

The manuscript is devoted to untargeted plasma metabolic profiling of patients with breast cancer. In general manuscript is interesting and provides some new metabolic features on plasma of breast cancer patients. However, some questions should be clarified to make the manuscript more understandable.

- 1) The title and abstract (line 27) contain misleading information about “Liquid Biopsy”, however, when reading Materials and Methods we see only plasma as biological sample (lines 252-253). Please, change this mistake and use Plasma both in the title and abstract.

Response 1: Thank you for your observation. Following your suggestion, we have replaced the term “liquid biopsy” by “plasma samples” in the abstract to avoid any confusion in this section. Is it worth noting that, generally, liquid biopsy refers to the evaluation and detection of any type of biomarkers, including metabolites, in a biological fluid [1–3]; taking this into account, we decided to leave the term “liquid biopsy” in the title and also include the specific type of liquid biopsy used along our research: plasma samples.

References response 1 – Reviewer 1:

1. Honoré, N.; Galot, R.; van Marcke, C.; Limaye, N.; Machiels, J.P. Liquid Biopsy to Detect Minimal Residual Disease: Methodology and Impact. *Cancers (Basel)*. **2021**, *13*, 1–25, doi:10.3390/cancers13215364.
2. Lone, S.N.; Nisar, S.; Masoodi, T.; Singh, M.; Rizwan, A.; Hashem, S.; El-Rifai, W.; Bedognetti, D.; Batra, S.K.; Haris, M.; et al. Liquid Biopsy: A Step Closer to Transform Diagnosis, Prognosis and Future of Cancer Treatments. *Mol. Cancer* **2022**, *21*, 1–22, doi:10.1186/s12943-022-01543-7.
3. Li, S.; Xin, K.; Pan, S.; Wang, Y.; Zheng, J.; Li, Z.; Liu, X.; Liu, B.; Xu, Z.; Chen, X. Blood-Based Liquid Biopsy: Insights into Early Detection, Prediction, and Treatment Monitoring of Bladder Cancer. *Cell. Mol. Biol. Lett.* **2023**, *28*, doi:10.1186/s11658-023-00442-z.

- 2) Lines 21-22 contain information that “Breast cancer (BC) is the first cause of death by tumour pathology in women worldwide”, then lines 38-39 contain some other information “Breast cancer (BC) ranks as the second most common cause of cancer-related deaths in women overall”. So, the first or the second?

Response 2: Thank you for pointing this out. We have corrected this information within the abstract (lines 22-23): “*Breast cancer (BC) is the most common tumour pathology in women and the second deathliest cancer worldwide...*”

Also, in the introduction (lines 40-41): “*In 2023, BC was the most frequently diagnosed cancer in women accounted for 31% worldwide [4,5] and the second in Spain for both sexes combined.*”

- 3) Figure 1. There are several questions to this figure. Are these chromatograms the sum of all healthy controls vs the sum of all breast cancer patients? Presumably not. If so, how did the authors decide which of the chromatograms was the most representative? In general, these chromatograms do not carry any valuable semantic meaning. If the authors want to discuss the retention time of some chemical groups of compounds, they should replace this figure

into Suppl. and discuss some features in general, or add some illustrative information on the figure to show where the compounds of interest were eluted.

Response 3: Thank you for your questions. In this regard, and also in response to the question 2 of Reviewer 2, we have replaced the TIC of Figure 1 by the TIC of the QCs in comparison with the gradient elution graphic. Following your suggestions, we thought this new figure may help to understand the subsequent discussion of the results and we kindly ask for your compliance.

- 4) Figure 2. My opinion is that Fig.2 b is the most important figure here, because it is clear and answer the question how well groups of patients and controls separate from each other. The meanings of other figures are unclear. Fig 2 a requires some more information than just lines 117-118. Figures 2 c-f are completely unclear. Figure 2 b has already answered the question on clustering patients from controls. The next interesting feature could be related with how four groups of patients could be separated from each other but not from controls. So, figures 2 c-f are unnecessary.

Response 4: Thank you for your kind suggestions. We have added some more information about Figure 2a for clarification: *"In this case, after the contamination filtering and the removal of features with an unacceptable variability, the close clustering of the QCs indicates that the natural behaviour of the samples and relevant separation between the groups of study are due to biological factors (Figure 2a)".*

We completely agree with you that it would be very interesting to see how the four groups of patients are separated from each other. Thus, we have performed the comparison between the four BC molecular subtypes where a natural separation can be inferred (Figure S2). As a result, you can observe the PLS-DA that shows how the four groups of patients could be separated from each other but not from the controls. Specifically, Luminal subtypes (light green circles – LA and dark green circles – LB) were clustered in the middle whereas the TN phenotype (red circles) and the HER2 positive patients (blue circles) appeared wide-spread in opposite directions.

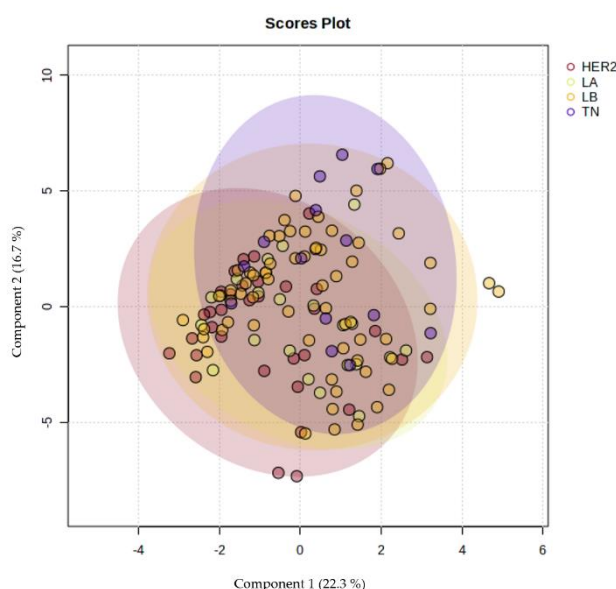


Figure S2. The PLS-DA score plots from the four molecular subtypes shows a clustering of the Luminal samples in the middle while the TN and HER2 samples are wide-spread in opposite directions.

Interestingly, we could observe that HER2 positive samples had a slight tendency to the Luminal B phenotype. We could expect his behaviour since in our stratification, the HER2 positive patients were not clinically classified as “pure” but some of them were also positive for the expression of hormone receptors (luminal - HER2 positive).

Finally, we have moved the original Figures 2c-f to the Supplementary Material and we have included two supplementary figures in accordance with the additional analyses performed.

- 5) If the authors decided to conduct some specific clustering on four groups of patients according to molecular subtypes it would also be interesting to analyze if the same clustering could be performed according to the stage of cancer since groups with stages IIA, IIB and IIIA are relatively similar in amount and could be statistically analyzed. The subsequent analysis of the biomarker profiles which were different in various groups seems very interesting and could significantly improve the manuscript.

Response 5: Thank you for your insightful observation. In order to test whether the clustering according to the BC stages (II and III) was similar to that obtained when comparing by the BC phenotypes, we performed an additional analysis. So, the stages mentioned above were compared, in which the metabolomics features appeared to cause a natural biological separation in the PLS-DA projection (Figure S3) similar to that obtained when comparing the BC molecular subtypes (Figure S2). These findings suggest that future analyses in larger cohort of patients may provide with the specific metabolome alterations during the progress of the disease in a subtype specific manner. So, a better strategy for further research would be targeting individual molecular subtypes classified by stages. In our current work, such approach could drive to misleading results due to the low sample size of the BC phenotypes when stratified by stages. Thus, we have decided not to include this analysis within the present manuscript, but to consider working in this line for future research.

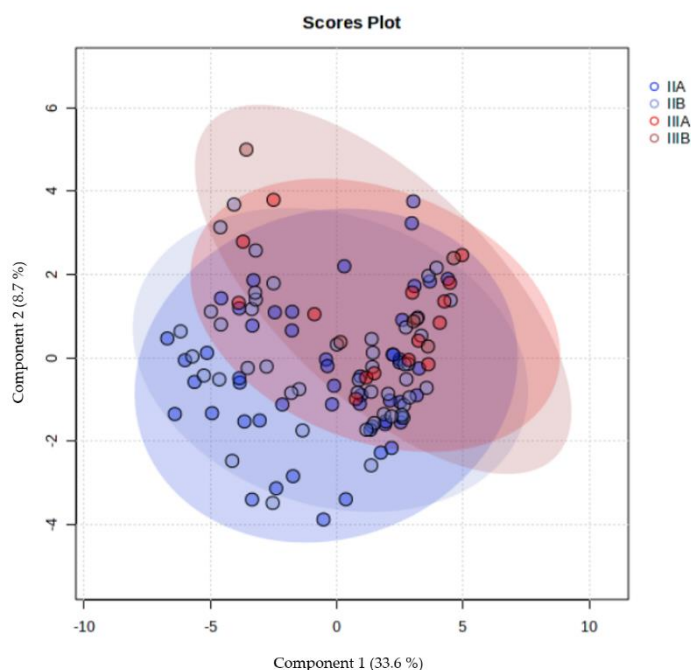


Figure S3. The PLS-DA score plots from BC samples stratified by stages shows a clustering according to the stages II (A- dark blue, B- light blue circles) and III (A – dark red, B – light red circles).

- 6) Lines 111-115 contain information about QC samples. When searching for some detailed information in Material and methods section, the following information was found (lines 290-292): “Blank solvents based on the MP samples and QCs were also analyzed during the run. The QCs were obtained by mixing small aliquots of all the biological samples under study. QCs evaluation provides with a measurement of the stability, quality and integrity of the system.” The mentioned approach seems to be unusual and the explanation of its use is important. Or some relevant references should be added, since [19] contain some general aspects. Were any internal standards added to the samples to control the analysis?

Response 6: Thank you for pointing this out. In accordance with your comment, we have added some more references in order to better support the mentioned approach which, indeed, is not unusual practice when performing untargeted metabolomics. In response to your question, these samples were previously analyzed in a RPLC positive mode, so the internal standards added to the samples to perform the peak alignment in the data pre-processing were: L-leucine, roxithromycin, L-alanine d³, caffeine d³, creatine d³. However, these molecules were not well retained in our HILIC negative conditions. In this case, the peak alignment was performed by monitoring mass error and retention time of well-retained compounds such as tryptophan, phenylalanine, threonine and arginine:

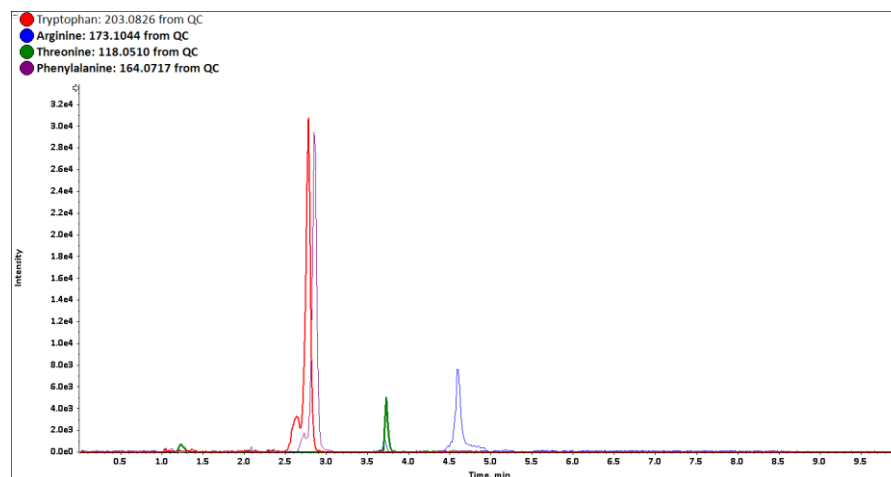


Figure Response 6. Extracted peaks from well retained compounds by using HILIC column and ESI negative.

- 7) Were the absolute peak areas used in the statistical analysis? No normalization procedure on any standard?

Response 7: Thank you for your question. The absolute peak areas were used and pre-processed by QC normalization, log transformation and pareto-scaling modes as stated in lines136-137, Section 2.1.1. Regarding the second question, as we explained in the previous response, we did not normalize based on the internal standards since they were not well retained under our LC-HRMS conditions.

- 8) Section 2.2.1. contains the molecular formulae of only 2 potential biomarkers from 27. What were the selection criteria?

Response 8: Thank you for pointing this out. Compounds' identification in untargeted metabolomics is considered the bottleneck of biological data interpretation and it can be

achieved by different levels of confidence as suggested by Schymanski et al. [24]. In our work, it could be claimed that the selection criteria followed for the identification was:

- 1.** To be statistically significant between the experimental groups (p-value < 0.05 and FC > 2)
- 2.** To have a MS/MS spectra
- 3.** To match with library spectrum data with an error mass < 5 ppm.

Therefore, we were able to find a tentative identity for 2 of the candidate metabolites with a level of confidence of 2 and 3 by the Schymanski classification.

References added to the main text according to response 8 – Reviewer 1:

24. Schymanski, E.L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H.P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48*, 2097–2098, doi:10.1021/es5002105.

- 9)** Sections 2.2.2 and 2.2.3 describe ROC analysis using all 27 biomarkers together, but not only two from section 2.2.1. I suppose that information about molecular formulae of all 27 potentially important molecules should be presented. Or the results of ROC analysis which demonstrate that only 2 biomarkers could be used to distinguish patients from controls should be demonstrated. However, all mentioned data could be interesting for readers.

Response 9: Thank you for your valuable observation. We completely agree that multivariate AUC analysis must be performed with only the biologically interesting candidates. Thus, the AUC-ROC analysis was repeated with the 2 identified metabolites and was changed in Section 2.2.3 (lines 198-200).