Research Article

Received: 17 October 2022 Revised: 10 July 2023 Accepted article published: 15 July 2023 Published online in Wiley Online Library: 2 August 2023

(wileyonlinelibrary.com) DOI 10.1002/jsfa.12861

Starch granule size and shape characterization of yam (*Dioscorea alata* L.) flour using automated image analysis

Mahugnon Ezékiel Houngbo,[a,b] Lucienne Desfontaines,[c] Jean-Luc Irep,[d] Konan Evrard Brice Dibi,[e] Maritza Couchy,[c] Bolanle O. Otegbayo[f] and Denis Cornet[a,b*]

Abstract

BACKGROUND: Roots, tubers and bananas (RTB) play an essential role as staple foods, particularly in Africa. Consumer acceptance for RTB products relies strongly on the functional properties of, which may be affected by the size and shape of its granules. Classically, these are characterized either using manual measurements on microscopic photographs of starch colored with iodine, or using a laser light-scattering granulometer (LLSG). While the former is tedious and only allows the analysis of a small number of granules, the latter only provides limited information on the shape of the starch granule.

RESULTS: In this study, an open-source solution was developed allowing the automated measurement of the characteristic parameters of the size and shape of yam starch granules by applying thresholding and object identification on microscopic photographs. A random forest (RF) model was used to predict the starch granule shape class. This analysis pipeline was successfully applied to a yam diversity panel of 47 genotypes, leading to the characterization of more than 205 000 starch granules. Comparison between the classical and automated method shows a very strong correlation ($R^2 = 0.99$) and an absence of bias for granule size. The RF model predicted shape class with an accuracy of 83%. With heritability equal to 0.85, the median projected area of the granules varied from 381 to 1115 μm^2 and their observed shapes were ellipsoidal, polyhedral, round and triangular.

CONCLUSION: The results obtained in this study show that the proposed open-source pipeline offers an accurate, robust and discriminating solution for medium-throughput phenotyping of yam starch granule size distribution and shape classification.

Supporting information may be found in the online version of this article.

Keywords: starch granule; root tuber and banana crops; high-throughput phenotyping; image analysis; *Dioscorea alata* L.

INTRODUCTION

Roots, tubers and bananas (RTB) such as yam crops are the main staple food in the diets of people in parts of sub-Saharan Africa, Asia and Latin America, making it vital for food security.[1] Yam plays an important role as a source of calories, nutrients and fiber in local diets, helping to ensure food security for people in production areas.[2,3] Once prioritizing yield, dry matter, and disease and pest resistance, breeding programs must now focus on end-product quality characteristics linked to processor and consumer preferences.[4] However, high-throughput phenotyping of RTB quality is still a challenging task and up to now has relied on indirect measures of composition or functional properties related to it.

Starch is the main component (60–90% of dry matter) of RTB[5] and the quality of food products is strongly correlated with its composition[6] and functional properties.[5,7-9] Padonou *et al.*[10] have

* *Correspondence to: D Cornet, CIRAD Montpellier Centre, UMR Agap Institut, F-34398 Montpellier, France. E-mail: denis.cornet@cirad.fr*

a *CIRAD, UMR AGAP Institut, F-34398 Montpellier, France*

b *UMR AGAP Institut, Université de Montpellier, CIRAD, INRAE, Institut Agronomie, F-34398 Montpellier, France*

c *INRAE, UR 1321 ASTRO Agrosystèmes tropicaux. Centre de recherche Antilles-Guyane, Petit-Bourg, France*

d *INRAE, UE 0805 PEYI, Centre de recherche Antilles-Guyane, Petit-Bourg, France*

e *CNRA, Station de Recherche sur les Cultures Vivrières, Bouaké, Côte d'Ivoire*

f *DFST, Bowen University, Iwo, Nigeria*

shown that mealiness, an important acceptance criteria for boiled cassava, could be predicted from starch functional properties such as apparent viscosity after pasting. Similarly, the higher the viscosity of the starch, the better the boiled and pounded yam,[11] while the gelatinization and retrogradation behavior of potato starch influences the textural changes occurring during thermal processing and cooking of products.[12]

The size distribution and shape of starch granules play a significant role in various functional properties such as viscosity, swelling power, retention capacity, mealiness, gelatinization and behavior during acid and enzymatic hydrolysis.[13-15] Starch granule size differences have been observed to impact the functional properties of sweet potato and banana.[16,17] According to Kouadio *et al*.,[18] mealy-cooking yam varieties typically have small starch granules (10–30 μm in diameter), while hard-cooking yam varieties have larger starch granules (35–40 μm in diameter). Additionally, Kang *et al*.[19] demonstrated that the shape of starch granules significantly affected the properties of rice flour and the quality of gluten-free bread made from it. The Seolgaeng rice variety, with its round starch structure, exhibits low granule hardness and damaged starch content, enabling the production of gluten-free rice bread with a well-built structure.[19] Therefore, to understand the relationship between starch biosynthesis, structure and properties, it is crucial to investigate the size and shape distributions. This is particularly important for nutritional and industrial applications of starch, as well as for understanding its digestibility and its association with nutritional disorders like diabetes and obesity.[20]

The most common method in the literature for characterizing starch granules of RTB products relies on microscopic photographs of flours or isolated starch colored with iodine. The characteristic parameters of granule size are then measured manually with software such as ImageJ and their shapes are estimated visually.[21] These measurements are tedious and only allow the analysis of a small number of granules while there is a large granule size and shape variability even within a given genotype.[22,23] Alternatively, the use of a laser light-scattering granulometer (LLSG) can hasten the characterization of starch granules.[24,25] This method is fast and accurate as it allows evaluation of more than 100 000 granules and measures those with a diameter as small as 0.1 μm.[26,27] However, it requires undamaged granules and is affected by the shape of the granules.[24] Non-spherical granules often have slightly smaller diameters than those obtained by image analysis techniques.[27] In addition, this method does not provide information on the shape of the granules apart from major and minor axis ratio.

Therefore, the objective of this study was to develop an open-source analysis pipeline that allows the automated and fast characterization of large number of yam (*Dioscorea alata* L.) starch granules providing size distribution and shape classification from microscopic imagery.

MATERIALS AND METHODS

Materials

Starch granules from 47 different yam (*Dioscorea alata* L.) genotypes produced in Guadeloupe in three locations (Duclos, Godet and Roujol) during 3 years (2016, 2017 and 2018) were used to develop the pipeline. After harvesting, two to six tubers of each genotype were peeled, washed and cut into 2 cm cubes. The cubes were then placed in an oven for a minimum of 48 h at a maximum temperature of 65 °C, allowing to preservation of the starch granule structure.[28] The dry material was ground to 0.25 mm with an SM100 knife grinder (Retsch GmbH, Haan, Germany). An optimal dilution suspension of 30 mg yam flour mixed in 1.5 mL of a 5% Dermal Betadine solution was found to maximize the number of starch granules observed for each picture, while avoiding granule overlaps (preliminary experiment, involving eight alternative dilution levels described in Supporting Information Appendix S1). A drop of the suspension was placed upon microscopic slide. The slides were then mounted on a light microscope (AxioScope A1, Carl Zeiss, Jena, Germany) for observation of the granules at a magnification of 100×. The microscope was equipped with a camera (AxioCam, ERc 5S, Carl Zeiss) connected to a computer for image acquisition. For each tuber, three to five microscopical images were taken. Details of images taken by year, site and genotype are given in Supporting Information Appendix S2.

Methods

Image segmentation and starch granule recognition

Image segmentation and starch granule recognition were done using R statistical language[29] with the *EBImage* library.[30] All image processing was done using an HP portable computer equipped with an Intel Core i7-6700HQ CPU 2.60 GHz with 16 GB RAM.

First, the raw image was converted to grayscale using the luminance weight.[31] Denoising was then applied in order to remove image speckle noise (i.e., isolated particles or small holes inside the granules) by operating a morphological closing.[32] After denoising, the image was converted into a binary image based on a gray threshold automatically detected by the Otsu segmentation method.[33]

Second, the algorithm looked for the region of interest (i.e., the starch granules). This step involved removing the background and separating overlapping starch granules. Granules were separated by calculating the distance map[34] of each pixel of the binary image. The distance map contained the distance to the nearest background pixel of each pixel. The granules were then separated using the watershed algorithm.[35,36] Once all starch granules were identified and isolated, a post-processing step was applied excluding objects touching the border of the image.

Supporting Information (Appendix S3) provides detailed steps of the pipeline and the corresponding R scripts.

Starch granule shape metrics

Prior to analysis, the images needed to be properly calibrated using an optical micrometer. A Thoma hemocytometer slide with 100 μm graduation was used to measure the field of view of the camera while keeping constant the microscope magnification and the camera zoom and resolution. This method enabled an accurate size calibration for each pixel. In this experiment, the image definition was 2560 × 1920 pixels and the pixel size corresponded to 0.44 μm.

After starch granule recognition, the functionalities of the *EBImage* package allowed us to extract some shape and size metrics:[37,38]

- the projected surface area (A_{proj}), which is the rectilinear projection of the starch granule onto the microscopic slide plane (μm^2);
- the estimated perimeter (P), which is the length of the boundary of the starch granule (μm);
- the minimum (r_{min}), maximum (r_{max}), average (r_{mean}) and standard deviation (r_{sd}) radius, which are respectively the minimum,

maximum, average and standard deviation of the distances between the granule center and the outer border (µm);
- the minor (θ_{min}) and major (θ_{max}) axis (µm) and the eccentricity of the ellipse, which is defined by $E = \sqrt{1 - \frac{\theta^2_{min}}{\theta^2_{max}}}$;
- the ratio (R_r) between the minimum and maximum radius: $R_r = \frac{r_{min}}{r_{max}}$;
- the ratio (R_θ) between the minor and major axis $R_\theta = \frac{\theta_{min}}{\theta_{max}}$;
- the shape factor 1 defined by $SF1 = \frac{\pi * \theta_{min} * \theta_{max}}{4A_{proj}}$;
- The shape factor 2 defined by $SF2 = \frac{\pi * \theta_{min} * \theta_{max}}{4} - A_{proj}$.

Shape classification model

The different classes of granule shape commonly observed in yam crops are ellipsoidal, polyhedral, round and triangular.[39-41] The shape of 512 starch granules of our dataset were classified accordingly. Shape classification examples are provided in Supporting Information (Appendix S4). The calibration and validation samples of the model contain data from 96 and 32 granules of each shape class, respectively. A random forest (RF) classification model was used to predict granule shape. The model was fit using the *randomForest* package,[42] with the 13 starch granule shape metrics defined previously as explanatory variables. In order to ensure model parsimony, the relative importance of each metric was calculated based on the mean decrease Gini.[43] The model was then simplified by discarding the least important variables as long as their removal did not impact model performance. The detailed procedure is provided in Supporting Information (Appendix S4). Finally, six parameters were kept (i.e., r_{sd}, E, R_r, R_θ, SF1 and SF2). The selected parameters then were used to calibrate the final model and test it using the independent validation dataset.

Pipeline performance

Granule size can be expressed as the average diameter, the average length of the major and minor axes, the mean maximum diameter, the mean granule volume or mean surface area.[24] In order to ensure accuracy of the starch granule size measurements provided by the pipeline, the A_{proj} of 100 starch granules were hand-measured using ImageJ[44] and compared to the estimations provided by the automated method.

To evaluate the possible effect of overlapping starch granules on size, we carried out robust analysis of variance (ANOVA) and a post hoc comparison test[45] with the R package *walrus*[46] using the median as estimator. The dependent variable was A_{proj} and the explanatory one was a categorical variable stating whether starch granules were isolated (i.e., not overlapping).

The performance of the shape classification model was evaluated on an independent validation set and quantified with the global accuracy and class-specific false positive (FP) and false negative (FN) rates.

Genotypic effect size and broad sense heritability

To assess the genotypic effect and the broad heritability of A_{proj}, we retained 20 common genotypes repeated over three years (2016, 2017 and 2018) at the Roujol site. The list of selected genotypes can be found in the Supporting Information (Appendix S5). Because A_{proj} distributions differ from the normal, the genotypic effect was tested using a robust ANOVA using the median as estimator.

The broad sense heritability of the median starch granule projected area (H^2_{Cullis}) was calculated using the formula of Cullis *et al.*,[47] as recommended by Covarrubias-Pazaran[48] using the *inti* package[49] with three years and three replicates (i.e., corresponding to the three tubers used per genotype):

$$H^2_{Cullis} = 1 - \frac{\overline{v}^{BLUP}_{\Delta}}{2\sigma^2_g} = 1 - \frac{\overline{PEV}}{\sigma^2_g}$$

where σ^2 refers to variance, g to genotype, $\overline{v}^{BLUP}_{\Delta}$ to the average BLUP (best linear unbiased predictor) difference or pairwise prediction error variance, and $\overline{PEV}$ refers to the average prediction error variance from genotypes.

Sample size

In order to guide future studies, two types of sample size were estimated. The first sample size allows the determination of the number n of starch granules that should be measured to have a statistically acceptable estimation of the median of A_{proj}. To achieve this, we calculated the median absolute deviation (MAD) of A_{proj} from every image. With 1214 samples (i.e., images), the central limit theorem[50] allowed us to approximate the MAD distribution to a normal distribution. Then, we determined n as follows:[51]

$$n = \frac{Z^2 \times \sigma^2}{d^2}$$

where Z is the value of standard normal deviation corresponding to the level of confidence. Here, we fixed the confidence level to 95%, which is equivalent to the value of Z equal to 1.96; d is the margin of error, fixed to 10% of the mean of MAD of A_{proj}; σ is the expected value of standard deviation, fixed here to be the mean of MAD of A_{proj}.

The second sample size corresponded to the minimum observations allowing to detect the mean difference between genotype A_{proj} median. In consequence, the effect size was calculated from the robust ANOVA by dividing the mean absolute value of the difference between the genotype medians by the pooled standard deviation.[52,53] This mean effect size was used to calculate the sample size (i.e., the number of granules per genotype) required to have a 95% chance of detecting a statistically significant difference at the 0.05 alpha level. Sample size was calculated using the R package *pwr*.[54]

RESULTS

Pipeline description and performances

Figure 1 illustrates the different steps in the analysis pipeline. The first step allows us to remove raw image noise through the closing morphology operation and to remove the image background (Fig. 1 (A–D)). The second step aims to recognize and isolate the starch granules using the watershed algorithm (Fig. 1 (E,F)). The third step removes granules located close to the border of the image (Fig. 1 (G,H)). Once isolated, the entire granules were characterized by their size and shape metrics. Thanks to the medium-throughput phenotyping pipeline developed, all of the 1214 previously acquired images were processed in less than 12 min, allowing the characterization of 205 256 starch granules.

The projected surface area of starch granules (A_{proj}) measured with ImageJ software was compared with those obtained with the developed automated pipeline. The coefficient of determination between the pipeline estimated and the ImageJ measured

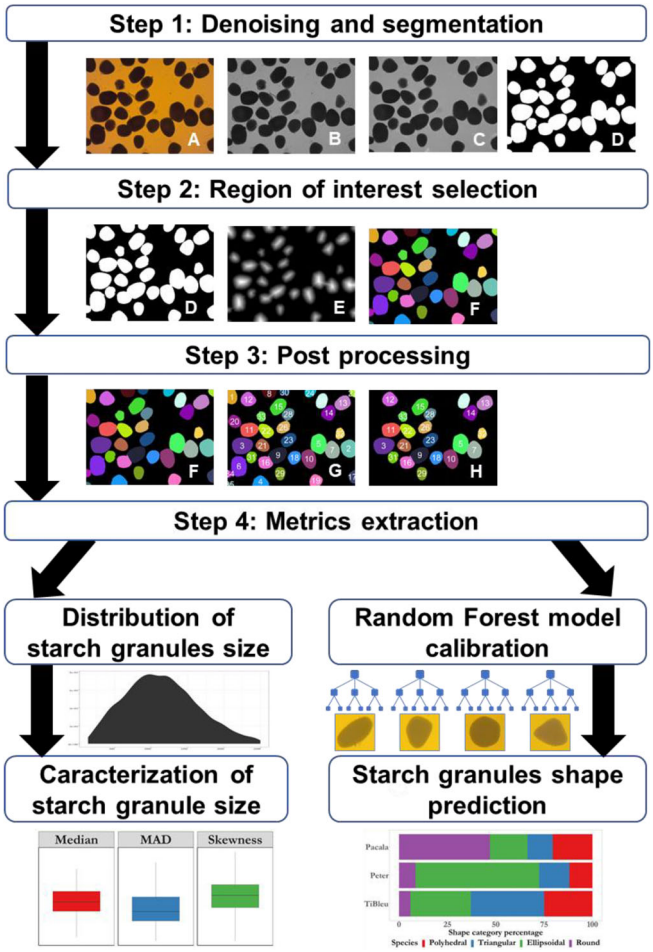

Figure 1. Proposed workflow for an automated image analysis pipeline to characterize the size and shape of yam starch granules.

starch granules size was 0.99 (Fig. 2). The root mean square error was 86.62 μm^2, demonstrating the low difference between the two methods.

The results of the robust ANOVA show a significant difference (P-value < 0.001) between the A_{proj} of isolated and overlapping granules. The median A_{proj} of isolated granules was, on average, 74.5 μm^2 smaller than that of overlapping granules.

Among the characteristic parameters of starch granule size and shape, the six most contributing to granule shape prediction were r_{sd}, E, R_r, R_θ, SF1 and SF2. Figure 3 presents the confusion matrix for the calibration and validation steps of the RF model. The model shows good accuracy of, respectively, 80% and 82.81% of correct class prediction. Specifically, the model correctly predicted 93.75% of ellipsoidal shapes, 62.50% of polyhedral shapes, 96.88% of round shapes and 75% of triangular shapes during the validation step.

Starch granule size distribution and sample size

The distribution of starch granule size of all genotypes is monomodal and positively skewed (Fig. 4). Indeed, the coefficient of asymmetry of the genotype size distribution varies between 0.32 and 1.74 (Fig. 5). Moreover, the distribution is heavy-tailed, exhibiting high median absolute deviation from 229 to 541 μm^2. The median of A_{proj} of the granules varies between 381 μm^2 (KL21) and 1115 μm^2 (Plimbite).

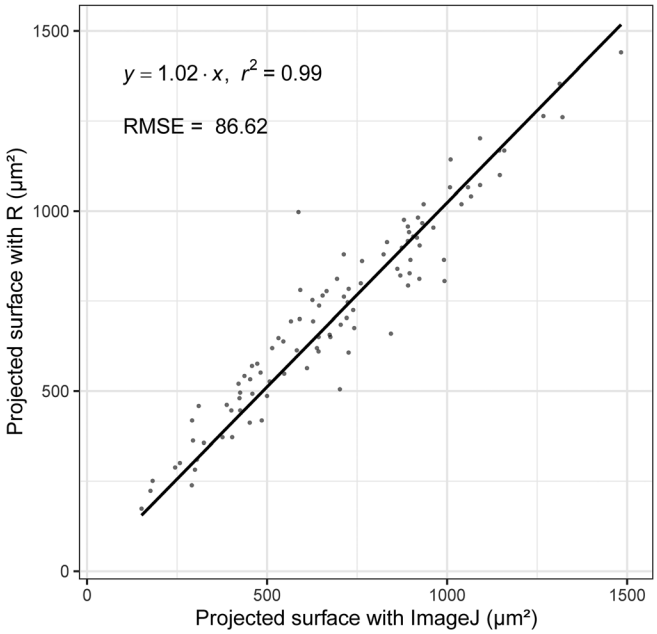

Figure 2. Regression analyses comparing the projected surface of starch granules estimated using the ImageJ reference method *versus* the proposed pipeline.

The results of the robust ANOVA showed a highly significant effect (P-value < 0.001) of the genotypes on A_{proj}. The mean difference between two genotypes was 190 μm^2. Finally, A_{proj} was found to be highly heritable, with a broad-sense heritability of 0.85 for the 20 genotypes over three years of cultivation (Supporting information Appendix S4). To have a 95% chance of detecting a difference corresponding to the mean genotypic difference with an alpha threshold of 0.05, it is necessary to observe at least 122 starch granules per genotype. As there are, on average, 165 starch granules per image, this sample size corresponds to one image acquisition. However, in order to correctly estimate the median value of the projected granule surface with less than 10% error in 95% of cases, the number of granules to be observed per experimental unit is 385, which corresponds to around two images.

Starch granule shape

The calibrated RF model was used to predict the shape of the 205 256 phenotyped granules. Figure 6 shows the proportion of each shape class by genotype. The main shape class varies between round (e.g., Pacala) and oval (e.g., Divin). Ellipsoidal shape proportion varied between 15.25% (Ptris) and 66.58% (Peter); polyhedral shape proportion varied between 11.16% (Peter) and 36.41% (H4X14M); round shape proportion per genotype varied between 3.53% (Divin) and 42.46% (Pacala); and triangular shape proportion varied between 7.77% (Roujol) and 43.68% (Tahiti).

DISCUSSION

The development of an open-source pipeline allowed us to segment an image, identify individual objects and extract the characteristic metrics of starch granules precisely in less than 1 s per image (0.53 s). Starch granule size seemed to vary widely across genotypes. Because all observed distribution of granules size were skewed, the median and median absolute deviation should

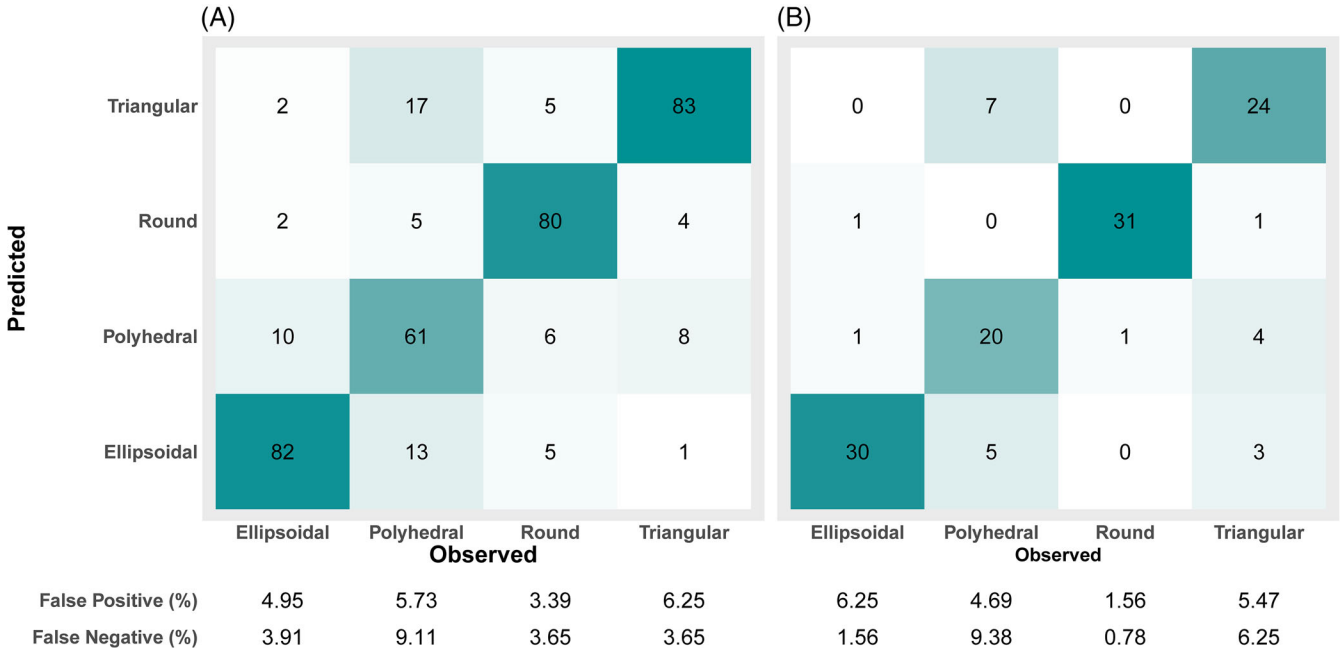

Figure 3. Confusion matrix of random forest model calibration (A) and validation (B).

be preferred over the mean and standard deviation to characterize the position and dispersion of starch granules from each genotype. Also, we will need a sample size of approximately 385 granules to estimate the median with a precision of 10% and a confidence level of 95%. With 165 granules per picture, this corresponds to approximately two pictures (Supporting Information, Appendix S5). In addition, one image acquisition is sufficient to capture the differences between the genotypes in the study.

The presence of overlapping granules suggests that there is a bias underestimating the size of these granules. In fact, attributing overlapping pixels to just one of the two granules reduces the median size estimate. However, the results of the robust ANOVA show the opposite; that is, the overlapping granules are significantly larger than the isolated granules. This could be explained by the fact that large granules have a greater tendency to aggregate than small ones. Under these conditions we recommend keeping the overlapping granules for size estimation.

Using low-angle LLSG, Farhat *et al.*[55] found similar sizes of *D. alata* starch granules (10–35 and 25–40 μm, for minor and major axis respectively) compared to this study (18–32 and 29–43 μm; Supporting Information, Appendix S6). Amani *et al.*[56] reported lower values (18.6–29.3 μm, for major axis). This discrepancy may be due to impurities, which can be considered as particles using low-angle LLSG.[55] Image analysis can visually verify that only starch granules are counted, whereas LLSG does not allow this distinction. This is why the literature refers to particle size rather than granule size.

According to Fauziah *et al.*,[9] who distinguish three types of starch depending on the size of the granules, every studied yam genotype has predominantly type A starch granules (Supporting Information, Appendix S6). This is in accordance with a previous study from Emiola and Delarosa[8] but in contradiction to Fauziah *et al.*,[9] who found type B granules for *D. alata*.

In previous studies, quantification of the shape of starch granules relies on only one parameter, called the form factor,[57,58] which equals 1 for a circle and tends to 0 for more irregular shapes. However, as Russ[59] explained, there is more than one way to deviate from the notion of 'circle', since a shape can be stretched into an ellipse but have an irregular edge or become angular like a polygon or star. The numerous metrics derived in the developed pipeline characterize the shape of starch granules better than the single form factor. As shown in Fig. 3, these parameters were able to properly discriminate the class of the starch granule shapes. This method thus offers the possibility to accurately and rapidly study the shape diversity between genotypes.

The shapes of the starch granules of the studied *D. alata* genotypes were ellipsoidal, polyhedral, round or triangular, as was the case in previous studies.[25,40] Because human perception is sometimes subjective and shape covers a continuous spectrum without a clear distinction between classes, we had difficulty classifying some granules into one of the four granule shape classes. These difficulties became more important with triangular and polyhedral shapes, explaining the higher false positive and false negative rates for these two classes (Fig. 3). However, these difficulties also demonstrate the advantage of a stable algorithm based on quantitative shape metrics compared to an often subjective visual classification.

The most dominant starch granule shapes of *D. alata* vary from study to study. According to Riley *et al.*,[40] they were the ellipsoidal and triangular shapes. For Fauziah *et al.*,[9] they were the triangular and polyhedral shapes. As for our study, the most dominant starch granule shapes are ellipsoidal and polyhedral. On the other hand, the proportion of shapes varies considerably from one genotype to another.

A satisfactory size determination technique should evaluate all granules, avoid missing small granules, prevent aggregation, distinguish between starch granules and non-starch particles, and include an intrinsic allowance for granule shape.[24,60] The results obtained in this study show that the proposed method for measuring physical parameters of starch granules is sufficiently

Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

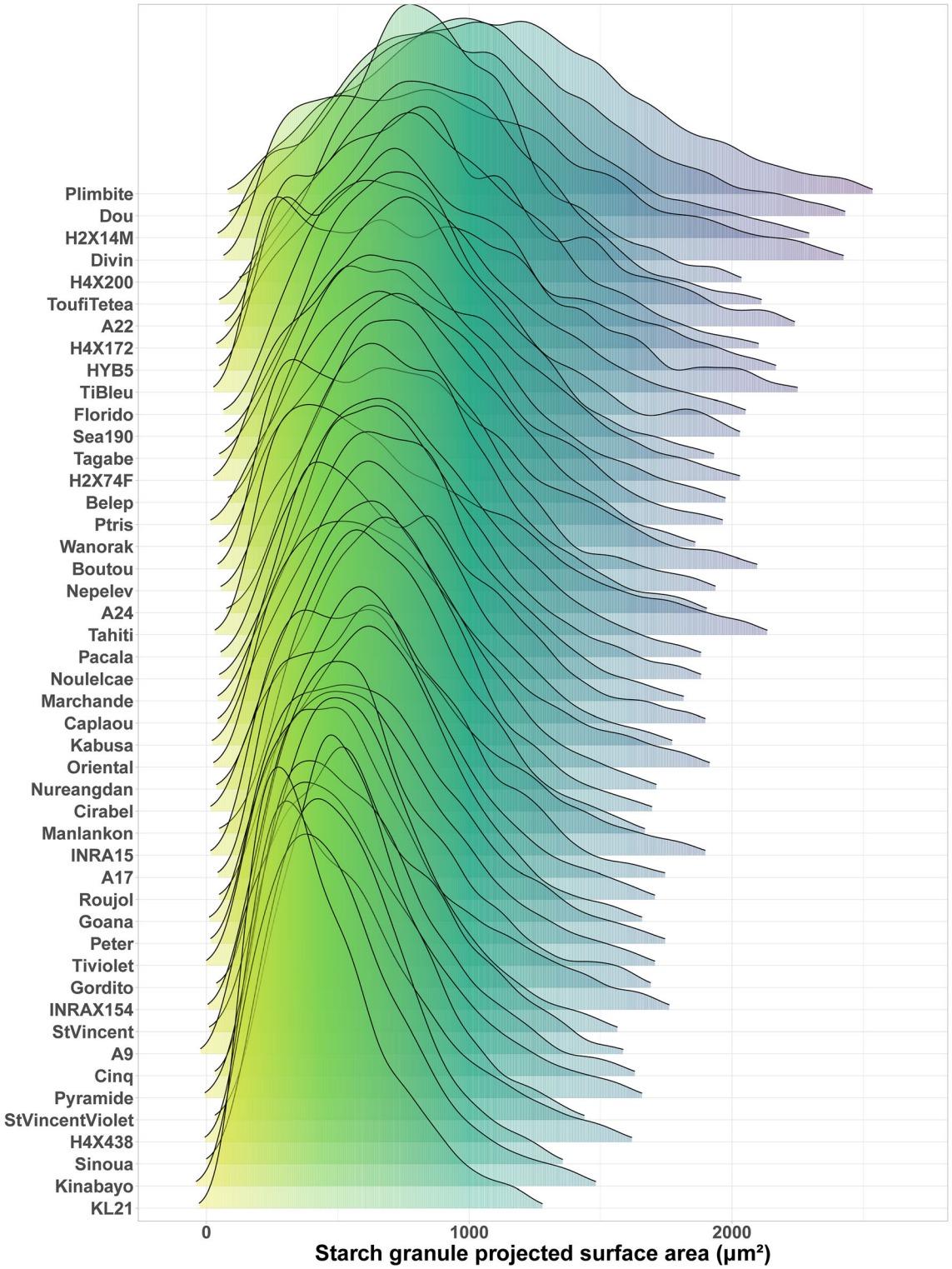

Figure 4. Starch granule size distribution of 47 selected yam (*Dioscorea alata* L.) genotypes.

accurate, robust and discriminating to be adopted in medium-throughput phenotyping studies. The application of this method reveals both similarities in monomodal distribution and differences in average size and shape proportions of starch granules present in yam flours, depending on genotype. The projected area of starch granules offers good prospects for effective selection, with a heritability of 0.85. These characteristics should now be linked to functional properties of the yam product. As suggested by Kang *et al.*[19] the gelatinization of round-shaped granules of rice occurs more rapidly than other shape classes and results in higher viscosity. Moreover, Kouadio *et al.*[18] proposed a possible link between starch granule size, proportion of

Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

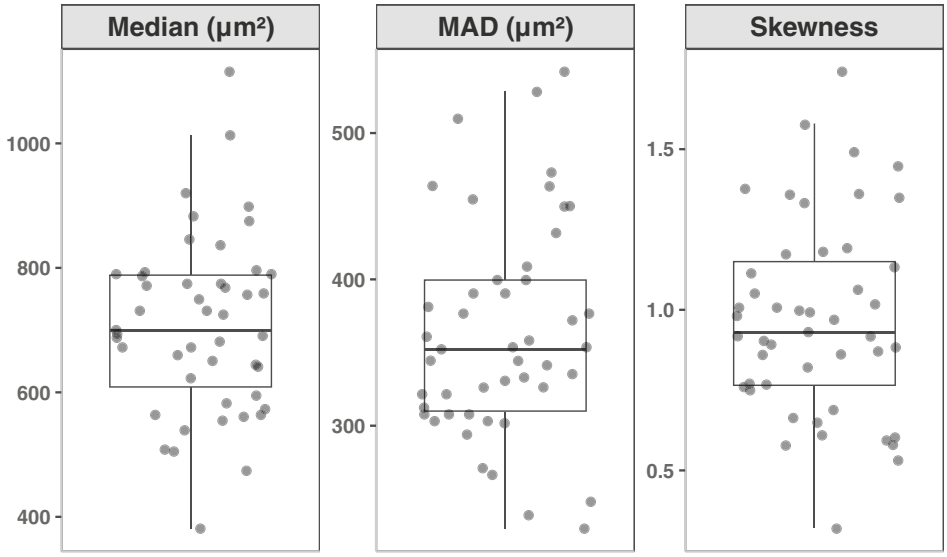

Figure 5. Variability of median, median absolute deviation (MAD) and skewness of the distribution of starch granule size of 47 yam (*Dioscorea alata* L.) genotypes.

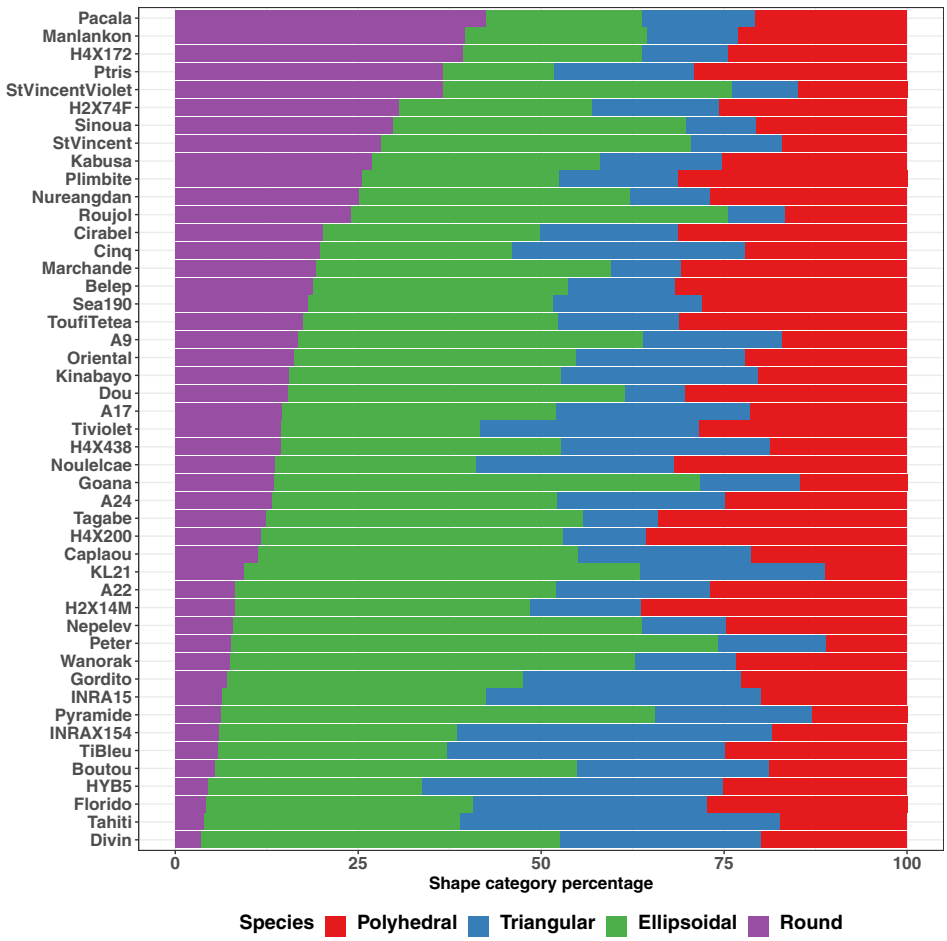

Figure 6. Proportion of granule shape class for 47 yam (*Dioscorea alata* L.) genotypes.

yam-resistant starch and product quality (i.e., mealiness). Further studies should be conducted to assess how the most contrasting genotypes in this study compare in starch functional properties (e.g., visco-amilograms) and, more importantly, cooking quality.

AUTHOR CONTRIBUTIONS

Conceptualization: CORNET Denis, DESFONTAINES Lucienne, DIBI Konan Evrard Brice. Data curation: HOUNGBO Mahugnon Ezékiel, IREP Jean-Luc. Formal analysis: HOUNGBO Mahugnon

Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

Ezékiel, CORNET Denis. Funding acquisition: CORNET Denis. Investigation: IREP Jean-Luc, HOUNGBO Mahugnon Ezékiel. Methodology: HOUNGBO Mahugnon Ezékiel, DESFONTAINES Lucienne, COUCHY Maritza, IREP Jean-Luc, CORNET Denis. Project administration: CORNET Denis. Resources: CORNET Denis. Supervision: CORNET Denis, OTEGBAYO Bolanle O. Writing ± original draft: HOUNGBO Mahugnon Ezékiel, CORNET Denis.

ACKNOWLEDGEMENTS

The authors are grateful to Pauline Mell, Elie Nudol (CIRAD), Christophe Perrot (CIRAD), David Lange (INRAE) and Jocelyne Leinster (INRAE) for field operation and sample preparation, and the grant opportunity ID OPP1178942 (Breeding RTB Products for End User Preferences, RTBfoods), to the French Agricultural Research Centre for International Development (CIRAD), Montpellier, France, by the Bill & Melinda Gates Foundation (BMGF): https://rtbfoods.cirad.fr.

CONFLICT OF INTEREST

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Kennedy G, Raneri JE, Stoian D, Attwood S, Burgos G, Ceballos H *et al.*, Roots, Tubers and Bananas contributions to Food Security Article. Ref Module Food Sci (2018).
2 Adifon FH, Yabi I, Vissoh P, Balogoun I, Dossou J and Saïdou A, Écologie, systèmes de culture et utilisations alimentaires des ignames en Afrique tropicale : synthèse bibliographique. *Cah Agric* **28**:22 (2019).
3 Honfozo L, Adinsi L, Bouniol A, Adetonah S, Forsythe L, Kleih U *et al.*, Boiled yam end-user preferences and implications for trait evaluation. *Int J Food Sci Technol* **56**:1447–1457 (2021).
4 Dufour D, Hershey C, Hamaker BR and Lorenzen J, Integrating end-user preferences into breeding programmes for roots, tubers and bananas. *Int J Food Sci Technol* **56**:1071–1075 (2021).
5 Hoover R, Composition, molecular structure, and physicochemical properties of tuber and root starches: a review. *Carbohydr Polym* **45**:253–267 (2001).
6 Hegenbart S, Understanding Starch Functionality Available online at: https://www.naturalproductsinsider.com/foods/understanding-starch-functionality [accessed June 9, 2023] (1996).
7 Ehounou AE, Cornet D, Desfontaines L *et al.*, Predicting quality, texture and chemical content of yam (*Dioscorea alata* L.) tubers using near infrared spectroscopy. *J Near Infrared Spectrosc* **29**:128–139 (2021).
8 Emiola LO and Delarosa LC, Physicochemical characteristics of yam starches. *J Food Biochem* **5**:115–130 (1981).
9 Fauziah F, Mas'udah S and Hendrian H, Study on the starch granules morphology of local varieties of *Dioscorea hispida* and *Dioscorea alata*. *J Trop Life Sci* **6**:6–52 (2016).
10 Padonou W, Mestres C and Nago MC, The quality of boiled cassava roots: instrumental characterization and relationship with physicochemical properties and sensorial properties. *Food Chem* **89**:261–270 (2005).
11 Otegbayo B, Aina J, Asiedu R and Bokanga M, Pasting characteristics of fresh yams (*Dioscorea spp.*) as indicators of textural quality in a major food product – 'pounded yam'. *Food Chem* **99**:663–669 (2006).
12 Alvarez M, Canet W and Tortosa M, Kinetics of thermal softening of potato tissue (cv. Monalisa) by water heating. *Eur Food Res Technol* **212**:588–596 (2001).
13 Bhat FM and Riar CS, Effect of amylose, particle size & morphology on the functionality of starches of traditional rice cultivars. *Int J Biol Macromol* **92**:637–644 (2016).
14 Cornejo-Ramírez YI, Martínez-Cruz O, Del Toro-Sánchez CL, Wong-Corral FJ, Borboa-Flores J and Cinco-Moroyoqui FJ, The structural characteristics of starches and their functional properties. *CyTA-J Food* **16**:1003–1017 (2018).
15 Singh N and Kaur L, Morphological, thermal, rheological and retrogradation properties of potato starch fractions varying in granule size. *J Sci Food Agric* **84**:1241–1252 (2004).
16 Otegbayo B, Lana O and Ibitoye W, Isolation and physicochemical characterization of starches isolated from plantain (musa Paradisiaca) and cooking banana (musa Sapientum). *J Food Biochem* **34**:1303–1318 (2010).
17 Singh N, Chawla D and Singh J, Influence of acetic anhydride on physicochemical, morphological and thermal properties of corn and potato starch. *Food Chem* **86**:601–608 (2004).
18 Kouadio OK, N'dri DY, Nindjin C, Marti A, Casiraghi MC, Faoro F *et al.*, Effect of resistant starch on the cooking quality of yam (*Dioscorea* spp.) and cassava (*Manihot esculenta*) based paste products. *Int J Food Sci Nutr* **64**:484–493 (2013).
19 Kang T-Y, Sohn KH, Yoon M-R, Lee J-S and Ko S, Effect of the shape of rice starch granules on flour characteristics and gluten-free bread quality. *Int J Food Sci Technol* **50**:1743–1749 (2015).
20 Cave RA, Seabrook SA, Gidley MJ and Gilbert RG, Characterization of starch by size-exclusion chromatography: the limitations imposed by shear scission. *Biomacromolecules* **10**:2245–2253 (2009).
21 Wan ZW, Lin SP, Ju M, Ling CH, Jia YL, Jiang MX *et al.*, Morphotypological analysis of starch granules through discriminant method and its application in plant archeological samples. *Appl Ecol Environ Res* **18**:4595–4608 (2020).
22 Guo S, Automatic segmentation on multiple starch granules from microscopic images. *Microsc Res Tech* **75**:524–530 (2012).
23 Tong C-S, Choy S-K, Chiu S-N, Zhao Z-Z and Liang Z-T, Characterization of shapes for use in classification of starch grains images. *Microsc Res Tech* **71**:651–658 (2008).
24 Lindeboom N, Chang PR and Tyler RT, Analytical, biochemical and physicochemical aspects of starch granule size, with emphasis on small granule starches: a review. *Starch-Stärke* **56**:89–99 (2004).
25 Otegbayo B, Oguniyan D and Akinwumi O, Physicochemical and functional characterization of yam starch for potential industrial applications. *Starch-Stärke* **66**:235–250 (2014).
26 Chmelík J, Krumlová A, Budinská M, Kruml T, Psota V, Bohačenko I *et al.*, Comparison of size characterization of barley starch granules determined by electron and optical microscopy, Low angle laser light scattering and gravitational field-flow fractionation. *J Inst Brewing* **107**:11–17 (2001).
27 Raeker MÖ, Gaines CS, Finney PL and Donelson T, Granule size distribution and chemical composition of starches from 12 soft wheat cultivars. *Cereal Chem* **75**:721–728 (1998).
28 Hoover R and Vasanthan T, Effect of heat-moisture treatment on the structure and physicochemical properties of cereal, legume, and tuber starches. *Carbohydr Res* **252**:33–53 (1994).
29 R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ (2022).
30 Pau G, Fuchs F, Sklyar O, Boutros M and Huber W, EBImage-an R package for image processing with applications to cellular phenotypes. *Bioinformatics* **26**:979–981 (2010).
31 Smith T and Guild J, The C.I.E. colorimetric standards and their use. *Trans Opt Soc* **33**:73–134 (1931).
32 Haralick RM, Sternberg SR and Zhuang X, Image analysis using mathematical morphology. *IEEE Trans Pattern Anal Mach Intell* **9**:532–550 (1987).
33 Otsu N, A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* **9**:5–66 (1979).
34 Kolountzakis MN and Kutulakos KN, Fast computation of the Euclidian distance maps for binary images. *Inf Process Lett* **43**:181–184 (1992).
35 Kornilov AS and Safonov IV, An overview of watershed algorithm implementations in open source libraries. *J Imaging* **4**:123 (2018).
36 Serra J, *Image Analysis and Mathematical Morphology: Vol.: 2: Theoretical Advances*, 2nd edn. Academic Press, London, United Kingdom (1988).

37 Boland MV and Murphy RF, A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. *Bioinformatics* **17**:1213–1223 (2001).
38 Rodenacker K and Bengtsson E, A feature set for cytometry on digitized microscopic images. *Anal Cell Pathol* **25**:1–36 (2003).
39 Kringel DH, El Halal SLM, da Rosa Zavareze E and Dias ARG, Methods for the extraction of roots, tubers, pulses, Pseudocereals, and other unconventional starches sources: a review. *Starch-Stärke* **72**: 1900234 (2020).
40 Riley CK, Wheatley AO and Asemota HN, Isolation and characterization of starches from eight *Dioscorea alata* cultivars grown in Jamaica. *Afr J Biotechnol* **5**:1528–1536 (2006).
41 Sukhija S, Singh S and Riar CS, Isolation of starches from different tubers and study of their physicochemical, thermal, rheological and morphological characteristics. *Starch-Stärke* **68**:160–168 (2016).
42 Liaw A and Wiener M, randomForest: Breiman and Cutler's random forests for classification and regression. *R Package Version* **4.7–1.1** https://CRAN.R-project.org/package=randomForest (2022).
43 Akhmad A, Lukas L and Mahawan B, Improving performance loan fraud model prediction using mean decrease accuracy and mean decrease Gini. *ACMIT Proc* **6**:36–41 (2019).
44 Rasband WS, *ImageJ*. U.S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/, pp. 1997–2018 (2018).
45 Wilcox RR, *Introduction to Robust Estimation and Hypothesis Testing*. Academic Press, Southern California, USA (2011).
46 Love J and Mair P, *walrus*: Robust Statistical Methods. *R Package Version* **1.0.5** https://CRAN.R-project.org/package=walrus (2022).
47 Cullis BR, Smith AB and Coombes NE, On the design of early generation variety trials with correlated data. *J Agric Biol Environ Stat* **11**:381–393 (2006).
48 Covarrubias-Pazaran GE, Heritability: meaning and computation Online Manual. Available online at: https://excellenceinbreeding.org/sites/default/files/manual/Heritability_v6.pdf, [accessed June 9, 2023] (2021).
49 Lozano-Isla F, *inti*: Tools and Statistical Procedures in Plant Science. *R Package Version* **0.6.1** https://CRAN.R-project.org/package=inti (2023).
50 Kwak SG and Kim JH, Central limit theorem: the cornerstone of modern statistics. *Korean J Anesthesiol* **70**:144–156 (2017).
51 Sapra RL, How to calculate an adequate sample size? in *How to Practice Academic Medicine and Publish from Developing Countries?*, ed. by Nundy S, Kakar A and Bhutta ZA. Springer, Singapore (2022).
52 Cohen J, *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. L. *Erlbaum Associates*, Hillsdale, N.J., USA (1988).
53 Nakagawa S and Cuthill IC, Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev* **82**:591–605 (2007).
54 Champely S, Ekstrom C, Dalgaard P, Gill J, Weibelzahl S, Anandkumar A *et al.*, Package '*pwr*' *R Package Version* 1.3.0 https://CRAN.R-project.org/package=pwr (2020).
55 Farhat IA, Oguntona T and Neale RJ, Characterisation of starches from West African yams. *J Sci Food Agric* **79**:2105–2112 (1999).
56 Amani NG, Buléon A, Kamenan A and Colonna P, Variability in starch physicochemical and functional properties of yam (Dioscorea sp) cultivated in Ivory Coast. *J Sci Food Agric* **84**:2085–2096 (2004).
57 Devaux MF, Qannari EM and Gallant DJ, Multiple-correspondence analysis optical microscopy for determination of starch granules. *J Chemometr* **6**:163–175 (1992).
58 Guo S, Tang J, Deng Y and Xia Q, An improved approach for the segmentation of starch granules in microscopic images. *BMC Genomics* **11**:S13 (2010).
59 Russ JC, Image analysis of foods. *J Food Sci* **80**:E1974–E1987 (2015).
60 Rasper V, Investigations on starches from major starch crops grown in Ghana: III.-particle size and particle size distribution. *J Sci Food Agric* **22**:572–580 (1971).

Journal of The Science of Food and Agriculture published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.