Easy Leaf Area: Automated Digital Image Analysis for Rapid and Accurate Measurement of Leaf Area

Author(s): Hsien Ming Easlon and Arnold J. Bloom
Source: Applications in Plant Sciences, 2(7)
Published By: Botanical Society of America
https://doi.org/10.3732/apps.1400033
URL: http://www.bioone.org/doi/full/10.3732/apps.1400033

Applications in Plant Sciences 2014 2(7): 1400033

Software Note

Easy Leaf Area: Automated digital image analysis for rapid and accurate measurement of leaf area[1]

Hsien Ming Easlon[2,3] and Arnold J. Bloom[2]

[2]Department of Plant Sciences, University of California, One Shields Avenue, Davis, California 95616 USA

- *Premise of the study:* Measurement of leaf areas from digital photographs has traditionally required significant user input unless backgrounds are carefully masked. Easy Leaf Area was developed to batch process hundreds of *Arabidopsis* rosette images in minutes, removing background artifacts and saving results to a spreadsheet-ready CSV file.
- *Methods and Results:* Easy Leaf Area uses the color ratios of each pixel to distinguish leaves and calibration areas from their background and compares leaf pixel counts to a red calibration area to eliminate the need for camera distance calculations or manual ruler scale measurement that other software methods typically require. Leaf areas estimated by this software from images taken with a camera phone were more accurate than ImageJ estimates from flatbed scanner images.
- *Conclusions:* Easy Leaf Area provides an easy-to-use method for rapid measurement of leaf area and nondestructive estimation of canopy area from digital images.

Key words: *Arabidopsis*; digital images; leaf area; Python.

Accurate, rapid, and nondestructive leaf area estimates are critical in many plant physiological and ecological experiments. Now-ubiquitous digital scanners and cameras, in conjunction with digital image processing software, have largely replaced older methods using light obstruction to estimate leaf area. ImageJ, the most common software used for leaf area measurement, uses a threshold-based pixel count measurement to calculate leaf area (Orsini et al., 2010; Warman et al., 2011; Juneau and Tarasoff, 2012; Carins Murphy et al., 2012; Schneider et al., 2012; Easlon et al., 2014). ImageJ, however, can require significant user input and often has difficulty in distinguishing leaves from their background using thresholding alone (Davidson, 2011). Physical masking of soil using paper collars before photographing leaves or software removal of background from images (using, e.g., GNU Image Manipulation Program; Kimball and Mattis, 2012) can remove background artifacts from images before ImageJ analysis, but these approaches add considerable processing time to leaf area measurements (Campillo et al., 2008; Warman et al., 2011; Juneau and Tarasoff, 2012).

We developed Easy Leaf Area software to rapidly estimate leaf area from *Arabidopsis* (DC.) Heynh. images against complex backgrounds with little user input. Easy Leaf Area uses a combination of thresholding, color ratios, and connected component analysis to rapidly measure leaf area in individual images in seconds or batch process hundreds of images in minutes; results are saved to a spreadsheet-ready CSV file. Each analyzed image is also saved in lossless TIFF format to provide a visual record of leaf area measurement and to facilitate additional analyses (Figs. 1C, F; 2C, E). Easy Leaf Area was written in Python (http://www.python.org/), a free and open-source programming language with image processing and mathematical tools, and is easy to modify to suit specific experimental requirements; e.g., a "Crop Cover" version of the program was written to facilitate measurement of projected leaf area and percent crop canopy cover.

[1]Manuscript received 3 April 2014; revision accepted 17 June 2014.

This work was supported by the National Science Foundation (IOS-1358675). The authors thank Eli Carlisle and Madeline Perez for their feedback on early versions of Easy Leaf Area, Eli Carlisle for providing tomato images, and Nicolas Cobo for providing wheat images. We would also like to thank the editor and reviewers for their useful comments.

[3]Author for correspondence: heaslon@ucdavis.edu

doi:10.3732/apps.1400033

METHODS AND RESULTS

Easy Leaf Area uses a red calibration area of known area in each image as a scale to calibrate leaf area estimates regardless of image source, eliminating the need for assessing camera distance and focal length or measuring ruler length manually (Baker et al., 1996). Total counts of green leaf pixels and red calibration pixels are used to estimate leaf area, according to: leaf area = (green pixel count) × (calibration area/red pixel count). When possible, the calibration area should be kept in the same plane as the leaves to avoid perspective distortion. Leaf area and calibration area should also be located in similar regions of the image to minimize errors from lens distortion. Errors due to camera set up and lens distortion can be quantified by analyzing area of squares in photographs of the 'distortion sheet' of green squares surrounding a red square of the same area (available for download at https://github.com/heaslon/Easy-Leaf-Area/blob/master/DistortionSheet.jpg). A camera phone (iPhone 4, Apple, Cupertino, California, USA) image of the 'distortion sheet' taken without a tripod at a camera distance of 20 cm had a mean distortion of 0.17% (standard error [SE] ± 0.006). A digital single-lens reflex (DSLR) camera (18–55-mm lens, 25-mm focal length = *f*/4; EOS Rebel T2i, Canon, Melville, New York, USA) image of the 'distortion sheet' at a camera distance of 30 cm had a mean distortion of −2.94% (SE ± 0.008) due to significant barrel distortion. Alternatively, destructively harvested leaves can be scanned on a flatbed scanner to eliminate leaf overlap and minimize perspective and lens distortions. Scanner images (MFC-J425w, Brother International, Bridgewater, New Jersey, USA) had a mean distortion of 0.02% (SE ± 0.003). Leaf area analyses typically rely on thresholding of either grayscale images or the blue channel of RGB (red, green, and blue) images to distinguish leaf and calibration areas from their background (O'Neal et al., 2002; Bylesjo et al., 2008; Davidson, 2011). Easy Leaf Area uses

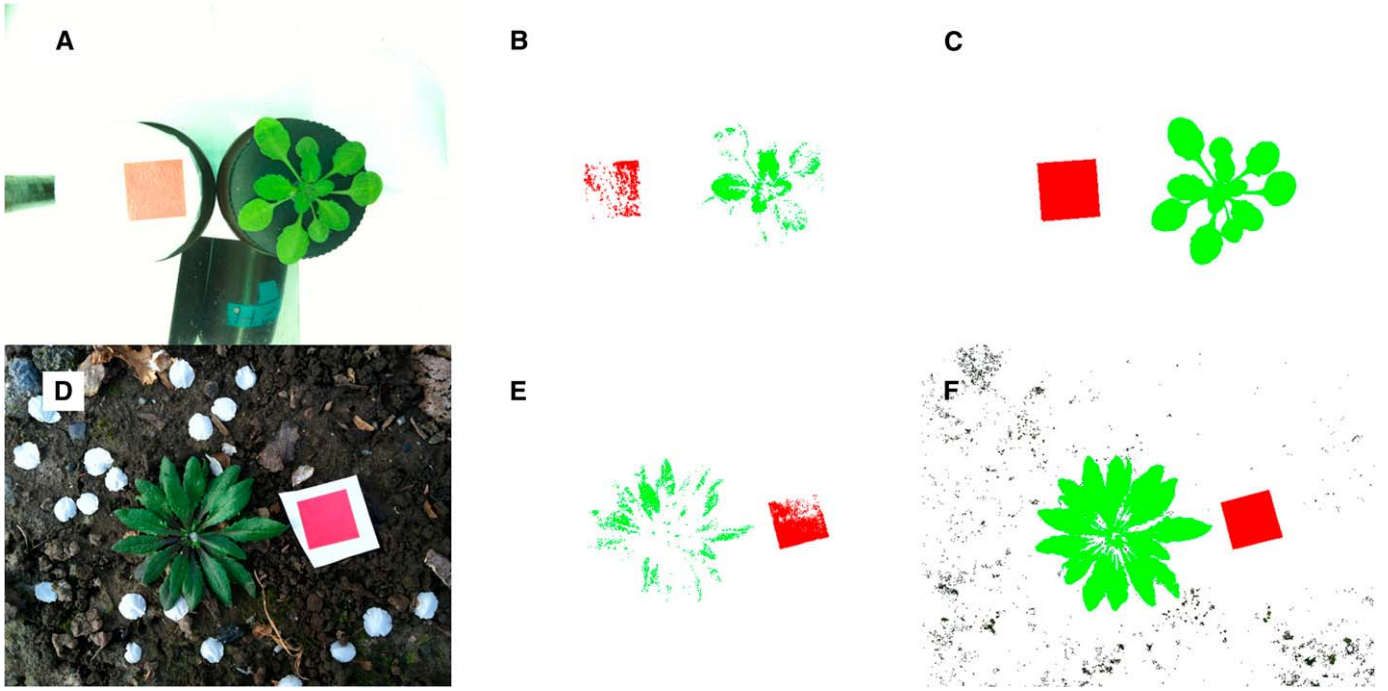

Fig. 1. Raw and processed photographs of *Arabidopsis*. Unprocessed images (A, D), images after greenest and reddest pixel selection (B, E), and images after final automated processing (C, F) with the delete background option selected. Areas recolored green were identified as leaves and areas recolored red were identified as calibration area. Darker nongreen components in the final image (F) fit pixel threshold and color ratio criteria, but were below the minimum component size, and so were not included in leaf area calculations.

thresholding combined with individual pixel RGB ratios to improve this process. For both green leaf pixel and red calibration pixel identification, two simple criteria are used. First, a minimal green or red threshold (i.e., a minimal green or red 8-bit RGB value [0–255]) is selected and any pixels with lower green or red values are not counted as leaf or calibration pixels. The second criteria uses ratios of green/red (G/R) and green/blue (G/B) or red/green and red/blue RGB values to determine which of the remaining pixels are leaf or calibration pixels. Pixel color ratios are similar to the modified excessive green index used in Lee and Lee (2011), but we found independent manipulation of G/R and G/B necessary for our *Arabidopsis* image set.

Easy Leaf Area uses an original algorithm based on *Arabidopsis* rosette images taken with a camera phone (iPhone 4, Apple) to automatically determine leaf area selection criteria without user input. This algorithm is derived from the relationship between the RGB values of the greenest leaf pixels compared to the optimal selection criteria for each image in a set of 50 *Arabidopsis* images of near-isogenic lines (NILs) from the NIL library described in Fletcher et al. (2013) (these NILs are based on chromosomal introgressions at quantitative trait loci for stomatal conductance or $\delta^{13}C$ from the Kas-1 accession in a Tsu-1 accession background) and naturalized *Arabidopsis* growing on the University of California, Davis, campus. G/R and G/B ratios for pixels of *Arabidopsis* leaves photographed under a variety of lighting and background settings were extracted with a modified version of Easy Leaf Area. Optimal selection criteria were determined from the 20 lowest green (G) values, G/R values, and G/B values of leaves in each image. The greenest leaf pixels in each image were determined from initial criteria of 75 minimum green (G), 1.8 green/blue ratio (G/B), and 2.0 green/red ratio (G/R). If the above initial criteria identified less than 200 leaf pixels, G/R and G/B were iteratively reduced by 6% until more than 200 leaf pixels were identified (Fig. 1B, E). There were strong correlations among greenest leaf pixel means and optimal selection criteria means for minimum G threshold ($R^2 = 0.899$, $p < 0.001$), G/R ($R^2 = 0.883$, $p < 0.001$), and G/B ($R^2 = 0.776$, $p < 0.001$). The algorithm uses linear regressions of these relationships to estimate optimal minimum G threshold, G/R ratio, and G/B ratio from the 200+ greenest leaf pixels in an image (Fig. 1B, E). For our *Arabidopsis* image set, the algorithm uses the following equations to calculate automated selection criteria:

$$\text{green threshold G} = 1.223 \times (\text{greenest mean G}) - 111$$

$$\text{G/R} = 0.360 \times (\text{greenest mean G/R}) + 0.589$$

$$\text{G/B} = 0.334 \times (\text{greenest mean G/B}) + 0.534$$

The same process was used to automatically calculate red calibration area selection criteria. The exact equations used to calculate automated selection criteria for calibration area are available in the Python code (https://github.com/heaslon/Easy-Leaf-Area). The accuracy of the automatic algorithm can be visually assessed for any leaf image. Pixels identified as leaf area or calibration area are recolored pure green or red for visual confirmation of leaf and calibration area identification; the background pixels can also be deleted for easier visual confirmation (Fig. 1B). For images that do not conform to the *Arabidopsis* automatic algorithm, manual adjustment of selection criteria using software sliders can be used to optimize selection criteria. These manual settings and the RGB values of the greenest leaf pixels can be saved to a new calibration file to calibrate the algorithm for an image set. During batch or individual image processing, pixel counts and leaf areas are output along with recolored images saved in lossless TIFF format to provide a record of leaf area measurement and to facilitate additional analyses (Figs. 1C, F; 2C, E).

The above method can result in many small groups of background pixels to be misidentified as leaves, especially in unmasked images of leaves with soil in the background (Figs. 1D, 2A), but these small groups of background pixels can be filtered prior to area calculation through connected component analysis (Figs. 1F, 2C). Connected component analysis identifies and labels connected leaf pixels as separate components. Small, nonleaf components can be filtered out if they are smaller than a user-selected minimum leaf size. Individual components can also be labeled with pixel counts if the area of multiple leaf components in a single image is desired.

A Windows executable "ela.exe" for automated leaf area measurement was built using PyInstaller (http://www.pyinstaller.org/) and is available for download with required supporting files at https://github.com/heaslon/Easy-Leaf-Area/blob/master/EasyLeafArea.zip (click on 'view the full file'). A Windows executable "elaCanopy.exe" for crop canopy cover and projected leaf

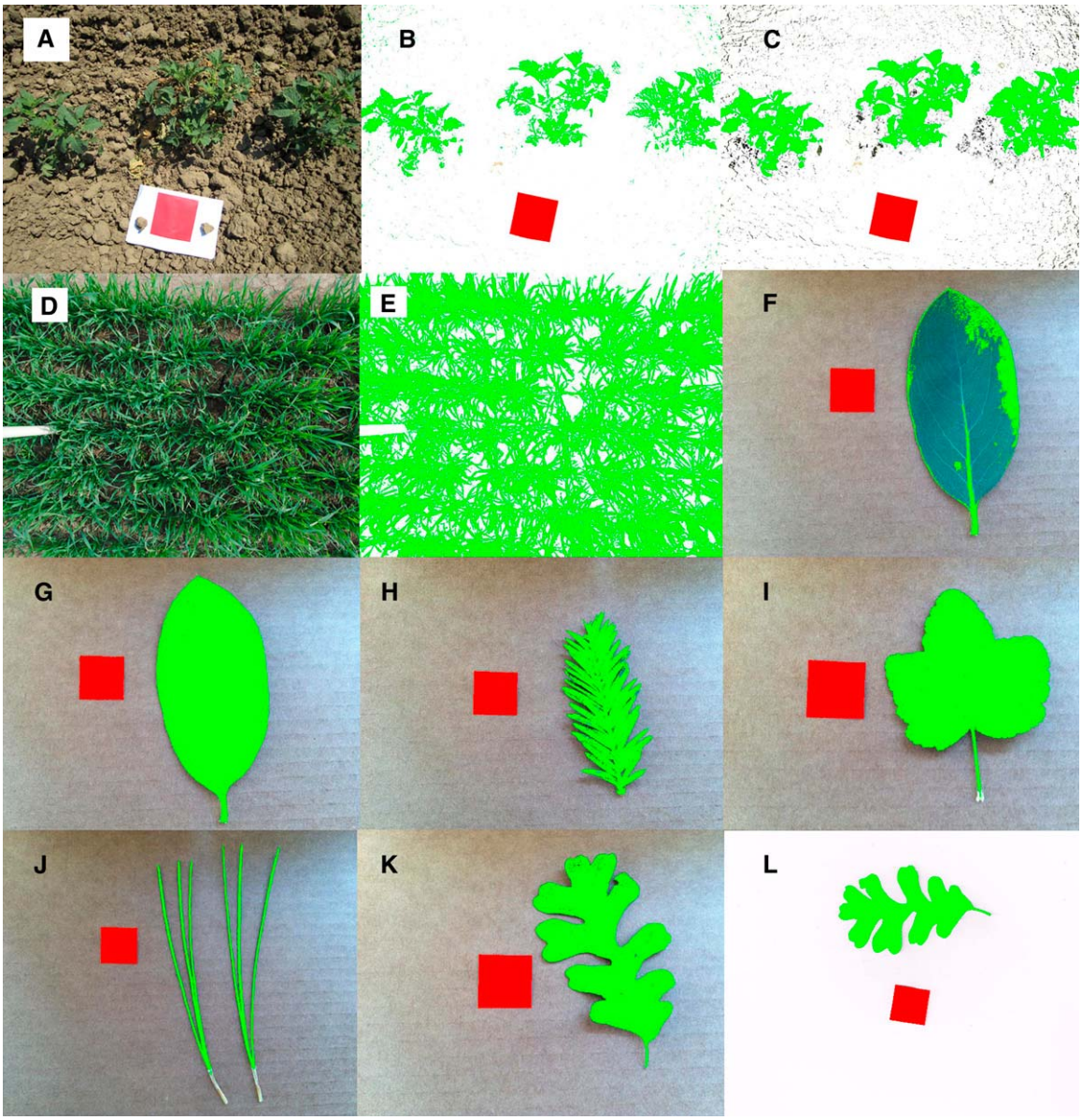

Fig. 2. Raw and processed photographs (A–K) and scans (L) of *Solanum lycopersicum* (A, B, C), *Triticum aestivum* (D, E), *Dendromecon harfordii* (F, G), *Sequoia sempervirens* (H), *Ribes malvaceum* (I), *Pinus jeffreyi* (J), and *Quercus lobata* (K, L). Images after *Arabidopsis*-based automated processing (B, F), and images after user-calibrated automated processing (C, E, G–K). Areas recolored green were identified as leaves and areas recolored red were identified as calibration area. Darker nongreen components in C and E fit pixel threshold and color ratio criteria, but were below the minimum component size, and so were not included in leaf area or canopy cover calculations. Minimum component size analysis was not used in F–L.

area estimation and "elaScanner.exe" for simplified threshold-based leaf area measurement from scanned leaves are in the same link. These executables can be run without installing additional software. The Python code "ela.py" and modified versions are available for download at https://github.com/heaslon/Easy-Leaf-Area and require installation of Python 2.7 (http://www.python.org/), Python imaging library (http://www.pythonware.com/products/pil/), SciPy (http://scipy.org/), and NumPy (http://www.numpy.org/). Instructions are available for download at https://github.com/heaslon/Easy-Leaf-Area/blob/master/ELAinstructions.txt. A video demonstration of the software is available at http://goo.gl/zZaonf. Printable calibration areas and more sample *Arabidopsis* images are available for download at https://github.com/heaslon/Easy-Leaf-Area.

To test the accuracy of Easy Leaf Area, it was compared with a LI-COR LI-3000 area meter with transparent belt conveyer accessory (LI-COR Biosciences, Lincoln, Nebraska, USA), and the commonly used image analysis software ImageJ (Schneider et al., 2012). All area estimates were compared using the traditional, weight-based, "paper doll" method. Outlines of 20 leaves, 0.5–18.5 cm^2 in area, were cut out of green paper and weighed on a microgram scale. A weight conversion factor for the paper was calculated from 10 rectangular cutouts of known area with uniform grammage. Actual leaf area was calculated as: leaf area = weight of leaf tracing × conversion factor. The paper tracings were measured with the LI-3000 leaf area meter, scanned with two flatbed scanners (CanoScan LiDE 20, Canon; MFC-J425w, Brother International) at 300 pixels per inch for ImageJ analysis, and photographed with 4-cm^2 red calibration area using a camera phone (iPhone 4, Apple) for analysis with Easy Leaf Area. Errors were calculated as: error = (estimated area − actual area)/actual area × 100. Easy Leaf Area estimates from iPhone images were essentially identical to weight-based estimates of leaf area, with a mean error of 1.18% (SE ± 0.62). Easy Leaf Area estimates from the DSLR camera (18–55-mm lens, 25-mm focal length = *f*/4; EOS Rebel T2i, Canon) images had a mean error of −4.89% (SE ± 0.88) due to significant barrel distortion. CanoScan images analyzed with ImageJ had a mean error of 7.21% (SE ± 0.55). The overestimation of leaf area in CanoScan images was caused by shadows generated

during scanning that can be minimized using a scanner with more diffuse lighting. MFC-J425w's diffuse lit scanner images analyzed with ImageJ had a mean error of 1.67% (SE ± 0.15). Shadows can also be removed manually using GIMP prior to ImageJ analysis (Warman et al., 2011; Juneau and Tarasoff, 2012), but this extra step is time consuming and was not taken in this comparison of streamlined methods. The LI-3000 area meter with transparent belt conveyer accessory is less suitable for small leaf areas and had a mean error of −11.56% (SE ± 2.16).

We also tested the automatic algorithm on photographs of field-grown tomatoes (*Solanum lycopersicum* L.) and wheat (*Triticum aestivum* L.) and photographs and scans of detached leaves of *Dendromecon harfordii* Kellogg, *Sequoia sempervirens* (D. Don) Endl., *Ribes malvaceum* Sm., *Pinus jeffreyi* Balf., and *Quercus lobata* Née. The *Arabidopsis* algorithm identified most of the tomato and wheat leaf area, but failed to select some shaded and less-green leaves (Fig. 2B). Shading in tomato and wheat crop canopies required a low minimum G threshold and G/R and G/B near 1.0 to select all leaf area (Fig. 2C, E). Calibrating the automatic algorithm for tomato and wheat canopy images allowed the automatic algorithm to identify leaf area in all examined canopy images, but background nonleaf components were also identified. In canopy images, leaf components were much larger than background nonleaf components, making them easy to filter out using minimum leaf size (Fig. 2C). In the "Crop Cover" version of Easy Leaf Area, canopy cover is estimated using green leaf pixels and total pixels, according to: % canopy cover = 100 × (green pixel count)/(total pixel count). The *Arabidopsis* algorithm failed to identify leaf area in blue-green leaves in the taxa listed above (Fig. 2F). The automatic algorithm's conservative approach to sampling only samples the greenest leaf pixels. In *D. harfordii*, the *Arabidopsis* algorithm only identified light green petioles and was not able to identify blue-green blade area (Fig. 2F). Saving manual adjustment of selection criteria from three images of each of the taxa listed above to a calibration file allowed the automatic algorithm to correctly identify leaf area in similar images of these taxa (Fig. 2G–K). Scans of leaves using white backgrounds can easily be analyzed with the "Scanner" version of Easy Leaf Area, which utilizes grayscale threshold to eliminate white background pixels from leaf area analysis similar to typical ImageJ leaf area measurement, but includes red scale measurement to simplify scanner calibration (Fig. 2L).

CONCLUSIONS

Easy Leaf Area software provides an accurate, free, and rapid tool to estimate leaf area from digital images. This tool's ability to separate leaf pixels from various backgrounds also allows it to be used for leaf area and crop canopy measurement without masking soil or relying on infrared images to distinguish leaf area from background soil. This will significantly improve rapid screening of large plant collections. The accuracy of this software was very high even with a commonly available camera phone, but care must be taken to avoid perspective and lens distortion. Images of *Arabidopsis* canopies are only approximations of leaf area and do not account for leaf overlap or leaf angle. Using scans of harvested leaves eliminates leaf overlap and minimizes camera distortion errors and is recommended for improved accuracy when destructive harvests are possible. User calibrations of the automated algorithm, like those made for tomato and wheat canopy images and detached leaves of diverse taxa, can be made to improve leaf area selection in taxa other than *Arabidopsis*. Additional selection criteria can also be added to improve selection of other color ranges for measurement of nongreen leaf area, but would require knowledge of Python. The time required using traditional methods (flatbed scanner and ImageJ analysis) is about 5 min for capturing an image and about 3 min for analyzing leaf area (Davidson, 2011). Using a digital camera and Easy Leaf Area required less than 30 s for capturing an image and less than 5 s for analyzing leaf area.

LITERATURE CITED

Baker, B., D. M. Olszyk, and D. Tingey. 1996. Digital image analysis to estimate leaf area. *Journal of Plant Physiology* 148: 530–535.

Bylesjo, M., V. Segura, R. Y. Soolanayakanahally, A. M. Rae, J. Trygg, P. Gustafsson, S. Jansson, and N. R. Street. 2008. LAMINA: A tool for rapid quantification of leaf size and shape parameters. *BMC Plant Biology* 8: 82.

Campillo, C., M. H. Prieto, C. Daza, M. J. Monino, and M. I. Garcia. 2008. Using digital images to characterize canopy coverage and light interception in a processing tomato crop. *HortScience* 43: 1780–1786.

Carins Murphy, M. R., G. J. Jordan, and T. J. Brodribb. 2012. Differential leaf expansion can enable hydraulic acclimation to sun and shade. *Plant, Cell & Environment* 35: 1407–1418.

Davidson, A. 2011 onward. 'Measuring leaf perimeter and leaf area.' Prometheus wiki. Website http://prometheuswiki.publish.csiro.au/tiki-index.php?%20page=Measuring+leaf+perimeter+and+leaf+area [accessed 23 June 2014].

Easlon, H. M., K. S. Nemali, J. H. Richards, D. T. Hanson, T. E. Juenger, and J. K. McKay. 2014. The physiological basis for genetic variation in water use efficiency and carbon isotope composition in *Arabidopsis thaliana. Photosynthesis Research* 119: 119–129.

Fletcher, R. S., J. L. Mullen, S. Yoder, W. L. Bauerle, G. Reuning, S. Sen, E. Meyer, et al. 2013. Development of a next-generation NIL library in *Arabidopsis thaliana* for dissecting complex traits. *BMC Genomics* 14: 655.

Juneau, K. J., and C. S. Tarasoff. 2012. Leaf area and water content changes after permanent and temporary storage. *PLoS ONE* 7: e42604.

Kimball, S., and P. Mattis. 2012. GNU Image Manipulation Program (GIMP). Website http://www.gimp.org/ [accessed 23 June 2014].

Lee, K. J., and B. W. Lee. 2011. Estimating canopy cover from color digital camera image of rice field. *Journal of Crop Science Biotechnology* 14: 151–155.

O'Neal, M. E., D. A. Landis, and R. Isaacs. 2002. An inexpensive, accurate method for measuring leaf area and defoliation through digital image analysis. *Journal of Economic Entomology* 95: 1190–1194.

Orsini, F., M. P. D'Urzo, G. Inan, S. Serra, D. H. Oh, M. V. Mickelbart, F. Consiglio, et al. 2010. A comparative study of salt tolerance parameters in 11 wild relatives of *Arabidopsis thaliana. Journal of Experimental Botany* 61: 3787–3798.

Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9: 671–675.

Warman, L., A. T. Moles, and W. Edwards. 2011. Not so simple after all: Searching for ecological advantages of compound leaves. *Oikos* 120: 813–821.