



# PROBIOL

Programa de Posgrado en Biología

*Trabajo de Tesis para optar al grado de*

**“Doctor en Ciencias Biológicas”**

*“Estudio comprensivo de las Proteínas  
de Golpe de Calor (HSPs) en cáncer de  
mama”*

**Med. Martin E. Guerrero Gimenez**

**Director: Dr. Felipe C. M. Zoppino**

**Codirector: Dr. Daniel R. Ciocca**



**UNCUYO**  
UNIVERSIDAD  
NACIONAL DE CUYO



Facultad de  
Ciencias Médicas

Laboratorio de Oncología  
IMBECU - Universidad Nacional de Cuyo  
CCT-Mendoza, CONICET, Argentina

**Noviembre 2020**

## DECLARACIÓN

Declaro aquí que la tesis titulada “Estudio comprensivo de las Proteínas de Golpe de Calor (HSPs) en cáncer de mama” presentada por mí, para la obtención del título de *Doctor en Ciencias Biológicas* otorgado por el programa de Posgrado en Biología de la Universidad Nacional de Cuyo, es un registro del trabajo realizado por mí bajo la supervisión del Dr. Felipe C.M. Zoppino y del Dr. Daniel R. Ciocca ambos investigadores del laboratorio de Oncología del Instituto de Medicina y Biología Experimental de Cuyo (IMBECU) del Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina (CONICET).

Declaro además que el trabajo reportado en esta tesis no ha sido presentado y no se presentará, ni en parte ni en su totalidad, para la concesión de ningún otro título en este instituto o en cualquier otro instituto o universidad.

**Martin E. Guerrero Gimenez**

15 de Noviembre de 2020

Mendoza, Argentina

## RESUMEN

El cáncer de mama (CM) es una enfermedad clínica y molecularmente heterogénea que causa la muerte de más de 500.000 mujeres anualmente a nivel mundial. Las Proteínas de Golpe de Calor (HSPs) son un conjunto de chaperonas moleculares que asisten en el correcto plegado y control de calidad de múltiples proteínas. En el cáncer, algunas HSPs han sido descritas como facilitadoras de la tumorigénesis, por lo que han sido propuestas como posibles objetivos terapéuticos en la enfermedad.

**Hipótesis:** Las HSPs presentan perfiles de expresión asociados a diferentes estados biológicos por lo que pueden ser utilizadas como marcadores pronósticos y/o dianas moleculares para tratar la enfermedad.

**Objetivo:** Identificar el estado de expresión y metilación de las HSPs en el CM para generar marcadores pronósticos y proponer micro ARNs (miRNAs) con la capacidad de interferir la actividad de las HSPs de mayor relevancia clínica.

**Métodos:** Utilizamos diferentes bases de datos ómicas para describir los perfiles transcriptómicos y metilómicos de las HSPs en CM. Se evaluaron los genes HSPs como marcadores pronósticos y se utilizaron estrategias de aprendizaje estadístico para detectar patrones de expresión de HSPs distintivos. Para perfeccionar la detección de patrones transcriptómicos desarrollamos “Galgo”, un algoritmo de software libre para detectar firmas génicas cohesivas y clínicamente significativas, con uso extensivo a múltiples tipos de cáncer. A través del minado de bases de datos, buscamos miRNAs con la capacidad de interferir múltiples HSPs simultáneamente. Finalmente, seleccionamos a hsa-miR-570, el cual fue evaluado mediante técnicas de biología molecular en dos líneas de CM de alta agresividad.

**Resultado:** Describimos diferentes patrones de expresión de HSPs con valor pronóstico. Desarrollamos un algoritmo capaz de automatizar la búsqueda de firmas génicas, el cual nos permitió refinar la firma génica de HSPs detectada inicialmente. Finalmente describimos a hsa-miR-570 como una molécula promisoría para interferir la expresión de las HSPs relacionadas a un peor pronóstico en el CM.

**Keywords:** *HSPs, Cáncer de Mama, Transcriptómica, Firmas génicas, miRNA.*

## AGRADECIMIENTOS

Quisiera dedicarles este trabajo a mis padres, **Ana y Mario**. Ellos, que significan todo para mí, además de acompañarme desde mis primeros pasos, me regalaron alas para ser libre y se quedaron conmigo hasta que pude aprender a volar, cuidándose siempre, listos para agarrarme si me caigo, con el corazón dispuesto cuando lo necesito, siempre sonriendo y amando, disfrutando el viaje al máximo, permanentemente mirando para adelante, en busca de nuevas aventuras y más vida por vivir. Los amo con mi corazón.

A mis maestros, **Martín y Daniel**, por creer en mí, por enseñarme de forma generosa, por estar a disposición en el trabajo y en la vida, por su apoyo incondicional siempre. He encontrado en ellos más que maestros, amigos para toda la vida y voy a estar agradecido siempre de haberlos cruzado en el camino.

A mi compañera **Constanza**, que me acompañó y apoyó siempre, en todos los sueños que perseguí, en todas las malas que también pasé. Por su amor, su hermosa sonrisa, su enorme paciencia y comprensión, no podría haber hecho este camino sin su preciosa compañía y su amor incondicional.

A mis hermanas, **María Sol y María Luz**, que son una fuente inagotable de amor, sonrisas y felicidad, siempre me hacen sentir bien cuando lo necesito y son un apoyo y una referencia para mí.

A mi sobrino **Ramiro**, que con su presencia ilumina todo de alegría y calidez, por que me recuerda siempre el niño que tenemos dentro y cuando jugamos puedo volver a aquellos tiempos felices.

A mis amigos del alma, mis hermanos, **Gabriel, Andrés, Rodrigo, Florencia, Milagros, Jorge y Pedro**, por todas las charlas interminables, la música, las discusiones de política, las cenas, los viajes, las idas y vueltas, por que pase lo que pase, sea lo que sea, siempre están.

A mi amigo y compañero, **Juan Manuel**, por que ceba los mejores mates de todos y por que con él realmente valía cada segundo ir a trabajar, sin dudas una de las mejores cosas que me llevo del doctorado. En ese sentido no se quedan atrás mis amigas y compañeras de trabajo **Gise P. y Mage**, con quienes compartí miles de buenos momentos y me levantaron siempre que me frustraba o andaba de mal humor.

A mis compañeros de aventuras científicas **Carlos “Harpo” Catania** y **Benjamin Lang**, por apoyarme en las ideas locas y aguantarme los e-mails infinitos y los mensajes en cualquier momento y a cualquier hora, por amar lo que hacen como yo, por enseñarme y por confiar en mi.

A mis tutores **Gustavo Stolovitzky**, **Rubén Carón** y **Laura Vargas**, por darme su apoyo, guía invaluable y preciado tiempo, realmente son en mi un faro en el cual me inspiro, como científicos, pero sobre todo como personas, son inmensos.

Mil gracias a mis compañeros y amigos de trabajo, que me hicieron sentir siempre en casa y con quienes compartimos tantas cosas **Mariel**, **Niubys**, **Eve**, **Gise C.** y **Darío** y a todo el equipo de **IMBECU** y del área de **Química Biológica** que son una gran familia para mi.

Finalmente quisiera agradecer a las personas que integran **PROBIOL**, por estar siempre a disposición, resolver cada uno de mis problemas y tenerme tanta paciencia.

A todos, inmensamente **GRACIAS!**

Lugar: Mendoza

Fecha: 01/12/2020

**MARTIN E. GUERRERO GIMENEZ**

# ÍNDICE

<b>RESUMEN</b> . . . . .	i
<b>AGRADECIMIENTOS</b> . . . . .	ii
<b>LISTA DE FIGURAS</b> . . . . .	ix
<b>LISTA DE TABLAS</b> . . . . .	xxxiv
<b>LISTA DE ABREVIACIONES</b> . . . . .	xxxvi
<b>1 Introducción</b> . . . . .	<b>1</b>
1.1 Cáncer, generalidades y epidemiología . . . . .	1
1.2 La biología del cáncer y las Proteínas de Golpe de Calor . . . . .	2
1.2.1 Las Proteínas de Golpe de Calor . . . . .	4
1.2.2 Autosuficiencia en las señales de crecimiento . . . . .	14
1.2.3 Insensibilidad a las señales inhibitoras del crecimiento . . . . .	16
1.2.4 Resistencia a la muerte celular . . . . .	19
1.2.5 Potencial de replicación ilimitado . . . . .	21
1.2.6 Angiogénesis sostenida . . . . .	22
1.2.7 Invasión de tejidos y metástasis . . . . .	23
1.2.8 Inestabilidad genómica . . . . .	26
1.2.9 Inflamación promotora de tumores . . . . .	29
1.2.10 Reprogramación del metabolismo energético . . . . .	31
1.2.11 Evasión de la destrucción inmunológica . . . . .	33
1.3 Cáncer de mama . . . . .	35
1.3.1 Evaluación de la expresión génica . . . . .	38
1.3.2 Los subtipos intrínsecos del cáncer de mama . . . . .	40
1.3.3 Subtipo intrínseco luminal . . . . .	43
1.3.4 Subtipo intrínseco de tipo basal . . . . .	45
1.3.5 Subtipo intrínseco enriquecido con HER2 . . . . .	46

1.4	Perspectivas terapéuticas . . . . .	48
1.4.1	Los miRNAs como moléculas promisorias para la regulación de la red de chaperonas . . . . .	51
<b>2</b>	<b>Descripción de la tesis</b>	<b>55</b>
2.1	Hipótesis . . . . .	55
2.2	Objetivo general . . . . .	55
2.3	Objetivos específicos . . . . .	55
2.4	Metodología general . . . . .	56
<b>3</b>	<b>Expresión y metilación de las HSPs</b>	<b>60</b>
3.1	Introducción . . . . .	60
3.2	Métodos . . . . .	61
3.2.1	Análisis de datos . . . . .	61
3.2.2	Clasificación de subtipos intrínsecos . . . . .	62
3.2.3	Expresión diferencial de genes en muestras de cáncer de mama . . . . .	65
3.2.4	Estado de metilación del ADN de las HSPs . . . . .	67
3.2.5	Análisis de sobrevida . . . . .	68
3.3	Resultados . . . . .	69
3.3.1	Evaluación de los niveles de expresión del ARN de las HSPs en el cáncer de mama . . . . .	69
3.3.2	Cambios en la expresión de las diferentes subfamilias de HSPs en cáncer de mama . . . . .	75
3.3.3	Análisis de metilación de las HSPs . . . . .	80
3.3.4	Impacto de la expresión de las HSPs en el pronóstico de los pacientes	85
3.4	Discusión . . . . .	86
<b>4</b>	<b>Perfiles de expresión de HSPs distintivos se asocian a diferente pronóstico en cáncer de mama</b>	<b>93</b>
4.1	Introducción . . . . .	93
4.2	Métodos . . . . .	94
4.2.1	Detección de perfiles de expresión de HSPs y construcción de un clasificador para subtipos HSPs . . . . .	94

4.2.2	Validación de los perfiles transcriptómicos y generación de una firma génica de HSPs . . . . .	94
4.2.3	Evaluación clínica de los perfiles de expresión de las HSPs . . . . .	96
4.3	Resultados . . . . .	96
4.3.1	Bloques de expresión de las HSPs en el cáncer de mama . . . . .	96
4.3.2	Subtipos transcriptómicos de HSPs . . . . .	99
4.3.3	Valor pronóstico de los HSP-Clusts en cáncer de mama . . . . .	102
4.3.4	Expresión de las HSPs de acuerdo a los HSP-Clusts . . . . .	103
4.3.5	Validación de los perfiles de expresión HSP-Clusts en la cohorte de METABRIC . . . . .	104
4.4	Discusión . . . . .	107
<b>5</b>	<b>Algoritmo Galgo</b>	<b>110</b>
5.1	Introducción . . . . .	110
5.2	Métodos . . . . .	115
5.2.1	Adquisición y pre-procesamiento de datos . . . . .	115
5.2.2	Implementación de Galgo . . . . .	116
5.2.3	Métricas de comparación . . . . .	118
5.2.4	Identificación de subtipos moleculares de cáncer utilizando Galgo	119
5.2.5	Clasificación de pacientes mediante firmas previamente publicadas	119
5.2.6	Distribución de los valores de C.Index obtenidos de firmas génicas aleatorias en comparación con las firmas de Galgo . . . . .	120
5.2.7	Comparación de Galgo con otras firmas génicas . . . . .	121
5.3	Resultados . . . . .	123
5.3.1	Identificación de subtipos moleculares de cáncer que predicen la supervivencia del paciente . . . . .	123
5.3.2	Comparación de Galgo con firmas de expresión genética aleatorias	133
5.3.3	Estabilidad y robustez de Galgo . . . . .	134
5.4	Discusión . . . . .	136
<b>6</b>	<b>Refinando los subtipos HSP con Galgo</b>	<b>142</b>
6.1	Introducción . . . . .	142
6.2	Métodos . . . . .	143

6.2.1	Adquisición y pre-procesamiento de datos . . . . .	143
6.2.2	Detección de firmas HSP-Galgo . . . . .	144
6.2.3	Comparación de la firma HSP-Galgo con otras firmas relevantes en CM . . . . .	145
6.2.4	Evaluación de la firma HSP-Galgo definitiva . . . . .	145
6.2.5	Análisis de las vías biológicas más relevantes en relación a los subtipos HSP-Galgo . . . . .	146
6.3	Resultados . . . . .	146
6.3.1	La firma HSP-Clust puede ser refinada en una firma con mayor poder predictivo . . . . .	146
6.3.2	HSP-Galgo en comparación con HSP-Clust y PAM50 . . . . .	149
6.3.3	Evaluación de las vías moleculares asociadas con los subtipos HSP-Galgo . . . . .	151
6.4	Discusión . . . . .	152
<b>7</b>	<b>Los microRNAs como posible moléculas colaboradoras en la inhibición de las HSPs en el cáncer de mama</b>	<b>156</b>
7.1	Introducción . . . . .	156
7.2	Métodos . . . . .	161
7.2.1	Búsqueda de subtipos HER2 con alteraciones en la expresión de genes de la matriz extracelular y/o proteínas asociadas al chaperoma	161
7.2.2	Evaluación de las alteraciones asociadas a los diferentes subtipos HER2 . . . . .	162
7.2.3	Búsqueda <i>In Silico</i> de posibles miRNAs para la modulación de las HSPs más relevantes en cáncer de mama . . . . .	162
7.2.4	Evaluación de la expresión, metilación y alteración en el número de copias de los miRNAs candidatos en los subtipos moleculares intrínsecos de Cáncer de Mama . . . . .	163
7.2.5	Evaluación biológica de los miRNAs en líneas celulares estable- cidas de cáncer de mama . . . . .	164
7.3	Resultados . . . . .	166
7.3.1	El subtipo tumoral HER2 se puede dividir en dos subtipos HER2- ECM con diferente pronóstico . . . . .	166

7.3.2	Los miRNAs 224, 522 y 570 son candidatos a regular la dupla HSPA1A y BAG3 en el cáncer de mama . . . . .	168
7.3.3	Evaluación biológica de miR-570 . . . . .	174
7.4	Discusión . . . . .	178
<b>8</b>	<b>Observaciones finales y perspectivas futuras</b>	<b>181</b>
	<b>REFERENCIAS</b> . . . . .	<b>185</b>
	<b>PUBLICACIONES</b> . . . . .	<b>214</b>

## **APÉNDICES**

<b>Apendice A</b>	<b>Cohortes evaluadas en el desarrollo de Galgo</b>	<b>217</b>
<b>Apendice B</b>	<b>Comparación de firmas Galgo con otras firmas génicas</b>	<b>224</b>
<b>Apendice C</b>	<b>Genes del complejo de silenciamiento inducido por ARNs (RISC) potencialmente inhibidos por miRNA-570</b>	<b>226</b>
<b>Apendice D</b>	<b>Figuras suplementarias</b>	<b>227</b>

# LISTA DE FIGURAS

- 1.1 **Los sellos del cáncer.** El esquema ilustra las condiciones necesarias para el desarrollo de la enfermedad maligna. Estos rasgos característicos pueden ser adquiridos en diferentes etapas en el desarrollo del cáncer, a través de distintivos mecanismos de acuerdo al tipo de cáncer. Modificado de *Hanahan and Weinberg* (2011). . . . . 4
- 1.2 **Superficie de energía libre en el plegamiento proteico.** Representación de la superficie de energía libre que las proteínas exploran a medida que avanzan hacia el estado nativo (verde) formando contactos intramoleculares. La dificultad del paisaje de energía libre da como resultado la acumulación de conformaciones atrapadas cinéticamente que necesitan atravesar barreras de energía libre para alcanzar un camino favorable. *In vivo*, estos pasos se aceleran mediante las chaperonas. Cuando varias moléculas se pliegan simultáneamente en el mismo compartimento, la superficie de plegado de energía libre puede solaparse con la de la agregación intermolecular, lo que da como resultado la formación de agregados amorfos, oligómeros tóxicos o fibrillas amiloides ordenadas (rojo). La agregación fibrilar ocurre típicamente por polimerización dependiente de nucleación. Puede iniciarse a partir de intermedios poblados durante el plegamiento *de novo* o después de la desestabilización del estado nativo (estados parcialmente plegados) y normalmente se evita mediante chaperonas moleculares. Modificado de *Hartl et al.* (2011). . . 6

- 1.3 **Red funcional de la familia HSPs necesaria para el mantenimiento del control de calidad de las proteínas celulares.** En la figura se pueden observar los ciclos de plegamiento de las diferentes familias de HSPs y sus sustratos, su funcionamiento mediante 'relevos' y la dependencia de algunos complejos HSPs de ATP y otros co-factores (Como factores intercambiadores de nucleótidos (NEF), Hop y prefoldinas, por mencionar algunas). Hsp70 (HSPA), Hsp110 (HSPH), sHsp (HSPB), Hsp90 (HSPC), Hsp40 (DNAJ), Hsp60 (CHAP I), TRiC/CCT (CHAP II). Extraído de <http://pds-lab.biochem.iisc.ernet.in/hspir/chaperone.php> (Ratheesh Kumar et al., 2012) . . . . . 8
- 1.4 **Las células tumorales pueden modular el microambiente tumoral secretando HSPs.** Se ha identificado que las células de carcinoma derivadas de varios tejidos secretan HSPs en el espacio extracelular en forma soluble o dentro de exosomas. Las HSPs secretadas de las células tumorales pueden modular la biología y las funciones de otras células en el microambiente tumoral. Específicamente, se ha demostrado que las HSPs secretadas modulan las actividades de las células supresoras mieloides, células dendríticas (CD), macrófagos asociados a tumores (MAT), células asesinas naturales (NK), otras células de carcinoma vecinas y células endoteliales. Las HSPs extracelulares pueden promover procesos tumorigénicos (enumerados en texto rojo). Alternativamente, las HSPs extracelulares también pueden promover la inmunidad tumoral (enumeradas en texto verde) al estabilizar los antígenos tumorales, estimular a los MAT para que secreten citocinas inflamatorias y mediar el procesamiento de antígenos por las células presentadoras de antígenos. Modificado de Lang et al. (2019), creado con <https://biorender.com/>. . . . . 31
- 1.5 **Algoritmo terapéutico en el cáncer de mama.** Esquema de decisiones terapéuticas de acuerdo a las preferencias de la paciente, el estadio tumoral y al fenotipo molecular del tumor. Fenotipo agresivo: Tumores triple negativos o HER2. Modificado de Cardoso et al. (2019). . . . . 37

1.6	<b>Elección del tratamiento sistémico (neo)-adyuvante de acuerdo a la expresión de marcador tumorales y los subtipos intrínsecos de CM.</b> Modificado de <i>Cardoso et al.</i> (2019). . . . .	38
1.7	<b>Proceso de cuantificación de ARNm mediante técnica de RNA-Seq.</b> Para cuantificar los diferentes transcritos de ARNm que posee una muestra primero se procede a diseccionar y homogeneizar el tejido a analizar. Posteriormente se extrae el ARN total y se aísla el ARNm mediante la selección de transcritos con extremos Poly-A. Los ARNm son fragmentados y retro-transcritos a ADNc, los cuales son acoplados a adaptadores específicos que permiten su fijado en las celdas de flujo (' <i>flowcell</i> ') del secuenciador. Una vez fijados los fragmentos en la <i>flowcell</i> , se procede a la formación de conglomerados de transcritos mediante el proceso de amplificación en puente. Posteriormente, una vez formados los conglomerados de transcritos, se inicia la secuenciación por síntesis, donde nucleótidos fluorescentes son incorporados de forma secuencial a las hebras de ADN fijadas en las celdas, lo cual produce una excitación lumínica que es captada por el secuenciador, que luego traduce dicha información en una secuencia de nucleótidos. Una vez obtenida la secuencia de cada uno de los fragmentos analizados, estos se alinean a un genoma de referencia y se cuentan la cantidad de copias obtenidas para cada uno de los genes del genoma humano. Modificado de <i>Momand et al.</i> (2016) . . . . .	40
1.8	<b>Imagen mediante microscopía óptica de muestras representativas de los diferentes subtipos de tumores mamarios.</b> Imagen teñida con HandE representativa de los tumores luminal A [A], luminal B [G], HER2 [M] y Basal [S]. Tinción de RE en tumores luminal A [B], luminal B [H], HER2 [N] y Basal [T]. Tinción RP en tumores luminal A [C], luminal B [I], HER2 [O] y Basal [U]. Tinción de HER2 en tumores luminal A [D], luminal B [J], HER2 [P] y Basal [V]. Tinción Ki67 en tumores luminal A [E], luminal B [K], HER2 [Q] y Basal [W]. Tinción de CK5 en tumores luminal A [F], luminal B [L], HER2 [R] y Basal [X]. Modificado de <i>Dias et al.</i> (2017). . . . .	43

1.9	<b>Descripción general de los procesos autónomos de células cancerosas mediados por HSPs que promueven la tumorigénesis.</b> El potencial transformador de los eventos autónomos de las células cancerosas, como las mutaciones oncogénicas, suele depender del apoyo de Hsp27, Hsp70 y Hsp90 intracelulares. Las actividades de los transductores de señales mitogénicas intracelulares clave también dependen de las HSPs intracelulares para la activación sostenida de la vía y, por lo tanto, las HSPs median cómo las células cancerosas responden a las señales de crecimiento que emanan del microambiente tumoral. Tales eventos de transformación requieren la capacidad de las células tumorales para superar la senescencia y perpetuar la inmortalidad replicativa. También representamos aquí otros procesos tumorigénicos clave respaldados por HSPs intracelulares, incluida la inhibición de la muerte celular programada, el metabolismo celular alterado para favorecer la glucólisis, la estimulación de la angiogénesis por células cancerosas y las propiedades invasivas y metastásicas. Modificado de <i>Lang et al.</i> (2019), creado con <a href="https://biorender.com/">https://biorender.com/</a> .....	49
1.10	<b>Resumen de los principales miRNAs e inhibidores farmacológicos de las HSPs en el cáncer.</b> Modificado de <i>Ban et al.</i> (2019) . . . . .	51
1.11	<b>Descripción general de la biogénesis y mecanismo de acción de los miRNAs.</b> La transcripción nuclear de un miRNA primario (pri-miRNA) por la ARN polimerasa II es seguida por su procesamiento mediante DGCR8 en una horquilla precursora (pre-miARN). Después de la exportación nuclear (EXPO5), el procesamiento del pre-miARN por Dicer ocurre en el citoplasma. Posteriormente, una de las hebras del dúplex maduro se carga en un complejo proteico de silenciamiento inducido por ARN (RISC), en donde el miRNA guía al complejo hacia su ARN diana (azul) de acuerdo al emparejamiento de bases de la secuencia semilla del miRNA en los nucleótidos 2-7 del extremo 5'(en rojo oscuro). Modificado de <i>Alberti and Cochella</i> (2017). . . . .	53

2.1	<b>Diagrama de flujo del trabajo realizado a lo largo de la tesis.</b> Si bien todos los capítulos se encuentran interconectados, para facilitar la lectura y la comprensión del esquema se separaron los diferentes diagramas en recuadros de acuerdo a las diferentes secciones del manuscrito	59
3.1	<b>Control de calidad de clasificación PAM50 de las muestras de TCGA I.</b> Cada punto en los gráficos representa una sola muestra. A) Análisis de componentes principales de los datos de entrenamiento ( <i>Bernard et al., 2009</i> ) y prueba 'The Cancer Genome Atlas' (TCGA). Se puede observar que tanto los datos de entrenamiento como los de prueba TCGA se distribuyen de forma similar. También es interesante observar la superposición entre las asignaciones de subtipos. B) Grado de correlaciones de las muestras con cada uno de los centroides para cada subtipo. En cada gráfico los muestras están divididas de acuerdo al subtipo asignado. También se muestra índice estimado de proliferación para cada muestra. . . . .	63
3.2	<b>Control de calidad de la clasificación PAM50 de las muestras de TCGA II.</b> Agrupación jerárquica no supervisada de muestras de acuerdo con la expresión del conjunto de genes PAM50. Lo que se puede observar es que hay una alta coherencia entre el subtipo asignado a cada muestra por el algoritmo PAM50 y la composición del grupo determinada por la técnica de agrupamiento no supervisada. . . . .	64

3.3	<b>Consistencia de los resultados entre los métodos EdgeR y DESeq2.</b>	
	A) Análisis de correlación entre el nivel de cambio (FC) estimado por ambos métodos. La figura muestra una estrecha tendencia lineal entre las estimaciones de los cambios según EdgeR y DESeq2. Los genes estimados con cambios significativos para ambos métodos están representados en círculos amarillos, en verde y rojo son genes expresados diferencialmente de forma significativa de acuerdo a uno de los dos métodos y en blanco, genes sin cambios significativos por ambas técnicas. B) Análisis de Bland-Altman que compara el FC estimado mediante ambos métodos (eje x) y la diferencia entre ellos (eje y). Este gráfico permite la identificación de diferencias sistemáticas entre métodos y posibles valores atípicos. Cada círculo representa un gen HSP y sus colores el subtipo para el que se calculó el FC. La línea de puntos azul representa la diferencia media entre ambas técnicas (0,02) y la línea de puntos azul claro representa los límites superior (0,88) e inferior (- 0,84) del intervalo de confianza del 95% de las diferencias. . . . .	66
3.4	<b>Comparación de la expresión de las HSPs en relación al resto de los genes en el CM.</b> A) La expresión media de cada gen del total de las muestras de cáncer se calculó y se ordenaron en orden decreciente. Los genes de HSPs se localizaron con una X roja. Se puede observar que seis genes HSPs están por encima de la línea naranja de los 100 genes más expresados en CM. B) Los gráficos muestran la distribución de expresión de ARN de los genes HSPs en la cohorte. . . . .	70

- 3.5 **Expresión diferencial de genes en CM.** Análisis de expresión diferencial de genes realizado mediante el método EdgeR. En el eje x se representa el nivel de cambio de la expresión en logaritmo con base 2 ( $Log_2FC$ ) entre el tejido tumoral y el tejido normal, mientras que en el eje y se muestra el  $-\log_{10}$  de FDR (los valores más altos implican FDR más pequeño, por lo tanto mayor significancia estadística). Los genes HSPs con  $Log_2FC > 1$  y  $FDR < 0,05$  se indican como círculos rojos. Los símbolos verdes en la parte superior de los gráficos indican genes con FDR muy pequeño ( $FDR < 5e - 324$ ). Los genes con  $Log_2FC$  significativos que no son HSPs están coloreados azul claro. . . . . 72
- 3.6 **HSPs desreguladas específicas de subtipo.** Diagramas de Venn que resumen la expresión diferencial de HSPs de forma específica o compartida en los diferentes subtipos intrínsecos de CM. Solo se muestran las HSPs que mostraron un  $Log_2FC > 1$  y valores  $FDR < 0,05$  por el método EdgeR y concordante con el método DESeq2. El grupo 'normal-like' se descartó por el bajo número de casos. A) Genes HSPs regulados negativamente. B) Genes HSPs regulados positivamente . . . 74
- 3.7 **Cambio de expresión de las HSPs de acuerdo a sus subfamilias y al subtipo molecular de CM.** En la figura, el diámetro de los círculos muestra el  $Log_2FC$  de las HSPs evaluado por el método EdgeR. Los círculos en verde muestran genes regulados negativamente y los rojos representan genes regulados positivamente. La opacidad del círculo está relacionada con los valores de FDR, los círculos con  $FDR > 0.05$  son transparentes y, por lo tanto, no se muestran. La figura hace énfasis en la magnitud de los cambios de expresión de las HSPs independientemente de cualquier umbral de cambio considerado . . . . . 76
- 3.8 **Tendencias de cambios en la expresión génica de las diferentes subfamilias de HSP en los subtipos moleculares de CM.** Los diagramas de caja representa los rangos de cambio ( $Log_2FC$ ) de la subfamilias de HSPs mediante el método EdgeR en los diferentes subtipos moleculares de cáncer de mama. . . . . 78

3.9	<b>Expresión diferencial de genes HSPs entre los diferentes subtipos tumorales de CM.</b> Se hizo una comparación de la expresión de las HSPs entre los diferentes subtipos PAM50 mediante un análisis EdgeR-ANOVA. Los genes HSPs se encuentran agrupados de acuerdo a sus familias correspondientes. Las líneas azules verticales representan la línea de base (subtipo de referencia), mientras que los puntos azul claro muestran el cambio en la expresión de acuerdo al subtipo de comparación. Los puntos rojos representan genes con cambios absolutos de $Log_2FC > 2$ . . . . .	79
3.10	<b>Metilación diferencial de regiones asociada a genes HSPs entre los diferentes subtipos tumorales de CM.</b> El gráfico de radar muestra las regiones CpG (Denominadas ' <i>bumps</i> ' e identificadas numéricamente entre paréntesis) de las HSPs con variaciones significativas de metilación en relación con el tejido normal (valor $p_{adj} < 0,05$ ) identificadas mediante el método <i>bumphunting</i> . La línea verde punteada representa el nivel de metilación de referencia del tejido mamario normal. Los picos hacia afuera representan cambios de metilación positivos (mayor metilación) en comparación con la misma región del tejido normal. Los picos hacia adentro representan desmetilación de la región estudiada. . .	81

- 3.11 **Estado de metilación de los genes HSPB2 / CRYAB en CM.** Esta y la figura siguiente (Figura 3.12) está compuesta por diferentes carriles de información integrada. En el carril “Beta”, los valores de metilación se muestran en el eje vertical (eje y), mientras que cada columna de puntos representa un sitio CpG determinado, ubicado en relación con su región cromosómica (mostrado en los carriles superiores). Cada punto representa una muestra y el valor de metilación medido para el sitio CpG correspondiente. En el carril de valores beta se añadió una curva de regresión suavizada al diagrama de dispersión para mostrar las diferencias de metilación entre los subtipos moleculares y el tejido normal. Los carriles a continuación muestran los *'bumps'* (protuberancias o jorobas) para cada subtipo de tumor, que representan los niveles de metilación en cada región en su conjunto. Obsérvese que a mayor altura de las *'bumps'*, mayor nivel de metilación en relación al nivel de referencia. En los dos últimos carriles se muestra un análisis de correlación entre la metilación y el estado de expresión. En el carril de “correlación” se muestra el valor de correlación de Pearson para cada sitio CpG sin tener en cuenta el subtipo molecular. Cada punto representa el impacto de un sitio CpG en la expresión génica (a mayores desviaciones de la línea horizontal mayor impacto en la expresión génica). En el último carril, un diagrama de dispersión muestra la relación entre los valores de metilación de un único sitio CpG (eje x) y el nivel de expresión de ARN del gen estudiado (eje y) en cada subtipo molecular. Observe que la metilación y los niveles de expresión tienen mayor asociación en las regiones donde se observaron las *'bumps'*. . . . . 84
- 3.12 **Estado de metilación del gen HSPB6 en CM.** La interpretación es equivalente a la figura 3.11. . . . . 85

- 4.1 **Agrupación jerárquica no supervisada de muestras de acuerdo con la expresión del conjunto de genes HSPs (I).** Mapa de calor de la expresión génica de las HSPs en pacientes de la cohortes TCGA de CM. En el presente mapa se representan los patrones de expresión de 89 genes HSPs (filas) en 1033 muestras tumorales (columnas) en el panel central (niveles de expresión bajos en azul y niveles de expresión altos en rojo). Se agregó una fila superior para indicar la clasificación PAM50 asignada a cada paciente y una columna a la izquierda que identifica la subfamilia a la que pertenece cada HSP. . . . . 97
- 4.2 **Agrupación jerárquica no supervisada de muestras de acuerdo con la expresión del conjunto de genes HSPs (II).** La presente figura tiene la misma interpretación que la figura 4.1, sin embargo se han identificado con recuadros y números romanos 6 diferentes bloques transcripcionales asociados a la expresión de las HSPs en CM. . . . . 99
- 4.3 **Análisis de 'Silhouette' del dendrograma del agrupamiento jerárquico basado en la expresión génica de las HSPs en CM (Figura4.1).** De acuerdo a la altura donde se corte el dendrograma la población se puede dividir en  $k$  grupos. El presente análisis calcula el coeficiente de *Silhouette* para cada una de las posibles divisiones. El coeficiente más alto corresponde al número óptimo de divisiones, en este caso  $k = 3$ . . . . 100

- 4.4 **Agrupación jerárquica no supervisada de muestras de acuerdo con la expresión del conjunto de genes HSPs (III).** La presente figura tiene la misma interpretación que la figura 4.1, sin embargo se ha agregado la identificación de los subtipos de HSP-Clust identificados mediante un algoritmo de agrupamiento jerárquico, los pacientes se agruparon en HSP-Clust I (rojo), HSP-Clust II (verde) y HSP-Clust III (naranja) (dendrograma superior). Se agregaron también filas para indicar: estado inmunohistoquímico de los receptores (ER, PR y HER2), tamaño del tumor ( $T > 2$  cm o  $T < 2$  cm), diseminación de los nódulos satélite (N positivo o N negativo) y clasificación PAM50. También se incluyen tres columnas que indican las subfamilias correspondientes de HSPs, los coeficientes univariados del modelo de regresión de Cox (el rosa representa los coeficientes positivos (mal pronóstico), mientras que el azul claro son los coeficientes negativos (buen pronóstico)) y sus valores correspondientes de FDR (los cuadros negros representan el valor de FDR para los coeficientes de Cox  $< 0,05$ ). . . . . 101
- 4.5 **Caracterización de los subtipos HSP.** A) Acuerdo entre la clasificación PAM50 y la de HSP-Clusts. El tamaño de las barras es proporcional al número de muestras en cada categoría. B) Supervivencia global de los subtipos de HSPs. Curvas de Kaplan-Meier correspondientes a HSP-Clust I, HSP-Clust II y HSP-Clust III. La significancia estadística se evaluó mediante la prueba Log-Rank. . . . . 102
- 4.6 **Modelo de Cox multivariable de HSP-Clusts.** A) Diagrama de 'Forest' que muestra la razón de riesgos de los subtipos de HSPs ajustados por factores de confusión (edad, estado de los ganglios, tamaño del tumor, estadio del tumor). Se representan la razón de riesgo (HR), el intervalo de confianza del 95% (95% CI) y los valores  $p$  correspondientes ( $p - val$ ). B) Mismo modelo de Cox, con la adición de subtipos PAM50 como covariables. . . . . 103

4.7	<b>Expresión diferencial de genes entre subtipos HSPs en la cohorte TCGA de CM.</b> Análisis tipo EdgeR-ANOVA que muestra las diferencias en la expresión de las HSPs en las diferentes comparaciones de los subtipos HSP-Clusts. Las líneas azules verticales representan el nivel de la línea de base del subtipo de referencia, mientras que los puntos azul claro muestran el cambio de veces de los genes HSPs en cada comparación por pares. Se representan puntos rojos para genes que tenían cambios absolutos en logaritmo en base 2 ( $\text{Log}_2$ ) > 2. . . . .	104
4.8	<b>Caracterización de los perfiles de expresión de HSPs.</b> A) Centroides prototípicos de los perfiles de expresión de los HSP-Clusts en la cohorte de TCGA y en los conjuntos de entrenamiento y de prueba de METABRIC. El color de las casillas representa la regulación hacia abajo (azul) o hacia arriba (rojo) del gen en el grupo correspondiente. La línea negra continua representa los valores medios de expresión de cada gen en el grupo en comparación con la media del mismo gen en todas las muestras (la línea punteada central). B) Concordancia entre los subtipos de PAM50 y HSP-Clusts para los conjuntos de entrenamiento y de prueba de METABRIC. El tamaño de las barras es proporcional al número de muestras en cada categoría. C) Supervivencia global de los subtipos HSP-Clusts para los conjuntos de entrenamiento y prueba para la cohorte de METABRIC. La significancia estadística se evaluó mediante la prueba de Log-Rank. . . . .	106
4.9	<b>Sobrevida global de los pacientes de acuerdo a los subtipos de PAM50 en las cohortes de TCGA y METABRIC.</b> . . . . .	108

- 5.1 **Descripción general del algoritmo Galgo.** La figura muestra los diferentes pasos dados por el algoritmo para encontrar firmas génicas robustas evolucionadas a lo largo de varias generaciones. (A) Para inicializar el algoritmo (primera generación), se establece un conjunto de genes aleatorios para crear una firma génica al azar. Cada 'cromosoma' está compuesto por una cadena de genes activados (que se incluirán en la firma) o desactivados (que no se incluirán), más un número del 2 al 10 que representa el número de particiones consideradas en cada firma. (B) Un conjunto de datos de entrenamiento se divide en  $k$  partes y se usa un esquema de validación cruzada para clasificar todas las muestras en el conjunto de datos usando cada firma generada. (C) Una vez que todas las muestras se clasifican usando cada firma, las mismas se evalúan teniendo en cuenta la consistencia del grupo ('*fitness 1*') y las diferencias de supervivencia entre los grupos ('*fitness 2*'). (D) Las soluciones se clasifican de acuerdo con sus funciones de aptitud, luego se seleccionan y se envían a los operadores de cruzamiento y mutación. (E) Todas las firmas de padres e hijos se agrupan, clasifican y se seleccionan las mejores soluciones para pasar a la siguiente generación. (F) Finalmente, las soluciones seleccionadas comienzan de nuevo el proceso desde B, a menos que se logren 500 generaciones. Si ese es el caso, el algoritmo se detiene y se informan las mejores soluciones. . . . . 114
- 5.2 **Frentes de Pareto después de la ejecución de Galgo en cada tipo de cáncer (CM, cáncer colorrectal (CCR), adenocarcinoma de pulmón (LUAD) y cáncer de ovario seroso de alto grado (HGSOC)).** Cada punto corresponde a una solución encontrada durante las 500 generaciones del algoritmo. La escala de azul claro a oscuro representa la generación donde se encontró cada solución. Las líneas negras muestran los frentes de Pareto estimados. Los puntos negros sobre la línea son soluciones que se encuentran en el frente de soluciones y los círculos rojos muestran las soluciones seleccionadas con el mejor rendimiento para cada caso. . . . . 124

5.3 **Evaluación de firmas para CM.** (A) Diagramas de bosque (*forest plot*) que representan las estimaciones (puntos) y los intervalos de confianza (líneas) del índice de concordancia (*C.Index*) para cada firma génica en cada cohorte estudiada. La línea discontinua representa el poder predictivo nulo ( $C.Index = 0.5$ ). Una estimación combinada de cada firma para todas las cohortes se muestra en un cuadro rojo. (B) Curvas de Kaplan-Meier que comparan las tendencias de supervivencia de los subgrupos de pacientes clasificados por el método PAM50 (línea de puntos) y por la firma Galgo-BRCA-4 (línea continua), tanto en el conjunto de entrenamiento como en el de pruebas combinado. Los valores *p* de la prueba Log-Rank para cada clasificador se representan en la parte inferior de cada gráfico. (C) Mapas de calor (*heatmaps*) que comparan los patrones de expresión génica de la firma Galgo-BRCA-4 en el conjunto de entrenamiento y el conjunto de pruebas combinado. En la parte superior de cada mapa de calor se muestran barras que representan la clasificación de cada muestra considerando las firmas PAM50 o Galgo-BRCA-4. (D) análisis de enriquecimiento de conjunto de genes o 'Gene Set Enrichment Analysis' (GSEA) de los subtipos Galgo-BRCA-4. Para cada subtipo, se representan los términos Ontología de genes (GO) alterados más significativos con respecto al resto de las muestras. Se utiliza una escala de color azul a rojo, donde el azul representa los términos de ontología regulados a la baja y el rojo representa los regulados al alza. . . . . 126

5.4 **Regularidad entre firmas Galgo de CM.** Se compararon las firmas de Galgo CM con mejor rendimiento (Galgo-BRCA-2, Galgo-BRCA-3, Galgo-BRCA-4, Galgo-BRCA-5 y Galgo-BRCA-7) para detectar la regularidad entre los diferentes genes involucrados en cada firma. A) Muestra un diagrama de Venn con el número compartido de genes en cada firma. B) Es una matriz de intersección que muestra los genes presentes (verdes) o ausentes (negros) en cada firma. Los bloques de colores de la derecha representan grupos de genes únicos para cada una de las firmas. . . . . 128

**5.5 Evaluación de firmas para CCR.** (A) Diagramas de bosque (*forest plot*) que representan las estimaciones (puntos) y los intervalos de confianza (líneas) del C.Index para cada firma génica en cada cohorte estudiada. La línea discontinua representa el poder predictivo nulo ( $C.Index = 0.5$ ). Una estimación combinada de cada firma para todas las cohortes se muestra en un cuadro rojo. (B) Curvas de Kaplan-Meier que comparan las tendencias de supervivencia de los subgrupos de pacientes clasificados por el método CMS (línea de puntos) y por la firma Galgo-CRC-2 (línea continua), tanto en el conjunto de entrenamiento como en el de pruebas combinado. Los valores  $p$  de la prueba Log-Rank para cada clasificador se representan en la parte inferior de cada gráfico. (C) Mapas de calor (*heatmaps*) que comparan los patrones de expresión génica de la firma Galgo-CRC-2 en el conjunto de entrenamiento y el conjunto de pruebas combinado. En la parte superior de cada mapa de calor se muestran barras que representan la clasificación de cada muestra considerando las firmas CMS o Galgo-CRC-2. (D) GSEA de los subtipos Galgo-CRC-2. Para cada subtipo, se representan los términos GO alterados más significativos con respecto al resto de las muestras. Se utiliza una escala de color azul a rojo, donde el azul representa los términos de ontología regulados a la baja y el rojo representa los regulados al alza. . 129

5.6 **Evaluación de firmas para LUAD.** (A) Diagramas de bosque (*forest plot*) que representan las estimaciones (puntos) y los intervalos de confianza (líneas) del *C.Index* para cada firma génica en cada cohorte estudiada. La línea discontinua representa el poder predictivo nulo (*C.Index* = 0.5). Una estimación combinada de cada firma para todas las cohortes se muestra en un cuadro rojo. (B) Curvas de Kaplan-Meier que comparan las tendencias de supervivencia de los subgrupos de pacientes clasificados por el método Wilkerson (línea de puntos) y por la firma Galgo-LUAD-4 (línea continua), tanto en el conjunto de entrenamiento como en el de pruebas combinado. Los valores *p* de la prueba Log-Rank para cada clasificador se representan en la parte inferior de cada gráfico. (C) Mapas de calor (*heatmaps*) que comparan los patrones de expresión génica de la firma Galgo-LUAD-4 en el conjunto de entrenamiento y el conjunto de pruebas combinado. En la parte superior de cada mapa de calor se muestran barras que representan la clasificación de cada muestra considerando las firmas Wilkerson o Galgo-LUAD-4. (D) GSEA de los subtipos Galgo-LUAD-4. Para cada subtipo, se representan los términos GO alterados más significativos con respecto al resto de las muestras. Se utiliza una escala de color azul a rojo, donde el azul representa los términos de ontología regulados a la baja y el rojo representa los regulados al alza. . . . . 130

5.7 **Evaluación de firmas para HGSOC.** (A) Diagramas de bosque (*forest plot*) que representan las estimaciones (puntos) y los intervalos de confianza (líneas) del C.Index para cada firma génica en cada cohorte estudiada. La línea discontinua representa el poder predictivo nulo ( $C.Index = 0.5$ ). Una estimación combinada de cada firma para todas las cohortes se muestra en un cuadro rojo. (B) Curvas de Kaplan-Meier que comparan las tendencias de supervivencia de los subgrupos de pacientes clasificados por el método ConsensusOv (línea de puntos) y por la firma Galgo-HGSOC-4 (línea continua), tanto en el conjunto de entrenamiento como en el de pruebas combinado. Los valores  $p$  de la prueba Log-Rank para cada clasificador se representan en la parte inferior de cada gráfico. (C) Mapas de calor (*heatmaps*) que comparan los patrones de expresión génica de la firma Galgo-HGSOC-4 en el conjunto de entrenamiento y el conjunto de pruebas combinado. En la parte superior de cada mapa de calor se muestran barras que representan la clasificación de cada muestra considerando las firmas ConsensusOv o Galgo-HGSOC-4. (D) GSEA de los subtipos Galgo-HGSOC-4. Para cada subtipo, se representan los términos GO alterados más significativos con respecto al resto de las muestras. Se utiliza una escala de color azul a rojo, donde el azul representa los términos de ontología regulados a la baja y el rojo representa los regulados al alza. . . . . 132

5.8 **Valores de C.Index de firmas Galgo vs. distribución de probabilidad del C.Index de firmas aleatorias.** Gráficos de violín que muestran la distribución de valores C.Index para 1000 firmas aleatorias de tamaño y número de particiones idénticas a las de las firmas Galgo. Cada color de los violines representa un tipo de cáncer diferente (violeta para HGSOC, azul claro para LUAD, verde para CCR (CRC) y rojo para CM (BRCA)). Los puntos rojos representan el C.Index obtenido por cada firma de Galgo. La línea negra representa un valor no predictivo ( $C.Index = 0.5$ ). A la derecha, los valores  $p$  de los C.Index de las firmas Galgo en relación a la distribución de C.Index de las firmas aleatorias. . . . . 134

5.9 **Robustez de Galgo en 30 evaluaciones.** El conjunto de datos METABRIC se dividió en un conjunto de entrenamiento y de prueba. Se aplicó Galgo en el conjunto de entrenamiento para encontrar un conjunto de soluciones no dominadas en 30 evaluaciones independientes, representadas en el eje  $x$ . A) Cada solución obtenida está representada por un punto de color con una clave de color que representa el número de particiones ( $k$ ) de la firma, mientras que el eje  $y$  representa el número de genes que constituyen cada firma. B) El diagrama de caja que representa la distribución de los valores de coeficiente de silueta (CS) de las soluciones no dominadas en el conjunto de entrenamiento y de prueba. El diagrama de caja en C) muestra la distribución de C.Index de las soluciones obtenidas en el conjunto de entrenamiento y prueba. D) Para cada solución en cada corrida, se calculó un coeficiente fuerza de predicción o '*Prediction Strength*' (PS) y se representan las distribuciones de los valores obtenidos. Las líneas correspondientes para cada cuadro conectan el valor medio de cada una de las evaluaciones. . . . . 135

- 5.10 **Robustez de Galgo en 30 evaluaciones.** El conjunto de datos METABRIC se dividió en conjunto de entrenamiento y prueba. Galgo se aplicó en el conjunto de entrenamiento para encontrar soluciones no dominadas en 30 evaluaciones independientes. A, B y C) Muestra la relación entre el número de particiones ( $k$ ) y la distribución del C.Index, la PS y los CS de las soluciones obtenidas. Para cada gráfico, una línea roja conecta el valor medio de cada una de las variables en cada  $k$ . En D) se analizan las relaciones pareadas entre el resto de las variables implicadas en Galgo mediante una matriz de correlación. En el triángulo inferior izquierdo de los gráficos en D) se pueden observar los gráficos de dispersión entre las variables de la intersección, donde cada punto representa una de las soluciones de las 30 evaluaciones de Galgo en CM. La línea roja representa el estimador suavizado de Loess que representa la curva de relación entre las variables. Los gráficos de la diagonal en D) muestran la curva de distribución de la variable en la intersección. Finalmente, los gráficos del triángulo superior representan los coeficientes de correlación de Spearman entre las variables en la intersección. . . . . 138
- 6.1 **Evaluación de firmas génicas de HSP para CM.** Mapas de calor (*heatmaps*) que comparan los patrones de expresión génica de la firma HSP-Clust y HSP-Galgo en el conjunto de entrenamiento y el conjunto de pruebas combinado. En la parte superior de cada mapa de calor se muestran barras que representan la clasificación de cada muestra considerando diferentes firmas génicas. En la figura el color azul representa expresión baja mientras que el color rojo representa expresión alta. . . . 147
- 6.2 **Comparación de las firmas HSP-Clust y HSP-Galgo con el resto de las firmas génicas para CM.** Diagramas de bosque (*forest plot*) que representan las estimaciones (puntos) y los intervalos de confianza (líneas) del C.Index para cada firma génica en cada cohorte estudiada. La línea discontinua representa el poder predictivo nulo ( $C.Index = 0.5$ ). Una estimación combinada de cada firma para todas las cohortes se muestra en un cuadro magenta. . . . . 148

6.3	<b>Concordancia entre las firma HSP-Galgo y las firmas HSP-Clust y PAM50.</b> El tamaño de las barras es proporcional al número de muestras en cada categoría. El color azul indica los pacientes clasificados como subtipo I, el rojo los del subtipo II, verde subtipo III y rosado el subtipo IV. . . . .	149
6.4	<b>Evaluación de sobrevida de los subtipos HSP-Galgo en comparación con HSP-Clust y PAM50.</b> Curvas de Kaplan-Meier que comparan las tendencias de supervivencia de los subgrupos de pacientes clasificados por el método PAM50 y por la firma HSP-Clust en comparación con la nueva firma HSP-Galgo, tanto en el conjunto de entrenamiento como en el de prueba combinado. Los valores <i>p</i> de la prueba Log-Rank para cada clasificador se representan en la parte inferior de cada gráfico. . . .	150
6.5	<b>Evaluación de las vías moleculares alteradas en los subtipos HSP-Galgo.</b> GSEA de los subtipos HSP-Galgo. Para cada subtipo, se representan los términos GO alterados más significativos con respecto al resto de las muestras. Se utiliza una escala de color azul a rojo, donde el azul representa los términos de ontología regulados a la baja y el rojo representa los regulados al alza. . . . .	152
7.1	<b>Resumen de la red de interacción proteína-proteína de las principales HSP</b> En la figura se representan las conexiones físicas entre el chaperoma y las diversas proteínas que pueden influir poderosamente en la tumorigénesis. En la red, los nodos representan los componentes más importantes del chaperoma, que están conectados con otras proteínas por bordes de diferente ancho. Los nodos rojos distinguen a las principales HSP y co-chaperonas, los azules son genes bien conocidos relacionados con el cáncer, los genes del complejo TRiC se muestran en verde y los genes del complejo Prp19 / CDC5L están en gris. El grosor de la línea de los bordes indica la fuerza de los datos experimentales que apoyan una interacción proteína-proteína. La red se construyó utilizando la base de datos STRING ( <a href="https://string-db.org">https://string-db.org</a> ) del Instituto Suizo de Bioinformática y el Laboratorio Europeo de Biología Molecular (EMBL). . . . .	157

- 7.2 **Descubrimiento y evaluación de la firma HER2-ECM.** En la figura **A** se muestra la firma HER2-ECM detectada por Galgo, en el conjunto de entrenamiento y reproducida en el conjunto de prueba. Los mapas de calor muestran el patrón de expresión de los genes de la firma, los cuales están asociados a la matriz extracelular y al plegamiento proteico. En **B** podemos observar que la firma puede distinguir de forma robusta dos grupos de pacientes con diferente pronóstico tanto en el conjunto de entrenamiento como en el conjunto de prueba. En **C**, se pueden observar las vías moleculares asociadas a los dos subgrupos detectados y por último en **D** se presenta una comparación del puntaje estimado de infiltración fibroblástica en los tumores mediante el método 'MCP-Counter'. . . . . 167
- 7.3 **Esquema de unión de miRNAs candidatos y BAG3.** En la figura se observa la secuencia de la región 3'UTR del gen BAG3 y los miRNAs predichos para unirse a la misma de acuerdo al algoritmo 'miRanda'. Arriba a la izquierda se pueden observar los puntajes mirSVR y PhatsCons predichos para cada uno de los miRNAs candidatos y abajo a la izquierda se puede ver el ranking de cada miRNA de acuerdo a su valor mirSVR predicho. Los números en azul indican la posición del nucleótido 3'UTR donde el miRNA predicho se une. . . . . 169
- 7.4 **Esquema de unión de miRNAs candidatos y HSPA1A.** En la figura se observa la secuencia de la región 3'UTR del gen HSPA1A y los miRNAs predichos para unirse a la misma de acuerdo al algoritmo 'miRanda'. Arriba a la izquierda se pueden observar los puntajes mirSVR y PhatsCons predichos para cada uno de los miRNAs candidatos y abajo a la izquierda se puede ver el ranking de cada miRNA de acuerdo a su valor mirSVR predicho. Los números en azul indican la posición del nucleótido 3'UTR donde el miRNA predicho se une. . . . . 170

7.5	<p><b>Esquema de unión de miRNAs candidatos y HSP90AA1.</b> En la figura se observa la secuencia de la región 3'UTR del gen HSP90AA1 y los miRNAs predichos para unirse a la misma de acuerdo al algoritmo 'miRanda'. Arriba a la izquierda se pueden observar los puntajes mirSVR y PhatsCons predichos para cada uno de los miRNAs candidatos y abajo a la izquierda se puede ver el ranking de cada miRNA de acuerdo a su valor mirSVR predicho. Los números en azul indican la posición del nucleótido 3'UTR donde el miRNA predicho se une. . . . .</p>	170
7.6	<p><b>Otras dianas de los miRNAs candidatos relacionados con el chaperoma y la matriz extracelular.</b> En la figura se presentan los genes de los procesos estudiados predichos como dianas de al menos uno de los tres miRNAs. En verde se identifican los genes dianas por el microRNA correspondiente de acuerdo a la fila. El negro implica que no hay una unión predicha entre las moléculas estudiadas. Además, los genes se han identificado de acuerdo a las subfamilias de genes más relevantes para nuestro estudio así como también se han identificado otros genes de interés que han sido relacionados con los procesos de modelado de la matriz extracelular. . . . .</p>	171
7.7	<p><b>Expresión y variación en el número de copias de los genes miRNAs candidatos en la cohorte de TCGA.</b> <b>A)</b> Se evaluaron los niveles de expresión total de los tres miRNAs en la cohorte de cáncer de mama comparando su expresión entre el tejido normal y los diferentes subtipos de tejido tumoral para los tres miRNAs. <b>B)</b> También se evaluaron las alteraciones en más y en menos de la variación en el número de copias en los diferentes subtipos tumorales. . . . .</p>	172

- 7.8 **Estado de metilación del gen hsa-miR-522 en CM.** Esta y la figura siguiente (Figura 7.9) está compuesta por diferentes carriles de información integrada. En el carril “Beta”, los valores de metilación se muestran en el eje vertical (eje y), mientras que cada columna de puntos representa un sitio CpG determinado, ubicado en relación con su región cromosómica (mostrado en los carriles inferiores). Cada punto representa una muestra y el valor de metilación medido para el sitio CpG correspondiente. En el carril de valores beta se añadió una curva de regresión suavizada al diagrama de dispersión para mostrar las diferencias de metilación entre los subtipos moleculares y el tejido normal. Los carriles a continuación muestran los ‘*bumps*’ (protuberancias o jorobas) para cada subtipo de tumor, que representan los niveles de metilación en cada región en su conjunto. Obsérvese que a mayor altura de las ‘*bumps*’, mayor nivel de metilación en relación al nivel de referencia. . . 173
- 7.9 **Estado de metilación del gen hsa-miR-570 en CM.** La interpretación es equivalente a la figura 7.8. . . . . 174
- 7.10 **Evaluación de la expresión de las chaperonas HSPA1A, HSP90AA1 y la co-chapreona BAG3 mediante RTq-PCR luego del tratamiento con miR-570 en líneas celulares de cáncer de mama.** A) La línea celular SKBR3 fue transfectada con miR-570 en una concentración de 10, 20 y 30nM y con una condición control de miRNA “scrambled”, no silente, para evaluar la expresión génica de los genes objetivo HSP90AA1, HSPA1A y BAG3 luego de 72hs de tratamiento. B) De la misma manera, la línea celular Hs578t fue transfectada con el miR-570 en dosis de 10 y 20nM y fue comparado con el miRNA control no silente, evaluando la expresión de los mismos genes también luego de las 72hs de tratamiento. El valor expresado es la cuantificación relativa al gen TMEM, utilizado como ‘*house-keeping*’ en cada condición. *ns* =  $p > 0.05$ , \* =  $p \leq 0.05$ , \*\* =  $p \leq 0.01$ , \*\*\* =  $p \leq 0.001$ , \*\*\*\* =  $p \leq 0.0001$ . Los datos expresan la media y una desviación estándar de la expresión relativa ( $n = 3$ ). . . . . 175

- 7.11 Análisis de Western Blot para evaluar la expresión de las proteínas Hsp72 y Bag3 en lisados de cáncer de mama luego del tratamiento con miR-570.** **A)** Las líneas celulares SKBR3 y Hs578t fueron transfectadas con miR-570 en diferentes concentraciones y la expresión proteica fue evaluada mediante la técnica de western blot. Para una mejor observación de las bandas correspondientes a cada proteína, las imágenes fueron cortadas y su contraste fue modificado, sin embargo las imágenes originales completas y sin cortar se encuentran en el apéndice D. **B)** Análisis cuantitativo de la expresión de Hsp72 en la línea celular SKBR3 y Hs578t sometidas al miR-570 o a una condición control. **C)** Para la proteína Bag3 se realizó el mismo procedimiento que para Hsp72. La expresión de las proteínas fue cuantificada mediante análisis densitométrico de las bandas en A) y normalizado contra la expresión total de proteína para su correspondiente calle. Las barras muestran el nivel de expresión proteica relativo al control para cada condición. ( $n = 1$ ). . 176
- 7.12 Curvas dosis-respuesta con valores de IC50 para evaluar la viabilidad celular de la línea SKBR3 luego del tratamiento con 17-alilamino-17-demetoxigeldanamicina (17-AAG) y miARN-570 10nM, 20nM y control no silente.** Los resultados se expresan como porcentaje de la medición de absorbancia del ensayo MTT. La curva de regresión calculada se indica como una línea coloreada. Los puntos indican la media, y las barras, una desviación estándar de las mediciones ( $n = 3$ ). . 177
- 7.13 Ensayo de proliferación celular con y sin inhibidor de Hsp90 (17-AAG) en células tratadas con miR-570.** Después del tratamiento con miR-570 o miRNA no silente (control), se evaluó la proliferación de las células SKBR3 con y sin tratamiento con el inhibidor de Hsp90 17-AAG en concentración 78nM ( $1.89 \log_{10}$  nM). Las células fueron contadas mediante citometría de imagen en campo claro al tiempo 0 y luego de 72hs. El valor expresado es la cuantificación celular relativa al tiempo 0 para cada condición.  $ns = p > 0.05$ ,  $* = p \leq 0.05$ ,  $** = p \leq 0.01$ ,  $*** = p \leq 0.001$ ,  $**** = p \leq 0.0001$ . Los datos expresan la media y una desviación estándar de las mediciones ( $n = 3$ ). . 178

- B.1 Patrones de expresión génica para la firma de mejor rendimiento en CM (BRCA) y CCR (CRC).** En cada mapa de calor (*heatmap*), las columnas representan a los pacientes y las filas a los genes que se encuentran en cada firma. Se utiliza una escala de colores de azul (bajo) a rojo (alto) para mostrar la expresión génica. En la parte superior de cada mapa de calor se muestran las barras que muestran la clasificación de los pacientes con diferentes firmas. Se utiliza una clave de color asociada al riesgo estimado para cada clasificación tumoral correspondientes. 224
- B.2 Patrones de expresión génica para la firma de mejor rendimiento en LUAD y HGSOC.** En cada mapa de calor (*heatmap*), las columnas representan a los pacientes y las filas a los genes que se encuentran en cada firma. Se utiliza una escala de colores de azul (bajo) a rojo (alto) para mostrar la expresión génica. En la parte superior de cada mapa de calor se muestran las barras que muestran la clasificación de los pacientes con diferentes firmas. Se utiliza una clave de color asociada al riesgo estimado para cada clasificación tumoral correspondientes. . . . 225
- D.1 Expresión de las proteínas Hsp72 y Bag3 en lisados de cáncer de mama luego del tratamiento con miR-570.** Las líneas celulares SKBR3 y Hs578t fueron transfectadas con miR-570 en diferentes concentraciones y la expresión proteica fue cuantificada mediante el análisis densitométrico de las bandas. Para una mejor observación de las bandas correspondientes a cada proteína, las imágenes fueron cortadas y su contraste fue modificado. . . . . 227

## LISTA DE TABLAS

1.1	Nomenclatura de la familia de las HSPA (HSP70 y HSP110) . . . . .	10
1.2	Nomenclatura de la familia de las DNAJ (HSP40) . . . . .	11
1.3	Nomenclatura de la familia de las HSPB (HSPs pequeñas) . . . . .	12
1.4	Nomenclatura de la familia de las HSPC (HSP90) . . . . .	13
1.5	Nomenclatura de la familia de las Chaperoninas . . . . .	14
1.6	Uso de biomarcadores en el cáncer de mama . . . . .	48
3.1	<b>Resumen de genes desregulados en CM.</b> Los datos tabulados muestran el número y los porcentajes de genes totales y genes HSPs que presentan un cambio absoluto $\geq 2$ veces, en las muestras totales y de acuerdo con los subtipos intrínsecos de CM . . . . .	71
3.2	Resumen de la clasificación PAM50 y las características inmunohistoquímicas de los tumores de CM de la cohorte de TCGA. . . . .	71
3.3	Estado de metilación de los sitios CpG de las HSPs . . . . .	80
3.4	<b>Regiones CpG asociadas a cambios en la expresión génica.</b> Para cada sitio CpG analizado se estimó la correlación de Pearson, Spearman y el Coeficiente de regresión lineal (Coef Reg) en relación a la expresión génica correspondiente al gen de la región analizada. En la tabla se muestran solo los resultados correspondientes a coeficiente de correlación de Pearson $< -0,3$ . Las regiones están identificadas de acuerdo a un número de identificación junto a su ubicación en el genoma de referencia (sitio de inicio, fin, longitud y el cromosoma donde se encuentra). . . . .	82

3.5	<b>Riesgos proporcionales de Cox univariado para las HSPs en CM.</b>	
	En la presente tabla se exponen la estimación del riesgo asociado a la expresión creciente de las diferentes HSPs, si el ' <i>Hazard Risk</i> ' ( <i>HR</i> ) es $< 1$ , la expresión de dicho gen está asociado a mejor pronóstico, si es $> 1$ , está asociado a peor pronóstico. A su vez se informan los valores de los coeficientes de regresión, el desvío estándar de los coeficientes ( <i>CoefSE</i> ), el valor <i>p</i> asociado ( <i>P - val</i> ) y el valor <i>p</i> ajustado para múltiples comparaciones ( <i>FDR</i> ). Solo se muestran los genes HSPs con $FDR < 0,05$ . . . . .	87
5.1	Parámetros de Galgo: se describe la configuración utilizada para ejecutar Galgo en todos los tipos de cáncer . . . . .	119
5.2	Resumen de resultados de la implementación de Galgo en los cuatro tipos de cáncer analizados . . . . .	127
6.1	Parámetros de Galgo: se describe la configuración utilizada para ejecutar Galgo para el conjunto de genes HSPs en cáncer de mama . . . . .	144
A.1	Cohortes de cáncer de mama . . . . .	217
A.2	Cohortes de adenocarcinoma de pulmón . . . . .	218
A.3	cohortes de cáncer colorectal . . . . .	218
A.4	Cohortes de carcinoma seroso de ovario . . . . .	219
C.1	<b>Predicción de unión de miRNA-570 a la región 3'UTR de genes relacionados al complejo de silenciamiento inducido por ARNs (RISC) mediante el algoritmo '<i>miRanda</i>' . . . . .</b>	226

## LISTA DE ABREVIACIONES

<b>17-AAG</b>	17-alilamino-17-demetoxigeldanamicina	xxxii, 27, 165, 176–178, 180, 183
<b>AG</b>	algoritmos genéticos	116
<b>APC</b>	células presentadoras de antígeno	30, 34, 155
<b>C.Index</b>	índice de concordancia	xxii–xxv, xxvii, 118–124, 126, 129–133, 135, 136, 138, 145–148
<b>CCR</b>	cáncer colorrectal	xxi, xxiii, xxv, xxxiii, 111, 115, 119, 120, 122–124, 128, 129, 131, 134, 139, 224
<b>CD</b>	células dendríticas	x, 31, 34, 35
<b>CDI</b>	carcinoma ductal invasivo	36, 43
<b>CLI</b>	carcinoma lobulillar invasivo	36, 43
<b>CM</b>	cáncer de mama	i, xi, xiv–xviii, xx–xxii, xxv, xxvii, xxxi, xxxiii–xxxv, 7, 15, 35–38, 41, 42, 44–46, 55–57, 60–62, 67–72, 74, 76–81, 84–93, 96–100, 104, 107, 109–112, 115, 116, 119, 120, 122–124, 126–128, 134, 138, 139, 142, 143, 145–148, 159, 161, 171, 173, 174, 181, 182, 224
<b>CMS</b>	subtipo molecular de consenso	111, 120, 122, 128
<b>CNV</b>	Variación en número de copias	163, 171, 172, 181
<b>CS</b>	coeficiente de silueta	xxvi, xxvii, 117, 118, 121, 123, 133, 135–138
<b>DDR</b>	respuesta al daño del ADN	27, 29
<b>ECM</b>	matriz extra-celular	24, 159–161, 166, 168, 179, 182, 183
<b>EMT</b>	Transición Epitelio Mesenquimal	24, 25, 29
<b>ERE</b>	estrés del retículo endoplásmico	89, 90
<b>FAC</b>	fibroblastos asociados a cáncer	24
<b>FC</b>	nivel de cambio	xiv, 66
<b>FDR</b>	<i>False Discovery Rate</i>	65, 69, 70, 73

<b>GED</b> genes expresados diferencialmente . . . . .	65, 70
<b>GO</b> Ontología de genes xxii–xxv, xxviii, 122, 126, 129, 130, 132, 146, 152, 161, 162	
<b>GOF</b> ganancia de función . . . . .	18
<b>GSEA</b> análisis de enriquecimiento de conjunto de genes o ' <i>Gene Set Enrichment Analysis</i> ' . . . . .	xxii–xxv, xxviii, 122, 126, 129, 130, 132, 146, 152, 162
<b>HGSOC</b> cáncer de ovario seroso de alto grado . . . . .	xxi, xxv, xxxiii, 110, 111, 115, 119, 120, 122–124, 128, 130–134, 139, 225
<b>HSR</b> respuesta al estrés térmico . . . . .	29
<b>IHC</b> inmunohistoquímica . . . . .	46, 47, 62, 71, 101
<b>Log2</b> logaritmo en base 2 . . . . .	xx, 62, 65, 70, 72, 73, 86, 94, 104, 116, 143
<b>LUAD</b> adenocarcinoma de pulmón . . . . .	xxi, xxiv, xxv, xxxiii, 110, 115, 119, 120, 122–124, 128–131, 134, 139, 225
<b>MAT</b> macrófagos asociados a tumores . . . . .	x, 30, 31
<b>MCP</b> muerte celular programada . . . . .	21, 24
<b>miRNAs</b> micro ARNs . . . . .	i, xii, 51–54, 56, 58, 161–164, 168–172, 174, 179
<b>MMT</b> tumor mamario murino MUC1.Tg MMTV-PyMT . . . . .	159, 160
<b>NCI-GDC</b> ' <i>National Cancer Institute Genomic Data Commons</i> ' . . . . .	115
<b>NGS</b> secuenciación de nueva generación . . . . .	38
<b>NK</b> natural killer . . . . .	35
<b>PAM</b> partición en torno a medoides . . . . .	116, 117
<b>PET</b> tomografía por emisión de positrones . . . . .	153
<b>PS</b> fuerza de predicción o ' <i>Prediction Strength</i> ' . . . . .	xxvi, xxvii, 121, 135–138
<b>RCT</b> receptor de células T . . . . .	34
<b>RDM</b> regiones diferencialmente metiladas . . . . .	68
<b>RE</b> receptor de estrógenos . . . . .	15, 36, 41–45, 47, 62, 94, 96, 111
<b>RI</b> radiación ionizante . . . . .	29

<b>RMST</b> tiempo medio de supervivencia restringido . . . . .	117
<b>RP</b> receptor de progesterona . . . . .	36, 62, 96
<b>RSEM</b> “ <i>RNA-Seq by Expectation-Maximization</i> ” . . . . .	62, 94
<b>SC</b> señales de crecimiento . . . . .	15
<b>SLP</b> supervivencia libre de progresión . . . . .	115
<b>SSP</b> predictor de muestra única o ‘ <i>Single Sample Predictor</i> ’	71, 95, 96, 104, 107, 109
<b>TCGA</b> ‘ <i>The Cancer Genome Atlas</i> ’ . . . . .	xiii, xviii, xx, xxxiv, 57, 61–64, 67, 68, 71, 94–97, 104, 106, 107, 115, 139, 163
<b>TMR</b> tumores mamarios de ratón . . . . .	24
<b>TNBC</b> tumores mamarios triple-negativo . . . . .	107
<b>WT</b> ‘ <i>Wild Type</i> ’ . . . . .	18, 22, 27

## CAPÍTULO 1

### Introducción

#### 1.1 Cáncer, generalidades y epidemiología

El cáncer, una enfermedad caracterizada por la rápida y descontrolada proliferación y diseminación celular, ha presentado importantes avances en cuanto a su comprensión y su concepción en los últimos 50 años, sin embargo, continúa siendo en el presente una de las principales causas de morbilidad y mortalidad para la humanidad, sin distinción geográfica o de raza.

Se estima que durante el 2018 hubieron 18,1 millones de nuevos casos de cáncer y 9,6 millones de muertes por cáncer a nivel mundial, siendo el cáncer de pulmón el más comúnmente diagnosticado (11,6% del total de casos) y la principal causa de muerte por cáncer (18,4% del total de muertes por cáncer), seguida de cerca por el cáncer de mama femenino (11,6%), el cáncer de próstata (7,1%) y el cáncer colorrectal (6,1%) (*Bray et al.*, 2018).

Entre las mujeres, el cáncer de mama es el cáncer más comúnmente diagnosticado y la principal causa de muerte por cáncer, seguido del cáncer colorrectal y de pulmón por incidencia y viceversa en mortalidad. Sin embargo, el cáncer diagnosticado con más frecuencia y la principal causa de muerte por cáncer varían sustancialmente entre países y dentro de cada país, según el grado de desarrollo económico y los factores sociales y de estilo de vida asociados (*Bray et al.*, 2018).

Según las últimas estimaciones realizadas por la Agencia Internacional de Investigación sobre Cáncer (IARC), Argentina presenta una tasa de incidencia de 212 casos por 100.000 habitantes, lo que corresponde a más de 125.000 casos nuevos de cáncer en ambos sexos por año, con porcentajes similares tanto en hombres como en mujeres (*Ministerio de Salud y Desarrollo Social de la Nación (Argentina)*, 2019b). Junto con su alta incidencia, en el año 2017 la mortalidad por cáncer registrada en Argentina fue de 118 y 87 defunciones por cada 100.000 varones y mujeres, respectivamente. Con algo más de 62.000 fallecimientos registrados, representando el 19,2% del total de las defunciones, ubicando al cáncer como la segunda causa de muerte en la población Argentina, precedidas por las enfermedades cardiovasculares (28,5%) y seguidas por las

respiratorias (19%)(*Ministerio de Salud y Desarrollo Social de la Nación (Argentina)*, 2019a).

En Argentina, durante el 2018, el cáncer de mama fue el de mayor ocurrencia con más de 21.000 casos anuales, representando el 17% de todos los tumores malignos y casi un tercio de los cánceres femeninos. En relación al cáncer colorrectal se estimaron 15.692 casos en ambos sexos, concentrando el 13% del total de tumores, seguidos por el cáncer de próstata y de pulmón (9,3%) (*Ministerio de Salud y Desarrollo Social de la Nación (Argentina)*, 2019b).

Puesto en perspectiva, esta enfermedad, junto a las enfermedades cardiovasculares y la diabetes, integra el grupo de las ENT (enfermedades no transmisibles) más relevantes; las cuales concentraron el 50% de todas las defunciones registradas en 2017 en Argentina (*Ministerio de Salud y Desarrollo Social de la Nación (Argentina)*, 2019a), por lo que representa un elevado costo social, económico y de calidad de vida, tanto para los pacientes que la padecen, sus familiares y para el sistema de salud, responsable de la atención de estos pacientes.

## 1.2 La biología del cáncer y las Proteínas de Golpe de Calor

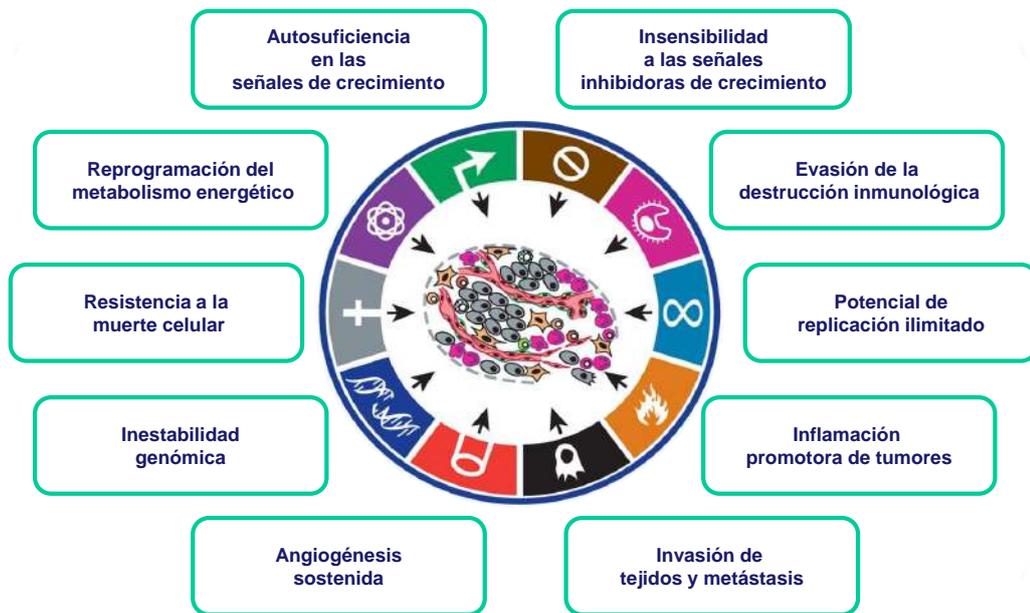
Desde la perspectiva biológica, el cáncer se puede interpretar como una enfermedad que involucra cambios dinámicos en el genoma como mutaciones, deleciones, amplificaciones y cambios epigenéticos producen “oncogenes” con ganancia de función dominante y “genes supresores de tumores” con una pérdida de función de carácter recesivo (*Hanahan and Weinberg*, 2000).

Varias líneas de evidencia indican que la tumorigénesis en humanos es un proceso de múltiples pasos y que estos pasos reflejan alteraciones genéticas que impulsan la transformación progresiva de células humanas normales en derivados malignos (*Balmain et al.*, 1993). Los análisis patológicos de varios sitios en órganos revelan lesiones que parecen representar los pasos intermedios en un proceso por el cual las células evolucionan progresivamente desde la normalidad, a través de una serie de estados pre-malignos, hacia cánceres invasivos (*Hanahan and Weinberg*, 2000).

Estas observaciones se han vuelto más concretas gracias a una gran cantidad de trabajos que indican que los genomas de las células tumorales se encuentran alterados en múltiples sitios, habiendo sufrido alteraciones a través de lesiones tan sutiles como mutaciones puntuales y tan obvias como cambios en el complemento cromosómico (*Kinzler and Vogelstein*, 1996). En conjunto, las observaciones de los cánceres humanos y los modelos animales argumentan que el desarrollo del tumor procede a través de un proceso formalmente análogo a la evolución darwiniana, en el que una sucesión de cambios genéticos, cada uno de los cuales confiere uno u otro tipo de ventaja de crecimiento, conduce a la conversión progresiva de células normales en células cancerosas

(Nowell, 1976). En este sentido, podemos considerar que las células cancerosas tienen defectos en los circuitos moleculares que gobiernan la proliferación y muerte celular normal y la homeostasis, dando lugar a numerosas manifestaciones fenotípicas. Así, hoy en día podemos describir numerosos tipos distintos de cáncer de acuerdo al órgano o tejido de origen, donde a su vez, podemos distinguir diferentes subtipos tumorales dentro de cánceres bien establecidos (Hoadley *et al.*, 2018; Guerrero-Gimenez *et al.*, 2020).

Dentro de la inmensa complejidad de los tumores malignos, ciertos criterios unificados han sido establecidos para lograr una mayor comprensión de los mecanismos biológicos subyacentes al comportamiento maligno de las células tumorales. De esta manera Hanahan and Weinberg (2011) pudieron resumir el vasto catálogo de alteraciones moleculares asociadas al cáncer en relación a 10 alteraciones esenciales de la fisiología celular. Estos puntos son denominados como los “Sellos del Cáncer” (Hanahan and Weinberg, 2011): Autosuficiencia en las señales de crecimiento, insensibilidad a las señales inhibitorias del crecimiento, resistencia a la muerte celular, potencial replicativo ilimitado, angiogénesis sostenida, desregulación de la energética celular, sensibilidad a las señales inflamatorias, evasión de la respuesta inmune, inestabilidad genómica y quizás una de las más características, la capacidad de invadir y hacer metástasis a otros tejidos (Fig.1.1).



**Fig. 1.1 Los sellos del cáncer.** El esquema ilustra las condiciones necesarias para el desarrollo de la enfermedad maligna. Estos rasgos característicos pueden ser adquiridos en diferentes etapas en el desarrollo del cáncer, a través de distintivos mecanismos de acuerdo al tipo de cáncer. Modificado de *Hanahan and Weinberg* (2011).

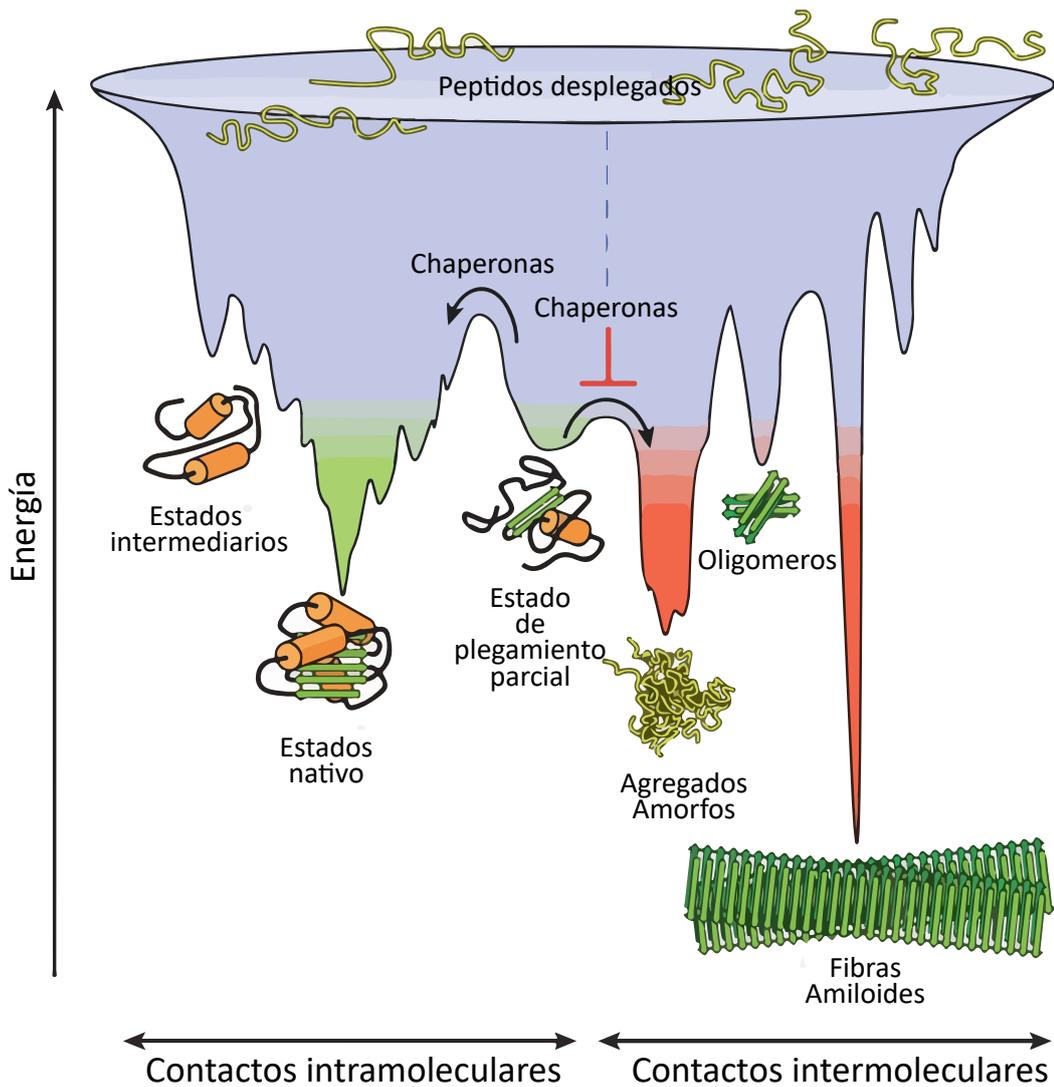
### 1.2.1 Las Proteínas de Golpe de Calor

Las proteínas, elementos centrales en los procesos biológicos, se sintetizan en los ribosomas como cadenas lineales de hasta varios miles de aminoácidos. Sin embargo, para funcionar adecuadamente, estas cadenas generalmente deben plegarse a su 'estado nativo', una estructura tridimensional determinada mediante la cual la proteína cumple su función (*Hartl*, 1996). Un aspecto fundamental de este problema es que las proteínas deben conservar cierta flexibilidad conformacional para funcionar y, por lo tanto, son sólo marginalmente estables termodinámicamente en su entorno fisiológico (*Hartl et al.*, 2011). En este sentido, una fracción sustancial de todas las proteínas en las células eucariotas adopta conformaciones plegadas solo después de la interacción con ciertos compañeros de unión (*Dunker et al.*, 2008). Por lo tanto, el control de la calidad de las proteínas y el mantenimiento de la homeostasis del proteoma (conocido como proteostasis) son cruciales para la salud celular y del organismo.

La proteostasis se logra mediante una red integrada de varios cientos de proteínas, que incluyen, de manera prominente, a las chaperonas moleculares denominadas HSPs (Heat shock proteins: proteínas de golpe de calor). Estas proteínas son moléculas alta-

mente conservadas presentes en todos los organismos vivos y si bien fueron descritas por su prominente expresión ante cambios bruscos de temperatura ('Heat Shock'), hoy sabemos que su presencia es ubicua en las células y su expresión se exagera ante diversos tipos de estrés tanto fisiológico como patológico (*Ciocca and Calderwood, 2005*). Las HSPs participan en numerosos eventos moleculares, destacándose por su relevancia en los sistemas de control de calidad de proteínas (PQC), vías de degradación (Sistema ubiquitina-proteosoma, degradación asociada a retículo endoplásmico), autofagia y regulación de la apoptosis entre otras funciones (*Lang et al., 2019*).

En general, es aceptado que las chaperonas actúan como facilitadoras del correcto plegado de proteínas celulares, modificando las barreras termodinámicas en los caminos de plegamiento proteico (Fig.1.2).



**Fig. 1.2 Superficie de energía libre en el plegamiento proteico.** Representación de la superficie de energía libre que las proteínas exploran a medida que avanzan hacia el estado nativo (verde) formando contactos intramoleculares. La dificultad del paisaje de energía libre da como resultado la acumulación de conformaciones atrapadas cinéticamente que necesitan atravesar barreras de energía libre para alcanzar un camino favorable. *In vivo*, estos pasos se aceleran mediante las chaperonas. Cuando varias moléculas se pliegan simultáneamente en el mismo compartimento, la superficie de plegado de energía libre puede solaparse con la de la agregación intermolecular, lo que da como resultado la formación de agregados amorfos, oligómeros tóxicos o fibrillas amiloides ordenadas (rojo). La agregación fibrilar ocurre típicamente por polimerización dependiente de nucleación. Puede iniciarse a partir de intermedios poblados durante el plegamiento *de novo* o después de la desestabilización del estado nativo (estados parcialmente plegados) y normalmente se evita mediante chaperonas moleculares. Modificado de *Hartl et al.* (2011).

Las HSPs constituyen una familia de genes conservada evolutivamente que incluyen 95 genes divididos en 6 subfamilias de acuerdo al peso molecular de las proteínas que codifican o su grado de homología de secuencia: las Hsp pequeñas de 12-43kDa (HSPB), Hsp40 (DNAJ), Hsp70 (HSPA), Hsp90 (HSPC) y las Hsp grandes de 100-110kDa (HSPH), finalmente los genes de las Hsp10, Hsp60 y proteínas CCT se agrupan bajo una misma subfamilia (Chaperoninas) y la familia HSPA y HSPH se agrupan

bajo un mismo grupo denominado HSP70 (*Kampinga et al.*, 2009). De forma aclaratoria, a partir de aquí, la terminología en minúscula (por ej. Hsp70) hará referencia a las proteínas codificadas por un determinado gen, mientras que la terminología en mayúscula (por ej. HSPA8), hará referencia a los genes o a una familia de genes (por ej. Hsp90 es codificada por el gen HSP90AA1, que pertenece a la familia de genes HSPC).

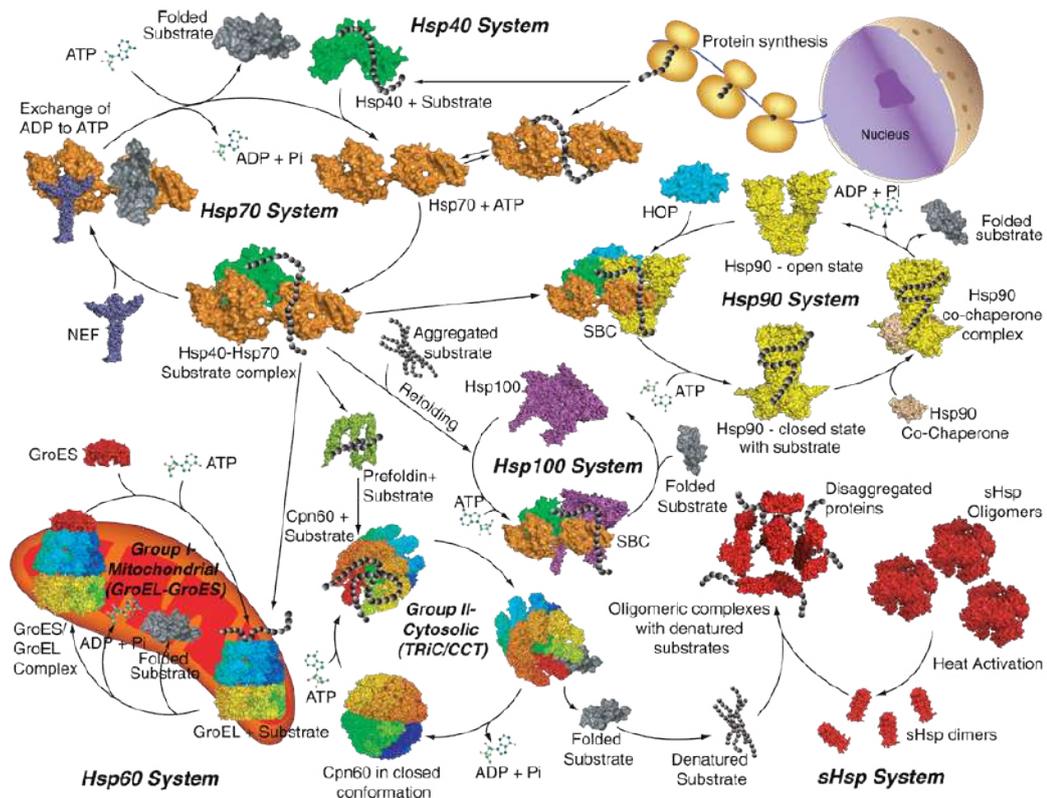
En este punto es importante destacar que, a pesar de que el nombre “*Heat Shock Proteins*” hace referencia a la capacidad de estos genes de ser inducibles por calor, no todos los integrantes de la familia HSP lo son. El motivo de esta divergencia es una consecuencia de que la nomenclatura utilizada combina la concepción histórica de los exponentes clásicos de la familia como HSP90AA1, HSPB1 y HSPA1A, que fueron descubiertos y descritos en relación a las observaciones de su potente inducción ante el golpe de calor, junto con el análisis computacional de homología de secuencia, el cual permite identificar genes pertenecientes a la misma familia pero que no necesariamente son inducibles por estrés térmico.

Las HSPs están sobre expresadas en una amplia variedad de cánceres en los cuales presentan diferentes acciones a veces pro-tumorales, promoviendo el crecimiento tumoral y la metástasis, y otras veces anti-tumorales por ejemplo inhibiendo la muerte celular por apoptosis (*Ciocca and Calderwood*, 2005; *Calderwood et al.*, 2006; *Ciocca et al.*, 2010)

Actualmente las HSPs están surgiendo como dianas moleculares en la terapia del cáncer debido a la diversidad de funciones que cumplen en las células cancerosas. Numerosos fármacos inhibidores de las HSPs está siendo sometidos a pruebas de ensayos clínicos para el tratamiento de varios tipos de cáncer entre el cuales se incluye el de mama (*Rappa et al.*, 2012; *Ciocca et al.*, 2012; *Wawrzynow et al.*, 2018). Los genes HSPs (y las proteínas que codifican) mayormente estudiados corresponden a los inducidos por el calor, tales como HSPA1A/B, DNAJB1, HSPB1 y HSP90AA1. Estos han sido evaluados en diversos modelos (cultivo de células, biopsias etc.), sin embargo muchas de las restantes HSPs todavía no han sido estudiadas en el contexto del cáncer y particularmente no se conoce el rol de muchas de ellas en el CM. La información recopilada de estudios de diversas fuentes brinda un panorama no pocas veces contradictorio respecto del rol de las HSPs en diferentes situaciones asociadas a la patología y al momento de comenzar nuestra investigación no habían estudios que integren las diversas plataformas “ómicas” disponibles en relación a estos genes.

Si bien las diferentes HSPs han sido agrupadas en diferentes familias de acuerdo a la homología entre sus constituyentes, sus funciones se encuentran altamente interrelacionadas y, de hecho, en muchos caso trabajan en forma de relevos como se puede observar en la figura 1.3. A continuación haremos una breve descripción de cada una de las características distintivas de las diferentes subfamilias de los genes HSPs y sus

principales funciones conocidas.



**Fig. 1.3 Red funcional de la familia HSPs necesaria para el mantenimiento del control de calidad de las proteínas celulares.** En la figura se pueden observar los ciclos de plegamiento de las diferentes familias de HSPs y sus sustratos, su funcionamiento mediante 'relevos' y la dependencia de algunos complejos HSPs de ATP y otros co-factores (Como factores intercambiadores de nucleótidos (NEF), Hop y prefoldinas, por mencionar algunas). Hsp70 (HSPA), Hsp110 (HSPH), sHsp (HSPB), Hsp90 (HSPC), Hsp40 (DNAJ), Hsp60 (CHAP I), TRiC/CCT (CHAP II). Extraído de <http://pdslab.biochem.iisc.ernet.in/hspir/chaperone.php> (Ratheesh Kumar et al., 2012)

### 1.2.1.1 Familia de las HSP70 (HSPA y HSPH)

El genoma humano codifica 13 miembros de la familia HSPA, excluyendo numerosos pseudogenes (Brocchieri et al., 2008) (Tabla 1.1). Los genes más estudiados son HSPA1A y HSPA1B, cuyos productos solo se diferencian en dos aminoácidos y se cree que son proteínas totalmente intercambiables. Junto con HSPA6, estos son los miembros de la familia más inducibles por calor. Durante mucho tiempo se consideró a HSPA7 como un pseudogén, sin embargo hoy sabemos que constituye un gen verdadero altamente homólogo a HSPA6. HSPA8 es otra análoga HSPA que se expresa de forma constitutiva, se designó anteriormente como Hsc70 (o HSP73). Es un miembro esencial de las HSPA de "mantenimiento" y participa en el plegamiento cotraduccional y la translocación de proteínas a través de las membranas intracelulares. HSPA1L y HSPA2 son dos miembros de la familia citosólica con alta expresión en los testículos. HSPA9 es un miembro de HSPA mitocondrial (también conocida como mortalin / mtHSP70 / GRP75

/PBP74). HSPA5 se encuentra localizada en el retículo endoplasmático (ER) y también es conocida como BiP. HSPA13, también conocida como Stch, se encuentra en microsomas. HSPA12A, HSPA12B y HSPA14 son miembros relacionados más lejanamente sobre los que hay muy pocos datos disponibles (*Kampinga et al.*, 2009).

Las formas de HSPA expresadas constitutivamente e inducibles por estrés son actores centrales en el plegamiento de proteínas y el control de la proteostasis. Las HSPA funcionan a través de un ciclo de reacción dependiente de ATP regulado por cochaperonas de la familia Hsp40 (DNAJs) y factores intercambiadores de nucleótidos o 'NEF' por sus siglas en inglés (como las proteínas BAG1 y SIL1 entre otras), que a partir de numerosos ciclos de interacción con sus proteínas clientes, conducen a una remodelación conformacional de las estructuras terciarias, modificando las barreras cinéticas del proceso de plegado proteico, favoreciendo a un correcto plegado de las mismas (Figura 1.3 (*Hartl et al.*, 2011)).

Por otro lado, tenemos cuatro genes que codifican para proteínas de mayor peso molecular, de alrededor de 110 KDa, agrupados en la subfamilia HSPH (Hsp110) (Tabla 1.1). Los mismos poseen alta homología con los miembros de HSPA, excepto por la existencia de un dominio intermedio más largo entre el dominio ATPasa N-terminal y el dominio de unión a péptidos C-terminal. De hecho, dos miembros, HSPA4 (HSPH2) y HSPA4L (HSPH3), han sido identificados como miembros de HSPA en la base de datos de Entrez Gene, nomenclatura que utilizaremos en el presente trabajo. Además de los tres miembros citosólicos, HYOU1 (Grp170) es un miembro de HSPH específico del compartimento del retículo endoplasmático, el cual también puede encontrarse como HSPH4 en la bibliografía. Evidencia reciente muestra que los miembros de HSPH funcionan como factores de intercambio de nucleótidos para la familia HSPA (*Dragovic et al.*, 2006; *Raviol et al.*, 2006).

**Table 1.1** Nomenclatura de la familia de las HSPA (HSP70 y HSP110)

	<b>Gen</b>	<b>Proteína</b>	<b>Otros nombres</b>	<b>Gene ID</b>
	<i>HSPA1A</i>	HSPA1A	HSP70-1; HSP72; HSPA1	3303
	<i>HSPA1B</i>	HSPA1B	HSP70-2	3304
	<i>HSPA1L</i>	HSPA1L	hum70t; hum70t; Hsp-hom	3305
	<i>HSPA2</i>	HSPA2	Heat-shock 70kD protein-2	3306
	<i>HSPA5</i>	HSPA5	BIP; GRP78; MIF2	3309
	<i>HSPA6</i>	HSPA6	Heat shock 70kD protein 6 (HSP70B')	3310
HSPA	<i>HSPA7</i>	HSPA7	Heat shock 70kD protein 7	3311
	<i>HSPA8</i>	HSPA8	HSC70; HSC71; HSP71; HSP73	3312
	<i>HSPA9</i>	HSPA9	GRP75; HSPA9B; MOT2; PBP74	3313
	<i>HSPA12A</i>	HSPA12A	FLJ13874; KIAA0417	259217
	<i>HSPA12B</i>	HSPA12B	RP23-32L15.1; 2700081N06Rik	116835
	<i>HSPA13</i>	HSPA13	Stch	6782
	<i>HSPA14</i>	HSPA14	HSP70-4; HSP70L1; MGC131990	51182
	<i>HSPH1</i>	HSPH1	HSP105	10808
	<i>HSPA4</i>	HSPA4	HSPH2; APG-2; HSP110	3308
HSPH	<i>HSPA4L</i>	HSPA4L	HSPH3; APG-1	22824
	<i>HYOU1</i>	HYOU1	Grp170; ORP150; HSP12A	10525

### 1.2.1.2 La familia de las DNAJ (HSP40)

La familia DNAJ (HSP40) es la familia más numerosa de HSPs (Tabla 1.2), la cual se identifica por la presencia de un dominio J conservado responsable del reclutamiento de HSPA y de la activación de su función ATPasa (Figura 1.3). *Hennessy et al.* (2005) dividieron esta familia en tres subfamilias basándose en su homología con la proteína DnaJ de *E. coli* (*Hennessy et al.*, 2005). En este sentido, el genoma humano codifica cuatro proteínas de tipo A que muestran homología con la proteína DnaJ de *E. coli* y contienen un dominio J N-terminal, una región rica en glicina / fenilalanina, una región rica en cisteína y un dominio C-terminal variable. El subgrupo B contienen también dominio J N-terminal y una región adyacente rica en glicina / fenilalanina, siendo DNAJB1 el miembro DNAJ humano más ampliamente expresado y e inducible por calor. Por último, las DNAJ de tipo C solo contienen el dominio J pero no necesariamente ubicadas en el extremo N-terminal (*Kampinga et al.*, 2009). Se ha sugerido que estos miembros recluten miembros de HSPA para subcompartimentos y/o funciones específicas, sin embargo es escaso el conocimiento que se tiene de gran parte de los miembros de esta familia de co-chaperonas.

**Table 1.2** Nomenclatura de la familia de las DNAJ (HSP40)

	<b>Gen</b>	<b>Proteína</b>	<b>Otros nombres</b>	<b>Gene ID</b>
DNAJA	<i>DNAJA1</i>	DNAJA1	DJ-2; DjA1; HDJ2; HSDJ; HSJ2; HSPF4	3301
	<i>DNAJA2</i>	DNAJA2	DNJ3; mDj3; Dnaj3; HIRIP4	10294
	<i>DNAJA3</i>	DNAJA3	Tid-1; Tid11	9093
	<i>DNAJA4</i>	DNAJA4	Dj4; Hsj4	55466
DNAJB	<i>DNAJB1</i>	DNAJB1	HSPF1; HSP40	3337
	<i>DNAJB2</i>	DNAJB2	HSJ1; HSPF3; Dnajb10; MDJ8	3300
	<i>DNAJB3</i>	DNAJB3	Hsj3; Msj1; MSJ-1; Hcg3	414061
	<i>DNAJB4</i>	DNAJB4	Hsc40	11080
	<i>DNAJB5</i>	DNAJB5	Hsc40; HSP40-3	25822
	<i>DNAJB6</i>	DNAJB6	Mrj; mDj4	10049
	<i>DNAJB7</i>	DNAJB7	Dj5; mDj5	150353
	<i>DNAJB8</i>	DNAJB8	mDj6	165721
	<i>DNAJB9</i>	DNAJB9	Mdg1; mDj7; ERdj4	4189
	<i>DNAJB11</i>	DNAJB11	Dj9; ABBP-2; Erdj3	51726
	<i>DNAJB12</i>	DNAJB12	Dj10; mDj10	54788
	<i>DNAJB13</i>	DNAJB13	Tsarg6; Tsarg 3 protein	374407
	<i>DNAJB14</i>	DNAJB14	EGNR9427; FLJ14281	79982
	DNAJC	<i>DNAJC1</i>	DNAJC1	MTJ1; ERdj1; ERj1p; Dnajl1
<i>DNAJC2</i>		DNAJC2	Zrf1; Zrf2; MIDA1; MPP11; zuotin; ZUO1	27000
<i>DNAJC3</i>		DNAJC3	p58; mp58; Prkri; Dnajc3; p58IPK; Dnajc3b	5611
<i>DNAJC4</i>		DNAJC4	HSPf2; Mcg18	3338
<i>DNAJC5</i>		DNAJC5	Csp	80331
<i>DNAJC5B</i>		DNAJC5B	CSP-beta	85479
<i>DNAJC5G</i>		DNAJC5G	MGC107182; gamma-CSP	285126
<i>DNAJC6</i>		DNAJC6	mKIAA0473; auxilin	9829
<i>DNAJC7</i>		DNAJC7	Ttc2; mDj11; mTpr2	7266
<i>DNAJC8</i>		DNAJC8	AL024084; AU019262; splicing protein (spf31)	22826
<i>DNAJC9</i>		DNAJC9	AU020082; RcDNAJ9	23234
<i>DNAJC10</i>		DNAJC10	JPDI; ERdj5; macrothioredoxin	54431
<i>DNAJC11</i>		DNAJC11	FLJ10737; dJ126A5.1	55735
<i>DNAJC12</i>		DNAJC12	Jdp1; mJDP1	56521
<i>DNAJC13</i>		DNAJC13	Rme8; RME-8; Gm1124	23317
<i>DNAJC14</i>		DNAJC14	HDJ3; LIP6; DRIP78	85406
<i>DNAJC15</i>		DNAJC15	Dnajd1; MCJ; Cell growth-inhibiting 22 protein	29103
<i>DNAJC16</i>		DNAJC16	mKIAA0962	23341
<i>DNAJC17</i>		DNAJC17	C87112	55192
<i>DNAJC18</i>		DNAJC18	MGC29463	202052
<i>DNAJC19</i>		DNAJC19	TIM14; TIMM14	131118
<i>HSCB</i>		HSCB	DNAJC20; JAC1; HSC20;	150274
<i>DNAJC21</i>		DNAJC21	GS3; JJJ1; DNAJA5	134218
<i>DNAJC22</i>		DNAJC22	FLJ13236; Wurst	79962
<i>SEC63</i>		SEC63	DNAJC23; AI649014	11231
<i>DNAJC24</i>		DNAJC24	DPH4; zinc finger, CSL-type containing 3	120526
<i>DNAJC25</i>		DNAJC25	AAH48318; LOC552891; G-protein gamma 10	548645
<i>GAK</i>		GAK	DNAJC26; cyclinG associated kinase; auxilin-2	2580
<i>DNAJC27</i>		DNAJC27	RBJ; RabJ	51277
<i>DNAJC28</i>		DNAJC28	Orf28; C21orf55; oculomedin	54943
<i>SACS</i>	SACS	DNAJC29; Sacsin	26278	
<i>DNAJC30</i>	DNAJC30	WBSCR18	84277	

### 1.2.1.3 La familia de las HSPB (HSPs pequeñas)

La familia de las HSPB consta actualmente de 11 miembros que se caracterizan por un dominio cristalino característico conservado flanqueado por extremos N y C variables (Tabla 1.3). Los miembros mejor estudiados son HSPB1 (HSP27), CRYAA ( $\alpha$ A crystallin) y CRYAB ( $\alpha$ B crystallin). Estas HSPs pequeñas se encuentran a menudo en complejos oligoméricos que involucran a uno o más miembros de la familia y, como tales, pueden proporcionar a la célula una gran diversidad en la especificidad de las chaperonas (Figura 1.3) (*Ciocca and Calderwood, 2005*).

**Table 1.3** Nomenclatura de la familia de las HSPB (HSPs pequeñas)

Gen	Proteína	Otros nombres	Gene ID
<i>HSPB1</i>	HSPB1	CMT2F; HMN2B; HSP27; HSP28; HSP25; HS.76067; DKFZp586P1322	3315
<i>HSPB2</i>	HSPB2	MKBP; HSP27; Hs.78846; LOH11CR1K; MGC133245	3316
<i>HSPB3</i>	HSPB3	HSPL27	8988
<i>CRYAA</i>	CRYAA	crystallin alpha A; HSPB4, CRYA1	1409
<i>CRYAB</i>	CRYAB	crystallin alpha B, HSPB5; CRYA2	1410
HSPB <i>HSPB6</i>	HSPB6	HSP20; FLJ32389	126393
<i>HSPB7</i>	HSPB7	cvHSP; FLJ32733; DK-FZp779D0968	27129
<i>HSPB8</i>	HSPB8	H11; HMN2; CMT2L; DHMN2; E2IG1; HMN2A; HSP22	26353
<i>HSPB9</i>	HSPB9	FLJ27437	94086
<i>ODF1</i>	ODF1	HSPB10; ODF; RT7; ODF2; ODFP; SODF; ODF27; ODFPG; ODFPGA; ODFPGB; MGC129928; MGC129929	4956
<i>HSPB11</i>	HSPB11	HSP16.2; C1orf41; PP25	51668

### 1.2.1.4 La familia de las HSPC (HSP90)

La familia de las HSPC es más conocida por su nombre asociado a su peso molecular HSP90 y está constituida por cuatro miembros (Tabla 1.4). La familia HSP90 se encuentra muy conservada entre especies y reinos (*Chen et al., 2006*) e incluso se han propuesto a estas chaperonas como capacitores del proceso evolutivo de otras proteínas (*Rutherford and Lindquist, 1998*).

Las HSP90 forman un homodímero a través de sus residuos carboxi-terminales mientras que la región amino-terminal, altamente conservada, contiene un bolsillo de unión a ATP. El dominio N-terminal va seguido de un dominio medio conservado y es-

tructuralmente flexible. El dímero HSP90 sufre un ciclo de reacción impulsado por ATP que se acompaña de un considerable reordenamiento estructural conformando lo que se denomina la “pinza molecular” HSP90. Esto da como resultado una compactación del dímero HSP90, en la que los monómeros individuales se retuercen entre sí. Después de la hidrólisis, los dominios de ATPasa se disocian y los monómeros de HSP90 se separan (Figura 1.3) (Hartl *et al.*, 2011).

Debido a que varios sustratos de HSP90 son quinasas con papeles bien documentados en el desarrollo de tumores, la inhibición de HSP90 con fármacos como la geldanamicina ha surgido como una estrategia prometedora para el tratamiento de ciertos cánceres (Yuno *et al.*, 2018).

**Table 1.4** Nomenclatura de la familia de las HSPC (HSP90)

	<b>Gen</b>	<b>Proteína</b>	<b>Otros nombres</b>	<b>Gene ID</b>
	<i>HSP90AA1</i>	HSP90AA1	HSPC1; HSPN; LAP2; HSP86; HSPC1; HSPCA; HSP89; HSP90; HSP90A; HSP90N; HSP-CAL1; HSPCAL4; FLJ31884	3320
HSPC	<i>HSP90AB1</i>	HSP90AB1	HSPC3; HSPC2; HSPCB; D6S182; HSP90B; FLJ26984; HSP90-BETA	3326
	<i>HSP90B1</i>	HSP90B1	HSPC4; ECGP; GP96; TRA1; GRP94; endoplasmín	7184
	<i>TRAP1</i>	TRAP1	HSPC5; HSP75; HSP90L	10131

### 1.2.1.5 La familia de las Chaperoninas (CHAP I, CHAP II y CHAP III)

Las chaperoninas (Tabla 1.5) son grandes complejos proteicos que funcionan al envolver globalmente proteínas de sustrato de hasta ~60 kDa para ordenar su correcto plegamiento. Las chaperoninas del grupo I (CHAP I) están conformadas por las proteínas HSPD1 (HSP60 o GroEL en bacterias) que forman un complejo en jaula de siete miembros, y HSPE1 (HSP10 o GroES en bacterias), que forman la tapa de la jaula plegable. Los dominios apicales de GroEL presentan residuos de aminoácidos hidrófobos para la unión del sustrato en el centro del anillo. La unión de GroES está regulada por ATP y está asociada con un marcado cambio conformacional de GroEL que conduce a la formación de una jaula con una pared interna altamente hidrófila con carga neta negativa. La proteína encapsulada puede plegarse libremente en este entorno hasta que la hidrólisis del ATP altera la conformación de la jaula en un proceso cíclico que facilita el plegado de los sustratos del complejo (Figura 1.3) (Hartl *et al.*, 2011).

El grupo de las chaperoninas II (CHAP II) está constituido por un complejo multiproteico de dos anillos superpuestos de 8 miembros cada uno denominado complejo TRiC, por su nombre en inglés *T-complex protein Ring Complex*. TRiC está formado

por las proteínas TCP1 y las CCT2 a 8. Hay dos genes que codifican CCT6 (subunidad zeta): CCT6A (zeta-1), que se expresa constitutivamente, mientras que CCT6B (zeta-2) se expresa de manera específica en testículo (*Kampinga et al., 2009*). Ninguno de estos genes ha demostrado ser inducible por calor. Este complejo citosólico funciona de manera similar al complejo HSPD1/HSPE1 donde sus componentes se abren y cierran en un ciclo de encapsulación de proteínas dependiente de ATP (Figura 1.3). TRiC interactúa con aproximadamente el 10% de las proteínas citosólicas recién sintetizadas, incluidas la actina y las tubulinas, y ha sido propuesto como una diana molecular para cáncer de pulmón y de mama (*Carr et al., 2017; Showalter et al., 2020*).

Por último, cabe destacar al grupo de las chaperoninas III (CHAP III), constituido por los genes MKKS/BBS6 BBS10 y BBS12, las cuales están involucradas en la formación del BBSoma, un complejo multiproteico involucrado en la formación de los cilios y el centrosoma (*Sokolov et al., 2019*).

**Table 1.5** Nomenclatura de la familia de las Chaperoninas

	<b>Gen</b>	<b>Proteína</b>	<b>Otros nombres</b>	<b>Gene ID</b>
CHAP I	<i>HSPD1</i>	HSPD1	HSP60; GroEL	3329
	<i>HSPE1</i>	HSPE1	HSP10; chaperonin 10; GroES	3336
CHAP II	<i>TCP1</i>	TCP1	CCT1; CCTA; CCT-alpha; TCP-1-alpha	6950
	<i>CCT2</i>	CCT2	CCTB; CCT-beta; TCP-1-beta	10576
	<i>CCT3</i>	CCT3	CCTG; CCT-gamma; TCP-1-gamma; TRiC-P5	7203
	<i>CCT4</i>	CCT4	CCTD; CCT-delta; TCP-1-delta; SRB	10575
	<i>CCT5</i>	CCT5	CCTE; CCT-epsilon; TCP-1-epsilon	22948
	<i>CCT6A</i>	CCT6A	CCT6; CCTZ; CCT-zeta; CCT-zeta1; TCP-1-zeta; HTR3; TCP20	908
	<i>CCT6B</i>	CCT6B	CCTZ2; CCT-zeta2; TSA303	10693
	<i>CCT7</i>	CCT7	CCTH; CCT-eta; TCP-1-eta	10574
	<i>CCT8</i>	CCT8	CCTQ; CCT-theta; TCP-1-theta; KIAA002	10694
CHAP III	<i>MKKS</i>	MKKS	McKusick–Kaufman syndrome; MKS; Bardet–Biedl syndrome 6; BBS6	8195
	<i>BBS10</i>	BBS10	Bardet–Biedl syndrome 10	79738
	<i>BBS12</i>	BBS12	Bardet–Biedl syndrome 12	166379

## 1.2.2 Autosuficiencia en las señales de crecimiento

Podría decirse que el rasgo más fundamental de las células cancerosas implica su capacidad para mantener una proliferación sostenida. Los tejidos normales controlan cuidadosamente la producción y liberación de señales promotoras del crecimiento que

instruyen la entrada y la progresión a través del ciclo de crecimiento y división celular, asegurando así una homeostasis del número de células y, por lo tanto, el mantenimiento de la arquitectura y función normales del tejido. Las células cancerosas, al desregular estas señales, se convierten en dueñas de sus propios destinos. Las señales habilitadoras son transmitidas en gran parte por factores de crecimiento que se unen a los receptores de la superficie celular, que normalmente contienen dominios de tirosina quinasa intracelulares los cuales, a través de vías de señalización intracelular ramificadas, regulan la progresión a través del ciclo celular así como el crecimiento celular. A menudo, estas señales influyen en otras propiedades biológicas de las células, como la supervivencia celular y el metabolismo energético (*Hanahan and Weinberg, 2011*).

La dependencia de las señales de crecimiento (SC) es evidente cuando se propagan células normales en cultivo, que típicamente proliferan sólo cuando se les suministran factores mitogénicos difusibles apropiados y un sustrato adecuado para sus integrinas. Este comportamiento contrasta fuertemente con el de las células tumorales, que invariablemente muestran una dependencia muy reducida de la estimulación exógena para el crecimiento.

En los tumores, son evidentes tres estrategias moleculares comunes para lograr la autonomía, que implican: la alteración de las señales de crecimiento extracelular, de los transductores transcelulares de esas señales o de los circuitos intracelulares que traducen esas señales en acción.

La mayoría de las células normales se encuentran bajo un crecimiento limitado debido a la falta de exposición a los factores de crecimiento esenciales para la proliferación. Los factores de crecimiento se unen a receptores, normalmente tirosina quinasas, y dan lugar a cascadas de señalización que permiten ondas de expresión génica, lo que permite la proliferación (*Paul and Mukhopadhyay, 2012*).

El escape de esta restricción requiere una mutación activadora de los receptores del factor de crecimiento o de los componentes de señalización posteriores. Por tanto, los genes receptores de EGFR, PDGF, FGF, HER2 y HER3, y las proteínas de señalización como Ras, PI3K y PTEN, a menudo se modifican por desregulación para actuar como oncogenes (*Paul and Mukhopadhyay, 2012; Hanahan and Weinberg, 2011*). Muchas de estas proteínas tienen conformaciones relativamente inestables, que están asociadas con una actividad catalítica competente, y son fuertemente dependientes de la chaperona de Hsp90 (*Kamal et al., 2003; Neckers and Ivy, 2003; Neckers and Workman, 2012*). Por lo tanto, los aumentos de Hsp90 estrechamente asociados con la malignidad pueden permitir el funcionamiento de las cascadas de señalización del receptor del factor de crecimiento y la proliferación en ausencia de estímulos de crecimiento externos (*Kamal et al., 2003; Miyata et al., 2012*). Un ejemplo establecido de la dependencia a HSPs de las células proliferantes en el CM es la asociación del receptor de estrógenos (RE) con Hsp90, en donde, en ausencia de estradiol, el RE, al igual que otros receptores de

esteroides, forma un complejo con Hsp90 y otras co-chaperonas como la inmunofilina FKBP51 y p23 (PTGES3), en donde el complejo mantiene una conformación estructural de alta afinidad para la unión a esteroides (Zgajnar *et al.*, 2019; Fliss *et al.*, 2000).

Este requisito para la estabilización de las oncoproteínas promotoras del crecimiento constituye la base para el desarrollo de inhibidores de Hsp90, que han pasado a la etapa de ensayo clínico (Trepel *et al.*, 2010; Workman, 2004). Dada la dependencia a Hsp90 de estas señales proliferativas claves, se puede considerar a esta proteína como al “padrino de la proliferación tumoral”, una relación que se ejemplifica por la potente restricción bien establecida de la proliferación de células cancerosas inducida por inhibidores de Hsp90 (Xu *et al.*, 2011). También, se ha demostrado que Hsp27 juega un papel clave en la progresión del ciclo celular, al reprimir las proteínas inhibitoras del ciclo celular E2F-4 y p130, y permitir la síntesis de múltiples proteínas del ciclo celular (CCNA2, CCNB1, CCNB, CDC25C, CDC3A, y CDK1) (Park *et al.*, 2018). Por otro lado, la perturbación de la actividad de Hsp70 también inhibe la estabilización de los transductores de señales mitógenos que incluyen BRAF (Budina-Kolomets *et al.*, 2016).

### **1.2.3 Insensibilidad a las señales inhibitoras del crecimiento**

Además de la capacidad característica de inducir y mantener señales estimulantes del crecimiento de acción positiva, las células cancerosas también deben eludir programas poderosos que regulan negativamente la proliferación celular, los cuales en su mayoría dependen de las acciones de los genes supresores de tumores.

Dentro de un tejido normal, operan múltiples señales antiproliferativas para mantener la inactividad celular y la homeostasis del tejido; estas señales incluyen tanto inhibidores del crecimiento solubles como inhibidores inmovilizados incrustados en la matriz extracelular y en las superficies de las células cercanas. Estas señales inhibitoras del crecimiento, al igual que sus contrapartes de acción positiva, son recibidas por receptores transmembrana de la superficie celular acoplados a circuitos de señalización intracelular (Hanahan and Weinberg, 2011).

Las señales anti-crecimiento pueden bloquear la proliferación mediante dos mecanismos distintos. Las células pueden verse forzadas a salir del ciclo proliferativo activo al estado inactivo (G0) del que pueden volver a emerger en alguna ocasión futura cuando las señales extracelulares lo permitan o, alternativamente, se puede inducir a las células a renunciar permanentemente a su potencial proliferativo induciéndolas a entrar en estados posmitóticos, generalmente asociados con la adquisición de rasgos específicos asociados a la diferenciación.

Las células cancerosas incipientes deben evadir estas señales antiproliferativas si quieren prosperar. Gran parte de los circuitos que permiten que las células normales respondan a las señales anti-crecimiento están asociados con el reloj del ciclo celular,

específicamente los componentes que gobiernan el tránsito de la célula a través de la fase G1 de su ciclo de crecimiento.

Los genes supresores de tumores prototípicos RB1 y TP53, codifican las proteínas pRB (asociadas al retinoblastoma) y p53, los cuales operan como nodos de control central dentro de dos circuitos reguladores celulares complementarios clave que gobiernan las decisiones de las células para proliferar o, alternativamente, activar programas de senescencia o apoptosis.

La pRB integra señales de diversas fuentes extracelulares e intracelulares y, en respuesta, decide si una célula debe continuar con su ciclo de crecimiento y división (*Burkhardt and Sage, 2008*). La interrupción de la vía pRb libera factores de transcripción E2F que controlan la expresión de los genes esenciales para la progresión de la fase G1 a la S y, por lo tanto, permite la proliferación celular, lo que hace que las células sean insensibles a los factores anti-crecimiento que normalmente operan a lo largo de esta vía para bloquear el avance a través de la fase G1 del ciclo celular (*Weinberg, 1995*). Por tanto, las células cancerosas con defectos en la función de la vía RB carecen de los servicios de un guardián fundamental de la progresión del ciclo celular cuya ausencia permite la proliferación celular persistente. Mientras que pRB transduce señales inhibitoras del crecimiento que se originan en gran parte fuera de la célula, p53 recibe entradas de sensores de estrés y anomalías que funcionan dentro de los sistemas operativos intracelulares de la célula: si el grado de daño al genoma es excesivo o si los niveles de nucleótidos, las señales que promueven el crecimiento, la glucosa o la oxigenación son subóptimas, p53 puede detener la progresión del ciclo celular hasta que estas condiciones se hayan normalizado (*Hanahan and Weinberg, 2011*). Alternativamente, ante las señales de alarma que indican un daño abrumador o irreparable en dichos subsistemas celulares, p53 puede desencadenar la apoptosis. En particular, los diversos efectos de p53 activado son complejos y altamente dependientes del contexto, y varían según el tipo de célula, así como según la gravedad y persistencia de las condiciones de estrés celular y daño genómico.

Aunque los dos supresores canónicos de la proliferación, p53 y pRB, tienen una importancia preeminente en la regulación de la proliferación celular, varias líneas de evidencia indican que cada uno opera como parte de una red más grande que está determinada para poseer redundancia funcional.

En este sentido, la amplia red proteica que participa en el crecimiento celular y en los procesos involucrados en su limitación están estrechamente relacionados con las HSPs por lo que resulta razonable que las mismas constituyan elementos claves en la regulación y en el correcto funcionamiento de los mecanismos que actúan como frenos del ciclo celular.

Por ejemplo, *quan Wei et al. (1995)* observó que la administración de un RNA antisentido de hsp70 interrumpe la proliferación celular y el progreso a través de la

fase G1 y S en células cancerosas humanas. Estudios han demostrado que las fases G1 y S del ciclo parecen acortarse considerablemente por la sobreexpresión de Hsp70 (Helmbrecht *et al.*, 2000). Las Hsp70 pueden influir en la proliferación celular a través de su asociación con varias proteínas, tales como c-myc (Henriksson *et al.*, 1992), pRb (Inoue *et al.*, 1995) y p53 (Hainaut and Milner, 1992; Selkirk *et al.*, 1996). Un estudio reciente encontró que HSPA8 forma complejos con p27Kip1, un inhibidor de la quinasa dependiente de ciclina. Esta interacción parece apuntar a la degradación de p27Kip1 durante la fase de transición G1 / S del ciclo celular (Nakamura *et al.*, 1999).

Desde un punto de vista termodinámico, la estructura terciaria del supresor de tumores p53 es relativamente inestable (Brandt *et al.*, 2012). El proceso de plegado de la estructura de múltiples dominios de p53 es muy complejo y requiere de la formación de plegados intermedios que, en última instancia, brinda a p53 su estructura funcional final (Wawrzynow *et al.*, 2018).

La ganancia de función (GOF) de p53 mutante se demostró por primera vez en 1984, donde se demostró que la inducción de mut p53 transformaba células que carecen de p53 (Wolf *et al.*, 1984). Una característica crítica de los fenotipos GOF es la estabilización de mut p53, que se manifiesta por una gran acumulación de mut p53 en las células cancerosas (Brosh and Rotter, 2009).

La abundancia tanto de 'Wild Type' (WT) como de mut p53 está estrechamente controlada por la degradación mediada por la ruta del proteosoma de ubiquitina. En este proceso, se involucra la proteína MDM2, que controla los niveles de estado estacionario de p53 en células no estresadas al actuar como una ligasa de ubiquitina E3 de dominio de dedo RING (Honda and Yasuda, 2000), lo que resulta en la poliubiquitinación de p53 y su consiguiente degradación proteosomal. Por otro lado, WT p53 es responsable de regular al alza la expresión de MDM2, formando así un bucle de retroalimentación autorregulado que no existe en el caso de mut p53. En este proceso, se ha demostrado que los niveles elevados de HSP90 (Muller *et al.*, 2008; Alexandrova *et al.*, 2015) y HSP70 (Wiech *et al.*, 2012) son factores de gran relevancia que estabilizan mut p53. Por otro lado, también se ha demostrado anteriormente que en las células cancerosas, donde los niveles endógenos de HSPs ya están elevados, la maquinaria HSP reconoce que mut p53 es una proteína aberrante mal plegada, lo que provoca la formación de un complejo estable de múltiples chaperonas (Pinhasi-Kimhi *et al.*, 1986; King *et al.*, 2001; Tracz-Gaszewska *et al.*, 2017). La proteína mut p53 es reconocida por la maquinaria chaperona HSP70 / HSP40 / HSP90 donde HSP90 y HSP70, al unirse a mut p53, despliegan parcialmente este sustrato proteico y facilitan su plegamiento espontáneo a conformaciones con un mínimo de energía divergente.

Desafortunadamente, este mecanismo de control de calidad impulsado por las HSPs es adoptado por las células cancerosas (Mendillo *et al.*, 2012) que, al sobreexpresar en gran medida las HSPs, evitan las señales de inhibición del crecimiento, la inanición

metabólica, renuncian a la inducción de apoptosis y la senescencia y regulan y ajustan la red de señalización que impulsa el potencial proliferativo de la célula y la motilidad celular.

#### **1.2.4 Resistencia a la muerte celular**

La capacidad de las poblaciones de células tumorales para expandirse en número está determinada no solo por la tasa de proliferación celular sino también por la tasa de muerte celular. La muerte celular programada (apoptosis) representa una fuente importante de este desgaste. El programa apoptótico está presente en forma latente en prácticamente todos los tipos de células en todo el cuerpo. Una vez activado por una variedad de señales fisiológicas, este programa se desarrolla en una serie de pasos coreografiados con precisión. Las membranas celulares se rompen, los esqueletos citoplasmático y nuclear se descomponen, el citosol se extruye, los cromosomas se degradan y el núcleo se fragmenta. Al final, los restos celulares remanentes son fagocitados por las células cercanas y desaparecen, típicamente dentro de las 24 horas (*Hanahan and Weinberg, 2011*).

De acuerdo a *Hanahan and Weinberg (2011)*, la maquinaria apoptótica se puede dividir ampliamente en dos clases de componentes: sensores y efectores. Los sensores son responsables de monitorear el ambiente extracelular e intracelular en busca de condiciones de normalidad o anormalidad que influyan en si una célula debe vivir o morir. Estas señales regulan la segunda clase de componentes, que funcionan como efectores de muerte apoptótica. Los centinelas incluyen receptores de superficie celular que se unen a factores de supervivencia o muerte. Los ejemplos de estos pares ligando / receptor incluyen señales de supervivencia transmitidas por IGF-1 / IGF-2 a través de su receptor, IGF-1R, y por IL-3 y su receptor afín, IL-3R. Las señales de muerte son transmitidas por el ligando FAS que se une al receptor FAS y por el TNF  $\alpha$  que se une al TNF-R1 (*Hanahan and Weinberg, 2000*). Los sensores intracelulares monitorean el bienestar de la célula y activan la vía de la muerte en respuesta a la detección de anomalías, incluido el daño del ADN, el desequilibrio de señalización provocado por la acción de los oncogenes, la insuficiencia de factores de supervivencia o la hipoxia (*Evan and Littlewood, 1998*). Además, la vida de la mayoría de las células se mantiene en parte mediante señales de supervivencia basadas en la adherencia célula-matriz y célula-célula cuya abrogación provoca la apoptosis (*Ishizaki et al., 1995; Giancotti and Ruoslahti, 1999*). Las señales reguladoras apoptóticas solubles e inmovilizadas probablemente reflejan las necesidades de los tejidos de mantener sus células constituyentes en configuraciones arquitectónicas apropiadas.

Muchas de las señales que provocan la apoptosis convergen en las mitocondrias, que responden a las señales proapoptóticas liberando el citocromo c (cyt c), un potente

catalizador de la apoptosis (Lin, 1999). Los miembros de la familia de proteínas Bcl-2, cuyos miembros tienen función proapoptótica (Bax, Bak, Bid, Bim) o antiapoptótica (Bcl-2, Bcl-XL, Bcl-W), actúan en parte regulando la señalización de la muerte mitocondrial a través de liberación de citocromo C. La proteína supresora de tumores p53 puede provocar apoptosis regulando al alza la expresión de Bax proapoptótico en respuesta a la detección de daños en el ADN; Bax, a su vez, estimula a las mitocondrias para que liberen el citocromo C.

Los efectores finales de la apoptosis incluyen una serie de proteasas intracelulares denominadas caspasas (Thornberry and Lazebnik, 1998) donde particularmente dos caspasas “guardianas”, la 8 y la 9, son activadas por receptores de muerte como FAS o por el citocromo C liberado de las mitocondrias, respectivamente. Estas caspasas proximales desencadenan la activación de una docena o más de caspasas efectoras que ejecutan el programa de muerte, a través de la destrucción selectiva de estructuras subcelulares y orgánulos, y del genoma.

En este aspecto, las células cancerosas pueden adquirir resistencia a la apoptosis mediante una variedad de estrategias donde, sin dudas, la pérdida más común de un regulador proapoptótico por mutación involucra al gen supresor de tumores p53. La inactivación funcional resultante de su producto, la proteína p53, se observa en más del 50% de los cánceres humanos y da como resultado la eliminación de un componente clave del sensor de daño del ADN que puede inducir la cascada efectora apoptótica (Harris, 1996). Las señales evocadas por otras anomalías, incluida la hipoxia y la hiperexpresión de oncogenes, también se canalizan en parte a través de p53 a la maquinaria apoptótica; estos también están alterados para provocar la apoptosis cuando se pierde la función de p53 (Levine, 1997). Además, la vía PI3 quinasa-AKT / PKB, que transmite señales de supervivencia antiapoptóticas, se encuentra involucrada en la mitigación de la apoptosis en una fracción sustancial de los tumores humanos. La autofagia representa una importante respuesta fisiológica celular que, como la apoptosis, normalmente opera a niveles basales bajos en las células, pero puede ser fuertemente inducida en ciertos estados de estrés celular, el más obvio de los cuales es la deficiencia de nutrientes (Levine and Kroemer, 2008). El programa autofágico permite a las células descomponer proteínas y los orgánulos celulares, como los ribosomas y las mitocondrias, lo que permite que los catabolitos resultantes se reciclen y, por lo tanto, se utilicen para la biosíntesis y el metabolismo energético.

La inanición de nutrientes, la radioterapia y ciertos fármacos citotóxicos pueden inducir niveles elevados de HSPs y autofagia en respuesta al estrés, que redundan en un efecto citoprotector para las células cancerosas (Dokladny et al., 2015). Además, se ha demostrado que las células cancerosas severamente estresadas se contraen mediante autofagia a un estado de latencia reversible (White and DiPaola, 2009). Esta respuesta de supervivencia puede permitir la persistencia y el eventual recrecimiento de algunos

tumores en etapa tardía después del tratamiento con potentes agentes anticancerosos.

Las HSPs están implicadas en estos mecanismos donde, como se mencionó anteriormente, la Hsp70 puede unirse y estabilizar los alelos p53 mutantes, lo que les permite inhibir los alelos p53 de tipo salvaje e imitar los efectos de la mutación homocigótica y la pérdida de función (Zylicz *et al.*, 2001). Además, los niveles elevados de Hsp90, Hsp70 y Hsp27 pueden inhibir directamente varios pasos en las vías de la muerte celular programada (MCP) de forma directa (Budina-Kolomets *et al.*, 2014; Powers *et al.*, 2008; Takayama *et al.*, 2003). Por tanto, las HSPs refuerzan los efectos de la mutación de p53, además de conducir a la inhibición directa de la MCP. Las funciones de las HSPs en la inhibición de la MCP en el cáncer pueden, por tanto, reflejar sus actividades en la respuesta al choque térmico, inhibiendo la muerte celular para permitir la reparación del daño letal de las proteínas. Además, tanto Hsp70 como Bag3 pueden suprimir MCP corriente abajo del factor de transcripción HSF1 en el cáncer (Antonietti *et al.*, 2017). Es probable que la activación simultánea de la proliferación, como se describió anteriormente, combinada con la inhibición de la apoptosis, aumente la tasa de crecimiento tumoral. Donde es posible que, con tal marcado desbalance, las células tumorales mueran por vías relacionadas a la catástrofe mitótica y a la necrosis, procesos en los cuales las HSPs no están tan estudiadas (Weaver and Cleveland, 2005).

### **1.2.5 Potencial de replicación ilimitado**

Las células cancerosas requieren un potencial replicativo ilimitado para generar tumores macroscópicos. Esta capacidad contrasta notablemente con el comportamiento de las células en la mayoría de los linajes celulares normales del cuerpo, que son capaces de atravesar solo un número limitado de ciclos sucesivos de crecimiento y división celular. Esta limitación se ha asociado con dos barreras distintas a la proliferación: la senescencia, una entrada típicamente irreversible en un estado no proliferativo pero viable, y la muerte celular (Hanahan and Weinberg, 2011).

Múltiples líneas de evidencia indican que los telómeros que protegen los extremos de los cromosomas están involucrados de manera central en la capacidad de proliferación ilimitada (Blasco, 2005). Los telómeros, compuestos por múltiples repeticiones de hexanucleótidos en tándem, se acortan progresivamente en las células no inmortalizadas propagadas en cultivo, perdiendo finalmente la capacidad de proteger los extremos de los ADN cromosómicos de la degradación y de las fusiones de un extremo a otro; tales fusiones generan cromosomas dicéntricos inestables cuya resolución da como resultado una alteración del cariotipo que amenaza la viabilidad celular. En consecuencia, la longitud del ADN telomérico en una célula dicta cuántas generaciones sucesivas de células puede atravesar su progenie antes de que los telómeros se erosionen en gran medida y, en consecuencia, hayan perdido sus funciones protectoras, lo que

desencadena la entrada en crisis.

La telomerasa, una enzima ADN polimerasa especializada que agrega segmentos de repetición de telómeros a los extremos del ADN telomérico, está casi ausente en las células no immortalizadas, pero se expresa a niveles funcionalmente significativos en la gran mayoría de las células immortalizadas espontáneamente ( $\sim 90\%$ ), incluidas las células cancerosas humanas. Al extender el ADN telomérico, la telomerasa es capaz de contrarrestar la erosión progresiva de los telómeros que de otro modo ocurriría en su ausencia. La presencia de actividad de la telomerasa, ya sea en células immortalizadas espontáneamente o en el contexto de células diseñadas para expresar la enzima, se correlaciona con una resistencia a la inducción tanto de senescencia como de crisis / apoptosis; a la inversa, la supresión de la actividad de la telomerasa conduce al acortamiento de los telómeros y a la activación de una u otra de estas barreras proliferativas (*Hanahan and Weinberg, 2011*).

Los niveles elevados de HSPs tienen una influencia considerable en este proceso. Por ejemplo, se ha observado que Hsp82p, homólogo de Hsp90 en levaduras, facilita la unión de la telomerasa al ADN y favorece su función en levaduras (*Toogun et al., 2008*), así como también se ha determinado que la inmunofilina FKBP51, proteína fuertemente asociada a Hsp90, se encuentra sobreexpresada en múltiples líneas celulares tumorales, co-precipitan con hTERT y potencia la actividad de la telomerasa en conjunto con Hsp90, que facilitarían la translocación de FKBP51 al núcleo y su asociación con hTERT (*Lagadari et al., 2016; Zgajnar et al., 2019*). Además, Hsp70 y Hsp27 pueden bloquear las vías descendentes que van desde el acortamiento de los telómeros hasta la senescencia mediante mecanismos que incluyen la inhibición de la capacidad de WT p53 para iniciar la vía de la senescencia (*O'Callaghan-Sunol et al., 2007; Yaglom et al., 2007*). Hsp72 inhibe tanto la vía dependiente de p53 desencadenada por la expresión de PI3K oncogénica como la vía de senescencia independiente de p53 inducida por la expresión del oncogén Ras (*Gabai et al., 2009*). De hecho, en el modelo HER2 / neu de carcinogénesis mamaria espontánea de ratón, la inactivación de Hsp72 por recombinación homóloga previno la formación de tumores en el 100% de los ratones debido al inicio de la senescencia (*Meng et al., 2011*).

### **1.2.6 Angiogénesis sostenida**

El oxígeno y los nutrientes suministrados por la vasculatura son cruciales para la función y la supervivencia de las células, lo que obliga a prácticamente todas las células de un tejido a residir dentro de los 100  $\mu\text{m}$  de un vaso sanguíneo capilar. Una vez que los tumores crecen y comienzan a invadir sus tejidos de origen, se alejan de la microcirculación. De hecho, se observa hipoxia pronunciada en muchos tumores macroscópicos (*Brown, 2007*). Sin embargo, la mayoría de los tumores pueden ensamblar una especie

de microcirculación mediante angiogénesis *de novo*, a través de la detección de hipoxia por los factores de transcripción de la familia del factor inducible por hipoxia (HIF), y la síntesis y liberación de citocinas angiogénicas, como VEGF, que inducen la formación de nuevos vasos sanguíneos estimulando la proliferación de células endoteliales (*Semenza, 2002; Eichhorn et al., 2007*).

Durante la organogénesis, el suministro de oxígeno está asegurado por el crecimiento coordinado de vasos y parénquima. Una vez que se forma un tejido, el crecimiento de nuevos vasos sanguíneos, el proceso de angiogénesis, es transitorio y está cuidadosamente regulado. Por el contrario, durante la progresión del tumor, un “interruptor angiogénico” casi siempre se activa y permanece encendido, lo que hace que la vasculatura, normalmente inactiva, brote continuamente nuevos vasos que ayudan a sostener los crecimientos neoplásicos en expansión (*Hanahan and Folkman, 1996*).

Resulta de interés mencionar que la señalización de VEGF está mediada a través de tres receptores tirosina quinasa (VEGFR-1–3) que a su vez, está regulada en múltiples niveles, lo que lleva a que la expresión del gen VEGF pueda regularse positivamente tanto por hipoxia como por señalización oncogénica (*Carmeliet, 2005*). La capacidad para inducir y mantener la angiogénesis parece adquirirse en un paso (o pasos) discretos durante el desarrollo del tumor, a través de un “cambio angiogénico” de la quiescencia vascular. HSF1 y las HSPs Hsp27, Hsp70 y Hsp90 pueden contribuir a la angiogénesis a través del aumento de la expresión de HIF1, el acompañamiento de los receptores de VEGF y el aumento de la potencia de la señalización de VEGF (*Colvin et al., 2014; Thuringer et al., 2013; Shiota et al., 2010; Bruns et al., 2012*) por lo que las mismas resultan determinantes ya sea, en relación a la resistencia celular ante la hipoxia y el estrés celular como durante la inducción de la angiogénesis y la regulación del “cambio angiogénico”.

### **1.2.7 Invasión de tejidos y metástasis**

Tarde o temprano, durante el desarrollo de la mayoría de los tipos de cáncer humano, las masas de tumores primarios generan células pioneras que se mueven, invaden los tejidos adyacentes y, de ahí, viajan a sitios distantes donde pueden tener éxito en la fundación de nuevas colonias. Estos asentamientos distantes de células tumorales (metástasis) son la causa del 90% de las muertes por cáncer en humanos (*Sporn, 1996*). La capacidad de invasión y metástasis permite a las células cancerosas escapar de la masa tumoral primaria y colonizar un nuevo terreno en el cuerpo donde, al menos inicialmente, los nutrientes y el espacio no son limitantes. Las metástasis recién formadas entonces surgen como amalgamas de células cancerosas y células de soporte normales reclutadas del tejido huésped.

El proceso de múltiples pasos de invasión y metástasis se ha esquematizado como

una secuencia de pasos discretos, a menudo denominada cascada de invasión-metástasis (*Talmadge and Fidler, 2010*). Esta representación visualiza una sucesión de cambios biológicos celulares, comenzando con la invasión local, luego la intravasación de las células cancerosas en los vasos sanguíneos y linfáticos cercanos, el tránsito de las células cancerosas a través de los sistemas linfático y hematógeno, seguido por el escape de las células cancerosas de la luz de tales vasos en el parénquima de tejidos distantes (extravasación), la formación de pequeños nódulos de células cancerosas (micrometástasis) y finalmente el crecimiento de lesiones micrometastásicas en tumores macroscópicos. Las células epiteliales están estrechamente unidas en su ubicación apropiada a través de una serie de moléculas de unión y adhesión para formar tejidos estables y altamente organizados (*Puliafito et al., 2017*). Sin embargo, en alguna etapa de la transformación, las células transformadas se vuelven móviles, capaces de migrar a través del tejido de origen e invadir las estructuras circundantes (*Brabletz, 2012*). Estas propiedades parecen implicar un cambio en el fenotipo, conocido como transición epitelio-mesenquimal Transición Epitelio Mesenquimal (EMT), en el que una célula epitelial puede perder su morfología columnar característicamente cuboide y puede adoptar una apariencia más fusiforme, común a las células de un origen mesenquimatoso (*Voutsadakis, 2016*). La EMT subyacente es un cambio en la expresión de genes marcadores epiteliales como E-cadherina, claudina y beta-catenina de superficie a marcadores mesenquimales, más típicamente vimentina, indicadores de un cambio transcripcional orquestado por una serie de factores de transcripción que pueden incluir SNAIL, SLUG, TWIST o ZEB1, según el tipo de célula y el contexto (*Brabletz, 2012*). La activación de un programa EMT puede conferir a las células tumorales propiedades relativamente móviles y resistentes a la MCP que, en última instancia, aumentan su potencial para progresar a través del órgano de origen y permite reemplazar los tejidos normales diferenciados. Este proceso suele ir acompañado de la secreción de las células del estroma tumoral, como los fibroblastos asociados a cáncer (FAC), de factores de crecimiento y proteínas de la matriz extra-celular (ECM) que incluyen colágenos fibrilares, que forman “vías de metástasis” en el tejido. La producción de enzimas remodeladoras de ECM, como MMP2, por las propias células tumorales, o los FAC, permite luego el paso de las células a través de la ECM y fuera del tejido local (*Gascard and Tlsty, 2016; Ma et al., 2019*). Se ha demostrado que Hsp70 es esencial para la invasión y metástasis en los tumores mamarios de ratón (TMR) y en líneas mamarias humanas, lo que sugiere un papel fundamental para la chaperona en este proceso (*Gong et al., 2015; Kasioumi et al., 2019*). Estos eventos parecen implicar la expresión del oncogén c-MET, así como su ligando HGF, ya que el fosfo-MET, la forma activada de c-MET, estaba ausente de los tumores TMR deficientes en Hsp72. Asimismo, se requirió una elevación de Hsp27 para la invasión y metástasis desencadenadas por HGF, y presumiblemente la activación de MET aguas abajo (*Musiani et al., 2014; Pavan et al., 2014*), así como en otros mecan-

ismos relacionados (*Gibert et al.*, 2012). c-MET es un oncogén potente en muchos tipos de células, y su dependencia de la expresión de HSPs para la activación apunta a uno de los mecanismos subyacentes a las funciones de las HSPs en el cáncer (*Organ and Tsao*, 2011). También se demostró que Hsp27 participa en el procesamiento de la metaloproteínasa 9 de la matriz (MMP9), un mediador clave de la invasión (*Choi et al.*, 2014a). La Hsp27 parece tener una importancia clave en varios pasos de la cascada metastásica, incluida la EMT, la metástasis y las células tumorales circulantes en el cáncer de próstata, lo que sugiere en general un papel versátil de las chaperonas en la diseminación secundaria del cáncer (*Shiota et al.*, 2013). La perturbación de Hsp90 o Hsp70 interrumpe proteínas clave para la migración y la invasión, incluidas FAK y WASF3 (*Budina-Kolomets et al.*, 2016; *Liu et al.*, 2011; *Teng et al.*, 2012). Las funciones de Hsp90 en las propiedades metastásicas autónomas de las células tumorales, como la migración y la invasión, están mediadas por el acompañamiento de una serie de dianas, más a menudo tirosina quinasas, incluidas Src y FAK (*Liu et al.*, 2011; *Teng et al.*, 2012; *Tsutsumi et al.*, 2009; *Xiong et al.*, 2019).

Hasta ahora, el papel descrito de las HSPs es, por supuesto, como chaperonas moleculares intracelulares, y hemos enumerado una serie de moléculas oncogénicas dependientes de la sobreexpresión de HSPs, tanto en células tumorales como estromales. Sin embargo, está cada vez más claro que las HSPs escapan de la célula en forma libre y en exosomas para interactuar con los receptores en sí mismos, en células adyacentes del mismo linaje o en diferentes tipos de células (*Lang et al.*, 2019). De forma similar a muchas HSPs intracelulares, los niveles de producción extracelular de varias HSPs aumentan por las células cancerosas de diversos orígenes y, en algunos casos, se han relacionado con mayores grados de carga de enfermedad (*Lang et al.*, 2019). Una cuestión adicional clave para comprender las funciones de las HSPs en el cáncer es la importancia relativa de las HSPs intracelulares y extracelulares (*Li et al.*, 2013; *Calderwood*, 2018). Cada una de las familias de genes HSPs contiene múltiples parálogos, algunos de los cuales podrían estar especializados para funciones extracelulares. En el caso de Hsp90, que existe como dos parálogos diferentes expresados en el citosol, Hsp90-alfa (HSP90AA1) y Hsp90-beta (HSP90AB1), el primero parece especializado para la secreción extracelular, mientras que el segundo generalmente se retiene en las células y parece más importante en el plegamiento y limpieza de clientes dentro de la célula (*Li et al.*, 2013). La Hsp90 secretada juega un papel importante en la invasividad de las células cancerosas mediante la unión a los receptores de superficie como LRP1 / CD91 y la interacción con la metaloproteasa 2 (MMP2) de la matriz en la superficie celular para mediar la invasividad, la EMT y un fenotipo iniciador de tumor (*Lang et al.*, 2019). Curiosamente, los inhibidores de Hsp90 impermeables a las células pueden bloquear estos procesos (*Tsutsumi et al.*, 2008).

### 1.2.8 Inestabilidad genómica

La adquisición de los múltiples sellos enumerados anteriormente depende en gran parte de una sucesión de alteraciones en los genomas de las células neoplásicas. Representados de manera simple, ciertos genotipos mutantes confieren una ventaja selectiva a los subclones de células, lo que permite su desarrollo y eventual dominio en un entorno de tejido local. En consecuencia, la progresión del tumor en varios pasos puede describirse como una sucesión de expansiones clonales, cada una de las cuales se desencadena por la adquisición fortuita de un genotipo mutante habilitante.

La extraordinaria capacidad de los sistemas de mantenimiento del genoma para detectar y resolver defectos en el ADN asegura que las tasas de mutación espontánea sean muy bajas durante cada generación celular. En el curso de la adquisición de los genes mutantes necesarios para orquestar la tumorigénesis, las células cancerosas a menudo aumentan las tasas de mutación (*Hanahan and Weinberg, 2011*). Esta mutabilidad se logra mediante una mayor sensibilidad a los agentes mutagénicos, mediante la ruptura de uno o varios componentes de la maquinaria de mantenimiento genómico, o ambos. Además, la acumulación de mutaciones puede acelerarse comprometiendo los sistemas de vigilancia que normalmente controlan la integridad genómica y fuerzan a las células dañadas genéticamente a la senescencia o la apoptosis (*Jackson and Bartek, 2009*).

Una gama diversa de defectos que afectan a varios componentes de la maquinaria de mantenimiento del ADN, a menudo denominados “cuidadores” del genoma (*Kinzler and Vogelstein, 1997*), se han documentado. El catálogo de defectos en estos genes cuidadores incluye aquellos cuyos productos están involucrados en i) detectar daños en el ADN y activar la maquinaria de reparación, ii) reparar directamente el ADN dañado e iii) inactivar o interceptar moléculas mutagénicas antes de que hayan dañado el ADN (*Negrini et al., 2010; Ciccia and Elledge, 2010; Jackson and Bartek, 2009; Kastan, 2008*).

Aunque los aspectos específicos de la alteración del genoma varían drásticamente entre los diferentes tipos de tumores, la gran cantidad de defectos de reparación y mantenimiento del genoma que ya se han documentado en los tumores humanos, junto con la abundante evidencia de una desestabilización generalizada del número de copias de genes y la secuencia de nucleótidos, nos persuaden de que la inestabilidad del genoma es inherente a la gran mayoría de las células cancerosas humanas. Esto lleva, a su vez, a la conclusión de que los defectos en el mantenimiento y reparación del genoma son selectivamente ventajosos y, por lo tanto, instrumentales para la progresión tumoral, aunque solo sea porque aceleran la velocidad a la que las células premalignas en evolución pueden acumular genotipos favorables. Como tal, la inestabilidad del genoma es claramente una característica habilitadora que está asociada causalmente con la adquisición de capacidades distintivas.

La inestabilidad del genoma puede producir la acumulación de copias adicionales de ADN, así como otras alteraciones importantes como translocaciones cromosómicas, inversiones, deleciones, roturas de cadena simple y doble en el ADN y otros cambios anormales en la estructura terciaria del ADN que pueden causar la expresión errónea de genes. La respuesta al daño del ADN (DDR) comprende rutas coordinadas desde la detección de daños en el ADN hasta la detención del ciclo celular, la reparación del ADN o incluso la apoptosis (*Jackson and Bartek, 2009*). El término DDR comprende vías complejas que implican la detección del daño, la activación de las redes de señalización y el punto de control del ciclo celular, y la reparación del ADN o la inducción de la muerte celular (*Ciccia and Elledge, 2010*). Generalmente se distinguen cinco mecanismos de reparación del ADN: (a) reversión directa del daño del ADN, (b) reparación de escisión de bases (*base excision repair* o BER), (c) reparación de errores de apareamiento (MMR), (d) reparación por escisión de nucleótidos (*nucleotide excision repair* o NER) y (e) recombinación homóloga (*homologous recombination* o HR) y Unión de extremos no homólogos (*non-homologous end joining* o NHEJ). En este sentido, es importante destacar que las HSPs han sido asociadas de forma directa a cada uno de estos mecanismos (*Sottile and Nadin, 2018*) entre lo que destacamos:

- La mutación p53 ha sido identificada como un actor clave del contexto de estabilidad genómica en el cáncer. Se ha demostrado que tanto p53 WT como mutado comparten vías de promoción de tumores con HSF1, ya sea evitando la senescencia inducida por tumores en tumores con p53 WT o alimentando un circuito de retroalimentación positiva donde mut p53 estimula las vías MAPK y PI3K – AKT– mTOR, induciendo la fosforilación de HSF1 en S326 y también interactuando directamente con HSF1 activado aumentando su afinidad por la cromatina, lo que permite una expresión más eficiente de HSPs, lo que facilita que las células se adapten al estrés proteotóxico (*Toma-Jonik et al., 2019*).
- Las HSPs no solo participan en el correcto plegado de la proteína p53 y en sus vías de señalización efectoras, sino también se ha demostrado que Hsp90, en complejo con la inmunofilina FKPB52, actúan de mediadores en el transporte de p53 al núcleo, donde la p53 lleva a cabo la mayoría de sus funciones guardianas (*Galigniana et al., 2004*).
- La interacción entre CHK1 y Hsp90: Esta fue demostrada mediante inmunoprecipitación. A su vez, se observó que el inhibidor de Hsp90, 17-AAG induce el agotamiento de CHK1 y sensibiliza las células cancerosas a la gemcitabina (*Arlander et al., 2003*).
- Dentro de la familia de las HSPs pequeñas, HSPB1 está vinculada con ATM, donde favorece la reparación de rupturas de doble cadena (*Cosentino et al., 2011*;

*Guttmann et al., 2013).*

- Se ha demostrado que ATR (sensor de daño de ADN), es una proteína cliente de Hsp90, la cual facilita el reclutamiento de ATR, CHK1 y 53BP1 a los sitios de daño del ADN, por lo que facilita la respuesta de reparación del ADN (*Ha et al., 2011*).
- La metiltransferasa MGMT, capaz de detectar y corregir el daño por alquilación del ADN se asocia fuertemente con Hsp90 (HSP90AA1 y HSP90AB1) (*Tubbs and Nussenzweig, 2017*).
- Se ha observado que HSPA1A se une a la endonucleasa APEX1 (HAP1 o APE1) y potencia su actividad endonucleasa específica de 10 a 100 veces (*Kenny et al., 2001*). Por otro lado, estimula la reparación mediada por Pol $\beta$  (*Mendez et al., 2003, 2000*) y se asocia con el complejo PARP1/XRCC1 en los dominios de unión a BRCA1, protegiendo ambas proteínas y permitiendo la reparación de rupturas de simple cadena del ADN (*Kotoglou et al., 2009*).
- HSPB1 y MSH2 aumentan y co-localizan en células de glioma expuestas a temozolomida (*Castro et al., 2015*). Además, MLH1 y MSH2 interactúan con HSPB1 y HSPA1A en células de cáncer de colon humano (*Sottile et al., 2015*).
- HSPB1 está involucrada en la escisión del daño del ADN inducido por la luz ultravioleta donde la inhibición de HSPB1 suprime el proceso de reparación del daño del ADN por UVc (*Wano et al., 2004*). También se ha demostrado que HSPB1 participa en este proceso al mantener las proteínas reparadoras de NER en su estado correctamente plegado en *E. coli* (*Zou et al., 1998*). En esta misma dirección, HSPA1A también ha sido asociada a la reparación NER en células de adenocarcinoma de pulmón (*Niu et al., 2006*).
- HSPA1A también parece estar involucrado en el proceso de reparación del ben-zopireno, donde se ha visto que co-localiza con las proteínas reparadoras XPA y XPG, (*Xiao et al., 2002; Yang et al., 2009; Duan et al., 2014*).
- Numerosas proteínas involucradas en la reparación de rupturas de doble cadena han sido descritos como clientes de, o que interactúan con, proteínas HSPs como por ej.: DNA-PK (*Solier et al., 2012; Choi et al., 2014a*), ku80 (*Katsogiannou et al., 2014b*), RAD51 (*Yao et al., 2007*), BRCA1 (*Stecklein et al., 2012*), BRCA2 (*Noguchi et al., 2006*), CHK1, complejo MRE11 / RAD50 / NBS1 (MRN) (*Dote et al., 2006*), DCLRE1B (Apollo o hSNM1B) (*Schmiester and Demuth, 2017; Anders et al., 2009*).

- Diferentes inhibidores de Hsp90, como NVP-AUY922, NVP-BEP800 y 17-DMAG, solos o en combinación con radiación ionizante (RI), conducen a una radiosensibilización de las líneas celulares tumorales. La inhibición de HSP90AA1 aumenta las rupturas de doble cadena inducidas por RI y provoca un retraso en la reparación del daño del ADN (*Stingl et al., 2010*).

Por toda la evidencia recogida, se hace evidente que las HSPs son cruciales para estabilizar los componentes centrales de los mecanismos de reparación del ADN. Los niveles de expresión alterados de las HSPs podrían conducir a una detección deficiente, así como a una reparación prematura del ADN dañado.

Por otro lado, la respuesta al estrés térmico (HSR) mediada principalmente por HSF1 está profundamente integrada con los sistemas de reparación de daños en el ADN y la transcripción de HSPs mediada por HSF1 permite el funcionamiento correcto de las proteínas DDR y los oncogenes mutados que de otro modo serían no funcionales, albergando un importante papel protector hacia las células de cáncer hipermutantes (*Sottile and Nadin, 2018*).

### **1.2.9 Inflamación promotora de tumores**

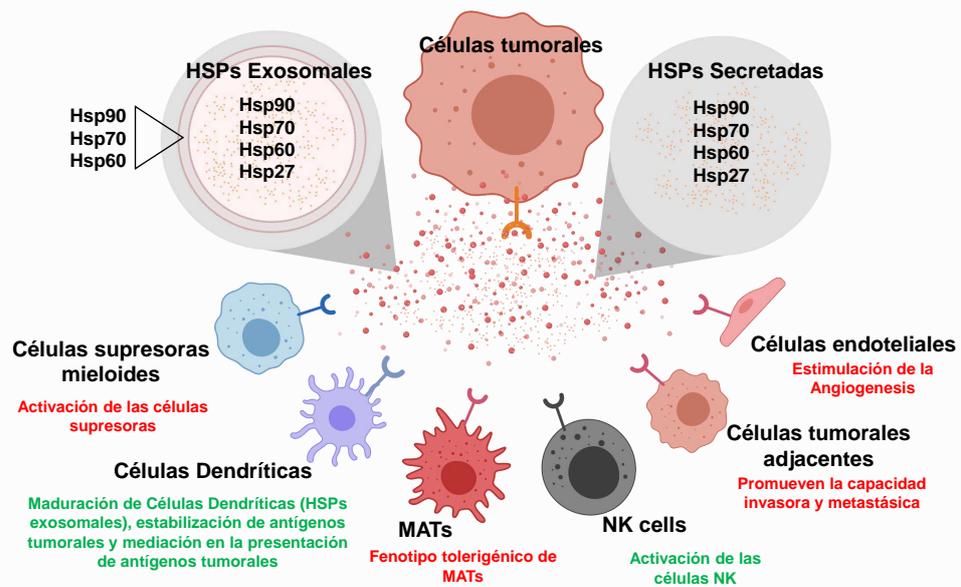
Los patólogos han reconocido desde hace mucho tiempo que algunos tumores están densamente infiltrados por células de los sistemas innato y adaptativo del sistema inmunológico y, por lo tanto, reflejan las condiciones inflamatorias que surgen en tejidos no neoplásicos (*Flier et al., 1986*). Con el advenimiento de mejores marcadores para identificar con precisión los distintos tipos de células del sistema inmunológico, ahora está claro que prácticamente todas las lesiones neoplásicas contienen células inmunes presentes en densidades que van desde infiltraciones sutiles detectables solo con anticuerpos específicos del tipo celular hasta inflamaciones graves que son aparentes incluso mediante técnicas de tinción histoquímica estándar (*Pagès et al., 2010*).

Esta respuesta inmunitaria resultó paradójica durante un largo tiempo, donde por un lado, observábamos que existía una respuesta inmune del organismo a la nueva progresión tumoral, donde el sistema inmunológico genera presión sobre el tumor, que a su vez evoluciona para evadir la destrucción inmunitaria, pero por otro lado, se observaba que en muchos casos, la respuesta inflamatoria local tenía el efecto paradójico de mejorar la tumorigénesis y la progresión. La inflamación puede contribuir a múltiples capacidades distintivas al suministrar moléculas bioactivas al microambiente tumoral, incluidos factores de crecimiento que mantienen la señalización proliferativa, factores de supervivencia que limitan la muerte celular, factores proangiogénicos, enzimas modificadoras de la matriz extracelular que facilitan la angiogénesis, la invasión y la metástasis, y señales que conducen a la activación de EMT y otros programas facilitadores de distintos (*Grivennikov et al., 2010; Qian and Pollard, 2010*).

Es importante destacar que, en algunos casos, la inflamación es evidente en las primeras etapas de la progresión neoplásica y es demostrablemente capaz de fomentar el desarrollo de neoplasias incipientes en cánceres en toda regla (*Qian and Pollard, 2010*). Además, las células inflamatorias pueden liberar sustancias químicas, en particular especies reactivas de oxígeno, altamente tóxicas a nivel proteico y activamente mutagénicas para las células cancerosas cercanas, acelerando su evolución genética hacia estados de malignidad intensificada (*Grivennikov et al., 2010*). Como tal, la inflamación puede considerarse una característica habilitadora por sus contribuciones a la adquisición de capacidades distintivas centrales.

Las propiedades tumorigénicas de diferentes miembros extracelulares de HSPs han sido implicados en la inmunomodulación. Por ejemplo, se encontró que la Hsp72 exosómica promueve la activación de células supresoras mieloides a través de la unión al receptor tipo toll-2 (TLR2), estimulando la liberación de IL-6 y la activación subsiguiente de STAT3, promoviendo la invasión y migración de células de carcinoma y la angiogénesis (*Chalmin et al., 2010*).

Alternativamente, las HSPs extracelulares también pueden promover la inmunidad tumoral (Figura 1.4) estabilizando los antígenos tumorales, estimulando los MAT para que secreten citocinas inflamatorias y mediar en el procesamiento de antígenos por las células presentadoras de antígeno (APC) (*Calderwood et al., 2005*).



**Fig. 1.4 Las células tumorales pueden modular el microambiente tumoral secretando HSPs.** Se ha identificado que las células de carcinoma derivadas de varios tejidos secretan HSPs en el espacio extracelular en forma soluble o dentro de exosomas. Las HSPs secretadas de las células tumorales pueden modular la biología y las funciones de otras células en el microambiente tumoral. Específicamente, se ha demostrado que las HSPs secretadas modulan las actividades de las células supresoras mieloides, CD, MAT, células asesinas naturales (NK), otras células de carcinoma vecinas y células endoteliales. Las HSPs extracelulares pueden promover procesos tumorigénicos (enumerados en texto rojo). Alternativamente, las HSPs extracelulares también pueden promover la inmunidad tumoral (enumeradas en texto verde) al estabilizar los antígenos tumorales, estimular a los MAT para que secreten citocinas inflamatorias y mediar el procesamiento de antígenos por las células presentadoras de antígenos. Modificado de *Lang et al.* (2019), creado con <https://biorender.com/>.

En esta dirección se observó que la Hsp27 está presente en niveles elevados en el líquido intersticial de los tumores de mama humanos, y se demostró que la diferenciación de monocitos humanos en macrófagos en presencia de Hsp27 soluble induce propiedades tolerogénica y proangiogénica en los MAT (*Banerjee et al.*, 2011).

Hsp27 también puede desempeñar un papel importante en la programación del entorno extracelular mediante la activación de la liberación de IL1-beta, TNF-alfa, IL-10, PGE2 y VEGF-A (*Lang et al.*, 2019). Por tanto, la Hsp27 secretada puede influir en la inmunidad tumoral y la angiogénesis, así como en las funciones relacionadas a metástasis mencionadas anteriormente.

### 1.2.10 Reprogramación del metabolismo energético

La proliferación celular crónica y a menudo incontrolada que representa la esencia de la enfermedad neoplásica implica no solo el control desregulado de la proliferación celular, sino también los correspondientes ajustes del metabolismo energético con el fin

de impulsar el crecimiento y la división de las células. En condiciones aeróbicas, las células normales procesan la glucosa, primero a piruvato a través de la glucólisis en el citosol y luego en dióxido de carbono en las mitocondrias; en condiciones anaeróbicas, se favorece la glucólisis y se envía relativamente poco piruvato a las mitocondrias que consumen oxígeno. Otto Warburg observó por primera vez una característica anómala del metabolismo energético de las células cancerosas (*Weinhouse et al.*, 1956; *Hsu and Sabatini*, 2008) : incluso en presencia de oxígeno, las células cancerosas pueden reprogramar su metabolismo de la glucosa y, por lo tanto, su producción de energía, al limitar su metabolismo energético en gran parte a la glucólisis, lo que lleva a un estado que se ha denominado “glucólisis aeróbica”.

La existencia de este cambio metabólico en las células cancerosas se ha comprobado en las décadas siguientes. Tal reprogramación del metabolismo energético es aparentemente contradictorio, ya que las células cancerosas deben compensar la eficiencia ~ 18 veces menor de producción de ATP que proporciona la glucólisis en relación con la fosforilación oxidativa mitocondrial. Lo hacen en parte regulando al alza los transportadores de glucosa, en particular GLUT1 , que aumenta sustancialmente la importación de glucosa al citoplasma (*Jones and Thompson*, 2009; *DeBerardinis et al.*, 2008; *Hsu and Sabatini*, 2008).

En este sentido, se ha demostrado que la alimentación glicolítica está asociada con oncogenes activados (p. Ej., RAS , MYC ) y supresores de tumores mutantes (p. Ej., TP53 ) (*DeBerardinis et al.*, 2008; *Jones and Thompson*, 2009), cuyas alteraciones en las células tumorales se han seleccionado principalmente por sus beneficios al conferir las capacidades distintivas de la proliferación celular, evitar los controles citostáticos y atenuar la apoptosis. Esta dependencia de la glucólisis puede acentuarse aún más en las condiciones hipóxicas que operan dentro de muchos tumores: el sistema de respuesta a la hipoxia actúa pleiotrópicamente para regular al alza los transportadores de glucosa y múltiples enzimas de la vía glucolítica (*Semenza*, 2010; *Jones and Thompson*, 2009; *DeBerardinis et al.*, 2008). Por tanto, tanto la oncoproteína Ras como la hipoxia pueden aumentar de forma independiente los niveles de los factores de transcripción HIF1 $\alpha$  y HIF2 $\alpha$ , que a su vez regulan al alza la glucólisis (*Semenza*, 2010). El aumento de la glucólisis permite la desviación de intermedios glucolíticos hacia diversas vías biosintéticas, incluidas las que generan nucleósidos y aminoácidos; esto facilita, a su vez, la biosíntesis de las macromoléculas y orgánulos necesarios para el ensamblaje de las nuevas células que resultan de una proliferación celular activa.

El metabolismo energético alterado está demostrando estar tan extendido en las células cancerosas como muchos de los otros rasgos asociados al cáncer que se han aceptado como distintivos del cáncer. De hecho, la redirección del metabolismo energético está orquestada en gran medida por proteínas que están involucradas de una forma u otra en la programación de las características principales del cáncer. Cuando

se ve de esta manera, la glucólisis aeróbica es simplemente otro fenotipo que está programado por oncogenes que inducen la proliferación.

En este sentido, las HSPs, como se ha mencionado previamente, cumplen un rol fundamental en la adaptación de las células a un medio ácido e hipóxico, que altera notablemente el paisaje energético de plegamiento proteico. Independientemente de este factor preponderante, también se ha demostrado que niveles elevados tanto de Hsp90 como de Hsp70 reducen la fosforilación oxidativa y aumentan la glucólisis aeróbica en las células tumorales, lo que implica un rol potenciador de las HSPs en relación a este metabolismo (Xu *et al.*, 2017; Wang *et al.*, 2012).

### **1.2.11 Evasión de la destrucción inmunológica**

El cáncer se caracteriza por la acumulación de un número variable de alteraciones genéticas y la pérdida de los procesos reguladores celulares normales (Hanahan and Weinberg, 2011). Se sabe desde hace mucho tiempo que estos eventos dan como resultado la expresión de neoantígenos o antígenos de diferenciación, que pueden conducir a la presentación de péptidos unidos a moléculas del complejo mayor de histocompatibilidad clase I (MHCI) en la superficie de las células cancerosas, distinguiéndolas de sus contrapartes normales. Desde el trabajo de Boon *et al.* (1994), hemos sabido que estos complejos de péptido-MHCI específicos del cáncer pueden ser reconocidos por células T CD8 + reclutadas en los tumores de pacientes con cáncer (Boon *et al.*, 1994). Sin embargo, incluso cuando se produjeron respuestas de células T, raras veces proporcionan inmunidad protectora donde la eliminación continua de células cancerosas que expresan dianas de células T (edición inmunitaria) puede permitir la selección clonal de células con la capacidad de evitar los ataques de parte de las células inmunes (Hanahan and Weinberg, 2011). A pesar de estos hallazgos, los resultados recientes del cáncer humano han demostrado que la superación de los reguladores negativos de las respuestas de las células T en los órganos linfoides (puntos de control) y en el lecho tumoral (función del inmunostato) probablemente explique el fracaso de la protección inmunitaria en muchos pacientes (Mullard, 2013). Los factores en el microambiente del tumor pueden actuar para modular la respuesta inmune de las células T antitumorales activadas existentes, actuando como un reóstato inmune o “inmunostato”. Esta clase de moléculas, incluidas las interactuantes PD-L1 y PD -1 (Chen *et al.*, 2012; Topalian *et al.*, 2012), enfatiza que la respuesta inmune en el cáncer refleja una serie de eventos cuidadosamente regulados.

Para que una respuesta inmunitaria contra el cáncer conduzca a la destrucción eficaz de las células cancerosas, se debe iniciar una serie de eventos escalonados y permitir que prosiga y se expanda de forma iterativa. Lo que se ha referido como el ciclo de inmunidad al cáncer (Chen and Snyder, 2013). En el primer paso, los neoantígenos crea-

dos durante la oncogénesis son liberados y capturados por CD para su procesamiento (paso 1). Para que este paso produzca una respuesta de células T contra el cáncer, debe ir acompañado de señales que especifiquen la inmunidad para que no se induzca la tolerancia periférica a los antígenos tumorales. Tales señales inmunogénicas pueden incluir citocinas pro inflamatorias y factores liberados por células tumorales moribundas. A continuación, las CD presentan los antígenos capturados en las moléculas MHCI y MHCII a las células T (paso 2), lo que da como resultado el cebado y la activación de las respuestas de las células T efectoras contra los antígenos específicos del cáncer (paso 3) que se consideran extraños o contra los cuales la tolerancia ha sido incompleta. La naturaleza de la respuesta inmune se determina en esta etapa, siendo clave para el resultado final un equilibrio crítico que representa la proporción de células efectoras T versus células reguladoras T. Finalmente, las células T efectoras activadas se trasladan (paso 4) e infiltran el lecho tumoral (paso 5), reconocen específicamente y se unen a las células cancerosas a través de la interacción entre su receptor de células T (RCT) y su antígeno afín unido a MHCI (paso 6) y mata la célula cancerosa diana (paso 7). La muerte de la célula cancerosa libera antígenos adicionales asociados al tumor (paso 1 nuevamente) para aumentar la amplitud y profundidad de la respuesta en las revoluciones posteriores del ciclo. En pacientes con cáncer, el ciclo de inmunidad al cáncer no funciona de manera óptima. Es posible que no se detecten los antígenos tumorales, las CD y las células T pueden tratar los antígenos como propios en lugar de extraños, creando así respuestas de células reguladoras T en lugar de respuestas efectoras, las células T pueden no vivir adecuadamente en los tumores, pueden no lograr infiltrarse en el tumor o (quizás lo más importante) factores en el microambiente tumoral pueden suprimir las células T efectoras que intentan destruir las células tumorales (*Motz and Coukos, 2013*).

Las HSPs desempeñan funciones poderosas, aunque a menudo ambiguas, en la inmunidad tumoral (*Calderwood et al., 2016*). Debido a su capacidad para unirse a un amplio espectro de secuencias de péptidos, se han utilizado para preparar vacunas eficaces contra el cáncer, en las que los antígenos tumorales se unen a las HSPs, que luego pueden dirigirlos para su procesamiento por APC, como las CD (*Murshid et al., 2010; Gao et al., 2008; Ciocca et al., 2012, 2007*). Sin embargo, no está claro si las HSPs intrínsecas de las células tumorales son característicamente inmunogénicas o inmunosupresoras. TGF- $\beta$  y CXCL12 parecen modular las poblaciones relativas de linfocitos que se infiltran en el tumor y sus respectivas actividades; esto ha sido demostrado por estudios que muestran el bloqueo de TGF- $\beta$  o CXCL12 para potenciar las respuestas tumorales a bloqueos de puntos de control inmunes en múltiples modelos tumorales (*Chen et al., 2019; Sharma and Allison, 2015; Mariathasan et al., 2018; Sow et al., 2019*). Por lo tanto, se ha propuesto la hipótesis de que la activación del mecanismo mediado por HSF1, TGF- $\beta$  y CXCL12 suprime la inmunidad antitumoral

y podría suponer un blanco terapéutico en el cáncer (*Lang et al.*, 2019). Además, se sabe que las HSPs pueden liberarse de las células tumorales y pueden influir en el comportamiento de las células adyacentes (*Mambula and Calderwood*, 2006). Gran parte de la Hsp70 o Hsp90 liberada por las células tumorales está contenida en partículas exosomales (*Eguchi et al.*, 2018), y se ha demostrado que los exosomas que contienen Hsp72 pueden producir inmunosupresión (*Chalmin et al.*, 2010; *Diao et al.*, 2015). Por lo tanto, la mayoría de la evidencia apunta a un papel inmunorregulador de las HSPs liberadas desde tumores y células estromales en condiciones basales. Sin embargo, en el contexto de la terapia citotóxica, la sobreexpresión de Hsp70 en el tejido tratado conduce a una inmunidad profunda y específica del tumor a través de la expresión inducida de IL-6 (*Kottke et al.*, 2009, 2007). Por tanto, es probable que los efectos de las HSPs extracelulares sobre la inmunidad dependan en gran medida del contexto. La familia HSP70 es extensa y se ha informado de que varios miembros de la familia HSP70 son secretados por células tumorales (*Mambula and Calderwood*, 2006; *Santos et al.*, 2017). La evidencia actual indica que la Hsp70 extracelular puede promover procesos tumorigénicos o tener actividades limitantes de tumores, y puede ser dirigida o utilizada para potenciar las respuestas tumorales al tratamiento (*Gobbo et al.*, 2016; *Cordonnier et al.*, 2017). Las actividades limitantes de la actividad tumoral de las Hsp70 descritas hasta la fecha han implicado principalmente varios métodos para estimular la inmunidad tumoral. Por ejemplo, los exosomas derivados de tumores positivos para Hsp70 promovieron la actividad de las células citolíticas natural killer (NK) hacia las células tumorales de una manera dependiente de Hsp70 (*Gastpar et al.*, 2005). Las HSPs extracelulares también pueden promover la inmunidad tumoral facilitando las funciones de las CD mediante la estabilización de los antígenos tumorales y mejorando el procesamiento de los antígenos tumorales por las CD.

### 1.3 Cáncer de mama

El CM es una enfermedad extendida en todo el mundo, presenta más de 1,6 millones de casos nuevos y más de 500.000 muertes anuales a nivel mundial (*Bray et al.*, 2018). En Argentina, según las estimaciones de incidencia del Observatorio Global de Cáncer de la OMS (*Bray et al.*, 2018), el cáncer de mama es el de mayor magnitud en cuanto a ocurrencia para 2018: con un volumen de más de 21.000 casos al año, representando el 17% de todos los tumores malignos y casi un tercio de los cánceres femeninos, lo que significa una tasa de 73 casos por cada 100.000 mujeres.

En particular, en el 2018 se registraron 6.049 defunciones por CM, lo que representó el 20% del total de las defunciones por cáncer en el sexo femenino en Argentina, lo que ubica al CM como el principal responsable de muerte por cáncer en mujeres, seguido a lo lejos por el cáncer colorrectal, responsable de un 11,3% de las muertes por cáncer en

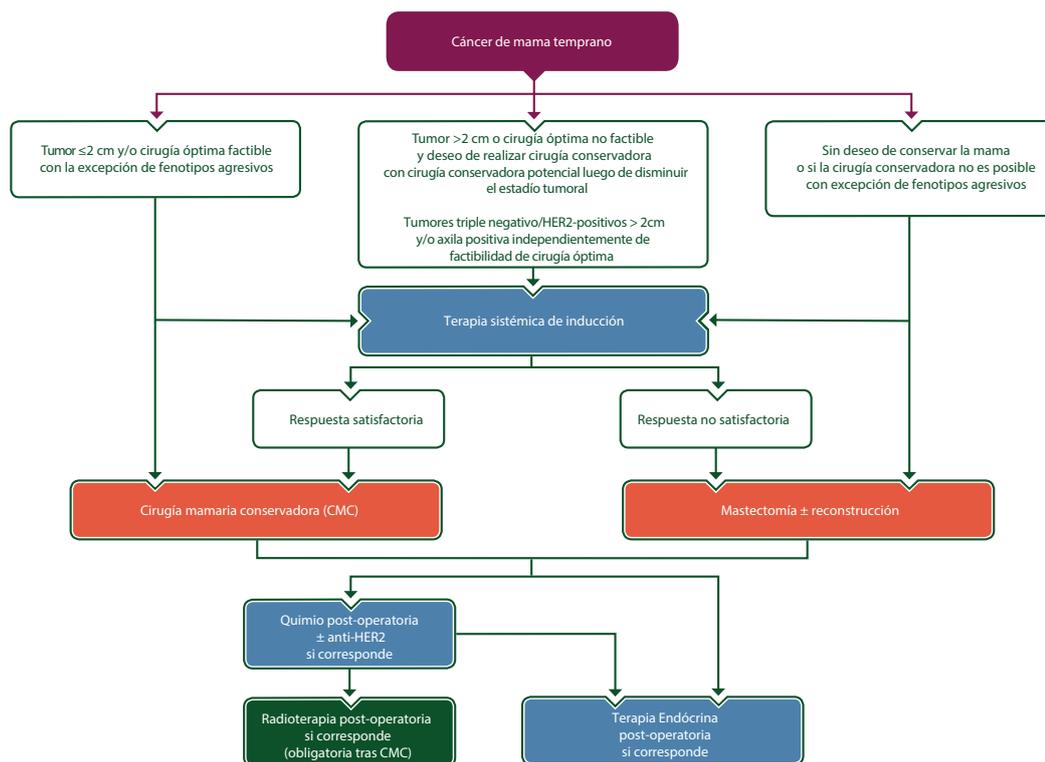
mujeres (*Ministerio de Salud y Desarrollo Social de la Nación (Argentina)*, 2019b).

Particularmente en la provincia de Mendoza entre el 2001 y el 2013 hubieron 3307 defunciones a causa de la enfermedad (*Ramírez et al.*, 2015) , siendo Mendoza la segunda provincia con más mortalidad asociada al CM con una tasa de mortalidad ajustada por edad de 20,5 cada 100.000 mujeres (*Ministerio de Salud y Desarrollo Social de la Nación (Argentina)*, 2019a).

El CM es una enfermedad altamente heterogénea desde el punto de vista clínico, genético y molecular. Históricamente, el CM se ha categorizado según criterios histomorfológicos en carcinoma ductal invasivo (CDI), carcinoma lobulillar invasivo (CLI) y otros subtipos menos comunes (*World Health Organization (WHO)*). En la mayoría de los centros médicos, a la hora de evaluar una neoplasia mamaria, se utiliza un esquema de clasificación de 3 niveles con respecto al patrón de crecimiento, el grado nuclear y la actividad proliferativa para subdividir aún más estos subtipos (*ELSTON and ELLIS*, 1991). Además, se lleva a cabo un análisis inmunohistoquímico del RE, receptor de progesterona (RP) y la expresión de HER2 para determinar si los pacientes califican para la terapia adyuvante y, en nuestro medio, se utiliza el marcador nuclear Ki-67 para distinguir tumores altamente proliferantes que necesitan de un tratamiento más agresivo.

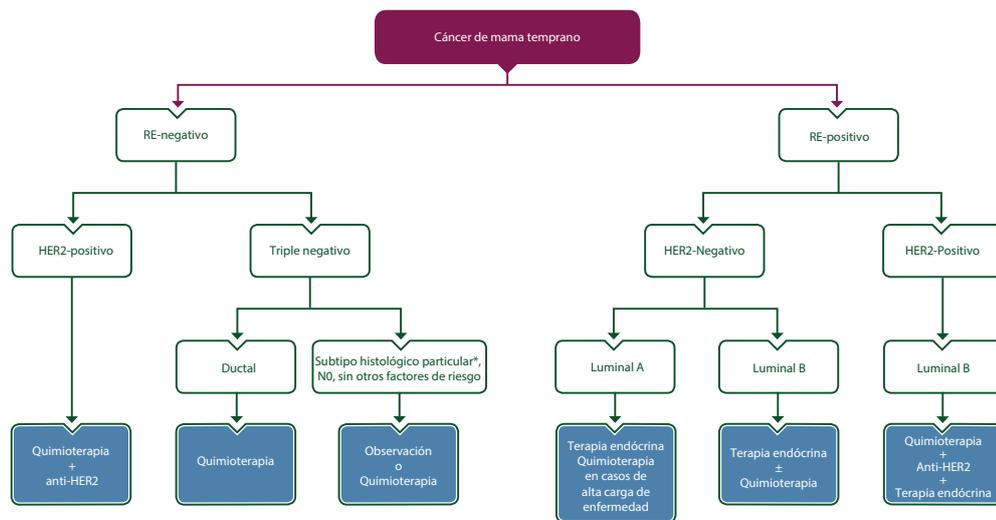
Aparte de lo criterios histológicos y moleculares, se incorpora en la evaluación del paciente información clínica de relevancia que incluye la edad, estado menopáusico, antecedentes familiares y patológicos, y finalmente, para estadificar la enfermedad, se incorporan conocimientos del tamaño del tumor (T), la cantidad de nódulos linfoides afectados (N) y la presencia o no de metástasis (M). Lo que nos permite realizar una evaluación integral de las características tanto clínicas de la paciente como de la identidad molecular del tumor en estudio (*Cardoso et al.*, 2019).

Este estudio detallado del paciente y del tumor, permite establecer diferentes alternativas de tratamiento que pueden variar de acuerdo a la combinación particular de variables que presente cada paciente. Así, en líneas generales, los pacientes pueden recibir cirugías más o menos conservadoras, terapias hormonales, diferentes tipos de quimioterapias adyuvantes o neo-adyuvantes, esquemas variados de radioterapia local y de acuerdo a la evolución de la paciente y sus características iniciales, también diferente seguimiento y control (*Cardoso et al.*, 2019)(Figura 1.5).



**Fig. 1.5 Algoritmo terapéutico en el cáncer de mama.** Esquema de decisiones terapéuticas de acuerdo a las preferencias de la paciente, el estadio tumoral y al fenotipo molecular del tumor. Fenotipo agresivo: Tumores triple negativos o HER2. Modificado de *Cardoso et al.* (2019).

Desde el punto de vista molecular, el CM es una enfermedad altamente heterogénea que posee diferentes clases de mutaciones en el ADN, inserciones/deleciones, re-arreglos cromosómicos y alteraciones en el número de copias que condicionan de forma diferencial el comportamiento tumoral, el pronóstico de los pacientes y por ende, las estrategias terapéuticas (*Koboldt et al.*, 2012)(Figura 1.6. El advenimiento de los chips de ADN (*'microarrays'*) y las nuevas técnicas de secuenciación génica de alto rendimiento (*'Next Generation Sequencing'* o *NGS*) permiten el estudio integral del CM teniendo en cuenta cambios genéticos y epigenéticos asociándolos a variables clínicas de interés en la práctica médica.



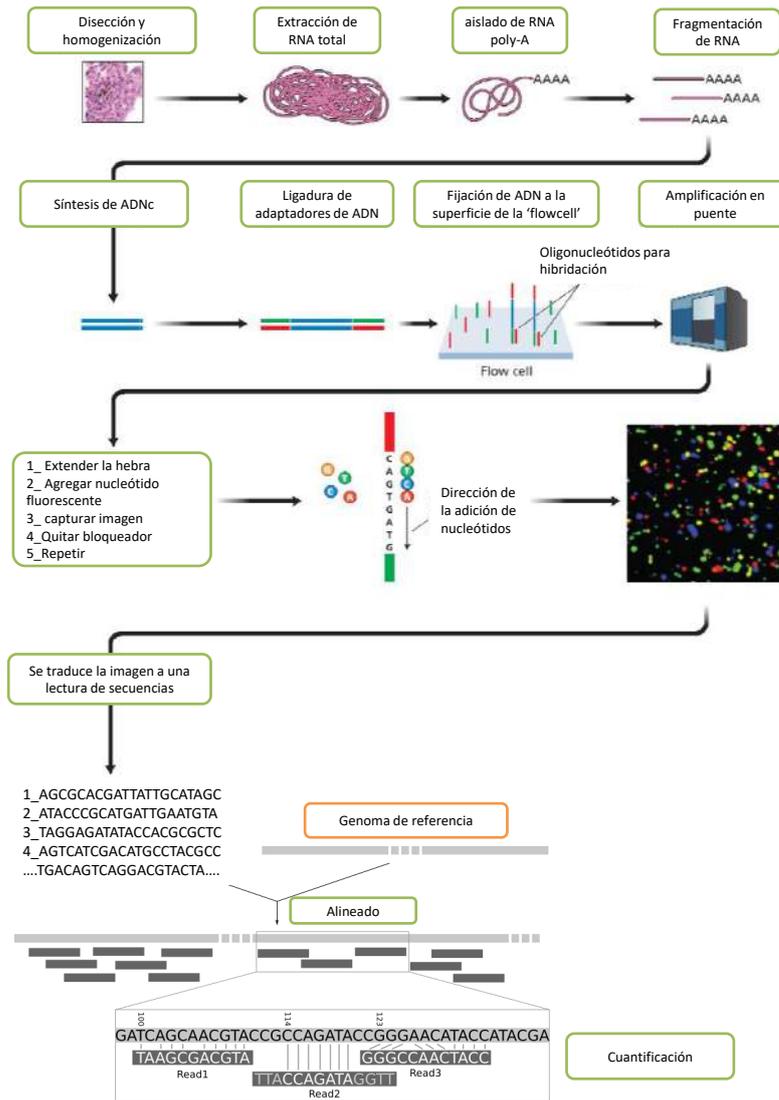
**Fig. 1.6 Elección del tratamiento sistémico (neo)-adyuvante de acuerdo a la expresión de marcador tumorales y los subtipos intrínsecos de CM.** Modificado de *Cardoso et al.* (2019).

### 1.3.1 Evaluación de la expresión génica

En los últimos años, se han utilizado *microarrays* de ADN para clasificar los carcinomas de mama invasivos independientemente de los criterios histomorfológicos e inmunohistoquímicos. Los *microarrays* de ADN representan una tecnología en la que el ARN se extrae de una muestra de tumor, se transcribe en ADNc y se hibridan en pocillos recubiertos con secuencias de ADN/ARN que representan genes distintos. Con ayuda del marcaje de fluorescencia, los niveles de expresión génica se determinan en relación con una referencia normal pudiéndose estimar la expresión relativa de miles de genes en una muestra tumoral (*Peppercorn et al.*, 2008).

De forma alternativa, en los últimos años se ha hecho cada vez más frecuente el uso de las tecnologías de secuenciación de nueva generación (NGS), donde, al igual que durante la técnica de *microarrays*, el ARN se extrae de una muestra biológica y se retrotranscribe a ADNc, luego estas moléculas se fragmentan a través de métodos enzimáticos o físicos y dichos fragmentos son ligados a diferentes adaptadores que permiten su unión a los pocillos presentes en las celdas de flujo (*'flowcell'*) donde se deposita la muestra a amplificar. Una vez que los diferentes fragmentos de ADNc se fijaron en los diferentes pocillos de la celda los diferentes fragmentos se amplifican en un proceso denominado “amplificación en puente” que genera millones de conglomerados de moléculas de ADN que luego serán secuenciadas en paralelo mediante un proceso de “secuenciación por síntesis”. Brevemente, durante este proceso, nucleótidos marcados mediante fluorescencia se van agregando de forma secuencial a los fragmentos

de DNA en las celdas del aparato, que mide la emisión lumínica en cada uno de los conglomerados durante la agregación secuencial de nucleótidos de forma masiva y paralela, lo que se traduce en una 'lectura' de la secuencia de nucleótidos correspondiente a cada uno de los pocillos. En este paso comienza el flujo bioinformático de las lecturas donde se alinean las diferentes secuencias obtenidas contra un genoma de referencia (el genoma humano en este caso), y mediante algoritmos específicos se cuentan la cantidad de lecturas obtenidas para cada uno de los genes del genoma humano en un proceso denominado '*alignment and quantification*' (Figura 1.7).



**Fig. 1.7 Proceso de cuantificación de ARNm mediante técnica de RNA-Seq.** Para cuantificar los diferentes transcritos de ARNm que posee una muestra primero se procede a diseccionar y homogeneizar el tejido a analizar. Posteriormente se extrae el ARN total y se aísla el ARNm mediante la selección de transcritos con extremos Poly-A. Los ARNm son fragmentados y retro-transcritos a ADNc, los cuales son acoplados a adaptadores específicos que permiten su fijado en las celdas de flujo (*'flowcell'*) del secuenciador. Una vez fijados los fragmentos en la *'flowcell'*, se procede a la formación de conglomerados de transcritos mediante el proceso de amplificación en puente. Posteriormente, una vez formados los conglomerados de transcritos, se inicia la secuenciación por síntesis, donde nucleótidos fluorescentes son incorporados de forma secuencial a las hebras de ADN fijadas en las celdas, lo cual produce una excitación lumínica que es captada por el secuenciador, que luego traduce dicha información en una secuencia de nucleótidos. Una vez obtenida la secuencia de cada uno de los fragmentos analizados, estos se alinean a un genoma de referencia y se cuentan la cantidad de copias obtenidas para cada uno de los genes del genoma humano. Modificado de *Momand et al.* (2016)

### 1.3.2 Los subtipos intrínsecos del cáncer de mama

En un trabajo cardinal, *Perou et al.* (2000) empleó *microarrays* de ADN para la investigación del perfil molecular de un colectivo de 65 muestras quirúrgicas de tumores