Targeting Metabolic Deregulation Landscapes in Breast Cancer Subtypes

Erandi A. Serrano-Carbajal [1], Jesús Espinal-Enríquez [1,2]* and Enrique Hernández-Lemus [1,2]*

[1] Computational Genomics Division, National Institute of Genomic Medicine, Mexico City, Mexico, [2] Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, Mexico City, Mexico

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/2lJJrgi).

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 $Log_2 FoldChange > 1$ and a B-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 $|\log_2 FoldChange| > 1$ and $B - statistic > 5$ 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
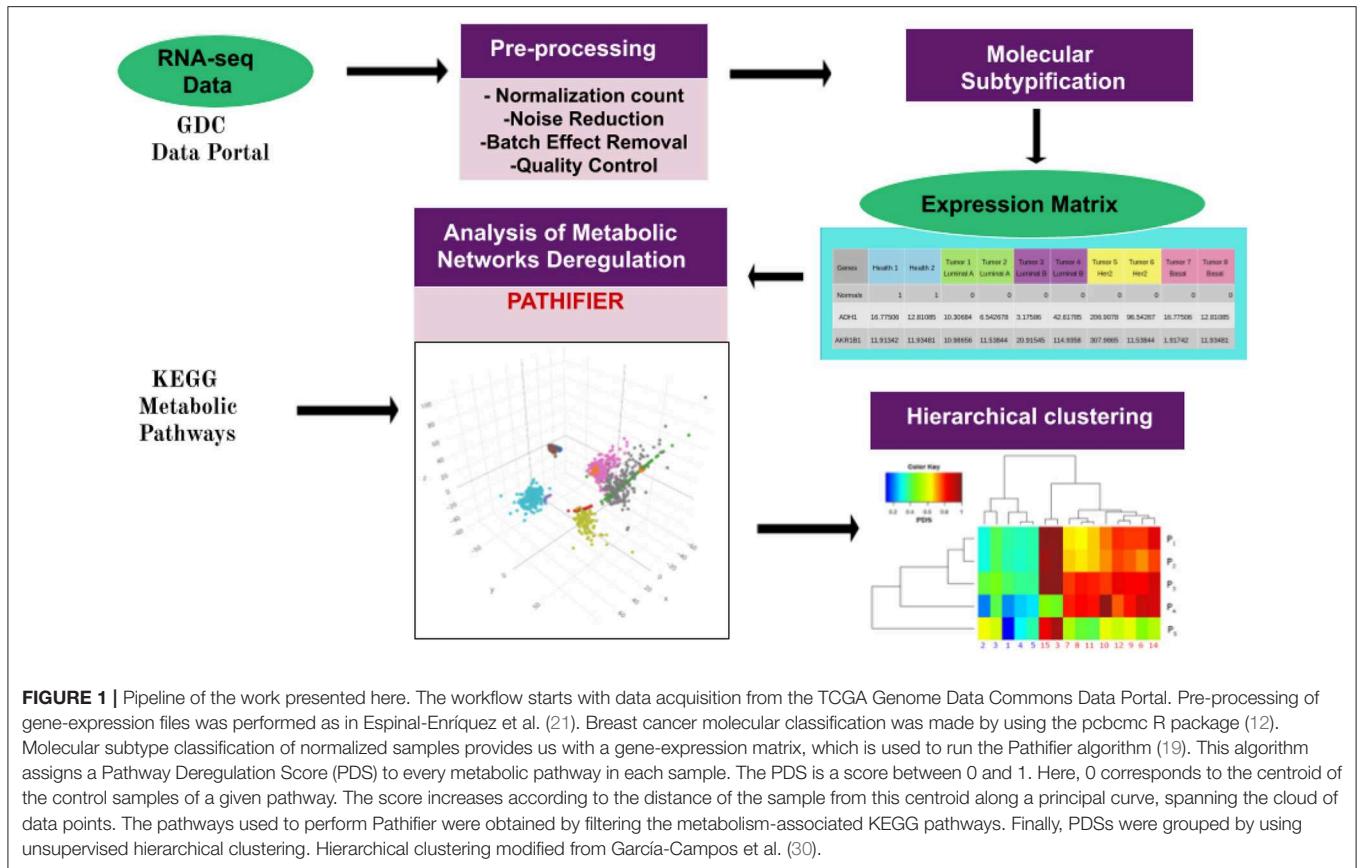
FIGURE 1 | Pipeline of the work presented here. The workflow starts with data acquisition from the TCGA Genome Data Commons Data Portal. Pre-processing of gene-expression files was performed as in Espinal-Enríquez et al. (21). Breast cancer molecular classification was made by using the pcbcmc R package (12). Molecular subtype classification of normalized samples provides us with a gene-expression matrix, which is used to run the Pathifier algorithm (19). This algorithm assigns a Pathway Deregulation Score (PDS) to every metabolic pathway in each sample. The PDS is a score between 0 and 1. Here, 0 corresponds to the centroid of the control samples of a given pathway. The score increases according to the distance of the sample from this centroid along a principal curve, spanning the cloud of data points. The pathways used to perform Pathifier were obtained by filtering the metabolism-associated KEGG pathways. Finally, PDSs were grouped by using unsupervised hierarchical clustering. Hierarchical clustering modified from García-Campos et al. (30).
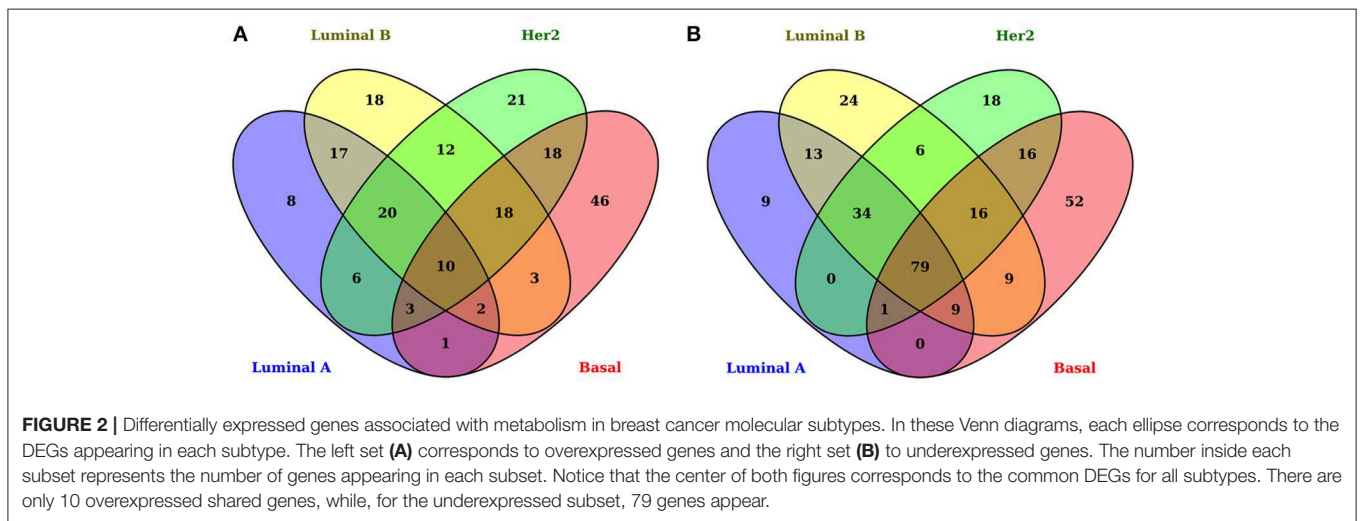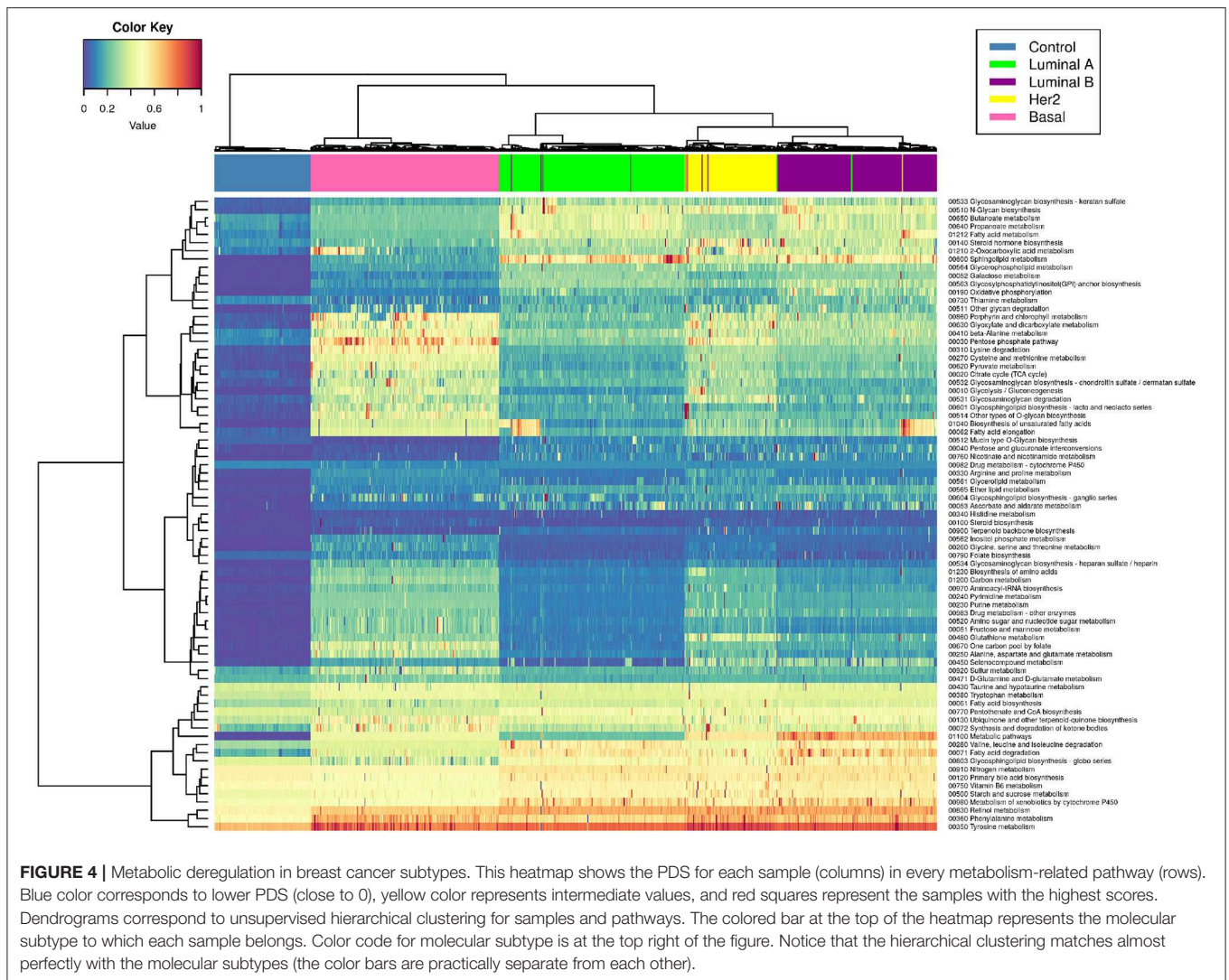

FIGURE 2 | Differentially expressed genes associated with metabolism in breast cancer molecular subtypes. In these Venn diagrams, each ellipse corresponds to the DEGs appearing in each subtype. The left set **(A)** corresponds to overexpressed genes and the right set **(B)** to underexpressed genes. The number inside each subset represents the number of genes appearing in each subset. Notice that the center of both figures corresponds to the common DEGs for all subtypes. There are only 10 overexpressed shared genes, while, for the underexpressed subset, 79 genes appear.

and shaping the immune microenvironment in melanoma (34). Autoimmune suppression and the inhibition of *CD*8+ 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 3 | Metabolic processes associated with DEGs. In these bipartite network representations, overexpressed **(A)** and underexpressed **(B)** genes are depicted in red and blue, respectively. Metabolic processes are shown in pink squares. Links between both types of nodes appear if the gene is present in the corresponding pathway. For this network, we kept only those DEGs that appear in the four subtypes.

(39), AK5, which catalyzes degradation reactions of ATP (40), or PDE and NPR, which control the proportion of second messengers, strongly implicated in signal transduction (41).

The majority of these genes are involved in the formation/degradation of ATP. Since cell proliferation is a hallmark of cancer, we argue that underexpression of these genes may enable the tumors to avoid ATP/GTP degradation, thus providing energetic fuel to cell proliferation.

3.2. Metabolic Deregulation Patterns Are Characteristic of Each Breast Cancer Subtype

Once it has been shown that common deregulated genes induce regulation patterns in some metabolic processes, the remaining question is whether variations in the whole gene-expression signature correspond to changes in specific metabolic deregulation.

FIGURE 4 | Metabolic deregulation in breast cancer subtypes. This heatmap shows the PDS for each sample (columns) in every metabolism-related pathway (rows). Blue color corresponds to lower PDS (close to 0), yellow color represents intermediate values, and red squares represent the samples with the highest scores. Dendrograms correspond to unsupervised hierarchical clustering for samples and pathways. The colored bar at the top of the heatmap represents the molecular subtype to which each sample belongs. Color code for molecular subtype is at the top right of the figure. Notice that the hierarchical clustering matches almost perfectly with the molecular subtypes (the color bars are practically separate from each other).
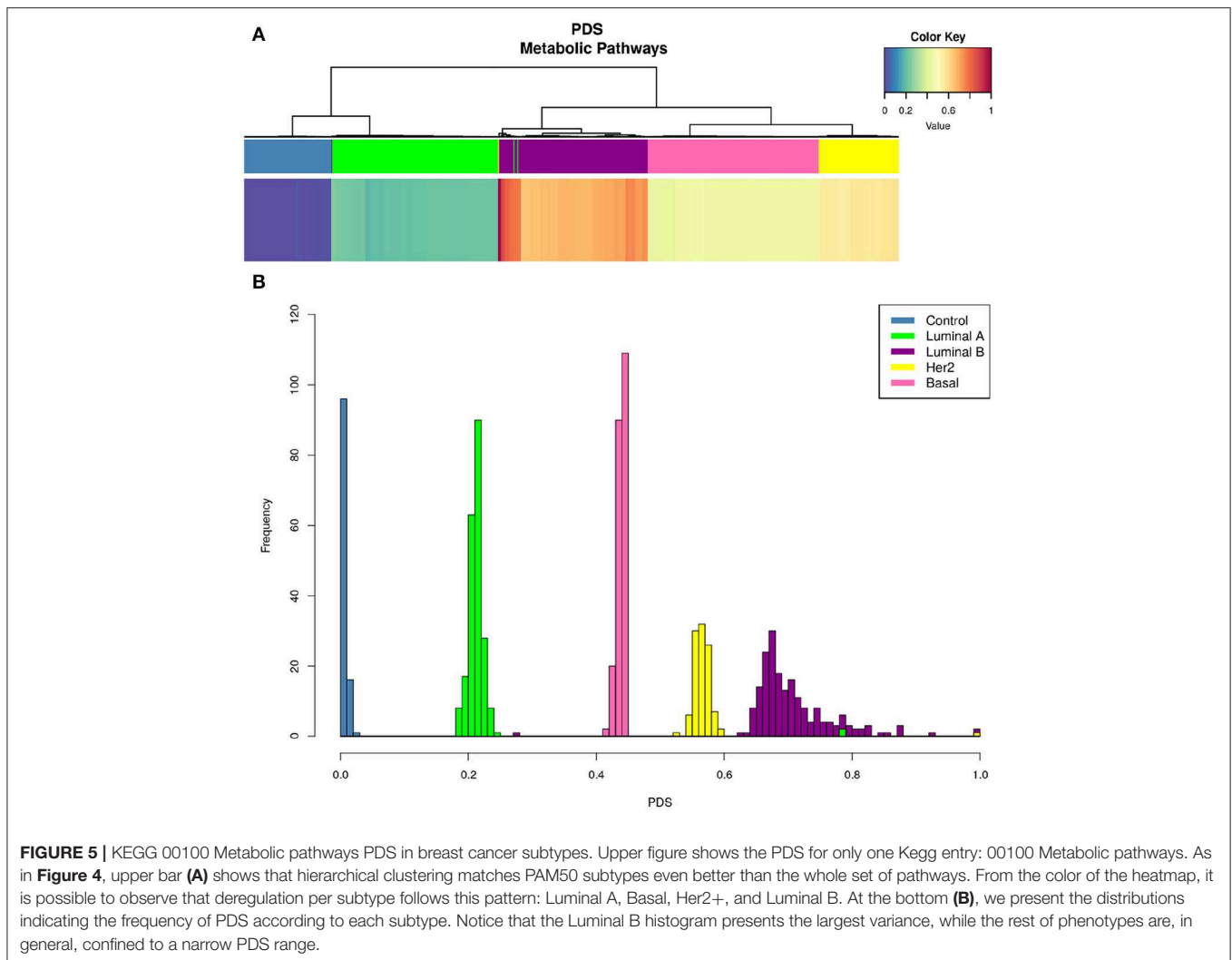
Figure 4 shows a heatmap of the PDS values (see methods) grouped by PDS similarity. Rows correspond to all pathways associated with metabolism, while columns correspond to samples. There are subsets of samples that present a similar metabolic deregulation among subgroups and differ from the other samples.

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 metabolic-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.

FIGURE 5 | KEGG 00100 Metabolic pathways PDS in breast cancer subtypes. Upper figure shows the PDS for only one Kegg entry: 00100 Metabolic pathways. As in **Figure 4**, upper bar **(A)** shows that hierarchical clustering matches PAM50 subtypes even better than the whole set of pathways. From the color of the heatmap, it is possible to observe that deregulation per subtype follows this pattern: Luminal A, Basal, Her2+, and Luminal B. At the bottom **(B)**, we present the distributions indicating the frequency of PDS according to each subtype. Notice that the Luminal B histogram presents the largest variance, while the rest of phenotypes are, in general, confined to a narrow PDS range.

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.
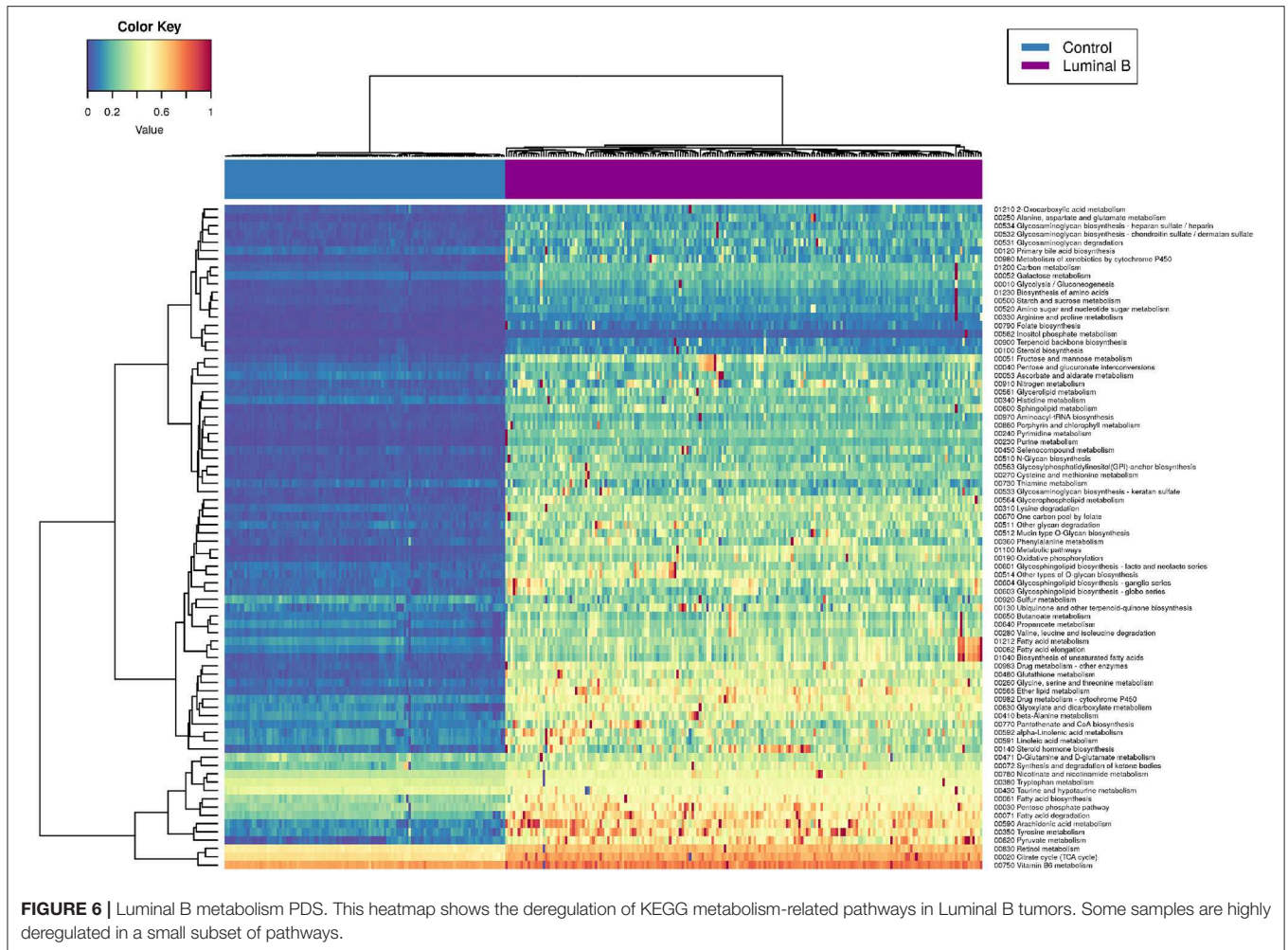
Regarding differences, Rosario et al. found different pathway scores for the Basal and Luminal A subtypes. However, as can be seen from Figure 6C of Rosario's paper, the low specificity of the average gene-expression Z-scores results in a non-conclusive depiction, as it is hard to distinguish signal 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

FIGURE 6 | Luminal B metabolism PDS. This heatmap shows the deregulation of KEGG metabolism-related pathways in Luminal B tumors. Some samples are highly deregulated in a small subset of pathways.

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	ATACIGUAT	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 Mejía-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.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc. 2020.00097/full#supplementary-material

Supplementary Figure 1 | PDS heatmap of validation cohort. This heatmap showing the deregulation of KEGG metabolism-related pathways in the METABRIC cohort. Color code is the same as in the figures in the main manuscript.

Supplementary Figure 2 | Distribution of PDS in the validation cohort. Distributions indicating the frequency of PDS according to each subtype in the METABRIC cohort.

Supplementary Table 1 | Comprehensive lists of drugs and their targets for commonly overexpressed breast cancer genes.

Supplementary Table 2 | Comprehensive lists of drugs and their targets for commonly underexpressed breast cancer genes.

REFERENCES

1. Melchor L, Honrado E, Garcia 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

2. 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

3. 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

4. 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

5. Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leibu 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

6. 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

7. 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/ncponc0908

8. Yersal 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

9. 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

10. Sørlie 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

11. 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

12. 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

13. Olsen SN, Wronski 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

14. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* (2011) 144:646–74. doi: 10.1016/j.cell.2011.02.013

15. 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

16. 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

17. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature.* (2011) 5:61–70. doi: 10.1038/nature11412

18. 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

19. 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

20. Curtis C, Shah SP, Chin SF, 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

21. 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

22. 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

23. 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

24. 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

25. 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, Martin-Garcia 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. *Oncoimmunology*. (2017) 6:e1278331. doi: 10.1080/2162402X.2016.1278331

35. Lasoudris F, Cousin C, Prevost-Blondel A, Martin-Garcia N, Abd-Alsamad 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 F, 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.tibs.2016.09.009

39. Klepinin A, Ounpuu L, Guzun R, Chekulayev V, Timohhina 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 F, 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, Lövrot 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/mdu065.1

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