Received: 7 March 2023 | Accepted: 4 September 2023

DOI: 10.1002/aps3.11552

APPLICATION ARTICLE

Applications in Plant Sciences

A standardized and efficient technique to estimate seed traits in plants with numerous small propagules

Christina Steinecke[1] | **Jeremiah Lee**[1,2] | **Jannice Friedman**[1]

[1]Biology Department, Queen's University, Kingston, Ontario K7L 3N6, Canada

[2]Department of Geography and Planning, Queen's University, Kingston, Ontario K7L 3N6, Canada

Correspondence

Christina Steinecke, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts 02138, USA.

Email: christinasteinecke@g.harvard.edu

Present address

Christina Steinecke, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA; Arnold Arboretum of Harvard University, Boston, Massachusetts 02131, USA.

This article is part of the special issue "Advances in Plant Imaging across Scales."

Abstract

Premise: Variation in seed traits is common within and among populations of plant species and often has ecological and evolutionary implications. However, due to the time-consuming nature of manual seed measurements and the level of variability in imaging techniques, quantifying and interpreting the extent of seed variation can be challenging.

Methods: We developed a standardized high-throughput technique to measure seed number, as well as individual seed area and color, using a derived empirical scale to constrain area in *Arabidopsis thaliana, Brassica rapa*, and *Mimulus guttatus*. We develop a specific rational model using seed area measured at various spatial scales relative to the pixel count, observing the asymptotic value of the seed area as the modeled number of pixels approaches infinity.

Results: We found that our model has high reliability in estimating seed traits and efficiently processes large numbers of images, facilitating the quantification of seed traits in studies with large sample sizes.

Discussion: This technique facilitates consistency between imaging sessions and standardizes the measurement of seed traits. These novel advances allow researchers to directly and reliably measure seed traits, which will enable tests of the ecological and evolutionary causes of their variation.

KEYWORDS

image processing, ImageJ, propagules, seed size, seed traits, spatial scale

In plants, seed number and seed size have important ecological and evolutionary implications. The reproductive fitness of an individual is commonly measured as the number of seeds produced and is therefore of fundamental importance to evolutionary questions. Additionally, seed size can affect fitness, through its impact on dispersal and predation (reviewed in Westoby et al., 1996; Eriksson, 2008; Bogdziewicz et al., 2019), as well as germination, establishment, and survival of the seedling (Krannitz et al., 1991; Eriksson, 1999; Elliott et al., 2007; Martínez-González et al., 2021).

Seed size has been studied extensively, with early emphasis on the extraordinary differences between species, with up to 10^6-fold variation within a region (Baker, 1972; Westoby et al., 1992; Leishman et al., 1995; Moles et al., 2005). Early research emphasized the relative constancy of seed size within species (Harper et al., 1970), but since then it has been well established that substantial variation exists within and among individuals (Janzen, 1977; Thompson, 1984; reviewed in Michaels et al., 1988; Eriksson, 1999; Gnan et al., 2014; Paczesniak et al., 2022). For example, seed mass varies positively with plant size (Hendrix, 1984; Aarssen and Jordan, 2001), with fruit maturation order (Fuller et al., 1983; Cavers and Steel, 1984; Hendrix, 1984; Torres et al., 2002; de Carvalho et al., 2021), with position within the ovary (Greenway and Harder, 2007), and with paternity (Stanton, 1984; Mazer et al., 1986; Andersson, 1990; Raunsgard et al., 2018). Seed size is thought to evolve as a compromise between producing numerous small seeds containing few resources and producing fewer large seeds

containing more resources (Westoby et al., 1992; Leishman and Murray 2001; Gnan et al., 2014). The variation in seed size and number within species can have important ecological and evolutionary consequences; however, quantifying this variation requires accurate and robust measurement of large sample sizes.

The importance of seed traits and standardized, reproducible measurements was emphasized by Saatkamp et al. (2019) in their call for a seed-trait functional ecology program—a research agenda to characterize seed-trait variation that is connected to plant functions and ecological strategies. Such a research agenda requires the development and maintenance of databases and compilation of standardized and useful seed traits at the global scale (Moles et al., 2007). The seed traits that feed into the axes of the seed ecological spectrum include morphological traits (e.g., size, number, shape, color), chemical traits (e.g., toxicity, nutrients), and physiological traits (e.g., light and moisture requirement to break dormancy) (Saatkamp et al., 2019). While seed morphological traits such as size, shape, and color are only one dimension along which seeds can vary, they are arguably the easiest to measure. Despite the relative simplicity of these traits, methods to quantify them are quite limited. Mussadiq et al. (2015) compared the accuracy of various programs (e.g., ImageJ, CellProfiler, P-TRAP, and Smartgrain) to estimate seed number, and found that custom ImageJ macros produced the best estimation of seed count. Their study does not estimate other traits, such as seed size, or include model-based approaches to allow for universal application. The use of high-throughput approaches to characterize seed traits has broad applicability to many agricultural, ecological, and evolutionary studies, particularly those requiring large sample sizes and/or studies of small-seeded species.

The extensive variation in seed morphological traits has ecological and evolutionary significance; however, their characterization is limited by the lack of standardized methods for quantifying seed trait variation. Here, we develop a high-throughput method that has the capacity to measure seed number, size, shape, and color using digital images and theoretical modeling of optimal parameters. Below we describe the protocol for measuring particle number and size; the procedure has the capacity to additionally estimate traits such as shape and relative color (not shown). We use seeds from three model species (*Arabidposis thaliana* (L.) Heynh. [Brassicaceae], *Brassica rapa* L. [Brassicaceae], and *Mimulus guttatus* DC. [synonym *Erythranthe guttata* (DC.) G. L. Nesom; Phyrmaceae]) to determine the limit in pixel count for reliable measurements and provide context for establishing camera configurations for bulk imaging of seeds (or particles more generally). Importantly, we assess the influence of camera resolution at a millimetric scale on accuracy and robustness of measurement, as well as provide a procedure to perform high-throughput particle size and number measurements using ImageJ and R software tools.

METHODS

Background on identifying objects in digital images

Digital photographs are typically stored as raster data, comprising a two-dimensional grid (x, y) of square pixels, with red, green, and blue (RGB) values for each pixel for color images and a single value for each pixel for grayscale images. The term "resolution" has several meanings, and can refer to the number of pixels making up a given image, the dimension of individual pixels in real-world units (e.g., millimeters, centimeters) within an image, or the scale at which objects of a given size can be resolved in image space. We use the last meaning. To identify objects (such as seeds) in an image, a surface made up of pixels that are distinct from their surroundings is manually or algorithmically delineated. This could be done by relying on differences in pixel value or the change in pixel values as a function of distance within an image. The ability to determine the precise border between an object and the surrounding space is limited by the resolution of the image.

The perimeter of an object identified in an image includes a jagged line of pixels defining the edge, which is not linear in the x or y dimension (Figure 1). The delineation of object borders is imperfect and granular and can impact measurements. The borders produce error in misidentifying pixels associated with the object and in the portion of the pixel covering non-object space alongside actual object surface area. The limits of the square pixel geometry become proportionally more important as the size of the object becomes smaller and the ratio of edge pixels becomes larger. Another source of error is introduced in converting pixels to real-world units of measurement, which requires a pixel-to-area conversion scale. Any error associated with the conversion scale will propagate. To reliably use digital images for object measurement, one needs to account for these sources of error. Resolution and spatial scale are well-studied in geographic information science (reviewed in Atkinson and Tate, 2000), and resolution influences the ability to measure features on a landscape scale such as vegetation changes (e.g., Wu et al., 2002; Dunwoody et al., 2013). Bringing image analysis to propagule measurement enables confident use of high-throughput image processing methods and facilitates measurement of propagules in large quantities or of small sizes.

Propagule imaging

The optimal imaging configuration includes bright, diffuse light from multiple angles to minimize shadows of the imaged object. Later processing can compensate for minor shadowing; however, consistency within and between imaging sessions is vital for meaningful comparison of propagule sizes. We arranged seed propagules on a sheet of white paper, with a camera mounted directly above. We

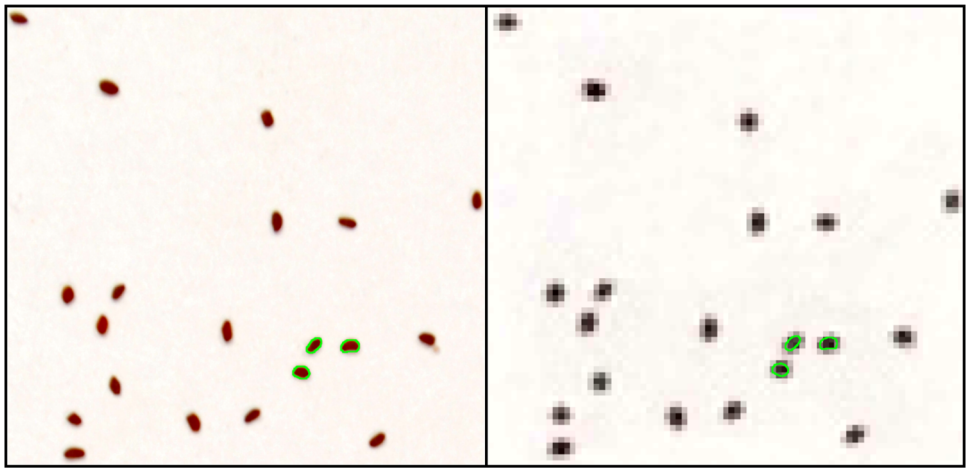

FIGURE 1 A sample of *Mimulus guttatus* seeds imaged using a Nikon D3100 digital SLR camera at a camera height of 30 cm and focal length of 55 mm (H30-FL55; left) and at a camera height of 69.7 cm and a focal length of 18 mm (H69.7-FL18; right). The green polygons indicate the isolated seed perimeter. At the H30-FL55 configuration, seeds consisted of an average of 600 px, while at the H69.7-FL18 configuration seeds were identified with lower resolution at approximately 10 px, introducing greater error in delineating the boundary.

checked that the seeds were within the field of view of the camera, positioned with no overlap. For reproducible imaging, we marked the paper to record the area occupied by the propagules within the field of view, along with the position of the background sheet on the mounting table.

To image seeds, we used manual settings on a digital single-lens reflex (SLR) camera (see Table 1 for camera components and parameters). We set the lens of the camera to either the maximum or minimum magnification, as those limits are reliably reproducible by reaching the physical limit of the lens adjustment. We used an f-stop and shutter speed appropriate for the room lighting, allowing for a sharp and well-lit image within the field of view of the propagules. We used a low ISO as the scene had sufficient illumination to clearly distinguish between the propagules and the white background. The images are captured in a lossless imaging format (e.g., .tif format); if a lossless format is not available, the highest quality, least compressed format should be used, as information will be lost with compressed filetypes such as .jpg files.

As a first step, we recommend users identify the camera configuration that results in a single propagule in the image being ≥100 px. This threshold ensures that the scale at which the seeds are resolved can be confidently converted to a physical area, as described below in the Validation section of the Methods. To achieve this, we used a camera height of 30 cm and focal length of 55 mm. With the established camera configuration, we took images with a precise scalebar to produce an empirical conversion ratio between image pixel size and physical size. We used a 12-inch ruler (product number 501-012; Products Engineering Corporation, Torrance, California, USA) placed in the center of the field of view and imaged four times, rotating the ruler by 45° after each image to capture length and diagonal scales. Deriving the empirical conversion ratio from the average of four ruler angles ensures that differences in distortion in *x*- and *y*-directions are minimized, to reduce any disparity in converted area for oblong particles. We do not account for distortions associated with un-orthorectified images.

TABLE 1 Camera configuration and parameters used for capturing seed images.

Camera components and parameters	Model or parameter value
Camera model	Nikon D3100
Sensor size (mm)	23.1 × 15.4
Sensor resolution (px)	4608 × 2304
Lens model	Nikon AF-S DX NIKKOR 18–55 mm f/3.5-5.6 G VR
Shutter speed (s)	1/500
f-stop	f/8
ISO	100
Focal lengths used (mm)	18, 55
Camera model	Leica Flexacam C1
Sensor size (mm)	6.25 × 4.69
Sensor resolution (px)	4000 × 3000
Microscope model	Leica S8APO
Shutter speed (s)	1/12
f-stop	f/4.9
ISO	500

Next, we imaged the propagules. We placed the propagules from a given sample within the designated frame, ensuring no overlap between propagules and the edge of the frame, and removing any detritus of a similar size to the seeds. The images of the seeds and the images of

the ruler were transferred into separate folders on the processing computer.

Image processing in ImageJ and R

Step-by-step instructions for using the ImageJ and R processing files are available in Appendix S1 and summarized here. Within ImageJ (Rasband, 2018), we first opened each image of the ruler and drew a line using the *Straight* tool across the graduation marks of a known distance, d_r. Using the *Set scale…* menu from the *Analyze* drop-down menu, the distance in pixels, d_{px}, was obtained for each ruler image. We calculated the image area conversion scale E according to the equation:

$$E = \left(\frac{d_r}{d_{px}}\right)^2$$

where the E scalar converts the count of square pixels to physical area. The maximum percent difference between the calculated mm^2/px scalars between each image was 1.7%.

Using the *Rectangle* tool, we created a cropping region of interest (ROI) template that contained the propagules in each image. Because we initially standardized the base sheet position on the imaging table and the propagules position within the drawn frame, the ROI template was the same for all images.

We developed an expanded macro (particleSizeID.txt, Appendix S2) from the batch processing instructions developed by Herbert (2011) to streamline image processing. Users select the folder containing the propagule images and provide the case-sensitive file extensions to ensure the appropriate images are analyzed. The ImageJ macro file associated with this procedure (Appendix S2) is opened within the macro window, and the destination folder for the results, in .csv format, is selected.

In processing the images, there are several settings that may be useful in some circumstances. The *Subtract Background* setting determines whether the objects in the image are emphasized from a background with inconsistent color intensity. This setting is helpful for obtaining propagule counts from noisy images when the propagules are less than 50 px in radius, but ideally is not used for accurate area measures as it alters pixel intensities around the border of objects. *Subtract Background* was not used here. The *Watershed concavity segmenting* setting is helpful for obtaining accurate propagule counts at the cost of accuracy in area measurements. The setting segments objects identified in an image at points of concavity, such as the boundary between two roughly circular seeds that are in contact with one another. Because shape does not influence object identification in this macro, the watershed segmenting tool enables the isolation of individual objects from groups that are in contact with one another. However, the boundary derived from this tool does not consistently match the exact propagule edge and so should only be used to improve counts with overlapping image features. The propagules here were not in contact or overlapping and thus the watershed segmentation was not used. Next, the thresholding method is chosen using either the default ImageJ setting or by manually selecting a value between 0–255 to isolate the propagules from the background. With a controlled imaging setup and consistent lighting, the default provided efficient identification using a modified version of the IsoData method described by Ridler and Calvard (1978).

The macro operates by cycling through the input image folder, performing a series of processing steps before saving the propagule measurement in a .csv file. First, each image is cropped to the ROI file boundaries. It then converts the three-channel RGB image to an 8-bit grayscale image using the mean of the three-color values of each pixel. Next, the image is converted from grayscale to a binary image, isolating the propagule features from the background as a mask image. If the watershed segmentation tool is used, it splits concave features in the binary mask into multiple features. Otherwise, each isolated feature in the mask is used to extract the number of pixels, the minimum and maximum caliper distance in pixels, and minimum, maximum, and mean RGB values of each propagule in the original input image, saved to the designated output folder as a .csv file.

The protocol uses the R programming language (R Core Team, 2021) to process the ImageJ results. The particleSizeProcess.R script (Appendix S3) takes the measurement output from ImageJ and performs an initial filter to exclude seed pixel values outside of a given input range; this is to immediately exclude any obvious non-seed particles, such as plant debris or soil. The script then performs additional statistical filtering using the area calculated from each seed (SD filter), excluding values outside an input number of standard deviations above and below the mean. This SD filter can be excluded by setting a null value, or performed on $\log_{10}$-transformed data to manage right-skewness. The output is a list class object with two entries. The first entry is a data frame with a row for each particle identified in the ImageJ macro, containing its source file name, number of pixels counted by ImageJ, converted area calculated by the input scale, the $\log_{10}$-transformed area, the minimum and maximum caliper distance, and the mean, minimum, and maximum of each red, green, and blue channel as captured by the camera sensor. The second list entry contains the same information but is limited to the particles falling within the SD filter included in the function call. If no size filters are applied, the output is the unfiltered data frame containing the above-mentioned data for each particle. With a controlled imaging process, the colors can be used for comparison between samples, although color standardization would be necessary to compare to other propagule data sets.

Validation using a test sample

To assess the accuracy of the protocol for measuring particle area, we took images at several camera configurations for

seeds from individuals of *M. guttatus* (for raw data of the seed measurements, see Appendix S4). For a given sample, the seeds were distributed within the frame of the Nikon D3100 camera (Nikon, Tokyo, Japan) and imaged according to the camera configurations in Table 2 (camera height and focal length) with the mean empirical scale calculated from multiple measurements of the imaged ruler. The minimum and maximum focal lengths of 18 and 55 mm were used as they are consistently attained as the limits of the lens mechanism. The magnification indicated in Table 2 indicates the approximate alteration of the size of the imaged object given the relation:

$$M = \frac{h_i}{h_o},$$

where M is magnification, h_i is the height of the image, and h_o is the height of the object. The vertical sensor dimension (found in the camera documentation) is used for image height, and the total length captured by the camera field of view in the corresponding dimension is used for object height, calculated by multiplying the empirical scale by the number of pixels in that dimension. Camera height above the propagules and focal length of the zoom lens are used to control the scale of the photographed propagules in image space.

Seed size robustness test

To determine the robustness of our protocol and whether imaging height had consistent effects, 10 seeds each of *A. thaliana*, *B. rapa*, and *M. guttatus* were tracked using multiple camera configurations, including those used for the test sample above, and different combinations of camera heights and focal lengths on the Nikon D3100 camera and a Leica S8APO dissection microscope (Leica, Wetzlar, Germany) with a Leica Flexacam C1 (Tables 1 and 2). The 10 seeds for each species were fixed to parafilm on a microscope slide to prevent any displacement that may affect which face of a seed was captured, ensuring the identical layout was captured at each configuration.

RESULTS

Validation using a test sample

To calculate seed area (mm^2) from the images, we used the empirical conversion scale outlined above. We applied an initial filter to remove outliers (e.g., soil or organic debris) that were identified as particles two orders of magnitude below the median seed size. We did not apply an SD filter to the data. The distributions of seed area are shown in Appendix S5; we applied a $\log_{10}$-transformation to account for the right-skewness of the data (at all camera configurations; Figure 2).

Increasing the number of pixels defining a particle may influence the measurement of seed area. To test this, we applied an inverse model to the $\log_{10}$-transformed seed area as a function of the number of pixels identified for each seed, using an iterative least-squares method:

$$f(x) = \frac{a}{x - b} + c$$

Where x is the seed area in number of pixels and $f(x)$ is the $\log_{10}$-transformed seed area. This formula best fit the transformed mean seed area represented as a function of pixel count at each configuration and aligned with the expectation that misidentified edge pixels would be of decreasing importance at higher pixel counts. Variable a represents the scalar of the equation, and b is the vertical asymptote. As the number of pixels decreases, the metric size of the particle increases due to edge pixel effects. The horizontal asymptote of the model, c, represents unlimited resolution and the area value that would not include any edge pixel misidentifications, and thus is a theoretical representation of the true area that could be derived from the image. In this case, the area described by the asymptotic value, back-transformed to metric space, was 0.143 mm^2 with a 95% confidence interval (CI) of 0.141–0.145 mm^2. This represents the best estimate of the true mean size of the sample.

The camera configuration with the greatest resolution was H30-FL55 (where H is the sensor height above the seeds [in centimeters] and FL is the focal length [in millimeters]). This camera configuration had a sample mean seed area of 0.145 mm^2 (95% CI: 0.141–0.149; Table 3, Figure 3). We compared all camera configurations using a linear model on the $\log_{10}$-transformed seed area, and none were significantly different from one another, except those involving H69.7-FL18 (the lowest magnification), which was significantly different from all other configurations.

Seed size robustness test

To assess the robustness of our protocol across multiple individual measurements, we tracked the area of 10 seeds at all camera configurations. We used both a Leica Flexacam C1 camera affixed to a Leica S8APO dissection microscope to capture each of the seeds at greater magnifications than possible with the Nikon camera, resulting in images composed of orders of magnitude more pixels. We delineated the area of each individual seed in ImageJ and scaled the sizes according to the empirical conversion value derived in the method above. We used the seed area attained from the microscope images as the reference value against which the camera images would be compared.

We converted each seed pixel count to a metric area using conversion ratios for the appropriate image configuration. Seed areas are shown in Figure 4, with the robustness evident in the relative size and rank order of the replicate

TABLE 2 Camera configurations used to image *Arabidopsis*, *Brassica*, and *Mimulus* seeds with the empirical scale (mm/px) derived from manual scale bar measurements and magnification calculated from the sensor sizes.

Configuration name	Camera components	Camera height above seeds (cm)	Focal length (mm)	Empirical scale (mm/px)	Approximate magnification	Relevant test
Scope (MA)	Flexacam C1, dissection microscope (*M. guttatus*, *A. thaliana*)	—	—	0.000305	5.12	Robustness test
Scope (B)	Flexacam C1, dissection microscope (*B. rapa*)	—	—	0.00252	0.621	Robustness test
H28-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	28	55	0.0139	0.361	Robustness test
H30-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	30	55	0.0159	0.316	Test sample, robustness test
H35-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	35	55	0.0207	0.242	Test sample, robustness test
H40-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	40	55	0.0255	0.197	Test sample, robustness test
H45-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	45	55	0.0302	0.166	Test sample, robustness test
H50-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	50	55	0.0349	0.144	Test sample, robustness test
H55-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	55	55	0.0396	0.127	Test sample, robustness test
H28-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	28	18	0.0411	0.122	Robustness test
H60-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	60	55	0.0443	0.113	Test sample, robustness test
H30-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	30	18	0.0472	0.106	Robustness test
H65-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	65	55	0.0490	0.102	Test sample, robustness test
H69.7-FL55	Nikon D3100, AF-S DX NIKKOR 18–55 mm	69.7	55	0.0535	0.094	Test sample, robustness test
H35-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	35	18	0.0614	0.082	Robustness test
H40-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	40	18	0.0744	0.067	Robustness test
H45-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	45	18	0.0883	0.057	Robustness test
H50-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	50	18	0.102	0.049	Robustness test
H55-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	55	18	0.115	0.043	Robustness test
H60-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	60	18	0.129	0.039	Robustness test
H65-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	65	18	0.143	0.035	Robustness test
H69.7-FL18	Nikon D3100, AF-S DX NIKKOR 18–55 mm	69.7	18	0.156	0.032	Test sample, robustness test

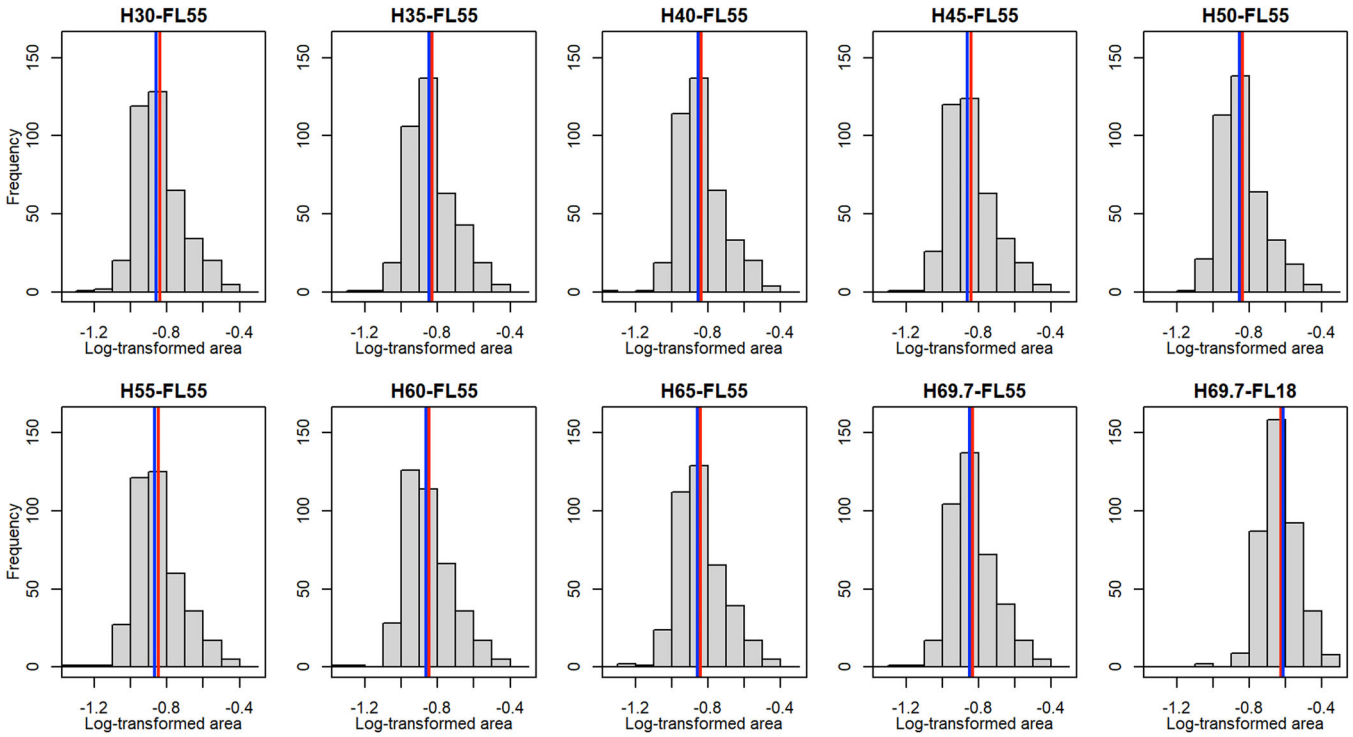

FIGURE 2 Histograms of converted seed area of the *Mimulus guttatus* samples imaged at camera configurations according to Table 2, $\log_{10}$-transformed to normalize the distribution. The vertical blue line represents the median, and the red line represents the mean seed area.

TABLE 3 Mean seed area (mm^2) and 95% confidence intervals (CI) for seed from *Mimulus guttatus* test data. Seed size was calculated on $\log_{10}$-transformed data from images at various camera configurations; the inverse model used seed size from all camera configurations. The mean number of pixels at each camera configuration is shown.

Model/ camera configuration	Log_{10}-transformed area	Back-transformed area (mm^2)	95% CI (mm^2)	Mean seed area (px)
Inverse model	−0.845	0.143	0.141–0.145	—
H30-FL55	−0.839	0.145	0.141–0.149	600
H35-FL55	−0.830	0.148	0.144–0.152	359
H40-FL55	−0.837	0.146	0.141–0.150	235
H45-FL55	−0.842	0.144	0.140–0.148	166
H50-FL55	−0.839	0.145	0.141–0.149	124
H55-FL55	−0.846	0.143	0.138–0.147	95
H60-FL55	−0.848	0.142	0.138–0.146	76
H65-FL55	−0.841	0.144	0.140–0.149	63
H69.7-FL55	−0.833	0.147	0.143–0.151	54
H69.7-FL18	−0.625	0.237	0.232–0.242	10

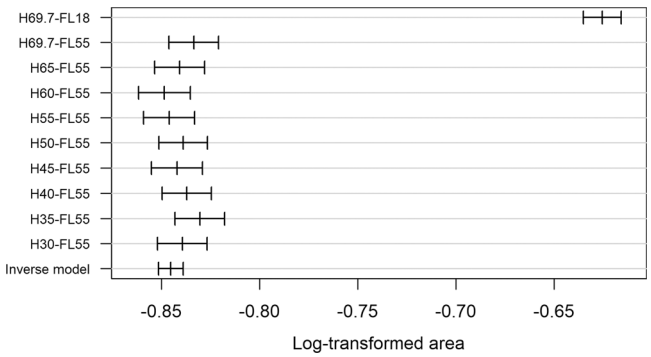

FIGURE 3 Mean area (±95% CI) of $\log_{10}$-transformed *Mimulus guttatus* seeds for the inverse model and for images captured at various camera configurations.

seeds at the different imaging configurations. The reliability of the measurements is high for all configurations except for those with the lowest magnification and pixel count. Specifically, measures of *B. rapa* seeds are robust until the final camera configuration, while measures of the smaller seeds of *A. thaliana* and *M. guttatus* become unreliable between the configurations of H28-FL18–H60-FL55, and H40-FL18–H45-FL18, respectively (Figure 5). At these threshold configurations, the mean number of pixels for *A. thaliana* (H60-FL55) and *M. guttatus* (H45-FL18) are 44.6 px and 19.0 px, respectively, while the mean number of pixels at the lowest magnification of *B. rapa* seeds (H69.7-FL18) is 83.0 px. The converted area of *A. thaliana* and *M. guttatus* seeds tends to increase at lower pixel values, with particular overestimation of the size of *M. guttatus* seeds with camera configuration H65-FL18. For raw data of seed pixel counts at each camera configuration, see

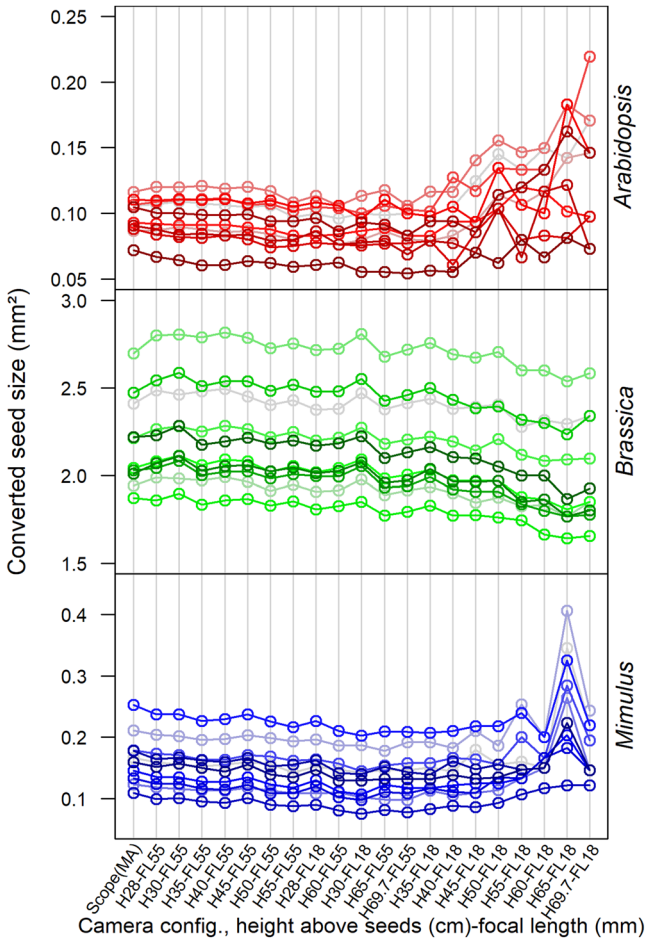

FIGURE 4 Seed size for 10 replicate seeds from *Arabidopsis thaliana*, *Brassica rapa*, and *Mimulus guttatus* imaged at various camera configurations. Open circles represent each of the replicate seeds, and lines connect a given seed across images from different configurations.

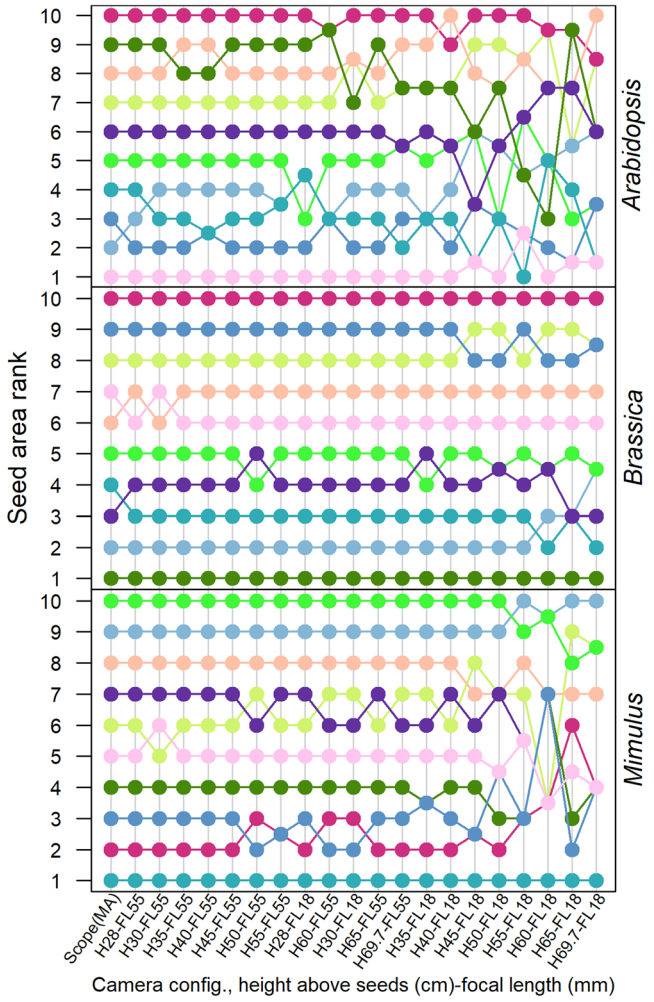

FIGURE 5 The rank order of seed size for 10 replicate seeds from *Arabidopsis thaliana*, *Brassica rapa*, and *Mimulus guttatus* imaged at various camera configurations. Closed circles represent each of the replicate seeds, and lines connect a given seed across images from different configurations. The actual sizes associated with each ranked seed are shown in Figure 4.

Appendices S6, S7, and S8 for *A. thaliana*, *B. rapa*, and *M. guttatus*, respectively. The pixel counts across species are compared in Appendix S9.

We created a heuristic for pre-emptively determining camera parameters by using a conservative lower bound of 100 px per particle captured in an image. The model converts the area of a particle comprising 100 px to the empirical conversion scale for the camera configurations (Table 2) using the following relation:

$$E_s = \sqrt{\frac{A}{100px}}$$

where E_s is the empirical scale and A is the area of the particle (Figure 6). To select an empirical scale for batch processing of samples, one would assess the approximate area of a novel particle and select a scale so that the area falls anywhere below the curve. While a greater number of pixels increases the accuracy of the area estimates, this gain diminishes as the number of pixels increases; furthermore,

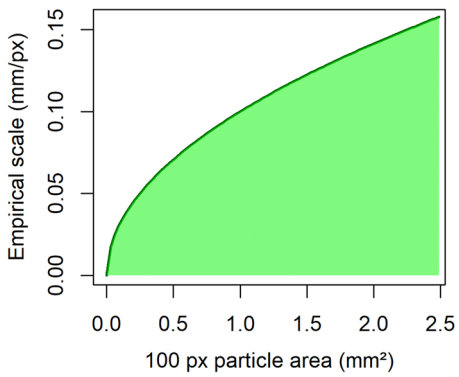

FIGURE 6 The relation between the area of a 100-px particle and the empirical scale at various camera configurations (Table 2). Given a novel particle of a specific area, any empirical scale that falls under the curve will give reliable area results.

greater pixels limit the number of propagules that fit within the camera frame.

DISCUSSION

Our study presents an efficient and standardized technique for using digital photographs to estimate seed traits, including seed number, size, and color. We applied the protocol to assess the accuracy of seed size estimates in three model plant species and found broad tolerance for a range of camera configurations. Our findings indicate that increasing the focal distance enhances the accuracy of seed size estimates; however, the improvement plateaus and compromises the ability to image large numbers of seeds at once. We demonstrate the empirical scale necessary to accurately estimate seed area for particles of specific sizes and connect this to the parameters of the imaging system (i.e., camera configuration). We recommend a conservative minimum pixel count of 100 px for imaged particles to ensure accuracy. In the following sections, we elaborate on the significance and future applications of the protocol.

Our results showed that seeds of *M. guttatus* have a right-skewed distribution, even after $\log_{10}$-transformation. The seed we used came from field-collected seed capsules, so the distribution could be caused by various factors including seed inviability or seed abortion due to inbreeding (e.g., Martin and Lee, 1993) or hybrid incompatibility (e.g., Coughlan et al., 2020). *Mimulus guttatus* can produce more than 1300 seeds per seed capsule (Waser et al., 1982; Searcy and Macnair 1990), and the relatively small seeds could result in right-skewness in the distribution of size.

Importantly for the accuracy and robustness of the protocol, the distribution and mean seed sizes remained consistent across all camera configurations, with the exception of H69.7-FL18 (i.e., camera height = 69.7 cm, focal length = 18 mm). This exception is most likely a result of poor resolution due to the camera configuration capturing insufficient pixel counts to precisely characterize seed areas. Taking that into consideration and omitting H69.7-FL18 from interpretation, we find that the mean seed area remained consistent across camera heights and focal distances. Therefore, it appears that most camera distances can robustly capture seed size in *M. guttatus*.

We tested our protocol on three model species—*A. thaliana*, *B. rapa*, and *M. guttatus*—and found the method to accurately and consistently measure seed size at most camera heights and focal lengths. However, error was introduced at species-specific camera configurations that depended on the size of their seeds. The breakdown occurs as camera magnification results in insufficient pixels to precisely resolve differences in imaged propagule area. For *A. thaliana*, we found that a camera height of 55 cm and a focal length of 55 mm was the threshold distance for accurate seed size measurements, and that any camera heights that were lower and of the same focal length did not differ and therefore were all appropriate for accurate estimation of seed size. The threshold distance for estimating seed size in *M. guttatus* was higher due to larger seed size relative to *A. thaliana* and could be estimated accurately with a camera distance up to 60 cm and a focal length of 55 mm. Finally, because *B. rapa* produces relatively large seeds, all camera distances and focal lengths produced similar seed size estimates. Notably, we also found that for all three species "manual" estimation of seed size (i.e., estimating seed size with a microscope) was not significantly different than seed size estimated with an appropriate camera distance and focal distance ranges described above, thereby validating our technique.

To select an empirical scale for an untested species, we developed a reference heuristic model that facilitates the selection of an appropriate camera configuration to image novel particles of an approximate area at 100 px. The simple model accounts for the empirical scales available according to our camera parameters and configurations and could be produced for any imaging bench.

Although the technique developed here is targeted to seeds, it could be used to estimate the number, size, and color of any small particle. For example, ImageJ has been used to manually estimate pollen number and size in flowering plant species (Costa and Yang, 2009; Kakui et al., 2021), colocalization of organelles within cells (Stauffer et al., 2018), and the abundance of scrub typhus cells multiplying under a microscope (Siritantikorn et al., 2012). The protocol developed here could be applied to a variety of particles and taxa.

CONCLUSIONS

Seed number and seed quality are crucial factors affecting fitness in flowering plants. However, a standardized and efficient technique for estimating these traits is lacking. Our study highlights the importance of validation and robustness when estimating seed size to make accurate comparisons between individual propagules or samples from populations. We developed customized methods for estimating seed size in three model species and found that while certain camera distance thresholds are necessary for capturing accurate seed size measurements, there is considerable flexibility. Our study demonstrates that it is possible to accurately estimate the number, size, and color of most particles efficiently using bulked seed imaging.

AUTHOR CONTRIBUTIONS

C.S., J.L., and J.F. conceived the project; C.S. and J.L. collected the data; C.S., J.L., and J.F. wrote the manuscript; J.L. developed the software and code; J.F. oversaw the methods and writing. All authors approved the final version of the manuscript.

ACKNOWLEDGMENTS

The authors thank colleagues from Queen's University, including I. Lewis and E. Gillette for assistance with photographing seeds, C. Smith and A. Van Natto for

insights on statistical methods, and J. Monaghan and A. Rooke for supplying *Arabidopsis* and *Brassica* seeds. This research was supported by the Natural Sciences and Engineering Research Council of Canada, through a Canada Graduate Scholarship (C.S.) and a Discovery Grant (J.F.).

DATA AVAILABILITY STATEMENT

All supporting data are provided with the article and Supporting Information.

ORCID

Christina Steinecke http://orcid.org/0000-0001-7946-5369
Jannice Friedman http://orcid.org/0000-0002-1146-0892

REFERENCES

Aarssen, L. W., and C. Y. Jordan. 2001. Between-species patterns of covariation in plant size, seed size, and fecundity in monocarpic herbs. *Ecoscience* 8: 471–477.

Andersson, S. 1990. Paternal effects on seed size in a population of *Crepis tectorum* (Asteraceae). *Oikos* 59: 3–8.

Atkinson, P. M., and N. J. Tate. 2000. Spatial scale problems and geostatistical solutions: A review. *Professional Geographer* 52: 607–623.

Baker, H. G. 1972. Seed weight in relation to environmental conditions in California. *Ecology* 53: 997–1010.

Bogdziewicz, M., J. M. Espelta, and R. Bonal. 2019. Tolerance to seed predation mediated by seed size increases at lower latitudes in a Mediterranean oak. *Annals of Botany* 123: 707–714.

Cavers, P. B., and M. G. Steel. 1984. Patterns of change in seed weight over time on individual plants. *American Naturalist* 124: 324–335.

Costa, C. M., and S. Yang. 2009. Counting pollen grains using readily available, free image processing analysis software. *Annals of Botany* 104: 1005–1010.

Coughlan, J. M., M. W. Brown, and J. H. Willis. 2020. Patterns of hybrid seed inviability in the *Mimulus guttatus* sp. complex reveal a potential role of parental conflict in reproductive isolation. *Current Biology* 30: 83–93.

de Carvalho, D. U., D. A. Boakye, T. Gasm, R. P. Leite Jr., and F. Alferez. 2021. Determining seed viability during fruit maturation to improve seed production and availability of new citrus rootstocks. *Frontiers in Plant Science* 12: 777078.

Dunwoody, E., A. Apan, and X. Liu. 2013. Effects of spatial resolution on measurement of landscape function using the landscape leakiness calculator. Proceedings of the Spatial Sciences & Surveying Conference 2013, Canberra, Australia.

Elliott, R. H., L. W. Mann, and O. O. Olfert. 2007. Effects of seed size and seed weight on seedling establishment, seedling vigour, and tolerance of summer turnip rape (*Brassica rapa*) to flea beetles, *Phyllotreta* spp. *Canadian Journal of Plant Science* 87: 385–393.

Eriksson, O. 1999. Seed size variation and its effect on germination and seedling performance in the clonal herb *Convallaria majalis*. *Acta Oecologia* 20: 61–66.

Eriksson, O. 2008. Evolution of seed size and biotic seed dispersal in angiosperms: Paleoecological and neoecological evidence. *International Journal of Plant Sciences* 169: 863–870.

Fuller, W., C. E. Hance, and M. J. Hutchings. 1983. Within-season fluctuations in mean fruit weight in *Leontodon hispidus* L. *Annals of Botany* 51: 545–549.

Gnan, S., A. Priest, and P. X. Kover. 2014. The genetic basis of natural variation in seed size and seed number and their trade-off using *Arabidopsis thaliana* MAGIC lines. *Genetics* 198: 1751–1758.

Greenway, C. A., and L. D. Harder. 2007. Variation in ovule and seed size and association size–number trade-offs in angiosperms. *American Journal of Botany* 94: 840–846.

Harper, J. L., P. H. Lovell, and K. G. Moore. 1970. The shapes and sizes of seeds. *Annual Review of Ecology, Evolution, and Systematics* 1: 327–356.

Hendrix, S. D. 1984. Variation in seed weight and its effects on germination in *Pastinaca sativa* L. (Umbelliferae). *American Journal of Botany* 71: 795–802.

Herbert, A. 2011. ImageJ Batch Processing. Available at: https://imagej.nih.gov/ij/docs/pdfs/ImageJBatchProcessing.pdf [accessed 12 December 2022].

Janzen, D. H. 1977. Variation in seed size within a crop of a Costa Rican *Mucuna andreana* (Leguminosae). *American Journal of Botany* 64: 347–349.

Kakui, H., E. Tsurisake, R. Shibata, and Y. Moriguchi. 2021. Factors affecting the number of pollen grains per male strobilus in Japanese Cedar (*Cryptomeria japonica*). *Plants* 10: 856.

Krannitz, P. G., L. W. Aarsen, and J. M. Dow. 1991. The effect of genetically based differences in seed size on seedling survival in *Arabidopsis thaliana* (Brassicaceae). *American Journal of Botany* 78: 446–450.

Leishman, M. R., M. Westoby, and E. Jurado. 1995. Correlates of seed size variation: A comparison among five temperate floras. *Journal of Ecology* 83: 517–530.

Leishman, M. R., and B. R. Murray. 2001. The relationship between seed size and abundance in plant communities: Model predictions and observed patterns. *Oikos* 94: 151–161.

Martin, M. E., and T. D. Lee. 1993. Self pollination and resource availability affect ovule abortion in *Cassia fasciculata* (Caesalpiniaceae). *Oecologia* 94: 503–509.

Martínez-González, I., L. R. Sánchez-Velázquez, B. Ruiz-Guerra, M. del Rosario Pineda-López, and N. Velázquez-Rosas. 2021. The role of seed size in the emergence and survival of seedlings in contrasting environments: The case of *Ceiba aesculifolia*. *New Forests* 52: 493–507.

Mazer, S. J., A. A. Snow, and M. L. Stanton. 1986. Fertilization dynamics and parental effects upon fruit development in *Raphanus raphanistrum*: Consequences for seed size variation. *American Journal of Botany* 73: 500–511.

Michaels, H. J., B. Benner, A. P. Hartgerink, T. D. Lee, S. Rice, M. F. Willson, and R. I. Bertin. 1988. Seed size variation: Magnitude, distribution, and ecological correlates. *Evolutionary Ecology* 2: 157–166.

Moles, A. T., D. D. Ackerly, C. O. Webb, J. C. Tweddle, J. B. Dickie, and M. Westoby. 2005. A brief history of seed size. *Science* 307: 576–580.

Moles, A. T., D. D. Ackerly, J. C. Tweddle, J. B. Dickie, R. Smith, M. R. Leishman, M. M. Mayfield, et al. 2007. Global patterns in seed size. *Global Ecology and Biogeography* 16: 109–116.

Mussadiq, Z., B. Laszlo, L. Helyes, and C. Gyuricza. 2015. Evaluation and comparison of open source program solutions for automatic seed counting on digital images. *Computers and Electronics in Agriculture* 117: 194–199.

Paczesniak, D., M. Pellino, R. Goertzen, D. Guenter, S. Jahnke, A. Fischbach, J. T. Lovell, and T. F. Sharbel. 2022. Seed size, endosperm, and germination variation in sexual and apomictic *Boechera*. *Frontiers of Plant Science* 13: 991531. https://doi.org/10.3389/fpls.2022.991531

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Website: https://www.R-project.org/ [accessed 13 September 2023].

Rasband, W. S. 2018. ImageJ. U.S. National Institutes of Health, Bethesda, Maryland, USA. Website: https://imagej.nih.gov/ij/ [accessed 13 September 2023].

Raunsgard, A., Ø. H. Opedal, R. K. Ekrem, and C. Pélabon. 2018. Intersexual conflict over seed size is stronger in more outcrossed populations of a mixed-mating plant. *Proceedings of the National Academy of Sciences, USA* 115: 11561–11566.

Ridler, T. W., and S. Calvard. 1978. Picture thresholding using an iterative selection method. *IEEE Transactions on Systems, Man, and Cybernetics* 8: 630–632.

Saatkamp, A., A. Cochrane, L. Commander, L. K. Guja, B. Jiménez-Alfaro, J. Larson, A. Nicotra, et al. 2019. A research agenda for seed-trait functional ecology. *New Phytologist* 221: 1764–1775.

Searcy, K. B., and M. R. Macnair. 1990. Differential seed production in *Mimulus guttatus* in response to increasing concentrations of copper in the pistil by pollen from copper tolerant and sensitive sources. *Evolution* 44: 1424–1435.

Siritantikorn, S., S. Jintaworn, S. Noisakran, Y. Suputtamongkol, D. H. Paris, and S. D. Blacksell. 2012. Application of ImageJ program to the enumeration of *Orienta tsutsugamushi* organisms cultered in vitro. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 106: 632–635.

Stanton, M. L. 1984. Seed variation in wild radish: Effect of seed size on components of seedling and adult fitness. *Ecology* 65: 1105–1112.

Stauffer, W., H. Sheng, and H. N. Lim. 2018. EzColocalization: An ImageJ plugin for visualizing and measuring colocalization in cells and organelles. *Scientific Reports* 25: 15764.

Thompson, J. N. 1984. Variation among individual seed masses in *Lornatium grayi* (Umbelliferae) under controlled conditions: Magnitude and partitioning of the variance. *Ecology* 65: 626–631.

Torres, C., M. C. Eynard, M. A. Aizen, and L. Galetto. 2002. Selective fruit maturation and seedling performance in *Acacia caven* (Fabaceae). *International Journal of Plant Sciences* 163: 809–813.

Waser, N. M., R. K. Vickery Jr., and M. V. Price. 1982. Patterns of seed dispersal and population differentiation in *Mimulus guttatus*. *Evolution* 36: 753–761.

Westoby, M., E. Jurado, and M. Leishman. 1992. Comparative evolutionary ecology of seed size. *Trends in Ecology and Evolution* 7: 368–372.

Westoby, M., M. Leishman, and J. Lord. 1996. Comparative ecology of seed size and dispersal. *Philosophical Transactions of the Royal Society B. Biological Sciences* 351: 1309–1318.

Wu, J., W. Shen, W. Sun, and P. Tueller. 2002. Empirical patterns of the effects of changing scale on landscape metrics. *Landscape Ecology* 17: 761–782.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1. Instructions for using the ImageJ and R scripts to perform high-throughput particle size and relative color characterization.

Appendix S2. The ImageJ macro file (particleSizeID.txt) for batch processing.

Appendix S3. An R script (particleSizeProcess.R) for reading and processing ImageJ output.

Appendix S4. Seed size estimates used to determine the accuracy of the protocol for measuring particle area. Only *Mimulus guttatus* seeds were used here and photographed at multiple heights to determine the optimal camera distance to capture the number of pixels per seed, the converted area (mm^2), log_{10}-transformation of the converted area, and the average and range of the relative red, green, and blue hue of the pixels.

Appendix S5. Histograms of converted seed area of the *Mimulus guttatus* samples imaged at camera configurations according to Table 2. The vertical blue line represents the median, and the red line represents the mean.

Appendix S6. *Arabidopsis thaliana* seed pixel areas from the seed size robustness test, where the areas of 10 seeds were estimated at all camera configurations and under a Leica S8APO stereo dissection microscope with an attached Leica Flexacam C1 camera.

Appendix S7. *Brassica rapa* seed pixel areas from the seed size robustness test, where the areas of 10 seeds were estimated at all camera configurations and under a Leica S8APO stereo dissection microscope with an attached Leica Flexacam C1 camera.

Appendix S8. *Mimulus guttatus* seed pixel areas from the seed size robustness test, where the areas of 10 seeds were estimated at all camera configurations and under a Leica S8APO stereo dissection microscope with an attached Leica Flexacam C1 camera.

Appendix S9. Mean seed area (point) and range (bars) in number of pixels for the 10 individual seeds of each *Arabidopsis*, *Brassica*, and *Mimulus* imaged at multiple camera configurations. The *x*-axis is categorical and refers to the height of the camera (H) and the focal length (FL). *Brassica* does not have data in the Scope (MA) category as those seeds were too large to fit in the Flexacam C1 frame at the same magnification used by *Arabidopsis* and *Mimulus*.

How to cite this article: Steinecke, C., J. Lee, and J. Friedman. 2023. A standardized and efficient technique to estimate seed traits in plants with numerous small propagules. *Applications in Plant Sciences* 11(5): e11552. https://doi.org/10.1002/aps3.11552