Peer Review File

Article information: http://dx.doi.org/10.21037/tcr-22-2686

Reviewer A

In this study, Lingli et al. propose a model of ferroptosis related genes in breast cancer using available data from TCGA and doing bioinformatic analysis through R. First, they search in the literature for genes related to ferroptosis and then analyze the differential expression of this genes in BRCA. Second, from those differential expressed they looked for the genes associated to survival. Finally, they propose a stratification of the BRCA patients in high risk and low risk according to the expression levels of the genes related to the ferroptosis. It is a study that bring some novelty and it is appropriate to the journal, but need to be improve.

Major concerns

1. The introduction is a little sloopy. The authors start a lot of phrases without uppercase. The statement "…in my country" line 32, needs to be edited to the name of the specific country.

The statement "BRCA is a highly heterogeneous tumor. Different molecular types have different prognosis, and there are also differences in the same molecular type. Therefore, it is very important to explore new treatment directions to improve survival prediction performance. Selective induction of cancer cell death is one of the most effective treatment methods for tumors." Is missing a reference.

Still in the introduction I would like to see a better explanation of BRCA classification and also about ferroptosis. In addition, some genes that is involved in ferroptosis.

Comment 1: The introduction is a little sloopy. The authors start a lot of phrases without uppercase.

Reply 1: Thank you for reviewing my artical in your busy schedule, thank you very much. This is my first time writing an article in English, so there will be many mistakes. I changed a lot of places, including the capitalization issue.

Changes in the text:see Page2,line29.36...

Comment 2: The statement "…in my country" line 32, needs to be edited to the name of the specific country.

Reply 2:I changed it to "China"

Changes in the text:see Page2,line32

Comment 3:The statement "BRCA is a highly heterogeneous tumor. Different molecular types have different prognosis, and there are also differences in the same molecular type. Therefore, it is very important to explore new treatment directions to improve survival prediction performance. Selective induction of cancer cell death is one of the most effective treatment methods for tumors." Is missing a reference.

Reply 3:Thanks to the teacher's careful examination, I added a Reference there

Changes in the text:see Page2,line39.

Comment 4:Still in the introduction I would like to see a better explanation of BRCA

classification and also about ferroptosis. In addition, some genes that is involved in ferroptosis.
Reply 4:In this section, I newly added breast cancer-related classifications, and related research on ferroptosis genes.
Changes in the text:see Page2,line41,49

2. The classification of BRCA in different subtypes is important to determine the prognosis of patients. It would be very interesting if the authors did the same analysis (of the ferroptosis genes) in each subtype: Luminal A, Luminal B, Her 2 positive and basal. It would aggregate value to the study.
Reply:Thank you for your suggestions. This is a valuable research direction. I try to classify them. I went back to review the literature, and many studies were done in three negative breast cancer. This is also where I further explore.
I have made a comparison of iron death genes in breast cancer with different molecular type.G1:LunimalA,G2:LunimalB,G3:HER2-riched,G4:basal like.See in supplementary picture

3. The explanation about how the genes where choose for the analysis should be describe in the methodology.
Reply:We added this part to introduce the source of related genes
Changes in the text:see Page3,line68.

4. In the line 57, the authors said that analysed 1097 BRCA tumors from TCGA, is it corrected? Or, it is 1097 samples divided in tumor and non-tumor? Also, the authors did not cite the number of non-tumor samples analysed.
Reply: Our description is inaccurate and has been changed to“ RNA-sequencing expression (level 3) profiles and corresponding clinical information for BRCA were downloaded from the TCGA dataset(https://portal.gdc.com).113 paracancerous samples were collected from TCGA database and 54 samples were collected from non diseased tissue sites in GTEx.”
Changes in the text: see Page2-3,line57.

5. The authors could add a supplementary table with the results of the cox regression analysis for each gene.
Reply:I added this part according to your request.See in supplementary table

6. The figure legends are also sloopy.
Reply: Thank you very much for your suggestion. We revised this part again. If necessary, we will ask professionals to modify this part
Changes in the text:see Page8,line226-243.

7. The discussion in simple, but I liked it. It brings everything that we should know.
Reply: Thank you very much for your encouragement. Your encouragement gives me motivation

Reviewer B

General comments
The study described in the manuscript addresses an interesting aspect that is the predictive value of ferroptosis-related genes in breast cancer, but the findings are limited to one dataset. In silico, data mining is an important and valuable tool to generate new hypotheses; however, the statistic and bioinformatic analysis should be well performed and the findings further validated. The present manuscript shows many inconsistencies, and the authors should adequately address some concerns before possible publication.
Specific comments
1.In the Material and Methods section, the authors should provide a valid link for TCGA dataset download. The link provided is unavailable.

Reply: Thank you for taking the time to review the article. Your suggestion is of great help to me. I revised the link again
Changes in the text: see Page3,line59.

2. The authors found that 19 out of the 25 (line 60) ferroptosis-related genes showed statistically significant differences in tumors compared to normal tissue however, data extracted from the UALCAN platform showed that all target genes showed expression differences in tumor tissue compared to normal tissue.

Reply: I went to verify it ,and it is right. All target genes of UALCAN platform showed expression differences in tumor tissue compared to normal tissue. We revisited the GSCA (http://bioinfo.life.hust.edu.cn/GSCA/#/)database and found that not all genes have survival differences(picture 1 to reviewer B). Compared with several data, we choose the least. TCGA database is one of the most commonly used databases for survival analysis.

3. The authors should re-evaluate their data or provide the complete list of expression differences of the ferroptosis-related genes compared between tumor and samples comparing their data with the UALCAN data as Supplementary data.
4. The authors found that only 5 ferroptosis-related genes (CISD1, ALOX15, CS, CARS1, and EMC2) showed significant differences in survival in BRCA, and there was no significant survival difference in the other related ferroptosis genes ($p < 0.05$) (lines 94-97). However, using the UALCAN platform, other ferroptosis-related genes showed survival differences, such as SLC7A11 ($p=0.0065$) and ATP5MC3 ($p=0.00011$).

Reply: Different databases have different data information. We should choose the common parts of the two databases. The UALCAN platform has seven genes, while the TCGA platform has five. Take the intersection part and choose the least
Changes in the text: see Page3,line59.

5. The authors should explain how they obtained the cut-off for stratification in high and low expression of the ferroptosis-related genes. CISD1, ALOX15, CS, CARS1, and EMC2 genes showed notable differences in transcripts abundance. How was the stratification of each gene to get exactly 548 tumors with high expression and 548 tumors with low expression for all these target genes?

Reply: Thank you for your advice. We revised this part and used lasso Cox Iterative Regression Method to reduce the dimension(picture 2 to reviewer B). We used median survival time to distinguish between high-risk group and low-risk group. The log rank test was used between groups. HR (high risk) represents the risk coefficient of the high-risk group relative to the low-risk group
Changes in the text: see Page3,line78

6. A larger cohort, second independent cohort should have been used to validate the findings. Other datasets, such METABRIC, should have been consulted as well to evaluate.

Reply: Thank you very much for your suggestion. Our plan is to screen genes in the database first, and then further verify them with our own tumor samples. Thank you for your understanding.

7. In silico analyses are limited to the TCGA dataset and are not addressing clinicopathological indicators such as tumor size, stage, Ki67/proliferation, nodal status, treatment information. The authors make no mention of the study's limitations and draw conclusions based on findings from one dataset with overinterpretation.
Reply: Thank you very much for your suggestion. We have added an explanation of the limitations of this result
Changes in the text: see Page6,line171

8.The English have to be reviewed to improve the quality and clarity of the manuscript.

Reply: This is my first time to write an article in English. Thank you for your understanding.