Grains of truth or transparent blindfolds?

Citation for published version:

Digital Object Identifier (DOI):
10.1007/s00334-011-0341-z

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Vegetation History and Archaeobotany

Publisher Rights Statement:

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
**Grains of truth or transparent blindfolds? A review of current debates in archaeological phytolith analysis**

Lisa-Marie Shillito

Abstract Phytolith analysis has become an increasingly popular archaeobotanical tool in the past few decades. Phytoliths have been used to support key hypotheses relating to the domestication of several food crops and in the study of ancient diet, and they are of particular importance in contexts where other plant remains are poorly preserved. However, the discipline has also been subject to controversy and debate. This paper gives an overview of the technique and three key case studies covering a range of geographical areas. Some of the problems that are common to each are discussed and suggestions are made for how these problems could be resolved in future research. It is suggested that further caution should be taken during interpretation, and a greater consideration given to taphonomy. Despite these criticisms it is concluded that there is still much potential in the technique, particularly when integrated with other lines of microarchaeological evidence.

Keywords Maize · Rice · Cereals · Taphonomy · Morphometrics · Microarchaeology

**Microfossil research in archaeology**

The term microfossil is used to describe biogenic particles which cannot be seen by eye, and covers materials including pollen, phytoliths, starch, diatoms and micro-charcoal. Phytoliths are opaline silica deposits that form within and between the cells of plants, forming ‘casts’ of the cells and intercellular spaces. Monocotyledonous plants are particularly prolific phytolith producers and can form large conjoined or ‘multi-celled’ phytoliths of sections of plant tissue. Silica is absorbed by the plant in a soluble form from groundwater and precipitated at different locations within the plants through a polymerisation process. It has been suggested that phytoliths provide structural support (Kaufman et al. 1985) or protection from herbivores (McNaughton et al. 1985; Cid et al. 1989; Massey and Hartley 2006; Massey et al. 2007; Reynolds et al. 2009) but the reasons for production and variability within and between taxa are still poorly understood. There is evidence that both genetic and environmental factors play a role, and that silica uptake is both an active and passive process, with particular cells becoming silicified first (Cooke and Leishman 2011).

‘Phytolitharia’ were identified over a century ago, commented on notably by Darwin, who recognised them as a component of wind-blown dust on the Beagle (Darwin 1846). The study of phytoliths has been applied to archaeological questions since the 1970s (Rovner 1971), and the first textbook by Piperno was published in 1988. From the 1990s the discipline saw a broadening of applications to include palaeoenvironmental reconstruction (Wallis 2001; Delhon et al. 2003), archaeobotanical studies of food and non-food use of plants (Piperno 1991; Tubb et al. 1993; Lentfer et al. 1997) and in reconstructing past agricultural processes (Rosen and Weiner 1994; Harvey and Fuller 2005). The past decade has seen a massive increase in the use of phytolith analysis in archaeological science, with recent developments integrating phytolith analysis with multiple microarchaeological techniques such as micromorphology and geochemistry (Weiner 2010; Albert et al. 2010; Cabanes et al. 2010).
Geographically, archaeological phytolith research can be broadly divided by continent. In the Americas a major focus has been investigating the origins of domesticated maize and other food crops (Piperno et al. 2009; Piperno 2009), East Asia has seen a similar focus on rice agriculture (Pearsall et al. 1995; Zhao et al. 1998; Zheng et al. 2003), whilst studies in West Asia (including the Near East and Mediterranean), have focused on major cereals such as wheat and barley (Rosen 1992; Tubb et al. 1993), and the non-food use of plants for fuel, bedding, basketry and matting for example (Gé et al. 1993; Madella et al. 2002; Rosen 2005; Albert et al. 2000, 2008, 2010).

The advantages and problems of phytolith analysis as an archaeobotanical tool

The major advantage of phytoliths over other botanical remains is that they do not need to be charred or waterlogged to be preserved. Studies of plant resource use in archaeology have traditionally focused on charred remains which are preserved only up to temperatures of c. 500°C, and as noted by Van der Veen (2007), the lighter components of plants such as chaff are readily lost on burning. Phytolith morphology is preserved to at least 800°C, and it has been demonstrated that their morphology can survive above this temperature, but they may become discoloured (Parr 2006) or partially melted (Matthews et al. 1997). Physical properties of phytoliths such as their refractive index can also be used as indicators of burning (Elbaum et al. 2003). The charred macrobotanical record is overwhelmingly a result of fuel use or the burning of food stores (Hillman 1981), whereas phytoliths can provide evidence of a more diverse range of activities. As well as remaining after the burning of plants as fuel, phytoliths can also enter the archaeological record when they are released during the decay of organic remains.

There has been a particular concern over the use of phytoliths to identify domesticated varieties of crop plants (Staller and Thompson 2002; Staller 2003; Fuller et al. 2010), and also the use of phytolith size as an indicator of past water availability and irrigation (Jenkins 2009; Shillito 2011a). This is a result of morphological criteria which some researchers see as problematic and difficult to reproduce, a lack of standardised methods for identification and quantification in different contexts, and a poor understanding of taphonomic processes affecting phytolith morphology and assemblages. Conflicts with other lines of evidence such as macrobotanical remains are also a concern. Despite these problems the technique is becoming increasingly applied to archaeological case studies, in many cases it seems without a clear understanding or consideration of the problems. This review attempts to address this, by providing an overview of three key case studies and the problems these have in common, and suggestions for some of the ways in which research could be improved in future studies.

Methods of analysis, identification and quantification in archaeological phytolith analysis

Phytolith investigations in archaeology can be divided into two broad categories; samples from specific archaeological features and artefacts such as ‘food crusts’ adhering to pottery, dental calculus, coprolites and ash deposits, and those extracted from bulk sediment and soil samples. With specific features, the assemblage can be more securely linked to a specific activity due to the associated contextual information. When small numbers are present, the entire assemblage is recorded, either as morphological types, or divided into categories from different parts of the plant (stems, leaves, husks etc.). With bulk sediment samples we can obtain a more general picture of plant use, as there may be multiple possible origins for the assemblage (Matthews 2010; Shillito 2011b), and counts are made on what is considered a ‘representative’ proportion of the assemblage.

Unlike similar disciplines such as palynology there are no firm guidelines for quantification, and this varies between studies depending on the research aims. Biases produced by insufficient count sizes are a concern, as this may result in errors in vegetation inference in palaeoecological studies, or errors in size statistics in studies of morphological quantification which are often used to distinguish between wild and domestic forms (Strömberg 2009).

Many studies have adopted the method of Albert and Weiner (2001), who calculated the numbers of phytoliths per gram of sediment in ash deposits, with minimum counts of 300 single and 100 conjoined cells suggested. However in contexts with lower abundances this leads to over inflation of very small numbers of phytoliths and can be misleading. A good review of quantification methods in phytolith analysis is given by Strömberg (2009), who concludes that the most suitable count size may vary considerably depending on the study aims and that appropriate count size needs to be determined on a case by case basis. The use of multiple samples is also stressed by Zurro et al. (2009) who found a high degree of variability within single contexts, suggested to be due to the diverse origins of phytoliths in archaeological deposits.

Until 2005 and the publication of the International Code for Phytolith Nomenclature (ICPN) (Madella et al. 2005), there was no standardised method for phytolith classification; researchers were using different terminology, sometimes for the same phytolith types. This approach makes comparison between different studies difficult. The ICPN was devised in an attempt to standardise the classification
and description of phytoliths, based on visually determined morphological criteria. Whilst this has begun to improve the way phytolith data is recorded, the links between the shape descriptions and how the phytolith forms relate to different plant cells and different parts of the plant is not always clear. In addition, some researchers still use their own classification systems. For example the distinct spiny phytoliths produced in date palm, described as globular echinate according to the ICPN, are referred to variably as spherical echinate (Albert et al. 2009), circular crenate (Ishida et al. 2002), and decorated spheres (Henry et al. 2011).

Many single cell phytoliths are common to a wide range of taxa and can be identified to family or genus rather than species, for example a common distinction is made between ‘rondel’ short cells associated with C3 plants, and ‘bilobate’ short cells associated with C4 plants (Twiss 1992). Identifications are made through visual comparisons of modern reference material with archaeological samples, with no standard identification key. Ball et al. (2009) discuss the distinction between typological and morphometric phytolith descriptions. The former (which includes the ICPN approach) describe phytoliths in terms of the general shape, for example ‘bilobate’ and ‘rondel’ shaped short cells, and in terms of the plant part from which they derive, for example ‘trichome’ or hair cell. The latter are used to describe and classify the same phytolith types which occur in several taxa, but have shape and size variations that may be used to distinguish between taxa, for example variations between papillae phytolith morphology in different cereals (Rosen 1992).

There are difficulties with both typological and morphometric approaches. Typologically, similar phytolith types occur in a wide variety of genera, and often the diagnostic criteria are based only on visual comparisons which may be prone to subjectivity, or at least variation between different analysts. Studies of the relationship between morphological data, the basis of systematics, and other variables in phytolith analysis, are almost entirely qualitative in this regard. The use of morphology as a diagnostic tool is complex in that shape criteria can be influenced by a huge variety of internal and external factors, many of which are also just components of random variation (MacLeod 2005). This is not just a problem in phytolith analysis but in many areas of systematics studies. The subjectivity of palynology for example has long been recognised (Stillman and Flenley 1996). MacLeod (2005) suggests this is due to the challenge in data analysis of morphological criteria. Geometric morphometric approaches offer an objective method for assessing phytolith shape, but the degree of random variation must also be assessed, which requires very large population sizes. The reasons why this is a particular concern in phytolith studies will become apparent by examining three case studies in “Case studies” section, which rely on very specific morphological characteristics.

Case studies

Domestication of maize in Central and South America

One of the most prolific areas of phytolith research has been the investigation of maize domestication in the Americas. The Zea genus consists of five species, including Z. mays, which itself is divided into four subspecies. Of these sub-species, Z. mays spp. mays is the modern domestic variety. All other species and sub-species of the genus are classed as teosinte (Kellogg and Bircher 1993).

The key argument is based on the difference in short cell phytoliths between teosinte and modern maize, initially cross shapes in leaves and more recently decorated rondels from glumes and cupules of cobs. It was first suggested by Pearseal (1978) that large cross phytoliths could be used to distinguish between teosinte and maize. Refined morphological parameters of 3D observations were described by Piperno (1984) and summarised in Piperno (2009). Variant 1 crosses from maize are said to usually have an average width between 12.7 and 15 µm, whereas a prevalence of wild grasses will contribute smaller cross-bodies with an average width between 10 and 12.5 µm. Maize leaves are also dominated by ‘mirror image’ cross shapes, compared to wild grasses which are dominated by other (non-specific) types of 3D structures.

The cross shaped 3D structure of teosinte is said to be dominated by shape Variants 2 and 6, whereas maize is dominated by Variant 1. One of the problems lies in the fact that these shapes are not actually confined to either species—teosinte can produce Variant 1, and maize can produce Variant 2 and 6, but ‘less frequently’. In an archaeological example presented by Piperno (1984) from Cueva de los Vampiros (6610 B.C.), 21 cross phytoliths were observed, with 19 Variant 1 and one Variant 6, interpreted as wild based on the size of the Variant 1 forms. The later deposits in this study at Aguadulce and Cueva de los Ladrones are interpreted as having maize due to a combination of size and morphological criteria, including ‘extra large’ crosses. These are only referred to with percentages in the discussion, but in Table 3 (p. 378) we can see that the argument is based on only 1–2 extra large crosses. It is likely that there are not enough individuals to confidently assign the assemblages to one category or the other, especially considering the degree of morphological variation even within an individual as discussed in “Methods of analysis, identification and quantification in archaeological phytolith analysis” section (MacLeod 2005).

Russ and Rovner (1989) criticised the initial study saying that environmental variables which may have
affected the size and frequency of phytoliths were not controlled sufficiently. They applied computer assisted image analysis to closely related maize and teosinte varieties grown under strict conditions, and concluded that the 2D shape of both cross and bilobate phytoliths could be used to distinguish maize and teosinte in modern populations, and that genetic factors were more important than environmental ones in controlling size.

Experiments on different types of maize were carried out by Doolittle and Frederick (1991) to assess the consistency of size variations, and to see if there were differences depending on the stage of plant growth. Their concern was that Piperno’s criteria, relying on a ‘high proportion’ of a specific size and shape, are too general. They did not find any of the diagnostic cross types described, but are criticised by Piperno (1998, 2003) and Iriarte (2003) who suggest that they simply were not experienced at identifying these types. Even if this is the case, it highlights a problem in a classification scheme that is essentially qualitative, subjective and not easily reproducible.

A study of modern grasses by Iriarte (2003) on taxa from South America uses discriminant analysis to assess the variability of cross shapes between wild and domestic types. The criteria given are more specific; Iriarte provides values of between 12 and 39% of cross shapes with a width of 16 μm or larger for maize, whereas wild types do not produce more than 8% of crosses with widths larger than 16 μm. Maize also produces an average of 3% ‘extra large’ crosses greater than 20.6 μm wide, whereas none of the wild types produce this size. Again there is considerable variability and overlap in the size ranges here, and it is unclear how these values could be translated to an archaeological assemblage, or how large the assemblage needs to be to confidently distinguish the presence of domestic types versus larger than average wild types.

A concern with the experimental studies by Piperno (1984) and Iriarte (2003) is that they are based on modern assemblages of known numbers of plants, and parts of plants, and provide only probabilities that a certain phytolith type belongs to a particular category. With archaeological samples we do not know how many individual plants or taxa we are dealing with, we do not know the stage of plant growth, the depositional origin of phytoliths in soils is unclear, and results are interpreted as definitive rather than as probabilities.

It is also unclear how useful the presence of leaf phytoliths is for identifying the food use of these plants. More recent work suggests that decorated rondels from corn cobs are also diagnostic of maize. ‘Wavy top’ rondels are said only to occur in maize and not teosinte. ‘Ruffle top’ types occur in maize and teosinte but not wild grass, and so it is said they can be used as maize indicators outside the known geographic range for wild teosinte.

‘Oblong half decorated bodies’ occur only in teosinte (Pearsall et al. 2003).

These shape criteria were subject to ‘blind testing’, whereby 84% of maize was correctly identified using the cob phytoliths, and 79% of teosinte. No non-Zea grasses were confused with Zea, suggesting that these phytoliths can at the very least be used to identify to genus level. The lack of overlapping size characteristics is also encouraging, as it suggests no possible overlap between different species. One concern with this however is the statement that any Zea outside the known geographical range for teosinte can automatically be considered maize. This makes assumptions about past geographic distributions of different grasses, and assumes that a wild grass type will never be found outside its normal geographical range. If ‘ruffle tops’ are present (maize and teosinte) but not ‘wavy tops’ (only maize) it could just as easily be a teosinte specimen occurring outside its preferred range.

Rovner (2004) criticised the applicability to archaeological assemblages in non-sealed contexts; he suggests that it is only in closely constrained contexts such as carbonized food residues in pottery, that there is a reasonable assurance of limited taxa. Rovner suggests that in archaeological contexts, the possibility of mixed taxa occurring would make the criteria suggested by Pearsall et al. (2003) unusable. However, unlike the cross-shapes in leaves, the ‘wavy tops’ are said to only ever occur in maize, so mixed assemblages would theoretically not be a problem if it could be demonstrated with confidence that (a) they are morphologically consistent and clearly distinguishable and (b) the shape criteria are not affected by post-depositional processes.

Even in closely constrained contexts, the small numbers of cells used by some researchers is a concern. For example Bozarth and Guderjan (2004) conclude that ears of maize were placed in Mayan vessels based on the presence of three cob phytoliths. A study of 17 stone tools by Pearsall et al. (2004) recovered starch grains and phytoliths, but in very small numbers; the most recovered from a single tool was 19 starch granules and 3 phytoliths, with a total of 91 granules and 17 phytoliths. It could be argued that even one phytolith in a sealed context is enough evidence for the plant’s presence in the past, but then we also have to question what information we can reconstruct about the use of a plant on the basis of such limited evidence.

Phytoliths and rice domestication

An excellent review of the current state of research into rice domestication is given in Zhao (2010) and Fuller et al. (2010). Here a brief overview is given of the issues related specifically to phytolith evidence. Three phytolith types are considered diagnostic of the Oryza genera—fan shaped
bulliforms with ornamentation and lateral protrusions, scooped bilobate short cells, and conjoined phytoliths from epidermal tissue (Pearsall et al. 1995; Chen and Jiang 1997; Zhao et al. 1998). Data suggest domesticated forms of rice have larger bulliforms, and morphometric analyses of the ‘double-peaked’ cells are used as domestic indicators (Zhao et al. 1998), though Zhao cautions that these morphological indicators of domestication do not necessarily equate with the beginning of rice farming (Zhao 2010). Zhao et al. (1998) applied linear discriminant analysis to double-peaked glume cells and found that, although there is a degree of difference between domestic and wild rice, there is still enough overlap in glume cell characteristics to make this a poor domestication indicator by itself. It is only by applying ‘prior probabilities’ that Zhao et al. are able to get a statistically satisfactory distinction between wild and domesticated types.

In general the data is treated with more caution than in maize studies, though the problems are similar. For example Fuller et al. (2010) argue that the large amount of variation and overlap in modern rice makes this an unreliable indicator, and that phytolith data should be treated with caution until the full range and extent of variability in cellular morphology has been determined (Fuller et al. 2007). Fuller et al. (2010) also argue that the contradiction between phytolith and macrobotanical evidence indicates that phytolith criteria are unreliable. Instead they suggest that variations in phytolith size may just as likely be due to climatic adaptation.

In addition to uncertainties over the use of morphological criteria, there are similar taphonomic problems as with other phytolith studies discussed. For example phytolith husks from peat deposits are cited as early evidence of rice in Korea (Lee and Woo 2003. However, it is questioned whether this is actually an intrusion of later Holocene material (Ahn 2010). As with other soil contexts, without supporting microstratigraphic analysis, it is impossible to identify whether post-depositional movement has occurred. Itzstein-Davey et al. (2007) repeat the assumption that rice phytoliths do not move far from their place of deposition, citing Piperno (2006) in support. Again, there is no evidence to support this assumption; in fact there is evidence to the contrary, as discussed in “Stratigraphic mixing/movement and potential contamination” section.

Phytolith analysis in the Mediterranean and Near East

As with other areas, phytolith analysis in this region has focused to a large extent on the origins of cereal cultivation and domestication. The presence of very large concentrations of phytoliths, often visible as white deposits at the macroscale, has also led to their application as non-food use indicators, and in identifying different activity areas within settlements and habitation areas (Ollendorf 1987; Albert et al. 2000; Rosen 1992, 2005; Jenkins and Rosen 2007; Shahack-Gross et al. 2005; Shahack-Gross and Finkelstein 2008; Cabanes et al. 2010).

Non-food use of plants

Phytoliths from secure, non-mixed contexts, such as in situ burning, are an important source of information on fuel use (Albert et al. 2000), especially where monocotyledonous plants have been used as fuel. Plants such as grasses and sedges burn quickly compared to wood, and do not tend to preserve in the charred macrobotanical record, but can survive in phytolith form. These have been observed as a significant component of ash deposits, for example at Kebara cave, Israel (Albert et al. 2000; Albert et al. 2010), the Sumerian city of Abu Salabikh, southern Iraq (Matthews et al. 1994) and the Neolithic settlement of Çatalhöyük, Turkey (Matthews 2005; Shillito et al. 2011). In some cases, phytoliths may be almost the only plant remains present, for example at the Dilmun settlement at Saar, Bahrain, a small quantity of seeds was recovered from flotation (Nesbitt 1993), whereas micromorphological observations of ash deposits showed that these consisted almost entirely of date palm phytoliths (Matthews et al. 1997; Matthews and French 2005).

It is the combination of secure context and ease of relative quantification that makes ash deposits particularly well suited to phytolith analysis, and clear distinctions can be made between different fuel inputs. By further combining phytolith analysis with micromorphology, the specific context of the phytoliths at the microscale can be observed. Micromorphology involves collecting an intact block of sediment directly from a section, impregnating it with resin, then cutting and grinding it to a thin section with a standard thickness of 30 µm, which allows individual fine layers and their structure to be resolved using a polarising microscope (Bullock et al. 1985). This method has the advantage that it enables the observation of different components of ash simultaneously, in their precise depositional context (Matthews 2010). The association of different materials is important in understanding their origin, for example whether plant remains have been used as fuel in their own right, or if they have entered the assemblage as inclusions in animal dung. Such problems have been noted for seed assemblages from the Neolithic sites of Makriyalos and Makri in Greece, where Valamoti (2006) suggests the use of micromorphology to help distinguish the presence of animal dung, which can be distinguished as distinct pellets, or by the presence of calcareous spherulites (Canti 1998; Portillo et al. 2009). This approach has been successfully applied in a wide range of fuel studies (Shahack-Gross et al. 2005; Goldberg et al. 2009).
extracting phytoliths for analysis, this association with other materials at a microscopic level is lost, and it is difficult to distinguish between assemblages with diverse depositional origins (Shillito 2011b).

In situ phytolith impressions of basketry and matting in buildings at Çatalhöyük have provided important insights into the use of plants such as reeds and sedges in craft activities (Rosen 2005; Ryan 2011). Visible layers of phytoliths on skeletal remains have also provided direct evidence for the use of bindings in burial contexts (Boz et al. 2007). Thick layers of monocotyledonous phytoliths have also been observed as highly articulated layers in floor thin sections at Çatalhöyük (Matthews 2005) and at Neolithic Makri, Greece (Karkanas and Esfritou 2009) which are also interpreted as matting. However, sampling from floors with no macro-scale context such as matting can be a problem, due to possible reworking from trampling (Matthews 2010), or contamination during excavation.

Analysis of floors near ovens at Çatalhöyük has identified wheat husk phytoliths associated with Phragmites reeds, which have been interpreted as fuel (Rosen 2005). Barley phytoliths have also been identified in small quantities on floors near storage bins both alone and with wheat husks in Building 5, suggested as remains of barley grain storage in one bin and alternating mixed storage in another. But it is also suggested that barley was entering the site primarily as a weed grass in dung and dung fuel from animal fodder due to its association with stem and leaf phytoliths and wild grass husks in Building 1, and the lack of correlation with wheat in Building 5 (Rosen 2005). Such problems demonstrate that interpreting the phytolith assemblage by itself can be problematic, without considering additional evidence that could narrow down possible interpretations.

As with ash deposits, the use of micromorphology can resolve some of these problems in building contexts. For example Matthews (2010) recognises that charred and phytolith remains of wheat husks are found together with burnt oven plaster fragments and may have been burnt as fuel that included sweepings from within buildings discarded in the hearths/ovens. Micromorphological observations have consistently shown a large variety of micro-contexts for phytoliths from cereals, reeds and grasses, including lime-burning, hearth rake out, animal dung and human coprolites (Matthews 2005, 2010; Shillito et al. 2008).

The origins of cereal agriculture

Cereal phytoliths are of particular importance, being central to questions concerning the origins, nature and timing of agricultural development in this region. Research at the early 9th millennium sites of Boncuklu in Turkey and Sheik-e Abad in Iran, for example, show very few if any indicators of cereal consumption in the charred macrobotanical record (Fairbairn personal communication; Matthews et al. 2010). Phytoliths offer a possible alternative way of investigating these remains, if they are present.

A number of studies have been carried out, particularly on the morphology of cereal husk phytoliths, which suggest that three characteristics of conjoined husks are distinguishable to the genus level in Triticum (wheat) and Hordeum (barley)—the number of pits in the papillae, the diameter of the papillae, and the wave pattern of the dendritic long cells (Rosen 1992). Rosen’s preliminary study (1992) examined husks (combining glume, lemma and palea) from specimens of T. dicoccum (emmer wheat), T. monococcum (einkorn), T. aestivum (bread wheat) and T. dicoccoides (wild emmer). The study determined that papilla sizes in domestic emmer ranged from 22 to 30 μm, with wild emmer ranging from 21 to 43 μm, and einkorn ranging from 25 to 50 μm. The study recognised that there is overlap between papilla sizes in these Triticum species. Rosen’s barley samples were from H. distichon (two row barley) and H. vulgare (6 row barley). The papilla sizes ranged from 18 to 25 μm for both species of barley. It is recognised that there is overlap in papilla measurements and the major difference Rosen notes between the two genera is the wave pattern of the long cells, described as square with even amplitude in barley, whilst wheat is more irregular. Both features therefore need to be used together.

A separate study by Tubb et al. (1993) investigated different cultivars of wheat and barley, and measured the width of 50 papillae, and the pit number of 100 papillae, for each species/cultivar. The study found major differences both within and between the Hordeum and Triticum genera. For example, of the 22 cultivars of T. aestivum studied, the average papilla widths (glume and lemma combined, \( n = 100 \)) varied from 16.6 ± 3.9 μm to 27.9 ± 4.6 μm. This demonstrates already the great variation even within a single species. Of archaeological relevance was that T. dicoccoides, wild emmer wheat and T. dicoccum, the cultivated form, could not be distinguished on the basis of these criteria. T. dicoccoides had average pit numbers of 8.9 ± 1.1 and papilla diameters of 19.6 ± 3.4 μm, whilst T. dicoccum had papilla diameters of 22.3 ± 3.3 μm and pit numbers 8.0 ± 0.9. Similarly, the eight H. vulgare cultivars studied had average pit number values ranging from 7.6 to 8.5 and papilla widths (glume + lemma combined) 14.1–17.0 μm. The values reported by Tubb et al. (1993) are all slightly smaller than values given for the same species by Rosen (1992). There also appeared to be some variation depending on whether the glume and lemma were analysed separately or together. It is worth noting that in archaeological samples, there is no way of
knowing whether husk phytoliths consist of glume cells, lemma cells, or a combination.

Ball et al. (1996) proposed that morphometric parameters of other phytolith types, including the dendritic cells, could also be used to distinguish between different species of wheat, using computer assisted image analysis. In Ball et al. (2009) a review is given of the full range of morphometric characteristics, including attributes such as roundness and convexity, which are difficult to assess visually, but which can be measured with image analysis software. Analysis using this complex range of measurements demonstrated that shape values were consistent within given taxa of wheat or barley, but size values demonstrated variance. Morphometric parameters do not seem to be routinely used in archaeological applications, which in general still use the average values for basic size measurements (Ishida et al. 2002; Emery-Barbier and Thiebault 2005; Tsartsidou et al. 2007).

It can be seen therefore that the basic criteria that have been established to identify wheat and barley are based on average values of size ranges which overlap in many cases. Although there appears to be a difference between wheat and barley in experimental samples, and a clear difference in the appearance of the wave pattern (Rosen 1992), there is little available data on other wild grasses. Rosen (1992) includes a limited discussion of Aegilops and Avena, and these genera have been used in other experimental studies (Tsartsidou et al. 2007) but no further morphological criteria have been published on them.

In order to match an ‘unknown’ archaeological sample to a particular category, enough papillae would need to be measured so that we could be statistically confident in the assignment. The discriminant functions described by Ball et al. (2009) again work well for known reference samples with large numbers of diagnostic elements, but transferring these to archaeological samples of unknown origin and composition, for example part of husks, mixed grass/cereal, is problematic. When we have overlaps in size ranges like this, it is likely that a larger number of papillae need to be assessed to see what the size range is within the individual being studied. This can be problematic with archaeological samples, which are often fragmented.

Both Rosen (1992) and Tubb et al. (1993) suggested that further work is needed to assess the full range of variation even within the Triticum and Hordeum genera, and also make the important point that further work is needed on inflorescence bracts from other species, which may overlap with the cereals. However such work has not been forthcoming and archaeological applications of the technique should be treated with caution, particularly in cases where other supporting evidence, such as charred macrobotanical remains, is lacking. The context of archaeological samples is also worth considering—if we have a very high concentration in a storage bin for example, an interpretation of cereal phytoliths is more convincing than small numbers from secondary deposition.

Cereal phytoliths as indicators of water availability and irrigation

In the Near East, a further point of interest has been the use of cereal phytoliths as proxies of water availability and irrigation. At Çatalhöyük, for example, the evidence from macrobotanical remains does not point clearly to a definitive wetland or dry land regime for cereal agriculture, with a weed seed flora being present from both environments, although a dry land regime is tentatively suggested (Fairbairn et al. 2002).

An experimental study of wheat by Rosen and Weiner (1994) suggested that conjoined phytoliths with more than 300 cells were an indicator of growth under conditions of high water availability in arid and semi-arid regions. This hypothesis is based on the idea that greater water availability means that plants have a greater rate of water and silica uptake, which allows formation of larger multi-celled phytoliths. The size criteria established by Rosen and Weiner (1994) have been applied to archaeological sites to suggest the presence or absence of irrigation agriculture (Katz et al. 2007). Interpretations have subsequently been used by other archaeologists seeking to understand spatial variation in settlement patterns and agricultural activity (Winter Livneh et al. 2010). At the Neolithic site of Çatalhöyük, believed to have been situated in a wetland area, the method has been used to suggest that wheat was growing under dry land conditions, and thus must have been brought to the site from several kilometres away (Roberts and Rosen 2009).

The use of conjoined phytolith size is problematic, firstly due to the context of the phytoliths and taphonomic processes, and secondly due to the impact of processing methods. A study of modern wheat grown under controlled conditions showed that the method of sample processing had a significant impact on the size of conjoined wheat husks (Jenkins 2009), and that husk size varied considerably on an annual basis even for plants growing under the same conditions, with phytoliths in excess of 800 cells being observed (Jenkins et al. 2011). Observations of archaeological wheat in situ in thin section have also suggested some of the possible taphonomic processes affecting the size of phytoliths such as trampling, sediment compaction, and bioturbation (Shillito 2011a). Madella et al. (2009) combined morphological and X-ray analyses of water molecules trapped within the wheat phytoliths, and demonstrate an association between chemical bonding in the opaline silica and water availability during plant growth. This combination of methods is perhaps more convincing than a single method approach.
Phytoliths—a summary of the common problems and possible solutions

A review of three applications of phytolith analysis has revealed some common problems that need to be addressed. Firstly, before we even approach archaeological samples, a clearer case needs to be made that certain phytolith types are really diagnostic. This requires objective classification of phytolith geometric morphometrics, using as large a population size as possible. This will enable researchers to assess how much is a component of random variation and how much is genus and species (and subspecies) specific. This will enable a clearer assessment of the particular circumstances in which phytoliths are suitable for supporting different hypotheses, and the probability values associated with an interpretation. Secondly, relative rates of phytolith production and the possible controlling factors on this should be addressed systematically. Thirdly, with archaeological samples it seems that taphonomy is often poorly understood or not considered and that inappropriately small numbers of phytoliths are used to support interpretations. Depositional and post-depositional impacts on size, shape and assemblage composition need further investigation.

Phytolith systematics—can individual cells be diagnostic of species?

Phytolith systematics needs to be addressed with urgency. At present morphological analysis, as with other areas of systematics, are carried out simply by visual examination by trained specialists, with conclusions being based on the number and quality of characteristic types observed during their careers (MacLeod 2005). In palaeontology, this type of analysis has been shown to have very low reproducibility. MacLeod argues for the adoption of a geometric paradigm. This approach has been investigated for over a decade in palynology, using automated identification and classification methods (France et al. 2000; Zhang et al. 2004; Li et al. 2004). More recently this has been applied to starch (Wilson et al. 2010), though it is noted that neither discipline has fully adopted these methods.

Despite the criticisms of ‘subjective identification’ by expert analysis, it is acknowledged that visual identification can be accurate to some extent, and that image analysis is not always feasible. Comparisons between image analysis and visual identification in medical research for example have shown that the latter can be reliable (Parker et al. 2008). Perhaps some of the criticisms of phytolith analysis could be addressed by such method validation experiments?

It is suggested that further caution should be applied to identifications based on visual examination. A major concern with phytolith evidence is that current statistical tests offer only a probability that a certain assemblage is domesticated rather than wild, rather than being able to say that one phytolith type is unique to one species or another. Such methods are only acceptable if researchers acknowledge the probability values during interpretation.

Factors controlling phytolith production

It has been suggested that the absolute production of phytoliths per unit weight of dried plant is genetically determined (Piperno 1988; Tsartsidou et al. 2007) and that plants producing large amounts of phytoliths will do so whatever their geographical region. However, this assumption has not been empirically tested and is based on qualitative comparisons between plants where no other variables have been considered. It is recognised that other factors such as soil substrate, temperature, water availability, pH and climate may also affect phytolith production, but these tend to be overlooked as being of less significance (Tsartsidou et al. 2007). Recent experimental studies suggest that phytolith formation is complex and that these ‘secondary’ factors are poorly understood (Mithen et al. 2008; Jenkins et al. 2011). A particular problem is how this translates into an archaeological signature under different depositional processes. Whereas for pollen we have estimates of the relative abundance of pollen for different taxa, with phytoliths this is only known on a crude basis. The rate of production needs to be considered especially when using the number per gram approach (Albert and Weiner 2001).

Morphological changes

When we get to the point where cellular variation within and between taxa has been properly addressed, further consideration needs to be given to the possibility of post-depositional morphological changes. Considering that phytolith identifications rely on very specific, sometimes subtle, morphological diagnostic criteria, the possibility of morphological changes is an essential area of investigation. It is known for example that phytoliths can dissolve at pH > 8.5 (Bartoli and Wilding 1980; Piperno 2006; Erlich et al. 2010). Breakage of conjoined forms has been established (Jenkins 2009; Shillito 2011a), and weathering of surface features of single cells has also been observed (Zucol et al. 2005; Osterrieth et al. 2009; French et al. 2009).

Morphological changes may also occur prior to deposition. For example, phytoliths in dental calculus have been used as dietary indicators (Henry et al. 2011). Although this is a ‘secure’ context, there has been no work to assess the impacts of mastication on phytolith morphology. There have been many studies into possible phytolith microwear...
patterns on teeth, and a recent study has suggested that in fact enamel is more likely to erode the phytoliths (Sanson et al. 2007). This calls into question interpretations such as that of Henry et al. (2011), who give specific identifications of plants in Neanderthal diet, based on seven poorly preserved phytoliths from dental calculus.

Stratigraphic mixing/movement and potential contamination

Stratigraphic mixing and moving is a concern when trying to relate a phytolith deposit to radiocarbon dates. Although Piperno (1984, 1985, 2006) discusses this and states that it is not a problem, it is not clear how widely applicable this conclusion is for different contexts. Piperno’s analysis suggests for example that the absence of phytoliths beneath house floors, storage pits and a burial area show that they do not move (Piperno 1985), and uses data from soil scientists in temperate areas to suggest that there is no significant movement in soils. Again we have the difference here between ‘sealed’ contexts such as deposits under floors, which micromorphology shows can have minimal post-depositional alteration, and soil deposits. Post-depositional processes in the latter are highly varied depending on the soil type, and the scale at which analysis is carried out. It is often assumed that the depositional nature of phytoliths is always from in situ decay (Friarte 2003; Itzstein-Davey et al. 2007). However this needs to be investigated separately for each study.

A study of bioturbation in an early modern burial complex in Thailand, combining micromorphology and phytolith analysis, suggested that although there was considerable bioturbation, the effects were localised, and a combination of the two techniques should be used to assess the extent to which vertical movement has occurred (Grave and Kealhofer 1999). Recent studies have investigated stratigraphic movement of phytoliths (Osterrith et al. 2009), and have shown that under some circumstances phytoliths are highly mobile in soil. One study of Phragmites australis (common reed) phytoliths showed an average of 4 cm movement after one year in cambisols and luvisols (Fishkis et al. 2010a, b). In a study of sandy sediment under conditions simulating high rainfall, 22% of phytoliths were leached from their layer of application (Fishkis et al. 2009).

Also to consider is the problem of potential contamination. It has long been recognised that phytoliths can be transported by wind (Darwin 1846), and when dealing with ash deposits it is likely that spreading and mixing of ash will have occurred. This ‘background noise’ is recognised at Çatalhöyük by Rosen (2005) who proposes that only significant peaks of particular types in an assemblage should be used as indicators of activity. However this does not seem to be considered in other studies.

Assumptions should not be made without very good stratigraphic controls, and it is not sufficient to assume that an assemblage is from in situ decay, unless supported by micro-contextual analysis such as thin section micromorphology (Goldberg et al. 2009; Albert et al. 2010; Shillito 2011b). This technique could also provide the stratigraphic control that would dispel some of the arguments against phytolith interpretations, and provide a direct demonstration of the depositional and post-depositional processes. Distinguishing between multiple depositional pathways can also be achieved by examining associated diagnostic indicators, such as spherulites to distinguish animal dung (Albert et al. 2008; Portillo et al. 2009).

Conclusions—what currently can and cannot be done with phytoliths in archaeology?

The abundance of phytoliths at many sites, and the lack of charred remains in many contexts, means that these microfossils have much potential as archaeological tools, and their study offers a hugely valuable technique for understanding human use of plant resources in prehistory. Phytolith analysis is still an evolving discipline, and like all relatively new approaches it has problems which will hopefully be addressed over time. Work on modern reference materials indicates that the assemblages do vary between different genera and in some cases species, but the degree and causes of variation even within an individual plant need to be fully understood before we can apply this to archaeological samples and make definitive conclusions. The conflict with other lines of evidence is a concern that clearly needs to be resolved. The problems can be summarised as (a) potential subjectivity in developing and applying morphological criteria (b) problems applying criteria from modern assemblages to archaeological assemblages (c) poorly understood taphonomy. Until these concerns are addressed satisfactorily, caution should be taken when making definitive statements based on phytolith analysis. Researchers need to be more realistic over what can and cannot be concluded on the basis of the data, and to avoid over-interpretation.

The discipline would also benefit from the publication of images of reference specimens and archaeological assemblages in addition to numerical tables, so that different datasets can be compared with more confidence. Although Piperno’s studies alone have analysed over 500 wild species and 40 maize varieties, no database of this material is available for comparison. The development of geometric morphometric methods and automated image analysis can only be improved with very large datasets, which would provide a robust dataset of benefit to researchers worldwide, which could continue to be improved and modified as more material is analysed.
Acknowledgments Thanks to the editor and three anonymous reviewers for providing helpful comments on this paper.

References


## References

Matthews W (2010) Geoarchaeology and taphonomy of plant remains and micromorphological residues in early urban environments in the ancient near east. Quat Int 214:98–113


Zhao Z, Pearsall DM, Bentfer RA Jr, Piperno DR (1998) Distin-
guishing rice (Oryza sativa, Poaceae) from Oryza species through phytolith analysis. J Archaeol Sci 35:526–531


Zhao Z, Pearsall DM, Bentfer RA Jr, Piperno DM (1998) Distin-
guishing rice (Oryza sativa, Poaceae) from Oryza species through phytolith analysis. J Archaeol Sci 35:526–531


Springer