

Short Communication: Analysis of Grain Size Distribution through Image Analysis

Dennis C. Gitz III, Jeffrey T. Baker, Paxton Payton, Zhanguo Xin, Robert J. Lascano

USDA-ARS, Cropping Systems Research Laboratory, Lubbock, TX, USA

Email: dennis.gitz@ars.usda.gov

How to cite this paper: Gitz III, D.C., Baker, J.T., Payton, P., Xin, Z. and Lascano, R.J. (2018) Short Communication: Analysis of Grain Size Distribution through Image Analysis. *American Journal of Plant Sciences*, 9, 2339-2346.

<https://doi.org/10.4236/ajps.2018.912169>

Received: September 25, 2018

Accepted: October 29, 2018

Published: November 2, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Herein we describe an approach to measure the volume and to characterize the volume distribution of large numbers of individual small regularly shaped seeds. The results of a preliminary investigation into the distribution of seed size of a multi-seeded sorghum mutant, as compared to the wild type from which it was developed are also reported, and are used as an example of the method's utility.

Keywords

Seed Size, Grain Size, Sorghum, Twin Seed, Multi-Seed, Image Analysis

1. Introduction

The maximal pre-harvest yield of a standing grain crop is the product of the number of grains per unit area and the average grain mass. It has been repeatedly reported that grain yield is more limited by numbers of sinks than by the photosynthate available for development [1] [2] [3]. It is generally accepted that although crop yield is more dependent on seed numbers, seed size is also an important determinate of yield. The physiological aspects of seed filling and control of seed size are still not completely understood [4]. The term "seed" is used herein to collectively describe the harvestable organs. While "seed" is accurate in the case of crops such as soybean, it should be recognized that in cases such as grains, achenes, or nuts, these organs are more precisely fruits.

Rigorous objective investigation of the physiological, developmental, and environmental determinants of seed size requires either mass determination or dimensional analyses of individual seeds. Mass determination is theoretically straight-forward, requiring little more than a laboratory balance. In the case of larger seeds, such as acorns, this has been done in ecological studies by weighing

numbers of seeds individually [5]. Relatively large numbers of seeds, around a thousand, can be weighed and the mass distribution described as a histogram or with a box and whisker plot. Determining the mass of individual smaller seeds is not possible, or is impractical. In the case of smaller seeds, a sub-sample of a number or a volume of seeds is weighed [6]. This results in an averaged seed mass of the sub-sample, which is extended to describe the population. Simple averages obtained in this way are useful but result in a less nuanced description of the variation in seed size.

The other simple direct approach to describe sampled seed size within a population is the physical determination of size through dimensional analysis. In principle dimensional analysis requires only a finely divided linear scale or set of calipers, some knowledge of the shape of the seed, and an analytic geometric expression relating seed volume to the measured dimensions. This is possible with small numbers of larger seeds. But, attempting to measure the dimensions of representative samples of smaller individual seeds might be difficult. In the case of sorghum [*Sorghum bicolor* (L.) Moench], a panicle of seeds contains upwards of 1500 rather small seeds and often considerably more. Measuring the length and width of even a small number of seeds of even a small sub-sample is time consuming because the seeds are rather small and difficult to manipulate. Attempting to measure a large enough number of seeds to adequately describe the size distribution from even a single panicle rapidly becomes intractable problem.

Such practical limitations and considerations of the accurate determination of seed size have likely impeded such investigations, especially with small seeds. It is likely that these and similar considerations have led to the qualitative nature of such studies. Traditionally, segregation or characterization and of seeds has been done by passing through successively finer screens or sieves with circular openings and subsequently weighing or counting [7] [8].

Image analysis has been used to examine and characterize seed uniformity and size as area. In practice, seeds are distributed on a digital scanner or are photographed and the resulting relative sizes of individual seeds determined as the length, width, or projected area of the seeds as pixels [9]. Other approaches examine seed shape as well as projected area. This is done by expressing the projected area of the seed as geometric shapes and comparing calculated parameters such as eccentricity, circularity, flatness, and roughness [10]. However, the practice of seed image analysis is not well developed. Attempting to estimate the volume of individual seeds as compared to the projected area has not been as extensively employed. Herein we describe a technique to analyze and characterize developmentally distinct cohorts of seeds developing of a sorghum panicle.

2. Materials and Methods

2.1. Materials

Seeds from sorghum plants with differing fruit development patterns were taken

from the seed storage room at the Lubbock, TX USDA-ARS Cropping Systems Research Laboratory. No effort was made to select envelopes containing similar volumes or masses of seeds. A technician was simply asked to provide envelopes for evaluation. The two envelopes held seed from either a normally developing public cultivar, BTx 623, or from a mutant line, MSD-P12, that exhibited altered panicle, floral, and fruit development [11] [12]. Each of the two envelopes contained seeds that had been mechanically threshed from a single panicle.

The mutant line was derived by chemical mutagenesis of the “wild type” BTx623. Briefly, the mutants were developed by soaking Btx623 seeds in an aqueous solution of ethyl methane sulfonate, thoroughly rinsed several times with water, air dried, planted, allowed to develop, mutants selected, and back-crossed with the wild type several times which resulted in homozygous stable lines with the recessive heritable multi-seed allele [11].

2.2. Methods

Digital image analysis was used to calculate the volume of individual seeds of each genotype and the frequency distribution characterized. Seeds were mechanically mixed by shaking the envelope, successive 8.5 ml scoops of seeds taken, and the samples placed in a petri dish. The number of scoops of each genotype was increased until numbers were great enough generate a smooth frequency distribution graph and then varied so that a similar number of seeds of each type was used for analysis. The numbers of BTx-623 seeds and MSD-P12 seeds used were 859 and 847 seeds, respectively. Any debris remaining from mechanical threshing was removed from the sub-sample with forceps. Each scoopful of seed was arranged on the platen of a bench top flat bed scanner (Fujitsu Model Fi-65f, Sunnyvale, CA), the seeds scanned at the native optical resolution of 600×600 dpi, and the captured images saved as jpg image files for analysis. Each seed within the image was analyzed (SigmaScan Pro, Golden Software) and the major and minor axes determined. The volume (mm^3) of each seed was calculated as a regular ellipsoid having two minor axes of equal length and a longer major axis. The data were imported into a spread-sheeting routine, sorted in increasing volume, and a frequency distribution generated from 2 to 75 mm^3 with 0.5 mm^3 buckets using the histogram function within the spread-sheeting routine (Quattro Pro X6, Corel Corp., Ottawa, ON). The resulting frequency distribution data were normalized setting the integrated area under the distribution curve to unity, smoothed [13] and the resulting curve deconvoluted into gaussian curves the sum of which (PeakFit, SeaSolve Software Inc., Framingham, MA), approximated the smoothed curve. The heights and widths of the gaussian curves were iteratively varied to minimize the residual unexplained error. For presentation, the resulting data were imported into a graphing program in which the raw data, frequency distribution curve, the deconvoluted histograms plotted as percentage of seeds with a class variable width of 0.5 mm^3 (Sigma Plot, Golden Software, Golden CO).

3. Results

Examples of sorghum seed images acquired with the benchtop scanner are shown in **Figure 1**. BTx623 seeds (Upper Panel Fig 1a) were clearly larger and less variable in apparent size as opposed to the MSD seeds. Scatterplots of the observed seed size distribution, the 95% prediction intervals about the fitted curve, gaussian distributions comprising the fitted curve are shown in **Figure 2** (Data were expressed as a scatterplot in the interest of clarity, but it should be remembered that although seed size is a continuous variable, the resulting histogram data are discrete variables of size classes.) All data were normalized so that the sum of individual measurements is unity (100%). The wide gray band is the 95% prediction interval. The prediction interval is shown because the confidence interval was rather small and was hidden behind the deconvoluted gaussian peaks (red and blue line plots), especially in the case of the BTx623 seeds (**Figure 2(A)**). Residual, unexplained error is not plotted.

The frequency distribution curves of both genotypes were assumed to be composed of two normally distributed sub-populations of seeds shown as red and blue line plots. In the case of BTx623 most seeds, 94%, were within a cohort having an average volume of 48.7 mm³. The bi-modal nature of the mutant sorghum seed size distribution (**Figure 2(B)**) was clearly visible as compared to that of the wild type BTx623 from which the mutant was derived (**Figure 2(A)**). The size distribution of multiseed mutant seeds were comprised of two well separated sub-populations having average sizes of 17.4 mm³ and 37.1 mm³ comprising

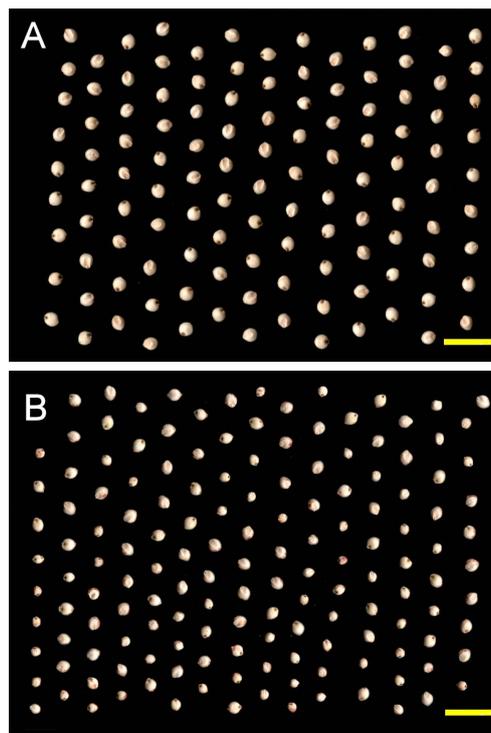


Figure 1. Examples of sorghum seed images collected with 600 × 600 dpi benchtop scanner. (A) BTx623 seeds (B) MSD seeds. Bar is 1 cm.

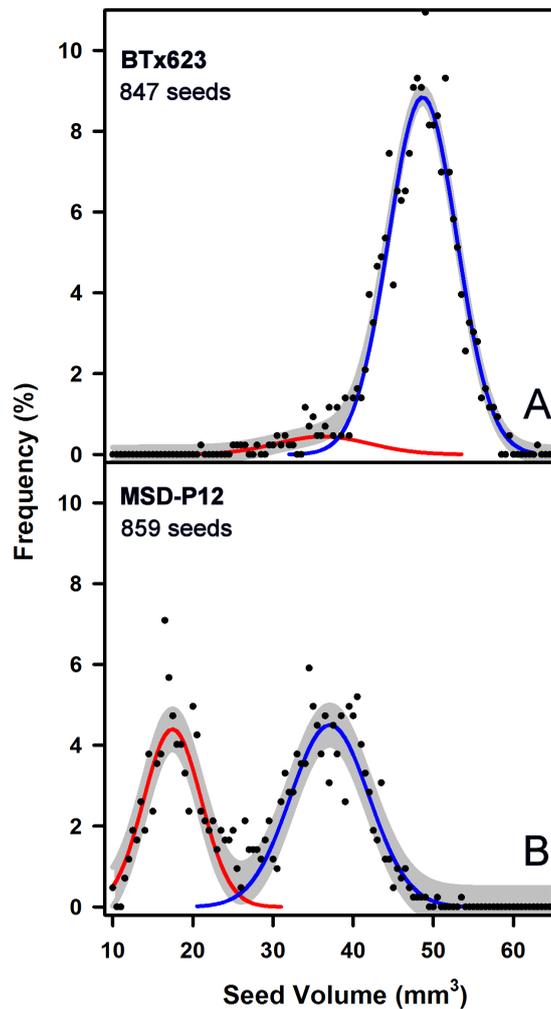


Figure 2. Frequency distribution (%) of individual sorghum seed volumes in 0.5 mm^3 volume classes increasing from 10 to 67 mm^3 . Top pane (A) is from BTx623. Bottom pane (B) is from a multiseed mutant, MSD-P12. Scatterplots (Solid circles) are raw data. Gray band is results of smoothing and 95% prediction interval. Red and blue line plots are gaussian distributions comprising smoothed plots. Numbers of seeds in subsamples are indicated.

41% and 58% of the seeds, respectively.

4. Discussion

The distribution of seed size has been described with large seeds such as acorns. In the case of acorns, the individual mass of up to 2000 seeds was determined by weighing each seed to within 1 mg [5]. A similar, but unsuccessful, approach was attempted during the present work with a manageable sample of MSD mutant sorghum seeds. It was later concluded that the weighing scales used probably did not have the needed sensitivity or resolution. Even if the scales would have been adequate a manageable sample size of 100 to 200 seeds would not have been large enough to characterize the frequency distributions of seed mass (not presented). In the case of acorns [5], the distribution of seed mass was used to infer

a range of evolutionary fitness components. Similarly, sorghum seed size affects agronomic fitness attributes such as emergence, and so, stand establishment [14], but also see [7]. In other crops, seed size affects a wide range of agronomically important crop fitness and yield characteristics [15]. Seed size uniformity can also affect post-harvest factors that directly influence harvested crop value. Soybean is an example of a crop in which seed size and uniformity affects post-harvest processing and crop value [9].

In the present work the power to detect differences was dependent on adequate numbers of individual seeds analyzed, and the small (0.5 mm³) volume classes used for analysis, although a systematic statistical sensitivity analysis was not done. Instead, the numbers of seeds were increased “scoop-wise” and the size classes decreased until details of the shapes of the distribution curves were easily resolved. Volume was used rather than projected area simply measuring the area of individual seeds as projections projected on the scanner platen might have detected differences. However, it was thought that since volume increases as the cube of the measurements rather than the square, as with area, that volume would be a more sensitive approach. Too, the amount of endosperm and maximal potential starch content within seeds is likely strongly correlated with volume. Hence, volume is biologically a more relevant measurement than projected area.

A limitation of the approach described here is the difficulty of extending the method to seeds that are irregularly shaped like wrinkled dried peas or corn, seeds that are difficult to model as with sunflower, or seeds that need an additional measurement in the axis perpendicular to the scanner platen. It seems that the method could be easily extended to other regularly shaped seed such as rice, mustard family seeds, and small legume seeds. Another limitation is that the seeds had to be manually separated so that they didn't touch each other and that each seed was located and identified manually. Automating both procedures would allow much greater sample throughput.

The selection of seed for analysis was based upon a long standing question that was recently re-examined. Earlier it was thought that the MSD mutants could provide a trait through which sorghum yield could be substantially increased [11]. This hypothesis was supported by a body of research that suggested increasing seed numbers led to increased yield (e.g. [1] [4] [3]). The concept of increasing sorghum yield by increasing seed numbers was suggested at least as early as the 1970's [16] [17]. No commercial lines with enhanced yield resulting from increased seed numbers have been developed as a result of the early research. An MSD mutant was subsequently developed and grown [11]. It was again thought that the trait could result in increased yields. However, careful work failed to detect expected yield increases [12]. It was concluded that simply increasing seed numbers does not in and of itself lead to increased yield in sorghum. Nevertheless, it remains possible that very small seeds might have been lost during mechanical threshing and passed through along with the trash. Seed size

analysis might resolve such questions.

Caution should be used in attempting to extending the results presented herein to draw conclusions about the mutant MSD sorghum lines. A single envelope containing seeds from a single mechanically threshed plant fails to address consistency and reproducibility. While the results presented herein are consistent with what is known about seed development in the two sorghum lines, further work with larger sample sets is needed to rigorously examine how seed development, seed size, and yield differ between the two lines.

5. Conclusion

Herein we detailed a procedure through which individual seed volumes of populations known to have different developmental histories can be described. Conversely, this suggests that digital seed volume analysis can be used to detect differences in seed development resulting from genotypic or environmental responses. The method could be useful in other respects especially if automated. Characterizing and controlling seed size variability might also lead to reductions in experimental variability (statistical error) and increase power to detect differences in field experiments. Selection of individuals with traits associated with seed size such as long-term seed viability, germination and emergence, low temperature emergence, seedling vigor, and stand establishment might benefit from controlling or eliminating error associated with seed volume.

Acknowledgements

The technical assistance of Messrs. Ryan Mounce and Kyle Tengler is acknowledged. Dr. Liu-Gitz is most gratefully acknowledged for generously providing materials for analysis and helpful discussions throughout the work. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The USDA is an equal opportunity employer.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Malik, A.S., Borrás, L., Slafer, G.A. and Otegui, M.E. (2004) Seed Dry Weight Response to Source-Sink Manipulations in Wheat, Maize and Soybean: A Quantitative Reappraisal. *Field Crops Research*, **86**, 131-146. <https://doi.org/10.1016/j.fcr.2003.08.002>
- [2] Borrás, L. and Gambín, B.L. (2010) Trait Dissection of Maize Kernel Weight: Towards Integrating Hierarchical Scales Using a Plant Growth Approach. *Field Crops Research*, **118**, 1-12. <https://doi.org/10.1016/j.fcr.2010.04.010>
- [3] Ordóñez, R.A., Savin, R., Cossani, C.M. and Slafer, G.A. (2018) Maize Grain Weight

- Sensitivity to Source-Sink Manipulations under a Wide Range of Field Conditions. *Crop Science*, **58**, 1-16.
- [4] Chang, T.-G. and Zhu, X.-G. (2017) Source-Sink Interaction: A Century Old Concept under the Light of Modern Molecular Systems Biology. *Journal of Experimental Botany*, **68**, 4417-4431. <https://doi.org/10.1093/jxb/erx002>
- [5] Gomez, J.M. (2004) Bigger Is Not Always Better: Conflicting Selective Pressures on Seed Size in *Quercus Ilex*. *Evolution*, **58**, 71-80. <https://doi.org/10.1111/j.0014-3820.2004.tb01574.x>
- [6] Wang, B., Phillips, J.S. and Tomlinson, K.W. (2018) Tradeoff between Physical and Chemical Defense in Plant Seeds Is Mediated by Seed Mass. *Oikos*, **127**, 440-447. <https://doi.org/10.1111/oik.04867>
- [7] Abdullahi, A. and Vanderlip, R.L. (1972) Relationships of Vigor Tests and Seed Source and Size to Sorghum Seedling Establishment. *Agronomy Journal*, **64**, 143-144. <https://doi.org/10.2134/agronj1972.00021962006400020004x>
- [8] Samarah, N., Mullen, R. and Cianzio, S. (2004) Size Distribution and Mineral Nutrients of Soybean Seeds in Response to Drought Stress. *Journal of Plant Nutrition*, **27**, 815-835. <https://doi.org/10.1081/PLN-120030673>
- [9] Shahin, M.A., Symons, S.J. and Poysa, V.W. (2006) Determining Soya Bean Seed Size Uniformity with Image Analysis. *Biosystems Engineering*, **94**, 191-198. <https://doi.org/10.1016/j.biosystemseng.2006.02.011>
- [10] Cervantes, E., Martín, J.J. and Saadaoui, E. (2016) Review Article: Updated Methods for Seed Shape Analysis. *Scientifica*, **2016**, Article ID: 5691825. <https://doi.org/10.1155/2016/5691825>
- [11] Burow, G., Xin, Z., Hayes, C. and Burke, J. (2014) Characterization of a Multiseeded (msd1) Mutant of Sorghum for Increasing Grain Yield. *Crop Science*, **54**, 2030-2037. <https://doi.org/10.2135/cropsci2013.08.0566>
- [12] Tolk, J.A. and Schwartz, R.C. (2017) Do More Seeds per Panicle Improve Grain Sorghum Yield? *Crop Science*, **57**, 490-496. <https://doi.org/10.2135/cropsci2016.04.0245>
- [13] Savitzky, A. and Golay, M.J.E. (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, **36**, 1627-1639. <https://doi.org/10.1021/ac60214a047>
- [14] Cisse, N. and Ejeta, G. (2003) Genetic Variation and Relationships among Seedling vigor Traits in Sorghum. *Crop Science*, **43**, 824-828. <https://doi.org/10.2135/cropsci2003.8240>
- [15] Tekroni, D.M. and Egli, D.B. (1991) Relationship of Seed Vigor to Crop Yield: A Review. *Crop Science*, **31**, 816-822. <https://doi.org/10.2135/cropsci1991.0011183X003100030054x>
- [16] Miller, F.R. (1979) Breeding of Sorghum. In: Harris, M.K., Ed., *Biology and Breeding for Resistance to Arthropods and Pathogens in Agricultural Plants*, Texas A&M University, Report #MP-1451, College Station, TX, 128-136.
- [17] Suh, H.W., Casady, A.J. and Vanderlip, R.L. (1974) Influence of Sorghum Seed Weight on the Performance of the Resulting Crop. *Crop Science*, **14**, 835-836. <https://doi.org/10.2135/cropsci1974.0011183X001400060017x>