



# Assessing Open Science Practices in Phytolith Research

RESEARCH PAPER

EMMA KAROUNE 

]u[ubiquity press

## ABSTRACT

Open science is an integral part of all scientific research, but the extent of open science 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. The use of phytoliths in archaeobotanical and palaeoecological studies has been increasing in recent years resulting in an upsurge in publications. The aims of this article are to assess open science practices in phytolith research 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). This study builds on similar studies conducted for zooarchaeology (Kansa *et al.* 2020) and macrobotanical remains (Lodwick 2019). This study shows that 53% of papers shared data in any format but only 4% of papers contained reusable data, 74% included some pictures of phytolith morphotypes for identification purposes, 69% had a fully described method, 47% used the International code for phytolith nomenclature (ICPN 1.0) and only 13% of articles were open access. Steps forward are then proposed, including planning for open projects, making more articles openly accessible and implementing the FAIR data principles, 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.

CORRESPONDING AUTHOR:

**Emma Karoune**The Alan Turing Institute, GB  
[ekaroune@googlemail.com](mailto:ekaroune@googlemail.com)

KEYWORDS:

Phytolith; Micro-botanical remains; Open Data; Open Science; Open Access Archaeobotany; Palaeoecology

TO CITE THIS ARTICLE:

Karoune, E. 2022. Assessing Open Science Practices in Phytolith Research. *Open Quaternary*, 8: 3, pp. 1–18. DOI: <https://doi.org/10.5334/oq.88>

## INTRODUCTION

### 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 plants (Madella & Lancelotti 2012). Monosilicic acid ( $H_4SiO_4$ ) in groundwater is absorbed by plant roots and is eventually deposited as solid silica dioxide ( $SiO_2$ ) in and between plants cells, forming distinctive shapes (morphotypes). These shapes can be used to identify plant tissues and also plant taxa to different taxonomic levels. When plants die, phytoliths become deposited naturally in soil or, if the plant has been exploited by humans, in archaeological layers. Although some research suggests phytoliths have in situ deposition (Iriarte 2003; Itzstein-Davey *et al.* 2007; Piperno 2006), the movement of phytoliths before they are deposited (transport) and once deposition has taken place is still under debate and needs further investigation (Madella & Lancelotti 2012; Shillito 2013:79). In any case, phytoliths are preserved over long time periods and can therefore be readily found in archaeological and palaeoecological samples. Consequently, phytoliths are used to address research questions in combination with, or instead of, other plant remains that do not preserve 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 (Hart 2016: fig 1). Archaeological phytoliths are examined from deposits using a variety of methods such as 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). Phytoliths are also extracted from ecofacts and artifacts such as dental calculus, coprolites, pottery and grinding stones (García-Granero *et al.* 2021, Jovanović *et al.* 2021, Shillito 2020). Applications in palaeoecology include reconstructing palaeoenvironments and palaeoclimates from modern reference plant material and core samples (Watling *et al.* 2020, Yost *et al.* 2020). The newer applications of carbon-14 dating and isotopic analysis of phytoliths are also being explored (see Zuo & Lu 2019 for a recent review).

With increased research comes the difficulty of how to standardise methods, share data and validate analyses (Cabanes 2020, Shillito 2013, Zurro *et al.* 2016). Phytolith research currently has less standardisation than closely-related disciplines such as 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 three categories – observational, experimental and computational data (Marwick & Pilaar Birch 2018). This wide-ranging data being produced must be made accessible and reusable so that other researchers can review, adapt, apply and collate their colleagues work for greater validation and sustainability of research.

The cost of phytolith research is high as it involves establishing a laboratory with large high-cost laboratory equipment and processing and analysing of phytolith samples takes many hours. The majority of phytolith research is conducted by large research institutions such as Universities and Museums as the costs involved often restrict its use in the commercial sector. Much of the phytolith research conducted concerns tropical regions where organic remains are often poorly preserved and therefore the higher research costs of phytolith research can be justified to enable archaeobotanical and palaeoecological questions to be adequately addressed. However, ensuring longevity of phytolith datasets by making them more openly available and easier to reuse will improve the cost effectiveness of phytolith research.

In the UK, where macro-botanical remains and pollen are used routinely to address the botanical elements of archaeological and palaeoecological studies, phytolith research still has much to offer. The use of phytolith analysis in commercial archaeological units is rare in the UK and hindered by a lack of comprehensive reference collections for the British Isles, which comes down again to the cost of establishing such a collection. Therefore, if we want to use phytolith research in new geographical locations, it becomes even more important to address issues of sustainability of data produced to justify the extra cost of analyses.

The types of phytolith research projects that are currently conducted determines that these phytolith studies, and the data produced, are rarely seen in commercial grey literature, apart from university dissertations, and are predominantly published in journal articles. It is therefore important that issues surrounding data sharing and research accessibility are comprehensively addressed to achieve the most value out of this current surge in research currently contained within journal articles.

The International Phytolith Society (IPS) have started to address some of the issues of data 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 publication of raw data (absolute and relative presence). The suggestions in Zurro *et al.*'s (2016) article are a much more holistic approach, but again, the extent to which these suggestions have been implemented by phytolith researchers needs to be assessed to move phytolith research into the age of open science.

### WHY IS OPEN SCIENCE SO IMPORTANT FOR PHYTOLITH RESEARCH?

Open science is more than just opening up scientific research. It is a multi-faceted movement to improve science based on four pillars (Masuzzo & Martens 2017): open data, open methods (including code), open access publications and open reviews. Open Science can be seen as a shift in research culture to a model that embraces practices that aid reproducible research and promote greater equity, diversity and inclusion in science. At every stage of the research lifecycle there is the opportunity to make work open. By embedding this approach in all scientific research, researchers can increase public trust in research, support scrutiny and validation, enable reuse and drive innovation (FOSTER 2018). Bringing some or all of these open practices into any scientific study will increase the sustainability of that research.

This article is coming out of a need for reform in archaeology to embed open science practices more deeply throughout the discipline to aid validation through reproducible research and greater sustainability. Kansa (2012) suggested that 'open access and open data lead to greater resilience for our profession but the open science approach needs more champions and remains at the edges of archaeological practice'. Therefore, we need to start to investigate how these practices are currently being applied in the many different archaeological and related disciplines.

Over the past decade there has been growing awareness of open science practices and a recent article summarising open science in archaeology (Marwick *et al.* 2017) sought to encourage these practices by highlighting their benefits to the archaeological community. These benefits include: open access publications achieving increased citations and being more accessible for students and non-academic collaborators; open data makes our own past research likely to be better documented and easier to reuse; and

open methods improve reproducibility and help us to publish our scientific workflows more easily. Other recent articles focus on computational archaeology (Schmidt & Marwick 2020), data sharing in zooarchaeology (Kansa *et al.* 2020) and innovative open field archaeological strategies (Marchetti *et al.* 2018).

Archaeological related studies in phytolith research are part of the wider discipline of Environmental Archaeology and there have been limited reviews of open science in this field. Zooarchaeologists are very much taking the lead (Kansa *et al.* 2020; Kansa & Kansa 2013, 2014; Lefebvre *et al.* 2019) with a focus on data sharing. 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 macro-botanical data, and much of this data was unusable. At present, the state of data sharing and other open science practices is not known in phytolith research and therefore there is a need to understand 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 validation of research, increased citation and increased research opportunities through collaborative projects.

In addition to journal requirements for open science practices, funding bodies such as the UK Research and Innovation (UKRI) and the European Research Council, that fund a large proportion of archaeological research in the UK and Europe, now require data management plans 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 management of data and block funding for open access has been

JOURNAL TITLE	PUBLISHER	DATA AVAILABILITY POLICY (ALL ACCESSED JUNE 2020)
Vegetation History and Archaeobotany	Springer	<b>Encourages</b> 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	<b>Encourages</b> authors, where possible and applicable, to deposit data that support the findings of their research in a public repository.
Environmental Archaeology	Taylor and Francis	<b>Encouraged</b> to supply data – list of positives given. Supplementary information to be put in Figshare.
The Holocene	SAGE publishing	<b>Acknowledges the need</b> 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 <b>requires</b> 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 – <b>requires</b> .
Journal of Ethnobiology	Society of Ethnobiology	No data guidance.
Quaternary International	Elsevier	Same as JAS above – <b>requires</b> .
Journal of Wetland Archaeology	Taylor and Francis	Authors are <b>encouraged</b> 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 – <b>requires</b> .
Journal of World Prehistory	Springer	<b>Encourages</b> authors, where possible and applicable, to deposit data that support the findings of their research in a public repository.
PLOS One– Open access journal.	(PLOS)	<b>Required</b> 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.	(PNAS)	Authors <b>must</b> make material, data and associated protocols including code and scripts available to readers of publication – from 2011. – Should follow FAIR (this part of the statement was added first in 2018). – Deposit data in community approved public repositories. Pre-2011 – authors <b>encouraged</b> to deposit data and use Supplementary information.

**Table 1** Data availability requirements of journals sampled in this study.

provided by the UKRI to research institutions to pay for article processing charges (APC) to enable Gold access publications for instant free publications. The new UKRI open access policy released in 2021, requires immediate article access either through Gold or Green access routes. Therefore, it is in every researcher's 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.

The aim of this study was to obtain a snapshot of where phytolith research is currently, in terms of open science practices, 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.

The objectives were to assess:

1. The accessibility of publications.
2. The type of data included in, attached to and linked to publications.
3. The metadata given within or linked to publications, specifically (i) use of the standard nomenclature (ICPN 1.0), (ii) inclusion of pictures and photos for identification purposes, and (iii) inclusion of a full method.

A further objective was to compare these assessments to a recent similar assessment for macro-botanical remains (Lodwick 2019) to determine how similar or disparate the two related disciplines are in terms of their adoption of open science practices, particularly concerning data sharing.

## METHODS

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).

To allow comparison to Lodwick (2019), this dataset was gathered from the same journals (see [Table 2](#) for the selection of journals), which are prominent archaeological and palaeoecological journals, and the same date range (2009–2018). 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’. I suggest 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, phytolith analysis is a sub-discipline of archaeobotany along with palynology, starch analysis, anthracology and other macro-botanical remains. I would also include aDNA and isotope analysis of plant remains from archaeological sites. Consequently, the data gathered by Lodwick (2019) will be referred to here as macro-botanical remains, rather than archaeobotanical remains, to distinguish it from this phytolith study.

The methodology used to retrieve the articles needed for this research on open science practices in phytolith research took the following steps:

1. The author accessed the journal website and searched the term ‘Phytolith’.
2. This was then refined to the 10-year period required (2009–2018).
3. Once the list of articles was found, each article was examined for primary phytolith data. Only articles that provided primary data were selected for the dataset. The articles could be archaeological, palaeoenvironmental or methodological. This was determined from the main focus of the article and the research questions being addressed. There is often overlap between palaeoenvironmental and archaeological studies and therefore some articles could have fallen into either category. In these cases, articles were put into the archaeological category as they were focused on samples from an archaeological site.

- *Primary phytolith data* was defined as being the presentation of new data from any sort of phytolith study.
- *Metadata* was defined as any data that provides information about other data. This includes descriptive metadata such as standard vocabularies used and pictures provided to describe identification categories used in the data.

ARCHAEOBOTANICAL JOURNALS	ARCHAEOLOGICAL SCIENCE JOURNALS	GENERAL ARCHAEOLOGY	CROSS-DISCIPLINARY JOURNALS
Vegetation History and Archaeobotany (VHA)	Archaeological and anthropological science (AAS)	Antiquity (ANT)	PLoS One (PO)
	Environmental Archaeology (EA)	Journal of Field Archaeology (JFA)	Proceedings of the National Academy of Sciences (PNAS)
	The Holocene (TH)	Oxford Journal of Archaeology (OJA)	
	Journal of Archaeological Science (JAS)	Journal of Anthropological Archaeology (JAA)	
	Journal of Archaeological Science reports (JASR)	Journal of World Prehistory (JWA)	
	Journal of Ethnobiology (JE)		
	Quaternary International (QI)		
	Journal of Wetland Archaeology (JWA)		

**Table 2** List of journals and categories used by both Lodwick 2019 and this study.

Many researchers still do not have access to all research outputs due to paywalls of academic journals. Access to knowledge, especially research that has been publicly funded, is an issue that must be addressed throughout the scientific community. Thus, I completed this research with no access to an academic library and consequently all articles have been gathered through open access (known as Gold open access), pre-prints and post-prints in open repositories (known as Green open access), academic social media sites such as ResearchGate (RG) or Academia.edu, emailing authors and help from 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.

Once the articles with primary phytolith data were retrieved, information was recorded for each article concerning region of study, period of study, type of study, data location, re-usability of the data, presence of pictures or photos for identification, accessibility of the article, use of ICPN, presence of a full methodology and more detail about the type and content of the data in the article. The categories and codes used in the dataset can be found in [Table 2](#) in Karoune (2020).

A full description of the methodology used in this study can be found in the accompanying data paper (Karoune 2020) along with the link to the full open access dataset (found here: <https://osf.io/8p3bn/>) and also for further transparency a research compendium to explain the research workflow and steps of analysis (found here: <https://osf.io/9wa2f/>).

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

1. In contrast to Lodwick (2019), which used only archaeological studies, this current study used all papers containing primary phytolith data, regardless of discipline. This current study therefore included archaeological, palaeoecological and methodological studies.
2. Only data regarding dataset availability and quality were collected here. Information about data citations and meta-analysis included in Lodwick (2019) were not considered.
3. The assessment of data sharing is specifically with reusability in mind and, therefore, some differences between the current study and Lodwick (2019). First, how the data was presented and 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 could be re-used. A judgement was made that if data was raw counts (and weights) and in an 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. It is incredibly important that data is in an interoperable form so that other researchers can reuse data in different ways and easily merge it with other datasets. Therefore, making an assessment of data content and format in this study will highlight the current ability of phytolith researchers to reuse published data. 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 1.0 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 research of others.

The use of programmatic statistical frameworks such as R or SciPy have low uptake among phytolith researchers. Reusability must also consider the target audience, and as such Excel was used for both data collection and analysis since it is widely used, easy to obtain (although does come with a cost) and is compatible with free and open software such as Libre Office. The steps of the analysis taken in Excel, involving very simple summations and transformations of the data into percentages for the relevant data categories, have been documented fully in the research compendium to produce a fully transparent record.

## RESULTS

### PRIMARY DATA SHARING

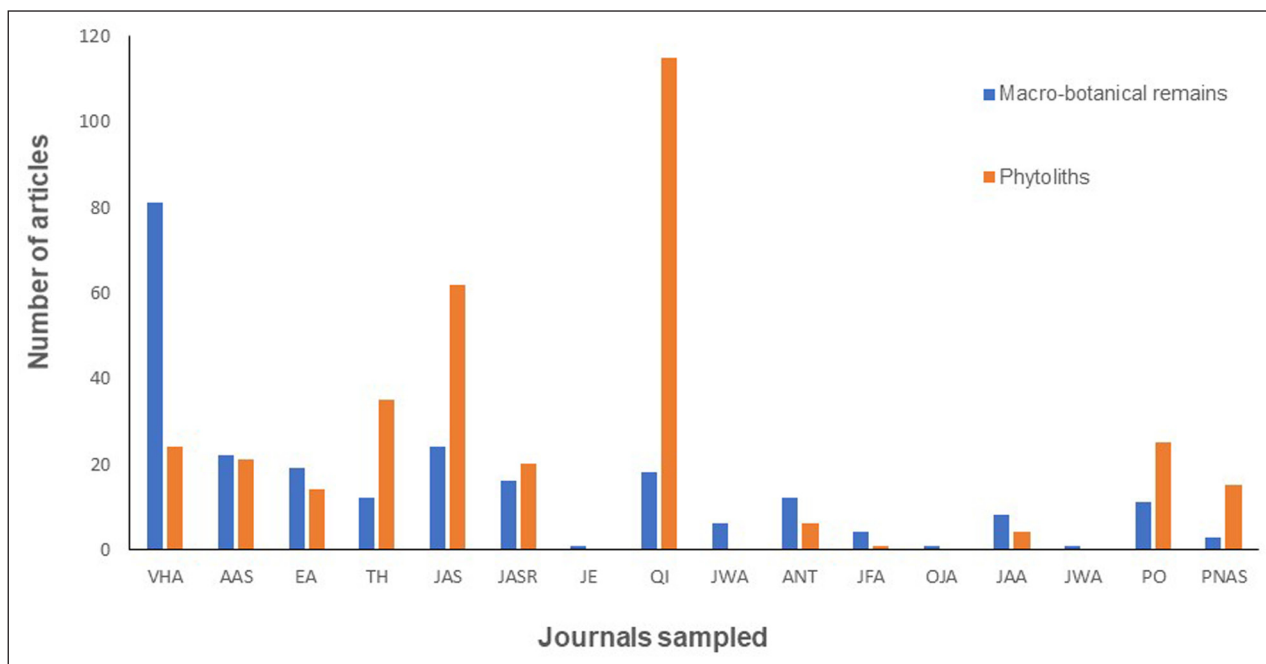
#### Overall data sharing and comparisons to macrobotanical-remains

During collection of the articles, no articles with primary phytolith data were found in the *Journal of Ethnobiology*, the *Oxford Journal of Archaeology*, the *Journal of Wetland Archaeology* or the *Journal of World Prehistory* for the time period of this study – 2009–2018 (see [Table 3](#)). These journals also had low occurrences of macrobotanical articles in Lodwick (2019).

In total, I found 341 journal articles containing primary phytolith data in the 16 journals sampled. These articles only come from 12 of the 16 journals that I sampled, compared to all 16 in the Lodwick (2019) study (see [Figure 1](#)). This included 214 (63%) archaeological, 53 (16%) palaeoenvironmental and 74 (22%) methodological articles. [Table 3](#) shows the number of articles per journal for this study compared to Lodwick (2019) for macrobotanical remains. Phytolith research is in a period of expansion and this can clearly be seen when compared to the number of macrobotanical articles in the same period (n = 239). In the decade sampled, although there are a similar number of archaeological articles, there are

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
<b>Total number of articles</b>	<b>239</b>	<b>341</b>

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



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

more than 100 more articles containing primary phytolith data than macro-botanical data.

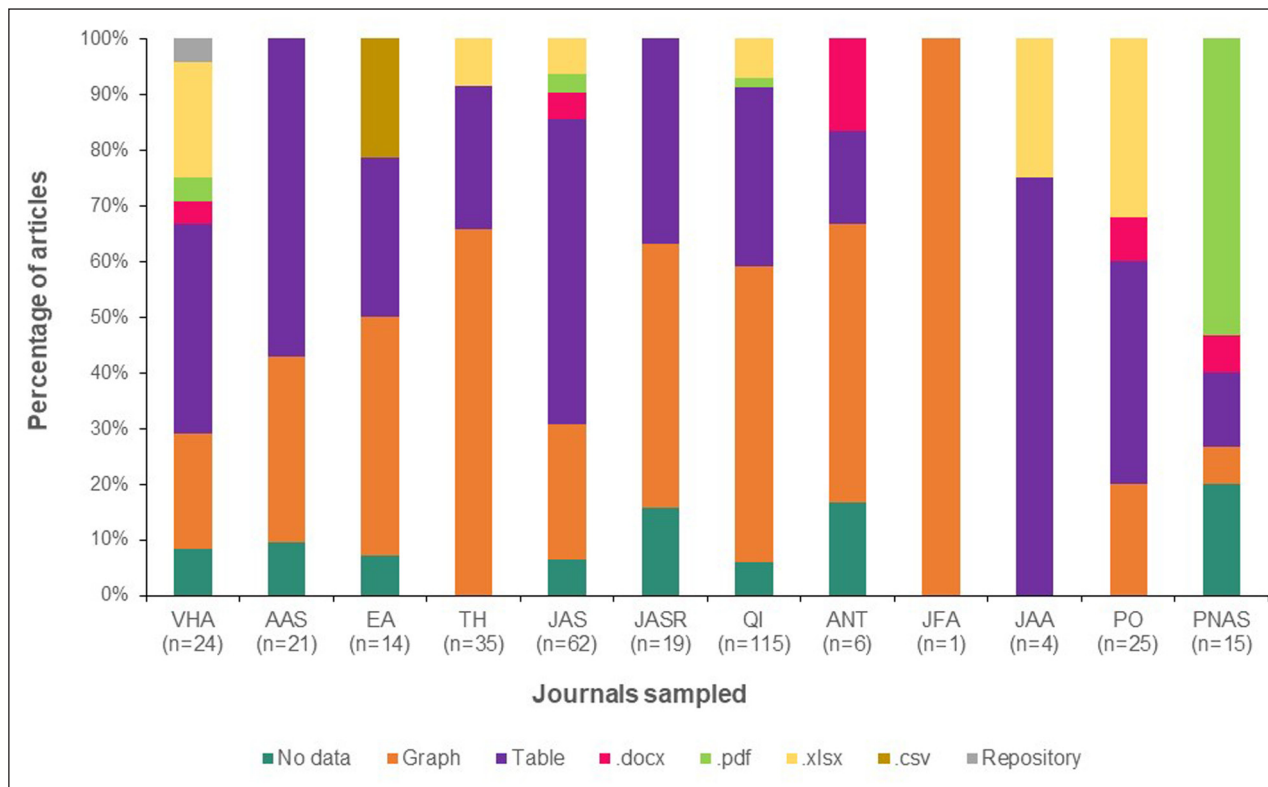
The occurrence of articles in the journals sampled is different (see **Figure 1**); phytolith research was predominantly found in QI, JAS and The Holocene. Whereas primary macro-botanical data, was most prevalent in VHA and then JAS.

**Figure 2** shows results of the location of data sharing in the articles grouped by journal in this study. Overall, for the 12 journals with articles containing primary phytolith data, 53% (n = 182) showed some form of data sharing, including any data in the form of a table, spreadsheet or in an open online repository. This means 47% of the articles did not share primary data – this was a sum of the

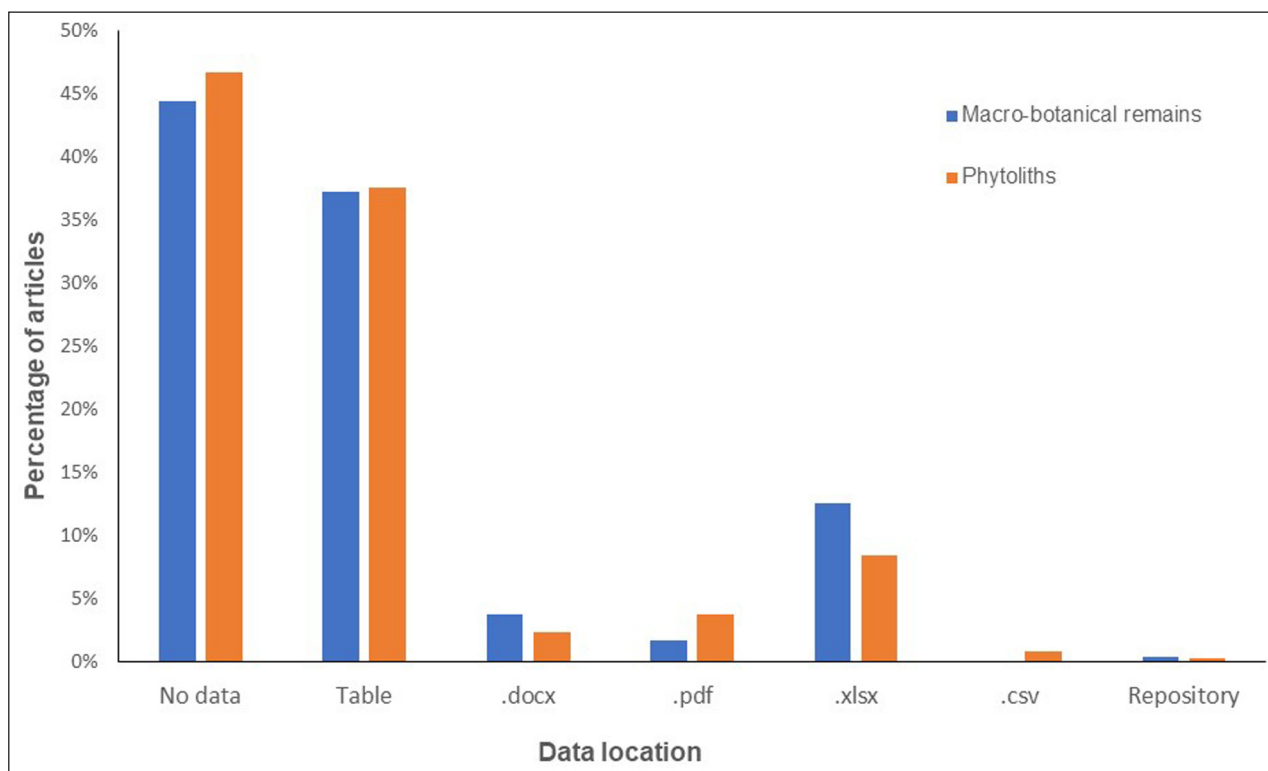
N (no data given in the article, although results would be discussed in the text) and results in graphs. Graphs were deemed to not be data sharing as it is very difficult to extract any reusable data from them. Graphs were in fact the most frequently found format of data presentation in this dataset – making up 40% of all data shown in the

articles. This is a large proportion of data to be unusable.

The level of data sharing in this study 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



**Figure 2** Chart showing the location of primary phytolith data in journals sampled (excludes the four journals that contained no articles with primary phytolith data).



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



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 higher in archaeobotanical and archaeological science publications.

Lodwick (2019) reported that only 52% of articles contained full raw data. The current study found only 15% of phytolith articles contained or referred to raw data (data in any location and could be raw counts, absolute presence or relative presence). The current study also determined that reusable raw data (raw counts of all morphotypes identified in Excel, csv files or a repository) was only 4%. Additional metadata to help with the interpretation of raw counts was often missing. For example, the number of fields counted on the slide and the weights recorded during processing. We need to make sure we supply the data that would be needed in the future for other researchers to be able to reuse our data in other ways. Therefore, it is suggested here that raw counts, number of slide fields, weights recorded during extraction and any transformations of this data are provided to allow the greatest opportunity of reuse.

To be counted as reusable data in this dataset, it was judged to be raw 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 study and therefore reflects a true measure of the reusability of phytolith data. This unfortunately is only 4% and therefore suggests poor data reusability for the studies included in this dataset.

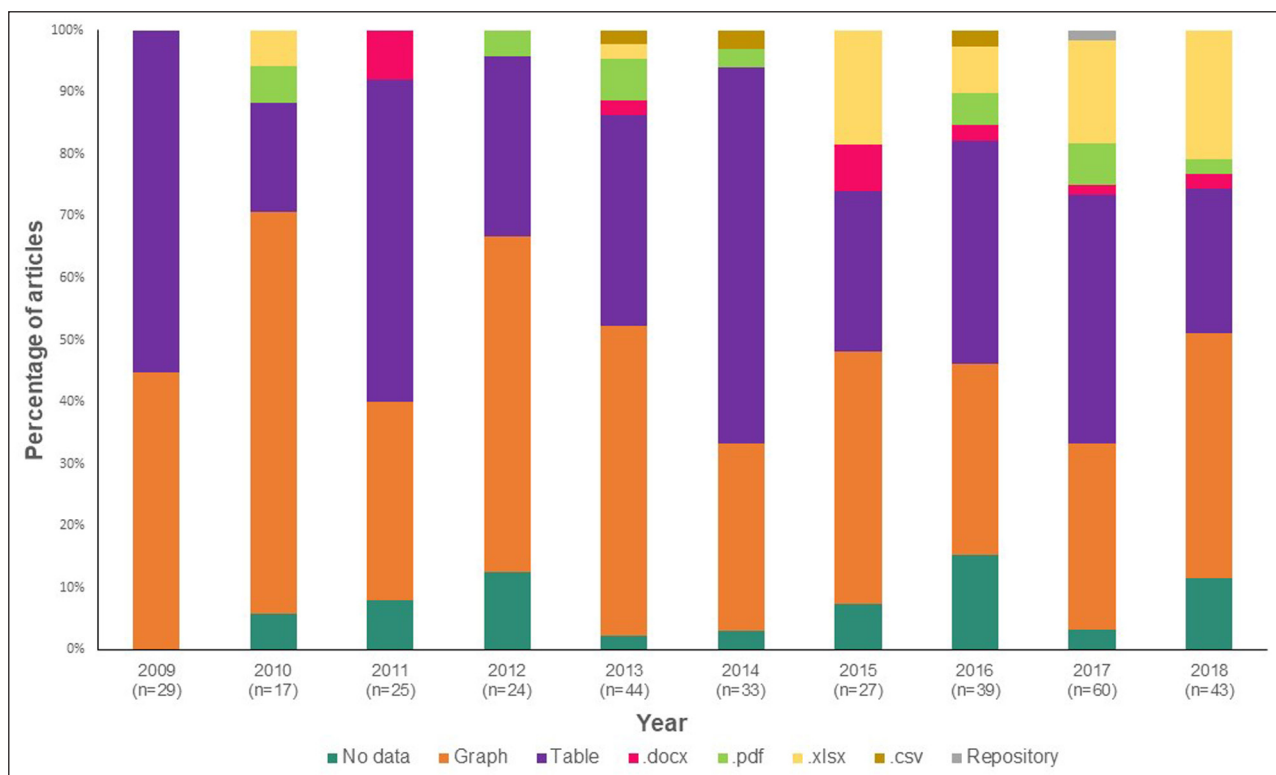
### Data location

Graphs were found to be the most common way to include data in articles 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 data visualisation 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 poor labelling; it was unclear what the graphs 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 was in tables contained within the text of articles (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 was 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



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

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 extract 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 Libre 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 supplementary file use gradually increased in this studies dataset, 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 87% of the articles in this study, then supplementary files are usually also not open. Some journals do allow access to data even if the article is paywalled, 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 putting data in supplementary files limits the accessibility and re-use potential of these data.

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 articles and therefore to find out what the file contained it had to be downloaded. This problem hinders the findability of data and metadata contained in articles.

There was only one article (Lancelotti *et al.* 2017) in this current study that used a repository to archive their phytolith data. Lodwick (2019) also found one article in her study of published macro-botanical data and therefore this demonstrates a lack of awareness of the need for long term storage and accessibility of data in archaeobotany. It clearly shows that much work is needed to address this issue and training must be put in place to improve data sharing more widely than just within phytolith research.

### Data availability and journal data policies

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

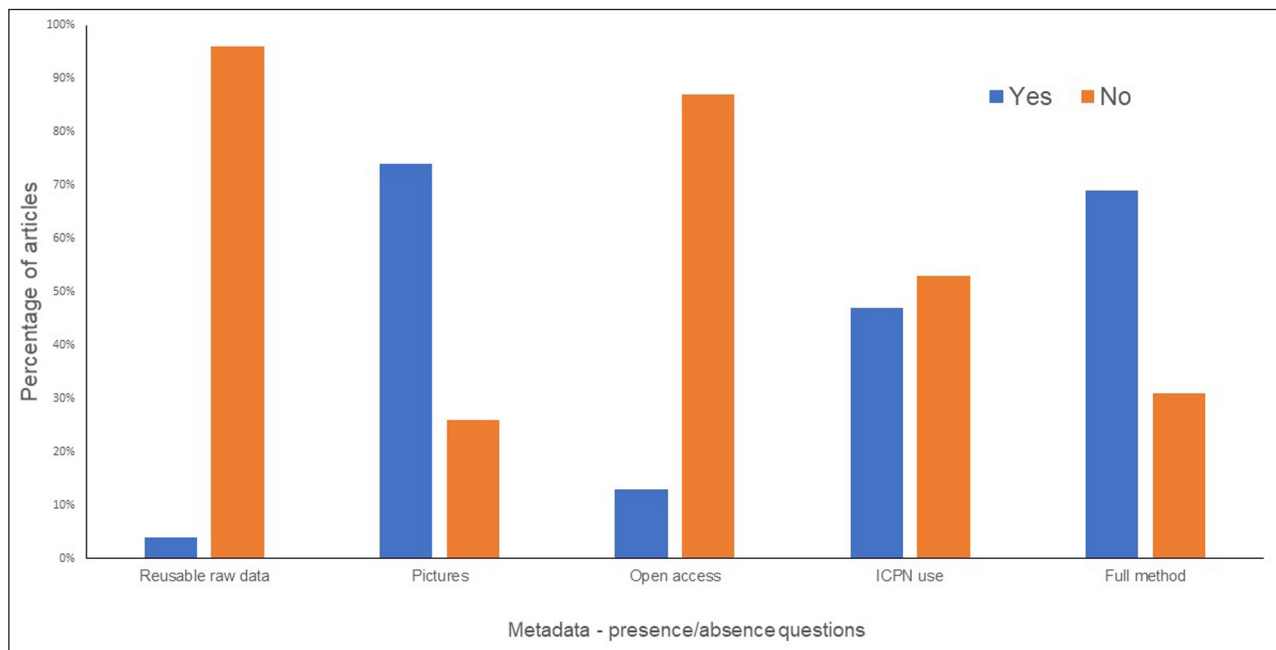
and associated protocols available and should follow the FAIR principles (PNAS 2018). This study found no correspondence between a stricter data policy and better data sharing practices. Journals with a high percentage of articles providing no data or having only data presented in graphs were not just 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 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 VHA in 2017, however VHA's 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 current study that also followed good data sharing practices. Therefore, it seems the authors' knowledge and experience influenced data sharing in this case rather than the journals data availability policy.

The open access journals PLOS One and PNAS, that currently have the strictest data policies, have fewer articles that have no data or that present data only as graphs, and PLOS One 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 the wider archaeological community needs to address if we are to implement better data sharing practices consistently. 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.

### METADATA

[Figure 5](#) presents data gathered for the purposes of assessing the sharing of metadata. Providing pictures for confirmation of identifications, using ICPN 1.0 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 section 4.

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 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.



**Figure 5** Chart showing percentages for presence/absence data categories (recorded as Yes or No): Reusable raw data, Pictures, Open access, ICPN use, Full method.

Likewise, most articles (69%) 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 such as giving more than one methodological reference, referencing a general textbook on phytoliths that contains more than one method and also writing that one method was used with some additions or differences and then not stating these changes. 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. Authors need to consider that the reason for providing a method in a scientific paper is for peer reviewer to assess the validity of the method, for other researchers to be able to replicate experiments to check for reproducibility and for other researchers to assess if their dataset could be combined with a published dataset that used the same methodology. Ideally, the method in an article should be given briefly in the methods section and then a full procedure included in the supplementary information.

Only 47% of the articles used ICPN 1.0. 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.

### OPEN ACCESS

The initial findings for open access are low in this dataset. Only 13% 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 standard procedure for PNAS. 54% of articles were found on academic social media sites such as ResearchGate or Academia.edu. Another 2% of the articles were found in institutional repositories. Therefore, in total 72% of the articles were accessible somewhere on the internet. For the other 28% of articles, they were procured through requesting articles through ResearchGate, 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.

### STEPS FORWARD

This review of open science practices demonstrates there is still much to improve in phytolith research.

While it has been found that general data sharing needs improvement, and therefore the findings are consistent in many ways with Lodwick (2019), 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. Immediate 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?

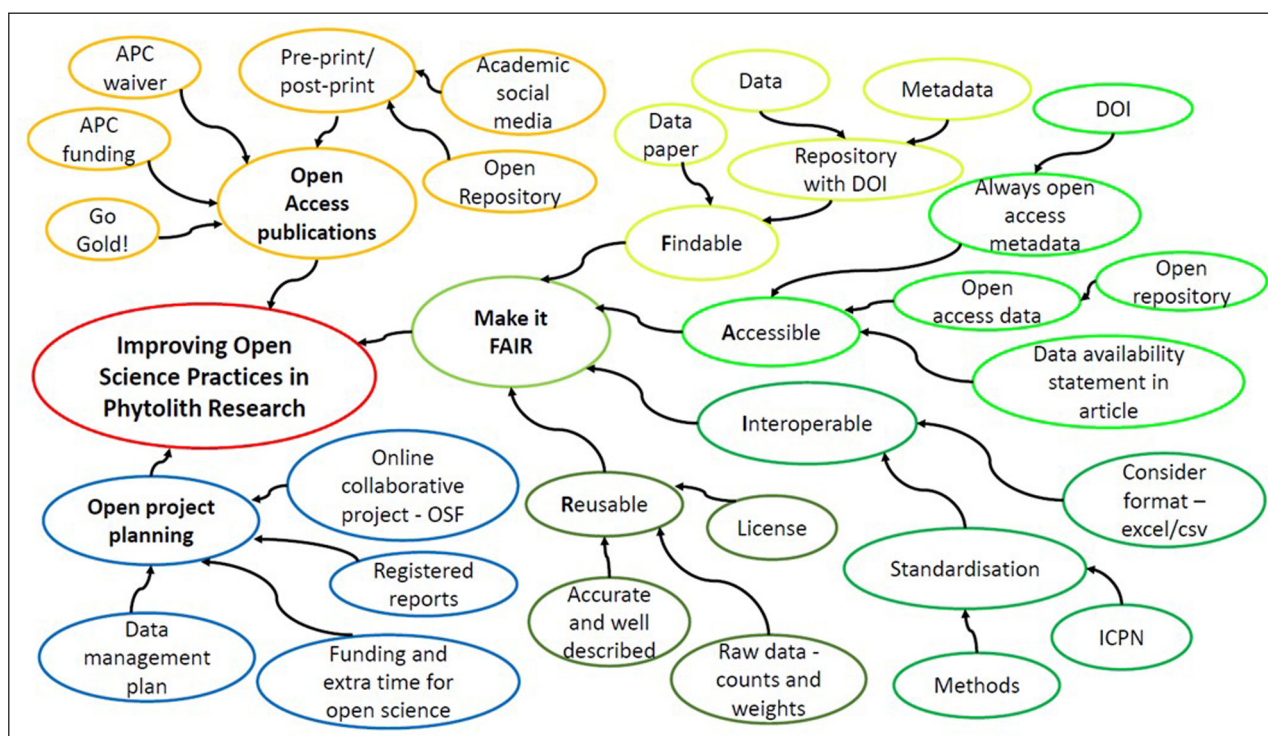
**Figure 6** is a summary of the steps to improve open science practices in phytolith research. As a discipline we must start planning new projects with open science at their heart and strive to always make our publications open. We must also improve our data sharing, and this can be done by adhering 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.

### OPEN PROJECT PLANNING

Improving open science practices throughout the phytolith community requires planning. It takes extra time and consequently more money to work openly, and therefore initial project planning needs to take this into account. 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 the expertise needed to make it happen. For example, 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). However, data storage can be free as there are many free open repositories that can be used to store any type of data or research output such as Figshare, Open Science Framework (OSF) or Zenodo.

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 or tool to create an open workflow, such as using OSF or through Github linked to Zenodo. These platforms can be used to project manage and 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. Both OSF and



**Figure 6** Summary of proposed steps for improving open science practices in phytolith research. An alternative version of this diagram as a table can be found here: <https://osf.io/9wa2f/>.

Zenodo assign a citable DOI to the data archive when any aspect is opened publicly. DOI's can be used to make the research output, such as data, more findable and will increase citation of your work. Every output can also have a license to tell other researchers how your output can be reused. Open workflow tools are an efficient way to work with other colleagues and can be used to link together different applications such as Google docs and Github. 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 project design and emphasize the importance of research questions and robust methodology (Centre for Open Science 2020). These articles are to be written after gathering 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.

### OPEN ACCESS PUBLICATIONS

It has been found in this project that the overall open access to publications needs improvement. However, with added access gained through academic social media sites and informal arrangements, such as emailing authors, access could be gained to the majority of publications in this dataset. Much still needs to be improved in this regard as 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 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 researchers in 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 agreements 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*. 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 journals allow 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-print 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.

### 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 more findable as separate entities, phytolith data and metadata needs to be deposited 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 both a map and a sample list, linking entries to publications. However, being linked to published articles does not mean the data is accessible. Access could be improved if all of these datasets were put into open repositories and the links added to the database. Improving the use of open repositories 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. Specific data journals provide a template for data papers and these articles 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 the dataset in this study 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 the 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 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 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 and code written for analysis along with datasets, although few phytolith researchers are currently using open code analysis.

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, more work is needed,

in particular, a comprehensive study of extraction methods to assess the differences in recovery from all methods currently utilised by phytolith researchers would be of great benefit. A study of extraction methods 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 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 readers to know what is contained in each file, however there was also issues with 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 headings and tabs. However, when you are providing a dataset for other people to view, and potentially reuse, 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 of measurements and did not make sample numbers or names clear. This is where good metadata

plays a role in the reusability of a dataset. Keys or data dictionaries need to be provided that explain any codes used and care should be taken when adding any headings to spreadsheets.

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 many 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.

## CONCLUSIONS

There are a number of positive findings 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 significant work required for a more complete adoption of open science practices in phytolith research. 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 STATEMENT

There is a data paper to accompany this research project. The DOI for this is: <http://doi.org/10.5334/joad.67>.

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

The research compendium for this project can be found here: <https://osf.io/9wa2f/>. It includes the raw data csv, an excel data analysis file, a readme file that explains the steps of data collection and data analysis, a workflow diagram and an alternative **Figure 6** that provides more information about steps forward and relevant resources in a simple table.

The dataset from Lodwick (2019) can be found here: <http://doi.org/10.5287/bodleian:7QkmEEp9Y>.

## 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. I would also like to thank the two reviewers that helped to improve this article.

## COMPETING INTERESTS

The author has no competing interests to declare.

## AUTHOR AFFILIATIONS

**Emma Karoune**  [orcid.org/0000-0002-6576-6053](https://orcid.org/0000-0002-6576-6053)  
The Alan Turing Institute, GB

## 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/guidelinesForDepositores.xhtml> [Last accessed 12<sup>th</sup> 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: <https://doi.org/10.1016/j.jas.2015.03.023>
- Cabanes, D.** 2020. Phytolith Analysis in Paleoecology and Archaeology. In: Henry, A. (ed.), *Handbook for the Analysis of Micro-Particles in Archaeological Samples*. Interdisciplinary Contributions to Archaeology. Cham: Springer. DOI: [https://doi.org/10.1007/978-3-030-42622-4\\_11](https://doi.org/10.1007/978-3-030-42622-4_11)
- 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 12<sup>th</sup> 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. DOI: <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 23<sup>rd</sup> June 2020).
- García-Granero, JJ, Hatzaki, E, Tsafou, E, Ayala, G, Serpetsidaki, I and Bogaard, A.** 2021. From Storage to Disposal: a Holistic Microbotanical Approach to Domestic Plant Preparation and Consumption Activities in Late Minoan Gypsades, Crete. *Journal of Archaeological Method Theory*, 28: 307–331. DOI: <https://doi.org/10.1007/s10816-020-09456-9>
- 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: <https://doi.org/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>
- Jovanović, J, Power, RC, de Beccdelièvre, C, Goude, G and Stefanović, S.** 2021. Microbotanical evidence for the spread of cereal use during the Mesolithic-Neolithic transition in the Southeastern Europe (Danube Gorges): Data from dental calculus analysis. *Journal of Archaeological Science*, 125. DOI: <https://doi.org/10.1016/j.jas.2020.105288>
- Kansa, E.** 2012. Openness and archaeology's information ecosystem. *World Archaeology*, 44(4): 498–520. DOI: <https://doi.org/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: <https://doi.org/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*, 8(1): 6. DOI: <https://doi.org/10.5334/joad.67>
- 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. DOI: <https://doi.org/10.1093/aob/mci172>
- Madella, M and Lancelotti, C.** 2012. Taphonomy and phytoliths: a user manual. *Quaternary International*, 275: 76–83. DOI: <https://doi.org/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: <https://doi.org/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: <https://doi.org/10.1093/aob/mcz064>
- Pearsall, DM.** 2019. *Case studies in Paleoethnobotany*. New York: Routledge. DOI: <https://doi.org/10.4324/9781351009683>
- 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: <https://doi.org/10.1007/s00334-011-0341-z>
- Shillito, L-M, Blong, JC, Green, EJ and van Asperen, EN.** 2020. The what, how and why of archaeological coprolite analysis. *Earth-Science Reviews*, 207. DOI: <https://doi.org/10.1016/j.earscirev.2020.103196>
- Watling, J, Castro, M/T, Simon, MF, Rodrigues, FO, Brilhante de Medeiros, M, De Oliveira, PE and Neves, EG.** 2020. Phytoliths from native plants and surface soils from the Upper Madeira river, SW Amazonia, and their potential for paleoecological reconstruction. *Quaternary International*, 550: 85–110. DOI: <https://doi.org/10.1016/j.quaint.2020.03.045>
- Wilkinson,** et al. 2016. The FAIR guiding principles for scientific data management and stewardship. *Scientific data*, 3: 160018. DOI: <https://doi.org/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: <https://doi.org/10.1371/journal.pone.0143460>
- Yost, CL, Ivory, SJ, Deino, AL, Rabideaux, NM, Kingston, JD and Cohen, AS.** 2020. Phytoliths, pollen, and microcharcoal from the Baringo Basin, Kenya reveal savanna dynamics during the Plio-Pleistocene transition. *Palaeogeography, Palaeoclimatology, Palaeoecology*. DOI: <https://doi.org/10.1016/j.palaeo.2020.109779>
- Zuo, X and Lu, H.** 2019. Phytolith radiocarbon dating: a review of previous studies in China and the current state of the debate. *Frontiers in Plant Science*, 10: 1302. DOI: <https://doi.org/10.3389/fpls.2019.01302>
- Zurro, D, Garcia-Granero, JJ, Lancelotti, C and Madella, M.** 2016. Directions in current and future phytolith research. *Journal of Archaeological Science*, 68: 112–117. DOI: <https://doi.org/10.1016/j.jas.2015.11.014>

---

TO CITE THIS ARTICLE:

Karoune, E. 2022. Assessing Open Science Practices in Phytolith Research. *Open Quaternary*, 8: 3, pp.1–18. DOI: <https://doi.org/10.5334/oq.88>

Submitted: 04 August 2020 Accepted: 24 January 2022 Published: 10 March 2022

COPYRIGHT:

© 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

*Open Quaternary* is a peer-reviewed open access journal published by Ubiquity Press.

