# Phylogeographic patterns and migration history of Garry oak (*Quercus garryana*) in western North America

by

Rande Kanne
B.Sc., University of Victoria, 2014

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of a

MASTER OF SCIENCE

in the Department of Biology

# Supervisory Committee

Phylogeographic patterns and migration history of Garry oak (*Quercus garryana*) in western North America

by

Rande Kanne
B.Sc., University of Victoria, 2014

**Supervisory Committee**

Dr. Geraldine A. Allen, Department of Biology
**Co-Supervisor**

Dr. Barbara J. Hawkins, Department of Biology
**Co-Supervisor**

Dr. Terri Lacourse, Department of Biology
**Departmental Member**

# Abstract

Garry oak (*Quercus garryana* Douglas ex. Hook) is a white oak (*Quercus* sect. *Quercus*) with a geographic range extending from southwestern BC to south-central California. It is the only native white oak in BC and Washington, and is the northernmost species of the California Floristic Province-Pacific Northwest white oak clade. I used molecular methods to address the following questions: 1) What are the patterns of genetic variation within Garry oak? 2) How do these patterns vary geographically, and how did the spatial distribution of the gene lineages come to occupy its current geographical range? 3) Does Garry oak show evidence of genetic interaction with other white oak species in western North America? 4) Is there morphological or genetic evidence to support the three described varieties of Garry oak?

I obtained samples of Garry oak from 117 localities over its geographic range, as well as samples of two other California white oaks (*Q. lobata* and *Q. douglasii)* and a Rocky Mountain species (*Q. gambelii*). Analyses of DNA sequence data from four plastid DNA regions revealed 24 distinct molecular variants (haplotypes) in Garry oak. These show a strong south-to-north decrease in genetic diversity, consistent with post-glacial northward expansion. Haplotypes present in the northern part of the range provide evidence of two separate northward migrations, only one of which reached the northern range limit of Garry oak in BC. I found that Garry oak shared plastid DNA haplotypes with two other white oak species, indicating that it hybridizes with other oaks in the southern part of its range. The nuclear ribosomal ITS phylogeny showed poor resolution, but both cpDNA and nrDNA may indicate that *Q. garryana* is more closely related to the white oaks of central North America than was previously thought. My findings also suggest that the three currently recognized varieties of Garry oak (var. *garryana*, *breweri* and *semota*) are not well differentiated genetically, but show morphological variation at the regional level. This study shows the phylogeographic patterns within *Q. garryana*. In addition, it contributes to conservation efforts in Garry oak ecosystems by indicating regions of high genetic diversity in Garry oak, including genetically unique populations that may be especially worthy of preservation.

# Table of Contents

**Supervisory Committee** ii
**Abstract** iii
**Table of Contents** iv
**List of Tables** v
**List of Figures** vi
**List of Appendices** viii
**Acknowledgments** ix
**Dedication** x
**Chapter 1.0: Introduction** 1
1.1 Biogeography and molecular evolution 1
1.2 Landscape history of western North America and its implications for plant geographic patterns 2
1.3 An overview of *Quercus* 4
1.4 Garry oak (*Quercus garryana*) 8
1.5 Research objectives 9
**Chapter 2.0: Materials and Methods** 11
2.1 Site selection and sampling 11
2.2 DNA extraction protocol 11
2.3 PCR protocols and DNA sequencing 13
2.4 Sequence assembly and alignment 16
2.5 Analyses of DNA sequence data 16
2.6 Geographic patterns and mapping 18
2.7 Assessment of morphological variation 19
**Chapter 3.0: Results** 21
3.1 Characteristics of the plastid and nuclear DNA regions 21
3.2 Plastid haplotype diversity and distribution of *Quercus garryana* 21
3.3 Plastid haplotype distributions in all sampled *Quercus* species 25
3.4 Phylogenetic analyses of *Quercus garryana* and outgroups 27
3.4.1 Phylogenetic analyses of plastid DNA sequences 27
3.4.2 Phylogenetic analyses of nuclear DNA sequences 29
3.4.3 Correspondence between the plastid and nuclear phylogenetic trees 32
3.5 Morphological variation among Garry oak varieties 33
3.5.1 Analysis of morphological characters from voucher specimens 33
3.5.2 Geographic patterns of morphological variation 36
3.5.3 Relationships between morphological and molecular variation 38
**Chapter 4.0: Discussion** 40
4.1 Geographic history of *Quercus garryana* 41
4.2 Introgression and hybridization between sympatric white oaks 44
4.3 Phylogenetic relationships of the California clade to other white oaks 45
4.4 Identification of Garry oak varieties 47
**Chapter 5.0: Conclusions** 50
**Bibliography** 51
**Appendices** 64

# List of Tables

Table 1. Current subgenera and sections of *Quercus* according to Hipp *et al*., 2018..... 6
Table 2. Phylogenetic groups within *Quercus* section *Quercus*. ..... 7
Table 3. Plastid regions amplified to test quality of sequence traces and quantity of variable characters. ..... 14
Table 4. Primer pairs tested for the internal transcribed spacer (ITS) region of the nuclear ribosomal gene complex. ..... 15
Table 5. Characters used for principal component analysis (PCA) of *Q. garryana* using PC-ORD. ..... 19
Table 6. Diversity and distribution of *Quercus* cpDNA haplotypes (28 haplotypes in total) found during this study. ..... 22
Table 7. Diversity of *Quercus* (section *Quercus*) haplotypes by region..... 24
Table 8. Categorical values for morphological characters used in identification of *Q. garryana* and varieties. ..... 34
Table 9. Quantitative characters used form morphological identification of *Q. garryana* and varieties. ..... 35
Table 10. PCA axes 1 and axes 2 r-values, tau-values and eigenvectors for seven morphological characters used in morphometric analysis of *Q. garryana*. ..... 35

# List of Figures

Figure 1. (a) Geographic distribution of *Q. garryana* with sampled localities shown (blue dots). (b) Localities (including *Q. garryana*) and ranges indicated for the other sampled *Quercus* species: *Q. lobata*, *Q. douglasii*, and *Q. gambelii* (Little, 1971). ..... 12
Figure 2. Statistical parsimony network showing relationships of sampled *Q. garryana* cpDNA haplotypes. Notches indicate mutations between haplotypes..... 21
Figure 3. Haplotype distribution and regions of diversity for 24 *Q. garryana* cpDNA haplotypes. Pie charts depict haplotype diversity in four geographical regions: BC/WA, OR, n CA, and s CA..... 23
Figure 4. Statistical parsimony network of *Q. garryana* haplotypes showing geographic distributions of haplotypes. ..... 24
Figure 5. SplitsTree network showing reticulate relationships of the 28 *Quercus* haplotypes found for all species sampled in this study (including the 24 haplotypes found in *Q. garryana*). ..... 25
Figure 6. Statistical parsimony network of 125 *Quercus* plastid sequences revealed 28 haplotypes. Twenty-one haplotypes were found only in *Q. garryana*, three are common to *Q. garryana* and *Q. lobata*, and one was common to *Q. lobata* and *Q. douglasii*. ..... 26
Figure 7. *Quercus garryana* and other sampled *Quercus* haplotype distributions and haplotype network of Group 1. (a) Geographic distributions of haplotypes. (b) Numbers of total and unique haplotypes. (c) Haplotype network. ..... 26
Figure 8. *Quercus garryana* and other sampled *Quercus* haplotype distributions and haplotype network of Group 2. (a) Geographic distributions of haplotypes. (b) Number of total and unique haplotypes. (c) Haplotype network. ..... 27
Figure 9. Phylogenetic trees for *Q. garryana* and related species haplotypes based on cpDNA sequences. (a) Bayesian analysis with posterior probabilities indicated at nodes. (b) Maximum Likelihood with bootstrap support indicated at nodes and line thickness indicates bootstrap support threshold above (thick lines) or below (thin lines) 50. ..... 28
Figure 10. Phylogenies of sampled *Quercus* and outgroups based on the nuclear ITS region; (a) Bayesian tree with posterior probabilities indicated at nodes; (b) PhyML Maximum Likelihood tree with clade bootstrap support indicated at nodes. ..... 30
Figure 11. Phylogenies of sampled *Quercus* and outgroups based on the nuclear ITS region; (a) Bayesian tree with posterior probabilities indicated at nodes; (b) PhyML Maximum Likelihood tree with clade bootstrap support indicated at nodes. ..... 31
Figure 12. Tanglegram showing correspondence between nuclear (ITS) and plastid Bayesian phylogenetic trees..... 32
Figure 13. PCA of *Q. garryana* based on seven morphological characters that have been used to distinguish the varieties. The first two axes capture 61.7 % of total variance. *Quercus garryana* var. *garryana* and var. *semota* were identifiable using these morphological characters; var. *breweri* could not be separated from the other varieties. 33
Figure 14. PCA plots for *Q. garryana*, indicating variation in individual morphological variables. Symbol size corresponds to the value (class or quantitative) of the morphological character. Axes for all plots correspond to those in Figure 13. ..... 34
Figure 15. PCA of *Q. garryana* based on seven morphological characters, with samples colour-coded by ecoregion region. ..... 36

Figure 16. Latitude of sampled *Q. garryana* vs. PCA axis 1 scores..... 37
Figure 17. Geographic variation in qualitative and quantitative morphological characters used to identify *Q. garryana* to variety: a) bud shape, b) terminal bud colour, c) terminal bud length, d) bud hair density, e) twig hair density, f) mean number of stellate rays, and g) mean length of stellate rays. ..... 38
Figure 18. PCA with plastid haplotype subclades (cluster 1 and cluster 2) and two divergent haplotypes indicated. Subclades correspond to Figure 9a and b. ..... 39

# List of Appendices

Appendix A. Sample localities. Locality information (site number, latitude, longitude and elevation), locality code, and species identification of 125 sampled *Quercus* individuals ..... 64
Appendix B. Plastid primers. Primers used in this study for cpDNA PCR amplification of *Quercus* species ..... 72
Appendix C. Nuclear ITS primers. Primers used in this study for nrDNA PCR amplification of *Quercus* species ..... 73
Appendix D. Fagaceae species retrieved from GenBank and used in the plastid and nuclear ITS phylogenetic analyses ..... 74

# Acknowledgments

There were many minds and hands contributing to this research project, and I want to acknowledge and thank all involved. First and foremost, I want to thank my supervisor, Dr. Geraldine Allen, for her many hours of time and support, her knowledge and expertise, and her willingness to help me. Her professional input and advice were invaluable. Thank you for your patience, time, and adventures in the field. Thank you to our laboratory technician, Laurie McCormick, who taught me all that I needed to know in the lab. Her assistance was vital to me and this project. Thank you to my co-supervisor, Dr. Barbara Hawkins, for her support and making this opportunity a reality for me. Thank you to my committee member, Dr. Terri Lacourse, for her encouragement and questions during committee meetings that guided my understanding of ordinations. Thank you, Dr. Ken Marr, for collecting the *Quercus gambelii* sample while visiting Colorado. Thank you to everyone else who contributed.

# Dedication

This thesis is dedicated to my husband, Charles Noseworthy, and my mom, Reayane Kanne, for their constant support and words of encouragement. Thank you both so much! I would also like to dedicate this thesis to my two cats, Pepe and Rumble. They were my constant companions during the enumerable hours of analyses and writing!

# Chapter 1.0 Introduction

## 1.1 Biogeography and molecular evolution

For over two centuries, the geographic distributions of organisms have been of interest to scientists. In the 18th century, Buffon and Daubenton (1749-1767) observed that regions that are ecologically similar but geographically separated have their own distinct species assemblages, a generalization that came to be known as Buffon's Law. In the 19th century, Darwin's (1839) observations of geographic patterns formed an essential underpinning for his theory of evolution by natural selection. Currently, the field of biogeography addresses the distributions of organisms and the processes involved in shaping patterns of spatial distributions (Rosen, 1978; Platnick and Nelson, 1984; Lomolino *et al*., 2010). It is an interdisciplinary field, including elements of molecular biology, ecology and historical geology. Biogeography may be subdivided into two areas of study, ecological and historical biogeography. The time scale is the defining difference. Ecological biogeography considers the distributions of living species as a function of their environment, whereas historical biogeography examines geological events and phylogenetic relationships over long time scales (Posadas *et al*., 2006).

One widely used approach to historical biogeography is phylogeography (Posadas *et al*., 2006). The term phylogeography, first introduced by Avise *et al*. (1987), describes the investigation of historical biogeographic processes using information from the phylogeny and spatial distributions of gene lineages. It examines processes of dispersal, vicariance and extinction, in some combination, at the inter- or intra-specific gene level (Avise, 1987, 2007). It combines the microevolutionary perspective of population genetics with the macroevolutionary perspective of phylogenetics (Avise, 2007; Avise *et al*., 2016; Cutter, 2013), typically using DNA sequence data to examine the evolutionary history of species. Allele or haplotype trees are used to infer relationships, the units of interest being the variable haplotypes that are identified (Castelloe and Templeton, 1994). DNA polymorphisms are used to assess the genetic variation within species and populations (Julio, 2009; Pleines *et al*., 2009). Information derived from DNA polymorphisms provides evolutionary insight into the demographic factors affecting species and populations (Julio, 2009). Inferences can then be drawn about biodiversity

(Gibson *et al.*, 2015; Pei *et al*., 2017), genealogical histories (Muller *et al*., 2005; Wu, 2011), and in some cases natural selection (da Silva, 2017; Hiroshi, 1999). Geographic regions of high genetic diversity are often older populations, but can also be the result of migration from multiple source populations. Landscape genetic patterns can then be used to infer routes of historical migrations.

Spatial distributions of species or populations are largely influenced by migration across the landscape, which is typically associated with decreasing genetic diversity caused by founder effects or bottleneck events (Taberlet *et al*., 1998; Knowles and Maddison, 2002). An understanding of population genetic processes can be combined with phylogenetic analyses to infer the direction of gene flow across landscapes (Marsico *et al*., 2009; Nevill *et al*., 2010). Refugial areas or core populations are expected to have higher genetic diversity than recently colonized areas or peripheral populations (Marsico *et al*., 2009; Nevill *et al*., 2010). These characteristic patterns of diversity reflect influences by a variety of climatic and environmental factors, including glacial and interglacial cycles. To comprehend spatial and temporal distributions, it is crucial to understand historic geological events.

## 1.2 Landscape history of western North America and its implications for plant geographic patterns

During the Cenozoic, terrestrial landscapes in North America (NA) became gradually cooler and drier. Glaciation began in the Miocene, starting in Alaska, Greenland, and Iceland (Ehlers and Gibbard, 2007) and continued into the Pliocene, with continental ice sheets reaching sea levels by the early Pleistocene (Ehlers and Gibbard, 2007). The Pleistocene epoch (*c*. 2.6 Ma to 11.7 ka) was characterized by cyclical climatic fluctuations, with long glacial periods interspersed with short interglacial phases (Ehlers and Gibbard, 2007; de Boer *et al*., 2010). During the Last Glacial Maximum (LGM), 21-18 ka, the Cordilleran and Laurentide ice sheets covered approximately the north half of NA (Ehlers and Gibbard, 2007; Izumi and Bartlein, 2016). The Cordilleran ice sheet in western NA was centred on what is now British Columbia (BC), extending north into present-day Yukon and south into northwestern Washington (Clark and Mix, 2002). During the LGM, ice and permafrost dominated the terrestrial landscape of

northern NA, with marginal ice-sheets extending over small areas of the Pacific coast (Clark and Mix, 2002). At 19 ka BP, the sea level was 118 to 130 m below present-day levels (Bard *et al*., 1990; Clark and Mix, 2002; Yokoyama *et al*., 2018).

Continental ice sheets, together with colder conditions over most of NA (Boer *et al*. 2010), led to range shifts or range reductions for many boreal and temperate plant species. Some of these persisted in southern refugia, and others in Beringia (Alaska, northwestern Yukon and land exposed in the Bering Strait). A few may have persisted in cryptic refugia surrounded by ice or on exposed coastal plains along the north Pacific coast (Shafer *et al*., 2010; Beatty and Provan, 2011; Allen *et al*., 2012; Guest and Allen, 2014). Some North American plant species may have been relatively unaffected by glacial advancement, if populations were situated south of the ice sheets. Because of its glacial history and other factors, the diversity of plant species in NA (as elsewhere) is inversely related to latitude. It is highest in tropical forests, decreasing in temperate forests and further declining in boreal forests (Usinowicz *et al*., 2017). These patterns can be explained by responses to repeated glaciation and seasonal climatic conditions (Qian and Ricklefs, 2007; Usinowicz *et al*., 2017). Following the LGM, plant species migrated into newly favourable habitats in northern NA. The presence of northern refugial populations may have increased dispersal and recolonization rates (Snell and Cowling, 2015). Dispersal rates are affected by low temperatures, drought and short growing seasons, leading to longer generational intervals (Brochmann and Brysting, 2008; Rhode, 1992). Beyond environmental conditions, other factors influence dispersal rates, such as competition and life history traits. Many plant species situated south of the ice sheet during the LGM were likely inhibited by competition for limited seedling sites. Post-Pleistocene revegetation may have initially been a slow process, at least with regard to primary succession, but then recolonization may have become more rapid, and increasingly so in the presence of refugia (Snell and Cowling, 2015). From that standpoint, many highly dispersible and fast-growing species may have had early success in establishment. The slower growing species, such as the deciduous hardwoods, may have been out-competed by evergreen softwoods (Fuchs *et al*., 2000; Roy *et al*., 2000; Dunwiddie *et al*., 2011). *Quercus garryana* Dougl. ex Hook. is an example, with its northern range limited by the fast-growing *Pseudotsuga menziesii* (Mirb.) Franco

(Dunwiddie *et al*., 2011). The colonial front of *Q. garryana* may have been less restricted by temperature than by competition for light.

Plants from various taxonomic groups (angiosperms and gymnosperms) and life history strategies (herbaceous, perennial, shrub or tree) may have had similar historic migration patterns. Many plant species in NA exhibit a south to north decrease in genetic variability as a result of northward post-glacial recolonization, a pattern consistent with the leading-edge model of Hewitt (1993) in which rapid range expansion is accompanied by decreased genetic diversity resulting from genetic drift. Examples of this are *Cornus nuttallii* (Keir *et al*., 2011), *Pinus contorta* (Cwynar and MacDonald, 1987), *Salix melanopsis* (Brunsfeld *et al*., 2007), *Thuja plicata* (O'Connell *et al*., 2008), *Tiarella trifoliata* (Soltis *et al*., 1992), and *Tolmiea menziesii* (Soltis *et al*., 1989), and many other species of the Pacific Northwest. Beatty and Provan (2010) found that the perennial *Orthilia secunda*, a circumboreal species, showed high diversity in western Washington and Oregon (areas that provided a southern refuge for many temperate and a few boreal species) and lower diversity in its northern range. Beatty and Provan (2011) also investigated *Monotropa hypopitys*, an understory temperate species with a disjunct population east-west in NA. They concluded that at least two southern refugial sites in both eastern and western regions provided a haven during the LGM (Beatty and Provan, 2011). *Acer macrophyllum*, a deciduous hardwood tree with a broad latitudinal range in the Pacific Northwest, was found to have high genetic diversity in the southern populations (Iddrisu and Ritland, 2004). Plant species with more southern ranges experienced climatic shifts (*e.g. Quercus lobata*), undergoing elevational expansion and contraction in response to glacial cycles (Gugger *et al*., 2013). A diverse set of geological and environmental factors influenced many species of western NA. Phylogeographic inquiries into the complex history of previously glaciated landscapes can provide insight into the present-day distribution and abundance of arctic, boreal and temperate species.

### 1.3 An overview of *Quercus*

Oaks (*Quercus*) are woody plants belonging to the angiosperm family Fagaceae, which also includes chestnuts (*Castanea*) and beeches (*Fagus*). The family Fagaceae is distributed globally and contains nine genera and *c*. 900 species. The largest genus is

*Quercus*, comprising *c.* 500 species. It belongs to the subfamily Quercoideae, along with *Castanea*, *Castanopsis*, *Chrysolepis*, *Lithocarpus* and *Trigonobalanus*. Oaks occupy various climatic regions from tropical montane through Mediterranean to cool temperate. They are especially diverse in North and Central America, Eurasia, and northern Africa (Manos *et al*., 1999; Kremer *et al*., 2012; Barron *et al*., 2017).

The first fagaceous fossils (identifiable to family, though not to genus) were found in eastern NA, dating to the Late Cretaceous epoch (Herendeen *et al*., 1995; Sims *et al.*, 1999). The North American *Quercus* fossil record dates to the early Tertiary (Paleogene), 50-55 million years ago (Crepet and Nixon 1989a, b). Nixon (2006) suggested that *Quercus* spp. (predominantly evergreen oaks) historically dominated forests of the Northern Hemisphere. This is supported by pollen evidence found in western Greenland and the Baltic amber region of northern Europe dated to the middle Eocene (Grimsson *et al*., 2015). In NA, fossil pollen samples (McIntyre, 1991) and fossiliferous siltstone (Basinger *et al*., 1994), also dating from the middle Eocene, indicate that oaks were present on the landscape as far north as Axel Heiberg Island, Nunavut, *c*. 45-55 Ma (Crepet and Nixon, 1989a, b). Denk *et al*. (2017) suggest that both fossil pollen and fossiliferous siltstone found on Axel Heiberg Island belong to section *Quercus*, whereas Grimsson *et al*. (2015) were able to identify pollen grains from the same site belonging to *Quercus* sections *Quercus*, *Lobatae* and *Protobalanus*. Hipp *et al*. (2018) suggest that the current North American oak sections (*Quercus*, *Lobatae*, *Protobalanus*, *Ponticae* and *Virentes*) rapidly diverged, but were still diversifying, by *c*. 33 Ma. Oaks were then pushed southward to their current distributions following decreasing temperatures (by *c*. 3-5 °C) during the Eocene-Oligocene climate change (Hipp *et al*., 2018). Hipp *et al*. (2018) hypothesized that the contemporary American oak clades originated and expanded *c.* 10-20 Ma in both eastern and western temperate regions of NA. The two major oak sections, *Quercus* sections *Lobatae* and *Quercus*, then underwent parallel vicariance, diverging into sister clades to the east and west. From there, oaks migrated south into Mexico (where they underwent extensive diversification) and into South America (Hipp *et al*., 2018).

The systematics and classification of oaks have proved both challenging and controversial. Since the advent of scanning electron microscopes and molecular genetics,

oaks have undergone extensive reclassification (Manos *et al*., 1999; Hipp *et al*., 2014; Hubert *et al*., 2014; Hipp *et al*., 2018). Most recently, *Quercus* is divided into two subgenera (Table 1): *Cerris*, with three sections (*Cyclobalanopsis*, *Ilex* and *Cerris*) and *Quercus*, with five sections (*Quercus*, *Lobatae*, *Protobalanus*, *Ponticae* and *Virentes*). The white oaks (sect. *Quercus*) are the largest and most widespread clade, with the greatest diversity (*c.* 65% of all oak species) found in the Americas (Hipp *et al*., 2018). In order to resolve the classification of oaks in the Americas, the phylogeny of the subgenus and section *Quercus* has been extensively explored using molecular approaches (Manos *et al*., 1999; Manos *et al*., 2001; Bellarosa *et al*., 2005; Oh and Manos, 2008; Denk and Grimm, 2010; Kremer *et al*., 2012; McVay *et al*., 2017a; Hipp *et al*., 2018).

**Table 1.** Current subgenera and sections of *Quercus* according to Hipp *et al*., 2018**.**

| **Subgenus and Section** | **No. of Species** | **Geographic Distribution** |
|---|---|---|
| **Subgenus *Cerris*** | | |
| Sect. *Cyclobalanopsis* | *c.* 90 | Tropical and subtropical Asia |
| Sect. *Ilex* | 35 | Eurasia, North Africa |
| Sect. *Cerris* | 10-12 | Eurasia, North Africa |
| **Subgenus *Quercus*** | | |
| Sect. *Quercus* | *c.* 150 | North and Central America, Eurasia, North Africa |
| Sect. *Lobatae* | *c.* 120 | North and South America |
| Sect. *Protobalanus* | 5 | Southwest North America and Northwest Mexico |
| Sect. *Ponticae* | 2 | North America, Caucasus Mountains |
| Sect. *Virentes* | 7 | Southeast North America, Central America |

White oaks (*Quercus* section *Quercus*), comprising *c.* 150 species, occur in NA, Mexico, Central America, western Eurasia, East Asia and North Africa (Denk *et al*., 2017). They reach their greatest diversity first in the mountains of southern Mexico, and secondly in the southeastern United States (Nixon, 2006; Hubert *et al*., 2014: McVay *et al.,* 2017b). White oaks occupy various habitats from dry chaparral to temperate forests and wet tropical cloud forests (Nixon, 2006). They are primarily deciduous trees, with

growth forms ranging from tall trees that dominate the forest canopy to small shrubs grouped together on drier sites (Nixon, 2006). Recent studies show six groups of white oaks (Table 2): Roburoid, *Dumosae*, *Prinoideae*, *Albae*, *Stellatae* and *Leucomexicanae* (Manos, 2016; McVay *et al*., 2017a; McVay *et al*., 2017b). The *Dumosae* group contains nine species of white oak primarily distributed in the California Floristic Province (CFP) with a single species, *Q. garryana*, reaching southwestern BC (Manos, 2016). *Quercus gambelii* is a white oak distributed in central NA, but its classification has been problematic. McVay *et al*. (2017a) determined that *Q. gambelii* and the CFP white oak *Q. lobata* (group *Dumosae*) formed a clade, which was sister to the remaining eastern North American white oaks and were more distantly related to the CFP white oaks. However, historical introgression between *Q. gambelii* and *Q. macrocarpa* (and possibly *Q. lobata*) may also be responsible for this pattern.

**Table 2.** Phylogenetic groups within *Quercus* section *Quercus*.

| **Group (informal name)** | **No. of Species** | **Geographic Distribution** |
|---|---|---|
| Roburoid | *c.* 25 | Eurasia |
| *Leucomexicanae* | *c.* 100 | South-central North America to Central America |
| *Stellatae* | 9 | Southeastern United States of America |
| *Prinoideae* | 5 | Central to southeastern United States of America |
| *Albae* | 3 | Central to southeastern United States of America |
| *Dumosae* | 9 | Western United States of America and southwestern Canada |

The *Dumosae* group of white oaks includes three tree oaks (*Q. lobata*, *Q. douglasii*, *Q. garryana*) and six scrub oaks (*Q. dumosa*, *Q. berberidifolia*, *Q. john-tuckeri*, *Q. pacifica*, *Q. cornelius-mulleri*, and *Q. durata*) (Nixon, 2002). It is apparent that the two tree oaks *Q. lobata* and *Q. garryana* are genetically distinct from the scrub white oaks (Sork *et al*., 2016; Fitz-Gibbon *et al*., 2017; Kim *et al*., 2018), which all form a single clade. Recent phylogenetic work on this group placed *Q. garryana* as the sister species to the scrub oaks, and *Q. lobata* as the sister species to both Garry oak and the scrub oaks (Fitz-Gibbon *et al*. 2017). The phylogenetic placement of Garry oak is of

particular interest because of its broad north-south distribution. The majority of California white oaks have much more restricted ranges; they occur in close proximity to one other, or occasionally in mixed stands (Ortego *et al*., 2015).

Hybridization of oaks is well documented within all sections of *Quercus*, particularly in regions of range overlap (Stebbins *et al*., 1947; Hardin, 1975; Belahbib *et al*., 2001; Zeng *et al*., 2010; Hauser *et al*., 2017; Fitzek *et al*., 2018; Ortego *et al*., 2018). The Dumosae clade of white oaks, including *Q. garryana*, is known to hybridize in regions of sympatry. Analysing hybridization provides insight into the introgression of alleles between species, and the processes of reproductive isolation and speciation (Barton and Hewitt, 1989; Payseur, 2010). Introgression was originally thought to counter speciation by the gene flow of hybrid offspring into both parental populations (Arnold, 1992), while hybridization promoted species divergence (Abbott *et al*., 2013). Introgression tends to interfere with the reconstruction of species phylogenies, as is the case with oaks (McVay *et al*., 2017a). However, with increased availability of next generation sequencing methods, oak phylogenies are being constructed group by group (Cavender-Bares *et al*., 2015; Eaton *et al*., 2015; Owusu *et al*., 2015; Fitz-Gibbon *et al*., 2017; Hauser *et al*., 2017; Leroy *et al*., 2017; McVay *et al*., 2017a; Kim *et al*., 2018).

### 1.4 Garry oak (*Quercus garryana*)

Garry oak (*Quercus garryana*), also known as Oregon white oak, has the broadest latitudinal distribution of any native oak species in western NA (Flora North America Editorial Committee, 1993+). It ranges from south-central California (*Quercus garryana* var. *breweri* and var. *semota*) along the western foothills of the Sierra-Nevada to the Pacific coast of northwestern California, and north to southeastern Vancouver Island, B.C. (*Quercus garryana* var. *garryana*) (Otvos *et al*., 2012). This broadleaved deciduous tree is the single native oak species in B.C. and Washington, and is one of the few deciduous tree species in the Pacific Northwest (Bakker *et al*., 2012; Otvos *et al*., 2012). Garry oak occurs in savannah and woodland habitats (Bakker *et al*., 2012) that range from dry and well-drained soils to mesic sites with deeper soils (Smith, 2007). In BC, Garry oak ecosystems provide habitat to the greatest number of rare plant species in the province (Devine and Harrington, 2013; Otvos *et al*., 2012), and these habitats are under

duress due to urban expansion and competition by both invasive and native species (Otvos *et al*., 2012).

Garry oak habitat often coincides with agricultural, rural or urban development. Landscape fragmentation, invasive species and intensive herbivory indicate the need for conservation (Pellatt and Gedalof, 2014). Garry oaks and other species found in Garry oak ecosystems rely on animal dispersers, such as Steller's jays (*Cyanocitta stelleri*), western gray squirrels (*Sciurus griseus*), and eastern gray squirrels (*Sciurus carolinensis*) (Michalak, 2011). Michalak (2011) demonstrated the effects of canopy cover on seedling survival, which indicated seedling survival increased beyond the oak canopy, and therefore a reliance on animal dispersers for oak regeneration. The effects of urban development revealed significantly lower abundance of oak saplings, as a result of high herbivore populations and different climatic, hydraulic and chemical components in urban settings (Michalak, 2011). Clements *et al*. (2011) determined the detrimental effects of herbivory, specifically ungulate herbivores, on Garry oak regeneration; however, desiccation appears to be the greatest threat to seedling and sapling mortality (Fuchs *et al*., 2000; Bakker *et al*., 2012). This is particularly crucial in the face of global climate change where intensified summer droughts and warming temperatures in the Pacific Northwest will alter plant phenologies (Lindh *et al*., 2018).

### 1.5 Research objectives

Climate change and anthropocentric development threatens many habitats, including that of *Q. garryana*. Assessing patterns of genetic variability and genetic differences among regions is important for conservation and management. Although the population genetic structure of *Quercus garryana* has previously been investigated (Ritland *et al*., 2005; Marsico *et al*., 2009; Degner, 2014), its intraspecific phylogeny and range-wide phylogeographic patterns are not well understood. Previous studies (Ritland *et al*., 2005; Marsico *et al*., 2009) were more regional and sampled fewer populations than this study; they also lacked a specific phylogenetic focus. Sampling from a large number of populations over the whole range allows investigation of the geographical origins and migration patterns of Garry oak. In this study I used DNA sequences and other data to address the following questions. 1) What are the patterns of genetic

variation within Garry oak? 2) How do these patterns vary geographically, and how did the spatial distribution of the gene lineages come to occupy its current geographical range? 3) Does Garry oak show evidence of genetic interaction with other white oak species in western NA? 4) Is there morphological or genetic evidence to support three described varieties of Garry oak?

# Chapter 2.0 Materials and Methods

## 2.1 Site selection and sampling

I sampled *Quercus garryana* throughout its geographic range in May and June of 2016 and in May of 2017. Single individuals were sampled from each of 117 different localities in British Columbia (BC), Washington (WA), Oregon (OR), northern California (n CA) and southern California (s CA) (Figure 1a; for detailed locality information, see Appendix A). I also sampled three other white oak species, including two closely related California species, *Quercus lobata* (5 samples) and *Quercus douglasii* (2 samples), and one more distantly related species, *Quercus gambelii* (1 sample) (Table 1). The two California species co-occur geographically with *Q. garryana,* whereas *Q. gambelii* occurs further east, in the southern Rocky Mountains (Figure 1b).

Candidate sample localities for *Quercus garryana* were compiled from herbarium records available online through the portals of the Consortium of Pacific Northwest Herbaria (http://www.pnwherbaria.org/) and the Consortium of California Herbaria (http://ucjeps.berkeley.edu/consortium/). Collection sites were chosen to be more or less uniformly distributed across the range of *Quercus garryana*; localities were usually ≥ 20 km apart. At each sample site, I selected a single individual tree and collected two mature leaf specimens for DNA analysis, along with a voucher specimen. Other information collected at each site included elevation and geographical position (latitude and longitude, using a GPS unit) (Table 1), a brief site description, a list of associated woody plant species, and a photographic image.

## 2.2 DNA extraction protocol

In the field, leaf tissue samples were collected in coin envelopes, then placed in plastic freezer bags each with approximately 250 mL of silica gel. Once they were completely dry, the leaf samples were stored at -20 °C. Genomic DNA was extracted using a modified cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1987). DNA extraction was initially carried out using DNeasy Plant Mini Kits (Qiagen, Mississauga, ON, Canada), but gave very low DNA yield. Therefore, the modified CTAB

method was used for all samples. *Quercus garryana* leaf samples were incised to remove major leaf veins, and 15-25 mg of dried leaf tissue was used. The dried tissue was ground in liquid nitrogen using nylon pestles; 1 mL of liquid nitrogen was added to each 1.5 mL

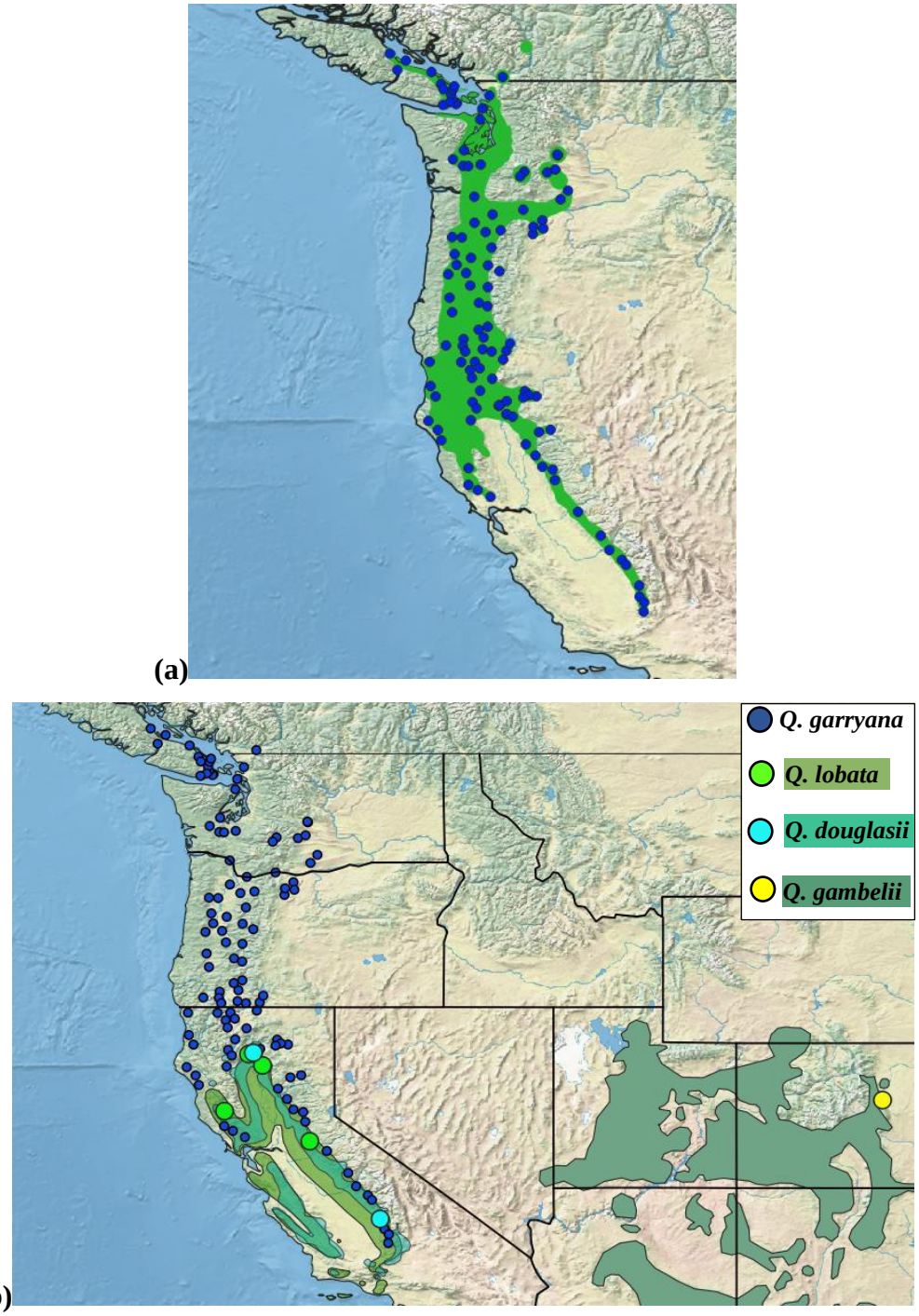



**Figure 1.** (a) Geographic distribution of *Q. garryana* with sampled localities shown (blue dots). (b) Localities (including *Q. garryana*) and ranges indicated for the other sampled *Quercus* species: *Q. lobata*, *Q. douglasii*, and *Q. gambelii* (Little, 1971).

microtube of pre-weighed leaf tissue. For a batch of 24 samples, the extraction buffer consisted of 15 mL of CTAB stock solution, 0.6 g of polyvinylpyrrolidone (PVP), and 75 µL β-mercaptoethanol.

For each sample, 500 µL CTAB buffer was added to the powdered leaf tissue and incubated at 65 °C for 2.5 h. Samples were then placed in a microcentrifuge for 1 min at maximum speed (13,300 rpm/17,000 x g). For the chloroform extraction, 500 µL of 24:1 chloroform:isoamyl alcohol was added to each sample; it was vortexed, then replaced in the microcentrifuge for 5 min at maximum speed. The upper aqueous phase was transferred into new 1.5 mL microtubes. The chloroform extraction was repeated once more. To precipitate the DNA, 7.5 M ammonium acetate (0.08X volume of aqueous phase) and 4 °C isopropyl alcohol (0.58X volume of aqueous phase) were added to the aqueous phase, mixed, and left at 4 °C overnight. Samples were placed in the microcentrifuge at maximum speed for 3 mins, then the aqueous phase was carefully poured out, ensuring to keep the DNA pellet in the microtube. To wash the DNA pellet, 1 mL 70 % cold ethanol was added; the sample was vortexed, then microcentrifuged at maximum speed for 3 mins. The wash was repeated using 85 % cold ethanol. Samples were left to air-dry for approximately 5 h. To resuspend DNA, 50 µL UltraPure™ Distilled PCR water was added and the DNA was incubated for 0.5 h at 65 °C before use. Extracted DNA stock solutions were stored at -20°C and used for all subsequent amplification and sequencing.

### 2.3 PCR protocols and DNA Sequencing

To identify plastid DNA regions that were sufficiently variable to reveal geographic patterns in *Q. garryana*, I selected 18 non-coding regions for amplification using primers previously described by Shaw *et al*. (2005; 2007; 2014). To test primers (Appendix B, Appendix C), four to eight DNA samples were selected from divergent geographic localities. The samples were prepared for PCR and multiple reactions were performed to confirm quality of PCR products. High quality PCR products were sequenced; products were obtained from ten primer pairs in total (Table 3). From this preliminary test set, four chloroplast regions that showed the greatest sequence variation (including single nucleotide polymorphisms [SNPs], indels, inversions and single

nucleotide repeats) were chosen for sequencing in all samples. These were the *accD-psaI75*, *psbD-trnT*, *rpoB-trnC*, and *trnH-psbA* intergenic spacers.

The internal transcribed spacer regions ITS-1 and ITS-2 of the 18S-5.8S-26S nuclear ribosomal cistron were amplified and sequenced using selected primer sets (Martin and Rygiewicz, 2005; Cheng *et al*., 2016). Eight primer pairs were tested (Table 4). The plant-specific *NSIP-NLBP* primer pair showed the best results with adequate amplification of the ITS region and reasonably good sequence quality. These were used in sequencing of all samples. The ITS-1, 5.8S, and ITS-2 regions of the ribosomal gene complex were sequenced as a single fragment.

**Table 3.** Plastid regions amplified to test quality of sequence traces and quantity of variable characters.

| Amplified plastid regions | Amplified PCR product (DNA bands) | Variable characters present? |
|---|---|---|
| accD-psaI75 | + | + |
| ndhC-trnV | + | --- |
| ndhF-rpl32 | + | --- |
| petA-psbJ | - | --- |
| rpl16 intron | + | - |
| rpl32-trnL | -- | --- |
| rpoB-trnC | + | + |
| rps16 intron | + | - |
| rps16-trnQ | - | --- |
| TabA-TabB | - | --- |
| TabE-TabF | + | - |
| trnC-ycf6 | + | - |
| 3'trnG-trnS | - | --- |
| trnD-trnT | + | + |
| trnH-psbA | + | --- |
| trnS-trnfM | - | --- |
| trnS-trnG2S | + | - |
| trnT-psbD | + | + |

Amplified DNA bands: + strong, - weak, -- double bands
Variable characters: + high character variability, - low character variability, --- not sequenced.

**Table 4.** Primer pairs tested for the internal transcribed spacer (ITS) region of the nuclear ribosomal gene complex.

| Amplified ITS primer combinations | Amplified PCR product (DNA bands) | Specificity of sequence results |
|---|---|---|
| ITS1-ITS4 | + | u, o, m |
| ITSu1-ITSu4 | + | u, o, m |
| ITSp5-ITSu4 | + | p, q, m |
| ITSp5-ITSp4 | + | p, q, m |
| NNC18S-C26A | + | p, q, m |
| 1406-307 | - | Not sequenced |
| NSIP-NLBP | + | p, q, c |

DNA band quality: + strong; - weak.
Specificity of sequence results: universal (u), plant-specific (p), amplified *Quercus* DNA only (q), amplified other DNA as well as *Quercus* DNA (o), produced overlapping sequence traces (m), produced primarily clean, non-overlapping sequence traces (c).

PCR amplifications for all four plastid non-coding regions and for the ITS region were carried out in 50 µL reaction volumes. PCR reagents for the plastid regions included the following: 5 µL 10X PCR buffer (Invitrogen by Thermo Fisher Scientific), 5 µL 2 mM dNTPs (Invitrogen by Thermo Fisher Scientific), 2.5 µL 5 µM forward primer, 2.5 µL 5 µM reverse primer, 1.5 µL 50 mM $MgCl_2$ (Invitrogen by Thermo Fisher Scientific), 0.75 µL 20 mg/mL bovine serum albumin (BSA) (New England BioLabs), 0.25 µL *Taq* polymerase (Invitrogen by Thermo Fisher Scientific), 27.5 µL $dH_2O$ (UltraPure Distilled Water; Invitrogen by Life Technologies), and 5 µL template DNA. PCR reagents for the nuclear ITS region followed the same protocols as above but differed in the following respects: 2.5 µL 10 µM forward primer, 2.5 µL 10 µM reverse primer, 2.5 µL 99.9% dimethyl sulfoxide (DMSO) (Sigma), 0.75 µL *Taq* polymerase, 24.5 µL $dH_2O$.

The PCR amplification protocols differed between DNA regions. For the cpDNA (plastid DNA) regions *accD-psaI75* and *psbD-trnT*, the PCR program settings were as follows: initial denaturation at 80 °C for 5 min; 30 cycles of denaturation at 95 °C for 1 min, annealing at 50 °C for 1 min, extension at 65 °C for 4 min; and the final extension at 65 °C for 5 min. For the cpDNA region *rpoB-trnC*, the protocol was as follows: initial denaturation at 80 °C for 5 min; 30 cycles of denaturation at 96 °C for 1 min, annealing at 55 °C for 2 min, extension at 72 °C for 3 min; final extension at 72 °C for 5 min. For the cpDNA region *trnH-psbA*, the protocol was as follows: initial denaturation at 95 °C for 2 min; 30 cycles of 95 °C for 1 min, 56 °C for 1 min, 72 °C for 2 min; final extension at 72

°C for 10 min. For the nuclear ITS region, the protocol was as follows: initial denaturation at 94 °C for 5 min; 35 cycles of denaturation at 94 °C for 30 s, annealing at 58 °C for 40 s, and extension at 72 °C for 1 min; final extension at 72 °C for 10 min.

PCR products were visualized using ethidium bromide and fluoresced images were produced using a MultiImage™ Light Cabinet (Alpha Innotech by Cell Biosciences) at 595 nm (+/- 50 nm) transmittance. Images were viewed using AlphaImager HP 3.4.0 (ProteinSimple). All PCR products were purified prior to sequencing, using the PureLink™ PCR Purification Kit (Invitrogen™ by Thermo Fisher Scientific). The plastid regions were sequenced with the same primers used for amplification. The nuclear rDNA ITS region was sequenced using universal primers *ITSu1* and *ITSu4* (Cheng *et al*., 2016). DNA sequencing was performed by Macrogen Inc. (Seoul, Korea), using ABI3730XL capillary sequencers (Applied Biosystems, Foster City, CA, USA). The selected regions were sequenced in both directions.

### 2.4 Sequence assembly and alignment

Sequence traces from all DNA regions were inspected and assembled using SeqMan ™ II 5.07 (DNASTAR Inc.). For each DNA region, sequences from individual samples were aligned using Jalview 2.10.3b1 (Waterhouse *et al*., 2009) and the web-alignment program Muscle (Edgar 2004). Aligned sequence lengths were trimmed to uniform length before analysis. The chloroplast intergenic spacer regions (*accD-psaI75*, *rpoB-trnC*, *psbD-trnT* and *trnH-psbA*) were concatenated into a single sequence for each sample prior to analysis. Single nucleotide repeats (SNRs) were omitted from all analyses because these regions are often subject to homoplasy, making interpretation difficult (Pleines *et al*., 2009). For the ITS region, there was one SNR present in the sequence, which was excluded from the analysis. In preparation for analysis, length variations larger than one base pair were reduced to a single character change and base substitutions that were inside indels were manually retained.

### 2.5 Analyses of DNA sequence data

Phylogenetic analyses were conducted on all 125 sampled *Quercus* for the plastid and nuclear ITS regions. Plastid samples were analysed using statistical parsimony

software, TCS 1.21 (Clement *et al*., 2000) and PopART (Leigh and Bryant, 2015), to produce an unrooted haplotype network for the concatenated chloroplast regions, and to determine haplotype abundance and relationships. A split decomposition network was constructed using SplitsTree 4.14.6 (Huson and Bryant, 2006). In preparation for phylogenetic analyses, sequences of red and white oaks distributed in NA and Asia, as well as more distantly related species of Fagaceae, were retrieved from GenBank using a BLASTn search (NCBI). These sequences (Appendix D) were then added to the plastid and nuclear ITS sequence alignments.

The plastid and nuclear ITS DNA sequence data sets were each analysed using both Bayesian and Maximum Likelihood (ML) methods. The generalized time reversible (GTR) nucleotide substitution model was selected for use in the Bayesian and ML analyses. Model selection was based on the Akaike Information Criterion (AIC) using MEGA X v10.0.1 (Kumar *et al*., 2018).

MrBayes v3.2 was used to perform Bayesian inference on the chloroplast and ITS sequence data respectively, using Markov chain Monte Carlo (MCMC) methods to determine posterior distribution of model parameters (Ronquist and Huelsenbeck, 2003). Sequence gaps (indels) were coded as binary (0/1) characters for Bayesian analysis. Two simultaneous, independent runs were performed on partitioned data sets, each initiated from different random trees. Parameters for the DNA dataset were set to a GTR nucleotide substitution model with gamma-distributed rate variation across sites and a proportion of invariable sites. Runs were performed for one million generations with sampling every $100^{th}$ generation, and a burn-in value of 25%. Not all nucleotide substitution models are available in MrBayes v3.2 (Ronquist and Huelsenbeck, 2003), but the algorithm can sample across the entire GTR model space and produce weighted averages of the best substitution models.

Maximum likelihood (ML) analyses were performed using SEAVIEW v4.7 (Galtier *et al*., 1996) for the chloroplast and ITS sequence data. SEAVIEW uses the program PhyML for ML analyses (Guindon *et al.*, 2010). PhyML parameters were set using a GTR nucleotide substitution model, optimising equilibrium frequencies, estimating proportion of invariable sites, setting the number of substitution rate categories to six, and estimating the gamma distribution parameter (Guindon *et al.*, 2010). Tree

topology search operations used subtree pruning and regrafting (SPR). Tree search was initiated from five random trees. One thousand bootstrap runs were used to test the statistical significance of the phylogenetic branch outputs. Gap characters are treated as unknown characters in PhyML, and therefore I removed them from the ML data sets prior to analysis (Guindon *et al.*, 2010).

Phylogenetic graphical outputs were viewed and edited using FigTree v1.4.3 (Rambaut, 2016). I examined correspondence between the Bayesian cpDNA and nuclear ITS phylogenetic trees using Dendroscope v3.5.10.0 (Huson and Scornavacca, 2012) to perform a Tanglegram neighbour-net heuristic analysis. Files containing cpDNA and ITS sequences for the same set of samples were analyzed in MrBayes v3.2, viewed in Figtree v1.4.3 and saved as Newick files to load into Dendroscope.

### 2.6 Geographic patterns and mapping

The geographical distributions of 125 *Quercus* samples were mapped using QGIS v3.0.0 (QGIS Development Team, 2009). The public domain mapping dataset Natural Earth v2.1.0 was used as the map base (Kelso *et al*., 2018). The mixed raster and vector input file (NE1_LR_LC_SR_W_DR) is a satellite-derived land cover layer with shaded relief, water and drainages. A world borders dataset was input as a shapefile (TM_WORLD_BORDERS-0.3) to outline national divisions (Sandvik *et al*., 2008). Latitude and longitude values of sampled *Quercus* were added as a delimited text layer. The *Quercus garryana* distribution layer was drawn using the Add Polygon Feature in QGIS (Little, 1971). Distances between sampled localities were determined using the QGIS command Measure Line.

To plot geographical distributions of individual plastid haplotypes, I added the latitude and longitude values of sampled *Quercus* spp. to the QGIS map as delimited text layers. I also constructed a haplotype diversity map using PopART (Leigh and Bryant, 2015). Haplotype sequences were clustered into five geographic groups to show regional diversity. Pie diagrams based on the regional haplotype distributions produced in PopART were constructed in Excel and overlaid on the haplotype distribution map constructed in QGIS.

## 2.7 Assessment of morphological variation

Three varieties of *Quercus garryana*, var. *garryana*, var. *breweri* (Engelm.) Jeps., and var. *semota* Jeps., are currently recognized on the basis of morphological differences. I examined morphological variation in *Q. garryana* using available traits taken from the voucher specimens collected during this study. Leaf and twig morphological characteristics were selected based on the Flora of NA key to *Quercus garryana* varieties (Flora North America Editorial Committee, 1993+) (Table 5). These included terminal bud shape, colour and length; hair density of buds and twigs; and leaf trichome measurements. On each voucher specimen, I selected three rays of the stellate

**Table 5.** Characters used for principal component analysis (PCA) of *Q. garryana* using PC-ORD.

| Character | Description |
|---|---|
| **Bud shape** | Qualitative assessment of terminal bud shape.<br>1=fusiform; 2=fusiform-ovoid; 3=ovoid. |
| **Terminal bud colour** | Terminal bud colour determined by colour of bud hairs or bud epidermis if hairs lacking.<br>1=yellowish-cream; 2=reddish yellow-cream; 3=brownish yellow-cream; 4=red-brown. |
| **Bud hair density** | Bud covered by dense, short hairs, covered by short hairs with some glandular hairs, or few hairs bearing glands.<br>1=densely pubescent; 2=densely pubescent with some glands; 3=sparsely glandular-puberulent. |
| **Twig hair density** | Twig hair density and orientation: covered with short, fine spreading hairs, sparsely covered in spreading hairs, or smooth and hairless.<br>1=puberulent with spreading hairs; 2=sparsely puberulent; 3=glabrate without spreading hairs. |
| **Terminal bud length** | Length of the terminal bud (mm). |
| **Mean number of rays on stellate trichomes** | Number of rays in a stellate trichome varied on a single leaf, the mean number of rays per leaf were calculated. |
| **Mean length of stellate rays** | Length of randomly selected rays of the stellate trichomes (mm). |

trichomes at random and measured their lengths. Quantitative traits were analysed by calculating the mean values for each individual. Qualitative traits were assigned to trait classes. Ordinations were performed to determine the overall extent and pattern of morphological variation. Preliminary analysis of the morphological data was performed using detrended correspondence analysis to test the total variance in the dataset (total $\sigma^2$=0.1356) and the gradient length of the first axis (SD=1.016). Principal Component Analysis (PCA) was then selected for an ordination method. PCA was selected because the dataset approximates multivariate normality and all variables approximate a normal distribution. All traits were assembled into a matrix and analysed using PC-ORD v7.0 (McCune and Mefford, 2011). The analysis was based on a correlation matrix and scores were calculated using a distance-based biplot. A randomization test with 999 runs was performed for PCA.

I overlaid the PCA scatterplot (of Axis 1 vs. Axis 2 scores) with each morphological variable in turn, to identify the relationships of these variables with the PCA axes. I also overlaid a variable indicating geographic ecoregions (BC and the Puget Trough, west Cascades, east Cascades, Willamette Valley, Klamath Mountains and Sierra Nevada Mountain Range) on the PCA plot to identify regional patterns associated with morphological differences. To investigate geographical patterns further, I examined the distributions of class and binned-quantitative variables by geographic region. I also plotted the first PCA axis values against sample latitude in order to assess morphological variation with latitude. Finally, to examine if there was correspondence between morphological patterns and cpDNA haplotype groups, I overlaid four plastid DNA haplotype groups (the two main plastid haplotype clusters as well as additional outlying haplotypes) on the PCA plot.

# Chapter 3.0 Results

## 3.1 Characteristics of the plastid and nuclear DNA regions

The aligned lengths for the four plastid intergenic spacer regions used for analysis were 841 base pairs (bp) for *accD-psaI75*, 1681 bp for *psbD-trnT*, 1014 bp for *rpoB-trnC*, and 563 bp for *trnH-psbA*. Total concatenated aligned length for all four regions was 4,099 base pairs for all sampled *Quercus* species. The sequenced regions contained 57 unique character variations, of which 37 characters were parsimony informative. These included 38 single nucleotide polymorphisms (SNPs), one inversion, 18 indels and 12 single nucleotide repeats (SNRs). The SNRs were removed from the DNA data set because, although these regions are highly variable, they tend to exhibit homoplasy, thus their similarities do not always indicate close relationship (Pleines *et al*., 2009). The aligned length for the ITS region (ITS1, 5.8S, and ITS2) was 599 bp, with 126 variable characters (excluding SNRs); of these, 52 characters were parsimony informative. The variable characters included 112 SNPs and 14 indels (indels varied from 1 to 20 bp).

## 3.2 Plastid haplotype diversity and distribution of *Quercus garryana*

Statistical parsimony analysis of 117 *Quercus garryana* plastid sequences revealed 24 haplotypes within the sampled range (Figure 2, Table 6).

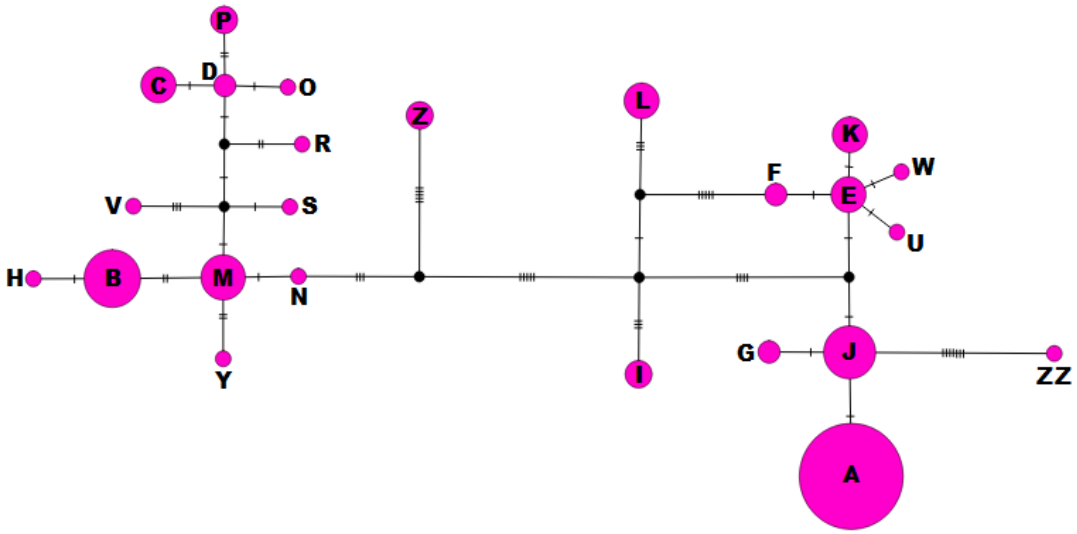



**Figure 2.** Statistical parsimony network showing relationships of sampled *Q. garryana* cpDNA haplotypes. Notches indicate mutations between haplotypes.

**Table 6.** Diversity and distribution of *Quercus* cpDNA haplotypes (28 haplotypes in total) found during this study.

| Haplotype | BC | WA | OR | n CA (38.3-42.0°) | s CA (<38.3°) | *Quercus* species |
|---|---|---|---|---|---|---|
| A | 15 | 17 | 12 | 0 | 0 | *Q. garryana* |
| B | 0 | 2 | 11 | 0 | 0 | *Q. garryana* |
| C | 0 | 0 | 0 | 5 | 0 | *Q. garryana* |
| D | 0 | 0 | 0 | 2 | 0 | *Q. garryana; Q. lobata* |
| E | 0 | 0 | 0 | 1 | 4 | *Q. garryana* |
| F | 0 | 0 | 0 | 0 | 2 | *Q. garryana* |
| G | 0 | 0 | 1 | 1 | 0 | *Q. garryana* |
| H | 0 | 0 | 0 | 1 | 0 | *Q. garryana* |
| I | 0 | 0 | 1 | 2 | 0 | *Q. garryana* |
| J | 0 | 0 | 7 | 4 | 0 | *Q. garryana* |
| K | 0 | 0 | 1 | 4 | 0 | *Q. garryana* |
| L | 0 | 0 | 1 | 4 | 0 | *Q. garryana; Q. lobata; Q. douglasii* |
| M | 0 | 0 | 0 | 8 | 0 | *Q. garryana* |
| N | 0 | 0 | 0 | 1 | 0 | *Q. garryana* |
| O | 0 | 0 | 0 | 1 | 0 | *Q. garryana* |
| P | 0 | 0 | 0 | 3 | 0 | *Q. garryana* |
| Q | 0 | 0 | 0 | 1 | 0 | *Q. lobata* |
| R | 0 | 0 | 0 | 1 | 0 | *Q. garryana* |
| S | 0 | 0 | 0 | 1 | 0 | *Q. garryana* |
| T | 0 | 0 | 0 | 0 | 1 | *Q. lobata* |
| U | 0 | 0 | 0 | 0 | 1 | *Q. garryana* |
| V | 0 | 0 | 0 | 0 | 1 | *Q. garryana* |
| W | 0 | 0 | 0 | 0 | 1 | *Q. garryana* |
| X | 0 | 0 | 0 | 0 | 1 | *Q. douglasii* |
| Y | 0 | 0 | 0 | 1 | 0 | *Q. garryana* |
| Z | 0 | 0 | 0 | 3 | 0 | *Q. garryana; Q. lobata* |
| ZZ | 0 | 0 | 0 | 1 | 0 | *Q. garryana* |
| ZZZ | 0 | 0 | 0 | 1 | 0 | *Q. gambelii* |

The diversity of haplotypes varied strongly with latitude (Figure 3). The greatest haplotype diversity was concentrated in n CA, southern OR and s CA, in decreasing order. The northern range of *Q. garryana* had the lowest haplotype diversity.

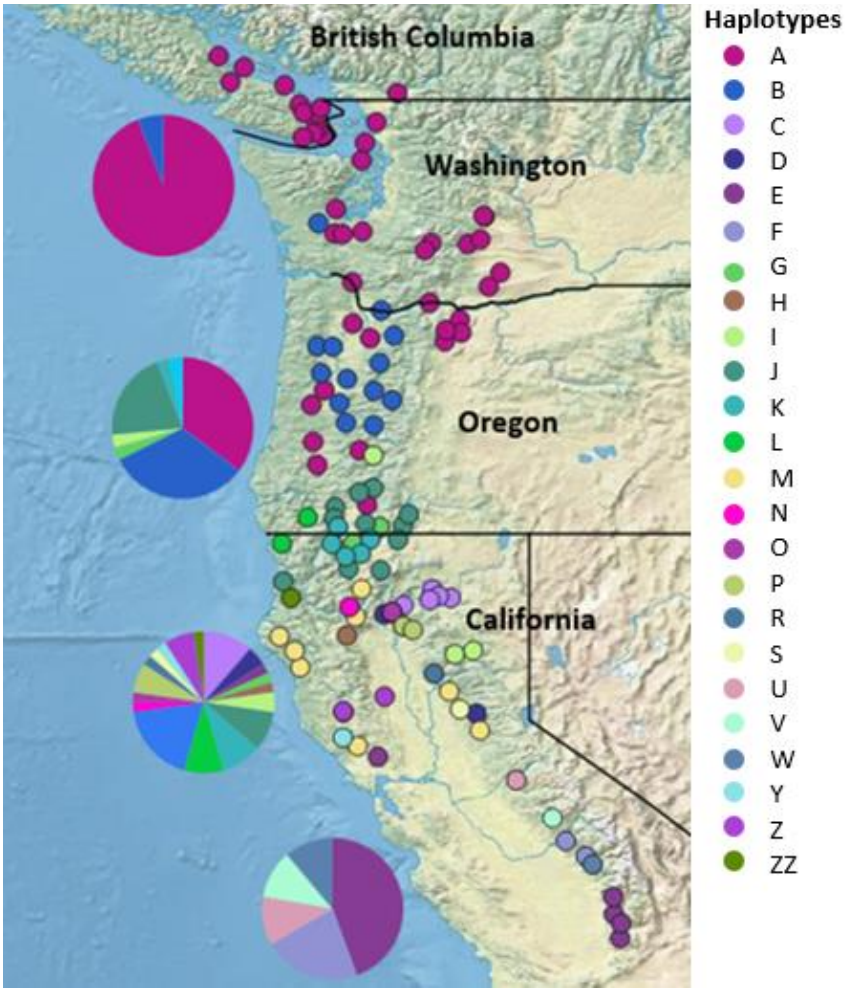



**Figure 3.** Haplotype distribution and regions of diversity for 24 *Q. garryana* cpDNA haplotypes. Pie charts depict haplotype diversity in four geographical regions: BC/WA, OR, n CA, and s CA.

Unique haplotypes (found in a single sampled population) were found only in n CA and s CA. Northern California (38.3° - 42.0° N) had 12 unique haplotypes (Table 7). In s CA, of the five haplotypes found south of 38.3° N., four haplotypes were unique to this region. A single haplotype was found in BC, two haplotypes in WA, and seven in OR; none of these haplotypes were unique to these regions.

**Table 7.** Diversity of *Quercus* (section *Quercus*) haplotypes by region.

| | Total | BC | WA | OR | N CA (38.3-42.0°) | S CA (<38.3°) | CO |
|---|---|---|---|---|---|---|---|
| **No. of samples** | 125 | 15 | 19 | 34 | 45 | 11 | 1 |
| **Total haplotypes** | 28 | 1 | 2 | 7 | 19 | 7 | 1 |
| **No. of unique haplotypes** | 28 | 0 | 0 | 0 | 13 | 6 | 1 |
| **No. of *Q. garryana* haplotypes** | 24 | 1 | 2 | 7 | 18 | 5 | 0 |
| **No. of unique *Q. garryana* haplotypes** | 24 | 0 | 0 | 0 | 12 | 4 | 0 |

Statistical parsimony and split decomposition analyses showed that *Q. garryana* haplotypes were clustered into two main groups (Figures 4 and 5), with three outlying haplotypes (I, L and Z). Group 1 included 12 haplotypes distributed in WA, OR, n CA and s CA, and showed the greatest haplotype diversity (especially in n CA). Group 2 was widespread; it was represented by fewer (9) haplotypes but had a greater latitudinal distribution than the first group.

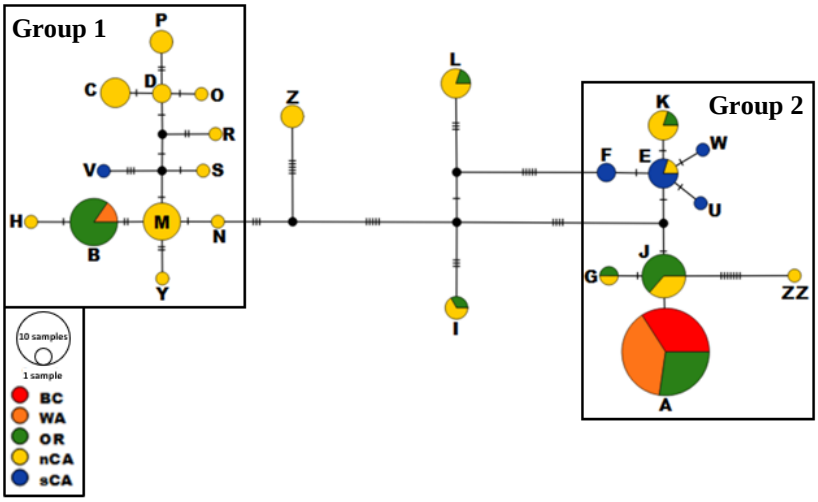



**Figure 4.** Statistical parsimony network of *Q. garryana* haplotypes showing geographic distributions of haplotypes.

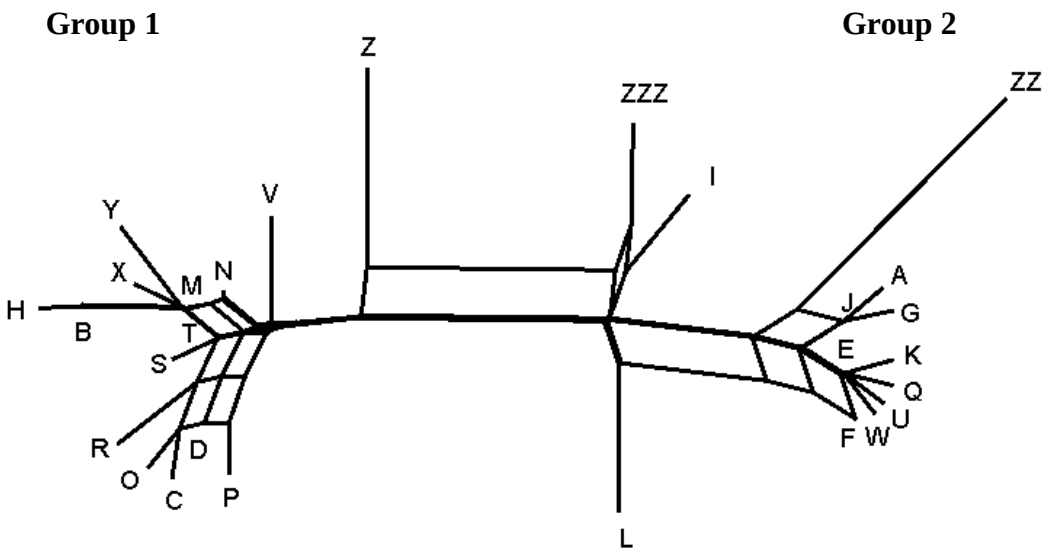



**Figure 5.** SplitsTree network showing reticulate relationships of the 28 *Quercus* haplotypes found for all species sampled in this study (including the 24 haplotypes found in *Q. garryana*).

### 3.3 Plastid haplotype distributions in all sampled *Quercus* species

Statistical parsimony analysis of all *Quercus* plastid sequences analyzed for this study (125 in total, including 117 *Quercus garryana,* five *Q. lobata,* two *Q. douglasii*, and one *Q. gambelii*) revealed 28 haplotypes (Figure 6). Both clusters of plastid haplotypes (as shown in Figure 4) also included samples of *Q. lobata* and *Q. douglasii*. *Quercus gambelii* (sampled in Colorado) was included in the analysis as an outgroup but yielded a haplotype similar to those of *Q. garryana* nearest to Group 2. Analysis of *Q. lobata* yielded three haplotypes shared with *Q. garryana* (haplotypes D, L and Z) and two haplotypes found only in *Q. lobata* (haplotypes Q and T). *Quercus douglasii* shared one haplotype with *Q. garryana* (haplotype L) and one haplotype was found only in *Q. douglasii* (haplotype X). Even those haplotypes not found in *Q. garryana* were separated by no more than one difference from *Q. garryana* haplotypes (Figure 6).

Within *Q. garryana,* the two haplotype clusters both showed the north-south pattern apparent overall and other *Quercus* haplotypes were embedded in the *Q. garryana* haplotype network, mainly corresponding to geographic distance. In the first cluster (Figure 7), haplotype diversity was primarily concentrated in n CA, except for haplotypes B and X. Haplotype B, found in 13 *Q. garryana* sampled localities, was similar to n CA

haplotypes, but occurred primarily in northern OR and WA. Haplotype X was found in a single sample of *Q. douglasii*.

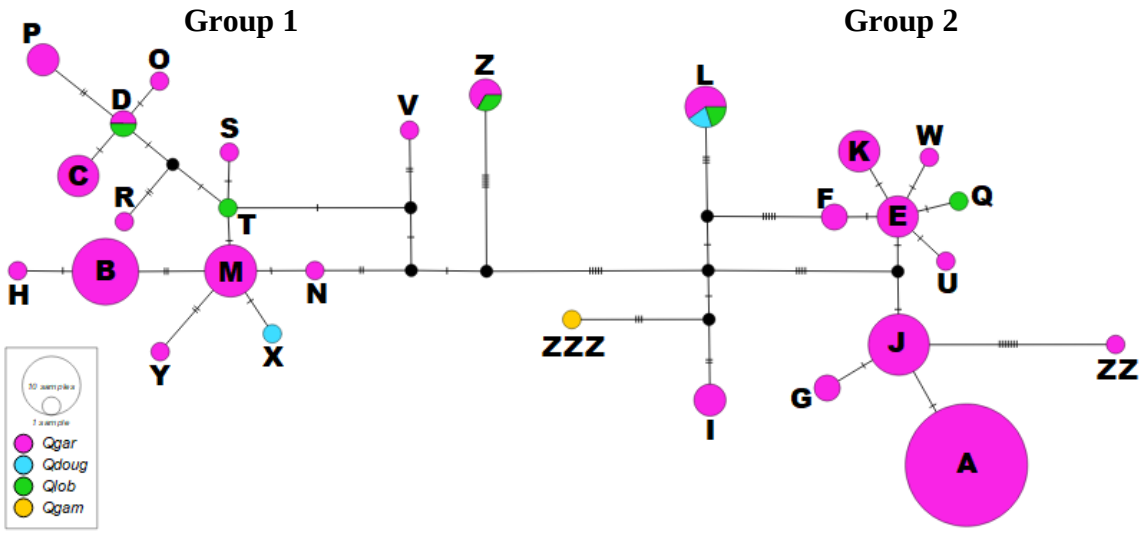
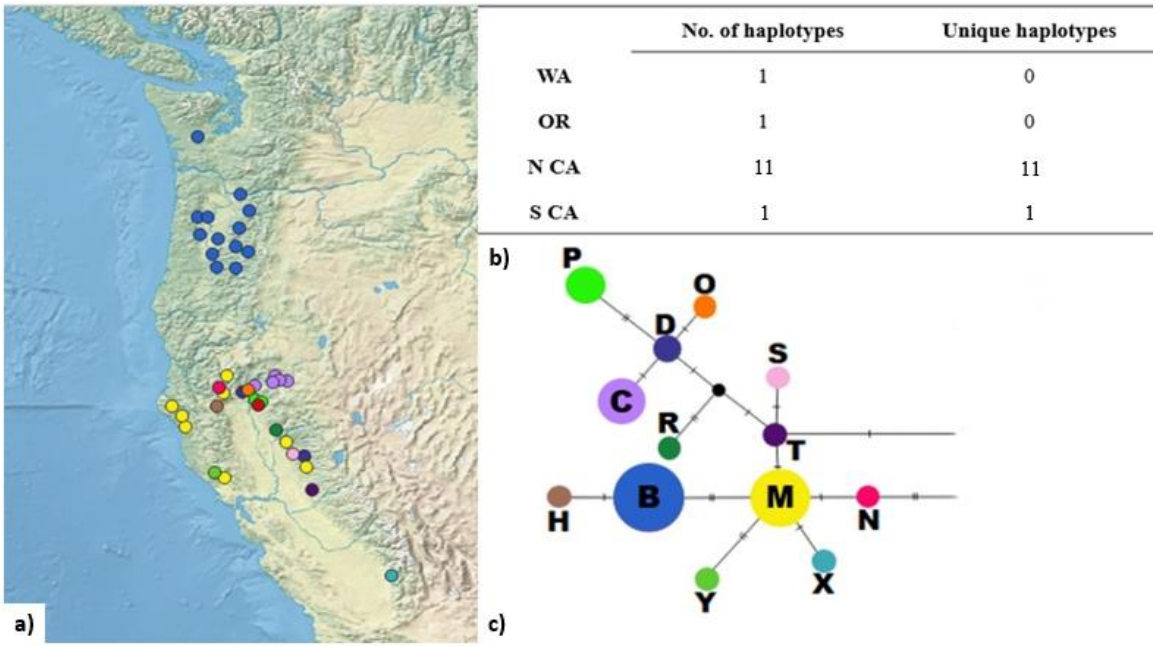


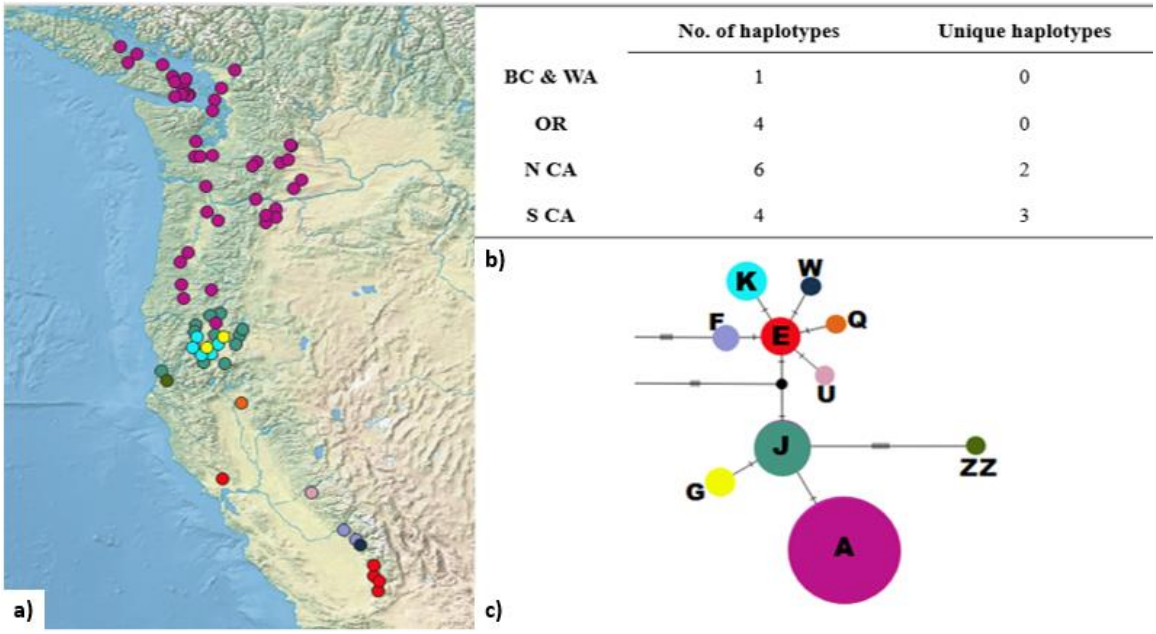

**Figure 6.** Statistical parsimony network of 125 *Quercus* plastid sequences revealed 28 haplotypes. Twenty-one haplotypes were found only in *Q. garryana*, three are common to *Q. garryana* and *Q. lobata*, and one was common to *Q. lobata* and *Q. douglasii*.




| | No. of haplotypes | Unique haplotypes |
|---|---|---|
| WA | 1 | 0 |
| OR | 1 | 0 |
| N CA | 11 | 11 |
| S CA | 1 | 1 |

**Figure 7.** *Quercus garryana* and other sampled *Quercus* haplotype distributions and haplotype network of Group 1. (a) Geographic distributions of haplotypes. (b) Numbers of total and unique haplotypes. (c) Haplotype network.

The second cluster of haplotypes (Figure 8) could be separated further into two subgroups, if the node between E and J represents a common ancestor. The first subgroup (haplotypes A, G, and J) occurred primarily in southern OR and n CA. Haplotype A was the most widely distributed haplotype (with a latitudinal range of 42.5° - 49.7°N) and was found from southern OR to southwestern BC (the northern extent of *Q. garryana*). The second subgroup (haplotypes E, F, K, Q, U, and W) was primarily distributed in s CA, apart from haplotype K, which was found in n CA. Haplotype Q was found in a single sample of *Q. lobata* collected from the foothills of the Sierra Nevada. The remaining four haplotypes found in this group extended to the southern range of *Q. garryana*.




| | No. of haplotypes | Unique haplotypes |
|---|---|---|
| BC & WA | 1 | 0 |
| OR | 4 | 0 |
| N CA | 6 | 2 |
| S CA | 4 | 3 |

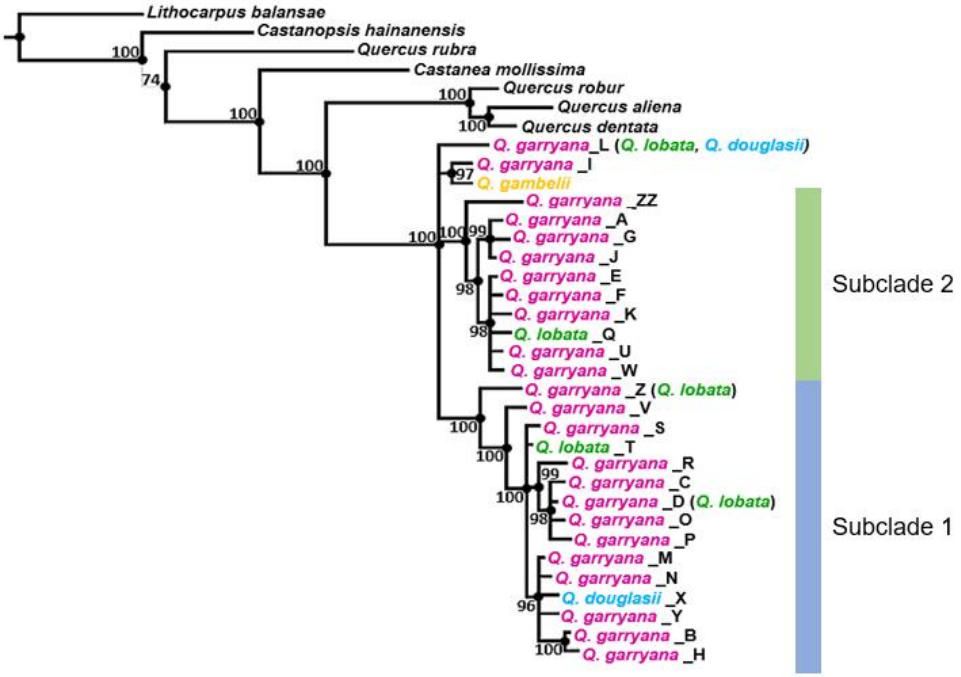
**Figure 8.** *Quercus garryana* and other sampled *Quercus* haplotype distributions and haplotype network of Group 2. (a) Geographic distributions of haplotypes. (b) Number of total and unique haplotypes. (c) Haplotype network.

### 3.4 Phylogenetic analyses of *Quercus garryana* and outgroups

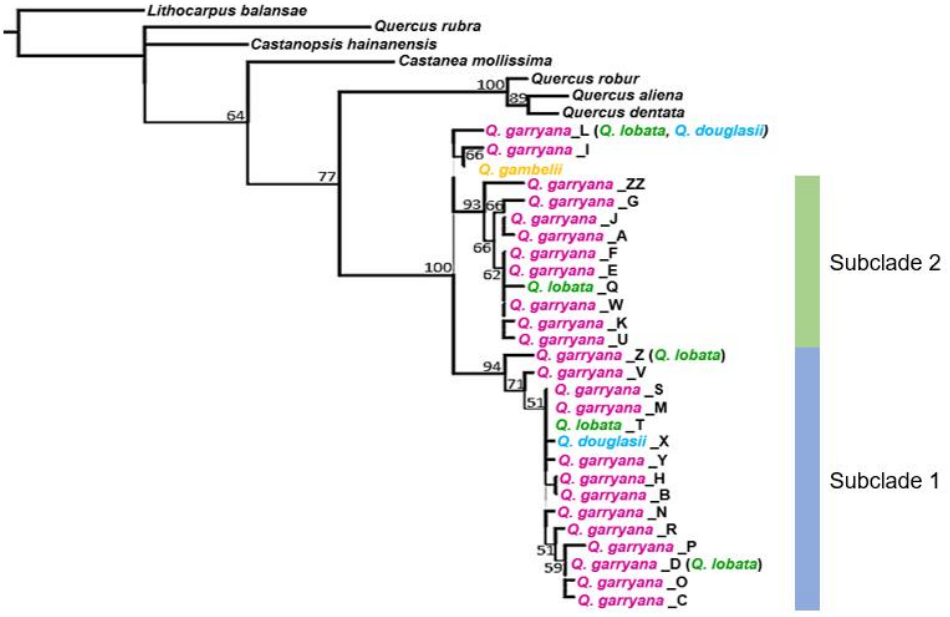
#### 3.4.1 Phylogenetic analyses of plastid DNA sequences

Bayesian analysis of *Q. garryana* cpDNA sequences resulted in two subclades (Subclade 1 and Subclade 2) of haplotypes, corresponding to the clusters apparent in the statistical parsimony and split decomposition analyses (Figure 9a). The phylogenetic tree

showed high clade support at the genus level and separation of red oaks (*Quercus* section *Lobatae*, represented here by *Q. rubra*) from white oaks (*Quercus* section *Quercus*).




**(a)**




**(b)**

**Figure 9.** Phylogenetic trees for *Q. garryana* and related species haplotypes based on cpDNA sequences. (a) Bayesian analysis with posterior probabilities indicated at nodes. (b) Maximum Likelihood with bootstrap support indicated at nodes and line thickness indicates bootstrap support threshold above (thick lines) or below (thin lines) 50.

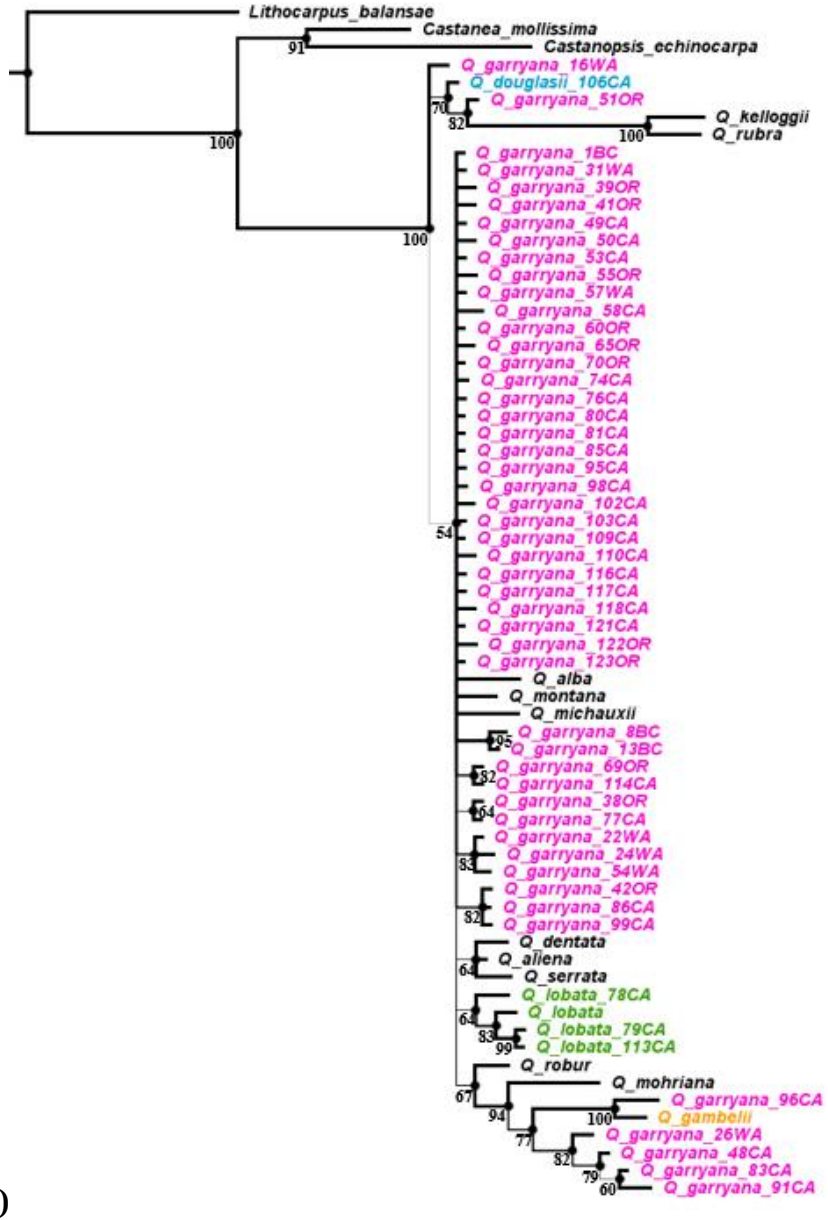
The phylogenetic relationships among the white oaks were less well resolved. *Quercus gambelii*, initially chosen and sampled as an outgroup, was grouped with *Q. garryana* and other white oaks of western NA. The *Q. garryana* haplotype I was most similar to the *Q. gambelii* sample. Haplotype I was found in three samples identified morphologically as *Q. garryana*, of which two were located in the northern Sierra Nevada (approximately 27 km apart), and the third was in the southern Cascade Mountains (approximately 373 km from the first two). The haplotypes found in the two California white oaks *Q. lobata* and *Q. douglasii* were grouped with *Q. garryana*, whether they were shared or not. Haplotype L is divergent from both of the subclades containing the remaining haplotypes. It was found in five samples (three *Q. garryana*, one *Q. lobata*, and one *Q. douglasii*); of these, three (one of each species) were collected less than 2.0 km apart, suggesting possible hybridization. The remaining two *Q. garryana* samples of haplotype L were geographically distinct; they were located 193 km and 204 km away from the previous three samples, and 58 km from one other, but this could also reflect the low chances of encountering sparsely distributed haplotypes.

The maximum likelihood tree showed a similar topology to the Bayesian tree, with lower clade support (Figure 9b). Maximum likelihood analysis was unable to resolve relationships among outgroups (*Castanopsis hainanensis* and *Quercus rubra*), perhaps because the ML analyses did not include indels. *Quercus gambelii*, *Q. lobata* and *Q. douglasii* remained included within the *Q. garryana* phylogeny. The haplotype clusters were similar in both analyses.

### 3.4.2 Phylogenetic analyses of nuclear DNA sequences

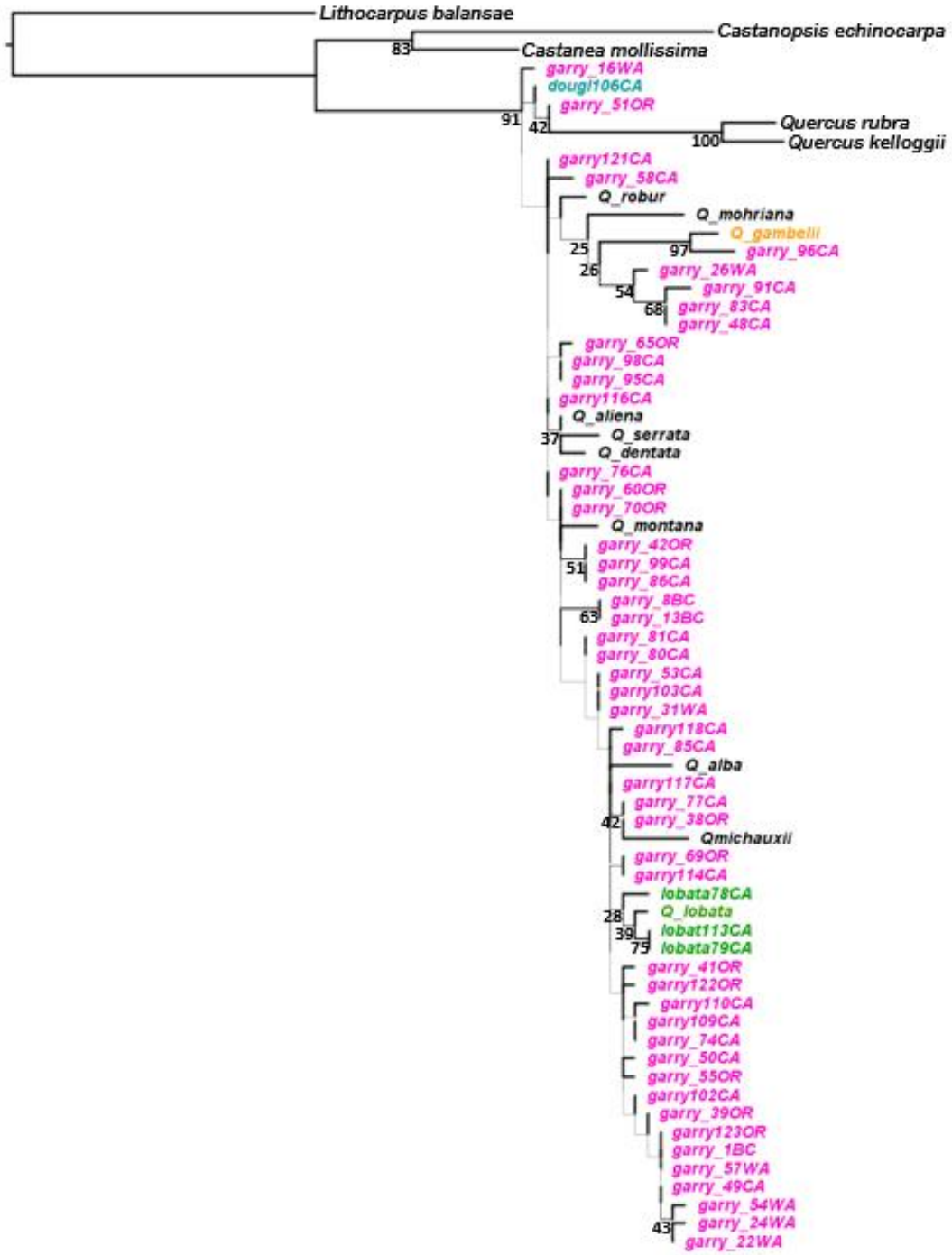
Analyses of the ITS region yielded inconclusive results (Figures 10, 11). *Quercus kelloggii* and *Q. rubra* (*Quercus* section *Lobatae*) formed a weakly supported sister clade to the main group of white oaks (*Quercus* section *Quercus*). Two samples of *Q. garryana* and a single sample of *Q. douglasii* were found to be more closely related to the red oaks (*Quercus* section *Lobatae*), with moderate clade support, which could be a result of convergent mutations or insufficient variation in the ITS region to resolve the tree. Within *Quercus* section *Quercus*, relationships were poorly resolved. Eastern and central North American white oaks (*Q. alba*, *Q. montana*, and *Q. michauxii*) were not separated

from the western species. Asian white oaks (*Q. dentata*, *Q. aliena* and *Q. serrata*) clustered together but with low clade support. The four samples of *Q. lobata* included in this analysis formed a clade but also with weak support. The remaining North American samples did not resolve into monophyletic clades, except for the well supported clade with *Q. gambelii*. Maximum likelihood analysis showed similar phylogenetic relationships, with little resolution within *Quercus* section *Quercus.*




**(a)**

**Figure 10.** Phylogenies of sampled *Quercus* and outgroups based on the nuclear ITS region; (a) Bayesian tree with posterior probabilities indicated at nodes; (b) PhyML Maximum Likelihood tree with clade bootstrap support indicated at nodes.




**(b)**

**Figure 11.** Phylogenies of sampled *Quercus* and outgroups based on the nuclear ITS region; (a) Bayesian tree with posterior probabilities indicated at nodes; (b) PhyML Maximum Likelihood tree with clade bootstrap support indicated at nodes.

### 3.4.3 Correspondence between the plastid and nuclear phylogenetic trees

A comparison of the plastid and nuclear ITS phylogenies of *Quercus*, including three distantly related Fagaceae used as outgroups, is shown in Figure 12. This comparison can reveal differences in plastid and nuclear phylogenies that may suggest hybridization. Although the plastid Bayesian tree had good support for some clades, the ITS phylogeny was poorly resolved.

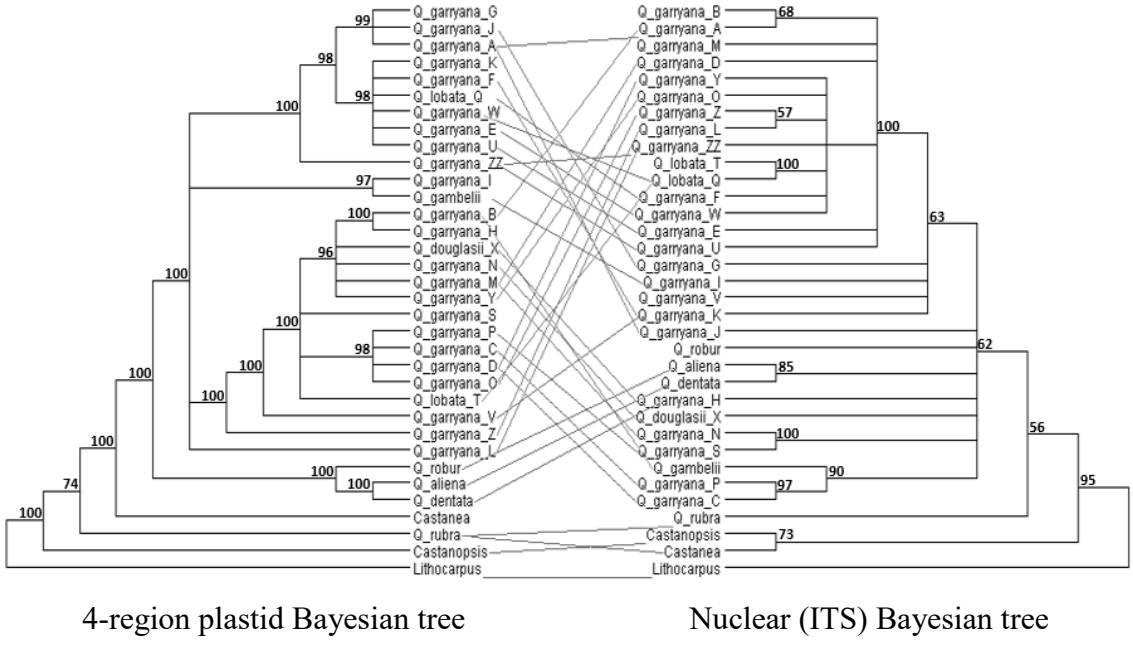



**Figure 12.** Tanglegram showing correspondence between nuclear (ITS) and plastid Bayesian phylogenetic trees.

The ITS and cpDNA Bayesian (and ML) trees showed some differences in clusters of *Q. garryana* samples, which is evident by diagonal lines indicating regions of incongruence between the trees. These differences are consistent with hybridization among oak species in this group. However, because of the lack of resolution in these phylogenies (especially the ITS tree), these results should be treated with caution.

### 3.5 Morphological variation among Garry oak varieties

#### 3.5.1. Analysis of morphological characters from voucher specimens

Three varieties of *Q. garryana* have been described on the basis of morphological differences: var. *garryana*, *breweri* and *semota*. A PCA was carried out to assess patterns of variation and detect differences among groupings (Figure 13). Although I did not detect distinct morphological clusters, some differences corresponding to previously described varieties were found.

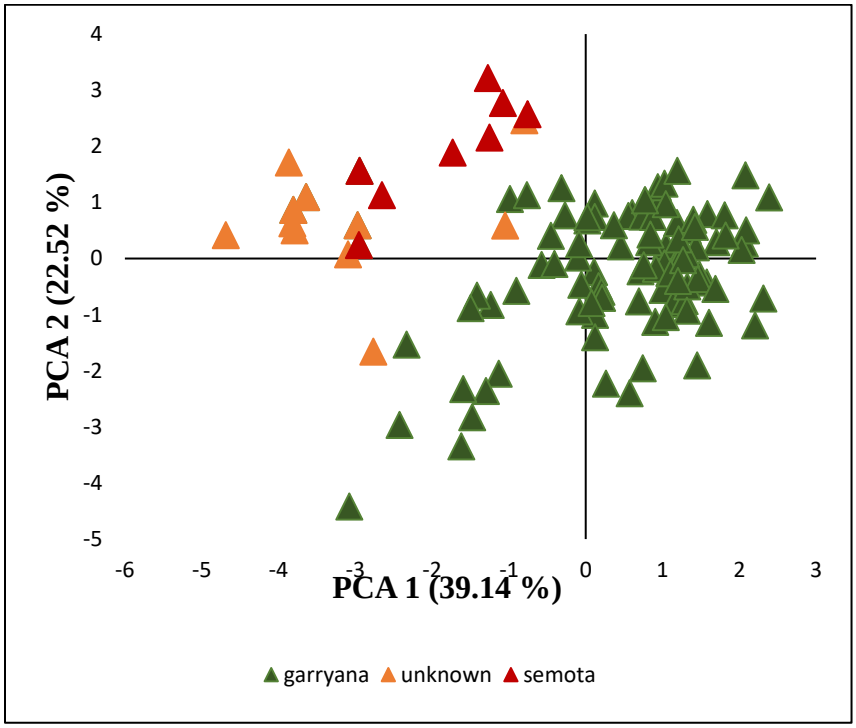



**Figure 13.** PCA of *Q. garryana* based on seven morphological characters that have been used to distinguish the varieties. The first two axes capture 61.7 % of total variance. *Quercus garryana* var. *garryana* and var. *semota* were identifiable using these morphological characters; var. *breweri* could not be separated from the other varieties.

The most southern samples of *Q. garryana* individuals clustered together based on the following characters: ovoid bud shape, short terminal bud length, a larger number of rays per stellate trichome, and shorter stellate rays (Figure 14; see tables 8 and 9 for character descriptions). These characters likely represent variety *semota*. Variety *garryana* was identified primarily based on growth-form (tree-like) and variety *breweri* was not distinguished from the other varieties in the PCA.

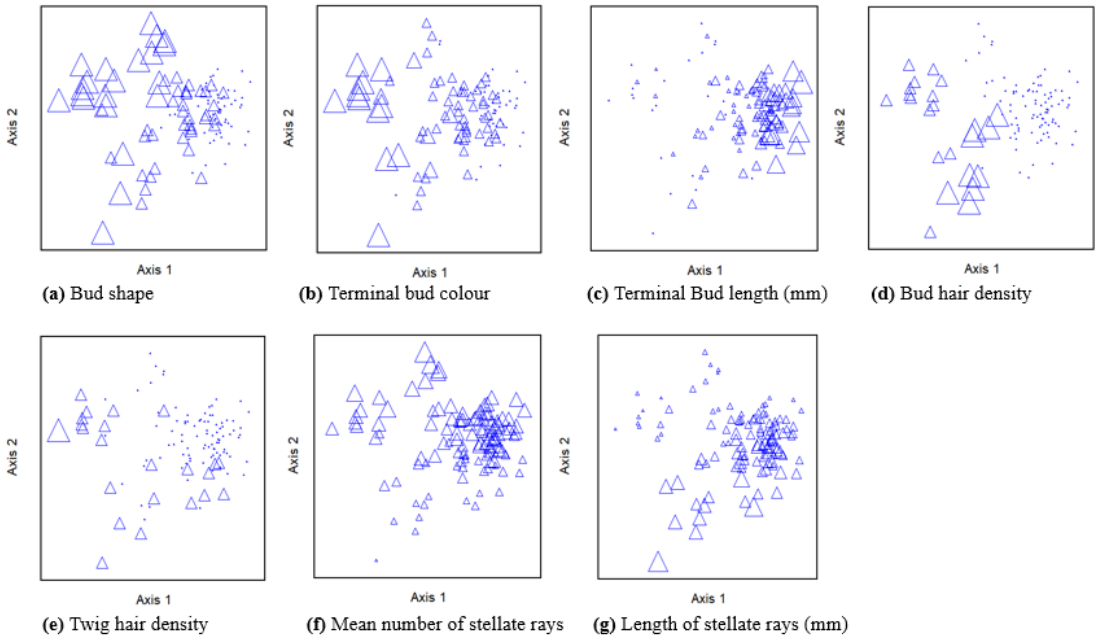



**Figure 14.** PCA plots for *Q. garryana*, indicating variation in individual morphological variables. Symbol size corresponds to the value (class or quantitative) of the morphological character. Axes for all plots correspond to those in Figure 13.

**Table 8.** Categorical values for morphological characters used in identification of *Q. garryana* and varieties.

| **Morphological character** | **Character description** |
|---|---|
| Bud shape | 1 = fusiform<br>2 = fusiform-ovoid<br>3 = ovoid |
| Terminal bud colour | 1 = yellowish-cream<br>2 = reddish yellow-cream<br>3 = brownish yellow-cream<br>4 = red-brown |
| Bud hair density | 1 = densely pubescent<br>2 = densely pubescent with some glands<br>3 = sparsely glandular-puberulent |
| Twig hair density | 1 = puberulent with spreading hairs<br>2 = sparsely puberulent<br>3 = glabrate without spreading hairs |

**Table 9.** Quantitative characters used form morphological identification of *Q. garryana* and varieties.

| Morphological character | Character description | N sample size | Mean | SD | Min - Max |
|---|---|---|---|---|---|
| Terminal bud length | Length of the terminal bud (mm). | 116 | 4.31 | +/- 1.98 | 2 - 10 |
| Mean no. of rays on stellate trichomes | Number of rays in a stellate trichome varied on a single leaf, the mean determined from 20 counts. | 116 | 4.28 | +/- 0.96 | 1.5 – 7.3 |
| Mean length of stellate rays | Length of randomly selected rays of the stellate trichomes (mm). | 116 | 0.20 | +/- 0.05 | 0.11 – 0.36 |

Correlations of individual morphological characters with PCA axes 1 and 2 are shown in Table 10. Terminal bud shape, colour and length make the highest contributions to the first axis, whereas the mean number of stellate hairs and length of stellate rays contribute most to the second axis. These correlations are also visible in Figure 13.

**Table 10.** PCA axes 1 and axes 2 r-values, tau-values and eigenvectors for seven morphological characters used in morphometric analysis of *Q. garryana*.

| Morphological characters | r-value | | tau-value | | Eigenvector | |
|---|---|---|---|---|---|---|
| | Axis 1 | Axis 2 | Axis 1 | Axis 2 | Axis 1 | Axis 2 |
| **Bud shape** | -0.820 | 0.171 | -0.666 | 0.145 | -0.496 | 0.136 |
| **Terminal bud colour** | -0.838 | -0.026 | -0.647 | 0.002 | -0.506 | -0.021 |
| **Terminal bud length** | 0.730 | -0.049 | 0.672 | -0.083 | 0.441 | -0.039 |
| **Bud hair density** | -0.662 | -0.347 | -0.512 | -0.126 | -0.400 | -0.276 |
| **Twig hair density** | -0.581 | -0.251 | -0.362 | -0.145 | -0.351 | -0.200 |
| **Mean no. of stellate hairs** | -0.048 | 0.873 | -0.005 | 0.698 | -0.029 | 0.696 |
| **Mean length of stellate rays** | 0.232 | -0.774 | 0.145 | -0.597 | 0.140 | -0.616 |

### 3.5.2 Geographic patterns of morphological variation

To determine if there was a geographic component to the morphological variation, samples were assigned to one of six ecoregions: BC and the Puget Trough, west Cascades, east Cascades, Willamette Valley, Klamath Mountains and Sierra Nevada Mountain Range (Figure 15). There was some morphological variation among ecoregions. In particular, southern samples (from the Sierra Nevada and Klamath regions) appeared to be different from the others. The remaining samples from further north and over the remainder of the geographic range were all fairly similar.

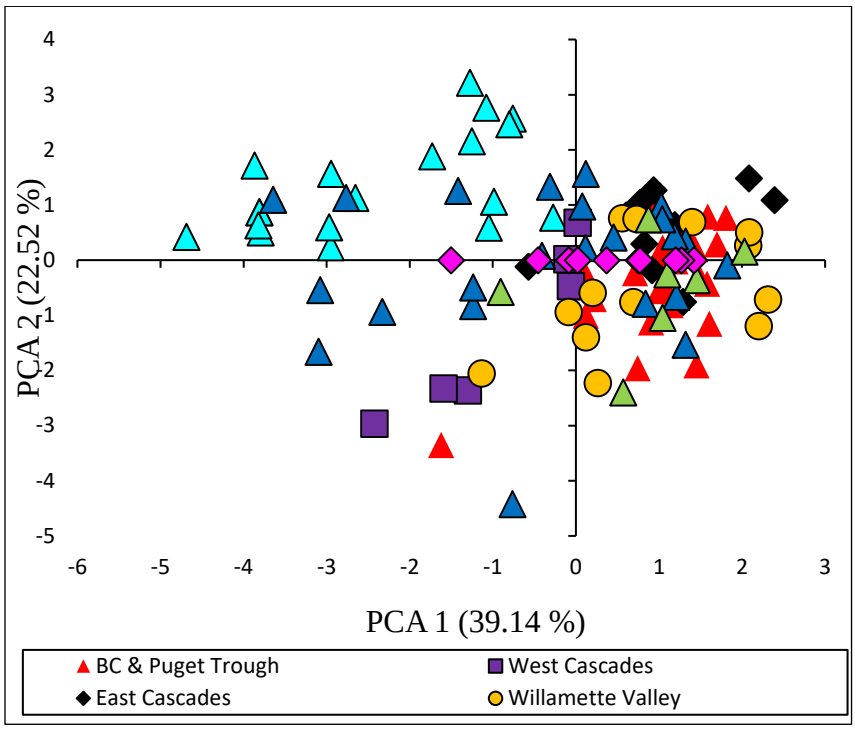



**Figure 15.** PCA of *Q. garryana* based on seven morphological characters, with samples colour-coded by ecoregion region.

To investigate these geographic patterns further, I examined the correlation between the first PCA axis and latitudes of sampled *Q. garryana* (Figure 16). The morphological variation in the character dataset was correlated with latitude ($r = 0.503$, ($P < 0.01$). I also examined the frequency distributions of the individual qualitative and quantitative morphological characters by geographic region (Figure 17a-g). From these it was evident that the morphological characters showed distinct geographic patterns.

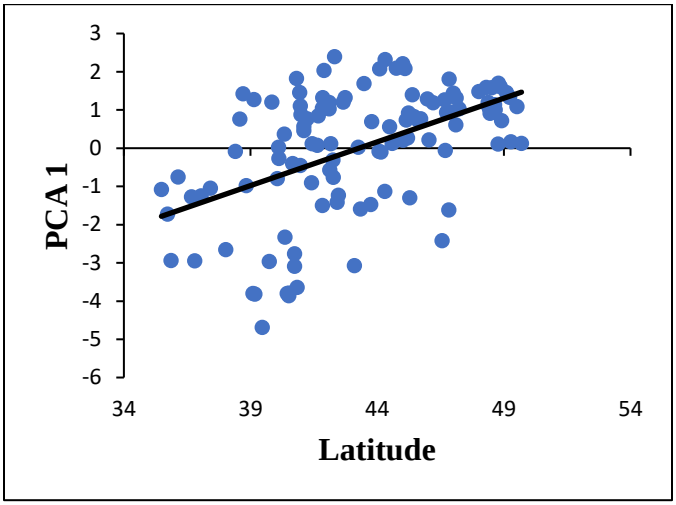
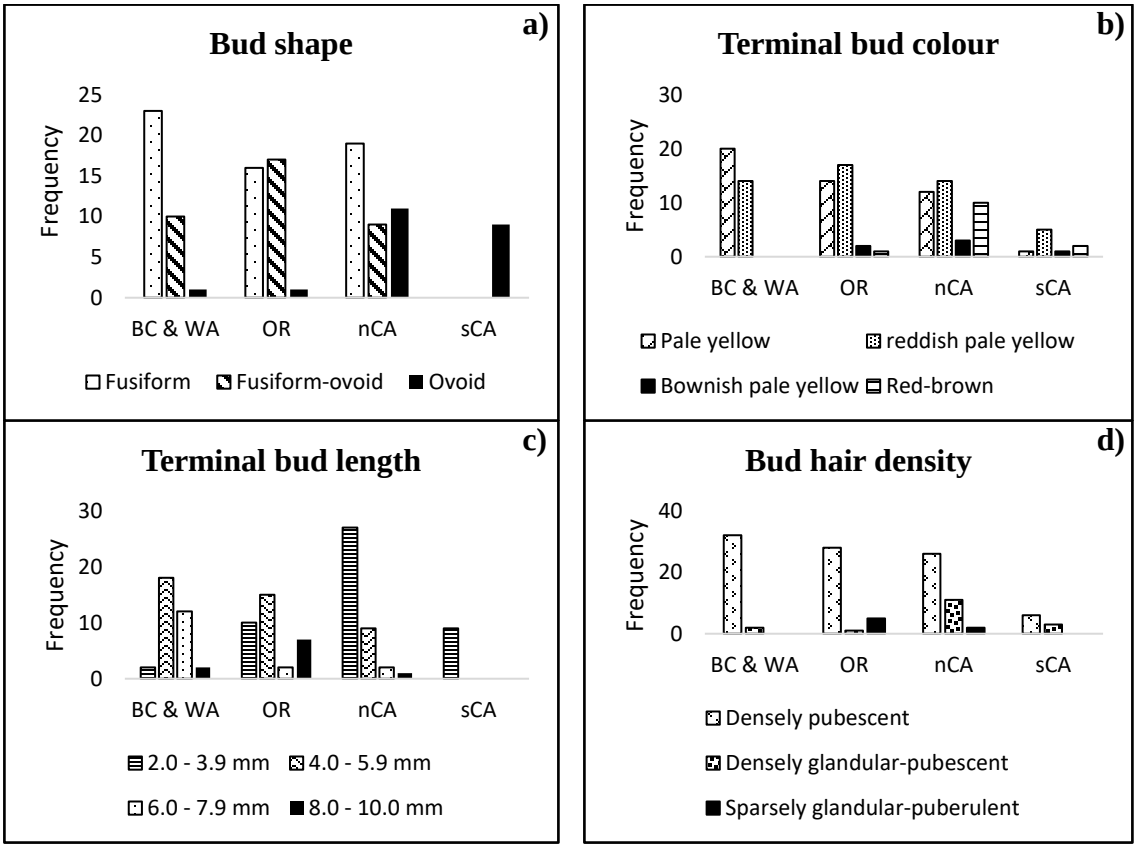



**Figure 16.** Latitude of sampled *Q. garryana* vs. PCA axis 1 scores.



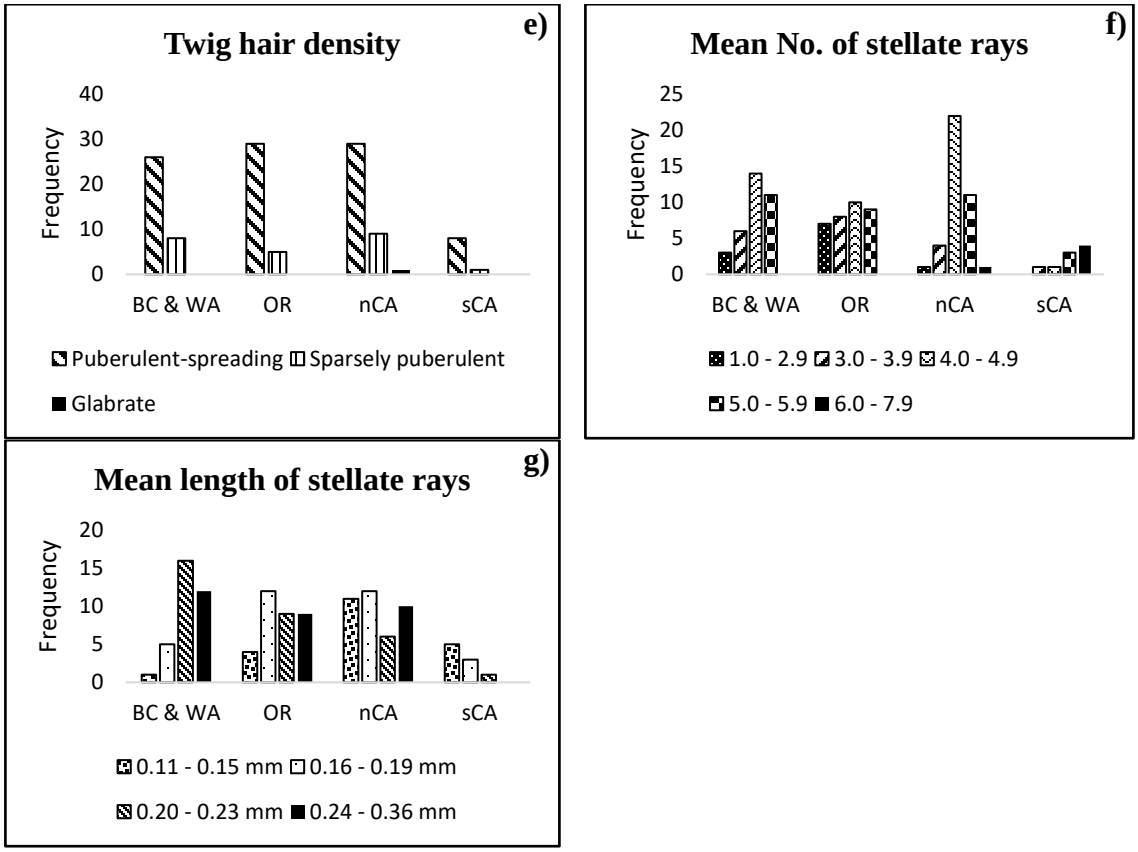



**Figure 17.** Geographic variation in qualitative and quantitative morphological characters used to identify *Q. garryana* to variety: a) bud shape, b) terminal bud colour, c) terminal bud length, d) bud hair density, e) twig hair density, f) mean number of stellate rays, and g) mean length of stellate rays.

### 3.5.3 Relationship between morphological and molecular variation

Plastid haplotype groups were overlaid on the PCA to determine if there were any patterns of morphological variation associated with molecular differences (Figure 18). Symbols indicate the two clades determined using Bayesian and ML phylogenetic analyses, as well as the two divergent haplotypes I and L. There was no obvious correspondence between morphological variation and haplotype cluster.

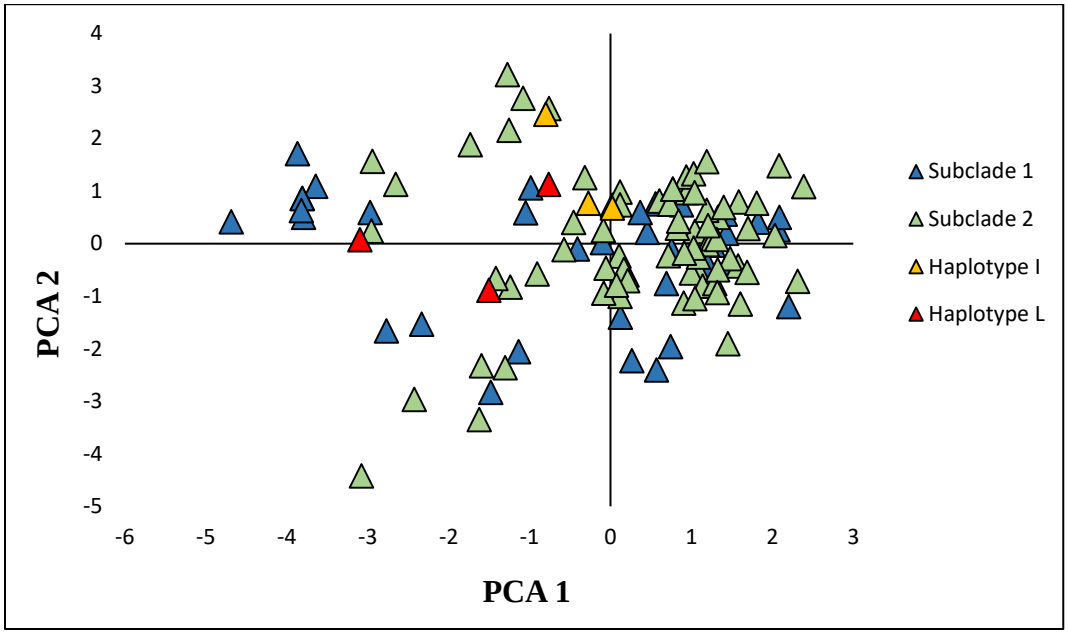



**Figure 18.** PCA with plastid haplotype subclades (cluster 1 and cluster 2) and two divergent haplotypes indicated. Subclades correspond to Figure 9a and b.

# Chapter 4.0 Discussion

In NA during the Pleistocene epoch, climatic oscillations and the associated advances and retreats of continental ice sheets had a profound effect on plant distributions. During alternating glacial and interglacial periods, geographic ranges contracted or expanded, shifted south or north, and sometimes became subdivided in multiple refugia. These shifts affected both species directly displaced by glacial advances (Shafer *et al*., 2010; Roberts and Hamann, 2015) and those located far south of the glacial boundary (Gugger *et al*., 2013). Following the LGM, many plant species expanded their ranges and gradually migrated northward. Post-glacial range shifts have been well documented for a variety of species with both palynological and molecular evidence; higher latitudes are generally associated with more recent arrival times in pollen records (Cwynar and MacDonald, 1985) or lower levels of genetic variability (Soltis *et al*., 1997; Beatty and Provan, 2011; Nadeau *et al*., 2015).

Prior to the Pleistocene, in the Late Cretaceous and early Tertiary, oaks (*Quercus*) in NA may have extended as far north as Axel Heiberg Island (79° 26′ N, 90° 46′ W) and dominated much of what is now boreal forest (Crepet and Nixon, 1989a, b; McIntyre, 1991; Basinger *et al*., 1994; Grimsson *et al*., 2015; Denk *et al*., 2017). As temperatures decreased, oaks migrated south and diversified in the temperate regions of NA and Mexico (Cavender-Bares *et al*., 2018; Hipp *et al*., 2018). In the California Floristic Province, the white oaks (*Quercus* sect. *Quercus*) diversified into 10-15 species (Nixon, 2002). Of these, *Quercus garryana* is of particular interest. It has the largest geographic range of any species in the California white oak group, and is the only species of this group that extends as far north as the LGM glacial boundary. It occurs in diverse habitats and climatic conditions, including marine-influenced mesic environments along the Pacific coast, inland continental habitats east of the Cascades in Washington, and dry forest and chaparral in the western foothills of the Sierra Nevada.

### 4.1 Geographic history of *Quercus garryana*

In this study I showed that the genetic diversity of *Q. garryana* (as inferred from cpDNA sequence data) is highest in northern California. Of the 24 Garry oak plastid haplotypes identified in this study, half were unique to northern California (Figure 4). This region includes the Klamath Mountains and North Coast Ranges, which are geologically old and known for high levels of species diversity and endemism (Baldwin *et al*., 2017). The Siskiyou Mountains (a sub-range of the Klamath) have previously been proposed as a geographic region for refugia during the Pleistocene (Whittaker, 1961; Smith and Sawyer, 1988). *Quercus garryana* sampled in the northern half of its range had lower genetic diversity, being represented by only two haplotypes. This pattern is consistent with the leading-edge model of range expansion (Cwynar and MacDonald, 1987; Hewitt, 1993; Soltis *et al*., 1997) and supports the hypothesis of post-glacial recolonization by *Q. garryana* from populations in the southern part of its range.

My analyses revealed two distinct groups or subclades of cpDNA haplotypes (Groups 1 and 2; see Figures 4, 7, 8, 9), each showing a north-south gradient in genetic diversity. The patterns of diversity in each group were very similar, with the northern populations in each case having a single haplotype, and all remaining haplotypes concentrated in the southern part of the range. These data suggest that *Q. garryana* made two separate migrations north following the LGM, each originating from different source groups within the species. As populations from each of these groups moved north, haplotypes were gradually lost via dispersal and genetic drift, until only one haplotype from each group remained (haplotypes A and B respectively; see Figure 3). Haplotypes A and B have different and somewhat separate distributions. Haplotype A has the larger latitudinal range (42.5° - 49.7° N) and extends as far north as southwestern BC, the northern range limit of *Q. garryana*. Individuals with this haplotype occupy open landscapes west of the Cascades, extending to foothills east of the Cascades in Washington and Oregon (where they probably spread via the Columbia Gorge), and north to protected coasts of the Puget Trough and southeastern Vancouver Island. Haplotype B occurs primarily in the Willamette Valley of Oregon, with a few (possibly more recent) populations further north and west. Haplotype A was more abundant in my sample (44

samples) than haplotype B (13 samples), suggesting that Group 2 haplotypes may have represented an earlier wave of recolonization.

Other recent genetic studies of *Q. garryana* (Ritland *et al*., 2005; Marsico *et al*., 2009; Degner, 2014) revealed genetic patterns generally consistent with those reported in this study. Marsico *et al*. (2009), using plastid and nuclear microsatellite markers, sampled 22 populations from southwestern BC to southern Oregon. They found low within-population variation in all populations sampled, but identified decreased allele richness with increasing latitude, reflecting the pattern of haplotype distributions for *Q. garryana* found in this study. Degner (2014) assessed genetic variation in *Q. garryana* using genotyping-by-sequencing (GBS), sampling 11 populations from southern BC to California. Degner (2014) grouped the sampled populations into a northern group (southern BC to southern Oregon) and a southern group (southern Oregon and California). He noted that these groups corresponded with the geographic distributions of *Q. garryana* varieties, one of which (var. *garryana*) is northern, whereas the other two (var. *breweri* and var. *semota*) are southern. Degner (2014) found a high degree of admixture within the GBS dataset, indicating substantive gene flow between the two groups of *Q. garryana* populations.

Ritland *et al*. (2005) investigated genetic variation in Garry oak using isozymes. Although these authors primarily investigated within-population genetic structure, they also showed that their sampled populations were grouped into northern and southern clusters. They speculated that one of these clusters could have been associated with a refugium on the west coast of Vancouver Island during the LGM. However, it seems unlikely that such a refugium could have harboured *Q. garryana*. Roberts and Hamann (2015) constructed species distribution models for 22 tree species to reconstruct possible refugia in western NA. For Vancouver Island, a refugium having a boreal-like climate was postulated for the north-central west coast of the island. If *Q. garryana* were present in such refugial sites during the LGM, this should be indicated by higher haplotype diversity in the northern range than was found in my study or other recent studies.

Low genetic diversity in the northern sampled populations of *Q. garryana* suggests relatively recent northward migration of this species, a pattern also supported by palynological evidence. Pollen records from Upper Klamath Lake (42.39°N 121.9°W) in

southern Oregon and Clear Lake (39.06°N 122.82°W) in northern California indicate that *Quercus* spp. have been present dating back to *c.* 26,000 cal yr BP (Hakala and Adam, 2004) and *c.* 130, 000 cal yr BP (Adam and Robinson, 1988), respectively. The long history of oaks in California is consistent with the high genetic haplotype diversity that I observed in *Q. garryana*, although *Quercus* was not identified to species in the studies by Hakala and Adam (2004) and Adam and Robinson (1988). In the northern range of *Q. garryana*, Pellatt *et al*. (2001) examined Holocene vegetation history and noted the arrival of *Quercus* pollen (likely *Q. garryana*) on southern Vancouver Island *c*. 8300 cal yr BP. Lucas and Lacourse (2013) found similar results in the southern Gulf Island region, with *Q. garryana* pollen increasing after *c.* 8000 cal yr BP. These findings could reflect long-distance pollen dispersal. However, Sork *et al*. (2002) estimated that effective pollen dispersal averaged 65 m for *Q. lobata* in open savannah landscapes in California, although non-effective pollen flow was documented to 200 km. It is important to account for the over-representation of pollen deposits by wind-pollinated tree species, as was concluded by Roy *et al*. (2018) in the western Himalayas; pollen abundance does not necessarily infer species abundance (Goring *et al*., 2013). However, it seems unlikely that *Q. garryana* pollen was transported over very large distances. Snell and Cowling (2015) simulated migration rates using long-distance seed dispersal and found that Beech (*Fagus* spp.) migrated *c.* 20 m $yr^{-1}$. Beeches and oaks share similar life history traits, including slow growth, long generation times, and large fruit. It seems reasonable to conclude that *Q. garryana* migrated slowly northward and is a relatively recent arrival in its current northern range.

Regions of high genetic diversity within a species are likely to most strongly reflect the recent events in its history, and do not always correspond to regions of origin. Other processes, such as hybridization, may also make an important contribution. Local patterns of dispersal may be superimposed on regional patterns. Under cooler climatic conditions, *Q. garryana* may have occurred further south with populations dispersed throughout the California Central Valley, and as temperatures increased it may have migrated to higher elevations. Similar local elevational migration has been suggested for *Q. lobata* during the Pleistocene glaciations, based on nuclear and plastid microsatellite variation (Gugger *et al*., 2013). Thus, it may be difficult to identify the geographical

origins of *Q. garryana*. However, studies using combinations of genetic markers can reveal complex patterns and help to further illuminate the histories of individual species.

### 4.2 Introgression and hybridization between sympatric white oaks

Oaks worldwide are well known for hybridization and introgression, so much so that they are often used as classic examples in discussions of plant evolution and speciation (Burger, 1975; Van Valen, 1977). Hybridization in oaks occurs primarily between sympatric species (Whittemore and Schaal, 1991), though they can still maintain species boundaries even over short distances (Craft *et al*., 2002; Lyu *et al*., 2018). In the white oaks, hybridization occurs predominantly within subsections: *Quercus, Lobatae*, *Protobalanus*, *Ponticae* and *Virentes*. First-generation (F1) hybrids, even if they do not persist, may backcross to parental populations, leading to introgression. Currently available evidence suggests that introgression is a widespread phenomenon in oaks (Heiser, 1949; Sork *et al*., 2016; Kim *et al*., 2018).

In this study I identified shared plastid haplotypes among *Q. garryana, Q. lobata*, and *Q. douglasii* in northern California, demonstrating sympatric hybridization among these three species. Out of 24 *Q. garryana* haplotypes detected in total, three were shared (two with *Q. lobata*, and a third with both *Q. lobata* and *Q. douglasii*). The white oaks of California are known to be interfertile, yet maintain morphological and ecological distinctions. Various studies have investigated the extent of hybridization and introgression in these oaks (Craft *et al*. 2002; McVay *et al*. 2017b; Kim *et al*. 2018; Ortego *et al*., 2018). Oney-Birol *et al*. (2018) found that among the California scrub oaks, *Q. engelmannii* shared alleles with *Q. berberidifolia* and *Q. cornelius-mulleri.* These authors were unable to differentiate whether these shared alleles were from a common ancestor or reflected recent introgression, and this was also the case with my study. The ranges of most white oak species in California do not overlap with that of *Q. garryana* (see Ortego *et al*., 2015). However, *Q. lobata* and *Q. douglasii* do overlap in range, particularly in northern California. Using nuclear DNA microsatellite markers, Craft *et al*. (2002) applied genetic differentiation and Bayesian clustering methods to characterize hybridization between *Q. lobata* and *Q. douglasii*. Although Craft *et al*. (2002) could clearly distinguish the species by their allele distributions, they found evidence for low

levels of historical introgression. Ancient introgression may be more common in the California white oaks than recent gene flow, which was also suggested by Kim *et al*. (2018) in the examination of the California scrub oaks. As Kim *et al*. (2018) noted, introgression is not pervasive within the California scrub oaks, but complete reproductive isolation is non-existent. Haplotype sharing among species could lead to increased estimates of genetic diversity, thus complicating the interpretation of within-species patterns. This emphasizes the need for detailed sampling of multiple species across their collective geographic ranges (including species that are related but might not currently have overlapping ranges) to resolve questions about the importance of hybridization.

### 4.3 Phylogenetic relationships of the California clade to other white oaks

The phylogenetic trees derived from plastid (Figure 9) and nuclear ITS (Figure 10) sequence data indicated that *Q. garryana* has not segregated into distinct geographic subgroups at the infraspecific level. The cpDNA phylogeny showed two subclades within *Q. garryana*, corresponding to the clusters produced by the statistical parsimony network (Figure 6). It also showed that the other white oaks included in the study, *Q. gambelii*, *Q. lobata* and *Q. douglasii*, were not separated from *Q. garryana,* but had chloroplast sequences in common. The plastid phylogenetic signal in *Q. garryana* is likely confounded by historical introgression, contemporary gene flow or both, involving *Q. gambelii*, *Q. lobata* and *Q. douglasii* (and possibly other species).

The ITS phylogeny was even more poorly resolved, with species intermingled. Two samples of *Q. garryana* and one of *Q. douglasii* clustered together, and were positioned next to two red oak samples (*Q. kelloggii* and *Q. rubra*). Red oaks are considered quite distinct from white oaks; thus, this may reflect convergent mutations in the ITS region, or simply insufficient sequence variation to resolve these groups. The ITS region has been used extensively in molecular plant phylogenies. However, it is less than 700 bp long, and since the ribosomal gene complex (including ITS) is present in multiple copies on one or more chromosomal loci, paralogues may be present due to incomplete concerted evolution (Buckler *et al*., 1997; Alvarez and Wendel, 2003). This may confound analyses of organismal divergence and lead to phylogenetic incongruence, especially in closely related species.

Initially, the Rocky Mountain species *Q. gambelii* was selected as the outgroup in this study, but it was found to be more closely related to *Q. garryana* than suspected. Basing my selection on the study by McVay *et al*. (2017a), I anticipated that *Q. gambelii* would act as a sufficient root, especially in conjunction with other eastern North American white oaks. However, the sample of *Q. gambelii* (even though collected from its eastern distribution) was grouped with several samples of *Q. garryana* in the cpDNA phylogeny. Two of the three samples of *Q. garryana* (haplotype I) that clustered with *Q. gambelii* are on the eastern edge of the *Q. garryana* range. In the ITS phylogeny, one of these *Q. garryana* individuals was also grouped with the single *Q. gambelii* sample. *Quercus mohriana*, a south-central North American white oak, also clustered with *Q. gambelii* and five *Q. garryana* samples, suggesting ancestral or contemporary allele sharing between *Q. mohriana* and *Q. gambelii*. Although McVay *et al*. (2017a) found little evidence of gene flow between these two species and other south-central white oaks, often this can only be uncovered by extensive sampling of multiple species. My study identifies *Q. garryana* individuals that may have been the historical recipients of alleles obtained from *Q. gambelii* and other south-central and eastern white oaks. Sister to this entire cluster was *Q. robur*, a European white oak commonly planted in urban areas. Hybridization has previously been observed between *Q. robur* and native North American white oaks (Hardin, 1975). The other notable cluster in the ITS phylogeny was formed by *Q. lobata*, which has been determined using RADseq data to be the sister species of *Q. garryana* (Fitz-Gibbon *et al*., 2017). The California oaks, including *Q. garryana*, have an interesting history and may be more closely related to the central North American white oaks than previously suspected.

Because oaks are proficient at hybridizing and have few reproductive barriers, they frequently form large species complexes rather than individual species. Variable oak morphology often appears as an ecological and genetic gradient on the landscape, with hybrid zones separating parent species (Tovar-Sanchez and Oyama, 2004; McCauley *et al*., 2012). McCauley *et al*. (2012) observed morphological variation in *Q. gambelii*, *Q. turbinella*, *Q. harvardii* and a hybrid species *Q. x undulata*, forming such a gradient. This was considered to reflect hybridization predominantly in close sympatry, a circumstance probably occurring in my sample populations as well. *Quercus gambelii* showed two

modes of morphology: a high montane and a low desert mode that differed in leaf surface area and depth of lobes (McCauley *et al*., 2012). Similarly, Tovar-Sanchez and Oyama (2004) observed that morphological variation between two hybridizing species, *Q. crassifolia* and *Q. crassipes*, differed among localities and both among and within populations. The hypothesis that oaks comprise a species complex is supported by the phylogenetic relationships demonstrated by McVay *et al*. (2017b). McVay *et al*. (2017b) found that *Q. lobata* (California), *Q. gambelii* (occurring in central NA) and *Q. macrocarpa* (in eastern NA) shared introgressed alleles when analysed together using RADseq data. In their analyses, *Q. lobata* and *Q. gambelii* formed a sister clade to the eastern North American lobed white oaks. Thus, ancient hybridization and introgression may have occurred between the California tree oaks and more distant white oak groups. My study demonstrated that *Q. gambelii* and a few *Q. garryana* individuals were similar in both plastid and nuclear sequences, even though these species are now allopatric. Historical range overlap with hybridization could explain shared alleles among now widely separated species.

### 4.4 Identification of Garry oak varieties

*Quercus garryana* is generally divided into three morphologically distinct varieties: var. *garryana*, var. *breweri*, and var. *semota* (Flora North America Editorial Committee, 1993+). *Quercus garryana* was first recorded by David Douglas in the early 1800s and described in 1838 (International Plant Names Index, 2012). The two southern varieties *semota* and *breweri* were described in 1909 (Jepson, 1909; International Plant Names Index, 2012). With the availability of molecular data, taxa described on the basis of morphological traits can be assessed to see if morphological differences correspond with genetic divergence. Genetic and phenotypic variation are not necessarily correlated (Houle, 1992; Scheiner, 1993; Ghalambor *et al*., 2007; Bossdorf *et al*., 2008; Pfennig *et al*., 2010). In oaks, for example, leaf morphology is a highly variable trait with both genetic and environmental components, and illustrates how plastic species can be in appearance (Sork *et al*., 2016). As a result of the complex interactions of hereditary genes and environmentally driven epigenetic forces, it is difficult to distinguish the reliability of morphological traits. For example, Fortini *et al.* (2009), sampling five European white

oaks, compared eight micromorphological traits to five nuclear microsatellite markers and found correspondence between morphological and genetic differences. In contrast, Rodriguez-Gomez *et al*. (2018) determined that in populations of the Mexican oak *Q. deserticola*, leaf morphological variation was significantly associated with climate variation and not with cpDNA markers.

I investigated whether 1) morphological traits measured on the collected voucher specimens could be used to identify *Q. garryana* to variety, and 2) there was correspondence between morphological traits and the molecular differences found in this study. A PCA based on morphological variables did not differentiate the sampled *Q. garryana* individuals into three distinct groups corresponding to the described varieties. Instead, the three varieties seem to reflect a morphological or genetic gradient (Figure 13). There was morphological differentiation between California specimens located in the Sierra Nevada and Klamath Mountains and those from more northern populations (Figure 15); these corresponded approximately (as determined from my field observations and notes) with the shrub and tree forms of this species. The morphological traits examined in this study showed variation along a latitudinal gradient (Figure 16), although there was variation among individuals at a given latitude. Specimens from southern localities had stellate hairs with fewer, shorter rays and terminal buds that were smaller, darker and more ovoid than specimens from northern localities (Figure 17); these are morphological traits that also differentiate the two southern varieties from var. *garryana*.

I found no obvious correspondence between cpDNA patterns in this study (as indicated by main haplotype groups) and morphological variation of voucher specimens. (Figure 18). This is perhaps not surprising, as I evaluated variation in Garry oak using noncoding sequences in the chloroplast genome (which are neutral markers that are not under selection). However, morphological features are likely to be adaptive traits that are determined by nuclear genes. Thus, variation in the nuclear genome may be more likely to correspond to the morphological traits used in this study. The genetic variation found by Degner (2014) based on GBS markers showed a pattern that appeared to more closely reflect morphological differences among varieties. A common garden experiment with samples of *Q. garryana* from throughout the range would also be valuable for demonstrating the genetic basis of morphological differences in this species.

# Chapter 5.0 Conclusion

In this study I used plastid and nuclear DNA sequences and morphological data to address the following questions: What are the patterns of genetic variation within Garry oak? How do these patterns vary geographically, and how did Garry oak come to occupy its current geographical range? Does Garry oak show evidence of genetic interaction with other white oak species in western NA? Is there morphological or genetic evidence to support the three described varieties of Garry oak? I demonstrated that patterns of genetic variation within Garry oak show highest genetic diversity in southern Oregon to south-central California. The genetic patterns vary geographically with fewer haplotypes found in the northern part of its range, which is indicative of post-glacial colonization. I identified two distinct subgroups of haplotypes, each with lower diversity to the north, indicating that Garry oak likely came to occupy its current geographical range by making two separate migrations north following the LGM. I also found that Garry oak shows evidence of shared cpDNA haplotypes with other white oak species in western NA, particularly in sympatric zones with Blue oak (*Q. douglasii*) and Valley oak (*Q. lobata*), indicating hybridization among these California white oaks. Finally, morphological and genetic evidence shows some correspondence with infraspecific groupings but does not fully support the three described varieties of Garry oak. Hybridization may have contributed to the diverse morphotypes of the Garry oak varieties currently described, but I found only weak correspondence between patterns of genetic variation (as evidenced by cpDNA haplotypes) and morphological variation associated with the three Garry oak varieties. To conserve genetic variability, geographic regions that include high haplotype diversity or unique genetic variants, especially the Klamath Mountains, the North Coast Ranges and the western foothills of the Sierra Nevada, should be protected.

# Bibliography

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., mallet, J., Martinez-Rodriguez, P., Most, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Vainola, R., Wolf, J.B.W., and Zinner, D. 2013. Hybridization and speciation. Journal of Evolutionary Biology **26**(2): 229-246.

Adam, D.P., and Robinson, S.W. 1988. Palynology of two upper Quaternary cores from Clear Lake, Lake County, California. U.S. Geological Survey Professional Paper 1363. https://doi.org/10.3133/pp1363.

Allen, G.A., Marr, K.L., McCormick, L.J., and Hebda, R.J. 2012. The impact of Pleistocene climate change on an ancient arctic-alpine plant: multiple lineages of disparate history in *Oxyria digyna*. Ecology and Evolution **2**(3): 649-665.

Alvarez, I., and Wendel, J.F. 2003. Ribosomal ITS sequence and plant phylogenetic inference. Molecular Phylogenetics and Evolution **29**(3): 417-434.

Arnold, M.L. 1992. Natural hybridization as an evolutionary process. Annual Review of Ecology and Systematics **23**: 237-261.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., and Saunders, N.C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics **18**(1):489-522.

Avise, J.C. 2007. Twenty-five key evolutionary insights from the phylogeographic revolution in population genetics. *In* Phylogeography of southern European refugia: Evolutionary perspectives on the origins and conservation of European biodiversity. *Edited by* S. Weiss and N. Ferrand. Springer, Dordrecht. pp. 7-21.

Avise, J.C., Bowen, B.W., and Ayala, F.J. 2016. In the light of evolution X: comparative phylogeography. PNAS **113**(29): 7957-7961. doi/10.1073/pnas.1604338113.

Bakker, J. D., Colasurdo, L. B., and Evans, J. R. 2012. Enhancing Garry oak seedling performance in a semiarid environment. Northwest Scientific Association. **86**(4): 300-309.

Baldwin, B.G., Thornhill, A.H., Freyman, W.A., Ackerly, D.D., Kling, M.M., Morueta-Holme, N., and Mishler, B.D. 2017. Specie richness and endemism in the native flora of California. American Journal of Botany **104**(3): 487-501.

Bard, E., Hamelin, B., Fairbanks, R.G. 1990. U/Th ages obtained by mass spectrometry in corals from Barbados. Sea level during the past 130,000 years. Nature **346**: 456-458.

Barron, E., Averyanova, A., Kvacek, Z., Momohara, A., Pigg, K.B., Popova, S., Postigo-Mijarra, J.M., Tiffney, B.H., Utescher, T., and Zhou, Z.K. 2017. The fossil history of *Quercus*. *In* Tree Physiology Volume 7: Oak physiological ecology. Exploring the functional diversity of genus *Quercus* L. *Edited by* E. Gil-Pelegrin, J.J. Peguero-Pina, D. Sancho-Knapik. Springer, Switzerland. pp 39-105. https://doi.org/10.1007/978-3-319-69099-5.

Basinger, J.F., Greenwood, D.R., and Sweda, T. 1994. Early Tertiary vegetation of Arctic Canada and its relevance to paleoclimatic interpretation. NATO ASI Series **127**: 175-198.

Beatty, G.E., and Provan, J. 2010. Refugial persistence and postglacial recolonization of North America by the cold-tolerant herbaceous plant *Orthilia secunda*. Molecular Ecology **19**: 5009-5021.

Beatty, G.E., and Provan, J. 2011. Phylogeographic analysis of North American populations of the parasitic herbaceous plant *Monotropa hypopitys* L. reveals a complex history of range expansion from multiple late glacial refugia. Journal of Biogeography **38**: 1585-1599.

Belahbib, N., Pemonge, M.H., Ouassou, A., Sbay, H., Kremer, A., and Petit, R.J. 2001. Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus suber* and *Q. ilex* in Morocco. Molecular Ecology **10**(8): 2003-2012.

Bellarosa, R., Simeone, M.C., Papini, A., and Schirone, B. 2005. ITS sequence data for phylogenetic reconstruction of Italian *Quercus* spp. Molecular Phylogenetics and Evolution **34**(2): 355-370.

Bossdorf, O., Richards, C.L., and Pigliucci, M. 2008. Epigenetics for ecologists. Ecology Letters **11**(2): 106-115.

Brochmann, C., and Brysting, A.K. 2008. The Arctic – an evolutionary freezer? Plant Ecology & Diversity **1**(2): 181- 195.

Brunsfeld, S.J., Miller, T.R., and Carstens, B.C. 2007. Insights into the biogeography of the Pacific Northwest of North America: evidence from the phylogeography of *Salix melanopsis*. Systematic Botany **32**(1): 129-139.

Buckler, E.S., Ippolito, A., and Holtsford, T.P. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics Society of America **145**(3): 821-832.

Buffon, G.L.L. de Comte.1749-1767. Histoire naturelle, generale et particuliere, avec la description du Cabinet du Roi par Mrs. de Buffon and Daubenton. Vol. 1-15. Paris, France.

Castelloe, J., and Templeton, A.R. 1994. Root probabilities for intraspecific gene trees under neutral coalescent theory. Molecular Phylogenetics and Evolution **3**(2): 102-113.

Cavender-Bares, J., Gonzalez-Rodriguez, A., Eaton, D.A.R., Hipp, A.A.L., Beulke, A., and Manos, P.S. 2015. Phylogeny and biogeography of the American like oaks (*Quercus* subsection *Virentes*): a genomic and population genetics approach. Molecular Ecology **24**(14): 3668-3687.

Cavender-Bares, J., Kothari, S., Meireles, J.E., Kaproth, M.A., Manos, P.S., and Hipp, A.L. 2018. The role of diversification in community assembly of the oaks (*Quercus* L.) across the continental U.S. American Journal of Botany **105**(3): 565-586.

Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., and Zhou, S. 2016. Barcoding the kingdom Plantae: new PCR primers and ITS regions of plants with improved universality and specificity. Molecular Ecology Resources **16**(1): 138-149.

Clark, P.U., and Mix, A.C. 2002. Ice sheets and sea level of the Last Glacial Maximum. Quaternary Science Reviews **21**: 1-7.

Clement, M., Derington, J., Woolley, and Posada, D. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology **9**(10): 1657-1660.

Clements, D.R. 2011. Techniques to promote Garry oak seedling growth and survival in areas with high levels of herbivory and competition. Northwest Science **85**(2): 172-181.

Craft, K.H., Ashley, M.V., and Koenig, W.D. 2002. Limited hybridization between *Quercus lobata* and *Quercus douglasii* (Fagaceae) in a mixed stand in central coastal California. American Journal of Botany **89**(11): 1792-1798.

Crepet, W.L., and Nixon, K.C. 1989a. Earliest megafossil evidence of Fagaceae: phylogenetic and biogeographic implications. American Journal of Botany **76**(6): 842-855.

Crepet, W.L., and Nixon, K.C. 1989b. Extinct transitional Fagaceae from the Oligocene and their phylogenetic implications. American Journal of Botany **76**(10): 1493-1505.

Cutter, A.D. 2013. Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. Molecular Phylogenetics and Evolution **69**(3): 1172-1185.

Cwynar, L.C., and MacDonald, G.M. 1985. A fossil pollen based reconstruction of the late Quaternary history of lodgepole pine (*Pinus contorta* spp. *latifolia*) in the western interior of Canada. Canadian Journal of Forestry and Resources **15**: 1039-1044.

Cwynar, L.C., and MacDonald, G.M. 1987. Geographical variation of lodgepole pine in relation to population history. The American Naturalist **129**(3): 463-469.

da Silva A.G. 2017. Measuring natural selection. *In* Bioinformatics Volume I: Data, sequence analysis and evolution 2nd edition. Methods in Molecular Biology vol 1525. *Edited by* J.M. Keith. Humana Press, New York, NY. pp. 315-347.

Darwin, C.R. (1839) 1909. The voyage of the Beagle. *Edited by* C.E. Eliot. P.F. Collier & Son, New York.

de Boer, B., Van de Wal, R.S.W., Bintanja, R., Lourens, L.J., and Tuenter, E. 2010. Cenozoic global ice-volume and temperature simulation with 1-D ice-sheet models forced by benthic $\delta^{18}O$ records. Annals of Glaciology **51**(55): 23-33.

Degner, J. C. 2014. Using a genotyping-by-sequencing (GBS) approach to elucidate population structure in Garry oak (*Quercus garryana*). M.Sc. thesis, The Faculty of Forestry, The University of British Columbia, Vancouver, B.C.

Denk, T., and Grimm, G.W. 2010. The oaks of western Eurasia: traditional classifications and evidence from two nuclear markers. Taxon **59**(2): 351-366.

Denk, T., Grimm, G.W., Manos, P.S., Deng, M., and Hipp, A.L. 2017. An updated infrageneric classification of the oaks: review of previous taxonomic schemes and synthesis of evolutionary patterns. *In* Oaks physiological ecology. Exploring the functional diversity of genus *Quercus* L. Geological Survey of Canada. *Edited by* E. Gil-Pelegrin, J.J. Peguero-Pina and D. Sancho-Knapik. Tree Physiology. pp. 13-38.

Devine, W. D., and Harrington, C. A. 2013. Restoration release of overtopped Oregon white oaks increases 10-year growth and acorn production. Forest Ecology and Management **291**: 87-95.

Dunwiddie, P.W., Bakker, J.D., Almaguer-Bay, M., and Sprenger, C.B. 2011. Environmental history of a Garry oak/Douglas-fir woodland on Waldron Island, Washington. Northwest Science **85**(2): 130-140.

Eaton, D.A.R., Hipp, A.L., Gonzalez-Rodriguez, A., and Cavender-Bares, J. 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. Evolution **69**(10): 2587-2601.

Flora of North America Editorial Committee, eds. 1993+. Flora of North America North of Mexico (20+ vols.). New York and Oxford. http://floranorthamerica.org/ [Accessed 16 April 2018].

Ehlers, J., and Gibbard, P.L. 2007. The extent and chronology of Cenozoic global glaciation. Quaternary International **164-165**: 6-20.

Fitzek, E., Delcamp, A., Guichoux, E., Hahn, M., Lobdell, M., and Hipp, A.L. 2018. A nuclear DNA barcode for eastern North American oaks and application to a study of hybridization in an Arboretum setting. Ecology and Evolution **8**(11): 5837-5851.

Fitz-Gibbon, S., Hipp, A.L., Pham, K.K., Manos, P.S., and Sork, V.L. 2017. Phylogenomic inferences from reference-mapped and de novo assembled short-read sequence data using RADseq sequencing of California white oaks (*Quercus* section *Quercus*). Genome **60**(9): 743-755.

Fuchs, M.A., Krannitz, P.G., and Harestad, A.S. 2000. Factors affecting emergence and first-year survival of seedlings of Garry oak (*Quercus garryana*) in British Columbia, Canada. Forest Ecology and Management **137**(1-3): 209-219.

Galtier, N., Gouy, M., and Gautier, C. 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Bioinformatics **12**(6):543-548.

Ghalambor, C.K., McKay, J.K., Carroll, S.P., and Reznick, D.N. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Functional Ecology **21**(3): 394-407.

Gibson, J.F., Shokralla, S., Curry, C., Baird, D.J., Monk, W.A., King, I., and Hajibabaei, M. 2015. Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. PLOS ONE **10**(10): e0138432.

Goring, S., Lacourse, T., Pellatt, M.G., and Mathewes, R.W. 2013. Pollen assemblage richness does not reflect region plant species richness: a cautionary tale. Journal of Ecology **101**(5): 1137-1145.

Grimsson, F., Zetter, R., Grimm, G.W., Pedersen, G.K., Pedersen, A.K., and Denk, T. 2015. Fagaceae pollen from the early Cenozoic of west Greenland: revisiting Engler's and Chaney's Arcto-Tertiary hypothesis. Plant Systematics and Evolution **301**(2): 809-832.

Guest, H.J, and Allen, G.A. 2014. Geographical origins of North American *Rhodiola* (Crassulaceae) and phylogeography of the wester roseroot, *Rhodiola integrifolia*. Journal of Biogeography **41**: 1070-1080.

Gugger, P.F., Ikegami, M., and Sork, V.L. 2013. Influence of late Quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* Nee. Molecular Ecology **22**(13): 3598-3612.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biology **59**(3): 307-321.

Hakala, K.J., and Adam, D.P. 2004. Late Pleistocene vegetation and climate in the southern Cascade Range and the Modoc Plateau region. Journal of Paleolimnology **31**(2): 189-215.

Hardin, J.W. 1975. Hybridization and introgression in *Quercus alba*. Journal of the Arnold Aboretum **56**(3): 336-363.

Hauser, D.A., Keuter, A., McVay, J.D., Hipp, A.L., and Manos, P.S. 2017. The evolution and diversification of the red oaks of the California Floristic Province (*Quercus* section *Lobatae*, series *Agrifoliae*). American Journal of Botany **104**(10): 1581-1595.

Heiser, C.B. 1949. Natural hybridization with particular reference to introgression. Botanical Review **15**(10): 645-687.

Herendeen, P.S, Crane, P.R., and Drinnan, A.N. 1995. Fagaceous flowers, fruits, cupules from the Campanian (Late Cretaceous) of central Georgia, USA. International Journal of Plant Sciences **156**(1): 93-116.

Hewitt, G.M. 1993. Postglacial distribution and species substructures: lessons from pollen, insects and hybrid zones. *In* Evolutionary patterns and processes. *Edited by* D.R. Lees and D. Edwards. Pp. 97-123. Academic Press, London.

Hipp, A.L., Eaton, D.A.R., Cavender-Bares, J., Fitzek, E., Nipper, R., and Manos, P.S. 2014. A framework phylogeny of the American oak clade based on sequenced RAD data. PLOS ONE **9**(4): e93975.

Hipp, A.L., Manos, P.S., Gonzalez-Rodrigruez, A., Hahn, M., Kaproth, M., McVay, J.D., Valencia Avalos, S., and Cavender-Bares, J. 2018. Sympatric parallel diversification of major oak clades in the Americas and the origins of Mexican species diversity. New Phytologist **217**(1): 439-452.

Hiroshi, A. 1999. Within- and between-species DNA sequence variation and the 'footprint' of natural selection. Gene **238**(): 39-51.

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. Genetics **130**(1): 195-204.

Hubert, F., Grimm, G.W., Jousselin, E., Berry, V., Franc, A., and Kremer, A. 2014. Multiple nuclear genes stabilize the phylogenetic backbone of the genus *Quercus*. Systematics and Biodiversity **12**(4): 405-423.

Huson, D.H., and Bryant, D. 2006. Application of the phylogenetic networks in evolutionary studies. Molecular Biology and Evolution **23**(2): 254-267.

Huson, D.H., and Scornavacca, C. 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Systematic Biology **61**(6): 1061-1067.

Iddrisu, M.N., and Ritland, K. 2004. Genetic variation, population structure, and mating system in bigleaf maple (*Acer macrophyllum* Pursh). Canadian Journal of Botany **82**: 1817-1825.

International Plant Names Index. 2012. Internet http://www.ipni.org. [Accessed 20 February 2019].

Izumi, K., and Bartlein, P.J. 2016. North American paleoclimate reconstructions from the Last Glacial Maximum using an inverse modeling through iterative forward modeling approach applied to pollen data. Geophysical Research Letters **43**(20): 10965-10972.

Jepson, W.L. 1909. The trees of California. Cunningham, Curtis & Welch, San Francisco.

Julio, R. 2009. Bioinformatics for DNA sequence analysis. Methods in Molecular Biology No. 537. pp. 337-350.

Knowles, L.L., and Maddison, W.P. 2002. Statistical phylogeography. Molecular Ecology **11**(12): 2623-2635.

Keir, K.R., Bemmels, J.B., and Aitken, S.N. 2011. Low genetic diversity, moderate local adaptations and phylogeographic insights in *Cornus nutallii* (Cornaceae). American Journal of Botany **98**(8): 1327-1336.

Kelso, N.V., Patterson, T., Furno, D., Buckingham, T., Buckingham, B., Springer, N., and Cross, L. 2018. Natural Earth: raster and vector map data. https://www.naturalearthdata.com/. [Accessed 23 March 2018].

Kim, B.Y., Wei, X., Fitz-Gibbon, S., Lohmueller, K.E., Ortego, J., Gugger, P.F., and Sork, V.L. 2018. RADseq data reveal ancient, but not pervasive, introgression between California tree and scrub oak species (*Quercus* sect. *Quercus*: Fagaceae). Molecular Ecology **27**(22): 4556-4571.

Kremer, A., Abbott, A.G., Carlson, J.E., Manos, P.S., Plomion, C., Sisco, P., Staton, M.E., Ueno, S., Vendramin, G.G. 2012. Genomics of Fagaceae. Tree Genetics and Genomes **8**(3): 583-610.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution **35**(6): 1547-1549. doi.org/10.1093/molbev/msy096.

Leigh, J.W., and Bryant, D. 2015. PopART: Full-feature software for haplotype network construction. Methods in Ecololgy and Evolution **6**(9):1110–1116.

Leroy, T., Roux, C., Villate, L., Bodenes, C., Romiguier, J., Paiva, J.A.P., Dossat, C., Aury, J.M., Plomion, C., and Kremer, A. 2017. Extensive recent secondary contacts between four European white oak species. New Phytologist **214**(2): 865-878.

Lindh, B.C., McGahan, K.A., and Bluhm, W.L. 2018. Changes in the urban plant phenology in the Pacific Northwest from 1959 to 2016: anthropogenic warming and natural oscillation. International Journal of Biometeorology **62**(9): 1675-1684.

Little, E.L. 1971. Atlas of United States trees. Volume 1. Conifers and important hardwoods. Misc. Publ. 1314. U.S. Department of Agriculture, Forest Service, Washington, DC.

Lomolino, M.V., Riddle, B.R., Brown, J.H., and Whittaker, R.J. 2010. Biogeography, 4th edition. Oxford University Press,

Lucas, J.D., and Lacourse, T. 2013. Holocene vegetation history and fire regimes of *Pseudotsuga menziesii* forests in the Gulf Islands National Park Reserve, southwestern British Columbia, Canada. Quaternary Research **79**(3): 366-376.

Lyu, J., Song, J., Liu, Y., Wang, Y., Li, J., and Du, K. 2018. Species boundaries between three sympatric oak species: *Quercus aliena*, *Q. dentata*, and *Q. variabilis* at the northern edge of their distribution in China. Frontiers in Plant Science **9**(414): 1-12. doi: 10.3389/fpls.2018.00414.

Manos, P.S., Doyle, J.J., and Nixon, K.C. 1999. Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). Molecular Phylogenetics and Evolution **12**(3): 333-349.

Manos, P.S., Zhou, Z-K., Cannon, C.H. 2001. Systematics of Fagaceae: phylogenetic tests of reproductive trait evolution. Journal of Plant Science **162**(6): 1361-1379.

Manos, P.S. 2016. Systematics and biogeography of the American oaks. International Oak Journal **27**: 23.36.

Marsico, T.D., Hellmann, J.J., and Romero-Severson, J. 2009. Patterns of seed dispersal and pollen flow in *Quercus garryana* (Fagaceae) following post-glacial climatic changes. Journal of Biogeography **36**: 929-941.

Martin, K.J., and Rygiewicz, P.T. 2005. Fungal-specific PCR primers developed for analysis of the ITS region of environmental DNA extracts. BMC Microbiology **5**(28): 1-11.

McCune, B., and Mefford, M.J. 2011. PC-ORD. Multivariate analysis of ecological data. Version 7. MjM Software, Oregon, U.S.A.

McIntyre, D.J. 1991. Pollen and spore flora of an Eocene forest, eastern Axel Heiberg Island, N.W.T. *In* Tertiary fossil forests of the Geodetic Hills, Axel Heiberg Island, Arctic Archipelago. Geological Survey of Canada. *Edited by* R.L. Christie and N.J. McMillan. Minister of Supply and Services Canada. pp. 83-97.

McVay, J.D., Hipp, A.L., and Manos, P.S. 2017a. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. Proceedings of the Royal Society B **284**: 20170300. http://dx.doi.org/10.1098/rspb.2017.0300.

McVay, J.D., Hauser, D., Hipp, A.L., and Manos, P.S. 2017b. Phylogenomics reveals a complex evolutionary history of lobed-leaf white oaks in western North America. Genome **60**(9): 733-742. doi.org/10.1139/gen-2016-0206.

Michalak, J. 2011. Effects of habitat and landscape structure on Oregon white oak (*Quercus garryana*) regeneration across an urban gradient. Northwest Science **85**(2): 182-193.

Muller, M.H., Poncet, C., Prosperi, J.M., Santoni, S., and Ronfort, J. 2005. Domestication history in the *Medicago sativa* species complex: inferences from nuclear sequence polymorphism. Molecular Ecology **15**(): 1589-1602.

Nadeau, S., Godbout, J., Lamothe, M., Gros-Louis, M.-C., Isabel, N., and Ritland, K. 2015. Contrasting patterns of genetic diversity across the ranges of *Pinus monticola* and *P. strobus*: a comparison between eastern and western North American postglacial colonization histories. American Journal of Botany **102**(8): 1342-1355.

NCBI (National Center for Biotechnology Information). Standard nucleotide BLAST. Available from https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn& PAGE_TYPE=BlastSearch &LINK_LOC=blasthome [Accessed 14 July 2018].

Nevill, P.G., Bossinger, G., and Ades, P.K. 2010. Phylogeography of the world's tallest angiosperm, *Eucalyptus regnans*: evidence for multiple isolated Quaternary refugia. Journal of Biogeography **37**: 179-192.

Nixon, K.C. 2002. The oak (*Quercus*) biodiversity of California and adjacent regions. *In* Proceedings of the Fifth Symposium on Oak Woodlands: Oaks in California's Challenging Landscape. *Edited by* R.B Standiford, *et al*. Gen. Tech. Rep. PSW-GTR-184, Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture: 3-20.

Nixon, K.C. 2006. Global and neotropical distribution and diversity of oak (genus *Quercus*) and oak forests. *In* Ecology and conservation of neotropical montane oak forests. *Edited by* M. Kappelle. Springer-Verlag, Berlin, Germany pp. 3-13.

Oh, S-H., and Manos, P.S. 2008. Molecular phylogenetics and cupule evolution in Fagaceae as inferred from nuclear CRABS CLAW sequences. Taxon **57**(2): 434-451.

Oney-Birol, S., Fitz-Gibbon, S., Chen, J.-M., Gugger, P.F., and Sork, V.L. 2018. Assessment of shared alleles in drought-associated candidate genes among southern California white oak species (*Quercus* sect. *Quercus*). BMC Genetics **19**(88): 1-12. https://doi.org/10.1186/s12863-018-0677-9.

Ortego, J., Noguerales, V., Gugger, P.F., and Sork, V.L. 2015. Evolutionary and demographic history of the California scrub white oak species complex: an integrative approach. Molecular Ecology **24**(24): 6188-6208.

Ortego, J., Gugger, P.F., and Sork, V.L. 2018. Genomic data reveal cryptic lineage diversification and introgression in California golden cup oaks (section *Protobalanus*). New Phytologist **218**(2): 804-818.

Otvos, I. S., Mills, D. A., and Conder, N. 2012. Within-crown distribution, attack, and germination of *Curulio occidentis*-damaged and *Cydia latiferreana*-damaged Garry oak acorns in Victoria, British Columbia, Canada. Canadian Entomology. **144**: 419-434.

Owusu, S.A., Sullivan, A.R., Weber, J.A., Hipp, A.L., and Gailing, O. 2015. Taxonomic relationships and gene flow in four North American *Quercus* species (*Quercus* section *Lobatae*). Systematic Botany **40**(2): 510-521.

Participants of the Consortium of Pacific Northwest Herbaria. 2017-2018. University of Washington. Burke Museum of Natural History and culture, Seattle, Washington. http://www.pnwherbaria.org/index.php.

Payseur, B.A. 2010. Using differential introgression in hybrid zones to identify genomic regions involved in speciation. Molecular Ecology Resources **10**(5): 806-820.

Participants of the Consortium of California Herbaria. 2018. Regents of the University of California. *ucjeps.berkeley.edu/consortium/.*

Pei, N., Chen, B., and Kress, W.J. 2017. Advances of community-level plant DNA barcoding in China. Frontiers in Plant Sciences **8**: 225.

Pellatt, M.G., Hebda, R.J., and Mathewes, R.W. 2001. High-resolution Holocene vegetation history and climate from Hole 1034B, ODP leg 169S, Saanich Inlet, Canada. Marine Geology **174**(1-4): 211-226.

Pellatt, M. G., and Gedalof, Z. 2014. Environmental change in Garry oak (*Quercus garryana*) ecosystems: the evolution of an eco-cultural landscape. Biodiversity and Conservation **23**(8): 2053-2067.

Pfennig, D.W., Wund, M.A., Snell-Rood, E.C., Cruickshank, T., Schlichting, C.D., and Moczek, A.P. 2010. Phenotypic plasticity's impacts on diversification and speciation. Trends in Ecology & Evolution **25**(8): 459-467.

Platnick, N.I., and Nelson, G. 1984. Composite areas in vicariance biogeography. Systematic Zoology **33**(3): 328-335.

Pleines, T., Jakob, S.S., and Blattner, F.R. 2009. Application of non-coding DNA regions in intraspecific analyses. Plant Systematics and Evolution **282** (3-4):281-294.

Posadas, P., Crisci, J.V., and Katinas. 2006. Historical biogeography: a review of its basic concepts and critical issues. Journal of Arid Environments **66**: 389-403.

QGIS Development Team. 2009. QGIS Geographic Information System. Open Source Geospatial Foundation. URL http://qgis.osgeo.org.

Rambaut, A. 2016. FigTree v1.4.3. Institute of Evolutionary Biology, University of Edinburgh. http://tree.bio.ed.ac.uk/software/figtree/ [Accessed 26 January 2018].

Ritland, K., Meagher, L.D., Edwards, D.G.W., and El-Kassaby, Y.A. 2005. Isozyme variation and the conservation genetics of Garry oak. Canadian Journal of Botany. **83**: 1478-1487.

Rhode, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. Oikos **65**(3): 514-527.

Roberts, D.R., and Hamann, A. 2015. Glacial refugia and modern genetic diversity of 22 western North American tree species. Proceedings of the Royal Society B – Biological Sciences **282**(1804): 20142903. DOI: 10.1098/rspb.2014.2903.

Ronquist, F., and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572–1574.

Rosen, D.E. 1978. Vicariant patterns and historical explanation in biogeography. Systematic Zoology **27**(2): 159-188.

Roy, V., Ruel, J.-C., and Plamondon, A.P. 2000. Establishment, growth and survival of natural regeneration after clearcutting and drainage on forested wetlands. Forest Ecology and Management **129**(1-3): 253-267.

Roy, I., Ranhotra, P.S., Shekhar, M., Bhattacharyya, A., Pal, A.K., Sharma, Y.K., Singh, S.P., and Singh, U. 2018. Over-representation of some taxa in surface pollen analysis misleads the interpretation of fossil pollen spectra in terms of extant vegetation. Tropical Ecology **59**(2): 339-350.

Sandvik, B., Erle, S., and Gilles, S. 2008. Thematic mapping. <thematicmapping.org>. [Accessed 23 March 2018].

Scheiner, S.M. 1993. Genetics and evolution of phenotypic plasticity. Annual Review of Ecology and Systematics **24**: 35-68.

Sims, H.J., Herendeen, P.S., Lupia, R., Christopher, R.A., and Crane, P.R. 1999. Fossil flowers with Normapolles pollen from the Upper Cretaceous of southeastern North America. Review of Palaeobotany and Palynology **106**: 131-151.

Shafer, A.B.A., Cullingham, C.I., Cote, S.D., and Coltman, D.W. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. Molecular Ecology **19**(21): 4589-4621.

Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W.S., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E., and Small, R.L. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. American Journal of Botany **92**(1): 142-166.

Shaw, J., Lickey, E.B., Schilling, E.E., and Small, R.L. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. American Journal of Botany **94(**3): 275-288.

Shaw, J., Shafer, H.L., Leonard, O.R., Kovach, M.J., Schorr, M., and Morris, A.B. 2014. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: the tortoise and the hare IV. American Journal of Botany **101**(11): 1987-2004.

Snell, R.S., and Cowling, S.A. 2015. Consideration of dispersal processes and northern refugia can improve our understanding of past plant migration rates in North America. Journal of Biogeography **42**(9): 1677-1688.

Soltis, D.E., Ranker, T.A., and Ness, B.D. 1989. Chloroplast DNA variation in a wild plant, *Tolmiea menziesii*. Genetics **121**: 819-826.

Soltis, D.E., Soltis, P.S., Kuzoff, R.K., and Tucker, T.L. 1992. Geographic structuring of chloroplast DNA genotypes in *Tiarella trifoliata* (Saxifragaceae). Plant Systematics and Evolution **181**: 203-216.

Soltis, D.E., Gitzendanner, M.A., Strenge, D.D., and Soltis, P.S. 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Plant Systematics and Evolution **206**(1-4): 353-373.

Sork, V.L., Davis, W., Smouse, P.E., Apsit, V.J., Dyer, R.J., Fernandez-M, J.F., and Kuhn, B. 2002. Pollen movement in declining populations of California Valley oak,

*Quercus lobata*: where have all the fathers gone? Molecular Ecology **11**(9): 1657-1668.

Sork, V.L., Gugger, P.F., Chen, J-M., and Werth, S. 2016. Evolutionary lessons from California plant phylogeography. PNAS **113**(29): 8064-8071.

Stebbins, G.L., Matzke, E.B., and Epling, C. 1947. Hybridization in a population of *Quercus marilandica* and *Quercus ilicifolia*. Evolution **1**(1-2): 79.88.

Taberlet, P., Fumagalli, L., Wust-Saucy, A-G, Cossons, J-F. 1998. Comparative phylogeography and postglacial colonization routes in Europe. Molecular Ecology. **7**: 453-464.

Tovar-Sanchez, E., and Oyama, K. 2004. Natural hybridization and hybrid zones between *Quercus crassifolia* and *Quercus crassipes* (Fagaceae) in Mexico: morphological and molecular evidence. American Journal of Botany **91**(9): 1352-1363.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A, Clamp, M. and Barton, G. J. 2009. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. Bioinformatics doi: 10.1093/bioinformatics/btp033.

Whittemore, A.T., and Schaal, B.A. 1991. Interspecific gene flow in sympatric oaks. Proceedings of the Natural Academy of Science in the United States of America **88**(): 540-2544.

Wu, Y. 2011. New methods for inference of local tree topologies with recombinant SNP sequences in populations. IEEE-ACM Transactions on Computational Biology and Bioinformatics **8**(1): 182-193.

Yokoyama, Y., Esat, T.M., Thompson, W.G., Thomas, A.L., Webster, J.M., Miyairi, Y., Sawada, C., Aze, T., Matsuzaki, H., Okuno, J., Fallon, S., Braga, J.C., Humblet, M., Iryu, Y., Potts, D.C., Fujita, K., Suzuki, A., and Kan, H. 2018. Rapid glaciation and a two-step sea level plunge into the Last Glacial Maximum. Nature **559**(7715): 603-620.

Zeng, Y.-F., Liao, W.-J., Petit, R.J., and Zhang, D.-Y. 2010. Exploring species limits in two closely related Chinese oaks. PLOS ONE **5**(11): e15529.

# Appendices

**Appendix A.** Sample localities and voucher information. Locality information (collection number, latitude, longitude and elevation), locality code, and species identification of 125 sampled *Quercus* individuals. All samples were collected by Rande Kanne, except for *Q. gambelii* (sampled by Ken Marr). All voucher specimens will be deposited in the University of Victoria herbarium (UVIC).

| | Locality | Locality code | Latitude | Longitude | Elevation (m) | *Quercus* species | Haplotype |
|---|---|---|---|---|---|---|---|
| 1 | Cattle Point, BC | CPBC | 48.4383 | -123.2962 | 19 | *garryana* | A |
| 2 | Walbran Park, BC | WPBC | 48.4112 | -123.3205 | 59 | *garryana* | A |
| 3 | Sumas Mountain, BC | SMBC | 49.09894 | -122.1300 | 29 | *garryana* | A |
| 4 | Observatory Hill, BC | OHBC | 48.5204 | -123.4145 | 145 | *garryana* | A |
| 5 | Patricia Bay, BC | PBBC | 48.6558 | -123.4461 | 7 | *garryana* | A |
| 6 | Saltspring Isl., BC | SSIBC | 48.7769 | -123.4042 | 56 | *garryana* | A |
| 7 | Mill Hill Regional Park, BC | MHBC | 48.4537 | -123.4802 | 108 | *garryana* | A |
| 8 | Coopers Cove, Saseenos, BC | CCSBC | 48.3893 | -123.6514 | 16 | *garryana* | A |
| 9 | Pipers Lagoon Park, Nanaimo, BC | PLNBC | 49.2268 | -123.9474 | 15 | *garryana* | A |
| 10 | Bare Point Rd, Chemainus, BC | BPCBC | 48.9179 | -123.7022 | 33 | *garryana* | A |
| 11 | Mount Tzouhalem, Duncan, BC | MTDBC | 48.7861 | -123.6427 | 112 | *garryana* | A |
| 12 | Spruce St, Port Alberni, BC | SPABC | 49.2779 | -124.8172 | 100 | *garryana* | A |
| 13 | Helliwell Park, Hornby Isl, BC | HHIBC | 49.5179 | -124.5987 | 28 | *garryana* | A |

| | Locality | Locality code | Latitude | Longitude | Elevation (m) | *Quercus* species | Haplotype |
|---|---|---|---|---|---|---|---|
| 14 | Condensory Rd, Courtenay, BC | CCBC | 49.6958 | -125.0094 | 27 | *garryana* | A |
| 15 | Chuckanut Mountain, WA | CMWA | 48.63 | -122.4679 | 73 | *garryana* | A |
| 16 | Rotary Park, Oak Harbor, WA | ROHWA | 48.3016 | -122.6487 | 24 | *garryana* | A |
| 17 | Isthmus Beach, Indian Island, WA | IBIIWA | 48.0185 | -122.7036 | 13 | *garryana* | A |
| 18 | Shelton, WA | SWA | 47.2340 | -123.1168 | 72 | *garryana* | A |
| 19 | Elma, WA | EWA | 47.0036 | -123.4003 | 19 | *garryana* | B |
| 20 | Rochester, WA | RWA | 46.8367 | -123.1395 | 28 | *garryana* | A |
| 21 | Rainier, WA | RWA | 46.8687 | -122.6908 | 131 | *garryana* | A |
| 22 | Hwy 12, Gifford Pinchot N.F., WA | HGPFWA | 46.6881 | -121.5793 | 516 | *garryana* | A |
| 23 | Rimrock Retreat, WA | RRWA | 46.6712 | -120.9936 | 698 | *garryana* | A |
| 24 | Tieton, WA | TWA | 46.7442 | -120.7885 | 492 | *garryana* | A |
| 25 | Hwy 97, Yakima Indian Reserve, WA | HYRWA | 46.2075 | -120.4662 | 333 | *garryana* | A |
| 26 | Satus Pass Summit, WA | SPSWA | 45.984 | -120.6539 | 955 | *garryana* | A |
| 27 | Thorp, WA | THWA | 47.1059 | -120.7219 | 520 | *garryana* | A |
| 28 | Hwy 10, Yakima River, WA | HRWA | 47.1247 | -120.7391 | 546 | *garryana* | A |
| 29 | Mount Galiano, Galiano Isl., BC | MGGBC | 48.8638 | -123.3638 | 282 | *garryana* | A |
| 30 | Camas, WA | CAWA | 45.5977 | -122.3882 | 2 | *garryana* | B |
| 31 | Cooks, WA | CWA | 45.7203 | -121.6061 | 51 | *garryana* | A |

| | Locality | Locality code | Latitude | Longitude | Elevation (m) | *Quercus* species | Haplotype |
|---|---|---|---|---|---|---|---|
| 32 | Dufur, OR | DOR | 45.4507 | -121.1183 | 396 | *garryana* | A |
| 33 | White River Falls State Park, OR | WRSPOR | 45.2412 | -121.0937 | 268 | *garryana* | A |
| 34 | Pine Grove, OR | PGOR | 45.0995 | -121.3616 | 693 | *garryana* | A |
| 35 | Castle Rock, OR | CROR | 44.1507 | -122.2126 | 1124 | *garryana* | B |
| 36 | Eugene, OR | EOR | 44.1053 | -123.0617 | 129 | *garryana* | B |
| 37 | Veneta, OR | VOR | 44.0754 | -123.515 | 172 | *garryana* | A |
| 38 | Monroe, OR | MOR | 44.3137 | -123.3053 | 110 | *garryana* | A |
| 39 | Corvalis, OR | COOR | 44.5897 | -123.3555 | 221 | *garryana* | B |
| 40 | Lebanon, OR | LOR | 44.4949 | -122.9377 | 143 | *garryana* | B |
| 41 | Cascadia, OR | COR | 44.3012 | -122.5041 | 250 | *garryana* | B |
| 42 | Gates, OR | GOR | 44.7539 | -122.4138 | 293 | *garryana* | B |
| 43 | Buell Park, Buell, OR | BPBOR | 45.0222 | -123.4192 | 121 | *garryana* | B |
| 44 | Mollala, OR | MOOR | 45.1503 | -122.5653 | 148 | *garryana* | A |
| 45 | Clackamas Hwy, OR | CHOR | 45.196 | -122.1802 | 250 | *garryana* | B |
| 46 | Tualatin, OR | TOR | 45.3881 | -122.8491 | 44 | *garryana* | A |
| 47 | Kalama, WA | KWA | 46.0516 | -122.8543 | 467 | *garryana* | A |
| 48 | Shasta-Trinity N.F. (North), CA | STNCA | 40.8461 | -122.0265 | 475 | *garryana* | C |
| 49 | Tahoe N.F., CA | TNFCA | 39.1015 | -120.8534 | 844 | *garryana* | D |
| 50 | Sequoia N.F., CA | SNFCA | 35.8671 | -118.6385 | 1224 | *garryana* | E |
| 51 | Medford, OR | MEOR | 42.4659 | -122.6174 | 630 | *garryana* | A |
| 52 | Sierra N.F., CA | SNFCA | 36.8008 | -119.0891 | 1093 | *garryana* | F |
| 53 | Klamath N.F. - Yreka, CA | KFYCA | 41.847 | -122.8363 | 535 | *garryana* | G |

| | Locality | Locality code | Latitude | Longitude | Elevation (m) | *Quercus* species | Haplotype |
|---|---|---|---|---|---|---|---|
| 54 | Scatter Creek, WA | SCWA | 46.8282 | -123.0088 | 63 | *garryana* | A |
| 55 | Mt. Hood N.F., OR | MHNFOR | 45.2822 | -121.3463 | 650 | *garryana* | A |
| 56 | Salem, OR | SOR | 45.0129 | -123.1698 | 191 | *garryana* | B |
| 57 | Gifford Pinchot N.F., WA | GPNFWA | 46.5704 | -121.6836 | 467 | *garryana* | A |
| 58 | Shasta-Trinity N.F. (South), CA | STNFSCA | 40.359 | -122.9418 | 788 | *garryana* | H |
| 59 | Westfir, OR | WOR | 43.7480 | -122.508 | 334 | *garryana* | B |
| 60 | Dorena Lake, OR | DLOR | 43.7886 | -122.9571 | 288 | *garryana* | B |
| 61 | Steamboat Cr Rd and Hwy 138, OR | SCRHOR | 43.3446 | -122.7355 | 356 | *garryana* | A |
| 62 | Copeland Cr Rd, off Hwy 138, OR | CCRHOR | 43.2606 | -122.5168 | 663 | *garryana* | I |
| 63 | South of Prospect, Hwy 62, OR | SPHOR | 42.7397 | -122.5121 | 766 | *garryana* | J |
| 64 | Hwy 162, East of Rogue Elk, OR | HEREOR | 42.6631 | -122.7426 | 455 | *garryana* | J |
| 65 | North River Rd, OR | NRROR | 42.4248 | -123.1281 | 320 | *garryana* | J |
| 66 | Hwy 238 and Humbug Cr Rd, OR | HHCROR | 42.251 | -123.1481 | 399 | *garryana* | J |
| 67 | Upper Applegate Rd, OR | UAROR | 42.1091 | -123.0853 | 527 | *garryana* | K |

| | Locality | Locality code | Latitude | Longitude | Elevation (m) | *Quercus* species | Haplotype |
|---|---|---|---|---|---|---|---|
| 68 | Hwy 66 and Reitan Dr, OR | HRDOR | 42.1647 | -122.6409 | 614 | *garryana* | J |
| 69 | Lincoln, OR | LIOR | 42.1088 | -122.4119 | 1153 | *garryana* | G |
| 70 | Topsy Grade Rd, OR | TGROR | 42.1271 | -122.0378 | 1171 | *garryana* | J |
| 71 | Howard Bay Park, OR | HBPOR | 42.3133 | -121.9428 | 1259 | *garryana* | J |
| 72 | Meiss Lake Rd, CA | MLRCA | 41.905 | -122.1173 | 1302 | *garryana* | J |
| 73 | NW College Ave, Near Weed, CA | NWCANW CA | 41.4132 | -122.3955 | 1089 | *garryana* | J |
| 74 | A19, CA | ACA | 41.1022 | -121.5622 | 1000 | *garryana* | C |
| 75 | Little Valley Rd, Near Little Valley, CA | LVNLVCA | 40.9658 | -121.2669 | 1127 | *garryana* | C |
| 76 | Hwy 299, CA | HCA | 40.9969 | -121.4594 | 1037 | *garryana* | C |
| 77 | Hwy 299 and Hwy 89, CA | HHCA | 40.9371 | -121.6051 | 988 | *garryana* | C |
| 78 | Lower Jones Valley Campground, CA | LJVCCA | 40.7275 | -122.2293 | 345 | *lobata* | L |
| 79 | Holiday Rd, Shasta Lake, CA | HRSLCA | 40.7022 | -122.3387 | 268 | *lobata* | D |
| 80 | Hornbrook, CA | HOCA | 41.9120 | -122.5613 | 667 | *garryana* | K |
| 81 | Hwy 3, CA | HCA | 41.6796 | -122.7164 | 1193 | *garryana* | K |
| 82 | Scott River Rd, CA | SRRCA | 41.6355 | -122.9641 | 822 | *garryana* | K |
| 83 | Seiad Valley, CA | SVCA | 41.838 | -123.1878 | 459 | *garryana* | K |
| 84 | Etna, CA | ECA | 41.4362 | -122.9130 | 962 | *garryana* | J |

| | Locality | Locality code | Latitude | Longitude | Elevation (m) | *Quercus* species | Haplotype |
|---|---|---|---|---|---|---|---|
| 85 | Coffee Creek, CA | CCCA | 41.1089 | -122.7051 | 786 | *garryana* | M |
| 86 | Rush Creek Campsite, CA | RCCCA | 40.8187 | -122.8990 | 884 | *garryana* | N |
| 87 | Hwy 299 and turn-off to Lewiston, CA | HLCA | 40.6635 | -122.7975 | 739 | *garryana* | M |
| 88 | Clickapudi Trail, Jones Valley, CA | CTJVCA1 | 40.7386 | -122.2128 | 330 | *garryana* | L |
| 89 | Clickapudi Trail, Jones Valley, CA | CTJVCA2 | 40.7387 | -122.2115 | 335 | *garryana* | O |
| 90 | Clickapudi Trail, Jones Valley, CA | CTJVCA3 | 40.7389 | -122.2111 | 323 | *douglasii* | L |
| 91 | Near Inwood, CA | NICA1 | 40.5131 | -122.0323 | 471 | *garryana* | P |
| 92 | Near Inwood, CA | NICA2 | 40.5131 | -122.0323 | 471 | *garryana* | P |
| 93 | North of Manton, CA | NMCA | 40.4498 | -121.8772 | 632 | *garryana* | P |
| 94 | Lanes Valley Rd, CA | LVRCA | 40.3889 | -121.9595 | 421 | *lobata* | Q |
| 95 | Hwy 89, near Crescent Mills, CA | HCMCA | 40.1161 | -120.9071 | 1077 | *garryana* | I |
| 96 | Caribou Rd, CA | CRCA | 40.0583 | -121.2012 | 818 | *garryana* | I |
| 97 | Near Deadwood, CA | NDCA | 39.7464 | -121.5387 | 434 | *garryana* | R |
| 98 | Brownsville, CA | BCA | 39.4600 | -121.2917 | 670 | *garryana* | M |
| 99 | Sherwood Forest, CA | SFCA | 39.1716 | -121.1167 | 694 | *garryana* | S |

| | Locality | Locality code | Latitude | Longitude | Elevation (m) | *Quercus* species | Haplotype |
|---|---|---|---|---|---|---|---|
| 100 | Spanish Flat, CA | SPFCA | 38.8318 | -120.7963 | 605 | *garryana* | M |
| 101 | Jesus Maria, CA | JMCA | 38.2925 | -120.672 | 348 | *lobata* | T |
| 102 | Confidence Rd, CA | CRCA | 38.0261 | -120.2087 | 1214 | *garryana* | U |
| 103 | Lewis Creek Trail, CA | LCTCA | 37.4180 | -119.6264 | 1190 | *garryana* | V |
| 104 | Beal Fire Rd, CA | BFRCA | 37.0495 | -119.4100 | 1103 | *garryana* | F |
| 105 | Eshom Creek Campground, CA | ECCCA | 36.6718 | -118.9802 | 1172 | *garryana* | W |
| 106 | Hwy 190, CA | HCA | 36.147 | -118.7558 | 523 | *douglasii* | X |
| 107 | Near Camp Nelson, CA | NCNCA | 36.1462 | -118.6451 | 1267 | *garryana* | E |
| 108 | Hwy 155, E of Alta Sierra, CA | HEASCA | 35.7205 | -118.5194 | 1380 | *garryana* | E |
| 109 | Breckenridge Rd, CA | BRCA | 35.4817 | -118.5354 | 1178 | *garryana* | E |
| 110 | W of Oakville, CA | WOCA | 38.4066 | -122.44 | 187 | *garryana* | E |
| 111 | Chalk Hill Rd, CA | CHRCA | 38.5791 | -122.7670 | 100 | *garryana* | M |
| 112 | Skaggs Spring Rd, CA | SSRCA | 38.7126 | -123.0053 | 194 | *garryana* | Y |
| 113 | Scotts Valley Rd, CA | SVRCA1 | 39.127 | -122.9955 | 421 | *lobata* | Z |
| 114 | Scotts Valley Rd, CA | SVRCA2 | 39.1379 | -122.9996 | 413 | *garryana* | Z |
| 115 | Walker Rd, Near Fair Oaks, CA | WRNFOCA | 39.3756 | -122.3328 | 445 | *garryana* | Z |

| | Locality | Locality code | Latitude | Longitude | Elevation (m) | *Quercus* species | Haplotype |
|---|---|---|---|---|---|---|---|
| 116 | Near Leggett, CA | NLCA | 39.8468 | -123.6946 | 339 | *garryana* | M |
| 117 | Alderpoint Rd, CA | ARCA | 40.1112 | -123.7835 | 266 | *garryana* | M |
| 118 | Mattole Rd, Near Honeydew, CA | MRNHCA | 40.3411 | -124.0264 | 113 | *garryana* | M |
| 119 | Redwood Creek Ranch, CA | RCRCA | 40.9644 | -123.8406 | 218 | *garryana* | ZZ |
| 120 | Bald Hills, CA | BHCA | 41.2281 | -123.9678 | 694 | *garryana* | J |
| 121 | Gasquet, CA | GCA | 41.8451 | -123.9923 | 102 | *garryana* | L |
| 122 | Lake Selmac, OR | LSOR | 42.2633 | -123.5754 | 442 | *garryana* | L |
| 123 | Hwy 42, Near Winston, OR | HNWOR | 43.1087 | -123.4192 | 179 | *garryana* | A |
| 124 | Cougar Creek Rd, OR | CCROR | 43.479 | -123.4852 | 87 | *garryana* | A |
| 125 | Roxborough State Park, CO | RSCO | 39.4290 | -105.0679 | 1892 | *gambelii* | ZZZ |

**Appendix B.** Plastid primers. Primers used in this study for cpDNA PCR amplification of *Quercus* species.

| Category | Primer name | Primer sequence | Primer length (bp) |
|---|---|---|---|
| Plastid - forward | accD | AATYGTACCACGTAATCYTTTAAA | 24 |
| | ndhC | TATTATTAGAAATGYCCARAAAATATCATATTC | 33 |
| | ndhF | GAAAGGTATKATCCAYGMATATT | 23 |
| | petA | GTTATGCATGAACGTAATGCTC | 22 |
| | 5’rpl16 | GCTATGCTTAGTGTGTGACTCGTTG | 25 |
| | rpl32 | CAGTTCCAAAAAAACGTACTTC | 22 |
| | rpS16 F | AAACGATGTGGTARAAAGCAAC | 22 |
| | rpoB | CKACAAAAYCCYTCRAATTG | 20 |
| | TabA | CATTACAAATGCGATGCTCT | 20 |
| | TabE | GGTTCAAGTCCCTCTATCCC | 20 |
| | trnC$^{GCA}$ | CCAGTTCRAATCYGGGTG | 18 |
| | trnD$^{GUC}$ | ACCAATTGAACTACAATCCC | 20 |
| | trnH$^{GUG}$ | CGCGCATGGTGGATTCACAAATC | 23 |
| | trnS$^{UGA}$ | GAGAGAGAGGGATTCGAACC | 20 |
| | trnS$^{GCU}$ | AGATAGGGATTCGAACCCTCGGT | 23 |
| | trnT$^{GGU}$ | CCCTTTTAACTCAGTGGTAG | 20 |
| Plastid - reverse | 3'trnG$^{UUC}$ | GTAGCGGGAATCGAACCCGCATC | 23 |
| | psaI75 | AGAAGCCATTGCAATTGCCGGAAA | 24 |
| | psbA | GTTATGCATGAACGTAATGCTC | 22 |
| | psbD | CTCCGTARCCAGTCATCCATA | 21 |
| | psbJ | ATAGGTACTGTARCYGGTATT | 21 |
| | 3’rpl16 | CCCTTCATTCTTCCTCTATGTTG | 22 |
| | rpl32 | CCAATATCCCTTYYTTTTCCAA | 22 |
| | rpS16 R | AACATCWATTGCAASGATTCGATA | 24 |

| | | |
|---|---|---|
| TabB | TCTACCGATTTCGCCATATC | 20 |
| TabF | ATTTGAACTGGTGACACGAG | 20 |
| $trnC^{GCA}$ | CACCCRGATTYGAACTGGGG | 20 |
| $trnfM^{CAU}$ | CATAACCTTGAGGTCACGGG | 20 |
| trnG2S | TTTTACCACTAAACTATACCCGC | 23 |
| $trnL^{UAG}$ | CTGCTTCCTAAGAGCAGCGT | 20 |
| $trnQ^{UUG}$ | CTATTCGGAGGTTCGAATCCTTCC | 24 |
| $trnT^{GGU}$ | CTACCACTGAGTTAAAAGGG | 20 |
| $trnV^{UAC}$ | CCGAGAAGGTCTACGGTTCG | 20 |
| ycf6 R | GCCCAAGCRAGACTTACTATATCCAT | 26 |

**Appendix C.** Nuclear ITS primers. Primers used in this study for nrDNA PCR amplification of *Quercus* species.

| **Category** | **Primer name** | **Primer sequence** | **Primer length (bp)** |
|---|---|---|---|
| ITS - forward | 1406 | TGTACACACCGCCCGT | 17 |
| | NNC18S-10 | AGGAGAAGTCGTAACAAG | 18 |
| | ITS 1 | TCCGTAGGTGAACCTGCGG | 19 |
| | ITSp5 | CCTTATCAYTTAGAGGAAGGAG | 22 |
| | ITSu1 | GGAAGKARAAGTCGTAACAAGG | 22 |
| | NSIP | GATTGAATGATCCGGTGAAG | 20 |
| ITS - reverse | 307 | TTGGGCTGCATTCCCA | 17 |
| | C26A | GTTTCTTTTCCTCCGCT | 17 |
| | ITS 4 | TCCTCCGCTTATTGATATGC | 20 |
| | ITSp4 | CCGCTTAKTGATATGCTTAAA | 21 |
| | ITSu4 | RGTTTCTTTTCCTCCGCTTA | 19 |
| | NLBP | GCTGTCACCCTCTCAGGC | 18 |

**Appendix D.** Fagaceae species retrieved from GenBank and used in the plastid and nuclear ITS phylogenetic analyses.

| Category | Species | Accession number |
|---|---|---|
| Plastid – whole genome | *Castanea mollissima* | HQ336406.1 |
| | *Castanopsis hainanensis* | NC_037389.1 |
| | *Lithocarpus balansae* | KP299291.1 |
| | *Quercus aliena* | KU240007.1 |
| | *Quercus dentata* | MG967555.1 |
| | *Quercus robur* | LT799008.1 |
| | *Quercus rubra* | JX970937.1 |
| Nuclear - ITS | *Castanea mollissima* | AY040396.1 |
| | *Castanopsis echinocarpa* | AY040375.1 |
| | *Lithocarpus balansae* | EU195795.1 |
| | *Quercus alba* | EU646136.1 |
| | *Quercus aliena* | KX838217.1 |
| | *Quercus dentata* | KM051445.1 |
| | *Quercus kelloggii* | AF098416.1 |
| | *Quercus lobata* | AF098422.1 |
| | *Quercus michauxii* | AY456175.1 |
| | *Quercus mohriana* | KM200971.1 |
| | *Quercus montana* | AY040484.1 |
| | *Quercus robur* | EU628560.1 |
| | *Quercus rubra* | AF098418.1 |
| | *Quercus serrata* | JF980309.1 |