DELLINGER-JOHNSTON, REBECCA A., M.S. A New Method for Creating a Visual Plant Identification Key. (2015)
Directed by Dr. Bruce K. Kirchoff. 88 pp.

Taxonomic keys are essential tools for species identification, used by students and professional biologists. In recent years, advancements in photography have allowed these keys to host high-quality photographs for aid in identification. However, most modern keys still rely heavily on text rather than images. Using text alone limits the user to a discrete number of characters, often described in esoteric terms. In order to create more effective keys, we developed a new method for constructing image-based taxonomic keys. These keys replace written characters with images – allowing the user to identify species using visual pattern recognition, rather than interpreting written text. In addition, we constructed our visual key using data on how different users assess the visual similarities between plant species. To ensure the strength of this methodology, our key focuses on the morphologically diverse genus, *Quercus*.

A set of standardized photographs was taken of forty-three species of oak native or naturalized in the Southeast. These photographs were used to create a survey on how botanical experts and botanical novices rate the pair-wise similarity of different oak leaves. The mean of each rating was summarized into a distance matrix, which was then converted into a dendrogram. From the resulting dendrogram, a visual key was constructed using the standardized photographs of oak leaves. The key was then tested on against an existing dichotomous key using botanical novices and botanical experts.

The resulting two-sample *t*-tests between the two identification keys demonstrated that users with our visual key produced between 22-30% more correct answers than users with the traditional key. Using this method of key creation, innovative keys could be constructed for other fields of biology.

A NEW METHOD FOR CREATING A VISUAL PLANT IDENTIFICATION KEY

by

Rebecca A. Dellinger-Johnston

A Thesis Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Greensboro
2015

Approved by

_____
Committee Chair

To Terre Milton Smith

APPROVAL PAGE

This thesis written by Rebecca Ann Dellinger-Johnston has been approved by the

following committee of the Faculty of The Graduate School at The University of North Carolina

at Greensboro.

Committee Chair_____

Committee Members_____

_____

_____

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Identification Keys: Use and Importance

One of the most crucial elements of any biological study is the recognition and correct identification of species. In order to achieve this, professionals use a multitude of tools, such as floras, monographs, species keys, and field guides. These resources not only provide them with a correct identification, but a breadth of taxonomic information that is necessary to locate and describe species they encounter in the field (Pearson *et al*. 2011). Areas such as ecology and forestry routinely use this information to identify priority species, remove invasive species, and complete biodiversity surveys (Terlizzi *et al.* 2003, Wheeler *et al.* 2004, Joppa 2011). Likewise, microbiologists use identification instruments to recognize and assess pathogens and microorganisms. As a result, species identification is also a crucial factor in making decisions regarding public and environmental health (Walter and Winterton 2007, Pearson *et al.* 2011).

As the concern for global health grows, much of current research focus has been directed toward the assessment of biological system health. This includes the identification, assessment, and preservation of keystone species within those systems (Bacher 2012). Currently, it is estimated that 2 million, out of a proposed 7 - 15 million, species have been described by taxonomists (Pearson et al. 2011). In the meantime, species extinction rates continue to rise and conservationists are faced with the dilemma of being unable to identify priority species (Bacher 2012).

For the past two decades, concerns of a "biodiversity crisis" have left many scientists wondering if there are enough specialists to adequately recognize and describe species before

they are driven to extinction (Terlizzi et al. 2003, Bacher 2012, Pearson et al. 2011, Tancoigne and Dubois 2013). For the past two decades, concerns of a "biodiversity crisis" have left many scientists wondering if there are enough specialists to adequately recognize and describe species before they are driven to extinction (Terlizzi *et al.* 2003, Bacher 2012, Pearson et al. 2011, Tancoigne and Dubois 2013). Exacerbating the problem is a reported decline in the number of professional taxonomists equipped to study species (Boero 2001, Blackmore 2002, Godfray 2002, Wheeler et al 2004, 2013, Walter and Winterton 2007, Pearson *et al.* 2011, Sluys 2013, Tancoigne and Dubois 2013, Granjou *et al*. 2014). Publications on the phenomena also point to a decline in the number of classes teaching taxonomy, as well as a lack of sufficient funding for taxonomy research (Boero 2001, Godfray 2002, Tancoigne and Dubois 2013, Sluys 2013, Wheeler 2014).

As the need for species data increases, there is also a need to better equip current and future researchers to meet the challenge of maintaining and studying biodiversity. In recent years, there have been many efforts to encourage the general public to become involved in ecological studies (Stagg and Donkin 2013). Projects such as the Great Nature Project, the Great Backyard Bird Count, and Seeds have been utilizing data from citizen scientists. Taking a greater initiative, agricultural research has begun training citizens to become parataxonomists. These specialists are native to particular geographic regions, trained to identify local species in order to identify disease and control for invasive species (Pearson *et al.* 2011, Pertot *et al*. 2012).

While the encouragement of laypersons is admirable, there is a great concern to record accurate information. In order to do this effectively, researchers and their trainees need better training to identify species more quickly and with greater accuracy. This means that identification tools need to be easy to use and highly available to reach a broad audience.

Identification Keys: Current Methodology

There are many different resources for species identification. Floras and monographs, for example, provide detailed descriptions of species life history; whereas pocket guides and pamphlets have a more practical use in the field. Historically, the most widely used tool of these is the dichotomous key (Edwards and Morse 1995, Farnsworth *et al.* 2013, Stagg and Donkin 2013). Dichotomous keys provide users with a pair of written descriptions for a particular group of taxa. From these pairs, the user must choose which description best fits the specimen they are trying to identify. This selection leads the user to another subsequent pair of descriptions. The process is continued until a single taxon is identified (Figure 1).

Group 3: Leaf scars alternate, pith homogenous

1. Prickles or thorns present ....2.
1.' Prickles and thorns lacking..6.

2. Stems arching or trailing, longitudinally grooved; bundle scars not evident..*Rubus* spp.
2.' Stems erect, not grooved; bundle scars evident.........................3.

3. Bundle scars > 8; leaf scar deeply U-shaped and large ....
....*Aralia spinosa*, Fig. 9
3.' Bundle scars 3; leaf scar crescent-shaped, narrow ....4.

4. Prickles lacking; at least some branch stems modified as thorns; stems often forming dense thickets....................*Prunus angustifolia*
4.' Prickles present; branch stems not modified as thorns; stems not forming dense thickets ....5.

28

Figure 1**.** Sample Dichotomous Key. Text from the Winter Twig Key to Common, Native, Fully Deciduous Trees and Phaneophyte Shurbs of the North Carolina Eastern Piedmont (Stucky 2003).

Generally, these descriptive couplets provide comparable traits, such as the presence of a particular structure, texture, or color (Dallwitz 1992, Drinkwater 2009). This way a user identifies a species by analyzing and selecting indicative characters.

While dichotomous keys are some of the most commonly used keys, users are often frustrated by the limitations of the key. A reader using a dichotomous key is often faced with the

3

dilemma of only having one set of characters to start with (Drinkwater 2009, Pertot *et al.* 2012). This is especially problematic if the user possesses an atypical or incomplete specimen. Many botanical keys, for example, rely on the presence of multiple plant structures to correctly identify a species. Flowers are especially common diagnostic characters, but this limits a positive identification to the plant's reproductive season. If the key does not list alternative characters, the user is faced with making an uninformed decision and is more likely to misidentify a species. Often, this obstacle can be avoided through the use of multi-access keys, which provide a list of various, nonhierarchical characters (Figure 2). This unordered list allows the user to choose which characters are available and familiar to them (Tilling 1984, Drinkwater 2009, Pertot *et al.* 2012).
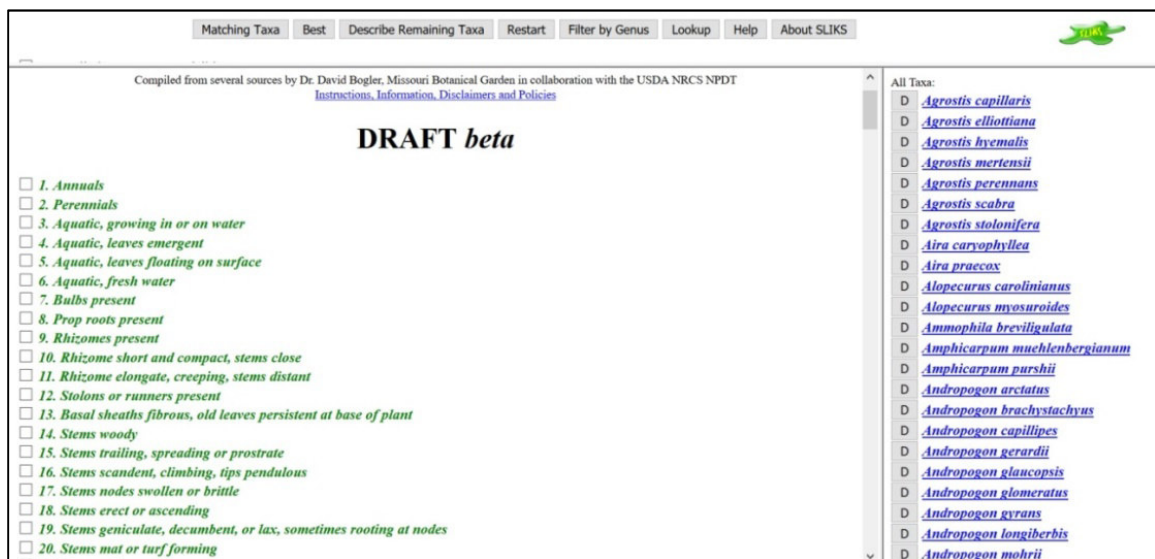
Figure 2. Sample Multi-access Key. Image taken from a screen-capture of the USDA SLIKS Key to North Carolina Native Grasses. Clicking on the characters present in the unknown (left side of the screen, highlighted in green) eliminates taxa (on the right, highlighted in blue) that do not exhibit those characteristics.

However, like dichotomous keys, multi-access keys are still limited by the user's familiarity with a species' diagnostic characters. If a user is not familiar with enough characters a daunting list of

possible species remains (Drinkwater 2009). In order to better understand how such keys can be improved, we must first examine some of the factors that hinder their ability to instruct users.

Limitations of Identification Keys

Part of the difficulty of species identification arises from the language keys use to describe species. Most identification keys use technical terminology, employed by experts who are familiar with them. For example, flowers are common structures used in botanical identification. Botanical keys often reference different parts of flower anatomy for specific diagnostic characters: petals, sepals, stamens, corolla, stigma, etc. While some of these features may be obvious for a novice, others are unfamiliar and require finding a description in a glossary. In the glossary of the Manual of Cultivated Trees and Shrubs (Rehder 2001) a flower structure called the hypanthium is described as "the cup-shaped or tubular receptacle on which the perianth and stamens are inserted." While a definition is provided, this description may be difficult for a novice to interpret because they do not yet have a concept for what is being described. Such a description may also contain additional terms, requiring the user to be familiar with multiple concepts. In this example, the user must also be familiar with the perianth and stamens in order to reach a conclusion on what the hypanthium might be.

The difficulty of using descriptive terminology is not limited to novices. Even experts, who are familiar with anatomical concepts, struggle with identifying new taxa (Hawthorne *et al*. 2014, Stevenson *et al.* 2003). This is due to the inherent discrepancies in the way different authors use or define the same terms. For example, there are distinct terms used to describe the variations of round structures: "ovate," "obovate," "orbicular," or "rotund" (Harris 1954). While there are crucial distinctions between these forms, the language used to describe them may vary between authors. In the Manual for Cultivated Trees and Shrubs, the term "ovate" is described as

5

"having an outline like a hen's egg" (Rehder 2001). The same term is defined as "descriptive of a flat organ flattest below the middle and broader than lanceolate" in the Manual of Vascular Plants of Northeastern United States and Adjacent Canada (Gleason and Cronquist 1963). This same guide defines the term "ovoid" as "egg-shaped". While some guides distinguish these terms with respect to two dimensional and three dimensional structures, this distinct is not always made apparent. As such, these inconsistencies demonstrate that authors often use terms interchangeably. Certainly this fluidity of language also poses a challenge to communicating characters in different languages, making species identification a challenge for international study.

Authors may also use vague adjectives, such as "slightly," or "minutely," that are open to interpretations. For example, Gleason and Cronquist (1963) describe "puberulent" as "minutely or sparsely pubescent". Using such a vague measure of pubescence leads the user to question at what point a term does or does not apply to their unknown specimen.

To ameliorate the problems of misinterpreting language, keys are often supplemented by images. Images enhance keys because they provide the user with visual information that is less confusing and easier to interpret (Baskauf and Kirchoff 2008, Kirchoff *et al*. 2008). While early illustrated guides had a limited number of images, modern advances in high quality printing and computer technology have allowed more images to be used in both printed and online keys (Hawthorne *et al*. 2014, Leggett and Kirchoff 2011). Images may be used in guides to exemplify a particular species or to illustrate term used within the key. For example, The Field Guide to the Orchids of Costa Rica and Panama features illustrations of botanical terms in addition to photographs of the species themselves (Dressler 1993). Adopting image use within the body of the key itself, allows the user to see actual examples of the structures defined within the text.

Images in keys can be photographs of herbarium specimens, photographs of live specimens, or illustrations (Figure 3). In particular, photographs of living specimens give the most accurate representation of a species. Color, season, and texture can be better preserved in photographs than pictures of preserved specimens (Baskauf and Kirchoff 2008). Illustrations, however, can often render images that are more detailed than photographs. Artists drawing a particular character can control the features that are most prominently displayed (Ang *et al*. 2013). Illustrations can also ensure that a typical specimen is presented to the user (Ang *et al*. 2013).

Even with advances in the availability of images, there is a great need for improvement to the application of image in keys. Most keys, with some exceptions, use images to illustrate the species itself, without providing illustrations of the terms used to reach that identification. Popular commercial guides, such as the Sibley Guide to Birds (Sibley 2000), commonly use illustrations in this manner, and do not provide an actual diagnostic key for identification. Highly illustrated guides such as this are also limiting in the number of taxa given – restricting species identification to only the most common or exotic species of interest.

In true diagnostic keys, the photographs used are often not large enough to show the detail that is necessary for a positive identification (Leggett and Kirchoff 2011). Photographs may also omit structures that are relevant to identification (Leggett and Kirchoff 2011). Likewise, illustrations are limiting because they only represent an ideal archetype. Using such a typical or a prototype specimen eliminates the variability within a species category. The same is true for the examples given in most photographs. Studies in psychology have suggested that using a prototype of an object limits the user's ability to gain an understanding of the concept of that object (Wiseniewski 2002). Most guides simply do not show enough images to demonstrate a character effectively. By limiting a species to a single image, the key cannot account for the inherent variation that exists in that character (Leggett and Kirchoff 2011).

GROWTH FORM: small, slow growing tree with a compact crown reaching a normal high of 26 feet (7.9 m), occasionally reaches 75 feet (22.9 m). **BARK:** gray to light brown, mature bark becomes scaly. **TWIGS and BUDS:** smooth red twigs with prominent light brown lenticels; buds are reddish-brown, ovoid with smooth scales that may be ciliated. **LEAVES:** petiole ¼ - ⅞ inch (6 - 22 mm) long, usually with a few hairs; leaf blade broadly elliptical and thin, 1 ½ - 5 ⅛ inches (38 - 130 mm) long, ¾ - 3 ½ inches (19 - 89 mm) wide, with a cuneate or obtuse base, margin with 3 - 5 pointed bristle-tipped lobes; surface is a shiny green above, pale green below with axillary tufts of tomentum. **ACORNS:** biennial; short-stalked cup, thin walled and saucer-shaped, outer surface slightly pubescent with smooth inner surface, covering ⅓ of the nut; brown subglobular nut, ⅜ - ½ inch (10 - 13 mm) in length.

- 32 -

Figure 3. Example of Illustrations and Photographs. Image is taken from The Field Guide to Native Oaks of Eastern North America (2003).

Keys are not only limited by the amount of character information they present, but also the perspective of the author, himself or herself. As Kirchoff (2001) observes, "Even with morphological analysis, the choice of characters is influenced by the approaches the scientist used as a student". Since keys are created by experts, they are limited by the information the expert

8

conveys as relevant. This becomes apparent when comparing keys that are written by different authors that cover the same taxa (Kirchoff 2001). As a result, keys are created for an audience that does not share the same perceptions as the author.

To improve the effectiveness of keys, the methodology behind key creation must include data on the perceptions of the users themselves. As studies in psychology point out, (Murphy and Medin 1985) the effectiveness of object identification is constrained by an observer's personal knowledge. This means that a user's ability to use an identification key is dependent on their own experiences, which are not accounted for in key creation. To our knowledge, there are no keys that account for the perspective of non-expert users, and the characters they find relevant to identification. With this in mind, keys should be created that utilize information on how users interpret differences between taxa.

Research Aims and Hypothesis

The aim of this project was to investigate a new method of creating identification keys. Our main focus was to create a plant identification key that would yield a higher user performance than conventional keys. In his commentary on key creation, Andrei Lobanov (http://www.zin.ru/Animalia/Coleoptera/eng/syst8.htm) reflects that "keys are compiled by those who do not need for those who cannot use them". Reflecting on this statement, our study sought to construct a key that would be useful to both the botanical novices and botanical experts that need them. In attempt to create a key for such a diverse audience, our project gathered data on the ability of various users to recognize key differences between species. To achieve this, we constructed a visual assessment test, using plant photographs, and administered it to potential users. The resulting information was used to construct the final framework for the key.

The characters of the resulting key were displayed as high quality photographs. These photos acted as visual characters – allowing the user to identify a species using a visual representation (Kirchoff *et al.* 2011). This technique replaced the traditional idea of a character – as a written description of specified elements in an organism's composition – with a photograph of a complete structure taken from the organism itself. Using this approach, we expect that information about a species will not be lost in written translation, but maintained in a visual concept (Kirchoff 2001). We hypothesize that this holistic approach to character representation will enable users to identify plants with greater accuracy compared to traditional dichotomous keys.

To ensure the photographs were highly comparable, we followed a series of best practices proposed by Leggett and Kirchoff (2011) and employed standards of photography created by Baskauf and Kirchoff (2008). According to these standards, each species requires a photograph of the complete specimen followed by an additional set of photographs of specific plant structures: bud, twig, bark, acorn, and leaf. Each species is photographed on a standard background, maintaining a standard direction of orientation. Following the recommended standards, photos of comparative characters were displayed directly beside one another to ensure a direct comparison. Previous work on a visual key to Fagaceae of the southeast (Kirchoff *et al* 2011) also utilized these same principles.

In order to create and test a substantial key, we chose to focus on a plant taxon that would prove challenging to botanical experts. This strategy ensures that the key is a useful tool, and the taxon itself is not simply easy to identify. As such, we chose to construct a key for a single, easily accessible genus that has been notoriously challenging for botanical experts, the genus *Quercus*. Highly recognized and commercially valuable, oaks are common in temperate regions of the globe. Although common, oaks display a high degree of morphological diversity,

within and between species. Furthermore, many species, such as *Quercus rubra* and *Quercus shumardii* display a high degree of visual similarity, making them difficult to distinguish. This degree of difficulty makes *Quercus* an ideal subject to test the effectiveness of a visual key.

The visual key was tested against another key – made using traditional means – to determine if these methods enhance a user's ability to correctly identify species. During the test, two user groups were asked to identify oak vouchers using one of the two keys. The results of each key were then compared using a two-sample *t*-test. We predicted that users with our key would be able to identify oaks with greater accuracy than those using the traditional key. Since our key was created using data from potential users, we expected users to be able to use our key with greater ease. The addition of more visual characters – such as bark, twig and acorn – should enable users to get a more holistic picture of what a species looks like. Unlike traditional keys, which still rely on descriptive text, our key will enable users to actively compare species in a holistic manner, instead of selecting character states. This should allow users to practice identification and begin to develop a species concept based on their own experience.

CHAPTER II

MATERIALS AND METHODS

Photographic Standards

The genus *Quercus* is a large, globally distributed genus comprised of more than five-hundred species of oak. For the creation of our visual key, we chose to concentrate our research on species that are native and naturalized to the Southeastern United States. We based our choice of taxa on one of the most recent, and definitive keys for Southeastern Flora, Weakley's Flora of the South and Mid-Atlantic States (2015). According to this source there are currently forty-three species of oaks native or naturalized to this region. This region includes Alabama, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, West Virginia, and Virginia.

From winter 2014 through fall 2015, we employed the botanical photography practices of Baskauf and Kirchoff (2008) to photograph forty-three different species of oaks, throughout the southeast United States. These practices include photographing a particular set of plant structures and maintaining the same orientation and blank background for each shot.

The suggested structures to be photographed for woody angiosperms include a picture of the entire plant, the bark, a horizontal shot of the twig, a horizontal shot of the winter bud, the floral inflorescence, the fruit, and the leaf (Baskauf and Kirchoff 2008). For certain structures, Baskauf and Kirchoff (2008) suggest that series of photographs should be taken from multiple perspectives. The leaves, for example, should be photographed to show both the upper and lower surfaces. To demonstrate a more distinct comparison, the upper surface of a leaf should be photographed on top of a leaf with the underside exposed. We maintained this particular practice

for our study, but did not include all of the suggested standards for flower and fruit photography. We did not photograph acorns while on the branch and instead chose to focus on the acorn separately in order to capture multiple perspectives for the acorns. Like the leaves, we photographed the acorns from various perspectives: the nut separate, the cap separate, and the nut and cap together. Since the flowers of oaks are short-lived, we chose not to photograph them for this particular key.

The vouchers for all species collected are archived at the herbarium of the University of North Carolina, Chapel Hill (Appendix A). These vouchers were reviewed for positive identification by USDA botanist, Doug Goldman, a taxonomic expert who verifies plant ID's for the herbarium at Chapel Hill. A complete list of species sampled and their location is available in Appendix A.

In some cases, trees that had been sampled throughout the summer of 2014 experienced little or no acorn production. In order to acquire acorn samples for these species, some photographs were taken using acorns from herbarium vouchers at the University of North Carolina, Chapel Hill. This included the acorns for species *Q. imbricaria*, *Q. minima*, *Q. prinoides* and *Q. similis*. Other acorns specimens were collected from members of the International Oak Society and shipped to our lab to be photographed: *Q. sinuata*, *Q. coccinea*, *Q. muehlenbergii*, *Q. pumila*, and *Q. inopina*.

All of the specimens were photographed on a black background (Figure 4a) with the exception of the bark and the entire tree itself. Using this standard eliminated background distractions and made the images more comparable. For this reason, the same direction and orientation of the subject was also maintained between each species.
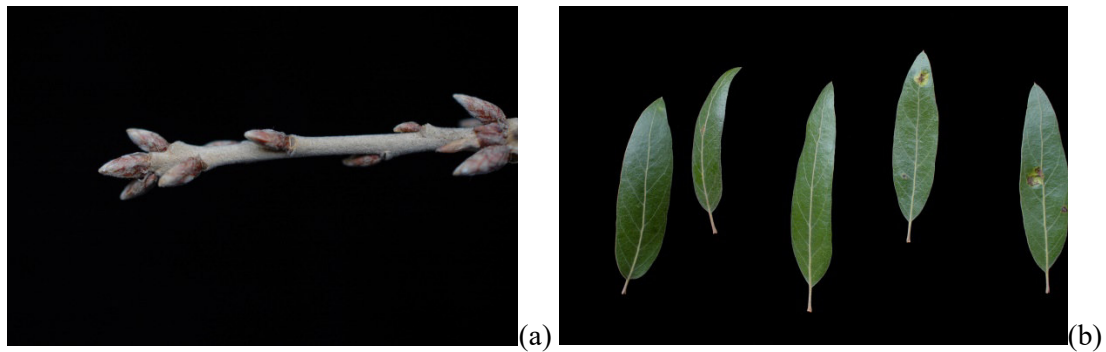
Figure 4. Photographic Standards of Oak Structures. (a) *Q. incana* winter bud.
(b) *Q. incana* leaf composite.

In addition to the Baskauf and Kirchoff standards (2008), composite photographs were taken of several of a species' leaves (Figure 4b). These composites served to demonstrate the morphological diversity within a species. Composite photographs were also taken of the acorns. Adobe Photoshop was used to create some of these composite photographs, and was also used to eliminate dust and debris from all photos. The resulting photographs served as visual concepts of a particular plant structure (Figure 4b). By using multiple examples, the user is able to formulate a more accurate concept of a particular character and the natural variability within a species can be expressed.

Taking photographs using a standard made the characters of our key more comparable. However, a challenge remained in how to best represent the comparative size of two homologous structures. Size is an important diagnostic character for plant structures and is often expressed in metric measurements or by the use of a scale bar in a photograph. Baskauf and Kirchoff (2008) suggest that the use of scale bar, especially a coin or ruler, should be such that it does not detract from the subject in the photograph itself. To overcome this obstacle, we chose to avoid using other objects for comparison and created a scaled version of the composite leaf photographs. This version would serve to show how the average sizes of the leaves are relative to one another.

We used Adobe Photoshop to make these adjustments. We determined the average height for all the oak leaves using the species' life history information available on eFloras (http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=127839). From this list we determined which species was the largest, on average, and used it as the template by which to scale the other species (Figure 5). This was found to be *Q. velutina*. A ratio was then determined between *Q. velutina* and all other species. The height of the pixels was then measured for the largest leaf in the composite for *Q. velutina*, as it appeared in Adobe Photoshop. The pixel height with which to adjust every other oak species was then calculated, using its ratio to *Q. velutina*. The photographs were then resized such that the largest leaf of the composite was relative to how it would naturally appear next to *Q. velutina*. This allowed our photographs to give the user a more intuitive perspective of the size differences between the leaves of two species.

(a)

(b)

Figure 5. Example of Leaf Resizing Procedure. Sizes adjusted for direct comparison using *Q. velutina* as a template to resize *Q. ilicifolia*. (5a) Composite photos before being resized. (5b) Composite photos after *Q. ilicifolia* is resized.

Assessment of Oak Leaf Similarities

Since there are multiple characters to consider in key creation, we chose to limit the investigation of oak similarity to leaves. Leaves are one of the most commonly used characters of botanical keys. In addition, many Southeastern oak species retain their leaves through October and November, making them highly accessible.

The construction on the framework of the key began by determining how to group individual species to be keyed out. In dichotomous keys, species that have traits in common appear together in a couplet – beginning with a broad measure of similarity. Each subsequent couplet presents the user with new characters, tapering to a fine measure of similarity. At the end of key, the species that share the most traits are displayed together in the final couplet. This process is typically done by taxonomists or field biologists who specialize in a particular set of taxa. Using their experience, specialists group species together based on their shared characteristics and construct a series of logical steps that end in identification. This often involves creating a dendrogram. Dendrograms are tree-like graphs that depict the relationships of objects in a progressive hierarchy of similarity. Alike objects are grouped into similar clusters, or branches. Object that are most similar appear together at the bottom of the tree. In biology, dendrograms were once used to depict theoretical phylogenetic relationships, but they have been also useful in key creation. Dendrograms can act as a map or a physical representation of the steps a user takes in a dichotomous key (Figure 10). Each branch represents a decision in the couplet with the terminal branches representing the identity of the species.

In order to achieve this, we created a dendrogram for our key using information gathered on how these potential users group oaks. We used a simple rating system to allow users to assign a value of similarity to pairs of different oak leaves. These similarity ratings were then

16

used to generate a distance matrix that could be turned into a dendrogram – a description for the key's progression.

Using the leaf composite photographs, we created an online test to observe how users with different levels of botanical experience asses the visual similarity of oak leaves. Two groups of potential key users were considered for the assessment: "non-expert" and "expert". Here we defined "non-expert" as those having little to no formal training in botanical taxonomy, such as undergraduate students and members of the general public. Conversely, we defined "expert" as individuals that have demonstrated formal training in plant identification – particularly botanists, graduate students and members of various botanical organizations, such as the Master Gardeners.

The test was conducted online and a submission was provided to the UNCG Internal Review Board (IRB) for human research subjects testing. Per standards of the UNCG IRB, our research was determined not to be human research subjects testing. Online test participants were recruited using two different methods. Participants from the "expert" group were recruited from biology programs and professional botanical organizations. These participants were sent an email, with a description of the project and a link to the survey (Appendix B). Members of the "non-expert" category were recruited using Amazon Mechanical Turk. This service is provided through Amazon as a means of crowd sourcing information. Members of Amazon Mechanical Turk can register as "requestors" – survey providers – or as "workers" – survey participants. Workers can freely browse and choose surveys that they wish to participate in. Requestors, in turn, can create and post surveys directly through Amazon or publish a link to another website. These surveys are posted as Human Intelligence Tests (HITs), along with a brief explanation of the task and a summary of the compensation received once the task is complete. Our workers were paid three dollars for their participation. Requestors can also recruit "Master Workers" who have demonstrated a high performance on particular types of tasks, such as categorization. These

17

workers are prescreened by Amazon Mechanical Turk and assigned a Master status if they have demonstrated a superior performance on tasks that require object categorization. To ensure quality performance, we chose this option for our own research.

The visual assessment test was created using Qualtrics software and distributed to either participatory group through an embedded hyperlink. Before the test began, participants were asked to rate their knowledge of Southeastern oaks by selecting a rating between "0" and "3" on a sliding bar (Figure 6). The test was composed of three practice questions, fifty-eight actual questions, and two catch trials.

Figure 6. Qualtrics Survey of Participant's Previous Knowledge. 0 = "Not Confident: not familiar with oaks at all"; 1 = "Not very confident: can ID 5 species or less"; 2 = "Somewhat confident: can ID between 6 and 10 species; and 3 = "Very confident: can ID more than 10 species".

Each of the fifty-eight test questions displayed two, randomly generated leaf composite photographs that were positioned side-by-side (Figure 7). Below the two photographs, a sliding bar was displayed with ratings zero through seven (Figure 7). During each test question, the user was asked to rate the degree of similarity between the two using a fixed scale between "0" and "7". A rating was required for each question in order for the participant to move on to the next question.

Figure 7. Qualtrics Visual Assessment Question. 0 = "Completely different objects: not a leaf"; 1= "Extremely Dissimilar"; 2 = "Somewhat Dissimilar"; 3 = "A Little Dissimilar"; 4 = "A Little Similar"; 5 = "Somewhat Similar"; 6 = "Extremely Similar"; and 7 = "No Difference: Exact same image"

In addition to the fifty-eight questions, participants were given two questions that served as catch trials (Figure 8). In psychological studies, catch trials serve to refocus a participant's attention or to identify users who may need to be eliminated from the study. Since the catch trial's purpose is to identify potential errors, specific answers were required in order for the user to continue. The first catch trial served as an initial warning to keep the user on task and asked the user to "try again" if they did not answer correctly. However, if the user answered the second catch trial incorrectly they were automatically eliminated from the study.

In our catch trials, the scale ratings "0" and "7" represented extreme cases that the user was instructed to select when presented with images that were appropriate. For example, one of the catch trials presented a photograph of an orange teapot next to a leaf composite (Figure 8). In

the introduction, the user was told to select the rating "0", or "Completely different objects: not a leaf" for these extreme dissimilar pairing. These objects were selected to have little in common with the leaf composites. All of the chosen objects were single, man-made items; were not green, and were displayed on a white background, instead of black.
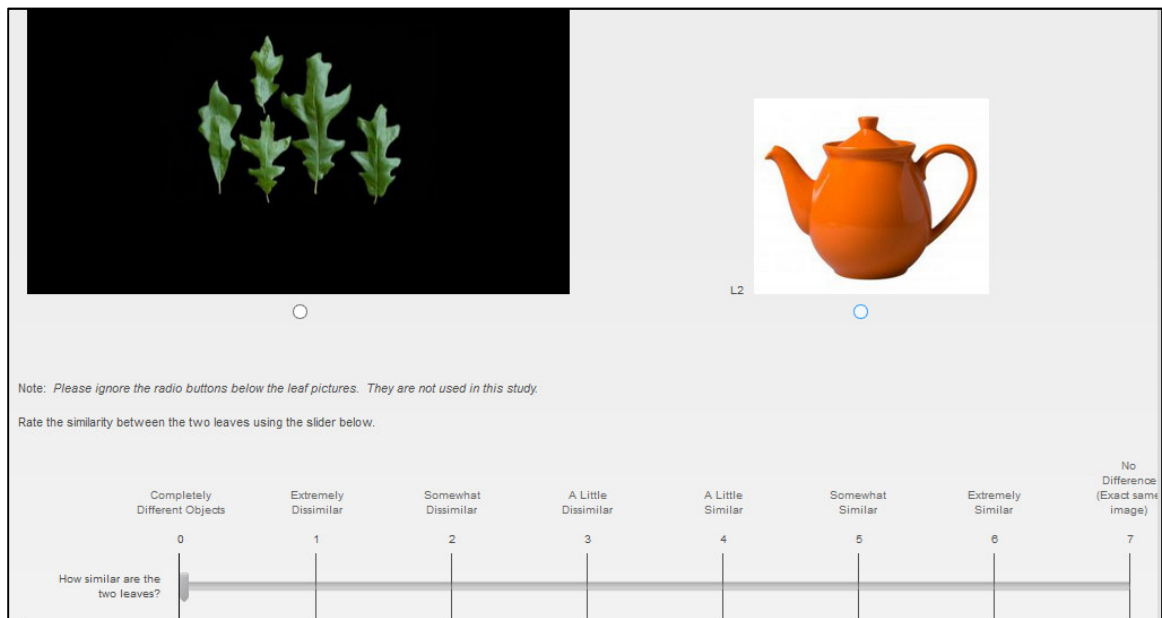
Figure 8. Sample Catch Trial Question. Participants were instructed to select 0 = "Completely different objects: not a leaf"; at the beginning of the session. 1= "Extremely Dissimilar"; 2 = "Somewhat Dissimilar"; 3 = "A Little Dissimilar"; 4 = "A Little Dissimilar"; 5 = "Somewhat Similar"; 6 = "Extremely Similar"; and 7 = "No Difference: Exact same image".

The catch trials were not only used to ensure the participant's attention, but served to identify data that may contain errors. Users who answered the catch trials correctly had their data recorded on an excel sheet in Qualtrics. This data was reviewed to determine if participants understood how to interpret the scale bar used in the survey. Participants were removed if they had the ratings "0" and "7" when not answering a catch trial question. For example, if a user was shown two dissimilar leaves and rated the pairing as a "0: Completely different object" then their ability to understand the scale bar was questioned and all of the participant's responses were

removed. To further determine the quality of responses, the answers to the first ten questions were removed from each participant. This ensured that the remaining participant's ratings were more accurate, given more time to acclimate to the test procedures.

Over 200 experts and 200 non-experts were polled for the visual assessment tests. The final eliminations from gleaning out the unusable data left 200 hundred participants polled for the "non-expert" and 241 participants polled for the "expert" group. Within each potential user group, at least ten ratings were collected for each species comparison, of which there were 903 possible species pairs. These response ratings were then averaged to create a mean rating for each species pair comparison. These means were then compiled into a distance matrix using Excel (Appendix C and D) and the resulting matrices were used to create two prototype dendrograms, one for non-expert and one for expert similarity ratings.

Construction of Visual Key Framework

A user's path through an identification key begins with the assumption that all species are separate individuals. As such, we chose to use an agglomerative hierarchy to create a dendrogram for our key. This method assumes all individual points of data to be initially unrelated, and gradually joins them together into similar groups, or clusters. The end result is a tree with all individuals belonging to the same cluster at the top of the tree, and all separate individuals partitioned out as solitary clusters at the bottom. This is often called a "bottom-up" approach to tree construction. To accomplish this, we imported the distance matrices into the program R, and created hierarchical clusters using the function *hclust*.

In order to cluster individuals, we explored three different methods of linking species: average linkage, centroid linkage, and Ward's minimum variance method. Each of these methods displays the data points in Euclidean space and gradually lumps individuals into clusters based on

specific criteria. The average method of linking clusters joins individuals into a cluster based on

the arithmetic mean of all objects within an extant cluster to the data point considered (Sneath and

Sokal 1973). In contrast, the centroid linkage method creates a centroid, or a group mean, for an

existing cluster. Individuals are then joined to the cluster by finding the nearest individual to the

centroid itself (Sneath and Sokal 1973).   Similarly, Ward's minimum variance method joins

clusters based on the centroid. In Ward's method, however, individuals are selected to join

clusters that minimize the standard deviation of the resulting cluster, once a new member has

joined. Thus the within-cluster variance is minimized (Sneath and Sokal 1973). Theoretically,

the Ward's minimum variance method should produce the most conservative dendrogram for our

purposes. By joining data points based on their effect upon the overall cluster, the result is

clusters that have a low degree of within-cluster variance. The R code for this procedure is listed

in Appendix F.

Upon comparing the three resulting dendograms, the centroid method produced a

dendrogram with overlapping clusters (Appendix E). The result was a highly complex tree with

no clear progression and was not considered for our final model. Both the average and Ward's

linkage methods produced dendrograms with clear progressive pathways. The average method,

however, produced clusters that demonstrated a high degree of separation between different

species groups (Figure 9). While this is not completely undesirable, eliminating species early

potentially leaves room for more error as the user may eliminate a species too early without

having the opportunity to see the potential similarities it has to other species. For example, the

average linkage method separated *Q. velutina* from other species, early in the dendrogram's

progression. While this elimination can be useful to lower the number of possible identifications,

it only eliminates one possible species, rather than showing a distinction between groups. By

contrast the Ward's clustering method produced a dendrogram that separated species into larger,

22

distinct groups – keeping a more tight association between species within those groups. This clearly demonstrates the mathematical process that Ward's method uses to clusters individuals. As a result, we chose the Ward's method of linkage to minimize outlier groups and give the user the opportunity to make more decisions.
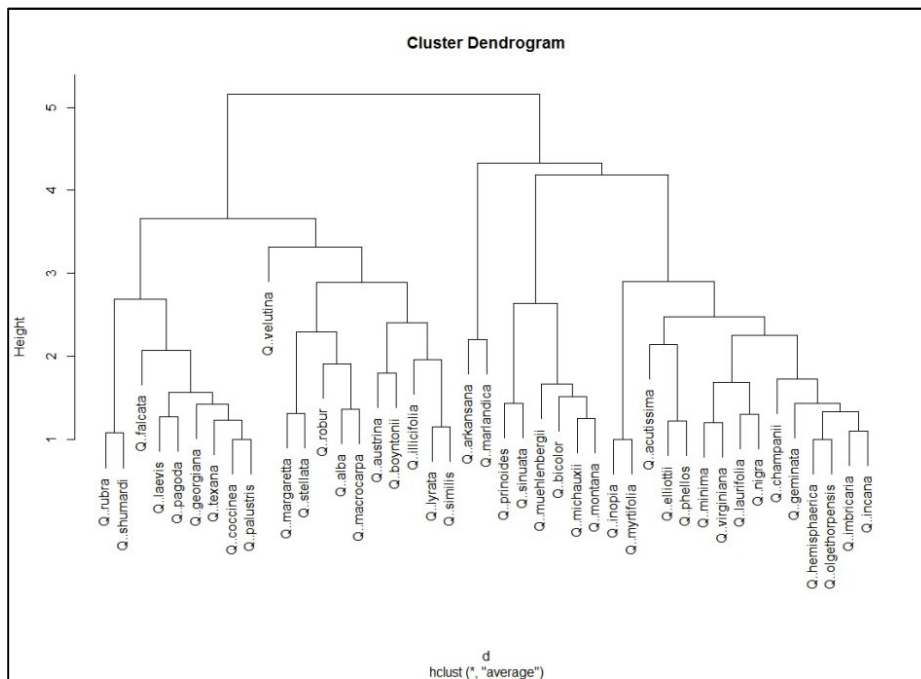
Figure 9. Dendrogram Created Using Average Linkage Method.

Dendrograms for both the "non-expert" (Figure 10) and "expert" data (Figure 11) were constructed using Ward's method. These dendrograms served as potential models for the final map of the key. To determine how similar the dendrograms were to one another, we created a consensus tree in the program Mesquite (mesquiteproject.org) (Figure 12). This consensus tree showed which branches the two dendrograms had in common by displaying them as they would appear in both dendrograms. All other branches that were not shared in common were displayed as single, terminating branches. The resulting consensus tree showed that the "expert" and "non-

23

expert" dendrograms shared eight branches in common (Figure 12). Therefore, these branches were maintained in the final model.
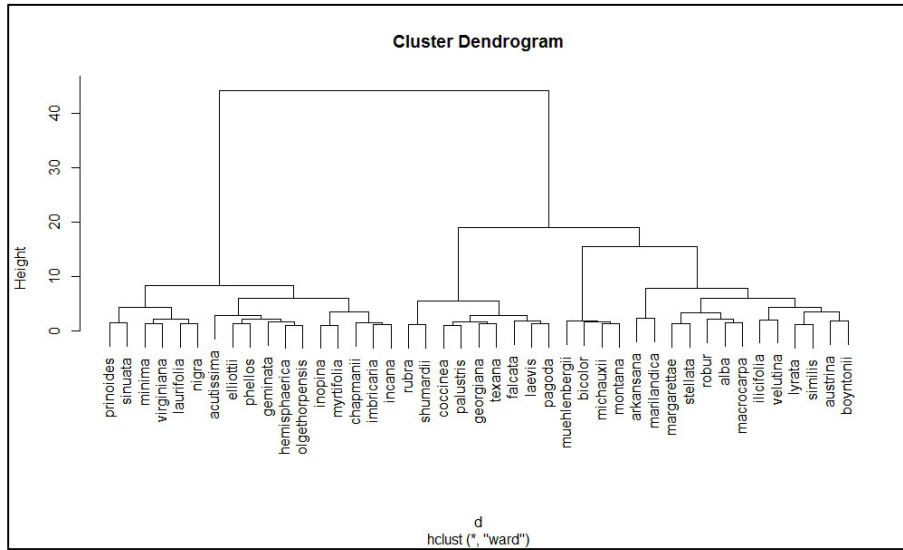
Figure 10. Non-expert Dendrogram. Results from "non-expert" distance matrix were used in conjunction with Wards clustering method to produce a dendrogram.
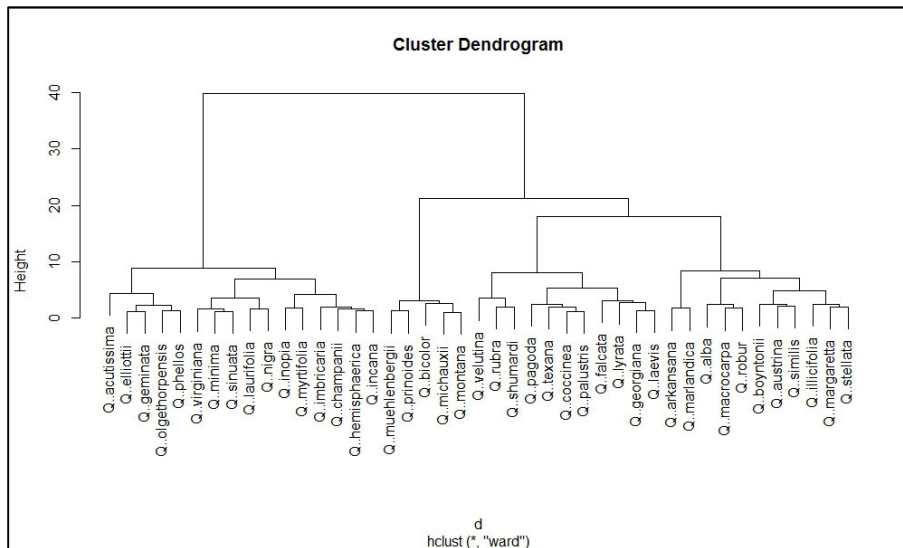
Figure 11. Expert Dendrogram. Results from "expert" distance matrix were used in conjunction with Wards clustering method to produce a dendrogram.

Figure 12. Consensus Tree. Red highlighted branches show shared branches between "expert" and "non-expert" dendrograms.

The remaining twenty-five branches represented individual species. Upon analyzing the different placement of these species, we noticed that the placement of some species appeared to differ only slightly between the two dendrograms. Looking at the branch lengths for both dendrograms, we observed that the height of some branches were less than two (Figure 10). This height represented the distance between two clusters, calculated by the Lance-Williams algorithm of Ward's method. For example, both dendrograms contained clades with the species *Q. incana*, *Q. imbricaria*, and *Q. chapmanii* (Figure 13b). However, the expert dendrogram contained *Q. hemispaherica* in this clade (Figure 13a). Regardless of members, the branch heights separating all members were less than two. This could indicate that the users did not detect a difference between several species. Therefore, we suspected that there may not be a significant difference between how the experts and novices were rating all three species. Looking at the similarity matrices of the expert and non-experts, we ran a *t*-test to determine if there were significant differences between the way each users group rated these four species. The results revealed that there was no significant differences (p>0.1) in the way non-expert and experts rated the four

25

species to one another. The only notable difference was the placement of *Q. hemisphaerica*, which was rated to be more similar to *Q. oglethorpensis* by the non-experts. As a result, we chose to keep *Q. hemisphaerica* in both clades and collapse the branches separating it from *Q. incana* and *Q. imbricaria*. This resulted in a trichotomy (Figure 13) which would give users more information by allowing them to observe three potential character concepts, rather than two.
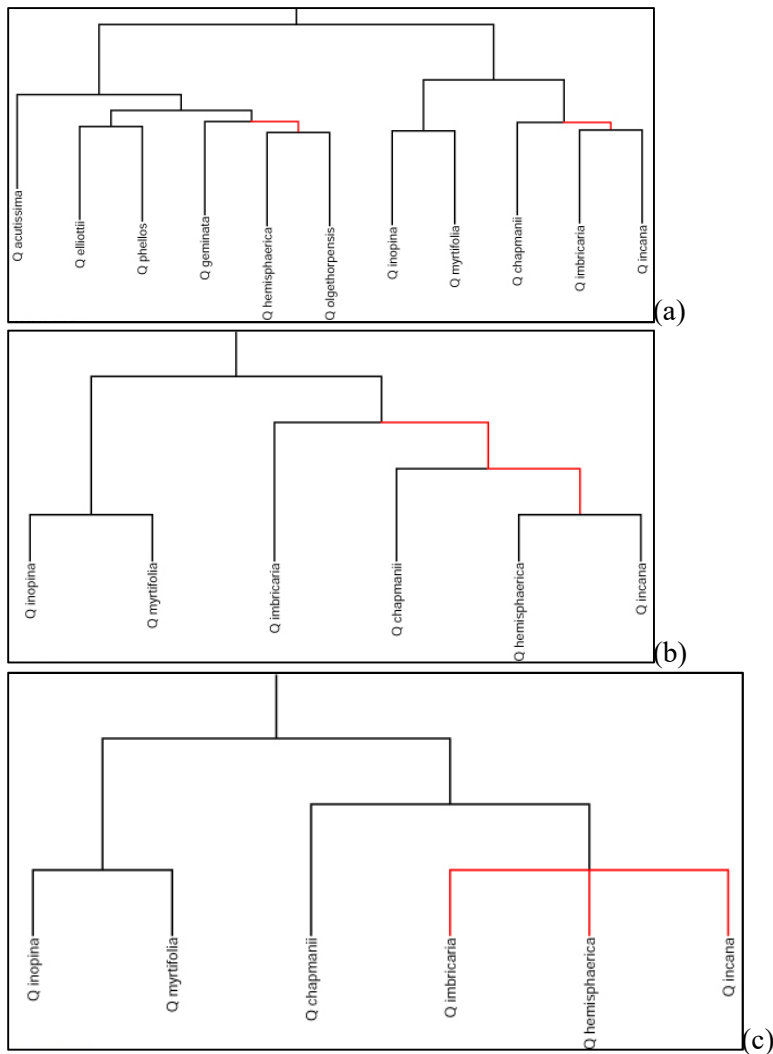
Figure 13. Example of Collapsing Branches into a Trichotomy. (a) Clade from Non-expert Dendrogram. Red highlights the branch heights that are less than 2. (b) Clade from Expert Dendrgoram. Red highlights the branch heights that are less than 2. (c) Trichotomy Resulting from Branch Collapse. Red highlights the final trichotomy from the collapsed branches.

We chose not to collapse the branch containing *Q. champanii* because we recognized it to be morphologically distinct from the other three species. Since *Q. champanii* has broad lobes that the other species do not demonstrate, we chose to emphasize these features in the final key to give the user a defining character of the species. Similarly, we did not collapse every branch that had a height less than two. For example, we noted distinct differences in the morphology of *Q. geminata* that could be emphasized to separate it from *Q. olgethorpensis* and *Q. hemispaherica* on the non-expert dendrogram (Figure 13). These decisions were based off of our expert knowledge of the species. This procedure was also used to investigate other clades whose members were separated by small branch heights of less than two. This resulted in six trichotomous branches and two tertachotomous branches present in the final dendrogram (Figure 16).

As with the case of *Q. hemisphaerica*, we considered allowing other species to exist in two different clades for the final dendrogram. We began by investigated how the two user dendrograms differed in the members contained in each clade. From this examination, we selected four additional species to be placed in two alternative clades in the final dendrgoram: *Q. lyrata*, *Q. prinoides*, *Q. velutina*, and *Q. virginiana* (Figure 14, highlighted in red). This technique of allowing a species to be identified along two separate paths in a key is common, especially if a species is polymorphic. *Q. viginiana*, for example, has two major morphological forms – a long, linear form and a pointed, holly-like form. Based on this observation, we decided to allow *Q. virginiana* to exist in two different clades. The result was a clade with *Q. virginiana*, *Q. laurifolia*, and *Q. nigra*, representing long, linear leaves (Figure 15a); and another clade with *Q. virginiana*, *Q. minima*, *Q. sinuata*, and *Q. prinoides*, representing the holly-like morphology (Figure 15b).
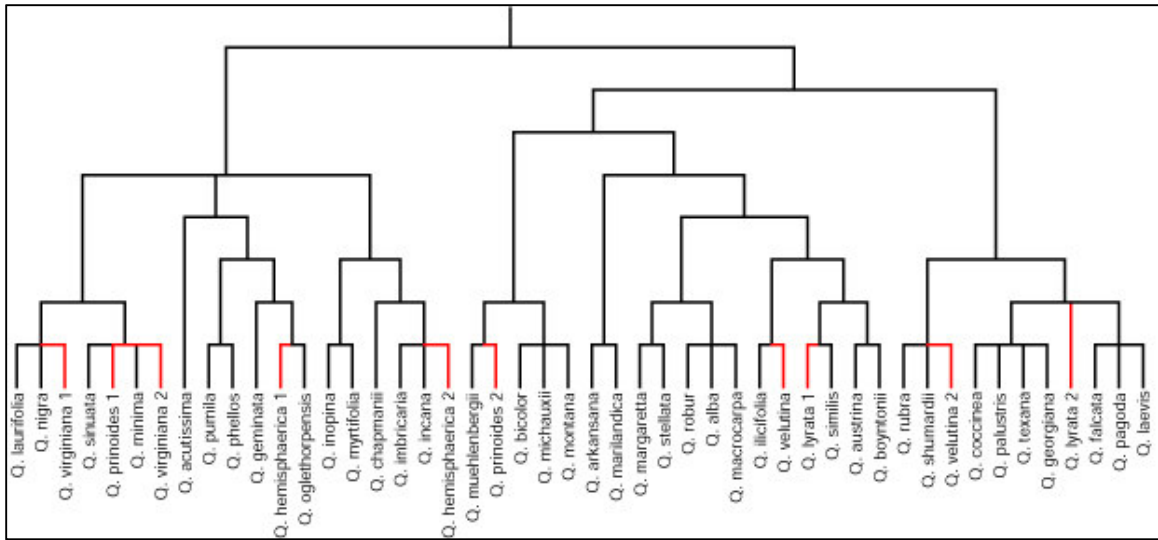
Figure 14. Species that Occur in Two Separate Branches. These species include *Q. virginiana*, *Q. prinoides*, *Q. hemisphaerica*, *Q. lyrata*, *Q. velutina*.

The final dendrogram for the key resulted in a combination of traits from both the "non-expert" and "expert" dendrograms, but with an emphasis on the similarities established by the non-experts. We chose to maintain much of the non-expert model because it represented the audience that could potentially benefit the most from the visual key. In addition, this model was a better fit with our understanding of the similarities of oak leaf morphology. The final model served as the map for the user's progression through our visual key (Figure 16). Each branch represented the point at which the user was asked to make a decision between two or more species concepts (Figure 19).
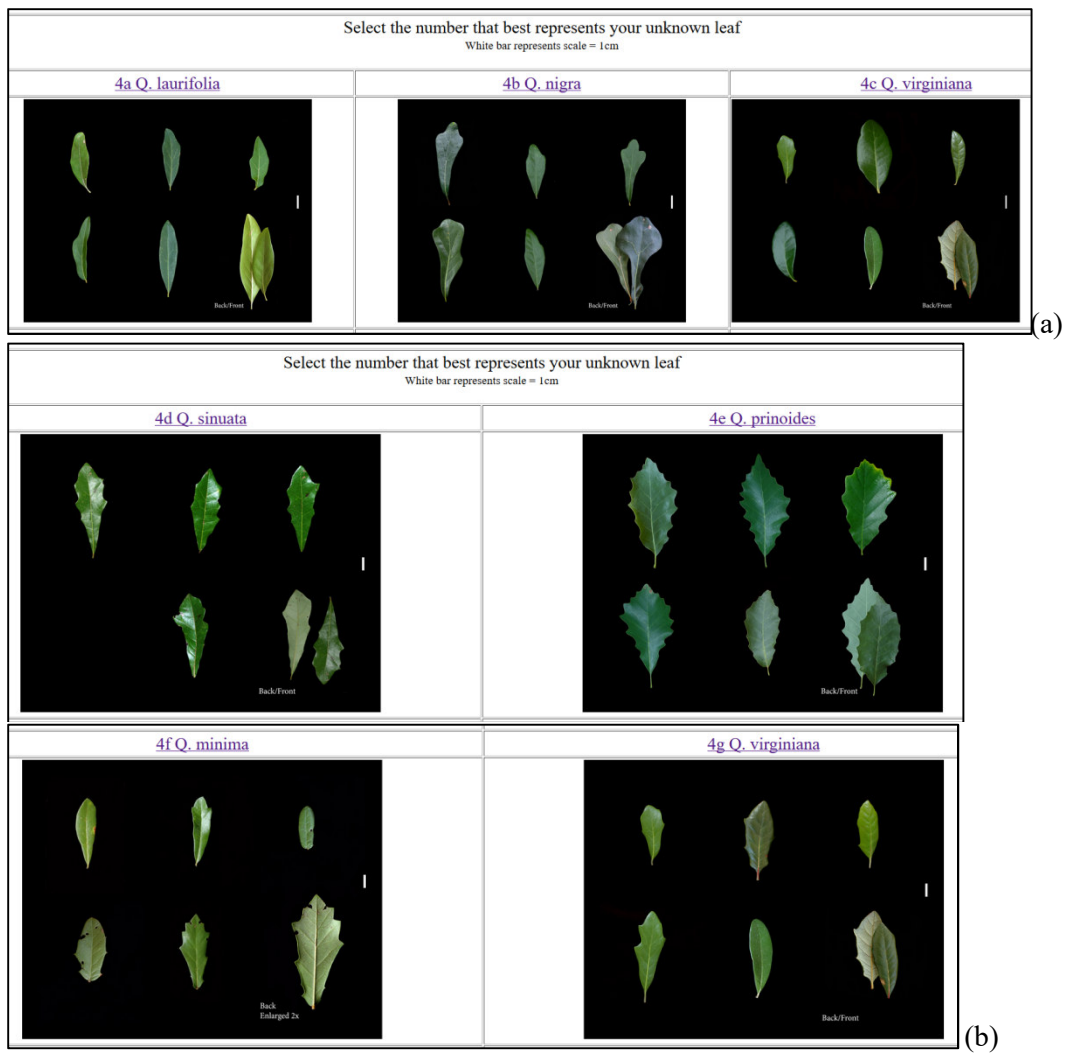
Figure 15. Clades in Final Key that Contain *Q. virginiana*. (a) Clade containing long, linear species with slight lobbing. (b) Clade containing "holly-like" leaves.
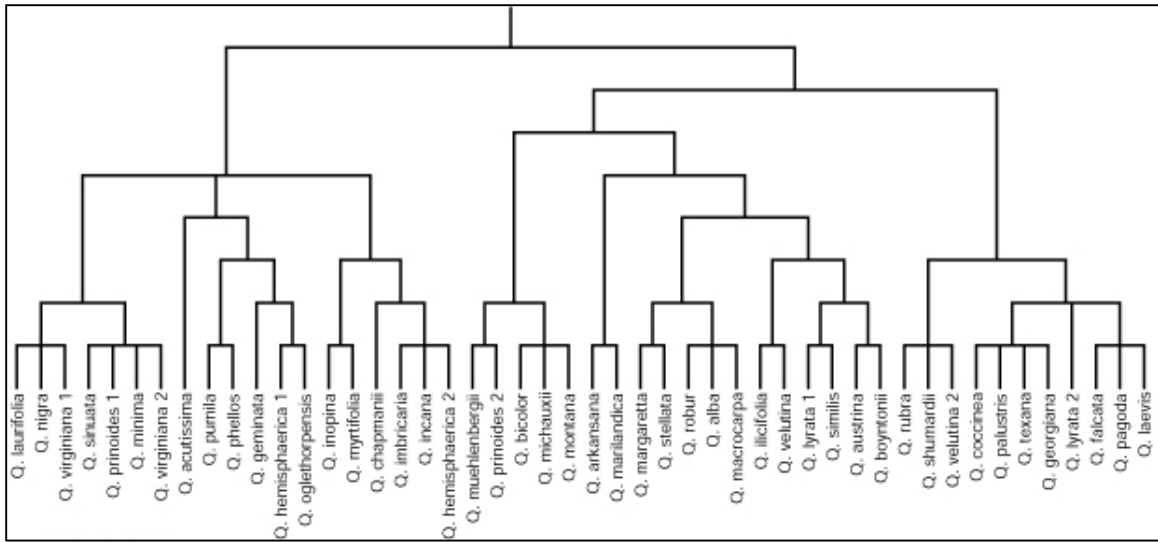
Figure 16. Final Dendrogram for the Visual Key.

The final key was written into html using the html editor, Kompozer (http://www.kompozer.net/download.php). Each step of the dendrogram displays as a webpage with two or more photographs for the user to consider (Figure 17). These photographs represent visual character concepts, replacing the text used by traditional keys. Each photograph consists of a composite image of leaves from possible species within that category (Figure 18). These composites were created in Adobe InDesign and displays six leaves evenly spaced on a black background. The leaves of all species shown in the photographs are scaled relative to one another to give a more accurate depiction of how similar they are in size. Each photo represents a concept of the species contained within the branches below it. For example, one node above *Q. hemisphaerica* and *Q. imbricaria* shows reprentations of three *Q. hemisphaerica* leaves and three *Q. oglethorpensis* leaves (Figure 18).
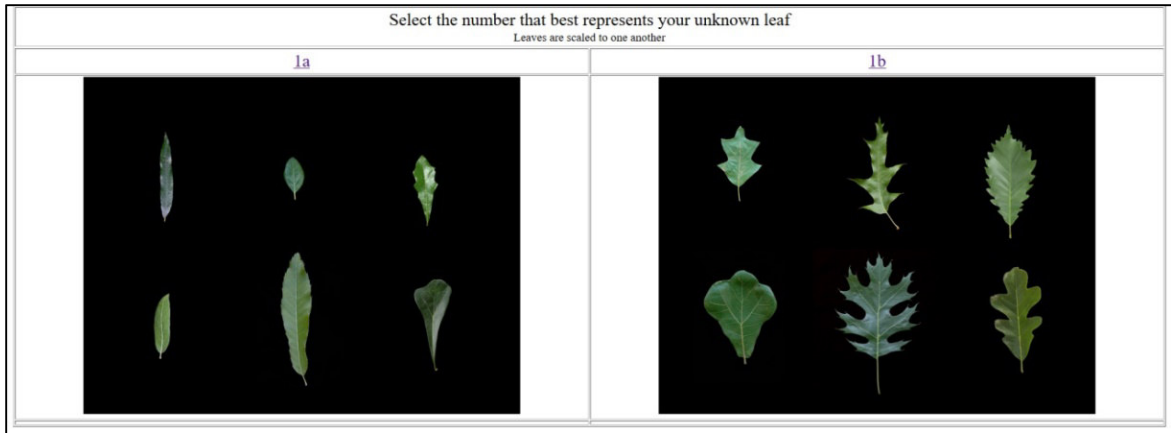
Figure 17. First Step in the Visual Key.

Figure 18. Node 5d in the Visual Key. Picture composed of *Q. hemispaherica* and *Q. olgethorpensis* leaves.

We tested our images in order to verify that the composites made were accurate representations of the species within a particular node. We printed a photocopy of each composite onto a card and assigned it a number based on it's position within the dendrogram. These cards were then spread over a table in the order they would appear in the final key. Volunteer participants were asked to use the cards to key out mounted voucher specimens. If a species was not keyed out correctly, we asked the participants to back track through the key to

31

determine which leaf in the composite lead to their decision. The composite photographs were then altered to better represent the taxa contained within them. A similar method of trial and revision was used by Kirchoff *et al*. (2010) in the creation of a key to Fagaceae. When the composite photographs were finalized, they were inserted into the key at the appropriate node they represented (Figure 19).
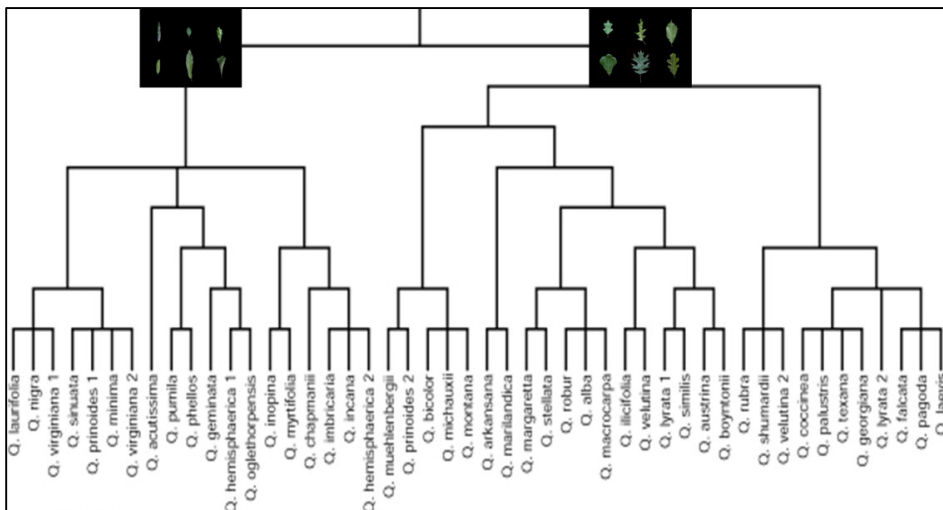
Figure 19. Composites Representing their Respective Node.

While the leaf composite photographs served as the main visual characters, supplemental photographs were placed into the key to represent additional characters. These include photographs of other oak structures – bud, bark, acorn, and twigs (Figure 20). Since the map of the key was created from a visual assesment data on leaf similarities, we could not place the additional photographs in the main body of the key. Therefore, these additional characters are given at the end of a final node to aid the user with the final identification (Figure 20).
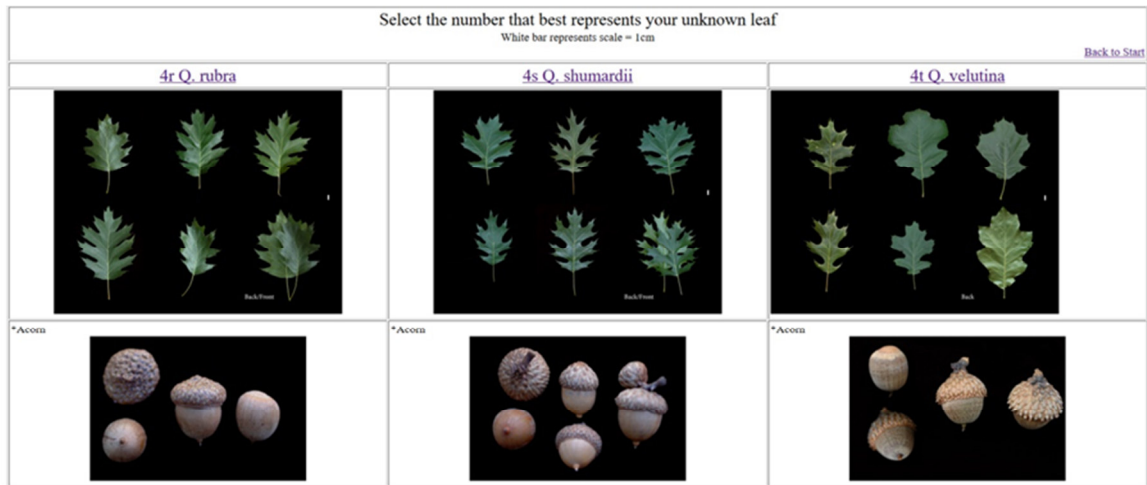
Figure 20. 4r-4t in the Visual Key. Picture displays a composite photo of leaves and acorns for three species: *Q. rubra*, *Q. shumardii*, and *Q. velutina*.

When a final selection is made, the user is asked to click on the name above the species they identified. These names are hyperlinked to a page containing additional photographs and a link to an outside page with supplemental species information. This link takes the user to a species page on eFloras website, provided by the the Missouri Botanical Garden. eFloras is the online home of the Flora of North America project, the defintive reference for North American plant taxa (http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id=127839).

Testing of the Key

The final visual key was tested against an illustrated key created by the U.S. Forest Service, Field Guide to Native Oak Species of the Eastern North America (Stein, Binion, and Acciavattii 2003) (Figure 21). This key provided some illustrated elements in both the key and the species descriptions. The key portion itself used few characters other than leaf and stem elements, which were the focus of our own key. The Forest Service key is also available in printed and pdf format, which enabled us to easily create an html version of the key, similar to

our own. The combination of these factors made the Forest Service key an appropriate key to be compared to our visual key.
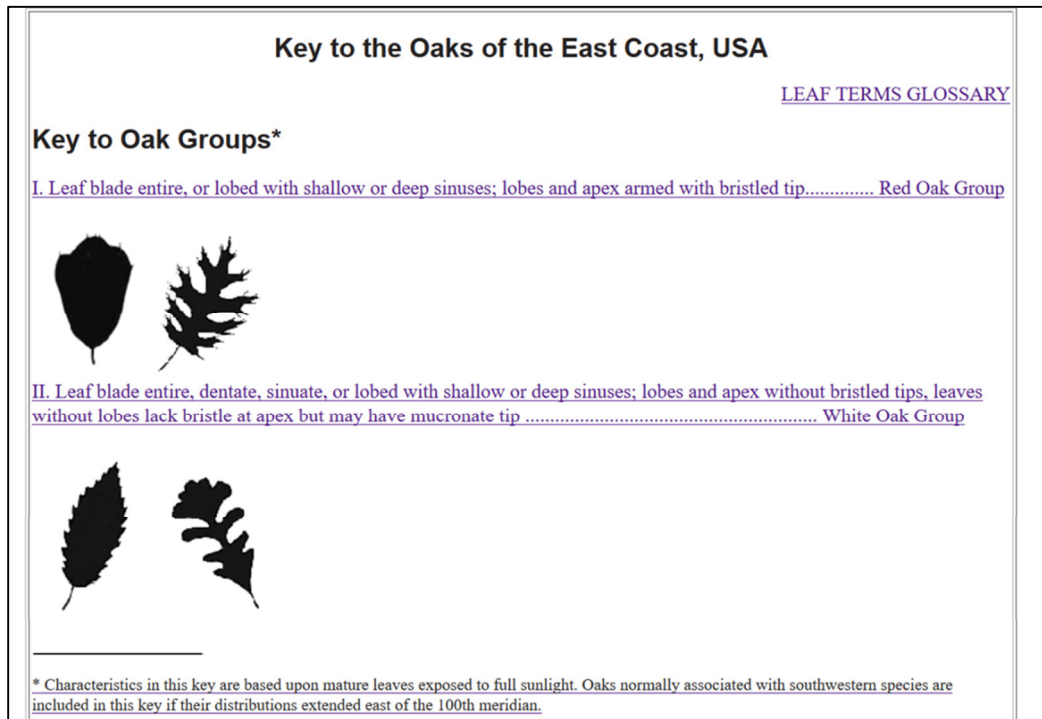
Figure 21. First Page of Forest Service Key. Forest Serivce key found in Field Guide to Native Oak Species of Eastern North America (2003).

Both keys were posted on the UNCG server and tested on two different groups of potential users: biology students and field experts. The first test recruited students enrolled in the BIO 354: Plant Diversity class at UNCG. Sixty-three students were tested in a controlled setting over one week. Each student was preassigned one of the keys, which were designated as "Key A", for our visual key, and "Key B", for the Forest Service guide. During the test, students were asked to identify ten different species of oak: *Q. bicolor*, *Q. falcata*, *Q. hemisphaerica*, *Q. lyrata*, *Q. macrocarpa*, *Q. montana*, *Q. muehlenbergii*, *Q. phellos*, *Q. rubra*, *Q. stellata*. Students were given one hour to identify the ten species and were allowed to back track through the key, if they

felt they had reached an incorrect identification. All ten species were represented by dried herbarium vouchers. We used dried vouchers in order to keep the same vouchers constant throughout the student and field expert test sessions. None of these vouchers were taken from a tree that had been photographed for the visual key. All voucher cuttings possessed acorns, with the exception of *Q. lyrata*, *Q. phellos*, and *Q. stellata*. All ten species were specifically chosen in an attempt to minimize the number of steps in each key that it took to achieve an identification. In our visual key, it was not possible for a species to take more than eight steps to identify. Therefore in preparing the test we chose species that could be correctly identified in eight steps or less, using the Forest Service Key (Table 1).

Table 1. Number of Steps to Positive Identification Between Both Keys. Species with two steps listed, appeared in two different places in the key.

Species	Steps in Key A	Steps in Key B
Q. bicolor	5	6
Q. falcata	5	6
Q. hemisphaerica	5 & 6	7
Q. lyrata	4 & 8	6
Q. macrocarpa	7	5
Q. montana	5	7
Q. muehlenbergii	5	6
Q. phellos	5	3
Q. rubra	4	7
Q. stellata	7	8

The second test examined the performance of botanical field experts using both keys. In particular, we tested experts that were familiar with *Quercus*, but who were not experts in identifying the oaks from the Southeastern United States. Twenty participants were recruited from attendees at the 8th International Oak Society Conference in October, 2015. As with the biology student test, participants were given the same ten voucher species to identify. These

vouchers were split into two different groups: Set 1 and Set 2 (Table 2). Each group contained one set of species that were morphologically similar (ex: *Q. phellos* and *Q. hemispaherica*). Using a within-subjects design, the expert participants were asked to identify a series of species labled as "Set 1" using either Key A or Key B. After completing this first set, they were then asked to switch to the opposing key to identify species of "Set 2" (Table 3).

Table 2. Species in Test Session 2. Species in "Set 1" and "Set 2" of Expert Identification Test.

Set 1	Set 2
Q. bicolor	Q. hemisphaerica
Q. lyrata	Q. phellos
Q. stellata	Q. falcata
Q. montana	Q. macrocarpa
Q. muehlenbergii	Q. rubra

Table 3. Within-subjects Contingency Table.

	Species Set 1-5	Species Set 6-10
10 participants	Key A	Key B
10 participants	Key B	Key A

Despite efforts, some of the experts were already familiar with some of the species; therefore, participants were asked to suspend their assumptions of species identity and adhere to the guidelines provided by either key. As with the student participants, the field experts were allowed to back up during the keying process, but not if they reached a final identification that they believed was incorrect. This stipulation ensured that the field experts were following the instructions provided by the key, not relying on their own knowledge of a voucher's identity.

In addition to the test itself, both user groups were given a survey after using each key. These surveys evaluated the users's previous experience with identification keys and allowed

them to give feedback on their experience using test key. Complete copies of both surveys are

listed in Appendix G and H.

CHAPTER III

RESULTS

Test 1: Results of Biology Student Identification Test

For each user group we tested, we ran a two-sample *t*-test to determine if the participant's scores were significantly different when using either key. First, we examined the results of the test session one, with the biology students. Results of the two-sample *t*-test showed that biology students using the visual key scored an average of 69.8% correct, compared to students using the Forest Service key, which scored an average of 46.9% correct. This yielded a significant difference between the scores of the two keys ($p \ll 0.001$), with students using the visual key scoring an average of 22.9% higher than those using the Forest Service key.

Despite efforts to minimize differences between the two tested keys, there was a difference in the number and identity of species contained within each key. The Forest Service key, for example, included species that were native to eastern Texas, such as *Q. fusiformis* (Table 4). Likewise, the visual key included two species that have been naturalized to the Southeast United States (Table 4). Furthermore, the Forest Service had a total of seven more species than the visual key. This discrepancy left room for a greater possibility of error with the users of the Forest Service key.
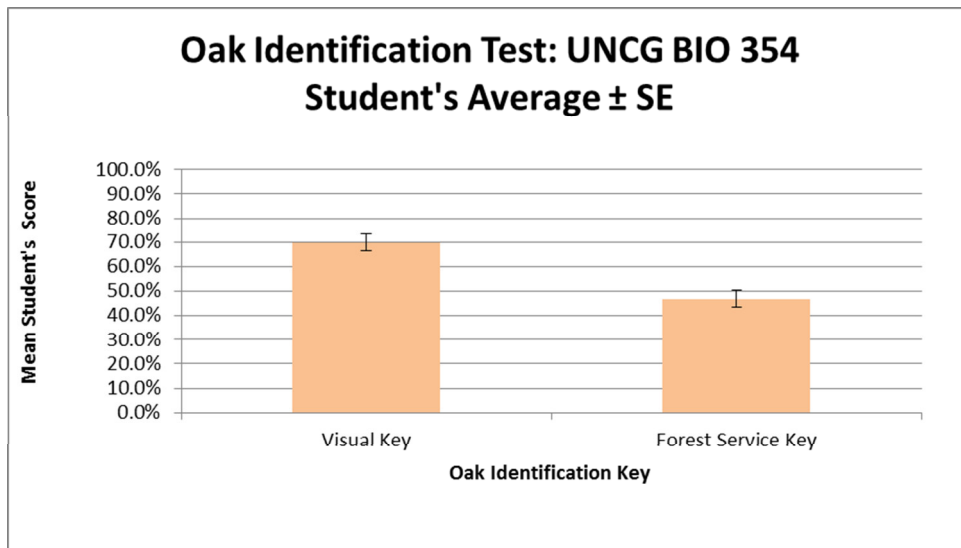
Figure 22. Results of Test Session 1.1. Mean scores of BIO 354 students at UNCG. Error bars indicate ± SE.

Table 4. List of Species Not Present in Both Keys.

Species in Key A not in Key B	Native Location	Species in Key B not in Key A	Native Location
Q. acutissima	Asia	*Q. acerifolia*	AK
Q. robur	Europe	*Q. buckleyii*	TX, OK
		Q. ellipsoidales	MI, OH, MN, WI, IA, IL, IN, ND, MO
		Q. fusiformis	TX, OK
		Q. havardii	TX, OK, NM
		Q. laceyi	TX
		Q. mohriana	TX, OK, NM
		Q. pungens	TX, AR, NM
		Q. vaseyana	TX

To test whether this inconsistency had an effect on the student's score, we eliminated all test questions in which a user chose a species from this list and recalculated the percent correct (Table 4). The results from this elimination (Mean visual key score: 70.3 ± SE; Mean Forest Service key score: 50.1 ± SE) showed that the mean score of the visual key remained significantly higher than the mean score of the Forest Service key (p<<0.001) (Figure 23).
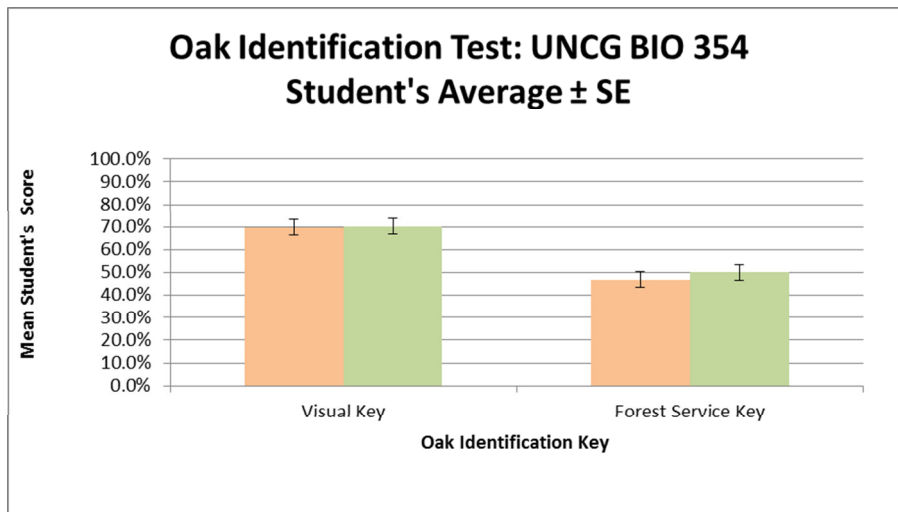
Figure 23. Results of Test Session 1.2. Green bars show mean scores of biology students at UNCG with species removed that did not occur in both keys (Table 4). Error bars indicate ± SE.

In addition to the two keys having a different number of species, there was also a discrepancy in the types of characters each key presented to the user. While both keys focused on leaf characters, our visual key provided photographs of additional structures at the end of the key, just before identification. One such addition was the inclusion of acorn photographs. Since over half of the vouchers themselves had acorns, the inclusion of acorn photographs in the visual key may have contributed to the high percentage of correct identifications with the visual key. To test whether or not the acorn photographs gave the visual key such an advantage, we reevaluated each user's identifications and counted a misidentified species as "correct" if a user's answer was in the same node or couplet as the true identification. For example, if a student using the Forest Service Key answered *Q. laurifolia*, for a *Q. hemisphaerica* voucher, we counted the answer as correct because they appear in the same couplet (Figure 24). This ensured that both keys were effectively leading the user to the correct node by leaf characters alone.

15 Leaf blade thin, apex obtuse or rounded; found on moist soil .... Q. laurifolia

15 Leaf blade thick and leathery, apex acute or acuminate; found on dry soil .... Q. hemisphaerica

[back to 13]
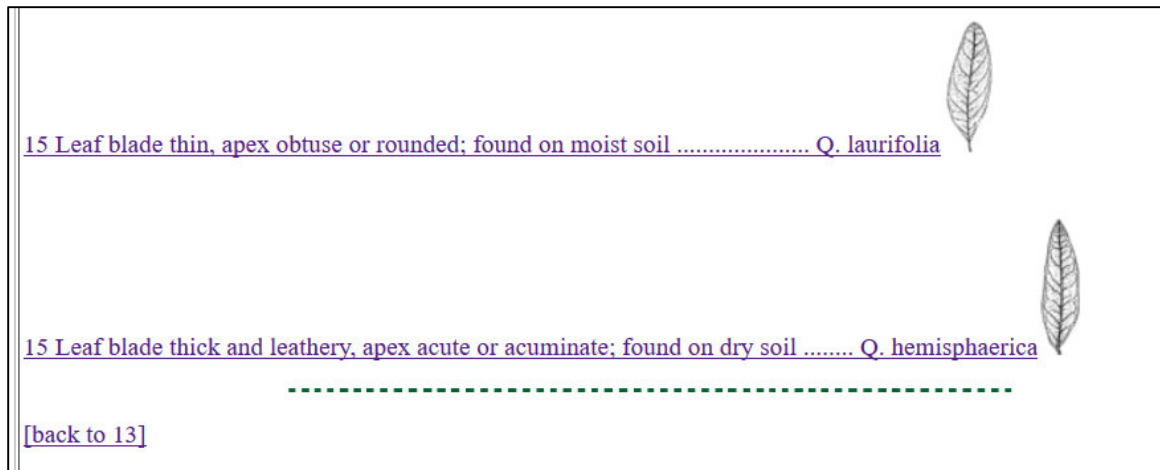
Figure 24. Node 15 of the Red Oak Group from Forest Service Key.
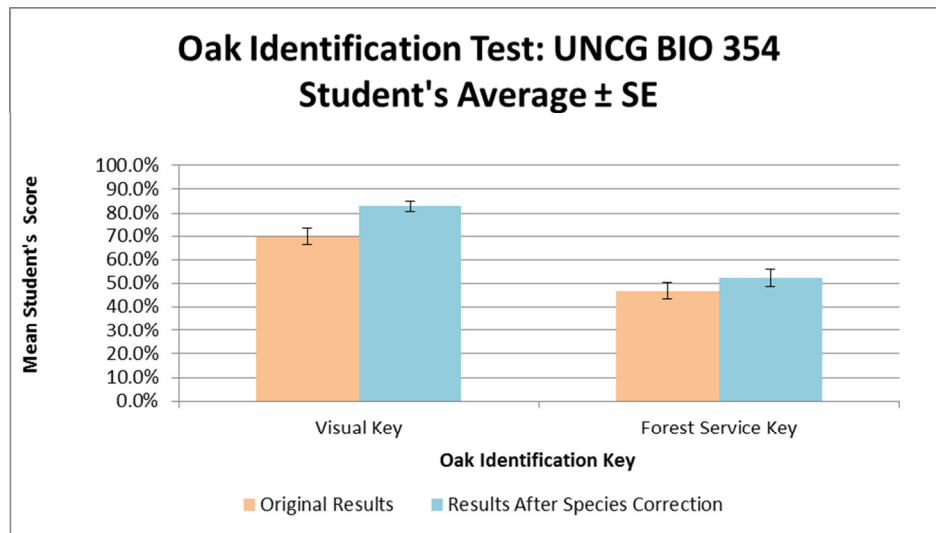
Figure 25. Results of Test Session 1.3. Blue bars indicate average scores after correcting for species misidentified in the final node. Error bars indicate ± SE.

As shown in Figure 25, the inclusion of the corrected species increased the students' average score by 13% while using the visual key (Mean visual key score: 82.8 ± SE). Likewise, this correction also increased the students' performance while using the Forest Key, but only by 2.2% (Mean Forest Key score: 52.2 ± SE). As a result, there was no change in the significant differences between the two key's performance ($p \ll 0.001$).

41

Test 2: Results of Botanical Field Expert Identification Test

The results of the second test were very similar to the results of the biology students. The botanical experts scored can average of 72.6% correct when using the visual key, while scoring an average of 42.1% correct when using the Forest Service key. These results demonstrate a significant difference ($p \ll 0.001$) between the performance of each key with the visual key scoring an average of 30.5% higher than the Forest Service key (Figure 26).
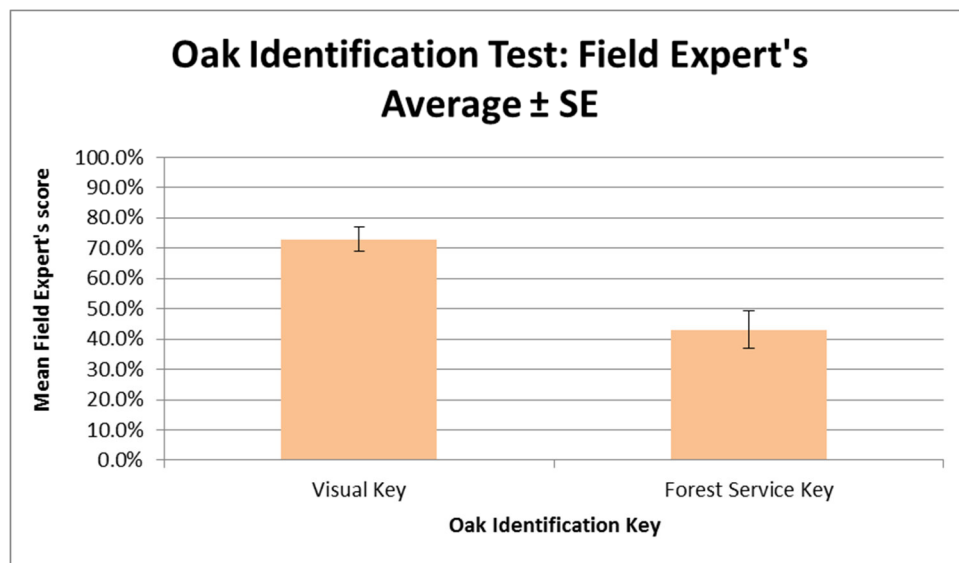
Figure 26. Results of Test Session 2.1. Mean scores of experts at 8th International Oak Conference Error bars indicate ± SE.

As with the data from test session 1, we removed the species that both keys did not share in common and recalculated the differences between the new score averages (Figure 27). The results showed an 8% increase in the average score for the Forest Service key and only a 0.4% increase in the scores for the visual key. However, the differences between the two keys remained significantly different ($p < <0.01$).
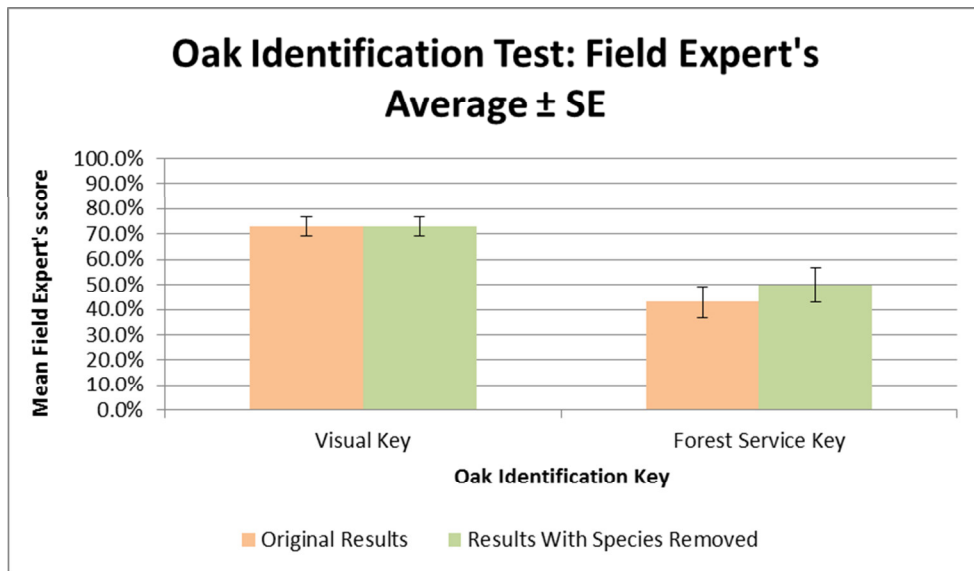
Figure 27. Results of Test Session 2.2. Green bars show mean scores of experts at 8th annual International Oak Society Conference with species removed that did not occur in both keys (Table 4). Error bars indicate ± SE.

The test answers for the field experts were also corrected for species that were misidentified at the same node (Figure 28). The results showed only a 1% increase for the Forest Service key and a 9% increase for the visual key (Mean visual key score: 82 ± SE; Mean Forest Service key score: 44 ± SE). As such, there was an even greater significant difference between the performance of the two keys ($p \ll 0.001$).
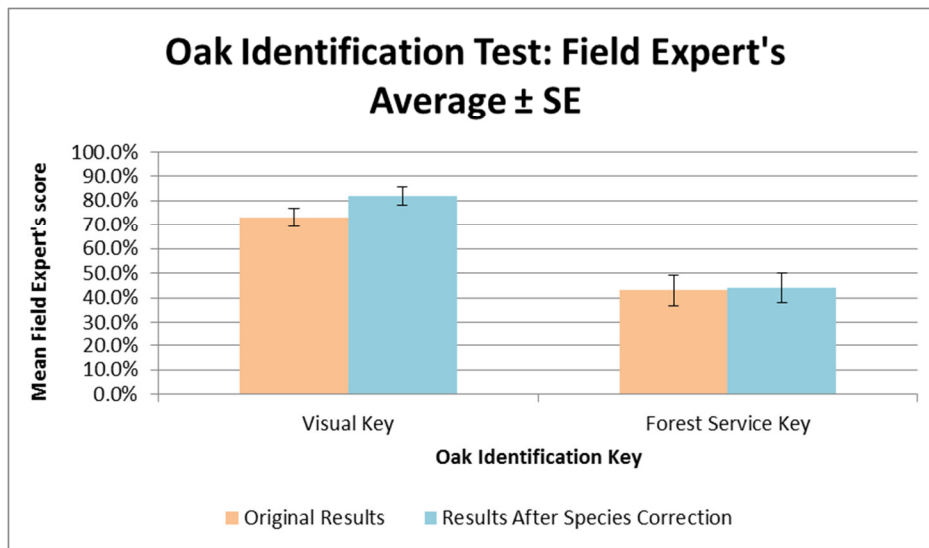
Figure 28. Results of Test Session 2.3. Blue bars indicate average scores after correcting for species misidentified in the final node. Error bars indicate ± SE.

We also investigated the species that were most commonly identified correctly using either key. The results for both students (Table 5a) and botanical experts (Table 5b) are listed in separate tables below. The species that were most often identified correctly are highlighted in red, while the species most often misidentified are highlighted in green.

Table 5. Frequency of Correct Identification for Individual Species. (a) Results from biology student test session. (b) Results from field expert test session. Species that were most commonly correctly identified are highlighted in red. Species that were most often misidentified are highlighted in green.

Test Session 1: Biology Students			
	Frequency of correct identification		
Species	Key A	Key B	Total
Q. bicolor	67.74%	53.33%	60.66%
Q. falcata	67.74%	43.33%	55.74%
Q. hemisphaerica	45.16%	13.33%	29.51%
Q. lyrata	70.97%	23.33%	47.54%
Q. mcrocaarpa	70.97%	50.00%	60.66%
Q. montana	77.42%	43.33%	60.66%
Q. muehlenbergii	83.87%	33.33%	59.02%
Q. phellos	61.29%	80.00%	70.49%
Q. rubra	83.87%	70.00%	77.05%
Q. stellata	58.06%	63.33%	60.66%
Total of paricipants using each key:	31	30	61

(a)

Test Session 2: Botanical Field Experts			
	Frequency of correct identification		
Species	Key A	Key B	Total
Q. bicolor	80.00%	30.00%	55.00%
Q. falcata	80.00%	60.00%	70.00%
Q. hemisphaerica	40.00%	30.00%	35.00%
Q. lyrata	50.00%	0.00%	25.00%
Q. mcrocaarpa	90.00%	50.00%	70.00%
Q. montana	50.00%	10.00%	30.00%
Q. muehlenbergii	100.00%	60.00%	80.00%
Q. phellos	90.00%	80.00%	80.00%
Q. rubra	100.00%	70.00%	85.00%
Q. stellata	50.00%	40.00%	45.00%
Total of paricipants using each key:	10	10	20

(b)

Both expert and non-expert tests subjects were given a survey to rate how easy the two

keys were to use. The results of these surveys are listed in the table below (Table 6).

Table 6. Results of the Test Participant Survey. (a) Results from biology students at UNCG. (b)
Results from experts at 8th International Oak Society conference.

Student Rating	Key A	Key B
Difficult	0%	13.33%
Somewhat difficult	22.58%	50%
Somewhat easy	67.74%	33.33%
Easy	9.68%	3.33% (a)

Expert Rating	Key A	Key B
Difficult	5.00%	15.00%
Somewhat difficult	25.00%	55.00%
Somewhat easy	10.00%	15.00%
Easy	60.00%	15.00% (b)

In both the student and expert surveys, the visual key was rated higher in its usability.

However, the field experts rated the visual key as "Easy" more often than the students, who most

often rated the visual key as only "Somewhat easy". Both user groups most often rated the Forest

Service key as "Somewhat difficult" to use. Again, a higher percentage of students found the

Forest Service key easier to user than the field experts.

CHAPTER IV

DISCUSSION

General Discussion of the Test Results

Throughout all tests, both user groups correctly identified more species while using the visual key opposed to the Forest Service key. This suggests that the visual key was just as useful for the expert audience as it was for the student. Surprisingly, there was also little difference in the overall performance between both user groups. While it was anticipated that the experts would score higher on average, the results showed that there was no significant difference between the performance of the experts and students ($p<0.01$; Visual key = 0.5481; Forest Service key = 0.5566). Students using the visual key scored and average of 69.8, while the experts scored and average of 73. Similarly, students using the Forest Service key scored an average of 43, with experts scoring 46.9, only 3.9% higher. These results may be attributed to a difference in the testing format and sample size between the two groups. While the students received only one key to identify specimens, the experts used both keys in a within-subjects test. This design, however, is particularly strong against the effects of small sample size. Effectively, this design doubles the sampling size of the test and reduces the possibility of error in regards to the individual variance of each test subject. However, this design is also weak against carry-over effects, where the results of one test may have a direct effect on the test subject's performance on the following task. While this may have had some affect our results, there was an effort to change the order that the tests were taken in per participant. Therefore, it is unlikely that the difference between the performances of the two keys is due to a carry-over effect. The differences

between the two identification tests also remained significant even when accounting for elements of dissimilarity between the two keys (Visual key p = 4.017-09; Forest Service key p = 8.584-06). The Forest Service key, for example, had seven more species than the visual key – leaving room for a greater possibility of error. We controlled for these differences by removing the species that the two keys did not have in common. Theoretically, this should have given the users of the Forest Service key a greater chance to achieve a correct answer. This alteration had little effect on the scores of the non-experts, and a slight effect on the scores of the experts (+ 7%). However, the average scores of the visual key and Forest Service key remained significantly different (p<0.01), demonstrating that the presence of additional species in the Forest Service key had no significant influence on the user's test score.

Likewise, the addition of the extra characters in the visual key had no significant influence on the participant's performance. While the Forest Service Key and the visual key focused primarily on leaf characters, there were some additional characters provided in the visual key. At the final node, where the user was presented with possible identifications for their species, the visual key provided the user with pictures of the species' bark and acorns. While the first character was inconsequential, since it was not provided by the voucher, an acorn was present in most of the vouchers. This presented a possible advantage for the visual key. To correct for this advantage, we recalculated the answers for each user and counted the identification correct if the user identified a species within the same final group as the correct answer. As a result, there was an increase in the average score of both keys. The effect was larger on the visual key – increasing the expert's score by 10% and the non-expert's scores by 13%. The effect was smaller for the experts using the Forest Service key (+1%), and for the students (+5.3%). This illustrates that those using the visual key were more likely to reach a correct answer than those using the Forest Service key.

The strong effect that these corrections had on the means of the visual key (+ 13%, +10%) may be attributed to some species appearing in two different places within the key itself. This was a strategy used in the visual key's creation to represent the differing perspectives of experts and non-experts. In fact, two of the vouchers tested appeared in two separate branches of the visual key: *Q. lyrata* and *Q. hemisphaerica* (Figure 14). This presents an advantage to users with the visual key –increasing their chances of keying out these species correctly. However, the Forest Service key also had alternative routes of identification for two test species: *Q. rubra* and *Q. lyrata*. This nullifies the possibility of the success of the visual key was attributed to alternative paths of species identification.

There remains one more difference between the two keys that could potentially account for the success of the visual key. Between the two keys there exists some differences between the number of species a user could view at one time. Unlike the dichotomous Forest Service key, the visual key possessed some polychotomous clades – containing more than two character states. This could present a possible advantage to identification since the user is given multiple characters states to choose from, rather than two. In effect, this also limits the number of steps it takes to identify a species. Given this information, the visual key may have out-performed the Forest Service key based on this advantage. However, the use of ploychotomous branches is not a new innovation. Future tests could control for this difference by testing the visual key against a text-based key that also utilizes polychotomous branches.

The two keys were also rated differently by the test participants. Overall, both test groups rated the visual key as easier to user than the Forest Service Key. However, the experts were more confident in asserting the key as easy, while the students were less confident. These results are most likely due to the students' inexperience with keys in general.

In conclusion, there is significant evidence to suggest that the visual key improves species identification, regardless of previous experience with a using identification keys. The majority of users also preferred using the visual key to the Forest Service key. While further testing is needed, the overall evidence suggests that the visual key is a more effective identification key.

Further Testing

Comments made by students about the strengths of the visual key indicate that the students found the images helpful. When asked to describe the strengths of the key, one student remarked "[h]aving an actual image, as opposed to a description can minimize misunderstandings with descriptive wording." Similarly, others described the key as having "easy navigation" with "good pictures", which they thought might be "good for people who have little to no knowledge with identification". What is most interesting, however, it the comments made by the students using the Forest Service key. This key possessed some, though limited, illustrations. Exactly half of students cited these illustrations as one of the key's most helpful features. In addition, some of the students recognized the need to show variety in the leaves they were asked to identify. One student remarked, "There should be a variety of pictures to choose from or look at". Comments such as this help affirm the users' preference for visual characters and the need to show variety using multiple photographs. However, it does not determine the effect that the photographs had on the user's performance.

While the results of the *t*-test demonstrates a strong performance using the visual key, the question remains whether it was the images alone that enhanced the user's performance or whether it was the way in which the key's map was constructed from potential user data. Could

50

alternative methods of creating a dendrogram yield similar results? Further research should be conducted to investigate the alternative possibilities.

One possible test could investigate the differences in clustering methods used to construct the dendrogram from the visual similarity data. Such a study could investigate if alternative clustering methods are just as effective for constructing a key model. Alternatively, the similarity matrices themselves could be created without human-subject data and constructed using computer algorithms of artificial intelligence programs. Future research could test our oak key against an oak key created from the artificial intelligence data matrix. The results of this artificial data matrix itself could be compared to see if the program places species in clusters similar to the human users.

Overall, the most provocative question is whether or not the dendrogram, created from human subject data, was a significant factor in enhancing user performance. Perhaps the images alone could be inserted into any oak key and the effect would remain the same. In order to test this hypothesis, a key could be constructed using our own expert knowledge, but still utilizes the same visual characters. This key would in-turn be tested against the original version of our visual key, created from user data. Alternatively, a key could be constructed from an existing dichotomous key that uses our standardized images in the place of written characters.

Since our key was only tested against an illustrated key, it is also worth exploring how the key would perform against other types of identification tools. Keys, such as matrix-based keys allow the user to start with any characters that are available. A key such as this could be tested against the visual key to determine if the addition of multiple starting points for character identification is an advantage over the visual key. Other identification tools, such as Leaf Snap, use visual recognition software to identify leaves (Farnsworth 2013). This popular application is available for download onto mobile devices. Since this is an easily accessible tool that relies on

visual recognition – albeit in a different manner – it would be of particular interest to test the visual key against such a device. Such a test would be especially useful to investigate the how well members of the general public can identify plants using our key.

Improvements to the Visual Key

So far we have discussed additional ways to test the effectiveness of the visual key. However, examining the results of the visual key's performance has also elucidated areas for improvement to the current key's design.

During our analysis, we identified weak points in both the visual and Forest Service keys by looking at the species that were most commonly missed between the two user groups. For both student and expert users of the visual key, Q. *hemisphaerica* was the species most often misidentified and both *Q. rubra* and *Q. muehlenbergii* were the species most often identified correctly (Table 5, highlighted in red). This consistency indicates potential points of strength and weakness of the visual key.

In particular, we were interested in whether or not *Q. hemisphaerica* was misidentified because it is a difficult species to recognize, or because there is a weakness in the visual key. To consider this further, we examined the results of the species most often misidentified by the users of the Forest Service Key. For students using the Forest Service key, the species that was most often misidentified was also *Q. hemisphaerica* (Table 5, highlighted in green). This is consistent of the results of the users with the visual key. While this was not true for the expert users, *Q. hemispaherica* was only correctly identified 30% of the time by botanical experts using the Forest Service key (Table 5). Such results may indicate that *Q. hemisphaerica* is indeed a difficult species to identify.

Given this possibility, improvements the visual key should begin by investigating what *Q. hemisphaerica* was most often misidentified as. Looking at the responses of the students using the visual key indicate *Q. oglethorpensis*. The node in the visual key that contained both of these species gave extra photographs that were meant to elucidate some of the differences between these species. In particular, a photograph was provided, contrasting the backside of both leaves. However, there is no additional instruction to guide users to this conclusion. Users who were unfamiliar with leaf pubescence, like the students, may not have been able to determine the significance of comparing these two photographs. In this instance, it may be prudent to add brief text or pointers to draw the user's attention to such structures. Looking at students' comments about the strengths of the two keys also emphasizes this point. A common comment made regarding the strengths of the Forest Service key was that it provided clear instructional text on characters, despite the student's inexperience with them. Conversely, students using the visual key cited the main weakness of the key was their inability to decide between species that were highly similar. As such, students may benefit from some concise instruction on what the structures being compared in the opposing photographs. One student also made the suggestion of adding a "zoom feature" to enable the user to look at more intricate details of the leaf. This addition, could help provide the specific character detail that some students were missing from the photographs.

In addition to improving the presentation of characters, our key could also be expanded to include characters for winter identification. In its current form, a user would be unable to use our visual key to identify oaks in winter. In this situation, some authors have created keys exclusively for winter identification, using characters of plant structures that are available during the winter months. In his key to Woody Plants of the Southeastern United States (2004), Ron Lance uses a combination of bud, bark, and acorn characters to identify trees during the winter.

53

Since our study included photographing these structures, it is reasonable to use them for creating a winter key. Data on winter structures, such as acorns, could be collected using the same visual assessment test we administered to leaves. Using this data, a dendrogram could be constructed to act as a map for a key to acorns.

Challenges and Future Applications of these Methods

In our study, we demonstrated that keys can be constructed from crowd-sourcing information on how different users discriminate characters, such as leaves. Through the use of survey engines like Mechanical Turk, non-expert data can be gathered relatively quickly on how users differentiate species. However, there are limitations of this method of gathering information. For instance, recruiting from a single website limits the sampling pool of participants to those who are members of that site. In our study we sampled using Amazon Mechanical Turk as our exclusive method of recruitment. While future studies could recruit members from a variety of sources, there is an inherent sampling bias when members are only recruited from sources where participants take surveys often. In fact, since we required a level of mastery for our test participants, we placed further restrictions on our participants by limiting them to those who are proficient in certain types of tasks. Future research could avoid these examples of bias by using a variety of different recruitment methods.

An even greater challenge of this method of key creation is the difficulty of sampling large amounts of visual assessment data in a short amount of time. While the non-expert data were obtained within twenty-four hours, the expert data was obtained over a period of five months. The tediousness of this process would certainly slow the collection process of visual assessment data for larger taxa. In order to create keys to plant families, for example, data would have to be gathered on how users assess similarities within and between species groups. For

large taxa, this would be a challenge to implement and would require a great deal of time. To ameliorate this, researchers could break species down into smaller groups to be tested. For example, in the visual assessment of oaks, both users groups partitioned out species groups into two major categories: lobed and unlobed. While some species remained ambiguous, such as the chestnut oaks, it was clear that some species groups were seen as similar by both experts and non-experts. Therefore, it may be beneficial to limit the number of taxa that users are asked to assess by preselecting species that are often mistaken for one another.

While the methodology of our visual key requires further refinement, there exists the exciting possibility of expanding our techniques to create keys for other species. Using data on visual assessment, improvements could also be made to identify other plant taxa. Since our key as created for a single genus, the next logical progression of developing visual plant keys would be the creation and testing of a key to identify more than one genus. As we have previously discussed, there is a concern for creating a key that incorporates a large number of taxa. Since our key was successful in enhancing user performance with a large and difficult large genus, it seems reasonable that a key that includes groups of smaller taxa would be achievable.

In addition, our methods could also be used to explore identification of other organisms such as fungi, birds, and insects. Fields like horticulture would benefit greatly from enabling users to quickly identify pathogens and disease using of visual characters.

Furthermore, using data on how users differentiate species could be investigated for non-visual characters, such as audio. Authors creating ornithology keys, for example, could benefit from investigating how different users interpret songs. While it is unlikely that such data could result in a dendrogram for a key, the data itself could prove useful in how experts present users with auditory characters.

Our research has shown evidence that the use of visual characters in identification keys has the potential to increase the likelihood for a correct identification. To our knowledge, this is the first key created using visual characters exclusively, and the first key created using visual assessment data from potential users. Although it is unclear if it is one, or a combination of these two factors, there is certainly room for further exploration of both methods. Only through exploring new methodology can improvements be made to identification tools, and subsequently to the various fields that rely on them.

REFERENCES

Ang, Y. L.J. Wong, and R. Meier. (2013). Using seemingly unnecessary illustrations to improve the diagnostic usefulness of descriptions in taxonomy–a case study on *Perochaeta orientalis* (*Diptera*, *Sepsidae*). *ZooKeys* 355: 9–27. doi: 10.3897/ zookeys.355.6013

Bacher, S., (2012) Still not enough taxonomists: reply to Joppa et al. Trends Ecol. Evol. 27, 65–66.

Baskauf S.J. and B.K. Kirchoff (2008). Digital plant images as specimens: Standards for photographing living plants. Vulpia 7, 16-30.

Blackmore, S. (2002) Biodiversity Update – Progress in taxonomy. *Science's Compass* 298: 365.

Boero, F. (2001) Light after dark: the partnership of enhancing expertise in taxonomy. *Trends in Ecology and Evolution* 16: 266.

Dallwitz, M.J., (1992) A comparison of matrix-based taxonomic identification systems with rule-based systems. In 'Proceedings of IFAC Workshop on Expert Systems in Agriculture', pp. 215-218.

Dallwitz, M.J. (1999 onwards). A comparison of formats for descriptive data. http://delta-intkey.com

Dressler, R.L. (1993) Field Guide to the Orchids of Costa Rica and Panama, Ithaca, NY: Comtock Publishing Associates.

Drinkwater, R. E. (2009) Insights into the development of online plant identification keys based on literature review; an exemplar electronic key to Australian Drosera. Bioscience Horizons 2: 90-96.

Edwards, M. and D.R. Morse (1995) The potential for computer-aided identification in biodiversity research. *Tree* 10: 153-158.

eFloras.org. Published on the Internet http://www.efloras.org [accessed 15 March 2014] Missouri Botanical Garden, St. Louis, MO & Harvard University Herbaria, Cambridge, MA.

Farnsworth, E.J., M. Chu, W.J. Kress, A. K. Neill, J.H. Best, J. Pickering, R.D. Stevenson, G.W. Courtney, J.K. VanDyk, and A.M. Ellison. (2013) Next Generation Field Guides. *Bioscience* 63: 891-899.

Gleason H.A., A. Cronquist (1963). Manual of Vascular Plants of Northeastern United States and Adjacent Canada, NY: D. Van Nostrand Company.

Godfray, H.C.J. (2002) Challenges for taxonomy. *Nature* 417: 17-19.

Granjou, C., I. Mauz., M. Barbier, P. Breucker. (2014) Making taxonomy environmentally relevant. Insights from All Taxa Biodiversity Inventory. Environmental Science and Policy 38: 354-362.

Harris, J.G. and M.W. Harris (1984). Plant Identification Terminology: An illustrated Glossary, Spring Lake, UT: Spring Lake Publishing.

Hawthorne, W.D., S. Cable, and C.A.M. Marshall (2014) Empirical Trials of Plant Field Guides. *Conservation Biology* 28:3 654-662.

Joppa, L.N., D.L. Roberts, and S. Pimm. (2011). The population ecology and social behaviour of taxonomists. *Trends Ecol. Evol*. 26, 551–553.

Kirchoff, B.K. (2001). Character description in phylogenetic analysis insights from Agnes Arber's Concept of the Plant. *Annals of Botany* 88: 1203-1214.

Kirchoff, B. K., D. L. Remington, L. Fu, F. Sadri. (2008) A New Type of Image-Based Key. Proceedings of the International Conference on BioMedical Engineering and Informatics, 2008. BMEI2008. 1: 825-829. DOI 10.1109/BMEI.2008.301

KompoZer: http://www.kompozer.net/download.php [accessed 12 February 2015].

Lance, R. (2004). Woody Plants of the Southeastern United States: A Winter Guide, Athens, GA: University of Georgia Press.

Leggett, R., and B. K. Kirchoff. 2011. Image use in field guides and identification keys: Review and recommendations. AoB Plants 2011 doi: *10.1093/aobpla/plr004*

Lobanov A.L. (2003) *Keys to beetles and biological diagnostics*. http://www.zin.ru/Animalia/Coleoptera/eng/syst8.htm. [accessed 10 June 2015].

Mesquite Project Team. 2015. Documentation for Mesquite: a modular system for evolutionary analysis. https://mesquiteproject.wikispaces.com [accessed 12 September 2015]

Murphy G.L. and D.L. Mendin (1985) The role of theories in conceptual coherence. *Psychological Review* 92: 289-316.

Pearson, D.L., A.L. Hamilton, and T.L. Erwin (2011) Recovery plan for the endangered taxonomy profession. Bioscience, 61, 58.

Pertot, I., T. Kuflik, I. Gordon, S. Freeman, and Y. Elad. (2012) Identificator: a web-based tool for visual plant disease identification, a proof of concept with a case study on strawberry. *Computers and Electronics in Agriculture* 84: 144-154.

Rehder, A. (2001). Manual of Cultivated Trees and Shrubs, 2nd ed. Caldwell, NJ: The Blackburn Press.

Sibley, D.A. (2000). The Sibley Guide to Birds, New York, NY: Alfred A. Knopf, Inc.

Sluys, R. (2013) The underappreciated, fundamentally analytical nature of taxonomy and the implications for biodiversity. Biodiversity and Conservation 22: 1095-105. DOI 10.1007/s10531-013-0472-x.Sneath, P.H.A., and R. R. Sokal. (1973). Numerical Taxonomy: the Principles and Practice of Numerical Classification, San Francisco, CA: W. H. Freeman and Company.

Stein, J., D. Binion, R. Acciavattii. (2003). Field Guide to Native Oak Species of Eastern North America, Morgantown, WV: United States Forest Service.

Stevenson, R.D., W.A. Haber, R.A. Morris. (2003) Electronic Field Guides and User Communities in the Eco-informatics Revolution. *Conservation Ecology*.

Stucky, J.M. (2003). Winter twig keys to common, native, fully deciduous trees and phanerophyte shrubs of the North Carolina eastern peidmmont. Vulpia 2: 23-42.

Tancoigne,E., and A. Dubois. (2013) Taxonomy: no decline, but inertia Cladistics 29: 567-570.

Terlizzi, A., S. Bevilacqua, S. Fraschetti, and F. Boero. (2003) Taxonomic sufficiency and the increasing insufficiency of taxonomic expertise. *Marine Pollution Bulletin* 46: 556-561.

Tilling, S. (1984) Keys to biological identification: their role and construction. *Journal of Biological Education*. 18: 293-304.

USDA: National Resources Conservation Services. (http://plants.usda.gov/plantkeys/north_carolina_grasses/NORTH_CAROLINA_GRASSES.html) [accessed 22 January 2014]

Walter, D.E., and S. Winterton. (2007) Keys and the crisis in taxonomy; extinction or reinvention? *Annual Review of Entomology* 52: 193-208.

Weakley A.S. (2015). Flora of the southern and mid-atlantic states. Chapel Hill, NC: UNC Chapel Hill.

Wheeler, Q., P.H. Raven, and E.O. Wilson. (2004) Taxonomy; Impediment or Expedient? *Science* 303: 285.

Wheeler, Q. (2014) Are reports of the death of taxonomy exaggeration? New Phytologist 201: 370-371.

Wisniewski, E.J. (2002). Concepts and categorization. In: Stevens' Handbook of Experimental Psychology, vol 2., Memory and Cognitive Processes, ed. D Medin, New York: Wiley, 467–531.

APPENDIX A

LIST OF SPECIES VOUCHERS

Species name	Study Collection Number	UNC accession number	Location species was collected/GPS coordinates	Date species(s) was collected
Q. acutissima	R5	634226	Foust Park, UNCG campus, Greensboro, Guilford Co., NC; N 36°03.998' W 079°48.416'	2/28/2014; 5/30/2014; 7/31/2014
Q. acutissima	R6		Spring Garden St & Forest St, UNCG campus, Greensboro, Guilford Co., NC N 36°03.947' W 079°48.614'	2/24/2014
Q. alba	R19	634217	Spring Garden St & Forest St, UNCG campus, Greensboro, Guilford Co., NC; N 36°04.005 W 079°48.628	3/11/2014; 7/31/2014
Q. arkansana	R42	634235	Park Ave, Aiken, Aiken Co., SC; N 33°33.252' W 081°42.391'	4/5/2014; 6/3/2020
Q. austrina	R37	634244	Beaufort & Coleton, Aiken, Aiken Co., SC N 33°33.106' W 081°42.354'	4/5/2014; 6/3/2015
Q. austrina	R74	634305	Fairfield St. and Abbeville Ave., Aiken, Aiken Co., SC; N 33°33.951 W 081°42.823	8/2/2014
Q. austrina	R134		New Bridge Bank, Highwoods Park, New Garden Rd. Greensboro, Guilford Co., NC; N 36°06.540 W 079°52.898	12/1/2014
Q. austrina	R146		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°04.940 W 079°47.931	12/28/2014
Q. bicolor	R37	634228	Foust Park UNCG campus, NC; N36°04.005' W 079°48.391'	2/28/2014; 7/31/14
Q. boyntonii	R40	634241	Parke Ave., Aiken, Aiken Co., SC N 33°33.077' W 081°41.868'	4/5/2014; 6/3/2018

Q. bicolor	R50		Park Ave., Aiken, Aiken Co., SC; N 33°33.195' W 081°42.225'	4/5/2014; 6/3/2028
Q. chapmanii	R23	634215	Beaufort St., Aiken, Aiken Co., SC N 33°33.410' W 081°42.210'	4/5/2014; 6/3/2014
Q. champanii	R100	634281	Sand Rd S3, John's Island, Tomoka State Park, Ormond Beach, Volusia Co., FL; N 29°18.414 W 081°04.850	8/19/2014
Q. chapmanii	R102	634279	Sand Rd S3, John's Island, Tomoka State Park, Ormond Beach, Volusia Co., FL; N 29°18.370 W 081°04.825	8/19/2014
Q. chapmanii	R103	634278	Sand Rd S3, John's Island, Tomoka State Park, Ormond Beach, Volusia Co., FL; N 29°18.356 W 081°04.820	8/19/2014
Q. coccinea	R30		Angel Rd., Mars Hill, Madison Co., NC; N 35°49.145 W 082°29.034	10/31/2014
Q. coccinea	R63	634265	Laurel Dormitory,Mars Hill University campus, Mars Hill, Madison Co. NC; N 35°49.202 W 082°32.915	7/24/2014
Q. coccinea	R148		Beside Kathleen Clay Edwards Library, Greensboro, Guilford Co., NC; N 36°06.314 W 079°52.682	12/28/2014
Q. falcata	R18	634240	Spring Garden St. & Forest St. near park, UNCG campus, Greensboro, Guilford Co., NC; N 36°03.984 W 079°48.592	3/10/2014; 7/31/2015
Q. falcata	R59	634315	Beaufort St., Aiken, Aiken Co., SC; N 33°33.121.' W 081°42.343	6/3/2014
Q. falcata	R132	634254	Near Mossman building on Forest St., UNCG Campus,; Greensboro, Guilford Co., NC; N 36°04.007 W 079°48.620	7/31/2014
Q. geminata	R48	634267	Sundy Ave., Aiken, Aiken Co., SC N 33°33.902' W 081°42.009'	4/5/2014; 6/3/2026

Q. geminata	R96	634283	Indian Mound Station Sanctuary, Mims, Brevard Co., FL; N 28°39.230 W 080°51.591	8/17/2014
Q. geminata	R97		Indian Mound Station Sanctuary, Mims, Brevard Co., FL; N 28°39.224 W 080°51.609	8/17/2014
Q. geminata	R108	634275	Junction of Steed Creek Rd. and Halfway Creek Rd., Francis Marion National Forest, Huger, Berkeley Co, SC; N 33°03.662 W 079°41.444	9/5/2014
Q. geminata	R139		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°05.014 W 079°47.801	12/28/2014
Q. georgiana	R45	634232	Park Ave., Aiken, Aiken Co., SC; N 33°33.213' W 081°42.285'	4/5/2014; 6/3/2023
Q. georgiana	R142		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°04.976 W 079°47.889	12/28/2014
Q. hemisphaerica	R20	634216	Jackson Library, UNCG campus, Greensboro, Guilford Co., NC; N 36°04.119 W 079°48.545	3/11/2014; 2/24/2014; 7/24/2014
Q. hemisphaerica	R121	634310	Hitchcock Woods, Aiken, Aiken Co., SC; N 33°33.286 W 081°45.152	10/12/2014
Q. hemisphaerica	R149		Between Hobbs rd. and Friendly Shopping Center, Greensboro, Guilford Co., NC; N 36°05.235 W 079°50.447	1/2/2015
Q. ilicifolia	R43	634234	Park Ave., Aiken, Aiken Co., SC; N 33°33.245' W 081°42.372'	4/5/2014; 6/3/2021
Q. ilicifolia	R89	634255	Crowder's Mountain Summit, Gaston Co., NC; N 35°14.053 W 081°16.452	8/10/2014
Q. ilicifolia	R91	634253	Crowder's Mountain Summit, Gaston Co., NC; N 35°14.084 W 081°16.429	8/10/2014

Q. imbricaria	R29	634237	Behind Margaret C. Moore School of Nursing Building, UNCG campus; N 36°04.171 W 079°48.439	8/4/2014
Q. imbricaria	R85	634256	Cornwallis Drive, Greensboro, Guilford Co., NC; N 36°05.748' W 079°49.640'	8/8/2014
Q. incana	R24		Hitchcock Woods, Aiken, Aiken Co., SC N 33°33.216' W 081°45.086'	4/5/2014
Q. incana	R54	634288	Ridge Mile Track, Hitchcock Woods, Aiken, Aiken Co., SC; N 33°33.234' W 081°45.201'	4/5/2014 & 6/3/2014
Q. incana	R61	634270	Hitchcock Woods, Aiken Co., SC; N 33°33.257' W 081°45.187'	6/3/2014
Q. incana	R123	634311	Hitchcock Woods, Aiken, Aiken Co., SC; N 33°33.250 W 081°45.197	10/12/2014
Q. incana	R138		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°04.960 W 079°47.834	12/28/2014
Q. inopina	R 39	634243	Park Ave., Aiken, Aiken Co., SC N 33°33.108' W 081°41.961'	4/5/2014; 6/3/2017
Q. inopina	R131		Lake Marion Creek Rd., Haines City, Polk Co., FL; N 28°07.083 W 081°32.967	11/8/2014
Q. laevis	R44	634233	Florence St., between Hamptonville Ave. and Abbeville Ave., Aiken, Aiken Co.,SC; N 33°34.225' W 081°43.380'	4/5/2014; 6/3/2022
Q. laevis	R77	634303	Hitchcock Woods, Rabbit Valley Trail, Aiken, Aiken Co., SC; N 33°33.281 W 081°45.123	8/3/2014
Q. laevis	R104	634277	Halfway Creek Rd., Francis Marion National Forest, Huger, Berkeley Co., SC; N 33°03.806 W 079°41.348	9/5/2014

Q. laevis	R137		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°04.970 W 079°47.846	12/28/2014; 6/2015
Q. laurifolia	R93	634251	Swamp Deer Rd., New Smyrna Beach, Volusia Co., FL; N 28°57.165 W 081°03.978	8/17/2014
Q. laurifolia	R157		Corner of Cridland Ave. and West Wendover, Greensboro, Guilford Co., NC; N 36°05.305 W 079°47.638	7/15/2015
Q. lyrata	R14	634220	Spring Garden St & Forest St, UNCG campus, Greensboro, Guilford Co., NC; N 36°03.980 W 079°48.577	2/24/2014
Q. lyrata	R56	634287	Friendly Center, Greensboro, Guilford Co., NC; N 36°05.266' W 079°50.295	5/30/2014
Q. lyrata	R57	634286	Friendly Center, Greensboro, Guilford Co., NC; N 36°05.296' W 079°50.340	5/30/2014
Q. macrocarpa	R2	634229	Cornwallis Drive, Greensboro, NC; N 36*05.748' W 079*49.649'	2/24/2014; 5/30/2014
Q. macrocarpa	R51	634290	Park Ave., Aiken, Aiken Co., SC; N 33°33.133' W 081°42.031'	4/5/2014; 6/3/2014
Q. macrocarpa	R147		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°04.893 W 079°47.708	12/28/2014
Q. margarettae	R25		Fairfield St. and Edgefield St., Aiken, Aiken Co., SC N 33°33.807' W 081°42.865'	4/5/2014
Q. margarettae	R62	634266	Ridge Mile Track, Hitchcock Woods, Aiken Co., SC; N 33°33.238' W 081°45.194'	6/3/2014
Q. margarettae	R73	634306	Fairfield St. and Abbeville Ave., Aiken, Aiken Co., SC; N 33°33.907 W 081°42.837	8/2/2014

Q. margarettae	R111	634298	Service road off of Cainhoy Rd., Francis Marion National Forest, Huger, Berkeley Co., SC; N 32°57.167 W 079°50.956	9/5/2014
Q. margarettae	R112	634299	Service road off of Cainhoy Rd., Francis Marion National Forest, Huger, Berkeley Co., SC; N 32°57.164 W 079°50.968	9/5/2014
Q. margarettae	R120	634312	Hitchcock Woods, Aiken, Aiken Co., SC; N 33°33.286 W 081°45.135	10/12/2014
Q.marilandica	R27	634239	Florence St., between Hamptonville Ave. and Abbeville Ave., Aiken, Aiken Co., SC; N 33°34.179' W 081°43.406'	4/5/2014 & 8/3/2014
Q. marilandica	R53		Florence St., between Hamptonville Ave. and Abbeville Ave., Aiken, Aiken Co., SC; N 33°33.332' W 081°44.904'	4/5/2014
Q. marilandica	R124	634308	Hill Crest Rd., Aiken, Aiken Co., SC; N 33°33.576 W 081°44.303	10/12/2014
Q.michauxii	R4	634227	Foust Park, UNCG campus, NC; N 36*03.990' W 079*48.401'	2/28/2014; 7/31/14
Q. michauxii	R75	634304	Langdon Rd., Aiken, Aiken Co., SC; N 33°34.331 W 081°33.249	8/2/2014
Q. michauxii	R92	634252	Cornwallis Drive, Greensboro, Guilford Co., NC; N 36°05.752' W 079°49.641'	8/15/2014
Q. michauxii	R126		Langdon Rd., Aiken, SC; N 33°34.342 W 081°33.335	10/12/2014
Q. michauxii	R150		Park inbetween W. Market and Greenway Dr., Greensboro, Guilford Co., NC; N 36°04.434 W 079°49.483	1/2/2015

Q. minima	R105	634272	Junction of Steed Creek Rd. and Halfway Creek Rd., Francis Marion National Forest, Huger, Berkeley Co., SC; N 33°03.791 W 079°41.342	9/5/2014
Q. minima	R109	634274	Francis Marion National Forest, Huger, Berkeley Co., SC; N 33°03.764 W 079°41.607	9/5/2014
Q. minima	R110	634273	Junction of Steed Creek Rd. and Halfway Creek Rd., Francis Marion National Forest, Huger, Berkeley Co., SC; N 33°03.801 W 079°41.596	9/5/2014
Q. minima	R115	634295	Junction of Steed Creek Rd. and Halfway Creek Rd., Francis Marion National Forest, Huger, Berkeley Co., SC; N 33°03.760 W 079°41.567	9/5/2014
Q. montana	R7		Pleasant Ridge Church, Greensboro, NC; N 36°08.938' W 079°56.399'	3/4/2014; 7/31/2014
Q. montana	R9	634225	Pleasant Ridge Church, Greensboro, NC; N 36°08.926' W 079°56.380'	3/4/2014; 7/31/2014
Q. montana	R65	634263	Pleasant Ridge Church, Greensboro, Guilford Co., NC; N 36°08.977' W 079°56.391'	7/31/2014
Q. muehlenbergii	R60	634314	Park Ave., Aiken, Aiken Co., SC; N 33°33.180' W 081°42.198'	6/3/2014
Q. muehlenbergii	R130		Old Fannings Field Rd., Mills River, Henderson Co., NC; N 35°24.915 W 082°33.208	10/30/2014
Q. muehlenbergii	R154		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°0.877 W 079°47.869	2/16/2015; 6/2015
Q. myrtifolia	R34	634247	Park Ave., Aiken, Aiken Co., SC N 33°33.097' W 081°41.929'	4/5/2014; 6/3/2014
Q. myrtifolia	R94	634285	Indian Mound Station Sanctuary, Mims, Brevard Co., FL; N 28°39.197 W 080°51.544	8/17/2014

Q. myrtifolia	R95	634284	Indian Mound Station Sanctuary, Mims, Brevard Co., FL; N 28°39.174 W 080°51.530	8/17/2014
Q. myrtifolia	R98	634282	Indian Mound Station Sanctuary, Mims, Brevard Co., FL; N 28°39.283 W 080°51.659	8/17/2014
Q. myrtifolia	R101	634280	Sand Rd S3, John's Island, Tomoka State Park, Ormond Beach, Volusia Co., FL; N 29°18.411 W 081°04.845	8/19/2014
Q. nigra	R10	634224	Foust Park UNCG campus, Greensboro, Guilford Co., NC N 36°04.009' W 079°48.461'	3/10/2014; 7/31/2014
Q. nigra	R67	634261	Foust Park, UNCG campus, Greensboro, Guilford Co., NC; N 36°03.997 W 079°48.477	7/31/2014
Q. nigra	R122	634309	Hitchcock Woods, Aiken, Aiken Co., SC; N 33°33.272 W 081°45.151	10/12/2014
Q. oglethorpensis	R38	634242	S. Boundary Ave., between Charleston St. and Berkley St., Aiken, Aiken Co., SC; N 33°32.950' W 081°42.280'	4/5/2014; 6/3/2016
Q. oglethorpensis	R69	634259	Langdon Rd., Aiken, Aiken Co., SC; N 33°34.196 W 081°33.514	8/1/2014
Q. oglethorpensis	R70	634258	Langdon Rd., Aiken, Aiken Co., SC; N 33°34.194 W 081°33.495	8/1/2014
Q. oglethorpensis	R83	634301	610A Service Rd., Sumter National Forest, Edgefield Co., SC; N 33°50.569 W 082°07.906	8/3/2014
Q. oglethorpensis	R143		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°04.882 W 079°47.929	12/28/2014; 6/2015
Q. oglethorpensis	R144		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36° W 079°	12/28/2014
Q. pagoda	R49	634291	Park Ave., Aiken, Aiken Co., SC; N 33°33.199' W 081°42.234'	4/5/2014; 6/3/2027

Q. pagoda	R113	634297	Canyon ln., near Francis Marion National Forest, Huger, Berkeley Co., SC; N 32°53.290 W 079°48.829	9/5/2014
Q. pagoda	R136		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°04.937 W 079°47.794	12/28/2014; /'2015
Q. palustris	R68	634260	In front of Stone Building UNCG, NC; N 36°04.057 W 079°48.502	7/31/2014
Q. phellos	R15		Foust Park, UNCG campus, Greensboro, Guilford Co., NC; N 36°03.983 W 079°48.481	3/10/2014
Q. phellos	R118	634293	Lowe's Foods Parkinglot, Old Oak Ridge Rd., Greensboro, Guilford Co., NC; N 36°07.714 W 079°56.055	9/9/2014
Q. phellos	R128		T. Wingate Andrew's High School, High Point, Guilford Co., NC; N 35°59.062 W 079°59.441	10/28/2014
Q. prinoides	R155		UNC Botanical Gardens, Chapel Hill, Orange County, NC; N 35°53.952 W 079° 02.042	2/12/2015; 6/29/2015
Q. pumila	R107	634276	Junction of Steed Creek Rd. and Halfway Creek Rd., Francis Marion National Forest, Huger, Berkeley Co., SC; N 33°03.834 W 079°41.326	9/5/2014
Q. pumila	R114	634296	Junction of Steed Creek Rd. and Halfway Creek Rd., Francis Marion National Forest, Huger, Berkeley Co., SC; N 33°03.750 W 079°41.593	9/5/2014
Q. pumila	R153		Madison Co., FL	2/4/2015
Q. robur	R33	634248	Park Ave., Aiken, Aiken Co., SC N 33°33.236' W 081°42.358'	4/5/2014
Q. robur	R71	634257	Langdon Rd., Aiken, Aiken Co., SC; N 33°34.155 W 081°33.554	8/1/2014

		634307	Langdon Rd., Aiken, Aiken Co., SC; N 33°34.342 W 081°33.335	
Q. robur	R72			8/2/2014
Q. rubra	R12		Foust Park, UNCG campus, Greensboro, Guilford Co., NC; N 36°04.004 W 079°48.441	3/10/2014
Q. rubra	R84	634300	Oaks at Lewisville Shopping Center, Lewisville, Forsyth Co., NC; N 36°05.637 W 080°25.661	8/8/2014
Q. rubra	R86		Pembroke Rd. & Bryan Blvd., Greensboro, Guilford Co., NC; N 36°05.532 W 079°49.719	8/8/2014
Q. rubra	R156		Beside Kathleen Clay Edwards Library, Greensboro, Guilford Co., NC; N 36°06.375 W 079°52.711	7/14/2015
Q. shumardii	R11	634223	Foust Park, UNCG campus, Greensboro, Guilford Co., NC N 36°04.012 W 079°48.490	3/10/2014
Q. shumardii	R13	634222	Foust Park, UNCG campus, Greensboro, Guilford Co., NC; N 36°03.989 W 079°48.501	3/10/2014; 7/25/2014
Q. shumardii	R64	634264	Foust Park, UNCG campus, Greensboro, Guilford Co., NC; N 36°01.011 W 079°48.411	7/25/2014
Q. shumardii	R151		Cliffside Dr. near Wendover Ave., Greensboro, Guilford Co., NC; N 36°03.900 W 079°50.734	1/2/2015
Q. similis	R41	634236	SC Dept. of Employment, Aiken, Aiken Co., SC N 33°33.173' W 081°41.961'	4/5/2014; 6/3/2019
Q. similis	R116	634294	Cypress Gardens, Monck's Corner, Berkeley Co., SC; N 33°03.281 W 079°57.331	9/6/2014
Q. sinuata (sinuata)	R80	634302	Turkey Rd., Sumter National Forest, Edgefield Co., SC; N 33°49.396 W 082°06.152	8/3/2014

Q. sinuata (sinuata)	R81		Turkey Rd., Sumter National Forest, Edgefield Co., SC; N 33°49.384 W 082°06.162	8/3/2014
Q. sinuata (var. sinuata)	R129		Old Fanning Fields Rd., Mills River, Henderson Co., NC; N 35°24.977 W 082°33.159	10/30/2014
Q. stellata	R17	634218	Spring Garden St & Forest St, UNCG campus, Greensboro, Guilford Co., NC; N 36°03.997 W 079°48.616	3/10/2014; 7/31/2014
Q. texana	R16	634219	Foust Park UNCG campus, Greensboro, Guilford Co., NC; N 36°04.010 W 079°48.409	3/10/2014; 7/25/2014
Q. velutina	R8		Pleasant Ridge Church, Greensboro, NC; N 36°08.929' W 079°56.395'	3/4/2014
Q. velutina	R58	634271	Private Residence, Pleasant Ridge Rd., Greensboro, Guilford Co., NC; N 36°09.199 W 079°55.923	5/30/2014
Q. velutina	R66	634262	Pleasant Ridge Church, Greensboro, Guilford Co., NC; N 36°08.976' W 079°56.387'	7/31/2014
Q. velutina	R152		Green Hill Cemetery, Greensboro, Guilford Co., NC; N 36°04.863 W 079°47.911	1/2/2015; 6/2015
Q. virginiana	R1	634230	Friendly Ave and Market St., Greensboro, NC	2/6/2014; 7/25/2014
Q. virginiana	R22		Mimosa Dr & Market St. Greensboro, NC; N 36°04.479 W 079°48.539	3/11/2014

RECRUITMENT LETTER

Greetings,

I am a graduate student at UNC Greensboro working with Dr. Bruce Kirchoff on a new method of creating visual identification keys. You can learn about Dr. Kirchoff's previous work on this subject by following this link: http://aobpla.oxfordjournals.org/content/2011/plr005. I am extending this work by developing a way of automating key creation by incorporating information on visual similarity garnered from an on-line survey that I have developed.

I am writing because I need your help in order to investigate how botanists assess visual similarity in plant characteristics. My work uses two different surveys – one for leaves and the other for acorns – to access the visual similarities between different species of oak. These surveys will be sent out at different times. This email is a request for your help with the first survey, comparing leaves. Each survey contains 60 questions, each consisting of two standardized pictures of leaves (or acorns) displayed side by side. You are asked to rate the visual similarity of the leaves/acorns with a sliding scale that will appear below the images. I will use the results of these aggregated similarity assessments to create a dendrogram, which will be used as the basis for creating the visual key. Each survey will take approximately 20 minutes to complete. Your responses will be stored under an anonymous ID. Neither Dr. Kirchoff nor I will not have access to your real name or any information about you. If you choose to participate, please click on the following link to go to the **leaf survey**: https://uncg.qualtrics.com/SE/?SID=SV_brUqLHqxOHpcgWF.

The survey site will ask you to confirm your consent to participate before beginning the survey.

If I do not receive sufficient responses to this first request, a reminder will be sent out in approximately one week, followed by a second reminder a few days later, if necessary. After the leaf survey is complete, I will send out a request for your participation in the acorn survey. You may choose to participate in one, both, or neither survey.

Thank you for helping with my work. With your help I plan to create and test a visual key to the oaks of the southeastern US. I expect this key to be more accurate and easier to use for both botanical novices and experts alike. If you have any further questions or want to be updated on the work, you may contact me via email at radellin@uncg.edu. I intend to post the results of my research to the listserv when it is published.

Best wishes,

Rebecca Dellinger-Johnston
radellin@uncg.edu

Dr. Bruce Kirchoff
Professor of Biology at UNCG
kirchoff@uncg.edu

APPENDIX C

EXPERT SIMILARITY MATRIX

Plant Species	Q. acutissima	Q. alba	Q. arkansana	Q. austrina	Q. bicolor	Q. boyntonii	Q. champanii	Q. coccinea	Q. elliottii	Q. falcata
Q. acutissima	0									
Q. alba	5.928571429	0								
Q. arkansana	5.4	5.8181818	0							
Q. austrina	5.545454545	4.0714286	3.826086957	0						
Q. bicolor	5.588235294	3.7647059	5.071428571	4	0					
Q. boyntonii	5.692307692	3.4166667	4.166666667	2.545454545	4.764705882	0				
Q. champanii	4.5	5.7857143	3.857142857	4.454545455	5.090909091	5.066666667	0			
Q. coccinea	6	4.2307692	5.833333333	5.461538462	5.923076923	5.125	6	0		
Q. ellottii	3.833333333	6	5.714285714	5.583333333	5.916666667	5.727272727	4.461538462	6	0	
Q. falcata	5.363636364	5.1578947	5.75	4.444444444	5.571428571	4.555555556	5.916666667	3.583333333	6	0
Q. geminata	3.642857143	6	5.428571429	5.25	5.826086957	5.833333333	3	6	1.2	5.8125
Q. georgiana	5.857142857	2.6666667	6	4.75	5.352941176	4.928571429	5.923076923	2.266666667	6	2.46153846
Q. hemisphaerica	4.769230769	5.7333333	4.166666667	5.166666667	4.736842105	5.333333333	1.4375	5.785714286	2.92857143	5.7
Q. ilicifolia	5.857142857	5.1	4.307692308	2.944444444	4.736842105	3.444444444	4.785714286	4.923076923	5.73333333	4.84615385
Q. imbricaria	3.545454545	5.9090909	4	5.181818182	5.818181818	5.692307692	2	5.923076923	2.25	5.85
Q. incana	4.125	6	3.75	5.066666667	5.8	5.357142857	1.666666667	5.833333333	2.5	6
Q. inopia	5.125	6	4.214285714	4.642857143	5.916666667	5.4375	2.285714286	6	3.52941176	5.93333333
Q. laevis	5.888888889	3.15	6	4.666666667	5.25	5.428571429	5.933333333	2.5625	5.94444444	2.33333333
Q. laurifolia	3.684210526	6	4.461538462	3.230769231	5.411764706	3.583333333	1.631578947	5.818181818	3.07142857	5.41666667
Q. lyrata	6	3.7692308	5.615384615	3	4.714285714	3.8	5.375	4.090909091	5.66666667	3.38461538
Q. macrocarpa	6	2.4285714	5.538461538	3.9375	4.85	3.142857143	5.636363636	4.923076923	6	4.84210526
Q. margaretta	6	2.5454545	5.470588235	3.833333333	5.083333333	2.666666667	5.714285714	4.647058824	5.91666667	4.8125
Q. marlandica	5.8	4.4166667	1.7	3.833333333	5	2.583333333	4.882352941	5.846153846	6	5.31578947
Q. michauxii	5.083333333	5.6470588	4.352941176	5.642857143	1.947368421	5.7	5.210526316	5.636363636	6	5.29411765
Q. minima	4.615384615	5.8461538	5	2.9375	4.111111111	4.333333333	1.3	5.470588235	2.875	5.61111111
Q. montana	4.916666667	5.5714286	4.6875	5.294117647	2.4	5.533333333	4.285714286	5.5	5.8	6
Q. muehlenbergii	3.615384615	5.7142857	4.583333333	4.466666667	2.9375	5.5625	3	5.727272727	5.625	5.41666667
Q. myrtifolia	5.090909091	6	3.529411765	5.6	5.785714286	5.545454545	3.133333333	6	4.31578947	6
Q. nigra	5	5.8333333	3.636363636	2.642857143	4.933333333	2.5	3	5.785714286	4.75	5.63636364
Q. olgethorpensis	3.25	5.75	4.65	5.384615385	5.6	5.357142857	3.307692308	5.933333333	2.4	5.86666667
Q. pagoda	5.923076923	3.8666667	5.933333333	5.5	5.416666667	5.357142857	6	1.461538462	6	3.19047619
Q. palustris	5.764705882	4.3181818	5.833333333	5.388888889	5.75	4.916666667	5.818181818	1.133333333	5.88235294	2.4
Q. phellos	2.428571429	5.9090909	5.454545455	5.307692308	6	5.3125	3.583333333	5.8125	1.26666667	6
Q. prinoides	4.117647059	5.3888889	4.5	3.444444444	1.8	5.384615385	3.923076923	5.785714286	5.88888889	5.65
Q. robur	5.933333333	1.9230769	5.2	3.214285714	3.545454545	3.315789474	5.153846154	4.916666667	5.8	5.11764706
Q. rubra	5.916666667	3.5	5.636363636	5.466666667	4.611111111	5.769230769	5.684210526	3.428571429	5.85	4.27272727
Q. shumardi	6	3.3571429	5.666666667		5.5	5.416666667	5.736842105	3.909090909	6	4.58333333
Q. similis	5.833333333	4.1176471	5.25	2.066666667	4.833333333	2.125	5.19047619	4.833333333	5.8	3.38461538
Q. sinuata	4.4	5.1875	4.666666667	2.4	3.625	3.944444444	2.533333333	5.647058824	4.42857143	5.42857143
Q. stellata	6	4.9166667	5.083333333	3.7	5.133333333	2.166666667	5.4	4.578947368	6	4.23076923
Q. texana	5.846153846	4.5833333	5.857142857	5.181818182	5.5	4.571428571	5.769230769	1.1875	5.93333333	3.25
Q. velutina	5.866666667	4.5	5.384615385	4.25	4.411764706	4.666666667	5.823529412	4.666666667	5.91666667	4.28571429
Q. virginiana	4.454545455	5.7857143	3.647058824	2.75	4.9375	4.133333333	1.772727273	5.909090909	3.57142857	5.83333333

Plant Species	Q. geminata	Q. georgiana	Q. hemisphaerica	Q. illicifolia	Q. imbricaria	Q. incana	Q. inopia	Q. laevis
Q. acutissima								
Q. alba								
Q. arkansana								
Q. austrina								
Q. bicolor								
Q. boyntonii								
Q. champanii								
Q. coccinea								
Q. ellottii								
Q. falcata								
Q. geminata	0							
Q. georgiana	5.842105263	0						
Q. hemisphaerica	2.416666667	5.705882353	0					
Q. ilicifolia	5.863636364	4	5.571428571	0				
Q. imbricaria	1.266666667	6	1.615384615	5.8	0			
Q. incana	1.842105263	5.8	1.272727273	5.454545455	1.588235294	0		
Q. inopia	2.3125	5.615384615	2.636363636	5.882352941	2.25	1.625	0	
Q. laevis	6	1.363636364	5.75	3.75	5.882352941	5.933333333	5.875	0
Q. laurifolia	2.692307692	5.4	2	5.086956522	1.846153846	1.909090909	2.555555556	5.533333333
Q. lyrata	5.933333333	2.142857143	5.466666667	2.866666667	5.8	5.928571429	5.722222222	2.6
Q. macrocarpa	5.923076923	4.714285714	5.615384615	4.714285714	5.777777778	5.916666667	6	4.615384615
Q. margaretta	5.6875	3.166666667	5.818181818	2.6875	6	5.5625	5.636363636	3.8
Q. marlandica	5.866666667	5.384615385	5.8	3.263157895	5.578947368	5.473684211	5.2	5.153846154
Q. michauxii	5.785714286	5.733333333	5.928571429	5.384615385	5.6	5.642857143	5.733333333	5.666666667
Q. minima	3.5625	5.789473684	1.642857143	4.666666667	2.076923077	2.875	4.1	5.461538462
Q. montana	5.666666667	5.736842105	5.2	5.533333333	5.545454545	5.3	5.230769231	6
Q. muehlenbergii	5.076923077	5.4375	4.533333333	5.071428571	4.7	4.733333333	4.5	5.785714286
Q. myrtifolia	3.1	5.9375	3.454545455	5.933333333	2.818181818	2.4	1.75	6
Q. nigra	3.166666667	5.583333333	2.75	4.538461538	3.375	3.363636364	4	5.785714286
Q. olgethorpensis	1.647058824	5.882352941	1.588235294	5.733333333	2.636363636	1.777777778	3	6
Q. pagoda	5.923076923	2.285714286	5.923076923	4.666666667	5.916666667	5.916666667	6	2.266666667
Q. palustris	5.95	2.083333333	5.705882353	3.923076923	6	6	5.933333333	1.857142857
Q. phellos	1.6875	5.666666667	2.571428571	5.875	2.181818182	2.142857143	3.470588235	5.933333333
Q. prinoides	5.666666667	5.666666667	5.0625	4.8	4.428571429	4.9	5.142857143	5.666666667
Q. robur	5.642857143	4.588235294	5.956521739	4.357142857	6	5.785714286	5.714285714	5
Q. rubra	6	4.4	5.916666667	4.785714286	6	5.9	5.789473684	3.785714286
Q. shumardi	5.846153846	3.25	5.909090909	4.571428571	6	5.909090909	5.6875	3.235294118
Q. similis	5.642857143	4.052631579	5.1875	2.411764706	5.733333333	5.25	5.785714286	4.166666667
Q. sinuata	3.846153846	5.615384615	2.416666667	4.666666667	3.6	3.4	4.111111111	5.631578947
Q. stellata	5.75	3.727272727	5.75	1.928571429	5.733333333	5.666666667	5.923076923	4.714285714
Q. texana	6	2.230769231	6	3.916666667	5.833333333	6	6	3.307692308
Q. velutina	5.846153846	3.923076923	6	2.647058824	6	6	5.866666667	3.933333333
Q. virginiana	2.818181818	5.5	1.45	4.266666667	3.285714286	2.625	2.833333333	5.25

Plant Species	Q. laurifolia	Q. lyrata	Q. macrocarpa	Q. margaretta	Q. marlandica	Q. michauxii	Q. minima	Q. montana	Q. muehlenbergii
Q. acutissima									
Q. alba									
Q. arkansana									
Q. austrina									
Q. bicolor									
Q. boyntonii									
Q. champanii									
Q. coccinea									
Q. ellottii									
Q. falcata									
Q. geminata									
Q. georgiana									
Q. hemisphaerica									
Q. ilicifolia									
Q. imbricaria									
Q. incana									
Q. inopia									
Q. laevis									
Q. laurifolia	0								
Q. lyrata	5.384615385	0							
Q. macrocarpa	5.307692308	3	0						
Q. margaretta	5.3	3.571428571	2.375	0					
Q. marlandica	5.615384615	5.583333333	3.941176471	4.25	0				
Q. michauxii	5.923076923	5.714285714	5.75	5.615384615	5.526315789	0			
Q. minima	1.818181818	4.928571429	5.058823529	5.307692308	5.307692308	5.1875	0		
Q. montana	4.909090909	5.8	5.692307692	5.714285714	5.666666667	1	5.166666667	0	
Q. muehlenbergii	4.666666667	5.642857143	5.583333333	5.857142857	5.368421053	2.272727273	4.111111111	1.923076923	0
Q. myrtifolia	4.076923077	5.947368421	5.923076923	5.923076923	5.823529412	5.692307692	5.142857143	4.714285714	4.647058824
Q. nigra	1.588235294	5.3125	4.846153846	5.166666667	3.375	5.6875	1.636363636	5.384615385	5.272727273
Q. olgethorpensis	1.705882353	5.785714286	5.9	5.8	5.4	5.5	2.076923077	5.466666667	4.636363636
Q. pagoda	5.933333333	3.111111111	4.533333333	4.928571429	5.8125	5.411764706	5.611111111	5.363636364	5.153846154
Q. palustris	5.75	3.846153846	4.571428571	3.866666667	5.769230769	5.733333333	5.846153846	5.909090909	5.875
Q. phellos	2.625	5.642857143	5.875	5.909090909	6	5.928571429	2.461538462	5.666666667	4.956521739
Q. prinoides	5	5	4.727272727	5.4	5.538461538	2.266666667	3.090909091	2	1.214285714
Q. robur	5.5	3.090909091	1.692307692	2.909090909	4	4.947368421	5.416666667	4.636363636	4.545454545
Q. rubra	5.8125	4.769230769	4.1875	5.095238095	5.181818182	4.933333333	5.615384615	4.769230769	5.333333333
Q. shumardi	5.636363636	4.363636364	4.2	4.19047619	5.5	5.333333333	5.666666667	5.636363636	5.3
Q. similis	4.153846154	2.315789474	3.388888889	3.3	3.923076923	5.714285714	3.583333333	5.538461538	5.666666667
Q. sinuata	2.363636364	4.133333333	5.277777778	5.642857143	5	4.933333333	1.090909091	4.636363636	3.083333333
Q. stellata	5.409090909	2.857142857	3.105263158	1.923076923	3.461538462	5.583333333	5.466666667	5.666666667	5.352941176
Q. texana	5.769230769	3.769230769	4.764705882	5	5.409090909	5.923076923	6	5.6875	5.857142857
Q. velutina	5.857142857	3.923076923	4.466666667	3.923076923	3.888888889	5.3125	5.642857143	5.3125	5.578947368
Q. virginiana	2.272727273	4.476190476	5.357142857	5.25	4.235294118	5.666666667	1.166666667	5.357142857	4.142857143

Plant Species	Q. myrtifolia	Q. nigra	Q. olgethorpensis	Q. pagoda	Q. palustris	Q. phellos	Q. prinoides	Q. robur	Q. rubra	Q. shumardi
Q. acutissima										
Q. alba										
Q. arkansana										
Q. austrina										
Q. bicolor										
Q. boyntonii										
Q. champanii										
Q. coccinea										
Q. ellottii										
Q. falcata										
Q. geminata										
Q. georgiana										
Q. hemisphaerica										
Q. ilicifolia										
Q. imbricaria										
Q. incana										
Q. inopia										
Q. laevis										
Q. laurifolia										
Q. lyrata										
Q. macrocarpa										
Q. margaretta										
Q. marlandica										
Q. michauxii										
Q. minima										
Q. montana										
Q. muehlenbergii										
Q. myrtifolia	0									
Q. nigra	3.875	0								
Q. olgethorpensis	4.8	3.6	0							
Q. pagoda	5.8	5.5	5.933333333	0						
Q. palustris	6	5.666667	5.818181818	2.09090909	0					
Q. phellos	4.545454545	3.8125	1.333333333	5.69230769	6	0				
Q. prinoides	5.714285714	5.038462	5.111111111	5.75	5.5	5.230769231	0			
Q. robur	5.9375	4.882353	6	5.46153846	4.7	5.727272727	4.666666667	0		
Q. rubra	5.923076923	5.909091	6	3.33333333	3.66666667	6	4.909090909	4.388889	0	
Q. shumardi	6	5.545455	5.866666667	3.05882353	3.6	5.833333333	5.857142857	4.428571	2	0
Q. similis	6	3.615385	5.454545455	4.38461538	4.58333333	5.5	5.125	3.071429	5.333333	5.333333333
Q. sinuata	4.5	2.933333	3.4375	5.58823529	5.30769231	4.692307692	1.727272727	4.5625	5.8	5.733333333
Q. stellata	5.944444444	4.294118	5.944444444	5.25	5.04761905	5.769230769	5.142857143	2.571429	5.3125	5.133333333
Q. texana	6	5.733333	6	2.70588235	2.25	6	5.272727273	4.857143	3.923077	2.647058824
Q. velutina	5.9	5.470588	6	4.21428571	3.83333333	6	5.428571429	3.411765	3.357143	2.875
Q. virginiana	4.111111111	2.789474	3.4	5.76470588	5.71428571	3.214285714	3.411764706	5.642857	5.923077	5.533333333

Plant Species	Q. similis	Q. sinuata	Q. stellata	Q. texana	Q. velutina	Q. virginiana
Q. acutissima						
Q. alba						
Q. arkansana						
Q. austrina						
Q. bicolor						
Q. boyntonii						
Q. champanii						
Q. coccinea						
Q. ellottii						
Q. falcata						
Q. geminata						
Q. georgiana						
Q. hemisphaerica						
Q. ilicifolia						
Q. imbricaria						
Q. incana						
Q. inopia						
Q. laevis						
Q. laurifolia						
Q. lyrata						
Q. macrocarpa						
Q. margaretta						
Q. marlandica						
Q. michauxii						
Q. minima						
Q. montana						
Q. muehlenbergii						
Q. myrtifolia						
Q. nigra						
Q. olgethorpensis						
Q. pagoda						
Q. palustris						
Q. phellos						
Q. prinoides						
Q. robur						
Q. rubra						
Q. shumardi						
Q. similis	0					
Q. sinuata	2.666666667	0				
Q. stellata	3.210526316	4.785714286	0			
Q. texana	4.833333333	5.857142857	4.642857143	0		
Q. velutina	4.8	5.214285714	3.75	4.07142857	0	
Q. virginiana	4.736842105	1.705882353	4.888888889	5.90909091	5.38461538	0

NON-EXPERT SIMILARITY MATRIX

Plant Species	Q. acutissima	Q. alba	Q. arkansana	Q. austrina	Q. bicolor	Q. boyntonii	Q. champanii	Q. coccinea	Q. elliottii	Q. falcata
Q. acutissima	0									
Q. alba	5.727272727	0								
Q. arkansana	3.909090909	5.5714286	0							
Q. austrina	4	5	3.642857143	0						
Q. bicolor	4.928571429	3.8333333	3.538461538	3.2	0					
Q. boyntonii	5.428571429	3.7692308	3.454545455	1.8	3.666666667	0				
Q. champanii	3.470588235	5.8666667	2.636363636	3.818181818	4.153846154	4.333333333	0			
Q. coccinea	5.8	2.8	5.615384615	4.733333333	4.692307692	5.083333333	5.875	0		
Q. ellottii	2.272727273	6	5.545454545	5.4	6	5.428571429	3.818181818	5.916666667	0	
Q. falcata	5.842105263	3.4285714	5.5	4	5	4.666666667	5.866666667	2.363636364	5.88888889	0
Q. geminata	2.538461538	6	4.454545455	4.928571429	5.733333333	4.769230769	2.3125	5.916666667	1.6	5.875
Q. georgiana	5.533333333	2.3571429	5.571428571	4.090909091	4.4	4.266666667	5.5	1.416666667	5.72727273	2.08333333
Q. hemisphaerica	2.166666667	5.9333333	3.8125	3.416666667	5	4.066666667	1.2	5.714285714	1.90909091	6
Q. ilicifolia	5.727272727	2.7333333	4.8	2.352941176	4	2.307692308	4.466666667	4.142857143	5.91666667	2.90909091
Q. imbricaria	2.75	5.875	3.363636364	4.470588235	5.666666667	4.666666667	1.5	5.846153846	2.5	5.88888889
Q. incana	2.294117647	5.9090909	2.941176471	4.266666667	5.538461538	3.818181818	1.285714286	5.923076923	2.58333333	5.75
Q. inopia	4.285714286	5.9285714	2.5	4.090909091	5.411764706	5.1	2	5.916666667	3.14285714	5.9
Q. laevis	5.7	2.0666667	5.4	4.6	4.923076923	4.875	5.8	1.6	6	1.58333333
Q. laurifolia	2.083333333	5.3125	3.571428571	3.25	4.538461538	3.733333333	2.333333333	5.25	2.57894737	5.125
Q. lyrata	5.333333333	2.3529412	4.9	3	4.05	2.866666667	5	3.181818182	5.54545455	2.8
Q. macrocarpa	5.272727273	1.3571429	4.714285714	3.666666667	4.117647059	2.384615385	5.333333333	3.727272727	5.9	4.3
Q. margaretta	5.416666667	2	5.214285714	3.533333333	3.941176471	2.769230769	5.583333333	3.916666667	5.76923077	3.85714286
Q. marlandica	5.545454545	4.7142857	2.2	2.833333333	3.714285714	2.266666667	4.470588235	5.25	5.72727273	4.85714286
Q. michauxii	3.928571429	4.6428571	4.076923077	4.466666667	1.272727273	5.083333333	4.583333333	5.75	5.92307692	5.47058824
Q. minima	2.666666667	5.6363636	4.3125	2.692307692	4.230769231	3.277777778	1.454545455	5.666666667	2	5.38461538
Q. montana	3	5.0833333	4.090909091	4.153846154	1.75	4.642857143	4.071428571	5.5	5.53846154	5.58333333
Q. muehlenbergii	2.692307692	5.4666667	3.846153846	3.75	1.928571429	4.466666667	2.636363636	6	5.41666667	5.41666667
Q. myrtifolia	4.5	6	2.6	4.470588235	5.384615385	5.5	1.833333333	5.882352941	3.08333333	5.58333333
Q. nigra	3.8	5.5882353	3.4	2.619047619	4	2.375	2.2	5.636363636	3.75	4.86666667
Q. olgethorpensis	2.214285714	5.8666667	4.25	3.631578947	5.235294118	4.928571429	2.333333333	5.7	1.95	5.71428571
Q. pagoda	6	2.2	6	4.533333333	4.583333333	5.1875	5.866666667	1.454545455	5.91666667	1.72727273
Q. palustris	5.692307692	3.0588235	5.571428571	5	5	5.3	5.642857143	1	5.5	1.75
Q. phellos	2	5.9090909	5.157894737	4	5.6	4.875	4	5.916666667	1.21428571	5.6
Q. prinoides	4	4.9	4.263157895	2.846153846	1.916666667	4.083333333	2.666666667	5.666666667	5.45454545	5.35294118
Q. robur	5.461538462	1.8125	4.238095238	2.333333333	2.928571429	2.714285714	5.083333333	4.4	5.75	4.6
Q. rubra	5.692307692	2.0833333	5.6	4.833333333	3.333333333	4.888888889	5.833333333	2.545454545	6	4
Q. shumardi	5.461538462	2.5714286	5.3	4.428571429	4.066666667	4.9	5.75	2.833333333	6	3.53333333
Q. similis	5.076923077	2.8461538	5.142857143	2	3.25	1.916666667	4.384615385	4.454545455	5.58333333	2.72727273
Q. sinuata	3.384615385	5.6666667	4.083333333	1.909090909	2.4	3	2.5	5.6	3.0625	5.2
Q. stellata	5.636363636	3.1875	4.571428571	3.583333333	4.357142857	2.25	5.071428571	3.611111111	6	3.45454545
Q. texana	5.705882353	2.8	5.428571429	5.076923077	5.142857143	4.823529412	5.428571429	1.3	5.875	2.90909091
Q. velutina	5.636363636	3	4.705882353	3.272727273	3.833333333	3.666666667	5.705882353	4.272727273	6	4.33333333
Q. virginiana	3.230769231	5.4285714	3.230769231	2.176470588	4.2	3.416666667	2.583333333	5.916666667	2.64285714	5.38888889

| Plant Species	Q. geminata	Q. georgiana	Q. hemisphaerica	Q. illicifolia	Q. imbricaria	Q. incana	Q. inopia	Q. laevis
Q. acutissima								
Q. alba								
Q. arkansana								
Q. austrina								
Q. bicolor								
Q. boyntonii								
Q. champanii								
Q. coccinea								
Q. ellottii								
Q. falcata								
Q. geminata	0							
Q. georgiana	5.818181818	0						
Q. hemisphaerica	1.166666667	5.272727273	0					
Q. ilicifolia	5.583333333	3	5.166666667	0				
Q. imbricaria	1.363636364	5.529411765	1.357142857	5.25	0			
Q. incana	1.583333333	5.75	1.111111111	5.1875	1.1	0		
Q. inopia	2.235294118	5.7	2.166666667	5.538461538	2.384615385	1.785714286	0	
Q. laevis	5.916666667	1.3125	5.470588235	2.818181818	6	5.833333333	6	0
Q. laurifolia	1.5	4.857142857	1.666666667	5	1.8	2.1	3	5.5
Q. lyrata	5.454545455	2	5.333333333	2	5.416666667	5.666666667	5.55	2.611111111
Q. macrocarpa	5.8	3.2	5.333333333	3.5	5.764705882	5.076923077	5.428571429	3.5
Q. margaretta	5.470588235	2.35	5.666666667	2.083333333	5.5	5.6	5.5	3.2
Q. marlandica	5.545454545	4.823529412	5.545454545	2.916666667	4.941176471	4.833333333	4.583333333	5.181818182
Q. michauxii	5.7	5.466666667	5.176470588	4.454545455	5.4	5	5.461538462	5.083333333
Q. minima	2.166666667	5.615384615	1.642857143	5.4	2	2.3125	3.25	5.555555556
Q. montana	4.842105263	5.461538462	4.933333333	4.5	4.818181818	3.727272727	5.125	5.583333333
Q. muehlenbergii	4.4	5.333333333	4.071428571	4.727272727	4.076923077	3.090909091	3.923076923	5.4
Q. myrtifolia	2	5.583333333	2.2	5.411764706	2.3	2.4	1	5.909090909
Q. nigra	3.071428571	4.727272727	2.8	3.833333333	3.285714286	3.833333333	3.785714286	5.058823529
Q. olgethorpensis	1.615384615	5.166666667	1	5.692307692	1.5	1.363636364	2.461538462	5.916666667
Q. pagoda	5.923076923	2	5.857142857	3.571428571	5.928571429	5.909090909	5.8	1.266666667
Q. palustris	5.9	1.6	5.833333333	4.153846154	5.533333333	5.733333333	5.785714286	1.666666667
Q. phellos	1.5	5.6	1.3	5.833333333	2.083333333	4	2.833333333	5.928571429
Q. prinoides	4.928571429	5.176470588	3.8125	3.357142857	4.823529412	3.416666667	3.909090909	5.727272727
Q. robur	5.733333333	3.1	5.615384615	2.75	5.5	5.416666667	5.636363636	3.25
Q. rubra	5.882352941	2.5625	6	4.055555556	6	6	5.916666667	2
Q. shumardi	5.928571429	2.75	5.882352941	3.545454545	5.916666667	5.8	5.916666667	2.611111111
Q. similis	5.142857143	2.909090909	4.714285714	1.916666667	4.928571429	5.5	4.933333333	3
Q. sinuata	3.357142857	5.090909091	1.785714286	4	2.916666667	2.25	3.5	5.3
Q. stellata	5.727272727	4.083333333	5.5	1.928571429	5.692307692	5.5	5.666666667	3.25
Q. texana	6	1.25	5.7	2.833333333	5.615384615	5.909090909	5.923076923	1.571428571
Q. velutina	5.916666667	4.083333333	5.818181818	2	5.363636364	5.833333333	5.75	3.294117647
Q. virginiana	1.636363636	5.545454545	1.5	3.75	2.583333333	2.833333333	2.083333333	5.8

Plant Species	Q. laurifolia	Q. lyrata	Q. macrocarpa	Q. margaretta	Q. marlandica	Q. michauxii	Q. minima	Q. montana	Q. muehlenbergii
Q. acutissima									
Q. alba									
Q. arkansana									
Q. austrina									
Q. bicolor									
Q. boyntonii									
Q. champanii									
Q. coccinea									
Q. ellottii									
Q. falcata									
Q. geminata									
Q. georgiana									
Q. hemisphaerica									
Q. ilicifolia									
Q. imbricaria									
Q. incana									
Q. inopia									
Q. laevis									
Q. laurifolia	0								
Q. lyrata	4.545454545	0							
Q. macrocarpa	5.105263158	2	0						
Q. margaretta	5.333333333	2.428571429	2.111111111	0					
Q. marlandica	5.1	5.076923077	4	4.333333333	0				
Q. michauxii	5.090909091	5.166666667	5.210526316	5	4.615384615	0			
Q. minima	1.529411765	3.833333333	5.352941176	4.5	4.571428571	4.666666667	0		
Q. montana	4.363636364	5.7	5.5	5.461538462	5	1.25	4.529411765	0	
Q. muehlenbergii	3.916666667	4.818181818	4.4	5.75	5.083333333	1.466666667	3	1.6	0
Q. myrtifolia	3.928571429	5.8	5.733333333	5.727272727	4.9	5.866666667	4.285714286	4.5	4.066666667
Q. nigra	1.3	4.636363636	5.181818182	5.090909091	4.076923077	4.933333333	2.333333333	4.8	4.333333333
Q. olgethorpensis	2.461538462	5.333333333	5.357142857	5.454545455	5.5	5.357142857	2	5.214285714	3.461538462
Q. pagoda	5.578947368	1.866666667	2.909090909	4.294117647	4.545454545	5.6	5.272727273	4.285714286	4.916666667
Q. palustris	5.5	2.75	3.294117647	4	5.153846154	5.454545455	5.166666667	5.294117647	5.5
Q. phellos	1.75	5.7	5.571428571	5.307692308	5.727272727	5.733333333	1.866666667	5.5	4.333333333
Q. prinoides	3.529411765	4.769230769	4.454545455	4.5	4.3125	1.833333333	3.461538462	1.9	1.666666667
Q. robur	5.235294118	3.5	2	2.153846154	2.9375	4.166666667	5	4.5	5.090909091
Q. rubra	5.882352941	3.315789474	3.764705882	4.307692308	4.75	4.222222222	5.75	4.7	4.785714286
Q. shumardi	5.363636364	3.277777778	2.7	3.352941176	5.181818182	4.727272727	5.6	5.363636364	4.875
Q. similis	3.411764706	1.15	3	3.117647059	3.916666667	5.083333333	3.25	4.916666667	4.153846154
Q. sinuata	2.133333333	3.583333333	4.6	5.1875	4.384615385	4.722222222	1.5	4.15	2.466666667
Q. stellata	5	3.3	2.25	1.3125	2.090909091	5.333333333	4.764705882	4.916666667	5.454545455
Q. texana	5.785714286	2.375	4	2.846153846	5.5	4.583333333	5.55	5.363636364	5.4
Q. velutina	5.529411765	3.666666667	3.833333333	3.818181818	3	4.230769231	5.578947368	4.571428571	4.533333333
Q. virginiana	1.454545455	4.9375	4.2	5	4.166666667	4.230769231	1.2	3.533333333	3.416666667

Plant Species	Q. myrtifolia	Q. nigra	Q. olgethorpensis	Q. pagoda	Q. palustris	Q. phellos	Q. prinoides	Q. robur	Q. rubra	Q. shumardi
Q. acutissima										
Q. alba										
Q. arkansana										
Q. austrina										
Q. bicolor										
Q. boyntonii										
Q. champanii										
Q. coccinea										
Q. ellottii										
Q. falcata										
Q. geminata										
Q. georgiana										
Q. hemisphaerica										
Q. ilicifolia										
Q. imbricaria										
Q. incana										
Q. inopia										
Q. laevis										
Q. laurifolia										
Q. lyrata										
Q. macrocarpa										
Q. margaretta										
Q. marlandica										
Q. michauxii										
Q. minima										
Q. montana										
Q. muehlenbergii										
Q. myrtifolia	0									
Q. nigra	4	0								
Q. olgethorpensis	3.416666667	2.3	0							
Q. pagoda	6	5.533333	5.846153846	0						
Q. palustris	5.933333333	5.5	5.833333333	1.41666667	0					
Q. phellos	3.555555556	3.166667	1.142857143	5.69230769	6	0				
Q. prinoides	4	4.1	4	5.18181818	5.6	4.466666667	0			
Q. robur	5.7	4.333333	5.857142857	4.6	4.46153846	5.933333333	3.818181818	0		
Q. rubra	5.857142857	5.916667	5.833333333	1.75	3.41666667	5.857142857	4.818181818	3.666667	0	
Q. shumardi	5.916666667	5.642857	5.714285714	2.16666667	3.21428571	5.909090909	4.8	3.25	1.071429	0
Q. similis	5.8	2.636364	5.285714286	3.09090909	3.6	5.071428571	4.666666667	2.25	4.066667	3.857142857
Q. sinuata	4.142857143	2.058824	2.15	4.81818182	5.21428571	3.9	1.428571429	4.666667	5.133333	5.5
Q. stellata	5.5625	4.3	5.384615385	5	4.05882353	5.909090909	4.9	2.071429	4.642857	4.5
Q. texana	5.769230769	5.466667	5.909090909	1.5	1.15384615	5.6875	4.928571429	3.5	2.666667	1.6
Q. velutina	6	5.133333	5.916666667	3.6	4	5.909090909	5.142857143	2.857143	3.333333	2.166666667
Q. virginiana	2.466666667	1.428571	1.909090909	5.76470588	5.63636364	2.857142857	2.1	4.916667	5.714286	5.833333333

Plant Species	Q. similis	Q. sinuata	Q. stellata	Q. texana	Q. velutina	Q. virginiana
Q. acutissima						
Q. alba						
Q. arkansana						
Q. austrina						
Q. bicolor						
Q. boyntonii						
Q. champanii						
Q. coccinea						
Q. ellottii						
Q. falcata						
Q. geminata						
Q. georgiana						
Q. hemisphaerica						
Q. ilicifolia						
Q. imbricaria						
Q. incana						
Q. inopia						
Q. laevis						
Q. laurifolia						
Q. lyrata						
Q. macrocarpa						
Q. margaretta						
Q. marlandica						
Q. michauxii						
Q. minima						
Q. montana						
Q. muehlenbergii						
Q. myrtifolia						
Q. nigra						
Q. olgethorpensis						
Q. pagoda						
Q. palustris						
Q. phellos						
Q. prinoides						
Q. robur						
Q. rubra						
Q. shumardi						
Q. similis	0					
Q. sinuata	2.954545455	0				
Q. stellata	2.357142857	5	0			
Q. texana	3.545454545	5.733333333	3.7	0		
Q. velutina	4	4.857142857	3	3.16666667	0	
Q. virginiana	3.5	1.375	4.846153846	5.83333333	5.33333333	0

APPENDIX E

DENDROGRAM RESULTING FROM CENTROID LINKAGE METHOD

Cluster Dendrogram

d
hclust (*, "centroid")
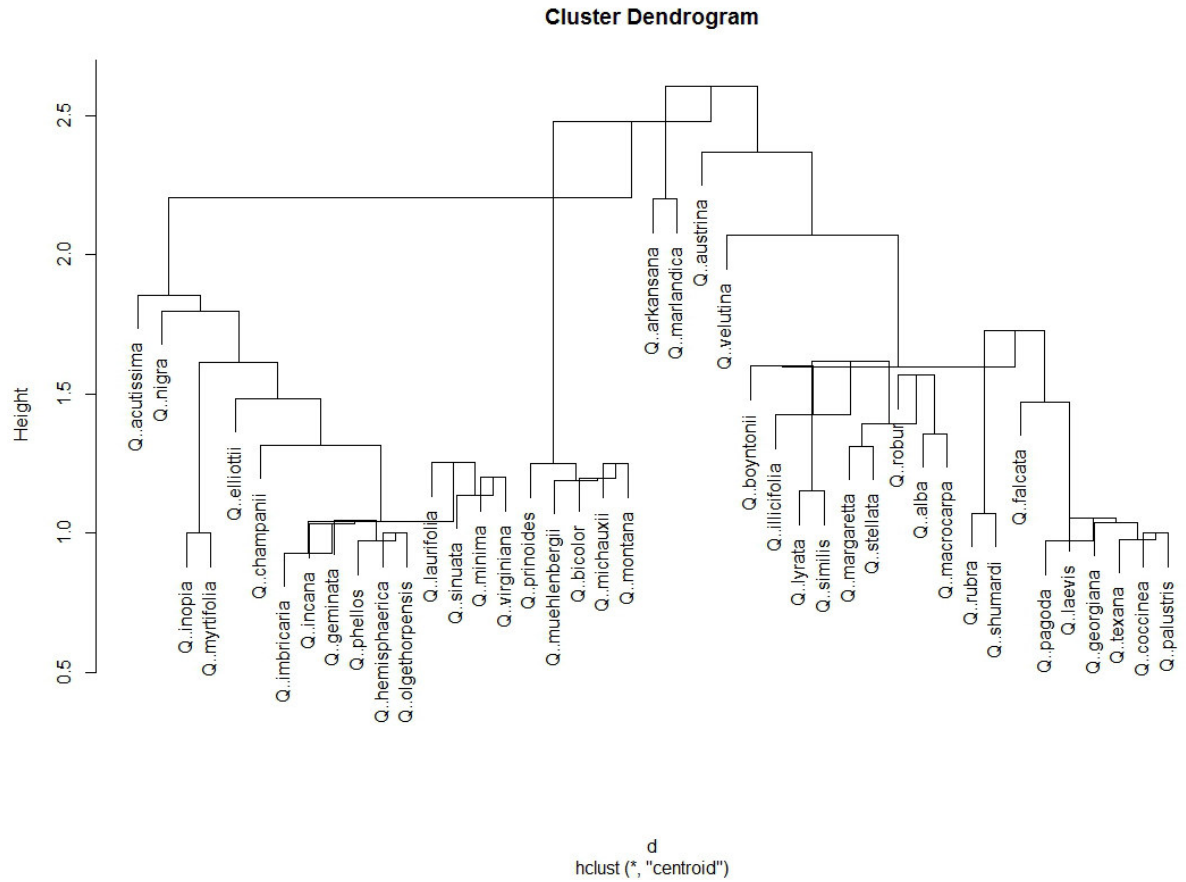
APPENDIX F

R CODE FOR DENDROGRAM CREATED USING WARD'S METHOD

```
d <- as.dist(leaf, diag = FALSE, upper = FALSE)

d

fit <- hclust(d, method="ward")

plot(fit)
```

POST-TEST EXPERT SURVEY

Post-test Survey

1. What is your occupation?

2. Approximate how many oaks, native to the southeast U.S., you can identify by sight?

    ○ None

    ○ Between 1 and 5

    ○ Between 6 and 10 species

    ○ Between10 and 20 species

    ○ More than 20 species

Please answer the following questions based on your experience using Keys A and B.

3. Using the following scale, place an X in the correct position for which key was easier to use.

Key A   Key B

_____ _____ _____ _____ _____ _____ _____

A easier   equal   B easier

4. How would you rate how easy Key A was to use?

     ○ Difficult

     ○ Somewhat difficult

     ○ Somewhat easy

     ○ Easy

5. How would you rate how easy Key B was to use?

     ○ Difficult

     ○ Somewhat difficult

     ○ Somewhat easy

     ○ Easy

6. Please describe some of the strengths of Key A

7. Please describe some of the weaknesses of Key A.

8. Please describe some of the strengths of Key B.

9. Please describe some of the weaknesses of Key B.

POST-TEST NON-EXPERT SURVEY

Student # _____

Post-test Survey

Please answer the following questions based on your personal experience using the identification key assigned to you.

1. Please indicate which key you were given

○ Key A

○ Key B

2. What was your experience identifying plants before participating in today's task? Check all that apply

○ None

○ I can identify less than 10 plants in the wild, by sight

○ I have used a commercial or picture guide

○ I have occasionally used a plant key that uses written descriptions

○ I am proficient with using a plant key that uses written descriptions

3. Have you taken BIO 354: **Plant Systematics** at UNCG?

○ Yes

○ No

4. Prior to this test, approximate how many native oaks you could identify by sight?

  ○ None

  ○ Between 1 and 5

  ○ Between 6 and 10 species

  ○ More than 10 species

5. How would you rate how easy the key was to use?

  ○ Difficult

  ○ Somewhat difficult

  ○ Somewhat easy

  ○ Easy

6. Please describe the difficulties you experienced while using the key.

7. Please describe what you think were the strengths of the key.