Supplement

Species composition

Table S1. Species composition for the four datasets. There are 8 genus level classes and 2 sub-generic classes (highlighted rows).

Class label	LD-train	LD-test	XT-train	XT-test
Celtis	*Celtis occidentalis*	*Celtis laevigata* *Celtis occidentalis* *Celtis reticulata*	*Celtis laevigata* *Celtis occidentalis* *Celtis reticulata*	*Celtis occidentalis* *Celtis* sp.
Fraxinus	*Fraxinus americana*	*Fraxinus pennsylvanica* *Fraxinus quadrangulata*	*Fraxinus americana* *Fraxinus nigra* *Fraxinus oregona* *Fraxinus pennsylvanica* *Fraxinus quadrangulata*	*Fraxinus nigra* *Fraxinus pennsylvanica* *Fraxinus quadrangulata*
Gleditsia	*Gleditsia triacanthos*	*Gleditsia aquatica* *Gleditsia triacanthos*	*Gleditsia aquatica* *Gleditsia triacanthos*	*Gleditsia triacanthos*
Maclura	*Maclura pomifera*	*Maclura pomifera*	*Maclura pomifera*	*Maclura pomifera* *Maclura* sp.
Morus	*Morus rubra*	*Morus rubra*	*Morus alba* *Morus rubra*	*Morus alba* *Morus rubra*
QuercusW	*Quercus alba*	*Quercus alba* *Quercus garryana*	*Quercus alba* *Quercus garryana*	*Quercus alba*
QuercusR	*Quercus rubra*	*Quercus rubra*	*Quercus falcata* *Quercus rubra*	*Quercus falcata*
Robinia	*Robinia pseudoacacia*	*Robinia pseudoacacia*	*Robinia neo-mexicana* *Robinia pseudoacacia*	*Robinia pseudoacacia*
Sassafras	*Sassafras albidum*	*Sassafras albidum*	*Sassafras albidum*	*Sassafras albidum* *Sassafras* sp.
Ulmus	*Ulmus americana*	*Ulmus americana* *Ulmus rubra*	*Ulmus americana* *Ulmus rubra*	*Ulmus americana*

Download link for LD-train: The dataset was downloaded on August 13, 2020, from https://ir.library.msstate.edu/handle/11668/18480. This link was invalid when the pre-publication proof was prepared. The current link for downloading LD-train is https://scholarsjunction.msstate.edu/cfr-publications/5/ (accessed 17 April 2022).

LD-test details

Wood specimens: Five specimens per class were selected from the population of 219 specimens used for XT-train (from the MADw and SJRw xylaria at the USDA Forest Products Laboratory) for imaging.

Imaging system: Wiedenhoeft (2020) identified the 14x Olloclip as comparable hardware for the imaging setup of Lopes et al. 2020 and tested the optical properties of the Olloclip lens system for comparison to the XyloTron and XyloPhone platforms. The Olloclip configuration used in that paper (Olloclip lens + XyloPhone white light LED array) was used to collect images at 1x zoom using the wide-angle camera on an iPhone XSMax to compare to the LD-train images. LD-test images were collected using the same optical hardware mounted to an iPhone XR. The iPhone XR's single rear facing camera is identical to the wide angle camera used on the XSMax.

Establishing image magnification: One specimen (from those selected for XT-test) of each of the species reported in Lopes et al. (2020) was selected and imaged with the above hardware, and then those images were compared, per taxon, to LD-train images with a similar number of growth rings per image. (Figure F.1). Individual earlywood vessels were measured in pixels in both the LD-train images and the LD-test images, and then vessel diameters were compared. This method allowed for a biologically informed attempt to normalize magnification by comparing cell metrics from similarly vigorous growth. This method was complicated by the lack of specimen preparation in the LD-train dataset, so the metrics must be considered approximate.

Establishing spatial resolution: Images of an Empire stainless steel flat ruler were captured and linear resolution per pixel was calculated across the central, least distorted portion of the image. Spatial resolution is thus reported for the best-performing portion of the image – substantial distortion imparted by the lens yields a disparity between calculated image area and actual area captured. Accuracy of the ruler's divisions was confirmed with a caliper micrometer.

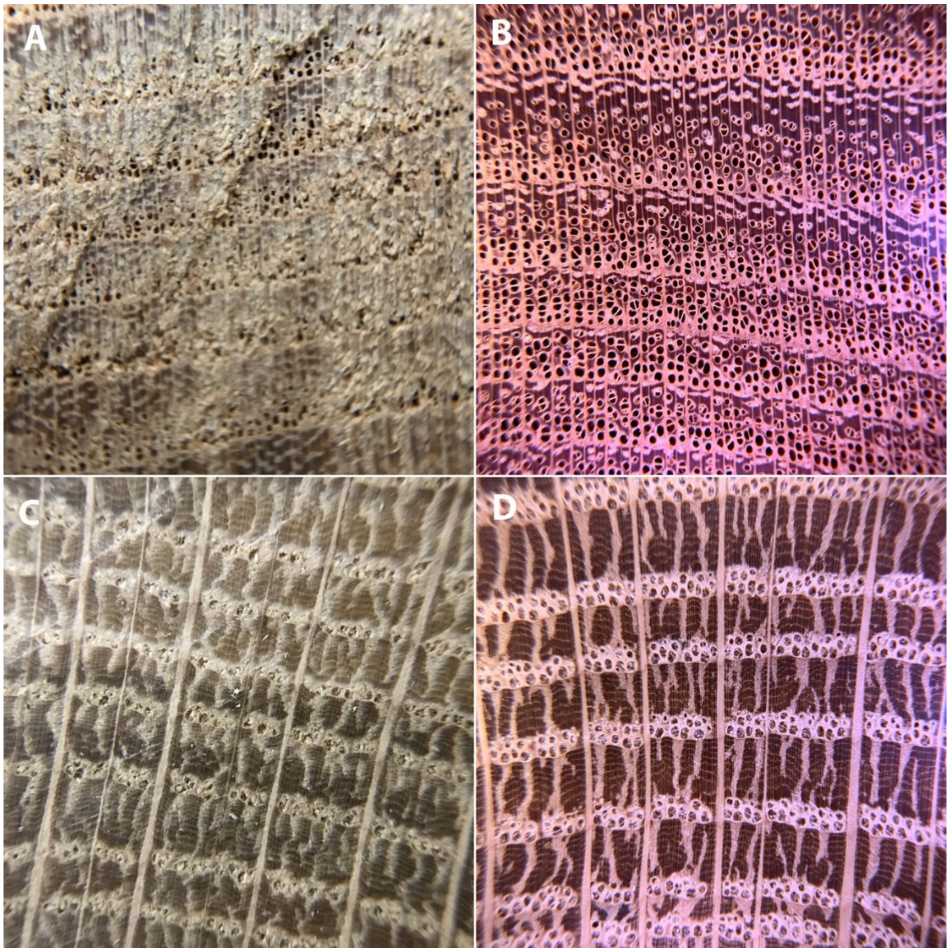

Figure S1. Side-by-side comparison of example LD-train images (A, C) and corresponding calibration images for the LD-test dataset (B, D). A and B are class Gleditsia, and C and D are class QuercusW. Note the distortion and spherical aberration in all images, but the anatomical detail readily observable in B, D.

Results and Discussion:

Image magnification: Table T.2 presents the average relative proportion of LD-train images /LD-test images – ranging from 0.843 to 1.111, with an average of 0.964. The imaging hardware used LD-test established a consistent distance between the specimen and the lens for each image, and also ensured perpendicularity between the wood surface and the light path, but these considerations were not reported for LD-train, and this, along with some biological variability, might account for some of the slight differences in apparent magnification.

Table S2. Class-wise proportional magnification comparing earlywood vessel diameters (LD-train image / LD-test image).

LD Class	Proportion
Celtis	0.987
Fraxinus	1.111
Gleditsia	0.952
Maclura	0.882
Morus	0.917
QuercusR	1.108
QuercusW	1.069
Robinia	0.918
Sassafras	0.857
Ulmus	0.843
Average	0.964

Spatial resolution: The LD-test dataset images have a pixel resolution of 4.89μm per pixel, and a calculated image dimension of 14.79mm per side, but the actual image dimensions are closer to 14.3mm on a side, as a result of image distortion.

Montages of central patches for the XT-train

Figure S2. Montage of central patches for class Celtis in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/1ijwmcur0iho0qccbqiscazyuf8y15ow

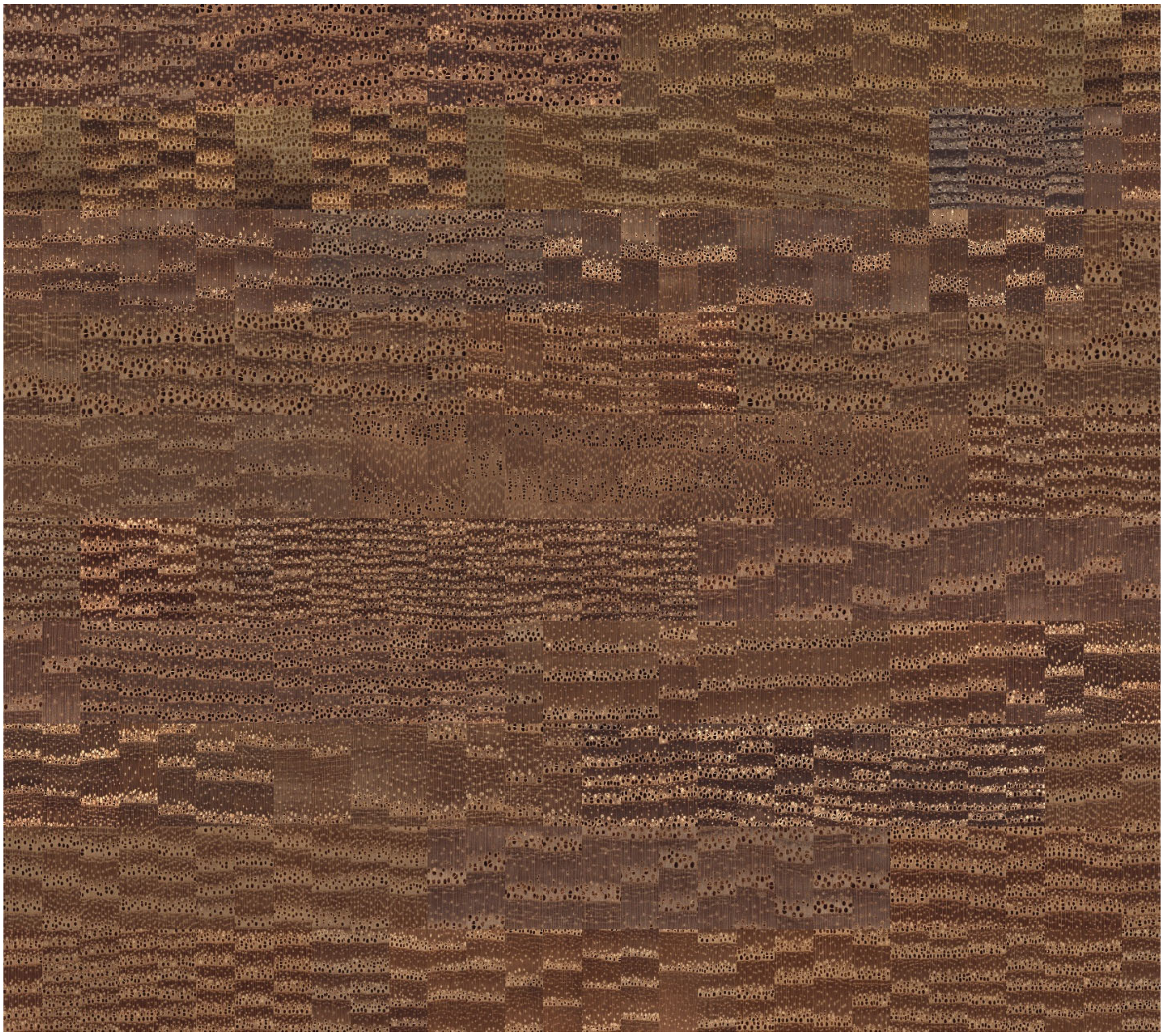

Figure S3. Montage of central patches for class Fraxinus in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/dh3x6ycmvrnuiua3emdj481yi14ehced

Figure S4. Montage of central patches for class Gleditsia in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/s07r6bdvfmw2aag1soyk8ef7sl63xn36

Figure S5. Montage of central patches for class Maclura in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/z8gvqnhs1dlegt3g9woiv536gz02li9l

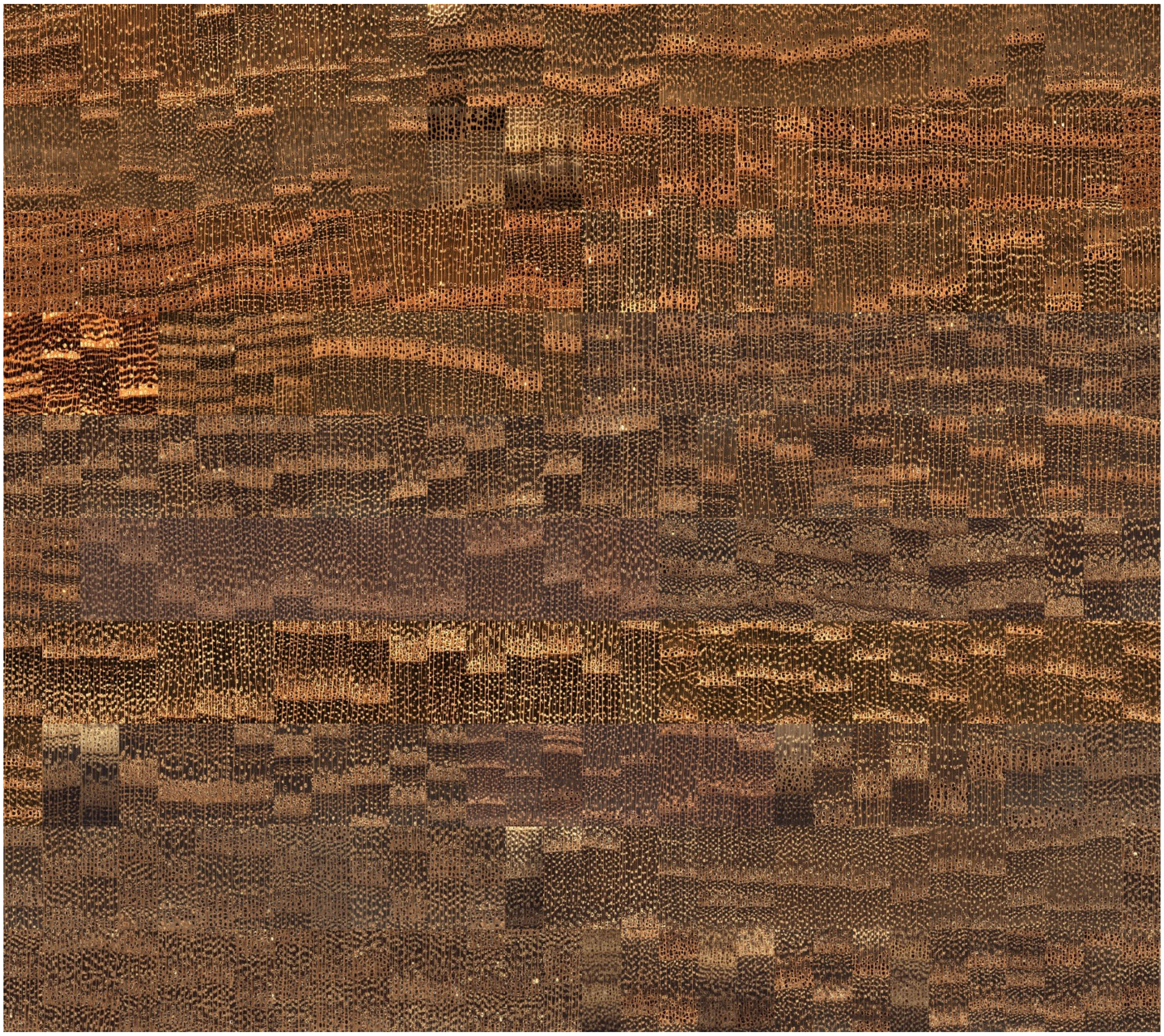

Figure S6. Montage of central patches for class Morus in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/mbnzpair5pwano43fr20gokh73102o8r

Figure S7. Montage of central patches for class QuercusR in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/41t1tki2568kb5ids4zyqjyj90l7zj0v

Figure S8. Montage of central patches for class QuercusW in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/yepg67whpd7h0saoljo4cr7k4wf17quz

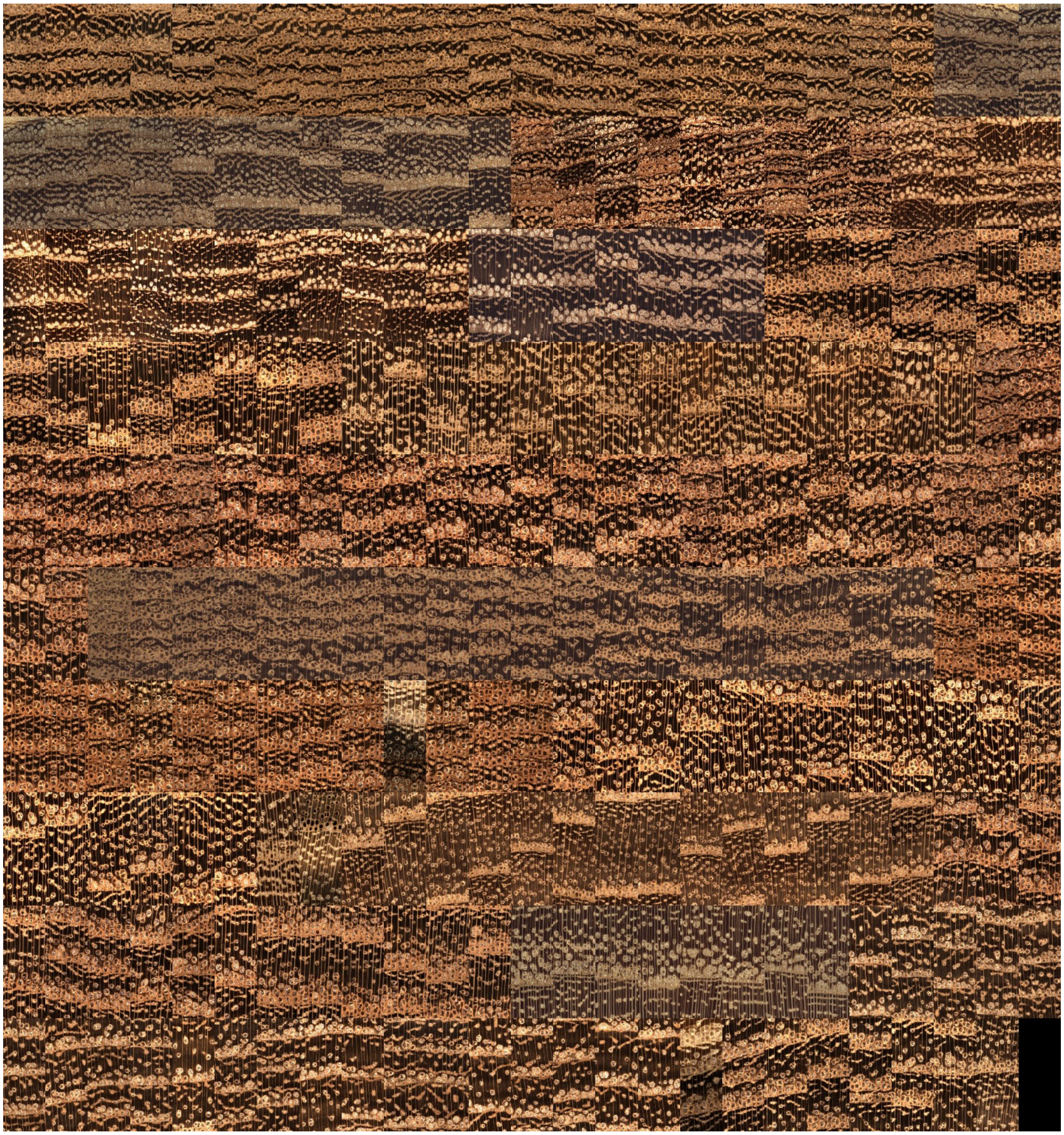

Figure S9. Montage of central patches for class Robinia in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/mvzikjj3j8dsy4g7zqg6gvg3b3im4oke

Figure S10. Montage of central patches for class Sassafras in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/hp4vohtfoqk94w350oxt8qb7a9oovzvi

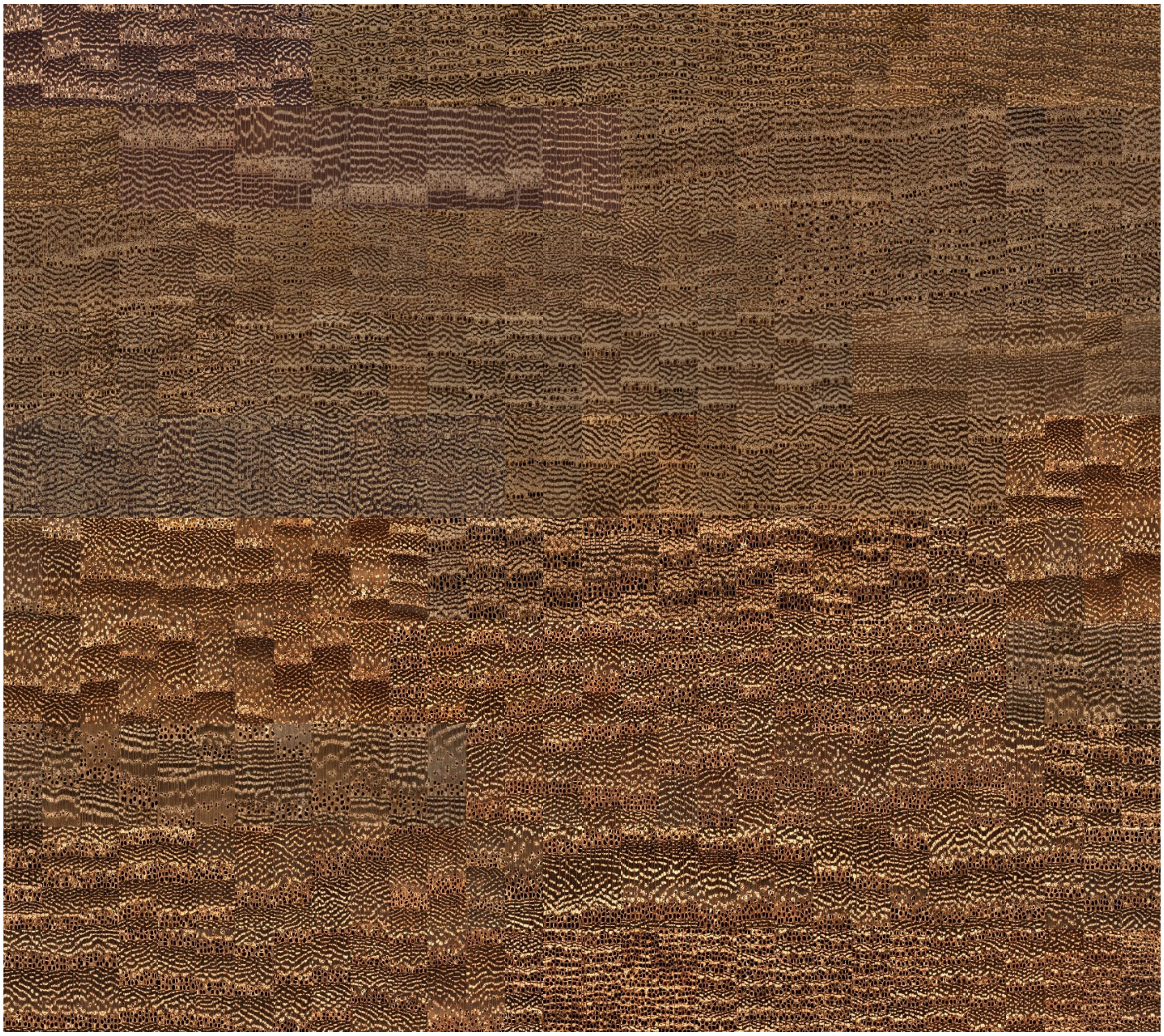

Figure S11. Montage of central patches for class Ulmus in XT-train. The full resolution montage is at: https://uwmadison.box.com/s/vhn4bwwcxu67qj1cjm57baqs5x9vnc7q

Montages of central patches for the LD-train

Figure S12. Montage of central patches for class Celtis in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/00eduxp00j5pox88aybx32clwfgqfeyx

Figure S13. Montage of central patches for class Fraxinus in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/7k4kgl9765fej3wis22benkwpwbe089j

Figure S14. Montage of central patches for class Gleditsia in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/vqu2icfytwmh1q29ocnaihblhi1dw8a2

Figure S15. Montage of central patches for class Maclura in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/xwcut9vyn0bzrz0sd23hlcox5bzprev2

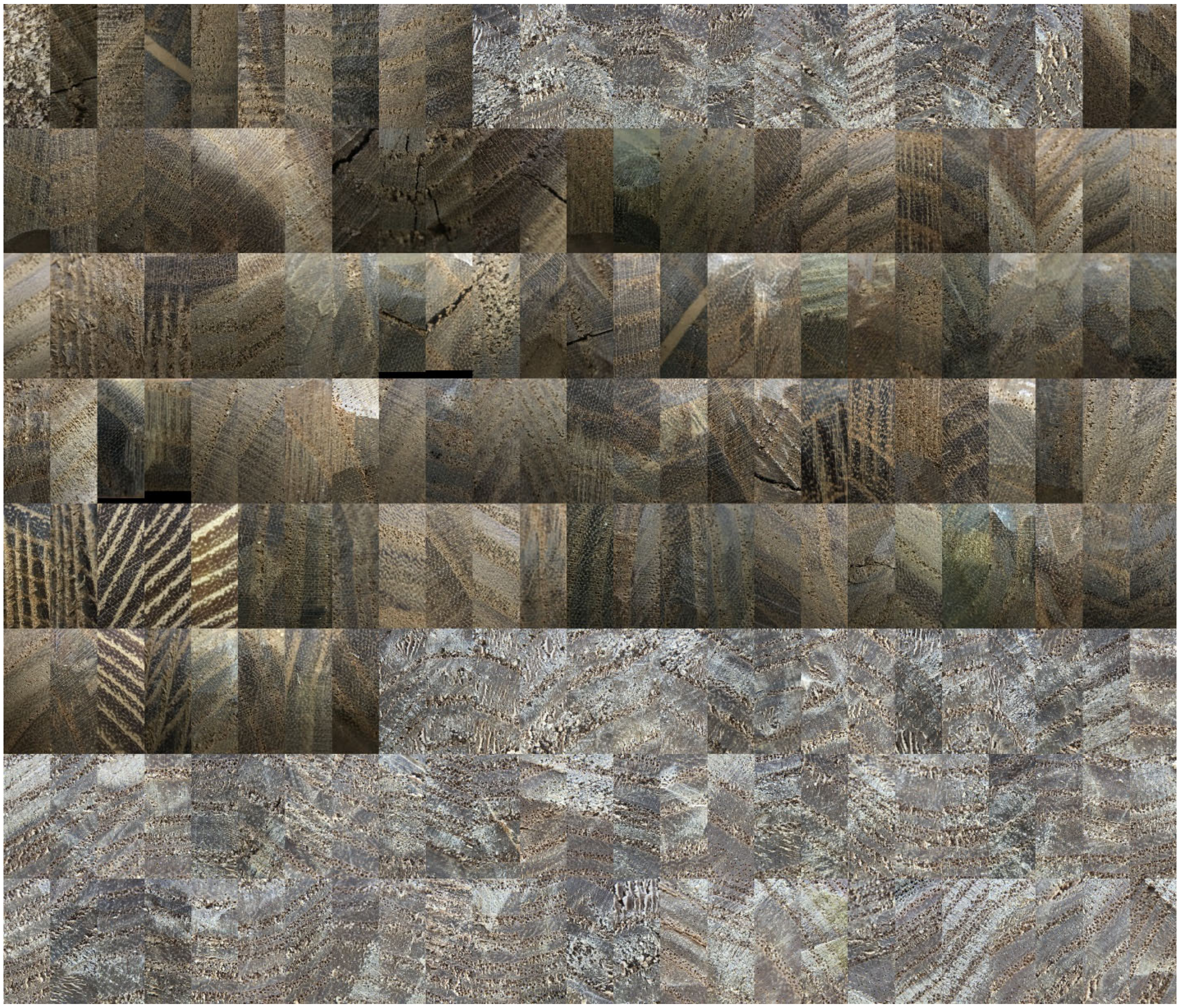

Figure S16. Montage of central patches for class Morus in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/dcjbljoxawmk9q6w57qkphmg5bq6woc2

Figure S17. Montage of central patches for class QuercusR in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/jjzmp9k4luoofmpiz7c82m3bd33zx8ao

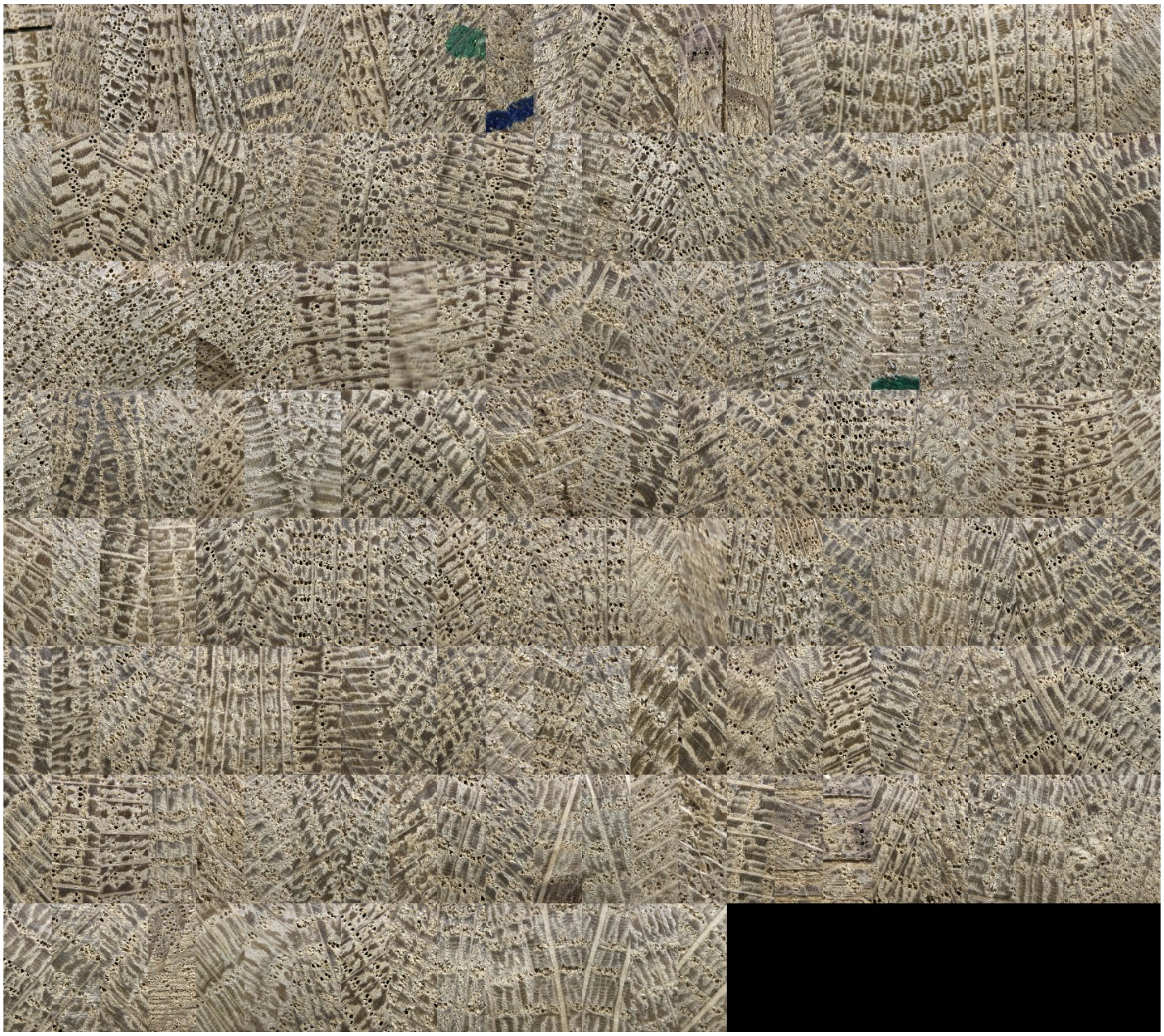

Figure S18. Montage of central patches for class QuercusW in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/bp0tt77qkww9xj7ltvznza25g7a4dhjz

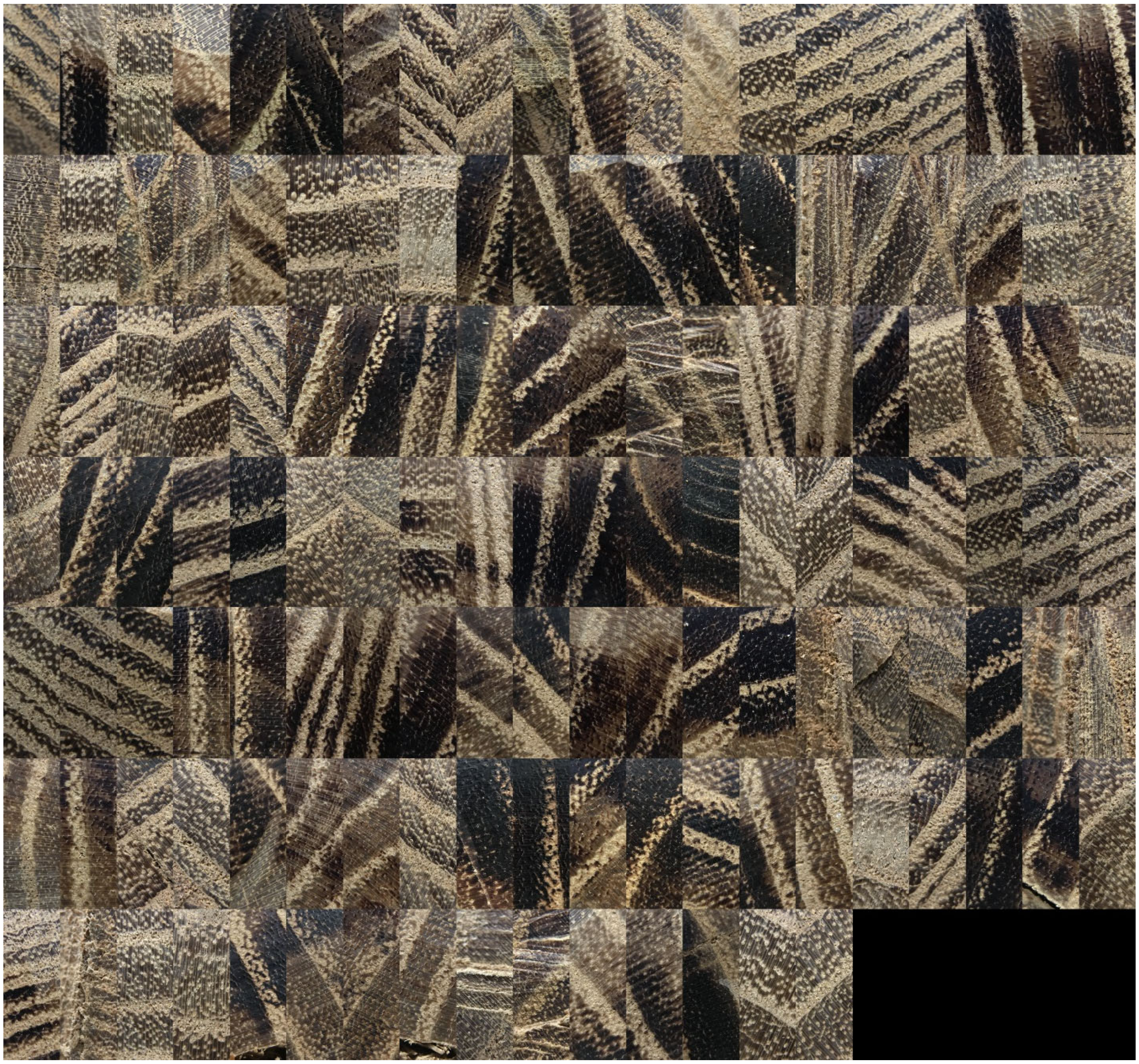

Figure S19. Montage of central patches for class Robinia in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/tosygsjjt1xxpj4z2y0nop1qf0niw2qs

Figure S20. Montage of central patches for class Sassafras in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/d68sy75llur2nwd3hnfn1qhneuzq0rlg

Figure S21. Montage of central patches for class Ulmus in LD-train. The full resolution montage is at: https://uwmadison.box.com/s/p1emnffmbhwtn7d25odpdrcfroijcxr5

Attempts to reconcile inconsistencies in the LD-train dataset:

1. The dataset subtending the Lopes et al. (2020) paper was downloaded using the URL: https://ir.library.msstate.edu/handle/11668/18480. The dataset contained 1709 images of (purportedly) 10 North American hardwoods from an unspecified number of distinct specimens.

2. Initial exploratory analysis of the data set revealed that there was no Exchangeable Image File Format (EXIF) data associated with the images, despite a cell-phone being used to capture the images. Additionally, in the absence of EXIF data, we asked if there were any image meta-data that was available for contextualizing the data and if any specimen identifiers were available – this information was not provided.

3. We sought clarification on the following, to perform data splits for 5-fold cross-validation analysis:

a) whether the cross-validation fold membership for each image was available,

b) should the image filenames with atypical file name patterns be handled in a specific way, and,

c) whether the random state used to seed the stratified cross validation splits was available?

4. In the absence of necessary details (3a-c) and the related meta-data (e.g., EXIF data, notes, random seed) for the images needed to perform specimen-level cross-validation splits, we did a random image-level cross-validation split of the Lopes et al. (2020) dataset for our analyses.

5. The lack of pixel resolution (microns per pixel) for the Lopes et al. (2020) dataset, precluded extraction of image patches that were commensurate with those of the XT patches (i.e., the same amount of tissue captured in each patch) and our analysis is based on patches of the same patch width and height in pixels.

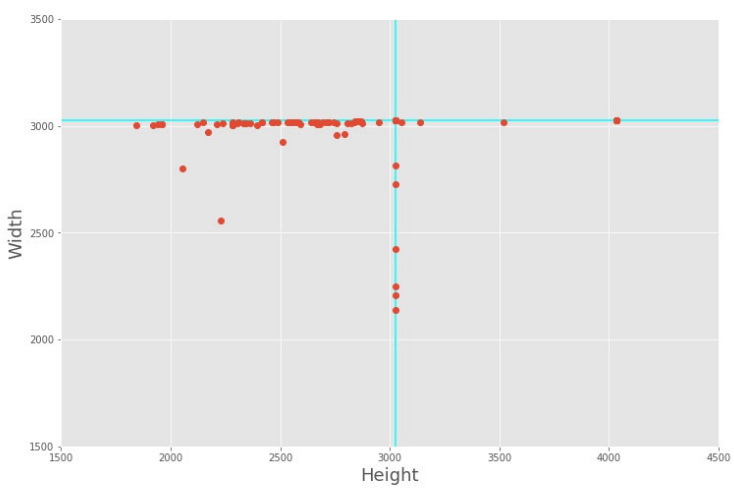

Figure S22. Scatter plot of image dimensions in the Lopes et al. (2020) dataset. The intersection of the cyan lines is the image size reported in the paper (3024 X 3024 pixels). The red dots are the image dimensions of the individual images.

Selecting up to 5 images per specimen level prediction for XT-trains

The XT-train dataset consisted 2635 images obtained from 219 specimens.

For the XT-train dataset, specimen level prediction was obtained as the majority of the predictions on up to 5 randomly chosen images contributed by the specimen. To study the variation in the prediction accuracies due to the random selection of 5 images per specimen, 10 repeats of the process of sampling (up to 5) images from a specimen to compute the specimen level cross-validation accuracy was performed. It should be noted that the sampling of up to 5 images per specimen was done on a fold only when it served as the validation set during cross-validation. The number of images per specimen to derive a specimen-level prediction was *ad hoc* and fixed *a priori* i.e., this value was not "tuned". The confusion matrices and the associated accuracies for the 10 repeats are presented here. The confusion matrix for a model with the highest accuracy and the averaged accuracy (and its standard deviation) over the 10 repeats was reported in the paper.

The accuracies obtained for the 10 repeats are (in %): 96.8, 97.7, 96.8, 97.3, 97.3, 96.8, 95.9, 97.3, 97.7, 97.7 which correspond to the confusion matrices in Figure F.23 in raster order.

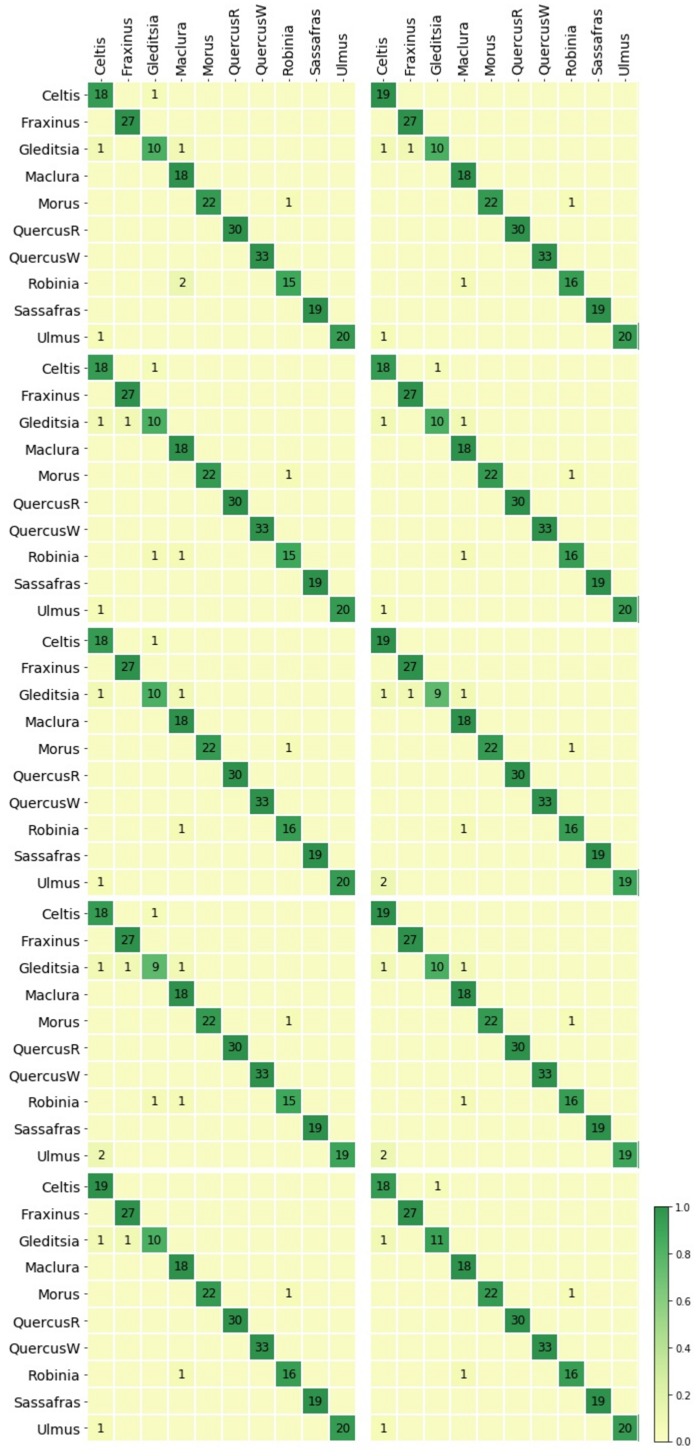

Figure S23. The confusion matrices from the 10 repeats of the sampling process. The confusion matrix corresponding to the best accuracy (top-right) was reported in the main paper.