



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
DOCTORADO EN CIENCIAS BIOMÉDICAS  
FACULTAD DE MEDICINA

**ESTUDIO GENÓMICO DE LAS CÉLULAS INICIADORAS TUMORALES DEL CÁNCER  
DE MAMA**

**TESIS**  
QUE PARA OPTAR POR EL GRADO DE:  
**DOCTORA EN CIENCIAS**

PRESENTA  
**ANGELA MARÍA SCHWARZ CRUZ Y CELIS**

DIRECTOR DE TESIS  
**DR. JORGE MELÉNDEZ ZAJGLA**  
FACULTAD DE MEDICINA

COMITÉ TUTOR  
**DRA. YOLANDA LÓPEZ VIDAL**  
FACULTAD DE MEDICINA  
**DR. RUBÉN LISKER YOURKOWITZKY<sup>†</sup>**  
FACULTAD DE MEDICINA



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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 se realizó en el laboratorio de “Genómica Funcional del Cáncer” en la  
División de Investigación Básica del  
Instituto Nacional de Medicina Genómica (INMEGEN)  
bajo la dirección y tutela del  
**Dr. Jorge Meléndez Zajgla.**

---

---

## RESUMEN

El cáncer es un grupo de enfermedades caracterizadas por el crecimiento descontrolado y anormal de diferentes tipos celulares que eventualmente invaden tejidos cercanos y distantes<sup>1</sup>. En los últimos años se ha planteado la posibilidad de que el mantenimiento de estos tumores sea dirigido por un grupo de células denominadas células iniciadoras tumorales (CIT) o células troncales cancerosas (CTC)<sup>1-17</sup>.

Estas células se caracterizan por regenerarse ilimitadamente, aunque a un ritmo bajo, presentar una división celular asimétrica (en donde cada célula genera dos células, una que es idéntica y mantiene las características iniciadoras tumorales y la segunda comprometida a un cierto linaje celular), así como también por su capacidad de restablecer el tumor en sitios distantes. Asimismo, éstas células son altamente resistentes a la terapia oncológica, por razones aún no totalmente esclarecidas. Se ha sugerido que estas células podrían ser responsables no solo de la metástasis, sino de la reincidencia de los tumores después de la terapia, por lo que son evidentemente un blanco terapéutico muy importante<sup>1-5,8-11,13,14,16-34</sup>.

Aunque no se conoce con certeza el origen de estas células, se piensa que podrían derivar de células madre normales multipotenciales, células progenitoras, o incluso de células parcialmente diferenciadas que sufrirían una dediferenciación. A pesar del reconocimiento de estas células y de su posible importancia, no existen reportes que describan sus características moleculares<sup>1-5,9,10,14,16,18-20,35-39</sup>. Es por esto que este estudio se dirigió a caracterizar y a analizar los perfiles genómicos de las CIT y no CIT de líneas celulares de cáncer de mama, para así en un futuro, poder desarrollar nuevos blancos terapéuticos y posibles marcadores pronósticos y diagnósticos de esta enfermedad. En particular, estamos interesados en utilizar modelos de cáncer de mama, debido a que esta neoplasia es la segunda causa de muerte en la población femenina de 30 a 54 años de edad de nuestro país y se ubica dentro de las primeras dos causas de mortalidad por tumores malignos a nivel mundial<sup>4,9,10,14,24,35,39-49</sup>.

---

Para conocer las diferencias entre los perfiles genómicos de las CIT y no CIT se utilizaron como modelo de estudio líneas celulares de origen basal y luminal de cáncer de mama. Ambas poblaciones fueron aisladas y caracterizadas fenotípicamente para realizar posteriormente el análisis de los perfiles estructurales, transcripcionales y de metilación mediante herramientas genómicas de alto rendimiento y cobertura, incluyendo microarreglos y secuenciación de nueva generación. A nivel estructural no se observaron diferencias en las poblaciones CIT y no CIT, mientras que el transcriptoma y metiloma presentaron cambios significativos, incluyendo activación de vías relacionadas con troncalidad, crecimiento y proliferación celular, ciclo celular y apoptosis, así como también, enriquecimientos en la actividad de factores clave de transcripción como NANOG, SOX2, KLF4 y OCT3/4, y niveles de expresión y metilación desregulados, tanto en genes previamente reportados por otros autores como en genes que aún no han sido asociados con nuestra población celular de interés como es el caso de UPK1B, PAX2, ATRX, entre otros. Al realizar el análisis de correlación entre el perfil de expresión y metilación observamos un posible vínculo entre BIRC5 (que es de nuestro particular interés) y las CIT el cual comenzamos a explorar en estudios posteriores. Por otro lado, se comprobó el potencial tumorigénico de las poblaciones caracterizadas como CIT en comparación con las no CIT en ratones *nu/nu*.

Estos resultados demuestran que las principales diferencias entre estos dos grupos celulares radican en su perfil de expresión y metilación y no en el estructural, concluyendo y comprobando así nuestra hipótesis en la que propusimos que el perfil genómico de las CIT es diferente al del resto de las poblaciones tumorales.

---

## SUMMARY

Cancer is a group of diseases characterized by the uncontrolled and abnormal growth of different cell types that are able to invade local and distant tissues<sup>1</sup>. In recent years it has been postulated that the maintenance of these tumors could be driven by a group of cells known as tumor or cancer initiating cells (CICs or TICs) as well as cancer stem cells (CSCs)<sup>1–17</sup>. These cells are able to divide asymmetrically and limitless, and to give rise to descendants that can re-establish the tumor at local or distant places. Likewise, it has been suggested that TICs could be responsible not only for tumor metastasis and drug resistance, but also for tumor recurrence after cancer therapy, turning them into a very important therapeutic target<sup>1–5,8–11,13,14,16–34</sup>.

Although nowadays we do not know the exact origin of CSCs, it has been proposed that they could derive from adult stem cells, progenitor cells or even from terminally differentiated cells that undergo dedifferentiation. In spite of the efforts made to date to unravel the functional and biological features of this group of cells, there are no reports that describe their molecular properties<sup>1–5,9,10,14,16,18–20,35–39</sup>. For this reason, we sought to analyze the genomic profile of TICs derived from breast cancer cell lines, so in the future there could be more results and information that could help to the development of new targeted therapies as well as prognostic and diagnostic markers for this disease. Our interest in this type of cancer relies on the fact that breast cancer in Mexico is the second cause of death in the female population from 30 to 54 years of age and is ranked within the first two causes of mortality by malignant tumors worldwide<sup>4,9,10,14,24,35,39–49</sup>.

In the present study, breast cancer cell lines derived from basal and luminal subtypes were used as model systems to analyze the genomic profile of TICs and non-TICs. Both cell populations were isolated using cell surface markers and then utilized to perform the transcriptome, genomic structure, and methylome analysis with high-throughput whole-genome technologies, including microarrays and Next-Generation Sequencing (NGS). We found that the genotype of breast TICs does not differ from the non-TICs. Nevertheless, the

---

transcriptome and methylome of TICs presented marked differences from the non-TICs including activation of pathways related to cell cycle, cellular growth, proliferation, apoptosis and stemness, enrichment in key transcription factors activity like NANOG, SOX2, KLF4 y OCT3/4, and deregulated expression and methylation of genes previously reported by other authors as in genes that have not been associated with this cell population such as UPK1B, PAX2, ATRX among others. Afterwards, we performed a correlation study using gene expression and methylation levels and we observed a possible link between BIRC5 (of our particular interest) and TICs which we began to explore in later studies. Finally, we validated TICs ability to form tumors in immunodeficient mice. These results demonstrate that the main differences between the genomic profiles of TICs and non-TICs are in their transcriptome and methylome and not in their genotype. In conclusion, in this study we showed that the genomic profile of breast TICs is different from that of the rest of the tumor populations.

---

*Para el Mano José Manuel,  
Diego, Santiago y Martín.*

---

## AGRADECIMIENTOS

A Dios, por todo lo que me ha dado, por ponerme aquí y ahora.

A mi esposo, José Manuel, por su cariño, por creer en mí, por sus ánimos, por sus consejos, por siempre estar, por ponernos siempre antes que nada y nadie, por nuestros hijos, por su paciencia, por ser un ejemplo a seguir pero sobretodo, por su apoyo incondicional. Sin ti, no lo hubiera logrado.

A mis hijos, Diego, Santiago y Martín, por existir, por su paciencia, por quererme tal y como soy y por sus sonrisas interminables.

A mis padres, Ángeles y Walter, por creer en mí, por su apoyo y enseñanzas, por su cariño, por haberme enseñado a volar, por haber hecho todo lo que estaba en sus manos para darme lo mejor.

A mis abuelos, Ángeles<sup>†</sup> y Agustín<sup>†</sup>, por todo el cariño y apoyo que me dieron, por haberme introducido a la medicina.

A mi familia, Andre, Grossi<sup>†</sup>, Ruedi<sup>†</sup>, Andrea, Luis Miguel, Uche, Andrés, Gabriela, Juan José, Patricio, Fernando, Gonzalo, Matías, Luciana, Catalina, Nadia, Rodrigo, Mariana, Juan Carlos, Paula, Paulina y Manuel<sup>†</sup>. Por que cada uno de ustedes fueron y son parte de este largo camino que recorrí, por su apoyo y cariño, por sus ánimos y simplemente por que siempre SIEMPRE han estado ahí.

A mis amigas, María José P., Moni, Ana, Bego, Lucía, María José H., Mariana, Gabriela, Nadia, Karin, Dana y Paola, por ser parte fundamental de mi vida y de este logro, por sus porras, por su hermandad, por sus consejos, por su cariño y apoyo incondicional, por sus risas y lágrimas, por ser, por estar y seguir estando. Viva la vida.

A Lourdes y Roberto, por su cariño y apoyo, por siempre hacerme sentir en familia.

A mi tutor, Dr. Jorge Meléndez Zajgla, por permitirme hacer mi posgrado con él y su equipo de trabajo, por sus enseñanzas, por haber creído en mí, por su apoyo, paciencia y tiempo.

A mis co-tutores, Dra. Vilma Maldonado Lagunas, Dra. Yolanda López Vidal y Dr. Rubén Lisker Yourkowitzky<sup>†</sup>, por su apoyo, por sus consejos y enseñanzas, por su tiempo.

A mis maestras y amigas del laboratorio, Magali, Gisela, Nelly y Floria, por sus enseñanzas, por su tiempo, por su cariño y apoyo, por su amistad, por su comprensión y paciencia.

---

A la Unidad de Genotipificación y Análisis de Expresión del INMEGEN (hoy en día llamada Unidad de Microarreglos o UMI), especialmente a Fabiola Morales y Raúl Mojica, gracias por haberme hecho sentir siempre bienvenida y por tener las puertas abiertas (24/7) de la Unidad para mí, por su apoyo y por sus enseñanzas.

A la Unidad de Citometría de Flujo (UCIF) del INMEGEN, especialmente a Nelly Patiño, por su apoyo, cariño y enseñanzas pero sobretodo por su interminable paciencia.

A la Unidad de Secuenciación (USEC) del INMEGEN, especialmente a Alfredo Mendoza por su ayuda y paciencia.

Al Laboratorio de Histología y Microscopía (LHIM) del INMEGEN, especialmente a la Bióloga Valeria Quintanar y a la Dra. Gabiño por su ayuda y paciencia.

Al Centro de Información del INMEGEN, especialmente al Lic. Alfredo Gallo por todo el apoyo bibliográfico que me brindó.

A la Unidad de Citometría de Flujo del CINVESTAV y a la Unidad de Producción y Experimentación de Animales de Laboratorio-Bioterio de la Universidad Autónoma de México (UPEAL-B) de la Unidad Xochimilco de la Cd.Mx., por haberme permitido utilizar sus instalaciones y servicios.

A mi ex-alumno, Juan Canales, por su paciencia, por haberme enseñado a enseñar, por haberme dejado ser su maestra.

A mis amigos, maestros, compañeros y cómplices de laboratorio, Denise Campos e Ivan Bahena, por su cariño, por su apoyo, por sus infinitas enseñanzas, por sus colaboraciones, por su amistad y por las famosas pláticas enriquecedoras a la hora de la comida.

A Evangelina Vargas Barrientos auxiliar del Posgrado de Ciencias Biomédicas de la UNAM, por todo el apoyo y cariño brindado hasta el día de hoy.

A la Unidad de Posgrado de Ciencias Biomédicas de la UNAM y al CONACYT, por todo su apoyo.

**GRACIAS**

---

## ÍNDICE

RESUMEN	I
SUMMARY	III
AGRADECIMIENTOS	VI
ÍNDICE	VIII
ÍNDICE DE FIGURAS	XI
ÍNDICE DE TABLAS	XIII
1. INTRODUCCIÓN	1
2. CÉLULAS INICIADORAS TUMORALES	4
2.1 ANTECEDENTES HISTÓRICOS	4
2.2 CARACTERÍSTICAS Y PROPIEDADES DE LAS CIT	5
2.3 ORIGEN	8
2.4 IDENTIFICACIÓN Y AISLAMIENTO DE LAS CIT	9
2.5. TERAPÉUTICA	11
3. CÁNCER DE MAMA	14
4. CÉLULAS INICIADORAS TUMORALES DEL CÁNCER DE MAMA	17
4.1. ANTECEDENTES HISTÓRICOS, CARACTERÍSTICAS Y PROPIEDADES DE LAS CITCM	17
4.2 ORIGEN Y AISLAMIENTO DE LAS CITCM	19
4.3. MARCADORES DE LAS CITCM	22
4.4 CITCM Y NCRNAS	26
4.5 VÍAS DE SEÑALIZACIÓN EN LAS CITCM	27
4.6 TERAPÉUTICA ACTUAL	30
5. SURVIVINA	34
6. JUSTIFICACIÓN	36

---

7. HIPÓTESIS	36
8. OBJETIVO GENERAL	36
9. OBJETIVOS ESPECÍFICOS	36
10. IMPACTO Y RELEVANCIA DEL PROYECTO EN EL ÁREA DE ESTUDIO	37
11. METODOLOGÍA	38
11.1 LÍNEAS CELULARES	38
11.2 CULTIVO CELULAR	39
11.3 AISLAMIENTO Y CARACTERIZACIÓN DE LAS CITCM	39
11.4 EXTRACCIÓN DE ARN TOTAL	40
11.5 EXTRACCIÓN DE ADN	41
11.6 ANÁLISIS DE EXPRESIÓN DE LAS CITCM	41
11.6.1 ANÁLISIS DE RESULTADOS DE LAS MUESTRAS DE LA PRIMERA ETAPA A	41
11.6.2 ANÁLISIS DE RESULTADOS DE LAS MUESTRAS DE LA PRIMERA ETAPA B	42
11.7 ANÁLISIS BIOINFORMÁTICO DE REDES, CASCADAS Y FUNCIONES CELULARES REGULADAS DIFERENCIALMENTE EN LAS CITCM	42
11.8 ESTUDIO ESTRUCTURAL DE LAS CITCM	43
11.9 ESTUDIO DEL PERFIL DE METILACIÓN DE LAS CITCM	44
11.10 FORMACIÓN DE MAMOESFERAS	45
11.11 FORMACIÓN DE COLONIAS EN AGAR SUAVE	45
11.12 INDUCCIÓN TUMORAL EN RATONES HEMBRA NU/NU	46
11.12.1 INOCULACIÓN DE CITCM Y NO CITCM PURIFICADAS POR FACS (2011)	46
11.12.2 INOCULACIÓN DE CITCM Y NO CITCM PURIFICADAS POR FACS (2012)	46
11.13 DIAGNÓSTICO HISTOPATOLÓGICO DE LAS TUMORACIONES FORMADAS EN LOS RATONES NU/NU	47
12. DIAGRAMA GENERAL DE TRABAJO	48
13. RESULTADOS	49

---

---

13.1 AISLAMIENTO Y CARACTERIZACIÓN FENOTÍPICA DE LAS CITCM A PARTIR DE LÍNEAS CELULARES	49
13.1.1 PRIMERA ETAPA A	49
13.1.2 PRIMERA ETAPA B	51
13.2 CARACTERIZACIÓN FUNCIONAL DE LAS CITCM	53
13.3 ESTUDIO DE LA ESTRUCTURA DEL GENOMA DE LAS CITCM	58
13.4 ESTUDIO DEL PERFIL DE METILACIÓN DE LAS CITCM	62
13.5 IDENTIFICACIÓN DE GENES DIFERENCIALMENTE EXPRESADOS EN LAS CITCM	66
13.5.1 PRIMERA ETAPA A	66
13.5.2 PRIMERA ETAPA B	78
14. DISCUSIÓN	88
14.1 AISLAMIENTO, CARACTERIZACIÓN FENOTÍPICA Y FUNCIONAL Y ANÁLISIS GENÓMICO DE LAS CITCM A PARTIR DE LÍNEAS CELULARES	89
14.1.1 PRIMERA ETAPA A	89
14.1.2 PRIMERA ETAPA B	94
15. CONCLUSIONES	99
16. REFERENCIAS	101
17. ANEXOS	107
ANEXO 1. ABREVIATURAS	107
ANEXO 2. VALIDACIÓN DE EPCAM CON UNIPARTS	114
ANEXO 3. PROYECTO DE TITULACIÓN DE LICENCIATURA DEL ALUMNO JUAN CANALES DELGADO	115
ANEXO 4. PUBLICACIONES Y ACTIVIDADES	117
ANEXO 5. ARTÍCULO	118

---

## ÍNDICE DE FIGURAS

FIGURA 1. EVOLUCIÓN DE LA MORTALIDAD POR CÁNCER ESTANDARIZADA POR EDADES.	1
FIGURA 2. TEORÍAS DEL DESARROLLO Y HETEROGENEIDAD TUMORAL.	3
FIGURA 3. AUTORENOVACIÓN Y DIFERENCIACIÓN DE CT Y CIT.	9
FIGURA 4. TERAPIAS ANTITUMORALES.	13
FIGURA 5. FORMACIÓN TUMORAL EN RATONES NOD/SCID.	19
FIGURA 6. MODELOS DE FORMACIÓN DE LAS CITCM.	21
FIGURA 7. POBLACIONES OBTENIDAS POSTERIOR A LA SEPARACIÓN CON MACS <sup>®</sup> DE LA LÍNEA CELULAR MDA-MB-231.	50
FIGURA 8. LECTURA Y SEPARACIÓN POR FACS DE LAS POBLACIONES CITCM Y NO CITCM DE MCF7.	50
FIGURA 9. LECTURA DE LA POBLACIÓN CITCM DE MDA-MB231 PREVIA AL AISLAMIENTO CON FACS.	52
FIGURA 10. INDUCCIÓN TUMORAL EN RATONES UTILIZANDO LA LÍNEA CELULAR MCF-7.	54
FIGURA 11. TINCIÓN CON HEMATOXILINA Y EOSINA DE LAS POBLACIONES DE MCF-7.	54
FIGURA 12. ENSAYO DE FORMACIÓN DE COLONIAS DE LAS POBLACIONES DE MDA-MB-231.	55
FIGURA 13. PRUEBAS DE CARACTERIZACIÓN FUNCIONAL CON MDA-MB-231.	56
FIGURA 14. NORMALIZACIÓN Y PCA DE LOS DATOS DE GENOTIPIFICACIÓN DE MDA-MB-231.	58
FIGURA 15. GRÁFICA DEL ANÁLISIS DE CNVS DE LAS POBLACIONES CITCM DE MDA-MB-231.	59
FIGURA 16. GRÁFICA DEL ANÁLISIS DE CNVS DEL CROMOSOMA 11 EN CITCM.	61
FIGURA 17. GRÁFICA DE LAS REGIONES DIFERENCIALMENTE METILADAS EN LAS CITCM.	63
FIGURA 18. GRÁFICAS DEL ESTADO DE METILACIÓN EN LAS CITCM Y NO CITCM.	63
FIGURA 19. AGRUPACIÓN JERÁRQUICA Y MAPA DE CALOR DE LOS GENES DIFERENCIALMENTE METILADOS EN LAS CITCM Y NO CITCM.	65
FIGURA 20. NORMALIZACIÓN Y PCA DE LOS DATOS DE EXPRESIÓN DE MDA-MB-231.	67
FIGURA 21. DIAGRAMA DE VENN DEL ANÁLISIS DE EXPRESIÓN DE CITCM.	68

---

FIGURA 22. VÍA DE EPCAM CON SUS INTERACCIONES Y UNIONES OBSERVADAS MEDIANTE EL PROGRAMA IPA.	71
FIGURA 23. REGULADORES TRANSCRIPCIONALES Y FACTOR DE TRANSCRIPCIÓN NANOG EN LAS CITCM.	73
FIGURA 24. DIAGRAMA DE VENN DEL ANÁLISIS DE EXPRESIÓN DE LAS CITCM POR LÍNEA CELULAR.	78
FIGURA 25. DIAGRAMA DE VENN DEL ANÁLISIS COMPARATIVO DE EXPRESIÓN Y METILACIÓN.	81
FIGURA 26. VÍAS DE SEÑALIZACIÓN ENRIQUECIDAS EN LAS CITCM.	83
FIGURA 27. FACTORES DE TRANSCRIPCIÓN ACTIVADOS EN EL ANÁLISIS COMPARATIVO DE EXPRESIÓN Y METILACIÓN.	86
FIGURA 28. GSEA DEL ANÁLISIS COMPARATIVO DE EXPRESIÓN Y METILACIÓN.	87
FIGURA 29. DIAGRAMA DE VENN DEL ANÁLISIS COMPARATIVO DE LOS LNCRNAS DIFERENCIALMENTE EXPRESADOS EN LAS CITCM.	97

---

## ÍNDICE DE TABLAS

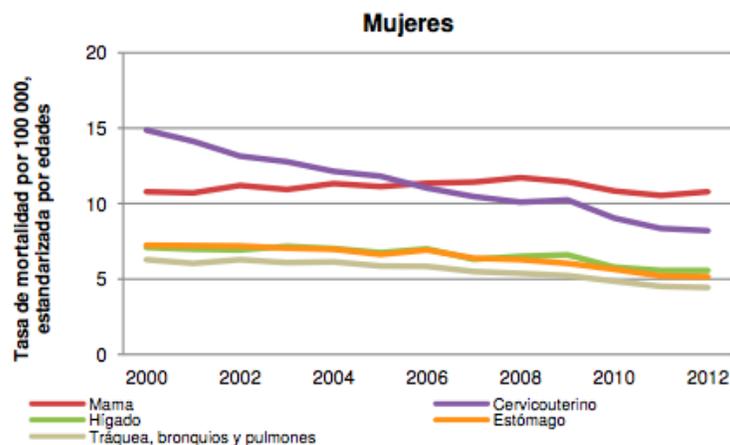
TABLA 1. CARACTERÍSTICAS PRINCIPALES DE LAS CIT.	7
TABLA 2. MARCADORES DE CIT EN DIVERSOS TIPOS TUMORALES.	10
TABLA 3. CLASIFICACIÓN MOLECULAR DEL CÁNCER DE MAMA.	17
TABLA 4. CARACTERÍSTICAS DE LAS LÍNEAS CELULARES DE ORIGEN BASAL Y LUMINAL.	38
TABLA 5. EXPRESIÓN DE LOS MARCADORES DE LAS CITCM EN LÍNEAS CELULARES DE CÁNCER DE MAMA DE ORIGEN BASAL.	51
TABLA 6. DIAGNÓSTICO HISTOPATOLÓGICO DE TUMORACIONES DE MDA-MB-231.	57
TABLA 7. CNVS OBSERVADAS EN LA POBLACIÓN CITCM DE LA LÍNEA CELULAR MDA-MB-231.	59
TABLA 8. CNVS OBSERVADAS EN LA POBLACIÓN CITCM DE LA LÍNEA CELULAR BT-20.	59
TABLA 9. CNVS OBSERVADAS EN LA POBLACIÓN CITCM DE LA LÍNEA CELULAR MDA-MB-468.	60
TABLA 10. LOS 20 CGS MÁS DIFERENCIALMENTE METILADOS EN LAS CITCM.	64
TABLA 11. LOS 10 GENES MÁS SOBRE Y SUB-REGULADOS EN LA POBLACIÓN CITCM (F+ DE MDA-MB-231).	68
TABLA 12. OTRAS MOLÉCULAS DIFERENCIALMENTE EXPRESADAS EN LA POBLACIÓN CITCM OBSERVADAS POSTERIOR AL ANÁLISIS DE EXPRESIÓN CON PGS.	69
TABLA 13. PREDICCIÓN GENERADA EN EL PROGRAMA IPA DEL ESTADO DE ACTIVACIÓN DE LOS REGULADORES TRANSCRIPCIONALES EN LAS CITCM.	70
TABLA 14. VÍAS CANÓNICAS DE SEÑALIZACIÓN Y ENRIQUECIMIENTO DE FUNCIONES MOLECULARES Y CELULARES OBSERVADAS EN LAS POBLACIONES CITCM.	72
TABLA 15. LOS 10 GENES MÁS SOBRE Y SUB-REGULADOS EN LA POBLACIÓN CITCM.	74
TABLA 16. GENES DIFERENCIALMENTE EXPRESADOS EN LA POBLACIÓN CITCM.	75
TABLA 17. LAS 50 MOLÉCULAS MÁS SOBRE-EXPRESADAS EN LAS CITCM PROVENIENTES DE LAS TRES LÍNEAS CELULARES.	79
TABLA 18. LAS 50 MOLÉCULAS MÁS SUB-EXPRESADAS EN LAS CITCM PROVENIENTES DE LAS TRES LÍNEAS CELULARES.	80
TABLA 19. GENES DE INTERÉS DEL ANÁLISIS COMPARATIVO DE EXPRESIÓN Y METILACIÓN.	82

---

## 1. INTRODUCCIÓN

A pesar de los grandes avances que se han generado durante décadas de estudio e investigaciones del cáncer, esta enfermedad sigue siendo un gran problema de salud pública a nivel mundial, representando las primeras causas de morbi-mortalidad en hombres y mujeres<sup>42-44,50,51</sup>. En nuestro país el cáncer de órganos respiratorios, digestivos, hematopoyéticos, mama, cérvicouterino y próstata constituyen los de mayor incidencia y mortalidad en la población adulta, mientras que en la edad pediátrica destaca el cáncer de órganos hematopoyéticos sin importar edad quinquenal (partiendo de 0 a 19 años de edad), sexo y estado socioeconómico<sup>50,52</sup>. A partir del año 2006, el cáncer de mama superó al cérvicouterino (Figura 1), se ubicó y se ha mantenido como la primera causa de mortalidad por tumores malignos y es considerado como la segunda causa de muerte en las mujeres de 30 a 54 años de edad<sup>41,51-53</sup>.

**Figura 1.** Evolución de la mortalidad por cáncer estandarizada por edades.



Observamos como en el 2006 la tasa de mortalidad del cáncer de mama superó a la del cáncer cérvicouterino. (Tomada de OMS; 2014)<sup>52</sup>.

---

A nivel biológico, el cáncer ha sido descrito como un conjunto de enfermedades genéticas complejas, que consta de cambios dinámicos en el genoma. La tumorigénesis es un proceso que cuenta con múltiples pasos en donde se generan alteraciones genéticas que conllevan a la transformación de las células normales a células cancerosas<sup>1,9,54</sup>. En el año 2000, Hanahan y Weinberg propusieron que durante este proceso (sin importar el tipo de tumor) se adquieren seis características principales que permiten que la tumorigénesis se lleve a cabo y permita el crecimiento tumoral y la diseminación metastásica. Estas características son: resistencia a muerte celular, activación de invasión y metástasis, inducción de angiogénesis, evasión de supresores de crecimiento, señales de proliferación sostenida y la activación de la inmortalidad replicativa. Once años después, los mismos autores agregaron dos nuevas características fundamentales a este modelo que son la evasión de la destrucción inmunológica y la reprogramación del metabolismo energético<sup>16,54–56</sup>.

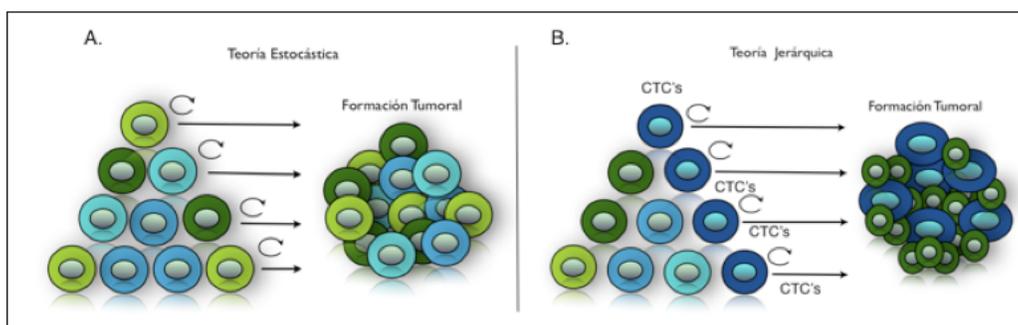
Una vez formado el tumor, las células cancerosas sufren cambios genéticos y epigenéticos, e interactúan con su microambiente (estroma, células inflamatorias, vasculatura, etc.) generando así múltiples tipos celulares dentro de una misma neoplasia, este resultado es conocido como “heterogeneidad tumoral”, el cual es fundamental para la progresión tumoral, resistencia terapéutica y a la larga, para la invasión y metástasis<sup>1,4,5,8,9,18,19,35,48,57,58</sup>.

Hoy en día a pesar de los grandes avances que se han logrado en el campo de la biología del cáncer, el origen y la progresión de la célula tumoral siguen siendo una interrogante<sup>1,5,18,59</sup>. Existen dos modelos los cuales no se excluyen mutuamente y que tratan de explicar el desarrollo, la heterogeneidad y la progresión tumoral: el modelo estocástico y el modelo jerárquico o de las células iniciadores tumorales (Figura 2). El modelo estocástico plantea que cualquier célula dentro del tumor cuenta con un potencial tumorigénico y de autorenovación, por lo tanto, cualquiera podrá formar y mantener la masa tumoral, mientras que el modelo jerárquico propone que solo una pequeña población celular con

---

determinadas capacidades de autorenovación y diferenciación ilimitada puede generar y mantener un tumor, estas células son denominadas “células iniciadoras tumorales” o “células troncales cancerosas” debido a que cuentan con un fenotipo troncal<sup>1,4,5,8,9,12,13,18–22,25,26,29,37,38,47,49,60–72</sup>.

**Figura 2.** *Teorías del desarrollo y heterogeneidad tumoral.*



**A.** Teoría Estocástica: plantea que cualquier célula dentro del tumor cuenta con un potencial tumorigénico y de autorenovación, por lo que cualquiera podrá formar y mantener la masa tumoral.  
**B.** Teoría Jerárquica o de las CIT o CTC: propone que solo las CIT pueden generar y mantener un tumor. (Modificado de Schwarz and Meléndez; 2011)<sup>5</sup>.

Las CIT o CTC fueron identificadas por primera vez en leucemias y sucesivamente han sido descritas en varios tipos de tumores sólidos, como en el de mama, ovario, cerebro, próstata, endometrio, colon, pulmón, páncreas, hígado, cabeza y cuello, piel, entre otros<sup>2,4–11,13,14,17,21,22,25,27–30,32,36–39,42,45,56,57,60,62–66,68–85</sup>. En el 2006, en un taller de CIT de la AACR (“American Association for Cancer Research”) definieron a este grupo celular como aquellas células tumorales con capacidad de autorenovarse y generar los linajes heterogéneos de células cancerosas que forman el tumor<sup>8</sup>. Por sus características y sus similitudes con las células troncales normales, las CIT han despertado un gran interés en el campo de estudio del cáncer y se ha propuesto que este grupo celular podría ser un nuevo blanco terapéutico que pudiera detener y eliminar la progresión, proliferación, resistencia terapéutica, reincidencia y metástasis tumoral<sup>1,3–6,8,14,18–22,25,30,36,45,47,49,55,59–66,73–75</sup>.

---

Para poder comprender mejor a las células iniciadoras tumorales, hemos querido enfocar este estudio en tratar de conocer la estructura genómica y transcripcional de este grupo celular, específicamente derivadas de cáncer de mama, no sólo por el gran impacto a nivel de salud pública que tiene este tipo de cáncer en la población femenina de nuestro país, sino que también para poder contar en un futuro con un enfoque global que permita lograr la prevención como el desarrollo de terapias efectivas contra este tipo de enfermedad.

## **2. CÉLULAS INICIADORAS TUMORALES**

### **2.1 ANTECEDENTES HISTÓRICOS**

Los primeros indicios de la existencia de las CIT surgieron desde 1855 cuando Rudolf Virchow planteó que el cáncer se generaba por la activación de células dominantes, estas eran células embrionarias remanentes que se encontraban en tejidos maduros. La hipótesis se conoce como “The embryonal rest hypothesis” o teoría de las células embrionarias durmientes y/o residuales<sup>62,64,76</sup>. Posteriormente esta teoría fue ampliada por su discípulo Julius Cohnheim, quien en 1875 propuso que las alteraciones de las células troncales durante el desarrollo embrionario podrían en un futuro conducir al desarrollo tumoral<sup>5,14,29,60,62–64,67,76,77</sup>. En 1988 Pierce y Speers establecen el concepto de que los carcinomas son caricaturas del proceso de renovación tisular compuestos por células troncales malignas<sup>62,78</sup>. Fue hasta 1994 cuando el Dr. John Dick y sus colaboradores identificaron por primera vez a las CIT en poblaciones purificadas de leucemias mieloides agudas humanas. Este grupo celular que expresaba marcadores de superficie específicos y parecidos a los de las células troncales hematopoyéticos (CD34<sup>+</sup>CD38<sup>-</sup>) podía formar tumores eficientemente posterior a su inoculación en ratones “SCID” (Inmunodeficiencia Combinada Severa) mientras que el resto de la población celular derivado del mismo tumor no<sup>4,5,7,9,17,19–22,30–32,36,37,39,60,62,66,79,82–86</sup>. Desde entonces y utilizando el mismo modelo

---

funcional, las células iniciadoras tumorales han sido descritas en varios tipos de tumores<sup>1,2,4,5,9,18–22,30,32,62–64,66,84,85</sup>.

## 2.2 CARACTERÍSTICAS Y PROPIEDADES DE LAS CIT

Las CIT o CTC han sido nombradas así debido a las similitudes que comparten con las células troncales normales (CT). Estas propiedades operan mediante vías de señalización moleculares parecidas aunque estas se encuentran alteradas en las células iniciadoras tumorales. La capacidad de autorenovarse indefinidamente (mediante división simétrica o asimétrica) y la capacidad de diferenciarse son las dos características fundamentales de estos grupos celulares<sup>2,4–6,8–10,12,14,17,23,24,26,28,31,32,34,36–38,47,49,59–62,64,68,69,73–75,82–92</sup>.

La autorenovación les permite generar una célula con el mismo fenotipo troncal, en el caso de las CIT serán células que podrán mantener el crecimiento tumoral mientras que en las CT serán aquellas que mantendrán el equilibrio de la población troncal a través del tiempo, permanecerán en su propio compartimiento conocido como “the stem cell niche” o nicho de las células troncales (compuesto por adipocitos, células inmunológicas, células endoteliales, células troncales mesenquimales, fibroblastos entre otras) y que usualmente se encontrarán quiescentes (la mayoría del tiempo se encontrarán en fase G0) dividiéndose con poca frecuencia. La diferenciación en cambio, les concede el potencial de producir células que continúan con sus divisiones celulares y que finalmente se encontrarán comprometidas a un cierto linaje celular, conocidas como células progenitoras en el caso de las CT mientras que en las CIT serán aquellas que formarán la masa o “bulk” tumoral<sup>1,2,4–6,8,9,13–15,20,22,25,29,31,32,35,47,60–62,68,74,75,83,85,86,88–91</sup>.

El tipo de división que les permite generar una célula idéntica a ellas y una más diferenciada es conocida como división asimétrica y se piensa que es la más utilizada por

---

estos grupos celulares. Así mismo, la división simétrica que produce dos células con el mismo destino celular, es utilizada por las CT cuando necesitan generar un gran número de células, por ejemplo, en los casos de lesiones y durante el desarrollo<sup>1,2,4-6,8-10,60,68,74,82,84,86,88,89</sup>.

Los dos tipos de división son utilizadas por las CT, cada una siendo requerida circunstancialmente y con el fin de mantener la homeostasis tisular, sin embargo hoy en día no se sabe si existe un tipo de división preferencial. Ya que el cáncer es considerado una enfermedad que cuenta con tasas de proliferación celular alteradas es necesario contar con una mejor comprensión de los procesos de autorenovación y diferenciación celular de las CT para así poder entender la regulación de estos procesos en las CIT<sup>1,4-6,68,86,88,89</sup>.

Otras características que comparten las CIT y las CT son que generalmente representan una población numéricamente limitada en comparación con los demás grupos celulares, expresan altas tasas de transportadores ABC (conocidos también como transportadores dependientes de ATP) que comúnmente se encuentran involucrados en el transporte al exterior de la célula de pequeñas moléculas y compuestos y que también tienen actividad de resistencia a fármacos como lo son ABCG2 (BCRP1) y ABCB1 (glicoproteína P o MDR1), que explica en las CIT la resistencia terapéutica pero por otro lado esta propiedad ha permitido purificar tanto a las CT como a las CIT mediante citometría de flujo con el fenotipo conocido hoy en día como "Side Population" (SP), ya que estas bombas de eflujo sacan el colorante Hoechst 33342 y Rodamina 123 de la célula permitiendo así su aislamiento y reconocimiento<sup>4,5,8,13,14,17,18,20,22-24,27,29,31,34,35,38,60,62,64,66,74,75,82,84,85,90,92</sup>. Las propiedades ya mencionadas de quiescencia relativa (baja tasa de división celular) y la resistencia a agentes tóxicos junto con la alta actividad de reparación de ADN, inducción de angiogénesis, alta actividad de telomerasa, telómeros largos y la resistencia a apoptosis (características que también presentan), les confiere a estas dos poblaciones un mayor índice de supervivencia<sup>4,5,13,20,22,30,31,34,47,60,64,74,75,90</sup>.

Así mismo, las CIT expresan marcadores típicos de pluripotencia presentes en las CT adultas, como lo son Oct4, Sox1, Sox2, Nanog, Fgfr1, CD133, c-kit, ALDH1, PROCR (receptor para la proteína C, también conocido como EPCR o CD201 e involucrado en la vía de la coagulación sanguínea), así como también expresan el receptor de quimiocinas CXCR4, el cual promueve la migración celular a tejidos que producen niveles altos de su quimiocina ligando CXCL12 (conocido también como SDF1), como lo son pulmón, hígado y médula ósea (sitios frecuentes de metástasis)<sup>1,4,5,8,12,16,18,20,26–28,32,34–36,38,45,47,60,64,66,71,75,84,92–95</sup>. Las vías de señalización Hedgehog (Hh), Wnt, TGF- $\beta$ , Hippo, TNF- $\alpha$ /NF- $\kappa$  $\beta$ , JAK/STAT y Notch que hoy en día se sabe que juegan un papel importante en la regulación de los procesos de autorenovación y diferenciación en las CT adultas y embrionarias también han sido observadas en las CTC, aunque con una activación y funcionamiento anormal (Figura 3)<sup>1,2,4,5,8,9,11,12,14,20,23,28,31,34,35,38,47,59,60,62,64,66,68,69,74,75,82,83,86,91,92</sup>. A continuación y para finalizar este apartado (Tabla 1) se resaltan las propiedades fundamentales de las CIT.

**Tabla 1. Características principales de las CIT.**

CARACTERÍSTICAS PRINCIPALES DE LAS CIT
Autorenovación
Diferenciación
División Asimétrica
Representan una población limitada dentro del tumor
Perfil de marcadores de superficie característico
Capacidad de formación tumoral (fenocopia del tumor original) en ratones inmunocomprometidos (estándar de oro).
Alta resistencia terapéutica
Regeneración ilimitada
Metástasis
Recurrencia

---

## 2.3 ORIGEN

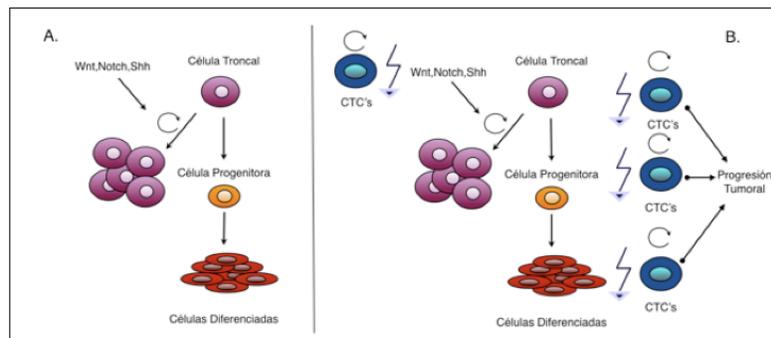
A pesar de que el origen de las CIT sigue siendo hoy en día incierto se han planteado varias hipótesis, la primera es que pudieran surgir de CT adultas, estas células que se encuentran en varios tejidos y que cuentan con muchas propiedades similares a las de las CIT, poseen un alto índice de supervivencia lo que les confiere la capacidad de acumular mutaciones genéticas y modificaciones epigenéticas suficientes para transformarse en una célula cancerígena. La segunda es que pudieran originarse a partir de células más diferenciadas como lo son las progenitoras, la tercera es que pudieran derivarse de células terminalmente diferenciadas y/o progenitoras maduras que adquieren propiedades de CT posterior a su transformación y dediferenciación. Por último, la cuarta hipótesis propone que algunos tipos de cáncer como los teratocarcinomas o sarcomas pediátricos se originan a partir de células parecidas a las CT embrionarias que se quedan presentes de forma anormal en tejidos durante la ontogénesis.

Para que se pueda llevar a cabo esta transformación (CT → CIT) son necesarios varios factores, entre ellos cambios en el tipo de división celular, en el ciclo celular, en el microambiente troncal, en la señalización celular, entre otros. Así mismo, se cree que la transición epitelio-mesénquima (TEM), proceso en el cual las células epiteliales adquieren un aumento de motilidad e invasividad (ya que se pierden los contactos célula-célula, la polaridad celular y la reorganización del citoesqueleto) por otro lado, se disminuye la expresión de algunas citoqueratinas epiteliales y de la E-caderina y se incrementa la expresión de marcadores mesenquimales (como de N-caderina, fibronectina y vimentina) así como también de los factores de transcripción FOXC2, ZEB1, TWIST1, TWIST2, Slug y Snail, todo este proceso en conjunto con la desregulación de las vías de autorenovación (Hh, Wnt, TGF- $\beta$ , TNF- $\alpha$ /NF- $\kappa\beta$ , Hippo, JAK/STAT y Notch) (Figura 3), mencionadas ya anteriormente juegan un papel importante en la transformación de las células progenitoras y

---

células terminalmente diferenciadas a CIT<sup>2-6,8-11,14,16,18-22,32,35,36,39,45,47,49,60,61,64,66,72,75,89,90,92,96,97</sup>.

**Figura 3. Autorenovación y diferenciación de CT y CIT.**



**A.** Proceso de autorenovación y diferenciación de las CT; observamos algunas de las vías de señalización que juegan un papel importante en la autorenovación de las CT. **B.** Alteraciones en el proceso de autorenovación y/o diferenciación de las CT que pueden dar origen a CTC's (CIT). (Modificada de Schwarz and Meléndez; 2011)<sup>5</sup>.

Por otro lado y como he comentado anteriormente la acumulación de modificaciones genéticas (mutaciones en genes supresores de tumores, oncogenes, entre otros) como las alteraciones epigenéticas (modificaciones de histonas, metilación anormal, etc) pueden participar en esta transformación. La pérdida del supresor tumoral Pten y sobre-regulación de la actividad de  $\beta$ -catenina han sido reportados como pasos fundamentales para esta conversión<sup>3,4,8,19,36,45,64,83</sup>.

## 2.4 IDENTIFICACIÓN Y AISLAMIENTO DE LAS CIT

Las CIT representan una población pequeña y limitada dentro de la mayoría de los tumores, por lo que se requiere de una separación y purificación precisa para poder lograr su identificación y con ello su análisis y estudio. El perfil o fenotipo de marcadores que

permite su reconocimiento varía dependiendo del tipo tumoral y hoy en día gracias a los avances realizados en este campo se han identificado diversos marcadores comunes (presentes en CIT de diversos tipos de cáncer) tanto de superficie como intracelulares, como es el caso de CD133, CD44, ALDH, YAP1, ABCG2, los factores de transcripción involucrados con troncalidad: Nanog, Oct4 y Sox2, entre otros; así como también se han descrito marcadores de superficie de un tipo tumoral en específico (Tabla 2)<sup>2,4,5,8-10,12-15,25-27,29-39,45,47-49,60,62,65,69,72,73,75,81-85,87,93,95,96,98</sup>. Más adelante describiré algunos de los marcadores que se han utilizado para enriquecer a las CIT del cáncer de mama, que son las de nuestro particular interés.

**Tabla 2. Marcadores de CIT en diversos tipos tumorales.**

Tipo de Cáncer	Fenotipo de Marcadores de CIT
Colon	CD133 <sup>+</sup> , CD44 <sup>+</sup> , CD166 <sup>+</sup> , EpCam <sup>+</sup> , CD24 <sup>+</sup> , CXCR4 <sup>+</sup> , CK20 <sup>+</sup> , CEA <sup>+</sup> , LGR5 <sup>+</sup>
Cabeza y cuello	CD44 <sup>+</sup> , ALDH <sup>+</sup> , YAP1 <sup>+</sup> , BMI1 <sup>+</sup>
Leucemia	CD34 <sup>+</sup> , CD38 <sup>-</sup> , HLA-DR <sup>-</sup> , CD71 <sup>-</sup> , CD90 <sup>-</sup> , CD117 <sup>-</sup> , CD123 <sup>+</sup>
Mama	ESA <sup>+</sup> , CD44 <sup>+</sup> , CD24 <sup>-low</sup> , Lineage <sup>-</sup> , ALDH1 <sup>high</sup>
Hígado	CD133 <sup>+</sup> , CD49f <sup>+</sup> , CD90 <sup>+</sup>
Cerebral	CD133 <sup>+</sup> , BCRP1 <sup>+</sup> , A2B5 <sup>+</sup> , SSEA1 <sup>+</sup>
Pulmón	CD133 <sup>+</sup> , ABCG2 <sup>high</sup>
Mieloma múltiple	CD138 <sup>-</sup>
Próstata	CD44 <sup>+</sup> , α2β1 <sup>high</sup> , CD133 <sup>+</sup>
Páncreas	CD133 <sup>+</sup> , CD44 <sup>+</sup> , EpCAM <sup>+</sup> , CD24 <sup>+</sup> , ABCG2 <sup>high</sup>

Modificado de Ajani *et al.*, 2015<sup>8</sup>

En la actualidad son dos los ensayos que se utilizan con mayor frecuencia para lograr el aislamiento y enriquecimiento de las CIT y son en base a su inmunofenotipo. El primero es mediante citometría de flujo con el método de FACS (Clasificación Celular Activada por Fluorescencia) y el segundo es vía perlas paramagnéticas unidas a anticuerpos (como son las microesferas MACS<sup>®</sup> de Miltenyi Biotec)<sup>2,5,8,9,20,25,32,34,39,66,83,84,92</sup>.

---

Existen otros ensayos que por medio de la citometría de flujo también permiten el enriquecimiento de estas poblaciones, como por ejemplo vía el fenotipo SP (mencionado ya anteriormente), el ensayo de ALDEFLUOR™ que mide la actividad de la enzima ALDH (involucrada en la oxidación de aldehídos intracelulares a ácidos carboxílicos y ácido retinóico así como también está implicada en la biosíntesis del ácido  $\gamma$ -amino butírico (GABA) y en la diferenciación de CT y CIT vía el ácido retinóico) y que en ciertas CIT se ha observado sobre-expresada y por último un ensayo de retención del colorante bromodesoxiuridina (BrDU), que se basa en el marcaje del ADN (ácido desoxirribonucleico) recién sintetizado utilizando este análogo de timidina<sup>4,8,9,17,20,23,24,26,29,31,32,34,35,38,47,60,62,65,66,75,82-86,91-93</sup>.

La caracterización funcional de las CIT se ha llevado a cabo *in vitro* mediante el ensayo de formación de colonias en agar suave y el ensayo de formación de esferas en cultivos líquidos (inicialmente utilizado para detectar CT neuronales y posteriormente para descubrir a las CT de mama y CIT de cáncer de mama). Estas dos técnicas permiten valorar la capacidad de autorenovación que posee este grupo celular mientras que la prueba y el estándar de oro es la demostración de la capacidad de formación tumoral *in vivo*, la cual en el caso de células humanas se lleva a cabo mediante un xenotransplante (donde por ejemplo un número limitado de CIT son inoculadas en ratones inmunocomprometidos)<sup>2,4,5,8,9,11,17,20,25,29,30,32,34,37,39,48,49,65,66,70,72,73,75,84,91-93</sup>.

## 2.5. TERAPÉUTICA

Hoy en día se ha propuesto que para poder lograr una erradicación efectiva del cáncer se deberán de eliminar específicamente a las CIT como a las poblaciones tumorales más diferenciadas, e idealmente, sin dañar a las CT ni a las células parenquimatosas, lo cual requiere de un amplio y específico conocimiento de los perfiles moleculares e inmunológicos de estos grupos celulares, así como de una combinación de tratamientos

---

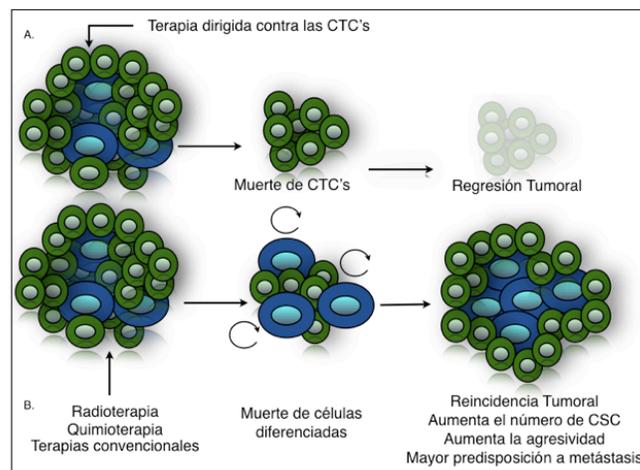
que incluya terapia dirigida y convencional<sup>1,3-6,8,13,18-20,22,25,30,38,45,60,64,66,83,86,93</sup>. Los tratamientos convencionales utilizados en la actualidad (inmuno, quimio y radioterapia) eliminan a las células tumorales diferenciadas, reduciendo así la masa tumoral pero dejando a las CIT vivas, lo cual termina en recurrencia y proliferación tumoral así como también en la generación de células tumorales más agresivas y resistentes a terapia oncológica (Figura 4)<sup>4-6,8,13,20,27,30,32,38,39,64,66,70,75</sup>. La radioresistencia de las CIT ha sido demostrada y reportada ya por varios investigadores y han observado una asociación de este fenómeno con niveles bajos de especies reactivas de oxígeno (ERO o ROS, por sus siglas en inglés: “Reactive Oxygen Species”) con la subsecuente disminución de formación de rupturas de la doble cadena de ADN<sup>5,8,20,24,35,47,69,70</sup>. Las CIT poseen varias características mencionadas ya anteriormente, que les confieren la capacidad de quimio y radio resistencia así como de resistencia a apoptosis, como lo son: su gran capacidad de reparar el ADN, la sobre-expresión de proteínas anti-apoptóticas (BCL-2 y survivina), lenta división celular que les permite escapar de las radio y quimioterapias tradicionales que van dirigidas a células que se multiplican rápidamente y la sobre-expresión de bombas de eflujo como lo son ABCG2 y ABCB1<sup>4,5,13,17,22,24,27,31,34,47,64,74,82,85</sup>.

Todo esto ha llevado a los investigadores a replantearse el potencial de la terapéutica actual, dirigiéndolos y fomentándolos a la búsqueda y desarrollo de nuevos blancos terapéuticos y tratamientos, para así poder lograr la eliminación no solamente de las CIT y células tumorales más diferenciadas, sino que también, poder lograr un bloqueo de las las vías implicadas en el mantenimiento de los tumores, reduciendo así la probabilidad de progresión y regresión tumoral<sup>8,38,47</sup>. Actualmente, ya se conocen varios compuestos que pueden llegar a afectar ciertas vías de señalización (y con ello bloquearlas y/o inducir la diferenciación de las CIT) como lo son ciclopamina y vismodegib (dirigidas en contra de la vía Hh), inhibidores de la  $\gamma$ -secretasa (en contra de la vía de Notch) y pamoato de pirvinio, IWR-1 y XAV939 dirigidos en contra de la vía de Wnt/ $\beta$ -catenina. A su vez, se están llevando a cabo varios ensayos que se encuentran en diferentes fases de estudio que

---

involucran inhibidores de la vía de Wnt/ $\beta$ -catenina como lo son NCT02178956, NCT02279719, NCT02231723, NCT01776307, NCT02069145, NCT02092363, NCT02050178 y también inhibidores de la vía Notch (NCT01647828, NCT01859741, NCT01277146, NCT01957007, NCT02259582, etc)<sup>4,5,8,20,38,45,47,64,66,74,82</sup>.

**Figura 4. Terapias antitumorales.**



**A.** Régimen terapéutico contra las CTC's (CIT). **B.** Terapias actuales anti-tumorales. (Modificada de Schwarz and Meléndez; 2011)<sup>5</sup>.

Por otro lado se ha observado que la fumitremorgina C (FTC), un producto natural, inhibe a la bomba de eflujo ABCG2, sin embargo se ha observado un cierto grado de toxicidad en células humanas y murinas lo cual no lo hace apto para futuros ensayos clínicos, pero abre una posibilidad de un nuevo agente terapéutico como lo es hoy en día el desarrollo de Ko143, un derivado químico de FTC que ha logrado disminuir la toxicidad y aumentar su especificidad<sup>5,35,38,90</sup>. Se ha observado también, que la biguanida metformina (el agente oral más utilizado para tratar diabetes mellitus tipo 2) y el ácido retinóico (análogo de la vitamina A) pueden llegar a bloquear la autorenovación e inhibir la proliferación de las

---

CIT, demostrando así una función anti-tumoral que habrá que valorar para futuras combinaciones de tratamientos<sup>4,20,64,66,74,75,82</sup>.

Más adelante profundizaremos un poco más en las implicaciones terapéuticas actuales de las CIT del cáncer de mamá, sin embargo es claro que se requiere identificar y caracterizar a este grupo celular de ser posible en la mayoría de los tumores malignos, esto proporcionará no solamente nuevos enfoques y estrategias que logren abatir a las CIT y no CIT de forma integral, sino que también ayudará al reconocimiento de un marcador universal<sup>9,14,20,75</sup>.

### 3. CÁNCER DE MAMA

El cáncer de mama es una enfermedad heterogénea, que engloba un conjunto de enfermedades con diferentes características biológicas y patológicas, distintas presentaciones, comportamientos y desenlaces clínicos, así como también diversas respuestas a tratamientos<sup>8,9,14,15,20,24,33,46,49,51,62,71,72,81,82,87,98–107</sup>. Esta heterogeneidad también se ve reflejada en las diversas clasificaciones que existen hoy en día para su estadificación, comprensión, tratamiento y estudio, desde la clásica histopatológica (la Organización Mundial de la Salud (OMS) ha reportado la existencia de hasta 21 tipos histológicos donde el carcinoma ductal invasivo es el tipo histológico invasivo más común y el carcinoma ductal *in situ* el tipo histológico no invasivo más común) que toma en cuenta la morfología celular, crecimiento y patrones arquitectónicos celulares, la clasificación inmunohistoquímica (IHQ) y la más reciente y moderna la clasificación molecular<sup>20,51,72,84,99,100,108</sup>.

Tomando en cuenta el análisis IHQ, el cáncer de mama es separado en tres grupos principales: receptores hormonales positivos (RH<sup>+</sup>, específicamente a estrógenos (RE) y progesterona (RP)), positivos al receptor 2 del factor de crecimiento epidérmico humano (“HER2<sup>+</sup>”, por sus siglas en inglés “Human Epidermal Growth Factor Receptor 2”) y triples

---

negativos (TN) (RH<sup>-</sup> y HER2<sup>-</sup>)<sup>9,15,24,46,51,81,82,99,105,109</sup>. El cáncer de mama con receptores hormonales positivos (RE<sup>+</sup>/ RP<sup>+</sup>) representa el 70% de los casos y las mujeres con este tipo de cáncer tiende a presentar un mejor pronóstico. El subtipo HER2<sup>+</sup> representa el 20% y cursa con un pronóstico desfavorable y un fenotipo más agresivo mientras que el TN constituye del 15-20% de los casos de cáncer de mama y cuenta con un curso agresivo y un pronóstico pobre<sup>4,9</sup>. Posteriormente a este panel IHQ se han sumado Ki67 (marcador de proliferación ki-67) la citoqueratina 5/6 (CK5/6) o citoqueratina 5 (CK5) y el receptor del factor de crecimiento epidérmico (EGFR) para determinar el índice de proliferación celular. La utilización y búsqueda de los biomarcadores mencionados ha sido de gran utilidad sobretodo en los casos o poblaciones donde no se cuenta con el material necesario para desarrollar alguna metodología que logre analizar los perfiles de expresión génica<sup>24,51,81,91,110</sup>. Las opciones terapéuticas van desde hormonoterapia para aquellos tumores RH<sup>+</sup> (moduladores selectivos de los receptores de estrógenos como el tamoxifeno y raloxifeno, inhibidores de la aromatasa entre otros), inhibidores de HER2 (trastuzumab, lapatinib) para los tumores que sobre-expresan HER2 en conjunto con quimioterapia y esta última opción para los tumores considerados TN<sup>4,30,31,51,56,82,93,99,109,111</sup>.

Estudios recientes (utilizando microarreglos de expresión) han permitido el desarrollo de una nueva clasificación molecular del cáncer de mama basada en diferentes perfiles de expresión y cada una con diferentes pronósticos y respuestas a tratamientos, los fenotipos reportados a la fecha incluyen: *luminal* (subtipo A y B, ambos son RH<sup>+</sup> pero el A es HER2<sup>-</sup> mientras que el B es HER2<sup>+</sup>, se caracterizan por expresar genes que se expresan en la luz de conductos mamarios como las citoqueratinas 8, 18 y 19 así como también se encuentran expresados los factores de transcripción ER, GATA3, FOXA1 y XBP1, el subtipo A no solo es la variante más común, sino que también se encuentra asociado con una mayor tasa de sobrevida, sobretodo en los primeros 5 años posterior al diagnóstico), *enriquecidos en HER2* (que expresan niveles altos de HER2 y son RH<sup>-</sup>), *parecido al normal*

---

(“normal like”, expresan varios genes propios del epitelio mamario normal, cuentan con alto contenido de estroma e infiltración linfocitaria, y son RH<sup>-</sup> y HER2<sup>-</sup>) y por último el *grupo basal*, “basal like” o TN ya que son RH<sup>-</sup> y HER2<sup>-</sup>, así como también se caracterizan por expresar genes propios de las células basales mamarias como lo son las citoqueratinas 5, 6, 14 y 17 y expresar niveles altos de marcadores de la TEM (Tabla 3)<sup>4,9,10,14,15,20,22,24,33,46,49,51,53,56,62,71,72,82,84,87,96–103,105–107,110–112</sup>. Se ha observado una asociación entre las mujeres portadoras de una mutación en el gen BRCA1 (“Breast Cancer 1” o cáncer de mama 1) y este subtipo molecular, proponiéndose como factor de riesgo para desarrollar este fenotipo molecular de cáncer de mama<sup>4,15,38,51,62,82,101,103,111,112</sup>. Los subtipos enriquecidos en HER2 y basal cuentan con una menor tasa de supervivencia (en específico el TN el cual es considerado el más agresivo y con peor pronóstico) en comparación con el subtipo luminal (el A cuenta con el pronóstico más favorable)<sup>4,15,22,24,33,51,53,71,82,87,96,100,101,103,107,111</sup>. Posteriormente, en el año 2007, se identificó un nuevo subtipo molecular denominado *bajo en claudina*, caracterizado por presentar una baja o ausente expresión de marcadores de diferenciación luminal, una alta expresión de genes mesenquimales incluyendo Vimentina, Snail1, Snail2, Twist1 y del factor de transcripción FOXC2 (proponiendo así una relación con la TEM), así como también se encuentra enriquecido con ciertas características de las CIT (inmunofenotipo CD44<sup>+</sup>CD24<sup>-</sup>, alta expresión de ALDH1, entre otras). Los tumores que caen dentro de este grupo tienen un pronóstico pobre y generalmente son RH<sup>-</sup> y HER2<sup>-</sup>, este subtipo es el menos aceptado y utilizado en la práctica en comparación con los mencionados anteriormente<sup>9,10,51,62,72,96–98,101,107,112</sup>.

El desarrollo y los avances de estas técnicas moleculares han logrado revelar los perfiles de expresión génica del cáncer de mama y estos a su vez nos han llevado a tener hoy en día una mejor comprensión de la heterogeneidad molecular e histológica que la caracteriza. La correlación clínica, radiológica e histopatológica en conjunto con la implementación de estas técnicas han permitido pronosticar el curso clínico y biológico de la

enfermedad así como también proporcionar un tratamiento individualizado y multidisciplinario al paciente para su mayor beneficio<sup>20,51,102,103,111</sup>.

**Tabla 3. Clasificación molecular del cáncer de mama.**

SUBTIPO	RE/RP (RH)	HER2	CARACTERÍSTICAS
Luminal A	RE <sup>+</sup> y/o RP <sup>+/-</sup>	-	Subtipo más común y con mejor pronóstico. Responden a la hormonoterapia. CK8 <sup>+</sup> , CK18 <sup>+</sup> , CK19 <sup>+</sup> .
Luminal B	RE <sup>+</sup> y/o RP <sup>+/-</sup>	+	Parecido al luminal A pero con peor pronóstico. Es más frecuente RE <sup>+</sup> RP <sup>-</sup> . CK8 <sup>+</sup> , CK18 <sup>+</sup> , CK19 <sup>+</sup> .
Enriquecido en HER2 o HER2 <sup>+</sup>	RE <sup>-</sup> RP <sup>-</sup>	+	Aproximadamente el 20% de los subtipos del cáncer de mama. Pronóstico desfavorable. Alrededor del 50% responde a trastuzumab.
Parecido al normal o "normal like"	RE <sup>-</sup> RP <sup>-</sup>	-	CK5/6 <sup>-</sup> , CK14 <sup>-</sup> y/o EGFR <sup>-</sup> . Expresa genes propios del epitelio mamario normal.
Triple negativo o "basal like"	RE <sup>-</sup> RP <sup>-</sup>	-	Subtipo más agresivo. Expresan niveles altos de TEM. Asociado a la mutación BRCA1. CK5/6 <sup>+</sup> , CK14 <sup>+</sup> , CK17 <sup>+</sup> , y/o EGFR <sup>+</sup> y/o c-kit <sup>+</sup> .
Bajo en claudina	RE <sup>-</sup> RP <sup>-</sup>	-	Subtipo menos utilizado y aceptado. Corresponde entre el 12 al 14% de los subtipos del cáncer de mama. Pronóstico pobre. Enriquecidos con características de las de las CIT y TEM. Los carcinomas metaplásicos mamarios tienden a presentar las características de este subtipo.

## 4. CÉLULAS INICIADORAS TUMORALES DEL CÁNCER DE MAMA

### 4.1. ANTECEDENTES HISTÓRICOS, CARACTERÍSTICAS Y PROPIEDADES DE LAS CITCM

Las células iniciadoras tumorales de cáncer de mama (CITCM) fueron descritas por primera vez en el año 2003 por Al-Hajj y sus colaboradores, y así mismo, estas fueron las primeras CIT aisladas a partir de un tumor sólido. Ellos separaron mediante FACS células

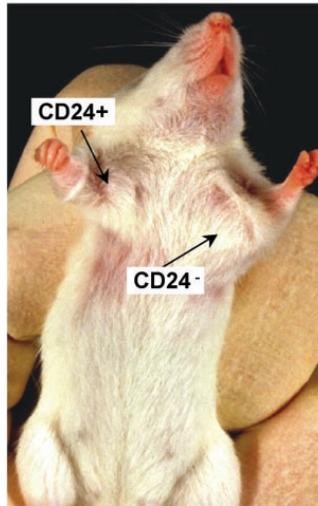
---

humanas con el fenotipo CD44<sup>+</sup>CD24<sup>-/low</sup>Lin<sup>-</sup> provenientes de tumores primarios y de sitios metastásicos de mama (efusiones pleurales) y observaron que consistentemente después de inyectar un mínimo de 1,000 células con este fenotipo en ratones NOD/SCID (Diabéticos No Obesos / Inmunodeficiencia Combinada Severa) podían generar tumores mientras que miles de células con otros fenotipos no (Figura 5). Así mismo, no solo lograron obtener un mayor enriquecimiento de la población CITCM al añadir EpCam (molécula de adhesión de células epiteliales, conocida también como CD326 o ESA) a su combinación de marcadores y generar tumores con tan solo 200 células con este fenotipo (CD44<sup>+</sup>CD24<sup>-/low</sup>Lin<sup>-</sup>EpCam<sup>+</sup>); sino que también pudieron comprobar el potencial de autorenovación y diferenciación de estas células mediante pases seriados<sup>2,4,6-10,15,20,22-24,30-35,38,45,47,49,59,60,62,63,68-73,75,82-85,87,91-93,101,112</sup>.

A partir de este estudio las CITCM han sido estudiadas y caracterizadas por varios grupos de investigadores con el fin de comprender mejor el comportamiento, la complejidad y la estructura heterogénea tumoral de este tipo de cáncer. Las CITCM al igual que las CIT y las CT, representan una pequeña y limitada población celular (0.1-1%), cuentan con la capacidad de autorenovación (demostrada funcionalmente al generar tumores posterior a su inoculación en ratones) y diferenciación celular (producen células dentro de la masa tumoral con y sin características de CT) y se han propuesto como las responsables de la resistencia terapéutica, metástasis y relapso tumoral<sup>4,7,9,10,14,15,20,23,30,35,45,47,69,72,75,82,83,85,93,101</sup>.

---

**Figura 5.** *Formación tumoral en ratones NOD/SCID.*



Tumor representativo en un ratón, en el sitio de inyección  $CD44^+CD24^{-/low}Lin^-$ , pero no en el sitio de inyección de  $CD44^+CD24^+Lin^-$ . (Tomado de Al-Hajj *et al.*, 2003)<sup>7</sup>.

#### **4.2 ORIGEN Y AISLAMIENTO DE LAS CITCM**

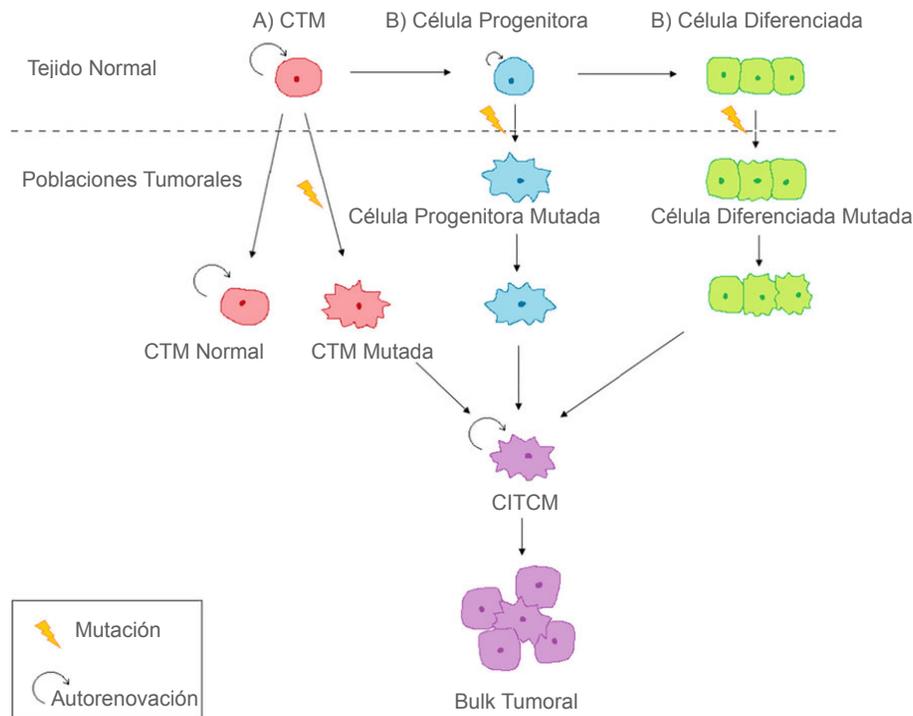
La glándula mamaria es un órgano dinámico que experimenta cambios constantes durante los ciclos de la vida en respuesta a la menarquia, embarazo, lactancia y menopausia. Está compuesta por una serie de ramificaciones ductales y unidades lóbulo-alveolares que a su vez se encuentran rodeadas por tejido graso. Cuenta con dos tipos de células epiteliales (luminales y mioepiteliales), dos tipos de estroma (inter e intra lobular) y dos estructuras principales (ductos largos y unidades lobulares ductal terminales). Las células mioepiteliales (conocidas también como basales) rodean a las células luminales y están en contacto con la membrana basal circundante que separa el parénquima del estroma tisular. Expresan el antígeno común de la leucemia linfoblástica aguda (CALLA o CD10), la  $\alpha$ -actina de músculo liso (SMA), vimentina, Thy-1(antígeno de diferenciación de timocitos, también conocido como CD90), citoqueratina 5 (CK5) y 14 (CK14) entre otras. La mama durante el embarazo y la lactancia presenta un aumento en el crecimiento celular y

---

formación de alveolos funcionales secretores de leche, esta última dirigida por el linaje celular luminal (así mismo, alínean las estructuras ductales que transportan la leche al pezón). Este grupo celular se caracteriza por expresar MUC1 (mucina 1), citoqueratinas 7 (CK7), 8 (CK8), 18(CK18) y 19 (CK19), mientras que EpCam se encuentra expresada en todas las células mamarias epiteliales<sup>4,10,38,39,45,59,60,72,91,92,106,112,113</sup>.

En el epitelio mamario existe una jerarquía celular de células troncales mamarias (CTM), células progenitoras y células diferenciadas que es de gran importancia para el desarrollo y la función normal de la mama y a su vez, puede ser fundamental para la carcinogénesis<sup>4,39,45,60,63,72,92,112</sup>. Las células troncales de la mama han sido aisladas durante años utilizando marcadores de superficie (como por ejemplo CD29<sup>high</sup>CD24<sup>+</sup> o CD49F<sup>high</sup>CD24<sup>+</sup>), los cuales (junto con sus perfiles de expresión) comparten con las células basales y por ello se piensa que las CTM radican en el compartimiento basal de la mama. Han observado sus capacidades de autorenovación y diferenciación mediante transplantes seriados y transplantes a “almohadillas” de tejido graso mamario (en donde demostraron que pueden reconstruir toda la glándula generando células con su mismo fenotipo así como células más diferenciadas)<sup>4,10,15,32,39,60,72,91</sup>. Estas mismas cualidades, como se mencionó anteriormente, y en conjunto con sus características fenotípicas y de marcadores de superficie similares, las comparten con las CIT y CITCM, y por ello, se han propuesto varias teorías (descritas previamente) en cuanto al origen de este grupo celular, en donde las CT (en este caso CTM), células progenitoras o células más diferenciadas de la mama podrían ser un modelo potencial para la formación de las células troncales del cáncer de mama (Figura 6). Sin embargo a la fecha, el origen real de las CITCM sigue siendo incierto<sup>4,20,23,35,45,47,60</sup>.

**Figura 6. Modelos de formación de las CITCM.**



**A.** Modelo 1: Las CITCM se originan a partir de las CTM posterior a la acumulación de mutaciones durante su estado quiescente e inicia la transformación tumoral. **B.** Modelo 2: Las CITCM se originan a partir de las células progenitoras posterior a la acumulación de mutaciones a este nivel. **C.** Modelo 3: Las CITCM se originan a partir de células mamarias diferenciadas las cuales sufren una desdiferenciación y adquieren características troncales. (Modificada de Sin and Lim; 2017)<sup>23</sup>

Las técnicas más utilizadas hoy en día para la detección y enriquecimiento de las CITCM (y que ya fueron mencionadas anteriormente) tanto de líneas celulares como de muestras tisulares de cáncer de mama y de sitios metastásicos son:

- Mediante su perfil inmunofenotípico (marcadores de superficie) utilizando citometría de flujo o perlas paramagnéticas.
- A través de SP, estas células sobre-expresan transportadores transmembranales (ABCG2, ABCB1) que funcionan como bombas de eflujo que expulsan de la

---

célula no solo fármacos sino que también colorantes lipofílicos como rodamina 123 y Hoechst 33342 permitiendo así su purificación mediante citometría de flujo.

- Utilizando el ensayo de ALDEFLUOR™.
- Usando la técnica de retención del colorante BrDU.
- Funcionalmente e *in vitro*, mediante los ensayos de formación de colonias en agar suave y formación de mamoesferas en cultivos líquidos utilizando placas de baja adherencia, e *in vivo*, mediante xenotransplantes con pases seriados, con los cuales se evalúa la capacidad de formación tumoral, la autorenovación y la diferenciación (desde la publicación de Al-Hajj y colaboradores, esta prueba es considerada el estándar de oro para caracterizar a las CITCM)<sup>4,9,14,16,17,20,24,30,32,34,38,39,47–49,62,65,69,73,75,82–85,92,93</sup>.

### 4.3. MARCADORES DE LAS CITCM

La identificación de biomarcadores ha sido y es de gran importancia para el aislamiento y definición de las CITCM, estos marcadores no solamente contribuyen a la interacción celular, sino que también les confiere propiedades únicas<sup>23,35</sup>. Desde el 2003, año en el cual las CITCM fueron descritas por primera vez, el fenotipo CD44<sup>+</sup>CD24<sup>-/low</sup>Lin<sup>-</sup> utilizado para este propósito, ha sido el más comúnmente utilizado para el enriquecimiento de esta población, ya sea con o sin la adición de EpCam<sup>+</sup><sup>4,7,14,15,20,23,30–32,34,35,38,39,45,47,49,60,62,75,82,83,85,93,96,98</sup>. Algunos otros marcadores usados para el aislamiento de las CITCM y que serán descritos posteriormente incluyen ALDH, CD133, claudinas e integrinas, PROCR, CXCR4, entre otros (para mayor información consultar Schwarz-Cruz y Celis *et al.*, 2015, en donde abundamos más sobre este tema en particular)<sup>4,9,14,17,20,23,30,32,34,35,38,39,45,47–49,62,65,71,75,82–85,87,96,98</sup>.

---

CD44 es una glicoproteína transmembranal tipo 1 que se expresa en algunos tipos celulares como el mesenquimal y del neuroectodermo y en varios tipos de cáncer (en los cuales su presencia se ha relacionado con un pronóstico pobre y resistencia terapéutica). Se une al ácido hialurónico (AH), laminina, colágeno y fibronectina (componentes de la matriz extracelular). En el tejido mamario normal y tumoral se ha localizado en las membranas basales de las células basales mioepiteliales y en algunas lumino-epiteliales. En el cáncer de mama juega un rol fundamental como molécula de adhesión, participa en las interacciones célula-célula, la proliferación celular, invasión, migración y angiogénesis. Han observado que la expresión de CD44 disminuye gradualmente durante la progresión tumoral así como su inhibición reduce la resistencia a agentes anti-tumorales<sup>4,7,9,10,13,15,16,20,23,24,34,35,47,59,62,63,71,81,84,106</sup>. CD24, también conocido como antígeno estable al calor (o HSA, por sus siglas en inglés “Heat Stable Antigen”) es una glicoproteína de superficie que incrementa la adhesión y promueve la proliferación y metástasis tumoral. En el tejido mamario sano se ha localizado en la membrana apical de las células luminales mientras que en muestras de cáncer de mama se ha encontrado principalmente en el citoplasma. Así mismo, se ha observado sobre-expresado en varios tipos de cáncer y se ha asociado con un curso de la enfermedad más agresivo<sup>7-10,15,20,23,35,47,59,62,71,98,106</sup>.

La expresión alta de CD24 o CD24<sup>high</sup> en el tejido mamario normal y tumoral corresponde a una firma de expresión más diferenciada, mientras que las células CD44<sup>+</sup> presentan un patrón de expresión parecidas al de las células troncales<sup>10,15,24,59,62,63</sup>. El fenotipo CD44<sup>+</sup>CD24<sup>-</sup> se ha observado con mayor frecuencia en los subtipos TN (“basal like”), en las líneas celulares y muestras tumorales que sufrieron una TEM y con menor frecuencia en los subtipos HER2<sup>+</sup> y luminales<sup>4,14,15,20,24,30,38,62,71,72,83,84,96,98</sup>. El uso de la combinación de la expresión de CD44 y CD24 (con o sin más marcadores) ha sido utilizada exitosamente para enriquecer las poblaciones de CITCM tanto de líneas celulares como de muestras tisulares, sin embargo existe mucha controversia con respecto a su tumorigenicidad, con lo cual resaltamos la importancia del uso de las pruebas funcionales (tanto *in vivo* como *in vitro*) para con ellas comprobar las características fundamentales de

---

esta población (autorenovación, diferenciación y potencial tumorigénico). El debate del uso de estas dos glicoproteínas llevó a la búsqueda de nuevos marcadores<sup>9,10,15,20,22,35,38,47,62,63,65,73,75,84,85,87,93</sup>. Tal es el caso de ALDH, su actividad ha sido identificada como un método para enriquecer tanto a las CTM, como también a las CITCM, esta subpoblación también cuenta con las capacidad de autorenovación y diferenciación, así como también puede restablecer la heterogeneidad del tumor original. Al combinarla con la expresión de CD44<sup>+</sup>CD24<sup>-</sup> se logra una fracción de CITCM más purificada, su expresión ha sido vinculada con un curso de la enfermedad pobre y con mayor resistencia a la quimioterapia<sup>4,10,14,20,22-24,30,32,35,38,45,47,48,60,62,65,71,75,82-85,87,93,98,112</sup>.

La prominina-1 o CD133 es un marcador utilizado para caracterizar a las CIT provenientes de varios tejidos tumorales (sangre, páncreas, colon, endometrio hígado, piel, próstata, mama, etc). La función exacta de este biomarcador en las células cancerígenas se desconoce, sin embargo, se sabe que está asociada con la unión del colesterol y por ello se piensa que se encuentra involucrada en la vía de señalización Hh siendo responsable de la diferenciación celular y de la TEM. Han demostrado que las células CD133<sup>+</sup> pueden formar mamoesferas y tumores en ratones NOD/SCID. También se ha observado en varios tumores sólidos, incluyendo el carcinoma mamario ductal invasivo subtipo TN y tumores BRCA-1<sup>4,9,13,14,20,23,33-35,38,45,47,62,75,84,112</sup>.

La glicoproteína transmembranal tipo 1 PROCR, es un marcador que también se ha utilizado aunque con menor frecuencia para enriquecer a las CITCM. Se ha localizado en la capa basal de conductos y alvéolos mamarios normales y se ha reportado altamente expresada en las células CD44<sup>+</sup> de cáncer de mama, así mismo, las células PROCR<sup>+</sup>ESA<sup>+</sup> derivadas de la línea celular MDA-MB-231 tienen una gran capacidad de formación de colonias en agar suave y un alto potencial tumorigénico (comprobado en ratones NOD/SCID) que las células PROCR<sup>-</sup>ESA<sup>-</sup> y que el resto de las poblaciones celulares, también

---

pueden dividirse asimétricamente y expresan niveles altos de NUMB y de marcadores de TEM como vimentina, SLUG2 y FOXC2<sup>20,45,59,84,98</sup>.

Las integrinas son glicoproteínas de superficie celular que funcionan como receptores para proteínas de la matriz extracelular. Estudios *in vivo* e *in vitro* han demostrado la presencia de las integrinas CD29 ( $\beta$ 1), CD49f ( $\alpha$ 6) y CD61 ( $\beta$ 3) solas y/o en conjunto con otros marcadores como CD24 y EpCam en el tejido mamario normal como tumoral y han introducido su uso (individual y combinado) para la purificación de las CITCM<sup>4,9,14,15,17,20,23,35,38,39,47,62,91,106,112</sup>. CD49f en el cáncer de mama está involucrada con la señalización y adhesión celular y ha sido asociada con un pronóstico pobre y una sobrevida corta, CD29 se encuentra involucrada también con la adhesión celular así como con la propagación metastásica de células tumorales mientras que CD61 con la iniciación tumoral<sup>4,9,17,23,38,62</sup>.

Por último, las claudinas (proteínas transmembranales asociadas con la formación de uniones estrechas entre las células epiteliales) 3, 4 y 7 han sido utilizadas junto a la ocludina y la e-cadherina (entre otras) para integrar un quinto subgrupo a la clasificación molecular de cáncer de mama, el “bajo en claudinas” (mencionado ya anteriormente) en el cual se han observado características de las CITCM y de la TEM; sin embargo estos marcadores son los menos utilizados y aprobados para el enriquecimiento de este grupo celular<sup>20,62,97,98,107</sup>.

Debido a la heterogeneidad intra e inter tumoral del cáncer, las CIT de diversos tipos tumorales cuentan con un perfil de expresión diferente, por ello la importancia de los estudios para la búsqueda de firmas moleculares los cuales contribuirán a la caracterización y separación subsecuente de las poblaciones de CITCM, por otra parte, el mejor entendimiento de los marcadores de CT expresados en el cáncer de mama pueden proveer

---

de una mejor comprensión de la biología y complejidad de las CITCM y esto permitirá el descubrimiento y desarrollo de nuevos blancos terapéuticos.

#### 4.4 CITCM y ncRNAs

En los últimos años han demostrado que los ARN no codificantes (ncRNAs) incluyendo a los ARNs largos no codificantes (lncRNAs) y a los micro ARNs conocidos como “miRNAs” (por sus siglas en inglés “microRNAs”) contribuyen a la carcinogénesis (mediante la regulación de la proliferación celular, invasión, metástasis y angiogénesis entre otros) así como también han reportado que estos pueden regular a las CIT de diversos tipos de cáncer. Los ncRNAs pueden dividirse según su tamaño en pequeños (menores a 200 nucleótidos (nt)) y largos (mayores a 200 nt). Dentro de los pequeños ncRNAs se encuentran los ARNs pequeños de interferencia “siRNAs” (19-23nt), “piRNAs” o ARNs asociados a Piwi (26-30nt) y los micro ARNs (miRNAs de 18-25nt). Los lncRNAs comprenden varios tipos de transcritos con una longitud desde cientos a miles de nt y se clasifican generalmente de acuerdo a su localización genómica relativa al gen codificante en: *sentido*, *antisentido*, *intrónico* e *intergénico*<sup>16,23,49,60,72,83</sup>.

En el año 2007, Yu y sus colaboradores reportaron por primera vez la actividad de un miRNA en la población de CITCM: let-7. Ellos observaron que este miRNA se encontraba regulado a la baja en este grupo celular y que este se encuentra involucrado en la regulación de varias propiedades de las CITCM como por ejemplo en la formación de mamoesferas y formación tumoral, en el mantenimiento de la proliferación e inhibición de la diferenciación celular, entre otros. A partir de este estudio, diferentes miRNAs han sido asociados como promotores o inhibidores de las CITCM, como por ejemplo la sobre regulación de miR-142, miR-21, miR-24, miR-29, miR-495, miR-181, miR-10b, miR-22, miR-221, miR-208a y miR-9 y la expresión baja de miR-200a, miR-200b, miR-200c, miR-93,

---

miR-100, miR-141, miR-34c, miR-30, miR-128, miR-140, miR-27a, miR-27b, miR-7 y miR-34a<sup>9,16,23,35,47,49,60,72,75,83</sup>.

A diferencia de los miRNAs que son altamente conservados y se encuentran generalmente implicados en la regulación negativa de la expresión génica a nivel post-transcripcional, los lncRNAs son pobremente conservados y pueden regular la expresión génica (positiva o negativamente) en varios niveles mediante diversos mecanismos. A la fecha un gran número de lncRNAs han sido identificados, sin embargo, solo una población limitada de estos han sido estudiados. Los lncRNAs HOTAIR, ROR, 00617, Hh y ATB han sido implicados con las CITCM al regular vías de señalización de la TEM. La progresión del estudio de los ncRNAs ayudará al mejor entendimiento de la regulación molecular de las CIT así como también al desarrollo de nuevos tratamientos tomando como blancos y factores pronóstico a los miRNAs y lncRNAs<sup>9,16,35,47,49,60,72,75,83</sup>.

#### **4.5 VÍAS DE SEÑALIZACIÓN EN LAS CITCM**

Los procesos de embriogénesis y organogénesis así como el mantenimiento de las CT y de la homeostasis tisular del adulto, son actividades muy complejas y altamente reguladas. Varias vías de señalización (principalmente Notch, Wnt y Hedgehog como mencioné en un inicio) juegan papeles clave en estos procesos, sus funciones e interferencias dirigen varias propiedades específicas de las CT (autorenovación y diferenciación) y cualquier alteración en estas vías en la glándula mamaria puede llevar a la transformación de las CT normales a CITCM<sup>4,14,23,31,35,45,47,60,62,82,83,91,92</sup>. Así mismo, estas vías se encuentran desreguladas en las CITCM y desempeñan un papel fundamental en la resistencia tumoral, recurrencia y metástasis<sup>9,47,60,62,83</sup>.

Notch es un receptor de señalización transmembranal necesario durante la embriogénesis e involucrado en la diferenciación celular, apoptosis, progresión del ciclo celular y en la determinación del destino de las CT. La familia Notch consta de los

---

receptores Notch1-4 los cuales a su vez se unen con cinco diferentes ligandos: ligandos tipo-delta (DLL1, DLL3 y DLL4) y las proteínas jagged (JAG1 y JAG2). Notch participa en la tumorigenesis mamaria vía el mantenimiento del fenotipo de las CITCM, así mismo han reportado que su actividad es mayor en las poblaciones enriquecidas de CITCM (CD44<sup>+</sup>CD24<sup>-/low</sup>ESA<sup>+</sup>) que en el resto de las poblaciones celulares tumorales ya diferenciadas. Al haber una activación aumentada de Notch se ha observado en las CITCM que la formación de esferas incrementa así como también la expresión de genes relacionados con troncalidad como lo son Oct4, Nanog, Sox2, ALDH y Klf4. Por último, esta vía juega un papel importante en la quimio y radioresistencia características de las CITCM<sup>4,8-10,14,23,31,35,38,45,47,60,62,69,73,82,83,93,95</sup>.

Wnt es una glicoproteína que sirve como ligando para Frizzled (FZD), receptor serpentina con siete dominios transmembrana y para la lipoproteína de baja densidad receptora relacionada con la proteína 5/6 (LRP5/6). A la fecha se han identificado diecinueve ligandos Wnt. Las proteínas de la familia Wnt activan la transferencia de  $\beta$ -catenina al núcleo y empieza la transcripción de genes que están implicados en la determinación de la diferenciación celular, división asimétrica, migración y polaridad celular. La vía de señalización de Wnt es una vía evolutivamente conservada que juega un papel clave en varias funciones fisiológicas y patológicas, incluyendo el desarrollo embrionario, homeostasis tisular y cáncer. Se ha reportado que está involucrada en la regulación de la autorenovación y diferenciación de las CT y CIT, en las CITCM han observado una alta actividad (especialmente en las ALDH<sup>+</sup>) en comparación con el resto de las poblaciones celulares tumorales (ALDH<sup>-</sup>), así como también se ha visto implicada en la radio y quimioresistencia de este grupo celular<sup>4,9,10,14,23,31,35,45,47,60,62,82</sup>.

La vía de señalización de Hedgehog (Hh) juega un rol muy importante en varios procesos celulares durante el desarrollo embrionario y homeostasis tisular, como también en la regulación de la proliferación, en el destino celular, troncalidad y en el mantenimiento

---

de las células progenitoras y su capacidad de renovación. En los mamíferos se han reportado tres ligandos homólogos de Hh: Sonic Hh (Shh), Indian Hh (Ihh) y Desert Hh (Dhh). La activación aumentada de esta vía se ha identificado en varios modelos de CIT, incluyendo las de cáncer de mama. Se ha reportado que Hedgehog está involucrada en la proliferación de las CITCM mediante la activación de los moduladores positivos de esta vía Gli1 y Ptch1. Por último, han observado que las poblaciones de CITCM con el fenotipo CD44<sup>+</sup>CD24<sup>-/low</sup>Lin<sup>-</sup>ALDH<sup>+</sup> se exhibe una alta señalización de Hh lo que provoca la retención del potencial de troncalidad<sup>4,9,23,31,35,47,60,62,82,83</sup>.

La vía de señalización JAK/STAT desempeña funciones muy importantes en procesos celulares como apoptosis, metástasis, proliferación, angiogénesis e invasión. Su desregulación ha sido observada en varios tipos de cáncer, así mismo está involucrada en el mantenimiento de las CIT y en las CITCM se ha propuesto a STAT3 como un posible marcador funcional de esta población, por lo que lo convierte en un nuevo blanco terapéutico para tomar en cuenta en el desarrollo del tratamiento dirigido. El bloqueo de esta vía resulta en la inhibición de la autorenovación y en la expresión de genes del metabolismo lipídico<sup>9,31,47,83</sup>. NF- $\kappa$ B es otra vía que se ha visto implicada en la expansión de la población CITCM mediante la activación de la vía de señalización Notch<sup>9,47</sup>.

Estas vías de señalización juegan papeles fundamentales en la regulación de las CITCM, nuevos estudios se requieren para poder comprender a fondo los mecanismos con los cuales intervienen en las funciones de esta población celular con la finalidad de poder desarrollar una terapéutica combinada y orientada a la eliminación específica de las CITCM sin la afcción de las funciones celulares normales<sup>4,23,31,47</sup>.

---

## 4.6 TERAPÉUTICA ACTUAL

La alta resistencia de las CITCM a la terapia convencional en comparación con otros grupos celulares tumorales destaca la necesidad de desarrollar nuevos tratamientos dirigidos a esta población celular. Las CITCM podrían ser eliminadas utilizando nuevos diseños terapéuticos eficientes (ya sea solos o en combinación con tratamientos convencionales) orientados al ataque de sus diferentes aspectos moleculares y funcionales como sus marcadores, vías de señalización responsables de la autorenovación, componentes moleculares que les confieren resistencia terapéutica (transportadores ABC, actividad de la familia ALDH, mecanismos de reparación del ADN, ROS, inhibidores de la apoptosis, etc), ncRNAs y su interacción tumor-estroma<sup>4,9,10,14,17,23,30,31,34,35,38,47,60,62,69,82</sup>. Hoy en día varios grupos de investigadores han dirigido sus estudios al desarrollo de nuevos compuestos que logren atacar estas propiedades y con ello conseguir no solamente la eliminación de las CITCM, sino que también impedir la resistencia y reincidencia tumoral.

El bloqueo de CD44, marcador de superficie comúnmente utilizado para el enriquecimiento de este grupo celular y receptor de AH induce la diferenciación de las CIT haciéndolas más susceptibles a varios regímenes terapéuticos, entre ellos a doxorubicina (antibiótico de la familia de las antraciclina). Así mismo, para su bloqueo se han implementado nuevas técnicas como la utilización de siRNAs, lentivirus así como la implementación del uso de nanopartículas con paclitaxel y rapamicina o salinomicina cubiertas con AH. Otra estrategia utilizada para el mismo fin ha sido mediante la inhibición de AH y su receptor utilizando pequeños oligosacáridos de AH que compiten con el polímero endógeno de AH o usando anticuerpos que bloquean el sitio de unión a AH de CD44. En el caso de prominina-1, se han utilizado inmunotoxinas para la destrucción de las moléculas CD133 y por lo tanto de las CITCM. Un ejemplo es AC133-saporina, AC-133 es un anticuerpo monoclonal anti-CD133 y la saponina es una toxina comúnmente utilizada para la generación de inmunotoxinas, la conjugación de AC-133 con saporina (AC133-

---

saporina) provoca arresto de la proliferación en las células CD133. La enzima ALDH1, no solamente es un marcador fenotípico, sino que también posee un rol funcional ya que también está involucrada en el mantenimiento de la troncalidad lo que la vuelve un gran blanco terapéutico para la erradicación de las CITCM. Para este objetivo hoy en día han utilizado witaferina A (lactona esteroidea derivada de una planta medicinal) la cual bloquea a ALDH1 (vía una regulación disminuida de Sox2, Oct4, Nanog y Bmi-1) y subsecuentemente ocasiona la pérdida de troncalidad de las CITCM, así como también han utilizado técnicas innovadoras como la terapia fototérmica que induce hipertermia localizada y con ello la eliminación de esta población celular. La actividad de esta enzima ha sido inhibida igualmente mediante el ácido holo-trans-retinóico (ATRA por sus siglas en inglés “All-trans Retinoic Acid”) y con el inhibidor específico de ALDH DEAB (del inglés “Diethylaminobenzaldehyde”) los cuales han demostrado que incrementan los efectos de la quimio (paclitaxel/doxorubicina) y radioterapia en células de cáncer de mama TN. Por otro lado, también han implementado el uso de nanopartículas con dosis bajas de decitabina (inhibidor de ADN metiltransferasas) con o sin doxorubicina para mejorar la respuesta de las CITCM con actividad alta de ALDH a la quimioterapia y para disminuir esta población celular. El factor de transcripción Nanog juega un papel muy importante en la autorenovación y en la pluripotencia de las CT embrionarias, se ha visto involucrada en la motilidad y en la resistencia a la quimioterapia y todas estas funciones las lleva a cabo mediante las interacciones con varios factores como OCT4, SOX2 y KLF4. La expresión aumentada de las proteínas de la familia Nanog se ha observado en varios tipos de cáncer incluyendo al de mama, la utilización de mitramicina A para alterar la expresión de este factor de transcripción junto con OCT4 logró sensibilizar a las CITCM a doxorubicina vía la supresión de genes relacionados con la autorenovación y quimioresistencia. En ámbitos generales el antibiótico inóforo salinomycinina (aislado de *S. Albus*) ha demostrado eliminar selectivamente a las CITCM al modificar la expresión de genes involucrados en la formación de esferas, TEM, diferenciación y supervivencia general. El uso de curcumina sola o en combinación con mitomicina C ha ayudado a sensibilizar a las CITCM a la quimioterapia,

---

como también el uso de eugenol con cisplatino o el extracto de castaña con paclitaxel. La cinasa Aurora A junto con Smad5 le confieren a las CITCM quimioresistencia mediante una vía no canónica, esta se ve abolida al utilizar alisertib, un inhibidor selectivo de la cinasa Aurora A. Han reportado también que la melatonina disminuye la capacidad de formación de mamoesferas y el fenetil ester del ácido caféico inhibe la capacidad de autorenovación y diferenciación de las CITCM. La metformina (descrita anteriormente) en combinación con trastuzumab inhibe la autorenovación y proliferación de las CITCM y en colaboración con doxorubicina disminuye esta población celular y el crecimiento tumoral<sup>9,10,20,23,30,31,47,48,62,82,93</sup>.

Previamente en el apartado de terapéutica de las CIT mencioné ya algunos tratamientos y ensayos que se han utilizado para el bloqueo de las vías de señalización. Algunos ejemplos de inhibidores de la  $\gamma$ -secretasa utilizados para bloquear a la vía de Notch y que hace a las CITCM más sensibles a agentes quimioterápicos como docetaxel incluyen a MK-0752 y PF-03084014. Otros agentes prometedores inhibidores de la  $\gamma$ -secretasa que se encuentran en fase I-III de ensayos clínicos incluyen a RO4929097, LY3039478, LY450139 y CB-103 así como también los anticuerpos monoclonales contra el ligando DLL4 MEDI0639 y REGN421 que se encuentran en fase I. La capsaicína es un derivado de los pimientos picantes el cual induce apoptosis en las CITCM al inhibir la translocación del dominio intracelular membranal de Notch en el núcleo, así mismo, el compuesto de la vitamina D  $1\alpha,25(\text{OH})_2\text{D}_3$  ha demostrado reducir la población de CITCM (regulando su diferenciación) por medio de la inhibición de Notch1-3 y JAG1-2<sup>4,8,9,20,31,38,47,62,82</sup>.

En el caso de la vía de Wnt, se ha reportado que su bloqueo ayuda a mantener el control de las CITCM, además de IWR-1 y XAV939 (inhibidores de la tanquirasa que ya fueron comentados) se han propuesto otros compuestos (que se encuentran la mayoría en fases de ensayos clínicos) para el bloqueo de la vía Wnt como LGK974 (inhibidor de “porcupine” que es una acetiltransferasa asociada a esta vía), celecoxib (anti-inflamatorio no

---

esteroideo), salinomicina, los anticuerpos monoclonales vantiactumab y cirmtuzumab (en combinación con paclitaxel), péptidos StAx, OMP-54F28, y Foxy-5, un péptido formilado que imita los efectos de Wnt5a y que actúa como un agente antimetastásico. Por otro lado, se ha reportado que el uso de nanopartículas con sulforafano en combinación con docetaxel, pamoato de pirvinio y fosfo-sulindaco inhibe la autorenovación de las CITCM, reprime la TEM y la resistencia terapéutica<sup>9,10,20,23,31,47,82</sup>.

Además de los compuestos vismodegib (antagonista del receptor smoothed o SMO que es parte de la vía de señalización de Hh) y ciclopamina (alcaloide esteroideal derivado de plantas que disminuye la capacidad de formación de mamoesferas) que han sido implementados para la represión de la vía Hh, han utilizado a sonidegib (antagonista de SMO que también se encuentra en ensayos clínicos) ya sea solo o combinado con paclitaxel o docetaxol y a GANT61 (inhibidor no canónico de Hh) el cual ha mostrado disminuir la población de CITCM en el cáncer de mama TN para el bloqueo de esta vía de señalización. Los efectos de el extracto acuoso de Huaier, la isoflavona de soya genisteína, el cloruro de nitidina y el efecto sinérgico de CDA-2 con tiotepa o paclitaxel han sido implementados en ensayos pre-clínicos para demostrar la represión de la vía Hh y la inhibición de la proliferación de las CITCM. La vía de señalización de Hippo es la vía más recientemente integrada que cuenta con propiedades para el mantenimiento de troncalidad de las CITCM, por lo que actualmente se encuentra bajo investigación para encontrar posibles blancos y agentes terapéuticos<sup>4,8-10,20,31,47,62,82</sup>.

Como mencionamos en el apartado de las CIT, FTC y Ko143 son compuestos utilizados para el bloqueo de ABCG2, su combinación con inhibidores de la fosfodiesterasa 5 pueden mejorar la respuesta de estos grupos celulares a agentes terapéuticos antineoplásicos. La leptina se encuentra involucrada en la interacción entre adipocitos y fibroblastos con las células cancerígenas en el microambiente tumoral mamario para inducir la TEM y expresión de genes relacionados con la troncalidad. Hoy en día se han utilizado

---

dos inhibidores de leptina Arg128Gln y péptidos cortos para lograr su bloqueo. Por último, la implementación de LM609 (inhibidor del receptor  $\alpha\beta 3$  de osteopontina) y aptámeros de osteopontina han logrado reducir la población de CITCM así como también han propuesto que el uso del inhibidor de ATM KU55933 puede lograr eliminar la radioresistencia en las CITCM<sup>9,35,38,47</sup>.

## 5. SURVIVINA

Survivina (codificada por el gen BIRC5), es una proteína multifuncional, con roles importantes en la regulación del ciclo celular como en la apoptosis. Se han reportado seis isoformas generadas mediante empalme o “splicing” alternativo: survivina DEx3, survivina 2B, survivina 2 $\alpha$ , survivina 3B, survivina 2B+32 y survivina. Es el miembro más pequeño de la familia de proteínas inhibidoras de la apoptosis IAPs (por sus siglas en inglés “Inhibitors of Apoptosis Proteins”) y junto con la cinasa Aurora B, la proteína centromérica interna INCENP (por sus siglas en inglés “Inner Centromere Protein”) y borealina (CDCA8 o Dasra) forman el complejo pasajero de los cromosomas CPC (por sus siglas en inglés “Chromosomal Passenger Complex”) el cual coordina varios eventos durante la mitosis como la segregación cromosómica y la citocinesis, esto, con el fin de lograr una división celular adecuada en donde no se vea comprometida la ploidía celular. Se expresa en tejidos embriónicos y posteriormente en tejidos normales está ausente. Han asociado la presencia de survivina citoplasmática con supervivencia celular mientras que la presencia de survivina a nivel nuclear con proliferación celular. En varios tipos de cáncer se encuentra sobre-regulada (incluyendo colon, pulmón, ovario, gástrico, mama, páncreas, hepático, cervical y próstata, entre otros) y se ha visto implicada en procesos como el mantenimiento de células troncales, angiogénesis, estrés celular, proliferación, disminución de la apoptosis, resistencia terapéutica y relapso tumoral. Varios estudios la han relacionado con la evolución clínica de diferentes tipos de cáncer y es hoy en día parte del panel de 21 genes

---

utilizado por la plataforma de predicción de recurrencia *Oncotype DX* y de la firma de clasificación de subtipos de cáncer de mama PAM50. Así mismo, varios autores han observado y establecido un vínculo entre la sobre-expresión de survivina y las CITCM<sup>101,114-122</sup>.

Durante varios años nuestro laboratorio ha estudiado a esta proteína, los mecanismos implicados en su empalme alternativo, específicamente de la isoforma DEx3, así como también hemos reportado que la sobre-expresión de esta isoforma disminuye la proliferación celular y la capacidad de formación de colonias y esferas tumorales en células HeLa, procesos fundamentales en el desarrollo tumoral<sup>116,117</sup>.

Por los antecedentes establecidos, survivina y BIRC5 son blancos terapéuticos potenciales y son particularmente de gran interés para nosotros ya que el conocer su diferentes perfiles y funciones nos permitirá comprender mejor los procesos de proliferación celular, apoptosis y resistencia terapéutica en el cáncer y específicamente en mi caso, en el cáncer de mama y las CITCM.

---

## **6. JUSTIFICACIÓN**

En la actualidad no existen reportes acerca del perfil genómico de las células iniciadoras tumorales, el estudiarlo nos permitirá comprender mejor las diferencias entre las subpoblaciones del cáncer de mama. Nos referimos a “perfil genómico” como a la integración de estudios y análisis que interrogan la estructura, expresión y/o función de un genoma de forma masiva y mediante plataformas genómicas de alto rendimiento, lo cual nos permitirá obtener una visión global de las alteraciones genómicas tumorigénicas (variaciones en el número de copias, variaciones transcripcionales, variaciones en el estado de metilación, etc) presentes en nuestra población de interés<sup>122–124</sup>.

## **7. HIPÓTESIS**

Las células iniciadoras tumorales presentan un perfil genómico diferente al de las células que componen el resto del tumor.

## **8. OBJETIVO GENERAL**

Conocer las diferencias entre el perfil genómico de las CIT y no CIT usando como modelo líneas celulares de cáncer de mama de origen basal y luminal.

## **9. OBJETIVOS ESPECÍFICOS**

1. Aislamiento y caracterización fenotípica de CITCM a partir de líneas celulares de cáncer de mama de origen basal y luminal.
2. Estudiar la estructura del genoma de las CITCM y compararla con las células no CITCM.

- 
3. Identificar a los genes que se encuentran diferencialmente expresados en las CITCM *versus* las no CITCM.
  4. \*Identificar genes que se encuentran diferencialmente metilados en las CITCM *versus* las no CITCM.
  5. Realizar un análisis bioinformático de las redes, cascadas y funciones celulares reguladas diferencialmente en estas células.
  6. Proponer posibles marcadores para estudios posteriores.

**\*Este objetivo se planteó una vez que tuvimos acceso a las diferentes plataformas.**

## **10. IMPACTO Y RELEVANCIA DEL PROYECTO EN EL ÁREA DE ESTUDIO**

La utilización de líneas celulares de cáncer de mama de origen luminal y basal (que representan los subtipos de cáncer de mama con mejor y peor pronóstico así como también subtipos ER<sup>+</sup> y ER<sup>-</sup> sucesivamente) como modelo de estudio, nos permitirá conocer las diferencias entre el perfil genómico de las CITCM y no CITCM. Las células iniciadoras tumorales y su papel en el crecimiento y propagación tumoral son hoy en día un tema de gran importancia en el ámbito de la investigación oncológica. Una vez teniendo en claro las características principales de las CIT podremos dar pie al desarrollo de nuevos blancos terapéuticos como al desarrollo de posibles marcadores diagnósticos, pronósticos y de preevalencia del cáncer de mama que continúa afectando drásticamente a la población femenina de nuestro país.

---

## 11. METODOLOGÍA

### 11.1 LÍNEAS CELULARES

Se utilizaron cuatro líneas celulares de cáncer de mama de origen basal y luminal (Tabla 4), las cuales adquirimos de la American Type Culture Collection (ATCC):

1. MDA-MB-231 (ATCC® HTB-26™)
2. MDA-MB-468 (ATCC® HTB-132™)
3. BT-20 (ATCC® HTB-19™)
4. MCF-7 (ATCC® HTB-22™)

**Tabla 4.** Características de las líneas celulares de origen basal y luminal.

LÍNEA CELULAR	ER	PR	HER2	SUBTIPO	TUMOR	ORIGEN
MDA-MB-231	-	-	-	BB	AC MET	EP
MDA-MB-468	-	-	-	BA	AC MET	EP
BT-20	-	-	-	BA	CDI	TP
MCF7	+	+	-	LA	AC MET	EP

**BA:** Basal A, **BB:** Basal B, **LA:** Luminal A, **CDI:** Carcinoma Ductal Invasivo, **AC MET:** Adenocarcinoma Metastásico, **EP:** Efusión Pleural y **TP:** Tumor Primario. (Modificado de Dai *et al.*, 2017 y Kao *et al.*, 2009)<sup>46,104</sup>.

---

## 11.2 CULTIVO CELULAR

Las células MDA-MB-231, MDA-MB-468 y MCF-7 se mantuvieron y se propagaron con medio DMEM (Medio de Eagle modificado por Dulbecco) (Gibco™, CA, EE. UU.) suplementado con 5% de suero fetal bovino (SFB) (Gibco™, CA, EE. UU.), mientras que para BT-20 utilizamos medio EMEM (Medio Mínimo Esencial de Eagle) (Gibco™, CA, EE. UU.) suplementado con 5% SFB (Gibco™, CA, EE. UU.). Todas las líneas celulares se incubaron a 37°C en una atmósfera húmeda con 5% de CO<sub>2</sub>.

## 11.3 AISLAMIENTO Y CARACTERIZACIÓN DE LAS CITCM

Se implementaron dos métodos para lograr el aislamiento y caracterización de las CITCM, citometría de flujo con FACS y separación celular activada por magnetos ((MACS®) Macs Miltenyi Biotec, Bergish Gladbach, DE.) los cuales se describen a continuación.

Por medio de la citometría de flujo (citómetro FACSAria™ y programa de análisis FACS Diva™ de Applied Biosystems, EE. UU.) con FACS (clasificación celular activada por fluorescencia, por sus siglas en inglés “Fluorescence Activating Cell Sorting”) se aislaron y caracterizaron a las CITCM de las cuatro líneas celulares. Los marcadores de superficie celular que utilizamos fueron: CD44, CD24 y EpCam (sus características fueron ya descritas en apartados anteriores) de Macs Miltenyi Biotec (Bergish Gladbach, DE.) y BD Biosciences (NJ, EE. UU.), la tinción se realizó siguiendo las recomendaciones de la casa productora. Las células se separaron con Accutase™ de BD Biosciences (NJ, EE. UU.), se mantuvieron y recuperaron en solución balanceada de Hanks con EDTA (ácido etilendiaminotetraacético) 0.1 mM (para la recolección agregamos 5% SFB y para la conservación 0.5% de albúmina de suero bovino (ASB)).

---

Así mismo, utilizando el sistema de separación MACS<sup>®</sup> se purificaron las CITCM provenientes de la línea celular de origen basal MDA-MB-231. Para este ensayo, las células se propagaron en monocapa, se separaron utilizando Accutase<sup>™</sup> de BD Biosciences (NJ, EE. UU.) y se resuspendieron en una solución específica (con PBS (buffer fosfato salino), ASB y EDTA) que recomienda la casa productora y que se utiliza durante todo el procedimiento de separación. Posteriormente, las células se pasaron por un filtro para la eliminación de agregados celulares, se realizó el marcaje de éstas con perlas paramagnéticas acopladas a EpCam (CD326 Microbeads, Macs Miltenyi Biotec, Bergish Gladbach, DE.) siguiendo las instrucciones del fabricante y por último se llevó a cabo la separación magnética utilizando las columnas MS y el sistema MiniMacs<sup>™</sup> ambas de Macs Miltenyi Biotec (Bergish Gladbach, DE.). Al finalizar el procedimiento se obtuvieron dos poblaciones, la primera, la fracción negativa a EpCam y la segunda, nuestro grupo de CITCM, la fracción positiva a EpCam.

La validación del aislamiento de las fracciones positivas y negativas a EpCam se realizó mediante citometría de flujo.

#### **11.4 EXTRACCIÓN DE ARN TOTAL**

Posterior a la recolección con FACS y MACS<sup>®</sup> de nuestras fracciones celulares, se utilizó el reactivo TRIzol<sup>™</sup> (Invitrogen<sup>™</sup>, ThermoFisher Scientific, MA, EE. UU.) para la extracción de ARN total de nuestras poblaciones siguiendo las recomendaciones de la casa productora. Una vez realizada la lisis celular con este reactivo, se añadió cloroformo, se centrifugó y se precipitó la fase acuosa con isopropanol. La pureza, concentración e integridad de las muestras se obtuvieron mediante las mediciones con nanodrop, la relación 260nm/280nm, el bioanalizador (Agilent Technologies, CA, EE. UU.) de la UMI que nos proporcionó el RIN (número de integridad de ARN) y mediante geles de agarosa al 1% teñidos con bromuro de etidio.

---

## 11.5 EXTRACCIÓN DE ADN

Se utilizó el Flexigene DNA kit y el QIAmp DNA Mini Kit (ambos de Quiagen Inc., Hilden, DE.) para la extracción del ADN de las poblaciones CITCM y no CITCM de MDA-MB-231, MDA-MB-468, BT-20 y MCF-7 aisladas previamente por FACS y MACS<sup>®</sup> siguiendo las instrucciones de la casa productora. Al final el ADN se resuspendió en agua Milli-Q<sup>®</sup> (Merck, Darmstadt, DE.) o TE1X, se cuantificó con Nanodrop y se observó la integridad mediante geles de agarosa al 1% y 1.5%.

## 11.6 ANÁLISIS DE EXPRESIÓN DE LAS CITCM

Se utilizaron de 250ng a 1.5µg de ARN total de las poblaciones CITCM y no CITCM de MDA-MB-231, MDA-MB4-68 y BT-20 aisladas previamente por FACS y MACS<sup>®</sup> para realizar el estudio de expresión con dos herramientas diferentes, la primera utilizando los microarreglos GeneChip<sup>™</sup> Human Gene ST 1.0 de Affymetrix<sup>®</sup> (ThermoFisher Scientific, MA, EE. UU.) y la segunda mediante RNAseq de la plataforma de Illumina<sup>®</sup> Genome Analyzer Iix (Illumina Inc., CA, EE. UU.). Ambos procedimientos se llevaron a cabo en el INMEGEN en la UMI y USEC respectivamente y siguiendo las recomendaciones de las unidades y casas productoras. Los datos obtenidos por ambas plataformas se analizaron con las herramientas bioinformáticas CLCBio (Quiagen Inc., Hilden, DE.) en el caso de RNAseq y Partek<sup>®</sup> Genomics Suite<sup>®</sup> ((PGS), Partek, MO, EE. UU.) y GenePattern<sup>125</sup> para los microarreglos. Los detalles específicos de los análisis realizados con las diferentes muestras y plataformas se describen a continuación.

### 11.6.1 ANÁLISIS DE RESULTADOS DE LAS MUESTRAS DE LA PRIMERA ETAPA A

El análisis de los datos generados con los microarreglos se realizó con PGS y GenePattern. Las muestras pasaron los controles de calidad con métricas de QC mayor a

---

0.8, se normalizaron los datos siguiendo las recomendaciones de PGS y GenePattern y se compararon las poblaciones CITCM y no CITCM utilizando los siguientes parámetros: “fold change” o razón de cambio (FC)  $> 1.3 / -1.3$  y un “p-value” o valor de  $p < 0.05$ . Al finalizar el estudio se obtuvo una lista con 2,306 genes sobre-expresados y 1,606 sub-expresados, el resto no tuvo cambios. Para el análisis de los datos generados con RNAseq se utilizó CLCBio con los siguientes parámetros: alineación con un máximo de dos desemparejamientos (“mismatches”), un mínimo de 0.9 para la longitud de fracción, un mínimo de 0.8 para la similitud de fracción y diez como máximo número de “hits” para un “read” (secuencias cortas de nt). Se filtraron todos los transcritos con un nivel relativo de expresión menor a 0.2 y 10 como un mínimo número de “reads”. Se calculó la expresión diferencial y aquellos genes con una diferencial de 1.3 fueron seleccionados para continuar nuestros siguientes análisis.

#### **11.6.2 ANÁLISIS DE RESULTADOS DE LAS MUESTRAS DE LA PRIMERA ETAPA B**

El análisis de los datos generados con RNAseq se realizó con CLCBio con los siguientes parámetros: alineación con un máximo de dos desemparejamientos (“mismatches”), un mínimo de 0.9 para la longitud de fracción, un mínimo de 0.8 para la similitud de fracción y diez como máximo número de “hits” para un “read”. Todos los transcritos se filtraron con un nivel relativo de expresión menor a 0.2 y 10 como un mínimo número de reads. La expresión diferencial se calculó y se seleccionaron los genes cuya diferencial fuera de 1.2 para los estudios subsecuentes.

#### **11.7 ANÁLISIS BIOINFORMÁTICO DE REDES, CASCADAS Y FUNCIONES CELULARES REGULADAS DIFERENCIALMENTE EN LAS CITCM**

Los resultados obtenidos del análisis de expresión y metilación con CLCBio y PGS se introdujeron en las herramientas Ingenuity® Pathways Analysis (IPA®, Quiagen Inc., Hilden,

---

DE.) para realizar la predicción de redes y cascadas y en GSEA (Gene Set Enrichment Analysis)<sup>107,126,127</sup> para el análisis de enriquecimiento de genes. Los parámetros de inclusión establecidos previamente con PGS, CLCBio se respetaron.

## 11.8 ESTUDIO ESTRUCTURAL DE LAS CITCM

Se utilizaron 500ng de ADN de 6 muestras CITCM y no CITCM de las líneas celulares MDA-MB-231, MDA-MB-468 y BT-20 aisladas previamente por FACS (muestras de la primera etapa B) y 500ng de ADN de seis muestras (dos replicas biológicas y una réplica técnica) CITCM (fracción positiva a EpCam) y no CITCM (fracción negativa a EpCam) de la línea celular MDA-MB-231 separadas previamente con MACS<sup>®</sup> (muestras de la primera etapa A) para realizar el estudio de genotipificación utilizando los microarreglos Genome-Wide Human SNP 6.0 de Affymetrix<sup>®</sup> (ThermoFisher Scientific, MA, EE. UU.). Este procedimiento se llevó a cabo siguiendo las recomendaciones de la casa productora en la UMI del INMEGEN. El análisis de los datos generados y la reconstrucción de la estructura genómica de las poblaciones CITCM y no CITCM provenientes de las tres líneas celulares se realizó utilizando las herramientas bioinformáticas GeneSpring ((GS) Agilent Technologies, CA, EE. UU.) y Partek<sup>®</sup> Genomics Suite<sup>®</sup> (Partek, MO, EE. UU.). Se normalizó mediante cuantiles y RMA. Se hizo un análisis pareado utilizando como base al grupo no CITCM para la observación de las variaciones en el número de copias de ADN (CNVs). Posteriormente, continuamos con la segmentación genómica para la detección de amplificaciones y deleciones utilizando diez como el parámetro mínimo de marcadores genómicos, un límite de p-value de 0.001, 0.3 de la relación señal/ruido (SNR) y 1.7 a 2.3 para el rango de número de copias diploide. Por último, se creó la lista de las regiones amplificadas o deletadas incluyendo valores mayores o iguales a y menores o iguales a dos.

---

## 11.9 ESTUDIO DEL PERFIL DE METILACIÓN DE LAS CITCM

Se utilizaron 500ng de ADN de tres experimentos diferentes y de las mismas muestras que usamos para el estudio estructural de las poblaciones CITCM y no CITCM de MDA-MB-231, MDA-MB-468 y BT-20 aisladas previamente por FACS para realizar el estudio de metilación con el microarreglo Infinium HumanMethylation450K BeadChip® (Illumina Inc., CA, EE. UU.), 1.0µg de ADN de las fracciones positivas y negativas a EpCam separadas previamente con MACS® y 1.0µg ADN de las cuatro poblaciones (CD44<sup>+</sup>CD24<sup>+</sup>, CD44<sup>-</sup>CD24<sup>-</sup>, CD44<sup>+</sup>CD24<sup>-</sup> y CD44<sup>-</sup>CD24<sup>+</sup>) de MCF-7 purificadas con FACS para realizar el estudio de metilación con el microarreglo Infinium HumanMethylation27 BeadChip® (Illumina Inc., CA, EE. UU.). La conversión del ADN con bisulfito se llevó a cabo con EZ DNA Methylation kit™ (Zymo Research, CA, EE. UU.) siguiendo las instrucciones de la casa productora<sup>128</sup>. Estos procedimientos se realizaron en la UMI del INMEGEN. El análisis de los datos generados se realizó con la herramienta bioinformática PGS. Inicialmente, los archivos crudos de intensidad (.idat) se importaron al módulo de metilación de PGS. Siguiendo las recomendaciones de Partek se normalizaron los datos y se corrigió el ruido de fondo. Para la detección de los niveles diferenciales de metilación entre las CITCM y no CITCM se hizo un análisis de ANOVA (análisis de varianza)<sup>93</sup> de una vía con los datos no transformados logarítmicamente (valores  $\beta$ ), así mismo se calcularon los valores de estimación. Se excluyeron las sondas de los cromosomas X y Y así como aquellas que contenían polimorfismos de un sólo nucleótido (SNPs). Finalmente se creó una lista de marcadores de 8,841 locus diferencialmente metilados en base a un p-value no ajustado de <0.05 los cuales fueron utilizados para los estudios subsecuentes.

La herramienta bioinformática HumMeth27QCReport también se utilizó, primeramente para el control de calidad y posteriormente para el análisis (sobretudo) del microarreglo HumanMethylation27. Se importaron tres archivos a esta plataforma “sample table”, “control table” y “BetaAverage table” generados por el software GenomeStudio de Illumina (Illumina Inc., CA, EE. UU.), se siguieron las instrucciones de

---

HumMeth27QCReport para seguir el proceso de análisis y para la generación de los siguientes reportes que contienen los resultados finales: lista de M values normalizados para la identificación de genes diferencialmente metilados, gráficas de controles internos, análisis de componentes principales (PCA), agrupación jerárquica, entre otros<sup>128</sup>.

### **11.10 FORMACIÓN DE MAMOESFERAS**

Se utilizaron  $80 \times 10^3$  células de las poblaciones CITCM y no CITCM de MDA-MB-231 aisladas previamente por FACS y  $80 \times 10^3$  células de la línea total (LT). Se sembraron en placas de agarosa al 2% con DMEM/F-12 (Gibco™, CA, EE. UU.) y sin suplementación de suero, se mantuvieron a 37°C durante seis días para que se formaran agregados, al séptimo día se trasladaron por tres días a una incubadora con movimiento orbital a 49 rpm conservando la misma temperatura, posteriormente se retiraron para medirlas y para sacar imágenes.

### **11.11 FORMACIÓN DE COLONIAS EN AGAR SUAVE**

Se utilizaron  $1 \times 10^3$  células de la línea total MDA-MB-231 y  $1 \times 10^3$  células de las fracciones positivas y negativas a EpCam de la línea celular de origen basal MDA-MB-231 (separadas previamente con MACS®) y se sembraron por triplicado en pozos con agar suave a una concentración de 0.33% con medio DMEM suplementado con 5% de SFB y se incubaron a 37°C en una atmósfera húmeda con 5% de CO<sub>2</sub>.

---

## **11.12 INDUCCIÓN TUMORAL EN RATONES HEMBRA NU/NU**

### **11.12.1 INOCULACIÓN DE CITCM Y NO CITCM PURIFICADAS POR FACS (2011)**

Se utilizaron  $1 \times 10^4$  células de las cuatro poblaciones (CD44<sup>+</sup>CD24<sup>+</sup>, CD44<sup>-</sup>CD24<sup>-</sup>, CD44<sup>+</sup>CD24<sup>-</sup> y CD44<sup>-</sup>CD24<sup>+</sup>) de MCF-7 aisladas previamente por FACS en la primera etapa A así como  $1 \times 10^4$  células de la LT MCF-7 y se inocularon por vía intradérmica en la espalda de ratones hembra *nu/nu* de aproximadamente 5 semanas de edad utilizando como vehículo Matrigel™ de BD Biosciences (NJ, EE. EE.) y PBS 1X. Se mantuvieron en observación durante 4 meses post-inoculación en los cuales analizamos tamaño, apariencia y condiciones generales de los ratones. Este experimento se llevó a cabo en el bioterio de la División de Investigación Básica del INCan (Instituto Nacional de Cancerología) de la Cd.Mx. (Ciudad de México).

### **11.12.2 INOCULACIÓN DE CITCM Y NO CITCM PURIFICADAS POR FACS (2012)**

Se utilizaron  $50 \times 10^3$  células de las poblaciones CITCM y no CITCM de MDA-MB-231 aisladas previamente por FACS en la primera etapa B y  $50 \times 10^3$  células de la LT. Estas se inocularon por triplicado y por vía intradérmica en la espalda de ratones hembra *nu/nu* de aproximadamente 12 semanas de edad utilizando como vehículo Matrigel™ de BD Biosciences (NJ, EE. EE.) y PBS 1X. Se mantuvieron en observación durante varios meses post-inoculación en los cuales analizamos tamaño, apariencia y condiciones generales de los ratones. Este experimento se llevó a cabo en la UPEAL-B de la Unidad Xochimilco de la Cd.Mx. .

---

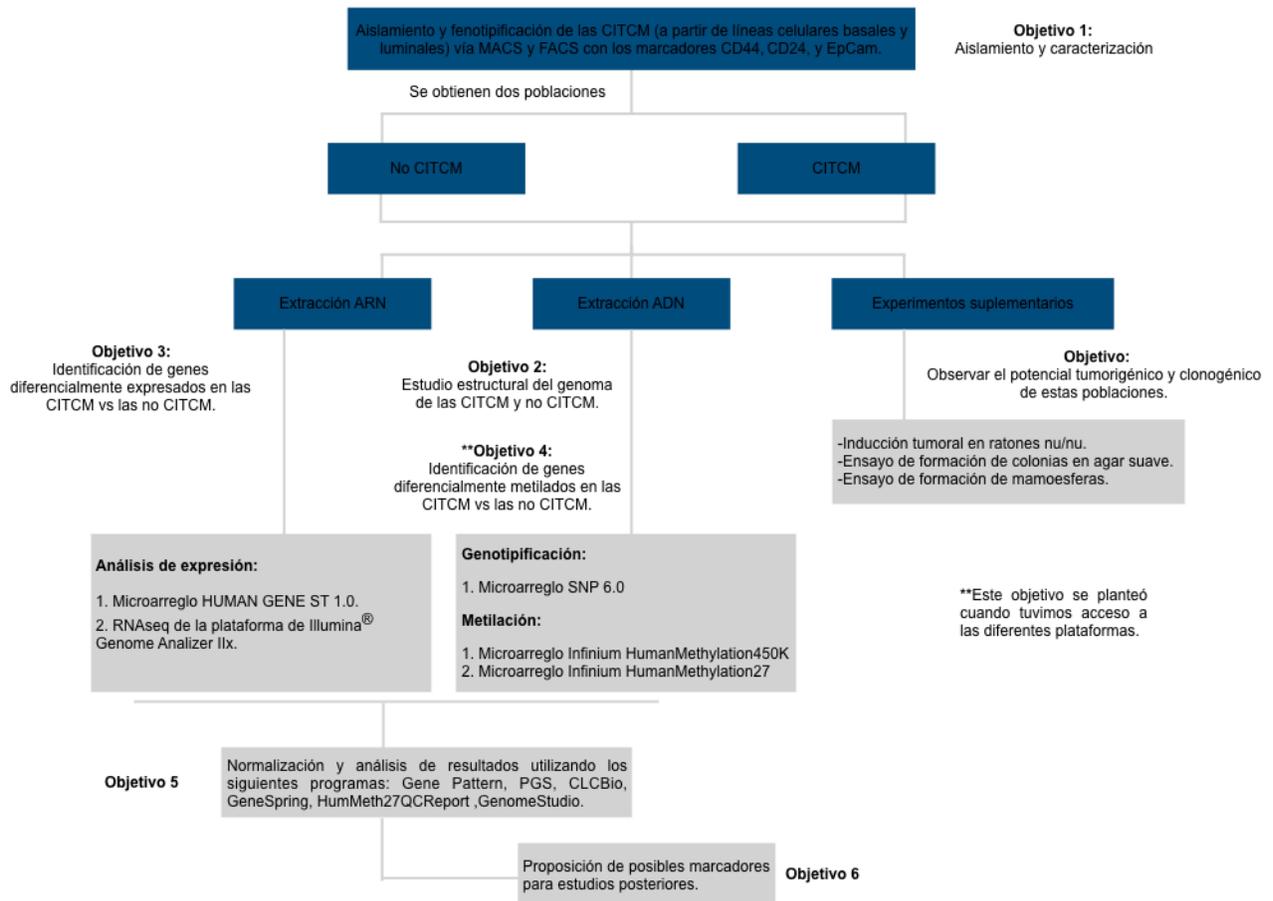
### 11.13 DIAGNÓSTICO HISTOPATOLÓGICO DE LAS TUMORACIONES FORMADAS EN LOS RATONES NU/NU

Se realizó el procesamiento de las muestras de los ratones *nu/nu*:

- Cinco tumores formados posterior a la inoculación de las poblaciones CD44<sup>+</sup>CD24<sup>+</sup>, CD44<sup>-</sup>CD24<sup>-</sup>, CD44<sup>+</sup>CD24<sup>-</sup> y CD44<sup>-</sup>CD24<sup>+</sup> y LT de la línea celular de origen luminal MCF-7 (2011).
- Tres tumores formados posterior a la inoculación de las poblaciones CITCM de MDA-MB231 y dos posterior a la inoculación con la LT MDA-MB231 (2012).

La tinción con hematoxilina y eosina y el análisis histopatológico se llevó a cabo en el LHIM del INMEGEN siguiendo sus recomendaciones.

## 12. DIAGRAMA GENERAL DE TRABAJO



---

## **13. RESULTADOS**

### **13.1 AISLAMIENTO Y CARACTERIZACIÓN FENOTÍPICA DE LAS CITCM A PARTIR DE LÍNEAS CELULARES**

Inicialmente, el aislamiento y caracterización fenotípica de las CITCM iban a realizarse a partir de la línea celular de origen basal MDA-MB-231 y de la línea celular de origen luminal MCF-7 utilizando el sistema de separación MACS<sup>®</sup> y FACS respectivamente. A esta primera etapa la denominaré A, ya que posteriormente modificamos la metodología a seguir (primera etapa B) para mejorar varios aspectos que se irán describiendo a lo largo de este apartado.

#### **13.1.1 PRIMERA ETAPA A**

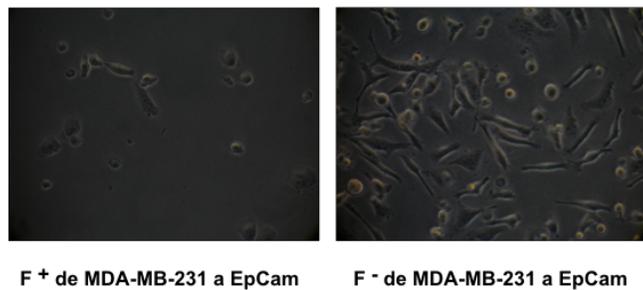
Se utilizaron las dos líneas celulares derivadas de adenocarcinomas metastásicos (específicamente de efusiones pleurales) MDA-MB-231 y MCF-7 para la purificación y caracterización de las CITCM. Con MDA-MB-231 se realizó la separación de la población CITCM utilizando el sistema MACS<sup>®</sup> y las perlas paramagnéticas acopladas a EpCam. Al finalizar se obtuvieron dos fracciones, la positiva a EpCam (F<sup>+</sup>) y a la que nombramos población CITCM y la negativa a EpCam (F<sup>-</sup>) denominada población no CITCM (Figura 7).

Para el aislamiento de las poblaciones CITCM y no CITCM de MCF-7 se utilizaron dos marcadores (cada uno acoplado a diferentes fluoróforos): CD44 y CD24, ambos utilizados frecuentemente por otros investigadores para este fin. Se realizaron las lecturas mediante citometría de flujo para cuantificar los niveles de expresión del doble marcaje en esta línea celular (Figura 8) y por último utilizando FACS se llevó a cabo la purificación de las cuatro poblaciones delimitadas por este marcaje: CD44<sup>+</sup>CD24<sup>+</sup>, CD44<sup>-</sup>CD24<sup>-</sup>,

CD44<sup>+</sup>CD24<sup>-</sup> (esta población corresponde a nuestra población CITCM) y CD44<sup>-</sup>CD24<sup>+</sup> (Figura 8).

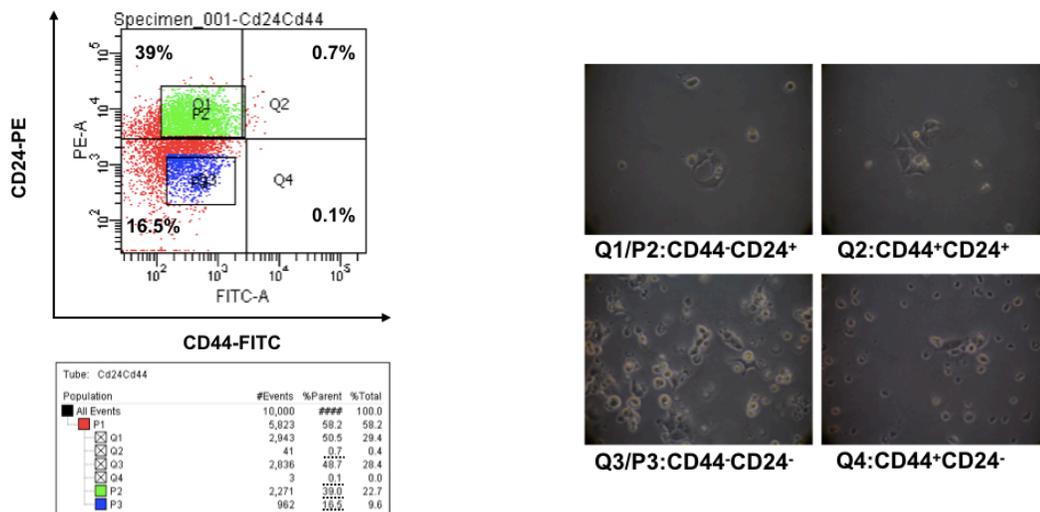
Posterior a la purificación, las poblaciones CITCM y no CITCM de las dos líneas celulares se utilizaron para la extracción de ARN, ADN y algunas muestras para la caracterización funcional de este grupo celular.

**Figura 7.** Poblaciones obtenidas posterior a la separación con MACS<sup>®</sup> de la línea celular MDA-MB-231.



En la imagen del lado izquierdo observamos la población CITCM (positiva a EpCam (F<sup>+</sup>)) y en la del lado derecho observamos a la que nombramos población no CITCM (negativa a EpCam (F<sup>-</sup>)).

**Figura 8.** Lectura y separación por FACS de las poblaciones CITCM y no CITCM de MCF7.



Las células se marcaron con los anticuerpos CD44 (isotiocianato de fluoresceína (FITC)) y CD24 (ficoeritrina (PE)) siguiendo las recomendaciones de la casa productora (BD Biosciences (NJ, EE. EE.)). En la gráfica observamos cuatro poblaciones: CD44<sup>+</sup>CD24<sup>+</sup> en el cuadrante Q2 y representa el 0.7%, C44<sup>-</sup>CD24<sup>-</sup> en el cuadrante Q3/P3 y representa el 16.5%, CD44<sup>+</sup>CD24<sup>-</sup> (nuestra población CITCM) en el cuadrante Q4 y representa el 0.1% y por último CD44<sup>-</sup>CD24<sup>+</sup> en el cuadrante Q1/P2 y representa el 39%. En la imagen derecha observamos las cuatro poblaciones posterior a la separación con FACS.

### 13.1.2 PRIMERA ETAPA B

En la primera etapa B, se utilizaron las tres líneas celulares de origen basal MDA-MB-231, MDA-MB-468 y BT-20, se cuantificaron los niveles de expresión mediante citometría de flujo de los marcadores previamente establecidos por varios grupos para purificar a las CITCM: CD44, CD24 y EpCam obteniendo los siguientes resultados, la expresión de la glicoproteína transmembranal tipo1 CD44 es mayor al 60% en las tres líneas celulares, mientras que de la molécula de adhesión CD24 observamos que no sigue un patrón común observándola desde un 0 a 90%, por último, la expresión de la molécula de adhesión de células epiteliales EpCam, conocida también como CD326 o ESA se encontró presente en más del 60% en las tres líneas celulares (Tabla 5).

**Tabla 5.** *Expresión de los marcadores de las CITCM en líneas celulares de cáncer de mama de origen basal.*

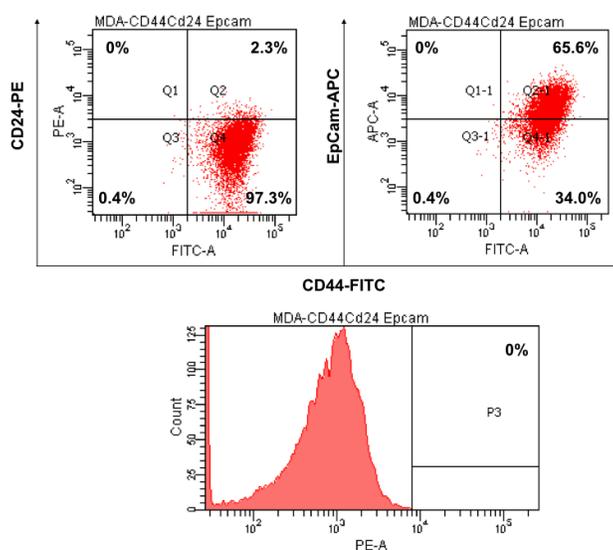
Porcentajes de citometría de flujo/ expresión/ IF\*

LÍNEA CELULAR	CD44	CD24	EPCAM	CITCM	NO CITCM
MDA-MB-231	>90	0-5	60-80	50-70	30-50
MDA-MB-468	70-90	80-90	>90	10-20	80-90
BT-20	60-70	70-80	>90	15-30	70-85

\*Los resultados están representados como el porcentaje de la intensidad de fluorescencia (IF)

Continuamos con el aislamiento de las CITCM utilizando el triple marcaje CD44+CD24-EpCam+ (cada uno acoplado a un fluoróforo diferente) mediante FACS (Figura 9) en las tres líneas celulares, obteniendo al final de la separación dos grupos, la descrita como CITCM y un segundo grupo representado con las siguientes poblaciones celulares: CD44+CD24-EpCam-, CD44+CD24+EpCam+, CD44+CD24+EpCam-, CD44-CD24-EpCam-, CD44-CD24-EpCam+, CD44-CD24+EpCam+ y CD44-CD24+EpCam- que llamaremos población no CITCM.

**Figura 9.** Lectura de la población CITCM de MDA-MB231 previa al aislamiento con FACS.



Las células se marcaron con los anticuerpos CD44 (FITC), CD24 (ficoeritrina (PE)) y EpCam (aloficocianina (APC)) siguiendo las recomendaciones de las casas productoras BD Biosciences (NJ, EE. EE.) y Macs Miltenyi Biotec (Bergish Gladbach, DE.). En la imagen inferior observamos la subpoblación positiva y negativa a CD24 presente en la población CD44+EpCam- (Q2-1).

Posterior a la purificación, las poblaciones CITCM y no CITCM de las tres líneas celulares se utilizaron para la extracción de ARN, ADN y algunas muestras para la

---

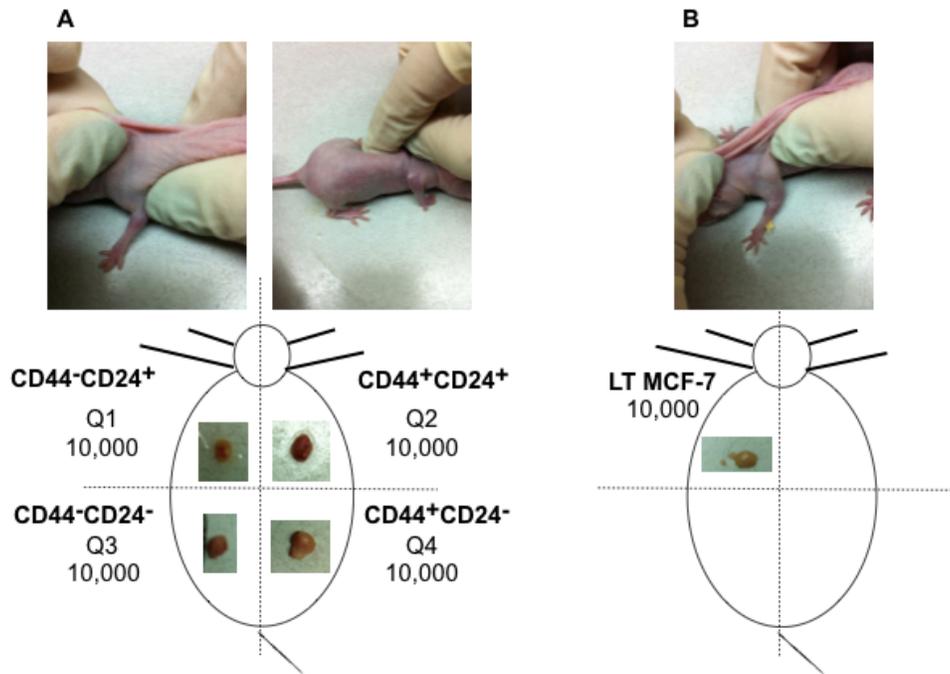
caracterización funcional de este grupo celular (inducción tumoral en ratones y ensayo de formación de mamoesferas).

### **13.2 CARACTERIZACIÓN FUNCIONAL DE LAS CITCM**

Para comprobar que en efecto estos grupos estuvieran enriquecidos con las poblaciones CITCM y no CITCM se realizaron tres pruebas con algunas de las muestras aisladas previamente con MACS® y FACS, la primera *in vitro* mediante el ensayo de formación de esferas tridimensionales (3D) en suspensión y en condiciones no adherentes, el segundo también *in vitro* mediante el ensayo de formación de colonias en agar suave y el tercero *in vivo* mediante la inducción tumoral en ratones hembras *nu/nu* (todos los procedimientos se efectuaron de acuerdo a lo descrito en la sección de metodología).

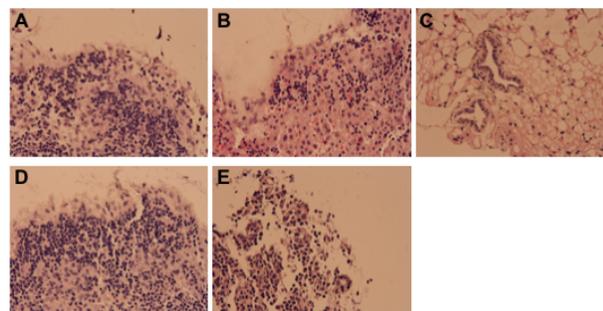
Como esperábamos y como ya publicado por otros autores, observamos y comprobamos que la población CITCM cuenta con mayor capacidad de formación de esferas 3D, mayor capacidad de formación de colonias y mayor capacidad de formación tumoral que la población no CITCM (Figuras 10-13). Por último y posterior al procesamiento de las muestras y tinción con hematoxilina y eosina (Figura 11), se realizó el diagnóstico histopatológico de las tumoraciones generadas en los ratones inoculados con las poblaciones CITCM, LT y no CITCM de la línea celular MDA-MB-231 (aisladas previamente como especifiqué en la primera etapa B), con los resultados obtenidos (Tabla 6) reafirmamos la capacidad tumorigénica de nuestras poblaciones CITCM (para mayor información consultar Bahena-Ocampo *et al.*, 2016, en donde profundizamos en la caracterización funcional de las CITCM)<sup>16</sup>.

**Figura 10.** Inducción tumoral en ratones utilizando la línea celular MCF-7.



Resultado de la inducción tumoral en ratones hembras *nu/nu* utilizando la población LT, CITCM (CD44+CD24-) y las poblaciones no CITCM (CD44+CD24+, CD44-CD24- y CD44-CD24+) provenientes de la línea celular MCF-7.

**Figura 11.** Tinción con hematoxilina y eosina de las poblaciones de MCF-7.

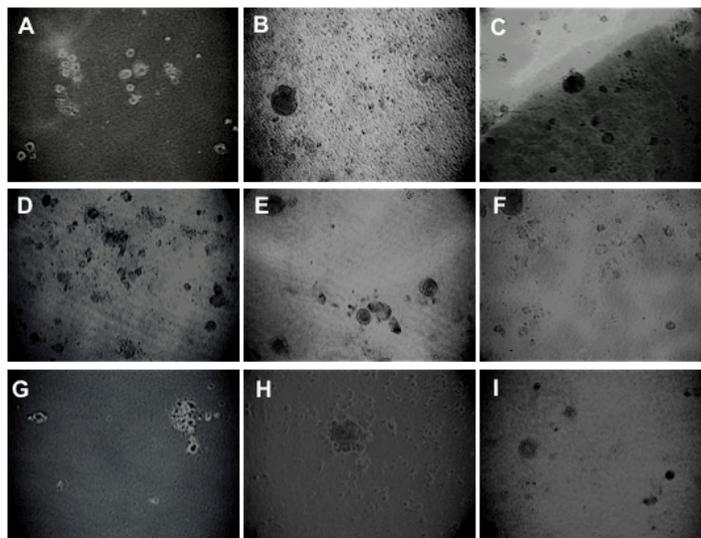


Resultado de la tinción con hematoxilina y eosina de las tumoraciones formadas con las distintas poblaciones de la línea celular MCF-7 en ratones hembras *nu/nu*. **A.** CD44-CD24+, 40X, diámetro:

---

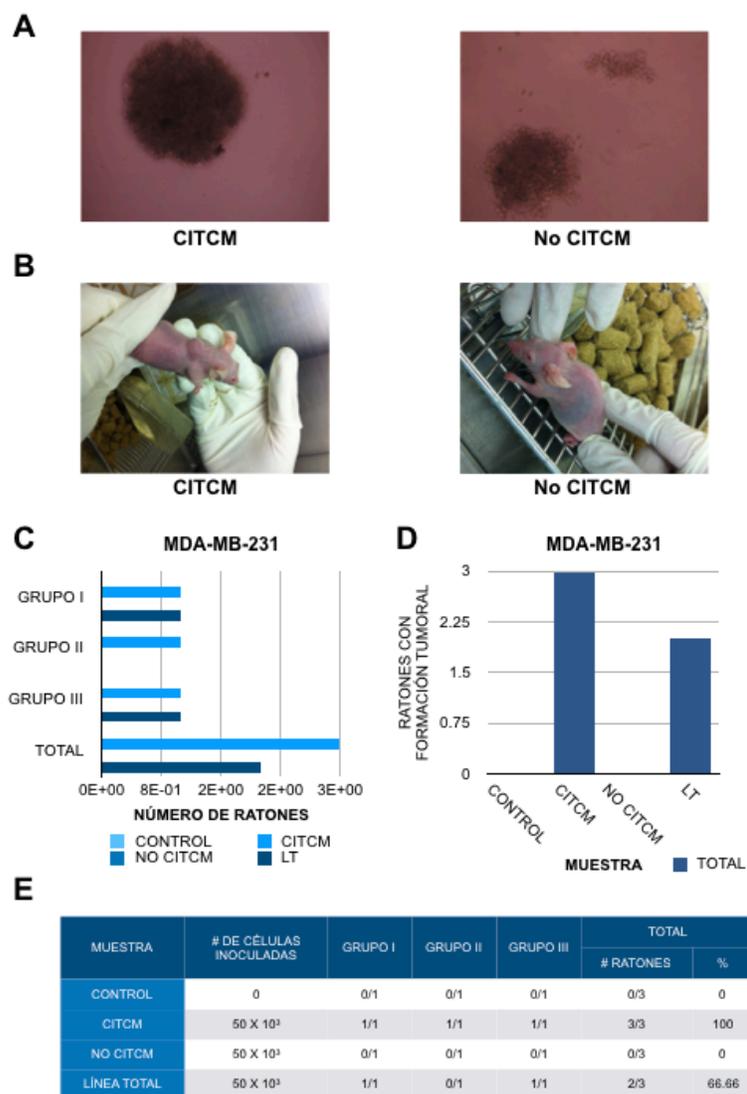
80.97  $\mu\text{m}$  **B.** CD44<sup>+</sup>CD24<sup>+</sup>, 40X, diámetro: 108.13  $\mu\text{m}$  **C.** CD44<sup>-</sup>CD24<sup>-</sup>, 40X, diámetro: 80.55  $\mu\text{m}$  **D.** CD44<sup>+</sup>CD24<sup>-</sup>, 40X, diámetro: 133.76  $\mu\text{m}$  **E.** LT MCF-7, 40X, diámetro: 54.21  $\mu\text{m}$ .

**Figura 12.** Ensayo de formación de colonias de las poblaciones de MDA-MB-231.



Resultado por triplicado del ensayo de formación de colonias utilizando poblaciones CITCM, LT y no CITCM de la línea celular MDA-MB-231 aisladas previamente con MACS<sup>®</sup>. **A, B y C** poblaciones LT de MDA-MB-231. **D, E y F** poblaciones F<sup>+</sup> de MDA-MB-231 a EpCam. **G, H e I** poblaciones F<sup>-</sup> de MDA-MB-231 a EpCam.

**Figura 13. Pruebas de caracterización funcional con MDA-MB-231.**



Pruebas de caracterización funcional utilizando CITCM y no CITCM provenientes de la línea celular MDA-MB-231 aisladas previamente por FACS. **A.** Resultado del ensayo de formación de mamoesferas en condiciones no adherentes. **B.** Resultado de la inducción tumoral en ratones hembras *nu/nu*. **C.** Representación gráfica de los resultados obtenidos por grupo (cada uno está compuesto por cuatro poblaciones celulares inoculadas de forma individual en un ratón) en la inducción tumoral. **D.** Representación gráfica de los resultados totales obtenidos en la inducción tumoral. **E.** Eficiencia individual y grupal de la inducción tumoral.

**Tabla 6. Diagnóstico histopatológico de tumoraciones de MDA-MB-231.**

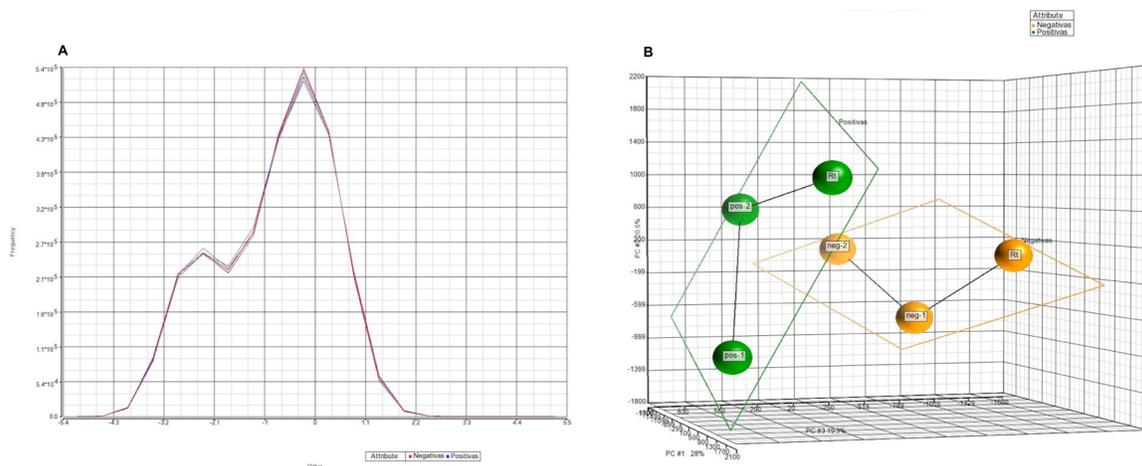
MUESTRA	DIAGNÓSTICO MICROSCÓPICO
CITCM 1	Se identifica piel, sin alteraciones histológicas, en otro fragmento se identifica lesión neoplásica constituida por células pequeñas con escaso citoplasma y núcleo hiper cromático que corresponde a carcinoma poco diferenciado.
CITCM 2	Se identifica escaso tejido compatible con tejido hepático, así como lesión de aspecto neoplásica constituida por pseudopapilas, revestidas por células grandes con citoplasma moderado-abundante eosinófilo y núcleo hiper cromático. Se observa abundante necrosis e infiltrado inflamatorio de predominio linfocitario y células plasmáticas. Compatible con carcinoma poco diferenciado.
CITCM 3	Se identifica escaso tejido adiposo multilocular y escasas fibras de músculo estriado esquelético, el tejido adiposo está infiltrado por lesión neoplásica constituida por láminas de células pequeñas con escaso-moderado citoplasma, algunas con núcleo hiper cromático y mitosis, se identifican pequeñas áreas de necrosis. Los cambios corresponden a carcinoma poco diferenciado.
LT 3A	Se identifica lesión neoplásica constituida por láminas de células pequeñas con escaso-moderado citoplasma, algunas con núcleo hiper cromático y mitosis, se identifican pequeñas áreas de necrosis. Los cambios corresponden a carcinoma poco diferenciado. No se identifica órgano residual.
LT 3B	Se identifica lesión neoplásica constituida por láminas de células pequeñas con escaso-moderado citoplasma, algunas con núcleo hiper cromático y mitosis, se identifican pequeñas áreas de necrosis. Los cambios corresponden a carcinoma poco diferenciado. No se identifica órgano residual.
LT 1	Se identifica lesión neoplásica constituida por grupos y láminas de células medianas a grandes con citoplasma moderado, núcleo grande irregular, hiper cromático, áreas de necrosis e infiltrado inflamatorio de tipo linfocitario correspondiente a carcinoma moderadamente diferenciado. No se identifica órgano residual.

Diagnóstico histopatológico de cinco tumoraciones formadas en los ratones hembra *nu/nu* (tres inoculados con las poblaciones CITCM de la línea celular MDA-MB-231 y dos inoculados con la línea total (LT 1 y 3)).

### 13.3 ESTUDIO DE LA ESTRUCTURA DEL GENOMA DE LAS CITCM

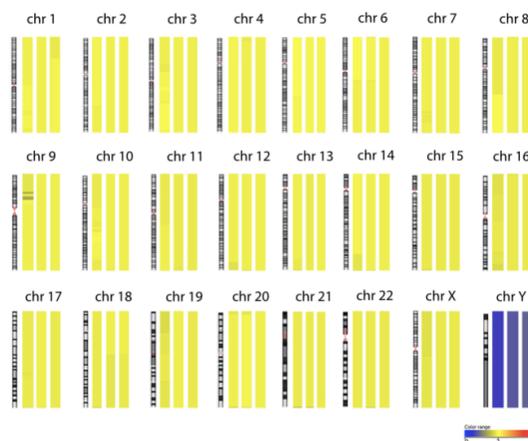
Para poder examinar la estructura genómica de las poblaciones CITCM usamos los microarreglos Genome-Wide Human SNP 6.0 de Affymetrix® utilizados hoy en día frecuentemente para este fin. Al concluir la normalización (Figura 14) y el análisis de los datos (procedimiento descrito previamente en el apartado de metodología) con las herramientas bioinformáticas PGS y GS observamos que no hay diferencias significativas en cuanto a las CNVs al compararse CITCM *versus* (vs) no CITCM de forma grupal (Figura 15), sin embargo, si encontramos algunas diferencias al hacer la comparación de estos mismo grupos pero a nivel individual de cada línea celular y utilizando las muestras de la primera etapa B (Tablas 7-9 y Figura 16).

**Figura 14.** Normalización y PCA de los datos de genotipificación de MDA-MB-231.



Normalización (**A**) y PCA (**B**) de los datos generados en el estudio de genotipificación con los microarreglos SNP 6.0 de Affymetrix® utilizando las muestras aisladas en la primera etapa A con MACS® (tres F<sup>+</sup> de MDA-MB-231 y tres F<sup>-</sup> de MDA-MB-231). La normalización se realizó mediante cuantiles y RMA. **Rt**: réplica técnica, **pos**: F<sup>+</sup> de MDA-MB-231 (verdes) y **neg**: F<sup>-</sup> de MDA-MB-231 (amarillas). Ambos análisis se realizaron con PGS.

**Figura 15.** Gráfica del análisis de CNVs de las poblaciones CITCM de MDA-MB-231.



Representación gráfica del análisis de las CNVs observadas en las poblaciones CITCM (F<sup>+</sup> de MDA-MB-231 a EpCam). El análisis se realizó con PGS, las pérdidas se observan en azul, ganancias en rojo y en amarillo sin cambios. No se encontraron cambios significativos.

**Tabla 7.** CNVs observadas en la población CITCM de la línea celular MDA-MB-231.

Chrom.	Start	Stop	Cytoband	MDA231B CN Status	Nearest Feature
8	5076079	5076344	8p23.2	del	CSMD1 (-)
11	126705738	126708817	11q24.2	del	KIRREL3-AS3 (+)

Se visualiza la posición de la variación (inicio y final), el estado de esta (*del* corresponde a delección y *ampl* a amplificación), junto con el gen más próximo. El análisis se realizó con PGS con las muestras de la primera etapa B.

**Tabla 8.** CNVs observadas en la población CITCM de la línea celular BT-20.

Chrom.	Start	Stop	Cytoband	BT20E CN Status	Nearest Feature
1	14603675	14619990	1p36.21	del	KAZN (+)
7	72063448	72095452	7q11.23	del	LOC100101148 (+)
7	73402908	73434896	7q11.23	del	CLIP2 (+)
11	25241937	25249045	11p14.3	ampl	LUZP2 (+)
11	25253376	25253488	11p14.3	ampl	LUZP2 (+)
11	67232189	67272921	11q13.2	del	ALDH3B2 (-)

Chrom.	Start	Stop	Cytoband	BT20E CN Status	Nearest Feature
17	18229846	18307302	17p11.2	del	KRT16P1 (+)
19	18701866	18749544	19p13.11	del	CRTC1 (+)
19	19473591	19497315	19p13.11	del	NDUFA13 (+)
22	22571602	22580740	22q11.23	del	MIF (+)
22	36510952	36511047	22q13.1	del	H1F0 (+)
22	39869762	39871431	22q13.2	del	EP300 (+)
X	41055055	41074406	Xp11.4	del	DDX3X (+)

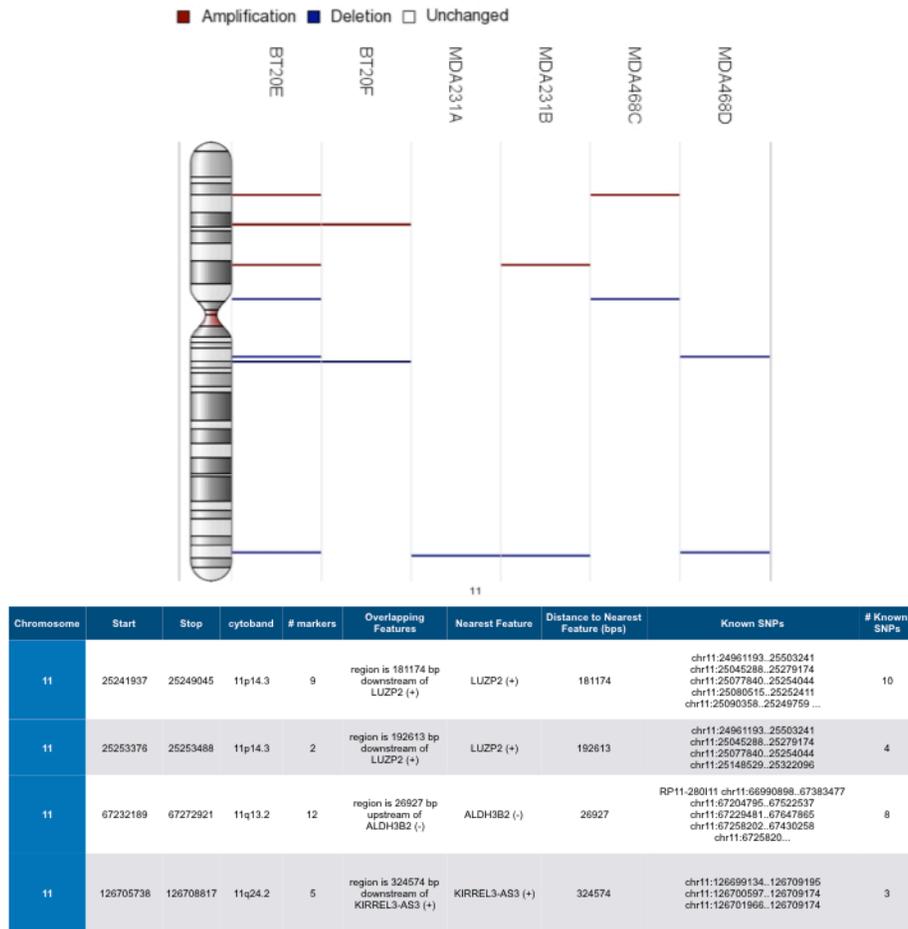
Se visualiza la posición de la variación (inicio y final), el estado de esta (*del* corresponde a delección y *ampl* a amplificación), junto con el gen más próximo. El análisis se realizó con PGS con las muestras de la primera etapa B.

**Tabla 9.** CNVs observadas en la población CITCM de la línea celular MDA-MB-468.

Chrom.	Start	Stop	Cytoband	MDA468C CN Status	Nearest Feature
2	107968943	107971090	2q12.3	ampl	SLC5A7 (+)
5	88035835	88039858	5q14.3	del	MEF2C (-)
7	109224334	109226581	7q31.1	del	EIF3IP1 (-)
8	10353613	10367823	8p23.1	del	MSRA (+)
9	87313507	87316423	9q21.33	del	AGTPBP1 (-)
20	1517410	1519397	20p13	del	SIRPB1 (-)
x	5782096	5783750	Xp22.32	ampl	NLGN4X (-)

Se visualiza la posición de la variación (inicio y final), el estado de esta (*del* corresponde a delección y *ampl* a amplificación), junto con el gen más próximo. El análisis se realizó con PGS con las muestras de la primera etapa B.

**Figura 16. Gráfica del análisis de CNVs del cromosoma 11 en CITCM.**



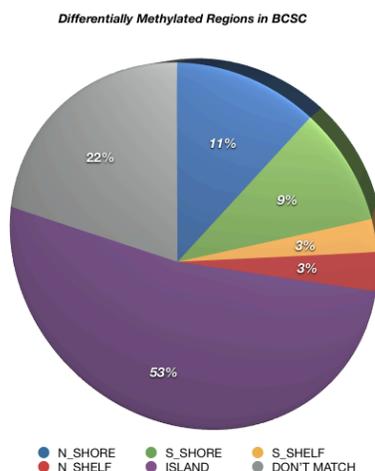
Representación gráfica y datos de CNVs del cromosoma 11 observadas en las poblaciones CITCM y no CITCM de las tres líneas celulares (muestras de la primera etapa B). En rojo se observan las amplificaciones, en azul las deleciones y en blanco no hay cambios. No encontramos diferencias significativas entre las poblaciones CITCM y no CITCM. El análisis se realizó con PGS.

---

### 13.4 ESTUDIO DEL PERFIL DE METILACIÓN DE LAS CITCM

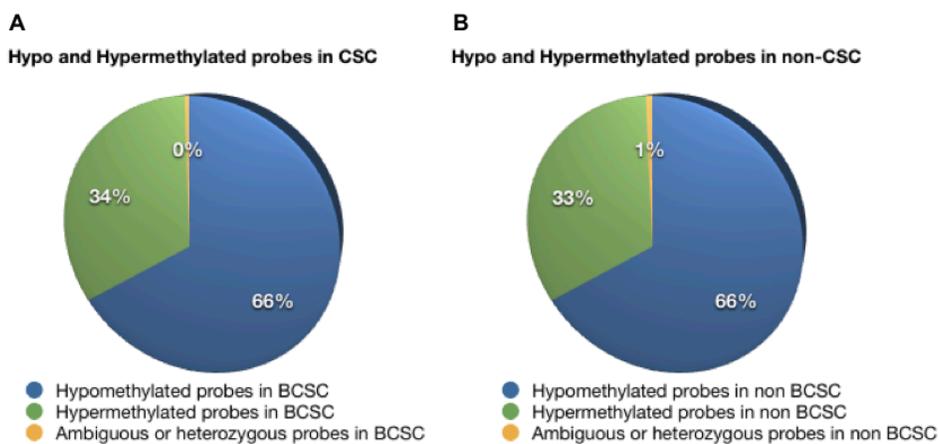
Para poder conocer el perfil de metilación de ambas poblaciones se utilizó inicialmente el microarreglo Infinium HumanMethylation27 BeadChip® con el ADN de las muestras purificadas en la primera etapa A, una F<sup>+</sup> de MDA-MB-231 a EpCam, una LT de MDA-MB-231, una F<sup>-</sup> de MDA-MB-231 a EpCam y una muestra de cada una de las cuatro poblaciones de MCF-7 purificadas con FACS (CD44<sup>+</sup>CD24<sup>+</sup>, CD44<sup>-</sup>CD24<sup>-</sup>, CD44<sup>+</sup>CD24<sup>-</sup> y CD44<sup>-</sup>CD24<sup>+</sup>). Posteriormente y por la disponibilidad del microarreglo Infinium HumanMethylation450K BeadChip® en la UMI del INMEGEN, se realizó un nuevo estudio utilizando esta herramienta que interroga más de 450,000 sitios CpG en comparación con el HumanMethylation27 que mide los niveles de metilación de únicamente 27,578 sitios CpG<sup>94,128</sup>. Las muestras de ADN que se usaron fueron las mismas utilizadas para el estudio estructural y aisladas previamente en la primera etapa B (seis muestras de CITCM y no CITCM de MDA-MB-231, MDA-MB-468 y BT-20). El procedimiento y el análisis de los datos generados se llevó a cabo como especifiqué en el apartado de metodología. Al finalizar, se obtuvo una lista de 8,841 loci diferencialmente metilados en la población CITCM, observamos la distribución de estos loci por región en la Figura 17 y en la Figura 18 contemplamos el estado de metilación de las poblaciones CITCM y no CITCM. En la Tabla 10 podemos ver los 20 cgs diferencialmente metilados en las CITCM, mientras que en la Figura 19 observamos el análisis de agrupación jerárquica y mapa de calor generado en base a los genes diferencialmente metilados. Por último analizamos estos resultados junto con los obtenidos en el análisis de expresión para observar si había una correlación entre los niveles de expresión y metilación específicamente en genes asociados con troncalidad, los resultados se describen y se muestran en el siguiente apartado.

**Figura 17.** Gráfica de las regiones diferencialmente metiladas en las CITCM.



Representación gráfica de los 8,841 loci diferencialmente metilados por región en la población CITCM (BCSC). N\_SHORE: azul, N\_SHELF: rojo, S\_SHORE: verde, ISLAND: morado, S\_SHELF: amarillo, DON'T MATCH: gris. El análisis se realizó con PGS.

**Figura 18.** Gráficas del estado de metilación en las CITCM y no CITCM.



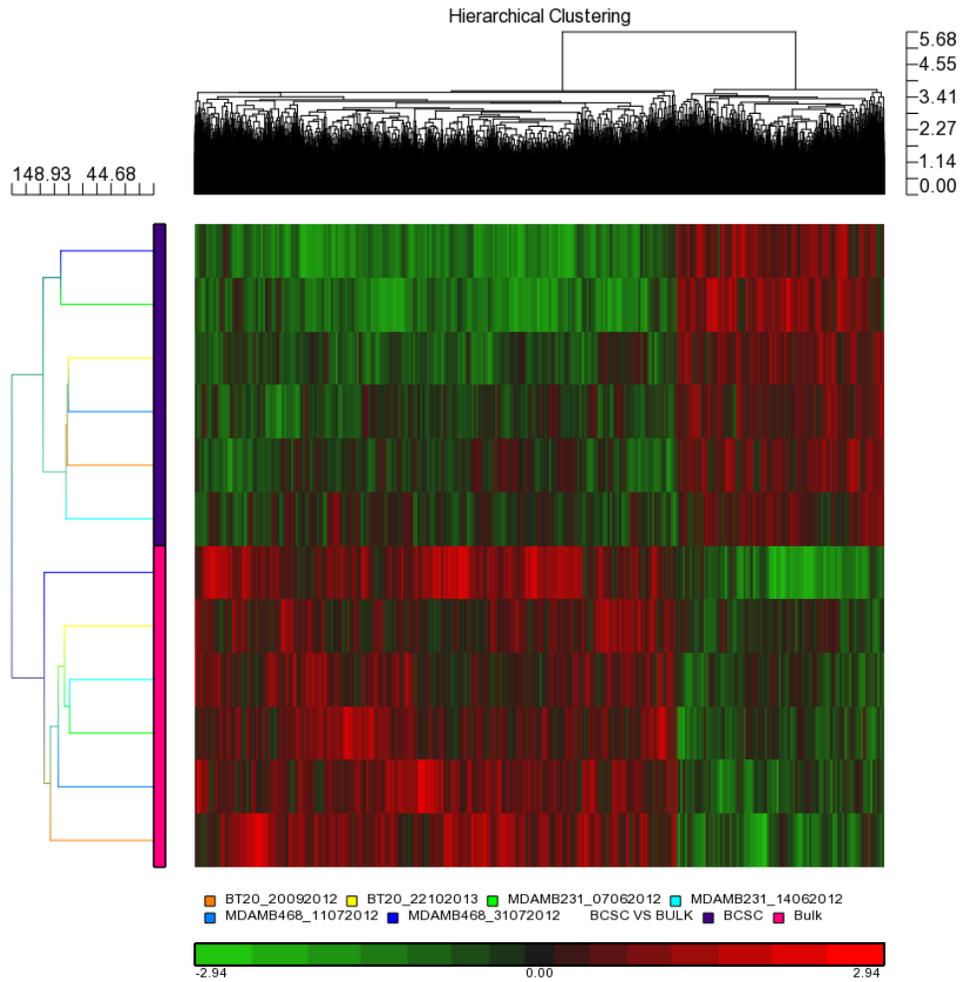
Gráficas ilustrativas del estado de metilación en la población CITCM ((CSC/BCSC) (A)) y en la no CITCM ((non-CSC/non-BCSC) (B)). Se considera una sonda hipometilada cuando su valor  $\beta$  es menor o igual a 0.3 e hipermetilada cuando su valor  $\beta$  es mayor o igual 0.7. El análisis se realizó con PGS.

**Tabla 10. Los 20 cgs más diferencialmente metilados en las CITCM.**

Estimate (CSC vs. Non CSC)	AVERAGE_B_CSC	AVERAGE_B_NonCSC	UCSC_REFGENE_NAME	Column ID
0.349598	0.649322	0.299723667	GSTT1	cg15254640
0.176693	0.813176667	0.636484	EWSR1	cg11942221
0.176469	0.9521405	0.775671667	HIST1H2BD	cg04492287
0.176212	0.851728	0.675515833		cg17474785
0.159485	0.890014833	0.73053	LBR;LBR	cg07954054
0.148651	0.875423333	0.7267725	FAM190A	cg15557457
0.138684	0.896253667	0.75757	DEFB114	cg19592003
0.137941	0.930568833	0.792627333	COL1A2	cg12563520
0.135699	0.833996333	0.698297333	LCORL	cg16328207
0.134979	0.8676065	0.732627833		cg25806491
0.13474	0.9154675	0.780728167	PDCD4	cg05999324
0.133359	0.924450167	0.791091167		cg18090431
0.132724	0.211758833	0.079035467	MCOLN1	cg00504457
0.129895	0.945423333	0.815528333	IQCH	cg19392994
0.125528	0.9142225	0.788694333	RIF1	cg26639076
0.125432	0.859980333	0.734548333	ZNF234	cg01798097
0.121996	0.936885667	0.8148895	SH3RF1	cg27416948
0.120812	0.8685205	0.7477085	VRK1	cg14398228
0.119806	0.892894833	0.773089167		cg24007445
0.118253	0.914939667	0.796686833	IGF1R	cg25709679

Los 20 cgs más diferencialmente metilados en las CITCM (CSC) en comparación con las no CITCM (Non CSC). El análisis comparativo se realizó con PGS y se utilizó un parámetro de estimación para determinar el rango de los niveles diferenciales.

**Figura 19.** Agrupación jerárquica y mapa de calor de los genes diferencialmente metilados en las CITCM y no CITCM.



Agrupación jerárquica y mapa de calor creados en PGS en base a los genes diferencialmente metilados en CITCM (morado) y no CITCM (rosa). La escala de color representa el nivel de metilación de cada gen, en verde hipometilados y en rojo hipermetilados. En las filas observamos las muestras y en las columnas los genes.

---

## 13.5 IDENTIFICACIÓN DE GENES DIFERENCIALMENTE EXPRESADOS EN LAS CITCM

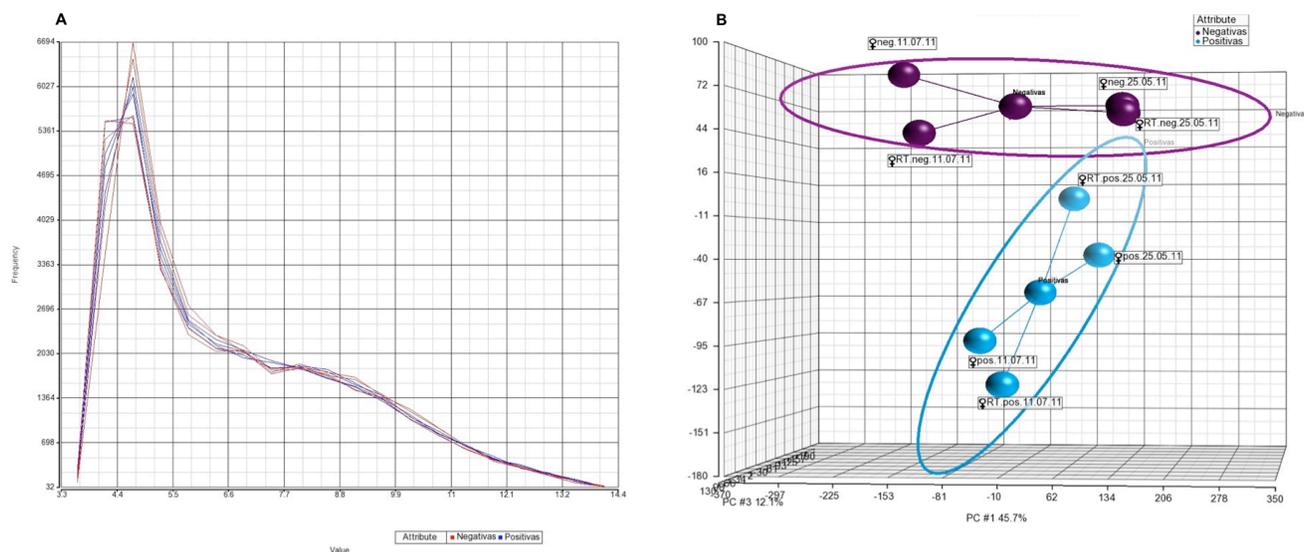
Para poder identificar genes diferencialmente expresados en la población CITCM se utilizaron dos plataformas, microarreglos Human Gene ST 1.0 de Affymetrix® y RNAseq de la plataforma de Illumina® Genome Analyzer Iix. De igual forma se usaron muestras de la primera etapa A y B, a continuación se describen los resultados.

### 13.5.1 PRIMERA ETAPA A

Se utilizaron 250ng de ARN de cuatro muestras (dos réplicas biológicas y dos réplicas técnicas) de las poblaciones CITCM (F<sup>+</sup> de MDA-MB-231 a EpCam) y no CITCM (F<sup>-</sup> de MDA-MB-231 a EpCam) aisladas previamente con MACS® para realizar el estudio de expresión con microarreglos. Para el estudio de expresión con RNAseq usamos aproximadamente 1.5µg de una muestra de la F<sup>+</sup> de MDA-MB-231 a EpCam y una muestra de F<sup>-</sup> de MDA-MB-231 a EpCam. Los procedimientos técnicos y de análisis se llevaron a cabo como mencionamos en el apartado de metodología. Los resultados se normalizaron y observamos su comportamiento por medio de un PCA (Figura 20). Al finalizar el análisis comparativo se obtuvieron 2,306 genes sobre-expresados, 1,606 sub-expresados y 29,385 sin cambios en la población CITCM vs la no CITCM (Figura 21). Dentro de las diez moléculas más sobre-expresadas (Tabla 11) observamos dos miRNAs que ya han sido previamente implicados como promotores de la población CITCM, miR-21 y miR-221, mientras que dentro de las diez más sub-expresadas encontramos a ATRX, gen de la familia SWI/SNF remodeladora de cromatina que mantiene papeles importantes en los mecanismos de reparación y recombinación del ADN, se encuentra involucrado en varios tipos de cáncer (incluyendo osteosarcoma, glioblastoma, neuroblastoma y tumores neuroendócrinos del páncreas) y enfermedades (alfa talasemia y síndrome de retraso mental) y recientemente su pérdida ha sido asociada a un crecimiento tumoral acelerado e inestabilidad genética<sup>23,83,129,130</sup>. Otras moléculas diferencialmente expresadas y/o

reguladas y que son de nuestro particular interés como CD44, CD24, ALDH y EpCam (con la cual podríamos validar el aislamiento previo con MACS<sup>®</sup>) puesto que se encuentran relacionadas con las poblaciones CITCM las podemos observar en la Tabla 12.

**Figura 20.** Normalización y PCA de los datos de expresión de MDA-MB-231.



Normalización (A) y PCA (B) de los datos generados en el estudio de expresión con los microarreglos GeneChip<sup>™</sup> Human Gene ST 1.0 de de Affymetrix<sup>®</sup> utilizando las muestras aisladas en la primera etapa A con MACS<sup>®</sup> (cuatro F<sup>+</sup> de MDA-MB-231 (azules) y cuatro F<sup>-</sup> de MDA-MB-231 (moradas)). La normalización se realizó mediante cuantiles y RMA con PGS al igual que el PCA.

Por último, los resultados obtenidos con PGS se introdujeron en la plataforma bioinformática IPA, los resultados mas relevantes contemplados fueron las predicciones del estado de activación de los reguladores transcripcionales (Tabla 13) ampliamente conocidos como factores fundamentales de regulación en las poblaciones CITCM así como la representación de la vía de sobre-expresión de EpCam con sus diversas interacciones (directas e indirectas) moleculares (Figura 22).

**Figura 21.** Diagrama de Venn del análisis de expresión de CITCM.

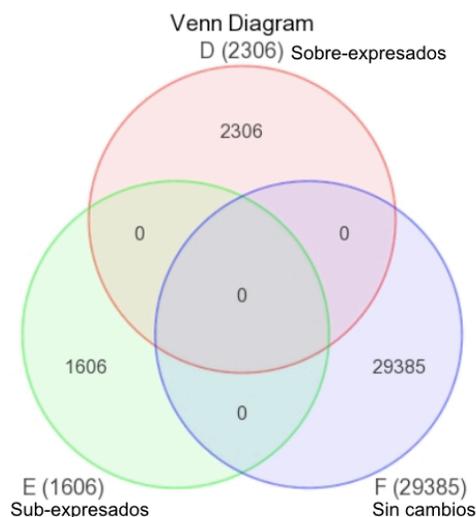


Diagrama de Venn generado con los resultados del análisis comparativo de expresión entre las CITCM y no CITCM. En el conjunto rojo o D observamos las poblaciones sobre-expresadas, en la verde o E los sub-expresados y en la F o azul aquellos que no sufrieron cambios.

**Tabla 11.** Los 10 genes más sobre y sub-regulados en la población CITCM ( $F^+$  de MDA-MB-231).

### Genes sobre-regulados

Fold Change up-regulated	
Moleculas	Expression Value
mir-21	4.877
NDUFA2	3.322
CXCL10	3.079
FOS	3.035
RNU4-1	3.016
mir-221	2.504
NDUFA3	2.485
CYB5A	2.423
RPS28	2.393
MT1A	2.389

## Genes sub-regulados

Fold Change down-regulated	
Molecules	Expression Value
ATRX	-3.946
CENPE	-2.810
AKAP9	-2.637
EEA1	-2.613
CEP290	-2.555
TPR	-2.485
CENPF	-2.313
SYNE2	-2.268
EIF5B	-2.255
GOLGB1	-2.254

**Tabla 12.** Otras moléculas diferencialmente expresadas en la población CITCM observadas posterior al análisis de expresión con PGS.

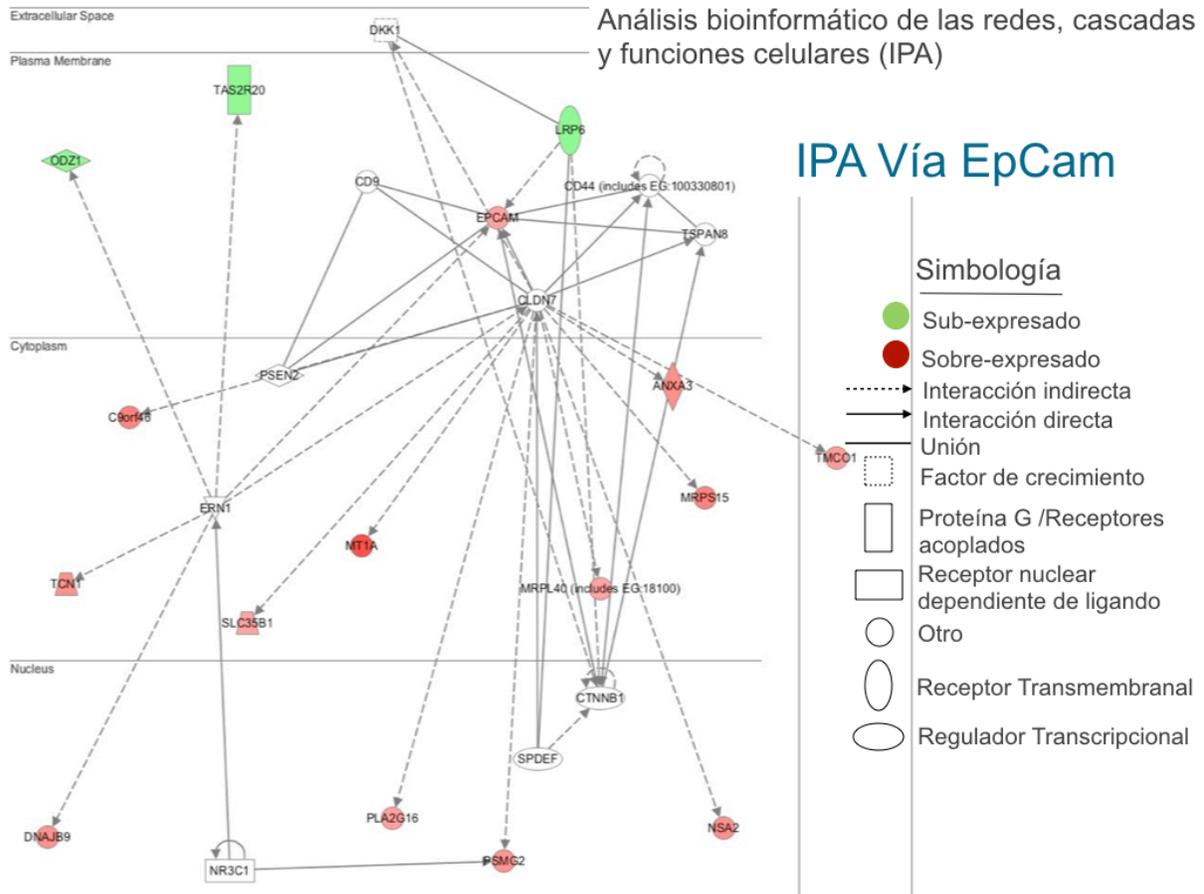
Symbol	Affymetrix	Fold Change	Networks	Location	Type(s)
ALDH7A1	8113773	1.342		Cytoplasm	Enzyme
ANXA3	8095986	1.609	25	Cytoplasm	Enzyme
ANXA4	8042468	1.507		Plasma Membrane	Other
APAF1	7957759	-1.323	1	Cytoplasm	Other
BCL2A1	7990818	1.416	12	Cytoplasm	Other
BRCA2	7968484	-1.983	7	Nucleus	Transcription Regulator
CASP1	7951397	1.361	14	Cytoplasm	Peptidase
CASP4	7951372	1.329	14	Cytoplasm	Peptidase
CASP8AP2	8121118	-1.679	2	Nucleus	Transcription Regulator
CD24			17	Plasma Membrane	Other
CD44			25	Plasma Membrane	Other
CDK5	8143850	1.386	11	Nucleus	Kinase
CDK6	8140955	-1.896	12	Nucleus	Kinase

CLDN7			25	Plasma Membrane	Other
DICER1	7981111	-1.378	5	Cytoplasm	Enzyme
EGFR			23	Plasma Membrane	Kinase
EPCAM	8098439	1.334	25	Plasma Membrane	Other
IGF1R	7986359	-1.552	13	Plasma Membrane	Transmembrane Receptor
IGF2R	8123181	-1.370	10	Plasma Membrane	Transmembrane Receptor
IL6	8131803	1.558	19	Extracellular Space	Cytokine
ITGA2	8105267	-1.460	14,21,23	Plasma Membrane	Other
MMP3	7951284	1.466	19	Extracellular Space	Peptidase
MMP1	7951271	1.300	19,22	Extracellular Space	Peptidase
NCAM2	8067985	-1.315	1	Plasma Membrane	Other
NDUFB3	8047372	1.903		Cytoplasm	Enzyme
NGRN	7985983	1.364		Nucleus	Other
SEMA3A	8140668	-1.569	2	Extracellular Space	Other
SMAD1	8097657	1.328	9	Nucleus	Transcription Regulator
TIMP1	8167185	1.627	19,23	Extracellular Space	Other

**Tabla 13.** Predicción generada en el programa IPA del estado de activación de los reguladores transcripcionales en las CITCM.

Transcription Regulator	Predicted Activation State	Regulation z-score	p-value of overlap	Target molecules in dataset dataset	Molecular Type
NFkB (complex)	Activated	3.679	2.65E-02	BAK1, BCL2A1, CCL2, CCL20, CCL3, CC	complex
EZH2	Activated	2.712	5.04E-02	C15orf48, CSF2, CSF3, CXCL10, CXCL2	transcription regulator
FOXL2	Activated	2.662	2.12E-03	ATF3, BCL2A1, CCL20, CCL3, CCL3L1/	transcription regulator
IRF3	Activated	2.460	8.39E-03	CCL2, CCL3, CCL5, CXCL10, IFI6, IL29	transcription regulator
NANOG	Activated	2.052	4.49E-01	MAP1B, NRIP1, PRRC2C, TBC1D3F	transcription regulator
Hdac	Inhibited	-2.119	2.40E-01	ATF3, DDIT3, FOS, MT1E, MT1F, MT1H	group
Nr1h	Inhibited	-2.330	4.32E-03	CCL3, CCL5, CSF2, CSF3, IL6, PTGS2	group

**Figura 22.** Vía de EpCam con sus interacciones y uniones observadas mediante el programa IPA.



Continuamos con el análisis de los datos generados con RNAseq, los resultados obtenidos por ambas tecnologías mostraron resultados semejantes, sin embargo los más relevantes de esta plataforma fueron el enriquecimiento (Tabla 14) en cascadas de transducción relacionadas con funciones celulares y moleculares como muerte y ciclo celular, proliferación y crecimiento, en las vías canónicas de señalización relacionadas con mecanismos moleculares de cáncer y receptores de estrógenos, la activación e inhibición de los factores de transcripción EPAS1, NANOG, OTX2, MXD1, SOX5 y SOX6 (Figura 23). En la Tabla 15 observamos las diez moléculas más sobre y sub reguladas calculadas con

esta plataforma, mientras que en la Tabla 16 (A -C) observamos genes con diversos niveles de expresión diferencial seleccionados y agrupados de esta manera por tener algún tipo de relación con la población CITCM y/o por ser de nuestro particular interés.

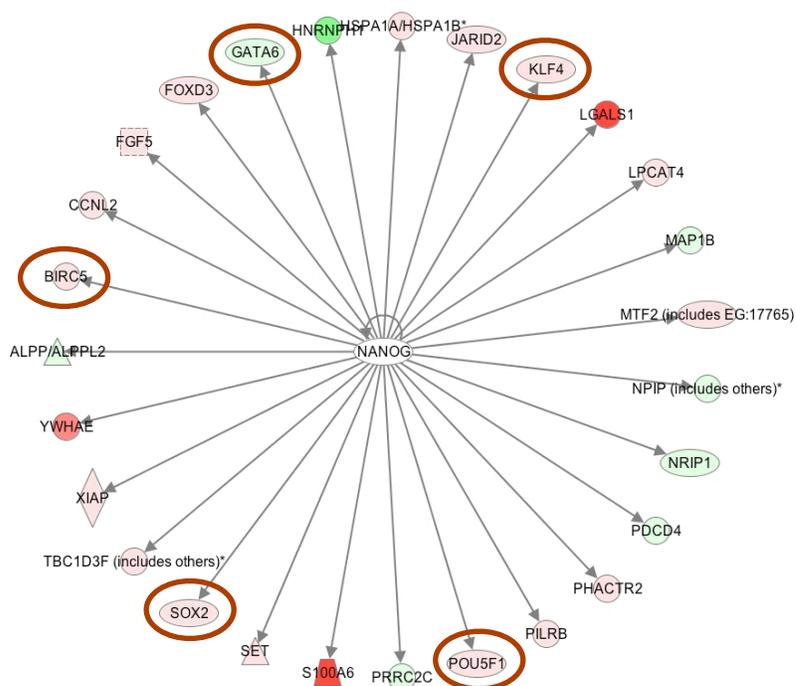
**Tabla 14.** *Vías canónicas de señalización y enriquecimiento de funciones moleculares y celulares observadas en las poblaciones CITCM.*

Top Canonical Pathways		
Name	p-value	Ratio
Molecular Mechanisms of Cancer	2.29E-27	301/378 (0.796)
Protein Ubiquitination Pathway	1.12E-26	235/269 (0.874)
Purine Metabolism	3.6E-16	227/393 (0.578)
Estrogen Receptor Signaling	4.72E-15	118/136 (0.868)
Huntington's Disease Signaling	7.04E-15	187/236 (0.792)

Molecular and Cellular Functions		
Name	p-value	# Molecules
Gene Expression	1.56E-84 - 1.13E-09	2200
Cell Death	2.08E-68 - 1.60E-05	2913
Cellular Growth and Proliferation	6.81E-50 - 1.01E-05	2931
Cell Cycle	9.32E-48 - 1.53E-05	1466
Cellular Assembly and Organization	1.07E-42 - 1.53E-05	1713

El análisis se realizó con los datos de RNAseq y con la herramienta bioinformática IPA. Observamos en la tabla superior las vías canónicas de señalización y en la inferior las funciones moleculares y celulares principales en nuestra población CITCM con sus respectivos valores y número de moléculas involucradas.

**Figura 23. Reguladores transcripcionales y factor de transcripción NANOG en las CITCM.**



© 2000-2012 Ingenuity Systems, Inc. All rights reserved.

Transcription Regulator	Predicted Activation State	Regulation z-score	p-value of overlap	Target molecules in dataset	Molecular Type
EPAS1	Activated	2.235	8.52E-04	ABCF2,ABCG2,A	transcription regulator
NANOG	Activated	2.211	2.78E-01	ALPP/ALPPL2,B	transcription regulator
MXD1	Activated	2.144	7.04E-02	BCL6,EDN1,EIF	transcription regulator
OTX2	Activated	2.028	8.40E-06	CDH4,DKK1,EF	transcription regulator
SOX6	Inhibited	-2.069	1.75E-01	BMP6,CCND1,F	transcription regulator
SOX5	Inhibited	-2.519	2.54E-01	BMP6,FGFR3,HS	transcription regulator

Predicción del estado de activación de los reguladores transcripcionales en las CITCM (tabla) y representación gráfica del factor de transcripción NANOG activado con sus diferentes interacciones moleculares sobre (en rojo) y sub-expresadas (en verde). Circulamos en rojo cinco moléculas de nuestro interés con expresión diferencial e interacción directa con NANOG (el análisis se realizó con los datos de RNAseq y con la herramienta bioinformática IPA).

**Tabla 15.** Los 10 genes más sobre y sub-regulados en la población CITCM.

### Genes sobre-regulados

Log Ratio up-regulated	
Moleculas	Expression Value
TMSB10/TMSB4X*	4954.937
S100A6	3486.005
RPL35A	3223.666
RPS3A	2624.782
RPL41	2531.259
RPLP0	2325.474
UBB	2312.043
RPS11	2295.787
RPL26	2287.982
RPS3	2001.941

### Genes sub-regulados

Log Ratio down-regulated	
Moleculas	Expression Value
RPS18	-78717.672
GAPDH	-12336.940
EEF1G	-10593.910
EIF5A	-7417.638
EEF1A1	-6032.692
ACTG1	-6024.935
RPS2	-5302.550
CALR	-5060.146
EIF4A1	-3475.361
TPI1	-3163.741

El análisis se realizó con los datos de RNAseq y con la herramienta bioinformática CLCBio e IPA. Los niveles de expresión corresponden al FC.

**Tabla 16. Genes diferencialmente expresados en la población CITCM.**

**A**

Symbol	Entrez Gene Name	Log Ratio	Fold Change
ABCA1	ATP-binding cassette	1.922	∞
ABCA3	ATP-binding cassette	1.415	∞
ABCA4	ATP-binding cassette	0.344	∞
ABCA7	ATP-binding cassette	0.687	∞
ABCA9	ATP-binding cassette	0.029	∞
ABCA12	ATP-binding cassette	0.019	∞
ABCB10	ATP-binding cassette	1.756	1.420
ABCC1	ATP-binding cassette	7.430	4.025
ABCC9	ATP-binding cassette	0.085	∞
ABCC4	ATP-binding cassette	3.690	∞
ABCD3	ATP-binding cassette	15.028	4.540
ABCD4	ATP-binding cassette	16.076	4.148
ABCF1	ATP-binding cassette	0.168	1.003
ABCG2	ATP-binding cassette	2.481	∞
ABCG4	ATP-binding cassette	0.046	∞
ACTA2	actin, alpha 2, smooth muscle	1.266	∞
AKT1	v-akt murine thymoma	-265.37	-4.239
AKT2	v-akt murine thymoma	14.042	∞
AKT3	v-akt murine thymoma	3.255	2.482
ALDH2	aldehyde dehydrogenase	19.755	∞
ALDH16A	aldehyde dehydrogenase	2.619	1.252
ALDH18A	aldehyde dehydrogenase	3.232	1.077
ALDH1A3	aldehyde dehydrogenase	3.042	∞
ALDH1B1	aldehyde dehydrogenase	22.464	∞
ALDH1L2	aldehyde dehydrogenase	0.072	∞
ALDH3A1	aldehyde dehydrogenase	1.917	∞
ALDH3A2	aldehyde dehydrogenase	2.623	1.208
ALDH3B1	aldehyde dehydrogenase	3.357	1.610

ALDH4A1	aldehyde dehydrogenase	1.862	∞
ALDH5A1	aldehyde dehydrogenase	1.813	∞
ALDH6A1	aldehyde dehydrogenase	3.550	∞

## B

Symbol	Entrez Gene Name	Log Ratio	Fold Change
BCL2	B-cell CLL/lymphoma 2	0.599	∞
BIRC2	baculoviral IAP repeat c	29.535	1.343
BIRC3	baculoviral IAP repeat c	20.933	2.698
BIRC5	baculoviral IAP repeat c	75.016	5.188
EPCAM	epithelial cell adhesion	13.429	∞
IL6	interleukin 6 (interferon,	64.807	2.587
IL8	interleukin 8	475.901	6.034
NOTCH2	notch 2	3.116	1.381
NOTCH3	notch 3	0.156	∞
NOTCH4	notch 4	0.027	∞
POU5F1	POU class 5 homeobox 1	0.082	∞
SNAI1	snail homolog 1 (Drosop	0.949	∞
SNAI2	snail homolog 2 (Drosop	1.072	1.141
SNAI3	snail homolog 3 (Drosop	0.421	∞
SOX2	SRY (sex determining re	0.072	∞
VEGFA	vascular endothelial gro	-15.380	-1.775
VEGFB	vascular endothelial gro	12.667	1.878
VEGFC	vascular endothelial gro	27.830	∞
VGF	VGF nerve growth factor	0.847	∞
VIM	vimentin	-1904.8	-2.258
WNT3	wingless-type MMTV int	1.204	∞
WNT11	wingless-type MMTV int	0.094	∞
WNT10A	wingless-type MMTV int	0.455	∞
WNT10B	wingless-type MMTV int	0.989	∞
WNT2B	wingless-type MMTV int	0.133	∞

WNT5A	wingless-type MMTV int	0.062	∞
WNT5B	wingless-type MMTV int	5.850	∞
WNT7B	wingless-type MMTV int	4.375	∞
WNT8B	wingless-type MMTV int	0.090	∞

### C

Symbol	Entrez Gene Name	Log Ratio	Fold Change
BMI1	BMI1 polycomb ring fin	4.024	1.856
BRCA1	breast cancer 1, early	4.401	2.985
BRCA2	breast cancer 2, early	-1.518	-2.155
CDH1	cadherin 1, type 1, E-cadheri	0.262	∞
CDH2	cadherin 2, type 1, N-cadheri	0.083	∞
CLDN7	claudin 7	-9.109	-1.626
CLDN15	claudin 15	-5.899	-5.590
CXCR4	chemokine (C-X-C moti	4.585	∞
EGFR	epidermal growth factor	2.339	1.238
FGF1	fibroblast growth factor	2.272	∞
HSPB1	heat shock 27kDa prot	25.346	1.705
IGF2	insulin-like growth facto	0.086	∞
MUC1	mucin 1, cell surface a	-25.821	-53.660
PROCR	protein C receptor, end	9.091	1.409
PTEN	phosphatase and tensi	6.582	3.265
TGFB2	transforming growth fa	5.890	3.176
TGFB3	transforming growth fa	0.170	∞
TGFB1 (include	transforming growth fa	27.165	4.707

**A, B y C.** No se realizó corte específico de FC por ser de carácter exploratorio en la población CITCM. Estos genes fueron seleccionados y agrupados de esta forma por ser de nuestro mayor interés durante el análisis generado con los datos de RNAseq por ejemplo en la **(A)** familia de ABCs, AKTs y ALDHs, **(B)** familia de BIRCs, NOTCH, WNT, etc y **(C)** familia de claudinas y TGFBs entre otras.

---

### 13.5.2 PRIMERA ETAPA B

Se utilizaron 1.5µg de ARN total de dos réplicas biológicas de las poblaciones CITCM (CD44<sup>+</sup>CD24<sup>-</sup>EpCam<sup>+</sup>) y no CITCM (CD44<sup>+</sup>CD24<sup>-</sup>EpCam<sup>-</sup>, CD44<sup>+</sup>CD24<sup>+</sup>EpCam<sup>+</sup>, CD44<sup>+</sup>CD24<sup>+</sup>EpCam<sup>-</sup>, CD44<sup>-</sup>CD24<sup>-</sup>EpCam<sup>-</sup>, CD44<sup>-</sup>CD24<sup>-</sup>EpCam<sup>+</sup>, CD44<sup>-</sup>CD24<sup>+</sup>EpCam<sup>+</sup> y CD44<sup>-</sup>CD24<sup>+</sup>EpCam<sup>-</sup>) de las líneas celulares MDA-MB-231, MDA-MB-468 y BT-20 aisladas previamente por FACS para realizar el estudio de expresión mediante RNAseq. Los procedimientos técnicos y analíticos se realizaron de acuerdo a lo descrito en el módulo de metodología. Al finalizar el estudio individual de las CITCM (línea celular por separado) se obtuvieron 9,896 genes diferencialmente expresados, de los cuales 507 compartían las CITCM provenientes de las tres líneas celulares, 1,992 se observaron en las CITCM de MDA-MB-231, 2299 en las de BT-20 y 5,605 en las CITCM de MDA-MB-468 (Figura 24). En el análisis general (CITCM vs no CITCM provenientes de las tres líneas celulares) se obtuvieron 1,032 genes sobre-regulados y 8,711 sub-regulados, en las Tablas 17 y 18 observamos los 50 genes más sobre (en donde observamos a ABCC12 y WNT1) y sub-expresados (a BIRC7 y MUC2).

**Figura 24.** Diagrama de Venn del análisis de expresión de las CITCM por línea celular.

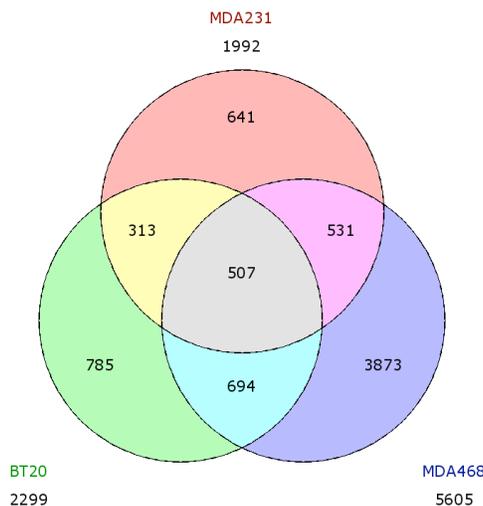


Diagrama de Venn generado con los resultados del análisis comparativo de expresión entre las CITCM y no CITCM de cada línea celular por separado. En el conjunto rojo observamos las poblaciones diferencialmente expresadas en las CITCM de MDA-MB-231, en la verde en las CITCM de BT-20 y en la azul en las CITCM de MDA-MB-468.

**Tabla 17.** Las 50 moléculas más sobre-expresadas en las CITCM provenientes de las tres líneas celulares.

Feature ID	Experiment - Fold Change (original values)	Feature ID	Experiment - Fold Change (original values)
1.UPK1B	1.7977E+308	26.AQP5	8.440999084
2.TPSD1	1.7977E+308	27.CLEC18C	8.389917159
3.TPH2	1.7977E+308	28.ALPL2	8.299904867
4.TNNT3	1.7977E+308	29.TRDC	8.201272592
5.TMEM150B	1.7977E+308	30.GDA	7.893067473
6.PCDHAC2	1.7977E+308	31.C1QL2	7.781382542
7.PCDHA5	1.7977E+308	32.FGF2BP2	7.729549527
8.LY6G6F	1.7977E+308	33.MYO1A	7.693372689
9.KRT74	1.7977E+308	34.ABCC12	7.603105046
10.KIAA1257	1.7977E+308	35.WNT1	7.467929183
11.HIST1H2AJ	1.7977E+308	36.OR4M2	7.379320398
12.HEPACAM2	1.7977E+308	37.PSG2	7.216440549
13.HCK	1.7977E+308	38.KIF26B	6.89950376
14.GRK7	1.7977E+308	39.TMEM59L	6.838438013
15.EMX1	1.7977E+308	40.GPR146	6.808691811
16.CEACAM20	1.7977E+308	41.KL	6.796838726
17.C6orf208	1.7977E+308	42.SLIT3	6.767630688
18.C4orf37	1.7977E+308	43.CLCA4	6.524686788
19.C3orf36	1.7977E+308	44.PRR9	6.502885485
20.C15orf55	1.7977E+308	45.PRLH	6.427409821
21.C14orf178	1.7977E+308	46.ELAVL3	6.386140473
22.C6orf15	11.50585526	47.CRYM	6.321684806
23.C14orf183	10.46397445	48.SRRM5	6.268668151
24.LGALS7B	8.969568392	49.TMPRSS11E	6.242415256
25.KRT20	8.735917946	50.DPT	6.054992525

Los pseudogenes y ncRNAs fueron excluidos de las listas.

**Tabla 18.** Las 50 moléculas más sub-expresadas en las CITCM provenientes de las tres líneas celulares.

Feature ID	Experiment - Fold Change (original values)	Feature ID	Experiment - Fold Change (original values)
1.WDR86	-1.7977E+308	26.CALML6	-1.7977E+308
2.UGT3A2	-1.7977E+308	27.C8orf56	-1.7977E+308
3.SLC17A4	-1.7977E+308	28.C4BPA	-1.7977E+308
4.SHANK1	-1.7977E+308	29.C2orf62	-1.7977E+308
5.PRSS54	-1.7977E+308	30.BIRC7	-1.7977E+308
6.P5.8	-1.7977E+308	31.ART4	-1.7977E+308
7.NAALADL1	-1.7977E+308	32.APLNR	-1.7977E+308
8.MYRIP	-1.7977E+308	33.AMY1C	-12.03705157
9.MRGPRF	-1.7977E+308	34.FGG	-11.70541529
10.MLNR	-1.7977E+308	35.LBP	-11.52860632
11.MFAP5	-1.7977E+308	36.SERPINB13	-11.32459703
12.LGSN	-1.7977E+308	37.GRRP1	-10.04335096
13.ISLR2	-1.7977E+308	38.DPCR1	-9.191441591
14.INE2	-1.7977E+308	39.PAX2	-8.866613152
15.IGLV1-51	-1.7977E+308	40.CSF3R	-8.634179429
16.HIST1H4K	-1.7977E+308	41.OTUD7A	-8.55249568
17.HIST1H4D	-1.7977E+308	42.STAC2	-6.014095374
18.GSG1L	-1.7977E+308	43.TAS1R1	-5.955007107
19.GRAP	-1.7977E+308	44.DUSP26	-5.940672577
20.GGTLC2	-1.7977E+308	45.GRAPL	-5.894330781
21.FMO1	-1.7977E+308	46.HIST2H2AC	-5.799575353
22.DLEU2L	-1.7977E+308	47.MAP1LC3A	-5.708699926
23.CSDC2	-1.7977E+308	48.KCNE4	-5.702587303
24.CEACAM16	-1.7977E+308	49.CETP	-5.652153924
25.CDH17	-1.7977E+308	50.MUC2	-5.602856166

Los pseudogenes y ncRNAs fueron excluidos de las listas.

---

Al finalizar el estudio de expresión diferencial juntamos los resultados con los obtenidos con el microarreglo Infinium HumanMethylation450K BeadChip® para observar si habían correlaciones entre los niveles de metilación y expresión, particularmente en genes relacionados con troncalidad, cáncer de mama, cáncer y/o alguno de nuestro especial interés. Encontramos 930 genes compartidos entre ambos estudios (Figura 25), de estos, veintinueve fueron los que más llamaron nuestra atención (Tabla 19). Posteriormente, realizamos el análisis de enriquecimiento de genes y predicción de redes y cascadas con IPA y GSEA con los cuales se encontró a las vías de señalización de Wnt/ $\beta$  catenina, Notch, de pluripotencia de las células troncales embrionarias humanas y de regulación de cáncer de mama por estatmina1 enriquecidas (Figura 26 A-D). Así mismo, observamos activados a los factores de transcripción SOX2, YY1 y POU5F1 (OCT3/4), en las Figura 27 A-C podemos observar la representación gráfica de estos con sus diversas interacciones moleculares, por ejemplo, SOX2 con HOXC4 y POU5F1 (OCT3/4 mantiene esta misma relación con SOX2) y YY1 con ERBB2 y Fas, entre otros. Por último, el resultado más relevante contemplado con la herramienta GSEA fue el enriquecimiento de la vía de Notch (Figura 28).

**Figura 25.** Diagrama de Venn del análisis comparativo de expresión y metilación.

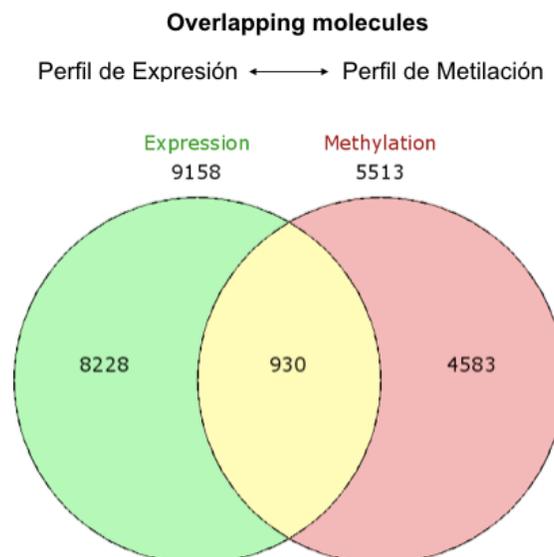


Diagrama de Venn generado con los resultados obtenidos del análisis comparativo de metilación y expresión entre las poblaciones CITCM y no CITCM provenientes de las tres líneas celulares. En el conjunto rojo observamos las poblaciones diferencialmente metiladas en las CITCM, en el verde las poblaciones diferencialmente expresadas en las CITCM y en el amarillo las moléculas que comparten ambos estudios.

**Tabla 19.** Genes de interés del análisis comparativo de expresión y metilación.

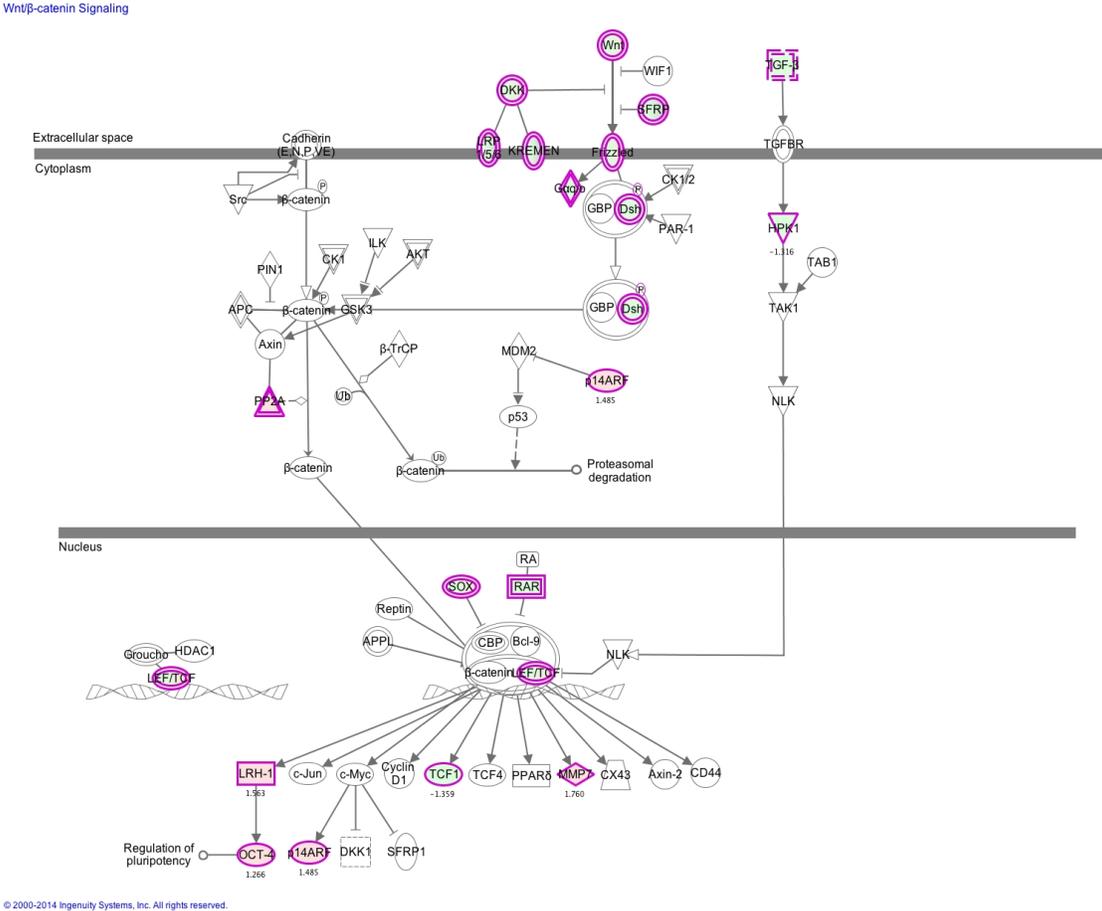
ID	Fold Change	Average_B
ALDH16A1	-1.243	0.05
ALDH3B1	-1.222	0.89
BIRC5	1.241	0.029
CASP4	-1.252	0.78
CAV1	1.3	0.03
CLDN1	1.202	0.0371
ERBB2	-1.209	0.05
FGFR3	-1.365	0.989
FOXO3	-1.203	0.85
FZR1	-1.301	0.024
GATA2	-1.306	0.98
GLI1	2.078	0.064
GLIPR2	-1.320	0.089
HOXA10	-1.278	0.96
HOXA2	1.536	0.986
HOXA9	2.018	0.83
HOXC4	-1.389	0.88
JAG2	-1.285	0.93
PCLO	1.2	0.02
PIK3R1	-1.585	0.92
POU5F1	1.266	0.98
SMO	-1.567	0.02
SMUG1	-1.202	0.91
SOX15	1.299	0.82
TGFB2	∞	0.04
TRIM22	-2.490	0.9
WNT1	7.468	0.91

WNT3A	-2.813	0.99
WNT7B	-1.251	0.88

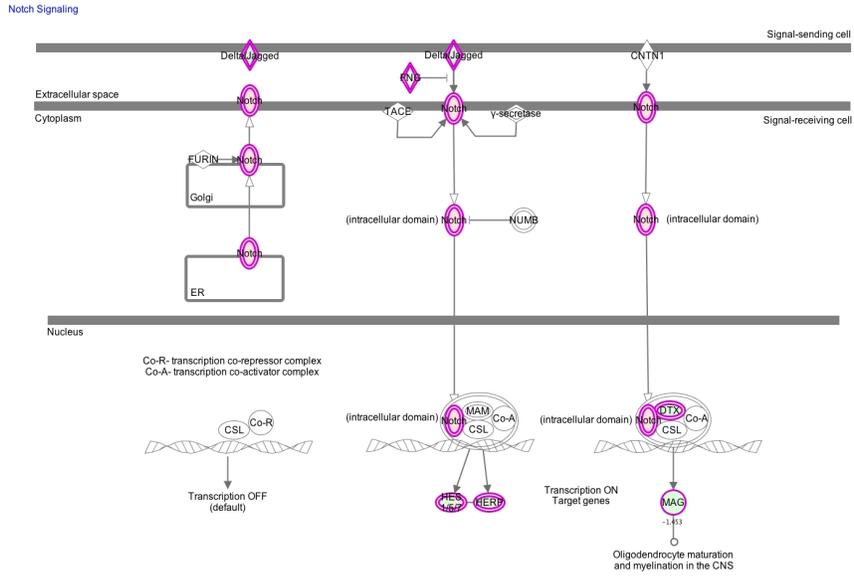
Moléculas con sus respectivos niveles de expresión (FC) y metilación (Average B) seleccionados por su relación con troncalidad, cáncer de mama y/o por ser de nuestro particular interés posterior al análisis de correlación entre los resultados del estudio de expresión con los resultados del estudio de metilación.

**Figura 26.** Vías de señalización enriquecidas en las CITCM.

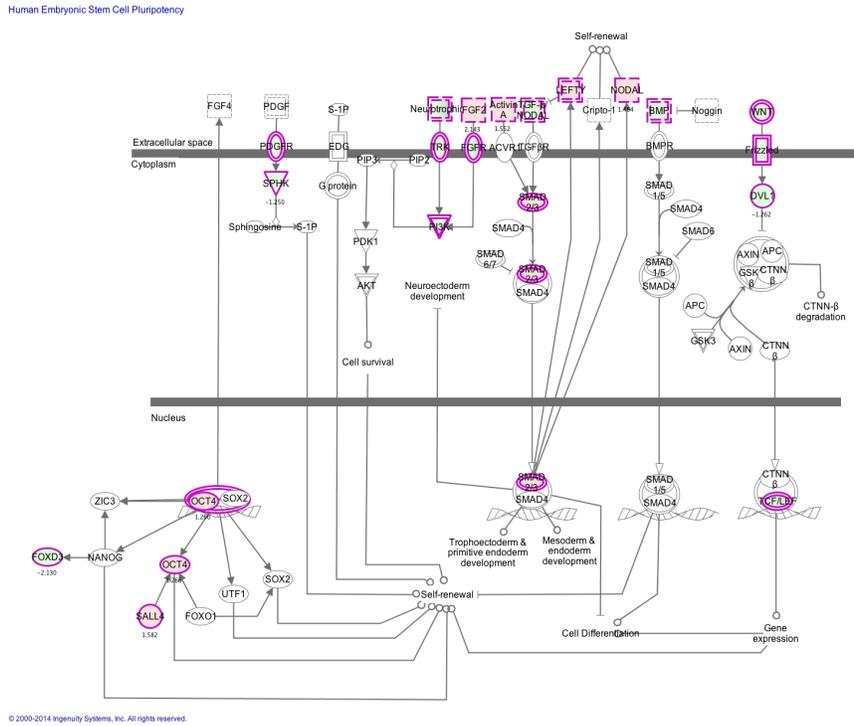
**A**



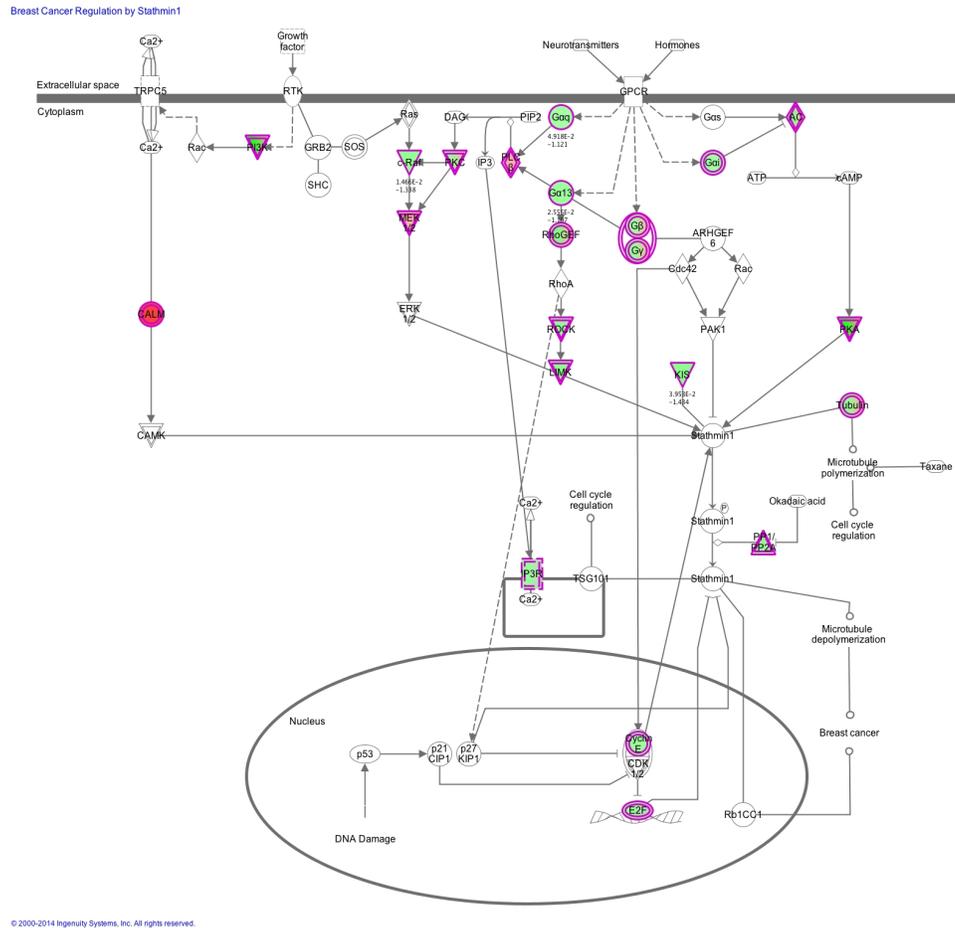
**B**



**C**



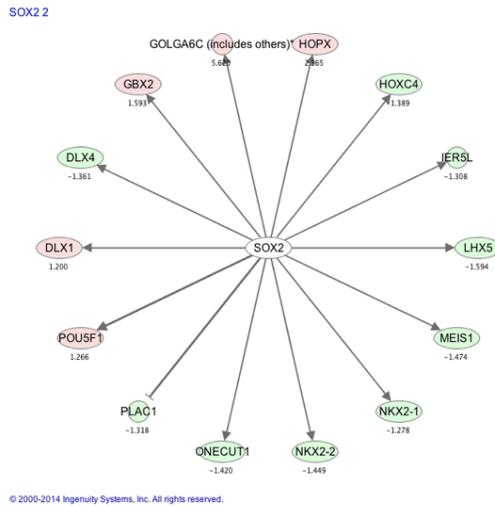
D



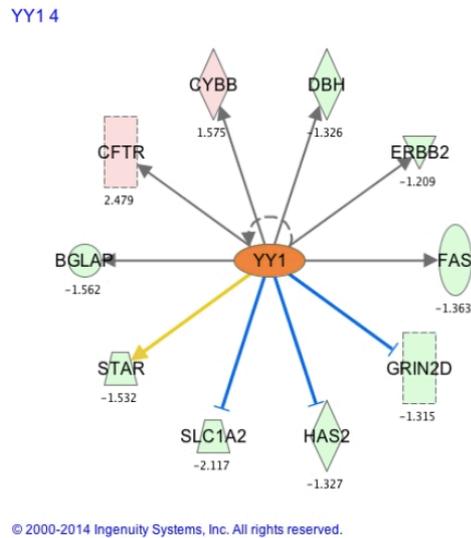
Imágenes generadas en el programa IPA del enriquecimiento de las vías de señalización de Wnt/ $\beta$ -catenina (A), Notch (B), de pluripotencia de las células troncales embrionarias humanas (C) y de regulación de cáncer de mama por estatmina1 (D).

**Figura 27.** Factores de transcripción activados en el análisis comparativo de expresión y metilación.

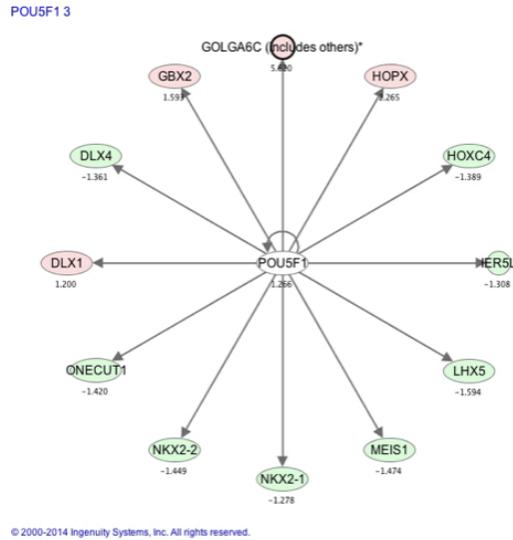
**A**



**B**



C



Interacción molecular de los factores de transcripción activados SOX2 (A), YY1 (B) y POU5F1 (C), predichos por la plataforma bioinformática IPA y generadas posterior al análisis del estudio de correlación entre los resultados de expresión y metilación.

**Figura 28.** GSEA del análisis comparativo de expresión y metilación.

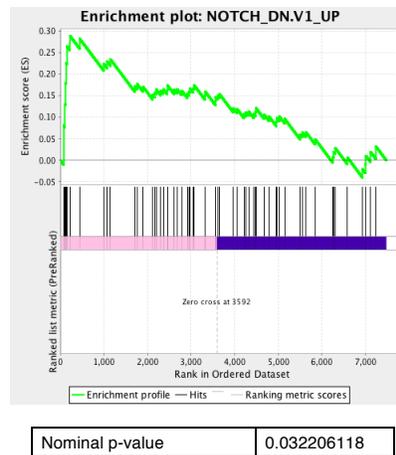


Imagen generada con GSEA del enriquecimiento de la vía de Notch lograda con un valor de p-nominal menor a 5%.

---

## 14. DISCUSIÓN

El cáncer de mama es considerado hoy en día la principal causa de muerte de mujeres a nivel mundial y el tipo de cáncer diagnosticado con mayor frecuencia en esta población proveniente de países desarrollados como en vías de desarrollo con aproximadamente un millón de nuevos casos anuales<sup>40,41,51</sup>. La introducción de nuevas tecnologías para el estudio del cáncer de mama, como los microarreglos de expresión y la secuenciación de nueva generación (NGS) han permitido generar grandes avances entorno a la mejor comprensión de la biología y complejidad tumoral, al reconocimiento de las diversas alteraciones genéticas, al pronóstico, a la respuesta al tratamiento y al reconocimiento de los diversos perfiles moleculares con los cuales se ha desarrollado una mejor clasificación tumoral que ha logrado la generación de nuevos blancos terapéuticos y que a su vez nos acerca más a la posibilidad del desarrollo de un tratamiento personalizado<sup>51,53,56,57,99,101,110,111,131</sup>.

Por las propiedades de autorenovación, diferenciación, alta resistencia (entre otras) que poseen las CITCM, estas han sido propuestas como la población celular responsable de la reincidencia, metástasis y heterogeneidad tumoral de este tipo de cáncer. Para su estudio y caracterización, las CITCM pueden ser aisladas tanto de líneas celulares como de muestras tisulares utilizando una combinación específica y ampliamente reportada de marcadores de superficie como CD44, CD24, ALDH, EpCam, CD133, etc<sup>20,32,34,65,73,85</sup>.

Conocer y analizar la estructura del genoma y los perfiles de expresión y metilación de esta población celular nos brindará nueva información para poder entender el origen, naturaleza y los mecanismos de acción de las CITCM, con lo cual se podrán mejorar las herramientas diagnósticas, pronósticas y terapéuticas del cáncer de mama.

---

Por lo ya mencionado, este trabajo se concentró en el estudio y análisis de los diversos perfiles genómicos de las CITCM provenientes de líneas celulares de cáncer de mama.

## **14.1 AISLAMIENTO, CARACTERIZACIÓN FENOTÍPICA Y FUNCIONAL Y ANÁLISIS GENÓMICO DE LAS CITCM A PARTIR DE LÍNEAS CELULARES**

Las CITCM para su estudio han sido previamente aisladas tanto de líneas celulares como de muestras tisulares utilizando varias metodologías y combinaciones de marcadores de superficie celular mencionadas ya anteriormente. Para esta investigación seleccionamos MACS<sup>®</sup> y citometría de flujo con FACS como herramientas de separación. El marcaje individual con EpCam (con el cual obtuvimos dos poblaciones, la fracción positiva y la negativa a EpCam), el doble marcaje con CD44 y CD24 (obteniendo las cuatro poblaciones CD44<sup>+</sup>CD24<sup>+</sup>, CD44<sup>-</sup>CD24<sup>-</sup>, CD44<sup>+</sup>CD24<sup>-</sup> y CD44<sup>-</sup>CD24<sup>+</sup>) y por último el triple marcaje con EpCam, CD44 y CD24 (con la cual obtuvimos las ocho poblaciones CD44<sup>+</sup>CD24<sup>-</sup>EpCam<sup>-</sup>, CD44<sup>+</sup>CD24<sup>+</sup>EpCam<sup>+</sup>, CD44<sup>+</sup>CD24<sup>+</sup>EpCam<sup>-</sup>, CD44<sup>-</sup>CD24<sup>-</sup>EpCam<sup>-</sup>, CD44<sup>-</sup>CD24<sup>-</sup>EpCam<sup>+</sup>, CD44<sup>-</sup>CD24<sup>+</sup>EpCam<sup>+</sup>, CD44<sup>-</sup>CD24<sup>+</sup>EpCam<sup>-</sup> y CD44<sup>+</sup>CD24<sup>-</sup>EpCam<sup>+</sup>) fueron los inmunofenotipos utilizados para purificar y enriquecer nuestras poblaciones CITCM.

### **14.1.1 PRIMERA ETAPA A**

Como comentamos previamente, en un inicio (primera etapa A), mi investigación estaba dirigida al enriquecimiento de las CITCM únicamente de dos líneas celulares utilizando los siguientes marcadores y sistemas de aislamiento:

- línea celular de origen basal MDA-MB-231 con MACS<sup>®</sup> y perlas paramagnéticas acopladas a EpCam.

- 
- línea celular de origen luminal MCF-7 con FACS y CD44/CD24 acoplados a diferentes fluoróforos.

Los marcadores de superficie celular empleados para el aislamiento de las CITCM en esta primera etapa A fueron seleccionados dependiendo del tipo de línea celular y en base al enriquecimiento de esta población previamente reportado por otros autores (fracción positiva a EpCam en MDA-MB-231 y fracción CD44<sup>+</sup>CD24<sup>-</sup> en MCF-7)<sup>59,84</sup>. Las separaciones (Figuras 7 y 8) y las pruebas funcionales (Figuras 10 y 12) de las CITCM provenientes de ambas líneas celulares se llevaron a cabo sin eventualidades, estas últimas nos permitieron comprobar tanto el potencial tumorigénico como las capacidades de autorenovación y diferenciación característicos de esta población celular<sup>2,4,10,47,62,65</sup>.

Posteriormente, nuestras poblaciones purificadas CITCM y no CITCM procedentes de ambas líneas celulares, fueron utilizadas para realizar la extracción de ADN y ARN, sin embargo, únicamente pudimos obtener las cantidades necesarias para iniciar los tres estudios (expresión, genotipificación y metilación) de las CITCM y no CITCM de la línea celular MDA-MB-231 mientras que de las de MCF-7 solo se logró realizar el estudio del perfil de metilación. El rendimiento de la extracciones, así como la integridad y pureza de las muestras de ARN y ADN que obtuvimos de las cuatro poblaciones provenientes de la separación de MCF-7 con CD44 y CD24 (con excepción de las muestras utilizadas para el estudio de metilación) no fueron nunca las óptimas (se intentó iniciar con un mayor número de células, cambiar los kits de extracción, utilizar otro citómetro, entre otros) sobretodo de las fracciones CD44<sup>+</sup>CD24<sup>+</sup> y CD44<sup>+</sup>CD24<sup>-</sup> (CITCM), las dos poblaciones más limitadas (Figura 8) dentro de esta separación.

A la par, las muestras de ARN y ADN de las F<sup>+</sup> y F<sup>-</sup> de MDA-MB-231 a EpCam se procesaron y se utilizaron para los tres estudios subsecuentes. El análisis estructural no

---

mostró diferencias significativas (Figura 15), mientras que el de expresión que se realizó con dos plataformas diferentes arrojó resultados muy interesantes. El estudio elaborado con los microarreglos Human Gene ST 1.0 de Affymetrix® mostró 2,306 genes sobre-expresados y 1,606 sub-expresados (Figura 21), dentro de estos, se encuentran varios ya relacionados con la población CITCM (con los cuales podemos reforzar lo ya reportado por varios investigadores así como validar nuestra separación) como es el caso de mir-21, mir-221, CD44, CD24, ALDH, EpCam, IL6 y reguladores transcripcionales como NF- $\kappa$ B y Nanog (Tablas 11-13)<sup>4,7,9,20,23,28,31,65,82,83</sup>. Por otro lado, observamos transcritos con niveles de expresión desregulados en las CITCM como ATRX (que como ya mencionamos se encuentra involucrado con inestabilidad genética y crecimiento tumoral acelerado, entre otros) algunas caspasas, anexinas, metalopeptidasas, CDKs, inhibidor tisular de metaloproteinasas 1 (TIMP1), IGF1-2R y EGFR (receptor del factor de crecimiento epidérmico que también ha sido vinculado con el subtipo TN de cáncer de mama) y los reguladores transcripcionales FOXL2, IRF3, Nr1h, EZH2 (histona metiltransferasa y subunidad del complejo represivo polycomb 2 (PRC2) que juega papeles importantes en las células troncales embrionarias, CT adultas y CITCM y se encuentra incrementada en varios tipos de cáncer incluyendo el subtipo TN de cáncer de mama) y Hdac (histona deacetilasas que trabajan en conjunto con EZH2 en el mismo complejo protéico para mediar la represión génica transcripcional) (Tablas 11-13) que nos abren la posibilidad de explorar sus funciones específicas dentro de esta población en estudios posteriores<sup>4,31,58,75,83,109,129,130,132</sup>.

En cuanto a los resultados generados con la plataforma de RNAseq, obtuvimos algunos datos similares a los observados con los microarreglos, como por ejemplo la activación del factor de transcripción NANOG, la expresión desregulada de EpCam, EGFR, ALDH, IL6, BCL2, BRCA2 entre otros. Dentro de los que más llamaron nuestra atención fueron las predicciones de la interacción directa de Nanog (Figura 23) con KLF4, POU5F1, SOX2 (estos últimos cuentan con niveles de expresión elevados y son tres de los cuatro factores de transcripción que en conjunto con c-Myc son necesarios para la reprogramación

---

de células diferenciadas a células troncales pluripotentes inducidas o iPS (del inglés “Induced Pluripotent Stem Cells”) además de estar involucrados en la regulación de apoptosis, migración, proliferación, diferenciación celular, organogénesis, etc) así como también con el factor de transcripción GATA6 y BIRC5 (descrita previamente), conocida también como survivina, es miembro de la familia de las IAPs, se encuentra involucrada en metástasis así como también en el proceso de autorenovación de las CITCM y es de nuestro particular interés en el laboratorio<sup>47,73,93,95</sup>. Otros de los factores de transcripción activados fueron EPAS1, OTX2, MXD1 e inhibidos SOX5 y SOX6 (Figura 23), mientras que las moléculas que observamos con un vínculo ya reportado con cáncer de mama y las CITCM se encuentran varios de los integrantes de la familia de transportadores ABC, NOTCH (2-4), SOX2, IL8, SNAI (1-3), VIM, VEGF (A-C), miembros de la familia WNT, TGFB (1-3), PTEN, BMI1, CXCR4, PROCR, MUC1, BRCA1 y nuevamente y por último las que llamaron más nuestra atención por los niveles de expresión obtenidos como por el nexo con nuestro grupo de trabajo fueron las IAP BIRC2, 3 y 5 (Tabla 16)<sup>4,9,16,20,28,34,38,45,47,59,62,65,73,75,82-84,93,129,131</sup>.

El estudio del perfil de metilación de las CITCM y no CITCM provenientes de las líneas celulares MDA-MB-231 y MCF-7 de la primera etapa A, se realizó con el microarreglo Infinium HumanMethylation27 BeadChip® junto con otras muestras, sin embargo, los resultados obtenidos fueron utilizados para otros proyectos del laboratorio (denominados “Perfil de metilación y expresión de miRNAs asociados a células troncales tumorales del cáncer de mama” y “Perfil de metilación y expresión de miR-10b en células troncales de cáncer de mama”), esto, debido a los contratiempos que tuvimos tanto con las muestras de MCF-7 (descritos previamente) como con la validación de MDA-MB-231 la cual se explica a continuación.

Una vez finalizados los tres estudios con las F<sup>+</sup> y F<sup>-</sup> de MDA-MB-231 a EpCam quisimos realizar una última validación utilizando una herramienta que nos permitiera

---

comprobar las purificaciones de las fracciones pero a nivel de superficie celular, ya que a nivel transcripcional ya contábamos con la sobre-expresión de EpCam en las F<sup>+</sup> en comparación con las F<sup>-</sup> demostrada mediante RNAseq y microarreglos de expresión. Para este fin, se utilizó citometría de flujo con la cual no obtuvimos los resultados que esperábamos ya que no observamos diferencias entre la fracción enriquecida con CITCM (F<sup>+</sup>) y la no enriquecida (F<sup>-</sup>) por lo que un representante en México de la casa productora de MACS<sup>®</sup>, Miltenyi Biotec realizó junto con nosotros la separación, aunque también sin éxito (Anexo 2). A la fecha seguimos en espera de recibir alguna respuesta por parte de Alemania que es donde se encuentran las oficinas centrales, pero mientras tanto los resultados y análisis de los tres estudios realizados no los podemos utilizar ni mucho menos publicar hasta que se demuestre el enriquecimiento y diferencias entre las fracciones mediante citometría de flujo.

A partir de estos inconvenientes decidimos, en conjunto con mi tutor principal y comité tutorial, iniciar el mismo proyecto pero utilizando tres líneas celulares de cáncer de mama de origen basal (MDA-MB-231, MDA-MB4-68 y BT-20), citometría de flujo con FACS como herramienta de purificación y únicamente un triple marcaje con CD44/CD24/EpCam para delimitar las poblaciones CITCM y no CITCM.

---

### 14.1.2 PRIMERA ETAPA B

El inmunofenotipo y herramienta de purificación seleccionados en esta etapa para aislar nuestras poblaciones CITCM (CD44<sup>+</sup>CD24<sup>-</sup>EpCam<sup>+</sup>) y no CITCM (CD44<sup>+</sup>CD24<sup>-</sup>EpCam<sup>-</sup>, CD44<sup>+</sup>CD24<sup>+</sup>EpCam<sup>+</sup>, CD44<sup>+</sup>CD24<sup>+</sup>EpCam<sup>-</sup>, CD44<sup>-</sup>CD24<sup>-</sup>EpCam<sup>-</sup>, CD44<sup>-</sup>CD24<sup>-</sup>EpCam<sup>+</sup>, CD44<sup>-</sup>CD24<sup>+</sup>EpCam<sup>+</sup> y CD44<sup>-</sup>CD24<sup>+</sup>EpCam<sup>-</sup>) nos permitieron obtener las cantidades necesarias de ADN y ARN de las tres líneas celulares para iniciar nuestros tres análisis. El porcentaje de las CITCM varió entre cada línea celular, obteniendo del 50-70% en MDA-MB-231, 10-20% en MDA-MB-468 y del 15-30% en BT20 (Tabla 5). Por consiguiente, utilizamos las CITCM provenientes de la línea MDA-MB-231 (mayor porcentaje y facilidad de recolección) para confirmar el enriquecimiento de esta población mediante ensayos funcionales *in vitro* e *in vivo* (Figura 13) con los cuales se comprobaron los altos potenciales tumorigénicos, de autorenovación y diferenciación característicos de las células iniciadoras tumorales<sup>2,4,10,38,47,62</sup>.

La estructura genómica de las CITCM se evaluó por medio de los microarreglos Genome-Wide Human SNP 6.0 de Affymetrix®. El análisis comparativo grupal (CITCM de las tres líneas celulares vs no CITCM de las tres líneas celulares) no mostró resultados significativos, sin embargo si observamos en la comparación individual, es decir, CITCM de MDA-MB-231 vs no CITCM de MDA-MB-231, algunas variaciones en el número de copias (CNVs), específicamente dos en MDA-MB-231, siete en MDA-MB-468 y trece en BT20 (Tablas 7-9). Aunque estas no son de nuestro particular interés, sería interesante en un futuro poder comprender y elucidar la naturaleza de estas diferencias.

En esta etapa, el estudio de perfil de metilación de las CITCM se llevó a cabo con el microarreglo Infinium HumanMethylation450K BeadChip® el cual tiene mayor cobertura que el Infinium HumanMethylation27 BeadChip®<sup>128</sup>. Se obtuvieron 8,841 loci diferencialmente metilados en la población CITCM en comparación con la no CITCM, de estos, el 53% se

---

encontraron distribuidos en islas, 11% en n\_shore, 9% en s\_shore, 3% en s\_shelf, 3% en n\_shelf y el restante 22% no se obtuvieron coincidencias (Figura 17). En cuanto a los estados individuales de metilación de las poblaciones CITCM y no CITCM no varió significativamente, observando un 66% de sondas hipometiladas (valor  $\beta$  menor o igual a 0.3) y 34-33% (valor  $\beta$  mayor o igual 0.7) de sondas hipermetiladas en ambas poblaciones (Figura 18). Dentro de los cgs más diferencialmente metilados (Tabla 10) encontramos a GSTT1, la cual presenta el parámetro de estimación más alto (0.3495), en las CITCM se encuentra cercano al estado hipermetilado (0.6493) en comparación con las no CITCM donde se encuentra hipometilado (0.2997), los demás cgs se encuentran dentro del mismo estado de metilación pero con parámetros de estimación significativos como es el caso de EWSR1, PDCD4, IGF1R entre otros, adicionalmente, entender que efectos tienen estas diferencias en las CITCM y analizar si existe también alguna repercusión de estas a nivel de expresión nos podría ayudar a entender la relación de estos patrones con el comportamiento de estos grupos celulares<sup>58,94,132-134</sup>.

Por medio del análisis de expresión realizado con RNAseq de la plataforma de Illumina® Genome Analyzer Iix se obtuvieron 1,032 genes sobre-regulados y 8,711 sub-regulados en la población CITCM al compararla con la no CITCM, dentro de los transcritos con los niveles más altos de expresión (Tabla 17) que encontramos, observamos a UPK1B (uroplaquina 1B, involucrada en el desarrollo y progresión tumoral, igualmente en la proliferación, invasión y metástasis de células cancerosas; en el cáncer de mama, cáncer pulmonar de células no pequeñas, colorectal y hepatocelular han observado una expresión alterada así como también se ha propuesto que este gen pudiera regular el cáncer de vejiga mediante la vía de señalización de Wnt/ $\beta$ -catenina), ABCC12, WNT1 y por el contrario PAX2 (factor de transcripción vinculado con la regulación de proliferación y supervivencia celular en el cáncer de mama), GRAP, MUC2 (mucina epitelial cuyos niveles bajos de expresión han sido reportados previamente en tumores mamarios), FMO1, DPCR1 y BIRC7 (IAP identificada en melanoma y en varios tipos de cáncer, se ha planteado que en conjunto con

---

survivina juegan un rol importante en la tumorigenesis y en el pronóstico clínico) contaron con algunos de los niveles más bajos de expresión (Tabla 18) en esta población derivada de las tres líneas celulares. Otras moléculas diferencialmente expresadas y que son de nuestro interés por la relación que sostienen con las CITCM, CIT, cáncer de mama y nuestras investigaciones del laboratorio y que no necesariamente se encontraron dentro de las 50 moléculas más diferencialmente expresadas las podemos observar en la Tabla 19, estas no solo fueron seleccionadas para realizar la correlación con los resultados de metilación sino que también dieron y darán pie a nuevas investigaciones en nuestro laboratorio alrededor de este grupo celular, tal es el ejemplo del artículo con el cual me titulo (Anexo 5) en donde trabajamos con BIRC5 que como mencionamos y demostramos a lo largo de la tesis está claramente relacionada con las CITCM como ya reportado anteriormente por Ponti *et al* (quienes demostraron niveles altos de expresión de survivina en las CITCM derivadas de la línea celular luminal MCF-7) y Siddarth *et al* (que proponen un posible rol de survivina como marcador de CITCM quiescentes en una fase pre-metastásica). Por los roles que mantiene survivina en los procesos normales del ciclo celular como en las de las células tumorales (apoptosis, proliferación celular y resistencia) y su relación con las CITCM, es de gran importancia conocer a fondo sus diferentes aspectos y perfiles moleculares para que con ello podamos elucidar sus diversos mecanismos de acción e implicaciones en estos grupos celulares y poder generar información que nos ayude a comprender mejor la heterogeneidad tumoral como también ayude al desarrollo de nuevas herramientas terapéuticas<sup>73,114,116,117</sup>. Otro proyecto que surgió a partir de este mismo análisis es el que denominamos “Long Non-Coding RNAs in Breast Cancer Stem Cells”, en este estudio se están investigando únicamente los lncRNAs con niveles de expresión desregulados en las CITCM (Figura 29), actualmente se encuentra en proceso<sup>118,131,135–137</sup>.

**Figura 29.** Diagrama de Venn del análisis comparativo de los lncRNAs diferencialmente expresados en las CITCM.

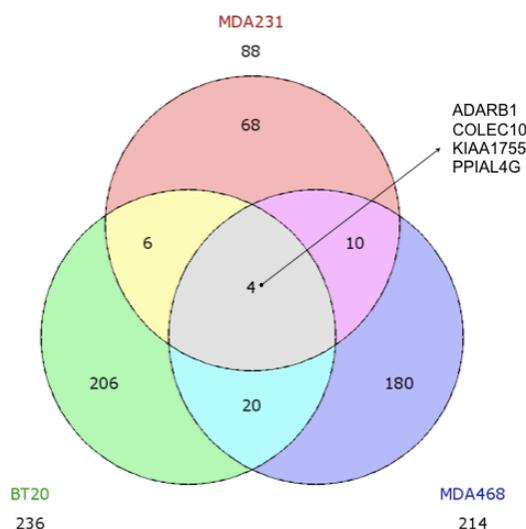


Diagrama de Venn generado con los resultados del análisis comparativo de los lncRNAs diferencialmente expresados en las CITCM de cada línea celular. En el conjunto rojo observamos los lncRNAs diferencialmente expresados en las CITCM de MDA-MB-231, en la verde en las CITCM de BT-20 y en la azul en las CITCM de MDA-MB-468. En total se obtuvieron 494 lncRNAs diferencialmente regulados, de los cuales únicamente cuatro compartían las CITCM de las tres líneas celulares.

Por último, el análisis de correlación entre los niveles de expresión y metilación nos permitió visualizar un probable vínculo entre estos dos mecanismos, específicamente en genes relacionados con troncalidad (Tabla 19), estos veintinueve posibles candidatos serán validados mediante RT-PCR (Retrotranscripción acoplada a la Reacción en Cadena de la Polimerasa) y MSP (PCR específica de metilación). Nuevamente y por ser de nuestro particular interés, hago un énfasis en los hallazgos observados en BIRC5 (valores de sobre-expresión e hipometilación) reforzando una vez más la asociación de survivina con las CITCM.

Estos resultados junto con los obtenidos posteriormente en los estudios de predicción de redes, cascadas, enriquecimiento de genes y estado de factores de

---

transcripción (Figuras 26-28) sustentan la idea de la existencia de un patrón de metilación activado seguido por cambios específicos y sucesivos a nivel transcripcional y reflejados a nivel de expresión génica que podrían ser los responsables del fenotipo troncal de las CITCM<sup>58,94,132-134</sup>.

---

## 15. CONCLUSIONES

- La citometría de flujo con FACS junto con el triple marcaje seleccionado (CD44<sup>+</sup>CD24<sup>-</sup>EpCam<sup>+</sup>) nos permitió lograr un enriquecimiento adecuado de las CITCM para su caracterización y estudio en comparación con las MACS<sup>®</sup> con las que experimentamos varios obstáculos y seguimos en espera de una solución por parte de la casa productora para realizar la validación de las separaciones previas.
- Se comprobó que las CITCM tienen mayor potencial tumorigénico, mayor capacidad de formación de mamosferas y colonias en condiciones no adherentes y en agar suave respectivamente, que las no CITCM.
- Se lograron los objetivos específicos planteados en un inicio. Algunos de los posibles marcadores propuestos ya se están y/o fueron utilizados en varios estudios como es el caso de BIRC5 (artículo de titulación), para el proyecto de lncRNAs : “Long Non-Coding RNAs in Breast Cancer Stem Cells”, para el de miRNAs: “Perfil de metilación y expresión de miRNAs asociados a células troncales tumorales del cáncer de mama” así como para el proyecto de titulación de licenciatura del alumno Juan Canales Delgado denominado: “Perfil de metilación y expresión de miR-10b en células troncales de cáncer de mama” el cual junto con el Dr. Meléndez Zajgla me dieron oportunidad de dirigir (Anexo 3). El resto de los probables marcadores serán validados posteriormente.
- Los resultados obtenidos demuestran que las diferencias entre las CITCM y no CITCM no radican en su perfil estructural, no obstante, encontramos cambios significativos en el perfil de expresión y metilación, apoyando así la hipótesis inicial en la que se propuso que las CITCM poseen un perfil genómico diferente

---

al de las células que componen el resto del tumor (no CITCM). En un futuro, sería de gran interés analizar e integrar los demás perfiles de este grupo celular (epigenéticos, miRNAs, lncRNAs, entre otros) para así poder comprender su origen, comportamiento y naturaleza y con ello lograr el desarrollo de nuevos blancos terapéuticos y marcadores diagnósticos y pronósticos dirigidos a esta enfermedad, el cáncer de mama .

---

## 16. REFERENCIAS

1. Gao, J.-X. Cancer stem cells: the lessons from pre-cancerous stem cells. *J. Cell. Mol. Med.* **12**, 67–96 (2008).
2. Ponnusamy, M. P. & Batra, S. K. Ovarian cancer: emerging concept on cancer stem cells. *J. Ovarian Res.* **1**, 4 (2008).
3. Mimeault, M. & Batra, S. K. Interplay of distinct growth factors during epithelial mesenchymal transition of cancer progenitor cells and molecular targeting as novel cancer therapies. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **18**, 1605–1619 (2007).
4. Morrison, B. J., Schmidt, C. W., Lakhani, S. R., Reynolds, B. A. & Lopez, J. A. Breast cancer stem cells: implications for therapy of breast cancer. *Breast Cancer Res. BCR* **10**, 210 (2008).
5. Schwarz-Cruz-y-Celis, A. & Meléndez-Zajgla, J. Cancer stem cells. *Rev. Investig. Clin. Organo Hosp. Enfermedades Nutr.* **63**, 179–186 (2011).
6. Ashkenazi, R., Gentry, S. N. & Jackson, T. L. Pathways to tumorigenesis—modeling mutation acquisition in stem cells and their progeny. *Neoplasia N. Y. N* **10**, 1170–1182 (2008).
7. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3983–3988 (2003).
8. Ajani, J. A., Song, S., Hochster, H. S. & Steinberg, I. B. Cancer stem cells: the promise and the potential. *Semin. Oncol.* **42 Suppl 1**, S3-17 (2015).
9. Palomeras, S. *et al.* Targeting Breast Cancer Stem Cells to Overcome Treatment Resistance. *Molecules* **23**, 2193 (2018).
10. Owens, T. W. & Naylor, M. J. Breast cancer stem cells. *Front. Physiol.* **4**, (2013).
11. Singh, B. N., Fu, J., Srivastava, R. K. & Shankar, S. Hedgehog Signaling Antagonist GDC-0449 (Vismodegib) Inhibits Pancreatic Cancer Stem Cell Characteristics: Molecular Mechanisms. *PLOS ONE* **6**, e27306 (2011).
12. Amini, S., Fathi, F., Mobalegi, J., Sofimajidpour, H. & Ghadimi, T. The expressions of stem cell markers: Oct4, Nanog, Sox2, nucleostemin, Bmi, Zfx, Tcl1, Tbx3, Dppa4, and Esrrb in bladder, colon, and prostate cancer, and certain cancer cell lines. *Anat. Cell Biol.* **47**, 1–11 (2014).
13. Jing, F. *et al.* Colon cancer stem cell markers CD44 and CD133 in patients with colorectal cancer and synchronous hepatic metastases. *Int. J. Oncol.* **46**, 1582–1588 (2015).
14. Filipova, A. *et al.* Breast cancer and cancer stem cells: a mini-review. *Tumori* **100**, 363–369 (2014).
15. Honeth, G. *et al.* The CD44+/CD24- phenotype is enriched in basal-like breast tumors. *Breast Cancer Res. BCR* **10**, R53 (2008).
16. Bahena-Ocampo, I. *et al.* miR-10b expression in breast cancer stem cells supports self-renewal through negative PTEN regulation and sustained AKT activation. *EMBO Rep.* **17**, 648–658 (2016).
17. Cariati, M. *et al.* Alpha-6 integrin is necessary for the tumourigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line. *Int. J. Cancer* **122**, 298–304 (2008).
18. Visvader, J. E. & Lindeman, G. J. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat. Rev. Cancer* **8**, 755–768 (2008).
19. Marotta, L. L. C. & Polyak, K. Cancer stem cells: a model in the making. *Curr. Opin. Genet. Dev.* **19**, 44–50 (2009).
20. Schwarz-Cruz Y Celis, A., Espinosa, M., Maldonado, V. & Melendez-Zajgla, J. Advances in the knowledge of breast cancer stem cells. A review. *Histol. Histopathol.* **31**, 601–612 (2016).
21. Foreman, K. E., Rizzo, P., Osipo, C. & Miele, L. The Cancer Stem Cell Hypothesis. in *Stem Cells and Cancer* 3–14 (Humana Press, 2009). doi:10.1007/978-1-60327-933-8\_1.
22. Hollier, B. G., Evans, K. & Mani, S. A. The epithelial-to-mesenchymal transition and cancer stem cells: a coalition against cancer therapies. *J. Mammary Gland Biol. Neoplasia* **14**, 29–43 (2009).
23. Sin, W. C. & Lim, C. L. Breast cancer stem cells—from origins to targeted therapy. *Stem Cell Investig.* **4**, (2017).

- 
24. Ricardo, S. *et al.* Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. *J. Clin. Pathol.* **64**, 937–946 (2011).
  25. Dalerba, P. *et al.* Phenotypic characterization of human colorectal cancer stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 10158–10163 (2007).
  26. Roudi, R., Korourian, A., Sharifabrizi, A. & Madjd, Z. Differential Expression of Cancer Stem Cell Markers ALDH1 and CD133 in Various Lung Cancer Subtypes. *Cancer Invest.* **33**, 294–302 (2015).
  27. Guzel, E. *et al.* Differential expression of stem cell markers and ABCG2 in recurrent prostate cancer. *The Prostate* **74**, 1498–1505 (2014).
  28. Karakasliotis, I. & Mavromara, P. Hepatocellular carcinoma: from hepatocyte to liver cancer stem cell. *Front. Physiol.* **6**, (2015).
  29. Allegra, E. & Trapasso, S. Cancer stem cells in head and neck cancer. *OncoTargets Ther.* **5**, 375–383 (2012).
  30. Paholak, H. J. *et al.* Elimination of epithelial-like and mesenchymal-like breast cancer stem cells to inhibit metastasis following nanoparticle-mediated photothermal therapy. *Biomaterials* **104**, 145–157 (2016).
  31. Nalla, L. V., Kalia, K. & Khairmar, A. Self-renewal signaling pathways in breast cancer stem cells. *Int. J. Biochem. Cell Biol.* **107**, 140–153 (2019).
  32. Ginestier, C. *et al.* ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* **1**, 555–567 (2007).
  33. Liu, T. J. *et al.* CD133+ cells with cancer stem cell characteristics associates with vasculogenic mimicry in triple-negative breast cancer. *Oncogene* **32**, 544–553 (2013).
  34. Wright, M. H. *et al.* Brca1 breast tumors contain distinct CD44+/CD24- and CD133+ cells with cancer stem cell characteristics. *Breast Cancer Res. BCR* **10**, R10 (2008).
  35. Bozorgi, A., Khazaei, M. & Khazaei, M. R. New Findings on Breast Cancer Stem Cells: A Review. *J. Breast Cancer* **18**, 303–312 (2015).
  36. Hubbard, S. A. *et al.* Evidence for cancer stem cells in human endometrial carcinoma. *Cancer Res.* **69**, 8241–8248 (2009).
  37. Singh, S. K. *et al.* Identification of human brain tumour initiating cells. *Nature* **432**, 396–401 (2004).
  38. Charafe-Jauffret, E., Ginestier, C. & Birnbaum, D. Breast cancer stem cells: tools and models to rely on. *BMC Cancer* **9**, 202 (2009).
  39. Vassilopoulos, A. *et al.* Identification and characterization of cancer initiating cells from BRCA1 related mammary tumors using markers for normal mammary stem cells. *Int. J. Biol. Sci.* **4**, 133–142 (2008).
  40. Bray, F., McCarron, P. & Parkin, D. M. The changing global patterns of female breast cancer incidence and mortality. *Breast Cancer Res. BCR* **6**, 229–239 (2004).
  41. Knaul Felicia Marie, Arreola-Ornelas Héctor, Lozano Rafael & Gómez Dantés Héctor. México: Numeralia de cáncer de mama. *Vigil. Epidemiológica Sem.* 26 2013 **30**, 5.
  42. Siegel, R., Ma, J., Zou, Z. & Jemal, A. Cancer statistics, 2014. *CA. Cancer J. Clin.* **64**, 9–29 (2014).
  43. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2016. *CA. Cancer J. Clin.* **66**, 7–30 (2016).
  44. Ferlay, J. *et al.* Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136**, E359–E386 (2015).
  45. Klonisch, T. *et al.* Cancer stem cell markers in common cancers - therapeutic implications. *Trends Mol. Med.* **14**, 450–460 (2008).
  46. Kao, J. *et al.* Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PLoS One* **4**, e6146 (2009).
  47. Butti, R., Gunasekaran, V. P., Kumar, T. V. S., Banerjee, P. & Kundu, G. C. Breast cancer stem cells: Biology and therapeutic implications. *Int. J. Biochem. Cell Biol.* **107**, 38–52 (2019).
  48. Li, S.-Y. *et al.* Combination therapy with epigenetic-targeted and chemotherapeutic drugs delivered by nanoparticles to enhance the chemotherapy response and overcome resistance by breast cancer stem cells. *J. Control. Release Off. J. Control. Release Soc.* **205**, 7–14 (2015).
  49. Schwarzenbacher, D., Balic, M. & Pichler, M. The role of microRNAs in breast cancer stem cells. *Int. J. Mol. Sci.* **14**, 14712–14723 (2013).
  50. INEGI. ESTADÍSTICAS A PROPOSITO DEL... DÍA MUNDIAL CONTRA EL CÁNCER (4 DE FEBRERO). [http://www.inegi.org.mx/saladeprensa/aproposito/2017/cancer2017\\_Nal.pdf](http://www.inegi.org.mx/saladeprensa/aproposito/2017/cancer2017_Nal.pdf).
-

- 
51. Romero Utrilla, A., Osuna-Ramos, J., Candanedo González, F., Guadalupe Zepeda Ramirez, M. & JE, P. Cáncer de Mama: Entidad Patológica de Biología Heterogénea. *Arch. Salud Sinaloa* **8**, 109 (2014).
  52. OMS. *Perfiles oncológicos de los países, 2014. México*. [http://www.who.int/cancer/country-profiles/mex\\_es.pdf?ua=1](http://www.who.int/cancer/country-profiles/mex_es.pdf?ua=1).
  53. Hidalgo-Miranda, A. & Jiménez-Sánchez, G. [Genomic basis for breast cancer: advances in personalized medicine]. *Salud Publica Mex.* **51 Suppl 2**, s197-207 (2009).
  54. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
  55. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
  56. Luen, S., Virassamy, B., Savas, P., Salgado, R. & Loi, S. The genomic landscape of breast cancer and its interaction with host immunity. *Breast Edinb. Scotl.* **29**, 241–250 (2016).
  57. Wood, L. D. *et al.* The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108–1113 (2007).
  58. Baylin, S. B. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum. Mol. Genet.* **10**, 687–692 (2001).
  59. Shipitsin, M. *et al.* Molecular definition of breast tumor heterogeneity. *Cancer Cell* **11**, 259–273 (2007).
  60. Nahas, G. R., Patel, S. A., Bliss, S. A. & Rameshwar, P. Can breast cancer stem cells evade the immune system? *Curr. Med. Chem.* **19**, 6036–6049 (2012).
  61. Friedmann-Morvinski, D. & Verma, I. M. Dedifferentiation and reprogramming: origins of cancer stem cells. *EMBO Rep.* **15**, 244–253 (2014).
  62. Schmitt, F., Ricardo, S., Vieira, A. F., Dionísio, M. R. & Paredes, J. Cancer stem cell markers in breast neoplasias: their relevance and distribution in distinct molecular subtypes. *Virchows Arch. Int. J. Pathol.* **460**, 545–553 (2012).
  63. Campbell, L. L. & Polyak, K. Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell Cycle Georget. Tex* **6**, 2332–2338 (2007).
  64. Soltanian, S. & Matin, M. M. Cancer stem cells and cancer therapy. *Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med.* **32**, 425–440 (2011).
  65. Charafe-Jauffret, E. *et al.* Breast cancer cell lines contain functional cancer stem cells with metastatic capacity and a distinct molecular signature. *Cancer Res.* **69**, 1302–1313 (2009).
  66. Prud'homme, G. J. Cancer stem cells and novel targets for antitumor strategies. *Curr. Pharm. Des.* **18**, 2838–2849 (2012).
  67. Bosch Barrera, J., López-Picazo González, J. M., García-Foncillas López, J. & Prósper Cardoso, F. [Stem cells and cancer: elucidating the origin of the cancer stem cell]. *Rev. Med. Univ. Navarra* **51**, 14–17 (2007).
  68. Cheshier, S. H. *et al.* A neurosurgeon's guide to stem cells, cancer stem cells, and brain tumor stem cells. *Neurosurgery* **65**, 237–249; discussion 249-250; quiz N6 (2009).
  69. Phillips, T. M., McBride, W. H. & Pajonk, F. The Response of CD24<sup>-/low</sup> /CD44<sup>+</sup> Breast Cancer-Initiating Cells to Radiation. *JNCI J. Natl. Cancer Inst.* **98**, 1777–1785 (2006).
  70. Lagadec, C. *et al.* Survival and self-renewing capacity of breast cancer initiating cells during fractionated radiation treatment. *Breast Cancer Res. BCR* **12**, R13 (2010).
  71. Park, S. Y. *et al.* Heterogeneity for Stem Cell-Related Markers According to Tumor Subtype and Histologic Stage in Breast Cancer. *Clin. Cancer Res.* **16**, 876–887 (2010).
  72. Tordonato, C., Di Fiore, P. P. & Nicassio, F. The role of non-coding RNAs in the regulation of stem cells and progenitors in the normal mammary gland and in breast tumors. *Front. Genet.* **6**, (2015).
  73. Ponti, D. *et al.* Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer Res.* **65**, 5506–5511 (2005).
  74. Chumsri, S. *et al.* Cancer stem cells and individualized therapy. *Cancer Genomics Proteomics* **4**, 165–174 (2007).
  75. Bao, B., Ahmad, A., Azmi, A. S., Ali, S. & Sarkar, F. H. Cancer Stem Cells (CSCs) and Mechanisms of Their Regulation: Implications for Cancer Therapy. *Curr. Protoc. Pharmacol. Editor. Board SJ Enna Ed.--Chief AI* **0 14**, Unit-14.25 (2013).
  76. Salaverry, O. La etimología del cáncer y su curioso curso histórico. *Rev. Peru. Med. Exp. Salud Publica* **30**, 137–141 (2013).
-

- 
77. Polyak, K. & Hahn, W. C. Roots and stems: stem cells in cancer. *Nat. Med.* **12**, 296–300 (2006).
  78. Pierce, G. B. & Speers, W. C. Tumors as caricatures of the process of tissue renewal: prospects for therapy by directing differentiation. *Cancer Res.* **48**, 1996–2004 (1988).
  79. Lapidot, T. *et al.* A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645–648 (1994).
  80. Wortham, M. & Yan, H. Apples to origins: Identifying brain tumor stem cell genes by comparing transcriptomes of normal and cancer stem cells. *Cancer Discov.* **2**, 492–494 (2012).
  81. Yang, Z. F. *et al.* Significance of CD90+ cancer stem cells in human liver cancer. *Cancer Cell* **13**, 153–166 (2008).
  82. Nguyen, N. P. *et al.* Molecular biology of breast cancer stem cells: potential clinical applications. *Cancer Treat. Rev.* **36**, 485–491 (2010).
  83. Zhao, Z., Li, S., Song, E. & Liu, S. The roles of ncRNAs and histone-modifiers in regulating breast cancer stem cells. *Protein Cell* **7**, 89–99 (2016).
  84. Hwang-Verslues, W. W. *et al.* Multiple Lineages of Human Breast Cancer Stem/Progenitor Cells Identified by Profiling with Stem Cell Markers. *PLOS ONE* **4**, e8377 (2009).
  85. Tanei, T. *et al.* Association of breast cancer stem cells identified by aldehyde dehydrogenase 1 expression with resistance to sequential Paclitaxel and epirubicin-based chemotherapy for breast cancers. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **15**, 4234–4241 (2009).
  86. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 (2001).
  87. Tsang, J. Y. S. *et al.* Cancer stem cell markers are associated with adverse biomarker profiles and molecular subtypes of breast cancer. *Breast Cancer Res. Treat.* **136**, 407–417 (2012).
  88. Morrison, S. J. & Kimble, J. Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature* **441**, 1068–1074 (2006).
  89. Wang, Q.-Z., Lu, Y.-H., Jiang, N., Diao, Y. & Xu, R.-A. The asymmetric division and tumorigenesis of stem cells. *Chin. J. Cancer* **29**, 248–253 (2010).
  90. Dean, M., Fojo, T. & Bates, S. Tumour stem cells and drug resistance. *Nat. Rev. Cancer* **5**, 275–284 (2005).
  91. Villadsen, R. *et al.* Evidence for a stem cell hierarchy in the adult human breast. *J. Cell Biol.* **177**, 87–101 (2007).
  92. Dontu, G. *et al.* In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.* **17**, 1253–1270 (2003).
  93. Kim, S.-H. & Singh, S. V. Mammary cancer chemoprevention by withaferin A is accompanied by in vivo suppression of self-renewal of cancer stem cells. *Cancer Prev. Res. Phila. Pa* **7**, 738–747 (2014).
  94. Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492 (2012).
  95. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
  96. Choi, Y. *et al.* Epithelial-mesenchymal transition increases during the progression of in situ to invasive basal-like breast cancer. *Hum. Pathol.* **44**, 2581–2589 (2013).
  97. Taube, J. H. *et al.* Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15449–15454 (2010).
  98. Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* **12**, R68 (2010).
  99. Dieci, M. V., Orvieto, E., Dominici, M., Conte, P. & Guarneri, V. Rare breast cancer subtypes: histological, molecular, and clinical peculiarities. *The Oncologist* **19**, 805–813 (2014).
  100. DeSantis, C. E. *et al.* Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA. Cancer J. Clin.* **66**, 31–42 (2016).
  101. Perou, C. M. & Børresen-Dale, A.-L. Systems Biology and Genomics of Breast Cancer. *Cold Spring Harb. Perspect. Biol.* **3**, (2011).
  102. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
-

- 
103. Sørlie, T. *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl. Acad. Sci.* **100**, 8418–8423 (2003).
  104. Dai, X., Cheng, H., Bai, Z. & Li, J. Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. *J. Cancer* **8**, 3131–3141 (2017).
  105. Stephens, P. J. *et al.* Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* **462**, 1005–1010 (2009).
  106. Keller, P. J. *et al.* Mapping the cellular and molecular heterogeneity of normal and malignant breast tissues and cultured cell lines. *Breast Cancer Res. BCR* **12**, R87 (2010).
  107. Herschkowitz, J. I. *et al.* Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol.* **8**, R76 (2007).
  108. Gathani, T., Bull, D., Green, J., Reeves, G. & Beral, V. Breast cancer histological classification: agreement between the Office for National Statistics and the National Health Service Breast Screening Programme. *Breast Cancer Res.* **7**, R1090 (2005).
  109. Huang, J. P. & Ling, K. EZH2 and histone deacetylase inhibitors induce apoptosis in triple negative breast cancer cells by differentially increasing H3 Lys27 acetylation in the BIM gene promoter and enhancers. *Oncol. Lett.* **14**, 5735–5742 (2017).
  110. Tamimi, R. M. *et al.* Traditional breast cancer risk factors in relation to molecular subtypes of breast cancer. *Breast Cancer Res. Treat.* **131**, 159–167 (2012).
  111. Imigo G., F., Mansilla S., E., Delama G., I., Poblete S., M. T. & Fonfach Z., C. Clasificación molecular del cáncer de mama. *Cuad. Cir.* **25**, 67–74 (2011).
  112. Lim, E. *et al.* Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat. Med.* **15**, 907–913 (2009).
  113. David G Hicks & Susan C. Lester. HISTOLOGY OF NORMAL BREAST. in *Diagnostic Pathology: Breast, 2nd Edition* 4–7 (Elsevier, 2016).
  114. Siddharth, S., Das, S., Nayak, A. & Kundu, C. N. SURVIVIN as a marker for quiescent-breast cancer stem cells-An intermediate, adherent, pre-requisite phase of breast cancer metastasis. *Clin. Exp. Metastasis* **33**, 661–675 (2016).
  115. Carmena, M., Wheelock, M., Funabiki, H. & Earnshaw, W. C. The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. *Nat. Rev. Mol. Cell Biol.* **13**, 789–803 (2012).
  116. Gaytan-Cervantes, J. *et al.* Protein Sam68 regulates the alternative splicing of survivin DEx3. *J. Biol. Chem.* **292**, 13745–13757 (2017).
  117. Espinosa, M. *et al.* Survivin isoform Delta Ex3 regulates tumor spheroid formation. *Cancer Lett.* **318**, 61–67 (2012).
  118. Altieri, B. *et al.* Livin/BIRC7 expression as malignancy marker in adrenocortical tumors. *Oncotarget* **8**, 9323–9338 (2016).
  119. Kaklamani, V. A genetic signature can predict prognosis and response to therapy in breast cancer: Oncotype DX. *Expert Rev. Mol. Diagn.* **6**, 803–809 (2006).
  120. Bernhardt, S. M. *et al.* Hormonal Modulation of Breast Cancer Gene Expression: Implications for Intrinsic Subtyping in Premenopausal Women. *Front. Oncol.* **6**, (2016).
  121. Kitagawa, M. & Lee, S. H. The chromosomal passenger complex (CPC) as a key orchestrator of orderly mitotic exit and cytokinesis. *Front. Cell Dev. Biol.* **3**, 14 (2015).
  122. Espinosa, E. *et al.* Gene profiling in breast cancer: time to move forward. *Cancer Treat. Rev.* **37**, 416–421 (2011).
  123. Kim, S. & Hahn, W. Cancer genomics: Integrating form and function. *Carcinogenesis* **28**, 1387–92 (2007).
  124. Balmain, A., Gray, J. & Ponder, B. The genetics and genomics of cancer. *Nat. Genet.* **33** Suppl, 238–244 (2003).
  125. Reich, M. *et al.* GenePattern 2.0. *Nat. Genet.* **38**, 500–501 (2006).
  126. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
  127. Mootha, V. K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).

- 
128. Mancuso, F. M., Montfort, M., Carreras, A., Alibés, A. & Roma, G. HumMeth27QCReport: an R package for quality control and primary analysis of Illumina Infinium methylation data. *BMC Res. Notes* **4**, 546 (2011).
  129. Roy, D., Guida, P., Zhou, G., Echiburu-Chau, C. & Calaf, G. m. Gene expression profiling of breast cells induced by X-rays and heavy ions. *Int. J. Mol. Med.* **21**, 627–636 (2008).
  130. Koschmann, C. *et al.* ATRX Loss Promotes Tumor Growth and Impairs Non-Homologous End Joining DNA Repair in Glioma. *Sci. Transl. Med.* **8**, 328ra28 (2016).
  131. Nik-Zainal, S. *et al.* Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47–54 (2016).
  132. Siegfried, Z. & Simon, I. DNA methylation and gene expression. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2**, 362–371 (2010).
  133. Razin, A. & Cedar, H. DNA methylation and gene expression. *Microbiol. Rev.* **55**, 451–458 (1991).
  134. Moore, L. D., Le, T. & Fan, G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* **38**, 23–38 (2013).
  135. Role of UPK1B in bladder cancer. *J Ma* 10.
  136. Silberstein, G. B., Dressler, G. R. & Van Horn, K. Expression of the PAX2 oncogene in human breast cancer and its role in progesterone-dependent mammary growth. *Oncogene* **21**, 1009–1016 (2002).
  137. Walsh, M. D., McGuckin, M. A., Devine, P. L., Hohn, B. G. & Wright, R. G. Expression of MUC2 epithelial mucin in breast carcinoma. *J. Clin. Pathol.* **46**, 922–925 (1993).

---

## 17. ANEXOS

### ANEXO 1. ABREVIATURAS

#### Números

**3D:** Tridimensionales

#### A

**AACR:** Asociación americana para la investigación del cáncer

**ABCB1:** Glicoproteína P o Proteína 1 de resistencia a multidroga, MDR1

**ABCC12:** Transportador dependiente de ATP subfamilia C miembro 12, Transportador ABC subfamilia C miembro 12, Proteína 9 de resistencia a multidroga, MRP9

**ABCG2 (BCRP1):** Proteína resistente al cáncer de mama o Transportador de membrana dependiente de ATP subfamilia G miembro 2

**AC MET:** Adenocarcinoma Metastásico

**AC-133:** Proliferin 1, PROM1, CD133

**AC133:** Anticuerpo monoclonal anti-CD133

**ADN:** Ácido desoxirribonucleico

**AH:** Ácido hialurónico

**ALDH:** Aldehído deshidrogenasa

**ALDH1:** Aldehído deshidrogenasa 1

**ANOVA:** Análisis de varianza

**APC:** Aloficocianina

**ARN:** Ácido ribonucleico

**ASB:** Albúmina de suero de bovino

**ATCC:** American Type Culture Collection

**ATM:** Gen Ataxia telangiectasia mutado o Quinasa ATM serina/treonina

**ATRA:** Por sus siglas en inglés "All-Trans Retinoic Acid"

**ATRX:** Remodelador de cromatina ATRX o Gen del síndrome de alfa-talasemia con retraso mental ligado al cromosoma X

#### B

**BA:** Basal A

**BB:** Basal B

**BCL-2:** Linfoma de células B-2 o regulador de apoptosis BCL2

**BIRC5:** Repetición IAP Baculoviral que contiene proteína 5, inhibidor de apoptosis survivina o Survivina variante 3 alfa

**BIRC7:** Repetición IAP Baculoviral que contiene proteína 7, Livin o Kiap

**Bmi-1:** Proteína 4 de dedo anular del grupo polycomb (PcG), protooncogen BMI1 del dedo anular del grupo polycomb o PCGF4.

**BRCA1:** Gen de cáncer de mama 1, BRCA1 asociado a la reparación de ADN

**BrDU:** Bromodesoxiuridina

#### C

---

**c-kit:** Proteína-Tirosina cinasa Kit, proto-oncogen c-kit, CD117

**c-Myc:** Proto-oncogen MYC / factor de transcripción BHLH, proto-oncogen c-Myc, factor de transcripción P64

**Cd.Mx:** Ciudad de México

**CALLA:** Gen Metaloendopeptidasa membranal, CD10, neprilisina o encefalinasa

**CD133:** Grupo de diferenciación CD133, Prominina 1, PROML1

**CD24:** Grupo de diferenciación 24 o antígeno estable al calor CD24 (HSA)

**CD326:** Grupo de diferenciación 326, Molécula de adhesión celular epitelial, EpCam, ESA o KSA

**CD34:** Grupo de diferenciación 34 o Antígeno celular progenitor hematopoyético CD34

**CD38:** Grupo de diferenciación 38 o ADP ribosa cíclica hidrolasa

**CD44:** Grupo de diferenciación 44, glicoproteína fagocítica 1, HCAM, Pgp1, antígeno Hermes, receptor de hialuronato

**CDI:** Carcinoma Ductal Invasivo

**CDKs:** Cinasas dependientes de ciclinas

**CIT:** Células iniciadoras tumorales, CTC

**CITCM:** Células iniciadoras tumorales de cáncer de mama

**CK:** Citoqueratina

**CK18:** Citoqueratina 18

**CK19:** Citoqueratina 19

**CK7:** Citoqueratina 7

**CK8:** Citoqueratina 8

**CNV:** Variaciones en el número de copias de ADN

**CPC:** Complejo pasajero de los cromosomas, por sus siglas en inglés “Chromosomal Passenger Complex”

**CSCs:** Del inglés “Cancer Stem Cells”, CIT, CTC

**CT:** Células troncales normales

**CTC:** Células troncales cancerosas, CIT

**CXCL12:** Quimiocina con motivo CXC ligando 12, SDF1, PBSF, IRH

**CXCR4:** Quimiocina con motivo CXC receptor 4, receptor de SDF1, antígeno CD184, Fusina

## D

**DEAB:** Del inglés “Diethylaminobenzaldehyde”

**DMEM:** Medio de Eagle modificado por Dulbecco

**DPCR1:** MUCL3, Parecida a Mucina tipo 3, PBLT

## E

**EDTA:** Ácido etilendiaminotetraacético

**EGFR:** Receptor del factor de crecimiento epidérmico, ERBB, HER1, ERBB1

**EP:** Efusión Pleural

**EPAS1:** Proteína del dominio endotelial PAS1, HIF2-alfa, HLF

**EpCam:** Molécula de adhesión celular epitelial, ESA, CD326 o KSA

---

**ERO:** Especies reactivas de oxígeno, ROS  
**EWSR1:** Proteína 1 de unión a RNA EWS, oncogen EWS (sarcoma de Ewing)  
**EZH2:** Potenciador del Zeste homólogo 2, Potenciador del Zeste 2 subunidad del complejo 2 represivo polycomb, ENX1, KMT6, metiltransferasa H3K27

## F

**FACS:** Clasificación celular activada por fluorescencia, por sus siglas en inglés “Fluorescence Activating Cell Sorting”  
**Fgfr1:** Receptor 1 del factor de crecimiento de los fibroblastos, FLT2, CEK, FLG  
**FITC:** Isotiocianato de fluoresceína  
**FMO1:** Monooxigenasa 1 dependiente de flavina, Monooxigenasa 1 dependiente de flavina hepática fetal  
**FOXC2:** Caja de cabeza de tenedor C2, Factor de transcripción FKH-14, MFH1, LD  
**FOXL2:** Caja de cabeza de tenedor L2, Factor de transcripción cabeza de tenedor FOXL2, BPES1, PINTO, POF3  
**FZD:** Frizzled

## G

**GABA:** Ácido  $\gamma$ -amino butírico  
**GATA6:** Proteína de unión GATA6, Factor de transcripción GATA6

**Gli1:** Dedo de zinc familia Gli1, Oncogen asociado a glioma homólogo 1 (proteína de dedo de zinc), Oncogen GLI1  
**GRAP:** Proteína adaptadora relacionada GRB2  
**GS:** GeneSpring herramienta bioinformática  
**GSTT1:** Glutation S-transferasa teta1

## H

**Hdac:** Histona deacetilasas  
**HER2:** Receptor 2 de factor de crecimiento epidérmico humano, por sus siglas en inglés “Human Epidermal Growth Factor Receptor 2”, ERBB2, CD340, Proto-oncogen Neu  
**Hh:** Hedgehog  
**HSA:** Antígeno estable al calor, CD24, grupo de diferenciación 24

## I

**IAPs:** Proteínas inhibidoras de la apoptosis, por sus siglas en inglés “Inhibitors of Apoptosis Proteins”  
**IF:** Intensidad de fluorescencia  
**IGF1-2R:** Receptor del factor de crecimiento 1 similar a la insulina (IGF1R), Receptor del factor de crecimiento 2 similar a la insulina (IGF2R)  
**IHQ:** Inmunohistoquímica  
**INCan:** Instituto Nacional de Cancerología  
**INCENP:** Proteína centromérica interna, por sus siglas en inglés “Inner Centromere Protein”

---

**INMEGEN:** Instituto Nacional de Medicina Genómica

**iPS:** Células troncales pluripotentes inducidas, del inglés "Induced Pluripotent Stem Cells"

**IRF3:** Factor regulador de interferon 3

## J

**JAK/STAT:** Janus Cinasas transductores de la señal y activadores de la transcripción

## K

**Ki67:** Marcador de proliferación Ki-67, MKI67, KIA

**KLF4:** Factor 4 parecido a Krüppel, GSKF, EZF

## L

**LA:** Luminal A

**LHIM:** Laboratorio de histología y microscopía del INMEGEN

**Lin:** Linaje

**InRNAs:** ARNs largos no codificantes

**LRP5/6:** Proteína relacionada al receptor de LDL-5/6, Proteína relacionada al receptor de la lipoproteína de baja densidad 5/6

**LT:** Línea Total

## M

**MACS<sup>®</sup>:** Separación celular activada por magnetos por sus siglas en inglés "Magnetic Activation Cell Sorting"

**miRNAs:** micro ARNs, por sus siglas en inglés "microRNAs"

**MSP:** PCR específica de metilación

**MUC1:** Mucina 1 asociada a superficie celular, Mucina asociada a carcinoma, Mucina polimórfica epitelial, Episialina

**MUC2:** Mucina 2 moco oligomérico/ formadora de gel, Mucina 2 Intestinal/Traqueal, MLP

**MXD1:** Proteína 1 dimerizadora MAX, MAD, BHLHC58

## N

**Nanog:** Proteína homeótica Nanog, Caja homeótica Nanog, Factor de transcripción de la caja homeótica Nanog

**ncRNAs:** ARNs no codificase

**neg:** F<sup>-</sup> de MDA-MB-231 a EpCam (en PCA)

**NGS:** Secuenciación de nueva generación

**NOD/SCID:** Diabéticos No Obesos / Inmunodeficiencia combinada severa

**Nr1h:** Receptores nucleares subfamilia 1 grupo H

**nt:** Nucleótidos

**NUMB:** Proteína adaptadora endocítica NUMB

## O

---

**Oct4:** Factor de transcripción-4 de unión a octámero, POU5F1, Oct-3

**OMS:** Organización Mundial de la Salud

**OTX2:** Caja homeótica ortodentículo 2, homólogo de ortodentículo 2, CPHD6, MCOPS5

## P

**PAX2:** Gen caja pareada 2, Gen pax 2, PAPRS, FSGS7,

**PBS:** Buffer fosfato salino, solución amortiguadora de fosfatos

**PCA:** Análisis de Componentes Principales

**PDCD4:** Muerte celular programada 4, antígeno nuclear H731

**PE:** Ficoeritrina

**PGS:** Partek<sup>®</sup> Genomics Suite<sup>®</sup> herramienta bioinformática

**piRNAs:** ARNs asociados a Piwi

**pos:** F<sup>+</sup> de MDA-MB-231 a EpCam (en PCA)

**PRC2:** Complejo represivo polycomb 2

**PROCR:** Receptor para la proteína C, EPCR, CD201

**Ptch1:** Dominio patched que contiene proteína 1, AUTSX4

**Pten:** Homólogo de fosfatasa y tensina, MMAC, TEP1

## R

**RE:** Receptor de estrógenos

**RH:** Receptores hormonales

**RIN:** número de integridad del ARN

**RMA:** Del inglés "Robust Multichip Average"

**RNAseq:** Secuenciación de ARN

**RP:** Receptor de progesterona

**Rt:** Réplica técnica (en PCA)

**RT-PCR:** Retrotranscripción (RT) acoplada a la Reacción en Cadena de la Polimerasa (PCR)

## S

**SCID:** Inmunodeficiencia combinada severa por sus siglas en inglés "Severe Combined Immunodeficiency"

**SFB:** Suero Fetal Bovino

**siRNAs:** ARNs pequeños de interferencia

**Slug (1-2) :** Represor transcripcional (1-3) de la familia snail, SNAIL (1-3), SNAI (1-3)

**Smad5:** Miembro 5 de la familia SMAD, JV5-1, MADH5

**SMO:** Smoothened receptor de la clase Frizzled, Miembro de la familia Frizzled 11, CRJS, Gx

**SNAI (1-3):** Represor transcripcional (1-3) de la familia snail, SNAIL (1-3), Slug (1-2)

**Snail (1-3):** Represor transcripcional (1-3) de la familia snail, Slug (1-2), SNAI (1-3)

**SNPs:** Polimorfismos de un sólo nucleótido

**SNR:** Relación señal/ruido

---

**SOX2:** Región determinante del sexo Y - Caja 2, ANOP3, Factor de transcripción SOX-2  
**SOX5:** Región determinante del sexo Y - Caja 5, LAMSHF, Factor de transcripción SOX-5  
**SOX6:** Región determinante del sexo Y - Caja 6, HSSOX6, Factor de transcripción SOX-6  
**SP:** Del inglés "Side Population"  
**SWI/SNF:** Complejo remodelador de cromatina switch/sacarosa no fermentable

## T

**TEM:** Transición epitelio-mesénquima  
**TGF- $\beta$ :** Factor de crecimiento transformante beta  
**Thy-1:** Antígeno de diferenciación de timocitos 1, CD90  
**TICs:** Del inglés "Tumor Initiating Cells", CTC, CIT  
**TIMP1:** Inhibidor tisular de metaloproteinasas 1, Inhibidor de colagenasa, CLGI, EPA, EPO  
**TN:** Triple negativo  
**TP:** Tumor Primario  
**TWIST1:** Factor de transcripción BHLH 1 de la familia Twist, CRS1, ACS3, BPES2  
**TWIST2:** Factor de transcripción BHLH 2 de la familia Twist, DERMO1, AMS, FFDD3

## U

**UCIF:** Unidad de Citometría de Flujo  
**UMI:** Unidad de Microarreglos

**UPEAL-B:** Unidad de Producción y Experimentación de Animales de Laboratorio-Bioterio de la Universidad Autónoma de México  
**UPK1B:** Uroplaquina 1B, tetraspanina 20, UPIB, TSPAN20  
**USEC:** Unidad de Secuenciación

## V

**VIM:** Vimentina  
**Vs:** versus

## W

**Wnt:** Del inglés "Wingless" e Int-1  
**WNT1:** Miembro 1 de la familia Wnt, Protooncogen Wnt-1, INT1, BMND16

## Y

**YAP1:** Proteína 1 asociada a YES, YAP65, COB1

## Z

**ZEB1:** Caja homeótica 1 de unión a la caja E con dedos de zinc, Factor de transcripción 8, AREB6, BZP

## Abreviaturas de medidas

**$\mu$ g:** Microgramos

---

**μm:** Micrómetros

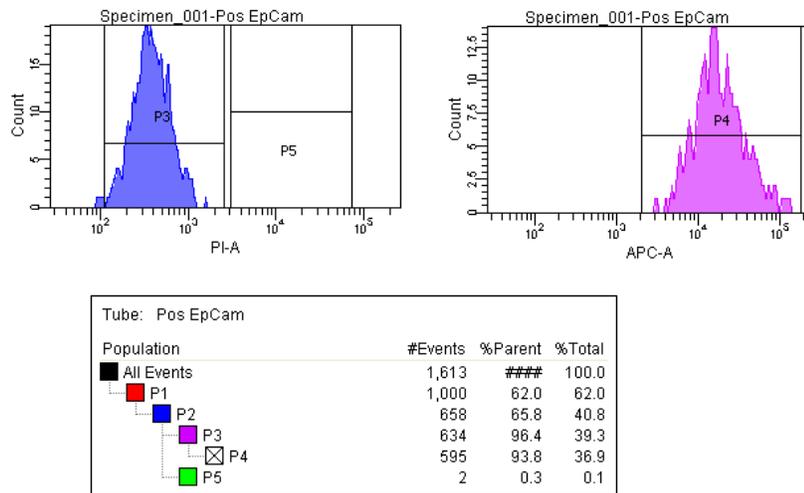
**mM:** Milimolar

**ng:** Nanogramos

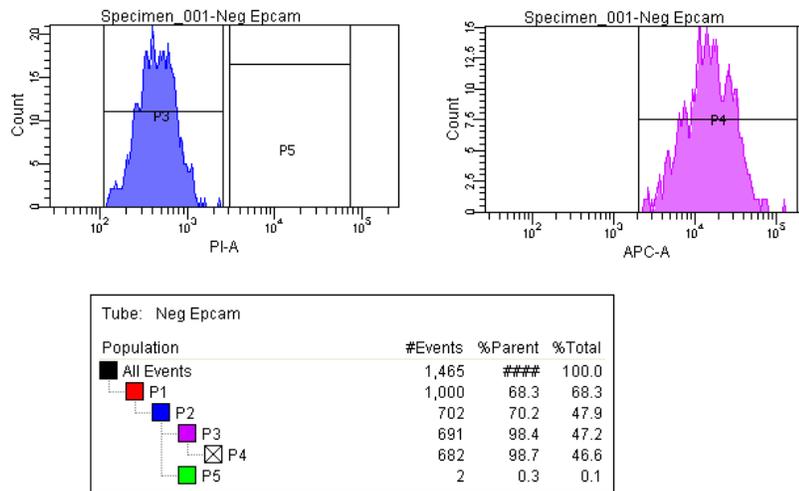
**nm:** Nanómetros

## ANEXO 2. VALIDACIÓN DE EPCAM CON UNIPARTS

**A**



**B**



Las células se marcaron con el anticuerpo de EpCam acoplado a APC. En la imagen superior **(A)** observamos la lectura de la población CITCM positiva a EpCam (P4) y en la inferior **(B)** la negativa a EpCam (no CITCM), sin embargo observamos positividad (P4) prácticamente total en esta última tomando como fallida la validación.

---

**ANEXO 3. PROYECTO DE TITULACIÓN DE LICENCIATURA DEL ALUMNO JUAN  
CANALES DELGADO**



---

---

**Universidad de Guanajuato**

Campus Guanajuato

División de Ciencias Naturales y Exactas

**“Perfil de metilación y expresión de miR-  
10b en células troncales de cáncer de  
mama”**

Tesis

Que para obtener el título de  
Químico Farmacéutico Biólogo

Presenta:

Juan Canales Delgado

Guanajuato, Guanajuato.

Febrero 2014

---

El presente trabajo se realizó en su totalidad en el Instituto Nacional de Medicina Genómica en México, en el Laboratorio de Genómica Funcional del Cáncer, bajo la completa dirección del Dr. Jorge Meléndez Zajgla y la Candidata a Doctora Angela Schwarz Cruz y Celis.

---

## ANEXO 4. PUBLICACIONES Y ACTIVIDADES

### ARTÍCULOS:

- Schwarz-Cruz-y-Celis, A. & Meléndez-Zajgla, J. Cancer stem cells. *Rev. Investig. Clin. Organo Hosp. Enfermedades Nutr.* **63**, 179–186 (2011).
- Lohr, J. G. *et al.* Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3879–3884 (2012).
- Crompton, B. D. *et al.* The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov.* **4**, 1326–1341 (2014).
- Bahena-Ocampo, I. *et al.* miR-10b expression in breast cancer stem cells supports self-renewal through negative PTEN regulation and sustained AKT activation. *EMBO Rep.* **17**, 648–658 (2016).
- Lizarraga, F. *et al.* Tissue inhibitor of metalloproteinases-4 (TIMP-4) regulates stemness in cervical cancer cells. *Mol. Carcinog.* **55**, 1952–1961 (2016).
- Schwarz-Cruz Y Celis, A., Espinosa, M., Maldonado, V. & Melendez-Zajgla, J. Advances in the knowledge of breast cancer stem cells. A review. *Histol. Histopathol.* **31**, 601–612 (2016).

### ACTIVIDADES:

**Agosto-noviembre 2011:** Estancia de tres meses en Broad Institute por parte de Iniciativa Slim en Medicina Genómica.

---

## **ANEXO 5. ARTÍCULO**

Article

# Basal-Type Breast Cancer Stem Cells Over-Express Chromosomal Passenger Complex Proteins

Angela Schwarz-Cruz y Celis <sup>1</sup>, Gisela Ceballos-Cancino <sup>1</sup>, Karla Vazquez-Santillan <sup>1</sup> ,  
Magali Espinosa <sup>1</sup> , Cecilia Zampedri <sup>1</sup>, Ivan Bahena <sup>1,2</sup>, Victor Ruiz <sup>3</sup>, Vilma Maldonado <sup>1</sup> and  
Jorge Melendez-Zajgla <sup>1,\*</sup> 

<sup>1</sup> Subdirección de Investigación Básica, Instituto Nacional de Medicina Genómica, Periférico Sur 4809, Col. Arenal Tepepan, Tlalpan, Mexico City 14610, Mexico; angelaschwarz.c@gmail.com (A.S.-C.y.C.); gceballos@inmegen.gob.mx (G.C.-C.); kivazquez@inmegen.gob.mx (K.V.-S.); mespinosa@inmegen.gob.mx (M.E.); mceciliazampedri@gmail.com (C.Z.); ivan.bahena@gmail.com (I.B.); vmaldonado@inmegen.gob.mx (V.M.)

<sup>2</sup> Present address Health Sciences department, Universidad Autónoma Metropolitana, San Rafael Atlixco No. 186, Col. Vicentina, Iztapalapa 09340, Mexico

<sup>3</sup> Instituto Nacional de Enfermedades Respiratorias. Calzada de Tlalpan 4502, Col. Sección XVI, Tlalpan, Mexico City 14080, Mexico; vicoruz@yahoo.com.mx

\* Correspondence: jmelendez@inmegen.gob.mx; Tel.: +52-5553-501920

Received: 13 February 2020; Accepted: 29 February 2020; Published: 13 March 2020



**Abstract:** (1) Aim: In the present paper we analyzed the transcriptome of CSCs (Cancer Stem Cells), in order to find defining molecular processes of breast cancer. (2) Methods: We performed RNA-Seq from CSCs isolated from the basal cell line MDA-MB-468. Enriched processes and networks were studied using the IPA (Ingenuity Pathway Analysis) tool. Validation was performed with qRT-PCR and the analysis of relevant genes was evaluated by overexpression, flow cytometry and in vivo zebrafish studies. Finally, the clinical relevance of these results was assessed using reported cohorts. (3) Results: We found that CSCs presented marked differences from the non-CSCs, including enrichment in transduction cascades related to stemness, cellular growth, proliferation and apoptosis. Interestingly, CSCs overexpressed a module of co-regulated Chromosomal Passenger Proteins including BIRC5 (survivin), INCENP and AURKB. Overexpression of BIRC5 increased the number of CSCs, as assessed by in vitro and in vivo zebrafish xenotransplant analyses. Analysis of previously published cohorts showed that this co-regulated module was not only overexpressed in basal breast tumors but also associated with relapse-free and overall survival in these patients. (4) Conclusions: These results underline the importance of Cancer Stem Cells in breast cancer progression and point toward the possible use of chromosomal passenger proteins as prognostic factors.

**Keywords:** breast cancer; networks; stem cells

## 1. Introduction

Breast cancer is the most commonly diagnosed cancer among women from both developing and developed countries and it remains one of the leading causes of death in women worldwide [1]. Among the described molecular subtypes, patients with basal breast tumors present the poorest prognosis, with shorter metastasis-free and overall survival [2]. These tumors have a gene expression pattern very similar to mammary stem cells [3], and reciprocally, normal mammary stem cells overexpress a set of genes common to basal breast tumors [4]. Over the past few years Breast Cancer Stem Cells (BCSCs) have been proposed to be one of the cell populations responsible for this type of cancer's heterogeneity, metastasis and tumor relapse. BCSCs have been isolated from tumor samples and breast

cancer cell lines using a combination of surface markers such as CD44, CD24, EpCAM, CD133 and ALDH, among others [5]. Even with a large number of published studies that analyze the phenotypic properties of BCSCs, genome-wide gene expression studies are needed to gain new insights for a better understanding of the complexity of breast cancer biology. In the present report, we analyzed the transcriptome of BCSCs derived from the basal breast cancer cell line MDA-MB-468 in which we found enrichment in gene networks associated with several signaling cascades related to stemness. Among the regulated networks we found a novel module consisting of genes codifying for chromosomal passenger proteins, including BIRC5, INCENP and AURKB, which proved to be associated with relapse-free and overall survival in triple-negative breast tumors.

## 2. Materials and Methods

### 2.1. Cell Line

MDA-MB-468 basal breast cancer cells were purchased from the American Type Culture Collection (ATCC, MD). These cells were grown in Dulbecco's Modified Minimal Medium (DMEM) supplemented with 5% Fetal Bovine Serum in a 37 °C and 5% CO<sub>2</sub>-humidified atmosphere.

### 2.2. Cancer Stem Cells Isolation

Cells were separated using a FACSAria BD cell sorter with the following cell surface markers: EpCam, CD44 and CD24, which have been reported previously [5]. The antibodies were bought from Miltenyi Biotec (Bergisch Gladbach, Germany) and BD Biosciences (Carlsbad, CA, USA). Briefly, cells were detached with Accutase (Life Technologies, Carlsbad, CA, USA) and washed with PBS 1%. A total of  $15 \times 10^6$  cells were resuspended in a stained buffer (PBS with FBS 1%) and then were incubated with antibodies for 30 min at 4 °C. FITC-IgG1 and PE-IgG1 (Miltenyi Biotec; Bergisch Gladbach, Germany) were used as isotypes controls. CSCs (CD44+/CD24<sup>low</sup>) and non-stem cells (CD44+/CD24<sup>+</sup>) were sorted on a FACSAria (fluorescence-activated cell sorter, Becton Dickinson, Franklin Lakes, NJ, USA). For validation, cells were seeded in MammoCult (StemCell Technologies; Vancouver, Canada) for 24 h before analyses. RNA was isolated using Trizol (Life Technologies, Mexico City, Mexico) and RNA was reverse transcribed with a Maxima First strand cDNA synthesis kit (Thermo Fisher Scientific, Rockford, IL, USA). The expression level of stem cell markers (ALDH1A1, OCT4, ALDH1A3 and KLF4) and survivin were assayed by real-time PCR using the SYBR or TaqMan Master Kit (Applied Biosystems, Mexico City, Mexico) and PCRs were performed on the QuantStudio 7 (Applied Biosystems).

### 2.3. RNA Analyses

Gene expression was analyzed by qRT-PCR using the double Delta-Ct method to compare the levels of transcripts. The sequence of primers and TaqMan assays are depicted in Table S1. For the transcriptome assays, two pooled (two biological replicates each to a total of four) breast CSCs and non-CSCs RNA samples isolated from the MDA-MB-468 cell line were subjected to RNA-Seq analysis. Approximately  $30\text{--}40 \times 10^6$  75 base-pair reads per sample were obtained. After quality control and trimming, reads were aligned against the GRCh38 human reference, using Gencode's version 84 GTF and the STAR aligner [6]. Normalization and transcript quantification was performed using RNA-Seq by Expectation Maximization (RSEM) [7] and differential expression calculated with EBSeq [8]. Transcripts with an FDR < 0.05 were considered Differentially Expressed (DE). Relative expression was depicted as absolute fold change. Raw data were deposited at the NCBI SRA archive (accession number PRJNA509033). The networks, functional and upstream analyses were generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City) and Gene Set Enrichment Analysis (GSEA) [9]. Correlation expression analysis was performed using COXPRESdb v6.0 [10], which aggregates 73,083 human microarray samples (Platform A-AFFY-44) from different tissues and experimental settings. Probe sets used were: 202095\_s\_at for BIRC5, 200853\_at for H2AFZ, 209464\_at for AURKB and 219769\_at for INCENP. R coefficient correlation was plotted for all platforms.

Since the analysis covered several platforms, COXPRESdb used supportability as a measure of the significance level, using the COXSIM values, as described in [10]. For this, the maxCOXSIM value derived from the Mutual Rank Score was compared with the null distribution generated under the same number of genes in the list of every platform employed. Expression levels of the chromosomal passenger proteins across breast cancer samples were obtained from the Gene expression-based Outcome for Breast Cancer Online (GOBO) tool [11], which includes 1881 breast samples from 11 public datasets. Expression levels for the analyzed genes in breast cancer subtypes were obtained from GEO (NCBI) and TCGA databases as normalized Fragments Per Kilobase Of Exon Per Million Fragments Mapped (FKPM; GSE58135 [12]) or Transcripts Per Kilobase Million (TPM; TCGA, database accessed 05/2016, [13]). The relapse-free and overall survival for the module was calculated using *kmplot* [14]. For this, the normalized value of the specific Affymetrix probe was used, the percentile between the lower and upper quartiles was computed and the best performing threshold was used as the final cutoff, as described in [15]. qRT-PCR validation: Stem and non-stem cells were isolated as described and seeded in MammoCult media (StemCell Technologies, BC, Canada) for 24 h.

#### 2.4. Xenotransplant Assays

Wild type AB Zebrafish (*Danio rerio*; donated from Dr. Hilda Lomely from IBT-UNAM) were bred and maintained in standard conditions. Two days' post-fertilization (dpf) embryos were dechorionated and anesthetized with tricaine (MS-222; Sigma, Mexico City, Mexico). Cells were isolated using Fluorescence-Activated Cell Sorting (FACS) or used directly. No tumors were observed in fibroblast-only injected animals. Embryos were microinjected in the middle of the embryonic yolk sac region at a final concentration of 37 cells for CSCs assays. For survivin-overexpressing assays 75, 125, 250 and 500 cells were injected. For each group of cells, 30 embryos were used, which were then monitored for five Days Post Injection (DPI). Only informative (e.g., surviving) embryos were assessed. In vivo limiting dilution assays were analyzed using the Extreme Limiting Dilution Assays (ELDA) software [16], which employs a generalized linear model used to compare active cell frequencies (stem cells) in cell populations. The project was revised and approved by INMEGEN's scientific and ethical committees (09/2015/I).

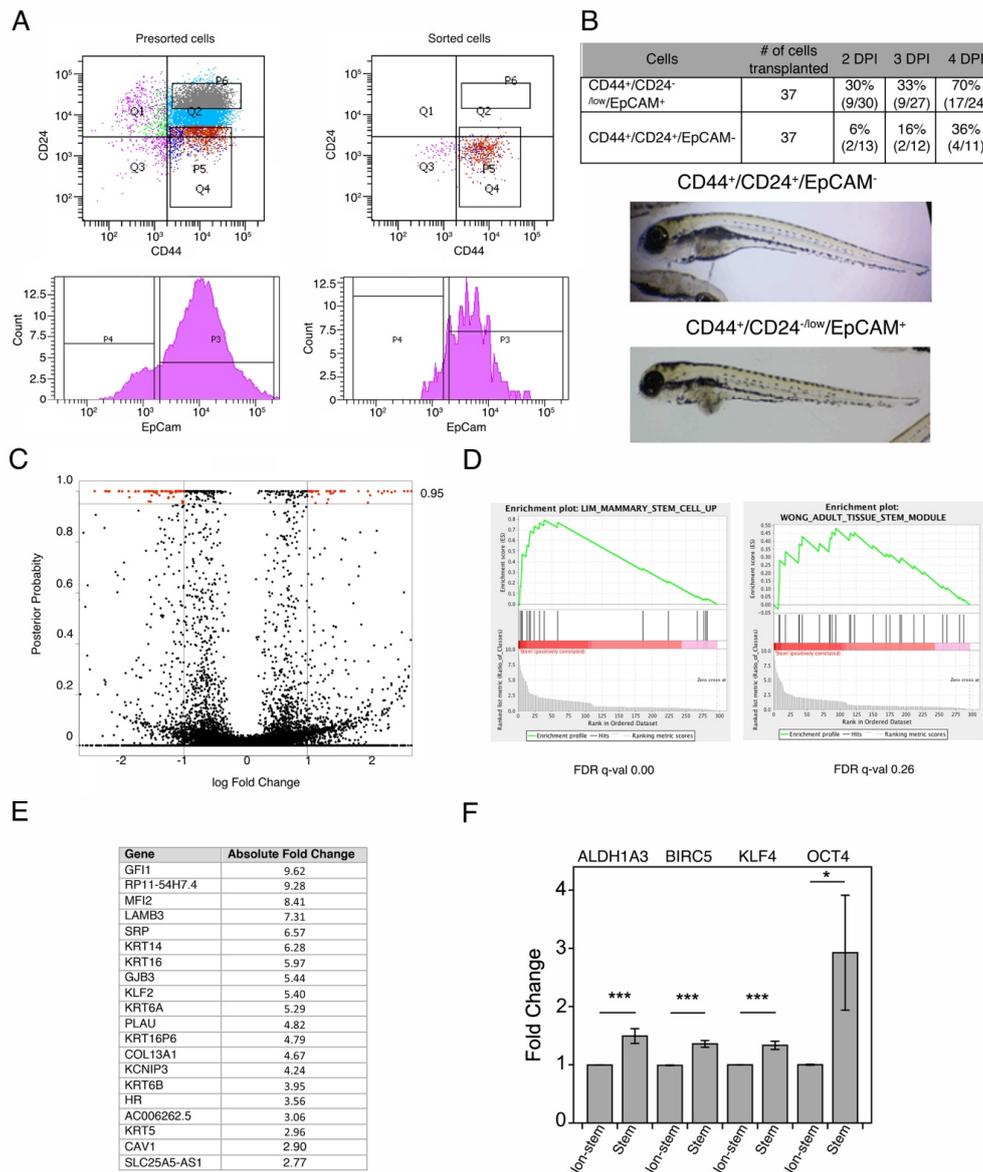
#### 2.5. BIRC5 Overexpression and Inhibition

To overexpress BIRC5 in MDA-MB-468 cells, the ORF of BIRC5 was cloned into the pQCXIP vector (Clontech, CA, USA). The construction was transiently transfected with Lipofectamine 2000 (Invitrogen, Mexico City, Mexico) and RNA was extracted 48 h after transfection. As a control an empty pQCXIP vector was used. To inhibit BIRC5 expression, the U6 promoter and a downstream 20 bp BIRC5-specific sequence (GACGACCCCATAGAGGAACA) were subcloned into the pQCXIP vector (Clontech). The construction was stably transfected in MDA-MB-468 cells. BIRC5 were decreased to approximately 75% (25% reduction) in these cells.

### 3. Results

#### 3.1. Isolation of Breast Cancer Stem Cells

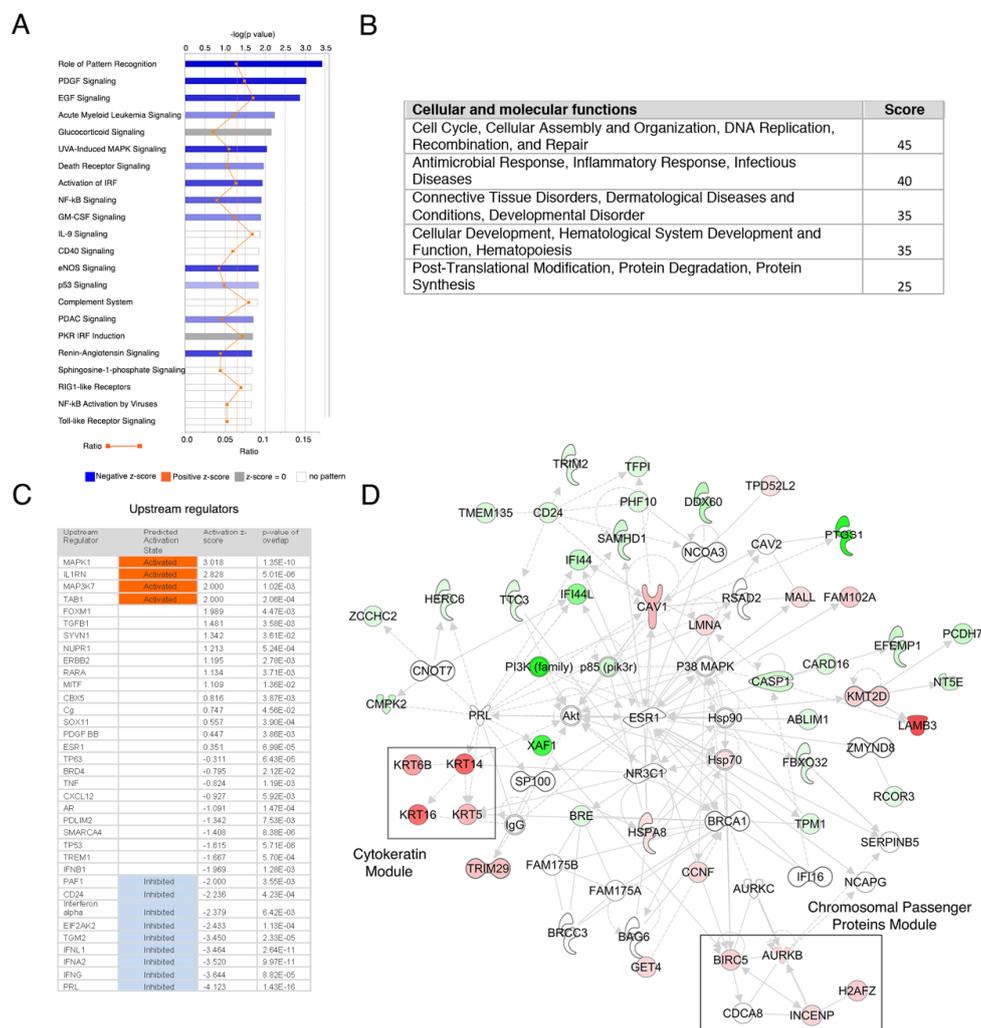
CSCs and non-CSCs were isolated using FACS from the triple negative breast cancer (TNBC) cell line MDA-MB-468 (Figure 1A). We used a combination of CD44, CD24 and EpCam markers to isolate CSCs since previous works have shown that these markers identify a subpopulation with increased tumor-initiating capabilities in triple-negative breast cancer cell lines [5] (Figure 1A). To verify that these surface markers indeed identified CSCs, we performed xenotransplant analyses in zebrafish. As expected, we found that CD44<sup>+</sup>/CD24<sup>-/low</sup>/EpCam<sup>+</sup> MDA-MB-468 cells had higher tumor-initiating ability in this model (Figure 1B).



**Figure 1.** Isolation of Cancer Stem Cells. (A) Left panels: Flow cytometry analysis of MDA-MB-468 cells previous to sorting. The upper panel shows a dot plot with the fluorescence intensity of CD44 versus CD24 and the selected fields for sorting (CD44 high and CD24 low). The lower panel shows the third used marker (EpCam). Right panels: These plots display the purified population obtained after sorting. (B) Tumorigenicity assays using a zebrafish xenotransplant model. Upper panel: Tumor developed by MDA-MB-468 non-Cancer Stem Cells (CSCs; CD44<sup>+</sup>CD24<sup>+</sup>EpCAM<sup>-</sup>) and CSCs (CD44<sup>+</sup>CD24<sup>-low</sup>EpCAM<sup>+</sup>) at 2, 3 and 4 DPI (Days Post Injection). For each experiment, 30 embryos were used and only informative (e.g., surviving) embryos were assessed. #: The number of cells transplanted to the embryos. Lower panel: Representative image taken at 2 DPI, showing abdominal tumors (arrows) in zebrafish embryos transplanted with MDA-MB-468 non-CSCs and MDA-MB-468 CSCs. (C) Volcano plot showing DE significant genes in red (posterior probability of being differentially expressed > 95%). The line shows the cutoff for significance. (D) GSEA analysis of the DE genes found. Left panel: Enrichment against a mammary stem cells signature (derived from [17]). Right panel: Enrichment against a tissue stem cell signature (derived from [18]). (E) Top 20 DE genes with a False Discovery Rate < 0.05. (F) Absolute fold change of representative genes in CD44<sup>+</sup>CD24<sup>-low</sup>EpCAM<sup>+</sup> (Stem) cells versus CD44<sup>+</sup>CD24<sup>+</sup>EpCAM<sup>-</sup> (Non-stem) cells analyzed by qRT-PCR. Data derived from three independent FACS isolation experiments; \* *p* < 0.05; \*\*\* *p* < 0.001.

### 3.2. Breast Cancer Stem Cells Transcriptome Analysis

We next performed RNA-Seq analysis of sorted cells. We found 297 DE genes with a False Discovery Rate (FDR) < 0.05 (Figure 1C and Supplementary Data). As expected, these Differentially Expressed (DE) genes presented an overrepresentation of mammary-stem and, to a lesser extent tissue-stem-associated gene (FDR q-value 0.26), as assessed by GSEA (Figure 1D). Among them we found commonly stem-associated genes such as SOX9, Oct1 and Myc. In addition, a group of cytokeratins including KRT5, KRT6 and KRT14 were among the top 20 DE genes in the stem cell population (Figures 1E and 2D). These proteins are part of the breast cancer’s molecular intrinsic gene expression classification that defines triple-negative breast tumors and are associated with worse prognosis [19]. As expected, qRT-PCR analyses validated the stem-related genes found to be upregulated in the RNAseq analyses (Figure 1F).



**Figure 2.** Network analysis of DE genes in CSCs from MDA-MB-468 cells. (A) Significant signaling cascades in the studied set. Bars length depicts the  $-\log(p)$  value of each cascade, the color of the bar represents the z-score and the orange line points the ratio of genes in the pathway (e.g., the proportion of the studied DE genes present in each signaling cascade). (B) Enriched cellular and molecular functions table. The main functions and scores are shown. (C) Upstream regulators. The table shows the upstream regulator gene, the predicted activation (based on an absolute score > 2) and the p value of the gene overlap. (D) Main network in CSCs, showing two important modules in squares (Cytokeratin module and Chromosomal Passenger Proteins module). Red represents increased expression of the molecule in CSCs and green decreased expression.

We then explored the implications of the DE genes found by performing network and pathway analysis using the IPA suite. As shown in Figure 2A, several canonical signaling pathways were predicted to be modulated in CSCs, including cascades previously reported as important regulators of breast cancer stem cell phenotype, such as the Death Receptor ( $p$  value:  $1.08^{-2}$ ) [20], p53 ( $p$  value:  $1.46^{-2}$ ) [21] and NF-kappaB ( $p$  value:  $1.24^{-2}$ ) [22] pathways, among others. This coincides with enrichment in genes associated with cellular functions indispensable for the stem cell phenotype such as cell death, growth and cellular development (Figure 2B,  $p$ -value ranges:  $7.41^{-3}$ – $1.73E^{-8}$ ,  $6.12^{-3}$ – $3.50^{-6}$  and  $6.46^{-3}$ – $3.50^{-6}$ , respectively). Additional support for the enrichment in stem cell genes was found by performing a network analysis of DE mRNAs isoforms (after excluding DE genes), which also showed activation of cascades associated with stem cells (Figure S1). We then explored the upstream regulator genes that could explain the expression changes and cellular functions in CSCs. Figure 2C shows that several upstream regulators were inhibited, including interferon-related pathways (IFNL1, IFNA2 and IFNG), Prolactin, PAF1, EIF2AK2 and, as expected, CD24. Due to the sheer number of upstream regulators found, we performed an additional comprehensive “master regulator” analysis, which employs hierarchical clustering to point potential regulators that could be explaining the upstream signaling in these cells. The top hits of this analysis were ERBB2, Prolactin, HGF, BECN1, FAS and EPHA2 ( $p$  value of overlap:  $8.17^{-17}$ ,  $1.43^{-16}$ ,  $1.69^{-15}$ ,  $2.16^{-15}$  and  $5.62^{-14}$ , respectively). All these cascades have been described as important regulators of the stem cell phenotype in normal or malignant epithelial cells [23–25].

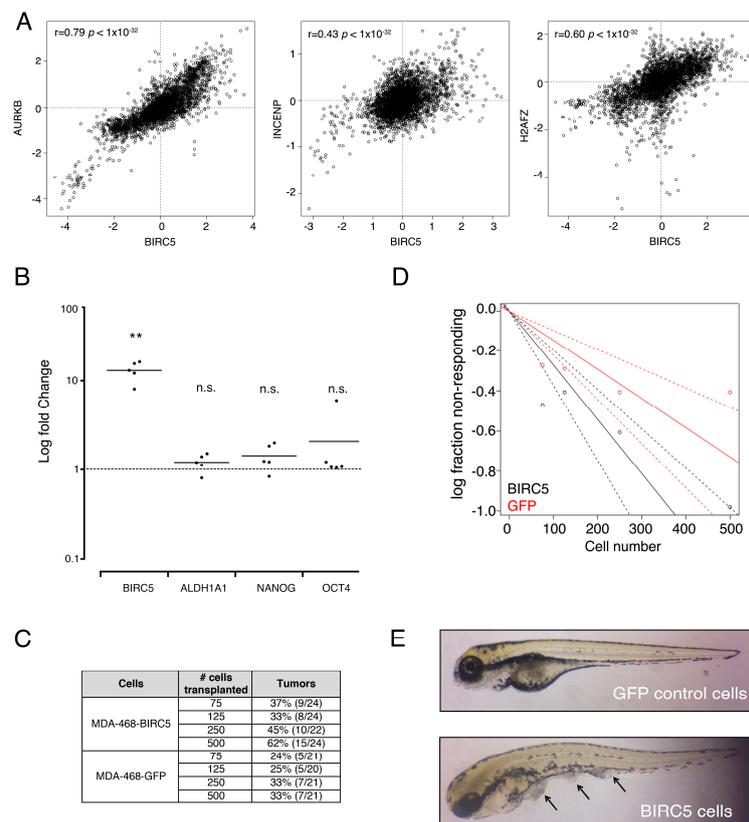
### 3.3. Genes from a Chromosomal Passenger Proteins Module (CPPM) are Overexpressed in CSCs

A more detailed analysis showed that one of the top general networks included a strongly induced expression module comprised by several basal cytokeratins (Figure 2D). As mentioned, these proteins are part of the breast cancer’s molecular intrinsic gene expression classification that defines triple-negative breast tumors [19], so we focused in it. Interestingly, in the same network we found deregulation of a Chromosomal Passenger Proteins Module (CPPM), which included BIRC5 (Survivin), INCENP, AURKB and H2AFZ. We validated the overexpression of BIRC5 and the stem cell marker ALDH1A3 by quantitative qRT-PCR (Figure 1F). BIRC5, INCENP and AURKB are three of the four proteins that constitute the chromosomal passenger complex (CPC), which is involved in maintaining a proper chromosome segregation and cytokinesis during mitosis (Figure 2D) [26]. To further support the existence of a CPPM co-regulated module in CSCs, we used a combined analysis of human gene expression datasets [10] and found that these proteins were positively co-expressed (BIRC5 Mutual Rank of co-expression of 14.4, 17 and 51.3 for AURKB, INCENP and H2AFZ, respectively; Figure 3A). Next, to explore the role of this module in the stem cell phenotype we overexpressed BIRC5 in MDA-MB-468 cells. Figure 3B shows that cells that overexpressed BIRC5 had a tendency to present higher expression levels of the stem markers ALDH1A1, NANOG and OCT4, although not statistically significant. As expected by these results, cells expressing BIRC5 showed an increased stem cell number, as assessed by xenotransplant dilution assays (Figure 3C–E). As expected, depletion of BIRC5 in MDA-MB-468 cells by an shRNA induced mitotic aberrations (20.01 aberrant mitosis per  $10^6$  cells).

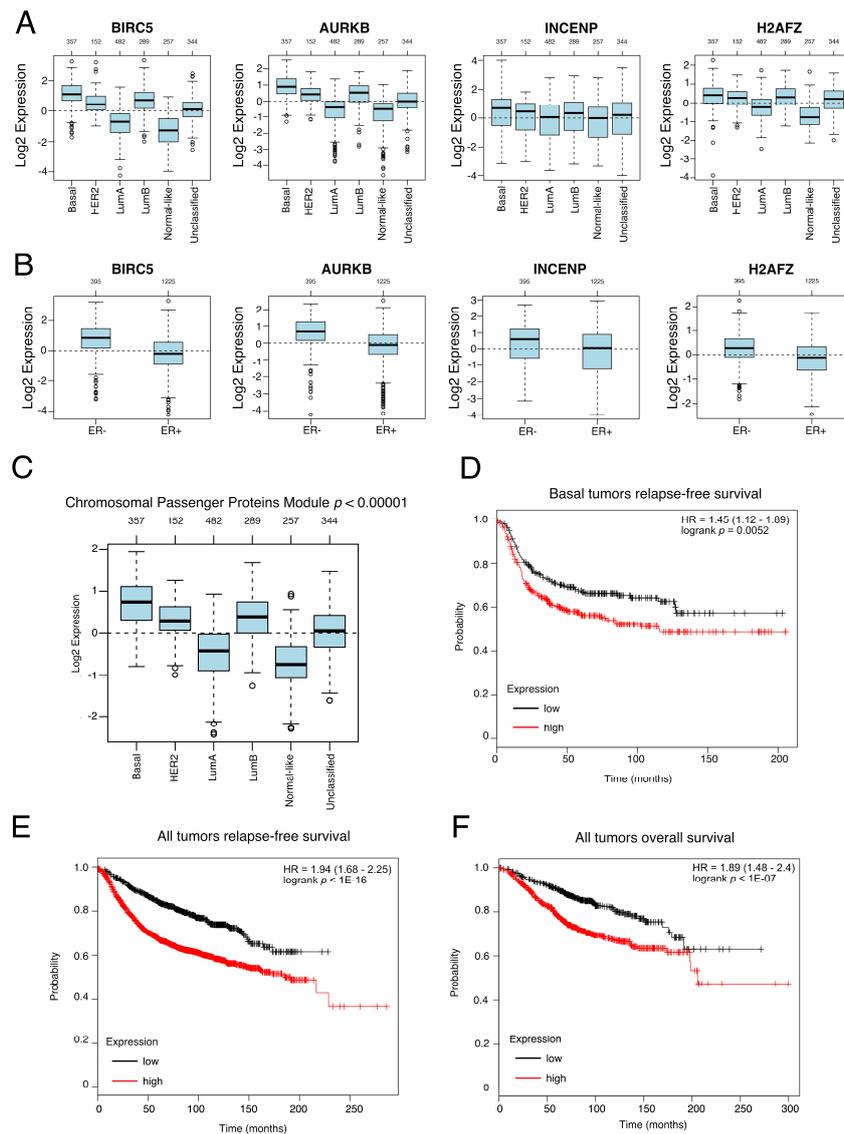
### 3.4. Genes from a CPPM Have Prognostic Significance in Breast Cancer Patients

Next, we analyzed the expression levels of this module in a previously reported breast cancer cohort [11] that includes 1881 breast cancer patients. All four genes showed higher expression levels in basal breast patients than the rest of the subtypes (Figure 4A) and, as expected, in ER negative tumors versus ER positive (Figure 4B). Concordantly, a higher level of the module’s composite expression was found in basal breast cancer patients (Figure 4C). It is important to note that basal breast tumors present stem expression signatures [3] and reciprocally, normal mammary stem cells overexpress a set of genes common to basal breast tumors [4]. Finally, we analyzed the prognostic significance of this gene expression module in a published set of 249 triple-negative breast cancer patients [14]. Figure 4D shows that basal breast tumor patients with overexpression of the CPPM had a worse prognosis (as

assessed by relapse-free survival), than those without it (logrank  $p = 0.0052$ , hazard ratio = 1.45 (range 1.12–1.89)). This was not limited to basal breast tumor patients since overexpression of this module was associated with lower relapse-free or overall survival independent of the molecular subtype. (Figure 4E, logrank  $p = 1 \times 10^{-16}$ , hazard ratio = 1.94 (range 1.68–2.25), Figure 4F, logrank  $p = 1.4 \times 10^{-17}$ , hazard ratio = 1.89 (range 1.48–2.4) and Figure S2). As expected, the hazard ratios were higher and the  $p$  value lower in the analyses using the composite module than in the separated genes (Figure S3). These results support the importance of the CPPM found in CSCs on breast cancer progression.



**Figure 3.** Participation of the Chromosomal Passenger Proteins Module (CPPM) in the stem cell phenotype of MDA-MB-468 cells. **(A)** Gene co-expression. Dot plots show the correlation between the expression of BIRC5 and AURKB (left panel), INCENP (middle panel) and H2AFZ (right panel). **(B)** MDA-MB-468 cells were transiently transfected with a vector containing a BIRC5 open reading frame and subjected to qRT-PCR for the genes shown. Log fold change refers to the logarithmic ratio between BIRC5-overexpressing cells versus empty vector-expressing cells [27], using the  $\Delta\Delta C_t$  method, with SDHA as a housekeeping gene. Each dot represents an independent transfection. BIRC5  $p = 0.0012$ , ALDH1A1  $p = 0.1547$ , Nanog  $p = 0.1137$  and Oct4  $p = 0.3128$  (\*\*  $p < 0.01$ , n.s. = non-significant). **(C)** Xenotransplant dilution assays using a zebrafish model. Table showing tumor frequencies 4 DPI. # cells: number of cells injected. For each experiment, 30 embryos were used and only informative (e.g., surviving) embryos were assessed. Percentages and number of fishes with or without tumors are shown. MDA-468-BIRC5 cells overexpress BIRC5 and MDA-468-GFP overexpress a GFP transgene (control cells). **(D)** Limiting dilution analysis obtained with the Extreme Limiting Dilution Assays (ELDA) software. The plot shows the percentage of the embryos with abdominal tumors injected with MDA-MB-468 cells overexpressing BIRC5 or GFP-expressing (control) cells at 4 DPI. The difference between the analyzed groups was significant (\*\* $p = 0.015$ ; estimated stem cell frequency 1/369 vs. 1/685 in BIRC5- and GFP-expressing cells, respectively; n.s. non-significant). **(E)** Examples of tumors formed in these assays (arrows).



**Figure 4.** Clinical relevance of the CPPM in breast cancer. **(A)** Boxplot showing the normalized expression of BIRC5, AURKB, INCENP and H2AFZ in a group of 1881 breast cancer patients by breast cancer subtypes. The number of patients is shown above the plot.  $p < 0.00001$  for all genes (ANOVA). **(B)** Boxplot showing the normalized expression of BIRC5, AURKB, INCENP and H2AFZ in a group of 1881 breast cancer patients by breast cancer ER status. The number of patients is shown above the plot.  $p < 0.00001$  for all genes (ANOVA). **(C)** Boxplot showing normalized expression of the CPPM in a group of 1881 breast cancer patients. The number of patients is shown above the plot,  $p < 0.00001$  (ANOVA). **(D)** Kaplan–Meier survival plot for relapse-free survival in patients with basal breast tumors derived from a published set of 249 triple-negative breast cancer patients [14], classified with the CPPM. High expression was defined as a mean expression of the module genes higher than 2093 logrank  $p = 0.0052$ , hazard ratio 1.45. **(E)** Kaplan–Meier survival plot for relapse-free survival in breast cancer patients derived from a published set of 249 triple-negative breast cancer patients [14], classified with the CPPM. High expression was defined as a mean expression of the module genes higher than 1223. logrank  $p = 1.10^{-16}$ , hazard ratio 1.94. **(F)** Kaplan–Meier survival plot for overall survival in breast cancer patients derived from a published set of 249 triple-negative breast cancer patients [14], classified with the CPPM. High expression was defined as a mean expression of the module genes higher than 1454. logrank  $p = 1.10^{-07}$ , hazard ratio 1.89.

#### 4. Discussion

In the present report, we sought to analyze the transcriptome of CSCs derived from basal breast cancer cells. Uncovering CSC expression pattern will aid to the comprehension of their cell of origin as well as improve the breast cancer diagnosis and treatment. As expected, we found enrichment in genes associated not only with somatic tissue stem cells, but specifically to mammary stem cells. In addition, we also found cellular processes commonly associated to stemness such as cell death and survival, proliferation and development. These findings support the similarities between breast normal and tumoral stem cells. We were also able to find a number of putative upstream regulators that need to be further explored. In particular, we found a group of regulators that are part of the shared TGF $\beta$  and IL1 signaling, including IL1RN, MAP3K7, TAB1 and TGF $\beta$ 1. This network plays a very important role in the responses elicited by environmental stress and, more important, it is able to induce a stem phenotype in breast cancer cells [20]. We also found a predicted decrease of several upstream regulators belonging to interferon signaling, including IFNL1, IFNA2, IFNG and Interferon alpha. It is interesting to note that IFN is able to inhibit breast cancer growth in *in vitro* and *in vivo* models [28,29], pointing toward the notion that this cytokine could be responsible for stem cells growth regulation. Indeed, we found that among the deregulated networks, cell growth was one of the most enriched GO functions. In our validation experiments, we isolated the stem population from MDA468 cells and seeded them in MammoCult<sup>TM</sup> media. MammoCult<sup>TM</sup> has been widely used in different types of cultures by many authors such as Choi HS et al. [30,31], Charpentier MS et al. [32] and Wolf J et al. [33], to evaluate and maintain breast cancer stem cells features while avoiding proliferation since it does not contain serum. Under these conditions, cells are not proliferating, so we cannot exclude that the enrichment in proliferation genes could be due to the cell culture conditions. Additional experiments using lineage tracing or *in vivo* visualization will be needed to address this result.

In addition to the signaling modules, we found the upregulation of a gene module consisting in BIRC5, AURKB, INCENP and H2AFZ. Three of these genes are part of the CPC, which is a key regulator and coordinator of several mitotic processes, such as activation of the spindle assembly checkpoint, correction of chromosome–microtubule attachment mistakes and regulation of the contractile structures that drives cytokinesis [26]. Chromosomal passenger proteins are the main molecular sensors of the spindle checkpoint [34] and thus, provide a key function to maintain ploidy in rapidly dividing cells. Since aneuploidy and the stemness phenotype are two of the main drivers of cancer progression [17], it has been proposed that these processes are tightly associated [18]. Indeed, modulating the expression of chromosomal passenger proteins, including BIRC5 and AURKB induces an increase in CSCs in various *in vitro* models [27,35]. In addition, it has been previously shown that BIRC5 is overexpressed in a subset of breast stem cancer cells [36,37]. These results support our findings.

Breast cancer is a heterogeneous disease in which molecular alterations, cellular composition, and clinical outcome provide a basis for diverse tumor classifications. In particular, Perou et al. identified four subtypes with clinical implications based on gene expression data: Luminal A (LumA), Luminal B (LumB), HER2-enriched and Basal-like [38]. In addition to classification, gene expression data has also been used as prognostic markers, as in the case of Oncotype DX® [39], a 21-gene signature that correlates with 10-year recurrence risk in ER-positive breast cancer. It is interesting to note that BIRC5 is one of the genes of the PAM50 signature (an intrinsic breast cancer subtype classifier) [40], and that high expression of this gene is one of the defining features of basal-like tumors, as defined by this approach. BIRC5 is also present in the Oncotype prognostic signature, which support its potential role as a potential prognostic marker. These findings are in concordance with the present results and help to provide a link between the expression of this gene and the main biological characteristic of basal breast cancer. Further analyses are required to establish if the CPPM could add additional value to the molecular classification of tumors, or, as we suggest, as a directed prognostic approach for basal-like breast tumors. Our finding that these proteins are co-regulated in CSCs further support this notion and points toward the presence of a still unknown common regulatory pathway that should be important in cancer progression, and an attractive therapeutic target. We also found that these proteins, alone or

in combination, are potential prognostic factors, underlying their importance in cancer progression and providing a link between stemness and ploidy regulation, which needs to be explored.

Finally, it is important to note that in the present study, we only worked with a basal breast cancer cell line. Additional experiments using cell lines derived from all four breast cancer subtypes will be required in order to ascertain if our results are applicable to breast cancer of all molecular subtypes or exclusively to basal-like breast tumors.

## 5. Conclusions

In conclusion, we performed a whole-genome transcriptome analysis of breast cancer stem cells derived from a TNBC cell line, which showed an enrichment of transduction cascades related to stemness, proliferation and apoptosis. Interestingly, a CPPM was found to be overexpressed in cancer stem cells. This module was overexpressed in triple-negative breast cancer tumors and was associated with poor prognosis. These results underline the importance of BCSCs in breast cancer progression and point toward the possible use of chromosomal passenger proteins as prognostic factors.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4409/9/3/709/s1>, Figure S1: Enriched signaling cascade of gene isoforms upregulated in Breast Cancer Stem Cells. Figure S2: Kaplan–Meier survival plot for relapse-free and overall survival in breast cancer patients' molecular subtypes. Figure S3: Kaplan–Meier survival plot for relapse-free and overall survival in breast cancer patients. Table S1: Primers used in the study. Supplementary Data: MDA-MB-468 mRNA Differential Expression with an FDR < 0.05.

**Author Contributions:** Conceptualization V.M. and J.M.-Z.; methodology, A.S.-C.y.C., M.E., G.C.-C., K.V.-S., I.B. and V.R.; Validation M.E., G.C.-C. and K.V.-S.; animal model, C.Z., bioinformatic analyses, J.M.-Z.; writing, A.S.-C.y.C., M.E., G.C.-C. and J.M.-Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Consejo Nacional de Ciencia y Tecnología (CONACyT) grant number A1-S-8462.

**Acknowledgments:** Angela Schwarz-Cruz y Celis is a doctoral student from Programa de Doctorado en Ciencias Biomedicas, Universidad Nacional Autónoma de México (UNAM) and received fellowship #351704 from Consejo Nacional de Ciencia y Tecnología (CONACyT). The results shown here are in part based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Perou, C.M.; Borresen-Dale, A.L. Systems biology and genomics of breast cancer. *Cold Spring Harb Perspect. Biol.* **2011**, *3*. [[CrossRef](#)] [[PubMed](#)]
2. Siegel, R.; Ma, J.; Zou, Z.; Jemal, A. Cancer statistics, 2014. *CA Cancer J. Clin.* **2014**, *64*, 9–29. [[CrossRef](#)] [[PubMed](#)]
3. Bertucci, F.; Finetti, P.; Birnbaum, D. Basal breast cancer: A complex and deadly molecular subtype. *Curr. Mol. Med.* **2012**, *12*, 96–110. [[CrossRef](#)] [[PubMed](#)]
4. Liu, R.; Wang, X.; Chen, G.Y.; Dalerba, P.; Gurney, A.; Hoey, T.; Sherlock, G.; Lewicki, J.; Shedden, K.; Clarke, M.F. The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl. J. Med.* **2007**, *356*, 217–226. [[CrossRef](#)] [[PubMed](#)]
5. Stingl, J.; Eirew, P.; Ricketson, I.; Shackleton, M.; Vaillant, F.; Choi, D.; Li, H.I.; Eaves, C.J. Purification and unique properties of mammary epithelial stem cells. *Nature* **2006**, *439*, 993–997. [[CrossRef](#)] [[PubMed](#)]
6. Schwarz-Cruz, Y.C.A.; Espinosa, M.; Maldonado, V.; Melendez-Zajgla, J. Advances in the knowledge of breast Cancer Stem Cells. A Review. *Histol. Histopathol.* **2016**, *31*, 601. [[CrossRef](#)]
7. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [[CrossRef](#)]
8. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinform.* **2011**, *12*, 323. [[CrossRef](#)]

9. Leng, N.; Dawson, J.A.; Thomson, J.A.; Ruotti, V.; Rissman, A.I.; Smits, B.M.; Haag, J.D.; Gould, M.N.; Stewart, R.M.; Kendzierski, C. EBSeq: An empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **2013**, *29*, 1035–1043. [[CrossRef](#)]
10. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U S A* **2005**, *102*, 15545–15550. [[CrossRef](#)]
11. Okamura, Y.; Aoki, Y.; Obayashi, T.; Tadaka, S.; Ito, S.; Narise, T.; Kinoshita, K. COXPRESdb in 2015: Coexpression database for animal species by DNA-microarray and RNAseq-based expression data with multiple quality assessment systems. *Nucleic Acids Res.* **2015**, *43*, D82–D86. [[CrossRef](#)] [[PubMed](#)]
12. Ringner, M.; Fredlund, E.; Hakkinen, J.; Borg, A.; Staaf, J. GOBO: Gene expression-based outcome for breast cancer online. *PLoS ONE* **2011**, *6*, e17911. [[CrossRef](#)] [[PubMed](#)]
13. Varley, K.E.; Gertz, J.; Roberts, B.S.; Davis, N.S.; Bowling, K.M.; Kirby, M.K.; Nesmith, A.S.; Oliver, P.G.; Grizzle, W.E.; Forero, A.; et al. Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res. Treat.* **2014**, *146*, 287–297. [[CrossRef](#)] [[PubMed](#)]
14. Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **2012**, *490*, 61–70. [[CrossRef](#)] [[PubMed](#)]
15. Gyorffy, B.; Lanczky, A.; Eklund, A.C.; Denkert, C.; Budczies, J.; Li, Q.; Szallasi, Z. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. Treat.* **2010**, *123*, 725–731. [[CrossRef](#)] [[PubMed](#)]
16. Mihaly, Z.; Kormos, M.; Lanczky, A.; Dank, M.; Budczies, J.; Szasz, M.A.; Gyorffy, B. A meta-analysis of gene expression-based biomarkers predicting outcome after tamoxifen treatment in breast cancer. *Breast Cancer Res. Treat.* **2013**, *140*, 219–232. [[CrossRef](#)]
17. Bharadwaj, R.; Yu, H. The spindle checkpoint, aneuploidy, and cancer. *Oncogene* **2004**, *23*, 2016–2027. [[CrossRef](#)]
18. Sansregret, L.; Swanton, C. The Role of Aneuploidy in Cancer Evolution. *Cold Spring Harb Perspect. Med.* **2017**, *7*. [[CrossRef](#)]
19. Hu, Y.; Smyth, G.K. ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **2009**, *347*, 70–78. [[CrossRef](#)]
20. Lehmann, B.D.; Bauer, J.A.; Chen, X.; Sanders, M.E.; Chakravarthy, A.B.; Shtyr, Y.; Pietenpol, J.A. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **2011**, *121*, 2750–2767. [[CrossRef](#)]
21. Asiedu, M.K.; Ingle, J.N.; Behrens, M.D.; Radisky, D.C.; Knutson, K.L. TGFbeta/TNF(alpha)-mediated epithelial-mesenchymal transition generates breast cancer stem cells with a claudin-low phenotype. *Cancer Res.* **2011**, *71*, 4707–4719. [[CrossRef](#)] [[PubMed](#)]
22. Memmi, E.M.; Sanarico, A.G.; Giacobbe, A.; Peschiaroli, A.; Frezza, V.; Cicalese, A.; Pisati, F.; Tosoni, D.; Zhou, H.; Tonon, G.; et al. p63 sustains self-renewal of mammary cancer stem cells through regulation of Sonic Hedgehog signaling. *Proc. Natl. Acad. Sci.* **2015**, *112*, 3499–3504. [[CrossRef](#)] [[PubMed](#)]
23. Shostak, K.; Chariot, A. NF-kappaB, stem cells and breast cancer: The links get stronger. *Breast Cancer Res.* **2011**, *13*, 214. [[CrossRef](#)]
24. Vermeulen, L.; De Sousa, E.M.F.; van der Heijden, M.; Cameron, K.; de Jong, J.H.; Borovski, T.; Tuynman, J.B.; Todaro, M.; Merz, C.; Rodermond, H.; et al. Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nat. Cell Biol.* **2010**, *12*, 468–476. [[CrossRef](#)]
25. Tam, W.L.; Lu, H.; Buikhuisen, J.; Soh, B.S.; Lim, E.; Reinhardt, F.; Wu, Z.J.; Krall, J.A.; Bierie, B.; Guo, W.; et al. Protein kinase C alpha is a central signaling node and therapeutic target for breast cancer stem cells. *Cancer Cell* **2013**, *24*, 347–364. [[CrossRef](#)] [[PubMed](#)]
26. Ceppi, P.; Hadji, A.; Kohlhapp, F.J.; Pattanayak, A.; Hau, A.; Liu, X.; Liu, H.; Murmann, A.E.; Peter, M.E. CD95 and CD95L promote and protect cancer stem cells. *Nat. Commun* **2014**, *5*, 5238. [[CrossRef](#)] [[PubMed](#)]
27. Pathak, S.; Multani, A.S. Aneuploidy, stem cells and cancer. *EXS* **2006**, *96*, 49–64.
28. Carmena, M.; Wheelock, M.; Funabiki, H.; Earnshaw, W.C. The chromosomal passenger complex (CPC): From easy rider to the godfather of mitosis. *Nat. Rev. Mol Cell Biol* **2012**, *13*, 789–803. [[CrossRef](#)] [[PubMed](#)]

29. Ling, X.; Marini, F.; Konopleva, M.; Schober, W.; Shi, Y.; Burks, J.; Clise-Dwyer, K.; Wang, R.Y.; Zhang, W.; Yuan, X.; et al. Mesenchymal Stem Cells Overexpressing IFN-beta Inhibit Breast Cancer Growth and Metastases through Stat3 Signaling in a Syngeneic Tumor Model. *Cancer Microenviron* **2010**, *3*, 83–95. [[CrossRef](#)]
30. Shen, C.J.; Chan, T.F.; Chen, C.C.; Hsu, Y.C.; Long, C.Y.; Lai, C.S. Human umbilical cord matrix-derived stem cells expressing interferon-beta gene inhibit breast cancer cells via apoptosis. *Oncotarget* **2016**, *7*, 34172–34179. [[CrossRef](#)]
31. Choi, H.S.; Kim, D.A.; Chung, H.; Park, I.H.; Kim, B.H.; Oh, E.S.; Kang, D.H. Screening of breast cancer stem cell inhibitors using a protein kinase inhibitor library. *Cancer Cell Int.* **2017**, *17*, 25. [[CrossRef](#)] [[PubMed](#)]
32. Choi, H.S.; Kim, J.H.; Kim, S.L.; Lee, D.S. Disruption of the NF-kappaB/IL-8 Signaling Axis by Sulconazole Inhibits Human Breast Cancer Stem Cell Formation. *Cells* **2019**, *8*, 1007. [[CrossRef](#)] [[PubMed](#)]
33. Charpentier, M.S.; Whipple, R.A.; Vitolo, M.I.; Boggs, A.E.; Slovic, J.; Thompson, K.N.; Bhandary, L.; Martin, S.S. Curcumin targets breast cancer stem-like cells with microtentacles that persist in mammospheres and promote reattachment. *Cancer Res.* **2014**, *74*, 1250–1260. [[CrossRef](#)] [[PubMed](#)]
34. Wolf, J.; Dewi, D.L.; Fredebohm, J.; Muller-Decker, K.; Flechtenmacher, C.; Hoheisel, J.D.; Boettcher, M. A mammosphere formation RNAi screen reveals that ATG4A promotes a breast cancer stem-like phenotype. *Breast Cancer Res.* **2013**, *15*, R109. [[CrossRef](#)]
35. Liang, Y.; Zhong, Z.; Huang, Y.; Deng, W.; Cao, J.; Tsao, G.; Liu, Q.; Pei, D.; Kang, T.; Zeng, Y.X. Stem-like cancer cells are inducible by increasing genomic instability in cancer cells. *J. Biol. Chem.* **2010**, *285*, 4931–4940. [[CrossRef](#)]
36. Nguyen, H.G.; Ravid, K. Tetraploidy/aneuploidy and stem cells in cancer promotion: The role of chromosome passenger proteins. *J. Cell Physiol.* **2006**, *208*, 12–22. [[CrossRef](#)]
37. Siddharth, S.; Das, S.; Nayak, A.; Kundu, C.N. SURVIVIN as a marker for quiescent-breast cancer stem cells-An intermediate, adherent, pre-requisite phase of breast cancer metastasis. *Clin. Exp. Metastasis* **2016**, *33*, 661–675. [[CrossRef](#)]
38. Yu, C.J.; Ou, J.H.; Wang, M.L.; Jialielihan, N.; Liu, Y.H. Elevated survivin mediated multidrug resistance and reduced apoptosis in breast cancer stem cells. *J. Buon.* **2015**, *20*, 1287–1294.
39. Perou, C.M.; Sorlie, T.; Eisen, M.B.; van de Rijn, M.; Jeffrey, S.S.; Rees, C.A.; Pollack, J.R.; Ross, D.T.; Johnsen, H.; Akslen, L.A.; et al. Molecular portraits of human breast tumours. *Nature* **2000**, *406*, 747–752. [[CrossRef](#)]
40. Kaklamani, V. A genetic signature can predict prognosis and response to therapy in breast cancer: Oncotype DX. *Expert Rev. Mol. Diagn* **2006**, *6*, 803–809. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).