

Manuscript reviewing 95688v2

1. Basic Reporting

The authors did their best to address all the raised issues from the first review of their manuscript, but are now encouraged to proof read thoroughly their manuscript as misspelling, convoluted sentences or sentences without verb were spotted. Consequently, a few inconsistencies in the way data are represented can also lead to misunderstanding. In its current state, the manuscript still requires key revisions. For examples:

- Line 281: "additionally" is misspelled
- Line 287: "according with" should be replaced by "according to"
- Lines 299, 304: Euclidean is misspelled twice.
- Line 417: "M344T129 are putatively", should be "M344T129 were putatively"
- Line 470-471: the sentence lacks a verb.
- Line 471: "the molecular ion produced by" should be replaced by "the molecular ion detected in". **BE careful!** *A molecule is produced by an organism and a molecular ion is detected by mass spectrometry in the sample.* **Check also line 466**
- Line 545: the sentence contains the word "tool" twice.
- Line 599-600: the last sentence should be revised what do the authors mean by chemical database? Is it spectral databases for dereplication?

The word "approach" is intensely used and should ideally be replaced, whenever possible by a more precise and appropriate one. Examples

- Line 515: "Using this integrative systematic approach" could be replaced by "Using this integrative systematic method"
- Line 546: "we propose the inclusion of new tool Molecular Network approach tool" could be replaced by "...we propose the inclusion of Molecular Network chemometrics to circumvent this key issue."
- What do the authors mean by biochemical data? It is not really clear in the manuscript. Do they refer to MS-based metabolomic data? Or to the combination of DNA based analysis and metabolomics?
- Line 540: the authors keep repetition in the paragraph both *metabolomic data* and *molecular data* often in the same sentence and should be more precise in the wording to avoid any confusion to the reader. Molecular barcoding sequence could be used instead of molecular data for instance. **Please check throughout.**
- Line 544: the word annotation should be replaced by identification
- Line 545, in the discussion section, the authors mention the Molecular Networking as a NEW tool. But the methodology exists since 2016. As such, it can't be considered as a new tool per se. It is now widely implemented in natural product chemistry for the structural dereplication of molecules through comparison of their MS2 spectra, notably with those available in the associated databases in the GNPS environment.
- Lines 550-551: the sentence should be revised. As written it is not a sentence a verb is missing.
- Line 552-555: The sentence should be split in two.

2. Experimental design

Nothing to add as previously reviewed the first time.

Raw data: The links to access to the raw data should be available through the manuscript unless there is a specific requirement from PeerJ.

3. Validity of the findings

All the required modifications or raised issues have been addressed in the revised version of the manuscript. Nevertheless, some key clarifications need to be brought as per the putative identifications are concerned for data accuracy. I acknowledge the efforts made by the authors for putative identification of molecules but there are key mistakes in table 3 and in the corresponding paragraph starting in line 411 that need to be corrected.

When an MS-based putative identification is proposed, confidence levels should be added following the generally accepted guidance (see Schymanski et al. 2014, and more recently Alseekh et al. 2021)

- Lines 415-416: the authors stated: “the limited chemical studies ... prevented from identifying the 13 molecular ions”. Nevertheless, a molecular formula without necessarily proposition a structure could possibly be given with a confidence level 4 (Schymanski et al. 2014).
- Line 420-421: “by manual inspection of the mass spectra”: please precise HR-MS1 and MS/MS spectra?

Table 3: MS-based putative identification of 29 molecules from Octocoral species

- title “molecular ion” is not correct. This title should “measured m/z ”
- Title “ m/z ” is not correct. This title should be “Ion Type” which should be presented [$M + H$]⁺
- The authors should add two columns; one for the “calculated m/z ” and the other for the “error in ppm”.
- This information is crucial as it will reflect the accuracy in the proposed molecular formula leading to the putative identification.
- For example in the case of muricenone B: the calculated m/z for [M+H]⁺ = 333.2424 but the measured one is 333.1621 so the associated error is too high for the proposed molecular formula, leading to incorrect data. This issue concerns the entire table 3.

What VIP scores for these molecules? . To which M..T.. feature these identifications pertained? Do they contribute to species discrimination? If not the authors should be careful these data lack accuracy or worse are incorrect as presented.

Figure 4 legend

- *M*= Parent mass number, should be correct as M = Parent mass integer
- *T*= retention time number, should be corrected as T= retention time in min (? or in sec?)

Figure S4: The M..T code for the features is not used on the Y axis of the VIP scores. Please check the consistency of your codification throughout your manuscript.