



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
DOCTORADO EN CIENCIAS BIOMÉDICAS
FACULTAD DE MEDICINA
INSTITUTO NACIONAL DE MEDICINA GENÓMICA

**DETERMINACIÓN DE LA PREVALENCIA DE SUBTIPOS TUMORALES DE CÁNCER
DE MAMA MEDIANTE PERFILES DE EXPRESIÓN GÉNICA Y MARCADORES
INMUNOHISTOQUÍMICOS**

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS BIOMÉDICAS

PRESENTA:
M EN C. ROSA GLORIA REBOLLAR VEGA

DIRECTOR DE TESIS
DR. ALFREDO HIDALGO MIRANDA
FACULTAD DE MEDICINA-INMEGEN

COMITÉ TUTOR

DRA. PATRICIA OSTROSKY SHEJET DR. SAMUEL CANIZALES QUINTEROS
INSTITUTO DE INVESTIGACIONES BIOMÉDICA FACULTAD DE QUÍMICA, UNAM-INMEGEN

CDMX, MARZO, 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.

Este trabajo fue realizado en el Instituto de Enfermedades de la Mama (FUCAM), Instituto Nacional de Medicina Genómica (INMEGEN) en el laboratorio de Genómica del Cáncer, bajo la supervisión del Dr. Alfredo Hidalgo Miranda.

El proyecto fue financiado por CONACYT, beca de doctorado no. 294330

El comité tutorial estuvo conformado por:

Dr. Alfredo Hidalgo Miranda

Tutor principal

Instituto Nacional de Medicina Genómica, INMEGEN.

Dr. Samuel Canizales Quinteros

Instituto de Química, UNAM-Instituto Nacional de Medicina Genómica, INMEGEN.

Dra. Martha Patricia Ostrosky Shejet

Instituto de Investigaciones Biomédicas, UNAM

AGRADECIMIENTOS

A TODAS LAS PACIENTES DEL FUCAM



AL LABORATORIO DE ONCOLOGÍA GENÓMICA, UNIDAD DE GENOTIPIFICACIÓN Y ANÁLISIS DE EXPRESIÓN AFFYMETRIX Y A LA UNIDAD DE VALIDACIÓN DE BIOMARCADORES DEL INSTITUTO NACIONAL DE MEDICINA GENÓMICA.



INDICE

RESUMEN	6
ABSTRACT	7
INTRODUCCIÓN.....	8
EPIDEMIOLOGÍA DEL CÁNCER DE MAMA	8
GLÁNDULA MAMARIA	10
FACTORES DE RIESGO.....	12
CÁNCER DE MAMA.....	13
CLASIFICACIÓN HISTOLÓGICA DEL CÁNCER DE MAMA.....	15
CARACTERIZACIÓN MOLECULAR DEL CÁNCER DE MAMA.....	18
MEDICINA DE PRECISIÓN EN CÁNCER DE MAMA	22
JUSTIFICACIÓN	30
HIPOTESIS.....	30
OBJETIVOS	31
GENERAL	31
PARTICULARES	31
MATERIALES Y MÉTODOS.....	32
INMUNOHISTOQUÍMICA	33
EXTRACCIÓN DE ÁCIDOS NUCLEICOS	34
PERFILES DE EXPRESIÓN GÉNICA DE MRNA Y SUBTIPOS PAM50	34
ANÁLISIS DE GENES EXPRESADOS DIFERENCIALMENTE Y ENRIQUECIMIENTO DE GENES.....	36
VALIDACIÓN DE GENES DIFERENCIALMENTE EXPRESADOS	36
ANÁLISIS DE NÚMERO DE COPIAS DE DNA	38
ANÁLISIS DE ANCESTRÍA	39
RESULTADOS	40
EVALUACIÓN DE INMUNOHISTOQUÍMICA.....	40
ANÁLISIS DE SUBTIPOS INTRÍNSECOS DE CÁNCER DE MAMA.....	43

CORRELACIÓN INMUNOHISTOQUÍMICA- SUBTIPOS INTRÍNSECOS	48
<u>DISCUSIÓN</u>	<u>60</u>
<u>CONCLUSIONES.....</u>	<u>65</u>
<u>PERSPECTIVAS</u>	<u>66</u>
<u>REFERENCIAS</u>	<u>67</u>
<u>PUBLICACIONES</u>	<u>72</u>

RESUMEN

El cáncer de mama es la neoplasia más común en mujeres a nivel mundial por lo que representa un grave problema de salud pública. En México, los tumores de mama representan la principal causa de cáncer en mujeres. En la práctica clínica, los marcadores inmunohistoquímicos, receptor de estrógeno; de progesterona y el receptor 2 del factor de crecimiento epidérmico humano se utilizan para orientar el diagnóstico y las decisiones de tratamiento. La clasificación molecular de los tumores de mama según los patrones de expresión génica también se ha traducido con éxito en pruebas clínicas para ayudar en la toma de decisiones médicas. El uso de perfiles de expresión génica, particularmente la firma para identificación de subtipos intrínsecos PAM50, ha revelado la existencia de al menos cuatro subcategorías de tumores de mama: Luminal A, Luminal B, enriquecida en Her2 y Basal. Por lo anterior, se realizó un análisis molecular completo de 105 tumores de mama y 21 tejidos adyacentes al tumor. Obteniendo los siguientes datos, la edad media de diagnóstico fue de 55 años, el 22% era menor de 45 años. La mayoría de los casos fueron principalmente en etapa temprana IIA y IIB. En promedio, el tamaño del tumor fue de 2,3 cm, el 67% del diagnóstico histológico fue cáncer ductal infiltrante, seguido de carcinoma lobular infiltrante y tumores mixtos (ductal y lobular). La mayoría de los tumores analizados presentan un grado histológico II y III. En cuanto al estatus de marcadores de inmunohistoquímica: 66.6% de los tumores fueron ER +, 56.19% PR +, 18.09% HER2/Neu positivo, 11% triple negativo (RE, RP y HER2/Neu negativo), 35.8% positivos a Citoqueratina 5/6 (CK5/6), 50% a Citoqueratina 14 (CK14) y 43.39% a Citoqueratina 17 (CK17). EGFR fue evaluado en los tumores del subtipo basal, encontrando un 33.3% de positividad, mientras que CLD1 y CLD3 tienen una frecuencia del 100%. Obtuvimos la frecuencia de subtipos intrínsecos en cáncer de mama en población mexicana de acuerdo a PAM50, 46% de tumores LumA, 28% de LumB, 15% de HER2 enriquecido y 12% de tipo basal. Sin embargo, no se encontró buena correlación entre la sub-clasificación basada en el PAM50 contra el perfil inmunohistoquímica. Se validaron los genes diferencialmente expresados. Se identificaron genes y se correlacionaron a nivel de alteraciones de número de copias que pudieran implicar en el desarrollo del cáncer de mama. Los marcadores inmunohistoquímicos han guiado tradicionalmente las decisiones de tratamiento en el cáncer de mama. Sin embargo, los avances en los perfiles de expresión génica y la disponibilidad de ensayos basados en genes han lanzado estas nuevas pruebas a la práctica clínica diaria. El análisis global de los perfiles de expresión génica, permitió obtener un mayor conocimiento acerca de las diferencias biológicas entre subtipos de tumores de

mama, además de obtener datos acerca de los mecanismos moleculares que se alteren en el cáncer de mama describiendo la prevalencia de subtipos de tumores de mama definidos molecularmente en nuestra población.

ABSTRACT

Breast cancer is the most common neoplasm in women worldwide and represents a serious public health problem. In Mexico, breast tumors represent the main cause of cancer in women. In clinical practice, immunohistochemical markers of estrogen receptor, progesterone receptor and human epidermal growth factor receptor 2 are used to guide diagnosis and treatment decisions. Molecular classification of breast tumors according to gene expression patterns has also been successfully translated into clinical trials to aid in medical decision making. The use of gene expression profiling, particularly the signature for identification of intrinsic PAM50 subtypes, has revealed the existence of at least four subcategories of breast tumors: luminal A, luminal B, Her2-enriched, basal-like. A comprehensive molecular analysis of 105 breast tumors and 21 tumor-adjacent tissues was performed. The following data were obtained, the mean age at diagnosis was 55 years, 22% were younger than 45 years. Most cases were mainly early stage IIA and IIB. On average, the tumor size was 2.3 cm, 67.61% of the histological diagnosis was infiltrating ductal, followed by infiltrating lobular carcinoma and mixed tumors (ductal and lobular). Most of the tumors analyzed presented histological grade II and III. Regarding the status of immunohistochemical markers: 66.6% of the tumors were ER +, 56.19% PR +, 18.09% HER2/Neu positive, 11% triple negative (ER, PR and HER2/Neu negative), 35.8% positive for cytokeratin 5/6 (CK5/6), 50% for cytokeratin 14 (CK14) and 43.39% for cytokeratin 17 (CK17). EGFR was evaluated in basal subtype tumors, finding 33.3% positivity, while CLD1 and CLD3 had a frequency of 100%. According to PAM50, it is subdivided into 46% LumA, 28% LumB, 15% HER2-enriched and 12% basal type tumors. Immunohistochemical markers have traditionally guided treatment decisions in breast cancer. However, advances in gene-expression profiling and availability of gene-based assays have launched these newer tests into everyday clinical practice. The global analysis of gene expression profiles allowed us to obtain a better understanding of the biological differences between breast tumor subtypes, as well as to obtain data about the molecular mechanisms that are altered in breast cancer, describing the prevalence of molecularly defined breast tumor subtypes in our population.

INTRODUCCIÓN

Epidemiología del cáncer de mama

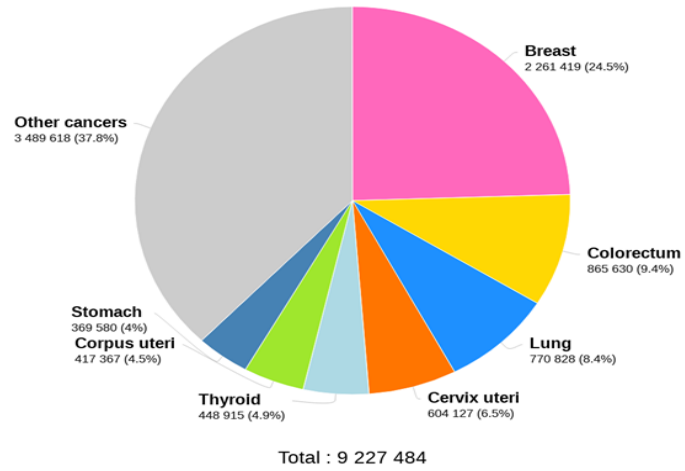
El cáncer es un problema de salud a nivel mundial, Globocan ha estimado un total de 9, 227, 48 de casos de cáncer en el 2020¹. El cáncer de mama es el tumor maligno más frecuente en las mujeres y la primera causa de muerte por cáncer (Fig 1). En 2018, se estimó que 2,1 millones de mujeres fueron diagnosticadas con cáncer de mama, aproximadamente un caso cada 18 segundos².

En México, el cáncer de mama ha tenido un incremento constante tanto en su incidencia como mortalidad en las últimas tres décadas³. Acorde al reporte del Departamento de Epidemiología de la Secretaría de Salud la incidencia se incrementó entre 2000 y 2013 llegando de 10.76 casos por 100,000 habitantes a 26.1 por cada 100,000 mujeres mayores de 25 años, estimando 23,873 nuevos casos en 2013⁴.

El cáncer de mama en México se diagnostica a una media de edad de 52.5 años, una década inferior a la población de América del Norte y Europa occidental, la incidencia aumenta proporcionalmente al incremento en la edad, lo que nos hace prever que solamente por el envejecimiento de la población, habrá también un aumento sustancial en el número de casos de cáncer de mama en los próximos años⁵.

Existen diferencias importantes en la incidencia y mortalidad de la enfermedad entre los estados del norte y centro de la República Mexicana. Esto es debido entre otras razones, al cambio del estilo de vida de la población, que ha adoptado el modelo donde las mujeres en general tienen un nivel de escolaridad más alto y trabajan fuera del hogar, con mayor ingesta también de grasas animales, tabaco y alcohol, sedentarismo y sobrepeso, retraso en el inicio de la reproducción, con primer embarazo tardío (> de 30 años) y ausencia de lactancias, y uso de agentes hormonales en la menopausia⁶.

Aunado a esto, la carencia de un programa nacional de detección oportuna, la atención en las instituciones públicas, la insuficiencia de recursos humanos, materiales y técnicos para el tratamiento y la falta de unidades mamarias especializadas⁴.



Data source: Globocan 2020
 Graph production: Global Cancer Observatory (<http://gco.iarc.fr>)

International Agency for Research on Cancer
 World Health Organization

Figura 1. Estimación del número de casos de cáncer en mujeres 2020. Globocan 2022²

Glándula mamaria

La glándula mamaria es un órgano complejo y altamente adaptable cuya función principal es producir leche durante la lactancia para el sustento de la descendencia. La glándula está compuesta por una variedad de tipos de células, que incluyen fibroblastos, adipocitos, células epiteliales, endoteliales e inmunes. Las células epiteliales se pueden subdividir en múltiples tipos de células que juntas forman la estructura ramificada de la glándula y constituyen los alvéolos secretores durante la lactancia (Figura 2)⁷.

Los conductos comprenden células luminales (capa interna) y células basales (capa externa) y están rodeadas por la membrana basal, fibroblastos y adipocitos. Al nacer, el sistema de conductos es rudimentario. Las hormonas como el estrógeno y la progesterona desencadenan en la pubertad: la diferenciación celular y la formación de brotes terminales (TEB) que se invaginan en la almohadilla adiposa para promover la expansión ductal. Con la concepción, interfiere otra hormona, la prolactina que ocasiona la diferenciación de las células alveolares Luminales y los conductos terciarios y las estructuras lóbulo alveolares comienzan a formarse durante el embarazo temprano. Así comienza la diferenciación de las células epiteliales mamarias (MEC)⁸. Al final de la gestación, las células alveolares Luminales aumentan la producción de leche, mientras que la oxitocina durante la lactancia desencadena la contracción de las células epiteliales, la constricción Luminal y la liberación de leche al lumen⁹. La fase inicial de involución es reversible, durante la siguiente fase irreversible se produce apoptosis y adipogénesis de MEC, los macrófagos eliminan los MEC muertos y las células alveolares sufren apoptosis. Mientras tanto, la matriz extracelular (ECM) se reconstituye con colágeno altamente fibrilar y la adipogénesis continúa¹⁰.

Varios mecanismos celulares, incluida la capacidad de respuesta a las hormonas circulantes, la composición del estroma y la especificación celular, contribuyen a los efectos de la edad y la paridad en el riesgo de cáncer de mama⁷.

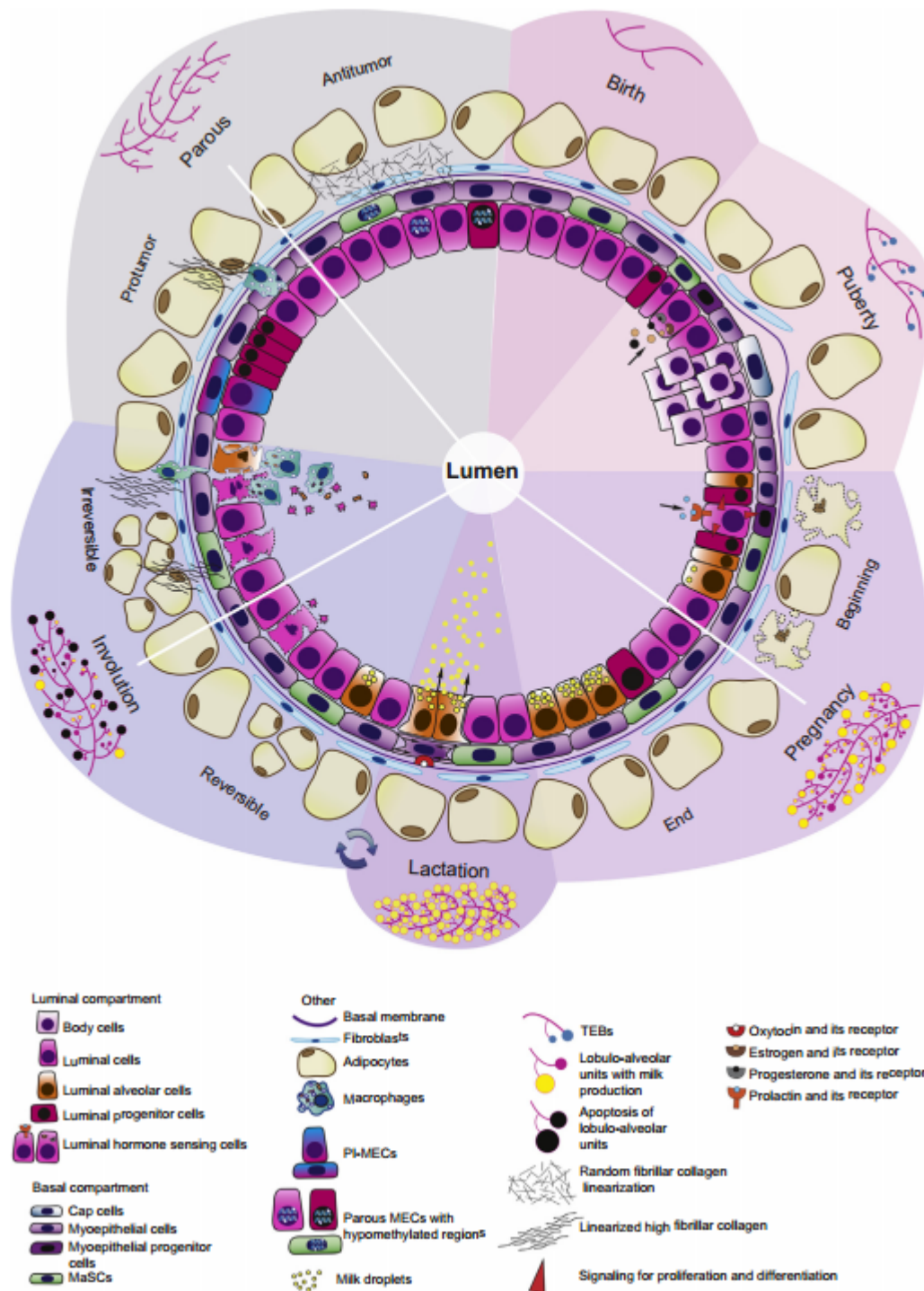


Figura 2. Glándula mamaria. Representación esquemática de un conducto mamario a lo largo del ciclo de desarrollo mamario: nacimiento, pubertad, embarazo, lactancia, involución y post-embarazo. Tomada de Pregnancy and Breast Cancer: Pathways to Understand Risk and Prevention⁷.

Factores de riesgo

Los factores de riesgo conocidos para el desarrollo del cáncer de mama son los siguientes:
Biológicos: sexo femenino, envejecimiento (a mayor edad mayor riesgo), antecedente personal o familiar de cáncer de mama en madre, hijas o hermanas. Antecedentes de hallazgos de hiperplasia ductal atípica, imagen radial o estrellada, así como carcinoma lobulillar in situ por biopsia. Vida menstrual mayor de 40 años (menarquía antes de los 12 años y menopausia después de los 52 años). Densidad mamaria. Ser portador de los genes BRCA1 o BRCA2⁴.

Ambientales: Exposición a radiaciones ionizantes, principalmente durante el desarrollo o el crecimiento. Radioterapia en el tórax.

Factores de riesgo relacionados con los antecedentes reproductivos: Nuligesta, primer embarazo a término después de los 30 años de edad. Terapia hormonal en la peri menopausia o posmenopausia por más de 5 años.

Factores de riesgo relacionados con el estilo de vida: Alimentación rica en carbohidratos y baja en fibra. Dieta rica en grasas tanto animales como ácidos grasos trans. Obesidad, Sedentarismo. Consumo de alcohol mayor de 15 g/día y Tabaquismo⁵.

Cáncer de mama

El mecanismo exacto por el cual se inicia el cáncer de mama se desconoce; sin embargo, se han hecho muchos esfuerzos para caracterizar molecularmente el cáncer de mama y proyectar su formación y progresión¹¹.

Las características distintivas del cáncer comprenden seis capacidades biológicas adquiridas durante el desarrollo de tumores en múltiples etapas. Estos hallmarks o sellos, constituyen un principio organizador para racionalizar las complejidades de esta enfermedad. Incluyen señales de proliferación, la evasión de los supresores de crecimiento, resistencia a la muerte celular, la posibilidad de la inmortalidad replicativa, inducción de angiogénesis y la activación de la invasión y la metástasis. Además está la inestabilidad del genómica, que genera la diversidad genética y la inflamación fomentando múltiples funciones distintivas¹². Se han añadido la reprogramación del metabolismo energético y la evasión de la destrucción inmunitaria. Aparte de las células cancerosas, los tumores exhiben otra dimensión de complejidad con un repertorio de células reclutadas, visiblemente normales, que contribuyen a la adquisición de rasgos distintivos al crear el microambiente tumoral. Recientemente se planteó la posibilidad de que la plasticidad fenotípica y la diferenciación alterada sean una capacidad distintiva, y que la reprogramación epigenéticas no mutacional y los microbiomas polimórficos constituyan características que faciliten la adquisición de capacidades distintivas. Además, las células senescentes de diversos orígenes, pueden añadirse a la lista de tipos celulares funcionalmente importantes en el microambiente tumoral (Fig. 3). El reconocimiento de la amplia aplicabilidad de estos conceptos afectará cada vez más el desarrollo de nuevos medios para tratar el cáncer¹³.

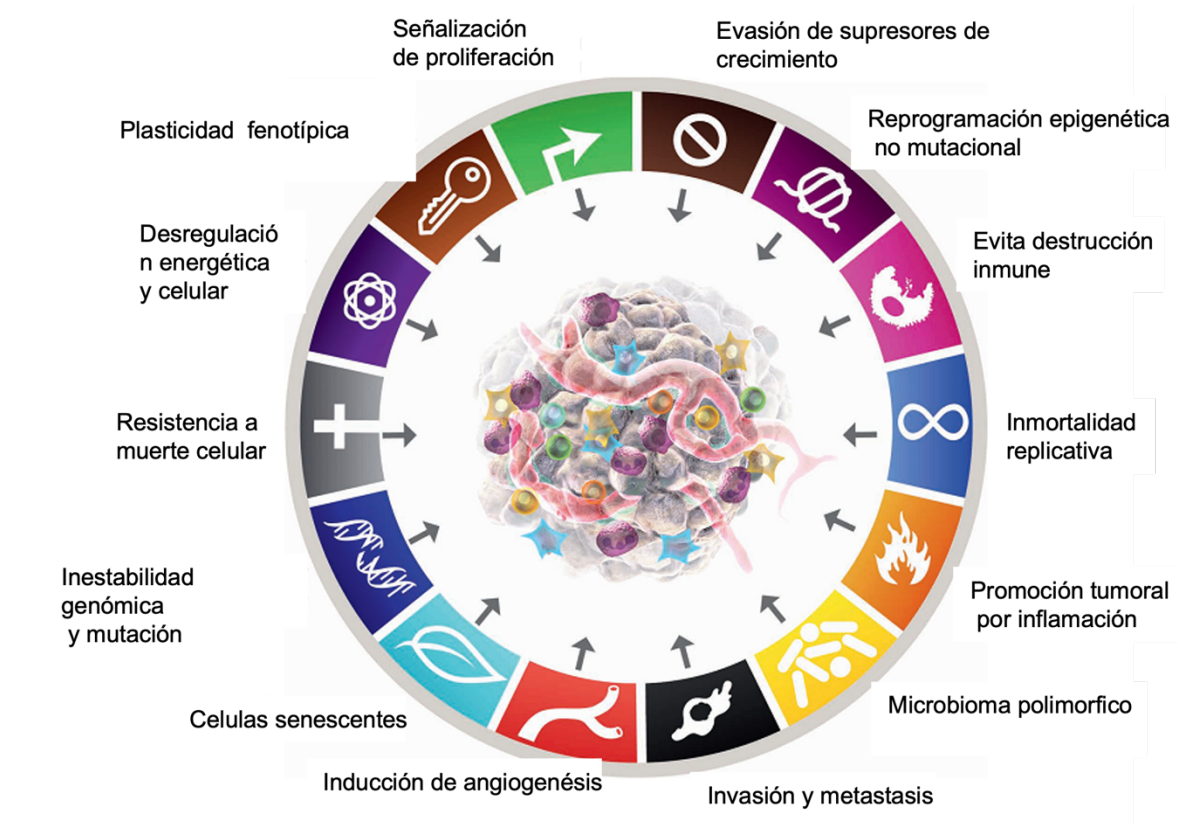


Figura 3. Características distintivas del cáncer. Modificada de Hanahan, 2022¹³

Clasificación Histológica del Cáncer de mama

El análisis histológico de las biopsias de mama combinado con análisis de sangre y mamografías permite clasificar los tumores de mama de acuerdo a su capacidad de invasión, tamaño, morfología, crecimiento y si han alcanzado los ganglios linfáticos (LN). Los tumores invasivos o no invasivos dependen de la capacidad o incapacidad de las células cancerosas para diseminarse en estructuras cercanas a los conductos de la mama. Un sistema de clasificación del cáncer de mama es el de estadios 0 a IV, lo que permite el pronóstico del tumor y la terapia potencial. De acuerdo a una variedad de parámetros, incluyendo el tamaño del tumor y la presencia de células cancerosas en los LN. TNM es otro sistema de clasificación basado en el tamaño del tumor (T), la participación de LN (N), y la metástasis (M) (www.breastcancer.org).

En cuanto a análisis histológico en el cáncer de mama se pueden definir varios tipos que se mencionan a continuación: El carcinoma ductal *in situ* es un grupo heterogéneo de neoplasias caracterizado por la presencia de células epiteliales malignas que crecen dentro de los conductos mamarios, sin rebasar la membrana basal, identificadas por microscopia de luz. Adopta diferentes patrones arquitectónicos de crecimiento intraductal y presenta características citológicas y de necrosis variables; generalmente es unifocal, se conoce también con el nombre de carcinoma intraductal.

Se reconoce que el carcinoma intraductal crece frecuentemente dentro de los ductos de manera discontinua. Estos carcinomas son inicialmente sospechados a raíz de un hallazgo mamográfico anormal (micro calcificaciones, masa o un área densa asimétrica) o por la existencia de un tumor palpable, o secreción por el pezón y la forma de presentación poco frecuente puede ser la enfermedad de Paget. El diagnóstico histológico y la determinación de la extensión (tamaño) son indispensables para la selección de la terapéutica adecuada.

El carcinoma lobulillar *in situ* es una lesión poco frecuente y su diagnóstico histológico y diferencial con hiperplasia atípica requiere la intervención de patólogos expertos. En general, no se relaciona con masa palpable o cambios mamográficos específicos. Esta lesión se considera como un marcador de riesgo y no como un cáncer que evoluciona hacia la forma invasora de manera directa. Alrededor de 10% a 15% de las pacientes presentará durante su vida un carcinoma invasor en cualquiera de las mamas, generalmente de tipo ductal infiltrante. El riesgo de aparición de cáncer de mama invasor es de aproximadamente

0.5% por año de seguimiento y cuando se vincula con carga genética de primer grado aumenta a 1% por año.

El carcinoma ductal in situ (CDIS) y el carcinoma lobular in situ (CLIS) son lesiones pre malignas no invasivas (estadio 0) pero que pueden llegar a ser invasivas si no se tratan. El cáncer de mama más comúnmente diagnosticado es el carcinoma ductal invasivo (CDI), en el cual las células cancerosas se filtran a través de la membrana basal de los conductos mamarios e infiltran el tejido adiposo de la glándula (estadios IA y IB). Cuando las células cancerosas se rompen a través de un vaso sanguíneo o linfático, es llamado metastásico porque estas células comúnmente invaden los vasos sanguíneos axilares (estadios I-III) y distales, los huesos, los pulmones, el hígado o la médula espinal y cerebro (estadio IV) Figura 4. Aunque el análisis histológico ha sido una herramienta valiosa, no proporciona una estratificación precisa de los pacientes ni de los posibles tratamientos.

ETAPAS DE CÁNCER DE MAMA

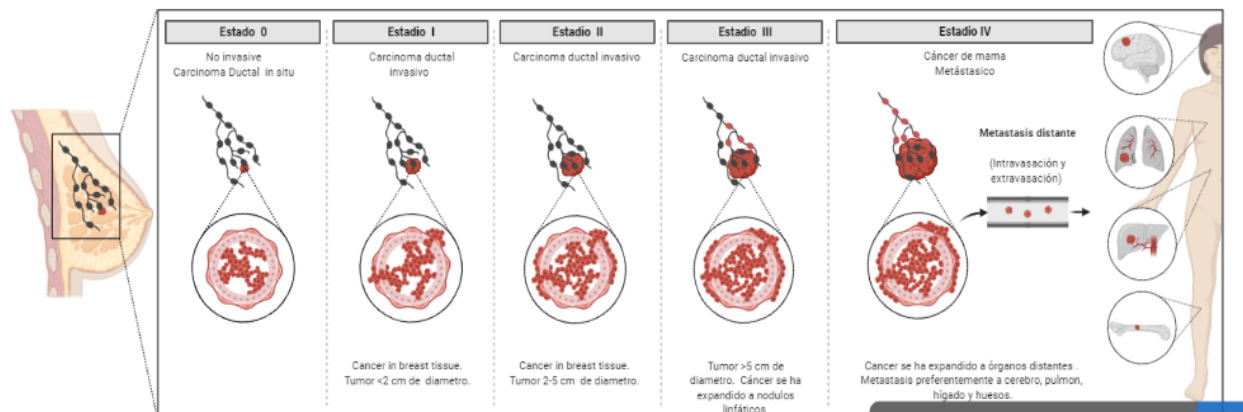


Figura 4. Etapas del cáncer de mama. Biorender 2021.

En la práctica clínica los receptores hormonales de estrógenos (RE), progesterona (RP) y la sobreexpresión de la onco proteína HER-2 y el Ki67 son factores pronósticos y predictivos indispensables en cáncer de mama, por lo que estos marcadores deben realizarse en todas las pacientes con este diagnóstico.

Posteriormente, se generó un sistema de clasificación que se ha basado en análisis inmunohistoquímica identificando subtipos subrogados, como Luminal A (ER+PR+HER2-Ki67-), Luminal B (ER +PR+HER2- o HER2+Ki67+), de tipo normal (ER+PR+HER2-Ki67-), enriquecido con HER2 (ER-PR-HER2+), y de tipo basal (ER-PR-HER2-K5/K6/K14+)⁴.

En la tabla 1 se muestran los subtipos moleculares de cáncer de mama, la aproximación por IHQ no cuenta aún con el nivel de evidencia suficiente para suplir a la clasificación genómica. Sin embargo, se recomienda identificar el fenotipo molecular porque favorece el avance en la identificación de biomarcadores y subgrupos de pacientes en la búsqueda de terapias blanco, para acercarse cada vez más a la práctica clínica ideal de la medicina de precisión⁴.

Tabla 1. Subtipos moleculares por IHQ.

Subtipo	%	Aproximación inmunohistoquímica
Luminal A	40	RE+, RP+, HER-2- Ki67< 14%
Luminal B	20	RE-/+, RP+, HER-2-/ + Ki67> 14%
HER-2+	10-15	RE-, RP-, HER-2+
Basal	15-20	RE-, RP-, HER-2- Expresión Ki67 alta EGFR+ CK5-6 CK14 y 17

EGFR: Receptor del factor de crecimiento epidérmico; RE: Receptor de estrógenos;

RP: Receptor de progesterona; Ki 67: Proliferación; Ck: Citoqueratina

Caracterización molecular del cáncer de mama

Han sido caracterizado en los últimos 15 años por el grupo del Dr. Perou en cuatro subtipos moleculares intrínsecos principales del cáncer de mama: (Luminal A, Luminal B, enriquecido en HER2 y Basal) mediante un algoritmo llamado PAM50^{14, 15} (Fig. 5 y 6).

A nivel de ARN y proteínas, los subtipos Luminal A y B se distinguen en gran medida por la expresión de dos procesos biológicos principales: proliferación celular y las vías lumínicas reguladas por hormonas^{16_17}. En comparación con los tumores Luminal A, los tumores Luminal B tienen una mayor expresión de genes o proteínas relacionados con la proliferación y ciclo celular (MKI67 y AURKA) y la baja expresión de varios genes o proteínas relacionados con el lumen, como RP y FOXA1, pero no con el receptor de estrógeno, que se encuentra similarmente expresado entre los dos subtipos lumínicos y puede ayudar a distinguir la enfermedad Luminal de la no Luminal. A nivel de ADN, los tumores Luminal A muestran un menor número de mutaciones en el genoma, un menor número de copias CNV (menores tasas de amplificación CCND1), menos mutaciones de TP53 (12% vs. 29%), mutaciones GATA3 similares (14% vs. 15%) y más mutaciones PIK3CA (45% vs. 29%) y MAP3K1 (13% vs. 5%) en comparación con los tumores Luminal B¹⁸. Interesantemente, un subgrupo de Luminal B se encuentran hipermetilados, y un subgrupo de Luminal A y B presentan amplificación/sobreexpresión de HER2¹⁹.

Dentro del subtipo Luminal, los Luminales B presentan características biológicas particulares que los diferencian de los Luminales A, tanto en los perfiles de expresión génica, de alteraciones en el número de copias, presencia de mutaciones somáticas, y perfiles de metilación del ADN; asimismo, los tumores Luminales B presentan un mayor grado de inestabilidad genómica que los Luminales A. Los tumores Luminales B se consideran y se tratan en su mayoría como positivos a receptor de estrógenos y tienen un curso clínico como los tumores basales y enriquecidos en HER2 durante los primeros 5 años después de diagnosticada la enfermedad, sin embargo, se ha demostrado que este subtipo tumoral tiene un comportamiento clínico incluso más agresivo que los mencionados, medida como tiempo de supervivencia libre de metástasis^{19_20}.

El subtipo enriquecido con HER2 se caracteriza por la alta expresión de HER2²¹ y proteínas relacionadas con la proliferación (ERBB2/HER2 y GRB7), expresión intermedia de genes y proteínas relacionados con el lumen (RE y RP) y baja expresión de genes relacionados con

la membrana basal y proteínas (queratina 5 y FOXC1). A nivel del ADN, estos tumores muestran el mayor número de mutaciones, y el 72% y el 39% de los tumores enriquecidos con HER2 son TP53 y PIK3CA mutado, respectivamente. Aunque la mayoría 68% de los tumores enriquecidos con HER2 tienen sobreexpresión/amplificación de ERBB2/HER2, debemos esperar identificar el subtipo enriquecido con HER2 dentro de la enfermedad HER2-negativa²². Curiosamente, el subtipo enriquecido con HER2 se ha encontrado especialmente enriquecido para los tumores con alta frecuencia de mutaciones asociadas al APOBEC3B. El APOBEC3B es una subclase de las deaminasas de citidina APOBEC, que convierte la citosina en uracilo y ha sido implicado como una fuente de mutaciones en muchos tipos de cáncer.

El subtipo de tipo Basal se caracteriza por la alta expresión de genes relacionados con la proliferación (MKI67) y queratinas, expresadas típicamente por la capa basal (queratinas 5, 14 y 17), expresión intermedia de genes relacionados con HER2 y expresión muy baja de genes relacionados con la luz del lumen. Los tumores con claudina baja se caracterizan por la baja o nula expresión de los marcadores de diferenciación luminal, el alto enriquecimiento de los marcadores de transición epitelial a mesenquimal, los genes de respuesta inmunitaria y los rasgos de las células madre del cáncer.²³ Desde el punto de vista clínico, la mayoría de los tumores con claudina baja son carcinomas ductales invasivos negativos para el RE, negativos para RP y negativos para HER2 (triple negativo), con una alta frecuencia de diferenciación metaplásica y medular²⁴. A nivel del ADN, estos tumores muestran el segundo número más alto de mutaciones a través del genoma, la mayoría son hipometiladas, y mutaciones en TP53 el 80% y el 9% PIK3CA. El cáncer de mama con mutación BRCA1 está asociado con este tipo tumoral. Finalmente, la sobreexpresión ERBB2/HER2/se encuentra en el 2.1-17.4% de este tipo tumoral. La mayoría de estos tumores tienen un comportamiento clínico agresivo²⁵.

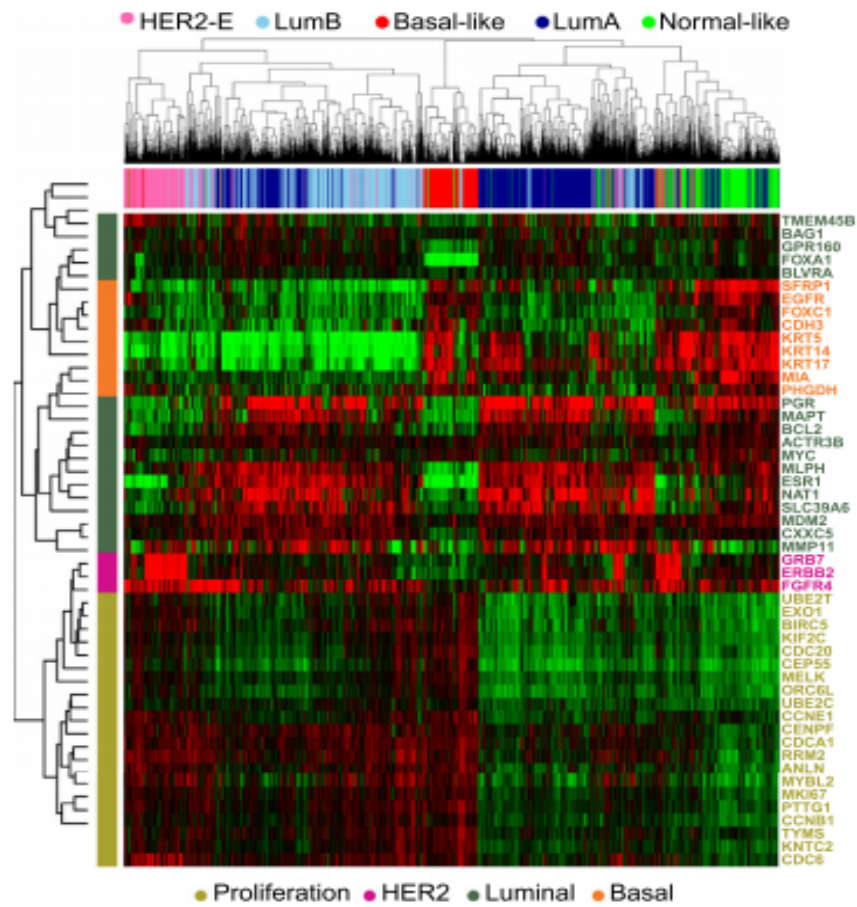


Figura 5. Identificación de subtipos intrínsecos. La clasificación se realizó utilizando el predictor PAM50.

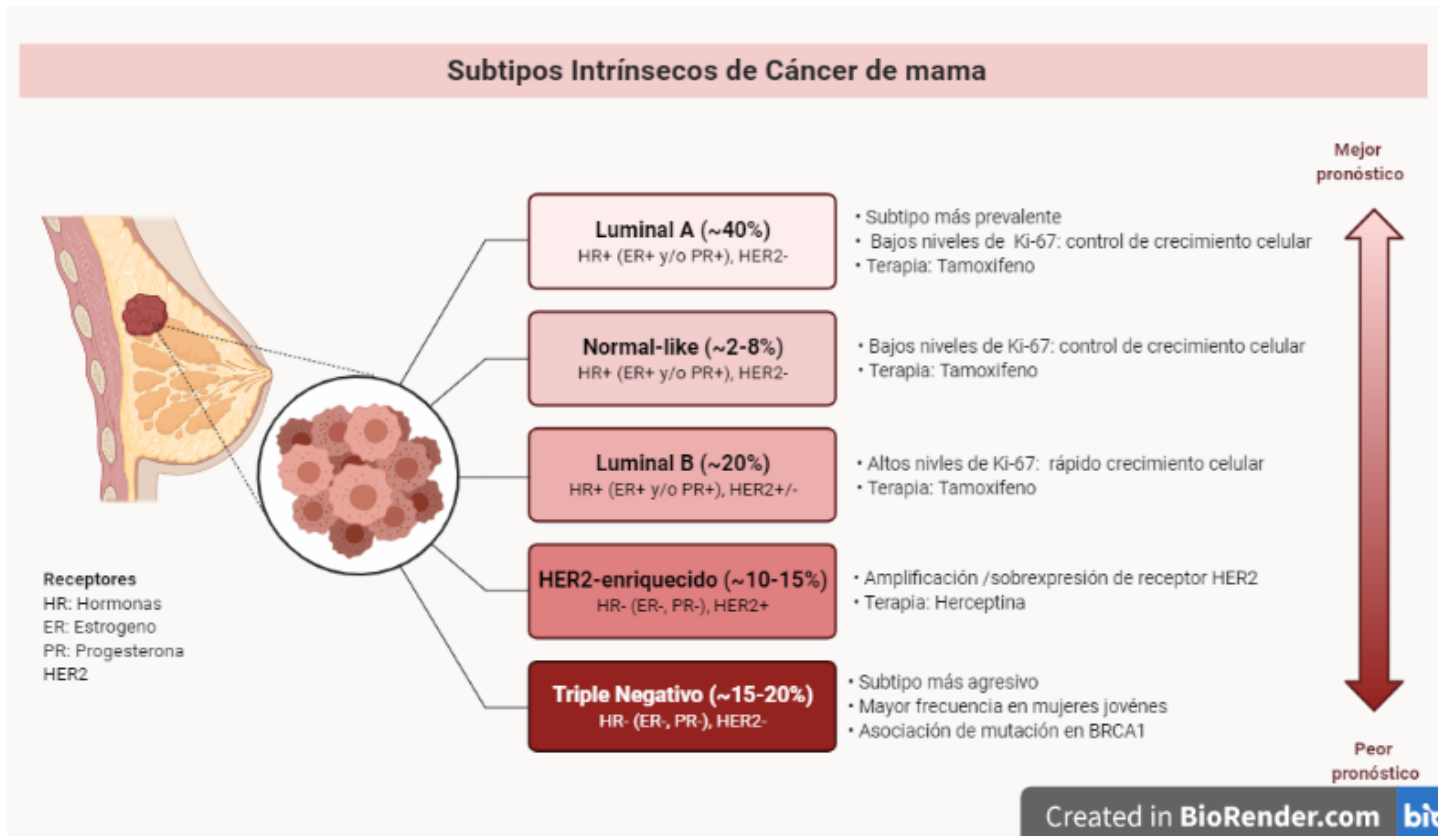


Figura 6. Características clínico-histopatológicas de los subtipos intrínsecos de cáncer de mama. Biorender 2021

Medicina de precisión en Cáncer de mama

La medicina de precisión es una estrategia para prevenir y tratar enfermedades considerando la variabilidad en genes, ambiente y estilo de vida. Recientes avances en tecnologías ómicas se han enfocado en tratamientos más precisos, un ejemplo de ellos son las firmas moleculares.

En los últimos años el cáncer de mama se ha tratado de acuerdo a biomarcadores como el estado de los receptores hormonales RE y RP así como HER2, se consideran los más importantes y se estiman mediante inmunohistoquímica (IHC) en la práctica clínica. Si al menos uno de los dos receptores (RE o RP) se encuentra positivo, la terapia hormonal será suficiente para un tratamiento eficaz, evitando los efectos secundarios mucho más graves de la quimioterapia.

Sin embargo, si el estado de los receptores se estima accidentalmente como falso positivo, el tratamiento hormonal no funcionará y la paciente puede verse privada de la quimioterapia que le salvaría la vida. Debido a su importancia focal, la calidad de la evaluación de los receptores ha sido analizada en estudios multicéntricos, sugiriendo hasta un 20% de errores de clasificación. Se han publicado directrices pero la mejora de la precisión de la evaluación sigue siendo una cuestión clave²⁶.

Recientemente se ha clasificado en subtipos intrínsecos basados en análisis moleculares. Estos subtipos han sido aceptados en la práctica clínica, la terapia endocrina adyuvante, quimioterapia y la terapia anti-HER2 se recomiendan basándose en estos subtipos.

Las firmas moleculares son una herramienta útil para la elección más personalizada del tratamiento sistémico en un grupo seleccionado de pacientes con enfermedad temprana. Esto ha permitido tener mayor certeza de la utilidad del tratamiento indicado, evitando la toxicidad de una terapia sin beneficio. Los efectos adversos de la quimioterapia son significativos por ello, surge la necesidad de predecir el riesgo de recaída y estimar ventajas o desventajas de administrar quimio terapias adyuvantes.

La mayoría de los ensayos de medicina de precisión se centran en las anomalías del ADN, pero sólo unos pocos tumores presentan alteraciones genómicas tratables. Hay una necesidad urgente de explorar objetivos terapéuticos más allá de la identificación de las

aberraciones genómicas conductoras²⁷. En consecuencia, basándose en las variantes oncogénicas somáticas y germinales mencionadas, las plataformas Cancer Genome Interpreter y PreMedKB proporcionaron una lista completa in silico de fármacos de terapia biológica destinada a mejorar la medicina de precisión en el cáncer de mama²⁸.

En particular, se ha estudiado ampliamente el tratamiento optimizado e individualizado basándose en la sub tipificación molecular o las firmas de expresión génica²⁹. Entre las más utilizadas están MammaPrint (Agendia), Oncotype DX (Genomic Health) y PAM50 (Prosigna) (Tabla 2 y Fig. 7)³⁰⁻³¹⁻³².

MammaPrint

MammaPrint es una prueba de genómica para evaluar el riesgo de metástasis de un tumor de mama. Se basa en la firma genómica de 70 genes utilizando tejido fresco para el análisis de microarreglos y tejido en parafina. La prueba es pronóstica más no predictiva. En la actualidad no debe emplearse para la elección del tratamiento médico o loco regional. Se encuentra en evaluación en el estudio MINDACT en comparación con marcadores clínico patológicos para la selección de quimioterapia adyuvante en pacientes^{33, 34}.

Oncotype

Oncotype DX es una prueba que cuantifica la probabilidad de recurrencia de la enfermedad en mujeres con cáncer de mama en estadio temprano, con ganglios negativos, RE+, HER2 normal (significado pronóstico) y evalúa el posible beneficio de cierto tipo de tratamiento, quimioterapia u hormonoterapia (significado predictivo). Oncotype DX analiza un panel de 21 genes del tumor para determinar una puntuación de recurrencia (RS ≤ 18 baja, de 18 a 30 intermedia y ≥ 31 alta)^{35, 36}.

La mayoría de las pacientes con cáncer de mama en México son tratadas a través del sistema público de salud y >80% reciben quimioterapia adyuvante. En 2015, Bargallo, et al³⁷, realizó un estudio prospectivo para caracterizar el impacto del ensayo Oncotype DX en la toma de decisiones sobre la terapia adyuvante y la confianza en esas decisiones entre los médicos del sector público en México, determinaron que el uso del ensayo de 21 genes en el sistema de salud pública mexicano tiene un impacto significativo en las recomendaciones de tratamiento adyuvante que puede reducir el uso global de la quimioterapia.

PAM50/ Prosigna

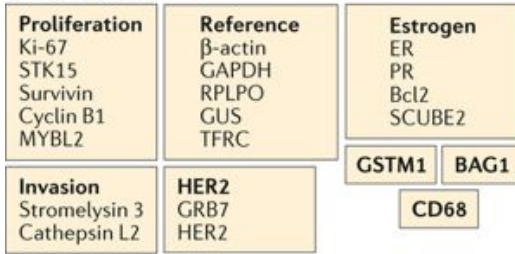
La prueba se basa en el algoritmo clasificador de los 50 genes, llamado PAM50, utiliza RNA extraído de muestras de tejido de tumor de mama fijado en formaldehído y embebido en parafina (FFPE). El algoritmo utiliza un perfil de expresión de 50 genes para clasificar el cáncer de mama dentro de una de las cuatro clases moleculares o subtipos intrínsecos siguientes: Luminal A, Luminal B, HER2 enriquecido o de tipo basal¹⁴. Los perfiles de expresión génica prototípicos (por ejemplo, centroide) de los cuatro subtipos intrínsecos se analizaron muestras de tumor de mama FFPE. El algoritmo computacional basado en la correlación de Pearson compara el perfil de expresión de 50 genes normalizados de la muestra de prueba del paciente con los perfiles de expresión prototípicos de los cuatro subtipos intrínsecos de cáncer de mama¹⁵. A la muestra del paciente se le asigna el subtipo con la correlación de Pearson más alta. El algoritmo indica el riesgo de recurrencia (ROR) en una escala del 0 al 100, que correlaciona con la probabilidad de recurrencia a distancia a los diez años para mujeres posmenopáusicas con cáncer de mama de receptores hormonales positivos en las primeras fases. También indica una categoría de riesgo (bajo, intermedio o alto). El ROR se calcula utilizando coeficientes de un modelo de Cox que incluye la correlación de Pearson con un subconjunto de 46 genes de los 50 genes utilizados para calcular cada centroide de subtipo intrínseco, un índice de proliferación y el tamaño del tumor. Las variables de la prueba se multiplican por los coeficientes correspondientes del modelo de Cox para generar el índice, que se ajusta a una escala de 0 a 100 basada en coeficientes generados a partir del conjunto de formación de muestras de tumor de mama FFPE. También se indican las categorías de riesgo basándose en valores de referencia de ROR determinados en un estudio de validación clínica. (<http://prosigna.com/>)

Tabla 2. Firmas de expresión génica

Prueba	Compañía	Número de genes	Tejido	Tecnología	Medición	Aprobado FDA
Oncotype DX	Genomic Health	16±5ctl	FFPE	qRT-PCR	RS:Low(\leq 18), intermediate (18-31), high (\geq 31) risk	No
MammPrint	Agendia	70	Fresco, FFPE	Microarreglos	Good risk and poor risk	Si
Breast Cancer Index	BioTheranostics	5±2 gene ratio	FFPE	qRT-PCR	Low, intermediate, and high risk	No
Prosigna	Nanostring	50±22ctl	FFPE	Análisis mRNA	ROR: Low (\leq 10), intermediate (10-20), high (\geq 20%) risk	Si
MapQuant	Ipsogen	97	Tejido congelado	Microarreglos	High-grade vs. low-grade tumors. It can resolve these "grade 2" tumors into either "grade 1" or "grade 3" tumors in 80 % of cases.	No
Mammostrat	Applied Genomics	5	FFPE	Inmunohisto química		
EndoPredict	Myriad	12	Tejido congelado	qRT-PCR		
TargetPrint	Agendia					
BluePrint	Agendia	80				

Abreviación: ctl, genes control; FFPE, tejido embebido en parafina; qRT-PCR, Reacción en cadena de la polimerasa; ROR, riesgo de recurrencia; RS marcador de recurrencia.

a Oncotype DX



b Prosigna



c Mammaprint

Cancer cell phenotype/function	Number of genes
Cell-cycle dysregulation	15
Angiogenesis	12
Proliferation and oncogenic transformation	11
Invasion and metastasis	8
Growth factor signal transduction	6
Resistance to apoptosis	2
Miscellaneous/unknown function	16

d Breast Cancer Index

- HOXB13
- IL17BR
- BUB1B
- CENPA
- NEK2
- RACGAP1
- RRM2
- H/I

e EndoPredict

Reference genes:	
CALM2	BIRC5
OAZ1	RBBP8
RPL37A	UBE2C
	IL6ST
	AZGP1
	DHCR7
	MGP
	STC2

Figura 7. Genes que se evalúan en las diferentes firmas de expresión génica a) Oncotype, b) Mamaprint, c) PAM50, d) Breast cáncer index y e) EndoPredict.

La oncología de precisión se centra en la adecuación del tratamiento más eficaz en función del perfil "ómico" de cada individuo o población. Existen algunas plataformas de las que podemos aprovechar la investigación relacionada con cáncer de mama. El CGI (<https://www.cancergenomeinterpreter.org/home>) señala biomarcadores genómicos de respuesta a fármacos con diferentes niveles de relevancia clínica. Huang et al., y el proyecto Pan-Cancer Atlas³⁸ llevaron a cabo la mayor investigación de variantes patogénicas de la línea germinal en el cáncer. Long et al., Cai et al., y Michailidou et al., realizaron estudios de asociación de todo el genoma para identificar variaciones de la línea germinal relacionadas con el desarrollo del cáncer de mama. PreMedKB (<http://www.fudan-pgx.org/premedkb/index.html#/home>) es una herramienta bioinformática que facilita la interpretación del significado clínico de las variantes genéticas de un paciente. PharmGKB (<https://www.pharmgkb.org/>) recopiló directrices completas para la aplicación de la farmacogenómica en la práctica clínica, según varios consorcios de todo el mundo. Por último, el Consorcio PCAWG (<https://dcc.icgc.org/>) reveló un análisis integrador de las alteraciones genómicas en regiones codificantes y no codificantes²⁸.

Los cánceres de mama son fenotípicamente diversos entre las pacientes en términos de tasa de crecimiento agresividad, dependencia hormonal y respuesta al tratamiento. Los actuales biomarcadores moleculares establecidos actualmente son los relacionados con la proliferación celular (Ki67) y el estado de los receptores: receptor de estrógeno (ER), receptor de progesterona (PR) y receptor 2 del factor de crecimiento epidérmico humano (HER2). Los receptores relacionados con el estrógeno (ERR) y los elementos extracelulares como los microRNAs (miRNAs), los exosomas (EXOs) y el ADN tumoral circulante (ctDNA) son biomarcadores moleculares emergentes que se están investigando (Fig. 8). Los biomarcadores genéticos del cáncer de mama incluyen variantes genéticas identificadas en pruebas genéticas por secuenciación, y perfiles de expresión génica, realizados por ensayos de genes. El uso de biomarcadores adecuados para el diagnóstico podría proporcionar información de la patogénesis del cáncer de mama y contribuir a la gestión personalizada del cáncer y a los enfoques terapéuticos y personalizados.³⁹

Aunque la tasa de supervivencia del cáncer de mama a los 5 años del diagnóstico es muy alta, sigue siendo una enfermedad difícil de diagnosticar y tratar, ya que su comportamiento difiere de una persona a otra. La capacidad para detectar el cáncer, definir su estadio y su agresividad se basa en múltiples herramientas de cribado y diagnóstico. Esto representa un reto multidisciplinar para los médicos, científicos, tecnológicos, ya que la comprensión

de la heterogeneidad del tumor y las variaciones fisiológicas entre los pacientes, requiere múltiples fuentes de información para ofrecer métodos más nuevos y precisos o mejorar los existentes. A medida que se van mejorando los métodos y herramientas actuales para detectar el cáncer, las nuevas tecnologías están en camino para proporcionar un aumento de la portabilidad, la conectividad, precio y gestión de datos que tendrán importantes consecuencias para el mantenimiento o la recuperación de la salud de cualquier individuo y paciente (Fig. 9). Por lo tanto, los profesionales de la medicina tienen que estar constantemente actualizados sobre las nuevas prácticas, y la interpretación de sus resultados.

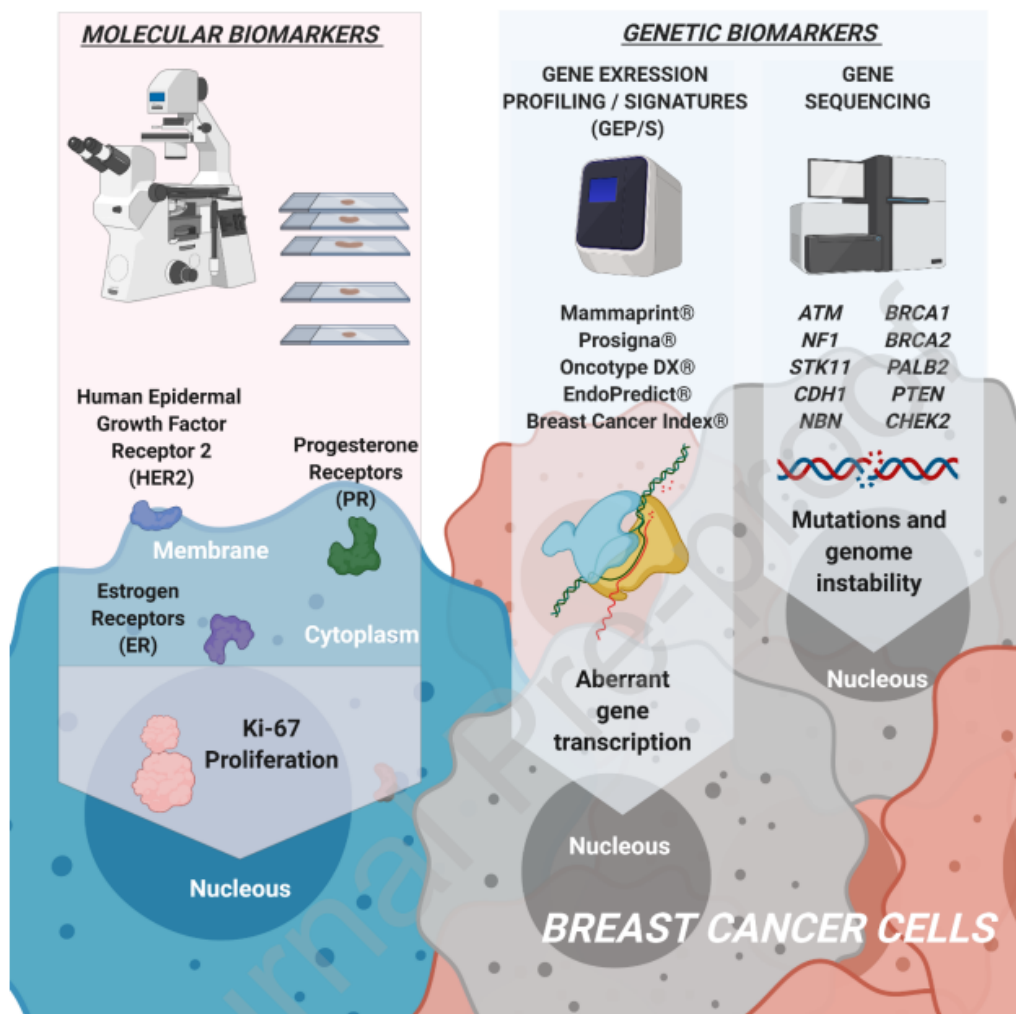


Figura 8. Biomarcadores moleculares y genéticos, localización y herramientas de análisis. Tras detectar el cáncer de mama, los biomarcadores moleculares y genéticos ayudan al médico a elegir la mejor estrategia terapéutica. Los biomarcadores moleculares se identifican mediante IHC, desde

la membrana celular (HER2), el citoplasma (PR y ER) y el núcleo (Ki-67). Estos biomarcadores proporcionan información sobre la respuesta hormonal y el estado proliferativo de las células tumorales. Los biomarcadores genéticos son clave para caracterizar las mutaciones que podrían aumentar la posibilidad de detectar un cáncer de mama agresivo. Los perfiles/firmas de expresión genética ayudan a comprender cómo se comporta un cáncer, lo que permite predecir mejor su agresividad y su respuesta a la terapia³⁹.

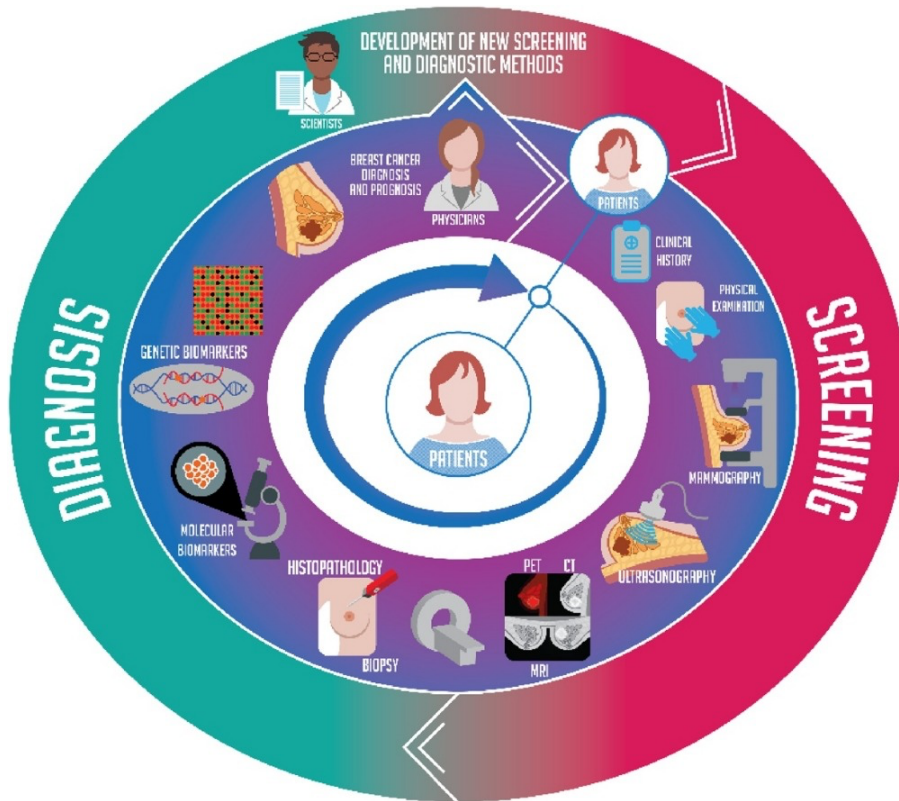


Figura 9. Es necesario un enfoque multidisciplinario para desarrollar herramientas de detección y diagnóstico en cáncer de mama, ofreciendo información clave para los profesionales de la medicina y la investigación.

JUSTIFICACIÓN

En la actualidad existe información muy limitada acerca de la arquitectura genómica de los tumores de mama en población mexicana, desconociéndose la prevalencia de los subtipos moleculares definidos por perfiles de expresión.

Por lo anterior, el establecimiento de las metodologías necesarias para llevar a cabo la clasificación de tumores mamarios en México con base en los perfiles de expresión nos permitirá de inicio determinar la frecuencia de los subtipos de tumores más agresivos en nuestra población y al mismo tiempo abre la posibilidad para que en el futuro se avance hacia el establecimiento de la medicina de precisión aplicada al tratamiento del cáncer de mama en México, además de ofrecernos un panorama más profundo acerca de las alteraciones genómicas involucradas en el desarrollo de esta enfermedad.

HIPOTESIS

El análisis global de los perfiles de expresión génica, permitirá obtener un mayor conocimiento acerca de las diferencias biológicas entre subtipos de tumores de mama, además de proporcionar nuevos datos acerca de los mecanismos moleculares que se alteren en el cáncer de mama y permitirá describir la prevalencia de subtipos de tumores de mama definidos molecularmente en nuestra población.

OBJETIVOS

General

Determinar la frecuencia de subtipos tumorales de cáncer de mama mediante perfiles de expresión génica y la evaluación de un panel de anticuerpos.

Particulares

- Determinar la frecuencia de la expresión de los marcadores: ER, PR, HER2, EGFR, CK 5/6, CK 14, CK 17 y CL1 y CL3 en los tumores analizados.
- Determinar la frecuencia de subtipos de tumores de mama por el PAM50.
- Evaluar la concordancia entre la sub clasificación basada en el PAM50 contra el perfil inmunohistoquímico.
- Analizar genes diferencialmente expresados en cada subtipo tumoral y tejido normal (adyacente al tumor).
- Evaluar la distribución de subtipos intrínsecos en población mexicana-hispana en comparación con otras poblaciones.

Materiales y Métodos

Tipo de Muestra: Se analizaron 105 muestras de tumores malignos de la glándula mamaria en estadios I-IV, provenientes de pacientes sin tratamiento previo y 35 tejidos normales (adyacentes al tumor).

Los criterios de inclusión de los pacientes fueron los siguientes: diagnóstico confirmado de cáncer de mama en estadios I-IV, atendidas en el FUCAM, sin tratamiento previo. El protocolo fue revisado y aprobado por las Comisiones de Investigación y ética tanto del INMEGEN como del FUCAM.

La colecta de tejido se realizó después de haber explicado a cada paciente el objetivo del estudio y los procedimientos involucrados en el mismo a través de una carta de consentimiento informado. Únicamente las pacientes que firmaron la carta de consentimiento informado fueron incluidas en el estudio. Todas las muestras se trabajaron de forma anónima, mediante el uso de códigos de barras, por lo que en ningún momento se entregaron resultados de pruebas específicas a pacientes particulares.

Protocolo para la obtención, almacenamiento y manejo de muestras de tejido

1) Previo a cualquier toma de muestra se obtuvo la carta de consentimiento informado debidamente llenada y firmada en presencia de dos testigos; 2) Las muestras de tejido fueron colectadas sin interferir con los procedimientos normales y sin que se comprometiera la valoración histopatológica o el diagnóstico de la lesión; 3) El patólogo identificó las regiones de tejido tumoral susceptibles de ser analizados; 4) Las muestras de tejido tumoral tenían un peso mínimo de 0.5 gramos; 5) Las muestras de tejido se congelan de forma inmediata en nitrógeno líquido; 6) Una vez congeladas, las muestras se mantienen en un contenedor de nitrógeno líquido o a -80°C hasta su procesamiento; 7) El contenido de células neoplásicas, así como la valoración histopatológica de los tejidos se llevó a cabo en cortes obtenidos con criostato o con cortes obtenidos a partir de tejido incluido en parafina de la misma pieza tumoral; 8) Únicamente las muestras con un contenido superior al 70% de células tumorales y sin presencia de necrosis se procesaron para la extracción de ácidos nucleicos.

Inmunohistoquímica

Las determinaciones de estatus para receptores hormonales y amplificación de *HER2*, se llevó a cabo mediante inmunohistoquímica, utilizando kits comerciales clínicamente aprobados para tal propósito (ER/PR PharmDX kit y HERcep Test kit de DAKO, (Fig. 10). Además de la detección de ER, PR y HER2, se analizó la expresión de un panel de anticuerpos para lograr objetivos. Se utilizó EGRF (Dako), Citokeratina 5/6 (Dako), Citokeratina 14 (Dako), Citokeratina 17 (Dako), Claudina 1 y Claudina 3. El primer panel se utilizó para la identificación de los tumores triples negativos “verdaderos”, es decir, aquellos tumores donde la falta de expresión de ER, PR y HER2 no se debe a algún artefacto técnico, lo cual se demostró mediante la positividad en la detección de alguna otra proteína del panel. El segundo objetivo de este panel es el de comparar los resultados obtenidos mediante clasificación de subtipos tumorales obtenidos con análisis de microarreglos de expresión, con los resultados obtenidos con este panel de anticuerpos, que de acuerdo a la literatura, puede resultar una opción adecuada y de bajo costo para llevar a cabo la subclasificación de los tumores de mama.

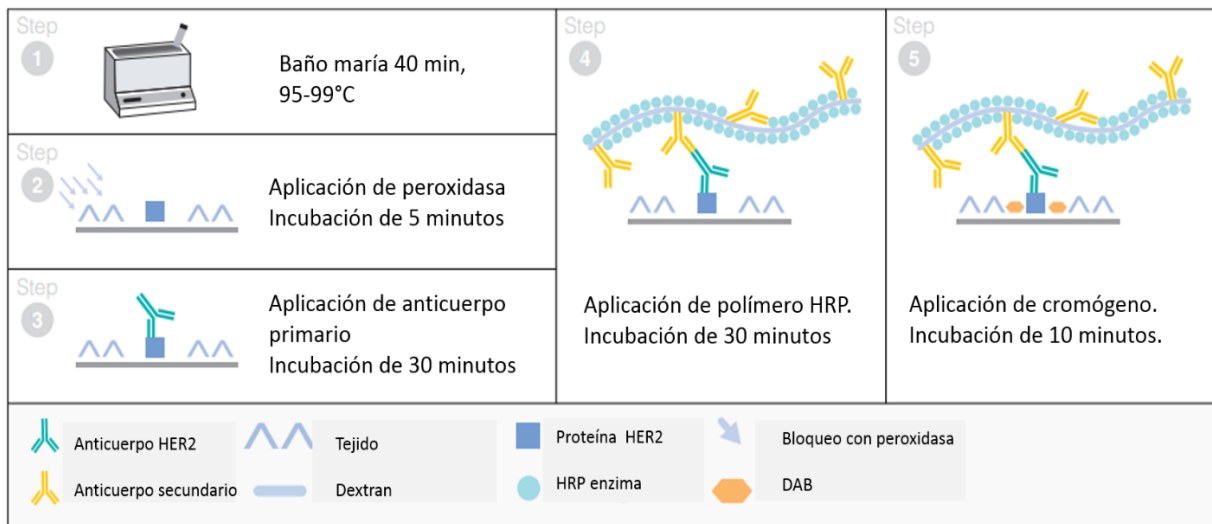


Figura 10. Procedimiento inmunohistoquímica para RE, RP, HER2, Citokeratina 5/6, Citokeratina 14, Citokeratina 17, Claudina 1 y Claudina 3.

Extracción de ácidos nucleicos

La extracción del DNA/RNA se llevó a cabo utilizando un kit comercial DNA/RNA all prep (Qiagen), basado en purificación por columnas. Una vez purificados, la concentración de los ácidos nucleicos fue evaluada mediante espectrofotometría y la integridad mediante electroforesis. La integridad del RNA total se evaluó mediante electroforesis capilar en el sistema Bioanalyzer (Agilent 2100 Bioanalyzer). Únicamente las muestras con RIN superior a 6 se utilizaron para los microarreglos de expresión.

Perfiles de Expresión Génica de mRNA y subtipos PAM50

Se analizaron los patrones de expresión de 35 muestras normales (adyacentes al tumor) y 105 tejidos tumorales con la plataforma de microarreglos de Affymetrix Human Gene 1.0 ST (Fig 11), esta tecnología evalúa 28,869 genes con 764,885 sondas distintas. El microarreglo tiene una cobertura superior al 99% de las secuencias presentes en la base de datos RefSeq.

Se utilizaron de 100 a 300 ng de RNA total obtenido a partir de los tejidos tumorales y tejido normal (adyacente al tumor), se procesaron de acuerdo al protocolo wt sense target manual del fabricante de los microarreglos Affymetrix, Inc. Los microarreglos se leen en un escáner de alta resolución modelo 7G y el análisis de imagen y control de calidad se llevó a cabo utilizando el software Expresión Console de Affymetrix. Posteriormente se utilizó el programa R de Bioconductor (<http://www.r-project.org/>).

Los valores de expresión de mRNA se normalizaron con el algoritmo RMA mediante el método de cuantiles contenido en el paquete de computo "Oligo". Posteriormente se realizó una normalización por lotes utilizando el algoritmo "ComBat", implementado en el paquete SVA. Todos los controles de calidad como diagramas de caja de valores de intensidad, histogramas de distribución de valores de intensidad y gráficos de MA se realizaron con herramientas de los paquetes "Oligo" y "limma". Para facilitar la comparación entre conjuntos de datos, aplicamos un ajuste de escala mediante la estandarización de valores de expresión (z-score) de cada gen. Todos los análisis se realizaron utilizando librerías del proyecto de Bioconductor dentro del ambiente R (<http://www.r-project.org/>).

Para obtener la clasificación de subtipos intrínsecos, se utilizó el algoritmo PAM50 como lo describe Parker et. al¹⁵, definiendo los siguientes subtipos moleculares: Luminal A y B, enriquecido en HER2, de tipo basal, de tipo normal y baja en Claudinas⁴⁰; para cada muestra en nuestra cohorte. El análisis con PAM50 se llevó a cabo utilizando el paquete pbcmc utilizando el parámetro de robustez.

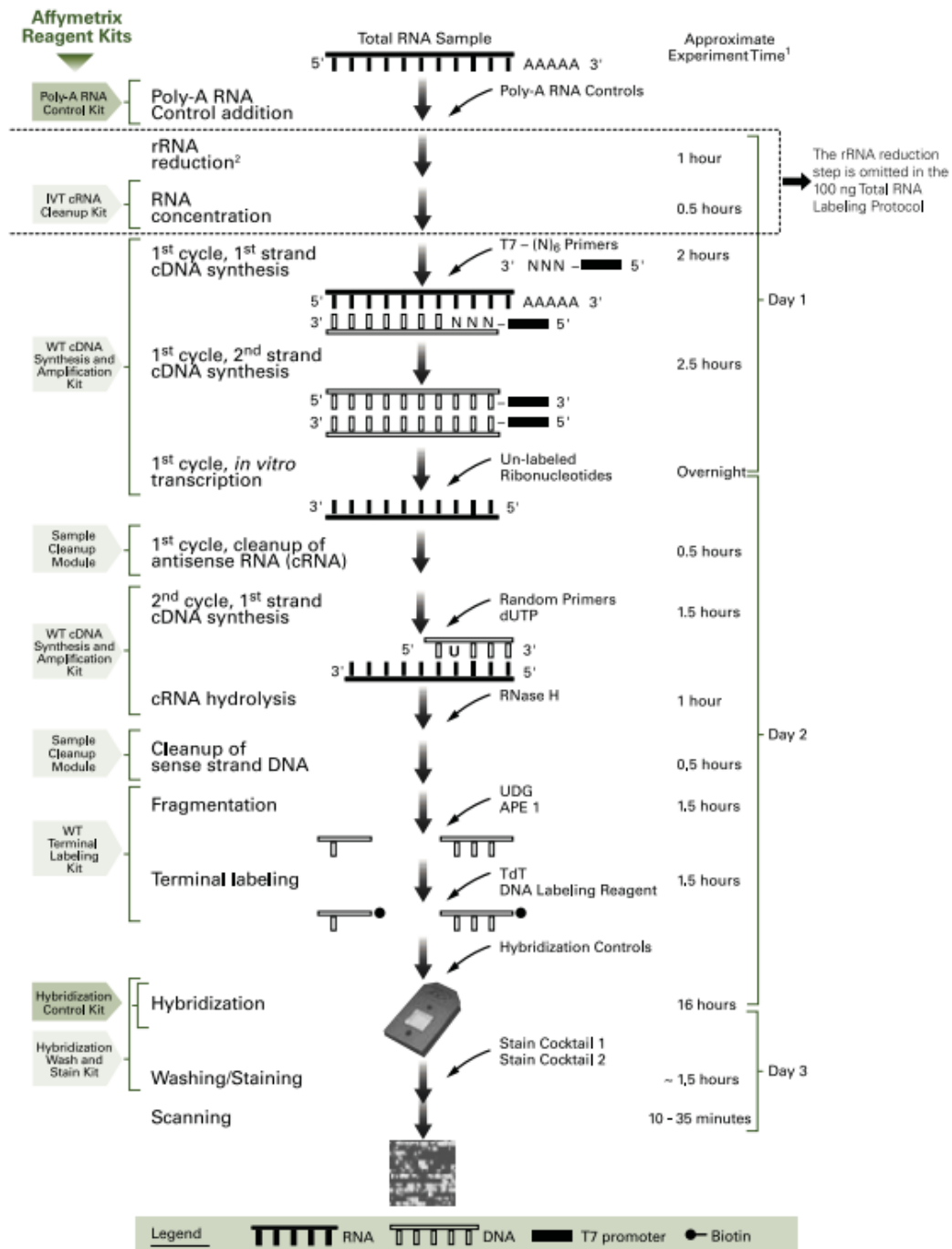


Figura 11. Protocolo para realizar microarreglos Human ST 1.0 Affymetrix

Análisis de genes expresados diferencialmente y enriquecimiento de genes

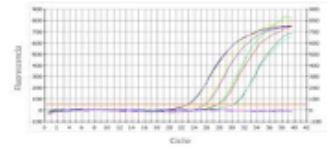
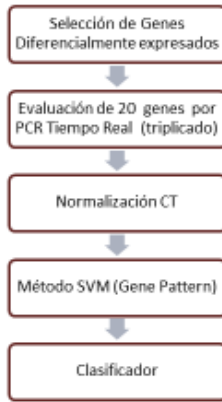
Para definir los perfiles de expresión de mRNA de cada subtipo intrínseco, así como los patrones de expresión diferencial únicos entre muestras normales y tumorales, utilizamos el paquete de estimación de modelos lineales ajustados “limma” de Bioconductor. Los genes expresados diferencialmente se consideraron con log FC: $<0.07 > -0.7$ y valor B mayor que 1. Posteriormente, las listas de genes con expresión diferencial más alta entre los subtipos de la clasificación de PAM50 fueron utilizados para realizar un análisis de enriquecimiento de vías y procesos biológicos, moleculares y patológicos con los algoritmos Enrichr (<http://amp.pharm.mssm.edu/Enrichr/>) y GSEA (<http://software.broadinstitute.org/gsea/index.jsp>). Un Valor p ajustado <0.05 y el número de recuento > 2 se eligieron como criterios de corte, para elegir vías y procesos enriquecidos significativamente.

Validación de genes diferencialmente expresados

Para validar los datos obtenidos se utilizó PCR en tiempo real analizando los genes diferencialmente expresados en las muestras de cáncer de mama (RE, RP, HER2, EGFR, Ki67, FOXA1, FOXM1, SOD1, Citoqueratina 14, GATA3, TPX2, AURKA).

Reacción en cadena de la polimerasa acoplada a retro transcripción (RT-qPCR). El cDNA fue sintetizado usando Superscript III RT-PCR (Invitrogen) siguiendo las recomendaciones del fabricante. 100 ng de RNA total fueron utilizados para sintetizar cDNA en una reacción final de 20 ul. La mezcla de PCR contuvo 1 ul de cDNA, 5 ul 2X Taqman Universal Master Mix, 0.5 ul de sondas Taqman y 3.5 ul de agua grado biología molecular. Los genes GAPDH y B-actina fueron usados como controles endógenos (Fig. 12) Posteriormente se utilizó el método de normalización CT. Se realizó el análisis con el programa de Gene Pattern.

GEN	Subtipo	Presentes en PAM50
PGR	LumA/B	SI
GATA3	LumA/B	NO
ESR1	LumA/B	SI
SOD1	Her2	NO
ERBB2	Her2	SI
CIT14	Basal	SI
ID4	Basal	NO
FOXA1	Basal	SI
EGFR	HER2/TN	SI
TPX2	LUMB	NO
AURKA	LUMB	NO
FOXM1	LUMB	NO
NUEVOS		
GRB7	HER2	SI
UBT2	HER2	SI
CDK1	Basal	NO
CDN1	Luminal A	SI
CYP39A1	Basal	NO
TNFAIP2	Basal	NO
MFAF4	Luminal A	NO
SLOC414	Basal	NO



GenePattern
SVM

Figura 12. Estrategia para realizar el clasificador de perfiles de expresión de cáncer de mama con los genes diferencialmente expresados.

Análisis de Número de copias de DNA

El ADN tumoral y de sangre periférica se procesó utilizando la plataforma de microarreglos Affymetrix Genome Wide Human SNP 6.0 (Affymetrix, Inc.) de acuerdo con los protocolos del fabricante. Brevemente, el ADN se digirió con las enzimas NspI y Styl (New England Biolabs), se ligó a los respectivos adaptadores Affymetrix utilizando la ADN ligasa T4 (New England Biolabs), fue amplificado (Clontech), purificado usando perlas magnéticas (Agencourt), marcado, fragmentado y finalmente hibridado al chip de microarreglos. Después de la hibridación, los microarreglos se lavaron y se tiñeron con estreptavidina-ficoeritrina (Invitrogen). La preparación de microarreglos y su escaneo se realizaron en el laboratorio de genotipificación del INMEGEN. La corrección de fondo y la extracción de valores crudos de intensidad de fluorescencia se realizaron con la consola de genotipado de Affymetrix.

Perfiles de número de copias de ADN

Para generar los perfiles de número de copia, utilizamos el método de "Inferencia del número de copias" implementada en la suite de genepattern: (<http://genepattern.broadinstitute.org/gp/pages/index.jsf>). En resumen, las intensidades de señal a nivel de la sonda de los microarreglos Affymetrix SNP6 (archivos cel) se normalizaron a un brillo uniforme y se combinaron para obtener valores para cada conjunto de sondas usando el algoritmo SNP File Creator. Posteriormente los valores normalizados fueron empleados para el llamado de número de copia mediante el modelo de Inferencia de número de copia. Después de reducir el ruido mediante la sustracción de la variación observada en las muestras normales (sangre periférica), el algoritmo de segmentación CBS identifica las regiones en el genoma que tienen un número de copia subyacente uniforme, creando un archivo de segmentación. Finalmente, utilizando los resultados de la segmentación cnv, las alteraciones somáticas significativas fueron identificadas con el algoritmo GISTIC.

Análisis de ancestría

Datos de SNP derivados de 100 muestras de nuestra cohorte fueron empleados para el análisis de ancestría. Se utilizaron genotipos de 299,411 SNP comunes a tres plataformas de genotipificación (SNP 6.0, Affymetrix 500k e Illumina 1M). Se incluyeron 95 muestras del proyecto HapMap (CEU: ascendencia del norte de Europa; YRI, africanos de Nigeria y CHB +JPT, población del este de Asia) y 37 muestras nativas de México como poblaciones parentales para el análisis, mientras que 161 muestras de mestizos mexicanos del Proyecto de Diversidad del Genoma de México (MGDP) [28, 29] se incluyeron para evaluar proporciones de ancestría en la población general mexicana. Se realizaron cuatro pruebas principales de control de calidad: 1) Estimación de la tasa de ausencia de genotipos por persona, excluyendo individuos con más de 5% de genotipos faltantes, 2) Estimación de la tasa de ausencia de SNP; solo se incluyeron SNP con una tasa de genotipificación del 95%, 3) Prueba de identidad por descendencia (IBD) para evaluar la calidad en el conjunto completo de muestras. Solo muestras de buena calidad y sin evidencia de relaciones familiares entre ellas fueron incluidas. Paralelamente, un análisis de componentes principales (PCA) se realizó para detectar estructuras poblacionales utilizando datos de todo el genoma y el software ADMIXTURE 1.22. Con base en el conocimiento a priori del origen de las muestras, se eligió un valor $K=4$, lo que significa cuatro grupos parentales fueron considerados para cuantificar la contribución ancestral: CEU, YRI, CHB + JPT y NAT MEX y para explicar la subestructura principal en el conjunto de 100 individuos. El componente ancestral se identificó en función de la frecuencia de las diferencias entre las poblaciones de padres relevantes (europea, asiática, africana y nativa). El programa Structure se utilizó para estimar las proporciones de mezcla individuales, así como las proporciones de mezcla promedio en cada muestra. La cohorte se dividió según el componente amerindio en 4 grupos, 1 (0-24% de presencia de componente amerindio), 2 (25-49%), 3 (50-74%) y 4 (75-100%).

RESULTADOS

Datos clínicos

La colecta de muestras se realizó de octubre 2008 a noviembre del 2012 en el FUCAM, se incluyeron tejido tumoral, tejido normal (adyacente al tumor) y sangre periférica de cada paciente. En total se colectaron 253 muestras, las cuales se encuentran en el banco de tejidos del FUCAM-INMEGEN.

Una vez procesadas mediante la técnica histológica, solamente 105 muestras de tejido tumoral cumplieron con todos los controles de calidad necesarios para su procesamiento (presencia de más de 80% de tejido tumoral, calidad y concentración de ácidos nucleicos). Los datos clínicos y patológicos de las muestras se presentan en la tabla 3. El rango de edad de las pacientes fue de 33-92 años, con un promedio de 54 años. En cuanto a los subtipos histológicos, el 9% presentó cáncer ductal *in situ*, 68% carcinoma ductal infiltrante, otros 23%. En relación a los grados histológicos de las muestras, el 7% fue grado I, 47% grado II y el 18% grado III. Las etapas clínicas de las muestras analizadas fueron: etapa I (10%), IIA (40%), IIB (27%), IIIA (8%), IIIB (5%), IIIC (3%).

Evaluación de Inmunohistoquímica

Las 105 muestras de tejido tumoral fueron analizadas mediante inmunohistoquímica para evaluar la expresión de receptores hormonales (receptor de estrógenos RE, receptor de progesterona RP) y para HER2/Neu, en la figura 13 se observan ejemplos de la positividad de las células dependiendo de la tinción y porcentaje de células de tejido tumoral teñidas, de acuerdo a las instrucciones de procesamiento y evaluación de los protocolos aprobados por la FDA.

Se obtuvieron los siguientes resultados: 77 fueron positivas para RE, 65 para RP, 27 HER2, 40 Citoqueratina 5/6, 15 triples negativas. Para las muestra triple negativas se evaluó el anticuerpo EGFR, de las cuales solo 5 fueron positivas, también se evaluaron las Claudinas 1 y 3 para las cuales resultaron todas positivas.

Tabla 3. Datos clínicos de las pacientes con Cáncer de mama, FUCAM

	n=105	%
Edad	Media (rango)	55
30-40	10	9.5
41-50	27	25.71
51-60	35	33.3
61-70	24	22.85
71-80	7	6.66
81-90	1	0.95
91-100	1	0.95
Stages		
I	7	6.66
II	65	61.90
III	18	17.14
IV	0	0
NA	15	14.28
Tumor grade		
I	8	7.61
II	55	52.38
III	17	16.19
NA	25	23.8
Tamaño de Tumor		
≤50mm	81	77.14
≥50mm	8	7.6
Otros	16	15.23
Tipo histológico		
Ductal	71	67.61
Lobular	9	8.57
Mixto ductal/lobular	5	4.76
Otros	5	4.76
NA	15	14.28
ICH		
ER+	70	66.66
RP+	59	56.19
HER2+	19	18.09
CK5/6+	44	41.90
CK14+	64	60.95
CK17+	55	52.38
EGFR+	11	33.3
CLD1+	11	100
CLD3+	11	100

ER positivo requiere al menos 10% de tinción de núcleos.

RP positivo requiere al menos 10% de tinción de núcleos.

HER2 positivo con 3+ IHC o 2+ y CISH confirmado

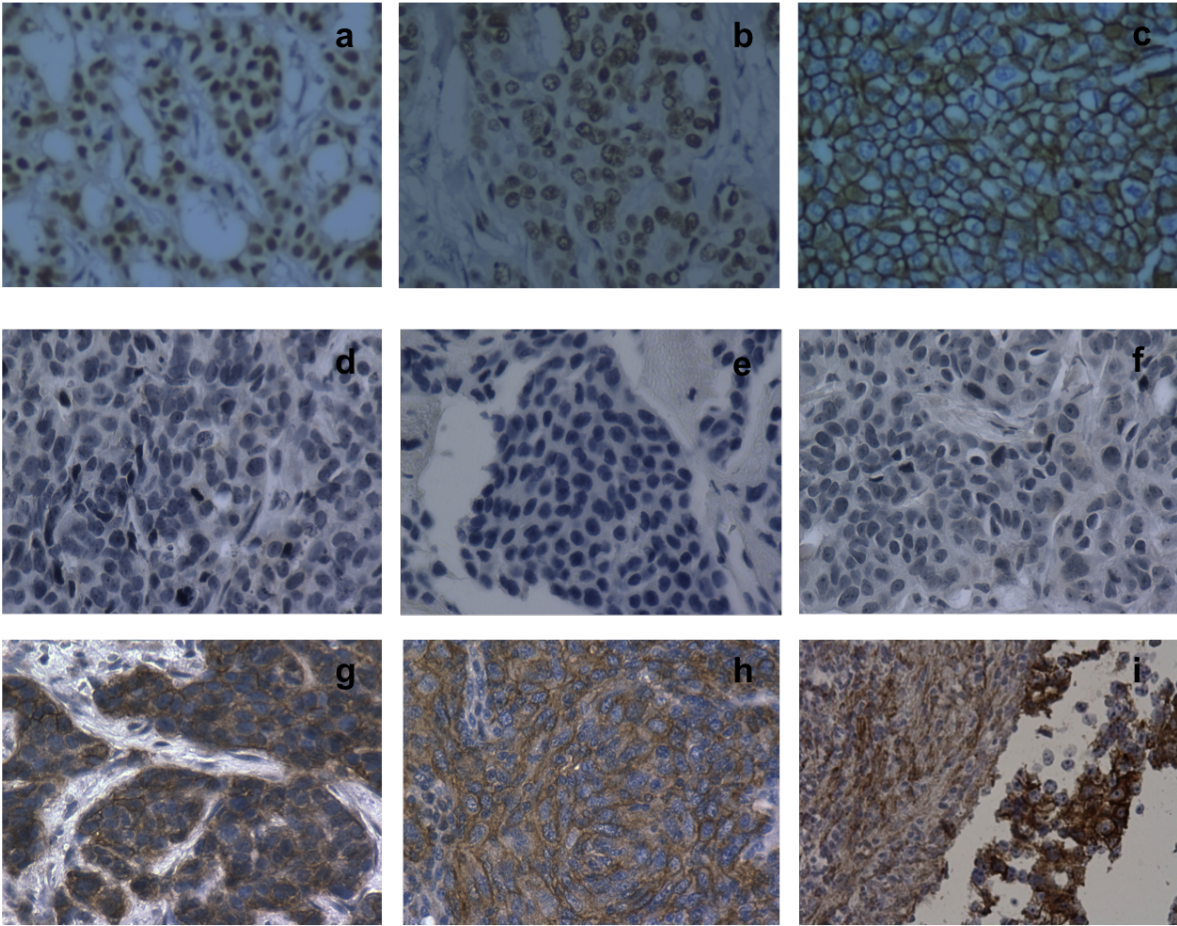


Figura 13. Inmunohistoquímica 40X. a) RE; b) RP; c) HER2; d) Cit 5/6; e) Cit 14; f) Cit 17; g) Claudina 1; h) Claudina 3 y i) EGFR. (Imágenes microscópicas representativas obtenidas con tinciones inmunohistoquímica (40X).

Análisis de Subtipos intrínsecos de cáncer de mama

La clasificación de los tumores se llevó a cabo con base al perfil de expresión intrínseco, utilizando el algoritmo PAM50, descrito por Perou, basado en el agrupamiento de las muestras mediante la evaluación de la expresión de 55 genes cuyo perfil permite identificar los cinco subtipos tumorales intrínsecos (Luminal A, Luminal B, Enriquecido en HER2, Basal y parecido a normal).

Al realizar el análisis con el clasificador PAM50 se obtuvieron los siguientes datos: 12% de las muestras corresponden al subtipo parecido a normal, 44% al Luminal A, 20% al Luminal B, 12% al enriquecido en HER2 y 12% al subtipo basal (Figura 14 y 15). Las muestras que se determinaron como parecidas a normal mediante perfiles de expresión se descartaron de los análisis posteriores, por posible contaminación con tejido normal, como ha sido reportado en la literatura.

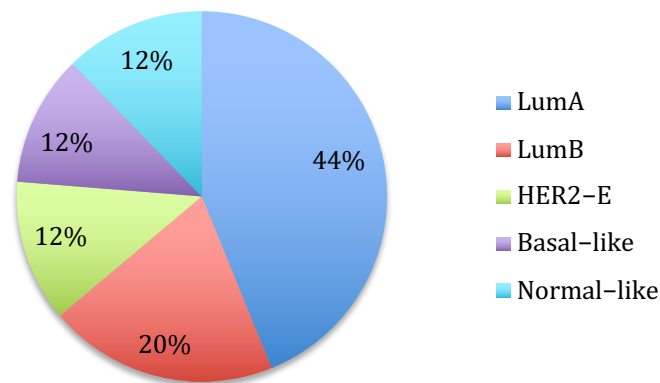


Figura 14. Subtipos intrínsecos obtenidos mediante el algoritmo PAM50.

Análisis de expresión diferencial

Al analizar los perfiles diferenciales de expresión del tejido tumoral (N=105) contrastado contra el tejido adyacente no tumoral (N = 35), se identificaron 121 genes expresados diferencialmente ($FC \geq 2$ y $FDR < 0.05$), involucrados en importantes vías oncogénicas como: adhesión celular, motivos extracelular, secreción de péptidos señal, transportadores ABC, región transmembrana, regulación de la proliferación celular, migración de leucocitos, regulación de la actividad de las MAP quinasas, proteína cromosómica, ciclina, regulación del ciclo celular mitótico, matriz extracelular proteica y respuesta inflamatoria (Fig. 16).

Las vías enriquecidas de los genes expresados diferencialmente recapitulan las características conocidas de cada subtipo, incluida la pérdida de regulación de las vías relacionadas con la señalización del receptor de hormonas en los tumores LumA. Del mismo modo, los tumores LumB también muestran cierta señalización de receptores hormonales y un aumento en la proliferación de rutas asociadas. Considerando que, los tumores enriquecidos con HER2 muestran una alta expresión de HER2 y los genes contenidos en el amplicón 17q12. Los tumores de tipo basal se caracterizan por una baja señalización del receptor hormonal y una alta expresión de la actividad de la vía de proliferación. Estos datos recapitulan diversas asociaciones biológicas entre vías y subtipos publicados previamente (Fig. 17).

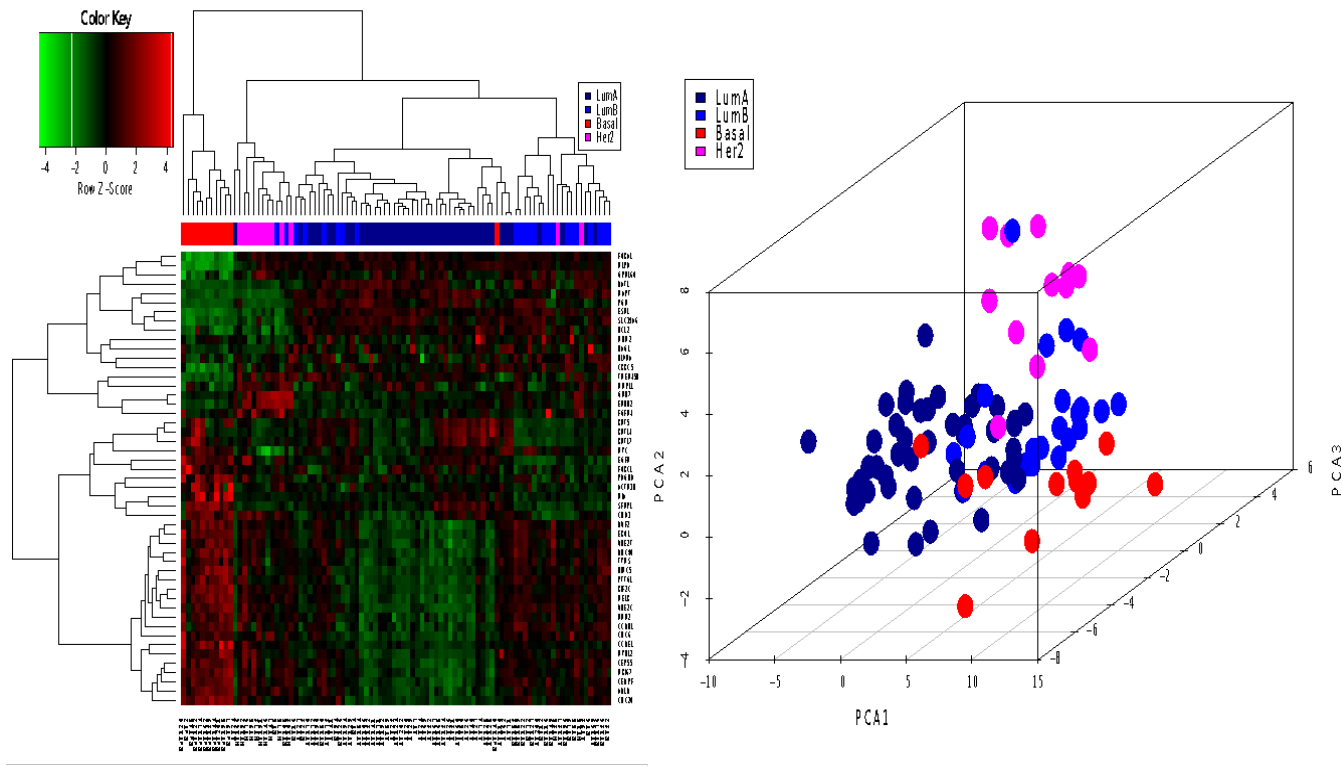


Figura 15. El mapa de calor representa la expresión génica de los 50 genes del predictor PAM50. Colores del mapa de calor corresponden a la expresión de RNAm como se indica en la clave de color: rojo (sobre-expresado) y verde (sub expresado). Luminales A se representan de color azul fuerte, Luminales B, azul claro, Her2, rosas y Basales rojos. Los mismos colores se representan en el PCA, en el que se muestran los diferentes subtipos tumorales cáncer de mama.

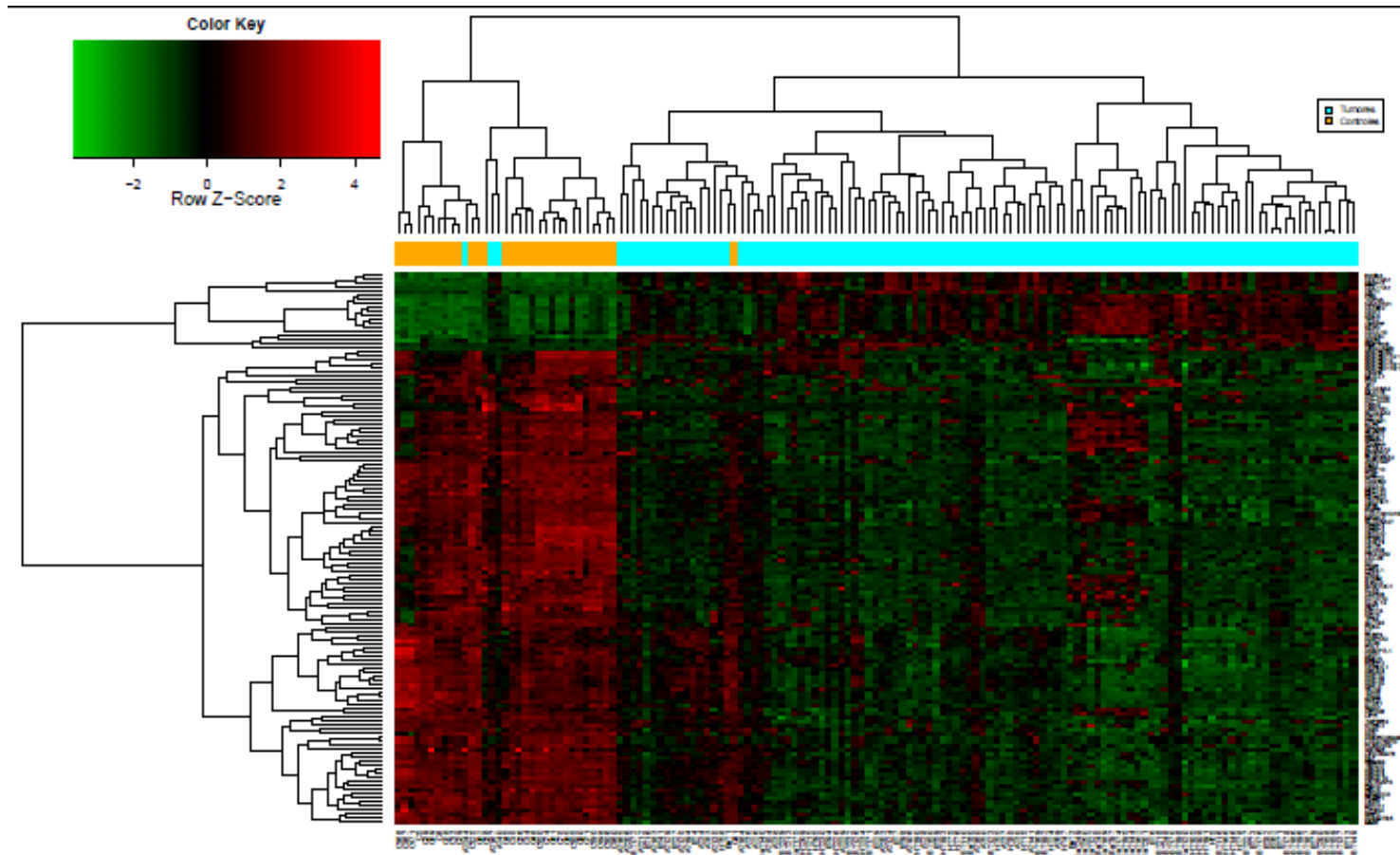


Figura 16. Análisis de agrupamiento jerárquico no supervisado con los 121 genes detectados como diferencialmente expresados entre muestras normales y tumorales. El mapa de calor (correlación de Spearman, distancia euclidiana) Colores del mapa de calor corresponden a la expresión de RNAm como se indica en la clave de color: rojo (sobre-expresado) y verde (sub expresado). Línea naranja: Las muestras de control, línea verde: las muestras tumorales.

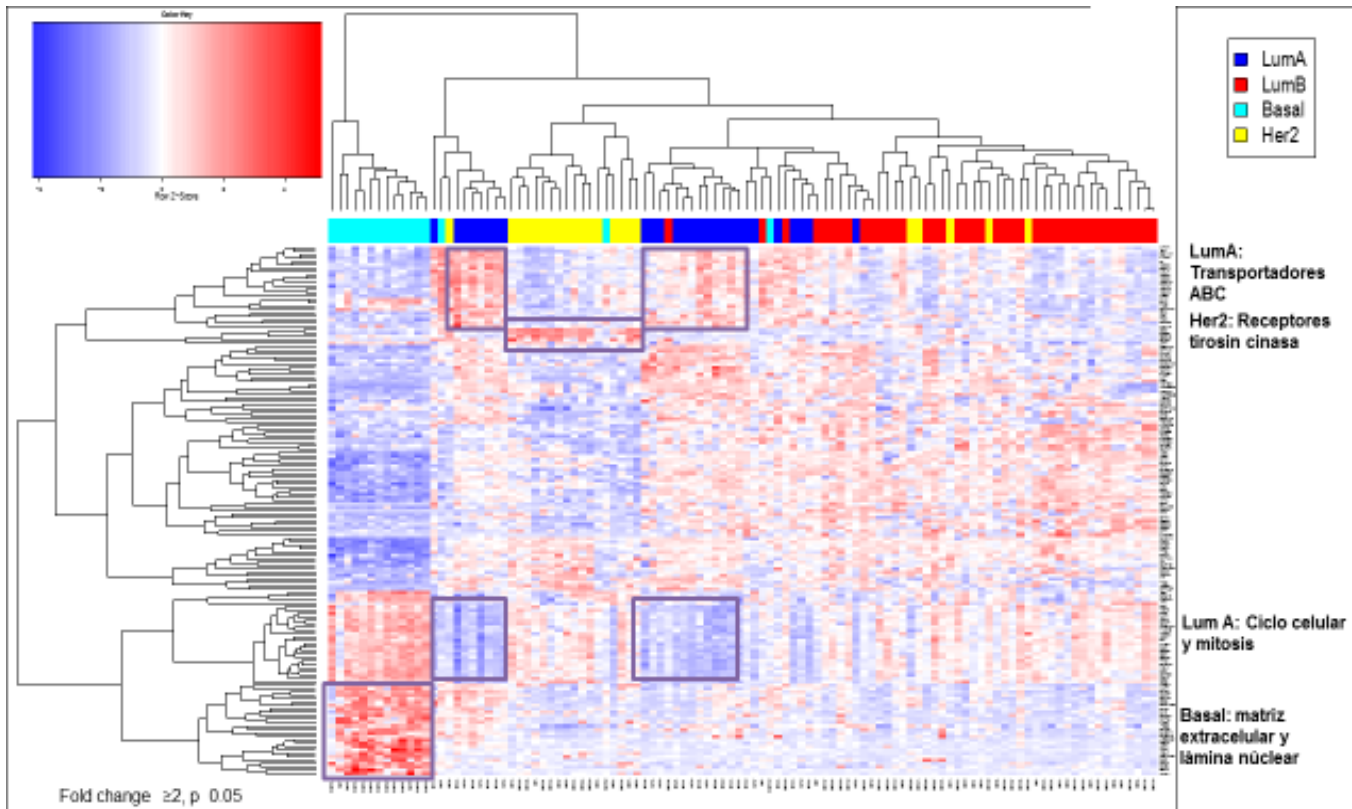


Figura 17. El mapa de calor representa la expresión génica de los genes diferencialmente expresados y representado con las rutas significativas en cada subtipo. Colores del mapa de calor corresponden a la expresión de mRNA como se indica en la clave de color: rojo (sobre-expresado) y azul (sub expresado).

Correlación inmunohistoquímica- subtipos intrínsecos

Con la finalidad de explorar la correlación entre los marcadores inmunohistoquímica y la clasificación obtenida mediante perfiles de expresión PAM50, se analizó la expresión de los receptores de estrógenos, progesterona y HER2 en las muestras de los diferentes subtipos (Figura 18). En los subtipos Lum A se obtuvo un 57% por IHQ vs 72% con PAM50, en Lum B, 17% vs 25%, para Her2 13% vs 64% y por último Basal 13% vs 82%. Los Lum B son los que menos se identifican y para este habría que incluir mas marcadores. Esto nos permitió describir la prevalencia de subtipos de tumores de mama definidos molecularmente en nuestra población y podemos observar que la clasificación con el algoritmo PAM50 es mucho más sensible como se esperaba en comparación con la inmunohistoquímica que se utilizan en la práctica clínica.

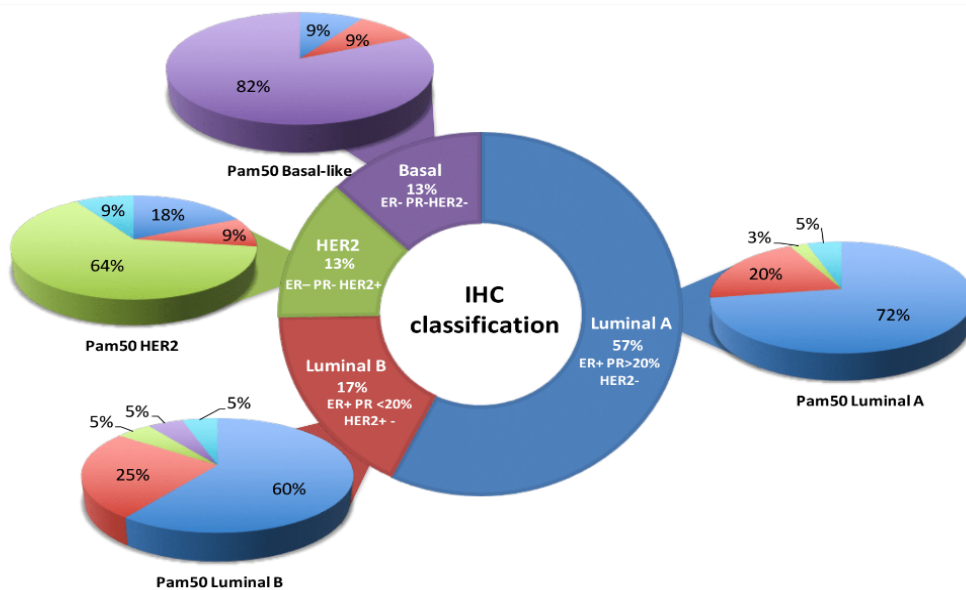


Figura 18. Distribución de subtipos intrínsecos PAM50 y su correlación con inmunohistoquímica.

Lum A en color azul, Lum B color rojo, Her2 verde y Basal morado.

Distribución de subtipos intrínsecos en población mexicana-hispana en comparación con otras poblaciones

Primero realizamos la clasificación de subtipos PAM50 definiendo los grupos intrínsecos con el algoritmo descrito por Perou et. al, utilizando los datos de microarreglos de expresión disponibles. La clasificación con PAM50 nos permitió identificar cuatro subtipos de tumores diferentes: LumA, LumB, Her2 y Basal. Mientras que las muestras clasificadas como parecido a normal se descartaron debido a la posible contaminación por tejido no tumoral, como se ha señalado anteriormente. Cabe destacar que, de acuerdo a esta clasificación, nuestra cohorte de pacientes hispano-mexicanos (N = 105) se subdivide en 46% de tumores LumA, 28% de LumB, 15% de HER2 enriquecido y 12% de tipo basal (Figura 19 A). Al comparar contra una cohorte HM más pequeña (GSE75678 N = 53), la frecuencia de los tumores de tipo basal parece estar enriquecida en comparación con nuestra cohorte (21% frente a 12% en nuestra cohorte), mientras que se observa una menor frecuencia de tumores LumA en este estudio (36% vs 46% en nuestra cohorte). Esto puede ser debido a la cantidad limitada de tumores evaluados en el estudio GSE75678⁴¹.

El análisis de la distribución de los subtipos intrínsecos en los tumores de mama entre todas las poblaciones humanas analizadas mostró que LumA era el subtipo más predominante en los diferentes grupos étnicos, excepto en la población de Afroamericana (AA) donde el subtipo de tipo basal mostró un claro enriquecimiento (30.4%), contrastando con la población asiática, que presentan la proporción más baja de este subtipo (15%). Respecto a los otros subtipos, se encontró una superposición sustancial en las frecuencias entre la población hispana y no hispana (Figura 19 B). De acuerdo con lo que se ha informado en otros estudios, los pacientes AA tuvieron un enriquecimiento significativo del subtipo basal en comparación con los pacientes con Hispano mexicana (HM) (OR: 2.988, IC 95%: 1.493-5.979). Las diferencias entre la población HM y el resto de las poblaciones en la proporción de los otros subtipos moleculares no alcanzaron significancia estadística. La edad en el momento del diagnóstico se correlacionó significativamente con el origen étnico, por ejemplo, las mujeres asiáticas y caucásicas tiene probabilidades menores de presentar un cáncer de mama en edades más jóvenes (<45 años) que las mujeres de la población hispano-mexicana (OR asiática: 0.115, IC 95%: 0.055 -0.241, caucásico OR: 0.547, IC 95%: 0.318-0.939); mientras que las mujeres Afroamericanas no presentaron ninguna diferencia significativa (Figura 19 C). En los modelos logísticos ajustados por edad, las pacientes afroamericanas con cáncer de mama tenían probabilidades significativamente más altas de

presentar un tumor de tipo basal que los pacientes con HM (OR: 2,43, IC 95%: 1.063-5.564.39), por otro lado, las mujeres AA mostraron menores probabilidades de presentar tumores LumA que las mujeres HM (OR: 0.328, IC 95%: 0.158-0.68). Además, entre las pacientes asiáticas se observó una reducción significativa del riesgo relativo de presentar tumores LumB con respecto a las pacientes HM (OR: 0.147, IC 95%: 0.033-0.652). Un resultado de interés, es el hecho de que no detectamos ninguna diferencia significativa en comparación con otras poblaciones hispanas conformadas principalmente por mujeres peruanas, brasileñas y puertorriqueñas.

Ancestría en subtipos intrínsecos en población mexicana

La historia de América Latina refleja una heterogeneidad de componentes de ancestría que son el resultado de la mezcla de poblaciones europeas, precolombinas y africanas; particularmente México tiene una importante diversidad genética. Como se espera, la inferencia de ancestría en pacientes mexicanas arroja predominantemente los componentes nativo-americanos y europeos. Mediante un enfoque genómico de integración investigamos las posibles asociaciones entre la estructura ancestral en la población mexicana y la distribución de subtipos intrínsecos. Dividimos nuestra cohorte según el porcentaje de componente amerindio de la siguiente manera: componente 1 (75-100%), componente 2 (50-74%), componente 3 (25-49%) y componente 4 (0-24%).

El subtipo LumA es el único que tiene pacientes que pertenecen al componente 4 (5%), mientras que el LumB muestra una mayor proporción de componente 2 (41%) y 3 (35%). El subtipo HER2 se encuentra enriquecido en el componente 2 (73%), mientras que el subtipo basal se enriqueció en el componente 2 (50%) y el componente 1 (25%) (Figura 20 A). Aunque no fue posible encontrar una asociación significativa entre los subtipos PAM50 y el componente amerindio, observamos una tendencia que correlaciona los subtipos más agresivos: Her2 y Basal, con la presencia de un componente amerindio superior y una ascendencia europea inferior (Fig. 20 B-C).

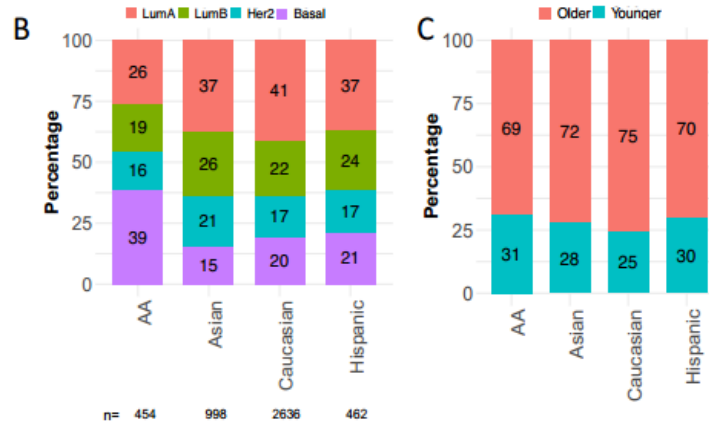
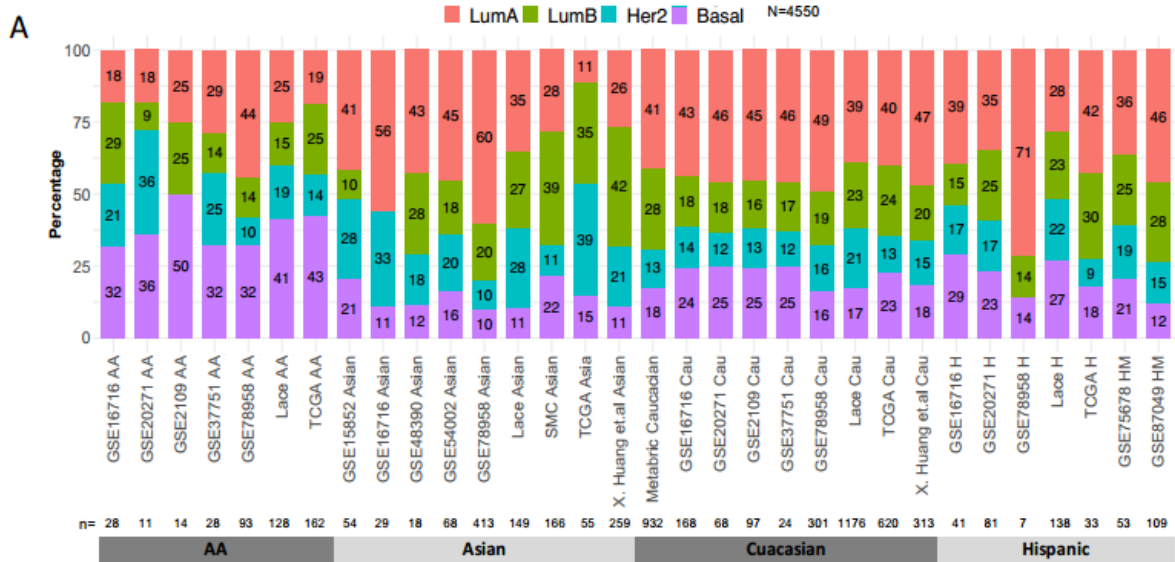


Figura 19. A. Distribución de los subtipos PAM50 en las poblaciones humanas: Afroamericana (AA) Asiática, Caucásica e Hispana (Hispana-Mexicana). **B.** Frecuencias de subtipos entre la población y **C.** Correlación entre edad y grupo étnico.

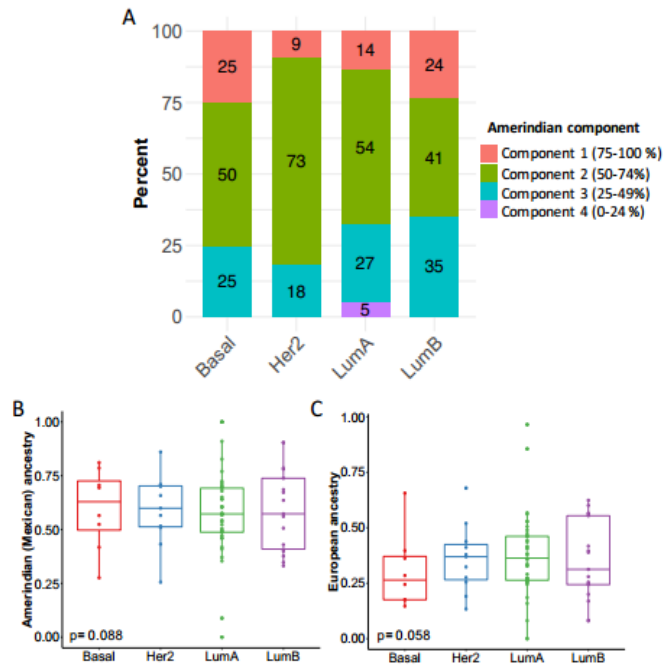


Figura 20. A Asociación entre la frecuencia del componente amerindio y europeo en subtipos intrínsecos PAM50 en población mexicana; **B.** Ancestría amerindio y **C.** Ancestría europea.

Análisis de alteraciones de número de copias entre subtipos

En cáncer, las alteraciones somáticas en el número de copias (SCNAs) afectan una fracción del genoma mucho mayor de lo que afectan otro tipo de alteraciones somáticas. Las SCNAs juegan un rol crítico en la activación de oncogenes y la inactivación de genes supresores de tumores. Con la intención de identificar y entender las SCNAs y sus consecuencias fenotípicas en nuestra cohorte de pacientes mexicanas analizamos un total de 100 tumores de mama pareados con muestra de sangre con la plataforma SNP 6.0 de Affymetrix, a continuación se realizó la segmentación de los datos con el programa CBS y se definieron las regiones significativamente alteradas en nuestro grupo de tumores con el algoritmo GISTIC (Fig. 21).

Como resultado, se identificó un perfil de alteraciones de SCNAs en un total de 165 regiones genómicas diferentes, afectando un total de 148 genes. Dentro de las alteraciones más significativas, encontramos amplificaciones en las regiones 1p, 1q, 8p, 17q (la región que alberga el amplicón de HER2) y 20p, que contienen genes como TRSP1, ACTN2, ERBB2, GRB7 y IKZF3; y deleciones significativas en las regiones 4q, 6q, 8p, 8q, 11p y 18q, conteniendo, entre otros, a los genes GRIK2, KRTAP9, ADAM32, ADAM18 y ZEB2. Es importante señalar, que los 20 genes con amplificaciones y 20 con deleciones en regiones más significativas muestran una frecuencia prácticamente dos veces mayor en la cohorte hispano-mexicana (HM) que en el resto de las poblaciones (Figura 22).

Además, se exploró el perfil de SCNA estratificando por subtipo de PAM50, aunque la cantidad de tumores en cada subtipo en nuestra cohorte es limitado, se encontraron datos consistentes con estudios previos. En particular, se observó que los tumores LumA se caracterizan por amplificaciones en los cromosomas 11q, 8p, 20q, 8q y 12q y deleciones en 17p, 6q y 8p. Por su parte, el subtipo LumB tiene una mayor incidencia de SCNA, mostrando amplificaciones en el cromosoma 11q, 17q, 8p, 8q, 1q y 14q y deleciones en 17q, 13q, 19p, 11q, 16q y 6q.

En el subtipo enriquecido con HER2, la amplificación en el HER2 fue evidente, junto con la amplificación en el cromosoma 6q y la eliminación en 14q. Finalmente, en el subtipo de tipo basal, uno de los tumores más agresivos, pero también uno de los menos comunes, no se detectaron amplificaciones probablemente por un sesgo atribuido al número de tumores

evaluados. Sin embargo, se definió eliminaciones en el cromosoma 10q, 4p, 5q y 8p. El contenido de genes en algunas de estas regiones son moléculas importantes para el establecimiento de cada subtipo, así como para las vías carcinógenas (Tabla 4).

Tabla 4. Alteraciones de número de copias en subtipos intrínsecos de cáncer de mama.

Subtipos	Amplificaciones	Deleciones
Luminal A	8q24.21, 8p11.23, 11q13.3, 20q13.2, 12q15	17p12, 6q27, 8p11.21
Luminal B	11q13.3, 17q21.32, 8p11.23, 8q21.13, 8p11.21, 1q21.3, 14q21.1	17p13.3, 13q14.2, 19p13.3, 11q25, 16q23.1, 6q21
HER2	17q12, 6q21	14q23.3
Basal	-----	10q26.3, 4p14, 5q11.2, 8p23.2, 10q23.31, 13q14.2, 2q37.3

Con el fin de profundizar sobre las posibles consecuencias biológicas de los cambios en el número de copias, los correlacionamos con la expresión de los genes contenidos en cada región alterada a lo largo de los subtipos PAM50. Interesantemente, la ganancia de 8p11.23 se correlaciona positivamente con la sobreexpresión de ZNF703 en el tumor Luminal A, y la pérdida de 5q12.1 con la reducción de la expresión de DEPDC1B. En los tumores Luminales B, identificamos la regulación a la baja de PROM1, junto con la deleción de la región 4p15.3 (Fig. 23). Además, observamos que en el subtipo HER2 existe sobreexpresión de GRB7, STARD3, PPP1R1B, PNMT, ABCC3 y PGAP3, que se corresponden con la amplificación del amplicón Her2 17q12 (Fig. 24). Por último, los tumores basales presentan una expresión reducida de TMEM30B (14q23.1) y SNED1 (2q37.3) en comparación con otros subtipos, de acuerdo con la eliminación de los cuales se han informado en otros artículos (Fig. 25).

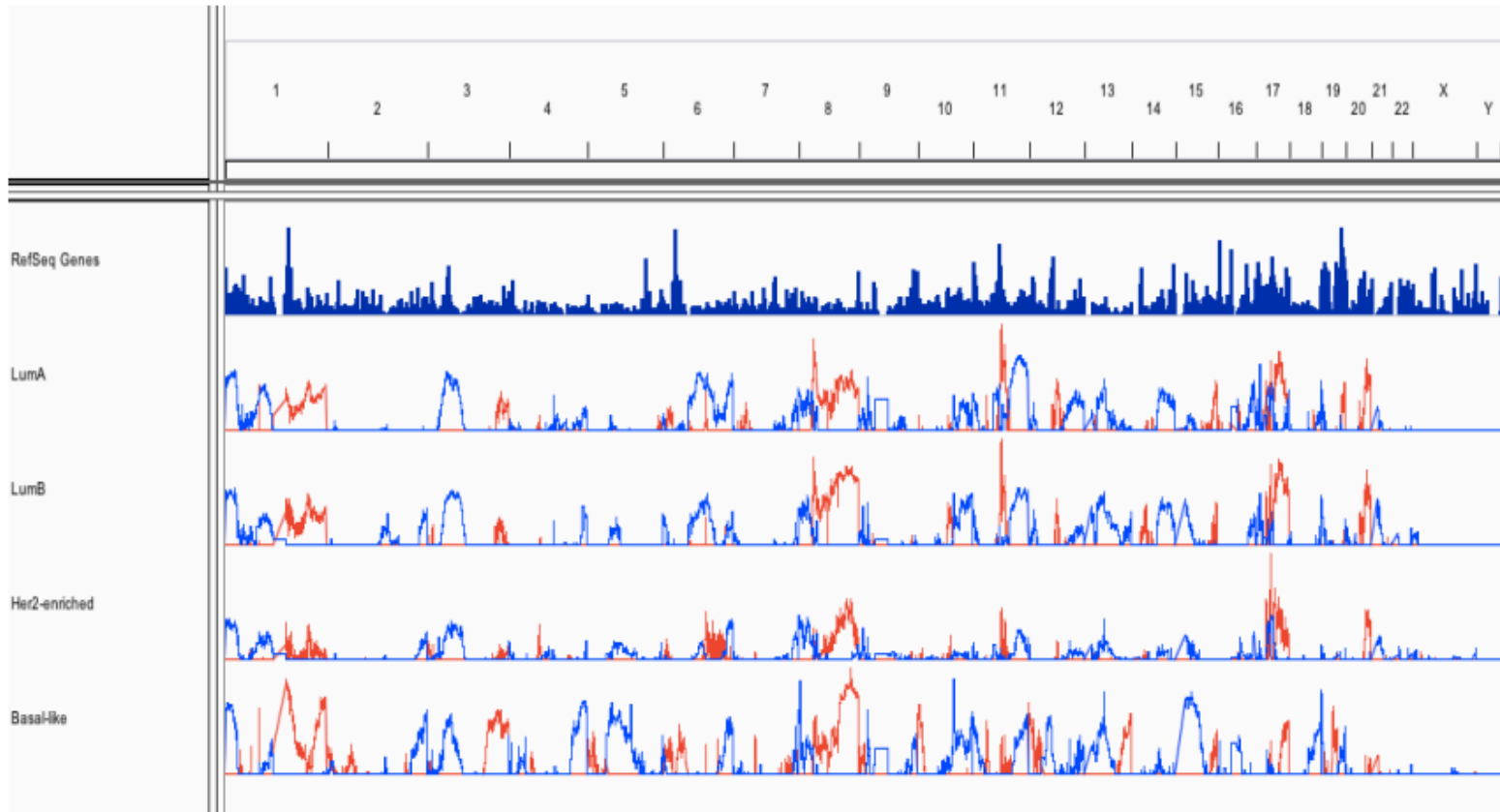


Figura 21. Alteraciones somáticas en el número de copias a través de los subtipos intrínsecos Pam50 en población hispano mexicana. Amplificaciones en rojo, deleciones en azul. El tamaño del pico representa la frecuencia de la amplitud de las aberraciones.

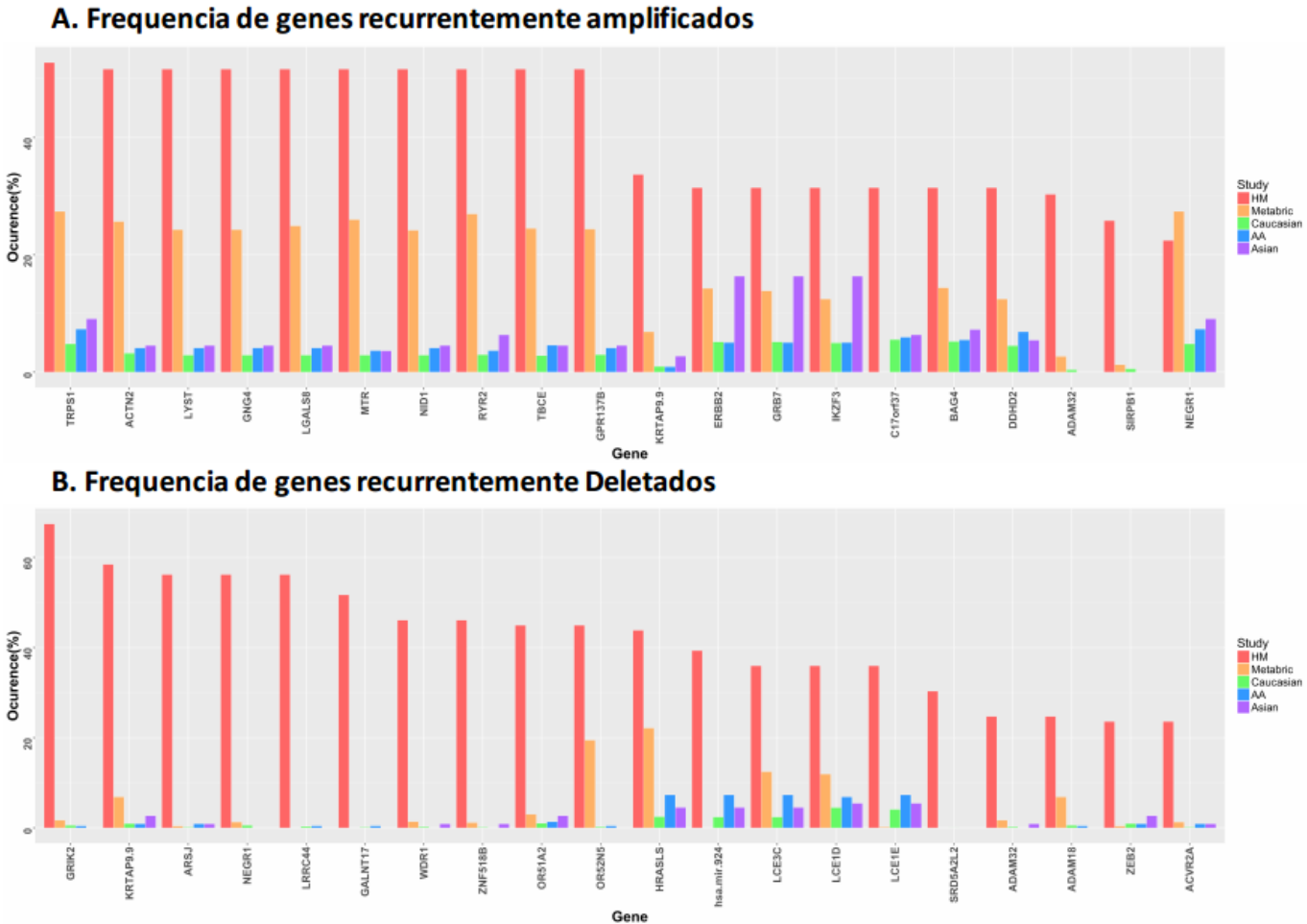


Figura 22. Frecuencia de genes recurrentemente y significativamente amplificados o deletados en población HM y su comparación con las frecuencias presentadas en otras poblaciones.

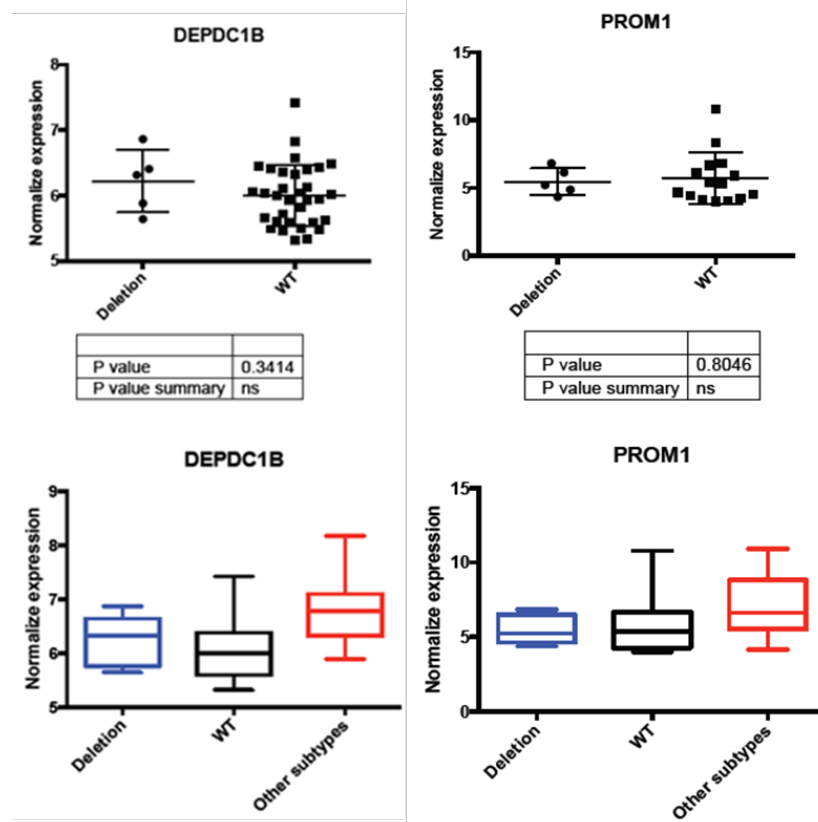


Fig. 23 Correlación número de copias y expresión de cada subtipo intrínseco de cáncer de mama. Luminal A (DEPDC1B), Luminal B (PROM1), en el set de datos (n=105) separados por número de copias. P valor determinado por prueba de student.

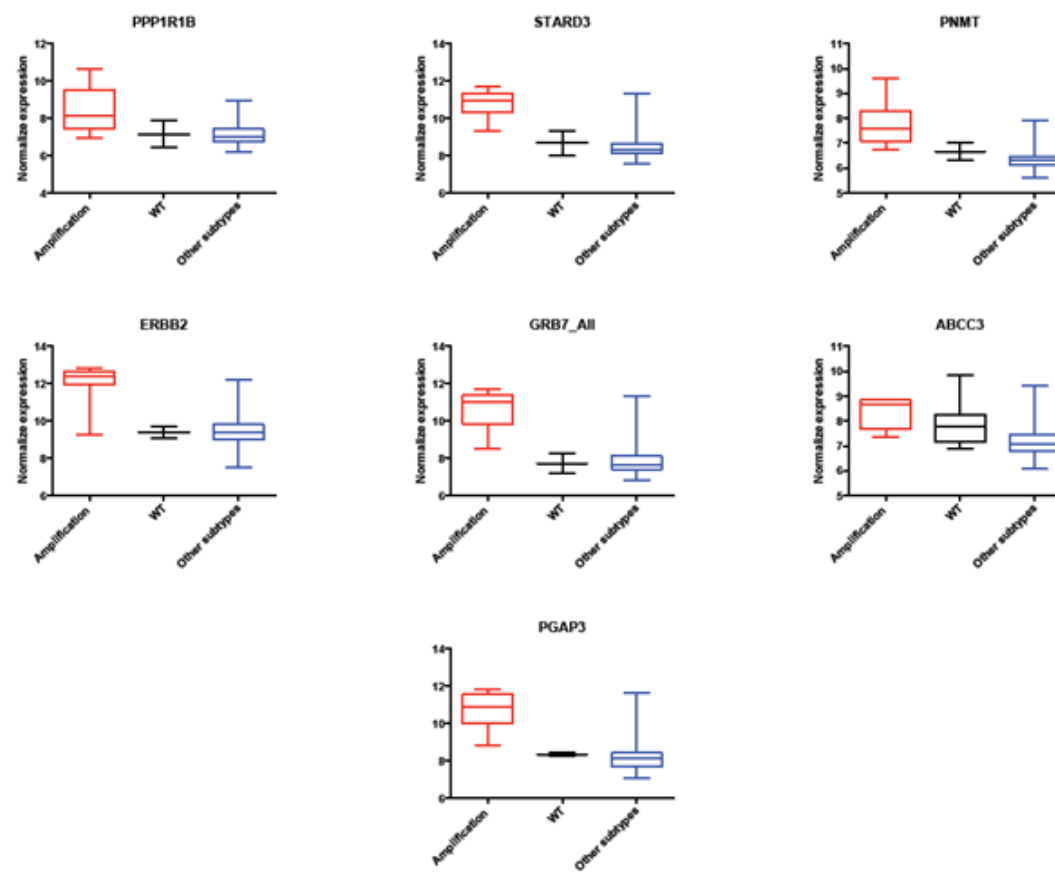


Fig. 24 Correlación número de copias y expresión de cada subtipo intrínseco de cáncer de mama. HER2-enriquecido (GRB7, STARD3, PPP1R1B, PNMT, ABCC3 y PGAP3PP) en el set de datos (n=105) separados por número de copias. P valor determinado por prueba de student.

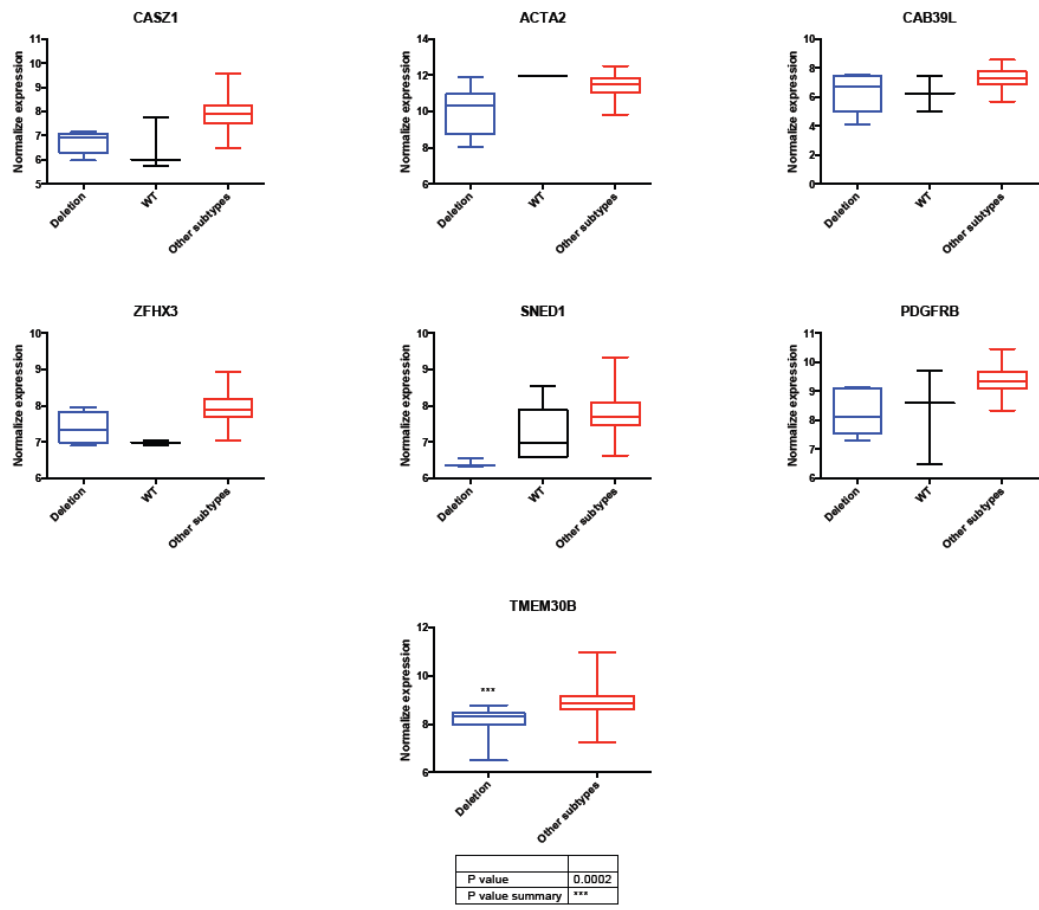


Fig. 25 Correlación número de copias y expresión de cada subtipo intrínseco de cáncer de mama. Basal-like (CASZ1, ACAT2, CAB39L, ZFH3, SNED1, PDGFRB, TMEM30B) en el set de datos (n=105) separados por número de copias. P valor determinado por prueba de student.

DISCUSIÓN

La publicación de datos sobre los subtipos moleculares de cáncer de mama por Perou et al¹⁴, hace más de una década revolucionó su comprensión, debido a la definición de distintas características clínicas, pronósticos y respuesta a terapia⁴². Sin embargo, los altos costos y la escasa disponibilidad tecnológica de los estudios de análisis molecular para determinar estos subtipos han llevado a la necesidad de clasificar los tumores de mama mediante inmunohistoquímica, en sustitución de la clasificación molecular⁴³.

Por ello, el objetivo de este estudio fue caracterizar la frecuencia de subtipos tumorales de cáncer de mama mediante perfiles de expresión génica e inmunohistoquímica. Este trabajo describe la prevalencia de subtipos intrínsecos en una cohorte mexicana basada en un ensayo de expresión génica, lo que se suma a los diversos esfuerzos de descripción de los subtipos de cáncer de mama en múltiples poblaciones.

Estos resultados confirman la existencia de subtipos intrínsecos de cáncer de mama en pacientes de México, como lo describe la literatura en otras poblaciones caracterizadas como Europa⁴⁴, América del Norte (EE. UU.)⁴⁵⁻⁴⁶, África⁴⁷ y Asia. La frecuencia estimada, excluyendo los tumores de tipo normal, demuestra que los tumores LumA corresponden al 46% de todos los casos de cáncer de mama, mientras que LumB, Her2 y Basal corresponden al 28%, 15% y 12%, respectivamente, en nuestro set de muestras.

Las diferencias en la prevalencia de subtipos moleculares de tumores de mama en poblaciones de distintos orígenes étnicos han sido ya reconocidas previamente. En el análisis, los tumores del subtipo Luminal representaron el grupo más numeroso en nuestras muestras, que concuerda con los reportes del TCGA¹¹. 70.1% de las muestras del TCGA corresponden a este subtipo (44.9% Luminal A- 25.2% Luminal B), comparado con el 64.3% en nuestras muestras (25.9% al Luminal A y el 38.4 al Luminal B).

En lo que respecta a pacientes Latinas, se ha reportado una mayor prevalencia de tumores positivos para la amplificación y sobreexpresión de HER2, en comparación con pacientes caucásicas^{18,48}. Esta observación se corrobora en nuestros resultados obtenidos mediante perfiles de expresión génica, presentando el subtipo enriquecido en HER2, casi el doble de prevalencia en nuestro grupo de muestras que en el reporte del TCGA (20.5% VS 11.3%).

El subtipo basal presentó una prevalencia de 18.6% en caucásicas, 24.60 % en Afroamericanas⁴⁸ y 12% en nuestras muestras. Se observa claramente que los subtipos tumorales clínicamente más agresivos, se presenta en mujeres más jóvenes⁴⁹ en comparación con otros subtipos⁵⁰. Este hallazgo podría relacionarse con la observación clínica de que los tumores que se presentan en mujeres pre menopáusicas en nuestro país por lo general, presentan cursos clínicos más agresivos.

El hecho de que este subtipo tumoral luminal A sea el más prevalente en las muestras que analizamos, aporta información relevante con posible relevancia clínica para el manejo de los tumores de mama en nuestra población, indicando que es necesario dar un seguimiento a largo plazo de las pacientes con este subtipo tumoral. Cada vez hay más datos que apuntan a que las diferencias significativas en la genética y la biología del cáncer de mama en las latinas contribuyen de forma importante a la mayor mortalidad, incluida una mayor incidencia de los cánceres de mama triplemente negativos. Es probable que otros factores sociales y ambientales también desempeñen un papel importante, como una menor tasa de mamografías de cribado y un acceso variable a la atención médica, entre otros⁵¹.

El uso de inmunohistoquímica se considera un estándar en el abordaje del cáncer de mama, al brindar una aproximación a los subtipos moleculares del cáncer de mama. Sin embargo, el acceso a las firmas genómicas necesarias para la clasificación molecular es limitado, por lo que en la práctica el uso de inmunohistoquímica es fundamental. Con la finalidad de explorar la correlación entre los marcadores inmunohistoquímicos y la clasificación obtenida mediante perfiles de expresión, se analizó la expresión de los RE, RP y HER2, EGFR, Ck 5/6, 14, 17, Claudinas 1,3, en las muestras de los diferentes subtipos. Esto permitió describir la prevalencia de subtipos de tumores de mama definidos molecularmente en la cohorte de luminales A RE+/RP+/HER2- en un 95%, los luminales B fueron RE+/RP+/HER2+ en un 9%, y los enriquecidos en HER2 RE-/RP-/HER2+ en un 83%, mientras que los tumores Basales fueron RE-/RP-/HER2- triples negativos en un 81%.

En México se llevado a cabo varios estudios para determinar los subtipo subrogados mediante inmunohistoquímica. Se evaluó una cohorte de 4315 pacientes basándose solo en RE, PR y HER2, encontraron que aunque las pacientes mexicanas jóvenes con cáncer

de mama tienen una enfermedad más agresiva en el momento de su presentación que las mujeres de mayor edad, sólo aquellas con tumores luminales B tienen una peor supervivencia⁵²⁻⁵³.

Por otro lado, en el Centro Médico Nacional 20 de noviembre, reportaron las siguientes frecuencias, RH (+)/HER2 (-) 69.6%, RH (+)/HER2(+) 17.5%, HER2(+) 5.3%, RH (-)/HER2 (-) 7.6%⁵⁴. Respecto a otros estudios en población mexicana, en una cohorte de pacientes del Seguro Popular, en el cual de 4,300 mujeres el 60.7% eran RH (+) / HER2 (-) y tenían un mejor pronóstico comparado con mujeres con cáncer de mama triple negativo⁵⁵. La frecuencia de cáncer de mama HER2 (+) fue de 23% y nosotros obtuvimos 18%. En un estudio en un hospital del Instituto Mexicano del Seguro Social (IMSS) por Pérez-Rodríguez⁵⁶ obtuvo una proporción de cáncer de mama Luminal A del 65% y Luminal B del 12%, lo cual concuerda con nuestros resultados, siendo los subtipos Luminales con expresión de RH los más frecuentes; sin embargo la proporción de triple negativos del 14% fue mayor a la reportada por nosotros, del 11%.

La mayor proporción de mujeres con cáncer de mama en nuestra población presentan un subtipo por inmunohistoquímica con expresión de receptores estrógeno y progesterona y este grupo de pacientes tienen una mejor supervivencia global comparada con las pacientes triple negativo. El desarrollo de nuevas opciones terapéuticas para este grupo de mal pronóstico es fundamental, buscando lograr un impacto como el que se ha logrado en las pacientes HER2 positivo.

En este trabajo también se evaluó la inestabilidad genómica, es una característica distintiva del cáncer que promueve la heterogeneidad inter e intratumoral para permitir la adaptación de las células cancerosas al estrés ambiental, impulsando así un comportamiento tumoral agresivo, y la resistencia a las terapias contra el cáncer⁵⁷. Con el fin de profundizar sobre las posibles consecuencias biológicas de los cambios en el número de copias, se correlaciono con la expresión de los genes contenidos en cada región alterada a lo largo de los subtipos PAM50.

Interesantemente, la ganancia de 8p11.23 se correlaciona positivamente con la sobreexpresión de ZNF703 en el tumor luminal⁵⁸ y la pérdida de 5q12.1 con la reducción de la expresión de DEPDC1B. Este gen se ha reportado sobreexpresado en el cáncer oral,

pulmón de células no pequeñas, el cáncer de próstata, el sarcoma de tejidos blandos, el cáncer de cuello de útero y el melanoma maligno⁵⁹. En los tumores luminales B, se identificó la regulación a la baja de PROM1, junto con la delección de la región 4p15.3. Se ha demostrado que PROM1, es un marcador de las células madre del cáncer, facilita la transición epitelio mesénquima en varios tipos de cáncer⁶⁰. Además, se observó que en el subtipo HER2 existe sobreexpresión de GRB7, STARD3⁶¹, PPP1R1B, PNMT, ABCC3 y PGAP3, que corresponden con la amplificación del amplicón Her2 17q12. Por último, los tumores basales presentan una expresión reducida de TMEM30B (14q23.1) el cual se ha reportado en cáncer de ovario⁶² y SNED1 (2q37.3)⁶³

Recientemente el grupo del Dr. Paniagua estudio la complejidad genómica y la heterogeneidad intrapaciente, así como sobre el desarrollo filogenético y evolutivo en muestras triple negativas sin respuesta a tratamiento, identificando los eventos troncales tempranos fueron amplificaciones de MCL1, las lesiones metastásico tenían delecciones en RB1 y PTEN, junto con TERT, AKT2 y amplificaciones de CCNE1, esto contribuye a dilucidar el fenómeno global de la evolución tumoral en el cáncer de mama triple negativo proporcionan datos genómicos que pueden ayudar a futuros estudios a definir alteraciones procesables para estas pacientes, permitiendo así una terapia más personalizada⁶⁴.

En los últimos años, la atención al paciente ha revolucionado ante la disponibilidad comercial de la secuenciación de nueva generación. Esta tecnología ha mejorado la identificación de mutaciones en la práctica diaria y la posterior prescripción de una terapia de diagnóstico. Las directrices actuales para los tumores sólidos recomiendan el uso de pruebas genómicas en la enfermedad progresiva avanzada para identificar posibles objetivos de tratamiento, como las fusiones y mutaciones de BRCA1/2, de receptores de tirosina neurotróficos (NTRK) y RET, la expresión de PD-(L)1, la reparación de desajustes del ADN, la inestabilidad de microsatélites y la carga mutacional. En el caso del cáncer de mama, las mutaciones de la línea germinal más frecuentes en BRCA1/2 se asocian a defectos de reparación de la recombinación homóloga del ADN. Así, estas pacientes son candidatas potenciales para el tratamiento con inhibidores de la poli-adenosina-difosforibosa polimerasa (PARP), ya que las células deficientes en BRCA1/2 son muy sensibles a la acumulación de roturas tóxicas de doble cadena, a la inestabilidad genómica y a la letalidad sintética causada por la inhibición de la PARP2, lo que representan un paso hacia el tratamiento personalizado⁶⁵.

Este tipo de estudios en los tumores a través de firmas genómicas o transcriptómicas se pueden aprovechar como herramienta predictiva del resultado clínico en el cáncer de mama. Hoy en día, la capacidad de comprender los principales aspectos de la biología del cáncer de mama y los métodos de detección más utilizados de detección permitirá desarrollar innovaciones para satisfacer las necesidades de las pacientes en tres aspectos: precisión, comodidad y accesibilidad.

CONCLUSIONES

El cáncer de mama es una enfermedad curable si se diagnostica en etapas tempranas, por lo que es fundamental mejorar los programas de tamizaje, diagnóstico y referencia de estas pacientes, para poder disminuir la frecuencia de casos en etapas avanzadas de la enfermedad, las cuales siguen siendo las más frecuentes en nuestra población.

El análisis global de los perfiles de expresión génica en cáncer de mama, permitió obtener un mayor conocimiento acerca de las diferencias biológicas y mecanismos moleculares. Así como, la descripción de la prevalencia de subtipos de tumores de mama definidos molecularmente en nuestra población.

La estratificación del cáncer de mama basada en la clasificación de subtipos intrínsecos no solo ha permitido una comprensión más profunda de las bases moleculares de la heterogeneidad de los tumores de mama, sino que también ha proporcionado medidas de puntaje clínicamente relevantes derivadas de características biológicas para evaluar la respuesta al tratamiento y el pronóstico de las terapias.

En este trabajo, se reporta una de las cohortes hispanas más grandes caracterizadas a nivel genómico y transcriptómico. A pesar de ello, una de las limitantes de nuestro estudio radica en la cantidad de muestras. Es por ello que con la finalidad de que los resultados encontrados pudieran reflejar con mayor precisión las tasas reales de prevalencia de subtipos y de las alteraciones moleculares subyacentes a los mismos en población hispana, aprovechamos la disponibilidad de datos públicos para tener una comparación multiétnica sólida y superar las limitaciones de tamaño.

Basándonos en nuestros resultados, obtuvimos la frecuencia de subtipos intrínsecos en cáncer de mama en población mexicana, Sin embargo, no se encontró buena correlación entre la sub-clasificación basada en el PAM50 contra el perfil inmunohistoquímica. Se validaron los genes diferencialmente expresados. Se identificaron genes y se correlacionaron a nivel de alteraciones de número de copias que pudieran implicar en el desarrollo del cáncer de mama.

Dado que el cáncer de mama es una enfermedad compleja y heterogénea, el estudio de diferentes enfoques onco-omics es una forma eficaz de revelar genes esenciales para comprender mejor el panorama molecular de los procesos que subyacen a la oncogénesis, y para desarrollar mejores tratamientos terapéuticos centrados en la farmacogenómica y la medicina de precisión.

PERSPECTIVAS

Nuestros resultados justifican la necesidad de futuros estudios que integren abordajes para el análisis de modificaciones genéticas de línea germinal y somáticas, así como la consideración del impacto de factores ambientales, para comprender mejor las causas de las disparidades en la evolución de las pacientes con cáncer de mama y poder desarrollar intervenciones más efectivas encaminadas a mejorar las oportunidades de supervivencia y mayor equidad en los esquemas de salud. Es necesario un enfoque multidisciplinario para desarrollar herramientas de detección y diagnóstico en cáncer de mama, ofreciendo información clave para los profesionales de la medicina y la investigación.

REFERENCIAS

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J Clin* **71**, 209–249 (2021).
2. Breast Globocan 2020.pdf.
3. Cárdenas-Sánchez, J. *et al.* Consenso Mexicano sobre diagnóstico y tratamiento del cáncer mamario. Octava revisión. Colima 2019. *GAMO* **18**, 2600 (2019).
4. FOLLETO_CONSENSO_DE_CANCER_DE_MAMA_9aRev2021a.pdf.
5. CONSENSO_DE_CANCER_DE_MAMA_8aRev2019a.pdf.
6. Chávarri-Guerra, Y. *et al.* Breast cancer in Mexico: a growing challenge to health and the health system. *The Lancet Oncology* **13**, e335–e343 (2012).
7. Slepicka, P. F., Cyrill, S. L. & dos Santos, C. O. Pregnancy and Breast Cancer: Pathways to Understand Risk and Prevention. *Trends in Molecular Medicine* **25**, 866–881 (2019).
8. Duivenvoorden, H. M. *et al.* Myoepithelial cell-specific expression of stefin A as a suppressor of early breast cancer invasion. *J. Pathol* **243**, 496–509 (2017).
9. Sharp, J. A., Lefevre, C., Brennan, A. J. & Nicholas, K. R. The Fur Seal—a Model Lactation Phenotype to Explore Molecular Factors Involved in the Initiation of Apoptosis at Involution. *J Mammary Gland Biol Neoplasia* **12**, 47–58 (2007).
10. Atabai, K., Sheppard, D. & Werb, Z. Roles of the Innate Immune System in Mammary Gland Remodeling During Involution. *J Mammary Gland Biol Neoplasia* **12**, 37–45 (2007).
11. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
12. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).
13. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov* **12**, 31–46 (2022).
14. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
15. Parker, J. S. *et al.* Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *JCO* **27**, 1160–1167 (2009).
16. Gao, J. J. & Swain, S. M. Luminal A Breast Cancer and Molecular Assays: A Review.

The Oncologist **23**, 556–565 (2018).

17. van 't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).

18. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).

19. Sorlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences* **98**, 10869–10874 (2001).

20. Creighton, C. The molecular profile of luminal B breast cancer. *BTT* 289 (2012) doi:10.2147/BTT.S29923.

21. Slamon, D. J. *et al.* Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **235**, 177–182 (1987).

22. Macrinici, V. & Romond, E. Clinical updates on EGFR/HER targeted agents in early-stage breast cancer. *Clin Breast Cancer* **10 Suppl 1**, E38-46 (2010).

23. Perou, C. M. Molecular stratification of triple-negative breast cancers. *Oncologist* **16 Suppl 1**, 61–70 (2011).

24. Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12**, R68 (2010).

25. Prat, A. *et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. *The Breast* **24**, S26–S35 (2015).

26. Kenn, M. *et al.* Decision theory for precision therapy of breast cancer. *Sci Rep* **11**, 4233 (2021).

27. Jiang, Y.-Z. *et al.* Molecular subtyping and genomic profiling expand precision medicine in refractory metastatic triple-negative breast cancer: the FUTURE trial. *Cell Res* **31**, 178–186 (2021).

28. López-Cortés, A. *et al.* OncoOmics approaches to reveal essential genes in breast cancer: a panoramic view from pathogenesis to precision medicine. *Sci Rep* **10**, 5285 (2020).

29. Griguolo, G. *et al.* Gene-expression signatures to inform neoadjuvant treatment decision in HR+/HER2– breast cancer: Available evidence and clinical implications. *Cancer Treatment Reviews* **102**, 102323 (2022).

30. Griguolo, G. *et al.* Gene-expression signatures to inform neoadjuvant treatment decision in HR+/HER2– breast cancer: Available evidence and clinical implications. *Cancer Treatment Reviews* **102**, 102323 (2022).

31. Sun, L., Wu, A., Bean, G. R., Hagemann, I. S. & Lin, C.-Y. Molecular Testing in Breast Cancer. *The Journal of Molecular Diagnostics* **23**, 1422–1432 (2021).
32. Kwa, M., Makris, A. & Esteva, F. J. Clinical utility of gene-expression signatures in early stage breast cancer. *Nat Rev Clin Oncol* **14**, 595–610 (2017).
33. Cardoso, F. *et al.* Clinical Application of the 70-Gene Profile: The MINDACT Trial. *JCO* **26**, 729–735 (2008).
34. Cardoso, F. *et al.* 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med* **375**, 717–729 (2016).
35. Paik, S. *et al.* A Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer. *N Engl J Med* **351**, 2817–2826 (2004).
36. Paik, S. *et al.* Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* **24**, 3726–3734 (2006).
37. Bargallo, J. E. *et al.* A study of the impact of the 21-gene breast cancer assay on the use of adjuvant chemotherapy in women with breast cancer in a Mexican public hospital. *J. Surg. Oncol.* **111**, 203–207 (2015).
38. Zack, T. I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nat Genet* **45**, 1134–1140 (2013).
39. Barba, D. *et al.* Breast cancer, screening and diagnostic tools: All you need to know. *Crit Rev Oncol Hematol* **157**, 103174 (2021).
40. Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12**, R68 (2010).
41. Tamez-Peña, J.-G. *et al.* Radiogenomics analysis identifies correlations of digital mammography with clinical molecular signatures in breast cancer. *PLoS One* **13**, e0193871 (2018).
42. Kwa, M., Makris, A. & Esteva, F. J. Clinical utility of gene-expression signatures in early stage breast cancer. *Nat Rev Clin Oncol* **14**, 595–610 (2017).
43. Limón, E. R. *et al.* Respuesta a la quimioterapia neoadyuvante (QTneo) según los subtipos intrínsecos determinados por inmunohistoquímica (IHQ) en pacientes con cáncer de mama. *Gaceta Médica de México*. 10.
44. Canello, G. *et al.* Prognosis and adjuvant treatment effects in selected breast cancer subtypes of very young women (<35 years) with operable breast cancer. *Ann Oncol* **21**, 1974–1981 (2010).
45. Hausauer, A. K., Keegan, T. H. M., Chang, E. T. & Clarke, C. A. Recent breast cancer trends among Asian/Pacific Islander, Hispanic, and African-American women in the US: changes by tumor subtype. *Breast Cancer Res* **9**, R90 (2007).

46. O'Brien, K. M. *et al.* Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. *Clin Cancer Res* **16**, 6100–6110 (2010).
47. Amankwaa-Frempong, E., Yeboah, F. A., Nguah, S. B. & Newman, L. A. Breast Cancer Genetic Testing Among African Patients With Breast Cancer: Deoxyribonucleic Acid Extraction From Tumor Tissue and International Multidisciplinary Partnerships. *JAMA Surg* **152**, 800–801 (2017).
48. Kurian, A. W., Fish, K., Shema, S. J. & Clarke, C. A. Lifetime risks of specific breast cancer subtypes among women in four racial/ethnic groups. *Breast Cancer Res* **12**, R99 (2010).
49. Schwartz, A. M., Henson, D. E. & Patel, A. Re: Age-specific incidence of breast cancer subtypes: understanding the Black-White crossover. *J Natl Cancer Inst* **105**, 368–370 (2013).
50. Narod, S. A. Breast cancer in young women. *Nat Rev Clin Oncol* **9**, 460–470 (2012).
51. Patel, T. A., Colon-Otero, G., Bueno Hume, C., Copland, J. A. & Perez, E. A. Breast Cancer in Latinas: Gene Expression, Differential Response to Treatments, and Differential Toxicities in Latinas Compared with Other Population Groups. *The Oncologist* **15**, 466–475 (2010).
52. Villarreal-Garza, C. *et al.* Molecular Subtypes and Prognosis in Young Mexican Women With Breast Cancer. *Clinical Breast Cancer* **17**, e95–e102 (2017).
53. Villarreal-Garza, C. & Castro-Sánchez, A. Preface. *Rev Invest Clin* **69**, 179–180 (2017).
54. Valle-Solís, A. E. *et al.* Supervivencia en cáncer de mama por subtipo mediante inmunohistoquímica: Un estudio retrospectivo. *GMM* **155**, 2041 (2019).
55. Reynoso-Noverón, N. *et al.* Clinical and Epidemiological Profile of Breast Cancer in Mexico: Results of the Seguro Popular. *JGO* **3**, 757–764 (2017).
56. Pérez-Rodríguez, G. Prevalencia de subtipos por inmunohistoquímica del cáncer de mama en pacientes del Hospital General Regional 72, Instituto Mexicano del Seguro Social. *Cirugía y Cirujanos* **83**, 193–198 (2015).
57. Kalimutho, M. *et al.* Patterns of Genomic Instability in Breast Cancer. *Trends Pharmacol Sci* **40**, 198–211 (2019).
58. Klæstad, E. *et al.* ZNF703 gene copy number and protein expression in breast cancer; associations with proliferation, prognosis and luminal subtypes. *Breast Cancer Res Treat* **186**, 65–77 (2021).
59. Zhang, S. *et al.* DEP Domain-Containing Protein 1B (DEPDC1B) Promotes Migration and Invasion in Pancreatic Cancer Through the Rac1/PAK1-LIMK1-Cofilin1 Signaling

Pathway. *Onco Targets Ther* **13**, 1481–1496 (2020).

60. Latorre, E. *et al.* The Ribonucleic Complex HuR-MALAT1 Represses CD133 Expression and Suppresses Epithelial-Mesenchymal Transition in Breast Cancer. *Cancer Res* **76**, 2626–2636 (2016).

61. Kauraniemi, P., Bärlund, M., Monni, O. & Kallioniemi, A. New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. *Cancer Res* **61**, 8235–8240 (2001).

62. Li, Y. *et al.* Differential gene expression identifies a transcriptional regulatory network involving ER-alpha and PITX1 in invasive epithelial ovarian cancer. *BMC Cancer* **21**, 768 (2021).

63. Naba, A., Clauser, K. R., Lamar, J. M., Carr, S. A. & Hynes, R. O. Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. *Elife* **3**, e01308 (2014).

64. Martínez-Gregorio, H. *et al.* The Evolution of Clinically Aggressive Triple-Negative Breast Cancer Shows a Large Mutational Diversity and Early Metastasis to Lymph Nodes. *Cancers (Basel)* **13**, 5091 (2021).

65. Bournal, M. T., Barragan-Carrillo, R., Mesa-Chavez, F. & Villarreal-Garza, C. Challenges of Treating Young Women With Cancer in the Era of New Oncologic Treatments. *RIC* **73**, 6784 (2021).

PUBLICACIONES

Romero-Cordoba SL, Salido-Guadarrama I, **Rebollar-Vega R**, Bautista-Piña V, Domínguez-Reyes C, Tenorio-Torres A, Villegas-Carlos F, Fernández-López JC, Uribe-Figueroa L, Alfaro-Ruiz L, Hidalgo-Miranda A. Comprehensive omic characterization of breast cancer in Mexican-Hispanic women. *Nat Commun.* 2021 Apr 14; 12(1):2245.

Cedro-Tanda A, Ríos-Romero M, Romero-Córdoba S, Cisneros-Villanueva M, **Rebollar-Vega RG**, Alfaro-Ruiz LA, Jiménez-Morales S, Domínguez-Reyes C, Villegas-Carlos F, Tenorio-Torres A, Bautista-Piña V, Beltrán-Anaya FO, Hidalgo-Miranda A. A lncRNA landscape in breast cancer reveals a potential role for AC009283.1 in proliferation and apoptosis in HER2-enriched subtype. *Sci Rep.* 2020 Aug 4; 10(1):13146. doi: 10.1038/s41598-020-69905-z.

Beltrán-Anaya FO, Romero-Córdoba S, **Rebollar-Vega R**, Arrieta O, Bautista-Piña V, Domínguez-Reyes C, Villegas-Carlos F, Tenorio-Torres A, Alfaro-Ruiz L, Jiménez-Morales S, Cedro-Tanda A, Ríos-Romero M, Reyes-Grajeda JP, Tagliabue E, Iorio MV, Hidalgo-Miranda A. Expression of long non-coding RNA ENSG00000226738 (LncKLHDC7B) is enriched in the immunomodulatory triple-negative breast cancer subtype and its alteration promotes cell migration, invasion, and resistance to cell death. *Mol Oncol.* 2019 Jan 16. doi: 10.1002/1878-0261.12446.

Fernández-López JC, Romero-Córdoba S, **Rebollar-Vega R**, Alfaro-Ruiz LA, Jiménez-Morales S, Beltrán-Anaya F, Arellano-Llamas R, Cedro-Tanda A, Ríos-Romero M, Ramírez-Florencio M, Bautista-Piña V, Domínguez-Reyes C, Villegas-Carlos F, Tenorio-Torres A, Hidalgo-Miranda A. Population and breast cancer patients' analysis reveals the diversity of genomic variation of the BRCA genes in the Mexican population. *Genomics.* 2019 Jan 10; 13(1):3. doi: 10.1186/s40246-018-0188-9.

Arellano-Llamas R, Alfaro-Ruiz L, Arriaga Canon C, Imaz Rosshandler I, Cruz-Lagunas A, Zúñiga J, **Rebollar Vega R**, Wong CW, Maurer-Stroh S, Romero Córdoba S, Liu ET, Hidalgo-Miranda A, Vázquez-Pérez JA. Molecular features of influenza A (H1N1) pdm09 prevalent in Mexico during winter seasons 2012-2014. *PLoS One.* 2017 Jul 10; 12(7):e0180419. doi: 10.1371/journal.pone.0180419.

Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, Lawrence MS, Sivachenko AY, Sougnez C, Zou L, Cortes ML, Fernandez-Lopez JC, Peng S, Ardlie KG, Auclair D, Bautista-Piña V, Duke F, Francis J, Jung J, Maffuz-Aziz A, Onofrio RC, Parkin M, Pho NH, Quintanar-Jurado V, Ramos AH, **Rebollar-Vega R**, Rodriguez-Cuevas S, Romero-Cordoba SL, Schumacher SE, Stransky N, Thompson KM, Uribe-Figueroa L, Baselga J, Beroukhim R, Polyak K, Sgroi DC, Richardson AL, Jimenez-Sanchez G, Lander ES, Gabriel SB, Garraway LA, Golub TR, Melendez-Zajgla J, Toker A, Getz G, Hidalgo-Miranda A, Meyerson M. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature*. 2012 Jun 20; 486(7403):405-9. doi: 10.1038/nature11154.

Romero-Cordoba S, Rodriguez-Cuevas S, **Rebollar-Vega R**, Quintanar-Jurado V, Maffuz-Aziz A, Jimenez-Sanchez G, Bautista-Piña V, Arellano-Llamas R, Hidalgo-Miranda A. Identification and pathway analysis of microRNAs with no previous involvement in breast cancer. *PLoS One*. 2012; 7(3):e31904. doi: 10.1371/journal.pone.0031904. Epub 2012 Mar 16. Erratum in: *PLoS One*. 2015; 10(9):e0137738.