Title: Assessing Open Science Practices in Phytolith Research

Author: Emma Karoune

Abstract:

Open science is becoming an integral part of all scientific research but the extent of these practices in phytolith research is unknown. Phytolith analysis examines silica bodies that are initially formed within and between plant cells during the life of the plant but become deposited in sediments once the plant dies. Their use in both archaeobotanical and palaeoecological studies has been ever increasing in recent years and consequently has seen an upsurge in publications. This article aims to assess open science practices in this field by reviewing data and metadata sharing, and open access, in a sample of journal articles containing primary phytolith data from 16 prominent archaeological and palaeoecological journals (341 articles). It builds on similar studies conducted for zooarchaeology (Kansa *et al.* 2020) and macro-botanical remains (Lodwick 2019). Data in this study was collected concerning data format, reusability of data, inclusion of phytolith morphotype pictures for identification purposes and a fully described method, use of the International code for phytolith nomenclature (ICPN) and whether the articles were open access. Steps forward are then suggested to use as a starting point for discussions in the wider phytolith and archaeological communities to develop guidelines for greater integration of open science practices.

Keywords: Phytolith, Micro-botanical remain, Open Data, Open Science, Open Access Archaeobotany, Palaeoecology

Lay summary:

Open science is the opening up of scientific research to the academic community but also the wider public. It has four strands: open data (the ability to access scientific data freely), open code (publication of computer coding used in data analysis), open papers (free open access to scientific publications) and open review (transparency and publication of reviews conducted by scientists of their colleagues work). These practices are becoming an integral part of all scientific research but the adoption of these practices in phytolith research is unknown. Phytolith analysis examines microscopic silica shapes formed within and between plant cells during the life of the plant. These remains are then deposited in sediments when the plant dies. They can be used to find out what plants grew in past environments and also what plants were used by ancient people. The use of phytoliths in archaeobotany (archaeobotany = the study of plant remains from archaeological sites) and palaeoecological studies (palaeoecology = the study of past ecosystems using fossil animals and plants) has increased in recent years and consequently has seen a growing number of publications. This article aims to assess open science practices in phytolith research by reviewing data sharing and its associated background information, and open access, in a sample of journal articles containing new phytolith data from 16 well-known archaeological and palaeoecological journals (341 articles). It builds on similar studies conducted for zooarchaeology (the study of animal remains from archaeological sites) and macro-botanical remains (the study of seeds and other larger plant remains found on archaeological sites). Data in this study was collected concerning data format, reusability of data, inclusion of phytolith pictures for identification purposes and a fully described method, use of the International Code for Phytolith Nomenclature (ICPN – a standard system for naming phytolith shapes) and whether the articles were open access publications. Steps forward are then suggested to use as a starting point for stimulating discussions in the wider phytolith and archaeological communities. It is hoped that growing awareness of the importance of these issues will help to develop community-wide guidelines and lead to greater adoption of open science practices.

Assessing Open Science Practices in Phytolith Research

Emma Karoune

1. Introduction

1.1. What are phytoliths and how are they used in archaeology and palaeoecology?

Phytoliths are silica-bodies that form within plant cells during the lifetime of the plant (Madella & Lancelotti 2012). Monosilicic acid (H_4SiO_4) in groundwater is absorbed by the plant roots and is eventually deposited as solid silica dioxide (SiO_2) in and between the plants cells forming distinctive shapes (morphotypes). These shapes can be used to identify plant taxa to different taxonomic levels. When the plant dies, phytoliths become deposited naturally in the soil or, if the plant has been exploited by humans, in archaeological layers. Phytoliths are often stated as having in-situ deposition and lack of movement within deposits (Iriarte 2003; Iltis-Davey *et al.* 2007; Piperno 2006) but the movement of phytoliths before they become deposited (transport) and once deposition has taken place is still under debate needing further investigation (Madella & Lancelotti 2012; Shillito 2013:79). In any case, these inorganic plant remains are preserved over long time periods and can therefore be readily found in archaeological and palaeoecological samples. Consequently, they are used to address research questions in combination with, or instead of, other plant remains that do not survive as well.

Phytoliths are being used for an ever-expanding list of research topics and applications (Hart 2016; Hodson *et al.* 2020). In recent years, there has been a substantial increase in research, and this has resulted in an upsurge in publications. Archaeological phytoliths are examined from deposits using a variety of methods such as the sub-sampling of bulk samples taken for flotation, column samples from exposed sections after excavation and increasingly a combination of soil micromorphology and then sub-sampling of the same column for sediment processing (Madella & Lancelotti 2012; Pearsall 2019; Piperno 2006). They are also extracted from ecofacts and artifacts such as dental calculus, coprolites, pottery and grinding stones. Applications in palaeoecology include reconstructing palaeoenvironments and palaeoclimates from core samples, and the newer applications of carbon-14 dating and isotopic analysis of phytoliths are also being increasingly explored.

With increased research comes the difficulty of how to standardise methods, share data and validate analyses. Phytolith research is a less established discipline than macro-botanical analysis or palynology. The stage of innovation currently seen in phytolith research is producing a wide range of data, both quantitative and qualitative, from archaeological, palaeoecological and methodological studies. Phytolith data falls into several categories - observational, experimental and computational data (Marwick & Pilaar Birch 2018). These data must be accessible and reusable so that other researchers can review, adapt, apply and collate their colleagues research.

The majority of phytolith research is conducted by large research institutions such as Universities and Museums. The cost of establishing a laboratory, processing and analysing samples for phytolith analysis is high and this restricts its use in the commercial sector. In the UK, the use of phytolith analysis in commercial archaeological units is rare and hindered by the lack of a comprehensive reference collection for the British Isles. Macro-botanical remains are the routinely analysed botanical element of archaeological sites in Britain. Much of the phytolith research conducted in UK research groups concerns more distant areas of the world such as tropical regions. In these geographic areas, organic remains are often poorly preserved in archaeological deposits and therefore the higher research costs of phytolith research can be justified to enable archaeobotanical questions to be adequately addressed. Consequently, phytolith data is rarely seen in the commercial grey literature, apart from university dissertations, and is therefore

predominantly published in journal articles. It is therefore important that the issue of data sharing, and particularly data communicated in journals, is comprehensively addressed to achieve the most value out of this current surge in research.

The International Phytolith Society (IPS) have started to address some of the issues of standardisation by setting up working groups for taxonomy (International Committee on Phytolith Taxonomy - ICPT) and morphometric methods (International Committee on Phytolith Morphometry - ICPM). However, despite these efforts, there is a need for these endeavours to be brought in line with open science practices to be more effective. The International Code for Phytolith Nomenclature was first published in 2005 (Madella *et al.* 2005), and recently updated by Neumann *et al.* (2019), and the first guidelines for standardisation of morphometric analysis were released in 2016 (Ball *et al.* 2016), but their impact throughout the discipline has not been thoroughly assessed. It has also not been made explicit how this form of standardisation can aid data sharing. Zurro *et al.* (2016) go further to suggest several minimum requirements that should be addressed in any phytolith study and publication including disclosure of full methods, use of ICPN and the publication of raw data (absolute and relative presence). This is a much more holistic approach, but again, the extent to which this has been implemented by researchers needs to be assessed to move phytolith research into the age of open science.

1.2. Why is open science so important for phytolith research?

Open science is more than just the opening up of scientific research. It is a multi-faceted movement to improve science based on four pillars (Masuzzo & Martens 2017): open data, open code, open papers and open reviews. At every stage of the research lifecycle there is the opportunity to make work open and embedding this approach in all scientific research has many benefits such as increasing public trust in research, supporting scrutiny and validation, enabling reuse and driving innovation (FOSTER 2018).

This article is coming out of a need for reform in archaeology to embed open science practices more deeply throughout the discipline. Kansa (2012) suggested that open access and open data lead to greater resilience for our Profession but that this approach needs more champions and remains at the edges of archaeological practice. Over the past decade there has been growing awareness of this issue and a recent article summarising open science in archaeology (Marwick *et al.* 2017) sought to encourage this approach by highlighting its benefits to the archaeological community. Subsequent articles include focuses on computational archaeology (Schmidt & Marwick 2020), data sharing in zooarchaeology (Kansa *et al.* 2020) and innovative open field archaeological strategies (Marchetti *et al.* 2018).

There are limited reviews of open science in Environmental Archaeology with zooarchaeologists very much taking the lead (Kansa *et al.* 2020; Kansa & Kansa 2013, 2014; LeFebvre *et al.* 2019). However, recently an assessment of data sharing and citation for archaeobotany (macro-botanical remains) was published by Lodwick (2019), which found a need for improvement. Only 56% of 239 papers provided primary data, and much of this data was unusable. At present, the state of data sharing and other open science practises is not known in phytolith research and therefore there is a need to discover the current situation.

In more general terms, as a researcher, it is increasingly important to implement open science practices as it is becoming a requirement in many areas of research, particularly data sharing and publication. Most journals require or encourage data to be deposited in some form, whether this is in the text, attached as a supplementary file or in an open repository. Data availability policies are included in most journals' author guidelines, although they vary in robustness (see Table 1 for the data availability requirements of the journals in this articles' dataset). Some journals encourage data sharing, but others have much stronger statements that require data to be made available in line with the FAIR principles (FAIR = Findable, Accessible, Interoperable and Reusable –

Wilkinson *et al.* 2016). Recent research has shown that there is variation in the extent to which authors provide data that conforms to these policies (Christensen *et al.* 2019; Womack 2015). Reasons suggested for this mismatch could be the result of varying enforcement by editors or an incomplete understanding of how to implement the requirements by authors, reviewers and editors (Marwick and Pilaar Birch 2018:131). Whatever the reason, the benefits of data sharing must be stressed to researchers as it offers increased citation, increased collaboration and even new funding opportunities.

In addition, funding bodies such as the UK Research and Innovation (UKRI) and the European Commission, that fund a large proportion of archaeological research in the UK, now require a data management plan that must set out what data will be created, how this will be stored and how it is to be shared. Funding can be allocated in applications for the management of data and block funding for open access has been provided by UKRI to research institutions to pay for the article processing charges (APC) to enable Gold access publications for instant free publications. Therefore, it is in every researchers' best interest to become aware, receive appropriate training and implement open science practices as a natural part of their research cycle before it becomes mandatory for all research activities.

Table 1: data availability requirements of journals sampled in this study.

Journal title (Publisher in brackets)	Data availability policy (all accessed June 2020)
Vegetation History and Archaeobotany (Springer)	Encourages authors, where possible and applicable, to deposit data that support the findings of their research in a public repository – from 2016 (Lodwick 2019).
Archaeological and anthropological science (Springer)	Encourages authors, where possible and applicable, to deposit data that support the findings of their research in a public repository.
Environmental Archaeology (Taylor and Francis)	Encouraged to supply data – list of positives given. Supplementary information to be put in Figshare.
The Holocene (SAGE publishing)	Acknowledges the need for data availability but does not require or encourage authors to submit primary data. If data is submitted, the data is made available whether article open access or not.
Journal of Archaeological Science (Elsevier)	'This journal requires and enables you to share data that supports your research publication and enables you to interlink the data with your published articles. Research data refers to the results of observations or experimentation that validate research findings. To facilitate reproducibility and data reuse, this journal also encourages you to share your software, code, models, algorithms, protocols, methods and other useful materials related to the project' -from 2018. Pre-2018 statement – all data must be available in supplementary files or repository but not explicit about facilitating reproducibility. - Marwick & Pilaar Birch (2018) – say this requirement is from at least 2013.
Journal of Archaeological Science reports (Elsevier)	Same as JAS above – requires .
Journal of Ethnobiology (Society of Ethnobiology)	No data guidance.
Quaternary International (Elsevier)	Same as JAS above – requires .
Journal of Wetland Archaeology (Taylor and Francis)	Authors are encouraged to share or make open the data supporting the results or analyses presented in their paper where this does not violate the protection of human subjects or other valid privacy or security concerns.
Antiquity (Antiquity publications)	No data guidance.
Journal of Field Archaeology (Taylor and Francis)	No data guidance.
Oxford Journal of Archaeology (Wiley)	No data guidance.
Journal of Anthropological Archaeology (Elsevier)	Same as JAS above – requires .

Journal of World Prehistory (Springer)	Encourages authors, where possible and applicable, to deposit data that support the findings of their research in a public repository.
PLOS One (PLOS) – Open access journal.	Required to make all data necessary to replicate study's findings publicly available without restriction – from 2014. Pre-2014 – all data not required but encouraged to deposit database in repository – Dryad.
Proceedings of the National Academy of Sciences – Open access after 6 months/free for developing countries to access.	Authors must make material, data and associated protocols including code and scripts available to readers of publication – from 2011. <ul style="list-style-type: none"> - Should follow FAIR (this part of the statement was added first in 2018). - Deposit data in community approved public repositories. Pre-2011 – authors encouraged to deposit data and use Supplementary information.

2. Methods

The aim of this research was to obtain a snapshot of where phytolith research is currently in terms of open science to establish a starting point that can then be used to raise awareness, develop and implement training and initiate discussion to move phytolith research forward. Therefore, as phytolith research is predominantly published in journal articles, a dataset was gathered from prominent archaeological and palaeoecological journals. This is the first study of its kind in phytolith research but builds on other such reviews within archaeology such as Marwick & Pilaar Birch (2018) and Lodwick (2019).

This dataset was gathered from the same journals (see Table 2 for the selection of journals) and the same date range (2009-2018) as the dataset published by Lodwick (2019). This dataset is therefore complementary and allows comparisons to be made between macro-botanical remains and phytolith research in terms of data sharing and the general state of open science practices.

For the purposes of this article, I take a wider view of archaeobotany than Lodwick (2019:2), who defines archaeobotanical data as 'the quantitative assessment of plant macrofossils present within a sample from a discrete archaeological context'. It is suggested here that archaeobotany is the study of any plant remain found on archaeological sites and this could be found in many different forms (Pearsall 2019:3). Therefore, I believe phytolith analysis is one of the sub-disciplines of archaeobotany along with macro-botanical remains, palynology, anthracology, and other microscopic analyses such as aDNA, starch analysis and isotope analysis. Consequently, the data gathered by Lodwick (2019) will be referred to as macro-botanical remains rather than archaeobotanical remains to distinguish it from other botanical remains.

There are several key differences in the methodology used to collect the dataset presented here, compared to Lodwick (2019):

- 1) All papers containing primary phytolith data were included – this is not just archaeological studies but palaeoenvironmental and methodological research. It is important to review the full extent of data available for phytolith research as reproducible studies need to be conducted on methods as well as more applied research. Studies of macro-botanical remains in recent years are rarely methodological as it is a much more established discipline than phytolith research, and those in Lodwick's (2019) study are all archaeological data papers.
- 2) There is no data here concerning meta-analysis or data citation as has been included in Lodwick (2019). Again, these types of studies are few and far between in phytolith research due to the developmental stage of the discipline and the lack of standardisation of methods that makes collation of data from other research groups difficult. Consequently, this dataset will be compared only to the dataset of data sharing produced by Lodwick (2019).

- 3) The assessment of data sharing is specifically with reusability in mind. Therefore, how the data was recorded has some differences. First, how the data was presented/located was noted; for example, in a graph, table or excel spreadsheet. The same categories have been used apart from the addition of a category for graphs. Then, an assessment was made as to whether this data sharing could be used. A judgement was made that if the data was raw counts (and weights) and in a excel spreadsheet, csv file or repository, then another researcher would be able to reuse it. This was recorded as Yes or No in the database.
- 4) Other measures required for data reuse were also recorded as presence/absence (Y/N) that are specific to phytolith data such as pictures of phytolith morphotypes for identification verification, use of ICPN and inclusion of a full extraction method (whether from soils or plant material). These extra pieces of information, or metadata, are required with datasets to allow other researchers to fully review, reuse or reproduce the results/methods/research of others.

Another interesting methodological aspect of this research (even if it was imposed on me) was my lack of access to academic publications. Therefore, I completed this research with no access to an academic library and consequently all the articles have been gathered by Gold and Green open access, academic social media sites such as ResearchGate or Academia.edu, emailing authors and the help of academic colleagues. The state of openness of publications and the associated data was recorded and will be assessed as part of the results with further discussion at the end of the article. The fact that many researchers still do not have access to all research outputs due to paywalls of academic journals, and especially research that has been publicly funded, is an issue that must be addressed throughout the scientific community.

A full description of the methodology used here can be found in the accompanying data paper (Karoune 2020) along with the link to the full open access dataset.

Table 2: List of journals and categories used by Lodwick 2019 and also in this study.

Archaeobotanical journals	Archaeological science journals	General Archaeology	Cross-disciplinary Journals
Vegetation History and Archaeobotany	Archaeological and anthropological science	Antiquity	PLoS One
	Environmental Archaeology	Journal of Field Archaeology	Proceedings of the National Academy of Sciences
	The Holocene	Oxford Journal of Archaeology	
	Journal of Archaeological Science	Journal of Anthropological Archaeology	
	Journal of Archaeological Science reports	Journal of World Prehistory	
	Journal of Ethnobiology		
	Quaternary International		
	Journal of Wetland Archaeology		

Table 3: Comparison of the number of articles for macro-botanical remains and phytoliths in the period 2009 to 2018.

Name of journal	Number of articles for macro-botanical remains (Lodwick 2019)	Number of articles for phytoliths
Vegetation History and Archaeobotany	81	24
Archaeological and anthropological science	22	21
Environmental Archaeology	19	14
The Holocene	12	35
Journal of Archaeological Science	24	62
Journal of Archaeological Science reports	16	20
Journal of Ethnobiology	1	0
Quaternary International	18	115
Journal of Wetland Archaeology	6	0
Antiquity	12	6
Journal of Field Archaeology	4	1
Oxford Journal of Archaeology	1	0
Journal of Anthropological Archaeology	8	4
Journal of World Prehistory	1	0
PLOS One	11	25
Proceedings of the National Academy of Sciences	3	15
Total number of articles	239	341

3. Results

3.1. Primary data sharing

In total, 341 journal articles were found to contain primary phytolith data in the 16 journals sampled. This included 214 (63%) archaeological, 53 (16%) palaeoenvironmental and 74 (22%) methodological articles. Table 3 shows the number of articles per journal compared to Lodwick (2019). As suggested above, phytolith research is in a period of expansion and this can clearly be seen when compared to the number of macro-botanical articles in the same period (n=239). In the decade sampled, although there are a similar number of archaeological articles, there are more than 100 more articles containing primary phytolith data than macro-botanical data. The spread of articles in the journals sampled is different (see Figure 1); phytolith research was predominantly found in QI, JAS and The Holocene. Whereas primary macro-botanical data, was most prevalent in VHA and then JAS. There were no phytolith articles with primary data in the Journal of Ethnobiology, the Oxford Journal of Archaeology, the Journal of Wetland Archaeology or the Journal of World Prehistory; journals that had low occurrences of macro-botanical articles. These journals are excluded from the graphs presented here.

Overall, data sharing of primary phytolith data was 53%. Figure 2 shows the results of the location of data sharing in the articles grouped by journal. Therefore, 47% of the articles did not share primary data – this was a sum of the N and Graph results. Graphs were deemed to not be data sharing as it is very difficult to extract any reusable data from them. They were in fact the most frequently found format of data presentation in this dataset – making up 40% of all data shown. This is a large proportion of data to be unusable.

Figure 1: Chart showing a comparison of the number of journal articles in the current study with Lodwick (2019).

Figure 2: Chart showing the location of phytolith data in journals sampled.

Figure 3: Chart showing the percentage of articles per data location in macro-botanical remains (data taken from Lodwick 2019) and phytolith analysis.

Figure 4: Chart showing the location of phytolith data per year (2009-2018) in primary phytolith data publications.

This level of data sharing is very similar to the data sharing found for macro-botanical remains (56%, Lodwick 2019: 5). Overall, the location of data sharing in published articles in both sub-disciplines is similar as shown in Figure 3, although per journal there is a different pattern. For phytoliths, sharing of data was found to be better in VHA, JAS and then the more open access cross-disciplinary journals (PLOS ONE and PNAS). This is different from the results found for macro-botanical remains, which showed data sharing was better in the archaeobotany and archaeological science publications.

In Lodwick (2019), it was reported that only 52% of articles contained the full raw data. This is in comparison to 15% for the phytolith articles in this study that had raw data in any form (this was data in any location and could be raw counts, absolute presence or relative presence). However, reusable raw data (raw counts of all morphotypes identified in excel, csv files or a repository) was only 4%. It is also important to provide the number of fields counted on the slide and the weights recorded during processing. It is suggested by Zurro *et al.* (2016) that absolute and relative presence are the raw data for phytoliths but this may well negate the use of these datasets in the future if a new form of analysis is developed that requires the raw count information. It is suggested here that raw counts, slide fields, weights recorded during extraction and any transformations of this data are provided to allow the greatest opportunity of reuse.

It is also incredibly important that the data is in an interoperable form. To be counted as reusable data in this dataset, it was judged to be data that is attached as a supplementary file (.xlsx or .csv file) or deposited in an open repository. This means that a much more stringent measure of reusable data has been made in this dataset and therefore reflects a true measure of the reusability of phytolith data. This unfortunately is only 4% and therefore suggests poor data reusability in this dataset.

Graphs were found to be the most common way to supply data in this study, however, these data are not reusable, so they are deemed here as not being data sharing. Most graphs encountered were stratigraphic graphs, similar to those presented in palynological studies. They are a good way to present data in archaeological and palaeoecological studies as we are often looking for changes in our results over specific time periods. However, they are a form of presentation of data analysis and should be supplemented with data tables of raw data preferably in an open repository. Another issue often found with graphs in this dataset was poorly labelling, meaning it was unclear what they were showing, again impeding data sharing. Consequently, graphs should not be accepted by journals as a data sharing method. This is a stipulation that reviewers and editors should make clear to authors and therefore should be enforcing better practices of raw data sharing.

The second most common form of data sharing is in tables contained within the text of the article ($n = 129$). This presentation of data is hard to extract and is not citable as a separate dataset, therefore making it much harder for other researchers to find and then access. Mistakes can also be made when copying data in this way and therefore this is not the most effective way to share data.

The next most common forms of data sharing are excel files and pdfs. The use of excel files has increased in recent years and the use of supplementary files in general (.docx, .pdf and .xlsx) have increased from 2015 by approximately 20%. This can be seen in Figure 4, which shows the location of phytolith data per year. The use of supplementary files has increased in journals that have recently updated their data availability statements such as PLOS One and PNAS, but this does not mean the data provided in these files is reusable. Word files and .pdf files may contain graphs and tables of data that are again hard to copy and prone to typing errors. Excel and Word files are proprietary file formats so not usually recommended as being compatible with open science practices, but users can open these files easily in free to use software such as Google Sheets or Open Office. Therefore, it is suggested here that Excel files are a good way to share

data, especially if the file is deposited in an open repository for long term storage. It is also easy to transform an Excel file into a .csv file and deposit both in the same place.

Although the use of supplementary files is gradually increasing, it is still not the best method of data sharing for a number of reasons. Most supplementary files in this study are not accessible because of pay walls. If the journal article is not open access, and this was the case for 86% of the articles in this study, then the supplementary files are usually also not open. Some journals do allow access, such as VHA in this study. However, it is rare that data in supplementary files is accessible if the article is not open access, therefore limiting the re-use potential of these data.

It was also found that supplementary files were often labelled inadequately to tell readers what information the file contained. This was despite most journals stipulating in their author guidelines that the title of the file must include a brief description. These files were also poorly referred to in the text of the articles and therefore to find out what the file contained it had to be downloaded. This problem hinders the findability of the data and metadata contained in articles.

Unfortunately, there was only 1 article (Lancelotti *et al.* 2017) in the whole study that used a repository to archive their data. This is the same finding as Lodwick (2019) for macro-botanical remains and demonstrates a lack of awareness of the need for long term storage and accessibility of data in archaeobotany. It demonstrates that much work is needed to address this issue and training must be put in place to improve data sharing across the discipline.

An issue found here, which corresponds with previous data sharing research (Lodwick 2019; Marwick & Pilaar Birch 2018), is the mismatch between journal data policies and data availability in articles. The current data policies of the journals in this study (see Table 1) range from having no policy (for example, Antiquity and the Journal of Field Archaeology) to having a very strict policy that states authors must make material, data and associated protocols available and should follow FAIR principles (PNAS 2018). It has been found here that there is no correspondence between a stricter data policy and better data sharing. Journals with a high percentage of articles providing no data or having only data presented in graphs were not just the journals with no data policies. They included JAS Reports and Quaternary International (the journal that contained the largest number of articles in this study), whose policies have required authors to share research data to facilitate reproducibility and data reuse since approximately 2013 (Marwick and Pilaar Birch 2018:131). VHA has seen a slight increase in data sharing after 2016 and this is when its current policy was updated to include an encouragement to share data in repositories. The one example of an article using a repository in this study comes from this journal in 2017, however this change in policy cannot be seen as having a major impact, as the authors concerned had a number of articles in other journals in this dataset that also followed good data sharing practices. Therefore, it seems the authors knowledge and experience influenced data sharing in this particular case rather than the journals data availability policy.

The open access journals PLOS One and PNAS, that currently have the strictest data policies, do have lower levels of no data and graphs in their articles, and the former has seen an increase in supplementary files since 2015. However, sharing of raw data in both journals is still low and PNAS has no articles that contain reusable data as defined in the stricter measure used in this dataset. This study confirms the findings that enforcement of data policies is currently poor. This is an issue that our discipline as a whole needs to address if we are to implement better data sharing practices. It may, in fact, be better for researchers to take personal responsibility for data sharing by routinely using data repositories for all their data rather than relying on journal editors to enforce data availability policies.

Figure 5: Pie charts of percentages for articles in the sample (recorded as yes or no answers): a) Reusable raw data; b) Pictures; c) Open access; d) ICPN use; e) Full method.

3.2. Metadata

Figure 5 presents data gathered for the purposes of assessing the sharing of metadata. Providing pictures for confirmation of identifications, using ICPN to standardise the nomenclature of phytolith morphotypes and providing a fully described methodology are important details that need to be included in publications to assist reuse of phytolith data. A detailed discussion of the importance of these aspects can be found in the next section. Most articles (74%) provided some pictures or diagrams to present identifications of the morphotypes found. However, not all of these articles provided pictures of all of the morphotypes presented in the data. In printed journals, it is unfeasible to include all pictures in the text and therefore they must be included in supplementary files. However, this is not an issue for online journals where inclusion of pictures is only limited by file size.

Likewise, the majority of articles (70%) gave a full description of the methods used; this might have been a description in the text or reference to one published method. All articles did contain a methods section, but there were several problems encountered that made the author discount the information. Authors need to consider that the reason for providing a method in a scientific paper is for the peer reviewer to assess the validity of the method and for other researchers to be able to replicate experiments to check for reproducibility. Some of the methods

provided were too brief and even an experienced phytolith analyst would not be able to perform an identical experiment due to the guess work needed to fill in the gaps. Other authors referenced several methodological papers to suggest that their work used a combination of these different methods. Again, this is not stating a clear method that could be followed. Ideally, the method should be given briefly in the text and then a full procedure included in the supplementary information.

A more disappointing finding was that only 51% of the articles used ICPN. Some of these articles expressed that ICPN was only used when possible therefore the full description of all morphotypes using ICPN is likely to be lower. It is an important step to find out the level of use of ICPN in the phytolith community and it is frustrating to find there is not a better level of implementation by researchers. The standardisation of terminology is crucial for the reuse of datasets, otherwise data becomes incredibly difficult to collate for larger meta-analyses.

3.3. Open access

The initial findings for open access are low in this dataset. Only 14% of articles are open from their initial date of publication (Gold open access). However, this is not the end of the story for open access because I was able to access the rest of the articles through other means to complete the dataset even though I had no access to an academic library. A further 3% of articles were made open access after six months of publication, which is the standard procedure for PNAS. 54% of articles were found on academic social media sites such as ResearchGate (RG) or Academia.edu. Another 2% of the articles were found in institutional repositories. Therefore, in total 73% of the articles were freely available somewhere on the internet. For the other 27% of articles, they were procured through requesting articles through RG, emailing the corresponding authors or they were kindly sent to me by colleagues. These findings are encouraging as it currently indicates that most articles are accessible. The academic social media sites, although not seen as best practice in open science for long term publication or data storage, are incredibly useful for providing access to researchers who have no other access to publications. However, there is still some way to go before all academic publications are freely available to all at the point of publication.

4. Steps forward

This review demonstrates there is still much to improve in phytolith research in terms of open science practices. While it has been found that general data sharing needs improvement, and therefore is consistent with other archaeological disciplines, in fact a more refined assessment indicates that overall reusable data sharing is poor. The assessment of providing metadata (pictures, ICPN use and methods) also demonstrated a need for improvement. These aspects are equally as important in publications, otherwise validation of research by stringent peer reviews, reuse of data for reproducible studies and extension of research cannot be conducted. Initial open access to publications was also poor with gold open access articles being uncommon in this study. However, the sharing of articles on academic social media sites and institutional repositories increased the overall level of open access.

What can be done to move forward from this point? Figure 6 is a summary of the steps set out here of how to improve open science practices in phytolith research. As a discipline we must start planning new projects with open science at their heart and strive to always make our publications open. We must also improve our data sharing, and this can be done by adhering more closely to the FAIR principles.

Some of these steps are similar to suggestions laid out by Zurro *et al.* (2016) and it is clear from this review that their guidelines are not being considered by many researchers. The steps set out here are only first steps to what I hope will become a growing awareness of the need for change, and I encourage colleagues to discuss and work towards refining a new set of guidelines as a community.

Figure 6: Summary of proposed steps for improving open science practices in phytolith research

4.1. Open project planning

If we are to improve open science practices throughout our discipline, this must be planned. It takes extra time and consequently more money to work openly, and therefore initial project planning needs to take this into account. As mentioned in the introduction, a data management plan is often required by funders, and therefore this is a good opportunity to think about how data will be handled in an open manner. It is imperative to consider where data will be stored and especially its long-term management so that any data being produced will conform to the storage guidelines of your chosen repository. The cost of data storage varies depending on the size of the project and these services, such as the Archaeological Data Service (ADS) or Open context, can be consulted prior to making funding applications to get advice and quotes for data management. A good example of guidelines on data storage and management can be found on the ADS website (ADS 2020).

If cost is an issue or you want to control your project yourself, the ideal way to manage a collaborative open project would be to use a free online platform, such as Open Science Framework (OSF) or Open AIRE, to store every element of a project from its initiation to publication. These platforms allow you to keep some work private and publish other parts openly including pre-prints and datasets. They assign a citable DOI when any aspect is opened publicly that can be used to make the project more findable and will increase citation of your work. They are an efficient way to work with other colleagues and can be used to link together different applications such as google docs and git hub. Increasing awareness and training of these online tools will aid their implementation in phytolith research projects, which will ultimately lead to their increased use.

A new development in open science publication that should become part of the planning process is registered reports. These are publications that describe the project design and

emphasize the importance of research questions and robust methodology (Centre for Open Science 2020). These articles are to be written after the gathering of initial ideas and design development, thus recording hypotheses. This can then be referred to in later publications, therefore adhering to a hypothetico-deductive model of scientific method. In addition, authors gain comments and peer review from the wider community that can be used to improve their initial project design and ultimately leads to better scientific design. It has other benefits such as increasing the number of publications and therefore improving citation for the project. This will ultimately raise the profile of new research and may also lead to new collaborations.

4.2. Open access publications

It has been found here that the overall open access to publications was needing improvement but with added access due to academic social media sites and informal arrangement such as emailing authors, access could be gained to the majority of publications in this dataset. However, there is much to improve, and informal access is too hit and miss to be an effective way to access the publications needed for efficient research. It is especially important to improve the publication access for those who do not have access to university libraries such as those returning to research, free-lancers or those between contracts. It is also imperative that colleagues in developing countries have the same access to research. Some journals do give discounts or free access to developing countries, which should be further encouraged.

What can we as researchers do to improve access to our publications? There are several ways that this can be done, and it does depend on who you work for. Always using Gold open access should be the route for those that have research funding to pay for the APC or colleagues that work at organisations that have open access deals with publishers. This option means the article is immediately accessible to all.

However, what do other researchers do to improve access to their publications? There is an increasing amount of funding available to pay for APC's through open access journals such as Internet Archaeology and organisations such as The Association of Environmental Archaeology. You can ask editors for a fee waiver if you are not in the position to pay for the APC and it is worth researching the cost of Gold open access as fees vary considerably meaning some journals have very low APC's such as Journal of Open Archaeology Data (just £100).

You can also use Green open access in which the journal allows you to self-archive a pre-print and/or post-print in an open repository. The accepted manuscript will have an embargo period put on it by the journal, but other versions can be made public at any time. This can be done through the academic social media sites or it is better practice to use an open repository, such as university repositories, Zenodo or Open Science Framework to name just a few. You could even use a specific pre-printer server, such as earthArxiv or bioRxiv. All these options provide a citable DOI for each version of the article that is deposited.

There is also the recently formed Peer Community in Archaeology (PCI Archaeology) that can be used instead of a traditional journal. After you deposit your pre-print, they will peer review the article and then recommend it. This can aid improvement of articles for early career researchers by receiving more feedback before submitting to a traditional journal. But in a sense, there is no need to publish the work anywhere else as it has been peer-reviewed and recommended, which completes the research cycle. This new type of publication format is likely to be popular with those that do not have to submit publications towards academic research output metrics.

Phytolith researchers are also encouraged to make publications available to the IPS. As a society, they maintain a bibliography of phytolith publications and can help members to access any publications that are hard to obtain. Although membership does come at a small cost, it is much cheaper than paying for access to all the journals that phytolith research is published in.

This is a positive element of openness within the phytolith community that helps to aid inclusiveness for all researchers and can be accessed through the society website.

4.3. FAIR principles

Bringing your data in line with the FAIR principles (Wilkinson *et al.* 2016) takes significant consideration. The FAIR principles do not just involve data but also metadata and each disciplines application of these principles will be unique. The basic principles are the same but there are separate requirements for each discipline because of how their data is created and the specific type of data that is produced. It is therefore important that each discipline makes a community decision about how to approach this matter so that data collection and storage is effective.

So, what does FAIR phytolith data look like?

Findable - To be findable as separate entities, the data and metadata must be in an online repository with a DOI. It is therefore not enough to put data into the supplementary material attached to a journal article. Using a repository is not yet a requirement of all journals but should be the best practice for our discipline to improve findability of datasets.

The IPS have constructed a database of modern and archaeological samples that is accessible on their website as a map to show the samples location and also a list to link each entry to a publication. However, being linked to published articles does not mean the data is accessible and therefore this could be improved if all of these datasets were put into open repositories and the links added to the database. This is a good example of how the implementation of this best practice would aid the improvement of phytolith data sharing.

To improve findability further, researchers can publish a data paper about each dataset. These articles describe the dataset, the methods used to construct it and where it can be found. The journal provides a template and they are therefore very simple to write. Consequently, they do not add much extra work but have great benefit if you want your data to be easy to find. There is a data paper available for this dataset in the Journal of Open Archaeology Data (Karoune 2020).

Accessible - To improve accessibility, it is ideal that data is open, but to be consistent with FAIR principles it does not have to be. Some researchers prefer to keep their datasets private and only make them available upon request. Lodwick (2019) has suggested this may be due to a number of reasons in macro-botanical research such as issues of commercial confidentiality in developer-funded projects, the need to publish data in excavation monographs and the need for more funds, training and knowledge of digital archiving resources. More awareness and training in data sharing best practices will hopefully mean this ad-hoc data sharing becomes less prevalent and data in open repositories becomes the norm.

However, whether your data is open or not, the metadata must be fully accessible. Providing the full metadata is important if others are to review or reuse your data and again the type of metadata needed is unique to each archaeological discipline. So, what metadata should be provided for phytoliths? Zurro *et al.* (2016) have provided much guidance on this matter. They have suggested sample context information, sampling strategy, taphonomic indicators (such as diagenesis, pH, etc.), full methodology for extraction, counting and data analysis, information on identification and preferably photographs of the identified morphotypes. All the journals in this study would allow this information to be provided as supplementary information.

It is important that full details are given for all metadata since, if reproducible studies are the aim of sharing data, this cannot be completed without such information as a step by step method for extraction. This type of full procedural method was rarely provided with the articles in this dataset, although authors often referred to the original methodological articles.

A more recent addition to the policies of some journals is the requirement to write a data availability statement. This allows the author to explain where their data can be found, whether it is provided with the article as supplementary files, by request, or by the link provided. It is likely to improve findability and therefore access to data. It should be adopted by all journals.

Interoperable - The interoperability of data is key to its potential reuse. It is strongly recommended here that all data is deposited in an open repository in several formats rather than in a supplementary data file. This is due to the potential findability and accessibility problems with supplementary files; poor labelling of files, lack of access due to paywalls and use of formats such as pdf or word docs that increase the likelihood of copying errors. Several different formats of the data can be added to a repository such as an excel and .csv file. There is also the potential to deposit software, scripts and code written for analysis along with datasets, although few phytolith researchers are currently using R stats.

Another aspect of interoperability is standardisation, and this is an issue in all aspects of phytolith research that still needs to be comprehensively addressed. The IPS has already started work on a few important areas of standardisation, taxonomy (ICPN) and morphometrics, although it has been found here that the implementation of ICPN is not widespread. It is important that the terminology used to identify different morphotypes is standardised with the use of ICPN to aid the reuse and collation of datasets. This can also be aided by the provision of pictures with all data for purposes of standardising identification.

In addition, an important resource available to phytolith researchers to aid standardisation of identification and use of ICPN is PhytCore (Albert *et al.* 2016). This is an open online database of phytolith images from modern reference material and archaeological samples that aims to stimulate discussion and aid the development of a common nomenclature. It is an incredibly useful tool for new researchers to develop their identification skills and more experienced researchers to calibrate their identifications to a set standard. The use of this tool should be encouraged by phytolith educators to increase the use of ICPN and therefore the interoperability of datasets.

Methods of sampling and extraction vary widely in phytolith research and currently make comparisons of different datasets problematic. Researchers have found some sediments, such as oxisols rich in iron oxides and hydroxides, do give different recovery rates but most of the other studies suggest similar recovery rates (Zurro *et al.* 2016: 113). However, I think more work is needed and in particular a comprehensive study of extraction methods to assess the differences in recovery from all methods currently utilised by phytolith researchers. This must be completed with enough replication by different research groups to prove that the results are valid. The results could then be confidently used to make suggestions on how to proceed with standardisation, whether this is the choosing of a particular standardised method, partial standardisation such as the addition of the calculation of the AIF (Acid insoluble fraction) in all extractions to add a standard reference unit of quantification (Zurro *et al.* 2016) or the development of a standard way to make adjustments to data to allow for any differences produced by the various extraction methods.

Reusable – The issue of raw data has been addressed above, with the conclusion that the actual raw data for phytoliths extracted from sediments is the raw counts of the number of different morphotypes, number of fields counted and extraction weights. Any other data provided has transformed this data into another form such as absolute presence or relative presence (percentage). For the longevity of a dataset, it is imperative that this true raw data is provided, which can be accompanied with any other data calculated by the researcher. Other data that should be provided as raw counts is data from modern reference collections (from plant material and sediment samples) and identification studies using morphometrics. These types of studies both tend to only provide summarised data such as presence/absence or ranges and averages. As phytolith research is still in a methodological developmental phase the ability to reuse data

from these new studies, whether the study of new plant taxa or a new geographic area, is incredibly important for validation of new methods. It also allows other researchers to build on new research and take the field into new areas of development.

A problem encountered during the gathering of data for this study was the poor labelling of datasets. This has already been mentioned in relation to the need for clear labelling of supplementary files so that it is easy for the reader to know what is contained in each file, however there was also issues with the labelling in tables and spreadsheets. It is very easy as a researcher to become over familiar with your own dataset and therefore use short cuts when labelling the headings and tabs. However, when you are providing a dataset for other people to view, and potentially use, you need to make sure that enough information is provided in the headings of tables/columns so that it is clear what the data is. There were many tables and spreadsheets that did not distinguish between absolute and relative presence, did not provide the units for measurement and did not make sample numbers or names clear. This is where good metadata plays a role in the reusability of a dataset. Keys need to be provided that explain any codes used and care should be taken when adding any headings.

Finally, all data provided should have a license so that others know what kinds of reuse are permitted. Licenses can be completely open, such as Creative Commons 0 (CC0), and usually open repositories require datasets to have these types of licenses. However, there are also lots of other variations that can be used to vary the user's rights to copy, distribute and make use of your data.

5. Conclusions

There are a number of positive finds from this first assessment of open science practices in phytolith research, such as the formative work on standardisation by the IPS that has resulted in the increasing use of the ICPN and the majority of researchers providing pictures to aid validation of identifications. Other work by this society, such as the development of a sample database and guide for morphometric analysis, has the potential to be substantially strengthened by using open science practices. However, there is clearly a plethora of work to tackle in terms of data sharing and open access. Raising awareness of the problems found here, associated with lack of open access to publications, poor data sharing and particularly sharing of raw data, is the first step in improving the situation. Training our community in all aspects of open science is vital to change attitudes to data sharing, which would result in a growing awareness of how to change current practices and develop a common strategy. Increasing knowledge and confidence in using open science tools will lead to researchers taking ownership of their data sharing and increase the use of open repositories for all research data, metadata and publications. We cannot wait to be told how to share data and publications by journal editors; we need to make the decisions about what is our own best practice. Only by working collaboratively can phytolith researchers address the issues raised in this study, and we must strive as a community to improve outcomes for all researchers in terms of increased validation and reproducibility of our studies to build more robust methodologies. We need to nurture an ethos of collaboration so that our discipline can move forward in a more open transparent manner.

Data Accessibility

There is a data paper to accompany this research project. The DOI for this is:

The full dataset is available at the following address as a csv file: <https://osf.io/8p3bn/>

Acknowledgements

The author would like to thank those in the phytolith community that answered emails to request their articles and supplementary files. A special thanks goes to Dorian Fuller, Ruth Pelling and Rand Evett, who helped to finish this dataset by retrieving those hard to access articles and data from behind paywalls. I am grateful to Juanjo Garcia-Granero and Arlene Rosen for commenting on early versions of the text.

Funding Information

No funding was provided for this research.

Competing Interests

The author has no competing interests to declare.

References

- Albert, RM., Ruiz, JA., and Sans, A. 2016. PhytCore ODB: A new tool to improve efficiency in the management and exchange of information on phytoliths. *Journal of Archaeological Science*, 68: 98-105. DOI: <https://doi.org/10.1016/j.jas.2015.10.014>
- Archaeological Data Service. 2020. Guidelines for depositors. Version 4.0. July 2020. Available at <https://archaeologydataservice.ac.uk/advice/guidelinesForDepositors.xhtml> [Last accessed 12th July 2020].
- Ball, T., Davis, A., Evett, RR., Ladwig, JL., Tromp, M., Out, WA., and Portillo, M. 2016. Morphometric analysis of phytoliths: recommendations towards standardization from the International Committee for Phytolith Morphometrics. *Journal of Archaeological Science*, 68:106-111. DOI: [dx.doi.org/10.1016/j.jas.2015.03.023](https://doi.org/10.1016/j.jas.2015.03.023)
- Centre for Open Science. 2020. Registered Reports: peer review before results are known to align scientific values and practice. Available at <https://www.cos.io/our-services/registered-reports> [last accessed 12th July 2020].
- Christensen, G., Dafoe, A., Miguel, E., Moore, DA., and Rose, AK. 2019. A study of the impact of data sharing on article citations using journal policies as a natural experiment. *PLOS ONE*, 14(12). e0225883. <https://doi.org/10.1371/journal.pone.0225883>
- FOSTER, 2018. What is Open Science? Open science training courses. Available at: <https://zenodo.org/badge/DOI/10.5281/zenodo.2629946.svg> (last accessed 23rd June 2020).
- Hart, TC. 2016. Issues and direction in phytolith analysis. *Journal of Archaeological Science*, 68: 24-31. DOI: <https://doi.org/10.1016/j.jas.2016.03.001>
- Hodson, MJ., Song, Z., Ball, TB., Elbaum, R. and Stuyt, E. 2020. Editorial: Frontiers in Phytolith

- Research. *Frontiers in Plant Science*, 11(454): 5-7. DOI: 10.3389/fpls.2020.00454
- Iriarte, L. 2003. Assessing the feasibility of identifying maize through the analysis of cross-shaped size and three-dimensional morphology of phytoliths in the grasslands of southeastern South America. *Journal of Archaeological Science*, 30: 1085-1094. DOI: [https://doi.org/10.1016/S0305-4403\(02\)00164-4](https://doi.org/10.1016/S0305-4403(02)00164-4)
- Itzstein-Davey, F., Taylor, D., Dodson, J., Atahan, P. and Zheng, H. 2007. Wild and domesticated forms of rice (*Oryza* sp.) in early agriculture at Qingpu, lower Yangtze, China: evidence from phytoliths. *Journal of Archaeological Science*, 3: 2101-2108. DOI: <https://doi.org/10.1016/j.jas.2007.02.018>
- Kansa, E. 2012. Openness and archaeology's information ecosystem. *World Archaeology* 44(4):498-520. DOI: 10.1080.00438243.2012.737575
- Kansa, EC. and Kansa, SW. 2013. Open Archaeology: we all know that a 14 is a sheep: data publications and professionalism in archaeological communication. *Journal of Eastern Mediterranean Archaeology & Heritage Studies*, 1(1): 88-97. DOI: <https://doi.org/10.5325/jeasmedarcherstu.1.1.0088>
- Kansa, SW., Atici, L., Kansa, EC. and Meadows RH. 2020. Archaeological analysis in the information age: guidelines for maximising the reach, comprehensiveness and longevity of data. *Advances in Archaeological Practice*, 8(1): 40-52. DOI: 10.1017/aap.2019.36
- Kansa, SW. and Kansa, E. 2014. Data publishing and Archaeology's information ecosystem. *Near Eastern Archaeology* (NEA), 77(3): 223-227. DOI: <https://doi.org/10.7183/0002-7316.79.1.5>
- Karoune, E. 2020. Data from assessing open science practices in phytolith research. *Journal of Open Archaeology Data*, submitted and will complete reference when published.
- Lancelotti, C., Ruiz-Perez, J. and Garcia-Granero, JJ. 2017. Investigating fuel and fireplaces with a combination of phytoliths and multi-element analysis; an ethnographic experiment. *Vegetation History and Archaeobotany*, 26: 75-83. DOI: <https://doi.org/10.1007/s00334-016-0574-y>
- LeFebvre, MJ., Brenskelle, L. Wiczorek, J., Kansa, SW. and Kansa, EC. 2019. ZooArchNet: Connecting zooarchaeological specimens to the biodiversity and archaeology data networks. *PLOS ONE*, 14(4): e0215369. DOI: <https://doi.org/10.1371/journal.pone.0215369>
- Lodwick, L. 2019. Sowing the seeds of future research: data sharing, citation and reuse in Archaeobotany. *Open Quaternary*, 5(7): 1-15. DOI: <https://doi.org/10.5334/oq.62>
- Madella, M., Alexandre, A., and Ball, T. 2005. International code for phytolith nomenclature 1.0. *Annals of Botany*, 96: 253-260.
- Madella, M. and Lancelotti, C. 2012. Taphonomy and phytoliths: a user manual. *Quaternary International*, 275: 76-83. DOI: 10.1016/j.quaint.2011.09.008
- Marchetti, N., Angelini, I., Artioli, G., Benati, G., Bitelli, G., Curci, A., Marfia, G. and Roccetti, M.

2018. NEARCHOS. Networked Archaeological Open Science: advances in archaeology through field analytics and scientific community sharing. *Journal of Archaeological Research*, 26: 447- 469. DOI: <https://doi.org/10.1007/s10814-017-9112-4>
- Marwick, B. et al. 2017. Open Science in Archaeology. *The SAA Archaeological Record*, 17(4):8-14. DOI: <https://doi.org/10.31235/osf.io/72n8g>
- Marwick, B. and Pilaar Birch, SE. 2018. A standard for the scholarly citation of archaeological data as an incentive to data sharing. *Advances in Archaeological Practice*, 6(2): 125-143. DOI:10.1017/aap.2018.3
- Masuzzo, P. and Martens, L. 2017. Do you speak open science? Resources and tips to learn the language. *PeerJ Preprints*, 5: e2689v1. DOI: <https://doi.org/10.7287/peerj.preprints.2689v1>
- Neumann, K., Stromberg, CAE., Ball, T., Albert, RM., Vrydaghs, L., and Scott Cummings, L. 2019. International Code for Phytolith Nomenclature (ICPN) 2.0. *Annals of Botany*, 124: 189-199. DOI: 10.1093/aob/mcz064
- Pearsall, DM. 2019. *Case studies in Paleoethnobotany*. New York: Routledge.
- Piperno, DR. 2006. *Phytoliths: a comprehensive guide for archaeologists and palaeoecologists*. Oxford: Altamira Press.
- PNAS. 2018. Editorial and Author policies – updated with FAIR principles statement August 2018. Available at <https://www.pnas.org/page/authors/journal-policies> [Last accessed 21st July 2020].
- Schmidt, SC. and Marwick, B. 2020. Tool-driven revolutions in Archaeological Science. *Journal of Computer applications in archaeology*, 3(1): 18-32. DOI: <https://doi.org/10.5334/jcaa.29>
- Shillito, L-M. 2013. Grains of truth or transparent blindfolds? A review of current debates in archaeological phytolith analysis. *Vegetation History and Archaeobotany*, 22: 71-82. DOI: 10.1007/s00334-011-0341-z
- Wilkinson et al. 2016. The FAIR guiding principles for scientific data management and stewardship. *Scientific data*, 3: 160018. DOI: 10.1038/sdata.2016.18
- Womack, RP. 2015. Research data in core journals in Biology, Chemistry, mathematics, and Physics. *PLOS ONE*, 10(12): e0143460. DOI: 10.1371/journal.pone.0143460
- Zurro, D., Garcia-Granero, J.J., Lancelotti, C. and Madella, M. 2016. Directions in current and future phytolith research. *Journal of Archaeological Science*, 68: 112-117. DOI: 10.1016/j.jas.2015.11.014