



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE QUÍMICA

**ANÁLISIS DE LA EXPRESIÓN DE UN RNA LARGO NO
CODIFICANTE ALTERADO EN CÁNCER DE MAMA
EMPLEANDO BASES DE DATOS GENÓMICOS.**

T E S I S

QUE PARA OBTENER EL TÍTULO DE

QUÍMICA

PRESENTA:

Vaitiare Padilla Palma



CIUDAD UNIVERSITARIA, CDMX.

AÑO 2018



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

JURADO ASIGNADO:

PRESIDENTE: **Dra. Marisol López López**
VOCAL: **Dra. Tzvetanka Dimitrova Dinkova**
SECRETARIO: **Dr. Alfredo Hidalgo Miranda**
1er. SUPLENTE: **Dra. Sara Margarita Garza Aguilar**
2do. SUPLENTE: **Dr. Alberto Ortega Vázquez**

SITIO DONDE SE DESARROLLÓ EL TEMA:

Instituto Nacional de Medicina Genómica, INMEGEN.
Laboratorio de Genómica del Cáncer.

ASESOR DEL TEMA:
Dr. Alfredo Hidalgo Miranda

SUPERVISOR TÉCNICO:
M. en C. Alberto Cedro Tanda

SUSTENTANTE:
Vaitiare Padilla Palma

INDICE

| | |
|--|-------------|
| Lista de tablas | VI |
| Lista de figuras | VII |
| Resumen | VIII |
| Marco teórico | 1 |
| EPIDEMIOLOGÍA DEL CÁNCER DE MAMA: ¿CUÁL ES LA MAGNITUD DEL PROBLEMA Y EL IMPACTO DE LA ENFERMEDAD? | 1 |
| CÁNCER DE MAMA | 2 |
| <i>Diagnóstico y pronóstico</i> | 3 |
| <i>Características Distintivas del Cáncer</i> | 12 |
| TECNOLOGÍAS GENÓMICAS EN LA INVESTIGACIÓN DEL CÁNCER | 14 |
| TECNOLOGÍAS GENÓMICAS APLICADAS EN LA INVESTIGACIÓN DEL CÁNCER DE MAMA..... | 15 |
| <i>Panorama de alteraciones en el ADN</i> | 15 |
| <i>Panorama transcripcional</i> | 17 |
| <i>Herramientas de genómicas de clasificación de tumores de mama—Ensayos de expresión multigénicos</i> | 19 |
| ARNs NO CODIFICANTES (NCRNAS) | 21 |
| <i>Organización genómica de los ncRNAs</i> | 23 |
| MICRORNAs (MIRNAS)..... | 24 |
| <i>microRNA 205 (mir-205)</i> | 25 |
| RNAs LARGOS NO CODIFICANTES (LNCRNAs)..... | 27 |
| <i>Biogénesis</i> | 28 |
| <i>Clasificación de los lncRNAs</i> | 29 |
| CÁNCER Y ONCOLNCRNAS | 40 |
| LABORATORIO DE GENÓMICA DEL CÁNCER DEL INMEGEN | 43 |

| | |
|--|-----------|
| Objetivo general | 44 |
| Objetivos particulares | 45 |
| Hipótesis | 47 |
| Material y métodos..... | 47 |
| MATERIALES. | 47 |
| <i>Instrumentación.....</i> | <i>47</i> |
| <i>Insumos.....</i> | <i>48</i> |
| <i>Kits de biología molecular.....</i> | <i>49</i> |
| <i>Reactivos.....</i> | <i>49</i> |
| <i>Amortiguadores y disoluciones.....</i> | <i>50</i> |
| <i>q-PCR cebadores.....</i> | <i>50</i> |
| <i>Líneas celulares.....</i> | <i>51</i> |
| <i>Software</i> | <i>52</i> |
| <i>Bases de datos</i> | <i>52</i> |
| Métodos..... | 54 |
| <i>Minado de bases de datos: identificación de los lncRNAs candidatos.....</i> | <i>54</i> |
| <i>Diseño de cebadores</i> | <i>56</i> |
| <i>Selección del modelo de estudio in vitro.....</i> | <i>58</i> |
| <i>Muestras.....</i> | <i>59</i> |
| <i>Extracción de RNA de muestras.....</i> | <i>59</i> |
| <i>Extracción de RNA de líneas celulares.....</i> | <i>60</i> |
| <i>Electroforesis en gel de agarosa.....</i> | <i>61</i> |
| <i>Determinación de la concentración y pureza del RNA.....</i> | <i>62</i> |
| <i>Síntesis del DNA complementario (cDNA).....</i> | <i>63</i> |
| <i>Reacción en cadena de la polimerasa cuantitativa (q-PCR).....</i> | <i>63</i> |
| Resultados..... | 66 |
| <i>Identificación de los lncRNAs candidatos.....</i> | <i>66</i> |
| <i>Diseño de cebadores.....</i> | <i>71</i> |

| | |
|---|------------|
| <i>Estandarización de la q-PCR</i> | 72 |
| <i>Curva de eficiencia relativa</i> | 74 |
| <i>Reacción en cadena de la polimerasa (qPCR)</i> | 75 |
| <i>Descripción in silico de MIR205HG</i> | 76 |
| Discusión de resultados | 81 |
| IDENTIFICACIÓN DE 4 LNCRNAs DESREGULADOS EN DIVERSOS TIPOS DE CÁNCER. | 82 |
| EL LNCRNA MIR205-HG ESTÁ SUB-EXPRESADO EN CÁNCER DE MAMA..... | 85 |
| LA SUB-EXPRESIÓN DE MIR205-HG ESTÁ ASOCIADA A MAYOR SUPERVIVENCIA GLOBAL DE PACIENTES. | 86 |
| MIR205-HG PUEDE ACTUAR COMO POSIBLE CENSA Y REGULAR A MIR205. | 87 |
| Conclusiones | 88 |
| Perspectivas | 89 |
| Referencias | 90 |
| ANEXOS | 105 |
| ESTANDARIZACIÓN DE LOS LNCRNAs PARA Q-PCR. | 105 |
| <i>Estandarización del lncRNA MIR205HG</i> | 105 |
| CURVA DE EFICIENCIA RELATIVA..... | 106 |

Lista de tablas

| | |
|--|----|
| TABLA 1 . LISTA DE EQUIPOS | 48 |
| TABLA 2. LISTA DE INSUMOS | 48 |
| TABLA 3. LISTA DE KITS DE BIOLOGÍA MOLECULAR | 49 |
| TABLA 4. LISTA DE REACTIVOS QUÍMICOS | 49 |
| TABLA 5. LISTA DE BUFFERS Y DISOLUCIONES..... | 50 |
| TABLA 6. LISTA DE CEBADORES EMPLEADOS EN LA Q-PCR..... | 50 |
| TABLA 7. LISTA DE LÍNEAS CELULARES EMPLEADAS..... | 51 |
| TABLA 8. SOFTWARE EMPLEADO | 52 |
| TABLA 9. BASES DE DATOS EXTERNAS AL LABORATORIO DE GENÓMICA DEL CÁNCER | 52 |
| TABLA 10. BASE DE DATOS DEL LABORATORIO DE GENÓMICA DEL CÁNCER. | 54 |
| TABLA 11. LISTA DE CEBADORES EMPLEADOS EN LA Q-PCR CON SUS CARACTERÍSTICAS. | 58 |
| TABLA 12. TAMAÑOS DE BANDA ESPERADOS EN ELECTROFORESIS DE AGAROSA. | 61 |
| TABLA 13. PROGRAMA DEL TERMOCICLADOR: CDNA | 63 |
| TABLA 14. REACTIVOS PARA Q-PCR | 64 |
| TABLA 15. RELACIÓN DE REACTIVOS PARA Q-PCR..... | 64 |
| TABLA 16. PROGRAMA DEL TERMOCICLADOR: QPCR..... | 65 |
| TABLA 17. PROGRAMA DEL TERMOCICLADOR: CURVA DE DISOCIACIÓN | 65 |
| TABLA 18. LOS LNCRNAs CANDIDATOS Y CARACTERÍSTICAS DE SELECCIÓN. | 66 |
| TABLA 19. LISTA DE CEBADORES EMPLEADOS EN LA Q-PCR CON SUS CARACTERÍSTICAS. | 71 |
| TABLA 20. CONDICIONES DE CONCENTRACIÓN DE CEBADORES PARA LAS CURVAS DE DISOCIACIÓN. | 72 |
| TABLA 21. FUENTES DE VARIACIÓN..... | 73 |
| TABLA 22. PREDICCIÓN DE MIRNAS ASOCIADOS A AL LNCRNA MIRS205HG | 79 |

Lista de figuras

| | |
|--|--------------------------------------|
| FIGURA 1. SUPERVIVENCIA RELATIVA A DIEZ AÑOS DE ACUERDO CON LA ESTADIFICACIÓN (TNM) PARA EL CÁNCER DE MAMA (SINGLETERY, 2006; AMY M. FREMGEN, 1999). | 7 |
| FIGURA 2. CURVA DE SUPERVIVENCIA GLOBAL DE 294 PACIENTES CON CÁNCER DE MAMA (SANDHU, PARKER, JONES, LIVASY, & COLEMAN, 2010). | 10 |
| FIGURA 3. CARACTERÍSTICAS CELULARES DISTINTIVAS DEL CÁNCER QUE LES CONFIEREN VENTAJAS COMPETITIVAS. MODIFICADA DE (HANAHAN & WEINBERG, 2011)..... | 13 |
| FIGURA 4 BIOGÉNESIS DE PEQUEÑOS NCRNAS (FRIEDMAN, FARH, BURGE, & BARTEL, 2009; LEWIS, BURGE, & BARTEL, 2005). | ¡ERROR! MARCADOR NO DEFINIDO. |
| FIGURA 5 LOCALIZACIÓN GENÓMICA Y CONTEXTO DE LOS LNCNRNAS (L. MA ET AL., 2013)..... | 31 |
| FIGURA 6 MECANISMOS FUNCIONALES DE LOS LNCNRNAS (L. MA ET AL., 2013). | ¡ERROR! MARCADOR NO DEFINIDO. |
| FIGURA 7 LNCNRNAS IMPLICADOS EN DIVERSOS TIPOS TUMORALES (AZULES: SUBEXPRESADOS; ROJO: SOBREENEXPRESADOS) MODIFICADA DE (BARTONICEK ET AL., 2016). | 42 |
| FIGURA 8 CURVA KAPLAN-MEIER DE LINC01614. (J. LI ET AL., 2015). | 67 |
| FIGURA 9 CURVA KAPLAN-MEIER DE SLC12A5. (J. LI ET AL., 2015). | 68 |
| FIGURA 10 CURVA KAPLAN-MEIER DE MIR205-HG. (J. LI ET AL., 2015). | 69 |
| FIGURA 11 CURVA KAPLAN-MEIER DE FGF14-AS2 EN CÁNCER DE MAMA. (J. LI ET AL., 2015). | 70 |
| FIGURA 12 CURVA KAPLAN-MEIER DE FGF14-AS2 EN CÁNCER DE MAMA. (J. LI ET AL., 2015). | 71 |
| FIGURA 13 EXPRESIÓN DE MIR205HG DATOS DE Q-PCR. | 76 |
| FIGURA 14 EXPRESIÓN DE MIR205HG DATOS DE EXPRESIÓN DE RNA- SEQ DE TCGA..... | 76 |
| FIGURA 15 POSIBLE ESTRUCTURA TRIDIMENSIONAL DEL TRANSCRITO MIR205HG-4 (VOLDERS ET AL., 2013). | 77 |
| FIGURA 16 EXPRESIÓN DEL TRANSCRITO MIR205HG:4 ENTRE DISTINTOS TIPOS DE CÁNCER (MITRANSCRIPTOME)..... | 78 |
| FIGURA 17 VISTA INTEGRADORA DE LOS SITIOS DE UNIÓN DE MI RNAs Y REGULACIÓN TRANSCRIPCIONAL DE MIR205HG. | 80 |
| FIGURA 18 EXPRESIÓN DE MIR205HG POR ETAPAS EN CÁNCER DE MAMA..... | 80 |

Resumen

Los RNAs largos no codificantes (lncRNAs) tienen un rol crítico en el establecimiento y desarrollo del cáncer. A medida que se profundiza en el conocimiento de los lncRNAs oncogénicos, se contribuye a la identificación de posibles marcadores biológicos que jueguen un papel importante en el diagnóstico y/o sirvan como blancos terapéuticos. A partir del minado de la base de datos genómicos TCLA —The Cancer LncRNome Atlas—, se seleccionaron aquellos lncRNAs que estuvieran presentes en cáncer de mama y otros tipos tumorales, y que además su expresión se encontrara alterada en cáncer de mama —sobreexpresados (1144) y subexpresados (1109). De estos, se seleccionaron sólo aquellos relacionados con la supervivencia del paciente (Cox <0.05 y de $P < 0.04$). Bajo estos criterios se refinó la búsqueda y se encontró asociación con 11 lncRNAs sobreexpresados y 7 lncRNAs subexpresados. De los cuales se eligieron 2 sobreexpresados —LINC01614 y SLC12A5-AS1— y 2 subexpresados —MIR205HG y FGF14-AS2— cuya asociación a la supervivencia global de los pacientes fuera la mayor. Además, los cuatro lncRNAs seleccionados se encuentran desregulados en otros tipos tumorales como los adenocarcinomas de colon (COAD), pulmón (LUAD) y próstata (PRAD); y en los carcinomas de células escamosas de cabeza y cuello (HNSC), de células claras de riñón (KIRC) y de células escamosas de pulmón (LUSC).

MIR205HG es un lncRNA desregulado en tumores en comparación a los tejidos normales de seno, próstata y pulmón; está asociado a la supervivencia global del paciente y además su expresión disminuye conforme la enfermedad avanza en cáncer de mama mostrando un mínimo en la etapa IV, lo que sugiere que MIR205HG pudiera estar involucrado en la carcinogénesis mamaria. En este trabajo se analizó la expresión del lncRNA MIR205HG por medio de q-PCR en muestras de tumores de cáncer de mama no pareadas (T=16 y N=16). De acuerdo con el análisis de expresión realizado, MIR205HG se encontró expresado 0.29 (P=0.0113) veces más en tejido normal adyacente con respecto al tumor. Este dato se corroboró en una cohorte de 105 muestras de tejido normal adyacente y 837 tumores de cáncer de mama, en datos de secuenciación de RNA obtenidos de la base de datos TCGA (The Cancer Genome Atlas). De acuerdo con lo anterior, se propone que su función en el contexto de cáncer de mama pueda estar involucrada en el desarrollo y progresión de la enfermedad.

Además, se identificaron por medio de bases de datos genómicos las familias de miRNAs con las que interactuaría MIR205HG, de las cuales miR-143/1721/4770, miR-205/205ab, miR-214/761/3619-5p y miR-194 se encuentran caracterizados y se sabe que están implicados en diversos procesos de carcinogénesis no solo en cáncer de mama. Debido a esto, se propone que MIR205HG actúa como esponja de miRNAs en el citoplasma.

Marco teórico

Epidemiología del cáncer de mama: ¿Cuál es la magnitud del problema y el impacto de la enfermedad?

A nivel mundial el cáncer de mama es el segundo tipo de neoplasia maligna más común en ambos sexos y la más frecuente en mujeres con 1.67 millones de casos diagnosticados en el 2012. Dicha cifra representa el 25% de todos los pacientes con cáncer; esta enfermedad contribuye con el 14% de las muertes relacionadas al cáncer (Ferlay et al., 2015). El cáncer de mama no sólo tiene una alta tasa de incidencia sino también se caracteriza por una alta tasa de mortalidad, tan sólo en el mismo año ocasionó 552,000 defunciones (Ferlay et al., 2015; Sinn & Kreipe, 2013).

En México, a partir del 2006 el cáncer de mama se colocó como la segunda causa de muerte en mexicanas de 30 a 54 años y como la primera por tumores malignos en mujeres (F. Knaul et al., 2009). Cada año son diagnosticados 20,444 nuevos casos (Ferlay et al., 2015). En el país existen barreras —económicas, culturales, geográficas y falta de recursos para la detección temprana— que contribuyen a que únicamente se detecten el 12.7% de los casos por cáncer de mama en etapa 1. La probabilidad de supervivencia dentro de los primeros 10 años después del diagnóstico en los estadios avanzados es menor y disminuye drásticamente (estadios I/II: 95/88% supervivencia; estadios III/IV: 40/10% de supervivencia). Asimismo, el fracaso del tratamiento está directamente relacionado con la etapa en la que se detecta el cáncer (Al-Amri, 2005; Singletary & Connolly, 2006; Tricarico et al., 2002). En México el diagnóstico de cáncer

de mama se realiza en estadios tardíos, aunado a que el acceso al diagnóstico y tratamiento se ve dificultado por la falta de prevención primaria y secundaria (Al-Amri, 2005). Únicamente el 22% de las mujeres entre 40 y 69 años tuvieron acceso a una mastografía en el 2006 (F. M. Knaul et al., 2009; Nigenda, Caballero, & González-Robledo, 2009; Secretaría de Salud México, 2013).

Por lo descrito anteriormente, el cáncer de mama es un grave problema de salud a nivel nacional por sus altas tasas de mortalidad —14.8% muertes por cada 100 mil mujeres— e incidencia —16% por cada 100 mil mujeres— (INEGI, 2015). Por ello es de vital importancia que las estrategias de prevención y reducción de riesgos en la población general sean las adecuadas para que no se desarrollen cánceres invasivos; así mismo es indispensable tener una mayor comprensión de su biología para ser capaces de proponer nuevos y más eficientes métodos de detección, pronóstico y de respuesta a tratamientos para las pacientes, de esta manera poder hacer frente a este grave problema de salud en el país.

Cáncer de mama

El cáncer de mama es un proceso por el cual las células del tejido epitelial glandular adquieren características que permiten el crecimiento descontrolado y la capacidad de invadir tejidos circundantes y otros órganos (Ramsay, Kent, Hartmann, & Hartmann, 2005). A pesar de que se suele hablar del cáncer de mama como una sola enfermedad, en realidad es un conjunto de enfermedades histológica y bioquímicamente heterogéneas debido a cambios genéticos, epigenéticos y transcriptómicos con consecuencias clínicas diversas en el

pronóstico, en la progresión de la enfermedad, el riesgo de recurrencia y en la respuesta al tratamiento farmacológico (Baarnrd, Boeke, & Tamimi, 2015; Bertram, 2000; BI, 2013; Dieci, Orvieto, Dominici, Conte, & Guarneri, 2014; Manuscript, 2012; Perou et al., 2000; X. R. Yang et al., 2011; Zardavas, Irrthum, Swanton, & Piccart, 2015).

La evidencia epidemiológica indica que los principales factores de riesgo asociados al cáncer de mama se vinculan a la mayor exposición a ciclos ovulatorios: menarca temprana y menopausia tardía, esterilidad, nuliparidad y lactancia escasa o nula (Cóppola, Nader, & Aguirre, 2005; Rakha, Reis-Filho, & Ellis, 2008).

Diagnóstico y pronóstico

El diagnóstico actual se realiza por medio de algún examen inicial ya sea una mastografía, una exploración médica, un examen de radiografía o ultrasonido donde se pueda detectar alguna masa densa —cabe mencionar que no todos los nódulos en el seno son malignos, en estos caen la fibrosis y los quistes simples. En caso de detectarse una lesión sospechosa se realiza una biopsia, la única manera definitiva de determinar que es cáncer. Paralelamente se realiza un historial médico familiar y personal, y se hace una valoración con más estudios si la biopsia es positiva (Cianfrocca & Goldstein, 2004).

El manejo clínico actual sigue dependiendo de los factores de pronóstico tradicionales que incluyen factores histológicos, clínicos (tamaño de tumor, nódulos linfáticos axilares positivos, invasión vascular y linfática) y algunos biológicos como marcadores de pronóstico y predicción como los receptores hormonales, de estrógenos y progesterona (RE y RP) y el

receptor del factor de crecimiento epidérmico (HER-2) (Cianfrocca & Goldstein, 2004; Osborne, Yochmowitz, Knight, & McGuire, 1980). A pesar de todas las asociaciones que existen entre estas variables con el pronóstico y el desenlace, presentan limitaciones al evaluar el comportamiento de esta enfermedad (P. Wu et al., 2014). El cáncer de mama es clasificado por diversos sistemas de estadificación, cada uno de estos factores influencia el pronóstico y afecta la estrategia de tratamiento. Este trabajo abordará de una manera general las distintas clasificaciones y se centrará en la clasificación molecular (Anderson, Rosenberg, Prat, Perou, & Sherman, 2014; Bernard et al., 2009; Rodríguez-Cuevas, Macías, Franceschi, & Labastida, 2001).

Histopatología

El cáncer de mama por lo general es clasificado principalmente por su aspecto histológico. De acuerdo con la Organización Mundial de la Salud (WHO, por sus siglas en inglés) el cáncer de mama puede clasificarse en 21 subtipos histológicos con base en la morfología celular, crecimiento y patrones estructurales. La mayor parte de los tumores malignos de mama corresponde a adenocarcinomas y sobre todo a los carcinomas ductal y lobulillar.

Carcinoma ductal in situ (DCIS)

La condición en la cual las células anormales invaden el tejido epitelial de los ductos mamarios y puede llegar hasta los lobulillos. El 90% de los carcinomas son DCIS. Es considerado una forma no invasiva de cáncer de mama siempre y cuando se mantenga

confinado al interior de los ductos. Es impalpable pero visible en una mamografía. La probabilidad de que se convierta en cáncer invasivo es del 30 – 50% (Cassidy, et al., 2010; Lagios & Silverstein, 2001; Pérez & Soto, 2003; Schwartz, Solin, et al., 2000).

Carcinoma lobulillar in situ (LCIS)

Se refiere a las células que crecen dentro de los lobulillos del seno. No es un precursor de cáncer invasivo. Se considera marcador de riesgo para ambas mamas (Cassidy, Jim. Bissett, Donald. Spence, Roy. Payne, 2010; Haagensen, 1978; Silverstein et al., 1996).

Carcinoma invasivo ductal (CDI)

El 75% de los tumores son de este tipo. Las células malignas están asociadas al estroma y el tumor invade el seno a través de las regiones vasculares y linfáticas hasta llegar a los nódulos y el sistema circulatorio (Cassidy, Jim. Bissett, Donald. Spence, Roy. Payne, 2010).

Carcinoma invasivo lobulillar

Del 5 – 10% de los cánceres de mama son de este tipo, y el 20% de las pacientes lo desarrollan cáncer de mama contralateral. (Cassidy, Jim. Bissett, Donald. Spence, Roy. Payne, 2010).

Grado

Las células cancerígenas pierden su grado de diferenciación, su agrupación se vuelve desorganizada y sus núcleos se observan menos uniformes. El patólogo evalúa el grado de diferenciación que se observa en las células cancerígenas en comparación a las normales y las clasifica como de bajo, intermedio o alto grado conforme menos diferenciadas se encuentran. Los cánceres muy poco diferenciados son los de peor pronóstico (Friedewald et al., 2014).

Estadificación: TNM/AJCC.

El sistema de estadificación más común en cáncer de mama es el TNM, de acuerdo al American Joint Committee on Cancer Staging System for Breast Cancer (AJCC) (Singletary & Connolly, 2006). En la **Figura 1** se observa una gráfica de la supervivencia de pacientes con cáncer de mama estadificadas de acuerdo con el sistema TNM, donde el porcentaje de supervivencia relativa disminuye conforme aumenta el estadio de la paciente. En una cohorte 1.3 millones de casos obtenidos de la Base de Datos Nacional de Cáncer de E.U.A. (NCDB, por sus siglas en inglés) se asocian al estadio I y II una supervivencia relativa a diez años del 95/88% y posteriormente en los estadios III y IV es del 40/10% respectivamente (Bland et al., 1998; Singletary & Connolly, 2006).

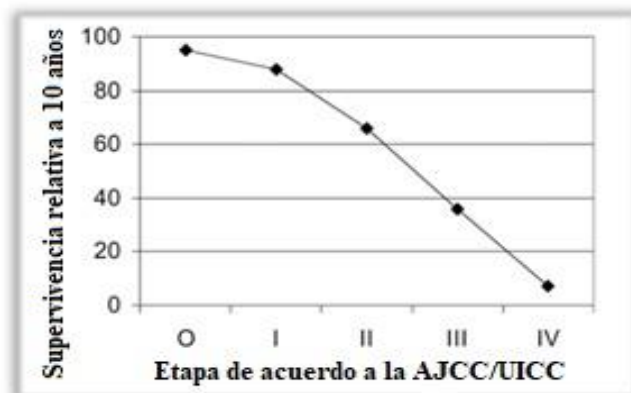


Figura 1. Supervivencia relativa a diez años de acuerdo con la estadificación (TNM) para el cáncer de mama (Singletary, 2006; Amy M. Fremgen, 1999).

- Tumor primario (T): hace referencia al tamaño y cuan extendido se encuentra el tumor primario. El tamaño del tumor se correlaciona directamente con la supervivencia. Sin embargo, el 20% de los tumores asintomáticos menores a 1 cm tienen metástasis nodulares.
- Compromiso de ganglios linfáticos regionales (N): alude al número de nódulos linfáticos cercanos con cáncer. La presencia o ausencia de nódulos linfáticos y el número de nódulos comprometidos son marcadores estrechamente relacionados con el curso clínico (Cianfrocca & Goldstein, 2004).
- Metástasis a distancia (M): proporciona información sobre si el cáncer se ha diseminado a alguna otra parte del cuerpo.

Posteriormente, se divide en etapas (0 al IV) de acuerdo a la progresión del tumor (American Joint Committee on Cancer Staging System for Breast Cancer, 2002; Shimm, 2008; Singletary & Connolly, 2006).

Expresión de receptores

La evaluación de tres receptores de crecimiento se realiza por inmunohistoquímica (IHC por sus siglas en inglés) y forman la base del diagnóstico actual: (1) receptores de estrógenos (RE) positivo —del 50 al 85% de los casos son positivos a RE— es asociado con desenlaces favorables y predice la respuesta a la terapia endócrina (Rosai J, 2011); (2) el grupo que co-expresa al receptor de progesterona (RP) refleja la vía de señalización del receptor de estrógenos debido a que el gen PGR se encuentra regulado por estrógenos; los tumores RE+/RP- o RE-/RP+ son menos propensos a responder al tamoxifen —sólo el 40% (Bardou, Arpino, Elledge, Osborne, & Clark, 2003; Cianfrocca & Goldstein, 2004; Elledge et al., 2000; Nieminen, 2008; Ravdin et al., 1992); (3) Receptor 2 del factor de crecimiento epidérmico humano (HER-2) un gen que se sabe es conductor del proceso de carcinogénesis, se encuentra amplificado en 17q21 y sobre-expresado en el 15% a 20% de los casos, su expresión se relaciona con un mal pronóstico y es marcador de predicción de respuesta a terapias dirigidas contra HER-2 (trastuzumab, pertuzumab, lapatinib) (Tanner et al., 2000). La amplificación del gen de HER- 2 (ERBB2) se detecta por medio de hibridación in situ (ISH, por sus siglas en inglés).

La terapia sistémica para tumores que no expresan ninguno de los marcadores anteriores [cáncer de mama triples negativos (TNBC, por sus siglas en inglés)] se encuentra limitada a quimioterapia, aunque se sabe que el desenlace en este grupo es altamente heterogéneo — algunos tumores exhiben una respuesta patológica completa, mientras que otros sufren de recaída dentro de los siguientes cinco años (Badve et al., 2011; Fulford et al., 2007). Este tipo de tumores son poco diferenciados, de alto grado histológico y tienen tasas bajas de

supervivencia. Probablemente los tumores triples negativos son el mejor ejemplo de la necesidad de nuevas terapias dirigidas que sean capaces de abordar subgrupos más definidos dentro del amplio espectro que son las categorías histológicas (Kalita-de Croft, Al-Ejeh, McCart Reed, Saunus, & Lakhani, 2016).

Clasificación molecular

Los análisis de expresión génica por microarreglos de ADN en tumores de cáncer de mama ha dado lugar al desarrollo de un sistema de clasificación basado en la biología tumoral más que en la morfología (Eliyatkin, Yalcin, Zengel, Aktaş, & Vardar, 2015). Las cuatro clases biológicas principales muestran diferencias significativas en la supervivencia global entre sí (Anderson et al., 2014; Rakha & Ellis, 2011; Eroles, Bosch, Alejandro Pérez-Fidalgo, & Lluch, 2012; Gerson Cwilich, Alban de la Torre, Villalobos Prieto, & Serrano Olvera, 2012; Gruvberger-Saal, Cunliffe, Carr, & Hedenfalk, 2006; Marchionni et al., 2008). La clasificación molecular actual incluye cuatro subtipos moleculares Luminal A y B, Her-2 y Basal (Rakha & Ellis, 2011).

En la **Figura 2.** se muestra una curva Kaplan Meier de supervivencia global correspondiente a 294 pacientes con cáncer de mama clasificados por los subtipos moleculares Luminal A, Luminal B, HER-2 y Basal. En dicha curva se puede discernir claramente entre los distintos subtipos moleculares donde se observa que el subtipo Luminal A es el de mejor pronóstico.

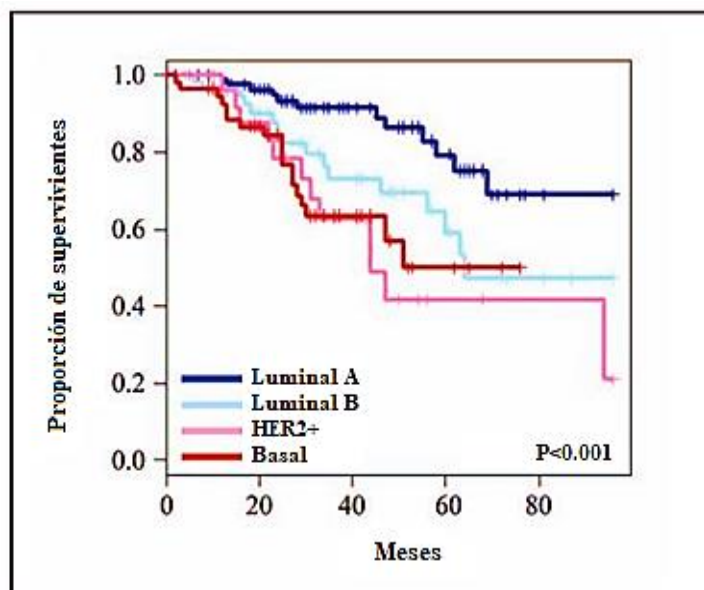


Figura 2. Curva de supervivencia global de 294 pacientes con cáncer de mama (Sandhu, Parker, Jones, Livasy, & Coleman, 2010).

Luminal A

Llamado así debido a que este grupo de tumores expresa las proteínas típicas de las células epiteliales luminales del seno. Es el subtipo molecular más común —50%-60% de los casos por cáncer de mama. Estos tumores presentan bajo grado histológico, buen pronóstico. Se define como tumores RE positivos o RP positivos, negativos a HER-2. El tratamiento se basa principalmente en la hormonoterapia (Haque et al., 2012; Howlader et al., 2000; Metzger-Filho et al., 2013; Siddiqui, Piyush, Faridi, Ahmad, & Khan, 2016; Yersal & Barutca, 2014).

Luminal B

Este subtipo comprende del 15% al 20% de los casos de cáncer de mama (Fan et al., 2006). Se caracteriza por un fenotipo más agresivo que el Luminal A, presenta un grado histológico mayor y un peor pronóstico. Este subtipo tiene altas tasas de recurrencia (Haque et al., 2012;

Howlader et al., 2000; Metzger-Filho et al., 2013; Reis-Filho, Weigelt, Fumagalli, & Sotiriou, 2010; Siddiqui et al., 2016; Yersal & Barutca, 2014). (Sotiriou et al., 2006; Yersal & Barutca, 2014).

HER-2

Esta categoría abarca del 15% al 20% de los casos de cáncer de mama. Tienen un comportamiento clínico muy agresivo. El 70% de los tumores pertenecientes a esta clase expresan en altos niveles a HER2 y genes asociados a esta vía (Perou et al., 2000) . Estos tumores tienen altas tasas de proliferación, el 75% de estos tumores presenta alto grado histológico y más del 45% tiene mutaciones en p53 (Dai, Chen, & Bai, 2014; Tsutsui et al., 2003; Yersal & Barutca, 2014).

Basal

Representa del 8% al 37% de los casos de cáncer de mama (Rakha et al., 2009). La mayoría de los tumores son carcinomas invasivos ductales, tienen un comportamiento clínico agresivo y altas tasas de metástasis al cerebro y pulmón (Heitz et al., 2009; Yersal & Barutca, 2014). Presentan mutaciones frecuentes en el gen de la proteína tumoral p53 (TP53), existe evidencia de inestabilidad genómica y activación de la vía del retinoblastoma (Rb). Se caracteriza por recurrencias en corto tiempo y patrones definidos de metástasis (Foulkes et al., 2004).

Características Distintivas del Cáncer

Conforme las células normales progresan a un estado neoplásico los genes implicados en los mecanismos de homeóstasis —aquellos que controlan la proliferación y muerte celular— sufren mutaciones. En consecuencia, se activan genes que estimulan la proliferación —oncogenes— y se inactivan aquellos que normalmente inhiben la proliferación —genes supresores tumorales (Bertram, 2000). Estas mutaciones permiten a la célula adquirir una serie de alteraciones distintivas y complementarias entre sí. Dichas alteraciones en conjunto hacen posible el crecimiento tumoral y la diseminación metastásica. Estos rasgos puntuales son mejor conocidos como las Características Distintivas del Cáncer (Hanahan & Weinberg, 2011), las cuales constan de ocho principios funcionales: señalización proliferativa sostenida, pérdida de respuesta ante inhibidores de crecimiento, inmortalidad replicativa, inducción de la angiogénesis, activación de la invasión y metástasis, reprogramación del metabolismo energético celular, evasión del sistema inmune; y dos que propician el desarrollo de las anteriores: inestabilidad genómica y mutación, e inflamación promovida por el tumor (**Figura 3**).



Figura 3. Características celulares distintivas del cáncer que les confieren ventajas competitivas. Modificada de (Hanahan & Weinberg, 2011)

Como se mencionó, la adquisición de las ocho características funcionales mencionadas depende en gran medida de las alteraciones que ocurran en el genoma de las células neoplásicas—inestabilidad genómica y mutaciones. Los genotipos mutados son capaces de proveer ventajas selectivas a ciertas subclonas celulares al acumular mutaciones que comprometan los mecanismos de reparación celular. Por otro lado, el sistema inflamatorio es capaz de ayudar inadvertidamente a las células cancerígenas al suplir factores de crecimiento, enzimas de degradación de la matriz celular, así como factores pro-angiogénicos (Denardo et al., 2009; Hamilton, 2008; E. Y. Lin, Nguyen, Russell, & Pollard, 2001).

Tecnologías genómicas en la investigación del cáncer

En las últimas décadas, el surgimiento de proyectos propiciado por el desarrollo de las tecnologías genómicas ha permitido el avance acelerado de la biología molecular del cáncer. El primero, fue el Proyecto del Genoma Humano establecido en 1990 por los Institutos Nacionales de Salud de E.U. y el Centro Sanger del Reino Unido (Beggs & Dilworth, 2015; Dulbecco, 1986). Este proyecto permitió establecer un genoma de referencia por medio de la secuenciación por fluorescencia, técnica descrita por Sanger y sus colegas (Sanger, Nicklen, & Coulson, 1977). El proyecto tomó 10 años en poder completar un genoma completo y costó aproximadamente tres mil millones de pesos, secuencia que fue liberada en el año 2003 (Beggs & Dilworth, 2015).

Los microarreglos y las técnicas de secuenciación de nueva generación han permitido estudiar el panorama de alteraciones genómicas en los distintos tipos de cáncer (Surrey, Luo, Chang, & Li, 2016) por medio de la creación de consorcios internacionales enfocados en la caracterización genómica —mutaciones, cambio en el número de copias, rearrreglos cromosómicos, entre otros— transcriptómica y epigenómica de por lo menos 50 tipos de cáncer en cientos de individuos; en la identificación de los genes conductores de la patología y de posibles blancos terapéuticos. Los dos proyectos principales son el Consorcio Internacional del Genoma del Cáncer (ICGC, por sus siglas en inglés) (McGovern et al., 2011) y el Atlas del Genoma del Cáncer (TCGA, por sus siglas en inglés) financiado por el National Cancer Institute (NCI, por sus siglas en inglés) y el National Human Genome Research Institute (NHGRI, por sus siglas en inglés) (Manuscript, 2012; Tomczak, Czerwińska, & Wiznerowicz, 2015). Una característica de este tipo de colaboraciones es su gran poder

estadístico al evaluar más de 11,000 pacientes (muestras pareadas de tejidos tumorales y normales) con altos estándares de calidad (O'Keefe, 2001) en 33 cánceres de distintos tejidos. Los datos y la información generada por los proyectos son de uso libre.

Tecnologías genómicas aplicadas en la investigación del cáncer de mama

Panorama de alteraciones en el ADN

Actualmente los proyectos internacionales se orientan a la creación de repertorios extensos que enlisten genes responsables de la iniciación y progresión del cáncer. Este tipo de estudios involucran la secuenciación de muestras tumor-normal pareadas y el posterior análisis matemático como estrategia de identificación de genes donde ocurran las mutaciones con mayor frecuencia —es decir, que la probabilidad de que ocurra alguna mutación sea mayor que en genes con mutaciones aleatorias y por ende sean genes conductores.

Este tipo de proyectos tienen como último fin el develar la vasta heterogeneidad en las alteraciones presentes en tumores de mama. Entre los más importantes se encuentran: el Consorcio Internacional del Genoma del Cáncer (ICGC, por sus siglas en inglés) el cual examinó el panorama mutacional de cien tumores e identificó cuarenta mutaciones somáticas conductoras —es decir aquellas que confieren ventajas selectivas a las subclonas portadoras y que están causalmente implicadas en la oncogénesis (Stratton, Campbell, & Futreal, 2009)— de las cuales nueve no se encontraban reportadas en la literatura (Stephens et al., 2012). Recientemente, el consorcio liberó los datos de 560 genomas donde identificó doce firmas

mutacionales y presentó una lista definitiva con 93 genes conductores en cáncer de mama (Nik-Zainal et al., 2016). En conjunto estos descubrimientos delinearon el panorama de las mutaciones somáticas en los procesos de reparación de daños al ADN y destacaron la inestabilidad del genoma —6 firmas de rearrreglos— como partícipe en la adquisición de mutaciones somáticas (Alexandrov, Nik-Zainal, Wedge, Campbell, & Stratton, 2013; Morganella et al., 2016; Nik-Zainal et al., 2016; Stephens et al., 2012).

El Laboratorio de Genómica del Cáncer del INMEGEN ha sido partícipe de diversos proyectos orientados al estudio de la heterogeneidad en cáncer en busca de nuevos genes conductores (Lawrence et al., 2013) y en el estudio de diversas alteraciones genómicas en distintos subtipos de cáncer de mama (Banerji et al., 2014). Este último fue una colaboración entre el Laboratorio de Genómica del Cáncer del INMEGEN y diversos institutos de EUA —The Broad Institute of MIT and Harvard, Dana-Farber Cancer Institute y el Beth Israel Deaconess Medical Center, entre otros— en el cual se realizó la secuenciación de exoma completo de 103 tumores pareados de mama y la secuenciación de genoma completo de 22 muestras pareadas tumor-normal de diversos subtipos de pacientes mexicanas y vietnamitas. En donde, además de confirmar mutaciones somáticas recurrentes en PIK3CA11, TP536, AKT112, GATA313 y MAP3K110, se descubrieron mutaciones comunes en el factor de transcripción CBFβ y deleciones en RUNX1. En específico, se identificó la fusión MAGI3-AKT3 enriquecida en los tumores triples negativos.

Panorama transcripcional

Los primeros estudios que aportaron en el conocimiento de la biología del cáncer de mama para su categorización y clasificación molecular fueron publicados en la década de los 90's, cuando por medio de análisis de pérdida de heterocigocidad (LOH, por sus siglas en inglés) y de hibridación genómica comparativa (CGH, por sus siglas en inglés) se pudieron identificar alteraciones —pérdidas, ganancias y amplificaciones de secuencias de ADN genómico— capaces de estratificar a los cánceres de mama en dos categorías, de alto y bajo grado (Lehmann et al., 2011; Sorlie et al., 2003). Estos análisis proveyeron el marco para el desarrollo de un sistema altamente sofisticado de clasificación transcriptómica del cáncer de mama relacionado con características clínicas particulares, a inicios del siglo XXI. A partir de entonces surgieron numerosos estudios que desarrollaron firmas génicas capaces de distinguir y reagrupar a los tumores en los subtipos moleculares (Curtis et al., 2012; Sorlie et al., 2001; Marc J van de Vijver et al., 2002).

Estudios como METABRIC (Curtis et al., 2012)—evaluó los perfiles de expresión por arreglos de 2500 tumores de mama; y TCGA (The Cancer Genome Atlas Network, 2012)—analizó 547 tumores de mama por medio de arreglos de expresión y 817 por secuenciación de ARN (Ciriello et al., 2015)— han permitido encontrar firmas mutacionales con aplicación clínica como indicadores de pronóstico: T-cell metagene, el cual está asociado a buen pronóstico en tumores HER2+/RE- (Rody et al., 2009); firmas asociadas a inestabilidad cromosómica (Carter, Eklund, Kohane, Harris, & Szallasi, 2006); y firmas asociadas al beneficio de quimioterapia como terapia adyuvante (Bernard et al., 2009).

Actualmente, existen distintos tipos de biomarcadores identificados —clasificados por su naturaleza química—, estos incluyen biomarcadores basados en ADN, en ARN, en proteínas y biomarcadores epigenéticos. A nivel de ADN se encuentran alteraciones cromosómicas que incluyen a los cambios de un solo nucleótido (SNVs), pequeños InDels —pequeñas inserciones o deleciones—, variantes estructurales (SVs) y alteraciones en el número de copias (CNVs) (Wing et al., 2011). A nivel de ARN, los biomarcadores incluyen mRNAs, miRNAs, lncRNAs y otros ncRNAs sobreexpresados o subexpresados. Los biomarcadores proteicos incluyen glicoproteínas como la Alfafetoproteína —AFP—, antígeno del cáncer 125 —CA125— y antígenos tumorales como el antígeno prostático específico —PSA, para cáncer de próstata— y el antígeno carcinoembrionario —CEA, para cáncer de colon. La creación de perfiles de expresión de mRNAs demostró que la expresión alterada de este tipo de mRNAs oncogénicos —como KRAS y MYC— y genes supresores tumorales (TSGs) —como APC, BRAF, TP53 entre otros— usualmente está relacionada con el desarrollo del cáncer. El primer estudio de perfiles de expresión de mRNAs que se realizó empleó microarreglos de Affymetrix con 6817 genes humanos e identificó clasificadores tumorales en las leucemias mieloide aguda (AML) y linfoblástica aguda (ALL) (Golub et al., 2016). En este estudio se identificó un panel de 50 genes generado a partir del análisis de 38 muestras de médula ósea —11 AML y 27 ALL— que demostró ser capaz de clasificar ambos tipos de leucemias (M. Li, Fu, & Xiao, 2015). Desde entonces se generó una ola de estudios que emplearon los microarreglos como herramienta en la identificación de firmas de biomarcadores en cáncer de pulmón (Botling et al., 2013), en cáncer gástrico (Hippo et al., 2002), en cáncer de próstata (Bull et al., 2001), cáncer de mama (van 't Veer et al., 2002; Marc J van de Vijver et al., 2002) y otros.

Herramientas de genómicas de clasificación de tumores de mama—Ensayos de expresión multigénicos

En E.U.A, las pruebas de diagnóstico son reguladas por la Administración de Alimentos y Medicamentos (FDA). A continuación, se explicarán brevemente algunos de los clasificadores genómicos para cáncer de mama aprobados por la FDA, los cuales se consideran herramientas de apoyo en el diagnóstico o pronóstico a los factores de pronóstico tradicionales.

Mamaprint

Desarrollada por Agendia es una herramienta de evaluación de pronóstico personalizado que calcula el riesgo de recurrencia de cáncer de mama dentro de los primeros cinco años al diagnóstico y como evaluador predictivo informa sobre la prescripción de quimioterapia. Evalúa la expresión de 70 genes de mRNAs por microarreglos en tejidos frescos congelados (FF). Las pacientes deben ser menores a 61 años, negativas a nódulos linfáticos (NL), estar en estadio I/II y el tamaño del tumor debe ser ≤ 5 cm (Buyse et al., 2006; Cardoso et al., 2008; van 't Veer et al., 2002; M. J. Van De Vijver et al., 2009; Marc J van de Vijver et al., 2002).

Oncotype DX

Desarrollado por Genomic Health asigna una puntuación de recurrencia (RS, Recurrence Score por sus siglas en inglés) a 10 años después del diagnóstico. El ensayo es una PCR en tiempo real (qRT-PCR) de 21 genes de mRNAs. Se encuentra incluido en los lineamientos de la Sociedad Americana de Oncología Clínica (ASCO, por sus siglas en inglés). Predice el beneficio que tendrá la quimioterapia en conjunto a la terapia endócrina. Pacientes RE+/NL-

y tejidos fijados en formalina, embebidos en parafina (FFPE) (Goldstein et al., 2008; Kim et al., 2004).

Prosigna

Desarrollada por nanoString Technologies, esta prueba se basa en PAM50 (firma de 50 genes de mRNAs) que define a los subtipos moleculares (véase clasificación molecular) en cáncer de mama. Valor pronóstico: es una escala de recurrencia basada en los subtipos moleculares, tamaño tumoral, número de nódulos linfáticos comprometidos e índice de proliferación. Valor predictivo: Evalúa la eficacia del Tamoxifen. Pacientes postmenopáusicas positivas a receptores hormonales (RH+), NL- (en estadio I o II), NL+ (en estadio II) (van 't Veer et al., 2002).

EndoPredict

Desarrollado por Myriad genetics esta firma de 8 genes de mRNAs sirve para evaluar una escala de quimioterapia con cis-platino y etopósido (EP), y de esta manera calcular el beneficio que puede acarrear al tratamiento. Su carácter pronóstico radica en valorar recurrencia a largo plazo en pacientes RE+/HER2-. Determina que pacientes no necesitan quimioterapia. Las pacientes deben ser ER+/HER2- y negativas a NL (Blank et al., 2015; Martin et al., 2016).

El que existan varias firmas génicas comerciales integradas en la práctica clínica refleja su valor percibido sobre el pronóstico tradicional (Sparano et al., 2015). El estudio de la expresión y función de microRNA (miRNA), largos no codificantes (lncRNA) y otros ARNs no codificantes es un ámbito que está cobrando importancia como blancos potenciales terapéuticos, de diagnóstico y predictivos. Los lncRNAs tienen funciones diversas pero críticas en los procesos celulares, y su contribución al proceso de oncogénesis ha sido demostrada,

como se detallará en las secciones posteriores (Huarte, 2015). Sun et al. han demostrado que una firma de nueve lncRNAs puede clasificar a los pacientes en grupos de alto y bajo riesgo de diseminación metastásica, independientemente de las variables clínicas evaluadas de rutina. Además, otra firma de lncRNAs es capaz de estimar e identificar aquellos pacientes que se podrían beneficiar de terapia adyuvante (Sørensen et al., 2015).

El estudio de la expresión y función de microRNA (miRNA), largos no codificantes (lncRNA) y otros ARNs no codificantes es un ámbito que está cobrando importancia como blancos potenciales terapéuticos, de diagnóstico y predictivos. Los lncRNAs tienen funciones diversas pero críticas en los procesos celulares, y su contribución al proceso de oncogénesis ha sido demostrada, como se detallará en las secciones posteriores (Huarte, 2015). Sun et al. han demostrado que una firma de nueve lncRNAs puede clasificar a los pacientes en grupos de alto y bajo riesgo de diseminación metastásica, independientemente de las variables clínicas evaluadas de rutina. Además, otra firma de lncRNAs es capaz de estimar e identificar aquellos pacientes que se podrían beneficiar de terapia adyuvante (Sørensen et al., 2015).

ARNs no codificantes (ncRNAs)

El genoma humano alberga 23,000 genes que codifican para proteínas, esta cifra representa menos del 2% de la secuencia del genoma total. Sin embargo, el 90% del genoma total es transcrito activamente de acuerdo al proyecto GENCODE (Hubbard et al., 2002; Sana, Faltejsova, Svoboda, & Slaby, 2012; Steijger et al., 2013) y del que diversos estudios han probado que los transcritos que no codifican para proteínas son funcionales más que ser

residuos transcripcionales (Djebali et al., 2012). Los resultados encontrados por proyectos como GENCODE han demostrado que el transcriptoma humano resulta ser mucho más complejo que el vasto repertorio de genes codificantes y sus isoformas, y que adicionalmente expresa ampliamente ARNs no codificantes, secuencias que se traslapan y algunas antisentido (Stein, 2004; Taft, Pang, Mercer, Dinger, & Mattick, 2010). Aunque inicialmente se consideraba a los ARNs no codificantes (ncRNAs) como desechos transcripcionales o evolutivos provenientes de la creación de genes y/o inserción de elementos móviles genéticos, pruebas recientes sugieren que estos jueguen papeles cruciales en diversos procesos biológicos y oncogénicos (Qu & Adelson, 2012).

Los ARNs no codificantes (ncRNAs) son una clase de moléculas de ARN transcritas sin el potencial de codificar para proteínas. Algunos ncRNAs se han conocido durante mucho tiempo y se encuentran muy bien caracterizados como los ARNs de transferencia (tRNAs), los ARNs ribosomales (rRNAs), los ARNs de corte y empalme (snRNAs), y los pequeños nucleolares (snoRNAs). Actualmente, los ncRNAs se clasifican de acuerdo con el tamaño del transcrito, ARNs pequeños no codificantes y ARNs largos no codificantes (lncRNAs). Los ncRNAs pequeños son representados por una amplia gama de especies de ARNs muy bien caracterizados. Esta clase incluye a los ARNs (siRNA), microRNAs (miRNAs), piwi-RNAs (piRNAs). En contraste con los miRNAs, los ARNs largos no codificantes (lncRNAs) son transcritos tipo mRNAs que varían en longitud de 200 nucleótidos (nt) a 100 kilobases (kb) y carecen de un marco de lectura abierto. Los niveles de expresión de los lncRNAs son bajos en comparación con los niveles de expresión de genes codificantes para proteínas y algunos lncRNAs son tejido-específicos. Los lncRNAs han sido asociados a procesos biológicos como

el de empalme alternativo de mRNAs o en la importación nuclear. Además, los lncRNAs han demostrado ser componentes estructurales, precursores de ARNs pequeños e incluso reguladores de la degradación de mRNAs. Aunque existe un número pequeño de lncRNAs anotados y bien caracterizados, muchos estudios sugieren que esta clase de ncRNA representa más de la mitad del transcriptoma (Redova, Svoboda, & Slaby, 2011; Sana, Hajduch, Michalek, Vyzula, & Slaby, 2011; Slaby, Bienertova-Vasku, Svoboda, & Vyzula, 2012; Slaby, Svoboda, Michalek, & Vyzula, 2009). En numerosos estudios se observa que la alteración de la expresión de estos lncRNAs (HOTAIR, MALAT1, HULC, T-UCRs, etc.) se comparte en diversos tipos de cáncer, lo que sugiere que su expresión aberrante puede ser un factor que contribuya al establecimiento y progresión del cáncer (Costa, 2005; Okamura et al., 2009; Qu & Adelson, 2012; Rinn et al., 2007)

Organización genómica de los ncRNAs

Los ncRNAs se originan de distintas regiones genómicas como las secuencias intrónicas —que constituyen más del 30% del genoma humano (Mattick and Gagen, 2001) y son altamente conservadas (Taft et al., 2007). Los ncRNAs también se derivan de las hebras sentido y antisentido de varias regiones genómicas —regiones que se traslapan con secuencias codificantes (Dinger et al., 2008) y regiones intergénicas (lincRNAs).

MicroRNAs (miRNAs)

La subclase más estudiada de los ncRNAs pequeños es la de microRNAs (miRNAs), descubierta por Victor Ambros en *Caenorhabditis elegans* (R. C. Lee, Feinbaum, & Ambros, 1993). Los miRNAs son moléculas de ARN de cadena simple de 18-25 nucleótidos, evolutivamente conservadas (Axtell & Bartel, 2005; Chen & Rajewsky, 2007; C.-T. Lee, Risom, & Strauss, 2007; Peterson, Dietrich, & McPeck, 2009; Tanzer & Stadler, 2004), e implicadas en la regulación de la expresión génica en eucariontes (Lagos-Quintana, Rauhut, Lendeckel, & Tuschl, 2001; R. C. Lee & Ambros, 2001; Pasquinelli et al., 2000; Reinhart et al., 2000). Se predice que los genes de miRNAs representan el 1-2 % del genoma humano y controlan la actividad del ~50% de todos los genes que codifican para proteínas (Friedman, Farh, Burge, & Bartel, 2009; Lewis, Burge, & Bartel, 2005).

Los miRNAs reprimen la expresión génica de su blanco post-transcripcionalmente, al parecer a nivel de traducción, por un complejo miRISC. La complementariedad perfecta entre el mRNA-miRNA ocasiona la ruptura del mRNA blanco, mientras que el mal apareamiento promueve la represión de la traducción del mRNA. Los miRNAs que usan este mecanismo reducen los niveles de proteína de sus genes objetivo, pero los niveles de mRNA de estos genes apenas son afectados (Barrangou et al., 2017; Chi, Zang, Mele, & Darnell, 2009; Kwak, Iwasaki, & Tomari, 2010; Lim et al., 2005; Ørom, Nielsen, & Lund, 2008; N.-P. Tsai, Lin, & Wei, 2009).

La mayor parte de los miRNAs descritos hasta el momento están implicados en los procesos celulares como proliferación, diferenciación, y apoptosis, y en el desarrollo y progresión de

varios trastornos incluyendo cáncer —donde se han encontrado miRNAs empleados en firmas génica de pronóstico y predicción (Cancer et al., 2012; Jansson & Lund, 2012; Sana et al., 2011; Satoh, 2010; Slaby et al., 2010).

microRNA 205 (mir-205)

El gen humano miR-205 se encuentra localizado dentro de la secuencia del largo no codificante miR-205HG (LINC00510) y se ubica en la región cromosómica 1q32.2. La secuencia de mir-205 es de 110pb y genera 2 miRNAs maduros: miR-205-5p y miR-205-3p. Diversos estudios han demostrado que la expresión de miR-205 se encuentra a la baja en cáncer de mama (Iorio et al., 2009; Wu et al., 2009) y el cual es capaz de interferir con la vía de supervivencia mediada por familia de receptores HER, directamente con el receptor HER3 (Iorio et al., 2009). Además, otros estudios indican que miR-205 es un regulador negativo de la Transición Epitelio-Mesenquima (EMT), al ser sus blancos los factores de transcripción ZEB1 y ZEB2, y que la expresión de mir-205 es perdida en líneas celulares de cáncer de mama mesenquimales (Gregory et al., 2008).

En cuanto a la expresión en los subgrupos diferentes de tumores de mama de mir-205, Sempere y sus colegas describieron que su expresión está restringida al compartimento celular mioepitelial/basal de conductos normales mamarios y lóbulos y reducido o eliminado en los tumores correspondientes (Sempere et al., 2007), posteriormente otro estudio divulgó que la expresión de miR-205 se encuentra a la baja en TNBCS primarios comparado a su contraparte normal. Todos estos estudios validan que mir205 en cáncer de mama podría estar actuando como un supresor tumoral. Más aún, un estudio llevado a cabo por C. Piotvan para identificar

las vías oncosupresoras dependientes de miR-205 en el cáncer de mama triple negativo encontró que mir-205 interactúa con E2F1, un factor transcripcional que juega un papel crucial en la regulación de la progresión del ciclo celular (Polager and Ginsberg, 2009), y con LAMC1, un componente de la matriz extracelular, implicada en una amplia variedad de procesos biológicos incluyendo adherencia celular, diferenciación, migración y metástasis (Martin and Timpl, 1987). Finalmente, ellos analizaron la unidad transcripcional de miR-205 para clarificar los mecanismos moleculares que afectan su expresión, y demuestran que el supresor tumoral p53, cuyas mutaciones y pérdida representan uno de los cambios genéticos más comunes en neoplasias malignas humanas incluyendo cáncer de mama, es capaz de modular positivamente los niveles de expresión de miR-205 en diferentes modelos celulares, y que este efecto es ejercido por la unión directa del factor de transcripción sobre secuencias reguladoras río arriba de la región codificante del microRNA. En cuanto a la regulación de su promotor, se encontró que el gen del lncRNA, ARN antisentido de transcripción de HOX (HOTAIR) puede controlar la expresión del gen de mir-205 al modificar las marcas de histonas en la región de su promotor H3K4me3 and H3K27me3 (I. Lee et al., 2009).

RNAs largos no codificantes (lncRNAs)

Los ARNs largos no codificantes (lncRNAs) se encuentran en una gran diversidad de especies, incluyendo animales (Brown et al., 1992; Clemson, McNeil, Willard, & Lawrence, 1996), plantas (Swiezewski, Liu, Magusin, & Dean, 2009), levaduras (Houseley, Rubbi, Grunstein, Tollervey, & Vogelauer, 2008), procariontes (Bernstein, Zopf, Freymann, & Walter, 1993), e incluso virus (Reeves, Davies, McSharry, Wilkinson, & Sinclair, 2006). Sin embargo, este tipo de ARNs son expresados en bajos niveles en comparación con mRNAs (Cabili et al., 2011; Derrien et al., 2012; van Bakel, Nislow, Blencowe, & Hughes, 2010). Aunque estos antecedentes en un inicio puedan cuestionan el rol biológico que juegan los lncRNAs existe suficiente evidencia que demuestra que están involucrados en procesos biológicos importantes (Clark & Mattick, 2011; Kapusta et al., 2013; Ponting, Oliver, & Reik, 2009): incluyendo la transcripción (Martianov, Ramadass, Serra Barros, Chow, & Akoulitchev, 2007; Rinn et al., 2007), empalme alternativo (Rintala-Maki & Sutherland, 2009; Tripathi et al., 2010), la traducción (Beltran et al., 2008; Muddashetty et al., 2002), la integridad de la estructura celular (Kloc et al., 2005; Sunwoo, Dinger, & Wilusz, 2009), impronta génica (Brockdorff et al., 1991; Rachmilewitz et al., 1992), el ciclo de celular (Ginger et al., 2006; Mourtada-Maarabouni, Hedge, Kirkham, Farzaneh, & Williams, 2010) y apoptosis (Huarte et al., 2010; Reeves et al., 2006), características pluripotenciales (Sheik Mohamed, Gaughwin, Lim, Robson, & Lipovich, 2010) y reprogramación celular (Loewer et al., 2011), y en la respuesta al choque térmico (Espinoza, Allen, Hieb, Kugel, & Goodrich, 2004; Shamovsky, Ivannikov, Kandel, Gershon, & Nudler, 2006). La sobreexpresión, la deficiencia o la mutación de genes de lncRNAs han sido implicados en numerosas

enfermedades humanas (Clemson et al., 1996) como la progresión de cáncer (Tano & Akimitsu, 2012). Por lo tanto, todas estas funciones y rasgos biológicos de los lncRNAs los hacen objeto de investigación interesante e importante en el progreso de cáncer, como genes reguladores de rutas oncogénicas o supresoras tumorales como p53, MYC y NFκB (Hung et al., 2011).

Actualmente, hay 73,370 lncRNAs de 1,239 organismos según NONCODE V3.0 50 —una base de datos bibliográfica de lncRNAs. De los cuales sólo una pequeña proporción —menos de 404 según lncRNADB 10, una base de datos de anotación de lncRNAs— han sido funcionalmente caracterizados. Esto refleja que las funciones de la mayoría de los lncRNA son desconocidas, aun así, la lista de los lncRNAs funcionalmente caracterizados —como el transcrito específico de inactivación del X (XIST; en la inactivación del cromosoma X), ARN antisentido de transcripción de HOX (HOTAIR, identidad posicional) y el ARN componente la telomerasa (TERC, en la elongación del telómero)— sigue creciendo.

Biogénesis

Los lncRNAs son transcritos por la ARN polimerasa II, y posteriormente sufren modificaciones post-transcripcionales en el extremo 3' —cola de Poli (A)— y empalme alternativo (Mercer et al., 2009). Los lncRNAs albergan los sitios canónicos de señales empalme, sin embargo, tienen menos exones que los mRNAs, lo que los hace generalmente más cortos que estos últimos —en promedio los lncRNA tienen 592 nt comparados a 2453 nt para mRNA (Derrien et 2012 Al-). Los lncRNAs no tienen marcos de lectura abiertos (ORFs)

y si los hay tienen codones de inicio escasos y contexto de los ORF es vago. Es por esto por lo que las funciones biológicas de los lncRNA se atribuyen más a su estructura que y propiedades bioquímicas. Los elementos regulatorios y promotores río arriba de los genes de los lncRNAs están marcados con marcas de histonas (H3K4me2/3, H3K9ac, H3K27ac) y a lo largo de los genes usualmente se encuentran marcados con la modificación H3K36me3, marcas características de regiones genómicas que son transcripcionalmente activas (Guttman et al., 2009; Derrien et al., 2012). Esto implica que los lncRNAs son objeto de regulación epigenética.

Clasificación de los lncRNAs

Por su localización y contexto genómicos

LncRNAs Intergénicos (lincRNAs) y lncRNAs intrónicos

Los lncRNAs son transcritos a partir de diversas regiones genómicas, aquellos que se transcriben a partir de las regiones intergénicas se llaman lncRNAs intergénicos (Figura 5A) y aquellos que se transcriben a partir de los intrones de genes codificantes son llamados lncRNAs intrónicos (Figura 5B). Ambos son probablemente regulados por mecanismos de activación de transcripción diferentes (Cao et al., 2012), pueden tener distintas modificaciones en el extremo 3' y manifestar actividades en diferentes compartimentos celulares (Cheng et al., 2005).

Poco se sabe de la función de los lncRNAs intrónicos, aunque el panorama de los lncRNAs intergénicos ha sido más explorado. Se sabe que los lncRNAs son capaces de operar mediante los mecanismos de regulación transcripcional en cis o en trans, en mecanismos de control traduccional, regulación del empalme alternativo y otros mecanismos de regulación de expresión post-transcripcionales (L. Ma, Bajic, & Zhang, 2013) .

Los lincRNAs contienen marcas de metilación “K4-K36” —trimetilaciones en la histona H3K4 en su extremo 5' y trimetilaciones H3K36 a lo largo del gen— un indicador de la transcripción activa en genes (Khalil et al., 2009; Ørom et al., 2014). Adicionalmente los lincRNAs se han encontrado conservados a través de varias especies de vertebrados (Cabili et al., 2011). Todas estas evidencias presentadas —transcripción activa, conservación evolutiva y la especificidad tejido-específica— son fuertes indicadores de la importancia funcional de los lincRNAs. Aún más, evidencia demuestra que los lincRNAs realizan funciones importantes en múltiples procesos celulares —desde procesos de células embrionarias pluripotentes a proliferación celular y progresión del cáncer (Cao et al., 2012; Guttman et al., 2009; Huarte et al., 2010; Ørom et al., 2014).

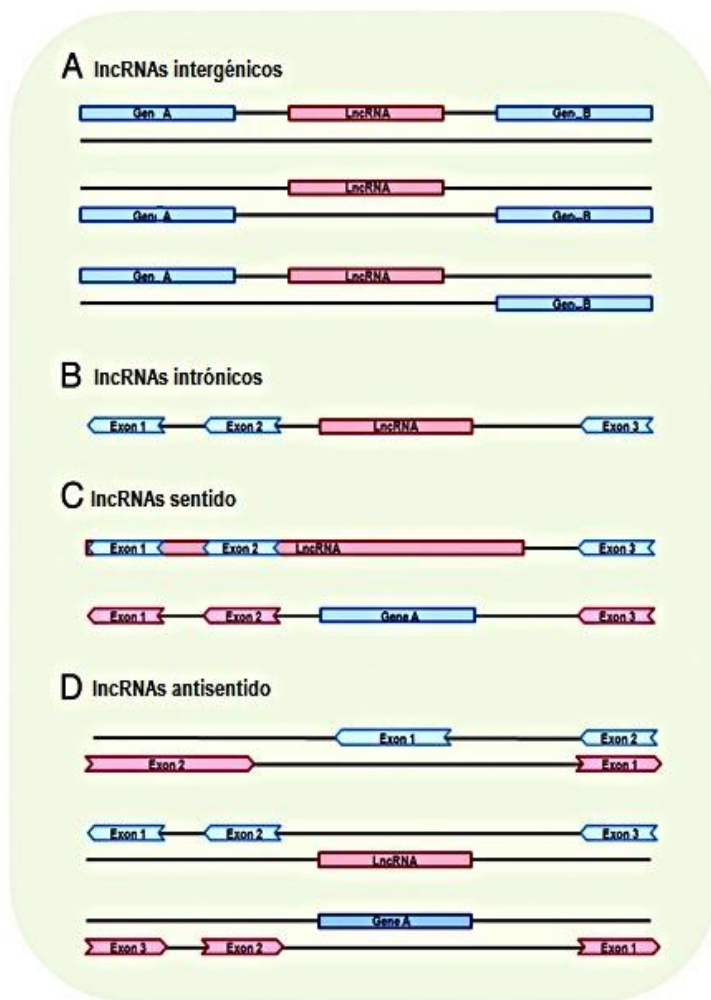


Figura 4 Localización genómica y contexto de los lncRNAs modificada de (L. Ma et al., 2013).

Los genes codificantes para proteínas (GCP) y sus exones están representados en color azul, mientras que los lncRNAs y sus exones están representados en color rojo. Los paneles de los lncRNAs están basados en la localización anotada de GENCODE 4. (A) lncRNA intergénico: se transcribe generalmente de ambas hebras. (B) lncRNA intrónico: Se transcribe casi en su totalidad de los intrones de GCP. (C) lncRNA sentido: Se transcribe de la hebra sentido de un gen codificante y contiene exones de genes GCP, traslapa con GCP o cubre la secuencia entera de un GCP mediante un intrón. (D) lncRNA antisentido: se transcribe de la hebra antisentido de un GCP, traslape con las regiones intrónicas y exónicas y/o cubre la región completa mediante un intrón.

lncRNAs sentido y antisentido

Los lncRNAs sentido se transcriben a partir de la hebra positiva de los genes codificantes para proteínas (GCP). Los lncRNAs sentido pueden traslaparse o cubrir una secuencia entera de los GCP (Figura 5C). Los lncRNAs antisentido, son transcritos de la hebra negativa de los GDP.

De acuerdo con GENCODE —una base de datos transcriptómica manualmente curada— los escenarios en los que se puede dar un lncRNA antisentido son: (1) el transcrito a partir de hebra antisentido del GCP traslapa con un exón del gen sentido a través de uno de los exones del lncRNA; (2) el transcrito del intrón de un gen sentido no tenga traslape exón-exón con el mismo gen sentido; (3) el transcrito cubra la secuencia entera a través de un intrón (Figura 5D).

Por sus efectos ejercidos en las secuencias de ADN

Se ha encontrado que los lncRNAs predominantemente se encuentran en el núcleo y asociados a la cromatina (Derrien et al., 2012), lo que sugiere que los lncRNAs pueden tener un impacto significativo en la secuencia de ADN. También una gran proporción de lncRNAs están involucrados en la regulación transcripcional (~42% de las 182 entradas de acuerdo con lncRNAdb). Basado en sus modos de acción sobre la secuencia de ADN los lncRNAs: se pueden clasificar como cis-lncRNAs (cis-acting lncRNAs) que regulan la expresión de genes próximos a su entorno y trans-RNAs (trans-acting lncRNAs) que regulan la expresión de genes distantes (**Figura 6**) (L. Ma et al., 2013).

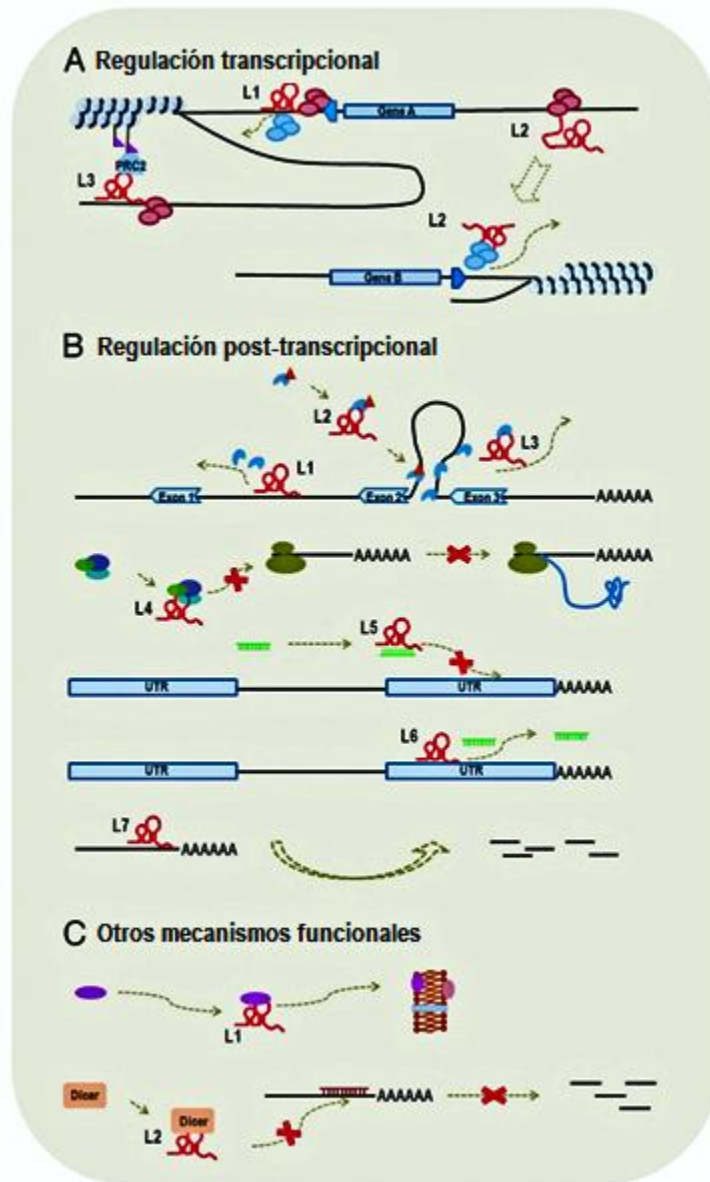


Figura 5 Mecanismos funcionales de los lncRNAs modificada de (L. Ma et al., El lncRNA está representado con la letra “L” y un número adjunto. (A) Regulación transcripcional. Se enlistaron los ejemplos de cis-lncRNA (L1 y L3) y trans-lncRNA (L2). L1 se transcribe de la región promotora del gen A y se encuentra unido al promotor del gen A. L3 modifica las proteínas acopladas a la cromatina por medio del reclutamiento del complejo PRC2. L2 influencia la transcripción del gen B desde una región distante por medio de la

interacción con factores de transcripción o por la ARN polimerasa. Por ende, L1 y L2 funcionan mediante la interferencia transcripcional, donde L3 lo hace mediante la modificación de la cromatina. (B) Regulación post-transcripcional: L1, L2 y L3 influyen en el proceso de empalme. Específicamente L1 se une a la región intrónica del factor de empalme, L2 modula la reserva de factores de empalme modificados —por ejemplo, fosforilados— y L3 se une a factores de empalme para bloquear la formación del complejo de empalme. L4

interactúa con factores de traducción para inhibir la propia traducción. L5 y L6 son dos ejemplos de ceRNAs, los cuales interactúan directamente o indirectamente con miRNAs. L5 se une a un miRNA y, por tanto inhibe la unión del miRNA a la región 3'UTR del mRNA blanco. L6 se une a la región 3'UTR del mRNA blanco, el cual bloquea la unión del miRNA al gen objetivo. L7 sirve como un inhibidor natural antisentido para promover la degradación de mRNA. (C) Otros mecanismos funcionales, L1 se encuentra involucrado en el transporte de proteínas y L2 se une a Dicer para influenciar la interferencia de ARN.

Tabla 1. Mecanismos funcionales y localización genómica de algunos lncRNAs (modificada de (M. F. Lin, Jungreis, & Kellis, 2011))

| <i>Función</i> | <i>Gen lncRNA</i> | |
|-----------------------------------|--------------------|---|
| <i>Regulación transcripcional</i> | Cis | Air ^d , ANRIL ^d , transcritos de Beta-globulina ^b , CAR Intergénico 10 ^d , COLDAIR ^e , COOLAIR ^d , Evf2 ^d , H19 ^b , HOTAIRM1 ^b , Hoxa11as ^d , Kcnq1ot1 ^d , Xist ^b |
| | Trans | 7SK ^b , B2 SINE RNA ^b , GAS5 ^b , HOTAIR ^d , |
| | No descrito | HOXA3as ^d , HOXA6as ^d |
| <i>Regulación traduccional</i> | | BC1 ^b , Gadd7, Zeb2NAT ^d |

Regulación del corte y empalme

Malat1^b, Zeb2NAT^d

Regulación post-transcripcional

Dio3os^b, antisentido E2F4^d, Gadd7, H19^b

Otros mecanismos funcionales

ENOD40^c, RNAs EBER1 y EBER2, TERRA, TERC^b, VAI y RNAs VAIL, VegT RNA^c.

Cis-lncRNAs

Es posible que los mecanismos de los *cis*-lncRNAs sean a través de la interferencia transcripcional o por medio de modificadores transcripcionales (Blume, Meng, Shrestha, Snyder, & Emanuel, 2003; Ponting et al., 2009) —*Xist*, X inactivación del transcrito específico— Los lncRNAs pueden influenciar la actividad transcripcional de genes blanco por medio de la unión a sus promotores para así bloquear la formación del complejo de pre-inicio (PIC) (Martianov et al., 2007) o mediante la interacción con factores de transcripción (Martens, Laprade, & Winston, 2004; Martianov et al., 2007; Thompson & Parker, 2007) (Figura 6A). Estos *cis*-lncRNAs usualmente son transcritos de las regiones promotoras. Por ejemplo, los transcritos de la región promotora DHFR —aproximadamente de 0.8–7.3 kb lncRNAs— pueden formar una estructura de triple hélice con el promotor de DHFR (Blume et al., 2003; Martianov et al., 2007) e interactuar eficientemente con TFIIB para disociar al complejo PIC (Martianov et al., 2007).

De igual manera, la transcripción de lncRNAs de regiones promotoras puede inducir modificaciones en la cromatina, y por ende activar la transcripción de genes río bajo —los ARNs del promotor de *fbp1* (Hirota et al., 2008). *Cis*-lncRNAs que funcionan por medio de modificaciones en la cromatina frecuentemente reclutan los complejos de modificación de la cromatina —PRC (complejo polycomb de represión).

Trans-lncRNAs

Los lncRNAs también son capaces de actuar en modo *trans* sobre un loci blanco distante. Por ejemplo, HOTAIR —un lncRNA de ~2.2 kb— es transcrito de la región HOXC (clúster homeobox C) con locus en el cromosoma (Clemson et al., 1996) y es transportado por la proteína Suz12 para regular la región homóloga HOXD (clúster homeobox D) en el locus del cromosoma 2 (Chu, Zhong, Artandi, & Chang, 2012; Rinn et al., 2007; M. Tsai et al., 2010).

Por sus mecanismos de funcionamiento

Regulación transcripcional

Existen diversos lncRNAs que regulan la expresión por medio de interferencia transcripcional —transcritos en la región promotora de DHFR, ARNs de SRG1, ARN B2 SINE— y modificadores de la cromatina —los ARNs del promotor de *fbp1* (fructosa-1,6-bifosfatasa-1), XIST, MEG3, GAL10-ncRNA, HOTAIR, HOTTIP and COLDAIR. Por lo tanto, los lncRNAs pueden ser sub-clasificados de acuerdo con su mecanismo de acción: (1) interferencia transcripcional y (2) modificadores de la cromatina (Figura 6A). Además, existe otra categoría

(3) por el efecto de su regulación: son lncRNAs que se transcriben de regiones amplificadoras llamados eRNAs (enhancer RNAs) y que regulan de manera positiva la transcripción de genes —ncRNA-a1 (Ørom et al., 2014), el ARN *Evf-2* (Feng et al., 2006).

Regulación post-transcripcional

Existen dos mecanismos post- transcripcionales en los cuales los lncRNAs participan — regulación del empalme y control traduccional (Figura 6B). Los lncRNAs que influyen en el empalme de mRNAs lo hacen a través de la unión (Tsuiji et al., 2011) y modulación (Tripathi et al., 2010) de los factores de empalme o a través de la hibridación directa con secuencias de mRNA para bloquear el empalme (Beltran et al., 2008; Rintala-Maki & Sutherland, 2009). Malat1 (transcrito 1 de metástasis asociado a adenocarcinoma de pulmón) —un lncRNA de ~7 kb—se puede unir al factor de empalme SR y regular su distribución nuclear (Tripathi et al., 2010), así mismo Malat1 puede modular los reservorios de SR y por ende, influenciar el empalme alternativo de mRNAs (Tripathi et al., 2010).

Adicionalmente, existen otros mecanismos a nivel de regulación post-transcripcional. LUST (transcrito LUCA 15-específico) —un lncRNA de ~1.4–2.4kb— es el transcrito antisentido de RBM5 (ARN de unión al motivo de la proteína 5) y se cree que regula la expresión de las isoformas de la proteína RBM5 a través de la hibridación a su secuencia (Rintala-Maki & Sutherland, 2009).

Los lncRNAs que participan en el control traduccional lo hacen por medio de la unión a factores traduccionales (D. Lin, Pestova, Hellen, & Tiedge, 2008; Muddashetty et al., 2002) o directamente al ribosoma (Crawford, Schools, Salmon, & Davies, 1996; Parrott et al., 2011).

Existen dos lncRNAs caracterizados, BC1 —ARN 1 citoplásmico cerebral— y BC200 —un

ARN citoplásmico de 200nt—, los cuales se pueden unir al eIF4A —factor de iniciación de la traducción 4A en eucariontes—, PABP —proteína de unión a poli(A)— y otros factores para reprimir la iniciación de la traducción mediante el bloqueo del ensamblado del complejo (D. Lin et al., 2008; Muddashetty et al., 2002). nAR — ARNs pequeños asociados a NF90— un lncRNA citoplasmático que se puede unir al ribosoma influenciando de esta manera la traducción de mRNAs (Parrott et al., 2011).

Los lncRNAs pueden funcionar como inhibidores de antisentido para promover la degradación de mRNA (Gong & Maquat, 2011; Pagano et al., 2007) (Figura 6B). Se ha encontrado que 21A —un lncRNA de 300bp y que su secuencia muestra alta homología a los intrones de la CENP-F (proteína F del centrómero)— es capaz de disminuir los niveles de expresión —tanto a nivel de mRNA como a nivel de proteína— al actuar como un inhibidor de antisentido (Pagano et al., 2007).

Por otro lado, existen lncRNAs que son capaces de interactuar tanto directa como indirectamente con miRNAs para estabilizar mRNAs blanco (Cesana et al., 2011; Salmena, Poliseno, Tay, Kats, & Pandolfi, 2011; Sumazin et al., 2011; Tay et al., 2012) (Fig. 2B). Estos lncRNAs son conocidos como ceRNAs (ARNs endógenos competentes). Por ejemplo, lincMD1 (un largo intergénico ncRNA asociado a la diferenciación muscular), actúa como esponja de los miR-133 y miR-135 con el fin de regular la expresión de dos factores de transcripción: MAML1 y MEF2C, los cuales activan la expresión de genes musculo-específicos (Cesana et al., 2011). Además, muchos lncRNAs antisentido son capaces de unirse a mRNAs para enmascarar los sitios de unión a miRNAs. BACE1AS —ARN antisentido BACE— es un lncRNA de ~2 kb lncRNA y en Alzheimer se encuentra activado en la forma de dúplex de

ARN con el mRNA de BACE1 — β -secretasa 1 (Faghihi, Modarresi, et al., 2010) —y el cual enmascara el sitio de unión de miR-485-5p y por lo tanto evita la represión traduccional del mRNA de BACE1 por miRNA(Faghihi, Zhang, et al., 2010).

Se han encontrado algunos lncRNAs que contienen miRNAs en su locus (Augoff, McCue, Plow, & Sossey-Alaoui, 2012; Cai & Cullen, 2007). H19 —un lncRNA de ~2.3 kb— contiene a mir-675 en un exón y sirve como precursor de este miRNA además de ser regulador de su actividad transcripcional(Cai & Cullen, 2007). Adicionalmente, lncRNA LOC554202 contiene en uno de sus intrones a mir-31, y ambos —mir-31 como el gen anfitrión— se encuentran expresados a la baja en pacientes con cáncer de mama triple negativo (Cai & Cullen, 2007).

Por su modo de acción

De acuerdo a su mecanismo de acción, los lncRNAs también pueden ser clasificados basado en su modo de acción (Wang & Chang, 2011), tal como, (1) señalización: muestran expresión tipo celular-específica y responden a diversos estímulos como Xist (Maenner et al., 2010), COLDAIR (Heo & Sung, 2011); (2) Señuelo: Se unen/titulan a un blanco proteico —como los transcritos río arriba de DHFR (Martianov et al., 2007), PANDA (Hung et al., 2011) — (3) guía: Se unen a proteínas y dirigen la ubicación del complejo ribonucleoproteico a blancos específicos —Xist (Blume et al., 2003), HOTAIR (Rinn et al., 2007)— (4) andamio: Sirven como plataformas para acercar múltiples proteínas y formas complejos ribonucleicos HOTAIR (Rinn et al., 2007; M. Tsai et al., 2010; Walter & Blobel, 1982; Zwieb, Nues, Rosenblad, Brown, & Samuelsson, 2005).

Cáncer y oncolncRNAs

La expresión desregulada de los primeros lncRNAs descubiertos en tejidos tumorales —h19 (Y. Zhang et al., 1993), MALAT1 (Luo et al., 2006) y PCA3 (Kok et al., 2002)— se realizó antes de que existieran las tecnologías de secuenciación de nueva generación. La secuenciación de ARN (RNA-seq) permitió la evaluación a gran escala de lncRNAs diferencialmente expresados en tumores en comparación con los tejidos normales y demostró que una cantidad significativa de lncRNAs se encontraban alterados (Bartoniccek, Maag, & Dinger, 2016). Las mutaciones somáticas en regiones regulatorias (Melton, Reuter, Spacek, & Snyder, 2016), las alteraciones epigenéticas (Yen et al., 2015) o las alteraciones somáticas en el número de copias (Beroukhim et al., 2010) son capaces de afectar la transcripción de los lncRNAs. Sin embargo, se cree que principal factor para la desregulación de la expresión de lncRNAs en tumores se debe a oncogenes que actúan como factores de transcripción. Dos genes que cabe resaltar actúan como factores de transcripción de diversos lncRNAs son Myc y p53 (Grossi, Sánchez, & Huarte, 2016; Hart, Roberts, Weinberg, Morris, & Vogt, 2014; Younger, Kenzelmann-Broz, Jung, Attardi, & Rinn, 2015).

En esta sección se presentarán algunos lncRNAs caracterizados e implicados en cáncer (FIGURA 7). HOTAIR es un lncRNA cuya sobreexpresión promueve la metástasis de células tumorales del seno mediante el silenciamiento epigenético de genes importantes en el clúster de HOXD, entre otros (Gupta et al., 2010). Estudios subsecuentes muestran que la desregulación de HOTAIR se asocia a la progresión del cáncer en 26 tipos tumorales (Block, Rawat, Brosgart, & Francisco, 2017) y es capaz de predecir la respuesta entre dos quimioterapias en pacientes con cáncer de ovario (Teschendorff et al., 2015); la

sobreexpresión de ANRIL —oficialmente conocido como ARN antisentido 1 CDKN2B (CDKN2BAS1)— silencia el locus supresor tumoral que consta de los genes INK4b, ARF e INK4a —oficialmente conocido como CDKN2B— y además se encuentra asociado a bajo pronóstico en los cánceres de próstata y gástrico (Cunnington, Koref, Mayosi, Burn, & Keavney, 2010; Yap et al., 2011; Yu et al., 2009; E. Zhang et al., 2014). PCA3 conocido como antígeno prostático, es un biomarcador en el cáncer de próstata (Bussemakers et al., 1999) aprobado por la FDA para el diagnóstico de este cáncer, el primer examen basado en un lncRNA. De manera similar, en análisis de secreciones gástricas de pacientes con cáncer gástrico se identificó al lncRNA AA174084, un biomarcador capaz de diferenciar entre el cáncer y desórdenes en el epitelio gástrico (Huarte, 2015; Schey, Luther, & Rose, 2016; Schmitt, Chang, Sloan, & Cancer, 2017; Shao et al., 2014).

En cáncer de mama se han encontrado lncRNAs desregulados en tumores, a continuación, se mencionarán algunos de estos genes caracterizados. ANRIL un lncRNA antisentido —que se encuentra también involucrado en cáncer de esófago, gástrico, en leucemia, melanoma, entre otros— regula a sus vecinos supresores tumorales CDKN2A/B por mecanismos epigenéticos, con lo que es capaz de intervenir en la proliferación celular (Wade, Li, & M. Wahl, 2013); ATB es un lncRNA intergénico sobreexpresado en cáncer de mama —involucrado en carcinoma hepatocelular— juega un papel importante en la transición epitelio-mesénquima (Shi et al., 2015). GAS5 es un lncRNA subexpresado en cáncer de mama —asociado a cáncer de hígado, linfoma y próstata— induce el arresto del ciclo celular y la apoptosis (Mourtada-Maarabouni, Pickard, Hedge, Farzaneh, &

Williams, 2009); H19 es un lncRNA cuya sobreexpresión aumenta la capacidad de las células cancerígenas en formar colonias y su crecimiento independiente de anclaje (Sun et al., 2015).

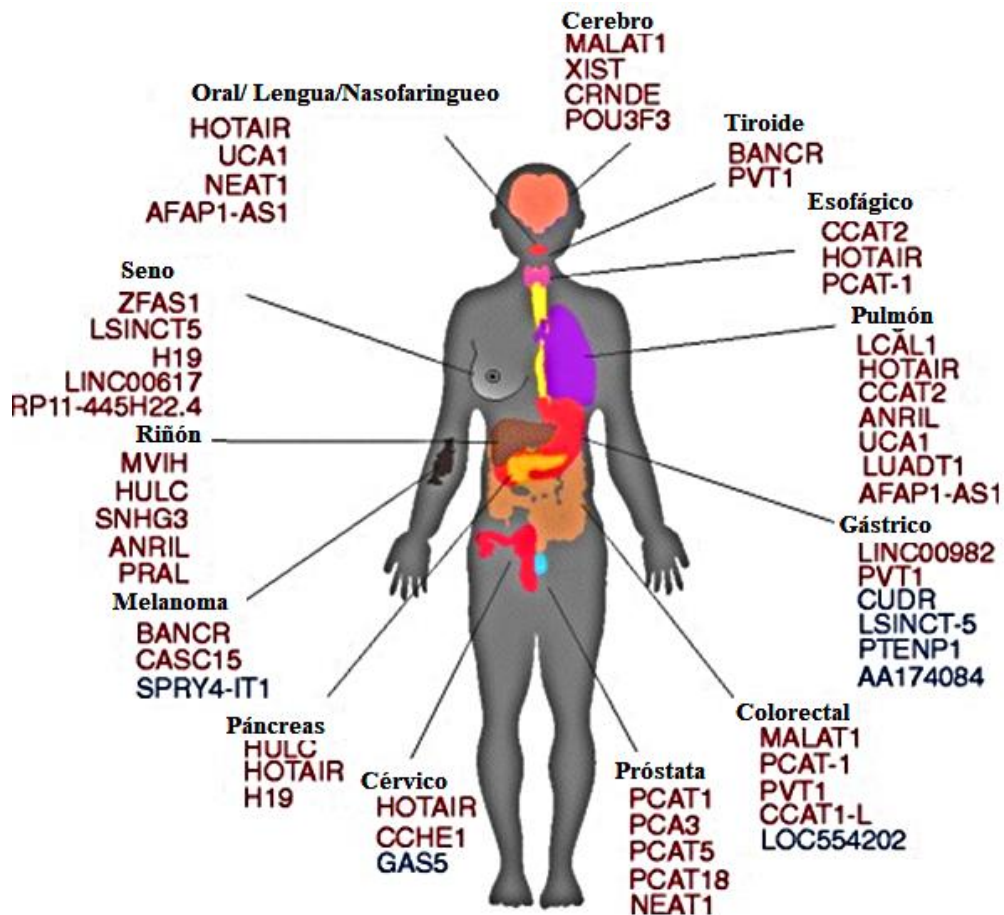


Figura 6 lncRNAs implicados en diversos tipos tumorales. (Azules: subexpresados; rojo: sobreexpresados) modificada de (Bartonicek et al., 2016).

MALAT1 se encuentra sobreexpresado en cáncer de mama —asociado a múltiples tipos tumorales— juega un papel crítico en el proceso de empalme alternativo de mRNA (Gutschner, Hämmerle, & Diederichs, 2013). SOX2OT se encuentra involucrado en la inducción y/o mantenimiento de la expresión de SOX2 (Zhao, Luo, & Jiao, 2014). ZFAS1 desregulado a la baja en cáncer de mama, incrementa la proliferación y la actividad metabólica (Askarian-Amiri et al., 2011). Además, existen diversos lncRNAs capaces de predecir la

supervivencia de los pacientes. Por ejemplo, un set de genes de lncRNAs —U79277, AK024118, BC040204 and AK000974— seleccionados por medio del minado de bases de datos, han mostrado ser buenos marcadores de supervivencia en cáncer de mama. Esta firma de expresión de genes de lncRNAs es independiente de la edad e incluso del subtipo de cáncer de mama (Sørensen et al., 2015).

De acuerdo con todo lo descrito los lncRNAs tienen un rol regulador importante en la biología del cáncer. Sin embargo, un porcentaje muy pequeño de estas moléculas han sido caracterizadas funcionalmente, de los 15,778 reportados por Gencode, y los hallazgos en sus funciones específicas demuestran la enorme complejidad que subyace en la regulación génica. A medida que se profundiza en el conocimiento de los lncRNAs oncogénicos, se contribuye a la identificación de posibles marcadores biológicos que jueguen un papel importante en el diagnóstico y/o sirvan como blancos terapéuticos.

Laboratorio de Genómica del cáncer del INMEGEN

En México el Laboratorio de Genómica del Cáncer del INMEGEN, se enfoca a la identificación de alteraciones somáticas en el genoma y transcriptoma de tumores humanos, con particular énfasis en tumores de cáncer de mama. Actualmente se están analizando los perfiles de alteraciones en el número de copias de DNA, perfiles globales de expresión de

RNA mensajeros y los perfiles de expresión de microRNAs y RNAs largos no codificantes en tumores y tejidos mamarios normales. Recientemente, a través del uso de secuenciación de nueva generación, se llevó a cabo el análisis de secuencia de genoma y exoma completo, así como la identificación de mutaciones en regiones no codificantes del genoma de tumores de mama. Estos análisis han resultado en la identificación de diversas alteraciones genómicas cuyo papel no había sido previamente descrito en este tumor. Actualmente en el laboratorio se está realizando la caracterización funcional de lncRNAs alterados en el subtipo de cáncer de mama triple negativo —artículo en revisión— y asociados a supervivencia de las pacientes en cáncer de mama y en leucemia linfoblástica infantil aguda (Beltran-Anaya, Cedro-Tanda, Hidalgo-Miranda, & Romero-Cordoba, 2016; Flores-Pérez et al., 2016; Rheinbay et al., 2017; Romero-Cordoba et al., 2012).

Objetivo General

Identificar a través del minado de bases de datos genómicos, lncRNAs asociados a la supervivencia de pacientes mexicanas con cáncer de mama y validar la expresión diferencial de uno de ellos mediante PCR en tiempo real en una muestra independiente de tumores.

Objetivos particulares

- Identificación de RNAs largos no codificantes (lncRNAs) cuya expresión se encuentre alterada en cáncer de mama mediante el minado de bases de datos genómicos.
- Selección de cuatro lncRNAs cuya expresión este alterada en cáncer de mama y asociada a sobrevida global de pacientes.
- Validación de la expresión de MIR205HG en una cohorte independiente de muestras (N=16 y T=16).
- Análisis funcional *in silico* de MIR205HG seleccionado.

Hipótesis

Se podrá validar la expresión de un RNA largo no codificante, seleccionado por el minado de bases de datos, por medio de la reacción en cadena de la polimerasa cuantitativa (q-PCR) en una cohorte independiente de tumores de cáncer de mama de pacientes mexicanas.

Material y métodos.

Materiales.

Instrumentación.

Todos los equipos empleados en este trabajo se listan en la Tabla 1.

Tabla 2 . Lista de equipos

| Equipo | Nombre | Fabricante |
|---------------------------------------|--|--------------------------|
| Agitador tipo vórtex | Maxi mix II | ThermoLyne |
| Balanza analítica | ABS-ABJ | Femto |
| Barra de agitación magnética | Barra de agitación magnética octagonal | Hach |
| Cámara de electroforesis y accesorios | Sub-Cell® Model 192 Cell | Bio-Rad |
| Centrífuga | Multifuge 3S Plus | Heraeus |
| Centrífuga | Spectrafuge C1301 Mini Centrífuga | Labnet |
| Congelador de laboratorio -20°C | REL2304 | REVCO |
| Congelador de laboratorio -80°C | ULT2586-9-A34 | REVCO |
| Equipo de PCR tiempo real | QuantStudio 7 Flex System | Applied Biosystems |
| Espectrofotómetro | NanoDrop ND-1000 | Thermo Fisher Scientific |
| Fuente de poder | Fuente de poder PowerPac™ Basic | Bio-Rad |
| Máquina de hielo | MHC-230/506 | Torrey |
| Microondas | MS0741CW | LG |
| Micropipetas | Pipeta Rainin Classic | Metler Toledo |
| Micropipetas | Pipetman G | Gilson |
| Potenciómetro | UB-10 | Denver Instrument |
| Refrigerador de laboratorio 4°C | LABO-468 | Kirsch |
| Sistema de agua ultrapura | Milli-Q® IQ 7000 | Merck |
| Sistema de PCR tiempo real | QuantStudio™ 7 Flex | Thermo Fisher Scientific |
| Termociclador | GeneAmp PCR System 9700 Base Module | Applied Biosystems |
| Transiluminador UV | Gel Doc XR+ System | Bio-Rad |

Insumos.

En la Tabla 2 se muestran los artículos empleados cotidianamente.

Tabla 3. Lista de insumos

| Producto | Fabricante |
|--------------------------------|-------------------|
| Charola de plástico para pesar | Prolab |

| | |
|---|--------------------------|
| Espátula analítica de acero inoxidable | Zelian |
| Frasco de vidrio con tapa | Kimax |
| Guantes de nitrilo desechables | Supreno |
| Microtubos | Thermo Fisher Scientific |
| Película adhesiva óptica MicroAmp | Thermo Fisher Scientific |
| Placas ópticas de 96 pozos para q-PCR MicroAmp® | Thermo Fisher Scientific |
| Puntas para micropipetas con filtro | Axygen |
| Soporte para microtubos | Detalab |
| Toallas de papel | Sanitas |
| Toallitas limpiadoras Kimwipes® | Kimberly-Clark |
| Tubos falcon 15 cónicos | Probiotek |

Kits de biología molecular.

Todos los kits que se emplearon en la tesis se enlistan en la Tabla 3.

Tabla 4. Lista de kits de biología molecular

| Kit | Fabricante |
|---|--------------------------|
| *AllPrep DNA-RNA® | Qiagen |
| High-Capacity cDNA Reverse Transcription Kit® | Thermo Fisher Scientific |
| PureLink® RNA Mini Kit | Thermo Fisher Scientific |
| PowerUp® SYBR® Green Master Mix | Applied Biosystems |

* Dicho kit no fue empleado por el sustentante, sin embargo, en el laboratorio se extrajo el RNA de las muestras y de las líneas celulares MCF-7 y MDA-468 por medio de este.

Reactivos.

Los reactivos que se emplearon para la elaboración de los experimentos se encuentran en la Tabla 4.

Tabla 5. Lista de reactivos químicos

| Reactivo | Fabricante |
|--|--------------------------|
| Agarosa | Sigma-Aldrich |
| Azul de bromofenol/Xileno cianol | Científica Senna |
| Buffer de carga 6X | Promega |
| Etanol | Probiotek |
| Glicerol | Merck |
| H ₂ O libre de nucleasas | Thermo Fisher Scientific |
| H ₂ O Milli-Q® | Merck |
| Marcador de peso molecular DNA Ladder Generuler 50pb | Promega |
| SYBR® gold | Thermo Fisher Scientific |
| Tris/Boratos/EDTA (TBE) buffer (10x) en polvo | Thermo Fisher Scientific |
| TRIzol™ | Thermo Fisher Scientific |

Amortiguadores y disoluciones.

Los buffers y disoluciones empleadas se enlistan en la Tabla 5.

Tabla 6. Lista de buffers y disoluciones

| Disolución | Composición |
|--------------------|---|
| Buffer TBE 10X | 890 mM Tris, 890 mM boric acid, 20 mM EDTA. |
| Buffer TBE 1X | Dilución (1:10) del Buffer TBE 10X; 89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.1-8.5. |
| Buffer de carga 6X | 10 mM Tris-HCl (pH 7.6), 60% (v/v) glicerol, 0.03% (m/v) azul de bromofenol, 0.03% xileno cianol FF y 60 mM EDTA, pH~7.6. |

*Todos los buffers se prepararon con agua Milli-Q®.

q-PCR cebadores

Los cebadores empleados se obtuvieron de Integrated DNA Technologies y de Invitrogen.

Los cebadores y secuencias correspondientes se encuentran en la Tabla 6.

Tabla 7. Lista de cebadores empleados en la q-PCR

| Nombre | Secuencia del oligo (5'→3') | Fabricante |
|---------------|---------------------------------------|--------------------------------------|
| FGF14-AS2 | FP: TCGTGGCGGATAAAGAGC | Integrated DNA Technologies (IDT) |
| | RP: GGGACTACAAAGGGCAGG | |
| MIR205HG | FP: AGGAGTCATTTCTGTTCCGC | Integrated DNA Technologies (IDT) |
| | RP: TCCAAGATGGGTACTTGAGAGA | |
| LINC01614 | FP: CCAGACCAAGGCTGCTGT | Integrated DNA Technologies (IDT) |
| | RP: CCTAGAAGAGTCATTTGAAGTCATGT | |
| SLC12A5 | FP: GCCCCAGGTGTTTAGGAGAC | Integrated DNA Technologies (IDT) |
| | RP: GGTCCTGAATCTGGCCACT | |
| ACTB | FP: TTCTACAATGAGCTGCGTGTG | Invitrogen |
| | RP: GGGGTGTTGAAGGTCTCAAA | |
| TBP | FP: TTCGGAGAGTTCTGGGATTGTA | Invitrogen |
| | RP: TGGACTGTTCTTCACTCTTGGC | |
| 18S rRNA | FP: ATG GCC GTT CTT AGT TGG TG | Invitrogen |
| | RP: TCC TCG TTC ATG TTT AAT AA | |

Líneas celulares.

Las líneas celulares empleadas en este trabajo se muestran en la Tabla 7.

Tabla 8. Lista de líneas celulares empleadas

| Línea celular | Descripción |
|----------------------|--------------------|
|----------------------|--------------------|

| | |
|------------|--|
| MCF-7 | Línea celular de carcinoma ductal mamario. Se aisló en 1970 de una mujer caucásica de 69 años. MCF-7 es el acrónimo del instituto en Detroit donde se estableció dicha línea, Michigan Cancer Foundation-7 (ahora el Barbara Ann Karmanos Cancer Institute). (Soule, Vazquez, Long, Albert, & Brennan, 1973) |
| MDA-MB-468 | Línea celular de cáncer de mama ductal derivada de un derrame pleural de una paciente negra de 51 años en 1977. (“The essenTials of life science ReseaRch Globally DeliveReD™,” n.d.) |

Software

Los softwars necesarios para cada etapa del proyecto se enlistan en la Tabla 8.

Tabla 9. Software empleado

| Software | Referencias |
|---------------------------------|--|
| AnnoLnc: Annotate human lncRNAs | http://annolnc.cbi.pku.edu.cn (Hou et al., 2016) |
| Beacon Designer™ Free Edition | Premier Biosoft |
| ChemiDoc™ MP Imager | Bio-Rad |
| Clustal Omega | (McWilliam et al., 2013; Sievers et al., 2011) |
| Expression Atlas | https://www.ebi.ac.uk/ |
| mFold | (Zuker, 2003) |
| NanoDrop™ 2000 | Thermo Fisher Scientific |
| Oligoanalyzer | Integrated DNA Technologies (Owczarzy et al., 2008) |
| Primer 3 | (Koressaar & Remm, 2007; Rozen & Skaletsky, 2000; Thornton & Basu, 2011; Untergasser et al., 2012) |
| Primer Blast | (Ye et al., n.d.) |
| QuantStudio™ Software V1.3 | Thermo Fisher Scientific |
| UNAFold | (Markham & Zuker, 2008) |

Bases de datos

Las bases de datos de acceso público empleadas en el proyecto se enlistan en la Tabla 9.

Tabla 10. Bases de datos externas al laboratorio de Genómica del Cáncer

| Nombre | Descripción | Web link | Referencia |
|---|---|---|---|
| Ensembl | Es una base de datos que provee un marco bioinformático para asociar la secuencia genómica a fenómenos biológicos. | http://www.ensembl.org/ | (Aken et al., 2016; Hubbard et al., 2002) |
| GENCODE | Enciclopedia de genes y sus variantes. | http://www.genecode.org/ | (Derrien et al., 2012; Harrow, Frankish, Gonzalez, Tapanari, Diekhans, Kokocinski, Aken, Barrell, Zadissa, Searle, Barnes, Bignell, Boychenko, Hunt, Kay, Mukherjee, Rajan, Despacio-Reyes, Saunders, Steward, Harte, Lin, Howald, Tanzer, Derrien, Chrast, Walters, Balasubramanian, Pei, Tress, Rodriguez, Ezkurdia, Van Baren, et al., 2012) |
| GeneCards v3 | Base de datos que brinda información detallada y actualizada sobre los genes humanos. | www.genecards.org | (Renhan, Chalifa-Caspi, & Prilusky, 1997) |
| Genome Browser | Buscador en línea que ofrece el acceso a la secuencia del genoma de diversos vertebrados e invertebrados. | http://genome.ucsc.edu/ | (Kent et al., 2002) |
| lncRNome | Banco de conocimientos sobre lncRNA en humanos orientado biológicamente. | http://genome.igib.res.in/lncRNome/ | (Bhartiya et al., 2013) |
| Lncipedia | Base de datos para las secuencias y estructuras de transcritos de lncRNA humanos anotados. | http://www.lncipedia.org/ | (Volders et al., 2013, 2015) |
| NCBI: National Center for Biotechnology Information | Banco de conocimientos que almacena y actualiza la información referente a secuencias genómicas además de publicaciones y datos de relevancia biotecnológica. | https://www.ncbi.nlm.nih.gov | Biblioteca Nacional de Medicina de los EUA (Evolution & Cancer, 2012) |
| PubMed | Base de datos desarrollada y dependiente del NCBI que contiene más de 22 millones de citas y resúmenes de literatura biomédica. | https://www.ncbi.nlm.nih.gov/pubmed/ | |
| Tanric: The Atlas of non-coding RNA in Cancer | Es una aplicación web que analiza lncRNAs en un contexto clínico y molecular entre distintos cánceres. | http://ibl.mdanderson.org/tanric/_design/basic/index.html | (J. Li et al., 2015) |
| TCGA: The Cancer Genome Atlas | Es una colaboración que ofrece mapas multidimensionales de cambios genómicos claves en 33 tipos de cánceres. | https://cancergenome.nih.gov/ | (Tomczak, Czerwińska, & Wiznerowicz, 2015) |

| | | | |
|---------------------------------|---|---|-----|
| TCLA: The Cancer LncRNome Atlas | Base de datos de las alteraciones en los lncRNA entre tumores de distintos tejidos. | http://tcla.fcgportal.org/ | TBD |
|---------------------------------|---|---|-----|

A continuación, en la Tabla 10 se encuentra una breve descripción de una base de datos del laboratorio de Genómica del Cáncer.

Tabla 11. Base de datos del laboratorio de Genómica del Cáncer.

| Nombre | Descripción |
|--|---|
| Análisis de expresión en líneas celulares de cáncer de mama. | Base de datos que proporciona información de expresión en múltiples líneas celulares de cáncer de mama por medio de microarreglos de expresión Human Transcriptome Array 2.0 de Affymetrix. |
| Banco de datos de cáncer de mama. | Base de datos que proporciona información sobre las muestras como la pureza y concentración, así como las características clínicas del paciente. |

Métodos.

Minado de bases de datos: identificación de los lncRNAs candidatos

Se utilizaron las bases de datos *TCLA: The Cancer LncRNome Atlas*, *GeneCards v3*, *Ensembl*, *Tanric: The Atlas of non-coding RNA in Cancer* y *Genome Browser* para la selección de cuatro lncRNAs. Finalmente se corroboró que los lncRNAs no hubiesen sido descritos o caracterizados previamente por medio de *NCBI* (Evolution & Cancer, 2012) en su base de datos *PubMed* (Canese & Weis, 2013).

Los criterios de selección de los lncRNAs comprendieron de las siguientes características:

1. Que existiera una expresión desregulada significativa al alza o baja en tumores comparada a la de tejido normal adyacente en cáncer de mama (BRCA) y se compartiera en otros tipos de cánceres.
2. Que se observara que la expresión del lncRNA estuviera asociada a una mayor o menor supervivencia global de las pacientes (Cox <0.05 y $p<0.04$).

TCLA: The Cancer LncRNome Atlas (Manuscript, 2012).

Se seleccionaron de las bases de datos de TCLA aquellos lncRNAs compartidos en BRCA y diversos tumores que se encontraran sobreexpresados (1144) y subexpresados (1109).

GeneCards v3 (Renhan et al., 1997) y **Ensembl** (Aken et al., 2016; Hubbard et al., 2002)

Se emplearon GeneCards v3 y Ensembl para obtener información detallada sobre los lncRNAs, así como sus diversos alias.

Tanric (Li et al., 2015).

De los 1144 lncRNAs sobreexpresados y 1109 lncRNA subexpresados compartidos en BRCA y diversos tumores, se seleccionaron sólo aquellos que se les relacionara con la supervivencia del paciente (Cox <0.05 y de P value < 0.04). Bajo estos criterios se refinó la búsqueda y se encontró asociación con 11 lncRNAs sobreexpresados y 7 lncRNAs subexpresados. De los cuales se eligieron 2 sobreexpresados y 2 subexpresados con mayor correlación a dicha característica clínica.

Genome Browser (Kent et al., 2002).

Genome Browser se empleó para corroborar que toda la información obtenida de cada lncRNA seleccionado correspondiera a la secuencia del ensamblado GRCh37/hg19.

Diseño de cebadores

Búsqueda de la secuencia blanco.

Los cebadores de los lncRNAs fueron diseñados en los exones que compartían el mayor número de isoformas. Se emplearon las bases de datos *Genome Browser* (Kent et al., 2002) y *LNCipedia* (Volders et al., 2015) para obtener las secuencias de los exones comunes entre los distintos transcritos a fin de empalmarlas con Clustal Omega (Sievers et al., 2011) y así obtener la región blanco donde fueran diseñados los cebadores.

Diseño de cebadores.

El diseño de los cebadores se basó en el artículo “Real-Time PCR (qPCR) Primer Design Using Free Online Software” (Thornton & Basu, 2011). Todos los cebadores fueron diseñados con la herramienta *Primer3* (Untergasser et al., 2012). Los objetivos principales con los que se diseñaron los cebadores fueron: que el producto de amplificación tuviera una longitud de 80 a 120pb, que la temperatura de disociación (T_m) fuera cercana a 60° y que entre las temperaturas de disociación (T_m) del primer forward y el reverse no hubiera más de 2°C de diferencia.

Detección de estructuras secundarias.

La detección de estructuras secundarias tanto entre cebadores (homodímeros y heterodímeros) como del amplicón se realizó mediante Beacon Designer™ Free Edition y Oligoanalyzer (Owczarzy et al., 2008).

Amplificación específica.

Se verificó que la amplificación ocurriera en la secuencia deseada sin falsos positivos por medio de Primer Blast (Ye et al., n.d.).

Cebadores de genes endógenos.

La secuencia de los cebadores para los genes endógenos ACTB (Lyng, Laenkholm, Pallisgaard, & Ditzel, 2008; Rho et al., 2010; Tricarico et al., 2002), TBP (Aigrain, Gu, & Quail, 2016; Muscarella et al., 2010) y 18S rRNA (Salameh et al., 2016; Jia Wang et al., 2016) se obtuvo de los artículos citados y se muestran en la Tabla 11.

Tabla 12. Lista de cebadores empleados en la q-PCR con sus características.

| Símbolo | Objetivo | Número de acceso GenBank | Sitio de unión | Secuencia del oligo (5'→3') | Tm (°C) | Longitud del amplicón |
|-------------|--|--------------------------|----------------|--|---------|-----------------------|
| ACTB | <i>Homo sapiens</i> β-actina, mRNA | NM_001101 | Exon 3 | FP: TTCTACAATGAGCTGCGTGTG | 63.2 | 211pb |
| | | | Exon 4 | RP: GGGGTGTGAAGGTCTCAAA | 60.1 | |
| TBP | <i>Homo sapiens</i> proteína de unión a TATA mRNA | NM_003194 | Exon 4 | FP: TTCGGAGAGTTCTGGGATTGTA | 61.0 | 137pb |
| | | | Exon 6 | RP: TGGACTGTCTTCACTCTTGGC | 62.0 | |
| 18S rRNA | <i>Homo sapiens</i> proteína de unión a TATA mRNA | NR_003286 | Exon 1 | FP: ATG GCC GTT CTT AGT TGG TG | 55.3 | 309pb |
| | | | Exon 1 | RP: TCC TCG TTC ATG TTT AAT AA | 53.2 | |

*Las Tm que se muestran son las temperaturas de disociación teóricas.

Selección del modelo de estudio in vitro.

Mediante la base de datos *Genome Browser* (Kent et al., 2002) y la base de datos del laboratorio de Genómica del Cáncer, *Análisis de expresión en líneas celulares de cáncer de*

mama se determinaron los transcript clusters correspondientes a los lncRNAs candidatos. Posteriormente se seleccionó el modelo celular en líneas de cáncer de mama donde la expresión de los lncRNAs candidatos fuese mayor, esto se evaluó de acuerdo al estadístico Z-score. Las líneas celulares de cáncer de mama empleadas en este trabajo fueron MCF-7 y MDA-468.

Muestras.

Las muestras de tumores y de tejido adyacente en fresco se obtuvieron del banco de muestras del FUCAM. El criterio de inclusión fueron pacientes diagnosticadas con cáncer de mama y como criterios de exclusión, que presentaran otras enfermedades y estuvieran bajo tratamiento de cáncer de mama. Dichas muestras fueron almacenadas a -80°C .

La selección de las muestras se realizó por medio de la base de datos del laboratorio de Genómica del Cáncer, *Banco de datos de cáncer de mama*. Se seleccionaron las muestras de tumores (16) y de tejido adyacente (16) cuyo RNA tuviera un alto grado de pureza ($A_{260/280} \geq 1.9$), alta concentración y no se hubieran analizado con microarreglos de expresión.

Extracción de RNA de muestras.

La extracción del RNA se llevó a cabo por medio del sistema AllPrep DNA-RNA™ de acuerdo a las instrucciones del fabricante. La extracción por medio de este kit permite la separación eficiente de DNA y RNA. En este apartado sólo se explicará lo concerniente a la extracción de RNA.

La muestra se fracciona y homogeneiza mediante disociación mecánica con un disruptor de tejidos en presencia del buffer de lisis. El buffer de lisis contiene tiocianato de guanidina, el cual inactiva las DNAsas y las RNAsas para mantener intactos al DNA y RNA, así como β -mercaptoetanol (β -ME) el cual facilita la disociación del tejido. Posteriormente el lisado se transfiere a la columna AllPrep DNA que es centrifugada. El RNA se encuentra en la fracción recuperada de la columna. Esta se transfiere a la columna RNeasy spin, no sin previamente agregar etanol al 70% para favorecer el proceso de unión del RNA a la membrana. Se lava la columna de residuos contaminantes y se eluye con agua el RNA.

En el laboratorio ya se había realizado la extracción del RNA de las muestras de tanto tumores como de tejido adyacente.

Extracción de RNA de líneas celulares.

La extracción del RNA de las líneas celulares MDA-468 y MCF-7 se realizó por medio de TRIzol™ y PureLink® RNA Mini Kit de acuerdo a las especificaciones del fabricante. Como primer paso se quita el medio a las células y se lavan con 1X PBS. Se agrega TRIzol™ (1 mL por cada 10 cm²) este reactivo liza a las células debido a que contiene guanidina de tiocianato y fenol. Se deja incubar por 5 minutos. Se agrega 0.2mL de cloroformo por cada 1mL de TRIzol™, se agita por 15 segundos, se deja reposar de 2-3 minutos y cetrigufa. Se traspasa la fase superior incolora a otro tubo y se agrega un volumen igual de etanol al 100%. Se transfieren 700μL una columna de silica donde es lavado y posteriormente eluído, listo para usar.

Electroforesis en gel de agarosa.

Se realizó electroforesis en gel de agarosa (1.5% m/v) del RNA extraído de las líneas celulares con el fin de corroborar su integridad (la relación de la intensidad de la banda 28S rRNA con respecto a la 18S rRNA es del doble) y como un control para realizar posteriormente la transcripción reversa en ADN complementario (cDNA).

En el proceso de estandarización de la qPCR se corrieron los productos de amplificación en una electroforesis en gel de agarosa (2.5% m/v) para corroborar la especificidad de los cebadores y las condiciones seleccionadas. Los tamaños esperados de las bandas se muestran en la Tabla 12.

Tabla 13. Tamaños de banda esperados en electroforesis de agarosa.

| Nombre | Tamaño de banda esperado (pb) |
|---------------|--------------------------------------|
| ACTB | 122 |
| TBP | 227 |
| FGF14-AS2 | 135 |
| Linc01614 | 98 |
| MIR205HG | 83 |
| SLC12A5-AS1 | 138 |

Al buffer de carga se le agregó SYBR Gold y almacenó a 4°C protegido de la luz. Cada muestra empleada en la electroforesis se diluyó en relación 1:6 con buffer de carga 6X. Las muestras se cargaron en geles de agarosa (1.5-2.5% agarosa en buffer TBE 1X). Se empleó como marcador de peso molecular el DNA Ladder Generuler 50pb de igual manera, diluido 1:6 con el buffer de carga 6X de acuerdo con las especificaciones del fabricante. Los geles se corrieron a 110V en TBE 1X. Las bandas de los geles se visualizaron por medio de un Transiluminador UV.

Determinación de la concentración y pureza del RNA.

La concentración y pureza del RNA en disolución acuosa se determinaron por espectrofotometría a 260nm por medio del NanoDrop ND-1000 y su integridad se comprobó mediante electroforesis en gel de agarosa. La pureza de la muestra se determina por medio de la relación que existe entre su absorbancia medida a 260 y 280nm. Se consideran de buena pureza aquellas muestras de RNA que presentan valores aproximados a 2.0.

Síntesis del DNA complementario (cDNA).

Las reacciones se prepararon con 200ng de RNA total. La transcripción reversa del RNA extraído se llevó a cabo por High-Capacity cDNA Reverse Transcription Kit™ de ThermoFisher® de acuerdo con las especificaciones del fabricante. El kit convierte de 0.2 a 2 µg de ARN total en ADN complementario monocatenario en una reacción de 20 µl. Incluye la transcriptasa reversa MultiScribe MuLV, dNTPs random octámeros y oligonucleótidos oligo-dT-16 para generar ADN complementario monocatenario.

El programa empleado en el termociclador para realizar la síntesis del cDNA se muestra en la siguiente Tabla 13.

Tabla 14. Programa del termociclador: cDNA

| Etapa | Temperatura (°C) | Tiempo |
|--------------|-------------------------|---------------|
| Paso 1 | 25 | 10 minutos |
| Paso 2 | 37 | 120 minutos |
| Paso 3 | 85 | 5 minutos |
| Paso 4 | 4 | ∞ |

Reacción en cadena de la polimerasa cuantitativa (q-PCR).

La PCR cuantitativa (q-PCR) se realizó por medio de Applied Biosystems™ PowerUp™ SYBR™ Green Master Mix. Este kit comprende de los componentes observados en la Tabla 14.

Tabla 15. Reactivos para q-PCR

| Compuesto | Características |
|---|---|
| Taq DNA polimerasa Dual-Lock™ | Polimerasa con dos anticuerpos que impiden que se active hasta pasados dos minutos a 95°C de tal manera que no produce amplificación residual. |
| Uracil-DNA-glicosilasa termolábil (UDG) | Impide la contaminación por re-amplificación de productos provenientes de otras PCR, debido a que desamina al uracilo del DNA rico en uracilo, monocatenario o de doble cadena. |
| dUTP/dTTP | Desoxiuridina trifosfato/ desoxitimidina trifosfato. |
| SYBR™ Green | Fluoróforo que emite su señal una vez que se une al surco menor de la doble hélice del DNA. |
| Referencia pasiva ROX™ | Fluoróforo de referencia interna que sirve para normalizar la señal del SYBR Green debido a fluctuaciones en volumen o concentración entre reacciones. |

La composición de cada reacción preparada se muestra en la Tabla 15.

Tabla 16. Relación de reactivos para q-PCR

| Reactivo | Volumen (μL) |
|--------------------------------------|---------------------|
| PowerUp™ SYBR™ Green Master Mix (2X) | 5 |
| H ₂ O libre de nucleasas | 2 |
| Forward primer | 1 |
| Reverse primer | 1 |
| cDNA | 1 |
| Volumen total por reacción | 10 |

Las reacciones de amplificación se realizaron en el sistema de PCR tiempo real QuantStudio™ 7 Flex. Se establecieron dos programas de temperaturas en el software QuantStudio™ V1.3, el primero corresponde a amplificaciones con cebadores con temperaturas de disociación (T_m) menores a 60°C y se muestra en la Tabla 16. El segundo fue un programa para realizar la curva de disociación (Melt curves), Tabla 17. El valor del umbral (C_T) fue el predeterminado por el software.

Tabla 17. Programa del termociclador: qPCR

| Etapa | Temperatura (°C) | Tiempo | Ciclos |
|--------------------------------------|-------------------------|---------------|---------------|
| Activación de la UDG | 50 | 2 minutos | 0 |
| Activación DNA polymerasa Dual-Lock™ | 95 | 2 minutos | 0 |
| Desnaturalización | 95 | 15 segundos | |
| Alineación | 55-60 | 15 segundos | 40 |
| Extensión | 72 | 1 minuto | |

Tabla 18. Programa del termociclador: Curva de disociación

| Etapa | Tasa de cambio de temperatura (°C/s) | Temperatura (°C) | Tiempo (s) |
|------------------------------|---|-------------------------|-------------------|
| Curva de disociación: Paso 1 | 1.6 | 95 | 15 |
| Curva de disociación: Paso 2 | 1.6 | 60 | 60 |
| Curva de disociación: Paso 3 | 0.15 | 95 | 15 |

Todas las reacciones se realizaron por triplicado. En la placa se incluyeron: controles negativos, es decir, sin cDNA de las muestras para el control de la especificidad de los cebadores y posible contaminación del mix.

Resultados.

Identificación de los lncRNAs candidatos.

En años recientes, diversas publicaciones han demostrado el potencial que tienen los lncRNAs en el desarrollo y establecimiento del cáncer. En este trabajo se identificaron lncRNAs con expresión diferencial significativa, de acuerdo con los análisis realizados por las bases de datos, entre tumores de cáncer de mama respecto al tejido normal adyacente, 1144 sobreexpresados y 1109 subexpresados, que además se comparten en diversos tumores como en adenocarcinoma de pulmón, de colon y de próstata, y los carcinomas de células escamosas de cabeza y cuello, de riñón y escamosas de pulmón, entre otros. Algunos de estos lncRNAs están asociados a menor supervivencia en cáncer de mama. Se seleccionaron dos lncRNAs sobreexpresados en cáncer de mama, LINC01614 y SLC12A5-AS1, y dos subexpresados, FGF14-AS2 y MIR205HG. En la Tabla 18 se muestran los lncRNAs candidatos seleccionados mediante el uso de bases de datos.

Tabla 19. Los lncRNAs candidatos y características de selección.

| Expresión | Tipos de cáncer en donde están desregulados | Nombre | Cox | P value |
|------------|---|-------------|-------|---------|
| Sobre-exp. | BRCA, COAD, HNSC, KIRC, LUAD, LUSC | LINC01614 | 0.044 | 0.024 |
| | BRCA; COAD; HNSC; KIRC; LUAD; LUSC | SLC12A5-AS1 | 0.040 | 0.057 |
| Sub-exp. | BRCA, COAD, HNSC, KIRC, LUAD, LUSC y PRAD | FGF14-AS2 | 0.013 | 0.0005 |
| | BRCA, LUSC y PRAD | MIR205HG | 0.016 | 0.008 |

BRCA: Cáncer de mama; COAD: Adenocarcinoma de colon; HNSC: Carcinoma de células escamosas de cabeza y cuello; KIRC: Carcinoma de células claras de riñón; LUAD: Adenocarcinoma de pulmón; LUSC: Carcinoma de células escamosas de pulmón; PRAD: Adenocarcinoma de próstata.

El gen del RNA Largo No Codificante Intergénico 1614 (LINC01614) se ubica en la región 2q35 y consta de 13 transcritos (Volders et al., 2015). El lncRNA se encuentra sobre-expresado en cáncer de mama y en otros 5 tipos de cáncer (adenocarcinoma de colon, carcinoma de células escamosas de cabeza y cuello, carcinoma de células claras de riñón, adenocarcinoma de pulmón y en carcinoma de células escamosas de pulmón) (Tomczak et al., 2015). LINC01614 también se asocia a supervivencia en cáncer de mama, en donde el grupo de pacientes que sobre-expresan el lncRNA tienen una mayor probabilidad de supervivencia global a 10 años. En la **Figura 8** se muestra una curva Kaplan-Meier donde se observa diferencia significativa en la supervivencia global (Cox=0.013 y P=0.0005) entre los grupos donde LINC01614 se encuentra sobre y sub-expresado, con un pronóstico más sombrío en el grupo donde está al alza.

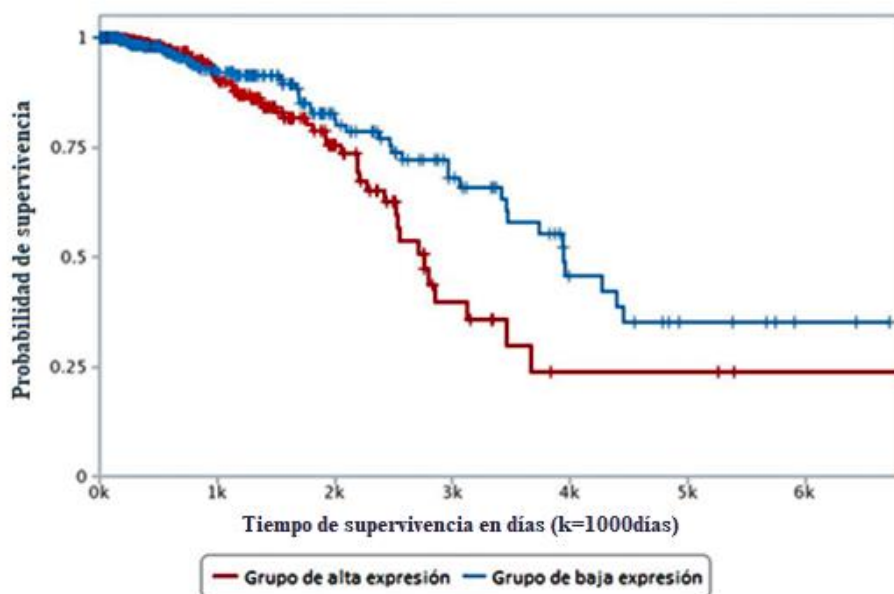


Figura 7 Curva Kaplan-Meier de LINC01614. Supervivencia global en pacientes con cáncer de mama (Cox=0.04 y P=0.024) donde LINC01614 se encuentra sobre-expresado (rojo) y se asocia a menor supervivencia. Modificada de (J. Li et al., 2015).

El segundo lncRNA candidato sobre-expresado en cáncer de mama es el RNA anti-sentido 1, SLC12A5 y MMP9 (SLC12A5-AS1). Este lncRNA está alterado al alza en los carcinomas de células escamosas de cabeza y cuello (HNSC), claras de riñón (KIRC) y escamosas de pulmón (LUSC); y en los adenocarcinomas de próstata (PRAD) y pulmón (LUAD). El gen SLC12A5-AS1 se encuentra en el 20q13.12 y tiene tres transcritos (Volders et al., 2013). SLC12A5-AS1 también se asocia a la supervivencia en pacientes con cáncer de mama. En la **Figura 9** se observa la curva de sobrevida global donde el grupo que sobre-expresa a SLC12A5 tiene menor supervivencia (Cox=0.04 y P=0.057).

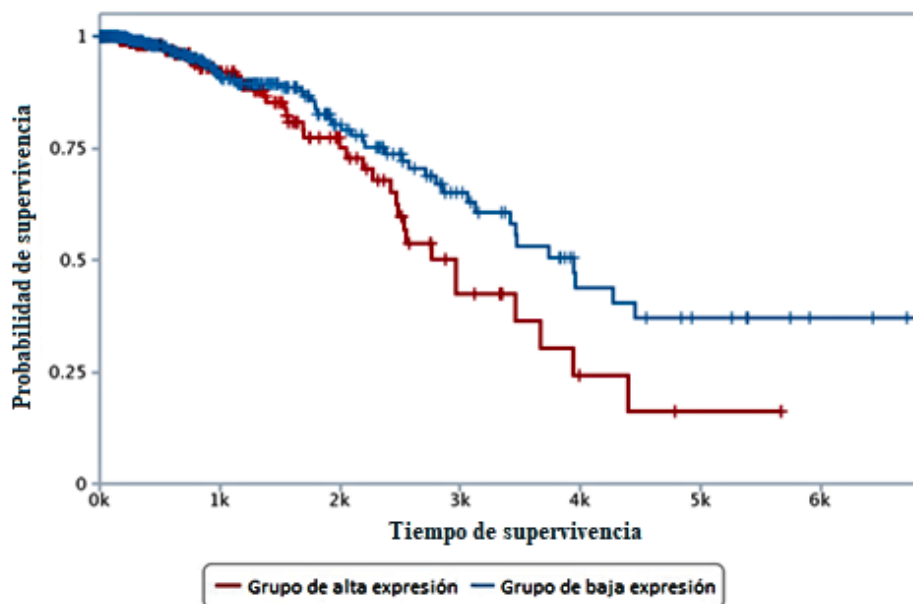


Figura 8 Curva Kaplan-Meier de SLC12A5. Curva de sobrevida global en pacientes con cáncer de mama de SLC12A5 (Cox=0.04 y P=0.57) donde se encuentra sobre-expresado (rojo) y asociado a un peor pronóstico. Modificada de (J. Li et al., 2015).

El microRNA 205 Host Gene (MIR205HG) es uno de los lncRNA candidatos seleccionados subexpresados en cáncer de mama, que además se encuentra alterado en cáncer de células escamosas de pulmón (al alza) y en cáncer de próstata (a la baja). El gen MIR205-HG es un RNA largo no codificante intergénico que se ubica en 1q32.2 y de acuerdo con LNCipedia cuenta con 27 transcritos. A MIR205HG se le asocia a menor supervivencia global en pacientes que lo expresan a la baja como se puede observar de la curva Kaplan-Meier en la **Figura 10**.

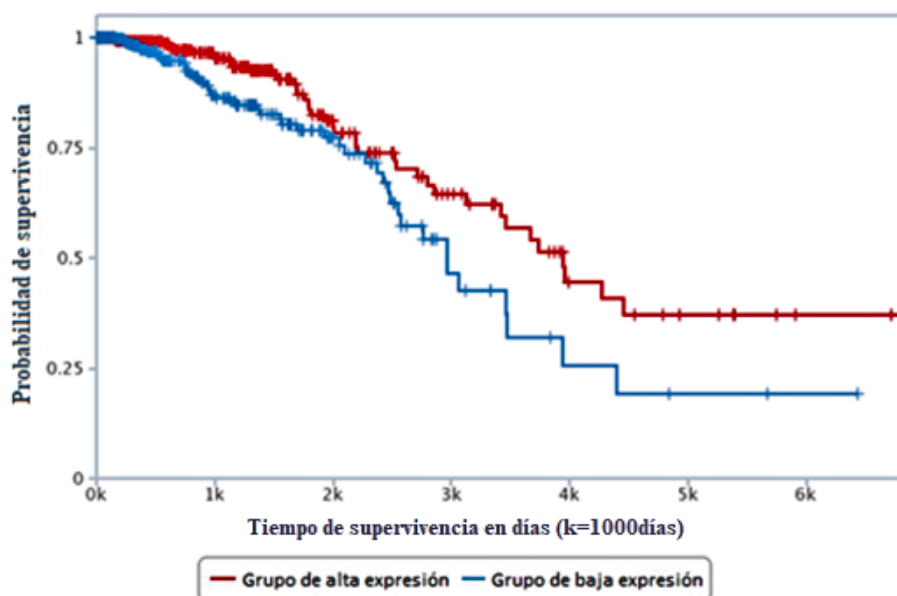


Figura 9 Curva Kaplan-Meier de MIR205-HG. Sobrevida global en pacientes con cáncer de mama (Cox=0.016 y P=0.008) donde MIR205HG se encuentra sobre-expresado (rojo) y asociado a un peor pronóstico. Modificada de (J. Li et al., 2015).

El RNA anti-sentido 2 o FGF14-AS2 es un lncRNA cuya expresión no sólo se encuentra alterada a la baja en cáncer de mama sino también en los adenocarcinomas de colon

(COAD), de pulmón (LUAD) y próstata (PRAD); y en los carcinomas de células escamosas de pulmón (LUSC) y claras de riñón (KIRC). El gen FGF14-AS2 se encuentra en el 13q33.1 y por splicing tiene dos transcritos. La sub-expresión de FGF14-AS2 está relacionada con menor supervivencia global en cáncer de mama (Cox=0.013 y P=0.0005) con un mayor número de decesos en el grupo de pacientes que lo sub-expresan, **Figura 11**.

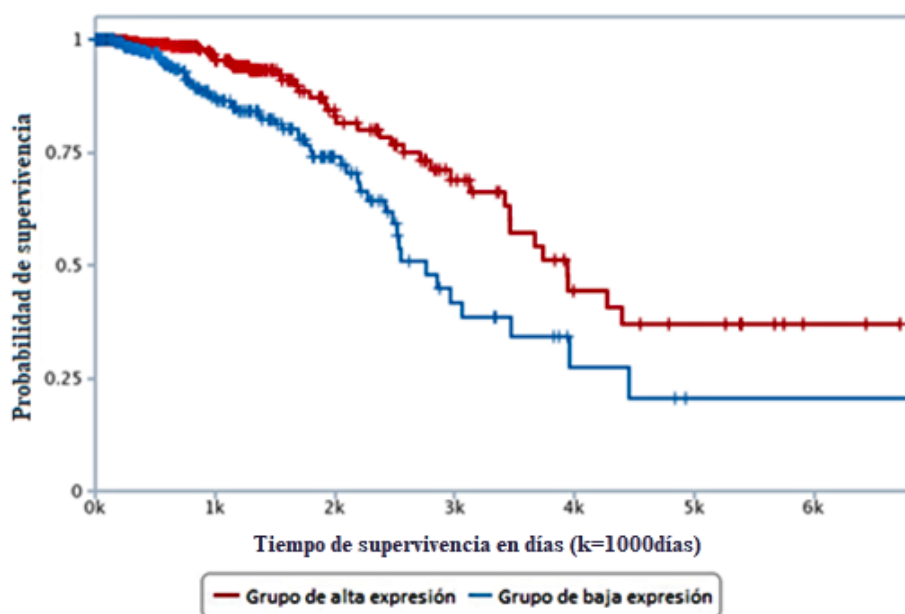


Figura 10 Curva Kaplan-Meier de FGF14-AS2 en cáncer de mama. Curva de supervivencia global en pacientes con cáncer de mama (Cox=0.013 y P=0.0005) donde FGF14-AS2 se encuentra sub-expresado (azul) y asociado a menor supervivencia global. Modificada de (J. Li et al., 2015).

La expresión regulada a la baja de FGF14-AS2 también se relaciona con menor supervivencia en el carcinoma de células claras de riñón (Cox=0.00050505 y P=0.0071138) donde los pacientes tienen mayor tasa de supervivencia global **Figura 12**. La selección del candidato FGF14-AS2 se realizó antes de su publicación (F. Yang et al., 2016).

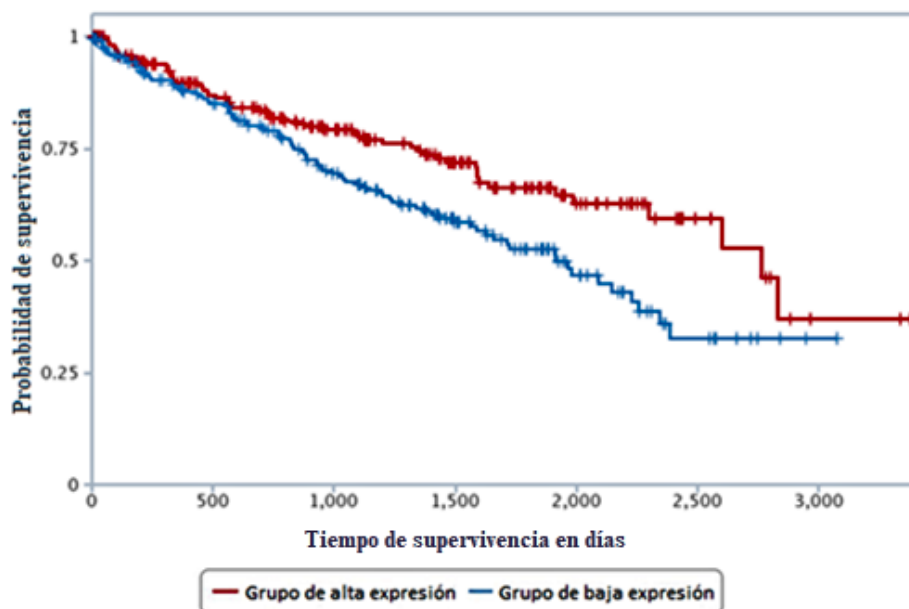


Figura 11 Curva Kaplan-Meier de FGF14-AS2 en cáncer de mama. Curva de sobrevivencia global en pacientes con carcinoma de células claras de riñón (Cox=0.04 y P=0.57) donde FGF14-AS2 se encuentra sub-expresado (azul) y asociado a un peor pronóstico. Modificada de Tanric (J. Li et al., 2015).

Diseño de cebadores.

En la Tabla 19 se muestran las secuencias diseñadas para los cuatro lncRNAs seleccionados mediante la herramienta Primer3.

Tabla 20. Lista de cebadores empleados en la q-PCR con sus características.

| Símbolo | Objetivo | Número de acceso GenBank | Sitio de unión del primer | Secuencia del oligo (5'→3') | T _m (°C) | Longitud del amplicón |
|-----------------|---|--------------------------|---------------------------|---------------------------------------|---------------------|-----------------------|
| FGF14-AS2 | <i>Homo sapiens</i> FGF14 RNA anti-sentido 2 | NR_036487.1 | Exón 2 | FP: TCGTGGCGGATAAAGAGC | 57.2 | 135pb |
| | | | | RP: GGGACTACAAAGGGCAGG | 56.9 | |
| MIR205HG | <i>Homo sapiens</i> MIR205 Host Gene | NR_145433.1 | Exón 1 | FP: AGGAGTCATTTCTGTTCCGC | 60.4 | 83pb |
| | | | | RP: TCCAAGATGGGTACTTGAGAGA | 45.5 | |
| LINC01614 | <i>Homo sapiens</i> RNA Largo No Codificante Intergénico 1614 | NR_132383 | Exón 2 | FP: CCAGACCAAGGCTGCTGT | 61.6 | 98pb |
| | | | | RP: CCTAGAAGAGTCATTTGAAGTC ATGT | 61 | |
| SLC12A5- AS1 | <i>Homo sapiens</i> RNA anti- sentido 1 SLC12A5 y MMP9 | ENSG000002040 44 | Exón 1 | FP: CATCCAGACAGAGCACGCC | 61.1 | 138pb |
| | | | | RP: CCCACCCACGCAGAACAA | 61.1 | |

Estandarización de la q-PCR.

La validación de ambos cebadores y del amplicón se realizó para los cuatro lncRNAs y los endógenos ACTB y TBP. El endógeno rRNA 18S ya había sido estandarizado en el laboratorio, por lo cual se emplearon cantidades equimolares de ambos cebadores (0.3mM).

Se realizaron curvas disociación donde se varió la proporción de ambos cebadores (forward y reverse) para encontrar aquella que favoreciera un solo producto de amplificación. Las concentraciones probadas fueron: 300nM, 500nM y 800nM, y sus posibles combinaciones con un total de 9 condiciones mostradas en la Tabla 20.

Tabla 21. Condiciones de concentración de cebadores para las curvas de disociación.

A

D

G

| | | |
|------------------|------------------|------------------|
| [FP/RP]= 300/300 | [FP/RP]= 500/300 | [FP/RP]= 800/300 |
| B | E | H |
| [FP/RP]= 300/500 | [FP/RP]= 500/500 | [FP/RP]= 800/500 |
| C | F | I |
| [FP/RP]= 300/800 | [FP/RP]= 500/800 | [FP/RP]= 800/800 |

Además, se corroboró la especificidad de los cebadores y las condiciones seleccionadas por medio de electroforesis en agarosa de los productos amplificados. Se seleccionaron aquellas condiciones en las que se observa un solo pico en la curva de disociación y a su vez un solo producto de amplificación por electroforesis en gel de agarosa.

Se estandarizaron las condiciones de amplificación de los genes endógenos ACTB y TBP en la línea celular MCF-7. Las curvas de disociación e imágenes de los geles de agarosa se muestran en la sección de ANEXOS. Las concentraciones óptimas de cebadores para el endógeno ACTB fueron [FP/RP]= 500/300 y para TBP, [FP/RP]= 800/300.

Se procedió a estandarizar los cuatro lncRNAs. Se realizaron las curvas de disociación con cada candidato, de acuerdo con la **Tabla 20**, y la electroforesis de los productos de amplificación. Donde se observaron múltiples productos de amplificación tanto en las curvas de disociación como en los geles. Este problema se solucionó corroborando experimentalmente las fuentes de variación posibles en una PCR, **Tabla 21**

Tabla 22. Fuentes de variación

| Variable | Abordaje |
|----------|----------|
|----------|----------|

| | |
|----------------------|--|
| Pipeteo del operador | Se realizaron experimentos de reproducibilidad con dos operadores preparando las reacciones de PCR con ACTB. |
| cDNA degradado | Las reacciones de PCR con ACTB se realizaron con el cDNA inicial y uno recién preparado con RNA fresco. Se evaluó la integridad del RNA por electroforesis y espectrofotometría. |
| SYBR Green | Se prepararon las reacciones de PCR con el reactivo usual PowerUp® SYBR® Green Master Mix y SYBR® Green Master Mix empleando el gen ACTB. |
| Cebadores | Se corrobora la especificidad del primer por Primer Blast. |

Finalmente se decidió cambiar el modelo celular MCF-7 al MDA-468 bajo la premisa de que el primero no expresaba los lncRNAs en una cantidad detectable por SYBR Green. Se eligió MDA-468 debido a que es la línea celular que expresa en mayor medida aquél lncRNA con el mayor nivel de expresión entre los candidatos, MIR205HG. Se estandarizaron las concentraciones de cebadores de MIR205HG en MDA-468, las cuales fueron [FP/RP]= 500/300.

Curva de eficiencia relativa.

A pesar de que, en teoría, la amplificación en la PCR sigue el modelo de 2^n , es decir cada molécula se supone se duplica en cada ciclo, en la práctica, llega a ser menor. Cuando se amplifican distintos genes blanco y estos no se amplifican con la misma eficiencia, la comparación entre ellos requiere de un factor de corrección. Es por ello que es necesario realizar una curva de eficiencia relativa. La curva de eficiencia de amplificación entre los lncRNA y los genes endógenos se realizó por medio de una curva de calibración que graficaba los valores de Ct con respecto al logaritmo base 2 de la concentración de RNA total. La curva

de calibración constó de 8 puntos en los que se varió la concentración de RNA inicial. Las curvas de eficiencia se pueden observar en el apartado de Anexos.

Validación de MIR205HG en MDA-468.

La validación del lncRNA MIR205HG se realizó en el modelo celular MDA-468 en 16 muestras de tumores de cáncer de mama y 16 muestras de tejido normal adyacente. Todas las reacciones se realizaron por triplicado. Se incluyeron, además de los controles negativos, controles positivos (con cDNA de MDA-468) para corroborar que la reacción realmente procediera. El valor de la tasa de cambio o fold-change se determinó por medio del método comparativo Ct ($2^{-\Delta\Delta C_t}$). Se empleó la prueba estadística T de Student de una sola cola.

Reacción en cadena de la polimerasa (qPCR).

La validación de MIR205HG se realizó en una cohorte de 16 muestras de tumores y 16 tejidos adyacentes, no pareadas. Fue necesario cambiar los genes endógenos ACTB y TBP por 18S rRNA debido a que resultaron no ser buenos normalizadores en las muestras de tumores. De acuerdo con los valores obtenidos, MIR205HG se encuentra 0.29 veces ($P=0.0113$) más expresado en tejido normal adyacente que en tumores de cáncer de mama. La expresión comparativa de MIR205HG entre tumor y tejido normal adyacente se muestra en la **Figura 13.**

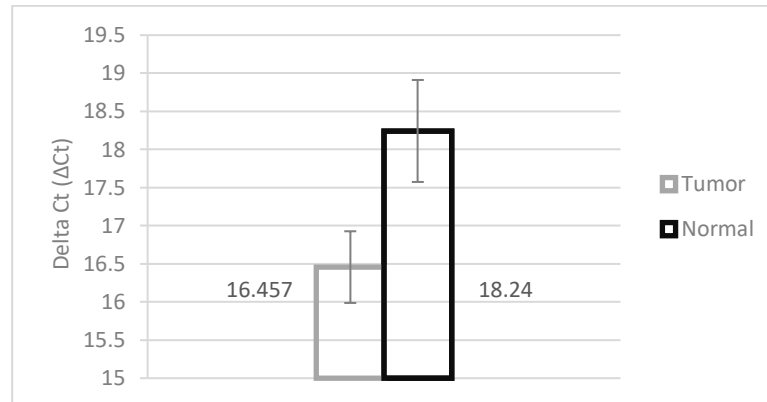


Figura 12 Expresión de MIR205HG datos de q-PCR.

Además, se corroboró que MIR205HG está sub-expresado en tumores respecto al normal en una cohorte independiente de 105 muestras de tejido normal adyacente y 837 tumores, datos de microarreglos obtenidos de TCGA (**Figura 14**).

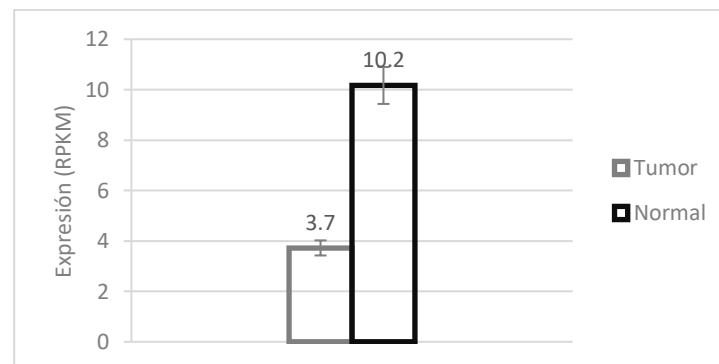


Figura 13 Expresión de MIR205HG datos de expresión de RNA-seq de TCGA.

Descripción in silico de MIR205HG.

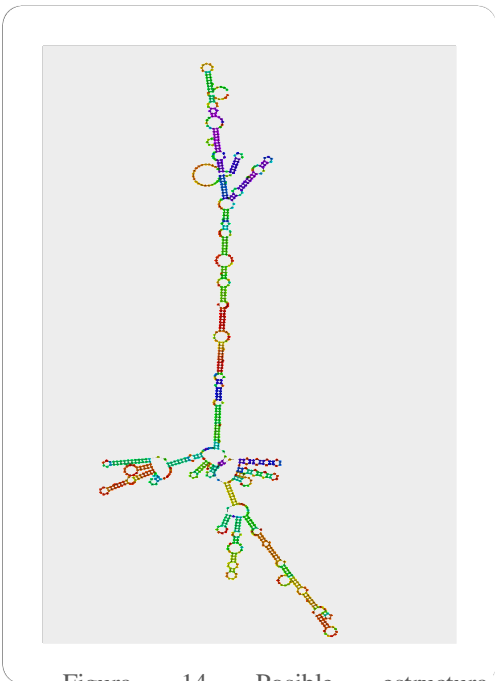


Figura 14 Posible estructura tridimensional del transcrito MIR205HG-4 (Volders et al., 2013).

El gen MIR205HG tiene 10 transcritos según Genecards y según LNCipedia, 27. Se caracterizó in silico la secuencia de la isoforma MIR205HG-4 mediante el programa AnnoLnc.

El transcrito MIR205HG-4 (Transcript ID: T029619; Gene ID: G006597) tiene una longitud de 888pb. Tiene cuatro exones y se ubica en chr1:209600864-209606783[+] (GRCh37/hg19). Posiblemente presenta

la siguiente estructura secundaria tridimensional (Figura 15) de acuerdo con AnnoLnc. Además de que

MIR205HG-4 comparte el mayor número de exones con las demás isoformas, se escogió este transcrito por ser uno de los que se observa desregulados, aparte de cáncer de mama, en el carcinoma de células escamosas de pulmón y en el adenocarcinoma de próstata (descritos por Tanric y TCGA); y en el carcinoma de cabeza y cuello, de acuerdo con la base de datos MiTranscriptome (Figura 16). De acuerdo con Illumina Body Map, un programa de expresión de tejido en personas sanas MIR205HG se encuentra enriquecido en los tejidos del seno, pulmón, próstata y testículos. Sin embargo, en la base de datos de AnnoLnc, se reporta que el

enriquecimiento de MIR205HG se observa únicamente en los tejidos del seno, pulmón y próstata. Lo anterior nos indica que MIR205HG es un lncRNA tejido-específico.

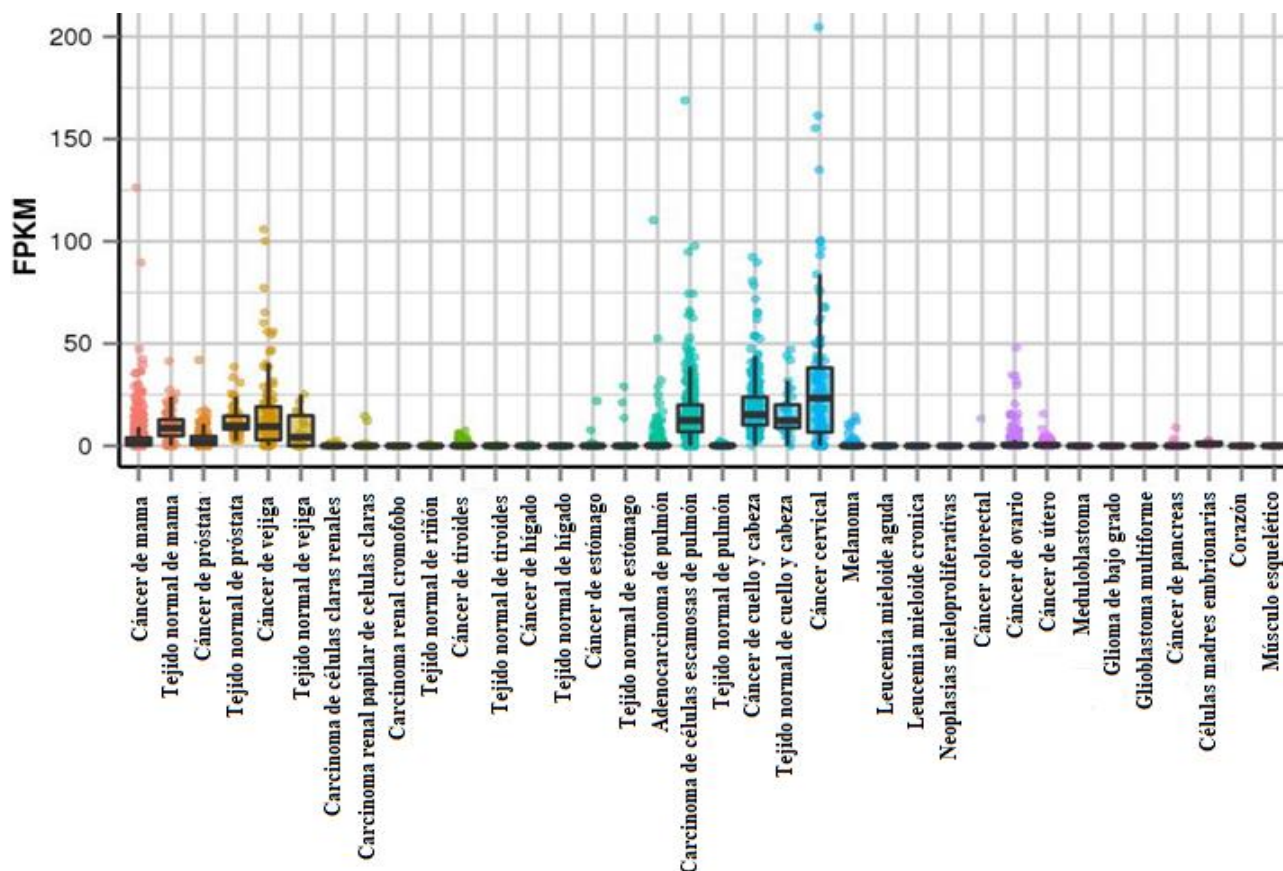


Figura 15 Expresión del transcrito MIR205HG:4 entre distintos tipos de cáncer (MiTranscriptome).

La regulación post-transcripcional en los lncRNAs es fundamental para controlar la expresión de lncRNAs. El MIR205HG interactuaría con las cinco familias de miRNAs: miR-143/1721/4770, miR-205/205ab, miR-214/761/3619-5p, miR-23abc/23b-3p y miR-194. Los miRNAs miR-143/1721/4770, miR-205/205ab, miR-214/761/3619-5p y miR-194 se encuentran caracterizados y se sabe que están implicados en diversos procesos de carcinogénesis no solo en cáncer de mama, esto se muestra resumido en la **Tabla 22**. Debido a esto, se propone que MIR205HG actúa como esponja de miRNAs, y por ende su expresión

podría estar controlada de esta manera tanto en cáncer de mama como en el carcinoma de células escamosas de pulmón y en el adenocarcinoma de próstata.

Tabla 23. Predicción de miRNAs asociados a al lncRNA MIRS205HG

| PCG | Características |
|---------------------|---|
| miR-143/1721/4770 | Se ubica en el cromosoma 5q33. Se reporta que está desregulado a la baja en cáncer de pulmón de células pequeñas (Cordes et al., 2009), cáncer colorectal (Slaby et al., 2008), cáncer de próstata (B. Xu et al., 2011) y cáncer de mama (Zhou, Dong, Huang, Sun, & Wu, 2017). MiRNA-143 ha probado su efecto anticancerígeno apuntando múltiples genes relacionados con la proliferación celular, apoptosis y migración, Bcl-2 (Park et al., 2011), MYO6 (Cordes et al., 2009), ELK1 (Cordes et al., 2009) y ERK5 (Akao, Nakagawa, Iio, & Naoe, 2009). |
| miR-205/205ab | El miR-205 es un inhibidor de apoptosis (Pang & Yue, 2017). Tiene potencial como biomarcador en el diagnóstico de cáncer de pulmón de células no pequeñas (Duan et al., 2017) y en cáncer cervical (Zhuo & Yu, 2017). |
| miR-214/761/3619-5p | MiR-214 es un supresor tumoral (He et al., 2016; F. Wang, Li, Chen, Zhu, & Gu, 2016), posible biomarcador de diagnóstico y pronóstico en cáncer de vejiga (Jinfeng Wang et al., 2015). MiR-214 aumenta la sensibilidad de las células cancerígenas al tamoxifen (X. Yu et al., 2015). |
| miR-23abc/23b-3p | No se encuentra caracterizado. |
| miR-194 | MiR23abc es un supresor tumoral por lo que inhibe la proliferación y metástasis en el melanoma (Guo, Hui, Zhang, Chang, & Tao, 2016). Es un inhibidor de la transición mesenquimal-epitelial en cáncer gástrico (Z. Li et al., 2014). |

En la **Figura 17** se muestra un esquema general de los sitios de unión de MIR205HG a miRNAs y sus factores transcripcionales.

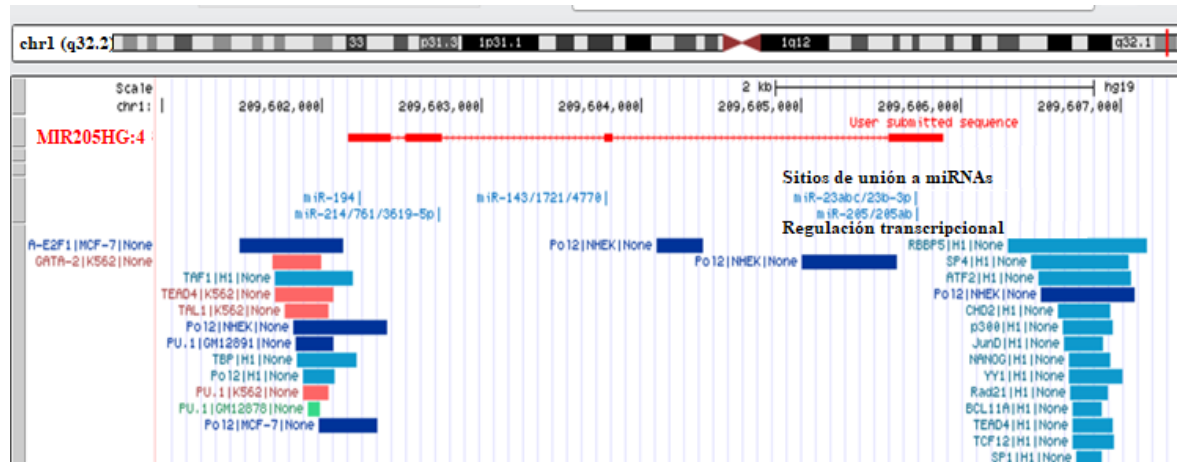


Figura 16 Vista integradora de los sitios de unión de miRNAs y regulación transcripcional de MIR205HG (Genome Browser).

De la **Figura 18** se observa la expresión de MIR205HG por etapa en cáncer de mama. En dicha gráfica de cajas se observa que la expresión de este lncRNA se encuentra en su mínimo en la etapa IV (valor de $P=0.064$) de acuerdo con Tanric.

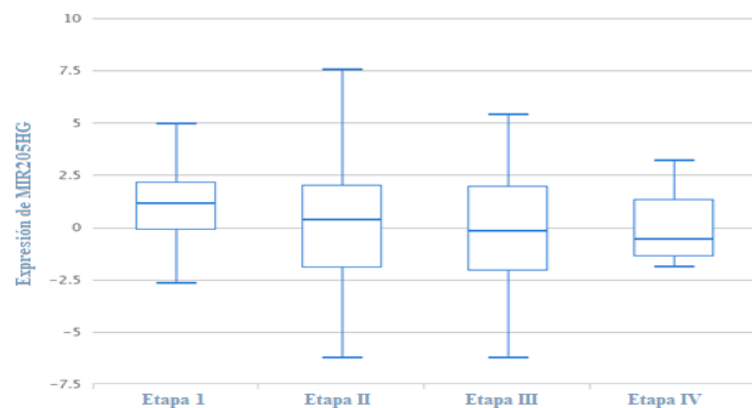


Figura 17 Expresión de MIR205HG por etapas en cáncer de mama.

Discusión de resultados

Hasta ahora encontrar un gen candidato había sido frecuentemente el producto secundario de diversos estudios y no el objetivo inicial de la investigación, esto posiblemente a que la generación de los perfiles de expresión involucraba técnicas extensas—como northern blot, hibridación *in situ* e inmunohistoquímica. Afortunadamente, los ensayos de arreglos a gran escala —microarreglos de cDNA (Schena et al. 1995), chips de cebadores (Lockhart et al. 1996), librerías de secuenciación de cDNA (Adams et al. 1993) y diversos análisis de expresión genómicos (Shyr & Liu, 2013; Yohe & Thyagarajan, 2017) —y el avance en el desarrollo las herramientas para análisis bioinformáticos han sentado las bases para una nueva etapa en la estrategia de identificación de genes blanco *in silico* (2,6). Diversos estudios de

expresión génicos en distintos tipos de cáncer pueden ser descargados de bases de datos públicas (J. Li et al., 2015; C. X. Ma & Ellis, 2013; Tomczak et al., 2015). Notablemente, los análisis genómicos realizados en diversas bibliotecas de cDNA de tejidos normales y tumores ofrecen un punto de partida atractivo para el descubrimiento de genes involucrados en cáncer. Sin embargo, es importante subrayar que, aunque la detección *in silico* sea prometedora no queda exenta de las limitantes de todo enfoque bioinformático —en el cual los resultados necesitan de la posterior validación experimental para evitar falsos derivados de valores atribuidos erróneamente.

Identificación de 4 lncRNAs desregulados en diversos tipos de cáncer.

Con el fin de determinar cuatro lncRNAs como posibles candidatos para su caracterización molecular en un subsecuente estudio en cáncer de mama, este proyecto se centró la selección de genes de lncRNAs mediante el minado de la base de TCLA la cual contiene datos de expresión por secuenciación de ARN (RNA-seq) de 5,037 tumores de 13 tipos tumorales y 424 tejidos normales de datos de TCGA. TCLA nos proporcionó información sobre aquellos lncRNAs que se encontraban desregulados en diversos tipos tumorales —1144 sobreexpresados y 1109 subexpresados. El hecho de que existan lncRNAs asociados a diversos tipos tumorales nos habla de que posiblemente están implicados y sean importantes para el establecimiento y/o progresión del proceso de carcinogénesis no aislado a un solo tipo de tejido. Posteriormente con el fin de seleccionar únicamente aquellos que estuvieran significativamente desregulados en tumores de cáncer de mama, se examinaron los datos de secuenciación de ARN (RNA-seq) de la base de datos de TCLA de su lista de genes

de lncRNAs desregulados en BRCA (tasa de cambio >2) —990 tumores y 106 tejidos normales. Con lo que finalmente fue posible obtener una lista de genes candidatos —5 sobreexpresados y 15 subexpresados en cáncer de mama— asociados a diversos tipos tumorales significativamente desregulados en cáncer de mama.

Con el fin de determinar si la expresión alterada de estos genes se encontraba asociada a una característica clínica, se evaluó la asociación de los 17 genes desregulados con la supervivencia global de las pacientes en la base de datos pública (Tanric) —datos de análisis de expresión RNA-seq de 837 tumor y 105 muestras de tejido normales de cáncer de mama— que nos proporciona las curvas Kaplan Meier y el análisis de regresión de Cox de cada gen de lncRNA. De acuerdo con lo anterior, se encontraron 15 posibles genes candidatos desregulados en distintos tipos tumorales asociados a la supervivencia global de la cual finalmente se decidió centrarse en 4 genes de lncRNAs estadísticamente más representativos (Cox <0.05 y de $P < 0.04$) de acuerdo con los valores proporcionados del análisis de regresión de Cox y sus curvas Kaplan Meier proporcionadas por Tanric. Por el método descrito de minado de bases de datos se encontró 4 nuevos lncRNAs cuyas características los hacen buenos candidatos de estudio para dilucidar sus mecanismos de acción y el papel que juegan en la carcinogénesis mamaria.

Los nuevos lncRNAs encontrados son: LINC01614 —desregulado en cáncer de mama, colon, carcinoma de células escamosas de cabeza y cuello, carcinoma de células claras de riñón, adenocarcinoma de pulmón y en carcinoma de células escamosas de pulmón—, SLC12A5-AS1 —desregulado en carcinomas de células escamosas de cabeza y cuello, claras de riñón y escamosas de pulmón; y en los adenocarcinomas próstata y pulmón, y cáncer de mama—,

FGF14-AS2 —desregulado en cáncer de mama, adenocarcinomas de colon, de pulmón y próstata—, MIR205-HG —desregulado en cáncer de mama, cáncer de células escamosas de pulmón y en cáncer de próstata. Más aún, de las curvas de Kaplan Meier proporcionadas por Tanric se puede observar que en cáncer de mama los lncRNAs LINC01614 y SLC12A5 se encuentran sobreexpresados en tumores de cáncer de mama y asociados con mejor supervivencia global; y los lncRNAs FGF14-AS2 y MIR205-HG, están subexpresados en tumores de cáncer de mama y asociados con mejor supervivencia global en las pacientes.

El valor de identificar estos 4 nuevos lncRNAs candidatos por medio del minado de bases de datos radica en las publicaciones, posteriores a su selección, de tres artículos de los lncRNAs: FGF14-AS2 y MIR205-HG. En el estudio que realiza F. Yang et al. sobre FGF14-AS2 implicado en cáncer de mama (Yang et al., 2016) reportan y validan que efectivamente es un lncRNA sub-expresado en cáncer de mama y cuya expresión anormal vinculada a un fenotipo clínico lo hace capaz de actuar como un biomarcador de pronóstico independiente, asociado a la supervivencia global de las pacientes. En cuanto a MIR205-HG, en un estudio de carcinoma de células escamosas de cabeza y cuello (Agostino et al., 2018) encuentran que la sub-expresión de este lncRNA reduce significativamente la proliferación, la migración celular y la formación de colonias de las células de cabeza y cuello. Además, en el mismo estudio encuentran que la expresión de MIR205-HG se encuentra al alza cuando una mutante de p53 está presente en comparación con el tejido normal pareado, e incluso en comparación con aquellos tejidos tumorales que no expresan la mutante en p53. En el estudio describen que MIR205-HG actúa como un competidor endógeno al unirse a miR-590-3p y así modula la expresión de las proteínas ciclina B, cdk1 y YAP, las cuales se observan enriquecidas en este

tipo tumoral. Por otro lado, se publicó un artículo por Zou et al. en el que evalúan un polimorfismo que reside en el cuarto exón de mir205-HG y en el que concluyen que no existe asociación del polimorfismo con el riesgo de cáncer de mama sin embargo sí con cáncer de pulmón (Zou et al., 2016).

De las observaciones anteriores se concluye y se valida el método de minado de bases de datos para la selección de genes de lncRNAs que se encuentren implicados en el proceso de carcinogénesis en diversos tejidos, asociados a una característica clínica como lo es la supervivencia global. Así mismo se reafirma que los lncRNAs pueden actuar por mecanismos diversos y estar involucrados en el proceso de carcinogénesis, como lo es en el caso de MIR205-HG.

El lncRNA MIR205-HG está sub-expresado en cáncer de mama

Hasta ahora muchos estudios se han enfocado en miR-205 como un importante regulador de la expresión en cáncer, sin embargo, no existen estudios que intenten esclarecer el mecanismo y función del lncRNA MIR205-HG en el proceso de carcinogénesis. En este estudio se reporta y se valida la expresión de MIR205-HG en cáncer de mama en una cohorte independiente de tumores y muestras de tejido normales de pacientes mexicanas con cáncer de mama (t=16; n=16). Se encontró que este lncRNA está expresado 0.29 (P=0.0113) veces más en tejido normal en comparación con el tumor. Además, se validó la expresión de MIR205-HG en una cohorte independiente con datos de secuenciación de RNA (RNA-seq) de tumores de mama de TCGA —990 tumores y 106 tejidos normales. En los cuales consistentemente este lncRNA

se encuentra sub-expresado. De acuerdo con los resultados de la validación podemos concluir que la sub-expresión de mir205-HG es un fenómeno biológico y que su alteración posiblemente es necesaria para el proceso de la oncogénesis en mama, como posible supresor tumoral.

La sub-expresión de MIR205-HG está asociada a mayor supervivencia global de pacientes.

Con el fin de demostrar si MIR205-HG estaba asociado a la supervivencia global en cáncer de mama, se analizó la base de datos de Tanric. El seguimiento clínico de las pacientes es de 1000 días. Como se mostró en la curva Kaplan Meier de MIR205-HG (FIGURA 10) la diferencia de supervivencia de ambos grupos es estadísticamente significativa ($P=0.008$). La supervivencia en el grupo de pacientes con valores bajos de expresión de MIR205-HG se relaciona con mejor supervivencia global. Respecto al valor predictivo de MIR205-HG mediante el análisis por regresión de Cox, el cual ajusta las variables clínicas, nos indica que MIR205-HG es un biomarcador de pronóstico independiente y que su expresión está asociada con un pronóstico no favorable en cáncer de mama ($Cox=0.016$).

Mir205-HG puede actuar como posible ceRNA y regular a mir205.

Además, se realizó una caracterización más profunda del lncRNA mir205-HG por medio del minado de bases de datos públicas. Con este fin, se eligió la secuencia de la isoforma de mir205-HG-4 —anotada como la desregulada en los diversos tipos de cáncer en MiTranscriptome— y se buscaron los posibles miRNAs a los cuales tuviera sitios de unión en la base de datos AnnoLnc. Uno de los miRNAs asociados fue su gen huésped mir-205 —un supresor tumoral implicado en la oncogénesis de diversos tejidos, incluido el de mama (De Cola et al., 2015; Duan et al., 2017; Elgamal et al., 2013; Nordby et al., 2017; Orang, Safaralizadeh, & Feizi, 2014; Piovan et al., 2012; Qin et al., 2013; H. Wu, Zhu, & Mo, 2017; Zhuo & Yu, 2017). De lo anterior, se propone que mir205-HG actúa como esponja de miRNAs y posiblemente regule la actividad de su gen huésped mir-205. Sin embargo, no sería el único mecanismo por el cual pudiera estar actuando MIR205-HG en cáncer de mama. Por otro lado se ha reportado que mir-205 se encuentra sub-expresado en cáncer de mama (Elgamal et al., 2013), esta observación en conjunto con nuestros resultados de validación donde MIR205-HG está sub-expresado en dicho cáncer dan indicios a pensar que posiblemente ambos genes sean co-transcritos.

Con todo lo dicho anteriormente, nuestros datos revelan que MIR205-HG está involucrado en el proceso de cáncer de mama en el cual podría actuar como un supresor tumoral. Además, se propone que MIR205-HG podría ser un biomarcador de pronóstico independiente con asociación a la supervivencia global de las pacientes en cáncer de mama, aunque más estudios *in vivo* e *in vitro* deben realizarse para poder validar los efectos de MIR205-HG y sus mecanismos como gen hospedero.

Conclusiones.

- El uso de las bases de datos genómicas de acceso público es útil para la obtención de genes candidatos desregulados en cáncer de mama asociadas a características clínicas como supervivencia global de los pacientes. Mediante su minado se identificaron cuatro RNAs largos no codificantes, LINC01614, SLC12A5-AS, FGF14-AS2 y MIR205HG alterados en cáncer de mama y otros tumores, asociados a supervivencia global.
- Se realizó la validación de un RNA largo no codificante, MIR205HG, en una cohorte independiente de 16 muestras de tumores de cáncer de mama y 16 de tejido normal

adyacente. MIR205HG se encontró expresado 0.3 ($P=0.0113$) veces más en tejido normal adyacente con respecto al normal. Este resultado se corroboró con los datos del TCGA, donde MIR205HG también se encuentra sub-expresado en tumores.

- Se realizó una descripción in silico de MIR205HG empleando bases de datos públicas.

Perspectivas

- Realizar la validación de los RNAs largos no codificantes LINC01614 y SLC12A5-AS en muestras independientes de tumores de mama.
- Realizar estudios de caracterización estructural y funcional de MIR205HG, LINC01614 y SLC12A5-AS.

Referencias.

- ACS. (2008). Breast Cancer Facts & Figures 2007-2008. *American Cancer Society*, 36. <https://doi.org/10.1046/j.1523-5394.2000.82001.x>
- Aigrain, L., Gu, Y., & Quail, M. A. (2016). Quantitation of next generation sequencing library preparation protocol efficiencies using droplet digital PCR assays - a systematic comparison of DNA library preparation kits for Illumina sequencing. *BMC Genomics*, 17(1), 458. <https://doi.org/10.1186/s12864-016-2757-4>
- Akao, Y., Nakagawa, Y., Iio, A., & Naoe, T. (2009). Role of microRNA-143 in Fas-mediated apoptosis in human T-cell leukemia Jurkat cells. *Leukemia Research*, 33(11), 1530–1538. <https://doi.org/10.1016/j.leukres.2009.04.019>
- Aken, B. L., Ayling, S., Barrell, D., Clarke, L., Curwen, V., Fairley, S., ... Searle, S. M. J. (2016). The Ensembl Gene Annotation System. *Database : The Journal of Biological Databases and Curation*, 2016, baw093. <https://doi.org/10.1093/database/baw093>
- Al-Amri, A. M. (2005). Prevention of breast cancer. *Journal of Family & Community Medicine*, 12(2), 71–4. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/23012078>
- Alcalá Pérez, D., Barrera Pérez, M., Yépiz, R. C., Luisa, A., & Pérez, C. (2015). Células madre cancerígenas: conceptos actuales, 24(Cm), 47–51.
- American Joint Committee on Cancer Staging System for Breast Cancer. (2002). Clasificación TNM. *AJCC Cancer Staging Manu*, 13–16.
- Anderson, W. F., Rosenberg, P. S., Prat, A., Perou, C. M., & Sherman, M. E. (2014). How many etiological subtypes of breast cancer: two, three, four, or more? *Journal of the National Cancer Institute*, 106(8), dju165-dju165. <https://doi.org/10.1093/jnci/dju165>

- Barnard, M. E., Boeke, C. E., & Tamimi, R. M. (2015). Established breast cancer risk factors and risk of intrinsic tumor subtypes. *Biochimica et Biophysica Acta - Reviews on Cancer*, 1856(1), 73–85. <https://doi.org/10.1016/j.bbcan.2015.06.002>
- Bartonicek, N., Maag, J. L. V., & Dinger, M. E. (2016). Long noncoding RNAs in cancer: mechanisms of action and technological advancements. *Molecular Cancer*, 15(1), 43. <https://doi.org/10.1186/s12943-016-0530-6>
- Beltran, M., Puig, I., Peña, C., García, J. M., Álvarez, A. B., Peña, R., ... Herreros, A. G. De. (2008). A natural antisense transcript regulates Zeb2 / Sip1 gene expression during Snail1-induced epithelial – mesenchymal transition. *Genes and Development*, 22, 756–769. <https://doi.org/10.1101/gad.455708.in>
- Bernard, P. S., Parker, J. S., Mullins, M., Cheung, M. C. U., Leung, S., Voduc, D., ... Perou, C. M. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of Clinical Oncology*, 27(8), 1160–1167. <https://doi.org/10.1200/JCO.2008.18.1370>
- Bertram, J. S. (2000). The molecular biology of cancer. *Molecular Aspects of Medicine*, 21, 167–223. [https://doi.org/10.1016/S0098-2997\(00\)00007-8](https://doi.org/10.1016/S0098-2997(00)00007-8)
- Bhartiya, D., Pal, K., Ghosh, S., Kapoor, S., Jalali, S., Panwar, B., ... Scaria, V. (2013). LncRNome: A comprehensive knowledgebase of human long noncoding RNAs. *Database*, 2013, 1–6. <https://doi.org/10.1093/database/bat034>
- Bl, D. (2013). 2 2013%.
- Bland, K. I., & Copeland, E. M. (2007). *La mama manejo multidisciplinario de las enfermedades benignas y malignas*. Editorial Médica Panamericana. Retrieved from https://books.google.es/books?id=01yvLPQD7QC&pg=PA23&hl=es&source=gbs_selected_pages&cad=2#v=onepage&q&f=false
- Bland, K. I., Menck, H. R., Scott-Conner, C. E., Morrow, M., Winchester, D. J., & Winchester, D. P. (1998). The National Cancer Data Base 10-year survey of breast carcinoma treatment at hospitals in the United States. *Cancer*, 83(6), 1262–1273.
- Bonnefoix, T., Bonnefoix, P., Verdiel, P., & Sotto, J. J. (1996). Fitting limiting dilution experiments with generalized linear models results in a test of the single-hit Poisson assumption. *Journal of Immunological Methods*, 194(2), 113–119. [https://doi.org/10.1016/0022-1759\(96\)00077-4](https://doi.org/10.1016/0022-1759(96)00077-4)
- Bosic, M. M., Brasanac, D. C., Stojkovic-Filipovic, J. M., Zaletel, I. V., Gardner, J. M., & Cirovic, S. L. (2016). Expression of p300 and p300/CBP associated factor (PCAF) in actinic keratosis and squamous cell carcinoma of the skin. *Experimental and Molecular Pathology*, 100(3), 378–385. <https://doi.org/10.1016/j.yexmp.2016.03.006>
- Caffarel, M. M., Moreno-Bueno, G., Cerutti, C., Palacios, J., Guzman, M., Mechta-Grigoriou, F., & Sanchez, C. (2008). JunD is involved in the antiproliferative effect of Delta9-tetrahydrocannabinol on human breast cancer cells. *Oncogene*, 27(August 2007), 5033–5044. <https://doi.org/10.1038/onc.2008.145>
- Caliskan, M., Gatti, G., Sosnovskikh, I., Rotmensch, N., Botteri, E., Musmeci, S., ... Luini, A. (2008). Paget's disease of the breast: The experience of the European institute of oncology and review of the literature. *Breast Cancer Research and Treatment*, 112(3), 513–521. <https://doi.org/10.1007/s10549-007-9880-5>
- Campbell, L. L., & Polyak, K. (2007). Breast tumor heterogeneity: Cancer stem cells or

- clonal evolution? *Cell Cycle*, 6(19), 2332–2338. <https://doi.org/10.4161/cc.6.19.4914>
- Cancer, T., Atlas, G., Muzny, D. M., Bainbridge, M. N., Chang, K., Dinh, H. H., ... Tarnuzzer, R. W. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*, 487(7407), 61–70. <https://doi.org/10.1038/nature11412>
- Canese, K., & Weis, S. (2013). PubMed: The bibliographic database. *The NCBI Handbook*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK153385/pdf/Bookshelf_NBK153385.pdf
- Cao, Q., Brenner, J. C., Laxman, B., Asangani, I., Kominsky, H. D., Cao, X., ... Siddiqui, J. (2012). Transcriptome Sequencing Identifies PCAT-1, a Novel lincRNA Implicated in Prostate Cancer Progression *John*, 29(8), 742–749. <https://doi.org/10.1038/nbt.1914>. Transcriptome
- Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., ... RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). (2005). The transcriptional landscape of the mammalian genome. *Science (New York, N.Y.)*, 309(5740), 1559–63. <https://doi.org/10.1126/science.1112014>
- Cassidy, Jim. Bissett, Donald. Spence, Roy. Payne, M. (2010). *Oxford Handbook of Oncology*.
- Cawley, S., Bekiranov, S., Ng, H. H., Kapranov, P., Sekinger, E. A., Kampa, D., ... Gingeras, T. R. (2004). Unbiased Mapping of Transcription Factor Binding Sites along Human Chromosomes 21 and 22 Points to Widespread Regulation of Noncoding RNAs. *Cell*, 116(4), 499–509. [https://doi.org/10.1016/S0092-8674\(04\)00127-8](https://doi.org/10.1016/S0092-8674(04)00127-8)
- Characterization, C., Moore, M. A. S., Williams, N., Cancer, D. M., Walter, T., & Hospital, R. M. (1973). In Vitro Colony Formation by Normal and Leukemic Human Hematopoietic. *Journal of the National Cancer Institute*, 50(3), 603–623.
- Chen, Q., Liang, Y., Zhang, Y., Jiang, M., Han, Z., Liang, Y., ... Zhong, W. (2017). Decreased expression of TCF12 contributes to progression and predicts biochemical recurrence in patients with prostate cancer. *Tumor Biology*, 39(6), 101042831770392. <https://doi.org/10.1177/1010428317703924>
- Cianfrocca, M., & Goldstein, L. J. (2004). Prognostic and Predictive Factors in Early-Stage Breast Cancer. *The Oncologist Breast Cancer*, 9, 606–616. <https://doi.org/10.1634/theoncologist.2008-0207>
- Ciriello, G., Miller, M. L., Aksoy, B. A., Senbabaoglu, Y., Schultz, N., & Sander, C. (2013). Emerging landscape of oncogenic signatures across human cancers. *Nature Genetics*, 45(10), 1127–1133. <https://doi.org/10.1038/ng.2762>
- Cóppola, F., Nader, J., & Aguirre, R. (2005). Metabolismo de los estrógenos endógenos y cáncer de mama. *Revista Médica Del Uruguay*, 21(Revisiones), 15–22.
- Cordes, K. R., Sheehy, N. T., White, M. P., Berry, E. C., Morton, S. U., Muth, A. N., ... Srivastava, D. (2009). miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature*, 460(7256), 705–710. <https://doi.org/10.1038/nature08195>
- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., ... Guigó, R. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Research*, 22(9), 1775–89. <https://doi.org/10.1101/gr.132159.111>

- Dieci, M. V., Orvieto, E., Dominici, M., Conte, P., & Guarneri, V. (2014). Rare Breast Cancer Subtypes: Histological, Molecular, and Clinical Peculiarities. *The Oncologist*, *19*(8), 805–813. <https://doi.org/10.1634/theoncologist.2014-0108>
- Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., ... Gingeras, T. R. (2012). Landscape of transcription in human cells. *Nature*, *489*(7414), 101–8. <https://doi.org/10.1038/nature11233>
- Duan, B., Guo, T., Sun, H., Cai, R., Rui, Q., & Xi, Z. (2017). miR-205 as a biological marker in non-small cell lung cancer. *Biomedicine & Pharmacotherapy*, *91*, 823–830. <https://doi.org/10.1016/j.biopha.2017.04.086>
- Ellsworth, R. E., Decewicz, D. J., Shriver, C. D., & Ellsworth, D. L. (2010). *Breast cancer in the personal genomics era. Current genomics.* <https://doi.org/10.2174/138920210791110951>
- Eroles, P., Bosch, A., Alejandro Pérez-Fidalgo, J., & Lluch, A. (2012). Molecular biology in breast cancer: Intrinsic subtypes and signaling pathways. *Cancer Treatment Reviews*, *38*(6), 698–707. <https://doi.org/10.1016/j.ctrv.2011.11.005>
- Evolution, C., & Cancer, I. N. (2012). NIH Public Access, *481*(7381), 306–313. <https://doi.org/10.1038/nature10762>. CLONAL
- Fan, C., Oh, D. S., Wessels, L., Weigelt, B., Nuyten, D. S. A., Nobel, A. B., ... Perou, C. M. (2006). Concordance among Gene-Expression–Based Predictors for Breast Cancer. *New England Journal of Medicine*, *355*(6), 560–569. <https://doi.org/10.1056/NEJMoa052933>
- Ferlay, J., Soerjomataram I, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., ... Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*, *136*(5), E359-386. <https://doi.org/10.1002/ijc.29210>
- Ford, D., Easton, D. F., Stratton, M., Narod, S., Goldgar, D., Devilee, P., ... Zelada-Hedman, M. (1998). Genetic Heterogeneity and Penetrance Analysis of the BRCA1 and BRCA2 Genes in Breast Cancer Families. *The American Journal of Human Genetics*, *62*(3), 676–689. <https://doi.org/10.1086/301749>
- Foulkes, W. D., Brunet, J.-S., Stefansson, I. M., Straume, O., Chappuis, P. O., Bégin, L. R., ... Akslen, L. A. (2004). The prognostic implication of the basal-like (cyclin E high/p27 low/p53+/glomeruloid-microvascular-proliferation+) phenotype of BRCA1-related breast cancer. *Cancer Research*, *64*(3), 830–5. <https://doi.org/10.1158/0008-5472.CAN-03-2970>
- Gao, J., Li, J., Du, J., & Li, X. (2016). Long non-coding RNA HOTAIR is a marker for hepatocellular carcinoma progression and tumor recurrence. *Oncology Letters*, 1791–1798. <https://doi.org/10.3892/ol.2016.4130>
- Gawlik-Rzemieniewska, N., & Bednarek, I. (2016). The role of NANOG transcriptional factor in the development of malignant phenotype of cancer cells. *Cancer Biology and Therapy*, *17*(1), 1–10. <https://doi.org/10.1080/15384047.2015.1121348>
- Gazon, H., Lemasson, I., Polakowski, N., Cesaire, R., Matsuoka, M., Barbeau, B., ... Peloponese, J.-M. (2012). Human T-Cell Leukemia Virus Type 1 (HTLV-1) bZIP Factor Requires Cellular Transcription Factor JunD To Upregulate HTLV-1 Antisense Transcription from the 3' Long Terminal Repeat. *Journal of Virology*, *86*(17), 9070–

9078. <https://doi.org/10.1128/JVI.00661-12>
- Gerson Cwilich, R., Alban de la Torre, L. F., Villalobos Prieto, A., & Serrano Olvera, J. A. (2012). Características clinicopatológicas, pronóstico e influencia en el tratamiento adyuvante en los grupos de riesgo de recurrencia determinados por el perfil de expresión de 21 genes, Oncotype Dx®, en cáncer de mama temprano. *Gaceta Médica de México*, 148(Artículo original), 117–124. Retrieved from <http://www.medigraphic.com/pdfs/gaceta/gm-2012/gm122a.pdf>
- Gruvberger-Saal, S. K., Cunliffe, H. E., Carr, K. M., & Hedenfalk, I. A. (2006). Microarrays in breast cancer research and clinical practice - The future lies ahead. *Endocrine-Related Cancer*, 13(4), 1017–1031. <https://doi.org/10.1677/erc.1.01246>
- Guo, B., Hui, Q., Zhang, Y., Chang, P., & Tao, K. (2016). MIR 194 is a negative regulator of GEF-H1 pathway in melanoma. *Oncology Reports*, 36(4), 2412–2420. <https://doi.org/10.3892/or.2016.5020>
- Gupta, R. a, Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., ... Chang, H. Y. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, 464(7291), 1071–6. <https://doi.org/10.1038/nature08975>
- Haagensen, C. D. (1978). Lobular neoplasia (so-called lobular carcinoma. *American Cancer Society Cancer*, 42, 737–769.
- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, 144(5), 646–674. <https://doi.org/10.1016/j.cell.2011.02.013>
- Haque, R., Ahmed, S. A., Inzhakova, G., Shi, J., Avila, C., Polikoff, J., ... Press, M. F. (2012). Impact of breast cancer subtypes and treatment on survival: An analysis spanning two decades. *Cancer Epidemiology Biomarkers and Prevention*, 21(10), 1848–1855. <https://doi.org/10.1158/1055-9965.EPI-12-0474>
- Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., ... Hubbard, T. J. (2012). GENCODE: The reference human genome annotation for the ENCODE project. *Genome Research*, 22(9), 1760–1774. <https://doi.org/10.1101/gr.135350.111>
- Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., ... Hubbard, T. J. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Research*, 22(9), 1760–74. <https://doi.org/10.1101/gr.135350.111>
- He, G. Y., Hu, J. L., Zhou, L., Zhu, X. H., Xin, S. N., Zhang, D., ... Liang, L. (2016). The FOXD3/miR-214/MED19 axis suppresses tumour growth and metastasis in human colorectal cancer. *British Journal of Cancer*, (June), 1367–1378. <https://doi.org/10.1038/bjc.2016.362>
- Heitz, F., Harter, P., Lueck, H. J., Fissler-Eckhoff, A., Lorenz-Salehi, F., Scheil-Bertram, S., ... Bois, A. du. (2009). Triple-negative and HER2-overexpressing breast cancers exhibit an elevated risk and an earlier occurrence of cerebral metastases. *European Journal of Cancer*, 45(16), 2792–2798. <https://doi.org/10.1016/j.ejca.2009.06.027>
- Heppner, G. H., & Miller, F. R. (1998). The cellular basis of tumor progression. *International Review of Cytology*, 177, 1–56. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9378615>
- Hou, M., Tang, X., Tian, F., Shi, F., Liu, F., & Gao, G. (2016). AnnoLnc: a web server for

- systematically annotating novel human lncRNAs. *BMC Genomics*, 17(1), 931. <https://doi.org/10.1186/s12864-016-3287-9>
- Howlader, N., Noone, A., Krapcho, M., Miller, D., Bishop, K., Altekruse, S., ... Cronin, K. (2000). SEER Cancer Statistics Review, 1975-2013. *National Cancer Institute*, http://seer.cancer.gov/csr/1975_2013/. Retrieved from http://seer.cancer.gov/csr/1975_2004/results_merged/topic_historical_mort_trends.pdf
- Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M. J., Kenzelmann-Broz, D., ... Rinn, J. L. (2010). A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*, 142(3), 409–419. <https://doi.org/10.1016/j.cell.2010.06.040>
- Huarte, M., & Rinn, J. L. (2010). Large non-coding RNAs: Missing links in cancer? *Human Molecular Genetics*, 19(R2). <https://doi.org/10.1093/hmg/ddq353>
- Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., ... Clamp, M. (2002). The Ensembl genome database project. *Nucleic Acids Research*, 30(1), 38–41. <https://doi.org/10.1093/NAR/30.1.38>
- Hung, T., Wang, Y., Lin, M. F., Koegel, A. K., Kotake, Y., Grant, G. D., ... Chang, H. Y. (2011). Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nature Genetics*, 43(7), 621–629. <https://doi.org/10.1038/ng.848>
- Ignatius, M. S., & Langenau, D. M. (2009). Zebrafish as a model for cancer self-renewal. *Zebrafish*, 6(4), 377–87. <https://doi.org/10.1089/zeb.2009.0610>
- Iyer, M. K., Niknafs, Y. S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., ... Chinnaiyan, A. M. (2015). The landscape of long noncoding RNAs in the human transcriptome. *Nature Genetics*, 47(3), 199–208. <https://doi.org/10.1038/ng.3192>
- Juilland, M., Gonzalez, M., Erdmann, T., Banz, Y., Jevnikar, Z., Hailfinger, S., ... Thome, M. (2016). CARMA1- and MyD88-dependent activation of Jun/ATF-type AP-1 complexes is a hallmark of ABC diffuse large B-cell lymphomas. *Blood*, 127(14), 1780–9. <https://doi.org/10.1182/blood-2015-07-655647>
- Kamstrup, M. R., Gniadecki, R., & Skovgaard, G. L. (2007). Putative cancer stem cells in cutaneous malignancies. *Experimental Dermatology*, 16(4), 297–301. <https://doi.org/10.1111/j.1600-0625.2007.00547.x>
- Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., & Lu, C. (2013). Mutational landscape and significance across 12 major cancer types. *Nature*, 503(7471), 333–339. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Kanitakis, J. (2007). Mammary and extramammary Paget's disease. *Journal of the European Academy of Dermatology and Venereology : JEADV*, 21(5), 581–590. <https://doi.org/10.1111/j.1468-3083.2007.02154.x>
- Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*, 12(6), 996–1006. <https://doi.org/10.1101/gr.229102>. Article published online before print in May 2002
- Kerlikowske, K., Miglioretti, D. L., Ballard-Barbash, R., Weaver, D. L., Buist, D. S. M., Barlow, W. E., ... Carney, P. A. (2003). Prognostic characteristics of breast cancer among postmenopausal hormone users in a screened population. *Journal of Clinical Oncology*, 21(23), 4314–4321. <https://doi.org/10.1200/JCO.2003.05.151>

- Khaitan, D., Dinger, M. E., Mazar, J., Crawford, J., Smith, M. A., Mattick, J. S., & Perera, R. J. (2011). The melanoma-upregulated long noncoding RNA SPRY4-IT1 modulates apoptosis and invasion. *Cancer Research*, *71*(11), 3852–3862. <https://doi.org/10.1158/0008-5472.CAN-10-4460>
- Knaul, F., Lopez-Carrillo, L., Lazcano-Ponce, E., Gómez-Dantés, H., Romieu, I., & Torres, G. (2009). Cáncer de mama : un reto para la sociedad y los sistemas de salud. *Salud Pública de México*, *51*(Suplemento 2), S138–S140.
- Knaul, F. M., Arreola-Ornelas, H., Velázquez, E., Dorantes, J., Méndez, Ó., & Ávila-Burgos, L. (2009). El costo de la atención médica del cáncer mamario: El caso del Instituto Mexicano del Seguro Social. *Salud Publica de Mexico*, *51*(SUPPL.2). <https://doi.org/10.1590/S0036-36342009000800019>
- Knaul, F. M., Nigenda, G., Lozano, R., Arreola-Ornelas, H., Langer, A., & Frenk, J. (2009). Cáncer de mama en México: Una prioridad apremiante. *Salud Publica de Mexico*, *51*(SUPPL.2), 335–344. <https://doi.org/10.1590/S0036-36342009000800026>
- Koressaar, T., & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, *23*(10), 1289–1291. <https://doi.org/10.1093/bioinformatics/btm091>
- L'abbate, A., Tolomeo, D., De Astis, F., Lonoce, A., Cunsolo, C. Lo, Mühlematter, D., ... Storlazzi, C. T. (2015). t(15;21) translocations leading to the concurrent downregulation of RUNX1 and its transcription factor partner genes SIN3A and TCF12 in myeloid disorders. *Molecular Cancer*, *1*, 10–14. <https://doi.org/10.1186/s12943-015-0484-0>
- Labreche, K., Simeonova, I., Kamoun, A., Gleize, V., Chubb, D., Letouzé, E., ... Houlston, R. S. (2015). TCF12 is mutated in anaplastic oligodendroglioma. *Nature Communications*, *6*, 7207. <https://doi.org/10.1038/ncomms8207>
- Lagios, M. D., & Silverstein, M. J. (2001). Sentinel node biopsy for patients with DCIS: a dangerous and unwarranted direction. *Ann Surg Oncol*, *8*(4), 275–277. <https://doi.org/10.1245/aso.2001.8.4.275>
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., ... International Human Genome Sequencing, C. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. <https://doi.org/10.1038/35057062>
- Lange, C. A., & Yee, D. (2008). Progesterone and breast cancer. *Women's Health (London, England)*, *4*(2), 151–162. <https://doi.org/10.2217/17455057.4.2.151>
- Li, H., Gui, C., & Yan, L. (2016). [Knock-down of BCL11A expression in breast cancer cells promotes MDA-MB-231 cell apoptosis]. *Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi = Chinese Journal of Cellular and Molecular Immunology*, *32*(11), 1532–1535. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/27774950>
- Li, J., Han, L., Roebuck, P., Diao, L., Liu, L., Yuan, Y., ... Liang, H. (2015). TANRIC: An interactive open platform to explore the function of lncRNAs in cancer. *Cancer Research*, *75*(18), 3728–3737. <https://doi.org/10.1158/0008-5472.CAN-15-0273>
- Li, Z., Ying, X., Chen, H., Ye, P., Shen, Y., Pan, W., & Zhang, L. (2014). MicroRNA-194 Inhibits the Epithelial-Mesenchymal Transition in Gastric Cancer Cells by Targeting FoxM1. *Digestive Diseases and Sciences*, *59*(9), 2145–2152. <https://doi.org/10.1007/s10620-014-3159-6>
- Life technologies. (2014). Real-time PCR handbook. *Applied Biosystems*, 1–70.

- <https://doi.org/CO320850812>
- Liu, H., Ding, L., Shen, Y., Zhong, F., Wang, Q., & Xu, X. (2016). RBQ3 participates in multiple myeloma cell proliferation, adhesion and chemoresistance. *International Journal of Biological Macromolecules*, *91*, 115–122.
<https://doi.org/10.1016/j.ijbiomac.2016.05.050>
- Lyng, M. B., Laenkholm, A.-V., Pallisgaard, N., & Ditzel, H. J. (2008). Identification of genes for normalization of real-time RT-PCR data in breast carcinomas. *BMC Cancer*, *8*, 20. <https://doi.org/10.1186/1471-2407-8-20>
- Manuscript, A. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*, *490*(7418), 61–70. <https://doi.org/10.1038/nature11412>
- Marchionni, L., Wilson, R., Wolff, A., Marinopoulos, S., Parmigiani, G., Bass, E., & Goodman, S. (2008). Annals of Internal Medicine Systematic Review : Gene Expression Profiling Assays in Early-Stage. *Ann Intern Med*, *148*, 358–369.
- Markham, N. R., & Zuker, M. (2008). UNAFold: Software for nucleic acid folding and hybridization. *Methods in Molecular Biology*, *453*, 3–31. <https://doi.org/10.1007/978-1-60327-429-6-1>
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., ... Lopez, R. (2013). Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Research*, *41*(Web Server issue), 597–600. <https://doi.org/10.1093/nar/gkt376>
- Mercer, T. R., Dinger, M. E., & Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nature Reviews. Genetics*, *10*, 155–159. <https://doi.org/10.1038/nrg2521>
- Merlo, L. M., Pepper, J. W., Reid, B. J., & Maley, C. C. (2006). Cancer as an evolutionary and ecological process. *Nat Rev Cancer*, *6*(12), 924–935.
<https://doi.org/10.1038/nrc2013>
- Metzger-Filho, O., Sun, Z., Viale, G., Price, K. N., Crivellari, D., Snyder, R. D., ... Cardoso, F. (2013). Patterns of recurrence and outcome according to breast cancer subtypes in lymph node-negative disease: Results from international breast cancer study group trials VIII and IX. *Journal of Clinical Oncology*, *31*(25), 3083–3090.
<https://doi.org/10.1200/JCO.2012.46.1574>
- Mori, K., Uchida, T., Fukumura, M., Tamiya, S., Higurashi, M., Sakai, H., ... Shibamura, M. (2016). Linkage of E2F1 transcriptional network and cell proliferation with respiratory chain activity in breast cancer cells. *Cancer Science*, *107*(7), 963–971.
<https://doi.org/10.1111/cas.12953>
- Muscarella, L. A., Guarnieri, V., Coco, M., Belli, S., Parrella, P., Pulcrano, G., ... D'Agostino, L. (2010). Small deletion at the 7q21.2 locus in a CCM family detected by Real-Time Quantitative PCR. *Journal of Biomedicine and Biotechnology*, *2010*(2009), 19622337. <https://doi.org/10.1155/2010/854737>
- Naik, R., Veldore, V. H., & Gopinath, K. S. (2015). Genetics and breast cancer - Oncologists perspectives. *Indian Journal of Surgical Oncology*, *6*(4), 415–419.
<https://doi.org/10.1007/s13193-016-0491-7>
- Narod, S. a., & Rodríguez, A. a. (2011). [Genetic predisposition for breast cancer: BRCA1 and BRCA2 genes]. *Salud Pública de México*, *53*(5), 420–9. <https://doi.org/S0036-36342011000500010> [pii]
- Nicolson, G. L., & Rosenberg, N. L. (1987). Diversification and progression of malignant

- tumors. *BioEssays*, 6(5), 204–208. <https://doi.org/10.1002/bies.950060503>
- Nieminen, A. (2008). *TISSUE MICROARRAY TECHNOLOGY IN BREAST CANCER SUBTYPING*. University of Jyväskylä.
- Nigenda, G., Caballero, M., & González-Robledo, L. M. (2009). Barreras de acceso al diagnóstico temprano del cáncer de mama en el Distrito Federal y en Oaxaca. *Salud Pública de México*, 51. <https://doi.org/10.1590/S0036-36342009000800016>
- Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science*, 194(4260), 23–8. <https://doi.org/10.1126/science.959840>
- Osborne, C. K., Yochmowitz, M. G., Knight, W. a., & McGuire, W. L. (1980). The value of estrogen and progesterone receptors in the treatment of breast cancer. *Cancer*, 46(12 Suppl), 2884–2888. <https://doi.org/3.0.CO;2-U>
- Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J. A., McQuisten, K. A., Almabrazi, H. G., ... Peek, A. S. (2008). IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Research*, 36(Web Server issue), 163–169. <https://doi.org/10.1093/nar/gkn198>
- Pang, H., & Yue, X. (2017). MiR-205 serves as a prognostic factor and suppresses proliferation and invasion by targeting insulin-like growth factor receptor 1 in human cervical cancer. *Tumor Biology*, 39(6), 101042831770130. <https://doi.org/10.1177/1010428317701308>
- Park, S.-S., Park, S.-K., Lim, J.-H., Choi, Y. H., Kim, W.-J., & Moon, S.-K. (2011). Esculetin inhibits cell proliferation through the Ras/ERK1/2 pathway in human colon cancer cells. *Oncology Reports*, 25, 223–230. <https://doi.org/10.3892/or>
- Pérez, J. A., & Soto, S. (2003). Análisis del período 1998-2002 en el Hospital Clínico Regional de Valdivia. *Cuad. Cir.*, (17), 23–29.
- Perou, C. M., Sùrlie, T., Eisen, M. B., Rijn, M. Van De, Jeffrey, S. S., Rees, C. A., ... Grant, S. (2000). Molecular Portraits Breast Cancer, 533(May), 747–752.
- Rahman, M., Deleyrolle, L., Vedam-Mai, V., Azari, H., Abd-El-Barr, M., & Reynolds, B. A. (2011). The cancer stem cell hypothesis: Failures and pitfalls. *Neurosurgery*, 68(2), 531–545. <https://doi.org/10.1227/NEU.0b013e3181ff9eb5>
- Rakha, E. A., & Ellis, I. O. (2011). Modern classification of breast cancer: should we stick with morphology or convert to molecular profile characteristics. *Advances in Anatomic Pathology*, 18(4), 255–267. <https://doi.org/10.1097/PAP.0b013e318220f5d1>
- Rakha, E. A., Reis-Filho, J. S., & Ellis, I. O. (2008). Basal-like breast cancer: A critical review. *Journal of Clinical Oncology*, 26(15), 2568–2581. <https://doi.org/10.1200/JCO.2007.13.1748>
- Ramsay, D. T., Kent, J. C., Hartmann, R. A., & Hartmann, P. E. (2005). Anatomy of the lactating human breast redefined with ultrasound imaging. *Journal of Anatomy*, 206(6), 525–534. <https://doi.org/10.1111/j.1469-7580.2005.00417.x>
- Reis-Filho, J. S., Weigelt, B., Fumagalli, D., & Sotiriou, C. (2010). Molecular profiling: moving away from tumor philately. *Science Transl. Med.*, 2(47), 47ps43. <https://doi.org/10.1126/scitranslmed.3001329>
- Renhan, M. (Weizmann I. of S., Chalifa-Caspi, V. (Weizmann I. of S., & Prilusky, J. (Weizmann I. of S. (1997). GeneCards : integrating information about genes , prote and diseases. *Genetwork*, 13(4), 163.

- Rho, H.-W., Lee, B.-C., Choi, E.-S., Choi, I.-J., Lee, Y.-S., & Goh, S.-H. (2010). Identification of valid reference genes for gene expression studies of human stomach cancer by reverse transcription-qPCR. *BMC Cancer*, *10*, 240. <https://doi.org/10.1186/1471-2407-10-240>
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Bruggmann, S. A., ... Chang, H. Y. (2007). Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell*, *129*(7), 1311–1323. <https://doi.org/10.1016/j.cell.2007.05.022>
- Rodr, D., Bretones, G., Villamor, N., Arango, J. R., Armando, L., Ramsay, A. J., ... Carlos, L. (2017). Mutations in CHD2 cause defective association with active chromatin and show prognostic value in chronic lymphocytic leukemia, *126*(2), 195–203. <https://doi.org/10.1182/blood-2014-10-604959.D.R>
- Rodríguez-Cuevas, S., Macías, C. G., Franceschi, D., & Labastida, S. (2001). Breast carcinoma presents a decade earlier in Mexican women than in women in the United States or European countries. *Cancer*, *91*(4), 863–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11241256>
- Rosai J. (2011). *Rosai and Ackerman's Surgical Pathology*. ed. MOSBY Elsevier. <https://doi.org/10.1016/B978-0-323-06969-4.00001-5>
- Rossella, F., Gabriele, G., Federica, T., Michela, A., Mario, G., Gabriele, P., ... Dapas, B. (2016). The Role of the Transcription Factor E2F1 in Hepatocellular Carcinoma. *Current Drug Delivery*, *13*(999), 1–1. <https://doi.org/10.2174/1567201813666160527141742>
- Rozen, S., & Skaletsky, H. (2000). Primer3 on the WWW for General Users and for Biologist Programmers. *Totowa, NJ. Humana Press*, *132*, 365–386. <https://doi.org/10.1385/1-59259-192-2:365>
- Salameh, A., Fan, X., Choi, B.-K., Zhang, S., Zhang, N., & An, Z. (2016). HER3 and LINC00052 interplay promotes tumor growth in breast cancer. *Oncotarget*, *8*(4), 6526–6539. <https://doi.org/10.18632/oncotarget.14313> salazar1975.pdf. (n.d.).
- Sandhu, R., Parker, J. S., Jones, W. D., Livasy, C. a., & Coleman, W. B. (2010). Microarray-Based Gene Expression Profiling for Molecular Classification of Breast Cancer and Identification of New Targets for Therapy. *Laboratory Medicine*, *41*(6), 364–372. <https://doi.org/10.1309/LMLIK0VIE3CJK0WD>
- Schwartz, G. F., Solin, L. J., Olivotto, I. a, Ernster, V. L., & Pressman, P. (2000). The Consensus Conference on the Treatment of in situ Ductal Carcinoma of the Breast, 22-25 April 1999. *Breast (Edinburgh, Scotland)*, *9*(4), 177–186. <https://doi.org/10.1054/brst.1999.0156>
- Secretaría de Salud México. (2013). México: numeralia de cáncer de mama. *Boletín Epidemiológico*, *30*(26), 1–5. Retrieved from <http://www.epidemiologia.salud.gob.mx/doctos/boletin/2013/semanas/sem26/pdf/edit2613.pdf>
- Shah, S. P., Morin, R. D., Khattra, J., Prentice, L., Pugh, T., Burleigh, A., ... Aparicio, S. (2009). Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution, *461*(October), 809–813. <https://doi.org/10.1038/nature08489>

- Shimm, D. S. (2008). *Harrison's Manual of Oncology. International Journal of Radiation Oncology*Biology*Physics* (Vol. 71). <https://doi.org/10.1016/j.ijrobp.2008.03.023>
- Siddiqui, B., Piyush, A., Faridi, S. H., Ahmad, S. S., & Khan, R. (2016). Trends in Molecular Classification of Breast Carcinoma in a tertiary Health care centre: A 5 year retrospective study. *International Archives of BioMedical and Clinical Research*, 2(2), 69–72. <https://doi.org/10.21276/iabcr.2016.2.2.15>
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., ... Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), 539. <https://doi.org/10.1038/msb.2011.75>
- Silverstein, M. J., Lagios, M. D., Craig, P. H., Waisman, J. R., Lewinsky, B. S., Colburn, W. J., & Poller, D. N. (1996). A prognostic index for ductal carcinoma in situ of the breast. *Cancer*, 77(11), 2267–2274. [https://doi.org/10.1002/\(sici\)1097-0142\(19960601\)77:11<2267::aid-cnrc13>3.0.co;2-v](https://doi.org/10.1002/(sici)1097-0142(19960601)77:11<2267::aid-cnrc13>3.0.co;2-v)
- Singletary, S. E., & Connolly, J. L. (2006). Breast cancer staging: working with the sixth edition of the AJCC Cancer Staging Manual. *CA: A Cancer Journal for Clinicians*, 56(1), 37-47-51. <https://doi.org/10.3322/canjclin.56.1.37>
- Sive, J., Basilico, S., Hannah, R., Kinston, S., Calero-Nieto, F., & Göttgens, B. (2015). Genome-scale definition of the transcriptional programme associated with compromised PU.1 activity in acute myeloid leukaemia. *Leukemia*, 30(May 2015), 14–23. <https://doi.org/10.1038/leu.2015.172>
- Slaby, O., Svoboda, M., Fabian, P., Smerdova, T., Knoflickova, D., Bednarikova, M., ... Vyzula, R. (2008). Altered expression of miR-21, miR-31, miR-143 and miR-145 is related to clinicopathologic features of colorectal cancer. *Oncology*, 72(5–6), 397–402. <https://doi.org/10.1159/000113489>
- Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J., ... Delorenzi, M. (2006). Gene expression profiling in breast cancer: Understanding the molecular basis of histologic grade to improve prognosis. *Journal of the National Cancer Institute*, 98(4), 262–272. <https://doi.org/10.1093/jnci/djj052>
- Soule, H. D., Vazquez, J., Long, A., Albert, S., & Brennan, M. (1973). A Human Cell Line From a Pleural Effusion Derived From a Breast Carcinoma 2. *JNCI: Journal of the National Cancer Institute*, 51(5), 1409–1416. <https://doi.org/10.1093/jnci/51.5.1409>
- Tamimi, R. M., Colditz, G. A., Hazra, A., Baer, H. J., Hankinson, S. E., Rosner, B., ... Collins, L. C. (2012). Traditional breast cancer risk factors in relation to molecular subtypes of breast cancer. *Breast Cancer Research and Treatment*, 131(1), 159–167. <https://doi.org/10.1007/s10549-011-1702-0>
- Tanner, M., Gancberg, D., Di Leo, A., Larsimont, D., Rouas, G., Piccart, M. J., & Isola, J. (2000). Chromogenic in situ hybridization: a practical alternative for fluorescence in situ hybridization to detect HER-2/neu oncogene amplification in archival breast cancer samples. *The American Journal of Pathology*, 157(5), 1467–72. [https://doi.org/10.1016/S0002-9440\(10\)64785-2](https://doi.org/10.1016/S0002-9440(10)64785-2)
- Terroso, G., Romão, V. C., & Rúa-figueroa, I. (2014). Characterizing Genetic Transitions of Copy Number Alterations and Allelic Imbalances in Oral Tongue Carcinoma Metastasis, 1–35. <https://doi.org/10.1002/acr>

- The essenTials of life science ReseaRch Globally DeliveReD™. (n.d.). Retrieved from <https://www.atcc.org/~media/PDFs/Cancer and Normal cell lines tables/Breast cancer and normal cell lines.ashx>
- Thornton, B., & Basu, C. (2011). Real-time PCR (qPCR) primer design using free online software. *Biochemistry and Molecular Biology Education*, 39(2), 145–154. <https://doi.org/10.1002/bmb.20461>
- Tomczak, K., Czerwińska, P., & Wiznerowicz, M. (2015). The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Wspolczesna Onkologia*, 1A, A68–A77. <https://doi.org/10.5114/wo.2014.47136>
- Tricarico, C., Pinzani, P., Bianchi, S., Paglierani, M., Distante, V., Pazzagli, M., ... Orlando, C. (2002). Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies• 1. *Analytical Biochemistry*, 309, 293–300. [https://doi.org/10.1016/S0003-2697\(02\)00311-1](https://doi.org/10.1016/S0003-2697(02)00311-1)
- Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., ... Prasanth, K. V. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Molecular Cell*, 39(6), 925–938. <https://doi.org/10.1016/j.molcel.2010.08.011>
- Tsang, D. P. F., Wu, W. K. K., Kang, W., Lee, Y. Y., Wu, F., Yu, Z., ... Cheng, A. S. L. (2016). Yin Yang 1-mediated epigenetic silencing of tumour-suppressive microRNAs activates nuclear factor-κB in hepatocellular carcinoma. *Journal of Pathology*, 238(5), 651–664. <https://doi.org/10.1002/path.4688>
- Tsutsui, S., Ohno, S., Murakami, S., Kataoka, A., Kinoshita, J., & Hachitanda, Y. (2003). Prognostic significance of the coexpression of p53 protein and c-erbB2 in breast cancer. *American Journal of Surgery*, 185(2), 165–167. [https://doi.org/10.1016/S0002-9610\(02\)01203-5](https://doi.org/10.1016/S0002-9610(02)01203-5)
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Research*, 40(15), 1–12. <https://doi.org/10.1093/nar/gks596>
- van 't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. a. M., Mao, M., ... Friend, S. H. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Letters To Nature*, 415(345), 530–536. <https://doi.org/10.1038/415530a>
- Volders, P. J., Helsen, K., Wang, X., Menten, B., Martens, L., Gevaert, K., ... Mestdagh, P. (2013). LNCipedia: A database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Research*, 41(D1), 246–251. <https://doi.org/10.1093/nar/gks915>
- Volders, P. J., Verheggen, K., Menschaert, G., Vandepoele, K., Martens, L., Vandesompele, J., & Mestdagh, P. (2015). An update on LNCipedia: A database for annotated human lncRNA sequences. *Nucleic Acids Research*, 43(D1), D174–D180. <https://doi.org/10.1093/nar/gku1060>
- Wang, F., Li, L., Chen, Z., Zhu, M., & Gu, Y. (2016). MicroRNA-214 acts as a potential oncogene in breast cancer by targeting the PTEN-PI3K/Akt signaling pathway. *International Journal of Molecular Medicine*, 37(5), 1421–1428. <https://doi.org/10.3892/ijmm.2016.2518>

- Wang, J., Yang, T., Xu, G., Liu, H., Ren, C., Xie, W., & Wang, M. (2016). Cyclin-dependent kinase 2 promotes tumor proliferation and induces radio resistance in glioblastoma. *Translational Oncology*, 9(6), 548–556. <https://doi.org/10.1016/j.tranon.2016.08.007>
- Wang, J., Zhang, X., Wang, L., Dong, Z., Du, L., Yang, Y., ... Wang, C. (2015). Downregulation of urinary cell-free microRNA-214 as a diagnostic and prognostic biomarker in bladder cancer. *Journal of Surgical Oncology*, 111(8), 992–999. <https://doi.org/10.1002/jso.23937>
- Wang, P., Ning, S., Zhang, Y., Li, R., Ye, J., Zhao, Z., ... Li, X. (2015). Identification of lncRNA-associated competing triplets reveals global patterns and prognostic markers for cancer. *Nucleic Acids Research*, 43(7), 3478–89. <https://doi.org/10.1093/nar/gkv233>
- Wang, Y., Waters, J., Leung, M. L., Unruh, A., Roh, W., Shi, X., ... Navin, N. E. (2014). Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature*, 512(7513), 1–15. <https://doi.org/10.1038/nature13600>
- Weigelt, B., Mackay, A., A'hern, R., Natrajan, R., Tan, D. S. P., Dowsett, M., ... Reis-Filho, J. S. (2010). Breast cancer molecular profiling with single sample predictors: A retrospective analysis. *The Lancet Oncology*, 11(4), 339–349. [https://doi.org/10.1016/S1470-2045\(10\)70008-5](https://doi.org/10.1016/S1470-2045(10)70008-5)
- Wu, P., Li, X., Wu, Y., Hu, W., Wang, Y., Gao, L., ... Zheng, W. (2014). Overexpression of Raf-1 in basal-like carcinoma of the breast: Correlation with clinicopathology and prognosis. *Wspolczesna Onkologia*, 18(6), 391–395. <https://doi.org/10.5114/wo.2014.47037>
- Xu, B., Niu, X., Zhang, X., Tao, J., Wu, D., Wang, Z., ... Wang, X. (2011). MiR-143 decreases prostate cancer cells proliferation and migration and enhances their sensitivity to docetaxel through suppression of KRAS. *Molecular and Cellular Biochemistry*, 350(1–2), 207–213. <https://doi.org/10.1007/s11010-010-0700-6>
- Xu, H., Yan, Y., Deb, S., Rangasamy, D., Germann, M., Malaterre, J., ... Ramsay, R. G. (2014). Cohesin Rad21 Mediates Loss of Heterozygosity and Is Upregulated via Wnt Promoting Transcriptional Dysregulation in Gastrointestinal Tumors. *Cell Reports*, 9(5), 1781–1789. <https://doi.org/10.1016/j.celrep.2014.10.059>
- Yang, F., Liu, Y. H., Dong, S. Y., Ma, R. M., Bhandari, A., Zhang, X. H., & Wang, O. C. (2016). A novel long non-coding RNA FGF14-AS2 is correlated with progression and prognosis in breast cancer. *Biochemical and Biophysical Research Communications*, 470(3), 479–483. <https://doi.org/10.1016/j.bbrc.2016.01.147>
- Yang, X. R., Chang-Claude, J., Goode, E. L., Couch, F. J., Nevanlinna, H., Milne, R. L., ... Garcia-Closas, M. (2011). Associations of breast cancer risk factors with tumor subtypes: A pooled analysis from the breast cancer association consortium studies. *Journal of the National Cancer Institute*, 103(3), 250–263. <https://doi.org/10.1093/jnci/djq526>
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (n.d.). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. Retrieved from <https://bmcbioinformatics.biomedcentral.com/track/pdf/10.1186/1471-2105-13-134?site=bmcbioinformatics.biomedcentral.com>
- Yersal, O., & Barutca, S. (2014). Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World Journal of Clinical Oncology*, 5(3), 412–24.

- <https://doi.org/10.5306/wjco.v5.i3.412>
- Yin, J., Zhang, F., Tao, H., Ma, X., Su, G., Xie, X., ... Yin, B. (2016). BCL11A expression in acute phase chronic myeloid leukemia. *Leukemia Research*, 47, 88–92. <https://doi.org/10.1016/j.leukres.2016.05.018>
- Yu, A. Q., Ding, Y., Li, C. L., Yang, Y., Yan, S. R., & Li, D. S. (2016). TALEN-induced disruption of Nanog expression results in reduced proliferation, invasiveness and migration, increased chemosensitivity and reversal of EMT in HepG2 cells. *Oncology Reports*, 35(3), 1657–1663. <https://doi.org/10.3892/or.2015.4483>
- Yu, X., Luo, A., Liu, Y., Wang, S., Li, Y., Shi, W., ... Schumacker, P. (2015). MiR-214 increases the sensitivity of breast cancer cells to tamoxifen and fulvestrant through inhibition of autophagy. *Molecular Cancer*, 14(1), 208. <https://doi.org/10.1186/s12943-015-0480-4>
- Yun, J., Song, S. H., Kang, J. Y., Park, J., Kim, H. P., Han, S. W., & Kim, T. Y. (2016). Reduced cohesin destabilizes high-level gene amplification by disrupting pre-replication complex bindings in human cancers with chromosomal instability. *Nucleic Acids Research*, 44(2), 558–572. <https://doi.org/10.1093/nar/gkv933>
- Zardavas, D., Irrthum, A., Swanton, C., & Piccart, M. (2015). Clinical management of breast cancer heterogeneity. *Nat Rev Clin Oncol*, 12(7), 381–394. <https://doi.org/10.1038/nrclinonc.2015.73>
- Zerbini, L. F., De Vasconcellos, J. F., Czibere, A., Wang, Y., Pancez, J. D., Gu, X., ... Libermann, T. A. (2011). JunD-mediated repression of GADD45 α and γ regulates escape from cell death in prostate cancer. *Cell Cycle*, 10(15), 2583–2591. <https://doi.org/10.4161/cc.10.15.16057>
- Zhang, J., Zhu, N., & Chen, X. (2015). A novel long noncoding RNA LINC01133 is upregulated in lung squamous cell cancer and predicts survival. *Tumor Biology*, 36(10), 7465–7471. <https://doi.org/10.1007/s13277-015-3460-9>
- Zhao, C., Chen, W., Dai, S., Zhang, X., Ban, N., Fan, S., ... Ren, J. (2015). Expression and clinical role of RBQ3 in gliomas. *Journal of the Neurological Sciences*, 359(1–2), 177–184. <https://doi.org/10.1016/j.jns.2015.10.058>
- Zhao, T., Wang, H., Ma, H., Wang, H., Chen, B., & Deng, Y. (2016). Starvation after Cobalt-60 γ -Ray Radiation Enhances Metastasis in U251 Glioma Cells by Regulating the Transcription Factor SP1. *International Journal of Molecular Sciences*, 17(4), 386. <https://doi.org/10.3390/ijms17040386>
- Zhou, L. L., Dong, J. L., Huang, G., Sun, Z. L., & Wu, J. (2017). MicroRNA-143 inhibits cell growth by targeting ERK5 and MAP3K7 in breast cancer. *Brazilian Journal of Medical and Biological Research*, 50(8), 1–8. <https://doi.org/10.1590/1414-431x20175891>
- Zhu, S., Zhao, L., Li, Y., Hou, P., Yao, R., Tan, J., ... Zhang, Y. (2016). Suppression of RAD21 Induces Senescence of MDA-MB-231 Human Breast Cancer Cells Through RB1 Pathway Activation Via c-Myc Downregulation. *Journal of Cellular Biochemistry*, 117(6), 1359–1369. <https://doi.org/10.1002/jcb.25426>
- Zhuo, Z., & Yu, H. (2017). miR-205 inhibits cell growth by targeting AKT-mTOR signaling in progesterone-resistant endometrial cancer Ishikawa cells. *Oncotarget*, 8(17), 28042–28051. <https://doi.org/10.18632/oncotarget.15886>

- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 31(13), 3406–3415. <https://doi.org/10.1093/nar/gkg595>
- SEER Training Modules, Module Name. U. S. National Institutes of Health, National Cancer Institute. Day Month Year (23, Mars 2017) <<https://training.seer.cancer.gov/>
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3 - new capabilities and interfaces. *Nucleic Acids Research* 40(15):e115
- Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3 *Bioinformatics* 23(10):1289-91

ANEXOS

Estandarización de los lncRNAs para q-PCR.

Se validaron ambos cebadores y el amplicón para los genes del lncRNA MIR205HG y los endógenos ACTB y TBP. Se realizaron curvas disociación donde se varió la proporción de ambos cebadores (forward y reverse). La especificidad de los cebadores de todos los lncRNAs se corroboró además por electroforesis en gel de agarosa. Las concentraciones probadas fueron: 300nM, 500nM y 800nM, y sus posibles combinaciones con un total de 9 condiciones mostradas en la **Tabla 20**, en el apartado de estandarización.

Estandarización del lncRNA MIR205HG.

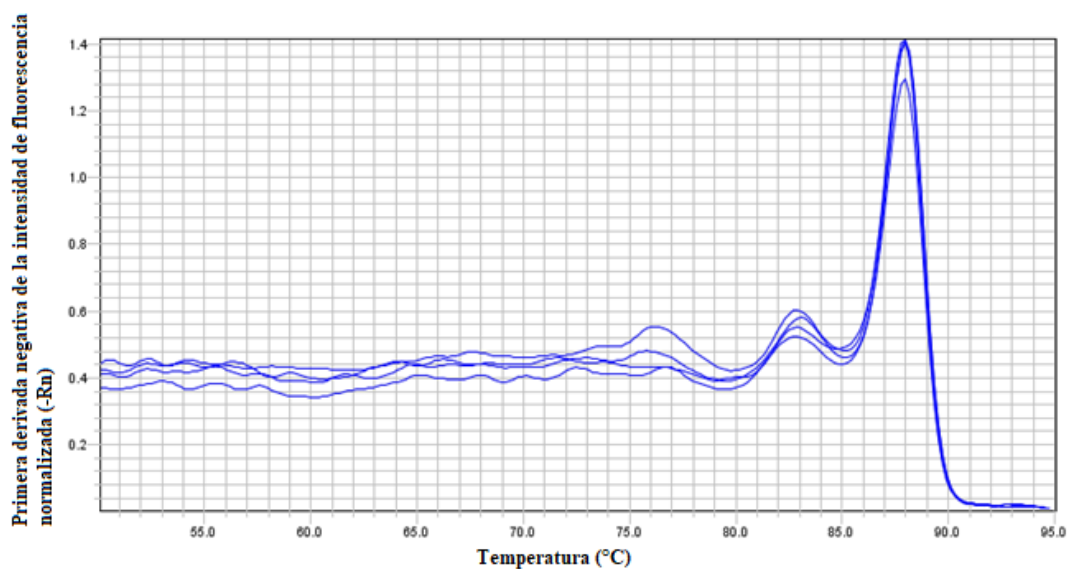


Figura 25. Curva de disociación de MIR205HG. Condición seleccionada [FP/RP]= 500/300.

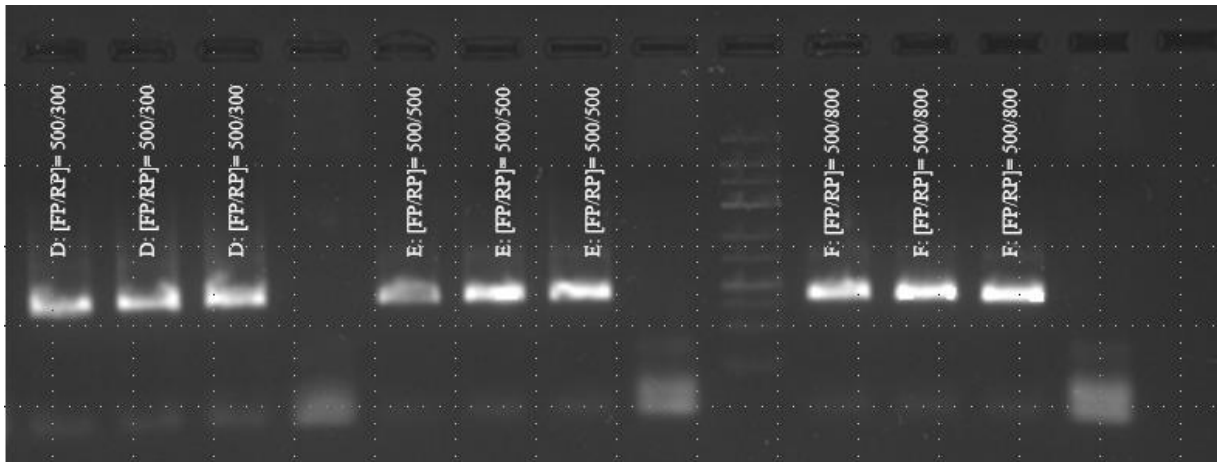


Figura 26. Banda esperada del lncRNA MIR205HG (D: [FP/RP= 500/300]) por electroforesis en gel de agarosa.

Las **Figuras 25** y **26** son imágenes representativas de los criterios bajo los cuales se eligieron las condiciones óptimas de amplificación de los genes ACTB y TBP.

Curva de eficiencia relativa.

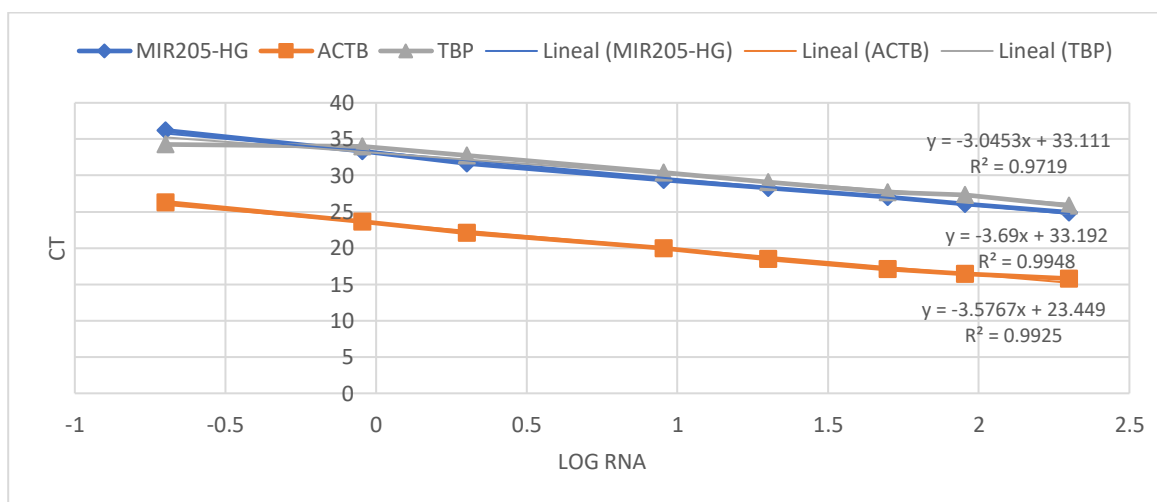


Figura 27. Eficiencia relativa en la amplificación de los genes ACTB, TBP y MIR205HG.

Donde se observa que la eficiencia de amplificación de acuerdo con la ecuación para cada gen es de: MIR205HG (112.99%), TBP (86.6%) y ACTB (90.3%). Donde la eficiencia es calculada por medio de la ecuación: $E = 10^{\left(-\frac{1}{pendiente}\right)} - 1$. De acuerdo con el Real Time PCR Handbook (Life technologies, 2014) no es necesario aplicar un factor de corrección para poder analizar los resultados de q-PCR.