METABOLOMIC PROFILING OF PLASMA AND TISSUE IN LUMINAL A AND B BREAST CANCER: BIOMARKER DISCOVERY AND TISSUE ASSOCIATIONS

Nerea Agud Lombarte

TRABAJO FINAL DE GRADO BIOTECNOLOGÍA

Tutor académico: Javier Capilla Luque, Departamento de Ciencias Médicas Básicas, (javier.capilla@urv.cat).

En cooperación con: Unidad de Investigación Biomédica (URB) del Instituto de Investigación Sanitaria Pere Virgili (IISPV).

Supervisor/es: Jorge Joven Maried, Doctor Responsable de la Unidad de Investigación Biomédica (URB), (jorge.joven@salutsantjoan.cat); Andrea Jiménez Franco, (andrea.jimenez@urv.cat), Técnica de laboratorio Predoctoral en la Unidad de Investigación Biomédica (URB).

Fecha de convocatoria; Junio 2023

Yo, Nerea Agud Lombarte, con DNI 73107334V, soy conocedora de la guía de prevención del plagio en la URV *Prevención, detección y tratamiento del plagio en la docencia: guía para estudiantes* (aprobada en julio 2017) (http://www.urv.cat/ca/vidacampus/serveis/crai/que-us-oferim/formacio-competencies-nuclears/plagi/) y afirmo que este TFG no constituye ninguna de las conductas consideradas como plagio por la URV.

Tarragona, 7 de junio de 2023

TABLE OF CONTENTS

CENTER DETAILS

This project is linked to the external curricular internships carried out at the Biomedical Research Unit (URB) of the Pere Virgili Health Research Institute (IISPV). This research group was created in 1990 and is located at the Sant Joan de Reus University Hospital. The unit is directed by Prof. Jorge Joven and Dr. Jordi Camps.

The URB is dedicated to basic and clinical research related to non-communicable and metabolic diseases such as morbid obesity, diabetes, non-alcoholic fatty liver disease, cardiovascular diseases and cancer. The lines of research are focused on the study of metabolic alterations through the multi-omic analysis of metabolic disorders.

ABSTRACT

Background and aims: Breast cancer (BC) is one the most diagnosed cancers and the main cause of cancer-related death among the women population. The metabolic reprogramming in BC increases proliferation and survival of cancer cells, conferring a unique metabolomic profile to different molecular subtypes of BC. In the present study, we aimed to investigate the tissue and systemic metabolomic alterations in luminal A (LA) and lumina B (LB) BC patients, discerning the distinctive metabolic signatures, and to identify potential biomarkers and tissue associations that may contribute to the development and progression of the disease.

Methods: 54 women with BC who had not yet received any oncological treatment were recruited from Hospital Universitari Sant Joan in Reus. Tissue samples were acquired from 13 of these patients who underwent surgery as their first cancer treatment. As a control group, we analyzed samples from 50 healthy women without any carcinogenic evidence. We performed targeted metabolomics in both plasma and tissue.

Results: Differential metabolites were identified between peritumoral and tumoral tissue, aligning with the metabolic disorder induced by tumor cells. Moreover, we identified biomarkers of luminal BC: xylonic acid, ethylmalonic acid, glycerol, 3-phosphoglyceric acid and maltose. Significant differences were found between LA BC and LB BC, illustrating the higher aggressiveness of LB BC and its metabolic functioning. The importance of clinical and pathophysiological covariates in the analyses was confirmed. Finally, correlations between plasma and tissue were found that may elucidate the metabolism of cancer.

Conclusion: Our findings revealed that patients with luminal BC exhibited extensive changes in metabolite levels, predominantly associated with carbohydrate pathways. The incorporation of clinical and histopathological variables in the analysis of luminal subtypes proved to be critical for appreciating their unique distinctions. The findings of this study may have significant implications for the diagnosis and treatment of LA and LB BC, ultimately improving the quality of life for women affected by this disease.

Keywords: tumor microenvironment; cell metabolism; molecular subtype; oncological tissue; targeted metabolomics.

ABBREVIATIONS

WHO: World Health Organization
ACS: American Cancer Society
DALYs: Disability-Adjusted Life Years
WCRF: World Cancer Research Fund International
BC: Breast cancer
DCIS: Ductal Carcinoma In Situ
LCIS: Lobular Carcinoma In situ
HRT: Hormone Replacement Therapy
LA: Luminal A
LB: Luminal B
HER2: Human Epidermal Growth Factor Receptor 2
TNBC: Triple-Negative Breast Cancer
ER: Estrogen Receptor
PR: Progesterone Receptor
SERMs: Selective Estrogen Receptor Modulators
CDK4/6: Cyclin-Dependent Kinase 4/6
TCA: Tricarboxylic Acid Cycle
TME: Tumor Microenvironment
NMR: Nuclear Magnetic Resonance
MS: Mass Spectrometry
CE: Capillary Electrophoresis
GC: Gas chromatography
EI: Electron impact ionization
MS: Mass spectrometry
QTOF: High-resolution time-of-flight
FC: Fold Change
FDR: False Discovery Rate
PCA: Principal Component Analysis
PLS-DA: Partial Least Squares Discriminant Analysis
VIP: Variable Importance in Projection
ROC: Receiver Operating Characteristic
AUC: Area Under the Curve
CV: Cross-Validation
ORs: Odds Ratios
CIs: Confidence Intervals
BMI: Body Mass Index
HDL: High-density lipoproteins
LDL: Low-density lipoproteins

VLDL: Very low-density lipoprotein
TG: Triglycerides
LDH: Lactate dehydrogenase

1. INTRODUCTION

1.1. Cancer Worldwide: Understanding the Disease and Its Impact on Global Health

Cancer is a complex disease that arises from genetic and epigenetic alterations that disrupt the standard control mechanisms governing cell proliferation, differentiation, and apoptosis. These alterations confer a growth advantage to a cell, resulting in uncontrolled proliferation, apoptosis evasion, and ultimately cancer development. The accumulation of additional genetic and epigenetic changes allows for the progression of the tumor to a more aggressive and metastatic phenotype [1].

As a leading cause of death worldwide, cancer significantly impacts global health. According to the official databases of *World Health Organization* (WHO) and *American Cancer Society* (ACS), cancer among all human diseases has the most significant impact on people's lives in terms of the number of years lost due to disability or early death, as measured by Disability-Adjusted Life Years (DALYs)[2]. Data extracted from *World Cancer Research Fund International* (WCRF) shows an estimated 18.1 million cancer cases around the world in 2020 (the last year available) being the breast cancer the most common, contributing 12.5% of the total number of new cases diagnosed in 2020.

The pervasiveness and severity of cancer underscore the importance of research efforts aimed at improving prevention, diagnosis, and treatment strategies. As our understanding of cancer biology and tumor progression deepens, novel therapeutic approaches and early detection methods can be developed, ultimately reducing the global burden of this devastating disease.

1.2. Breast cancer

Breast cancer (BC) is a type of cancer that arises in the cellular components of the breast tissue. The breast is comprised of glandular tissues that harbor the milk-producing lobules and ducts and the stromal tissues that consist of the adipose and fibrous connective tissues. Moreover, the breast contains lymphatic tissue that serves as part of the immune system and aids in removing cellular fluids and waste (**Figure 1**).

Most BC cases originate from the epithelial cells lining the ducts, called ductal carcinomas. A smaller subset arises from cells surrounding the lobules, known as lobular carcinomas, while a minority of BC cases develop from other tissues within the breast (**Figure 2**) [3]. The diverse

origins of BC contribute to its heterogeneity and complexity, which influence the diagnosis, prognosis, and treatment strategies for affected individuals.

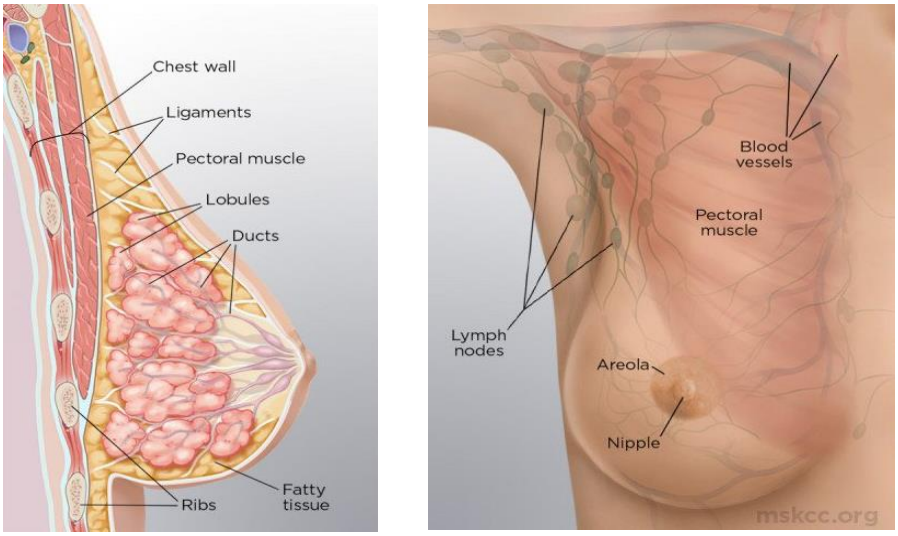


Figure 1. Anatomy of the female breast with a focus on structures relevant to breast cancer. Extracted from *Memorial Sloan Kettering Cancer Center.*

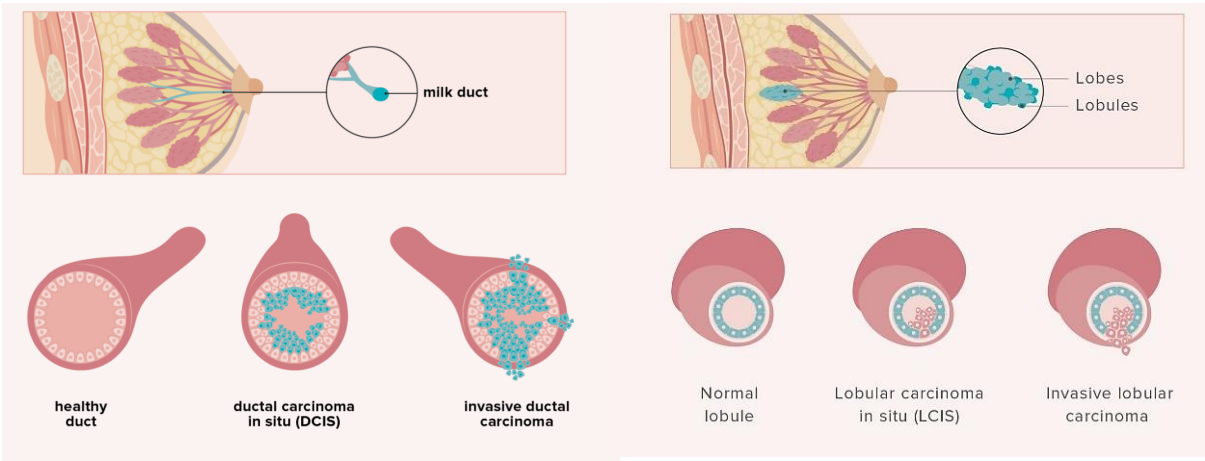


Figure 2. Schematic representation of the different types of breast cancer. Ductal Carcinoma In Situ (DCIS) and Lobular Carcinoma in situ (LCIS) are non-invasive forms of BC where abnormal cells are confined to the ducts or lobules, respectively. Invasive ductal carcinoma and invasive lobular carcinoma refer to cancer that has spread beyond the ducts or lobules into the surrounding breast tissue. Extracted from *Healthline.*

1.2.1. Incidence and impact on public health

BC is a major health concern worldwide, as it is the leading cause of morbidity and mortality among women. BC can affect both men and women; however, male BC cases are exceedingly rare. BC is the most frequently diagnosed malignancy in the female population, as evidenced by its prevalence in numerous countries, including Spain (**Figure 3**).

Furthermore, the economic impact of BC is also significant, placing a considerable financial burden on individuals, families, and society at large primarily due to direct medical costs, productivity losses, and informal care costs [4].

Therefore, it is crucial to identify cost-effective and efficient diagnostic and therapeutic methods to improve patient outcomes and alleviate the overall impact of BC. To achieve this goal, concerted research efforts are needed to understand the underlying pathophysiology of BC and the development and progression of the disease[5].

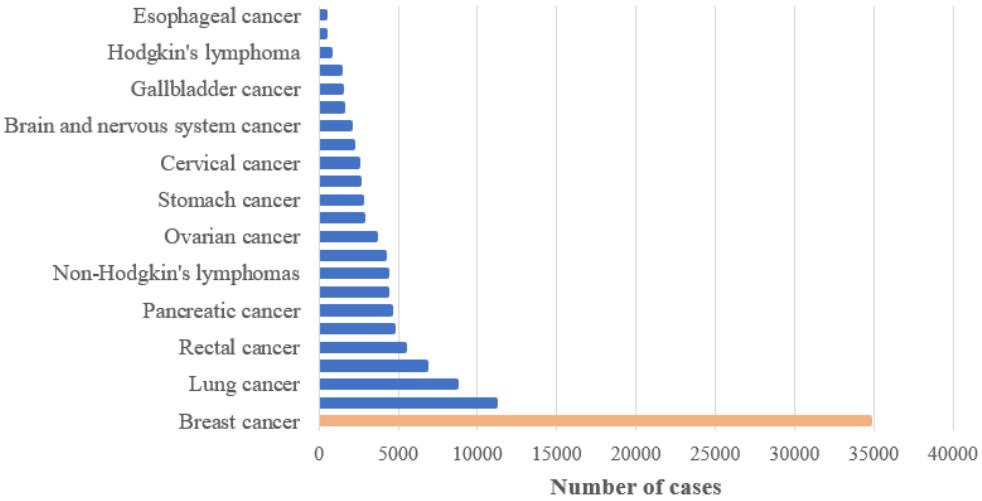


Figure 3. Estimated number of cancer incident cases in Spain by tumor type, 2022. Females*. Data extracted from the *Spanish Network of Cancer Registries* (REDECAN)

1.2.2. Risk factors

As BC development is influenced by genetic, environmental, and lifestyle factors, risk factors for BC can be classified into non-modifiable and modifiable factors[6].

Non-modifiable risk factors include age, with about 80% of BC cases occurring in women aged 50 years and older[7]. Gender is another non-modifiable factor, as women are at a much higher risk of developing this disease than men due to differences in breast tissue and hormonal exposure. A family history of BC, particularly in first-degree relatives (i.e., mother, sister, or daughter), increases the risk of developing the disease[8]. Approximately 5-10% of BC cases are due to inherited genetic mutations, such as BRCA1 and BRCA2[9]. Early menarche (before age 12) and late menopause (after age 55) are associated with an increased risk due to more prolonged exposure to estrogen[10]. Additionally, women who have never given birth or had their first child after age 30 have a higher risk of BC[11].

Modifiable risk factors also affect the likelihood of developing BC [12]. One is hormone exposure, as combined estrogen-progestin hormone replacement therapy (HRT) increases BC risk. However, the risk decreases after discontinuation of HRT[13]. Alcohol consumption has been positively associated with BC risk, with a 10% increase in risk for every 10 grams of alcohol consumed per day[14]. Multiple case–control studies observed an inverse association between leisure time activity and BC risk [15]. Additionally, consistent evidence suggests that regular physical activity can reduce the risk of disease by 10-25%[16]. Postmenopausal women who are obese have also an increase of risk due to higher circulating estrogen levels produced by adipose tissue[17]. A diet high in red and processed meat has been associated with a modest increase in BC risk, while a diet rich in fruits, vegetables, and whole grains may help reduce the risk[14].

1.3. Molecular subtypes and characteristics of BC

BC is a heterogeneous disease with numerous molecular subtypes, each with a unique gene expression profile, biological behavior, and therapeutic response. The classification of BC into molecular subtypes has enhanced the knowledge of the condition and prompted the development of more specialized strategies for treatment. The five main molecular subtypes of BC are luminal A (LA), luminal B-HER2-negative (LB-HER2-), luminal B-HER2-positive (LB-HER2+), HER2-positive and Triple-negative BC (TNBC)[18].

Classification of invasive BC into distinct molecular subtypes can be achieved through based on the cell proliferation index (Ki67%)[19] and immunohistochemical methods that evaluate the expression levels of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2)[20]. LA BC (ER+ and/or PR+, and HER2-) is the most common BC and typically carries a favorable prognosis. LB BC (ER+ and/or PR+, and HER2-/HER2+) is characterized by high levels of ki67 (>14%), a marker of cellular proliferation. LB-HER2+ is distinguished by a higher cell proliferation rate and is more aggressive than LB-HER2- cancer. HER2-positive BC (ER-, PR-, and HER2+) is associated with a poorer prognosis. TNBC (ER-, PR-, and HER2-) is often more aggressive, associated with a worse prognosis than other molecular subtypes and is commonly observed in younger women [21] (**Figure 4**).

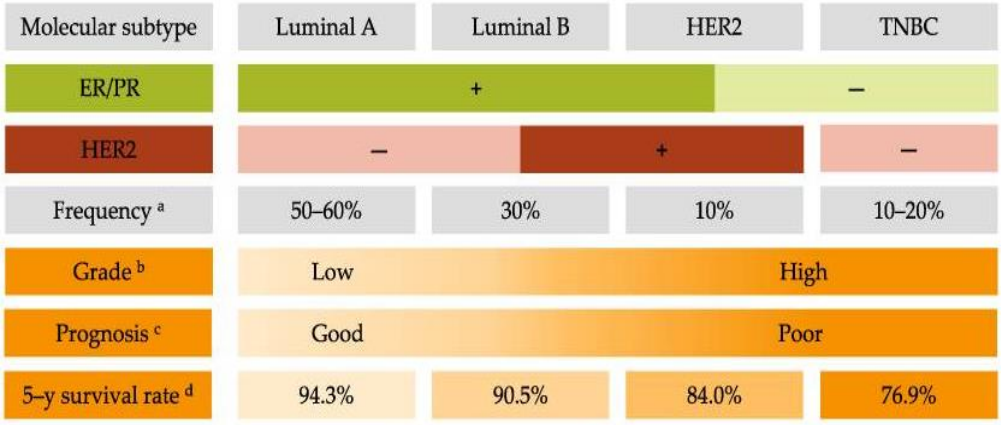


Figure 4. Molecular subtypes of BC. ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2; TNBC: triple-negative breast cancer. Figure extracted and adapted from *Burguin et al.* [21]

1.3.1. LA and LB BC

Despite their relatively less aggressive nature, the prevalence of LA and LB BC subtypes among women worldwide emphasizes the importance of comprehensive study and characterization. A thorough understanding of each subtype's shared and distinctive features holds significant value in selecting optimal and efficacious treatment strategies for patients. By unraveling the intricacies of these subtypes, healthcare professionals can enhance patient care and tailor interventions to achieve favorable outcomes.

LA and LB BC are characterized by the expression of hormone receptors, such as ER and/or PR. Although both categories are typically responsive to hormone therapies, they differ in their molecular characteristics, prognosis, and response to treatment. LA BC is the most prevalent subtype, accounting for approximately 40-50% of all breast cancers. It is typically associated with a lower histologic grade, slower growth rate, and better prognosis than other subtypes[22]. LA tumors are susceptible to hormonal therapies and have a lower risk of recurrence.

On the other hand, LB BC represents around 20-30% of all breast cancers. Compared to LA tumors, it has a higher histologic grade, a faster growth rate, and a worse prognosis[22]. LB cancers are also positive for hormone receptors but have a higher proliferation rate and may express higher levels of the HER2 protein. This subtype may be less sensitive to hormonal therapies and may benefit from additional treatments, such as chemotherapy and/or HER2-targeted therapies.

LA and LB BC treatment options primarily involve hormonal therapies, such as tamoxifen, aromatase inhibitors, or selective estrogen receptor modulators (SERMs). These treatments target the hormone receptors on the cancer cells, inhibiting their growth and proliferation. As

the LB tumors are more aggressive, there are many cases in which additional treatments like chemotherapy, HER2-targeted therapies (e.g., trastuzumab), or cyclin-dependent kinase 4/6 (CDK4/6) inhibitors may be recommended to improve outcomes[23].

1.3.2. Clinical and Histopathological Characteristics

In addition to molecular subtypes, several other classifications play a crucial role in characterizing luminal BC. These characteristics are tumor grade, tumor size, histological classification, lymph node involvement and the presence of hormone receptors and HER2 status. [24]

Tumor grade is determined according to the Nottingham grading system[25]. This system stratifies tumor grade into three levels: I, II, and III. The Nottingham grading system considers three main features of the tumor: tubule formation, nuclear pleomorphism, and the rate of mitosis [26]. These grades offer insight into the likely aggressiveness of the tumor and its potential response to treatment.

Tumor size was classified following the Tumor-Node-Metastasis (TNM)-based staging system, with classifications T1, T2, and T3 indicating increasing size or extent of the primary tumor [27]. The size of the tumor is a critical factor in staging and can influence the choice of treatment and prognosis.

Histological classification was identified and categorized into two major histological types, namely, ductal and lobular, referring to the origin of the cancer cells within the breast.

The presence and status of hormone receptors (ER and PR) and HER2 were also considered. These markers are vital in guiding treatment strategies, with different therapies targeted toward different receptor statuses.

1.4. Metabolic reprogramming

Numerous studies have shown that the metabolic profiles of BC and healthy breast tissues vary significantly [28] [29]. Changes in several metabolic processes, such as glycolysis, the TCA cycle, amino acid, nucleotide, and/or lipid metabolisms, may cause these variations. One of the main effects that change the metabolism of cancer cells is the Warburg effect[30]. The Warburg effect is a metabolic alteration in many cancer cells, including BC. In this effect, the primary source of energy shifts from the efficient process of cellular respiration to the less efficient method of aerobic glycolysis, which breaks down glucose to produce energy. In BC, this metabolic reprogramming is associated with increased expression of glucose transporters and glycolytic

enzymes and decreased activity of mitochondrial electron transport chain. These metabolic changes may provide a growth advantage to cancer cells and contribute to therapy resistance.

1.4.1. Role of the tumor microenvironment and peritumoral tissue

Not only metabolic reprogramming occurs in tumor cells but also in their surrounding tumor microenvironment (TME). The TME refers to the non-cancerous cells and extracellular matrix surrounding and supporting tumor growth and survival. This environment includes immune cells, stromal cells, blood vessels, nerve cells, and signaling molecules interacting with tumor cells. The TME is essential in tumor progression, therapy resistance, and metastasis.

The tissue directly surrounding a BC tumor, known as the peritumoral tissue, is crucial to the development and growth of the tumor. As mentioned in other studies, the body's immune response to the tumor may cause the tissue to become inflamed and supply nutrients and growth factors that support tumor development [31]. Additionally, peritumoral tissue can play a significant role in the process of angiogenesis –creating new blood vessels – which is essential for providing oxygen and nutrition to tumors. Angiogenesis can also provide a pathway for cancer cells to spread to other body parts[32] [33]. Due to the activity of stromal cells, the peritumoral tissue, in some cancer forms, may become fibrous and dense. This fibrosis may be a physical barrier that prevents immune cells and drugs from reaching the tumor[34]. Furthermore, the tumor can invade and replace the peritumoral tissue around it in later phases of BC. This invasion may aid the lymphatic or blood systems' ability to disseminate cancer cells.

Considering all this, peritumoral tissue can serve as a physical barrier, promote invasion, or provide nutrients, growth factors, and a vascular network that can significantly affect tumor development. This suggests that this tissue is also altered and not functioning properly. Therefore, knowing the intricate interaction between the tumor and its microenvironment is essential to develop efficient cancer treatments.

1.5. Metabolomics in cancer

Metabolomics, an omics science, is an emerging field in cancer research that has expanded rapidly over the last few years, focusing on the comprehensive, data-driven investigation of the low molecular weight compounds found in biological systems, known as the metabolome[35]. Specifically, the term metabolome refers to the total number of metabolites present within the cell, tissue, organ, or organism and have a wide range of functions. In the biological process, the information encoded in an organism's DNA (genome) is transcribed into RNA (transcriptome), which is further translated into proteins (proteome) and finally results in the formation of small molecules called metabolites (metabolome) (**Figure 5**).

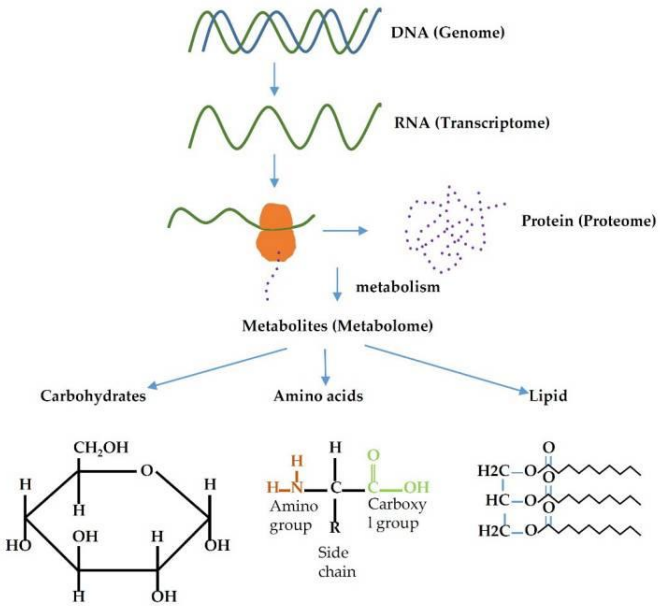


Figure 5. Relation flow of different omics family.
Figure extracted from *Subramani R et al.* [35]

Metabolomics or metabolic profiling measures the biological systems of low molecular weight metabolites and intermediates that reflect the dynamic response to genetic modification or metabolic changes. That is why, nowadays, metabolomics is considered a powerful and reliable tool that has high reproducibility and can have a significant impact in the field of cancer.

The most studied feature in this area includes central carbon metabolism and the relationship between glycolysis, the tricarboxylic acid (TCA) cycle, and oxidative phosphorylation[36]. By identifying altered metabolic pathways in cancer cells, metabolomics can reveal specific enzymes or metabolites crucial for tumor development and survival, leading to the discovery of potential biomarkers and therapeutic targets for BC[37].

The metabolome comprises both endogenous and external elements, such as extracellular or extra-organic compounds, which are catabolized or anabolized by the biological system. Due to its dynamic nature, constantly changing in response to various environmental and physiological stimuli, the metabolome can provide essential insights into the functional state of biological systems and the pathophysiology of diseases, including cancer, cardiovascular diseases, and neurological disorders. Several analytical techniques are employed in metabolomics studies, including Nuclear Magnetic Resonance (NMR) spectroscopy, Mass Spectrometry (MS), and Capillary Electrophoresis (CE), which can help identify multiple metabolites simultaneously, providing a comprehensive view of the metabolic profile[38].

Metabolomics has the potential to revolutionize personalized medicine by identifying unique metabolic signatures associated with individual patients, tumors, or treatment responses. This information can be used to tailor cancer treatments based on the specific metabolic characteristics of each patient, improving treatment outcomes and minimizing side effects. Biomarker discovery is a critical aspect of personalized medicine, as it enables the identification of molecules that can predict disease risk, monitor disease progression, or assess treatment response. By analyzing the metabolome through metabolomics approaches, novel biomarkers and therapeutic targets for various diseases, including breast cancer, can be revealed. These new biomarkers can aid in early detection and diagnosis, providing valuable information for developing targeted therapies and optimizing treatment strategies.

Nonetheless, the investigation of BC presents inherent complexities due to the concurrent presence of other influential characteristics that can skew comparison results. To overcome this challenge, rigorous inter-group studies or adjusted comparisons are indispensable to facilitate a comprehensive and accurate metabolomic analysis.

2. HYPOTHESIS AND OBJECTIVES

Currently, the classification of BC into LA or LB subtypes is determined through a biopsy and subsequent tissue analysis. As biopsies are invasive procedures that can cause discomfort and carry some risk of complications, the development of less invasive methods for subtype determination, such as the identification of plasma biomarkers, would be of great clinical significance.

Several studies have explored the potential of utilizing plasma biomarkers for differentiating between tumor subtypes[39], which can provide valuable information for cancer diagnosis and treatment planning. To fully understand the plasma markers' significance, it is crucial to investigate the tumor tissue characteristics of both LA and LB subtypes and the surrounding tumor microenvironment. This approach can shed light on the differences between these subtypes and provide insight into their unique metabolic and molecular profiles.

Furthermore, tumor less invasive diagnostic methods for BC subtyping can be developed by deepening the understanding of tumor tissue and microenvironment characteristics. This advancement could improve patient care and outcomes.

The main objective of the present study is to delineate differential metabolomic parameters for the precise diagnosis and distinction between LA and LB subtypes of BC. In pursuit of this primary objective, a series of investigations will be implemented, each with specific and targeted objectives that can be summarized as follows:

1. Characterize the metabolomic profiles of tumoral tissue and peritumoral regions in luminal BC and identify significant differences between LA and LB BC.
2. Discover metabolomic biomarkers for the detection of luminal BC.
3. Identify differences in LA and LB BC metabolomic profiles of plasma that could serve as diagnostic biomarkers for BC subtyping.
4. Establish possible associations between tissue-level metabolomics and plasma metabolomics in LA and LB BC subtypes.

3. MATERIALS AND METHODS

3.1. Study design and participants

The study enrolled 54 women with invasive BC at the Hospital Universitari Sant Joan in Reus between September 2020 and October 2021. Samples from 50 healthy women without any carcinogenic evidence were used as the control group.

The inclusion and exclusion criteria were carefully considered, and only patients with invasive BC LA or LB (HER2-negative) aged 18 years and above were included. Patients with a history of oncological disease, confirmed BC metastatic stage, Paget's nipple disease, vascular collagen disease, systemic lupus erythematosus and/or scleroderma, pregnancy or lactation, psychiatric disorders, or other conditions were excluded. Plasma samples used in the study were obtained at the time of BC diagnosis. Tissue samples were collected from patients who had not undergone any oncological treatment before surgery to avoid confounding factors. A total of 13 patients were included in the tissue study, with 13 peritumoral tissue samples and 8 tumoral tissue samples available for analysis (**Figure 6**). The lack of a control group is a limitation in the study of tissue biopsies.

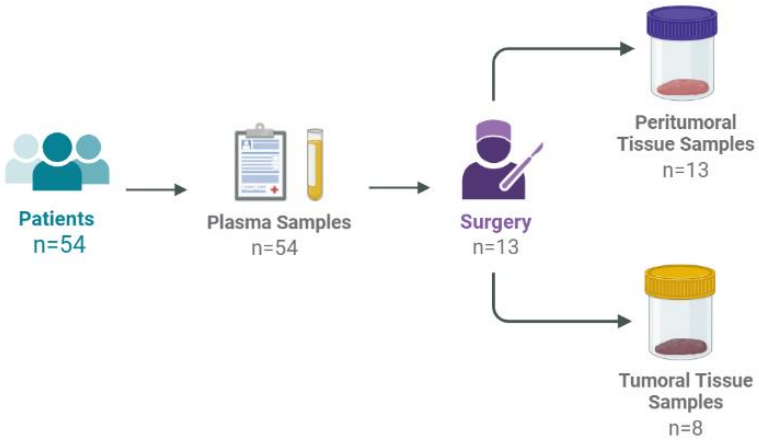


Figure 6. **Sample collection scheme of BC patients.**

3.2. Biological sample collection

At the time of diagnosis, two blood samples (10ml each) were collected from each participant before any treatment initiation. One of the blood samples was collected in an EDTA-containing tube to prevent coagulation. Samples were processed to obtain aliquots of serum and plasma via centrifugation (2.500xg, 15min, 4ºC) and stored at -80ºC until being analyzed. The samples obtained from the control group underwent identical processing procedures.

A surgical biopsy was performed to obtain tissue samples. During surgery, an incision was made in the affected region to get tumor tissue, and from the adjacent area in tissue to obtain peritumoral tissue. Samples of both types of tissue were taken using sterile surgical instruments. The samples obtained were immediately placed in suitable containers and kept at a controlled temperature of 4°C to preserve their integrity and avoid metabolic alterations. Tissue samples were processed by homogenization and extraction techniques, using appropriate solutions to ensure metabolites' stability in the tissue. The extracts obtained were subjected to a derivatization process to improve the detection and quantification of the metabolites. Derivatized extracts were stored under deep-freezing conditions at -80°C to ensure their integrity and stability until metabolomic analysis.

To ensure participant anonymity, all samples were labeled and identified using assigned numerical codes as part of the anonymization process.

3.3. Biological sample analyses

Targeted metabolomics was employed to determine plasma and tissue concentration of 74 metabolites involved in different metabolic pathways[40]. The metabolic categories analyzed were carbohydrate, amino acid, lipid, metabolism of cofactors and vitamins, nucleotide, xenobiotic biodegradation, and energy metabolism. The determination was done using a combination of gas chromatography (GC), electron impact ionization (EI) mass spectrometry (MS) and a high-resolution time-of-flight (QTOF) analyzer (GC-EI-QTOF-MS determination) to provide an accurate identification of the components present in the sample even among those with similar masses[41].

3.3.1. Sample preparation

A 50 µL aliquot of plasma samples was deposited into a 1.5 ml Eppendorf tube and amalgamated with 200 µL of an 8:2 (v/v) methanol:water solution inclusive of internal standards. Following the mixture, the samples were subjected to vortexing and centrifugation at 15000 rpm and a temperature of 4ºC for five minutes. Subsequently, supernatants (200 µL) were decanted into a fresh tube and subjected to evaporation in a SpeedVac at 45 ºC. The samples were then reconstituted using 30 µL of methoxyamine and placed in an incubator at 37ºC for 90 minutes. The final step involved silylating the samples with 45 µL of MSTFA supplemented with 1% TMCS at ambient temperature for 60 minutes.

Tissue samples were weighed prior to preparation. These samples were combined with 200 µL of a methanol:water (8:2, v/v) solution, which also contained internal standards. A bullet

blender equipped with stainless-steel balls was utilized for the homogenization of the samples. Post-homogenization, the samples underwent centrifugation at 15000 rpm and 4ºC for five minutes. The supernatants (200 µL) were subsequently transferred into a clean tube and evaporated using a SpeedVac set at 45 ºC. The samples were later reconstituted by adding 30 µL of methoxyamine and were incubated at 37ºC for 90 minutes. The final step of the process involved silylating the samples with a combination of 45 µL of MSTFA and 1% TMCS, held at room temperature for 60 minutes.

3.3.2. GC-(EI)qTOF methodology

The chromatographic separation process was facilitated using helium as the carrier gas, boasting a purity level exceeding 99.99%, at a consistent flow rate of 1.1 mL/min. The GC oven's initial temperature was maintained at 60°C. A minute after the injection, the oven's temperature progressively increased by 10°C per minute, reaching a peak temperature of 320°C, and was sustained at this temperature for 10 minutes. Sample injections were carried out in a 1:20 split mode, with the injection temperature set at 250ºC.

Detection was executed utilizing a mass spectrometer (MS) operating in an electron ionization mode at 70 eV. It was further monitored in full-scan mode within the mass-to-charge ratio (m/z) range of 50–600, achieving an acquisition rate of 5 spectra per second. The ion source temperature was maintained at 250°C, whereas the quadrupole temperature was set at 200°C.

The analytes' identification and semi-quantification were achieved using relative units (RU), determined by the ratio of the compound area to the internal standard area. More specifically, ions were selected and used for quantitation based on their impact electron spectra (70 eV), and the primary specific ions recorded in the Fiehn-pct-2013 spectral library.

To supplement the information, the equipment used was constituted of a 7890A gas chromatograph paired with a 7200-quadruple time-of-flight mass spectrometer equipped with an electron impact source. In addition, the system was equipped with a 7693 autosampler module, along with a J&W Scientific HP-5MS column (30ms 0.25 mm, 0.25µm) from Agilent Technologies (Santa Clara, CA, USA). The calibration curve was derived by correlating the standard concentrations with the peak area.

3.4. Statistical analysis

The central part of the study was done using the tool *MetaboAnalyst 5.0*. *MetaboAnalyst* is a web-based platform that offers a suite of computational tools for analyzing and interpreting metabolomics data. It can perform statistical analysis and visualize the metabolite profiles obtained from different group samples.

Univariate analyses were used as an exploratory approach to the data using one-factor studies. These analyses allow the identification of those metabolites that are significantly increased or decreased in one group compared to another. Univariate analyses such as *T-test*, *Fold Change* (FC), *False Discovery Rate* (FDR), and Volcano plot analyses were used. For this purpose, a *p-value* and FDR<0.05 were established. Data were normalized using logarithmic base 10 transformations, and *missing values* were replaced by 1/5 of the min positive values of their corresponding variables. The graphical representation of the Volcano plot shows that the x-axis indicates the fold change (log2 scale), and the y-axis represents the statistical significance (-log10 of *p-value*).

Correlation heatmaps were also performed, representing the correlations between multiple variables and showing patterns and associations. These associations can help identify strongly correlated metabolites that may be involved in the same metabolic pathway or in a common biological process.

Chemometric assays such as *Principal Component Analysis* (PCA) or *Partial Least Squares - Discriminant Analysis* (PLS-DA) were also used. These multivariant analyses aim to identify patterns and relationships between the data and extract relevant information, including classification, regression, and clustering methods. PLS-DA, a supervised regression technique, was used to find factors that best distinguished established groups of samples by connecting the predictor matrix (metabolomic data) and the response matrix (predefined categories). In contrast, PCA was used to reduce dimensionality and correlation, and to explain the variance in the original data. Therefore, both were used for a better understanding of the data.

Receptor Operating Characteristic (ROC) curves analysis was also used in our investigation. ROC curve analysis is a powerful tool used to assess the diagnostic ability of a binary classifier system. The degree of separability is represented by the Area Under the Curve (AUC), which aids in differentiating between positive and negative outcomes. The CI (Confidence Interval) is a measure of uncertainty associated with the AUC value. The performance of our classification model was also visualized using a confusion matrix, which allowed us to see the predictions that were correct and incorrect.

For further statistical evaluations, the sample groups were compared using *Rstudio* with R version V4.2.3 (2023-03-15). *Mann-Whitney U tests* applying the *Benjamini-Hochberg correction* (FDR) were performed. Differences between groups were considered significant when $p<0.05$. Table1 was generated for the recompilation of patient character characteristics. Quantitative variables were expressed as median variables (interquartile range), and qualitative variables as a percentage of the total number of participants. Further, R was also used to generate correlation networks and bubble plot visualizations. Specifically, these bubble plots were designed to group metabolites by their primary metabolic pathways to better comprehend the relative metabolic expression across different groups. This analysis was carried out with the package "*ggplot2"* in R.

Metabolites associated with LA and LB BC subtypes were identified using *unadjusted logistic regression* analysis. Multivariable logistic regression included potential confounding factors described previously (i.e., menopausal state, tumor size, tumor grade, and histological classification). The odds ratios (ORs) and corresponding 95% confidence intervals (CIs) were calculated, employing LA BC as the reference category. A logistic regression approach using an *elastic net penalty* was adopted considering the data's dimensionality and collinearity. This approach concurrently performs variable selection and regularization, effectively addressing multicollinearity concerns. The elastic net method merges the L1 and L2 penalties of the Lasso and Ridge techniques, respectively. To ensure the most accurate selection, we used internal 10-fold cross-validation (CV) during the analysis. Alpha and lambda values were 0.5 and 0.2, respectively. These analyses were conducted using the "*glmnet*" package in R.

4. RESULTS

4.1. Clinical characteristics of BC patients

The main clinical characteristics of BC patients can be summarized in **Table 1**. Most of the patients exhibited the onset of BC during the postmenopausal stage, with an age range between 40 and 75 years. The study population demonstrated a low prevalence of comorbidities, while there was a prevalence of familial oncologic antecedents. Ductal carcinoma emerged as the most frequently diagnosed pathological tumor type, with a predominance of histological grade II and the LB BC molecular subtype.

Table 1. Clinical characteristics of BC patients

	BC patients in plasma analysis (n=54)	BC patients in tissue analysis (n=13)
Clinical characteristics		
Diagnostic age (years)	57.0 (49.0-67.8)	57.0 (49.0-71.0)
Smoking, n (%)	6 (11.1)	2 (15.4)
Enolism, n (%)	5 (9.3)	3 (23.1)
Menopausal status, n (%)		
Premenopausal	15 (27.8)	4 (30.8)
Perimenopausal	4 (7.4)	2 (15.4)
Postmenopausal	35 (64.8)	7 (53.8)
Intake of oral contraceptives, n (%)	15 (27.8)	2 (15.4)
Parity, n (%)	48 (88.9)	11 (84.6)
Body Mass Index (BMI)	29.3 (24.7-33.7)	26.7 (25.0-32.0)
Diabetes Mellitus, n (%)	8 (14.8)	2 (15.4)
Hypertension, n (%)	15 (27.8)	4 (30.8)
Dyslipidaemia, n (%)	12 (22.2)	5 (38.5)
Pulmonary pathology, n (%)		
Tuberculosis (TBC)	1 (1.9)	1 (7.7)
Others	3 (5.6)	1 (7.7)
Cardiac pathology, n (%)	6 (11.1)	-
Oncologic family history, n (%)	31 (57.4)	7 (53.8)
Cancer characteristics		
Affected breast, n (%)		
Left	30 (55.6)	8 (61.5)
Right	21 (38.9)	4 (30.8)
Bilateral	3 (5.6)	1 (7.7)
Pathological anatomy of the tumor, n (%)		
Ductal	41 (77.4)	8 (61.5)
Lobular	9 (17.0)	4 (30.8)
Others	3 (5.7)	1 (7.7)
Histological Grade, n (%)		
I	13 (24.5)	3 (23.1)
II	38 (71.7)	9 (69.2)
III	2 (3.8)	1 (7.7)
Estrogen Receptors	98.0 (90.0-100.0)	97.0 (90.0-100.0)
Progesterone Receptors	70.0 (15.0-90.0)	70.0 (40.0-90.0)
Ki67 antigen tumor in biopsy	23.0 (12.0-30.0)	18.0 (12.0-30.0)
Molecular subtype		
Luminal A	17 (31.5)	4 (30.8)
Luminal B	37 (68.5)	9 (69.2)
Tumor Size, n (%)		
T1	21 (37.7)	1 (7.7)
T2	26 (49.1)	9 (69.2)
T3	7 (13.2)	3 (23.1)
Nodes (TNM system), n (%)		
N0	33 (62.3)	7 (53.8)
N1	17 (32.1)	5 (38.5)
N2	3 (5.7)	1 (7.7)

Normal distributed continuous data is presented as a mean (standard deviation), not normal distributed continuous data is presented as median (interquartile range); finally, for categorical variables data is presented as frequency percentage (%).

A comparison was also made by stratifying patients into LA and LB BC (**Supplementary table 1**). Significant differences were observed in the levels of Progesterone Receptor (PR) and Ki67 antigen in tumor biopsies, two fundamental factors in biopsy analyses for determining the patient's BC molecular subtype. Significant differences were further noted in tumor grade, with

LB patients more frequently presenting grade II and LA patients predominantly demonstrating grade I.

A comparative evaluation was also conducted, distinguishing between patients with peritumoral tissue samples and those with tumoral tissue samples (noting that patients with tumoral tissue are the same as those with peritumoral tissue). No substantial disparity was detected in the clinical attributes across these patient groupings (**Supplementary table 2**).

Biochemical main profiles of patients LA and LB BC showed no significant difference based on *p-value* (**Table 2**). However, urea levels were found at lower levels in LA BC patients.

Table 2. Biochemical characteristics of BC patients of the plasma study. Stratified by molecular subtype LA and Luminal B.

	Luminal A (n=17)	**Luminal B (n=37)**	**p value**
Glucose (mg / dL)	98.0 (85.5-105.0)	94.0 (87.5-110.0)	0.715
Creatinine (mg/dL)	0.7 (0.6-0.7)	0.7 (0.6-0.8)	0.400
Cholesterol total (mg/dL)	192.0 (187.0-206.3)	197.0 (177.0-227.0)	0.782
HDL (mg/dL)	59.0 (51.5-66.0)	57.5 (46.0-72.5)	0.820
LDL (mg/dL)	118.0 (104.0-126.5)	113.5 (97.2-133.7)	0.642
VLDL (mg/dL)	23.0 (15.5-29.0)	22.0 (16.0-31.0)	0.982
TG (mg/dL)	114.5 (75.5-145.0)	115.0 (81.5-153.0)	0.956
Urea (mg/dL)	29.0 (25.0-37.0)	34.0 (28.5-39.0)	0.156
Bilirubin (mg/dL)	0.4 (0.3-0.5)	0.4 (0.3-0.5)	0.577
LDH (U/L)	178.5 (163.5-194.5)	161.5 (148.7-189.2)	0.388
Haemoglobin (g/dL)	13.3 (12.8-14.2)	13.7 (12.8-14.4)	0.730

Values are provided as median (interquartile range). HDL: high-density lipoproteins; LDL: low-density lipoproteins; VLDL: very low-density lipoprotein; TG: triglycerides; LDH: lactate dehydrogenase.
The significance was determined by the Mann-Whitney U-test, $p < 0.05$.

4.2. Metabolomic tissue analysis

4.2.1. Group analysis: absence of significant differences

Inter-group comparisons of peritumoral and tumoral tissues were performed. It was first stratified by molecular subtype, LA and LB BC. Moreover, data were stratified using characteristics based on previous studies describing them as factors that may affect the diagnosis and progression of the disease: histological classification, histological grade, menopausal state, and tumor size. The specific characteristics of each patient are shown in **Supplementary Table 3**.

The inter-group comparisons applied in tumoral and peritumoral tissue samples yielded not significative differences (using *Mann Whitney Analysis p-value* < 0.05 with *FDR adjustment*) between all the groups analyzed, possibly due to the limited sample size.

4.2.2. Comparison of tumoral and peritumoral tissue

All metabolite concentrations were compared by categorizing the tissue samples into tumoral and peritumoral tissue. Univariate and multivariate analyses were performed to discern potential differences in metabolite concentrations between these two tissue subtypes. As evidenced in the PLSD-DA (**Figure 7A**), a clear differentiation between the groups can be seen, attributable to the distinct metabolic profiles. Univariate analysis Mann-Whitney U test with FDR adjustment showed significant differences in the concentrations of different metabolites between two groups: ribonic acid, taurine, 4-hydroxyproline, ethanolamine, a-ketoglutaric acid, xylonic acid, lactic acid (overexpressed in tumoral tissue) and ornithine (overexpressed in peritumoral tissue) (**Supplementary table 4**). The volcano plot displayed the expression of the significative metabolites (**Figure 7B**). As shown in **Figure 7C**, a general trend towards higher metabolite expression in tumor tissue is discernible.

Through the VIP score, the most important variables for the discrimination of the groups were ribonic acid, taurine and 4-Hydroxyproline (**Figure 7D**). According to the VIP score and *p-values*, ROC and matrix confusion analyses for ribonic acid, taurine and ornithine were the most predictive metabolites in BC tissue (AUC = 0.917) (**Supplementary Figure 1**) The differential concentration of these metabolites is shown in **Figure 7E**.

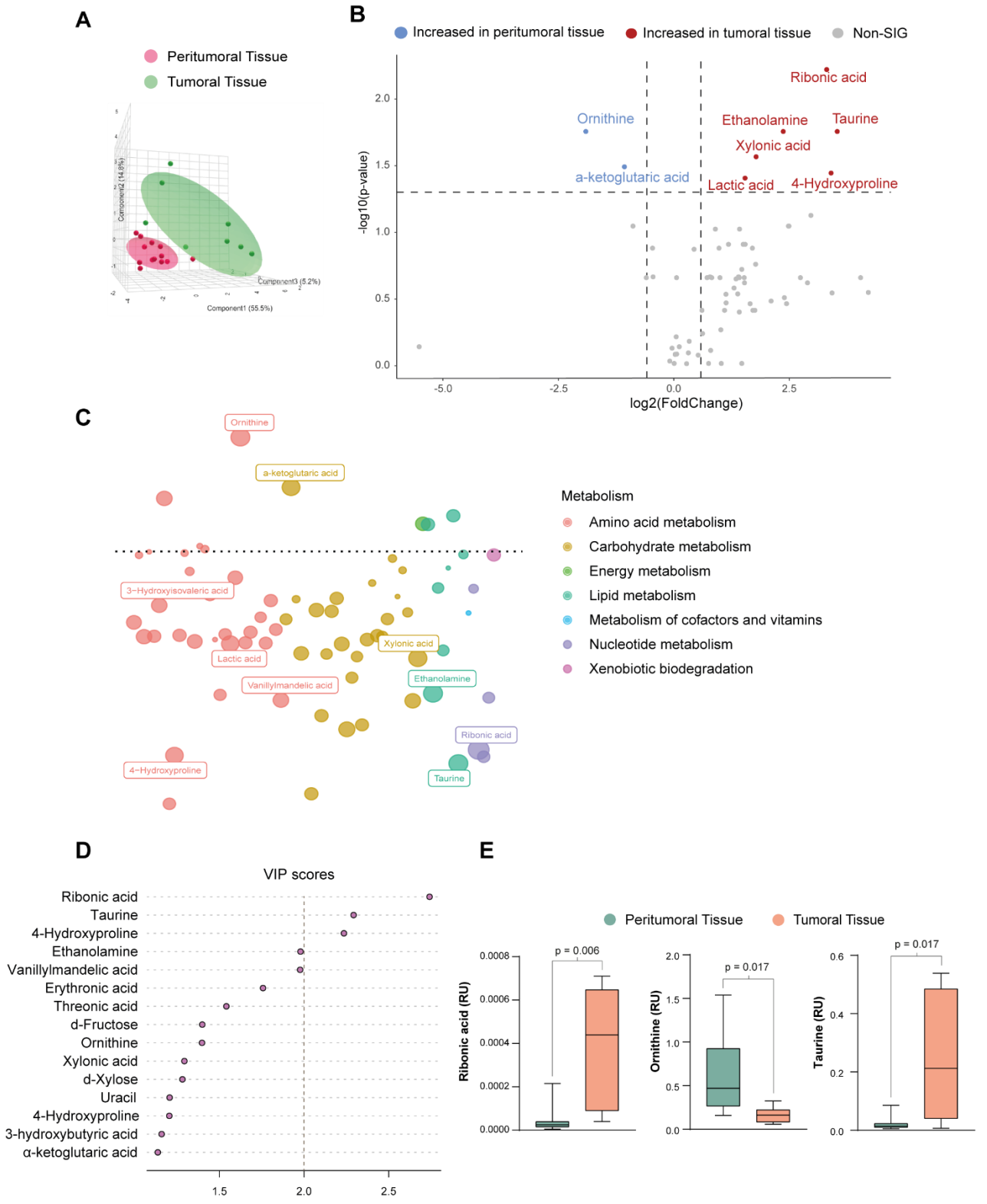


Figure 7. Evaluation of the overall differences and discriminative markers between tumoral and peritumoral luminal BC tissue.

(A) Three-dimensional (3D) PLS-DA, (B) Volcano Plot representing the expression of discriminative metabolites between BC tumoral tissue and BC peritumoral tissue, (C) Bubble plot illustrating the relative expression levels of metabolites, color-coded by metabolism. Metabolites above the x-axis are more expressed in peritumoral tissue, while those below are more expressed in tumoral tissue. The size of the bubbles corresponds to the statistical significance (p-value) of the differential expression, with larger bubbles indicating greater significance, (D) VIP score of significant metabolites and (E) Boxplot for significant metabolites.

The significance was determined by the Mann-Whitney U-test, $p < 0.05$.

4.3. Metabolomic plasma analysis

4.3.1 Comparison of BC patients with the control group

Univariate and multivariate analyses were conducted to uncover potential variations in metabolite concentrations between these groups, thereby revealing potential biomarkers of luminal BC.

The bi-dimensional (2D) PLS-DA (2D PLS-DA) distinctly delineates the two groups, underscoring a clear differentiation between them (**Figure 8A**).

Mann-Whitney U test revealed substantial global differences between BC patients and the control group (**Supplementary Table 5**). The differential expression of these metabolites in BC is visualized in the volcano plot (**Figure 8B**). A tendency of higher expression of metabolites in BC can be observed in **Figure 8C**, showing that carbohydrate metabolism is the most notably altered metabolic pathway in BC patients. Nevertheless, several metabolites with a high *p-value* were underexpressed in the BC group: ethylmalonic acid, xylonic acid and glycerol.

Based on VIP score, a group of metabolites - maltose, hypoxanthine, 3-phosphoglyceric acid, xylonic acid, ethylmalonic acid, and glycerol - stood out as the most discriminative variables characterizing BC patients (**Figure 8D**).

The model incorporating xylonic acid, ethylmalonic acid, glycerol, 3-phosphoglyceric acid, and maltose displayed accurate discriminative power for BC, as corroborated by ROC curve and confusion matrix analyses, achieving AUC = 1 (**Supplementary Figure 2**). **Figure 8E** illustrates the concentrations of the significant metabolites mentioned before.

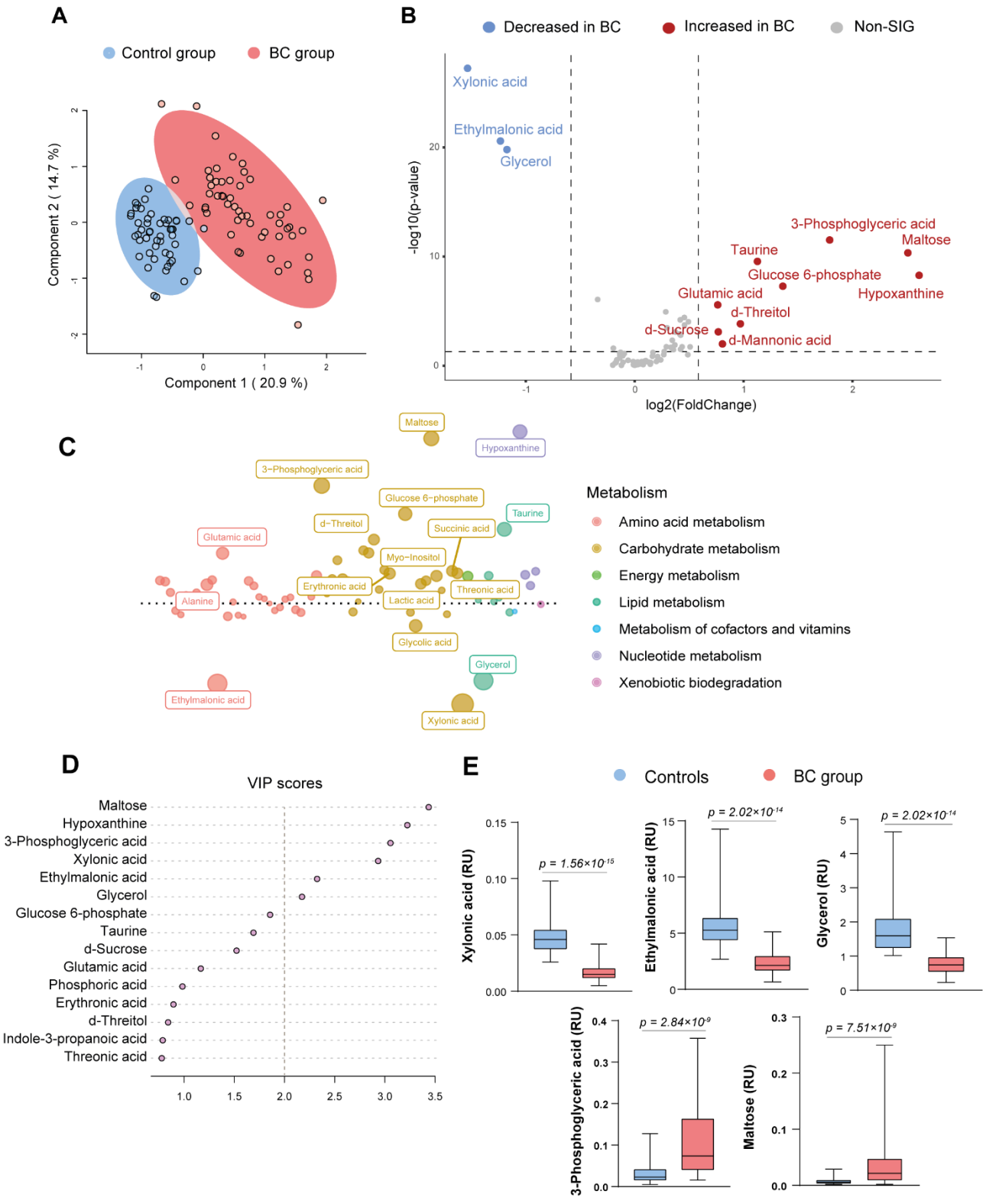


Figure 8. Analysis of significantly different metabolites between the control group and BC patients

(A) PLSDA 2D scores plot distinguishing BC group from the control group, (B) Volcano plot representing the expression of discriminant metabolites, (C) Bubble plot illustrating the relative expression levels of metabolites color-coded by metabolism. Metabolites above the x-axis are more expressed in BC patients, while those below are underexpressed in BC. The size of the bubbles corresponds to the statistical significance (p-value) of the differential expression, with larger bubbles indicating greater significance, (D) VIP score of significant metabolites and (E) Boxplot for significant metabolites.

The significance was determined by the Mann-Whitney U-test, $p<0.05$.

4.3.2. Delineation between LA and LB BC

Distinct metabolomic variations between BC molecular subtypes were not entirely evident through linear analysis. Upon individual examination through the Mann Whitney U test, phenylalanine and succinic acid displayed differential concentrations (**Supplementary Table 6** and **Figure 9A**). However, no metabolite differences were found to be significant after adjustment for FDR in the Mann-Whitney U test. Additionally, multivariate analysis also did not differentiate between the two groups. Nevertheless, a tendency of overexpression for most of the metabolites was noticed in the LB BC group (**Figure 9B**). The metabolisms that showed significant changes were carbohydrate and amino acid metabolism.

Maltose and 3-phosphoglyceric acid, as indicated by the VIP score (**Figure 9C**), emerged as potential discriminators between LA and LB BC. The ROC curve analysis for maltose, phenylalanine and succinic acid, resulted in an AUC of 0.45, indicating a poor performance of this model with a lack of reliability (**Supplementary Figure 3**).

In view of BC's inherent heterogeneity, we performed a multivariate analysis, incorporating key influential characteristics as discussed in section ***1.3.2. Clinical and Histopathological Characteristics*** (menopausal state, histological grade, pathological anatomy of the tumor, and tumor size). ANCOVA results showed the retention of succinic acid's significance and the emergence of two new significant metabolites, DL-2-Hydroxyglutaric acid and 2-Hydroxyisobutyric acid (**Supplementary Table 6**).

Supplementary Table 7 presents results for the univariate logistic regression analysis to explore associations of metabolites in LA and LB BC. Two metabolites, DL-2-hydroxyglutaric and succinic acid, showed increased concentration in LB BC when compared to LA BC: with odds ratio (OR) 2.05 (95% CI, 1.09-4.31, P=0.038) and 1.97 (95% CI, 1.06-3.98, P=0.042) respectively. Nonetheless, after adjusting for potential confounding factors, none of the metabolites maintained significance. **Supplementary Table 8** lists the 5 metabolites in elastic net regression analysis for LA and LB BC. DL-2-Hydroxyglutaric and succinic acid, following previous results, showed positive associations with LB BC. Aditionally, indolelactic acid, myo-inositol, and 3-phosphoglyceric acid, were also found in higher concentrations in LB BC (**Supplementary Table 7**). Nevertheless, all associations became insignificant after further covariates adjustment.

The differential concentration between groups of the significant metabolites mentioned is displayed in **Figure 9D**.

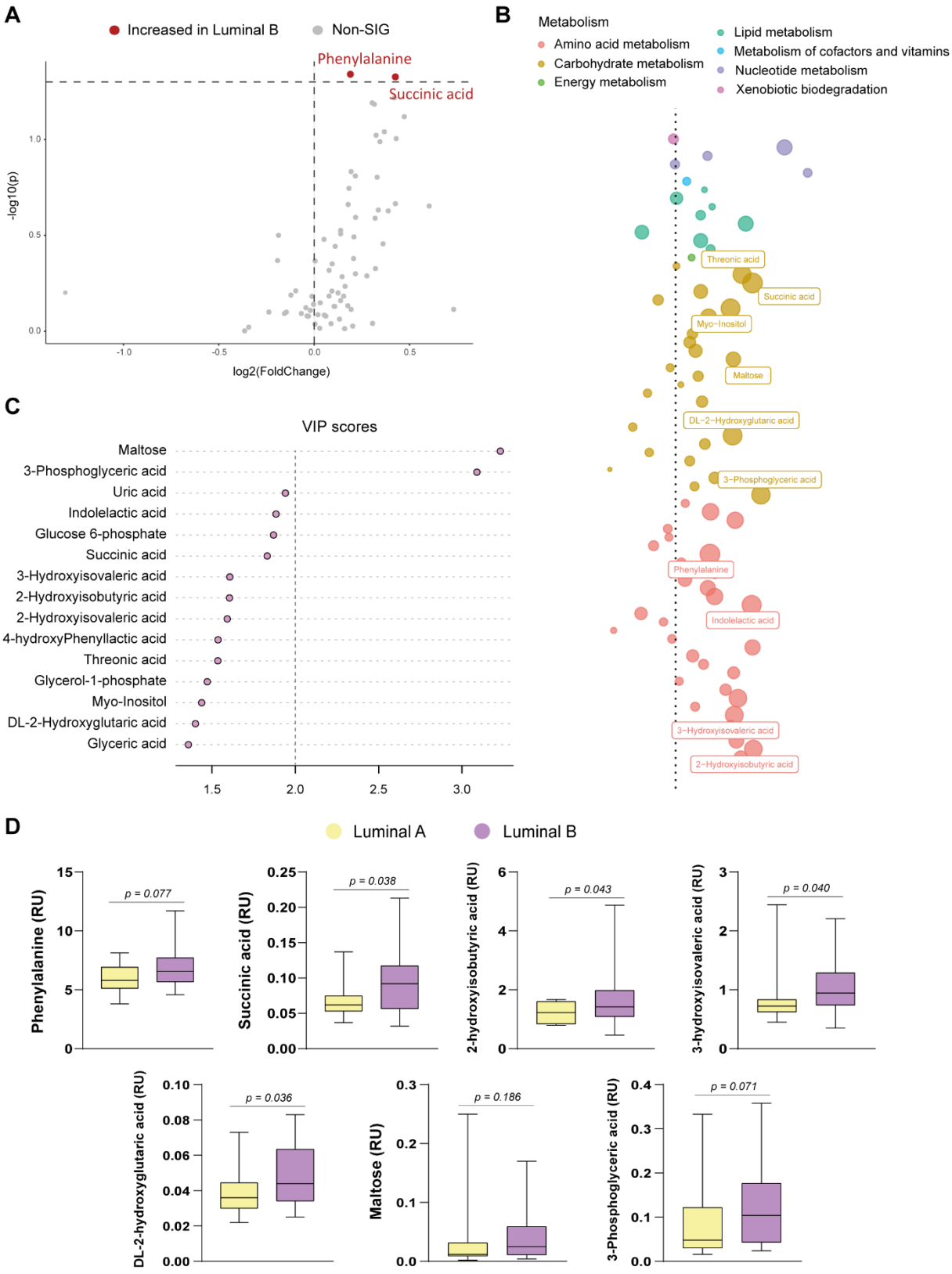


Figure 9. Analysis of significantly different metabolites between LA BC and LB BC.
(A) Volcano plot representing the expression of discriminant metabolites, (B) Bubble plot illustrating the relative expression levels of metabolites, color-coded by metabolism. Metabolites located to the left of the x-axis are more expressed in LA, while those to the right are more expressed in LB. The size of the bubbles corresponds to the statistical significance (p-value) of the differential expression, with larger bubbles indicating greater significance, (C) VIP score of significant metabolites (D) Boxplot for significant metabolites.

The significance was determined by the Mann-Whitney U-test, $p < 0.05$.

4.4. Correlation analysis between tumoral tissue, peritumoral tissue and plasma

The correlation analysis was performed to understand the interconnections and potential regulatory relationships among the metabolites in patients' tumoral and peritumoral tissue and plasma. This analysis used samples from 13 (4 patients LA BC and 9 patients LB BC) with both plasma and tissue samples (**Supplementary Table 3**).

The neural network representation revealed a dense network of connections among the tumoral tissue nodes, indicating a strong interplay between the metabolites in this type of tissue. In contrast, the network showed an increased number of connections between peritumoral tissue and plasma, with fewer connections between tumoral tissue and the other two groups (**Figure 10**).

A correlation heatmap was developed to further investigate these relationships and how the different BC molecular subtypes influence them (**Figure 11**). The heatmap compares patients' plasma in front of their tissue, specifically LA BC peritumoral tissue vs. plasma, LB BC peritumoral tissue vs. plasma, and LB tumoral tissue vs. plasma (this test could not be done for LA BC tumoral tissue due to the lack of samples). The objective was to investigate whether there are tissue-systemic correlation and to identify differential correlations between the LA and LB BC molecular subtypes. The majority of correlations were negative for LB BC tumor tissue vs. plasma, with fewer strong correlations than in the peritumoral tissue. The metabolites with the highest correlation were erythronic acid, threonic acid (negative correlation) and taurine (positive correlation). Most correlations for LB peritumoral tissue vs plasma were positive, with strong positive correlations observed with 4-hydroxyproline, 3-hydroxybutyric acid, and ethanolamine, and a strong negative correlation with phosphoric acid. In contrast, LA BC peritumoral tissue showed a mix of positive and negative correlations with plasma and displayed the highest overall correlation. The metabolites most negatively correlated were urea, tetradecanoic acid, ribonic acid, proline, oleic acid, dodecanoic acid, D-arabitol, and benzoic acid. Metabolites showing strong positive correlations included ribonic acid, malic acid, glycolic acid, Fumaric acid, benzoic acid, a-ketoglutaric acid, 3-hydroxyisovaleric acid, and 2-keto-3-methylvaleric acid.

Further analysis was carried out to assess the variation of specific metabolites across different groups. In the case of metabolites such as urea, dodecanoic acid, and 3-hydroxyisovaleric acid, a negative correlation was observed in peritumoral tissue of LA BC. In contrast, in LB BC, either no correlation was found, or a weak negative correlation was detected. On the other hand,

metabolites like phosphoric acid and 4-hydroxyproline displayed an inverse correlation pattern between the two subtypes, being negatively correlated in LB BC, in contrast to LA BC, where a positive correlation was noted. Another significant finding was threonic acid, which exhibited a negative correlation in tumor tissue, but a positive correlation in peritumoral tissue.

Focusing on tumor tissue, the metabolite taurine, which positively correlated with plasma, was interesting. This behavior was markedly different from its behavior in peritumoral tissue, where no correlation was observed.

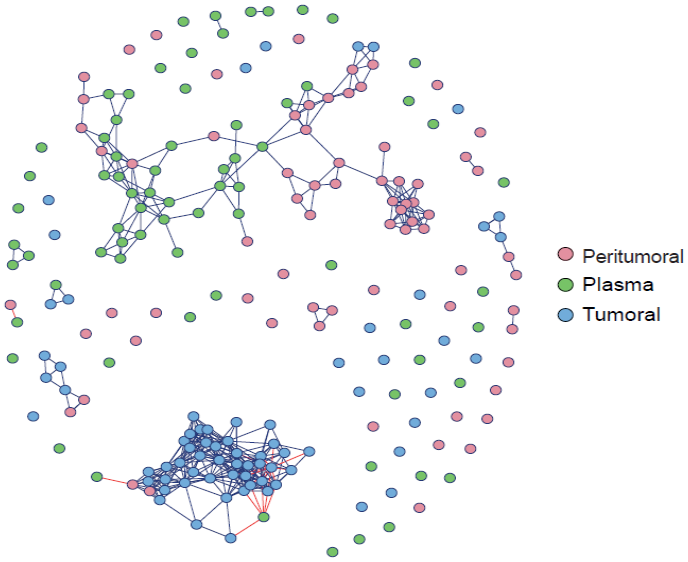


Figure 10. Correlation network between tumoral tissue, peritumoral tissue and plasma.
Each node represents a metabolite, with edges denoting significant correlations between them. Only correlation coefficients values greater than 0.3 are represented in the graph.

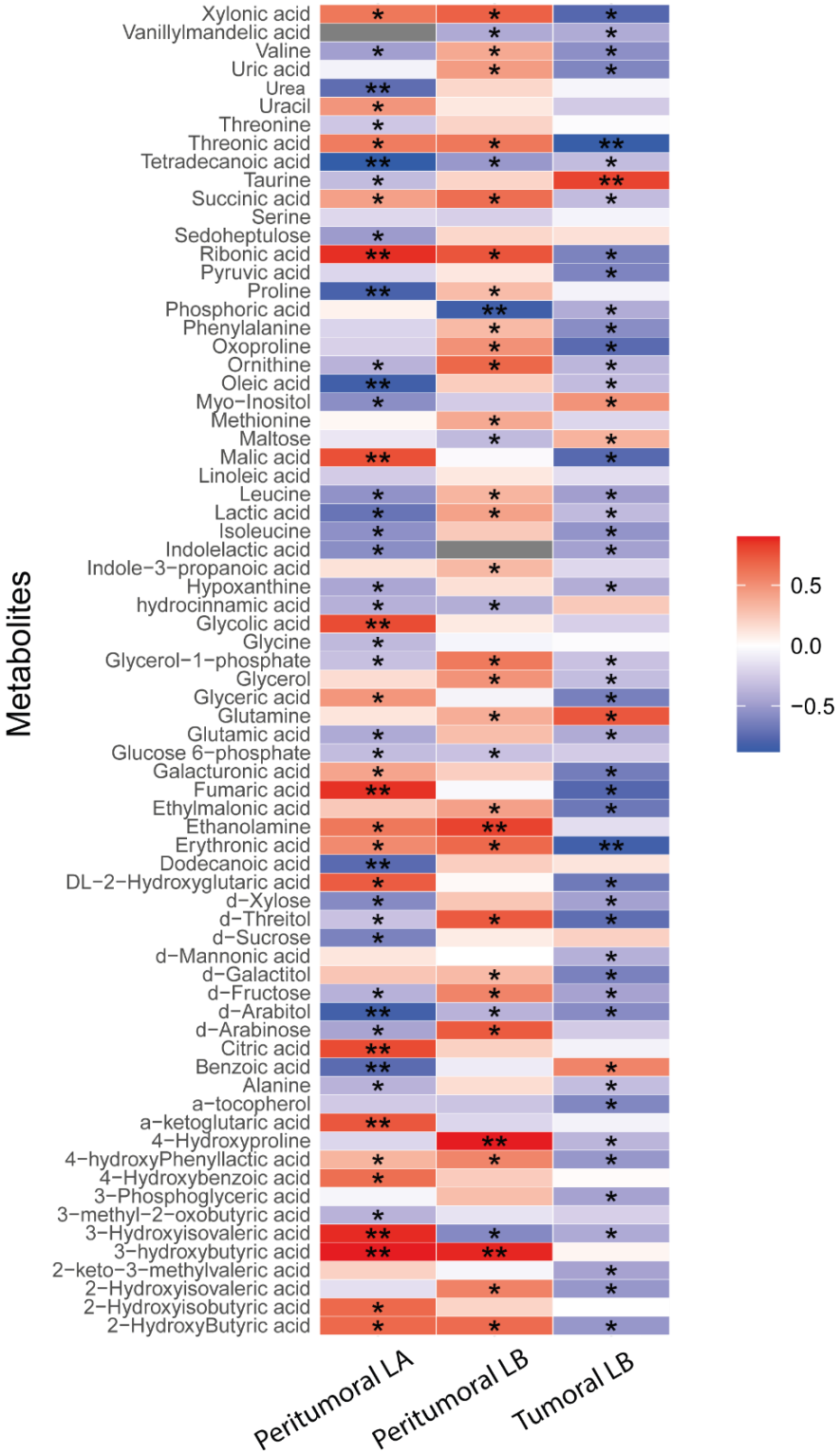


Figure 11. Correlation heatmap between tumoral tissue, peritumoral tissue and plasma.

Each cell corresponds to the correlation coefficient value, with warmer colors (red) indicating stronger positive correlations and cooler colors (blue) indicating stronger negative correlations. Correlation coefficients greater than 0.3 are marked with a single asterisk (*), whereas those greater than 0.8 are marked with double asterisks (**).

5. DISCUSSION

Metabolomic studies in cancer research, hold considerable potential for discovering new biomarkers and unraveling the intricate metabolic changes accompanying tumorigenesis. Metabolic alterations are an established hallmark of cancer, and metabolomic studies can provide essential information in this regard.

In the present study, we explore the metabolomic landscape of breast cancer, focusing on the most common molecular subtypes LA and LB, which display notable differences in diagnosis, prognosis, and therapeutic responses. A comprehensive understanding of subtype-specific biomarkers could improve therapeutic strategies adapted to each patient's metabolic profile. To accomplish this, we analyzed 74 metabolites for plasma and tissue samples (peritumoral and tumoral), distributed across seven distinct metabolic pathways. These metabolites were profiled to identify differential expressions, offering a detailed picture of metabolic variations in these BC subtypes. To further refine the comparative analyses, we incorporated a series of clinicopathological cofounding factors (menopausal status, tumor size, tumor grade and node status) to analyze their potential influence on the metabolic profiles of these BC molecular subtypes.

LA and LB are subtypes of BC distinguished by their gene expression and hormone receptor profiles. LA BC is known to have higher levels of PR compared to LB BC. PR is associated with a better prognosis and response to hormonal therapies due to the dependence of these cancer cells on hormonal signals for growth and survival. On the other hand, LB BC generally shows a lower expression of PR, which may contribute to their increased aggressiveness and resistance to hormonal therapies[42]. Furthermore, Ki67 antigen, a cell proliferation-associated protein found in all active cell cycle phases, is overexpressed in LB BC when compared to LA BC[42]. This further enhances the aggressivity of LB tumors by providing them with greater aggressiveness and increased proliferation potential. These findings are consistent with the values obtained in our patient’s study.

Regarding the biochemical variables obtained from blood tests, our study revealed no considerable differences between LA and LB BC subtypes. This suggests that despite existing molecular and cellular differences between these subtypes, these disparities do not manifest as observable changes in routine blood biochemical indicators.

For the **metabolomic study of tissue,** the low number of samples was a limiting factor when stratifying these groups by molecular subtype or critical characteristics without finding any

significant differences. Therefore, profiling of tumor tissue versus peritumoral tissue of LA and LB BC patients together was performed. A generalized higher elevation in metabolite concentrations in tumor tissue was found, indicating the heightened energy demand characteristic of cancer cells. This finding underscores the intricate relationship between metabolic reprogramming and cancer progression, highlighting the need for a deeper understanding of the altered metabolic landscape in tumor tissues.

In our study, we identified numerous differentially expressed metabolites, and multivariate analysis successfully differentiated the two groups. Notably, metabolites such as ribonic acid, ornithine, taurine, ethanolamine, and 4-hydroxyproline were the most significant in distinguishing between the groups. All of them showed a higher concentration in tumoral tissue except ornithine, whose concentration was higher in peritumoral tissue.

Ribonic acid is a monosaccharide sugar acid that can be enzymatically converted into ribose-5-phosphate by ribokinase, a reaction that can also occur in reverse [43]. Elevated ribonic acid might reflect an increased demand for ribose-5-phosphate, crucial for synthesizing of nucleotides, nucleic acids, and combating oxidative stress[44].

On the other hand, ornithine was found in higher concentrations in peritumoral tissue. This amino acid is a key product of the urea cycle. The increased concentration of ornithine in peritumoral tissue may be associated with tumor-induced metabolic alterations in the surrounding tissue. Polyamines, which are products of ornithine metabolism, have been implicated in cell proliferation and may play a role in promoting tumor growth[45].

Taurine, a sulfonic amino acid, is involved in numerous biological functions such as osmoregulation, detoxification, and modulation of cellular signalling[46]. In our study, we observed an elevated taurine concentration in tumoral tissue. Taurine participates in critical cellular processes like proliferation, oxidative stress regulation, and apoptosis, vital for cancer cell survival and growth. Previous studies underscore the importance of taurine in tumor development, with the taurine transporter being a significant target in other cancers, such as colorectal cancer, due to its association with enhanced survival and anti-apoptotic effects[47]. Ethanolamine is an organic compound that is part of many biological processes. It acts as a critical component of phospholipids, which are essential for the integrity and functionality of cell membranes. Specifically, ethanolamine is an integral part of phosphatidylethanolamines, a type of phospholipid particularly prevalent in cell membranes. In the context of cancer, it has been observed that changes in the phospholipid composition of cell membranes may be

characteristic of cancer cells. In particular, cancer cells often show an increased demand for phospholipids due to their rapid growth and proliferation, intensifying membrane biogenesis[48].

4-hydroxyproline, a key component of collagen, gives collagen its characteristic helical structure, which is essential for its function[49]. In BC, an increased 4-hydroxyproline concentration in tumoral tissues may indicate increased collagen turnover or remodeling rate, an important feature of the tumor microenvironment. In addition, increased collagen deposition and remodeling contribute to the desmoplastic reaction frequently observed in BC, which correlates with tumor progression, metastasis and worse prognosis[50]. An increase in 4-hydroxyproline could serve as a potential indicator of this reaction, thus providing valuable information on the underlying pathogenesis of breast cancer.

The metabolomic study of luminal BC patients compared to healthy individuals has been used as a baseline for understanding the metabolic reprogramming in patients presenting luminal BC subtypes. Many metabolites were found with large significant differences and most metabolites were found in higher concentrations in BC patients. However, three metabolites of high significance were found at a lower concentration than in control patients. These metabolites were xylonic acid, ethylmalonic acid, and glycerol, which are involved in carbohydrate metabolism, amino acid metabolism, and co-factor and vitamin metabolism, respectively. The metabolites xylonic acid, ethylmalonic acid, glycerol, 3-phosphoglyceric acid, and maltose have been identified as potential biomarkers for luminal BC according to the ROC curve (AUC = 1).

Xylonic acid is involved in the pentose phosphate pathway, which could potentially be downregulated in cancer patients as they rely more on glycolysis due to the Warburg effect[51] [52]. Ethylmalonic acid was also found at reduced levels in BC patients. It is a by-product of amino acid metabolism, and its reduced levels might indicate an increased consumption of amino acids for protein synthesis or alternative metabolic pathways in cancer cells[53]. Finally, reduced levels of glycerol could suggest a shift in lipid metabolism in cancer patients, potentially reflecting an increased utilization of glycerol for lipid biosynthesis to support rapid cell proliferation[54].

On the other hand, maltose, hypoxanthine, 3-phosphoglyceric acid, glycerol, and taurine were observed at higher concentrations in BC patients. Maltose, a disaccharide of glucose, could reflect an increased demand for glucose as an energy source in rapidly proliferating cancer cells, in line with the well-established Warburg effect[51]. Similarly, 3-phosphoglyceric acid, an intermediate in glycolysis, may also be indicative of the increased glycolytic activity in cancer

cells[55]. Hypoxanthine, a degradation product of purines, may indicate increased nucleotide turnover in cancer cells, reflecting the high replication rate and DNA repair mechanisms of these cells[56]. Taurine, a sulfonic amino acid, also exhibited higher concentrations in the plasma of BC patients. This observation aligns with our earlier findings in tumoral tissue, indicating the importance of taurine in the systemic metabolic alterations linked to BC. Consequently, consistently detecting elevated taurine levels in both tumoral tissue and plasma samples emphasizes its potential relevance in BC metabolism and progression. Furthermore, these data reaffirm previous studies which suggested that taurine could be used as a new biomarker for the early diagnosis of BC[57].

Referring to the **metabolomic comparison between LA and LB BC patients**, a noteworthy tendency towards elevated concentrations of most metabolites in the LB BC subtype has been discerned. This increase in metabolites could potentially signify the greater aggressiveness associated with LB BC[24].

Phenylalanine and succinic acid were observed in markedly higher concentrations in the LB subtype. Phenylalanine, an essential amino acid involved in protein biosynthesis, can be converted into tyrosine, another amino acid that plays a pivotal role in cellular signaling. Succinic acid is an essential component of the citric acid cycle, crucial for ATP generation. However, this may not indicate an increased usage in LB, but rather an accumulation, as cancer cells often exhibit metabolic dysregulation, with a reduced reliance on the citric acid cycle for energy production[30]. Moreover, maltose and 3-phosphoglyceric acid also appear to discriminate between these subgroups when evaluated conjointly, with both metabolites presenting a higher concentration in the LB BC subtype. Given that both these metabolites are associated with glycolysis, this could possibly be connected to an increased glucose demand and heightened glycolytic activity in LB BC cells[30].

When taking into account the previously mentioned covariates, some variations in the results were observed. The ANCOVA maintained the significance of succinic acid and revealed two additional metabolites, DL-2-hydroxyglutaric acid and 2-hydroxyisobutyric acid. These metabolites could be implicated in specific cellular processes within cancer cells, affected by the considered covariates. Particularly, L-2-hydroxyglutaric acid has been observed to promote cancer when accumulated in higher concentrations within cancerous cells and under hypoxic conditions. This occurs as it acts as an inhibitor of alpha-ketoglutarate-dependent enzymes, altering histone methylation patterns and affecting the expression of genes involved in cellular differentiation and tumor suppression. This would contribute to enhanced growth and survival

of LB BC cancer cells[58]. 2-Hydroxyisobutyric acid is a metabolite derived from the degradation of the amino acids arginine and proline, which could signify an increased demand for both amino acids. This suggests an increase in cellular proliferation, oxidative stress, and protein turnover[59]. Despite DL-2-hydroxyglutaric acid was found significant in the univariate logistic regression and elastic net regression, its significance did not sustain post adjustment for covariates. This suggests that its significance may be influenced by the interaction with these factors, underlining the importance of considering covariates when interpreting metabolic results.

Lastly, the analysis also identified myo-inositol and indole lactic acid as potentially relevant metabolites in LB BC. Myo-inositol, an essential component of phosphatidylinositol signals, is involved in various cellular functions, including cell growth and survival[60]. Indole lactic acid, a product of tryptophan metabolism, has been implicated in the modulation of immune response and resistance to apoptosis in cancer[61]. Both metabolites could be at higher concentrations in LB BC subtype due to an increased demand for cellular signaling and stress resistance.

We also conducted a **tissue-plasma correlation analysis** aimed at understanding how metabolic adaptations within local environments (tumoral and peritumoral tissue), trigger responses within the systemic, circulating environment. In this study, we found a more densely connected network within the metabolites of tumoral tissue, suggesting a higher level of metabolic interaction and interdependence within this environment. The high metabolic demand of tumor cells[48 62] and the complex reprogramming of metabolic pathways[63] that support tumor growth and survival could explain the findings. Metabolic biomarkers are released into circulation, either by tumor cells themselves or by surrounding cells in the tumor microenvironment, exhibiting varying levels of specificity. To explore the interrelationship between the tumoral, peritumoral, and plasma environments, provides insights into the systemic and localized metabolic alterations in BC patients.

Interestingly, more connections were observed between peritumoral tissue and plasma metabolites. This may occur due to the peritumoral tissue acting as an interface between the tumor and systemic environments. This suggests that metabolic changes in peritumoral tissue may be a sensitive indicator of both local and systemic metabolic status and may have potential as a new target for metabolic interventions and biomarker discovery for predicting cancer type and its prognosis.

The correlation map showed negative correlations for 86% of metabolites when comparing LB BC tumoral tissue versus plasma, which could reflect metabolic differences and altered metabolic exchange between the tumor and systemic environments in this subtype. Among others, positive correlation with taurine is noteworthy. Although taurine has been reported to play several roles in cancer, studies examining the impact of taurine on tumors are limited, and the precise mechanism underlying its antitumor capabilities remains unclear. It has been reported that taurine can protect cells against injury caused by potent oxidants and cytotoxic agents through scavenging activities[64]. Therefore, taurine, as an efficient antioxidant, may increase the amount of reactive oxygen species present in tumors, potentially slowing the spread of cancer[65]. However, further research is needed.

Contrary to tumoral tissue BC correlation with plasma, 73% of plasma metabolites correlated positively with peritumoral tissue metabolites. In our study, the levels of metabolites related to the TCA cycle, i.e., citric acid and succinic acid, were revealed to be highly positively associated between peritumoral tissue and plasma. Despite this, we failed on finding a positive correlation between malic acid and fumaric acid in, although their coefficients were close to zero (-0.030 and -0.037 respectively). Furthermore, for previous mentioned metabolites, we found a negative correlation between tumoral tissue and plasma. Cancer cells are under the Warburg effect, that leads to a major dependence on aerobic glycolysis instead of oxidative phosphorylation for energy production[66]. Our findings could be explained for this metabolic alteration.

On the other hand, peritumoral LA BC tissue showed an overall negative correlation with plasma. Ribonic acid, 3-hydroxyisovaleric acid, and 3-hydroxybutyric acid were found to be highly correlated within the analyzed variables. Previous studies have reported that 3-hydroxyisovaleric acid is found in high concentrations in the urine from ketoacidotic patients[67]. While ketoacidosis is not necessarily a consequence of cancer, there can be instances where cancer patients may develop ketoacidosis as a complication of their disease. Tetradecanoic acid, oleic acid, and urea, displayed the most significant negative correlations. A study with 22 patients with localized carcinoma of the bladder revealed an increased urea synthesis rate and net protein catabolism patients with bladder cancer[68]. Our results are consistent with these findings, as urea concentration in peritumoral tissue is higher when compared with plasma, and concentration of urea in plasma is decreased by renal excretion. However, it should be noted that the available data were limited in scope, and further investigation is required to validate and establish the significance of this correlation.

Finally, the possible limitations of all this study could be related to the low number of patients. Studies with a larger number of patients are necessary to validate these preliminary results.

6. CONCLUSIONS AND FUTURE PERSPECTIVES

This study highlights the ability to distinguish luminal-type BC patients from healthy individuals using xylonic acid, ethylmalonic acid, glycerol, 3-phosphoglyceric acid, and maltose as potential biomarkers. Despite this progress, a clear differentiation between the LA and LB molecular subtypes remains a significant challenge due to the profound molecular complexity and diversity inherent in BC subtypes.

The importance of metabolomic characterization of tissue within this context cannot be underestimated. This approach provides a unique insight into the metabolic dynamics when cancer is present, enabling the identification of potential biomarkers of therapeutic resistance and the uncovering of new therapeutic targets. Furthermore, it allows us to understand the changes occurring at the systemic level during the development of BC. In this study, we have shown that peritumoral and tumoral tissues of luminal BC patients have significant differences that may be of great importance in understanding how this disease develops.

Given the complexities in distinguishing between the LA and LB molecular subtypes it is clear that a multidimensional approach, utilizing more than just the metabolome, may be required for precise differentiation. Nevertheless, the metabolites identified in this study hold potential as biomarkers, especially when integrated with other clinical and pathological parameters in BC diagnosis.

Metabolomics brings the potential for less invasive techniques to differentiate molecular subtypes of breast cancer, a welcome alternative to current methods such as breast biopsy, which can cause patient discomfort and distress. This underscores the urgent need for continued research and development of more accurate, comprehensive methods for classifying and treating BC. Progress in these areas is key to enhancing treatment efficacy and ultimately improving the quality of life for BC patients.

7. SELF-ASSESMENT

The purpose of my stay at the Biomedical Research Unit was to expand my professional competence, applying the theoretical and practical knowledge I acquired throughout my studies in biotechnology and computer engineering. Not only have these objectives been achieved, but the learning I have acquired during this period has exceeded my initial expectations.

From the scientific point of view, I have had the opportunity to deepen and execute advanced techniques in the field of biotechnology and bioinformatics, allowing me to appreciate their potential in research. In addition, I have acquired skills in handling and interpreting data using advanced statistical methods, and in searching and analyzing information from scientific articles and biological databases. In addition, I have been able to combine the two fields I have always been passionate about, which are programming and new technologies, with experimental science.

This project has allowed me to develop my ability to work independently, encouraging the generation of innovative ideas and improving my ability to communicate them effectively. At the same time, I have reinforced essential skills such as responsibility, initiative, self-confidence and teamwork. These are aspects that I consider fundamental for my personal and professional development.

Finally, I would like to express my deep gratitude to all those who have participated in my training during this time. This whole process has culminated in a deep sense of personal and professional fulfilment that drives me to continue growing and contributing to the advancement of science and technology.

8. ACKNOWLEDGEMENTS

The journey of my double degree in Biotechnology and Computer Engineering, a demanding five-year adventure, has come to an end. These years have been intense, full of hard work and stress, but every moment has been worth it for the wealth of knowledge I have gained. My university experience has been marked by many people who have crossed my path, to all of whom I would like to express my deep gratitude.

I extend my most sincere thanks to my professors, whose teachings not only extended my knowledge, but also nurtured my passion for the scientific field, shaping me as a professional.

My sincere thanks to the entire team of the Biomedical Research Unit, who gave me the opportunity to undertake a project of this magnitude. A special mention goes to Prof. Jorge Jove for allowing me to participate in my curricular internship with his group, and to my mentor and project supervisor, Andrea, for her continuous support and guidance. I also thank Helena and my internship partners, who made the experience enjoyable and fulfilling.

I feel fortunate for the friends I have made during this trip. Their mutual help, encouragement in difficult times and shared moments of joy were crucial to my success. My boyfriend's unwavering support and innovative ideas enriched all my projects, for which I am deeply grateful.

Finally, to my family, who have always believed in me and celebrated my achievements, giving me the strength to pursue my objectives. I would also like to thank my lifelong friends for giving me the fresh air I sometimes needed.

The university years have left an indelible mark on me, a time that I will remember with affection. I am incredibly grateful for the people I have met, the knowledge I have gained and the personal growth I have experienced.

9. REFERENCES

1. Ilango, S., Paital, B., Jayachandran, P., Padma, P. R. & Nirmaladevi, R. Epigenetic alterations in cancer. *Frontiers in Bioscience - Landmark* **25**, 1058–1109 (2020).
2. Mattiuzzi, C. & Lippi, G. Current Cancer Epidemiology. *J Epidemiol Glob Health* **9**, 217 (2019).
3. Sharma, G. N., Dave, R., Sanadya, J., Sharma, P. & Sharma, K. K. Various types and management of breast cancer: An overview. *J Adv Pharm Technol Res* **1**, 109 (2010).
4. Luengo-Fernandez, R., Leal, J., Gray, A. & Sullivan, R. Economic burden of cancer across the European Union: A population-based cost analysis. *Lancet Oncol* **14**, 1165–1174 (2013).
5. Kashyap, D. *et al.* Global Increase in Breast Cancer Incidence: Risk Factors and Preventive Measures. *Biomed Res Int* **2022**, (2022).
6. Colditz, G. A. & Bohlke, K. Priorities for the primary prevention of breast cancer. *CA Cancer J Clin* **64**, 186–194 (2014).
7. Howlader, N. *et al.* US Incidence of Breast Cancer Subtypes Defined by Joint Hormone Receptor and HER2 Status. *JNCI Journal of the National Cancer Institute* **106**, (2014).
8. Liu, L. *et al.* Correlation between family history and characteristics of breast cancer. *Sci Rep* **11**, 6360 (2021).
9. Mahdavi, M. *et al.* Hereditary breast cancer; Genetic penetrance and current status with BRCA. *J Cell Physiol* **234**, 5741–5750 (2019).
10. Satish, S., Moore, J. F., Littlefield, J. M., Bishop, I. J. & Rojas, K. E. Re-Evaluating the Association Between Hormonal Contraception and Breast Cancer Risk. *Breast Cancer (Dove Med Press)* **15**, 227–235 (2023).
11. Fleay, B. *et al.* Exercise medicine considerations for pregnancy-associated breast cancer: Characteristics, prognosis and treatment. *Women's Health* **19**, (2023).
12. Cohen, S. Y., Stoll, C. R., Anandarajah, A., Doering, M. & Colditz, G. A. Modifiable risk factors in women at high risk of breast cancer: a systematic review. *Breast Cancer Res* **25**, 45 (2023).
13. Olsson, H., Bladström, A., Ingvar, C. & Möller, T. R. A population-based cohort study of HRT use and breast cancer in southern Sweden. *Br J Cancer* **85**, 674 (2001).

14. Buja, A., Pierbon, M., Lago, L., Grotto, G. & Baldo, V. Breast Cancer Primary Prevention and Diet: An Umbrella Review. *Int J Environ Res Public Health* **17**, 1–54 (2020).

15. Yang, D. *et al.* Physical activity and breast cancer risk among Asian-American women in Los Angeles. *Cancer* **97**, 2565–2575 (2003).

16. Verloop, J., Rookus, M. A., Van Der Kooy, K. D. & Van Leeuwen, F. E. Physical activity and breast cancer risk in women aged 20-54 years. *J Natl Cancer Inst* **92**, 128–135 (2000).

17. Lathigara, D., Kaushal, D. & Wilson, R. B. Molecular Mechanisms of Western Diet-Induced Obesity and Obesity-Related Carcinogenesis—A Narrative Review. *Metabolites 2023, Vol. 13, Page 675* **13**, 675 (2023).

18. Serrano-Gómez, S. J. *et al.* Molecular Profiles of Breast Cancer in Hispanic/Latina. *Advancing the Science of Cancer in Latinos* 103–109 (2019) doi:10.1007/978-3-030-29286-7_10.

19. Nielsen, T. O. *et al.* Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. *J Natl Cancer Inst* **113**, 808–819 (2021).

20. de Mooij, C. M. *et al.* The influence of receptor expression and clinical subtypes on baseline [18F]FDG uptake in breast cancer: systematic review and meta-analysis. *EJNMMI Res* **13**, (2023).

21. Burguin, A., Diorio, C. & Durocher, F. Breast Cancer Treatments: Updates and New Challenges. *J Pers Med* **11**, 808 (2021).

22. Sørlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* **98**, 10869 (2001).

23. Sebastian, W. *et al.* Genetics, Treatment, and New Technologies of Hormone Receptor-Positive Breast Cancer. *Cancers (Basel)* **15**, (2023).

24. Ahn, H. J., Jung, S. J., Kim, T. H., Oh, M. K. & Yoon, H. K. Differences in Clinical Outcomes between Luminal A and B Type Breast Cancers according to the St. Gallen Consensus 2013. *J Breast Cancer* **18**, 149–159 (2015).

25. ELSTON, C. W. & ELLIS, I. O. pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* **19**, 403–410 (1991).

26. Ellis, I. O., Rakha, E. A., Tse, G. M. & Tan, P. H. An international unified approach to reporting and grading invasive breast cancer. An overview of the International Collaboration on Cancer Reporting (ICCR) initiative. *Histopathology* **82**, 189 (2023).

27. Rakha, E. A., Tse, G. M. & Quinn, C. M. An update on the pathological classification of breast cancer. *Histopathology* **82**, 5 (2023).

28. Liang, J. *et al.* Non-coding RNAs in breast cancer: with a focus on glucose metabolism reprogramming. *Discover. Oncology* **14**, 72 (2023).

29. Wawrzkiewicz-Jałowiecka, A. *et al.* Potassium Channels, Glucose Metabolism and Glycosylation in Cancer Cells. *Int J Mol Sci* **24**, (2023).

30. Vaupel, P., Schmidberger, H. & Mayer, A. The Warburg effect: essential part of metabolic reprogramming and central contributor to cancer progression. *https://doi.org/10.1080/09553002.2019.1589653* **95**, 912–919 (2019).

31. Balkwill, F. & Mantovani, A. Inflammation and cancer: Back to Virchow? *Lancet* **357**, 539–545 (2001).

32. Carmeliet, P. & Jain, R. K. Principles and mechanisms of vessel normalization for cancer and other angiogenic diseases. *Nature Reviews Drug Discovery 2011 10:6* **10**, 417–427 (2011).

33. García Ortega, A. A. Infiltración peritumoral en glioblastomas multiformes, estudio de la angiogénesis tumoral mediante correlación inmuno-histo-radiológica. 1 (2018).

34. Piersma, B., Hayward, M. K. & Weaver, V. M. Fibrosis and cancer: A strained relationship. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1873**, 188356 (2020).

35. Subramani, R., Poudel, S., Smith, K. D., Estrada, A. & Lakshmanaswamy, R. Metabolomics of Breast Cancer: A Review. *Metabolites* **12**, (2022).

36. Gómez-Cebrián, N. *et al.* Multi-Omic Approaches to Breast Cancer Metabolic Phenotyping: Applications in Diagnosis, Prognosis, and the Development of Novel Treatments. *Cancers (Basel)* **13**, (2021).

37. Danzi, F. *et al.* To metabolomics and beyond: a technological portfolio to investigate cancer metabolism. *Signal Transduction and Targeted Therapy 2023 8:1* **8**, 1–22 (2023).

38. Wishart, D. S. *et al.* NMR and Metabolomics-A Roadmap for the Future. *Metabolites* **12**, (2022).

39. Díaz-Beltrán, L. *et al.* Human Plasma Metabolomics for Biomarker Discovery: Targeting the Molecular Subtypes in Breast Cancer. *Cancers (Basel)* **13**, 1–18 (2021).

40. Costanzo, M., Caterino, M. & Ruoppolo, M. Targeted metabolomics. *Metabolomics Perspectives: From Theory to Practical Application* 219–236 (2022) doi:10.1016/B978-0-323-85062-9.00006-4.

41. Klee, S. & Braekling, S. Simultaneous Electron Ionization & Chemical Ionization with the EC-TOF for GC-MS.

42. Kakudji, B. K., Mwila, P. K., Burger, J. R., du Plessis, J. M. & Naidu, K. Breast cancer molecular subtypes and receptor status among women at Potchefstroom Hospital: a cross-sectional study. *Pan Afr Med J* **38**, 1–13 (2021).

43. Park, J. & Gupta, R. S. Review Adenosine kinase and ribokinase-the RK family of proteins. doi:10.1007/s00018-008-8123-1.

44. Ge, T. *et al.* The Role of the Pentose Phosphate Pathway in Diabetes and Cancer. *Front Endocrinol (Lausanne)* **11**, 365 (2020).

45. Nowotarski, S. L., Woster, P. M. & Casero, R. A. Polyamines and cancer: implications for chemotherapy and chemoprevention. **15**, 3 (2013).

46. Duszka, K. Versatile Triad Alliance: Bile Acid, Taurine and Microbiota. *Cells* **11**, (2022).

47. Stary, D. & Bajda, M. Taurine and Creatine Transporters as Potential Drug Targets in Cancer Therapy. *Int J Mol Sci* **24**, (2023).

48. Guo, S., Wang, Y., Zhou, D. & Li, Z. Significantly increased monounsaturated lipids relative to polyunsaturated lipids in six types of cancer microenvironment are observed by mass spectrometry imaging. *Sci Rep* **4**, (2014).

49. Karabencheva-Christova, T. G., Christov, C. Z. & Fields, G. B. Collagenolytic Matrix Metalloproteinase Structure–Function Relationships: Insights From Molecular Dynamics Studies. *Adv Protein Chem Struct Biol* **109**, 1–24 (2017).

50. Martinez, J. & Smith, P. C. The Dynamic Interaction between Extracellular Matrix Remodeling and Breast Tumor Progression. (2021) doi:10.3390/cells10051046.

51. Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci* **41**, 211–218 (2016).

52. Kubik, J., Humeniuk, E., Adamczuk, G., Madej-Czerwonka, B. & Korga-Plewko, A. Targeting Energy Metabolism in Cancer Treatment. *Int J Mol Sci* **23**, (2022).

53. Pavlova, N. N. & Thompson, C. B. THE EMERGING HALLMARKS OF CANCER METABOLISM. *Cell Metab* **23**, 27 (2016).

54. Koundouros, N. & Poulogiannis, G. Reprogramming of fatty acid metabolism in cancer. *Br J Cancer* **122**, 4 (2020).

55. Yu, L., Chen, X., Sun, X., Wang, L. & Chen, S. The Glycolytic Switch in Tumors: How Many Players Are Involved? *J Cancer* **8**, 3430 (2017).

56. Gaillard, H., García-Muse, T. & Aguilera, A. Replication stress and cancer. *Nature Reviews Cancer 2015 15:5* **15**, 276–289 (2015).

57. El Agouza, I. M., Eissa, S. S., El Houseini, M. M., El-Nashar, D. E. & Abd El Hameed, O. M. Taurine: a novel tumor marker for enhanced detection of breast cancer among female patients. *Angiogenesis* **14**, 321–330 (2011).

58. Intlekofer, A. M. *et al.* Hypoxia Induces Production of L-2-Hydroxyglutarate. *Cell Metab* **22**, 304–311 (2015).

59. Pînzariu, O. & Georgescu, C. E. Metabolomics in acromegaly: a systematic review. *https://doi.org/10.1177/10815589231169452* 108155892311694 (2023) doi:10.1177/10815589231169452.

60. Montag, K., Ivanov, R. & Bauer, P. Role of SEC14-like phosphatidylinositol transfer proteins in membrane identity and dynamics. *Front Plant Sci* **14**, (2023).

61. Peyraud, F. *et al.* Targeting Tryptophan Catabolism in Cancer Immunotherapy Era: Challenges and Perspectives. *Front Immunol* **13**, (2022).

62. Patra, K. C. & Hay, N. The pentose phosphate pathway and cancer. *Trends Biochem Sci* **39**, 347–354 (2014).

63. Nong, S. *et al.* Metabolic reprogramming in cancer: Mechanisms and therapeutics. *MedComm (Beijing)* **4**, (2023).

64. Matés, J. M., Segura, J. A., Alonso, F. J. & Márquez, J. Oxidative stress in apoptosis and cancer: an update. *Arch Toxicol* **86**, 1649–1665 (2012).

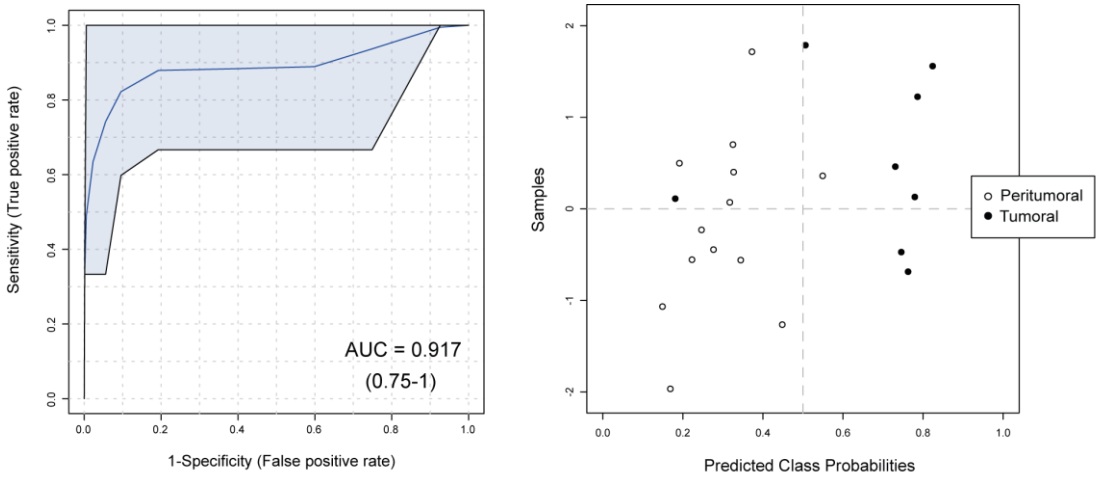
65. Mates, J. M., Segura, J. A., Alonso, F. J. & Marquez, J. Sulphur-containing non enzymatic antioxidants: Therapeutic tools against cancer. *Frontiers in Bioscience - Scholar* **4 S**, 722–748 (2012).

66. Heiden, M. G. V., Cantley, L. C. & Thompson, C. B. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* **324**, 1029 (2009).

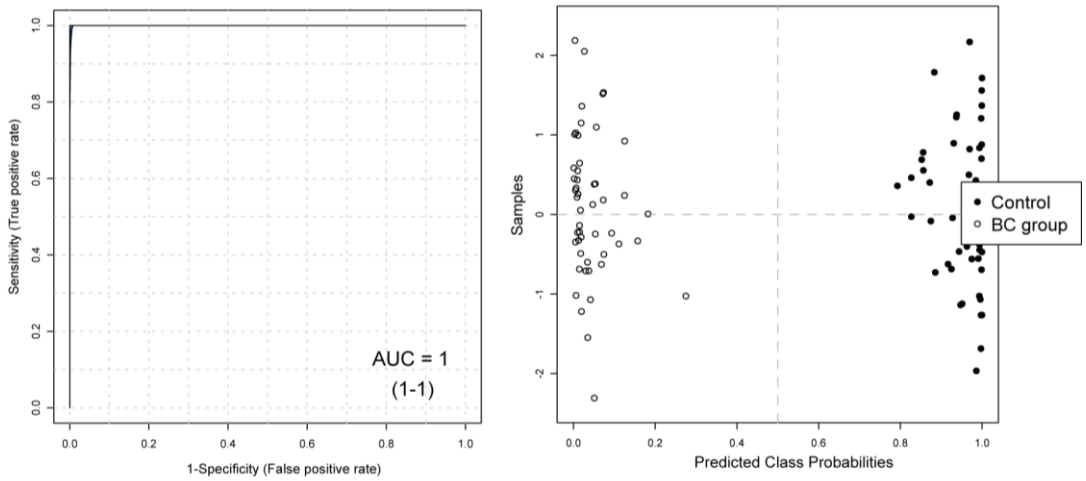
67. Landaas, S. Accumulation of 3-hydroxyisobutyric acid, 2-methyl-3-hydroxybutyric acid and 3-hydroxyisovaleric acid in ketoacidosis. *Clinica Chimica Acta* **64**, 143–154 (1975).

68. Lattermann, R. *et al.* Integrated analysis of glucose, lipid, and urea metabolism in patients with bladder cancer: Impact of tumor stage. *Nutrition* **19**, 589–592 (2003).

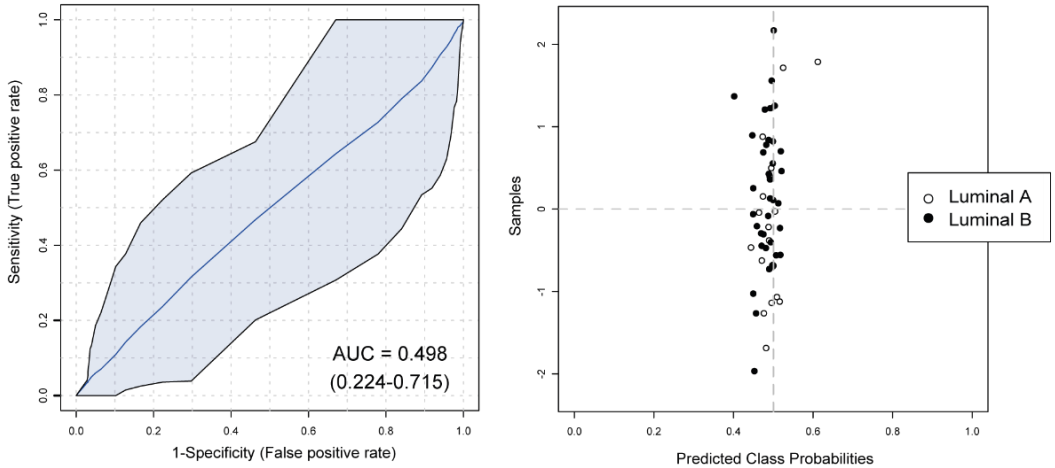
10. APPENDIX



Supplementary Figure 1. ROC Curve and matrix confusion of ribonic acid, taurine and ornithine. The shaded area represents the 95% confidence band.



Supplementary Figure 2. ROC Curve and matrix confusion of xylonic acid, ethylmalonic acid, glycerol, 3-phosphoglyceric acid and maltose.



Supplementary Figure 3. ROC Curve and matrix confusion of phenylalanine and succinic acid. The shaded area represents the 95% confidence band.

Supplementary Table 1. Participant anthropometric characteristics and relevant characteristics for the plasma study between LA and LB BC groups.

	Luminal A (n=18)	**Luminal B (n=39)**	**p-value**
Clinical characteristics			
Age, years, n (%)	53.50 (46.0 - 62.0)	58.00 (52.0 - 68.5)	0.137
Smoking, n (%)	2 (11.1)	5 (12.8)	1
Enolism, n (%)	1 (5.6)	4 (10.3)	0.937
Menopausal status, n (%)			0.137
Premenopausal	8 (44.4)	9 (23.1)	
Perimenopausal	0 (0.0)	4 (10.3)	
Postmenopausal	10 (55.6)	26 (66.7)	
Intake of oral contraceptives, n (%)	7 (38.9)	10 (25.6)	0.481
Parity, n (%)	13 (72.2)	35 (89.7)	0.195
BMI	31.62 (26.1 - 32.8)	27.64 (24.5 - 33.8)	0.608
Diabetes Mellitus, n (%)	1 (5.6)	7 (17.9)	0.400
Hypertension, n (%)	2 (11.1)	13 (33.3)	0.148
Dyslipidaemia, n (%)	5 (27.8)	8 (20.5)	0.789
Pulmonary pathology, n (%)			0.257
SAHS	0 (0.0)	1 (2.6)	
Tuberculosis	1 (5.6)	0 (0.0)	
Others	0 (0.0)	3 (7.7)	
Cardiac pathology, n (%)	2 (1.1)	4 (10.3)	1
Oncologic history, n (%)	0 (0.0)	1 (2.6)	1
Oncologic family history, n (%)	10 (55.6)	24 (61.5)	0.891
Cancer characteristics			
Affected breast, n (%)			0.716
Left	9 (50.0)	23 (59.0)	
Right	8 (44.4)	13 (33.3)	
Bilateral	1 (5.6)	3 (7.7)	
Pathological anatomy of the tumor, n (%)			0.085
Ductal	11 (61.1)	33 (86.8)	
Lobular	5 (27.8)	4 (10.5)	
Others	2 (11.1)	1 (2.6)	
Histological Grade, n (%)			**0.001**
I	10 (55.6)	3 (7.9)	
II	8 (44.4)	33 (86.8)	
III	0 (0.0)	2 (5.3)	
Strogen Receptors	95.0 (90.0 - 99.0)	100.00 (90.0 - 100.0)	0.093
Progesterone Receptors	90.0 (60.0 - 100.0)	70.00 (1.0 - 90.0)	**0.025**
Ki67 antigen tumor in biopsy	10.0 (7.0 - 12.0)	30.00 (22.5 - 39.0)	**<0.001**
Tumor Size, n (%)			0.159
T1	9 (52.9)	11 (28.2)	
T2	8 (47.1)	21 (53.8)	
T3	0 (0.0)	7 (18.0)	
Nodes (TNM system), n (%)			0.072
N0	14 (82.4)	20 (51.3)	
N1	2 (11.8)	17 (43.6)	
N2	1 (5.9)	1 (2.6)	

Normal distributed continuous data is presented as a mean (standard deviation), not normal distributed continuous data is presented as median (interquartile range]; finally, for categorical variables data is presented as frequency percentage (%).

Supplementary Table 2. Participant anthropometric characteristics and relevant characteristics for the tissue study stratified by tissues samples.

	Peritumoral Tissue Samples (n=13)	Tumoral Tissue Samples (n=8)	p-value
Clinical characteristics			
Diagnostic Age	57.0 (49.0 - 71.0)	56.5 (51.3 - 61.0)	0.942
Smoking, n (%)	2 (15.4)	2 (25.0)	1
Enolism, n (%)	3 (23.1)	2 (25.0)	1
Menopausal status, n (%)			0.856
Premenopausal	4 (30.8)	2 (25.0)	
Perimenopausal	2 (15.4)	2 (25.0)	
Postmenopausal	7 (53.8)	4 (50.0)	
Intake of oral contraceptives, n (%)	2 (15.4)	1 (12.5)	1
Parity, n (%)	11 (84.6)	7 (87.5)	1
BMI (median (IQR))	26.7 (25.0 - 32.0)	25.9 (24.8 - 31.3)	0.885
Diabetes Mellitus, n (%)	2 (15.4)	1 (12.5)	1
Hypertension, n (%)	4 (30.8)	2 (25.0)	1
Dyslipidaemia, n (%)	5 (38.5)	2 (25.0)	0.874
Pulmonary pathology, n (%)			0.862
Tuberculosis	1 (7.7)	1 (12.5)	
Cardiac pathology, n (%)	0 (0.0)	0 (0.0)	1
Oncologic history, n (%)	0 (0.0)	0 (0.0)	1
Oncologic family history, n (%)	7 (53.8)	3 (37.5)	0.781
Cancer characteristics			
Affected breast (%)			0.915
Left	8 (61.5)	5 (62.5)	
Right	4 (30.8)	2 (25.0)	
Bilateral	1 (7.7)	1 (12.5)	
Pathological anatomy of the tumor (%)			0.948
Ductal carcinoma	8 (61.5)	4 (50.0)	
Lobular carcinoma	4 (30.8)	3 (37.5)	
Others	1 (7.7)	1 (12.5)	
Histological Grade (%)			0.912
I	3 (23.1)	1 (12.5)	
II	9 (69.2)	6 (75.0)	
III	1 (7.7)	1 (12.5)	
Positive Strogen receptor	97.0 (90.0 - 100.0)	96.0 (93.8 - 98.5)	0.825
Positive Progesterone receptors	70.0 (40.0 - 90.0)	52.5 (12.5 - 75.0)	0.489
Ki67 antigen in tumor biopsy	18.0 (12.0 - 30.0)	24.0 (15.0 - 45.0)	0.488
Tumor molecular classification, n (%)			0.669
Luminal A	4 (30.8)	1 (12.5)	
Luminal B	9 (69.2)	7 (87.5)	
Tumor Size, n (%)			0.855
T1	1 (7.7)	1 (12.5)	
T2	9 (69.2)	4 (50.0)	
T3	3 (23.1)	3 (37.5)	
Nodes (TNM system), n (%)			0.759
N0	7 (53.8)	3 (37.5)	
N1	5 (38.5)	4 (50.0)	
N2	1 (7.7)	1 (12.5)	

Normal distributed continuous data is presented as a mean (standard deviation), not normal distributed continuous data is presented as median (interquartile range); finally, for categorical variables data is presented as frequency percentage (%).

Supplementary Table 3. Patients used in the tissue analysis with all studied samples and their main characteristics.

Peritumoral Tissue Sample	Tumoral Tissue Sample	Basal Sample	Molecular Subtype	Histological Classification	Histological Grade	Menopausal State	Tumor Size
✓		✓	Luminal A	Ductal	II	Premenopausal	T1
✓	✓	✓	Luminal B	Ductal	III	Perimenopausal	T1
✓	✓	✓	Luminal B	Ductal	II	Premenopausal	T2
✓	✓	✓	Luminal B	Ductal	II	Premenopausal	T2
✓	✓	✓	Luminal B	Lobular	II	Postmenopausal	T2
✓	✓	✓	Luminal B	Lobular	II	Perimenopausal	-
✓		✓	Luminal A	Ductal	I	Premenopausal	T2
✓	✓	✓	Luminal A	Lobular	I	Postmenopausal	T2
✓	✓	✓	Luminal B	Ductal	II	Postmenopausal	T3
✓		✓	Luminal B	Ductal	II	Postmenopausal	T1
✓		✓	Luminal B	Ductal	1	Postmenopausal	T2
✓	✓	✓	Luminal B	Other	II	Postmenopausal	T2
✓		✓	Luminal A	Lobular	II	Postmenopausal	T3

Supplementary Table 4. Energy metabolism characteristics of peritumoral and tumoral tissue samples.

	Peritumoral (n=13)	Tumoral (n=8)	p-value adjusted
Carbohydrate metabolism			
Fructose and mannose			
d-Fructose	1×10^{-3} (5×10^{-4} - 2×10^{-3})	4.65×10^{-3} (1.175×10^{-3}, 7.1×10^{-3})	0.218
d-Mannonic acid	3.05×10^{-4} (11.35×10^{-4} - 6.00×10^{-4})	5.43×10^{-4} (1.54×10^{-4} - 1.963×10^{-3})	0.384
Galactose metabolism			
d-Galactitol	3×10^{-5} (2×10^{-5} - 4.5×10^{-5})	7.5×10^{-5} (3.9×10^{-5} - 1.34×10^{-4})	0.131
Glycolysis			
3-Phosphoglyceric acid	7.86×10^{-3} (4.28×10^{-3} - 0.014)	3.67×10^{-3} (1.28×10^{-3} - 0.014)	0.338
Glucose 6-phosphate	1.20×10^{-3} (6.75×10^{-4} - 2.54×10^{-3})	3.28×10^{-3} (9.26×10^{-4} - 0.014)	0.239
Lactic acid	1.68 (1.15 - 2.16)	5.69 (3.12 - 8.68)	**0.039***
Nucleotide sugar			
d-Arabinose	3.00×10^{-5} (1.50×10^{-5} - 4.00×10^{-5})	9.8×10^{-5} (2.4×10^{-5} - 1.89×10^{-4})	0.131
d-Threitol	9.00×10^{-5} (8.00×10^{-5} - 1.30×10^{-4})	1.93×10^{-4} (1.15×10^{-4} - 3.1×10^{-4})	0.216
d-xylose	2.00×10^{-5} (1.50×10^{-5} - 3.50×10^{-5})	4×10^{-5} (2.9×10^{-5} - 1.04×10^{-4})	0.122
Erythronic acid	6.00×10^{-5} (3.00×10^{-5} - 8.00×10^{-5})	4.88×10^{-4} (6.6×10^{-5} - 8.76×10^{-4})	0.075
Threonic acid	9.35×10^{-4} (5.90×10^{-4} - 1.12×10^{-3})	3.67×10^{-3} (9.45×10^{-4} - 8.95×10^{-3})	0.089
xylonic acid	6.50×10^{-5} (5.00×10^{-5} - 1.00×10^{-4})	2.45×10^{-4} (1.11×10^{-4} - 4.13×10^{-4})	**0.027***
Pentose glucoronate interconversion			
d-Arabitol	2.00×10^{-5} (1.00×10^{-5} - 3.00×10^{-5})	3.8×10^{-5} (1.8×10^{-5} - 9.1×10^{-5})	0.720
Galacturonic acid	6.50×10^{-5} (3.00×10^{-5} - 1.05×10^{-4})	1.45×10^{-4} (3.4×10^{-5} - 3.5×10^{-4})	0.384
Myo-Inositol	0.771 (0.6929 - 0.943)	0.94 (0.56 - 1.06)	0.720
Sucrose metabolism			
d-Sucrose	5.00×10^{-5} (3.00×10^{-5} - 6.00×10^{-5})	1.5×10^{-4} (8.3×10^{-5} - 3.15×10^{-4})	0.219
Maltose	2.10×10^{-4} (1.05×10^{-4} - 6.80×10^{-4})	5.78×10^{-4} (2.21×10^{-4} - 2.15×10^{-3})	0.291
TCA			
a-ketoglutaric acid	0.045 (0.036, 0.064)	0.018 (0.016 - 0.021)	**0.032***
Citric acid	0.16 (0.01, 0.54)	0.24 (0.018 - 0.92)	0.768
DL-2-Hydroxyglutaric acid	0.075 (0.071, 0.083)	0.094 (0.069 - 0.17)	0.219
Fumaric acid	0.16 (0.12, 0.207)	0.28 (0.095 - 1.38)	0.327
Glutamine	0.14 (0.11, 0.205)	0.58 (0.29 - 0.81)	0.189
Malic acid	0.092 (0.074, 0.15)	0.22 (0.067 - 0.47)	0.396
Pyruvic acid	0.19 (0.16, 0.25)	0.19 (0.092 - 0.41)	0.961
Succinic acid	0.034 (0.029, 0.051)	0.033 (0.028 - 0.053)	0.538
Glyxolate and decarbocylate			
Glycolic acid	0.049 (0.029, 0.065)	0.039 (0.021 - 0.12)	0.833
Pentose phosphate pathway			
Sedoheptulose	7.45×10^{-4} (2.30×10^{-4} - 1.62×10^{-3})	1.49×10^{-3} (4.21×10^{-4} - 2.69×10^{-3})	0.801
Amino acid metabolism			
Alanine and aspartate			
Alanine	1.62 (1.23 - 2.01)	5.83 (1.015 - 8.76)	0.216
Glutamic acid	1.69 (1.29 - 2.24)	7.32 (1.28 - 9.84)	0.123
Arginine and proline			
2-Hydroxyisobutyric acid	3.14×10^{-3} (2.73×10^{-3} - 3.42×10^{-3})	3.01×10^{-3} (2.68×10^{-3} - 3.84×10^{-3})	0.818
Urea	1.08 (0.545 - 1.51)	2.24 (1.29 - 2.76)	0.219
Proline	1.60 (1.09 - 2.21)	7.037 (1.057 - 11.65)	0.173
4-Hydroxyproline	0.022 (0.016 - 0.057)	0.35 (0.054 - 0.8)	**0.036***
Ornithine	0.47 (0.28 - 0.86)	0.16 (0.109 - 0.19)	**0.017***
Oxoproline	0.14 (0.12 - 0.17)	0.57 (0.107 - 1.063)	0.218
Glyicine and Serine			
Glyceric acid	0.15 (0.075 - 0.25)	0.24 (0.073 - 0.27)	0.923
Glycine	1.33 (1.22 - 1.53)	1.12 (0.87 - 1.87)	0.738
Serine	1.14 (0.81 - 1.58)	3.028 (0.64 - 4.74)	0.384

Threonine	0.5 (0.37 - 0.76)	2.23 (0.3 - 3.7)	0.239
Tyrosin metabolism			
Vanillylmandelic acid	2.50×10^{-5} (1.00×10^{-5} - 3.40×10^{-5})	1.7×10^{-4} (9.3×10^{-5} - 1.91×10^{-4})	0.089
Valine - leucine and isoleucine			
2-HydroxyButyric acid	0.063 (0.031 - 0.083)	0.19 (0.082 - 0.25)	0.123
3-methyl-2-oxobutyric acid	4.89×10^{-4} (2.32×10^{-4} - 6.95×10^{-4})	1.03×10^{-4}) (1.03×10^{-4} - 1.93×10^{-4})	0.089
3-hydroxybutyric acid	0.017 (7.02×10^{-3} - 0.025)	0.056 (0.023 - 0.11)	0.810
2-Hydroxyisovaleric acid	2.06×10^{-3} (1.96×10^{-3} - 4.56×10^{-3})	7.61×10^{-3} (3.94×10^{-3} - 0.011)	0.095
2-keto-3-methylvaleric acid	2.24×10^{-3} (1.93×10^{-3} - 2.65×10^{-3})	2.27×10^{-3} (1.91×10^{-3} - 2.81×10^{-3})	0.961
3-Hydroxyisovaleric acid	2.21×10^{-3} (2.01×10^{-3} - 2.63×10^{-3})	3.14×10^{-3} (2.57×10^{-3} - 4.91×10^{-3})	0.094
Ethylmalonic acid	0.033 (0.025 - 0.035)	0.043 (0.024 - 0.055)	0.656
Isoleucine	0.23 (0.14 - 0.34)	0.94 (0.12 - 1.82)	0.289
Leucine	0.73 (0.47 - 1.04)	2.76 (0.49 - 4.92)	0.215
Valine	0.80 (0.49 - 1.03)	2.3 (0.41 - 4.3)	0.261
Cysteine and methionine metabolism			
Methionine	1.90 (1.86 - 2.01)	2.65 (1.88 - 3.46)	0.134
Tryptophan metabolism			
Indole-3-propanoic acid	9.00×10^{-4} (5.00×10^{-4} - 1.20×10^{-3})	1.55×10^{-3} (5×10^{-4} - 2.3×10^{-3})	0.573
Indolelactic acid	5.00×10^{-5} (3.30×10^{-5} - 7.00×10^{-5})	1.7×10^{-4} (1.4×10^{-5} - 3.4×10^{-4})	0.309
Phenylalanine metabolism			
4-hydroxyPhenyllactic acid	2.35×10^{-3} (1.16×10^{-3} - 2.95×10^{-3})	6.3×10^{-3} (9.33×10^{-4} - 2.29×10^{-2})	0.282
Benzoic acid	2.80×10^{-3} (2.65×10^{-3} - 3.20×10^{-3})	3.14×10^{-3} (2.71×10^{-3} - 3.47×10^{-3})	0.822
Hydrocinnamic acid	4.00×10^{-5} (2.50×10^{-5} - 7.50×10^{-5})	9.3×10^{-5} (6.9×10^{-5} - 1.4×10^{-4})	0.219
Phenylalanine	0.58 (0.38 - 0.95)	2.3 (0.35 - 4.35)	0.223
Lipid metabolism			
Lipids			
Dodecanoic acid	6.49×10^{-3} (4.69×10^{-3} - 8.16×10^{-3})	4.7×10^{-3} (3.52×10^{-3} - 5.83×10^{-3})	0.215
Linoleic acid	0.05032 (0.03157 - 0.08734)	0.055 (0.039 - 0.1)	0.965
Oleic acid	0.383705 (0.247125 - 0.48476)	0.25 (0.18 - 0.41)	0.219
Tetradecanoic acid	0.03236 (0.030965 - 0.03502)	0.025 (0.021 - 0.031)	0.607
Glycolipid metabolism			
Ethanolamine	0.1854 (0.1337 - 0.2471)	1.84 (1.32 - 2.098)	**0.017***
Glycerol	0.2526 (0.197 - 0.2856)	0.43 (0.19 - 0.61)	0.384
Glycerol-1-phosphate	8.88×10^{-3} (4.06×10^{-3} - 0.023)	0.014 (5.49×10^{-3} - 0.053)	0.343
Primary bile acid biosynthesis			
Taurine	0.014 (0.011 - 0.023)	0.21 (0.044 - 0.42)	**0.017***
Metabolism of cofactors and vitamins			
Cofactor biosynthesis			
a-tocopherol	7.01×10^{-3} (3.84×10^{-3} - 0.01134)	0.014 (6.09×10^{-3} - 0.023)	0.961
Nucleotide metabolism			
Purine and pirymidine			
Hypoxanthine	0.045 (0.036 - 0.094)	0.061 (0.029 - 0.39)	0.961
Ribonic acid	2.50×10^{-5} (1.50×10^{-5} - 3.50×10^{-5})	4.4×10^{-4} (1.8×10^{-4} - 6.23×10^{-4})	**0.006***
Uracil	0.085 (0.059 - 0.17)	0.49 (0.055 - 1.77)	0.284
Uric acid	0.028 (0.023 - 0.0401)	0.19 (0.022 - 0.41)	0.342
Xenobiotic biodegradation			
Benzoate degradation			
4-Hydroxybenzoic acid	0.022 (0.016 - 0.057)	0.35 (0.053 - 0.8)	0.219
Energy metabolism			
Oxidative phosphorylation			
Phosphoric acid	1.27 (1.12 - 1.29)	0.8 (0.39 - 1.33)	0.131

Values are provided as median (interquartile range). The units for the metabolites are given in Relative Units (RU). **p< 0.05*** by the Mann-Whitney U-test.

Supplementary Table 5. Energy metabolism characteristics of control group and BC group.

	Control group (n=50)	BC group (n=54)	p-value adjusted
Carbohidrate metabolism			
Fructose and mannose			
d-Fructose	0.015 (0.012-0.02)	0.019 (0.015-0.03)	0.0624
d-Mannonic acid	0.02 (0.016-0.029)	0.025 (0.02-0.029)	0.0624
Galactose metabolism			
d-Galactitol	0.02 (0.016-0.025)	0.017 (0.013-0.022)	**0.0249***
Glycolysis			
3-Phosphoglyceric acid	0.023 (0.016-0.039)	0.074 (0.041-0.158)	**2.84×10^{-9}****
Glucose 6-phosphate	0.005 (0.004-0.007)	0.01 (0.007-0.014)	**1.4×10^{-6}****
Lactic acid	56.04 (42.96-62.539)	62.469 (56.481-73.806)	**0.0029****
Nucleotide sugar			
d-Arabinose	0.017 (0.014-0.023)	0.021 (0.016-0.027)	0.1534
d-Threitol	0.019 (0.016-0.021)	0.025 (0.02-0.03)	**2.41×10^{-6}****
d-xylose	0.007 (0.005-0.01)	0.009 (0.006-0.015)	0.1668
Erythronic acid	0.02 (0.015-0.027)	0.028 (0.023-0.035)	**8.77×10^{-5}*****
Threonic acid	0.243 (0.199-0.293)	0.334 (0.269-0.426)	**3.51×10^{-4}****
xylonic acid	0.046 (0.038-0.054)	0.015 (0.012-0.02)	**1.56×10^{-15}****
Pentose glucoronate interconversion			
d-Arabitol	0.008 (0.007-0.01)	0.011 (0.009-0.014)	**0.0023****
Galacturonic acid	0.018 (0.015-0.021)	0.017 (0.014-0.021)	0.8121
Myo-Inositol	0.431 (0.367-0.519)	0.603 (0.459-0.683)	**2.49×10^{-4}*****
Sucrose metabolism			
d-Sucrose	0.003 (0.002-0.005)	0.007 (0.003-0.01)	**0.0023****
Maltose	0.006 (0.004-0.009)	0.021 (0.01-0.044)	**7.51×10^{-9}****
TCA			
a-ketoglutaric acid	0.139 (0.108-0.164)	0.143 (0.105-0.213)	0.9676
Citric acid	16.944 (13.965-19.619)	18.979 (15.176-27.806)	0.0787
DL-2-Hydroxyglutaric acid	0.04 (0.034-0.05)	0.041 (0.033-0.054)	0.9384
Fumaric acid	0.126 (0.106-0.146)	0.121 (0.099-0.161)	0.7679
Glutamine	35.475 (28.219-45.124)	34.025 (26.301-39.253)	0.1668
Malic acid	0.056 (0.046-0.068)	0.068 (0.053-0.094)	**0.0314***
Pyruvic acid	4.755 (3.287-6.172)	5.187 (3.183-6.967)	0.5648
Succinic acid	0.055 (0.048-0.065)	0.072 (0.056-0.105)	**0.0048****
Glyxolate and decarbocylate			
Glycolic acid	0.216 (0.189-0.243)	0.173 (0.145-0.195)	**4.54×10^{-6}****
Pentose phosphate pathway			
Sedoheptulose	0.015 (0.011-0.02)	0.014 (0.01-0.018)	0.3299
Amino acid metabolism			
Alanine and aspartate			
Alanine	20.574 (17.779-24.173)	24.797 (22.066-28.673)	**1.71×10^{-4}*****
Glutamic acid	3.656 (3.126-4.375)	5.71 (3.711-8.41)	**1.13×10-4****
Arginine and proline			
2-Hydroxyisobutyric acid	1.114 (0.904-1.334)	1.37 (1.014-1.74)	0.0729
Urea	31.449 (22.922-36.931)	34.539 (28.424-41.424)	0.1032
Proline	26.831 (21.587-33.734)	29.791 (24.867-38.652)	0.2685
4-Hydroxyproline	1.369 (1.084-1.892)	1.252 (1.02-1.795)	0.5866
Ornithine	13.051 (10.005-16.709)	13.197 (9.807-15.128)	0.8241
O×oproline	25.688 (21.665-29.701)	24.855 (21.028-28.635)	0.5866
Glyicine and Serine			
Glyceric acid	0.621 (0.475-0.797)	0.491 (0.381-0.676)	**0.0249***
Glycine	6.492 (5.985-7.034)	6.579 (6.104-7.148)	0.5421
Serine	8.841 (7.365-9.528)	7.624 (6.59-9.34)	0.1937
Threonine	8.737 (7.139-10.075)	7.939 (6.863-9.451)	0.181

Tyrosin metabolism			
Vanillylmandelic acid	0.006 (0.005-0.008)	0.008 (0.006-0.011)	**0.0138***
Valine - leucine and isoleucine			
2-HydroxyButyric acid	1.114 (0.904-1.334)	1.37 (1.014-1.74)	0.0729
3-methyl-2-oxobutyric acid	0.596 (0.45-0.701)	0.543 (0.421-0.756)	0.551
3-hydroxybutyric acid)	8.622 (5.27-20.664)	7.463 (4.369-17.414)	0.6205
2-Hydroxyisovaleric acid	1.114 (0.904-1.334)	1.37 (1.014-1.74)	0.0729
2-keto-3-methylvaleric acid	1.253 (0.947-1.513)	1.106 (0.889-1.326)	0.0818
3-Hydroxyisovaleric acid	0.891 (0.644-1.101)	0.848 (0.715-1.226)	0.8009
Ethylmalonic acid	5.269 (4.418-6.308)	2.135 (1.741-2.912)	**2.02×10^{-14}****
Isoleucine	4.465 (3.888-5.138)	4.632 (3.817-5.66)	0.7164
Leucine	8.896 (8.067-9.784)	8.923 (7.713-10.942)	0.8009
Valine	16.438 (15.402-18.283)	17.075 (14.81-20.331)	0.6929
Cysteine and methionine metabolism			
Methionine	3.643 (3.484-3.816)	3.623 (3.398-3.89)	0.6285
Tryptophan metabolism			
Indole-3-propanoic acid	0.122 (0.081-0.192)	0.155 (0.101-0.236)	0.176
Indolelactic acid	0.137 (0.107-0.168)	0.15 (0.128-0.197)	0.1534
Phenylalanine metabolism			
4-hydroxyPhenyllactic acid	0.147 (0.11-0.195)	0.174 (0.131-0.232)	0.1668
Benzoic acid	0.459 (0.332-0.644)	0.563 (0.451-0.683)	0.1394
Hydrocinnamic acid	0.068 (0.035-0.101)	0.061 (0.037-0.108)	0.8156
Phenylalanine	6.211 (5.884-6.678)	6.472 (5.557-7.532)	0.5866
Lipid metabolism			
Lipids			
Dodecanoic acid	0.497 (0.381-0.733)	0.548 (0.429-0.74)	0.3577
Linoleic acid	2.845 (1.548-4.194)	2.374 (1.326-3.921)	0.6285
Oleic acid	14.627 (7.916-25.588)	16.748 (10.582-22.653)	0.7679
Tetradecanoic acid	1.589 (1.172-2.095)	1.291 (0.906-2.183)	0.265
Glycolipid metabolism			
Ethanolamine	0.228 (0.2-0.27)	0.231 (0.196-0.276)	0.9676
Glycerol	1.593 (1.253-2.073)	0.743 (0.561-0.945)	**$2.02x10^{-14}$****
Glycerol-1-phosphate	0.158 (0.121-0.197)	0.177 (0.133-0.289)	0.0818
Primary bile acid biosynthesis			
Taurine	0.494 (0.403-0.65)	0.892 (0.672-1.26)	**5.33×10^{-8}****
Metabolism of cofactors and vitamins			
Cofactor biosynthesis			
a-tocopherol	0.237 (0.144-0.408)	0.221 (0.166-0.319)	0.7164
Nucleotide metabolism			
Purine and pirymidine			
Hypoxanthine	0.18 (0.137-0.269)	0.503 (0.286-1.499)	**2.09×10^{-8}****
Ribonic acid	0.003 (0.002-0.004)	0.004 (0.003-0.005)	**0.0443***
Uracil	0.018 (0.014-0.024)	0.022 (0.016-0.027)	0.1394
Uric acid	9.742 (6.279-15.21)	11.345 (6.704-20.288)	0.3464
Xenobiotic biodegradation			
Benzoate degradation			
4-Hydroxybenzoic acid	0.123 (0.12-0.126)	0.122 (0.115-0.127)	0.6285
Energy metabolism			
Oxidative phosphorylation			
Phosphoric acid	158.699 (67.025-208.949)	194.936 (158.452-237.615)	**0.0170***

Values are provided as median (interquartile range). The units for the metabolites are given in Relative Units (RU). **p< 0.05** by the Mann-Whitney U-test. The p-values have been adjusted using the *Benjamini-Hochberg correction* (FDR). The levels of significance are indicated with asterisks: ***** 0.01 ≤ Adj.p-value < 0.05; ****** 0.001 ≤ Adj.p-value < 0.01; *******: 0.0001 ≤ Adj.p-value < 0.001; ********: Adj.p-value < 0.0001

Supplementary Table 6. Energy metabolism characteristics of LA and LB BC groups.

	Luminal A (n=17)	**Luminal B** (n=37)	**p-value**	**p- value ANCOVA**
Carbohidrate metabolism				
Fructose and mannose				
d-Fructose	0.016 (0.015-0.027)	0.019 (0.015-0.031)	0.539	0.137
d-Mannonic acid	0.025 (0.021-0.037)	0.025 (0.02-0.029)	0.696	0.657
Galactose metabolism				
d-Galactitol	0.016 (0.013-0.022)	0.017 (0.014-0.022)	0.533	0.371
Glycolysis				
3-Phosphoglyceric acid	0.048 (0.03-0.112)	0.104 (0.043-0.176)	0.071	0.179
Glucose 6-phosphate	0.009 (0.006-0.012)	0.011 (0.007-0.017)	0.248	0.342
Lactic acid	61.147 (57.075-71.039)	62.512 (56.446-79.62)	0.740	0.525
Nucleotide sugar				
d-Arabinose	0.019 (0.016-0.025)	0.022 (0.017-0.03)	0.361	0.483
d-Threitol	0.022 (0.019-0.031)	0.026 (0.02-0.03)	0.306	0.157
d-xylose	0.01 (0.005-0.017)	0.009 (0.007-0.014)	0.759	0.635
Erythronic acid	0.026 (0.023-0.035)	0.028 (0.023-0.036)	0.867	0.746
Threonic acid	0.334 (0.218-0.367)	0.34 (0.285-0.47)	0.148	0.063
xylonic acid	0.015 (0.012-0.02)	0.015 (0.012-0.02)	0.615	0.967
Pentose glucoronate interconversion				
d-Arabitol	0.012 (0.009-0.014)	0.011 (0.009-0.013)	0.744	0.988
Galacturonic acid	0.017 (0.015-0.019)	0.018 (0.014-0.021)	0.787	0.964
Myo-Inositol	0.569 (0.373-0.642)	0.625 (0.464-0.77)	0.128	0.058
Sucrose metabolism				
d-Sucrose	0.007 (0.003-0.008)	0.007 (0.003-0.01)	0.816	0.710
Maltose	0.012 (0.009-0.023)	0.025 (0.012-0.055)	0.186	0.775
TCA				
a-ketoglutaric acid	0.134 (0.105-0.164)	0.146 (0.092-0.226)	0.737	0.471
Citric acid	17.864 (15.684-22.238)	20.901 (15.084-32.246)	0.438	0.212
DL-2-Hydroxyglutaric acid	0.036 (0.03-0.044)	0.044 (0.034-0.063)	0.094	**0.036***
Fumaric acid	0.113 (0.096-0.138)	0.125 (0.099-0.164)	0.506	0.380
Glutamine	34.672 (25.021-40.655)	33.988 (26.662-38.985)	0.726	0.412
Malic acid	0.068 (0.051-0.085)	0.07 (0.057-0.095)	0.347	0.199
Pyruvic acid	5.136 (3.421-6.687)	5.236 (2.846-7.134)	0.811	0.845
Succinic acid	0.062 (0.053-0.073)	0.092 (0.058-0.116)	**0.038***	**0.031***
Glyxolate and decarbocylate				
Glycolic acid	0.16 (0.144-0.197)	0.176 (0.146-0.189)	0.794	0.642
Pentose phosphate pathway				
Sedoheptulose	0.011 (0.009-0.015)	0.014 (0.011-0.018)	0.241	0.943
Amino acid metabolism				
Alanine and aspartate				
Alanine	26.125 (21.262-30.71)	24.624 (22.564-28.476)	0.985	0.977
Glutamic acid	4.709 (3.43-8.541)	5.994 (3.989-8.319)	0.365	0.613
Arginine and proline				
2-Hydroxyisobutyric acid	1.225 (0.846-1.6)	1.421 (1.116-1.978)	0.068	**0.043***
Urea	31.62 (26.49-35.235)	36.964 (30.472-42.743)	0.077	0.154
Proline	35.418 (25.234-43.247)	28.606 (24.818-36.867)	0.355	0.217
4-Hydroxyproline	1.217 (1.011-1.753)	1.272 (1.047-1.887)	0.671	0.316
Ornithine	12.424 (9.636-13.994)	14.033 (10.105-16.256)	0.132	0.181
Oxoproline	23.345 (19.639-28.061)	24.912 (21.185-28.86)	0.658	0.420
Glyicine and Serine				
Glyceric acid	0.499 (0.35-0.584)	0.483 (0.383-0.836)	0.336	0.152
Glycine	6.803 (6.059-7.205)	6.554 (6.116-7.031)	0.740	0.699
Serine	8.004 (6.851-9.296)	7.43 (6.586-9.392)	0.839	0.827

Threonine	8.389 (6.892-9.186)	7.921 (6.853-9.54)	0.839	0.578
Tyrosin metabolism				
Vanillylmandelic acid	0.008 (0.006-0.012)	0.008 (0.007-0.01)	0.933	0.624
Valine - leucine and isoleucine				
2-HydroxyButyric acid	3.784 (3.167-4.278)	4.046 (2.798-6.498)	0.355	0.107
3-methyl-2-oxobutyric acid	0.534 (0.422-0.671)	0.548 (0.421-0.833)	0.712	0.580
3-hydroxybutyric acid	7.9 (4.968-10.629)	6.606 (3.895-18.43)	0.811	0.340
2-Hydroxyisovaleric acid	1.225 (0.846-1.6)	1.421 (1.116-1.978)	0.068	0.043
2-keto-3-methylvaleric acid	1.005 (0.894-1.223)	1.139 (0.878-1.357)	0.309	0.204
3-Hydroxyisovaleric acid	0.724 (0.65-0.825)	0.945 (0.75-1.27)	0.052	0.176
Ethylmalonic acid	2.367 (1.921-2.635)	2.092 (1.63-3.227)	0.956	0.345
Isoleucine	4.461 (3.803-5.149)	4.859 (3.859-6.032)	0.214	0.230
Leucine	8.602 (7.699-10.271)	9.029 (7.754-11.771)	0.259	0.207
Valine	16.574 (14.652-18.918)	17.946 (15.285-22.69)	0.229	0.183
Cysteine and methionine metabolism				
Methionine	3.569 (3.461-3.854)	3.704 (3.398-3.938)	0.427	0.431
Tryptophan metabolism				
Indole-3-propanoic acid	0.155 (0.114-0.255)	0.161 (0.095-0.233)	0.503	0.354
Indolelactic acid	0.137 (0.122-0.159)	0.158 (0.132-0.205)	0.103	0.091
Phenylalanine metabolism				
4-hydroxyPhenyllactic acid	0.15 (0.12-0.185)	0.193 (0.143-0.242)	0.065	0.099
Benzoic acid	0.554 (0.47-0.647)	0.575 (0.447-0.697)	0.768	0.287
Hydrocinnamic acid	0.054 (0.037-0.099)	0.067 (0.038-0.11)	0.675	0.846
Phenylalanine	5.788 (5.351-6.909)	6.576 (5.727-7.714)	**0.050***	0.077
Lipid metabolism				
Lipids				
Dodecanoic acid	0.547 (0.456-0.681)	0.549 (0.428-0.75)	0.985	0.683
Linoleic acid	2.842 (2.202-3.834)	2.184 (1.137-4.34)	0.438	0.627
Oleic acid	15.352 (13.34-20.57)	16.877 (8.487-24.999)	0.912	0.344
Tetradecanoic acid	1.293 (1.036-1.875)	1.289 (0.768-2.343)	0.956	0.629
Glycolipid metabolism				
Ethanolamine	0.221 (0.195-0.263)	0.236 (0.196-0.29)	0.375	0.126
Glycerol	0.835 (0.692-0.96)	0.686 (0.551-0.926)	0.201	0.412
Glycerol-1-phosphate	0.18 (0.133-0.274)	0.174 (0.134-0.309)	0.310	0.177
Primary bile acid biosynthesis				
Taurine	0.746 (0.624-1.005)	0.901 (0.742-1.528)	0.188	0.831
Metabolism of cofactors and vitamins				
Cofactor biosynthesis				
a-tocopherol	0.233 (0.135-0.315)	0.219 (0.166-0.338)	0.956	0.931
Nucleotide metabolism				
Purine and pyrimidine				
Hypoxanthine	0.426 (0.295-0.948)	0.513 (0.283-1.607)	0.605	0.229
Ribonic acid	0.004 (0.003-0.006)	0.004 (0.003-0.005)	0.867	0.669
Uracil	0.022 (0.017-0.025)	0.023 (0.015-0.031)	0.662	0.245
Uric acid	9.696 (5.864-16.509)	12.66 (6.82-21.691)	0.292	0.132
Xenobiotic biodegradation				
Benzoate degradation				
4-Hydro×ybenzoic acid	0.124 (0.113-0.127)	0.122 (0.117-0.126)	0.703	0.918
Energy metabolism				
Oxidative phosphorylation				
Phosphoric acid	197.36 (167.41-227.94)	192.50 (150.82-241.44)	0.868	0.642

Values are provided as median (interquartile range). The units for the metabolites are given in Relative Units (RU). **p < 0.05*** by the Mann-Whitney U-test.

Supplementary Table 7. Logistic regression analysis for energy metabolism characteristics of LA and LB BC groups.

	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 1 p-value	Model 2 p-value
Metabolites				
Fructose and mannose				
d-Fructose	0.83 (0.47, 1.49)	0.53 (0.23, 1.08)	0.521	0.098
d-Mannonic acid	0.92 (0.52, 1.69)	0.69 (0.33, 1.39)	0.760	0.296
Galactose metabolism				
d-Galactitol	1.15 (0.64, 2.11)	0.73 (0.34, 1.49)	0.628	0.386
Glycolysis				
3-Phosphoglyceric acid	1.85 (1.01, 3.63)	1.87 (0.88, 4.53)	0.056	0.125
Glucose 6-phosphate	1.49 (0.82, 2.96)	1.92 (0.89, 4.87)	0.216	0.121
Lactic acid	1.12 (0.63, 2.07)	1.14 (0.5, 2.73)	0.692	0.767
Nucleotide sugar				
d-Arabinose	1 (0.57, 1.89)	0.75 (0.36, 1.53)	1.000	0.403
d-Threitol	0.85 (0.44, 1.57)	0.63 (0.14, 1.29)	0.560	0.349
d-xylose	1.07 (0.6, 1.96)	1.15 (0.54, 2.6)	0.807	0.718
Erythronic acid	1.07 (0.59, 1.91)	0.77 (0.33, 1.62)	0.827	0.503
Threonic acid	1.67 (0.92, 3.25)	1.17 (0.48, 2.97)	0.109	0.724
xylonic acid	0.98 (0.55, 1.77)	0.63 (0.29, 1.3)	0.949	0.219
Pentose glucoronate interconversion				
d-Arabitol	0.99 (0.55, 1.78)	0.58 (0.24, 1.26)	0.973	0.190
Galacturonic acid	0.98 (0.55, 1.77)	0.65 (0.3, 1.29)	0.947	0.228
Myo-Inositol	1.82 (0.99, 3.62)	1.4 (0.65, 3.25)	0.064	0.399
Sucrose metabolism				
d-Sucrose	0.91 (0.5, 1.62)	0.44 (0.17, 1)	0.737	0.066
Maltose	1.55 (0.86, 3.04)	1.86 (0.82, 5.02)	0.167	0.170
TCA				
a-ketoglutaric acid	1.2 (0.67, 2.25)	0.9 (0.41, 2.06)	0.544	0.794
Citric acid	1.28 (0.72, 2.38)	0.9 (0.36, 2.18)	0.403	0.816
DL-2-Hydroxyglutaric acid	2.05 (1.09, 4.31)	1.82 (0.83, 4.61)	**0.038***	0.159
Fumaric acid	1.18 (0.66, 2.13)	1.21 (0.58, 2.63)	0.572	0.617
Glutamine	1.33 (0.74, 2.46)	1.35 (0.65, 2.96)	0.345	0.425
Malic acid	1.36 (0.76, 2.52)	1.27 (0.59, 2.87)	0.308	0.549
Pyruvic acid	0.86 (0.47, 1.54)	0.79 (0.36, 1.61)	0.625	0.522
Succinic acid	1.97 (1.06, 3.98)	2.64 (1.07, 8.92)	**0.042***	0.063
Glyxolate and decarbocylate				
Glycolic acid	1.14 (0.64, 2.05)	1.19 (0.62, 2.44)	0.657	0.608
Pentose phosphate pathway				
Sedoheptulose	1.11 (0.62, 1.99)	0.98 (0.44, 2.08)	0.711	0.957
Amino acid metabolism				
Alanine and aspartate				
Alanine	0.98 (0.54, 1.75)	0.78 (0.37, 1.62)	0.935	0.513
Glutamic acid	1.36 (0.76, 2.57)	1.21 (0.59, 2.64)	0.324	0.602
Arginine and proline				
2-Hydroxyisobutyric acid	1.45 (0.81, 2.71)	1.86 (0.84, 4.91)	0.225	0.159
Urea	1.44 (0.79, 2.8)	1.63 (0.7, 4.41)	0.253	0.290
Proline	0.79 (0.42, 1.41)	0.53 (0.23, 1.09)	0.444	0.099
4-Hydroxyproline	1.26 (0.71, 2.41)	1.22 (0.62, 2.6)	0.450	0.584
Ornithine	1.36 (0.76, 2.49)	1.22 (0.6, 2.67)	0.301	0.591
Oxoproline	1.28 (0.72, 2.31)	1.85 (0.81, 4.58)	0.400	0.155
Glyicine and Serine				
Glyceric acid	1.41 (0.78, 2.68)	1.46 (0.68, 3.52)	0.272	0.361
Glycine	0.93 (0.51, 1.65)	1.03 (0.54, 1.98)	0.794	0.917
Serine	0.97 (0.54, 1.73)	1.08 (0.56, 2.12)	0.931	0.825

Threonine	0.86 (0.47, 1.53)	0.91 (0.43, 1.84)	0.603	0.803
Tyrosin metabolism				
Vanillylmandelic acid	1.24 (0.7, 2.27)	1.04 (0.5, 2.21)	0.468	0.910
Valine - leucine and isoleucine				
2-HydroxyButyric acid	1.45 (0.81, 2.71)	1.86 (0.84, 4.91)	0.225	0.159
3-methyl-2-oxobutyric acid	1.13 (0.63, 2.04)	1.07 (0.49, 2.48)	0.679	0.866
3-hydroxybutyric acid	1.23 (0.69, 2.34)	1.86 (0.87, 4.77)	0.500	0.138
2-Hydroxyisovaleric acid	1.51 (0.83, 2.98)	1.61 (0.74, 3.84)	0.194	0.246
2-keto-3-methylvaleric acid	1.31 (0.73, 2.46)	0.9 (0.43, 1.93)	0.375	0.779
3-Hydroxyisovaleric acid	1.64 (0.9, 3.28)	1.48 (0.72, 3.35)	0.128	0.305
Ethylmalonic acid	1.23 (0.69, 2.23)	1.72 (0.8, 4.24)	0.485	0.188
Isoleucine	1.4 (0.78, 2.67)	1.09 (0.51, 2.4)	0.271	0.824
Leucine	1.41 (0.79, 2.64)	1.25 (0.58, 2.83)	0.261	0.569
Valine	1.48 (0.82, 2.83)	1.3 (0.61, 3)	0.207	0.507
Cysteine and methionine metabolism				
Methionine	1.26 (0.71, 2.4)	1.23 (0.58, 2.81)	0.448	0.594
Tryptophan metabolism				
Indole-3-propanoic acid	0.73 (0.39, 1.3)	0.62 (0.26, 1.34)	0.293	0.235
Indolelactic acid	1.91 (1.01, 4.22)	2.1 (0.87, 6.95)	0.069	0.147
Phenylalanine metabolism				
4-hydroxyPhenyllactic acid	1.56 (0.87, 2.93)	1.28 (0.62, 2.79)	0.144	0.516
Benzoic acid	1.21 (0.68, 2.31)	1.43 (0.68, 3.41)	0.524	0.378
Hydrocinnamic acid	1.11 (0.62, 2)	1.26 (0.62, 2.63)	0.713	0.528
Phenylalanine	1.77 (0.97, 3.52)	1.25 (0.57, 2.88)	0.078	0.587
Lipid metabolism				
Lipids				
Dodecanoic acid	1.06 (0.6, 1.99)	1.17 (0.55, 2.74)	0.847	0.693
Linoleic acid	0.94 (0.52, 1.67)	1.25 (0.61, 2.69)	0.832	0.542
Oleic acid	1.14 (0.63, 2.01)	1.68 (0.8, 4.03)	0.661	0.194
Tetradecanoic acid	1.06 (0.6, 1.94)	1.28 (0.62, 2.91)	0.833	0.515
Glycolipid metabolism				
Ethanolamine	1.53 (0.85, 2.92)	1.44 (0.63, 3.56)	0.168	0.398
Glycerol	0.84 (0.45, 1.5)	1.18 (0.59, 2.51)	0.565	0.640
Glycerol-1-phosphate	1.47 (0.82, 2.81)	1.43 (0.66, 3.43)	0.216	0.387
Primary bile acid biosynthesis				
Taurine	1.53 (0.83, 3.13)	2.06 (1, 4.86)	0.200	0.066
Metabolism of cofactors and vitamins				
Cofactor biosynthesis				
a-tocopherol	1 (0.56, 1.8)	1 (0.44, 2.25)	0.988	0.997
Nucleotide metabolism				
Purine and pyrimidine				
Hypoxanthine	1.19 (0.67, 2.2)	1.59 (0.73, 3.63)	0.554	0.244
Ribonic acid	0.77 (0.41, 1.38)	0.48 (0.18, 1.04)	0.399	0.091
Uracil	1.27 (0.71, 2.36)	1.79 (0.89, 4.03)	0.432	0.119
Uric acid	1.57 (0.86, 3.07)	1.37 (0.6, 3.29)	0.156	0.455
Xenobiotic biodegradation				
Benzoate degradation				
4-Hydroxybenzoic acid	1.02 (0.56, 1.8)	0.84 (0.41, 1.64)	0.941	0.603
Energy metabolism				
Oxidative phosphorylation				
Phosphoric acid	1.02 (0.56, 1.8)	1.21 (0.53, 2.93)	0.937	0.652

Values are provided as median (interquartile range). The units for the metabolites are given in Relative Units (RU). Model 1 is unadjusted. Model 2 is adjusted with covariates (menopausal state, tumor grade, tumor size and histological classification). OR indicates odds ratio.

***p-value < 0.05**. None of them remained significant after Bonferroni correction for multiple comparisons.

Supplementary Table 8. Significant metabolites for the net regression analysis.

Metabolites	β
DL-2-Hydroxyglutaric acid	0.0759
Succinic acid	0.0515
Indolelactic acid	0.0228
Myo-Inositol	0.0146
3-Phosphoglyceric acid	4.42×10^{-4}