UC Irvine
UC Irvine Previously Published Works

Title
Classification of grass pollen through the quantitative analysis of surface ornamentation and texture

Permalink
https://escholarship.org/uc/item/10t8t1st

Journal
Proceedings of the Royal Society B: Biological Sciences, 280(1770)

ISSN
1471-2954

Authors
Mander, Luke
Li, Mao
Mio, Washington
et al.

Publication Date
2013-11-07

Peer reviewed
Classification of grass pollen through the quantitative analysis of surface ornamentation and texture

Luke Mander1,†, Mao Li2, Washington Mio2, Charless C. Fowlkes3 and Surangi W. Punyasena1

1Department of Plant Biology, University of Illinois, Urbana, IL 61801, USA
2Department of Mathematics, Florida State University, Tallahassee, FL 32306, USA
3Department of Computer Science, University of California, Irvine, CA 92697, USA

Taxonomic identification of pollen and spores uses inherently qualitative descriptions of morphology. Consequently, identifications are restricted to categories that can be reliably classified by multiple analysts, resulting in the coarse taxonomic resolution of the pollen and spore record. Grass pollen represents an archetypal example; it is not routinely identified below family level. To address this issue, we developed quantitative morphometric methods to characterize surface ornamentation and classify grass pollen grains. This produces a means of quantifying morphological features that are traditionally described qualitatively. We used scanning electron microscopy to image 240 specimens of pollen from 12 species within the grass family (Poaceae). We classified these species by developing algorithmic features that quantify the size and density of sculptural elements on the pollen surface, and measure the complexity of the ornamentation they form. These features yielded a classification accuracy of 77.5%. In comparison, a texture descriptor based on modelling the statistical distribution of brightness values in image patches yielded a classification accuracy of 85.8%, and seven human subjects achieved accuracies between 68.33 and 81.67%. The algorithmic features we developed directly relate to biologically meaningful features of grass pollen morphology, and could facilitate direct interpretation of unsupervised classification results from fossil material.

1. Introduction

The taxonomic identification of pollen and spores, in common with many other biological sciences that rely on morphological comparison, uses inherently qualitative descriptors of shape and ornamentation. As a result, identifications are restricted to taxonomic groupings that can be reliably classified by multiple analysts and subtle morphological differences are often ignored. A consequence of this conservative approach is the coarse taxonomic resolution of the pollen and spore record [1]. The development of intuitive, numerical measures of shape and ornamentation would directly address these limitations of pollen and spore identification, and allow researchers to translate morphological differences observed under the microscope into quantitative, repeatable measurements. Treating morphology as measurement, rather than description, allows a broader range of observations to be incorporated into the analysis and identification of pollen and spores.

Grass pollen presents a classic demonstration of the taxonomic limits of current analytical approaches. The grass family (Poaceae) is an exceptionally successful group of plants, and can be found in a wide range of habitats from the tropics to the arctic. However, although Poaceae contains more than 11,000 recognized species [2] and its genetic diversity is visible in the diverse morphology of grass flowers [3], the gross morphology of grass pollen is remarkably similar throughout the family. The pollen is generally spheroidal with a single pore surrounded by an annulus [4,5]. This simple morphology
has led researchers to suggest that pollen morphology is ‘uniform’ within the family [6] and has little to contribute to the reconstruction of grass evolution and diversification. Instead, direct evidence for the palaeoecology and evolutionary history of grasses has been provided mostly by other fossil groups such as phytoliths (microscopic silica bodies formed in plant tissues) [6,7].

Yet, owing to their high abundance in terrestrial and marine sediments, and standardized protocols that allow relative abundances of different plant groups to be directly compared through time, pollen grains provide a potentially rich source of information on the evolutionary history of grasses. As a result, palynologists have attempted to increase the taxonomic precision of grass pollen by measuring characters such as pollen grain length, grain width, pore diameter and annulus width [8–11], and by noting differences in the organization of the grass pollen exine [5,9]. High-resolution scanning electron microscopy (SEM) studies have revealed a diversity of surface ornamentation patterns that may have taxonomic significance [5,12–16] but are not visible when viewed using traditional light microscopy. However, exine ornamentation has not been widely used to classify grass pollen because of the difficulty in comparing the relatively small differences in surface patterning [5,13,14].

In this paper, we classify 240 specimens of grass pollen from 12 species in three subfamilies within Poaceae (table 1) by quantifying the size and density of sculptural elements and the complexity of the surface ornamentation that they form. In doing so, we develop a means of quantifying morphological features that have traditionally been described solely in qualitative terms. Our results provide a potential solution to the problem of classifying grass pollen, and demonstrate how computational image analysis, combined with high-resolution microscopy, holds the potential to dramatically increase the taxonomic resolution of pollen and spore records of Earth’s vegetation [18,19].

2. Scanning electron microscopy image acquisition

Grass pollen from the 12 species was prepared for SEM imaging using standard palynological methods (see the electronic supplementary information). Twenty grains of each species were imaged at ∗2000, ∗6000 and ∗12 000 magnification (see the electronic supplementary information, dataset S1). Analyses were undertaken on 400 × 400 pixel windows that were manually cropped from the ∗6000 images (see the electronic supplementary information, dataset S2). In these images, 1 pixel measures 16.(6).nm. SEM images of all grass pollen analysed in this paper can be found at https://www.ideals.illinois.edu/handle/2142/43358.

3. Grouping morphologically similar species

(a) Quantifying sculptural element size and density

SEM revealed that the surface ornamentation of the 12 species is considerably diverse. Using terminology from descriptive palynology [20], some species are characterized by scabrate ornamentation consisting of granula, whereas others have more complex areolate ornamentation, with polygonal areas separated by grooves that form a negative reticulum. We first clustered the species into groups with similar patterns of surface ornamentation. To do this, we developed a scalar feature \( \lambda \) that represents the size of the sculptural elements on the surface of each specimen, and a feature \( \nu \) that is related to the density of the granula on the surface of each specimen (figure 1).

The size of sculptural elements on the surface of each pollen grain was quantified from five windows measuring 100 × 100 pixels randomly chosen and cropped from each ×6000 SEM image of each specimen. Working with smaller windows decreased the influence of variations in the brightness and contrast of each image. A Sobel edge detector was applied to each 100 × 100 window, which turned each of these into a binary image with the white pixels representing the detected edges (figure 1b,f). The white pixels were then clustered by forming a graph with the white pixels as nodes and two nodes connected with an edge if one node

<table>
<thead>
<tr>
<th>subfamily</th>
<th>genus and species</th>
<th>species code</th>
<th>group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pooidae</td>
<td>Anthoxanthum odoratum</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Dactylis glomerata</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Phalaris arundinacea</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Poa australis</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Stipa tenuifolia</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Chloridoideae</td>
<td>Cynodon dactylon</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Eragrostis mexicana</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Sporobolus pyramidalis</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Triodia basedowii</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Panicoideae</td>
<td>Bothriochloa intermedia</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Digitaria insularis</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Opilimenus hirtellus</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>
in the analysis of sculptural element size to identify the connected components of $H$. We used the number of connected components $n_H$ to estimate the number of granula in each $100 \times 100$ pixel window, and used the average value $v$ over the five $100 \times 100$ windows as a feature related to the density of the granula on the surface of each specimen. To balance out the scales of $t$ and $v$ that are related to sculptural element size and density, respectively, we centred both features by subtracting their mean values and scaling them to have unit variance.

(b) Identification and validation of morphological groups
Using a scatterplot of features $t$ and $v$ as a guide, we clustered the 12 species into four groups (figure 2 and table 1). These four morphological groups were validated by a leave-one-out experiment with a $k$-nearest-neighbour classifier in which 228 of the 240 specimens were classified correctly (95% accuracy; we selected $k = 9$ based on classification performance). Group one contains five species that are characterized by granula spread relatively densely over the surface of the pollen grain and poorly defined areolae (figure 2). Group two contains three species that are characterized by areolae and granula spread relatively densely over the surface of the pollen grain (figure 2). Group three contains three species that are characterized by areolae and granula distributed relatively sparsely over the pollen surface (figure 2). The simple scabrate surface ornamentation of Stipa tenuifolia lacks areolae and has sparsely distributed granula. This surface ornamentation is sufficiently different to any of the other species that it clusters alone (figure 2).

4. Classification of species within each morphological group
Next, we classified the species of grass contained within each of the four morphological groups. To do this, we approached the pixels in each image of a pollen grain as nodes in a network, with two nodes connected with an edge using a neighbouring relation. We treated the sculptural elements on the surface of a pollen grain as foreground objects (white pixels in figure 3c), and used the notion of network centrality to define two 20-dimensional features, $\tau_1$ and $\tau_2$, that provide measures of the complexity of the surface ornamentation on each pollen grain. Centrality here refers to the relative importance of the nodes of a network for its local–global connectivity [23].

(a) Quantifying the complexity of grass pollen surface ornamentation and construction of features $\tau_1$ and $\tau_2$
Five windows measuring $40 \times 40$ pixels were randomly chosen and cropped from each $\times 6000$ SEM image of each specimen [22] within these groups. Minimum variance quantization was used to cluster the pixels in each window into four groups and turn each window into a binary image. For species within group one (figure 2), the pixels in the two clusters of lower intensity were turned into black pixels (background) and those in the two clusters of higher intensity were turned into white pixels (foreground; figure 3d). For species within groups two and three (figure 2), only pixels in the cluster of lowest intensity were turned into black pixels (background) and those in the other three were turned into white pixels.
the adjacency matrix of the network as follows: Let \( C_2 \) be an edge if they are immediate neighbours to the north, south, east or west. SC was used to rank the pixels of this network. A sequence of 20 expanding subregions of the background was formed, starting with the pixels ranked in the top 5% and adding the next 5% until the entire background of black pixels was covered (figure 4a). For each of the subregions, the number of connected components was calculated, which were recorded in a 20-dimensional feature vector \( \tau_m \). As pixels are added to the shape, the number of connected components may increase or decrease and existing components may coalesce (figure 4a). The last component of the vector \( \tau_m \) is the number of connected components of the background. This process was repeated using each of the five randomly cropped \( 40 \times 40 \) pixel windows from each specimen. The average was taken to obtain a 20-dimensional feature vector \( \tau_1 \).

Feature \( \tau_2 \) is derived from a weighted network having all 1600 pixels of each \( 40 \times 40 \) window as nodes, and two nodes connected with an edge using the same neighbouring relation as for feature \( \tau_1 \). Edges connecting two foreground (white)
pixels or two background (black) pixels received weight 1, whereas background–foreground transition edges received weight \( a \), \( 0 < a < 1 \). Experiments indicate that \( \tau_2 \) remains essentially unchanged for \( a \leq 0.01 \), so we used \( a = 0.01 \) in all calculations. As for feature \( \tau_1 \), SC was used to rank the pixels of this network and a sequence of 20 expanding sub-regions of the 40 \( \times \) 40 window was formed (figure 4b,c). The number of connected components was calculated at each stage of the sequence (figure 4b,c), which was recorded in a 20-dimensional feature vector. This process was repeated using each of the five randomly cropped 40 \( \times \) 40 pixel windows from each specimen. The average was taken to obtain a 20-dimensional feature vector \( \tau_2 \).

(b) Species classification of grass pollen
To optimize the combination of \( \iota \) (sculptural element size) and \( \upsilon \) (sculptural element density) with the feature \( \tau \) derived from centrality, we allow a scaling factor \( a > 0 \), and use the 22-dimensional feature vector \( Y = (a_i, a_\upsilon, \tau) \) to classify species. For each group, we experimented with a range of values of \( a \), and selected the value that yielded the highest classification accuracy. We employed either feature \( \tau_1 \) or feature \( \tau_2 \) depending on which feature yielded the highest classification accuracy. Feature vector \( Y \) was reduced to three dimensions using principal component analysis.

As \( S. \) tenuifolia clusters alone using features \( \iota \) and \( \upsilon \), we only further subdivide species in groups one, two and three (figure 2). Classification of species within these groups was validated with a leave-one-out experiment using a \( k \)-nearest-neighbour classifier. The choice of \( k \) was based on classification performance. Within group one, 78 out of 100 specimens were classified correctly at the species level (78%) using feature \( \tau_2 \), \( a = 15.8 \) and \( k = 5 \). For this group, feature \( \tau_2 \) quantifies the complexity of the patterning formed by the granula on the pollen surface (figure 4c). (Typically, for complex patterning, the vector \( \tau \) will have several larger values that reflect the presence of numerous connected components. In many cases, the values of \( \tau \) also show an oscillatory behaviour indicating that many new connected components are created and existing ones get merged as more pixels are added to the count.) Within group two, 50 out of 60 specimens were classified correctly (83%) using feature \( \tau_2 \), \( a = 8 \) and \( k = 6 \). For this group, feature \( \tau_2 \) quantifies the complexity of the patterning formed by the negative reticulum and the areolae (figure 4b). Within group three, 50 out of 60 specimens were classified correctly (83%) using feature \( \tau_1 \), \( a = 5 \) and \( k = 3 \). For this group, feature \( \tau_1 \) quantifies the complexity of the patterning formed by the negative reticulum (figure 4a).

Our classification of the 12 species investigated is shown schematically as a tree in figure 5. At the first level, we use features \( \iota \) and \( \upsilon \) to cluster the 12 species into groups that have similar patterns of surface ornamentation (figure 5). At the second level, we use the 22-dimensional feature vector \( Y \) to classify species within groups one, two and three (figure 5). Classification errors are introduced at the first and second levels of the classification. A leave-one-out experiment with all 240 specimens yielded a classification accuracy of 77.5% (see the electronic supplementary material, figure S1). Given that leave-one-out cross-validation was used to estimate model parameters and the predictive performance of \( k \)-nearest-neighbour classifiers, the performance rates reported in the paper are subject to the general limitations of the cross-validation procedure, as estimates so obtained may exhibit considerable variation (cf. [24]).

5. Comparison with an alternative texture descriptor and human subjects
As a baseline for the species-level classification accuracy of our approach, we evaluated a texture descriptor based on the modelled statistical distribution of brightness values in local image patches. This approach is common in computational image analysis and has been used previously for classifying materials [25] and detection of boundaries between textured regions [26]. Each cropped \( \times 6000 \) SEM image of grass pollen grain was down-sampled to a resolution of 120 \( \times \) 120. Each 5 \( \times \) 5 pixel window centred at every pixel location within the image (excluding boundaries) was quantized by subtracting the mean brightness of the patch and assigning it to the closest element in a pre-computed dictionary of canonical appearances. We used a generic dictionary containing 75 elements learned from a set of consumer photographs by minimizing sparse reconstruction error (graciously provided by Ren & Bo [27]). The relative frequencies of these appearances were stored in a 75-bin histogram and left-out samples were classified using \( k \)-nearest-neighbour with an \( L_1 \) distance (cityblock). \( k = 3 \) was selected by cross-validation. A histogram recording the frequencies of the different patch appearances in a sample image was used as the texture descriptor for the sample. \( k \)-nearest-neighbour classification using this histogram descriptor with \( k = 3 \) and leave-one-out cross-validation yielded a final species-level classification accuracy of 85.8% (see the electronic supplementary material, figure S2).

We next measured the ability of seven human subjects to classify the same SEM images of grass pollen that were used for our computational image analyses (see the electronic supplementary material, dataset S2). Each subject was provided with a reference library containing six images of each species of grass pollen, grouped and labelled by species. Each subject was then provided with a set of 120 unlabelled images of grass pollen containing 10 images of each of the 12 species. Two of these 10 images were duplicates, and one of these 10 images also appeared in the reference library. Images were classified by assigning each image to one of the 12 species. The images and classification key are provided in the electronic supplementary material, dataset S3.

The classification accuracy of the seven human subjects ranged from 68.33 to 81.67% (average 75.48%; figure 6).
Palynological studies of grass [5,13,14]. They include the...

However, classification consistency between subjects was low; only 28.33% of the specimens were classified correctly by all seven subjects. This consistency falls to 21.82% when the visually distinctive S. tenuifolia is excluded from the analysis.

6. Discussion

Quantitative image-based analyses can enumerate the small, subtle and diverse morphological differences among taxa that may be observed by the expert, but cannot easily be conveyed by the terminology used to describe pollen and spores [18]. This has been a significant barrier to the classification of grass pollen [5,13,14]. For example, the patterns of surface ornamentation revealed by SEM have been used to define morphotypes that contain many species of grass [13,14]. However, these morphotypes form a continuum rather than clearly defined boundaries between them [13], and this is partly responsible for their limited use in palynological studies of grass [5,13,14]. They include the Hordeum-type, Triticum-type, Avena-type and Setaria-type [14]. The Setaria-type, for example, is characterized by extensive field-like areolae of irregularly polygonal outlines. Their bulging surface is studded with very small pointed spinules, in most cases, (3–)5–8(–10) [sic] (14), p.139).

Our analyses demonstrate how the morphological characters that are described by such terminology can be quantified and used to classify grass pollen.

We have attempted to develop and use features that can be directly related to biologically significant characteristics of grass pollen (figures 1, 3 and 4). For example, the clustering step in our classification uses features related to sculptural element size and density (figures 1 and 2). The four morphological groups that are produced are consistent with visual perception of surface ornamentation (figure 2) and this step is analogous to the description of grass pollen morphotypes [5,13,14], but using measurements of morphology rather than qualitative descriptions.

Our approach is rooted in the accurate description of morphology, and achieves classification results that are comparable to the results of manual human classification (figure 6) and a more conventional computational image analysis based on the distribution of brightness values in local image patches [25,26]; see the electronic supplementary material, figures S1 and S2). These comparisons highlight that there are a variety of analytical techniques that could be used to classify grass pollen once sufficient morphological information has been recovered from individual specimens. However, in our experiments with human subjects, classification accuracy comes at the expense of consistency, highlighting that the taxonomic resolution of the pollen and spore record can be reduced by disagreement among multiple analysts [18,28]. The patch appearance histogram approach resulted in higher identification accuracies for 10 of the 12 species, but this increase in accuracy (approx. 8% on average) comes at the expense of interpretability. The histogram counts encode the image appearance in a distributed way and, in contrast to the features we developed in this paper (figures 1 and 4), are not easily understood in terms of basic morphological features of pollen grains. We anticipate that the use of features that can be directly related to biologically significant morphological characters will be critical in the interpretation of fossil samples, where extant reference specimens are unavailable, and all possible morphologies are not known. Interpretability is important in this context, as character-based features will produce unsupervised clusters that are more intuitively understood by palynological experts. This is not always true of the supervised learning approaches popular in automated pollen classification, which generally provide low interpretability of biological features [18,29].

Additionally, the development of quantitative measures of shape and ornamentation has wide applications in other branches of the biological sciences that rely on visual inspection for classification of phenotypic differences. The methods to quantify morphology that we have developed in this paper have immediate utility for the classification of other groups of organisms for which the patterns of surface ornamentation are an important taxonomic character, such as diatoms and ostracods, as well as the potential to quantify the complexity of other biological structures, such as venation in leaves and insect wings.

7. Conclusion

In this paper, we have used a combination of high-resolution microscopy and computational image analyses to classify 12 species of modern grass pollen. We have classified these species by developing features that quantify the size and density of sculptural elements (figure 1) and measure the complexity of the surface ornamentation that they form (figures 3 and 4). These features can be understood in terms of the basic morphological features of the grass pollen exine. In our experiments using this algorithmic method, 186 out of 240 specimens were classified correctly, yielding a classification accuracy of 77.5%. We also compared a baseline texture classification approach using histograms of local quantized image patches [25,26], which yielded an accuracy of 85.8% but provided low interpretability. Seven human subjects achieved classification accuracies between 68.33 and 81.67% (figure 6) on a subset of these images. However, classification consistency...
between subjects was low, and just 28.33% of the specimens were correctly classified by all subjects.

Our results support the view that a combination of high-resolution microscopy and computational image analyses can generate classifications at fine taxonomic levels that are beyond the capability of human experts [18,28]. This approach has the potential to dramatically increase the taxonomic resolution of pollen and spore records of ancient vegetation, which will in turn expand the range and depth of hypotheses that can be tested using the fossil record [18,19].

Acknowledgements. Michael Urban provided the specimens of grass pollen used in this paper. Margaret Collinson provided advice on the preparation of pollen grains for SEM. Claire Belcher, Sarah Baker, Derek Haselhorst, Jacqueline Rodriguez, Shivangi Tiwari and Cassandra Wesseln produced the human baseline data. We thank our two anonymous reviewers for their comments and Carlos Jaramillo for encouraging us to pursue the problem of grass pollen classification. This research was carried out in part in the Frederick Seitz Materials Laboratory Central Facilities, University of Illinois.

Data accessibility. SEM images of grass pollen and files used to test human classification performance: University of Illinois IDEALS digital archive (https://www.ideals.illinois.edu/handle/2142/43358).

Funding statement. We acknowledge funding from the National Science Foundation (DBI-1052997 to S.W.P., DBI-102942 to W.M. and DBI-1053036 to C.C.F.). L.M. was partly supported by a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme (PIIF-GA-2012-328245).

References

11. Flenley JR, King SM. 1984 Late Quaternary pollen records from Easter Island. Nature 307, 47—50. (doi:10.1038/307047a0)