


# A rapid method to quantify vein density in C<sub>4</sub> plants using starch staining

Conor J. C. Simpson<sup>1</sup> | Pallavi Singh<sup>1</sup>  | Deedi E. O. Sogbohossou<sup>2</sup> |  
M. Eric Schranz<sup>2</sup> | Julian M. Hibberd<sup>1</sup> 

<sup>1</sup>Department of Plant Sciences, University of Cambridge, Cambridge, UK

<sup>2</sup>Biosystematics Group, Wageningen University, Wageningen, The Netherlands

## Correspondence

Julian M. Hibberd, Department of Plant Sciences, University of Cambridge, Downing St, Cambridge CB2 3EA, UK.  
Email: [jmh65@cam.ac.uk](mailto:jmh65@cam.ac.uk)

## Funding information

Netherlands Organization for Scientific Research; Biotechnology and Biological Sciences Research Council; European Research Council

## Abstract

C<sub>4</sub> photosynthesis has evolved multiple times in the angiosperms and typically involves alterations to the biochemistry, cell biology and development of leaves. One common modification found in C<sub>4</sub> plants compared with the ancestral C<sub>3</sub> state is an increase in vein density such that the leaf contains a larger proportion of bundle sheath cells. Recent findings indicate that there may be significant intraspecific variation in traits such as vein density in C<sub>4</sub> plants but to use such natural variation for trait-mapping, rapid phenotyping would be required. Here we report a high-throughput method to quantify vein density that leverages the bundle sheath-specific accumulation of starch found in C<sub>4</sub> species. Starch staining allowed high-contrast images to be acquired permitting image analysis with MATLAB- and Python-based programmes. The method works for dicotyledons and monocotyledons. We applied this method to *Gynandropsis gynandra* where significant variation in vein density was detected between natural accessions, and *Zea mays* where no variation was apparent in the genotypically diverse lines assessed. We anticipate this approach will be useful to map genes controlling vein density in C<sub>4</sub> species demonstrating natural variation for this trait.

## KEYWORDS

bundle sheath, C<sub>4</sub> photosynthesis, high-throughput phenotyping

## 1 | INTRODUCTION

Photosynthesis is the basis of life on Earth and central to this process is the enzyme Ribulose 1•5-Bisphosphate Carboxylase/Oxygenase (RuBisCO) that operates in the Calvin–Benson–Bassham cycle to fix CO<sub>2</sub>. However, oxygenase activity RuBisCO results in the energy-intensive photorespiratory pathway (Driever & Kromdijk, 2013) and in some environments costs of photorespiration are thought to have driven the evolution of carbon

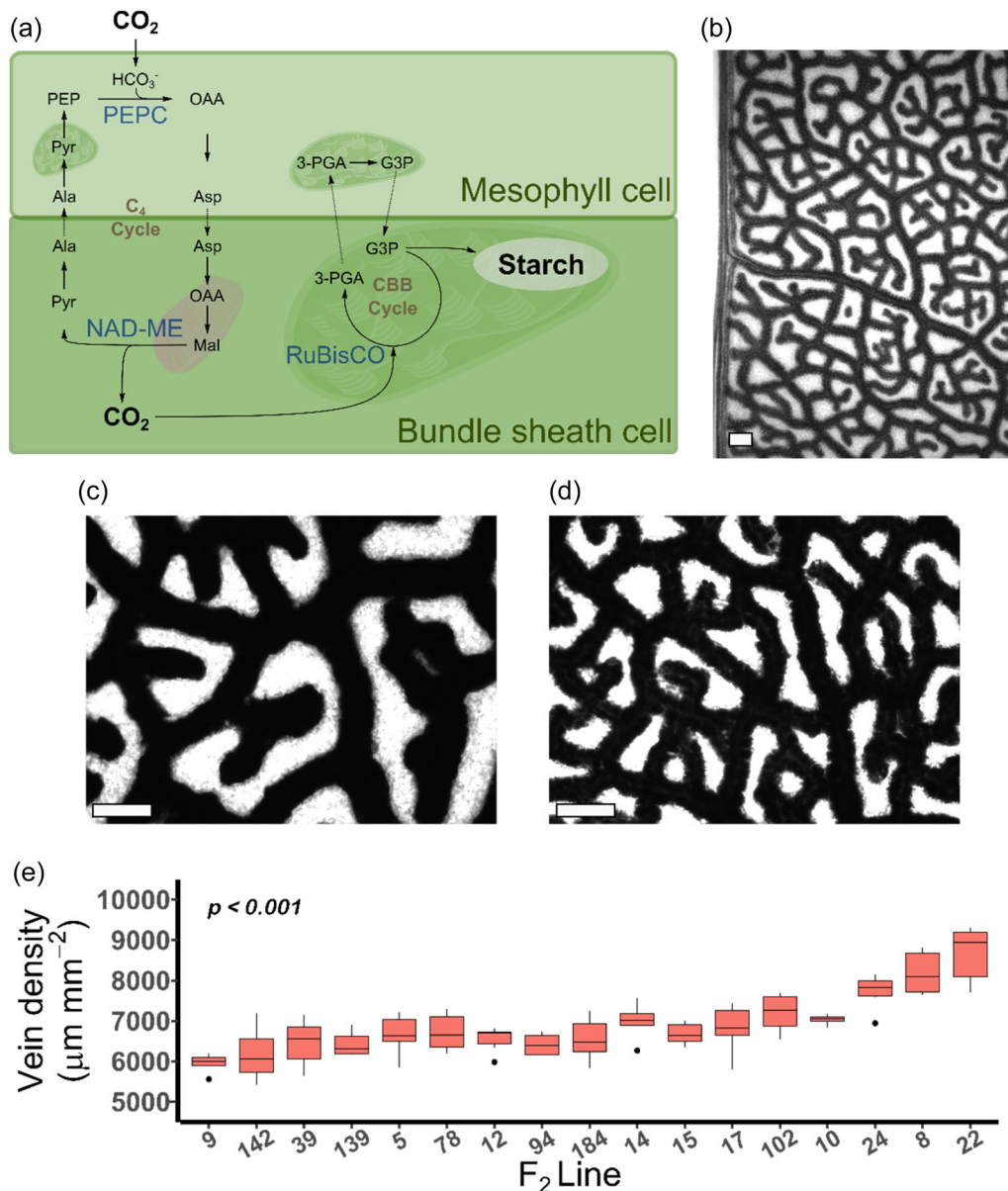
concentrating mechanisms such as C<sub>4</sub> photosynthesis and Crassulacean acid metabolism (CAM). Thus, C<sub>3</sub> photosynthesis is considered the ancestral state and despite its complexity, derived states involving carbon concentrating mechanisms are thought to have evolved repeatedly (Edwards, 2019). In the case of C<sub>4</sub> photosynthesis, current estimates are that it has evolved independently in more than 66 lineages (Sage et al., 2011). The vast majority of C<sub>4</sub> plants evolved specialized anatomy to compartmentalise photosynthesis between two cell types. Typically, this involves the

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Plant, Cell & Environment* published by John Wiley & Sons Ltd.

mesophyll being the site of initial  $\text{HCO}_3^-$  assimilation by phosphoenolpyruvate carboxylase (PEPC) and the bundle sheath being repurposed for the Calvin–Benson–Bassham cycle. Compared with the  $\text{C}_3$  state, the partitioning of photosynthesis between these two cell types is associated with changes to transcriptional, post-transcriptional and also posttranslational regulatory mechanisms (Hibberd & Covshoff, 2010; Reeves et al., 2017). Different  $\text{C}_4$  ‘subtypes’ have been defined by the predominant use of three  $\text{C}_4$  acid decarboxylases in the bundle sheath (Hatch et al., 1975).

Although there is growing support for the notion that species can modify the extent to which each  $\text{C}_4$  acid is engaged (Omoto et al., 2012; Sales et al., 2018; Sharwood et al., 2014). For some genes encoding components of the  $\text{C}_4$  cycle a detailed understanding has emerged such that changes in *cis* (Gowik et al., 2004, 2017; Nomura et al., 2000; Nomura, Higuchi, Ishida, et al., 2005; Nomura, Higuchi, Katayama, et al., 2005; Williams et al., 2016) or *trans* (Brown et al., 2011; Reyna-Llorens et al., 2018) are considered the driving forces behind alterations to gene expression.



**FIGURE 1** Starch staining in the bundle sheath of  $\text{C}_4$  *Gynandropsis gynandra* demonstrates natural variation in a diversity panel. (a) Schematic of the  $\text{C}_4$  pathway. Arrows indicate carbon flux. Enzymes involved in carboxylation and decarboxylation are shown in blue. Intracellular compartments represented schematically with mitochondria in pink and chloroplasts dark green. (b, c, d) Representative images of *Gynandropsis gynandra* leaves stained for starch at low magnification (b), higher magnification with low vein density (c) and higher magnification with high vein density (d). (e) Natural variation in vein density across 17 lines. Scale bars = 200  $\mu\text{m}$ . Six images per plant were assessed and  $p$  value calculated by ANOVA. Ala, alanine; Asp, aspartate; CBB, Calvin–Benson–Bassham cycle; G3P, Glyceraldehyde 3-phosphate; Mal, malate; NAD-ME, NAD-Malic Enzyme; OAA, oxaloacetate; PEP, phosphoenolpyruvate; PEPC, phosphoenolpyruvate carboxylase; Pyr, pyruvate; RuBisCO, Ribulose 1•5-Bisphosphate Carboxylase/Oxygenase; 3-PGA, 3-Phosphoglycerate.

In contrast, we have a more limited understanding of the genes underpinning the evolution of Kranz anatomy and how their expression has changed during the C<sub>3</sub>-C<sub>4</sub> transition (Sedelnikova et al., 2018).

Reports of significant natural variation in traits including components of Kranz anatomy (Kolbe & Cousins, 2018; Lundgren et al., 2016; Reeves et al., 2018; Yabiku & Ueno, 2017) provide an opportunity to probe the basis of this trait using quantitative genetics. For example, if it were possible to screen vein density or bundle sheath size rapidly in populations made up of genetically diverse individuals such as Multi-Advanced Generation Inter-Crossing (MAGIC) populations or natural accessions designed for Genome Wide Association Study (GWAS) it should be possible to start to link genes with Kranz anatomy. This approach would be facilitated by rapid phenotyping of these traits. Significant advances have been made in phenotyping at scale for protein activity assays (Gibon et al., 2004; Sulpice et al., 2010) or the use of robotics for data collection in the field (Virlet et al., 2017). Notably, using a phenotyping facility Guo et al. (2018) rapidly analysed 51 image-based traits in rice under drought stress. Correlating these traits with traditional drought resistance traits enabled novel genes linked to drought resistance to be identified via a GWAS. Such approaches may need support from noninvasive imaging techniques (Yang et al., 2020) and consequently, high-throughput image processing is key to alleviate the phenotyping bottleneck associated with laborious, error-prone manual image analysis (Moen et al., 2019). Image analysis using Convolutional Neural Networks (CNNs) has been used to map Quantitative Trait Loci (QTL) associated with stomata size and density in maize (*Zea mays*; Xie et al., 2021). Moreover, image segmentation tools have been developed to quantify vein density (Bühler et al., 2015; Dhondt et al., 2012; Parsons-Wingerter et al., 2014), but such automated approaches have not yet replaced manual tracing (Perez-Harguindeguy et al., 2016). Moreover, analysis of tissues such as cotyledons of *Arabidopsis thaliana* (Dhondt et al., 2012; Parsons-Wingerter et al., 2014) can require extensive manual supervision and may not be suited for tissues with higher vein density. Although, the powerful programme LeafVeinCNN (Xu et al., 2021) enables high-throughput analysis of traits including vein order, areole number, vein width and vein density it currently requires manual real-time supervision of individual images and can be computationally slow due to the number of networks in use and the number of measurements being attained per image (Xu et al., 2021). Further, to our knowledge, only one tool has been developed to assess vein structure in monocotyledons (Robil et al., 2021). As C<sub>4</sub> photosynthesis is found in both monocotyledons and dicotyledons it would be useful to have a tool that permits analysis of vein density in both clades.

Here we aimed to test whether preferential accumulation of starch in the bundle sheath of C<sub>4</sub> plants (Lunn & Furbank, 1997; Figure 1a) could be used to generate high-contrast images so that a segmentation-based method could rapidly assess vein density in these plants. We show that such an approach can be used to measure vein density in both monocotyledons and dicotyledons. Moreover,

this allowed us to develop a computational pipeline enabling high-contrast black-and-white images to be processed automatically using the image processing toolbox in MATLAB. The automated pipeline was also implemented in the open-source Python tool. Not only does the method allow rapid phenotyping but it also showed that whilst in genotypically diverse founders of a maize population (Dell'Acqua et al., 2015) there was little variation in vein density, in the C<sub>4</sub> dicotyledonous model *Gynandropsis gynandra* there is significant natural variation in vein density.

## 2 | METHODS

### 2.1 | Plant growth and leaf sampling

*G. gynandra* was grown under irrigated glasshouse conditions from April to July 2019. Temperatures were maintained at 24°C in the day and 20°C during the night, with artificial lighting maintaining a minimum light intensity of 300  $\mu\text{mol m}^{-2} \text{s}^{-1}$  and a photoperiod of 16 h light and 8 h dark. Relative humidity was 60%. Maize was grown under field conditions from April to August 2022 at NIAB. Plants were treated with pre-emptive herbicide (Stomp<sup>®</sup> Aqua) 1 day after sowing, fertilizer (ammonium nitrate) 5 days after sowing, molluscicide (SluXX) 1 month after sowing, and herbicide (Leystar<sup>®</sup>) 2 months after sowing. During the growing season, average temperature was 19.8°C during the day and 15.2°C during the night. Maximum and minimum daytime temperature was 39.9°C and 8.9°C and that during the night was 34.2°C and 2.9°C. Total precipitation was 89 mm and average relative humidity was 65%.

*G. gynandra* leaves were harvested over a 3-day period from 4-week-old plants. They were harvested at least 7 h into the photoperiod. Briefly, the youngest fully expanded leaf was selected and tissue was harvested from the central leaflet of each leaf and immediately placed into an embedding cassette submerged along with other cassettes in 3:1 100% (v/v) ethanol:acetic acid fixative solution for 4 h. The solution was then placed in 70% (v/v) ethanol for an hour at 37°C, and then left in fresh ethanol overnight. Samples were then cleared using 5% (w/v) NaOH for 3 h at 37°C before being washed and replaced in 70% (v/v) ethanol until imaging. Samples can also be stored in 70% (v/v) ethanol before clearing with NaOH. Maize leaves were harvested over a 2-week period as part of a larger field trial. All disks were taken from the most recently fully expanded leaf. Leaves were harvested in the field and kept in water until leaf disks were taken. One disk was taken per leaf and immediately submerged in 3:1 100% (v/v) ethanol:acetic acid fixative solution for 4 h before being transferred to 70% (v/v) ethanol solution and left overnight at 37°C. Samples were then cleared using 5% (w/v) NaOH for 6 h at 37°C before being washed with and replaced in 70% (v/v) ethanol until preparation for imaging.

*Setaria italica*, *Setaria viridis*, *Sorghum bicolor* and *Pennisetum glaucum* were sown directly onto the soil surface, covered with propagator lids to control moisture at 28°C in the absence of light for 1 day. After this, pots were transferred to a growth chamber with

12 h light, 12 h dark, maintaining a constant temperature of 28°C day and 20°C night, with ambient CO<sub>2</sub> concentration and a light intensity of 500  $\mu\text{mol m}^{-2} \text{s}^{-1}$ .

Before imaging, leaf disks were placed in Lugol's solution. Specifically, using forceps each disk was transferred to Lugol's solution and then removed immediately after submersion and placed into 70% (v/v) ethanol to minimize stain accumulation. If the stain did not accumulate rapidly the sample was replaced in Lugol's solution until it stain had been taken up. Excess Lugol's solution was washed off in 70% (v/v) ethanol. Each leaf disk was then mounted on a slide using water and imaged at a magnification of  $\times 100$  on an Olympus BX41 light microscope with a mounted Micropublisher 3.3 RTV camera (Q Imaging). Images were captured with Q-Capture Pro 7 software.

## 2.2 | Manual and automated analysis of vein density and statistical analysis

Veins were measured directly in ImageJ (Rasband, 2014) using the paint brush tool to quantify total length of veins in  $\mu\text{m}$ . Lengths were converted to density ( $\mu\text{m mm}^{-2}$ ) by dividing by the image's area in  $\text{mm}^{-2}$ . Table 1 summarises the important MATLAB functions and their function. In addition to this, pattern recognition was used to identify commissural veins in maize (Supporting Information: Figure S3). From a subset of 102 images, 27 were also analysed using LeafVeinCNN (Xu et al., 2021). These 27 images comprised three images per plant from nine plants. LeafVeinCNN was run using default parameters consisting of three ensemble CNNs. Appropriate threshold values were determined for each image to attain optimal vein traces through segmentation. To achieve results comparable with Starch4Kranz the output value 'VTotL' meaning total vein length in mm, was converted to  $\mu\text{m mm}^{-2}$ . To generate the Python version of Starch4Kranz, functions from the OpenCV (Cv2; Bradski, 2000), NumPy (Harris et al., 2020), Scikit-image (skimage; Van Der Walt

et al., 2014), SciPy (Jones et al., 2001), Matplotlib (Hunter, 2007), and PlantCV (Gehan et al., 2017) libraries were used. Important functions used are listed in Table 2.

All statistical analysis were carried out in R Studio (V: 4.1.3). Figures were generated using the ggpubr (Kassambara, 2020) package. To compare manual and automated quantification methods in *G. gynandra* a subset of 17 plants from a total of 199 were selected randomly. Six images were assessed per plant such that the data set was made up of 102 images. To attain a vein density measurement for each of the 17 lines, the mean of these six images was determined. We noted that using three images yielded no significant differences in plant vein densities when comparing manual (*t* test,  $p = 0.90$ ) and automated (*t* test,  $p = 0.85$ ) methods. Therefore for maize, six replicates were used for A632, F7, and W153R; five for B73, CML91, H99, HP301, and Mo17; and three for B96. The total data set of 138 images was assessed both manually and automatically. For both treatments, all lines were normally distributed (Shapiro–Wilk's test,  $p > 0.05$ ) and Bartlett's test revealed homogeneity of variance ( $p > 0.05$ ) meaning parametric analysis was carried by analysis of variance (ANOVA). Correlation analysis was carried out using Pearson's correlation coefficient. In total, a *G. gynandra* population of 199 plants, each with five or six image replicates meaning a total of 1102 images were assessed for variation.

## 2.3 | Raw script availability and code implementation

The Starch4Kranz pipeline was designed to be simple to use and can run in the background while other tasks take place. While it requires a manual check, this can be done rapidly (~5–10 s per image) as the user checks each image within a whole directory once the script has stopped running. Only three scripts are needed to run the full pipeline in MATLAB, Starch4Kranz.m contains all executable code used in the production of our pipeline, the function, blur.m contains

**TABLE 1** Main MATLAB base and Image Processing Toolbox functions used in Starch4Kranz.

Function	Use
rgb2gray	Converts Red Blue Green (RGB) image to greyscale
adapthisteq	Improves contrast of image using contrast-limited adaptive histogram equalization (Zuiderveld, 1994)
blur	Blurs the image by converting each pixel to the mean of pixel value in a submatrix of $2w + 1$ ( <a href="https://uk.mathworks.com/matlabcentral/answers/472573-write-a-function-called-blur-that-blurs-the-input-image?s_tid=mwa_osa_a">https://uk.mathworks.com/matlabcentral/answers/472573-write-a-function-called-blur-that-blurs-the-input-image?s_tid=mwa_osa_a</a> )
imbinarize	Replaces all pixels above a threshold to a value of 1 and makes all others 0. Threshold is based on Otsu's method (Otsu, 1979)
imcomplement	Complements an image
bwskel	Reduces binary objects to a 1-pixel wide curved line while conserving image structure
bwmorph	Applies a particular morphological operation to a binary image. We used the 'branchpoints' and 'endpoints' operations to create two separate masks that showed the skeleton's branch points and end points, these could be used in conjunction with the find function to pinpoint their positions and lengths
bwdistgeodesic	Calculates the distance (in pixels) between defined points in a binary image. Used here to calculate the lengths of branches, that is, length between a branch point and an end point. This value was important for the trimming step of the pipeline.

Function	Use
cv2.cvtColor	Converts Red Blue Green (RGB) image to greyscale
cv2.createCLAHE	Improves contrast of image using contrast-limited adaptive histogram equalization (Zuiderveld, 1994)
cv2.blur	Blurs image
cv2.threshold	Binarises image
cv2.bitwise_not	Complements an image
skimage.morphology.skeletonise	Reduces binary objects to a 1-pixel wide curved line while conserving image structure
plantcv.morphology.find_branch_pts	Identifies branch points of skeleton
plantcv.morphology.find_tips	Identifies end points of skeleton

**TABLE 2** Main Python functions used in Starch4Kranz; citations are in the text.

the blurring function, and running `Starch4Kranz.m` contains a short script that runs `Starch4Kranz.m` while saving results to a table easily transferable to software for further data analysis. For Python, just two scripts are required, `Starch4Kranz.py` and `running_Starch4Kranz.py`. Scripts are available at <https://github.com/plycs5/Starch4Kranz>. To run the script in MATLAB the user must have the ImageProcessingToolbox activated and in Python they must have the dependent libraries installed into their environment. There are 14 inputs the user can supply (Supporting Information: Table S1), five of which are necessary; filename, trim\_factor (see Results), pixel\_length\_um (the image scale), y\_pixels, and x\_pixels (image dimensions), while the other nine can be altered to optimise performance.

### 3 | RESULTS AND DISCUSSION

#### 3.1 | Natural variation for vein density in *C<sub>4</sub> G. gynandra* and an automated image processing tool

Staining leaves of *G. gynandra* that predominantly uses NAD-dependent Malic Enzyme as the  $C_4$  acid decarboxylase clearly indicated preferential accumulation of starch in the bundle sheath (Figure 1b). We, therefore, subjected an  $F_2$  population of *G. gynandra* comprising 199 lines derived from a cross between parents differing in vein density to starch staining (Reeves et al., 2018). Initially, a random subset of 17 plants were stained and this indicated significant natural variation in vein density amongst the  $F_2$  plants ( $p < 0.001$ ; Figure 1c–e). We reasoned that this variance in vein density between accessions provided a resource to develop and ground-truth a semiautomated procedure to quantify vein density.

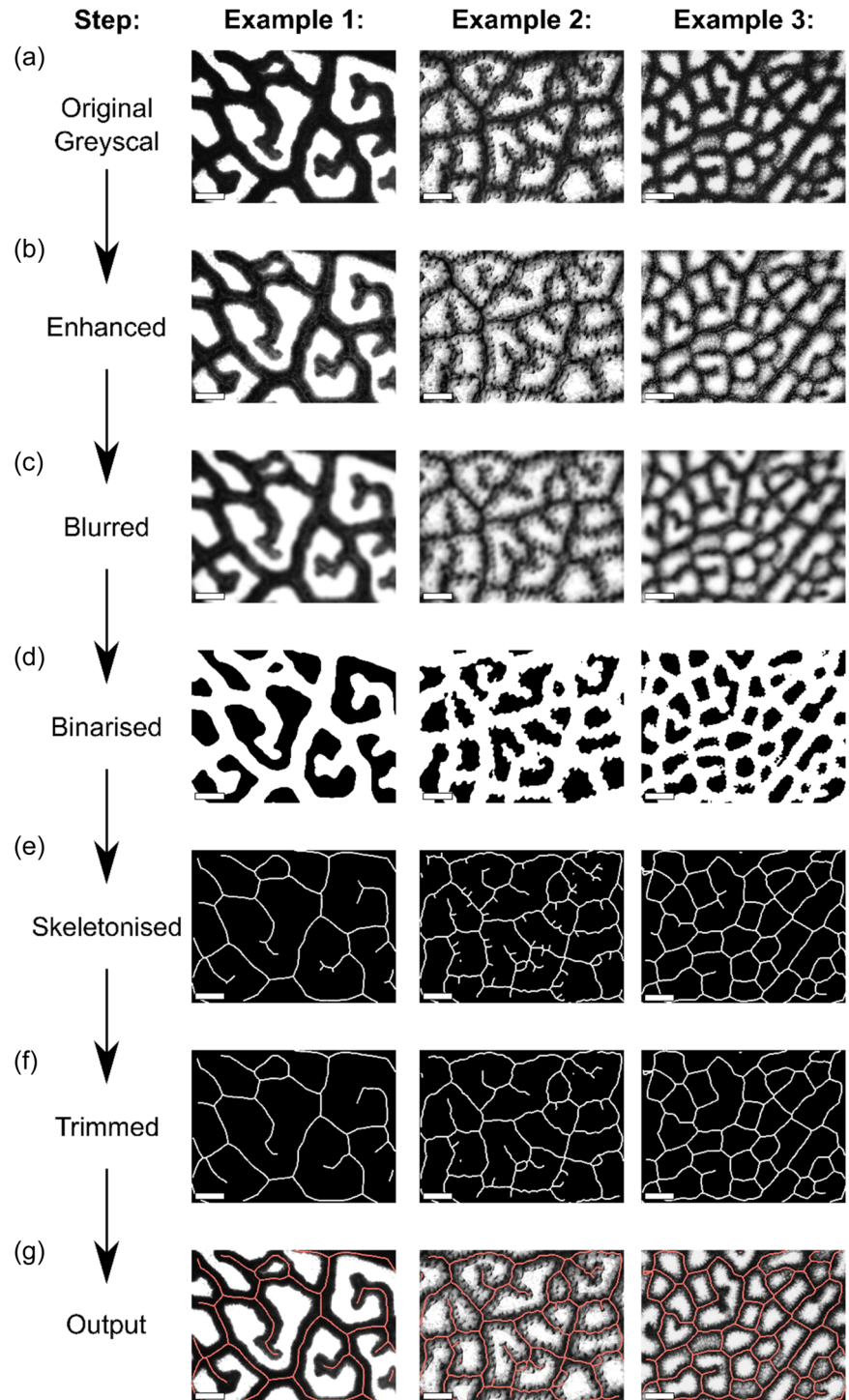
Using MATLAB's Image Processing Toolbox (2019a) an automated pipeline, hereafter Starch4Kranz, was developed to reliably quantify bundle sheath density. This pipeline involved converting each image to greyscale (Figure 2a), enhancing the contrast to improve differences between bundle sheath strands and the background (Figure 2b) and then blurring such that a given pixel is converted to the mean pixel intensity in a submatrix of  $2w + 1$ ,

where  $w$  is 20. Thus, for each pixel, its value is converted to the average of the neighbouring submatrix of  $20 \times 20$  pixels. This smoothing resulted in pixels in each bundle sheath strand having similar values, and the same was true for background pixels (Figure 2c). Images are then binarised (Figure 2d) such that stained regions appear white (with a pixel value of 1) and the background is black (with a pixel value of 0). Skeletonisation converts all pixels with a value of 1 to curved lines a single pixel in width (Figure 2e). Lastly, these skeletonised images are trimmed such that superfluous pixels are removed (Figure 2f, Supporting Information: Figure S1). The number of pixels trimmed is defined by the user as the trim\_factor, but if the programme detects a significantly larger number of superfluous pixels compared with expected vein density then the trim\_factor can be calculated automatically based on the mode length of connected superfluous pixels. At this stage, bundle sheath strands are represented by a line that is one pixel thick and so the total number of pixels provides an estimate of vein length. Vein density can then be determined from the total number of pixels and the image size.

Starch4Kranz processes all images in a given directory and saves results into a table containing image name and vein density. After each image is processed a window appears containing the vein trace superimposed on the original image in greyscale (Figure 2g). Compared with approximately 11 h required to manually trace around 102 images (Figure 1e) implementing the Starch4Kranz pipeline for the same samples took 15 min.

We further implemented a Python version of Starch4Kranz so it is available as open source. The blurring method implemented by the Python version is based on openCV's blur function (Bradski, 2000) in which the kernel size was defined by the given blur\_factor. To display the image, the libraries Matplotlib (Hunter, 2007), and PlantCV (Gehan et al., 2017) are required. Due to working with a large population, we adopted a sampling method that enabled rapid acquisition and concurrent fixing and clearing of large numbers of samples that could be acquired simply (see Section 2). However, due to the nature of the programme any method generating high-contrast images of leaf veins should be compatible with Starch4Kranz.

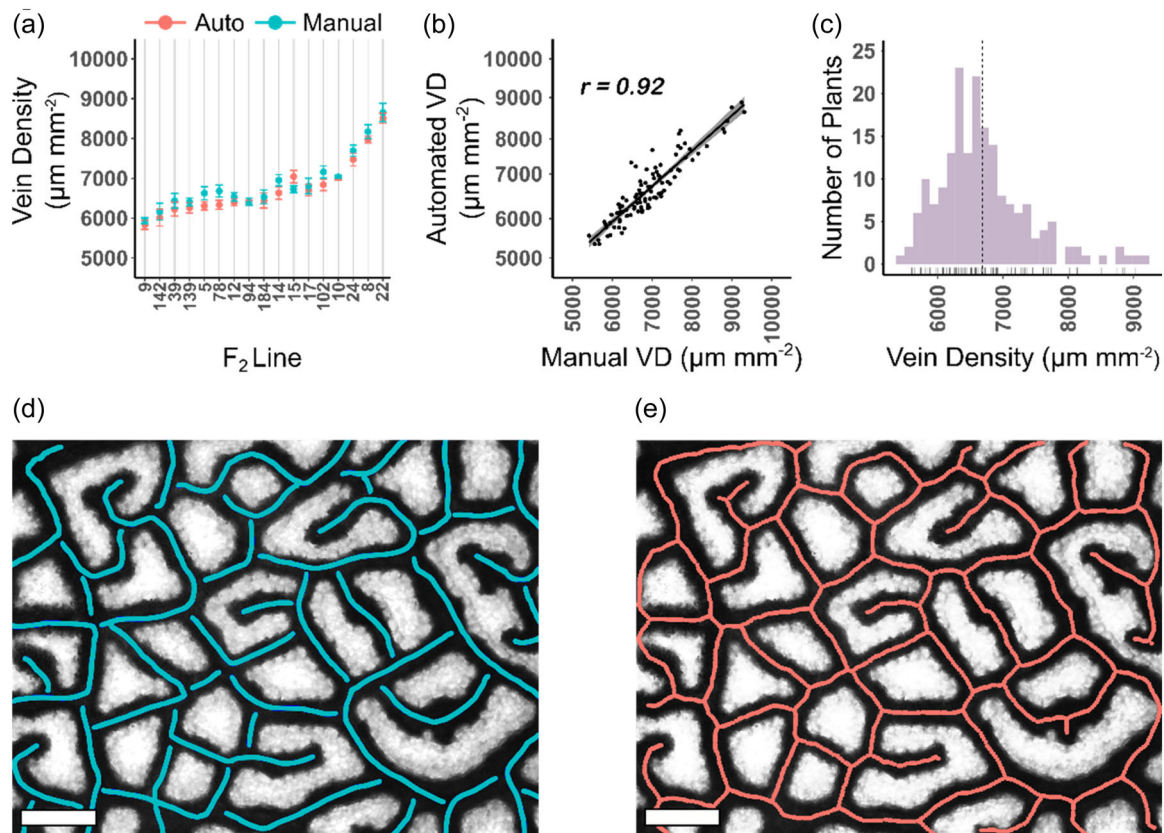
**FIGURE 2** Summary of critical automation steps. (a) A greyscale image is used as input and subjected to contrast enhancement (b) followed by blurring (c), binarization (d), skeletonization (e), and trimming (f), before producing the output for the user to check (g). Scale bars = 200  $\mu\text{m}$ . Skeletons have been thickened to ensure visibility. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



### 3.2 | Automatic quantification of high-contrast images for vein density

The Starch4Kranz pipeline was next implemented on data from the 17  $F_2$  plants for which vein density had been determined manually (Figure 3). From 102 images, the same pattern of variation in vein density was observed (Figure 3a) and estimates from manual and automated pipelines were highly correlated ( $r = 0.92$ ; Figure 3b). This was also the case for the Python version, which correlated strongly

with manual assessments ( $r = 0.91$ ; Supporting Information: Figure S1A), and with the MATLAB version ( $r = 0.97$ ; Supporting Information: Figure S1B). When Starch4Kranz was applied to 1101 images, 98.5% passed the manual check. Of the 17 images that did not, manual inspection indicated they were poorly stained or the images were of poor quality. To reduce chances of poor staining, sampling late in the photoperiod is helpful—we achieved maximal staining 9 h into a 16 h photoperiod (Supporting Information: Figure S3). We compared Starch4Kranz with existing software. With PhenoVein

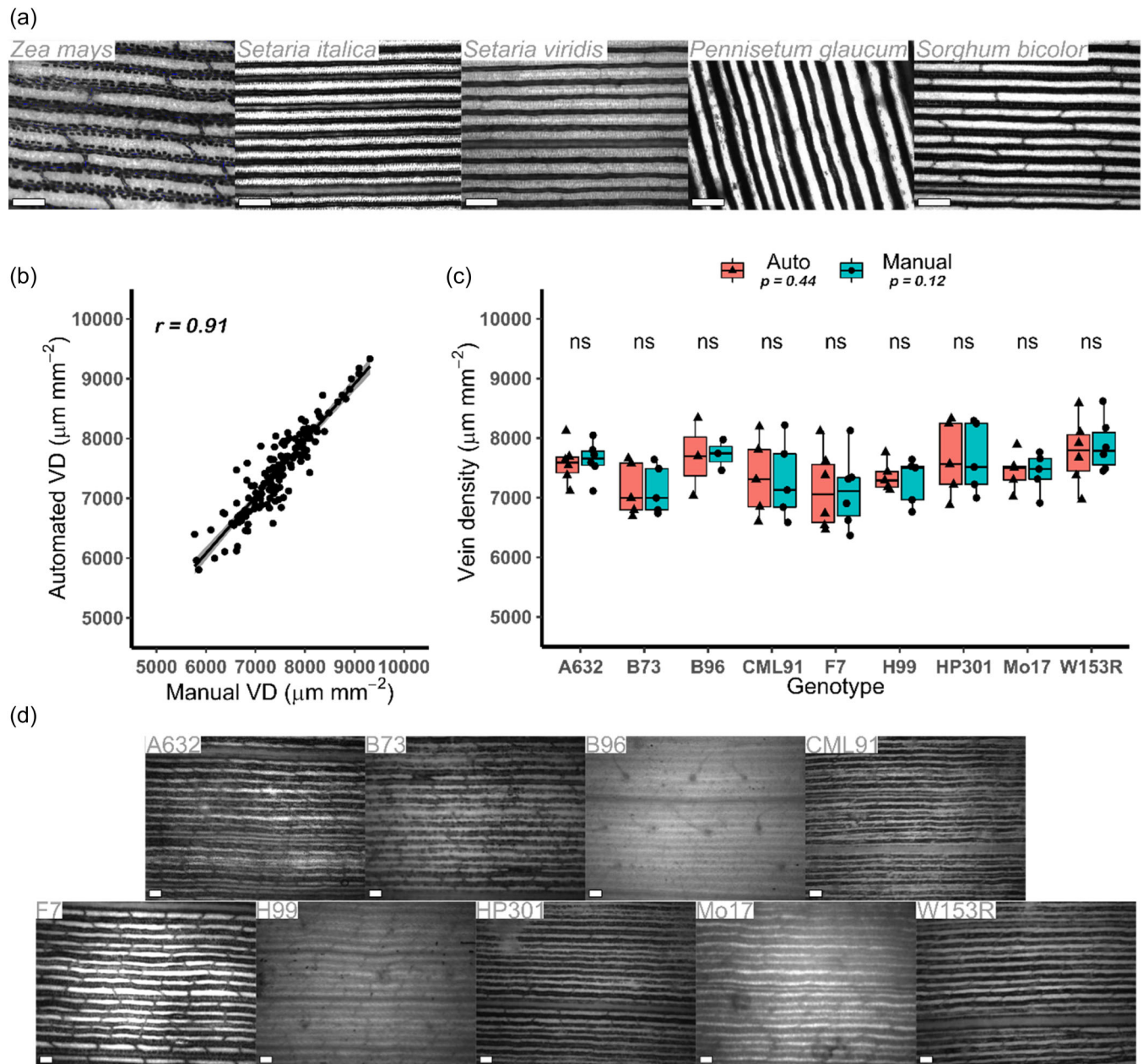


**FIGURE 3** Ground truthing of automated analysis. (a) Seventeen randomly selected lines from the population comprising 199 lines were assessed both automatically (red) and manually (blue) for vein density. (b) Correlation analysis between automatic and manual quantification of vein density. Line of best fit shown with grey shading representing 95% confidence intervals. (c) Distribution of vein density across 199 plants determined using Starch4Kranz. (d) Example of a manually traced image. (e) Example of the automated output. Scale bars = 200  $\mu\text{m}$ . Skeletons have been thickened to ensure visibility.  $r$ , Pearson's correlation coefficient; VD, vein density. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

(Bühler et al., 2015) installation requirements meant the software was awkward to implement, but also that it needed extensive manual input into each image. The requirement for manual inputs meant direct comparison with Starch4Kranz was limited and so we also compared our pipeline with that from LeafVeinCNN (Xu et al., 2021). Although LeafVeinCNN was designed for low-magnification, unstained leaves from trees its automation was more comparable with our pipeline. When 27 images were subjected to the LeafVeinCNN pipeline (Xu et al., 2021) 3.5 min per image were required to run the CNNs but repeated manual input was required to obtain optimal thresholding. Moreover, compared with LeafVeinCNN ( $r = 0.74$ ; Supporting Information: Figure S3B) the outputs from Starch4Kranz pipeline correlated more strongly with manual tracings ( $r = 0.95$ ; Supporting Information: Figure S4A). LeafVeinCNN performed less well when vein densities were low (Supporting Information: Figure S3C), which may be because it was trained on data derived from leaves at lower magnification.

We found that subdividing an image into multiple sections, and capturing slightly out-of-focus images could enhance contrast between veins and background. This was of particular use for samples with high vein densities, or when staining was poor. As we

found that some images did not allow accurate vein density to be determined, a manual check is incorporated into the pipeline (Figure 4b). Compared with other methods that require continual manual checking of images during processing (Bühler et al., 2015; Dhondt et al., 2012; Parsons-Wingert et al., 2014; Xu et al., 2021) Starch4Kranz saves time because it runs in the background across multiple images, and a manual check does not take place until the programme has finished. Machine learning techniques for image analysis are a useful tool for quantitative studies in plant sciences and have been employed to identify QTL associated with stomatal density in maize (Xie et al., 2021) and estimate heritability of vein density in sunflower (Earley et al., 2022). However, such approaches can require extensive computational understanding as well as power to build neural networks, and depending on the complexity of the phenotype may require manual annotation of thousands of images for network training (Moen et al., 2019; van Dijk et al., 2021). Thus, LeafVeinCNN is a powerful tool suitable for quantifying multiple traits associated with vein architecture, but Starch4Kranz offers a rapid and robust tool allowing vein density to be determined in  $C_4$  species. It also confirmed significant natural variation amongst  $F_2$  lines of *G. gynandra* (Figure 4c).



**FIGURE 4** Starch staining of vein density in monocotyledons. (a) Preferential accumulation of starch revealed through staining of bundle sheath strands of the  $C_4$  monocotyledons *Zea mays*, *Setaria italica*, *Setaria viridis*, *Pennisetum glaucum* and *Sorghum bicolor*. (b and c) No significant differences between founder lines of maize were detected from either manual (red) or automatic (blue) methods.  $p$  Values determined by ANOVA and Student's  $t$  test. (d) Representative images of maize leaves stained for starch. Scale bars = 200  $\mu\text{m}$ .  $r$ , Pearson's correlation coefficient; VD, vein density. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

### 3.3 | Natural variation in vein density was not detectable in founder lines of a maize MAGIC population

Aside from Grasvåg (Robil et al., 2021) we noted a lack of automated methods for the assessment of vein density in monocotyledons. Because the preferential accumulation of starch in bundle sheath cells had first been reported in monocotyledons (Lunn & Furbank, 1997), we tested the NADP-ME species *S. italica*, *S. viridis*, *P. glaucum*

and *S. bicolor* as well as maize itself (Figure 4a) and then subjected various maize accessions to the Starch4Kranz algorithm. For this, an additional step was added to the pipeline to remove commissural veins (Supporting Information: Figure S5), which are not considered important components of Kranz anatomy (Langdale, 2011). We also noted that they stained for starch inconsistently likely because they are considered a simplified vascular tissue not encircled by bundle sheath cells (Sakaguchi & Fukuda, 2008; Sakaguchi et al., 2010). If commissural veins are of interest, and sufficient contrast obtained,

then 'auto' can be selected as an input variable instead of 'monocot', or the 'monocot\_branch\_length\_to\_keep' can be set to 0 (see Section 2; Supporting Information: Table S1) and commissural veins will not be removed. Commissural veins were defined as the shortest branch stemming from a pixel with three branch points. Pattern recognition allowed these to be identified and removed (Supporting Information: Figure S5). The Python implementation of Starch4Kranz works slightly differently and detects commissural veins as the shortest Euclidean distance between two branch points and then removes this if it is below the 'monocot\_branch\_length\_to\_keep' variable (Supporting Information: Table S1). To ground-truth this approach, manual tracings that did not include these veins were undertaken. Images from starch staining of nine founder lines of a MAGIC population from maize (Dell'Acqua et al., 2015) were subjected to Starch4Kranz. Estimates were highly correlated with manual analysis ( $r = 0.91$ ; Figure 4a) but no statistically significant differences were observed between genotypes ( $p > 0.05$ ; Figure 4b). When ran through the Python version the correlation was slightly weaker ( $r = 0.81$ ; Supporting Information: Figure S2C) likely due to differences in commissural vein detection. Thus, Starch4Kranz can be used to quantify vein density in  $C_4$  monocotyledons in addition to dicotyledons (Figures 3 and 4; Supporting Information: Figure S6). In addition to this, we show that starch staining highlights the veins of both NAD-ME and NADP-ME species (Figures 1 and 4a). However, for the MAGIC maize founder lines studied here, we detected no significant differences in vein density (Figure 4c).

This finding emphasises the need to generate specific mapping populations designed for the trait of interest. *G. gynandra*  $F_2$  lines produced from a cross between two lines phenotypically distinct for vein density displayed abundant natural variation in this trait (Figure 3c).

Variation for vein density was not present in the MAGIC maize founders (Figure 4c) indicating that the MAGIC maize population (Dell'Acqua et al., 2015) is unlikely to be suitable for studying QTL underlying vein density. However, using a subset of lines from an association panel of 302 maize lines (Flint-Garcia et al., 2005), Kolbe and Cousins (2018) did identify significant variation for vein density. Considering the ease and speed with which Starch4Kranz permits accurate quantification of vein density means that further screening of this association panel or other  $C_4$  mapping populations such as NAM populations of maize (McMullen et al., 2009; Yu et al., 2008) or MAGIC populations of sorghum (Ongom and Ejeta, 2017) could prove valuable for future detection of QTL controlling vein density.

In summary, we report a simple protocol allowing vein density to be determined from starch staining of vascular tissue in  $C_4$  species. Automation of this pipeline allowed rapid and accurate quantification of vein density. As increased vein density is considered an important trait allowing  $C_4$  photosynthesis (Sage, 2001), where intraspecific natural variation for this trait exists our approach could provide insight into mechanisms evolution has co-opted to generate the complex  $C_4$  phenotype.

## ACKNOWLEDGEMENTS

The authors thank Angie Burnett, John Ferguson, and Johannes Kromdijk for provision of maize leaf material. The work was supported by a BBSRC DTP studentship to Conor J. C. Simpson, European Research Council Grant 694733 Revolution to Julian M. Hibberd, and Netherlands Organization for Scientific Research Grant W.08.270.350 to M. Eric Schranz. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

## ORCID

Pallavi Singh  <http://orcid.org/0000-0003-3694-6378>

Julian M. Hibberd  <http://orcid.org/0000-0003-0662-7958>

## REFERENCES

- Bradski, G. 2000. *The OpenCV Library*. Dr. Dobb's Journal of Software Tools.
- Brown, N.J., Newell, C.A., Stanley, S., Chen, J.E., Perrin, A.J., Kajala, K. et al. (2011) Independent and parallel recruitment of preexisting mechanisms underlying  $C_4$  photosynthesis. *Science*, 331, 1436–1439.
- Bühler, J., Rishmawi, L., Pflugfelder, D., Huber, G., Scharr, H. & Hülskamp, M. et al. (2015) phenoVein—a tool for leaf vein segmentation and analysis. *Plant Physiology*, 169, 2359–2370.
- Dell'Acqua, M., Gatti, D.M., Pea, G., Cattonaro, F., Coppens, F., Magris, G. et al. (2015) Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in *Zea mays*. *Genome Biology*, 16, 167.
- Dhondt, S., Van Haerenborgh, D., Van Cauwenbergh, C., Merks, R.M.H., Philips, W., Beemster, G.T.S. et al. (2012) Quantitative analysis of venation patterns of Arabidopsis leaves by supervised image analysis. *The Plant Journal*, 69, 553–563.
- Driever, S.M. & Kromdijk, J. (2013) Will  $C_3$  crops enhanced with the  $C_4$   $CO_2$ -concentrating mechanism live up to their full potential (yield)? *Journal of Experimental Botany*, 64, 3925–3935.
- Earley, A.M., Temme, A.A., Cotter, C.R. & Burke, J.M. (2022) Genomic regions associate with major axes of variation driven by gas exchange and leaf construction traits in cultivated sunflower (*Helianthus annuus* L.). *The Plant Journal*, 111, 1425–1438.
- Edwards, E.J. (2019) Evolutionary trajectories, accessibility and other metaphors: the case of  $C_4$  and CAM photosynthesis. *New Phytologist*, 223, 1742–1755.
- Flint-Garcia, S.A., Thuillet, A.C., Yu, J., Pressoir, G., Romero, S.M., Mitchell, S.E. et al. (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. *The Plant Journal*, 44, 1054–1064.
- Gehan, M.A., Fahlgren, N., Abbasi, A., Berry, J.C., Callen, S.T., Chavez, L. et al. (2017) PlantCV v2: image analysis software for high-throughput plant phenotyping. *PeerJ*, 5, e4088.
- Gibon, Y., Blaesing, O.E., Hannemann, J., Carillo, P., Höhne, M., Hendriks, J.H.M. et al. (2004) A robot-based platform to measure multiple enzyme activities in Arabidopsis using a set of cycling assays: comparison of changes of enzyme activities and transcript levels during diurnal cycles and in prolonged darkness. *The Plant Cell*, 16, 3304–3325.
- Gowik, U., Burscheidt, J., Akyildiz, M., Schlue, U., Koczor, M., Streubel, M. et al. (2004) cis-Regulatory elements for mesophyll-specific gene expression in the  $C_4$  plant *Flaveria trinervia*, the promoter of the  $C_4$  phosphoenolpyruvate carboxylase gene. *The Plant Cell*, 16, 1077–1090.
- Gowik, U., Schulze, S., Saladié, M., Rolland, V., Tanz, S.K., Westhoff, P. et al. (2017) A MEM1-like motif directs mesophyll cell-specific

- expression of the gene encoding the C<sub>4</sub> carbonic anhydrase in *Flaveria*. *Journal of Experimental Botany*, 68, 311–320.
- Guo, Z., Yang, W., Chang, Y., Ma, X., Tu, H., Xiong, F. et al. (2018) Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. *Molecular Plant*, 11, 789–805.
- Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D. et al. (2020) Array programming with {NumPy}. *Nature*, 585, 357–362.
- Hatch, M., Kagawa, T. & Craig, S. (1975) Subdivision of C<sub>4</sub>-pathway species based on differing C<sub>4</sub> acid decarboxylating systems and ultrastructural features. *Functional Plant Biology*, 2, 111.
- Hibberd, J.M. & Covshoff, S. (2010) The regulation of gene expression required for C<sub>4</sub> photosynthesis. *Annual Review of Plant Biology*, 61, 181–207.
- Hunter, J.D. (2007) Matplotlib: a 2D graphics environment. *Computing in Science & Engineering*, 9, 90–95.
- Jones, E., Oliphant, T. & Peterson, P. (2001) *SciPy: open source scientific tools for Python*.
- Kassambara, A. 2020. *ggpubr: 'ggplot2' based publication ready plots. R package version 0.4.0*.
- Kolbe, A.R. & Cousins, A.B. (2018) Mesophyll conductance in *Zea mays* responds transiently to CO<sub>2</sub> availability: implications for transpiration efficiency in C<sub>4</sub> crops. *New Phytologist*, 217, 1463–1474.
- Langdale JA. 2011. C<sub>4</sub> cycles: past, present, and future research on C<sub>4</sub> photosynthesis. *The Plant Cell*, 23, 3879–3892.
- Lundgren, M.R., Christin, P.A., Escobar, E.G., Ripley, B.S., Besnard, G., Long, C.M. et al. (2016) Evolutionary implications of C<sub>3</sub>–C<sub>4</sub> intermediates in the grass *Alloteropsis semialata*. *Plant, Cell & Environment*, 39, 1874–1885.
- Lunn, J.E. & Furbank, R.T. (1997) Localisation of sucrose-phosphate synthase and starch in leaves of C<sub>4</sub>. *Plants*, 61, 106–111.
- McMullen, M.D., Kresovich, S., Villeda, H.S., Bradbury, P., Li, H., Sun, Q. et al. (2009) Genetic properties of the maize nested association mapping population. *Science*, 325, 737–740.
- Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M. & Van Valen, D. (2019) Deep learning for cellular image analysis. *Nature Methods*, 16, 1233–1246.
- Nomura, M., Higuchi, T., Ishida, Y., Ohta, S., Komari, T., Imaizumi, N. et al. (2005) Differential expression pattern of C<sub>4</sub> bundle sheath expression genes in rice, a C<sub>3</sub> plant. *Plant and Cell Physiology*, 46, 754–761.
- Nomura, M., Higuchi, T., Katayama, K., Taniguchi, M., Miyao-Tokutomi, M., Matsuoka, M. et al. (2005) The promoter for C<sub>4</sub>-type mitochondrial aspartate aminotransferase does not direct bundle sheath-specific expression in transgenic rice plants. *Plant and Cell Physiology*, 46, 743–753.
- Nomura, M., Sentoku, N., Nishimura, A., Lin, J.H., Honda, C., Taniguchi, M. et al. (2000) The evolution of C<sub>4</sub> plants: acquisition of cis-regulatory sequences in the promoter of C<sub>4</sub>-type pyruvate, orthophosphate dikinase gene. *The Plant Journal*, 22, 211–221.
- Omoto, E., Taniguchi, M. & Miyake, H. (2012) Adaptation responses in C<sub>4</sub> photosynthesis of maize under salinity. *Journal of Plant Physiology*, 169, 469–477.
- Ongom, P.O. & Ejeta, G. (2017) Mating design and genetic structure of a Multi-Parent Advanced Generation Intercross (MAGIC) population of sorghum (*Sorghum bicolor* L.) moench). *Genetics*, 8, 331–341.
- Otsu, N. (1979) A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9, 62–66.
- Parsons-Wingenter, P., Vickerman, M.B., Paul, A.-L. & Ferl, R.J. (2014) Mapping by VESGEN of leaf venation patterning in *Arabidopsis thaliana* with bioinformatic dimensions of gene expression. *Gravitational and Space Research*, 2, 68–81.
- Perez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H. & Jaureguiberry, P. et al. (2016) New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany*, 61, 167–234. <https://doi.org/10.1071/BT12225>
- Rasband, W.S. (2014) *ImageJ: U. S. National Institutes of Health, Bethesda, Maryland, USA, 1997-2014*. <http://imagej.nih.gov/ij/>
- Reeves, G., Grangé-Guermente, M.J. & Hibberd, J.M. (2017) Regulatory gateways for cell-specific gene expression in C<sub>4</sub> leaves with Kranz anatomy. *Journal of Experimental Botany*, 68, 107–116.
- Reeves, G., Singh, P., Rossberg, T.A., Sogbohossou, E.O.D., Schranz, M.E. & Hibberd, J.M. (2018) Natural variation within a species for traits underpinning C<sub>4</sub> photosynthesis. *Plant Physiology*, 177, 504–512.
- Reyna-Llorens, I., Burgess, S.J., Reeves, G., Singh, P., Stevenson, S.R., Williams, B.P. et al. (2018) Ancient duons may underpin spatial patterning of gene expression in C<sub>4</sub> leaves. *Proceedings of the National Academy of Sciences*, 115, 1931–1936.
- Robil, J.M., Gao, K., Neighbors, C.M., Boeding, M., Carland, F.M., Bunyak, F. et al. (2021) Grasviq: an image analysis framework for automatically quantifying vein number and morphology in grass leaves. *The Plant Journal*, 107, 629–648.
- Sage, R.F. (2001) Environmental and evolutionary preconditions for the origin and diversification of the C<sub>4</sub> photosynthetic syndrome. *Plant Biology*, 3, 202–213.
- Sage, R.F., Christin, P.-A. & Edwards, E.J. (2011) The C<sub>4</sub> plant lineages of planet Earth. *Journal of Experimental Botany*, 62, 3155–3169.
- Sakaguchi, J. & Fukuda, H. (2008) Cell differentiation in the longitudinal veins and formation of commissural veins in rice (*Oryza sativa*) and maize (*Zea mays*). *Journal of Plant Research*, 121, 593–602.
- Sakaguchi, J., Itoh, J.I., Ito, Y., Nakamura, A., Fukuda, H. & Sawa, S. (2010) COE1, an LRR-RLK responsible for commissural vein pattern formation in rice. *The Plant Journal*, 63, 405–416.
- Sales, C.R.G., Ribeiro, R.V., Hayashi, A.H., Marchiori, P.E.R., Silva, K.I., Martins, M.O. et al. (2018) Flexibility of C<sub>4</sub> decarboxylation and photosynthetic plasticity in sugarcane plants under shading. *Environmental and Experimental Botany*, 149, 34–42.
- Sedelnikova, O.V., Hughes, T.E. & Langdale, J.A. (2018) Understanding the genetic basis of C<sub>4</sub> Kranz anatomy with a view to engineering C<sub>3</sub> crops. *Annual Review of Genetics*, 52, 249–270.
- Sharwood, R.E., Sonawane, B.V. & Ghannoum, O. (2014) Photosynthetic flexibility in maize exposed to salinity and shade. *Journal of Experimental Botany*, 65, 3715–3724.
- Sulpice, R., Trenkamp, S., Steinfath, M., Usadel, B., Gibon, Y., Witucka-Wall, H. et al. (2010) Network analysis of enzyme activities and metabolite levels and their relationship to biomass in a large panel of Arabidopsis accessions. *The Plant Cell*, 22, 2872–2893.
- Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P. & Hawkesford, M.J. (2017) Field Scanalyzer: an automated robotic field phenotyping platform for detailed crop monitoring. *Functional Plant Biology*, 44, 143–153.
- Van Der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N. et al. (2014) Scikit-image: image processing in python. *PeerJ*, 2, e453.
- van Dijk, A.D.J., Kootstra, G., Kruijer, W. & de Ridder, D. (2021) Machine learning in plant science and plant breeding. *iScience*, 24, 101890.
- Williams, B.P., Burgess, S.J., Reyna-Llorens, I., Knerova, J., Aubry, S., Stanley, S. et al. (2016) An untranslated cis-element regulates the accumulation of multiple C<sub>4</sub> enzymes in *Gynandropsis gynandra* mesophyll cells. *The Plant Cell*, 28, 454–65.
- Xie, J., Fernandes, S.B., Mayfield-Jones, D., Erice, G., Choi, M., E Lipka, A. et al. (2021) Optical topometry and machine learning to rapidly phenotype stomatal patterning traits for maize QTL mapping. *Plant Physiology*, 187, 1462–1480.
- Xu, H., Blonder, B., Jodra, M., Malhi, Y. & Fricker, M. (2021) Automated and accurate segmentation of leaf venation networks via deep learning. *New Phytologist*, 229, 631–648.

- Yabiku, T. & Ueno, O. (2017) Variations in physiological, biochemical, and structural traits of photosynthesis and resource use efficiency in maize and teosintes (NADP-ME-type C<sub>4</sub>). *Plant Production Science*, 20, 448–458.
- Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J.H., Batchelor, W.D. et al. (2020) Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. *Molecular Plant*, 13, 187–214.
- Yu, J., Holland, J.B., McMullen, M.D. & Buckler, E.S. (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics*, 178, 539–551.
- Zuiderveld, K. (1994) Contrast limited adaptive histogram equalization. In: Heckbert, P.S. (Ed.) *Graphic Gems IV*. Academic Press Professional, Inc., pp. 474–485.

## SUPPORTING INFORMATION

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

**How to cite this article:** Simpson, C.J.C., Singh, P., Sogbohossou, D.E.O., Eric Schranz, M. & Hibberd, J.M. (2023) A rapid method to quantify vein density in C<sub>4</sub> plants using starch staining. *Plant, Cell & Environment*, 46, 2928–2938. <https://doi.org/10.1111/pce.14656>