An automated, high-throughput method for standardizing image color profiles to improve image-based plant phenotyping

Jeffrey C. Berry[1]**, Noah Fahlgren**[1]**, Alexandria A. Pokorny**[1]**, Rebecca Bart**[1]**, and Kira M. Veley**[1]

[1]**Donald Danforth Plant Science Center, St. Louis, MO 63132**

Corresponding author:
Kira M. Veley

Email address: kveley@danforthcenter.org

ABSTRACT

High-throughput phenotyping has emerged as a powerful method for studying plant biology. Large image-based datasets are generated and analyzed with automated image analysis pipelines. A major challenge associated with these analyses is variation in image quality that can inadvertently bias results. Images are made up of tuples of data called pixels, which consist of R, G, and B values, arranged in a grid. Many factors, for example image brightness, can influence the quality of the image that is captured. These factors alter the values of the pixels within images and consequently can bias the data and downstream analyses. Here, we provide an automated method to adjust an image-based dataset so that brightness, contrast, and color profile is standardized. The correction method is a collection of linear models that adjusts pixel tuples based on a reference panel of colors. We apply this technique to a set of images taken in a high-throughput imaging facility and successfully detect variance within the image dataset. In this case, variation resulted from temperature-dependent light intensity throughout the experiment. Using this correction method, we were able to standardize images throughout the dataset, and we show that this correction enhanced our ability to accurately quantify morphological measurements within each image. We implement this technique in a high-throughput pipeline available with this paper, and it is also implemented in PlantCV.

INTRODUCTION

Imaging and computer vision approaches are powerful tools for plant phenotyping because they allow plant physical and physiological features to be measured non-destructively with relatively high quantitative and temporal resolution (Fahlgren et al., 2015b). In addition to robotic platforms, micro-computers are making large scale image acquisition more affordable and accessible (Tovar et al., 2018). From these datasets, it is possible to measure plant size, shape, color, and other features in an automated fashion and correlate phenotypes with experimental treatments (Honsdorf et al., 2014; Chen et al., 2014; Neilson et al., 2015; Fahlgren et al., 2015a; Acosta-Gamboa et al., 2016; Feldman et al., 2017; Veley et al., 2017; Liang et al., 2018)

The use of high-throughput imaging and computer vision analysis for plant phenotyping presents challenges in data acquisition, management, and analysis that require a team with diverse expertise in biology, engineering, and mathematics (Furbank and Tester, 2011; Fahlgren et al., 2015b). One of these challenges is assuring standardization of methods so proper inferences can be made (Wang and Zhang, 2010; Furbank and Tester, 2011). For example, variation within an image dataset can have a significant impact on phenotype inferences and so must be considered. For a given pixel that is stored in an image there are three values, one for each component (red, green, and blue; RGB) (Finlayson et al., 2015; Gunturk et al., 2005). Image brightness is an overview of how large the values are for each pixel, and image contrast is defined as the range of pixel values. Color profile is the range of colors observed in the image. What we perceive as image quality is a combination of contrast and color profile among other features (Livingstone and Hubel, 1988). Images with large contrast and color profile are considered high-quality images because they have larger numerical range relative to low-quality images.

Once an image set has been collected, image analysis is the process of extracting numerical data that describes the object in the image. First, the object must be separated from background pixels through a process called object segmentation. Accuracy of segmentation decreases as the image quality decreases. Various methods for segmentation have been proposed to handle varying image qualities including use of adaptive thresholding or learning algorithms (Yogamangalam and Karthikeyan, 2013). Rather than relying on mathematical models to segment the object in the image, we propose a method of image standardization based on a reference color palate and then altering the image so fixed-threshold segmentation becomes highly robust to image quality. Standardizing the image does not require a training set which is a major bottleneck in creating learning algorithms for segmentation and reduces the variability of autothresholding. We tested this method using a dataset from an automated imaging system containing approximately 24,000 images taken over an 11-day period. Each image contains a single plant and a reference color palate. Using a fixed-threshold segmentation routine to isolate the plant, various shapes and color profile were recorded for every image. We compare the quantification of plant features before and after image standardization and demonstrate that our method improves overall accuracy of image analysis.

MATERIALS AND METHODS

Software

Image analysis was performed with the C++ source file included with this manuscript and must be compiled against OpenCV (only tested on 3.1). Statistical analyses and graphics were done using R version 3.4.4 with the following packages: ggplot 2.2.1, ggthemes 3.4.0, reshape2 1.4.3, plyr 1.8.4, grid 3.4.4, gridextra 2.2.1, vegan 2.4-6. T-test, and Wilcoxon signed-rank test were done with base R functions. Constrained analysis of principal coordinates was done with vegan.

Plant Growth and Imaging Conditions

The plant growth phenotyping data used in this experiment is a part of a large dataset. Plant growth conditions were based on a previous experiment (Veley et al., 2017) with adjustments. Details are as follows: Two sorghum genotypes (*Sorghum bicolor* (L.) Moench, genotypes BTx623 and China 17) were germinated between pieces of damp filter paper in petri dishes at 30°C for 3 days prior to planting. The genotype information was collapsed throughout this analysis. Square pots (7.5 cm wide, 20 cm high) fitted with drainage trays were pre-filled with Profile® Field & Fairway™ calcined clay mixture (Hummert International, Earth City, Missouri). On day 4 (4 DAP), germinated seeds were transplanted into pre-moistened pots, barcoded (including genotype identification, treatment group, and a unique pot identification number), randomized and scanned onto the Bellwether Phenotyping System (Conviron, day/night temperature: 32°C/22°C, day/night humidity: 40%/50%, day length: 14hr (ZT 0-14 = hr 0700-2100), night length: 9hr (ZT 14-24 = hr 2100-0700), light source: metal halide and high pressure sodium, light intensity: 400 μmol/m^2/s). Plants received 40 mL fertilizer treatments (see below) daily, and were watered to weight (640g starting dry weight, 1070g target weight DAP 4 - 11, 960g DAP 12 - 25, not including weight of phenotyping system carrier) daily by the system using distilled water. Fertilizer treatment information was collapsed for all analysis except where indicated, and were as follows:

Fertilizer Treatments:

Control fertilizer treatment (100% N): 6.5 mM KNO$_3$, 4.0 mM Ca(NO$_3$)$_2$·4H$_2$O, 1.0 mM NH$_4$H$_2$PO$_4$, 2.0 mM MgSO$_4$·7H$_2$O, micronutrients, pH 4.6.

Low nitrogen fertilizer treatment (10% N): 0.65 mM KNO_3, 4.95 mM KCl, 0.4 mM $Ca(NO_3)_2 \cdot 4H_2O$, 3.6 mM $CaCl_2 \cdot 2H_2O$, 0.1 mM $NH_4H_2PO_4$, 0.9 mM KH_2PO_4, 2.0 mM $MgSO_4 \cdot 7H_2O$, micronutrients, pH 5.0.

The same micronutrients were used for both treatments: 4.6 μM H_3BO_3, 0.5 μM $MnCl_2 \cdot 4H_2O$, 0.2 μM $ZnSO_4 \cdot 7H_2O$, 0.1 μM $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$, 0.2 μM $CuSO_4 \cdot 5H_2O$, 71.4 μM Fe-EDTA.

Standardization Method

The image standardization method used here is based on a color transfer method that can adjust the colors in an image to match a target image color profile (Gong et al., 2016). The goal is to create a transform such that when applied to the values of every pixel in a source image, it returns values mapped to a target image profile. It was shown that the color transfer is a single homography from source to target (Gong et al., 2016). Color data from the source and the target image are used to define the homography and the color data must be sampled from homologous points, so a reference is included in the images from which the values of R, G, and B can be measured. There are many references to choose from and the choice of reference matters (Ilie and Welch, 2005). We chose to include the ColorChecker Passport Photo (X-Rite, Inc.), which has a panel of 24 industry standard color reference chips, within each image. A target image (or reference) is declared and the process of computing the homography and applying it to a given source image is as follows:

1. Let T and S be matrices containing measurements for the R, G, and B components of each of the ColorChecker reference chips pictured in the target image and source image respectively.

$$T = \begin{bmatrix} T_R & T_G & T_B \end{bmatrix} = \begin{bmatrix} t_{r_1} & t_{g_1} & t_{b_1} \\ t_{r_2} & t_{g_2} & t_{b_2} \\ \vdots & \vdots & \vdots \\ t_{r_{24}} & t_{g_{24}} & t_{b_{24}} \end{bmatrix} \text{ and } S = \begin{bmatrix} S_R & S_G & S_B \end{bmatrix} = \begin{bmatrix} s_{r_1} & s_{g_1} & s_{b_1} \\ s_{r_2} & s_{g_2} & s_{b_2} \\ \vdots & \vdots & \vdots \\ s_{r_{24}} & s_{g_{24}} & s_{b_{24}} \end{bmatrix}$$

2. Extend S to include the square and cube of each element.

$$S = \begin{bmatrix} s_{r_1} & s_{g_1} & s_{b_1} & s_{r_1}^2 & s_{g_1}^2 & s_{b_1}^2 & s_{r_1}^3 & s_{g_1}^3 & s_{b_1}^3 \\ s_{r_2} & s_{g_2} & s_{b_2} & s_{r_2}^2 & s_{g_2}^2 & s_{b_2}^2 & s_{r_2}^3 & s_{g_2}^3 & s_{b_2}^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{r_{24}} & s_{g_{24}} & s_{b_{24}} & s_{r_{24}}^2 & s_{g_{24}}^2 & s_{b_{24}}^2 & s_{r_{24}}^3 & s_{g_{24}}^3 & s_{b_{24}}^3 \end{bmatrix}$$

3. Calculate M, the Moore-Penrose inverse matrix of S (* denotes transpose).

$$M = (S^*S)^{-1}S^*$$

4. Estimate standardization vectors of each R, G, and B channel by multiplying M with each column of T.

$$\begin{bmatrix} R_h \\ G_h \\ B_h \end{bmatrix} = \begin{bmatrix} MT_R \\ MT_G \\ MT_B \end{bmatrix} = \begin{bmatrix} \hat{r}_r & \hat{r}_g & \hat{r}_b & \hat{r}_{r2} & \hat{r}_{g2} & \hat{r}_{b2} & \hat{r}_{r3} & \hat{r}_{g3} & \hat{r}_{b3} \\ \hat{g}_r & \hat{g}_g & \hat{g}_b & \hat{g}_{r2} & \hat{g}_{g2} & \hat{g}_{b2} & \hat{g}_{r3} & \hat{g}_{g3} & \hat{g}_{b3} \\ \hat{b}_r & \hat{b}_g & \hat{b}_b & \hat{b}_{r2} & \hat{b}_{g2} & \hat{b}_{b2} & \hat{b}_{r3} & \hat{b}_{g3} & \hat{b}_{b3} \end{bmatrix}$$

5. To standardize the R, G, and B components of each pixel, i, in the source image, apply each standardization vector to S_i, the linear, quadratic, and cubic RGB components of i.

$$S_{i,standardized} = \begin{pmatrix} S_iR_h & S_iG_h & S_iB_h \end{pmatrix}$$

$$\text{where } S_i = \begin{pmatrix} r_i & g_i & b_i & r_i^2 & g_i^2 & b_i^2 & r_i^3 & g_i^3 & b_i^3 \end{pmatrix}$$

The resulting image will have the same color profile, contrast, and brightness of the target image. To quantify the deviance, denoted as D, between the source and target images, extend T to include quadratic and cubic terms for each R, G, and B exactly as (2) and estimate standardizations for the 6 new columns in T.

6. Construct matrix H using all 9 standardizations vectors.

$$H_{9x9} = \begin{bmatrix} R_h \\ G_h \\ B_h \\ R_h^2 \\ G_h^2 \\ B_h^2 \\ R_h^3 \\ G_h^3 \\ B_h^3 \end{bmatrix} = \begin{bmatrix} MT_R \\ MT_G \\ MT_B \\ MT_{R^2} \\ MT_{G^2} \\ MT_{B^2} \\ MT_{R^3} \\ MT_{G^3} \\ MT_{B^3} \end{bmatrix}$$

7. Compute deviance from reference.

$$D = 1 - \det(H)$$

If the source image has an identical color profile as the target image, H is the identity matrix and the matrix determinant is 1. So that no change in profile is reported as a 0, one minus the determinant is the returned value.

RESULTS

Identifying previously unknown source of variance

The Bellwether Phenotyping Facility at the Donald Danforth Plant Science Center is a controlled-environment growth facility that utilizes a Scanalyzer 3D-HT plant-to-sensor system (LemnaTec GmbH) to move plants through stationary imaging cabinets (Fahlgren et al., 2015b). The design of the imaging cabinets is intended to produce consistent images with uniform backgrounds. To test the consistency of images collected by the Bellwether system, we imaged sorghum plants for 11 days and included a ColorChecker Passport panel in the field of view so image color profiles could be compared. We designated an image taken at noon on the first day of the experiment to be the reference, and we calculated the deviance (D) for every other image. We observed that there was large variance in D in the image set (median = -0.370, mean = -8.756, variance = 530.991). By manually inspecting images that showed a large and small D, we observed that a large, negative D corresponded to a decrease image brightness. Furthermore, we observed that D had a periodic distribution with the time of day the image was taken. Collapsing all the images taken over the 11-day span to a 24-hour window shows that there is a clear time of day effect (Fig. 1A). Images taken during experimental nighttime hours had decreased brightness compared to images taken during daytime hours. (Wilcoxon rank-sum test, p-value <0.001). We tested two hypotheses that could account for the observed diurnal pattern: 1) light leakage into the imaging cabinet through its vents during the day, or 2) the temperature of the room was influencing the brightness of the fluorescent bulbs that illuminate the cabinet. To test the former hypothesis, we took two pictures containing a ColorChecker Passport panel: one with the vents open and another with the vents covered. Setting one of images as the reference and standardizing the other to it resulted in D = -0.454, which is insufficient to explain the large absolute D values observed in the plant image set. To test the latter hypothesis, a two-day test was done with a 14 hr/10 hr (day/night) photoperiod. For the first 24 hrs the temperature cycled with the photoperiod (32°C/22°C, day/night). For the second 24 hrs temperature was held at a constant 32°C. Five thermometers were placed throughout the imaging cabinet and recorded temperature every minute. Images containing a ColorChecker Passport panel were captured every hour (ZT0-ZT18; to mimic the period of time images were collected in the full experiment) and used to calculate D for each image. Temperatures within the chamber tracked closely with those in the room (Fig. 1B). We observed that under diurnal conditions, D varies with temperature, and in particular, large absolute values of D were observed for temperatures below 28°C (Fig. 1C). Additionally, the variation of D was consistent with the variation observed for the plant image set both in frequency and magnitude (Fig. 1). Furthermore, when the ambient temperature remained constant, D did not change (Fig. 1C). The imaging cabinet is illuminated with fluorescent bulbs and it is known that light output from this type of bulb is affected by ambient temperature (Bleeker and Veenstra, 1990). Using the measure of deviance defined in Eq 7, we were able to detect and identify a source of variance within our plant imaging facility.

Standardization of images improves shape and color measurements

Given that there was variable brightness in the plant image set, the commonly used fixed-threshold segmentation routines produced noisy results. Dark, low-quality images that were segmented using a fixed-threshold routine optimized for bright, high-quality images resulted in either background pixels being classified as object, object pixels being classified as background or both (Fig. 2A). For example, segmentation of an uncorrected sample image take during the experimental night fails to remove a large amount of background near darker parts of the image and salt and pepper noise around the plant. In contrast, segmentation of the transformed image results in clean segmentation of the plant without any additional filtering needed.

To further explore how our standardization method affects phenotype analysis, a set of morphological shapes of each plant was measured before and after standardization. We calculated the difference between the measurements obtained from the original images and the standardized images and grouped the results into experimental day and night (Fig. 2B). The values of the shapes are on vastly different scales so normalization of those values by subtracting the grand mean and dividing by the variance was done for each image and for each shape. The effects in this scale are equivalent to the effects in the native scale for each shape. Many commonly reported measurements (area, convex hull area, width, perimeter, circularity, solidity, center of mass x-coordinate, height, ellipse minor axis, fractal dimension, roundness, eccentricity, center of mass y-coordinate, ellipse angle, convex hull vertices) are significantly different using an unequal variance t-test (p-value <0.05) between images taken during daytime and nighttime temperatures. Other common measurements such as ellipse major axis, ellipse center x-coordinate, aspect ratio, and ellipse center y-coordinate were not significantly different (Fig. 2B). Plant area (cm^2) was previously shown to correlate to fresh weight biomass on the Bellwether platform (Fahlgren et al., 2015b), and we observed a similar relationship in the current experiment with a small subset of plants (n=162) that were manually weighed on the last

day of the experiment (Fig. 2C). Plant area responds to salt and pepper noise in a one-to-one relationship, so a relatively small amount of noise compared to total plant area will not produce large effects. In contrast, shapes like perimeter are more sensitive and produce larger differences because inclusion of one background pixel can add one or more pixels to the resulting measurement (Fig. 2D).

Next, we sought to understand how the color profile of the object was affected by our standardization method. Hue is a measure of the color in degrees around a circle of an object in an image, and we previously used hue to measure plant responses to abiotic stress Moroney et al. (2002); Veley et al. (2017). We compared the hue profiles of plants from original and standardized images and observed an increase in signal above 120° (green) in the nighttime temperature unstandardized images relative to the daytime temperature unstandardized images (Fig. 3A). Additionally, we observed a large amount of noise below 90° (red-yellow) in the same comparison. Visual inspection of a subset of images suggested that the former is due to misinterpretation of the true object pixels' shade of green and the latter is due to background being included in the hue histogram. The standardized image sets result in consistent color profiles between the nighttime and daytime images (Fig. 3A). To statistically test for differences between hue histograms, a two-sample Kolmogorov-Smirnov test was applied to each pairwise comparison. Hypothesis testing was done using the following:

$$8. \qquad K > c(\alpha)\sqrt{\frac{n+m}{nm}}$$

$$9. \qquad c(\alpha) = \sqrt{-\frac{1}{2}\ln(\frac{\alpha}{2})}$$

$$10. \qquad K = \sup \left| F_x(\text{hue}) - F_y(\text{hue}) \right|$$

F is a cumulative distribution of samples x and y, each with sample sizes n and m, respectively. In Eq 8, if the left-hand side is greater than the right-hand side at $\alpha = 0.05$, then we reject the null hypothesis that the two histograms are the same distribution. Using this test, we determined that the standardization procedure significantly alters the color profile of the images taken under nighttime temperatures but not the images taken under daytime temperatures, and the standardized image set are no longer significantly different between day and night temperatures (Fig. 3B). This suggests that the standardization method accurately maps the color space from the dark, low-quality images to the same color space as the bright, high-quality images. Changes in hue spectra have been shown to be associated to varying levels of nitrogen in sorghum and rice (Veley et al., 2017; Wang et al., 2014). Using only the nighttime images and the entire hue channel, constrained analysis of principal coordinates conditioning on time, in days, indicates that unstandardized nighttime images appearing to have relatively similar hue spectra across nitrogen treatments, whereas standardized nighttime images show clear separation (Fig. 3C).

DISCUSSION

Variation within an image-based dataset can inhibit the accurate measurement of shape-based attributes. Additionally, color profiling is important for understanding plant disease, chlorophyll content, and transpiration rate (Mutka and Bart, 2015) and these phenotypes would likely be affected by image set variation similar to what we observe here. Although the image variation described here was specifically found in the Bellwether Phenotyping Facility at the Donald Danforth Plant Science Center, it serves as an example of the potential for variability in even tightly controlled imaging systems. In many cases, imaging conditions may be far less controlled, contributing to image variability and measurement error. In our case, the source of variation could potentially be removed by using a different light source such as LEDs, although these have been shown to have temperature-dependent light output as well (Narendran et al., 2007). Regardless, it is often impractical to remove all sources of variation, so having a methodical way to standardize an image set allows anyone using image analysis to reduce the variance of the numerical output. Reducing variance will increase statistical power, allowing for observation and inference of smaller effects that would otherwise not be quantifiable. While it might not be possible to include color reference cards in every image for all applications, our results suggest that it is an important consideration for plant phenotyping systems, and a key aspect of our method is that it can be used as a tool to assess whether significant variability exists by doing a relatively straightforward pilot experiment.

It is important to note that there are other methods that exist to solve the problem of segmenting objects from an image set of varying quality. Some methods include Bayes classifiers, neural networks, and autothresholders (Yogamangalam and Karthikeyan, 2013). All object segmentation methods aim to produce a binary image wherein all object pixels are denoted as 1 and all background are denoted 0. Autothresholders attempt to alleviate misclassification by varying the value of the threshold depending on other characteristics of the image. Neural networks and Bayes classifiers work in a similar fashion by "learning" what is object and what is background and making decisions based on those models. While these methods have been proven to show better classification of object to background vs fixed-threshold methods, these methods do not change the values of the pixels, so standardized color analysis would still be problematic under variable conditions. Another downside to methods such

as these is that they require training data from across the spectrum of observed variation so that robust classification can be done. Instead, we propose that image standardization is a key first step in image analysis that will simplify the computational complexity of downstream steps, regardless of choice in segmentation method, and will improve the reliability of results and inferences.

CONCLUSIONS

We show that using the image standardization method described here improves fixed-threshold object segmentation and creates an environment where detailed color measurements are possible. Improving segmentation means shape measurements of the object are more accurate to physical, ground truth measurements. This standardization method also changes the values of the pixels of in the image, and we show it accurately maps the color space from a given image to a high-quality reference image. Finally, this method supplies a measure of deviation from an image to a reference and can be used to identify sources of variance in an image set. Future research will determine the limits of this method of standardization but it is not likely to be limited to high-throughput plant phenotyping and will be applicable to many image-based experiments.

AVAILABILITY

This method is available for anyone to use and comes in two forms. The first being C++ source code that is supplied with this paper and must be compiled against OpenCV (only tested against version 3.1), and second being a submodule that comes with PlantCV (versions >3.0.dev1). PlantCV is available on GitHub at https://github.com/danforthcenter/plantcv. PlantCV v3.0.dev1 is archived on Zenodo at https://doi.org/10.5281/zenodo.1296371. Original images are available to download at https://plantcv.danforthcenter.org/pages/data.html. R scripts and corresponding data files are provided in supplemental and can be used to reproduce all graphs and statistics.

AUTHOR CONTRIBUTIONS

Jeffrey C.Berry wrote the paper, analyzed the data, prepared figures and/or tables, reviewed drafts of the paper, conceived and designed the experiments, contributed reagents/materials/analysis tools. Noah Fahlgren reviewed drafts of the paper, conceived and designed the experiments,contributed reagents/materials/analysis tools. Alexandria A. Pokorny contributed reagents/materials/analysis tools, reviewed drafts of the paper. Rebecca Bart reviewed drafts of the paper, conceived and designed the experiments,contributed reagents/materials/analysis tools. Kira M. Veley reviewed drafts of the paper, conceived and designed the experiments, prepared figures and/or tables, contributed reagents/materials/analysis tools.

ACKNOWLEDGEMENTS

REFERENCES

Acosta-Gamboa, L. M., Liu, S., Langley, E., Campbell, Z., Castro-Guerrero, N., Mendoza-Cozatl, D., and Lorence, A. (2016). Moderate to severe water limitation differentially affects the phenome and ionome of *Arabidopsis*. *Funct. Plant Biol.*

Bleeker, N. and Veenstra, W. (1990). The performance of four foot fluorescent lamps as a function of ambient temperature on 60hz and high frequency ballasts. In *Annual Illuminating Engineering Society Conference, Paper A*, volume 45.

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., and Klukas, C. (2014). Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *Plant Cell*, 26(12):4636–4655.

Fahlgren, N., Feldman, M., Gehan, M. A., Wilson, M. S., Shyu, C., Bryant, D. W., Hill, S. T., McEntee, C. J., Warnasooriya, S. N., Kumar, I., Ficor, T., Turnipseed, S., Gilbert, K. B., Brutnell, T. P., Carrington, J. C., Mockler, T. C., and Baxter, I. (2015a). A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in setaria. *Mol. Plant*, 8(10):1520–1535.

Fahlgren, N., Gehan, M. A., and Baxter, I. (2015b). Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Curr. Opin. Plant Biol.*, 24:93–99.

Feldman, M. J., Paul, R. E., Banan, D., Barrett, J. F., Sebastian, J., Yee, M.-C., Jiang, H., Lipka, A. E., Brutnell, T. P., Dinneny, J. R., Leakey, A. D. B., and Baxter, I. (2017). Time dependent genetic analysis links field and controlled environment phenotypes in the model C4 grass setaria. *PLoS Genet.*, 13(6):e1006841.

Finlayson, G. D., Mackiewicz, M., and Hurlbert, A. (2015). Color correction using root-polynomial regression. *IEEE Trans. Image Process.*, 24(5):1460–1470.

Furbank, R. T. and Tester, M. (2011). Phenomics–technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.*, 16(12):635–644.

Gong, H., Finlayson, G., and Fisher, R. (2016). Recoding color transfer as a color homography. In *Procedings of the British Machine Vision Conference 2016*, pages 17.1–17.11. British Machine Vision Association.

Gunturk, B. K., Glotzbach, J., Altunbasak, Y., Schafer, R. W., and Mersereau, R. M. (2005). Demosaicking: color filter array interpolation. *IEEE Signal Process. Mag.*, 22(1):44–54.

Honsdorf, N., March, T. J., Berger, B., Tester, M., and Pillen, K. (2014). High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. *PLoS One*, 9(5):e97047.

Ilie, A. and Welch, G. (2005). Ensuring color consistency across multiple cameras. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, pages 1268–1275 Vol. 2. IEEE.

Liang, Z., Pandey, P., Stoerger, V., Xu, Y., Qiu, Y., Ge, Y., and Schnable, J. C. (2018). Conventional and hyperspectral time-series imaging of maize lines widely used in field trials. *Gigascience*, 7(2):1–11.

Livingstone, M. and Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, 240(4853):740–749.

Moroney, N., Fairchild, M. D., Hunt, R. W. G., Li, C., Luo, M. R., and Newman, T. (2002). The CIECAM02 color appearance model. *Color and Imaging Conference*, 2002(1):23–27.

Mutka, A. M. and Bart, R. S. (2015). Image-based phenotyping of plant disease symptoms. *Frontiers in Plant Science*, 5.

Narendran, N., Gu, Y., Jayasinghe, L., Freyssinier, J. P., and Zhu, Y. (2007). Long-term performance of white LEDs and systems. In *Proceeding of First International Conference on White LEDs and Solid State Lighting*, pages 174–179. pdfs.semanticscholar.org.

Neilson, E. H., Edwards, A. M., Blomstedt, C. K., Berger, B., Møller, B. L., and Gleadow, R. M. (2015). Utilization of a high-throughput shoot imaging system to examine the dynamic phenotypic responses of a C4 cereal crop plant to nitrogen and water deficiency over time. *J. Exp. Bot.*, 66(7):1817–1832.

Tovar, J. C., Hoyer, J. S., Lin, A., Tielking, A., Callen, S. T., Elizabeth Castillo, S., Miller, M., Tessman, M., Fahlgren, N., Carrington, J. C., Nusinow, D. A., and Gehan, M. A. (2018). Raspberry pi-powered imaging for plant phenotyping. *Appl. Plant Sci.*, 6(3):e1031.

Veley, K. M., Berry, J. C., Fentress, S. J., Schachtman, D. P., Baxter, I., and Bart, R. (2017). High-throughput profiling and analysis of plant responses over time to abiotic stress. *Plant Direct*, 1(4):e00023.

Wang, X. and Zhang, D. (2010). An optimized tongue image color correction scheme. *IEEE Trans. Inf. Technol. Biomed.*, 14(6):1355–1364.

Wang, Y., Wang, D., Shi, P., and Omasa, K. (2014). Estimating rice chlorophyll content and leaf nitrogen concentration with a digital still color camera under natural light. *Plant Methods*, 10:36.

Yogamangalam, R. and Karthikeyan, B. (2013). Segmentation techniques comparison in image processing. *International Journal of Engineering and Technology*, 5:307–313.
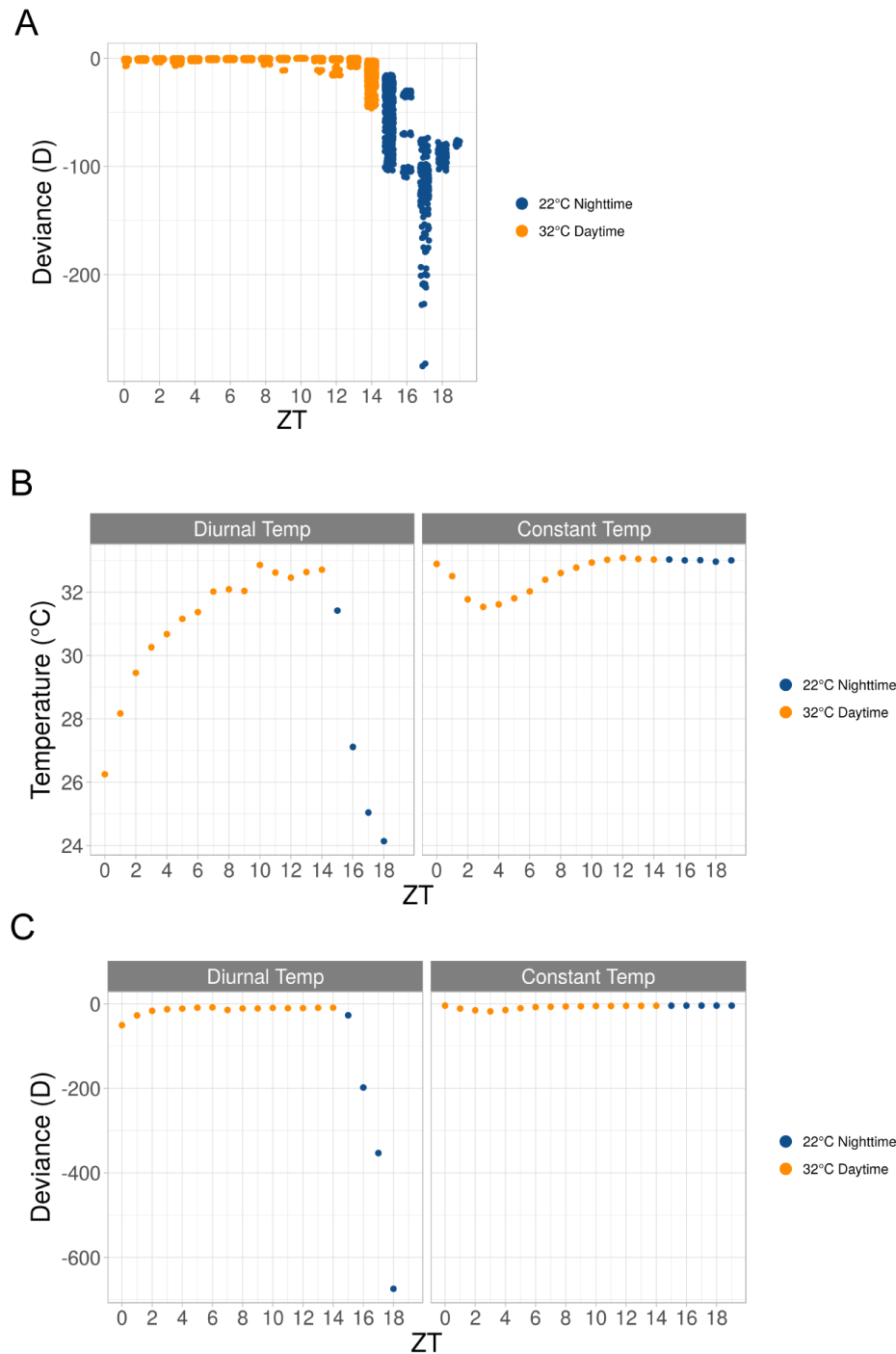
Figure 1. Effect of temperature on light intensity and image variability (A) Graph of calculated deviance (y-axis) within the plant image set and the hour the image was taken in ZT time (x-axis). Each dot represents a single image in the plant image set. The deviance from reference was calculated in Eq 7. (B) Graph of temperature readings (y-axis) taken at hours at which imaging was done in ZT time (x-axis) in either constant (right) or diurnal (left) conditions. (C) Graph of calculated deviance (y-axis) within the "temperature image set" and the hour the image was taken in ZT time (x-axis) under either constant (right) or diurnal (left) conditions. Each dot represents a mean of two images in the "temperature image set." The deviance from reference was calculated in Eq 7. (A–C). Daytime ZT 0-14 = hr 0700-2100 and nighttime ZT 14-24 = hr 2100-0700 (A-C). Temperature settings at particular times of day are indicated, either 22°C (blue) or 32°C (gold).
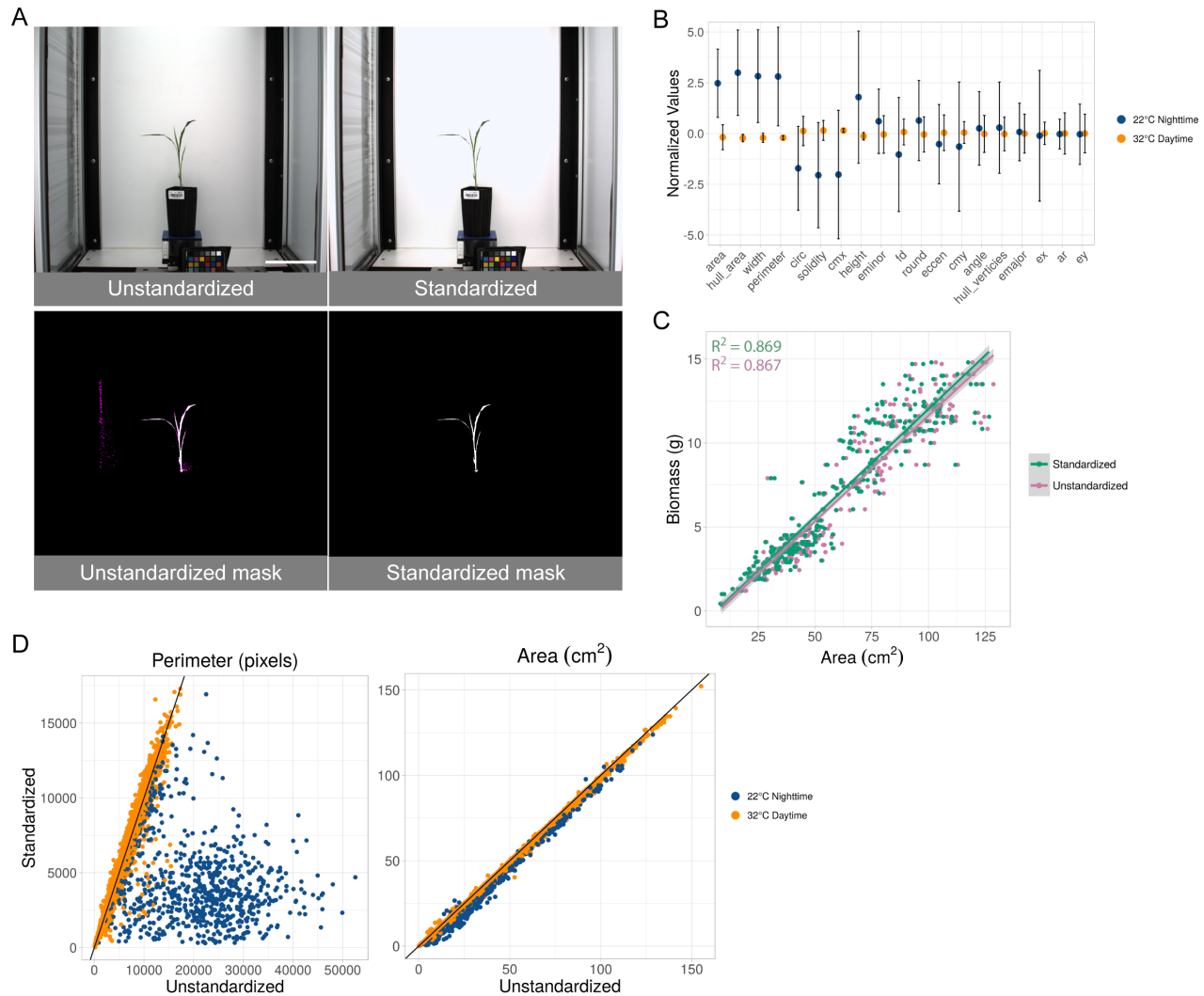
Figure 2. Standardization effects on shape measurements (A) Example of the effect of standardization. Upper left: original unedited image. Upper right: same image after applying standardization. Bottom row: Plant-isolation masks generated without (left) and with (right) standardization. The magenta coloring in the unstandardized mask indicates pixels that are not contained in the standardized mask. Scale bar = 15 cm. (B) Effect size (y-axis) of each of the shapes (x-axis) before and after standardization grouped by experimental day and night. Values of each shape per image are normalized by subtracting the mean and dividing by the variance. All shapes except ellipse major axis, ellipse center x-coordinate, aspect ratio, and ellipse center y-coordinate are significantly different between daytime and nighttime temperatures (Unequal variance t-test, p-value ¡ 0.05). Shapes are sorted by increasing p-values and are: area, convex hull area, width, perimeter, circularity, solidity, center of mass x-coordinate, height, ellipse minor axis, fractal dimension, roundness, eccentricity, center of mass y-coordinate, ellipse angle, convex hull vertices, ellipse minor axis, ellipse center x-coordinate, aspect ratio, and ellipse center y-coordinate. (C) Association of biomass (y-axis) to area (x-axis) before and after standardization. Every point is an image taken on the last day within the plant image set for which weight was recorded. Linear fit and R2 is displayed for each condition. (D) Effect of standardization on perimeter (left) and area (right). Every point is an image in the entire plant image set. Displayed are the measured perimeter and area from before standardization (x-axis) and after standardization (y-axis). Temperature settings at particular times of day are indicated, either 22ºC (blue) or 32ºC (gold). Black line indicates y = x.
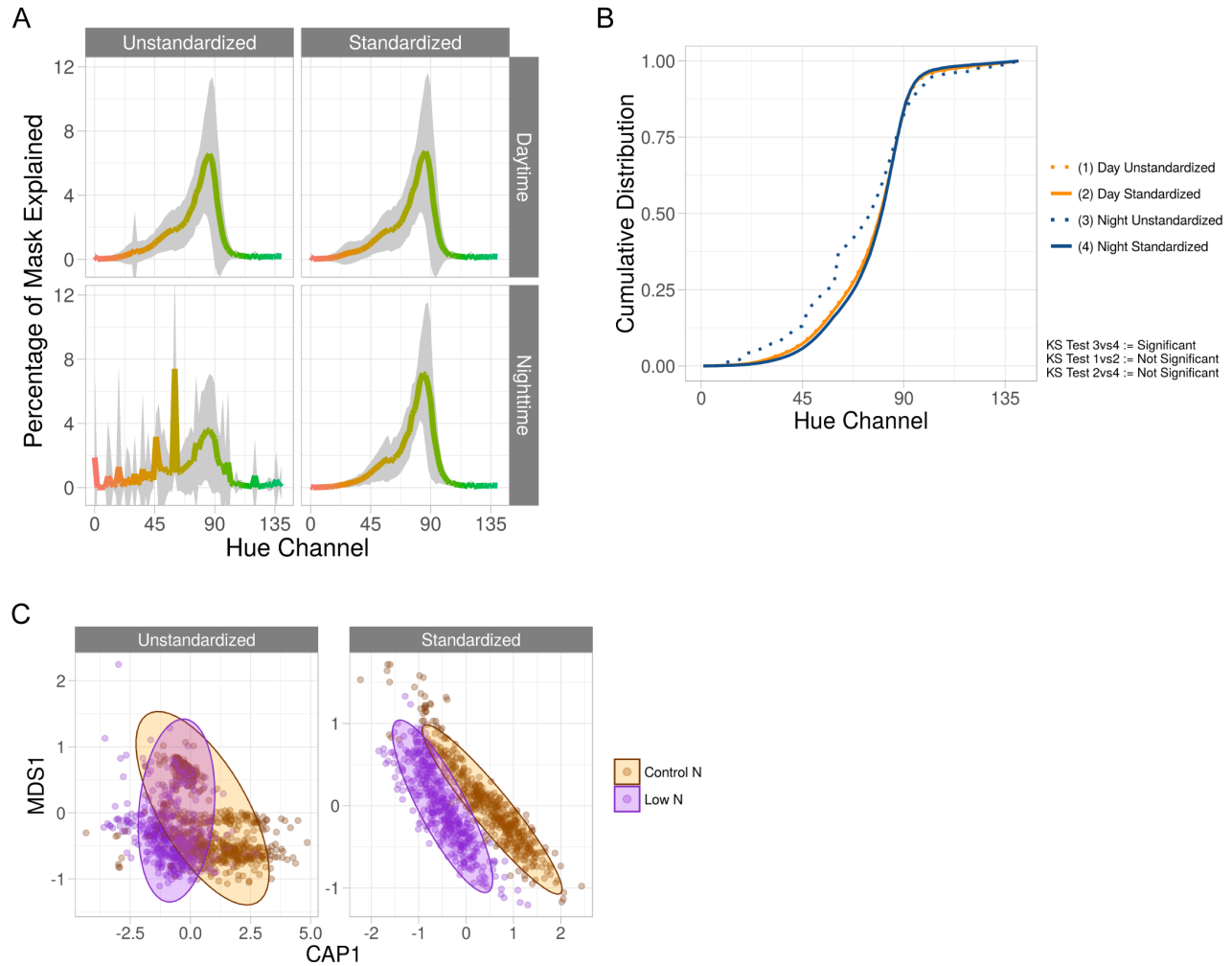
Figure 3. Standardization effects on color measurements (A) Average hue histograms of plants 19 days after planting. Shown are the average histograms of unstandardized daytime and nighttime images as well as standardized daytime and nighttime images. For each facet, the x-axis is the hue channel from 0 to 140 degrees and the y-axis is the percentage of the mask at a particular degree. The color of each degree is shown as the color of the line. Gray areas are the 95% confidence interval. (B) Using the same histograms as in (A), cumulative distributions of each are shown on the y-axis, and the hue range from 0 to 140 are on the x-axis. As the legend indicates, each line is given a number and the Kolmogorov-Smirnov test as described in Eqs 8, 9, and 10 are displayed using the numbers that are assigned to each condition. (C) With only nighttime images from the plant image set, constrained analysis of principal coordinates conditioning on time, in days, for both unstandardized and standardized images colored by nitrogen treatment. Constrained axis 1 (CAP1, x-axis) is conditioned on time while multidimensional scaled axis 1 (MDS1, y-axis) is not conditioned on any variables.