



DEPARTAMENTO DE MEDICINA Y CIRUGÍA

TESIS DOCTORAL

**Valor pronóstico del análisis de marcadores
moleculares mediante proteómica dirigida en cáncer de
mama precoz con receptores hormonales positivos**

María Purificación Martínez del Prado

Dirigida por:

Dr. Enrique Espinosa Arranz

Dr. Juan Ángel Fresno Vara

Madrid, 2018



Departamento de Medicina y Cirugía



TESIS DOCTORAL

**Valor pronóstico del análisis de marcadores moleculares
mediante proteómica dirigida en cáncer de mama precoz con
receptores hormonales positivos**

Memoria para optar al grado de Doctor presentada por

María Purificación Martínez del Prado

Dirigida por:

Dr. Enrique Espinosa Arranz

Dr. Juan Ángel Fresno Vara

Madrid, 2018

**Memoria presentada para la obtención del grado de Doctor en
Medicina por M^a Purificación Martínez del Prado**

Dirigida por el Dr. Enrique Espinosa Arranz y el Dr. Juan Ángel Fresno Vara

D. Enrique Espinosa Arranz, Médico Jefe de Sección del Servicio de Oncología Médica del Hospital Universitario La Paz y clínico colaborador docente de la Facultad de Medicina de la Universidad Autónoma de Madrid.

D. Juan Ángel Fresno Vara, investigador de la Unidad de Oncología Traslacional del Servicio de Oncología Médica del Hospital Universitario La Paz.

Certifican:

Que Dña. M^a Purificación Martínez del Prado ha realizado durante los últimos tres años, bajo nuestra dirección, el trabajo titulado “**Valor pronóstico del análisis de marcadores moleculares mediante proteómica dirigida en cáncer de mama precoz con receptores hormonales positivos**”, para optar al grado de Doctor en Medicina.

Estimamos que su trabajo es novedoso y original, además de reunir los requisitos necesarios de interés científico y rigor metodológico para ser aceptado para su presentación.

Madrid, 15 de enero de 2018

Dr. Enrique Espinosa Arranz

Dr. Juan Ángel Fresno Vara

*A mis padres Andrés e Irene:
por haberme dado la vida,
por sus enseñanzas sobre la vida ,
y por ser el mejor modelo a seguir en mi vida.*

*“El no y el sí son breves de decir
pero piden pensar mucho”*

Baltasar Gracián

(1601-1658)

*“La verdadera ciencia enseña,
por encima de todo, a dudar y ser ignorante”*

Miguel de Unamuno

(1864-1936)

AGRADECIMIENTOS

A mis directores de Tesis, Dr. Enrique Espinosa y Dr. Juan Ángel Fresno, porque me brindaron la oportunidad de poder colaborar con ellos y por introducirme en un tema tan complejo, pero a la vez tan apasionante como es la proteómica aplicada a la medicina y en concreto al cáncer de mama. Gracias por hacer que un sueño se haya hecho realidad.

Al Dr. Angelo Gámez, por su profesionalidad, su capacidad de trabajo, su implicación y su ayuda desinteresada, por todo lo que me ha enseñado a lo largo de la duración de este trabajo. Y sobre todo muchas gracias por tu ayuda a nivel personal.

A Lucía Trilla, porque siempre tiene una sonrisa amable y por responder a mis requerimientos al momento. Gracias por tu trabajo.

A Mariana Díaz Almirón, miembro de la Unidad de Bioestadística del Instituto de Investigación del HU La Paz por el análisis estadístico.

Mi agradecimiento al resto de colaboradores y personal del laboratorio. Tengo una deuda con vosotros. Gracias.

A la Dra. Pilar Zamora Auñón, Médico adjunto del Servicio de Oncología Médica del Hospital Universitario La Paz por conseguir implicarme en este proyecto de investigación, que ha sido todo un reto. Gracias por tu amistad y cariño.

A la Dra. Elena Galve, Dra. M^a Ángeles Sala del Servicio de Oncología Médica del Hospital Universitario Basurto, por recordarme cada día que tras el esfuerzo diario conseguiría alcanzar la meta y sobre todo por su ayuda.

A la Dra. Sara Fernández del Servicio de Anatomía Patológica por la revisión de todas las muestras tumorales para mejorar la calidad del proyecto.

A mis padre Andrés e Irene, a mis hermanos Andrés y Begoña, a Yolanda, a mis sobrinos Fernando y Sete, por vuestros ánimos y apoyo durante esta larga andadura. Gracias de corazón por quererme siempre. Os quiero.

RESUMEN

Introducción: El cáncer de mama (CM) es una enfermedad heterogénea a nivel molecular. La proteómica dirigida permite la investigación de potenciales biomarcadores, factores pronósticos, predictivos, diagnósticos y nuevas dianas terapéuticas en el cáncer de mama. En una serie inicial de pacientes (cohorte de descubrimiento) con CM, receptores hormonales positivos (RH +) y afectación ganglionar, se analizó en muestras de tejido fijado en formalina y embebido en parafina (FFEP) la expresión proteica mediante espectrometría de masas (*label free shotgun*). Se objetivaron dos subgrupos moleculares: Triple Negativo-*like* (TN-*like*) y ER-true, con distinta evolución clínica y que presentaban diferencias en proteínas relacionadas con la adhesión celular por lo que a este conjunto de proteínas diferenciadoras se le denominó “Firma de Adhesión” (FdA).

Objetivo: Validar la existencia del subgrupo molecular denominado TN-*like* en una cohorte independiente de CM, para confirmar que la FdA define a un subtipo dentro del grupo luminal con características moleculares y clínicas similares al triple negativo.

Material y Métodos: Estudio retrospectivo de una cohorte de pacientes con diagnóstico de CM, con RH + y afectación ganglionar. Se incluyeron 46 muestras de tejido FFEP. A su vez se analizaron 18 muestras de la cohorte de descubrimiento. Se realizó un análisis de proteómica dirigida vía monitorización de reacción selectiva (SRM).

Resultados: SRM fue capaz de detectar diferencias entre el subgrupo TN-*like* y el ER-True en las muestras de la serie de descubrimiento. Se realizó un predictor de clase con 14 proteínas para asignar a la cohorte de validación la categoría ER-True y TN-Like. El predictor de clase asignó el 34,7% (16) de las muestras al grupo ER-true y el 65,3% (30) al grupo TN-*like* en la cohorte de validación. El subgrupo TN-Like presenta mayor tamaño tumoral, mayor grado histológico y mayor porcentaje de metástasis a distancia y segundos tumores (77%). La supervivencia libre de enfermedad a distancia a los 5, 10 y 15 años fue del 93,73%, 87,05% y 67,70% y del 68,79%, 50,45% y 40,36% para los subgrupos ER-true y TN-*like* respectivamente, sin significación estadística ($p=0,0777$).

Conclusión: La cuantificación de proteínas en muestras fijadas en formol y embebidas en parafina es posible mediante la técnica SRM. La “firma de adhesión” es capaz de identificar dos grupos de pacientes (ER-true y TN-*like*) con distintas características moleculares y evolución clínica.

SUMMARY

Introduction: Breast cancer (BC) is a heterogeneous disease at the molecular level. Targeted proteomics allows the characterization of biomarkers with prognostic, diagnostic and predictive value, and new molecular targets for the treatment of breast cancer. In an initial cohort of patients with hormone-receptor positive (HR+) BC and lymph node involvement, protein expression was analyzed in formalin-fixed and paraffin-embedded tissue samples (FFPE) by mass spectrometry (label free shotgun). Two molecular subgroups with different molecular characteristics and clinical evolution were observed (Triple Negative-like (TN-*like*) and ER-*true*). These groups differed in proteins related to cellular adhesion, so this set of proteins was named "Adhesion Signature" (FdA).

Objective: To validate the existence of the molecular subgroup known as TN-*like* in a breast cancer independent cohort, to confirm that FdA defines a new subtype within the luminal group with molecular and clinical characteristics similar to the triple negative subtype.

Material and Methods: A retrospective study of a cohort of BC patients diagnosed with, with HR + BC and showing lymph node involvement was performed. 46 FFPE tissue samples were included. Additionally, 18 samples from the discovery cohort were analyzed. A targeted proteomics analysis was carried out via selective reaction monitoring (SRM).

Results: SRM detected differences at the protein expression level between the TN-*like* and the ER-*true* subgroups from the discovery cohort. A class predictor with 14 proteins was built to assign these categories in the validation cohort. The class predictor assigned 34.7% (16) to the ER-true group and 65.3% (30) to the TN-like group in the validation cohort. Tumors in the TN-Like subgroup had larger size, higher histological grade and higher percentage of metastatic spread and second tumors (77%), Disease-free survival rates at 5, 10 and 15 years were 93.73%, 87.05% and 67.70% for ER-true and 68.79%, 50.45% and 40.36% for TN-like tumors ($p = 0.0777$).

Conclusion: SRM technique allows proteins quantification in FFPE samples. The "adhesion signature" identifies two groups of patients (ER-true and TN-like) with different molecular characteristics and clinical evolution.

ÍNDICE

RESUMEN	I
SUMMARY	II
ÍNDICE.....	III
ÍNDICE DE FIGURAS	V
ÍNDICE DE TABLAS.....	VI
ABREVIATURAS	VII
1. INTRODUCCIÓN.....	1
1.1 EPIDEMIOLOGÍA DEL CÁNCER DE MAMA.....	1
1.2 FACTORES DE RIESGO.....	3
A) FACTORES DEMOGRÁFICOS:	3
B) FACTORES HORMONALES Y REPRODUCTIVOS.....	3
C) FACTORES GENÉTICOS	4
D) ENFERMEDAD BENIGNA DE LA MAMA	4
E) DENSIDAD MAMARIA	4
F) FACTORES MEDIOAMBIENTALES	4
G) FACTORES RELACIONADOS CON EL ESTILO DE VIDA	5
1.3. DIAGNÓSTICO Y ESTADIFICACIÓN DEL CÁNCER DE MAMA	5
A) DIAGNÓSTICO	5
B) ESTADIFICACIÓN.....	6
1.4 CLASIFICACIÓN DEL CÁNCER DE MAMA.....	8
A) PARÁMETROS INCLUIDOS EN LA CLASIFICACIÓN HISTOPATOLÓGICOS	8
B) CLASIFICACIONES MOLECULARES DEL CÁNCER DE MAMA.....	9
1.5. FACTORES PRONÓSTICOS Y PREDICTIVOS	13
A) EDAD.....	14
B) TAMAÑO DEL TUMOR.....	14
C) GANGLIOS AXILARES	14
D) VARIEDAD HISTOLÓGICA.....	14
E) GRADO HISTOLÓGICO.....	15
F) INVASIÓN LINFOVASCULAR.....	15
G) RECEPTORES HORMONALES.....	15
H) Ki67	15
I) PERFILES MOLECULARES	16
1.6. GENERALIDADES DEL TRATAMIENTO DEL CÁNCER DE MAMA.....	24
1.7 PROTEÓMICA EN CÁNCER DE MAMA.....	26
A) GENERALIDADES	26
B) PLATAFORMAS PARA EL ANÁLISIS DE PROTEÍNAS	27
C) ESPECTROMETRÍA DE MASA DE PROTEÍNAS.....	28
D) TÉCNICAS DE SEPARACIÓN O FRACCIONAMIENTO DE PROTEÍNAS	29
E) TÉCNICAS DE ANÁLISIS DE PROTEÍNAS	30
F) TÉCNICAS DE IDENTIFICACIÓN DE PROTEÍNAS	30
G) TÉCNICAS DE CUANTIFICACIÓN DE PROTEÍNAS	31
H) ANÁLISIS DE LOS DATOS	32

1.8 PERFILES DE EXPRESIÓN PROTEICA EN CÁNCER DE MAMA	33
1.9 FIRMA DE ADHESIÓN: EL SUBGRUPO TN-LIKE	35
2. HIPÓTESIS Y OBJETIVOS	39
2.1 HIPÓTESIS.....	39
2.2 OBJETIVOS	39
A. OBJETIVO PRIMARIO:	39
B. OBJETIVOS SECUNDARIOS:.....	39
3. MATERIAL Y MÉTODOS	41
3.1 COHORTE DE DESCUBRIMIENTO	41
A) SELECCIÓN DE LA MUESTRA	41
B) SELECCIÓN DE LAS VARIABLES.....	41
3.2 COHORTE DE VALIDACIÓN: PROYECTO ÁLAMO	41
A) SELECCIÓN DE LA MUESTRA	41
B) SELECCIÓN DE LAS VARIABLES.....	44
3.3 ANÁLISIS ESTADÍSTICO	45
3.4 METODOLOGÍA PARA EL ANÁLISIS PROTEÓMICO	46
A) PREPARACIÓN Y EXTRACCIÓN DE PROTEÍNAS A PARTIR DE LAS MUESTRAS CLÍNICAS	46
B) DETECCIÓN DE LA EXPRESIÓN DE ERBB2 POR qPCR.....	46
C) <i>SELECTED REACTION MONITORING</i> (SRM).....	47
D) DESARROLLO DEL PREDICTOR DE CLASE.....	48
4. RESULTADOS	51
4.1 DESCRIPCIÓN DE LA COHORTE DE VALIDACIÓN	51
4.2 ANÁLISIS DE SUPERVIVENCIA.....	52
4.3 EXPRESIÓN DE ERBB2 -HER2- POR qPCR	53
4.4 RESULTADOS SRM	54
A) DISEÑO DEL EXPERIMENTO DE SRM	54
B) MEDICIÓN DE LA EXPRESIÓN DE PROTEÍNAS POR SRM.....	55
C) VALIDACIÓN DEL PREDICTOR DE CLASE EN LA COHORTE DE DESCUBRIMIENTO (18 MUESTRAS).....	55
D) CORRELACIÓN ENTRE LA ESPECTROMETRÍA DE MASAS Y LOS SRM.....	56
4.5 DESARROLLO DEL PREDICTOR DE CLASE.....	57
4.6 ASIGNACIÓN DE LAS MUESTRAS DE VALIDACIÓN A LOS GRUPOS ER-TRUE O TN-LIKE A PARTIR DEL PREDICTOR DE CLASE.....	58
4.7 SUPERVIVENCIA LIBRE DE ENFERMEDAD	60
4.8 ANÁLISIS UNIVARIANTE Y MULTIVARIANTE	61
5. DISCUSIÓN.....	63
6. CONCLUSIONES.....	71
7. BIBLIOGRAFÍA	73
8. ANEXOS	95

ÍNDICE DE FIGURAS

FIGURA 1. GENES INCLUIDOS EN LA PLATAFORMA PAM 50	22
FIGURA 2. <i>CLUSTER</i> JERÁRQUICO	35
FIGURA 3. <i>CLUSTER</i> JERÁRQUICO DE LAS 44 PROTEÍNAS EXPRESADAS DIFERENCIALMENTE ENTRE <i>ER-TRUE</i> Y <i>TN-LIKE</i>	36
FIGURA 4. DIAGRAMA DE FLUJO	43
FIGURA 5. SUPERVIVENCIA LIBRE DE ENFERMEDAD	52
FIGURA 6. SUPERVIVENCIA LIBRE DE ENFERMEDAD A DISTANCIA	53
FIGURA 7. MEDICIÓN DE ERBB2 EN LA SERIE DE REFERENCIA Y LA COHORTE DE VALIDACIÓN.....	53
FIGURA 8. CURVA ROC PARA LA PREDICCIÓN POR QPCR DEL ESTATUS DE ERBB2.....	54
FIGURA 9. FLUJO DE TRABAJO EXPERIMENTAL EMPLEADO EN NUESTRA SERIE	55
FIGURA 10. CLÚSTER JERÁRQUICO DE LAS 18 MUESTRAS DE LA SERIE DE DESCUBRIMIENTO	56
FIGURA 11. CORRELACIÓN ENTRE LAS MEDICIONES OBTENIDAS POR EM Y LOS SRM PARA LAS 33 PROTEÍNAS SELECCIONADAS	57
FIGURA 12. SUPERVIVENCIA LIBRE DE ENFERMEDAD A DISTANCIA EN LOS DOS GRUPOS <i>ER-TRUE</i> Y <i>TN-LIKE</i>	60
FIGURA 13. SUPERVIVENCIA LIBRE DE ENFERMEDAD EN LOS <i>ER-TRUE</i> Y LOS <i>TN-LIKE</i>	60

ÍNDICE DE TABLAS

TABLA 1. FACTORES DE RIESGO CONOCIDOS DE CÁNCER DE MAMA Y SU RIESGO RELATIVO.....	3
TABLA 2. CLASIFICACIÓN TNM DEL CÁNCER DE MAMA	7
TABLA 3. AGRUPACIÓN POR ESTADIOS	8
TABLA 4. CARACTERÍSTICAS DE LOS SUBTIPOS <i>INTEGRATIVE CLUSTER</i> DEL PROYECTO METABRIC.....	13
TABLA 5. GENES INCLUIDOS EN LA PLATAFORMA ONCOTYPE DX.....	18
TABLA 6. FUNCIÓN DE LOS GENES INCLUIDOS EN LA PLATAFORMA MAMMAPRINT	20
TABLA 7. CARACTERÍSTICAS DE LAS PLATAFORMAS DE EXPRESIÓN GÉNICA MÁS UTILIZADAS.....	23
TABLA 8. SUBTIPOS DE CÁNCER DE MAMA DEFINIDOS POR IHQ Y TRATAMIENTO	25
TABLA 9. ESTUDIOS DE PROTEÓMICA EN CÁNCER DE MAMA RECEPTORES HORMONALES POSITIVOS Y/O NEGATIVOS	34
TABLA 10. FUNCIONES DE LAS PROTEÍNAS QUE COMPONEN LA FIRMA DE ADHESIÓN.....	37
TABLA 11. . CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS BASALES DE LAS PACIENTES.....	51
TABLA 12. PODER DISCRIMINANTE DE LAS 14 PROTEÍNAS QUE COMPONEN EL PREDICTOR DE CLASE	57
TABLA 13. CARACTERÍSTICAS DE LAS PROTEÍNAS QUE DEFINEN EL PREDICTOR DE CLASE.....	58
TABLA 14. CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS DE LOS SUBGRUPOS <i>TN-LIKE</i> Y <i>ER-TRUE</i>	59
TABLA 15. RECURRENCIA DE LAS PACIENTES EN LA COHORTE DE VALIDACIÓN POR SUBGRUPOS.....	59
TABLA 16. ANÁLISIS UNIVARIANTE Y MULTIVARIANTE	61

ABREVIATURAS

- 2D-DIGE: Electroforesis bidimensional diferencial en gel
- ABCSG: *Austrian Breast & Colorectal Cancer Study Group*
- ADN: ácido desoxirribonucleico
- ADNc: ácido desoxirribonucleico complementario
- AJCC: American Joint Committee on Cancer
- ARN/RNA: ácido ribonucleico
- ARNm/RNAM: ácido ribonucleico mensajero.
- ATAC: *Arimidex, tamoxifen alone or in combination*
- BRCA1: *Breast Cancer Type 1 susceptibility protein*
- BRCA2: *Breast Cancer Type 2 susceptibility protein*
- CK: citoqueratinas
- CM: cáncer de mama
- CID: Disociación inducida por colisión, del inglés *Collision-Induced Dissociation*
- Cq: ciclo de cuantificación, del inglés *Quantification cycle*
- Ct: equivale al Cq anterior, según la nomenclatura de Applied Biosystems, del inglés *Threshold cycle*
- DAVID: Base de datos para la anotación, visualización y el descubrimiento integrado, del inglés *Database for Annotation, Visualization and Integrated Discovery*
- DIGE: Electroforesis diferencial en gel
- ECD: Disociación por captura de electrones, del inglés *Electron Capture Dissociation*
- EM: Espectrómetro de masas
- EP: EndoPredict
- PCR: reacción en cadena de la polimerasa (en inglés, *Polymerase Chain Reaction*)
- ER: Receptor de Estrógenos
- ESI: Ionización por electrospray, del inglés *Electrospray Ionization*
- FAC5: Fluoracilo, Adriamicina, ciclofosfamida
- FDR: Tasa de falsos positivos, del inglés *False Discovery Rate*
- FF: Congeladas en fresco, del inglés *Fresh Frozen*
- FFEP: Fijadas en Formol y Embebidas en Parafina, del inglés *Formalin-Fixed, Parafin-Embedded*
- FISH: inmunofluorescencia de hibridación *in situ*
- FN: factor de normalización

G: grado nuclear

GEICAM: Grupo Español de Investigación de Cáncer de Mama

HBAS: Hospital de Basurto

HER2: receptor 2 de factor de crecimiento epidérmico humano

HPLC: Cromatografía líquida de alta eficiencia, del inglés *High-Performance Liquid Chromatography*

HULP: Hospital Universitario La Paz

HT: hormonoterapia

IHQ: inmunohistoquímica

IR: índice de recaída

ICAT: Marcaje de afinidad codificado por isótopos

Itraq: Marcaje isobárico para cuantificación relativa y absoluta

ILV: invasión linfovascular

LCM: Microdissección por captura con láser

LOOCV: Validación cruzada dejando uno fuera, del inglés *Leave-One Out Cross-Validation*

LTQ: Trampa lineal

M: Metástasis a distancia

m/z: Relación masa / carga

MALDI: Ionización por desorción láser asistida por matriz, del inglés *Matrix-Assisted Laser Desorption / Ionization*.

METABRIC: *Molecular Taxonomy of Breast Cancer International Consortium*

MINDACT: *Microarray in Node-Negative Disease May Avoid Chemotherapy*

MIQE: *Minimum Information for Publication of Quantitative Real-Time PCR Experiments*

MS: Espectrometría de masas, del inglés *Mass Spectrometry*

MS/MS: Espectrometría de masas en tándem

N: afectación ganglionar

NIH: Instituto Nacional del Cáncer, del inglés *National Institute of Health*

PCR: reacción en cadena de la polimerasa

QHT: tratamiento combinado de quimioterapia y hormonoterapia

QT: quimioterapia

PANTHER: Análisis de proteínas mediante relaciones evolutivas, del inglés *Protein Analysis Through Evolutionary Relationships*

PCR: Reacción en cadena de la polimerasa

PEG: Plataformas de expresión génica

RE: receptor de estrógenos

RH: receptores hormonales

ROC: Característica operativa del receptor, del inglés *Receiver Operating Characteristic*

RPg: receptor de progesterona

RPPA: Matrices de proteínas de fase inversa

RPMI-1460: Medio de cultivo desarrollado en el Roswell Park Memorial Institute

RS: *Recurrence Score* (Oncotype DX[®])

ROR: riesgo de recurrencia, del inglés *Risk of Recurrence Score*

RT-PCR: reacción en cadena de la polimerasa con transcripción inversa

RT-qPCR: reacción en cadena de la polimerasa con transcripción inversa en tiempo real

SAM: Análisis de significación de microarrays, del inglés *Significance Analysis of Microarrays*

SILAC: Marcaje con isótopos estables mediante aminoácidos en cultivos celulares

SDS: Dodecilsulfato de sodio

SRM: Monitorización selectiva de reacción, del inglés *Selected Reaction Monitoring*

T: Tamaño tumoral

TAILORx: *Trial Assigning Individualized Options for Treatment*

TRANSBIG: Proyecto de acercamiento del trabajo de laboratorio a los tratamientos mediante la investigación traslacional

TNBC: Cáncer de Mama Triple Negativo, del inglés, *Triple Negative Breast Cancer*

TNM: es el sistema de estadificación del cáncer, basado en el tamaño tumoral (T), afectación ganglionar (N) y afectación metastásica (M)

TOF: Tiempo de vuelo, del inglés *Time Of Flight*

1. INTRODUCCIÓN

1. INTRODUCCIÓN

1.1 EPIDEMIOLOGÍA DEL CÁNCER DE MAMA

Según los datos del proyecto GLOBOCAN 2012 de la *International Agency for Research on Cancer* (IARC), el cáncer de mama (CM) es el segundo tumor más frecuente en el mundo tras el cáncer de pulmón, con 1,67 millones de nuevos casos (11,9%) diagnosticados en 2012 en ambos sexos. Es el tumor más frecuente entre las mujeres (25%), con más casos en los países en vías de desarrollo (883.000 casos) que en los países desarrollados (794.000). Presenta una tasa ajustada por edad (100.000 personas/año) del 43,3 y el riesgo acumulado (0-74 años) del 4,6% en 2012 (1).

La tasa de incidencia en el mundo es mayor en Norte América, Australia/Nueva Zelanda y noreste de Europa y menor en Asia y en África subsahariana (2). La incidencia puede variar incluso en 4 veces dependiendo de las regiones del mundo, con unas tasas que oscilan entre el 27 por 100.000 en África Central y Este de Asia y el 96 en el Oeste de Europa. En las décadas de los 80 y 90 se incrementó en los países desarrollados cerca del 30%. Esta tendencia se explicó por la implementación de los programas de cribado, los cambios en los patrones de reproducción, el uso de la terapia hormonal sustitutiva (THS) en mujeres post-menopáusicas, la exposición medioambiental y la dieta (3,4). A partir del año 2003 se produjo una ligera disminución o estabilización de la incidencia, tras la disminución de la prescripción en mujeres post-menopáusicas de la THS a raíz de la publicación de un estudio que demostró el incremento del riesgo de cáncer de mama entre las mujeres post-menopáusicas con THS (5) y quizás por una estabilización en la participación en los programas de cribado (6). Por el contrario, en los países en vías de desarrollo la tasa de incidencia continúa aumentando, siendo las causas probables la mayor sensibilización y la detección precoz, así como los cambios en los patrones de reproducción. Respecto a la mortalidad, el cáncer de mama es la 5ª causa de muerte global, con 522.000 muertes en 2012 (6,4%) y la causa más frecuente por cáncer en mujeres en los países en vías de desarrollo con 324.000 muertes, (14,3%). En los países desarrollados es la segunda causa de muerte por cáncer con 98,000 muertes (15,4%) después del cáncer de pulmón. El rango en la tasa de mortalidad entre las diferentes regiones del mundo es menor que para la incidencia por la mejor supervivencia en los países desarrollados a pesar de una mayor incidencia (1). Desde 1990, la tasa de mortalidad ha disminuido en la mayoría de los países desarrollados, y se atribuye a la detec-

ción precoz y a los avances en el tratamiento (3). Aunque esta disminución se produce en los países con mayor incidencia, sin embargo está aumentando en países con incidencia menor, como en Latinoamérica, Caribe y parte de Asia, probablemente debido a los cambios en los factores de riesgo y el acceso limitado a la detección precoz y los tratamientos (7).

La estimación de la incidencia general de cáncer para el año 2017 en Estados Unidos es de 1.688.780 casos nuevos, de los cuales 255.180 corresponden a cáncer de mama en ambos sexos, siendo 252.710 en mujeres, lo que representa el 30 % del total. Respecto a las muertes por cáncer en 2017 se estiman en 600.920 casos totales, de los cuales 41.070 corresponde a muertes por cáncer de mama en ambos sexos, siendo 40.610 casos en mujeres (14%) (8). La edad media del diagnóstico es de 61 años y cerca del 19% se diagnostican en mujeres en edades comprendidas entre 30-49 años, y el 44% en mujeres de 65 años o mayores (9). En los últimos años se ha objetivado una disminución de la incidencia relacionada con la instauración de los programas de cribado, la mejora en los tratamientos sistémicos y la disminución en el uso de THS (4).

En Europa, el cáncer de mama es el más frecuente en mujeres con una incidencia estandarizada por edad de 71/100.000 en 2012 (10). La supervivencia relativa o específica a 5 años (SR) en el periodo 2000–2007 fue del 82%, con un rango entre el 85% (países nórdicos) y el 74% (Este de Europa). La supervivencia fue más elevada en los rangos de edad de 45 a 65 años y disminuyó en las pacientes ancianas (11).

En España, según los datos publicados por REDECAN, en el año 2015 se diagnosticaron 27.747 casos nuevos de cáncer de mama, siendo el tumor más frecuente entre las mujeres, con una tasa cruda de 117,5/100.000 personas/año. La probabilidad de desarrollar un cáncer de mama en la población femenina en España durante el periodo 2003-2005, entre los 0-84 años fue del 8,99. Respecto a la prevalencia, el cáncer de mama es un tumor con una prevalencia elevada, ya que está relacionada con la supervivencia. En el año 2012 se diagnosticaron 25.215 pacientes con cáncer de mama, y su mortalidad fue de 6.075 (5,9%) casos por lo que su prevalencia a los 5 años fue de 104.210 (17,9%) casos en los últimos 5 años (1). La mortalidad por cáncer de mama en España en el año 2015 según el Instituto Nacional de Epidemiología fue de 6.212 casos, con una tasa bruta del 15,71 (ajustada por 100.000 personas/año) (12). La supervivencia relativa a 5 años, durante el periodo comprendido entre 2000 y 2007, fue del 82%, similar a la descrita por EUROCAR 5 para Europa (13).

1.2 FACTORES DE RIESGO

A los factores de riesgo del cáncer de mama establecidos, en los últimos años están asociándose nuevos factores como la obesidad. Algunos de ellos pueden ser modificables si mejoramos nuestro estilo de vida (Tabla 1).

Tabla 1. Factores de riesgo conocidos de cáncer de mama y su riesgo relativo (RR)(14).

RR < 2	RR 2-4	RR >4
Menarquia precoz	Familiar de primer grado con cáncer de mama	Mutación BRCA1/BRCA2
Menopausia tardía	Mutación CHEK2	Carcinoma lobulillar in situ
Nuliparidad	Primer embarazo a término > 35 años	Hiperplasia atípica
Terapia hormonal sustitutiva combinada	Enfermedad proliferativa de la mama	Exposición a radioterapia antes de los 40 años.
Ingesta de alcohol	Mamas densas	
Obesidad postmenopáusica		

a) Factores demográficos:

La edad es el factor de riesgo más importante. La incidencia se incrementa con la edad y es mayor a partir de la menopausia (15). Respecto al género, es 100 veces más frecuente en las mujeres que en los varones (15).

b) Factores hormonales y reproductivos

La mayoría de los tumores de mama están relacionados con las hormonas reproductoras femeninas. Estudios epidemiológicos han demostrado que el incremento de la exposición a los estrógenos endógenos debido a factores como la menarquia precoz, la menopausia tardía, la nuliparidad o la edad tardía del primer parto incrementan el riesgo de cáncer de mama (16). La THS en mujeres postmenopáusicas incrementa el riesgo comparado con las mujeres no consumidoras y se incrementa con la duración del mismo. El riesgo se reduce tras la suspensión del mismo (17). En relación al tipo de THS, la terapia hormonal combinada con estrógenos y progestágenos se asocia con mayor riesgo de cáncer que la monoterapia con estrógenos o placebo (5). Según el re-análisis de Oxford el uso de anticonceptivos orales no incrementa el riesgo (18), pero un análisis más reciente indica que por cada 10 años de uso se incrementa el riesgo en el 14% siendo este significativo (19).

c) Factores genéticos

La mayoría de las mujeres con diagnóstico de cáncer de mama no tienen un familiar con este diagnóstico. Solo el 5-10% tienen una predisposición hereditaria para el cáncer de mama. La historia familiar es un factor de riesgo heterogéneo, dependiendo del número de familiares afectados, el grado de relación familiar, la edad al diagnóstico y el número de familiares no afectados. El riesgo de desarrollar un cáncer de mama se incrementa en 1,5-3 veces si una mujer tiene una madre o una hermana con diagnóstico de cáncer de mama. La edad al diagnóstico menor de 30 años de un familiar de primer grado afectado incrementa el riesgo 3 veces y solo 1,5 veces si es mayor de 60 años (20). Entre los genes conocidos de predisposición al cáncer de mama destacan los siguientes: BRCA1/BRCA2 (21), P53 (22,23), PTEN (22,24), CDH1 (25,26) STK11/LKB1 (27) y otros genes de moderada penetrancia como CHEK2, ATM, BRIP1 y PALB2 (28).

d) Enfermedad benigna de la mama

La enfermedad proliferativa sin atipia aumenta ligeramente el riesgo (RR 1,5 a 2) siendo mayor el riesgo (RR 4-5) de la enfermedad proliferativa con hiperplasia ductal o lobular atípica (29). Un reciente meta-análisis demuestra que la enfermedad benigna de la mama con o sin atipia está asociada a un incremento significativo del riesgo de cáncer de mama (30).

e) Densidad mamaria

El sistema BI-RADS (*The Breast Imaging Reporting and Data System*) clasifica el tejido mamario como extremadamente denso cuando la mama presenta un tejido glandular mayor del 75% (31). Comparado con mujeres con una densidad mamaria menor del 10% en la mamografía, las mujeres con densidad igual o mayor al 75% tienen un incremento del riesgo de cáncer de mama de 4,7 si se detecta por cribado o menos de 12 meses tras una mamografía negativa (32).

f) Factores medioambientales

Las radiaciones ionizantes son el único factor de riesgo medioambiental reconocido relacionado con el riesgo de cáncer de mama. El riesgo está inversamente relacionado con la edad de la mujer en el momento de la exposición. Se incrementa en pacientes

jóvenes tratadas con radioterapia en Mantle antes de los 15 años (33). El riesgo absoluto acumulado tras la finalización del tratamiento es del 29% a los 30 años, estando relacionado con la asociación de agentes alquilantes (34).

g) Factores relacionados con el estilo de vida

La obesidad en mujeres post-menopáusicas está asociada a un incremento de cáncer de mama. Para un índice de masa corporal (IMC) previo al diagnóstico comparado con un peso normal, el RR para mortalidad total fue 1,41 para un IMC >30, 1,07 para un IMC entre 25-30 y 1,10 para un IMC <18,5. En mujeres obesas (IMC >30) pre-menopáusicas el RR para mortalidad es de 1,75 y 1,34 para post-menopáusicas (35). No se ha demostrado que la asociación entre la obesidad con los resultados del cáncer difiera por estado hormonal y estado menopáusico (36). La ingesta de alcohol es el único factor de riesgo relacionado con la dieta que se asocia con un incremento moderado del riesgo de cáncer de mama siendo su RR de 1,61 (37). Su asociación demuestra una relación lineal dosis respuesta (38).

1.3. DIAGNÓSTICO Y ESTADIFICACIÓN DEL CÁNCER DE MAMA

El cáncer de mama se diagnostica la mayoría de las veces por una biopsia de un nódulo mamario palpable o por una biopsia de un hallazgo sospechoso en la mamografía. Cualquier alteración mamaria debe investigarse hasta su resolución. Las alteraciones mamarias en el varón deben de valorarse igual que en la mujer. Los nódulos mamarios en una mujer embarazada o lactante no deben atribuirse a cambios hormonales de entrada, por lo que deben ser investigados.

a) Diagnóstico

Debe realizarse una exploración física, una mamografía y una ecografía mamaria. Si se objetiva una lesión sospechosa de malignidad se realizará una biopsia de la misma ya sea con aguja fina, con aguja gruesa o una biopsia excisional. Una mamografía negativa no excluye un cáncer de mama, por lo que deben realizarse otras pruebas diagnósticas como la resonancia nuclear magnética mamaria o PET-TAC. Las pacientes que se presentan con una mamografía de cribado anormal en ausencia de masa palpable

también requieren una evaluación, la cual depende de la interpretación de la mamografía por el radiólogo que categoriza la anormalidad mamográfica utilizando el sistema BI-RADS (31).

b) Estadificación

El objetivo de la estadificación de cualquier tumor es clasificar a los pacientes diagnosticados de cáncer según la extensión de su enfermedad con el objeto de establecer grupos de similar pronóstico, seleccionar el tratamiento más adecuado en cada estadio y poder comparar los resultados de diferentes series. En 1986 el *American Joint Committee on Cancer (AJCC)* propuso una estadificación basada en el sistema TNM (T: tamaño del tumor, N: ganglios linfáticos, M: metástasis a distancia) de la UICC. Se realizan sucesivas ediciones que van incorporando los avances originados en el diagnóstico y el tratamiento del cáncer de mama como consecuencia de los programas de cribado, incorporación del ganglio centinela, el uso de la inmunohistoquímica y las técnicas moleculares. Se realiza una estadificación clínica (cT cN cM) basada en la exploración física y pruebas radiológicas y tras la resección del tumor primario una estadificación patológica (pT pN pM). Se complementa con un estudio de extensión para descartar enfermedad a distancia. En la tabla 2 se describe la clasificación TNM 5ª Edición de 1997 utilizada en este proyecto (39,40) y en la tabla 3 la agrupación por estadios.

Tabla 2. Clasificación TNM del cáncer de mama (Modificada de *AJCC Cancer Staging Manual*, 5ª ed., 1997).

TUMOR PRIMARIO (T) Similar clasificación clínica y patológica	
Tis	Carcinoma in situ.
T1	≤ 2 cm en su diámetro mayor
T1a	> 0,1 cm y ≤ 0,5 cm
T1b	> 0,5 cm y ≤ 1 cm
T1c	> 1 cm y ≤ 2 cm
T2	> 2 a 5 cm en su diámetro mayor.
T3	> 5 cm en su diámetro mayor.
T4	Extensión directa a pared costal o piel
T4a	Extensión a la pared torácica
T4b	Ulceración o nódulos satélites ipsilaterales confinados a la mama o edema (incluida piel de naranja)
T4c	Ambos:T4a y T4b
T4d	Carcinoma inflamatorio
GANGLIOS LINFÁTICOS (N) Clasificación anátomo-patológica	
pN0	No ganglios regionales con metástasis.
pN1	Metástasis movibles en ganglio(s) axilares ipsilaterales
pN1a	Sólo micrometástasis (<0,2 cm)
pN1b	Metástasis > de 2 mm (alguna de ellas)
pN1bi	Metástasis en 1 a 3 ganglios, >0,2 cm/<2 cm
pN1bii	Metástasis en 4 ó más ganglios >0,2 cm/<2 cm
pN1biii	Extensión extracapsular en un ganglio linfático < 2 cm
pN1biv	Metástasis en ganglios linfáticos ≥ 2 cm
pN2	Metástasis en ganglios axilares ipsilaterales con fijación entre ellos o con otras estructuras
pN3	Metástasis en ganglios mamarios internos
METÁSTASIS A DISTANCIA (M) Similar clasificación clínica y patológica	
M0	No hay evidencia de metástasis a distancia.
M1	Metástasis a distancia (incluidos los ganglios supraclaviculares)

Tabla 3. Agrupación por estadios (AJCC 5ª Ed.).

ESTADIFICACIÓN	
Estadio 0	Tis N0 M0
Estadio I	T1 N0-N1a M0
Estadio IIA	T0-1,N1 M0; T2 N0 M0
Estadio IIB	T2 N1 M0; T3 N0 M0
Estadio IIIA	T0-T2 N2 M0 T3 N1-2 M0
Estadio IIIB	T4, cualquier N, M0 Cualquier T, N3, M0
Estadio IV	Cualquier T, Cualquier N, M1

1.4 CLASIFICACIÓN DEL CÁNCER DE MAMA

La clasificación clásica del cáncer de mama invasivo se ha venido realizando según criterios histológicos basados en la morfología celular y en el análisis inmunohistoquímico (IHQ). Los parámetros morfológicos como el tipo histológico, tamaño tumoral, grado histológico y afectación ganglionar se han correlacionado con los resultados clínicos y son factores pronósticos de la enfermedad (41). Los marcadores IHQ como la expresión de receptores de estrógenos (RE) y de progesterona (RPg) y la sobreexpresión y/o amplificación del receptor del factor de crecimiento epidérmico (HER2) proporcionan además un valor predictivo y son importantes para la toma de decisión en el tratamiento (42).

Por otra parte, en los últimos años se han descrito clasificaciones moleculares, normalmente basadas en el análisis de expresión génica con tecnología de alto rendimiento. A continuación ambos tipos de clasificación: la clásica histopatológica y la molecular.

a) Parámetros incluidos en la clasificación histopatológicos

- *Subtipos histológicos*

Según la clasificación de los tumores de la Organización Mundial de la Salud y basándose en las características citológicas y en los patrones arquitecturales, el 70-80% de los carcinomas de mama se describen como carcinoma invasivo de tipo no especial (CINST). El resto se consideran subtipos histológicos específicos, como lobular invasivo,

medular, tubular, metaplásico etc. El diagnóstico se realiza por exclusión de tipos específicos reconocidos de cáncer de mama (43).

- ***Grado histológico***

El grado histológico está basado en el grado de diferenciación del tejido tumoral. En el cáncer de mama se realiza una evaluación semicuantitativa de las características morfológicas. En la actualidad se recomienda el Sistema de Gradación de Nottingham (SGN) que valora la diferenciación celular y la actividad proliferativa. Está basado en tres características morfológicas: grado de formación glandular o tubular, pleomorfismo nuclear, e índice mitótico (41). Este SGN permite la estratificación tumoral proporcionando información pronóstica (44).

- ***RE, RPg y HER-2***

Además de la valoración histopatológica, de forma rutinaria se realiza la valoración de la expresión de los RE, RPg y HER 2 en tejido tumoral por IHQ. El 75-80% de los casos de cáncer de mama presenta receptores hormonales (RH) positivos, el 10-15% tiene HER2 sobre-expresado y/o amplificado y la mitad de estos expresan además receptores hormonales (45). El 10-15% restante se definen como triple negativo (ausencia de expresión de receptores hormonales y HER2) (46).

b) Clasificaciones moleculares del cáncer de mama

- ***Subtipos intrínsecos***

El análisis de la expresión génica en el cáncer de mama por tecnología basada en matrices de ADN ha permitido la identificación de subtipos moleculares con relevancia clínica y ha proporcionado información sobre la heterogeneidad molecular de esta enfermedad. Estos subtipos son el luminal A, luminal B, enriquecido en HER2, basal-like y normal.

El subtipo luminal A representa el 50-60% del total de los tumores de mama y se caracteriza por la expresión de genes activados por el receptor de estrógenos como factor de transcripción y presentan una baja expresión de genes relacionados con la proliferación celular (47,48)

A nivel patológico e inmunohistoquímico presentan un Ki 67 y grado histológico bajo y expresan RE, RPg, Bcl-2 y citoqueratinas (CK) del epitelio luminal CK8/18. El punto

de corte de expresión de Ki 67 para diferenciar entre luminal A y B se estableció en 13,25% (49). Presentan mejor pronóstico con menor número de recaídas y con una mayor supervivencia. La tasa de recaídas es del 27,8% con una supervivencia prolongada de 2,2 años. El patrón de recaída es diferente al resto de los subtipos con más recaídas a nivel óseo y con metástasis viscerales en menos del 10% de los casos. Responden al tratamiento con hormonoterapia (50).

El subtipo luminal B representa un 10-30% de los tumores de cáncer de mama. Su fenotipo es más agresivo. Son tumores con un grado histológico mayor, un índice proliferativo más elevado (Ki 67 superior a 14%) (49) y peor pronóstico que los luminales A. Tienen un patrón de recaídas a distancia diferente, siendo la localización en el hueso cercana al 30% y las metástasis viscerales a nivel hepático representan el 13,8%. Desde el diagnóstico de la recidiva tienen una peor supervivencia que los luminales A con una mediana de 1,6 años (50). Dado que simultáneamente pueden sobre-expresar HER2, el carcinoma luminal B se divide en dos subgrupos: luminal B/HER2 negativo y luminal B/HER2 positivo. A nivel de expresión génica, presentan niveles bajos de receptores estrógenos y baja expresión de genes relacionados con RE, alta expresión de genes relacionados con el ciclo celular y proliferación (47,48)

El subtipo HER2 enriquecido representa el 4-7%. Se caracteriza por la amplificación o sobreexpresión del oncogén Her2 y por genes localizados en el amplicón Her2 del cromosoma 17q12. A nivel histológico e IHQ presenta sobreexpresión de Her2, RE y RPg negativos y un alto grado histológico. Comparado con los subtipos luminales presenta una mayor incidencia de enfermedad metastásica al diagnóstico y mayor riesgo de metástasis cerebrales, pulmonares y hepáticas, con un porcentaje mayor de recaídas y peor supervivencia (51,52). La introducción de dianas terapéuticas anti-Her2, principalmente trastuzumab y pertuzumab en combinación con quimioterapia, ha mejorado los resultados de este subtipo, tanto en enfermedad localizada como metastásica (53,54).

El subtipo *normal-breast* representa alrededor del 5-10% de los casos de cáncer de mama (48). Expresan genes característicos de tejido adiposo y presentan un pronóstico intermedio entre el subtipo luminal y el *basal-like*. No expresan receptores hormonales ni HER2, por lo que pueden clasificarse como tumores triple negativos, pero no son *basal-like*. En la actualidad este subtipo se cree que podría ser un artefacto técnico en los estudios de expresión génica por elevada contaminación por tejido normal (55).

El subtipo *basal-like* representa entre el 10-20% de todos los tumores de mama. Se ca-

racteriza por la sobreexpresión de genes relacionados con las células de la membrana basal de los ductos mamarios (mioepiteliales) incluyendo CK5/6, CK17, integrina b4, laminina, proteína 7 unida a ácidos grasos y receptor del factor de crecimiento epidérmico (EGFR) (47,48). Son de alto grado histológico, y por IHQ son positivos para CK5/6, CK17 y presentan un índice Ki 67 elevado. No suelen expresar receptores hormonales ni HER2, por lo que a veces se les define como carcinoma de mama triple negativo (CMTN). Sin embargo, los grupos basal y triple negativo no son equivalentes: además del hecho de que uno se refiere a una clasificación molecular y el otro a una clasificación histopatológica, se ha descrito una discordancia entre ambos (56). Desde el punto de vista clínico, los tumores basales se presentan en mujeres más jóvenes, con mayor incidencia en mujeres africanas, con un tamaño tumoral mayor de 2 cm en más del 50% de los casos, grado histológico elevado y mayor afectación ganglionar. A nivel morfológico, la mayoría suelen ser carcinomas ductales infiltrantes, con un elevado índice mitótico, necrosis tumoral asociada, márgenes expansivos y con abundante infiltración linfocítica estromal (57). El mayor riesgo de diseminación a distancia es en los 2-3 primeros años desde el diagnóstico disminuyendo a partir de los 5 años (58). La recurrencia es más frecuente a nivel pulmonar, sistema nervioso central y metástasis ganglionar a distancia (50,59), Presenta un mal pronóstico y mayor agresividad que otros subtipos intrínsecos, probablemente por el elevado índice de mutaciones en p53 (60). A nivel IHQ se realizan 5 paneles de anticuerpos para clasificar este subtipo tumoral: RE, RPg, Her2, EGFR y CK5/6 con una especificidad del 100% y una sensibilidad del 76% (61). Pero esta definición de *basal-like*, no tiene en el momento actual implicaciones terapéuticas, considerándose como triple negativo en general ya que más del 90% de los cáncer de mama *basal-like* son triple negativos (62). Por el contrario, el cáncer de mama *basal-like* representa entre el 55-81% del subtipo triple negativo (63,64). En los últimos años se han publicado varias clasificaciones moleculares de la enfermedad triple negativa basadas en los resultados de estudios de genómica y de transcriptómica. Estas clasificaciones son biológicamente diversas y todavía no se sabe si alguna de ellas puede llegar a tener utilidad clínica (65–67).

Para añadir más complejidad al tema, los tumores de mama que se originan en pacientes portadoras de la mutación en línea germinal BRCA1 tienen características moleculares del cáncer de mama esporádico subtipo *basal-like* (68), con elevada probabilidad de

presentar un alto grado, receptores hormonales negativos, her2 negativo y mutaciones en TP53 (64).

El subtipo bajo en claudina (*claudin-low*) representa menos del 5% de todos los tumores de mama. Se describió en el año 2007 por Herschkowitz y cols. (69). En la agrupación jerárquica se localiza junto al subtipo *basal-like*, lo que implica que ambas subtipos tumorales comparten algunas características en cuanto a expresión génica. Se caracterizan por expresión de marcadores de diferenciación luminal baja o ausente, marcadores de la transición epitelio mesénquima (TEM) altamente enriquecidos, genes de respuesta inmune y características similares a las células troncales del cáncer de mama (*stem cell-like*). Presenta una baja expresión de genes de adhesión celular como las claudinas 3, 4, 7, la cingulina y la ocludina, localizadas en las uniones estrechas y la e-cadherina, una proteína de adhesión celular dependiente de calcio. Clínicamente, la mayoría de estos tumores son carcinomas ductales infiltrantes, presentan con frecuencia una elevada diferenciación medular y metaplásica, ausencia de RE, RPg y Her 2 (triple negativos) y un elevado grado histológico (70). Sin embargo, cerca de un 20% expresan receptores hormonales por IHC y presentan mal pronóstico (71).

- ***Otras clasificaciones moleculares***

El proyecto METABRIC ha descrito otros subtipos moleculares combinando los datos obtenidos por análisis del genoma y del transcriptoma utilizando diferentes plataformas como las matrices de número de copias de ADN genómico, la metilación de ADN, la secuenciación de exoma, las matrices de ARNm, la secuenciación de microARN y las matrices de proteínas de fase inversa (RPPA) (72) (Tabla 4).

La red *The Cancer Genome Atlas* (TCGA), utilizando las mismas plataformas que el proyecto METABRIC, ha identificado cuatro clases principales de cáncer de mama. Las mutaciones somáticas en los genes TP53, PIK3CA y GATA3 presentan una incidencia > 10%. Se identificaron dos nuevos subgrupos definidos por la expresión de proteínas, posiblemente relacionados con elementos estromales y microambientales y se identificaron rutas de señalización específicas dominantes en cada subtipo molecular (73). Estas clasificaciones demuestran de nuevo la gran heterogeneidad genética del cáncer de mama y en nuestro caso de los tumores con receptores hormonales positivos.

Tabla 4. Características de los subtipos *Integrative Cluster* del Proyecto METABRIC (72)

Subtipo	%Pacientes	Características	Grado	Pronóstico	SCE (5 años)
IntClust 1	7%	Luminales B	Alto	Intermedio	
IntClust 2	4%	Luminal A y B Inestabilidad genómica elevada Genes Ciclo celular G1/S enriquecidos		Pobre	50%
IntClust3	15%	Luminal A Afectación ganglionar baja Inestabilidad genómica baja Baja prevalencia de CNA y mutación p53	Bajo	Bueno	90%
IntClust4	17%	25% TN CNA escasa	Bajo	Bueno	
IntClust 5	10%	HER2 Edad más precoz Afectación ganglionar elevada	Alto	Pobre	
IntClust 6	4%	Luminales B No características diferentes		Intermedio	
IntClust 7	10%	Luminales A Mujeres ancianas	Bajo	Bueno	80%
IntClust 8	15%	Luminales A Mujeres ancianas	Bajo	Bueno	80%
IntClust 9	7%	Luminales B	Alto	Intermedio	
IntClust 10	11%	<i>Basal-like</i> TN, G3, Mutaciones p53 elevadas	Alto	Pobre	

CNA: número de copias aberrantes; SCE: Supervivencia cáncer específica.

1.5. FACTORES PRONÓSTICOS Y PREDICTIVOS

La selección del tratamiento adyuvante sistémico en las pacientes con cáncer de mama precoz se realiza en función de los factores pronósticos y predictivos. Los factores pronósticos son aquellos que se obtienen en el momento del diagnóstico, o en el momento de la cirugía, que en ausencia de tratamiento adyuvante están asociados con el riesgo de recurrencia, muerte y otros resultados clínicos y los factores predictivos son los asociados con el grado de respuesta a un tratamiento específico. Cualquier factor puede ser pronóstico y predictivo y su importancia depende del criterio de valoración clínica y del método de comparación del tratamiento (74).

a) Edad

La edad es un factor pronóstico independiente. La edad media del diagnóstico de cáncer de mama es entre los 60 y 65 años. Estudios de base poblacional relacionan a las mujeres menores de 35 años con peor supervivencia, porque presentan tumores más agresivos, de mayor tamaño, con invasión vascular y alto grado histológico. Las mujeres mayores de 70 años presentan tumores con receptores de estrógenos positivos, con un crecimiento tumoral más lento y mayor respuesta a la hormonoterapia (75). Hay un subgrupo de pacientes, que si se diagnostican en edad joven y no reciben tratamiento con quimioterapia presentan un mal pronóstico (76).

b) Tamaño del tumor

El tamaño tumoral es un factor pronóstico independiente del estado de los ganglios axilares. Tumores menores de 2 cm con ganglios negativos presentan unas tasas de supervivencia alrededor del 95% y menor al 46% si el tamaño tumoral es de 5 cm y presenta afectación axilar (77). Los tumores menores de 1 cm sin afectación axilar presentan un excelente pronóstico, con supervivencias del 92% después de 8 años, según un subanálisis conjunto realizado de las pacientes incluidas en varios ensayos clínicos con dichas características (78).

c) Ganglios axilares

La presencia o ausencia de infiltración tumoral de los ganglios axilares en el cáncer de mama es el factor pronóstico más importante. El número de ganglios axilares afectados se relaciona de forma inversa con la supervivencia a largo plazo. De hecho, la supervivencia a 10 años disminuye del 75% en ganglios no afectados al 25-30% en ganglios afectados. Tanto la afectación ganglionar como el número de ganglios afectados es indicativo de mal pronóstico (79).

d) Variedad histológica

El tipo histológico es otro factor pronóstico. Existen tipos histológicos minoritarios como el carcinoma tubular, mucinoso y el adenoide quístico que tienen un buen pronóstico (80,81).

e) Grado histológico

El grado histológico (G) es un factor pronóstico independiente bien establecido. El método más utilizado es el propuesto por Scarff-Bloom-Richardson (82), modificado de Nottingham propuesto por Elston y Ellis (41). Los tumores con alta grado histológico (grado 3) presentan un peor pronóstico. El grado histológico se asocia significativamente con la supervivencia libre de enfermedad y la supervivencia específica por cáncer de mama en relación con el tamaño tumoral y la afectación axilar (83).

f) Invasión linfovascular

La invasión linfovascular es un factor pronóstico negativo en pacientes con ganglios linfáticos negativos (84) así como un predictor independiente de supervivencia libre de metástasis a distancia y de la supervivencia específica por cáncer de mama (85).

g) Receptores hormonales

Los receptores hormonales de estrógenos y de progesterona junto con el receptor de crecimiento epidérmico Her2 son los factores pronósticos y predictivos más importantes. Las pacientes que presentan receptores positivos tienen una probabilidad de supervivencia libre de enfermedad un 10% superior que las que presentan receptores negativos, con una mediana de seguimiento a 5 años (86). Como factores predictivos los tumores que expresan los receptores se benefician del tratamiento con hormonoterapia, no así los tumores con receptores negativos. El 20% de las pacientes presentan amplificación o sobreexpresión de Her2 y se asocia con tumores de mayor grado tumoral, ausencia de receptores y mayor proliferación y peor pronóstico. Her2 es el factor predictivo de respuesta a trastuzumab (54,87).

h) Ki67

Ki67 es un marcador de proliferación celular que se ha asociado con peor pronóstico (88). Se ha investigado como un potencial factor pronóstico y predictivo (89), pero no se recomienda como guía para la selección del tratamiento adyuvante con quimioterapia (42). Se ha establecido como un parámetro para dividir al subgrupo luminal en A o B con un punto de corte de 13,25 % (49). Pero existe controversia respecto a este punto de corte para valorar o no el tratamiento con quimioterapia en las pacientes con tumores

hormonales positivos. Estos puntos de corte se han ido modificando en los sucesivos consensos de St. Gallen recomendando un Ki 67 <14 en el año 2011 (90), un Ki 67 < 20 en el año 2013 (91) y en el consenso de 2015 se aceptó la definición de luminal B con una determinación de Ki 67 superior al 20-29% (92). La elección del punto de corte influye en la proporción de pacientes que se consideran de bajo riesgo y subsecuentemente en la decisión del tratamiento (93). Se recomienda que el punto de corte se interprete de acuerdo a los resultados del laboratorio local, teniendo en cuenta la mediana de cada centro, considerándolos elevados o disminuidos si están por encima o por debajo del 10% de esta mediana (92). Además hay que tener en cuenta que la determinación por IHQ presenta una variabilidad intra e interobservador elevada (94).

i) Perfiles moleculares

El nivel de expresión de múltiples genes (perfiles moleculares) medido por plataformas específicas se ha venido empleando desde hace más de diez años en la clínica. Estas plataformas de expresión génica (PEG) deben estar validadas de acuerdo a los criterios del Grupo de Trabajo de Evaluación de Aplicaciones Genómicas en la Práctica y Prevención (de sus siglas en inglés EGAPPWG) (95,96). Se comentan las más importantes: OncoType DX, MammaPrint, Prosigna y EndoPredict.

- ***OncoType DX***

OncoType DX se desarrolló tras analizar los datos de 3 estudios clínicos independientes de pacientes con cáncer de mama y largo seguimiento (97). El test está basado en la expresión de 16 genes relacionados con el tumor y 5 genes de referencia (Tabla 5) expresando el resultado (rango entre 0-100) como un *Recurrence Score* (RS) tras la realización de un modelo matemático. Este RS estima la probabilidad de recurrencia de la enfermedad a 10 años. Clasifica a las pacientes en tres categorías de riesgo: bajo (RS <18), intermedio (RS 18-30) y alto (RS \geq 31) (97). La mayor expresión de genes asociados con un buen pronóstico (ER, GSTM1, BAG1) se relacionan con un RS bajo, mientras que la mayor expresión de los genes del grupo de proliferación, del grupo Her2 positivo, el grupo de invasión y el CD 68 se relacionan con un RS más elevado y peor pronóstico.

En pacientes con RE+, N0 y tratamiento con hormonoterapia (tamoxifeno o anastrozol) la validación clínica del papel pronóstico de OncoType DX, se evaluó en 4 estudios (97–

100). En todos ellos, el riesgo de recurrencia a distancia a los 10 años fue menor en las pacientes con RS bajo y más elevado en las pacientes con un RS alto, resultando independiente de la edad y del tamaño tumoral (97).

La validación clínica como predictor de beneficio a la quimioterapia en N0 se realizó en muestras tumorales de un subgrupo de pacientes del estudio NSABP-B20. Las pacientes que habían recibido QT con un RS alto (≥ 31) presentaron una reducción del riesgo de recurrencia a 10 años (RR= 0,26) y una reducción absoluta del riesgo de recidiva sistémica a 10 años del 27,6%. Las pacientes con RS bajo no se beneficiaron de la QT (RR= 1,31). En el grupo de riesgo intermedio el beneficio de la quimioterapia no puede excluirse (101).

Para validar la utilidad clínica de Oncotype Dx en pacientes con N0 se diseñó el ensayo clínico fase III (TAILORx). El objetivo primario era determinar si las pacientes con un RS intermedio (11-25) se benefician de añadir QT al tratamiento HT (102). En 2015 se publicaron los resultados en relación a las pacientes de bajo riesgo (RS < 11), tratadas únicamente con HT. Con un seguimiento de 5 años, tienen un pronóstico excelente con una SLE del 98,7% y SG del 98% (103).

OncoType DX también se ha validado en pacientes con ganglios axilares positivos tratadas con hormonoterapia adyuvante dentro de los estudios trans-ATAC (104), SWOG 8817 (105) y ECOG E2197 (106). Otros ensayos están analizando la plataforma en esta misma población con 1 a 3 ganglios axilares: RxPONDER/SWOG S1007 (107), West German Study Group PlanB trial (WGSG PlanB trial) (108). Además se han publicado los datos de dos registros de base poblacional basados en la práctica clínica real (109,110) que corroboran los datos de los estudios de ganglios positivos (104,111). OncoType DX cuenta con la aprobación de la Sociedad Americana de Oncología –ASCO– para pacientes con receptores de estrógenos positivos y ganglios negativos. (42).

Tabla 5. Genes incluidos en la plataforma Oncotype DX

Proliferación	Invasión	Estrógenos	Her2	Otros	Referencia
KI 67	Stromelisin 3	RE	GRB7	GSTM1	Beta Actina
STK15	Catepsina L2	Rpg	HER2	BAG1	GAPDH
Survivina		Bcl2		CD68	RPLPO
Ciclina B1		SCUBE2			GUS
MYBL2					TFRC

- **MammaPrint**

MammaPrint analiza 70 genes mediante tecnología de *microarrays* de ADN para la cuantificación de expresión génica (Tabla 6). Se desarrolló en una cohorte de 78 pacientes con cáncer de mama precoz menores de 55 años, T<5 cm, N0 y sin tratamiento adyuvante la mayoría de las pacientes. Clasifica a las pacientes en dos grupos: alto y bajo riesgo. (112).

La primera validación clínica de la *firma de 70 genes* como factor pronóstico se realizó en 295 muestras tumorales de tejido fresco, de las cuales el 49% eran N+. Las pacientes clasificadas de bajo riesgo presentaron una mediana de supervivencia a 10 años del 94,6% y las de alto riesgo del 54,6%. A su vez se comparó con los criterios clínico patológicos utilizados en ese momento como St. Gallen (113) y del NIH (74) y se confirmó que seleccionaba a las pacientes mejor que los factores clínico patológicos clásicos, identificando pacientes de buen pronóstico en el grupo de alto riesgo clasificados según las recomendaciones clásicas y predijo mejor la recurrencia a distancia a 5 años (114).

La primera validación multicéntrica independiente se realizó en 307 muestras de pacientes incluidas por el consorcio TRANSBIG. El riesgo clínico se estimó con la herramienta informática Adjuvant!Online. Las pacientes de bajo riesgo por perfil genómico tenían una tasa de supervivencia a los 10 años del 88% para el grupo de bajo riesgo clínico y del 89% para el de alto riesgo clínico. Las pacientes de alto riesgo por perfil genómico tuvieron una supervivencia a los 10 años del 69%, tanto para el bajo y alto riesgo clínico. Estos datos sugerían que el perfil de 70 genes predice la evolución de la enfermedad con independencia de los criterios clínico-patológicos. (115)

El primer estudio prospectivo de validación de esta firma génica en N0 fue el estudio

RASTER (116). El 81% con alto riesgo y el 15% de bajo riesgo por perfil génico recibieron quimioterapia adyuvante. La supervivencia a 5 años sin recurrencia a distancia en las pacientes de bajo riesgo por perfil génico y alto riesgo por criterios clínicos que no recibieron QT fue del 97% comparada con el 91,7% de aquellas de alto riesgo, lo que confirmó su valor pronóstico. También se ha confirmado que esta firma génica es un factor pronóstico independiente en estudios más recientes (117) y en varios estudios con pacientes postmenopáusicas. (118)

La validación clínica en pacientes con N+ (1-3) se ha realizado en estudios retrospectivos (119) objetivándose una supervivencia libre de metástasis a distancia a los 10 años (SLMD) del 91% para el grupo de bajo riesgo y del 76% para el de alto riesgo según la firma génica. La firma génica fue superior que los factores pronósticos clásicos en predecir la supervivencia específica por cáncer de mama (119).

La validación de la utilidad clínica se ha realizado con el ensayo clínico MINDACT (120). El riesgo genómico de las pacientes se determinó por MammaPrint y el riesgo clínico utilizando una versión modificada de Adjuvant! Online. Se realizaron 4 subgrupos: el grupo de bajo riesgo clínico y genómico recibió hormonoterapia; el grupo de alto riesgo clínico y genómico recibió quimioterapia; cuando los resultados no coincidían, las pacientes se aleatorizaban a recibir o no quimioterapia. La supervivencia libre de metástasis a distancia (SLMD) a los 5 años en el subgrupo de alto riesgo clínico y bajo riesgo genómico que no recibían quimioterapia fue del 94,7%. Los resultados muestran que alrededor del 46% de mujeres con cáncer de mama en estadio temprano N0 o N+ (1-3) pueden evitar de forma segura la QT adyuvante tras el análisis de la expresión génica de su enfermedad. Con estos datos se ha recomendado por la guía ASCO que el test MammaPrint puede ser usado en la toma de decisión del tratamiento adyuvante con quimioterapia en pacientes con receptores positivos, N0 y en pacientes seleccionados con N+ (121).

Tabla 6. Función de los genes incluidos en la plataforma MammaPrint

Fenotipo de las células cancerosas/Función	Número de genes
Desregulación del ciclo celular	15
Angiogénesis	12
Proliferación y transformación oncogénica	11
Invasión y metástasis	8
Transducción de la señal del factor de crecimiento	6
Resistencia a la apoptosis	2
Miscelánea/Función desconocida	16

- **EndoPredict**

El test genómico EndoPredict (EP) analiza 8 genes asociados y 3 genes de control (Tabla 8). Se ha desarrollado a partir de una serie de 964 muestras tumorales de cáncer de mama precoz, con receptores hormonales positivos y HER2 negativo y tratadas solo con tamoxifeno. Se analizaron más de 20.000 genes, identificándose doce genes que ofrecían una información más detallada sobre el grado de agresividad del tumor. A partir de determinados niveles de actividad de los 8 genes seleccionados (BIRC5, RBBP8, UBE2C, IL6ST, AZGP1, DHCR7, MGP y STC2), mediante una fórmula matemática se calcula una puntuación EP mediante una escala del 0 al 15. Una puntuación > de 5 clasifica a las pacientes en el grupo de alto riesgo (122).

EndoPredict también incluye en el cálculo del riesgo el tamaño del tumor y el estado de los ganglios linfáticos. La combinación de EP con estos dos factores de riesgo clínicos resulta en el score EPclin. El umbral establecido de EPclin es 3.3. Las pacientes con un EP score <5 (EPclin score < 3.3) se clasificaron como de riesgo bajo para recurrencia a distancia, mientras que con un EP >5 (EPclin>3.3) se clasificaron como alto riesgo.

La validación clínica se efectuó en un análisis conjunto de dos ensayos clínicos independientes realizados por el Grupo Austriaco de Estudio del Cáncer de Mama y Colo-rectal (ABCSG): el ABCSG-6 (123) y el ABCSG-8 (124). Todas las pacientes recibieron tratamiento con tamoxifeno 5 años o tamoxifeno 2 años seguido de anastrozol por 3 años. El objetivo primario era la recurrencia a distancia precoz (<5 años) o tardía (>5 años). Se incluyeron 1.702 pacientes. Con un seguimiento de 7.12 años, 998 pacientes estaban en riesgo después de 5 años. Las pacientes del grupo de bajo riesgo presentaron

una tasa del 5% de metástasis a distancia entre los 5 y 10 años de seguimiento (125). El score EPclin es capaz de predecir un pronóstico favorable en la mayoría de las pacientes a las que las guías clínicas asignan un riesgo intermedio o alto (126). Se ha realizado un estudio descentralizado que corrobora la reproducibilidad de estos resultados (127).

Asimismo se ha realizado la validación clínica del test en muestras tumorales de pacientes con N0 en el ensayo clínico GEICAM 9906. Es un estudio prospectivo-retrospectivo, en el cual se incluyeron 1.246 pacientes, aleatorizadas a dos esquemas de QT. El estudio concluye que el test genómico Endopredict proporciona un valor pronóstico independiente en la población analizada.(128).

- **Prosigna**

Prosigna (PAM50) (Prediction Analysis of Microarrays 50) (129) es un predictor genómico basado en la expresión de 50 genes (Figura 1) que identifica los subtipos moleculares luminal A, luminal B, HER2 enriquecido y basal (48). También identifica la categoría *normal-like*. No identifica el subtipo *claudin-low* (70). Inicialmente se realizó utilizando RT-qPCR, pero en la actualidad se realiza con tecnología nCounter Analysis System, denominado test Prosigna (130). Además del subtipo intrínseco del tumor, proporciona un score del riesgo de recurrencia a distancia a 10 años (ROR) en una escala del 0-100: pacientes de bajo riesgo ROR < 40 (menos del 10% de riesgo); riesgo intermedio ROR 40-60 (10-20% riesgo de recaída); alto riesgo ROR > 60 (más del 20% de riesgo de recaída) y realiza un cálculo individual del riesgo de recaída a distancia a 10 años (48,129). Incluye el tamaño del tumor, el subtipo molecular intrínseco y el estado de proliferación del tumor.

La validación clínica se ha realizado en dos estudios, TransATAC (131) y ABCSG-8 (124). El estudio TransATAC incluyó 1.017 muestras de pacientes con cáncer de mama precoz con N0 y N+ incluidas en el ensayo ATAC, que comparaba el tratamiento adyuvante con anastrozol frente a tamoxifeno. El resultado del estudio mostró que el ROR score añadía información pronóstica significativa respecto a la recurrencia a distancia sobre la aportada por los factores pronósticos clínico patológicos clásicos en la población total y en todos los subgrupos analizados. La relación entre ROR score y el riesgo de recaída a distancia varió dependiendo del estado ganglionar (131). También se ha demostrado su capacidad pronóstica de recaída a distancia en el periodo entre los 5 y 10 años tras la cirugía (132). Del estudio ABCSG-8, se incluyeron 1.478 muestras. Las pacientes habían recibido como tratamiento tamoxifeno o tamoxifeno seguido de anas-

trozol con un seguimiento a 10 años. Los resultados confirmaron que el ROR score proporciona valor pronóstico adicional significativo sobre los factores clínico patológicos clásicos para predecir la recurrencia a distancia (133). Un análisis conjunto de ambos estudios, demostró el valor pronóstico independiente de PAM50, además de lo aportado por las características clínicas y patológicas clásicas, en mujeres con carcinoma de mama con afectación ganglionar y receptores hormonales positivos (134).

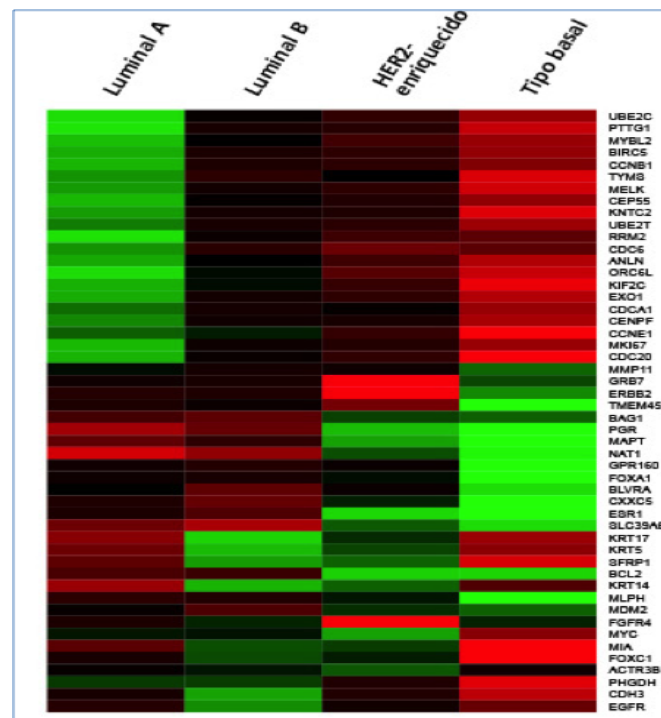


Figura 1. Genes incluidos en la plataforma PAM 50 (www.prosigna.es)

Se han descrito otros perfiles moleculares, como Breast Cancer Index (135,136), Genomic Grade Index (137) o IHC Score (94,138), pero se emplean poco y no tienen ventajas sobre los ya referidos. Se proporciona un resumen comparativo de los perfiles moleculares más utilizados en la tabla 7.

Tabla 7. Características de las Plataformas de Expresión Génica más utilizadas (Modificado de (139,140)).

Plataformas de Expresión Génica	Oncotype DX	MammaPrint	Endopredict	Prosigna	Breast Cancer Index
Genes	21 genes (16 genes relacionados +5 controles)	70 genes	11 Genes (8 relacionados +3 controles) + T y N	PAM50 (50 genes clasificadores + 5 controles)	2 genes + 5 genes (MGI)
Marcador	Molecular	Molecular	Molecular	Molecular	Molecular
Tecnología	qRT-PCR	Matrices DNA	qRT-PCR	qRT-PCR (nCounter RNA)	qRT-PCR
Muestra Biológica	FFPE	Tejido Fresco; FFPE (2012)	FFPE	FFPE	FFPE
Análisis Centralizado	Si	Si	No	No	Si
Índice Pronóstico	RS	Índice Mammaprint	EP/EPclin	ROR	BCI
Indicación	Pronóstico Predictivo	Pronóstico Predictivo	Pronóstico	Pronóstico Clasificación por subtipos	Pronóstico Predictivo (Recur tardía)
Población a estudio	N0-1; RE+	N0-1; RE +; < 62 años	N0-1; RE +	N0-1; RE +; PostM	RE;+ N0
Categorías de Riesgo	BR: RS<18. RI: RS 18-30. AR: RS>=31.	BR AR	BR <5/<3.3 AR >5/>3.3	BR <40 RI 40-60 AR >60	BR:<5 RI: 5-6.4 AR: >6.4
Validación clínica	Si	Si	Si	Si	Si
Evidencia Prospectiva/Retrospectiva	Kaiser NSABP B-B14/B-20 ECOG 9127 (N+) TransATAC TAYLORx WSG-PlanB, ADAP RxPONDER	TRANSBIG RASTER MINDACT	ABCSG6 ABCSG8 GEICAM 9906	NCIC MA.12 NCIC MA.5 ABCSG8/TransATAC RxPONDER	MA.17
Recomendación ASCO	SI	SI	SI	SI	SI

PEG: Plataformas de expresión génica; RE, receptor de estrógenos; FFPE, fijado en formalina, embebido en parafina; FISH, hibridación fluorescente in situ; IHQ, inmunohistoquímica; RPg, receptor de progesterona; N:D: No disponible; RT-PCR, reacción en cadena de polimerasa de transcripción inversa. N0: ganglios negativos; N1: 1-3 ganglios positivos; BR: bajo riesgo; RI: riesgo intermedio; AR: alto riesgo MGI: Índice de grado molecular.

1.6. GENERALIDADES DEL TRATAMIENTO DEL CÁNCER DE MAMA

Las herramientas disponibles para tratar el cáncer de mama son la cirugía y la radioterapia como tratamiento local y la quimioterapia, la hormonoterapia y las terapias biológicas como tratamientos sistémicos. La elección de estos tratamientos se realiza dependiendo de si la enfermedad es loco-regional, es decir limitada a la mama y/o ganglios linfáticos regionales o es una enfermedad metastásica. En la enfermedad localizada la cirugía puede ser conservadora o radical (mastectomía) asociada a la realización de una biopsia selectiva del ganglio centinela y/o linfadenectomía axilar. Esta puede realizarse previa a o tras un tratamiento sistémico. La radioterapia está indicada tras la realización de la cirugía como tratamiento complementario si se cumplen una serie de criterios clínico patológicos y como tratamiento paliativo en la enfermedad metastásica (92). La decisión del tratamiento sistémico se realiza teniendo en cuenta determinados factores como el riesgo individual de recaída y la sensibilidad a un tratamiento específico. Para la selección del tratamiento se tienen en cuenta las características clínicas clásicas (el estadio, la edad, el estado menstrual), la biología tumoral (expresión de los receptores hormonales y Her2), el grado y la expresión de Ki 67 y el subtipo molecular. En la toma de decisión también se valoran las comorbilidades de la paciente, la toxicidad del tratamiento, la expectativa de vida y las preferencias de la paciente.

El meta-análisis del *Early Breast Cancer Trialists' Collaborative Group* (EBCTCG) de 2005 ha demostrado el beneficio del tratamiento sistémico en la supervivencia de las pacientes con cáncer de mama tanto con ganglios positivos (N+) como con ganglios negativos (N0) con el empleo de quimioterapia y hormonoterapia. Respecto a los esquemas de tratamiento con quimioterapia, en la actualidad se utilizan aquellos basados en combinaciones de antraciclinas y taxanos que han demostrado ser más eficaces que las combinaciones con antraciclinas (141). Para facilitar la toma de decisiones, se dispone de guías clínicas nacionales e internacionales realizadas por expertos, y que reflejan la evidencia científica del momento. También se dispone de herramientas informáticas que tras incluir los datos clínico-patológicos y la opción de tratamiento, proporcionan la probabilidad de recaída y supervivencia a 5 y 10 años como Predict (<http://www.predict.nhs.uk>) (142).

En la conferencia de consenso de St. Gallen 2017 los expertos están de acuerdo en que la quimioterapia no tiene papel en los casos de bajo riesgo clínico como pT1a/b, grado 1, RE elevados y N0. Se recomienda que las PEG sirvan como marcadores pronósticos en el contexto de la terapia endocrina adyuvante en cáncer de mama sin afectación ganglionar. Todos los casos identificados de bajo riesgo tienen un excelente pronóstico y no precisarían tratamiento con quimioterapia. Independientemente de la expresión génica, el estado ganglionar es un factor pronóstico. Los expertos concluyen que las PEG pueden redefinir el pronóstico del cáncer de mama con afectación ganglionar, pero no hubo un consenso respecto a su utilización en la toma de decisión del tratamiento con quimioterapia adyuvante en los casos con ganglios positivos.

A grandes rasgos el tratamiento del cáncer de mama infiltrante con receptores hormonales positivos y HER2 negativo consiste en hormonoterapia asociada o no a quimioterapia. En los tumores triple negativos el tratamiento sistémico es la quimioterapia basada en antraciclinas y taxanos, tanto en la enfermedad localizada (adyuvancia, neoadyuvancia) como en la enfermedad metastásica. En los tumores con expresión o amplificación de HER2 el tratamiento sistémico con quimioterapia debe incluir además tratamiento con anticuerpos anti-HER2: trastuzumab con o sin pertuzumab (54,143–146) (Tabla 8).

Tabla 8. Subtipos de Cáncer de Mama definidos por IHQ y tratamiento

Subtipo	Características IHQ	Tratamiento
Luminal A	RE positivos, RPg positivos >20%, HER-2 negativo y Ki 67 bajo.	Hormonoterapia
Luminal B	RE positivos, RPg <20%, HER2 positivo o negativo y Ki 67 bajo.	Hormonoterapia +- Quimioterapia
HER 2 enriquecido	HER 2 positivo y RE /RPg negativos.	Quimioterapia + Fármacos antiHER2
Triple negativo:	RE negativo, RPg negativo y HER2	Quimioterapia

1.7 PROTEÓMICA EN CÁNCER DE MAMA.

a) Generalidades

El genoma humano tiene aproximadamente 25.000 genes codificantes de proteínas (147). Un gen puede codificar múltiples proteínas diferentes (isoformas) las cuales se traducen a más de 1,5 millones de proteínas como resultado del procesamiento alternativo (*alternative splicing*), la presencia de polimorfismos secuenciales, modificaciones postraduccionales y otros mecanismos de procesamiento de proteínas (148). Sin embargo la aproximación genómica no captura las modificaciones postraduccionales que afectan a la función y estabilidad de las proteínas (149). En el año 2000 se inició el proyecto *Human Proteomics Initiative* (HPI) cuyo objetivo es registrar todas las proteínas conocidas y describir su estructura, función, localización, modificaciones postraduccionales y semejanzas con otras proteínas de mamíferos (150).

La proteómica se define como el estudio a gran escala, del conjunto de proteínas que se expresan en células, tejidos u órganos en diferentes condiciones espacio-temporales. Abarca diferentes aspectos como su ontología, clasificación, niveles de expresión y sus propiedades. La proteómica funcional estudia la activación de las proteínas, la interacción proteína-proteína, la implicación de las vías de señalización, la cuantificación y su análisis funcional. También implica la identificación de subgrupos de proteínas, tales como quinasas (quinoma), proteínas secretadas (secretoma), proteínas fosforiladas (fosfoproteómica), proteínas exosomales (exosoma) y proteasas (degradoma). A este conjunto de proteínas que se expresan en la célula en unas condiciones determinadas se le denomina proteoma. El análisis de proteínas se puede realizar en muestras biológicas (suero, plasma, saliva, secreción del pezón, orina, líquido cefalorraquídeo, etc.), en tejido (fresco o parafinado) y en cultivos celulares (151). Las proteínas son los efectores moleculares últimos de la función celular. Al contrario que el genoma, el proteoma es dinámico. Tras su síntesis sufre modificaciones denominadas postraduccionales, en determinados momentos y por las interacciones ambientales, que afectan a su estabilidad y función (149).

El objetivo de los estudios de proteómica en el cáncer de mama es poder descifrar las vías de señalización que conllevan el inicio y progresión del cáncer de mama, identificar nuevos biomarcadores moleculares con valor pronóstico y predictivo y descubrir nuevas dianas terapéuticas (152). En la actualidad existen biomarcadores proteicos identificados mediante análisis de proteómica clásica, de utilidad en la práctica clínica, con

diferentes aplicaciones como la expresión de uPA/PAI-1 y el receptor Her-2 (tejido) para la toma de decisión del tratamiento, los receptores de estrógenos y progesterona (tejido) como factor pronóstico y de respuesta al tratamiento, y los biomarcadores en suero CA 15.3 y CA 27-29 para monitorizar la respuesta al tratamiento (42).

b) Plataformas para el análisis de proteínas

Las plataformas proteómicas pueden dividirse en aquellas que precisan del conocimiento previo de la proteína que queremos testar y posteriormente validar con anticuerpos y las que no (149). Las plataformas de validación basadas en anticuerpos incluyen:

- ***Western blot***

Detecta proteínas específicas en una muestra dada de extracto de tejido. Utiliza la electroforesis en gel para separar las proteínas mediante la estructura tridimensional o la longitud del polipéptido. Las proteínas se transfieren posteriormente a una membrana, y se tiñen con los anticuerpos específicos de la proteína. Examina uno o un número limitado de proteínas al tiempo. Es una técnica de bajo rendimiento, laboriosa y cuantitativa (153).

- ***Ensayo inmunoabsorbente ligado a enzima (ELISA)***

Es una técnica altamente específica usada principalmente para análisis en suero de antígenos o anticuerpos. Es adecuada para validar biomarcadores proteicos, pero es cara y de bajo rendimiento y requiere una gran cantidad de equivalentes celulares para ser sensible (154).

- ***Inmunohistoquímica (IHC)***

Es una técnica convencional altamente específica que se realiza en tejido tumoral para evaluar la expresión de proteínas. Proporciona información celular y la localización espacial de las proteínas. Asociada con micromatrices de tejidos (TMA), se pueden analizar un grupo de marcadores de proteínas en múltiples fragmentos de tejido tumoral. Es una técnica de bajo rendimiento, laboriosa, semicuantitativa, y es dependiente del observador (149).

- ***Micromatrices de proteínas (PMA)***

Es un método de alto rendimiento usado para valorar la expresión, interacción y actividad de las proteínas a gran escala. Se pueden analizar una gran cantidad de proteínas en paralelo. Pueden dividirse en matriz de proteínas de fase directa (*Forward Phase Protein Micro-Arrays*) para la fase sólida y matriz de proteínas de fase inversa (*Reverse Phase Protein micro-Arrays*) para la fase líquida de los analitos incluyendo lisados de células tumorales o material de muestras tumorales humanas. La primera también se denomina analítica y la segunda matriz de lisado (154).

c) Espectrometría de masa de proteínas

La espectrometría de masas (EM) de proteínas es una herramienta analítica de alto rendimiento que permite el examen y cuantificación de un gran número de proteínas inicialmente desconocidas. Después de la extracción y separación de las proteínas de la muestra biológica, la EM permite la cuantificación e identificación de las proteínas para poder identificar biomarcadores de relevancia clínica o perfiles proteicos para generar firmas multi-proteicas (155).

El espectrómetro de masas está compuesto por una fuente de iones que fragmenta la muestra en iones generalmente con carga positiva, un analizador de masas que dispersa los iones en función de su relación masa/carga (m/z) y un detector que convierte el haz de iones en una señal eléctrica que puede ser procesada, almacenada y registrada. Los iones generados en la fuente de iones son acelerados hasta el analizador y separados en función de su m/z aplicando campos eléctricos o magnéticos para determinar la masa del ion en el detector. El espectrómetro posteriormente representa las concentraciones de los diferentes componentes de la muestra en un gráfico donde cada línea representa la abundancia de un ion respecto a su relación m/z (156).

El flujo de trabajo del análisis de proteínas por espectrometría de masas consiste en tres pasos: 1) las proteínas son aisladas de la muestra biológica y opcionalmente fraccionadas. La muestra de proteínas es digerida y el péptido resultante es posteriormente fraccionado. 2) Los péptidos son analizados por el espectrómetro de masas de forma cualitativa y cuantitativa. 3) La gran cantidad de datos generados son analizados a través de herramientas informáticas para identificar la secuencia de aminoácidos y si es factible, la cuantificación de las proteínas de la muestra. Los péptidos identificados se cotejan con las bases de datos de proteínas empleando los motores de identificación (148).

En los últimos años, el desarrollo tecnológico y la bioinformática han permitido que la proteómica sea de gran interés para la investigación básica y clínica. Pero la metodología actual no permite realizarla en un solo paso. Dada la complejidad de las técnicas utilizadas en proteómica, a continuación se describe de una manera sucinta la metodología utilizada en el análisis de proteínas por EM.

d) Técnicas de separación o fraccionamiento de proteínas

La complejidad de las muestras biológicas hace necesario que, previo al análisis de las proteínas por EM, se realice la extracción y separación de las proteínas siguiendo unos protocolos determinados para cada tipo de muestra. Este es el primer paso en cualquier estudio de proteómica. Entre los métodos más utilizados para la separación de proteínas se encuentran:

- ***Electroforesis bidimensional (2D-E)***

Permite la separación de proteínas según su punto isoeléctrico (gradiente del pH) en diferentes fracciones en función de la masa molecular. La más utilizada ha sido la electroforesis bidimensional en gel de poliacrilamida (2D-PAGE) que permite el estudio simultáneo de varios cientos de proteínas en un único gel (157). Para visualizar las proteínas es necesario teñir el gel con colorantes (158,159) o realizar un marcaje con fluorocromos característicos de la electroforesis diferencial en gel (*Differential in gel electrophoresis*, DIGE de sus siglas en inglés) (160). Esta técnica mejora la sensibilidad, reproducibilidad y rapidez de la técnica convencional 2D-PAGE, así como la variabilidad intergel y mejora la seguridad de la cuantificación.

- ***Cromatografía líquida (LC) multidimensional***

Permite el fraccionamiento de proteínas con puntos isoeléctricos extremos (ácido o básico), de proteínas hidrofóbicas y de proteínas poco abundantes, de muestras complejas en diferentes fracciones, mediante el acoplamiento de varias columnas cromatográficas. Consiste en una fase estacionaria no polar (columna) y una fase móvil líquida. Puede ser de diferentes tipos: por afinidad, de intercambio iónico, normal, o reversa. Los péptidos en solución se inyectan a través de columnas cromatográficas. Las especies retenidas dependen de las propiedades cromatográficas de la columna, posteriormente se elu-

yen y se evalúan mediante instrumentos de EM (161).

e) Técnicas de análisis de proteínas

Entre las técnicas más utilizadas para el análisis de proteínas tenemos:

- ***EM en “tiempo de vuelo” mediante Ionización-Desorción por Laser Asistida por Matriz (de sus siglas en inglés MALDI-TOF MS)***

Los péptidos ionizados son generados de muestras en fase sólida y analizados por la medida de su tiempo de vuelo, la cual se correlaciona con su masa. La matriz más utilizada es el ácido hidroxicinámico. Puede acoplarse a analizadores de masa TOF, trampa de iones cuadropolar o *ion trap* (QIT), cuadrupolo tiempo de vuelo (Q-TOF) orbitrap y resonancia ciclotrónica de iones con transformación de Fourier (FT-ICR) (162).

- ***EM en “tiempo de vuelo” mediante la ionización por pulverización eléctrica (de sus siglas en inglés ESI-TOF MS)***

La ionización de los péptidos se produce en fase líquida. La ionización por ESI de péptidos y proteínas genera iones multicargados siendo la extensión de la carga dependiente del pH de la solución y del número de grupos básicos o ácidos de la molécula. El ESI puede utilizarse acoplado a cualquier analizador de masas, especialmente a trampas lineales, orbitraps y FT-ICR (162,163).

f) Técnicas de identificación de proteínas

La identificación de proteínas puede realizarse mediante:

- ***Mapeo peptídico o huella de masas peptídicas***

Una proteína aislada se convierte en péptidos de tamaño medio mediante la acción de enzimas específicas, entre otras, la tripsina. La huella es generada por la masa/carga de cada uno de los péptidos ionizados que se compara con la teórica masa/carga de la digestión *in silico* de secuencias de proteínas existentes en bases de datos. El mapeo peptídico se realiza habitualmente por técnicas de MALDI TOF y permite analizar decenas de muestras por hora de forma automatizada (162).

- ***Espectrometría de masas en tándem (tándem MS/MS)***

Permite la secuenciación directa de los aminoácidos que constituyen la proteína. El procedimiento puede aplicarse a la identificación de una única proteína o a la de una mezcla compleja de proteínas, mediante análisis de los péptidos obtenidos tras su digestión. Esta es la base de la técnica *shotgun proteomics*. Los analizadores más utilizados son las trampas de iones y los sistemas híbridos de Q-TOF y Q-LIT.

Respecto a la proporción estimada de falsos positivos en una población de péptidos identificados con una determinada confianza se expresa por la tasa de asignaciones incorrectas o FDR (*False Discovery Rate*) y suelen ser valores $< 2\%$ (162,163).

g) Técnicas de cuantificación de proteínas

Tras la identificación de las proteínas, el siguiente paso es la cuantificación absoluta o relativa de las mismas. Existen diferentes técnicas:

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

Realizan una cuantificación relativa basada en la intensidad de la señal (164) o en la cantidad de espectros (165). En la primera, se comparan las intensidades de los péptidos de una determinada proteína (166). Los cromatogramas de todos los péptidos se obtienen de la carrera de LC-MS, que combina las capacidades de separación física de cromatografía líquida (LC) con las capacidades de análisis de masa de MS. Se calcula el área de la señal obtenida y estos valores se comparan entre muestras para ver su expresión diferencial (164). La segunda técnica se basa en que las proteínas más abundantes tienen más espectros de fragmentación asociados porque sus péptidos se analizan más veces. La cuantificación implica la comparación del número de espectros de la misma proteína entre diferentes muestras (165).

- ***Técnicas con marcaje isotópico***

Permiten la comparación directa de diferentes muestras usando isótopos estables con diferentes masas, un isótopo con masa ligera y otro con masa pesada. Incorporan el isótopo en las proteínas y posteriormente se identifican a partir de las diferencias de sus masas por EM. Pueden identificar cuantitativamente proteínas que no pueden analizarse por electroforesis 2D (163). Los métodos más utilizados son el marcaje con isótopos

estables mediante aminoácidos en cultivos celulares (*Stable Isotope Labeling with Aminoacids in Cell culture*, SILAC), el marcaje de afinidad codificado por isótopos (*Isotope-Coded Affinity Tags*, ICAT) y el marcaje isobárico para cuantificación relativa y absoluta. (*isobaric Tagging for Relative and Absolute Quantification*, iTRAQ) (155). Estas técnicas tienen las desventajas de no permitir la comparación de más de 2 ó 3 muestras en paralelo (salvo el método iTRAQ, en el que se pueden comparar hasta 8), el coste elevado de los reactivos, el marcaje es incompleto, y presentan una gran complejidad de la preparación de la muestra, así como la necesidad de mayores concentraciones de la muestra (166).

- **Monitorización de reacción selectiva (*Selected Reaction Monitoring*, SRM)**

Es el método más fiable en cuanto a la cuantificación de proteínas (165). Requiere un espectrómetro con un analizador de masas triple cuadrupolo. El primer cuadrupolo Q1, selecciona un ion precursor de masa conocida. Este ion pasa al Q2, donde se fragmenta mediante disociación inducida por colisión (CID). Los fragmentos resultantes se analizan en el Q3, y las transiciones precursor-fragmento se emplean para cuantificar el péptido (167). Con este método se pueden realizar cuantificaciones absolutas de proteínas si se añaden péptidos sintéticos marcados. Permite analizar un elevado número de muestras al mismo tiempo (148). En este proyecto de tesis se ha utilizado SRM como técnica de cuantificación de proteínas.

h) Análisis de los datos

Una vez analizadas y cuantificadas las proteínas se identifican en repositorios de acceso público. Pueden identificarse por el nombre de la proteína o por el identificador que recibe cuando se registra en una base de datos determinada (162). No todas almacenan la misma información. Como ejemplo SwissProt (www.uniprot.org) o NCBI (www.ncbi.nlm.nih.gov) almacenan información relacionada con la secuencia de aminoácidos de la proteína así como información funcional y bibliográfica.

1.8 PERFILES DE EXPRESIÓN PROTEICA EN CÁNCER DE MAMA

Los estudios de análisis de proteómica en cáncer de mama por EM descritos (Tabla 9) están realizados en muestras de tejido. No se hace referencia a estudios realizados en otras muestras biológicas o líneas celulares ya que el material de este trabajo son muestras de tejido parafinado. Los estudios de proteómica por EM en tejido parafinado son escasos, con pequeño número de muestras y realizados con diferentes técnicas. Existe discordancia entre los estudios debido a las diferencias en el flujo de trabajo experimental, pero principalmente por la heterogeneidad del tumor y la contaminación de la muestra por el tejido mamario adyacente al tumor. Algunos autores realizan microdissección con láser (de sus siglas en inglés LCM) (168–172), para solventar estos problemas (173). En la mayoría de los estudios no se ha realizado una distinción entre los subtipos definidos por IHQ entre luminal A y B por lo que la expresión de los receptores hormonales es determinante en los subtipos luminales (151). Se han realizado estudios de expresión de proteínas comparando diferentes características del tejido mamario tanto benigno como maligno y se han encontrado patrones de expresión diferencial de proteínas entre ellos. Algún estudio ha realizado una aproximación proteómica para comparar los subtipos histológicos ductal y lobulillar a nivel proteico con resultados limitados (174,175). Se han comparado tumores mamarios receptores positivos con tejido mamario benigno (176,177), tejido mamario adyacente al tumor (168), con tumores mamarios receptores hormonales negativos (169–171,178,179). También se ha realizado algún estudio para valorar el análisis diferencial de proteínas comparando diferentes estadios de cáncer de mama y su relación con la recurrencia de la enfermedad (180). Hay también algún estudio que se ha realizado para valorar la expresión diferencial de proteínas relacionadas con la resistencia al tratamiento en tumores de mama con receptores hormonales positivos (172). En 2016 se validó el primer predictor de proteínas en relación con los resultados del tratamiento con tamoxifeno en cáncer de mama recurrente con receptores hormonales positivos (181).

En los últimos años se ha realizado algún estudio para identificar perfiles proteicos en los diferentes subtipos de cáncer de mama definidos por IHQ, encontrando diferencias funcionales entre ellos (182). Así mismo, Liu y colaboradores han validado la primera firma proteómica que valora el curso clínico natural de la enfermedad triple negativa

(183). También se han identificado patrones proteicos que confieren mayor riesgo de metástasis loco-regionales en pacientes con o sin afectación ganglionar en el subgrupo triple negativo (184). Un estudio ha realizado un análisis de expresión de proteínas por EM y expresión de microRNAs comparando muestras de cáncer de mama con receptores positivos y muestras triple negativas encontrando diferencias moleculares con implicaciones clínicas (185).

Tabla 9. Estudios de proteómica en cáncer de mama receptores hormonales positivos y/o negativos (186)

Estudio	Tipo de Tejido	N	Comparación	Técnica	Nº Prots. Identificadas	Validación/ Verificación	Proteínas Candidatas
Weitzel 2010 (176)	Congelado	18	RE + vs. Tejido benigno adyacente	2D-DIGE MALDI-TOF MS	50	ND/WB	PPlaseB, Rho-GDI alfa, TPM4
Cha 2010 (168)	Congelado	18	RE+ vs. Tejido sano	1D PAGE LC-MS/MS	25	ND/IHQ	
Traub 2010 (174)	Fresco y FFPE	80	RE+ vs RE -	HPLG MALDI-TOF/MS	ND	IHQ/ND	Timosina alfa 1
Neubauer 2008 (169)	Congelado	16	RE+ vs. RE-	2DPAGE/ MALDI TOF/MS	16	IFA/ND	RGRMC1
Rezaul 2010 (170)	Congelado	6	RE α + vs. RE α -	GeLC-MS/MS	236	IHQ/ND	Liprin alfa 1, Fascin, DAP5, Beta arrestin1
Semaan 2012 (179)	Congelado	ND	RE+ vs.RE-	2DIGE LCLTQ/FT-ICR/MS	11 glicoproteínas diferenciales	ND/WB	Alfa 1 antitripsina
Sahab 2010 (171)	Fresco	ND	RE+ Vs. RE-	2D-PAGE MALDI-TOF/MS	22 diferenciales	ND/WB	SOD, Ral, ABP, Rho-GDI alfa, Galectina-1, UP2, CRABP1, S100-A11, NDPK A
Gómez-Pozo 2015 (187)	FFPE	106	RE+ vs.TN	RT-qPCR LC/MS-MS	100 diferenciales	Cohorte externa	PHGD11, LDH beta, S100A8/A9, GRP78, AGR 2-3, PIP,GPD1
Liu 2014 (183)	Fresco	63	TN	nLC-MS/MS	981	Cohorte externa	Firma de 11 proteínas

1.9 FIRMA DE ADHESIÓN: EL SUBGRUPO TN-LIKE

La firma de adhesión (FdA) es una firma de expresión proteica formada por 36 proteínas, la mayoría de ellas relacionadas con procesos de adhesión celular. Clasifica a las pacientes con diagnóstico de cáncer de mama luminal en dos grupos: *ER-true* con un pronóstico favorable a largo plazo y *TN-like* con un peor pronóstico, similar a las pacientes con cáncer de mama triple negativo (TN).

El desarrollo de la firma de adhesión se realizó en el IdiPaz (187) utilizando tecnología de alto rendimiento para el análisis de proteínas mediante la espectrometría de masas (*label free shotgun*). Se incluyeron 96 muestras de cáncer de mama, de las cuales 71 eran de tumores con receptores hormonales positivos y 25 muestras de tumores triple negativos, todas ganglios positivos. Se encontraron 224 proteínas expresadas de forma diferencial entre ambos grupos de muestras ($FDR < 5\%$). Tras la construcción de un *cluster* jerárquico (Figura 2) se observaron dos grupos: a la izquierda la mayoría de los tumores receptores hormonales positivos y a la derecha la mayoría de los tumores TN. En este grupo se encontraron, además, un grupo de tumores receptores hormonales positivos cuyo perfil de expresión era similar a los TN y se les denominó *TN-like*. El resto de tumores receptores hormonales positivos pasamos a denominarlos *ER-true*.

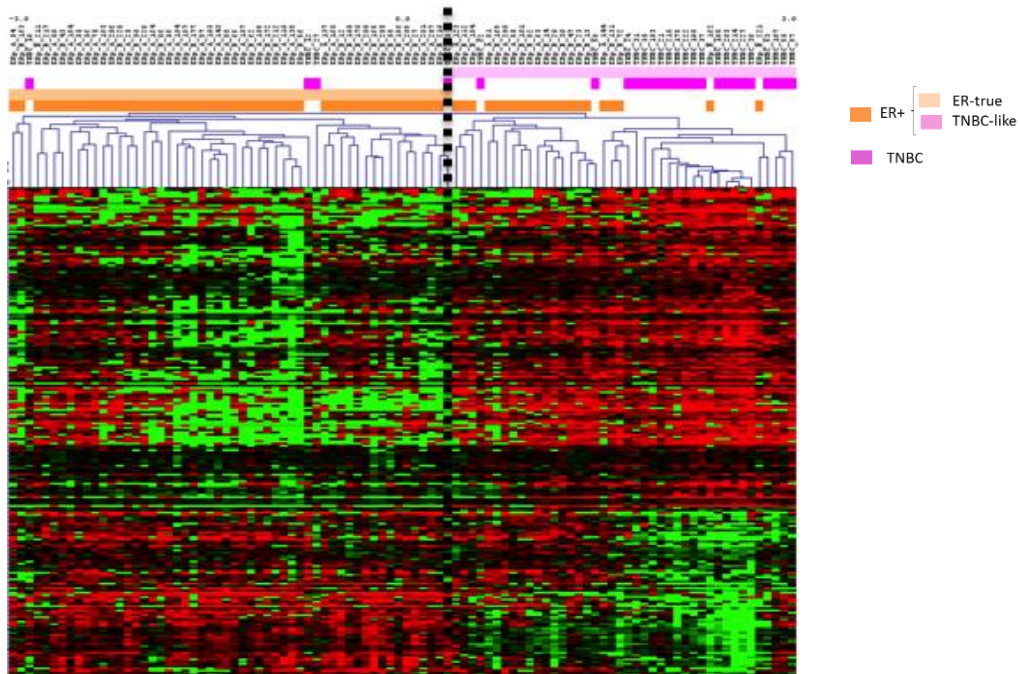


Figura 2. Cluster jerárquico. ER + (naranja oscuro) y TN (rosa oscuro). Nueva clasificación: *TN-like* (rosa claro), *ER-true* (naranja claro).

Para tratar de reducir el número de proteínas diferenciales entre los tumores ER-*true* (verde) y TN-*like* (azul), se realizó de nuevo un análisis de expresión diferencial con estos 2 grupos excluyendo los TN y centrando el estudio en los ER+. Se identificaron 44 proteínas que presentaban una expresión diferencial entre los grupos ER-*true* y TN-*like* (Figura 3).

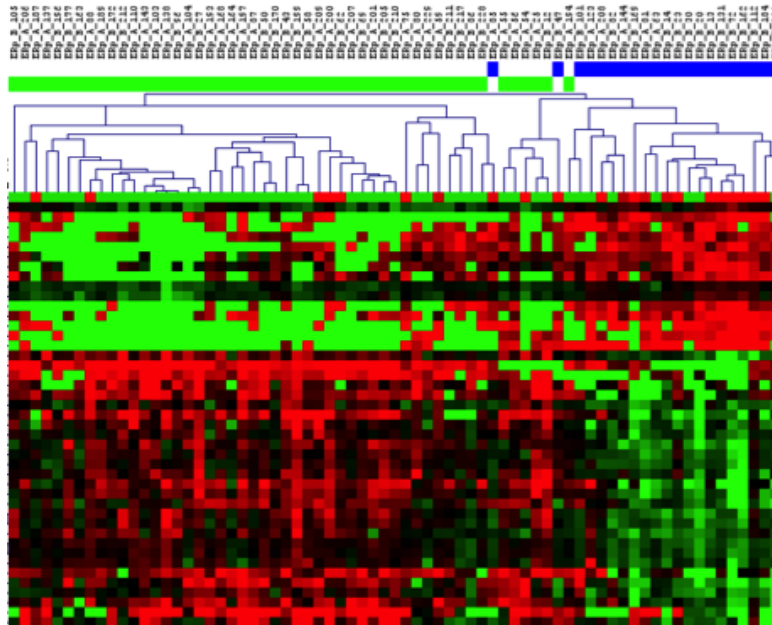


Figura 3. Cluster jerárquico de las 44 proteínas expresadas diferencialmente entre ER-*true* (verde) y TN-*like* (azul).

Para evaluar el valor pronóstico de esta nueva clasificación proteica del CM, se realizó un análisis de supervivencia de los 3 grupos (Kaplan-Meier). La supervivencia libre de metástasis a distancia (DMFS) a los 5 años fue del 65,4% para el grupo TN, del 71,1% para el TN-*like* y del 88% para el grupo ER-*true*, siendo la diferencia no significativa ($p=0.1752$). Estos datos sugirieron que el pronóstico de las pacientes con características TN-*like* era similar al del grupo triple negativo, además de presentar un perfil molecular parecido.

Con la población de tumores RH positivos, clasificados por el *cluster* en ER-*true* y TN-*like* se diseñó un “predictor de clase”. La supervivencia libre de enfermedad fue mejor para el grupo ER-*True*, pero esta diferencia no fue estadísticamente significativa ($p=0.0860$). Este “predictor de clase” se validó en una cohorte externa de expresión génica de 421 pacientes con RE + y N+. La SLE a los 5 años fue del 81,9% para el grupo ER-*true* y del 72,5% para el grupo TN-*like*, siendo esta diferencia significativa ($p=0.0047$) con una HR del 0.577, lo que significa que el riesgo de recaída para las pa-

cientes *TN-like* es del 43%. En las bases de datos disponibles solo se pudieron asignar 38 proteínas a los genes, con solo valores de expresión disponibles para 36. La mayoría de estas proteínas estaban relacionadas con el proceso de adhesión celular, (Tabla 10) por lo que este predictor de clase se denominó “Firma de Adhesión”.

Tabla 10. Funciones de las proteínas que componen la Firma de Adhesión

Categoría	Función asociada	p-valor
GOTERM_CC_FAT	GO:0031012~extracellular matrix	6.18E-12
GOTERM_CC_FAT	GO:0005581~collagen	4.17E-05
GOTERM_BP_FAT	GO:0030198~extracellular matrix organizat.	5.47E-04
KEGG_PATHWAY	Has:04512: ECM-receptor interaction	0.0017
GOTERM_BP_FAT	GO:0007155~cell adhesion	0.0167
GOTERM_BP_FAT	GO:0022610~biological adhesion	0.0167
KEGG_PATHWAY	Has:04510: Focal adhesion	0.0201
PANTHER_PATHWAY	P00034: Integrin signaling pathway	0.0285

El objetivo principal de esta tesis es la validación de la firma de adhesión en una cohorte independiente de pacientes con cáncer de mama precoz con receptores hormonales positivos.

2. HIPÓTESIS Y OBJETIVOS

2. HIPÓTESIS Y OBJETIVOS

2.1 HIPÓTESIS

Los tumores de mama TN-*like* constituyen un subgrupo dentro de los tumores con receptores hormonales positivos con entidad propia, tanto desde el punto de vista clínico como molecular. En consecuencia, deberían reproducirse los resultados previamente conseguidos en una población independiente de pacientes.

2.2 OBJETIVOS

a. Objetivo Primario:

Validación de la denominada “Firma de Adhesión” en una cohorte independiente de cáncer de mama con receptores hormonales positivos a partir de biopsias, identificando aquellas que tienen peor pronóstico. El parámetro principal de medida será la supervivencia libre de metástasis a distancia.

b. Objetivos Secundarios:

- Descripción de las características clínico patológicas de las pacientes según la Firma de Adhesión.
- Correlación entre la *técnica label-free shotgun* (cohorte de descubrimiento) y SRM (cohorte de validación).
- Analizar la supervivencia libre de enfermedad asociada a la firma de adhesión en la cohorte de validación.
- Verificar si la Firma de Adhesión ofrece una información pronóstica independiente de otros factores clínico-patológicos (tamaño tumoral, grado de diferenciación, presencia de ganglios) en la cohorte de validación.

3. MATERIAL Y MÉTODOS

3. MATERIAL Y MÉTODOS

El primer paso consistió en la selección de dos cohortes de pacientes, la primera de descubrimiento y la segunda de validación.

3.1 COHORTE DE DESCUBRIMIENTO

a) Selección de la muestra

De las 71 pacientes con cáncer de mama RH positivos incluidas en la cohorte de descubrimiento (ver apartado 1.9 Firma de Adhesión) se seleccionaron 18 pacientes: 9 del grupo *ER-true* y 9 del grupo *TN-like*, representativas y con suficiente cantidad de proteína extraída. El objetivo de reanalizar estas muestras fue confirmar los resultados previos (es decir, si realmente correspondían a muestras *ER-true* y *TN-like*) y poner a punto la SRM. Una vez puesta a punto la SRM, se procedería a trabajar con la cohorte de validación.

b) Selección de las variables

De la muestra de descubrimiento se ha seleccionado la expresión de proteínas de la denominada “Firma de Adhesión”, obtenidas por espectrometría de masas (*label-free shotgun*).

3.2 COHORTE DE VALIDACIÓN: PROYECTO ÁLAMO

a) Selección de la muestra

El proyecto Álamo es un estudio descriptivo, retrospectivo y transversal en el que se recogen datos demográficos, clínicos, terapéuticos y evolutivos de mujeres diagnosticadas de cáncer de mama invasivo en hospitales integrados en el Grupo Español de Investigación en Cáncer de Mama (GEICAM). Los datos se recogen en periodos evolutivos de 4 años: El Álamo I (AI: 1990-1993) incluyó un total de 4.532 pacientes (10% del total de las pacientes diagnosticadas en España en ese intervalo) (188); el Álamo II (AII:1994-1997) incluyó un total de 10.849 pacientes (18% del total) (189) Álamo III (AIII: 1998-2001) incluyó un total de 10.675 pacientes (17% del total). En la actualidad está en marcha el proyecto Álamo IV que incluye las pacientes diagnosticadas de cáncer

de mama invasivo en el periodo 2002-2005. Para nuestro estudio se seleccionó un subgrupo del total de las 1.070 pacientes incluidas por el Servicio de Oncología Médica (SOM) del HUB en el proyecto AII y AIII que cumplían los siguientes criterios de inclusión:

- ✓ Mujeres diagnosticadas de cáncer de mama invasivo (excluido carcinoma inflamatorio) entre el 1-1-1994 y el 31-12-2001.
- ✓ Estadio T1-3 N positivos, M0.
- ✓ Receptores de estrógenos y/o progesterona positivos por inmunohistoquímica (IHQ).
- ✓ Cirugía del tumor primario: mastectomía o cirugía conservadora con linfadenectomía axilar o biopsia selectiva del ganglio centinela.
- ✓ Tratamiento adyuvante con quimioterapia (basada en antraciclinas y/o taxanos) y hormonoterapia (tamoxifeno, inhibidores de la aromatasa o secuenciación de ambos).
- ✓ Tratamiento con radioterapia adyuvante cuando se hubiera realizado cirugía conservadora o, en caso de mastectomía, para casos seleccionados de alto riesgo según práctica clínica habitual del centro.
- ✓ Seguimiento mínimo de 5 años o hasta el fallecimiento si este ocurrió antes.
- ✓ No tener otro diagnóstico previo de cáncer excepto tumor de piel no melanoma o carcinoma *in situ* de cérvix o mama adecuadamente tratados.
- ✓ Disponibilidad de muestra de tejido tumoral parafinada.

De las 1.070 pacientes incluidas por nuestro centro, 85 cumplían los criterios de inclusión descritos previamente. El primer paso fue la desagregación de los datos clínico patológicos respecto a su código de identificación para su posterior revisión. Se realizó una revisión individualizada de los datos clínico-patológicos de todos los casos, registrados en la base de datos de ambos proyectos (AII-AIII). Se actualizaron las fechas de las recidivas, la fecha del último evento y la situación de la paciente a fecha 30 de Marzo de 2016. Así mismo se realizó una revisión de las muestras tumorales parafinadas por el patólogo referente en patología mamaria del centro, para la confirmación del diagnóstico y el contenido de células tumorales. Se incluyeron aquellas muestras cuyo contenido de células tumorales era mayor del 50%. El tamaño de la muestra definitivo

de la cohorte de validación fue de 46 muestras tumorales. Las muestras FFEP fueron aportadas por el BioBanco Vasco /O+EHUN (RD09/0076/00140) (Figura 4).

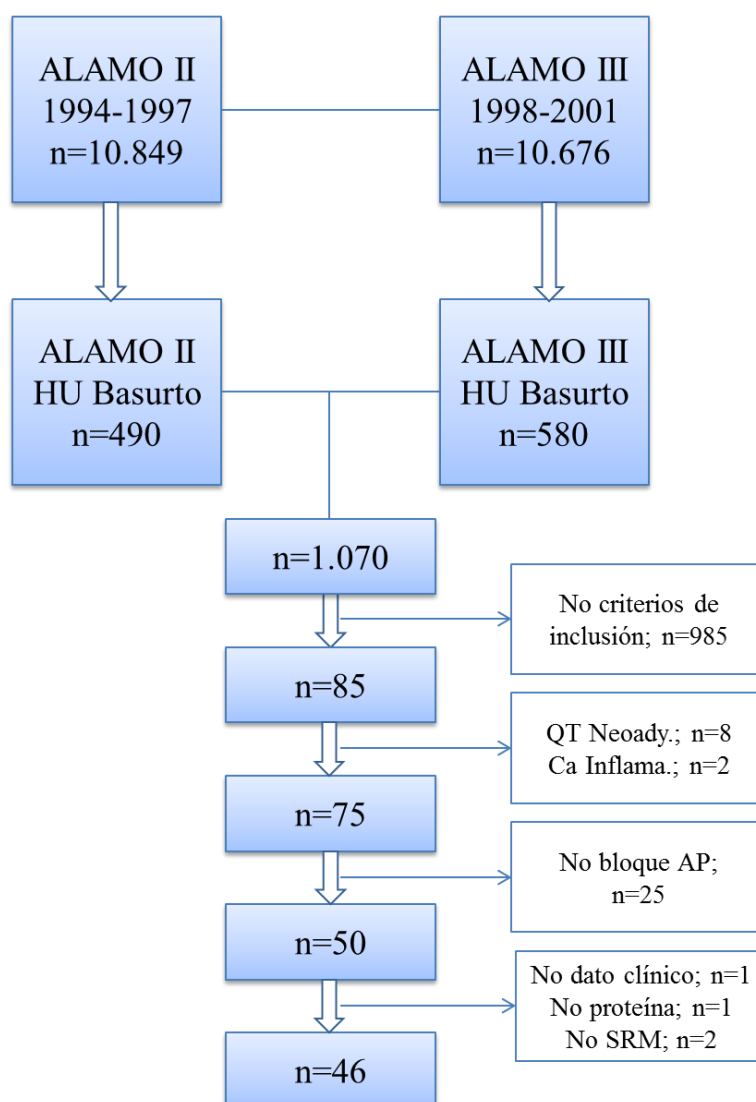


Figura 4. Diagrama de flujo de la selección de pacientes de la cohorte de validación

El estudio fue aprobado por el Comité de Ética de Euskadi con fecha 17 de Enero de 2014 (Acta 01/2014) para la cesión de muestras por parte del BIO-Banco Vasco para la Investigación al INGEMM (Instituto de Genética Médica y Molecular) del Hospital de la Paz -IdiPAZ, como parte de un proyecto de investigación.

b) Selección de las variables

De la tabla original de los datos de las pacientes del proyecto AII y AIII del SOM-HUB, se seleccionaron las siguientes variables y se trasladaron a una nueva base de datos excel.

- ✓ Edad al diagnóstico.
- ✓ Estado menopáusico: premenopáusica, postmenopáusica, desconocido.
- ✓ Fecha del diagnóstico del tumor primario.
- ✓ Fecha de la cirugía del tumor primario.
- ✓ Tipo de cirugía: Conservadora/mastectomía con linfadenectomía axilar o biopsia selectiva de ganglio centinela con/sin linfadenectomía axilar.
- ✓ Estadio clínico al diagnóstico: IIA, IIB, IIIA, IIIB, IIIC, según la clasificación TNM de la UICC (*Internacional Union Against Cancer*) (40) coincidente con la clasificación por estadios de la enfermedad de la AJCC (*American Joint Cancer Committee*) (39).
- ✓ Tamaño tumoral categorizado: <2 cm, 2,1-5 cm, >5 cm y T del TNM clínico (T₁, T₂, T₃).
- ✓ N del TNM clínico (N₁, N₂, N₃).
- ✓ Número de ganglios axilares afectados.
- ✓ Tipo histológico: ductal, lobulillar, otros o desconocido.
- ✓ Grado: I, II, III o desconocido.
- ✓ Receptores hormonales: positivos y/o negativos. Diferenciando entre receptores de estrógenos y progesterona
- ✓ HER2/neu: en el periodo de diagnóstico de estos casos (1994-2001) su determinación no se había incorporado de forma rutinaria a la clínica, por lo que no se dispone de este marcador al diagnóstico. Se realizó en todas las pacientes que presentaron una recaída a lo largo del seguimiento. Se clasificaron como positivas para la expresión del receptor HER2/ *neu*, aquellas pacientes con un resultado de IHQ +++, y aquellas con IHQ ++ con FISH positivo. Se consideró como expresión positiva del receptor, aquellos casos con resultado de FISH positivo, independientemente del resultado de IHQ. A su vez se realizó la determinación de HER2 por qPCR en todas las muestras.
- ✓ Tratamiento adyuvante: quimioterapia basada en antraciclinas y/o taxanos; hormonoterapia con tamoxifeno y/o inhibidores de la aromatasas.

- ✓ Fecha de la recaída.
- ✓ Localización de la recaída: local, regional, a distancia, segundo tumor primario infiltrante no mamario, cáncer de mama infiltrante contralateral,
- ✓ Fecha del último evento.
- ✓ Situación de la paciente en la fecha de último evento: viva sin enfermedad (VSE), viva con enfermedad (VCE), muerta con enfermedad (MCE), (incluye las muertas por toxicidad) o muerta sin enfermedad, (MSE).
- ✓ Fecha del último seguimiento.
- ✓ Fecha de la muerte.

3.3 ANÁLISIS ESTADÍSTICO

El análisis de las funciones de supervivencia de la cohorte de validación se realizó mediante el test de Kaplan-Meier. La comparación de las funciones de supervivencia de los distintos subgrupos se realizó mediante el test exacto de log-rank

Se realizó un análisis univariante y multivariante de regresión de riesgos proporcionales de Cox para evaluar los factores pronósticos. La correlación se realizó utilizando una regresión lineal y r de Pearson.

Para el análisis estadístico se utilizaron los programas BRB-ArrayTools, SPSS v16, GraphPad Prism v6 y R v3.2.5. Todas las pruebas estadísticas se consideraron bilaterales y, como valores significativos, aquéllos p -valores con $p < 0,05$. Los análisis de proteínas diferenciales se llevaron a cabo con el software MeV (Saeed 2003).

Supervivencia libre de enfermedad (SLE): Definida como el periodo de tiempo comprendido entre la fecha del diagnóstico del tumor primario (fecha del primer informe de anatomía patológica que confirma el diagnóstico de tumor mamario) y la fecha del primer evento, considerándose como primer evento los siguientes: recurrencia local o loco-regional, aparición de metástasis a distancia, presencia de un segundo tumor primario infiltrante, incluido el cáncer de mama, fecha de la muerte si no ha sucedido ninguno de los eventos previos, o fecha del último contacto de no suceder ninguno de los eventos anteriores, en cuyo caso se censura en ese momento

Supervivencia libre de enfermedad a distancia (SLED) definida como el periodo de tiempo comprendido entre la fecha del diagnóstico del tumor primario y la fecha del primer evento a distancia. Se excluye la recidiva local ipsilateral, la recurrencia infil-

trante loco-regional, cáncer de mama contralateral y todos los cánceres *in situ* (190).

3.4 METODOLOGÍA PARA EL ANÁLISIS PROTEÓMICO

Para el análisis proteómico de las muestras tumorales de la cohorte de validación se ha seguido la metodología empleada previamente en otros estudios del equipo de investigación (191,192).

a) Preparación y extracción de proteínas a partir de las muestras clínicas

Se extrajeron las proteínas a partir de muestras fijadas y embebidas en parafina (FFEP). Las muestras se desparafinaron con xilol y se lavaron dos veces con etanol absoluto. Los extractos de proteínas provenientes de las muestras FFEP se prepararon en una disolución de 2% de SDS mediante un protocolo basado en la recuperación de antígeno inducida por calor. La concentración de proteína se determinó con el kit MicroBCA Protein Assay Kit (Pierce-Thermo Scientific). Seguidamente, se digirieron 10 µg de extracto con tripsina (1:50) y se eliminó el SDS mediante las columnas Detergent Removal Spin Columns (Pierce). Las muestras se desalaron con ZipTips (Millipore), secan y resuspendieron en 15 µL de una disolución de 0.1% de ácido fórmico y 3% de acetonitrilo antes de su análisis por proteómica

b) Detección de la expresión de ERBB2 por qPCR

Las muestras fijadas en formol y embebidas en parafina (FFEP) se cortaron en secciones seriadas con un grosor de 10 µm. El RNA se aisló mediante el RecoverAll Total Nucleic Acid Isolation Kit (Life Technologies). Se cuantificó y midió la pureza del RNA extraído en un espectrofotómetro ND-1000 NanoDrop (Thermo Fisher Scientific). La expresión de ERBB2 se midió con una sonda TaqMan (ERBB2-Hs00170433_m1). Además, se midió la expresión de cuatro genes de referencia (IPO8-Hs00183533_m1, HMBS-Hs00609297_m1, POLR2A-Hs00172187_m1, SDHA-Hs00188166_m1), identificados con el método Normean (193). La concentración de RNA se ajustó a 166.7 ng/µL. Seguidamente, para la síntesis de ADNc, 500g del RNA total se transcribieron con TaqMan MicroRNA Reverse Transcription Kit combinado con stem-loop Megaplex Primer Pools, de acuerdo al protocolo propuesto por el fabricante (Life Technologies).

Las reacciones de PCR se llevaron a cabo en un sistema de detección ABI Prism 7900HT Sequence (Life Technologies). Los valores de cada ciclo de amplificación (Cq) se obtuvieron con el software SDS 2.2 (Applied Biosystems) empleando los ajustes por defecto y un límite de 0.2. El máximo valor Cq se ajustó a 40 y los valores Cq se normalizaron usando la expresión de los genes de referencia. La expresión relativa de cada gen se expresó como $\Delta Cq = Cq_{\text{ref}} - Cq_{\text{gene}}$. Cq ref se obtuvo calculando la media geométrica de los genes de referencia. Se empleó como punto de corte el valor $Cq=0.95$, definido previamente en una serie de 218 tumores de mama para los que fue evaluada la expresión de ERBB2 por IHC/FISH y por qPCR (193).

c) *Selected Reaction Monitoring (SRM)*

El diseño de los SRM se basó en análisis previos de “*shotgun proteomics* (185) y en las bases de datos PeptideAtlas (<http://www.peptideatlas.org/>) (194) y SRMATlas (www.srmatlas.org) (195). La adquisición de los espectros MS2 de los experimentos de SRM se realizó en un QTRAP 5500 (ABSciex, Concord, Ontario) y los resultados obtenidos se analizaron en un espectrómetro de masas TSQ Vantage Triple Quadrupole (ThermoFisher, San Jose, CA, USA), ambos equipados con una fuente iónica de nanoelectrospray. La separación cromatográfica de los péptidos se llevó a cabo empleando un NanoLC-2D HPLC (Eksigent, Dublin, CA) acoplado a un emisor de sílice fundido de 15cm y 75 μm de diámetro, empaquetado con un ReproSil-Pur C18-AQ 120 y 1.9 μm de resina (Dr. Maisch HPLC GmbH). Los péptidos se cargaron en la columna desde un Eksigent autosampler a 4°C y se separaron con un gradiente lineal de acetonitrilo/agua, compuesto por 0.1% de ácido fórmico, con una tasa de flujo de 300 nl/min. Se empleó un gradiente de 5 a 35% de acetonitrilo en un tiempo de 40 minutos. Los espectros MS2 se registraron a partir de la detección por SRM de 1000 iones. Se usaron una media de 100 transiciones (tiempo de exploración 10 ms/transición) por cada carrera y se obtuvieron Q1 y Q3 a 0.7 unidades de resolución de masas. Los espectros MS2 se registraron mediante el modo de activación del ión producto (EPI) para las transiciones MRM más altas, empleando un tiempo de rellenado dinámico, una unidad de resolución Q1 y una velocidad de exploración de 10000 amu/s, con un rango m/z de 300 a 1000. Las energías de colisión usadas por los dos modelos de adquisición se calcularon de acuerdo a la siguientes fórmulas: $CE = 0.044 * m/z + 5.5$ y $CE = 0.051 * m/z + 4$, siendo CE la energía de colisión, y m/z el ratio de masa-carga del ión precursor, para iones precursor-

res con doble y triple carga respectivamente. En los SRM, el espectrómetro de masas se utilizó en el modo “SRM scan”, donde Q1 y Q3 se obtuvieron a 0.7 unidades de resolución de masas. Las energías de colisión de cada transición se calcularon de acuerdo a las siguientes ecuaciones: $CE = 0.034 * (m/z) + 3.314$ y $CE = 0.044 * (m/z) + 3.314$ (CE= energía de colisión, m/z= ratio masa-carga) para iones precursores con doble y triple carga respectivamente. Se monitorizaron tres transiciones distintas para cada péptido endógeno (los péptidos ligeros) y para cada péptido usado como standard interno (los péptidos pesados). Los datos obtenidos se procesaron mediante el software SRM Skyline. Finalmente, para la cuantificación de proteínas se emplearon los siguientes criterios:

1. Correlación entre los ratios entre los péptidos ligeros y pesados.
 2. Correlación entre los ratios de los iones de ambas formas y los ratios de los iones obtenidos a partir de los espectros de MS/MS presentes en la librería de espectros de SRM.
 3. Intensidad de las transiciones de la forma ligera y la forma pesada mayores que 10.
- Las tres transiciones de cada péptido ligero y de cada péptido pesado se emplearon para cuantificar los péptidos excepto en el caso en el que se detectaron interferencias al coeluir. Se ignoraron las mediciones puntuales de péptidos ligeros menores que la medición de fondo. Por último, se calculó un ratio entre el péptido ligero y el péptido pesado (ratio L/H) para todas las transiciones. Los valores de expresión de las proteínas se calcularon a partir de la mediana de la expresión de las tres transiciones para cada par de péptidos ligero y pesado.

d) Desarrollo del predictor de clase

Un predictor de clase es un algoritmo que permite clasificar, en este caso, las muestras clínicas en una o más categorías. Los predictores de clase se desarrollaron mediante el método “*Support Vector Machine*” (196). Al modelo predictivo se incorporaron aquellas proteínas que presentaban diferencias significativas en su expresión mediante un t-test de varianza aleatoria y en las que hubiese podido ser previamente identificado su gen correspondiente en la base de datos de meta-validación. El error predictivo de cada modelo se estimó usando una validación cruzada “*leave-one-out*” (LOOCV). Para cada grupo de validación LOOCV, se repitió la construcción del modelo, incluido el proceso de selección de genes. Además, se evaluó si el error estimado para cada modelo era significativamente menor que la predicción al azar. El nivel de significación corresponde a

la proporción de permutaciones al azar en las que se obtiene un error estimado menor que el error estimado obtenido con los datos originales. Este análisis se llevó a cabo en BRB-ArrayTools v4.2.1, desarrollado por el Dr. R. Simon.

4. RESULTADOS

4. RESULTADOS

4.1 DESCRIPCIÓN DE LA COHORTE DE VALIDACIÓN

Un total de 46 pacientes con cáncer de mama del HUB se incluyeron en la cohorte de validación independiente. Todas las pacientes presentaban un carcinoma ductal infiltrante con afectación ganglionar, receptores hormonales positivos (receptores de estrógenos y/o de progesterona). Así mismo, todas las pacientes recibieron tratamiento adyuvante con quimioterapia basada en antraciclinas seguido de tratamiento hormonal. Las características clínicas de la población se presentan en la Tabla 11.

Tabla 11. Características clínico-patológicas basales de las pacientes

Característica clínico-patológica		Cohorte Validación N=46
Edad media (rango)		55 (39-70)
Estado Menstrual (%)	Pre-menopáusicas	19 (41)
	Post-menopáusicas	26 (57)
	Desconocido	1 (2)
Estadio clínico (%)	I	13 (28)
	IIA	14 (30)
	IIB	6 (13)
	IIIA	7 (15)
	Desconocido	6 (14)
Cirugía (%)	Conservadora + LA	20 (44)
	Mastectomía radical + LA	26 (56)
Estadio patológico (%)	IIA	17
	IIB	16
	IIIA	13
pT(%)	T1	19 (41)
	T2	21 (46)
	T3	6 (13)
	T4	0 (0)
pN (%)	N1	39 (85)
	N2	7 (15)
Grado Tumoral (%)	G1	6 (13)
	G2	22 (48)
	G3	12 (26)
	Desconocido	6 (13)
Receptores Estrógenos (%)	Positivos	44 (96)
	Negativos	2 (4)
Receptores Progesterona (%)	Positivos	39 (85)
	Negativos	5 (11)
	Desconocido	2 (4)

LA: Linfadenectomía Axilar.

La edad media al diagnóstico era de 55 años (39-70), el 57% post-menopáusicas. El 43% presentaban un estadio clínico II. Se realizó cirugía conservadora en el 44%. El 46% T2, 85% N1, 48% G2. Los datos clínicos utilizados para el análisis fueron actualizados hasta Abril de 2016. La mediana de seguimiento para la serie fue de 160,50 meses (rango 12-261). Durante el período de seguimiento hubo 32 eventos (70%): 6 recidivas loco-regionales, realizándose una exéresis quirúrgica a 5, 19 recurrencias a distancia y 7 segundos tumores (1 carcinoma in situ de mama contralateral, 1 carcinoma lobulillar de mama contralateral, 1 carcinoma de esófago, 1 melanoma, 1 linfoma no Hodgking, 1 carcinoma de tiroides, 1 aplasia medular, 1 carcinoma de colón).

4.2 ANÁLISIS DE SUPERVIVENCIA

La mediana de seguimiento para la serie fue de 160,50 meses (rango 12-261). La supervivencia libre de enfermedad (SLE) a 5, 10 y 15 años fue del 74%, 43% y el 29%, siendo la mediana de 7,85 años (Figura 5).

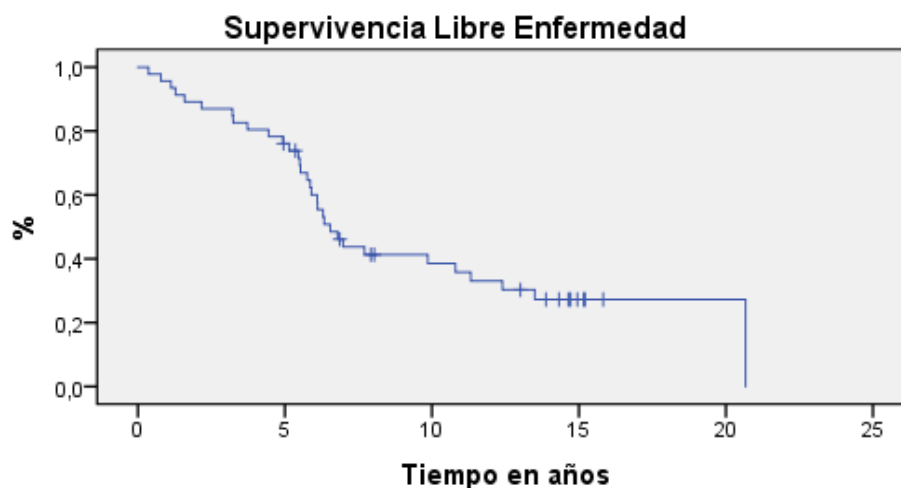


Figura 5. Supervivencia libre de enfermedad a los 5, 10 y 15 años.

La media de supervivencia libre de enfermedad a distancia (SLED) fue de 14,86 años (IC 95% 12,374-17,347). La mediana no se ha alcanzado. La SLED a 5, 10 y 15 años fue del 78%, 67% y 50% respectivamente (Figura 6).

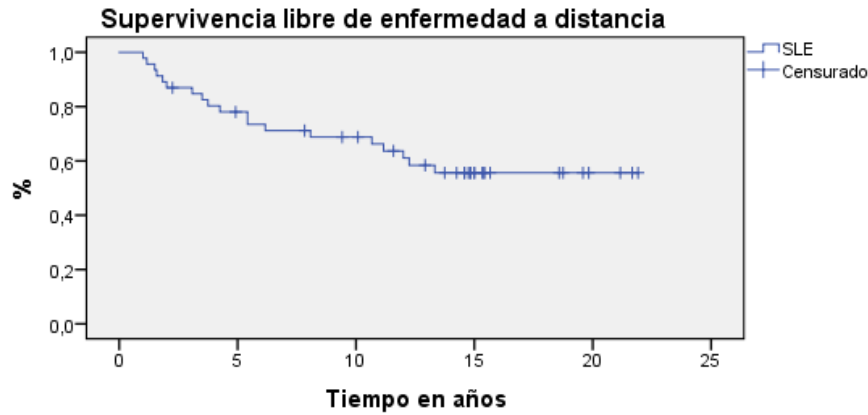


Figura 6. Supervivencia libre de enfermedad a distancia a los 5, 10 y 15 años.

4.3 EXPRESIÓN DE ERBB2 -HER2- POR qPCR

Se realizó la determinación de Her2 por qPCR a 284pacientes para establecer el punto de corte para determinar estatus positivo de ERBB2, incluyendo las 46 de la muestra de validación. De estas 46 muestras, 41 (89 %) fueron negativas y 5 (11%) fueron positivas. En esta cohorte de validación, a 13 muestras tumorales se le había realizado a la recaída la determinación de HER2 por IHQ/FISH siendo negativa y en todas se confirmó la expresión negativa de ERBB2 por qPCR. La coincidencia entre la técnica de IHQ/FISH y qPCR fue, por tanto, del 100% (Figura 7).

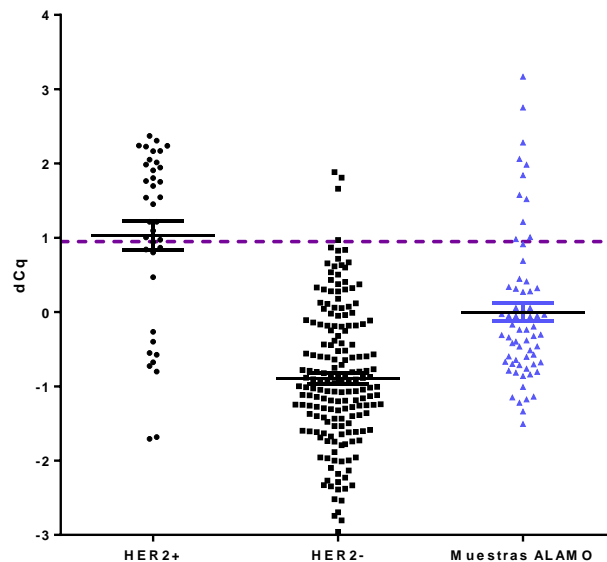


Figura 7. Medición de ERBB2 en la serie de referencia y la cohorte de validación.

El área bajo la curva ROC para la predicción por qPCR del estatus de ERBB2 fue 0.88 ($p < 0.0001$) con una sensibilidad y especificidad para el punto de corte elegido (0,95) de 97,78 y 63,16 respectivamente (Figura 8).

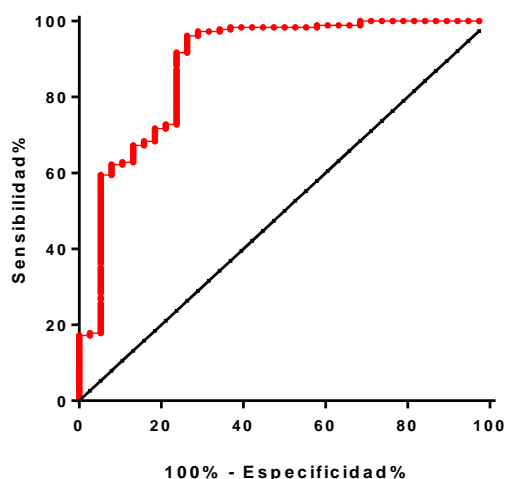


Figura 8. Curva ROC para la predicción por qPCR del estatus de ERBB2.

4.4 RESULTADOS SRM

a) Diseño del experimento de SRM

A partir de los datos de proteómica obtenidos mediante espectrometría de masas (EM) *shotgun label free*, se seleccionaron 33 proteínas diferenciales entre las dos clases de tumores previamente establecidas, *ER-true* y *TN-like*. Además, se analizó otro grupo de 17 proteínas con valor pronóstico definidas previamente (185). Para la selección de las transiciones se emplearon por un lado los datos de 152 péptidos de los experimentos previos (185) y se completaron empleando la base de datos SRMatlas para las proteínas en las que no se disponía de 3 péptidos únicos identificados en la base de datos experimental previa (80 péptidos). Finalmente, el experimento diseñado permitió la medición de 50 proteínas y 159 péptidos (Figura 9).

- 159 péptidos * 2 precursores (*Heavy & Light*) * 3 transiciones=954 mediciones.
- 64 muestras * 2 carreras por muestra=132 carreras.

TOTAL: 477 mediciones*64 muestras=30528 variables a analizar.

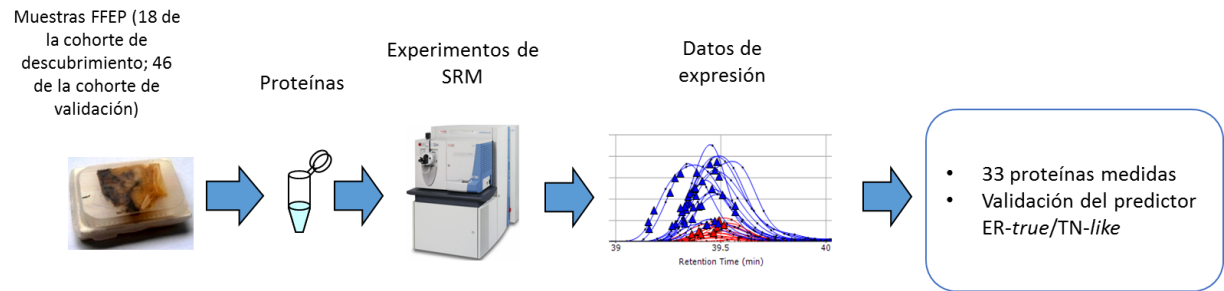


Figura 9. Flujo de trabajo experimental empleado en nuestra serie.

b) Medición de la expresión de proteínas por SRM

Se midió la expresión de estas 33 proteínas por “*Selecting Reaction Monitoring*” (SRM) en la cohorte de validación: 46 pacientes ER+. Una muestra fue excluida del análisis debido a la poca cantidad de proteína extraída y dos debido a la baja calidad de la medición. También se midió esta expresión en las 18 muestras adicionales de la cohorte de descubrimiento (9 ER-true y 9 TN-like) (Anexo 1).

c) Validación del predictor de clase en la cohorte de descubrimiento (18 muestras)

Con el fin de comprobar si los datos obtenidos por SRM eran capaces de discriminar entre los subtipos ER-true y TN-like, se analizaron las 18 muestras de la cohorte de descubrimiento. Efectivamente, la técnica SRM fue capaz de detectar diferencias entre los subtipos TN-like y ER-true en las 18 muestras provenientes de esta cohorte de descubrimiento, evaluadas mediante agrupamiento jerárquico (Figura 10).

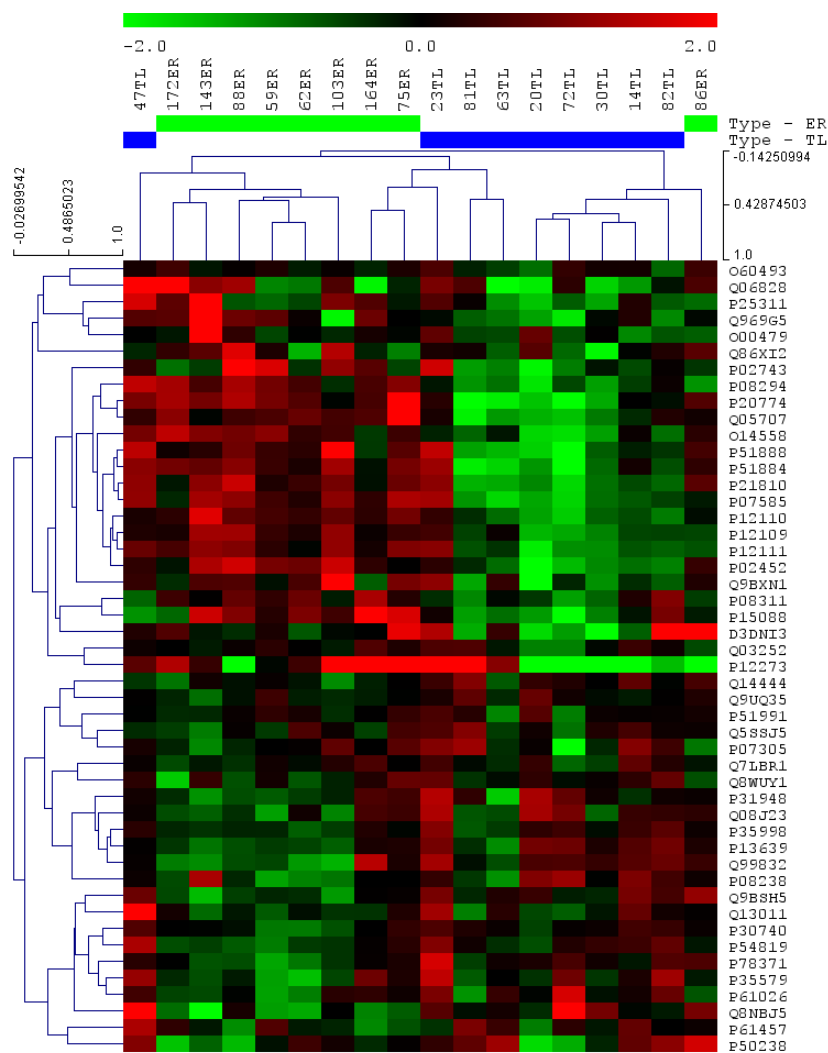


Figura 10. Clúster jerárquico de las 18 muestras de la serie de descubrimiento. ER-true (verde) y TN-like (azul).

d) Correlación entre la espectrometría de masas y los SRM

Se calculó el coeficiente de correlación de Pearson entre cada una de las mediciones obtenidas por EM y las obtenidas por SRM para cada una de las 33 proteínas seleccionadas. En algunos casos, la correlación fue muy alta, siendo las mediciones comparables. Sin embargo, hay proteínas en las que la correlación entre las dos mediciones fue muy baja. El 36% presentaron una correlación mayor de 0,8 (P21810, P12109, P12273, P02452, P02743, P12111, P51884, O14558, P07585, P08311, P08294 y P12110) y el 64% menos de 0,8 (P25311, Q05707, P15088, P13639, P20774, P51888, P61026, Q06828, Q8WUY1, Q9BSH5, Q8NBJ5, P08238, Q08J23, P31948, Q99832, Q99832, Q9UQ35, P35998, Q9BXN1, Q86XI2, Q969G5 y Q14444) (Figura 11).

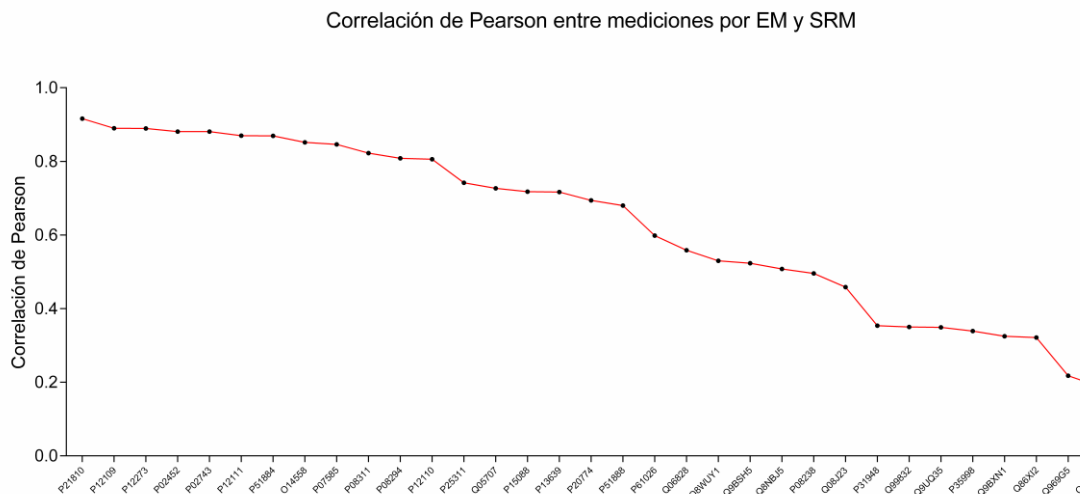


Figura 11. Correlación entre las mediciones obtenidas por EM y los SRM para las 33 proteínas seleccionadas.

4.5 DESARROLLO DEL PREDICTOR DE CLASE

Para asignar las muestras de la cohorte de validación, se definió un clasificador de *ER-true/TN-like* que incluía 14 proteínas de las 33 proteínas medidas (Tabla 12). El algoritmo de predicción se define como la suma de los pesos (w_i) y la expresión (x_i) de los genes o, en este caso, de las proteínas significativas. La expresión es el logaritmo de la intensidad normalizada. Una muestra se clasifica como *ER-true* si la suma es mayor que el punto de corte, es decir, $\sum_i w_i x_i > \text{punto de corte}$. El punto de corte calculado por BRB para el predictor por “Support Vector Machine” es de -3,2.

Tabla 12. Poder discriminante de las 14 proteínas que componen el predictor de clase

Proteína	p-valor	Media geométrica clase1	Media geométrica clase2	Vez de cambio (cl. 1 / cl. 2)	SVM
Q99832	0.0047	0.2	0.41	0.48	-0.3411
P08238	0.0073	2.59	4.61	0.56	-0.3224
Q8NBJ5	0.0092	0.26	0.48	0.54	-0.5478
P12111	0.0090	17.09	8.52	2.01	-0.3864
Q06828	0.0081	0.52	0.19	2.72	0.3348
P21810	0.0080	3.71	1.6	2.31	-0.1182
P51884	0.0034	7.47	2.59	2.89	-0.118
P07585	0.0027	2.2	0.84	2.63	-0.2396
O14558	0.0019	0.88	0.34	2.55	-0.0216
P12109	0.0013	8.49	3.65	2.33	-0.1776
Q05707	0.0010	0.98	0.35	2.77	0.4686
P20774	0.0007	2.35	0.65	3.6	0.3767
P02452	0.0002	0.73	0.22	3.27	0.8191
P12110	0.0001	0.81	0.37	2.18	0.2296

Clase1: *ER-true*; Clase2: *TN-like*. SVM: Pesos asignados por el análisis *Support Vector Machine*.

Entre las proteínas con mayor expresión en tumores *ER-true*, encontramos 5 proteoglicanos canónicos extracelulares pequeños ricos en leucina (de sus siglas en inglés SLRP): decorina, biglican, mimecan, lumican y la fibromodulina; las tres cadenas polipeptídicas principales de Colágeno VI - $\alpha 1(VI)$, $\alpha 2(VI)$ y $\alpha 3(VI)$, la cadena alfa 1 del colágeno I, y la proteína de choque térmico (HSP) o proteínas de stress HSPB6, también denominada HSP20. Por otro lado, los tumores de tipo *TN-like* tenían una expresión más alta de proteínas HSP90AB1, cadena alfa del colágeno 14 y la procolágeno galactosiltransferasa 1 y la subunidad eta del complejo proteico T (Tabla 13).

Tabla 13. Características de las proteínas que definen el predictor de clase

Código	Uniprot	Nombre de la proteína	Gen	Sobre-expresion
O14558	HSPB6_HUMAN	Heat shock protein beta-6	HSPB6	ER-true
P02452	CO1A1_HUMAN	Collagen alpha-1(I) chain	COL1A1	ER-true
P07585	PGS2_HUMAN	Decorin (Bone proteoglycan II)	DCN	ER-true
P08238	HS90B_HUMAN	Heat shock protein HSP90-beta	HSP90AB1	TN-like
P12109	CO6A1_HUMAN	Collagen alpha-1(VI) chain	COL6A1	ER-true
P12110	CO6A2_HUMAN	Collagen alpha-2(VI) chain	COL6A2	ER-true
P12111	CO6A3_HUMAN	Collagen alpha-3(VI) chain	COL6A3	ER-true
P20774	MIME_HUMAN	Mimecan	OGN	ER-true
P21810	PGS1_HUMAN	Biglycan	BGN	ER-true
P51884	LUM_HUMAN	Lumican	LUM	ER-true
Q05707	COEA1_HUMAN	Collagen alpha-1(XIV) chain	COL14A1	TN-like
Q06828	FMOD_HUMAN	Fibromodulin	FMOD	ER-true
Q8NBJ5	GT251_HUMAN	Procollagen galactosyltransferase 1	COLGALT1	TN-like
Q99832	TCPH_HUMAN	T-complex protein 1 subunit eta	CCT7	TN-like

4.6 ASIGNACIÓN DE LAS MUESTRAS DE VALIDACIÓN A LOS GRUPOS ER-TRUE O TN-LIKE A PARTIR DEL PREDICTOR DE CLASE

De las 46 muestras tumorales de la cohorte de validación el predictor de clase asignó el 34,7% (16 muestras) al grupo *ER-true* y el 65,3% (30 muestras) al grupo *TN-like*.

Las características clínico-patológicas de ambos subgrupos se describen en la tabla 14.

Tabla 14. Características clínico-patológicas de los subgrupos TN-like y ER-true

		ER-true, n=16 (%)	TN-like, n=30 (%)
Edad media (rango)		54 (39-67)	54 (39-70)
Estadio patológico (%)	IIA	5 (31)	12 (40)
	IIB	6 (38%)	10 (33)
	IIIA	5 (31)	8 (27)
pT (%)	T1	6 (38)	13 (43)
	T2	9 (56)	12 (40)
	T3	1 (6)	5 (18)
pN (%)	N1	12 (75)	27 (90)
	N2	4 (25)	3 (10)
Grado Tumoral (%)	G1	2 (13)	4 (13)
	G2	9 (56)	13 (43)
	G3	4 (25)	8 (27)
	Desconocido	1 (6)	5 (17)
Receptores Estrógenos (%)	Positivos	16 (100)	28 (93)
	Negativos	0 (0)	2 (7)
Receptores Progesterona (%)	Positivos	14 (88)	25 (83)
	Negativos	2 (13)	3 (10)
	Desconocido	0 (0)	2 (7)

Respecto a los eventos observados en la cohorte de validación una vez asignada la categoría ER-true/TN-like por el predictor de clase, 23 (77%) eventos sucedieron en el subgrupo TN-like y 9 (23%) en el subgrupo ER-True. El subgrupo TN-like presentó un mayor porcentaje de metástasis a distancia y de segundos tumores, mientras que el subgrupo ER-true presentó un porcentaje mayor de recidivas locales. (Tabla 15).

Tabla 15. Recurrencia de las pacientes en la cohorte de validación por subgrupos

Eventos	ER-true (%), n=9 (23)	TN-like (%), n=23 (77)
Loco-regional	4	2
Metástasis a distancia	4	15
2° Tumores	1	6

La supervivencia libre de enfermedad a distancia a los 5, 10 y 15 años fue del 93,73%, 87,05% y 67,70% respectivamente para el subgrupo ER-true y 68,79%, 50,45% y 40,36% para el subgrupo TN-like, con una $p=0,0777$ (Figura xx). La mediana de supervivencia se alcanzó a los 12,18 años en el caso del subgrupo TN-like. No se ha alcanzado en el subgrupo de ER-true cuando se realizó el análisis (Figura 12).

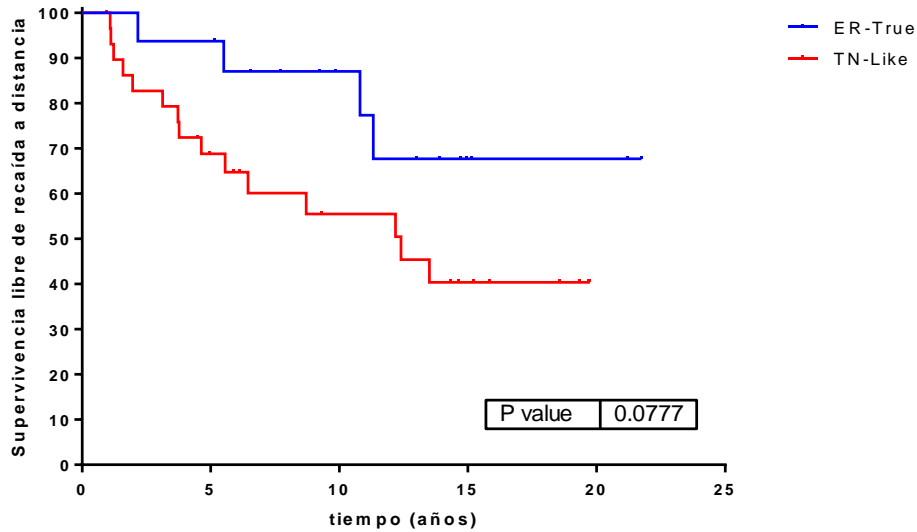


Figura 12: Supervivencia libre de enfermedad a distancia en los dos grupos ER-true y TN-like.

4.7 SUPERVIVENCIA LIBRE DE ENFERMEDAD

La SLE a los 5, 10 y 15 años fue del 87,5%; 56,25% y 43,75% para el subgrupo ER-true y del 56,47%; 35,29% y 35,29% para el TN-like respectivamente, con una $p=0,0954$. La mediana de la SLE se alcanzó a los 10,8 años en el subgrupo ER-true y a los 6,1 años en el subgrupo TN-like respectivamente (Figura 13).

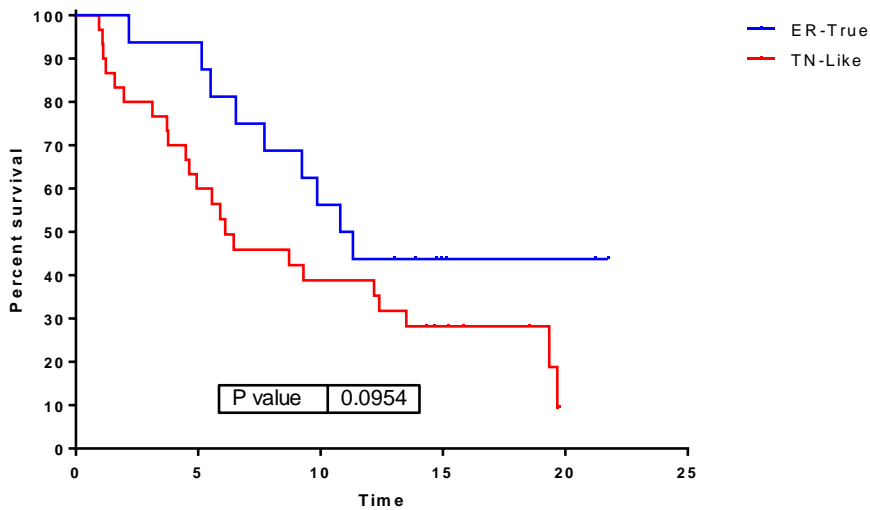


Figura 13: Supervivencia libre de enfermedad en los ER-true y los TN-like.

4.8 ANÁLISIS UNIVARIANTE Y MULTIVARIANTE

Con el fin de establecer si la clasificación en *ER-true* y *TN-like* aporta información adicional a los parámetros clínicos en esta cohorte, se realizaron análisis univariantes y multivariante con el tamaño tumoral, el estatus ganglionar y el grado de diferenciación. Ninguno de los parámetros clínicos ni el propio predictor aportaron información significativa (Tabla 16).

Tabla 16. Análisis Univariante y Multivariante entre el predictor y el tamaño tumoral, estatus ganglionar y grado de diferenciación con sus valores pronósticos (n=46)

Cohorte Validación	Supervivencia Libre de Metástasis a distancia			
	Análisis Univariante		Análisis Multivariante	
	Hazard Ratio (IC 95%)	p	Hazard Ratio (IC 95%)	p
Predictor <i>ER-true</i> vs. <i>TN-like</i>	1.955 (0.607-6.301)	0.261	1.955 (0.607-6.301)	0.261
Tamaño Tumoral <2 cm vs > 2 cm	0.740 (0.300-1.823)	0.512	0.689 (0.195-2.433)	0.563
Ganglios 1-3 vs >=4	1.777 (0.588-5.373)	0.309	1.388 (0.414-5.324)	0.595
Grado G3 vs G2-G1	0.946 (0.266-3.357)	0.931	1.163 (.0254-5.324)	0.846

5. DISCUSIÓN

5. DISCUSIÓN

Durante los últimos años, se ha hecho cada vez más patente la importancia de dividir el cáncer de mama en distintas entidades. Si bien la heterogeneidad de la enfermedad se daba por sentada desde hacía mucho tiempo, habida cuenta de la variedad de evolución clínica, no había sido posible una adecuada caracterización hasta la llegada de técnicas de análisis molecular de alto rendimiento. La oncología está pasando ahora de una clasificación de la enfermedad basada en la anatomía y la IHQ a una subdivisión basada en la expresión de genes y la presencia de mutaciones.

La subdivisión del cáncer de mama no solo posee interés académico. Esta nueva visión de la enfermedad pretende, en el caso de los tumores localizados, ofrecer una información más exacta sobre el pronóstico de cada paciente y sobre las posibilidades de tratamiento complementario a la cirugía.

La estadificación TNM y la expresión de receptores hormonales y HER2 permiten tomar decisiones terapéuticas en el ámbito adyuvante (Dings 2013). El uso combinado de quimioterapia y hormonoterapia ha permitido mejorar progresivamente las tasas de supervivencia de los tumores localizados (141,197). Sin embargo, los médicos que atienden a las pacientes son conscientes de que cada grupo definido según estos parámetros clásicos sigue siendo heterogéneo. Ello implica que muchas pacientes no reciben la mejor opción de tratamiento. En el caso de los tratamientos adyuvantes, algunas pacientes sufren la toxicidad de los mismos sin necesitarlos, mientras que otras recaen a pesar de recibirlos.

Una clasificación precisa con valor pronóstico afecta a la elección del tratamiento adyuvante, pero también al seguimiento de la paciente. Los esquemas de seguimiento han variado durante las últimas décadas, habiéndose pasado de un modelo maximalista muy poco rentable a otro basado en la realización de las pruebas imprescindibles. Resulta fácil intuir que cada paciente debería someterse a un plan de seguimiento ajustado a su riesgo de recaída, pero es muy difícil llegar a un consenso al respecto mientras no dispongamos de marcadores pronósticos más fidedignos.

Las anteriores consideraciones remarcan la necesidad de desarrollar clasificaciones que superen las limitaciones y permitan un abordaje más personalizado. La clasificación por subtipos moleculares con el perfil PAM50 ha supuesto un primer paso en este sentido

(63,72), porque ofrece información pronóstica independiente de la proporcionada por los parámetros clásicos. Sin embargo, hasta el momento la clasificación por subtipos moleculares ha carecido de utilidad terapéutica. Los tumores luminales siguen tratándose con hormonoterapia y los HER2 con trastuzumab, pero la decisión se toma simplemente con la información ofrecida por la IHQ.

Las plataformas génicas que ya se emplean en la rutina asistencial (OncoType, MammaPrint, Prosigna o EndoPredict) suponen un avance significativo, porque permiten evitar la quimioterapia en pacientes con tumores hormono-sensibles de bajo riesgo de recaída. Esto ya se ha constatado en casos sin afectación axilar y también en casos con hasta 3 ganglios axilares positivos (42).

El trabajo presentado en esta tesis doctoral da un paso más allá al subdividir el grupo de tumores con receptores hormonales positivos y HER2 negativo en dos entidades con un pronóstico diferente. Utilizando proteómica de alto rendimiento, los tumores con estas características pertenecían bien a un grupo de pronóstico excelente *-ER-true-* o a un segundo grupo con un pronóstico similar al cáncer de mama triple negativo *-TN-like-*. Se confirman así los resultados del estudio previo que describió por primera vez el grupo *TN-like* (187).

La proteómica ofrece una visión complementaria a la de la genómica, informando sobre procesos que pueden pasar desapercibidos al análisis masivo de expresión génica. Por este motivo, se está prestando cada vez más atención a este tipo de tecnología. La proteómica del cáncer de mama ha seguido un desarrollo más lento que el de la genómica, en gran parte debido a limitaciones técnicas. Los estudios previos con muestras de tejido parafinado han sido escasos, heterogéneos y con un pequeño número de pacientes. El estudio actual ha pretendido superar estos problemas.

La población seleccionada en este trabajo para realizar la validación externa independiente de la Firma de Adhesión formaba parte del proyecto de investigación del grupo GEICAM Álamo, cuyos datos cumplen un estándar de calidad importante. Esto supera una de las limitaciones habituales en los estudios retrospectivos: la fiabilidad de los datos clínicos recogidos. Además, la serie era homogénea en cuanto a la extensión de la enfermedad (todas las pacientes presentaban ganglios positivos) y el tratamiento. Éste incluyó cirugía seguida de quimioterapia basada en antraciclinas (esquema FAC) y hormonoterapia con tamoxifeno. Además, el periodo de seguimiento era suficientemen-

te largo. La supervivencia global a 5 y 10 años de la serie fue del 74% y el 47%, respectivamente.

El grupo *TN-like* fue inicialmente descrito mediante espectrometría de masas en un estudio previo. Los resultados debían confirmarse con un nuevo trabajo. El presente trabajo respalda la existencia de este grupo, no solo porque demuestra que el tipo *TN-like* puede ser identificado y tiene mal pronóstico, sino porque se ha empleado para ello una técnica alternativa, la SRM. El uso de técnicas moleculares de alto rendimiento puede originar subdivisiones espurias, fruto de la agrupación estadística de datos pero sin traslación biológica. Cuando el hallazgo se confirma con una segunda técnica, tenemos datos sólidos para considerar al grupo como una entidad biológicamente definida.

Tanto la espectrometría de masas como la SRM son complejas y muy difíciles de implantar en la práctica clínica. Esto indica que la proteómica de alto rendimiento, si bien tiene un gran potencial, corresponde más al ámbito de la investigación y el descubrimiento. En el caso de que se quisiera detectar el grupo *TN-like* en clínica de forma rutinaria, debería realizarse un traslado de tecnología y plataforma, por ejemplo a expresión génica con qPCR.

El hallazgo de un tipo molecular con valor pronóstico es solo un primer paso dentro del estudio de las proteínas del tumor. En el futuro, deberá analizarse el papel que tienen las proteínas más representativas del subgrupo y buscar posibles dianas terapéuticas.

El predictor de clase se compone de 14 proteínas capaces de clasificar a las muestras ER+ en *ER-true* y *TN-like*. De las proteínas más expresadas en tumores *ER-true*, cabe destacar que cinco pertenecen al grupo de los proteoglicanos (la decorina, el mimecan, el biglican, el lumican y la fibromodulina). Los proteoglicanos (PG) controlan numerosos procesos normales y patológicos, entre los que se encuentran la morfogénesis, la reparación de tejidos, la inflamación, la vascularización y la metástasis del cáncer. Durante el desarrollo y crecimiento tumoral, la expresión de los PG se modifica en el microambiente tumoral. La expresión alterada de proteoglicanos en las membranas de las células tumorales y estromales afecta la señalización, el crecimiento y la supervivencia de las células cancerosas, la adhesión celular, la migración y la angiogénesis (198).

La decorina es un PG *pequeño rico en leucina* (SLRP, del inglés *small leucine rich proteoglycan*), siendo un factor clave en la dinámica de la matriz celular que da como resul-

tado una multitud de fenotipos celulares y biológicos. Es un inhibidor pan-receptor tirosina quinasa (RTK) que afecta a la función del receptor en múltiples niveles, incluida la modulación y biodisponibilidad de los ligandos del receptor, para la supresión tumorigénica y metastásica. Se le considera el "guardián de la matriz" ya que la deficiencia de decorina es permisiva para la tumorigénesis y aumenta la actividad basal de receptores múltiples de tirosina quinasa. La disminución de la expresión de decorina y lumican en los tejidos de cáncer de mama se asocia con un mayor tamaño tumoral, mayor riesgo y tasa de recaída, estado negativo de ER / PRg y baja supervivencia en pacientes con cáncer de mama con ganglios negativos (199). En pacientes con RH positivos la alta expresión de decorina se asocia con metástasis ganglionares, mayor número de ganglios linfáticos positivos y peor supervivencia general y la alta expresión de HSP90B1 se asocia con metástasis a distancia y disminuye la supervivencia general (200). Las SLRP se expresan de forma ubicua en la mayoría de las matrices extracelulares y afectan la progresión del cáncer de mama de múltiples maneras (201).

El mimecan, también llamado osteoglicina, es una proteína humana que está codificada por el gen OGN. Induce la formación ectópica ósea en conjunción con el factor de crecimiento transformante beta. Los niveles de expresión de este gen están relacionados con la hipertrofia ventricular izquierda (202). Los biglicanos se consideran inhibidores de la proliferación celular y supresores del ciclo celular, promueven la angiogénesis, la migración celular y la inflamación (203–206). El lumican se expresa de manera ubicua y se localiza principalmente en tejidos mesenquimatosos y estroma tumoral del cáncer de mama (207)

La fibromodulina se expresa en el tejido conjuntivo denso. Tiene un papel en el ensamblaje y el mantenimiento del colágeno y retrasa la formación de fibrillas del colágeno. Se une a los colágenos I y II. En tumores sólidos, la fibromodulina parece modular el estroma tumoral al aumentar el volumen del líquido extracelular y la disminución de la presión del líquido intersticial (208). Tanto *in vitro* como *in vivo* promueve la angiogénesis (209).

La mayoría de las proteínas de colágeno mostraron una expresión más alta en tumores ER-true. La cadena alfa 1 del colágeno 1 es una proteína codificada por el gen COL1A1. Es el componente principal del colágeno tipo I, el colágeno fibrilar que se encuentra en la mayoría de los tejidos conectivos, incluido el cartílago. Aunque el colágeno tipo I induce la muerte celular apoptótica en células de carcinoma de mama lumi-

nal pero no en células de carcinoma de mama de tipo basal (210), parece tener un papel importante en el desarrollo y la iniciación de metástasis en células de cáncer de mama (211). El colágeno VI es una proteína de la matriz extracelular (ECM) compuesta por tres cadenas polipeptídicas, codificadas por los genes COL6A1, COL6A2 y COL6A3, que forman una red de microfilamentos que interactúan con otras moléculas de ECM y proporcionan soporte estructural para las células. Además, desencadena vías de señalización relacionadas con el cáncer, como la apoptosis, la autofagia, la proliferación, la angiogénesis y la inflamación (212)

Por último, la baja expresión de HSPB6, también llamada HSP20, está asociada a progresión tumoral en diversos cánceres (ovario, hepatocarcinoma, colon, etc.) (213–215). No se conoce relación previa con el cáncer de mama.

Los tumores TN-*like* mostraron características moleculares que podrían conferir un mal pronóstico. En nuestra cohorte, en el subgrupo TN-*like* se objetivó una sobreexpresión de la HSP90AB1 o endoplasmína, de la cadena alfa del colágeno 14 (más expresada en el grupo de tumores ER-*true* en la cohorte de descubrimiento), la proteína procolágeno galactosiltransferasa 1 y la subunidad ϵ del complejo proteico T. Las proteínas de choque térmico o proteínas de stress (HSP) son miembros de las chaperonas moleculares, un grupo de proteínas que desempeñan un papel esencial en el plegamiento de un gran número de proteínas celulares. Intervienen en la supervivencia celular durante la hipertermia al inhibir la muerte celular programada y la senescencia celular. Parece que intervienen en la carcinogénesis al permitir que las células escapen de las vías de supresión tumoral, para promover la progresión, resistencia al tratamiento y facilitar las metástasis (216). Las proteínas de choque térmico HSP90 junto con la HSP70 facilitan el plegamiento de proteínas, pero también pueden dirigir proteínas para la degradación mediada por ubiquitina (217). El mecanismo que regula estas funciones opuestas es la unión a otras co-chaperonas. Se ha encontrado una expresión elevada de HSP90 en el carcinoma de mama invasivo, pero no en el carcinoma *in situ* cuando se han comparado con el tejido normal adyacente y en carcinoma de mama ductal invasivo con receptores positivos comparados con los triple negativos (216). Sin embargo, en otros estudios no se ha demostrado una expresión diferencial entre los tumores triple negativos y receptores de estrógenos positivos de la proteína HSP90 (185). La expresión elevada de HSP90 está asociada a un mayor riesgo de muerte por cáncer de mama en pacientes con receptores positivos y HER2 negativo y mayor recurrencia a distancia en pacientes con

tumores triple negativos (218,219), lo cual es consistente con nuestros resultados, mostrando una mayor expresión de esta proteína en tumores TN-*like*, los cuales presentaron mayor recurrencia a distancia que el subgrupo ER-*true*.

La cadena alfa del colágeno de tipo XIV desempeña un papel adhesivo al integrar haces de colágeno y se sobre-expresa en pacientes con cáncer de mama con afectación ganglionar múltiple (pN3) en comparación con la no existencia de afectación ganglionar (pN0) (220). Esta proteína estaba más expresada en el grupo de tumores ER-*true* en la cohorte de descubrimiento. *T-complex protein 1 subunit eta* está codificada por el gen CCT7. Es una chaperona molecular que ayuda al plegamiento de proteínas relacionadas con la hidrólisis de ATP. Conocida por desempeñar un papel, *in vitro*, en el plegamiento de actina y tubulina. Los estudios publicados están realizados en muestras celulares. No hay datos en tejidos (221). La proteína procolágeno galactosiltransferasa 1 está codificada por el gen COLGALT1. Está relacionada con el colágeno. No hay estudios en tejido mamario (221).

El presente trabajo presenta varias limitaciones. En primer lugar, el número de muestras analizadas fue relativamente pequeño. El número estuvo limitado por las muestras disponibles, pero se prefirió utilizar aquellas accesibles desde ALAMO, por ser esta una base de datos de calidad contrastada.

Por otra parte, la SRM tiene sus propias limitaciones: la dificultad para la cuantificación de proteínas (46 péptidos), la existencia de decaje en el tiempo de retención en la columna de la cromatografía líquida y la no detección de algunas proteínas que sí pueden ser identificadas mediante otras técnicas. En nuestro caso en particular, el perfil original de 33 proteínas detectado por espectrometría de masas quedó reducido a 14 proteínas con SRM porque la correlación del resto era muy baja. Confirmamos que estas 14 proteínas identificaban correctamente al grupo TN-*like*, pero sin duda se pierde parte de la información con esta técnica.

Finalmente, por el momento la definición del tipo TN-*like* proporciona información pronóstica, pero no indica cómo debe tratarse. En el futuro se deberá investigar el papel que las principales proteínas juegan para ver si pueden convertirse en dianas terapéuticas.

En definitiva, este trabajo ha validado la existencia de un subgrupo de tumores de mama de mal pronóstico dentro del grupo general de receptores hormonales positivos y HER2

negativo. El subgrupo se caracteriza por una firma de adhesión detectada mediante proteómica de alto rendimiento. El trabajo subraya la necesidad de profundizar en la biología molecular del cáncer de mama y la posibilidad de incorporar la proteómica a otras plataformas de alto rendimiento, con objeto de proporcionar una información más completa que refleje toda la complejidad de la enfermedad.

6. CONCLUSIONES

6. CONCLUSIONES

1. La “firma de adhesión” identifica a dos grupos de pacientes con cáncer de mama de características luminales: *ER-true* y *TN-like*. Estos dos grupos tienen distintas características moleculares y distinta evolución clínica.
2. El análisis del estatus de HER2 por PCR cuantitativa ha demostrado ser un método de eficacia comparable a la práctica clínica habitual de determinación mediante IHC y FISH.
3. Es posible la cuantificación de proteínas en muestras fijadas en formol y embebidas en parafina mediante la técnica SRM.
4. La correlación de dicha cuantificación entre espectrometría de masas no dirigida y SRM es muy heterogénea para las proteínas analizadas, lo que sugiere una discordancia parcial entre ambas técnicas.
5. Es posible generar un predictor de clase basado en la “firma de adhesión” que divide al cáncer de mama con receptores hormonales positivos y HER negativo en dos grupos, *ER-true* y *TN-like*, empleando valores de cuantificación de proteínas por SRM.
6. El análisis multivariante muestra que las diferencias en la evolución clínica observadas entre los grupos *ER-true* y *TN-like* son independientes de las variables clínicas analizadas (tamaño tumoral, estatus ganglionar y grado de diferenciación).

7. BIBLIOGRAFÍA

7. BIBLIOGRAFÍA

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* . 2015;136(5):E359-E386.
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65(2):87-108.
3. Althuis MD, Dozier JM, Anderson WF, Devesa SS, Brinton LA. Global trends in breast cancer incidence and mortality 1973-1997. *Int J Epidemiol*. 2005;34(2):405–12.
4. Ravdin PM, Cronin KA, Howlader N, Berg CD, Chlebowski RT, Feuer EJ, et al. The Decrease in Breast-Cancer Incidence in 2003 in the United States. *N Engl J Med*. 2007;356(16):1670–4.
5. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women’s Health Initiative randomized controlled trial. *JAMA*. 2002;288(3):321–33.
6. Youlten DR, Cramb SM, Dunn NAM, Muller JM, Pyke CM, Baade PD. The descriptive epidemiology of female breast cancer: An international comparison of screening, incidence, survival and mortality. *Cancer Epidemiology*. 2012;36(3):237-48.
7. Torre LA, Siegel RL, Ward EM, Jemal A. Global Cancer Incidence and Mortality Rates and Trends--An Update. *Cancer Epidemiol Biomarkers Prev*. 2016;25(1):16–27.
8. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin*. 2017;67(1):7–30.
9. Miller KD, Siegel RL, Lin CC, Mariotto AB, Kramer JL, Rowland JH, et al. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin*. 2016;66(4):271–89.
10. Ferlay J. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer*. 2013;49:1374–403.
11. Sant M, Chirlaque Lopez MD, Agresti R, Sánchez Pérez MJ, Hollecsek B, Bielska-Lasota M, et al. Survival of women with cancers of breast and genital

- organs in Europe 1999-2007: Results of the EUROCORE-5 study. *Eur J Cancer*. 2015;51(15):2191–205.
12. Mortalidad de cáncer en España. 2015;80–1.
 13. Sant M, Chirlaque Lopez MD, Agresti R, Sánchez Pérez MJ, Holleccek B, Bielska-Lasota M, et al. Survival of women with cancers of breast and genital organs in Europe 1999-2007: Results of the EUROCORE-5 study. *Eur J Cancer*. 2016;51(15):2191–205.
 14. DeVita VT, Lawrence TS RS. DeVita, Hellman, and Rosenberg's Cancer: Principles and Practice of Oncology. Wolters Kluwer Health/Lippincott Williams & Wilkins; 2011. 924-54 p.
 15. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin*. 2016;66(1):7–30.
 16. Collaborative Group on Hormonal Factors in Breast Cancer J, Shipley M, Lewington S, al. et, Bull D, Collaborators for the MWS. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol*. 2012;13(11):1141–51.
 17. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52 705 women with breast cancer and 108 411 women without breast cancer. *Lancet*. 1997;350(9084):1047–59.
 18. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. *Lancet*. 1996 Jun 22;347(9017):1713–27.
 19. Zhu H, Lei X, Feng J, Wang Y. Oral contraceptive use and risk of breast cancer: A meta-analysis of prospective cohort studies. *Eur J Contracept Reprod Heal Care*. 2012;17(6):402–14.
 20. Collaborative Group on Hormonal Factors in Breast Cancer R, Lopez A, Boreham J, Thun M, Heath C, Persson I. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet*. 2001;358(9291):1389–99.
 21. Fackenthal JD, Olopade OI. Breast cancer risk associated with BRCA1 and

- BRCA2 in diverse populations. *Nature Reviews Cancer*. 2007;7(12):937-48.
22. Walsh T, King M-C, Nakanishi K, Ohashi A, Wu J, Christ N, et al. Ten genes for inherited breast cancer. *Cancer Cell*. 2007;11(2):103–5.
 23. Mouchawar J, Korch C, Byers T, Pitts TM, Li E, McCredie MRE, et al. Population-based estimate of the contribution of TP53 mutations to subgroups of early-onset breast cancer: Australian breast cancer family study. *Cancer Res*. 2010;70(12):4795–800.
 24. Tan MH, Mester JL, Ngeow J, Rybicki LA, Orloff MS, Eng C. Lifetime cancer risks in individuals with germline PTEN mutations. *Clin Cancer Res*. 2012;18(2):400–7.
 25. Pharoah PDP, Guilford P, Caldas C, Jolley D, Giles G, McCredie M, et al. Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. *Gastroenterology*. 2001;121(6):1348–53.
 26. Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, et al. E-cadherin germline mutations in familial gastric cancer. *Nature*. 1998;392(6674):402–5.
 27. Lim W, Olschwang S, Keller JJ, Westerman AM, Menko FH, Boardman LA, et al. Relative frequency and morphology of cancers in STK11 mutation carriers1. *Gastroenterology*. 2004;126(7):1788–94.
 28. Foulkes WD. Inherited Susceptibility to Common Cancers. *N Engl J Med*. 2008;20359(13):2143–53.
 29. Hartmann LC, Sellers T a., Frost MH. Benign Breast Disease and the Risk of Breast Cancer. *N Engl J Med*. 2005;353(3):981–92.
 30. Dyrstad SW, Yan Y, Fowler AM, Colditz GA. Breast cancer risk associated with benign breast disease: systematic review and meta-analysis. *Breast Cancer Res Treat*. 2015;149(3):569–75.
 31. Sickles, EA, D’Orsi CJ, Bassett LW et al. ACR BI-RADS® Mammography. Reston, VA AC of R 2013., editor. ACR BI-RADS® Atlas, Breast Imaging Report Data Syst. 5th ed. 2013;135.
 32. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic Density and the Risk and Detection of Breast Cancer. *N Engl J Med*. 2007;356(3):227–36.
 33. De Bruin ML, Sparidans J, Van ’t Veer MB, Noordijk EM, Louwman MWJ,

- Zijlstra JM, et al. Breast Cancer Risk in Female Survivors of Hodgkin's Lymphoma: Lower Risk After Smaller Radiation Volumes. *J Clin Oncol*. 2009;27:4239–46.
34. Travis LB, Hill D, Dores GM, Gospodarowicz M, Van Leeuwen FE, Holowaty E, et al. Cumulative Absolute Breast Cancer Risk for Young Women Treated for Hodgkin Lymphoma. *J Natl Cancer Inst*. 2005;97(19).
 35. Chan DSM, Vieira AR, Aune D, Bandera E V, Greenwood DC, McTiernan A, et al. Body mass index and survival in women with breast cancer-systematic literature review and meta-analysis of 82 follow-up studies. *Ann Oncol*. 2014;25(10):1901–14.
 36. Niraula S, Ocana A, Ennis M, Goodwin PJ. Body size and breast cancer prognosis in relation to hormone receptor and menopausal status: a meta-analysis. *Breast Cancer Res Treat*. 2012;134(2):769–81.
 37. Bagnardi V, Rota M, Botteri E, Tramacere I, Islami F, Fedirko V, et al. Alcohol consumption and site-specific cancer risk: a comprehensive dose–response meta-analysis. *Br J Cancer*. 2015;112(3):580–93.
 38. Scoccianti C, Lauby-Secretan B, Bello P-Y, Chajes V, Romieu I. Female breast cancer and alcohol consumption: a review of the literature. *Am J Prev Med*. 2014;46(3 Suppl 1):S16-25.
 39. Fleming, ID. Cooper, JS. Henson, DE. Hutter, RV. Kennedy, BJ. Murphy G et al. *AJCC Cancer Staging Manual*. 5th ed. Philadelphia: Lippincott-Raven, editor. 1997. 31-46 p.
 40. Sobin LH, Fleming ID. TNM classification of malignant tumors, Fifth edition (1997). *Cancer*. John Wiley & Sons, Inc.; 1997;80(9):1803–4.
 41. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*. 1991;19(5):403–10.
 42. Harris LN, Ismaila N, McShane LM, Andre F, Collyar DE, Gonzalez-Angulo AM, et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol*. 2016;34(10):1134–50.
 43. Lakhani, S.R., Ellis. I.O., Schnitt, S.J., Tan, P.H., van de Vijver MJ. WHO Classification of Tumours of the Breast, Fourth Edition - WHO - OMS - .

- 2012;240.
44. Rakha E a, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res.* 2010;12(4):207.
 45. Konecny G, Pauletti G, Pegram M, Untch M, Dandekar S, Aguilar Z, et al. Quantitative association between HER-2/neu and steroid hormone receptors in hormone receptor-positive primary breast cancer. *J Natl Cancer Inst.* 2003;95(2):142–53.
 46. Dawson SJ, Provenzano E, Caldas C. Triple negative breast cancers: Clinical and prognostic implications. *Eur J Cancer.* 2009;45(SUPPL. 1):27–40.
 47. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci.* 2001;98(19):10869–74.
 48. Perou C. Molecular portraits of human breast tumours. *Nature.* 2000;406:747.
 49. Cheang MCU, Chia SK, Voduc D, Gao D, Leung S, Snider J, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst.* 2009;101(10):736–50.
 50. Kennecke H, Yerushalmi R, Woods R, Cheang MCU, Voduc D, Speers CH, et al. Metastatic behavior of breast cancer subtypes. *J Clin Oncol.* 2010;28(20):3271–7.
 51. Kennecke H. Metastatic Behavior of Breast Cancer Subtypes. *J Clin Oncol.* 2010;28:3271.
 52. Voduc KD, Cheang MCU, Tyldesley S, Gelmon K, Nielsen TO, Kennecke H. Breast cancer subtypes and the risk of local and regional relapse. *J Clin Oncol.* 2010;28(10):1684–91.
 53. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. *N Engl J Med.* 2001;344(11):783–92.
 54. Slamon D, Eiermann W, Robert N, Pienkowski T, Martin M, Press M. Breast Cancer International Research Group. Adjuvant trastuzumab in HER2-positive breast cancer. *N Engl J Med.* 2011;365(14):1273-83.
 55. Weigelt B, Mackay A, A'hern R, Natrajan R, Tan DSP, Dowsett M, et al. Breast cancer molecular profiling with single sample predictors: A retrospective analysis. *Lancet Oncol.* 2010;11(4):339–49.

56. Kreike B, van Kouwenhove M, Horlings H, Weigelt B, Peterse H, Bartelink H, et al. Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas. *Breast Cancer Res.* 2007;9(5):R65.
57. Livasy CA, Karaca G, Nanda R, Tretiakova MS, Olopade OI, Moore DT, et al. Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Mod Pathol.* 2006;19(2):264–71.
58. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, et al. Triple-negative breast cancer: Clinical features and patterns of recurrence. *Clin Cancer Res.* 2007;13(15):4429–34.
59. Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JGM, et al. Subtypes of breast cancer show preferential site of relapse. *Cancer Res.* 2008;68(9):3108–14.
60. Sørli T. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *PNAS.* 2001;98:10869.
61. Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res.* 2004;10(16):5367–74.
62. Cheang MCU, Martin M, Nielsen TO, Prat A, Voduc D, Rodriguez-Lescure A, et al. Defining Breast Cancer Intrinsic Subtypes by Quantitative Receptor Expression. *Oncologist* [Internet]. 2015 [cited 2017 Sep 5];20:474–82. Available from: <http://theoncologist.alphamedpress.org/content/20/5/474.full.pdf>
63. Prat A, Adamo B, Cheang MCU, Anders CK, Carey LA, Perou CM. Molecular Characterization of Basal-Like and Non-Basal-Like Triple-Negative Breast Cancer. *Oncologist.* 2013;18(2):123–33.
64. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol.* 2016;13(11):674–90.
65. Lehmann BDB, Bauer J a J, Chen X, Sanders ME, Chakravarthy a B, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest.* 2011;121(7):2750–67.
66. Lehmann BD, Jovanović B, Chen X, Estrada M V, Johnson KN, Shyr Y, et al. Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One.* 2016;11(6):e0157368.
67. Burstein MD, Tsimelzon A, Poage GM, Covington KR, Contreras A, Fuqua SAW, et al. Comprehensive genomic analysis identifies novel subtypes and

- targets of triple-negative breast cancer. *Clin Cancer Res.* 2015;21(7):1688–98.
68. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci.* 2003;100(14):8418–23.
 69. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol.* 2007;8(5):R76.
 70. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* 2010;12(5):R68.
 71. Prat A. Deconstructing the molecular portraits of breast cancer. *Mol Oncol.* 2011;5:5.
 72. Dawson SJ, Rueda OM, Aparicio S, Caldas C. A new genome-driven integrated classification of breast cancer and its implications. *EMBO Journal.* 2013. p. 617–28.
 73. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumors. *Nature.* 2012;490(7418):61–70.
 74. Eifel P, Axelson JA, Costa J, Crowley J, Curran WJ, Deshler A, et al. National Institutes of Health Consensus Development Conference Statement: adjuvant therapy for breast cancer, November 1-3, 2000. *J Natl Cancer Inst.* 2001;93(13):979–89.
 75. Winchester DP, Osteen RT, Menck HR. The National Cancer Data Base report on breast carcinoma characteristics and outcome in relation to age. *Cancer.* 1996;78(8):1838–43.
 76. Kroman N, Jensen MB, Wohlfahrt J, Mouridsen HT, Andersen PK, Melbye M. Factors influencing the effect of age on prognosis in breast cancer: population based study. *BMJ.* 2000;320(7233):474–8.
 77. Carter C. Relation of tumor size, lymph node status and survival in 24,740 breast cancer cases. *Cancer.* 1989;63:181–7.
 78. Fisher B, Dignam J, Tan-Chiu E, Anderson S, Fisher ER, Wittliff JL, et al. Prognosis and treatment of patients with breast tumors of one centimeter or less and negative axillary lymph nodes. *J Natl Cancer Inst.* 2001;93(2):112–20.
 79. Fisher ER, Costantino J, Fisher B, Redmond C. Pathologic findings from the National Surgical Adjuvant Breast Project (Protocol 4). Discriminants for 15-

- year survival. National Surgical Adjuvant Breast and Bowel Project Investigators. *Cancer*. 1993;71(6 Suppl):2141–50.
80. Rosen PP, Groshen S, Saigo PE, Kinne DW, Hellman S. Pathological prognostic factors in stage I (T1N0M0) and stage II (T1N1M0) breast carcinoma: a study of 644 patients with median follow-up of 18 years. *J Clin Oncol*. 1989;7(9):1239–51.
81. Ellis IO, Galea M, Broughton N, Locker A, Blamey RW, Elston CW. Pathological prognostic factors in breast cancer. II. Histological type. Relationship with survival in a large study with long-term follow-up. *Histopathology*. 1992;20(6):479–89.
82. Bloom HJ RW. Histological grading and prognosis of breast cancer. *Br J Cancer*. 1957;11:359–77.
83. Rakha EA, El-Sayed ME, Lee AHS, Elston CW, Grainge MJ, Hodi Z, et al. Prognostic significance of nottingham histologic grade in invasive breast carcinoma. *J Clin Oncol*. 2008;26(19):3153–8.
84. Lee AHS, Pinder SE, Macmillan RD, Mitchell M, Ellis IO, Elston CW, et al. Prognostic value of lymphovascular invasion in women with lymph node negative invasive breast carcinoma. *Eur J Cancer*. 2006;42(3):357–62.
85. Rakha EA, Martin S, Lee AHS, Morgan D, Pharoah PDP, Hodi Z, et al. The prognostic significance of lymphovascular invasion in invasive breast carcinoma. *Cancer*. 2012;118(15):3670–80.
86. Clark GM, McGuire WL. Steroid receptors and other prognostic factors in primary breast cancer. *Semin Oncol*. 1988;15(2 Suppl 1):20–5.
87. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*. 1987;235(4785):177–82.
88. Ross JS, Fletcher JA. The HER-2/neu Oncogene in Breast Cancer: Prognostic Factor, Predictive Factor, and Target for Therapy. *Stem Cells*. 1998;16(6):413–28.
89. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA. Ki67 in breast cancer: prognostic and predictive potential. *The Lancet Oncology*. 2010. p. 174–83.
90. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ. Strategies for subtypes-dealing with the diversity of breast cancer: Highlights of

- the St Gallen international expert consensus on the primary therapy of early breast cancer 2011. *Ann Oncol.* 2011;22(8):1736–47.
91. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, et al. Personalizing the treatment of women with early breast cancer: Highlights of the st gallen international expert consensus on the primary therapy of early breast Cancer 2013. *Ann Oncol* 2013;24(9):2206–23.
 92. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, et al. Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol.* 2015;26(8):1533–46.
 93. Focke CM, van Diest PJ, Decker T. St Gallen 2015 subtyping of luminal breast cancers: impact of different Ki67-based proliferation assessment methods. *Breast Cancer Res Treat.* 2016;159(2):257–63.
 94. Polley MYC, Leung SCY, McShane LM, Gao D, Hugh JC, Mastropasqua MG, et al. An international ki67 reproducibility study. *J Natl Cancer Inst.* 2013;105(24):1897–906.
 95. Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) initiative: Methods of the EGAPP Working Group. *Genet Med.* 2009;11(1):3–14.
 96. Botkin JR, Teutsch SM, Kaye CI, Hayes M, Haddow JE, Bradley LA, et al. Outcomes of interest in evidence-based evaluations of genetic tests. *Genet Med.* 2010;12(4):228–35.
 97. Paik S. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med.* 2004;351(27):2817–26.
 98. Habel LA, Shak S, Jacobs MK, Capra A, Alexander C, Pho M, et al. A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. *Breast Cancer Res.* 2006.;8(3):R25.
 99. Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, et al. Prediction of Risk of Distant Recurrence Using the 21-Gene Recurrence Score in Node-Negative and Node-Positive Postmenopausal Patients With Breast Cancer Treated With Anastrozole or Tamoxifen: A TransATAC Study. *J Clin Oncol.* 2010;28:1829–34.
 100. Toi M, Iwata H, Yamanaka T, Masuda N, Ohno S, Nakamura S, et al. Clinical significance of the 21-gene signature (Oncotype DX) in hormone receptor-

- positive early stage primary breast cancer in the Japanese population. *Cancer*. 2010;116(13):3112–8.
101. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol*. 2006;24.
 102. Sparano J a. TAILORx: trial assigning individualized options for treatment (Rx). *Clin Breast Cancer*. 2006;7(4):347–50.
 103. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med*. 2015;373(21):2005–14.
 104. Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, et al. Prediction of Risk of Distant Recurrence Using the 21-Gene Recurrence Score in Node-Negative and Node-Positive Postmenopausal Patients With Breast Cancer Treated With Anastrozole or Tamoxifen: A TransATAC Study. *J Clin Oncol*. 2010;28(11):1829–34.
 105. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT. Breast Cancer Intergroup of North America. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomise. *Lancet Oncol*. 2010;11(1):55-65.
 106. Goldstein LJ, Gray R, Badve S, Childs BH, Yoshizawa C, Rowley S, et al. Prognostic Utility of the 21-Gene Assay in Hormone Receptor-Positive Operable Breast Cancer Compared With Classical Clinicopathologic Features. *Clin Oncol J Clin Oncol*. 2008;26(25):4063–71.
 107. Ramsey SD, Barlow WE, Gonzalez-Angulo AM, Tunis S, Baker L, Crowley J, et al. Integrating comparative effectiveness design elements and endpoints into a phase III, randomized clinical trial (SWOG S1007) evaluating oncotypedX-guided management for women with breast cancer involving lymph nodes. *Contemp Clin Trials*. 2013;34(1):1–9.
 108. Nitz U, Gluz O, Christgen M, Kates RE, Clemens M, Malter W, et al. Reducing chemotherapy use in clinically high-risk, genomically low-risk pN0 and pN1 early breast cancer patients: five-year data from the prospective, randomised phase 3 West German Study Group (WSG) PlanB trial. *Breast Cancer Res Treat*. 2017;165(3):573–83.

109. Stemmer SM, Steiner M, Rizel S, Geffen DB, Nisenbaum B, Peretz T, et al. Clinical outcomes in ER+ HER2 -node-positive breast cancer patients who were treated according to the Recurrence Score results: evidence from a large prospectively designed registry. *npj Breast Cancer*. 2017;3(1):32.
110. Petkov VI, Miller DP, Howlader N, Gliner N, Howe W, Schussler N, et al. Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study. *npj Breast Cancer*. 2016;2(1):16017.
111. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *Lancet Oncol*. 2010;11(1):55–65.
112. Van't Veer L. Gene expression profiling predicts clinical outcome of breast cancer . *Nature*. 2002;(415):530.
113. Goldhirsch A, Glick JH, Gelber RD, Coates AS, Senn H-J. Meeting Highlights: International Consensus Panel on the Treatment of Primary Breast Cancer. *J Clin Oncol*. 1991;19(18):3817–27.
114. van de Vijver MJ, He YD, van 't Veer LJ, Dai H, Hart AAM, Voskuil DW, et al. A Gene-Expression Signature as a Predictor of Survival in Breast Cancer. *Nejm*. 2002;347(25):1999–2009.
115. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, et al. Validation and Clinical Utility of a 70-Gene Prognostic Signature for Women With Node-Negative Breast Cancer. *JNCI J Natl Cancer Inst*. 2006;98(17):1183–92.
116. Drukker CA, Bueno-De-Mesquita JM, Retèl VP, Van Harten WH, Van Tinteren H, Wesseling J, et al. A prospective evaluation of a breast cancer prognosis signature in the observational RASTER study. *Int J Cancer*. 2013;133(4):929–36.
117. Bueno-de-Mesquita JM, Linn SC, Keijzer R, Wesseling J, Nuyten DS, van Krimpen C et al., Linn SC, Keijzer R, Wesseling J, Nuyten DSA, Van Krimpen C, et al. Validation of 70-gene prognosis signature in node-negative breast cancer. *Breast Cancer Res Treat*. 2009;117(3):483–95.
118. Winner BS, Sgroi DC, Ryan PD, Bruinsma TJ, Glas AM, Male A, et al. Analysis of the mamma print breast cancer assay in a predominantly postmenopausal cohort. *Clin Cancer Res*. 2008;14(10):2988–93.
119. Mook S. the 70-gene prognosis signature predicts disease outcome in breast

- cancer patients with 1-3 positives lymph nodes in an independent validation study. *Breast Cancer Res Treat.* 2009;116(2):295–302.
120. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med.* 2016; 376(8):717-29.
121. Krop I, Ismaila N, Andre F, Bast RC, Barlow W, Collyar DE, et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American society of clinical oncology clinical practice guideline focused update. *J Clin Oncol.* 2017;35(24):2838–47.
122. Filipits M, Rudas M, Jakesz R, Dubsy P, Fitzal F, Singer CF, et al. A New Molecular Predictor of Distant Recurrence in ER-Positive, HER2-Negative Breast Cancer Adds Independent Information to Conventional Clinical Risk Factors. *Clin Cancer Res.* 2011;17(18):6012–20.
123. Schmid M, Jakesz R, Samonigg H, Kubista E, Gnant M, Menzel C, et al. Randomized trial of tamoxifen versus tamoxifen plus aminoglutethimide as adjuvant treatment in postmenopausal breast cancer patients with hormone receptor-positive disease: Austrian Breast and Colorectal Cancer Study Group Trial 6. *J Clin Oncol.* 2003;21(6):984–90.
124. Dubsy PC, Jakesz R, Fitzal F, Singer CF, Rudas M, Bartsch R, et al. Tamoxifen and anastrozole as a sequencing strategy: A randomized controlled trial in postmenopausal patients with endocrine-responsive early breast cancer from the Austrian Breast and Colorectal Cancer Study Group. *J Clin Oncol.* 2012;30(7):722–8.
125. Dubsy P, Brase JC, Jakesz R, Rudas M, Singer CF, Greil R, et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2– breast cancer patients. *Br J Cancer.* 2013;109(12):2959–64.
126. Dubsy P, Filipits M, Jakesz R, Rudas M, Rudas M, Greil R, et al. Endopredict improves the prognostic classification derived from common clinical guidelines in ER-positive, HER2-negative early breast cancer. *Ann Oncol.* 2013;24(3):640–7.
127. Kronenwett R, Bohmann K, Prinzler J, Sinn B V, Haufe F, Roth C, et al. Decentral gene expression analysis: analytical validation of the Endopredict genomic multianalyte breast cancer prognosis test. *BMC Cancer.* 2012;12:456–66.

128. Martin M, Brase JC, Calvo L, Krappmann K, Ruiz-Borrego M, Fisch K, et al. Clinical validation of the EndoPredict test in node-positive, chemotherapy-treated ER+/HER2- breast cancer patients: results from the GEICAM 9906 trial. *Breast Cancer Res.* 2014;16(2):R38.
129. Parker J. Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *J Clin Oncol.* 2009;27:1160.
130. Nielsen T, Wallden B, Schaper C, Ferree S, Liu S, Gao D, et al. Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. *BMC Cancer.* 2014;14(1):177.
131. Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J Clin Oncol.* 2013;31(22).
132. Sestak I, Dowsett M, Zabaglo L, Lopez-Knowles E, Ferree S, Cowens JW, et al. Factors predicting late recurrence for estrogen receptor-positive breast cancer. *J Natl Cancer Inst.* 2013;105(19):1504–11.
133. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. *Ann Oncol.* 2014;25(2):339–45.
134. Gnant M, Sestak I, Filipits M, Dowsett M, Balic M, Lopez-Knowles E, et al. Identifying clinically relevant prognostic subgroups of postmenopausal women with node-positive hormone receptor-positive early-stage breast cancer treated with endocrine therapy: A combined analysis of ABCSG- 8 and ATAC using the PAM50 risk of recurrence. *Ann Oncol.* 2015;26(8):1685–91.
135. Xiao-Jun M, Salunga R, Dahiya S, Wang W, Carney E, Durbecq V, et al. A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer. *Clin Cancer Res.* 2008;14(9):2601–8.
136. Jerevall P-L, Ma X-J, Li H, Salunga R, Kesty NC, Erlander MG, et al. Prognostic utility of HOXB13:IL17BR and molecular grade index in early-stage breast cancer patients from the Stockholm trial. *Br J Cancer.* 2011;104(11):1762–9.

137. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al. Gene Expression Profiling in Breast Cancer: Understanding the Molecular Basis of Histologic Grade To Improve Prognosis. *JNCI J Natl Cancer Inst.* 2006;98(4):262–72.
138. Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the genomic health recurrence score in early breast cancer. *J Clin Oncol.* 2011;29(32):4273–8.
139. Harbeck N, Sotlar K, Wuerstlein R, Doisneau-Sixou S. Molecular and protein markers for clinical decision making in breast cancer: Today and tomorrow. *Cancer Treatment Reviews.* 2014;40(3):434-44.
140. Kwa M, Makris A, Esteva FJ. Clinical utility of gene-expression signatures in early stage breast cancer. *Nature Reviews Clinical Oncology.* 2017;14(10):595-610.
141. Peto R, Davies C, Godwin J, Gray R, Pan HC. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. *Lancet.* 2012;379(9814):432-44.
142. Wishart GC, Azzato EM, Greenberg DC, Rashbass J, Kearins O, Lawrence G, et al. PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Res.* 2010;12(1):R1.
143. Gianni L, Pienkowski T, Im YH, Roman L, Tseng LM, Liu MC, et al. Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): A randomised multicentre, open-label, phase 2 trial. *Lancet Oncol.* 2012;13(1):25–32.
144. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I. Herceptin Adjuvant (HERA) Trial Study Team. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med.* 2005;353(16):1659-72.
145. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE, Davidson NE. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med.* 2005;353(16):1673-84.
146. Swain SM, Kim SB, Cortés J, Ro J, Semiglazov V, Campone M. Pertuzumab,

- trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA study): overall survival results from a randomised, double-blind, placebo-controlled, phase 3 study. *Lancet Oncol.* 2013;14. (6):461-71.
147. Consortium IHGS. Finishing the euchromatic sequence of the human genome. *Nature.* 2004;431(7011):931–45.
 148. Domon B. Mass Spectrometry and Protein Analysis. *Science.* 2006;312(5771):212–7.
 149. Chae YK, Gonzalez-Angulo AM. Implications of Functional Proteomics in Breast Cancer. *Oncologist.* 2014;19(4):328–35.
 150. O’Donovan C, Apweiler R, Bairoch A. The human proteomics initiative (HPI). *Trends in Biotechnology.* 2001;19(5):178–81.
 151. Lam SW, Jimenez CR, Boven E. Breast cancer classification by proteomic technologies: Current state of knowledge. *Cancer Treatment Reviews.* 2014;40(1):129-38.
 152. Zeidan BA, Townsend PA, Garbis SD, Copson E, Cutress RI. Clinical proteomics and breast cancer. *Surg.* 2015;13(5):271–8.
 153. Burnette WN. “Western Blotting”: Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem.* 1981;112(2):195–203.
 154. Kamel D, Brady B, Tabchy A, Mills G, Hennessy B. Proteomic Classification of Breast Cancer. *Curr Drug Targets.* 2012;13(12):1495–509.
 155. Goncalves A, Bertucci F. Clinical application of proteomics in breast cancer: State of the art and perspectives. *Med Princ Pract.* 2010;20(1):4–18.
 156. Last RL, Jones AD, Shachar-Hill Y. Towards the plant metabolome and beyond. *Nature Reviews Molecular Cell Biology.* 2007;8(2):167-74.
 157. O’Farrell P. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem.* 1975;250(10):4007–21.
 158. De St. Groth SF, Webster RG, Datyner A. Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. *Biochim Biophys Acta.* 1963;71(14):377–91.
 159. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal Chem.* 1996;68(5):850–8.
 160. Ünlü M, Morgan ME, Minden JS. Difference gel electrophoresis. A single gel

- method for detecting changes in protein extracts. *Electrophoresis*. 1997;18(11):2071–7.
161. Stults JT, Arnott D. Proteomics. *Methods in Enzymology*. 2005;402:245–89.
 162. Abián J, Carrascal M GM. Introducción a la Espectrometría de Masas para la caracterización de péptidos y proteínas en Proteómica. *Proteomica*. 2008;2:16–35.
 163. Espejel F, Carlos Roa J. Surface enhanced laser desorption/ionization (SELDI): tecnología proteómica y su aplicación en oncología. *Med Clin (Barc)*. 2008 Sep;131(8):312–7.
 164. Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: A tutorial. *Molecular Systems Biology*. 2008;4:222.
 165. Wasinger VC, Zeng M, Yau Y. Current Status and Advances in Quantitative Proteomic Mass Spectrometry. *Int J Proteomics*. 2013;2013:1–12.
 166. Zhu W, Smith JW, Huang CM. Mass spectrometry-based label-free quantitative proteomics. *Journal of Biomedicine and Biotechnology*. 2010; 840518.
 167. Pan S, Aebersold R, Chen R, Rush J, Goodlett DR, McIntosh MW, et al. Mass spectrometry based targeted protein quantification: Methods and applications. *Journal of Proteome Research*. 2009;8(2) 787–97.
 168. Cha S, Imielinski MB, Rejtar T, Richardson EA, Thakur D, Sgroi DC, et al. In Situ Proteomic Analysis of Human Breast Cancer Epithelial Cells Using Laser Capture Microdissection: Annotation by Protein Set Enrichment Analysis and Gene Ontology. *Mol Cell Proteomics*. 2010;9(11):2529–44.
 169. Neubauer H, Clare SE, Wozny W, Schwall GP, Poznanović S, Stegmann W, et al. Breast cancer proteomics reveals correlation between estrogen receptor status and differential phosphorylation of PGRMC1. *Breast Cancer Res*. 2008;10(5):R85.
 170. Rezaul K, Thumar JK, Lundgren DH, Eng JK, Claffey KP, Wilson L, et al. Differential Protein Expression Profiles in Estrogen Receptor-Positive and -Negative Breast Cancer Tissues Using Label-Free Quantitative Proteomics. *Genes Cancer*. 2010;1(3):251–71.
 171. Sahab ZJ, Man YG, Semaan SM, Newcomer RG, Byers SW, Sang QXA. Alteration in protein expression in estrogen receptor alpha-negative human breast cancer tissues indicates a malignant and metastatic phenotype. *Clin Exp Metastasis*. 2010;27(7):493–503.

172. Umar A, Kang H, Timmermans AM, Look MP, Meijer-van Gelder ME, den Bakker MA, et al. Identification of a Putative Protein Profile Associated with Tamoxifen Therapy Resistance in Breast Cancer. *Mol Cell Proteomics*. 2009;8(6):1278–94.
173. De Marchi T, Braakman RBH, Stingl C, van Duijn MM, Smid M, Foekens JA, et al. The advantage of laser-capture microdissection over whole tissue analysis in proteomic profiling studies. *Proteomics*. 2016;16(10):1474–85.
174. Traub F, Feist H, Kreipe HH, Pich A. SELDI-MS-based expression profiling of ductal invasive and lobular invasive human breast carcinomas. *Pathol Res Pract*. 2005;201(12):763–70.
175. Oliveira NCS, Gomig THB, Milioli HH, Cordeiro F, Costa GG, Urban CA, et al. Comparative proteomic analysis of ductal and lobular invasive breast carcinoma. *Genet Mol Res*. 2016;15(2):1–10.
176. Weitzel LRB, Byers T, Allen J, Finlayson C, Helmke SM, Hokanson JE, et al. Discovery and verification of protein differences between Er positive/Her2/neu negative breast tumor tissue and matched adjacent normal breast tissue. *Breast Cancer Res Treat*. 2010;124(2):297–305.
177. Abdullah Al-Dhabi N, Srigopalram S, Ilavenil S, Kim YO, Agastian P, Baaru R, et al. Proteomic analysis of stage-II breast cancer from formalin-fixed paraffin-embedded tissues. *Biomed Res Int*. 2016;2016.
178. Traub F, Jost M, Diger Hess R, Schorn K, Menzel C, Budde P, et al. Peptidomic analysis of breast cancer reveals a putative surrogate marker for estrogen receptor-negative carcinomas. *Lab Investig*. 2006;86:246–53.
179. Semaan SM, Wang X, Marshall AG, Sang QXA. Identification of potential glycoprotein biomarkers in estrogen receptor positive (ER+) and negative (ER-) human breast cancer tissues by LC-LTQ/FT-ICR mass spectrometry. *J Cancer*. 2012;3(1):269–84.
180. Bateman NW, Sun M, Bhargava R, Hood BL, Darfler MM, Kovatich AJ, et al. Differential proteomic analysis of late-stage and recurrent breast cancer from formalin-fixed paraffin-embedded tissues. *J Proteome Res*. 2011;10(3):1323–32.
181. De Marchi T, Liu NQ, Stingl C, Timmermans MA, Smid M, Look MP, et al. 4-Protein Signature Predicting Tamoxifen Treatment Outcome in Recurrent Breast Cancer. *Mol Oncol*. 2016;10(1):24–39.
182. Tyanova S, Albrechtsen R, Kronqvist P, Cox J, Mann M, Geiger T. Proteomic

- maps of breast cancer subtypes. *Nat Commun.* 2016;7:1–11.
183. Liu NQ, Stingl C, Look MP, Smid M, Braakman RBH, De Marchi T, et al. Comparative proteome analysis revealing an 11-protein signature for aggressive triple-negative breast cancer. *J Natl Cancer Inst.* 2014;106(2):djt376.
 184. Greenwood C, Metodieva G, Al-Janabi K, Lausen B, Alldridge L, Leng L, et al. Stat1 and CD74 overexpression is co-dependent and linked to increased invasion and lymph node metastasis in triple-negative breast cancer. *J Proteomics.* 2012;75(10):3031–40.
 185. Gámez-Pozo A. Combined Label-Free Quantitative Proteomics and microRNA Expression Analysis of Breast Cancer Unravel Molecular Differences with Clinical Implications. *Cancer Res.* 2015;75:2243–53.
 186. Lam SW, Jimenez CR, Boven E. Breast cancer classification by proteomic technologies: Current state of knowledge. *Cancer Treat Rev.* 2014;40(1):129–38.
 187. Gámez-Pozo A, Trilla-Fuertes L, Prado-Vázquez G, Chiva C, López-Vacas R, Nanni P, et al. Prediction of adjuvant chemotherapy response in triple negative breast cancer with discovery and targeted proteomics. *PLoS One.* 2017;12(6):1–16.
 188. Martín M, Llombart-Cussac A, Lluch A, Alba E, Munárriz B, Tusquets I, et al. Epidemiological study of the GEICAM group about breast cancer in Spain (1990-1993): El Alamo project. *Med Clin (Barc).* 2004;122(1):12–7.
 189. Martín M, Mahillo E, Llombart-Cussac A, Lluch A, Munarriz B, Pastor M, et al. The «El Álamo» project (1990-1997): Two consecutive hospital-based studies of breast cancer outcomes in Spain. *Clin Transl Oncol.* 2006;8(7):508–18.
 190. Hudis CA, Barlow WE, Costantino JP, Gray RJ, Pritchard KI, Chapman JAW, et al. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: The STEEP system. *J Clin Oncol.* 2007;25(15):2127–32.
 191. Sánchez-Navarro I, Gámez-Pozo A, Pinto Á, Hardisson D, Madero R, López R, et al. An 8-gene qRT-PCR-based gene expression score that has prognostic value in early breast cancer. *BMC Cancer.* 2010;10(336):1–10.
 192. Espinosa E, Sánchez-Navarro I, Gámez-Pozo A, Marin ÁP, Hardisson D, Madero R, et al. Comparison of prognostic gene profiles using qRT-PCR in paraffin samples: A retrospective study in patients with early breast cancer. *PLoS One.* 2009;4(6):e5911.
 193. Sánchez-Navarro I, Gámez-Pozo A, González-Barón M, Pinto-Marín Á,

- Hardisson D, López R, et al. Comparison of gene expression profiling by reverse transcription quantitative PCR between fresh frozen and formalin-fixed, paraffin-embedded breast cancer tissues. *Biotechniques*. 2010;48(5):389–97.
194. Desiere F. The PeptideAtlas project. *Nucleic Acids Res*. 2006;34(90001):D655–8.
195. Kusebauch U, Campbell DS, Deutsch EW, Chu CS, Spicer DA, Brusniak MY, et al. Human SRMAtlas: A Resource of Targeted Assays to Quantify the Complete Human Proteome. *Cell*. 2016;166(3):766–78.
196. Vapnik V. No Title Support Vector Machine. In: Springer-Verlag, editor. *The Nature of Statistical Learning*. 1995.
197. Fisher B, Ravdin RG, Ausman RK, Slack NH, Moore GE, Noer RJ. Surgical adjuvant chemotherapy in cancer of the breast: results of a decade of cooperative investigation. *Ann Surg*. 1968;168(3):337–56.
198. Theocharis AD, Skandalis SS, Neill T, Mulhaupt HAB, Hubo M, Frey H, et al. Insights into the key roles of proteoglycans in breast cancer biology and translational medicine. *Biochimica et Biophysica Acta - Reviews on Cancer*. 2015;1855(2):276–300.
199. Troup S, Njue C, Kliewer E V, Parisien M, Roskelley C, Chakravarti S, et al. Reduced expression of the small leucine-rich proteoglycans, lumican, and decorin is associated with poor outcome in node-negative invasive breast cancer. *Clin Cancer Res*. 2003;9(1 I):207–14.
200. Cawthorn TR, Moreno JC, Dharsee M, Tran-Thanh D, Ackloo S, Zhu PH, et al. Proteomic analyses reveal high expression of decorin and endoplasmic (HSP90B1) are associated with breast cancer metastasis and decreased survival. Aziz SA, editor. *PLoS One*. 2012;7(2):e30992.
201. Iozzo R V., Schaefer L. Proteoglycan form and function: A comprehensive nomenclature of proteoglycans *Matrix Biology*. 2015;42:11–55.
202. Petretto E, Sarwar R, Grieve I, Lu H, Kumaran MK, Muckett PJ, et al. Integrated genomic approaches implicate osteoglycin (Ogn) in the regulation of left ventricular mass. *Nat Genet*. 2008;40(5):546–52.
203. Xing X, Gu X, Ma T. Knockdown of biglycan expression by RNA interference inhibits the proliferation and invasion of, and induces apoptosis in, the HCT116 colon cancer cell line. *Mol Med Rep*. 2015;12(5):7538–44.
204. Weber CK, Sommer G, Michl P, Fensterer H, Weimer M, Gansauge F, et al.

- Biglycan is overexpressed in pancreatic cancer and induces G1-arrest in pancreatic cancer cell lines. *Gastroenterology*. 2001;121(3):657–67.
205. Sun H, Wang X, Zhang Y, Che X, Liu Z, Zhang L, et al. Biglycan enhances the ability of migration and invasion in endometrial cancer. *Arch Gynecol Obstet*. 2016;293(2):429–38.
206. Yamamoto K, Ohga N, Hida Y, Maishi N, Kawamoto T, Kitayama K, et al. Biglycan is a specific marker and an autocrine angiogenic factor of tumour endothelial cells. *Br J Cancer*. 2012;106(6):1214–23.
207. Leygue E, Snell L, Dotzlaw H, Hole K, Hiller-Hitchcock T, Roughley PJ, et al. Expression of lumican in human breast carcinoma. *Cancer Res*. 1998;58(7):1348–52.
208. Oldberg A, Kalamajski S, Salnikov A V, Stuhr L, Morgelin M, Reed RK, et al. Collagen-binding proteoglycan fibromodulin can determine stroma matrix structure and fluid balance in experimental carcinoma. *Proc Natl Acad Sci*. 2007;104(35):13966–71.
209. Jian J, Zheng Z, Zhang K, Rackohn TM, Hsu C, Levin A, et al. Fibromodulin promoted in vitro and in vivo angiogenesis. *Biochem Biophys Res Commun*. 2013;436(3):530–5.
210. Maquoi E, Assent D, Detilleux J, Pequeux C, Foidart J-M, Noël A. MT1-MMP protects breast carcinoma cells against type I collagen-induced apoptosis. *Oncogene*. 2012;31(4):480–93.
211. Kim SH, Lee HY, Jung SP, Kim S, Lee JE, Nam SJ, et al. Role of secreted type I collagen derived from stromal cells in two breast cancer cell lines. *Oncol Lett*. 2014;8(2):507–12.
212. Chen P, Cescon M, Bonaldo P. Collagen VI in cancer and its biological mechanisms. *Trends Mol Med*. 2013;19(7):410–7.
213. Matsushima-Nishiwaki R, Toyoda H, Nagasawa T, Yasuda E, Chiba N, Okuda S, et al. Phosphorylated Heat Shock Protein 20 (HSPB6) Regulates Transforming Growth Factor- α -Induced Migration and Invasion of Hepatocellular Carcinoma Cells. Guan X-Y, editor. *PLoS One*. 2016;11(4):e0151907.
214. Nahomi RB, Dimauro MA, Wang B, Nagaraj RH. Identification of peptides in human Hsp20 and Hsp27 that possess molecular chaperone and anti-apoptotic activities. *Biochem J*. 2015;465(1):115–25.
215. White MY, Murray CI, Kane LA, Fu Q, Stanley BA, Van Eyk JE, et al.

- Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. *Nat Methods*. 2013;10(1):28–34.
216. Zagouri F, Bournakis E, Koutsoukos K, Papadimitriou CA. Heat shock protein 90 (Hsp90) expression and breast cancer. *Pharmaceuticals*. 2012;5(9):1008–20.
217. Muller P, Ruckova E, Halada P, Coates PJ, Hrstka R, Lane DP, et al. C-terminal phosphorylation of Hsp70 and Hsp90 regulates alternate binding to co-chaperones CHIP and HOP to determine cellular protein folding/degradation balances. *Oncogene*. 2013;32(25):3101–10.
218. Cheng Q, Chang JT, Geradts J, Neckers LM, Haystead T, Spector NL, et al. Amplification and high-level expression of heat shock protein 90 marks aggressive phenotypes of human epidermal growth factor receptor 2 negative breast cancer. *Breast Cancer Res*. 2012;14(2):R62.
219. Pick E, Kluger Y, Giltane JM, Moeder C, Camp RL, Rimm DL, et al. High HSP90 expression is associated with decreased survival in breast cancer. *Cancer Res*. 2007;67(7):2932–7.
220. Goto R, Nakamura Y, Takami T, Sanke T, Tozuka Z. Quantitative LC-MS/MS analysis of proteins involved in metastasis of breast cancer. Coleman WB, editor. *PLoS One*. 2015;10(7):e0130760.
221. www.uniprot.org. www.uniprot.gov.

8. ANEXOS

8. ANEXOS

Anexo 1. Dato crudo de la medición por SRM (se proporciona en formato electrónico).

Anexo 2. Publicación relacionada con el trabajo de tesis.