

The Maintenance of Species Boundaries of an Endangered Rare Species, Maple-leaf Oak (*Quercus acerifolia*), under Selection and Gene Flow

Introduction: Interbreeding with close relatives can obscure the boundaries of rare and endemic species, and increase their extinction risk [1-5]. The chances of extinction for rare species can be as high as 90% when pure rare species are replaced by fertile and viable hybrids [1-2,4,6]. However, gene flow between parental species might be suppressed by strong selection against hybrids [7-9]. Such divergent selection can limit gene flow by filtering out early-generation hybrids through hybrid disadvantage or competitive exclusion [10-11]. Thus, strong ecological selection can counter gene flow between close relatives and maintain the species boundaries of rare species. **This project will study the effects of these two counteracting forces, selection and gene flow, in maintaining the integrity of a rare outcrossing species. By combining both morphological and genetic data, this study will establish an exemplar protocol for delineating rare species, which is crucial for rare species conservation [12].**

I will study the rare and endangered Maple-leaf Oak (*Quercus acerifolia*), and the two close relatives with which it hybridizes. *Q. acerifolia* is an endangered species, considered to be restricted to only four mountain ridges in Arkansas [13-14] (Fig. 1), although recent surveys have found additional locations (see “Methods”). Experienced naturalists have long suspected that *Q. acerifolia* hybridizes with its sister species, *Q. shumardii* (Shumard Oak), and their close relative, *Q. rubra* (Northern Red Oak) (Fig. 1). Despite its rarity and potential for hybridization, the processes maintaining species boundaries in *Q. acerifolia* are unknown.

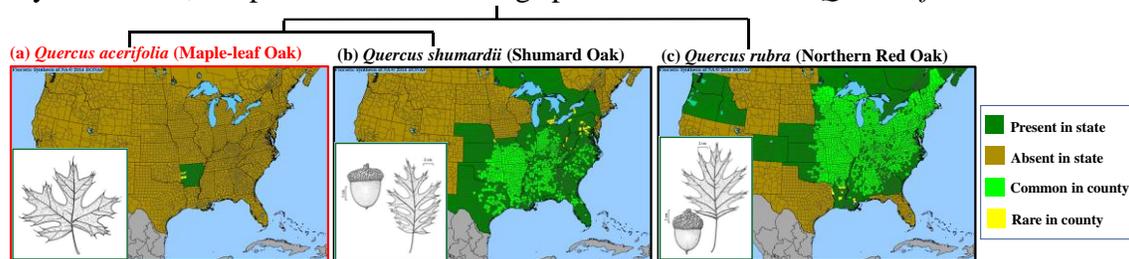


Fig. 1 Distribution [14], morphology [16], and phylogeny [17] of study species. Note that map is outdated for Maple-leaf Oak.

Limited investigations of Maple-leaf Oak suggest that both selection and gene flow may influence the populations of this endangered species. *Q. acerifolia* is phenotypically distinct from its close relatives [15], which has been attributed to divergent selection for soil type [16]: *Q. acerifolia* is restricted to dry sandstone or novaculite soil at high elevations, while its sister species, *Q. shumardii*, inhabits moist valleys, and *Q. rubra* prefers mesic soil. Molecular analysis suggests potential gene flow between *Q. acerifolia* and sympatric populations of *Q. shumardii* [17]. Nevertheless, limited sampling has prevented a comprehensive understanding of the relative importance of gene flow and selection in shaping the species boundary of *Q. acerifolia*.

Recent genomic studies suggest that moderate gene flow is common in oaks, yet divergence occurs [18]. Oaks are known for frequent hybridization, yet oak species maintain morphological distinctiveness (*i.e.*, “oak syngameon”) [8,18-19]. Genetic evidence suggests that ecological selection is maintaining the boundaries of oak species: genes differentiated among closely related species are clustered within the genome [20] forming “genomic islands” that contain important genes relevant to adaptation to specific environmental conditions.

However, other genomic regions are homogenized by gene flow and do not differ among species. Thus, genomic islands are a good indicator of ecological divergence despite gene flow. With the recent availability of an annotated oak genome to identify functional genes [21], it is possible now to directly investigate this process of “divergence-despite-gene-flow” in *Q. acerifolia*. As a critical part of my Ph.D. thesis on rare species evolution, this project will combine both morphological and genetic evidence to test the following hypotheses.

Hypotheses: **H1 - Morphological distinctiveness:** *Q. acerifolia* and its two closest relatives exhibit morphological distinctiveness from each other despite ongoing hybridizations. **H2 - Strong gene flow:** *Q. acerifolia* and its two closest relatives exhibit substantial gene flow in much of their genomes. **H3 - Ecological selection under gene flow:** *Q. acerifolia* and its two closest relatives show differentiation in genes associated with soil adaptation, and exhibit “genomic islands” of differentiation despite gene flow.

Preliminary results: 1) With the collection of mainly leaf samples in 2019 (Table 1), I found that *Q. acerifolia* does exhibit an overall morphological differentiation from its close relatives, although 25% of *Q. acerifolia* (defined by current taxonomy) samples fall in two morphological groups representing *Q. shumardii* and *Q. rubra*. However, a lack of acorn samples prevented me from incorporating important reproductive traits to test

H1. Thu. I will need to expand my sampling in 2020 to include acorn samples and more sites in order to fully cover the purported species range. **2)** By mapping *PstI* RAD-seq (restriction site-associated DNA sequencing) loci of one *Q. shumardii* individual [17] to the latest annotated haploid genome of *Q. robur* (Pedunculate Oak) [21], I found 84,540 successful alignments distributed evenly across the genome, suggesting that RAD-seq mapping in oaks will provide insights into differentiated genes and ecological selection.

Methods and Expectations:

Sampling: Building upon my sampling in 2019, I will collect leaf and acorn samples of *Q. acerifolia* and its two close relatives from eight sites in Arkansas (AR) and one site in Alabama (AL) in 2020 (see shaded rows in Table 1). For each tree, I will collect five to ten mature leaves and mature acorns. This study will use multiple sites with sympatric populations as “independent experiments”, including four historic sites in AR [13] (no. 1-4 in Table 1), one newly discovered site in AR [22] (no. 5), three new sites in Tennessee (TN) [23] (no. 6-8), one new site in AL (no. 11; pers. comm. with Dr. Richard Condit). This broad collection will enable me to test the

Table 1 Collection sites of *Quercus acerifolia* and its two close relatives. Numbers before the parentheses are the sample size for the measurement of morphological traits, while the numbers in the parentheses are the sample size for genetic analysis. “NA” means that the species is absent at the site. Collection tasks in 2020 are highlighted in grey background color. In summary, I will have more than 500 samples for morphological analysis and 95 DNA samples by the end of 2020 field season.

No	Site	State	Collection status	Sample size		
				<i>Quercus acerifolia</i>	<i>Quercus shumardii</i>	<i>Quercus rubra</i>
1	Magazine Mountain	AR	Leaf and acorn samples needed	36 (4)	NA	16 (4)
2	Porter Mountain	AR	Acorn samples needed	32 (4)	NA	16 (4)
3	Pryor Mountain	AR	Acorn samples needed	30 (4)	NA	30 (4)
4	Sugarloaf Mountain	AR	Acorn samples needed	42 (4)	NA	18 (4)
5	Kessler Mountain	AR	Acorn samples needed	8 (4)	15 (4)	15 (4)
6	Duck River Complex	TN	Completed in 2019	4 (4)	24 (4)	NA
7	Cumberland Plateau	TN	Completed in 2019	6 (4)	25 (4)	NA
8	Cedar Glade Natural Area	TN	Completed in 2019	1 (1)	21 (4)	NA
9	Tyson Research Center	MO	Completed in 2019	NA	NA	29 (4)
10	Shaw Nature Reserve	MO	Completed in 2019	NA	30 (4)	30 (4)
11	Paint Rock	AL	Leaf and acorn samples needed	20 (4)	30 (4)	NA
TBD	Three sites of <i>Q. shumardii</i> (TBD)	AR	Leaf and acorn samples needed	NA	60 (10)	NA
Total				179 (33)	205 (34)	154 (28)

generality of ecological selection and gene flow at multiple locations, ensuring a stringent evaluation of the hypotheses. In addition, samples of the two close relatives have been collected from two sites in Missouri (MO) (no. 9-10), which are allopatric to *Q. acerifolia*. I will also collect samples of *Q. shumardii* from three sites in AR that are adjacent to the known distributions of *Q. acerifolia*. By the end of 2020, I will have > 500 samples for morphological analyses and 95 samples for genetic analyses. **Morphology (H1):** I will measure leaf, acorn, and bud traits (e.g., lobe length, abaxial leaf domatia, acorn length, bud pubescence) that differ to some degree among the study species according to previous research [15,24]. Normal mixture models (NMMs) will be used to examine the clustering pattern in morphological space without *a priori* species assignment [25]. I expect NMMs to identify three morphological clusters corresponding to species taxa. **Gene flow (H2):** I will use RAD-seq, a well-established and effective genomic approach for studying genetics in oaks [17,26-28]. Admixture will be identified using the STRUCTURE analysis [29], and the extent of gene flow will be estimated using the *D*-statistics [30]. I expect to detect high admixture and substantial gene flow among species. **Ecological selection under gene flow (H3):** First, RAD-seq loci will be mapped to the latest annotated *Q. robur* genome [21] using *Magic-BLAST* [31-32]. I will identify genes under divergent selection using F_{st} values, Q_{st} - F_{st} comparison [33], and the McDonald-Kreitman test [34]. The functions of divergent genes will be inferred from genome annotations. I expect that genes associated with adaptation to soil properties are under divergent selection. I will also test spatial overdispersion of divergent genes within the genome, and I expect differentiated genes to cluster spatially within the genome, forming “genomic islands”.

Scientific merit: This study will become a model for rare species delimitation in the field of plant taxonomy. Accurate delineation of rare species is crucial to their conservation [35]; however, species hybridization, which occurs frequently in plants [36], can obscure the boundary between a rarer species and its common close relatives. This study is designed in response to the urgent call for synthesizing morphological and genomic data to define species taxa in assisting rare species conservation [12,37].

Broader impact: 1) This project will train undergraduates and high-school students in plant taxonomic research. Specifically, I will collaborate with the [Missouri Botanical Garden's NSF Research Experiences for Undergraduates program](#) to give one to two undergraduate students experience with field trip planning and logistics, species identification, and herbarium vouchering. I will also mentor high-school students of minority

groups in species identification and morphological analysis through the [Shaw Institute for Field Training Program](#). 2) This study will also produce findings that are informative and applicable to landowners and park managers, who share common interests in tree conservation. This project has already involved collaboration with staff and scientists from the Missouri Botanical Garden, the Morton Arboretum (Chicago), the Arnold Arboretum (Boston), and public and individual landowners in the Midwest who assisted my field collections. These organizations and individuals have long participated in the conservation of this putative endangered species; however, their work is limited by the unclear species status and under-studied evolutionary history of *Q. acerifolia*. The results of this study will inform these stakeholders about the conservation of this rare species, while also providing a framework for further delimitation and conservation of other rare species.

In summary, funding from the International Association for Plant Taxonomy will greatly facilitate this project by providing research opportunities for students and helping landowners as well as park managers to make informed decisions concerning rare species conservation.

Literature cited

- [1] KS Burgess et al., *New Phytologist* **177** (1), 276 (2008).
- [2] M Todesco et al., *Evolutionary Applications* **9** (7), 892 (2016).
- [3] JM Rhymer and D Simberloff, *Annual Review of Ecology and Systematics* **27** (1), 83 (1996).
- [4] O Lepais et al., *Molecular Ecology* **18** (10), 2228 (2009).
- [5] O Seehausen et al., *Molecular Ecology* **17** (1), 30 (2008).
- [6] L Lagache et al., *Molecular Ecology* **22** (2), 423 (2013).
- [7] S Via, *Philosophical Transactions of the Royal Society B: Biological Sciences* **367** (1587), 451 (2012).
- [8] J Cavender-Bares and A Pahlich, *American Journal of Botany* **96** (9), 1690 (2009).
- [9] J Ortego et al., *Journal of Biogeography* **41** (10), 1895 (2014).
- [10] JW Hardin, *Journal of the Arnold Arboretum* **56** (3), 336 (1975).
- [11] AR Templeton, in *The Units of Evolution: Essays on the Nature of Species*, edited by Marc Ereshefsky (The MIT Press, 1989), pp. 159.
- [12] J Johnson et al., *Systematic Biology* **53** (6), 841 (2004).
- [13] K Wenzell, Kenny, L. & Beckman, E, *The IUCN Red List of Threatened Species 2016*: e.T35039A2858212. (2016).
- [14] J Kartesz, *North American Plant Atlas - The Biota of North America Program (BONAP)* (2015).
- [15] WJ Hess and NA Stoyloff, *Systematic Botany*, 89 (1998).
- [16] JD Stein, *Field guide to native oak species of eastern North America*. (DIANE Publishing, 2003).
- [17] AL Hipp et al., *New Phytologist* **217** (1), 439 (2018).
- [18] AL Hipp et al., *Annals of the Missouri Botanical Garden* **104** (3), 455 (2019).
- [19] CH Cannon and RJ Petit, *New Phytologist* (2019).
- [20] C Scotti-Saintagne et al., *Genetics* **168** (3), 1615 (2004).
- [21] C Plomion et al., *Nature Plants* **4** (7), 440 (2018).
- [22] T Witsell, *A rapid terrestrial ecological assessment of Kessler Mountain Reserve* (2015).
- [23] AN Mummaw, *Master Thesis at Austin Peay State University* (2018).
- [24] N Stoyloff and WJ Hess, *SIDA, Contributions to Botany*, 267 (1990).
- [25] CD Cadena et al., *Systematic Biology* **67** (2), 181 (2018).
- [26] J Ortego et al., *New Phytologist* **218** (2), 804 (2018).
- [27] E Fitzek et al., *Ecology and Evolution* **8** (11), 5837 (2018).
- [28] XL Jiang et al., *Molecular Phylogenetics and Evolution* **132**, 36 (2019).
- [29] L Porrás-Hurtado et al., *Frontiers in Genetics* **4**, 98 (2013).
- [30] Y Zheng and A Janke, *BMC Bioinformatics* **19** (1), 10 (2018).
- [31] GM Boratyn et al., *BMC Bioinformatics* **20** (1), 405 (2019).
- [32] AL Hipp et al., *New Phytologist* (**Early View**) (2019).
- [33] T Leinonen et al., *Nature* **14** (3), 179 (2013).
- [34] JH McDonald and M Kreitman, *Nature* **351** (6328), 652 (1991).
- [35] JJ Wiens, *Systematic Biology* **56** (6), 875 (2007).
- [36] KD Whitney et al., *Perspectives in Plant Ecology, Evolution and Systematics* **12** (3), 175 (2010).
- [37] GS Lim et al., *Systematic Biology* **61** (1), 165 (2012).