...ular tools for ...f oak forests

...udies on indigenous *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. populations

Erica G. Bakker

ALTERRA SCIENTIFIC CONTRIBUTIONS 5

ALTERRA

Towards molecular tools for management of oak forests

Genetic studies on indigenous *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. populations

Alterra is the main Dutch centre of expertise on rural areas and water management. It was founded 1 January 2000. Alterra combines a huge range of expertise on rural areas and their sustainable use, including aspects such as water, wildlife, forests, the environment, soils, landscape, climate and recreation, as well as various other aspects relevant to the development and management of the environment we live in. Alterra engages in strategic and applied research to support design processes, policymaking and management at the local, national and international level. This includes not only innovative, interdisciplinary research on complex problems relating to rural areas, but also the production of readily applicable knowledge and expertise enabling rapid and adequate solutions to practical problems.

The many themes of Alterra's research effort include relations between cities and their surrounding countryside, multiple use of rural areas, economy and ecology, integrated water management, sustainable agricultural systems, planning for the future, expert systems and modelling, biodiversity, landscape planning and landscape perception, integrated forest management, geo-information and remote sensing, spatial planning of leisure activities, habitat creation in marine and estuarine waters, green belt development and ecological webs, and pollution risk assessment.

Alterra is part of Wageningen University and Research Centre (Wageningen UR) and includes two research sites, one in Wageningen and one on the island of Texel.

Towards molecular tools for management of oak forests

Genetic studies on indigenous *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. populations

Erica G. Bakker

ALTERRA SCIENTIFIC CONTRIBUTIONS 5

ALTERRA GREEN WORLD RESEARCH, WAGENINGEN

2001

This research presented in this volume was jointly conducted at the Laboratory of Plant Breeding, Wageningen University and at Alterra Green World Research in Wageningen, The Netherlands.

This volume was also published as a doctoral thesis of the University of Wageningen

ISBN 90 327 0310 2

Contents

Abstract

This thesis describes the genetic composition of indigenous oak populations in the Netherlands using two different molecular marker techniques, AFLP and microsatellites. No diagnostic AFLP marker was found that could differentiate between *Q. robur* and *Q. petraea*, but five species-indicative AFLP markers were observed. This indicates that the two oak species are closely related. AFLP and microsatellite analysis of an ancient woodland described large *Q. robur* and *Q. petraea* clones with diameters up to 5.8 m. Such large clones are an indication for the old age of such woodlands that have a history of coppicing and grazing. As indigenous oak populations in the Netherlands suffered from large scale deforestation and have been coppiced and grazed for centuries, it was expected that the genetic variation of such populations would be low. However, the genetic variation observed for two indigenous *Q. robur* populations was as high as for unmanaged French and German *Q. robur* populations. Molecular markers that are closely linked to quantitative trait loci (QTLs) controlling eco-physiological and morphological traits can be used in ecology and forestry. Therefore, a genetic linkage map of *Q. robur* was constructed and QTLs controlling eco-physiological and morphological traits were located.

1 General introduction

The genetic composition of indigenous *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. populations

Indigenous or autochthonous - both words are used as synonyms - populations of *Quercus robur* L. (pedunculate oak) and *Q. petraea* (Matt.) Liebl. (sessile oak) are populations that have occurred in an area since their establishment after the last ice age (14,000 years B.P.). About 8,000 years ago *Q. robur* and *Q. petraea* established themselves in the Netherlands. Such populations are assumed to be valuable as they are expected to be adapted to the local environmental and climatic conditions. Moreover, the disappearance of indigenous populations will always be a genetic loss (Maes 1993). For that reason, interest grew in conservation and propagation of indigenous *Q. robur* and *Q. petraea* trees. However, knowledge about indigenous *Q. robur* and *Q. petraea* populations has so far been poor, especially for populations in the Netherlands. So far, indigenous populations in the Netherlands have been recognized based on a few guidelines determined by Maes (1993): 1. Cultivars can never be indigenous. 2. Trees must be old or there must be old coppices present. 3. The trees or shrubs do not seem to be planted (no clear rows). 4. The area lies within the natural range of dispersion of the species. 5. The area was covered with forest on maps of 1830-1850. 6. The area shows ecological characteristics of the natural area of prevalence of the species. 7. There are indicator species indicative for old forest (undisturbed soils) growing in the area. 8. In the surroundings the species grows on several similar areas. Besides, for the Netherlands it is known that *Q. petraea* has almost never been planted during the course of history. Therefore Dutch *Q. petraea* populations can be stated with more certainty to be autochthonous than *Q. robur* populations. So far, based on Maes' (1993) indications, several *Q. robur* and *Q. petraea* populations in the Netherlands have been pointed out as indigenous populations. However, as Maes' (1993) method for the identification of indigenous populations only gives indications, it is necessary to analyze these presumed autochthonous populations with other (objective) tools in order to analyze them for their autochthonicity, history, age, structure, and quality. This implies that ecological studies only will not be sufficient. There is a need to study the genetic composition of these populations. The best way to study this genetic composition is by means of molecular DNA markers.

Studies of indigenous populations of *Q. robur* and *Q. petraea* in several European countries using molecular markers have revealed information about postglacial recolonization routes (Petit et al. 1993; Dumolin-Lapègue et al. 1997; Petit et al. 1997), genetic variation and structure (Kremer et al. 1991; Petit et al. 1993; Moreau et al. 1994; Zanetto et al. 1994; Samuel et al. 1995; Zanetto & Kremer 1995; Bodénès et al. 1997a,b; Streiff et al. 1998; Degen et al. 1999), and differentiation

between *Q. robur* and *Q. petraea* (Kremer et al. 1991; Petit et al. 1993; Zanetto et al. 1994; Moreau et al. 1994; Bacilieri et al. 1995; Kleinschmit et al. 1995; Samuel et al. 1995; Barreneche et al. 1996; Bodénès et al. 1997a,b; Streiff et al. 1998; Muir et al. 2000).

However, for indigenous populations of *Q. robur* and *Q. petraea* in the Netherlands no such studies have been done so far. As the situation of indigenous populations in the Netherlands might be different as compared to other European countries it is not possible to extrapolate the results for those studies to the Dutch populations. Moreover, molecular markers need to confirm that the assumed autochthonous populations are truly indigenous. First indications about autochthonicity of Dutch populations were obtained after chloroplast DNA (cpDNA) analysis of 13 populations that were selected based on the guidelines of Maes (1993) by Van Dam & De Vries (1998) and Konig et al. (*in press*). This thesis will continue from this point onwards and aims to describe the genetic composition of indigenous *Q. robur* and *Q. petraea* populations in the Netherlands.

Taxonomy

Quercus robur L. (pedunculate oak) and *Q. petraea* (Matt.) Liebl. (sessile oak) (Fagaceae) belong to a species-rich genus consisting of between 300 and 600 (depending of the authors) oak species spread over the northern hemisphere (Europe, Africa, North America, and Asia) from boreal zones to semi-arid areas (Camus 1954; Schwarz 1964). The white oak section (*Lepidobalanus*), to which *Q. robur* and *Q. petraea* belong, is the most species-rich section of the genus consisting of more than 150 species spread over Europe, Africa, North America, and Asia. In the Netherlands only *Q. robur* and *Q. petraea* occur naturally. However, in arboreta (e.g. in Wageningen and Rotterdam) a high variety of *Quercus* species can be observed. The American northern red oak (also called eastern red oak or grey oak; black oaks) *Q. rubra* L. is very popular for road plantations. Until recently, these oaks were also popular for forest plantations.

Q. robur and *Q. petraea* are widespread throughout whole Europe. *Q. robur* and *Q. petraea* are two closely related sympatric species occupying different ecological niches (Rushton 1979; Grandjean & Sigaud 1987). *Q. robur* occurs on a wide range of soils, whereas *Q. petraea* prevails on elevated, drier soils. The two oak species can be recognized based on their leaf and fruit morphology (Rushton 1978, 1979; Van der Meijden 1990; Fig. 1.1). The two species are similar in their wood anatomy, however, significant differences between the two species have been found for the surface proportion of earlywood vessels, the number of earlywood vessels, and the size and surface proportion of fiber zones (Feuillat et al. 1997).

Q. robur and *Q. petraea* are diploid species (2n=2x=24) with a genome size of ±

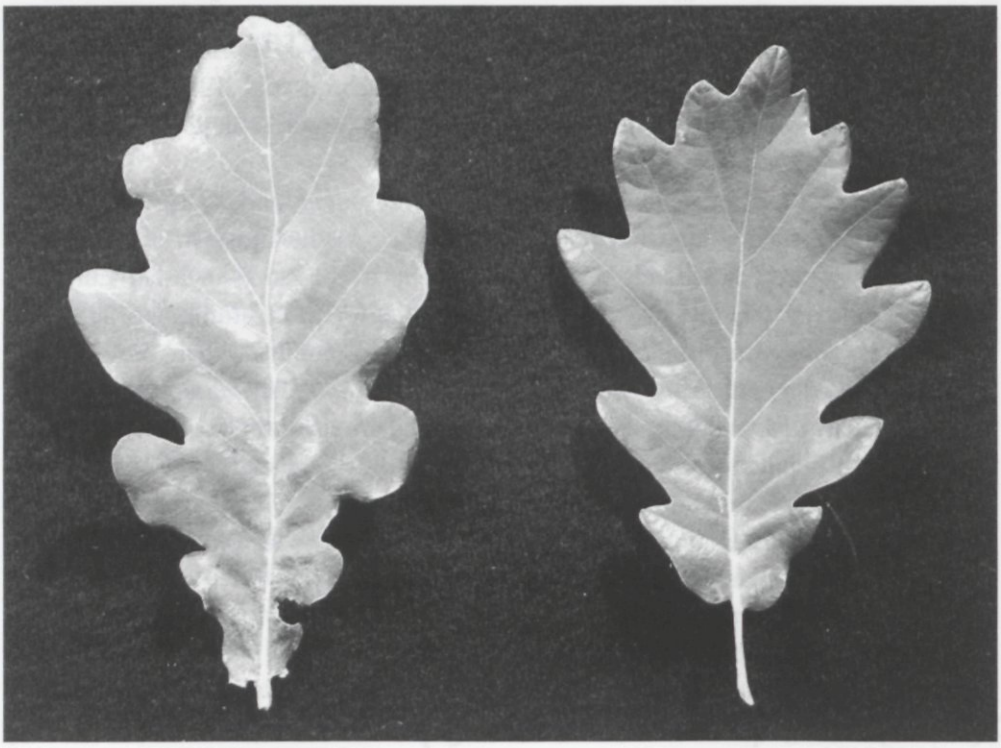


Figure 1.1
Q. robur and *Q. petraea* leaves (left and right leaf, respectively) differ in several morphological traits. (photo Alterra)

870 Mb (1.88 pg/2C) (Bennett & Leitch 1995; Favre & Arnould 1996; Favre & Brown 1996). They are obligatory outcrossing, wind pollinated, long-lived tree species. Both *Q. robur* and *Q. petraea* are monoecious species containing separate male and female flowers (Fig. 1.2). Reproduction takes place only by seeds, as the species do not vegetatively reproduce either by stem cuttings or by apomixis (Asker & Jerling 1992). Fruits are transported by blue jays (large distances) and rodents (small distances). In general the seeds that fall down the tree do not develop into trees, as seedlings cannot survive under a closed canopy (Oosterbaan & Van Hees 1989). The next generation of *Q. robur* and *Q. petraea* trees will establish itself in open areas. Therefore the long and short distance transport by animals is very important for the survival of the species. The chance of development of a next generation of *Q. robur* and *Q. petraea* is greatest when the fruit production is very high. This is the case when there is a mast year. Mast years occur on a regular basis with intervals of several years.

Quercus species are known for their ability to hybridize in many combinations, which has resulted in the existence of intermediate zones between species (Grant 1971; Whittemore & Schaal 1991). Hybridization between *Q. robur* and *Q. petraea* in natural populations has been reported (Streiff 1999) and controlled crosses between the two

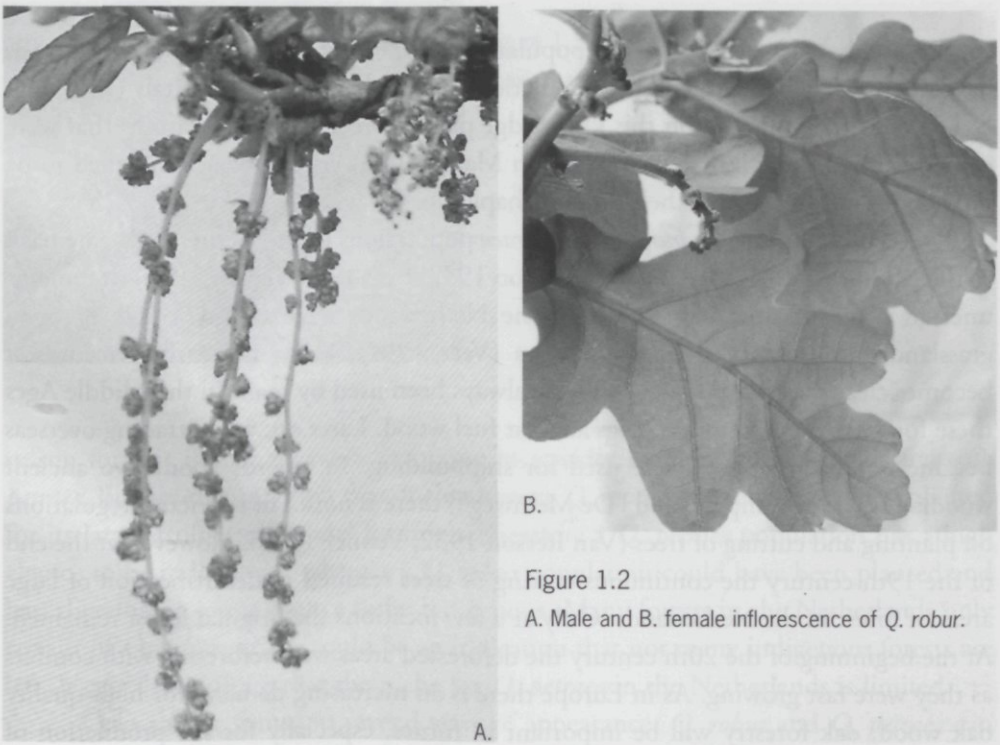


Figure 1.2
A. Male and B. female inflorescence of *Q. robur*.

species were successful (Steinhof 1993; Kleinschmit et al. 1995). Hybrids in natural populations are not easily recognized, as there is much confusion about the hybrid status of trees with intermediate characteristics (Gardiner 1970; Rushton 1978, 1979; Ietswaart & Feij 1989; Kleinschmit et al. 1995). Knowledge about hybridization and introgression in indigenous populations may be useful for a proper management and conservation of these populations. In forestry a clear identification of species and hybrids is required for wood and seed certification purposes.

History

After the last glaciation *Q. robur* and *Q. petraea* migrated northward and established themselves in the Netherlands from 8,000 B.P. onwards, as was found based on pollen records in the soil (Huntley & Birks 1983). Analyses of cpDNA haplotypes revealed the recolonization routes of *Q. robur* and *Q. petraea* for whole Europe since the last ice age (Petit et al. 1993; Dumolin-Lapègue et al. 1997; Petit et al. 1997). Oaks migrated from different refugia in the south of Europe northward. For the different refugia different cpDNA haplotypes were found. In the Netherlands the Atlantic haplotypes originating from southern Spain and a middle European haplotype originating from

southern Italy meet each other: the populations in the western part of the Netherlands have migrated from Spain and the populations in the eastern part from Italy (Van Dam & De Vries 1998). Based on this knowledge part of the 13 oak populations that were earlier claimed to be indigenous based on Maes' (1993) criteria were confirmed to be indigenous on the basis of their cpDNA haplotypes.

Records about *Q. robur* and *Q. petraea* populations in the Netherlands date back to the Middle Ages (1100 A.D.; Van Iterson 1932; Buis 1985;Venner 1985). It remains unclear if the original oak forests in the Netherlands were closed forests or open grassland with solitary trees and bushes (Vera 1997). From the earliest records it becomes clear that oak populations have always been used by man. In the Middle Ages these forests were used to feed pigs and for fuel wood. Later on, when trading overseas became popular the trees were used for shipbuilding. In records about two ancient woodlands ("De Stompert" and "De Meinweg") there is notice of restrictive regulations on planting and cutting of trees (Van Iterson 1932; Venner 1985). However, at the end of the 19th century the continuous cutting of trees resulted in desertification of large areas of forest in the Netherlands. Only at a few locations the original forest remained. At the beginning of the 20th century the deforested areas were reforested with conifers as they were fast growing. As in Europe there is an increasing demand for high-quality oak wood, oak forestry will be important in future, especially for the production of high-quality timber (Flies 1993).

Ecological values

Individual *Q. robur* and *Q. petraea* trees can reach ages up to about 600 years. Old (about 500-600 years old) *Q. robur* trees in the Netherlands can be found as single trees in Vorden, Ruurlo (Gelderland), and Oosterhout (Noord-Brabant) (Maes 1993). During their life oaks host a vast amount of epiphytes, insects, birds, and small animals. Acorns are eaten by blue jays, wild pigs, and rodents. Dead oak wood can serve for decades as a substrate for fungi.

Q. robur and *Q. petraea* can be distinguished based on their leaf morphology (Rushton 1978, 1979) (Fig. 1.1). Some of these leaf morphology traits could be eco-physiological adaptations of the two species to their different niches. The presence of abaxial hairs on *Q. petraea* leaves is expected to be an adaptation to a dry environment (reduction of water evaporation through leaves) as it has been shown that shading decreased hair density per unit leaf area, hair length, and maximum hair breath (Igboanugo 1992; Fig. 1.3). The long petiole of *Q. petraea* leaves could be an adaptation of the leaves to a more efficient reception of sunlight (less overshadowing of other leaves) (Takenaka 1994). This could explain why *Q. petraea* has been found to be more shade tolerant than *Q. robur* (Jones 1959).

Q. robur is more common in the Netherlands as compared to *Q. petraea*. One

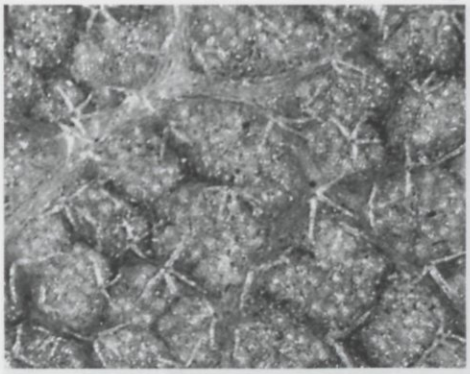


Figure 1.3
Abaxial hairs on *Q. petraea* leaf.

reason for this is that *Q. robur* can grow in a wider range of environments than *Q. petraea*. But another reason is that during history *Q. robur* has been favored and planted for its larger fruits (for animal feeding). Therefore a *Q. petraea* population has a high chance to be indigenous, whereas a *Q. robur* population could have been planted and has, therefore, a lower chance to be indigenous. Many forests in the Netherlands only consist of *Q. robur*, which could be an indication that not many indigenous forests are left. It can also indicate that the niche for *Q. petraea* in the Netherlands is limited.

Oaks can be found in several ways of appearance. *Q. robur* and *Q. petraea* can occur as a solitary tree reaching large sizes and old ages (e.g. solitary *Q. robur* tree in Vorden, The Netherlands). When managed in a proper way both *Q. robur* and *Q. petraea* can grow very tall with straight stems (e.g. "Het Liesbos" in the Netherlands and commercial oak forests in Croatia). When the trees are not managed properly they can develop many side branches and epicormic shoots (= small branches occurring directly from the main stem at lower heights). These characteristics are negative for forestry, but positive for ecology as they create niches for all kinds of plants and animals. *Q. robur* and *Q. petraea* can also be found as coppice stools in ancient woodlands (Fig. 1.4). Coppicing is the tradition of cutting trees every 7-10 years after which the trees sprout. Coppice practice has been common in many areas in Western Europe. This practice was used in the Netherlands from the Middle Ages onwards till the beginning of the 20th century (Van Iterson 1932; Buis 1985; Venner 1985). Coppice wood was used by farmers for fuel wood and wood for fences. In the 19th century oak coppice wood was extensively used for the leather industry (tannins in the bark). Due to coppicing the oak trees developed into clusters of thin, bend stems. The age of a coppiced population is hard to determine but some indications can be deduced from the size of the clusters of identical genotypes. Most of the time coppice practice was combined with animal grazing in the forest. This practice deteriorated the oak forests: trees disappeared and were replaced by heather or drift-sand (sand dunes) (Fig. 1.5). However, these landscapes possess an ecological value as well as a cultural value. They are ecologically very valuable, as these areas host endangered plants and animals. These areas are

Figure 1.4
Coppice stool in population "De Stompert".

culturally very valuable, as the landscape of coppice clusters surrounded by heather and sand dunes is a unique landscape that shows land-use in past centuries.

Natural rejuvenation of *Q. robur* and *Q. petraea* does not take place under closed canopy (Oosterbaan & Van Hees 1989). Only when a sufficiently large gap becomes available, oak seedlings can colonize it and reach full maturity. However, another tree species – beech (*Fagus silvatica*) – is shade tolerant and its seedlings survive many years under a closed canopy. When a tree dies and sunlight becomes available beech seedlings can increase rapidly in size and so outcompete the recently emerged oak seedlings. Therefore it is commonly thought that oak forests will eventually be replaced by beech forest.

Autochthonous *Q. robur* and *Q. petraea* populations could be valuable as they are assumed to be well adapted to the local environmental conditions. As autochthonous *Q. robur* and *Q. petraea* populations are rare, these populations could be conserved in order to stop the possible loss of genetic variation. In order to develop an appropriate management and conservation plan, the genetic structure and quality of these populations needs to be known. Molecular markers are most suitable for such kind of studies. For several European countries molecular marker studies involving allozymes (Zanetto et al. 1994; Samuel et al. 1995; Zanetto & Kremer 1995; Streiff et al. 1998; Degen et al. 1999), cpDNA polymorphisms (Petit et al. 1993), RAPD

Figure 1.5
Coppice stool surrounded by heather in population "De Meinweg". (photo Alterra)

(Moreau et al. 1994; Bodénès et al. 1997a), SCAR (Bodénès et al. 1997a,b), and microsatellites (Streiff et al. 1998; Degen et al. 1999) have been conducted in order to study the genetic variation and structure of indigenous *Q. robur* and *Q. petraea* populations.

Information about the genetic variation of a population tells us not only about the present state of a population, but also about the expected variation in the next generation when mating between trees is random. A high level of variation is required for future generations to be able to adapt themselves to a changing environment. So far, genetic variation has been studied by means of molecular markers that are assumed to be neutral. However, there is much debate about neutrality as many molecular data do not fit the neutral theory very well (Mortitz & Hillis 1996). Genetic information obtained from neutral markers gives an indication about the level of genetic variation, inbreeding, population differentiation (number of migrants), and population structure. In order to be able to study genetic variation in eco-physiological traits it is necessary to find genetic markers that are closely linked to the trait of interest. This can be done by either multiple regression analysis of molecular markers on eco-physiological traits or by constructing a genetic linkage map and mapping quantitative trait loci (QTLs) on the map.

Economical values

Since the earliest civilization oaks were used by man. For the Netherlands there are records of pottery filled with charred acorns found in Bronze Age settlements (Van Zeist 1970). The economical values of oaks involved fruits for animal feeding, wood for fuel and timber (in the 17th century a lot of oak wood was used for the ship building industry), and bark for the production of tannins for the leather industry (19th century). The wood of properly managed oak forests can be sold at a high price for veneer in the furniture industry. As oak is a hardwood species, the wood can be used for several sustainable applications. In the Netherlands oak is used in forest production areas (about 16% of the total forested area is covered with oak as main tree species), roadside plantations (about 25% of all lane- and roadside plantations are planted with oak), and in landscaping. Dutch seed stands are used for oak afforestation and reforestation in large parts of Western Europe. About 400,000 kg of acorns is harvested per year. About half is exported as seed and the other half is sown in Dutch nurseries. The majority of the young plants is exported to neighboring countries (Jensen et al. 1997; De Vries & Van Dam 1998). Seed stands are tested for a production of high quality seeds by comparing the offspring in different provenance trials. In the Netherlands there is one seed orchard of oak. For a seed orchard seeds are collected from selected 'plus' trees. These trees are expected to pollinate each other, which may result in superior offspring (De Vries & Van Dam 1998). In Austria efforts have been made in order to certify seed batches (Lexer et al. 1999, 2000). The application of the microsatellite technique for the analysis of seed batches was successful. Microsatellites could identify the number of mother trees included in seed harvests and could also detect seed contamination with unrelated seedlings (Lexer et al. 1999). Based on microsatellite analysis the number of contributing pollen donors in commercial seedlots could be monitored (Lexer et al. 2000).

In order to be able to unambiguously certify *Q. robur* and *Q. petraea* wood or seedlings, there is a need for an objective technique that can distinguish between the two species, especially in the case of intermediate type of trees. As molecular markers can detect genetic differences it is expected that molecular markers might be able to unambiguously discriminate between the two species. So far, studies involving molecular markers were not successful to detect a diagnostic marker (Kremer et al. 1991; Petit et al. 1993; Zanetto et al. 1994; Moreau et al. 1994; Bacilieri et al. 1995; Kleinschmit et al. 1995; Samuel et al. 1995; Barreneche et al. 1996; Bodénès et al. 1997a,b; Streiff et al. 1998; Muir et al., 2000). These studies show that the two species are highly related and that it will take, therefore, a lot of effort to find a diagnostic marker that can differentiate the two species.

Certification of autochthonicity is important for trade in autochthonous plant material. As autochthonicity of oak seedlings is hard to be proven, molecular marker

techniques like PCR-RFLP (for the detection of cpDNA polymorphisms) and microsatellites are important tools for this matter. Based on cpDNA studies for oaks in the Netherlands it became clear that in the western part of the Netherlands the Spanish lineage occurs and in the eastern part the Italian lineage (Van Dam & De Vries 1998). Based on this information it is possible to find out about plantation from the Balkans or other remote areas (Dumolin-Lapègue et al. 1997; Konig et al. *in press*). If any other cpDNA haplotype – besides the Spanish and the Italian haplotypes – is found within a batch of seedlings, it will be clear that seeds were not harvested from indigenous trees. However, this test will not rule out the possibility of plantings of the Spanish and Italian haplotypes from southern Spain or southern Italy. However, it is not expected that there has been much transport of planting material from Southern Europe to the Netherlands as southern material is less tolerant to frost. Seedlings can be tested to be true descendants from a claimed indigenous tree by means of maternity analysis. By means of paternity analysis the seedlings can be tested for the paternal contribution. Paternity analysis can find out about the percentage of pollinations with father trees from outside the indigenous population.

In forestry it takes many years to select trees, as some traits are visible only at a later age. This is one of the reasons why a plantation is generally started with many oak seedlings that are thinned after several selection rounds. Another practice in forestry is provenance testing: seeds are harvested from different locations and seedlings are planted in a trial field where the different provenances can be compared (in a block design) and tested for their quality. In order to calculate heritabilities open-pollinated progeny tests can be used. Such tests involve a randomized block design with several replications containing the offspring of individual trees from different provenances (Jensen et al. 1997). When this test is done on several locations genotype by environment interaction can be studied. However, when seeds from different trees are used and these are planted at different locations the test does not contain replications of the same genetic material. Besides, it takes many years before the seedlings are sufficiently large to do observations. A more efficient approach could be the construction of a genetic linkage map on which traits of interest (tree height, tree growth rate, flushing stadium, etc.) will be mapped as quantitative trait loci (QTLs). In this way a better understanding of the heredity of these traits can be obtained and for some traits – if a marker is closely linked to the QTL – selection of trees and provenance testing might be done at the seedling stage. When the full-sib family is vegetatively propagated it can be transplanted in different environments. In this way genotype by environment interaction will be understood better.

Future

It is feared that in future the survival of *Q. robur* and *Q. petraea* trees in the Netherlands

will face problems. This fear is partly based on predictions of climate change due to global warming. These predictions indicate that when temperature increases due to global warming many tree species in the Netherlands will face problems and will be outcompeted by species that are better adapted to the changed climatological circumstances in the Netherlands (Nabuurs et al. 1997). Besides, there are records of premature oak decline due to drought and insect attack dating from 1983 onwards (Moraal 1997; Oosterbaan et al. 2001). Premature oak decline mainly occurs at areas where the groundwater level is high and the roots of the trees are continuously submerged in the groundwater. In this way, parts of the root system decay, which leads to damage of the trees in drought periods. Trees that are weakened in this way are attacked by insects, which often leads to the death of the trees. Reports about a new insect (*Agrilus biguttatus*) originating for Southern Europe indicate that this insect develops an epidemic growth and has so far not been susceptible to any action taken (Moraal 1997). As due to climatic change the temperature in the Netherlands will rise, it is expected that the problems of drought and insect attack will increase.

As indigenous planting material has been adapted to the local situation over hundreds of years it is likely to perform better as compared to imported planting material. There are examples of bad performance in the Netherlands of imported planting material that was of a high quality in the country of origin. However, the opposite can also happen, which indicates that a general rule about indigenous planting material cannot be given (Maes 1993). If we assume that adaptation is indeed a key-issue, new plantations should use indigenous planting material.

Molecular markers

In above sections questions concerning ecology, conservation, and forestry of oak were met that could be answered with the help of molecular data. For this purpose many molecular markers are available (Hillis et al. 1998). However, for each specific question another type of molecular marker technique should be chosen. Different molecular marker systems describe different levels of genetic variation. Low levels of genetic variation can be found in highly conserved regions of the genome. Higher levels of genetic variation can be found in coding regions for adaptive traits (adaptation to environmental change). High levels of genetic variation can be found for non-coding repeats: microsatellites and minisatellites, which differ for the length of the repeat: microsatellites are built from short motifs, minisatellites are built from longer motifs. Minisatellites or VNTRs (variable number tandem repeats) are short (9-64 bp) tandemly repeated sequence units (Jarman and Wells 1989). Due to a process called unequal crossing over (between sister chromatids or homologous chromosomes) the number of repeated sequences can increase or decrease after mitosis or meiosis. Within the genome the repeats are monomorphic and highly conserved by a process called

concerted evolution. Allelic variation between the number of repeats can be high. Microsatellites are short di- or tri-nuleotide repeats (less abundant are repeats consisting of up to six nucleotides). Different individuals can differ for the number of repeats in the fragment. This high variation in number of repeats is caused by strand slippage of DNA polymerase during DNA replication (Kashi et al. 1997).

In this thesis two molecular marker techniques are used: AFLP and microsatellites. The AFLP technique is based on DNA fragments generated after digestion with a combination of restriction enzymes. These restriction enzymes cut DNA at many places randomly dispersed over the genome. Different individuals can differ for the recognition sequences for the restriction enzymes leading to varying fragment lengths after digestion with restriction enzymes. In this way genetic variation over the whole genome is described (Vos et al. 1995). The microsatellite technique is based on DNA fragments amplified from regions on the genome that consist of microsatellite repeats. Primers are designed to anneal to areas on the genome surrounding the microsatellite repeat. Different individuals can differ for the length of the amplified repeat. As the microsatellite technique describes microsatellite repeat variation it is expected that this technique describes higher levels of genetic variation than the AFLP technique that describes genetic variation at random locations over the genome. Besides the level of genetic variation detected, the two molecular marker techniques differ in their dominance. AFLP is a dominant molecular marker technique, which means that – without sophisticated image analysis software – no difference can be detected between homozygote dominant and heterozygote genotypes. Markers can only be scored for the presence or absence of a band (Fig. 1.6). Only when conditions can be completely controlled, it is possible to score the AFLP bands for their intensity: bands with high intensity are homozygote dominant, bands with low intensity are heterozygote bands. Especially for the construction of genetic linkage maps where the parents of the population are known and the segregation pattern can be predicted this way of scoring AFLP bands can be attempted. The microsatellite technique is a co-

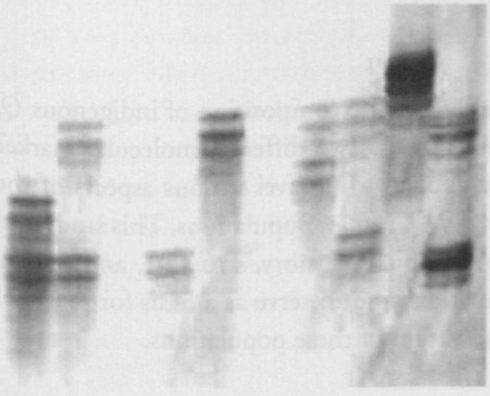


Figure 1.7
Microsatellite fingerprints

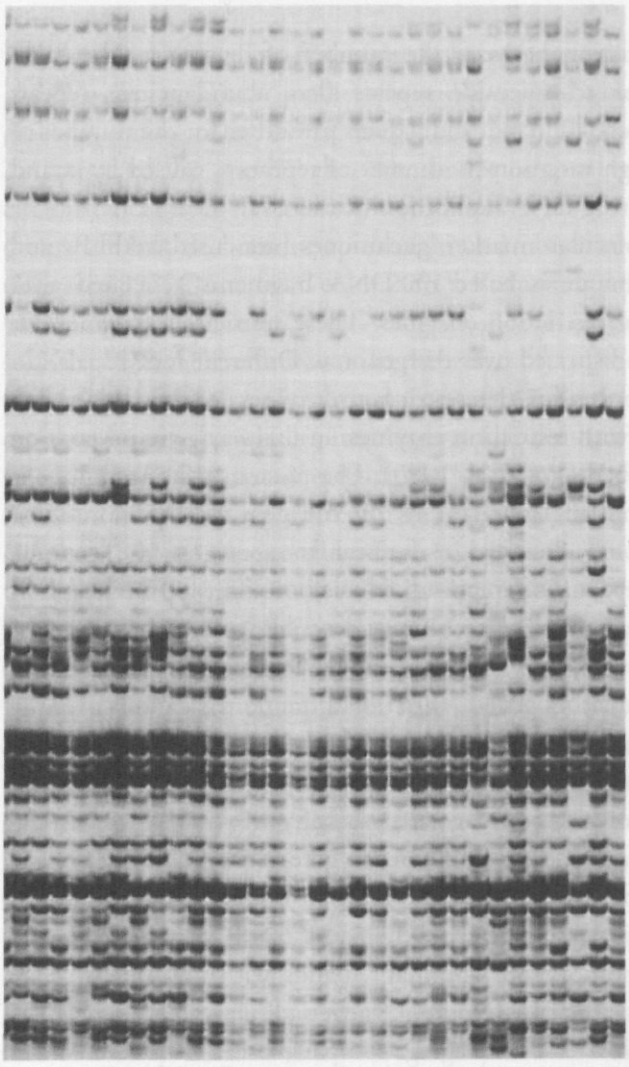


Figure 1.6
AFLP fingerprints.

dominant molecular marker technique, which means that all combinations of alleles can be detected (Fig. 1.7).

Aim of the study

The aim of this Ph.D. thesis was to describe the genetic composition of indigenous *Q. robur* and *Q. petraea* populations in the Netherlands. Two different molecular marker techniques (AFLP and microsatellites) were applied to unravel various aspects of the genetic composition of indigenous *Q. robur* and *Q. petraea* populations. This study was conducted in order to understand the autochthonicity, history, structure, and quality of indigenous *Q. robur* and *Q. petraea* populations and to serve as a basis for practical guidelines for future management and conservation of these populations.

In the following chapters results of several studies of AFLP and microsatellites in indigenous *Q. robur* and *Q. petraea* populations with possible applications to ecology and forestry are presented. In chapter 2 the genetic aspects of species differentiation between *Q. robur* and *Q. petraea* is studied by means of AFLP. In the same chapter a search for diagnostic markers is conducted as diagnostic markers can be used for ecological studies about the hybrid nature of intermediate type trees and for forestry (seed and wood certification). In chapter 3 the size, number, and distribution of oak clones in coppiced and grazed ancient *Q. robur* and *Q. petraea* woodlands is studied by means of AFLP and microsatellites. As most indigenous oak populations have a history of coppicing and grazing, information about clones in these populations is necessary for management and conservation. In chapter 4 the genetic variation of indigenous oak populations is described and compared with results based on the same microsatellite markers obtained for two half-sib families. Information about the genetic variation of a population gives an indication for the variability of next generations and therefore the possibility of the population to adapt itself to changing environmental and climatological conditions. In chapter 5 the inheritance of eco-physiological and morphological traits is studied by means of the construction of a genetic linkage map of oak. Markers that are found to be closely linked to a trait of interest may be used in ecology and forestry. In the last chapter (chapter 6) the use of the two different molecular marker techniques and their applications to ecology, forestry, and tree breeding are discussed.

2 A discrimination between *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. based on species-indicative AFLP markers

Erica G. Bakker[1,2], *Barbara C. Van Dam*[1], *Herman J. Van Eck*[2], *Evert Jacobsen*[2]

Summary

In natural populations *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. comprise a morphological continuum due to hybridization and/or an overlap in variation between the two species. In order to obtain diagnostic markers to investigate this morphological continuum, leaf morphology and AFLP were evaluated for their ability to discriminate between these two species in an autochthonous population in the Netherlands. Multivariate statistical analyses revealed a differentiation between the species based on leaf morphology data as well as AFLP data. Discriminant analysis resulted in the detection of only three out of 13 studied leaf morphology traits to be involved with species discrimination. None of the 92 polymorphic AFLP markers were diagnostic, however, there were 13 markers exhibiting significant marker band frequency differences among which there were five species-indicative markers. Regression analyses of AFLP markers on each of the 13 leaf morphology characteristics resulted in significant associations between groups of AFLP markers and leaf morphology traits. This study indicates that *Q. robur* and *Q. petraea* are closely related and probably only differ for a few genes coding for leaf morphology traits.

Keywords: AFLP, leaf morphology, species discrimination, *Quercus petraea*, *Quercus robur*

Introduction

Quercus robur L. (pedunculate oak) and *Q. petraea* (Matt.) Liebl. (sessile oak) (Fagaceae) are two closely related species occurring in most parts of Europe as sympatric species occupying different ecological niches (Rushton 1979; Grandjean & Sigaud 1987). Various studies have tried to discriminate between these two species based on morphological characteristics (Rushton 1978, 1979; Grandjean & Sigaud 1987; Ietswaart & Feij 1989). However, none of these studies were able to distinguish

[1] *Department of Ecology and Environment, Alterra, P.O. Box 47, 6700 AA, Wageningen, The Netherlands*
[2] *Laboratory of Plant Breeding, Wageningen University, P.O. Box 386, 6700 AJ, Wageningen, The Netherlands*

between the two species without ambiguity. The reason for this is the occurrence of intermediate types possessing morphological characteristics of either of the two species. It is assumed that these intermediate types have resulted from hybridization and backcrossing between the two species (Rushton 1978; Ietswaart & Feij 1989). However, in view of the infrequent occurrence of hybridization, a part of the intermediate types probably represents the wide range and overlap of variation of the two species (Gardiner 1970). On the other hand, hybrids and introgression products might not necessarily possess intermediate morphological characteristics, as they may also resemble one of the parents (Kleinschmit et al. 1995).

Promising tools for the description of species differentiation are molecular marker techniques. Therefore many molecular marker techniques (allozymes, cpDNA, RAPD, SCAR, two-dimensional gel electrophoresis of proteins, microsatellites) have been used in order to obtain markers that might discriminate between *Q. robur* and *Q. petraea* trees. So far, none of these marker techniques could provide a diagnostic marker (a marker that is present in all individuals of one species, but is absent from all individuals of another species) for species identification (Kremer et al. 1991; Petit et al. 1993; Zanetto et al. 1994; Moreau et al. 1994; Bacilieri et al. 1995; Kleinschmit et al. 1995; Samuel et al. 1995; Barreneche et al. 1996; Bodénès et al. 1997a,b; Streiff et al. 1998; Muir et al. 2000). However, for some marker loci the species were discriminated based on significant marker band frequency differences (markers occur at a low frequency in one species and at a high frequency in another species) between the two species (Zanetto et al. 1994; Moreau et al. 1994; Bacilieri et al. 1995; Barreneche et al. 1996; Bodénès et al. 1997a). In none of these studies significant species-indicative markers (markers that occur in only one species, but not in all individuals, and are absent in all individuals in another species) were found.

A new highly reproducible technique for DNA fingerprinting, AFLP™, is able to efficiently generate large numbers of markers (Vos et al. 1995). The AFLP technique has been suitable for biosystematic studies in *Pisum* and *Solanum* as it is discriminative from the individual genotypic level to the species level (Lu et al. 1996; Kardolus et al. 1998). Next, AFLP was successfully used to analyze hybridization and invasion in populations of weedy *Onopordum* thistles (O'Hanlon et al. 1999).

In this study, leaf morphology variation and AFLP polymorphisms are compared with each other for their ability to differentiate between *Q. robur* and *Q. petraea* trees in an autochthonous population (the population has occurred in the area since its establishment after the last ice age), "De Meinweg". First, the subset of leaf morphology traits involved in species discrimination will be identified. This will be followed by a search for diagnostic AFLP markers that can be used to investigate trees with intermediate morphological traits. Finally, the associations between leaf morphology variation and AFLP polymorphisms will be analyzed.

Materials and methods

Study area

The study area, "De Meinweg" (state survey co-ordinates x/y 207.5/354.8), is located in the south of the Netherlands on the slope of an old river-bed. "De Meinweg" is an autochthonous population that can be characterized as an old (devastated) woodland (Venner 1985; Maes 1993; Van Dam & De Vries 1998). *Q. robur* occurred in the entire area, while *Q. petraea* was only found in the higher parts. *Q. robur* and *Q. petraea* trees (N=48 for both species) were sampled after an evaluation in the field based on a few species characteristics (basal shape of the lamina and the level of abaxial hairiness) and habitat characteristics (*Q. petraea* only occurs on elevated, nutrient poor soils) as described in Van der Meijden (1990). In this way putative hybrid trees with intermediate characteristics (*Q. robur* type basal shape and *Q. petraea* type level of abaxial hairiness and vice versa) were excluded. The positions of the sampled trees were recorded. The map was subsequently digitized with ArcView® GIS (Environmental Systems Research Institute Inc., Redlands, USA; Fig. 2.1).

Leaf morphology analysis

Five fully expanded leaves were sampled from different sides from the crown of each tree. The leaves were evaluated for their level of abaxial hairiness (HR) on a scale from 1 (no hairs at all) to 6 (densely hairy), and subsequently dried and stored. After drying, another eight leaf morphology characteristics were scored: lamina length (LL); petiole length (PL); lobe width (LW); sinus width (SW); length of lamina from the lamina base to the widest part (WP); number of lobes (NL); number of intercalary veins (NV); basal shape of the lamina (BS; measured on a scale from 1 (wedge-shaped) to 9 (two clear lobes)). Subsequently, four derived characteristics were calculated: lamina shape, OB=LL/WP; petiole ratio, PR=PL/(LL+PL); lobe depth ratio, LDR=LW/(LW-SW); percentage venation, PV=NV*100/NL (Rushton 1978).

AFLP analysis

DNA was extracted from fresh leaves or buds with a DNA extraction kit (Puregene®, Gentra Systems, Minneapolis, USA) including 4% PVP-40 to remove phenolic compounds. AFLP analysis was performed according to Vos et al (1995). Primers including one selective nucleotide (Eco+A and Mse+C) were used for pre-amplification of the template. The selective amplification was performed with γ33P-ATP labeled primers Eco+AAG or Eco+ATA in combination with the unlabeled primer Mse+CCC. After electrophoresis the gels were vacuum dried on 3 mm Whatman paper and

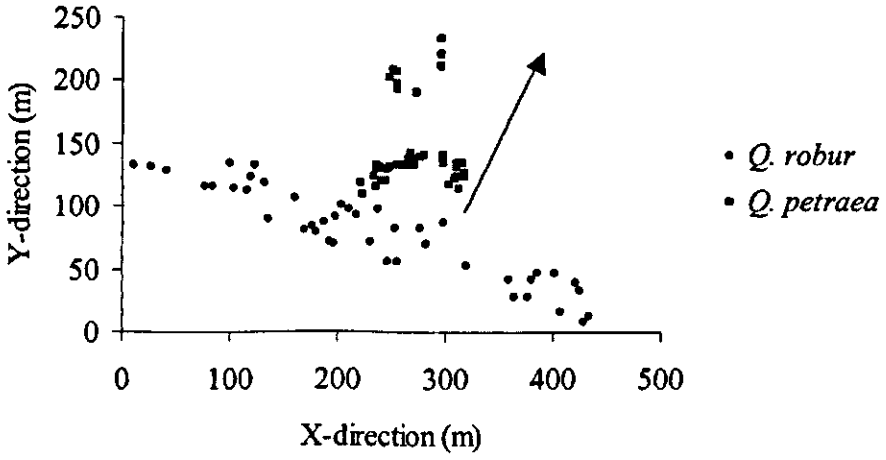


Figure 2.1
Map of "De Meinweg" with the locations of the trees (dots are *Q. robur*, squares are *Q. petraea*). The arrow indicates the direction of the slope with a height difference of 10 m.

subsequently exposed to X-ray films (Kodak).

Reproducibility of AFLP was tested on five individuals (two *Q. robur* and three *Q. petraea*). For each of the five individuals AFLP fingerprints obtained from DNA samples from two different tissues (buds and leaves) from different years (1996 and 1997) were compared.

Associations between the two species and leaf morphology variation

The averages of the leaf morphology traits per tree were plotted for each species separately in histograms. Subsequently, significant differences were tested between the averages of the two species distributions per trait by means of t-tests or Mann-Whitney tests dependent on the assumption of normality of the data using Genstat 5 (Release 4.1; Lawes Agricultural Trust, Rothamsted, UK). The data were analyzed for a sub-structuring in species based on the nine directly measured traits by means of a principal component analysis (PCA) with a correlation matrix using Genstat 5 (Release 4.1; Lawes Agricultural Trust, Rothamsted, UK). This was followed by a discriminant analysis of all 13 leaf morphology traits based on a pooled covariance matrix with a stepwise selection procedure of variables ($F_{IN}>3.84$, $F_{OUT}<2.71$, which is comparable with an approximate threshold value α for $F_{IN}<0.05$, for $F_{OUT}>0.10$) with SPSS 9.0 (SPSS Inc., Chicago, USA). Subsequently, the leaf morphology data were normalized

(subtracted with the minimum and this divided by the range). Averages of the Euclidean distances between all possible pairs of trees within and between species were calculated using program Phyltools (J. Buntjer, Laboratory of Plant Breeding, Wageningen University, Wageningen, The Netherlands).

Associations between the two species and AFLP polymorphisms

The AFLP fingerprints were scored for the presence (1) and absence (0) of bands. The data were analyzed for a sub-structuring in species by means of a correspondence analysis (CA) using Genstat 5 (Release 4.1; Lawes Agricultural Trust, Rothamsted, UK). Subsequently, a discriminant analysis based on a pooled covariance matrix with a stepwise selection procedure of variables was carried out (F_{IN}>3.84, F_{OUT}<2.71, which is comparable with an approximate threshold value α for F_{IN}<0.05, for F_{OUT}>0.10) using SPSS 9.0 (SPSS Inc., Chicago, USA).

Averages of Jaccard's (1908) genetic distances between all possible pairs of trees within and between species were calculated with program PhylTools (J. Buntjer, Laboratory of Plant Breeding, Wageningen University, Wageningen, The Netherlands). Significant band frequency differences between the two species were analyzed using the Fisher exact procedure of Genstat 5 (Release 4.1; Lawes Agricultural Trust, Rothamsted, UK). The threshold of 5% was adjusted according to the sequential Bonferroni test as proposed by Rice (1989).

Associations between leaf morphology variation and AFLP polymorphisms

Regression analyses of AFLP markers on each of the 13 leaf morphology traits were performed in order to analyze associations between leaf morphology variation and AFLP polymorphisms. This was done for each species separately. Associations between AFLP markers and quantitative leaf morphology characteristics were analyzed by means of multiple linear regression analyses (F_{IN}>3.0, F_{OUT}<3.0, which is comparable with an approximate threshold value α for F_{IN}<0.10, for F_{OUT}>0.10) with the STEP procedure of Genstat 5 (Release 4.1; Lawes Agricultural Trust, Rothamsted, UK). Level of abaxial hairiness (HR) and basal shape (BS) were analyzed by means of GLM analyses assuming a poisson distribution (log as link function) of these response variables (F_{IN}>3.0, F_{OUT}<3.0, which is comparable with an approximate threshold value α for F_{IN}<0.10, for F_{OUT}>0.10) with the STEP procedure of Genstat 5 (Release 4.1; Lawes Agricultural Trust, Rothamsted, UK).

Results

AFLP analysis

A total data set was obtained with 110 different amplification products of which 16% were monomorphic for all samples. AFLP resulted in fingerprints with an average number of 43 marker bands per sample. The reproducibility study showed that AFLP fingerprints were 100% identical within one genotype and not affected by tissue source, year of sampling, or day-to-day variation in the lab. Identical AFLP fingerprints were obtained from different pairs of trees suggesting clonal maintenance that originated due to coppicing and animal grazing (unpublished results). From each pair of trees with identical genotypes one tree was removed. To avoid an unbalanced data set of *Q. robur* and *Q. petraea* genotypes, two additional *Q. robur* genotypes were removed, thus leading to a total sample size of 43 *Q. robur* and 43 *Q. petraea.*

Associations between the two species and leaf morphology

All leaf morphology traits showed overlapping bimodal distributions for the two species. For leaf morphology traits BS, HR, LDR, LL, NL, NV, PL, PR, PV, and SW highly significant differences were observed between *Q. robur* and *Q. petraea* ($p<0.005$). Whereas for leaf morphology traits LW and WP the difference observed between *Q. robur* and *Q. petraea* was less significant ($p<0.05$). No difference between *Q. robur* and *Q. petraea* was found for OB. PCA based on the nine directly measured leaf morphology traits resulted in two clear groups corresponding to the previously identified species (Fig. 2.2). The first axis was responsible for the separation of the two species. All nine variables showed high loadings. The highest loadings were observed for: SW, HR, PL, and NL. Discriminant analysis resulted in the inclusion of HR, BS, and NV in the discriminant function. The average Euclidean distance between *Q. robur* trees was significantly lower ($p<0.05$) than between *Q. petraea* trees (Tab. 2.1). The average Euclidean distance between trees from different species was significantly higher ($p<0.05$) than between trees within each of the two species (Tab. 2.1).

Associations between the two species and AFLP polymorphisms

Correspondence analysis (CA) based on AFLP data resulted in a differentiation between the two species (Fig. 2.3). However, one individual that was previously assigned as *Q. robur* was after CA located in the *Q. petraea* cluster. The first axis was responsible for the differentiation between two groups corresponding to the two species. Discriminant analysis revealed that 19 AFLP markers were involved in the differentiation between the two species. The average Jaccard's (1908) genetic distance

between *Q. robur* trees was significantly lower ($p<0.05$) than between *Q. petraea* trees (Tab. 2.1). The average Jaccard's (1908) genetic distance between trees from different species was significantly higher ($p<0.05$) than between trees within each of the two species (Tab. 2.1).

The Fisher exact test for association between the presence/absence of an AFLP band and a classification of *Q. robur*/ *Q. petraea* resulted in the detection of 13 markers with significant differences in marker band frequency between the two species ($p<0.05$). From among these 13 markers, five were significant species-indicative markers (Tab. 2.2). However, only seven out of the 19 AFLP markers that were included in the discriminant function were also observed as significant markers in the Fisher exact test. For both analyses the same marker was the most significant for species discrimination: Eco+AAG/Mse+CCC-235. This marker was not species-indicative.

Associations between leaf morphology variation and AFLP polymorphisms

As for regression analyses the number of variables should not be higher than the number of individuals tested (Montgomery and Peck 1992), the set of 76 and 72 polymorphic AFLP markers for *Q. robur* and *Q. petraea*, respectively, had to be reduced. Therefore all AFLP markers with less than five present or absent bands were removed as these markers were expected to be less informative. After removing all markers with less than five present or absent bands we retained a set of 43 markers for both *Q. robur* and *Q. petraea*. Correlation analysis did not show any pairs of AFLP markers with correlation coefficients higher than 0.9. Therefore no AFLP markers were further removed from the data set. Three leaf morphology characteristics (NV, NL, and LDR) were transformed (square root, square root, and 10log, respectively) in order to

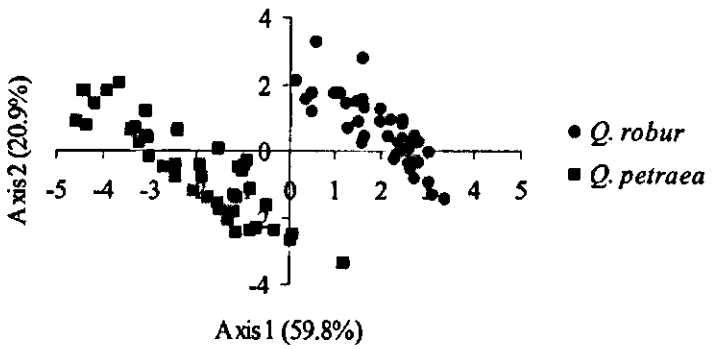


Figure 2.2 PCA
based on nine leaf morphology traits (dots are *Q. robur*, squares are *Q. petraea*).

Table 2.1

Average Euclidean distances based on 13 leaf morphology traits (D_{EU}) and average Jaccard's (1908) genetic distances based on 110 AFLP markers (D_{JC}) within and between *Q. robur* and *Q. petraea*.

SPECIES	D_{EU}	D_{JC}
Q. robur	0.65 ± 0.034 (SE)a[2]	0.35 ± 0.008 (SE)d
Q. petraea	0.84 ± 0.041 (SE)b	0.38 ± 0.009 (SE)e
Between[1]	1.63 ± 0.025 (SE)c	0.41 ± 0.005 (SE)f

1) Between *Q. robur* and *Q. petraea*.
2) Average Euclidean distances or Jaccard's genetic were marked with different letters in order to indicate that they differed significantly ($p<0.05$).

obtain a homogeneous residual variation. Regression analysis of the 43 *Q. robur* and 43 *Q. petraea* AFLP markers on the 13 leaf morphology characteristics resulted in models consisting of up to 8 AFLP markers with percentages of explained variation up to 59.5% (Tab. 2.3). For *Q. robur* from among the 43 analyzed AFLP markers 25 were significantly ($p<0.05$) associated with up to four leaf morphology traits. For *Q. petraea* from among the 43 analyzed AFLP markers 27 were significantly ($p<0.05$) associated with up to four leaf morphology traits.

Among the 19 AFLP markers that were in the discriminant function, three and seven AFLP markers were found to be significantly associated with leaf morphology traits for *Q. robur* and *Q. petraea*, respectively. However, 13 *Q. robur* and five *Q. petraea* AFLP markers that were in the discriminant function were not included in the

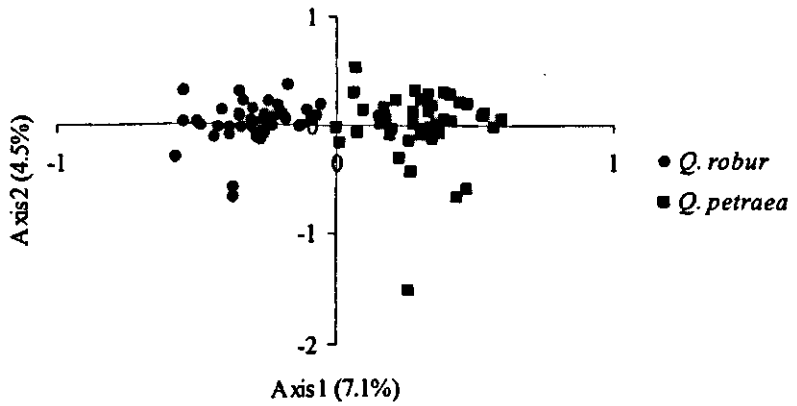


Figure 2.3

CA based on 110 AFLP markers (dots are Q. robur, squares are Q. petraea).

Table 2.2
Five species-indicative AFLP markers and their frequencies of occurrence in *Q. robur* and *Q. petraea*.

MARKER	*Q. robur*	*Q. petraea*
Eco+AAG/Mse+CCC-138	0	0.42
Eco+AAG/Mse+CCC-186	0.35	0
Eco+ATA/Mse+CCC-120	0.37	0
Eco+ATA/Mse+CCC-257	0	0.44
Eco+ATA/Mse+CCC-360	0.53	0

regression analyses as these markers possessed less than five present or absent bands in the data set.

Discussion

Associations between the two species and leaf morphology variation

The species distributions of the leaf morphology characteristics were all overlapping. This indicates that for a discrimination between *Q. robur* and *Q. petraea* it is necessary to take a multivariate approach. PCA based on nine leaf morphology traits could make a clear grouping between the two species. Our study shows that only three leaf morphology characteristics were sufficient for a discrimination between *Q. robur* and *Q. petraea* trees. However, previous studies used many leaf morphology characteristics in order to define *Q. robur* and *Q. petraea* populations (Rushton 1978, 1979; Grandjean & Sigaud 1987; Bacilieri et al. 1996). This has probably been done as when intermediate type trees are included, it might be necessary to analyze more leaf morphology characteristics.

Associations between the two species and AFLP polymorphisms

Correspondence analysis (CA) based on 110 AFLP markers resulted in a grouping of the two species. However, one *Q. robur* individual was located in the *Q. petraea* cluster. This could be an introgression product that resembled the *Q. robur* parent as has been observed by Kleinschmit et al. (1995).

The average Jaccard's (1908) genetic distance within *Q. robur* was significantly lower than the average Jaccard's (1908) genetic distance within *Q. petraea*. The same result was observed after two-dimensional protein gel, RAPD, and SSR analyses of *Q. robur* and *Q. petraea* trees in natural mixed populations (Moreau et al. 1994; Kleinschmit et al. 1995; Barreneche et al. 1996; Bodénès et al. 1997a; Streiff et al.

Table 2.3

Number of AFLP markers explaining the various leaf morphology traits in regression models.

TRAIT	NUMBER OF AFLP MARKERS		% EXPLAINED VARIATION	
	Q. robur	*Q. petraea*	*Q. robur*	*Q. petraea*
HR	0	0	-	-
LL	5	1	42.8	20.7
PL	4	1	32.8	11.2
LW	4	6	36.3	43.1
SW	2	4	23.4	30.0
WP	5	0	29.9	-
SqrtNL	2	7	15.8	55.9
SqrtNV	7	5	46.1	50.6
BS	0	0	-	-
OB	1	8	10.0	56.4
PR	3	0	29.9	-
LogLDR	7	6	51.1	42.0
PV	5	6	50.8	59.5

1998). However, allozyme analyses gave contradictory results: whereas Zanetto et al. (1994), Kleinschmit et al. (1995), and Streiff et al. (1998) showed a higher genetic diversity for *Q. petraea* than for *Q. robur*, Bacilieri et al. (1995) and Samuel et al. (1995) showed a lower genetic diversity for *Q. petraea* than for *Q. robur*. This lower genetic diversity for *Q. petraea* was also found for SCAR analysis (Bodénès et al., 1997b). This discrepancy in genetic variation observed for the two species is probably caused by the different characteristics of the various molecular marker methods. It appears that molecular markers that are not coding like two-dimensional protein gel analysis, RAPD, microsatellites, and AFLP give more consistent results than the protein and DNA markers that represent coding regions. However, as these studies used different sampling schemes, this could be another reason for the discrepancy in genetic variation observed for the two species.

Although no diagnostic markers were observed, we found significant band frequency differences (for 13 AFLP markers) between the two species. The same has been reported for allozymes (Zanetto et al. 1994; Bacilieri et al. 1995), two dimensional gel electrophoresis of proteins (Barreneche et al. 1996), and RAPD (Moreau et al. 1994; Bodénès et al. 1997a). Like in this study, these studies excluded intermediate types. However, it is possible that the species definitions varied between these studies. Therefore, it is not possible to compare the effectiveness of AFLP to discriminate between the two species in our study with the effectiveness of other molecular marker techniques in previous studies. Although no diagnostic markers could be found, we were able to find five significant species-indicative markers. The band frequencies for these markers were high and ranged between 0.35 and 0.53. Previous studies based on other molecular markers did not detect significant species-

indicative markers for *Q. robur* and *Q. petraea*. Therefore, we conclude that AFLP is so far the best molecular marker method for discrimination between *Q. robur* and *Q. petraea*. Still, the band frequencies of the species-indicative markers are probably not high enough to be used for hybrid studies like the study of a hybrid zone between *Q. grisea* and *Q. gambelii* by Howard et al. (1997) where the RAPD band frequencies for species-indicative markers were all above 0.61. Moreover, hybrids can best be studied with diagnostic markers. This was done by O'Hanlon et al. (1999) where diagnostic AFLP markers were used to investigate hybrids between *Onopordum* species.

Associations between leaf morphology variation and AFLP polymorphisms

By means of regression analyses significant associations between leaf morphology traits and AFLP markers were found for both species. For all analyzed leaf morphology traits different AFLP markers were found to be associated for *Q. robur* as compared to *Q. petraea*. This observation can be explained by the fact that for each species different markers had been excluded from the analysis. Other explanations are marker band frequency differences, differences in leaf morphology traits, and differences in variation of these leaf morphology traits between the two species. The percentage of explained variation varied between the leaf morphology traits and between the two species. The percentage of explained variation reached levels as high as 59.5% for PV in *Q. petraea*. Differences in levels of explained variation can be pointing to polygenically and monogenically coded traits, respectively. Especially for the leaf morphology traits with lower levels of explained variation it is necessary to study the genetic loci involved in the inheritance of leaf morphology in order to get better models. The here presented regression analyses give a preliminary indication about linkage between leaf morphology traits and AFLP markers. However, a genetic linkage map will reveal the true linkage and the location of the traits on the genome.

In this study it was observed that species could be as well differentiated based on leaf morphology as based on AFLP. However, a clearer separation of species was obtained based on leaf morphology traits. Other studies give the opposite result: AFLP could distinguish between *Eragrostis pilosa* accessions better than morphological traits, that varied too much due to environmental fluctuations (Ayele et al. 1999). Mace et al. (1999) showed that AFLP could distinguish between taxa of the tribe Datureae more clearly than isozyme, morphology, and ITS-1 data. As *Q. robur* and *Q. petraea* are usually recognized based on leaf morphology characteristics and not based on AFLP polymorphisms, it was to be expected that therefore the groups would be better differentiated based on leaf morphology than based on AFLP.

Our study shows that average genetic and phenotypic distances within and between species were consistent. However, for some individuals larger genetic distances might be observed as compared to phenotypic distances. This is the result of the fact

that most phenotypic traits are polygenic, which leads to no clear correlation between genetic and phenotypic distances. This effect will increase when the population consists of unrelated individuals (Burstin & Charcosset 1997).

The fact that so far no diagnostic marker has been found indicates the high relatedness between the two species. Previous studies of controlled crosses and paternity analysis have already shown that to a certain extent the species can exchange their genetic material (Steinhoff 1993; Streiff et al. 1999). Evidence about speciation is hard to obtain and therefore it is not known if the high relatedness between the two species can also be pointing to the fact that speciation is still going on. Although this study does not intend to be a taxonomic study, we have indications that *Q. robur* and *Q. petraea* are probably not pure species and therefore can better be called morphotypes (Baverstock & Moritz 1996) or semispecies (Grant 1971). This point of view on *Q. robur* and *Q. petraea* is shared and discussed by Kleinschmit & Kleinschmit (2000). A greater differentiation between the two species based on leaf morphology variation than based on AFLP polymorphisms can now be explained according to plant speciation theories: due to hybridization the two species can exchange their genetic material, but certain combinations of leaf morphology characteristics are being maintained as these combinations of characteristics are under a continuous selection pressure for adaptation to the niche of the semispecies (Grant 1971). Certain AFLP markers are linked with these leaf morphology characteristics and have therefore been found to be associated with these characteristics in regression studies. From this study it becomes clear that in a Dutch autochthonous mixed *Q. robur* and *Q. petraea* population these two species behave as sympatric morphotypes or semispecies. Still it will be necessary to treat these two groups separately. For conservation purposes it is recommendable to conserve mixed *Q. robur* and *Q. petraea* populations as two sympatric morphotypes or semispecies cover more variation than only one of the two. Moreover, still only little is known about ecological processes within and between both morphotypes or semispecies. In order to enable future adaptations to changing environments it will be necessary to conserve the complete syngameon.

Acknowledgements

We thank Cor Das, Gert Kranenborg, Sielke van der Werf, Connie van Doorn, and Jan Bovenschen for collecting the material and their technical help; Bert van der Werf for statistical support and Ronald van den Berg for comments on the manuscript.

3 The description of clones in an ancient woodland of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. with microsatellites and AFLP

Erica G. Bakker[1,2], Barbara C. Van Dam[1], Herman J. Van Eck[2], and Evert Jacobsen[2]

Summary

In densely populated areas autochthonous *Quercus robur* L. (pedunculate oak) and *Q. petraea* (Matt.) Liebl. (sessile oak) (Fagaceae) populations have been maintained as ancient devastated woodlands. The continuous cutting, grazing, and resprouting of such woodlands enabled the development of clonal structures. For conservation purposes an analysis of the actual number, size, and spatial distribution of clones is necessary, especially when there is an interest in the genetic variation of the population. This study describes for the first time – based on microsatellite and AFLP™ analysis – clones in an autochthonous mixed *Q. robur* and *Q. petraea* population that has been coppiced and grazed for several centuries. Based on six microsatellite loci and 69 polymorphic AFLP markers only 14 unique genotypes were detected in a plot that consisted of 80 trees. Clones were observed for both *Q. robur* and *Q. petraea*. The largest clone diameters were observed for *Q. robur* with distances up to 5.8 m. The observed clone sizes may indicate the old age of the trees.

Key words: AFLP; clones; leaf morphology; microsatellites; *Quercus petraea*; *Quercus robur*

Introduction

In the densely populated area of the Netherlands indigenous *Quercus robur* L. (pedunculate oak) and *Q. petraea* (Matt.) Liebl. (sessile oak) populations almost disappeared due to human action from as early as the beginning of the middle ages onwards (Van Iterson 1932). Some indigenous pedunculate and sessile oaks may have survived, most probably as devastated ancient woodland. Such woodland was used for coppice wood and sheep grazing by local farmers. These ancient woodlands differ from unmanaged natural populations because cutting and grazing can induce stem suckers that will result in a clonal structure. Besides this, for tree species fire or uprooting can induce clonal structures (Koop 1987). However, little is known about the number, size,

[1] *Department of Ecology and Environment, Alterra, P.O. Box 47, 6700 AA, Wageningen, The Netherlands; and*
[2] *Laboratory of Plant Breeding, Wageningen University, P.O. Box 386, 6700 AJ, Wageningen, The Netherlands*

and spatial distribution of clones of *Q. robur* and *Q. petraea*. Observations in West Wood (South Essex) report expected clonal structures of *Q. robur* and *Q. petraea* with diameters up to 2.5 m (Rackham 1980). So far, all observations were based on visual connections between the trees. However, for several reasons we expect that these measurements do not reveal the true size of the clones. First, as the coppicing and grazing could have lasted for many centuries, the clones are probably very large and the sprouted stems may have developed to such an extent that they might look like separate trees. When the connecting branches are decomposed or covered with sand dunes, sprouts are no longer visibly connected. Second, the genetic identity of two trees cannot be demonstrated easily by digging up the root system. Parts of the formerly connecting root system may have died. Moreover, natural grafting between roots of neighboring oak trees may occur (Graham & Bormann 1966).

Isozymes have been used successfully to analyze the spatial structure of clones in plant populations (Ellstrand & Roose 1987). For species of the genus *Quercus* (*Quercus chrysolepis*, *Q. laevis*, and *Q. margaretta*) isozyme studies of natural populations detected clones (Berg & Hamrick 1994, 1995; Montalvo et al. 1997). However, for *Q. robur* and *Q. petraea* molecular marker studies (isozymes and microsatellites) of a natural population have not detected any clones so far (Bacilieri et al. 1994; Streiff et al. 1998). As these studies describe an unmanaged natural population, clones are not expected to occur at a high frequency.

Although isozymes can be used for the detection of clones in a population, the relatively small number of low variable loci that can be studied may result in false positives. For this reason the highly variable microsatellite technique and the multi-locus AFLP technique (Vos et al. 1995) are more suitable techniques to recognize clones more precisely. These techniques have been used to successfully detect clones in natural populations of black poplar (*Populus nigra* L.) and eelgrass (*Zostera marina* L.) (Arens et al. 1998; Reusch et al. 1999a,b).

The objective of this study is to get an indication about the number, size, and spatial distribution of clones in ancient woodlands that have been under human influence for many centuries. An ancient mixed *Q. robur* and *Q. petraea* woodland is selected as a representative for such ancient woodlands in the Netherlands. Microsatellites and AFLP are evaluated for their ability to detect clones.

Materials and methods

Study area

The selected mixed *Q. robur* and *Q. petraea* population, "De Stompert", located in the center of the Netherlands (state survey co-ordinates 148.7/460.7), is assumed to be indigenous (Maes 1993; Van Dam & De Vries 1998). Centuries of cutting (coppice)

and overgrazing turned the oak population into a devastated ancient woodland. One circular plot (0.3 ha) was chosen irrespective of the species composition. The studied plot was coppiced for the last time in the first half of the 20th century. *Q. robur* and *Q. petraea* were distributed throughout the area; they occurred on brown podzolic soils. All 80 trees in the plot were sampled. The positions of the sampled trees were recorded. The map was subsequently digitized with ArcView® GIS (Environmental Systems Research Institute Inc., Redlands, USA).

Leaf morphology analysis

Five fully expanded leaves were sampled from different sides from the crown of each tree. The leaves were evaluated for their level of abaxial hairiness (HR) on a scale from 1 (no hairs at all) to 6 (densely hairy), and subsequently dried and stored. After drying, another eight leaf morphology characteristics were scored: lamina length (LL); petiole length (PL); lobe width (LW); sinus width (SW); length of lamina from the lamina base to the widest part (WP); number of lobes (NL); number of intercalary veins (NV); basal shape of the lamina (BS; measured on a scale from 1 (wedge-shaped) to 9 (two clear lobes)) (Rushton 1978).

A principal component analysis (PCA) was performed based on the standardized variables in order to differentiate between the two species, as well as to recognize the clones, using the Genstat 5 package (Release 4.1; Lawes Agricultural Trust, Rothamsted, UK).

DNA-extraction

DNA was extracted from fresh leaves or buds with a DNA extraction kit (Puregene®, Gentra Systems, Minneapolis, USA) including 4% PVP-40 to remove phenolic compounds. PCR amplification, electrophoresis on 6% standard denaturing polyacrylamide gels, and silverstaining of the gels were done according to Streiff et al. (1998).

Microsatellite analysis

All trees were genotyped using six microsatellite loci. Four microsatellite loci (ssrQpZAG104, ssrQpZAG9, ssrQpZAG1/5, ssrQpZAG36), that were developed for *Q. petraea*, amplify at $(AG)_n$ dinucleotide repeats and showed Mendelian inheritance in controlled crosses of *Q. robur* (Steinkellner et al. 1997). Two other microsatellite loci (MSQ4 and MSQ13) were developed for *Q. macrocarpa* (belongs to the section *Lepidobalanus*, which contains also *Q. robur* and *Q. petraea*) and amplify at $(AG)_n$ and $(TC)_n$ dinucleotide repeats, respectively. Both loci showed Mendelian inheritance as

was observed after maternity testing (Dow et al. 1995).

AFLP analysis

AFLP analysis was performed according to Vos et al. (1995). Primers including one selective nucleotide (Eco+A and Mse+C) were used for pre-amplification of the template. The selective amplification was performed with γ^{33}P-ATP labeled primers Eco+AAG or Eco+ATA in combination with the unlabeled primer Mse+CCC. After electrophoresis the gels were vacuum dried on 3 mm Whatman paper and subsequently exposed to X-ray films (Kodak). Reproducibility of AFLP was tested on five individuals (two *Q. robur* and three *Q. petraea*). For each of the five individuals AFLP fingerprints obtained from DNA samples from two different tissues (buds and leaves) from different years (1996 and 1997) were compared.

Description of clones

After microsatellite analysis, individuals were ordered into groups sharing an identical genotype over all six microsatellite loci. Identical genotypes, as characterized based on six microsatellite loci, were identified as clones when the probability for a second encounter of the same multi-locus genotype, p_{se}, was lower than 0.05 (Parks & Werth 1993). The probability p_{se} was calculated according to Parks & Werth (1993) based on the probability of occurrence of a diploid genotype in the population, p_{gen}, under the assumptions of absence of mutation, random mating, and absence of linkage disequilibrium. p_{gen} was calculated after Parks & Werth (1993) as

$$p_{gen} = \prod_{i=1}^{L} g_i \qquad (3.1)$$

where g_i is the Hardy-Weinberg expected frequency of the genotype at locus i and L is the number of loci. Allele frequencies were calculated for each locus based on the number of unique genotypes that could be detected after exclusion of that particular locus from the data set (sub-sampling approach) according to Parks & Werth (1993). The probability for a second encounter of the same multilocus genotype among G genotypes was calculated after Parks & Werth (1993) as

$$p_{se} = 1 - \left(1 - p_{gen}\right)^{G} \qquad (3.2)$$

The AFLP fingerprints were scored for the presence (1) or absence (0) of bands. The number of polymorphisms between pairs of trees were analyzed for each species separately based on simple matching of the AFLP fingerprints using Treecon (Van de Peer et al. 1994). Average Jaccard's (1908) genetic distances were calculated for pairs of

trees of each species using program Phyltools (J. Buntjer, Laboratory of Plant Breeding, Wageningen University, Wageningen, The Netherlands) as

$$GD_{xy} = \frac{N_{xv}}{N_{xy} + N_x + N_y} \qquad (3.3)$$

where N_x is the number of bands in individual x and not in individual y. N_y is the number of bands in individual y and not in individual x, and N_{xy} is the number of bands shared in individuals x and y.

Results

Leaf morphology analysis

Principal component analysis (PCA) based on all nine leaf morphology characteristics resulted in a separation of the two species (Fig. 3.1). The first axis was responsible for the separation of the two species. All nine variables showed equally high loadings. The highest loadings were observed for: SW, NV, and BS. BS and NV showed a positive loading, whereas the other seven morphological traits showed negative loadings. This observation agrees with Rushton's (1978) species separation. Among 80 trees 59 *Q. robur* and 21 *Q. petraea* trees were identified. PCA did not reveal a clear clustering of genotypes within either of the two species (Fig. 3.1).

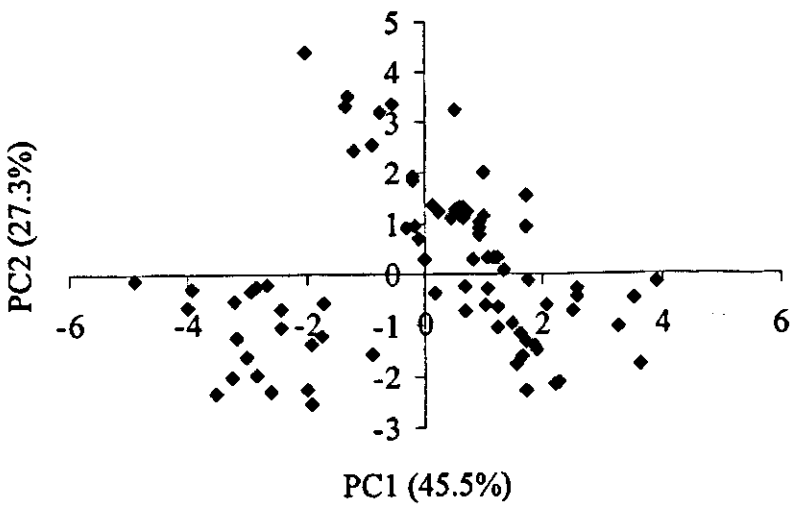


Figure 3.1

Principal component analysis of *Q. robur* and *Q. petraea* based on nine leaf morphology characteristics.

Microsatellite analysis

For the group of 59 *Q. robur* trees the number of different alleles per microsatellite locus varied between five and 11. For the group of 21 *Q. petraea* trees this number varied between six and eight. Allele frequencies calculated based on the sub-sampling approach (calculation of allele frequencies for each locus based on the number of unique genotypes that could be detected after exclusion of that particular locus from the data set; Parks and Werth 1993) were equal to the original allele frequencies for both the *Q. robur* group and the *Q. petraea* group. Therefore, the p_{se} values could be calculated based on the original allele frequencies. The chance of a second occurrence of the same genotype was very low for both species: 3.3E-06 and 1.0E-06 for *Q. robur* and *Q. petraea*, respectively. Therefore, all identical genotypes were assumed to be part of a single clone. In total 14 different genotypes were identified: nine *Q. robur* genotypes and five *Q. petraea* genotypes.

AFLP analysis

In the two AFLP fingerprints generated with two different primer combinations we detected 69 polymorphic markers out of a total of 93 markers (74% polymorphism). Separate analysis of *Q. robur* and *Q. petraea* resulted in 58 and 38 polymorphic markers corresponding with 62% and 41% polymorphism, respectively. As the reproducibility test showed complete repeatability over different tissues and years of the AFLP fingerprints for all five tested individuals, identical banding patterns could safely be regarded as compelling evidence for identical genotypes. Still, there is a chance of the occurrence of co-segregating bands. However, it was expected that this chance was reduced by generating fingerprints consisting of a low density of bands.

AFLP analysis revealed in total 14 groups that consisted of highly related individuals. These 14 groups corresponded with the nine *Q. robur* and five *Q. petraea* clones that were identified based on microsatellite analysis. However, some of the individuals within these groups showed additional AFLP bands. This is illustrated by the bimodal frequency distribution of the number of polymorphisms between pairs of trees: within and between groups (Fig. 3.2). The number of polymorphic AFLP bands between two trees from the same group varied between zero and three for *Q. robur* and between zero and six for *Q. petraea*. Between two trees from different groups AFLP fingerprint differences varied between 11 and 30 bands for *Q. robur* and between 14 and 23 bands for *Q. petraea* (Fig. 3.2). The average Jaccard's (1908) genetic distances within groups varied between 0 and 0.039 for *Q. robur* and between 0 and 0.063 for *Q. petraea*. The average Jaccard's (1908) genetic distances between groups amounted 0.36 ± 0.07 (SD) for *Q. robur* and 0.33 ± 0.05 (SD) for *Q. petraea*.

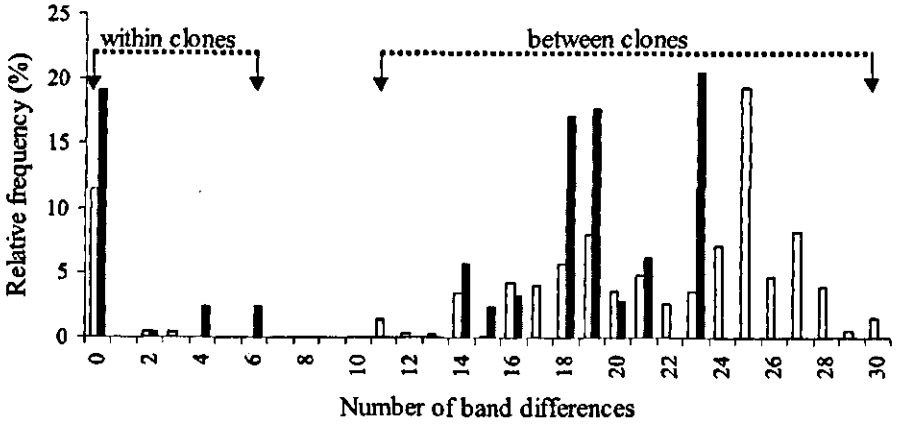


Figure 3.2

Distribution of the number of band differences in AFLP fingerprints between pairs of trees within and between clones represented as relative frequencies for *Q. robur* (white bars) and *Q. petraea* (black bars).

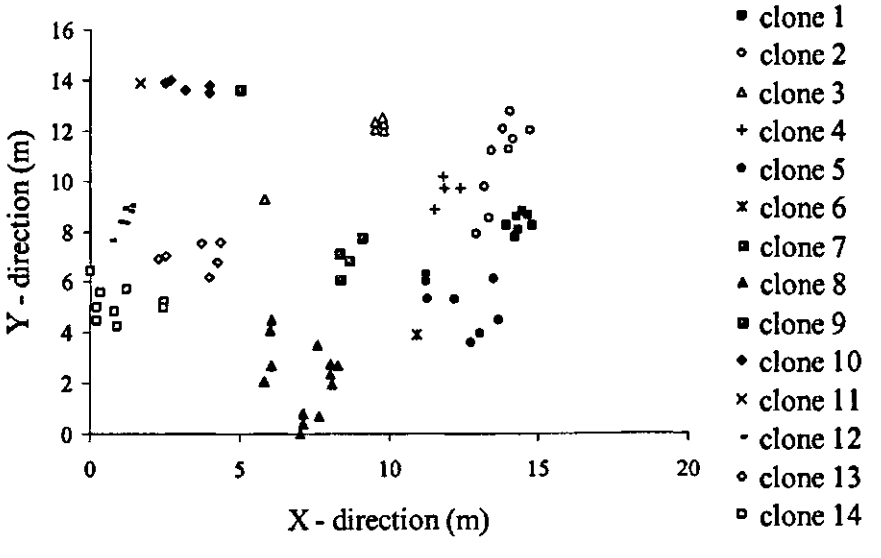


Figure 3.3

Map of the sampled *Q. robur* and *Q. petraea* trees in "De Stompert". Each clone is represented with a different symbol. Clones 2,3,4,5,6,7,8,10, and 14 are *Q. robur*. Clones 1,9,11,12, and 13 are *Q. petraea*.

Description of clones

Among the 59 *Q. robur* trees were nine genotypes of which eight were clones. Among the 21 *Q. petraea* trees were five genotypes of which three were clones. The average number of trees within a clone was 7 ± 3 trees (SD) for *Q. robur* and 6 ± 0.6 trees (SD) for *Q. petraea*. The highest number of trees per clone was 13 trees for *Q. robur*. The average distance between trees within a clone was 1.5 ± 0.5 m (SD) for *Q. robur* and 0.9 ± 0.4 m (SD) for *Q. petraea*. The average maximum distance between a pair of trees within a clone was 3.6 ± 1.9 m (SD) for *Q. robur* and 2.0 ± 0.7 m (SD) for *Q. petraea*. The maximum distance between two trees that belong to one clone was 5.8 m for *Q. robur*. The clones were growing in aggregations, but were irregularly shaped and were growing intermingled (Fig. 3.3).

Discussion

We were able to describe the number, size, and spatial distribution of clones of *Q. robur* and *Q. petraea* in an autochthonous population by means of microsatellite and AFLP analysis. Among 80 sampled trees only 14 unique genotypes were detected. Clones were observed for both *Q. robur* and *Q. petraea*.

Q. robur and *Q. petraea* could be distinguished on the basis of the nine leaf morphology characteristics in a principal component analysis (PCA). A clear-cut separation between the clones was not obtained, but individuals that were found to be part of a clone based on microsatellite and AFLP results clustered to some extent. As these clusters overlapped each other, it was not possible to detect clones based on leaf morphology analysis. Apparently the variation of leaf morphology characteristics among trees constituting a clone was substantial as compared to the variation among clones. Due to environmental heterogeneity (variability in light, microclimate, or soil conditions encountered by different plant modules) the leaf morphology characteristics of the different trees that are part of a clone can vary (Whitham 1981). Besides, there is a low chance of somatic mutations in meristematic tissues, which can cause a variation in morphological characteristics (Antolin & Strobeck 1985).

Based on six microsatellite loci 14 different genotypes (nine *Q. robur* and five *Q. petraea*) could be identified as the chance for a second occurrence of a genotype among the studied group was much lower than 0.05 (Parks & Werth 1993). The groups of highly related individuals that were detected by means of AFLP analysis corresponded with the clones observed with microsatellite analysis. However, up to six additional bands have been observed between trees from the same group. These few polymorphisms between some trees within a clone were no arbitrary amplification products, but could be observed in other oak genotypes. Moreover, these polymorphic AFLP bands within groups could be reproduced. This does not mean that these AFLP

bands are no artifacts as Jones et al. (1997) observed AFLP artifacts that could be reproduced. The observed within group variation corresponded with a similarity ranging between 95% and 100%, which is within the bounds of scoring errors for AFLPs (Huys et al. 1996; Arens et al. 1998; Winfield et al. 1998). Alternative explications for the observed within group variation could be the presence of highly similar full-sibs or somatic mutations in some of the clones. It is not likely that trees differing for up to six bands separated by distances of up to 5.8 m are all full-sibs as there is evidence that *Q. robur* and *Q. petraea* only exhibit weak spatial structures of diversity (Streiff et al. 1998). It is possible that some of the observed polymorphisms are the result of somatic mutations, as it has been described that AFLP is able to detect somatic mutations (Chen et al. 1999).

Although the mutation rate of microsatellites is high (Kashi et al. 1997), this marker technique can detect clonal structures of more than 1,000 years old (Reusch et al. 1999a). This was confirmed by our study where nine different *Q. robur* and five different *Q. petraea* clones could be identified based on six microsatellite loci. The fact that AFLP detected a low level of polymorphism within these clones shows that AFLP is a less suitable technique for detecting clones as it is more susceptible to artifacts, although the possibility of somatic mutations cannot be ruled out (Huys et al. 1996; Jones et al. 1997; Arens et al. 1998; Winfield et al. 1998; Chen et al. 1999).

Clones have not been observed in the studies of Bacilieri et al. (1994) and Streiff et al. (1998). Both studies dealt with a wood that originated from natural regeneration. Although no clones were detected, there could be clones present due to uprooting or fire (Koop 1987). Apparently, clones that originate due to natural causes occur at a low frequency. As "De Stompert" has a history of coppicing and grazing, we suppose that the clones were majorily formed because of these two factors. We propose that identical genotypes are sprouts from the same trunk as apomixis has not been found in oak (Asker and Jerling 1992). The observed maximum clone diameters for *Q. robur* and *Q. petraea* in this study were larger than those observed by Rackham (1980) and Koop (1987). Based on the large size of the clones, it is expected that the trees are of a considerable age.

As the clones in "De Stompert" were growing intermingled, visual detection of clones cannot be advised. Neither can a leaf morphology analysis be advised. The best method for the detection of the full spatial structure of clones in a plot is genotyping all trees with molecular markers, of which the microsatellite technique is most suitable for this goal.

Our results have implications for the analysis of genetic variation of ancient woodlands. The presence of clones should be taken into account in order to get a correct estimation of the level of genetic variation in an ancient woodland. Therefore, it is advisable to first analyze ancient woodlands for the presence of clones before the genetic variation will be studied. Knowledge about the prevalence and size of clones can

be used to make proper decisions about thinning, rejuvenation, and conservation.

Acknowledgements

The authors thank Cor Das, Gert Kranenborg, and Triye Huibers for technical assistance; Marie-Claire Boerwinkel and Jan Bovenschen for help in the laboratory; and Nelleke Kreike, Henk Koop, Rienk-Jan Bijlsma, and Ad van Hees for fruitful discussions and useful comments on the manuscript.

4 Genetic variation and structure of indigenous *Quercus robur* L. populations

Erica G. Bakker[1,2], Barbara C. Van Dam[1], Fred A. Van Eeuwijk[2], Evert Jacobsen[2]

Summary

In the Netherlands indigenous *Q. robur* L. populations are rare and have been maintained as patches ancient woodland. For a proper conservation of these populations, information about genetic variation and population structure is necessary. In order to assess the genetic variation and the structure of these populations, microsatellite polymorphisms were studied in two of these autochthonous populations. The two autochthonous populations differed slightly for their gene diversity, which was as high as was observed for *Q. robur* populations in France and Germany. No clear indications for a population structure were found. Subsequently, the gene diversity of a half-sib family obtained from a tree from one of the two studied autochthonous populations was compared with the gene diversity of a half-sib family obtained from a tree located at a roadside in urban area. The gene diversity of the two half-sib families was similar, but was relatively low. The implications of these results for the management and conservation of autochthonous *Q. robur* populations are discussed.

Keywords: *Quercus robur*, conservation, gene diversity, population structure, half-sib family

Introduction

So far, history, structure, and quality of Dutch indigenous oak populations have been investigated based on old documents and indicator species (Maes 1993). An indigenous – also called autochthonous – population has occurred in an area since its establishment after the last ice age. In this way a few ancient woodlands in the Netherlands were indicated as autochthonous *Quercus robur* L. (pedunculate oak) and *Q. petraea* (Matt.) Liebl. (sessile oak) populations (Maes 1993). However, additional information is necessary to state with certainty that those populations are truly autochthonous and to decide about the right conservation management plan. First, it is necessary to find out if the populations have naturally occurred in those areas since

[1] *Department of Ecology and Environment, Alterra, P.O. Box 47, 6700 AA, Wageningen, The Netherlands*
[2] *Laboratory of Plant Breeding, Wageningen University, P.O. Box 386, 6700 AJ, Wageningen, The Netherlands*

the last ice age. Pollen records in the soil indicate the presence of both oak species in the Netherlands dating back from 8,000 B.P. onwards (Huntley & Birks 1983). However, this is not a direct evidence for the autochthonicity of a particular population. Recently, chloroplast DNA (cpDNA) analyses revealed the re-colonization routes of *Q. robur* and *Q. petraea* since the last ice age (Petit et al. 1993; Dumolin-Lapègue et al. 1997; Petit et al. 1997). Based on cpDNA analysis of a number of Dutch oak populations that were earlier characterized as autochthonous according to Maes' (1993) criteria, indications were obtained about re-colonization routes of oaks in the Netherlands (Konig et al. *in press*). The populations in the western part of the Netherlands have migrated from Spain and the populations in the eastern part from Italy (Van Dam & De Vries 1998). Based on this information it is possible to find out about plantation from the Balkans or other remote areas (Dumolin-Lapègue et al. 1997; Konig et al. *in press*). Plantation with material of the same cpDNA haplotype cannot be detected, but the presence of recent plantations is less likely when the population consists of old trees. Nothing is known about the age of the trees in Dutch ancient woodlands, as they consist of coppices. Because of the continuous sprouting of the coppices the main trunk has disappeared and therefore year ring counting is not possible. Microsatellite and AFLP analyses have shown that these coppices are of a considerable size, which is an indication for the old age of these trees (chapter 3 of this thesis). An indication for natural rejuvenation is the existence of a family structure. However, for tree species like *Q. robur* where extensive pollen flow can act as a homogenizing factor, only a weak population structure can be expected. Such a weak population structure was found in an autochthonous mixed *Q. robur* and *Q. petraea* population in North France using allozymes and microsatellites. A stronger spatial autocorrelation was observed for *Q. petraea* for the first distance class (0-20 m) (Bacilieri et al. 1994; Streiff et al. 1998).

Indigenous oak populations in the Netherlands are rare. Due to large scale deforestation at the end of the 19th century only a few patches of woodland remained surrounded by heather and drift-sand. Because of the small number of trees in these relict populations inbreeding could have occurred and genetic variation could have been reduced. Moreover, most of these indigenous oak populations have been managed. The continuous coppicing and grazing might have resulted in a selection for trees that are able to sprout, which could have led to an additional reduction in genetic variation. An indication for the quality of a population is the genetic variation. When there is a high genetic variation in the population, then the next generations are likely to be highly variable when mating is random. A high genetic variation in next generations enables adaptation of the population to changing environmental conditions. Therefore conservation management aims to maintain a high genetic variation within indigenous populations. The genetic variation can be investigated by means of several molecular marker techniques. For *Q. robur* and *Q. petraea* studies have

been done based on allozymes (Zanetto et al. 1994; Samuel et al. 1995; Zanetto & Kremer 1995; Streiff et al. 1998; Degen et al. 1999), cpDNA polymorphisms (Petit et al. 1993), RAPD (Moreau et al. 1994; Bodénès et al. 1997a) and SCAR (Bodénès et al. 1997a,b). All these studies indicated that in the investigated populations the two species exhibited a high genetic variation. A high genetic variation was expected as the studied populations have neither been reduced in size due to large scale deforestation, nor have they been managed. Besides, *Q. robur* and *Q. petraea* are both obligatory outcrossing, wind pollinated, long-lived tree species, which are life history traits that have a positive effect on the genetic variation of a species (Hamrick et al. 1979; Hamrick & Godt 1996). Recently, microsatellites have been used for studies of genetic variation and population structure because of their high mutation frequency, which makes them very suitable to detect very recent population processes. Studies by Streiff et al. (1998) and Degen et al. (1999) showed based on microsatellite results a high genetic variation present within populations of *Q. robur* and *Q. petraea* in France and Germany.

There is not only an interest in the genetic variation of Dutch indigenous oak populations, but also in the genetic variation of the offspring of indigenous trees. For reforestation purposes the genetic variation of half-sib families should be high in order to ensure a high genetic variation in the planted populations. Besides, there is an interest in the genetic variation present in half-sib families obtained from trees located at roadsides in urban area, as in the Netherlands roadside plantations are used as a seed source. Therefore in our study we aimed to describe the genetic variation present within two half-sib families: one obtained from a tree located in an autochthonous population and another obtained from a tree located at a roadside in urban area. Depending on the number of fathers contributing to the offspring, the genetic variation of a half-sib family can vary. The level of genetic variation observed for the half-sib families will give an indication about the number of trees from which seeds should be collected for reforestation purposes.

In this study two autochthonous *Q. robur* populations will be studied for their family structure and gene diversity. Subsequently, the results will be compared with a half-sib family from one of these two autochthonous populations and a half-sib family obtained from a tree at a roadside in urban area.

Materials and methods

Populations

Two autochthonous populations and two half-sib families were investigated. The two autochthonous populations were mixed *Q. robur* and *Q. petraea* ancient woodlands that have been continuously coppiced and grazed for centuries: "De Meinweg" (Italian

lineage), located in the south of the Netherlands, and "De Stompert" (Spanish lineage), located in the center of the Netherlands (Van Iterson 1932; Venner 1985; Van Dam & De Vries 1998). Due to the continuous coppicing and grazing the trees in these two populations have sprouted, which led to the formation of clones (chapter 3 of this thesis). Attention was paid only to include unique genotypes in the sample. Species were identified based on their leaf morphology (Rushton 1978). In total 61 trees (48 *Q. robur* and 13 *Q. petraea*) were sampled within a rectangular plot (0.3 ha) in "De Stompert". In "De Meinweg" 81 trees (48 *Q. robur* and 33 *Q. petraea*) were sampled within a rectangular plot (0.4 ha). Because of the low number of *Q. petraea* sampled in "De Stompert" (N=13), it was decided to only study *Q. robur* in both populations (N=48). The first half-sib family was obtained by collecting seeds (N=48) from one *Q. robur* tree in "De Meinweg". The second half-sib family was obtained by collecting seeds (N=384) from one *Q. robur* tree located at a roadside in urban area (Amsterdam).

Microsatellite analysis

DNA was extracted from fresh leaves or buds with a DNA extraction kit (Puregene®, Gentra Systems, Minneapolis, USA) including 4% PVP-40 to remove phenolic compounds.

All trees were genotyped using six microsatellite markers. Four microsatellite markers (ssrQpZAG104, ssrQpZAG9, ssrQpZAG1/5, ssrQpZAG36), that were developed for *Q. petraea*, amplify at (AG)*n* dinucleotide repeats and showed Mendelian inheritance in controlled crosses of *Q. robur* (Steinkellner et al. 1997). Two other microsatellite markers (MSQ4 and MSQ13) were developed for *Q. macrocarpa* (belongs to the section *Lepidobalanus*, which contains also *Q. robur* and *Q. petraea*) and amplify at (AG)*n* and (TC)*n* dinucleotide repeats, respectively. Both markers showed Mendelian inheritance as was observed after maternity testing (Dow et al. 1995). PCR amplification, electrophoresis on 6% standard denaturing polyacrylamide gels, and silverstaining of the gels was done according to Streiff et al. (1998).

Population genetic analyses

Linkage equilibrium and Hardy-Weinberg equilibrium

Various population genetic analyses assume linkage equilibrium and Hardy-Weinberg equilibrium (HWE). Therefore the two indigenous populations were first analyzed for the presence of genotypic linkage disequilibria according to Weir (1996, pp. 127-128) and Raymond & Rousset (1995b). Subsequently, the two autochthonous populations were tested for Hardy-Weinberg proportions of the genotype frequencies. As the sample size was small and the table of genotypes was sparse due to the high number of

alleles per locus the standard goodness-of-fit tests can sometimes lead to false rejection or acceptance of Hardy-Weinberg equilibrium (HWE). Therefore a test for HWE as proposed by Guo & Thompson (1992) was performed based on Levene's (1949) P-value.

F-statistics

Departure from HWE can also be tested by means of the *F*-statistics: F_{is}, F_{it}, and F_{st}. The *F*-statistics quantify the level of inbreeding per locus (F_{is} and F_{it}) and the population differentiation (F_{st}). Results from HWE tests and *F*-statistics were compared.

F_{is} is the inbreeding coefficient and it measures the proportionate reduction in heterozygosity relative to random mating in a subpopulation (Hartl & Clark 1997). F_{is} can reach values up to 1, which corresponds with complete inbreeding. An F_{is} value of zero indicates that the subpopulation is random mating. Besides F_{is} values per subpopulation ("De Meinweg" and "De Stompert"), average F_{is} values over both subpopulations, F_{it}, and F_{st} were calculated according to Weir & Cockerham (1984). F_{it} describes inbreeding for the subpopulations together. F_{st} describes the differentiation between subpopulations. F_{st} can reach values between 0 and 1, corresponding to no differentiation and complete population differentiation, respectively. Significance levels for the *F*-statistics were obtained according to Rousset & Raymond (1995).

Gene diversity

The two autochthonous populations and the two half-sib families were analyzed for their gene diversity, which is a good measure of genetic variation in the population (Nei & Roychoudhury 1974). The gene diversity is equal to the chance that two genes sampled at random from a population are not identical. If the population is in HWE the gene diversity estimates the heterozygosity of the population. The gene diversity gives an indication for the genetic variation present in the population and for the level of variation in next generations when mating is random. Unbiased estimates of the gene diversity, $\hat{h}_k$ were calculated according to Nei & Roychoudhury (1974) for each locus k separately as

$$\hat{h}_k = \left\{n\left(1-\sum x_i^2\right)-1\right\}/(n-1) \qquad (4.1)$$

where n is the number of genes ($n/2$ individuals in diploid organisms) and x_i is the sample gene frequency of the i-th allele. In order to obtain the standard errors (SE) for the gene diversities, the intra-locus variance per locus, $V_s(\hat{h}_k)$, was calculated according to Nei and Roychoudhury (1974) as

$$V_s(\hat{h}_k) = \frac{2(n-1)}{n^3}\left\{(3-2n)\left(\sum x_i^2\right)^2 + 2(n-2)\sum x_i^3 + \sum x_i^2\right\} \qquad (4.2)$$

When a population is not in HWE an indication about the actual heterozygosity can be obtained from the observed heterozygosity ($\hat{h}_{obs}$). The observed heterozygosity was estimated for the two autochthonous populations by counting the proportion of heterozygotes per locus directly. The sampling variance of this estimate was calculated according to Nei & Roychoudhury (1974) as $\hat{h}_{obs}(1 - \hat{h}_{obs})/N$, where N is the number of individuals examined.

Another way to describe the genetic variation of a population is the effective number of alleles, $A_e = 1/(1 - \hat{h}_k)$. This is the number of equally frequent alleles that would be required to produce the same homozygosity as was found in an actual population (Hartl & Clark 1997 pp. 176-177). The effective number of alleles was calculated in order to be able to compare the values found for the Dutch populations with another study that calculated A_e values but did not calculate $\hat{h}_k$ values (Degen et al. 1999).

Spatial structure

The spatial structure of the two autochthonous populations was analyzed based on the spatial coancestry, ρ_{ij}, between all possible pairs of individuals i and j at certain distance intervals. This statistic measures the correlation between the frequencies of homologous alleles, p_i and p_j, at a locus of two individuals i and j that are separated by a certain geographical distance (Loiselle et al. 1995) as

$$\hat{\rho}_{ij} = \frac{\sum_{ij}(p_i - \bar{p})(p_j - \bar{p})}{k\bar{p}(1-\bar{p})} + \frac{2}{(8k+1)^{0.5} - 1} \qquad (4.3)$$

where $k = n(n-1)/2$ is the total number of possible pairwise connections between n individuals located at a discrete number of map units away from each other. Significance values for ρ_{ij} were obtained after 1000 random permutations of the individuals.

The method of spatial coancestry tests for a significant clustering of related individuals. For *Q. robur* populations this means that it is possible that this method will detect only a weak spatial structure or no structure at all. As natural rejuvenation of *Q. robur* does not take place under a closed canopy, this means that next generations only develop at places where there is a sufficient level of light (Oosterbaan & Van Hees 1989). This fact prevents the formation of clusters of siblings around maternal trees. Therefore next to the method of spatial coancestry another method was used to describe the spatial structure of the two autochthonous populations. This method

involved the identification of family relationships (parent-sibling, half-sib, and full-sib) between trees by means of an analysis of pairs of individuals for at least one common allele per locus. In this way we aimed to describe identity by descent. These family relationships were tested for cryptic gene flow, the probability that two individuals that share common alleles are not related (identical by state; Westneat & Webster 1994; Hartl & Clark 1997 pp. 141-142). Based on the observed frequencies of these common alleles within a population, the expected frequency of genotypes carrying a common allele for all studied microsatellite loci was calculated according to Westneat & Webster (1994) under the assumption of HWE and linkage equilibrium as

$$x = \prod_{i=1}^{L} \left(2p_i q_i + q_i^2\right) \tag{4.4}$$

where q is the observed frequency in the population for a common allele between two individuals in the studied population at locus i ($p = 1 - q$) and L is the number of loci. The probability that at least two individuals out of the population carried by chance a common allele for all studied loci was calculated according to Westneat & Webster (1994) as

$$p_m = 1 - (1 - x)^n \tag{4.5}$$

where n is the population size. A family relationship between two individuals was assigned to cryptic gene flow when p_m was higher than 0.05 (Westneat & Webster 1994; Dow & Ashley 1996).

The family structure and cryptic gene flow were subsequently represented by drawing connections between trees that were identified to share family relationships using ArcView® GIS (Environmental Systems Research Institute Inc., Redlands, USA).

Computer programs

The population genetic analyses were done with Genepop 3.1c (Raymond & Rousset 1995a). Spatial coancestry, ρ_{ij}, was calculated using program AutocorG 2.1 (O. Hardy and X. Vekemans, Laboratoire de Génétique et d'Ecologie Végétales, Université Libre de Bruxelles, Bruxelles, Belgium).

Results

Linkage equilibrium and Hardy-Weinberg equilibrium

For both autochthonous populations no significant genotypic linkage disequilibrium (α=0.05) was observed (data not shown). Hardy-Weinberg equilibrium (HWE) was observed for all six microsatellite loci in population "De Meinweg" and for all loci

except locus AG36 in population "De Stompert".

F-statistics

The significantly ($p<0.05$) F_{is} value observed for AG36 for "De Stompert" confirmed the departure from HWE that was observed for this locus. For "De Meinweg" for three loci significant ($p<0.05$) F_{is} values have been observed, whereas there was no significant departure from HWE (Tab. 4.1). F_{it} values and average F_{is} values were close to zero for all six microsatellite loci except for MSQ4. Differentiation (F_{st}) among the allele frequencies in the two autochthonous populations was close to zero for all loci, except for MSQ4 ($p<0.005$) and AG1/5 ($p<0.05$). Negative estimates were interpreted as the absence of inbreeding and population differentiation for F_{is}, F_{it}, and F_{st}, respectively.

Gene diversity

Estimates for the gene diversity are represented in Tab. 4.2. The average gene diversity of the two autochthonous populations differed significantly ($p<0.05$), but the difference in average gene diversity was small as compared to the difference in average gene diversity between the autochthonous populations and the half-sib families. The half-sib family from "De Meinweg" and the half-sib family from Amsterdam differed significantly ($p<0.05$) for their average gene diversity, but the difference in average gene diversity was small as compared to the difference in average gene diversity between the autochthonous populations and the half-sib families. All six loci showed a significantly ($p<0.05$) lower gene diversity for the half-sib families as compared to the autochthonous populations. The average observed heterozygosity was similar to the average gene diversity and was the same for both autochthonous populations (Tab. 4.2).

Table 4.1

F_{is}, F_{it}, and F_{st} values for the two autochthonous populations, "De Meinweg" (Pop 1) and "De Stompert" (Pop 2), according to Weir & Cockerham (1984).

LOCUS	F_{is} POP 1	F_{is} POP 2	AVERAGE F_{is}	F_{is}	F_{st}
AG1/5	+0.027	-0.081	-0.025	-0.010	0.015*
AG9	-0.018	+0.049	0.016	0.016	-0.001
AG36	-0.062	+0.189*	0.060	0.059	-0.002
AG104	+0.059*	-0.035	0.012	0.013	0.001
MSQ4	+0.178*	+0.019	0.100	0.118	0.020**
MSQ13	+0.087*	-0.079	0.004	0.008	0.004

*P<0.05; **P<0.001. P values estimated for F_{is} values within each of the two populations and for F_{it} values between the two populations with Genepop 3.0 (Rousset and Raymond 1995a,b).

Table 4.2

Unbiased estimates of the gene diversity ($\hat{h}_k$) per locus, according to Nei & Roychoudhury (1974) for two autochthonous populations, "De Meinweg" (Pop 1) and "De Stompert" (Pop 2), a half-sib population from "De Meinweg" (HS 1) and a half-sib population from Amsterdam (HS 2).

LOCUS	Pop 1		Pop 2		HS 1	HS 2
$\hat{h}_k$	$\hat{h}_k$	$\hat{h}_{obs}$	$\hat{h}_k$	$\hat{h}_{obs}$	$\hat{h}_k$	$\hat{h}_k$
AG1/5	0.87	0.88	0.81	0.88	0.73	0.79
AG9	0.84a[1]	0.88	0.85a	0.81	0.73	0.76
AG36	0.86	0.92	0.82	0.67	0.76	0.73
AG104	0.93b	0.88	0.93b	0.96	0.76	0.83
MSQ4	0.89	0.73	0.85	0.83	0.72	0.66
MSQ13	0.82c	0.75	0.81c	0.88	0.52	0.60
Average	0.87	0.84	0.85	0.84	0.70	0.73

[1]) Values that do not differ significantly (α=0.05) between populations are indicated with identical letters.

Table 4.3

Effective number of alleles in two autochthonous *Q. robur* populations, "De Meinweg" (Pop 1) and "De Stompert" (Pop 2), a half-sib population from "De Meinweg" (HS 1) and a half-sib population from Amsterdam (HS 2).

LOCUS	POP 1	POP 2	HS 1	HS 2
AG1/5	8.0	5.3	4.8a[1]	3.7a
AG9	6.2b	6.9b	4.2c	3.7c
AG36	7.3	5.6	3.7d	4.2d
AG104	14.2e	13.5e	5.9	4.2
MSQ4	8.7	6.6	2.9f	3.6f
MSQ13	5.6g	5.3g	2.5h	2.1h
Average	8.3	7.2	4.0i	3.6i

[1]) Populations that are not significantly different (α=0.05) have the same letter.

The effective number of alleles, A_e, was significantly lower ($p<0.05$) for the half-sib families as compared to the autochthonous populations. The two autochthonous populations differed significantly ($p<0.05$) for their effective number of alleles for all loci, except for AG104. The two half-sib populations did not differ significantly ($p<0.05$) for their effective number of alleles (Tab. 4.3).

Population structure

Based on spatial coancestry analysis with distance intervals of 10 m no significant (α = 0.05) genetic structure was observed in either of the two autochthonous populations (data not shown). Based on the common alleles approach for "De Stompert" 19 family

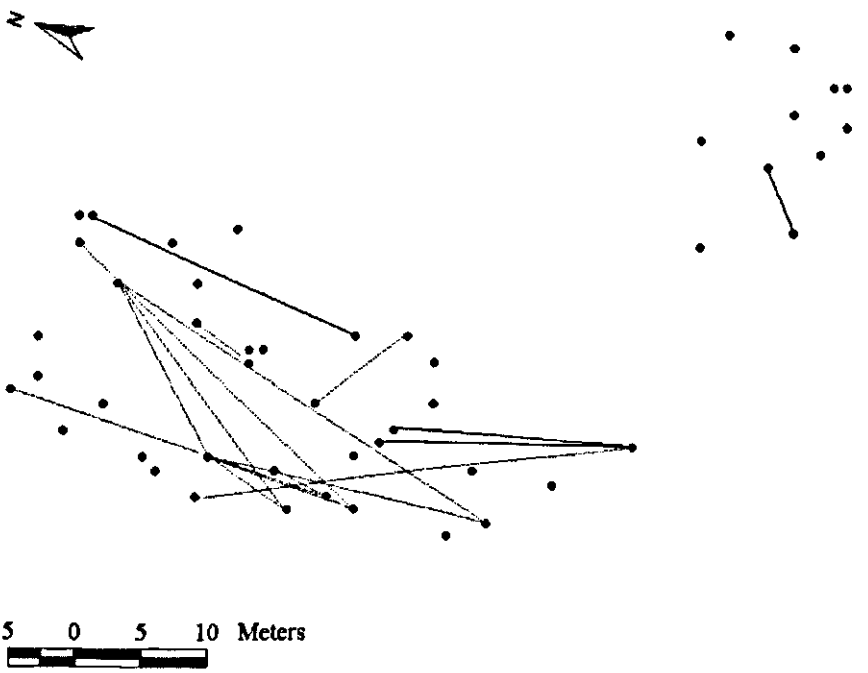


Figure 4.1
Map with the locations of the trees in "De Stompert". Trees that share a family relationship are connected with a line. Family relationships that can be attributed to cryptic gene flow are indicated as dotted lines. In the middle of the plot is a piece of heath where no oaks occurred.

relationships were found with distances ranging from 2.5 m to 33.5 m (Fig. 4.1). Out of these 19 family relationships 15 could be attributed to cryptic gene flow. For "De Meinweg" seven family relationships were found with distances ranging from 9.0 m to 57.5 m (Fig. 4.2). Out of these seven family relationships three could be attributed to cryptic gene flow.

Discussion

Linkage equilibrium

The observed genotypic linkage equilibrium between microsatellite loci confirms the observation of Barreneche et al. (1997) that MSQ4 and MSQ13 are located on different chromosomes and that AG9 and AG1/5 are located on the same chromosome at a distance large enough to ensure independent segregation (37 cM). Barreneche et al. (1997) could not determine the distance between AG36 and AG104 as they were assigned to the maternal and the paternal linkage map, respectively. As we did not detect any linkage disequilibrium between these two loci, it is to be expected that these

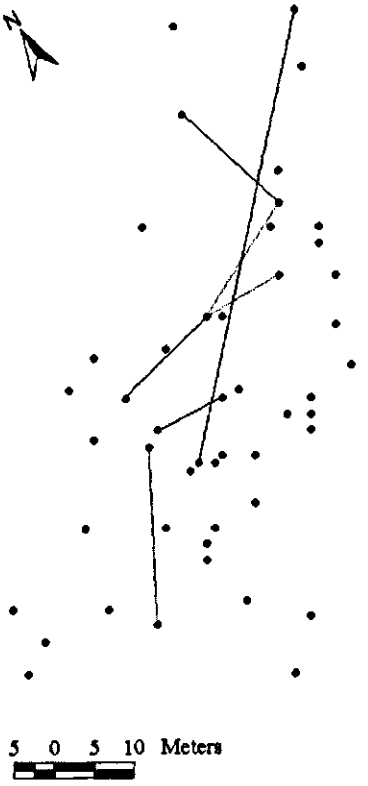


Figure 4.2
Map with the locations of the trees in "De Meinweg". Individuals that share a family relationship are connected with a line. Family relationships that can be attributed to cryptic gene flow are indicated as dotted lines.

two loci are either located at some distance from each other on the same chromosome or at two different chromosomes. Hartl & Clark (1997 p. 105) describe that linkage disequilibrium can be found even when the loci are not closely linked. In such a case linkage disequilibrium indicates an admixture of subpopulations. The absence of linkage disequilibrium indicates, therefore, that the studied plots in "De Meinweg" and "De Stompert" did not consist of two or more subpopulations.

Hardy-Weinberg equilibrium and F-statistics

As linkage equilibrium and HWE (for most of the loci in both autochthonous populations) were observed, it was possible to do all population genetic analyses that assume linkage equilibrium and HWE. Although HWE was observed, it is possible that several assumptions are violated simultaneously in such a way that their effects cancel each other out. Moreover, there are indications that in oak forests mating is not random (Streiff et al. 1999). Therefore a HWE test does not provide much relevant information.

The *F*-statistics did not completely confirm the conclusions based on the HWE test. The significant ($p<0.05$) F_{is} values found for three loci in "De Meinweg" that were

previously found to be in HWE can be attributed to the way of testing. The F_{is} values were tested for heterozygote deficiency. This test was directional as opposed to the two-sided test for HWE. High F_{is} values indicate inbreeding only when all studied loci show equally high F_{is} values. High F_{is} values observed for a part of the studied loci indicate selection for these loci, selection for genes closely linked to these loci (combined with epistatic effects), or null alleles. Null alleles have been reported for MSQ4 in *Quercus macrocarpa* (Dow & Ashley 1996). In another study of *Q. robur* – with the same six microsatellite loci – in a forest in North France similar F_{is} values were found. The highest observed F_{is} value was 0.14 for MSQ4 (Streiff et al. 1998). A study of *Q. robur* in two forests in North Germany based on four microsatellite loci (AG36, AG1/5, AG9, AG104) showed F_{is} values close to zero for all studied loci (Degen et al. 1999).

Remarkably, no inbreeding was detected in the two Dutch autochthonous *Q. robur* populations. Although the two populations were small relict populations this did not lead to inbreeding within these populations. Average F_{is} and F_{it} values over both autochthonous populations were close to zero for all six microsatellite loci values except for MSQ4. This indicates selection or null alleles for MSQ4. Although microsatellites are assumed to be neutral there are indications that natural selection plays a role in the formation of tandem-repetitive non-coding DNA (Stephan and Cho 1994). Kashi et al. (1997) state that the repeat-number variation in simple sequence repeats (SSR) enables a rapid evolutionary adaptation to ecological changes, as the SSR repeats are involved in gene expression and gene function. So far, no such studies have been done for the six microsatellite loci that have been used in our study.

The two autochthonous populations did not differ much for their environmental and climatological characteristics. As the two autochthonous populations have originated from different glacial refugia (Spain and Italy; Van Dam & De Vries 1998; Konig et al. *in press*), effects of selection or null alleles that have occurred in these different refugia might still have their effects even after several generations of intermating between the two lineages. Based on the F_{is} values it is not possible to tell on what trait selection is acting. In order to find out about this a genetic linkage map needs to be constructed. Selective forces on quantitative trait loci (QTLs) located close to these microsatellite loci can then be investigated.

Population differentiation between the two autochthonous populations was very low, which indicates a high number of migrants. As the two populations were separated at a distance of more than 100 km this observation is remarkable. A possible explanation could be long distance pollen flow as has been suggested based on the small pollen grain diameter (reviewed by Ducousso et al. 1993) and paternity analysis (Streiff et al. 1999). For AG1/5 and MSQ4 a significant population differentiation was found. This is an indication for either selection or null alleles for these loci.

Genetic variation in autochthonous populations

The gene diversity observed for the two Dutch autochthonous *Q. robur* populations was as high as the gene diversity observed for a French indigenous *Q. robur* population based on the same microsatellite loci (Streiff et al. 1998). Besides, the effective number of alleles found for the two Dutch autochthonous *Q. robur* populations was as high as was found for two German *Q. robur* populations (Degen et al. 1999). These results indicate that the genetic variation of Dutch autochthonous *Q. robur* populations is as high as the genetic variation of *Q. robur* populations in neighboring countries. As the French and German populations did not undergo coppicing, grazing, and reduction in size due to deforestation, our results indicate that these factors did not reduce the genetic variation of the indigenous *Q. robur* populations.

The gene diversity in "De Meinweg" differed only very little from the gene diversity in "De Stompert". Apparently, differences between the two populations due to their different location or different migration lineage had no impact on the genetic variation within the two populations. Values for the observed heterozygosity were high and similar to the values for the gene diversity. This was expected, as when there is HWE, the observed heterozygosity and the gene diversity (= expected heterozygosity) are identical. High values for the observed heterozygosity indicate that the populations consisted of large numbers of heterozygous individuals.

Genetic variation in half-sib families

The half-sib family from "De Meinweg" and the half-sib family from Amsterdam differed significantly for their gene diversity, but this difference was much smaller as compared to the difference in gene diversity between the half-sibs and the autochthonous populations. The significantly lower ($p<0.05$) genetic variation observed for the half-sib families as compared to the autochthonous populations was expected as half-sib families consist for 50% of one or two alleles inherited from a mother tree. The difference in gene diversity between the two half-sib families was not because of the different sample size, as no significant difference in gene diversity ($\alpha=0.05$) was observed between the full sample (N=384) and a smaller sample (N=48) from the same half-sib family (E.G. Bakker unpublished results). The significantly lower ($p<0.05$) average gene diversity observed for the half-sib family from a tree in a forest as compared to a tree in urban area could be caused by the fact that in a forest, where natural rejuvenation takes place, there are more related trees that can act as pollinators than in urban area, where all trees are planted. However, as this difference in gene diversity was relatively small and as the two half-sib families did not differ significantly for their effective number of alleles, these observations indicate that location (forest or urban area) did not influence the genetic variation of the half-sib

families. The level of genetic variation observed for a half-sib family depends on the number of different fathers that contributed to the offspring. Apparently, the number of different fathers pollinating a tree in a forest did not differ much from the number of different fathers pollinating a tree in urban area.

Population structure

Spatial coancestry analysis revealed non-significant (α=0.05) spatial coancestry (ρ_{ij}) values close to zero for both autochthonous populations (data not shown). By looking for at least one common allele per locus between two individuals some family relationships could be detected, but a large number of family relationships could be attributed to cryptic gene flow. Therefore it can be concluded that no population structure could be detected in either of the two indigenous *Q. robur* populations. This absence of population structure confirms the observation that the populations were in HWE. In a French mixed *Q. robur* population the average Moran's I value (an autocorrelation measure very similar to spatial coancestry analysis; Sokal & Oden 1978) for the same six microsatellite loci showed a weak population structure (average Moran's I values up to 0.035) with significant ($p<0.05$) positive autocorrelation between 20 and 40 m (Streiff et al. 1998). Allozyme studies of other *Quercus* species (*Q. macrocarpa, Q. laevis, Q. chrysolepis*) revealed weak population structures (average Moran's I values up to 0.1 ;spatial coancestry values up to 0.25) for distances up to 10 m (Berg & Hamrick 1995; Montalvo et al. 1997). Such weak spatial genetic structures are typical for forest tree species as these species are characterized by high levels of gene dispersal and immigration. Moreover, *Q. robur* and *Q. petraea* do not rejuvenate under a closed canopy, which prevents the formation of clusters of related individuals (Oosterbaan & Van Hees 1997). Micro-environmental selection did not take place, as there were no loci that showed a significant spatial structure (Epperson 1992). Another explanation for an absence of a spatial structure is the small sample size of the plots in "De Meinweg" and "De Stompert". It is possible that for larger samples over larger areas a weak structure will be found as for *Q. robur* in France (Streiff et al. 1998).

Implications for management and conservation

This study has shown that there was no inbreeding within two indigenous *Q. robur* populations in the Netherlands and that the genetic variation observed for these populations was as high as for unmanaged populations in France and Germany. This indicates that these indigenous *Q. robur* populations – and probably also the rest of the indigenous *Q. robur* populations in the Netherlands – do not run any risk of extinction due to inbreeding and loss of genetic variation. The observed high genetic variation guarantees a highly genetically variable offspring when mating is random. As it is not

possible to start the coppice practice again (due to the long abandonment of the coppice system the trees cannot sprout after they have been cut, J. Riemens personal communication) natural rejuventation will be necessary in order to preserve these populations. Natural rejuvenation only takes place at large gaps: in both "De Stompert" and "De Meinweg" oak seedlings have been observed in patches of heather. As there are indications for long-distance pollen transfer, contaminations of the indigenous material with non-indigenous material cannot be excluded (Streiff et al. 1999). It should be noted that seeds from indigenous trees are only for a part of their genetic material indigenous. The paternal half of the genetic material could come for a part from non-indigenous trees. Paternity analysis will give more information about gene flow in oak populations in the Netherlands. The observation that the genetic variation observed for a half-sib family obtained from a tree at a roadside in urban area was as low as for a half-sib family from an autochthonous population indicates that care should be taken to harvest many different trees at different locations – both for an autochthonous population as for a roadside plantation – in order to ensure a high genetic variation in new plantations.

The low genetic differentiation between the two indigenous populations indicates that – based on theoretical calculations – pollen flow can reach large distances. This indicates that "De Meinweg" and "De Stompert" (and probably all *Q. robur* populations in the Netherlands) are possibly one large intermating population. Although this indicates that the genetic variation will in this way be maintained at a high level, this does not mean that all indigenous populations in the Netherlands are the same. As these populations are very old, they have been adapted to the present local environmental and climatological circumstances. The number of microsatellites used in this study was too small to detect adaptational differences between populations. When more information becomes available about linkage between markers and quantitative trait loci (QTLs) from genetic linkage maps, it will be possible to study genetic variation and selection for eco-physiological traits.

Acknowledgements

The authors thank Marie-Claire Boerwinkel and Jeanine Lagendijk for their help with collecting leaf material and DNA extractions; Jan Bovenschen for his technical help in the laboratory; Kuke Bijlsma and Joukje Buiteveld for their comments on the manuscript; Bert Maes, Jaap Riemens, and G.J.H. Hendriks for information and discussion about the indigenous populations.

5 Mapping of quantitative trait loci for eco-physiological traits on a full-sib genetic linkage map of *Quercus robur* L. based on microsatellite and AFLP markers

Erica G. Bakker[1,2], Herman J. Van Eck[2], Barbara C. Van Dam[1], Evert Jacobsen[2]

Summary

A genetic linkage map of *Quercus robur* L. was constructed based on 101 full-sib individuals – obtained after paternity analysis of 397 half-sib seedlings – using six microsatellites and 343 AFLP markers obtained from 24 *Eco*RI/*Mse*I and 22 *Pst*I/*Mse*I primer combinations. Maternal (496 cM, 13 linkage groups) and paternal linkage map (566 cM, 13 linkage groups) were constructed according to the double pseudo-testcross mapping strategy. The maternal and paternal linkage groups could be integrated using 58 'allelic bridge' markers. One remarkable linkage group was observed that contained 118 paternal AFLP markers(48% of the total number of 247 paternal markers). On the integrated map quantitative trait loci (QTLs) were mapped by means of interval mapping and Kruskal-Wallis test. These QTLs controlled flushing stadium, several leaf morphology traits, length abaxial hairs on leaf, plant height and growth rate, number of branches, double top, mildew resistance, leaf coloration in autumn, and number of old leaves staying in winter. Highest LOD scores/ Kruskal-Wallis K-values indicating major QTLs were observed for leaf width : sinus width ratio (LOD = 3.58), number of veins (LOD = 4.60), plant height (LOD = 4.30 and 4.64 for chromosomes 4 and 12, respectively), plant growth rate (LOD = 4.17), number of lobes (p = 0.0001), and flushing stadium (p = 0.0001). The QTLs were unequally distributed over the genome. The relationship between the observed QTLs for leaf morphology traits and the taxonomic distinction between *Q. robur* and *Q. petraea* is discussed.

Key words: *Quercus robur*, linkage map, double pseudo-testcross, AFLP, microsatellites, QTL-analysis, paternity analysis, full sib

[1] *Department of Ecology and Environment, Alterra, P.O. Box 47, 6700 AA, Wageningen, The Netherlands*

[2] *Laboratory of Plant Breeding, Wageningen University, P.O. Box 386, 6700 AJ, Wageningen, The Netherlands*

Introduction

Quercus robur L. (pedunculate oak) is a widespread oak species in Europe and of importance for its ecological and economic values. Studies based on leaf morphology (Rushton 1979) and molecular markers (Moreau et al. 1994; Bodénès et al. 1997; chapter 2 of this thesis) showed that *Q. robur* is closely related to another oak species, *Q. petraea* (Matt.) Liebl.. Population genetic studies on indigenous populations of these species revealed a high level of genetic variation within populations and a low differentiation between populations (Zanetto & Kremer 1995; Streiff et al. 1998; chapter 4 of this thesis). These observations can be explained by the fact that *Q. robur* and *Q. petraea* are wind pollinated, obligatory outcrossing, long-lived tree species (Hamrick et al. 1979; Hamrick & Godt 1996). The two oak species are diploid (2n=2x=24) and possess similar karyotypes as revealed by fluorochrome banding and FISH patterns (Zoldos et al. 1999). The physical genome size of *Q. robur* is larger (1.88 pg/2C equals 870 Mb) as compared to other tree species (*Populus*, *Eucalyptus*, *Acacia*, *Pyrus*, *Fraxinus*) but smaller as compared to many other plant species (Arumuganathan & Earle 1991; Favre & Arnould 1996; Favre & Brown 1996; Zoldos et al. 1998).

As *Q. robur* and *Q. petraea* majorily differ for their leaf morphology and as there are indications that the species differentiation is based on only a few genes (Bodénès et al. 1997a; chapter 2 of this thesis), it is expected that the difference between the two species is majorily based on a few genes coding for leaf morphological traits. However, it remains unclear which genes code for the morphological differences between these two closely related oak species and which genes are involved in other eco-physiological characteristics of the species. When a genetic linkage map is available with the locations of QTLs controlling eco-physiological and morphological traits, the markers that are closely linked to these QTLs can be used to study natural *Q. robur* and *Q. petraea* populations. When trees of a different age, corresponding with different generations by means of natural rejuvenation, within a population are analyzed with these markers, selection for eco-physiological and morphological traits can be studied. For eco-physiological and morphological traits involved in the differentiation between *Q. robur* and *Q. petraea* this means that in this way a better insight will be gained in the hybrid nature of intermediate type trees.

So far, a genetic linkage map of *Q. robur* has been constructed based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers (Barreneche et al. 1998). However, this map does not have a high marker density and no morphological or eco-physiological traits have been mapped. In this study AFLP was used (Vos et al. 1995), as it generates 50-100 bands per primer combination, of which 10-20 are polymorphic markers. Although AFLP is a co-dominant marker technique, the difference between single and double intensities of heterozygous and homozygous genotypes are usually not sufficient for reliable scoring. Therefore it is commonly used

as a dominant marker technique for the construction of genetic linkage maps because of its high multiplex rate in markers per primer combination and its high reproducibility (Vos et al. 1995). Moreover, in the offspring of non-inbred parents the issue of dominance of the marker is of minor importance, because 3:1 segregating markers are a minority.

The power of the detection of quantitative trait loci (QTLs) depends on the heritability of the trait, the number of loci involved, and the mapping population offspring size. Previous research on progeny and provenance testing of *Q. robur* has not been very successful in finding traits with high heritabilities. For *Q. robur* the highest heritabilities were found for stem straightness, epicormic shoots, and tree height (Jensen et al. 1997). Heritabilities were calculated based on open-pollinated offspring families and on bulked provenance samples, which are not exact repetitions of the same genetic material (Jensen et al. 1997). In this study by Jensen et al. (1997) no information could be found about high heritabilities of other traits. For other tree species high heritabilities were observed for leaf morphology (*Populus, Crepis*, and *Eucalyptus*; Potts & Jordan 1994; Wu & Stettler 1996; Andersson 1999), disease resistance (*Eucalyptus* and *Populus*; Bradshaw & Grattapaglia 1994), and flushing (*Pseudotsuga, Eucalyptus, Populus*, and *Picea*; El-Kassaby & Park 1993; Bradshaw & Grattapaglia 1994; Niensteadt 1984).

This study aims to construct a genetic linkage map of *Q. robur* and to locate QTLs controlling morphological traits involved in species differentiation as a verification of the markers involved in the genetic differentiation between the sympatric species *Q. robur* and *Q. petraea*. Furthermore, QTLs controlling other eco-physiological traits and traits that are important for forestry and biodiversity conservation will be studied.

Material and methods

Plant material

A full-sib mapping population (N=101) was selected from half-sib seedlings (N=397) from a wind pollinated single *Q. robur* mother tree. This mother tree was selected because of its solitary position in an urban area (corner A. Loosjesstraat and Zuideinde in Amsterdam, The Netherlands). Another *Q. robur* tree – from now onwards referred to as the paternal tree – at 8 m distance from the maternal tree, was selected to act as a pollen donor. Within a range of 10 m two more *Q. robur* trees were located near the maternal and paternal tree (Fig. 5.1). The trees were isolated by buildings from other *Q. robur* or *Q. petraea* trees (>100 m). Acorns were collected from the maternal *Q. robur* tree and were sown directly outside after collection in September in small pots containing sandy soil.

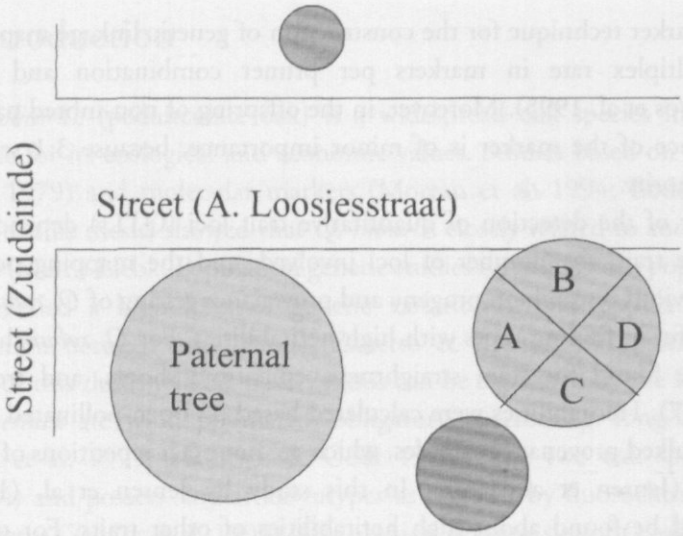


Figure 5.1

Location of the parental trees and two other trees standing isolated in an urban area along a roadside. For the maternal tree the locations of the seed batches have been indicated with letters A-D, corresponding with the different seed batches that were collected from this tree.

DNA extraction

In June, young leaves of the recently emerged seedlings were collected and DNA was extracted with a DNA extraction kit (Puregene®, Gentra Systems, Minneapolis, USA) including 4% PVP-40 to remove phenolic compounds.

Microsatellite analysis

Four microsatellite markers (ssrQpZAG104, ssrQpZAG9, ssrQpZAG1/5, ssrQpZAG36) developed by Steinkellner et al. (1997) and two other microsatellite markers (MSQ4 and MSQ13) developed by Dow et al. (1995) were used. These markers amplify at (AG)*n* or (TC)*n* dinucleotide repeats. PCR amplification, electrophoresis on 6% standard denaturing polyacrylamide gels, and silver-staining of the gels was done according to Streiff et al. (1998).

Table 5.1

Fisher exact test results of pollination proportions of different seed batches (A-D; see Fig. 5.1) collected from a tree isolated next to the pollen donor tree in an urban area. Batch A is the seed batch that was sampled from the side of the maternal tree closest to the paternal tree.

BATCH	HALF-SIB[1]	FULL-SIB[2]
A	135	60
B	129	27
C	58	9
D	57	3
Unknown	18	2

1) The number of half-sib seedlings obtained for each of the four seed batches.
2) The number of full-sib seedlings that have been pollinated by the paternal tree.

AFLP analysis

AFLP analysis was performed according to Vos et al. (1995). Two restriction enzyme combinations (*Eco*RI/*Mse*I and *Pst*I/*Mse*I) were used in order to achieve a uniform coverage of the genetic linkage map. *Eco*RI- and *Mse*I-based primers containing one selective nucleotide (Eco+A and Mse+C) and *Pst*I-based primers without a selective nucleotide were used for pre-amplification of the template. The selective amplification was performed with 24 and 22 fluorescently labeled Eco- and Pst-primers, respectively, in combination with unlabeled Mse-primers. Electrophoresis and visualization of marker bands was performed on a LiCor sequencing machine.

Paternity analysis of the F_1 population

Paternity analysis was performed with six microsatellite loci. Only those trees that matched the presumed father for all alleles on all six loci were considered to be full-sibs. Exclusion probability according to Smouse & Chakraborty (1986) was expected to be as high as 99.99% as was observed by Streiff et al. (1999) for the same six microsatellite loci. To investigate the process of wind pollination seeds were kept in four separate batches from four different positions in the mother tree (Fig. 4.1). Differences in pollination success of the paternal tree between the different seed batches were calculated by means of a chi-square goodness-of-fit test (Tab. 5.1).

Map construction

AFLP markers were scored for the absence (aa) or presence (ab for 1:1 segregating

markers or b- for 3:1 segregating markers) of a segregating marker band using the image analysis computer program CrossChecker (Buntjer 2000). AFLP markers nomenclature was based on the first letter of the restriction enzyme combination, the letters of the selective nucleotides at the 3' end of the primers, and the mobility of the amplification products relative to the Sequamark 10-base ladder (Research Genetics, Huntsville, AL, USA). All scored markers were split into 1:1 and 3:1 segregating markers according to the genetic models <abXaa>, <aaXab>, and <abXab>, which corresponds to markers heterozygous in the female or male parent, or in both, respectively, using the software package Splitloc (http://www.spg.wau.nl/pv/pub/splitloc/). These maternal and paternal data sets were used for the construction of the maternal and paternal map, respectively. The 3:1 segregating AFLP markers and the 1:1:1:1 and the 1:1:1 segregating microsatellite markers served as 'allelic bridge' markers. These markers were mapped on the maternal and paternal maps, while keeping the order of 1:1 segregating markers fixed. They were used to align the maternal and the paternal map with each other.

Linkage analysis of the markers was done according to the double pseudo-testcross approach for cross pollinating populations (Grattapaglia & Sederoff 1994) using program JoinMap 2.0 (Stam 1993; Stam & Van Ooijen 1995). Mendelian segregation of all AFLP and microsatellite markers was tested in the offspring using Chi-square goodness-of-fit tests with program JoinMap 2.0 (Stam 1993; Stam & Van Ooijen 1995). AFLP markers that showed significant deviation from the expected Mendelian segregation were checked on the gel image for the presence of two superimposed bands with the same electrophoretic mobility. If this was the case, the AFLP marker was discarded from the data set. Markers that truly displayed a significant segregation distortion could indicate the presence of a lethal allele or close linkage to a locus that is under strong selection.

Markers were first assigned to linkage groups using a LOD score threshold of 4.0. Per linkage group a preliminary marker order was determined. According to this order the raw data were displayed in Excel where conditional formatting of the cells resulted in graphical genotypes that displayed recombination events as a change of color of a column (personal communication H. van Os, Laboratory of Plant Breeding, Wageningen University, Wageningen, The Netherlands). In this way the data could be checked for singletons. Singletons are single markers that appear to be flanked at either side by a recombination event. Such double recombination events in a short interval are not likely due to negative chiasma interference. Therefore singletons are likely to be scoring errors. As scoring errors can seriously distort the marker order of a linkage map (Van Os et al. 2000), singletons were re-evaluated by inspection of these data points on the original gel images. Subsequently, linkage maps were constructed based on pairwise recombination values obtained with threshold values for LOD > 0.001 and for REC < 0.499. Map distances in centiMorgans (cM) were obtained by using the Kosambi

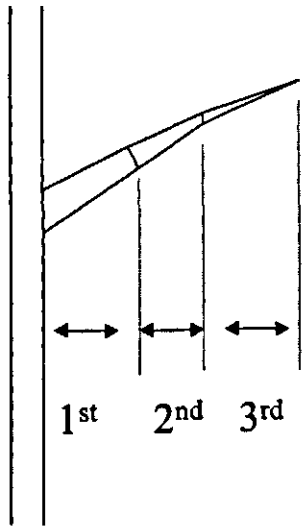


Figure 5.2
Schematic presentation of a branch of a seedling that has experienced second growth and formed lammas shoots. Indicated are the first, second, and third growth (= lammas shoot).

(Kosambi 1944) mapping function. The maternal and paternal map were drawn using the freeware package MapChart 2.0 (Voorrips 2001; http://www.joinmap.nl) (Fig. 5.2). Subsequently, 'allelic bridge' markers were connected with lines between the aligned maternal and paternal linkage groups. Because QTL analysis should not be based on individual parental maps, an integrated map was constructed after combining the aligned maternal and paternal linkage groups. It was taken care of that the order of the 1:1 segregating markers was the same as in the maternal and the paternal maps. In the case of changes in this order, because of limitations in the JoinMap algorithm, the option of 'fixed-order-files' was used. In the case that JoinMap could not calculate the order due to negative distances the 'allelic bridge' markers were located manually on the map.

QTL mapping

The full-sib family and the parents were analyzed for 35 traits (Tab. 5.2). Besides flushing stadium, all traits were measured at the end of the growing season of two consecutive years (1999 and 2000). Leaf morphology traits LL, PL, LW, WP, NL, NV, OB, PR, LDR, and PV were analyzed for five dried leaves randomly sampled from different sides of each tree (two-year-old seedlings) according to Rushton (1978). As abaxial pubescence could be more clearly scored for dried leaves and it appeared after comparison between fresh and dried leaves that no hairs had disappeared after drying, abaxial pubescence was analyzed for two dried leaves of each tree (two-year-old seedlings). An area of 5 mm^2 (80× magnification) next to the midrib of each leaf was analyzed for the length of the hairs (HRL), which was described in three categories: 1 (small), 2 (intermediate), and 3 (large). Subsequently, the number of small and

Table 5.2
Description of traits.

TRAIT CODE	TRAIT NAME	MEASURED IN:	RANGE	MOTHER	FATHER
LL	Leaf length	mm	48.4-80.2	72.4	82.0
PL	Length of leaf petiole	mm	1-4.4	6.4	6.0
LW	Leaf width	mm	14.6-25	26.8	27.8
SW	Sinus width	mm	4.4-11.8	12.4	12.4
WP	Leaf length till widest part	mm	28.2-50.2	43.4	46.6
NL	Number of lobes	number	5.4-12.6	7.0	5.6
NV	Number of intercalary veins	count	0.4-8.2	1.4	4.8
OB	= LL/WP	-	1.4-1.94	1.7	1.8
PR	= PL/(LL + PL)	-	0.02-0.06	0.082	0.068
LDR	= LW/(LW – SW)	-	1.31-2.68	1.9	1.9
PV	= NV*100/NL	-	6.67-88.8	20.8	86.7
HRL	Length hairs on leaf	categorical	1-3	1	3
HRSN	Number of small hairs	categorical	1-4	1	0
HRLN	Number of large hairs	count	0-40	0	90
HRMV	Number of hairs on main vein	categorical	0-2	0	2
LE1	Length stem seedlings	cm	7-30	-	-
LE2	Length stem second year	cm	20-79	-	-
LGR	Length growth = LE2 – LE1	cm	0-64	-	-
DM1	Diameter stem seedlings	mm	2.6-7.2	-	-
DM2	Diameter stem second year	mm	6-15	-	-
DGR	Diameter growth = DM2 – DM2	mm	1.1-11.3	-	-
NB1	Number of branches seedlings	count	0-6	-	-
NB2	Number of branches second year	count	2-23	-	-
DT1	Double top seedlings	presence/absence	0-1	-	-
DT2	Double top second year	presence/absence	0-1	-	-
EPI	Number of epicormic shoots	count	0-13	0	0
SB	Number of shoots at the base of the stem	count	0-3	0	0
FLU	Flushing stadium	categorical (1-5)	1-5	-	-
SFLU	Second flushing stadium	categorical (1-5)	1-5	-	-
SGR	Second growth	cm	4-28	-	-
LAS	Length longest lammas shoot	cm	0-38	0	25
MD	Mildew infection	categorical (0-5)	0-5	0	3
FU	Infection with fungus *Telephora terrestris*	categorical (0-3)	0-3	-	-
LCOL	Leaf coloration in autumn	categorical (1-5)	1-5	-	-
OL	Number of old leaves at 06/03/01	categorical (0-3)	0-3	0	0

intermediate hairs (HRSN) and the number of large hairs (HRLN) were counted. In addition, the number of hairs on the midrib of the leaf were counted (HRMR; 0 (no hairs), 1 (<10 hairs) to 3 (>10 hairs)). The seedlings were measured in both their first and second year for the length of the main stem (LE1 and LE2) and root-neck diameter (DM1 and DM2). Based on the observations of tree height and diameter in two subsequent years, length growth rate (LGR) and diameter growth rate (DGR) could be calculated for the period of one year. The number of branches was counted in both the first (NB1) and the second year (NB2). The presence of a double top was observed in both the first (DT1) and the second year (DT2). The number of epicormic shoots (EPI) was counted in the second year as the number of short, thin branches below the

first branch. The number of large shoots originating from the base of the stem (SB) was counted in the second year only. The length of the longest lammas shoot (LAS = shoots developed in end June/July and can be recognized based on the different leaf morphology; Fig. 5.2) was measured in the second year only. Mildew infection (MD) was measured in the second year in classes from 0 (no infection) to 5 (completely infected). Fungal infection with *Thelephora terrestris* at the base of the stem (FU) was measured after the second year in classes from 0 (no infection), 1 (parts of the base of the stem are infected) to 3 (complete infection). Flushing stadium (FLU) was measured in the second year on 28 April in classes from 1 (bud closed) to 5 (leaves have emerged). When observing the growth of the seedlings it appeared that the initial flushing stopped in June, after which some seedlings started flushing again. This was called the second flushing stadium. At the end of the growing season some of those trees that flushed for a second time developed lammas shoots, which is in fact a third flushing (Fig. 5.2). Second flushing (SFLU) was measured in the second year on 15 June in classes from 1 (no second flushing), 2 (sprouts < 1 cm), 3 (sprouts 1 – 5 cm), 4 (sprouts 5 – 10 cm), to 5 (sprouts > 10 cm). Second growth rate (SGR) was measured in the second year as the part of the main stem that was formed during the second flushing till the end of the growing season or till the start of the lammas shoot. Leaf discoloration (LCOL) in autumn was measured in the first year on 29 October in classes from 1 (completely green) to 5 (completely yellow). The number of old leaves that did not fall off the plants during winter was measured in the winter after the second year in classes from 0 (no old leaves) to 3 (many old leaves).

QTLs were mapped on the integrated linkage map of *Q. robur*. Only when the maternal and paternal map could not be integrated unambiguously, QTL mapping was also performed on the separate maternal and paternal maps. Quantitative traits were tested for normality and, when necessary, transformed (square root or 10log). Normally distributed quantitative traits were analyzed using the interval mapping approach (Lander & Botstein 1989) with program MapQTL 4.0 (Van Ooijen 2000). Significance thresholds for QTLs as detected after interval mapping were defined after 1000 permutations according to Churchill and Doerge (1994) with program MapQTL 4.0 (Van Ooijen 2000). Subsequently, all traits were analyzed by means of a Kruskal-Wallis test according to Kruglyak & Lander (1995) with program MapQTL 4.0 (Van Ooijen 2000). Significance thresholds for QTLs as detected after Kruskal-Wallis test were defined by dividing significance threshold α=0.05 by 24 (the total number of chromosome arms). QTLs were detected when for either the interval mapping or the Kruskal-Wallis test the significance was above the pre-specified threshold.

Results

Paternity analysis of the F_1 population

To identify a full-sib offspring from the half-sib family harvested from a single mother tree, the 397 half-sib seedlings were analyzed with microsatellite markers to detect alleles different from the tree growing at 8 m distance (the paternal tree of the full-sib family). Paternity analysis by means of six microsatellite loci revealed that 102 seedlings (26%) were the result of pollination of the mother tree by the father tree. The other two *Q. robur* trees that were standing in close vicinity of the two parental trees contributed each 2% to pollination. This indicates that 70% of the pollen came from trees that were located at distances >100 m. Per microsatellite locus between 11 (for MSQ13) and 30 (for AG104) different alleles were observed for the total half-sib population (N=397). Only when an individual matched the parental alleles for all loci, this individual was assigned to be offspring from that particular combination of parents. As one seedling immediately died after genotyping, this seedling was excluded from further analyses, leaving 101 full-sib individuals for the construction of the genetic linkage map.

Chi-square goodness-of-fit test of the different seed batches collected from four different sides of the mother tree (Fig. 5.1) showed a significant ($p<0.005$) batch-effect. This means that the paternal tree contributed significantly ($p<0.005$) more pollen to the side of the maternal tree that was closest (batch A) as compared to the other sides of the maternal tree (batches B, C, and D).

Map construction

In total 108 primer combinations were tested to select a subset of primer combinations with good looking fingerprints of desired complexity and large numbers of polymorphic markers for the efficient generation of a sufficient number of segregating markers. The number of segregating markers per primer combination ranged between three and 27. A total of 24 *Eco*RI/*Mse*I and 22 *Pst*I/*Mse*I primer combinations were analyzed, resulting in an average of 10 and 11 markers per primer combination, respectively. In total 96 <ab×aa>, 189 <aa×ab>, and 58 <ab×ab> segregating AFLP markers were obtained. The microsatellite loci from the paternity analysis, which belonged to the full-sib population, were used along with the AFLP data for map construction. One microsatellite (MSQ4) did not segregate. Another (MSQ13) segregated for the male parent only according to the genetic model <aa×bc>. Four loci segregated from both parents and allowed full classification: three loci (AG1/5, AG9, and AG36) fitted a three allele genetic model <ab×ac>; one locus (AG104) fitted a four allele genetic model <ab×cd>. The raw data were used to construct a preliminary

marker order to identify data points that were not supported by the flanking markers. These 'singletons' were used to hint scoring errors. Re-evaluation of the gel images revealed that for the maternal map only five singletons were scoring errors and 30 singletons were most likely the result of misinterpretation of unclear bands, which were subsequently replaced with missing values. For the paternal map only seven singletons were scoring errors and 46 singletons most likely resulted from misinterpretation of unclear bands. However, 42 singletons observed for the maternal map and 89 singletons observed for the paternal map were confirmed by the original gel image and were kept unchanged in the data. After singleton checking the data were grouped for a second time.

Grouping of the markers into linkage groups was tested at increasing LOD thresholds. For both the maternal and the paternal map a stable grouping of 13 linkage groups was obtained for LOD scores between 4.0 and 7.0. In contrast to the haploid chromosome number of 12 chromosomes of *Q. robur* (2n=2x=24) 13 maternal linkage groups could be aligned with 13 paternal linkage groups based on the <ab×ab> markers which are in common to the maternal and paternal linkage groups, and thus serve as allelic bridges. One of the paternal linkage groups was highly dissimilar to the other linkage groups in terms of marker density. This linkage group contained 105 out of the 189 <aa×ab> segregating markers and 13 of the 58 <ab×ab> segregating markers, which is almost half (48%) of all paternal markers and 22% of all <ab×ab> segregating markers. This remarkably marker-dense linkage group was homologous to one of the maternal linkage groups that remarkably was composed exclusively of 13 <ab×ab> segregating markers. All 118 paternal markers and all 13 maternal markers in this exceptional linkage group were linked in coupling phase. Besides, a strong overrepresentation of *Pst*I/*Mse*I generated markers has been observed for this dense linkage group. In linkage groups 1-12 there is a slight 1.7 fold overrepresentation of *Eco*RI/*Mse*I markers relative to *Pst*I/*Mse*I generated markers . In the exceptionally dense linkage group this ratio was 0.24, which shows an overrepresentation of *Pst*I (χ^2=98.9; p<0.0001). The corresponding overrepresentation of *Pst*I recognition sites in this linkage group suggests a higher CG content of this exceptionally dense linkage group. This conclusion is based on the preferential occurrence of the *Pst*I recognition sites only. The selective nucleotides have not contributed to this effect, because there was hardly a difference in the CG-content of the selective nucleotides in the dense linkage group (0.48) as compared to this ratio for the rest of the genome (0.43).

The total map length of the maternal map and the paternal map was 496 cM and 566 cM, respectively. Map density was one marker per 4.85 cM and 6.48 cM for the maternal map and the paternal map without the exceptionally dense linkage group, respectively. The map length of the exceptionally dense maternal and paternal linkage groups was 30 cM (density of one marker per 2.31 cM) and 27 cM (density of one marker per 0.26 cM) respectively, which is remarkably short in view of the huge

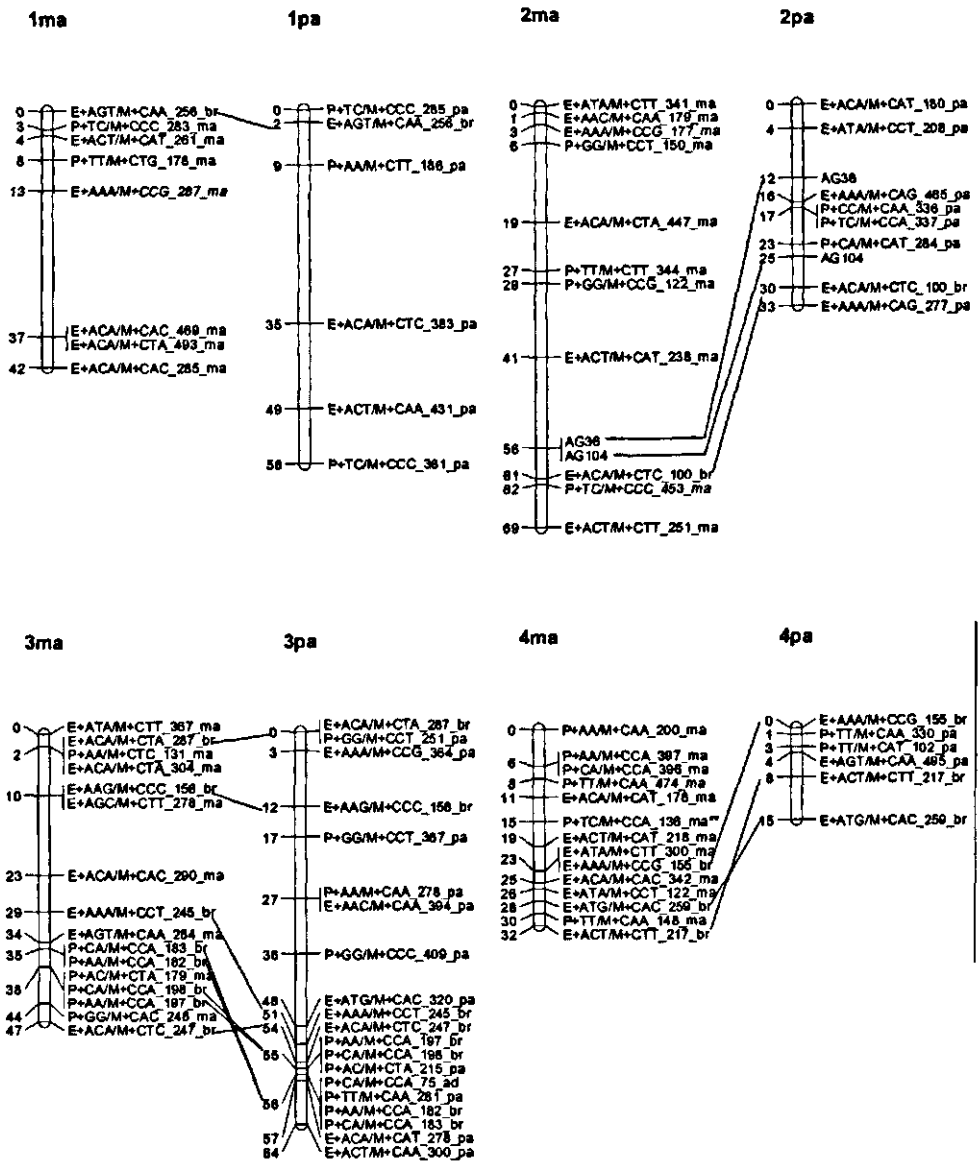


Figure 5.3

Maternal and paternal genetic linkage maps aligned and connected with lines based on 'allelic bridge' markers (indicated with extension "br" behind marker name). Maternal markers are indicated with extension "ma" behind marker name. Paternal markers are indicated with extension "pa" behind marker name. Segregation distortion of markers (chi-square test) is indicated with asterisks: * – ***** corresponding with $p<0.05$ – $p<0.0001$.

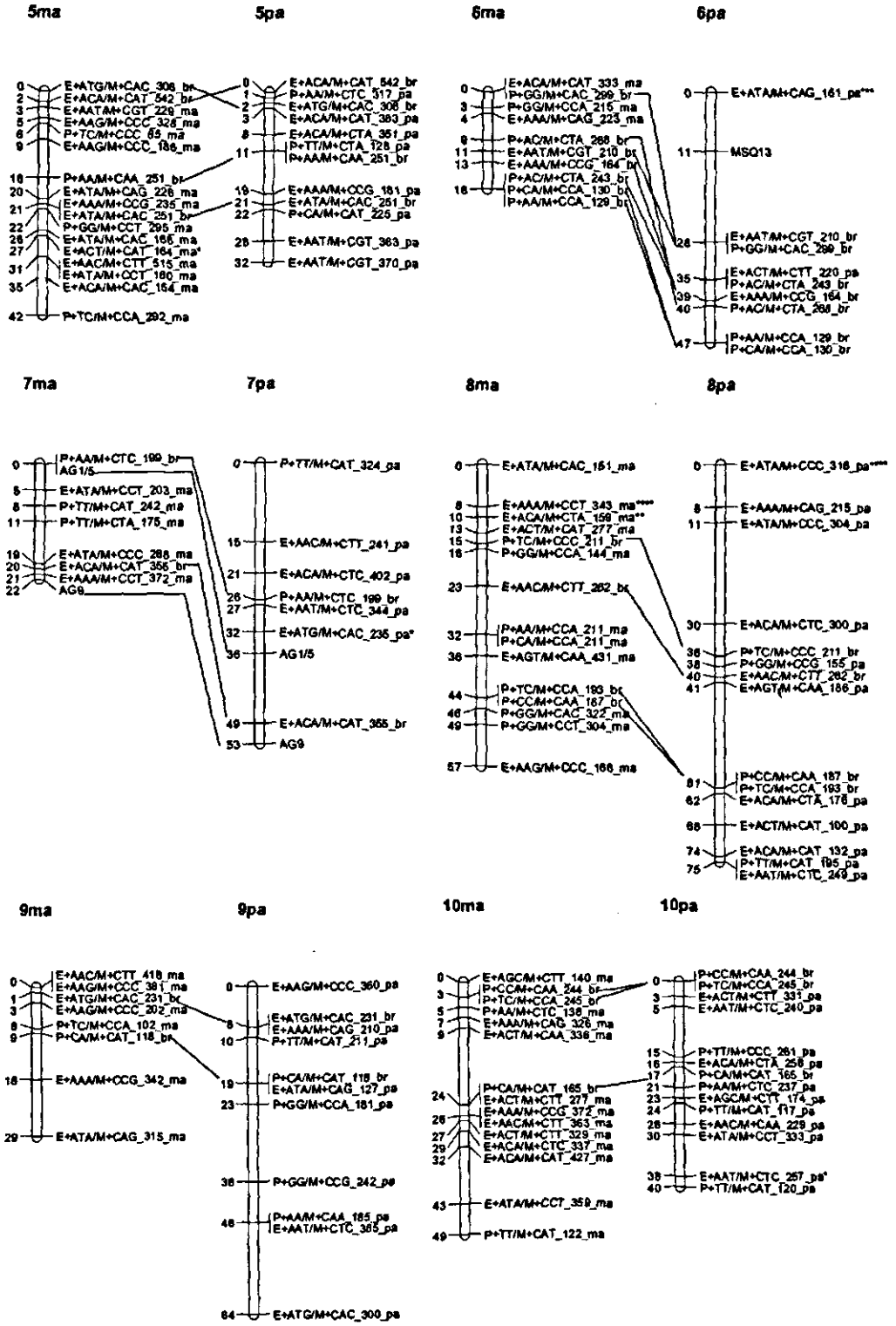


Figure 5.3
(continued)

numbers of markers (Fig. 5.3).

For eight linkage groups (linkage group 1, 2, 4, 5, 8, 10, 12, and 13) the integrated map could be constructed straightforwardly, after merging the corresponding maternal and paternal raw data (loc files). For three linkage groups (linkage group 3, 6, and 9) the integrated map did not reflect the marker order of the separate maternal and paternal map. As this is not acceptable the integrated map was

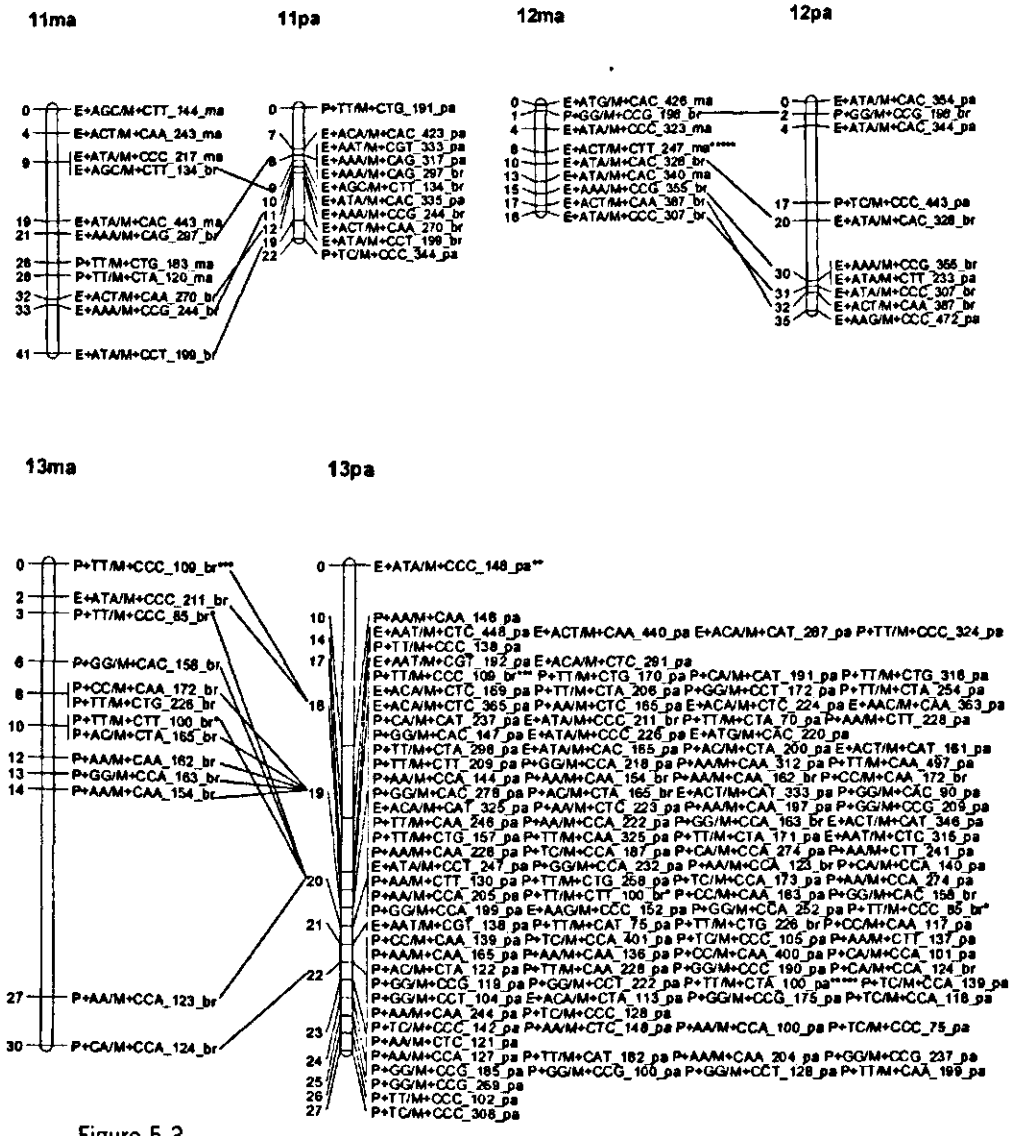


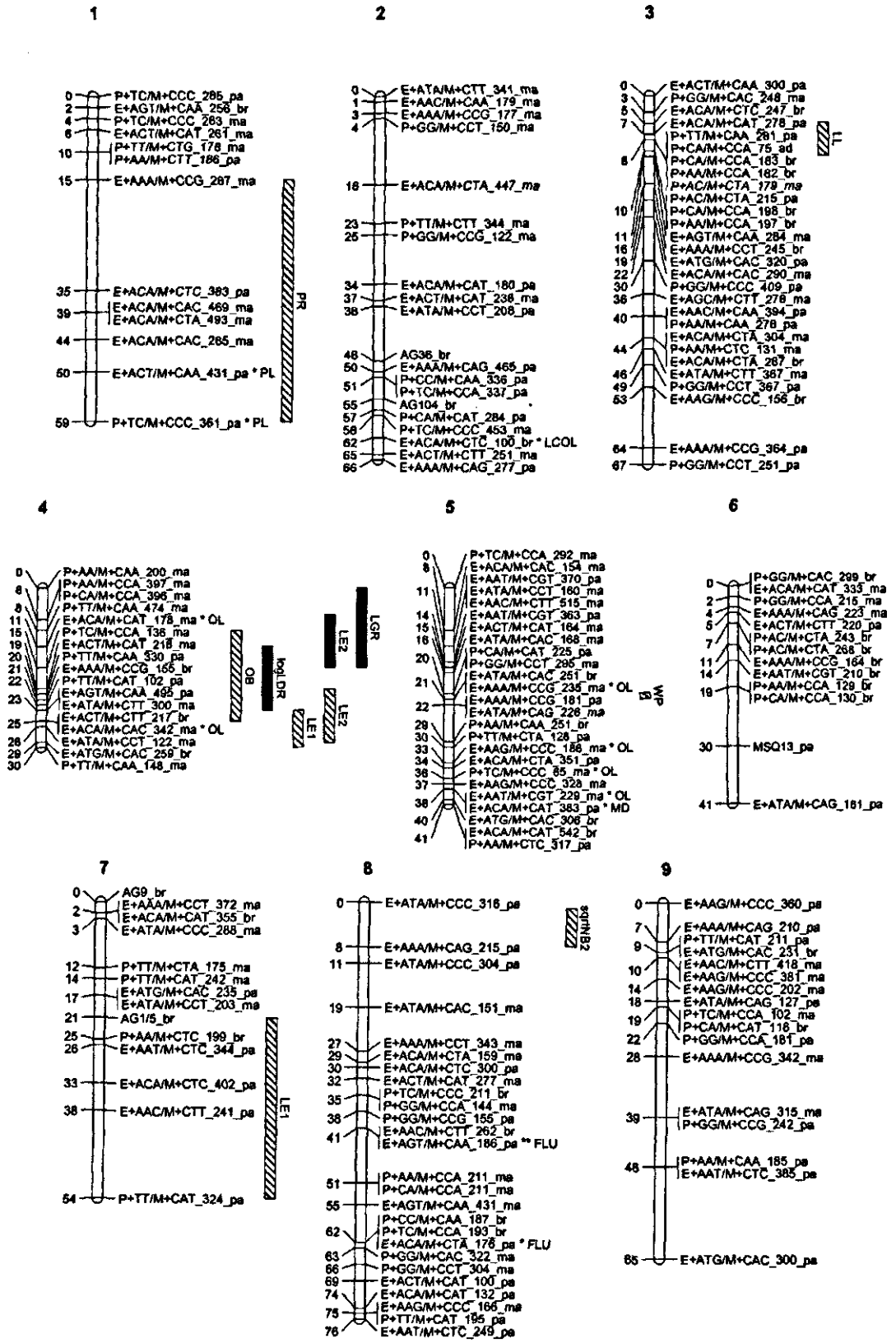
Figure 5.3
(continued)

constructed by forcing JoinMap to accept the marker order of the separate maternal and paternal linkage groups as a fixed order. For six linkage groups (linkage group 3, 4, 5, 6, 11, and 12) some 'allelic bridge' markers could not be placed unambiguously on the map as the positions of the 'allelic bridge' markers differed in their order between the maternal map and the paternal map. In order to construct the integrated map these 'allelic bridge' markers were added only after the construction of the parental maps based on the 1:1 segregating markers. In this way 10 out of the12 linkage groups could be aligned. Two linkage groups (7 and 11) could not be aligned as JoinMap inverted one parental map relative to the other parental map during construction of the integrated map. Removal of 3:1 segregating markers from the set of markers in the linkage group and putting a fixed order of the maternal and paternal markers did not result in the right order of markers in the integrated map. Therefore these two linkage groups (7 and 11) consisted of maternal and paternal markers in opposite order. The total map length of the integrated map was 641 cM. Map density was one marker per 2.74 cM for the map without taking the exceptionally dense linkage group into account (Fig. 5.4). Both *Eco*RI/*Mse*I generated markers and *Pst*I/*Mse*I generated markers were equally distributed over the entire map. No clear clustering of AFLP markers was observed in any of the linkage groups.

QTL mapping

Among the 35 traits measured in the 101 offspring genotypes, ten quantitative traits (LL, SW, WP, OB, PR, LE1, LE2, LGR, DM1, and SGR) showed a normal distribution of the phenotypic values, three other quantitative traits (PV, EPI, and NB2) showed a normal distribution after a square root transformation. Another trait (LDR) was normally distributed after a ^{10}log transformation. The range of the trait values for the F_1 population and the parental values are presented in Tab. 5.2. It should be noted that age and growing conditions of the parental trees and the offspring does not correspond. For 20 traits parental and offspring phenotypic values could be compared. The offspring phenotypes of LL, PL, LW, SW, and PR exceeded the parental trait values. For the remaining six leaf morphology traits (WP, NL, NV, PV, LDR, and OB) progeny phenotypic values were below the parental trait values. Transgressive segregation was observed for six other traits (HRSN, HRLN, EPI, SB, LAS, MD, and OL). Only for three traits (HRL, HRLN, and HRMV) progeny values ranged between the parental values.

Marker-trait associations identified by interval mapping are represented in Tab. 5.3. For logLDR, sqrtPV, LE2, and LGR QTLs were observed with a global (= for the whole genome) significance $p<0.05$. For LL, WP, OB, PR, sqrtNB2, LE1, and SGR QTLs were observed with a local (= per chromosome) significance $p<0.05$. Only few markers that were significantly associated with a trait after interval mapping were also


1
0 P+TC/M+CCC_285_pa
2 E+AGT/M+CAA_256_br
4 P+TC/M+CCC_263_ma
6 E+ACT/M+CAT_261_ma
10 P+TT/M+CTG_178_ma
P+AA/M+CTT_186_pa
15 E+AAA/M+CCG_287_ma
35 E+ACA/M+CTC_383_pa
39 E+ACA/M+CAC_469_ma
E+ACA/M+CTA_493_ma
44 E+ACA/M+CAC_265_ma
50 E+ACT/M+CAA_431_pa * PL
59 P+TC/M+CCC_361_pa * PL
PR
2
0 E+ATA/M+CTT_341_ma
1 E+AAC/M+CAA_179_ma
3 E+AAA/M+CCG_177_ma
4 P+GG/M+CCT_150_ma
18 E+ACA/M+CTA_447_ma
23 P+TT/M+CTT_344_ma
25 P+GG/M+CCG_122_ma
34 E+ACA/M+CAT_180_pa
37 E+ACT/M+CAT_238_ma
38 E+ATA/M+CCT_208_pa
48 AG36_br
50 E+AAA/M+CAG_465_pa
51 P+CC/M+CAA_336_pa
P+TC/M+CCA_337_pa
55 AG104_br
57 P+CA/M+CAT_284_pa
58 P+TC/M+CCC_453_ma
62 E+ACA/M+CTC_100_br * LCOL
65 E+ACT/M+CTT_251_ma
66 E+AAA/M+CAG_277_pa
3
0 E+ACT/M+CAA_300_pa
3 P+GG/M+CAC_248_ma
5 E+ACA/M+CTC_247_br
7 E+ACA/M+CAT_278_pa
P+TT/M+CAA_281_pa
P+CA/M+CCA_75_ad
8 P+CA/M+CCA_183_br
P+AA/M+CCA_182_br
P+AC/M+CTA_179_ma
P+AC/M+CTA_215_pa
10 P+CA/M+CCA_198_br
P+AA/M+CCA_197_br
11 E+AGT/M+CAA_284_ma
16 E+AAA/M+CCT_245_br
19 E+ATG/M+CAC_320_pa
22 E+ACA/M+CAC_290_ma
30 P+GG/M+CCC_409_pa
36 E+AGC/M+CTT_276_ma
40 E+AAC/M+CAA_394_pa
P+AA/M+CAA_278_pa
E+ACA/M+CTA_304_ma
44 P+AA/M+CTC_131_ma
E+ACA/M+CTA_287_br
46 E+ATA/M+CTT_367_ma
49 P+GG/M+CCT_367_pa
53 E+AAG/M+CCC_156_br
64 E+AAA/M+CCG_364_pa
67 P+GG/M+CCT_251_pa
T1
4
0 P+AA/M+CAA_200_ma
6 P+AA/M+CCA_397_ma
P+CA/M+CCA_396_ma
8 P+TT/M+CAA_474_ma
11 E+ACA/M+CAT_178_ma * OL
15 P+TC/M+CCA_136_ma
19 E+ACT/M+CAT_218_ma
20 P+TT/M+CAA_530_pa
21 E+AAA/M+CCG_155_br
22 P+TT/M+CAT_102_pa
23 E+AGT/M+CAA_495_pa
E+ATA/M+CTT_300_ma
25 E+ACT/M+CTT_217_br
E+ACA/M+CAC_342_ma * OL
28 E+ATA/M+CCT_122_ma
29 E+ATG/M+CAC_259_br
30 P+TT/M+CAA_148_ma
OB
holDR
LE2
LE1
LGR
5
0 P+TC/M+CCA_292_ma
8 E+ACA/M+CAC_154_ma
E+AAT/M+CGT_370_pa
11 E+ATA/M+CCT_160_ma
E+AAC/M+CTT_515_ma
14 E+AAT/M+CGT_363_pa
15 E+ACT/M+CAT_164_ma
16 E+ATA/M+CAC_168_ma
P+CA/M+CAT_225_pa
20 P+GG/M+CCT_295_ma
21 E+ATA/M+CAC_251_br
E+AAA/M+CCG_235_ma * OL
22 E+AAA/M+CCG_181_pa
E+ATA/M+CAG_226_ma
29 P+AA/M+CAA_251_br
30 P+TT/M+CTA_128_pa
33 E+AAG/M+CCC_186_ma * OL
34 E+ACA/M+CTA_351_pa
36 P+TC/M+CCC_85_ma * OL
37 E+AAG/M+CCC_328_ma
38 E+AAT/M+CGT_229_ma * OL
E+ACA/M+CAT_383_pa * MD
40 E+ATG/M+CAC_306_br
41 E+ACA/M+CAT_542_br
P+AA/M+CTC_317_pa
WIP
6
0 P+GG/M+CAC_299_br
E+ACA/M+CAT_333_ma
2 P+GG/M+CCA_215_ma
4 E+AAA/M+CAG_223_ma
5 E+ACT/M+CTT_220_pa
7 P+AC/M+CTA_243_br
P+AC/M+CTA_268_br
11 E+AAA/M+CCG_164_br
14 E+AAT/M+CGT_210_br
19 P+AA/M+CCA_129_br
P+CA/M+CCA_130_br
30 MSQ13_pa
41 E+ATA/M+CAG_181_pa
7
0 AG9_br
2 E+AAA/M+CCT_372_ma
E+ACA/M+CAT_355_br
3 E+ATA/M+CCC_288_ma
12 P+TT/M+CTA_175_ma
14 P+TT/M+CAT_242_ma
17 E+ATG/M+CAC_235_pa
E+ATA/M+CCT_203_ma
21 AG1/5_br
25 P+AA/M+CTC_199_br
26 E+AAT/M+CTC_344_pa
33 E+ACA/M+CTC_402_pa
38 E+AAC/M+CTT_241_pa
54 P+TT/M+CAT_324_pa
LE1
8
0 E+ATA/M+CCC_316_pa
8 E+AAA/M+CAG_215_pa
11 E+ATA/M+CCC_304_pa
19 E+ATA/M+CAC_151_ma
27 E+AAA/M+CCT_343_ma
29 E+ACA/M+CTA_159_ma
30 E+ACA/M+CTC_300_pa
32 E+ACT/M+CAT_277_ma
35 P+TC/M+CCC_211_br
P+GG/M+CCA_144_ma
38 P+GG/M+CCG_155_pa
41 E+AAC/M+CTT_262_br
E+AGT/M+CAA_186_pa ** FLU
51 P+AA/M+CCA_211_ma
P+CA/M+CCA_211_ma
55 E+AGT/M+CAA_431_ma
P+CC/M+CAA_187_br
62 P+TC/M+CCA_193_br
E+ACA/M+CTA_176_pa * FLU
63 P+GG/M+CAC_322_ma
66 P+GG/M+CCT_304_ma
69 E+ACT/M+CAT_100_pa
74 E+ACA/M+CAT_132_pa
75 E+AAG/M+CCC_166_ma
76 P+TT/M+CAT_195_pa
E+AAT/M+CTC_249_pa
sqdNDB2
9
0 E+AAG/M+CCC_360_pa
7 E+AAA/M+CAG_210_pa
9 P+TT/M+CAT_211_pa
E+ATG/M+CAC_231_br
10 E+AAC/M+CTT_418_ma
14 E+AAG/M+CCC_381_ma
E+AAG/M+CCC_202_ma
18 E+ATA/M+CAG_127_pa
19 P+TC/M+CCA_102_ma
P+CA/M+CAT_116_br
22 P+GG/M+CCA_181_pa
28 E+AAA/M+CCG_342_ma
39 E+ATA/M+CAG_315_ma
P+GG/M+CCG_242_pa
48 P+AA/M+CAA_185_pa
E+AAT/M+CTC_385_pa
65 E+ATG/M+CAC_300_pa

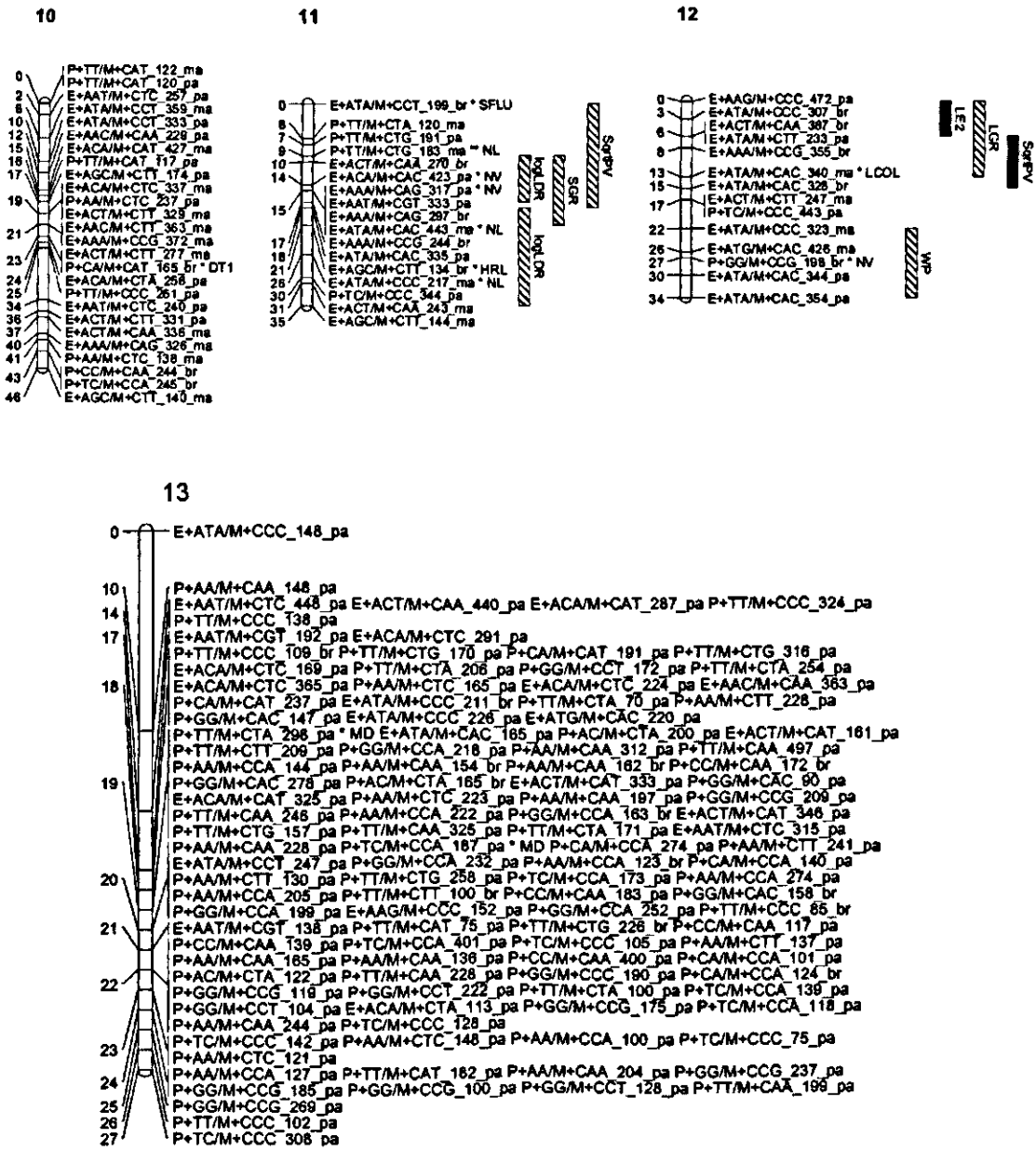


Figure 5.4

Integrated genetic linkage map with the positions of QTLs. Black bars represent the 1 LOD support interval for QTLs. Striped bars represent 1 LOD support intervals for QTLs that are only significant within a linkage group. Traits marked with one asterisk represent a QTL. Traits marked with two asterisks represent a highly significant ($p<0.0001$) QTL.

Table 5.3

QTLs detected by interval mapping of traits that displayed a normal distribution

TRAIT CODE	L[1]	POS. (cm)[2]	MARKER	m[3]	Add. Eff.[4]	Var. (%)[5]	LOD	Sign.[6]	K[7]	P-value[8]	Add. Eff.[9]
LL	3	8.4	P+AC/M+CTA_179	ma	-2.75	13.8	3.02	L*	-	-	-
WP	5	21.3	E+ATA/M+CAC_251	br	-1.62	12.7	2.69	L*	-	-	-
WP	12	30.3	E+ATA/M+CAC_344	pa	0.16	13.7	2.66	L*	-	-	-
OB	4	15.2	P+TC/M+CCA_136	ma	0.01	15.6	2.92	L*	-	-	-
PR	1	58.5	P+TC/M+CCC_361	pa	0.01	14.9	2.76	L*	-	-	-
LogLDR	4	15.2	P+TC/M+CCA_136	ma	-0.03	18.9	3.58	G*	-	-	-
LogLDR	11	15.4	E+AAT/M+CGT_333	pa	0.03	11.9	2.72	L*	-	-	-
LogLDR	11	15.4	E+AAA/M+CAG_297	br	0.03	11.9	2.72	L*	-	-	-
LogLDR	11	26.0	E+ATA/M+CCC_217	ma	0.03	14.2	2.88	L*	-	-	-
SqrtPV	11	15.4	E+AAT/M+CGT_333	pa	-0.80	13.6	3.13	L*	12.29	0.0005	-0.85
SqrtPV	11	15.4	E+AAA/M+CAG_297	br	-0.48	13.6	3.13	L*	-	-	-
SqrtPV	12	12.5	E+ATA/M+CAC_340	ma	1.01	20.9	4.60	G**	13.54	0.0005	-1.00
SqrtNB2	8	0.0	E+ATA/M+CCC_316	pa	0.21	36.4	3.92	L**	-	-	-
LE1	4	25.1	E+ACT/M+CTT_217	br	-1.35	14.2	3.15	L*	-	-	-
LE1	7	32.5	E+ACA/M+CTC_402	pa	0.39	37.9	3.57	L*	-	-	-
LE1	12	3.3	E+ATA/M+CCC_307	br	-	-	-	-	11.04	0.001	-3.87
LE2	4	10.8	E+ACA/M+CAT_178	ma	-0.57	31.2	4.30	G*	-	-	-
LE2	4	22.8	E+ATA/M+CTT_300	ma	-2.57	14.0	2.75	L*	-	-	-
LE2	12	0.0	E+AAG/M+CCC_472	pa	10.14	21.4	4.64	G*	13.57	0.0005	10.29
LGR	4	6.1	P+AA/M+CCA_397	ma	1.57	39.9	3.88	L**	-	-	-
LGR	4	10.8	E+ACA/M+CAT_178	ma	0.34	31.0	4.17	G*	-	-	-
LGR	4	19.6	P+TT/M+CAA_330	pa	-	-	-	-	12.88	0.0005	7.94
LGR	12	0.0	E+AAG/M+CCC_472	pa	8.86	16.4	3.46	L**	12.42	0.0005	8.77
LGR	12	8.4	E+AAA/M+CCG_355	br	2.18	11.1	2.77	L*	-	-	-
SGR	11	15.4	E+ATA/M+CAC_443	ma	-2.27	14.1	3.05	L*	-	-	-

1) Linkage group
2) Position of the QTL peak on the linkage group in cM
3) Map source of the markers: maternal (ma: <ab×aa>), paternal (pa: <aa×ab>, or 'allelic bridge' (br: <ab×ab>.
4) Additive effect after interval mapping is calculated for the maternal and paternal markers as (mean aa genotype (ma or pa) – mean ab genotype (ma or pa)) and for the 'allelic bridge' markers as ((mean (aa – bb) genotype (ma)) – (mean (aa – bb) genotype (pa)))/2.
5) Percentage of explained variance
6) LOD significance has been defined by permuation tests. L= (local) significane per linkage group, G= (global) significance for the entire genome, *=p<0.05, **=p<0.005.
7) K value obtained after Kruskal-Wallis test
8) P-value corresponding with the K value
9) Additive effect after Kruskal-Wallis test is calculated for the maternal and paternal markers as (mean aa genotype (ma or pa) – mean ab genotype (ma or pa)) and for the 'allelic bridge' markers as (mean aa genotype – mean b-genotype).

Table 5.4

QTLs detected by Kruskal-Wallis test of traits that did not show a normal distribution of trait values.

TRAIT CODE	LG[1]	POS. (cm)[2]	MARKER	Map[3]	K[4]	P-value[5]	Add. Effect[6]
PL	1	49.5	E+ACT/M+CAA_431	pa	12.58	0.0005	0.52
PL	1	58.5	P+TC/M+CCC_361	pa	13.34	0.0005	0.53
NL	11	8.8	P+TT/M+CTG_183	ma	16.10	0.0001	-1.03
NL	11	15.4	E+ATA/M+CAC_443	ma	14.53	0.0005	0.98
NL	11	26.0	E+ATA/M+CCC_217	ma	14.91	0.0005	1.04
NV	11	14.2	E+ACA/M+CAC_423	pa	11.41	0.001	1.06
NV	11	15.4	E+AAA/M+CAG_317	pa	13.10	0.0005	1.13
NV	12	27.1	P+GG/M+CCG_198	br	11.46	0.001	1.18
HRL	11	21.2	E+AGC/M+CTT_134	br	12.12	0.0005	1.77
OL	4	10.8	E+ACA/M+CAT_178	ma	10.98	0.001	0.80
OL	4	25.1	E+ACA/M+CAC_342	ma	11.41	0.001	-0.81
OL	5	21.3	E+AAA/M+CCG_235	ma	12.51	0.0005	-0.88
OL	5	32.8	E+AAG/M+CCC_186	ma	11.92	0.001	-0.95
OL	5	35.9	P+TC/M+CCC_65	ma	13.26	0.0005	0.92
OL	5	38.4	E+AAT/M+CGT_229	ma	12.31	0.0005	-0.86
DT1	10	22.6	P+CA/M+CAT_165	br	10.86	0.001	0.22
SFLU	11	0.0	E+ATA/M+CCT_199	br	11.19	0.001	-0.91
MD	5	38.4	E+ACA/M+CAT_383	pa	11.08	0.001	0.99
MD	13	5.4	P+TT/M+CTA_298	pa	11.96	0.001	1.47
MD	13	6.5	P+TC/M+CCA_187	pa	12.51	0.0005	1.07
LCOL	2	62.1	E+ACA/M+CTC_100	br	13.53	0.0005	-0.72
LCOL	12	12.5	E+ATA/M+CAC_340	ma	11.21	0.001	-0.60
FLU	8	41.3	E+AGT/M+CAA_186	pa	15.78	0.0001	-0.83
FLU	8	62.4	E+ACA/M+CTA_176	pa	11.69	0.001	-0.83

1) Linkage group
2) Position on the linkage group in cM
3) Map source of the markers: maternal (ma: <ab×aa>), paternal (pa: <aa×ab>, or 'allelic bridge' (br: <ab×ab>.
4) K value obtained after Kruskal-Wallis test
5) P-value corresponding with the K value
6) Additive effect is calculated for the maternal and paternal markers as (mean aa genotype (ma or pa) – mean ab genotype (ma or pa)) and for the 'allelic bridge' markers as (mean aa genotype – mean b- genotype).

detected by the Kruskal-Wallis test. From among the 21 traits that were not normally distributed for ten traits (PL, NL, NV, HRL, OL, DT1, SFLU, MD, LCOL, and FLU; Tab. 5.4) a QTL was detected. For two of these traits (NL and FLU) the QTL was highly significant ($p<0.0001$).

As for linkage groups 7 and 11 the integrated map consisted of maternal and paternal markers in reverse order, these linkage groups were investigated separately for the maternal and the paternal map. Interval mapping of all normally distributed traits on these two linkage groups resulted in the detection of different positions of a QTL as

was found for the integrated map. However, in both cases the QTL was located within the 1 LOD support. Two QTLs were found for SW and sqrtNB2 on linkage group 11 of the maternal map. One QTL was found for OB on linkage group 7 of the paternal map. These three QTLs were not found on the integrated map.

The distribution of QTLs controlling for a particular trait over the genome was unequal ($p<0.01$) after doing a chi-square goodness-of-fit test of the distribution of QTLs controlling for a particular trait over the linkage groups. On two linkage groups no QTLs were detected (linkage groups 6 and 9), on five linkage groups one QTL was detected (linkage groups 2, 3, 7, 10, and 13) (Fig. 5.4). Loci involved in two or more traits were detected for linkage groups 1 and 8 (2 traits), linkage group 5 (3 traits), linkage groups 4 and 12 (6 traits), linkage group 11 (7 traits).

Discussion

Paternity analysis of the F_1 population

The new approach for obtaining a full-sib family based on paternity analysis of a half-sib family proved to be successful. For *Q. robur* this new approach can be advised as controlled crosses in this species are laborious and the chance of success is low due to mast years and incompatibility between trees (Lumaret et al. 1991; Streiff et al. 1999). The disadvantage of the paternity analysis approach is that choice of parents is limited, as only isolated trees are suitable. Moreover, there is a chance that the isolated trees are incompatible or flower at different times.

Paternity analysis of the half-sib progeny that was collected from a tree in Amsterdam (referred to as the maternal tree) revealed that the tree standing at a distance of 8 m next to the maternal tree (referred to as the paternal tree) pollinated 26% of the progeny of the maternal tree. However, two other *Q. robur* trees that were standing within a range of 10 m distance from the maternal tree contributed each only 2% pollen to the progeny of the maternal tree. These observations correspond with observations of Streiff et al. (1999) in a mixed oak stand of *Q. robur* and *Q. petraea* where similar differences in percentages of pollination for short distances (0-20 m) were found. These differences in pollination success between trees could be attributed to wind direction and velocity, meteorological conditions and local disturbances, fecundity, phenology, and incompatibility (Streiff et al. 1999). It has been observed that all four trees flowered at about the same time. Wind direction is not likely to be important in urban area, as the buildings will break the wind. It is likely that the maternal and paternal tree were compatible as the pollination success was high. Besides, an additional factor could have enhanced the pollination success: isolation of the two compatible *Q. robur* trees from other compatible *Q. robur* trees. Pollen clouds from surrounding trees are likely to be diluted to only a few grains when they reach the

isolated *Q. robur* trees over distances exceeding 100 m (Streiff et al. 1999). Therefore, the highly concentrated pollen cloud of the paternal tree had a higher chance to fertilize the female flowers of the maternal tree. Paternity analysis of seed batches sampled from four different sides of the maternal tree revealed that the paternal tree contributed most of its pollen to the side of the maternal tree that was closest to the paternal tree.

Six microsatellite loci were sufficient to discriminate between full-sib and half-sib as the theoretical exclusion probability (Smouse & Chakraborty 1986) was expected to be as high as the value of 99.99% as was found by Streiff et al. (1999) for the same six microsatellite loci. This method of paternity analysis of a half-sib in order to obtain a full-sib can be recommended instead of controlled crossing of trees, because it is less labor intensive than controlled crossing and success rate is high. The disadvantage of this method is limitation in parent choice.

AFLP and microsatellite markers

Initially we had the impression that the level of polymorphism between the two parental clones was very low. Analysis of pairwise similarity between *Q. robur* trees in a natural population ("De Meinweg") based on simple matching of AFLP bands generated by two primer combinations (E+ATA/M+CCC and E+AAG/M+CCC) that were obtained in a previous study (chapter 2 of this thesis) resulted in a distribution of pairwise similarity values. Based on the analysis of this distribution the upper 5% threshold for highest similarity was 0.88 and the lower 5% threshold for lowest similarity was 0.79. In view of these values the observed pairwise similarity between the maternal and paternal tree of 0.81 is not an exceptional value. Therefore it can be concluded that the maternal and paternal tree were not more genetically identical than any pair of two trees from a natural population.

Map construction

This is the second report on the construction of a *Q. robur* linkage map (Barreneche et al. 1998). However, the phenotype of the paternal tree does not completely match the taxonomic description of *Q.robur* as it possesses too many abaxial hairs on its leaves. In view of this *Q. petraea* resembling character, the paternal tree might be a hybrid between *Q. robur* and *Q. petraea*. Consequently, this could be the first map based on a cross between a hybrid tree and a *Q. robur* tree. The map described in this paper can be aligned with the map of Barreneche et al. (1998) after mapping of additional microsatellite markers (planned for future research in the European project OAKFLOW). However, alignment over many loci is more easily obtained by comparative AFLP analysis of both maps. The locus specificity of AFLP markers has

been proved for many species (e.g. Rouppe van der Voort et al. 1997; Waugh et al. 1997), and was successfully for map alignment (Rouppe van der Voort et al. 1998).

In total 13 maternal and paternal linkage groups could be aligned. However, *Q. robur* has a chromosome number of 12 (2n=2x=24). The data suggest that one linkage group is exceptional. For several reasons this linkage group was remarkable. 1. This linkage group contained a high number of markers. 2. This linkage group contained much more paternal markers (118, which is 48% of all paternal markers) than maternal markers (13). 3. All maternal markers were <abXab> segregating markers. 4. The number of <abXab> segregating markers was high (13, which is 22% of all <abXab> segregating markers). 5. The observed overrepresentation of *Pst*I recognition sites suggests a higher CG content of this linkage group. 6. All markers were in coupling phase, indicating that all markers were derived from one homologous chromosome and were absent for the other homologue. This suggests that for this linkage group both the male and female parent have a homologous chromosome without any marker or the absence of a homologous chromosome. The complete absence of the other homologue within the maternal and paternal genotype would not be expected as its absence would also prevent recombination, which is not the case, although the length of the linkage group is not large. Besides, the absence of the other homologue is not substantiated with cytogenetic observations in this material. Karyotype analysis of full-sib individuals containing all 118 AFLP fragments of the exceptionally dense linkage group did not show an additional chromosome (aneuploidy) or a B-chromosome (J.H. de Jong unpublished results). A remarkable observation was that some of the small chromosomes in the complement showed strikingly weak DAPI fluorescence, which points at great variation in chromatin condensation or heterochromatin - euchromatin differentiation, and so at great differences in DNA content between the individual chromosomes. A possible explanation for an additional linkage group could be the occurrence of a small chromosome region with obligate crossovers (recombination hotspot) on any of the very long chromosomes separating the loci of one chromosome into two linkage groups. Although there are reports about B-chromosomes in *Q. robur*, this can be ruled out as Zoldos et al. (1999) proved that the B-chromosomes as described by Ohri & Ahuja (1990) were in fact satellites separated by long secondary constrictions from the remaining part of the nucleolar organiser chromosomes.

Genome expansion has been reported as the result of increase in the number of transposable elements (Vicient et al. 1999), but genome increase in the paternal tree by rapid duplication of a certain DNA sequence is very unlikely in our case. The observations that all AFLP markers on the exceptionally dense linkage group are unique single copy amplification products, and the lack of high copy AFLP fragments – observed as thick bands within a fingerprint – do not support such hypothesis.

To our knowledge, such a linkage group of short map distance and long physical

distance (assuming uniform distribution of AFLP markers on the genome) has not been observed in earlier mapping studies.

The 'allelic bridge' markers were scored as dominant markers. By scoring them in this way information was lost, which can result in inaccurate estimates of the recombination frequencies. Therefore 3:1 segregating markers are likely to be placed at an inaccurate position on the genetic linkage map (Maliepaard et al. 1997). For this reason less attention was paid to the placement of these markers when constructing the integrated map. Fortunately, these markers were mapped on the integrated map at exactly the same position as on the maternal and paternal maps for several linkage groups. In the case of inconsistencies in marker order of the 3:1 markers on the maternal and paternal version of a linkage group we examined the mapping process of Joinmap carefully and obtained more consistent results by the addition of the 3:1 markers after the 1:1 marker order was obtained and fixed.

Linkage groups 7 and 11 could not be aligned. However, for linkage group 7 there were four 'allelic bridge' markers that were in the same order in both the maternal and the paternal map. Nevertheless, JoinMap inverted one parental map relative to the other parental map during construction of the integrated map, in spite of highly informative microsatellite markers that segregated as <ab×ac> markers. We conclude that this is the result of shortcomings of the algorithm of JoinMap. For linkage group 11 the 'allelic bridge' markers were not in the same order for the maternal and the paternal map. As there were no microsatellite markers located on this linkage group, and due to the inconsistensies in marker order of the 3:1 segregating markers it was not possible to align the maternal and paternal linkage groups properly.

As there was only one gap between two markers exceeding 15 cM, it was expected that the genetic linkage map was sufficiently covered in order to be able to detect quantitative trait loci (QTLs).

QTL mapping

The disadvantage of tree species like *Q. robur* is that the time between two generations is long (about 20 years). This results in the fact that parents and offspring of a mapping population can not easily be compared for many traits, which complicates the description of a Mendelian model of the heredity. Attempts to vegetatively propagate the parents by means of cuttings and grafting failed (E.G. Bakker unpublished results). However, vegetative propagation of seedlings was successful. Therefore the mapping population can be vegetatively reproduced, which enables repetitions and destructive experiments.

The lower values for LL and SW (small leaves) obtained for the F_1 population as compared to the parental values (larger leaves) can be attributed to the juvenile stadium of the F_1 plants. Apparently, two-year-old *Q. robur* seedlings are too small to

bear leaves of a normal size. The lower values observed for the other nine leaf morphology traits (PL, PR, SW, WP, NL, NV, PV, LDR, and OB) could as well be explained by the juvenile stadium of the plants. However, the lower and higher F_1 values as compared to the parental values for WP, NL, NV, PV, LDR, OB, HRSN, HRLN, EPI, SB, LAS, MD, and OL could be the result of heterozygosity in the parents for QTL alleles with positive as well as negative effects.

Permutation analyses (1000 permutations) according to Churchill & Doerge (1994) for LOD thresholds resulted in similar global thresholds as the global thresholds that were recommended by Van Ooijen (1999) using simulation studies. As permutation analysis takes a lot of computing time, the global thresholds as described by Van Ooijen (1999) can be used. However, Van Ooijen (1999) does not give detailed threshold information for local QTLs. As he sets the threshold for detection of local QTLs at 3.2, many local QTLs would not be detected.

Although the two parental trees differed for their abaxial pubescence and the length of the longest lammas shoot and the progeny segregated for these traits, no QTLs were detected for these traits. The reason for this could be that abaxial pubescence and length of the longest lammas shoot are complex traits coded by many different genes located at different positions on the genome. In that case the progeny size was too small in order to be able to detect these QTLs. An approximate multiple-QTL model ('MQM mapping' Jansen 1993) can be used in order to detect these QTLs, as it is a more sensitive method for detecting QTLs. As MQM was not programmed in MapQTL 4.0 for CP (= cross pollinator), this was not possible.

As for chromosomes 7 and 11 the orders of the maternal and paternal markers on the integrated map were opposite to each other, for these two chromosomes QTL mapping was done on the maternal and paternal maps separately. Comparison between QTLs located on the integrated map and the separate maternal and paternal maps showed that QTLs located on the integrated map were also located on the separate maternal and paternal maps. Therefore the QTLs detected on chromosomes 7 and 11 of the integrated map were assumed to be true QTLs.

QTLs controlling plant length

The detection of QTLs for LE1, LE2, and LGR on different positions on chromosome 4 indicates that on this chromosome QTLs controlling plant length and length growth rate are located. However, due to variation in length and length growth rate between years the exact location of the QTL was not found. Next to chromosome 4, chromosomes 7 and 12 seem to be involved in length and length growth rate. The absence of QTLs for some of the following traits LE1, LE2, or LGR on these chromosomes can be attributed to variation in plant length between years. As during the first few years, the length of *Q. robur* seedlings is completely dependent on the size

of the acorn, it is expected that measurements at an older age will reveal the true QTLs for length and length growth rate. Still, the observed QTLs for length and length growth rate based on observations for seedlings could be true QTLs for these traits as all acorns were of the same size and from the same tree. The amount of endosperm inside the acorn should be constant as this trait is coded by the maternal genome. The only variation in acorn size can have arisen due to the orientation of the fruit to the sun when it was developing.

QTLs controlling related traits

The QTLs coding for PL and PR were located at the same position on linkage group 1. In this way the two traits that describe petiole length confirmed each other. The QTLs coding for OB and logLDR were located at exact the same position on linkage group 4. Both traits are related as they both describe the leaf shape: OB describes the ratio between length and width of the leaf, LDR describes the ratio between the width and the sinus of the leaf. Other related traits are sqrtPV and NV as both traits describe the number of intercalary veins. On linkage group 12 QTLs controlling these two related traits were located at a larger distance. It is possible that the QTL for one of these two related traits was not a true QTL. SFLU and SGR are also related as they both describe the second growth of the seedlings in June. The fact that the QTLs controlling for these traits were not located at the same location on the linkage group can be explained by the time interval between measuring the two traits.

Distribution of QTLs over the genome

The distribution of QTLs coding for a specific trait over the genome was unequal. On two linkage groups no QTLs were detected (linkage groups 6 and 9), on four linkage groups one QTL was detected (linkage groups 2, 3, 7, and 10). QTLs controlling for two or more traits were detected for linkage groups 1 and 8 (2 traits), linkage group 5 (3 traits), linkage groups 4 and 12 (6 traits), linkage group 11 (7 traits). On the exceptionally dense linkage group (linkage group 13) two QTLs controlling mildew resistance were detected. This suggests that this additional part of a linkage group contained coding regions.

Clustering of several QTLs controlling unrelated traits was most evident for linkage groups 4 and 11. As the resolution of the QTLs was low (long 1 LOD intervals covering up to 2/3 of a linkage group) it is not clear if on these two linkage groups there are closely linked genes or pleiotropic loci. No evidence was found for clustering of QTLs controlling for unrelated leaf morphology traits. Such a clustering of leaf morphology QTLs was expected as *Q. robur* and *Q. petraea* were found to be genetically very similar based on molecular marker analyses (Bodénès et al. 1997a;

chapter 2 of this thesis). As *Q. robur* and *Q. petraea* are majorily discriminated based on their leaf morphology (Rushton 1979) it is expected that the two species differ only for genes that code for leaf morphology traits. Clusters of closely linked genes or pleiotropic genes coding for several leaf morphology traits would explain the high genetic similarity found based on randomly amplifying molecular markers between the two species.

Power of QTL detection

LOD scores detected in the here presented study were all low (highest LOD score of 4.64). It has been observed that for F_1 populations from non-inbred parents LOD scores do not reach as high values as compared to F_2 or BC_1 populations from inbred parents (Grattapaglia et al. 1995; Yin et al. 1999). The reason for this can be the high number of allele combinations that are possible when both the maternal and the paternal parent are heterozygous (Han 2001).

Additive effects were in general small. Only for LE2 and LGR on linkage group 12 large additive effects of 10.14 cm and 8.86 cm, respectively, were detected. The small additive effects for the other QTLs indicate that these traits are controlled by several QTLs, each one contributing a small additive effect.

The high heritability found for tree height of *Q. robur* by Jensen et al. (1997) was reflected by a QTL for this trait on the map. Although the heritability for tree height was assessed for 17-year-old trees, it appears that one- and two-year-old seedlings reveal a high heritability for this trait as well. For the presence of epicormic shoots no QTL was detected, although Jensen et al. (1997) described a high heritability for this trait. Apparently, the seedlings were too young to be properly evaluated for this trait, moreover, the two parents did not possess many epicormic shoots. High heritabilities for leaf morphology, disease resistance, and flushing based on findings for other tree species (leaf morphology: *Populus*, *Crepis*, and *Eucalyptus*; Wu & Stettler 1996; Andersson 1999; Potts & Jordan 1994; disease resistance: *Eucalyptus* and *Populus*; Bradshaw & Grattapaglia 1994; flushing: *Pseudotsuga*, *Eucalyptus*, *Populus*, and *Picea*; El-Kassaby & Park 1993; Bradshaw & Grattapaglia 1994; Niensteadt 1984) were expected to apply as well for *Q. robur* as QTLs for these traits were located on the *Q. robur* genetic linkage map. Therefore it appears that traits with high heritabilities in other tree species are likely to display high heritabilities in *Q. robur*. This knowledge can be useful in selecting traits for mapping. However, it might be necessary to look for QTLs at different ages of the tree as selection for these traits might not be very successful at a later age as has been shown for tree height: low correlation between response to early selection based on height at age 15 years and height at age 120 years (Lambeth 1980).

The full-sib population that was established in this study can be used for the

next couple of decades for genetic analysis of traits that are displayed at later ages. Besides, after vegetative multiplication additional (destructive) experiments can be done investigating QTLs that control for tolerance to drought- or waterstress. Together with the other genetic linkage maps of *Q. robur* and *Q. petraea* (inter- and intraspecific) constructed in the OAKFLOW project, this genetic linkage map and the mapping population will be a valuable source of information about speciation and all kinds of traits of importance for oak breeding, forestry and ecology.

Acknowledgments

Erica Bakker was financed in part by EU project FAIROAK. The authors would like to thank Petra van den Berg and Fien Meijer-Dekens for their technical assistance in the lab, Hans van Os for his help with map construction, Fred van Eeuwijk for fruitful discussion about the manuscript, and Hans de Jong for karyotype analysis.

6 General discussion

The genetic composition of indigenous Q. robur and Q. petraea populations

This Ph.D. thesis describes the population genetics, molecular ecology, and inheritance aspects of the genetic composition of indigenous *Q. robur* and *Q. petraea* populations in the Netherlands. This study was conducted as an extension of the EU-project "Synthetic maps of gene diversity of oak resources in Europe" (FAIROAK), which resulted in the description of post-glacial re-colonization routes of oaks in Europe. This thesis was partially financed by the Ministry of Agriculture, Nature, and Fisheries (program 320) and the EU-project "Intra and interspecific gene flow in oaks as mechanisms promoting genetic diversity and adaptive potential" (OAKFLOW).

The genetic composition of indigenous *Q. robur* and *Q. petraea* populations could be described using the molecular marker techniques AFLP and microsatellites. These molecular marker techniques differ in some aspects and are therefore suitable for different analyses. In four chapters it has been shown how molecular markers can describe the genetic information of indigenous *Q. robur* and *Q. petraea* populations in order to get a better understanding of the autochthonicity, history, structure, and quality of these populations.

In chapter 2 of this thesis the AFLP technique has been used for species discrimination of trees from an indigenous mixed *Q. robur* and *Q. petraea* population ("De Meinweg"). Comparison between leaf morphology data and AFLP data indicated that the difference between the two species was small and probably only based on a few genes coding for traits (probably leaf morphology traits) involved in the adaptation to the different ecological niches of the two species. This was already suggested based on RAPD and SCAR results by Bodénès et al. (1997a). It is expected that for this reason no diagnostic marker for *Q. robur and Q. petraea* has been found so far. Other studies searching for diagnostic markers between closely related species faced the same problem. A study of six *Cichorium* species based on 184 AFLP markers showed that diagnostic markers were only found for one species (*C. bottae*) that was located outside the geographic range of the other species (Kiers et al. 2000). Only one of the other five *Cichorium* species (*C. endivia*) possessed one unique AFLP band. A study of a hybrid zone between *Q. grisea* and *Q. gambelii* based on RAPD markers by Howard et al. (1997) did not detect any diagnostic markers, but RAPD band frequencies for species-indicative markers – markers that are completely absent in one species and present at a

high frequency in the other species – were all above 0.61. In the study presented in chapter 2 of this thesis we also found species-indicative markers. This is the first report on the observation of species-indicative markers for *Q. robur* and *Q. petraea*. AFLP was used in this study because this technique generates many markers per primer combination. Microsatellites could be used as well in order to find a diagnostic marker. Future studies could make use of the 49 microsatellite markers that are now available in order to detect a diagnostic microsatellite marker (Steinkellner et al. 1997; Kampfer et al. 1998). However, the microsatellite technique only generates one marker per PCR assay, which is more time consuming as compared to AFLP. Therefore it can be advised to continue with AFLP. As more microsatellites will be placed on the genetic linkage map in the OAKFLOW project, a selection could be made of microsatellite and/or AFLP markers that are closely linked to loci or clusters of loci on the genome that are different between *Q. robur* and *Q. petraea*. Information about loci or clusters of loci controlling the differences between *Q. robur* and *Q. petraea* will be obtained in the next few years from the OAKFLOW project.

As similar results as in chapter 2 of this thesis have been obtained for other mixed *Q. robur* and *Q. petraea* populations in Europe (Moreau et al. 1994; Bodénès et al. 1997a,b) it is expected that this high similarity between the two oak species can be observed in all indigenous mixed *Q. robur* and *Q. petraea* populations, including those in the Netherlands. Because of their high relatedness and hybridization/introgression taking place, *Q. robur* and *Q. petraea* are probably not pure species and therefore can better be called morphotypes (Baverstock & Moritz 1996), semispecies (Grant 1971), or ecotypes (Kleinschmit & Kleinschmit 2000). This indicates that for management and conservation purposes *Q. robur* and *Q. petraea* populations should be treated as a whole. The existence of the two 'species' together in the same area enables hybridization and introgression to take place, which will increase the genetic variation and therefore the adaptational abilities of the next generations.

As it has become clear from the study presented in chapter 2 of this thesis and the study by Bodénès et al. (1997a) that only a small part of the genome is involved in species differentiation, the search for a diagnostic marker could be done in a different approach. As there is an interest in obtaining diagnostic markers for certification of wood and planting material, future studies could be more focussed to the detection of a diagnostic marker. Therefore, future studies could attempt to use the bulked segregant analysis (BSA; Michelmore et al. 1991) method that has so far been used for populations obtained from a controlled cross (F2, BC1, etc.). BSA is a method that involves comparison of two pooled DNA samples of individuals. Within each pool the individuals are identical for the trait of interest. Subsequently, the two pools are analyzed for markers that distinguish them. Markers that are polymorphic between the two pools will be genetically linked to the trait of interest (based on which the pools were designed). To our knowledge BSA has so far not been applied to a population

consisting of unrelated individuals, like the indigenous population studied in chapter 2 of this thesis. To identify the genetic loci involved in the morphological differences between the two species BSA would only work if linkage disequilibrium is about 100%. For a population of unrelated individuals it is expected that a marker and a QTL are separated by much more recombinations as compared to a population obtained from a single cross. This means that for a population of unrelated individuals it is expected that BSA would only detect those markers that are closely linked to the QTL of interest. In such a case it might be better to use an approach called "selective genotyping". In this approach only the most extreme phenotypes from a large mapping population are used for BSA. As only the most extreme phenotypes are investigated, there is a higher chance to find markers that are closely linked to these traits of interest.

In chapter 3 of this thesis genetically identical individuals (clones) were detected and described for an ancient mixed *Q. robur* and *Q. petraea* woodland ("De Stompert") by means of microsatellites and AFLP. To our knowledge, this is the first time that clones of sizes larger than 3 m in diameter were described in ancient woodland that has been coppiced and grazed for centuries. The observed clone diameter (up to 5.8 m) was larger than was earlier expected based on visual observations of clusters of *Q. robur* and *Q. petraea* trees (Rackham 1980). The large clone size suggests that the genotypes are very old, which confirms the assumptions by Maes (1993) that we are dealing with ancient woodland, probably dating back about 1000 years ago when coppice culture started in the Netherlands. Within-clone polymorphisms (up to 5%) were detected by AFLP, but not by microsatellites. Although there are reports on up to 5% artifacts for AFLP (Huys et al. 1996; Jones et al. 1997; Arens et al. 1998; Winfield et al. 1998), it is very well possible that AFLP detected somatic mutations in these trees. In that case this would be another indication for the old age of the trees. However, in order to find out about the possibility of somatic mutations detected by AFLP additional studies need to be conducted involving additional DNA extractions from the same trees at different branches in the trees. In this way periclinal chimeras might be detected.

The observation of large clones in ancient *Q. robur* and *Q. petraea* woodlands will have implications for the future study of such populations not only in the Netherlands, but also in other European countries where coppice culture and animal grazing has been practiced in the past. When the woodland has been abandoned for many years after the last coppice cycle, the trees might have developed to such an extent that no indication for past coppice practice is left. In such a case the presence of clones in the population might be overlooked even when molecular markers are used. In that case the gene diversity for the population might be underestimated.

In chapter 4 of this thesis microsatellite analysis revealed that the genetic variation in two indigenous *Q. robur* populations was as high as was observed for unmanaged

French and German *Q. robur* populations (Streiff et al. 1998; Degen et al. 1999). This indicates that coppice practice, animal grazing, and reduction in population size due to deforestation did not result in a reduction of the genetic variation of the Dutch indigenous *Q. robur* populations. Population differentiation between the two indigenous *Q. robur* populations ("De Meinweg" and "De Stompert") was very low. As "De Meinweg" and "De Stompert" differed for their cpDNA haplotype this indicates that populations differing in cpDNA haplotype do not necessarily also differ for their nuclear DNA. The observed low population differentiation was remarkable as the two populations were separated at a distance of more than 100 km. This indicates that (based on theoretical calculations of population differentiation) pollen transfer can reach large distances as has earlier been suggested based on the small pollen grain diameter (reviewed by Ducousso et al. 1993) and paternity analysis (Streiff et al. 1999).

Although *Q. petraea* was not included in this study, it is expected that the genetic variation for this species will be as high as for *Q. robur* as both species are obligatory outcrossing, wind pollinated, long-lived tree species, which are life history traits that have a positive effect on the genetic variation of a species (Hamrick et al. 1979; Hamrick & Godt 1996). The only concern about a lower genetic variation within the *Q. petraea* populations in the Netherlands could be due to the fact that *Q. petraea* is not very common in the Netherlands. However, analysis of 33 *Q. petraea* trees with microsatellites in "De Meinweg" resulted in similar values for the gene diversity as was observed for *Q. robur* in "De Meinweg" and "De Stompert" (E.G. Bakker unpublished results). The same was observed for *Q. petraea* trees in an indigenous French population (Streiff et al. 1998).

The microsatellite results of "De Meinweg" could not be compared with the AFLP results for the same population (chapter 2), as different plots were investigated. The reason for this is the different research questions of the two studies. The AFLP study aimed to investigate *Q. robur* and *Q. petraea* trees in order to find a diagnostic marker to distinguish between the two species. In order to prevent inclusion of hybrids, *Q. robur* and *Q. petraea* trees were sampled in their ecological niches – it is expected that hybrids grow in areas with intermediate environmental characteristics between both niches (Rushton 1979; Van der Meijden 1990). In contrast, the microsatellite study aimed to investigate the total genetic variation present in *Q. robur*. As hybrids cannot be detected so far, this study included all trees with *Q. robur* traits that clustered together in a principal component analysis (PCA) based on leaf morphology traits. The plot was chosen in an area where *Q. robur* and *Q. petraea* occurred together. This makes that microsatellite and AFLP results could not be compared for the same individuals. Nevertheless, it appears that the AFLP results (Jaccard's (1908) genetic distance) reflect the microsatellite results (gene diversity according to Nei & Roychoudhury (1974)). For both *Q. robur* and *Q. petraea* the genetic variation was high.

Based on the results obtained for two indigenous *Q. robur* populations in the

Netherlands an indication of the genetic variation and structure of other indigenous *Q. robur* and *Q. petraea* populations in the Netherlands can be given. As many indigenous *Q. robur* and *Q. petraea* populations in the Netherlands have a similar history of coppicing and animal grazing, the genetic variation and structure of these populations are expected to be very similar to the genetic variation and structure as observed for "De Meinweg" and "De Stompert". If sprouting ability is a genetic factor, then genotypes that cannot sprout are expected to disappear from the population. However, so far – based on neutral markers – no evidence for such a selection pressure has been found, as the genetic variation observed for coppiced Dutch woodlands was as high as observed for a natural French population (Streiff et al. 1998). Moreover, due to long distance pollen flow seemingly isolated woodlands exchange their genetic material and therefore maintain their high level of genetic variation.

In the same chapter the gene diversity of a half-sib family harvested from a tree standing at a roadside in urban area was compared with the gene diversity of a half-sib family harvested from a tree standing in an autochthonous population ("De Meinweg"). As both half-sib families showed similar gene diversities this indicates that there was no difference between location (urban area or ancient woodland) for the genetic variation and the number of different father trees that contributed to the offspring.

In chapter 5 of this thesis QTLs for eco-physiological and morphological traits of *Q. robur* were detected. A new approach for the selection of a full-sib family proved to be successful. This method involved paternity analysis of a half-sib family. Although this approach was chosen in order to save time on controlled crossings, paternity analysis of a large half-sib family was time consuming. Another disadvantage is that not every parental combination can be realized, because two isolated trees are required. It appeared to be troublesome to find two completely isolated trees. Several days were spent looking for isolated trees in three major cities of the Netherlands (Amsterdam, Rotterdam, and Utrecht) and in a pine forest with hills where isolated oaks occurred ("Leuvenumse bos"). Isolated oaks were only found at a few locations in the visited cities and at one spot in the pine forest. All these oaks turned out to be *Q. robur* and were very similar to each other for their leaf morphology. Only at the location in Amsterdam the two *Q. robur* trees differed from each other for an important eco-physiological trait: abaxial pubescence. Based on this difference between the two trees in Amsterdam it was assumed that both trees would be sufficiently different for the construction of a genetic linkage map and the detection of QTLs. Another drawback of this approach for establishing a full-sib family is that many acorns need to be harvested, as the percentage of pollination from the neighboring tree is unknown. All half-sib seeds (about 1000) were sown and the next spring they needed to be analyzed for their microsatellite fingerprints as soon as possible in order to be able to quickly

transplant the full-sib seedlings to larger pots. As Streiff et al. (1999) report about low percentages of pollination between neighboring oaks and high percentages of pollination from oaks outside the studied plot (>100 m), it was expected that many seedlings should be tested in order to obtain a full-sib population of at least 100 individuals. Fortunately the percentage of pollination by the tree next to the mother tree was high: 26%. Two trees that were standing in the neighborhood of the mother tree (<10 m) contributed together about 4% of pollination. This means that still a percentage as high as 70% of the seedlings was the product of pollination from trees that were located at a distance larger than 100 m from the mother tree. This observation of long distance pollination for oaks in an urban area is similar to earlier observations for oaks in a forest where percentages as high as 73% for *Q. robur* and 81% for *Q. petraea* were observed (Streiff et al. 1999).

Maternal (496 cM, 13 linkage groups) and paternal linkage map (566 cM, 13 linkage groups) were constructed according to the double pseudo-testcross mapping strategy. The maternal and paternal linkage groups could be integrated using 58 'allelic bridge' markers. One remarkable linkage group was observed that contained 118 paternal AFLP markers(48% of the total number of 247 paternal markers) in coupling phase was observed. Karyotype analysis did not detect any additional chromosome or a B-chromosome (J.H. de Jong unpublished results).

The here presented genetic linkage map will be extended with more microsatellite markers in order to enable comparisons between this map and other maps (intra and inter-species maps for *Q. robur* and *Q. petraea*) that are being developed in the framework of the recently started (January 2001) OAKFLOW project. More eco-physiological traits (e.g. level of photosynthesis; drought tolerance; tannin content as indication for insect resistance) will be measured in this full-sib family and mapped during the OAKFLOW project.

None of the studies described in this thesis claims to be an ecological analysis, however, these studies will be a contribution to ecological research of indigenous *Q. robur* and *Q. petraea* populations. Besides, there are possibilities to apply the results of this thesis in forestry. The results described in chapter 5 are a first step towards the application of molecular markers in ecology and oak breeding: based on markers that are closely linked to QTLs on the genome, genetic variation and selection for a specific trait can be studied.

In this study major QTLs controlling for leaf width : sinus width ratio (LOD = 3.58), number of veins (LOD = 4.60), plant height (LOD = 4.30 and 4.64 for chromosomes 4 and 12, respectively), plant growth rate (LOD = 4.17), number of lobes ($p = 0.0001$), and flushing stadium ($p = 0.0001$) were observed. Besides these few highly significant QTLs several less significant QTLs were found. These QTLs controlled for leaf morphology (lamina length, length lamina to widest part, leaf length : leaf width ratio, petiole length, petiole length : (petiole length + lamina length) ratio),

length abaxial hairs on leaf, number of branches, double top, mildew resistance, leaf coloration in autumn, and number of old leaves staying in winter. Repetition of the observations on replicates of the full-sib population and comparison with other intra and interspecific *Q. robur* and *Q. petraea* maps will show which ones of those QTLs could be repeated and are therefore true QTLs. Only markers that are closely linked to a QTL may be used for ecology and forestry. Possible applications of QTLs are the analysis of genetic variation and selection of an eco-physiological trait like drought tolerance in a natural population. Because of information about heredity and genotype by environment interaction obtained from the genetic linkage map and the full-sib family, forestry and oak breeding can become more efficient.

Neutral molecular markers vs. coding molecular markers

Molecular markers like AFLP and microsatellites are assumed to be neutral, which means that they are not coding. A part of the AFLP and microsatellite markers could be located in or near coding genes and are therefore not neutral. Besides, AFLP and microsatellite markers located in non-coding areas of the genome are possibly not neutral when it appears that these non-coding areas are in fact coding or interact with coding regions. Therefore the term "neutral marker" should be used with caution (Moritz & Hillis 1996). Neutral markers are useful for population genetic analyses, as some population genetic analyses assume Hardy-Weinberg equilibrium (HWE). HWE only applies when certain requirements are fulfilled, among which the presence of random mating and absence of selection. These two requirements are also characteristics of a neutral marker. As there are many different requirements for HWE, absence of HWE cannot easily be attributed to the absence of one requirement for HWE. On the other hand does the presence of HWE not mean that all requirements are fulfilled, as opposite actions of several requirements can cancel each other out.

There is an interest in markers that are closely linked to coding genes or located in coding genes in order to study population genetics of those coding genes. However, markers that are found to be closely linked to coding genes in a mapping population are expected to be unlinked or very less linked in a natural population. This is because of the unrelatedness between the parents of a mapping population and the individuals in a natural population. Due to random crossing over events linkage disequilibrium between a marker and a QTL disappears gradually after many generations since the coancestry of oaks in the Netherlands. If we assume that for oaks in the Netherlands coancestry dates back to the end of the last ice age when the first migrants entered the Netherlands (8,000 years B.P.), then the number of generations can be estimated to be at least about 30. Close linkage between a marker and a QTL could possibly not yet

have been diminished in this number of generations. Therefore it is possible that some of the markers that have been found to be closely linked to certain QTLs can be used to do population genetic studies in natural populations.

AFLP vs. microsatellites

The use of different molecular marker techniques depends on the research question as well as on economical and practical issues of data collection. In this thesis two molecular marker techniques – AFLP and microsatellites – have been used. The most important difference between the two techniques is that AFLP is a dominant marker technique and the microsatellite technique is a co-dominant marker technique. As AFLP is a dominant marker technique, only the presence or absence of a marker band is scored. As no distinction can be made between homozygote dominant genotypes and heterozygote genotypes, a large part of the information generated by this molecular marker technique is not available. This fact has important implications for data analysis. Most population genetic analyses require co-dominant marker data. Although some population genetic software packages (Popgene; Yeh et al. 1997) claim to be able to handle dominant marker data, it is doubted if the results are not highly biased when the population size and number of markers are not sufficiently large (Lynch & Milligan 1994). Therefore many AFLP markers and many individuals are necessary in order to reduce the bias of the population genetic estimates. It is better to use a co-dominant marker technique like microsatellites, as such a technique generates data that can be analyzed for population genetic purposes without any problem. However, AFLP has two advantages over microsatellites: 1. AFLP generates 50-100 markers per electrophoresis as compared to only one marker per electrophoresis for microsatellites. 2. AFLP primers are readily available and can be applied to any species as opposed to microsatellites where the primers need to be developed for each species separately. Development of microsatellite primers can take up to six months. When microsatellite primers are not yet available for a particular species, it is advisable to start with AFLP analysis. However, by choosing for AFLP the number of options for population genetic analyses are limited unless the population size and the number of AFLP markers studied are large enough (Lynch & Milligan 1994). When AFLP is used for a small or moderate population size, then the data can only be analyzed for genetic distances between individuals. AFLP is therefore suitable for the detection of differences between individuals, groups of individuals, or species. When genetic distances between individuals are averaged per population, an indication will be obtained about the genetic variation present per population. In this thesis the AFLP technique was used for the analysis of species differentiation (chapter 2) and the description of clones (chapter 3). In both cases genetic distances between individuals were calculated using Jaccard's (1908) index. In this way the number of matching marker bands between two

individuals are divided by the total number of marker bands that are matching or different between these two individuals. Instead of Jaccard's (1908) index other descriptors for genetic distance can be used, like Nei & Li's (1979) genetic distance. Nei & Li's (1979) index for genetic distance is similar to Jaccard's (1908) index as both indices only differ in a linear way: Nei & Li's (1979) index weights the number of matching marker bands between two individuals. Another measure for genetic distance can be obtained by means of simple matching. For simple matching the number of matching marker bands between two individuals are divided by the total number of bands that are matching or different between these two individuals, including the cases where a band is absent in both individuals. Genetic distances between individuals as detected by AFLP can be graphically represented by means of a correspondence analysis (CA; Lebart et al. 1984). This multivariate statistical method is very similar to principal component analysis (PCA). The major difference between both methods is that CA is used for categorical data and PCA is used for quantitative data.

In chapter 2 it has been shown that AFLP is a suitable marker technique to detect species differentiation. Based on AFLP data the two oak species were separated in CA and species-indicative markers could be found. In chapter 3 it has become clear that clones can be identified unambiguously by means of microsatellite analysis, but that by means of AFLP analysis a few within-clone polymorphisms are detected. Although it is possible that these polymorphisms are caused by somatic mutations, they could also be attributed to artifacts (Huys et al. 1996; Jones et al. 1997; Arens et al. 1998; Winfield et al. 1998). About experimental errors two viewpoints prevail in literature at this moment. Those who wish to identify somaclonal variation or study cultivar identity for plant breeders rights (essentially derived variety issues) need technology to visualize differences below 1% of the genome up to much less than 1%. In that case a 1% experimental error would make AFLP results without value. In contrast, other studies that aim to confirm that two genotypes are identical, easily accept that a 1% experimental error should be a proper threshold that does not violate conclusions on the identity of two samples. Besides these two points of view, there could be an intermediate point of view that does not only look at quantitative differences between AFLP fingerprints in terms of numbers of polymorphisms. The quality of the differences should be regarded as well. In the study presented in chapter 3 of this thesis the presence of a new AFLP band as the basis of a polymorphism between trees within a clone was never observed. In all cases the within-clone polymorphic bands were shared by other clones. The observation of novel AFLP bands (or non-oak AFLPs) is essential to allow the conclusion of artifacts, whereas the observation of absence polymorphism would indicate mutations. The chance that an existing marker band disappears due to mutation is greater than the chance that a new marker band appears. Further studies should be conducted in order to find out about the possibility of somatic mutations in old *Q. robur* and *Q. petraea* clones detected by

AFLP. As Chen et al. (1999) could detect somatic mutations with AFLP in orchids, it should be possible to detect somatic mutations with AFLP in old *Q. robur* and *Q. petraea* clones.

A characteristic of the microsatellite technique is that this marker technique is highly susceptible to mutations. Microsatellites mutate at a rate of one nucleotide per 104 nucleotides, whereas other parts of the genome mutate at a rate of one nucleotide per 106 nucleotides (Kashi et al. 1997). For this reason, microsatellites should be used with caution, especially in the case of an analysis over several generations or when the individuals are of old age. Based on the observed large clone sizes, the age of the clones was assumed to be high (chapter 3). Mutations in these clones would be noticed as deviations in band length of the microsatellite loci. However, no such deviations were observed for all clones studied. As much less microsatellite markers were used as compared to AFLP markers, it is possible that the number of markers was too low to detect mutations. It is possible that the mutation rate is not high enough to result in visible mutations in the few microsatellite loci studied, even when the individuals are up to 1000-year old. This was also observed for microsatellite fingerprints of an enormous eelgrass clone in the Baltic Sea that was assumed to be 1000-year old (Reusch et al. 1999). Microsatellites were also very suitable for the analysis of genetic variation and structure in *Q. robur* populations. On the genetic linkage map of *Q. robur* no significant linkage was observed between a microsatellite locus and a trait of interest. Still, microsatellites could be located in coding regions as suggested by Kashi et al. (1997).

Multiple linear regression vs. genetic linkage map

Associations between molecular markers and phenotypic traits can be analyzed in two different ways: 1. By means of multiple linear regression analysis. 2. By means of QTL mapping on a genetic linkage map. The advantage of the first method is that the analysis can be done on data of any population, whereas the second method requires the construction of a genetic linkage map. Because of their height, incompatibility, and irregular seed production (mast years) controlled crossing of oak trees can be troublesome. The attempt used in chapter 5 to circumvent controlled crossing by doing paternity analysis of a half-sib family was successful, but still took a lot of time. Because of the long generation time it will take long before QTLs can be found for traits that are visible at later ages of the trees. Moreover, the construction of a genetic linkage map takes additional time and expenses. Therefore, when no genetic linkage map is available, multiple regression analysis is a good alternative to detect associations between markers and traits as has been shown by Virk et al. (1996). In this way QTL polymorphisms can be found for all possible polymorphic loci that are present in a population consisting of unrelated individuals. Whereas in a mapping population only

the polymorphic QTLs from the parents will contribute to the signal. The disadvantages of multiple linear regression analysis of AFLP markers on phenotypic traits are: 1. No information is available about the location of the markers on the genome. 2. For the multiple linear regression analysis as was used in chapter 2 the number of markers should not exceed the number of individuals studied. Therefore the number of markers that can be analyzed is limited. Although there are reports about genome wide searches of markers in human genetics, for association studies in human genetics only a few candidate genes are tested on sample sizes of up to a few thousand individuals (Risch 2000). Association studies are not suitable for genome wide searches in large, mixed populations because due to linkage disequilibrium over short distances tens of thousands of markers would be necessary to cover the whole genome which also increases the required significance level (Lander and Schork 1994).

As the parents of the full-sib family did not segregate for most of the markers that were found to be associated with a phenotypic trait after multiple regression analysis, it was not possible to compare most multiple regression analysis results with the genetic linkage map. This confirms that the two parents contain only a part of the total variation. Only a few markers that showed association with a phenotypic trait in the multiple linear regression analysis segregated in the full-sib family, but those markers did not show any linkage with a QTL on the genetic linkage map. This can be explained by the long coalescence time between the full-sib family and the natural population, which leads to disappearance of linkage disequilibrium. Besides, it should be mentioned that for multiple regression analysis chance associations can result from correlated allele frequencies in small samples (Virk et al. 1996).

Another approach could be tried in order to circumvent the troublesome crossing and the limited parent choice of the paternity analysis approach. It is possible to construct a genetic linkage map based on a half-sib population. QTL analysis will in that case be done based on methods developed in human genetics ("allele sharing"; Lander and Schork 1994).

Population genetic analyses

The number of studies in the field of population genetics increased rapidly after the introduction of allozymes. With these co-dominantly inherited markers all kinds of genetic phenomena can be studied (Murphy et al. 1996). Population genetic parameters were developed in order to analyze the allozyme data (Nei 1987). These parameters are nowadays used for the analysis of other molecular markers like microsatellites and AFLP. Although the microsatellite technique is a co-dominant marker technique like the allozyme technique, it differs from allozymes in the number of alleles per locus. Whereas allozymes exhibit maximally about five alleles per locus, microsatellites can reach numbers of more than 20 alleles per locus. This feature of

microsatellites has an effect on several statistics. The most important consequence of multiple allelism is that when the sample size is small the genotype table becomes sparse. Therefore the conventional chi-square test for Hardy-Weinberg proportions (HWP) cannot be applied, as it will lead to over- or underestimation of HWP (Guo & Thompson 1992). There are many computer programs available that can calculate population genetic statistics. However, they all differ in their calculation methods and assumptions. As these computer programs do not warn for erroneous analyses, care should be taken when selecting a population genetic computer program. For chapter 4 the computer package Genepop (Raymond & Rousset 1995) was used, as it was considered best for the data available. Genepop was compared with Popgene (Yeh et al. 1997) and Arlequin (Schneider et al. 2000). In the first place Popgene was chosen as it is the most user-friendly program among all. The program is Windows based and calculation of the population genetic statistics is fast. However, the test for HWE makes use of a chi-square test based on a table of genotypes. As for a relatively small data set of microsatellite data such a table is sparse, the test for HWE would be biased. Genepop and Arlequin both use the method as proposed by Guo & Thompson (1992) for testing HWE. Another point of difference between the three population genetic computer programs is the calculation of the *F*-statistics. Popgene calculates the *F*-statistics, but does not give significance levels. Genepop and Arlequin provide significance values for the *F*-statistics (Excoffier et al. 1992; Rousset & Raymond 1995). For the estimation of the gene diversity (Nei & Roychoudhury 1974) any computer program could be used, however, Popgene and Genepop do not provide the standard errors of the values. Arlequin uses the formulas as described by Nei & Roychoudhury (1974) for the unbiased estimate of gene diversity and its intra-locus variance.

Population structure was analyzed with another program, AutocorG 2.1 (O. Hardy and X. Vekemans, Laboratoire de Génétique et d'Ecologie Végétales, Université Libre de Bruxelles, Bruxelles, Belgium). AutocorG can calculate Moran's I (Sokal & Oden 1978) and the spatial coancestry coefficient (Loiselle et al. 1995). Both approaches are used for the analysis of population structure, but spatial coancestry analysis is more suitable as it is based on population genetics and it uses allelic information (Loiselle et al. 1995). Spatial autocorrelation (Moran's I) analysis is done per allele, after which the results are averaged to obtain the values per locus. Spatial coancestry analysis is done per locus. In this way no information is lost due to averaging over alleles. As both methods test for a significant clustering, family relationships that are not clustered will remain unnoticed. The common allele approach as described in chapter 4 appears to be able to show a family relationships to some extent when there is no clustering of related individuals. In this way the real connections and the cryptic gene flow between individuals becomes clear. However, as for the two studied autochthonous populations high levels of cryptic gene flow were observed, the

common allele approach could not detect a population structure in these populations.

Applications of molecular markers in ecology

The importance of the molecular ecology lies in the fact that molecular markers go beyond where the conventional ecology ends. Molecular markers can be successfully applied in ecology when information is available about close linkage with QTLs controlling for eco-physiological traits. Chapter 5 of this thesis is a first step towards this goal. So far, several eco-physiological and morphological traits have been investigated and for some (flushing stadium, leaf morphology (lamina length, length lamina to widest part, leaf length : leaf width ratio, leaf width : sinus width ratio, number of intercalary veins, petiole length, petiole length : (petiole length + lamina length) ratio; number of lobes), length abaxial hairs on leaf, plant height and growth rate, number of branches, double top, mildew resistance, leaf coloration in autumn, and number of old leaves staying in winter) QTLs were found. Few of these QTLs were found to be highly significant. These QTLs are expected to be true QTLs. The other QTLs might turn out to be true QTLs when they can be confirmed after repetition and on other genetic linkage maps. The application of marker-QTL associations in ecology and forestry is limited as the association found in the full-sib family is very likely not present in a natural population. Due to long coancestry, linkage has disappeared due to crossing over events. Only closely linked marker-QTL combinations could still be linked in a natural population.

When the full-sib family is vegetatively propagated (stem cuttings from sprouted branches) the plants can be planted in controlled experiments testing for the performance of the plants under different ecological circumstances (e.g. very wet or very dry). The full-sib family will be especially very valuable when there are several clones that grow up in different regions in the Netherlands, representing different ecological, environmental, and climatological conditions. In this way genotype by environment interaction can be tested during the life of the trees. Insight will be obtained about plasticity of traits and how traits change due to certain environmental factors.

In the other chapters molecular markers were used for genetic studies and the results were put in an ecological context. In this way molecular markers could be used for several purposes: 1. Species identification (chapter 2). 2. Clone identification (chapter 3). 3. Description of genetic variation and structure of populations (chapter 4; Van Dam & Bakker 2001). 4. Paternity analysis (chapter 5; Bakker & Van Dam 1999). The implications for ecology of these few applications are large.

Although *Q. robur* and *Q. petraea* could not be identified based on a diagnostic AFLP marker, the species-indicative AFLP markers can be a useful tool in combination with leaf morphology analysis for investigation of *Q. robur* and *Q. petraea* and their

hybrids in natural populations. The study presented in chapter 2 is a first step in the direction of understanding the genetic background of *Q. robur*, *Q. petraea*, and their hybrids. When diagnostic markers (or more species-indicative markers) become available in future studies, shifts in marker band frequencies can be studied over an environmental gradient where the two species and intermediate types occupy different ecological niches.

Paternity analysis will provide information about incompatibility between species and hybrids. Clone detection by means of molecular markers can be applied to other ancient woodlands in order to get an impression about clone size and therefore age of the trees. This especially is the case for ancient woodlands where coppice practice was abandoned more than 70 years ago, where clusters are no longer visible and the population looks like a normal forest. If one is not aware of the presence of clones in a population, population genetic studies might reveal too negative results about the genetic variation. Clone detection can also be used for the purpose of finding out which trees in a population are really old (clones) and which trees are possible later generations (single genotypes).

Gene diversity estimates for indigenous populations and half-sib families can be compared with observations about gene diversity for other populations. As genetic variation of *Q. robur* in two ancient woodlands ("De Meinweg" and "De Stompert") was as high as a French indigenous *Q. robur* population, it is expected that the genetic variation of other ancient woodlands in the Netherlans will be as high. Besides, populations could be compared for private alleles, which might indicate selection for specific alleles in some populations.

Applications of molecular markers in forestry and oak breeding

The results that are described in this thesis have besides an application in ecology an application in forestry and oak breeding. The five species-indicative AFLP markers will be a basis to continue the search for a diagnostic marker. However, based on several species-indicative AFLP markers it should be possible to devise a certification of the rare *Q. petraea* woods and seedlings. Only those trees that possess a marker band for any of the species-indicative AFLP markers can be certified to be *Q. petraea*. In the recently started OAKFLOW project more individuals and AFLP primer combinations will be analyzed in order to detect more species-indicative markers and to possibly find diagnostic markers.

As *Q. robur* and *Q. petraea* are no vegetatively propagated species, clone detection of these species will probably only be limited to ancient woodlands. An application of clone detection in oak forestry is for the establishment of seed orchards. Confusion about the location of clones in such seed orchards could be cleared up after

doing an analysis with molecular markers. Paternity analysis of seed orchards will give results about which trees pollinate each other. In this thesis there are indications that different parts of a single tree are pollinated by different trees (chapter 5). This would mean that in order to obtain a seed batch to which as many fathers as possible have contributed, seeds should be harvested from all sides of a tree. Analysis of the gene diversity of two half-sib families showed that the gene diversity of a half-sib family harvested from a tree at a roadside in urban area was about as high as observed for a half-sib harvested from a tree in an autochthonous population. This indicates that it does not matter for the genetic variation where to harvest the seeds.

The full-sib family and the genetic linkage map are a valuable source of information when maintained, vegetatively propagated, and transplanted in different environments. In that way during the next decades information can be gathered about traits important for forestry and oak breeding like tree height, diameter, flushing, branching, forking, epicormics, frost tolerance, etc. For some of these traits a close linkage with a molecular marker might be found on the genetic linkage map. Information about heritability and genotype by environment interaction obtained from the genetic linkage map and the full-sib family can be used for forestry and tree breeding. Besides, in future there might be an application of molecular markers that are closely linked to a trait of interest in order to select seed batches or to speed up tree selection.

Management of indigenous *Q. robur* and *Q. petraea* populations

Based on the results of molecular marker analysis of indigenous *Q. robur* and *Q. petraea* populations, several guidelines for management of these populations can be made.

1. Because of the close relatedness between *Q. robur* and *Q. petraea*, one can hardly speak of two different species. *Q. robur* and *Q. petraea* should be treated as two sympatric ecotypes that hybridize. Management of populations where both species occur should be directed towards a co-existence of both species in the same area. Due to hybridization and backcrossing the genetic variation will be maintained. As *Q. petraea* grows on drier sites, it is expected that this species carries genes that code for drought resistance. By means of hybridization and backcrossing these genes might spread over the *Q. robur* population as well. In this way the total population is expected to be better prepared to cope with a dryer and warmer climate in the future. Management of *Q. robur* populations should be directed towards the introduction of indigenous *Q. petraea* in the population in order to enable inter-specific matings.
2. Management of ancient *Q. robur* and *Q. petraea* woodlands should be based on the average size of clones in these populations. A forest manager could choose to

preserve the clonal structures in order to maintain the cultural heritage. When the forest manager wants to do thinning, he might then decide to remove an entire clone. However, in order to maintain a high genetic variation within the population the forest manager should remove only a few individuals per clone. In this way all genetic material will be conserved.

3. The genetic variation within an indigenous oak population is high. This means that next generations will be highly variable when mating is random. Seeds harvested from autochthonous populations are therefore suitable to be used for reforestation. It should be mentioned that it is expected that a high percentage (70-80%) of the seedlings is the product of pollinations from (possibly non-indigenous) trees from outside the ancient woodland. In that case the seedlings are only for a part of their genetic material autochthonous. The effect of such an admixture of genetic material on adaptation has so far not been studied. By means of paternity analysis a good indication about the percentage of pollinations from outside the autochthonous population can be obtained. Besides, paternity analysis can tell for some of the seedlings which of the trees from inside the indigenous population contributed to the pollination.
4. If it is true that indigenous planting material is better adapted to the local circumstances, then natural rejuvenation or planting with local (indigenous) material is favored above planting with foreign material. The genetic variation of indigenous populations is high and it is expected that highly variable offspring will be generated in the next generations. If natural rejuvenation takes place on a large scale it is expected that due to natural selection of the highly variable offspring, next generations of forest will consist of well adapted trees that can cope with changing environmental and climatological conditions.
5. The full-sib family and the genetic linkage map provided information about the heredity of certain traits and about markers that can possibly be used to study or select for those traits in natural populations. In future years the full-sib family will be vegetatively multiplied and planted at different locations in the Netherlands in order to obtain information about genotype by environment interaction. When the trees grow up, more information about traits at older ages becomes available. All this information can be incorporated in the genetic linkage map and subsequently more closely linked markers to QTLs controlling for traits of interest will be found. This means that the full-sib family and the genetic linkage map are a valuable source of genetic information of oak that can be applied in ecology and forestry.

Summary

The description of the genetic composition of indigenous *Quercus robur* L. (pedunculate oak) and *Q. petraea* (Matt.) Liebl. (sessile oak) populations is of importance to gain insight in the history, structure, and quality of these populations. Results of population genetic analyses of indigenous *Q. robur* and *Q. petraea* populations can help forest managers decide about the best strategies in order to manage and conserve these populations.

Q. robur and *Q. petraea* are two related sympatric oak species that can hybridize in nature. Based on their leaf morphology *Q. robur* and *Q. petraea* can be differentiated, however, it is unclear if intermediate type trees are species-hybrids or part of the overlap of variation in leaf morphology between the two species. In chapter 2 of this thesis species differentiation between *Q. robur* and *Q. petraea* in an ancient mixed population in the south of the Netherlands was investigated with the AFLP technique in order to find diagnostic markers. Based on AFLP markers the same differentiation was found as based on leaf morphology data. This is a clear indication for association between several leaf morphology traits and a part of the AFLP markers, which was confirmed by regression analysis. There were 13 markers that showed significant frequency differences between the two species. Moreover, five species-indicative AFLP markers were found, which means that these AFLP markers were absent in one species and present at a high percentage in the other species. This is the first indication of the existence of species-indicative markers for *Q. robur* and *Q. petraea*. As no diagnostic marker has been found so far, this is an indication for frequent hybridization and introgression between the two species during the course of history. Therefore *Q. robur* and *Q. petraea* are likely to be interbreeding semi-species or ecotypes.

In the Netherlands indigenous *Q. robur* and *Q. petraea* populations can be found mainly as ancient woodlands that have been used for coppice practice and animal grazing. The structure of such populations, which is characterized by clusters of genetically identical trees, is different as compared to natural, unmanaged *Q. robur* and *Q. petraea* populations. The true clonal structure can easily be described by means of molecular marker analysis. In chapter 3 of this thesis the number, size, and spatial distribution of *Q. robur* and *Q. petraea* clones in an ancient mixed *Q. robur* and *Q. petraea* woodland in the center of the Netherlands were described with two molecular marker techniques (microsatellites and AFLP). Based on six microsatellite markers and 69 polymorphic AFLP markers 14 unique genotypes were detected in a plot that consisted of 80 trees. Clones were observed for both *Q. robur* and *Q. petraea*. The maximum distance between two trees that belonged to the same clone was 5.8 m for *Q. robur*. To our knowledge this is the largest *Q. robur* clone reported so far. The observed large clone size may indicate the assumed old age of the trees. Knowledge about the presence of clones in *Q. robur* and *Q. petraea* populations is of importance

when the genetic variation of such populations is studied. As clones can result in an underestimation of the genetic variation in a population, their existence should be known or anticipated. Management and conservation strategies should take account of these observations for strategies for thinning of populations including clones in order to maintain the genetic variation in the population.

A high degree of genetic diversity enhances the chances for adaptation to changing environmental conditions in future. In chapter 4 of this thesis the genetic diversity and structure of two indigenous *Q. robur* populations were investigated using the six earlier mentioned microsatellite markers. For the population in the center of the Netherlands, as well as for the population in the south, as high values for the gene diversity were observed as compared to a natural *Q. robur* population in France. Similar analyses for two half-sib families from a tree in the population in the south of the Netherlands and a tree in an urban area, respectively, revealed significantly lower values for the gene diversity. However, these gene diversity values were similar, indicating that the genetic variation in *Q. robur* half-sibs is independent of the location of the mother tree: forest or urban area. In the two indigenous *Q. robur* populations no clear population structure could be detected, which was expected for a tree species that is characterized by long distance pollen dispersal and natural rejuvenation only at large gaps in the forest. As the genetic diversity in ancient *Q. robur* woodlands is high we assume, therefore, that this should allow adaptation to changing environmental conditions in future.

The results about the genetic composition of indigenous *Q. robur* and *Q. petraea* populations can be applied to ecological research of these populations when molecular markers are used that are closely linked with eco-physiological traits. In chapter 5 of this thesis a genetic linkage map of *Q. robur* with map positions of QTLs controlling several eco-physiological and morphological traits is presented. This chapter also describes a new method that can be used in order to select a full-sib family. For tree species like *Q. robur* controlled crossing is troublesome because of the size of the trees, unknown incompatibility between trees, and fluctuating seed harvests because of mast years. The new method involves the selection of a full-sib family from a half-sib population by means of paternity analysis using the set of earlier mentioned six microsatellite markers. The maternal and paternal tree (of the full-sib family) were standing isolated in an urban area together with two other oak trees (distances <10 m). It turned out that the highest percentage of pollinations was from the paternal tree (26%), which pollinated mostly at the side of the maternal tree closest to the paternal tree. The two other *Q. robur* trees only contributed each 2% pollination to the maternal tree. Therefore 70% of the offspring originated from pollination by trees that were located at >100 m distance. Paternity analysis of 397 half-sib individuals resulted in the detection of in total 102 full-sib individuals. A genetic linkage map was constructed based on this full-sib family using six microsatellite markers and 343

segregating AFLP markers (24 *Eco*RI/*Mse*I and 22 *Pst*I/*Mse*I). Remarkably, an exceptionally dense linkage group consisting of 118 AFLP markers (48% of the total number of 247 paternal markers) in coupling phase was observed. QTLs (quantitative trait loci) were detected for several eco-physiological and morphological traits: flushing stadium, leaf morphology (lamina length, length lamina to widest part, leaf length : leaf width ratio, leaf width : sinus width ratio, number of intercalary veins, petiole length, petiole length : (petiole length + lamina length) ratio, number of lobes), plant height and growth rate, number of branches, double top, mildew resistance, leaf coloration in autumn, and number of old leaves staying in winter. This means that markers closely linked to these QTLs may be used to study population genetics of these eco-physiological and morphological traits. The full-sib population will – especially after vegetative reproduction and planting in different ecological circumstances – be a valuable resource for more genetic studies. Phenotypic information can be obtained of more eco-physiological traits that will appear at later stages in the development of the trees. Therefore the full-sib family and the genetic linkage map may be used for new research questions during the next few years.

Samenvatting

Het beschrijven van de genetische samenstelling van inheemse *Quercus robur* L. (zomereik) en *Q. petraea* Matt. (Liebl.) (wintereik) populaties is belangrijk om inzicht te kunnen krijgen in de geschiedenis, structuur en kwaliteit van deze populaties. Gegevens over de populatiegenetische achtergronden van inheemse *Q. robur* en *Q. petraea* populaties kunnen de bosbeheerder helpen beslissen over de strategieën om deze populaties op de beste manier te kunnen beheren en behouden.

Q. robur en *Q. petraea* zijn twee verwante sympatrische eikensoorten die in de natuur kunnen hybridiseren. Gebaseerd op hun bladmorfologische kenmerken kunnen *Q. robur* en *Q. petraea* van elkaar worden onderscheiden, maar het is onduidelijk of intermediaire types het gevolg zijn van interspecifieke hybridisatie of van de overlap in morfologische variatie tussen de twee soorten. In hoofdstuk 2 van dit proefschrift is de AFLP techniek gebruikt om diagnostische merkers te vinden die *Q. robur* en *Q. petraea* in een oude gemengde populatie in het zuiden van Nederland (Limburg) van elkaar zouden kunnen onderscheiden. Gebaseerd op AFLP merkers werd eenzelfde onderscheid gevonden tussen *Q. robur* en *Q. petraea* als op basis van bladmorfologische kenmerken. Dit is een duidelijke aanwijzing voor associatie tussen enkele van de bladmorfologische eigenschappen en een gedeelte van de AFLP merkers, dat bevestigd werd door middel van regressie analyse. Er zijn 13 AFLP merkers gevonden die significant in frequentie verschilden tussen de twee soorten. Bovendien werden er vijf soorts-indicerende merkers gevonden, hetgeen betekent dat vijf AFLP merkers afwezig waren in de ene soort en voor een hoog percentage aanwezig in de andere soort. Dit is de eerste waarneming van het aanwezig zijn van soorts-indicerende merkers voor *Q. robur* en *Q. petraea*. Maar aangezien er tot nu toe nog geen diagnostische merker is gevonden, duidt dit er waarschijnlijk op dat de twee soorten in de loop der tijd frequent hybridisatie en introgressie hebben ondergaan. In dat geval zouden *Q. robur* en *Q. petraea* met elkaar kruisende halfsoorten of ecotypes kunnen zijn.

In Nederland zijn de meeste inheemse *Q. robur* en *Q. petraea* populaties bewaard gebleven in de vorm van oude bosplaatsen die zijn gebruikt voor hakhout en het weiden van schapen en varkens. De structuur van zulke populaties, die is gekenmerkt door clusters van genetisch identieke bomen, is anders dan natuurlijke *Q. robur* en *Q. petraea* populaties die niet door de mens zijn beïnvloed. De werkelijke klonale structuur kan alleen worden vastgesteld door middel van moleculaire merker analyse. In hoofdstuk 3 van dit proefschrift werden twee moleculaire merker technieken (microsatellieten en AFLP) gebruikt om het aantal, de omvang en de ruimtelijke verdeling van *Q. robur* en *Q. petraea* klonen te beschrijven in een oude gemengde *Q. robur* en *Q. petraea* bosplaats in het midden van Nederland (Utrecht). Gebaseerd op zes microsatelliet merkers en 69 polymorfe AFLP merkers werden 14 unieke genotypen gevonden in een gebied dat bestond uit 80 bomen. Klonen werden gevonden voor

zowel *Q. robur* als *Q. petraea*. De maximale afstand tussen twee bomen die tot dezelfde kloon behoorden was 5.8 m voor *Q. robur*. Bij ons weten is dit tot nu toe de grootste *Q. robur* kloon die is waargenomen. De geobserveerde grote omvang van de klonen is een aanwijzing voor de vermeende hoge leeftijd van de bomen. Kennis over het optreden van klonen in *Q. robur* en *Q. petraea* populaties is van belang voor studies naar de genetische variatie van eikenpopulaties. Aangezien de aanwezigheid van klonen de genetische variatie van een populatie kan verlagen, is het nodig dat de aanwezigheid van zulke klonen in een populatie bekend is. Toekomstige beheersplannen zouden deze kennis mee moeten nemen bij het maken van beslissingen over het dunnen van populaties met daarin klonen om zodoende de aanwezige genetische variatie in de populatie te bewaren.

Een hoge genetische variatie vergroot de kans voor een populatie om zich in de toekomst aan te kunnen passen aan veranderende omgevingscondities. In hoofdstuk 4 van dit proefschrift werden de genetische variatie en structuur van twee inheemse *Q. robur* populaties onderzocht met behulp van de zes eerder genoemde microsatelliet merkers. Voor zowel de populatie in het midden van Nederland (Utrecht) als de populatie in het zuiden van Nederland (Limburg) werden even hoge waarden voor de genetische diversiteit waargenomen als in een natuurlijke *Q. robur* populatie in Frankrijk. Echter, overeenkomstige analyses voor twee half-sib families afkomstig van een boom uit de populatie in het zuiden van Nederland en respectievelijk afkomstig van een boom in stedelijk gebied (Amsterdam) lieten significant lagere waarden voor de genetische diversiteit zien. Maar deze waarden voor de genetische diversiteit waren gelijk, hetgeen laat zien dat de genetische variatie van *Q. robur* half-sib families onafhankelijk is van de standplaats van de moederboom: bos of stedelijk gebied. Geen duidelijke familie-structuur is gevonden in de twee inheemse *Q. robur* populaties, wat verwacht kan worden voor een boomsoort waar het pollen over lange afstanden wordt verspreid en natuurlijke verjonging alleen plaatsvindt op open plekken in het bos. Deze resultaten laten zien dat de genetische variatie in oude *Q. robur* bosplaatsen hoog genoeg is om voldoende genetische variatie in de nakomelingen te genereren zodat toekomstige aanpassingen van de populatie aan veranderende omgevingscondities mogelijk zijn.

De resultaten van de genetische component van inheemse *Q. robur* en *Q. petraea* populaties kunnen worden toegepast in ecologisch onderzoek van deze populaties als er moleculaire merkers gebruikt worden die sterk gekoppeld zijn met eco-fysiologische kenmerken. In hoofdstuk 5 van dit proefschrift wordt een genetische koppelingskaart gepresenteerd van *Q. robur* met daarop gekarteerde QTLs die eco-fysiologische en morfologische kenmerken controleren. In hetzelfde hoofdstuk wordt ook een nieuwe methode voorgesteld die gebuikt kan worden om tot de selectie van een full-sib populatie te komen. Voor boomsoorten als *Q. robur* is het maken van kruisingen problematisch vanwege de hoogte van de bomen, onbekendheid over incompatibiliteit

tussen bomen en fluctuerende zaadproductie (mastjaren). De nieuwe methode behelst de selectie van een full-sib familie uit een half-sib familie door middel van ouderschapsanalyse met behulp van de zes eerder genoemde microsatelliet merkers. Zowel de moederlijke als de vaderlijke boom (van de full-sib familie) stonden geïsoleerd in stedelijk gebied (Amsterdam) samen met nog twee andere *Q. robur* bomen (afstanden <10 m). Het bleek dat het hoogste percentage bestuivingen afkomstig was van de vaderboom (26%), die voor het grootste deel de kant van de moederboom aan de kant van de vaderboom bestoof. De twee andere *Q. robur* bomen droegen ieder slechts 2% bij aan de bestuiving van de moederboom. Daarom is 70% van het nakomelingschap van de moederboom afkomstig van bestuiving door bomen die verder dan 100 m weg stonden. Ouderschapsanalyse van 397 half-sib individuen resulteerde in de detectie van in totaal 102 full-sib individuen. Een genetische kaart werd van deze full-sib populatie gemaakt met behulp van zes microsatelliet merkers en 343 segregerende AFLP merkers (24 *Eco*RI/*Mse*I en 22 *Pst*I/*Mse*I). Opmerkelijk was de detectie van een uitzonderlijk merker-dichte koppelingsgroep die bestond uit 118 AFLP merkers (48% van het totaal aan 247 vaderlijke merkers) in dezelfde koppelingsfase. QTLs (quantitative trait loci) werden gevonden voor een aantal verschillende eco-fysiologische en morfologische kenmerken: uitloopstadium, blad morfologie (blad lengte, lengte blad tot het breedste gedeelte, blad lengte : blad breedte verhouding, blad breedte : sinus breedte verhouding, aantal tussennerven, bladsteel lengte, bladsteel lengte : (bladsteel lengte + blad lengte) verhouding, aantal lobben), plant hoogte en groei, aantal zijtakken, dubbele top, meeldauw resistentie, blad verkleuring in de herfst, en het aantal oude bladeren dat in de winter aan de boom blijft hangen. Dit betekent dat merkers die sterk gekoppeld zijn met deze QTLs zouden kunnen worden gebruikt om de populatiegenetica van deze eco-fysiologische en morfologische kenmerken te bestuderen. De full-sib populatie zal een waardevolle bron van informatie zijn – vooral na vegetatieve vermeerdering en planten in verschillende ecologische omstandigheden – voor het bestuderen van meer eco-fysiologische kenmerken die op latere stadia van de ontwikkeling van de bomen te zien zullen zijn. Daarom kan de komende jaren de genetische kaart continue worden uitgebreid met nieuwe QTLs als er voortdurend waarnemingen aan deze populatie worden gedaan.

References

Antolin, M.F. and Strobeck, C. (1985) The population genetics of somatic mutations in plants. Am. Nat. 126: 52-62.

Arens, P., Coops, H., Jansen, J., and Vosman, B. (1998) Molecular genetic analysis of black poplar (Populus nigra L.) along Dutch rivers. Mol. Ecol. 7: 11-18.

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Rep 35: 1439-1445.

Asker, S.E. and Jerling, L. (1992) Apomixis in plants. Boca Raton: CRC Press.

Ayele, M., Tefera, H., Assefa, K., and Nguyen, H. (1999) Genetic characterization of two *Eragrostis* species using AFLP and morphological traits. Hereditas 130: 33-40.

Bacilieri, R., Labbe, T., and Kremer, A. (1994) Intraspecific genetic structure in a mixed population of *Quercus petraea* (Matt.) Liebl. and *Q. robur* L. Heredity 73: 130-141.

Bacilieri, R., Ducousso, A., and Kremer, A. (1995) Genetic, morphological, ecological and phenological differentiation between *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. in a mixed stand of Northwest of France. Silv. Genet. 44: 1-10.

Bacilieri, R., Ducousso, A., and Kremer, A. (1996) Comparison of morphological characters and molecular markers for the analysis of hybridization in sessile and pedunculate oak. Ann. Sci. For. 53: 79-91.

Bakker, E.G., and Van Dam, B.C. (1999) Vaderschapsanalyse bij eik: eikenstruifmeel komt van ver. Nederlands Bosbouw Tijdschrift 71: 35-38.

Barreneche, T., Bahrman, N., and Kremer, A. (1996) Two dimensional gel electrophoresis confirms the low level of genetic differentiation between *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. For. Genet., 3, 89-92.

Barreneche, T., Bodénès, C., Lexer, C., Trontin, J.-F., Fluch, S., Streiff, R., Plomion, C., Roussel, G., Steinkellner, H., Burg, K., Favre, J.-M., Glössl, J., and Kremer, A. (1998) A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. Theor Appl Genet 97: 1090-1103.

Baverstock, P.R., and Moritz, C. (1996) Project design. In: . Hillis, D.M., Moritz, C., and Mable, B.K. (eds.) Molecular systematics, 2nd edn., pp. 17-27. Sinauer Associates, Massachusetts.

Bennett, M.D. and Leitch, I.J. (1995) Nuclear DNA amounts in Angiosperms. Annals of Botany 76: 113-176.

Berg, E.E. and Hamrick, J.L. (1994) Spatial and genetic structure of two sandhills oaks: *Quercus laevis* and *Quercus margaretta* (Fagaceae). Am. J. Bot. 81: 7-14.

Berg, E.E. and Hamrick, J.L. (1995) Fine-scale genetic structure of a turkey oak forest. Evolution 49: 110-120.

Bodénès, C., Joandet, S., Laigret, F., and Kremer, A. (1997a) Detection of genomic regions differentiating two closely related oak species *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. Heredity 78: 433-444.

Bodénès, C., Labbé, T., Pradère, S., and Kremer, A. (1997b) General vs. local differentiation between two closely related white oak species. Molecular Ecology 6: 713-724

Bradshaw, H.D. Jr. and Grattapaglia, D. (1994) QTL mapping in interspecific hybrids of forest trees. Forest Genetics 1: 191-196.

Buis, J. (1985) Historia forestis. Nederlands bosgeschiedenis (deel 1&2). Landbouwhogeschool Wageningen en HES Uitgevers B.V., Utrecht.

Buntjer, J.B. (2000) CROSS CHECKER: COMPUTER ASSISTED SCORING OF GENETIC AFLP DATA. Abstract Plant & Animal Genome VIII Conference, San Diego, CA, January 9-12, 2000. http://www.intl-pag.org/pag/8/abstracts/pag8664.html

Burstin, J. and Charcosset, A. (1997) Relationship between phenotypic and marker distances: theoretical and experimental investigations. Heredity 79: 477-483.

Camus, A. (1954) Les Chênes - Monographie due genre *Quercus*. Editions Le Chevalier, Paris.

Chen, X., Lim, S.H., Wong, S.M., Lee, Y.H., Kuo, J., Yam, T.W., and Lin, J.-J. (1999) Amplified fragment length polymorphism analysis of vandaceous orchids. Plant Sci. 141: 183-189.

Churchill, G.A. and Doerge, R.W. (1994) Empirical threshold values for quantitative trait mapping. Genetics 138: 963-971.

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: 597-603.

De Vries, S.M.G. and Van Dam, B.C. (1998) Selection programme of Oak in the Netherlands. In: K.C. Steiner (ed). Diversity and adaptation of oak species. Proceedings of the Second Meeting of Working Party 2.08.05. Genetics of Quercus, of the International Union of Forest Research Organizations. October 12-17, 1997. University Park (State College), Pennsylvania: 201-208

Dow, B.D., Ashley, M.V., and Howe, H.F. (1995) Characterization of highly variable (GA/CT)n microsatellites in the bur oak, *Quercus macrocarpa*. Theor Appl Genet 91: 137-141.

Dow, B.D. and Ashley, M.V. (1996) Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. Molecular Ecology 5: 615-627.

Ducousso, A., Michaud, H., and Lumaret, R. (1993) Reproduction and gene flow in the genus *Quercus* L. Ann. Sci. For. 50 (Suppl. 1): 91s-106s.

Dumolin-Lapègue, S., Demesure, B., Fineschi, S., Le Corre, V., and Petit, R.J. (1997) Phylogeographic structure of white oaks throughout the European continent. Genetics 146: 1475-1487.

El-Kassaby, Y.A. and Park, Y.S. (1993) Genetic variation and correlation in growth, biomass, and phenology of Douglas-fir diallel progeny at different spacings. Silvae Genetica 42: 289-297.

Ellstrand, N.C. and Roose, M.L. (1987) Patterns of genotypic diversity in clonal plant species. Am. J. Bot. 74: 123-131.

Epperson, B.K. (1992) Spatial structure of genetic variation within populations of forest trees. New Forests 6: 257-278.

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: 479-491.

Favre, J.M. and Arnould, M.F. (1996) Nuclear DNA content and interspecific ribosomal variation between *Q. robur* and *Q. petraea*. In: Kremer A, Mühs H (eds) Inter- and intra-specific variation in European oaks: evolutionary implications and practical consequences. European Commission, EUR 16717EN, Luxembourg, pp 41-51.

Favre, J.M. and Brown, S. (1996) A flow cytometric evaluation of the nuclear DNA content and GC percent in genomes of European oak species. Ann. Sci. For. 53: 915-917.

Feuillat, F., Dupouey, J.L., Sciama, D., and Keller, R. (1997) A new attempt at discrimination between Quercus petraea and Quercus robur based on wood anatomy. Canadian Journal of Forest Research 27: 343-351.

Flies, R. (1993) Opening address. Ann. Sci. For. 50 (Suppl 1):19-20.

Gardiner, A.S. (1970) Pedunculate and sessile oak (*Quercus robur* L. and *Quercus petraea* (Mattuschka) Liebl.). Forestry 43: 151-160.

Graham, B.F. and Bormann, F.H. (1966) Natural root grafts. Bot. Rev. 32: 255-292.

Grandjean, G. and Sigaud, P. (1987) Contribution à la taxonomie et à l'ecologie des chênes du Berry. Ann. Sci. For.: 44, 35-66.

Grant, V. (1971) Plant speciation. Columbia University Press, New York.

Grattapaglia, D. and Sederoff, R. (1994) Genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137: 1121-1137.

Grattapaglia, D., Bertolucci, F.L., and Sederoff, R.R. (1995) Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross strategy and RAPD markers. Theor. Appl. Genet. 90: 933-947.

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

Han, T.-H. (2001) Use of genetic markers in Alstroemeria. Thesis Wageningen University and Research Centre, Wageningen, The Netherlands.

Hamrick, J.L., Linhart, Y.B., and Mitton, J.B. (1979) Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. Ann. Rev. Ecol. Syst. 10: 173-200.

Hamrick, J.L. and Godt, M.J.W. (1996) Effects of life history traits on genetic diversity in plant species. Phil. Trans. R. Soc. Lond. B 351: 1291-1298.

Hartl, D.L. and Clark, A.G. (1997) Principles of population genetics. Third edition. Sinauer Associates, Inc. Publishers Sunderland, Massachussetts.

Hillis, D.M., Moritz, C., and Mable, B.K. (eds.) Molecular systematics, 2nd edn.. Sinauer Associates, Massachusetts.

Howard, D.J., Preszler, R.W., Williams, J., Fenchel, S., and Broecklen, W.J. (1997) How discrete are oak species? Insights from a hybrid zone between *Quercus grisea* and *Quercus gambelii*. Evolution 51: 747-755.

Huntley, B. and Birks, H.J.B. (1983) An atlas of past and present pollen maps for Europe, 0-13,000 years ago. Cambridge University Press, Cambridge, UK.

Huys, G., Coopman, R., Janssen, P., and Kersters, K. (1996) High resolution genotyping analysis of the genus Aeromonas by AFLP fingerprinting. Intern. J. Syst. Bact. 46: 572-580.

Ietswaart, J.H. and Feij, A.E. (1989) A multivariate analysis of introgression between *Quercus robur* and *Quercus petraea* in the Netherlands. Acta Bot. Neerl. 38: 313-325.

Igboanugo, A.B.I. (1992) Characteristics of leaf structure and photosynthetic apparatus within the crown of systematically shaded Quercus petraea and Nothofagus procera seedlings. Biologia Plantarum 34: 401-407.

Jaccard, P. (1908) Nouvelles recherches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. 44, 223-270.

Jansen, R.C. (1993) Interval mapping of multiple quantitative trait loci. Genetics 135: 205-211.

Jarman, A.P., and Wells, R.A. (1989) Hypervariable minisatellites: recombinators or innocent bystanders? Trends in Genetics 5: 367-371.

Jensen, J.S., Wellendorf, H., Jager, K., De Vries, S.M.G. and Jensen, V. (1997) Analysis of a 17-year old open-pollinated progeny trial with *Quercus robur* L. Forest Genetics 4: 139-147.

Jones, E.W. (1959) Biological flora of the British Isles: Quercus L. Journal of Ecology 47: 169-216.

Jones, C. J., Edwards, K. J., Castiglione, S., Winfield, M. O., Sala, F., Wiel, C. Van de, Bredemeijer, G., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschi, A., Marmiroli, N., Aert, R., Volckaert, G., Rueda, J., Linacero, R., Vazquez, A., and Karp, A. (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories, Mol. Breeding 3, 381-390.

Kampfer, S., Lexer, C., Glössl, J., and Steinkellner, H. (1998) Brief report characterization of (GA)n microsatellite loci from Quercus robur. Hereditas 129: 183-186.

Kardolus J.P., Van Eck, H.J. and Van den Berg, R.G. (1998) The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy. Plant Syst. Evol. 210: 87-103.

Kashi, Y., King, D. and Soller, M. (1997) Simple sequence repeats as a source of quantitative genetic variation. Trends in Genetics 13: 74-78.

Kiers, A.M., Mes, T.H.M., Van der Meijden, R., and Bachmann, K. (2000) A search for diagnostic AFLP markers in Cichorium species with emphasis on endive and chicory cultivar groups. Genome 43: 470-476.

Kleinschmit, J.G.R., Bacilieri, R., Kremer, A., and Roloff, A. (1995) Comparison of morphological and genetic traits of pedunculate oak (*Q. robur* L.) and sessile oak (*Q. petraea* (Matt.) Liebl.). Silv. Genet. 44: 256-269.

Kleinschmit, J. and Kleinschmit, J.G.R. (2000) *Quercus robur - Quercus petraea*: a critical review of the species concept. Glasn. Sum. Pok. 37: 441-452.

Konig, A.O., Van Dam, B.C., Csaikl, U.M., Coart, E., Burg, K., Degen, B., De Vries, S.M.G., Petit, R.J., Ziegenhagen, B. (*in press*) Chloroplast DNA variation in oak in western central Europe and the genetic consequences of human influences. Forest Ecology and Management.

Koop, H. (1987) Vegetative reproduction of trees in some European natural forests. Vegetatio 72: 103-110.

Kosambi, D.D. (1944) The estimation of map distance from recombination values. Ann. Eugen. 12: 172-175.

Kremer, A., Petit, R., Zanetto, A., Fougère, V., Ducousso, A., Wagner, D., and Chauvin, C. (1991) Nuclear and organelle gene diversity in *Quercus robur* and *Q. petraea*. In: Genetic variation in European populations of forest trees. G. Müller-Starck and M. Ziehe (eds.). Sauerländer's Verlag, Frankfurt am Main.

Kruglyak, L. and Lander, E.S. (1995) A nonparametric approach for mapping quantitative trait loci. Genetics 139: 1421-1428.

Lambeth, C. (1980) Juvenile-mature correlations in Pinaceae and implications for early selection. Forest Science 26: 571-580.

Lander, E.S. and Botstein, D. (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185-199.

Lander, E.S. and Schork, N.J. (1994) Genetic dissection of complex traits. Science 265: 2037-2048.

Lebart, L., Morineau, A., and Warwick, K.M. (1984) Multiple descriptive statistical analysis: correspondence analysis and related techniques for large matrices. New York, Wiley.

Levene, H. (1949) On a matching problem arising in genetics. Annals of Mathematical Statistics 20: 91-94.

Lexer, C., Heinze, B., Steinkellner, H., Kampfer, S., Ziegenhagen, B., and Glössl, J. (1999) Microsatellite analysis of maternal half-sib families of *Quercus robur*, pedunculate oak: detection of seed contaminations and inference of the seed parents from the offspring. Theor. Appl. Genet. 99: 185-191.

Lexer, C., Heinze, B., Gerber, S., Macalka-Kampfer, S., Steinkellner, H., Kremer, A., and Glössl, J. (2000) Microsatellite analysis of maternal half-sib families of *Quercus robur*, pedunculate oak: II. Inferring the number of pollen donors from the offspring. Theor. Appl. Genet. 100: 858-865.

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: 1420-1425.

Lu, J., Knox, M.R., Ambroxe, M.J., Brown, J.K.M., and Ellis, T.H.N. (1996) Comparative analysis of genetic diversity in pea assessed by RFLP- and PCR-based methods. Theor. Appl. Genet. 93: 1103-1111.

Lumaret, R., Yacine, A., Berrod, A., Romane, F., and Li, T.X. (1991) Mating system and genetic diversity in holm oak (*Quercus ilex* L. Fagaceae). In: (eds Fineschi, S., Malvoti, M.E., Cannata, F., Hattemer, H.H.) Biochemical markers in the population genetics of forest trees. SPB Academic Publishers, The Hague. Pp. 149-153.

Lynch, M., and Milligan, B.G. (1994) Analysis of population genetic structure with RAPD markers. Molecular Ecology 3: 91-99.

Mace, E.S., Gebhardt, C.G., and Lester, R.N. (1999) AFLP analysis of genetic relationships in the tribe Daturae (Solanaceae). Theor. Appl. Genet. 99: 634-641.

Maes, N.C.M. (1993) Genetische kwaliteit van inheemse bomen en struiken. Deelproject: Randvoorwaarden en knelpunten bij behoud en toepassing van inheems genenmateriaal. IKC-NBLF en IBN-DLO, Wageningen. Rapport nr. 020.

Maliepaard, C., Jansen, J., and Van Ooijen, J.W. (1997) Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. Genet. Res. Camb. 70: 237-250.

Michelmore, R.W., Paran, I., and Kesseli, R.V. (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA 88: 9828-9832.

Montalvo, A.M., Conard, S.G., Thompson Conkle, M., and Hodgskiss, P.D. (1997) Population structure, genetic diversity, and clone formation in *Quercus* chrysolepis (Fagaceae). Am. J. Bot. 84: 1553-1564.

Montgomery, D.C., and Peck, E.A. (1992) Introduction to linear regression analysis, 2nd edn. John Wiley and sons, Inc., New York.

Moraal, L.G. (1997) Eikenprachtkever, *Agrilus biguttatus*, en eikensterfte: een literatuurstudie over aantastingen, levenswijze en verspreiding. Instituut voor Bos- en Natuuronderzoek (IBN-DLO), Wageningen. IBN-rapport 320.

Moreau, F., Kleinschmit, J., and Kremer, A. (1994) Molecular differentiation between *Q. petraea* and *Q. robur* assessed by random amplified DNA fragments. Forest Genetics 1: 51-64.

Moritz, C. and Hillis, D.M. (1996) Molecular systematics: context and controversies. In: Hillis, D.M., Moritz, C., and Mable, B.K. (eds.). Molecular systematics. 2nd edition. Sunderland, MA, Sinauer Associates, Inc. pp. 1-13.

Muir, G., Fleming, C.C., and Schlötterer, C. (2000) Species status of hybridizing oaks. Nature 405: 1016.

Murphy, R.W., Sites, J.W.Jr., Buth, D.G., and Haufler, C.H. (1996) Proteins: Isozyme electrophoresis. In: Hillis, D.M., Moritz, C., and Mable, B.K. (eds.). Molecular systematics. 2nd edition. Sunderland, MA, Sinauer Associates, Inc. pp. 51-120.

Nabuurs, G.J., Kramer, K., and Mohren, G.M.J. (1997) Effecten van klimaatverandering op het Nederlandse bos en bosbeheer. Instituut voor Bos- en Natuuronderzoek, IBN-DLO. IBN-rapport 256.

Nei, M. and Roychoudhury, A.K. (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76: 379-390.

Nei, M. and Li, W.-H. (1979) Mathematical model for studying genetic variation in terms of restriction nucleases. Proc. Natl. Acad. Sci. USA 76: 5269-5273.

Nei, M. (1987) Molecular evolutionary genetics. New York, Columbia University Press.

Nienstaedt, H. (1984) Inheritance and correlations of frost injury, growth, flowering, and cone characteristics in white spruce, *Picea glauca* (Moench) Voss. Can J. For. Res. 15: 498-504.

O'Hanlon, P.C., Peakass, R., and Briese, D.T. (1999) Amplified fragment length polymorphism (AFLP) reveals introgression in weedy Onopordum thistles: hybridization and invasion. Mol. Ecol. 8: 1239-1246.

Oosterbaan, A. and Van Hees, A.F.M. (1989) Resultaten van een lichtingsproef in een Beuken-wintereikenbos. De Dorschkamp, Wageningen, Rapport, nr. 551.

Oosterbaan, A., Van den Berg, C.A., Maas, G.J., and Moraal, L.G. (2001) Eikensterfte in Nederland. Omvang en de rol van enkele achterliggende factoren. Alterra, Wageningen. Alterra-rapport 236.

Parks, J.C. and Werth, C.R. (1993). A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). Am. J. Bot. 80: 537-544.

Paul, A.D., Foster, G.S., Caldwell, T., and McRae, J. (1995) Trends in genetic and environmental parameters for height, diameter, and volume in a multilocation clonal study with Loblolly Pine. Forest Science 43: 87-98.

Petit, R.J., Kremer, A., and Wagner, D.B. (1993) Geographic structure of chloroplast DNA polymorphisms in European oaks. Theor. Appl. Genet. 87: 122-128.

Petit, R.J., Pineau, E., Demesure, B., Bacilieri, R., Ducousso, A., and Kremer, A. (1997) Chloroplast DNA footprints of postglacial recolonization by oak. Proc. Natl. Acad. Sci. USA 94:9996-10001.

Potts, B.M. and Jordan, G.J. (1994) Genetic variation in the juvenile leaf morphology of *Eucalyptus globulus* Labill. ssp. *globulus*. Forest Genetics 1: 81-95.

Rackham, O. (1980) Ancient woodland. Its history, vegetation and uses in England. London: Edward Arnold (Publishers) Ltd..

Raymond, M. and Rousset, F. (1995a) GENEPOP Version 1.2, a population genetics software for exact tests and ecumenicism. J. Hered. 86: 248-249.

Raymond, M. and Rousset, F. (1995b) An exact test for population differentiation. Evolution 49: 1280-1283.

Reusch, T.B.H., Bostrom, C., Stam, W.T., and Olsen, J.L. (1999a). An ancient eelgrass clone in the Baltic. Marine Ecol. Progr. Series 183, 301-304.

Reusch, T.B.H., Stam, W.T., and Olsen, J.L. (1999b). Microsatellite loci in eelgrass Zostera marina marked polymorphism within and among populations. Mol. Ecol. 8, 317-321.

Rice, W.R. (1989) Analyzing tables of statistical tests. Evolution 43: 223-225.

Risch, N.J. (2000) Searching for genetic determinants in the new millennium. Nature 405: 847-856.

Rouppe van der Voort, J.N.A.M., Van Zandvoort, P., Van Eck, H.J., Folkertsma, R.T., Hutten, R.C.B., Draaistra, J., Gommers, F.J., Jacobsen, E., Helder, J., and Bakker, J. (1997) Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol. Gen. Genet. 255: 438-447.

Rouppe van der Voort, J.N.A.M., Van Eck, H.J., Draaistra, J., Van Zandvoort, P.M., Jacobsen, E., and Bakker, J. (1998) An online catalogue of AFLP markers covering the potato genome. Molecular Breeding 4:73-77.).

Rousset, F. and Raymond, M. (1995) Testing heterozygote excess and deficiency. Genetics 140: 1413-1419.

Rushton, B.S. (1978) *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl.: a multivariate approach to the hybrid problem, 1. Data acquisition, analysis and interpretation. Watsonia 12: 81-101.

Rushton, B.S. (1979) *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl.: a multivariate approach to the hybrid problem, 2. The geographical distribution of population types. Watsonia 12: 209-224.

Samuel, R., Pinsker, W., and Ehrendorfer, F. (1995) Electrophoresis analysis of genetic variation within and between populations of *Quercus cerris*, *Q. pubescens*, *Q. petraea* and *Q. robur* (Fagaceae) from eastern Austria. Bot. Acta

108: 290-299.

Schneider, S., Roessli, D., and Excoffier, L. (2000) Arlequin ver. 2000. A software for population genetic data analysis. http://anthropologie.unige.ch/arlequin

Schwarz, O. (1964) *Quercus*. In: Tutin TG, Heywood VH, Burges NA, Valentine DH, Walters SM, Webb DA (eds) Flora Europea, vol. 1. Cambridge University Press, London, pp 61-64.

Smouse, P.E. and Chakraborty, H.H. (1986) The use of restriction fragment length polymorphisms in paternity analysis. American Journal of Human Genetics 38: 918-939.

Sokal, R.R. and Oden, N.L. (1978) Spatial autocorrelation in biology. 1. Methodology. Biological Journal of the Linnean Society 10: 199-228.

Stam, P. (1993) Construction of integrated linkage maps by means of a new computer package: JOINMAP. Plant J. 3: 739-744.

Stam, P. and Van Ooijen, J.W. (1995) JoinMap (tm) version 2.0: Software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen, The Netherlands.

Steinhoff, S. (1993) Results and species hybridization with *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. Ann. Sci. For. 50 (Suppl 1): 137s-143s.

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: 1093-1096.

Stephan, W. and Sho, S. (1994) Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. Genetics 136: 333-341.

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: 317-328.

Streiff, R., Ducousso, A., Lexer, C., Steinkellner, H., Gloessl, J., and Kremer, A. (1999) Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. Molecular Ecology 8: 831-841.

Takenaka, A. (1994) Effects of leaf blade narrowness and petiole length on the light capture efficiency of a shoot. Ecological Research 9: 109-114.

Van Dam, B.C. and De Vries, S.M.G. (1998) In de voetsporen van de eik, post-glaciale herkolonisatie routes. De Levende Natuur 99: 38-41.

Van Dam, B.C. and Bakker, E.G. (2001) Genetische diversiteit in gemengde populaties van Zomer- en Wintereiken. De Levende Natuur 102: 26-29.

Van de Peer, Y. and De Wachter, R. (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Computer Applic. Biosci. 10: 569-570.

Van der Meijden, R. (1990) Heukel's flora van Nederland, 21st edn. Wolters-Noordhoff, Groningen.

Van Iterson, W. (1932) De historische ontwikkeling van de rechten op de grond in de provincie Utrecht. Dl. 1. Band 2. Inleiding: Markeverhoudingen in het Overkwartier en Eemland. Thesis, Rijksuniversiteit Leiden. Leiden: Brill.

Van Ooijen, J.W. (1999) LOD significance thresholds for QTL analysis in experimental populations of diploid species. Heredity 83: 613-624.

Van Ooijen, J.W. (2000) MapQTL (r) Version 4.0: User friendly power in QTL mapping. Plant Research International, http://www.plant.wageningen-ur.nl/products/mapping/MapQTL/mqintro.htm.

Van Os, H., Buntjer, J., Stam, P., Van Eck, H.J. (2000) EVALUATION OF TWO ALGORITHMS FOR THE CONSTRUCTION OF GENETIC LINKAGE MAPS. Abstract Plant & Animal Genome VIII Conference, San Diego, CA, January 9-12, 2000. http://www.intl-pag.org/pag/8/abstracts/pag8621.html

Van Zeist, W. (1970) Chapters 5 (Schipborg) and 11 (Ommen). Paleohistoria 14, pp. 66 and 100.

Venner, G.H.A. (1985) De Meinweg: onderzoek naar de rechten op gemene gronden in het voormalige Gelders-Gulikse

grensgebied, circa 1400-1822. Maaslandse monografiën nr. 40. Van Gorcum, Assen.

Vera, F.W.M. (1997) Metaforen voor de wildernis. Eik, hazelaar, rund en paard. Ministerie van Landbouw, Natuurbeheer en Visserij, 's-Gravenhage.

Vicient, C.M., Suoniemi, A., Anamthawat-Jónsson, K., Tanskanen, J., Beharav, A., Nevo, E., and Schulman, A.H. (1999) Retrotransposon *BARE-1* and its role in genome evolution in the genus *Hordeum*. The Plant Cell 11: 1769-1784.

Virk, P.S., Ford-Lloyd, B.V., Jackson, M.T., Pooni, H.S., Clemeno, T.P., and Newbury, H.J. (1996) Predicting quantitative variation within rice germplasm using molecular markers. Heredity 76: 296-304.

Voorrips, R.E., 2001. MapChart version 2.0: Windows software for the graphical presentation of linkage maps and QTLs. Plant Research International, Wageningen, The Netherlands.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, Th., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995) AFLP: a new technique for DNA fingerprinting. Nucl. Acids Res. 23, 4407-4414.

Waugh, R., Bonar, N., Baird, E., Thomas, B., Graner, A., Hayes, P., and Powell, W. (1997) Homology of AFLP products in three mapping populations of barley. Mol. Gen. Genet. 255: 311-321

Weir, B.S. and Cockerham, C.C. (1984) Estimating F-statistics for the analysis of population structure. Evolution 38: 1358-1370.

Weir, B.S. (1996) Genetic Data Analysis II. Sinauer Associates, Inc. Massachusetts.

Westneat, D.F. and Webster, M.S. (1994) Molecular analysis of kinship in birds: interesting questions and useful techniques. In: Molecular Ecology and Evolution: Approaches and Applications (eds. Schierwater, B., Streit, B., Wagner, G.P., and DeSalle, R.), pp. 91-126. Birkhäuser Verlag, Basel.

Whitham, T.G. (1981) Individual trees as heterogeneous environments: adaptations to herbivory or epigenetic noise. In: Insect life history patterns: habitat and geographic variation (eds. Denno, R.F. and Dingle, R.F.) New York: Springer-Verlag, pp. 9-27.

Winfield, M.O., Arnold, G.M., Cooper, F., Le Ray, M., White, J., Karp, A., and Edwards, K.J. (1998) A study of genetic diversity in Populus nigra subsp. Betufolia in the Upper Severn area of the UK using AFLP markers. Mol. Ecol. 7: 3-10.

Whittemore, A.T., Schaal, B.A. (1991) Interspecific gene flow in sympatric oaks. Proc. Natl. Acad. Sci. USA 88: 2540-2544.

Yeh, F.C., Yang, R.-C., and Boyle, T. (1997) Popgene version 1.21. Microsoft Window-based Freeware for Population Genetic Analysis. http://www.ualberta.ca/~fyeh

Yin, X., Stam, P., Dourleijn, C.J., and Kropff, M.J. (1999) AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. Theor. Appl. Genet. 99: 244-253.

Zanetto, A., Roussel, G., and Kremer, A. (1994) Geographic variation of inter-specific differentiation between *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. Forest Genetics 1: 111-123.

Zanetto, A. and Kremer, A. (1995) Geographic structure of gene diversity in *Quercus petraea* (Matt.) Liebl. I. Monolocus patterns of variation. Heredity 75: 506-517.

Zoldos, V., Papes, D., Brown, S., Panaud, O., and Siljak-Yakovlev, S. (1998) Genome size and base composition of seven *Quercus* species: inter- and intra-population variation. Genome 41: 162-168.

Zoldos, V., Papes, D., Cerbah, M., Panaud, O., Besendorfer, V., and Siljak-Yakovlev, S. (1999) Molecular-cytogenetic studies of ribosomal genes and heterochromatin reveal conserved genome organization among 11 *Quercus* species. Theor Appl Genet 99: 969-977.