



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE ESTUDIOS SUPERIORES IZTACALA
BIOMEDICINA

**“Evolución genómica del cáncer de mama triple negativo durante
diferentes etapas de la quimioterapia y su impacto con variables clínicas”**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS BIOLÓGICAS

PRESENTA:

M. en C. Héctor Martínez Gregorio

TUTOR PRINCIPAL DE TESIS: Dr. Felipe Vaca Paniagua

Facultad de Estudios Superiores Iztacala, UNAM

COMITÉ TUTOR: Dr. Alejandro Emmanuel González Voyer

Instituto de Ecología, UNAM

Dr. Pedro Eduardo Miramontes Vidal

Facultad de Ciencias, UNAM

Los Reyes Iztacala, Tlalnepantla, Estado de México, 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE ESTUDIOS SUPERIORES IZTACALA
BIOMEDICINA

**“Evolución genómica del cáncer de mama triple negativo durante
diferentes etapas de la quimioterapia y su impacto con variables clínicas”**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS BIOLÓGICAS

PRESENTA:

M. en C. Héctor Martínez Gregorio

TUTOR PRINCIPAL DE TESIS: Dr. Felipe Vaca Paniagua

Facultad de Estudios Superiores Iztacala, UNAM

COMITÉ TUTOR: Dr. Alejandro Emmanuel González Voyer

Instituto de Ecología, UNAM

Dr. Pedro Eduardo Miramontes Vidal

Facultad de Ciencias, UNAM

Los Reyes Iztacala, Tlalnepantla, Estado de México, 2022

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE ESTUDIOS SUPERIORES IZTACALA

OFICIO CPCB/912/2022

ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
Presente

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día **22 de agosto de 2022** se aprobó el siguiente jurado para el examen de grado de **DOCTOR EN CIENCIAS** del estudiante **MARTÍNEZ GREGORIO HÉCTOR** con número de cuenta **515015342** con la tesis titulada **“Evolución genómica del cáncer de mama triple negativo durante diferentes etapas de la quimioterapia y su impacto con variables clínicas.”**, realizada bajo la dirección del **DR. FELIPE VACA PANIAGUA**, quedando integrado de la siguiente manera:

Presidente: DRA. CARMEN YOLANDA ACEVES VELASCO
Vocal: DR. CRISTIAN GABRIEL OLIVERIO ARRIAGA CANON
Vocal: DRA. VILMA ARACELÍ MALDONADO LAGUNAS
Vocal: DR. JOSÉ BUSTOS ARRIAGA
Secretario: DR. ALEJANDRO EMMANUEL GONZÁLEZ VOYER

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
“POR MI RAZA HABLARÁ EL ESPÍRITU”
Ciudad Universitaria, Cd. Mx., a 06 de octubre de 2022

COORDINADOR DEL PROGRAMA



DR. ADOLFO GERARDO NAVARRO SIGÜENZA



Agradecimientos

Al Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México por la formación académica recibida.

Al Consejo Nacional de Ciencia y Tecnología al otorgarme la beca para cursar mis estudios de doctorado (CVU:630737). Este trabajo fue financiado con los apoyos de los proyectos UNAM PAPIIT IN219217, CONACyT Fondo Sectorial 272573, Fondo SEP CONACyT 285879 y CONACYT Laboratorios Nacionales 2021 315817.

Al Dr. Felipe Vaca Paniagua, por la dirección del presente trabajo.

Al Dr. Alejandro Emmanuel González Voyer y al Dr. Pedro Eduardo Miramontes Vidal, por la asesoría brindada durante el transcurso de mis estudios de doctorado.

Agradecimiento personal

A la UNAM, por abrimé las puertas al conocimiento y darme las oportunidades para mi formación académica.

Al Dr. Felipe Vaca Paniagua, por aceptarme en su equipo de trabajo y guiarme en la realización de este proyecto.

A la Dra. Clara Estela Díaz Velásquez, por todas las enseñanzas durante mi estancia en el laboratorio tanto académicas como de vida.

Al Dr. Hugo Aldo De La Cruz, por su apoyo en el laboratorio y los consejos prácticos durante la escritura de la tesis.

A la Dra. Carmen Yolanda Aceves Velasco, al Dr. Cristian Gabriel Oliverio Arriaga Canon, al Dr. Alejandro Emmanuel González Voyer, a la Dra. Vilma Araceli Maldonado Lagunas y al Dr. José Bustos Arriaga, por la revisión y corrección de la tesis.

A mis compañeros de laboratorio, por la amistad brindada hacia mi persona, por los consejos dados y por las múltiples enseñanzas aprendidas de ellos.

A mis padres Héctor y Catalina, por todo su apoyo y amor incondicional.

A mis hermanos Rocío Celeste y Helmer Domitilo, y a mi cuñado Fidel Desiderio Martínez por su cariño y consejos.

Dedicatoria

A mis padres Héctor y Catalina, por ser el pilar fundamental en mi vida y en mi formación académica, por su amor y apoyo incondicional en todo momento

Índice

<i>Índice de figuras</i>	<i>VIII</i>
<i>Índice de abreviaturas</i>	<i>IX</i>
<i>Resumen</i>	<i>1</i>
<i>Abstract</i>	<i>3</i>
1. Introducción	4
1.1 Cáncer	4
1.2 Epidemiología del cáncer de mama	8
1.3 Cáncer de mama	9
1.4 Clasificación molecular del cáncer de mama triple negativo	11
1.5 Heterogeneidad intratumoral e intertumoral en el cáncer de mama triple negativo	14
1.6 Firmas mutagénicas presentes en el cáncer de mama triple negativo	15
1.7 Factores de riesgo para el cáncer de mama	16
1.7.1 Edad	16
1.7.2 Historia familiar y predisposición genética	16
1.7.3 Factores de historia de vida.....	17
1.7.4 Estilo de vida.....	18
1.8 El cáncer como proceso evolutivo	19
1.8.1 Heterogeneidad intratumoral y evolución.....	20
1.8.2 Microambiente tumoral y evolución	22
1.8.3 Mecanismos evolutivos en el desarrollo tumoral.....	24
1.8.4 Modelos de evolución	25
1.8.5 Evolución en el cáncer de mama.....	28
2. Justificación	30
3. Preguntas de investigación	31
4. Hipótesis	31
5. Objetivos	31
5.1 Objetivo general	31
5.2 Objetivos particulares	31
6. Metodología	32
6.1 Información de la muestra	32
6.2 Extracción de DNA	32
6.3 Preparación de bibliotecas genómicas	33
6.4 Identificación de SNV e indels	33
6.5 Análisis de carga mutagénica tumoral	34
6.6 Identificación de la variación del número de copias	34

6.7	Análisis de las firmas mutagénicas.....	34
6.8	Análisis de vías de señalización.....	35
6.9	Análisis global de alteraciones accionables.....	35
6.10	Análisis filogenético	35
7.	<i>Resultados</i>	36
7.1	Las pacientes con cáncer de mama triple negativo exhiben gran variedad de escenarios clínicos y genómicos.....	36
7.2	Las pacientes con cáncer de mama triple negativo muestran una amplia heterogeneidad genética	39
7.3	Procesos mutagénicos actúan en diferentes estadios del cáncer de mama triple negativo.....	41
7.4	El cáncer de mama triple negativo presenta divergencia temprana.....	42
7.5	Implicaciones clínicas potenciales de las mutaciones detectadas en el cáncer de mama triple negativo	45
8.	<i>Discusión</i>	46
9.	<i>Conclusiones</i>	54
10.	<i>Referencias bibliográficas</i>	55
11.	<i>Anexos</i>	66
	Anexo 1. Flujo de trabajo bioinformático.....	66
	Anexo 2. Mutaciones conductoras de un solo nucleótido	67
	Anexo 3. Mutaciones conductoras en variaciones del número de copias.....	69
	Anexo 4. Artículo publicado en el doctorado	72

Índice de figuras

Figura 1. Alteraciones genéticas y cambios histológicos en el desarrollo del cáncer colorrectal	4
Figura 2. Mutaciones somáticas y procesos mutagénicos en diferentes tipos de neoplasia	5
Figura 3. Características distintivas del cáncer.....	8
Figura 4. Incidencia y mortalidad del cáncer a nivel mundial y en México para ambos sexos.....	9
Figura 5. Características clínicas y moleculares del cáncer de mama	10
Figura 6. Modelo de evolución darwiniana del cáncer	20
Figura 7. Modos de evolución del cáncer.....	27
Figura 8. Historia clínica de las pacientes con cáncer de mama triple negativo ...	38
Figura 9. Composición alélica y alteraciones en el número de copias en genes conductores	40
Figura 10. Perfil de firmas mutagénicas en las pacientes con cáncer de mama triple negativo	41
Figura 11. Árboles filogenéticos de las pacientes con cáncer de mama triple negativo	43
Figura 12. Alteraciones accionables en cáncer de mama triple negativo de acuerdo con la base de datos de oncoKB	46
Figura 13. Resumen de las conclusiones claves de nuestro estudio.....	52

Índice de abreviaturas

APOBEC	del inglés apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
BL1	del inglés Basal-like 1
BL2	del inglés Basal-like 2
CCR	Cáncer colorrectal
CM	Cáncer de mama
CMTN	Cáncer de mama triple negativo
CN	del inglés copy number
CNV	del inglés copy number variation
DNA	del inglés deoxyribonucleic acid
FDA	del inglés Federal Drug Administration
FFPE	del inglés formalin-fixed, paraffin-embedded
FISH	del inglés fluorescence in-situ hybridization
HER2	del inglés human epidermal growth factor receptor 2
IM	del inglés immunomodulatory
indels	inserciones y deleciones
ITH	del inglés intratumoral heterogeneity
INCan	Instituto Nacional de Cancerología
LAR	del inglés luminal androgen receptor
M	del inglés mesenchymal
Mb	Megabase
MRCA	del inglés the most recent common ancestor
MSL	del inglés mesenchymal stem-like
NAC	del inglés neoadjuvant chemotherapy
OMS	Organización mundial de la salud
PARP	del inglés population-attributable risk proportion
pCR	del inglés pathological complete response
RE	Receptor de estrógeno
RP	Receptor de progesterona
SNP	del inglés Single Nucleotide Polymorphism
SNV	del inglés Single Nucleotide Variation
TMB	del inglés Tumor Mutational Burden
WES	del inglés whole-exome sequencing

Resumen

Las pacientes con cáncer de mama triple negativo (CMTN) son tratadas con quimioterapia neoadyuvante debido a la ausencia de terapia dirigida. Sin embargo, el 30% de ellas no logran una respuesta patológica completa después del tratamiento y más del 90% de ellas fallecen debido a la aparición de metástasis. La respuesta diferencial en el tratamiento y el desarrollo de metástasis son atribuidas en parte a la heterogeneidad genética y la evolución tumoral de estos tumores. Actualmente, existe una dificultad persistente para el estudio de la evolución del CMTN debido a la dificultad para la obtención de muestras múltiples de seguimiento que permitan evaluar el perfil genómico y el efecto de la evolución tumoral. Por tal motivo, este trabajo tiene como objetivo central evaluar las alteraciones genéticas y la evolución tumoral de pacientes con CMTN agresivos que no respondieron a diferentes líneas de terapia. Se realizó la secuenciación masiva en paralelo de exomas, de 16 muestras totales de cuatro pacientes con falla terapéutica, que incluyen múltiples muestras de seguimiento, es decir, muestras tomadas en el momento del diagnóstico sin tratamiento previo, muestras después del tratamiento, así como metástasis a ganglios linfáticos y a la piel. Se identificó una notable heterogeneidad intrapaciente con una carga mutagénica tumoral variable. Se identificaron eventos mutacionales conductores tempranos en las pacientes con CMTN como la amplificación del gen *MCL1*. Las lesiones metastásicas tuvieron deleciones en *RB1* y *PTEN*, así como las amplificaciones en *TERT*, *AKT2* y *CCNE1*. Se identificaron 11 de 30 firmas mutagénicas descritas, de las cuales la firma 06 (asociadas a defectos en la reparación del ADN por bases mal apareadas) y 12 (de etiología desconocida) fueron detectadas en cinco de seis muestras analizadas que corresponden a muestras de metástasis de piel y ganglios linfáticos. Por otro lado, el análisis filogenético mostró que las metástasis a ganglios linfáticos ocurren en un estadio temprano del desarrollo del CMTN en las pacientes analizadas. Finalmente, cada paciente tuvo de tres a ocho mutaciones que podrían ser blancos potenciales para el tratamiento dirigido. Este estudio demuestra la complejidad genómica y el desarrollo evolutivo del CMTN clínicamente agresivo, sugiere el desarrollo de

metástasis en etapas tempranas e identifica alteraciones genéticas específicas asociadas con una respuesta a terapias dirigidas potenciales.

Abstract

In triple-negative breast cancer (TNBC), only 30% of patients treated with neoadjuvant chemotherapy achieve a pathological complete response after treatment and more than 90% die due to metastasis formation. The diverse clinical responses and metastatic developments are attributed to extensive intrapatient genetic heterogeneity and complex tumor evolution. For years, it has been challenging to study the evolution of TNBC due to the difficulty in obtention of follow-up samples that allow evaluating the genomic profile and the tumor evolution. In this work, we aimed to evaluate genomic alterations and tumor evolution in TNBC patients with an aggressive disease. We sequenced the whole exome of 16 malignant lesions from four patients who did not respond to diverse lines of therapy, and took several follow-up samples, including samples from tumors before and after treatment, as well as from the lymph nodes and skin metastases. We found substantial intrapatient genetic heterogeneity, with a variable tumor mutational composition. Early truncal events were *MCL1* amplifications. Metastatic lesions had deletions in *RB1* and *PTEN*, along with *TERT*, *AKT2*, and *CCNE1* amplifications. Mutational signatures 06 and 12 were mainly detected in skin metastases and lymph nodes. According to the phylogenetic analysis, the lymph node metastases occurred at an early stage of TNBC development. Finally, each patient had three to eight candidate driving mutations for targeted treatments. This study delves into the genomic complexity and the phylogenetic and evolutionary development of aggressive TNBC, supporting early metastatic development, and identifies specific genetic alterations associated with a response to targeted therapies.

1. Introducción

1.1 Cáncer

El cáncer es un término genérico que se le da a un conjunto de enfermedades que se caracterizan por el crecimiento anormal de las células del cuerpo debido a la acumulación progresiva de alteraciones genéticas y epigenéticas en los genes conductores del cáncer, las cuales proporcionan ventajas de crecimiento selectivo a las células que las portan y consecuentemente su expansión clonal (Martincorena & Campbell, 2015; Vogelstein *et al.*, 2013).

El cáncer es un proceso de múltiples pasos que involucra cambios moleculares e histológicos a través del tiempo (figura 1). Uno de los modelos que ha servido para explicar el desarrollo tumoral, es el cáncer colorrectal (CCR). Estos tumores se originan a partir de una mutación inicial inactivante del gen *APC*, la cual permite la formación de focos de criptas aberrantes, las primeras lesiones precursoras para el desarrollo del CCR. El segundo alelo de *APC* se pierde por delección del locus o sufre una mutación de pérdida de función, causando la inactivación bialélica del gen. Un segundo evento de mutación en *KRAS* es requerido para la formación de microadenomas, y en la etapa localmente avanzada, se genera una serie de mutaciones en *PIK3CA* y *TGF- β* , así como la alteración en diferentes vías de señalización que promueven la invasión de estas células tumorales a la matriz extracelular y favoreciendo la formación de metástasis hacia órganos locales o distantes (Martincorena & Campbell, 2015; Vogelstein *et al.*, 2013). Por otra parte, el desarrollo de modelos que ejemplifiquen la evolución tumoral en los tumores de mama son complejos debido a la alta heterogeneidad intertumoral presente en estos pacientes.

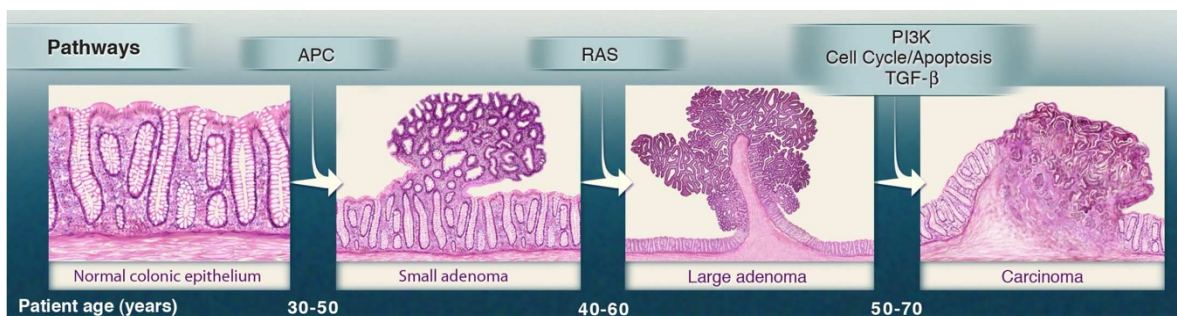


Figura 1. Alteraciones genéticas y cambios histológicos en el desarrollo del cáncer colorrectal. Se presentan las alteraciones genéticas y las vías de señalización (parte superior) afectadas en el CCR, los cambios a nivel histológico (en medio), así como el tiempo necesario para la formación del tumor (parte inferior). Imagen tomada de Vogelstein *et al.*, 2013.

Por otra parte, el número y el tipo de alteraciones genéticas encontradas en una muestra tumoral es variable; depende del tejido afectado y los procesos mutagénicos que han operado en ella. Por ejemplo, el melanoma cutáneo y los tumores de pulmón contienen el número más alto de mutaciones puntuales e inserciones y deleciones (indels) debido a la exposición a los mutágenos luz ultravioleta y al cigarro, respectivamente; en comparación con los tumores pediátricos como el glioblastoma, el neuroblastoma o las leucemias que presentan el número más bajo de mutaciones puntuales e indels (Martincorena & Campbell, 2015; Vogelstein *et al.*, 2013) (figura 2). Más del 95% de las mutaciones encontradas en los tumores son variaciones de un solo nucleótido (SNV, por sus siglas en inglés) y el resto son indels, así como variaciones en el número de copias (CNV, por sus siglas en inglés) (Martincorena & Campbell, 2015; Vogelstein *et al.*, 2013). Los estudios clásicos sugieren que los tumores se forman a partir de la adquisición de más de 6 mutaciones en diferentes genes conductores del cáncer (Armitage & Doll, 1954; Nordling, 1953), sin embargo, los modelos matemáticos y la secuenciación masiva han establecido que para la formación del CCR, de pulmón y tumores de mama sólo se necesitan tres mutaciones en diferentes genes conductores del cáncer (Anandakrishnan *et al.*, 2019; Tomasetti *et al.*, 2015).

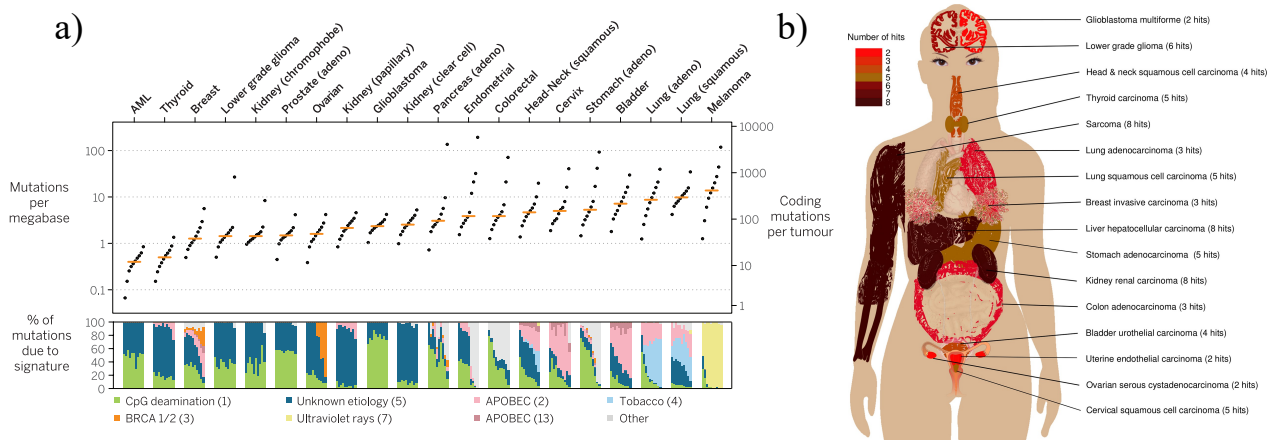


Figura 2. Mutaciones somáticas y procesos mutagénicos en diferentes tipos de neoplasias. a) La imagen muestra el número de total mutaciones somáticas por genoma y exoma encontradas en diferentes tipos de

tumores, y los procesos mutagénicos que han sido asociados a ellos. (b) El número de alteraciones requeridas para la formación de diferentes tipos de tumores. Imágenes tomadas de Martincorena & Campbell, 2015 y de Anandakrishnan *et al.*, 2019.

Las células tumorales presentan características o sellos distintivos que adquieren durante el desarrollo tumoral (Hanahan & Weinberg, 2011). Estas características se describen a continuación.

- Mantenimiento de la señalización proliferativa. El rasgo fundamental de las células tumorales es el mantenimiento de proliferación crónica que se genera mediante diferentes mecanismos que actúan simultáneamente. Por ejemplo, las células tumorales pueden producir y secretar ligandos de factores de crecimiento por sí mismas, generando señales autocrinas de proliferación. También, las células malignas envían señales a las células normales dentro del estroma asociado al tumor para estimular la secreción de factores de proliferación. Por otra parte, la señalización de receptores de membrana puede ser desregulada mediante la elevación de sus niveles de expresión en las células tumorales o a través de su activación constitutiva o de los componentes río debajo de la vía, mediante mutaciones en regiones hipermutables (hot spots).
- Evasión de vías de regulación del crecimiento. Las células tumorales evaden los programas que regulan de manera negativa la proliferación celular. Varios de estos programas dependen de las acciones de los genes supresores de tumor como *RB1* y *TP53*, los cuales funcionan como nodos de control central dentro de dos circuitos reguladores celulares complementarios que definen la activación de vías proliferativas o de programas moleculares de senescencia y apoptosis.
- Resistencia a la muerte celular programada (apoptosis). Las células tumorales desarrollan una gran variedad de estrategias moleculares para evadir la apoptosis. Algunos de los mecanismos por los cuales evaden la apoptosis son: la pérdida de la función de *TP53*, lo que elimina el sensor de daños críticos del circuito que induce la apoptosis; la sobreexpresión de reguladores antiapoptóticos como Bcl-2 y Bcl-XL; la sobreexpresión de señales de sobrevivencia (por ejemplo, Igf1/2); o la baja expresión de factores proapoptóticos como Bax, Bim y Puma.

- Incremento en la capacidad replicativa. Las células normales tienen un número limitado de ciclos de crecimiento y división celular, conocido como el límite de Hayflick (40-60 veces). En contraste, las células tumorales requieren un potencial replicativo ilimitado para generar tumores macroscópicos. El control del potencial replicativo ha sido asociado a dos diferentes barreras de la proliferación: (i) la senescencia, un estado no proliferativo pero viable, y (ii) la crisis, que promueve la muerte celular. Por otra parte, la sobreexpresión de los componentes de la telomerasa aumenta permite la proliferación celular y la resistencia a la apoptosis.
- Angiogénesis. Los tumores requieren sustratos en la forma de nutrientes y oxígeno, así como la capacidad para eliminar desechos metabólicos y dióxido de carbono. La neovasculatura generada por la activación del proceso de angiogénesis, cubre estas necesidades. Los principales factores que estimulan la angiogénesis son los factores de crecimiento vascular endotelial, sobreexpresados durante la hipoxia y por las señales oncogénicas, y la trombospondina 1, un contrapeso clave durante el cambio angiogénico.
- Invasión y metástasis. La invasión y metástasis son procesos de múltiples pasos, en los cuales ocurre una sucesión de cambios biológicos, iniciando con la invasión local, la intravasación de las células tumorales a los vasos sanguíneos y ganglios linfáticos, el tránsito y sobrevivencia de las células tumorales aisladas o en grupos a través del sistema linfático y sanguíneo, la extravasación, y finalmente la colonización de estas células tumorales a los sitios locales o distantes.

Adicionalmente, se han propuesto dos características más de las células tumorales: (i) la desregulación energética, principalmente la del metabolismo de glucosa; y (ii) la evasión de la destrucción celular por el sistema inmune, el cual reconoce y elimina a las células tumorales. Todas estas características que presenta el cáncer se podrían englobar en dos categorías generales que contribuyen a la adquisición de todas las anteriores: i) la promoción del tumor mediada por la inflamación y la ii) inestabilidad genómica causante del aumento de

la tasa de mutaciones, alterando la función celular y contribuyendo a la progresión del tumor. En conjunto, estas características moleculares y celulares hacen que el desarrollo tumoral sea un proceso complejo, orquestado por diferentes alteraciones genéticas, epigenéticas, metabólicas, inmunológicas y cambios a nivel histológico que se reflejan en estos sellos distintivos del cáncer (figura 3).



Figura 3. Características distintivas del cáncer. Las células tumorales presentan diferentes características que van adquiriendo a través del tiempo. Estas características permiten a las células tumorales crecer, reproducirse y desarrollar metástasis. Imagen tomada y modificada de Hanahan y Weinberg 2011.

1.2 Epidemiología del cáncer de mama

El cáncer de mama (CM) es la neoplasia con mayor frecuencia en la población mundial y mexicana. En el reporte epidemiológico realizado por la Organización Mundial de la Salud (OMS) en el 2020, se reportaron 19,292,789 nuevos casos y 9,958,133 muertes por cáncer a nivel mundial (Sung *et al.*, 2021), figura 4. De acuerdo con el tipo de tumor, el CM fue la principal neoplasia maligna con la mayor tasa de incidencia y la quinta con la mayor tasa de mortalidad tanto en hombres como en mujeres. Esto significa que se detectaron 2,261,419 casos (11.7% de todos los casos de cáncer) y fallecieron 684,996 pacientes (6.9% de todos los casos de cáncer) a nivel global. En cuanto a la población mexicana, se detectaron 195,499 nuevos casos y 90,222 muertes por todos los tipos de cáncer, de los cuales el CM presentó la mayor tasa de incidencia y mortalidad. De acuerdo

a estos datos, se reportaron 29,929 nuevos casos (15.3% de todos los casos de cáncer) y 7,931 muertes (8.8% de todos los casos de cáncer). Estos datos epidemiológicos indican que los tumores de mama representan un problema de salud pública a nivel mundial y en nuestra población mexicana.

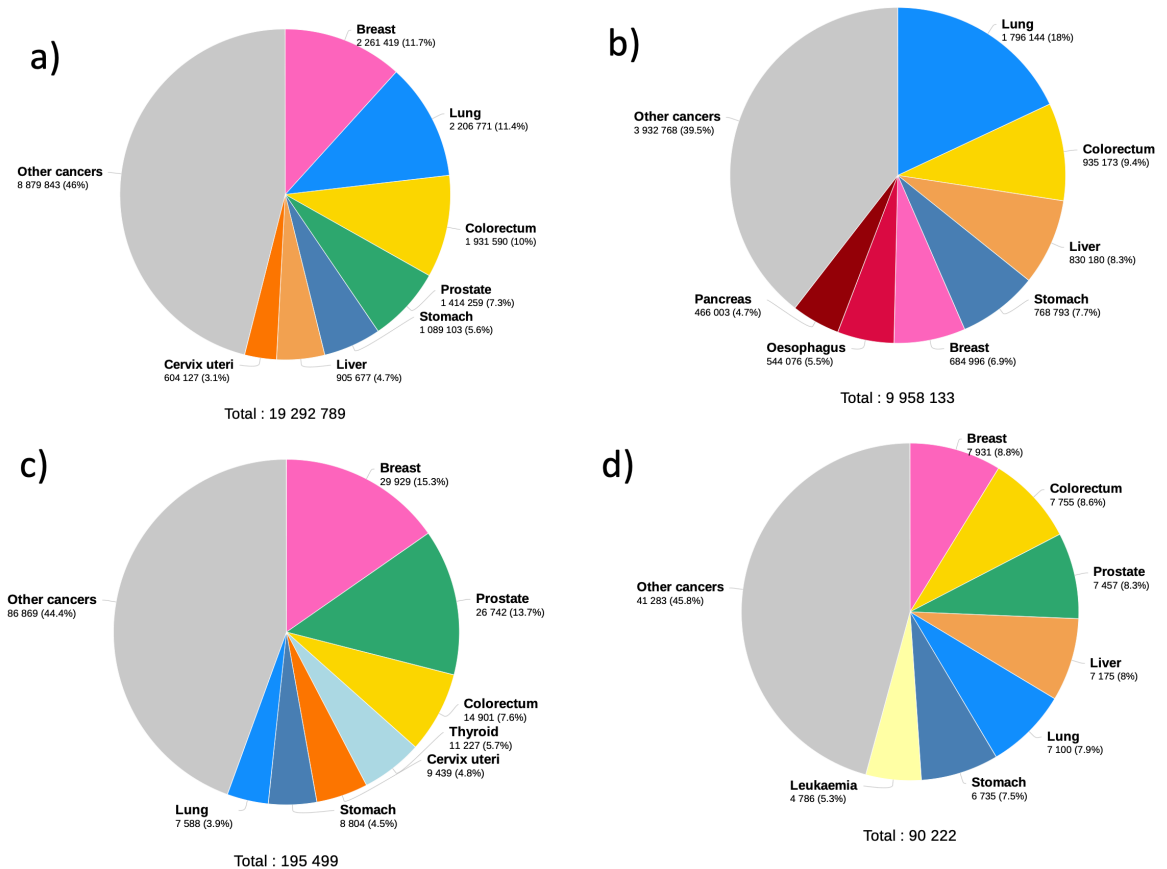


Figura 4. Incidencia y mortalidad del cáncer a nivel mundial y en México para ambos sexos. a) Incidencia y b) mortalidad del cáncer a nivel mundial; c) incidencia y d) mortalidad del cáncer en México. Imágenes tomadas de la base de datos de GLOBOCAN (<https://gco.iarc.fr/>).

1.3 Cáncer de mama

El CM es una enfermedad heterogénea que comprende numerosas entidades clínicas que difieren a nivel histológico, biológico y clínico (Ng *et al.*, 2015). Clínicamente, los tumores de mama se clasifican con base en la expresión de tres receptores farmacológicos mediante inmunohistoquímica y FISH: (i) receptores de estrógenos (RE), (ii) receptores de progesterona (RP), que son diagnósticas por inmunohistoquímica y (iii) el receptor 2 del factor de crecimiento epidérmico humano

(HER2), (figura 5). En este sentido, los tumores de mama que expresan RE y RP son considerados tumores de mama positivo a receptores hormonales, por el contrario, los tumores que no expresan RE, RP y HER2 son denominados como cáncer de mama triple negativo (CMTN).

	Triple negativo	HER2+	Luminal B	Luminal A	
Expresión de receptores		HER2		ER/PR	
% de cáncer de mama	15 - 20 %	10 - 15 %	25 - 30%	40 - 45%	
Grado histológico	Alto grado		Bajo grado		
Genes alterados	TP53	75%	84%	32%	
	PIK3CA	7%	7%	32%	
	PTEN	8%		MAPK3K1	5%
	BRCA1/2	30/6%		GATA3	14%
Terapia dirigida disponible	Sin terapia dirigida disponible	Anti-HER2 (trastuzumab)		Terapia endocrina (tamoxifeno, fulvestrant)	
Pronóstico	Pobre		Bueno		
Sobrevida a los 5 años	30%	35 - 50%	50 - 70%	95%	

Figura 5. Características clínicas y moleculares del cáncer de mama. La imagen muestra la clasificación y las características clínicas de los tumores de mama, los tratamientos asociados a cada subtipo y las mutaciones detectadas en cada uno de ellos. Imagen tomada y modificada de (Harbeck *et al.*, 2019)

Las pacientes que expresan receptores hormonales presentan un bajo grado histológico (baja formación de microtúbulos; pleomorfismo nuclear pequeño, regular y uniforme en la mayoría de los casos, así como un bajo índice mitótico), tratadas con terapias endocrinas o inhibidores de la aromatasa y tienen el mejor pronóstico, ya que la sobrevida global de los pacientes es del 95% a los 5 años de su diagnóstico. Las pacientes que expresan HER2 pueden tener un alto o bajo grado histológico y el tratamiento más efectivo con terapia dirigida (trastuzumab) contra este receptor. Estas pacientes tienen una sobrevida global del 50% a los 5 años de su diagnóstico. Por el contrario, las pacientes con CMTN tienen el peor pronóstico de todos los subtipos clínicos, presentan un alto grado histológico, pueden ser tratadas con quimioterapia basada en antraciclinas y taxanos que se combinan con otras quimioterapias convencionales. Las pacientes con CMTN tienen una sobrevida global del 30% a los 5 años de su diagnóstico (Bergin & Loi, 2019; Bianchini *et al.*, 2016).

En términos de alteraciones genéticas, los tres subtipos de CM presentan mutaciones en *TP53* y *PIK3CA*, con frecuencias alélicas mutantes variadas. Sin embargo, cada subtipo presenta mutaciones específicas en algunos genes. Por ejemplo, las pacientes con receptores hormonales frecuentemente desarrollan mutaciones en *GATA3* y *ESR1*. Por el contrario, los tumores de mama triple negativo se asocian más a mutaciones en *BRCA1/2* (Cancer Genome Atlas Network, 2012).

1.4 Clasificación molecular del cáncer de mama triple negativo

La única opción de tratamiento para el CMTN es la quimioterapia neoadyuvante y con la finalidad de entender las bases moleculares de esta neoplasia y ampliar las opciones terapéuticas, varios grupos de investigadores clínicos y moleculares lo han clasificado en diferentes subgrupos basados en estudios genómicos y transcriptómicos (Curtis *et al.*, 2012; Lehmann *et al.*, 2011; Perou *et al.*, 2000). En 2000, Perou y colaboradores realizaron una subclasificación basada en diferencias en la expresión genética intrínseca de los tumores de mama y que resultó en 5 subgrupos: (i) Luminal A, caracterizado por la expresión de RE+, HER2-, expresión de redes de señalización de RE y citoqueratinas de bajo peso molecular CK8/18; (ii) Luminal B, por RE+, HER2+/-, expresión de redes de señalización de RE y genes asociados a la proliferación celular; (iii) HER2, por HER2+, RE +/-, sobre-expresión o amplificación de *HER2* y sobreexpresión de los genes en el cromosoma 17q.22.24; (iv) basal-like, frecuentemente triple negativo, presentan un alto grado molecular y expresión de marcadores mioepiteliales (EGFR, CK5/6, ck14), c-kit y FOXC1 y (v) Normal-like, caracterizado por RE-/+ , HER2-, bajo grado histológico y expresión de genes mioepiteliales y adiposos, baja expresión de genes luminales, agrupados con fibroadenomas y muestras de mama normal (Perou *et al.*, 2000). Las mutaciones encontradas en tumores de mama usando esta clasificación fueron variables y con frecuencias alélicas mutantes diferentes, con la excepción del normal-like. El Luminal A se caracterizó por mutaciones en *PIK3CA* (45%), *GATA3* (14%), *MAP3K1* (13%), *TP53* (12%) y *CDH1* (9%). Por su parte, el subtipo Luminal B presenta mutaciones en *PIK3CA* (29%), *TP53* (29%) y *GATA3*

(15%). Los tumores HER2+ tienen mutaciones en *TP53* (72%) y *PIK3CA* (39%). El subtipo Basal-like presenta alteraciones genéticas en *TP53* (80%) y *PIK3CA* (9%), así como mutaciones en otros genes conductores con una frecuencia alélica mutante menor del 5% (Cancer Genome Atlas Network, 2012).

Por otra parte, en el 2011 Lehmann y colaboradores clasificaron al CMTN en seis subtipos, con base en la expresión de genes evaluada por ensayos de microarreglos de expresión. Este estudio mostró 2 subtipos tipo basal, basal-like, (BL1 y BL2, por sus siglas en inglés), un subtipo inmunomodulador (IM, por sus siglas en inglés), un subtipo mesenquimal (M, por sus siglas en inglés), un subtipo mesenquimal stem-like (MSL, por sus siglas en inglés) y un subtipo luminal con expresión de receptores de andrógeno (LAR, por sus siglas en inglés) (Lehmann *et al.*, 2011). Cada subtipo presentó características moleculares, clínicas y de respuesta terapéutica diferentes. El subtipo BL1 tuvo una elevada expresión de genes asociados al ciclo celular y respuesta al daño al DNA, mientras que el BL2 mostró enriquecimiento de marcadores mioepiteliales y factores de crecimiento. El subtipo IM expresa genes que codifican para antígenos inmunes y citocinas. El subtipo M y MSL comparten una expresión elevada de genes asociados a la transición epitelio mesenquima y vías de factores de crecimiento, sin embargo, el subtipo MSL presenta una baja expresión de genes involucrados en la proliferación celular con respecto al subtipo M. El subtipo LAR se caracterizó por la expresión de genes luminales y por una alta dependencia a la señalización mediada por el receptor de andrógenos. Empleando esta clasificación, se probó la respuesta al tratamiento con quimioterapia agentes neoadyuvantes (NAC, por sus siglas en inglés) en pacientes con CMTN usando antraciclinas y ciclofosfamidias seguido por taxanos. La respuesta al tratamiento fue variable entre los subtipos, por ejemplo, los pacientes con el subtipo BL1 lograron la respuesta patológica completa (pCR, por sus siglas en inglés) más alta (en el 52% de los pacientes), mientras que el subtipo BL2 y LAR tuvieron la pCR más baja, 0 y 10%, respectivamente. Este estudio puso de manifiesto la notable heterogeneidad molecular del CMTN.

En el 2012, Curtis y colaboradores clasificaron al CM en 10 subtipos moleculares (IntClust) combinando datos genómicos y transcriptómicos (Curtis *et al.*, 2012). El subtipo IntClust 1 se caracterizó por la amplificación en 17q23, presentó grado histológico alto y una inestabilidad genómica alta. IntClust 2 por amplificación en 11q13/14, sin características clínicas distintivas y una inestabilidad genética alta. IntClust 3 por falta de número de copias, grado histológico bajo y una inestabilidad genómica baja. El subtipo IntClust 4 no mostró alteraciones de CNV, tuvo grado histológico bajo y una inestabilidad genómica reducida. IntClust 5 mostró amplificaciones en *HER2*, diagnóstico a edad temprana, grado histológico alto y una inestabilidad genética intermedia. IntClust6 presentó amplificaciones en 8p12, sin características clínicas distintivas y una inestabilidad genómica alta. El subtipo IntClust 7 se caracterizó por la ganancia en 16p, pérdida en 16q y amplificación en 8q, así como por pacientes diagnosticadas a una edad avanzada, tumores con grado histológico alto y con inestabilidad genómica intermedia. El subtipo IntClust8 tuvo ganancia de 1q, pérdida de 16q, diagnóstico a edad avanzada, grado histológico bajo y una inestabilidad genética intermedia. El subtipo IntClust9 mostró ganancia de 8q, amplificación de 20q, grado histológico alto y una inestabilidad genómica alta. El último subtipo, IntClust10, presentó pérdida de 5q, ganancia de 8q, 10p y 12p, diagnóstico a edad temprana, grado histológico alto, tumores grandes y una inestabilidad genómica intermedia. Dentro de esta clasificación, los tumores de mama triple negativo se agruparon en los subtipos InstClust4 e InstClust 10.

Todas las clasificaciones representadas en estos estudios reflejan que el CM es muy heterogéneo y especialmente el CMTN es una entidad clínica y molecular compleja. Aunque para fines de establecimiento de pronóstico y de respuesta al tratamiento, el CMTN se ha clasificado con base en la nula expresión de tres receptores farmacológicos, a nivel molecular este subtipo es altamente heterogéneo, presentando, según el estudio de Lehmann (2011), hasta seis subtipos.

1.5 Heterogeneidad intratumoral e intertumoral en el cáncer de mama triple negativo

Los tumores de mama triple negativos presentan una amplia heterogeneidad intratumoral e intertumoral. La heterogeneidad intratumoral (ITH, por sus siglas en inglés) consiste en que en dentro de diferentes regiones anatómicas de un tumor existen diferentes subclonas o poblaciones de células tumorales, que difieren tanto a nivel molecular como a nivel fenotípico. En la heterogeneidad intertumoral, estas diferencias se observan entre las distintas pacientes.

Los tumores de mama triple negativo presentan una respuesta baja al tratamiento con la NAC, en parte debido su alta ITH. A nivel genómico, se ha observado que el CMTN presenta en promedio 1.6-5.5 mutaciones somáticas por mega base en la región codificante, alrededor de 60 mutaciones somáticas codificantes por cada caso (Cancer Genome Atlas Network, 2012; Kandoth *et al.*, 2013; Rojas-Jiménez *et al.*, 2020; Vaca-Paniagua *et al.*, 2015). Sin embargo, la carga mutagénica en estos tumores no es uniforme, algunos de ellos pueden presentar una carga mutagénica alta de hasta 6.2 mutaciones somáticas por mega base en la región codificante (Kandoth *et al.*, 2013; Rojas-Jiménez *et al.*, 2020).

En el 2012, el consorcio de la TCGA realizó la secuenciación de exomas de 507 tumores de mama que incluyeron tumores positivos a RE, RP y HER2, así como tumores de mama triple negativo. Todos los subtipos de CM presentaron heterogeneidad intertumoral. En el caso del CMTN, las SNV más frecuentes fueron en *TP53* (80%), *PIK3CA* (10%) y una serie de mutaciones con una frecuencia de menos del 5% (Cancer Genome Atlas Network, 2012). En ese mismo año, Shah y colaboradores realizaron la secuenciación de genoma y exoma completo de 65 pacientes con CMTN sin tratamiento previo y observaron la misma heterogeneidad intertumoral presente en el trabajo de la TCGA, siendo *TP53* (62%) el gen más mutado, seguido de *PIK3CA* (10.2%), *USH2A* (9.2%), *MYO3A* (9.2%), *PTEN* y *RB1* (7.7%), además de otros genes con menor frecuencia (Shah *et al.*, 2012). Este trabajo mostró que las mutaciones en *TP53* son eventos tempranos en el CMTN.

Por otra parte, Wang y colaboradores en el 2014 realizaron la secuenciación de célula única en 3 pacientes RE+ y 1 paciente con CMTN. Las pacientes con RE+ mostraron homogeneidad en términos de SNV y CNV, sin embargo, en la paciente con CMTN encontraron ITH, ellos identificaron tres poblaciones que compartieron una serie de SNV y CNV, a su vez, cada una de ellas presentó diferentes SNV y CNV (Y. Wang *et al.*, 2014). Estos estudios describen ampliamente la heterogeneidad presente en los tumores de mama triple negativo.

1.6 Firmas mutagénicas presentes en el cáncer de mama triple negativo

Diferentes procesos mutagénicos operan durante el desarrollo de un tumor. Estos procesos mutagénicos pueden ocurrir por la exposición a mutágenos exógenos (por ejemplo, la luz ultravioleta y/o los compuestos genotóxicos del humo de tabaco) o por mecanismos endógenos (como mutaciones somáticas asociadas a la edad, el estrés oxidativo, entre otros). Estos procesos mutagénicos dejan huellas o patrones específicos en el genoma, los cuales son conocidos como firmas mutagénicas (Alexandrov *et al.*, 2013). En el estudio realizado por Alexandrov y colaboradores en el 2013, describieron 30 firmas mutagénicas en un contexto de 96 trinucleótidos, cada una de ellas asociadas a diferentes procesos mutagénicos.

Las firmas mutagénicas que se han encontrado en los tumores de mama incluyen: las firmas 1 y 5 (asociadas a la edad) presentes en el 75% de los tumores de mama; la firma 3 (asociada a defectos en la reparación del DNA por recombinación homóloga), presente en el 20% de los casos, sin embargo, se ha encontrado que en los pacientes con CMTN este porcentaje se eleva hasta un 75%; las firmas 2 y 13 (asociadas a la actividad de las enzimas APOBEC, por sus siglas en inglés apolipoprotein B mRNA editing enzyme, catalytic) se encuentran con frecuencias superiores al 75% de los casos; las firmas 6, 20 y 25 (asociadas a defectos en la reparación del DNA por bases mal apareadas); y las firmas 17, 18 y 30 (de etiología molecular desconocida), las cuales están presentes en menos del 10% de los casos (Nik-Zainal & Morganella, 2017).

Todas estas firmas identificadas en el CMTN pueden actuar de manera individual o conjunta durante las diferentes etapas del desarrollo tumoral. Las frecuencias pueden variar a medida que se detecten otras firmas mutagénicas, es decir, durante la detección del CM las firmas 1 y 5 son las que se presentan con mayor frecuencia, a medida que va progresando la enfermedad otros procesos mutagénicos son activados, los cuales modifican la frecuencia de las firmas anteriormente mencionadas, manteniéndolas a una baja frecuencia.

1.7 Factores de riesgo para el cáncer de mama

El CM es una enfermedad multifactorial y está asociada a diferentes factores causales, que pueden ser ambientales, genéticos y de comportamiento. Varios estudios epidemiológicos han demostrado que muchos factores pueden influir en el desarrollo de esta neoplasia como la edad, el estilo de vida y las mutaciones germinales, entre otras (Martincorena & Campbell, 2015).

1.7.1 Edad

La edad es un factor de riesgo para el desarrollo de cualquier tipo de cáncer. El riesgo de sufrir algún tipo de cáncer antes de los 40 años es aproximadamente del 2%, sin embargo, a los 80 años el riesgo puede aumentar hasta el 50% (Martincorena & Campbell, 2015). La incidencia de algunos tumores se eleva aproximadamente a la cuarta o hasta la sexta potencia en función de la edad. Teniendo en cuenta sólo la edad como factor de riesgo, los tumores de mama presentan un riesgo acumulado del 2% a los 40, 5% a los 60 y 15% a los 80 años de edad (Martincorena & Campbell, 2015).

1.7.2 Historia familiar y predisposición genética

Aproximadamente un cuarto de los tumores de mama está relacionados a la historia familiar. Las mujeres que tienen parientes cercanos con CM son más propensas a desarrollar esta neoplasia. En un estudio realizado en Reino Unido con una cohorte de más de 113,000 mujeres, se encontró que las mujeres con una pariente de primer grado con CM tienen 1.77 veces la probabilidad de desarrollar

CM (95% CI, 1.58-1.97, $p < 0.0001$) en comparación con aquellas que no tienen parientes con esta neoplasia. El riesgo puede aumentar 2.52 veces (95% CI, 1.83-3.47, $p < 0.0001$) si la paciente tienen dos o más parientes cercanos con CM (Brewer *et al.*, 2017).

La historia familiar que predispone al desarrollo de CM es atribuida principalmente a mutaciones en los genes *BRCA1*, *BRCA2*, *PALB2*, *ATM*, *CHEK2*, *BARD1*, *RAD51C*, *RAD51D*, *TP53*, entre otros (Breast Cancer Association Consortium *et al.*, 2021; C. Hu *et al.*, 2021). Recientemente, se reevaluó el riesgo asociado en genes que predisponen al CM en dos estudios diferentes, uno realizado por el consorcio del CM (Breast Cancer Association Consortium *et al.*, 2021) con una cohorte de 113,000 pacientes y otro por Hu y colaboradores en una cohorte de aproximadamente 65,000 pacientes (C. Hu *et al.*, 2021) basado en paneles de genes. El consorcio de CM encontró 5 genes asociados con un riesgo significativo alto para el desarrollo de CM, *BRCA1* con una razón de momios de 10.57 (8.02-13.93), *BRCA2* con 5.85 (4.85 -7.06), *PALB2* con 5.02 (3.73-6.76), *CHEK2* con 2.54 (2.21-2.91) y *ATM* con 2.10 (1.71-2.57). Adicionalmente, se encontraron otros 7 genes con evidencia moderada para el desarrollo del CM como *BARD1* razón de momios de 2.09 (1.35 - 3.23), *RAD51C* de 1.93 (1.20 – 3.11), *RAD51D* de 1.8 (1.11-2.94), *PTEN* de 2.25 (0.85 - 6.00), *NF1* de 1.76 (0.96 - 3.21), *TP53* de 3.06 (0.63 - 14.91) y *MSH6* de 1.96 (1.15 - 3.33).

Los resultados obtenidos por Hu y colaboradores fueron similares a los del estudio realizado por el consorcio de CM, dos genes fueron asociados con un alto riesgo para CM, *BRCA1* con una razón de momios de 7.62 (5.33-11.27) y *BRCA2* con 5.23 (4.09-6.77). Las variantes patogénicas en *PALB2* y *CHEK2* fueron asociadas con un riesgo moderado, con una razón de momios de 3.83 de (2.68-5.63) y 2.47 (2.02-3.05), respectivamente.

1.7.3 Factores de historia de vida

Los factores de historia de vida como la menarca temprana, la menopausia tardía, el primer embarazo a una edad tardía y la nuliparidad incrementan el riesgo

de CM. Se ha estimado que por cada año de retraso en la menopausia se incrementa el riesgo de desarrollar CM en un 3%. Por el contrario, por cada año de retraso en la menarca o por cada nacimiento adicional se disminuye el riesgo de CM hasta un 5% y 10%, respectivamente (Dall & Britt, 2017). Los factores reproductivos están fuertemente asociados con el estado de los RE, con diferencias en las razones de momios entre CM con RE+ y RE- para la paridad (OR: 0.7 vs 0.9 para ≥ 3 nacimientos vs nuliparidad) y la edad del primer nacimiento (OR: 1.6 vs. 1.2 a la edad ≥ 30 vs <25 años) (Rosato *et al.*, 2014)

1.7.4 Estilo de vida

El estilo de vida puede predisponer a la formación de diferentes tipos de tumores malignos. El consumo de alcohol y grasas pueden incrementar el riesgo de desarrollar CM. El consumo de alcohol puede elevar los niveles de hormonas relacionadas al estrógeno en la sangre y desencadenar la vía de receptores de estrógeno. Un meta análisis basado en 53 estudios epidemiológicos indicó que la ingesta de 35-44 gramos de alcohol por día puede aumentar el riesgo de CM en un 32%, con un incremento del 7.1 en el riesgo relativo por cada 10 gramos adicionales de alcohol por día (Hamajima *et al.*, 2002). Por otra parte, muchos estudios sugieren que una mayor ingesta de ácidos grasos saturados antes del diagnóstico se ha asociado con un riesgo incrementado de desarrollar CM. La ingesta de ácidos grasos trans posdiagnóstico ha sido asociada a un incremento de 45% del riesgo de muerte para el CM (Makarem *et al.*, 2013). Aunque la relación entre fumar y el riesgo de CM permanece controversial, se ha demostrado que el riesgo de CM es elevado en mujeres que fuman y beben (RR = 1.54, 95% CI 1.18-1.99) (Knight *et al.*, 2017). Por otro lado, la obesidad puede ser un factor protector o de riesgo para el CM. En un estudio epidemiológico donde se incluyeron 58,146 mujeres con CM antes de la menopausia y 144,600 después de la menopausia se evaluó la proporción de riesgo atribuible a la población (PARP, por sus siglas en inglés), es decir, la proporción de casos de una enfermedad en una población que podrían no haber ocurrido en la ausencia de un factor de riesgo. El sobrepeso y la obesidad (grado I, II y III) tuvieron una PARP de 16.4% (95% CI, 14.4-18.4) en las pacientes

posmenopáusicas, en cuanto a las pacientes premenopáusicas, no se encontró ninguna asociación. Otro punto interesante fue la asociación de la densidad mamaria para el desarrollo del CM. Pacientes premenopáusicas con alta densidad mamaria tienen una PARP de 28.9% (95% CI, 25.3-32.5) y posmenopáusicas de 14.4% (95%, 12.6-16.0), aunque hay ciertas diferencias en el grado de densidad que se presenta. Por ejemplo, en pacientes premenopáusicas con densidad heterogénea o extrema densidad tienen una razón de momios de 1.57 (95% CI, 1.46-1.69) y 1.81 (95% CI, 1.65-1.99), respectivamente, a diferencia de las pacientes posmenopáusicas con razón de momios de 1.40 (95% CI, 1.34-1.45) para densidad mamaria heterogénea y 1.58 (95% CI, 1.46-1.71) para senos extremadamente densos (Engmann *et al.*, 2017).

1.8 El cáncer como proceso evolutivo

En un estudio nominal publicado por Nowell en 1976, se postuló que el cáncer es un proceso evolutivo que se origina a partir de una célula ancestral única que adquiere una serie de mutaciones en los genes conductores, las cuales le proporcionan ventajas de crecimiento selectivo sobre las demás células dentro de su nicho, llevándola a una expansión clonal (Nowell, 1976) (figura 6). Nowell reconoció que las células tumorales son caracterizadas por una inestabilidad genética y que a lo largo del tiempo nuevas mutaciones son acumuladas en un proceso de múltiples pasos, generando subpoblaciones tumorales o subclonas. Sin embargo, debido a la selección natural, muchas subclonas son eliminadas y sólo una de ellas logra convertirse en la subclona dominante que dirige el desarrollo del tumor. El modelo propuesto por Nowell recapitula el concepto establecido de la evolución Darwiniana, donde se sugiere que los individuos reproductivos, en este caso las células tumorales, adquieren diferentes alteraciones genéticas seguidas de una selección natural de las variantes con la mayor eficacia biológica. Sin embargo, en la última década y con la llegada de tecnologías de alto rendimiento como la secuenciación masiva en paralelo, se ha demostrado que la evolución de las poblaciones tumorales es más compleja de lo que originalmente propuso Nowell.

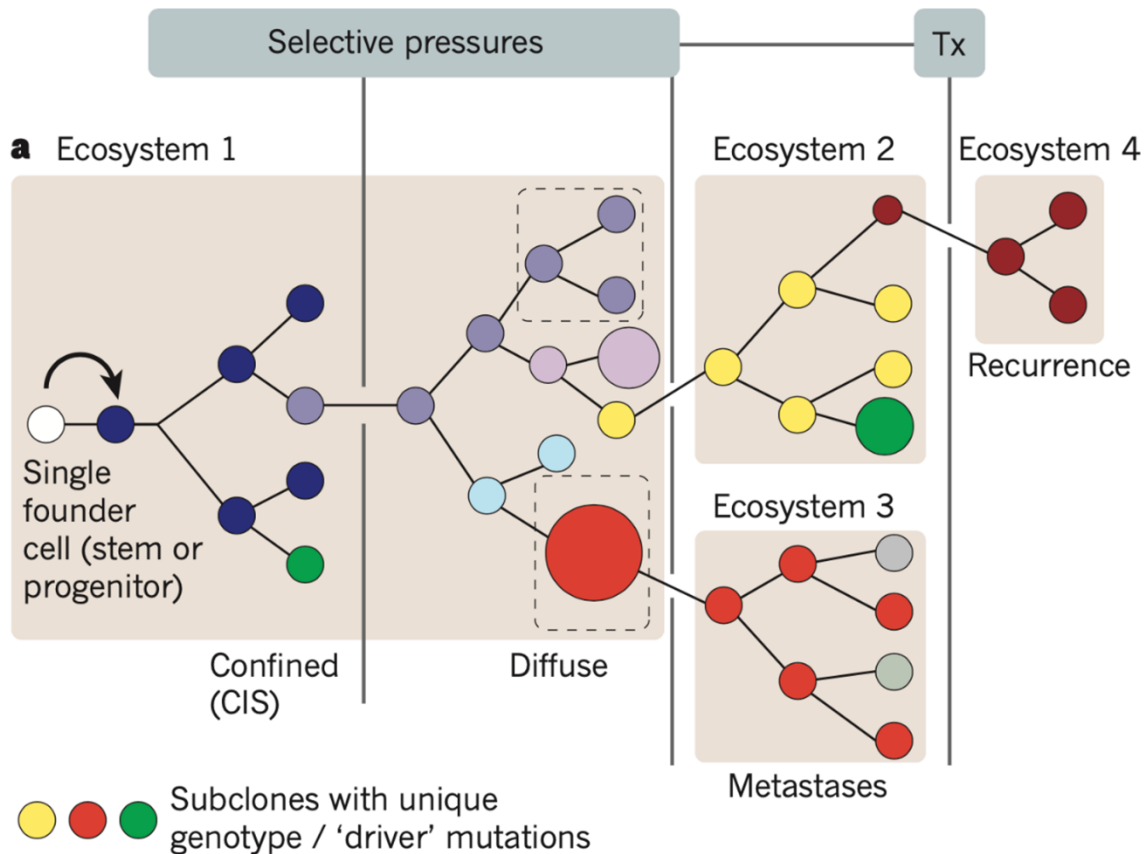


Figura 6. Modelo de evolución darwiniana del cáncer. Las células tumorales se originan a partir de una mutación conductora del cáncer en una sola célula ancestral dentro de un tejido (ecosistema). Durante la evolución tumoral se generan diferentes subclonas (poblaciones tumorales que son representadas por los círculos de colores) que compiten por espacio y nutrientes, y que son seleccionadas de manera positiva o negativa por el microambiente tumoral y los tratamientos antineoplásicos. Aquellas subclonas que son favorecidas por las diferentes presiones de selección son las que dirigen el desarrollo del tumor en la etapa avanzada de la enfermedad y metástasis. Al momento del diagnóstico de un tumor, este se ha sometido a varias rondas de división y selección de subclonas. Imagen tomada de Greaves & Maley, 2012.

1.8.1 Heterogeneidad intratumoral y evolución

El proceso de inestabilidad genómica resulta en una tasa elevada de aberraciones genómicas que van desde mutaciones puntuales hasta variaciones del número de copias y la duplicación o triplicación del genoma completo. Esta inestabilidad contribuye a la ITH que proporciona un conjunto de mutaciones, en la cual la selección natural actúa en un microambiente determinado (Burrell & Swanton, 2014).

Las mutaciones generadas por diferentes procesos endógenos y exógenos son asociadas a procesos específicos de inestabilidad genómica, que reflejan el tipo

de daño en el DNA que se ha producido. Por ejemplo, los tumores de pulmón de células no pequeñas de fumadores muestran una preponderancia de transversiones C>A y CNV en comparación con los no fumadores (Govindan *et al.*, 2012), mientras que los tumores colorrectales con deficiencias en los mecanismos de reparación del DNA por bases mal apareadas exhiben un enriquecimiento de transiciones C>T, particularmente en los sitios CpG y un bajo nivel de alteraciones cromosómicas (McGranahan & Swanton, 2015).

En la mayoría de las muestras tumorales analizadas, se han detectado al menos dos tipos de procesos mutagénicos. La firma 1, identificada en 25 tipos de tumores diferentes. Esta firma se caracteriza por transiciones C>T en los sitios CpG, lo cual es ocasionado por la desaminación hidrolítica espontánea de la 5-metilcitosina, un cambio molecular asociado con la edad. Gran parte de estas mutaciones son adquiridas antes de la tumorigénesis. El segundo proceso mutagénico lo comprenden las firmas 2 y 13, asociadas a la actividad de las enzimas de la familia APOBEC, identificadas en 15 tumores diferentes y caracterizadas por mutaciones de C>T y C>G en los sitios TpC. La mutagénesis mediada por las enzimas APOBEC está asociada a la adquisición de mutaciones conductoras, enfatizando la importancia de este proceso mutagénico en la conformación de la evolución del genoma del cáncer (Alexandrov *et al.*, 2013). Un ejemplo de la importancia de las firmas 2 y 13 es que en tumores de pulmón de células no pequeñas se ha descrito que las mutaciones tempranas son dominadas por el efecto del tabaco, induciendo transversiones C>A, mientras que la mutagénesis mediada por las enzimas APOBEC fue la fuerza mutagénica dominante en la etapa tardía en la evolución de este tumor, proporcionando un sustrato fértil para la adaptación del tumor a las presiones de selección del microambiente tumoral y las inducidas por los tratamientos citotóxicos o la radioterapia (de Bruin *et al.*, 2014).

Por otra parte, la terapia actúa como una fuente exógena de inestabilidad genómica (McGranahan & Swanton, 2015). Por ejemplo, en el estudio realizado por Ding y colaboradores en el 2012, donde estudiaron la evolución clonal de leucemia mieloide aguda en tumores primarios y recurrentes, encontraron un incremento en

la tasa de transversiones después de la terapia citotóxica en las muestras de recurrencia (46%), comparada con las mutaciones antes del tratamiento (30.7%) (Ding *et al.*, 2012). Los resultados de Ding y colaboradores fueron consistentes con un estudio realizado en un modelo de *C. elegans* con tratamiento de cisplatino. En este trabajo se observó un aumento de transversiones C>A, así como una elevada tasa de sustituciones de 2 nucleótidos, indels y variantes estructurales (Meier *et al.*, 2014). Por otra parte, pacientes que han sido tratados con temozolomida presentan una tasa elevada de transiciones C>T en los sitios CpC y CpT (Johnson *et al.*, 2014). En el 2014, Johnson y colaboradores realizaron una comparación del perfil genómico de gliomas en el diagnóstico y la recurrencia, esta última tratada con temozolomida. Ellos encontraron que 6 de 10 tumores recurrieron como glioblastoma, un tumor de alto grado con pronóstico pobre. Las 6 pacientes mostraron hipermutación, con una carga mutagénica aumentada de 7 a 540 veces más en comparación con el glioma primario. Además, muchas mutaciones fueron seleccionadas de manera positiva, principalmente en los genes *RB1* y *CDKN2A* (Johnson *et al.*, 2014). En este ejemplo, la terapia no actuó como una fuente exógena de mutaciones, sino como una barrera de selección, que estableció la trayectoria evolutiva de los tumores y su progresión a una fase más agresiva. Todos estos procesos exógenos y endógenos, ambientales, fisiológicos y terapéuticos contribuyen a la inestabilidad genómica en los tumores, generando un conjunto de mutaciones donde la selección actúa sobre la evolución tumoral.

1.8.2 Microambiente tumoral y evolución

Las poblaciones o subclonas tumorales crecen en un tejido complejo y dinámico, con múltiples componentes que pueden imponer restricciones en la evolución de un tumor. Tales restricciones surgen debido a las limitaciones de recursos, depredación mediada por el sistema inmune y condiciones adversas de crecimiento en la forma de hipoxia, acidosis y tratamientos para el cáncer, entre otros factores. Además, cada subclona compete con las otras por oxígeno, nutrientes o espacio (Marusyk & Polyak, 2010).

En un estudio realizado por Yuan y colaboradores, donde integraron datos de expresión de genes, CNV y análisis de células estromales (muestras fijadas en portaobjetos y teñidas con hematoxilina & eosina), encontraron que la distribución espacial de las células estromales se asocia a una sobrevida global baja en pacientes con CM RE- (Yuan *et al.*, 2012). Similarmente, en tumores de ovario, el porcentaje de las células estromales ha sido asociado a una sobrevida global pobre y una supervivencia libre de enfermedad baja (Natrajan *et al.*, 2016). En tumores de próstata, un índice que evalúa el nivel de la inestabilidad genómica junto con el de la hipoxia, mostró mayor precisión para establecer el pronóstico, más allá de los parámetros clínicos convencionales (Lalonde *et al.*, 2014).

Algunos componentes del microambiente del tejido promueven la cooperación subclonal y la progresión tumoral. En un estudio realizado por Marusky y colaboradores se usó un modelo de ratones con xenotransplantes para observar las interacciones subclonales en los tumores. Ellos observaron que las subpoblaciones tumorales promovieron la supervivencia y el crecimiento de las células tumorales a través de los cambios en el microambiente mediadas por la IL-11 (Marusyk *et al.*, 2014). Por otra parte, Cleary y colaboradores demostraron que la cooperación subclonal puede ser esencial para el mantenimiento del tumor en un modelo de CM (Cleary *et al.*, 2014). Utilizaron ratones que contenían subclonas basales, luminales y una mezcla de ambos en una proporción 1:1. En los ratones monoclonales no se observó la formación del tumor, en cambio, los ratones biclonales fueron altamente tumorigénicos.

El ecosistema del tejido canceroso puede alterarse radicalmente por la quimioterapia o radioterapia. Muchas células tumorales pueden ser destruidas, pero la remodelación del ecosistema del microambiente tumoral crea nuevas presiones de selección, fuentes y oportunidades que permiten que surjan nuevas subclonas a partir de células preexistentes con nuevas variantes genéticas que les permiten adaptarse a las presiones impuestas por el tratamiento.

1.8.3 Mecanismos evolutivos en el desarrollo tumoral

La mutación, la deriva génica, la migración y la selección son fuerzas evolutivas fundamentales que colectivamente definen la diversidad genética dentro de una población (Z. Hu *et al.*, 2017; Lipinski *et al.*, 2016). La evolución clonal, el proceso de acumulación de mutaciones y adaptación en las células somáticas, involucran cada una de estas fuerzas.

Las mutaciones son la primera fuente de diversidad genética en las especies, así como en las poblaciones tumorales. El tipo de mutaciones identificadas en las poblaciones tumorales se compone de SNV, CNV y rearrreglos genómicos complejos, los cuales establecen las características moleculares y clínicas a un tumor. Varios estudios genéticos han proporcionado un catálogo de alteraciones somáticas entre diferentes tipos de tumores y se ha observado que la tasa de mutación por división celular varía durante el curso de la progresión de la enfermedad (Lipinski *et al.*, 2016).

La deriva génica se refiere a los cambios en las frecuencias alélicas de una población debido a eventos aleatorios de muerte y nacimiento. En este contexto, cada célula nueva perteneciente a una subclona tumoral tiene una cierta probabilidad de morir como resultado de los factores aleatorios y ocasionalmente todas las células de una subclona pequeña o mayoritaria muere, incluso si esta clona específica porta mutaciones benéficas. La deriva génica tiene un mayor impacto en las poblaciones pequeñas, por ello existe una mayor probabilidad de erradicar una sola célula o una sola clona que aún no se ha expandido significativamente. El efecto de la deriva génica en la composición subclonal es más pronunciado después de un cuello de botella poblacional. Por ejemplo, cuando pocas células o células únicas colonizan un nuevo nicho metastásico, o después de una reducción masiva en el tamaño de la población a través de tratamientos citotóxicos o extirpación quirúrgica (Lipinski *et al.*, 2016).

La migración es otro elemento que altera la diversidad genética de una población. La migración celular tiene un papel dual en la generación de la diversidad

genética tumoral. Por un lado, la migración dentro del tumor relaja las restricciones espaciales, aumentando la eficacia de la selección dando lugar a una menor diversidad genética dentro de un tumor. Por otro lado, la migración fuera del tumor y en las metástasis aumenta la divergencia genética entre los tumores presentes en el mismo paciente (heterogeneidad intertumoral) (Lipinski *et al.*, 2016).

Una nueva mutación que incrementa la habilidad de las células para sobrevivir y reproducirse bajo las condiciones ambientales particulares del microambiente tumoral y que han escapado de la deriva génica podrían gradualmente incrementar su abundancia dentro de las poblaciones. En los tumores existen múltiples subclonas que portan diferentes mutaciones conductoras, mostrando distintos fenotipos y evolucionando de diversas formas. Las células tumorales pueden ser seleccionadas de manera positiva o de manera negativa. En esta última hay un decremento en la diversidad poblacional a través de una selección purificadora. Todas estas fuerzas evolutivas forman y moldean el genoma de las células tumorales (Z. Hu *et al.*, 2017).

1.8.4 Modelos de evolución

Antes de describir los modelos de evolución tumoral es necesario presentar algunos conceptos evolutivos importantes. Una clona se define como un grupo de células que comparte perfiles genéticos idénticos, mientras que las subclonas son un grupo de células que divergen en el linaje evolutivo y han adquirido mutaciones adicionales. Las mutaciones troncales son mutaciones ancestrales en el tronco del árbol filogenético y que son compartidas por todas las clonas, mientras que las mutaciones subclonales definen un linaje que ha divergido del tronco. Las mutaciones privadas son las mutaciones presentes en una sola muestra. Otro punto importante es el concepto de eficacia biológica que se refiere a la capacidad de una célula tumoral para sobrevivir y proliferar, de tal forma que pueda propagar su genotipo en el tumor. Las subclonas tumorales con eficacia biológica incrementada son las más prevalentes en la masa tumoral a lo largo del tiempo. Un barrido selectivo describe el proceso en el cual un genotipo emerge con una eficacia

biológica extremadamente alta, la cual le permite superar la concentración de las otras subclonas en un tumor (Z. Hu *et al.*, 2017).

Hay cuatro modelos que describen la evolución tumoral, el modelo lineal, el ramificado, el neutral y el de evolución puntuada (figura 7). El modelo lineal postula que a partir de un ancestro común las mutaciones se adquieren linealmente en un proceso gradual que conduce a estados malignos del cáncer. En este modelo se describe que nuevas mutaciones conductoras proporcionan ventajas selectivas tan fuertes que las nuevas subclonas superan a las subclonas anteriores a través de un barrido selectivo que ocurre durante la evolución tumoral. El barrido selectivo ocurre después de que las mutaciones conductoras son adquiridas. El árbol filogenético resultante de la evolución tumoral muestra una clona dominante con intermediarios raros con bajas proporciones que persisten de los barridos selectivos anteriores.

En la evolución ramificada, las clonas divergen de un ancestro común y evolucionan de manera paralela en la masa tumoral, resultando en múltiples linajes subclonales. No hay barrido selectivo en la evolución ramificada, sino que múltiples subclonas se expanden simultáneamente debido a que presentan una eficacia biológica incrementada que les permite coexistir. En este modelo, la ITH fluctúa durante la progresión del tumor y se observan múltiples subclonas presentes en el momento del muestreo clínico. El árbol filogenético de la evolución ramificada presenta taxas intermedias, así como taxas que se han expandido clonalmente debido a la selección positiva de las mutaciones conductoras en los linajes subclonales.

La evolución neutral es un caso extremo de la evolución ramificada que plantea la hipótesis de que no hay selección o cambios en la eficacia biológica producto de las mutaciones durante la mayor parte de la vida de un tumor. Este modelo asume la acumulación de mutaciones aleatorias a lo largo del tiempo lo que conduce a la deriva génica y a una amplia ITH. En este modelo evolutivo es importante destacar que la ITH es un subproducto de la progresión del tumor y no tiene importancia funcional para impulsar el crecimiento tumoral. El árbol filogenético resultante de la evolución neutral consiste en muchos nodos

intermedios y una estructura altamente ramificada, sin ninguna evidencia de taxones únicos que se hayan expandido a causa de la selección natural.

Los modelos lineal, gradual y neutral comúnmente asumen que las mutaciones son adquiridas secuencialmente y gradualmente a lo largo del tiempo. Sin embargo, hay algunos tumores que presentan una gran cantidad de aberraciones genómicas que ocurren en un corto lapso en los estadios tempranos de la progresión tumoral. Este proceso es explicado mediante la evolución puntuada. Este tipo de evolución establece una ITH alta en los estadios iniciales de la iniciación de un tumor, después una o pocas clonas dominantes se expanden para formar una masa tumoral. El árbol filogenético resultante de la evolución puntuada es definido por la ausencia de clonas intermedias durante la evolución tumoral, resultando en un nodo de raíz largo y uno o pocas clonas dominantes. En este modelo, las células tumorales son más invasivas, metastáticas y resistentes al tratamiento. En contraste con el modelo de la evolución lineal y ramificada que son sustentados por mutaciones puntuales, la evolución puntuada se sustenta por aberraciones en el número de copias del DNA y reordenamientos cromosómicos.

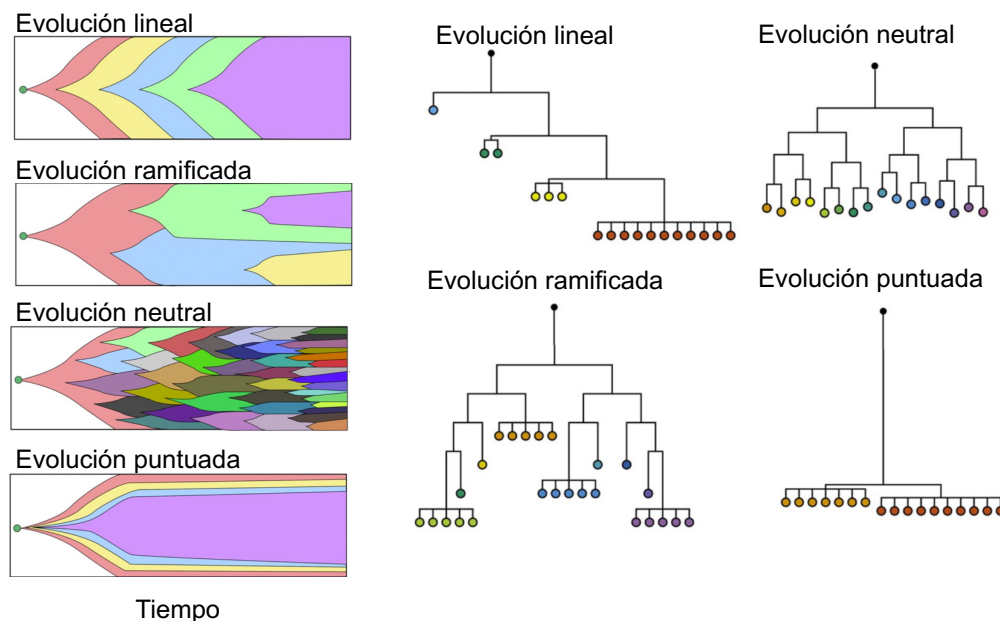


Figura 7. Modos de evolución del cáncer. a) Representación gráfica de los diferentes modelos de evolución usando gráficas de “fishplot”. b) Linajes clonales y árboles filogenéticos. Los colores representan las diferentes clonas que se generan durante el desarrollo del tumor.

1.8.5 Evolución en el cáncer de mama

Diferentes estudios han tratado de elucidar la evolución tumoral del CM y se ha observado que estas neoplasias pueden presentar evolución lineal, ramificada y puntuada. La evolución neutral no se ha detectado en CM debido a que es una característica peculiar detectada en los tumores colorrectales.

En el 2014, Wang y colaboradores realizaron la secuenciación de genomas y exomas de células únicas (single-cell) a cuatro pacientes con CM, tres de ellos positivos a receptores hormonales (RE-, RP-, HER2+) y una con CMTN. Las pacientes positivas a receptores hormonales mostraron una evolución lineal con SNV en *PIK3CA*, *CASP3*, *FBN2* y *PPP2R5E*, y CNV similares en las células tumorales únicas, presentando una población monoclonal. Por su parte, las células únicas de los tumores de mama triple negativo mostraron una evolución ramificada con SNV en diferentes genes conductores del cáncer, así como varias deleciones cromosómicas y una amplificación focal en el cromosoma 19p13.2. Posteriormente, usaron un modelado matemático para determinar la tasa de mutaciones (M_R) y encontraron que el tumor de mama triple negativo tuvo una $M_R=8$, lo cual sugiere que estos tumores de mama tienen 13.3 veces más mutaciones por ronda de división celular (tasa de mutación de una célula normal es de 0.6 por división celular) (Y. Wang *et al.*, 2014).

En el 2015, Gao y colaboradores usaron secuenciación de genoma completo de células únicas de 12 pacientes con CMTN sin tratamiento previo que mostraron evidencia inicial de CNV. Durante el estudio, confirmaron que estas pacientes presentaron altos niveles de CNV y que presentaron 1-3 subclonas tumorales en las muestras analizadas. El análisis filogenético mostró un nodo de raíz largo con múltiples alteraciones de CNV, interpretadas como los eventos fundadores del desarrollo tumoral en estas pacientes. Los datos demostraron la presencia de una evolución puntuada dirigido por CNV, donde una gran cantidad de CNV son adquiridas en un estadio temprano en la evolución tumoral y que permanecen de forma altamente estable durante la expansión de la masa tumoral. A pesar de que ellos analizaron cientos de células únicas, no detectaron ningún perfil intermedio del

número de copias, la cual descartó el desarrollo de una evolución gradual (R. Gao *et al.*, 2016)

En el 2017, Yates y colaboradores realizaron la secuenciación de genoma completo de 17 pacientes (40 muestras tumorales totales, de 2 a 4 muestras por paciente) con diferentes subtipos de tumores de mama. Este estudio comprendió tres escenarios clínicos: metástasis a ganglios linfáticos axilares sincrónicos, metástasis distantes y recurrencia local. Dieciséis de 17 muestras fueron clonalmente relacionadas, es decir compartieron entre 12 a 98% de todas las mutaciones detectadas. Además, las metástasis a ganglios linfáticos mostraron alta similitud al tumor primario. Las mutaciones conductoras se concentraron en el tronco del árbol filogenético, a pesar de que se presentaron 1 o 2 genes conductores con mutaciones patogénicas adicionales en las clonas de metástasis o recurrencia (Yates *et al.*, 2017). En este estudio no se estableció un modo de evolución debido a que la mayoría de los pacientes solo tuvieron dos muestras de tumor.

En el 2018, Barry y colaboradores estudiaron a 11 pacientes con CM de diferentes subtipos sin tratamiento previo y con metástasis a ganglios linfáticos. Ellos identificaron dos principales rutas evolutivas: una divergencia temprana y una ausencia completa de divergencia. La divergencia temprana fue observada en cuatro de 11 pacientes, en la cual los ganglios linfáticos divergieron en los estadios iniciales del desarrollo del CM. En el resto de la cohorte se encontró una topología tipo palma, con ramas cortas y una ausencia de divergencia entre los ganglios linfáticos y los tumores de mama (Barry *et al.*, 2018)

Estos trabajos sobre la evolución del CMTN son los más representativos de la literatura. Los dos primeros estudios involucran el uso de células únicas con un corte transversal. A diferencia de las dos últimas publicaciones, las cuales usan tejidos tumorales y un corte longitudinal, donde los tumores de mama triple negativo son minoritarios. En el estudio del CM, los tumores primarios y las muestras de metástasis son los más frecuentes, sin embargo, regularmente no hay muestras de seguimiento que permitan recapitular los procesos biológicos y moleculares presentes en estas neoplasias. Por tal motivo, este trabajo de investigación se

centró en el estudio del CMTN en diferentes estadios de la enfermedad. Para aumentar el impacto de nuestro trabajo en la comprensión del CMTN, incluimos muestras de seguimiento, muestras de tumores antes y después del tratamiento, así como muestras de ganglios linfáticos y metástasis a piel. En este conjunto de muestras múltiples por paciente realizamos la identificación de las alteraciones genéticas conductoras, evaluamos la composición de las firmas mutagénicas y describimos la evolución tumoral del CMTN en un contexto filogenético.

2. Justificación

El CMTN es una enfermedad altamente heterogénea a nivel clínico y molecular, sin terapia dirigida y sólo el 30% de los pacientes logran una respuesta patológica completa con la quimioterapia neoadyuvante y más del 90% de los casos fallecen debido a la aparición de metástasis y la falla terapéutica a múltiples líneas de tratamiento.

Actualmente, existe poca información acerca de la evolución tumoral del CMTN a lo largo de sus diferentes etapas de desarrollo, especialmente en los tumores clínicamente agresivos que no responden a la terapia. Por lo cual, es necesario entender como los distintos mecanismos evolutivos interactúan para remodelar el genoma tumoral durante el curso de la enfermedad, así como describir detalladamente a nivel molecular las etapas de la evolución del CMTN, desde el tumor primario hasta la formación de metástasis.

Por tal motivo, este trabajo tiene como objetivo realizar una descripción integrativa de la evolución del CMTN clínicamente agresivo durante diferentes etapas de la quimioterapia y su impacto en las variables clínicas, a fin de identificar la composición y distribución de las alteraciones genéticas y firmas mutagénicas, y para definir potenciales blancos terapéuticos.

3. Preguntas de investigación

1. ¿Qué cambios genómicos ocurren durante la evolución del CMTN?
2. ¿Existen alteraciones genéticas o vías de señalización en estas pacientes con CMTN que pueden ser candidatas a terapia dirigida a lo largo de la historia evolutiva de la enfermedad?

4. Hipótesis

Debido a que el CMTN presenta una evolución ramificada, al analizar muestras múltiples de seguimiento de las mismas pacientes, existirán alteraciones genéticas o vías de señalización en común o privadas que puedan ser candidatas a terapia dirigida.

5. Objetivos

5.1 Objetivo general

Analizar la evolución genómica del CMTN, desde la etapa del diagnóstico hasta el desarrollo de metástasis, para la identificación de alteraciones genéticas o vías de señalización candidatas a terapia dirigida.

5.2 Objetivos particulares

1. Identificar alteraciones genéticas antes y después del tratamiento en múltiples biopsias de seguimiento, para definir genes o vías de señalización candidatas a terapia dirigida.
2. Evaluar firmas mutacionales antes y después del tratamiento para identificar que procesos mutagénicos actúan en el desarrollo del CMTN.
3. Realizar reconstrucciones filogenéticas para evaluar la similitud genética entre las muestras de cada paciente y la divergencia que existe entre ellos.

6. Metodología

6.1 Información de la muestra

Se partió de un total de 97 pacientes que fueron diagnosticadas con CMTN y tratadas en el Instituto Nacional de Cancerología (INCan, México) entre abril de 2007 y abril 2010. Los tumores de mama triple negativo fueron diagnosticados por inmunohistoquímica basados en el criterio de menos del 1% de expresión en los puntajes de tinción para los RE, RP y HER2. Para nuestro estudio seleccionamos casos con una enfermedad clínicamente agresiva, definida como la ausencia de pCR a la NAC, muerte dentro de los tres años después del diagnóstico y con carcinoma ductal. Después del control de la calidad del DNA, solo 16 muestras provenientes de cuatro pacientes fueron amplificables y útiles para el análisis de secuenciación masiva de exomas, las cuales incluyeron tumores primarios sin tratamiento, tumores tratados, ganglios linfáticos y metástasis a piel, pareadas con sus respectivas muestras normales. Este estudio, contó con la aprobación del comité de ética y de investigación del INCan (protocol #016/013/IBI CEI/1021/16) y fue conducido bajo la declaración de Helsinki.

6.2 Extracción de DNA

Las muestras de DNA se obtuvieron a partir de tejidos embebidos en parafina, fijados en formaldehído (FFPE, por sus siglas en inglés) y extraídas a partir de cortes de 50 μm correspondiente a 50-200 mg de tejido. Las muestras fueron desparafinadas con baños de xilol y etanol, y digeridas toda la noche con proteinasa K. El entrecruzamiento entre el DNA y las demás biomoléculas fue revertido a través de una incubación de 90°C por una hora. La extracción del DNA fue realizada con el kit Dneasy Blood & Tissue (Qiagen, Hilden, Germany), siguiendo las recomendaciones del fabricante. El DNA fue cuantificado por fluorimetría en un equipo Qubit (Invitrogen, Carlsbad, USA). La integridad y la pureza del DNA fueron verificadas por electroforesis en gel de agarosa y espectrofotometría, respectivamente.

6.3 Preparación de bibliotecas genómicas

La preparación de bibliotecas para la secuenciación de exoma completo (WES, por sus siglas en inglés) fue realizada con el kit de Illumina Nextera Rapid Capture Exome Kit v.1.2 (Illumina, San Diego, Ca, USA), el cual cubre 37 Mb de las bases codificantes para proteínas, siguiendo las recomendaciones del fabricante. La secuenciación masiva paired-end fue realizada en el equipo de Illumina HiSeq 2500 con 2 x 150 ciclos, con una profundidad de 19.08X para los exomas germinales y 19.39X para los exomas tumorales en la Red de Apoyo a la Investigación del Instituto Nacional de Nutrición Salvador Zubirán, Ciudad de México, México.

6.4 Identificación de SNV e indels

Las lecturas de secuenciación fueron alineadas al genoma de referencia hg19 usando bwa mem v.1.7.17 (H. Li & Durbin, 2009), Anexo1. El programa GATK v.4.1.8.1 fue usado para ordenar las lecturas alineadas, marcar duplicados y realizar el recalibramiento de las bases de acuerdo a criterios internacionales (Van der Auwera *et al.*, 2013). Las SNV fueron detectadas con Mutect2 – incluido en el paquete de GATK v.4.1.8.1 - (Cibulskis *et al.*, 2013) y anotadas con ANNOVAR (K. Wang *et al.*, 2010) en la versión disponible en la web (<https://wannovar.wglab.org/>) para su interpretación biológica. Las variantes fueron filtradas bajo los siguientes criterios: (i) se seleccionaron aquellas variantes con fracción alélica <0.001 en las bases de datos de 1000 genomas, ExAC y ESP6500, (ii) SNPs comunes fueron excluidos, a menos que fueran reportados como variantes patogénicas en ClinVar, (iii) se eligieron aquellas variantes con al menos dos lecturas en la cadena sentido y antisentido en las muestras tumorales y (iv) se incluyeron aquellas variantes con fracción alélica ≥ 0.05 . Se usó la base de datos ExAC para excluir variaciones naturales no patogénicas, considerando que este depósito incluye la secuenciación de datos de 5789 exomas de latinos.

Los genes conductores fueron definidos usando dos listas de genes: Intogen v.2019 (<https://www.intogen.org/>) y la lista de Yates (Yates *et al.*, 2017) (ver sección 1.8.5 Evolución en el CM). Las mutaciones conductoras fueron definidas como

patogénicas si están clasificadas como patogénicas en ClinVar o si al menos dos de los algoritmos SIFT, PolyPhen2 o MutationTaster – información disponible en el anotado de ANNOVAR - las predicen como mutaciones deletéreas, o si son mutaciones conductoras identificadas en la base de datos de COSMIC (Tate *et al.*, 2019) y cBioportal (J. Gao *et al.*, 2013). Todos los genes conductores fueron curados manualmente verificando su evidencia experimental en los archivos BAM en el programa IGV (Thorvaldsdóttir *et al.*, 2013).

6.5 Análisis de carga mutagénica tumoral

La carga mutagénica tumoral (TMB, por sus siglas en inglés), fue definida como el número de mutaciones somáticas codificantes, incluyendo mutaciones codificantes de sustituciones de base e indels por cada megabase del genoma examinado. Para calcular la TMB por megabase, el número total de las mutaciones contadas fue dividida por el tamaño de la región codificante analizada. El tamaño del exoma utilizado para el análisis fue de 37 Mb.

6.6 Identificación de la variación del número de copias

Las CNV fueron identificadas con el software CNVkit v.0.9.9 (Talevich *et al.*, 2016) usando un panel de muestras normales con los parámetros por default. El análisis se concentró en los genes conductores del cáncer y enriquecidos en el CM descritos en las listas de Yates e Intogen, descritos previamente (ver sección 6.4 Identificación de SNV e indels). Las variantes encontradas fueron filtradas por direccionalidad, considerando la biología de los genes afectados bajo el siguiente criterio: (i) CNV con número de copias (CN, por sus siglas en Inglés) = 2 fueron excluidas, (ii) oncogenes con $CN < 2$ fueron excluidos y $CN > 2$ fueron incluidos; y (iii) genes supresores de tumores con $CN > 2$ fueron excluidos y $CN < 2$ fueron incluidos.

6.7 Análisis de las firmas mutagénicas

Los patrones mutagénicos o firmas mutagénicas fueron evaluados en el contexto de 96 trinucleótidos usando el paquete deconstructSigs v.1.9.0 de R (Rosenthal *et al.*, 2016). Para evaluar la distribución de las firmas mutagénicas en

las muestras, se realizó un análisis de agrupamiento jerárquico no supervisado calculando distancias euclidianas.

6.8 Análisis de vías de señalización

El análisis de enriquecimiento de las vías de señalización fue realizado con la herramienta de anotación funcional DAVID v.6.8 (Dennis *et al.*, 2003), usando los genes conductores de la base de datos de vías de señalización KEGG. Las vías de señalización no informativas fueron eliminadas, así como las enfermedades distintas a cáncer.

6.9 Análisis global de alteraciones accionables

Los tratamientos aprobados por la FDA para genes accionables en todos los tipos de cáncer fueron seleccionados usando la base de datos de OncoKB (Chakravarty *et al.*, 2017).

6.10 Análisis filogenético

Los árboles filogenéticos fueron construidos a partir de alteraciones específicas de SNV, indels y CNV, usando el principio de máxima parsimonia (Fitch, 1971). Las variantes fueron codificadas con 0 y 1 para indicar la ausencia y presencia de las mutaciones, respectivamente, y procesadas en mesquite (<http://www.mes-quiteproject.org/>) para generar archivos nexus. Los árboles filogenéticos fueron analizados en PAUP v.4.0 (<http://paup.phylosolutions.com/>) usando un método heurístico y enraizado con las muestras normales. La confiabilidad de los árboles fue evaluada con el método de bootstrap (Efron *et al.*, 1996). Los árboles generados fueron guardados en formato Newick y visualizadas con el programa de FigTree v.1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

7. Resultados

7.1 Las pacientes con cáncer de mama triple negativo exhiben gran variedad de escenarios clínicos y genómicos

Para este estudio, se revisó la base de datos de los expedientes de las pacientes diagnosticadas con CMTN en el INCan durante el periodo abril de 2007 a abril de 2010. Se seleccionaron 16 muestras de FFPE de cuatro pacientes con enfermedad agresiva y con DNA de alta calidad para realizar la secuenciación masiva de exoma completo. Cuatro escenarios clínicos fueron representados por las muestras: (i) tumor primario sin tratamiento, (ii) tumor tratado, (iii) ganglios linfáticos y (iv) metástasis a piel. Ninguna paciente presentó la pCR y todas fallecieron dentro de los tres primeros años a partir del diagnóstico (Figura 8a,b).

La paciente TNBC11 fue una mujer diagnosticada a la edad de 53 años con un tumor de mama localmente avanzado. Se obtuvo una primera biopsia del tumor primario y tuvo una variante patogénica en *TP53* y *EP300*, así como amplificaciones de *MCL1* e *IGF1R* y una delección de *BAP1* (Figura 8 y 9; Anexo 2 y 3). Posteriormente, una segunda biopsia fue extraída del mismo tumor primario en el segundo mes del diagnóstico y la paciente fue tratada con quimioterapia. Esta segunda muestra tuvo diferentes alteraciones conductoras de las cuales solo compartió la amplificación *IGF1R* con la primera biopsia. En el noveno mes, se le realizó una mastectomía y se obtuvieron dos biopsias del ganglio linfático. Ambos nodos linfáticos presentaron una delección en *RB1* y una TMB baja. El ganglio linfático 1 mostró amplificaciones de *NOTCH3*, *AKT2* y *CCNE1*, y el ganglio linfático 2 tuvo una pérdida en el número de copias en *CDKN1B*. En el decimocuarto mes, después de 5 meses de radioterapia, una última mastectomía fue realizada y se obtuvo una muestra de recurrencia, la cual mostró una composición de genes conductores muy similar a la detectada en el tumor primario 1, incluyendo mutaciones en *EP300* y *TP53*, así como amplificaciones de *MCL1*, *CCND1*, *MDM4*, *IGF1R*, *MYC*, *AKT3*, *CDK6*, *JAK2* y *KRAS*. La muestra de recurrencia también compartió la delección de *BAP1* con el tumor primario 1. Globalmente, las lesiones

neoplásicas tuvieron un total de 57 mutaciones conductoras: 41 CNV, 15 SNV y 1 indel. La paciente CMTN11 murió después del mes 16 a partir del diagnóstico.

La paciente CMTN18 fue una mujer diagnosticada a la edad de 54 años con dos muestras disponibles, un tumor primario sin tratamiento y un tumor tratado. El tumor primario sin tratamiento fue obtenido en el primer mes del diagnóstico. Después de 16 meses de tratamiento con quimioterapia neoadyuvante y radioterapia, se realizó una cirugía para obtener el tumor tratado. Ambas muestras mostraron un patrón similar de alteraciones genéticas y una TMB > 7.5 mutaciones/MB de la región codificantes. Nueve deleciones fueron compartidas, incluyendo *RB1*, *ARID1A*, *BAP1*, *BRCA1*, *CDKN1B*, *PTEN*, *CDKN2A*, *CDKN2B* y *PBRM1*. La paciente CMTN18 tuvo una carga total de 28 alteraciones conductoras 19 CNV y 9 SNV, falleció en el mes 26 a partir del diagnóstico.

La paciente CMTN41 fue diagnosticada a la edad de 45 años y tuvo tres muestras: un tumor tratado, un ganglio linfático y metástasis a piel. El tumor primario sin tratamiento no estuvo disponible. Después de doce meses de tratamiento con la quimioterapia neoadyuvante, se realizó una mastectomía y se obtuvieron dos muestras: un tumor tratado y una metástasis a ganglio linfático. El tumor tratado mostró varias CNV en genes conductores, incluyendo amplificaciones de *MCL1*, *TERT*, *CCND1*, *AKT1*, *MYC*, *EGFR* y *FGFR3*; deleciones de *CDK1B*, *CDKN2A* y *CDKN2B*; una mutación puntual en *PIK3CA*; y una TMB de 5.43 mutaciones/ MB de la región codificante. Por el contrario, la metástasis a ganglio linfático tuvo solamente cuatro alteraciones patogénicas y una TMB de 2.75 mutaciones/ MB de la región codificante. En el veintavo mes, se llevó a cabo una cirugía y una biopsia de metástasis a piel fue obtenida. La muestra de metástasis a piel mostró un patrón similar de alteraciones en genes conductores con el tumor tratado, incluyendo una mutación puntual en *PIK3CA*, las amplificaciones de *CCND1*, *EGFR*, *FGFR3*, y deleciones de *CDK1B*, *CDKN2A* y *CDKN2B*. Esta metástasis a piel también presentó la pérdida de *PTEN*. Las tres muestras compartieron SNV en *JAK3*, así como las amplificaciones en *TERT* y *AKT1*. La paciente CMTN41 tuvo un total de 33 genes conductores 26 CNV y 7 SNV; falleció 26 meses después del diagnóstico.

Por otro lado, la paciente CMTN84 fue diagnosticada a los 60 años y tuvo 6 muestras. El tumor primario sin tratamiento fue extraído en el diagnóstico y tuvo amplificaciones de *MCL1* y *NOTCH3*, y la delección de *BRCA1*. Después de 9 meses de quimioterapia y radioterapia se realizó una cirugía y se obtuvo una biopsia de ganglio linfático. Esta biopsia tuvo amplificación de *TERT*, delección de *BRCA1* y una mutación patogénica en *TSC2*. El tumor tratado y la metástasis a piel fueron obtenidos en el mes 16, y ambos mostraron la amplificación de *MCL1*, *TERT* y *MDM4*. Finalmente, una cirugía adicional fue realizada en el mes 24 y se obtuvieron dos ganglios linfáticos, codificados como ganglio linfático 2 y ganglio linfático 3. Ambos ganglios linfáticos mostraron la delección de *RB1*, así como la amplificación de *MCL1*, *NOTCH3*, *AKT2* y *CCNE1*. Interessantemente, todas las lesiones, excluyendo la muestra de ganglio linfático 1, tuvieron una amplificación en *MCL1*. La paciente CMTN84 presentó 39 mutaciones conductoras: 28 CNV, 9 SNV y 2 indes; falleció en el mes 30 después del diagnóstico. Todas las pacientes analizadas en este estudio fueron detectadas en un estadio avanzado y ninguna de ellas tuvo una respuesta patológica completa con la quimioterapia neoadyuvante, todas fallecieron antes de los tres años después de su diagnóstico.

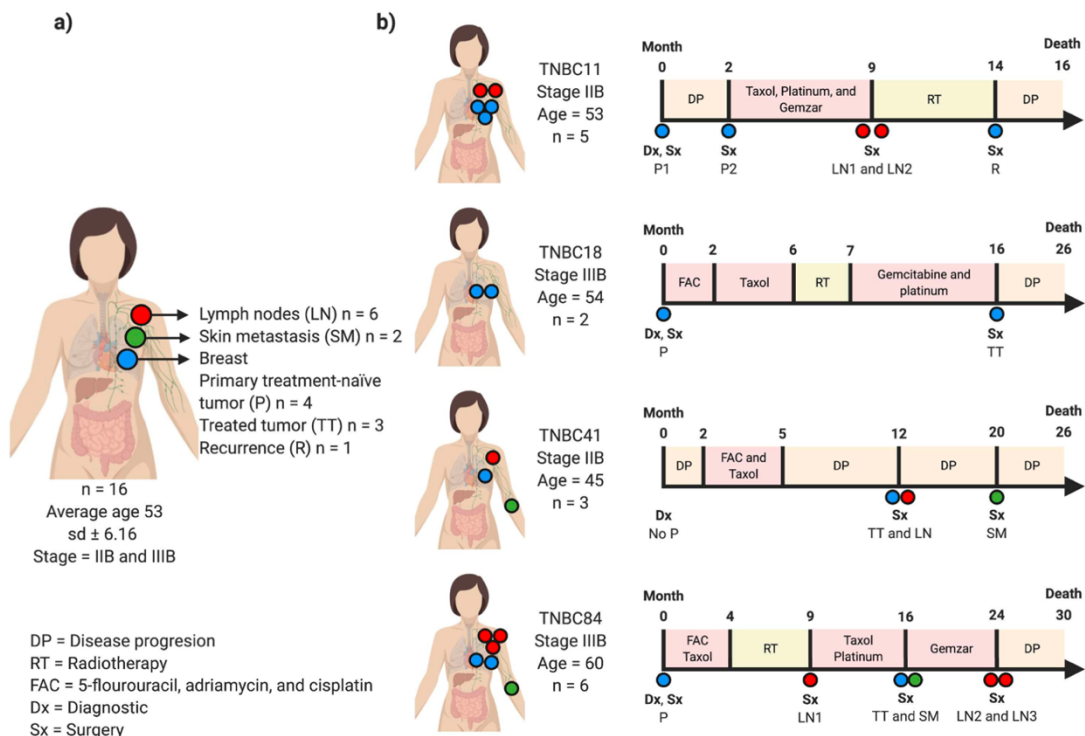


Figura 8. Historia clínica de las pacientes con cáncer de mama triple negativo. a) La cohorte de CMTN con 16 muestras de cuatro pacientes incluye tumores antes y después del tratamiento, ganglios linfáticos y metástasis a piel. b) La línea del tiempo describe las características clínicas de todas las pacientes analizadas y las muestras tomadas durante la progresión de la enfermedad. Las cajas de colores muestran los diferentes tratamientos recibidos por las pacientes (azul: seno; rojo: metástasis a ganglio linfático; verde: metástasis a piel).

7.2 Las pacientes con cáncer de mama triple negativo muestran una amplia heterogeneidad genética

Nuestro análisis de secuenciación identificó una amplia heterogeneidad genética interpaciente. Se identificaron un total de 2264 SNV en las 16 muestras analizadas, incluyendo 611 mutaciones sinónimas, 1539 mutaciones no sinónimas y 114 deleciones. La TMB fue amplia, con un intervalo de 0.4-8.9 mutaciones por Mb, con un promedio de 3.8 mutaciones/Mb de la región codificante (figura 9). Cada paciente mostró una TMB única. Por ejemplo, para la paciente CMTN11 la muestra de recurrencia tuvo la TMB con 4.72 mutaciones/Mb de la región codificante, mientras las pacientes CMTN18, CMTN41 y CMTN84 la TMB más alta fue identificada en las muestras de los tumores tratados con 8.86, 5.43 y 5.08 mutaciones/Mb de la región codificante. Por otra parte, las muestras de ganglios linfáticos mostraron la TMB más baja menos de 1.5 mutaciones/Mb de la región codificante en dos de cuatro pacientes analizadas (CMTN11 y CMTN84), menos para la muestra de ganglios linfáticos 1 de la paciente CMTN que tuvo 3.67 mutaciones / MB de la región codificante.

En total, se identificaron 157 mutaciones en 67 genes conductores, de los cuales 114 fueron CNV (promedio de 6.5, intervalo de 2-15 por muestra) y 43 fueron SNV (promedio 2.5, intervalo de 0-7 por muestra). Las mutaciones más prevalentes en nuestra cohorte fueron las CNVs, incluyendo aquellas en *MCL1* (50%), *RB1* (43.75 %), *TERT* (37.5%), *ARID1A*, *BAP1*, *BRCA1*, *CCND1*, *CDKN1B* y *PTEN*, todas ellas presentes en el 31.25% de las muestras analizadas.

La ontología de los genes reveló el enriquecimiento de 8 vías de señalización importantes: factores de transcripción (22.4%), remodeladores de la cromatina (20.9%), ciclo celular (19.4%), PI3K/mTOR (16.4%), RAS/MAPK (11.9%), JAK-STAT (2.9%) y adhesión celular (1.5%). La mayoría de los genes conductores y vías de señalización fueron compartidas por la paciente CMTN18 (especialmente tumor primario sin tratamiento y tumor tratado) y la paciente CMT41 (tumor tratado y metástasis a piel). Por el contrario, las muestras más heterogéneas fueron para las pacientes CMTN11 y CMTN84, las cuales compartieron pocas mutaciones y vías de señalización (figura 9). Estos resultados demuestran que existe una amplia heterogeneidad intertumoral en los tumores de mama triple negativo.

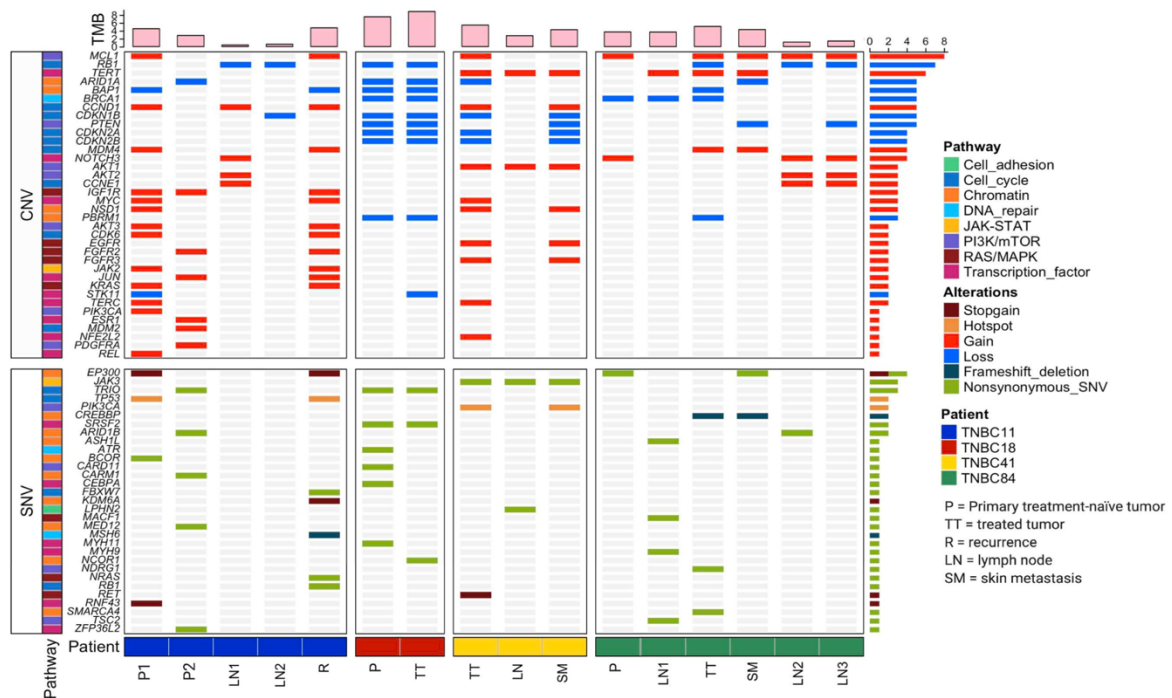


Figura 9. Composición alélica y alteraciones en el número de copias en genes conductores. Las muestras fueron segmentadas por pacientes y ordenadas de manera cronológica. La TMB es indicada en la parte superior de la gráfica con barras de color rosa. Las alteraciones en los genes conductores se clasificaron en dos categorías: SNV (en el panel más bajo) y CNV (en el panel medio). Las vías de señalización en las cuales están involucradas las mutaciones conductoras, así como el tipo de mutaciones se muestran en diferentes colores.

7.3 Procesos mutagénicos actúan en diferentes estadios del cáncer de mama triple negativo

Para identificar qué procesos mutagénicos actúan en el CMTN, se realizó el análisis de las firmas mutagénicas. Se identificaron 11 de 30 firmas mutagénicas previamente descritas por (Alexandrov *et al.*, 2013), con una media de 3.5 firmas mutagénica por paciente (intervalo 1-5). Cinco de estas firmas han sido previamente identificadas en CM, mientras las seis restantes es la primera vez que se describen en CM (figura 10). La firma mutagénica S01 (asociada a la edad) fue prevalente en todas las muestras y contribuyó a la tasa más alta de mutaciones. La firma mutagénica S03 (asociada a deficiencia en la recombinación homóloga) fue detectada en el 81%, la firma S06 (asociada a la deficiencia de bases mal apareadas) en el 44% y la firma S12 (etiología desconocida) en el 31% de las muestras. Otras firmas fueron encontradas en menos del 20% de las muestras.

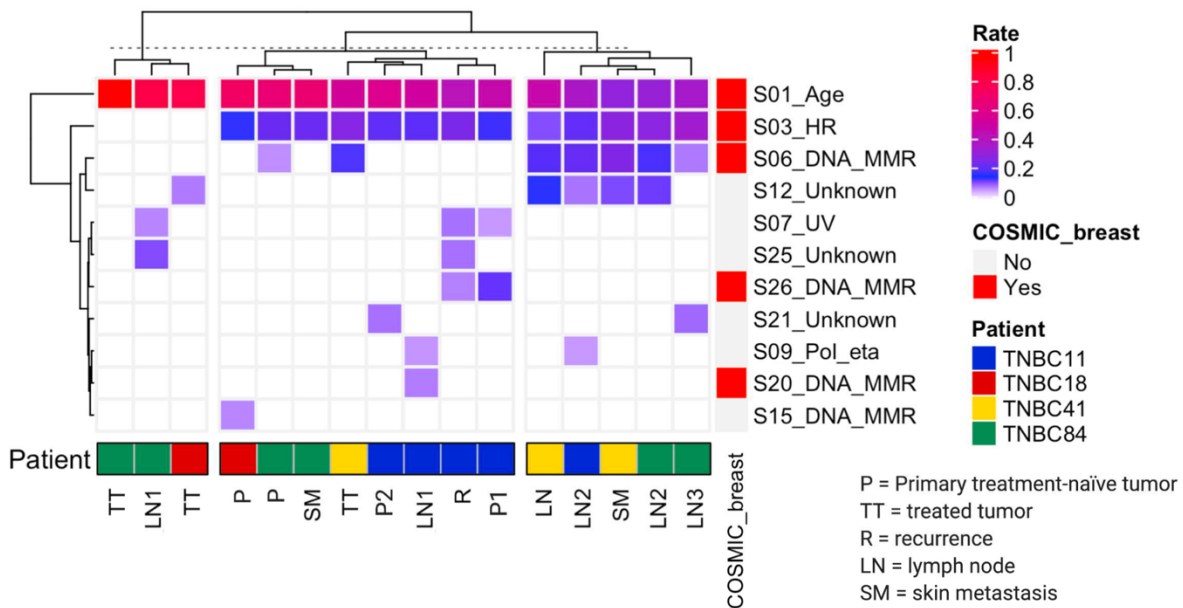


Figura 10. Perfil de firmas mutagénicas en las pacientes con cáncer de mama triple negativo. Las muestras se ordenaron de acuerdo al agrupamiento jerárquico no supervisado, usando distancias euclidianas. La banda de COSMIC describe las firmas encontradas en la base de datos de COSMIC Sanger y la escala de colores de la derecha muestra la frecuencia relativa de las firmas en cada muestra. La tasa se expresa como la frecuencia relativa de las mutaciones asociadas con las firmas correspondientes. Las firmas mutagénicas previamente identificadas en las muestras de cáncer de mama en la base de datos de COSMIC se presentan en color rojo. Las pacientes están codificadas con colores. El tejido de origen de las muestras se muestra en la esquina inferior derecha y para cada paciente.

Para identificar qué procesos mutagénicos se activan entre las muestras, se realizó un análisis de agrupamiento jerárquico no supervisado, calculando distancias euclidianas para determinar la distribución de las firmas mutagénicas de todas las muestras. Este análisis clasificó a las muestras en tres grupos: (i) muestras que comparten una firma mutagénica en común (3/16), (ii) muestras que comparten dos firmas mutagénicas en común y (iii) muestras que comparten tres o más firmas en común. Cada grupo tuvo características únicas. Por ejemplo, el grupo 2 compartió dos firmas mutagénicas (S01 y S03, asociadas a la edad y defectos en la reparación del ADN por recombinación homóloga, respectivamente) en diferentes muestras de cada paciente. Por otra parte, el grupo 3 incluyó las firmas S01 y S03, pero a su vez encontramos que estos tumores adquirieron dos firmas adicionales, S06 (asociada a mecanismos de reparación del ADN por bases mal apareadas) y S12 (de etiología desconocida). Interesantemente, este grupo incluyó ganglios linfáticos y metástasis a piel de tres pacientes (CMTN11, CMTN41 y CMTN84). Estos hallazgos demuestran que hay algunas firmas mutagénicas importantes al inicio del cáncer y otras que se adquieren en etapas avanzadas de la enfermedad.

7.4 El cáncer de mama triple negativo presenta divergencia temprana

Para determinar la historia evolutiva de los tumores de CMTN se realizó un análisis filogenético. Se clasificaron las SNVs, indels y CNVs en tres grupos: (i) troncales (mutaciones encontradas en todas las muestras provenientes del mismo paciente), (ii) compartidas (mutaciones encontradas en más de una muestra del mismo paciente, pero no en todas) y (iii) privadas (mutaciones encontradas en una sola muestra del mismo paciente) (figura 11a). La filogenia fue estimada con el principio de máxima parsimonia (Fitch, 1971) y el método de bootstrapping (Efron et al., 1996) para evaluar la confiabilidad del árbol filogenético.

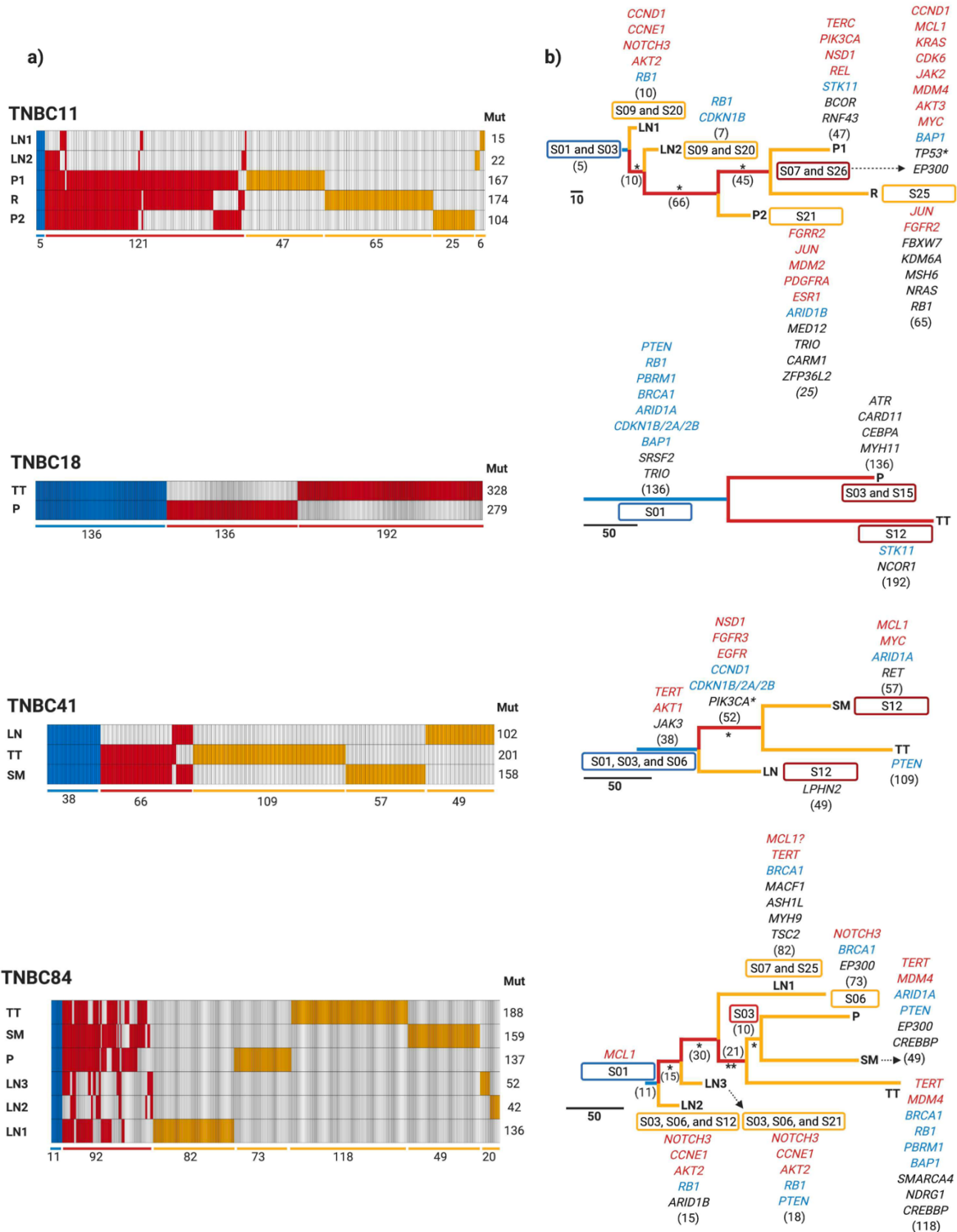


Figura 11. Árboles filogenéticos de las pacientes con cáncer de mama triple negativo. a) Las gráficas de distribución describen las mutaciones troncales (color azul), compartidas (color rojo) y privadas (color amarillo), encontradas en cada tumor. El número de mutaciones se muestra en la parte inferior y el total de las mutaciones para cada muestra, a la derecha. b) Los árboles filogenéticos describen la relación de las muestras de cada

paciente. La longitud de las ramas es proporcional al número de las mutaciones detectadas. Las alteraciones en los genes conductores se presentan codificadas por tipo en distintos colores: SNVs, color negro; amplificaciones, rojo; y deleciones en azul. Las firmas mutagénicas se muestran para cada muestra. P: tumor primario sin tratamiento, TT: tumor tratado, R: recurrencia, GL: ganglio linfático y MP: metástasis a piel. * = valor de bootstrap >95% y ** = valor e bootstrap >65% y <95%.

Cada paciente tuvo una combinación única de mutaciones troncales, compartidas y privadas. Las muestras de recurrencia y tumores tratados presentaron el número más alto de mutaciones privadas y las metástasis a ganglios linfáticos tuvieron el número más bajo de mutaciones privadas. Estos hallazgos muestran una heterogeneidad genética intrapaciente importante en todas las lesiones neoplásicas. Por ejemplo, la paciente CMTN11 tuvo cinco mutaciones troncales, 121 mutaciones compartidas y cada muestra tuvo un número diferente de mutaciones privadas, siendo la muestra de recurrencia la que tuvo el número más alto de mutaciones privadas. Por otra parte, la paciente CMTN18 acumuló 136 mutaciones troncales, 136 y 192 mutaciones privadas en el tumor primario sin tratamiento y en el tumor tratado, respectivamente. La paciente CMTN41 tuvo 38 mutaciones troncales, 66 mutaciones compartidas y la muestra del tumor tratado fue la que tuvo el mayor número de mutaciones privadas con 109. La paciente CMTN84 tuvo 11 mutaciones troncales, 92 mutaciones compartidas y la muestra del tumor tratado fue la más mutada, mostrando 118 mutaciones.

Con base en la estructura del árbol filogenético, tres pacientes (CMTN11, CMTN41 y CMTN84) tuvieron el mismo patrón evolutivo (figura 11b). El tumor primario sin tratamiento, el tumor tratado, la recurrencia y la metástasis a piel fueron agrupados en el mismo clado y tuvieron el ancestro común más reciente (MRCA, por sus siglas en inglés). La metástasis a piel y el tumor primario sin tratamiento de la paciente CMTN84 fueron agrupadas en el mismo clado debido a que estas dos muestras compartieron 76 mutaciones. Interesantemente, los ganglios linfáticos fueron agrupadas en otro clado, sugiriendo que tuvieron una divergencia temprana. Por otra parte, los ganglios linfáticos tuvieron una media de 4.6 conductores alterados, indicando posiblemente un cuello de botella evolutivo. Adicionalmente, se identificaron 136 mutaciones troncales en las dos muestras obtenidas de la paciente CMTN18. Una limitante del estudio fue que no se pudo evaluar el patrón evolutivo

en esta paciente debido a que al menos se requieren tres muestras para este análisis. Estos análisis permiten entender la evolución del CMTN para identificar modos de evolución, selección de las mutaciones conductoras en la etapa avanzada de la enfermedad, así como la agresividad de la misma.

7.5 Implicaciones clínicas potenciales de las mutaciones detectadas en el cáncer de mama triple negativo

Para definir si los tumores de estas pacientes que no respondieron a ninguna terapia presentaban mutaciones específicas asociadas con la respuesta a tratamientos de medicina de precisión se realizó una búsqueda puntual en la base de datos de OncoKB empleando datos de fármacos aprobados por la Federal Drug Administration de Estados Unidos (FDA) (Chakravarty et al., 2017). La idea central de este análisis fue conocer si estas pacientes hubieran tenido opciones terapéuticas que mejoraran su pronóstico. Los tumores de todas las pacientes presentaron mutaciones específicas tratables con fármacos dirigidos. Se identificaron 17 blancos terapéuticos potenciales (3-8 blancos por pacientes) distribuidos en 7 vías de señalización. Las dos vías de señalización con el número más alto de blancos accionables fueron RAS/MAPK y PIK3/mTOR (figura 12).

Posteriormente, se agruparon las alteraciones accionables basadas en las mutaciones detectadas por tipo de tumor y se observaron siete grupos diferentes de tumores y su tratamiento asociado. Por ejemplo, las pacientes con CMTN que presentan mutaciones en *PIK3CA* (alpelisib, fulvestrant), *AKT1* (AZD5363) y *ESR1* (AzD9496, fulvestrant), así como las amplificaciones en *EGFR* (afitinib, erlotinib, gefitinib), podrían ser tratadas con fármacos que son usados en otros subtipos de tumores de mama y cáncer de pulmón de células no pequeñas, respectivamente (figura 12). Esto abre una nueva ventana terapéutica en CMTN basada en los genes candidatos potenciales para la precisión en los tratamientos terapéuticos.

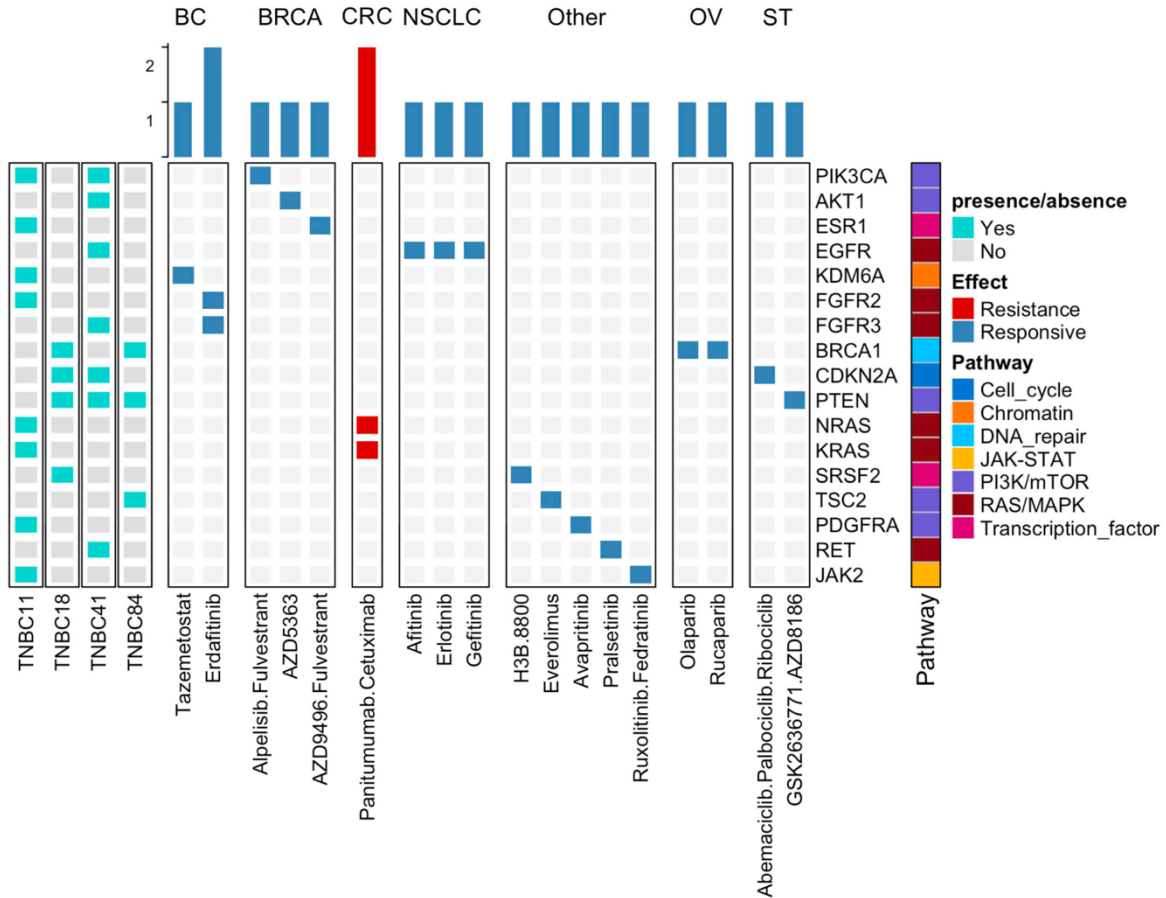


Figura 12. Alteraciones accionables en cáncer de mama triple negativo de acuerdo con la base de oncoKB. Se presentan los fármacos segmentados de acuerdo con el tipo de tumor para el cual cuentan con aprobación de uso por la FDA. Las vías de señalización se indican a la derecha y la identificación de los pacientes a la izquierda. BC: Bladder Cancer; BRCA: Breast Cancer; CRC: Colorectal Cancer; NSCLC: Non-Small Cell Lung Cancer; OV: Ovarian Cancer; and ST: Solid Tumors.

8. Discusión

El propósito de este estudio fue analizar de manera integrativa las características moleculares y la evolución de los tumores de pacientes con CMTN clínicamente agresivo. El CMTN muestra el peor pronóstico de todos los subtipos de CM, con aspectos moleculares que aún no han sido revelados en su totalidad. Para este estudio seleccionamos a cuatro pacientes con varias muestras de seguimiento de DNA de alta calidad. Para evaluar las alteraciones en la región codificante completa del genoma tumoral se empleó una estrategia de secuenciación de exoma completo. Con esta aproximación fue posible evaluar la heterogeneidad tumoral, el perfil de las firmas mutagénicas, las alteraciones en

genes conductores y la evolución tumoral en un total de 16 muestras con seguimiento clínico. Todas las pacientes fueron diagnosticadas en un estadio localmente avanzado y cada una presentó escenarios clínicos únicos. Estos escenarios concuerdan con la complejidad reconocida de esta neoplasia y la dificultad para el desarrollo de tratamientos específicos, dada la amplia heterogeneidad genética observada entre las pacientes (Bianchini *et al.*, 2016). En nuestra cohorte, ninguna de estas pacientes logró una pCR con NAC y todas fallecieron dentro de los primeros tres años después del diagnóstico, después de múltiples líneas de tratamiento. Nuestros hallazgos son consistentes con reportes previos sobre la tasa de mortalidad en América Latina y a nivel mundial (X. Li *et al.*, 2017; Reynoso-Noverón *et al.*, 2017).

A nivel molecular, descubrimos una gran heterogeneidad genética inter- e intrapaciente en términos de CNV y SNV, lo cual concuerda con estudios previos (Barry *et al.*, 2018; Ciriello *et al.*, 2013; Gerstung *et al.*, 2020; Ullah *et al.*, 2018; Yates *et al.*, 2017). Las SNVs en *TP53* y amplificaciones en *MCL1* encontradas en tumores primarios sin tratamiento de dos pacientes (CMTN11 y CMTN84), sugieren que estos son eventos tempranos en la carcinogénesis del CMTN. *TP53* es el gen más frecuentemente mutado en el CMTN y, en concordancia con nuestros hallazgos, se ha reportado como una alteración que ocurre en etapas tempranas del desarrollo del CMTN (Barry *et al.*, 2018; Shah *et al.*, 2012; Yates *et al.*, 2017). En cambio, las amplificaciones en *MCL1* han sido asociadas con resistencia a la NAC y están enriquecidas frecuentemente en estadios avanzados de la enfermedad (Balko *et al.*, 2014), lo cual podría explicar por qué las pacientes CMTN11, CMTN41 y CMTN84 no tuvieron pCR, aunque no se descartan otras causas como alteraciones no detectadas en este estudio.

Las deleciones de *RB1* fueron predominantemente encontradas en los ganglios linfáticos de dos pacientes (CMTN11 y CMTN84); las deleciones en *PTEN* en tres (CMTN18, CMTN41 y CMTN84) de cuatro pacientes, principalmente en muestras de metástasis a piel. Estos hallazgos son similares a los de Bertucci y colaboradores, quienes identificaron cinco genes, incluyendo *RB1* y *PTEN*

enriquecidos en las metástasis de CMTN (Bertucci *et al.*, 2019). Las amplificaciones de *TERT* fueron detectadas en tumores tratados, ganglios linfáticos y metástasis a piel de dos pacientes (CMTN41 y CMTN84), así como las amplificaciones de *AKT2* y *CCNE1* en ganglios linfáticos de dos pacientes (CMTN11 y CMTN84). Estas CNVs han sido reportadas en otros estudios de tumores de mama en estadios avanzados y metástasis. Por ejemplo, Watkins y colaboradores encontraron que *TERT* y *CCNE1* fueron cambios subclonales que aparecieron en estadios intermedios y avanzados durante la progresión de 22 tumores diferentes, incluyendo CM (Watkins *et al.*, 2020). De manera global, los tumores de mama triple negativo presentaron una amplia heterogeneidad genética con varios genes involucrados en diferentes etapas de la carcinogénesis del CMTN (figura 13a).

Los pacientes con CMTN presentaron una TMB heterogénea, con un promedio de 3.8 mutaciones/megabase de la región codificante (intervalo 0.4-8.9). Estos resultados son comparables con estudios previos (Kandoth *et al.*, 2013; Rojas-Jiménez *et al.*, 2020). Los tumores tratados y la recurrencia en nuestra cohorte tuvieron la TMB más alta en comparación con el tumor primario sin tratamiento, ganglios linfáticos y metástasis a piel. Estas diferencias en la TMB entre los diferentes sitios tumorales pueden ser el resultado del desarrollo tumoral y del efecto mutagénico de la quimioterapia. De hecho, ha sido demostrado que la NAC altera el perfil genético de los tumores y promueve la progresión de la enfermedad incrementando la TMB en pacientes que no responden al tratamiento (Pich *et al.*, 2019; Szikriszt *et al.*, 2016). Por el contrario, los ganglios linfáticos muestran la TMB más baja, la cual puede deberse a un efecto de cuello de botella evolutivo durante la diseminación de las células tumorales debido a las restricciones del microambiente tumoral, un número reducido de células invasoras (una baja diversidad genética), o a una celularidad tumoral reducida en estas muestras.

El análisis de las firmas mutagénicas reveló la presencia de la firma S01 en el 100% de las muestras; S03 en el 83%; S06 y S12 en menos del 50%; y otras firmas mutagénicas en menos del 20%. Un análisis de agrupamiento jerárquico no supervisado usando distancias euclidianas identificó tres grupos de muestras de

acuerdo con sus firmas mutagénicas en común. El grupo dos compartió dos firmas mutagénicas, S01 y S03, ambas detectadas en varias muestras tumorales. En cambio, el grupo tres mostró más firmas mutagénicas en común, las cuales incluyeron S01, S03, S06 y S12. Las lesiones metastásicas de los ganglios linfáticos y a piel fueron los sellos distintivos del grupo tres. En general, el análisis de las firmas mutagénicas reveló que las lesiones avanzadas tratadas con fármacos antitumorales son influenciadas por un ambiente mutagénico más complejo que las lesiones primarias sin tratar. Cinco de las 11 firmas mutagénicas identificadas en nuestra cohorte han sido asociadas con el CM, incluyendo S01, S03, S06, S20 y S26 (Nik-Zainal & Morganella, 2017), mientras que las seis restantes no han sido detectadas en CM. En efecto, la firma S01 y S03 fueron detectadas en casi todas las muestras, mientras que las firmas S06 y S13 fueron encontradas en los ganglios linfáticos y metástasis a piel (figura 13b). La firma mutagénica S01 se presenta en todos los tipos de tumores malignos humanos debido a su asociación con la desaminación hidrolítica espontánea de la 5-metilcitosina acumulada a través del tiempo, por la edad. La firma S03 ha sido identificada en el 20% de todos los subtipos moleculares de CM; sin embargo, en el CMTN esta firma constituye más del 70% de todos los casos (Lefebvre *et al.*, 2016; Nik-Zainal & Morganella, 2017). Las firmas mutagénicas S06, S20 y S26, asociadas al déficit en mecanismos de reparación de bases mal apareadas, fueron detectadas en nuestro estudio y son prevalentes en CM (Lefebvre *et al.*, 2016; Nik-Zainal & Morganella, 2017). Otros estudios como el Gerstung y colaboradores han reportado que las firmas asociadas a deficiencias en los mecanismos de reparación de bases mal apareadas fueron desarrolladas en los estadios avanzados de 38 diferentes tipos de cáncer, incluyendo el CM (Gerstung *et al.*, 2020), y se asociaron con tumores de mama de alto grado (Cheng *et al.*, 2020). Por el contrario, no hay evidencia sobre la prevalencia de la firma S12 en CM. Esta firma ha sido detectada en el hepatocarcinoma celular, el subtipo más común de cáncer de hígado. Por otra parte, se ha reportado que los pacientes con diferentes tipos de neoplasia que no logran la pCR después del tratamiento con la NAC basada en sales de platino presentan algunas firmas asociadas con la exposición a las sales de platino (Pich *et al.*, 2019),

firmas S04, S24 y S29, las cuales no fueron observadas en nuestra cohorte. Sin embargo, no descartamos la prevalencia de las firmas asociadas con la resistencia a la NAC basada en sales de platino, las cuales podrían ser evidenciadas experimentalmente mediante secuenciación de genoma completo.

Las células del CM se diseminan a otros tejidos a través de vías linfáticas y hematológicas (Chitty *et al.*, 2018). Entendiendo los patrones evolutivos se podría mejorar la comprensión de la progresión de los tumores y la invasión a otros tejidos. En este estudio, se observó un patrón evolutivo definido en tres de cuatro pacientes analizados. El tumor primario sin tratamiento, el tumor tratado, la recurrencia y la metástasis a piel se agruparon en el mismo clado, indicando que comparten el MRCA. Durante muchos años se ha hipotetizado que los ganglios linfáticos axilares constituyen un reservorio de células malignas que les permite diseminarse después del tratamiento a órganos distantes (Halsted, 1894; Patey & Dyson, 1948). Sin embargo, nuestro estudio mostró que la metástasis a piel proviene directamente del tumor primario sin tratamiento, a pesar del estadio de los ganglios linfáticos. Adicionalmente, los ganglios linfáticos formaron un clado independiente con pocas mutaciones compartidas entre el tumor primario sin tratamiento y otras lesiones, sugiriendo que la diseminación podría ocurrir de manera temprana en estos tejidos. Este mismo patrón de invasión en etapas tempranas ha sido reportado en otros estudios (Barry *et al.*, 2018; Ullah *et al.*, 2018; Venet *et al.*, 2020). Asimismo, distintos estudios moleculares y clínicos del CMTN y del CM han mostrado la diseminación a ganglios linfáticos a partir de tumores pequeños en etapa T1 (<2 cm) y la ocurrencia de diseminación metastásica en los estadios tempranos del crecimiento tumoral (Chen *et al.*, 2013; Z. Hu *et al.*, 2020; X.-X. Wang *et al.*, 2016). Por otra parte, los ganglios linfáticos tuvieron pocos genes conductores alterados, sugiriendo un cuello de botella evolutivo. Este cuello de botella podría ser causado por la diseminación de un pequeño número de células invasivas en los ganglios linfáticos, lo cual podría resultar en una pérdida de la diversidad genética. La heterogeneidad subclonal reducida en los ganglios linfáticos después de la NAC ha sido reportada en pacientes con CM, cáncer colorrectal y de pulmón (Z. Hu *et al.*, 2020). Tomando en cuenta lo anterior, nuestros hallazgos soportan el modelo de

metástasis temprana e implican una rápida progresión metastásica de lesiones tempranas en estas pacientes con neoplasias agresivas (Hunter *et al.*, 2018).

El único tratamiento sistémico aprobado para las pacientes con CMTN es la quimioterapia neoadyuvante. En las pacientes analizadas encontramos alta cantidad de alteraciones genéticas en 67 genes conductores (2-15 SNV y 0-7 CNV) distribuidas de manera heterogénea en las diferentes lesiones tumorales de cada paciente. Esta amplia diversidad genética de alteraciones que aumentan la eficacia biológica es uno de los principales retos de la terapia contra el CMTN (Bianchini *et al.*, 2016). Dado el alto grado de heterogeneidad genética del CMTN, es posible usar los perfiles moleculares para identificar genes candidato para la terapia dirigida. Nuestros análisis identificaron 17 alteraciones potencialmente tratables agrupadas en siete vías de señalización primarias, siendo las más importantes en RAS/MAPK (*EGFR*, *FGFR2*, *FGFR3*, *NRAS*, *KRAS* y *RET*) y PIK3/mTOR (*PIK3CA*, *AKT1*, *PTEN*, *TSC2* y *PDGFRA*), las cuales son consistentes con estudios previos de nuestro grupo de trabajo y otros (Rojas-Jiménez *et al.*, 2020; Sanchez-Vega *et al.*, 2018; Yates *et al.*, 2017). Se identificaron siete grupos de tumores con tratamientos disponibles contra las mutaciones específicas detectadas en las muestras que analizamos. Aunque es necesario que el reposicionamiento de fármacos para su uso en pacientes con CMTN sea confirmado mediante ensayos clínicos aleatorizados, algunos de estos tratamientos son también usados en otros subtipos de CM. En efecto, nuestro análisis identificó alteraciones genéticas potencialmente importantes que podrían ser usadas en el tratamiento del CMTN. Es necesario que ensayos clínicos demuestren la utilidad de estos fármacos en el CMTN, Figura 12.

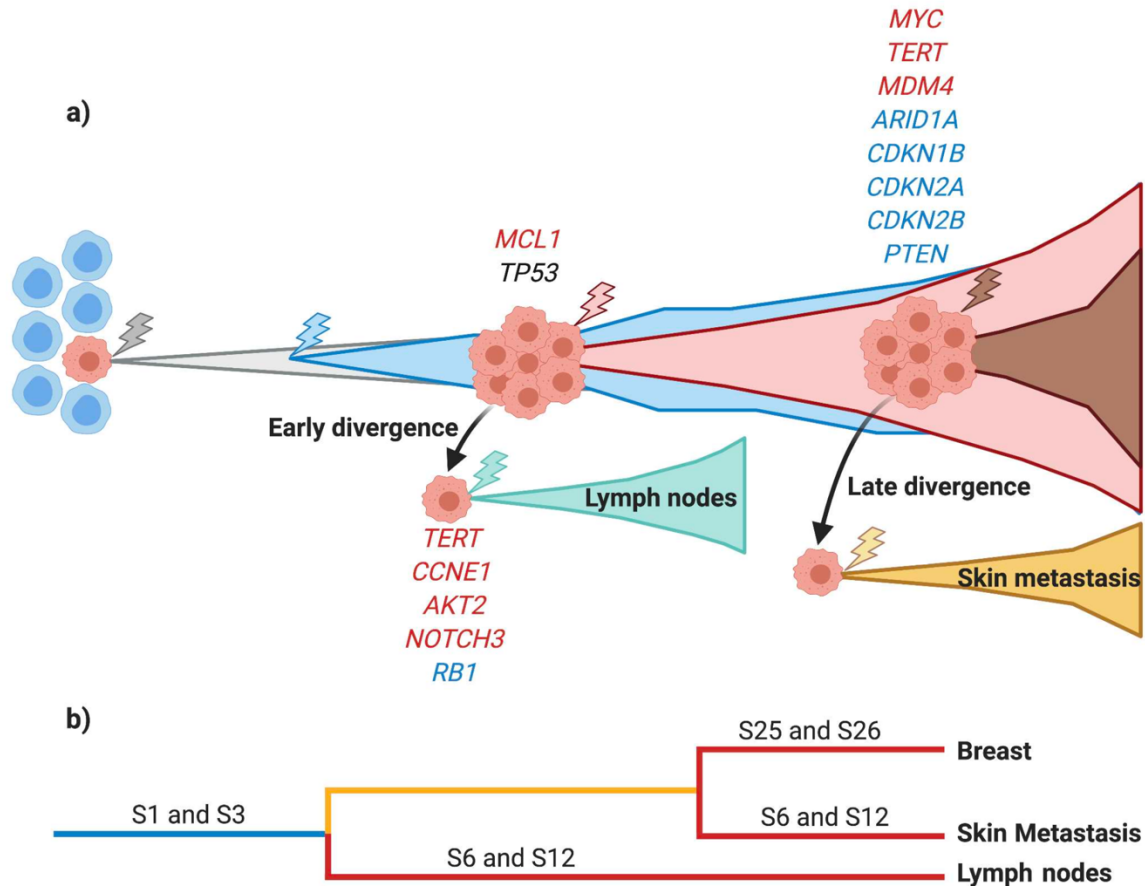


Figura 13. Resumen de las conclusiones claves de nuestro estudio. a) Evolución tumoral de CMTN observado en nuestra cohorte. Se identificó una divergencia temprana a ganglios linfáticos y tardía a metástasis a piel. Los genes conductores con SNV (color negro), las amplificaciones (color rojo) y deleciones (color azul) están representadas en los diferentes estadios durante la carcinogénesis del CMTN. b) Árbol filogenético de las principales firmas mutagénicas adquiridas en cada muestra. Las firmas mutagénicas troncales se indican en color azul, las compartidas en amarillo y las privadas en rojo.

Nuestro estudio proporciona un análisis detallado de cuatro pacientes con CMTN. Los tumores fueron caracterizados molecularmente y la amplia heterogeneidad genética que mostraron podría impactar en la falta de respuesta al tratamiento y la evolución de esta enfermedad. Sin embargo, las limitaciones de este estudio incluyen: la imposibilidad de correlacionar características clínicas con las moleculares debido a la limitación del tamaño muestral; y debido a que los casos no son homogéneos en términos de tiempo de progresión y tratamiento, es decir, cada paciente presentó características clínicas únicas y diferentes estadios de la enfermedad, por lo cual los tratamientos fueron heterogéneos. Otra limitación de este estudio es que el diseño no permite una evaluación del papel de la etnicidad

en la evolución de la enfermedad. Adicionalmente, es posible que el uso de muestras FFPE y una potencial baja celularidad en ganglios linfáticos hayan disminuido la sensibilidad de detección de mutaciones en estos tejidos. La relevancia a nivel clínico de los análisis de exoma completo en pacientes que no tuvieron pCR radica en la capacidad de detección de la prevalencia y composición de las alteraciones en genes conductores que pueden ser usados como blancos potenciales tratables de manera específica. Sin embargo, la ausencia de un grupo de pacientes con pCR limitó la comparación de la evolución de los tumores con aquellos que respondieron al tratamiento.

A pesar de estas limitaciones, nuestros datos soportan hallazgos previos de otros estudios y evidencian una notable heterogeneidad genética intrapaciente, así como la estructura de la arquitectura molecular y la evolución genómica de estas neoplasias agresivas. El análisis longitudinal realizado permitió establecer las relaciones filogenéticas entre las neoplasias de las pacientes y la identificación de una divergencia temprana hacia los ganglios linfáticos. Las firmas mutagénicas S06 y S12 fueron encontradas exclusivamente en ganglios linfáticos y metástasis a piel, lo cual sugiere que las células metastásicas se adaptan a las presiones selectivas de estos nuevos nichos ecológicos mediante la acumulación de nuevas alteraciones producidas por procesos mutagénicos distintos a los presentes en el tumor primario. Finalmente, fueron identificadas mutaciones específicas tratables y que son efectivas en otras neoplasias, lo cual podría servir para la realización de ensayos clínicos en pacientes con CMTN que no responden a la terapia neoadyuvante. De acuerdo con nuestro conocimiento, este es el primer estudio que integra las alteraciones genómicas con procesos evolutivos en una cohorte de CMTN en América Latina, la cual es una región notablemente subrepresentada en los estudios genómicos. Colectivamente, estos hallazgos muestran que existe una amplia heterogeneidad genética intrapaciente, distribuida en todas las lesiones malignas, hecho que podría jugar un papel determinante en la falla terapéutica, evolución adaptativa y el comportamiento clínico agresivo del CMTN.

9. Conclusiones

Este estudio proporciona una visión detallada de la complejidad genómica y heterogeneidad intrapaciente, así como el desarrollo evolutivo del CMTN. Mediante el análisis de bases de datos describimos las características genómicas que han sido asociadas con el tratamiento dirigido en otras neoplasias y que podrían ser evaluadas en el CMTN mediante ensayos clínicos. Integramos de forma longitudinal las alteraciones genéticas con técnicas evolutivas para entender la evolución y la variabilidad genética, así como la descripción de cada estadio de la progresión de la enfermedad en los pacientes con CMTN. Estos hallazgos contribuyen a elucidar el fenómeno global de la evolución del CMTN y proporcionan datos genómicos que podrían dirigir estudios futuros para definir alteraciones específicas tratables para estas pacientes, permitiendo así una terapia más personalizada.

10. Referencias bibliográficas

- Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A.-L., Boyault, S., Burkhardt, B., Butler, A. P., Caldas, C., Davies, H. R., Desmedt, C., Eils, R., Eyfjörd, J. E., Foekens, J. A., ... Stratton, M. R. (2013). Signatures of mutational processes in human cancer. *Nature*, *500*(7463), 415-421. <https://doi.org/10.1038/nature12477>
- Anandakrishnan, R., Varghese, R. T., Kinney, N. A., & Garner, H. R. (2019). Estimating the number of genetic mutations (hits) required for carcinogenesis based on the distribution of somatic mutations. *PLoS Computational Biology*, *15*(3), e1006881. <https://doi.org/10.1371/journal.pcbi.1006881>
- Armitage, P., & Doll, R. (1954). The age distribution of cancer and a multi-stage theory of carcinogenesis. *British Journal of Cancer*, *8*(1), 1-12. <https://doi.org/10.1038/bjc.1954.1>
- Balko, J. M., Giltane, J. M., Wang, K., Schwarz, L. J., Young, C. D., Cook, R. S., Owens, P., Sanders, M. E., Kuba, M. G., Sánchez, V., Kurupi, R., Moore, P. D., Pinto, J. A., Doimi, F. D., Gómez, H., Horiuchi, D., Goga, A., Lehmann, B. D., Bauer, J. A., ... Arteaga, C. L. (2014). Molecular profiling of the residual disease of triple-negative breast cancers after neoadjuvant chemotherapy identifies actionable therapeutic targets. *Cancer Discovery*, *4*(2), 232-245. <https://doi.org/10.1158/2159-8290.CD-13-0286>
- Barry, P., Vatsiou, A., Spiteri, I., Nichol, D., Cresswell, G. D., Acar, A., Trahearn, N., Hrebien, S., Garcia-Murillas, I., Chkhaidze, K., Ermini, L., Huntingford, I. S., Cottom, H., Zabaglo, L., Koelble, K., Khalique, S., Rusby, J. E., Muscara, F., Dowsett, M., ... Sottoriva, A. (2018). The Spatiotemporal Evolution of Lymph Node Spread in Early Breast Cancer. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *24*(19), 4763-4770. <https://doi.org/10.1158/1078-0432.CCR-17-3374>
- Bergin, A. R. T., & Loi, S. (2019). Triple-negative breast cancer: Recent treatment advances. *F1000Research*, *8*, F1000 Faculty Rev-1342. <https://doi.org/10.12688/f1000research.18888.1>
- Bertucci, F., Ng, C. K. Y., Patsouris, A., Droin, N., Piscuoglio, S., Carbuccia, N., Soria, J. C., Dien, A. T., Adnani, Y., Kamal, M., Garnier, S., Meurice, G., Jimenez, M., Dogan, S., Verret, B., Chaffanet, M., Bachelot, T., Campone, M., Lefeuvre, C., ... André, F. (2019). Genomic characterization of metastatic breast cancers. *Nature*, *569*(7757), 560-564. <https://doi.org/10.1038/s41586-019-1056-z>

- Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E., & Gianni, L. (2016). Triple-negative breast cancer: Challenges and opportunities of a heterogeneous disease. *Nature Reviews. Clinical Oncology*, 13(11), 674-690. <https://doi.org/10.1038/nrclinonc.2016.66>
- Breast Cancer Association Consortium, Dorling, L., Carvalho, S., Allen, J., González-Neira, A., Luccarini, C., Wahlström, C., Pooley, K. A., Parsons, M. T., Fortuno, C., Wang, Q., Bolla, M. K., Dennis, J., Keeman, R., Alonso, M. R., Álvarez, N., Herraes, B., Fernandez, V., Núñez-Torres, R., ... Easton, D. F. (2021). Breast Cancer Risk Genes—Association Analysis in More than 113,000 Women. *The New England Journal of Medicine*, 384(5), 428-439. <https://doi.org/10.1056/NEJMoa1913948>
- Brewer, H. R., Jones, M. E., Schoemaker, M. J., Ashworth, A., & Swerdlow, A. J. (2017). Family history and risk of breast cancer: An analysis accounting for family structure. *Breast Cancer Research and Treatment*, 165(1), 193-200. <https://doi.org/10.1007/s10549-017-4325-2>
- Burrell, R. A., & Swanton, C. (2014). Tumour heterogeneity and the evolution of polyclonal drug resistance. *Molecular Oncology*, 8(6), 1095-1111. <https://doi.org/10.1016/j.molonc.2014.06.005>
- Cancer Genome Atlas Network. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*, 490(7418), 61-70. <https://doi.org/10.1038/nature11412>
- Chakravarty, D., Gao, J., Phillips, S. M., Kundra, R., Zhang, H., Wang, J., Rudolph, J. E., Yaeger, R., Soumerai, T., Nissan, M. H., Chang, M. T., Chandarlapaty, S., Traina, T. A., Paik, P. K., Ho, A. L., Hantash, F. M., Grupe, A., Baxi, S. S., Callahan, M. K., ... Schultz, N. (2017). OncoKB: A Precision Oncology Knowledge Base. *JCO Precision Oncology*, 2017. <https://doi.org/10.1200/PO.17.00011>
- Chen, X., Yu, X., Chen, J., Zhang, Z., Tuan, J., Shao, Z., Guo, X., & Feng, Y. (2013). Analysis in early stage triple-negative breast cancer treated with mastectomy without adjuvant radiotherapy: Patterns of failure and prognostic factors. *Cancer*, 119(13), 2366-2374. <https://doi.org/10.1002/cncr.28085>
- Cheng, A. S., Leung, S. C. Y., Gao, D., Burugu, S., Anurag, M., Ellis, M. J., & Nielsen, T. O. (2020). Mismatch repair protein loss in breast cancer: Clinicopathological associations in a large British Columbia cohort. *Breast Cancer Research and Treatment*, 179(1), 3-10. <https://doi.org/10.1007/s10549-019-05438-y>
- Chitty, J. L., Filipe, E. C., Lucas, M. C., Herrmann, D., Cox, T. R., & Timpson, P. (2018). Recent advances in understanding the complexities of metastasis.

- Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E. S., & Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnology*, 31(3), 213-219. <https://doi.org/10.1038/nbt.2514>
- Ciriello, G., Miller, M. L., Aksoy, B. A., Senbabaoglu, Y., Schultz, N., & Sander, C. (2013). Emerging landscape of oncogenic signatures across human cancers. *Nature Genetics*, 45(10), 1127-1133. <https://doi.org/10.1038/ng.2762>
- Cleary, A. S., Leonard, T. L., Gestl, S. A., & Gunther, E. J. (2014). Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature*, 508(7494), 113-117. <https://doi.org/10.1038/nature13187>
- Curtis, C., Shah, S. P., Chin, S.-F., Turashvili, G., Rueda, O. M., Dunning, M. J., Speed, D., Lynch, A. G., Samarajiwa, S., Yuan, Y., Gräf, S., Ha, G., Haffari, G., Bashashati, A., Russell, R., McKinney, S., METABRIC Group, Langerød, A., Green, A., ... Aparicio, S. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, 486(7403), 346-352. <https://doi.org/10.1038/nature10983>
- Dall, G. V., & Britt, K. L. (2017). Estrogen Effects on the Mammary Gland in Early and Late Life and Breast Cancer Risk. *Frontiers in Oncology*, 7, 110. <https://doi.org/10.3389/fonc.2017.00110>
- de Bruin, E. C., McGranahan, N., Mitter, R., Salm, M., Wedge, D. C., Yates, L., Jamal-Hanjani, M., Shafi, S., Murugaesu, N., Rowan, A. J., Grönroos, E., Muhammad, M. A., Horswell, S., Gerlinger, M., Varela, I., Jones, D., Marshall, J., Voet, T., Van Loo, P., ... Swanton, C. (2014). Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science (New York, N.Y.)*, 346(6206), 251-256. <https://doi.org/10.1126/science.1253462>
- Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*, 4(5), P3.
- Ding, L., Ley, T. J., Larson, D. E., Miller, C. A., Koboldt, D. C., Welch, J. S., Ritchey, J. K., Young, M. A., Lamprecht, T., McLellan, M. D., McMichael, J. F., Wallis, J. W., Lu, C., Shen, D., Harris, C. C., Dooling, D. J., Fulton, R. S., Fulton, L. L., Chen, K., ... DiPersio, J. F. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*, 481(7382), 506-510. <https://doi.org/10.1038/nature10738>

- Efron, B., Halloran, E., & Holmes, S. (1996). Bootstrap confidence levels for phylogenetic trees. *Proceedings of the National Academy of Sciences of the United States of America*, 93(14), 7085-7090. <https://doi.org/10.1073/pnas.93.14.7085>
- Engmann, N. J., Golmakani, M. K., Miglioretti, D. L., Sprague, B. L., Kerlikowske, K., & Breast Cancer Surveillance Consortium. (2017). Population-Attributable Risk Proportion of Clinical Risk Factors for Breast Cancer. *JAMA Oncology*, 3(9), 1228-1236. <https://doi.org/10.1001/jamaoncol.2016.6326>
- Fitch, W. M. (1971). Toward Defining the Course of Evolution: Minimum Change for a Specific Tree Topology. *Systematic Biology*, 20(4), 406-416. <https://doi.org/10.1093/sysbio/20.4.406>
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., Cerami, E., Sander, C., & Schultz, N. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science Signaling*, 6(269), p11. <https://doi.org/10.1126/scisignal.2004088>
- Gao, R., Davis, A., McDonald, T. O., Sei, E., Shi, X., Wang, Y., Tsai, P.-C., Casasent, A., Waters, J., Zhang, H., Meric-Bernstam, F., Michor, F., & Navin, N. E. (2016). Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nature Genetics*, 48(10), 1119-1130. <https://doi.org/10.1038/ng.3641>
- Gerstung, M., Jolly, C., Leshchiner, I., D'Ente, S. C., Gonzalez, S., Rosebrock, D., Mitchell, T. J., Rubanova, Y., Anur, P., Yu, K., Tarabichi, M., Deshwar, A., Wintersinger, J., Kleinheinz, K., Vázquez-García, I., Haase, K., Jerman, L., Sengupta, S., Macintyre, G., ... PCAWG Consortium. (2020). The evolutionary history of 2,658 cancers. *Nature*, 578(7793), 122-128. <https://doi.org/10.1038/s41586-019-1907-7>
- Govindan, R., Ding, L., Griffith, M., Subramanian, J., Dees, N. D., Kanchi, K. L., Maher, C. A., Fulton, R., Fulton, L., Wallis, J., Chen, K., Walker, J., McDonald, S., Bose, R., Ornitz, D., Xiong, D., You, M., Dooling, D. J., Watson, M., ... Wilson, R. K. (2012). Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell*, 150(6), 1121-1134. <https://doi.org/10.1016/j.cell.2012.08.024>
- Halsted, W. S. (1894). I. The Results of Operations for the Cure of Cancer of the Breast Performed at the Johns Hopkins Hospital from June, 1889, to January, 1894. *Annals of Surgery*, 20(5), 497-555. <https://doi.org/10.1097/00000658-189407000-00075>

- Hamajima, N., Hirose, K., Tajima, K., Rohan, T., Calle, E. E., Heath, C. W., Coates, R. J., Liff, J. M., Talamini, R., Chantarakul, N., Koetsawang, S., Rachawat, D., Morabia, A., Schuman, L., Stewart, W., Szklo, M., Bain, C., Schofield, F., Siskind, V., ... Collaborative Group on Hormonal Factors in Breast Cancer. (2002). Alcohol, tobacco and breast cancer—Collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *British Journal of Cancer*, 87(11), 1234-1245. <https://doi.org/10.1038/sj.bjc.6600596>
- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, 144(5), 646-674. <https://doi.org/10.1016/j.cell.2011.02.013>
- Harbeck, N., Penault-Llorca, F., Cortes, J., Gnant, M., Houssami, N., Poortmans, P., Ruddy, K., Tsang, J., & Cardoso, F. (2019). Breast cancer. *Nature Reviews. Disease Primers*, 5(1), 66. <https://doi.org/10.1038/s41572-019-0111-2>
- Hu, C., Hart, S. N., Gnanaolivu, R., Huang, H., Lee, K. Y., Na, J., Gao, C., Lilyquist, J., Yadav, S., Boddicker, N. J., Samara, R., Klebba, J., Ambrosone, C. B., Anton-Culver, H., Auer, P., Bandera, E. V., Bernstein, L., Bertrand, K. A., Burnside, E. S., ... Couch, F. J. (2021). A Population-Based Study of Genes Previously Implicated in Breast Cancer. *The New England Journal of Medicine*, 384(5), 440-451. <https://doi.org/10.1056/NEJMoa2005936>
- Hu, Z., Li, Z., Ma, Z., & Curtis, C. (2020). Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nature Genetics*, 52(7), 701-708. <https://doi.org/10.1038/s41588-020-0628-z>
- Hu, Z., Sun, R., & Curtis, C. (2017). A population genetics perspective on the determinants of intra-tumor heterogeneity. *Biochimica Et Biophysica Acta. Reviews on Cancer*, 1867(2), 109-126. <https://doi.org/10.1016/j.bbcan.2017.03.001>
- Hunter, K. W., Amin, R., Deasy, S., Ha, N.-H., & Wakefield, L. (2018). Genetic insights into the morass of metastatic heterogeneity. *Nature Reviews. Cancer*, 18(4), 211-223. <https://doi.org/10.1038/nrc.2017.126>
- Johnson, B. E., Mazar, T., Hong, C., Barnes, M., Aihara, K., McLean, C. Y., Fouse, S. D., Yamamoto, S., Ueda, H., Tatsuno, K., Asthana, S., Jalbert, L. E., Nelson, S. J., Bollen, A. W., Gustafson, W. C., Charron, E., Weiss, W. A., Smirnov, I. V., Song, J. S., ... Costello, J. F. (2014). Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science (New York, N.Y.)*, 343(6167), 189-193. <https://doi.org/10.1126/science.1239947>
- Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J. F., Wyczalkowski, M. A., Leiserson, M. D. M., Miller, C. A.,

- Welch, J. S., Walter, M. J., Wendl, M. C., Ley, T. J., Wilson, R. K., Raphael, B. J., & Ding, L. (2013). Mutational landscape and significance across 12 major cancer types. *Nature*, *502*(7471), 333-339. <https://doi.org/10.1038/nature12634>
- Knight, J. A., Fan, J., Malone, K. E., John, E. M., Lynch, C. F., Langballe, R., Bernstein, L., Shore, R. E., Brooks, J. D., Reiner, A. S., Woods, M., Liang, X., Bernstein, J. L., & WECARE Study Collaborative Group. (2017). Alcohol consumption and cigarette smoking in combination: A predictor of contralateral breast cancer risk in the WECARE study. *International Journal of Cancer*, *141*(5), 916-924. <https://doi.org/10.1002/ijc.30791>
- Lalonde, E., Ishkanian, A. S., Sykes, J., Fraser, M., Ross-Adams, H., Erho, N., Dunning, M. J., Halim, S., Lamb, A. D., Moon, N. C., Zafarana, G., Warren, A. Y., Meng, X., Thoms, J., Grzadkowski, M. R., Berlin, A., Have, C. L., Ramnarine, V. R., Yao, C. Q., ... Bristow, R. G. (2014). Tumour genomic and microenvironmental heterogeneity for integrated prediction of 5-year biochemical recurrence of prostate cancer: A retrospective cohort study. *The Lancet. Oncology*, *15*(13), 1521-1532. [https://doi.org/10.1016/S1470-2045\(14\)71021-6](https://doi.org/10.1016/S1470-2045(14)71021-6)
- Lefebvre, C., Bachelot, T., Filleron, T., Pedrero, M., Campone, M., Soria, J.-C., Massard, C., Lévy, C., Arnedos, M., Lacroix-Triki, M., Garrabey, J., Boursin, Y., Deloger, M., Fu, Y., Commo, F., Scott, V., Lacroix, L., Dieci, M. V., Kamal, M., ... André, F. (2016). Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Medicine*, *13*(12), e1002201. <https://doi.org/10.1371/journal.pmed.1002201>
- Lehmann, B. D., Bauer, J. A., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., & Pietenpol, J. A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *The Journal of Clinical Investigation*, *121*(7), 2750-2767. <https://doi.org/10.1172/JCI45014>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754-1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, X., Yang, J., Peng, L., Sahin, A. A., Huo, L., Ward, K. C., O'Regan, R., Torres, M. A., & Meisel, J. L. (2017). Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. *Breast Cancer Research and Treatment*, *161*(2), 279-287. <https://doi.org/10.1007/s10549-016-4059-6>
- Lipinski, K. A., Barber, L. J., Davies, M. N., Ashenden, M., Sottoriva, A., & Gerlinger, M. (2016). Cancer Evolution and the Limits of Predictability in Precision

- Cancer Medicine. *Trends in Cancer*, 2(1), 49-63.
<https://doi.org/10.1016/j.trecan.2015.11.003>
- Makarem, N., Chandran, U., Bandera, E. V., & Parekh, N. (2013). Dietary fat in breast cancer survival. *Annual Review of Nutrition*, 33, 319-348.
<https://doi.org/10.1146/annurev-nutr-112912-095300>
- Martincorena, I., & Campbell, P. J. (2015). Somatic mutation in cancer and normal cells. *Science (New York, N.Y.)*, 349(6255), 1483-1489.
<https://doi.org/10.1126/science.aab4082>
- Marusyk, A., & Polyak, K. (2010). Tumor heterogeneity: Causes and consequences. *Biochimica Et Biophysica Acta*, 1805(1), 105-117.
<https://doi.org/10.1016/j.bbcan.2009.11.002>
- Marusyk, A., Tabassum, D. P., Altrock, P. M., Almendro, V., Michor, F., & Polyak, K. (2014). Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature*, 514(7520), 54-58.
<https://doi.org/10.1038/nature13556>
- McGranahan, N., & Swanton, C. (2015). Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell*, 27(1), 15-26.
<https://doi.org/10.1016/j.ccell.2014.12.001>
- Meier, B., Cooke, S. L., Weiss, J., Bailly, A. P., Alexandrov, L. B., Marshall, J., Raine, K., Maddison, M., Anderson, E., Stratton, M. R., Gartner, A., & Campbell, P. J. (2014). *C. elegans* whole-genome sequencing reveals mutational signatures related to carcinogens and DNA repair deficiency. *Genome Research*, 24(10), 1624-1636. <https://doi.org/10.1101/gr.175547.114>
- Natrajan, R., Sailem, H., Mardakheh, F. K., Arias Garcia, M., Tape, C. J., Dowsett, M., Bakal, C., & Yuan, Y. (2016). Microenvironmental Heterogeneity Parallels Breast Cancer Progression: A Histology-Genomic Integration Analysis. *PLoS Medicine*, 13(2), e1001961. <https://doi.org/10.1371/journal.pmed.1001961>
- Ng, C. K. Y., Schultheis, A. M., Bidard, F.-C., Weigelt, B., & Reis-Filho, J. S. (2015). Breast cancer genomics from microarrays to massively parallel sequencing: Paradigms and new insights. *Journal of the National Cancer Institute*, 107(5), djv015. <https://doi.org/10.1093/jnci/djv015>
- Nik-Zainal, S., & Morganella, S. (2017). Mutational Signatures in Breast Cancer: The Problem at the DNA Level. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 23(11), 2617-2629.
<https://doi.org/10.1158/1078-0432.CCR-16-2810>

- Nordling, C. O. (1953). A new theory on cancer-inducing mechanism. *British Journal of Cancer*, 7(1), 68-72. <https://doi.org/10.1038/bjc.1953.8>
- Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science (New York, N.Y.)*, 194(4260), 23-28. <https://doi.org/10.1126/science.959840>
- Patey, D. H., & Dyson, W. H. (1948). The prognosis of carcinoma of the breast in relation to the type of operation performed. *British Journal of Cancer*, 2(1), 7-13. <https://doi.org/10.1038/bjc.1948.2>
- Perou, C. M., Sørlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lønning, P. E., Børresen-Dale, A. L., Brown, P. O., & Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature*, 406(6797), 747-752. <https://doi.org/10.1038/35021093>
- Pich, O., Muiños, F., Paul Lolkema, M., Steeghs, N., Gonzalez-Perez, A., & Lopez-Bigas, N. (2019). The mutational footprints of cancer therapies. *Nature genetics*, 51(12), 1732-1740. <https://doi.org/10.1038/s41588-019-0525-5>
- Reynoso-Noverón, N., Villarreal-Garza, C., Soto-Perez-de-Celis, E., Arce-Salinas, C., Matus-Santos, J., Ramírez-Ugalde, M. T., Alvarado-Miranda, A., Cabrera-Galeana, P., Meneses-García, A., Lara-Medina, F., Bargalló-Rocha, E., & Mohar, A. (2017). Clinical and Epidemiological Profile of Breast Cancer in Mexico: Results of the Seguro Popular. *Journal of Global Oncology*, 3(6), 757-764. <https://doi.org/10.1200/JGO.2016.007377>
- Rojas-Jiménez, E., Mejía-Gómez, J. C., Díaz-Velásquez, C., Quezada-Urban, R., Martínez Gregorio, H., Vallejo-Lecuona, F., de la Cruz-Montoya, A., Porras Reyes, F. I., Pérez-Sánchez, V. M., Maldonado-Martínez, H. A., Robles-Estrada, M., Bargalló-Rocha, E., Cabrera-Galeana, P., Ramos-Ramírez, M., Chirino, Y. I., Alonso Herrera, L., Terrazas, L. I., Oliver, J., Frecha, C., ... Vaca-Paniagua, F. (2020). Comprehensive Genomic Profile of Heterogeneous Long Follow-Up Triple-Negative Breast Cancer and Its Clinical Characteristics Shows DNA Repair Deficiency Has Better Prognostic. *Genes*, 11(11), E1367. <https://doi.org/10.3390/genes11111367>
- Rosato, V., Bosetti, C., Negri, E., Talamini, R., Dal Maso, L., Malvezzi, M., Falcini, F., Montella, M., & La Vecchia, C. (2014). Reproductive and hormonal factors, family history, and breast cancer according to the hormonal receptor status. *European Journal of Cancer Prevention: The Official Journal of the European Cancer Prevention Organisation (ECP)*, 23(5), 412-417. <https://doi.org/10.1097/CEJ.0b013e3283639f7a>
- Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S., & Swanton, C. (2016). DeconstructSigs: Delineating mutational processes in single tumors

distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biology*, 17, 31. <https://doi.org/10.1186/s13059-016-0893-4>

Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W. K., Luna, A., La, K. C., Dimitriadou, S., Liu, D. L., Kantheti, H. S., Saghafeinia, S., Chakravarty, D., Daian, F., Gao, Q., Bailey, M. H., Liang, W.-W., Foltz, S. M., Shmulevich, I., Ding, L., Heins, Z., ... Schultz, N. (2018). Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell*, 173(2), 321-337.e10. <https://doi.org/10.1016/j.cell.2018.03.035>

Shah, S. P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., Turashvili, G., Ding, J., Tse, K., Haffari, G., Bashashati, A., Prentice, L. M., Khattra, J., Burleigh, A., Yap, D., Bernard, V., McPherson, A., Shumansky, K., Crisan, A., ... Aparicio, S. (2012). The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*, 486(7403), 395-399. <https://doi.org/10.1038/nature10933>

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 71(3), 209-249. <https://doi.org/10.3322/caac.21660>

Szikriszt, B., Póti, Á., Pipek, O., Krzystanek, M., Kanu, N., Molnár, J., Ribli, D., Szeltner, Z., Tusnády, G. E., Csabai, I., Szallasi, Z., Swanton, C., & Szüts, D. (2016). A comprehensive survey of the mutagenic impact of common cancer cytotoxics. *Genome Biology*, 17, 99. <https://doi.org/10.1186/s13059-016-0963-7>

Talevich, E., Shain, A. H., Botton, T., & Bastian, B. C. (2016). CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Computational Biology*, 12(4), e1004873. <https://doi.org/10.1371/journal.pcbi.1004873>

Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., Boutselakis, H., Cole, C. G., Creatore, C., Dawson, E., Fish, P., Harsha, B., Hathaway, C., Jupe, S. C., Kok, C. Y., Noble, K., Ponting, L., Ramshaw, C. C., Rye, C. E., ... Forbes, S. A. (2019). COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research*, 47(D1), D941-D947. <https://doi.org/10.1093/nar/gky1015>

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14(2), 178-192. <https://doi.org/10.1093/bib/bbs017>

- Tomasetti, C., Marchionni, L., Nowak, M. A., Parmigiani, G., & Vogelstein, B. (2015). Only three driver gene mutations are required for the development of lung and colorectal cancers. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(1), 118-123. <https://doi.org/10.1073/pnas.1421839112>
- Ullah, I., Karthik, G.-M., Alkodsi, A., Kjällquist, U., Stålhammar, G., Lövrot, J., Martinez, N.-F., Lagergren, J., Hautaniemi, S., Hartman, J., & Bergh, J. (2018). Evolutionary history of metastatic breast cancer reveals minimal seeding from axillary lymph nodes. *The Journal of Clinical Investigation*, *128*(4), 1355-1370. <https://doi.org/10.1172/JCI96149>
- Vaca-Paniagua, F., Alvarez-Gomez, R. M., Maldonado-Martínez, H. A., Pérez-Plasencia, C., Fragoso-Ontiveros, V., Lasa-Gonsebatt, F., Herrera, L. A., Cantú, D., Bargallo-Rocha, E., Mohar, A., Durand, G., Forey, N., Voegelé, C., Vallée, M., Le Calvez-Kelm, F., McKay, J., Ardin, M., Villar, S., Zavadil, J., & Olivier, M. (2015). Revealing the Molecular Portrait of Triple Negative Breast Tumors in an Understudied Population through Omics Analysis of Formalin-Fixed and Paraffin-Embedded Tissues. *PloS One*, *10*(5), e0126762. <https://doi.org/10.1371/journal.pone.0126762>
- Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*, 11.10.1-11.10.33. <https://doi.org/10.1002/0471250953.bi1110s43>
- Venet, D., Fimereli, D., Rothé, F., Boeckx, B., Maetens, M., Majjaj, S., Rouas, G., Capra, M., Bonizzi, G., Contaldo, F., Galant, C., Piccart, M., Pruneri, G., Larsimont, D., Lambrechts, D., Desmedt, C., & Sotiriou, C. (2020). Phylogenetic reconstruction of breast cancer reveals two routes of metastatic dissemination associated with distinct clinical outcome. *EBioMedicine*, *56*, 102793. <https://doi.org/10.1016/j.ebiom.2020.102793>
- Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., & Kinzler, K. W. (2013). Cancer genome landscapes. *Science (New York, N.Y.)*, *339*(6127), 1546-1558. <https://doi.org/10.1126/science.1235122>
- Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, *38*(16), e164. <https://doi.org/10.1093/nar/gkq603>
- Wang, X.-X., Jiang, Y.-Z., Li, J.-J., Song, C.-G., & Shao, Z.-M. (2016). Effect of nodal status on clinical outcomes of triple-negative breast cancer: A population-

based study using the SEER 18 database. *Oncotarget*, 7(29), 46636-46645. <https://doi.org/10.18632/oncotarget.9432>

Wang, Y., Waters, J., Leung, M. L., Unruh, A., Roh, W., Shi, X., Chen, K., Scheet, P., Vattathil, S., Liang, H., Multani, A., Zhang, H., Zhao, R., Michor, F., Meric-Bernstam, F., & Navin, N. E. (2014). Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature*, 512(7513), 155-160. <https://doi.org/10.1038/nature13600>

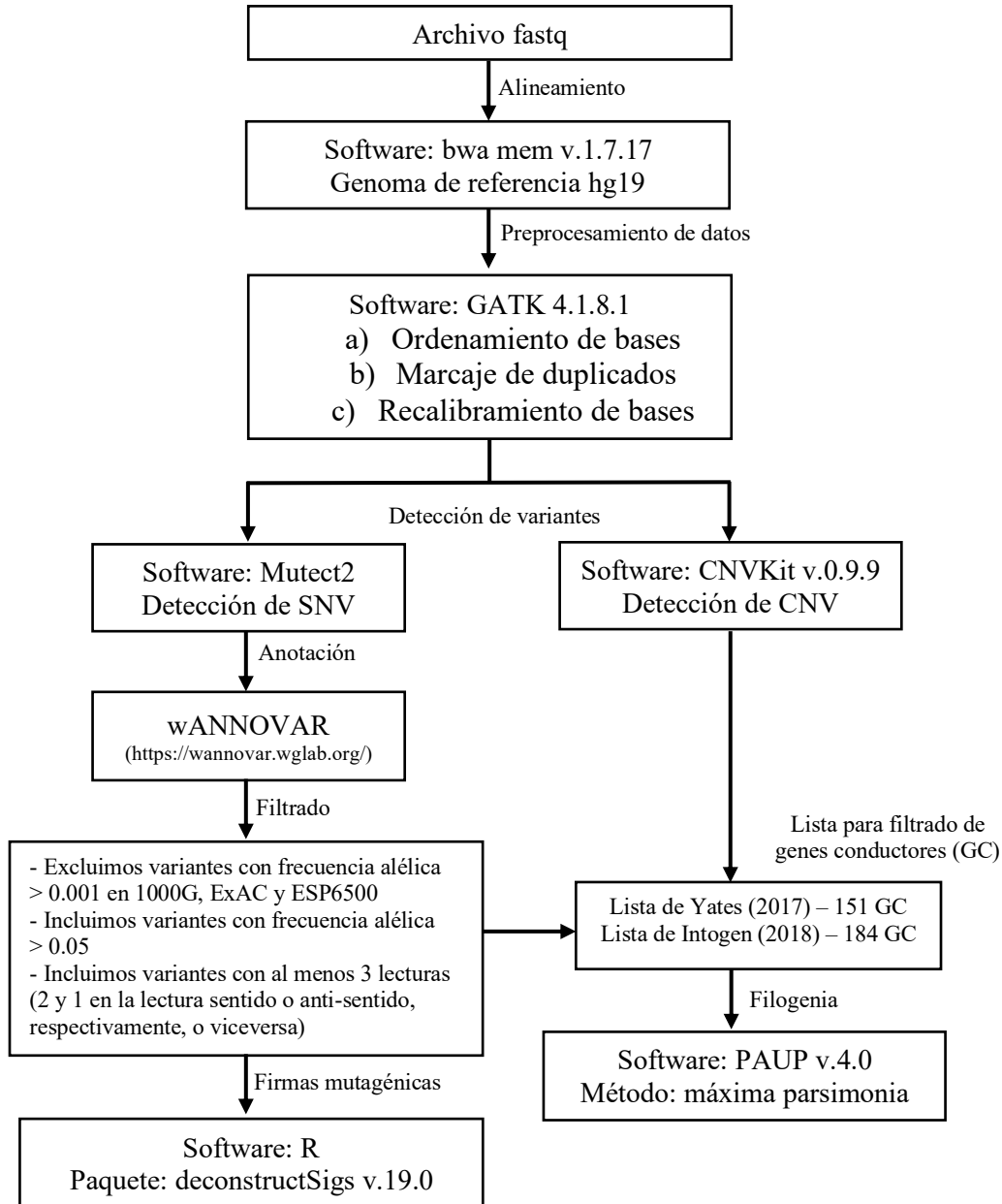
Watkins, T. B. K., Lim, E. L., Petkovic, M., Elizalde, S., Birkbak, N. J., Wilson, G. A., Moore, D. A., Grönroos, E., Rowan, A., Dewhurst, S. M., Demeulemeester, J., Dentre, S. C., Horswell, S., Au, L., Haase, K., Escudero, M., Rosenthal, R., Bakir, M. A., Xu, H., ... Swanton, C. (2020). Pervasive chromosomal instability and karyotype order in tumour evolution. *Nature*, 587(7832), 126-132. <https://doi.org/10.1038/s41586-020-2698-6>

Yates, L. R., Knappskog, S., Wedge, D., Farmery, J. H. R., Gonzalez, S., Martincorena, I., Alexandrov, L. B., Van Loo, P., Haugland, H. K., Lilleng, P. K., Gundem, G., Gerstung, M., Pappaemmanuil, E., Gazinska, P., Bhosle, S. G., Jones, D., Raine, K., Mudie, L., Latimer, C., ... Campbell, P. J. (2017). Genomic Evolution of Breast Cancer Metastasis and Relapse. *Cancer Cell*, 32(2), 169-184.e7. <https://doi.org/10.1016/j.ccell.2017.07.005>

Yuan, Y., Failmezger, H., Rueda, O. M., Ali, H. R., Gräf, S., Chin, S.-F., Schwarz, R. F., Curtis, C., Dunning, M. J., Bardwell, H., Johnson, N., Doyle, S., Turashvili, G., Provenzano, E., Aparicio, S., Caldas, C., & Markowitz, F. (2012). Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Science Translational Medicine*, 4(157), 157ra143. <https://doi.org/10.1126/scitranslmed.3004330>

11. Anexos

Anexo 1. Flujo de trabajo bioinformático



Anexo 2. Mutaciones conductoras de un solo nucleótido

Paciente	Muestra	Mutación	Gen	Proteína	Referencia
CMTN11	Tumor primario 1	chrX:g.39923020C>A	BCOR	p.G1230C	NM_001123385
CMTN11	Tumor primario 1	chr22:g.41573560C>T	EP300	p.Q1949X	NM_001429
CMTN11	Tumor primario 1	chr17:g.56436149G>A	RNF43	p.R330X	NM_017763
CMTN11	Tumor primario 1	chr17:g.7577534C>A	TP53	p.R249S	NM_000546
CMTN11	Tumor primario 2	chr6:g.157469779C>A	ARID1B	p.A845E	NM_017519
CMTN11	Tumor primario 2	chr19:g.11024679G>A	CARM1	p.E266K	NM_199141
CMTN11	Tumor primario 2	chrX:g.70354253C>T	MED12	p.T1555M	NM_005120
CMTN11	Tumor primario 2	chr5:g.14290880C>T	TRIO	p.T199I	NM_007118
CMTN11	Tumor primario 2	chr2:g.43451775C>A	ZFP36L2	p.A390S	NM_006887
CMTN11	Recurrencia	chr22:g.41573560C>T	EP300	p.Q1949X	NM_001429
CMTN11	Recurrencia	chr4:g.153271243G>A	FBXW7	p.R179C	NM_033632
CMTN11	Recurrencia	chrX:g.44929157C>T	KDM6A	p.Q753X	NM_021140
CMTN11	Recurrencia	chr2:g.48030640C>-	MSH6	p.T1085fs	NM_000179
CMTN11	Recurrencia	chr1:g.115256533C>T	NRAS	p.G60R	NM_002524
CMTN11	Recurrencia	chr13:g.48954341G>C	RB1	p.A488P	NM_000321
CMTN11	Recurrencia	chr17:g.7577534C>A	TP53	p.R249S	NM_000546
CMTN18	Tumor primario	chr3:g.142231221G>A	ATR	p.T1578I	NM_001184
CMTN18	Tumor primario	chr7:g.2962867T>C	CARD11	p.T681A	NM_032415
CMTN18	Tumor primario	chr19:g.33793107C>T	CEBPA	p.A72T	NM_004364
CMTN18	Tumor primario	chr16:g.15815380C>T	MYH11	p.E1493K	NM_002474
CMTN18	Tumor primario	chr17:g.74732517C>A	SRSF2	p.R131L	NM_001195427
CMTN18	Tumor primario	chr5:g.14508370C>T	TRIO	p.L3045F	NM_007118
CMTN18	Tumor tratado	chr17:g.15938219C>T	NCOR1	p.S2332N	NM_006311
CMTN18	Tumor tratado	chr17:g.74732517C>A	SRSF2	p.R131L	NM_001195427
CMTN18	Tumor tratado	chr5:g.14508370C>T	TRIO	p.L3045F	NM_007118
CMTN41	Tumor tratado	chr19:g.17952511G>C	JAK3	p.R308G	NM_000215
CMTN41	Tumor tratado	chr3:g.178952085A>G	PIK3CA	p.H1047R	NM_006218
CMTN41	Tumor tratado	chr10:g.43622131C>T	RET	p.R1050X	NM_020975
CMTN41	Ganglio linfático	chr19:g.17952511G>C	JAK3	p.R308G	NM_000215
CMTN41	Ganglio linfático	chr1:g.82409181G>A	LPHN2	p.R309H	NM_001297704
CMTN41	Metástasis a piel	chr19:g.17952511G>C	JAK3	p.R308G	NM_000215
CMTN41	Metástasis a piel	chr3:g.178952085A>G	PIK3CA	p.H1047R	NM_006218
CMTN84	Tumor primario	chr22:g.41572929C>G	EP300	p.C1738W	NM_001429
CMTN84	Ganglio linfático 1	chr1:g.155340754G>A	ASH1L	p.S2118F	NM_018489
CMTN84	Ganglio linfático 1	chr1:g.39788600C>T	MACF1	p.R1391C	NM_012090
CMTN84	Ganglio linfático 1	chr22:g.36714284C>T	MYH9	p.D399N	NM_002473
CMTN84	Ganglio linfático 1	chr16:g.2131748G>A	TSC2	p.V1255M	NM_000548
CMTN84	Tumor tratado	chr16:g.3843560G>-	CREBBP	p.P348fs	NM_004380

CMTN84	Tumor tratado	chr8:g.134251332C>T	NDRG1	p.R325Q	NM_001135242
CMTN84	Tumor tratado	chr19:g.11106911G>A	SMARCA4	p.R539H	NM_001128844
CMTN84	Metástasis a piel	chr16:g.3843560G>-	CREBBP	p.P348fs	NM_004380
CMTN84	Metástasis a piel	chr22:g.41572929C>G	EP300	p.C1738W	NM_001429
CMTN84	Ganglio linfático 1	chr6:g.157528726G>A	ARID1B	p.V2138M	NM_017519

Anexo 3. Mutaciones conductoras en variaciones del número de copias

Paciente	Muestra	Gen	Tipo	CN
CMTN11	Tumor primario 1	<i>AKT3</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>BAP1</i>	Deleción	1
CMTN11	Tumor primario 1	<i>CCND1</i>	Amplificación	9
CMTN11	Tumor primario 1	<i>CDK6</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>IGF1R</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>JAK2</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>KRAS</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>MCL1</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>MDM4</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>MYC</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>NSD1</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>PIK3CA</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>REL</i>	Amplificación	3
CMTN11	Tumor primario 1	<i>STK11</i>	Deleción	1
CMTN11	Tumor primario 1	<i>TERC</i>	Amplificación	3
CMTN11	Tumor primario 2	<i>ARID1A</i>	Deleción	1
CMTN11	Tumor primario 2	<i>ESR1</i>	Amplificación	3
CMTN11	Tumor primario 2	<i>FGFR2</i>	Amplificación	4
CMTN11	Tumor primario 2	<i>IGF1R</i>	Amplificación	3
CMTN11	Tumor primario 2	<i>JUN</i>	Amplificación	3
CMTN11	Tumor primario 2	<i>MDM2</i>	Amplificación	3
CMTN11	Tumor primario 2	<i>PDGFRA</i>	Amplificación	3
CMTN11	Ganglio linfático 1	<i>AKT2</i>	Amplificación	3
CMTN11	Ganglio linfático 1	<i>CCND1</i>	Amplificación	7
CMTN11	Ganglio linfático 1	<i>CCNE1</i>	Amplificación	3
CMTN11	Ganglio linfático 1	<i>NOTCH3</i>	Amplificación	3
CMTN11	Ganglio linfático 1	<i>RB1</i>	Deleción	1
CMTN11	Ganglio linfático 2	<i>CDKN1B</i>	Deleción	1
CMTN11	Ganglio linfático 2	<i>RB1</i>	Deleción	1
CMTN11	Recurrencia	<i>AKT3</i>	Amplificación	3
CMTN11	Recurrencia	<i>BAP1</i>	Deleción	1
CMTN11	Recurrencia	<i>CCND1</i>	Amplificación	9
CMTN11	Recurrencia	<i>CDK6</i>	Amplificación	3
CMTN11	Recurrencia	<i>FGFR2</i>	Amplificación	3
CMTN11	Recurrencia	<i>IGF1R</i>	Amplificación	3
CMTN11	Recurrencia	<i>JAK2</i>	Amplificación	3
CMTN11	Recurrencia	<i>JUN</i>	Amplificación	3

CMTN11	Recurrencia	<i>KRAS</i>	Amplificación	3
CMTN11	Recurrencia	<i>MCL1</i>	Amplificación	3
CMTN11	Recurrencia	<i>MDM4</i>	Amplificación	3
CMTN11	Recurrencia	<i>MYC</i>	Amplificación	3
CMTN18	Tumor primario	<i>ARID1A</i>	Delección	0
CMTN18	Tumor primario	<i>BAP1</i>	Delección	1
CMTN18	Tumor primario	<i>BRCA1</i>	Delección	1
CMTN18	Tumor primario	<i>CDKN1B</i>	Delección	0
CMTN18	Tumor primario	<i>CDKN2A</i>	Delección	1
CMTN18	Tumor primario	<i>CDKN2B</i>	Delección	1
CMTN18	Tumor primario	<i>PBRM1</i>	Delección	1
CMTN18	Tumor primario	<i>PTEN</i>	Delección	0
CMTN18	Tumor primario	<i>RB1</i>	Delección	0
CMTN18	Tumor tratado	<i>ARID1A</i>	Delección	0
CMTN18	Tumor tratado	<i>BAP1</i>	Delección	1
CMTN18	Tumor tratado	<i>BRCA1</i>	Delección	1
CMTN18	Tumor tratado	<i>CDKN1B</i>	Delección	0
CMTN18	Tumor tratado	<i>CDKN2A</i>	Delección	1
CMTN18	Tumor tratado	<i>CDKN2B</i>	Delección	1
CMTN18	Tumor tratado	<i>PBRM1</i>	Delección	1
CMTN18	Tumor tratado	<i>PTEN</i>	Delección	0
CMTN18	Tumor tratado	<i>RB1</i>	Delección	1
CMTN18	Tumor tratado	<i>STK11</i>	Delección	1
CMTN41	Tumor tratado	<i>AKT1</i>	Amplificación	3
CMTN41	Tumor tratado	<i>ARID1A</i>	Delección	1
CMTN41	Tumor tratado	<i>CCND1</i>	Amplificación	3
CMTN41	Tumor tratado	<i>CDKN1B</i>	Delección	1
CMTN41	Tumor tratado	<i>CDKN2A</i>	Delección	0
CMTN41	Tumor tratado	<i>CDKN2B</i>	Delección	0
CMTN41	Tumor tratado	<i>EGFR</i>	Amplificación	3
CMTN41	Tumor tratado	<i>FGFR3</i>	Amplificación	3
CMTN41	Tumor tratado	<i>MCL1</i>	Amplificación	3
CMTN41	Tumor tratado	<i>MYC</i>	Amplificación	3
CMTN41	Tumor tratado	<i>NFE2L2</i>	Amplificación	3
CMTN41	Tumor tratado	<i>NSD1</i>	Amplificación	3
CMTN41	Tumor tratado	<i>TERC</i>	Amplificación	3
CMTN41	Tumor tratado	<i>TERT</i>	Amplificación	5
CMTN41	Ganglio linfático	<i>AKT1</i>	Amplificación	3
CMTN41	Ganglio linfático	<i>TERT</i>	Amplificación	3
CMTN41	Metástasis a piel	<i>AKT1</i>	Amplificación	3

CMTN41	Metástasis a piel	<i>CCND1</i>	Amplificación	3
CMTN41	Metástasis a piel	<i>CDKN1B</i>	Delección	1
CMTN41	Metástasis a piel	<i>CDKN2A</i>	Delección	1
CMTN41	Metástasis a piel	<i>CDKN2B</i>	Delección	1
CMTN41	Metástasis a piel	<i>EGFR</i>	Amplificación	3
CMTN41	Metástasis a piel	<i>FGFR3</i>	Amplificación	3
CMTN41	Metástasis a piel	<i>NSD1</i>	Amplificación	3
CMTN41	Metástasis a piel	<i>PTEN</i>	Delección	1
CMTN41	Metástasis a piel	<i>TERT</i>	Amplificación	5
CMTN84	Tumor primario	<i>BRCA1</i>	Delección	1
CMTN84	Tumor primario	<i>MCL1</i>	Amplificación	3
CMTN84	Tumor primario	<i>NOTCH3</i>	Amplificación	3
CMTN84	Ganglio linfático 1	<i>BRCA1</i>	Delección	1
CMTN84	Ganglio linfático 1	<i>TERT</i>	Amplificación	3
CMTN84	Tumor tratado	<i>BAP1</i>	Delección	1
CMTN84	Tumor tratado	<i>BRCA1</i>	Delección	1
CMTN84	Tumor tratado	<i>MCL1</i>	Amplificación	3
CMTN84	Tumor tratado	<i>MDM4</i>	Amplificación	3
CMTN84	Tumor tratado	<i>PBRM1</i>	Delección	1
CMTN84	Tumor tratado	<i>RB1</i>	Delección	1
CMTN84	Tumor tratado	<i>TERT</i>	Amplificación	3
CMTN84	Metástasis a piel	<i>ARID1A</i>	Delección	1
CMTN84	Metástasis a piel	<i>MCL1</i>	Amplificación	3
CMTN84	Metástasis a piel	<i>MDM4</i>	Amplificación	3
CMTN84	Metástasis a piel	<i>PTEN</i>	Delección	1
CMTN84	Metástasis a piel	<i>TERT</i>	Amplificación	3
CMTN84	Ganglio linfático 2	<i>AKT2</i>	Amplificación	3
CMTN84	Ganglio linfático 2	<i>CCNE1</i>	Amplificación	3
CMTN84	Ganglio linfático 2	<i>MCL1</i>	Amplificación	3
CMTN84	Ganglio linfático 2	<i>NOTCH3</i>	Amplificación	3
CMTN84	Ganglio linfático 2	<i>RB1</i>	Delección	1
CMTN84	Ganglio linfático 3	<i>AKT2</i>	Amplificación	3
CMTN84	Ganglio linfático 3	<i>CCNE1</i>	Amplificación	3
CMTN84	Ganglio linfático 3	<i>MCL1</i>	Amplificación	4
CMTN84	Ganglio linfático 3	<i>NOTCH3</i>	Amplificación	3
CMTN84	Ganglio linfático 3	<i>PTEN</i>	Delección	1
CMTN84	Ganglio linfático 3	<i>RB1</i>	Delección	1

Anexo 4. Artículo publicado en el doctorado







Martínez-Gregorio H., Rojas-Jiménez E., Mejía-Gómez JC., Díaz-Velásquez C., Quezada-Urban R., Vallejo-Lecuona F., De la Cruz-Montoya A., Porras-Reyes FI., Pérez-Sánchez VM., Maldonado-Martínez HA., Robles-Estrada M., Bargalló-Rocha E., Cabrera-Galeana P., Ramos- Ramírez M., Chirino YI., Herrera LA., Terrazas LI., Frecha C., Oliver J., Perdomo S. and Vaca- Paniagua F. (2021). **“The Evolution of Clinically Aggressive Triple-Negative Breast Cancer shows a Large Mutational Diversity and Early Metastasis to Lymph Nodes”**. *Cancer*, 13, 5091.

doi: <https://doi.org/10.3390/cancers13205091>

Dirrección electrónica: <https://www.mdpi.com/2072-6694/13/20/5091>

Article

The Evolution of Clinically Aggressive Triple-Negative Breast Cancer Shows a Large Mutational Diversity and Early Metastasis to Lymph Nodes

Héctor Martínez-Gregorio ^{1,2,3}, Ernesto Rojas-Jiménez ^{2,3} , Javier César Mejía-Gómez ⁴, Clara Díaz-Velásquez ², Rosalía Quezada-Urban ^{2,3,5,6}, Fernando Vallejo-Lecuona ^{2,3} , Aldo de la Cruz-Montoya ³ , Fany Iris Porras-Reyes ⁷, Víctor Manuel Pérez-Sánchez ⁷, Héctor Aquiles Maldonado-Martínez ⁷, Maybelline Robles-Estrada ⁸, Enrique Bargalló-Rocha ⁷, Paula Cabrera-Galeana ⁷, Maritza Ramos-Ramírez ⁷, Yolanda Irasema Chirino ³ , Luis Alonso Herrera ^{7,9,10}, Luis Ignacio Terrazas ^{2,3}, Cecilia Frecha ¹¹, Javier Oliver ¹² , Sandra Perdomo ¹³  and Felipe Vaca-Paniagua ^{2,3,7,*}

- ¹ Posgrado en Ciencias Biológicas de la Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Iztacala, UNAM, Mexico City 54090, Mexico; mag_hector@hotmail.com
- ² Laboratorio Nacional en Salud, Diagnóstico Molecular y Efecto Ambiental en Enfermedades Crónicas-Degenerativas, Facultad de Estudios Superiores Iztacala, Tlalnepantla 54090, Mexico; erarroji@gmail.com (E.R.-J.); cdiazvelasquez@aol.com (C.D.-V.); rosalia.quezadaurban@petermac.org (R.Q.-U.); fernando.vallejo.lecuon@gmail.com (F.V.-L.); literrazas@unam.mx (L.I.T.)
- ³ Unidad de Biomedicina, Facultad de Estudios Superiores Iztacala, UNAM, Tlalnepantla 54090, Mexico; audelacm@gmail.com (A.d.l.C.-M.); irasemachirino@gmail.com (Y.I.C.)
- ⁴ Division of Breast Cancer, Department of Medical Oncology, Mt. Sinai Hospital, University of Toronto, Toronto, ON M5G 1X5, Canada; javier.mejiagomez@mail.utoronto.ca
- ⁵ Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, VIC 3000, Australia
- ⁶ Cancer Research Division, Peter MacCallum Cancer Centre, Melbourne, VIC 3000, Australia
- ⁷ Instituto Nacional de Cancerología, Mexico City 14080, Mexico; fany.porras@gmail.com (F.I.P.-R.); manuelps98@gmail.com (V.M.P.-S.); arzhaus@yahoo.com (H.A.M.-M.); ebargallo@yahoo.com (E.B.-R.); drapaulacabrera@gmail.com (P.C.-G.); maritzaramos1304@gmail.com (M.R.-R.); metil@hotmail.com (L.A.H.)
- ⁸ Hospital General de Pachuca SSA, Pachuca 42070, Mexico; roem1825@hotmail.com
- ⁹ Instituto Nacional de Medicina Genómica, Mexico City 14610, Mexico
- ¹⁰ Unidad de Investigación Biomédica en Cáncer, Instituto de Investigaciones Biomédicas—Instituto Nacional de Cancerología, Mexico City 14080, Mexico
- ¹¹ Unidad de Producción Celular del Hospital Regional Universitario de Málaga—IBIMA—Málaga, 29010 Málaga, Spain; frechacecilia@gmail.com
- ¹² Medical Oncology Service, Hospitales Universitarios Regional y Virgen de la Victoria, Institute of Biomedical Research in Málaga, CIMES, University of Málaga, 29010 Málaga, Spain; javiom@gmail.com
- ¹³ Genomic Epidemiology Branch, International Agency for Research on Cancer (IARC/WHO), 150 Cours Albert Thomas, Lyon 69372, France; perdomos@iarc.fr
- * Correspondence: felipe.vaca@iztacala.unam.mx; Tel.: +52-55-5623-1333 (ext. 39788)



Citation: Martínez-Gregorio, H.; Rojas-Jiménez, E.; Mejía-Gómez, J.C.; Díaz-Velásquez, C.; Quezada-Urban, R.; Vallejo-Lecuona, F.; de la Cruz-Montoya, A.; Porras-Reyes, F.I.; Pérez-Sánchez, V.M.; Maldonado-Martínez, H.A.; et al. The Evolution of Clinically Aggressive Triple-Negative Breast Cancer Shows a Large Mutational Diversity and Early Metastasis to Lymph Nodes. *Cancers* **2021**, *13*, 5091. <https://doi.org/10.3390/cancers13205091>

Academic Editor: Samuel C. Mok

Received: 11 August 2021

Accepted: 15 September 2021

Published: 12 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: Triple-negative breast cancer (TNBC) is a clinically, phenotypically, and molecularly heterogeneous disease. This heterogeneity is a factor that negatively impacts therapy response. To analyze evolutionary patterns and the genomic alterations in patients with clinically aggressive disease who did not respond to treatment, we performed whole-exome sequencing in multiple longitudinal samples from diagnosis to distant metastasis. We found an extensive inpatient and interpatient genetic heterogeneity, mutational signature composition at different stages, and, interestingly, an early lymph node metastasis formation during the evolution of aggressive TNBC. This study provides detailed insights into the genomic complexity, and the phylogenetic and evolutionary development of TNBC, as well as identifying specific mutations associated with targeted treatments in TNBC.

Abstract: In triple-negative breast cancer (TNBC), only 30% of patients treated with neoadjuvant chemotherapy achieve a pathological complete response after treatment and more than 90% die due to metastasis formation. The diverse clinical responses and metastatic developments are attributed to extensive inpatient genetic heterogeneity and tumor evolution acting on this neoplasm. In

this work, we aimed to evaluate genomic alterations and tumor evolution in TNBC patients with aggressive disease. We sequenced the whole exome of 16 lesions from four patients who did not respond to therapy, and took several follow-up samples, including samples from tumors before and after treatment, as well as from the lymph nodes and skin metastases. We found substantial inpatient genetic heterogeneity, with a variable tumor mutational composition. Early truncal events were *MCL1* amplifications. Metastatic lesions had deletions in *RB1* and *PTEN*, along with *TERT*, *AKT2*, and *CCNE1* amplifications. Mutational signatures 06 and 12 were mainly detected in skin metastases and lymph nodes. According to phylogenetic analysis, the lymph node metastases occurred at an early stage of TNBC development. Finally, each patient had three to eight candidate driving mutations for targeted treatments. This study delves into the genomic complexity and the phylogenetic and evolutionary development of aggressive TNBC, supporting early metastatic development, and identifies specific genetic alterations associated with a response to targeted therapies.

Keywords: triple-negative breast cancer; tumor evolution; metastasis; lymph nodes; early divergence; targeted therapies

1. Introduction

Triple-negative breast cancer (TNBC) is defined by a lack of expression of estrogen receptors (ER), progesterone receptors (PR), and an amplification or overexpression of human epidermal growth factor receptor type 2 (HER2), and accounts for 10–15% of breast cancer cases [1,2]. TNBC is treated with chemotherapy based on anthracyclines, taxanes, and platinum salts [3]. Despite optimal management of TNBC treatment, only 30% of patients achieve a pathological complete response (pCR) with neoadjuvant chemotherapy (NAC) [3]. This differential response with NAC has been attributed, at least in part, to intratumoral heterogeneity and continuous tumor evolution, which facilitate the growth of treatment-resistant subclones and their subsequent spread through metastatic events [4,5].

Breast cancer (BC) associated mortality is a consequence of the metastatic spread of a primary tumor to different organs and accounts for 90% of all deaths from breast cancer [6]. BC has different clinical scenarios and metastatic spread patterns. Lymph nodes are the most common sites of metastases in BC, and they are associated with a bad prognosis and recurrence [7,8]. Skin metastases, on the other hand, occur in up to 20% of all cases and these are also associated with a poor prognosis [9,10]. Several studies have demonstrated that primary tumors share many mutations with their metastases, but at the same time each metastatic lesion has clear differences in the number and type of affected genes as a result of tumor evolution and selective treatment pressures [5,11–13].

Usually, primary tumors and metastases are the most available samples [5,11,12,14]. However, studying evolution in TNBC has been challenging due to the difficulty of obtaining samples at all stages of the disease. In this context, we performed whole-exome sequencing (WES) on 16 samples from four patients with clinically aggressive disease who did not respond to several lines of treatment in order to analyze the evolutionary patterns of the cancerous cells, and to describe their genomic allelic composition and mutational patterns. Our analysis included a set of follow-up samples, such as tumor samples (before and after treatment), as well as samples from lymph nodes, skin metastases, and normal tissue. Our data show an extensive intra- and interpatient genetic heterogeneity. Additionally, these data indicate an early divergence to lymph nodes during TNBC carcinogenesis. Finally, we identified potential candidate genes that could be suitable for targeting with therapeutic precision agents which may contribute to preventing treatment failure in the future.

2. Materials and Methods

2.1. Sample Information

A total of 97 patients were diagnosed with TNBC and treated at the National Cancer Institute (INCAN) of Mexico between April 2007 and April 2010. TNBCs were diagnosed by immunohistochemistry (IHC) based on the criteria of less than one percent expression of the staining score for the ER, PR, and HER2 receptors. We selected cases with clinically aggressive disease, defined as the absence of pCR to NAC, a lack of response to any treatment, death within three years of diagnosis, and with ductal carcinoma (Table S1). After DNA quality control, only 16 samples from four patients were amplifiable and suitable for WES analysis, including primary treatment-naïve tumors, treated tumors, lymph nodes, and skin metastases, all of which were paired with their respective normal samples. For this study, the protocol was approved by the Research and Ethics Committees (protocol #016/013/IBI CEI/1021/16) and conducted by the Declaration of Helsinki.

2.2. Sample Preparation and DNA Extraction

Samples were obtained from formalin-fixed paraffin embedded (FFPE) tissues and extracted from three 50 μm slices corresponding to about 50–200 mg of tissue. Samples were dewaxed with xylol and ethanol baths and digested overnight with proteinase K. Crosslinking formation by formaldehyde was reversed through a 90 °C incubation for 1 h. DNA extraction was done with the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. DNA concentration was quantified by Qubit (Invitrogen, Carlsbad, USA). DNA integrity and purity were verified by agarose gel electrophoresis and spectrophotometry, respectively, to evaluate DNA quality.

2.3. Library Preparation

The library preparation was done on the Illumina Nextera Rapid Capture Exome Kit v.1.2 (Illumina, San Diego, CA, USA), which covers 37 Mb of protein-coding bases, following the manufacturer's instructions. Pair-end sequencing was performed on an Illumina HiSeq 2500 for 2×150 cycles, at the *Red de Apoyo a la Investigación* of the National Institute of Nutrition Salvador Zubirán, Mexico City, Mexico.

2.4. SNV e Indels Identification

Sequencing reads were aligned to the human genome reference hg19 with BWA-MEM [15]. GATK tools were used to sort reads, mark duplicates, and for base recalibration [16]. Single nucleotide variations (SNV) were called with Mutect2 [17] and annotated with ANNOVAR [18]. Variants were filtered as follows: (i) variants with an allelic frequency of less than 0.001 in 1000 genomes, ExAC and ESP6500 databases, were included; (ii) common SNP were excluded, unless they were reported as pathogenic in ClinVar; (iii) variants with at least two reads in both DNA chains in the tumor samples were kept; and (iv) variants with an allelic fraction of ≥ 0.05 were also kept. We used ExAC to exclude non-pathogenic natural variation, considering that this repository includes the sequencing data of 5789 exomes of Latin and admixed Americans.

Driver genes were identified using two well-curated gene lists: Intogen v.2019 (<https://www.intogen.org/>) and the list from Yates [5]. Then, driver mutations were defined as pathogenic if they were previously described in ClinVar as pathogenic, if at least two out of three algorithms—SIFT, PolyPhen2 or MutationTaster—predicted them as deleterious, and if they were known hotspot mutations in COSMIC [19] and cBioportal [20] (Table S2). All driver variants were manually curated by inspection of the BAM files using the IGV software [21].

2.5. Tumor Mutational Burden Analysis

Tumor mutational burden (TMB) was defined as the number of somatic, coding, and indel mutations, and base substitutions per megabase of the examined genome. To calculate the TMB per megabase, the total number of counted mutations was divided by the size

of the coding region of the targeted territory. An estimated exome size of 37 Mb was considered for the analyses.

2.6. CNV Identification

Copy number variation (CNV) was identified with CNVkit [22], using a panel of normal samples with default parameters. We defined the driver genes according to the Intogen and Yates lists previously described. The found variants were filtered by directionality, considering the biology of the affected gene as follows: (i) CNV with CN = 2 were excluded; (ii) oncogenes with copy number (CN) < 2 were excluded and CN > 2 were kept; (iii) tumor suppressor genes with CN > 2 were excluded and CN < 2 were kept.

2.7. Mutational Signature Analysis

Mutation patterns were evaluated in the context of 96 trinucleotides using the R package deconstructSigs [23]. To evaluate the distribution of the samples with mutational signatures, an unsupervised hierarchical clustering analysis was performed by calculating Euclidean distances.

2.8. Pathway Enrichment Analysis

Pathway and network enrichment analysis was defined with the David Functional Annotation Tool 6.8 [24], using driver genes in the KEGG pathway database. Uninformative pathways were eliminated, such as general cancer and non-cancer diseases.

2.9. Global Actionable Alterations

Treatments approved by the FDA for actionable genes in all types of cancer were selected using the OncoKB database [25].

2.10. Evolutionary Tree Construction

Phylogenetic trees were constructed with SNV, indels, and CNV, using maximum parsimony principles [26]. Variants were coded with 0 and 1 to indicate the absence and presence of mutations, respectively, and processed in mesquite (<http://www.mesquiteproject.org/>) to generate nexus files. The phylogenetic trees were analyzed in PAUP (<http://paup.phylosolutions.com/>) using a heuristic method and rooted with normal samples. The reliability of trees was evaluated with the Bootstrap Method [27]. The trees generated were saved in Newick format and visualized with the FigTree software (<http://tree.bio.ed.ac.uk/software/figtree/>).

3. Results

3.1. Patients with TNBC Exhibit a Variety of Clinical Scenarios

For this study, we reviewed the database of patients diagnosed with TNBC and treated at the INCAN during April 2007 and April 2010. We chose 16 FFPE samples from four patients with aggressive disease and high-quality DNA to perform WES. Four clinical settings were represented by the samples: (i) primary treatment-naïve tumor, (ii) treated tumor, (iii) lymph nodes, and (iv) skin metastasis. All patients were diagnosed at advanced stage and received NAC treatment. None of them achieved a pCR and all died within three years of their diagnosis (Figure 1a,b; Table S1).

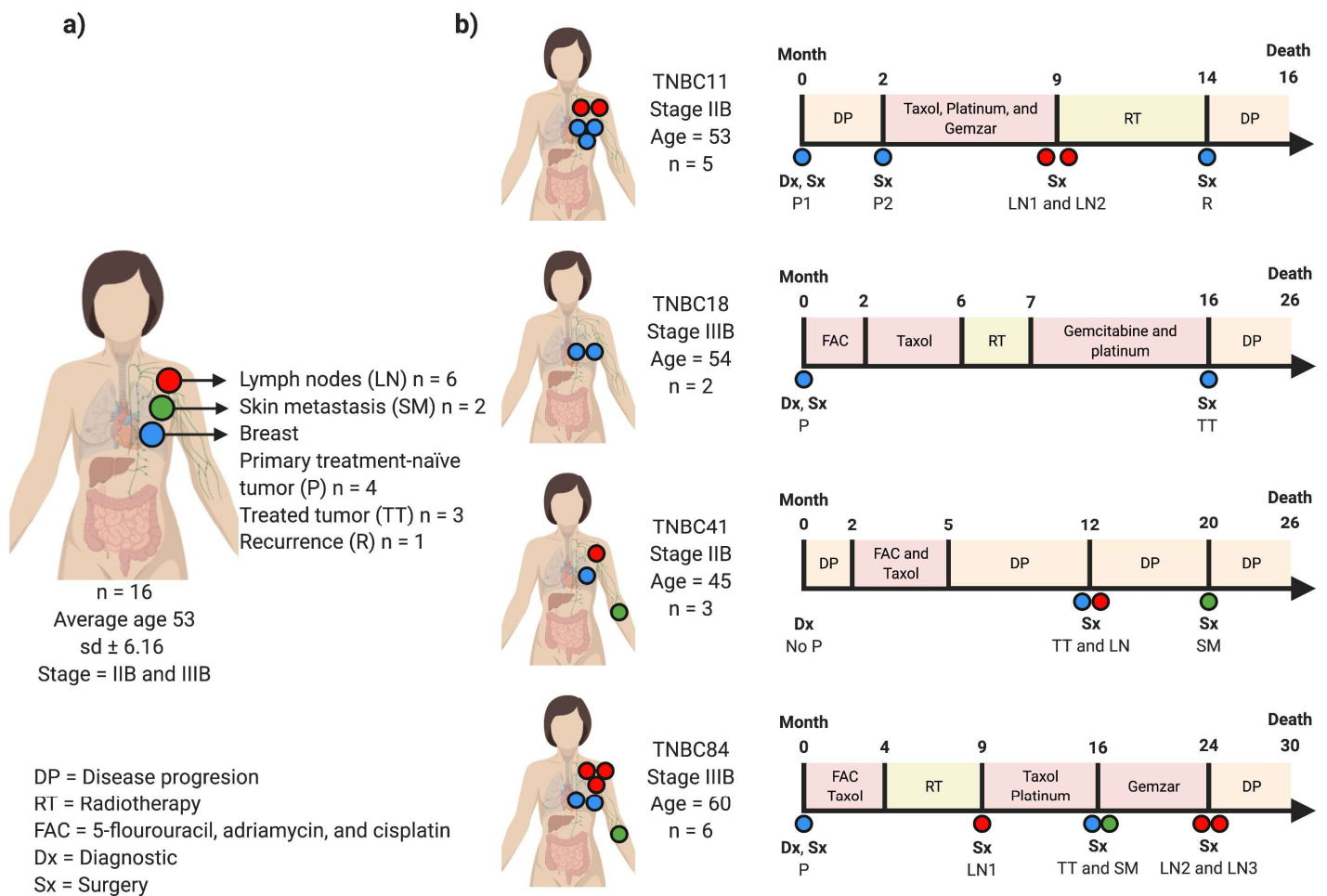


Figure 1. Clinical history of TNBC patients. (a) The cohort of TNBC with 16 samples from four patients includes tumors obtained before and after treatments, lymph node- and skin metastasis. (b) The timeline depicts the clinical characteristics of all patients analyzed and samples taken during the progression of the disease. Colored boxes show the different treatments received by patients (blue: breast; red: lymph node metastases; green: skin metastases).

Patient TNBC11 was a woman diagnosed at age 53 with a locally advanced TNBC. A first biopsy of the primary tumor was obtained and had pathogenic variants in *TP53* and *EP300*, as well as amplifications in *MCL1* and *IGF1R* (Figures 1 and 2). Then, a second biopsy was taken from the same primary tumor at month two, and the patient started chemotherapy treatment. This second sample had different driving alterations and only shared the *IGF1R* amplification with the previous sample. At the ninth month, a mastectomy was performed, and two lymph node biopsies were obtained. Both lymph nodes showed *RB1* deletions and low TMB. Lymph node 1 showed *NOTCH3*, *AKT2*, and *CCNE1* amplifications, and lymph node 2 had a CNV loss in *CDKN1B*. At month 14, after five months of radiotherapy, a last mastectomy was performed and a recurrence sample was obtained, which showed a composition of the altered driver genes very similar to those detected in the primary tumour 1, including the *EP300* and *TP53* mutations, and amplifications in *MCL1*, *CCND1*, *MDM4*, *IGF1R*, *MYC*, *AKT3*, *CDK6*, *JAK2*, *KRAS*. The recurrence also shared a *BAP1* deletion with the primary tumor 1. Globally, all the neoplastic lesions had 57 driver alterations. Patient TNBC11 died after the 16th month of diagnosis.

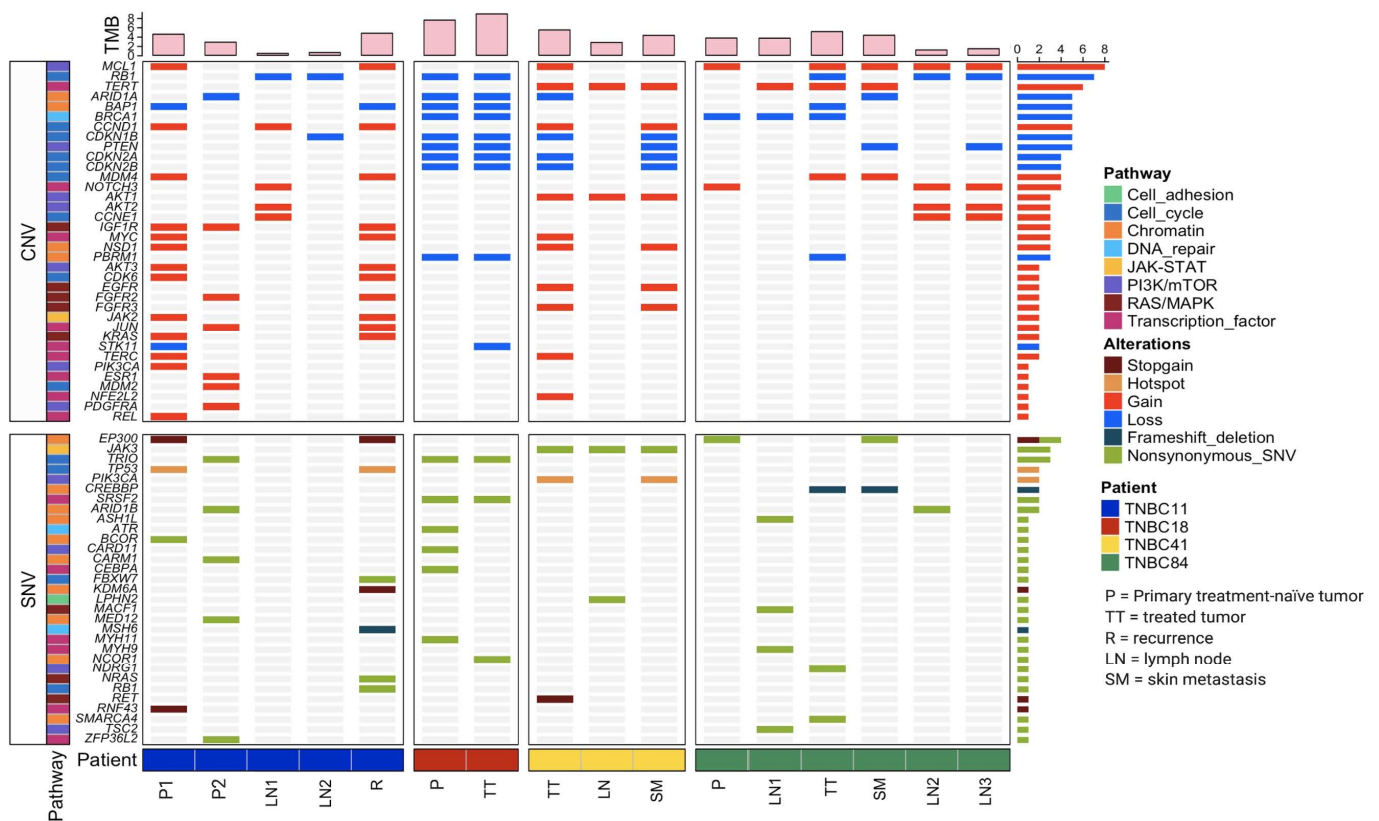


Figure 2. Allelic composition and CNV alterations in driver genes. The samples were segmented and color-coded by patient and sorted chronologically. TMB is indicated on the top panel with pink bars. Driver genes are classified into two categories: SNV (lower panel) and CNV (upper panel). Mutation signaling pathways in which driver genes are involved, as well as mutation types, are denoted by distinct colors.

Patient TNBC18 was a woman diagnosed at age 54 with two available samples, a primary treatment-naïve tumor and a treated tumor. The primary treatment-naïve tumor was obtained in the first month of diagnosis. After 16 months of treatment with chemo- and radiotherapy, surgery was performed to obtain the treated tumor. Both samples showed a very similar pattern of alterations and a high TMB. Nine deletions were shared, including *RB1*, *ARID1A*, *BAP1*, *BRCA1*, *CDKN1B*, *PTEN*, *CDKN2A*, *CDKN2B*, and *PBRM1*. Patient TNBC18 had a total load of 28 driver alterations and died in the 26th month of diagnosis.

Patient TNBC41 was diagnosed at age 45 and had three samples taken from a treated tumor, a lymph node and a skin metastasis. The primary treatment-naïve tumor was not available. After 12 months of treatment with chemotherapy, a mastectomy was performed, and two samples were obtained: a treated tumor and a lymph node metastasis. The treated tumor showed several CNVs in driver genes, including amplifications in *MCL1*, *TERT*, *CCND1*, *AKT1*, *MYC*, *EGFR*, and *FGFR3*; deletions in *CDK1B*, *CDKN2A* and *CDKN2B*; a *PIK3CA* hotspot mutation; and a high TMB. In contrast, the lymph node metastasis had only four pathogenic alterations and a lower TMB. In the 20th month, further surgery was carried out, and a skin metastasis biopsy was obtained. The skin metastasis sample showed a very similar pattern of alterations in driver genes as compared with the treated tumor, including the *PIK3CA* hotspot mutation, the amplifications in *CCND1*, *EGFR*, and *FGFR3*, and the deletions in *CDK1B*, *CDKN2A*, and *CDKN2B*. This skin metastasis also had a private *PTEN* deletion. The three samples shared SNV in *JAK3* as well as *TERT*, and *AKT1* amplifications. Patient TNBC41 had a total number of 33 driver genes affected and died in the 26th month after diagnosis.

Patient TNBC84 was diagnosed at age 60 and had six samples taken. The primary treatment-naïve tumor was extracted in the diagnosis and had *MCL1* and *NOTCH3* amplifications, and a *BRCA1* deletion. After nine months of chemo- and radiotherapy, surgery was performed, and a lymph node biopsy was obtained. This had a *TERT* amplification, a *BRCA1* CNV loss, and a *TSC2* pathogenic mutation. A treated tumor and a skin metastasis sample were obtained in the 16th month, and both showed *MCL1*, *TERT*, and *MDM4* amplifications. Finally, additional surgery was performed in the 24th month, and samples were obtained from two lymph nodes: lymph nodes 2 and 3. Both lymph nodes showed *RB1* deletions, as well as *MCL1*, *NOTCH3*, *AKT2*, and *CCNE1* amplifications. Interestingly, all the lesions excluding lymph node 1 had *MCL1* amplifications. Patient TNBC84 harbored 39 driving alterations and died in the 30th month after diagnosis.

3.2. Patients with TNBC Display an Extensive Genetic Heterogeneity

Our sequencing analysis identified an extensive genetic interpatient heterogeneity. We found a total of 2264 SNVs in the 16 samples, including 611 synonymous, 1539 non-synonymous mutations, and 114 indels. TMB was broad, ranging from 0.4–8.9 mutations per megabase, with an overall mean of 3.8 (Figure 2). Each patient's sample showed a unique TMB. For instance, for patient TNBC11 the recurrence sample had the highest TMB, while for patients TNBC18, TNBC41, and TNBC84 the highest TMB was found in samples coming from a treated tumor. On the other hand, lymph nodes showed the lowest TMB in three out of four patients analyzed (TNBC11, TNBC41, and TNBC84).

We identified 157 mutations in 67 driver genes, from which 114 were CNV (median 6.5, range 2–15 for each sample) and 43 were SNV (median 2.5, range 0–7 for each sample). The full list of driving mutations is presented in Tables S2 and S3. The most prevalent alterations in our cohort were CNVs, *MCL1* (50%), *RB1* (43.75%), *TERT* (37.5%), *ARID1A*, *BAP1*, *BRCA1*, *CCND1*, *CDKN1B*, and *PTEN* (31.25%) being the most frequently affected genes.

The gene ontology analysis revealed enrichment in eight important signaling pathways: transcription factors (22.4%), chromatin (20.9%), cell cycle (19.4%), PIK3/mTOR (16.4%), RAS/MAPK (11.9%), JAK-STAT (2.9%), and cell adhesion (1.5%). Most of the driver genes and signaling pathways were shared by patients TNBC18 (primary treatment-naïve tumor and treated tumor) and TNBC41 (treated tumor and skin metastases). In contrast, the most heterogenous samples were TNBC11 and TNBC84, which shared only a few signaling pathways (Figure 2).

3.3. Mutational Processes Act on Different Stages of TNBC

To identify the mutational processes acting on TNBC, a mutational signature analysis using deconstructSigs in the R language was performed. Eleven of the 30 mutational signatures previously described by Alexandrov [28] were detected with a median of 3.5 per sample (range 1–5). Five of these signatures have been previously identified in BC, while the remaining six are identified for the first time in BC patients (Figure 3). Signature S01 (age-related) was prevalent in all samples and contributed the highest mutational rate. Signature S03 (homologous recombination deficiency) was detected in 81% of the samples, S06 (DNA mismatch repair deficiency) was detected in 44%, and S12 (unknown etiology) in 31%. Other signatures were found in fewer than 20% of the samples.

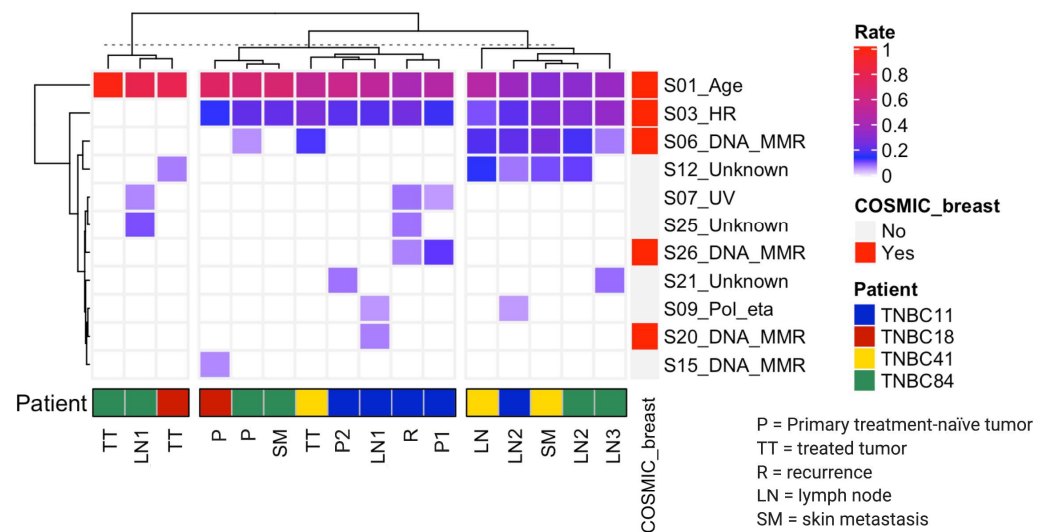


Figure 3. Mutational signature profile in TNBC patients. The samples are ordered according to the hierarchical grouping by mutational signature composition described above. The band of COSMIC breast describes the signatures found in the database of COSMIC Sanger, and the color palette shows the frequency of signatures in each sample. Rate is expressed as the relative frequency of the mutations associated with the corresponding signature. Mutational signatures previously identified in breast cancer samples in the COSMIC database are shown in red boxes. Patients are color-coded as in Figure 2. The tissue origin of the sample is indicated.

To identify potential common mutational processes active between all samples, we performed an unsupervised hierarchical clustering analysis, calculating Euclidean distances to determine the sample distribution of mutational signatures. This analysis classified the samples into three groups: (i) samples sharing a single mutational signature (3/16), (ii) those sharing two (8/16), and (iii) those sharing three or more (5/16) mutational signatures. Each group had unique characteristics. For instance, group two shared two mutational signatures (S01 and S03), albeit in different types of samples for each patient. On the other hand, group three includes signatures S01 and S03, but also acquired two additional signatures, S06 and S12. Interestingly, this group included lymph node and skin metastasis samples from three patients (TNBC11, TNBC41, TNBC84).

3.4. Early Divergence in TNBC

To determine the evolutionary history of the TNBC tumors, a phylogenetic analysis was performed. We classified SNV, indels, and CNV into three groups: (i) truncal (mutations found in all samples coming from the same patient), (ii) shared (mutations discovered in more than one sample of the same patients but not in all of them), and (iii) private (mutations found only in one sample of the same patient) (Figure 4a). A maximum parsimony algorithm [26] and the nonparametric bootstrapping methods [27] were used to estimate tumor phylogenies.

Each patient had a unique combination of truncal, shared, and private mutations, but we could observe that recurrence and treated tumors exhibited the highest number of private mutations, and lymph nodes had the fewest private mutations. These findings show a large genetic inpatient heterogeneity distributed in all the neoplastic lesions. For example, TNBC11 had five truncal mutations, 121 shared, and each sample had a different number of private mutations, with the recurrence sample containing the most private mutations. Patient TNBC18, on the other hand, had 136 truncal mutations, with 136 and 192 private mutations in the primary treatment-naïve tumor and treated tumor, respectively. TNBC41 had 38 truncal mutations, 66 shared, and the treated tumor was the sample that had the most private mutations (109). TNBC84 contained 11 truncal mutations, 92 shared, and the treated tumor was the most mutated, showing 118 mutations.

Based on the phylogenetic tree, three patients (TNBC11, TNBC41, and TNBC84) had the same evolutionary patterns (Figure 4b). Primary treatment-naïve tumor, treated tumor, recurrence, and skin metastasis samples were all identified in the same clade and shared the most recent common ancestor (MRCA). Skin metastasis and primary treatment-naïve tumor from TNBC84 were clustered in the same clade since these two samples shared numerous mutations. Interestingly, lymph nodes were clustered together suggesting an early divergence event in TNBC patients. On the other hand, lymph nodes had fewer driver genes, possibly indicating an evolutionary bottleneck. Notably, we could identify 136 truncal mutations in the two samples obtained from TNBC18. Unfortunately, we could not evaluate the evolutionary patterns of this patient due to the low number of samples available.

3.5. Potential Clinical Implications of Mutations Detected in TNBC

A specific search in the precision oncology knowledge base OncoKB [25] was performed to identify specific mutations associated with treatment response. Of note, TNBC is almost exclusively treated with chemotherapy [3]. Interestingly, we could identify 17 targets (3–8 targets per patient), distributed in seven signaling pathways. RAS/MAPK and PIK3/mTOR were the two signaling pathways with the highest number of actionable targets (Figure 5).

When we clustered actionable alterations based on detected mutations by tumor type, we observed seven distinct groups of tumors and their associated treatments. For instance, TNBC patients bearing mutations in *PIK3CA*, *AKT1*, and *ESR1*, as well as *EGFR* amplifications, may be treated with different drugs to those that are used to treat other subtypes of breast cancer (BRCA) and Non-Small Cell Lung Cancer (NSCLS), respectively. This helps to open new therapeutic windows in TNBC based on the potential candidate genes suitable for therapeutic precision agents.

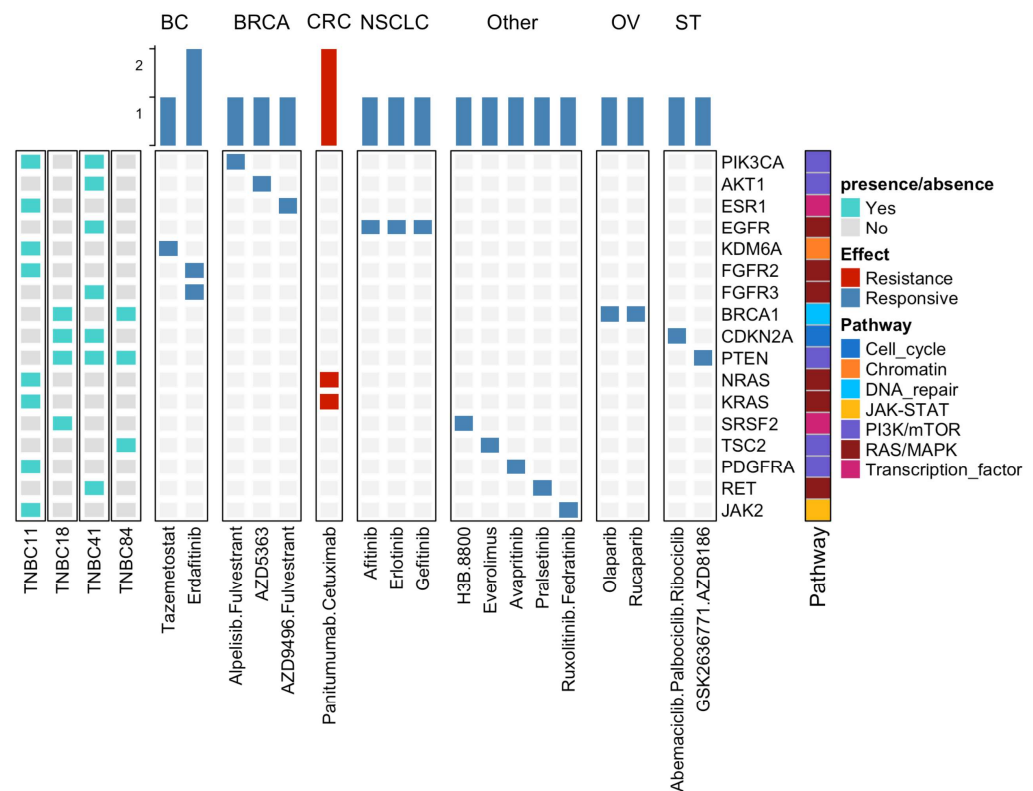


Figure 5. Actionable alterations in TNBC according to the oncoKB database. The figure depicts the actionable alterations found in our study and the corresponding FDA-approved drug. The drugs were segmented according to tumor type. Signaling pathways are indicated on the right and the patient identification on the left. BC: Bladder Cancer; BRCA: Breast Cancer; CRC: Colorectal Cancer; NSCLC: Non-Small Cell Lung Cancer; OV: Ovarian Cancer; and ST: Solid Tumors.

4. Discussion

The purpose of this study was to characterize the molecular features and evolution of tumors in patients with clinically aggressive TNBC. TNBC shows the worst prognosis of all breast cancers, with still unrevealed molecular aspects. We chose four patients from whom several follow-up samples were taken. We used a WES approach to characterize tumor heterogeneity, mutational signature profiles, driver genes, and tumor evolution in a total of 16 samples. Each patient was diagnosed at an advanced stage and presented a unique clinical scenario. These scenarios are in line with the complexity of this neoplasm and the difficulty for the development of specific treatments, given the wide genetic heterogeneity observed among patients [3]. In our cohort none of the patients achieved a complete response with NAC and all of them died within three years. Our findings are consistent with and similar to previously reported mortality rates in Latin America and worldwide [29–31].

In addition, we discovered an extensive inter- and inpatient genetic heterogeneity in terms of CNV and SNV, in line with previous studies [5,13,32–34]. SNV in *TP53*, found in the primary treatment-naïve tumor in TNBC11, and *MCL1* amplifications in the primary treatment-naïve tumors in TNBC11 and TNBC84, suggest that these alterations are early events in TNBC carcinogenesis. *TP53* is the most frequently encountered SNV in TNBC and is considered to occur early in TNBC carcinogenesis [5,32]. Conversely, *MCL1* amplifications have been associated with resistance to neoadjuvant chemotherapy and are frequently enriched in late-stage patients [35], which may explain why TNBC11, TNBC41, and TNBC84 patients did not achieve a PCR, although other alterations not detected in this study cannot be ruled out.

Deletions in *RB1* were predominantly found in the lymph nodes of two patients (TNBC11 and TNBC84) and deletions in *PTEN* were found in three (TNBC18, TNBC41

and TNBC84) of four patients, primarily in skin metastasis. These findings are similar to those of Bertucci et al., who identified five genes, including *RB1* and *PTEN*, enriched in metastatic TNBC [36], in addition to *ARID1A* alterations that were found to be enriched in metastatic TNBC lesions [14]. *TERT* amplifications were detected in treated tumors, lymph nodes, and skin metastasis of two patients (TNBC41 and TNBC84), as well as *AKT2* and *CCNE1* amplifications in lymph nodes of two patients (TNBC11 and TNBC84). These CNVs have been reported in other studies of advanced and metastatic breast cancer. For example, Watkins et al. found that *TERT* and *CCNE1* were subclonal and appeared at intermediate and late stages during the progression of 22 tumors, including breast cancers [37]. Overall, TNBC exhibited an extensive genetic heterogeneity, with numerous genes involved in various stages of TNBC carcinogenesis (Figure 6a).

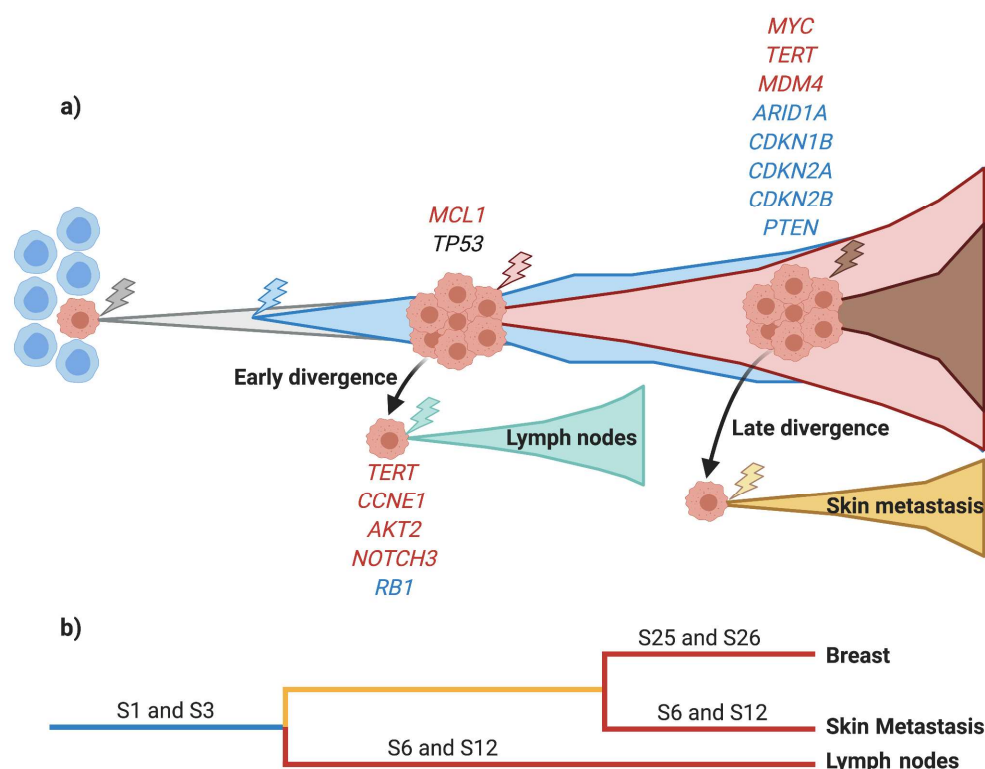


Figure 6. Summary of the key conclusions from our study. (a) Tumor evolution of TNBC seen in our cohort. We identified an early and late divergence to lymph nodes and skin metastasis, respectively. Driver genes with SNV (black), amplifications (red), and deletions (blue) are shown at each stage of TNBC carcinogenesis. (b) Phylogenetic tree of mutational signatures shows the main signatures acquired in each sample. Truncal mutational signatures are indicated in blue, shared signatures in yellow, and private signatures in red.

TNBC patients showed a heterogeneous TMB, with an average of 3.8 mutations (range 0.4–8.9). These results are comparable to other studies [31,38]. The treated tumors and recurrences in our cohort had the highest TMB in comparison to primary treatment-naïve tumor, lymph node, and skin metastasis samples. These TMB differences among tumor sites may be a result of tumor development per se, or the effect of neoadjuvant chemotherapy. In fact, it has been demonstrated that neoadjuvant chemotherapy alters the genetic landscape of tumors and promotes disease progression by increasing TMB in non-responder patients [39,40]. By contrast, lymph nodes show the lowest TMB, which may be an effect of an evolutionary bottleneck during cancer cell spread due to the tissue specific environmental constraints, a reduced number of invading cells (low genetic diversity), or reduced cellularity in these samples.

The analysis of mutational signatures revealed the presence of signature S01 in 100% of samples, S03 in 83%, S06 and S12 in less than 50%, and other signatures in less than 20% of samples. Unsupervised hierarchical clustering analysis based on Euclidean distances identified three groups of samples according to their common mutational signatures. Group two shared two signatures, S01 and S03, each of them detected in various tumor samples. Conversely, group three was the one showing more shared signatures, namely, S01, S03, S06, and S12. Metastatic lesions from lymph nodes and skin were the hallmarks in group three. In general, the mutational signature analysis revealed that advanced lesions are influenced by a more complex mutational environment. Five out of the 11 mutational signatures identified in our cohort had been associated with breast cancer, including S01, S03, S06, S20, and S26 [41], while the remaining six had not been detected in breast cancer until now. We discovered that certain signatures are active at various stages of the disease. Indeed, S01 and S03 were found in almost all samples, whereas S06 and S12 were found in lymph node and skin metastasis samples (Figure 6b). Mutational signature S01 is frequent in all type of tumors due to its association with spontaneous 5-methyl cytosine deamination accumulated with age. S03 has been found in at least 20% of all molecular types of breast cancer; however, in TNBC this signature constitutes more than 70% of the cases [41,42]. The mutational signatures S06, S20, and S26, associated with mismatch repair deficiency, were detected in our study, and are also prevalent in breast cancer [41,42]. Other studies, like the one from Gerstung et al., also reported that signatures associated with DNA mismatch repair deficiency were developed in the late stages of 38 different types of cancer, including breast cancer [34], and were associated with high-grade breast tumors [43]. Although there is no for evidence for the S12 signature in breast cancer, this signature has been detected in liver cancer, accounting for a small proportion of the total number of mutations (approximately 20%) found in this neoplasm. On the other hand, patients not achieving pCR after platinum-based NAC have been reported to exhibit some signatures associated with platinum exposure [39], including S04, S24, and S29, but these signatures were not observed in our cohort. Nevertheless, we do not rule out the prevalence of signatures associated with resistance to platinum-based NAC, which could be evidenced by whole genome sequencing.

Breast cancer cells spread to other tissues through hematogenous and lymphatic routes [44]. Understanding evolutionary patterns is fundamental to improving the comprehension of the progression of tumors and metastatic invasion to other tissues. In this study, we observed a defined evolutionary pattern in three out of four patients analyzed. The primary treatment-naïve tumor, treated tumors, recurrences, and skin metastases formed clusters in the same clade, indicating that they shared an MRCA. It has been hypothesized for many years that axillary lymph nodes constitute a reservoir for further dissemination to distant organs [45,46]. However, our study showed that skin metastasis was seeded directly from the primary treatment-naïve tumor, regardless of lymph node stage. In addition, lymph nodes formed an independent clade with a few mutations shared between the primary treatment-naïve tumor and other lesions, suggesting that spreading may occur early in these tissues. This same pattern has been reported in other molecular studies [13,32,47]. Clinical and molecular analyses in TNBC and breast cancer have also shown dissemination to lymph nodes from small T1 tumors (<2 cm), and the occurrence of metastatic spread in the early stages of tumor growth [48–50]. On the other hand, lymph nodes had fewer altered driver genes, indicating an evolutionary bottleneck. This bottleneck may be caused by the seeding of a small number of invasive cells into lymph nodes, which may result in a loss of genetic diversity. Reduced clonal heterogeneity in metastatic lymph nodes after NAC has been reported in breast, colorectal, and lung cancer patients [49]. Taken together, these findings support the early metastasis model and imply a rapid metastatic progression from early lesions in those patients with aggressive disease [51].

The only current systemic treatment approved for TNBC patients is chemotherapy. Given the high degree of genetic heterogeneity in TNBC, it is possible to use this molecular trait to identify candidates for targeted therapy. We detected 17 potential actionable alter-

ations clustered in seven altered signaling pathways, which are consistent with previous findings [5,31,52]. We identified seven groups of tumors and their associated treatments based on the type of mutations detected; these treatments are also used in other subtypes of breast cancer and other tumor types. Indeed, our analysis identified potential valuable genetic alterations that could be used in the treatment of TNBC.

Our study provides a detailed analysis of four patients with TNBC. The cancers were characterized molecularly and the wide genetic heterogeneity they showed may be impacting on the clinical treatment and the evolution of this disease. However, the limitations of this study include: the impossibility to correlate clinical and molecular features due to sample size limitations; and the fact that the cases were not homogenous in terms of time of progression and treatment, as each patient exhibited unique clinical characteristics and different stages of the disease, and so the treatments were heterogeneous. Another limitation of this study is that its design does not allow an evaluation of the role of ethnicity in the evolution of the disease. In addition, the use of FFPE samples and a potential low tumor cellularity in lymph node samples may have hampered the identification of the totality of mutations in these tissues. The clinical relevance of using WES in non-pCR patients longitudinally may allow an evaluation of the prevalence and composition of genetic driver alterations that can be used as potential actionable targets, and which evolve as resistance mechanisms. However, the lack of a pCR group limited the comparison of the evolution of tumors with better prognoses.

Despite these limitations, our data further supports previous findings from other studies and evidences the notorious inpatient genetic heterogeneity associated with all tumor lesions, as well as elucidating the molecular architecture and the genomic evolution of these aggressive neoplasms. The performed longitudinal analysis enabled the establishment of phylogenetic relationships between patients' neoplasms and the identification of an early divergency in lymph nodes. The mutational signatures S06 and S12 were found to be exclusive to lymph nodes and skin metastasis, and, finally, actionable alterations were identified for TNBC treatment. To our very best knowledge, this is the first study that integrates genomic alterations with evolutionary processes in a TNBC cohort from Latin America, which is, in turn, an underrepresented ethnicity in most genomic studies. Taken together, these findings suggest that the extensive inpatient genetic heterogeneity found in all malignant lesions may play a determinant role in the therapeutic failure, evolutionary adaptability, and aggressive clinical behavior of TNBC.

5. Conclusions

This study provided detailed insights into the genomic complexity and inpatient heterogeneity, as well as the phylogenetic and evolutionary development of TNBC. It also identified genomic features that have been associated with beneficial targeted treatments. We integrated the comprehensive genomic alterations with evolutionary techniques to understand the evolution and the genetic variability, as well as the description of each stage of the disease progression in the TNBC patients longitudinally. These findings contribute to the elucidation of the global phenomenon of tumor evolution in TNBC and provide genomic data that may aid future studies in defining actionable alterations for these patients, thereby enabling a more personalized therapy.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/cancers13205091/s1>, Table S1: Clinical characteristics; Table S2: Driver SNV; Table S3: Driver CNV.

Author Contributions: Conceptualization, H.M.-G. and F.V.-P.; methodology, H.M.-G., C.D.-V., R.Q.-U., E.R.-J., F.V.-L., F.I.P.-R., V.M.P.-S., H.A.M.-M. and M.R.-E.; validation, A.d.I.C.-M.; formal analysis, H.M.-G. and F.V.-P.; investigation, H.M.-G., J.C.M.-G., F.I.P.-R., V.M.P.-S., H.A.M.-M., M.R.-E., E.B.-R., P.C.-G., M.R.-R. and F.V.-P.; resources, E.B.-R. and L.A.H.; data curation, H.M.-G. and M.R.-E.; writing—original draft, H.M.-G. and F.V.-P.; writing—review and editing, H.M.-G., J.C.M.-G., Y.I.C., L.I.T., J.O., C.F., S.P. and F.V.-P.; visualization, H.M.-G. and F.V.-P.; supervision, C.D.-V. and F.V.-P.; project administration, F.V.-P.; funding acquisition, F.V.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by grants from UNAM PAPIIT IN219217, CONACyT Fondo Sectorial 272573, Fondo SEP CONACyT 285879, CONACyT National Laboratories 2021 project 315817.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and was approved by the Ethics Committee of National Cancer Institute of Mexico (protocol 016/013/IBI-CEI/1021/16).

Informed Consent Statement: Patient consent was waived due to the observational and retrospective design of the study, which did not affect patient management.

Data Availability Statement: The data presented in this study are available in the article and Supplementary Material. Raw data is available upon request.

Acknowledgments: Héctor Martínez Gregorio is a doctoral student from the Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México (UNAM) and has received a scholarship from CONACyT. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/World Health Organization.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Foulkes, W.D.; Smith, I.E.; Reis-Filho, J.S. Triple-negative breast cancer. *N. Engl. J. Med.* **2010**, *363*, 1938–1948. [[CrossRef](#)]
2. Harbeck, N.; Penault-Llorca, F.; Cortes, J.; Gnant, M.; Houssami, N.; Poortmans, P.; Ruddy, K.; Tsang, J.; Cardoso, F. Breast cancer. *Nat. Rev. Dis. Primer* **2019**, *5*, 66. [[CrossRef](#)]
3. Bianchini, G.; Balko, J.M.; Mayer, I.A.; Sanders, M.E.; Gianni, L. Triple-negative breast cancer: Challenges and opportunities of a heterogeneous disease. *Nat. Rev. Clin. Oncol.* **2016**, *13*, 674–690. [[CrossRef](#)]
4. Jones, S.; Chen, W.-D.; Parmigiani, G.; Diehl, F.; Beerenwinkel, N.; Antal, T.; Traulsen, A.; Nowak, M.A.; Siegel, C.; Velculescu, V.E.; et al. Comparative lesion sequencing provides insights into tumor evolution. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4283–4288. [[CrossRef](#)] [[PubMed](#)]
5. Yates, L.R.; Knappskog, S.; Wedge, D.; Farmery, J.H.R.; Gonzalez, S.; Martincorena, I.; Alexandrov, L.B.; Van Loo, P.; Haugland, H.K.; Lilleng, P.K.; et al. Genomic Evolution of Breast Cancer Metastasis and Relapse. *Cancer Cell* **2017**, *32*, 169–184. [[CrossRef](#)] [[PubMed](#)]
6. Valastyan, S.; Weinberg, R.A. Tumor metastasis: Molecular insights and evolving paradigms. *Cell* **2011**, *147*, 275–292. [[CrossRef](#)] [[PubMed](#)]
7. Buonomo, O.C.; Caredda, E.; Portarena, I.; Vanni, G.; Orlandi, A.; Bagni, C.; Petrella, G.; Palombi, L.; Orsaria, P. New insights into the metastatic behavior after breast cancer surgery, according to well-established clinicopathological variables and molecular subtypes. *PLoS ONE* **2017**, *12*, e0184680. [[CrossRef](#)] [[PubMed](#)]
8. Colleoni, M.; Zahrieh, D.; Gelber, R.D.; Holmberg, S.B.; Mattsson, J.E.; Rudenstam, C.-M.; Lindtner, J.; Erzen, D.; Snyder, R.; Collins, J.; et al. Site of primary tumor has a prognostic role in operable breast cancer: The international breast cancer study group experience. *J. Clin. Oncol.* **2005**, *23*, 1390–1400. [[CrossRef](#)] [[PubMed](#)]
9. Hussein, M.R.A. Skin metastasis: A pathologist's perspective. *J. Cutan. Pathol.* **2010**, *37*, e1–e20. [[CrossRef](#)]
10. Kalmykov, B.; Walker, S. Cutaneous metastases in breast cancer. *Clin. J. Oncol. Nurs.* **2011**, *15*, 99–101. [[CrossRef](#)] [[PubMed](#)]
11. Brastianos, P.K.; Carter, S.L.; Santagata, S.; Cahill, D.P.; Taylor-Weiner, A.; Jones, R.T.; Van Allen, E.M.; Lawrence, M.S.; Horowitz, P.M.; Cibulskis, K.; et al. Genomic Characterization of Brain Metastases Reveals Branched Evolution and Potential Therapeutic Targets. *Cancer Discov.* **2015**, *5*, 1164–1177. [[CrossRef](#)]
12. Brown, D.; Smeets, D.; Székely, B.; Larsimont, D.; Szász, A.M.; Adnet, P.-Y.; Rothé, F.; Rouas, G.; Nagy, Z.I.; Faragó, Z.; et al. Phylogenetic analysis of metastatic progression in breast cancer using somatic mutations and copy number aberrations. *Nat. Commun.* **2017**, *8*, 14944. [[CrossRef](#)]
13. Ullah, I.; Karthik, G.-M.; Alkods, A.; Kjällquist, U.; Stålhammar, G.; Lövrot, J.; Martinez, N.-F.; Lagergren, J.; Hautaniemi, S.; Hartman, J.; et al. Evolutionary history of metastatic breast cancer reveals minimal seeding from axillary lymph nodes. *J. Clin. Investig.* **2018**, *128*, 1355–1370. [[CrossRef](#)]

14. Angus, L.; Smid, M.; Wilting, S.M.; van Riet, J.; Van Hoeck, A.; Nguyen, L.; Nik-Zainal, S.; Steenbruggen, T.G.; Tjan-Heijnen, V.C.G.; Labots, M.; et al. The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nat. Genet.* **2019**, *51*, 1450–1458. [[CrossRef](#)]
15. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinform. Oxf. Engl.* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
16. Van der Auwera, G.A.; Carneiro, M.O.; Hartl, C.; Poplin, R.; Del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J.; et al. From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **2013**, *43*, 1–11. [[CrossRef](#)]
17. Cibulskis, K.; Lawrence, M.S.; Carter, S.L.; Sivachenko, A.; Jaffe, D.; Sougnez, C.; Gabriel, S.; Meyerson, M.; Lander, E.S.; Getz, G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **2013**, *31*, 213–219. [[CrossRef](#)] [[PubMed](#)]
18. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [[CrossRef](#)]
19. Tate, J.G.; Bamford, S.; Jubb, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creatore, C.; Dawson, E.; et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **2019**, *47*, D941–D947. [[CrossRef](#)]
20. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **2013**, *6*, p11. [[CrossRef](#)]
21. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **2013**, *14*, 178–192. [[CrossRef](#)] [[PubMed](#)]
22. Talevich, E.; Shain, A.H.; Botton, T.; Bastian, B.C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput. Biol.* **2016**, *12*, e1004873. [[CrossRef](#)] [[PubMed](#)]
23. Rosenthal, R.; McGranahan, N.; Herrero, J.; Taylor, B.S.; Swanton, C. DeconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* **2016**, *17*, 31. [[CrossRef](#)] [[PubMed](#)]
24. Dennis, G.; Sherman, B.T.; Hosack, D.A.; Yang, J.; Gao, W.; Lane, H.C.; Lempicki, R.A. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **2003**, *4*, R60. [[CrossRef](#)] [[PubMed](#)]
25. Chakravarty, D.; Gao, J.; Phillips, S.M.; Kundra, R.; Zhang, H.; Wang, J.; Rudolph, J.E.; Yaeger, R.; Soumerai, T.; Nissan, M.H.; et al. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis. Oncol.* **2017**, *1*, 1–16. [[CrossRef](#)]
26. Fitch, W.M. Toward Defining the Course of Evolution: Minimum Change for a Specific Tree Topology. *Syst. Biol.* **1971**, *20*, 406–416. [[CrossRef](#)]
27. Efron, B.; Halloran, E.; Holmes, S. Bootstrap confidence levels for phylogenetic trees. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13429–13434. [[CrossRef](#)]
28. Alexandrov, L.B.; Nik-Zainal, S.; Wedge, D.C.; Aparicio, S.A.J.R.; Behjati, S.; Biankin, A.V.; Bignell, G.R.; Bolli, N.; Borg, A.; Børresen-Dale, A.-L.; et al. Signatures of mutational processes in human cancer. *Nature* **2013**, *500*, 415–421. [[CrossRef](#)]
29. Li, X.; Yang, J.; Peng, L.; Sahin, A.A.; Huo, L.; Ward, K.C.; O'Regan, R.; Torres, M.A.; Meisel, J.L. Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. *Breast Cancer Res. Treat.* **2017**, *161*, 279–287. [[CrossRef](#)]
30. Reynoso-Noverón, N.; Villarreal-Garza, C.; Soto-Perez-de-Celis, E.; Arce-Salinas, C.; Matus-Santos, J.; Ramírez-Ugalde, M.T.; Alvarado-Miranda, A.; Cabrera-Galeana, P.; Meneses-García, A.; Lara-Medina, F.; et al. Clinical and Epidemiological Profile of Breast Cancer in Mexico: Results of the Seguro Popular. *J. Glob. Oncol.* **2017**, *3*, 757–764. [[CrossRef](#)]
31. Rojas-Jiménez, E.; Mejía-Gómez, J.C.; Díaz-Velásquez, C.; Quezada-Urban, R.; Martínez Gregorio, H.; Vallejo-Lecuona, F.; de la Cruz-Montoya, A.; Porras Reyes, F.I.; Pérez-Sánchez, V.M.; Maldonado-Martínez, H.A.; et al. Comprehensive Genomic Profile of Heterogeneous Long Follow-Up Triple-Negative Breast Cancer and Its Clinical Characteristics Shows DNA Repair Deficiency Has Better Prognostic. *Genes* **2020**, *11*, 1367. [[CrossRef](#)] [[PubMed](#)]
32. Barry, P.; Vatsiou, A.; Spiteri, I.; Nichol, D.; Cresswell, G.D.; Acar, A.; Trahearn, N.; Hrebien, S.; Garcia-Murillas, I.; Chkhaidze, K.; et al. The Spatiotemporal Evolution of Lymph Node Spread in Early Breast Cancer. *Clin. Cancer Res.* **2018**, *24*, 4763–4770. [[CrossRef](#)]
33. Ciriello, G.; Miller, M.L.; Aksoy, B.A.; Senbabaoglu, Y.; Schultz, N.; Sander, C. Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* **2013**, *45*, 1127–1133. [[CrossRef](#)] [[PubMed](#)]
34. Gerstung, M.; Jolly, C.; Leshchiner, I.; D'Entro, S.C.; Gonzalez, S.; Rosebrock, D.; Mitchell, T.J.; Rubanova, Y.; Anur, P.; Yu, K.; et al. The evolutionary history of 2,658 cancers. *Nature* **2020**, *578*, 122–128. [[CrossRef](#)] [[PubMed](#)]
35. Balko, J.M.; Giltner, J.M.; Wang, K.; Schwarz, L.J.; Young, C.D.; Cook, R.S.; Owens, P.; Sanders, M.E.; Kuba, M.G.; Sánchez, V.; et al. Molecular profiling of the residual disease of triple-negative breast cancers after neoadjuvant chemotherapy identifies actionable therapeutic targets. *Cancer Discov.* **2014**, *4*, 232–245. [[CrossRef](#)] [[PubMed](#)]
36. Bertucci, F.; Ng, C.K.Y.; Patsouris, A.; Droin, N.; Piscuoglio, S.; Carbuccia, N.; Soria, J.C.; Dien, A.T.; Adnani, Y.; Kamal, M.; et al. Genomic characterization of metastatic breast cancers. *Nature* **2019**, *569*, 560–564. [[CrossRef](#)] [[PubMed](#)]
37. Watkins, T.B.K.; Lim, E.L.; Petkovic, M.; Elizalde, S.; Birkbak, N.J.; Wilson, G.A.; Moore, D.A.; Grönroos, E.; Rowan, A.; Dewhurst, S.M.; et al. Pervasive chromosomal instability and karyotype order in tumour evolution. *Nature* **2020**, *587*, 126–132. [[CrossRef](#)] [[PubMed](#)]

38. Kandoth, C.; McLellan, M.D.; Vandin, F.; Ye, K.; Niu, B.; Lu, C.; Xie, M.; Zhang, Q.; McMichael, J.F.; Wyczalkowski, M.A.; et al. Mutational landscape and significance across 12 major cancer types. *Nature* **2013**, *502*, 333–339. [[CrossRef](#)]
39. Pich, O.; Muiños, F.; Lolkema, M.P.; Steeghs, N.; Gonzalez-Perez, A.; Lopez-Bigas, N. The mutational footprints of cancer therapies. *Nat. Genet.* **2019**, *51*, 1732–1740. [[CrossRef](#)]
40. Szikriszt, B.; Póti, Á.; Pipek, O.; Krzystanek, M.; Kanu, N.; Molnár, J.; Ribli, D.; Szeltner, Z.; Tusnády, G.E.; Csabai, I.; et al. A comprehensive survey of the mutagenic impact of common cancer cytotoxics. *Genome Biol.* **2016**, *17*, 99. [[CrossRef](#)] [[PubMed](#)]
41. Nik-Zainal, S.; Morganella, S. Mutational Signatures in Breast Cancer: The Problem at the DNA Level. *Clin. Cancer Res.* **2017**, *23*, 2617–2629. [[CrossRef](#)] [[PubMed](#)]
42. Lefebvre, C.; Bachelot, T.; Filleron, T.; Pedrero, M.; Campone, M.; Soria, J.-C.; Massard, C.; Lévy, C.; Arnedos, M.; Lacroix-Triki, M.; et al. Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Med.* **2016**, *13*, e1002201. [[CrossRef](#)] [[PubMed](#)]
43. Cheng, A.S.; Leung, S.C.Y.; Gao, D.; Burugu, S.; Anurag, M.; Ellis, M.J.; Nielsen, T.O. Mismatch repair protein loss in breast cancer: Clinicopathological associations in a large British Columbia cohort. *Breast Cancer Res. Treat.* **2020**, *179*, 3–10. [[CrossRef](#)]
44. Chitty, J.L.; Filipe, E.C.; Lucas, M.C.; Herrmann, D.; Cox, T.R.; Timpson, P. Recent advances in understanding the complexities of metastasis. *F1000Research* **2018**, *7*. [[CrossRef](#)] [[PubMed](#)]
45. Halsted, W.S.I. The Results of Operations for the Cure of Cancer of the Breast Performed at the Johns Hopkins Hospital from June, 1889, to January, 1894. *Ann. Surg.* **1894**, *20*, 497–555. [[CrossRef](#)]
46. Patey, D.H.; Dyson, W.H. The prognosis of carcinoma of the breast in relation to the type of operation performed. *Br. J. Cancer* **1948**, *2*, 7–13. [[CrossRef](#)] [[PubMed](#)]
47. Venet, D.; Fimereli, D.; Rothé, F.; Boeckx, B.; Maetens, M.; Majjaj, S.; Rouas, G.; Capra, M.; Bonizzi, G.; Contaldo, F.; et al. Phylogenetic reconstruction of breast cancer reveals two routes of metastatic dissemination associated with distinct clinical outcome. *EBioMedicine* **2020**, *56*, 102793. [[CrossRef](#)]
48. Chen, X.; Yu, X.; Chen, J.; Zhang, Z.; Tuan, J.; Shao, Z.; Guo, X.; Feng, Y. Analysis in early stage triple-negative breast cancer treated with mastectomy without adjuvant radiotherapy: Patterns of failure and prognostic factors. *Cancer* **2013**, *119*, 2366–2374. [[CrossRef](#)]
49. Hu, Z.; Li, Z.; Ma, Z.; Curtis, C. Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nat. Genet.* **2020**, *52*, 701–708. [[CrossRef](#)]
50. Wang, X.-X.; Jiang, Y.-Z.; Li, J.-J.; Song, C.-G.; Shao, Z.-M. Effect of nodal status on clinical outcomes of triple-negative breast cancer: A population-based study using the SEER 18 database. *Oncotarget* **2016**, *7*, 46636–46645. [[CrossRef](#)]
51. Hunter, K.W.; Amin, R.; Deasy, S.; Ha, N.-H.; Wakefield, L. Genetic insights into the morass of metastatic heterogeneity. *Nat. Rev. Cancer* **2018**, *18*, 211–223. [[CrossRef](#)] [[PubMed](#)]
52. Sanchez-Vega, F.; Mina, M.; Armenia, J.; Chatila, W.K.; Luna, A.; La, K.C.; Dimitriadoy, S.; Liu, D.L.; Kantheti, H.S.; Saghafinia, S.; et al. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **2018**, *173*, 321–337. [[CrossRef](#)] [[PubMed](#)]