

## Principles of visual key construction—with a visual identification key to the Fagaceae of the southeastern United States

By: Bruce K. Kirchoff, Roxanne Leggett, Va Her, Chue Moua, Jessica Morrison and Chamika Poole

Kirchoff, B.K., Leggett, R., Her, V., Moua, C., Morrison, J., & Poole, C. (2011). Principles of visual key construction—with a visual identification key to the Fagaceae of the southeastern United States. *AoB Plants*, 2011, plr005. DOI:10.1093/aobpla/plr005.

**\*\*\*Note: This version of the document is not the copy of record. Made available courtesy of the Oxford University Press. Link to Full Text:**

**<http://aobpla.oxfordjournals.org/content/2011/plr005.full>**

**\*\*\*Note: Photographs are not included in this version of the document.**

### **Abstract:**

*Background and aims.* Advances in digital imaging have made possible the creation of completely visual keys. By a visual key we mean a key based primarily on images, and that contains a minimal amount of text. Characters in visual keys are visually, not verbally defined. In this paper we create the first primarily visual key to a group of taxa, in this case the Fagaceae of the southeastern USA. We also modify our recently published set of best practices for image use in illustrated keys to make them applicable to visual keys.

*Methodology.* Photographs of the Fagaceae were obtained from internet and herbarium databases or were taken specifically for this project. The images were printed and then sorted into hierarchical groups. These hierarchical groups of images were used to create the ‘couplets’ in the key. A reciprocal process of key creation and testing was used to produce the final keys.

*Principal results.* Four keys were created, one for each of the parts—leaves, buds, fruits and bark. Species description pages consisting of multiple images were also created for each of the species in the key. Creation and testing of the key resulted in a modified list of best practices for image use visual keys.

*Conclusions.* The inclusion of images into paper and electronic keys has greatly increased their ease of use. However, virtually all of these keys are still based upon verbally defined, atomistic characters. The creation of primarily visual keys allows us to overcome the well-known limitations of linguistic-based characters and create keys that are much easier to use, especially for botanical novices.

### **Article:**

#### **INTRODUCTION**

The creation of easy-to-use identification tools has long been a goal of the systematics and biodiversity communities (Tilling 1984; Stevenson *et al.* 2003; Lawrence and Hawthorne 2006; Walter and Winterton 2007). Although most older keys are text based with relatively few illustrations, recent advances in digital technology have made the creation of visually enhanced identification guides a reality. In a companion paper we review the current state of the art of image use in printed and electronic guides, and formulate a series of best practices for image use in these tools (Leggett and Kirchoff 2011). In this paper we create the first primarily visual key

to a set of taxa, and modify our set of best practices to adapt them to primarily visual keys. By a primarily visual key we mean a key based almost exclusively on images, and which contains a minimal amount of text. We selected the Fagaceae of the southeastern USA as the target group for this key because members of the genus *Quercus* (Fagaceae) are difficult to identify, and almost impossible to key out successfully. By creating a visual key to the Fagaceae, we demonstrate the viability of our approach on one of the most difficult species groups in the southeastern USA.

## MATERIALS AND METHODS

The majority of the photographs were obtained from Steve Baskauf's Bioimages website (Baskauf 2003–2010), from the online image database maintained by the University of Tennessee Vascular Plant Herbarium (Wofford 2010) or were taken by one of the authors. A few additional images came from the other photographers and websites listed in the Acknowledgements. Our original photographs were taken either with a Wild M5 dissecting scope equipped with a Leica DFC420 digital camera using Media Cybernetics Image-Pro Express software or with a Nikon D90 digital camera equipped with an AF-S Micro Nikkor 105 mm macro lens. As much as possible, only standardized photographs were used for this project (Baskauf and Kirchoff 2008). Sets of images, including multiple photographs of leaves, buds, bark, inflorescences and fruit, were collected for each species. Only species for which multiple images of each type were available were included in the key. The need for multiple images restricted the number of taxa that could be included (Table 1).

The key covers 13 of the 20 common *Quercus* ssp. that occur in the southeastern USA (Weakley 2010), and 6 of the 19 rare or uncommon species (Table 1). Of the three species that have varieties in the southeastern USA, *Quercus sinuata* var. *sinuata* is not included in the key. The other two (*Quercus marilandica* var. *marilandica* and *Quercus shumardii* var. *shumardii*) are listed without their varietal names (Table 1), as only one variety occurs in the region. *Quercus rubra* has two subspecies (*rubra* and *ambigua*), but these subspecies are not distinguished in the key due to the lack of required photographs. The application of the name *Quercus prinus* has been controversial. Although it has nomenclatural priority, its providence remains in doubt (Weakley 2010). We follow Weakley (2010) in using *Q. montana* for clarity. One of the two common *Castanea* species is included in the key. The remaining two species are either uncommon (*Castanea ozarkensis*) or cultivated (*Castanea mollissima*) (Table 1). We do not recognize subspecies in *Fagus grandifolia*, the only species of *Fagus* in the southeastern USA (Weakley 2010).

There were three stages in building the key: acquiring images (described above), printing and arranging the images (creating the key), and testing. The latter two stages were done in alternation, so that information from the tests was fed back into the arrangement of the images. To print the images, they were first arranged into contact sheets using Adobe Photoshop CS3. The contact sheets were printed on an HP 4600 Color LaserJet printer and cut into individual 6.2 × 4.2 cm photographs. The photographs were then sorted into groups by image type (leaves, buds, bark, inflorescences and fruit). Within a type, the photographs were sorted into hierarchical similarity groups (Kirchoff *et al.* 2007). These similarity groups serve as holistic characters, and replace more traditional verbally defined, atomistic characters (Kirchoff *et al.* 2004, 2007). The first sort produced two to three highest-level (most inclusive) groups. These initial groups were

**Table 1:** Taxonomic coverage of the key. Abundance and resemblance group data are from Weakley (2010).

Genus	Species	In key	Abundance	Resemblance group
<i>Castanea</i>	<i>dentate</i>	X	Common	Leaves < 15 cm
<i>Castanea</i>	<i>ozarkensis</i>		Restricted	Leaves < 15 cm
<i>Castanea</i>	<i>mollissima</i>		Cultivated	Leaves 8-30 cm
<i>Castanea</i>	<i>pumila</i>		Common	Leaves 8-30 c
<i>Fagus</i>	<i>grandifolia</i>	X	Common	
<i>Quercus</i>	<i>acutissima</i>		Cultivated	Chestnut oaks
<i>Quercus</i>	<i>michauxii</i>		Common	Chestnut oaks
<i>Quercus</i>	<i>montana</i>	X	Common	Chestnut oaks
<i>Quercus</i>	<i>muehlenbergii</i>	X	Rare	Chestnut oaks
<i>Quercus</i>	<i>prinoides</i>		Rare	Chestnut oaks
<i>Quercus</i>	<i>bicolor</i>	X	Uncommon	Chestnut/white oaks
<i>Quercus</i>	<i>chapmanii</i>		Uncommon	Laurel oaks
<i>Quercus</i>	<i>elliottii</i>		Uncommon/rare	Laurel oaks
<i>Quercus</i>	<i>geminata</i>		Common	Laurel oaks
<i>Quercus</i>	<i>hemisphaerica</i>		Common	Laurel oaks
<i>Quercus</i>	<i>imbricaria</i>	X	Uncommon	Laurel oaks
<i>Quercus</i>	<i>incana</i>		Common	Laurel oaks
<i>Quercus</i>	<i>laurifolia</i>	X	Common	Laurel oaks
<i>Quercus</i>	<i>minima</i>		Uncommon	Laurel oaks
<i>Quercus</i>	<i>myrtifolia</i>		Uncommon	Laurel oaks
<i>Quercus</i>	<i>oglethorpensis</i>		Rare	Laurel oaks
<i>Quercus</i>	<i>phellos</i>	X	Common	Laurel oaks
<i>Quercus</i>	<i>virginiana</i>		Uncommon	Laurel oaks
<i>Quercus</i>	<i>marilandica</i>	X	Common	Laurel/red oaks
<i>Quercus</i>	<i>nigra</i>	X	Common	Laurel/red oaks
<i>Quercus</i>	<i>arkansana</i>		Rare	Laurel/white oaks
<i>Quercus</i>	<i>sinuata</i> var. <i>sinuato</i>		Rare	Laurel/white oaks
<i>Quercus</i>	<i>coccinea</i>	X	Common	Red oaks
<i>Quercus</i>	<i>falcata</i>	X	Common	Red oaks
<i>Quercus</i>	<i>georgiana</i>		Rare	Red oaks
<i>Quercus</i>	<i>ilicifolia</i>		Common	Red oaks
<i>Quercus</i>	<i>laevis</i>		Common	Red oaks
<i>Quercus</i>	<i>pagoda</i>	X	Common	Red oaks
<i>Quercus</i>	<i>palustris</i>	X	Uncommon	Red oaks
<i>Quercus</i>	<i>rubra</i>	X	Common	Red oaks
<i>Quercus</i>	<i>shumardii</i>	X	Uncommon	Red oaks
<i>Quercus</i>	<i>velutina</i>	X	Common	Red oaks
<i>Quercus</i>	<i>alba</i>	X	Common	White oaks
<i>Quercus</i>	<i>austrina</i>		Rare	White oaks
<i>Quercus</i>	<i>boyntonii</i>		Restricted	White oaks
<i>Quercus</i>	<i>lyrata</i>	X	Common	White oaks
<i>Quercus</i>	<i>macrocarpa</i>	X	Rare	White oaks
<i>Quercus</i>	<i>margarettae</i>		Common	White oaks
<i>Quercus</i>	<i>robur</i>		Cultivated	White oaks
<i>Quercus</i>	<i>similis</i>		Rare	White oaks
<i>Quercus</i>	<i>stellata</i>	X	Common	White oaks

each sorted into two to four smaller subgroups, and this process was continued until only two species remained in each group, or we were unable to further distinguish between the remaining taxa. The resulting hierarchical groups were used to create the key. Each level in the hierarchy was used to create one level of the key. In the remaining portion of the paper, we refer to each level of the key as a couplet, even though there may be more than two choices at some levels.

The keys were created by choosing three or four pictures (to account for variation) from the  $N$  set of groups at a given hierarchical level, and affixing them to a piece of paper in a way that preserved their group relationships (see page 1L of the Leaf Key). These  $N$  sets of images served as a couplet of the key. This process was repeated at each level of the hierarchy until the full key

was created. Four separate keys were created, one for each of the standardized types of photographs: leaves, buds, bark and fruit. There were insufficient photographs of inflorescences to create an inflorescence key. The choice of which photographs to use at a given hierarchical level proved to be crucial for the effective operation of the key. We adjusted our choices through a process of trial and revision, always making sure that we included photographs from multiple species in non-terminal couplets.

To test a key, the two or three authors who did not participate in its creation used it to identify live specimens collected for this purpose, or to identify images from online databases that did not appear in the key. The tester worked through the key at his or her own pace, and was then asked to explain which choice he or she had made at each division in the key, and why. This process allowed us to determine where error(s) had occurred, and to correct them by reworking the couplets.

Once the basic keys were created and tested, we added standardized photographs of complementary plant parts to the final couplets to ease the identification process. For instance, *F. grandifolia*, *Quercus macrocarpa* and *Castanea dentata* come out together in the Fruit Key (Fruit Key, page 2F). Inclusion of a leaf photograph to each terminal helps distinguish these species. We also created species description pages consisting of all of the photographs for each taxon. We did not include text-based species descriptions, both because these are available elsewhere (Stein *et al.* 2003), and because our work is intended to demonstrate the principles of visual key construction, not to produce a fully field-ready key. Following testing and revision, the final keys were laid out using Adobe InDesign CS3 and CS5.

## RESULTS

The four visual keys are reproduced in the Appendix (key pages 1L–11L, 1B–4B, 1T–4T, 1F–12F). They are followed by the description pages for the 21 species we cover. The four keys are laid out in a split-page design so that two of the keys can be used simultaneously. Leaf (L) and bud (B) keys occur on the same full page, as do twig (T) and fruit (F) keys. The different lengths of the keys mean that, other than on their first pages, they do not start and end together. Coloured bands below each non-terminal couplet help direct the user to the next couplet, which is marked with a similarly coloured band on its upper right margin. Curved arrows within each couplet indicate alternative choices. When couplets extend over more than one page, the continuation is marked by yellow-highlighted curved arrows on both pages, and parenthetical numbers added to each page designation (see e.g. pages 3L(1)–3L(2), 7L(1)–7L(2)).

The full pages containing the keys can be printed in the normal way, stapled along their left margin and cut (incompletely) apart along the dotted lines so that the upper and lower halves of each page remain together, but turn independently. This allows the upper and lower keys to be used independently, or together. When cut apart, either the Leaf or Bark Key (upper section) can be used simultaneously with either the Bud or Fruit Key (lower section). The species description pages can also be split, although this is not essential. The upper halves of these pages show bud and fruit images, while the lower contain leaf and bark. This allows the user to check his or her identifications against a fuller set of images of the same type. For instance, a user reaching page 5L in the Leaf Key must choose between *Quercus palustris*, *Q. falcata* and *Q. pagoda*, species

that can be difficult to distinguish. The arrangement of images in the species section allows him or her to see more images of these species without losing his or her place in the key.

Scale bars are only used when they provide information relevant to a specific couplet. That is, when size is a deciding factor in making a decision and shape is not. For instance, the terminal couplet on 8L asks the user to distinguish between *Q. marilandica* and *Quercus nigra*, taxa that are easy to distinguish in the field, but difficult to tell apart from the images in the key. Scale bars added to two of the images show that the leaves are different sizes. Use of the species identification pages will also help the user distinguish similar species.

Abundance information is used in the same way as scale bars. It is included above final couplets where the user can expect to find one of the species much more frequently than the other. For instance, a user arriving at page 10L must choose between *C. dentata* and *F. grandifolia*. The abundance information provided above the images allows the user to make a more informed choice between the species. These species can also be easily distinguished by their bark, photographs of which are available on their species pages.

The degree of structure in the keys varies. The Leaf Key is the most highly structured, with relatively few tri- or quadracotomies. At the other extreme is the Bark Key, where the user must choose among 17 possibilities at the final level (key pages 4T(1)–4T(4)).

## DISCUSSION

A full understanding of the principles embodied in these keys can only be gained through their use. For testing outside the southeastern USA, images of appropriate members of the Fagaceae can be downloaded from the web, printed and used to test the keys. Testing is better accomplished if the user does not know the identity of the unknown before using the key.

When first approaching a key of this type, botanical novices tend to consider aspects of the images that are not relevant to identification. For instance, when using the Leaf Key it is very common for novices to attend to the colour or degree of reflectance (shininess) of leaves, rather than characteristics of leaf shape, which are more diagnostic. The easiest way to overcome these initial predilections is to instruct novice users that colour and reflectance are not good indicators of taxon identity. Directing a user's attention to shape instead of the surface features will improve his or her initial performance with the keys. We have found it best for novice users to begin with the Leaf Key as they are more accustomed to looking at leaves than buds, fruits or bark. Once they have made several correct identifications using this key, they can move on to the other keys. Use of the other keys is more difficult not only because the subject matter is less familiar, but also because we were not able to obtain a sufficient number of photographs of these features to show the full range of variation. This is especially true of the Bark Key, which has very little hierarchical structure compared with the other keys. We were only able to obtain a single bark photograph for most species. The general opinion among botanists is that trees cannot be distinguished by bark alone, so there are few photographs of bark available. However, our experiments with botanical novices show that they can make correct identifications based only on bark, even using our relatively ill-structured key. We hope that collections of standardized bark photographs will be available in the future so that more complete bark keys can be created.

### *Best practices in image use in visual keys*

We have reviewed image use in field guides and keys elsewhere, and have formulated a set of best practices for image use in these publications (Leggett and Kirchoff 2011). Many of these best practices have been implemented in this key. For instance, we have used multiple standardized photographs on non-distracting backgrounds to illustrate each character state, have used coloured bands to help users find their way through the key, and have used scale bars with discretion. In this section, we refine our set of best practices to make them more applicable to visual keys. The following list contains our best practices for image use in primarily visual keys:

1. The use of multiple standardized images is essential in visual keys. Multiple images of the same part should be used in each couplet so that the user can form a concept of variation represented by the couplet (Wisniewski 2002; Leggett and Kirchoff 2011).
2. The degree of variation represented by the images in a couplet must encompass the variation in the taxa that are identified by following the leads in that couplet. This can be verified by testing the key during its construction. When a target species is misidentified during testing, it is usually an indication that the degree of variation in the selected images does not adequately represent variation in the taxa.
3. The variation represented in the alternative choices of a couplet must not overlap. Where it is impossible to remove the overlap (e.g. in cases of polymorphic taxa), it must be possible to identify the overlapping taxa by following more than one path through the key.
4. The use of arrows and other indicators on the images should be restricted to situations where taxa can only be distinguished by the use of technical characters (Leggett and Kirchoff 2011). In these cases, arrows or other indicators provide a useful service by focusing the user's attention on the part of the image that contains the relevant character (see Lung *et al.* (2001) for examples).
5. Scale bars are not essential on most images, and should only be used in cases where alternative choices cannot be distinguished by form alone.

### CONCLUSIONS AND FORWARD LOOK

With modern digital imaging technology it is now possible to create completely visual keys that avoid the well-known problems with technical terminology (Stevens 1991; Fallshore and Schooler 1995; Gift and Stevens 1997; Kirchoff *et al.* 2004, 2007, 2008a; Vogt *et al.* 2010). These types of keys may be either electronic (Kirchoff *et al.* 2008b) or paper based, but in either case provide an alternative to terminology-based keys that are in common use today and which are very difficult to use, especially for novices. The development of illustrated electronic and of multi-access keys has greatly increased ease of key use (Edwards and Morse 1995; Stevenson *et al.* 2003; Farr 2006), as has the addition of illustrations to traditional text-based keys (Leggett and Kirchoff 2011). The possibility of developing completely visual keys extends this potential in ways that we find exciting.

### CONTRIBUTIONS BY THE AUTHORS

B.K.K. conceived the project, arranged for or took the photographs, helped assemble the key, wrote the manuscript and codified the best practices. R.L., V.A., C.M., J.M. and C.P. helped assemble and test the key.

## ACKNOWLEDGEMENTS

In addition to the authors, the following photographers provided images for the keys: Steve Baskauf (<http://bioimages.cas.vanderbilt.edu>), Dan Tenaglia (<http://missouriplants.com/>), John Seiler (<http://www.dendro.cnre.vt.edu>), Will Cook (Duke University, <http://www.duke.edu/~cwcook>), David Lemke (<http://botany.csd.tamu.edu/FLORA/swts/faga005.jpg>), Steve Hurst (<http://plants.usda.gov>), Robert Vidéki (Doronicum Kft, <http://bugwood.org/>), Bill Cook (Michigan State University, <http://bugwood.org/>), Keith Kanoti (Maine Forest Service, <http://bugwood.org/>), Wendy VanDyk Evans (<http://bugwood.org/>), Chris Evans (River to River CWMA, <http://bugwood.org/>), Linda Haugen (USDA Forest Service, <http://bugwood.org/>), Paul Wray (Iowa State University, <http://bugwood.org/>), and John Beck, Edward W. Chester and B. Eugene Wofford (all <http://tenn.bio.utk.edu/vascular/vascular.shtml>). The authors thank them for permission to use their images.

## REFERENCES

- Baskauf SJ. 2003–2010. Bioimages website. Nashville, TN.  
<http://www.cas.vanderbilt.edu/bioimages> (12 August 2010).
- Baskauf SJ, Kirchoff BK. 2008. Digital plant images as specimens: standards for photographing living plants. *Vulpia* 7: 16–30.
- Edwards M, Morse DR. 1995. The potential for computer-aided identification in biodiversity research. *Trends in Ecology & Evolution* 10: 153–158.
- Fallshore M, Schooler JW. 1995. Verbal vulnerability of perceptual expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21: 1608–1623.
- Farr DF. 2006. On-line keys: more than just paper on the web. *Taxon* 55: 589–596.
- Gift N, Stevens PF. 1997. Vagaries in the delimitation of character states in quantitative variation—an experimental study. *Systematic Biology* 46: 112–125.
- Kirchoff BK, Richter SJ, Remington DL, Wisniewski E. 2004. Complex data produce better characters. *Systematic Biology* 53: 1–17.
- Kirchoff BK, Richter SJ, Remington DL. 2007. Characters as groups: a new approach to morphological characters in phylogenetic analysis. *Taxon* 56: 497–492.
- Kirchoff BK, Pfeifer E, Rutishauser R. 2008a. Plant structure ontology: how should we label plant structures with doubtful or mixed identities? *Zootaxa* 1950: 103–122.
- Kirchoff BK, Remington DL, Fu L, Sadri F. 2008b. A new type of image-based key. 2008 *International Conference on BioMedical Engineering and Informatics (BMEI 2008)*. Hainan, China: Sanya.

- Lawrence A, Hawthorne W. 2006. *Plant identification: creating user-friendly guides for biodiversity management*. London, UK: Earthscan.
- Leggett R, Kirchoff BK. 2011. Image use in field guides and identification keys: review and recommendations. *AoB Plants* 2011 plr004 doi:10.1093/aobpla/plr004.
- Lung M, Sommer S, Paulson D. 2001. Visual key to dragonflies & damselflies. In: Link P, Peterson C, Welhan J, Sommer S, McNamara J, Schiappa T eds. *Digital atlas of Idaho*. Pocatello, ID: Idaho Museum of Natural History.  
<http://imnh.isu.edu/digitalatlas/bio/insects/drgnfly/idkey/keyfr.htm> (3 January 2011).
- Stein J, Binion D, Acciavatti R. 2003. *Field guide to native oak species of Eastern North America*. Morgantown, WV: United States Forest Service.
- Stevens PF. 1991. Character states, morphological variation, and phylogenetic analysis: a review. *Systematic Botany* 16: 553–583.
- Stevenson RD, Haber WA, Morris RA. 2003. Electronic field guides and user communities in the eco-informatics revolution. *Conservation Ecology* 7: article 3.  
<http://www.consecol.org/vol7/iss1/art3/> (4 January 2011).
- Tilling S. 1984. Keys to biological identification—their role and construction. *Journal of Biological Education* 18: 293–304.
- Vogt L, Bartolomaeus T, Giribet G. 2010. The linguistic problem of morphology: structure versus homology and the standardization of morphological data. *Cladistics* 26: 301–325.
- Walter DE, Winterton S. 2007. Keys and the crisis in taxonomy: extinction or reinvention? *Annual Review of Entomology* 52: 193–208.
- Weakley AS. 2010. *Flora of the southern and mid-atlantic states*. Chapel Hill, NC: UNC Chapel Hill.
- Wisniewski EJ. 2002. Concepts and categorization. In: Medin D, ed. *Stevens' handbook of experimental psychology. vol 2. Memory and cognitive processes*, 3rd edn. New York, NY: Wiley, 467–531.
- Wofford BE. 2010. *Vascular Plant Herbarium* (online). Knoxville, TN: University of Tennessee.  
<http://tenn.bio.utk.edu/vascular/vascular.shtml> (6 January 2011).