View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Helsingin yliopiston digitaalinen arkisto

GENETIC DIVERSITY AND ITS EXTENDED EFFECTS - FRAGMENTED PEDUNCULATE OAK (*QUERCUS ROBUR*) POPULATIONS IN SOUTH-WESTERN FINLAND

Tähti Pohjanmies
Master's thesis
University of Helsinki
Department of Agricultural Sciences
Agroecology
January 2014

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department
Faculty of Agriculture and Forestry	Department of Agricultural Sciences

Tekijä — Författare — Author
Tähti Pohjanmies

Työn nimi — Arbetets titel — Title
Genetic diversity and its extended effects – fragmented pedunculate oak (*Quercus robur*) populations in south-western Finland

Oppiaine —Läroämne — Subject
Plant Production Sciences, Agroecology

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Master's thesis	January 2014	34 p.

Tiivistelmä — Referat — Abstract

Genetic variation within a population is shaped by the life history traits of the species and the properties of the surrounding ecosystem. It is an important factor in the preservation of populations. According to the emerging field of community genetics, genetic variation within a population of one species may also influence the dynamics and diversity of associated species, extending the conservational relevance of intraspecific genetic diversity.

Finnish populations of pedunculate oak (*Quercus robur*) offer an interesting study system for population genetics. *Q. robur* grows in south-western Finland at the northern limit of its natural range. Here, its distribution has been shaped by long-term climatic and geological changes as well as by human disturbance, and the current populations are small and strongly fragmented. As *Q. robur* supports a high diversity of associated species, it is considered to have great ecological and conservational importance.

In this thesis, I studied the amount and distribution of genetic diversity within and among three *Q. robur* populations in south-western Finland using population genetic parameters. I also described the spatial and temporal sub-population structure of one population, on the island of Wattkast. The genetic data was based on 15 nuclear microsatellite loci. Additionally, I examined the effect of the genetic diversity and genotypic identity of the oaks within Wattkast on associated herbivore communities. In the analysis, I used observational data from two years.

As predicted for widespread, long-lived tree species, the microsatellite loci showed high levels of diversity within the populations, but also significant differentiation among them. This may be due to fragmentation and to the marginality of the populations. Within the population on Wattkast, I observed patterns of spatial and temporal sub-population differentiation. The characteristics of the site, including the ongoing shift to less extensive land use, suggest that the population is in genetic disequilibrium. As both the genetic distance and the community dissimilarity between pairs of trees increased with increasing geographic distance, I could not conclude the genotypic identity of the host trees to have an effect on the herbivore community structure. However, higher heterozygosity was associated with higher richness and abundance of species. This result supports the notion that intraspecific genetic variation may increase associated species richness.

Based on the results of my study, both the life history traits of the species and the historic habitat changes may be observed in the genetic structure of *Q. robur* populations in Finland. The results also suggest that preservation of genetic variation within the remaining stands may be a factor not only in the preservation of these populations, but also in the conservation of associated species diversity.

Avainsanat — Nyckelord — Keywords
Quercus robur, population genetics, microsatellites, genetic diversity, community genetics

Säilytyspaikka — Förvaringsställe — Where deposited
Department of Agricultural Sciences and Viikki Campus Library

Muita tietoja — Övriga uppgifter — Further information
Supervisor(s): Tomas Roslin

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department
Maatalous-metsätieteellinen tiedekunta	Maataloustieteiden laitos

Tekijä — Författare — Author
Tähti Pohjanmies

Työn nimi — Arbetets titel — Title
Geneettinen monimuotoisuus ja sen laajennetut vaikutukset – pirstoutuneet tammipopulaatiot Lounais-Suomessa

Oppiaine —Läroämne — Subject
Kasvintuotantotieteet, agroekologia

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Maisterintutkielma	Tammikuu 2014	34 s.

Tiivistelmä — Referat — Abstract
Populaation sisäisen perinnöllisen muuntelun määrään vaikuttavat lajin elinkierron piirteet sekä ympäröivän ekosysteemin ominaisuudet. Perinnöllinen muuntelu on tärkeää populaation säilymisen kannalta. Genetiikkaa, evolutiikkaa ja yhteisöekologiaa yhdistelevän teorian mukaan yhden lajin sisäinen perinnöllinen muuntelu voi vaikuttaa myös assosioituneiden lajien dynamiikkaan, mikä laajentaa lajinsisäisen geneettisen monimuotoisuuden suojelullista merkitystä.

Suomalaiset tammipopulaatiot tarjoavat hyvän kohteen populaatiogeneettiselle tutkimukselle. Metsätammi (*Quercus robur*) kasvaa Lounais-Suomessa luontaisen esiintymisalueensa pohjoisrajalla. Pitkäaikaiset ilmastolliset ja geologiset muutokset sekä viime vuosisatoina ihmisen toiminta ovat muovanneet lajin esiintymispaikkoja voimakkaasti, ja nykyiset populaatiot ovat verraten pieniä ja pirstoutuneita. Tammella elää suuri kirjo muita lajeja.

Tässä työssä tutkin perinnöllisen muuntelun määrää ja jakautumista kolmen lounais-suomalaisen tammipopulaation sisällä ja välillä käyttäen populaatiogeneettisiä tunnuslukuja. Kuvailin myös yhden populaation (Wattkastin saarella) sisäistä geneettistä rakennetta tutkimalla muuntelun jakautumista populaation sisällä sekä tilassa että ajassa. Geneettinen aineisto oli määritelty 15 mikrosatelliittilokuksen perusteella. Lisäksi tutkin Wattkastin saarella kasvavien puiden geneettisen monimuotoisuuden ja genotyypin vaikutusta puilla eläviin kasvinsyöjä-yhteisöihin. Analyysissä käytin havaintoaineistoa kahdelta vuodelta.

Havaitsemani populaatioiden sisäinen geneettinen monimuotoisuus oli korkeaa, mikä on tyypillistä pitkäikäiselle, tehokkaasti leviävälle lajille. Toisaalta populaatioiden välinen eriytyminen oli myös merkittävää. Tämä selittynee suomalaisten tammipopulaatioiden pirstoutuneisuudella ja sijainnilla lajin levinneisyysalueen reunalla. Wattkastin populaatiossa havaitsin populaation sisäistä maantieteellistä ja ajallista eriytymistä. Lajin ominaisuuksien sekä alueen piirteiden perusteella pidän todennäköisenä, että populaatio on geneettisessä epätasapainossa. Koska sekä puiden väliset geneettiset etäisyydet että niiden hyönteisyhteisöjen erilaisuus kasvoivat maantieteellisten etäisyyksien kasvaessa, en voinut päätellä isäntäpuiden genotyypillä olevan vaikutusta kasvinsyöjäyhteisöiden rakenteeseen. Puiden heterotsygotian, joka kuvaa yksilön geneettistä monimuotoisuutta, sen sijaan havaitsin lisäävän niillä esiintyvien kasvinsyöjälajien monimuotoisuutta ja yksilörunsautta.

Tulosteni perusteella sekä lajin elinkierron piirteet että elinympäristöjen historialliset muutokset ovat havaittavissa suomalaisten tammipopulaatioiden geneettisessä rakenteessa. Lisäksi tulokseni viittaavat siihen, että kantojen sisäisen perinnöllisen muuntelun säilymisellä voi olla merkitystä paitsi tammipopulaatioiden myös niillä elävien lajien monimuotoisuuden suojelussa.

Avainsanat — Nyckelord — Keywords
Quercus robur, populaatiogenetiikka, mikrosatelliitit, geneettinen monimuotoisuus

Säilytyspaikka — Förvaringsställe — Where deposited
Maataloustieteiden laitos ja Viikin kampuskirjasto

Muita tietoja — Övriga uppgifter — Further information
Työtä ohjasi(vat) Tomas Roslin

Contents

1 Introduction

The amount of neutral genetic variation within and among populations is shaped by population genetic processes such as random genetic drift and gene flow. The nature of these processes is determined by the life history traits of the species and the properties of the surrounding ecosystem. In small, isolated populations, genetic drift is strong and gene flow limited, which may lead to pronounced genetic differentiation of populations and loss of genetic variation (Allendorf and Luikart, 2007). Human land use affects the population dynamics of organisms by changing the size and connectivity of habitat patches. Loss of genetic variation due to reduced population size and increased isolation are among the consequences of habitat fragmentation, which as a whole is one of the leading causes of biodiversity loss (e.g. Pimm and Raven, 2000).

Genetic variation is a prerequisite for the evolutionary adaptation of organisms and thereby an important factor in the preservation of populations (Allendorf and Luikart, 2007). Higher levels of heterozygosity, a measure of genetic diversity, are also typically associated with higher population fitness (Reed and Frankham, 2003). Furthermore, according to the theory of community genetics, genetic variation within a population of one species may influence the dynamics and diversity of associated species - a notion for which support has been found in several empirical studies (Hersch-Green et al., 2011, and citations therein). If the genetic variation within a basal species significantly affects associated species diversity, its preservation becomes a factor in the conservation of all of those species (e.g. Wimp et al., 2004).

Oaks (*Quercus* spp.) are characterized by high levels of genetic variation both within and among species (Kremer and Petit, 1993; Curtu et al., 2007), and higher genetic variation within than among populations has been observed in several *Quercus* taxa (Kremer and Petit, 1993; Lind and Gailing, 2013). More genetic variation within than among populations is a typical pattern for long-lived woody species with large geographic ranges, outcrossing breeding systems, and wind-mediated pollination (Hamrick et al., 1992), and these species may be relatively resilient to loss of genetic variation even in fragmented habitats (Hamrick, 2004).

Due to their biology, phylogeography, and large range, *Quercus* species may indeed be relatively resistant to fragmentation. For example, no significant effect of landscape structure on chloroplast DNA (cpDNA) diversity of four different *Quercus* species was found in France (Petit et al., 2002b). However, reductions in the allelic diversity of young trees in fragmented forest areas were observed in *Q. humboldtii*

in Colombia (Fernández-M. and Sork, 2007). Many tree species exhibit enhanced pollen immigration into isolated populations; however, less is known about the impact of landscape change on seed flow between populations and fragments (Sork and Smouse, 2006). Isolation and reduction in sources of gene flow may increase the level of inbreeding and the risk of future genetic bottlenecks also in fragmented tree populations (see also Honnay and Jacquemyn, 2007).

Oaks may be considered foundation species in the temperate forests of the northern hemisphere, and the ongoing decline in their abundance has caused concern for the preservation of associated species and ecosystem functions (Ellison et al., 2005; Lindbladh and Foster, 2010). Pedunculate oak (*Quercus robur*) is the only oak species occurring naturally in Finland. It supports high levels of associated biodiversity (Southwood, 1961), including several red-listed specialist species (Rassi et al., 2010). The current distribution of *Q. robur* in Finland is a result of climatic and geological changes after the last glacial period (Ferris et al., 1998) and, more recently, of human land and resource use. As a result, the present populations are relatively small and strongly fragmented (Ollinmaa, 1952; Vakkari et al., 2006). There has been a considerable decline in the abundance of *Quercus* species in southern Scandinavia over the last 4000 years, and, with the expansion of forestry and agriculture, the decline has been particularly rapid since the 18th century (Lindbladh and Foster, 2010).

Q. robur grows in south-western Finland at the northern limit of its natural range. Genetic diversity is hypothesized to be typically lower and differentiation higher in populations at the margin of a species' geographic range compared with geographically central populations. This pattern has been repeatedly observed in real populations, but the causes and consequences of the phenomenon have been explored less extensively (Eckert et al., 2008). Marginal populations may also diverge from central populations and become reservoirs of distinct genetic variations with potential importance for genetic conservation (Lesica and Allendorf, 1995).

Based on cpDNA data, among-population differentiation in *Q. robur* has been found to be higher in northern Europe than central Europe, potentially due to fragmentation of oak forests but also to the absence of *Q. petraea* and consequent interspecific gene flow in much of the region (Petit et al., 2002a). *Q. petraea* is also absent from all of Finland. Chloroplast DNA variation of *Q. robur* populations in Finland is on average lower and differentiation among populations higher than in central Europe (Vakkari et al., 2006). Additionally, comparably small population sizes and the location at the margin of the natural range is suggested to magnify the effects of fragmentation on the genetic structure of Finnish oak populations (Vakkari et al., 2006).

The objective of this thesis is to examine the genetic diversity of these marginal *Quercus robur* populations and the relevance for a community genetics perspective in this context. To achieve this, I examine the distribution of genetic variation of pedunculate oak in south-western Finland at two hierarchical scales: 1) its spatial distribution between three populations, and 2) its fine-scale spatial and temporal distribution within one population. Within one population, I investigate the effect of the genetic diversity and genotypic identity of host trees on the structure of associated herbivore communities.

2 Materials and methods

2.1 Study populations

In Finland, *Quercus robur* occurs naturally near the coast in the south-western part of the country (Ferris et al., 1998). After colonizing Finland following the last glacial period, the species has decreased in range due to climatic changes and competition from Norway spruce (Vakkari et al., 2006). Oaks are wind-pollinated and highly self-incompatible with a high pollen-dispersal ability (Ducousso et al., 1993). Acorns are dispersed over short distances by small mammals, and over long distances particularly by jays.

The three populations chosen for this study are located in south-western Finland: on the mainland in Inkoo and Salo, and on the island of Wattkast in Korppoo (Figure 2.1). The population in Inkoo is a scattered stand of approximately 100 trees on the southern shoreline of the Torplandet peninsula in the lake Brukträsket, covering an area of approximately 1000 m by 200 m. The Salo population is a stand of nearly 300 trees on the southern and south-eastern slope of Linnamäki hill, encompassing an area of approximately 600 m by 50 m between dry rock and cultivated fields.

The island of Wattkast has an area of approximately 5 km^2. The location and size of 1868 oak trees on the island have been mapped (encompassing all trees with a height of at least 50 cm in 2003 - 2004; see Gripenberg and Roslin, 2005). Within the island, the distribution of the oaks seems to be limited by the trees' dispersal ability, not by lack of suitable habitats (Kettunen, 2010).

In Wattkast and Salo, a set of 194 and 35 trees, respectively, were sampled for genotyping. In Inkoo, all trees with diameter at breast height (DBH) > 5 cm ($n = 96$) were sampled. In Inkoo and Salo, samples were collected in the autumn of 2004 by Pekka Vakkari, and in Wattkast in September 2011 by members of the Spatial Food Web Ecology group, University of Helsinki.

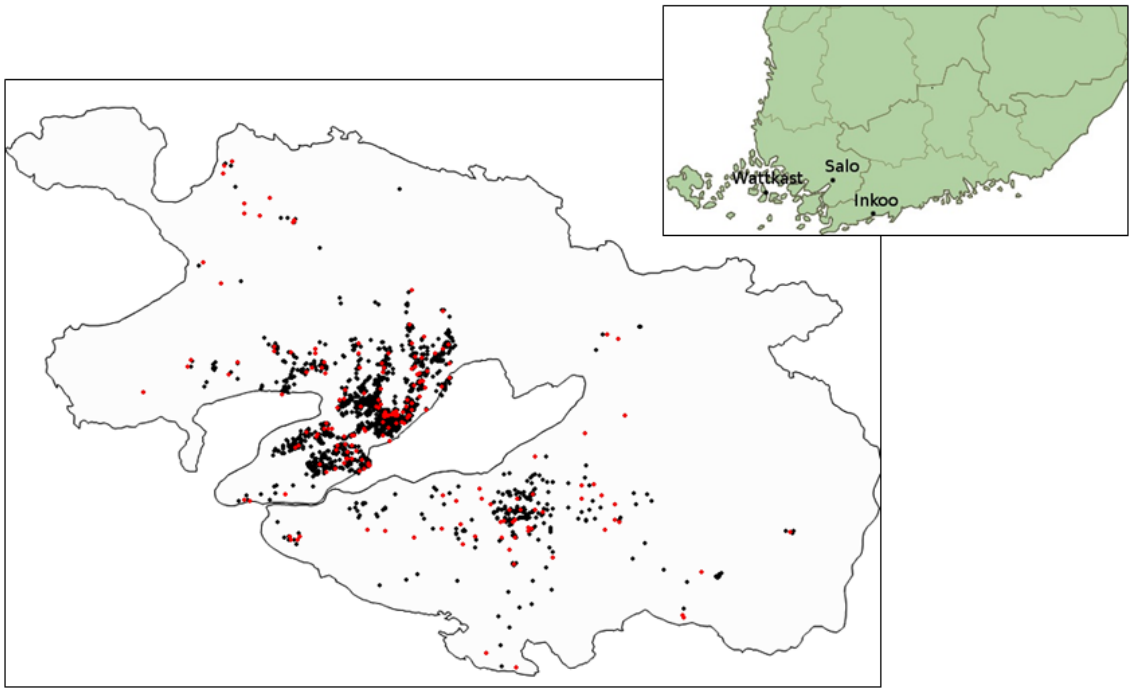

Figure 2.1: Locations of the three study populations in south-western Finland (small map), and of individual oaks within Wattkast (large map). All oak trees in Wattkast are marked on the map with black dots, and the 194 trees sampled for genotyping with red dots.

2.2 Genotyping

Sampled individuals were genotyped using 15 nuclear microsatellite loci (Dow et al., 1995; Steinkellner et al., 1997; Kampfer et al., 1998) by Sakina Elshibli, Finnish Forest Research Institute. DNA was extracted from leaf samples using the E.Z.N.A.™ extraction protocol (E.Z.N.A.™ SP Plant Mini Kit D5511-02; Omega Bio-Tek Inc., Norcross, GA, USA). Twelve pairs of primers were used to amplify the microsatellite loci in the extracted DNA. The PCR amplification reactions were performed in a volume of 13 μL. Four different multiplex reaction mixtures were made using QIAGEN Multiplex PCR Kit. Forward primers were fluorescently labeled with Beckman's WeI1RED (Beckman Coulter Inc., Brea, CA, USA). Genotyping was carried out using a CEQ 8000 genetic analyzer (Beckman Coulter Inc.); 1 μl of amplified DNA and 24 μl of DNA Size Standard were loaded into 96-well plates. Allele visualization and scoring were performed using the fragment analysis module CEQ 8000 software (Beckman Coulter Inc.).

2.3 Genetic differentiation between Wattkast, Inkoo and Salo

To quantify the levels of genetic diversity within the focal populations, the number of alleles detected (A), expected heterozygosity (H_e), and observed heterozygosity (H_o) were calculated for each locus and each population in Arlequin v3.5.1.3 (Excoffier and Lischer, 2010). Exact tests of Hardy-Weinberg equilibrium (Guo and Thompson, 1992) were performed to evaluate and compare deviation from its conditions among the three populations.

Measures of allelic richness can be highly dependent of sample size, especially when using microsatellites (Leberg, 2002). When the loci studied are highly polymorphic, less rare alleles are detected in small samples, and the estimated allelic richness is biased downward. For this reason, the estimates of allelic richness (numbers of alleles per locus) were corrected for unequal sample sizes by the rarefaction method of Hurlbert (1971) as described by Kalinowski (2004). According to this method, the expected number of alleles in a random sample of n genes selected from a sample of N genes is

$$\hat{A}_n = \sum_{i=1}^{m} [1 - \frac{\binom{N-a_i}{n}}{\binom{N}{n}}], \tag{2.1}$$

where a_i = the number of copies of the ith allele in the sample, and m = the total number of alleles detected. Because the sample size was smallest in the Salo population ($n = 35$), the method was used to calculate estimates of the expected number of alleles in each locus in a random sample of 35 individuals from the Wattkast and Inkoo populations.

Genetic differentiation between the three populations was partitioned with a standard analysis of molecular variance (AMOVA; Excoffier et al., 1992) using Arlequin with default settings. The significance of the fixation index (F_{ST}) was tested by a comparison of the observed value to those generated by 10,000 permutations of individual genotypes among the populations. 10% missing data per locus was allowed.

Population differentiation was further examined with a Bayesian clustering analysis implemented in Structure 2.3.4 (Pritchard et al., 2000). The method implemented by the program detects genetic clusters in the data, and determines the probable ancestry of each individual in these clusters based on allele frequencies. The correlated allele frequency model (Falush et al., 2007), which assumes that allele frequencies in different populations are correlated, and an admixture model, allowing for mixed ancestry among individuals, were applied. These modeling choices are supposed to improve detection of distinct clusters also when the populations are closely related (Pritchard et al., 2010). A LOCPRIOR model (Hubisz et al., 2009),

which integrates informative data about sampling location (in this case the population of origin in the data set) into the clustering process, was also applied. The analysis was repeated for different values of K (assumed number of populations, 1 to 10) for ten times with a burn-in length of 10,000 and for 100,000 steps after the burn-in.

As described by Pritchard et al. (2010), inferring the correct number of distinct clusters, i.e. the real value of K, from the results of STRUCTURE simulations is not entirely straightforward. Thus, in addition to the values of $\ln P(X|K)$, the patterns in which individuals were assigned to different populations were used in interpreting which value of K best describes the structure in the data. As a supplementary method to identify the most likely number of genetic clusters, rate of change values (ΔK) were calculated according to Evanno et al. (2005). ΔK is an *ad hoc* statistic based on the second order rate of change of the likelihood function of the data estimated by STRUCTURE. It is expected to show a modal value at the correct K. ΔK is calculated as

$$\Delta K = \frac{m|L''(K)|}{s[L(K)]}, \tag{2.2}$$

where *L(K)* refers to the posterior probability of data for K (same as $\ln P(X|K)$), $m|L''(K)|$ refers to the mean of the absolute values of $L''(K)$ averaged over independent runs of the analysis, and the denominator $s[L(K)]$ refers to the standard deviation of $L(K)$ (i.e. the standard deviation from the mean over all runs).

2.4 Genetic structure of the Wattkast population

To identify genetic sub-structuring within the Wattkast population, a Bayesian clustering analysis was performed using STRUCTURE with the same settings as described above. The simulation was run for different values of K ranging from 1 to 10 for ten times with a burn-in length of 10,000 and for 100,000 steps after the burn-in. Rate of change values (ΔK) were calculated for each value of K. To explore the factors contributing to the observed structure, the results of the simulation were further analyzed by logistic regression, with the probability of an individual belonging to one of the groups as the response, and individual size and location (geographic coordinates and their interaction) as the explanatory variables.

Fine-scale spatial genetic structure within the population was examined by linear regressions of pairwise relatedness coefficients on geographic distances computed with SPAGeDi v.1.4 (Hardy and Vekemans, 2002). The analysis was performed over 10 distance classes, with the maximal distance of each class defined so that the number of pairwise comparisons within each distance class is approximately constant. The upper limits of the distance classes ranged from 162 m to 2863 m,

increasing by approximately 100-150 m for classes 1-8 and more steeply for the last two classes. For comparison, acorns from *Q. robur* have been observed to be commonly dispersed over distances of more than 500 m (Olrik et al., 2012. The kinship coefficient according to J. Nason (described in Loiselle et al., 1995) was used as the measure of genetic relatedness in the analysis. The significance of the coefficients computed was tested by comparing observed values to those expected based on 10,000 permutations of individual locations among all individuals.

The effect of the landscape context on the individual-level inbreeding was examined by calculating correlation estimates between measures of individual inbreeding and connectivity. Individual connectivity was described with a connectivity metric (S250) adopted from Tack et al. (2010), as well as with the simple sum of distances to other oaks on the island. The S250 value for an individual tree is calculated based on the distances to other trees and the sizes of the other trees, so that it is increased by proximity to other trees and proximity to large trees in particular (for details, see Tack et al., 2010). Individual inbreeding was described by uncorrected homozygosity (H_o) and internal relatedness (IR). H_o is simply the proportion of homozygous loci, whereas IR incorporates allele frequencies into the measure. IR was calculated for each individual as

$$IR = \frac{(2H - \sum f_i)}{(2N - \sum f_i)}, \tag{2.3}$$

where H is the number of homozygous loci, N is the number of loci, and f_i is the frequency of the ith allele contained in the genotype (as described by Aparicio et al., 2006). Values of IR can vary between 1 and -1, with negative IR values indicating higher heterozygosity and positive values higher homozygosity. Allele frequencies were calculated using Genepop 4.2 (Rousset, 2008).

Finally, to test for differences in genetic diversity and heterozygosity between the young and old generations of oaks on the island, correlations between tree size and the two measures of individual inbreeding were calculated, and an AMOVA between juvenile and adult trees was performed. Here, a loss of genetic diversity in the young generation and/or excess of heterozygotes in the old generation would potentially indicate selection against certain genotypes or homozygotes within the population. Division of the sampled trees into juveniles and adults was based on the size distribution of the trees, with the limit set at a circumference of 15 cm. This resulted in sample sizes of $n = 103$ for the juvenile trees and $n = 91$ for the adult trees. The analysis was implemented in Arlequin, and the significance of the fixation index was tested with 10,000 permutations of individual genotypes among populations. All correlations were calculated in R v.2.15.3 (R Core Team, 2013).

2.5 Surveys of herbivore communities

The presence and abundance of more than twenty insect herbivore species on the oaks of Wattkast have been annually surveyed for more than ten years by the Spatial Food Web Ecology group, University of Helsinki. In the surveys, data is collected on a comprehensive random sample of trees which, at the initiation of the surveys, were small enough for their entire canopy to be examined. For each tree, twenty primary shoots are selected randomly, and the number of leaves in the shoot and the presence and abundance of each herbivore species is recorded. Within this sample of trees, 100 are the same as in the sample used for the genetic analysis here. Data on these 100 trees from two years, 2006 and 2007, were used in this study. 21 of the species surveyed were observed within this sample of 100 trees (Table 2.1). Given large variation in abundance between generations of gallers, the sexual and asexual generations were treated as separate for *Andricus pseudoinflator* and *Neuroterus quercusbaccarum* taxa.

Table 2.1: List of insect herbivore species encountered in a sample of 100 oak trees on Wattkast in 2006 and 2007, and their feeding mode (from Tack et al., 2010).

Order	Family	Genus and species	Feeding mode
Diptera	Cecidomyiidae	*Macrodiplosis dryobia*	Galler
Homoptera	Triozidae	*Trioza remota*	Galler
Hymenoptera	Cynipidae	*Andricus callidoma*	Galler
Hymenoptera	Cynipidae	*Andricus curvator*	Galler
Hymenoptera	Cynipidae	*Andricus pseudoinflator*	Galler
Hymenoptera	Cynipidae	*Cynips longiventris*	Galler
Hymenoptera	Cynipidae	*Neuroterus numismalis*	Galler
Hymenoptera	Cynipidae	*Neuroterus quercusbaccarum*	Galler
Hymenoptera	Tenthredinidae	*Profenusa pygmaea*	Miner
Lepidoptera	Coleophoridae	*Coleophora* spp.	Miner
Lepidoptera	Eriocraniidae	*Dyseriocrania subpurpurella*	Miner
Lepidoptera	Gracillariidae	*Caloptilia alchimiella*	Miner
Lepidoptera	Gracillariidae	*Phyllonorycter* spp.	Miner
Lepidoptera	Nepticulidae	*Ectoedemia albifasciella*	Miner
Lepidoptera	Nepticulidae	*Stigmella* spp.	Miner
Lepidoptera	Tischeriidae	*Tischeria dodonea*	Miner
Lepidoptera	Tischeriidae	*Tischeria ekebladella*	Miner
Homoptera	Asterolecaniidae	*Asterolecanium variolosum*	Other
Lepidoptera	Bucculatricidae	*Bucculatrix ulmella*	Other
Lepidoptera	Heliozelidae	*Heliozela sericiella*	Other
Lepidoptera	Tortricidae	*Ancylis mitterbacheriana*	Other

2.6 Association between host plant genotype and herbivore community structure

The effect of the heterozygosity of host trees on their herbivore community structure was analyzed by four separate sets of analyses. First, generalized linear regression models were constructed to test for an effect of the internal relatedness (IR) of a tree on four descriptors of its herbivore community structure: total species richness, total herbivore abundance, leaf miner abundance, and galler abundance. A Poisson distribution was assumed for each response variable. Species diversity was further described by the Shannon-Wiener index, and the effect of IR on these values was examined by standard linear regression.

Second, the effect of IR on species-specific abundance was tested with Poisson regression.

Third, a generalized linear mixed model (GLMM) was constructed of species-specific abundances as a function of species identity, IR, and their interaction as fixed effects, and tree identity as a random effect. A Poisson distribution was assumed for the species abundances. The model was built and fitted using the SAS software (SAS Institute, Cary, NC, USA).

Given that the size of the tree and the sampling intensity may influence the number and abundance of species present and encountered, the number of leaves examined during the surveys was included in all of the models above as an explanatory variable.

Fourth, a potential association between the genotypic identity of the host and its herbivore community structure was examined by Mantel tests of the correlation between the geographic distance, the genetic distance, and the dissimilarity of herbivore communities between individual trees. Geographic and genetic distance matrices were computed with SPAGeDi, with Rousset's $\hat{a}$ (Rousset, 2000) as the measure of genetic distance between individuals. For building the community dissimilarity matrices, data on species abundances were transformed into relative abundances by dividing each species-specific abundance with the total herbivore abundance observed in the tree. The community dissimilarity measure used was the Bray-Curtis index (Faith et al., 1987). The community dissimilarity matrices were computed using the vegan package for R (Oksanen et al., 2013). Trees in which no species had been observed were excluded from the data, as empty sites may lead to erroneous results (Oksanen et al., 2013), leaving $n = 92$ trees in 2006 and $n = 93$ trees in 2007. Mantel tests between community dissimilarity and geographic distance, between community dissimilarity and genetic distance, and a partial Mantel test between all three matrices were computed using the vegan package for R. The significances of the computed

Mantel statistics were tested with 1,000 permutations of the community dissimilarity matrix.

All of the analyses described above were repeated separately for data from two years, 2006 and 2007, and all but the Mantel tests also for a set combining the data from both years. When data from both years were used, year was included as an explanatory variable in the models, and data on three species not encountered in both years were excluded (*Cynips longiventris*, *Neuroterus numismalis*, and *Tischeria dodonea*). In the GLMM, species × year interaction was added as a fixed effect, to account for the variation in abundance between years.

A Mantel test was used to check for spatial autocorrelation in the observed *IR* values. The computation was performed in R using the ade4 package (Chessel et al., 2004) with 10,000 permutations of the matrices. No significant spatial autocorrelation in the *IR* values was detected ($P = 0.24$).

3 Results

3.1 Genetic variation within and among three Finnish oak populations

All 15 microsatellite loci examined showed high levels of polymorphism. In Wattkast all but one locus (Z112), in Inkoo six loci, and in Salo five loci showed significant ($P < 0.05$) deviation from Hardy-Weinberg equilibrium due to an excess of homozygotes (Table 3.1). The average expected heterozygosity over all loci varied between 0.72 in Salo and 0.82 in Inkoo (Table 3.1). The number of different alleles detected ranged from 3 (at locus A15 in Salo) to 29 (at locus Z11 in Inkoo; Table 3.1). The average number of alleles per locus was greatest in Inkoo and smallest in Salo. After rarefaction, however, the Wattkast and Salo populations showed comparable levels of genetic diversity. Genetic diversity was still clearly higher in the Inkoo population than in the other two.

Standard AMOVA (Table 3.2) showed the majority of the variation to occur within populations (88.33%), with significant genetic differentiation between the populations (11.67%, $P < 0.001$).

3.2 Genetic groups within the three populations

The results of the clustering analysis did not clearly indicate the correct number of distinct clusters (K) within the focal populations. When all three populations were analyzed together, the value of $\ln P(X|K)$ seemed to plateau at $K = 5$, but

Table 3.1: Genetic variation (number of alleles (A), expected heterozygosity (H_e), and observed heterozygosity (H_o)) at 15 microsatellite loci in the Wattkast, Inkoo, and Salo populations.[a] The P-value refers to the significance of departure from Hardy-Weinberg Equilibrium based on the exact test (Guo and Thompson, 1992). For the Wattkast and Inkoo populations, $\hat{A}_{35}$ refers to the expected number of alleles after rarefaction.[b]

	Wattkast					Inkoo					Salo			
Locus	A	$\hat{A}_{35}$	H_e	H_o	P-value	A	$\hat{A}_{35}$	H_e	H_o	P-value	A	H_e	H_o	P-value
A110	20	8.53	0.71	0.50	< 0.001	16	13.18	0.66	0.66	0.23	11	0.70	0.79	0.93
A36	15	9.66	0.81	0.64	< 0.001	11	10.30	0.85	0.82	< 0.001	9	0.83	0.97	0.03
Z112	15	8.29	0.70	0.68	0.18	19	16.09	0.91	0.97	0.96	11	0.84	0.91	0.22
Z7	19	11.16	0.85	0.80	< 0.001	20	15.73	0.90	0.90	0.19	10	0.86	0.91	0.86
A9	13	9.39	0.81	0.75	< 0.001	12	9.95	0.87	0.77	0.01	11	0.73	0.74	0.74
Z11	20	13.34	0.88	0.75	< 0.001	29	20.88	0.91	0.88	0.03	15	0.90	0.91	0.16
A15	9	6.08	0.68	0.60	< 0.001	11	8.38	0.52	0.53	0.37	3	0.25	0.09	< 0.001
A1-5	13	7.38	0.72	0.57	< 0.001	12	10.44	0.88	0.95	0.56	11	0.88	0.85	0.03
Q4	10	8.07	0.74	0.31	< 0.001	9	7.76	0.83	0.62	< 0.001	5	0.75	0.72	0.75
Z87	28	13.89	0.86	0.87	< 0.001	23	16.32	0.90	0.89	0.08	13	0.86	1.00	0.04
A16	20	11.64	0.78	0.72	< 0.001	20	15.56	0.91	0.89	0.03	13	0.86	1.00	0.03
Q13	11	5.45	0.27	0.25	0.01	12	9.03	0.70	0.67	0.09	5	0.70	0.58	0.34
Z108	14	8.51	0.70	0.65	< 0.001	18	13.34	0.74	0.60	< 0.001	-	-	-	-
Z101	13	7.60	0.71	0.60	< 0.001	20	15.71	0.88	0.85	0.76	11	0.81	0.78	0.67
Z104	22	11.43	0.83	0.80	0.02	24	18.92	0.90	0.88	0.10	14	0.90	0.90	0.77
Mean	16.13	9.37	0.74	0.63		17.07	13.44	0.82	0.79		10.14	0.72	0.80	
s.d.	5.21	2.50	0.15	0.17		5.81	4.00	0.12	0.14		3.55	0.26	0.24	

[a]For the Salo population, data is missing for locus Z108.

[b]The rarefaction method of Hurlbert (1971) was used to calculate estimates of the expected numbers of alleles in each locus in a random sample of 35 individuals from these populations, corresponding to the sample from Salo ($n = 35$).

Table 3.2: Standard analysis of molecular variance (Excoffier et al., 1992) for the Wattkast, Inkoo and Salo populations.

Source of variation	d.f.	Sum of squares	Variance component	% variation	P-value
Among populations	2	253.35	0.69	11.67	< 0.001
Within populations	647	3360.83	5.19	88.33	
Total	649	3614.18	5.88		

the highest value of ΔK was at $K = 2$ (Figure 3.1). When K was set to 2, there was a clear distinction between Wattkast and the two other populations (Figure 3.2a). At $K = 3$, two clusters were identified within Wattkast, while Inkoo and Salo still appeared undifferentiated, indicating a stronger subpopulation structure within Wattkast alone than between the Inkoo and Salo populations together (Figure 3.2b). Only at $K = 4$ or greater was there a distinction between the Inkoo and Salo populations, but now a clear one (Figure 3.2c). These patterns were mostly consistent between all 10 runs.

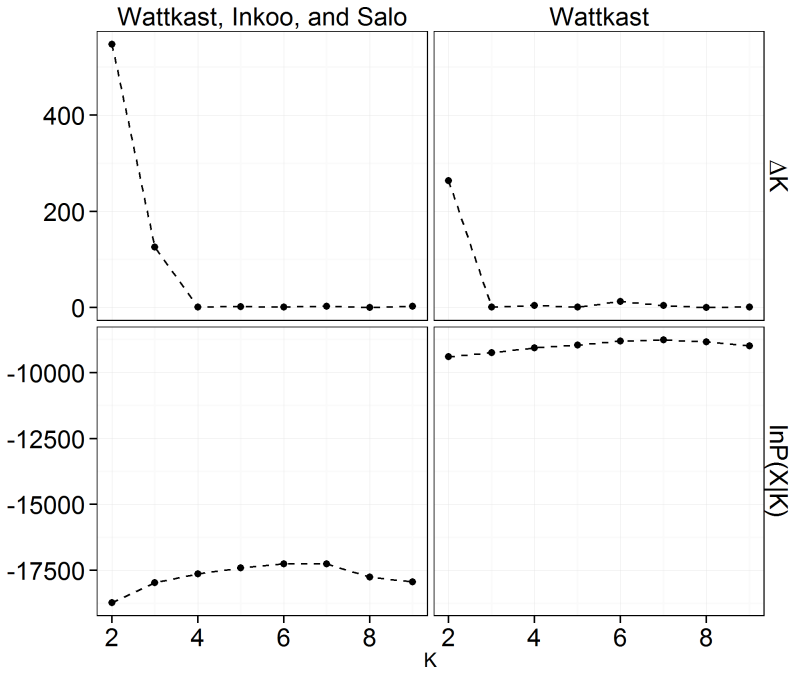

Figure 3.1: Metrics reflecting the number of genetic groups (K) offering the best fit to the data for all populations and for the Wattkast population alone. Shown are average values of $\ln P(X|K)$ and ΔK for different values of K. ΔK is a function of the second order rate of change of $\ln P(X|K)$ and its variance, and is expected to peak at the most likely number of genetic groups (Evanno et al., 2005).

The reliability of the analysis may also be improved by analyzing some of the study populations individually (Pritchard et al., 2010), which was done for the Wattkast population. Here, there was no clear plateau for $\ln P(X|K)$, as the absolute changes in the parameter remained relatively small as K increased (ranging approximately from 50 to 150 as K went from 2 to 9). The highest value of ΔK was found at $K = 2$ (Figure 3.1).

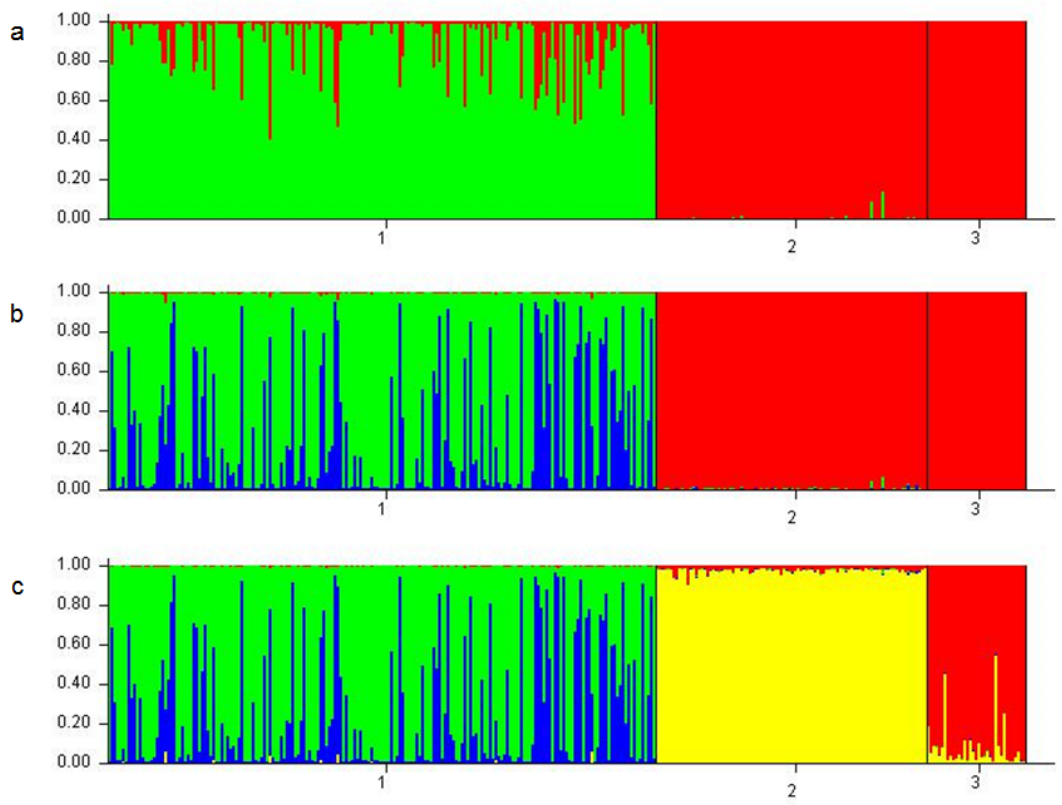

Figure 3.2: Bar plots demonstrating the ancestry of individuals with different numbers of assumed populations (K) as estimated by the software STRUCTURE (Pritchard et al., 2000). (a) $K = 2$; (b) $K = 3$; (c) $K = 4$. Each plot is an example of one run of the simulation. In the plots, each vertical bar represents an individual, the different colors represent the genetic groups identified by the analysis, and the proportion of colors shows the proportion of ancestry derived from these groups. Source populations of the samples are separated by black vertical lines. 1 stands for Wattkast, 2 for Inkoo, and 3 for Salo.

3.3 Genetic structure of the Wattkast population

Within the Wattkast population, the assignment of individuals into two ($K = 2$) genetic groups was further examined. In a logistic regression analysis, location was found to have a significant effect and size a marginally significant effect on the

probability of an individual belonging to one of the groups (Table 3.3). The majority of small trees belonged to the same group, and most of the trees in this group were located on the western side of the island (Figure 3.3).

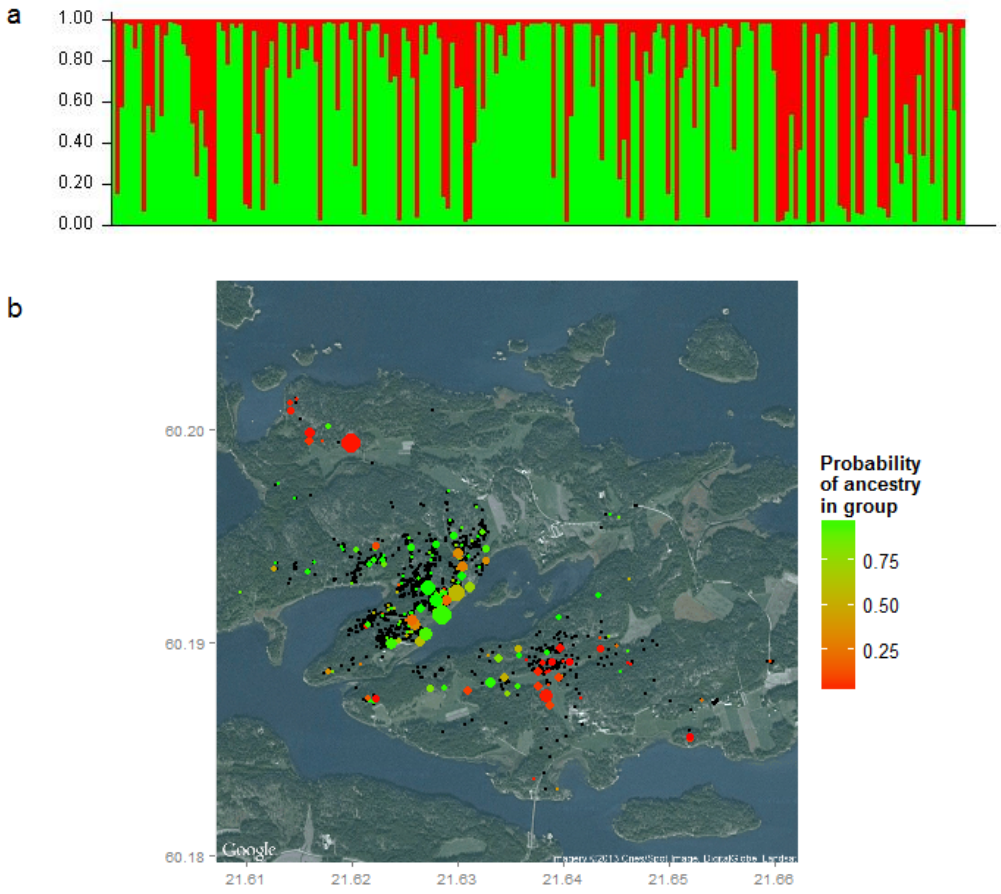

Figure 3.3: Assignment of individuals from Wattkast to two genetic groups identified by the software STRUCTURE (Pritchard et al., 2000). (a) Bar plot demonstrating the estimated ancestry of individuals in the two groups. Each vertical bar represents an individual, the two colors represent groups identified by the analysis, and the proportion of each color shows the probability of ancestry in the corresponding group. (b) A map of individual trees on Wattkast as colored by the genetic groups identified by the analysis. Colors correspond to those in the bar plot above, and the size of the dot to the circumference of the tree. For comparison, black dots show the locations of all oak trees on the island.

The genetic relatedness among pairs of individuals significantly decreased as the distance between them increased (Figure 3.4a). Observed values of pairwise kinship coefficients were significantly greater than expected in a random population ($P[obs>exp] < 0.05$) in the first three distance classes and significantly smaller ($P[obs<exp] < 0.05$) in the last five classes (Figure 3.4b).

Table 3.3: Generalized linear model of factors influencing the probability of an individual belonging to one of two genetic groups identified by the Bayesian clustering analysis implemented in STRUCTURE (Pritchard et al., 2000).

	Estimate	Std. Error	z-value	P-value
Circumference	−0.01	0.003	-1.78	0.07
East	−29.64	6.74	-4.40	< 0.0001
North	−14.19	3.23	-4.40	< 0.0001
East × North	4.4E-06	1.0E-08	4.40	< 0.0001

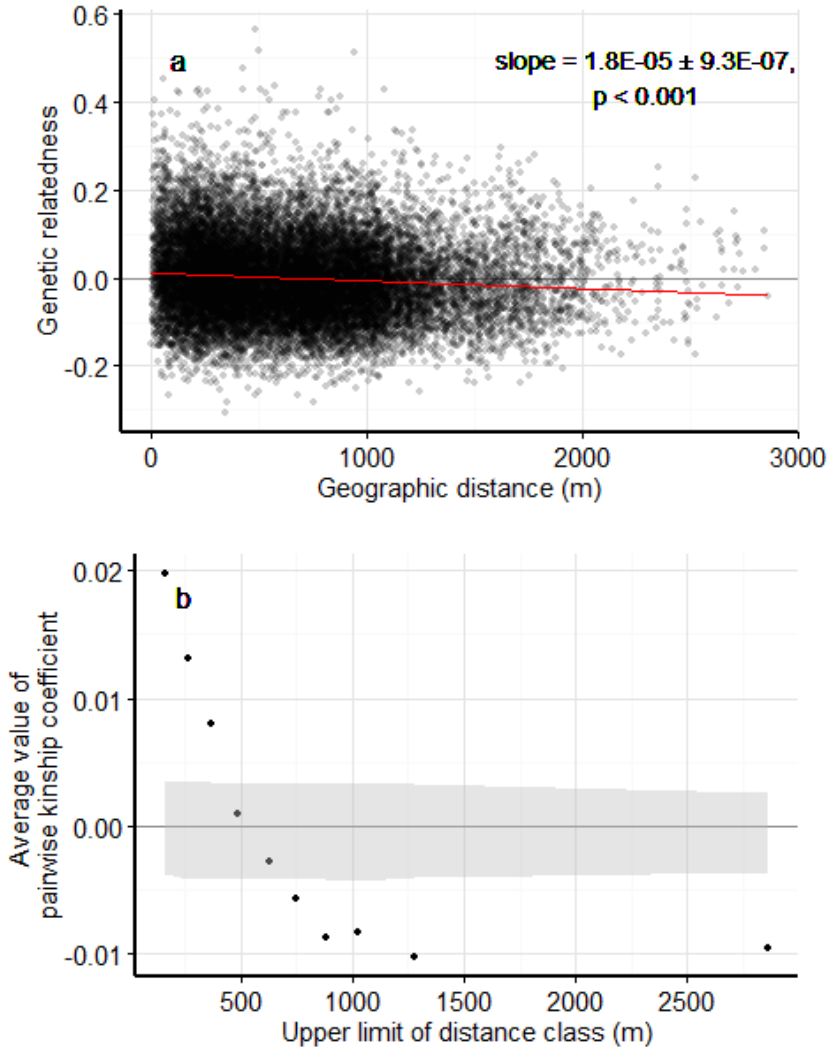

Figure 3.4: The association of genetic relatedness and geographic distance between pairs of individual trees on Wattkast. (a) Pairwise values of genetic relatedness plotted against pairwise geographic distances. The red line represents the line of best fit. For graphical comparison, a horizontal line has been added at $y = 0$. (b) Average values of pairwise kinship coefficients within the 10 distance classes. Grey area shows 95% confidence interval based on randomized distributions within the distance class.

Bigger and therefore older trees were significantly more heterozygous than smaller trees (Figure 3.5a-b). A comparison of the numbers of alleles per locus detected and the observed heterozygosity between the juvenile and adult trees also suggested a loss of genetic variation in the younger generation, although these patterns were not statistically significant (Figure 3.6). However, an analysis of molecular variance confirmed a slight but significant differentiation between the two groups (F_{ST} = 0.004, P = 0.02).

There was no significant effect of spatial connectivity on the measures of individual inbreeding (Figure 3.5c-d).

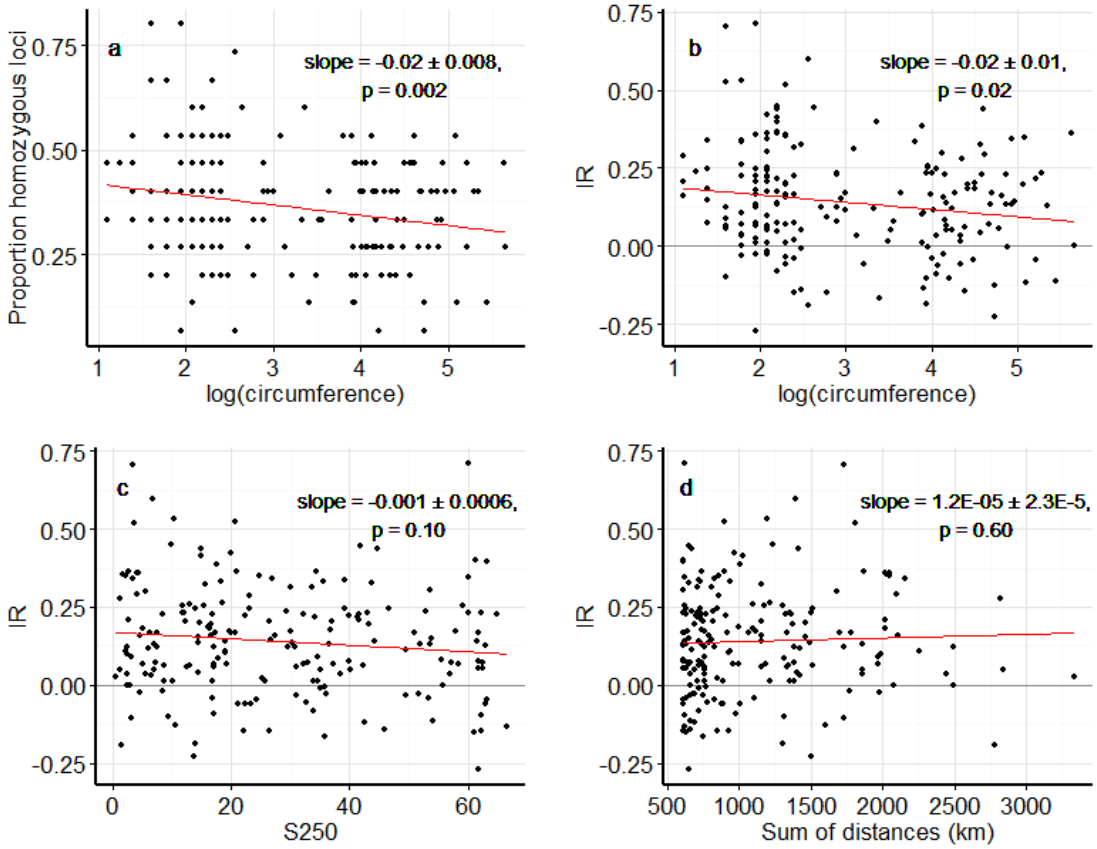

Figure 3.5: Measures of individual inbreeding plotted against (a) and (b) the size (natural logarithm of circumference), and (c) and (d) the connectivity of trees on Wattkast. Red lines represent lines of best fit. Negative *IR* values indicate higher heterozygosity and positive values higher homozygosity.

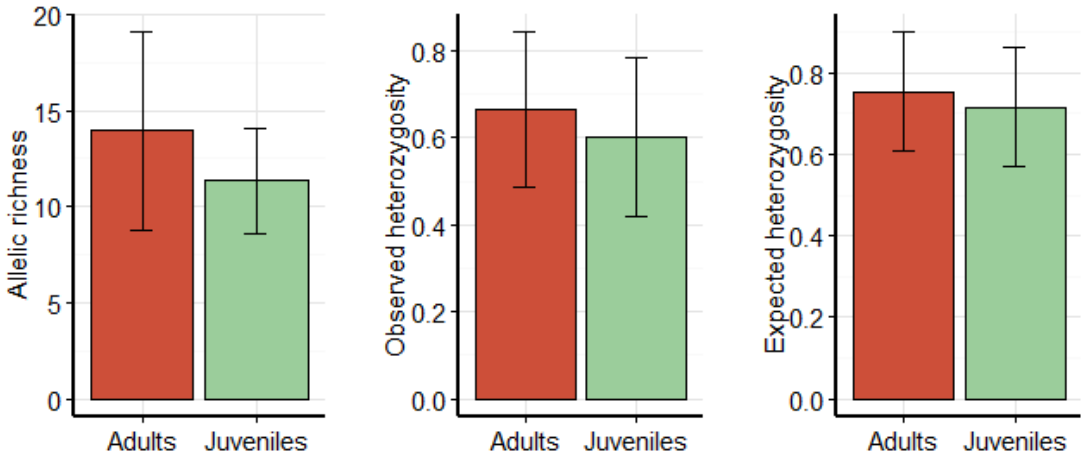

Figure 3.6: Allelic richness, observed heterozygosity, and expected heterozygosity in adult and juvenile trees on Wattkast. Values represent the average across all loci ± s.d.

3.4 Association between oak genotype and herbivore community structure on Wattkast

The heterozygosity of the trees (as measured by *IR*) was found to have a statistically significant effect on all of the descriptors of herbivore community structure except Shannon diversity (Table 3.4). The patterns were consistent between both years and when data from the two years were analyzed together. The effect of *IR* on species richness and herbivore abundance was negative: higher *IR* values were associated with lower richness and abundance (Figure 3.7). As higher *IR* values indicate higher homozygosity, a negative correlation between species richness and *IR* implies that an increase in the heterozygosity of the host increases the number of associated species. Similarly, a negative correlation between abundance and *IR* suggests that herbivores are, on average, more abundant on more heterozygous trees.

The negative correlation between abundance and *IR* was also observed on the species level. Species-specific abundances were mostly negatively correlated with *IR*: in both years, a clear majority of the correlations were negative, and most of the negative correlations were statistically significant. When data from the two years were combined, none of the few positive correlations were significant. Higher *IR* values were found to reduce the abundance of herbivores also in the GLMM of species-specific abundances. The species × *IR* interaction was non-significant ($F = 0.55$, $P = 0.93$), suggesting that most species respond to *IR* in the same way. When this interaction was dropped, the effect of *IR* was statistically significant (Table 3.5).

Table 3.4: Generalized linear models of descriptors of herbivore communities on individual trees as a function of the number of leaves examined during the survey and of the *IR* of the tree. Shown are the results for data combining both years (2006 and 2007), with year included as an explanatory variable. Species richness, total herbivore abundance, and miner and galler abundance were modeled with Poisson regression; Shannon diversity with standard linear regression.

	n leaves		Year		*IR*	
	Z	*P*	*Z*	*P*	*Z*	*P*
Richness	16.23	<2E-16	−3.13	0.002	−2.93	0.003
Abundance	95.75	<2E-16	−18.70	<2E-16	−14.16	<2E-16
Miners	50.07	<2E-16	−2.79	0.005	−5.29	1.19E-07
Gallers	69.66	<2E-16	−32.98	<2E-16	−11.26	<2E-16
	t	*P*	*t*	*P*	*t*	*P*
Shannon	7.09	2.80E-11	−2.34	0.02	−1.65	0.10

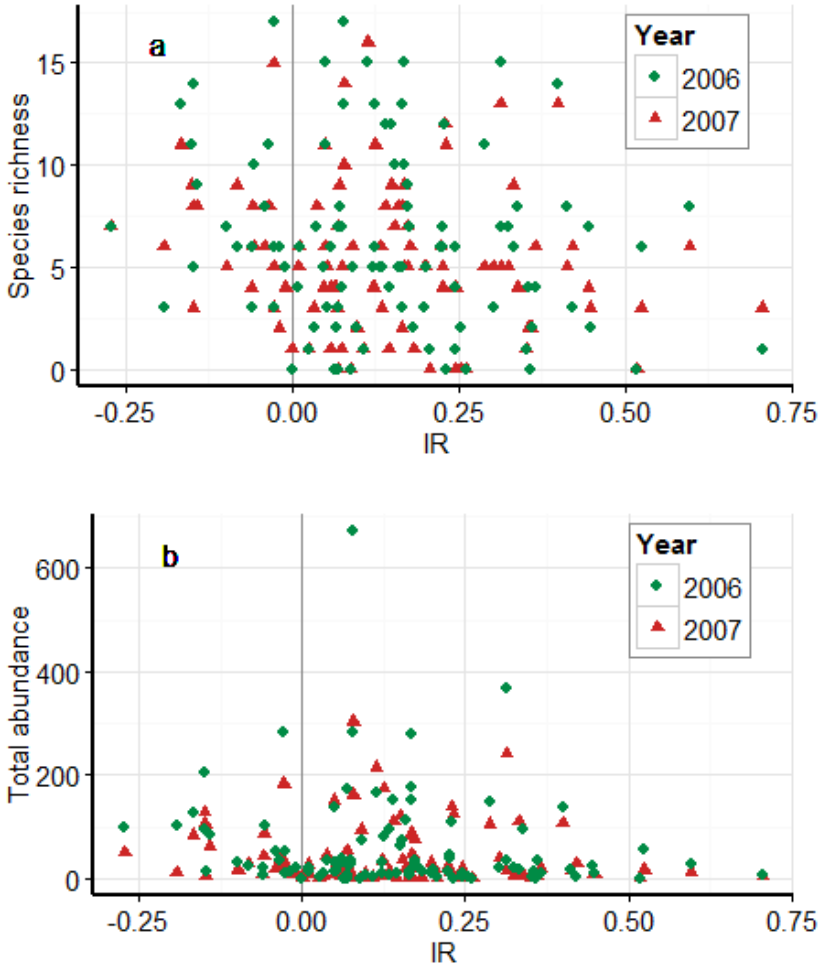

Figure 3.7: Species richness and total herbivore abundance plotted against *IR*.

Table 3.5: GLMM of species-specific abundances as a function of species identity, number of leaves examined in the survey, year, and *IR*. In the model, *IR* had a significant negative effect on the species-specific abundances (slope = -0.86 ± 0.21). The abundances of species varied between years, as indicated by the significant species × year interaction.

Source of variation	d.f.	Den d.f.	F-value	P-value
Species	18	1782	29.53	<0.0001
n leaves	1	3662	801.87	<0.0001
Year	1	99	0.73	0.40
IR	1	98	17.54	<0.0001
Species×year	18	1782	8.27	<0.0001

In both years, there was a statistically significant positive correlation between both the pairwise values of geographic distance and herbivore community dissimilarity, and of genetic distance and herbivore community dissimilarity (Table 3.6). The former, however, was stronger and more significant than the latter. The correlation between the three distance matrices was the weakest and non-significant.

Table 3.6: Mantel tests of correlation between herbivore community dissimilarity, geographic distance, and genetic distance of trees on Wattkast. Community dissimilarity was described by the Bray-Curtis index (Faith et al., 1987), and genetic distance by Rousset's $\hat{a}$ (Rousset, 2000). The partial Mantel test between the three matrices (bottom row) compared the first two matrices (community dissimilarity and genetic distance) while controlling for the third (geographic distance). r = the Mantel statistic computed, and P = its significance based on 1,000 permutations.

	2006		2007	
Distance matrices tested	r	P	r	P
Community × Geographic	0.36	0.001	0.39	0.001
Community × Genetic	0.10	0.02	0.15	0.004
Community × Genetic × Geographic	0.02	0.30	0.06	0.10

4 Discussion

Among the three *Quercus robur* populations studied here, 15 microsatellite loci showed high levels of diversity within the populations and significant differentiation among them. Within one population, on the island of Wattkast, patterns of spatial and temporal subpopulation differentiation were observed. The genetic variation and identity of the trees within the island were found to have a significant effect

on insect herbivore community structure: higher heterozygosity was associated with higher richness and abundance of species, and genetic relatedness between pairs of trees with more similar community compositions.

4.1 Genetic diversity and differentiation in marginal *Q. robur* populations

Genetic diversity is hypothesized to be typically lower and differentiation higher in marginal than in central populations (Eckert et al., 2008). *Q. robur* grows in northern Europe at the northern limit of its natural range, and cpDNA variation has been found to be lower and differentiation higher in *Q. robur* populations in northern Europe than in central Europe (Petit et al., 2002a; Vakkari et al., 2006). Maternally inherited cpDNA may show somewhat different levels of variation and differentiation than nuclear microsatellites (Petit et al., 2005). However, the current results are in agreement with the earlier findings.

The levels of nuclear microsatellite diversity detected in this study showed a similar trend as those reported for cpDNA: numbers of alleles detected were, on average, lower than those recorded in more central populations in Europe (Table 4.1). Comparing estimated allelic richness between populations is problematic when sample sizes are not equal (Leberg, 2002), but, based on the available data, genetic diversity appears to be, on average, lower in these marginal populations than elsewhere in Europe.

The majority of genetic variation (88.33%) occurred within populations, which is a common pattern among *Quercus* species (Kremer and Petit, 1993; Lind and Gailing, 2013). However, the level of population differentiation was statistically significant, and the proportion of among-population genetic variation (11.67%, or $F_{ST} = 0.12$) seems relatively high. For comparison, using 14 microsatellite loci, Neophytou et al. (2010) observed a noticeably lower F_{ST} of 0.039 among three *Q. robur* populations located in Germany, Greece and Bulgaria, respectively, i.e. across a much larger geographic range than the one covered by this study. F_{ST}-values lower or only slightly higher than the one observed here have also been recorded at the level of different *Quercus* species: e.g. F_{ST} of 0.096, 0.113, and 0.168 between *Q. robur* and *Q. petraea*, *Q. pubescens*, and *Q. frainetto*, respectively, in Romania as based on six microsatellite loci (Curtu et al., 2007). These results give support to the earlier findings of higher differentiation among oak populations in northern than central Europe; however, this differentiation may be in some part explained by fragmentation as opposed to geographic marginality alone (see also Petit et al., 2002a).

Table 4.1: Numbers of alleles detected in the three oak populations studied here (top part of table) as compared with results for the same loci detected in *Q. robur* populations elsewhere in Europe (lower part of table). Note that sample sizes (n) vary.

		Locus				
Location	n	A36	A9	A1-5	Z104	Q13
Finland - Wattkast	194	15	13	13	22	11
Finland - Inkoo	96	11	12	12	24	12
Finland - Salo	35	9	11	11	14	5
Mean		11.67	12	12	20	9.33
Romania (Curtu et al., 2007)	65	16	13	14	30	8
Germany (Degen et al., 1999)	228	17	14	15	33	-
	85	13	11	12	28	-
UK (Cottrell et al., 2003)	388	20	12	20	31	12
	58	28	16	20	25	15
France (Streiff et al., 1998)	183	19	14	17	29	12
Turkey (Yücedağ and Gailing, 2013)	50	15	-	12	17	-
	50	15	-	12	17	-
Mean		17.88	13.33	15.25	26.25	11.75

Excess of homozygotes was more pronounced in the Wattkast population than in Inkoo and Salo; however, genetic diversity was not lower in Wattkast than in the other two. Since there was stronger subpopulation structure within Wattkast alone than between Inkoo and Salo together, the more pronounced heterozygote deficiency within Wattkast could be due to a Wahlund effect: the presence of subpopulations with different allele frequencies reduces the overall heterozygosity of the population (Allendorf and Luikart, 2007). The subpopulation structure of the Wattkast oak population is discussed in more detail below.

4.2 Spatial genetic structure of the Wattkast population

Within Wattkast, some subpopulation differentiation was observed among young and old trees and among trees located in different parts of the island. As genetic relatedness among pairs of individuals significantly decreased as the distance between them increased, the population is not completely admixed and shows signs of isolation-by-distance. In the Netherlands, a significant negative correlation between genetic relatedness and geographic distance was likewise found among individuals in autochthonous stands of *Q. robur* and *Q. petraea*, but not in planted stands (Buiteveld and Koelewijn, 2006). This association between relatedness and geographic distance was also revealed by the distribution of the two genetic groups identified by the Bayesian clustering analysis. The probability of an individual hav-

ing an ancestry in one of the two groups identified was significantly affected by the individual's location.

Spatial subpopulation structure may be caused by isolation-by-distance and/or by inbreeding (Allendorf and Luikart, 2007). However, such an interpretation would assume that the population is in genetic equilibrium. Several issues give reason to question this state within the Wattkast population:

South-western Finland was colonized by oaks relatively recently, about 8000 BP, from both the east and the west (Ferris et al., 1998). As Ferris et al. (1998) describe, during the first centuries of colonization, the distribution of oaks in Finland was significantly different from the current state, and the entire present range of oaks was still submerged in the sea. For example, as recently as in the 16th century, the shoreline of Wattkast was almost three meters higher than today (Zilliacus, 2001). Therefore, the current oak populations in Finland can be considered relatively young. Cottrell et al. (2003) suggest incomplete genetic homogenization due to the recency of the post-glacial colonization may be a factor in explaining high homozygote excess in oak populations in northern Britain, and that mixing of nuclear genomes from different oak refugia after the last glacial period may be a common pattern in northern Europe. As the generation length of oaks is long and the generations are overlapping in time, reaching genetic equilibrium may take a long time.

The landscape of Wattkast has also experienced intensifying human disturbance since the 15th century up to recent decades, when human land use on the island has become less extensive (Zilliacus, 2001). In the past, the island has been inhabited by more than 80 people with agriculture as the main source of livelihood. When the population was growing, arable land on the island was scarce and even small and scattered patches of land suitable for farming were cleared. The remaining woods and reeds were grazed by livestock, and small-scale forestry was also practiced. Today, the population of Wattkast is less than 30 people (Zilliacus, 2001).

Within the island, oaks are absent from several suitable habitats, and the occurrence of the oaks seems to be limited by the trees' dispersal ability (Kettunen, 2010). Due to its physiological and biological properties, *Q. robur* possesses a high colonizing ability (Ducousso et al., 1993). Distance to uncolonized habitats alone seems like an inadequate explanation to the observed distribution. Other potential explaining factors include dispersal barriers and movements of seed-dispersing animals (see also the discussion in Kettunen, 2010). Some of these factors may be still adjusting to the changing land use on the island.

Unbalanced dispersal of the trees within the island may also explain the patterns of differentiation observed among juvenile and adult trees. The degree of differentiation (F_{ST}) between juvenile and adult trees was 0.004 and statistically significant at a 5%

confidence level, and the majority of small trees belonged to the same genetic group (green in Figure 3.3). The larger trees in this group are centered in the large cluster of trees on the north-western side of the Wattkast bay. It is thus possible, for example, that the activity of seed-dispersing animals is also centered on this part of the island. An alternative explanation is that selection favors certain genotypes, causing them to be proportionally more common among adult than juvenile trees.

In conclusion, the results of the analyses combined with the characteristics of the site, including the ongoing shift to less extensive land use, give reason to suspect the population is in genetic disequilibrium.

4.3 *Q. robur* and its herbivores in the community genetics framework

The diversity of herbivore communities has been found to increase with increasing individual and population level genetic diversity of host plants in several studies (Hersch-Green et al., 2011, and citations therein). Such patterns have also been reported in the case of oaks (Tovar-Sánchez and Oyama, 2006; Tovar-Sánchez et al., 2013). In the present study, the heterozygosity of the host tree showed a significant positive effect on the richness and abundance of insect herbivore species. Similarly, Tovar-Sánchez et al. (2013) found more heterozygous *Q. castanea* and *Q. crassipes* hosts in Mexico to harbor, on average, more diverse and dense insect communities. These authors found total arthropod biomass, then again, to decrease with increasing heterozygosity of the host, which the authors interpreted as a potential sign of weaker defense in more homozygous hosts. Measures of biomass and abundance are not directly comparable, but, in the present study, increasing heterozygosity of the host was found to increase both the total herbivore abundance and the abundances of individual species.

A statistically significant association was observed between herbivore community dissimilarity and genetic distance between pairs of trees. The correlation of host genetic relatedness and associated community similarity, coined the 'genetic similarity rule', has been observed, for instance, in cottonwood-based systems (Bangert et al., 2006a,b), and in tropical epiphytic bromeliads even when the spatial autocorrelation of relatedness was accounted for (Zytynska et al., 2012). The pairwise genetic distances of trees correlated with their geographic distances also in the present study, but, unlike in the study of Zytynska et al. (2012), the association of community similarity and genetic relatedness became non-significant when this was accounted for. Additionally, the correlation between herbivore community dissimilarity and geographic distance was stronger and more significant than that between community dissimilarity and genetic distance. Thus, it is possible that the apparent effect of the

genotype observed here is, in fact, due to spatial autocorrelation. These results may be interpreted as the effect of the spatial context overriding the effect of the genotypic identity. This interpretation also corresponds to those from previous studies on oak-based food webs in Finland, which have suggested that the spatial structure of the habitats, i.e. distribution of the host tree, plays a major role in structuring the associated insect communities, whereas host genotype has been considered to have a secondary effect (Gripenberg and Roslin, 2008; Tack et al., 2010).

Quantification of the importance of genetic variation relative to environmental factors has been proposed as one of the next, necessary steps in the development of the community genetics theory (Hersch-Green et al., 2011; Tack et al., 2012). Here, when the association of the heterozygosity of the host and the associated species richness and abundance was examined, the spatial context was not included in the models. The distribution of the host trees is only one of the aspects of environmental variation that may affect the presence and abundance of insect herbivore species. Other potentially important aspects include abiotic conditions, inter- and intraspecific interactions, and level of disturbance (Hersch-Green et al., 2011).

The variation in environmental factors should be taken into account in order to conclusively determine the ecological significance and conservation value of the patterns observed here. However, the current results show that, within this marginal, fragmented *Q. robur* population encompassing an area of several km^2, the heterozygosity of the host has a detectable effect on the structure and composition of associated insect herbivore communities. These findings give support to the notion that the conservational relevance of intraspecific genetic diversity may extend beyond the preservation of populations also in the case of *Q. robur* in Finland.

4.4 Conclusions

The results of this study suggest that, despite the geographically marginal location, the absence of other, interfertile oak species, and the significant historic fragmentation of habitats, Finnish *Q. robur* populations may host relatively high levels of genetic diversity, and that variation in heterozygosity among individuals in a population translates into detectable effects on the species diversity of associated insect herbivores. The high levels of genetic diversity observed may signal the viability of these populations; however, the significant and comparably high differentiation among the populations indicates that fragmentation has potentially had an effect on the incoming gene flow. The genetic structure of the Wattkast population illustrates that tree populations may adjust to environmental change with an often considerable time lag. Despite the high genetic diversity within populations and the

strong gene flow typical to *Quercus* species, *Q. robur* populations are not immune to habitat reduction and fragmentation. The current Finnish populations may still be adjusting to the drastic changes of the past centuries. Preservation of genetic variation within the remaining stands is an important factor in the conservation of *Q. robur* in Finland and, based on the results obtained here, potentially a factor also in the preservation of associated species diversity.

Acknowledgements

Foremost, I would like to thank the supervisor of this work, Tomas Roslin, for his invaluable guidance and support. I would also like to thank Sakina Elshibli and Pekka Vakkari for advice regarding the genetic data, and Bess Hardwick for assistance with the herbivore community data. I owe great thanks to Ayco Tack, Helena Korpelainen, and Pertti Pulkkinen for their valuable input and assistance in the method selection and incisive comments on the text. I appreciate all of the free and open-source software used in the work. Finally, I would like to thank the entire Spatial Food Web Ecology group for offering an inspiring and welcoming work environment.

Bibliography

Allendorf, F. and Luikart, G. (2007). *Conservation and the Genetics of Populations*. Blackwell Publishing. 642p.

Aparicio, J. M., Ortego, J., and Cordero, P. J. (2006). What should we weigh to estimate heterozygosity, alleles or loci? *Molecular Ecology*, 15(14):4659–4665.

Bangert, R. K., Allan, G. J., Turek, R. J., Wimp, G. M., Meneses, N., Martinsen, G. D., Keim, P., and Whitham, T. G. (2006a). From genes to geography: a genetic similarity rule for arthropod community structure at multiple geographic scales. *Molecular Ecology*, 15(13):4215–4228.

Bangert, R. K., Turek, R. J., Rehill, B., Wimp, G. M., Schweitzer, J. A., Allan, G. J., Bailey, J. K., Martinsen, G. D., Keim, P., Lindroth, R. L., and Whtiham, T. G. (2006b). A genetic similarity rule determines arthropod community structure. *Molecular Ecology*, 15(5):1379–1391.

Buiteveld, J. and Koelewijn, H. (2006). CpDNA haplotype variation reveals strong human influence on oak stands of the Veluwe forest in The Netherlands. *Forest Ecology and Management*, 228(1–3):160 – 167.

Chessel, D., Dufour, A., and Thioulouse, J. (2004). The ade4 package-I - One-table methods. *R News*, 4:5–10.

Cottrell, J., Munro, R., Tabbener, H., Milner, A., Forrest, G., and Lowe, A. (2003). Comparison of fine-scale genetic structure using nuclear microsatellites within two British oakwoods differing in population history. *Forest Ecology and Management*, 176(1–3):287–303.

Curtu, A. L., Gailing, O., Leinemann, L., and Finkeldey, R. (2007). Genetic variation and differentiation within a natural community of five oak species (*Quercus* spp.). *Plant Biology*, 9(1):116–126.

Degen, B., Streiff, R., and Ziegenhagen, B. (1999). Comparative study of genetic variation and differentiation of two pedunculate oak (*Quercus robur*) stands using microsatellite and allozyme loci. *Heredity*, 83(5):597–603.

Dow, B., Ashley, M., and Howe, H. (1995). Characterization of highly variable $(GA/CT)_n$ microsatellites in the bur oak, *Quercus macrocarpa*. *Theoretical and Applied Genetics*, 91(1):137–141.

Ducousso, A., Michaud, H., and Lumaret, R. (1993). Reproduction and gene flow in the genus *Quercus* L. *Ann. For. Sci.*, 50:91s–106s.

Eckert, C. G., Samis, K. E., and Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, 17(5):1170–1188.

Ellison, A. M., Bank, M. S., Clinton, B. D., Colburn, E. A., Elliott, K., Ford, C. R., Foster, D. R., Kloeppel, B. D., Knoepp, J. D., Lovett, G. M., Mohan, J., Orwig, D. A., Rodenhouse, N. L., Sobczak, W. V., Stinson, K. A., Stone, J. K., Swan, C. M., Thompson, J., Holle, B. V., and Webster, J. R. (2005). Loss of foundation species: Consequences for the structure and dynamics of forested ecosystems. *Frontiers in Ecology and the Environment*, 3(9):pp. 479–486.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14(8):2611–2620.

Excoffier, L. and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3):564–567.

Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131(2):479–91.

Faith, D. P., Minchin, P. R., and Belbin, L. (1987). Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio*, 69(1-3):57–68.

Falush, D., Stephens, M., and Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, 7(4):574–578.

Fernández-M., J. F. and Sork, V. L. (2007). Genetic variation in fragmented forest stands of the andean oak *Quercus humboldtii* Bonpl. (Fagaceae). *Biotropica*, 39(1):72–78.

Ferris, C., King, R. A., Väinölä, R., and Hewitt, G. M. (1998). Chloroplast DNA recognizes three refugial sources of European oaks and suggests independent eastern and western immigrations to Finland. *Heredity*, 80(5):584–593.

Gripenberg, S. and Roslin, T. (2005). Host plants as islands: Resource quality and spatial setting as determinants of insect distribution. *Annales Zoologici Fennici*, 42(4):335–345.

Gripenberg, S. and Roslin, T. (2008). Neither the devil nor the deep blue sea: larval mortality factors fail to explain the abundance and distribution of *Tischeria ekebladella*. *Ecological Entomology*, 33(3):346–356.

Guo, S. W. and Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, 48(2):361–372.

Hamrick, J. (2004). Response of forest trees to global environmental changes. *Forest Ecology and Management*, 197(1–3):323–335.

Hamrick, J., Godt, M., and Sherman-Broyles, S. (1992). Factors influencing levels of genetic diversity in woody plant species. *New Forests*, 6(1-4):95–124.

Hardy, O. J. and Vekemans, X. (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, 2:618–620.

Hersch-Green, E. I., Turley, N. E., and Johnson, M. T. J. (2011). Community genetics: what have we accomplished and where should we be going? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1569):1453–1460.

Honnay, O. and Jacquemyn, H. (2007). Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conservation Biology*, 21(3):823–831.

Hubisz, M., Falush, D., Stephens, M., and Pritchard, J. (2009). Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour*, 9(5):1322–32.

Hurlbert, S. H. (1971). The nonconcept of species diversity: A critique and alternative parameters. *Ecology*, 52(4):pp. 577–586.

Kalinowski, S. (2004). Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. *Conservation Genetics*, 5(4):539–543.

Kampfer, S., Lexer, C., Glössl, J., and Steinkellner, H. (1998). Characterization of $(GA)_n$ microsatellite loci from *Quercus Robur*. *Hereditas*, 129(2):183–186.

Kettunen, H. (2010). Tammen alueellisen levinneisyyden rajoitteet: Tapaustutkimus Länsi-Turunmaalla. Master's thesis, University of Helsinki, Finland.

Kremer, A. and Petit, R. (1993). Gene diversity in natural populations of oak species. *Annals of Forest Science*, 50:186–202.

Leberg, P. L. (2002). Estimating allelic richness: Effects of sample size and bottlenecks. *Molecular Ecology*, 11(11):2445–2449.

Lesica, P. and Allendorf, F. W. (1995). When are peripheral populations valuable for conservation? *Conservation Biology*, 9(4):753–760.

Lind, J. F. and Gailing, O. (2013). Genetic structure of *Quercus rubra* L. and *Quercus ellipsoidalis* E. J. Hill populations at gene-based EST-SSR and nuclear SSR markers. *Tree Genetics & Genomes*, 9(3):707–722.

Lindbladh, M. and Foster, D. R. (2010). Dynamics of long-lived foundation species: the history of *Quercus* in southern Scandinavia. *Journal of Ecology*, 98(6):1330–1345.

Loiselle, B. A., Sork, V. L., Nason, J., and Graham, C. (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, 82(11):pp. 1420–1425.

Neophytou, C., Aravanopoulos, F., Fink, S., and Dounavi, A. (2010). Detecting interspecific and geographic differentiation patterns in two interfertile oak species (*Quercus petraea* (Matt.) Liebl. and *Q. robur* L.) using small sets of microsatellite markers. *Forest Ecology and Management*, 259(10):2026–2035.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., and Wagner, H. (2013). *vegan: Community Ecology Package*. R package version 2.0-9.

Ollinmaa, P. J. (1952). Jalot lehtipuumme luontaisina ja viljeltyinä. *Silva Fennica*, 77:1–73.

Olrik, D. C., Hauser, T. P., and Kjaer, E. D. (2012). Natural colonisation of an open area by *Quercus robur* l. – from where did the vectors disperse the seed? *Scandinavian Journal of Forest Research*, 27(4):350–360.

Petit, R. J., Csaikl, U. M., Bordács, S., Burg, K., Coart, E., Cottrell, J., van Dam, B., Deans, J. D., Dumolin-Lapègue, S., Fineschi, S., Finkeldey, R., Gillies, A., Glaz, I., Goicoechea, P. G., Jensen, J. S., König, A. O., Lowe, A. J., Madsen, S. F., Mátyás, G., Munro, R. C., Olalde, M., Pemonge, M.-H., Popescu, F., Slade, D., Tabbener, H., Taurchini, D., de Vries, S. G., Ziegenhagen, B., and Kremer, A. (2002a). Chloroplast DNA variation in European white oaks: Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management*, 156(1–3):5–26.

Petit, R. J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., and Vendramin, G. G. (2005). Invited review: Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, 14(3):689–701.

Petit, R. J., Latouche-Hallé, C., Pemonge, M.-H., and Kremer, A. (2002b). Chloroplast DNA variation of oaks in France and the influence of forest fragmentation on genetic diversity. *Forest Ecology and Management*, 156(1–3):115–129.

Pimm, S. L. and Raven, P. (2000). Biodiversity: Extinction by numbers. *Nature*, (6772):843–845.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–959.

Pritchard, J. K., Wen, X., and Falush, D. (2010). Documentation for STRUCTURE software: Version 2.3.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rassi, P., Hyvärinen, E., Juslén, A., and Mannerkoski, I., editors (2010). *Suomen lajien uhanalaisuus – Punainen kirja 2010*. Ympäristöministeriö & Suomen ympäristökeskus. 685p.

Reed, D. H. and Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, 17(1):230–237.

Rousset, F. (2000). Genetic differentiation between individuals. *Journal of Evolutionary Biology*, 13(1):58–62.

Rousset, F. (2008). GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, 8(1):103–106.

Sork, V. and Smouse, P. (2006). Genetic analysis of landscape connectivity in tree populations. *Landscape Ecology*, 21(6):821–836.

Southwood, T. R. E. (1961). The number of species of insect associated with various trees. *Journal of Animal Ecology*, 30(1):1–8.

Steinkellner, H., Fluch, S., Turetschek, E., Lexer, C., Streiff, R., Kremer, A., Burg, K., and Glössl, J. (1997). Identification and characterization of $(GA/CT)_n$ - microsatellite loci from *Quercus petraea*. *Plant Molecular Biology*, 33(6):1093–1096.

Streiff, R., Labbe, T., Bacilieri, R., Steinkellner, H., Glössl, J., and Kremer, A. (1998). Within-population genetic structure in *Quercus robur* l. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Molecular Ecology*, 7(3):317–328.

Tack, A. J., Ovaskainen, O., Pulkkinen, P., and Roslin, T. (2010). Spatial location dominates over host plant genotype in structuring an herbivore community. *Ecology*, 91(9):2660–2672.

Tack, A. J. M., Johnson, M. T. J., and Roslin, T. (2012). Sizing up community genetics: it's a matter of scale. *Oikos*, 121(4):481–488.

Tovar-Sánchez, E. and Oyama, K. (2006). Effect of hybridization of the *Quercus crassifolia*×*Quercus crassipes* complex on the community structure of endophagous insects. *Oecologia*, 147(4):702–713.

Tovar-Sánchez, E., Valencia-Cuevas, L., Castillo-Mendoza, E., Mussali-Galante, P., Pérez-Ruiz, R., and Mendoza, A. (2013). Association between individual genetic diversity of two oak host species and canopy arthropod community structure. *European Journal of Forest Research*, 132(1):165–179.

Vakkari, P., Blom, A., Rusanen, M., Raisio, J., and Toivonen, H. (2006). Genetic variability of fragmented stands of pedunculate oak (*Quercus robur*) in Finland. *Genetica*, 127(1-3):231–241.

Wimp, G. M., Young, W. P., Woolbright, S. A., Martinsen, G. D., Keim, P., and Whitham, T. G. (2004). Conserving plant genetic diversity for dependent animal communities. *Ecology Letters*, 7(9):776–780.

Yücedağ, C. and Gailing, O. (2013). Morphological and genetic variation within and among four *Quercus petrea* and *Q. robur* natural populations. *Turkish Journal of Botany*, 37:619–629.

Zilliacus, L. (2001). Visst om Wattkast. Unpublished manuscript.

Zytynska, S. E., Khudr, M. S., Harris, E., and Preziosi, R. F. (2012). Genetic effects of tank-forming bromeliads on the associated invertebrate community in a tropical forest ecosystem. *Oecologia*, 170(2):467–475.