Morphometrics of Starch Granules From Sub-Saharan Plants and the Taxonomic Identification of Ancient Starch

Julio Mercader*, Matthew Abtosway, Robert Bird, Mariam Bundala, Siobhan Clarke, Julien Favreau, Jamie L. Inwood, Makarius Itambu, Fergus Larter, Patrick Lee, Robert Patalano, Maria Soto, Laura Tucker and Dale Walde

Department of Anthropology and Archaeology, University of Calgary, Calgary, AB, Canada

The assumption that taxonomy can be ascertained by starch granule shape and size has persisted since the late nineteenth and early twentieth century biochemistry. More recent work has established that granule morphological affinity is scattered throughout phylogenetic branches, morphotype proportions vary within the genus, granules from closely related genera can differ dramatically in shape, and size variations do not reflect phylogenetic relationships. This situation is confounded by polymorphism at the species and tissue level, resulting in redundancy and multiplicity. This paper classifies morphological features of starch granules from 77 species, 31 families, and 22 orders across three African ecoregions. This is the largest starch reference collection published to date, rendering the dataset uniquely well-suited to explore (i) the diagnostic power of unique morphometric classifiers and their frequency, (ii) morphotypes that cut across taxonomic boundaries, and (iii) issues surrounding the minimum counts needed to accurately reflect granule polymorphism, variability, and identification. In a collection of 23,100 granules, taxonomic identification occurred very rarely. In the instances it did, it was at the species level, with no occurrences of a single morphotype or complement identifying all species within a family or genus. Some families cannot be uniquely identified, and morphometric types are shared despite taxonomic distance for three quarters of the taxa. However, this reference collection boasts 98 unique identifiers located in the Arecaceae, Convolvulaceae, Cyperaceae, Dioscoreaceae, Fabaceae, Musaceae, Pedaliaceae, Poaceae, and Zamiaceae.

Keywords: starch granule morphometry, reference collection, starch identification, ancient starch research, sub-saharan ethnobotany, unique identifiers

INTRODUCTION

For decades, scientists have relied on the comparative study of starch granules from reference collections to infer phylogenetics and species identification for both fresh and ancient materials (Cortella and Pochettino, 1995; Parr, 2002; Langejans, 2006; Lentfer, 2009; Yang et al., 2012; Musaubach et al., 2013; Hart, 2014; Louderback et al., 2016). The assumption that taxonomy can be ascertained by granule shape and size dates back to late nineteenth and early twentieth century

biochemistry (Nageli et al., 1858; Meyer, 1895; Reichert, 1913; but see Czaja, 1978; Jane et al., 1994). Recently, Matsushima et al. (2013) have established that granule morphological affinity is scattered throughout phylogenetic branches, morphotype proportions vary within the genus, granules from closely related genera can differ dramatically in shape, and size variations do not reflect phylogenetic relationships (Shapter et al., 2008). Moreover, polymorphism at the species and tissue level yields a marked redundancy and multiplicity (Parr, 2002; Langejans, 2006).

Taxonomic identification of ancient starch granules often underpins reconstructions of human palaeoecology and ethnobotany, but remains problematic. To paraphrase Copeland and Hardy (2018), plant identification usually involves the visual comparison of individual archeological granules with modern reference material using a method that was pioneered over 100 years ago; however, this approach has major limitations as the inspection of images of starch granules from multiple plant sources reveals considerable similarities in morphology between and within species. While some plant genera do produce morphologically characteristic starch granules, there is significant overlap between many species. A very limited number of studies have tackled automated systems of identification (Torrence et al., 2004; Wilson et al., 2010; Coster and Field, 2015; Arráiz et al., 2016). And although the degree of subjectivity and successful classification rates from visual vs. computerized identification are controversial, direct and automated comparisons both face three unsolved methodological problems (Mercader et al., 2018):

i)   The unreliability of small reference collections.
ii) The frequency of unique morphometric classifiers in a pool of targets, and issues surrounding minimum counts needed to accurately reflect granule polymorphism and variability per taxon.
iii) The existence of many morphotypes that cut across taxonomic boundaries.

This paper is a methodological contribution that examines the accuracy of using starch granule morphotypes and morphometrics to identify plant of origin and the chances for retrieval of unique identifiers: a diagnostic type is defined here as one that allows for differentiation of the members of a certain species from other taxa. Blind testing was not conducted, as the present study was aimed at establishing criteria for future work. Moreover, very few atlases or reference collections for starch identification have been published to date (Nageli et al., 1858; Meyer, 1895; Reichert, 1913; Czaja, 1978; Jane et al., 1994; Lindeboom et al., 2004). It is important for the discipline to publish these reference collections, so that researches working in this area can access the classification, morphometric analysis, and images of a range of materials that might be present in archeological assemblages. This work contributes *circa* 4,300 images of starch granules from taxa commonly exploited by Indigenous societies in Sub-Saharan Africa. It is the largest starch reference collection published to date, classifying 23,100 granules and morphological features from 77 species, 31 families, and 22 orders across the angiosperms and in one basal gymnosperm

Order	Families	Part Sampled
Lamiales	Lamiaceae (2) Pedaliaceae (1)	USO Seed
Gentianales	Apocynaceae (2)	USO
Solanales	Convolvulaceae (1)	USO
Ericales	Sapotaceae (1)	Fruit
Caryophyllales	Polygonaceae (1)	USO
Malvales	Bombacaceae (1)	Mesocarp
Myrtales	Combretaceae (1)	Legume
Cucurbitales	Cucurbitaceae (2)	Seed
Rosales	Cannabaceae (1)	Fruit
Fabales	Fabaceae (16)	Legume USO
Malpighiales	Euphorbiaceae (2)	Seed
Proteales	Proteaceae (1)	USO
Ranunculales	Menispermaceae (1)	USO
Zingiberales	Musaceae (1) Strelitziaceae (1)	USO
Poales	Cyperaceae (3) Poaceae (4) Typhaceae (1)	USO Grain
Arecales	Arecaceae (1)	Trunk
Asparagales	Amaryllidaceae (1) Asphodelaceae (1) Hypoxidaceae (3) Iridaceae (3) Orchidaceae (1)	USO
Pandanales	Pandanaceae (1)	Trunk
Dioscoreales	Dioscoreaceae (14) Taccaceae (1)	USO
Alismatales	Araceae (2)	USO
Nymphaeales	Nymphaeaceae (1)	USO
Cycadales	Zamiaceae (5)	Trunk

FIGURE 1 | Family phylogenetic grouping of studied samples classified from derived to primitive (top to bottom). USO, Underground Storage Organ.

lineage, presented in phylogenetic order from primitive to derived (**Figure 1**). Selection followed ethnobotanical criteria in three African ecoregions: the Guineo-Congolian, Zambezian, and Somalia-Masai zones.

MATERIALS AND METHODS

The botanical samples originated from 13 Sub-Saharan countries that fall within three phytochoria (**Table 1**) (White, 1983), with about three-quarters of the analyzed specimens native to the Zambezian woodlands of Mozambique and northern South Africa. With one exception, all specimens came from wild species. Due to the human ecology focus, species and materials were selected for their ethnobotanical significance. More specifically, we prioritized plants used as food and medicine. As a result, almost half of the species fall within two families (Fabaceae and Dioscoreaceae), as they play a key role in Indigenous utilization
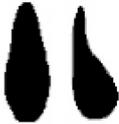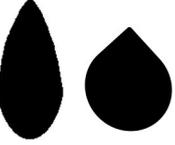
TABLE 1 | Sample provenance.

Species	Provenance	Species	Provenance
Acacia nilotica	South Africa	*Eriosema preptum*	South Africa
Acacia tortilis	Tanzania	*Faidherbia albida*	Mozambique
Acanthosicyos naudinianus	South Africa	*Fockea angustifolia*	South Africa
Adansonia digitata	Mozambique	*Gladiolus gueunzii*	South Africa
Afzelia quanzensis	Mozambique	*Gladiolus* spp.	Mozambique
Amaryllis belladonna	South Africa	*Gladiolus undulatus*	South Africa
Bulbine sedifolia	South Africa	*Hyphaene petersiana*	South Africa
Cassia petersiana	Mozambique	*Hypoxis nyasica*	Mozambique
Celtis africana	South Africa	*Hypoxis rigidula*	South Africa
Ceropegia rendalii	South Africa	*Hypoxis zeiheri*	South Africa
Cajanus cajan	Mozambique	*Ipomoea longituba*	Tanzania
Combretum spp.	Mozambique	*Jateorhiza palmata*	Ethiopia
Cordyla africana	South Africa	*Lablab purpureus*	South Africa
Cyperus esculentus	India	*Lablab* spp.	Mozambique
Cyperus rotundus	Tanzania	*Lotus discolor*	South Africa
Cyperus tomaiophyllus	Cameroon	*Momordica cardiospermoides*	South Africa
Dioscorea alata	Cameroon	*Mucuna pruriens*	Mozambique
Dioscorea bulbifera	South Africa	*Mucuna* spp.	Equatorial Guinea
Dioscorea burkilliana	Gabon	*Nymphaea caerulea*	Botswana
Dioscorea cayenensis	South Africa	*Oxytenanthera abyssinica*	Mozambique
Dioscorea diversifolia	South Africa	*Pandanus livingstonianus*	Mozambique
Dioscorea dumetorum	Gabon	*Plectranthus esculentus*	Mozambique
Dioscorea mangenotiana	Central African Republic	*Polygonum senegalense*	South Africa
Dioscorea praehensillis	Cameroon	*Protea welwitschii*	South Africa
Dioscorea rupicola	South Africa	*Ricinodendron heudelotii*	Equatorial Guinea
Dioscorea sansibarensis	Central African Republic	*Ricinodendron rautanenii*	Namibia
Dioscorea schimperiana	Cameroon	*Sesamum angolense*	Mozambique
Dioscorea smilacifolia	Central African Republic	*Solenostemon rotundifolius*	South Africa
Dioscorea spp.	Mozambique	*Sorghum bicolor*	Mozambique
Dioscorea sylvatica	South Africa	*Sphenostylis stenocarpa*	Malawi
Disa spp.	Lesotho	*Sporobolus panicoides*	South Africa
Eleusine africana	Tanzania	*Strelitzia nicolai*	South Africa
Encephalartos friderici-guilielmi	South Africa	*Tacca leontopetaloides*	Zimbabwe
Encephalartos laevifolius	South Africa	*Typha capensis*	South Africa
Encephalartos natalensis	South Africa	*Vigna lobatifolia*	South Africa
Encephalartos princeps	South Africa	*Vigna subterranea*	Mozambique
Encephalartos senticosus	South Africa	*Zantedeschia aethiopica*	South Africa
Englerophytum magalismontanum	South Africa	*Zantedeschia rehmannii*	South Africa
Ensete ventricosum	South Africa		

of natural resources (Van Wyk and Gericke, 2000; Laden and Wrangham, 2005; Maundu and Tengnas, 2005; Williamson, 2005; Dharani and Yenesew, 2010). Taxonomic grouping and nomenclature come from the index of accepted names for the flowering plants of Sub-Saharan Africa (Klopper et al., 2006). Samples were collected from herbaria (South African National Biodiversity Institute, Tropical Pesticides Research Institute of Tanzania) and in the field, in which case we relied on identification provided by local botanists. When species level identification of a specimen was uncertain and could belong to several species within the genus, we listed it as spp.

The method used to observe native starches is the extraction by powdering and suspension in aqueous media for light microscopy. Microscopic observations were conducted with two apparatuses: Motic BA410E (Photomicrography, Moticam 5+ camera, MoticImagesPlus 3.0) and Olympus BX51 (Photomicrography, SC50 camera, cellSens Standard). Granules were measured to an accuracy of 0.01 μm. The discovery and elimination process needed to find unique identifiers required counts per taxon >300 to ensure that variability and proportions are within 5% of the true value (counts of 500 achieve 95% confidence: Thompson, 1992). We did not use or follow the so-called "International Code for Starch Nomenclature" (http://fossilfarm.org/ICSN/Code.html), as this system is not peer reviewed and rarely utilized. We followed instead a geometric morphometric typology

TABLE 2 | Starch granule typology.

3D classifier	3D name	3D description	2D	Schematic outline
Conoid (Co)	Clavate	Club-shaped; gradually thickening from slender base	Irregular (Irr) Oblong Elongate (OE) Ovate (Ov) Triangular (Tri)	
	Conical, Sub	Overall triangular shape with inflection and tapering edge, ending in a cone-shape	Irregular (Irr) Triangular (Tri)	
	Composite Hemisphere/Cone	Hemisphere base with narrow, conical tip; rounded base	Cuneiform (Cu) Irregular (Irr)	
	Fusiform	Elongate with tapering ends	Fusiform (F) Irregular (Irr)	
	Guttiform	Drop-shaped; large, bulbous base with small, narrow tip	Irregular (Irr) Oblong Elongate (OE) Orbicular (Orb)	
	Tower	Inverted hemisphere base with conical projection; flat base	Bell (B) Irregular (Irr)	
Cylindroid (Cy)	Cylindroid	Elongate with circular cross-section	Elliptical (Ell) Irregular (Irr) Oblong Elongate (OE) Ovate (Ov) Trapeziform (Tra)	
Globular (G)	Globular	Spherical	Irregular (Irr) Orbicular (Orb) Trapeziform (Tra)	
Pear-Shaped (PS)	Pear-Shaped	Globular base with slightly narrower tip	Irregular (Irr) Ovate (Ov) Shell (S)	
Hemisphere (H)	Hemisphere	Half sphere	Irregular (Irr) Orbicular (Orb)	
Irregular (Irr)	Irregular	Irregular	Irregular (Irr)	

(Continued)

TABLE 2 | Continued

3D classifier	3D name	3D description	2D	Schematic outline
Lenticular (Le)	Lenticular Biconvex	Narrow edges, wide in middle	Fusiform (F) Irregular (Irr) Oblong Elongate (OE) Shell (S)	
	Lenticular Planoconvex	Lens-shaped; flat side with opposite side slightly rounded, but not forming full half-sphere	Irregular (Irr) Orbicular (Orb)	
Lobate (Lo)	Lobate	Having lobes	Irregular (Irr) Shell (S)	
Parabolic Prism (PP)	Parabolic Prism	In the shape of a parabola; sides either parallel or slightly diverging. Sub-tabular in cross-section	Irregular (Irr) Oblong Elongate (OE) Parabolic (P) Shell (S) Trapeziform (Tra) Triangle (Tri)	
Prismatic (Pri)	Orthogonal	Many sides; orthogonal both in above-view and cross-section	Irregular (Irr) Polygonal (Pol)	
	Quadrangular Prism	Four-sided; tabular in cross-section	Irregular (Irr) Oblong Elongate (OE) Quadratic (Q)	
Reniform (R)	Reniform	Bean-shaped	Irregular (Irr) Reniform (R)	

that includes 50 possible morphotypes (**Table 2**), which can be broken down in 11 major classes, 18 three-dimensional descriptors, and 49 two-dimensional outlines. As is standard in the morphological classification of starch granules, we studied the position of the hylum, multiple textural features, faceting, slit presence, depressions, grooves, vacuoles, and fissures. These characteristics can be found in the original quantification matrix. **Appendix** shows the studied taxa and plant part (fruit: 3%, grain: 4%, legume: 19%, mesocarp:1%, seed: 6%; trunk: 9%, Underground Storage Organ–USO: 58%); as well as images of the granules. Morphotype richness per taxon is in **Table 3**. Elimination isolated primary morphometric types, defined as 3D/2D combinations and maximum length range. There are two kinds of unique identifiers: one combines morphology and size with volumetric and textural features; the other is restricted to only morphometry. The original quantitative matrix, recording all observations and cases as well as all photomicrographs taken for granule measures, is with the Federated Research Data Repository (doi: 10.20383/101.017). In addition, a preprint of this article has been uploaded to the Open Science Framework and is available online (doi: 10.17605/OSF. IO/MF6GH).

FAMILY MORPHOMETRICS AND IDENTIFIERS

Zamiaceae

Represented by trunk samples from five *Encephalartos* species, with a morphotype complement broken down in **Supplementary Table 1**. The relative morphotype proportions differ within the family, but nil or low counts prevent valid statistical tests of difference. *Encephalartos friderici-guilielmi* is

TABLE 3 | Lead morphotypes and richness per species.

Family	Taxon	Morphotype richness	Lead morphotype
Amaryllidaceae	Amaryllis belladonna	6	PS-Ov
Apocynaceae	Ceropegia rendalii	3	G-Orb
Apocynaceae	Fockea angustifolia	5	G-Orb
Araceae	Zantedeschia aethiopica	4	G-Tra
Araceae	Zantedeschia rehmannii	2	O-Pol
Arecaceae	Hyphaene petersiana	9	Cy-El
Asphodelaceae	Bulbine sedifolia	1	Cy-El
Bombacaceae	Adansonia digitata	1	G-Orb
Cannabaceae	Celtis africana	1	G-Orb
Combretaceae	Combretum spp.	3	Cy-El
Convolvulaceae	Ipomoea longituba	1	G-Orb
Cucurbitaceae	Acanthosicyos naudinianus	3	G-Orb
Cucurbitaceae	Momordica cardiospermoides	4	Cy-El
Cyperaceae	Cyperus esculentus	3	PS-Ov
Cyperaceae	Cyperus rotundus	3	Lo-Irr
Cyperaceae	Cyperus tomaiophyllus	5	G-Orb
Dioscoreaceae	Dioscorea alata	5	PP-P
Dioscoreaceae	Dioscorea bulbifera	3	PP-P
Dioscoreaceae	Dioscorea burkilliana	1	PP-P
Dioscoreaceae	Dioscorea cayenensis	2	PP-T
Dioscoreaceae	Dioscorea diversifolia	3	Cy-El
Dioscoreaceae	Dioscorea dumetorum	3	Cy-El
Dioscoreaceae	Dioscorea mangenotiana	1	PP-P
Dioscoreaceae	Dioscorea praehensillis	13	PP-P
Dioscoreaceae	Dioscorea rupicola	3	Cy-OE
Dioscoreaceae	Dioscorea sansibarensis	6	PP-P
Dioscoreaceae	Dioscorea schimperiana	7	PP-P
Dioscoreaceae	Dioscorea smilacifolia	4	PS-Ov
Dioscoreaceae	Dioscorea spp.	5	PP-P
Dioscoreaceae	Dioscorea sylvatica	4	Co-Cu
Euphorbiaceae	Ricinodendron heudelotii	3	G-Orb
Euphorbiaceae	Ricinodendron rautanenii	3	G-Orb
Fabaceae	Acacia nilotica	1	G-Orb
Fabaceae	Acacia tortilis	2	Cy-El
Fabaceae	Afzelia quanzensis	1	G-Orb
Fabaceae	Cassia petersiana	1	Cy-El
Fabaceae	Cajanus cajan	7	Cy-El
Fabaceae	Cordyla africana	4	H-Orb
Fabaceae	Eriosema preptum	13	G-Orb
Fabaceae	Faidherbia albida	4	Cy-El
Fabaceae	Lablab purpureus	2	Cy-El
Fabaceae	Lablab spp.	8	Cy-El

(Continued)

TABLE 3 | Continued

Family	Taxon	Morphotype richness	Lead morphotype
Fabaceae	Lotus discolor	2	G-Orb
Fabaceae	Mucuna pruriens	5	PS-Ov
Fabaceae	Mucuna spp.	2	Cy-El
Fabaceae	Sphenostylis stenocarpa	3	G-Tra
Fabaceae	Vigna lobatifolia	7	Cy-El
Fabaceae	Vigna subterranea	6	G-Orb
Hypoxidaceae	Hypoxis nyasica	2	G-Orb
Hypoxidaceae	Hypoxis rigidula	2	G-Orb
Hypoxidaceae	Hypoxis zeiheri	2	G-Orb
Iridaceae	Gladiolus gueunzii	3	G-Tra
Iridaceae	Gladiolus spp.	3	G-Tra
Iridaceae	Gladiolus undulatus	4	H-Orb
Lamiaceae	Plectranthus esculentus	3	G-Tra
Lamiaceae	Solenostemon rotundifolius	3	G-Tra
Menispermaceae	Jateorhiza palmata	7	Cy-El
Musaceae	Ensete ventricosum	12	Co-Irr
Nymphaeaceae	Nymphaea caerulea	4	G-Orb
Orchidaceae	Disa spp.	3	G-Orb
Pandanaceae	Pandanus livingstonianus	2	G-Orb
Pedaliaceae	Sesamum angolense	1	Lo-Irr
Poaceae	Eleusine africana	4	O-Pol
Poaceae	Oxytenanthera abyssinica	3	G-Tra
Poaceae	Sorghum bicolor	1	Pri-Pol
Poaceae	Sporobolus panicoides	3	G-Orb
Polygonaceae	Polygonum senegalense	1	G-Orb
Proteaceae	Protea welwitschii	2	G-Orb
Sapotaceae	Englerophytum magalismontanum	3	H-Orb
Strelitziaceae	Strelitzia nicolai	4	Co-F
Taccaceae	Tacca leontopetaloides	3	G-Orb
Typhaceae	Typha capensis	4	G-Orb
Zamiaceae	Encephalartos friderici-guilielmi	8	G-Orb
Zamiaceae	Encephalartos laevifolius	5	G-Orb
Zamiaceae	Encephalartos natalensis	5	G-Orb
Zamiaceae	Encephalartos princeps	3	G-Tra
Zamiaceae	Encephalartos senticosus	1	G-Tra

uniquely identified by Globular–Irregulars with a centric hilum and marked lamellae ($n = 1$), as well as Lenticular–Orbiculars ($n = 46$). Morphometrically, Globular–Trapeziforms with a maximum length range of 41.00–46.86 µm ($n = 4$) uniquely identify the family, while the species *Encephalartos laevifolius* is

uniquely identified by Hemisphere–Orbiculars with a maximum length range of 36.74–38.10 μm ($n = 2$).

Nymphaeaceae

Represented by the USO of a single species: *Nymphaea caerulea*. The complement has Cylindroid–Ellipsoidals ($n = 24$), Globular–Orbiculars ($n = 265$), Globular–Trapeziforms ($n = 7$), and Hemisphere–Orbiculars ($n = 4$). The Cylindroid–Ellipsoidals with centric hilum and porous surface ($n = 1$) and the Cylindroid–Ellipsoidals with centric hilum and psilate texture, with maximum length between 19.05 and 19.58 μm ($n = 2$), have diagnostic power. However, if the hilum visibility and textural characteristics were not preserved, the species would not have unique identifiers.

Araceae

Represented by USOs of two *Zantedeschia* species. Morphotypes in these taxa differ significantly in content and proportion for the species *aethiopica* and *rehmannii*. Overall, the complement is Globular–Orbicular [$n = 100/8$ (Z.a./Z.r.)], Globular–Trapeziform ($n = 172/0$), Hemisphere–Orbicular ($n = 27/0$), Hemisphere–Trapeziform (1/0), and Prismatic–Polygonal (0/292). *Zantedeschia aethiopica* is uniquely identified by a Hemisphere–Trapeziform morphotype with centric hilum, psilate surface, and a faceted orbicular volume ($n = 1$).

Zantedeschia rehmannii is uniquely identified by a Prismatic–Polygonal morphotype with centric hilum, psilate surface texture, and a maximum length range of 3.68–12.22 μm ($n = 292$). If we only had morphometric information and no textural or additional volumetric data, *Zantedeschia aethiopica* is still uniquely identifiable by Hemisphere–Trapeziforms with a length of 9.48 μm ($n = 1$), while *Zantedeschia rehmannii* is uniquely identifiable by Prismatic–Polygonals with a maximum length range of 3.68–5.31 μm ($n = 24$).

Taccaceae

Represented by the USO of *Tacca leontopetaloides*. The complement consists of Globular–Orbicular ($n = 209$), Globular–Trapeziform ($n = 39$), and Hemisphere–Orbicular ($n = 52$). It is uniquely identified by Hemisphere–Orbiculars with centric hilum, creased surface texture, and faceting, with a maximum length range of 12.09–25.81 μm. If degradation was to delete textural and volumetric characteristics, then there would not be unique identifiers for this species.

Dioscoreaceae

Represented by USOs from 14 *Dioscorea* species with variable complements and proportions (**Table 4**), although most assemblages are dominated by Parabolic Prism–Prisms [cf. *Dioscorea cayenensis* ($n = 6$), *D. diversifolia*, *D. dumetorum*, *D. rupicola*, *D. smilacifolia*, and *D. sylvatica* ($n = 0$)]. This very large sample produced many unique identifier morphotypes for either individual or multiple species. When only morphometrics are considered, the Dioscoreaceae are uniquely identified by Conoid–Triangular [maximum length >26.00 μm ($n = 11$)], Lobate–Shell [maximum length 35.71–47.05 μm ($n = 2$)], Parabolic Prism–Irregular ($n = 39$), Parabolic Prism–Oblong

Elongate ($n = 22$), Parabolic Prism–Parabolic–Parabolic [maximum length <14.90 μm ($n = 22$) >53.9 μm ($n = 195$)], Prism–Shell ($n = 89$), Parabolic Prism–Trapeziform ($n = 3$), and Parabolic Prism–Triangular ($n = 303$).

Pandanaceae

Represented by specimens from the trunk of *Pandanus livingstonianus*. Morphotypes include Globular–Orbicular ($n = 289$), and Globular–Trapeziform ($n = 11$). All granules have centric hilum and psilate texture. Faceting is common. There are no unique morphometric identifiers within this species.

Orchidaceae

Represented by granules from the USO of *Disa* spp. with types that include Globular–Orbicular ($n = 272$), Globular–Trapeziform ($n = 16$), and Hemisphere–Orbicular ($n = 12$). All granules have centric hilum and psilate texture, with some faceting ($n = 214$). There are no unique morphometric identifiers within this species.

Iridaceae

Represented by USOs from three *Gladiolus* species. The complement proportions amongst the species (**Supplementary Table 2**) are significantly different (Chi–square = 201.623, df = 6, $p = 0.000$), likely to result from the proportions seen in *G. undulatus*. However, this significance remains when *G. undulatus* is removed, confirming significant differences in the complement (Chi–square = 12.815, df = 2, $p = 0.002$). *G. undulatus* is uniquely identified by Hemisphere–Orbiculars with a centric hilum, creased surface texture, and with maximum length 34.00–36.00 μm ($n = 2$). This species is also uniquely identified by Hemisphere–Orbiculars with centric hilum, psilate surface, centric slit (–), and maximum length 26.82–26.92 μm ($n = 2$). *G. spp.* is uniquely identified by Hemisphere–Orbiculars, centric hilum, psilate surface, and centric slit (v) ($n = 35$). Removing volumetric and textural features with low taxonomic value would imply that there are no unique identifiers within this family.

Hypoxidaceae

Represented by USOs of three *Hypoxis* species (**Supplementary Table 3**). Most granules share features with other families: Globular–Orbicular with centric hilum, psilate texture, faceted/non–faceted; Globular–Orbicular with centric hilum, faint lamellae, faceted/non–faceted, and; Cylindroid–Ellipsoidal type with centric hilum, psilate texture, though sometimes with creasing or faint lamellae, and weak birefringence. Morphotype proportions in *H. nyasica* differ from those in *H. rigidula* and *H. zeiheri* (Chi–square = 56.379, $p = 0.000$. df = 2), but these two latter species do not differ significantly between themselves (Chi–square = 0.293, $p = 0.394$, df = 1). If morphometry is considered exclusively, then there are no unique identifiers within this family.

Asphodelaceae

Represented by the USO of *Bulbine sedifolia*. All granules belong to a widely shared identifier [Cylindroid–Ellipsoidal (eccentric

TABLE 4 | Dioscoreaceae morphotypes.

Species	Co-Cu	Co-Tri	Cy-OE	Cy-El	Cy-Ov	G-Orb	Lo-Irr	Lo-S	PP-Irr	PP-OE	PP-P	PP-S	PP-Tra	PP-Tri	Pri-Q	PS-Ov	Total
D. alata				1							294	3	1		1		300
D. bulbifera										2	295	3					300
D. burkilliana											300						300
D. cayenensis											6			294			300
D. diversifolia			29	221		50											300
D. dumetorum			5	150		145											300
D. mangenotiana											300						300
D. praehensillis		8	19	38	1	3	1	1	9	4	190	6			1	19	300
D. rupicola			246	5		49											300
D. sansibarensis						1			6	11	225	55	2				300
D. schimperiana		2				1		1		1	274	19			2		300
D. smilacifolia			38											9	1	252	300
D. spp.		5							24	4	264	3					300
D. sylvatica	130		75	94												1	300
Total	130	15	412	509	1	249	1	2	39	22	2148	89	3	303	5	272	4200

hilum, faint lamellae)] and maximum length overlaps with many others as well, resulting in no unique morphometric identifiers.

Amaryllidaceae

Represented by the USO of *Amaryllis belladonna*. Its complement includes Cylindroid–Oblong Elongates ($n = 80$), Cylindroid–Ellipsoidals ($n = 53$), Globular–Orbiculars ($n = 3$), Parabolic Prism–Parabolics ($n = 2$), Pear Shaped–Irregulars ($n = 8$), and Pear Shaped–Ovates ($n = 154$). No other taxon shares this complement, and the taxon is uniquely identified by Parabolic Prism–Parabolics with eccentric hilum, marked lamellae, and maximum length 52.65–53.89 μm ($n = 2$), as well as by Pear Shaped–Irregulars with eccentric hilum and marked lamellae ($n = 8$). From a purely morphometric point of view, the species is uniquely identified by the presence of Cylindroid–Oblong Elongates with maximum length 60.80–89.70 μm ($n = 42$).

Arecaceae

Represented by a trunk sample from *Hyphaene petersiana*. The complement includes Conoid–Bells ($n = 5$), Conoid–Cuneiforms ($n = 27$), Conoid–Oblong Elongates ($n = 2$), Conoid–Orbiculars ($n = 3$), Cylindroid–Oblong Elongates ($n = 2$), Cylindroid–Ellipsoidals ($n = 119$), Globular–Orbiculars ($n = 81$), Pear Shaped–Irregulars ($n = 1$), and Pear Shaped–Ovates ($n = 60$). No other taxa share these morphotypes in content or proportions. The species has unique identifiers such as Conoid–Bells with centric hilum and faint lamellae ($n = 5$), Conoid–Cuneiforms with centric hilum ($n = 27$), Conoid–Oblong Elongate with centric hilum ($n = 2$), Conoid–Orbicular with centric hilum and faint lamellae ($n = 3$), Pear Shaped–Irregulars with centric hilum ($n = 1$), and Pear Shaped–Ovates with centric hilum and creased surface ($n = 1$). If we consider morphometrics only, *Hyphaene petersiana* is uniquely identified by Conoid–Bell with maximum length 16.41–19.23 μm ($n = 5$), Conoid–Cuneiforms with maximum length <13.00 μm ($n = 27$), Conoid–Oblong

Elongates with maximum length 11.03–15.58 μm ($n = 2$), and Conoid–Orbiculars with maximum length 19.83–23.86 μm ($n = 3$).

Typhaceae

Represented by the USO of *Typha capensis*. The complement consists of Cylindroid–Ellipsoidal ($n = 21$), Globular–Orbicular ($n = 264$), Hemisphere–Orbicular ($n = 2$), and Pear Shaped–Ovate ($n = 13$). There are no unique morphometric identifiers within this species.

Poaceae

Represented by specimens from *Eleusine africana* (grain), *Oxytenanthera abyssinica* (USO), *Sorghum bicolor* (grain), and *Sporobolus panicoides* (grain). Their complements are detailed in **Supplementary Table 4**. *Sorghum bicolor* is uniquely identified by Prismatic–Polygonals with centric hilum and creasing ($n = 3$) or granulation ($n = 2$), and Prismatic–Polygonals with centric hilum and psilate texture with maximum length 12.25–30.34 μm ($n = 295$). *Eleusine africana* is uniquely identified by Prismatic–Polygonals with centric hilum and faint lamellae ($n = 255$). If morphometric criteria are considered, *Eleusine africana* and *Sorghum bicolor* are mutually identified by Prismatic–Polygonals with a maximum length 12.25–18.17 μm ($n = 121$), while individually, *Sorghum bicolor* is uniquely identified by Prismatic–Polygonals with maximum length 18.26–30.34 μm ($n = 108$).

Cyperaceae

Represented by the USOs of three *Cyperus* species: *C. esculentus*, *C. rotundus*, and *C. tomaiophyllus*. They demonstrate dramatic differences in complements (**Supplementary Table 5**).

C. rotundus and *C. tomaiophyllus* are uniquely identified by Lobate–Irregular granules with centric hilum and psilate texture [$n = 207/11$ (C.e./C.t.)]. *C. esculentus* is uniquely identified by Lobate–Irregular granules with centric hilum and creased surface

($n = 16$), Pear Shaped–Ovates with centric hilum and creased surface with maximum length 6.45–22.79 μm ($n = 214$), and Pear Shaped–Ovates with centric hilum location and pitted surfaces ($n = 1$). *C. rotundus* is uniquely identified by Pear Shaped–Ovates with centric hilum and psilate texture ($n = 14$). *C. tomaiophyllus* is uniquely identified by Globular–Irregulars with centric hilum and psilate texture ($n = 6$) and by Lobate–Shells with centric hilum and psilate texture with maximum length 6.45–22.79 μm ($n = 6$). If we look at morphometrics only hoping to discriminate among unknown granules, only *C. tomaiophyllus* could be uniquely identified [Lobate–Shells with max. size range 14.45–21.65 μm ($n = 6$)].

Strelitziaceae

Represented by fruit mesocarp from *Strelitzia nicolai* with a granule complement consisting of Conoid–Fusiforms ($n = 227$), Conoid–Irregulars ($n = 26$), Conoid–Orbiculars ($n = 14$), and Conoid–Ovates ($n = 33$). Uniquely identified by Conoid–Fusiforms ($n = 227$), Conoid–Irregulars with centric hilum and psilate texture ($n = 26$), Conoid–Orbiculars with centric hilum and psilate texture ($n = 14$), and Conoid–Ovates with centric hilum and psilate texture ($n = 33$).

Morphometrically, this species is uniquely identified by Conoid–Fusiforms with a maximum length 7.61–30.32 ($n = 227$), Conoid–Irregulars with maximum length 11.64–18.85 μm ($n = 18$), Conoid–Orbiculars with maximum length 8.12–13.62 μm ($n = 14$), and Conoid–Ovates with a maximum length 10.12–20.90 μm ($n = 33$).

Musaceae

Represented by the USO of *Ensete ventricosum*. The starch granule complement comprises Conoid–Bells ($n = 2$), Conoid–Irregulars ($n = 77$), Conoid–Oblong Elongates ($n = 42$), Conoid–Ovates ($n = 6$), Cylindroid–Oblong Elongates ($n = 3$), Cylindroid–Ellipsoidals ($n = 10$), Cylindroid–Irregulars ($n = 34$), Irregular–Irregulars ($n = 52$), Lobate–Irregulars ($n = 57$), Pear Shaped–Irregulars ($n = 4$), Pear Shaped–Shells ($n = 29$), and Pear Shaped–Ovates ($n = 4$). It is uniquely identified by Pear Shaped–Irregulars with hyper-eccentric hilum ($n = 4$), Pear Shaped–Shells with hyper-eccentric hilum and marked lamellae ($n = 4$), and Pear Shaped–Ovates with hyper-eccentric hilum ($n = 29$). If we look at morphometrics only, unique identifiers include Conoid–Bells with maximum length 33.35–38.01 μm ($n = 2$), Conoid–Irregulars with maximum length 30.08–82.36 μm ($n = 77$), Conoid–Oblong Elongates with maximum length 31.47–90.42 μm ($n = 42$), Conoid–Ovates with maximum length 35.82–48.43 μm ($n = 5$), Cylindroid–Irregulars with maximum length 51.78–67.74 μm ($n = 7$), Irregular–Irregulars with maximum length 38.93–93.33 μm ($n = 45$), Lobate–Irregulars with maximum length 50.43–82.32 μm ($n = 15$), and Pear Shaped–Shells with maximum length 31.49–53.46 μm ($n = 21$).

Menispermaceae

Represented by the USO of *Jateorhiza palmata*. The complement includes Conoid–Ovates ($n = 1$), Cylindroid–Oblong Elongates

($n = 36$), Cylindroid–Ellipsoidals ($n = 186$), Cylindroid–Irregulars ($n = 1$), Globular–Orbiculars ($n = 15$), Pear Shaped–Irregulars ($n = 4$), and Pear Shaped–Ovates ($n = 57$). It is uniquely identified by Conoid–Ovates with eccentric hilum and psilate texture ($n = 1$), Pear Shaped–Irregulars with eccentric hilum and psilate textures ($n = 4$), and Pear Shaped–Ovates with eccentric hilum and psilate texture ($n = 57$), although if we restrict parameters to morphometric identifiers, there would be none within the species.

Proteaceae

Represented by the USO of *Protea welwitschii*. The two morphotypes present, Globular–Orbicular ($n = 274$) and Hemisphere–Orbicular ($n = 26$), are ubiquitous across families. There are no unique morphometric identifiers within this species.

Euphorbiaceae

Represented by seeds from two *Ricinodendron* species: *heudelotii* and *rautanenii*. The complement includes Globular–Orbiculars [$n = 271/169$ (R.h./R.r.)], Globular–Trapeziforms ($n = 10/13$), and Hemisphere–Orbiculars ($n = 19/118$). While the morphotype complement is consistent within the Euphorbiaceae, as represented by these two species, the proportions are not. *R. heudelotii* is uniquely identified by Hemispere–Orbiculars with granulate texture ($n = 19$) while *R. rautanenii* is uniquely identified by Hemisphere–Orbiculars with psilate textures ($n = 118$). Unique identifiers within this family would be none if we restricted criteria to morphometry alone.

Fabaceae

Represented by the legume or USO of sixteen species (**Table 5**). Granules can be identified using 15 unique morphotypes combined with hilum location and surface texture. From a purely morphometric standpoint, *Eriosema preptum* (USO) is uniquely identified by Conoid–Triangulars with maximum length 16.13–21.65 μm ($n = 7$) as well as by Prismatic–Irregulars with maximum length 18.11–25.20 μm ($n = 2$), and Prismatic–Quadratics with maximum length 20.77–38.64 μm ($n = 80$). *Cassia petersiana* (legume) is uniquely identified by Cylindroid–Ellipsoidals maximum length 3.48–4.90 μm ($n = 31$). *Cajanus cajan* (legume) is uniquely identified by Cylindroid–Ovates with maximum length 41.90–41.99 μm ($n = 2$). *Lablab* spp. (legume) and *Vigna lobatifolia* (legume) are jointly identified by Reniform–Reniforms with maximum length 20.77–38.84 μm ($n = 14$).

Cannabaceae

Represented by the fruit of *Celtis africana*. The granule assemblage is dominated by Globular–Orbicular granules ($n = 299$) with granulate texture. The single remaining granule is a Cylindroid–Ellipsoidal with granulate texture. These morphotypes are ubiquitous across families and length ranges, resulting in no unique morphometric identifiers.

Cucurbitaceae

Represented by the seeds of *Acanthosicyos naudinianus* and *Momordica cardiospermoides*. The morphotype complements and proportions of the two species differ considerably: Cylindroid–Oblong Elongate [$n = 0/15$ (A.n./M.c.)],

TABLE 5 | Fabaceae morphotypes.

Species										Morphotype									
	Co-Irr	Co-Tri	Cy-OE	Cy-El	Cy-Irr	Cy-Ov	G-Irr	G-Orb	G-Tra	H-Orb	Irr-Irr	Lo-Irr	PP-P	PS-Irr	PS-Ov	Pri-Irr	Pri-Q	R-R	Total
Acacia nilotica								300											300
Acacia tortilis				193				107											300
Afzelia quanzensis				300															300
Cassia petersiana				300															300
Cajanus cajan			1	243	23	5	4	24											300
Cordyla africana				16				58	44	182	1								300
Eriosema preptum	1	7		38	2		2	98	7	3			21		9	2	93		300
Faidherbia albida			3	248				47		18									300
Lablab purpureus				220				80		2									300
Lablab spp.			1	202	1			39							4			2	300
Lotus discolor								163		2		49							300
Mucuna pruriens			2	123				49		137				1	125				300
Mucuna spp.				192				108											300
Spherostylis stenocarpa								11	149	140									300
Vigna lobatifolia				204	1			49				21	1		12			12	300
Vigna subterranea			5	140				146		3					4				300
Total	1	7	12	2419	27	5	6	1279	200	487	1	70	22	1	154	2	93	14	4800

Cylindroid–Ellipsoidal ($n = 3/171$), Globular–Orbicular ($n = 289/96$), Prismatic–Quadratic ($n = 8/0$), and Pear Shaped–Ovate ($n = 0/18$). *Acanthosicyos naudinianus* is uniquely identified by Prismatic–Quadratics with psilate textures ($n = 8$). *Momordica cardiospermoides* is uniquely identified by Cylindroid–Ellipsoidals with eccentric hilum and faint lamellae ($n = 4$). From a morphometric point of view, *Acanthosicyos naudinianus* is uniquely identified by Prismatic–Quadratics with maximum length 4.72–8.49 μm ($n = 8$).

Combretaceae

Represented by the legume of *Combretum* spp. The granule assemblage is dominated by Cylindroid–Ellipsoidals ($n = 207$) with declining representation of Globular–Orbiculars ($n = 62$) and Pear Shaped–Ovates ($n = 31$). Morphologically it is uniquely identified by Pear Shaped–Ovates with granulate texture ($n = 31$), but there are no unique morphometric identifiers.

Bombacaceae

Represented by the fruit mesocarp of *Adansonia digitata*. All granules belong to a single morphotype: Globular–Orbicular with centric hilum, psilate texture, sometimes faceted. This is the single most widely distributed morphotype in the collection and several species from different families share the full complement of Globular–Orbicular granules. Additionally, the maximum length overlaps completely with several families noted here, resulting in no unique morphometric identifiers.

Polygonaceae

Represented by the USO of *Polygonum senegalense*. A single 3D-2D morphotype is present in the Polygonaceae sample: Globular–Orbicular with centric hilum and psilate texture. These identifiers are nearly ubiquitous and maximum length overlap amongst the various families, preventing this species from being uniquely identified morphometrically.

Sapotaceae

Represented by the fruit of *Englerophytum magalismontanum*. The morphotypes include Globular–Orbiculars ($n = 34$), Globular–Trapeziforms ($n = 128$), and Hemisphere–Orbiculars ($n = 138$). All granules have centric hilum and psilate texture, with 266 faceted cases that include Orbiculars and Globular–Trapeziforms. These morphotypes and length ranges are ubiquitous across families and cannot be used in unique morphometric identification.

Convolvulaceae

Represented by the USO of *Ipomoea longituba*. All granules belong to a single morphotype: Globular–Orbiculars with psilate texture and equatorial groove. This morphotype is a unique identifier for this species; however, there are no unique morphometric identifiers within this species.

Apocynaceae

Represented by the USOs of *Ceropegia rendalii* and *Fockea angustifolia*. Their granule morphotype complement differ

TABLE 6 | Presence/absence of unique identifiers per family.

Family	Number of granules observed	Unique morphological identifiers	Unique morphometric identifiers
Zamiaceae	1500	Yes	Yes
Nymphaeaceae	300	Yes	No
Araceae	600	Yes	Yes
Taccaceae	300	Yes	No
Dioscoreaceae	4200	Yes	Yes
Pandanaceae	300	No	No
Orchidaceae	300	No	No
Iridaceae	900	Yes	No
Hypoxidaceae	900	No	No
Asphodelaceae	300	No	No
Amaryllidaceae	300	Yes	Yes
Arecaceae	300	Yes	Yes
Typhaceae	300	No	No
Poaceae	1200	Yes	Yes
Cyperaceae	900	Yes	Yes
Strelitziaceae	300	Yes	Yes
Musaceae	300	Yes	Yes
Menispermaceae	300	Yes	No
Protaceae	300	No	No
Euphorbiaceae	600	Yes	No
Fabaceae	4800	Yes	Yes
Cannabaceae	300	No	No
Cucurbitaceae	600	Yes	Yes
Combretaceae	300	Yes	No
Bombacaceae	300	No	No
Polygonaceae	300	No	No
Sapotaceae	300	No	No
Convolvulaceae	300	Yes	No
Apocynaceae	600	Yes	No
Pedaliaceae	300	Yes	No
Lamiaceae	600	Yes	No

significantly in content and proportion to include Pear Shaped–Ovates [$n = 0/89$ (C.r./F.a.)], Globular–Orbiculars ($n = 262/192$), Globular–Trapeziforms ($n = 32/7$), Hemisphere–Orbiculars ($n = 6/10$), and Pear Shaped–Shells ($n = 0/2$). *Fockea angustifolia* is uniquely identified by Hemisphere–Orbiculars with centric hilum and marked lamellae ($n = 8$) and by Pear Shaped–Shells with centric hilum ($n = 2$). There are no unique morphometric identifiers within this family.

Pedaliaceae

Represented by the seeds of *Sesamum angolense* yielding a unique identifier: Lobate–Irregulars with psilate texture. This species is not morphometrically unique.

Lamiaceae

Represented by the USOs of *Plectranthus esculentus* and *Solenostemon rotundifolius*.

The morphotype complement is noted in **Supplementary Table 6**. The complement proportions in the two species differ significantly (Chi–square = 68.328, df = 2, $p = 0.000$). *Plectranthus esculentus* is uniquely identified by Hemisphere–Orbiculars with centric hilum, psilate texture, facets and bifid pleating ($n = 74$). There are no unique morphometric identifiers within this family, however, if we were to include morphology and size only.

DISCUSSION

A recent overview of 176 publications in archeological starch from the last two decades shows that one of the most pressing needs in the field is to establish reference collections that evaluate the reliability of taxonomic identification and highlight diagnostic morphotypes (Barton and Torrence, 2015). The vast majority of ancient starch attributions to plant of origin are still conducted using starch granule morphotypes and morphometrics directly recognized and classified under the microscope, such as we did here. Only four publications

to date have explored computerized manipulation, and they present divergent results and degrees of success (Torrence et al., 2004; Wilson et al., 2010; Coster and Field, 2015; Arráiz et al., 2016). A persistent problem with both visual and automated classification of reference collections is small size, often utilizing <30 taxa. Paradoxically, as the pool of target species grows, so do difficulties and mistaken identifications, making even automated classification unsatisfactory (Torrence et al., 2004; Wilson et al., 2010; Arráiz et al., 2016).

In our analysis of the data, 10 of the 31 families cannot be identified (**Table 6**) by combining morphometrics with volumetric and/or textural features. Using morphometrics exclusively results in the removal of an additional 10 families. Morphotypes are shared despite taxonomic distance for 67.7% of the taxa. This cohort of confuser granules encompasses 37 morphotypes (**Supplementary Table 7**). Conversely, this reference collection boasts 98 unique identifiers, most of which concentrate in five families (Dioscoreaceae, Fabaceae, Musaceae, Cyperaceae, and Arecaceae) (**Figure 2A**). The relationship between total number of granules per family and the appearance of unique identifiers is not significant
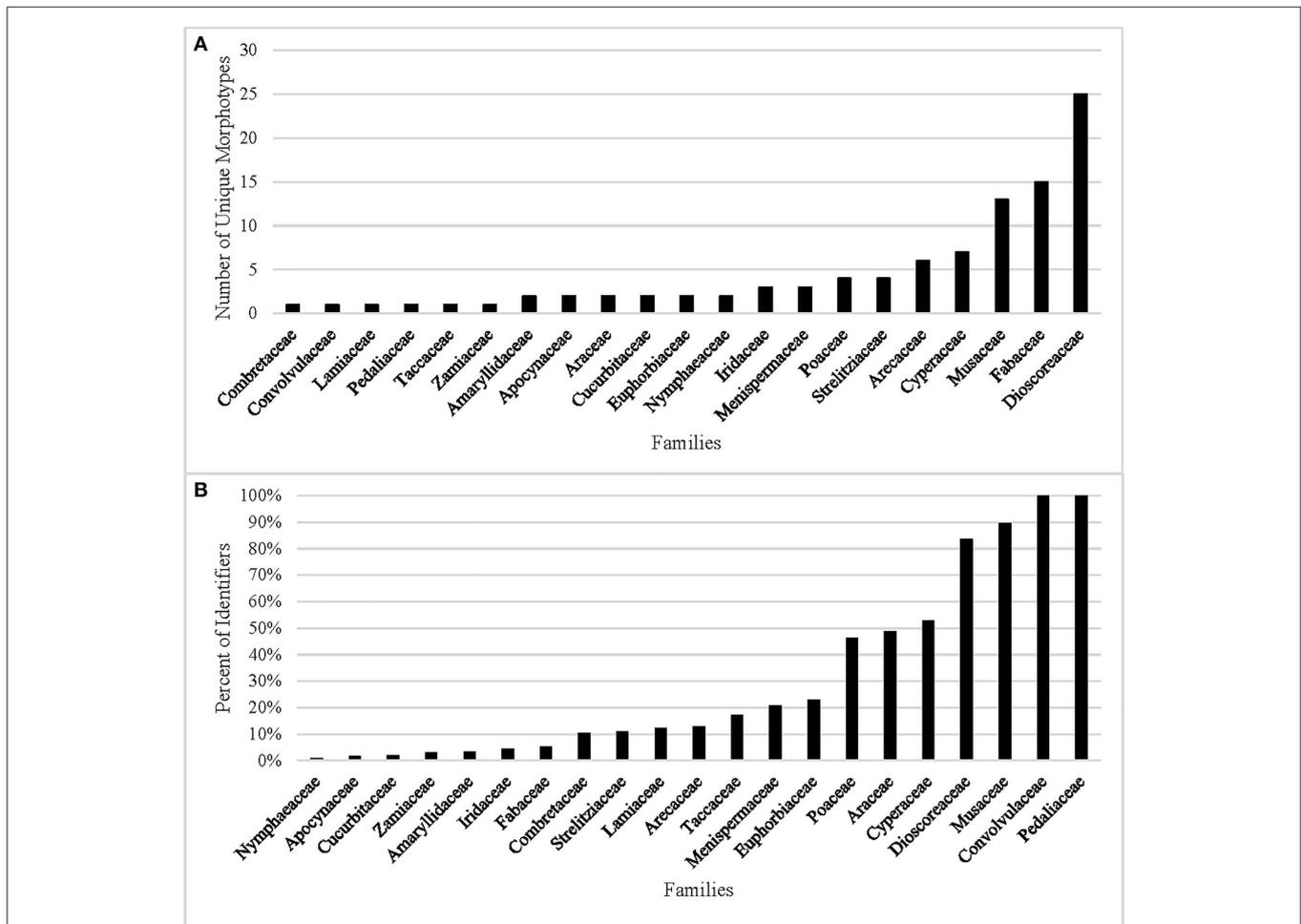
FIGURE 2 | (A) Number of morphometric identifiers per family, **(B)** Percent of granules displaying unique morphology from the total count per family.

($r2 = 0.028$, $p = 0.371$). Even if we were to accept outlier bias and include families represented by a large number of species, the linear relationship between them and the identification of unique identifiers would be only weakly positive ($r2 = 0.629$, $p = 0.000$). Two of the families with the fewest species (Convolvulaceae, Pedaliaceae) have the highest percentage of identifier granules (**Figure 2B**). The Fabaceae, on the other hand, has the highest granule count ($n = 4,800$) but only 5.3% are unique to the family. Comparing species with similar granule counts, the Poaceae ($n = 1,200$) had 46.3% unique granules, while the Zamiaceae ($n = 1,500$) had three.

Reniforms are extremely rare, only 0.3% of all Fabaceae and 0.06% of all granules.

Reniforms in the Fabaceae consist of two examples from *Lablab* spp. and 12 from *Vigna lobatifolia*. Planoconvex Lenticulars were recovered only amongst the Zamiaceae ($n = 46$), specifically from the trunk of *Encephalartos friderici-guilielmi*. The other four Zamiaceae did not yield lenticular granules. Lenticulars are 3% of all Zamiaceae and 0.2% of the full collection. Irregulars constitute another rare morphotype in the collection ($n = 53$, 0.2%). Fifty-two of these (98.1%) were recovered from the Musaceae (17.3% of all Musaceae granules) and one (1.9%) from a Fabaceae (a *Cajanus cajan* legume). Parabolic Prisms occur frequently ($n = 2,628$, 11.3%), and are overwhelmingly associated with the Dioscoreaceae (99.1%), although 0.08% were recovered from *Amaryllis belladonna* ($n = 2$) and 0.8% from the Fabaceae ($n = 22$, 21 from an *Eriosema preptum* USO and one from a *Vigna lobatifolia* legume). Parabolic Prisms represent 62% of all Dioscoreaceae granules, 0.6% of all Amaryllidaceae granules, and 0.4% of all Fabaceae granules. Parabolic Prisms with distal bifid shape are unique to eight of the 15 *Dioscorea*. Parabolic Prisms vary in maximum length 8.49–83.81 μm. In the present sample, all Parabolic Prisms 26.40–52.00 μm in length are from *Dioscorea* species. Amaryllidaceae specimens are >52 μm in length and all 22 Fabaceae granules are <26.40 μm.

We examined size distribution and found that maximum lengths overlap ubiquitously. When creating a subsample with the upper 10 percent sizes only (12 families, $n = 2,310$), no patterning or exclusive identificatory characteristics could be observed. Subsampling for the largest granules was detrimental to identification, as it eliminated 19 families and 40 species from consideration and blocked previously recognized unique identifiers (10 of 22). The procedure also affected sample size and morphotype proportions from each family differentially.

Distributions of maximum lengths amongst the 77 species were examined for normality using the SPSS Statistics 25 Explore and Descriptives modules. Where distributions were found not to be normal, the data were transformed using log10 and/or square root calculations and re-examined for normality of distribution (**Supplementary Table 8**; **Supplementary Figure 1**). The results indicate that 36% of the taxa follow a non-normal distribution (**Supplementary Table 9**), which demonstrates that

using mean or median maximum length as an identificatory feature is an invalid way to describe central tendency in a granule population; at least not without previous confirmation that a given taxon actually yields a normal size distribution (see **Data Sheet 1, Supplementary Table 10**).

CONCLUSIONS

The size of a reference collection is key to ascertaining the uniqueness of a morphotype, which must have characteristic shapes and sizes to allow for solid identification. In a collection consisting of 23,100 granules, firm taxonomic identification occurred very rarely (7%), and when it did, it was at the species level, with no instances in which a morphotype identifies all species within a family or genus. We found no evidence that starch granule morphology, in isolation, is able to reveal plant part, species, family or order unambiguously: There are no diagnostic 3-D or 2-D shapes, and granule length alone is not a useful variable in identification, but rather must be assessed in combination with morphology.

Positively, some species and families produce more morphometrically unique granules than others do. Therefore, the identification of starch granules is possible, but not guaranteed, depending on the type of plant that was exploited in the past and whether the specific taxa and ecoregion contain unique identifiers. If early humans utilized plants that bear starch granules without demonstrable diagnostic morphometrics, then identification will be improbable to impossible. The problem isolated here is not with the type of inspection conducted, visual, or automated, rather it is with the loss of unique identifiers when using large datasets, and reflect an overwhelming morphometric overlap between many species, phylogenetically related or not (see also Mercader et al., 2018).

AUTHOR CONTRIBUTIONS

JM and SC carried out experimental conceptualization, conducted background research, and sample analysis. MA, RB, MB, JF, PL, MI, FL, JI, RP, MS, LT and DW performed research and analysis. JM, DW, and SC wrote the manuscript text with input from all authors. All authors contributed to research design and planning.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2018.00146/full#supplementary-material

REFERENCES

Arráiz, H., Barbarin, N., Pasturel, M., Beaufort, L., Domínguez-Rodrigo, M., and Barboni, D. (2016). Starch granules identification and automatic classification based on an extended set of morphometric and optical measurements. *J. Archaeol. Sci. Rep.* 7, 169–172. doi: 10.1016/j.jasrep.2016.03.039

Barton, H., and Torrence, R. (2015). Cooking up recipes for ancient starch: assessing current methodologies and looking to the future. *J. Archaeol. Sci.* 56, 194–201. doi: 10.1016/j.jas.2015.02.031

Copeland, L., and Hardy, K. (2018). Archaeological starch. *Agronomy* 8, 4. doi: 10.3390/agronomy8010004

Cortella, R., and Pochettino, M. (1995). Comparative morphology of starch of three andean tubers. *Starch Starke* 47, 455–461.

Coster, A. C. F., and Field, J. H. (2015). What starch grain is that? – A geometric morphometric approach to determining plant species origin. *J. Archaeol. Sci.* 58, 9–25. doi: 10.1016/j.jas.2015.03.014

Czaja, A. (1978). Structure of starch grains and the classification of vascular plant families. *Taxon* 27, 463–470.

Dharani, N., and Yenesew, A. (2010). *Medicinal Plants of East Africa: An Illustrated Guide*. Cape Town: Struik Nature.

Hart, T. (2014). Analysis of starch grains produced in select taxa encountered in Southwest Asia. *Ethnobot. Lett.* 5, 135–145. doi: 10.14237/ebl.5.2014.251

Jane, J. L., Kasemsuwan, T., Leas, S., Zobel, H., and Robyt, J. F. (1994). Anthology of starch granule morphology by scanning electron microscopy. *Starch Starke* 46, 121–129. doi: 10.1002/star.19940460402

Klopper, R. R., Chatelain, C., Habashi, C., Gautier, L., and Spichiger, R. E. (2006). *Checklist of the Flowering Plants of Sub-Saharan Africa: An Index of Accepted Names and Synonyms*. Pretoria: Sabonet.

Laden, G., and Wrangham, R. (2005). The rise of the hominids as an adaptive shift in fallback foods: plant underground storage organs (USOs) and australopith origins. *J. Hum. Evol.* 49, 482–498. doi: 10.1016/j.jhevol.2005.05.007

Langejans, G. H. J. (2006). Starch grain analysis on Late Iron Age grindstones from South Africa. *S. Afr. Humanit.* 18, 71–91.

Lentfer, C. J. (2009). Building a comparative starch reference collection for indonesia and its application to palaeoenvironmental and archaeological research. *Terra Austalis* 30, 80–101. doi: 10.22459/TA30.07.2009.07

Lindeboom, N., Chang, P. R., and Tyler, R. T. (2004). Analytical, biochemical and physicochemical aspects of starch granule size, with emphasis on small granule starches: a review. *Starch* 56, 89–99. doi: 10.1002/star.200300218

Louderback, L. A., Herzog, N. M., and Pavlik, B. M. (2016). A new approach for identifying starch granules of wild food plants from arid western North America. *Starch Starke* 68, 1–7. doi: 10.1002/star.201600167

Matsushima, R., Yamashita, J., Kariyama, S., Enomoto, T., and Sakamoto, W. (2013). A phylogenetic re- evaluation of morphological variations of starch grains among poaceae species. *J. Appl. Glycosci.* 60, 37–44. doi: 10.5458/jag.jag.JAG-2012_006

Maundu, P., and Tengnas, B. (2005). *Useful Trees and Shrubs for Kenya*. Nairobi: World Agroforestry Centre.

Mercader, J., Akeju, T., Brown, M., Bundala, M., Collins, M. J., Copeland, L., et al. (2018). Exaggerated expectations in ancient starch research and the need for new taphonomic and authenticity criteria. *Facets* 3, 777–798. doi: 10.1139/facets-2017-0126

Meyer, A. (1895). *Untersuchangen Über die Stäkekorner*. Jena: Fisher.

Musaubach, M., Plos, A., and Bobot, P. (2013). Differentiation of archaeological maize (*Zea mays* L.) from native wild grasses based on starch grain morphology. Cases from the Central Pampas of Argentina. *J. Archaeol. Sci.* 40, 1186–1193. doi: 10.1016/j.jas.2012.09.026

Nageli, K. W., Cramer, K. E., and Wartmann, B. (1858). *Die Stärkekörner: Morphologische, Physiologische, Chemisch-Physicalische und Systematisch-Botanische*. Zurich: F. Schulthess.

Parr, J. (2002). The identification of xanthorrhoea resins by starch morphology: prospects for archaeological and taxonomic applications. *Econ. Bot.* 56, 260–270. doi: 10.1663/0013-0001(2002)056[0260:TIOXRB]2.0.CO;2

Reichert, E. T. (1913). *The Differentiation and Specificity of Starches in Relation to Genera, Species, Etc.* Washington, DC: Carnegie Institution of Washington.

Shapter, F. M., Henry, R. J., and Lee, L. S. (2008). Endosperm and starch granule morphology in wild cereal relatives. *Plant Genet. Resour.* 6, 85–97. doi: 10.1017/S1479262108986512

Thompson, S. K. (1992). *Sampling*. New York, NY: John Wiley & Sons.

Torrence, R., Wright, R., and Conway, R. (2004). Identification of starch granules using image analysis and multivariate techniques. *J. Archaeol. Sci.* 31, 519–532. doi: 10.1016/j.jas.2003.09.014

Van Wyk, B., and Gericke, N. (2000). *People's Plants: A Guide to Useful Plants of Southern Africa*. Pretoria: Briza Publications.

White, F. (1983). *The Vegetation of Africa: A Descriptive Memoir to Accompany the UNESCO/ AETFAT/UNSO Vegetation Map of Africa*. Paris: UNESCO.

Williamson, J. (2005). *Useful Plants of Malawi*. Limbe: Momfort Press.

Wilson, J., Hardy, K., Allen, R., Copeland, L., Wrangham, R., and Collins, M. (2010). Automated classification of starch granules using supervised pattern recognition of morphological properties. *J. Archaeol. Sci.* 37, 594–604. doi: 10.1016/j.jas.2009.10.024

Yang, X., Zhang, J., Perry, K., Ma, Z., Wan, Z., Li, M., et al. (2012). From the modern to the archaeological: starch grains from millets and their wild relatives in China. *J. Archaeol. Sci.* 39, 247–254. doi: 10.1016/j.jas.2011.09.001

APPENDIX

TABLE A1 | Taxonomic classification of the granules counted and plant parts studied.

Order	Family	Species	Plant part	Brightfield	Darkfield
Cycadales	Zamiaceae	*Encephalartos friderici-guilielmi*	Trunk	20µm	
Cycadales	Zamiaceae	*E. laevifolius*	Trunk	20µm	
Cycadales	Zamiaceae	*E. natalensis*	Trunk	20µm	
Cycadales	Zamiaceae	*E. princeps*	Trunk	20µm	
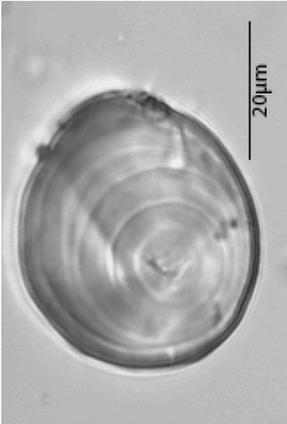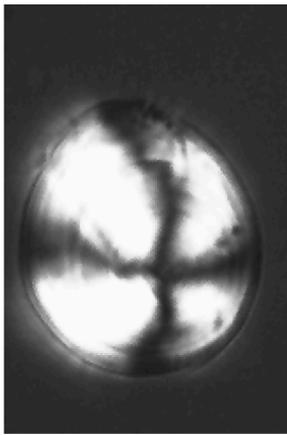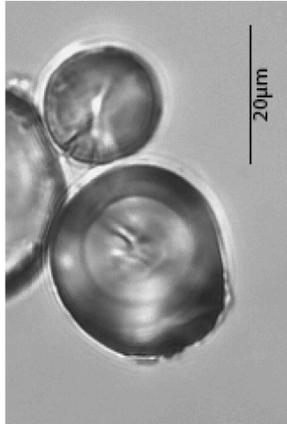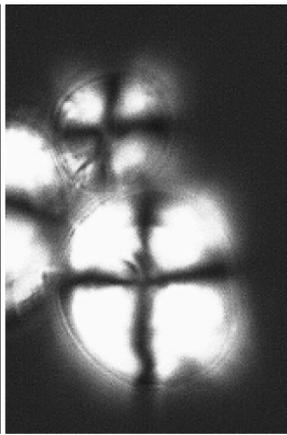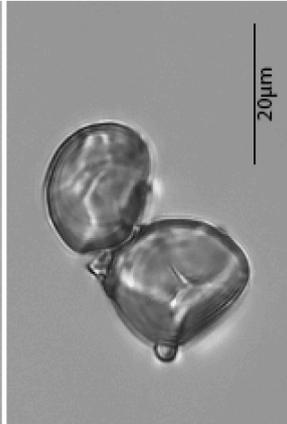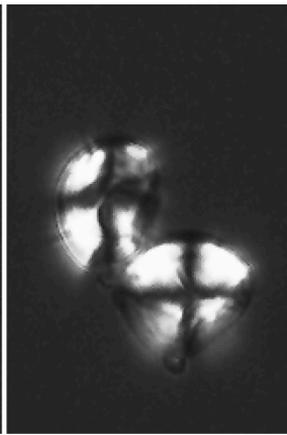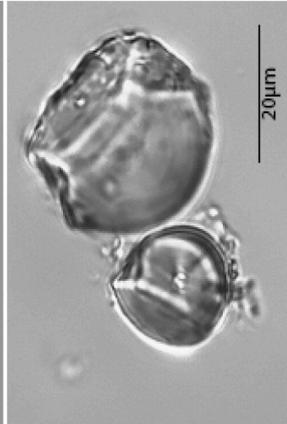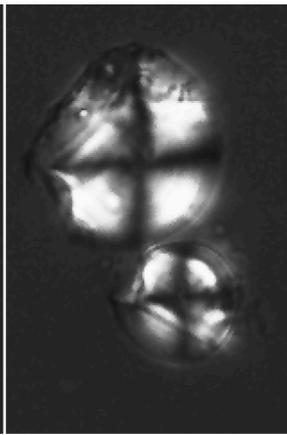
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Cycadales	Zamiaceae	*E. senticosus*	Trunk		
Nymphaeales	Nymphaeaceae	*Nymphaea caerulea*	USO		
Alismatales	Araceae	*Zantedeschia aethiopica*	USO		
Alismatales	Araceae	*Zantedeschia rehmannii*	USO		
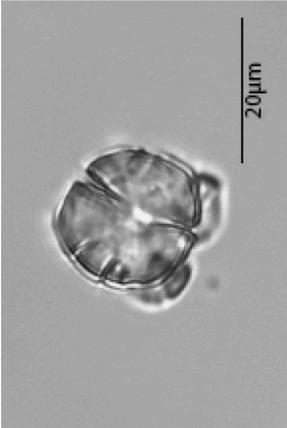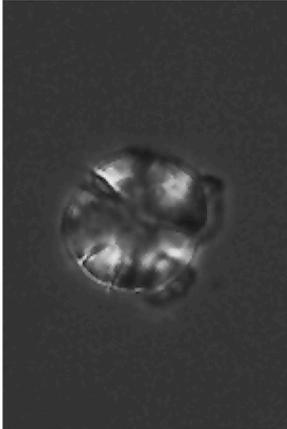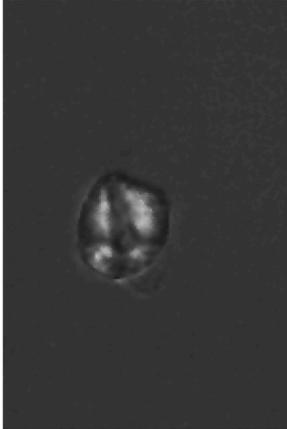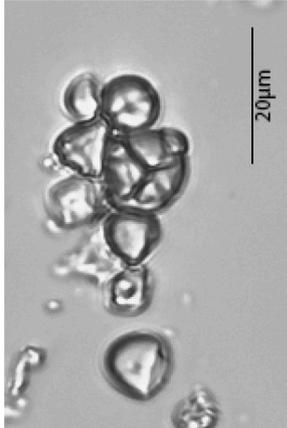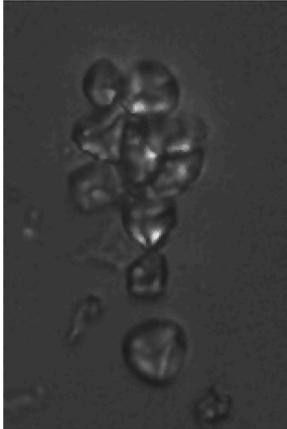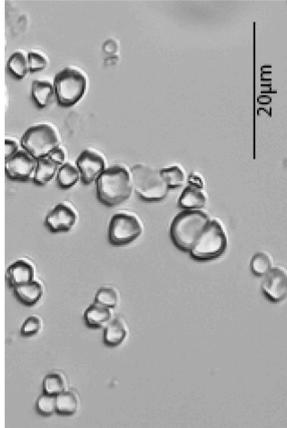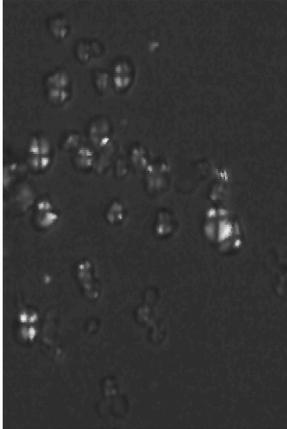
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Dioscoreales	Dioscoreaceae	*Dioscorea alata*	USO		
Dioscoreales	Dioscoreaceae	*D. bulbifera*	USO		
Dioscoreales	Dioscoreaceae	*D. burkilliana*	USO		
Dioscoreales	Dioscoreaceae	*D. cayenensis*	USO		
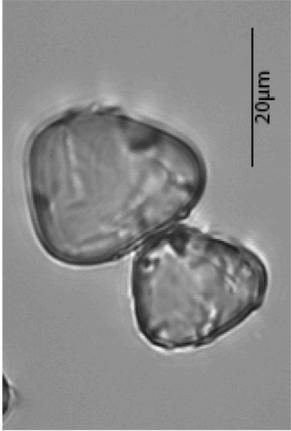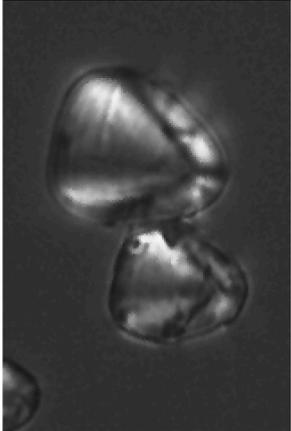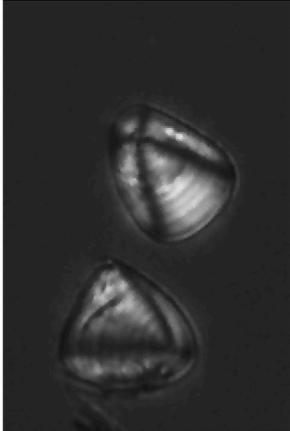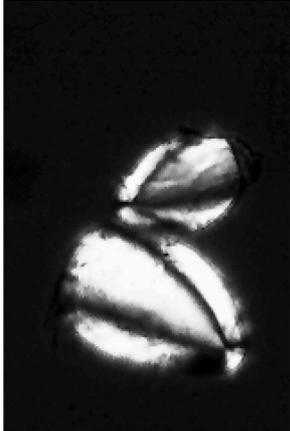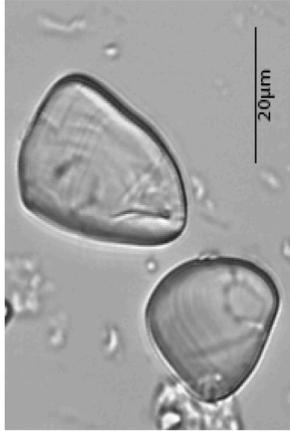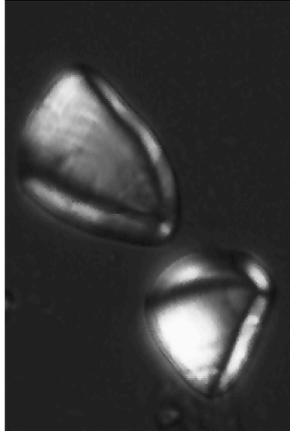
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Dioscoreales	Dioscoreaceae	*D. diversifolia*	USO		
Dioscoreales	Dioscoreaceae	*D. dumetorum*	USO		
Dioscoreales	Dioscoreaceae	*D. mangenotiana*	USO		
Dioscoreales	Dioscoreaceae	*D. praehensilis*	USO		
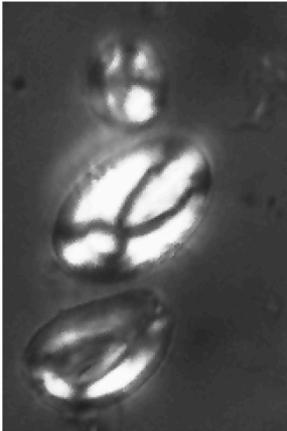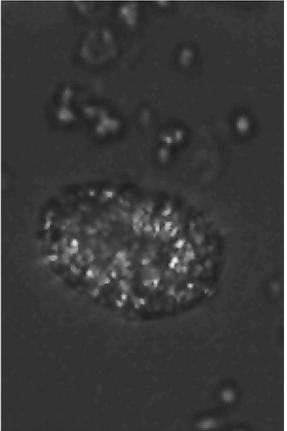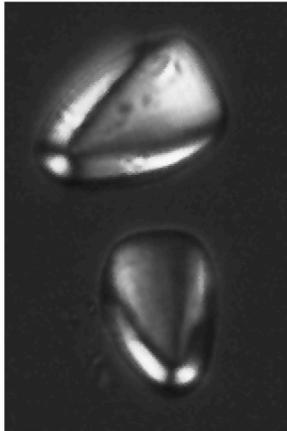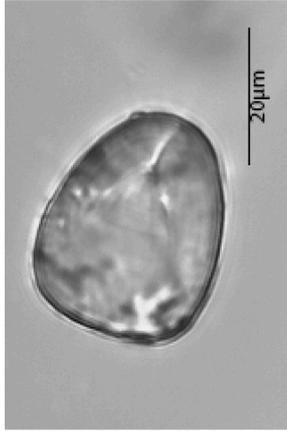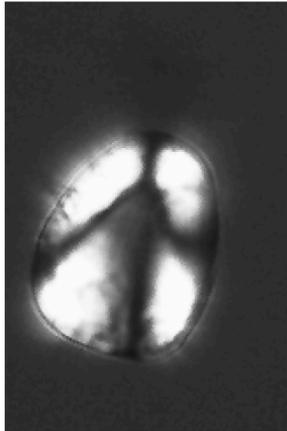
(Continued)

TABLE A1 | Continued

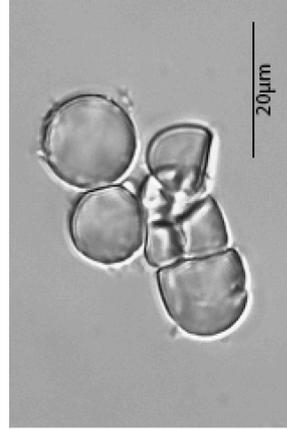| Order	Family	Species	Plant part	Brightfield	Darkfield
Dioscoreales	Dioscoreaceae	*D. rupicola*	USO		
Dioscoreales	Dioscoreaceae	*D. sansibarensis*	USO		
Dioscoreales	Dioscoreaceae	*D. schimperiana*	USO		
Dioscoreales	Dioscoreaceae	*D. smilacifolia*	USO		
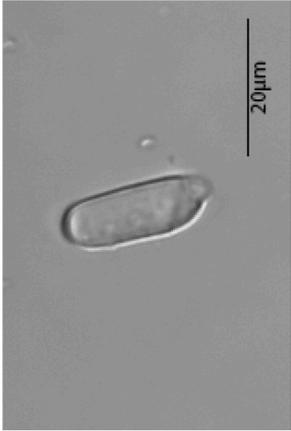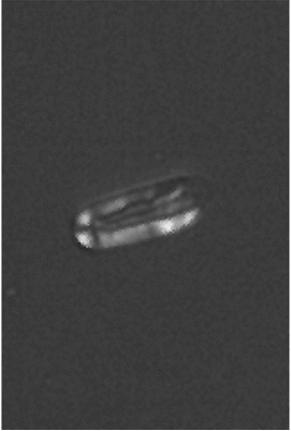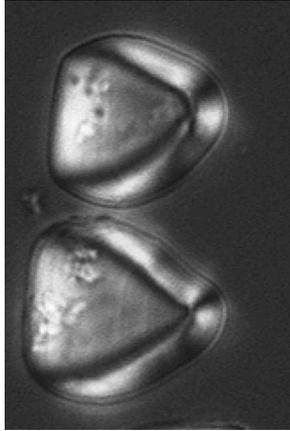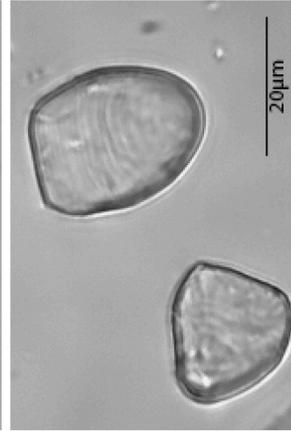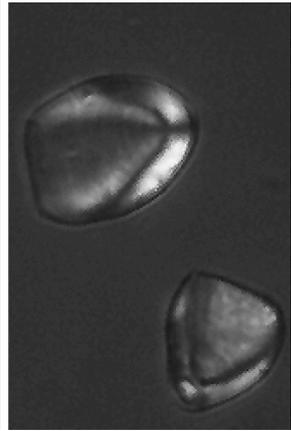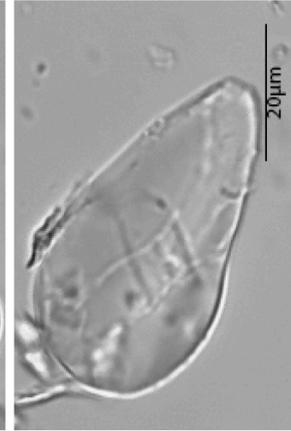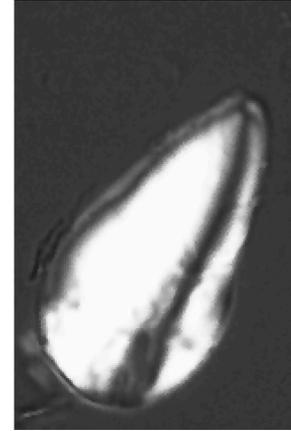
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Dioscoreales	Dioscoreaceae	Dioscorea spp.	USO		
Dioscoreales	Dioscoreaceae	D. sylvatica	USO		
Dioscoreales	Taccaceae	Tacca leontopetaloides	USO		
Pandanales	Pandanaceae	Pandanus livingstonianus	Trunk		
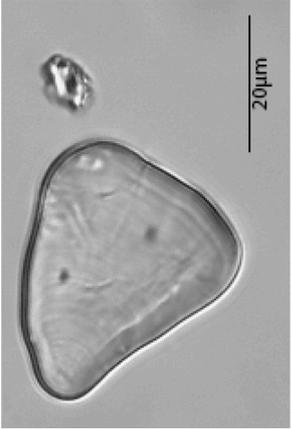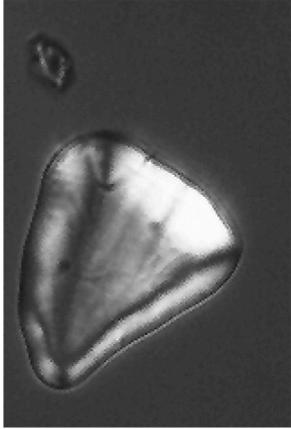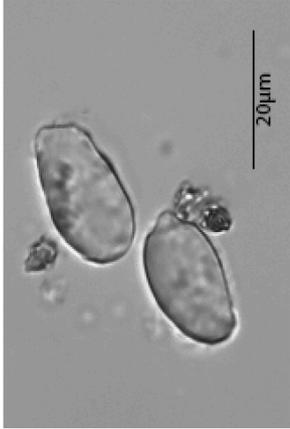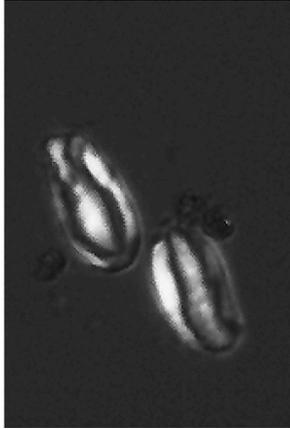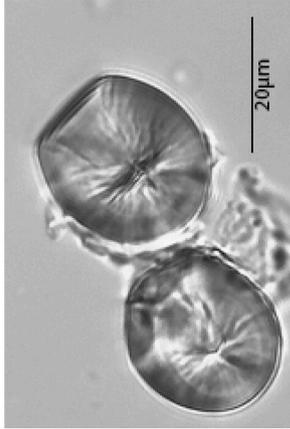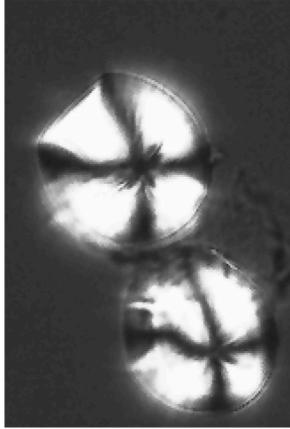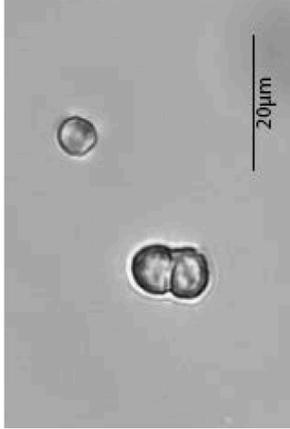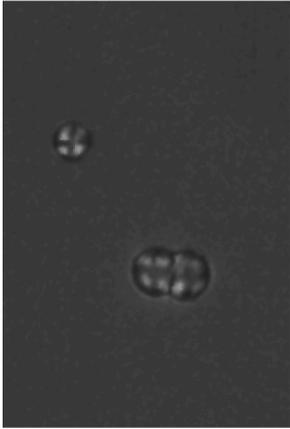
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Asparagales	Amaryllidaceae	*Amaryllis belladonna*	USO		
Asparagales	Asphodelaceae	*Bulbine sedifolia*	USO		
Asparagales	Hypoxidaceae	*Hypoxis nyasica*	USO		
Asparagales	Hypoxidaceae	*H. rigidula*	USO		
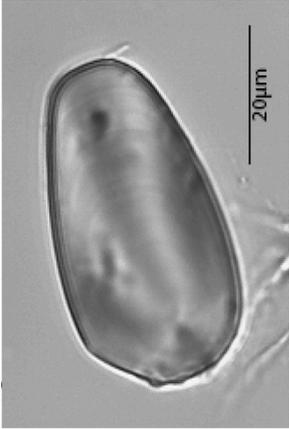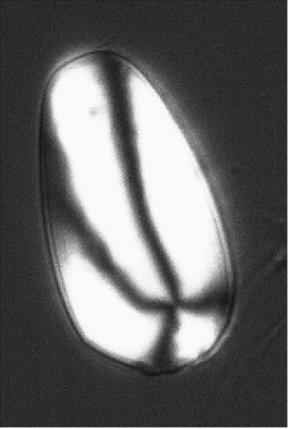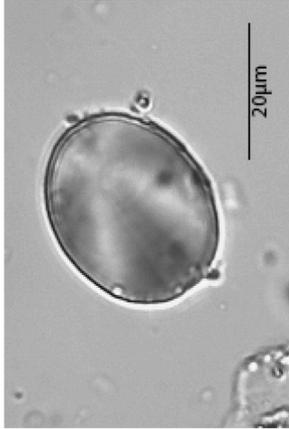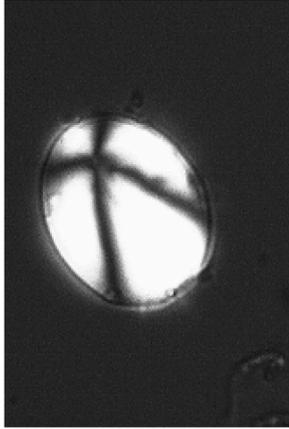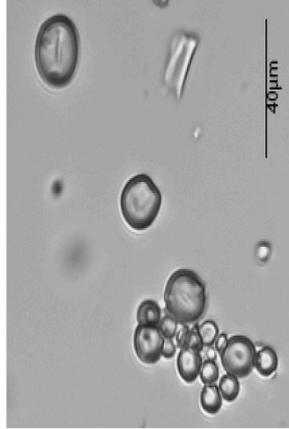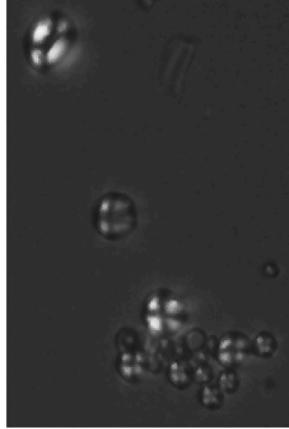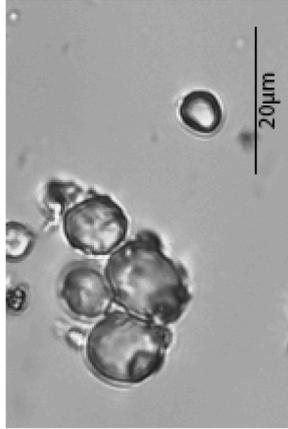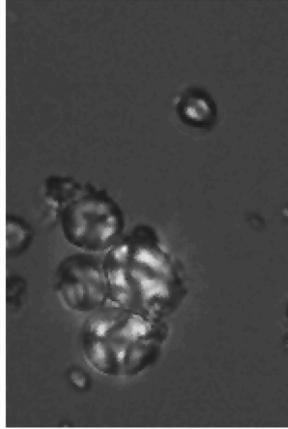
(Continued)

(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Asparagales	Hypoxidaceae	*H. zeiheri*	USO		
Asparagales	Iridaceae	*Gladiolus gueinzii*	USO		
Asparagales	Iridaceae	*Gladiolus* spp.	USO		
Asparagales	Iridaceae	*Gladiolus undulatus*	USO		
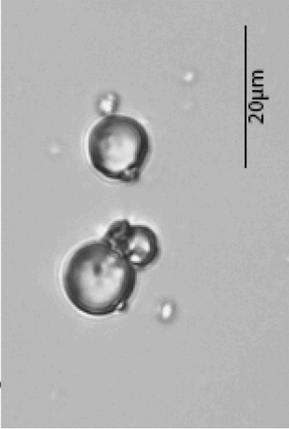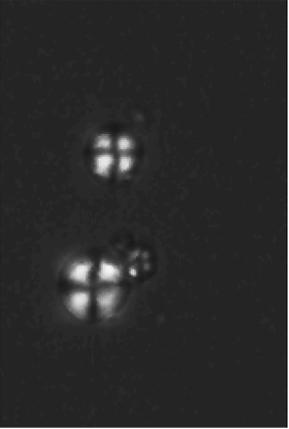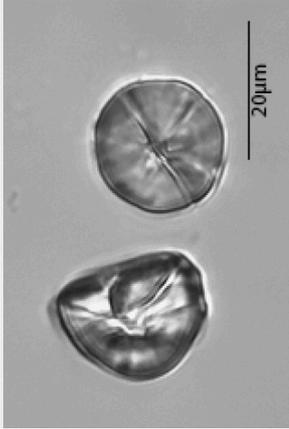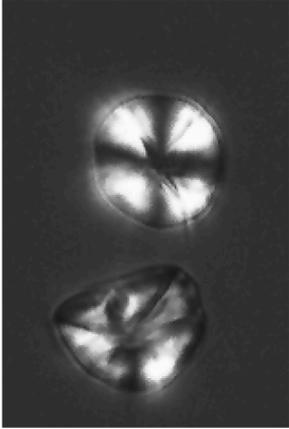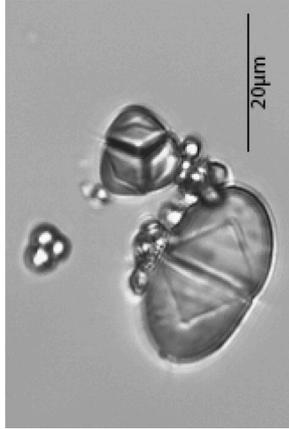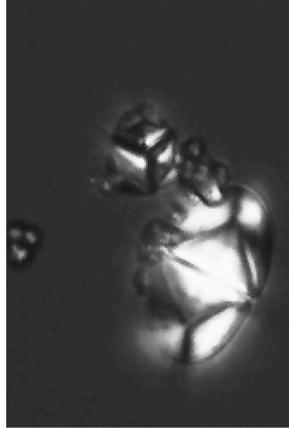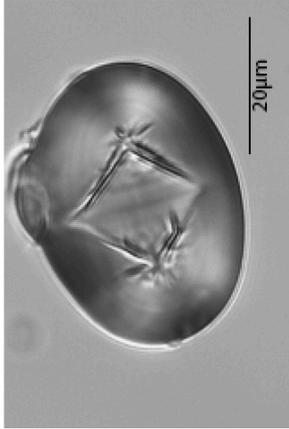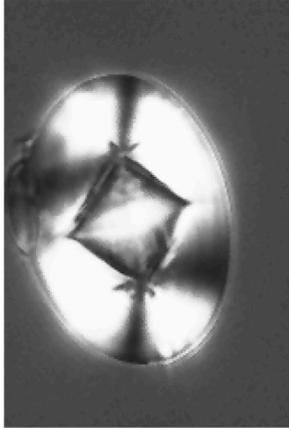
TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Asparagales	Orchidaceae	*Disa spp.*	USO		
Arecales	Arecaceae	*Hyphaene petersiana*	Trunk		
Poales	Cyperaceae	*Cyperus esculentus*	USO		
Poales	Cyperaceae	*Cyperus rotundus*	USO		
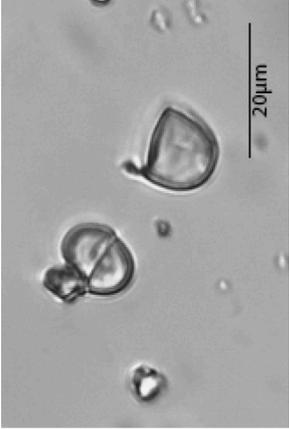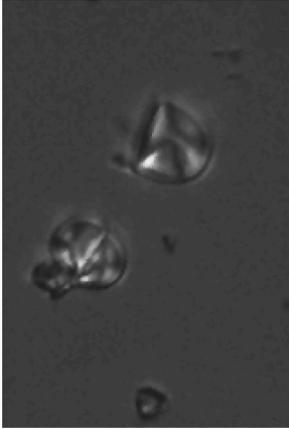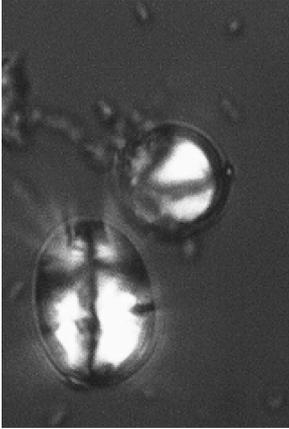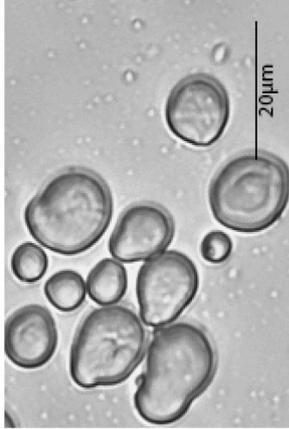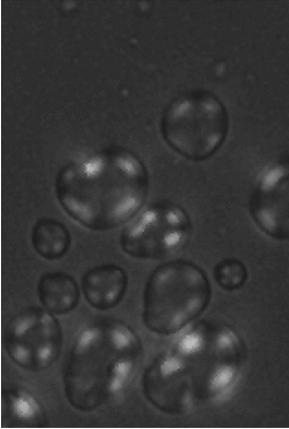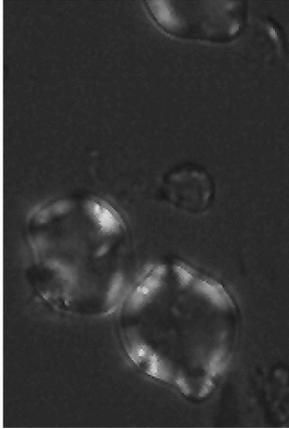
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Poales	Cyperaceae	*Cyperus tomaiophyllus*	USO		
Poales	Poaceae	*Eleusine africana*	Grain		
Poales	Poaceae	*Oxytenanthera abyssinica*	USO		
Poales	Poaceae	*Sorghum bicolor*	Grain		
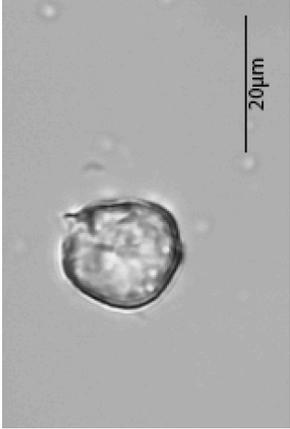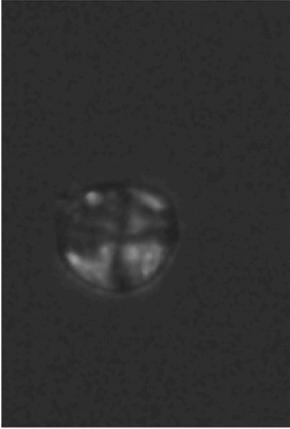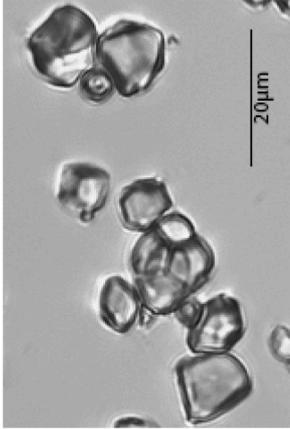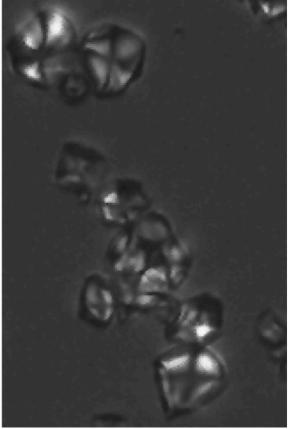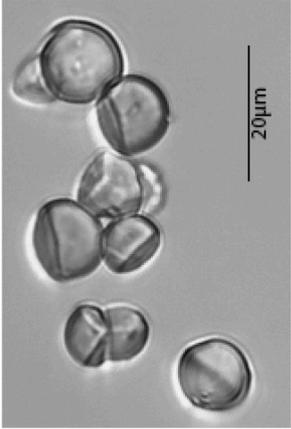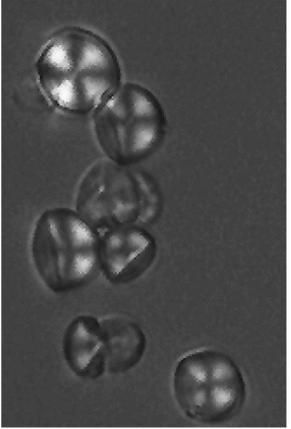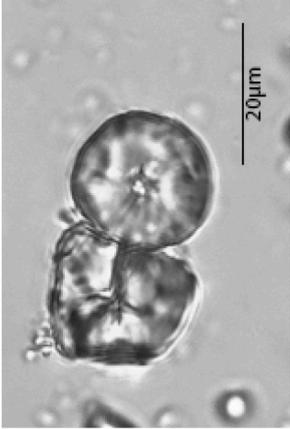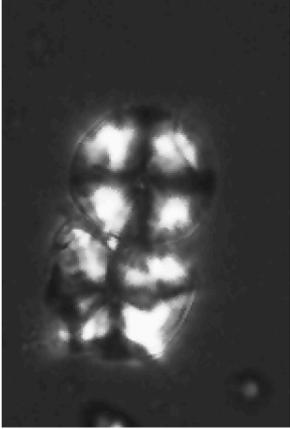
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Poales	Poaceae	*Sporobolus panicoides*	Grain		
Poales	Typhaceae	*Typha capensis*	USO		
Zingiberales	Musaceae	*Ensete ventricosum*	USO		
Zingiberales	Strelitziaceae	*Strelitzia nicolai*	USO		
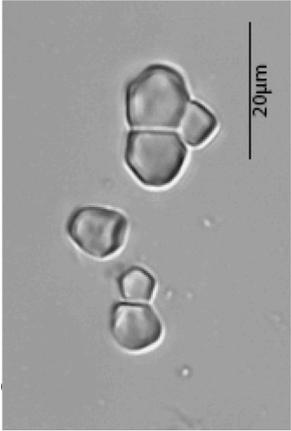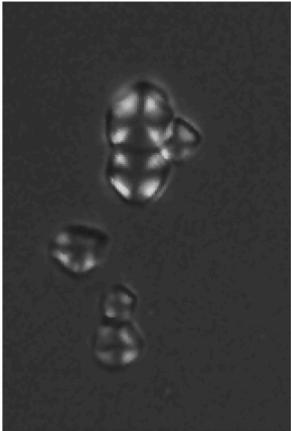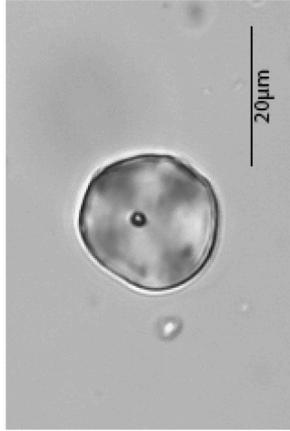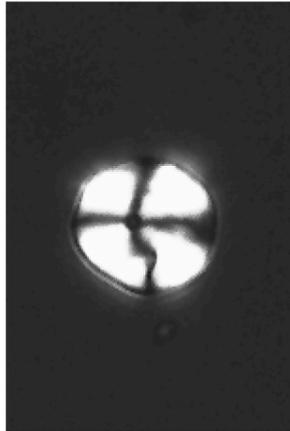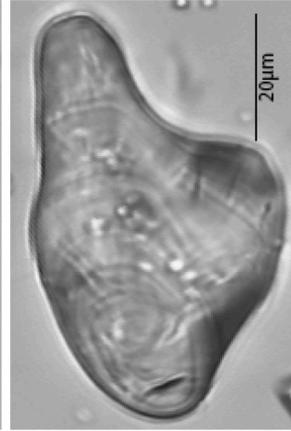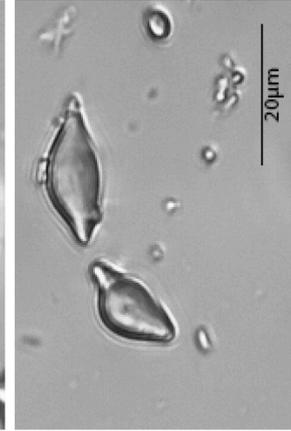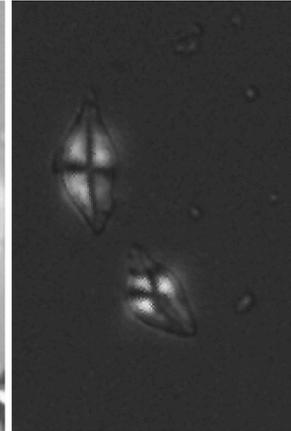
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Ranunculales	Menispermaceae	*Jateorhiza palmata*	USO		
Proteales	Proteaceae	*Protea welwitschii*	USO		
Malpighiales	Euphorbiaceae	*Ricinodendron heudelotii*	Seed		
Malpighiales	Euphorbiaceae	*R. rautanenii*	Seed		
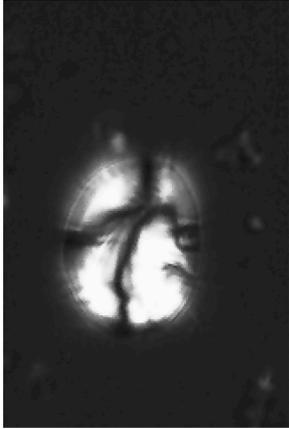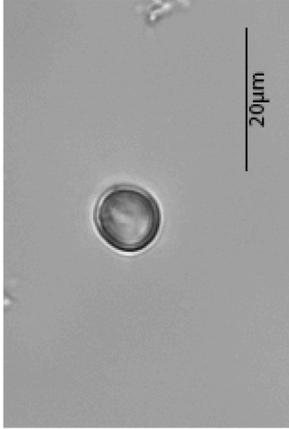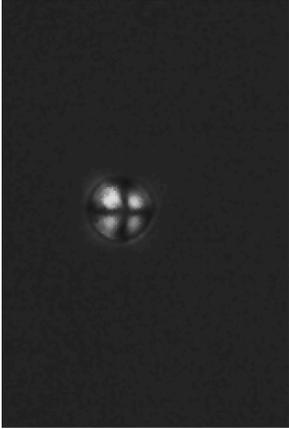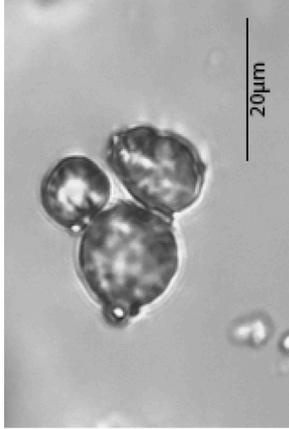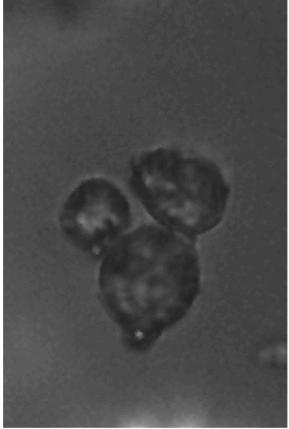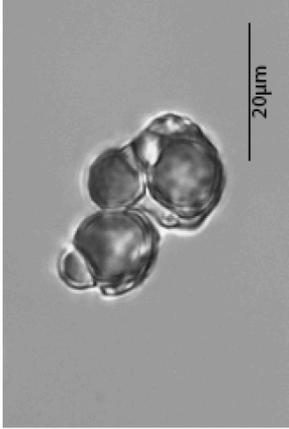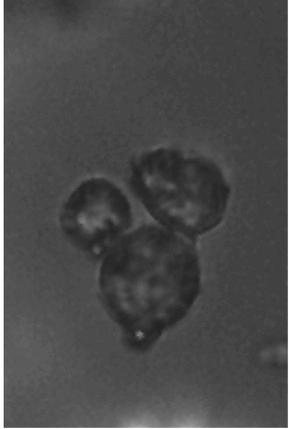
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Fabales	Fabaceae	*Acacia nilotica*	Legume		
Fabales	Fabaceae	*A. tortilis*	Legume		
Fabales	Fabaceae	*Afzelia quanzensis*	Legume		
Fabales	Fabaceae	*Cajanus cajan*	Legume		
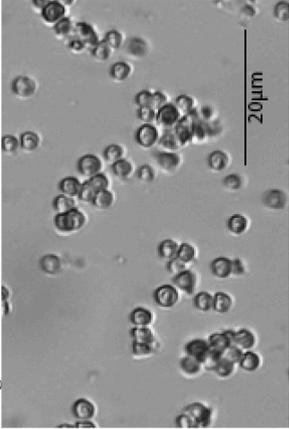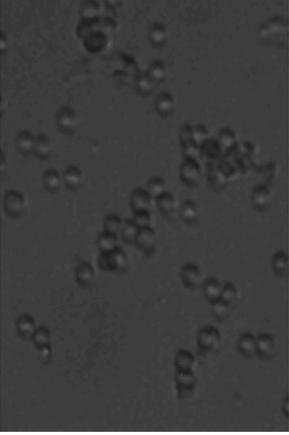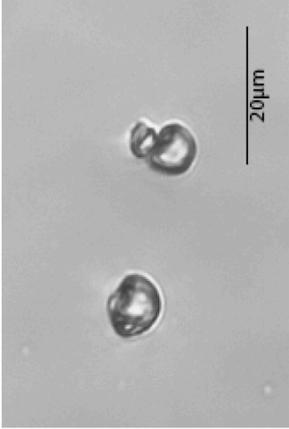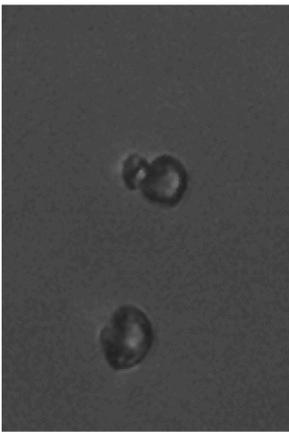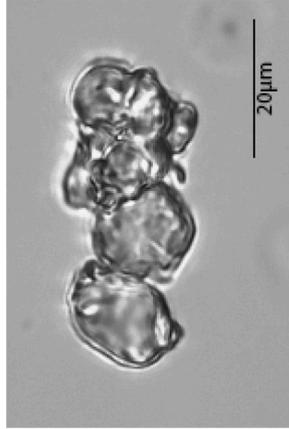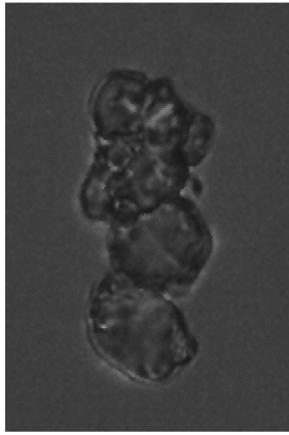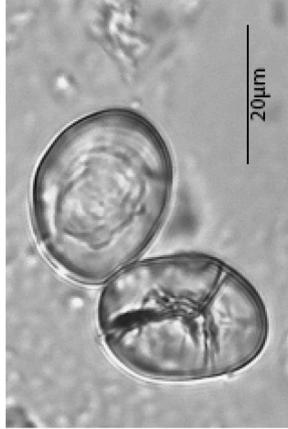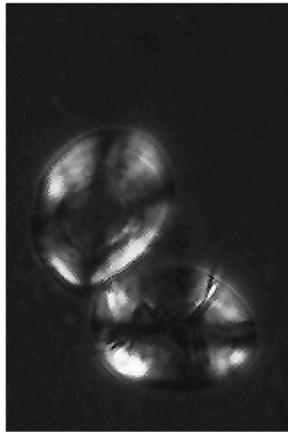
TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Fabales	Fabaceae	*Cassia petersiana*	Legume		
Fabales	Fabaceae	*Cordyla africana*	Legume		
Fabales	Fabaceae	*Eriosema preptum*	USO		
Fabales	Fabaceae	*Faidherbia albida*	Legume		
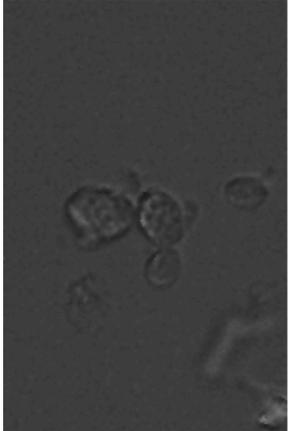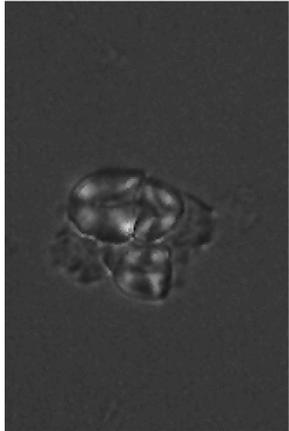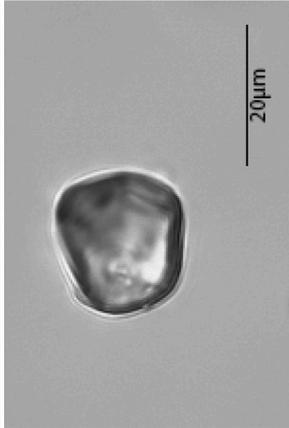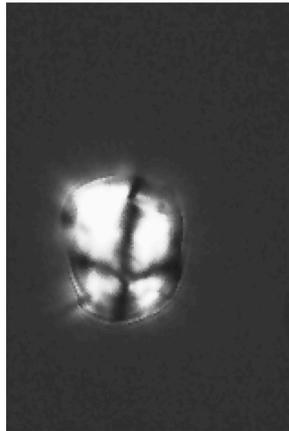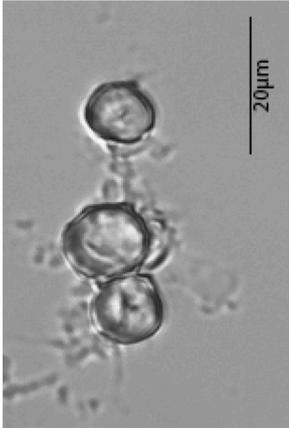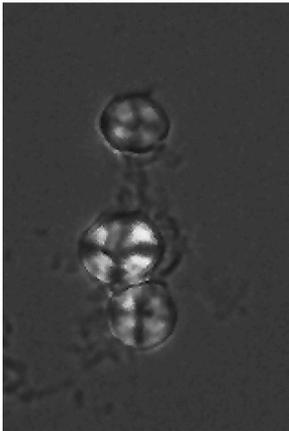
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Fabales	Fabaceae	*Lablab purpureus*	Legume		
Fabales	Fabaceae	*Lablab* spp.	Legume		
Fabales	Fabaceae	*Lotus discolor*	USO		
Fabales	Fabaceae	*Mucuna pruriens*	Legume		
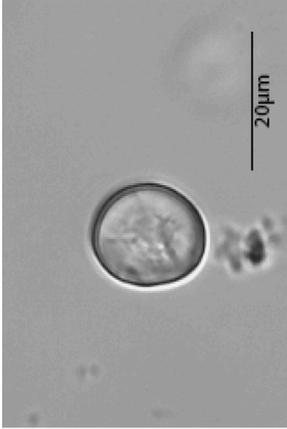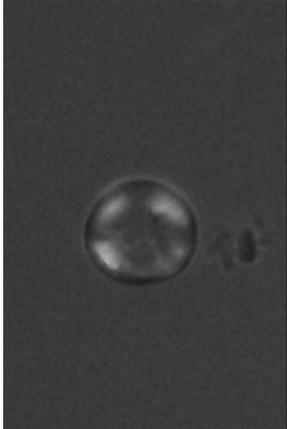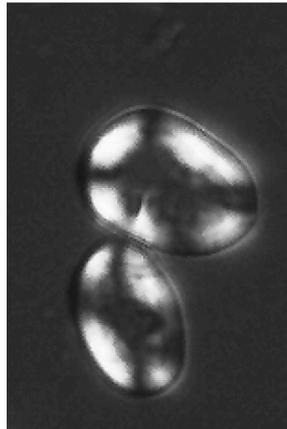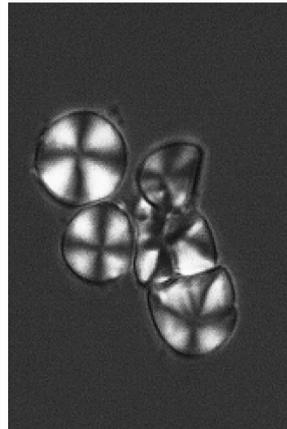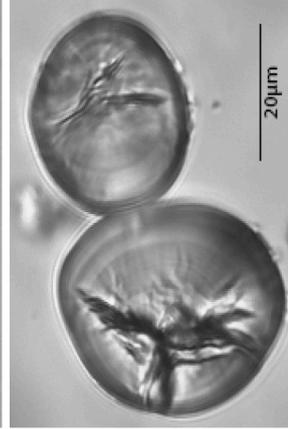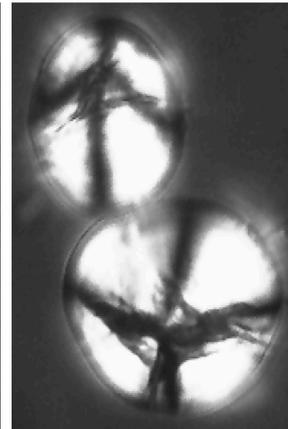
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Fabales	Fabaceae	Mucuna spp.	Legume		
Fabales	Fabaceae	Sphenostylis stenocarpa	USO		
Fabales	Fabaceae	Vigna lobatifolia	Legume		
Fabales	Fabaceae	V. subterranea	Legume		
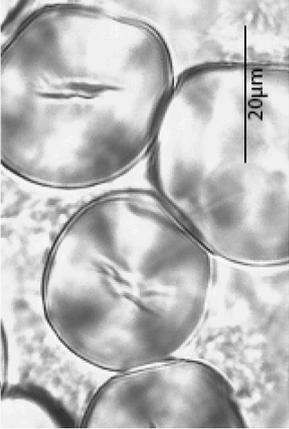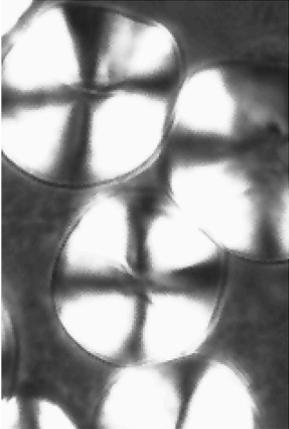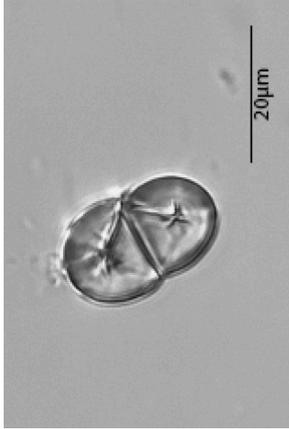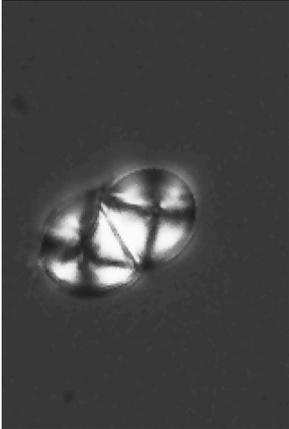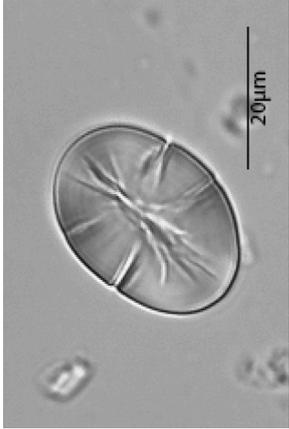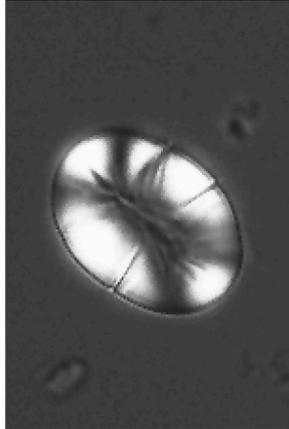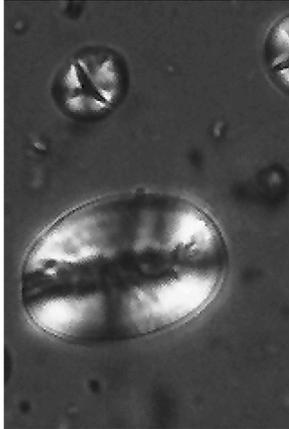
(Continued)

TABLE A1 | Continued

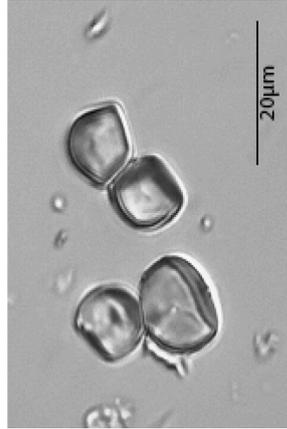| Order	Family	Species	Plant part	Brightfield	Darkfield
Rosales	Cannabaceae	*Celtis africana*	Fruit		
Cucurbitales	Cucurbitaceae	*Acanthosicyos naudinianus*	Seed		
Cucurbitales	Cucurbitaceae	*Momordica cardiospermoides*	Seed		
Myrtales	Combretaceae	*Combretum* spp.	Legume		
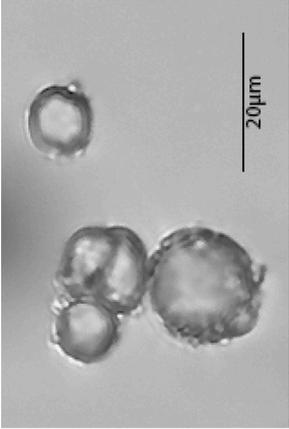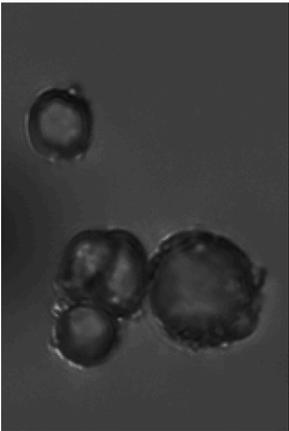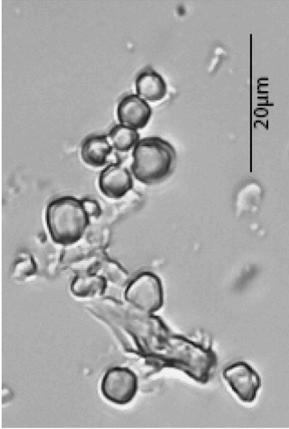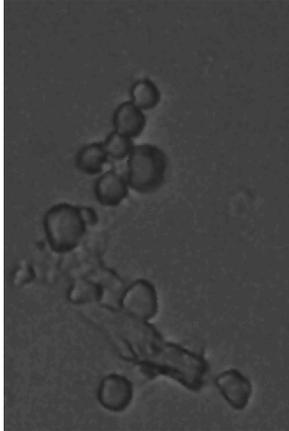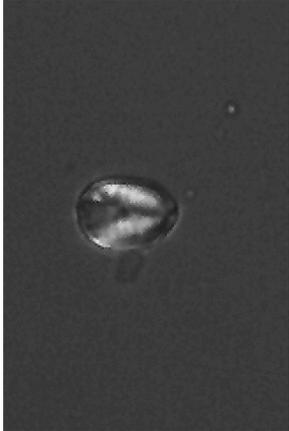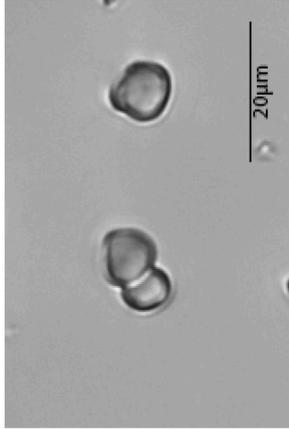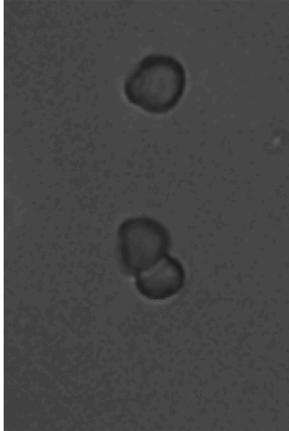
(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Malvales	Bombacaceae	*Adansonia digitata*	Mesocarp		
Caryophyllales	Polygonaceae	*Polygonum senegalense*	USO		
Ericales	Sapotaceae	*Englerophytum magalismontanum*	Fruit		
Solanales	Convolvulaceae	*Ipomoea longituba*	USO		
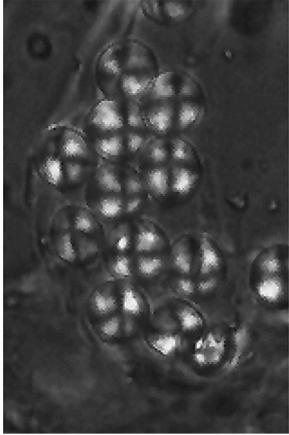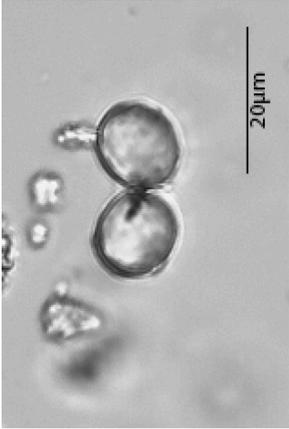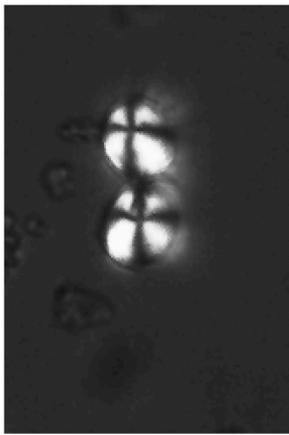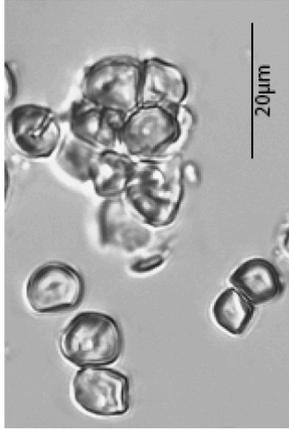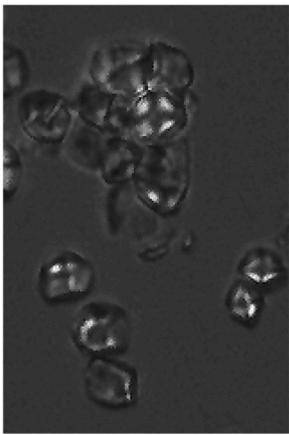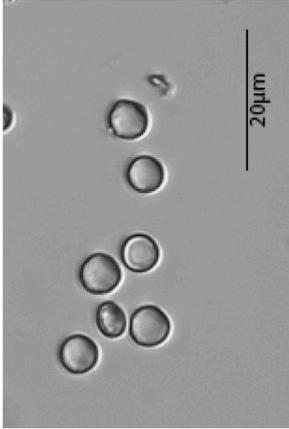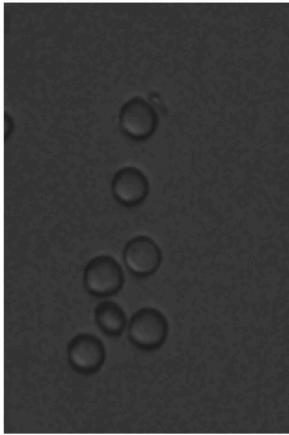
(Continued)

(Continued)

TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Gentianales	Apocynaceae	*Ceropegia rendalii*	USO		
Gentianales	Apocynaceae	*Fockea angustifolia*	USO		
Lamiales	Lamiaceae	*Plectranthus esculentus*	USO		
Lamiales	Lamiaceae	*Solenostemon rotundifolius*	USO		
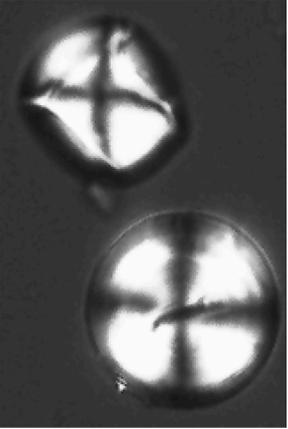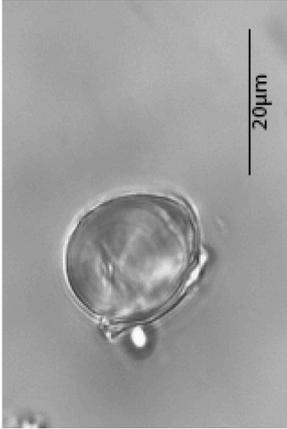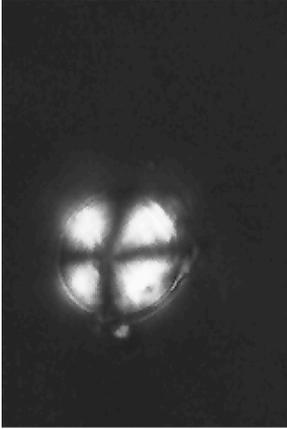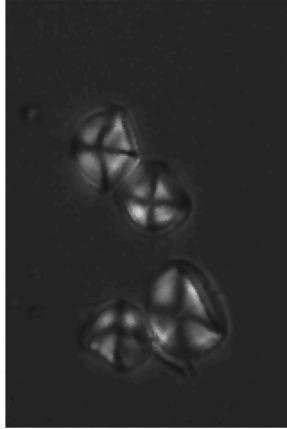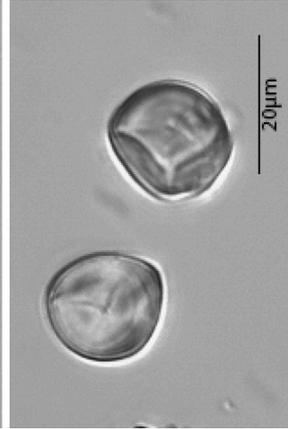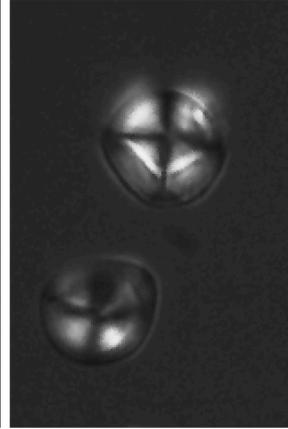
TABLE A1 | Continued

Order	Family	Species	Plant part	Brightfield	Darkfield
Lamiales	Pedaliaceae	*Sesamum angolense*	Seed		