Applications in Plant Sciences

3D segmentation of plant bud CT scans support the bud packing hypothesis

--Manuscript Draft--

Abstract:	Premise of the study: Three-dimensional imaging technology allow for a rich exploration of 3-d anatomy of organismal structures. Though quantitative studies require that different structures be readily distinguishable from each other, manual segmentation of high-resolution image stacks is a laborious task. Though off-the-shelf methods can automate segmentation for tissues of different appearance, these methods fail to segment tissues of similar grayscale intensity. Methods: We propose a pipeline composed of computer-vision and machine-learning techniques to analyze the anatomy of leaf buds across Viburnum. We demonstrate how the proposed method successfully distinguishes leaves from bud scales, a problem that available software packages fail to solve. After successful segmentation, we calculate various bud metrics (leaf-to-bud volume ratio, compactness, fractal dimension) and use phylogenetically independent contrasts to evaluate correlations between extracted metrics and species-level characteristics. Results: We find that bud compactness differs significantly based on leaf margin type (entire, lobed, toothed). Entire-leaved species have less leaf tissue (lower leaf-to-bud volume ratios) and leaf tissue is more tightly packed together (higher compactness values) relative to species with lobed or toothed leaves. Discussion: These results lend preliminary support to the bud packing hypothesis, indicating that the packing of leaves within buds correlates with mature leaf form.

	University of Vienna brian.metscher@univie.ac.at	
	Mason Earles UC Davis jmearles@ucdavis.edu	
Opposed Reviewers:		
Funding Information:	National Science Foundation (NSF IOS-1256706)	Dr Michael Donoghue
	National Science Foundation (NSF IOS-1257262)	Dr Erika Edwards

Dear Prof. Culley,

It is my pleasure to submit the manuscript titled '3D segmentation of plant bud CT scans support the bud packing hypothesis' for review to the Journal of Applications in Plant Sciences as an Application Article on behalf of my co-authors. My name is Dr. Andras Zsom and I am a Lead Data Scientist at Brown University in the Center for Computation and Visualization (CCV). My team at CCV are not biologists but data scientists and we help faculty members at Brown with their data-intensive research projects.

This project started roughly two years ago when Prof. Erika Edwards (then at Brown University currently at Yale) approached me because her team was struggling to segment microCT scans of plant buds into bud scale and leaf with open source or commercial image segmentation softwares (like ITK-SNAP). The main reason why those tools failed was because the scale and the leaf tissues have a similar intensity.

I was intrigued by the problem and I was happy to collaborate. Initially I tried to develop supervised machine learning models (neural networks) to automatically segment the buds based on tissue structural patterns. Ultimately these approaches were deemed inaccurate. After a couple of failed attempts, we found that the buds can be accurately segmented by a custom image processing pipeline that requires minimal user input. This approach and the preliminary results are described in the manuscript.

After successful segmentation, we calculate various bud metrics some of which are novel in the field to the best of my knowledge (e.g., fractal dimension and compactness). We then use phylogenetically independent contrasts to evaluate correlations between extracted metrics and species-level characteristics. Our results and findings lend preliminary support to the bud packing hypothesis, indicating that the packing of leaves within buds correlates with mature leaf form and suggesting new avenues of research.

I believe the manuscript is a good fit for the journal because we study the structure and development of leaves in the Viburnum genus. Our code is publicly available in a GitHub repository and we hope the repository will make this a tool researchers can actually use. The manuscript describes original research, no part of this manuscript is published elsewhere, and we have no conflicts to disclose.

Thanks you for your consideration!

Best wishes,
Andras Zsom
Lead Data Scientist
Center for Computation and Visualization
Brown University, Computing and Information Systems
3 Davol Square, Providence, RI 02903, USA
https://ccv.brown.edu/

Applications in Plant Sciences Author Agreement Form
Click on the first gray box and start typing. You may use the Tab to get to the next field.

Corresponding Author's Name: Andras Zsom
Date: July, 17th, 2019

Respond to ALL the statements below either by typing your initials or checking the appropriate box. After you have completed this form, save it to your desktop as a Word document, and upload it with your manuscript submission on the "Attach Files" page in Editorial Manager (http://www.edmgr.com/apps).

1. All Contributing Authors know of and concur with the submission of this manuscript. Single authors please also initial.
Initials: AZ

2. All authors of this research paper have directly participated in the planning, execution, or analysis of the study;
AND
All authors of this paper have read and approved the final version submitted.
Initials: AZ

3. The contents of this manuscript have not been copyrighted or published previously and are not now under consideration for publication elsewhere. The contents of this manuscript will not be copyrighted, submitted, or published elsewhere while acceptance by the Journal is under consideration.
Initials: AZ

4. *APPS* operates under a Creative Commons Attribution 3.0 license. Please specify which license should be applied to publication of this article:

CC-BY-NC-SA: This is the default license under which articles are published in *APPS*. Under the terms of this license, authors retain ownership of the copyright of their articles. Users are allowed to download, print out, extract, reuse, archive, and distribute the article for noncommercial use; this use is contingent upon appropriate credit being given to the authors and the source of the work and upon the new creation being licensed under the same terms.

CC-BY: This license should be selected if mandated by article funding sources. CC-BY permits use, distribution, and reproduction in any medium, provided the original work is properly cited. There is no distinction between reuse for commercial or noncommercial purposes.

CC0: This is the Creative Commons public domain license. Papers authored by one or more US government employees should use this license, which allows unlimited distribution and reuse of the article for any lawful purpose.

Each of these licenses ensure wide availability of the article and that the article can be included in any scientific archive. No permission is required from the author or the Botanical Society of

America. If a CC-BY-NC-SA license is used, those wishing to use material for commercial purposes should contact the author.

Please select the license that should be applied to this article:

☒ **CC-BY-NC-SA**

☐ **CC-BY**

☐ **CC0**

4a. If you select the CC0 license, please confirm the following: One or more authors of this manuscript are US government employees, and this work was prepared as part of the author's official duties as a US government employee. Note: If this work was not prepared as part of the author's official duties as a US government employee, it should be published under a CC-BY-NC-SA or CC-BY Creative Commons license.
Initials:

5. Authors are responsible for recognizing and disclosing any duality of interest that could be perceived to bias their work, acknowledging all financial support and any other personal connections.

☒**No**, there is no duality of interest that I should disclose, having read the above statement.

☐**Yes**, having read the above statement, there is potential duality of interest. This has been fully detailed in my cover letter.

6. Have the results/data/figures in this manuscript been published or are they under consideration for publication elsewhere?

☒**No**, the results/data/figures in this manuscript have not been published elsewhere, nor are they under consideration (from you or one of your Contributing Authors) by another publisher.

☐**Yes**, some portion of the results/data/figures in this manuscript has been published or is under consideration for publication elsewhere.

6a. If you select Yes, please identify results/data/figures taken from other published/pending manuscripts in the textbox below and explain why this does not constitute dual publication. (Note: The existence of pending or previously published articles that use or have used any of the same results presented in the submitted manuscript does not generally prejudice review and acceptance.)

7. All *APPS* papers are published as Open Access articles. Article Processing Charges (APCs) are as follows:

Non-BSA members: $1400 per article
BSA members: $800 per article
BSA members with an active membership for longer than two years: $450 per article

This policy applies to all articles accepted, except for invited articles for which the Editorial Board has agreed to waive fees at the time of article submission. **Authors will be invoiced and payment must be received before online publication.**

☐ **No,** No authors of this manuscript are members of the BSA.

☒ **Yes,** I confirm that at least one author is a BSA member

Please note here the name(s) of author(s) who are BSA members. Membership status will be verified by the editorial office.

Erika Edwards and Michael Donoghue

8. Authors are responsible for covering charges for author-initiated changes made in proof. Authors will be charged when more than 10 changes are requested per article, as well as for figure replacements in proof ($25.00 per figure). This policy applies to BSA members and nonmembers. I have read and understand this policy.
AZ

9. *APPS* requires that genetic information, such as DNA, RNA, or protein sequences, be submitted to an appropriate data bank (i.e., GenBank or EMBL). Alignments used to produce phylogenies must be submitted to TreeBase or to *APPS* to be published with the paper as supplementary material. Authors are encouraged to archive all sequences generated from next-generation sequencing techniques in a suitable public depository (e.g., the Sequence Read Archive of NCBI, the Sequence Read Archive of ENA, or Dryad). I have read and understand this policy.
AZ

[last revised 21 June 2017]

3D segmentation of plant bud CT scans

support the bud packing hypothesis

Andras Zsom[1], Center for Computation and Visualization, Brown University, Providence, RI

02912, andras_zsom@brown.edu

Joel Abraham, Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect

Street, New Haven, CT 06511 joel.abraham@yale.edu

Lucina Schwarz, Princeton University, lucinas@princeton.edu

Maria I Restrepo, Center for Computation and Visualization, Brown University, Providence RI

02912, maria_restrepo@brown.edu

David Chatelet, Biomedical Imaging Unit, University of Southampton, Southampton,

SO16 6YD, UK

Michael Donoghue, Department of Ecology and Evolutionary Biology, Yale University, 21

Sachem St, New Haven, CT 06511 michael.donoghue@yale.edu

Yannick Stadler, Department of Botany & Biodiversity Research, University of Vienna

Jürg Schönenberger, Department of Botany & Biodiversity Research, University of Vienna

Mark Howison, Watson Institute for International and Public Affairs, Brown University,

Providence, RI 02906 mhowison@ripl.org

Erika J. Edwards, Department of Ecology and Evolutionary Biology, Yale University, 165

Prospect Street, New Haven, CT 06511 erika.edwards@yale.edu

---

[1] Corresponding author

Abstract

Premise of the study: Three-dimensional imaging technology such as CT Scans allow for a rich exploration of the 3-d anatomy of organismal structures. Though quantitative studies require that different structures be readily distinguishable from each other, manual segmentation of high-resolution image stacks is a laborious task for humans. Though off-the-shelf methods can be used to automate segmentation for tissues of very different appearance, these methods fail to segment tissues of similar grayscale intensity.

Methods: We propose a pipeline composed of computer-vision and machine-learning techniques to analyze the anatomy of leaf buds across *Viburnum*, a morphologically diverse clade of woody plants. We demonstrate how the proposed method successfully distinguishes leaves from bud scales, a problem that available software packages fail to solve. Our approach is general, makes no strong assumptions about the underlying data, and is therefore extensible to other segmentation problems in bioinformatics. After successful segmentation, we calculate various bud metrics (leaf-to-bud volume ratio, compactness, fractal dimension) and use phylogenetically independent contrasts to evaluate correlations between extracted metrics and species-level characteristics.

Results: We find that bud compactness differs significantly based on leaf margin type (entire, lobed, toothed). Entire-leaved species have less leaf tissue (lower leaf-to-bud volume ratios) and leaf tissue is more tightly packed together (higher compactness values) relative to species with lobed or toothed leaves.

Discussion: These results lend preliminary support to the bud packing hypothesis, indicating that the packing of leaves within buds correlates with mature leaf form and suggesting new avenues of research.

Availability: https://github.com/brown-ccv/bud_segment

Introduction

Advances in three-dimensional imaging technology allow for a rich exploration of the internal anatomy of organisms, in a way that has never before been possible. A particularly promising imaging technology is Computed Tomography, commonly referred to as CT scanning. CT scanning is a low-cost X-ray imaging technique that produces many cross-sectional images, or slices, of a scanned object (tomography). The technique allows for the non-invasive exploration of the interior of an object, and thus CT scans have been applied widely in medicine and industry. Recent improvements in CT technology have increased the applicability of this technique to other disciplines; for instance, utilizing a new, high-resolution CT scanning technology -- micro-Computed Tomography (μCT) -- it is possible to monitor the growth of root systems in the soil through time in a non-damaging way (Metzner et al. 2015; Lafond et al. 2015). This high-throughput and non-destructive technology is rapidly opening up new areas of investigation, elucidating long-standing biological questions, and transforming our understanding of the natural world.

However, for these advances to be realized, the technologies used to process CT scans must be able to readily separate distinct anatomical structures, a process known as segmentation. Thus far, segmentation has largely been done manually, where the images are divided into different tissue types by hand. However, manual segmentation of the high-resolution image-stacks that result from CT scanning is a laborious and time consuming task. Because segmentation is a major bottleneck in the CT post-processing methodology and represents the greatest barrier to widespread applicability of CT scanning, many techniques have been proposed for automated segmentation, particularly within the realm of medical and brain imaging (Sharma and Aggarwal 2010; Jain et al. 2010)

Most of the classical segmentation algorithms are simply generalizations of classical segmentation techniques for regular images, and are based on gray-level features of the CT scan voxels, or 3-dimensional (volumetric) pixels. Segmentation techniques can be divided into three general categories: (i) segmentation based on histogram-features; (ii) edge-based segmentation; and (iii) region-based segmentation. The first class of algorithms (i) apply gray-level thresholds to an image to separate regions of interest. The optimal threshold can be found using global or local histograms of gray-level intensities (Ramesh 1995; Lai 2006; Brink 1995). If the contrast between different objects is high, these algorithms can be sufficient and successful. However, in practice, finding an effective threshold that consistently separates different tissue types may not be possible. Edge-based algorithms (ii) have been developed to circumvent some of this difficulty. These algorithms operate in two stages. In the first stage, the boundaries between objects are detected using gradients in intensity, color, or texture (Sobel and Feldman 1968; Marr and Hildreth 1980; Sobel et al. 1980), the Canny Edge Detector (Canny 1986), or more sophisticated methods using Markov Random Fields (Geman and Geman 1984; Li 1994), Graph Cuts, and Level Sets (Osher and Fedkiw 2001; Jin and Angelini 2005). In the second stage the final segmentation is labeled using connectivity properties of the voxels. Finally, the third class of algorithms (iii) mostly consist of clustering together pixels or voxels of similar properties and applying a series of splitting and merging techniques to refine the clusters (Sonka et al. 1993).

Most of the aforementioned methods can be utilized effectively when segmenting structures that are very different in appearance, but are less stable when segmenting tissues of similar color intensity -- a common problem for biological tissues where the x-ray values of different tissues or organs in µCT scans often overlap. Low-contrast or intensity-homogeneity between objects is sometimes handled by modifying the imaging condition; for example, when segmenting root systems, adjusting soil humidity can optimize contrast between root tissue and the surrounding soil (Metzner et al. 2015). However, it is not always possible to achieve contrast between tissues

simply by modifying the imagine conditions. Thus, attempts have been made to modify classical techniques in order to segment organic tissues of similar contrast. These techniques, however, tend to be very specific to the application, highly susceptible to algorithm parameters, and highly dependent upon the imaging conditions and composition of the anatomical structures being segmented. For instance, when segmenting images of a thorax region, an algorithm needs to handle the motion artifacts that result from breathing (Sharma and Aggarwal 2010). Thus, there is a pressing need for an automated segmentation technique that is both widely-applicable and flexible to tissue-specific restrictions.

The resting buds of plants represent an ideal system for developing such a methodology, because the distinct tissues inside developing buds have complex outlines, are often overlapping and interdigitated with one another, and are low-contrast with respect to each other, precluding the use of off-the-shelf segmentation techniques. Yet, as noted above, these difficulties are pervasive within biology; most biological structures are of similar contrast and have complex outlines. Thus, a methodology designed for segmenting buds would be widely-applicable and could be readily modified to deal with system-specific difficulties.

The internal structure of plant resting buds is biologically important, as well. Buds likely play a role in determining the shape of the organs within them, as the limitations associated with developing within a bounded space may impose constraints on adult leaf form (Couturier et al. 2012). This idea forms the basis of the bud packing hypothesis, which suggests that selection for efficient packing of leaves within overwintering buds may explain the widely-observed latitudinal gradient in leaf form, with complex leaf shapes over-represented in deciduous woody plants of colder climates (Edwards et al. 2016; Edwards et al. 2017). It has long been observed that lobed and toothed leaves are prevalent among temperate plant species (Bailey and Sinnott 1916; Little et al. 2010; Peppe et al. 2011) but mechanistic explanations of this pattern have been elusive

(Edwards et al. 2017). The bud packing hypothesis posits that lobed and toothed leaves relate to the increased preformation of leaves inside the resting buds of temperate species. The efficient packing of leaf tissue within buds (Couturier et al. 2011) and the more rapid unfolding of leaves in the spring (Kobayashi et al. 1998) may provide a competitive advantage to species occupying highly-seasonal environments (Edwards et al. 2016; Edwards et al. 2017). We have not been able to empirically evaluate this and other hypotheses regarding bud development, due largely to methodological limitations; precise measurements of bud packing characteristics are difficult and imprecise using traditional dissections. A technique that allows for the application of μCT scanning to plant buds would provide access to the internal structure of buds and provide new information about how bud development and preformation influences leaf form, while also being widely-applicable across disciplines.

Here, we present an analytical pipeline composed of computer-vision techniques with minimal manual input that distinguishes between tissues of similar contrast, and we apply it to developing plant buds. This approach exploits the structure of the local geometry in images derived from μCT scanning to analyze the buds of *Viburnum*, a clade of woody plants that spans tropical-temperate mesic forest biomes and exhibits great diversity in leaf form (Schmerler et al. 2012; Spriggs et al. 2017). Specifically, we demonstrate how the proposed method can be used to segment developing leaves from bud scales, two tissues with similar color intensity, addressing a problem that available software packages fail to solve.

Methods

Manual segmentation of high-resolution image stacks is a laborious and time consuming task for humans. Our work seeks to minimize human effort while accurately segmenting 3D image stacks of buds into leaf, bud scale, and background using an interpolation-based approach. While the algorithm was specifically developed for the described task, it is extensible to any segmentation

problem where the various parts can be separated by closed contours in 2D image slices. Since the pipeline requires user interaction, it is only of practical use when the number of objects to segment is not very large (100 or less per user). The source code is publicly available on GitHub[2].

Bud collection and CT scanning

Buds were collected from 39 plants representing 32 species of *Viburnum* (Adoxaceae), a well studied model clade of plants (e.g., Schmerler et al. 2012; Spriggs et al. 2015; Scoffoni et al. 2016; Spriggs et al. 2018). *Viburnum* spans a gradient of mesic temperate to mesic tropical environments, and features multiple transitions from the tropics into temperate zones (Spriggs et al. 2015). These tropical-temperate transitions have been marked by dramatic phenotypic changes in leaf form that parallel broader latitudinal patterns in leaf form (Schmerler et al. 2012; Bailey and Sinnott 1916): tropical *Viburnum* species tend to have elliptical and elongate leaves with entire margins, whereas temperate species tend to have more orbicular or rounded leaves with lobes or toothed margins (Schmerler et al. 2012; Spriggs et al., 2018). These repeated changes in leaf phenotype within *Viburnum* present an opportunity to test whether bud characteristics correlate with adult leaf characteristics.

We collected buds from several tropical or warm temperate forests of Asia (*V. clemensiae, V. glabberimum*, and *V. awabuki*), and from 30 deciduous *Viburnum* species representing a variety of bud characteristics in the Arnold Arboretum of Harvard University. Buds collected in the Arnold Arboretum were fixed in FAA, whereas the Asian accessions that were field collected were first preserved in 70% EtOH and were later fixed in FAA. Our species sampling captured multiple independent cold temperate origins, and provided six clear phylogenetic independent contrasts between tropical and temperate lineages (Fig. 3)

---

[2] https://github.com/brown-ccv/bud_segment

Before CT scanning, the buds were stained for two weeks in 1% phosphotungstic acid in FAA (w/vol) (Staedler et al. 2013). They were then transferred to a solution of 1% phosphotungstic acid in 96% Ethanol (w/vol) and left to soak overnight. Finally, the they were transferred to acetone for 30 minutes and were then critical point-dried with an autosamdri®-815 (tousimis res. corp.). The buds were then mounted on aluminum stubs with epoxy glue (UHU PLUS®), and scanned in a MicroXCT-200 system from Zeiss Microscopy. This methodology resulted in a stack of 728 high-resolution cross-sectional images of each individual bud.

Image stack segmentation

Once the image stacks were generated as above, segmentation could begin. All steps of the segmentation process are done in Python (V 3.7). The first step in the process of bud segmentation consists of manually marking the boundary between the leaf and the bud scale. The process is simple and intuitive. A frame is displayed and the user clicks ~20-30 points along the boundary (an illustration is shown in Fig. 1, left). We found that the user needs to click through 40 frames per bud to accurately segment buds that have small leaves and the user also needs minimal training to identify the different structures. The process took roughly 5 minutes per bud.

During the second step, background voxels were separated from plant tissue. We used k-means clustering to group the voxels into two categories based on their intensity. The dark intensity voxels were designated as empty space (no plant tissue) and the bright intensity voxels as plant tissue (either leaf or bud scale; they are not distinguishable at this stage).

In the last step of the pipeline, the user's clicks were interpolated into closed curves that separated the leaf from the bud scale. Interpolation itself can be broken down into two stages. In the first, a third order periodic spline interpolation (Dierckx 1982) was performed on each of the 40 clicked

frames to produce a smooth (differentiable) and closed contour (the yellow line in Fig. 1, left). Then, a non-periodic third order spline interpolation was performed along the axis of the scan (the two yellow lines in Fig. 1, middle). In contrast to periodic interpolation, a non-periodic interpolation returns a non-differentiable open contour where the first and the last points are not connected. The final outcome was the boundary contour on each frame or the boundary surface which separated plant tissue voxels inside and outside the contour (leaf and bud scale, respectively). The final segmented bud is illustrated in Fig. 1, right.

Bud feature derivations and analysis

After the buds were segmented into bud scale and leaf, we quantified various properties of the buds and the leaves developing inside. We first calculated straightforward features, like the volume of the entire bud and the volume of the leaves developing inside, from which we derived the leaf-to-bud ratio, the fraction of leaf volume inside the bud to the volume of the entire bud. We also calculated some more complex features, such as the 3-d compactness of the bud and the fractal dimension of the developing leaves, both of which are described below.

We first calculated the 3-d compactness of the buds. In the area of computer vision, 3-d compactness (Martinez-Ortiz 2010) is commonly used to measure how topologically similar 3-d shapes are to a compact sphere. 3-d compactness is defined by

$$C(S) = \frac{3^{5/3}}{5(4\pi)^{2/3}} \frac{\mu_{0,0,0}(S)^{5/3}}{\mu_{2,0,0}(S) + \mu_{0,2,0}(S) + \mu_{0,0,2}(S)},$$

where $C(S)$ is the compactness of a shape S, and $\mu_{x,y,z}$ are central moments, or weighted averages of voxel intensities calculated using 'skimage' (Kilian 2001). Compactness ranges between 0 and 1, with 1 being most similar to a compact sphere.

We then computed the fractal dimension of the developing leaves inside of the buds. The fractal dimension (or mass dimension) describes how the leaf voxels are distributed radially around the leaf's center of mass and it is between 1 and 3. A straight line has a fractal dimension of 1, a uniform surface has a fractal dimension of 2, and a uniform sphere has a fractal dimension of 3. A snowflake or any dissected planar surface would have a fractal dimension between 1 and 2, and any 3-d object that is not perfectly spherical and uniform would have a fractal dimension between 2 and 3. In order to compute the fractal dimension, we first determined the leaf's center of mass, then we counted how many leaf voxels were present within concentric spheres about this center of mass. The radii of the concentric spheres was increased by one voxel consecutively outward from the center to the leaf voxel furthest from the center of mass. The leaf voxel count as a function of the sphere radius is X(r). We fitted a function of the form cr^D to the inner third of the voxel count X(r), where c is a scaling constant, and D is the fractal dimension. The outer two thirds of X(r) were discarded because most of the volume was outside of the leaf and thus did not represent the mass distribution within the leaf (see Fig. 2 for an example in 2-d).

Once these metrics were computed, data were outputted for further analysis in R (V 3.5.1) using the package 'ape', which allows for analysis of data with phylogenetic structure (Paradis et al. 2004). First, we subsetted the dataset for only the species with bud scales, as the metrics calculated above are only relevant for those buds with bud scales. This generated a subset (n=23) of the 32 species included in this study. We scanned multiple buds of some species and averaged the values for those species. Then we modeled with a linear model the various bud characteristics (leaf volume, bud compactness, leaf-to-bud ratio, and fractal dimension) on species-level traits (here, leaf margin type) individually. We then used analysis of variance (ANOVA) in order to determine the direction and significance of the relationships between bud characteristics and these species traits. To control for the non-independence of phylogenetically structured data, we

performed phylogenetically independent contrasts (PICs) for all the above traits. The PICs of all the traits were again modeled with linear models and ANOVAs were again used to ensure that the relationships remained significant when controlling for phylogenetic non-independence of samples.

Results

The data collection and our segmentation pipeline are described in the Methods section along with the leaf and bud metrics that we extract after the segmentation is performed. We give only a brief overview of the methods here before we describe the results of our analyses and the correlations between the extracted metrics and species-level characteristics. The goal of the segmentation algorithm is to minimize human effort with little to no loss of segmentation accuracy as assessed by manual inspection. The user manually marks the boundary between the leaf and bud tissue by clicking on a displayed cross-sectional image of the stack. The process takes *ca.* 5 minutes per bud and the user clicks are interpolated to separate the leaf from the bud scale. Once segmentation is complete, we calculate bud metrics: leaf-to-bud volume ratio, the leaf tissue volume relative to the volume of the entire bud; compactness, a measure of how topologically similar the developing leaf tissue is to a sphere; and fractal dimension, a measure of the radial distribution of leaf mass tissue.

The bud metrics we calculated here differed significantly based on species-level leaf characteristics, particularly leaf margin type. Both lobed and toothed leaves had significantly higher leaf-to-bud ratios than did entire leaves (Fig. 4a; $F_{2,20} = 14.75$, $p < 0.005$), indicating that there is significantly more leaf tissue in the buds of toothed and lobed *Viburnum* species. These relationships were even more significant when phylogenetically independent contrasts (PICs) of these traits were used ($F_{1,20} = 59.69$, $p < 0.0001$). Also, buds of lobed and toothed leaves were significantly less compact than buds with entire leaves (Fig. 4b; $F_{2,20} = 3.889$, $p < 0.05$). This

relationship was likewise more significant when PICs were used ($F_{1,20}$ = 8.256, p < 0.001). As compactness essentially reflects how similar the leaf mass is to a solid sphere, the lower compactness of toothed and lobed leaves developing inside buds implies that the leaf tissue is more dispersed throughout the bud. There was no difference in fractal dimension between the different leaf margin types, with PICs ($F_{1,20}$=0.1411, p = 0.7112) or without (Fig. 4c; $F_{2,20}$=0.1921, p = 0.8267).

Discussions

We present, to our knowledge, the first "in-silico" analysis of resting bud anatomy, and demonstrate the power of non-destructive visualization in understanding the arrangement and filling of bud volume *in-situ*. Our results, though preliminary, lend support to the bud packing hypothesis. This hypothesis suggests that the complex leaf shapes observed in many temperate plant species may be due in large part to their early development inside the constrained space of a resting bud, which yields complicated patterns of folding of the leaf primordia and boundaries between tissues that would not exist in neo-formed leaves that developed outside of a bud later in the growing season (Edwards et al. 2016; Edwards et al. 2017). A key assumption of this hypothesis is that, during the transition to a temperate, deciduous habit, species evolved to pack more tissue inside of their buds. If true, lineages in the temperate zone should therefore exhibit more leaf development inside of their buds as compared to their tropical relatives. Here we found that the buds of the temperate species that we examined, which have toothed or lobed leaves, did have a significantly higher volume of leaf tissue packed within than did the three tropical species we measured with entire leaf margins. The development of more leaf tissue within buds suggests that leaves will mature more rapidly when the buds open in the spring (Kobayashi et al. 1998; Edwards et al. 2016; Edwards et al. 2017).

Leaf-to-bud volumetric ratio is but one measure of packing efficiency. We also found that lobed and toothed leaves were significantly less compact in bud than entire-margined leaves. Counter-intuitively, this result is also consistent with the bud packing hypothesis. Compactness measures how similar the leaf mass is to a solid sphere, or how tightly aggregated the leaf mass is inside the bud. The higher compactness values of entire-margined leaf buds are a product of the fact that these buds contain hardly any leaf tissue inside them. The leaf primordia that they do contain are not very well developed and densely aggregated at the center of the bud. Thus, the lower compactness values of the lobed and toothed *Viburnum* species reflect the fact that the leaves inside their buds are comparatively more developed, and also folded in complex ways (Couturier et al. 2011; Kobayashi et al. 1998; Edwards et al. 2016; Edwards et al. 2017).

The relatively few samples processed in this first experiment provide a "proof of concept" and preliminary support for the bud packing hypothesis. Now that we are confident in the methods presented here, these results can be tested with orders of magnitude more data in *Viburnum* and in many other relevant plant groups.

The bud packing hypothesis highlights but one way in which bud development relates to plant form and might influence evolutionary trajectories (Lubbock 1899; Couturier et al. 2011). Yet, few studies of plant morphology have explicitly considered the potential influence of bud structure and development on plant form and function. This lack of attention is at least partly due to methodological difficulties associated with extracting quantitative characteristics from buds; traditional dissection methods are time-consuming and destructive. The pipeline outlined here makes analyzing plant buds much more straightforward, and opens up exciting new avenues in the study of plant development and evolution. Furthermore, we believe that this pipeline can be widely utilized for segmentation across diverse biological systems.

Literature Cited

Bailey, I.W. and Sinnott, E.W. 1916. The climatic distribution of certain types of angiosperm leaves. *American Journal of Botany* 3(1), pp. 24–39.

Brink, A.D. 1995. Minimum spatial entropy threshold selection. *IEE Proceedings - Vision, Image, and Signal Processing* 142(3), p. 128.

Canny, J. 1986. A Computational Approach to Edge Detection. *IEEE transactions on pattern analysis and machine intelligence* PAMI-8(6), pp. 679–698.

Clement, W., M. Arakaki, P. Sweeney, E. J. Edwards, and M. J. Donoghue. 2014. A chloroplast tree for *Viburnum*(Adoxaceae) and its implications for phylogenetic classification and character evolution. Amer. J. Bot. 101: 1029-1049.

Couturier, E., Brunel, N., Douady, S. and Nakayama, N. 2012. Abaxial growth and steric constraints guide leaf folding and shape in Acer pseudoplatanus. *American Journal of Botany* 99(8), pp. 1289–1299.

Couturier, E., du Pont, S.C. and Douady, S. 2011. The filling law: a general framework for leaf folding and its consequences on leaf shape diversity. *Journal of Theoretical Biology* 289, pp. 47–64.

Dierckx, P. 1982. Algorithms for smoothing data with periodic and parametric splines. *Computer Graphics and Image Processing* 20(2), pp. 171–184.

Eaton, D. A. R., E. L. Spriggs, B. Park, and M. J. Donoghue. 2017. Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. Syst. Biol. 66: 399-412.

Edwards, E.J., Chatelet, D.S., Spriggs, E.L., Johnson, E.S., Schlutius, C. and Donoghue, M.J. 2017. Correlation, causation, and the evolution of leaf teeth: A reply to Givnish and Kriebel.

American Journal of Botany 104(4), pp. 509–515.

Edwards, E.J., Spriggs, E.L., Chatelet, D.S. and Donoghue, M.J. 2016. Unpacking a century-old mystery: Winter buds and the latitudinal gradient in leaf form. *American Journal of Botany* 103(6), pp. 975–978.

Geman, S. and Geman, D. 1984. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE transactions on pattern analysis and machine intelligence* 6(6), pp. 721–741.

Jain, V., Seung, H.S. and Turaga, S.C. 2010. Machines that learn to segment images: a crucial technology for connectomics. *Current Opinion in Neurobiology* 20(5), pp. 653–666.

Jin, Y. and Angelini, E.D. 2005. AF Laine in Handbook of Medical Image Analysis: Advanced Segmentation and Registration Models.

Kilian, J. 2001. Simple Image Analysis By Moments. 2. Available at: http://breckon.eu/toby/teaching/dip/opencv/SimpleImageAnalysisbyMoments.pdf [Accessed: 9 April 2019].

Kobayashi, H., Kresling, B. and Vincent, J.F.V. 1998. The geometry of unfolding tree leaves. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 265(1391), pp. 147–154.

Lafond, J.A., Han, L. and Dutilleul, P. 2015. Concepts and analyses in the CT scanning of root systems and leaf canopies: A timely summary. *Frontiers in plant science* 6, p. 1111.

Lai, C.C. 2006. A novel image segmentation approach based on particle swarm optimization. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* E89–A(1), pp. 324–327.

Li, S.Z. 1994. Markov random field models in computer vision. In: Eklundh, J.-O. ed. *Computer vision — ECCV '94*. Berlin/Heidelberg: Springer-Verlag, pp. 361–370.

Little, S.A., Kembel, S.W. and Wilf, P. 2010. Paleotemperature proxies from leaf fossils reinterpreted in light of evolutionary history. *Plos One* 5(12), p. e15161.

Lubbock, J. 1899. *On buds and stipules.* London: K. Paul, Trench, Trübner.

Marr, D. and Hildreth, E. 1980. Theory of edge detection. *Proceedings of the Royal Society B: Biological Sciences* 207(1167), pp. 187–217.

Martinez-Ortiz, C.A. 2010. 2D and 3D Shape Descriptors.

Metzner, R., Eggert, A., van Dusschoten, D., Pflugfelder, D., Gerth, S., Schurr, U., Uhlmann, N. and Jahnke, S. 2015. Direct comparison of MRI and X-ray CT technologies for 3D imaging of root systems in soil: potential and challenges for root trait quantification. *Plant Methods* 11, p. 17.

Osher, S. and Fedkiw, R.P. 2001. Level set methods: an overview and some recent results. *Journal of Computational Physics* 169(2), pp. 463–502.

Paradis, E., Claude, J. and Strimmer, K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20(2), pp. 289–290.

Peppe, D.J., Royer, D.L., Cariglino, B., Oliver, S.Y., Newman, S., Leight, E., Enikolopov, G., Fernandez-Burgos, M., Herrera, F., Adams, J.M., Correa, E., Currano, E.D., Erickson, J.M., Hinojosa, L.F., Hoganson, J.W., Iglesias, A., Jaramillo, C.A., Johnson, K.R., Jordan, G.J., Kraft, N.J.B. and Wright, I.J. 2011. Sensitivity of leaf size and shape to climate: global patterns and paleoclimatic applications. *The New Phytologist* 190(3), pp. 724–739.

Ramesh, N. 1995. Thresholding based on histogram approximation. *IEE Proceedings - Vision, Image, and Signal Processing* 142(5), p. 271.

Scoffoni, C., D. S. Chatelet, J. Pasquet-kok, M. Rawls, M. J. Donoghue, E. J. Edwards, L. Sack. 2016. Hydraulic basis for the evolution photosynthetic productivity. Nature Plants 2 (DOI: 10.1038/NPLANTS.2016.72)

Schmerler, S.B., Clement, W.L., Beaulieu, J.M., Chatelet, D.S., Sack, L., Donoghue, M.J. and Edwards, E.J. 2012. Evolution of leaf form correlates with tropical-temperate transitions in Viburnum (Adoxaceae). *Proceedings. Biological Sciences / the Royal Society* 279(1744), pp. 3905–3913.

Sharma, N. and Aggarwal, L.M. 2010. Automated medical image segmentation techniques. *Journal of medical physics / Association of Medical Physicists of India* 35(1), pp. 3–14.

Sobel, I. and Feldman, G. 1968. An Isotropic 3x3 Image Gradient Operator.

Sobel, I., Levinthal, C. and Macagno, E.R. 1980. Special techniques for the automatic computer reconstruction of neuronal structures. *Annual review of biophysics and bioengineering* 9, pp. 347–362.

Sonka, M., Hlavac, V. and Boyle, R. 1993. *Image processing, analysis and machine vision*. Boston, MA: Springer US.

Spriggs, E.L., Clement, W.L., Sweeney, P.W., Madriñán, S., Edwards, E.J. and Donoghue, M.J. 2015. Temperate radiations and dying embers of a tropical past: the diversification of Viburnum. *The New Phytologist* 207(2), pp. 340–354.

Spriggs, E.L., Schmerler, S.B., Edwards, E.J. and Donoghue, M.J. 2018. Leaf Form Evolution in Viburnum Parallels Variation within Individual Plants. *The American Naturalist* 191(2), pp. 235–249.

Staedler, Y.M., Masson, D. and Schönenberger, J. 2013. Plant tissues in 3D via X-ray tomography: simple contrasting methods allow high resolution imaging. *Plos One* 8(9), p. E75295.
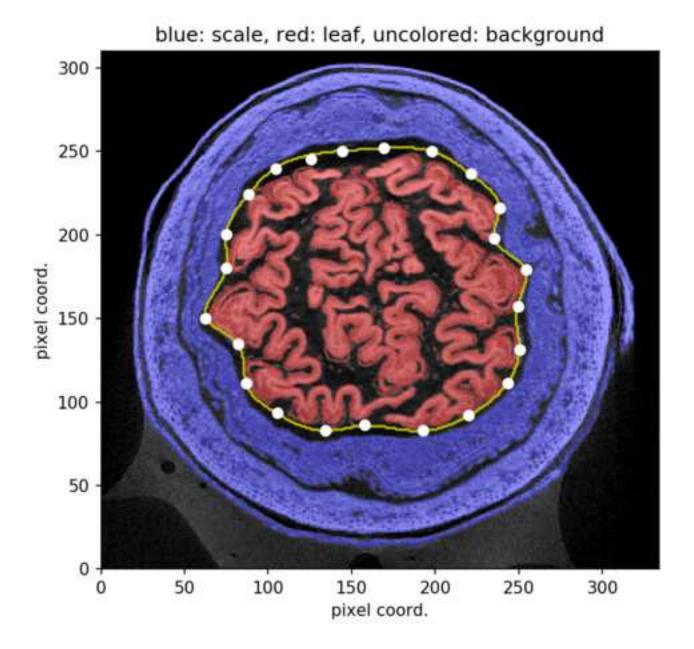
Figure titles

Fig. 1. Left: One segmented frame along the axis of scan. The white dots are the user's clicks, the yellow solid line is the interpolated, smooth boundary. Pixels shaded blue and red are bud scale and leaf pixels, respectively. Unshaded (dark grey) pixels are background. **Middle**: One segmented frame parallel to the axis of scan. The thin white lines depict frames that the user clicked through, the two yellow solid lines are the interpolated non-periodic boundaries along the axis of scan. **Right**: 3D representation of the same bud. The background is transparent, the bud scale is semi-transparent green, and the leaf is non-transparent green. The red dots are the user's clicks.

Fig. 2. The method to calculate fractal dimension is illustrated in 2-d. The actual values described in the Results section are calculated in 3-d. **Left**: One frame parallel to the axis of scan is shown. The white areas are leaf pixels, the red dot is the center of mass of the leaf pixels, the red circles are a few of the concentric circles used to calculate the pixel count as a function of radius ($X(r)$ in 2-d). The circles inside the thick red circle (the inner third) are used to estimate the fractal dimension. **Right**: The blue line is $X(r)$, the thick red line marks the inner third region, and the black dashed line is the fit to the inner third of $X(r)$. The fit has a form of cr^D, where c is a scaling constant, and D is the fractal dimension. It is 1.86 in the 2-d example.

Fig. 3. Phylogenetic relationships of the *Viburnum* species used in our analyses. This tree was obtained by pruning a tree for 110 (of a total of ~165) species of *Viburnum* based on chloroplast DNA (Clement et al. 2014) and nuclear RAD-seq data (Eaton et al. 2017), with branch lengths for our PIC analyses based on the chloroplast DNA (Landis et al., in prep.). The reconstruction shown here for biome state (tropical versus temperate) is based on the full tree. "Temperate" refers to cold temperate climates, which experience prolonged freezing temperatures during the winter;

"Tropical" includes both tropical and warm temperate (lucidophyllous) forests, which rarely experience freezing (Edwards et al. 2017).   The independent transitions to the temperate zone that are captured by our sampling are indicated on the phylogeny by the numbers in circles; leaf margin type is indicated by the color of the circles at the tips, and the sizes of the circles correspond to standardized values of leaf-to-bud ratio, compactness, and fractal dimension.

Fig. 4. (a) Leaf-to-bud ratio, (b) bud compactness, and (c) fractal dimension by leaf margin type. Letters indicate significant difference between margin types.

blue: scale, red: leaf, uncolored: background

blue: scale, red: leaf, uncolored: background