**Supplementary Figure 1.** Results of the principal component analysis (PCA) showing variation across PC3 and PC4. The entire *Skiagia* dataset has been included (N = 191). The proportion of variation explained by each PC is shown along their respective axes. Filled and empty symbols represent ‘unambiguous’ and ‘ambiguous’ specimens, respectively, and are color-coded according to their (tentative) species attribution. Inset on the bottom-left corner shows the PC3–PC4 biplot.

**Supplementary Figure 2.** Results of the principal component analysis (PCA) with convex hulls encompassing specimens from each sample. The proportion of variation explained by each PC is shown along their respective axes. Filled and empty symbols represent ‘unambiguous’ and ‘ambiguous’ specimens, respectively, and are color-coded according to their (tentative) species attribution. Blue: *S. scottica*; black: *S. ciliosa*; purple: *S. orbiculare*; green: *S. ornata*, yellow: *S. compressa*.

**Supplementary Figure 3.** Loadings of the first two discriminant functions (DF) calculated based on the reduced dataset of 83 ‘unambiguous’ specimens. Positive and negative coefficients are shown as black and white bars, respectively.

**Supplementary Figure 4.** Linear discriminant analysis (LDA) of ‘unambiguous’ specimens from each sample, grouped per species. A, LDA of sample 184002 (n = 36). B, LDA of sample 184003 (n = 86). C, LDA of sample 184004 (n = 69). ‘Ambiguous’ specimens (empty symbols, light-colored convex hulls) were plotted a posteriori and color-coded according to their tentative species attribution. The proportion of between-group variance explained by each discriminant function (DF) is shown along their corresponding axes. Biplots are shown as insets.

**Supplementary Figure 5.** Linear discriminant analysis (LDA) of the complete *Skiagia* dataset (N = 191), grouped per sample. Empty and filled symbols represent ‘ambiguous’ and ‘unambiguous’ specimens, respectively, and are color-coded according to their (tentative) species attribution. Blue: *S. scottica*; black: *S. ciliosa*; purple: *S. orbiculare*; green: *S. ornata*, yellow: *S. compressa*. All specimens contributed to the calculation of discriminant functions, considering sample attribution as the categorical variable. The proportion of between-group variance explained by each discriminant function (DF) is shown along their corresponding axes.

**Supplementary Figure 6.** Linear discriminant analysis (LDA) of the complete *Skiagia* dataset (N = 191), grouped per species. Both ‘unambiguous’ and ‘ambiguous’ specimens (filled and empty symbols, respectively) were used to calculate discriminant functions (DF), and were color-coded according to their (tentative) species attribution. The proportion of between-group variance explained by each discriminant function (DF) is shown along their corresponding axes.

**Supplementary Figure 7.** Relative abundance of the main *Skiagia* species in all three samples. Row data is available in Supplementary Table 2.

**Supplementary Table 1.** Correlation coefficients and *p*-values for each pair of variables used for PCAs and LDAs. *P*-values are shown in the upper right triangle of the matrix. Significant values (*p* < 0.05) are highlighted. Spearman's *rs* coefficients are shown in the lower left triangle of the matrix. Grey cells are figures referred to in the main text.

**Supplementary Table 2.** Counts of life cycle-related features among *Skiagia* species and samples. ‘inner bodies’ refer to dark translucent to opaque, clearly delimited subcircular areas within the vesicle cavity. ‘opened’ vesicles are vesicles showing a median split. 'Closed empty' vesicles are immature vesicles lacking an inner body.