo 2017	+	+	+	+	+	+	+
Han 2016	+	+	+	+	+	+	+
He 2009	+	+	+	+	+	+	+
He 2016	?	-	+	+	+	+	+
He 2016 (2)	?	-	+	+	+	+	+
He 2017	?	+	+	+	+	+	+
Hong 2019	+	-	+	+	+	+	+
Hu 2012	+	+	+	+	+	+	+

Huang 2012	+	+	+	+	+	+	+
Ishii 2015	+	+	+	+	+	+	+
Ito 2018	+	+	+	+	+	+	+
Jenkins 2017	+	+	+	+	-	+	+
Jiang 2011	+	+	+	+	+	+	+
Jing 2013	+	+	+	+	+	+	+
Karlovich 2016	-	?	+	?	+	+	+
Kasahara 2017	?	-	+	+	+	?	+
Kim HR 2013	?	?	+	+	+	+	+
Kim ST 2013	+	+	+	+	+	+	+
Kimura 2006	+	+	+	+	+	+	+
Kimura 2006 (2)	+	+	+	+	+	?	?
Kimura 2007	+	+	+	+	+	?	?
Kobayashi 2018	?	+	+	+	+	+	+
Kuang 2009	-	?	+	+	+	+	+
Lam 2015	+	+	+	+	+	+	+
Leighl 2019	+	+	?	+	+	+	?
Li 2014	?	+	+	+	+	+	+
Li 2019	?	-	+	+	+	+	+
Liu 2013	+	+	+	+	+	+	+
Liu 2018	+	+	+	+	+	+	+
Lv 2013	+	+	+	+	+	+	+
Ma 2016	+	+	+	+	+	+	+
Ma 2017	+	+	+	+	+	+	+
Mellert 2017	?	-	+	?	+	+	+
Nakamura 2012	-	?	+	+	?	?	+

Newman 2016	?	?	+	+	+	?	+
Oxnard 2016	?	+	+	+	+	+	+
Pasquale 2015	?	-	+	+	+	+	+
Paweletz 2016	?	+	+	+	+	+	+
Pecuchet 2016	+	-	+	+	+	+	+
Punnoose 2012	+	-	?	+	+	?	+
Que 2016	+	+	+	+	+	+	+
Rachiglio 2016	+	+	+	+	+	+	+
Reck 2016	+	+	+	+	+	+	+
Reckamp 2016	?	-	+	+	+	+	+
Sacher 2016	?	+	+	+	+	+	+
Schwaederle 2017	+	+	?	+	+	+	+
Seki 2016	?	+	+	+	+	+	+
Shi 2018	+	+	+	+	+	+	+
Sriram 2011	+	+	+	+	+	+	+
Sundaresan 2016	+	-	+	+	+	+	+
Suzawa 2017	?	+	+	+	+	+	+
Takahama 2016	+	+	+	+	+	+	+
Taniguchi 2011	?	+	+	+	+	+	+
Thompson 2016	+	+	+	+	+	+	+
Thress 2016	?	+	+	?	+	+	+
Uchida 2015	+	?	+	+	+	+	+
Vázquez 2016	?	+	+	+	+	+	+
Veldore 2018	+	+	+	+	+	+	+
Wang 2014	+	+	+	+	+	+	+
Wang 2017	+	+	+	+	+	+	+
Weber 2014	+	+	+	+	+	+	+

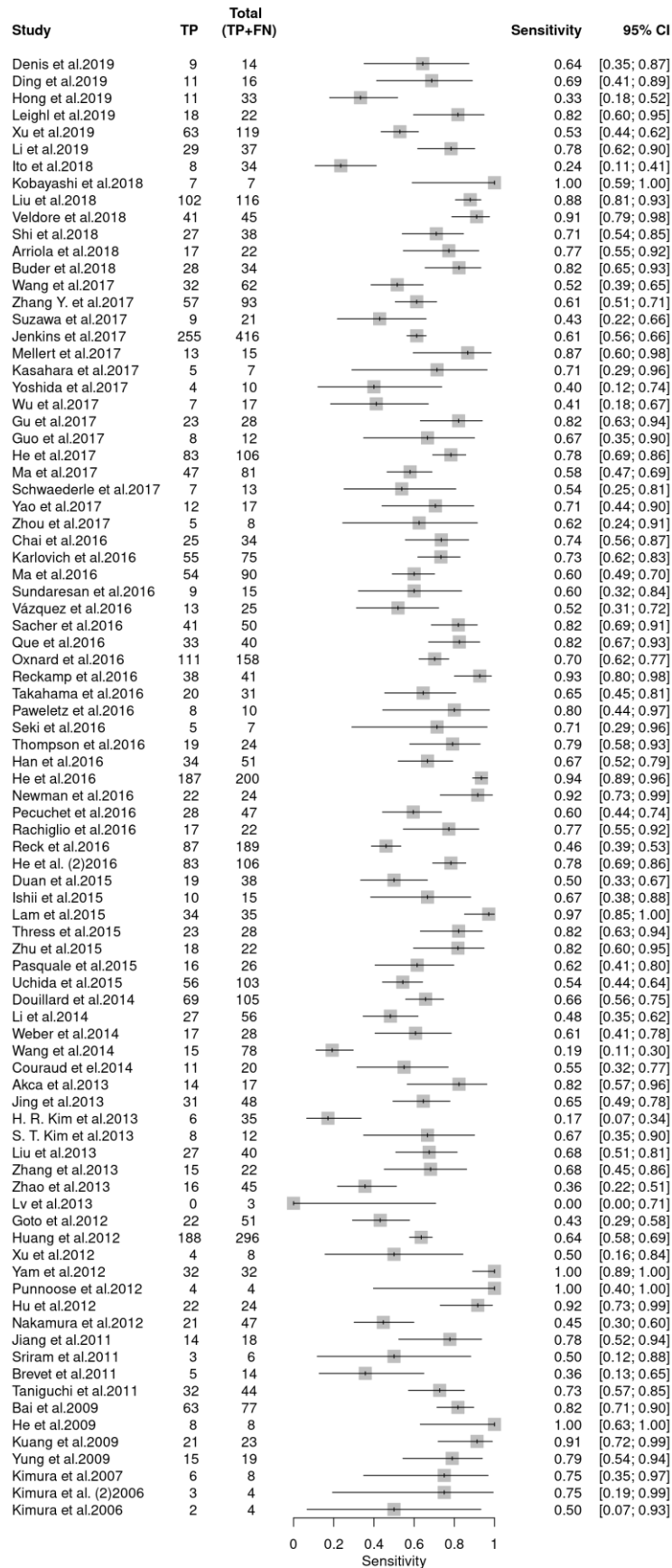
Wu 2017	+	+	+	+	+	+	+
Xu 2012	+	+	+	+	+	+	+
Xu 2019	+	+	+	+	+	+	+
Yam 2012	+	+	+	+	+	+	+
Yao 2017	+	+	+	+	+	+	+
Yoshida 2017	?	-	+	+	+	+	+
Yung 2009	+	+	+	+	+	+	+
Zhang 2013	+	+	+	+	+	+	+
Zhang X 2017	+	?	+	+	+	+	+
Zhang Y 2017	+	+	+	+	+	+	+
Zhao 2013	+	+	+	+	+	+	+
Zhou 2017	+	+	+	+	+	+	+
Zhu 2015	+	+	+	+	+	+	+

 High
  Unclear
  Low

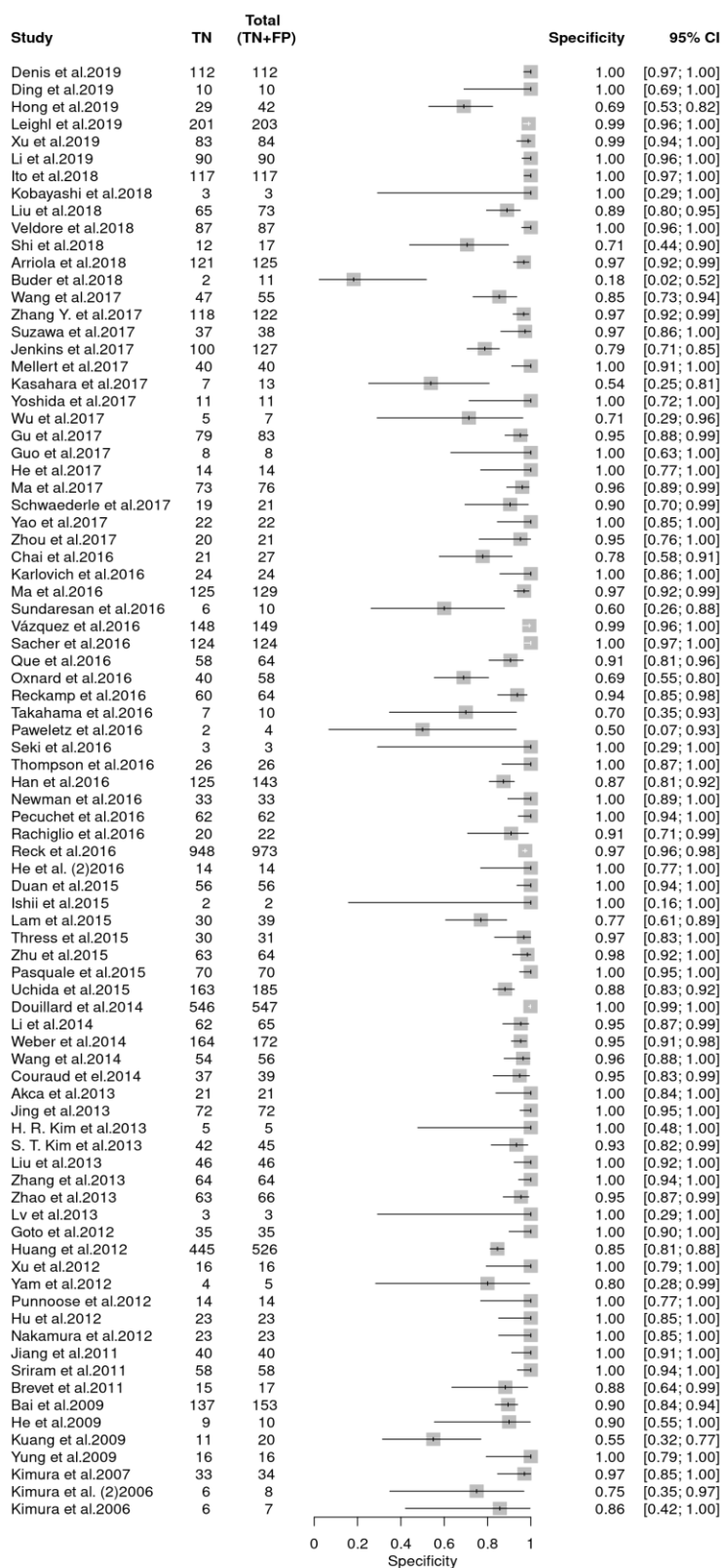
Anexos

Anexo 9: *Forest plots* de la sensibilidad y especificidad obtenida en el metanálisis todos los estudios clínicos incluidos para evaluar la efectividad de la adecuación del tratamiento del análisis de las mutaciones en *EGFR* en muestras de biopsia líquida en pacientes de NSCLC al compararlas con biopsias tisulares.

- Representación *forest plot* de la sensibilidad



- Representación *forest plot* de la especificidad



Anexo 10: Listado de los ARNlnc con expresión diferencial comunes en los 4 tipos de muestras.

<i>FENDRR</i>	<i>DRAIC</i>
<i>FIRRE</i>	<i>LINC01290</i>
<i>A2M-AS1</i>	<i>SFTA1P</i>
<i>ADAMTS9-AS1</i>	<i>LINC00968</i>
<i>CASC19</i>	<i>AC020907.1</i>
<i>AC098617.1</i>	<i>AC011700.1, PROX1-AS1</i>
<i>ADIRF-AS1, AL136982.7</i>	<i>AC244100.2</i>
<i>AC027228.2</i>	<i>ADAMTS9-AS2</i>
<i>PARAL1</i>	<i>MIR22HG</i>
<i>AP001189.1</i>	<i>VIPR1-AS1</i>
<i>CASC9</i>	<i>AC009630.2</i>
<i>AC044810.2, AC044810.3</i>	<i>AC018647.1</i>
<i>AC023593.1</i>	<i>AP003501.1</i>
<i>STARD13-AS</i>	<i>PVT1</i>
<i>ZEB2-AS1</i>	<i>LANCL1-AS1</i>
<i>ZFPM2-AS1</i>	<i>TFAP2A-AS2</i>
<i>AL021026.1</i>	<i>LINC01980</i>



UNIVERSIDAD
DE MÁLAGA

