The manuscript entitled: "Biochemical, anatomical, and histochemical characterization of cachichín (Oecopetalum mexicanum Greenm. & C.H. Thomps.) seeds" seeks to evaluate alterations in the carbohydrates and proteins composition of O. mexicanum submitted to different types of heating processing: raw form (not heated); boiling, commercial toasting, and controlled toasting. The presented results are interesting because they demonstrate that the chemical and structural composition of the seeds can vary depending on the different thermal treatments, which can directly affect the nutritional power seed. Studies of this nature are important as they expand the possibilities of using new species as nutritional sources for global food.

Despite the study relevance, the work has serious deficiencies that make it unfeasible to publish it in the PeerJ journal now.
Below I make some more targeted comments.

BASIC REPORTING

In general, the manuscript presents a clear language with professional English. However, the introduction is very superficial and contextualizes very little with the background information available in classical literature. Many essential aspects for understanding the work should be presented in the introduction. An example of this are the possible chemical interactions between the studied biomacromolecules and the Maillard reaction, which were abruptly presented in the discussion. Authors need to consider that not all readers know the minutiae of the topic addressed in the paper, requiring clearer writing to be adopted to contemplate a more diverse audience. Furthermore, I believe that the number of figures is unnecessary. Many figures could be dropped, and others condensed into a single figure. I believe that the understanding of the results becomes even clearer. Finally, the raw data is inaccessible. As the raw data are in .csv format, it was not possible to access them because the data was misconfigured. I believe that the authors, anticipating possible formatting problems, could have been more careful and converted the raw data to .docx or .xlsx format.

EXPERIMENTAL DESIGN

The work is original and presents a relevant and clear objective, and practically all methods have been partially carried out and described with rigor and clarity. Below I point out some problems and suggestions detected in the methodological questions.

Although I think that the data presented in the manuscript are still very superficial, I consider it to be a starting point for future research to provide relevant discoveries for academia and society in general. However, I believe that many aspects need to be improved, mainly from the statistical perspective and anatomical analyzes to make the work with an acceptable data set for publication. In addition, many aspects presented and discussed are guided by a speculative approach that do not match the real contributions of the study.

STANDOUT REVIEWING TIPS

Next, I will make some guided comments and suggestions for specific points in the manuscript.

Manuscript title

Include botanical family name.

Abstract

The abstract is the cover letter of any paper. Therefore, it needs to be written very carefully and accurately. The abstract of the manuscript presents a serious flaw in my opinion, as the object of the work is not presented. I suggest that the authors review this point in the abstract.

Introduction

Line 57 - Include the name of the botanical family only in the first citation of the species in the text, from that moment on the abbreviated form must be used (eg. *O. mexicanum*)

As mentioned earlier, the introduction is superficial and does not lead the reader to a complete understanding of the studied subject. I suggest that the authors provide a major revision in the introduction and present the biochemical and structural aspects of seeds in general. Discuss the interactions between biomacromolecules and the Maillard reaction.

Materials and methods

Topic: Seed anatomy by scanning electron microscopy (SEM)

Just a suggestion on the microscopy processing. Whenever one chooses to use aldehyde-based chemical fixatives, it is convenient to use a mixture of 2.5% glutaraldehyde and 4% formaldehyde (Karnovsky, 1965). As the principle of chemical fixation is through the diffusion of fixative agents in the sample, it is convenient to use formaldehyde because it is a simpler molecule (CH2O) which facilitates its ability to penetrate the sample. However, its ability to cross-link with the -NH3 proteins groups is limited by the low aldehyde group amount in the molecule (1 group). To compensate for this, it is convenient to always use glutaraldehyde, as it has two aldehyde groups in its molecular formula (C5H8O2). However, the larger size of the glutaraldehyde molecule may compromise its penetration into the sample. Many cases of observed structural alterations can characterize artifacts generated by poor fixation of the samples.

Another note that I would like to highlight is that even though the sample fixation was carried out under vacuum conditions, it is convenient to leave the sample for at least 24 hours in immersion in the fixative in a refrigerated environment. In the case of seeds, this immersion time in the fixative is essential because the thickness of the seed coat makes it difficult for chemical fixatives to penetrate.

Histochemical stains

Were the seeds used for the histochemical tests also submitted to the fixation step? If yes, this needs to be informed, as the PAS test must always be done with fresh samples, since the fixing agents can interfere in the efficiency test.

Experimental design and statistical analysis

I consider the methodology used in the statistical analysis inadequate. To apply the ANOVA, some assumptions need to be respected**:** Initially, it is important to analyze the data normality using the Shapiro Wilk test (n<30). If the distribution is not normal, it is convenient to transform the data through the Log10 (decimal values) or by the square root (integer numbers). If, even after the transformation, the data remain with a non-normal distribution (non-parametric), it is convenient to use a non-parametric variance test (eg. Kruskal-Wallis).

It is convenient to analyze the variance homogeneity using Levene's test. If the variance is not homicedastic, it is convenient to adjust the degrees of freedom (df) using the Welch approximation. In this case, during the script construction of the variance analysis, just add the argument "var.equal = F" and R automatically makes the correction.

Finally, it is very important to evaluate the outlier's presence in the data set, in which case it is interesting to build a boxplot to identify leveraging outliers.

Despite not being clear in the methodological description, the authors implied that the statistical analyzes were carried out on a data set with six repetitions. This sample N is too low for the application of a parametric variance test (eg. Anova). In this situation it is usually applied direct non-parametric variance tests (eg. Kruskal-Wallis).

Authors need to pay attention to these statistical issues, as the use of wrong statistical methods can inflate or hide differences between treatments.

Results

Line 187 to 198 - The description of results for Concentration of total and reducing sugars and Concentrations of proteins and amino acids is very confusing. The authors use codes to refer to the treatments (T1, T2, T3 and T4) and put the respective descriptions for each code in parentheses (raw, Boiled, Commercial toasting, Controlled toasting), this is long-winded and makes reading confusing. This becomes even more complicated, when looking at Table 1 and the treatment codes are omitted. The authors need to decide whether to use the codes or describe the treatments.

Furthermore, the authors describe these results as follows: *"For total sugars, the highest means were observed in treatments T1 (Raw) and T4 (Controlled toasting)...."*. This is not a suitable way to present a result that seeks to establish a comparison of means. In this case, the authors need to add the mean values and standard error, followed by the test significance (p-value) to the text so that the reader can immediately identify the magnitude of the differences between the means.

Eg. "For total sugars, the highest means were observed in treatments T1 (488 ± 28.53, p = XXX) and T4 (491 ± 12.69, p = XXX)..."

Another point that needs to be considered is the fact that the authors adopted percentage values to show how much one treatment increases or decreases in relation to another. This is an interesting way to present the results, but I suggest that the authors add a column to Table 1 demonstrating the delta variation between treatments.

Line 202 to 203 – The authors mention information about the ovoid shape, length, and width of the seeds, but do not present any image of these seeds by stereomicroscopy. I suggest that these images be provided.

Line 204*, topic: Internal structure of the seed* – The authors said: “A papyraceous tissue covers the endosperm, probably the endotesta. A developmental study is needed to establish its origin and nomenclature.”

That's too vague. What is this papyraceous tissue? They do not present any image of this tissue type.

In my understanding, this papyraceous tissue is basically the tegument of the seeds. The origin of this tissue type is very clear in the literature, and an ontogenetic study is not necessary to determine the nomenclature or origin this tissue.
I request additional explanations from the authors.

Line 242 to 247*, topic: Structure of the external integument or testa of the seed*

In general, SEM images are of excellent quality and add a lot to the work. However, I believe that SEM imaging to assess changes in the organization of macrosclereids, brachysclereids and osteosclereids is not the best option. Changes in these cell types are best observed using light microscopy (LM) and transmission electron microscopy (TEM). The alterations indicated by the authors in these cell types are not so evident through the SEM and perhaps could have been provoked during the manual section of the seeds. I advise authors to adopt historesin blocking techniques for LM and Epoxy or LR White resin for TEM to certify that such alterations are the result of treatments and not of manipulation of the material.

Topic: Histochemical staining and effect of thermal processing on the biochemical composition of the seeds

Histochemical tests present interesting results, however, the images have low quality. Most images show artifacts resulting from folding and rips in the sample. Some images have problems in the background that could be easily resolved in image processing programs without altering the result fidelity. Under these conditions, I suggest that authors perform new histochemical tests in the search for better quality images. An alternative might be historesin blocking, or depending on the hardness of the seed coat, it might be convenient to use slide microtomes that are used to section wood samples.

Discussion

Line 339 to 340 – The authors to affirm that “Therefore, heat treatments do not decrease the protein concentration, but change their organization in the intra and extracellular space of the cachichín seed cells”.

However, this is a speculation that cannot be substantiated based on the results obtained. Based on the data obtained, it is impossible to state that heat treatments can influence the intra and extra cellular reorganization of proteins.

Line 380 to 381 – The authors to affirm that “lipid staining (Figure 7a) revealed the presence of spherosomes covered by a unit of endoplasmic reticulum, where lipids are stored in endosperm cells”. This is a purely speculative statement and impossible to prove with the results presented. This would only have been possible if the authors had presented transmission electron microscopy data.

Conclusion

Line 400 to 402 – Remove this sentence.

Figures and Table

Figure 1 is the same image as Figure 2A. So, I don't see the need for Figure 1.

Figure 2. I suggest that authors identify the seed structures in Figure 2A, and ask them to choose more intuitive acronyms to identify such structures (eg. Emb = Embryo; Rad = radicle; etc...). Do this for all images.

Figure 3. Identify the starch grains in the images, remember that not all readers are aware about the seeds structure.

Figures 4 and 5. I see no validity in these figures without a light microscopy image to prove such changes.

Figure 6, 7, 8, 9 and 10. They are of poor quality and need to be improved for publication. Also, I think they could be grouped into a single Figure.

Table 1. Decide whether to refer to treatments by codes (T1, T2, T3 and T4) or by their respective names.