

Who am I this time?
The affinities and misbehaviors of Hill's oak.
(*Quercus ellipsoidalis* E. J. Hill).

Andrew Hipp^{1,2,3}, Jaime Weber¹, and Alka Srivastava¹

¹ The Morton Arboretum / 4100 Illinois Route 53 / Lisle IL 60532-1293, USA

² The Field Museum / 1400 S. Lakeshore Drive / Chicago IL 60605, USA

³ Author for correspondence (ahipp@mortonarb.org). This paper was originally presented in modified form at the 6th International Oak Society Conference in Puebla, Mexico, 20 October 2009.

Introduction

Oaks, it has long been recognized, readily hybridize. They do not obey the limit to interspecific hybridization that is the hallmark of the biological species concept [1], and they have in fact been described by two of the leaders in the field of speciation as a “worst case scenario for the biological species concept” [2]. Pioneering work by James Hardin [3] showed hybridization among 14 of the 16 species recognized for eastern North America within the subgenus *Quercus* (white oaks), with hybridization occurring almost anywhere that white oaks grow in sympatry. In the molecular era, hybridization is often suggested by chloroplast sharing among species [4-6]. But in spite of this fact, nuclear markers in sufficient numbers can distinguish oak species [7-9], and the genetic groupings that result generally accord closely with our understanding of oak species limits based on morphology, geography, and ecology. The congruence among these data sources strongly suggests that oak species are genetically coherent across broad ranges.

In the North American Great Lakes region, the taxonomy of Hill's oak (*Quercus ellipsoidalis* E.J.Hill) and scarlet oak (*Q. coccinea* Münchh.) has long been recognized as problematic [8, 10-16]. The two species are largely allopatric but overlap in characters of the end buds, leaves, and acorns. In their purest expressions, there is no confusing them, but the overlap raises the question of whether the two are better treated as separate species [8, 17-19] or as endpoints on a morphological continuum [11, 12, 14, 15, 20, 21]. A second issue is the degree to which these species hybridize with black oak (*Quercus velutina* Lam.), which is widespread throughout the region occupied by both species and morphologically entangled with them [19].

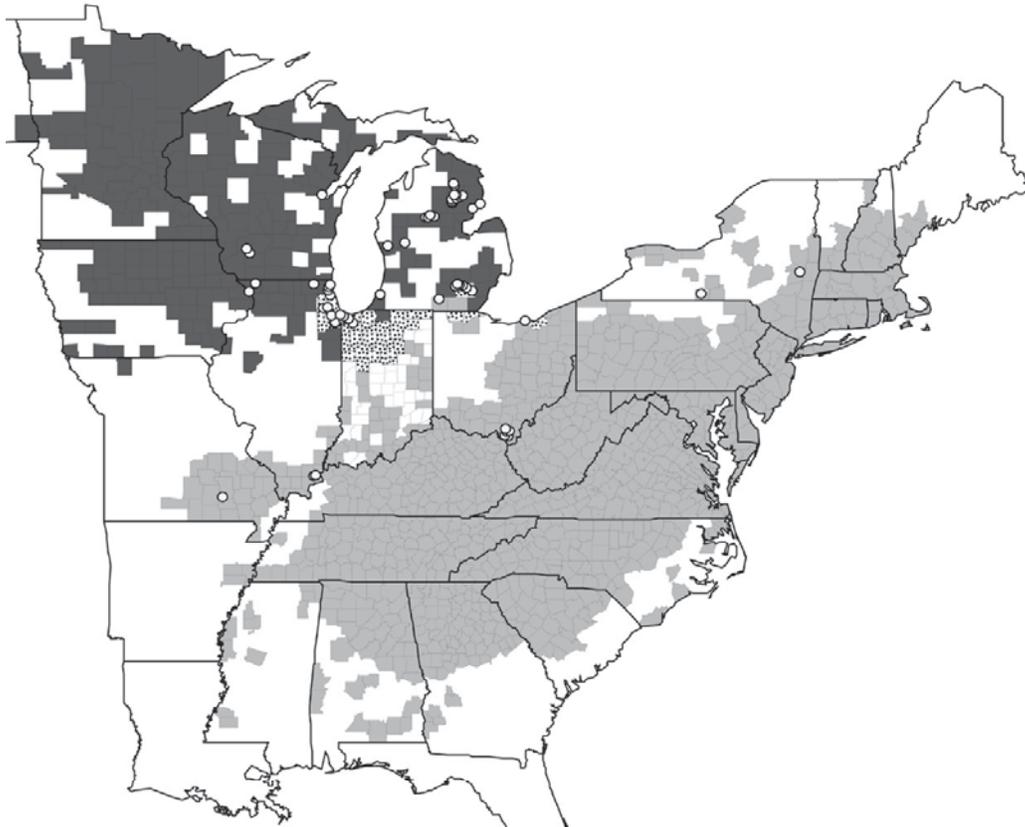
Previous molecular genetics work on these species demonstrated that scarlet oak is distinct from both Hill's oak and black oak, that black oak and Hill's oak are genetically similar to one another and may hybridize, and that genetic intermediacy between black oak and Hill's oak correlates poorly with morphological intermediacy [8]. In demonstrating that Hill's oak and black oak grade into each other morphologically and genetically in the Great Lakes region, while Hill's oak and scarlet oak grade into each other morphologically but apparently not genetically, the study raised the intertwined issues of the evolutionary origins of Hill's oak and relatives and their ongoing gene flow. Is

Hill's oak derived from within black oak, its similarity to scarlet oak an expression of shared ancestral characters that have been lost in black oak? Alternatively, is Hill's oak more closely related to scarlet oak, and genetically similar to black oak due to interspecific gene flow? The alternative effects of gene flow and phylogeny in explaining genetic similarity can be difficult to tease apart [22, 23], and in this study we utilize a combination of population-genetic, phylogenetic, and biogeographic approaches to try to do so.

Our current work is focused on three primary questions. First, we are interested in the phylogeny (evolutionary relationships) of the black oak species of eastern North America, particularly in order to know whether Hill's oak and scarlet oak are in fact sister species. Second, we are interested in deepening our understanding of the genetic disjunction between Hill's oak and scarlet oak by sampling populations of the former from northwest Indiana and southern Michigan, where the two species have been thought to be sympatric, and of the latter from northeastern North America, where the morphology is quite different from the forms observed in southern Missouri, Illinois, and Ohio. Finally, we are interested in evaluating the patterns of gene flow within Hill's oak and black oak and among populations of the two species, to understand whether the limited genetic structure that we found within each species in our previous work holds up with additional sampling and whether there is local gene flow between the species.

Methods

Our study is based on a genotyped sample of 803 individuals from six species—black oak, Hill's oak, scarlet oak, pin oak (*Quercus palustris* Münchh.), red oak (*Q. rubra* L.), and Shumard's oak (*Q. shumardii* Buckley)—collected from 58 sites, including 803 individuals from the three focal species and 24 from red oak (Figure 1). Individuals were assigned to populations based on their identification and site; thus, individuals of black oak from a given site were considered one population, while individuals of Hill's oak or scarlet oak from that same site were considered a second population. Sites were defined by groups of individuals found within an area of 4 km in radius about a common central point, which was geolocated using a GPS (global positioning system) unit. Genotypes were characterized using amplified fragment length polymorphism (AFLP) data [24] using methods described in [8]. Genetic disjunctions were identified both at the individual level and at the population level. Gene flow among and within species was assessed only at the population level. Although our morphological determinations largely correspond with the genetic clusters identified in this study, in several cases there is a mismatch between genetic and morphological identifications of individuals. Consequently, population-based analyses were conducted on three sets of individuals: all individuals identified based on morphology and reliably assigned to one species; all individuals identified based on genetic clustering with $\geq 65\%$ assignment to a single species based on Bayesian clustering; and all individuals that had morphological and genetic identifications the same and $\geq 65\%$ assignment to a single species based on Bayesian clustering. This last grouping, referred to hereafter in the paper as the ID-consensus group,



Phylogenetics and Evolution in R language. Bioinformatics 20:289-290.

Figure 1. Map of species distributions, with sampling localities.

The distribution of Hill's oak (*Quercus ellipsoidalis*) is mapped in dark grey, the distribution of scarlet oak (*Q. coccinea*) in light grey. Speckling indicates counties in which both species have been reported. Dots indicate sites where species were sampled for the current study. Note that only pin oak (*Q. palustris*) was sampled from the northern Ohio locality. Base map adapted from [8], with Indiana distribution according to [26].

is the most stringent and is the only one reported in most of this paper. Results of analyses using the other groupings had no effect on the conclusions of this study.

Results and Discussion

Phylogeny. Our first major finding is that with broadened sampling and additional data, we find that Hill's oak and scarlet oak are sister species and genetically coherent (Figure 2). The sister species relationship between Hill's oak and scarlet oak is borne out with additional sampling of 10 AFLP loci and well supported (bootstrap support = 0.85; data not shown). This is a satisfying result, given the long-standing taxonomic uncertainty regarding Hill's oak and scarlet oak. It also answers a question raised in our previous work about the degree to which the genetic similarity of Hill's oak and black oak might be due to recent divergence versus introgression ([8], p. 155). As Hill's oak and scarlet oak are

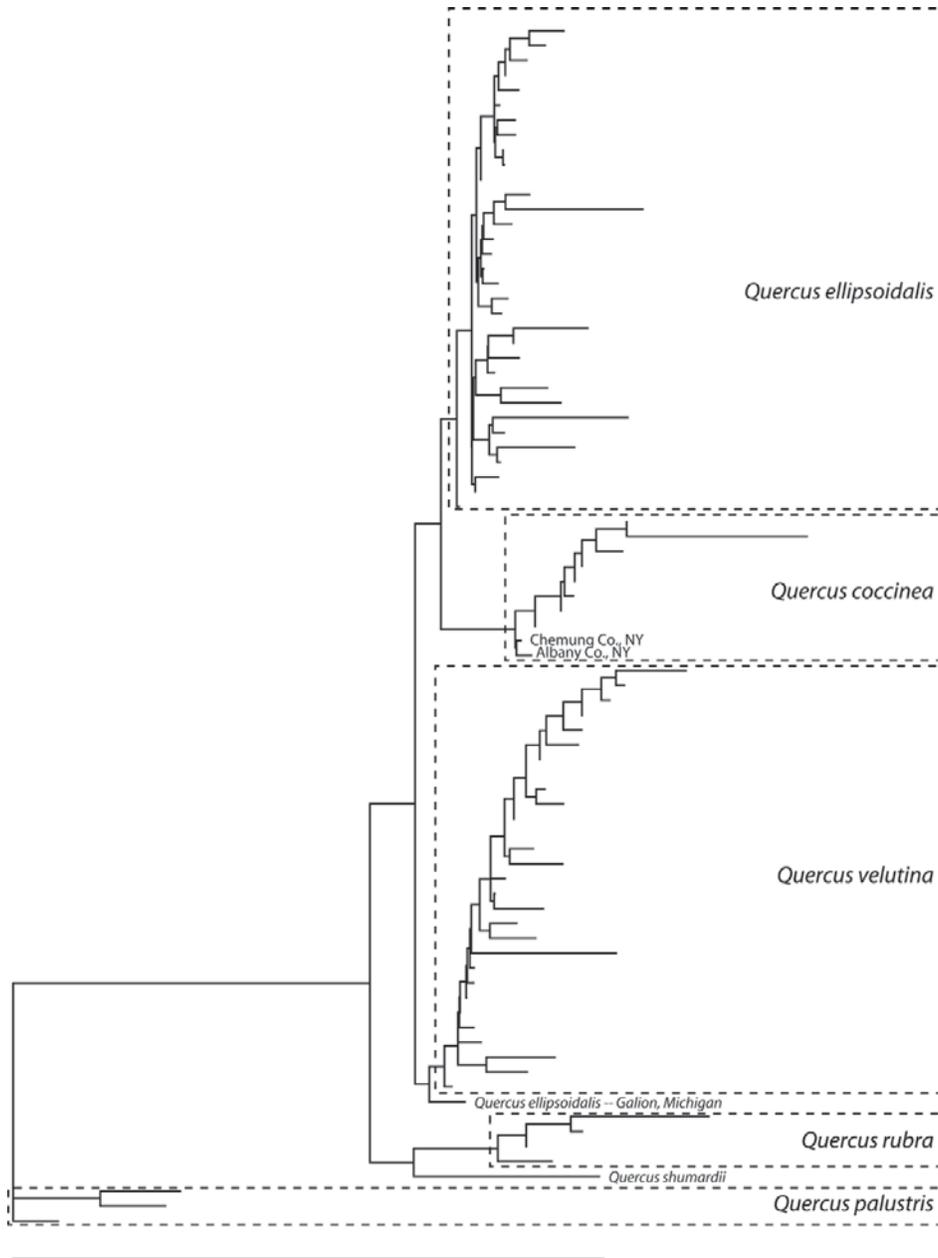


Figure 2. Relationships among populations of Hill's oak (*Quercus ellipsoidalis*) and allies.

Genetic distances among all populations were estimated using Nei's genetic distance, with distances corrected for dominant markers using the method of Lynch and Milligan [27] as implemented in AFLP-SURV [28]. The resulting pair wise genetic distance matrix was used to estimate genetic relationships among populations using the unweighted minimum evolution algorithm of Desper and Gascuel [29] as implemented in ape [30]. For sake of reconstructing relationships among populations, individuals included in analysis were limited to the ID-consensus group with population assignment $\geq 80\%$ to one species. Sensitivity of the topology and population distinctions will be investigated in subsequent work.

sister species, their morphological similarity is most parsimoniously explained as a consequence of shared derived characteristics. The genetic similarity between Hill's oak and black oak is likely due to ongoing gene flow between the two.

One aspect of our finding is especially interesting in light of recent work by Dave Shepard [15] that demonstrates the morphological similarity between populations of scarlet oak in northeastern North America and Hill's oak of the Great Lakes region. We now find from the standpoint of molecular genetics that the New York populations of scarlet oak that we sampled are genetically more similar to Hill's oak than any of the other scarlet oak populations we sampled (Figure 2). The genetic disjunction between the species still appears to be broad, as our previous work suggested. Whether morphologically intermediate populations of Hill's oak and scarlet oak are consistently genetically intermediate requires additional study.

Species distinctions and hybridization. Our second major finding is that when we analyze individuals rather than populations, the species separate out cleanly (data not shown), though with some remarkable misclassifications between black oak and Hill's oak, i.e., incongruence between our identifications based on morphology and the population assignments based on genetic data (e.g., 14 *Q. velutina* out of 286 sampled with > 0.20 assignment to *Q. ellipsoidalis*). This mismatch between genetic and morphological species assignments is a hallmark of introgressive hybridization and has been reported previously in oaks [25], and the presence of such individuals supports the hypothesis of gene flow between the two species. Gene flow between the two species is also supported by preliminary study of pubescence on the inner surface of the acorn cap, which appears to correlate with genetic admixture estimates (Figure 3, next page).

This work is preliminary, and we are investigating further to determine whether the correlations will hold up with additional sampling. There is also some evidence of allele-sharing between Hill's oak and scarlet oak (3 *Q. coccinea* out of 101 sampled with > 0.20 assignment to *Q. ellipsoidalis*; 16 *Q. ellipsoidalis* out of 370 sampled with > 0.20 assignment to *Q. coccinea*). However, the presumably recent divergence between these two species begs the question of whether this apparent allele sharing is a consequence of shared ancestral polymorphism or ongoing gene flow in spite of allopatry.

Related to the hypothesis of introgressive gene flow between Hill's oak and black oak, we find some evidence for allele sharing between local populations of the two species in the upper Great Lakes region, where the two occur in sympatry. This finding is based on our finding of a positive correlation between geographic distance and genetic distance in pairwise comparisons between black oak and Hill's oak populations ($r = 0.1087$, 2-tailed $P = 0.073$ based on 2000 permutations). The significance of this correlation is not strong, and while the correlation is still positive even when only individuals with $\geq 95\%$ assignment to one species are included in analysis ($r = 0.1316$, 2-tailed $P = 0.216$ based on 2000 permutations; note that the stronger correlation but weaker significance may be a consequence of the reduced sample size in the second test), the finding bears further study. However, if borne out, localized gene flow between the species would accord with previous morphological work by Jensen [19] suggesting that hybridization between Hill's oak and black oak may be relatively common.

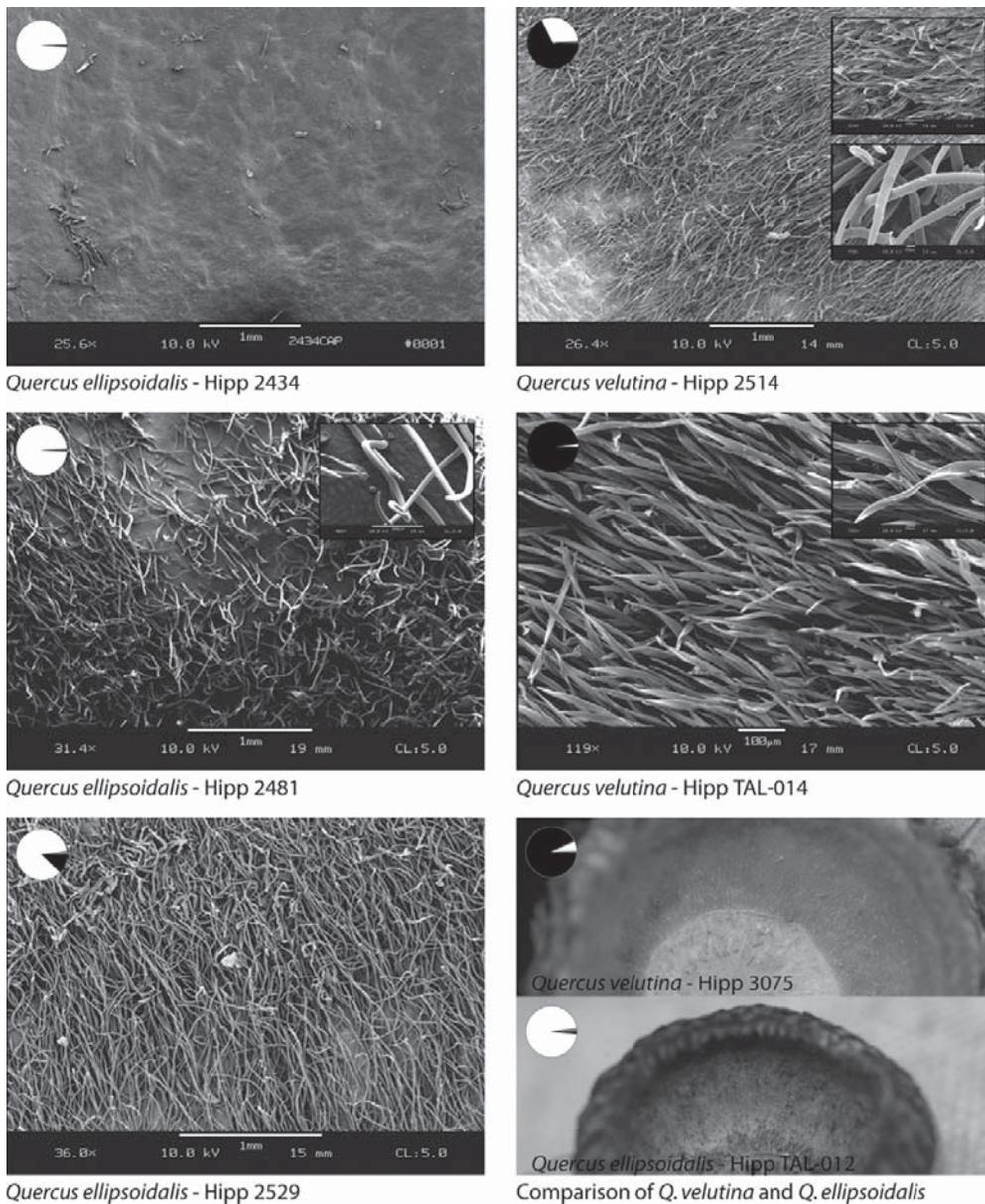


Figure 3. Pubescence on the inner surface of the acorn cap in Hill's oak (*Quercus ellipsoidalis*) and black oak (*Q. velutina*).

Images show scanning electron microscope images (A–E) or binocular light microscope images (F, G) of the inner surface of the acorn cap in Hill's oak (A–C, F) and black oak (D, E, G). Pie charts in the upper left-hand corner of each panel indicate the percentage of each individual's genome estimated to have derived from each of four ancestral populations: scarlet oak, *Q. coccinea* (red); Hill's oak (white), red oak, *Q. rubra* (yellow); and black oak (black). The figure demonstrates that pubescence on the inner surface of the acorn cap in the Hill's oak individuals sampled ranges from absent (A, G) to moderate (panel B); in the latter case, trichomes are slender and approximately cylindrical, often

(caption information continued on next page)

curling. In most of the black oaks we have sampled, pubescence is dense (E, F), the trichomes flattened and more or less appressed. In individuals that are genetically intermediate moderate pubescence may be found (C, D), along with both trichome types (D). This illustration represents a preliminary investigation of the relationship between genetic admixture and pubescence on the inner surface of the acorn cap, and bears further study. In particular, ongoing work in northern Michigan suggests that the pubescence type typical of black oak may be found in some individuals that are admixed or genotype as Hill's oak.

Genetic variance among populations within Hill's oak. Our third major finding is that there is significant population genetic structure in Hill's oak, and that genetic structure has a geographic component. Our previous work suggested that a four-population model fits our three-species dataset significantly better than a three-population model, and that the fourth population which we inferred captured geographic variation primarily within Hill's oak ([8], p. 152 and Figure 7B). In the current study, we similarly find that a seven-population model best fits our four-species dataset (comprising red, black, Hill's, and scarlet oak), and that four of the seven inferred populations are predominantly found within our concept of Hill's oak (data not shown). The population genetic structure suggests a geographic differentiation within Hill's oak, with many populations in Michigan and northwestern Indiana separating from Wisconsin and northeastern Illinois. This geographic structure has moderate statistical support based on correlations of geographic distances and genetic distances between populations ($r = 0.0916$, 1-tailed $P = 0.0535$ based on 2,000 Mantel permutations). This geographic genetic structure is particularly remarkable given that the maximum geographic distance between the Hill's oak populations that we sampled is just over 550 km, whereas the geographic distance between the most distant black oak populations sampled is over 1100 km. No geographic signal was detected in black oak ($r = 0.0783$, 1-tailed $P = 0.2409$ based on 2,000 Mantel permutations).

Conclusions

The work reported in this paper contributes to our understanding of the taxonomy and evolution of Hill's oak, black oak, and scarlet oak in three ways. First, it demonstrates that Hill's oak and scarlet oak are sister species (Figure 2), which suggests that their morphological similarity is a consequence of shared ancestry, not convergence or gene flow. Variance in the morphology of both species, such that they overlap morphologically, may be reflected in their population genetic structure. Strong genetic disjunction between the two is consistent with recognition at the species level. Second, it demonstrates that there is substantial geographically-structured genetic variation within Hill's oak that is not present in the other two species. This may reflect the biogeography of Hill's oak, which is essentially wrapped around Lake Michigan (Figure 1); differential rates or patterns of gene flow among Hill's oak, black oak, and red oak in different regions; or a higher rate of population differentiation within Hill's oak for other causes. Finally, the paper provides strong evidence for introgression

between Hill's oak and black oak based on morphological data, and preliminary morphological and molecular evidence for ongoing local gene flow between the species. However, evidence for the latter is preliminary and bears further study.

Acknowledgements

Funding for this study came from The American Philosophical Society, the Hanes Fund of the Michigan Botanical Club, and The Morton Arboretum. Conversations with M. Bowles, R. Jensen, A. Reznicek, and D. Shepard have yielded insights into this work, and assistance in numerous aspects of field work and herbarium curation were provided by J. Sturner. Additional assistance in the field was provided by L. Anderson, D. Bassett, K. Bolger, G. Fewless, M. Bowles, R. Ewert, A. Gapinski, J. Hitz, C. Kirschbaum, D. Ladd, J. Mendelson, M. Nee, B. Olker, J. Wibbenmeyer, J. Yunger, the Forest Preserve Districts of Cook, Lake, DuPage Counties, Illinois, Albany Pine Bush Preserve, Indiana DNR, and numerous others. Allan Taylor assisted with review of the manuscript. All AFLP reactions were run out on the automated sequencer of the Field Museum's Pritzker Laboratory with assistance from K. Feldheim and E. Sackett-Hermann.

References cited

1. Grant, V.E. 1981. *Plant Speciation*. New York: Columbia University Press.
2. Coyne, J.A. and H.A. Orr. 2004. *Speciation*. Sunderland, MA: Sinauer Associates.
3. Hardin, J.W. 1975. *Hybridization and introgression in Quercus alba*. Journal of the Arnold Arboretum 56:336-363.
4. Whittemore, A.T. and B.A. Schaal. 1991. *Interspecific gene flow in sympatric oaks*. Proceedings of the National Academy of Sciences of the United States of America 88:2540-2544.
5. Curtu, A., O. Gailing, and R. Finkeldey. 2007. *Evidence for hybridization and introgression within a species-rich oak (Quercus spp.) community*. BMC Evolutionary Biology 7(1):218.
6. Dumolin-Lapegue, S., et al. 1997. *Phylogeographic structure of white oaks throughout the European continent*. Genetics 146:1475-1487.
7. Muir, G., C.C. Fleming, and C. Schlötterer. 2000. *Species status of hybridizing oaks*. Nature (London) 405:1016.
8. Hipp, A.L. and J.A. Weber. 2008. *Taxonomy of Hill's Oak (Quercus ellipsoidalis: Fagaceae): Evidence from AFLP Data*. Systematic Botany 33:148-158.
9. Dodd, R.S. and N. Kashani. 2003. *Molecular differentiation and diversity among the California red oaks (Fagaceae; Quercus section Lobatae)*. Theoretical and Applied Genetics 107(5):884-892.
10. Wadmond, S.C. 1933. *The Quercus ellipsoidalis-Quercus coccinea complex*. Transactions of the Wisconsin Academy of Sciences, Arts and Letters 28:197-203.

11. Overlease, W.R. 1977. *A study of the relationship between Scarlet Oak (Quercus coccinea Muenchh.) and Hill Oak (Quercus ellipsoidalis E.J. Hill) in Michigan and nearby states.* Journal of the Pennsylvania Academy of Science 51:47-50.
12. Overlease, W.R. 1991. *Genetic relationships between three species of oaks as determined by common garden studies with populations from Michigan, Indiana and Wisconsin.* Journal of the Pennsylvania Academy of Science 65(2):71-74.
13. Jensen, R.J., R. Depiero, and B.K. Smith. 1984. *Vegetative characters, population variation, and the hybrid origin of Quercus ellipsoidalis.* American Midland Naturalist 111(2):364-370.
14. Shepard, D.A. 1993. *The Legitimacy of Quercus ellipsoidalis Based on a Populational Study of Quercus coccinea in Illinois.* P. 168 in *Biological Sciences Department, Western Illinois University, Macomb, IL.*
15. Shepard, D.A. 2009. *A review of the taxonomic status of Quercus ellipsoidalis and Quercus coccinea in the Eastern United States.* International Oak Journal 20:65-84.
16. Trelease, W. 1919. *The Jack Oak (Quercus ellipsoidalis).* Transactions of the Illinois State Academy of Science 12:108-118.
17. Jensen, R.J. 1977. *Numerical analysis of the scarlet oak complex (Quercus subgen. Erythrobalanus) in the eastern United states: relationships above the species level.* Systematic Botany 2:122-133.
18. Jensen, R.J. 1997. *Quercus Linnaeus sect. Lobatae Loudon. Flora of North America North of Mexico* 3:447-468. New York: Oxford University Press.
19. Jensen, R.J. 1977. *A preliminary numerical analysis of the red oak complex in Michigan and Wisconsin USA.* Taxon 26(4):399-407.
20. Shepard, D.A. 1991. *The occurrence of the scarlet oak, Quercus coccinea in northern Illinois.* Transactions of the Illinois State Academy of Science 84, Suppl. 35, Abstract.
21. Overlease, W.R. 1964. Page 71 in *A study of Variation in the Red Oak Group in Michigan and Nearby States.* East Lansing: Michigan State University.
22. Muir, G. and C. Schlötterer. 2005. *Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (Quercus spp.).* Molecular Ecology 14(2):549-561.
23. Lexer, C., A. Kremer, and R.J. Petit. 2006. *Shared alleles in sympatric oaks: recurrent gene flow is a more parsimonious explanation than ancestral polymorphism.* Molecular Ecology 15:2007-2012.
24. Vos, P., et al. 1995. *AFLP: A new technique for DNA fingerprinting.* Nucleic Acids Research, 23:4407-4414.
25. Cavender-Bares, J. and A. Pahlich. 2009. *Molecular, morphological, and ecological niche differentiation of sympatric sister oak species, Quercus virginiana and Q. geminata (Fagaceae).* American Journal of Botany 96(9):1690-1702.

26. Biagi, A. and R.J. Jensen. 1995. *The Genus Quercus (Fagaceae) in Indiana: Phytogeography and a Key to the Species*. Indiana Academy of Science 104:11-24.
27. Lynch, M. and B.G. Milligan. 1994. *Analysis of population genetic structure with RAPD markers*. Molecular Ecology 3:91-99.
28. Vekemans, X. 2002. *AFLP-SURV, version 1.0*. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles (Belgium).
29. Desper, R. and O. Gascuel. 2002. *Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle*. Journal of Computational Biology 9(5):687-705.
30. Paradis, E., J. Claude, and K. Strimmer. 2004. *APE: Analyses of Phylogenetics and Evolution in R language*. Bioinformatics 20:289-290.