Carolyn Gigot

OEB 107

Andrew Knoll

December 10, 2015

Measuring Leaf Venation: Methods and Implications

1. Abstract

Studies of modern and fossil leaf vein density (LVD), the total length of measured veins per unit of leaf area, suggest that angiosperms show significantly denser venation than other present or past groups (Boyce et al, 2009). This characteristic might explain angiosperms' diversification and proliferation across many modern environments. In addition to its implications for angiosperm success, better understanding leaf vein architectures might aid understanding of plant systematics, physiology, development, and ecology (Roth-Kebelsick et al, 2001).

Despite apparent patterns in vein architectures and their significance for the study of modern and fossil plants- where molecular data are unavailable and LVD would represent an important quantitative tool- this measure presents a variety potential confounding factors. In addition to impracticality for field study, Price et al found that the field of view used for automated image analysis, magnification, and image resolution each affected the result of LVD measurements, explaining some of the variation within specific taxa recorded in previous studies (2014). Green et al instead studied differences in vein architecture through the semi quantitative leaf "rank" and automated areole measurement systems, achieving relatively equivelant and stable results across the two methods (2014).

This paper aims to investigate a third quantitative classification of leaf vein architecture, the distance from each pixel in a leaf image to a "vein" pixel, as classified by a binary mask generated through adaptive threshholding. This method seems most physiologically applicable, as the distance photosynthate or water must travel to or from the nearest vein applies directly to its ability to participate in metabolism. We hypothesize that progressively degrading a series of leaf images will affect this measurement in a similar manner to LVD as studied by Price et al (2014), and that measured distance maps will begin to deviate significantly around where image quality breaks down, at 25% of images' original number of pixels.

2. Introduction

Angiosperm eudicot leaves show a characteristic and highly reticulated organization of veins, important for the transportation of water, photosynthate, and nutrients, as well as for leaf structure and photosynthetic tissue support. The high variation in leaf venation patterns and plasticity of many leaf vein features suggest strong selective pressure on this aspect of leaf morphology (Roth-Kebelsick et al, 2001). Leaf vein architectural diversity allows this characteristic to act as an important tool for systematic classifications, especially for fossilized plants where molecular data cannot be employed (Hickey and Wolfe, 1975). In addition to their importance for systematics, features of leaf vein architecture allow closer study of present and past leaf physiology. The hydraulic conductance of leaves across plant species varies more than 65-fold, and scales with differences in vein architecture as well as anatomy. Because leaf hydraulics affect rates of gas exchange and thus photosynthesis as well as changes through the day and year, or responses to environmental variation, understanding vein architecture represents a key component of understanding metabolism (Sack and Holbrook, 2006). Developmental botanists also require information on vein architecture and its interactions with phylogeny and

physiology; the development of vein networks remains an important and relatively unanswered problem (Sack and Scoffoni, 2013).

Furthermore, because leaves are among the most commonly and best preserved organs of fossil plants, the study of leaf venation allows insights into ancient plant phylogeny and physiology largely unavailable through other estimates, as well as closer study of paleoclimate and paleoenvironments. Boyce et al measured the ratio of leaf vein lengths to areas using digital image analysis to estimate that angiosperm leaf vein densities vary from 8 mm/mm^2 while non-angiosperms average closer to 2 mm/mm^2, possibly implicating denser veins in the success of these taxa. Boyce et al's study of vein densities also allowed the authors to consider events such as the advent of modern tropical rainforests. In studying the impacts of denser leaf veins on transpiration-driven water cycles, the authors predicted larger-scale ecological changes in line with evidence from sedimentary geology and molecular phylogenies (Boyce et al, 2009).

Despite the enormous potential the study of leaf veins holds for the study of systematics, plant physiology, developmental botany, paleoecology and geobiology, quantifying leaf venation remains difficult. Much work on leaf venation patterns remain largely qualitative, due to the time consuming nature of quantitative methods; the specificity of architectures to particular taxa; impracticality for field identification; and the difficulty of applying many methods to the often limited and fragment-based fossil record (Ellis et al, 2009).

Some analytical or numerical methods have been borrowed from geomorphological classifications of river systems. The Horton-Stahler method, which considers a network as a directed graph and gives each node a degree by its number of children, can analytically and numerically classify vein networks, but results often differs from other methods. Other mathematical methods consider vein networks as particular types of graphs, naming internodes

or segments by their connectivity, with either the largest (midrib) considered primary or the the finest as primary and others counted upwards towards the periole. However, the time required to characterize even a small leaf is large, and the three-dimensional nature of leaves adds further complexity. Additionally, even defining what makes up a vein presents difficulties: some authors would include any strand of vascular tissue, while others demand both xylem and phloem, and some also add the adjacent fibers and bundle sheath cells. Choosing the simple functional definition of a vein as a vascular bundle that stains darker than surrounding tissue when treated with common biological stains avoids these confusions, but leaves the biological definition of a vein for debate (Green et al, 2014).

Like the functional definition of a vein, the semi-quantitative vein "order "system allows classification of vein structure and density in a consistent but somewhat arbitrary manner. Ash et al measure four to seven orders: the thickest vein is primary, and the next less than ¾ the size of the primary is secondary, and the veins which cover the lamina are tertiary; for higher orders than three, smaller branches than ¾ again receive a denomination as the next highest orders (2009). The 12-15 leaf "ranks" measured by Hickey et al classify vein organizations in a similar manner; this method is also repeatable but shows limited correlations to phylogeny, environment, and ecological significance (1975). Though no complete developmentally or physiologically based system classifies leaf veins, some evidence suggests that vein density correlates with physiological variables such as water uptake (Green et al, 2014).

In order to provide a more meaningful, repeatable, and easily measured quantification for leaf vein architecture, Green et al examined several automated image analysis methods (2014). All use Matlab and EBImage to analyze cleared leaf specimens by translating the image to a black and white binary mask of vein and non-vein or areole. The first measure describes the

absolute size of the areole, or non-leaf vein area, and the second executes a sizing transform on this information. Areole sizes show a significant correlation with rank, though rank explains only 14% of variation in areole size. Thus, ranks might serve as a quick method for classifying differences in the field or the fossil record where information is comparatively sparse, but automated image analysis provides an increasingly important and efficient classification system (Green et al, 2014).

Previous automated analysis of leaf vein architecture largely focuses on leaf vein density (LVD), a measure of the total length of veins per unit area. This quantity shows correlations with photosynthetic rates, hydraulic conductance, leaf size and conductance, and leaf shape or allometry. Because higher vein density means water must travel shorter distances through mesophyll, vein density provides a mechanism to change leaf physiological efficiency. However, different methods of measuring this quantity as well as different initial images might affect LVD as well. Price et al examined how different magnifications, fields of view, and image resolutions affected LVD (2014). The authors found that small fields of view caused high estimates of LVD with a high variance, but that the measure leveled off asymptotically with larger fields of view. Higher magnification and resolution also increased estimates of LVD. These results explain much of the variation across LVD studies, and emphasize the importance of considering the role of additional factors or controlling to build consistent comparisons across leaves (Price et al, 2014).

If the distance water must travel from vein to location of photosynthesis affects metabolism, then that distance provides the most physiologically useful quantity. Directly creating the distance map for distances between non-vein and vein pixels for each pixel using automated image analysis could reduce some of the issues with the LVD measurement or the

areole area model. However, initial image quality and measurement method might still affect this quantity and call into question phylogenetic, physiological, developmental, and ecological implications. This paper will therefore examines this method and the impact of selectively degrading the images to be analyzed.

3. Materials and Methods

129 0.8 by 1.2 cm cleared and stained angiosperm leaf sections photographed with a Nikon D50 at an initial size of 3008*2000 pixels were taken from the Dryad Digital Repository Green et al, 2014. Leaf venation was measured using the software Matlab and the program EBImage; image was first converted to a binary, black and white mask using adaptive thresholding, then cleaned for black or white locations unconnected to others smaller than a particular size. The distmap function was used to generate a matrix containing the distance from each pixel to a pixel classified as "vein" by the binary mask (see Appendix for script).

Test image number of pixel images was reduced by a factor of 2, 4, 8, and 16, creating new images with 50%, 25%, 12.5%, and 6.25% of the original pixels. The new images were then resized to their original dimensions. The effect on distance map results were observed for a series of five leaves across the phylogeny: the log(distmap) results were plotted on a series of histograms, and a table of mode, mean, and standard deviation for the distmap results were created (see Appendix for script).

4. Results

Figure 1. Sample analyses for five leaves across the angiosperm phylogeny. The cleared leaf section used for each analysis is shown at the top. The histograms below represent the distribution of numeric values for distances from any given pixel to a vein, appearing similar for the original, 50%, and 25% pixel images, but deviating at 12.5%. The images beneath each

histogram show the appropriate image degradation level; vein patterning and visible structure also becomes unclear with roughly 12.5% of the original pixels. See attached files for full-resolution original and degraded images.

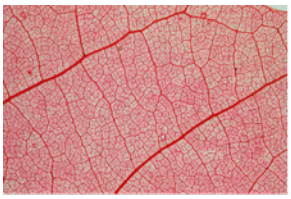
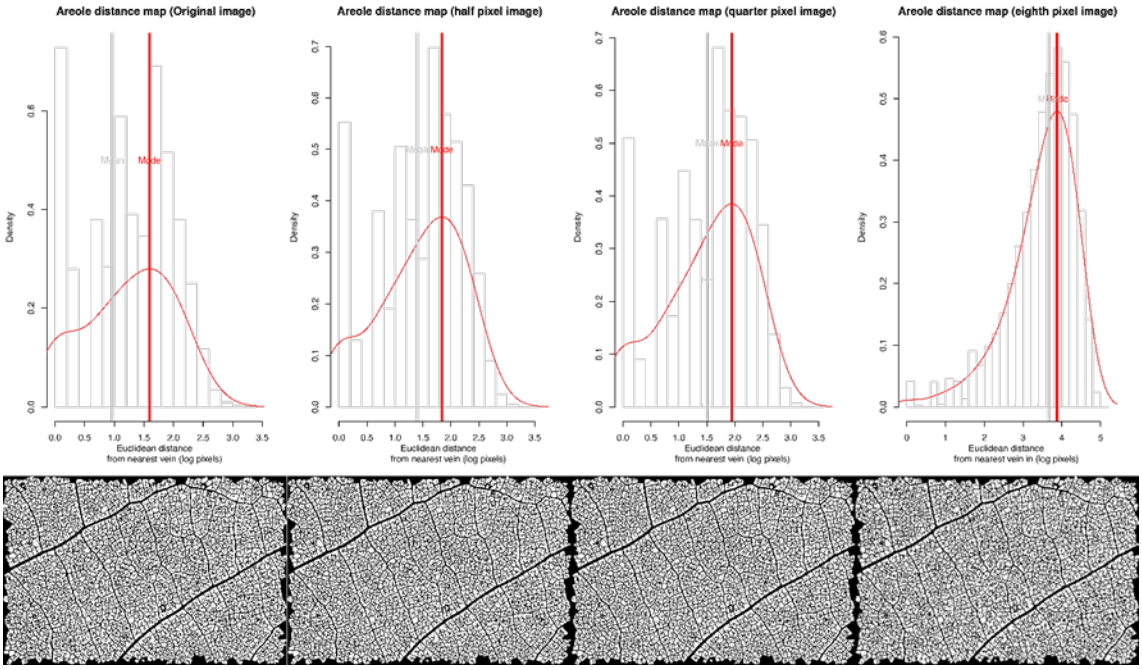


Image 1, *Uraria clarksei* (family Fabaceae)



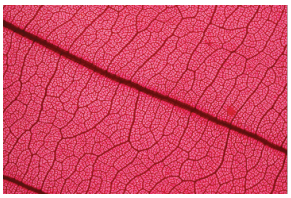
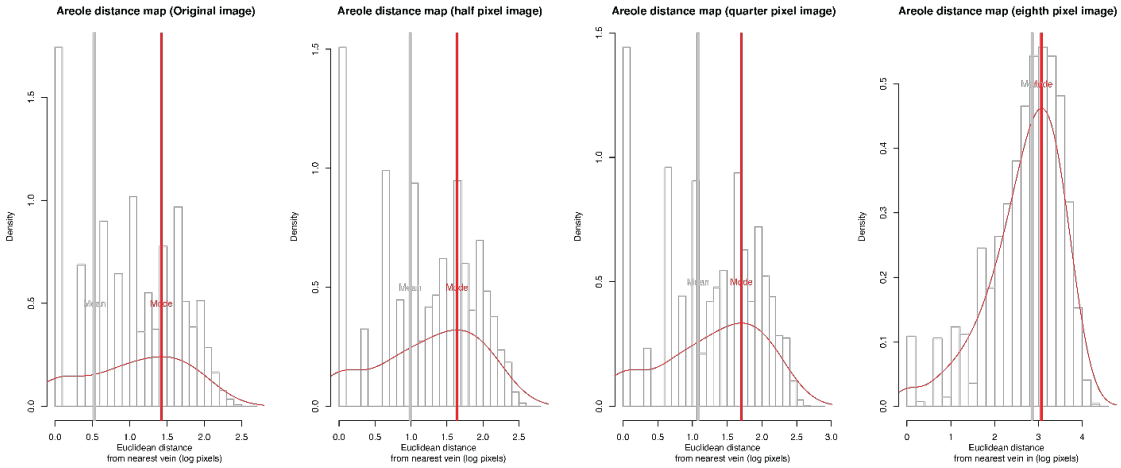


Image 2, *Chrysophyllum bonkokoensis* (family Sapotaceae)



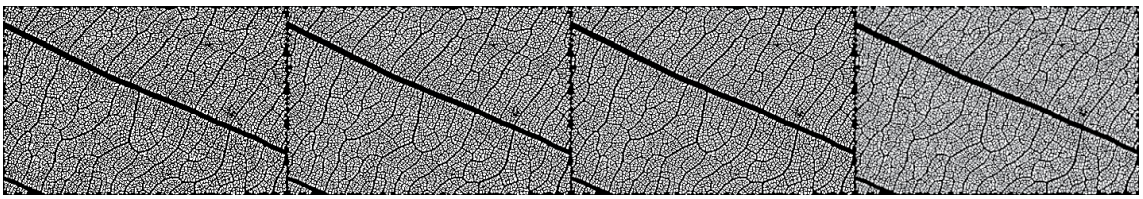
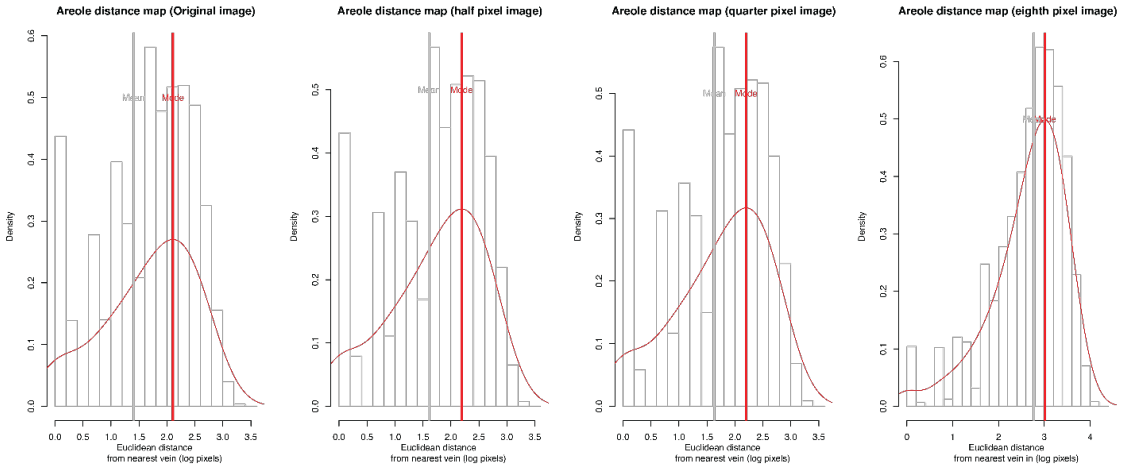
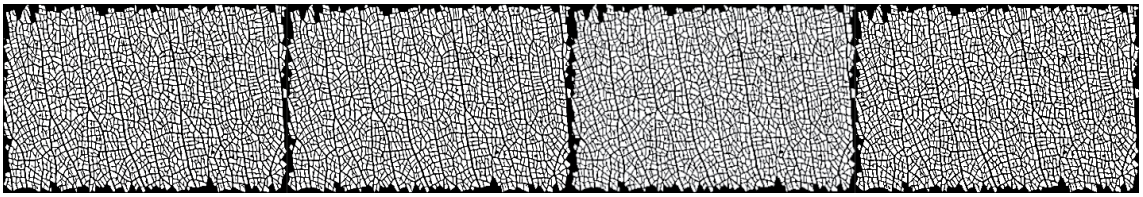


Image 3, *Salix humilis* (family Salicaceae)





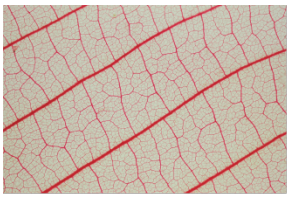
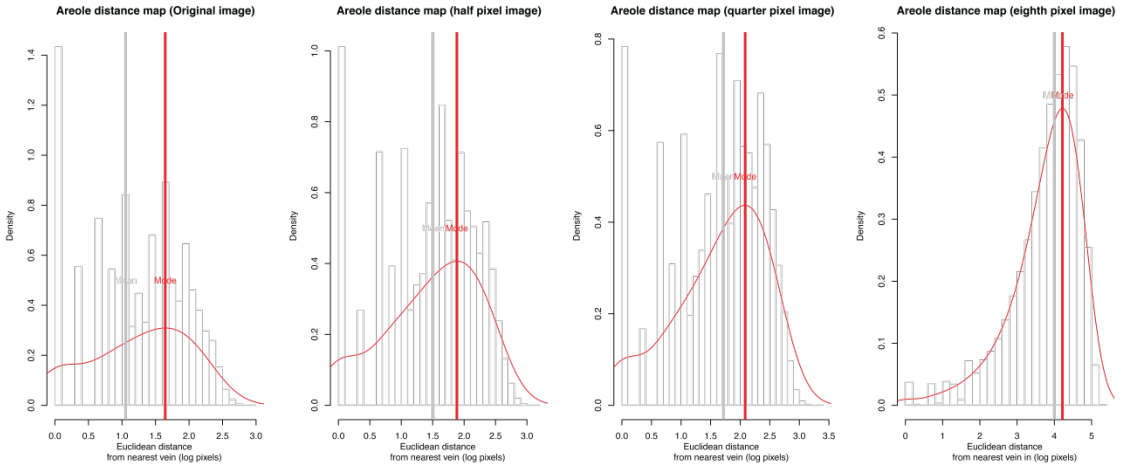
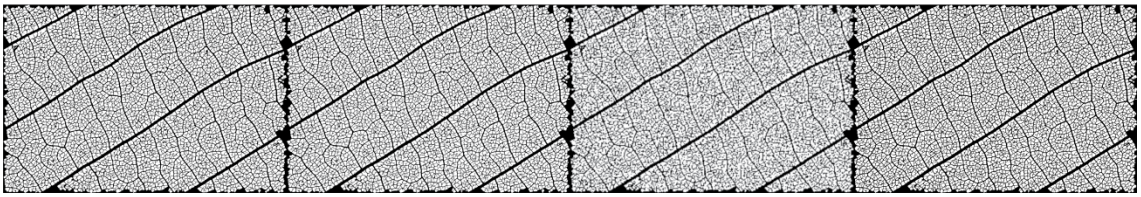


Image 4, *Coussapoa villosa* (family Cecropiaceae)





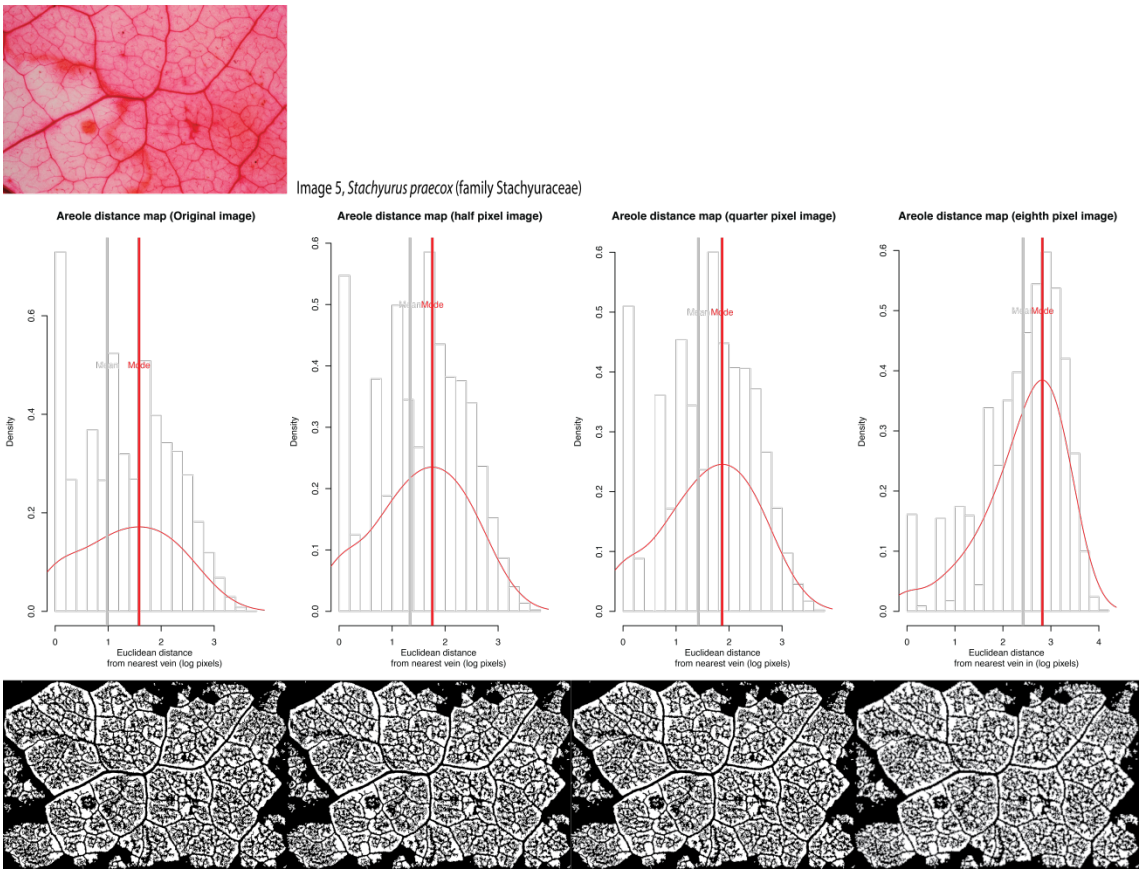
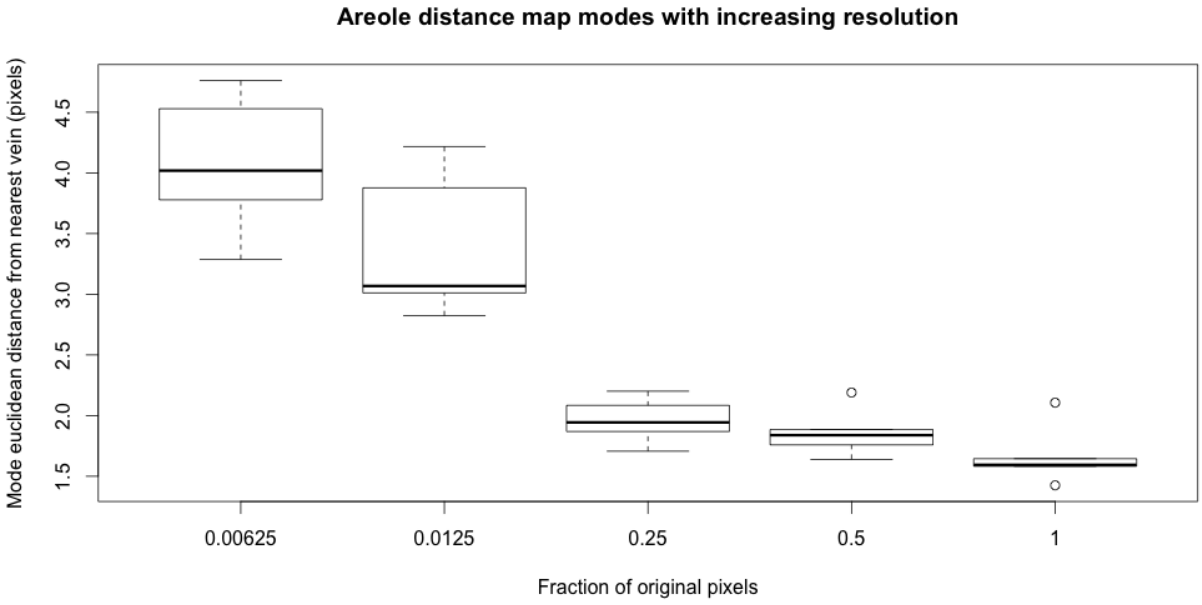
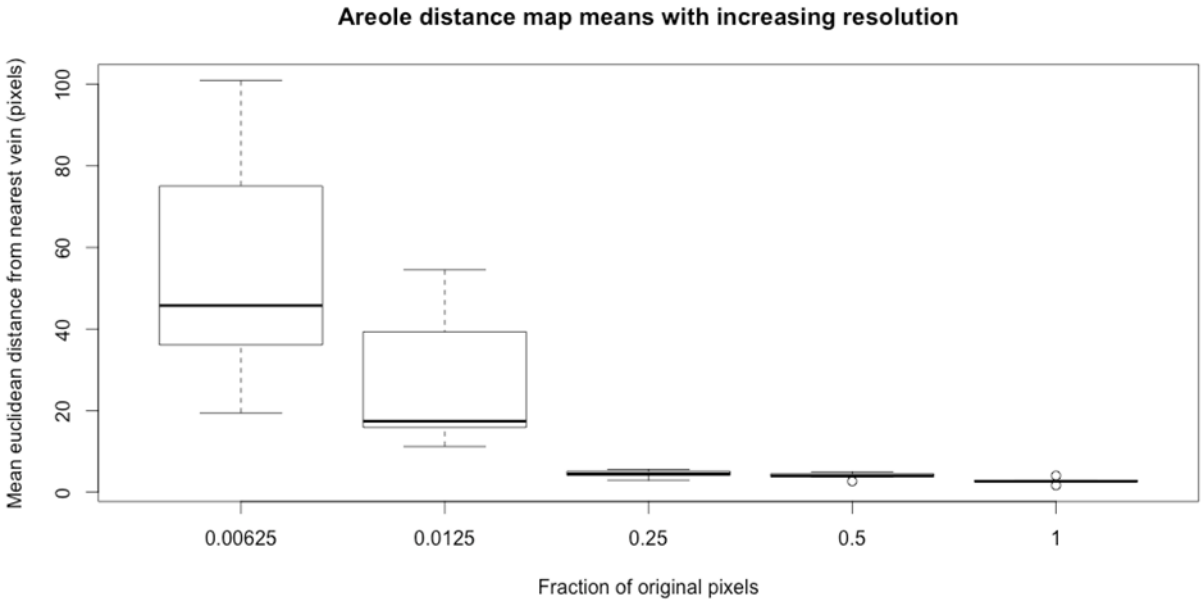
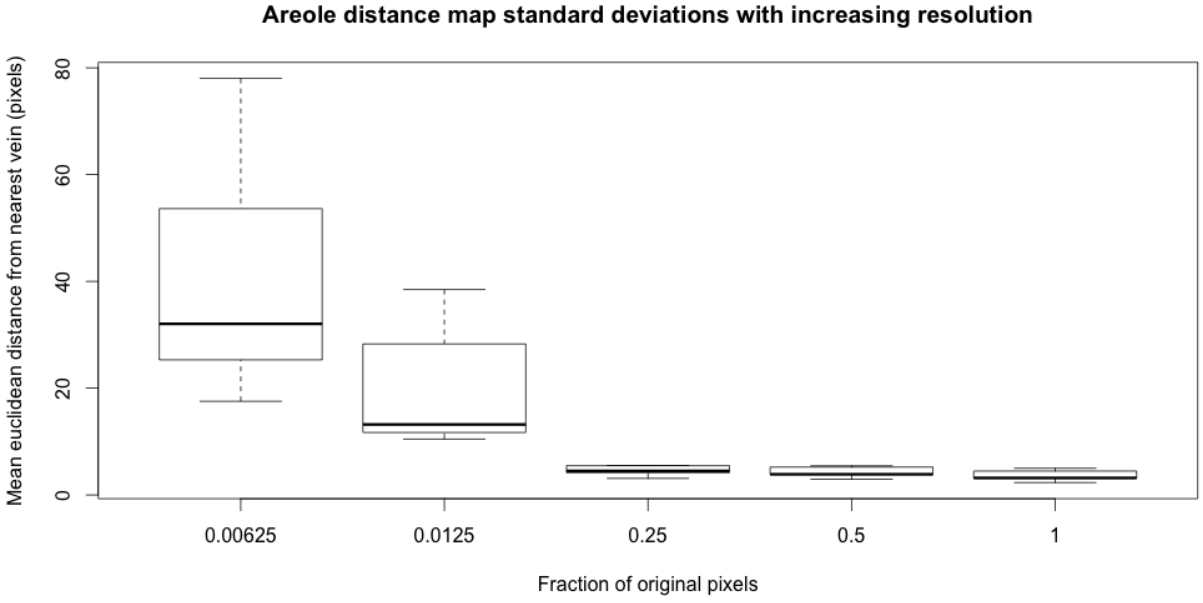


Figure 2.

Boxplots show the distribution of modes, means, and standard deviations for the images studied at each level of image quality reduction.


Areole distance map modes with increasing resolution
Mode euclidean distance from nearest vein (pixels)
1.5
2.0
2.5
3.0
3.5
4.0
4.5
0.00625
0.0125
0.25
0.5
1
Fraction of original pixels





5. Discussion

As hypothesized, the distance from the nearest vein as described by mean, mode, and distribution all appear to shift as image quality deteriorates; these measures diverge significantly from those found in the original image at a pixel reduction to 12.5% of the original number of pixels, the point at which leaf vein patterns in the original images also begin to break down

(Figure 1). These results suggest that the log-transformed distance map representation of leaf vein density provides a measure which corresponds reasonably well to human judgment.

Although we initially predicted that degrading image quality would both increase and decrease the mean and mode of the distance maps, our results showed increased distances from veins with decreasing image quality for all leaves studied (Figure 2). This might be due to the image-degrading method selected: by keeping the region of interest constant but reducing the number of pixels, the process chosen effectively removes the smallest, then larger and larger veins from the image. The most degraded images then eventually reveal only the largest veins, and distances from background pixels to these vein pixels are thus relatively larger.

Thus, considering other ways of reducing image size- perhaps choosing a smaller region of interest- might also affect the consistency of leaf vein distance maps with decreasing image quality. The region chosen for analysis might also affect our results. Our results appear consistent across taxa, further supporting the distance map measure of leaf vein density. However, only angiosperms were studied; adding other major groups to our study of image robustness might add to confidence in this measure. Additionally, all images studied were cleared and prepared in the same way; further investigating different preparation or preservation methods provides an important next avenue of research.

6. Conclusion

In summary, distance maps from background image pixels to pixels classified as veins through the generation of a binary pixel/vein mask provide a physiologically useful and consistent measure of leaf vein architecture, a metric important to botanists studying the systematics, physiology, development, and ecology of modern and ancient plants. However, the resolution of the image selected also affects this metric, especially as resolution decreases below

752,000 square pixels per square centimeter. Studying additional non-angiosperm taxa, different leaf clearing methods, and the factors affecting the initial adaptive thresholding classifying pixels as vein or non-vein all represent important avenues to further evaluate this metric.

7. Appendix

```
#Leaf Venation Analysis II
#December 10, 2015
#Carolyn Gigot (modified from Tinker Green)

###Load libraries
library(EBImage)
source('http://bricol.net/temp/image_analysis_functions.R')

###Set parameters
PATH = "/Users/carolyn/Documents/OEB107Project/images/testREAL/"
THRESH_WIN = 30 #Size of window for adaptive thresholding
THRESH_SENSE = 0.01 #threshhold sensitivity for adaptive thresholding
SCHMUTZ_MIN = 81 #Maximum size dismissed as noise

leafID={}
factor={}
mode={}
mean={}
sd={}

###Read and clean images
ims=dir(PATH)
for(i in seq_along(ims)){ #loop through all images
  #i=1
  raw <- readImage(paste(PATH, ims[i], sep = ''))
  writeImage(raw, 'raw.tiff', bits.per.sample = 16)
  bin.cl <- clean(raw, thresh_win = THRESH_WIN, thresh_sense = THRESH_SENSE, schmutz
= SCHMUTZ_MIN) #Makes the binary mask
  writeImage(bin.cl, 'cleaned.tiff', bits.per.sample = 16)

  ###Defines by what factors I will reduce the size of the pictures
  one=0.5 #I will initially reduce the number of pixels in the image by a factor of 2 (*0.5)
  two=0.25
  three=0.0125
  four=0.00625

  ###Decreases the number of pixels in the image
  dimensions=dim(bin.cl)
```

```
  bin1.cl=resize(bin.cl, dimensions[1]*sqrt(one), dimensions[2]*sqrt(one)) #Reduces the number of pixels by sqrt(one), that is sqrt (0.5), making the image half as large
  bin1.cl=resize(bin1.cl, dimensions[1], dimensions[2]) #Brings the image size back up so distance measures are the same

  bin2.cl=resize(bin.cl, dimensions[1]*sqrt(two), dimensions[2]*sqrt(two))
  bin2.cl=resize(bin2.cl,dimensions[1],dimensions[2])

  bin3.cl=resize(bin.cl, dimensions[1]*sqrt(three), dimensions[2]*sqrt(three))
  bin3.cl=resize(bin3.cl, dimensions[1],dimensions[2])

  bin4.cl=resize(bin.cl, dimensions[1]*sqrt(four), dimensions[2]*sqrt(four))
  bin4.cl=resize(bin4.cl, dimensions[1],dimensions[2])

  ###Optional: displays what the reduction of image size did
  #display(bin.cl)
  #display(bin1.cl)
  #display(bin2.cl)
  #display(bin3.cl)
  #display(bin4.cl)

  ###Distance transform of areole size
  bin.dm <- distmap(bin.cl)
  dm.bar <- mean(bin.dm)
  dens <- density(log(as.numeric(bin.dm)), adjust = 10) #Not sure what this does?

  ###Same transform for the reduced images
  bin1.dm <- distmap(bin1.cl)
  dm1.bar <- mean(bin1.dm)
  dens1 <- density(log(as.numeric(bin1.dm)), adjust = 10)

  bin2.dm <- distmap(bin2.cl)
  dm2.bar <- mean(bin2.dm)
  dens2 <- density(log(as.numeric(bin2.dm)), adjust = 10)

  bin3.dm <- distmap(bin3.cl)
  dm3.bar <- mean(bin3.dm)
  dens3 <- density(log(as.numeric(bin3.dm)), adjust = 10)

  bin4.dm <- distmap(bin4.cl)
  dm4.bar <- mean(bin4.dm)
  dens4 <- density(log(as.numeric(bin4.dm)), adjust = 10)

  #This will save the plot
  name=paste("histograms",i,".pdf")
  pdf(name, width=14, height=6)
```

```
  #Allows me to plot multiple histograms side by side
  old.par <- par(mfrow=c(1, 4))

  #Plots log tranformed results on a histogram
  hist(log(as.numeric(bin.dm)), main = 'Areole distance map (Original image)', freq = FALSE,
border = 'grey',
      xlab = 'Euclidean distance
      from nearest vein (log pixels)')
  lines(dens, col = 'red')
  abline(v = dens$x[dens$y == max(dens$y)], col = 'red', lwd = 3)
  text(x = dens$x[dens$y == max(dens$y)], y = 0.5, 'Mode', col = 'red', xpd = NA)
  abline(v = log(dm.bar), col = 'grey', lwd = 3)
  text(x = log(dm.bar), y = 0.5, 'Mean', col = 'grey', xpd = NA)

  #Plots log transformed results for the degenerated images
  hist(log(as.numeric(bin1.dm)), main = 'Areole distance map (half pixel image)', freq = FALSE,
border = 'grey',
      xlab = 'Euclidean distance
      from nearest vein (log pixels)')
  lines(dens1, col = 'red')
  abline(v = dens1$x[dens1$y == max(dens1$y)], col = 'red', lwd = 3)
  text(x = dens1$x[dens1$y == max(dens1$y)], y = 0.5, 'Mode', col = 'red', xpd = NA)
  abline(v = log(dm1.bar), col = 'grey', lwd = 3)
  text(x = log(dm1.bar), y = 0.5, 'Mean', col = 'grey', xpd = NA)

  hist(log(as.numeric(bin2.dm)), main = 'Areole distance map (quarter pixel image)', freq =
FALSE, border = 'grey',
      xlab = 'Euclidean distance
      from nearest vein (log pixels)')
  lines(dens2, col = 'red')
  abline(v = dens2$x[dens2$y == max(dens2$y)], col = 'red', lwd = 3)
  text(x = dens2$x[dens2$y == max(dens2$y)], y = 0.5, 'Mode', col = 'red', xpd = NA)
  abline(v = log(dm2.bar), col = 'grey', lwd = 3)
  text(x = log(dm2.bar), y = 0.5, 'Mean', col = 'grey', xpd = NA)

  hist(log(as.numeric(bin3.dm)), main = 'Areole distance map (eighth pixel image)', freq =
FALSE, border = 'grey',
      xlab = 'Euclidean distance
      from nearest vein in (log pixels)')
  lines(dens3, col = 'red')
  abline(v = dens3$x[dens3$y == max(dens3$y)], col = 'red', lwd = 3)
  text(x = dens3$x[dens3$y == max(dens3$y)], y = 0.5, 'Mode', col = 'red', xpd = NA)
  abline(v = log(dm3.bar), col = 'grey', lwd = 3)
  text(x = log(dm3.bar), y = 0.5, 'Mean', col = 'grey', xpd = NA)
```

```
  #hist(log(as.numeric(bin4.dm)), main = 'Areole distance map (sixteenth pixel image)', freq =
FALSE, border = 'grey',
      # xlab = 'Euclidean distance
       #from nearest vein (log pixels)')
  #lines(dens4, col = 'red')
  #abline(v = dens4$x[dens4$y == max(dens4$y)], col = 'red', lwd = 3)
 # text(x = dens4$x[dens4$y == max(dens4$y)], y = 0.5, 'Mode', col = 'red', xpd = NA)

  #abline(v = log(dm4.bar), col = 'grey', lwd = 3)
  #text(x = log(dm4.bar), y = 0.5, 'Mean', col = 'grey', xpd = NA)

  par(old.par)
  dev.off()

  ###Builds a dataframe with the maximum distance, mean distance, and standard deviation for
each resolution of the image
  ###defines the mode of the log-transformed areole distancemap

  leafID=c(leafID,rep(i,5)) #defines which leaf was studied
  factor=c(factor,1, one, two, three, four) #fraction of original pixels
  mode=c(mode, dens$x[dens$y == max(dens$y)],dens1$x[dens1$y ==
max(dens1$y)],dens2$x[dens2$y == max(dens2$y)],dens3$x[dens3$y ==
max(dens3$y)],dens4$x[dens4$y == max(dens4$y)])
  mean=c(mean, dm.bar, dm1.bar, dm2.bar, dm3.bar, dm4.bar)
  sd=c(sd, sd(bin.dm), sd(bin1.dm), sd(bin2.dm), sd(bin3.dm), sd(bin4.dm))
  df=data.frame(leafID, factor, mode, mean, sd)}

save(df)

#makes the most basic plot of factor degraded with the mean
#plot(df$factor,df$mean, main="Areole distance map with increasing resolution",xlab="Fraction
of original pixels",ylab="Mean euclidean distance from nearest vein (pixels)")
#abline(lm(df$mean~df$factor),col="blue")

#makes the same plot for just one leaf
#leaf1=df[1:5,]
#factor=leaf1$factor
#factor2=factor^2
#plot(factor,leaf1$mean)
#abline(lm(leaf1$mean~factor+factor2),col="blue")
#save(df,file="dataFull.Rda")
#write.table(df, file="dataFull.txt")

#makes a series of boxplots at each factor for the means, modes, and sds
```

```
boxplot(df$mean~df$factor, data=df, main="Areole distance map means with increasing resolution",xlab="Fraction of original pixels",ylab="Mean euclidean distance from nearest vein (pixels)")
boxplot(df$mode~df$factor, data=df, main="Areole distance map modes with increasing resolution",xlab="Fraction of original pixels",ylab="Mode euclidean distance from nearest vein (pixels)")
boxplot(df$sd~df$factor, data=df, main="Areole distance map standard deviations with increasing resolution",xlab="Fraction of original pixels",ylab="Mean euclidean distance from nearest vein (pixels)")
```

8. Works Cited

Ash,A., B. Ellis, L. J. Hickey, K. Johnson, P.Wilf, S.Wing.1999. *Manual of leaf architecture*. Smithsonian Institution, Washington, D.C., USA.

Boyce, K., Brodribb, T., Field, T., and Zwieniecki, M. 2009. "Angiosperm leaf vein evolution was physiologically and environmentally transformative." *Proceedings of the Royal Society Biology* 276, 1771-1776. https://canvas.harvard.edu/courses/6502/files/folder/Readings?preview=1015539

Green, W., Little, S., Price, C., Wing,S., Smith, S., Kotre, B, and Doria, G. 2014. "Reading the Leaves: A Comparison of Leaf Rank and Automated Areole Measurement for Quantifying Aspects of Leaf Venation." *Applications in Plant Sciences,* 2(8). http://bricol.net/downloads/pubs/Green_et_al.2014.APPS.pdf

Hickey, L. and Wolfe, J.1975. "The Bases of Angiosperm Phylogeny: Vegetative Morphology." *Annals of the Missouri Botanical Garden*62(3): 538-589. http://www.jstor.org/stable/2395267?seq=1#page_scan_tab_contents

Price, C., Munro, P., and Weitz, J. "Estimates of Leaf Vein Density are Scale Dependent." 2014. *Plant Physiology* 164(1):173-180. http://www.plantphysiol.org/content/164/1/173.full

Roth-Kebelsick et al, 2001. "Evolution and Function of Leaf Venation Architecture: A Review."

Annals of Botany 87:533-566. http://aob.oxfordjournals.org.ezp-prod1.hul.harvard.edu/content/87/5/553.full.pdf+html

Sack, L. and Scoffoni, C. 2013. "Leaf venation: structure, function, development, evolution, ecology, and applications in the past and future." *New Phytologist* 198(4): 983-1000. http://onlinelibrary.wiley.com.ezp-prod1.hul.harvard.edu/doi/10.1111/nph.12253/abstract

Sack, L., and Holbrook, M. 2006. "Leaf Hydraulics." *Annual Review of Plant Botany.* 57: 361-381.http://apps.webofknowledge.com.ezp-prod1.hul.harvard.edu/full_record.do?product=WOS&search_mode=GeneralSearch&qid=9&SID=3AWf4bjccSYWZbcJtRu&page=1&doc=1

Strahler, A. N. 1957. "Quantitative analysis of watershed geomorphology", *Transactions of the American Geophysical Union* **38** (6): 913–920, doi:10.1029/tr038i006p00913.