

Manuscript Title:

Ordovician to Silurian graptolite specimen images for global correlation and shale gas exploration

Corresponding Author:

Hong-He Xu

REASON TO PUBLISH

As in my first review, I welcome this approach to sharing the contents of paleontological collections in museums and research institutes. It could save considerable travel costs for research based in part upon remote collections. It may allow the verification or updating of identifications published in the papers cited in the XLSX file. I imagine that the quality and quantity of the new images far exceeds that in some of the original publications. Naturally, they cannot exceed the quality of the specimens themselves.

The accompanying manuscript does not need to be long or detailed. It is sufficient in its current form. Most of the data users will likely have their own expert reasons to access the Zenodo database. The manuscript serves to advertise its existence. The 5-6 short paragraphs on the Zenodo website are an adequate introduction and well written. The ESSD manuscript has the added benefit of some references and several helpful illustrations.

AN IMPROVEMENT and A SERIOUS NEW PROBLEM

The revision notes indicate that the authors have now added the essential scale bars that were promised but missing on images in Version 1 (August 16 2021) of the Xenodo data repository. The new version is presumably enlarged by adding a second set of images that include scales. Although this would be a very welcome correction, I am still struggling to verify it by opening the Version 2 (Jan 31, 2022) ZIP files. The downloaded ZIP files are not readable or extractable by my Windows computer. Neither the file explorer nor the customary unzipping utility can read them. I have experimented with different utilities and tried renaming the files, but to no avail. I can still read and extract the Version 1 files. So, my computer is presumably not the problem. Until this failing is addressed, there are effectively no scaled images in the database. It should not be published in this condition. I would have expected the ESSD staff to at least check that the data are accessible before sending the manuscript for review. Perhaps they did and this problem simply does not arise for every user.

CONTINUING CONCERNS

The data-rich XLSX file is the index to the image collection. Surely it should indicate for each taxon or specimen, which of the ZIPPED files contains the corresponding images. It should not be necessary to download, extract and examine all the ZIP files to find images for a single taxon, publication, or locality. Perhaps I have somehow overlooked this information. Downloading the large ZIP files may be slow, but not unreasonable, unless one needs to download all eleven files when perhaps only one is really needed.

There is potential confusion about the Zenodo web address. In the authors' response it is misprinted – the last digit is missing. The revised text provides the address in two places, but they are different. The abstract has been edited to give the new web address for Version 2 of the data. The "Data Availability" section still has the old address; that is for Version 1 of the database which lacks the scales.

Although the data are surely intended for use by experts, the manuscript includes some very elementary facts about graptolites and Paleozoic stratigraphy. Not all are strictly correct. The local first appearances of graptolite taxa, for example, are a means to correlate with locations away from GSSPs for many Ordovician and Silurian stage boundaries. The graptolite taxa are indicative, not definitive. First appearances may be diachronous and earlier occurrences may even be found at stratotype sections. The “spikes” are definitive, even if less practical.

The various sections of the paper and its figure captions tend to be repetitive and still contain many common errors of English. The repetitiveness is irritating and serves no purpose. The grammatical flaws do not obscure meaning and are not surprising for non-native English-speaking authors. The usual mismatches of singular subjects with plural verbs occur in the first two words of section 1 – “Graptolite was . . .” The database managers want readers to trust taxonomic revisions made by un-named experts of their choosing. It would be unfortunate to undermine this by failing to find an English language proof-reader. At one time, trained journal staff would undertake such corrections, but that is less common in large publishing houses today. Fortunately, universities are more cosmopolitan.

REASON TO REQUIRE FUNDAMENTAL REVISION

Given the concerns about quality-control, which range from trivial to nullifying, it is possible that readers will lose trust in the authors’ attention to other aspects of their data. Certainly, it begins to shake my confidence in the project. Am I making an elementary unzipping mistake or are the database managers careless? Because the credibility of the data is at risk, it is important to address even minor concerns. I recommend easy remedies for the authors. They should enlist a set of volunteers with different computer expertise and hardware to test the accessibility of their data files. It should also be a simple matter to find a native English-speaking paleontologist who can quickly correct the English grammar and syntax, given an editable text file. I can correct English grammar, but I am worried about database managers and editors whose revised datafiles are unreadable on a very standard Windows computer that can still read the previous version.

Publishing the data in its current format would be a serious disservice to the admirable effort of compiling this large set of illustrations.