



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**Maestría y Doctorado en Ciencias Bioquímicas**

**Mecanismos de desregulación del metabolismo energético  
asociados a cáncer de mama**

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
Maestro en Ciencias

PRESENTA:

Erandi Aketzalli Serrano Carbajal

TUTOR PRINCIPAL

Dr. Enrique Hernández Lemus  
[Instituto Nacional de Medicina Genómica](#)

MIEMBROS DEL COMITÉ TUTOR

Dra. María Elena Ibarra Rubio  
[Facultad de Química](#)

Dra. Lorena Aguilar Arnal  
[Instituto de Investigaciones Biomédicas](#)

Ciudad de México. Agosto, 2021



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.

PMDCB/1066/2019

**Serrano Carbajal Erandi Aketzalli**

Estudiante de Maestría en Ciencias Bioquímicas

**P r e s e n t e**

Los miembros del Subcomité Académico en reunión ordinaria del día 8 de abril del presente año, conocieron su solicitud de asignación de **JURADO DE EXAMEN** para optar por el grado de **MAESTRO EN CIENCIAS**, con la réplica de la tesis "**Mecanismos de desregulación del metabolismo energético asociados a cáncer de mama**", dirigida por el/la Dr(a). Hernández Lemus Enrique.

De su análisis se acordó nombrar el siguiente jurado integrado por los doctores:

PRESIDENTE	Morett Sánchez Juan Enrique
VOCAL	Saavedra lira Emma Cecilia
VOCAL	Pedraza Chaverri José
VOCAL	Hidalgo Miranda Alfredo
SECRETARIO	Uribe Carvajal Salvador

Sin otro particular por el momento, aprovecho la ocasión para enviarle un cordial saludo.

**Atentamente**

**"POR MI RAZA, HABLARÁ EL ESPÍRITU"**

Cd. Universitaria, Cd. Mx., a 8 de abril de 2019.

**COORDINADORA**



**Dra. ANA BRÍGIDA CLORINDA ARIAS ÁLVAREZ**

*A mis padres*

# Índice general

Índice de figuras

Índice de tablas

Abreviaturas

Resumen

<b>1. Introducción</b>	<b>1</b>
1.1. Cáncer de mama	1
1.1.1. Características del cáncer	2
1.1.2. Reprogramación metabólica.	4
1.1.2.1. Alteración genómica como mecanismo de reprogramación metabólica.	6
1.1.3. Subtipos en cáncer de mama	8
1.2. Biología de sistemas.	10
1.2.1. Análisis de vías.	11
<b>2. Planteamiento del problema</b>	<b>14</b>
<b>3. Objetivos</b>	<b>15</b>
3.1. Objetivo general	15
3.1.1. Objetivos particulares	15
<b>4. Materiales y Métodos</b>	<b>16</b>
4.1. Datos.	16
4.2. Preprocesamiento.	16
4.2.1. Normalización.	17
4.2.2. Corrección por efecto de lote	17
4.2.3. Subtipificación.	18
4.3. Análisis de expresión diferencial	18
4.3.1. Comparación de genes diferencialmente expresados entre subtipos.	19
4.3.2. Relación entre genes diferencialmente expresados y vías metabólicas.	20
4.4. Análisis de desregulación de vías metabólicas.	20
4.4.1. Agrupamiento jerárquico.	22

<b>5. Resultados</b>	<b>24</b>
5.1. Genes diferencialmente expresados	24
5.1.1. Relación genes diferencialmente expresados y vías metabólicas.	29
5.2. Desregulación de vías metabólicas y agrupamiento jerárquico	36
5.3. Metabolismo de purinas como blanco potencial en todos los subtipos de cáncer de mama	46
<b>6. Discusión</b>	<b>48</b>
<b>7. Conclusión</b>	<b>57</b>
<b>8. Perspectivas</b>	<b>58</b>
<b>9. Referencias</b>	<b>60</b>
<b>10. Anexos</b>	<b>76</b>
Anexo A	76
Anexo B	79

# Índice de figuras

<b>Figura 1.</b> Número estimado de casos nuevos de cáncer en 2018.	1
<b>Figura 2.</b> Número estimado de muertes debidas a cáncer en 2018.	2
<b>Figura 3.</b> Características distintivas del cáncer.	3
<b>Figura 4.</b> Diagrama de flujo.	23
<b>Figura 5.</b> Diagrama de Venn. Se muestran el número de genes sobreexpresados en los diferentes subtipos tumorales.	25
<b>Figura 6.</b> Diagrama de Venn. Se muestra el número de genes subexpresados en los diferentes subtipos tumorales.	27
<b>Figura 7.</b> Red que relaciona los genes sobreexpresados comunes para los cuatro subtipos tumorales con las vías metabólicas a las que se asocian.	30
<b>Figura 8.</b> Red que relaciona los genes subexpresados comunes para los cuatro subtipos tumorales con las vías metabólicas a las que se asocian.	35
<b>Figura 9.</b> Mapa de calor. Representa el PDS del análisis de Pathifier. Para vías de metabolismo (KEGG).	37
<b>Figura 10.</b> Mapa de calor. Representa el PDS del análisis de Pathifier para el conjunto de genes denominado “Metabolic Pathways” (KEGG) que engloba los genes involucrados en todas las vías metabólicas.	39
<b>Figura 11.</b> Histograma de frecuencias. Muestra la frecuencia de los valores PDS del análisis de Pathifier para el conjunto de genes denominado “Metabolic Pathways” (KEGG).	40
<b>Figura 12.</b> Gráfico de la curva principal utilizada para calcular el valor de PDS del conjunto de genes denominado “Metabolic Pathways”.	41
<b>Figura 13.</b> Mapa de calor. Representa el PDS del análisis de Pathifier para las vías de metabolismo de ácido linoléico, metabolismo de ácido alfa-linoléico, metabolismo de ácido araquidónico y biosíntesis de esteroides (KEGG).	43
<b>Figura 14.</b> Mapa de calor. Representa el PDS del análisis de Pathifier para las muestras con subtipo luminal B.	45

# Índice de tablas

<b>Tabla 1.</b> Matriz de expresión génica.	21
<b>Tabla 2.</b> Matriz de conjunto de datos empleados para el análisis de desregulación de vías.	22
<b>Tabla 3.</b> Genes sobreexpresados únicos en cada subtipo tumoral.	26
<b>Tabla 4.</b> Genes subexpresados únicos en cada subtipo tumoral.	28
<b>Tabla 5.</b> Número total de genes diferencialmente expresados para cada subtipo tumoral.	29
<b>Tabla 6.</b> Número de genes diferencialmente expresados únicos de cada subtipo tumoral.	29
<b>Tabla 7.</b> Valores de expresión diferencial de los genes sobreexpresados comunes para los cuatro subtipos tumorales.	32
<b>Tabla 8.</b> Función de los 10 genes que se encuentran sobreexpresados en los 4 subtipos tumorales.	33
<b>Tabla 9.</b> Genes diferencialmente expresados con inhibidores o activadores aprobados por la FDA para regular el metabolismo de las purinas.	47

## Abreviaturas

ACACB	Acetil-coenzima A carboxilasa 2
ACOX1	Acil-Coenzima A Oxidasa 1
ADCY4	Adenilato ciclasa 4
ADCY5	Adenilato ciclasa 5
ADH1	Alcohol deshidrogenasa 1
AK5	Adenilato cinasa 5
AKR1B1	Aldo-ceto reductasa familia 1 miembro B1
ALDH2	Aldehído deshidrogenasa 2
ALG1	Chitobiosildifosfodolicol beta-manosiltransferasa
ALOX12	Araquidonato 12-lipoxigenasa, tipo 12S
ARSyN	Identificación y Remoción de Ruido Sistémico
ATP	Adenosín trifosfato
ATP6V1H	ATPasa 6 protón tipo V subunidad H
ATP6V1B1	ATPasa 6 protón tipo V subunidad B1
CHPT1	Colina fosfotransferasa 1
CHST15	Carbohidrato sulfotransferasa 15
COX17	Citocromo C oxidasa chaperona de cobre
CPT1C	Carnitina Palmitoiltransferasa 1C
CYP21A2	Citocromo P450 familia 21 subfamilia A miembro 2
DHDH	Dihidrodiol deshidrogenasa
DNA	Ácido desoxirribonucleico
EGFR	Receptor del factor de crecimiento epidérmico
ENO1	Enolasa 1
ENO2	Enolasa 2
ENPP4	Ectonucleótido pirofosfatasa/fosfodiesterasa 4
ER	Receptor de estrógeno
ERBB2	Receptor 2 del factor de crecimiento epidérmico humano
EZH1	Potenciador de Zeste 1 del complejo represivo Polycomb subunidad 2
EZH2	Potenciador de Zeste 2 del complejo represivo Polycomb subunidad 2
FASN	Ácido graso sintasa
FDA	Administración de Fármacos y Alimentos de Estados Unidos
GAGE	Análisis de Vías para Enriquecimiento de Conjunto de Genes Generalmente Aplicable
GALM	Galactosa mutarotasa
GALNT5	Polipéptido N-acetilgalactosaminiltransferasa 5
GALT	Galactosa-1-fosfato uridililtransferasa
GDC	Datos Genómicos Comunes
GGH	Gama-glutamil hidrolasa

GSTP1	Glutación S-transferasa pi 1
Her2	Receptor 2 del factor de crecimiento epidérmico humano
HK2	Hexocinasa 2
HSD17B6	Hidroxiesteroide 17-beta deshidrogenasa 6
IL4I1	Inducida por [interleucina 4] 1
INPP4B	Inositol polifosfato 4-fosfatasa tipo 2
JAK-STAT	Janus cinasa - Transductor de señales y activador de la transcripción
KEGG	Enciclopedia de Genes y Genomas de Kyoto
LCLAT1	Lisocardiopina aciltransferasa 1
LCT	Precursor de lactasa-florizina hidrolasa
LDHA	Lactato deshidrogenasa A
LDHB	Lactato deshidrogenasa B
limma	Modelos Lineales para Microarreglos y RNA-seq
LPCAT4	Lisofosfatidilcolina aciltransferasa 4
MAT1A	Metionina adenosiltransferasa 1A
MBOAT1	Dominio que contiene la membrana unida a O-aciltransferasa 1
MBOAT2	Dominio que contiene la membrana unida a O-aciltransferasa 2
MTHFD1L	Metilentetrahidrofolato deshidrogenasa [dependiente de NADP] 1
NADH	Nicotinamida adenina dinucleótido
NADPH	Nicotinamida adenina dinucleótido fosfato
NAGS	N-acetilglutamato sintasa, mitocondrial
NCI	Instituto Nacional de Cáncer
NDST3	N-deacetilasa y N-sulfotransferasa 3
NME1-NME2	Nucleósido difosfato cinasa A y B
NMNAT3	Nicotinamida nucleótido adenililtransferasa 3
NPR1	Receptor de péptido natriurético 1
NPR2	Receptor de péptido natriurético 2
OMS	Organización Mundial de la Salud
OXSM	3-oxoacil-[proteína acarreadora de acilo] sintasa, mitocondrial
PAFAH1B3	Factor activador de plaquetas acetilhidrolasa 1b subunidad catalítica 3
PAH	Fenilalanina hidroxilasa
PAK1	Proteína cinasa de serina/treonina
pbcmc	Clasificación Molecular basada en Permutaciones
PDE7A	Fosfodiesterasa 7A
PDS	Puntuación de Desregulación de Vía
PFK	Fosfofructocinasa
PIGQ	Fosfatidilinositol N-acetilglucosaminiltransferasa clase Q
PIGX	Fosfatidilinositol N-acetilglucosaminiltransferasa clase X
PLA2G10	Fosfolipasa secretora A2 grupo 10
POLR2K	RNA polimerasa II, I y III subunidad K
PSAT1	Fosfoserina aminotransferasa 1
PR	Receptor de progesterona
RBKS	Ribocinasa
RNA	Ácido ribonucleico
RNA-seq	Secuenciación de RNA
RRM2	Ribonucleótido reductasa reguladora subunidad M2

SDS	Serina deshidratasa
ST3GAL1	Supresor de tumorigenicidad 3 beta-galactósido alfa-2,3-sialiltransferasa 1
ST6GALNAC5	Supresor de tumorigenicidad 6 N-acetilgalactosamina alfa-2,6-sialiltransferasa 5
SUV39H2	Supresor de abigarramiento 3-9 homólogo 2
TCGA	Atlas del Genoma del Cáncer
TMM	Media Recortada de Valores M
UDP	Uridin difosfato
UGCG	Uridin difosfato-Glucosa Ceramida glucosiltransferasa
UGT8	Uridin difosfato glicosiltransferasa 8
UPP2	Uridina fosforilasa 2

## Resumen

El cáncer de mama es una enfermedad heterogénea caracterizada por una desregulación de los flujos metabólicos que conllevan a una reprogramación de la célula. Las aberraciones genómicas son uno de los factores que repercuten directa o indirectamente en el nivel de expresión de los genes y contribuyen en los cambios de la actividad metabólica oncogénica, induciendo el proceso carcinogénico. Entre estas aberraciones destacan las mutaciones puntuales, las alteraciones en el número de copias de los genes y alteraciones epigenómicas.

Por lo anterior, en este estudio se analizaron los perfiles de expresión de pacientes con cáncer de mama, tomando en cuenta el subtipo tumoral al que pertenecen (luminal A, luminal B, her2 y basal).

Primero se realizó un análisis de expresión diferencial de los genes relacionados con el metabolismo, el cual reveló que en total de los cuatro subtipos, 204 genes se encuentran sobreexpresados, mientras que 287 genes se encuentran subexpresados. También se encontraron 10 genes sobreexpresados y 79 subexpresados comunes a los cuatro subtipos y fueron relacionados a vías de metabolismo de lípidos, carbono y nucleótidos.

Además, se identificaron los genes que sólo se expresaron en un sólo subtipo tumoral, siendo el subtipo tumoral basal el que presentó la mayor cantidad de genes diferencialmente expresados (98 genes).

Por otro lado, se realizó un análisis de desregulación de vías metabólicas que engloba la información de los perfiles de expresión de los genes pertenecientes a una vía en particular. El análisis calcula la variación de las muestras tumorales en comparación con los controles asignando una puntuación de desregulación (PDS, pathway deregulation score, por sus siglas en inglés) para cada una de las muestras, en cada una de las vías metabólicas a evaluar. Se observó que los subtipos tumorales de cáncer de mama presentan niveles de desregulación diferentes cuando se evalúa todo el metabolismo como una gran vía, y se destaca que el subtipo luminal B es el que presenta mayor nivel de desregulación.

En conclusión, se identificaron las coincidencias y disimilitudes entre los perfiles de expresión diferencial en los genes de metabolismo para cada uno de los cuatro subtipos tumorales. Se asociaron algunos de estos genes a posibles mecanismos bioquímicos involucrados en el proceso carcinogénico. Asimismo, se encontraron diferentes patrones de desregulación en las vías metabólicas que se equiparán con las características moleculares de los subtipos.

Por último, se relacionaron las vías más desreguladas en cada muestra y los genes diferencialmente expresados implicados en las vías seleccionadas, para considerarlos como posibles blancos farmacológicos.

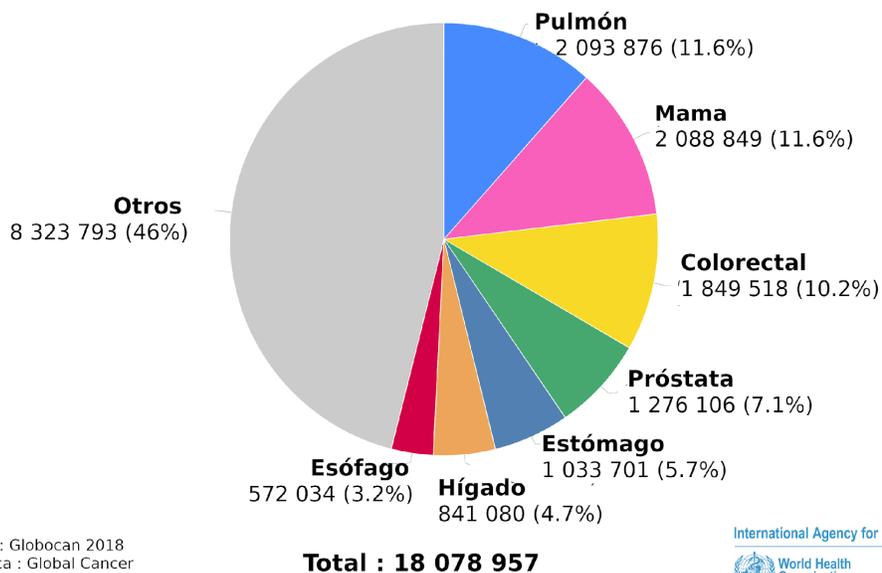
Algunos de los resultados de la investigación desarrollada en esta tesis están publicados en el siguiente artículo científico: Serrano-Carbajal, E. A., Espinal-Enríquez, J., & Hernández-Lemus, E. (2020). Targeting metabolic deregulation landscapes in breast cancer subtypes. *Frontiers in Oncology*, 10, 97. Para conveniencia de los lectores el artículo completo se encuentra reproducido al final de esta tesis como Anexo B.

# 1. Introducción

## 1.1. Cáncer de mama

El cáncer de mama es el tipo de cáncer más frecuente en mujeres a nivel mundial. De acuerdo con la Organización Mundial de la Salud (OMS) la prevalencia de cáncer de mama ha aumentado en los últimos años, representando el 16% de todos los tipos de cáncer femeninos (OMS, 2017). Tiene una incidencia del 11.6% (Figura 1) y un porcentaje de mortalidad del 6.6% del total de tipos de cáncer descritos en GLOBOCAN, 2018 (Figura 2), siendo el de mayor mortalidad en mujeres (Bray *et al.* 2018).

**Número de casos nuevos estimados en 2018, en todo el mundo, ambos sexos**



Fuente de datos : Globocan 2018  
Producción gráfica : Global Cancer  
Observatorio (<http://gco.iarc.fr>)

International Agency for Research on Cancer  
World Health Organization

Figura 1. Número estimado de casos nuevos de cáncer en 2018. (Modificado de Globocan 2018)

### Número de muertes estimadas en 2018, en todo el mundo, ambos sexos

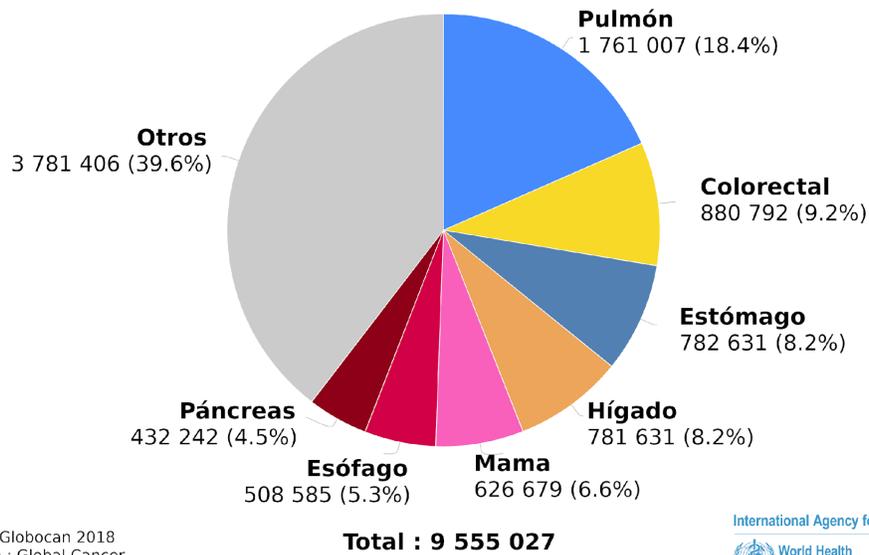


Figura 2. Número estimado de muertes debidas a cáncer en 2018.(Modificado de Globocan 2018)

Por lo anterior, es necesario ampliar el conocimiento acerca de los mecanismos carcinogénicos involucrados en esta enfermedad, con el propósito de encontrar nuevos tratamientos terapéuticos.

#### 1.1.1. Características del cáncer

El cáncer es una enfermedad compleja que se caracteriza por la heterogeneidad de las células tumorales, las cuales aceleran su crecimiento y comienzan a proliferar sin control aparente (Cooper, 2000). Esto es causado por alteraciones en el DNA genómico, desregulación en la expresión de genes, desorden hormonal, alteraciones metabólicas y en vías de señalización (Espinal-Enríquez *et al.* 2017).

Originalmente se describieron seis características distintivas dentro de los diferentes tipos de cáncer, incluido el de mama (Hanahan y Weinberg, 2000), las cuales incluyen: autosuficiencia en señales de crecimiento, insensibilidad a las señales de crecimiento, evasión de la muerte celular programada, inducción de la angiogénesis, potencial replicativo ilimitado, e invasión del tejido y metástasis (Hanahan y Weinberg, 2000). En 2011, se sumaron dos características distintivas involucradas en la patogénesis de diferentes tipos de cáncer: la evasión de la respuesta inmune y la desregulación en los flujos energéticos (Figura 3) (Hanahan y Weinberg, 2011).

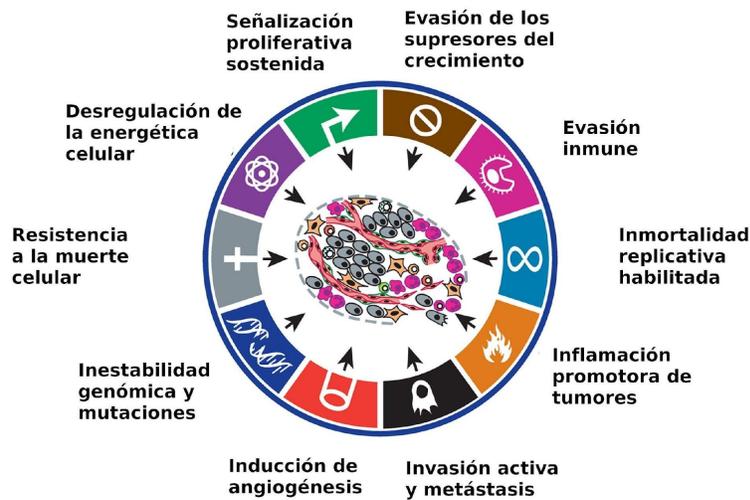


Figura 3. Características distintivas del cáncer. (Adaptado de Hanahan y Weinberg, 2011).

Específicamente en esta última característica, las células tumorales reprograman su metabolismo generando mayor síntesis de macromoléculas y modificando la producción de energía (DeBerardinis *et al.* 2016). Todo esto resulta en una mayor tasa de crecimiento, proliferación y sobrevivencia de las células cancerosas.

### **1.1.2. Reprogramación metabólica.**

Las células normales llevan a cabo, para su funcionamiento, reacciones bioquímicas que convierten compuestos de carbono en macromoléculas como hidratos de carbono, lípidos y aminoácidos. Además, mantienen flujos metabólicos que aseguran el suministro constante de la energía libre de Gibbs mediante la hidrólisis del adenosín trifosfato (ATP) y la oxidación de equivalentes de óxido-reducción (Nicotinamida adenina dinucleótido [NADH], dihidronicotinamida-adenina dinucleótido fosfato [NADPH]) (Nielsen, 2003).

En cambio, en células neoplásicas, esta actividad metabólica se altera a consecuencia de mutaciones cancerígenas y cambios en la expresión de oncogenes y genes supresores de tumores, entre otros factores (DeBerardinis y Chandel, 2016). Los tumores aumentan la biosíntesis de macromoléculas que permiten su crecimiento y proliferación en etapas tempranas del desarrollo neoplásico (DeBerardinis y Chandel, 2016). Posteriormente, cuando los recursos

energéticos comienzan a escasear, la célula oncogénica entra en un estado de estrés oxidante; para su sobrevivencia, se propicia un aumento en el flujo de vías metabólicas que generan moléculas de óxido-reducción (NADH, NADPH, etc.) y contrarrestan el efecto de dicho estrés contribuyendo al equilibrio químico del sistema (Acharya *et al.*, 2010). Estas modificaciones en el flujo metabólico conducen a lo que se conoce como reprogramación metabólica (Yoshida, 2015; DeBerardinis y Chandel, 2016).

El ejemplo más conocido de reprogramación metabólica es el efecto Warburg, el cual describe que en células tumorales el flujo de la glucólisis incrementa en comparación con células normales, generando lactato incluso en presencia de oxígeno. La glucólisis genera ATP con menor eficiencia, pero más rápidamente que la fosforilación oxidante, lo que promueve una proliferación más rápida de los tumores. Además, la alta tasa glicolítica favorece el crecimiento tumoral mediante la biosíntesis de lípidos, nucleótidos, NADPH y aminoácidos (Lunt y Vander Heiden, 2011; Zhang y Yang, 2013).

Sin embargo, el efecto Warburg no es el único mecanismo que describe la reprogramación metabólica, se han descrito otros procesos involucrados en cambios de los flujos metabólicos en células cancerosas. Estos cambios interfieren y desregulan diversas vías del metabolismo de glucosa, lípidos y nucleótidos debido, entre otros factores, a alteraciones genómicas (Hirschey *et al.*, 2015; Pavlova y Thompson, 2016).

### **1.1.2.1. Alteración genómica como mecanismo de reprogramación metabólica.**

Una de las características distintivas del cáncer son las aberraciones a nivel genómico que repercuten directa o indirectamente en el nivel de expresión de los genes (Shao *et al.* 2019; Baylin y Ohm, 2016) y contribuyen en los cambios de la actividad metabólica oncogénica (Eason y Sadanandam, 2016; Jia *et al.* 2019). Estas aberraciones pueden ser mutaciones tumorigénicas, translocaciones y variación en el número de copias, que se define por la pérdida o ganancia de fragmentos de DNA genómico (amplificaciones, deleciones, etc.)(Kumaran *et al.* 2017), así como, alteraciones en el RNA mensajero, en la metilación de DNA y en la modificación de histonas (cambios epigenómicos) (Eifert y Powers, 2012; Chakravarthi *et al.*, 2016).

Algunos estudios han evidenciado que una expresión anormal de genes que codifican enzimas glicolíticas están relacionados con diversos tipos de cáncer (Sheng *et al.*, 2012; Wolf *et al.*, 2011; Le *et al.*, 2010). Por ejemplo, se ha reportado que los genes codificantes para la enzima lactato deshidrogenasa, *LDHA* y *LDHB*, presentan una sobreexpresión que contribuye con el efecto Warburg (Ždravlević *et al.*, 2017) y se asocia a crecimiento tumoral, metástasis, proliferación celular, migración e invasión en neoplasias de mama, pancreáticas, gástricas, nasofaríngeas y de endometrio entre otras (Mishra y Banerjee, 2019; Schlichtholz

*et al.*, 2005; Su *et al.*, 2017; Jiang *et al.*, 2015; Giatromanolaki *et al.*, 2006; Dennison *et al.*, 2013).

Por otro lado, se ha descrito que eventos de represión transcripcional provocados por un aumento en la metilación de histonas pueden producir una pérdida de la función o disminución en la expresión de los genes (Baylin y Ohm, 2016). El gen GSTP1 codifica para la enzima glutatión S-transferasa y se ha encontrado hipermetilado en cáncer de próstata (Gonzalzo *et al.*, 2004). La enzima glutatión S-transferasa está involucrada en la desintoxicación del metabolismo celular, así como en la regulación de la muerte celular (Schnekenburger *et al.*, 2014). Es por ello que una disminución de la expresión de GSTP1 debido a la represión transcripcional de este gen puede aumentar la sensibilidad a toxinas metabólicas y ambientales, de este modo se promueven mutaciones que podrían desencadenar el desarrollo del cáncer (Coughlin y Hall, 2002; Schnekenburger *et al.*, 2014).

Debido a la heterogeneidad tumoral, estas alteraciones varían entre tipos de cáncer, e incluso, en un mismo tejido tumoral pueden ser diferentes. Específicamente en cáncer de mama, la expresión de genes y su contribución con la reprogramación metabólica cambian dependiendo de las características moleculares de cada tumor (Sims *et al.* 2007), es por ello que se desarrolló un clasificación de subtipos de cáncer de mama, basada en el perfil transcriptómico de estas células (Parker *et al.*, 2009).

### 1.1.3. Subtipos en cáncer de mama

Como ya se ha mencionado, el cáncer de mama es una enfermedad heterogénea con diferentes características clínicas, biológicas y moleculares (Simpson *et al.* 2005). Dadas las diferentes características moleculares presentes en el cáncer de mama, se ha clasificado a esta enfermedad en cuatro subtipos intrínsecos: Luminal A, Luminal B, HER2-enriquecido y Basal (Goldhirsch *et al.*, 2011).

A continuación se describen los subtipos:

Luminal A. Es el subtipo más común, representa el 40% del total de tumores mamarios. En general, estas células cancerosas manifiestan un aumento de receptores de estrógeno (ER) y progesterona (PR), y de genes que regulan estos receptores. Por otro lado, mantienen una actividad proliferativa baja y son negativas para el *factor de crecimiento epidérmico 2* o *ERBB2* (HER2-). Los pacientes con cáncer luminal A tienen el mejor pronóstico (Hu *et al.*, 2006), las menores tasas de recurrencia, y una buena respuesta al tratamiento hormonal (Laible *et al.*, 2019)

Luminal B. Aproximadamente el 20% de los tumores de cáncer de mama pertenecen a este subtipo. Tiene una variabilidad mayor en la expresión de receptores de estrógeno y genes relacionados al receptor HER2, en relación con

el subtipo luminal A. Existe un mayor riesgo de recaídas asociadas a este subtipo respecto al luminal A (Sorlie *et al.*, 2003).

HER2 enriquecido (HER2 +). Se define por la sobreexpresión del receptor HER2 (derivado de su nombre en inglés, *human epidermal growth factor receptor 2*), involucrado en la regulación de la fisiología, crecimiento y proliferación celular (Burstein, 2005). Estos tumores suelen ser negativos a receptores de estrógeno y progesterona, y tiene una peor prognosis en comparación con los subtipos luminales (Burstein, 2005). La sobreexpresión del oncogen HER2 ocurre en el 25-30% de los subtipos de cáncer de mama y se ha asociado con una mala respuesta a tratamientos hormonales (Flowers y Thompson, 2009; Ross y Flecher, 1998).

Basal. Tiene patrones de expresión similares a los del epitelio mamario basal. Presenta subexpresión de los receptores a estrógeno, progesterona y HER2. Este tipo de tumor se ha asociado a inestabilidad genómica, alta expresión de genes proliferativos y grados histológicos más altos (Bayraktar y Glück, 2013).

## 1.2. Biología de sistemas.

La biología de sistemas es un área de investigación que integra diferentes disciplinas conectando componentes moleculares dentro de una escala biológica sencilla para lograr un entendimiento global de los procesos biológicos incluyendo interacciones de los elementos, tanto internos como externos, que influyen en el desarrollo del proceso (Tavassoly et al., 2018). Este enfoque pretende comprender integralmente el funcionamiento de los sistemas biológicos y profundizar en el entendimiento de cómo sus interacciones internas y con otros sistemas conllevan a la aparición de nuevas propiedades.

En ese sentido, la biología de sistemas busca desarrollar nuevas aproximaciones para integrar la gran cantidad de datos obtenidos por las nuevas tecnologías de alto rendimiento, por ejemplo, la técnica de RNA-seq que nos permite secuenciar fragmentos de RNA que codifican para genes y de esta manera obtener perfiles de expresión génica en un fenotipo específico (Kitano, 2000; Wang *et al.* 2009). Como resultado de la simulación, al poner a funcionar los modelos matemáticos con los que se representa al proceso, se obtiene una serie de predicciones del estado del proceso biológico que corresponden a los resultados experimentales esperados, lo que conlleva a la generación de nuevas hipótesis. Así, en lugar de concentrarnos en las propiedades fisicoquímicas y biológicas por separado o limitarnos al estudio de un número de moléculas, la biología de sistemas nos permite realizar un

análisis cuantitativo de todos los componentes de un sistema biológico que interactúan entre sí.

Un caso particular de esta aproximación se describe en el trabajo realizado por Wu y colaboradores (2017). Ellos realizaron un análisis sistémico de datos de pacientes con cáncer mamario obtenidos del Atlas del Genoma del Cáncer (TCGA, por sus siglas en inglés) e identificaron 136 genes diferencialmente expresados entre los cuatro subtipos tumorales. Posteriormente, analizaron los patrones de coexpresión de estos genes y evaluaron alteraciones en 6 vías biológicas asociadas (vía de señalización de JAK-STAT, carcinoma de células basales, regulación mediada por inflamación, cáncer de pulmón de células no pequeñas, sinapsis glutamatérgica y esclerosis lateral amiotrófica) basándose en un análisis de enriquecimiento funcional de vías (Wu *et al.*, 2017).

### **1.2.1. Análisis de vías.**

El análisis de enriquecimiento funcional de vías o análisis de vías, tiene como objetivo principal buscar genes candidatos que contribuyan, por ejemplo, a determinar cómo se desarrolla el cáncer. No solo eso, este tipo de análisis aporta un panorama global de los cambios que ocurren en cierto fenómeno, en este caso

cómo podrían modificarse los flujos en las vías metabólicas (García-Campos et al., 2015).

Una vía metabólica es un conjunto de reacciones químicas que se llevan a cabo en la célula para mantener la homeostasis (Schilling *et al.*, 1999). Si se desregulan estas vías pueden verse alterados otros procesos biológicos, como sucede en el cáncer (Hirschey *et al.*, 2015; Wu et al., 2010). Para evaluar dicha desregulación, es posible emplear herramientas computacionales que nos permitan estudiar a gran escala todas las rutas metabólicas (García-Campos et al., 2015).

En este contexto, Pathifier (Drier et al., 2013) es un algoritmo de análisis de vías que integra datos de tecnologías de alto rendimiento (datos genómicos, transcriptómicos, proteómicos etc.) de cada muestra de estudio. Este análisis cuantifica un valor de desregulación en una vía metabólica para cada muestra de manera específica. Livshits y colaboradores (2015) estudiaron la heterogeneidad del cáncer de mama mediante Pathifier, basándose en la clasificación molecular antes descrita, y encontraron nueve tipos de tumores con distintos perfiles de desregulación agrupados en siete conjuntos de vías (de señalización; de sistema inmune; relacionadas a cáncer, ciclo celular y reparación de DNA; señalización/inmune; relacionadas con la proteína cinasa A; relacionadas con el receptor del factor de crecimiento epidérmico (EGFR) y; relacionadas con el gen PAK1 (proteína cinasa de serina/treonina)). Ellos concluyen que sus análisis indican que en el cáncer de mama es más probable que se obtengan firmas

pronósticas confiables tratando por separado los diferentes subgrupos de pacientes (Livshits *et al.*, 2015).

Derivado de lo anterior, el presente estudio plantea un análisis específico para vías de metabolismo en pacientes con cáncer de mama, que nos permitan asociar diferentes perfiles de expresión génica con la desregulación metabólica.

## 2. Planteamiento del problema

La desregulación metabólica es una característica distintiva del cáncer. Los patrones que se alteran en las vías metabólicas derivan en una exacerbada síntesis de macromoléculas, incremento en la proliferación celular y resistencia al tratamiento vía la alteración del procesamiento de fármacos. Además la heterogeneidad molecular confiere una mayor dificultad para el tratamiento de esta enfermedad. Esta amplia variación en el metabolismo molecular del cáncer de mama constituye, simultáneamente, un recurso de pronóstico y retos terapéuticos, así como una herramienta para nuevas intervenciones. En este trabajo se investigó el panorama de la desregulación metabólica en los subtipos moleculares del cáncer de mama para 735 muestras de cáncer Luminal A, Luminal B, Her2 y Basal, así como 113 de tejido sano adyacente obtenidas del portal de datos genómicos comunes (Genomic Data Commons, GDC) del Instituto Nacional de Cáncer (NCI, por sus siglas en inglés). Esta aproximación visualiza a las firmas regulatorias que están detrás de las características metabólicas específicas de cada subtipo. Por medio de un algoritmo basado en una sola muestra, se busca determinar la desregulación de cada vía metabólica para cada muestra y asociarla al perfil de expresión diferencial de los genes implicados.

## 3. Objetivos

### 3.1. Objetivo general

Identificar **patrones de desregulación** en **vías metabólicas** a nivel **transcripcional** en muestras de pacientes con **cáncer de mama** mediante herramientas computacionales.

#### 3.1.1. Objetivos particulares

-Identificar los **genes** que presentan **expresión diferencial** en cáncer de mama por subtipo tumoral.

-Asociar los **genes diferencialmente expresados** con las **vías metabólicas** involucradas

-Evaluar la **desregulación** de las **vías metabólicas** por subtipo tumoral de cáncer de mama comparando muestras de tejido tumoral contra muestras de tejido sano .

## **4. Materiales y Métodos**

### **4.1. Datos.**

Se descargaron del portal de datos Genomic Data Commons (GDC, <https://gdc.cancer.gov/>) del Instituto Nacional de Cáncer (NCI, por sus siglas en inglés), los datos de expresión de 1,200 genes obtenidos experimentalmente mediante la técnica de RNA-Seq. Estos genes se clasifican como genes asociados a metabolismo que codifican enzimas, de acuerdo con la base de datos la Enciclopedia de Genes y Genomas de Kyoto (KEGG, por sus siglas en inglés) (Kanehisa *et al.*, 2019).

Se analizaron 735 muestras de tejido mamario tumoral primario y 113 de tejido sano adyacente.

### **4.2. Preprocesamiento.**

Se realizaron controles de calidad a los datos de expresión para corregir sesgos ocasionados por diferencia de longitudes entre las lecturas, y/o el contenido de guanina/citosina, así como sesgos propios de la técnica de secuenciación.

Posteriormente se transformaron los datos de expresión para obtener una distribución normal de los datos y con ello se creó una matriz de expresión.

#### **4.2.1. Normalización.**

Se realizó normalización de los datos para remover efectos sistémicos y reducir sesgos. Se empleó el método estadístico de *Media Recortada de Valores M* (TMM, por sus siglas en inglés), el cual supone que la mayoría de los genes en las muestras no se encuentra diferencialmente expresado y considera iguales los niveles generales de expresión de genes entre las muestras (Robinson y Oshlack, 2010). TMM estima la proporción de RNA usando una media recortada ponderada de las proporciones de expresión en logaritmo base 2 (Robinson y Oshlack, 2010).

#### **4.2.2. Corrección por efecto de lote**

La técnica de RNA.-seq presenta variaciones debidas a efectos del lote (tipo de secuenciador, laboratorio, fecha de ejecución, entre otros) que impiden observar únicamente la variabilidad de la muestra debida al fenómeno de estudio. Para reducir este sesgo se empleó el método estadístico de identificación y remoción de ruido sistémico (ARSyN, por sus siglas inglés) (Nueda *et al.*, 2012), que se basa

en la combinación del análisis de varianza junto con el análisis de componentes principales para analizar un conjunto de datos ómicos multifactoriales.

#### **4.2.3. Subtipificación.**

Para la subtipificación de las muestras tumorales, se empleó el paquete de R para clasificación molecular basada en permutaciones (pbcmc, por sus siglas en inglés) (Fresno *et al.*, 2016), el cual evalúa la incertidumbre en los clasificadores de expresión de genes (e.g. PAM50).

### **4.3. Análisis de expresión diferencial**

Consiste en determinar si existe diferencia significativa de la expresión de los genes de las muestras de tejido tumoral en comparación con las muestras de tejido sano.

Para el análisis se empleó el paquete de R-Bioconductor, *Modelos Lineales para Microarreglos y RNA-seq* (limma) (Smyth, 2005), basado en un ajuste de modelos lineales generalizados (Ritchie *et al.*, 2015). Los genes que se consideran diferencialmente expresados son aquellos que presentan lo siguiente:

-Un log fold change mayor a 1. El log fold change es la proporción de cambio en el nivel de expresión entre condiciones experimentales en logaritmo base 2.

-Un estadístico B mayor a 5. Indica la consistencia estadística en los contrastes entre grupos de niveles de expresión.

El análisis de expresión diferencial se llevó a cabo para cada subtipo tumoral, en comparación con las muestras control, por separado.

#### **4.3.1. Comparación de genes diferencialmente expresados entre subtipos.**

Para identificar coincidencias entre los perfiles de expresión de los subtipos tumorales, se identificaron los genes que presentaron expresión diferencial en todos los subtipos tumorales.

Por otro lado, para determinar las diferencias entre dichos perfiles de expresión se clasificaron los genes que sólo presentan expresión diferencial en un subtipo tumoral y no en los demás. Así, se obtuvieron los genes diferencialmente expresados únicos para cada subtipo.

Estas comparaciones se hicieron para genes sobreexpresados y genes subexpresados.

#### **4.3.2. Relación entre genes diferencialmente expresados y vías metabólicas.**

Se analizaron los genes con expresión diferencial relacionados a metabolismo. A partir de estos genes se realizó una búsqueda manual del conjunto de bases de datos de KEGG, para asociar los genes sobreexpresados o subexpresados con las vías en las que se encuentran involucrados .

#### **4.4. Análisis de desregulación de vías metabólicas.**

Para el cálculo de los niveles de desregulación de las vías metabólicas en cada muestra se utilizó el algoritmo Pathifier (Drier *et al.*, 2013), el cual integra los datos de expresión de los genes involucrados en una sola vía metabólica, en un único valor de desregulación para cada muestra. Este algoritmo asigna una puntuación de 0 a 1, en donde los valores cercanos a 0 son calculados para muestras con niveles de expresión similares a las muestras control, mientras que muestras con valores mayores difieren en los niveles de expresión en comparación a este mismo grupo. Pathifier calcula el nivel de desregulación de una vía metabólica en

una muestra tumoral midiendo la desviación de dicha muestra con respecto al comportamiento normal (Drier *et al.*, 2013).

Pathifier emplea una matriz de la expresión de los genes involucrados en las vías a estudiar, un ejemplo de cómo está construida la matriz se puede observar en la tabla 1.

Tabla 1. Matriz de expresión génica. En las columnas se observan las muestras, y en las filas los genes con sus respectivos niveles de expresión para cada muestra.

Genes	Sano 1	Sano 2	Tumor 1	Tumor 2	Tumor 3
Normales	1	1	0	0	0
ADH1	16.775063	12.810858	10.306848	6.542678	10.596081
AKR1B1	11.913423	11.934814	10.986564	11.53844	11.147704

Además, se descargó un conjunto de datos que relaciona las vías metabólicas a estudiar con los genes presentes en cada vía (Tabla 2). Dichas vías con sus respectivos genes se obtuvieron de la colección de bases de datos KEGG (Kanehisa *et al.* 2019). Las vías que se seleccionaron fueron las que esta base de datos nombra como vías de metabolismo. El conjunto de datos se descargó del paquete de R, GAGE (del inglés, *Generally Applicable Gene-set Enrichment for Pathway Analysis*), para el organismo *Homo sapiens* (Luo *et al.*, 2009).

El análisis de vías se llevó a cabo con el paquete Pathifier (Drier *et al.*, 2013) descargado de Bioconductor.

Tabla 2. Ejemplo de la matriz de conjunto de datos empleados para el análisis de desregulación de vías

Glucólisis/ gluconeogénesis	ADH1	GALM	ENO1	HK2	...
Metabolismo de galactosa	GALM	AKR1B1	GALT	LCT	...
Metabolismo de ácidos grasos	CPT1C	OXSM	ACOX1	FASN	...

#### 4.4.1. Agrupamiento jerárquico.

Para conocer las similitudes y diferencias entre las muestras de estudio, se realizó un agrupamiento jerárquico no supervisado, mediante el paquete de R `gplots` (R core team, 2018), el cual calcula la distancia euclidiana de los valores de PDS y usa el método de agrupamiento aglomerativo no supervisado de Ward (Murtagh & Legendre, 2014). Este análisis da una jerarquía de grupos donde se minimiza la varianza de cada grupo.

A continuación se presenta en la figura 4 el diagrama de flujo a seguir en nuestra metodología:

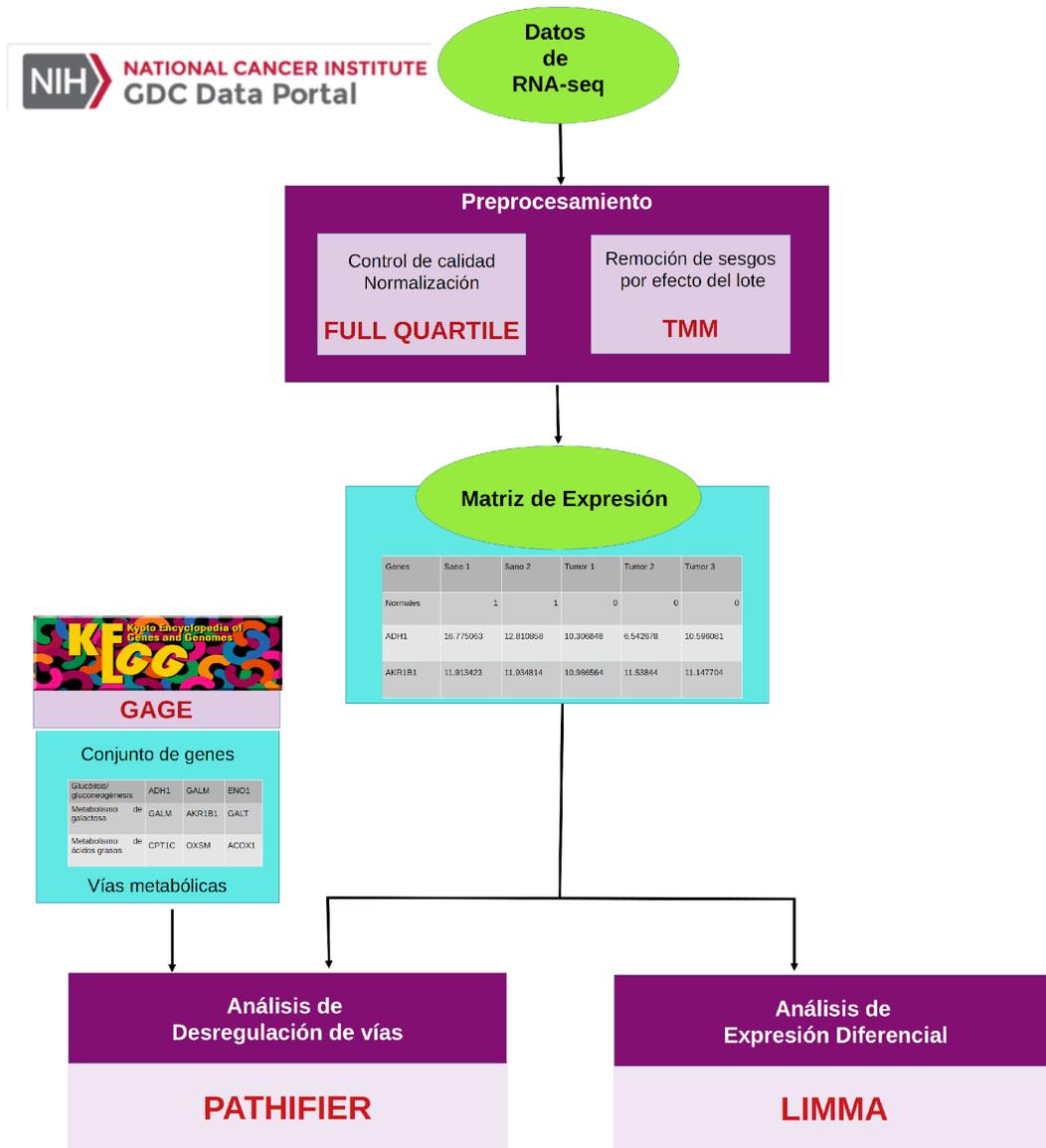


Figura 4. Diagrama de flujo.

## 5. Resultados

### 5.1. Genes diferencialmente expresados

El análisis de expresión diferencial nos permite comparar la variación que existe entre el fenotipo tumoral y las muestras control. Este análisis se realizó para cada subtipo tumoral. Se encontraron en total 204 genes sobreexpresados de los cuatro subtipos tumorales (suma total de genes en la figura 5), de los cuales 10 genes se encuentran sobreexpresados en los cuatro subtipos (centro de la figura 5).

En la figura 5 se muestra el número de genes sobreexpresados que se compartieron por subtipo tumoral. Luminal A compartió más de la mitad de los genes sobreexpresados con luminal B (47 genes comunes entre luminal A y luminal B), mientras que con el subtipo basal, sólo compartió 15 genes sobreexpresados (Figura 5). El subtipo basal compartió el mayor número de genes sobreexpresados con Her2 (en total 49 genes) pero Her2 compartió más (50) con luminal B.

Además, se observó que el subtipo luminal A presentó 8 genes sobreexpresados únicos de dicho subtipo (en total luminal A presentó 67 genes sobreexpresados

totales). Por otro lado, se vio que el subtipo con mayor número de genes sobreexpresados únicos fue el subtipo basal con 46 (en total el subtipo basal presentó 101 genes sobreexpresados totales). Los subtipos luminal B y her2 presentaron 18 y 21 genes sobreexpresados únicos, respectivamente (en total luminal B presentó 100 sobreexpresados y el subtipo her2 108).

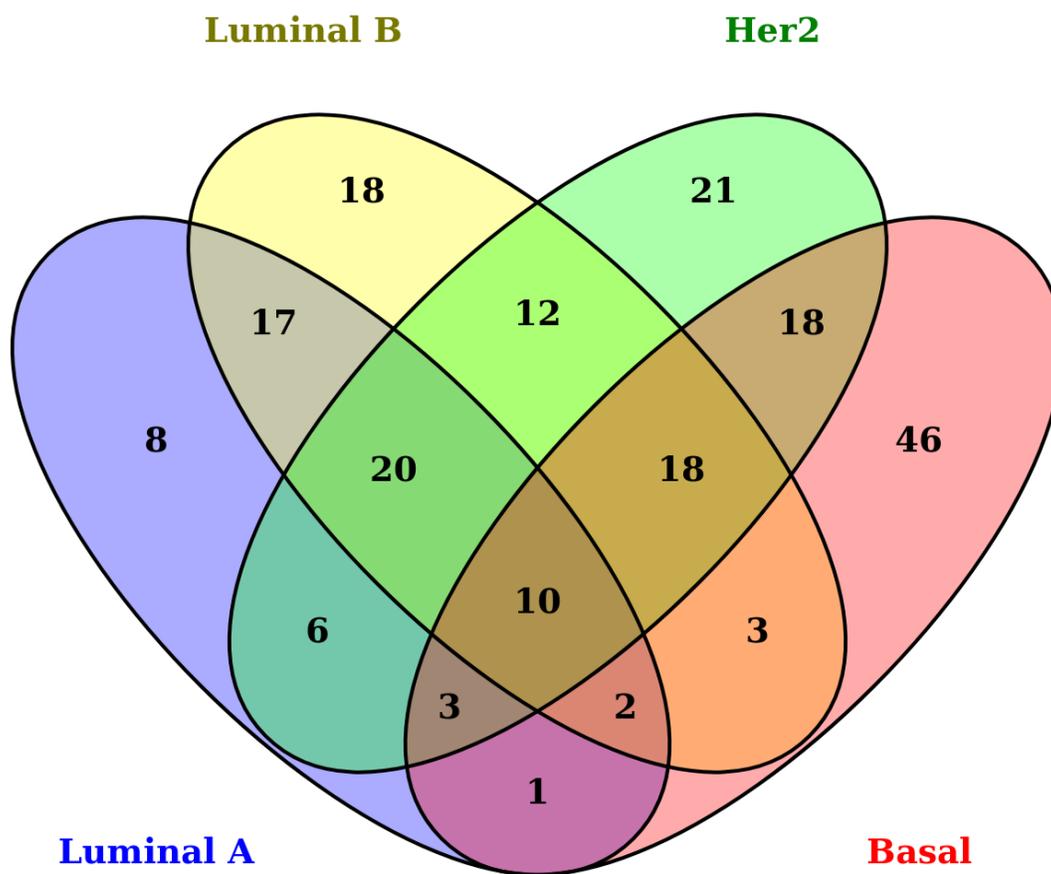


Figura 5. Diagrama de Venn. Se muestran el número de genes sobreexpresados en los diferentes subtipos tumorales.

En la tabla 3 se presentan los 10 genes con mayor expresión diferencial de cada subtipo tumoral

Tabla 3. Genes sobreexpresados únicos en cada subtipo tumoral. Se muestran solo los 10 genes sobreexpresados con mayor expresión diferencial. El subtipo luminal A solo presenta 8 genes sobreexpresados únicos de este subtipo.

Luminal A	Luminal B	Her2	Basal
CHST15	NME1-NME2	ATP6V1B1	GGH
CYP21A2	ENO2	EZH1	MTHFD1L
GALNT5	ATP6V1H	NMNAT3	PSAT1
INPP4B	PAH	CHPT1	PDE7A
PIGQ	ST3GAL1	MAT1A	SUV39H2
PLA2G10	COX17	ALOX12	NDST3
RBKS	POLR2K	NAGS	DHDH
UGCG	ALG1	UPP2	UGT8
	MBOAT1	LPCAT4	ST6GALNAC5
	PIGX	ENPP4	PFK

Por otro lado, el total de los genes subexpresados fue de 287, de los cuales 79 se encuentran subexpresados en todos los subtipos. Los genes subexpresados (figura 6) para luminal A fueron 145 de los cuales 135 también se encontraron subexpresados también en luminal B, siendo luminal B, el subtipo con más genes subexpresados, en total 190. Her2 presentó 179 genes subexpresados de los cuales compartió la mayor parte con luminal B, en total 135, únicos entre los diferentes subtipos tumorales. Basal mostró 46 genes sobreexpresados y 52 genes subexpresados; mientras que Her2 presentó 18 genes subexpresados y 21

sobreexpresados únicos para este subtipo. Luminal A y B tuvieron 9 y 24 genes subexpresados diferentes de los otros subtipos tumorales, respectivamente. En el apéndice se muestran los niveles de expresión diferencial de todos los subtipos tumorales.

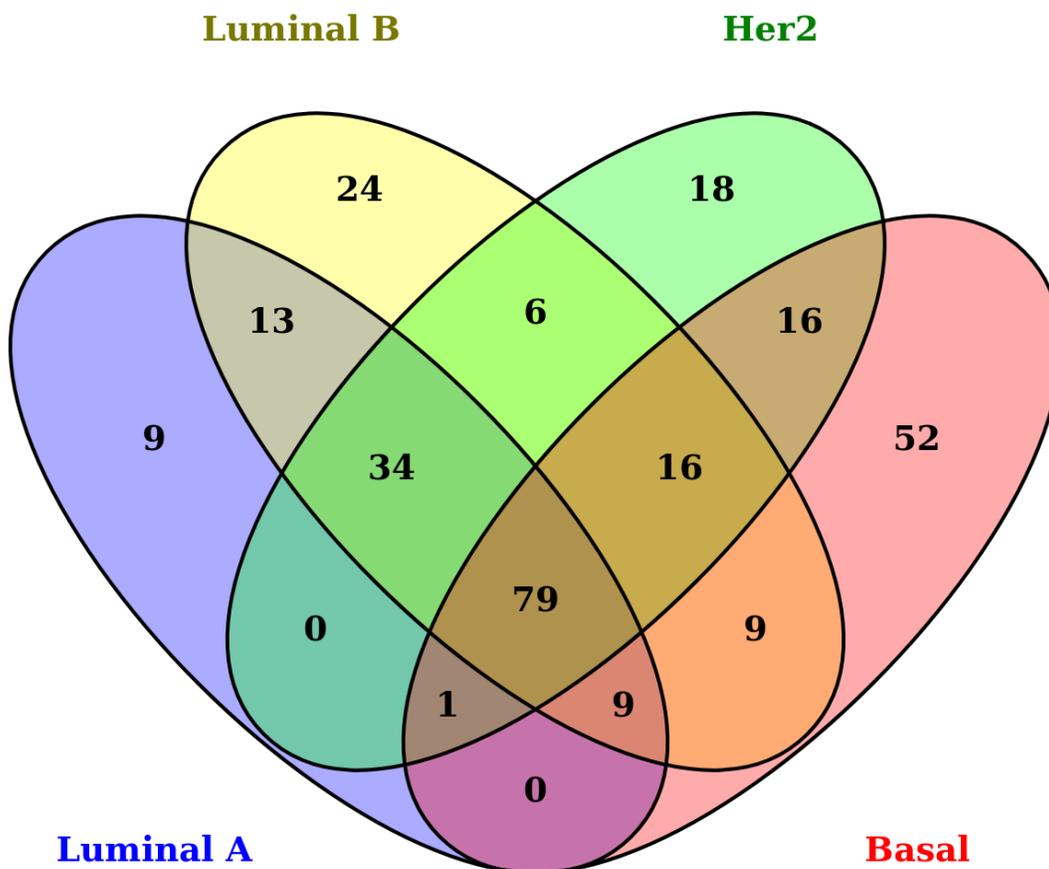


Figura 6. Diagrama de Venn. Se muestra el número de genes subexpresados en los diferentes subtipos tumorales.

Tabla 4. Genes subexpresados únicos en cada subtipo tumoral. Se muestran solo los 10 genes sobreexpresados con mayor expresión diferencial. El subtipo luminal A solo presenta 9 genes subexpresados únicos de este subtipo.

Luminal A	Luminal B	Her2	Basal
CKB	LIPG	QPRT	ACOX2
FOLH1	GALNACT1	PLCH1	CA12
GALT	NNMT	MGAT5B	GSTM3
GBE1	DGKI	GSR	GPD1L
GPT2	FMO3	SRD5A3	INPP4B
GSTA4	FUT4	G6PD	INPP5J
OXCT1	PDE4B	PGM2L1	ALDH6A1
PC	PLCB4	P4HA2	ECI2
PLCH1	GCNT2	CHSY3	CMBL
	CYP1B1	GLB1	TM7SF2

En la tabla 5 se resume la cantidad de genes diferencialmente expresados en todos los subtipos tumorales; mientras que en la tabla 6 el número de genes diferencialmente expresados únicos de cada subtipo tumoral, es decir, los genes que sólo se expresan en un subtipo tumoral y no en los demás.

Tabla 5. Número total de genes diferencialmente expresados para cada subtipo tumoral.

Genes	Luminal A	Luminal B	Her2	Basal
Sobreexpresados	67	100	108	101
Subexpresados	145	190	170	182

Tabla 6. Número de genes diferencialmente expresados únicos de cada subtipo tumoral.

Genes	Luminal A	Luminal B	Her2	Basal
Sobreexpresados	8	18	21	46
Subexpresados	9	24	18	52

### 5.1.1. Relación genes diferencialmente expresados y vías metabólicas.

Después de obtener la expresión diferencial en los diferentes subtipos de cáncer de mama el siguiente objetivo fue identificar la relación de los genes con mayor expresión diferencial y las vías metabólicas en las que se encuentran involucrados. Los 10 genes sobreexpresados comunes, se eligieron para observar su relación con la vías metabólicas de estudio. La figura 7 representa la relación de los genes sobreexpresados, en color rojo y las vías metabólicas asociadas a dichos genes, en color rosa.

Los genes relacionados con mayor número de vías fueron el gen IL411 y el gen SDS ambos involucrados en vías de aminoácidos. Estos dos genes, tienen en común que están presentes en la vía del metabolismo de cisteína y metionina, sin embargo SDS también se asocia con vías del metabolismo de carbono.

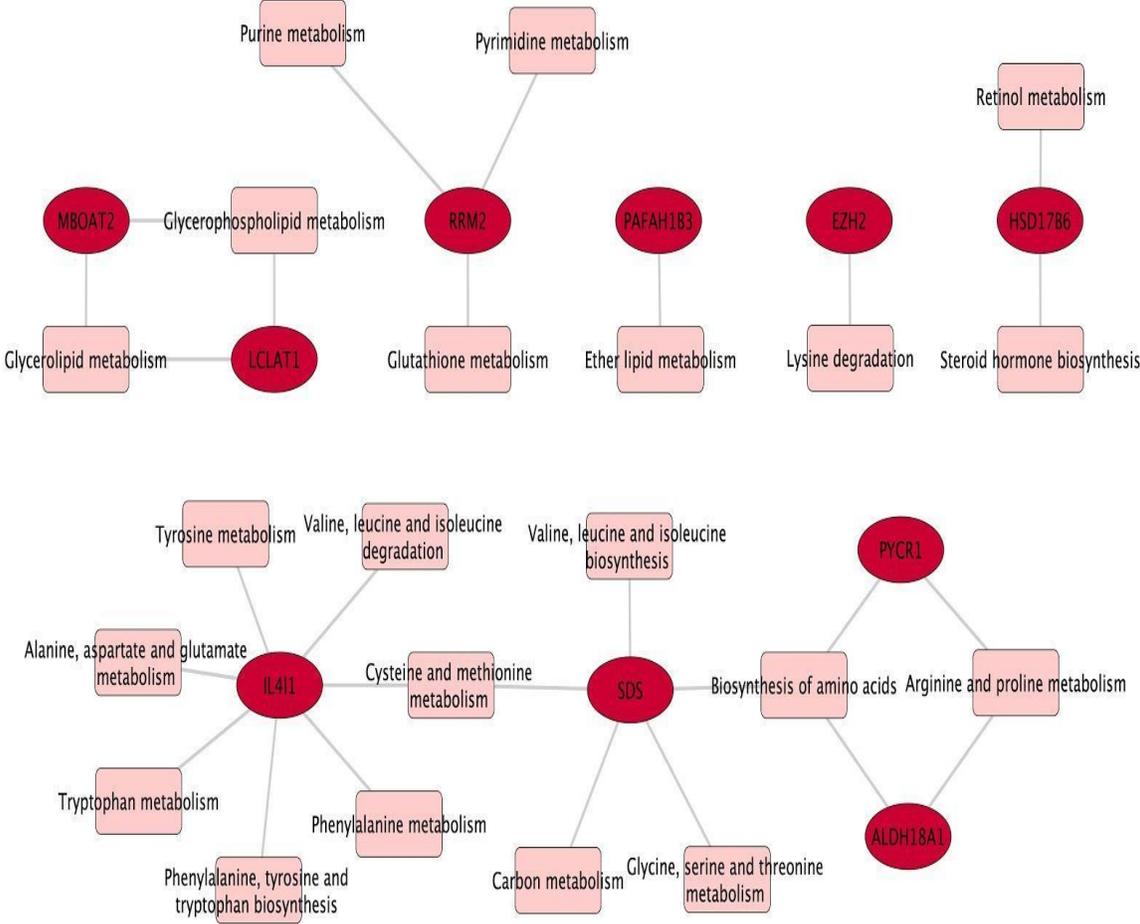


Figura 7. Red que relaciona los genes sobreexpresados, en color rojo, comunes para los cuatro subtipos tumorales, con las vías metabólicas a las que se asocian, en rosa

También se observan genes involucrados en diferentes vías del metabolismo de lípidos, como son LCLAT1, MBOAT2, HSD17B6, PAFAH1B3. A diferencia de los demás genes RRM2 es un gen involucrado en vías de metabolismo de purinas, metabolismo de pirimidinas y metabolismo de glutatión (Figura 7).

La tabla 7 presenta los valores de expresión de los genes sobreexpresados comunes a los 4 subtipos. Algunos genes como EZH2, RRM2 y HSD17B6 se encuentran entre los genes con mayor valor de expresión diferencial consistentemente entre los diferentes subtipos (el valor máximo de expresión diferencial cambia en cada subtipo).

A continuación en la tabla 8 se describen las funciones de cada uno de los 10 genes sobreexpresados comunes a los 4 subtipos tumorales.

Tabla 7. Valores de expresión diferencial de los genes sobreexpresados comunes para los cuatro subtipos tumorales. La tabla muestra los valores de log fold change para cada subtipo tumoral

Gen	Luminal A	Luminal B	Her2	Basal
ALDH18A1	1.037675	1.085271	1.441966	1.147434
EZH2	1.831320	3.002159	3.197941	3.887844
HSD17B6	2.715453	2.795574	3.231483	2.552097
IL4I1	1.570941	1.904331	2.836210	3.139105
LCLAT1	1.021394	1.475294	1.584314	1.734744
MBOAT2	1.755001	1.928848	2.163166	1.612596
PAFAH1B3	2.043799	2.450379	2.115915	1.747070
PYCR1	1.628938	2.104200	2.458958	2.022777
RRM2	1.945293	3.002159	4.188379	3.818371
SDS	2.319184	2.678921	2.988193	2.372645

Tabla 8. Función de los 10 genes que se encuentran sobreexpresados en los 4 subtipos tumorales.

Gen	Función	Referencia
ALDH18A1	Codifica una enzima aldehído deshidrogenasa, asociado a enfermedades del tejido conectivo	Bicknell <i>et al.</i> , 2008
EZH2	Enhancer of Zeste 2 polycomb repressive complex 2. Codifica lisin -N metiltransferasa, metila la lisina para degradación. Pertenece a Polycomb, grupo proteico encargado de la represión transcripcional	Yan <i>et al.</i> 2017
HSD17B6	Hidroxiesteroide 17-Beta deshidrogenasa 6. La enzima presenta actividad epimerasa y oxidoreductasa	Muthusamy <i>et al.</i> 2011
IL4I1	Produce la enzima L-aminoácido oxidasa, que cataliza la reacción de producción de oxaloacetato, entre otras reacciones. La expresión de este gen se induce por interleucina 4. Regula el crecimiento y la función de diferentes tipos celulares del sistema inmune.	Boulland <i>et al.</i> 2007
LCLAT1	Liso Cardiolipin aciltransferasa, transfiere grupos acilo. Implicada en enfermedades relacionadas con estrés oxidante	Li <i>et al.</i> 2012
MBOAT2	O-aciltransferasa unida a membrana. Implicado en el aumento de expresión de fosfatidilcolina en cáncer colorrectal	Kurabe <i>et al.</i> 2013
PAFAH1B3	Codifica una enzima acetil hidrolasa que remueve un grupo acetilo de esqueletos de glicerol del factor activador de plaquetas	Adachi <i>et al.</i> , 1995
PYCR1	Pirrolin-5- carboxilato reductasa 1. Involucrado en la respuesta celular al estrés oxidante. Sobreexpresión del transcrito relacionado con tumores luminal B.	Craze <i>et al.</i> , 2018
RRM2	Ribonucleósido difosfato reductasa subunidad M2. Enzima esencial en la síntesis y reparación del DNA	Pavloff <i>et al.</i> , 1991
SDS	Serina deshidratasa que cataliza la reacción de L-serina o L-glicina a piruvato	Yamada <i>et al.</i> , 2008

Asimismo, se determinó la relación entre los genes subexpresados y las vías metabólicas en las que se encuentran involucrados. Se encontraron 79 genes subexpresados comunes a todos los subtipos y 64 vías asociadas con estos genes. La figura 8 muestra la relación de los genes con las vías metabólicas en las que actúan.

Las vías metabólicas que se asociaron a los genes están involucradas en el metabolismo de lípidos, carbono, aminoácidos, nitrógeno y purinas.

Destaca en esta relación la vía del metabolismo de purinas que está relacionada con 10 genes subexpresados. Esta vía es la que presenta mayor número de genes asociados y además estos genes no actúan en ninguna otra vía.

Por otro lado, ALDH2 es el gen asociado a mayor número de vías. Algunas de las 12 vías con las que se relaciona este gen son glicólisis, interconversión de pentosas y glucuronato, degradación de ácidos grasos, y algunos aminoácidos, metabolismo de piruvato, glicerolípidos, ascorbato y aldarato, y aminoácidos como histidina y triptófano.



## 5.2. Desregulación de vías metabólicas y agrupamiento jerárquico

En este estudio se realizó un análisis de desregulación de vías metabólicas que nos da como resultado un valor de desregulación de vía (PDS) inferido a partir del algoritmo Pathifier.

El análisis de desregulación de vías se presenta en la figura 9. En la parte superior de la figura se observa el dendrograma del agrupamiento jerárquico no supervisado, el cual nos dio como resultado que las muestras de estudio se agrupaban por subtipo tumoral, salvo algunas excepciones (figura 9).

El mapa de calor de la figura 9 expone la diferencia de desregulación en los diferentes subtipos tumorales. Se observan patrones definidos del valor de PDS para subtipo tumoral en cada vía. Sólo en algunos casos el agrupamiento jerárquico situó en otro grupo algunas muestras que no pertenecían al subtipo tumoral. Se aprecia que 10 muestras que se clasificaron como luminal B se agruparon dentro de las muestras de luminal A y de Her2. El mismo caso para luminal A y Her2 en donde algunas muestras se agruparon dentro de luminal B. Además, de acuerdo al agrupamiento jerárquico, las muestras con subtipo tumoral luminal B y Her2 se encuentran mayormente relacionadas entre sí, en comparación con el luminal A. Por el contrario, todas las muestras con subtipo

tumoral basal, forman un único grupo diferente a los otros 3 subtipos mencionados anteriormente.

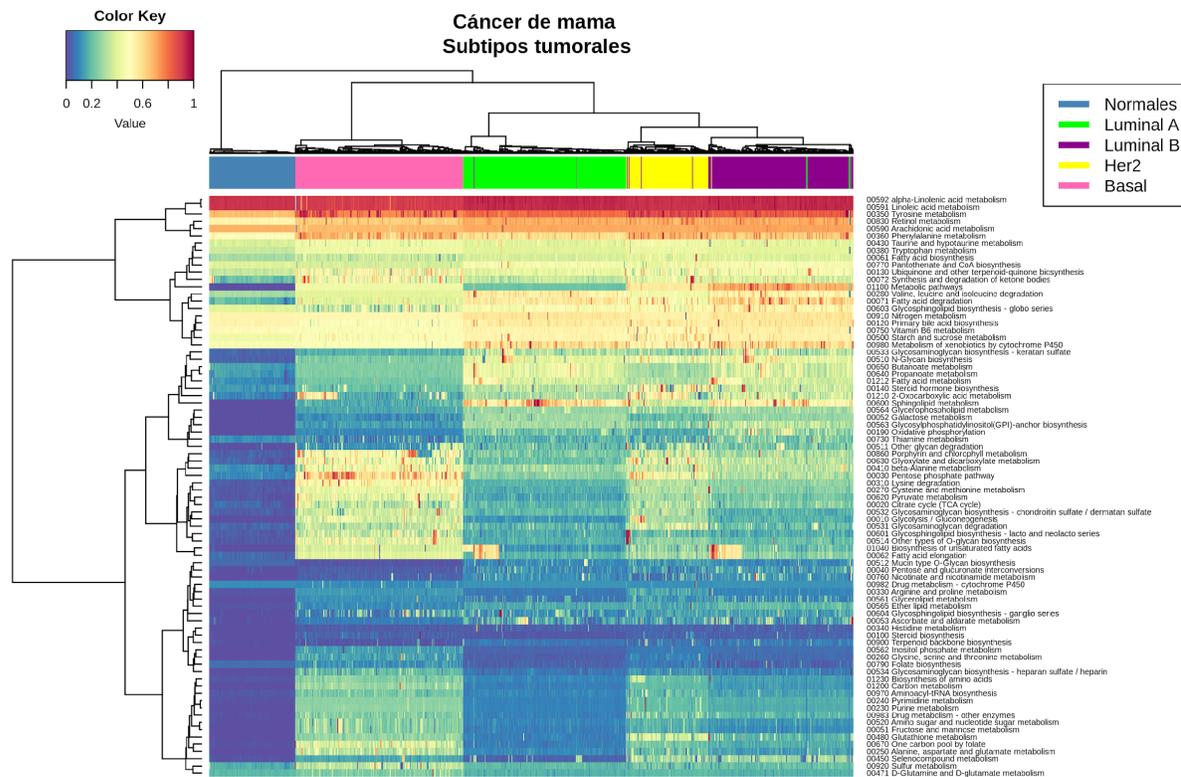


Figura 9. Mapa de calor. Representa el PDS (Pathway Deregulation Score) del análisis de Pathifier. Para vías de metabolismo (KEGG). Cada columna corresponde a una muestra y cada fila corresponde a una vía. Las vías y las muestras se agrupan de acuerdo a su valor de PDS. El color azul representa un bajo puntaje (no desregulación) y el color rojo representa un alto puntaje (mayor desregulación). La barra superior representa el subtipo tumoral comparados entre sí y con las muestras de sanos. Cabe resaltar que el agrupamiento basado en las muestras define a los subtipos tumorales en diferentes grupos.

Además se observa el nivel de desregulación de las muestras por cada vía metabólica, destacando que para el conjunto de genes de la vía denominada “Metabolic Pathways” los cambios en el valor de PDS están claramente definidos por el subtipo tumoral (figura 9). Se realizó un nuevo mapa de calor con esta vía.

La figura 10 es una representación únicamente de la vía “Metabolic Pathways” que presenta un agrupamiento jerárquico que define claramente a los subtipos de cáncer de mama, cuando se analizan en conjunto todos los genes involucrados en todas las vías metabólicas. Se puede observar que el nivel de desregulación va en aumento dependiendo del subtipo tumoral, es decir, luminal A parece ser el subtipo con menor desregulación a nivel de vías metabólicas, seguido del subtipo basal. Por otro lado, Her2 presenta mayor desregulación que los dos subtipos mencionados anteriormente, y finalmente el subtipo más desregulado es luminal B.

En este mapa de calor (figura 10) se realizó un agrupamiento específico para la vía “Metabolic Pathways” (en la figura 10 el agrupamiento toma en cuenta todas las vías de estudio), dicho agrupamiento muestra una relación más cercana entre los subtipos her2 y basal. El subtipo luminal B presenta más variaciones en sus valores de PDS y por tanto forma más subgrupos en el agrupamiento, este subtipo tiene mayor relación con los subtipos her2 y basal en comparación a luminal A. Luminal A es el subtipo con menor puntaje de desregulación, sus valores de PDS

son cercanos a 0.2 (figura 10), se asocia más cercanamente a las muestras control que a los otros subtipos tumorales.

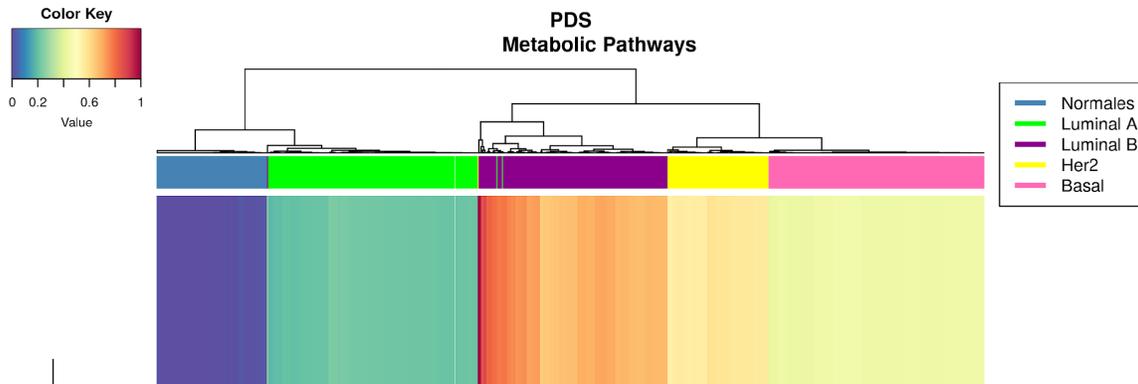


Figura 10. Mapa de calor. Representa el PDS (Pathway Deregulation Score) del análisis de Pathifier para el conjunto de genes denominado "Metabolic Pathways" (KEGG) que engloba los genes involucrados en todas las vías metabólicas. Cada columna corresponde a una muestra y cada fila corresponde a una vía. Las vías y las muestras se agrupan de acuerdo a su valor de PDS. El color azul representa un bajo puntaje (no desregulación) y el color rojo representa un alto puntaje (mayor desregulación). La barra superior representa los subtipos tumorales comparados entre sí y con las muestras de sanos. Cabe resaltar que el agrupamiento jerárquico basado en las muestras define a los subtipos tumorales en diferentes grupos.

Dadas las diferencias tan notorias de los subtipos tumorales, se decidió mostrar la frecuencia de los valores de PDS. En la figura 10 observamos las diferencias entre los valores de PDS de los subtipos tumorales.

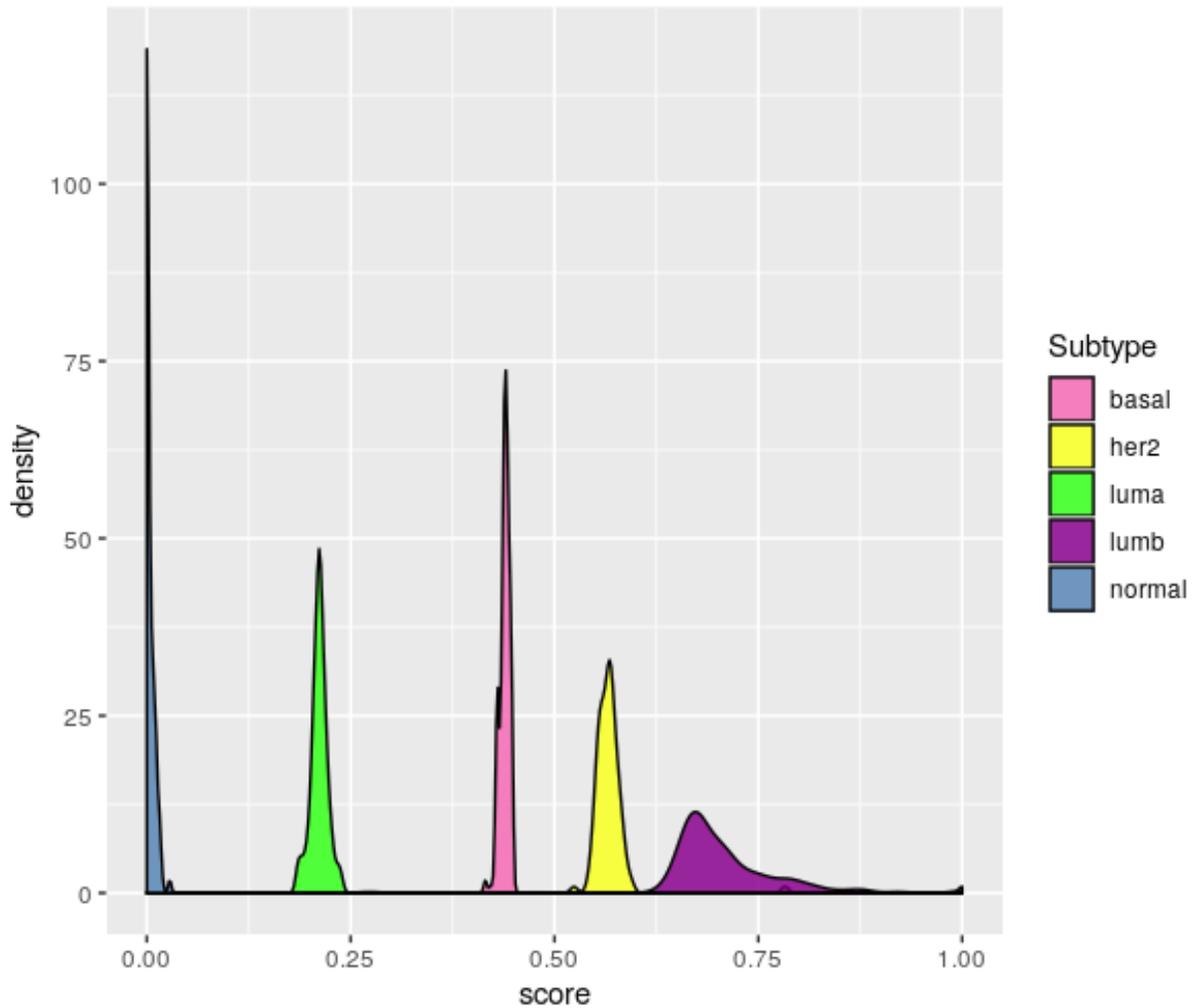


Figura 11. Histograma de frecuencias. Muestra la frecuencia de los valores PDS (Pathway Deregulation Score) del análisis de Pathifier para el conjunto de genes denominado “Metabolic Pathways” (KEGG) que engloba los genes involucrados en todas las vías metabólicas. La curva de color azul, son los PDS de las muestras normales, en color verde los valores de PDS del subtipo luminal A, en color morado luminal B, en color amarillo her2 y finalmente en color rosa el subtipo basal. Los picos de cada subtipo tumoral se encuentran bien definidos, es decir, el traslape de las distribuciones es casi nulo.

Los PDS de las muestras normales (en azul) no presentan variación, es decir, sus valores son aproximadamente cero. Por otro lado, como ya se había descrito anteriormente, luminal A es el subtipo con menor desregulación, seguido de basal y her 2, respectivamente y finalmente el subtipo con mayor desregulación luminal B (figura 11). Destaca en este figura la gran variación de PDS en luminal B que oscila de 0.6 hasta 1.

Principal curve of 01100 Metabolic pathways

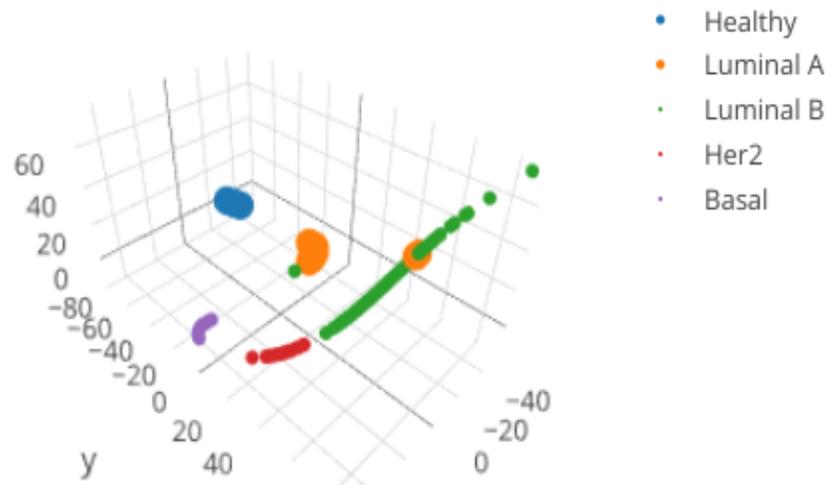


Figura 12. Gráfico de la curva principal utilizada para calcular el valor de PDS del conjunto de genes denominado "Metabolic Pathways". En color azul los controles, en color naranja las muestras con subtipo tumoral Luminal A, en color verde las muestras Luminal B, en color rojo las muestras Her2 y en color morado las muestras Basal. Se observa en este conjunto de genes una diferencia entre las muestras de diferentes subtipos tumorales, así como en comparación con los controles.

La figura 12 es una representación de la curva principal diseñada para cuantificar el valor de PDS, esta figura representa el nivel de desregulación entre cada subtipo como se describió en el párrafo anterior. En esta figura se observa la distribución espacial de cada uno de los valores de PDS para cada muestra, resaltando que cada subtipo se encuentra separado de los demás, así como del grupo control.

Otro de los resultados a resaltar son las vías que parecen tener un mismo valor de PDS tanto en controles como en tumores (Figura 13). Para visualizar estas vías tomamos un subconjunto de datos de las vías metabólicas.

Algunas de las vías que presentan un valor similar de PDS a lo largo de todas las muestras (tanto muestras controles como muestras tumorales) son las vías de metabolismo de ácido linoléico, metabolismo de ácido alfa-linoléico, metabolismo de ácido araquidónico y biosíntesis de esteroides.

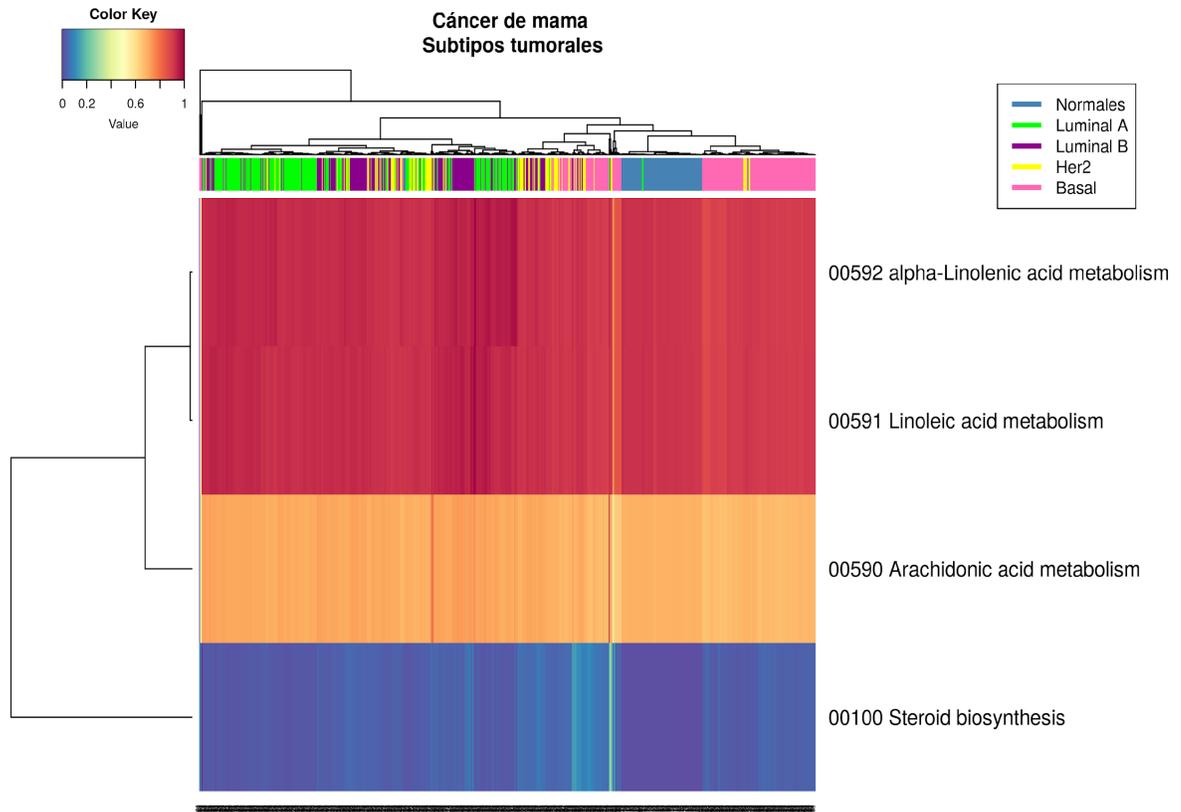


Figura 13. Mapa de calor. Representa el PDS (Pathway Deregulation Score) del análisis de Pathifier para las vías de metabolismo de ácido linoléico, metabolismo de ácido alfa-linoléico, metabolismo de ácido araquidónico y biosíntesis de esteroides (KEGG). Cada columna corresponde a una muestra y cada fila corresponde a una vía. Las vías y las muestras se agrupan de acuerdo a su valor de PDS. El color azul representa un bajo puntaje (no desregulación) y el color rojo representa un alto puntaje (mayor desregulación). La barra superior representa el subtipo tumoral comparados entre sí y con las muestras de sanos. Cabe resaltar que el agrupamiento basado en las muestras solo agrupa a los controles y al subtipo basal.

La figura 13 es la representación del subconjunto de vías descritas anteriormente. En esta figura, se observa que los valores de PDS son diferentes entre las vías, pero es interesante notar que dentro de las vías todas las muestras tienen un valor de PDS muy cercano entre sí, y que estas vías pertenecen a un subgrupo de vías metabólicas que están involucradas en general en el metabolismo de lípidos.

Con respecto al agrupamiento jerárquico, es notorio que las muestras se agrupan de manera diferente a su subtipo tumoral manteniendo en un mismo subgrupo a los grupos controles y en otro subgrupo definido a la mayoría de las muestras con subtipo basal (figura 13).

Posteriormente se realizó un nuevo análisis de vías, para cada subtipo tumoral.

La figura 14 representa los valores de PDS del análisis de vías para el subtipo luminal B. Se observa que las vías con mayor número de muestras desreguladas en este subtipo son el metabolismo de piruvato, el metabolismo de tirosina, el metabolismo de ácido araquidónico, degradación y biosíntesis de ácidos grasos y la vía de las pentosas fosfato. Algunas de las vías presentan sólo desregulación en subgrupos de muestras, lo que describe la heterogeneidad de las muestras en este subtipo.

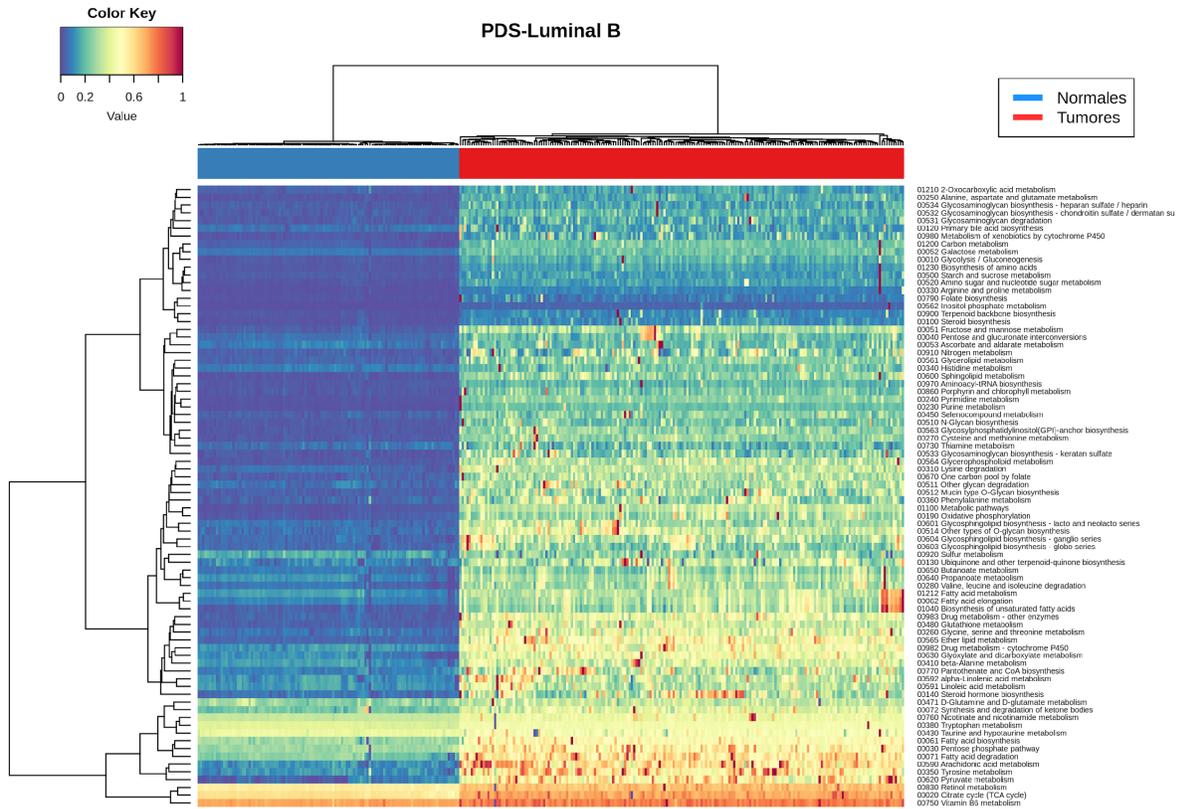


Figura 14. Mapa de calor. Representa el PDS (Pathway Deregulation Score) del análisis de Pathifier para las muestras con subtipo luminal B. Cada columna corresponde a una muestra y cada fila corresponde a una vía. Las vías y las muestras se agrupan de acuerdo a su valor de PDS. El color azul representa un bajo puntaje (no desregulación) y el color rojo representa un alto puntaje (mayor desregulación). La barra superior representa, en azul, a las muestras controles y en rojo a las muestras tumorales.

### **5.3. Metabolismo de purinas como blanco potencial en todos los subtipos de cáncer de mama**

La observación de que un subconjunto de genes están comúnmente diferencialmente expresados en todos los subtipos de cáncer, sugiere que deberían existir fármacos comunes para cualquier tumor de mama, independientemente del subtipo. Para evaluar esta idea, se empleó un método computacional de minería de datos transcriptómicos (desarrollado por nuestro grupo de trabajo) con el fin de encontrar, a partir de los genes diferencialmente expresados, blancos terapéuticos para fármacos aprobados por la Administración de Fármacos y Alimentos de Estados Unidos (FDA, por sus siglas en inglés) (Mejía-Pedroza *et al.*, 2018).

En la tabla 9 se describen los genes sobreexpresados pertenecientes al metabolismo de purinas que podrían ser inhibidos por algún fármaco así como, los genes subexpresados que podrían ser activados. RRM2 es un gen que participa, además, en procesos relacionados con el metabolismo de pirimidinas y del glutatión, y presentó la mayor cantidad de inhibidores.

Tabla 9. Genes diferencialmente expresados con inhibidores o activadores aprobados por la FDA para regular el metabolismo de purinas.

Genes	Fármaco	Tipo de Interacción
RRM2	Fludarabín fosfato	inhibidor
RRM2	Nitrato de galio	inhibidor
RRM2	Cladribina	inhibidor
RRM2	Clofarabina	inhibidor
RRM2	Fludarabina	inhibidor
RRM2	Gemcitabina*	inhibidor
RRM2	Hidroxiurea	inhibidor
RRM2	Motexafin gadolinium	inhibidor
RRM2	Tezacitabina	inhibidor
RRM2	Hidrocloruro de gemcitabina	inhibidor
EZH2	CHEMBL3287735	inhibidor
ACACB	Metformina	activador
NPR1	Ataciguat	activador
PDE1C	Bepridil	activador
PDE2A	CHEMBL395336	activador

## 6. Discusión

Una de las características distintivas del cáncer de mama es la desregulación de la energética celular (Hanahan y Weinberg, 2011), en donde las células tumorales reprograman su metabolismo para sustentar su crecimiento y proliferación (DeBerardinis et al., 2016; Pavlova y Thompson, 2016).

Dado que el cáncer en general es una enfermedad relacionada con daños y modificaciones en el genoma humano (Stratton et al., 2009), en este estudio se analizó cómo es que la expresión de genes podría influir en el fenotipo metabólico desde dos enfoques.

Nuestro primer enfoque se basa en el estudio de las diferencias de expresión de los genes relacionados con el metabolismo (en específico los genes que codifican para enzimas) en el fenotipo tumoral entre los subtipos de cáncer de mama y comparado con muestras control. Además se integró la información que relaciona a los genes diferencialmente expresados con las vías metabólicas implicadas.

Se llevó a cabo un análisis de expresión diferencial para cada subtipo tumoral, de aproximadamente 1200 genes relacionados con el metabolismo. En un estudio anterior, Wu y colaboradores (2017), encontraron 136 genes diferencialmente expresados en los 4 subtipos. Por el contrario, en este trabajo los genes diferencialmente expresados para los 4 subtipos tumorales fueron 89, ya que el

conjunto de genes de estudio se acotó a los relacionados con el metabolismo, específicamente los que codifican para enzimas.

Los genes con expresión diferencial fueron separados en sobreexpresados y subexpresados debido a que cada subtipo tumoral presenta un perfil de expresión diferente y algunos genes sobreexpresados en un subtipo pueden estar subexpresados en otro, por ejemplo, el gen UGT8 que codifica para la enzima UDP glicosiltransferasa 8 y que está involucrado en las vías de metabolismo de lípidos de éter y metabolismo de esfingolípidos (KEGG, 2019), presentó sobreexpresión en el subtipo basal, mientras que en luminal B se encontró subexpresado. De los 204 genes sobreexpresados en total de los cuatro subtipos, sólo 10 genes fueron comunes en los 4 subtipos.

Uno de los genes sobreexpresados comunes involucrado en mayor número de vías es el gen IL4I1, el cual está relacionado con vías de metabolismo de diferentes aminoácidos como lo son aspartato, alanina, valina, fenilalanina, entre otros. El gen IL4I1 codifica para una enzima con actividad L-aminoácido oxidasa (Boulland et al., 2007) que cataliza reacciones de desaminación de aminoácidos produciendo amoníaco, peróxido de hidrógeno y metabolitos precursores de moléculas relacionadas con cáncer, como el oxaloacetato que mantiene el flujo metabólico del ciclo de Krebs durante el crecimiento tumoral (DeBerardinis et al., 2007). Además, este gen está implicado en procesos inmunológicos, ya que es

inducido por interleucina 4 y su expresión se da también en células dendríticas y macrófagos asociados a tumores (Bod *et al.*, 2017).

Otro gen interesante, es el gen EZH2 que interviene en la vía de degradación de la lisina. Este gen codifica una enzima llamada histona lisina metil transferasa, que cataliza la metilación de colas de lisina en procesos dependientes de cromatina, como la transcripción de genes, la replicación del DNA y la reparación por daño al DNA (Hirschey *et al.* 2015; Gan *et al.*, 2018). Se han encontrado a EZH2 sobreexpresado en diferentes tipos de cáncer, entre ellos el de mama (Holm *et al.*, 2012), asociando la sobreexpresión de este gen a tumores de mama más agresivos (Kleer *et al.*, 2003; Chang *et al.*, 2011).

En este caso, la vía metabólica con mayor número de genes asociados fue la vía de biosíntesis de aminoácidos. De los 10 genes sobreexpresados comunes, 3 tienen relación con esta vía, de los cuales el gen SDS resaltó debido a que está involucrado con el metabolismo de carbono y con la biosíntesis de aminoácidos.

La sobreexpresión de SDS implica un posible aumento en la concentración de la enzima serina deshidratasa que conlleva a la formación de piruvato a partir de serina o viceversa, según los requerimientos de la célula tumoral. Esta idea se basa en lo descrito por Pavlova y Thompson (2016), en donde se explica que las células tumorales aumentan flujos metabólicos específicos, generando mayor

cantidad de macromoléculas para su crecimiento. Estudios experimentales validarían la idea y podría definir hacia donde está definida la reacción, hacia la producción de piruvato para formar esqueletos de carbono o hacia la producción de L-serina para formación de proteínas.

Por otro lado, se encontró que de estos 10 genes sobreexpresados comunes que pertenecen a 21 vías metabólicas, solo RRM2 se relaciona con vías de síntesis de DNA y otros eventos de regulación celular como lo son el metabolismo de pirimidinas, purinas y glutatión (Wu *et al.*, 2004). La sobreexpresión de este gen ha sido asociada a cáncer de mama, de cabeza y cuello, de pulmón, de vejiga entre otros (Chen *et al.*, 2019; Sun *et al.*, 2019; Rahman *et al.*, 2013; Morikawa *et al.*, 2010), y se ha demostrado que la sobrerregulación de la expresión de RRM2 acelera la división celular mediante el aumento en la acumulación de dinucleótidos (Sun *et al.*, 2019 ).

Sin embargo, mientras que RRM2, un gen que pertenece a la vía del metabolismo de purinas, se encuentra sobreexpresada en todos los subtipos tumorales, se presentaron 10 genes subexpresados comunes, en esta misma vía (figura 8). Los genes ADCY4 y ADCY5 codifican para enzimas adenilato ciclasa, las cuales se encargan de regular las proporciones de nucleótidos (Klepinin *et al.* 2016); los genes AK5 que codifica adenilato cinasa catalizan reacciones de hidrólisis de ATP (Sadana y Dessauer, 2008); finalmente NPR1 y NPR2, y 5 genes de la familia

fosfodiesterasas, que codifican enzimas que controlan la proporción de segundos mensajeros y están implicados en la transducción de señales (Jeon *et al.* 2005).

Lo descrito anteriormente nos da una pauta de posibles mecanismos que podrían estar involucrados en la desregulación del metabolismo globalmente en el cáncer de mama. Sin embargo, para conocer las diferencias intrínsecas de cada subtipo es necesario un análisis más minucioso. En este estudio se identificó a los genes que se expresan diferencialmente en un subtipo tumoral pero no en los demás, a los que nombramos “genes diferencialmente expresados únicos” de cada subtipo tumoral (Tabla 6). Las diferencias entre estos genes indican que los posibles cambios dados en el metabolismo podrían estar relacionados a nivel molecular por el subtipo tumoral .

UGT8 es un gen que codifica para la enzima UDP glicosiltransferasa 8 (NCBI, 2019) involucrado en las vías de metabolismo de lípidos de éter y metabolismo de esfingolípidos (KEGG, 2019), el producto de este gen cataliza la reacción de transferencia de galactosa a ceramidas, liberando uridina difosfato. Este gen solo se sobreexpresa en el subtipo tumoral basal; se ha descrito que en la mayoría de los casos subtipo basal tiene un fenotipo triple negativo (Lachapelle y Foulkes, 2011), por lo que la sobreexpresión del gen UGT8 en el subtipo basal concuerda con la evidencia reportada en la literatura (Santuario-Facio *et al.* 2017). La expresión de UGT8 tiene un mal pronóstico en los pacientes ya que proporciona

ventajas tumorigénicas y metastásicas a las células tumorales con subtipo tumoral basal (Cao et al. 2018). Estudios relacionados con este gen describen que inhibir la proteína expresada impide la metástasis tumoral, sin afectar la expresión del gen (Cao et al. 2018).

Este enfoque del análisis de expresión diferencial de genes nos permitió observar diferencias y similitudes entre los subtipos de cáncer de mama. Sin embargo, el análisis gen a gen se complica por la gran cantidad de genes de estudio.

Debido a lo anterior, se realizó un análisis de desregulación de vías para darle un enfoque diferente al estudio del metabolismo basado en expresión génica. Este segundo enfoque engloba todos los genes involucrados en una misma vía metabólica. A diferencia de lo realizado por Wu y colaboradores (2017), que toman la media del valor de expresión de sus muestras, en este estudio el análisis de desregulación de vías se calcula para cada muestra, empleando Pathifier como método para cuantificar un valor de desregulación.

Livshits y colaboradores (2015) analizaron mediante Pathifier muestras de cáncer de mama encontrando una clasificación de 9 tipos tumorales con distintos perfiles transcriptómicos en 7 grupos de vías (de señalización; de sistema inmune; relacionadas a cáncer, ciclo celular y reparación de DNA; señalización/inmune; relacionadas con la proteína cinasa A; relacionadas con el receptor del factor de

crecimiento epidérmico (EGFR) y; relacionadas con el gen PAK1 (proteína cinasa de serina/treonina)). En este estudio se incluyeron únicamente las vías de metabolismo y los valores de PDS agruparon 4 tipos diferentes de perfiles transcriptómicos que correlacionan con los subtipos tumorales. Esto sugiere que la desregulación en los genes del metabolismo podría depender de la clasificación molecular descrita en la literatura (Perou *et al.*, 2000; Sorlie *et al.*, 2001) .

La figura 10 nos muestra que el subtipo tumoral con mayor desregulación dentro del conjunto de genes denominado “metabolic pathways” es el subtipo luminal B, este es un importante hallazgo ya que el subtipo luminal B es el más difícil de tratar de acuerdo con el Panel Internacional de Consenso de St. Gallen en 2015 (Capelletti *et al.* 2017; Coates *et al.* 2015) lo que sugiere que la expresión de este conjunto de genes podría estar ligada al fenotipo tumoral, es decir, se sugiere que una mayor desregulación en las vías metabólicas supone una mayor dificultad para tratar el tumor. Análisis posteriores a este conjunto de datos (muestras de tejido tumoral mamario con subtipo tumoral luminal B), así como pruebas experimentales, se requieren para un mejor entendimiento de la desregulación del metabolismo en el subtipo luminal B.

Posteriormente, se analizaron las vías metabólicas que no presentaron diferencias entre las muestras tumorales y las muestras controles, encontrando que las vías

sin diferencias significativas entre los grupos de estudio son las relacionadas con metabolismo de lípidos.

Una de las vías que se encontró más interesante, es la vía relacionada con el metabolismo de ácido linoléico. El ácido linoléico y el ácido  $\alpha$ -linolénico, son ácidos grasos esenciales precursores de metabolitos más complejos como el ácido araquidónico (Lu *et al.* 2010) y se ha descrito que podría tener funciones protectoras contra el cáncer (Lee *et al.* 2005).

En el análisis de desregulación de vías, la vía de metabolismo del ácido linoléico, presentó una dispersión muy grande en sus valores de PDS tanto en muestras control como en muestras tumorales, calculando un puntaje de desregulación de vía cercano a 1 sin diferenciar tumores de controles, por lo que esta vía no se puede considerar dentro del análisis de desregulación de vías global (figura 13).

Sin embargo, al realizar el análisis de vías por subtipo tumoral se encontró que en el subtipo basal el metabolismo de ácido linoléico no se encuentra desregulado, mientras que para el subtipo her2 parece tener mayor desregulación comparado con los demás subtipos (anexo A ). Esta vía podría estar implicada en un proceso de protección contra el cáncer en el subtipo her2, por lo que es relevante para futuras investigaciones.

En general, el análisis de desregulación de vías nos permitió observar patrones de expresión diferentes entre los subtipos tumorales, así como la variación del nivel de desregulación entre las muestras. El metabolismo de purinas es una vía que se encuentra desregulada en algunas muestras, sin embargo podemos observar que presenta genes diferencialmente expresados que coinciden en todos los subtipos tumorales. Debido a lo anterior, se decidió buscar posibles tratamientos farmacológicos para esta vía. El gen RRM2, que presentó sobreexpresión en los 4 subtipos tumorales y se encuentra involucrado en las vías de pirimidina y el metabolismo de glutatión, puede ser inhibido por diferentes fármacos aprobados por la FDA (Mejía-Pedroza *et al.*, 2018). Cabe resaltar, que los fármacos gemcitabina que inhibe el gen RRM2 y la metformina, activador del gen ACACB, ya se emplean para tratamientos en pacientes con cáncer de mama (Xie *et al.*, 2018; Saraei *et al.*, 2019).

Por lo anterior, consideramos que el análisis de vías para cada muestra permite identificar, para cada paciente, las vías con mayor desregulación para luego realizar un estudio más profundo en las vías seleccionadas y considerar los genes con expresión diferencial que podrían ser empleados como blancos farmacológicos. Desde esta aproximación, el análisis de vías podría contribuir a la búsqueda de tratamientos terapéuticos personalizados.

## 7. Conclusión

La heterogeneidad es un factor crucial que impide el entendimiento, el diagnóstico y el tratamiento del cáncer de mama. Las manifestaciones de esta heterogeneidad pueden ser observadas a nivel genómico, histológico o clínico. En este trabajo se ha proporcionado otro caso de esta heterogeneidad: la desregulación metabólica.

Cada subtipo de cáncer tiene su propio patrón de desregulación en el metabolismo, siendo luminal B el que presenta mayores puntajes de desregulación. Este subtipo exhibe alteraciones en procesos metabólicos como el metabolismo de piruvato, de tirosina, la degradación de ácidos grasos y la vía de las pentosas fosfato.

Hasta donde sabemos, esta es la primera vez que un análisis de vías basado en una sola muestra ha sido utilizado para identificar las diferencias de regulación metabólica. Y al mismo tiempo, este trabajo nos permite diseñar un esquema terapéutico común aprobado por la FDA, para regular el metabolismo de purinas independientemente del subtipo. Con este tipo de aproximación, es posible determinar patrones de desregulación globales y, a su vez, encontrar firmas individuales que podrían representar un paso más allá hacia la medicina personalizada.

## 8. Perspectivas

Los resultados obtenidos en este estudio abren la puerta a nuevas interrogantes. En primer lugar, observamos que el **metabolismo de ácido linoléico** no presenta un perfil de desregulación, además un análisis minucioso explica que esta ruta metabólica no presenta genes diferencialmente expresados por lo que nuevos estudios podrían enfocarse identificar si esta vía es propensa a intervención farmacológica.

Siguiendo con esta idea, un **análisis de balance de flujos** podría sugerir que vías metabólicas presentan mayor actividad y en cuales disminuye, incluida la ruta del ácido linoléico. Un análisis de balance de flujo es una simulación que se basa en un modelo matemático para estudiar el metabolismo de un sistema biológico y que puede emplear datos, genómicos, proteómicos y/o metabolómicos.

Por otro lado, un análisis proteómico y metabolómico podría ampliar y explicar algunos de los resultados descritos en el presente trabajo, puesto que evalúan las vías metabólicas a niveles de interacción más complejos.

Finalmente, los resultados de esta investigación deben validarse experimentalmente, tanto en modelos *in vitro* como en modelos *in vivo*. Aunque

nos basamos en muestras de pacientes con cáncer, estos sólo reflejan el fenotipo de la población de estudio y es necesario reproducirlos en otros modelos.

## 9. Referencias

- Acharya, A., Das, I., Chandhok, D., & Saha, T., 2010. Redox regulation in cancer: a double-edged sword with therapeutic potential. *Oxidative medicine and cellular longevity*, 3, 23-34.
- Adachi, H., Tsujimoto, M., Hattori, M., Arai, H., y Inoue, K., 1995. cDNA cloning of human cytosolic platelet activating factor acetylhydrolase  $\gamma$ -subunit and its mRNA expression in human tissues. *Biochemical and Biophysical Research Communications*, 214, 180-187.
- Agrawal, A., Gandhe, M. B., Gupta, D., & Reddy, M. V. R., 2016. Preliminary study on serum lactate dehydrogenase (LDH)-prognostic biomarker in carcinoma breast. *Journal of clinical and diagnostic research: JCDR*, 10, BC06.
- Baylin, S. y Ohm, E., 2016. Epigenetic gene silencing in cancer – A mechanism for early oncogenic pathway addiction? *Nature Reviews Cancer*, 6, 107-116.
- Bayraktar, S. y Glück, S., 2013. Molecularly targeted therapies for metastatic triple-negative breast cancer. *Breast Cancer Research and Treatment*, 138, 21-35.
- Bediaga, N.G., Acha-Sagredo, A., Guerra, I., Viguri, A., Albaina, C., Ruiz Diaz, I., Rezola, R., Alberdi, M.J., Dopazo, J., Montaner, D., de Renobales, M., Fernandez, A.F., Field, J.K., Fraga, M.F., Liloglou, T., de Pancorbo, M.M., 2010. DNA methylation epigenotypes in breast cancer molecular subtypes. *Breast Cancer Research*. 12, 1-12.

- Bicknell, L. S., Pitt, J., Aftimos, S., Ramadas, R., Maw, M. A. y Robertson, 2008. A missense mutation in *ALDH18A1*, encoding D1-pyrroline-5-carboxylate synthase (P5CS), causes autosomal recessive neurocutaneous syndrome. *European Journal of Human Genetics*, 16, 1176-1186.
- Bod, L., Lengagne, R., Wrobel, L., Ramspott, J. P., Kato, M., Avril, M., Castellano, F., Molinier-Frenkel, V. y Prévost-Blondel, A., 2017. IL4-induced gene 1 promotes tumor growth by shaping the immune microenvironment in melanoma. *Oncoimmunology*, 6, e1278331.
- Boulland, M. L., Marquet, J., Molinier-Frenkel, V., Möller, P., Guiter, C., Guiter, C., Lasoudris, F., Copie-Bergman, C., Baia, M., Gaulard, P., Leroy, K. y Castellano, F., 2007. Human IL4I1 is a secreted L-phenylalanine oxidase expressed by mature dendritic cells that inhibits T-lymphocyte proliferation. *Blood Journal*, 110, 220-227.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R., Torre, L., Jemal, A., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A cancer journal for clinicians*, 68, 394-424.
- Burstein, H., 2005. The distinctive nature of HER2-positive breast cancers. *The New England Journal of Medicine*, 353, 1652-1654.
- Cao, Q., Chen, X., Wu, X., Lioa, R., Huang, P., Tan, Y., Wang, L., Ren, G., Huang, J. y Dong, C., 2018. Inhibition of UGT8 suppresses basal-like breast cancer progression by attenuating sulfatide-  $\alpha$ V $\beta$ 5 axis. *Journal of Experimental Medicine*, 215, 1679-1692.

- Cappelletti, V., Iorio, E., Miodini, P., Silvestri, M., Dugo, M., y Daidone, M., 2017. Metabolic footprints and molecular subtypes in breast cancer. *Disease Markers*, 2017.
- Chakravarthi, B., Nepal, S. y Varambally, S., 2016. Genomic and epigenomic alterations in cancer. *The American Journal of Pathology*, 186, 1724-1735.
- Chang, C.J., Yang, J. Y., Xia, W., Chen, C. T., Xie, X., Chao, C. H., Woodward, W., Hsu, J. M., Hortobagyi, G. y Hung, M. C., 2011 EZH2 promotes expansion of breast tumor initiating cells through activation of RAF1- $\beta$ -Catenin signaling. *Cancer Cell*, 19, 86-100.
- Chen, W., Yang, L., Xu, L., Cheng, L., Qian, Q., Sun, L. y Zhu, Y., 2019. Bioinformatics analysis revealing prognostic significance of RRM2 gene in breast cancer. *Bioscience Reports*, 39 1-9.
- Coates, A. S., Winer, E. P., Goldhirsch, A., Gelber, R. D., Gnant, M., Piccart-Gebhart, M., Thürlimann, B., Senn, H. J.; Panel Members, 2015. Tailoring therapies-- improving the management of early breast cancer: St Gallen International Consensus on the primary therapy of early breast cancer 2015. *Annals of Oncology*, 26. 1533-1546.
- Cooper GM. The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates; 2000. The Development and Causes of Cancer. Disponible en: <https://www.ncbi.nlm.nih.gov/books/NBK9963/>
- Coughlin, S. S. y Hall, I. J., 2002. A review of genetic polymorphisms and prostate cancer risk. *Annals of Epidemiology*, 12, 182-196. *Bioinformatics*

- Craze, M.L., Cheung, H., Jewa, N., Coimbra, N. D. M., Soria, D., El-Ansari, R., Aleskandarany, M. A., Wai Cheng, K., Diez-Rodriguez, M., Nolan, C. C., Ellis, I. O., Rakha, E. A. y Green, A. R., 2018. MYC regulation of glutamine-proline regulatory axis is key in luminal B breast cancer. *British Journal of Cancer*, 118, 258-265.
- De Anda Jauregui, G., 2016. Búsqueda de blancos terapéuticos asociados a fenotipos de cáncer de mama. Tesis doctoral.
- DeBerardinis, R. J., y Chandel, N. S., 2016. Fundamentals of cancer metabolism. *Science advances*, 2, e1600200.
- DeBerardinis, R. J., Macuso, A., Daikhin, E., Nissim, I., Yudkoff, M., Wehrli, S. y Thompson, C., 2007. Beyond aerobic glycolysis: Transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *PNAS*, 104, 19345-19350.
- DeBerardinis, R., Chandel, N., 2016. Fundamentals of cancer metabolism. *Science Advances*, 2, e1600200.
- Dennison, J. B., Molina, J. R., Mitra, S., González-Angulo, A. M., Balko, J. M., Kuba, M. G., ... y Brown, R. E., 2013. Lactate dehydrogenase B: a metabolic marker of response to neoadjuvant chemotherapy in breast cancer. *Clinical Cancer Research*, 19, 3703-3713.
- Drier, Y., Sheffer, M., Domany, E., 2013. Pathway-based personalized analysis of cancer. *PNAS*, 110, 6388-6393.

- Eason, K., y Sadanandam, A., 2016. Molecular or metabolic reprogramming: what triggers tumor subtypes?. *Cancer research*, 76, 5195-5200.
- Eifert, C. y Powers, S., 2012. From cancer genomes to oncogenic drivers, tumour dependencies and therapeutic targets. *Nature Reviews Cancer*, 12, 572-578.
- Espinal-Enríquez, J., Fresno, C., DeAnda-Jauregui, G. y Hernández-Lemus, E., 2017. RNA-Seq based genome-wide analysis reveals loss of inter-chromosomal regulation in breast cancer. *Scientific Reports*, 7.
- Flowers, M., y Thompson, P. A., 2009. t10c12 conjugated linoleic acid suppresses HER2 protein and enhances apoptosis in SKBr3 breast cancer cells: possible role of COX2. *PLoS One*, 4, e5342.
- Fresno, C., González, G. A., Llera, A. S., & Fernández, E. A., 2016. pbcmc: Permutation-Based Confidence for Molecular Classification. *R package version*, 1.
- Gan, L., Yang, Y., Li, Q., Feng, Y., Liu, T. y Guo, W., 2018. Epigenetic regulation of cancer progression by EZH2: from biological insights to therapeutic potential. *Biomarker research*, 6.
- García-Campos, M.A., Espinal-Enríquez, J., y Hernández-Lemus, E., 2015. Pathway analysis: state of the art. *Frontiers in Physiology*, 6.
- García-Campos, M.A., 2016. Pa3BC: Pathway Analyses of Apoptosis and1. Autophagy on Breast Cancer, First release, DOI: 10.5281/zenodo.166107
- Gentric, G., Mieulet, V. y Mechta-Grigoriou, F., 2017. Heterogeneity in cancer metabolism: New concepts in an old field. *Antioxidant & Redox Signaling*, 26, 462-485.

- Giatromanolaki, A., Sivridis, E., Gatter, K. C., Turley, H., Harris, A. L., & Koukourakis, M. I., 2006. Lactate dehydrogenase 5 (LDH-5) expression in endometrial cancer relates to the activated VEGF/VEGFR2 (KDR) pathway and prognosis. *Gynecologic oncology*, 103, 912-918.
- Goldhirsch, A., Wood, W. C., Coates, A. S., Gelber, R. D., Thürlimann, B., Senn, H. J. and Panel members, 2011. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Annals of Oncology*, 22, 1736-1747.
- Gonzalzo, M., Nakayama, M., Lee, S., De Marzo, A. y Nelson W., 2004. Detection of GSTP1 methylation in prostatic secretions using combinatorial MSP analysis. *Urology*, 63, 414-418.
- Hanahan, D., & Weinberg, R., 2000. The hallmarks of cancer. *Cell Press*, 100, 57-60.
- Hanahan, D., & Weinberg, R., 2011. Hallmarks of cancer: The next generation. *Cell Press*, 144, 646-674.
- Hastie, T. & Stuetzle, W., 1989. Principal curves. *Journal of the American Statistical Association*, 84, 502–516.
- Hirschey, M., DeBerardinis, R., Diehl, A., Drew, J., Frezza, C., *et al.*, 2015. Dysregulated metabolism contributes to oncogenesis. *Seminars in Cancer Biology*, 35, S129-S150.

- Hu, Z., Fan, C., Oh, D. S., Marron, J. S., He, X., Qaqish, B. F., ... y Nobel, A., 2006. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC genomics*, 7, 96.
- Jeon, Y., Heo, Y. S., Kim, C., Hyun, Y. L., Lee, T., Ro, S., *et al.*, 2005. Phosphodiesterase: overview of protein structures, potential therapeutic applications and recent progress in drug development. *Cellular and Molecular Life Sciences*, 62, 1198-1220.
- Jia, D., Lu, M., Jung, K. H., Park, J. H., Yu, L., Onuchic, J. N., ... y Levine, H. 2019. Elucidating cancer metabolic plasticity by coupling gene regulation with metabolic pathways. *Proceedings of the National Academy of Sciences*, 116, 3909-3918.
- Jiang, W., Zhou, F., Li, N., Li, Q., y Wang, L., 2015. FOXM1-LDHA signaling promoted gastric cancer glycolytic phenotype and progression. *International journal of clinical and experimental pathology*, 8, 6756.
- Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., and Tanabe, M., 2019. New approach for understanding genome variations in KEGG. *Nucleic Acids Research*. 47, D590-D595.
- Kitano, H., 2000. Perspectives on systems biology. *New Generation Computing*, 18, 199–216.
- Kittaneh, M., Montero, A.J., y Glück, S., 2013. Molecular profiling for breast cancer: a comprehensive review. *Biomarkers in Cancer*, 5, 61-70.

- Kleer, C.G., Cao, Q., Varambally, S., Shen, R., Ota, I., Tomlins, S.A., Ghosh, D., Sewalt, R.G., Otte, A.P., Hayes, D.F., *et al.*, 2003. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proceedings of the National Academy of Sciences, USA* 100, 11606–11611.
- Klepinin, A., Ounpuu, L., Guzun, R., Chekulayev, V., Timohhina, N., Tepp, K., *et al.* 2016. Simple oxygraphic analysis for the presence of adenylate kinase 1 and 2 in normal and tumor cells. *Journal of bioenergetics and biomembranes*, 48, 531-548.
- Kumaran, M., Cass, C., Graham, K., Mackey, J., Hubaux, R., Lam, W., Yasui, Y. y Damaraju, S., 2017. Germline copy number variations are associated with breast cancer risk and prognosis. *Scientific Reports*, 7, 1-15.
- Kurabe, N., Hayasaka, T., Ogawa, M., Masaki, N., Ide, Y., Waki, M., Nakamura, T., Kurachi, K., Kahyo, T., Shinmura, K., Midorikawa, Y., Sugiyama, Y., Setou, M. y Sugimura, H., 2013. Accumulated phosphatidylcholine (16:0/16:1) in human colorectal cancer; possible involvement of LPCAT4. *Cancer Science*, 104, 1295-1302.
- Lachapelle, J. y Foulkes, W. D., 2011. Triple-negative and basal-like breast cancer: implications for oncologists. *Current oncology*, 18, 161-164.
- Laible, M., Hartmann, K., Gürtler, C., Anzeneder, T., Wirtz, R., Weber, S., ... y Ramaswamy, A., 2019. Impact of molecular subtypes on the prediction of distant recurrence in estrogen receptor (ER) positive, human epidermal growth factor

receptor 2 (HER2) negative breast cancer upon five years of endocrine therapy. *BMC cancer*, 19, 694.

- Le, A., Cooper, C. R., Gouw, A. M., Dinavahi, R., Maitra, A., Deck, L. M., ... y Dang, C. V. 2010. Inhibition of lactate dehydrogenase A induces oxidative stress and inhibits tumor progression. *Proceedings of the National Academy of Sciences*, 107, 2037-2042.

- Lee, K. W., Lee, H. J., Cho, H. Y. Y Kim, Y. J., 2005. Role of the conjugated linoleic acid in the prevention of cancer. *Critical Reviews in Food Science and Nutrition*, 45, 135-144.

- Li, J., Liu, X., Wang, H., Zhang, W., Chan, D. y Shi, Y., 2012. Lysocardiolipin acyltransferase 1 (ALCAT1) controls mitochondrial DNA fidelity and biogenesis through modulation of MFN2 expression. *PNAS*, 18, 6975-6980

- Livshits, A., Git, A., Fuks, G., Caldas, C., y Domany, E. (2015). Pathway-based personalized analysis of breast cancer expression data. *Molecular oncology*, 9(7), 1471-1483.

- Lu, X., He, G., Yu, H., Ma, Q., Shen S., Das, U., 2010. Colorectal cancer cell growth inhibition by linoleic acid is related to fatty acid composition changes. *Journal of Zhejiang University Science B*, 1, 923-930.

- Lunt, S. Y., y Vander Heiden, M. G., 2011. Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. *Annual review of cell and developmental biology*, 27, 441-464.

- Luo, W. y Brouwer, C., 2013. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics*, 29, 1830-1831.
- Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D. y Woolf, P.J., 2009. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*
- Mejía-Pedroza, R. A., Espinal-Enríquez, J., y Hernández-Lemus, E., 2018. Pathway-based drug repositioning for breast cancer molecular subtypes. *Frontiers in pharmacology*, 9, 905.
- Mishra, D., y Banerjee, D., 2019. Lactate dehydrogenases as metabolic links between tumor and stroma in the tumor microenvironment. *Cancers*, 11, 750.
- Morikawa, T., Maeda, D., Kune, H., Homma, Y. y Fukuyama, M., 2010. Ribonucleotide reductase M2 subunit is a novel diagnostic marker and a potential therapeutic target in bladder cancer. *Histopathology*, 57, 885-892.
- Murtagh, F. and Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion?. *Journal of Classification*, 31, 274-295.
- Muthusamy, S., Andersson, S., Kim, H. J., Buttler, R., Waage, L., Bergerhelm, U. y Gustafsson, J. A., 2011. Estrogen receptor  $\beta$  and 17 $\beta$ -hydroxysteroid dehydrogenase type 6, a growth regulatory pathway that is lost in prostate cancer. *PNAS*, 108, 20090-200904.
- Nielsen, J., 2003. It is all about metabolic fluxes. *Journal of bacteriology*, 185, 7031-7035.

- Nueda, M. J., Ferrer, A., & Conesa, A., 2012. ARSyN: a method for the identification and removal of systematic noise in multifactorial time course microarray experiments. *Biostatistics*, 13, 553-566.
- Parker, J. S., Mullins, M., Cheang, M. C., Leung, S., Voduc, D., Vickery, T., ... y Quackenbush, J. F., 2009. Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of clinical oncology*, 27, 1160.
- Pavlova, N., y Thompson, C., 2016. The emerging hallmarks of cancer metabolism. *Cell metabolism*, 23, 27-47.
- Perou, C. M., Sørlie, T., Eisen, M. B., Van De Rijn, M., Jeffrey, S. S., Rees, C. A., ... y Fluge, Ø., 2000. Molecular portraits of human breast tumours. *Nature*, 406, 747-752.
- Rahman, M. A., Amin, R. M., Wang, D. S., Koenig, L., Nannapanemi, S., Chen, Z. J. *et al.*, 2013. RRM2 regulates Bcl-2 in head and neck and lung cancers: A potential target for cancer therapy. *Clinical Cancer Research*, 19, 3416-3428.
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., Smyth G. K., 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43.
- Robinson, M. y Oshlack, A., 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, 11, R25
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

- Ross, J. S., y Fletcher, J. A., 1998. The HER-2/neu oncogene in breast cancer: prognostic factor, predictive factor, and target for therapy. *Stem cells*, 16, 413-428.
- Santuario-Facio, S., Cardona-Huerta, S., Perez-Paramo, Y., Trevino, V., Hernandez-Cabrera, F., Rojas-Martinez, A., et al., 2017. A new gene expression signature for triple-negative breast cancer using frozen fresh tissue before neoadjuvant chemotherapy. *Molecular medicine*, 23, 101-111.
- Sadana, R. y Dessauer, C. W., 2009. Physiological roles for G protein-regulated adenylyl cyclase isoforms: insights from knockout and overexpression studies. *Neurosignals*, 17, 5-22.
- Saraei, P., Asadi, I., Kakar, M. A., y Moradi-Kor, N., 2019. The beneficial effects of metformin on cancer prevention and therapy: a comprehensive review of recent advances. *Cancer Management and Research*, 11, 3295–313.
- Schilling, C., Schuster, S., Palsson, B. y Heinrich, R., 1999. Metabolic pathway analysis: Basic concepts and scientific applications in the post-genomic era. *Biotechnology Progress*, 15, 296-303.
- Schlichtholz, B., Turyn, J., Goyke, E., Biernacki, M., Jaśkiewicz, K., Śledziński, Z., y Swierczynski, J. 2005. Enhanced citrate synthase activity in human pancreatic cancer. *Pancreas*, 30, 99-104.
- Schnekenburger, M., Karius, T. y Diederich, M., 2014. Regulation of epigenetic traits of the glutathione S-transferase P1 gene: From detoxification toward cancer prevention and diagnosis. *Frontiers in Pharmacology*, 5, 170.

- Shao, X., Lv, N., Liao, J., Long, J., Xue, R., Ai, N., ... y Fan, X. 2019. Copy number variation is highly correlated with differential gene expression: a pan-cancer study. *BMC medical genetics*, 20, 175.
- Sheng, S. L., Liu, J. J., Dai, Y. H., Sun, X. G., Xiong, X. P., & Huang, G., 2012. Knockdown of lactate dehydrogenase A suppresses tumor growth and metastasis of human hepatocellular carcinoma. *The FEBS journal*, 279, 3898-3910.
- Sims, A. H., Howell, A., Howell, S. J., y Clarke, R. B., 2007. Origins of breast cancer subtypes and therapeutic implications. *Nature Clinical Practice Oncology*, 4, 516-525.
- Simpson, P., Reis-Filho, J., Gale, T., Lakhani, S., 2005. Molecular evolution of breast cancer. *The Journal of Pathology*, 205, 248-254.
- Smith, G. K., 2004. Linear models and empirical bayes methods for assessing expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3, 1-25.
- Smyth, G. K. (2005). Limma: linear models for microarray data. In *Bioinformatics and computational biology solutions using R and Bioconductor* (pp. 397-420). Springer, New York, NY.
- Sørlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S. et al., 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences*, 98, 10869–10874.

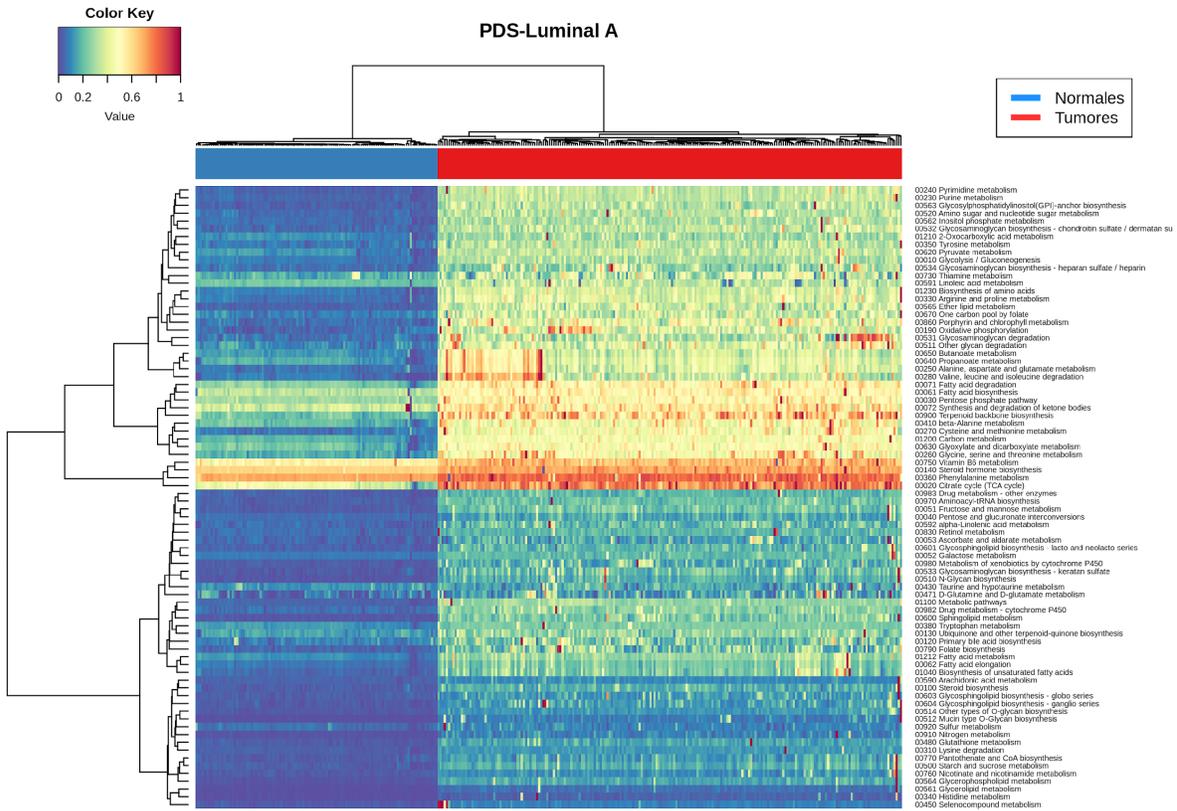
- Sørli, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S. et al., 2003. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proceedings of the National Academy of Sciences*, 100, 8418–8423.
- Su, Y., Yu, Q. H., Wang, X. Y., Yu, L. P., Wang, Z. F., Cao, Y. C., y Li, J. D., 2017. JMJD2A promotes the Warburg effect and nasopharyngeal carcinoma progression by transactivating LDHA expression. *BMC cancer*, 17, 477.
- Sun, H., Yang, B., Zhang, H., Song, J., Zhang, Y., Xing, J., Yang, Z., Wei, C., Xu, T., Yu, Z., Xu, Z., Hou, M., Ji, M. y Zhang, Y., 2019. RRM2 is a potential prognostic biomarker with functional significance in glioma. *International Journal of Biological Sciences*, 15, 533-543.
- Stratton, M., Campbell, P. and Futreal, A., 2009. The cancer genome. *Nature*, 458, 719-724.
- Tavassoly, I., Goldfarb, J. y Iyengar, R., 2018. Systems biology primer:the basic methods and approaches. *Essays in Biochemistry*, 62, 487-500.
- Valvona, C. J., Fillmore, H. L., Nunn, P. B., y Pilkington, G. J., 2016. The regulation and function of lactate dehydrogenase a: therapeutic potential in brain tumor. *Brain pathology*, 26, 3-17.
- Vander Heiden, M., Cantley, L., and Thompson, C., 2009. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science*, 324, 1029-1033.

- Vogelstein, B. y Kinzler, K., 2004. Cancer genes and the pathways they control. *Nature Medicine*, 10, 789-799.
- Wang, Z., Gerstein, M., y Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, 10, 57-63.
- Wolf, A., Agnihotri, S., Micallef, J., Mukherjee, J., Sabha, N., Cairns, R., ... y Guha, A., 2011. Hexokinase 2 is a key mediator of aerobic glycolysis and promotes tumor growth in human glioblastoma multiforme. *Journal of Experimental Medicine*, 208, 313-326.
- Wu, G., Fang, Y. Z., Yang, S., Lupton, J. R. y Turner, N. D., 2004. Glutathione metabolism and its implications for health. *The Journal of Nutrition*, 134, 489-492.
- Wu, G., Feng, X. & Stein, L., 2010. A human functional protein interaction network and its application to cancer data analysis. *Genome Biology*, 11, 1-23.
- Wu, T., Wang, Y., Jiang, R., Lu, X., y Tian, J., 2017. A pathways-based prediction model for classifying breast cancer subtypes. *Oncotarget*, 8, 58809.
- Xie, Z., Zhang, Y., Jin, C., y Fu, D., 2018. Gemcitabine-based chemotherapy as a viable option for treatment of advanced breast cancer patients: a meta-analysis and literature review. *Oncotarget*, 9, 7148.
- Yamada, T., Komoto, J., Kasuya, T., Takata, Y., Ogawa, H., Mori, H. y Takusagawa, F., 2008. A catalytic mechanism that explains a low catalytic activity of serine dehydratase like-1 from human cancer cells: Crystal structure and site-directed mutagenesis studies. *Biochimica et Biophysica Acta – General Subjects*, 1780, 809-18.

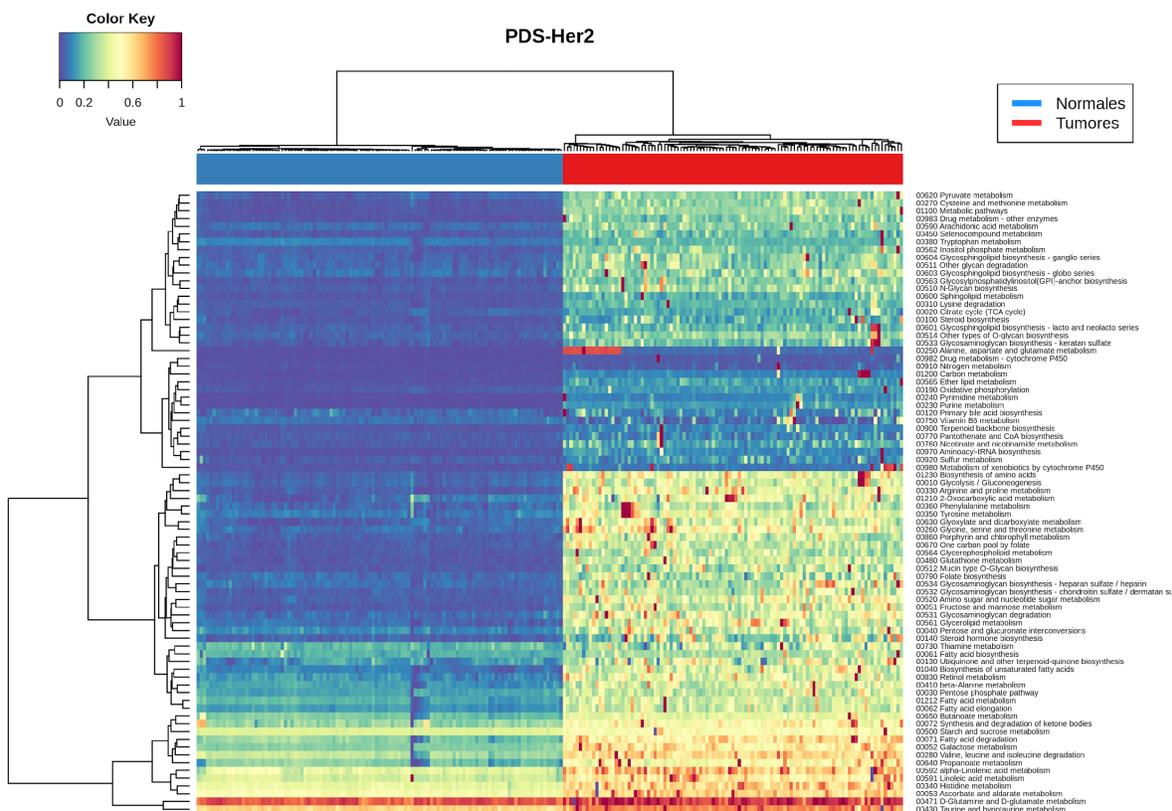
- Yan, K. S., Lin, C. Y., Liao, T.W., Peng, C.M., Lee, S. C., Liu, Y. J., Chan, W. P. y Chou, R. H., 2017. EZH2 in cancer progression and potential application in cancer therapy: A friend or foe? *International Journal of Molecular Sciences*, 18, 1172.
- Yoshida, G. J. 2015. Metabolic reprogramming: the emerging concept and associated therapeutic strategies. *Journal of experimental & clinical cancer research*, 34, 1-10.
- Ždravlević, M., Marchiq, I., de Padua, M. M. C., Parks, S. K., y Pouyssegur, J. 2017. Metabolic plasticity in cancers—distinct role of glycolytic enzymes GPI, LDHs or membrane transporters MCTs. *Frontiers in oncology*, 7, 313.
- Zhang, Y., y Yang, J.M. 2013. Altered energy metabolism in cancer. A unique opportunity for therapeutic intervention. *Cancer Biology & Therapy*, 14, 81-89.

# 10. Anexos

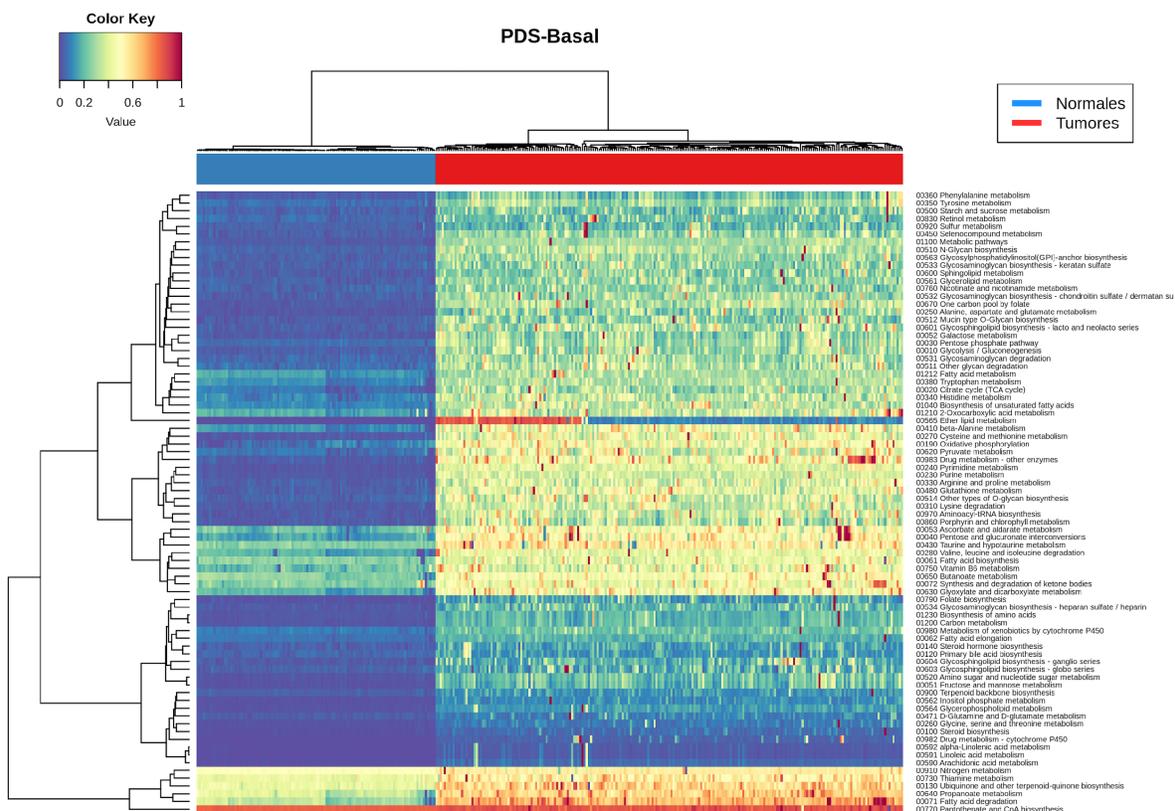
## Anexo A



Mapa de calor. Representa el PDS (Pathway Deregulation Score) del análisis de Pathfier para las muestras con subtipo luminal A. Cada columna corresponde a una muestra y cada fila corresponde a una vía. Las vías y las muestras se agrupan de acuerdo a su valor de PDS. El color azul representa un bajo puntaje (no desregulación) y el color rojo representa un alto puntaje (mayor desregulación). La barra superior representa, en azul, a las muestras controles y en rojo a las muestras tumorales.



Mapa de calor. Representa el PDS (Pathway Deregulation Score) del análisis de Pathfier para las muestras con subtipo Her2. Cada columna corresponde a una muestra y cada fila corresponde a una vía. Las vías y las muestras se agrupan de acuerdo a su valor de PDS. El color azul representa un bajo puntaje (no desregulación) y el color rojo representa un alto puntaje (mayor desregulación). La barra superior representa, en azul, a las muestras controles y en rojo a las muestras tumorales.



Mapa de calor. Representa el PDS (Pathway Deregulation Score) del análisis de Pathfier para las muestras con subtipo basal. Cada columna corresponde a una muestra y cada fila corresponde a una vía. Las vías y las muestras se agrupan de acuerdo a su valor de PDS. El color azul representa un bajo puntaje (no desregulación) y el color rojo representa un alto puntaje (mayor desregulación). La barra superior representa, en azul, a las muestras controles y en rojo a las muestras tumorales.



# Targeting Metabolic Deregulation Landscapes in Breast Cancer Subtypes

Erandi A. Serrano-Carbajal<sup>1</sup>, Jesús Espinal-Enriquez<sup>1,2\*</sup> and Enrique Hernández-Lemus<sup>1,2\*</sup>

<sup>1</sup> Computational Genomics Division, National Institute of Genomic Medicine, Mexico City, Mexico, <sup>2</sup> Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, Mexico City, Mexico

## OPEN ACCESS

### Edited by:

Nadia Judith Jacobo-Herrera,  
Instituto Nacional de Ciencias  
Médicas y Nutrición Salvador Zubirán  
(INCMNSZ), Mexico

### Reviewed by:

Cinzia Antognelli,  
University of Perugia, Italy  
Lajos Pusztai,  
School of Medicine, Yale University,  
United States

### \*Correspondence:

Jesús Espinal-Enriquez  
jespinal@inmegen.gob.mx  
Enrique Hernández-Lemus  
ehernandez@inmegen.gob.mx

### Specialty section:

This article was submitted to  
Cancer Metabolism,  
a section of the journal  
Frontiers in Oncology

**Received:** 17 August 2019

**Accepted:** 20 January 2020

**Published:** 11 February 2020

### Citation:

Serrano-Carbajal EA,  
Espinal-Enriquez J and  
Hernández-Lemus E (2020) Targeting  
Metabolic Deregulation Landscapes in  
Breast Cancer Subtypes.  
*Front. Oncol.* 10:97.  
doi: 10.3389/fonc.2020.00097

Metabolic deregulation is an emergent hallmark of cancer. Altered patterns of metabolic pathways result in exacerbated synthesis of macromolecules, increased proliferation, and resistance to treatment via alteration of drug processing. In addition, molecular heterogeneity creates a barrier to therapeutic options. In breast cancer, this broad variation in molecular metabolism constitutes, simultaneously, a source of prognostic and therapeutic challenges and a doorway to novel interventions. In this work, we investigated the metabolic deregulation landscapes in breast cancer molecular subtypes. Such landscapes are the regulatory signatures behind subtype-specific metabolic features.  $n = 735$  breast cancer samples of the Luminal A, Luminal B, Her2+, and Basal subtypes, as well as  $n = 113$  healthy breast tissue samples were analyzed. By means of a single-sample-based algorithm, deregulation for all metabolic pathways in every sample was determined. Deregulation levels match almost perfectly with the molecular classification, indicating that metabolic anomalies are closely associated with gene-expression signatures. Luminal B tumors are the most deregulated but are also the ones with higher within-subtype variance. We argued that this variation may underlie the fact that Luminal B tumors usually present the worst prognosis, a high rate of recurrence, and the lowest response to treatment in the long term. Finally, we designed a therapeutic scheme to regulate purine metabolism in breast cancer, independently of the molecular subtype. This scheme is founded on a computational tool that provides a set of FDA-approved drugs to target pathway-specific differentially expressed genes. By providing metabolic deregulation patterns at the single-sample level in breast cancer subtypes, we have been able to further characterize tumor behavior. This approach, together with targeted therapy, may open novel avenues for the design of personalized diagnostic, prognostic, and therapeutic strategies.

**Keywords:** cancer metabolism, pathway deregulation, breast cancer subtypes, therapeutic targets, steroid and fatty acid metabolism, purine metabolism

## 1. INTRODUCTION

Breast cancer is a complex, heterogeneous disease. Manifestations of this heterogeneity can be observed at the transcriptomic, molecular, or histological level (1). The origins of such manifestations can be traced back by looking at different levels of molecular control within the cells and tissues. The mechanisms behind gene expression, cell signaling, and metabolism are highly

intertwined, and cross-regulation patterns appear (2, 3), which strongly determine the phenotypic variance observed in clinical practice (4–6). In fact, this broad variance in molecular metabolism in breast cancer constitutes, simultaneously, a source of prognostic and therapeutic challenges and a doorway to novel interventions (7–9).

In order to face the challenges posed by tumor heterogeneity, it is customary to classify or *subtype* tumors according to their feature similarity. One currently used classification method in breast cancer, which has been particularly useful for capturing biological functional features, is the so-called molecular subtyping (10). The default classification scheme in this regard is given by the PAM50 (10, 11) algorithm, which groups breast tumors into molecular classes or subtypes according to a gene-expression signature of 50 genes relevant to the patho-physiology of the tumor. These subtypes are *Luminal A*, *Luminal B*, *Her2+*, and *Basal*. Some authors include a fifth subtype, the so-called *Normal-like*, but its use is controversial, and its use has been in decline lately (12).

These subtypes have been able to capture relevant differences in the origin, prognosis, response to treatment, and relapse probability of breast tumors. In general, it is considered that luminal subtypes are less aggressive and have better prognosis and better response to treatment than non-luminal ones (11). However, under certain circumstances, Luminal B tumors may have a higher recurrence, less response to treatment, and worse long-term prognosis (13). This variation in response is not clear and is of the utmost importance for the understanding of the disease at the personalized level.

Genomic alterations (mutations, copy number variations, chromosomal aberrations) often derive into anomalous cell functioning, including deregulation of metabolism—an important emergent hallmark of cancer (14) via abnormal gene regulatory programs. Aberrant gene-expression patterns are currently studied using next-generation sequencing (NGS) techniques such as RNA-Seq.

The analysis of these gene deregulation signatures provides a comprehensive (genome-wide) approach to dig into the molecular basis of disease. In the case of tumor metabolism, one may argue that metabolomics and phospho-proteomics would be closer proxies to the actual underlying molecular mechanisms. However, despite important advances in experimental-omic techniques, comprehensive metabolomic mapping and fluxomics are still under-developed for the task of describing cellular metabolic processes comprehensively, although this should change in the upcoming years. Approaches to analyzing metabolic deregulation in cancer based on gene expression have been developed (15, 16). Those extensive studies used differentially expressed genes for more than 20 types of cancer to distinguish deregulated metabolic pathways. In both cases, specific pathways were identified as deregulated in particular types of cancer. However, those studies performed phenotype-specific analyses and did not focus on single-sample deregulation.

To overcome this issue, an appealing way to study deregulation of metabolism is by analyzing metabolism-related gene-expression signatures at a single-sample level. In this work, we used TCGA gene-expression data from 735 tumor samples

(17, 18), classified according to their molecular signature, to investigate the pathway deregulation patterns for the four PAM50 molecular subtypes, to determine subtype-specific metabolic landscapes. We used a single-sample-based algorithm (19) to quantify metabolic anomalies. This algorithm provides a pathway deregulation score for each pathway at a sample level. For validation purposes, we used a 2,000-sample cohort (20) with the same pipeline. Analyzing metabolic deregulation patterns at the subtype and individual sample levels provides a means of characterizing tumor behavior with a view to designing personalized diagnostic, prognostic, and therapeutic strategies.

## 2. MATERIALS AND METHODS

### 2.1. RNASeq Data Acquisition and Processing

Data were acquired from the Genome Data Commons Data Portal (<https://bit.ly/2lJrJgi>).

Briefly, 1,102 primary breast tumors and 113 normal solid tissues (normal solid tissue refers to healthy tumor-adjacent tissue taken from some of the tumors) samples were acquired and pre-processed to obtain  $\log_2$  normalized gene-expression values (21). Data were pre-processed to eliminate intrinsic experimental biases (22).

#### 2.1.1. Integration

The following pipeline was already used and reported in Espinal-Enríquez et al. (21). Basically, an integrity check had to be carried out on raw expression files to ensure that all of them both had the same dimensions and provided TCGA identifiers before complementary annotation could be incorporated.

#### 2.1.2. Quality Control

The NOISEq R library was used for global quality control (23, 24). All samples reached saturation for the number of detected features at the corresponding sequencing depth. Global expression quantification for each experimental condition yielded a feature sensitivity >60% for 10 counts per million (CPM). Bias detection assessment showed the presence of gene length, %GC, and RNA patterns.

The EDASeq R library was used for batch-effect removal (25). Before normalization, genes with mean counts <10 were filtered, resulting in 17,215 genes, as suggested in Risso et al. (25). Different within/between **normalization** strategies were tested to remove bias.

Exploration of sample  $\log_2$ (normalized count) expression densities showed a consistent bi-modal pattern, corresponding to **noisy** lower-expressed genes and global sample behavior. Filtering out features with low counts ( $CPM < 10$  cut-off) retained 15,281 genes, removing the undesired lower density peak. Finally, ARSYN R library was used for multidimensional noise reduction using default parameters (22).

#### 2.1.3. Subtyping

We classified the 1,112 breast cancer samples into the four molecular subtypes using the pbcmc R package (26), a variation of the PAM50 algorithm, which characterizes the assessment

of the uncertainty in gene-expression-based classifiers (e.g., PAM50) based on permutation tests (12). Tumor samples with a non-reliable breast cancer subtype call were removed from the analysis. The number of removed samples was 377, giving a final number of 735 reliable samples.

## 2.2. Differential Expression Analysis and Pathway Discrimination

To determine overexpressed or underexpressed genes, we used the *limma* R package (27), considering an absolute difference of  $\text{Log}_2\text{FoldChange} > 1$  and a  $B\text{-statistic} > 5$ . The False Discovery Rate-adjusted  $p$ -value threshold was  $10^{-3}$ . Since the main goal of this work is to establish the extent of deregulation in the metabolism for each breast cancer sample/subtype, we kept 80 metabolic pathways present in the KEGG database (28) (the Pathifier algorithm needs a minimum number of molecules to be performed).

## 2.3. Pathway Deregulation Analysis

Metabolic pathway deregulation in each sample was quantified by using the Pathifier algorithm (19). This algorithm integrates the expression data of genes involved in a given metabolic pathway into a single deregulation value at the individual-sample level. The algorithm assigns a score between 0 and 1, called the Pathway Deregulation Score (PDS). Values close to 0 correspond to samples whose expression levels are similar to controls (29). Samples with higher values present higher differences in expression levels compared to the control group. Pathifier quantifies the level of deregulation of a metabolic pathway in a single tumor sample by measuring the deviation of said sample from control behavior.

In some cases, a single sample with extreme gene-expression changes (majorly different from those of other samples) for genes in a given pathway may give rise to a really high (assigning PDS = 1 to that sample) score, making all other deregulated samples (with large but comparatively low gene-expression changes) close to zero, thus appearing to be minorly deregulated. In other words, several deregulated pathways/samples would be missing. In such cases, the outlier sample was removed from the analysis. Finally, an unsupervised clustering method was used to group samples with similar PDS. A graphical representation of the pipeline is presented here in Figure 1.

## 2.4. Identification of Potential Pharmacological Targets

Genes being commonly over/underexpressed in all breast cancer subtypes would suggest that there should be subtype-independent drugs. In order to assess this idea, we performed data mining on transcriptomic/drug data by using a previously developed (by our group) computational pipeline to find differentially expressed pharmacological targets of FDA-approved drugs (31) for those shared DEGs. This tool performs all possible combinations of differentially expressed targets and FDA-approved drugs in public pharmacological databases, as well as their two-drug interactions. So, for the more than 2611 drugs annotated in the DrugBank database and the 660

drugs annotated in PharmGKB, all subtype-specific differentially expressed genes were interrogated.

## 2.5. Validation

For validation purposes, we used 2,000 microarray samples from the METABRIC cohort (20), performed the same analysis with the already classified samples, obtained the single-sample PDS, and compared them with the TCGA cohort.

## 3. RESULTS AND DISCUSSION

### 3.1. Subtype-Specific Deregulated Genes Are Associated With Characteristic Metabolic Pathways

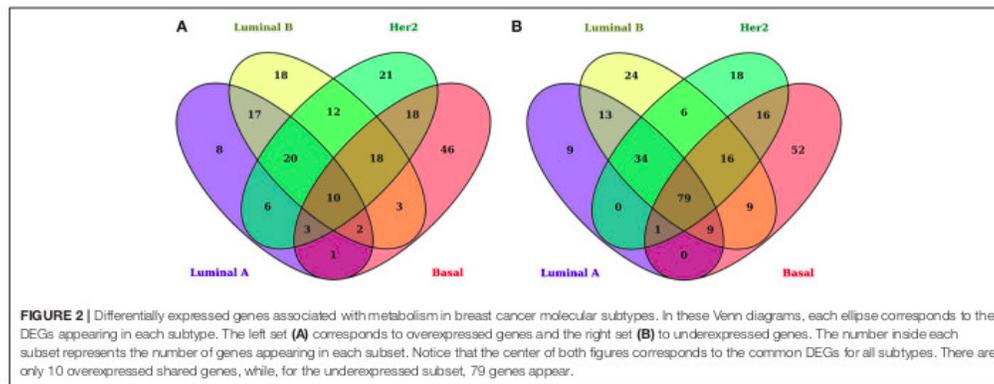
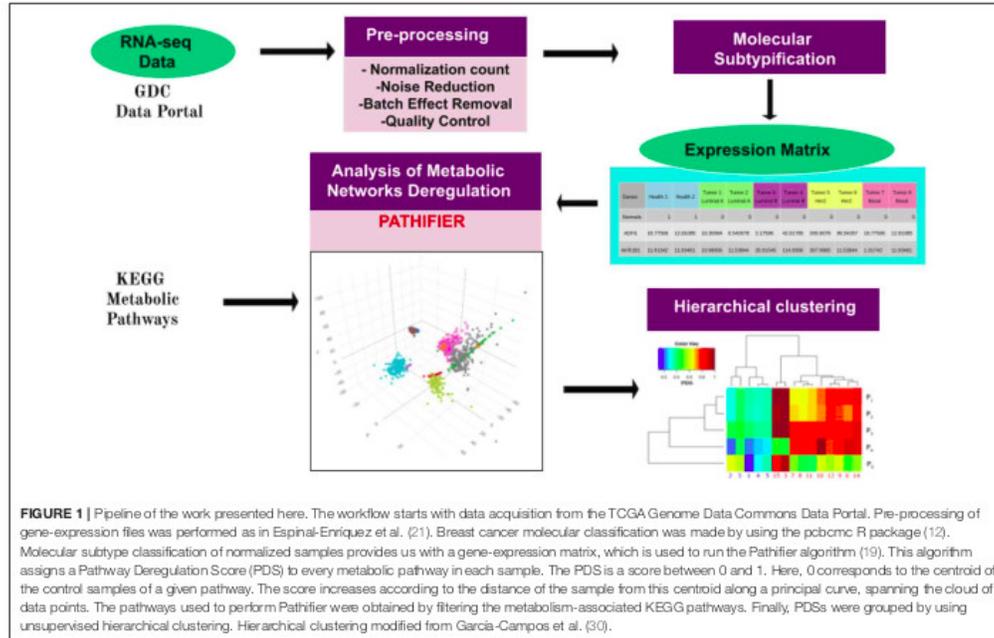
As has been observed previously (1, 6), gene-expression signatures differ between all subtypes (Figure 2). The signatures presented here include only genes associated with metabolic pathways. Figure 2 shows the overexpressed and underexpressed metabolism-associated genes for each subtype in the form of a Venn diagram. It can be observed that all subtypes have a non-shared set of differentially expressed genes (DEGs) but also a small subset of shared deregulated genes.

By using  $|\log_2\text{FoldChange}| > 1$  and  $B\text{-statistic} > 5$  as significance thresholds, the number of DEGs in all the tumors is 204 overexpressed and 287 underexpressed. The numbers of overexpressed and underexpressed genes for each subtype are very similar. Interestingly, the subset of shared overexpressed genes ( $n = 10$ ) is substantially smaller than that of the underexpressed genes ( $n = 79$ ). This difference between the number of shared underexpressed and overexpressed genes may be associated with the fact that some metabolic pathways are silenced or decreased in all subtypes; on the other hand, metabolic pathways with incremental activity are subtype-specific.

To evaluate whether shared overexpressed genes influence the regulation of metabolism, we associated them with the metabolic processes in which they participate. Figure 3 shows the relationships between the overexpressed genes (in red), and their associated metabolic processes (in pink) in the form of a bipartite network—a network composed by nodes of different nature, in this case, genes and pathways. Analogously, we constructed a network composed of the common underexpressed genes and their associated metabolic pathways.

As can be seen from the structure of the bipartite network, there are central molecules involved in several interrelated metabolic processes, giving rise to the so-called pathway-crosstalk events. This is a result of the utmost importance, since crosstalk phenomena have been associated with anomalous therapeutic responses and pharmacological resistance in breast cancer subtypes (32).

We can see, for instance, how the Interleukin 4-induced 1 gene (IL4I1) is the one with the most associated metabolic processes ( $n = 7$ ), all related to amino acid biosynthesis (Figure 3A). This gene is often overexpressed in B-cell lymphomas (33) and has also been associated with cancer by promoting tumor growth



and shaping the immune microenvironment in melanoma (34). Autoimmune suppression and the inhibition of *CD8+* cells are also pro-tumor-associated mechanisms regulated by IL4I1 (35, 36). Such processes are ultimately linked to the metabolic activity of IL4I1 as a phenylalanine oxidase. Crosstalk events involving cross-regulation via IL4I1 and non-coding RNAs have also been reported to play a role in triple-negative breast cancer (37).

As can be observed from **Figure 3B**, common underexpressed genes participate collectively in specific metabolic processes, such as purine metabolism. This pathway provides the metabolites needed for survival and cell proliferation and DNA and RNA production (38). ATP and GTP are also products of this metabolic pathway.

Among the underexpressed genes, we may find *ADCY* genes (*ADCY4* and *ADCY5*), which regulate the nucleotide proportion



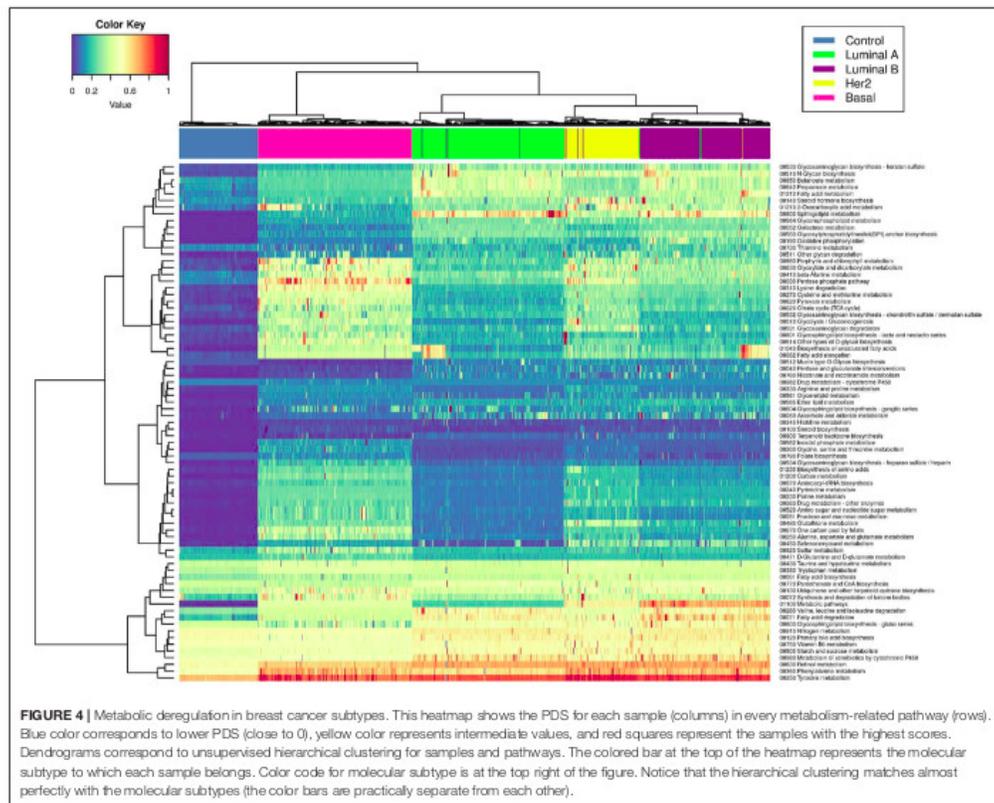


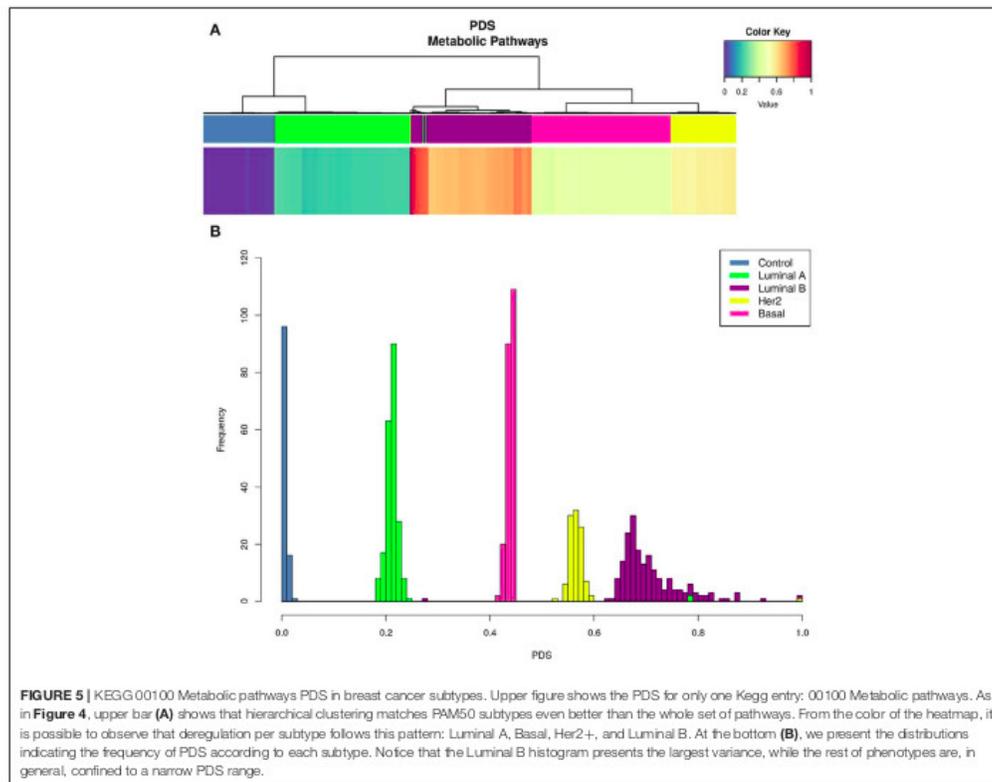
Figure 4 shows a heatmap of the PDS values (see methods) grouped by PDS similarity. Rows correspond to all pathways associated with metabolism, while columns correspond to samples. There are subsets of samples that present a similar metabolic deregulation among subgroups and differ from the other samples.

Interestingly, unsupervised hierarchical clustering of PDS coincides almost absolutely with the PAM50 classification. The colored bars in the upper part of the figure correspond to each subtype, and, as can be appreciated, each color of the bar is grouped together. This phenomenon reflects the high specificity of metabolic deregulation for each molecular subtype.

Figure 5 shows that only one KEGG pathway: 01100 Metabolic Pathways" contains the full set of 1,142 genes present in every metabolism-associated KEGG pathway. Hence, the PDS in this particular process summarizes (to a certain degree) information about the rest of the metabolism-related pathways. The PDS for each subtype again presents a subtype-specific behavior, but more widespread than in Figure 4.

The PDS values are different between all subtypes, but more importantly, it is clear to observe that Luminal B is the subtype with the highest PDS. This result was unexpected, since it is usually considered that the most aggressive and with worst prognosis is the Basal subtype (42). In this case, the order of deregulation is as follows: Luminal A, Basal, HER2+, and, finally, Luminal B (Figure 5). From the PDS distributions, it can be noticed that the Luminal B subtype has the highest values but also the largest variance between samples. The rest of the subtypes are highly concentrated in a narrow range of PDS.

Previous reports have also analyzed the relationship between transcriptional deregulation and metabolic changes in cancer (15, 16). From these studies, some commonalities and differences arise. The work of Rosario uses differentially expressed genes for several phenotypes, breast cancer subtypes included. There, a score is based on LogFoldChange and adjusted  $p$ -values, measures that have not been derived with pathway-level assessment in mind, in contrast with the PDS, which is a specific pathway-level measure.



Regarding commonalities, metabolic pathways are found to be differentially regulated in all subtypes in both manuscripts, in spite of the different approaches to pathway scoring. Purine and retinol metabolism are also found to be highly deregulated in both studies, particularly in the Luminal B and Basal subtypes. Interestingly enough, the Luminal B and Basal subtypes are the most deregulated phenotypes in both studies. This is reflected in Figure 6d from Rosario's paper and in **Figure 4** in our manuscript.

Another point in common between both studies is the coincidence of the Citric acid cycle as a unique pathway observed in the Basal subtype, with the TCA cycle found in our Basal samples (**Figure 4**). Interestingly, the categories reported in Figures 6d-f of Rosario's paper correspond to those of the Reactome database and not the ones described in the KEGG database. This is relevant since the categories are similar but not identical. This may be an additional source of some apparent discrepancies between Rosario's results and ours.

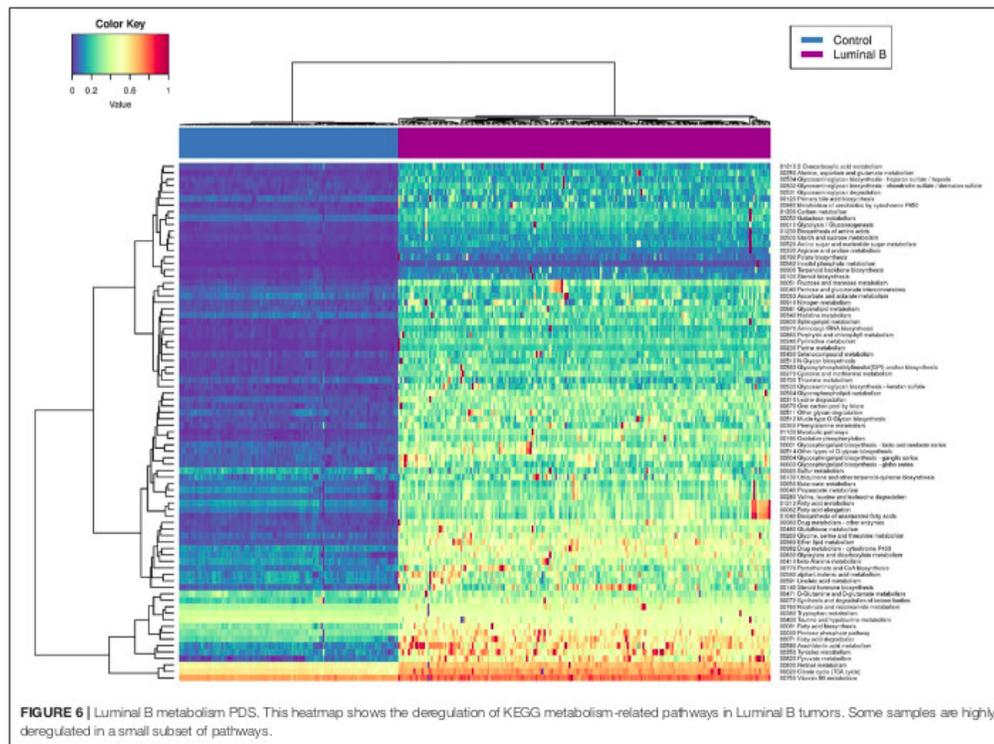
Regarding differences, Rosario et al. found different pathway scores for the Basal and Luminal A subtypes. However, as can

be seen from Figure 6C of Rosario's paper, the low specificity of the average gene-expression Z-scores results in a non-conclusive depiction, as it is hard to distinguish signal from background noise. This is also reflected in the density plot of the Figure 6C heatmap. Additionally, the hierarchical clustering on top of the heatmap reflects a large degree of heterogeneity, resulting from the broad variance of the average gene-expression profiles. However, a clear phenotypic fingerprint of basal tumors is actually captured in terms of average gene-expression profiles, likely due to a reduced heterogeneity in these tumors.

### 3.3. Luminal B Tumors Present Higher Pathway Deregulation Scores

**Figure 6** represents the PDSs for the Luminal B subtype only. It can be observed that several metabolic processes are highly deregulated (reddish rows), such as is the case of pyruvate metabolism, tyrosine metabolism, fatty acid degradation, and the pentose phosphate pathway.

In some cases, only a small subgroup of samples presents high PDSs (scattered red pixels), which in turn reflects the intrinsic



heterogeneity of samples, even if they belong to the same subtype. In the following, we will make some remarks regarding the most deregulated metabolic pathways observed in Luminal B tumors.

Pyruvate-related metabolic reprogramming has been associated with metastatic potential and treatment resistance in cancer (43). Pyruvate is a central metabolite for glucose, lactate, lipids, and amino acids. In breast cancer, liver-metastatic breast cancer cells exhibit a unique metabolic program compared to bone- or lung-metastatic cells, converting glucose-derived pyruvate into lactate, with a concomitant reduction in glutamine. This metabolic reprogramming results in a higher metastatic potential (44). Deregulation of fatty acid metabolism is crucial for malignant transformation in breast cancer. Proteins involved in the synthesis and oxidation of fatty acids play a pivotal role in the proliferation, migration, and invasion of breast cancer cells. Additionally, it has been shown that molecular subtypes display specific fatty acid metabolism (45). Deregulation of fatty acid metabolism has been associated with non-luminal tumors. Luminal subtypes rely on a balance between de novo fatty acid synthesis and oxidation as sources for biomass and energy. On the other hand, triple-negative basal breast cancer often uses exogenous fatty acids. In terms of targeted,

personalized therapy, it is desirable to take such differences into account. In the case of the pentose phosphate pathway (PPP), it has been shown that PPP-associated proteins, such as 6PGL, 6PGDH, or NRF2, are not differentially expressed among breast cancer subtypes but are overexpressed relative to control samples (46). Glucose 6-phosphate dehydrogenase G6PD has been closely associated with prognosis in Basal tumors (47). It has been demonstrated that G6PD silencing increases the glycolytic flux, reduces lipid synthesis, and increases glutamine uptake in breast cancer cells. This effect has also been strongly related to poor prognosis (48). Her2-positive Luminal B tumors present overexpression of G6PDH (49). However, even if the presence of PPP-related proteins in Luminal B breast cancer has been established, a global analysis of this pathway is still lacking.

As we have said, the Luminal B subtype is the one with the highest metabolic deregulation. It is known that, in the long-term, the Luminal B subtype presents higher drug resistance, metastasis, and relapses (50, 51). This could be, in part, due to the individual heterogeneity at the gene-expression level. The metabolic deregulation in this subtype could also underlie drug resistance.

**TABLE 1** | Overexpressed genes with FDA-approved inhibitors to regulate purine metabolism.

Search term	Drug	Interaction type
RRM2	FLUDARABINE PHOSPHATE	Inhibitor
RRM2	GALLIUM NITRATE	Inhibitor
RRM2	CLADRIBINE	Inhibitor
RRM2	CLOFARABINE	Inhibitor
RRM2	FLUDARABINE	Inhibitor
RRM2	GEMCITABINE	Inhibitor
RRM2	HYDROXYUREA	Inhibitor
RRM2	MOTEXAFIN GADOLINIUM	Inhibitor
RRM2	TEZACITABINE	Inhibitor
RRM2	GEMCITABINE HYDROCHLORIDE	Inhibitor
EZH2	CHEMBL3287735	Inhibitor

**TABLE 2** | Underexpressed genes with FDA-approved activators to regulate purine metabolism.

Search term	Drug	Interaction type
ACACB	METFORMIN	Activator
NPR1	ATACIGIJAT	Activator
PDE1C	BEPRIDIL	Activator
PDE2A	CHEMBL395336	Activator

To our knowledge, a profound study regarding metabolism in the Luminal B subtype is still necessary. However, we suggest that the long-term malignancy and poor prognosis of the Luminal B subtype are due, in part, to global metabolic deregulation more than to any single-molecule alteration. Further analyses in this regard are required to assess the metabolic deregulation patterns observed here with higher accuracy.

### 3.4. Purine Metabolism as a Potential Target in All Breast Cancer Subtypes

For the more than 2,611 drugs annotated in the DrugBank database and the 660 drugs annotated in PharmGKB, all subtype-specific differentially expressed genes were matched. The top 20 identified potential pharmacological targets obtained by the pipeline performed in Mejia-Pedroza et al. (31) are reported in **Table 1**. It contains those drugs that inhibit overexpressed genes. **Table 2** lists those drugs that activate underexpressed ones.

As can be observed in **Table 1**, RRM2, which participates in purine, pyrimidine, and glutathione metabolism, is the most targeted gene. EZH2, involved in lysine degradation, is another target that may be inhibited.

It is worth noticing that this computational tool provides all FDA-approved drugs that target a list of molecules, together with the effect that is produced in the target. **Supplementary Tables 1, 2** contain comprehensive lists of drugs and their targets for commonly overexpressed and underexpressed breast cancer genes.

In the case of underexpressed genes, three of the four targets of activator drugs participate in purine metabolism:

NPR1, PDE1C, and PDE2A. This result appears to be relevant in terms of the potential therapeutic options that breast cancer patients may have. There is a common deregulated metabolic pathway (purine metabolism) that can be targeted by specific drugs that have activator and inhibitory actions over underexpressed/overexpressed genes, respectively.

### 3.5. Deregulation of Metabolism Is Consistent in a Different Cohort

We performed a comparison with data from METABRIC (20), another large breast cancer cohort study. Our validation analysis shows a separation between groups as in the discovery group. A heatmap of the validation cohort is presented in **Supplementary Figure 1**, and the distribution of PDS in the METABRIC dataset is presented in **Supplementary Figure 2**. Some of our findings replicate those of METABRIC, although there were also differences, some of which may be attributable to METABRIC being a microarray-based experimental approach, whereas TCGA included data from RNA-sequencing experiments.

## 4. CONCLUSIONS

Heterogeneity is a crucial factor that impedes the understanding, diagnosis, and treatment of breast cancer tumors. Manifestations of this heterogeneity can be observed at the genomic, histological, or clinical level. In this work, we have provided another instance of this heterogeneity: metabolic deregulation.

Each breast cancer subtype has its own pattern of deregulation in metabolism, with Luminal B having the highest deregulation scores. This subtype presents alterations to metabolic processes such as pyruvate metabolism, tyrosine metabolism, fatty acid degradation, and the pentose phosphate pathway.

To our knowledge, this is the first time that a single-sample-based pathway analysis in breast cancer subtypes has been performed to identify differences in metabolic regulation. At the same time, this work has allowed us to design a common therapeutic FDA-approved scheme to regulate purine metabolism, independently of the subtype. With this kind of approach, it is possible to determine global deregulation patterns while, at the same time, finding individual signatures that may represent a further step toward personalized medicine.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the (Genomic Data Commons Data Portal).

## AUTHOR CONTRIBUTIONS

EH-L and JE-E contributed to the conception and design of the study. ES-C collected, organized the database, preprocessed the data, performed the computational analysis, and performed results visualization. ES-C, JE-E, and EH-L discussed and contextualized the results. JE-E and EH-L wrote the first draft of the manuscript. All authors

contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

This work was supported by the Consejo Nacional de Ciencia y Tecnología (SEP-CONACYT-2016-285544 and FRONTERAS-2017-2115) and the National Institute of Genomic Medicine, México. Additional support has been granted by the Laboratorio Nacional de Ciencias de la Complejidad from the Universidad Nacional Autónoma de México.

## ACKNOWLEDGMENTS

The authors thank Diana García-Cortés for her invaluable help with carrying out the validation analysis and improving the design of the computational strategy. We also thank Raúl A. Mejía-Pedroza for his help in the elaboration of Tables 1, 2. ES-C is a graduate student in the Biochemical Sciences program at the Universidad Nacional Autónoma de México and received a scholarship from the Consejo Nacional de Ciencia y Tecnología.

## REFERENCES

- Melchor L, Honrado E, García M, Alvarez S, Palacios J, Osorio A, et al. Distinct genomic aberration patterns are found in familial breast cancer associated with different immunohistochemical subtypes. *Oncogene*. (2008) 27:3165. doi: 10.1038/sj.onc.1210975
- Alcalá-Corona SA, de Anda-Jáuregui G, Espinal-Enríquez J, Hernández-Lemus E. Network modularity in breast cancer molecular subtypes. *Front Physiol*. (2017) 8:915. doi: 10.3389/fphys.2017.00915
- de Anda-Jáuregui G, Velázquez-Caldelas TE, Espinal-Enríquez J, Hernández-Lemus E. Transcriptional network architecture of breast cancer molecular subtypes. *Front Physiol*. (2016) 7:568. doi: 10.3389/fphys.2016.00568
- Habermann JK, Doering J, Hautaniemi S, Roblick UJ, Bündgen NK, Nicorici D, et al. The gene expression signature of genomic instability in breast cancer is an independent predictor of clinical outcome. *Int J Cancer*. (2009) 124:1552–64. doi: 10.1002/ijc.24017
- Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leib E, et al. Novel patterns of genome rearrangement and their association with survival in breast cancer. *Genome Res*. (2006) 16:1465–79. doi: 10.1101/gr.5460106
- Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell*. (2006) 10:529–41. doi: 10.1016/j.ccr.2006.10.009
- Sims AH, Howell A, Howell SJ, Clarke RB. Origins of breast cancer subtypes and therapeutic implications. *Nat Rev Clin Oncol*. (2007) 4:516. doi: 10.1038/nrnc0908
- Yensal O, Barutca S. Biological subtypes of breast cancer: prognostic and therapeutic implications. *World J Clin Oncol*. (2014) 5:412. doi: 10.5306/wjco.v5.i3.412
- Telang N. Stem cell targeted therapeutic approaches for molecular subtypes of clinical breast cancer. *World Acad Sci J*. (2018) 1:20–4. doi: 10.3892/wasj.2018.3
- Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA*. (2001) 98:10869–74. doi: 10.1073/pnas.191367098
- Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol*. (2011) 5:5–23. doi: 10.1016/j.molonc.2010.11.003
- Fresno C, González GA, Merino GA, Flesia AG, Podhajcer OL, Llera AS, et al. A novel non-parametric method for uncertainty evaluation of correlation-based molecular signatures: its application on PAM50 algorithm. *Bioinformatics*. (2017) 33:693–700. doi: 10.1093/bioinformatics/btw704
- Olsen SN, Wrónski A, Castaño Z, Dake B, Malone C, De Raedt T, et al. Loss of RasGAP tumor suppressors underlies the aggressive nature of luminal B breast cancers. *Cancer Discovery*. (2017) 7:202–17. doi: 10.1158/2159-8290.CD-16-0520
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. (2011) 144:646–74. doi: 10.1016/j.cell.2011.02.013
- Hu J, Locasale JW, Bielas JH, O'Sullivan J, Sheahan K, Cantley LC, et al. Heterogeneity of tumor-induced gene expression changes in the human metabolic network. *Nat Biotechnol*. (2013) 31:522. doi: 10.1038/nbt.2530
- Rosario SR, Long MD, Affronti HC, Rowsam AM, Eng KH, Smiraglia DJ. Pan-cancer analysis of transcriptional metabolic dysregulation using The Cancer Genome Atlas. *Nat Commun*. (2018) 9:5330. doi: 10.1038/s41467-018-07232-8
- Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. (2011) 561:70. doi: 10.1038/nature11412
- Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contempor Oncol*. (2015) 19:A68–77. doi: 10.5114/wo.2014.47136
- Drier Y, Sheffer M, Domany E. Pathway-based personalized analysis of cancer. *Proc Natl Acad Sci USA*. (2013) 110:6388–93. doi: 10.1073/pnas.1219651110
- Curtis C, Shah SP, Chin SE, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. (2012) 486:346–52. doi: 10.1038/nature10983
- Espinal-Enríquez J, Fresno C, Anda-Jáuregui G, Hernández-Lemus E. RNA-Seq based genome-wide analysis reveals loss of inter-chromosomal regulation in breast cancer. *Sci Rep*. (2017) 7:1760. doi: 10.1038/s41598-017-01314-1
- Nueda MJ, Ferrer A, Conesa A. ARSYN: a method for the identification and removal of systematic noise in multifactorial time course microarray experiments. *Biostatistics*. (2012) 13:553–66. doi: 10.1093/biostatistics/kxr042
- Tarazona S, García-Alcalde F, Dopazo J, Ferrer A, Conesa A. Differential expression in RNA-seq: a matter of depth. *Genome Res*. (2011) 21:2213–23. doi: 10.1101/gr.124321.111
- Tarazona S, García F, Ferrer A, Dopazo J, Conesa A. NOIseq: a RNA-seq differential expression method robust for sequencing depth biases. *EMBnet J*. (2012) 17:18–9. doi: 10.14806/ej.17.B.265
- Risso D, Schwartz K, Sherlock G, Dudoit S. GC-content normalization for RNA-Seq data. *BMC Bioinform*. (2011) 12:480. doi: 10.1186/1471-2105-12-480

26. Fresno C, González GA, Llera AS, Fernández EA. *pbcmc: Permutation-Based Confidence for Molecular Classification*. Saitama: R package version 1. (2016).
27. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007
28. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* (2000) 28:27–30. doi: 10.1093/nar/28.1.27
29. Drier Y. Quantify deregulation of pathways in cancer. *Bioconductor.* (2013). doi: 10.18129/B9.bioc.pathifier
30. García-Campos MA, Espinal-Enríquez J, Hernández-Lemus E. Pathway analysis: state of the art. *Front Physiol.* (2015) 6:383. doi: 10.3389/fphys.2015.00383
31. Mejía-Pedroza RA, Espinal-Enríquez J, Hernández-Lemus E. Pathway-based drug repositioning for breast cancer molecular subtypes. *Front Pharmacol.* (2018) 9:905. doi: 10.3389/fphar.2018.00905
32. de Anda-Jáuregui G, Mejía-Pedroza RA, Espinal-Enríquez J, Hernández-Lemus E. Crosstalk events in the estrogen signaling pathway may affect tamoxifen efficacy in breast cancer molecular subtypes. *Comput Biol Chem.* (2015) 59:42–54. doi: 10.1016/j.compbiolchem.2015.07.004
33. Carbonnelle-Puscian A, Copie-Bergman C, Baia M, Martín-García N, Allory Y, Haioun C, et al. The novel immunosuppressive enzyme IL4I1 is expressed by neoplastic cells of several B-cell lymphomas and by tumor-associated macrophages. *Leukemia.* (2009) 23:952. doi: 10.1038/leu.2008.380
34. Bod L, Lengagne R, Wrobel L, Ramspott JP, Kato M, Avril MF, et al. IL4-induced gene 1 promotes tumor growth by shaping the immune microenvironment in melanoma. *Oncimmunology.* (2017) 6:e1278331. doi: 10.1080/2162402X.2016.1278331
35. Lasoudris F, Cousin C, Prevost-Blondel A, Martín-García N, Abd-Allah I, Ortonne N, et al. IL4I1: an inhibitor of the CD8+ antitumor T-cell response *in vivo*. *Eur J Immunol.* (2011) 41:1629–38. doi: 10.1002/eji.201041119
36. Castellano E, Molinier-Frenkel V. IL4I1: an emerging target to reinvigorate antitumor immune responses. *Immunother Open Access.* (2017) 3:132–6. doi: 10.4172/2471-9552.1000132
37. Yuan N, Zhang G, Bie F, Ma M, Ma Y, Jiang X, et al. Integrative analysis of lncRNAs and miRNAs with coding RNAs associated with ceRNA crosstalk network in triple negative breast cancer. *Oncotargets Therapy.* (2017) 10:5883. doi: 10.2147/OTT.S149308
38. Pedley AM, Benkovic SJ. A new view into the regulation of purine metabolism: the purinosome. *Trends Biochem Sci.* (2017) 42:141–54. doi: 10.1016/j.tbs.2016.09.009
39. Klepinin A, Ounpuu L, Guzun R, Chekulayev V, Timohina N, Tepp K, et al. Simple oxygraphic analysis for the presence of adenylate kinase 1 and 2 in normal and tumor cells. *J Bioenerget Biomembran.* (2016) 48:531–48. doi: 10.1007/s10863-016-9687-3
40. Sadana R, Dessauer CW. Physiological roles for G protein-regulated adenylyl cyclase isoforms: insights from knockout and overexpression studies. *Neurosignals.* (2009) 17:5–22. doi: 10.1159/000166277
41. Jeon Y, Heo YS, Kim C, Hyun YL, Lee T, Ro S, et al. Phosphodiesterase: overview of protein structures, potential therapeutic applications and recent progress in drug development. *Cell Mol Life Sci CMLS.* (2005) 62:1198–220. doi: 10.1007/s00018-005-4533-5
42. Sasmita A, Wong Y. Organoids as reliable breast cancer study models: an update. *Int J Oncol Res.* (2018) 1:008. doi: 10.23937/ijor-2017/1710008
43. Corbet C. Stem cell metabolism in cancer and healthy tissues: pyruvate in the limelight. *Front Pharmacol.* (2018) 8:958. doi: 10.3389/fphar.2017.00958
44. Dupuy E, Tabariès S, Andrzejewski S, Dong Z, Blagih J, Annis MG, et al. PDK1-dependent metabolic reprogramming dictates metastatic potential in breast cancer. *Cell Metabol.* (2015) 22:577–89. doi: 10.1016/j.cmet.2015.08.007
45. Monaco ME. Fatty acid metabolism in breast cancer subtypes. *Oncotarget.* (2017) 8:29487. doi: 10.18632/oncotarget.15494
46. Choi J, Kim ES, Koo JS. Expression of pentose phosphate pathway-related proteins in breast cancer. *Disease Mark.* (2018) 2018:9369358. doi: 10.1155/2018/9369358
47. Dong T, Kang X, Liu Z, Zhao S, Ma W, Xuan Q, et al. Altered glycometabolism affects both clinical features and prognosis of triple-negative and neoadjuvant chemotherapy-treated breast cancer. *Tumor Biol.* (2016) 37:8159–68. doi: 10.1007/s13277-015-4729-8
48. Benito A, Polat IH, Noé V, Ciudad CJ, Marin S, Cascante M. Glucose-6-phosphate dehydrogenase and transketolase modulate breast cancer cell metabolic reprogramming and correlate with poor patient outcome. *Oncotarget.* (2017) 8:106693. doi: 10.18632/oncotarget.21601
49. Cha YJ, Jung WH, Koo JS. Differential site-based expression of pentose phosphate pathway-related proteins among breast cancer metastases. *Disease Mark.* (2017) 2017:7062517. doi: 10.1155/2017/7062517
50. Tobin N, Harrell J, Lövtrot J, Egyhazi Brage S, Frostvik Stolt M, Carlsson L, et al. Molecular subtype and tumor characteristics of breast cancer metastases as assessed by gene expression significantly influence patient post-relapse survival. *Ann Oncol.* (2014) 26:81–8. doi: 10.1093/annonc/mdl065.1
51. Prat A, Pineda E, Adamo B, Galván P, Fernández A, Gaba L, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast.* (2015) 24:S26–35. doi: 10.1016/j.breast.2015.07.008

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Serrano-Carbajal, Espinal-Enríquez and Hernández-Lemus. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.