UNIVERSITEIT GENT

Faculteit Landbouwkundige en

Toegepaste Biologische Wetenschappen

Academiejaar 2002-2003

MOLECULAR CONTRIBUTIONS TO THE CONSERVATION OF FOREST GENETIC RESOURCES IN FLANDERS: GENETIC DIVERSITY OF *MALUS SYLVESTRIS*, *QUERCUS SPP*. AND *CARPINUS BETULUS*

MOLECULAIRE BIJDRAGEN TOT HET BEHOUD VAN GENETISCHE BRONNEN IN VLAAMSE BOSSEN: GENETISCHE DIVERSITEIT VAN *MALUS SYLVESTRIS*, *QUERCUS SPP*. EN *CARPINUS BETULUS*

door

ir. Elisabeth COART

Thesis submitted in fulfilment of the requirements for the degree
of Doctor (Ph.D.) in Applied Biological Sciences

Proefschrift voorgedragen tot het bekomen van de graad van
Doctor in de Toegepaste Landbouwkundige en Biologische Wetenschappen

Op gezag van
Rector: **Prof. dr. A. DE LEENHEER**

Decaan:
Prof. dr. ir. H. VAN LANGENHOVE

Promotor:
Prof. dr. ir. E. VAN BOCKSTAELE
Dr. I. ROLDAN-RUIZ

The author and the promoter give the authorization to consult and to copy parts of this work for personal use only. Any other use is limited by the Laws of Copyright. Permission to reproduce any material contained in this work should be obtained from the author.

Auteur en promotor geven toelating dit doctoraatswerk voor consultatie beschikbaar te stellen en delen ervan te kopiëren voor persoonlijk gebruik. Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot uitdrukkelijk de bron te vermelden bij het aanhalen van resultaten uit dit werk.

The promoter,

Prof. dr. ir. Erik Van Bockstaele

The author,

ir. Elisabeth Coart

Preface – Woord vooraf

Als pas afgestudeerd ‘groentje’ ben ik beginnen werken op het Instituut voor Bosbouw en Wildbeheer. Daar is mijn interesse in genetische diversiteit begonnen. Het waren vooral de interesse en gedrevenheid van Jos Van Slycken die mij aan het denken zetten. Hij heeft mij ook de vrijheid gegeven om mij in de materie in te werken en mijn eigen weg te zoeken. In onze discussies samen met Danny Maddelein van de afdeling Bos en Groen kreeg de aanpak van de problematiek rond genetische diversiteit van autochtone bomen en struiken stilaan concrete vorm. Ik weet dat dit onderzoek geen antwoord biedt op al die vragen die we toen gesteld hebben maar ik hoop dat het toch een goede aanzet is om verder de weg uit te stippelen voor het behoud van autochtone bomen en struiken. Dankjewel Jos en Danny voor jullie vele constructieve bijdragen en steun!

Toegekomen op het Departement Plantengenetica en –verdeling kwam ik terecht in een goed geoliede organisatie: het labo biotech. Ik was verwonderd over de inzet en expertise die er in de groep aanwezig waren en meer nog, de openheid waarmee de kennis zomaar gedeeld werd. Iedereen in het labo heeft zo z’n specialiteit en bleek telkens weer bereid om mij daarmee te helpen. De betrokkenheid bij het werk is groot maar ook de onderlinge betrokkenheid tussen collega’s mag er zijn. Ik heb met veel plezier in deze groep, geleid door Marc De Loose, gewerkt en wil jullie allemaal heel erg bedanken voor de leuke jaren (waar we er hopelijk nog een aantal van te goed hebben!).

Er zijn een paar collega’s die ik nog extra in de bloemen wil zetten. Om te beginnen: Nancy Mergan. Zij is er in geslaagd om mijn chaotische geest en bruuske bewegingen compatibel te maken met het fijne labowerk. Stap voor stap heeft ze me geduldig alles geleerd in ’t labo, niet alleen praktisch maar ook hoe je er in slaagt om een goede ‘boekhouding’ bij te houden zodat je op het einde van de analyses de vele staaltjes nog uit elkaar kan houden. En wat een motivatie om al die stalen te verwerken, 1000 AFLP’s lopen is immers sneller gezegd dan gedaan maar bij Nancy konden er altijd nog wel een paar bij.

En dan kwam Katrien Liebaut erbij, die al snel geconfronteerd werd met het eiken-DNA dat zich niet zo graag zuiver laat extraheren. Al die foto’s van lege gels waar zoveel tijd in zat, altijd opnieuw proberen. Maar dankzij haar doorzettingsvermogen zijn er toch mooie resultaten uitgekomen.

Sabine Van Glabeke begon als studente op ’t labo. Al snel had ik door dat het een studente uit de duizend was: interesse, gedrevenheid, inzicht. Het was een plezier om samen aan haar thesis te werken. Het deed me deugd dat haar inzet kon ‘beloond’ worden met een contract op het DvP en dat heeft niemand zich beklaagd.

En dan is er Isabel Roldán-Ruiz. Aan haar heb ik zoveel te danken dat ik niet weet waar te beginnen. Plannen van experimenten, inzamelen van materiaal, organisatie in ’t labo,

scoren van gels, dataverwerking, op alles wist ze raad. Na een altijd even constructieve als heldere discussie met Isabel had ik weer het gevoel dat ik op de goede weg was.

Ook de bureaugenootjes waren een echte steun: Veerle Lamote, Hilde Muylle, Inge Van Daele, Ellen De Keyser, Isabelle De Grieck en Wendy Aartsen. Het is er gezellig toekomen en zelfs met het tekort aan computers was er altijd veel begrip voor elkaars werk en situatie.

En natuurlijk mag mijn promotor Erik Van Bockstaele niet ontbreken. Onder zijn leiding bloeit het DvP, zijn gedrevenheid motiveert om er ook zelf voor te gaan.

Het was en is me een waar genoegen om met jullie allemaal samen te werken!

Ook buiten het DvP zijn er nog vele mensen die veel hebben bijgedragen aan dit onderzoek, door discussies en samenwerking: Xavier Vekemans die tijd vrijmaakte om mij mijn weg te helpen vinden in de dataverwerking, René Smulders die in onze samenwerking graag z'n ruime expertise ter beschikking stelde, Bert Maes en Chris Rövenkamp met hun uitgebreide terreinkennis van autochtone bomen en struiken, Paul Vandecasteele met z'n schitterend programma om snel bladparameters op te meten, Beatrijs Van der Aa en Bart Vandecasteele voor de verspreidingskaartjes, iedereen die mij materiaal toestuurde of samen materiaal inzamelde, de studenten die een groot deel van het labowerk op zich genomen hebben, …

Het is voor mij heel duidelijk: onderzoek doe je niet alleen. Het is dankzij het team op het DvP en de stimulerende samenwerking met andere onderzoeksinstellingen dat het allemaal de moeite waard is.

Tenslotte wil ik ook familie, vriend en vrienden danken voor hun steun. Ik hoop dat ik genoeg heb laten blijken wat jullie voor mij betekenen. Net zoals vele andere dingen was dit doctoraat zonder jullie steun niet mogelijk geweest!

Acronyms and abbreviations

AFLP	Amplified Fragment Length Polymorphism
APS	Ammonium Per Sulphate
ATP	Adenosine Tri Phosphate
bp	base pairs
BSA	Bovine Serum Albumine
CLO	Centrum voor Landbouwkundig Onderzoek / Agricultural Research Centre
cp	chloroplast
CRA	Centre de Recherches Agronomiques
CTAB	Cetyl Trimethyl Ammonium Bromide
DNA	Deoxyribo Nucleic Acid
DvP	Departement Plantengenetica en –veredeling / Department Plant Genetics and Breeding
dNTP	deoxyribo Nucleoside Tri Phosphate
EDTA	Ethyleen Diamine Tetra Acetic acid
ESU	Evolutionary Significant Unit
EUFORGEN	European Forest Genetic Resources program
IAM	Infinite Allele Model
ISSR	Inter Simple Sequence Repeat
NIR	Near Infra Red
NTSYSpc	Numerical Taxonomy and multivariate analysis SYStem
*mat*K	maturase K
PAGE	polyacrylamide gel electroforesis
PCO	Principal COordinates analysis
PCR	Polymerase Chain Reaction
PRI	Plant Research International
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
SMM	Stepwise Mutation Model
SPAGeDi	Spatial Pattern Analysis of Genetic Diversity
SPSS	Statistical Package for the Social Sciences
SSR	Simple Sequence Repeat
Taq	*Thermus aquaticus*
TBE	Tris Borate EDTA

Abstract

As a result of widespread deforestation and overexploitation, many tree species have become the focus of growing conservation concern. At the Flemish scale, recent protection efforts have concentrated on the conservation of autochthonous populations. A systematic inventory of autochthonous trees and shrubs, although not yet completed, has revealed that autochthonous populations of most woody species are very rare in Flanders. The research presented in this thesis consists of a molecular approach to identify and study autochthonous populations and the present population genetic structure in tree species in order to support the set up of sound conservation programs. This information was necessary if these relics are to be used as seed sources for future forestations, what is seen as an important part of their conservation. Species with very different histories and contemporary situations were selected for this pilot study.

Wild apple ***(Malus sylvestris)***, the apple species native to western and central Europe, is one of the most endangered tree species in Flanders and the remaining individuals are very scattered. Therefore, the construction of a gene bank that can be used as a new interbreeding population is seen as a necessary step for its future conservation. However, this implies the discrimination between 'genuine' wild and cultivated apple trees, what is not straightforward because of the presence of phenotypic intermediates in the forests. The molecular markers applied in this thesis provided a valuable approach to ascertain the identity of putative wild genotypes. Both AFLP and microsatellite markers revealed a very clear differentiation among the wild gene pool, edible and ornamental apple cultivars, despite the fact that individuals derived from edible cultivars were found in the wild. A very low admixture between the wild and edible cultivar group was detected, with only three genotypes (or 4%) being identified as hybrids. The degree of hairiness of leaves was shown to be of use as a first indication of the origin of an apple tree but the resolution of assignment of individuals to the wild and/or cultivated gene pool reached by molecular markers is much higher.

The chloroplast marker typed (duplication in *mat*K region) resulted in contrasting data and showed possible evidence of introgression with edible cultivars for up to 39% of Belgian wild apples. No introgression was detected in the German wild gene pool. However, hybridisation might have been underestimated since the *mat*K-marker was not present in all edible cultivars. Most (nuclear) genetic variation was present within the locations sampled but a significant differentiation was detected within and between Belgian and German origins. Furthermore, the discovery of different chloroplast types in Belgian populations implies the different history of Belgian and German wild apple trees. Based on the results of this study, it can be advised to create a regional gene bank including all Belgian trees with

nuclear fingerprints typical of *M. sylvestris*. Further decisions on the choice of genotypes to include in a future seed orchard can be made soon when results on more Belgian genotypes will become available.

In contrast to wild apple, the indigenous oak species ***Quercus petraea*** (sessile oak) and ***Quercus robur*** (pedunculate oak) are very common in Flanders. Conversely this is not true for autochthonous populations of these often planted species of high economic importance. In this study it was attempted to evaluate the autochthonous character of Flemish oak populations based on their chloroplast-DNA diversity. The detected patterns of chloroplast DNA variation in putative (as determined by field-evaluation) autochthonous populations reflected the original post-glacial distribution of haplotypes and proved helpful to determine the origin of oak populations. The congruence between the field assessment and the chloroplast data emphasises the applicability of both evaluation methods.

As is expected for outcrossing species, only a weak population genetic structure was observed for both oak taxa on a Flemish scale when studied with six (nuclear) microsatellite markers. Within-population diversities were high for autochthonous populations and were not significantly different between the two species. Diversity estimates for the relatively small Flemish forests are comparable to those found for large oak forests in Europe and no evidence for recent bottlenecks was found in the autochthonous relics. A small but significant excess of homozygotes has been found in many populations of both species. No significant genetic differentiation between autochthonous populations and selected provenances was observed within both species. For both taxa, selected stands exhibited slightly lower within-population diversities than autochthonous populations and had higher heterozygote deficiencies but these differences were not significant. This study revealed that in order to safeguard the genetic variability in future oak forests, there is no reason to choose selected provenances over autochthonous populations as seed sources. In view of the established population genetic structure of Flemish oaks, the delineation of conservation units on the sole basis of chloroplast DNA polymorphisms may not reflect the enormous levels of diversity present nowadays in oak populations and would ignore the possible local adaptations that might have arisen after recolonisation.

This study also provided information on the differentiation between *Q. robur* and *Q. petraea*. AFLP markers showed a clear differentiation between both oak taxa and were able to assign an individual oak tree to a species, but no species-specific marker was found. Assignment based on AFLP markers and morphological parameters were congruent. The application of six microsatellite loci could to some extent also differentiate sessile and pedunculate oak populations but more loci should be studied to reach the resolution of

AFLP markers. These results are in accordance with the knowledge that both oak species are only differentiated for a few loci. If required, AFLP markers could be applied to determine the taxonomic oak species present in a forest or nursery. Taking into account that the hybridisation between both oak species is an occasional and natural occurring phenomenon, there is no reason to treat populations that include *Q. x rosaceae* individuals differently in conservation issues.

Finally, a study based on AFLP markers was conducted in order to define the present population genetic structure in ***Carpinus betulus*** (hornbeam). In accordance with the wind-pollinated and outcrossing breeding system, high within-population diversity and little (but significant) genetic differentiation were detected at Flemish and European scales. However, only on a European scale the genetic structure was correlated with the geographic structure. The weak geographic pattern that was possibly present on a Flemish scale in ancient times might have been disrupted by the translocation of individuals over limited distances to create the many line plantations in Flanders. At European scale within-population genetic diversities were shown to be significantly correlated with the distance from the glacial refugium, suggesting that gene diversity has increased during postglacial recolonisation, despite the bottleneck at the outset of recolonisation. However, more research including co-dominant markers would be necessary to study the current patterns of genetic diversity into greater detail. The detection on a European scale of a geographic structure of nuclear diversity patterns suggests that possibly adaptive changes might exist between hornbeam populations from distant locations. Furthermore, the within-population variation was shown to be lower in Flemish than in European populations. It is therefore advisable for future plantations (i) to join seeds harvested on different locations and (ii) to use reproductive material from own or nearby provenances.

Samenvatting

In het kader van het behoud van bomen werd de aandacht recent gericht op autochtone genenbronnen. In Vlaanderen werd een systematische inventarisatie van autochtone bomen en struiken gestart in 1997. Hoewel deze inventarisatie nog niet werd afgerond, is het reeds duidelijk dat autochtone populaties van de meeste houtige soorten zeldzaam zijn. Het onderzoek in deze thesis stelt een moleculaire aanpak voor om autochtone populaties te identificeren en om de populatiegenetische structuur van boomsoorten te bepalen. Deze informatie was essentieel, te meer omdat het gebruik van autochtone populaties als zaadbronnen voor toekomstige bebossingen gezien wordt als een belangrijk onderdeel van de behoudsmaatregelen. Voor deze pilootstudie werden soorten uitgekozen met een totaal verschillende geschiedenis en huidige situatie.

Wilde appel ***(Malus sylvestris)*** is, als inheemse soort in west en centraal Europa, één van de meest bedreigde boomsoorten in Vlaanderen. De laatste relicten beperken zich tot kleine populaties of solitaire bomen waardoor de noodzaak ontstaat om een genenbank aan te leggen. Deze kan in de toekomst eventueel aangewend worden als zaadboomgaard. Cruciaal hierbij is echter het onderscheid tussen Wilde appels, cultivars en hybriden, wat bemoeilijkt wordt door de aanwezigheid van fenotypisch intermediaire individuen. De in deze thesis aangewende moleculaire merkers bleken in staat dit onderscheid te maken. Zowel AFLP- als microsatelliet merkers lieten een duidelijke differentiatie toe tussen de wilde genenpool, eetappel en sierappels, ondanks het feit dat eetappels in de natuur werden vastgesteld. Een hybridisatie tussen Wilde appels en eetappels werd slechts in 3 genotypen (of 4%) geconstateerd. De beharing van de bladeren bleek een eerste indicatie te geven van de oorsprong van een appelboom maar bereikte niet dezelfde resolutie als de toegepaste nucleaire merkers.

De typering van een chloroplast merker (een duplicatie in de *mat*K regio) resulteerde in contrasterende data en toonde mogelijke introgressie met eetappels aan voor 39% van de Belgische Wilde appels. In de Duitse Wilde appels werd geen introgressie vastgesteld. De hybridisatie werd mogelijks onderschat aangezien de matK-merker niet aanwezig was in alle eetappels. Het grootste deel van de vastgestelde (nucleaire) variatie was aanwezig in de bemonsterde populaties, maar een significante differentiatie werd vastgesteld tussen en binnen Belgische en Duitse herkomsten. Ook het verschil in aanwezige chloroplast typen toont aan dat de Duitse Wilde appels mogelijk een andere geschiedenis kenden. Vandaar dat, op basis van de huidige informatie, kan aanbevolen worden om een regionale genenbank aan te leggen met alle Belgische appelbomen die een nucleaire ‘fingerprint’ vertonen die typisch is voor de *M. sylvestris* genenpool. In de nabije toekomst zal informatie

beschikbaar zijn voor een groter aantal Belgische appelbomen en kan beslist worden welke individuen in de zaadboomgaard kunnen opgenomen worden.

De inheemse eiken ***Quercus petraea*** (Wintereik) en ***Quercus robur*** (Zomereik) komen algemeen voor in Vlaanderen. Autochtone populaties van deze vaak aangeplante soorten zijn echter zeldzaam. In deze studie werd getracht om het autochtone karakter van Vlaamse eikenpopulaties te evalueren op basis van de diversiteit in het chloroplast-DNA. De geobserveerde patronen van chloroplast-DNA variatie weerspiegelden de oorspronkelijke postglaciale verdeling van haplotypes en bleek een bruikbaar referentiekader om de oorsprong van eikenbestanden te achterhalen. De analoge resultaten van de plaatselijke *in-situ* veldevaluatie en de chloroplast-analyse tonen aan dat beide methoden geschikt zijn om het autochtone karakter van eikenpopulaties te beoordelen.

Zoals verwacht voor kruisbestuivende soorten, werd enkel een zwakke populatiegenetische structuur vastgesteld voor beide eikensoorten op Vlaamse schaal. Diversiteit binnen populaties was hoog voor autochtone populaties en waren niet significant verschillend tussen beide eikentaxa. De diversiteitstatistieken voor de relatief kleine Vlaamse eikenpopulaties waren ook niet significant verschillend van deze aangetroffen in grote Europese eikenbossen; er werden geen aanwijzingen gevonden van een recente ‘bottleneck’. Een klein maar significant overschot aan homozygoten werd vastgesteld in beide soorten. Er werd geen significant verschil vastgesteld tussen autochtone herkomsten en erkende zaadbestanden, wederom voor beide soorten. Deze studie toonde aan dat er geen reden is om erkende zaadbestanden boven autochtone eikenpopulaties te verkiezen om de diversiteit van toekomstige eikenbossen veilig te stellen. Op basis van de vastgestelde populatiegenetische structuur in Vlaamse eiken, lijkt het aflijnen van behoudséénheden enkel op basis van chloroplastdiversiteit niet aangewezen. De chloroplastdiversiteit geeft immers niet de enorme diversiteit aan zoals die in hedendaagse eikenbossen aanwezig is. Deze werkwijze zou verder de mogelijke locale aanpassingen negeren die na de herkolonisatie zijn opgetreden.

AFLP-merkers toonden een duidelijk onderscheid tussen Zomer- en Wintereik en maakten het bovendien mogelijk om een individuele eik aan een soort toe te wijzen. Soortspecifieke merkers werden echter niet gevonden. De toewijzing op basis van AFLP-merkers was in overéénstemming met de karakterisering op basis van morfologische parameters. De toepassing van zes microsatellieten kon tot op zekere hoogte ook eikenpopulaties van beide soorten onderscheiden maar meer loci dienen getypeerd te worden om de resolutie van AFLP te bereiken. Deze resultaten stroken met de bestaande eikenliteratuur die

aangeeft dat beide soorten slechts gedifferentieerd zijn voor een beperkt aantal loci. Indien gewenst kan AFLP worden toegepast om de aanwezige eikensoorten in een populatie of op een kwekerij te bepalen. Uitgaande van het feit dat hybridisatie tussen beide eikensoorten een occasioneel en natuurlijk fenomeen is, is er geen reden om populaties met *Q. x rosaceae* anders te behandelen in behoudsmaatregelen.

Een studie op basis van AFLP werd uitgevoerd om de populatiegenetische structuur van ***Carpinus betulus*** (Haagbeuk) te onderzoeken. In overéénstemming met het uitkruisende karakter van deze windbestuiver werd een hoge diversiteit binnen populaties vastgesteld en een zwakke (maar significante) genetische differentiatie tussen populaties zowel op Vlaamse als Europese schaal. Enkel op Europese schaal bleek deze genetische structuur gecorreleerd met de geografische structuur. Het zwakke geografische patroon dat mogelijk aanwezig was op Vlaamse schaal kan gemaskeerd zijn door verplaatsing van Haagbeuken over beperkte afstanden om de lijnvormige landschapselementen aan te leggen. De diversiteit binnen populaties bleek significant (positief) gecorreleerd met de geografische afstand tot het glaciale refugium, wat suggereert dat de genetische diversiteit is toegenomen tijdens de postglaciale herkolonisatie ondanks de ‘bottleneck’ aan het begin ervan. Het huidige patroon van genetische diversiteit dient echter verder onderzocht, bij voorkeur aan de hand van co-dominante merkers. De vaststelling van een geografische structurering van de genetische diversiteit suggereert dat mogelijk adaptieve verschillen bestaan tussen ververwijderde Haagbeukpopulaties. Verder werd ook aangetoond dat de genetische diversiteit binnen populaties gemiddeld lager was in Vlaamse dan in Europese populaties. Op basis van de informatie uit deze studie is het daarom aan te raden om bij toekomstige beplantingen (i) zaden op verschillende locaties te oogsten en samen te brengen en (ii) lokale of nabij gelegen herkomsten te gebruiken.

Contents

Preface – Woord vooraf iii

Acronyms and abbreviations v

Abstract vii

Samenvatting xi

Contents xv

1 General introduction 1

1.1 Conservation of forest genetic resources 3

1.1.1 In practice 3

1.1.2 Molecular contributions 3

1.1.3 International context 5

1.1.4 Conservation of forest genetic resources in Flanders 6

1.1.5 The legal context 8

1.2 Occasion of the present research and outline of this thesis 9

2 General Description of Methods 11

2.1 Molecular techniques 13

2.1.1 DNA extractions 13

2.1.2 Amplified fragment length polymorphisms (AFLP) 14

2.1.3 Microsatellites (or Short Sequence repeats, SSR) 15

2.1.4 Study of chloroplast DNA variation of oaks using PCR-RFLP 17

2.1.5 Study of chloroplast DNA variation in the *mat*K region of apples 22

2.2 Analysis of molecular data 24

2.2.1 AFLP data 24

2.2.2 SSR data 25

2.2.3 Estimation of correlation between different types of data 27

2.2.4 Chloroplast variation in oaks 27

2.2.5 Chloroplast variation in apples 28

2.3 Morphological evaluations 28
2.3.1 Leaf parameters in oak 28
2.3.2 Leaf parameters in apple 30

3 Genetic diversity of *Malus sylvestris* 31
3.1 Introduction 33
3.1.1 *Malus sylvestris* (L.) Mill. 33
3.1.2 Occurrence of *Malus sylvestris* in Flanders 35
3.1.3 Relationship with *Malus x domestica* 36
3.2 Present apple research and objectives 38
3.3 Studied genotypes 41
3.4 Results 43
3.4.1 Description of hairiness of inferior leaf surfaces 43
3.4.2 Allelic variation at AFLP and microsatelite loci 43
3.4.3 Relationships among genotypes using nuclear markers 44
3.4.4 Relationships among samples using nuclear markers 48
3.4.5 Relationships among genotypes using the cp-marker *mat*KdupII 51
3.4.6 Assignment of genotypes to gene pools combining morphological, nuclear and cp-information 54
3.5 Discussion 58
3.5.1 Comparison of results from AFLP and microsatellite marker systems 58
3.5.2 Delineation of the *Malus sylvestris* gene pool 58
3.5.2 Chloroplast variation at the matK region: is duplication II diagnostic for *M. x domestica*? 59
3.5.4 Occurrence of hybrids and cultivated genotypes in the wild 60
3.5.5 Population genetic structure in *Malus sylvestris* 62
3.6 Conclusions 63

4 Genetic diversity of *Quercus robur* and *Q. petraea* 65

4.1 Introduction 67

4.1.1 *Q. petraea* and *Q. robur* 67

4.1.2 Genetic diversity and evolutionary history at the European level 68

4.1.3 Hybridisation between *Q. petraea* and *Q. robur* 70

4.1.4 Occurrence and management of oak stands in Flanders 72

4.2 Present oak research and objectives 74

4.3 Sampled oak stands 75

4.4 Chloroplast DNA diversity in Flanders and patterns of diversity 78

4.4.1 Results: Occurrence of cpDNA haplotypes in Flanders 78

4.4.2 Discussion 80

4.5 Nuclear DNA diversity in Flanders and patterns of diversity 84

4.5.1 Allelic variation at AFLP and microsatellite loci 84

4.5.2 Nuclear DNA diversity in Flanders and patterns of diversity studied with microsatellite markers 86

4.5.3 Comparison of diversity and differentiation measures as detected with AFLP and microsatellite marker systems 100

4.6 Conclusions 112

5 Genetic diversity of *Carpinus betulus* 113

5.1 Introduction 115

5.1.1 *Carpinus betulus* 115

5.1.2 Genetic diversity and evolutionary history 116

5.1.3 Occurrence of hornbeam in Flanders 117

5.2 Occasion of the present research and objectives 118

5.3 Sampled *Carpinus* populations 118

5.4 Results 121

5.4.1 Choice of AFLP primer combinations 121

5.4.2 Reproducibility of AFLP patterns over two years 122

5.4.3 Allelic variation at AFLP loci 123

5.4.4 Relationships among individuals 124
5.4.5 Relationships among populations 128
5.4.6 Within-population diversity 131
5.5 Discussion 134
5.5.1 Use of AFLP for population genetic study of *Carpinus* 134
5.5.2 Relationship between *C. betulus* and *C. orientalis* 134
5.5.3 Structuring of nuclear genetic diversity 135
5.6 Conclusions 138

6 General Discussion 141
6.1 AFLP versus microsatellite markers 143
6.1.1 Use of dominant markers in population genetic studies 143
6.1.2 Estimates of genetic diversity 144
6.1.3 Estimates of inter-specific and inter-population genetic differentiation 146
6.2 Nuclear and cytoplasmic genetic diversity in tree species 147
6.3 The relationship between neutral markers, genetic diversity and adaptive potential 148
6.4 Dealing with hybrids in conservation issues 150

7 General conclusions and future prospects 153
7.1 *Malus sylvestris* 155
7.2 *Quercus robur* and *Q. petraea* 156
7.3 *Carpinus betulus* 158
7.4 Importance of a European context 159

8 References 161

List of Publications 181
Curriculum vitae 185

1 General introduction

1.1 Conservation of forest genetic resources

As a result of widespread deforestation and overexploitation, many tree species have recently become the focus of growing conservation concern. These concerns have drawn attention to the within-species variation, as the extent and distribution of genetic variation within a species are of fundamental importance to its evolutionary potential and determine its chances of survival (Newton *et al*. 1999). The within-species diversity is of particular importance for securing evolutionary flexibility for sessile and long-living organisms as trees (Gregorius & Kleinschmit 1999).

1.1.1 In practice

Forest genetic resources conservation may be either dynamic or static. Dynamic (or *in situ*) conservation maintains the genetic variability of evolving populations thanks to the combined effects of environmental pressure and sexual reproduction. *In situ* conservation preserves the species' adaptation potential over the long term and lets them evolve in and with their natural environment (Lefèvre 2001). The set up of an *in situ* conservation strategy has two steps: the selection of the units that will make up the conservation network and the definition of management guidelines for each unit. Static (or *ex situ*) conservation is a possible alternative when dynamic conservation is not feasible. The main advantages of *ex situ* conservation are twofold: it can be put into place rapidly and the material protected is immediately available for use, for example in genetic improvement programmes or to rebuild populations that have disappeared from their native range. The inconvenience over the long term is that the protected individuals are excluded from the interplay of recombination and selection, which creates new diversity (Lefèvre & Collin 2001). The set up of an *ex situ* conservation strategy has three steps: harvesting, conservation and regeneration. Prior to the three treatment stages, a selection of genotypes has to be made in order to conserve a representative sample of a particularly precious, rare or threatened part of a species' genetic diversity.

1.1.2 Molecular contributions

Molecular approaches can be of value to conservation efforts by providing tools for measuring and managing genetic diversity and for investigating the processes that influence it (Morritz 1994). One of their main advantages is that by screening different marker types, different levels of organisation can be analysed (Schaal *et al*. 1998). Phylogeographic studies based on chloroplast DNA markers have proven a powerful approach for analysing and interpreting patterns of differentiation between populations and regions in a historical perspective (e.g. Demesure *et al*. 1996 for *Fagus sylvatica*; King & Ferris 1998 for *Alnus glutinosa*; Dumolin-Lapègue *et al*. 1997 for *Quercus* spp.). Without a clear knowledge of the historical factors that shaped variation, interpretations of contemporary geographical distribution of

genetic diversity may be difficult (Schaal *et al.* 1998). On the other hand, to define current population genetic structure, techniques targeting the whole nuclear genome, such as isozyme, microsatellite and AFLP markers, are more appropriate (Haig 1998). Studies of different tree species using nuclear markers have identified far less population differentiation than studies of the same species using organelle encoded markers (e.g. Comps *et al*. 2001 for *Fagus sylvatica* using isozymes; King & Ferris 2000 for *Alnus glutinosa* using ISSR; Zanetto *et al*. 1994 and Mariette *et al*. 2002b for *Quercus spp*. using isozymes and a combination of AFLP and microsatellites respectively). These results are consistent with theoretical predictions of higher population differentiation in maternally inherited markers, partially for the lower potential for dispersal in seeds than in pollen, and partly because of the greater genetic drift for the haploid genome in small refugial populations (Ennos 1994).

Molecular information can thus be applied as a rationale basis for identifying units for conservation and can, in contrast to classical taxonomic tools, fully consider intraspecific variants (Newton *et al*. 1999). A sound definition of a conservation unit should ideally be supported by data from both organelle and nuclear loci, in order to reflect both historical and recent differentiation processes and because conclusions based on a single marker system could give misleading indications of the extent of within-species variation that exists (Schaal *et al*. 1998). In this regard, different concepts of evolutionary significant units (ESUs) have been proposed (e.g. Ryder 1986; Morritz 1994 and 1999; see Fraser & Bernatchez 2001 for a review). Differences in the various concepts lie in the criteria used to define ESUs themselves, e.g. dependence on monophyly, the time scales involved or different ways of gene flow reduction. Nevertheless, conflicting ESU concepts are all essentially aiming to define the same thing: segments of species whose divergence can be measured or evaluated by putting different emphasis on the role of evolutionary forces at varied temporal scales (Fraser & Bernatchez 2001). These concepts, originally developed for animal taxa, will need to be refined for application to tree species because of the widespread occurrence of hybridisation, introgression and reticulate evolution (Newton *et al*. 1999).

However, caution is required when using molecular evidence as a basis for conservation programs. The greatest understanding of molecular markers information will come when it is used in conjunction with ecological, demographic or physiological data collected in the field (Haig 1998; Cruzan 2000). Furthermore, differences found in the often used random (selectively neutral) molecular markers can be indicative of adaptive differentiation between populations but this is by no means always the case (Hedrick 2001). A discrepancy between neutral and adaptive diversity can be due to the nature of the marker technique and data

analysis applied (Hedrick 1999). But also different evolutionary scenarios might be responsible for different amounts of neutral and adaptive variation. The variation in adaptive genome regions might reflect the past influences of selection, which can be different for each gene, superimposed on the pattern of variation as a result of history, migration and drift, that is expected to affect all markers in similar ways (Hedrick 2001; van Tienderen *et al*. 2002).

1.1.3 International context

Also policymakers acknowledged the importance of the conservation of forest genetic resources. This was formulated in resolution S2 of the Strasbourg Conference (Ministerial Conference on the Protection of Forests in Europe 1990): 'Above and beyond the conservation of forest species, the essential objective is the conservation of the genetic diversity of these species, which are an essential part of mankind's heritage'. The further recognition of the vital importance of biodiversity and the worldwide loss of diversity resulted in the Convention on Biological Diversity, adopted in Rio de Janeiro in 1992. The European Forest Genetic Resources Program (EUFORGEN, www.ipgri.cgiar.org/networks/euforgen) was established as a mechanism to implement resolution S2 of the Strasbourg Ministerial Conference. EUFORGEN is a collaborative program among European countries aimed at ensuring the effective conservation and the sustainable use of forest genetic resources in Europe. EUFORGEN operates through networks in which forest geneticists and other forestry specialists meet and work together to analyse needs, exchange experiences and develop conservation methods for selected species. The networks also contribute to the development of conservation strategies for the ecosystems to which these species belong. The five operating gene conservation EUFORGEN-networks are: conifers, *Populus nigra*, Mediterranean oaks, noble hardwoods and social broadleaves (temperate oaks and beech). Belgium is an active member of all networks except the Mediterranean oaks network.

1.1.4 Conservation of forest genetic resources in Flanders

At the Flemish scale, actions towards the protection of native woody plant species were undertaken by the Forest and Green Areas Division[1], together with the Institute for Forestry and Game Management[2]. Recently, protection efforts have focused on the conservation of the last autochthonous relic populations. Autochthonous material can be defined as 'material that since its spontaneous colonisation after the last ice age has only rejuvenated naturally or has been artificially rejuvenated with material of strict local origin' (Heybroeck 1992). The term 'autochthonous' thus applies on a specific population or individual, in contrast to 'indigenous' what refers to species that reached a certain region without direct or indirect human actions. Autochthonous populations can thus be seen as the direct descendants of the plants that colonised a particular region during the last post-glacial migration northwards. It is generally assumed that local selection pressures have adapted the established populations to the local environmental and climatic conditions, what is reflected in the current patterns of differentiation for phenotypic traits (Kremer *et al*. 2002).

While the importance of autochthonous genetic resources is generally recognised, the discrimination between autochthonous populations and forests that have been created by man is anything but straightforward, especially in the Flemish region that has suffered an exceptionally high human impact. Ancient Flemish forests have been largely reduced in size during two major deforestation periods (Middle Age and between 1730 and 1880, Tack *et al*. 1993). After the massive deforestations in the Middle Age, forest areas increased again starting around the beginning of the 14th century. It is known that these substantial reforestations would not have been possible without the supply of plant material grown in tree nurseries (Tack *et al*. 1993). However, it is difficult to envisage that storage and transportation of seeds occurred on any appreciable scale during the Middle Age and the following centuries. Even the vast areas of the next large deforestation, that occurred in the

[1] **The Forest and Green Areas division** (Ministry of the Flemish Community) aspires at forest ecosystems with a high level of naturalness. Half of the Flemish forest is currently planted with exotic species and it is the intention of the Division to turn the public forests into mixed broadleaved forests with a substantial part of indigenous species. For future forestations it is stated that 80% of the planted material must be of indigenous species. Furthermore, the importance of the genetic origin of forest reproductive material used for plantations is also acknowledged and the conservation and use of autochthonous material is stimulated.

[2] **The Institute for Forestry and Game Management** is active in the field of forest genetic resources. The mission of the section 'Conservation and use of genetic resources in forestry' is twofold. In the conservation of genetic resources, emphasis is laid upon the inventory, collection and characterisation of genetic resources and conservation of autochthonous gene sources. In the selection and breeding of tree species the goal is to create economically superior reproductive material of valuable tree species.

18th and 19th century, were probably regenerated using seed from local forests near farmers' settlements (König *et al*. 2002). Large long-distance forest seed transfers started during the 19th century and continued in the 20th century, with the aid of the newly developed railroad system. Records of seed transports from Austria and Hungary to Germany and between Belgium, France and Germany are available for this time period (König *et al*. 2002).

Based on this historical information, Maes (1993) developed criteria to evaluate the autochthonous character of a tree or shrub population, including characteristics of both the plants and their growth site.

Important criteria regarding the location are:

- ✓ The location was present (e.g. as forest or hedge) on the topographic map of 'de Ferraris' (drawn from 1771 till 1778) or on other old topographic maps;
- ✓ The location gives an undisturbed impression;
- ✓ The ecological conditions of the location are similar to the conditions in the natural area of prevalence of the species;
- ✓ Plant species indicative for old growth forest (Tack *et al*. 1993) are present in herbaceous, shrub or tree layer;
- ✓ The location lies within the natural distribution range of the species.

Important criteria regarding the tree or shrub are:

- ✓ The tree or shrub is a wild individual and no cultivated variety;
- ✓ The tree is old or old coppice stools are present.

In addition, historical sources and oral information can be used to further establish the autochthonous character of the location. However, as the author points out, in most situations this set of criteria shall be only partially applicable and the autochthonous character of a population can only be established with absolute certainty in a very limited number of cases.

The identification of autochthonous populations is the first step towards the preservation of the autochthonous genetic heritage. Therefore, the Forest and Green Areas Division started in 1997 with a systematic inventory of growth sites harbouring autochthonous trees and shrubs in Flanders, applying the criteria defined above. This inventory is still unfinished but has already revealed that putative autochthonous populations of almost all indigenous woody species are very rare (Maes & Rövekamp 1998; Rövekamp & Maes 1999; Maes & Rövekamp 2000; Rövekamp & Maes 2000; Rövekamp *et al*. 2000; Opstaele 2001). If necessary, *in situ* conservation measures can be installed immediately together with the owner of the site. Furthermore, these autochthonous resources could also be officially acknowledged as seed sources to allow commercialisation of harvested seeds and the marketing of autochthonous forest reproductive material (see 1.1.5). This will help to protect the autochthonous populations *in situ*. In parallel, elaboration of living *ex situ* collections of the autochthonous genomes is seen as an important

conservation act for the long-term preservation of the remaining genetic diversity. If necessary, seed harvest in the living gene banks can be used to produce autochthonous plant material (Vander Mijnsbrugge *et al*. submitted). However, the populations should also be genetically divers in order to safeguard diversity in future forests.

1.1.5 The legal context

Until recently only stands selected for their superior forestry characteristics (e.g. stem form, vigorous growth) could be legally used as seed sources for the production of forest reproductive material for economically important forest species as oaks and beech. A list of recommended provenances, including both Belgian and foreign seed stands, was established for use in the Flemish region. Although their exact origin is often unknown, it is assumed that most of these provenances are from non-autochthonous origins. For many other tree species, no legislation was present and thus no control on the origin of reproductive material was possible.

During the last few years, the importance of autochthonous populations as seed sources for future forestations has been recognised as an important part of the conservation of autochthonous relics. Therefore, possibilities for trade in autochthonous forest reproductive material have been recently generated in the European directive on forest reproductive material (1999/105/EG) under the newly created category ‘source identified’. This category enables trade in reproductive material from forest stands that are not selected on the basis of superior forestry traits and thus the trade in reproductive material of autochthonous stands can be legally organised in the framework of this directive. In order to implement this new European legislation at a Flemish scale, the adaptation of the regional legislation (article 42 of the Flemish forest decree) on the trade of forest reproductive material was necessary. This modification was recently completed and it is expected that the new legislation will become in force in February 2003. The new Flemish legislation will also enlarge the number of species that have to be certified compulsory (47 species) or facultative (33 species). The use of reproductive material from autochthonous origin, both by private owners and by public organisations, will be stimulated by the subvention policy of the Flemish Community.

1.2 Occasion of the present research and outline of this thesis

The inventory of autochthonous populations, the new legislation and the planned conservation actions were the direct occasion for the present research. It had previously been shown that molecular markers were useful tools both in refining the evaluation process of the origin of populations and in the establishment of the population genetic structure present. It was therefore decided to examine the possibilities of a molecular approach to identify autochthonous populations of trees and shrubs in Flanders and to support the set up of sound conservation programs for these populations. Species with very different histories and contemporary situations were selected for a pilot study.

Malus sylvestris, the apple tree native to western and central Europe, is extremely endangered. Recent inventories revealed that the wild apple is also one of the most rare tree species in Flanders and the remaining individuals are very scattered. Therefore, the construction of a gene bank is urgently needed in order to safeguard these last genotypes. This *ex situ* collection can be seen as a new interbreeding population that provides genetically variable seeds for future restocking of the small populations. However, spontaneous hybridisation with the omnipresent cultivated apple is thought to occur and apple trees with intermediate phenotypes have been recorded in Flemish forests (Maes & Rövekamp 1998; Rövekamp & Maes 1998; Coart *et al*. 1998). This has raised doubts about the wild nature of some of the apple trees present in our forests. It is clear that the discrimination between 'genuine' wild genotypes and genotypes derived from or related to cultivars is the key issue that has to be solved prior to the creation of the living gene bank. A study on the genetic diversity of apple is presented in **chapter 3**.

In contrast to apple, the indigenous oak species ***Q. petraea*** (sessile oak) and ***Q. robur*** (pedunculate oak) are very common in Flanders. Conversely, this is not true for putative autochthonous populations of these often planted species of high economic interest. Also acknowledged oak seed stands are limited and cannot produce the necessary material for all forestations in the region. As a result, foreign material which adaptation to local conditions is unknown, is often imported. Therefore, the Institute for Forestry and Game Management and the Forest and Green Areas Division are working on the acknowledgement of new seed sources of both autochthonous populations and stands selected based on superior economic forestry traits (Vander Mijnsbrugge *et al*. submitted). This would allow the production of reproductive material that is potentially better adapted to the regional conditions than material of foreign origin. Although the identification of autochthonous populations is not always straightforward, recent European research based on chloroplast DNA information provided a methodology that might be of help in the

evaluation of the origin of Flemish stands. However, taking into account that the number of oak locations that are putatively autochthonous is very limited (Maes & Rövekamp 1998, Rövekamp & Maes 1999) and that these populations are mostly of small size, the effects of historical bottleneck processes (induced by deforestations and subsequent recolonisations) on the genetic variation present in these autochthonous relic populations were also investigated. This kind of information is required before these populations can be safely used as seed sources. Also knowledge on the current population genetic structure was required for the delineation of provenances and for the definition of conservation units. The oak research is presented in **chapter 4**.

The last species investigated in this thesis, hornbeam (***Carpinus betulus***), is also a common species in many Flemish regions. However, as hornbeam had in contrast to oak, only limited economic importance in the past, it is thought that trade in its reproductive material was less significant and remained local. Hence, most of the old hornbeam individuals are regarded as autochthonous, including the many old pollard trees planted in hedges and along fields. Until now, it was not possible to certify reproductive material for hornbeam and often no information on the origin of the planted material was available. However, once the new Flemish legislation will become into force in 2003 (see 1.1.5), only trade in certified reproductive material of hornbeam will be allowed. Therefore, the acknowledgement of seed sources and the delineation of regions of provenance are urgently needed. Molecular markers can, as described for the oak research, support these actions by providing information on the population genetic structure. The current research on hornbeam is described in **chapter 5**.

Before embarking on the case studies, general information on the methods applied is given in **chapter 2**. This chapter provides details on the molecular and morphological techniques used in this thesis and on the subsequent treatment of the data.

2 General Description of Methods

2.1 Molecular techniques

2.1.1 DNA extractions

2.1.1.1 DNA extraction protocol for Quercus

Five mature oak leaves were harvested on each sampled tree, frozen in liquid nitrogen and lyophilised for 48 hours. The dried material was further stored under vacuum conditions until DNA extraction. Forty mg of the dried material was ground with a mill (Retsch MM200) and a DNA-isolation was carried out according to the CTAB-extraction procedure of Lefort & Douglas (1999), adapted from Doyle & Doyle (1990). The extraction buffer was prepared containing 50mM Tris HCl pH 8, 20mM EDTA pH8, NaCl 0.7 mM, LiCl 0.4 mM, 1% PVP, 2% SDS, 1% CTAB and 1% ß-mercaptho-ethanol. One ml of extraction buffer was added to each sample and the tubes were incubated at 65°C for 15 min. After incubation, 500µl of chloroform/isoamylalcohol (24:1) was added and the mixture was agitated thoroughly until an emulsion was formed. After 5 min of centrifugation (17,000g), the upper phase was transferred to a new tube and centrifuged again for 1 min (17,000g) in order to pellet possible debris. The supernatant was transferred to a new tube and mixed with an equivalent volume of cold isopropanol. After precipitation of the DNA, the tubes were centrifuged for 1 min (17,000g) and the supernatant was discarded. DNA pellets were washed in 1 ml of 70% ethanol, centrifuged for 1 minute (17,000g) and the washing solution was discarded. Finally, the pellets were vacuum-dried at room temperature for 15 minutes and 50 µl of pure water was added. This method yielded up to 25 µg of genomic DNA per extraction. DNA concentrations were determined, relative to uncut lambda DNA, on 1.5% agarose gels.

2.1.1.2 DNA extraction protocol for Malus and Carpinus

Five leaves were collected on each sampled tree and immediately, frozen in liquid nitrogen and lyophilised for 48 hours. The DNA extraction method for apple followed the CTAB protocol (adapted from Doyle & Doyle 1990) as described in Dumolin *et al*. (1995) with some modifications. Total genomic DNA was extracted from 40 mg dried leaf tissue. Each sample was ground in a polypropylene tube using a mill (Retsch MM200). The extraction buffer contained 100mM Tris HCl pH 8, 20mM EDTA pH8, NaCl 1.4 mM, 1% PVP, 2% CTAB and 0.4% ß-mercaptho-ethanol. One ml of extraction buffer was added to each sample and the tubes were incubated at 55°C for 1h. They were allowed to cool down before adding 400µl of chloroform/isoamylalcohol (24:1). After 5 minutes of centrifugation at 4°C (11,000g), the upper phase was transferred to a new tube, mixed with 400 µl cold isopropanol and stored at –20°C for at least one hour. The tubes were centrifuged for 10 min at 4°C (17,000g) and the supernatant removed. Then 1 ml of 76% ethanol was added to the pellet and the DNA was precipitated as before. Finally, the supernatant was

discarded and the pellet was vacuum-dried at room temperature for 15 min and 50 µl of distilled water was added. This method yielded up to 20 µg of genomic DNA per extraction. DNA concentrations were determined, relative to uncut lambda DNA, on 1.5% agarose gels.

2.1.2 Amplified fragment length polymorphisms (AFLP)

AFLP was performed according to Vos *et al*. (1995) with available products and kits (Preamp primer mix and AFLP™ Core reagent kits, Invitrogen). 250 ng genomic DNA was digested for 2h at 37°C in a final volume of 25 µl containing 10mM MgOAc, 50mM KOAc, 10mM Tris-HCl pH 7.5, 2.5 U *EcoR*I (Invitrogen) and 2.5 U *Mse*I (Invitrogen). Two adaptors, one for the *EcoR*I ends and one for the *Mse*I ends, designed to avoid the reconstruction of the restriction sites, were ligated to the restriction fragments by adding 25 µl of a mix containing 5 pmol *EcoR*I adaptor, 50 pmol *MseI* adaptor, 10 mM ATP, 10 mM Tris-HCL, 10 mM MgOAc, 50 mM KOAc and 1 U T4 DNA ligase (Invitrogen). The ligation mixture was incubated for 2 hours at 37°C. The resulting primary template was diluted to 100 µl with 10 mM Tris-HCl (pH 8.0) and 0.1 mM EDTA.

A pre-amplification step was performed with primers complementary to the *EcoR*I and *Mse*I adaptors with an additional selective 3' nucleotide. The PCR reactions were performed in a 50 µl volume of 10 mM Tris-HCL pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 0.2 mM of each dNTP, 25 ng of each primer (Invitrogen), 1 U *Taq* DNA polymerase (Applied Biosystems) and 5 µl of the diluted restriction/ligation mix. The PCR amplifications were carried out in a Hybaid Omni Gene cycler (or Perkin Elmer Geneamp PCR system 9600 for **hornbeam**) using 20 cycles, each cycle consisting of 30s at 94°C, 60s at 56°C and 60s at 72°C. Primers with six selective bases were used for the selective amplification.

For the generation of AFLP fingerprints of **apple** and **oak**, the *Eco*RI primer was labelled with a fluorescent near-infrared group (IRD-700 or IRD-800) for detection with a LI-COR 4200 DNA Analyzer Sequencer. For the generation of AFLP fingerprints of **hornbeam** the *Eco*RI primer was labelled with a fluorescent label for detection with an ABI Prism 377 DNA sequencer (Perkin Elmer). The PCR amplification mixture was composed of 3 µl diluted pre-amplification product (1/10 of their initial concentration), 1 µl *Mse*I primer at 5 µM, 1 µl *EcoR*I primer at 1 µM, 1 U *Taq* DNA polymerase (Applied Biosystems), 15 µl 10x PCR Buffer (Applied Biosystems) and 1 µl of dNTP's (20 mM each, Amersham Biosciences). The selective amplification was carried out in a Perkin Elmer Geneamp PCR system 9600 with the following parameters: 1 cycle of 2 min at 94°C, 30s at 65°C, 2 min at

72°C, followed by 9 cycles in which the annealing temperature decreases 1°C per cycle, followed by 23 cycles of 1 s at 94°C, 30s at 56°C and 2 min at 72°C.

After PCR-amplification, the **oak** and **apple** samples were denatured by adding 1 µl of IR^2 stop solution (LI-COR) to 1 µl template DNA for detection on the 700 channel, 1.5 µl template DNA for the 800 channel, and by heating for 3 min at 90°C. AFLP fragments were separated by PAGE on a LI-COR 4200 DNA Analyzer Sequencer on 25 cm gels using 6.5% denaturing polyacrylamide gels (KB-plus solution, LI-COR, Inc.). Near-infrared labelled size standards (LI-COR, Inc.) were loaded on each gel at regular intervals for sizing of the AFLP fragments. Gels were run for 2.5 h at 1,500V. Each gel was used for two subsequent PAGE runs. Gene ImagIR software 3.55 (LI-COR) was used to score the fragment size and to define marker bins. Each marker was coded as present (1) or absent (0) for each plant and a binary data matrix was created using Microsoft Excel and Microsoft Access. AFLP bands were scored within the size range of 65 to 540 bp.

At the end of the selective PCR, the **hornbeam** samples were denatured by adding a mix of 0.9 µl of formamide buffer (Amresco), 0.35 µl Blue dye (Applied Biosystems) and 0.25 µl GS-500 Rox (Applied Biosystems) to 1.5 µl template DNA, and by heating for 3 min at 90°C. Of each sample 1.5 µl was loaded on 5% polyacrylamide/bisacrylamide 19:1 (Biorad), 7.5 M urea (Invitrogen) and 1x TBE gels. Gels (25 x 36 cm) were run in 1 x TBE electrophoresis buffer using an ABI Prism 377 DNA sequencer (Perkin Elmer). GS-500 Rox labelled size standard was loaded in each lane in order to allow the automatic analysis of the data. GeneScan Analysis Software 3.1.2 (Perkin Elmer) was used to translate the information collected by the ABI377 into fragment sizing information. In all cases a peak amplitude threshold of 50 was set for the analysis. The GeneScan files were further scored with Genotyper 2.0 (Perkin Elmer). The presence or absence of the markers in each plant analysed was recorded with Genotyper and a 1/0 matrix was generated using Microsoft Excel and Microsoft Access. AFLP bands were scored within the size range of 75 to 450 bp.

2.1.3 Microsatellites (or Short Sequence repeats, SSR)

2.1.3.1 Amplification of SSR loci

All amplification reactions were performed using the PCR-Core-kit I (Promega). Amplifications were carried out in a reaction volume of 15 µl containing 5 ng of genomic DNA, 10mM Tris-HCl, 0.2mM of each dNTP, 1.5mM $MgCl_2$, 50mM KCl and 1 U *Taq*, according to the protocol of Gianfranceschi *et al*. (1998). Forward primers of each primer pair were labelled with a fluorescent near infrared dye (IRD-700 or IRD-800). Primer concentrations were optimised for PAGE on a LI-COR 4200 DNA

Analyzer; optimised concentrations ranged from 0.02 µM to 0.25 µM depending on the locus amplified and the channel used for visualisation (in general the IRD700 laser is more sensitive than IRD800 and less primer is needed when the fragments are detected on the IRD700 channel).

The amplification was carried out in a Perkin Elmer Geneamp PCR system PE9600 or PE9700 with the following parameters: 4 min at 94°C, followed by 35 cycles of 40 s at 94°C, 40 s at T_A°C and 20 s at 72°C and a final extension of 10min at 72°C. Annealing temperatures (T_A) were chosen according to the original publications (Guilford *et al*. 1997 and Gianfranschesi *et al*. 1998 for apple; Steinkellner *et al.* 1997 and Dow *et al.* 1995 for oak). Sample denaturation and gel electrophoresis were performed as described for apple and oak AFLPs, but the gels were run for only 1.5 h. Each gel was used for three subsequent PAGE runs. Gene ImagIR software (Li-Cor, Inc.) was used to determine fragment sizes and to define allele bins. Alleles were scored according to their molecular weight and the occurrence of between-gel differences was checked using reference samples, which were present on each gel (two Flemish oak samples for oak SSRs and cultivar 'Elstar' for apple SSRs).

2.1.3.2 Oak microsatellite loci

For **oak**, six primer pairs of microsatellite loci (Table 2.1) which had previously been shown to display easy to read band patterns, which mapped to different linkage groups (Barrenche *et al.* 1998) and which displayed a high degree of polymorphism in *Quercus robur* and *Q. petraea*, were used (Dow *et al.* 1995; Steinkellner *et al.* 1997). Four SSR loci (MSQ4, AG15, AG104, AG110) did only amplify if the ramp time (time it takes the instrument to change from one temperature to another during PCR-cycling) was increased to 50% of the maximum speed (on PE 9700) or to 2 minutes (on PE 9600). Size ranges of amplification products overlapped for almost all oak loci and only one locus could be detected by each laser during one run.

Table 2.1: SRR loci used for diversity screening in oak.

Locus	Repeat motif	Linkage group
MSQ4[b]	$(GA)_{17}$	G4
MSQ13[b]	$(GA)_{14,11}$	G6
AG104[a]	$(AG)_{16}AT(GA)_3$	G2
AG110[a]	$(AG)_{15}$	G8
AG15[a]	$(AG)_{23}$	G9
AG9[a]	$(AG)_{12}$	G7

[a]: Steinkellner *et al*. 1997; [b]: Dow *et al*. 1995.

2.1.3.3 Apple microsatellite loci

For **apple**, twelve primer pairs of microsatellite loci (Table 2.2) which had previously been shown to display easy to read band patterns and which mapped to different chromosomes were chosen (Guilford *et al*. 1997, Gianfranceschi *et al*. 1998, Maliepaard *et al*. 1998, Liebhard *et al.* 2002). Two PCR products with the same IRD-label and different size ranges, were multiplexed after PCR and loaded on the gel together, enabling the visualisation of 4 microsatellite loci in a single run (2 loci by each laser).

Table 2.2: SRR loci used for diversity screening in apple. Normalised nomenclature of the loci follows Liebhard *et al*., 2002.

Locus	Repeat motif	Chromosome
NZ02b01[a]	$(GA)_{14}$	15 [b]
NZ04h11[a]	$(GA)_{23}$	9 [b]
NZ05g08[a]	$(GA)_{18}$	4 [b]
NZ23g04[a]	$(GA)_{19}$	6 [b]
NZ28f04[a]	$(GA)_{18}$	12 [b]
CH01h10[c]	$(AG)_{21}$	8 [d]
CH01e12[c]	$(AG)_{32}$	8 [e]
CH01f02[c]	$(AG)_{22}$	12 [d]
CH01h01[c]	$(AG)_{25.5}$	17 [d]
CH02b12[c]	$(GA)_{26}$	5 [d] + 10 [e]
CH02c06[c]	$(GA)_{21}(GA)_{17}$	2 [d]
CH02d12[c]	$(GA)_{19}$	11 [d]

[a] From Guilford *et al*. 1997; [b] Mapped in Maliepaard *et al.* 1998; [c] From Gianfranceschi *et al*. 1998; [d] Mapped in Liebhard *et al*. 2002; [e] Mapped by Eric van de Weg *et al*. (in preparation).

2.1.4 Study of chloroplast DNA variation of oaks using PCR-RFLP

A set of conserved primers homologous to the most conserved coding regions of chloroplast DNA (cpDNA) and that allow the amplification of the more variable non-coding regions, according to Dumolin-Lapègue *et al*. (1997) were used to identify the chloroplast haplotypes present in Flemish oak populations. Three cpDNA fragments were used in this study, each in combination with a restriction enzyme: *trn*D/*trn*T with *Taq*I, *psb*C/*trn*D with *Taq*I and *trn*T/*trn*F with *Alu*I. The first two primer pairs are described in Demesure *et al*. (1995), the last one in Taberlet *et al*. (1991). Although in the publications of Demesure *et al*. (1995) and Dumolin-Lapègue *et al*. (1997) more primer pairs are available for the determination of oak cp-haplotypes, the use of the three primer pairs mentioned above was sufficient to assign the Flemish oak samples analysed a cp-haplotype.

Diluted DNA (5ng/µl) was used as template in PCR reactions. The PCR reaction mixture contained 67mM Tris HCl, 2 mM $MgCl_2$ 100 µM of each dNTP, 0.2 µM of

each primer and 0.2 U of *Taq* polymerase (all reaction components Promega, PCR-Core-kit I). The amplification was carried out using 1 cycle of 4 min at 94°C, 30 cycles of 45s at 92°C, 45 s at T_A, 2 or 4 min at 72°C (depending on the length of the fragment analysed) followed by a final elongation of 10 min at 72°C. Annealing temperatures (T_A) as found in the original publications were applied. Five µl of the amplification product was added to 15 µl digestion mix, containing 5 U of the restriction enzyme (all enzymes from Fermentas). PAGE was carried out on an electrophoresis apparatus model Protean® IIxiCell of BIO-RAD. A 1-Kb ladder (Invitrogen) was used as molecular weight marker. The migration was performed on 1.5 mm thick, 8% polyacrylamide gels (Protogel, National Diagnostics). Gels were run for approximately 4h at 300V and stained with ethidium bromide afterwards.

The cpDNA fragment amplified with the primers *trn*D and *trn*T was digested with the restriction enzyme *Taq*I. This allowed us to classify the trees into four categories: cp-haplotype 10 or 11, cp-haplotype 1, 2 or 32, cp-haplotype 12 or cp-haplotype 7 (figure 2.1A). The only haplotype of Balkan origin present in the sampled trees (cp-haplotype 7), and haplotype 12, of Iberian origin could be identified unambiguously using only one primer/enzyme combination. For other types, a second primer-enzyme system (*trn*T-*trn*F to distinguish between haplotypes 10 and 11 and *psb*C-*trn*D to distinguish between haplotypes 1, 2 and 32, Figure 2.1B and 2.1C) was required in order to define the exact haplotype. Therefore, for each individual tree included in this study, the analysis of up to two cpDNA fragments enabled the unambiguous determination of the haplotype. Haplotypes were identified following Dumolin-Lapegue *et al*. (1997) and Petit *et al*. (2002a).

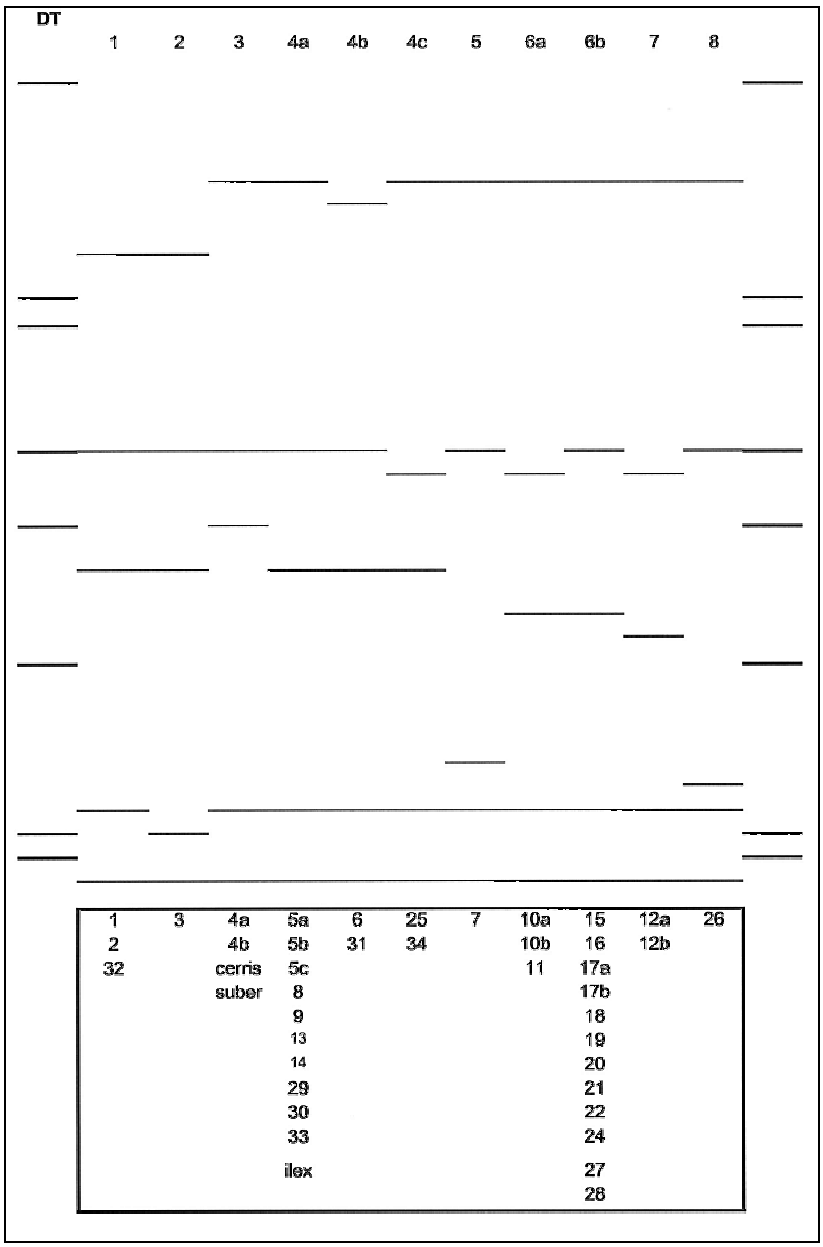


Figure 2.1A: Restriction diagram *trn*D/*trn*T-*Taq*I. Molecular weight markers are indicated on both sides of the diagram. The other lanes correspond to the different patterns observed with this particular PCR fragment/restriction enzyme. One pattern may correspond to one or several haplotypes, which are listed below under the corresponding lane. All are haplotypes found within white oaks, except were mentioned (*Q.) cerris*, (*Q.) suber* and (*Q). ilex*. The patterns shown in lanes 1, 5, 6a and 7 were found in this study. Figure taken from FAIROAK website: http://www.pierroton.inra.fr/Fairoak.

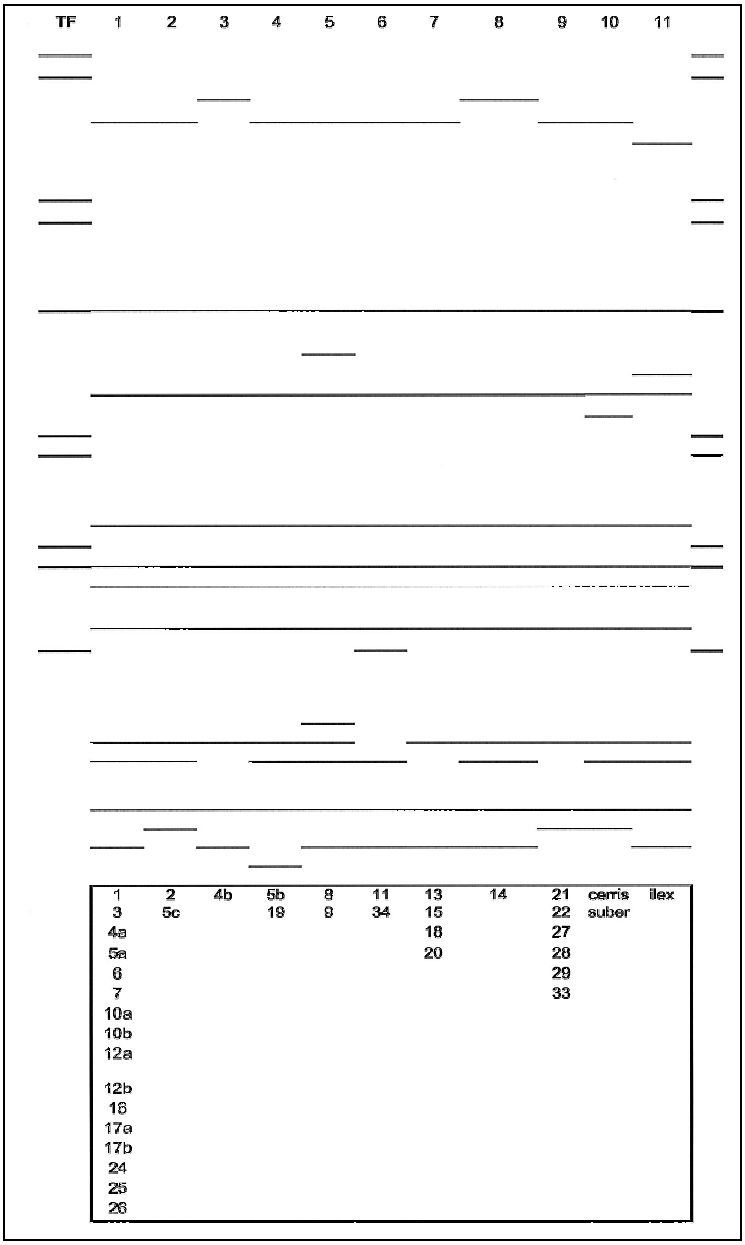


Figure 2.1B: Restriction diagram *trn*T/*trn*F-*Alu*I. Molecular weight markers are indicated on both sides of the diagram. The other lanes correspond to the different patterns observed with this particular PCR fragment/restriction enzyme. One pattern may correspond to one or several haplotypes, which are listed below under the corresponding lane. All are haplotypes found within white oaks, except were mentioned (*Q.) cerris*, (*Q.) suber* and (*Q). ilex*. The patterns shown in lanes 1 and 6 were found in this study. Figure taken from FAIROAK website: http://www.pierroton.inra.fr/Fairoak.

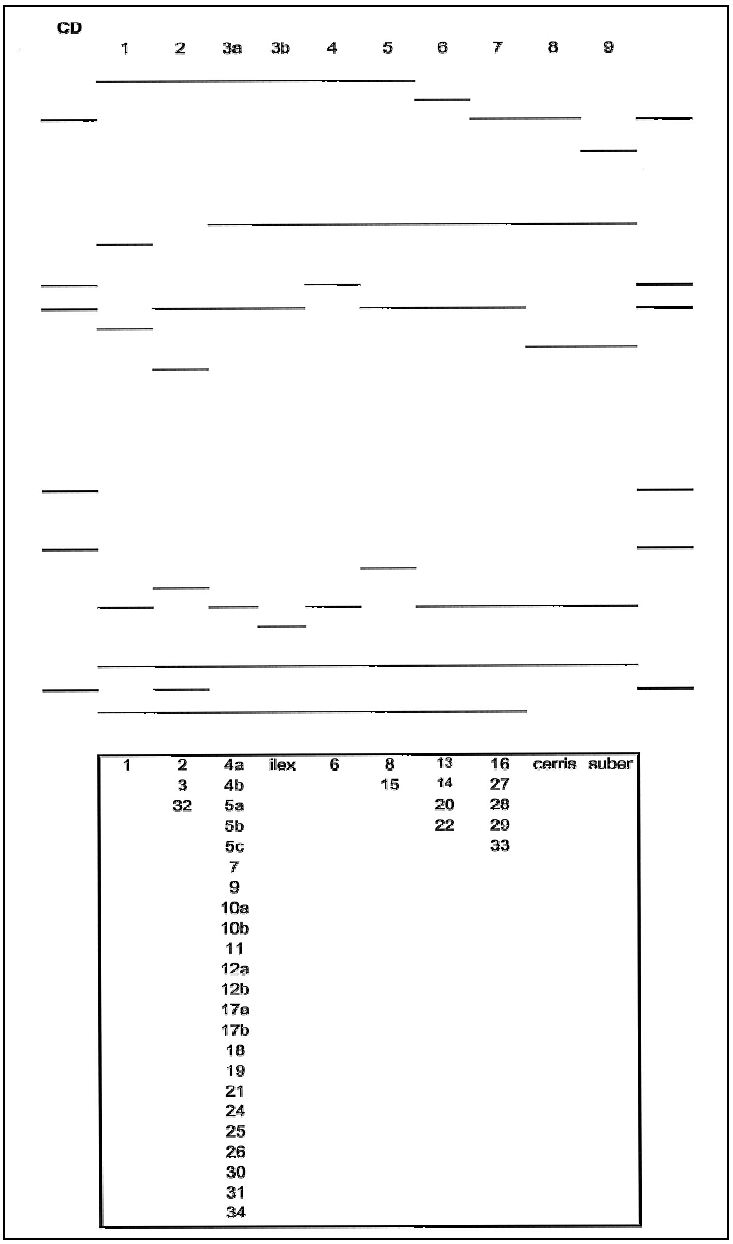


Figure 2.1C: Restriction diagram *psb*C/*trn*T-*Taq*I. Molecular weight markers are indicated on both sides of the diagram. The other lanes correspond to the different patterns observed with this particular PCR fragment/restriction enzyme. One pattern may correspond to one or several haplotypes, which are listed below under the corresponding lane. All are haplotypes found within white oaks, except were mentioned (*Q.) cerris*, (*Q.) suber* and (*Q). ilex*. Only the pattern shown in lanes 1 was found in this study. Figure taken from FAIROAK website: http://www.pierroton.inra.fr/Fairoak.

2.1.5 Study of chloroplast DNA variation in the *mat*K region of apples

2.1.5.1 Sequencing of the matK region

The *mat*K gene was amplified as described in Robinson *et al*. (2001) using primers developed by Sang *et al*. (1997), with modifications to the PCR program (Dendauw *et al*. 2002): 10 min at 95°C, 5 cycles of 1 min at 95°C, 30 s at 53°C, 1 min at 72°C, 35 cycles of 30 s at 95°C, 1 min at 53°C, 75 s at 72°C and finally 7 min at 72°C. The amplifications were carried out in a Perkin Elmer Geneamp PCR system PE9600 or PE9700. In order to sequence the complete gene, new primers were developed (using the software primer Express version 2.0.0, Applied Biosystems) starting from partial sequences that were obtained by using the primers of Sang *et al*. (1997). In addition, two primers developed for sequencing the *mat*K gene in Azalea were used (Dendauw *et al*. 2002). This resulted in three primer pairs that together amplified the complete *mat*K gene. Primer pairs and sequences are given in Table 2.3 and 2.4. Following amplification, the PCR products were checked for homogeneity on an agarose gel. PCR fragments were directly sequenced, using the PCR primers as sequencing primers (forward and reverse) and the ABI Prism BigDye Terminator cycle Sequencing kit (Applied Biosystems) according to the manufacturer's protocols. Sequencing reactions were analyzed on an ABI Prism 377 DNA Sequencer (Applied Biosystems).

Table 2.3: Primers used for amplification and sequencing of the matK locus.

Primerpair	Forward primer	Reverse primer	Length fragment amplified (bp)
1	P1F[1]	1RNL1A	643
2	1FWNL1B	4FDR153R	518
3	3MF[2]	*trn*K2R[2]	*776*

[1]: from Sang *et al*. 1997; [2]: from Dendauw *et al*. 2002; other primers developed in this study.

Table 2.4: Sequences of primers used for amplification and sequencing of *mat*K region.

Primer*	Start position	Sequence
P1F (F)	0	ACTGTATCGCACTATGTATCA
1RNL1A (R)	643	GGAGGCAAGAATAATCGTGGAT
1FWNL1B (F)	551	CTTCGCTATTGGGTGAAAGATCC
4FDR153R (R)	1069	AATTTTCTAGCATTTGACCCCG
3MF (F)	992	GTGGTCTCAACCAAGAAGG
*trn*K2R (R)	1802	AACTAGTCGGATGGAGTAG

*: F= forward primer; R= reverse primer

2.1.5.2 Marker development

As had previously been shown in the study by Robinson *et al*. (2001), the presence of an interesting duplication in the *mat*K region, located 39 bp upstream from the 3' end of the *mat*K-coding region might be used to discriminate between wild (*M. sylvestris*) and modern *M. x domestica* cultivars. In order to investigate the presence of this insertion in more *Malus* genotypes, a primer '*mat*KdupF' (sequence ATAGAGGTCGAATTTGGTATTTGGAT) was developed. The fragment containing the duplication was amplified with the primers *mat*KdupF and *trn*K2R (further referred to as *mat*KdupII locus) and using the reaction components and PCR program described above. *TrnK2*R is one of the primers originally developed by Sang *et al*. (1997). PCR fragments were loaded on 8% polyacrylamide gels as explained for oak PCR-RFLP analysis and the fragments were scored as duplication present (PCR-product 351 bp) or duplication absent (PCR-product 333 bp), as shown in Figure 2.2.

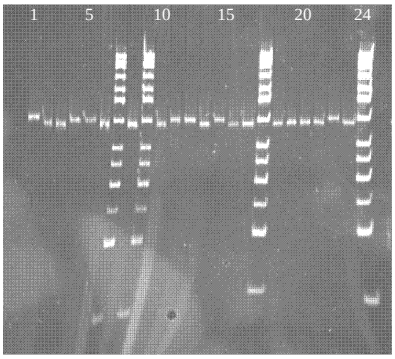


Figure 2.2: Analysis of the *mat*KdupII locus. Four molecular weight markers (1 kb ladder) are loaded in lanes 7, 9, 17 and 24. The other lanes correspond to the amplified *mat*KdupII fragment of different apple samples. Long fragments, indicating the presence of duplication II are found in lanes 1, 4, 5, 11, 12, 14 and 22. Short fragments, indicating the absence of duplication II are found in lanes 2, 3, 6, 8, 10, 13, 15, 16, 18, 19, 20, 21 and 23.

2.2 Analysis of molecular data

2.2.1 AFLP data

After elimination of monomorphic markers, Spearman correlations were calculated between pairs of markers. Using the software AFLP-SURV v1.0 (Vekemans *et al.* 2002, available at http://www.ulb.ac.be/sciences/lagev/), estimates of pairwise relatedness coefficients (*r*) were calculated between individuals according to Lynch & Milligan (1994). A principal co-ordinate analysis (PCO) was performed based on this matrix using NTSYSpc (Rohlf 2000). The first two axes were plotted graphically using SPSS 10.1.0 (SPSS inc. 1989-2000), for **apple** together with 90% confidence ellipses for each group (STATISTICA 6.0, StatSoft Inc., Tulsa, OK, USA). After exclusion of outliers, pairwise genetic distances (Nei 1987) were calculated between samples. A Neighbour-joining tree was computed, based on these distance measurements. One thousand bootstraps were performed over AFLP loci using AFLP-SURV and PHYLIP (Felsenstein 1993).

Furthermore, allelic variation, genetic diversity and differentiation statistics were computed using the software AFLP-SURV v1.0. Allelic frequencies at AFLP loci were calculated from the observed frequencies of fragments using the Bayesian approach proposed by Zhivotovsky (1999) for diploid species and assuming some deviation from Hardy-Weinberg genotypic proportions as estimated from SSR data (if available). A non-uniform prior distribution of allelic frequencies was assumed with its parameters derived from the observed distribution of fragment frequencies among loci (see note 4 in Zhivotovsky 1999). The frequency of the null allele at each locus is computed from two numbers, the sample size and the number of individuals in the sample that lack the AFLP fragment, using a Bayesian method that also estimates the distribution of allele frequencies based on the variation over loci of the frequencies of AFLP fragments in the sample. When the data concerns several populations, the distribution of allele frequencies is estimated separately for each population. This method is supposed to give the most accurate results (Zhivotovsky 1999).

These allelic frequencies were used as input for the analysis of genetic diversity within and between samples following the method described in Lynch & Milligan (1994). Gene diversity (*Hj*) is calculated as the probability that two genes, randomly drawn from population j, differ at the loci typed. This measure is equivalent to the expected heterozygosity under Hardy-Weinberg equilibrium. In order to analyse the population genetic structure the following statistics were calculated:

- *Ht*: the total gene diversity;
- *Hw*: the mean gene diversity within populations;
- *Hb*: the average gene diversity among populations in excess of that observed within populations;

- *Fst*: Wright's fixation index, measuring the genetic correlation between pairs of genes sampled within a population relative to pairs of genes sampled within the overall set of populations (also interpreted as the proportion of the total gene diversity that occurs among as opposed to within populations).

The significance of the genetic differentiation between groups was tested by comparison of the observed F_{ST} with a distribution of F_{ST} under the hypothesis of no genetic structure, obtained by means of 1000 random permutations of individuals among groups. In addition, average fragment size (with standard deviation) and Pearson correlation coefficients between fragment sizes and fragment frequencies (together with the P-value associated with the correlation) were calculated on the overall sample. The significance of the differences between groups of populations for the level of within-population variation was tested using a Student-t test (SPSS 10.1.0, SPSS inc. 1989-2000).

A model-based clustering method was applied on AFLP data to infer genetic structure and define the number of clusters (gene pools) in the dataset using the software STRUCTURE (Pritchard *et al.* 2000). For AFLP data, each class of genotypes was treated as being a haploid allele. Individuals were assigned probabilistically to inferred gene pools.

2.2.2 SSR data

Relatedness coefficients between pairs of individuals were estimated by computing the multilocus Moran's I statistic (Hardy & Vekemans 1999) using the program SPAGeDi 0.0 (Hardy & Vekemans 2002, available at http://www.ulb.ac.be/sciences/lagev/). A principal co-ordinate analysis (PCO) was performed based on this similarity matrix using NTSYSpc (Rohlf 2000); the first two axes were plotted graphically using SPSS 10.1.0 (SPSS inc. 1989-2000), for **apple** together with 90% confidence ellipses for each group (STATISTICA 6.0, StatSoft Inc., Tulsa, OK, USA). After exclusion of outliers, pairwise distances (standard genetic distance, Nei (1978) and $\delta\mu^2$, Goldstein *et al*. (1995)) were calculated between samples (SPAGeDi 0.0) and neighbour-joining trees with bootstraps over SSR loci computed, based on both distance measures using PHYLIP (Felsenstein 1993) and MICROSAT (http://hpgl.stanford.edu/projects/microsat). Deviations of genotypic frequencies from Hardy-Weinberg proportions were tested with the program GENEPOP version 3.3 (Raymond & Rousset 1995).

The following statistics of genetic variation within samples were computed as average over loci with the software GEN-SURVEY (Vekemans & Lefèbvre 1997):

- A: mean number of alleles per locus;
- H_0: average observed heterozygosity;

- H_E: mean expected heterozygosity calculated on the assumption of random mating or average gene diversity, computed according to Nei (1978);
- F_{IS}: Wright's inbreeding coefficient, corrected for small sample sizes (Kirby 1975). F_{IS} is a measure of the deviation of genotypic frequencies from panmictic frequencies in terms of heterozygous deficiency or excess. It conventionally defined as the probability that two alleles in an individual are identical by descent. It is calculated in a single population as $(H_E - H_0) / H_E$. It shows the degree to which heterozygosity is reduced below the expectation. The value of F_{IS} ranges between -1 and +1. Negative FIS values indicate heterozygote excess (outbreeding) and positive values indicate heterozygote deficiency (inbreeding) compared with Hardy-Weinberg expectations. The overall mean value of F_{IS} was used as input for the estimation of AFLP-marker allelic frequencies (see above).

Furthermore, the analysis of gene diversity was performed using the corrections for small sample sizes following Nei & Chesser (1983), including

- H_t: the gene diversity in the total set of populations;
- H_s: the average gene diversity within populations;
- D_{st}: the average gene diversity between populations;
- G_{st}: the coefficient of gene differentiation.

Ninety-five percent confidence intervals for the means were computed by 1,000 bootstraps over loci.

Using the same software, two additional tests based on numerical re-sampling methods were performed (only for the **oak** data set): a test of genetic differentiation between two defined groups and a test for differences between both groups in their level of within-population variation. In each case, the statistical significance of the test was assessed by computing the distribution of the test statistics using a numerical resampling method in which new groups of populations were formed by randomly combining the populations into two groups, and estimates of the test statistic were computed for each new combination of populations. All re-sampling statistics are based on 1,000 samples unless otherwise specified.

The population genetic structure among and within defined groups was analysed using hierarchical *F*-statistics computed with the software ARLEQUIN (Schneider *et al*. 2000). The significance of each variance component was tested with permutation tests (Excoffier *et al*. 1992).

The software STRUCTURE (Pritchard *et al.* 2000) was also applied on multilocus SSR data to infer genetic structure and to define the number of clusters (gene pools) in the dataset. Individuals were assigned probabilistically to inferred gene

pools, or jointly to two or more gene pools if their genotypes indicated that they were admixed. This allowed the identification of hybrids between inferred gene pools and the detection of genotypes that were outliers in their sample of origin and in fact belonged to another gene pool.

For **oak**, the occurrence of recent bottlenecks was inferred using the software BOTTLENECK (Cornuet & Luikart 1996). This program is designed for detection of recent effective population size reductions from allele frequency data. The program is based on the principle that populations which have experienced a recent reduction of their effective population size (i.e. a 'bottleneck') exhibit a correlative reduction of the allele numbers and heterozygosities at polymorphic loci. But the allelic diversity is reduced faster than the heterozygosity, or the observed heterozygosity is larger than the heterozygosity expected from te observed allele number were the locus at mutation-drift equilibrium (i.e. the effective population size has remained constant in the past). The test applied determines whether the proportion of loci with heterozygosity excess is significantly larger than expected at mutation-drift equilibrium, assuming subsequently an infinite allele (IAM) and stepwise mutation model (SMM) underlying the microsatellite loci.

2.2.3 Estimation of correlation between different types of data

Mantel tests (Mantel 1967) with 999,999 permutations (using the software NTSYSpc, Rohlf 2000) were performed to check for significant correlation (i) between the AFLP and SSR datasets of **apple** and (ii) between pair-wise geographic distances and population divergence for the **hornbeam** dataset. Furthermore, the correlation between within-population diversity and geographic distance to the glacial refuge was computed for **hornbeam** populations. For **oak**, populations were ranked according to the diversity parameters A, H_O and H_E for microsatellites and according to H_j for AFLP. The value of these parameters were then compared by computing Spearman's rank coefficient correlation.

2.2.4 Chloroplast variation in oaks

The within-population genetic diversity (h_s), the total diversity (h_t) and the coefficient of genetic differentiation (G_{st}) were calculated according to the method of Pons & Petit (1995). Only those populations represented by at least three trees were included in the analysis. Diversity statistics were calculated for the whole dataset, for autochthonous populations and selected provenances separately and for each species separately.

2.2.5 Chloroplast variation in apples

Raw ABI sequence files were edited and assembled for each genotype with AutoAssembler 2.1.1 software (Applied Biosystems). Alignment was carried out with the software CLUSTAL W 1.81 (Thompson *et al*. 1994) and afterwards the sequences were exported to GeneDoc 2.6.002 (Nicholas *et al*. 1997), where the alignments were checked manually for polymorphisms.

2.3 Morphological evaluations

2.3.1 Leaf parameters in oak

In one *Q. robur* (QRAUTW) and one *Q. petraea* (QPAUTKB) population, leaf material was collected on the same trees sampled for the DNA analysis (30 trees per population). Ten mature leaves were collected from the upper part of the crown of each sampled tree. Leaves were dried and stored in a herbarium. Five intact leaves of each tree were chosen for further analysis. The image analysis system used for the morphological characterisation of the leaves consisted of a camera (KY-F55B, JVC, Japan), frame grabber card (Flashbus, Integral Technologies, USA), IBM-compatible PC and image analysis software (Wit version 5.31, Logical Vision, Canada). A programme was developed to measure the leaf characteristics automatically. Figure 2.3 illustrates the measured leaf parameters. Additional estimations of lamina pubescence were made with a binocular microscope (magnification 40x) and the number of intercalary veins was counted by eye. Table 2.6 summarises the measured parameters and the derived characters used for further computations. A Mann-Whitney U test was performed to estimate the power of the morphological parameters for species discrimination. In addition, a PCO analysis was performed and the first two axes were presented graphically (using NTSYSpc and SPSS).

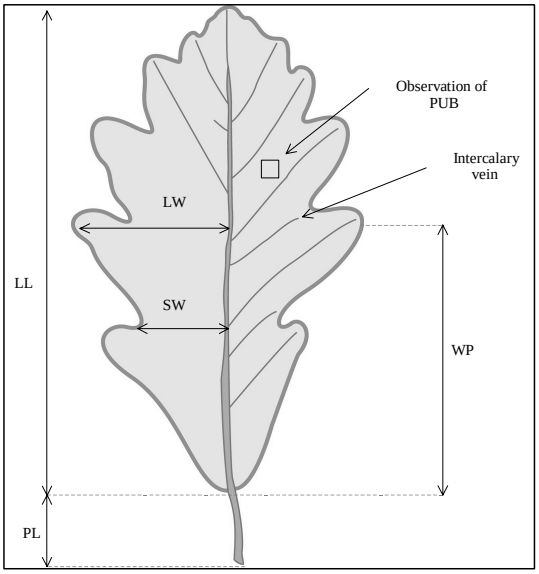


Figure 2.3: Leaf parameters measured on oak leaves. See Table 2.6 for abbreviations.

Table 2.6: List of directly measured and derived morphological characters.

Measured variables		
Abbreviation	**Variable**	**Measure**
LL	Lamina length	Pixel (1/3 mm)
PL	Petiole length	Pixel (1/3 mm)
LW	Lobe width: width of the deepest lobe from the mid rib to the top	Pixel (1/3 mm)
SW	Sinus width: lamina width from the mid rib to the base of the sinus, at the widest part of the lamina	Pixel (1/3 mm)
WP	Widest part: length of lamina from the lamina base to the widest part of the lamina	Pixel (1/3 mm)
NL	Number of lobes	Count
NV*	Number of intercalary veins*	Count
BW_0, BW_{10}, BW_{20}, BW_{-10}, BW_{-20}	Basal width of the lamina measured at the lamina base and 20 and 10 pixels up and down	Pixel (1/3 mm)
PUB*	Abaxial lamina pubescence*	Categorical, from 1 to 6

Derived characters		
Abbreviation	**Character**	**Calculation formula**
LS	Shape of lamina	LL/WP
PR	Petiole ratio	PL/(LL+PL)
LDR	Lobe depth ratio	LW/(LW-SW)
PV	Percentage venation	NV*100/NL
BS	Basal shape	$(BW_{-20} + BW_{-10} + BW_0 + BW_{10} + BW_{20})/5$

*: character evaluated by eye.

2.3.2 Leaf parameters in apple

The hairiness of leaves was described in autumn for all wild apple trees and edible cultivars grafted on the DvP-nursery, according to Wagner (1998). Ornamental trees were not described as they belong to different species. The hairiness of the inferior leaf surfaces was scored from 0 to 3:

Score 0: hairless

Score 1: sparsely hairy, difficult to recognize (magnifying lens needed); hairs only on main nerves

Score 2: moderately hairy, easy to recognize by eye; hairs on nerves and entire leaf surface

Score 3: felted leaf surface

The grafted trees were too young to describe other morphological traits that are thought to discriminate between wild and cultivated trees, such as thorns on twigs and several flower and fruit characters.

3 Genetic diversity of *Malus sylvestris*

3.1 Introduction

3.1.1 Malus sylvestris (L.) Mill.

The name apple applies to the species of the genus *Malus* in the subfamily *Maloideae* of the *Rosaceae* family. This subfamily also includes the native European genera *Pyrus*, *Crataegus*, *Sorbus*, *Cotoneaster*, *Mespilus* and *Cydonia.* Although there is widespread agreement that the base chromosome number of *Maloideae* (x = 17, which is high in comparison with other *Rosaceae* species with x = 7, 8, or 9) indicates that the group had a polyploid origin, there is much less agreement regarding the parental lineages (see Phipps *et al*. 1990 for a review). The current opinion is that *Maloideae* arose entirely from spiraeoid progenitors (x = 9), by autopolyploidy or allopolyploidy, or that hybridisation between *Spiraeoideae* (as maternal lineage) and another lineage (*Amygdaloideae*, x = 8) produced the x = 17 maloid ancestor, and that these polyploidisation events took place before the origin of the pome (Morgan *et al*. 1994). The genus *Malus* is variously reported to consist of 25 to 47 species, with the number of species still a matter of debate and dependent on the authority (Phipps *et al*. 1990). Species within the genus are widely distributed although generally they are found in the northern temperate zones of North America, Europe and Asia.

Malus sylvestris (L.) Mill. is an indigenous wild apple to most European countries. Its natural range is not known exactly mainly because cultivated apples are planted outside the natural range of the wild species and the wild and cultivated apples cannot be discriminated unambiguously (see section 3.1.3). In addition, animal-pollinated plants typically leave no significant traces of their former distribution in the pollen record and pollen of *M. sylvestris* cannot be distinguished from pollen of related *Rosaceae* species. However, European evidence for the collection of apples ‘from the wild’ can be found in Neolithic (ca 11,000 years ago) and Bronze Age (ca 4,500 years ago) settlements suggesting the long time presence of wild apples in central-western Europe (Harris *et al*. 2002).

Wild apple can grow on nearly all soils except on highly acidic soils. Best growth occurs on fresh and basic soils. Its optimum probably is on the wet edge of the forest (Stephan *et al*. in press). *M. sylvestris* is a very light demanding species and has weak competitive abilities. Therefore the species mostly occur in forest edges, hedges or on extreme dry or wet sites where other competitors do not survive (Kleinschmit *et al*. 1998).

For apple trees it takes five to six years from seed to flowering tree. *M. sylvestris* is an outbred species. It has a gametophytic self-incompatibility system, which prevents fertilisation by pollen carrying a self-incompatibility allele if the allele is also

present in the tree being pollinated (Janssens *et al*. 1995). It is an insect pollinated species, with honeybees as its most important pollinator. Fruits are mainly eaten and dispersed by mammals (Weeda *et al*. 1985).

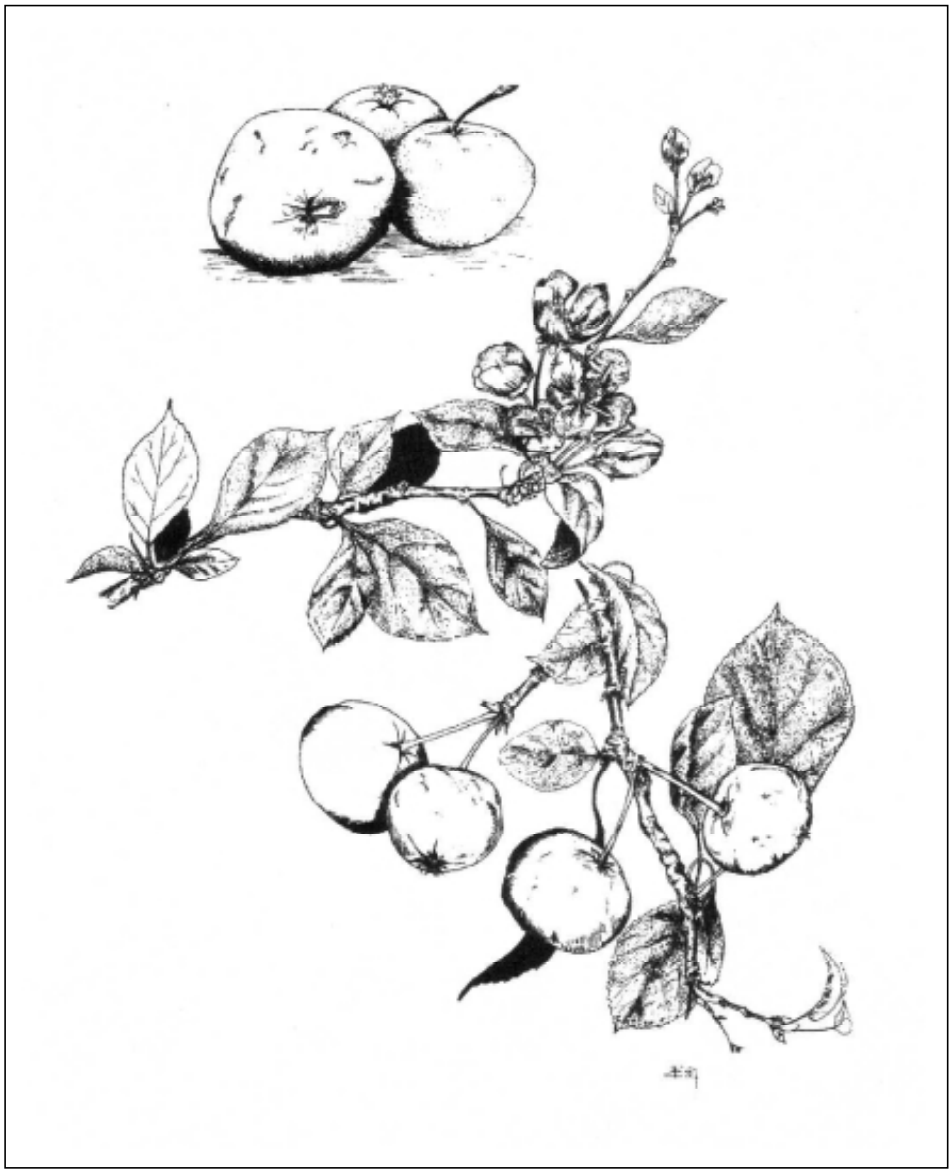


Figure 3.1: *M. sylvestris* (drawing: Filip Koopman, Institute for Forestry and Game Management).

To our knowledge, no study with molecular markers has been published up to date to analyse the amount or geographic structure of genetic diversity in *M. sylvestris*. The majority of molecular work in *Malus* has concentrated on differentiating cultivars (e.g. Hokanson *et al*. 1998 using microsatellites; Marquard & Chan 1995 using Isozymes) and phylogenetic relationships between *M. x domestica* and wild species (e.g. Savolainen *et al.* 1995, Morgan *et al*. 1994, Wagner & Weeden 2000 and Robinson *et al*. 2001). It is supposed that within species genetic diversity is high in *M. sylvestris* because (i) it is an obligatory cross-pollinator (ii) the species has a large natural distribution range, (iii) it grows under different environmental conditions, e.g. calcareous and siliceous soils (Kleinschmit *et al*. 1998).

3.1.2 Occurrence of Malus sylvestris *in Flanders*

Many European natural populations of wild apple species are being eliminated or reduced in size due to anthropogenic activities, especially during the last decennia (Hokanson *et al.* 1998, Stephan *et al.* in press). Recent inventories revealed that *M. sylvestris* has become one of the most endangered tree species in the Flemish region (Coart *et al.* 1998, Maes & Rövekamp 1998; Maes & Rövekamp 2000; Opstaele 2001; Rövekamp & Maes 1999; Rövekamp *et al*. 2000). Table 3.1 shows all currently known locations of *M. sylvestris* in Flanders (present in the database of the above mentioned inventories, provided by Kristine Vander Mijnsbrugge, Institute for Forestry and Game Management). Except for the large group of apple trees in Meerdaalwoud, wild apples are found as solitary trees or small groups of trees. No detailed inventory of Meerdaalwoud is available, but the population size is estimated at circa 100 trees. It is thought that wild apples were planted in this forest in order to provide additional food for the game and to increase the size of hunting populations (pers. comm. B. Meuleman, Forest and Green Areas Division). Furthermore, the genetic identity of the *M. sylvestris* trees listed in Table 3.1 is uncertain due to possible hybridisation with apple cultivars.

The conservation of wild apples is not only important from a nature conservation point of view but this species might also serve as source of novel genes in apple breeding programs. It can for instance be expected that natural populations, subjected to selection pressure during several generations, will be enriched for genes involved in stress and disease resistance (Schlosser *et al*. 1991), which are highly desirable traits in apple breeding programs.

Table 3.1: Locations of known prevalence for *Malus sylvestris* in Flanders. Inv NB: inventory number; Lamb X, lamb Y: Lambert X and Y coordinates; *M. x sylvestris*: phenotype intermediate between *M. sylvestris* and *M. x domestica*; *M. sp*.: not further identified *Malus* tree; Nb: number of *Malus* individuals present; n.a.: information not available.

Inv NB	Lamb X	Lamb Y	Location	Name	Species	Nb
115	179.5	184.9	Aarschot	Vorsdonkbos noord	*M. sylvestris*	2
150	208.9	192.2	Beringen	Holle weg Hek	*M. sylvestris*	2
BL05	231.1	170.4	Bilzen	Alden Biezen	*M. sylvestris*	1
601	40.8	163.7	Heuvelland	Voorbos	*M. sylvestris*	n.a.
377	89.1	162.4	Kluisbergen	Kluisbos	*M. sylvestris*	3
379	88.5	161.8	Kluisbergen	Doveleenbos	*M. sylvestris*	2
KB04	162.6	174.5	Kortenberg	Eikelenhof	*M. sylvestris*	2
326	98.3	170.7	Oudenaarde	Bos 't Ename	*M. sylvestris*	2
166	174.3	165.6	Vaalbeek	Meerdaalwoud	*M. sylvestris*	±100
146	194.6	181.5	Bekkevoort	Elzenhout/Begijnenbeek	*M. x sylvestris*	n.a.
BR23	168.0	169.3	Bertem	Bertembos	*M. x sylvestris*	1
370	92.9	162.5	Kluisbergen	Beiaardbos	*M. x sylvestris*	1
372	91.4	162.5	Kluisbergen	Feelbos	*M. x sylvestris*	1
162	169.9	167.8	Leuven	Putten v/d IJzeren Weg	*M. x sylvestris*	1
192	111.4	182.4	Oosterzele	Siemensbos	*M. x sylvestris*	n.a.
ZE02	152.1	186.9	Zemst	Bos Kollinten	*M. x sylvestris*	1
101	100.7	195.4	Gent	Regenboog	*M. sp.*	n.a.
99	100.4	195.2	Gent	Meerskant	*M. sp.*	n.a.
131	220.2	199.4	Hechtel-Eksel	Hoef	*M. sp.*	n.a.
136	218.2	199.0	Hechtel-Eksel	Spiekelspade	*M. sp.*	n.a.
144	106.9	187.5	Merelbeke	Mijleke	*M. sp.*	n.a.

3.1.3 Relationship with Malus x domestica

Hybridisation and introgression are supposed to have played important roles, both in the evolution of the genus *Malus* as a whole and in the origin of the domesticated edible apple (*M. x domestica* Borkh) (Phipps *et al*. 1990). Different species have been proposed as progenitors for the domesticated apple, including *M. sylvestris*, *M. sieversii*, *M. orientalis*, *M. baccata*, *M. mandsurica, M. prunifolia, M. floribunda, M. praecox* and *M. dasyphyllus* (e.g. Morgan *et al*. 1994, Hokanson *et al*. 1998, see Robinson *et al.* 2001 for a review). However, recent research pointed out that *M. sylvestris* probably might not have had much, if any, contribution to the domesticated apple gene pool. *M. sieversii* from the Kazachstan region has been identified as the most probable maternal ancestor of the domesticated apple (based on sequence information of the *mat*K region, Robinson *et al*. 2001) and *M. sieversii* is also more closely related to *M. x domestica* than *M. sylvestris,* as was revealed based on isozymes (Wagner & Weeden 2000). This information refuted the

widespread, intuitively attractive idea that edible apple cultivars had been derived from our native wild species with small and sour fruits. Nevertheless, although the earliest selections of domesticated apples could have come directly from *M. sieversii* without involvement of other species, later hybridisations could have been important in the creation of new cultivars. Until more variable, phylogenetic informative markers are found, the hypothesis about hybridisation and the origin of the domesticated apple cannot be rejected completely (Harris *et al*. 2002).

However, crosses between *M. sylvestris* and *M. x domestica* are possible (Korban 1986) and therefore it is currently assumed that the gene flow from *M. x domestica* to the *M. sylvestris* gene pool is more significant than in the opposite direction. The occurrence in nature of many phenotypic intermediate forms that resemble *M. x domestica* has led to the hypothesis that by frequent hybridisation, the wild species has been replaced by a hybrid swarm, consisting of hybrid genotypes with a significant admixture of the cultivated genetic information (Kleinschmit *et al*. 1998). However, the morphological variability within *M. sylvestris* is large, what makes it difficult to discriminate between 'genuine' wild individuals and hybrids with the cultivated gene pool. The main morphological characters that discriminate between wild and cultivated apple are the hairiness of inferior leaf surfaces, presence of thorns on twigs and form, colour and taste of the fruits (see Wagner 1996 for a detailed review in German on this topic). Since fruits on apple trees in dense forests are mostly absent, fruit characteristics can seldom be used for classification. Therefore, the hairiness of inferior leaf surfaces is considered one of the most useful discriminatory phenotypic characteristics between wild apples and edible cultivars (*M. x domestica*: felted leaf surface, *M. sylvestris*: sparsely hairy (on veins) in spring and hairless in autumn; Remmy & Gruber 1993; Wagner 1996 and 1998). Individuals with different degrees of leaf hairiness are known to occur in Belgian forests and it is sometimes difficult to apply this criterion to classify a specific tree. A preliminary study, conducted with isozymes, showed promising results to discriminate between wild and cultivated apple genotypes (Wagner & Weeden 2000) but most information that molecular markers could provide on this issue, remains untapped.

3.2 Present apple research and objectives

This study aimed at tackling the issues that have to be clarified prior to the development of a sound conservation program for the endangered wild apple in Belgium. The objectives of the study were to

(1) investigate the level of genetic differentiation between Belgian wild apple trees and apple cultivars;
(2) specifically test for the occurrence of cultivated genotypes or hybrid genotypes in the wild;
(3) investigate the population genetic structure of wild apple trees from Belgian and other origins and test whether the Belgian trees form a distinct gene pool.

With these objectives in mind, apple trees were inventoried and sampled in Belgian forests. Also genotypes with felted hairy leaves were sampled in the wild. Additional apple trees from four origins were also analysed in order to generate reference groups: (1) edible apple cultivars; (2) ornamental cultivars; (3) wild individuals of *M. sylvestris* from Germany and (4) wild individuals from a private collection.

It is clear that the discrimination between genuine wild genotypes, cultivated genotypes and hybrids is the key issue that has to be solved if the objectives listed above are to be met. Like the majority of studies concerning hybridisation between crops and wild relatives, hypotheses regarding hybridisation between *M. sylvestris* and cultivated apples are based on morphological evidence. However, (introgressive) hybridisation is not necessarily indicated by the phenotypic occurrence of the characters of one taxon in another. Identical characters in cultivated varieties and their wild relatives may occur as a result of either phenotypic plasticity, convergent evolution, or simply common ancestry (Linder *et al*. 1998). Furthermore, individuals from hybrid swarms that obtained most of their genes from one of the parental taxa are often morphologically indistinguishable from that parental taxon (Allendorf *et al*. 2001). Molecular markers have become powerful tools to establish the extent of hybridisation processes (Linder *et al*. 1998; see Jarvis & Hodgkin 1999 for a review) as well as to provide the background knowledge necessary to implement conservation genetics programs (Haig 1998, Smulders *et al.* 2000). The advantages of using molecular markers in these fields are that most molecular markers do not exhibit plasticity and can generate information on many different loci. Several studies, however, have demonstrated that the results obtained using a given marker system may not fully agree with the results obtained using a different marker system, especially for loosely related genotypes (Roldán-Ruiz *et al.* 2001). This could be due for instance to differences among marker systems in the targeted genome (cytoplasmic vs. nuclear) and

genomic regions, in mutation rate, or in dominance characteristics. King & Burke (1999) recommended combining the information derived from different types of molecular markers to minimise the risks related to misinterpreting the data or to the inevitable drawbacks of each marker technique.

It was therefore decided to investigate the possibilities of two currently often used marker systems, namely AFLP and microsatellites, in relation to the objectives defined. AFLP markers (Vos *et al.* 1995) have recently been adopted in the fields of ecological genetics and conservation of plant genetic resources. The qualities of this technique can be found in the high reproducibility of the generated fingerprint patterns and the high numbers of markers produced without much prior information. This makes the technique well suited for studying genetic variation (e.g. Roldán-Ruiz *et al*. 2000; De Riek *et al*. 1999) and hybridisation processes (Jarvis & Hodgkin 1999; Coulibaly *et al.* 2002). However, AFLP suffers from several drawbacks: (1) if scored as dominant markers, difficulties arise for estimation of population genetic parameters (Mueller & Wolfenbarger 1999); (2) the occurrence of size homoplasy (i.e. non-homologous AFLP fragments of the same size) which varies as a function of fragment size (Vekemans *et al*. 2002).

Due to their co-dominant mode of inheritance and amenability to high throughput analysis, microsatellite markers have become an important tool for studying genetic diversity and identity in both natural populations and in gene bank collections (Djè *et al*. 2000; Heuertz *et al*. 2001; Hokanson *et al.* 2001). Given these characteristics and the ample availability of primer sequence information for *M. x domestica* (Guilford *et al*. 1997; Hokanson *et al*. 1998; Gianfrancheschi *et al*. 1998), this technique represents a valuable approach for studying the relationship between apple crops and their wild relatives. Furthermore, many SSR loci for apple have been mapped (Gianfrancheschi *et al*. 1998; Maliepaard *et al*. 1998; Liebhard *et al*. 2002), enabling the choice of a set of unlinked microsatellite loci.

An additional requirement for the correct interpretation of DNA-marker data is the use of statistics that can be compared between dominant (AFLP) and co-dominant (microsatellite) markers (Lynch & Milligan 1994; Vekemans *et al*. 2002). Information on the marker systems and data analysis applied is provided in Chapter 2.

Furthermore, a combination of nuclear and chloroplast data may provide complementary views of the hybridisation process between *Malus* species. In particular, Robinson *et al*. (2001) hypothesized that a specific region of the chloroplast *mat*K gene could be used to distinguish between edible apple cultivars and *M. sylvestris* individuals. In this gene, two duplications, located 39bp upstream from the 3'-end of the *mat*K-coding region are present. Duplication I is an imperfect

8bp duplication, duplication II is a perfect 18bp duplication. Duplication I seems to be present in most apples of the section *Malus* and is therefore not useful to discriminate between *M. sylvestris* and *M. x domestica* individuals. However, all but one of the nine *M. x domestica* trees included in the study by Robinson *et al.* (2001) were shown to have duplication II and this duplication was absent in the two *M. sylvestris* individuals included in the study. To establish the validity of this trait to differentiate between *M. sylvestris* and *M. x domestica*, we sequenced this region of the *mat*K gene for 18 additional individuals. Subsequently, a marker able to detect the presence or absence of duplication II as a length polymorphism was developed. All wild individuals and edible cultivars were typed for this locus (further referred to as *mat*KdupII locus, see chapter 2 for details on the techniques applied).

Finally, the chloroplast information is contrasted with the results derived from AFLP and SSR genotyping and the conclusions based on DNA-marker information are compared with the trait of hairiness of inferior leaf surfaces.

Part of this research has been published:
E. Coart, X. Vekemans, M.J.M. Smulders, I. Wagner, J. Van Huylenbroeck, E. Van Bockstaele and I. Roldán-Ruiz (2003). Genetic variation in the endangered Wild apple (*Malus sylvestris* (L.) Mill.) in Belgium as revealed by AFLP and microsatellite markers. Molecular Ecology, 12, 845-857.

3.3 Studied genotypes

Locations and sample sizes of the samples included in this study are presented in Table 3.2:

- Forty-five Belgian putative wild apple trees were sampled at three different locations in forests, including trees with different degrees of leaf hairiness.
- Six presumed hybrids, all with felted hairy leaves, were collected. These individuals were sampled near forests, on roadsides and in meadows.
- Supplementary genotypes of *M. sylvestris* were obtained from various sources: 28 individuals from two German gene banks, and three European presumed *M. sylvestris* genotypes from a private Belgian collection.
- In addition, grafts from 11 old edible apple cultivars were obtained from the gene bank of CRA-Gembloux: 'Gris Braibant', 'Belle Fleur De Brabant', 'Oogstappel', 'Gueule de Mouton', 'Court Pendu Rose', 'Calville Des Prairies', 'Pomme De Sucre', 'Reinette De Wattripont', 'Pomme De Douce', 'Blanc Braibant' and 'Jerusalem'. These ancient cultivars were chosen because they were cultivated at the time of germination of the putative wild Belgian apple trees sampled (the average age of the sampled apple trees in forests is estimated at circa 100 years). One additional cultivar, 'Elstar', was added as a standard genotype for the microsatellite analysis (allele sizes of 'Elstar' are available for all used microsatellite loci, K. Kenis, KULeuven).
- Finally, 28 ornamental apple cultivars were sampled from the *Malus* collection of the Department of Plant Genetics and Breeding (DvP). These ornamental cultivars have been derived from different *Malus* species, and are used in the breeding program of ornamental apple trees. The sampled ornamental cultivars included many different species: *M. mandsurica*, *M. x zumi* 'Calocarpa', *M .*'Van Eseltine', *M.* 'Golden Hornet', *M.* 'Maypole', *M.* 'Adirondack', *M. baccata* 'Street parade', *M. baccata* 'Yellow Syberian', *M.* 'Professor Sprenger', *M.* 'Red Sentinel', *M .*'John Downie', *M.* 'Liset', *M.* 'Butterball', *M.* 'Profusion', *M.* '*D.V.P. Obel*' *(*Red Obelisk®), *M.* 'Eleyi', *M.* 'Evereste', *M. floribunda*, *M.* 'Hartuigii', *M. pumila* 'Hopa New', *M.* 'Makamik', *M.* 'Neville Copeman', *M.* 'Red Jade', *M.* 'Royalty', *M. toringo* 'Rosea' *M. toringo* var. *sargentii* and *M. tschonoskii.*

Most of the genotypes included in this study were grafted at the nursery of the DvP (column 5 of Table 3.2). For genotypes that were not grafted in the nursery, DNA had to be extracted out of often badly preserved leaves collected in the field, and this was of insufficient quality for AFLP analysis. The DNA was of appropriate quality for SSR analysis and for amplification of the *mat*KdupII locus, which are less sensitive to DNA quality. As a result, for the 10 individuals of WBVi and WPr no clean AFLP fingerprints could be obtained after several trials and only information on *mat*KdupII and SSR loci is available (column 6 of Table 3.2).

Table 3.2: Description of the sampled apple genotypes. N: number of genotypes sampled; Mat: G (grafts) or L (leaves) available.

Origin	Code	Coor-dinates	N	Mat	Typed for	Description of sampling site
WILD BELGIAN						
- Meerdaal	WBM	4°41'/ 50°48'	29	G	AFLP, SSR, *mat*KdupII	*M.sylvestris* genotypes sampled in one forest, more genotypes present (estimated at 100 individuals, no inventory available), total area of collection: 6 km^2
- Voeren	WBVo	5°52'/ 50°44'	9	G	AFLP, SSR, *mat*KdupII	*M.sylvestris* genotypes sampled in forests, all known genotypes in neighbourhood sampled, total area of collection: 3 km^2
- Viroin	WBVi	4°28'/ 50°04'	7	L	SSR, *mat*KdupII	*M.sylvestris* genotypes sampled in forests, presumably more genotypes present (but no inventory of area available), total area of collection: +/- 20 km^2
BELGIAN HYBRIDS						
- Limburg	WBL	5°15'/ 51°03'	6	G	AFLP, SSR, *mat*KdupII	Presumed hybrids between *M.sylvestris* and *M. x domestica,* sampled near forests, total area of collection: 6 km^2
WILD GERMAN						
- Gene bank NFV Escherode (G)	WGL	/	22	G/L	AFLP, SSR, *mat*KdupII	*M.sylvestris* genotypes of German Federal District Lower Saxony and Shleswic Holstein, 9 genotypes grafted
- Gene bank Forstgenbank Arnsberg (G)	WGF	/	6	G	AFLP, SSR, *mat*KdupII	*M.sylvestris* genotypes of German Federal District North Rhein West Phalia
WILD PRIVATE COLLECTION*						
- Private collection (B, F, GR)	WPr	/	3	L	SSR, *mat*KdupII	*M.sylvestris* genotypes from the south of Belgium, south of France and Greece
EDIBLE APPLE CULTIVARS						
- Gene bank CRA-Gembloux	ECV	/	11	G	AFLP, SSR, *mat*KdupII	Old apple cultivars
ORNAMENTAL APPLE CULTIVARS						
- Gene bank DvP	OCV	/	28	G	AFLP, SSR	Ornamental cultivars, including many different species

*Origin: B: Belgium, F: France, GR: Greece.

3.4 Results

3.4.1 Description of hairiness of inferior leaf surfaces

Leaf hairiness was described in autumn for the 70 apples trees (wild trees and edible cultivars) that were grafted in the nursery at DvP. The results are summarised in Table 3.3. The majority of putative wild trees had hairless (score 0) or sparsely hairy (score 1) leaves; only three putative wild individuals were given score 2 (moderately hairy). All edible cultivars and all but one hybrid had felted inferior leaf surfaces (score 3), one hybrid was given score 2.

Table 3.3: Description of hairiness of inferior leaf surfaces (according to the scale described by Wagner 1998). Sample codes correspond with column 2 of Table 3.2.

Sample	Score 0	Score 1	Score 2	Score 3
Belgian Wild apples				
WBM	20	8	1	0
WBVo	5	3	1	0
German Wild apples				
WGF	5	1	0	0
WGL	6	0	1	0
Belgian hybrids				
WBL	0	0	1	5
Edible cultivars				
ECV	0	0	0	11

Score 0: hairless; score 1: sparsely hairy; score 2: moderately hairy; score 3: felted hairy leaf surface. Figures are number of trees with a given score for hairiness.

3.4.2 Allelic variation at AFLP and microsatelite loci

Three primer combinations that had previously been chosen for AFLP analysis using ^{33}P labelled primers (results not shown) were tested using fluorescently labelled primers and this resulted in clear and scoreable fingerprints. The primer combinations used were *Eco*RI-ACA + *Mse*I-CAG, *Eco*RI-ACT + *Mse*I-CTG and *Eco*RI-AGC + *Mse*I-CAT. A reproducibility test was performed for five genotypes with these three primer combinations. For each genotype, three independent DNA extractions were carried out and AFLP fingerprints were generated using all different DNA templates (a total of 45 AFLP fingerprints). Mean reproducibility values (calculated as the percentage of markers that were identical in the three repeats for the same plant) were high with respectively 97.5, 97 and 98% of reproducibility for the different primer combinations. The use of three AFLP primer combinations on 110 *Malus* genotypes resulted in 139 scoreable markers, of which 126 (91%) were polymorphic (least common state at least 5%). All trees were characterised by a unique banding pattern. A negative correlation between fragment sizes and frequencies (-0.29, $p<0.01$) was detected, which indicates that some degree of

homoplasy might be present in the dataset (Vekemans *et al*. 2002). Therefore, data analysis was repeated with only AFLP fragments larger than 150 bp. For this reduced data set (90 markers), correlation between fragment sizes and frequencies was also negative (-0.17, p=0.10) and differentiation values for all groups considered were similar to those obtained with the complete dataset (results not shown), suggesting that the potential presence of size homoplasy of AFLP fragments does not result in underestimating genetic divergence between samples. The low average pair-wise correlation value between markers of 0.014±0.174 (S.D.) indicates that only a limited amount of information in the data set is redundant.

All 12 SSR loci analysed in this study revealed to be highly polymorphic, displaying many alleles (minimum 11 and maximum 35 alleles per locus) and a wide size range of PCR products (Table 3.4). More alleles and wider size ranges were recorded than in the original publications, where only a limited set of cultivated varieties and species was used.

Table 3.4: Allelic diversity at microsatellite loci scored in 114 apple genotypes.

			Original publication		**This study**	
Locus	**Repeat motif**	**Chromosome**	**Number of alleles**	**Range of sizes (bp)**	**Number of alleles**	**Range of sizes (bp)**
NZ02b01[a]	$(GA)_{14}$	15 [b]	7	212-238	18	194-246
NZ04h11[a]	$(GA)_{23}$	9 [b]	6	201-233	13	200-242
NZ05g08[a]	$(GA)_{18}$	4 [b]	6	115-141	19	90-156
NZ23g04[a]	$(GA)_{19}$	6 [b]	9	84-116	17	82-125
NZ28f04[a]	$(GA)_{18}$	12 [b]	4	98-112	11	91-123
CH01h10[c]	$(AG)_{21}$	8 [d]	7	93-119	26	88-147
CH01e12[c]	$(AG)_{32}$	8 [e]	8	243-248	14	223-275
CH01f02[c]	$(AG)_{22}$	12 [d]	11	168-222	23	153-227
CH01h01[c]	$(AG)_{25.5}$	17 [d]	9	107-141	21	89-145
CH02b12[c]	$(GA)_{26}$	5 [d]+10 [e]	8	124-142	18	109-159
CH02c06[c]	$(GA)_{21}$	2 [d]	10	216-254	35	206-309
CH02d12[c]	$(GA)_{17}$ $(GA)_{19}$	11 [d]	9	175-205	20	175-219

Normalised nomenclature of the loci follows Liebhard *et al*., 2002. [a] From Guilford *et al*., 1997; [b] Mapped in Maliepaard *et al.*, 1998; [c] From Gianfranceschi *et al*., 1998; [d] Mapped in Liebhard *et al*., 2002; [e] Mapped by Eric van de Weg *et al*. (in preparation).

3.4.3 Relationships among genotypes using nuclear markers

3.4.3.1 Overall relationships

Both PCO plots (based on respectively AFLP and SSR data, Figures 3.2a and 3.2b) show congruent groupings. The first two PCO axes explain 6 and 4% of the variation for the PCO derived from AFLP markers and 9 and 5% for the PCO calculated from SSR data. Individuals from ornamental cultivars and those from

edible apple cultivars clearly form two separate groups. Most individuals from the German gene banks, from the Belgian wild samples and the three European wild individuals from the private collection (collected at different geographical locations and typed only for SSR) are positioned within a third group, suggesting that this group represents the wild genotypes of *Malus sylvestris*. Within this group, the individuals were not distributed according to their geographic origin and not even German and Belgian origins could be differentiated with either SSR or AFLP markers. The two PCO plots also clearly indicate that the presumed hybrid trees from Limburg (WBL) are indistinguishable from edible apple cultivars. Moreover, one individual from the German gene bank (from WGL) and one individual from the Belgian forests (from WBM) seem to be more related to the edible apple cultivars than to the *M. sylvestris* group on both PCO plots.

The main difference between AFLP-derived and the SSR-derived PCO plots is the larger heterogeneity of the groups, particularly the wild group, in the plot based on SSR data. This may be due to the high number of SSR alleles present within this group: the number of alleles present at each locus is higher than for the edible cultivar group and the average number of alleles per locus is the highest within the wild gene pool (the average number of alleles per locus is 13, 12.5 and 6.6 within the wild, ornamental cultivar and edible cultivar group, respectively). This leads to very low similarities between genotypes.

The correlation between the pairwise relatedness values calculated on the basis of AFLP data and SSR data for 103 genotypes was tested using a Mantel test. A highly significant positive correlation was found between both datasets ($r = 0.214$, P-value = 0.0001).

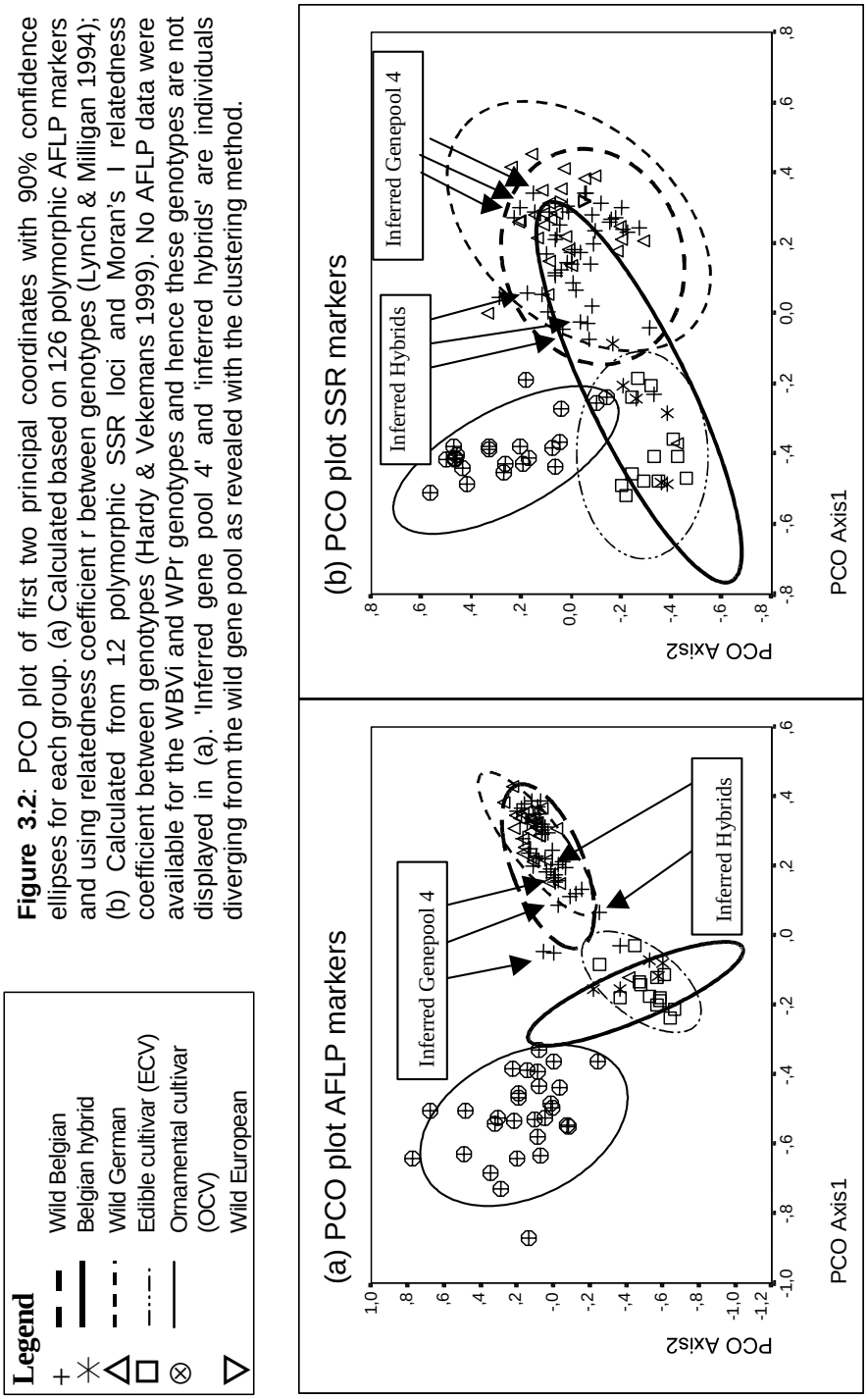


Figure 3.2: PCO plot of first two principal coordinates with 90% confidence ellipses for each group. (a) Calculated based on 126 polymorphic AFLP markers and using relatedness coefficient r between genotypes (Lynch & Milligan 1994); (b) Calculated from 12 polymorphic SSR loci and Moran's I relatedness coefficient between genotypes (Hardy & Vekemans 1999). No AFLP data were available for the WBVi and WPr genotypes and hence these genotypes are not displayed in (a). 'Inferred gene pool 4' and 'inferred hybrids' are individuals diverging from the wild gene pool as revealed with the clustering method.

3.4.3.2 Inference of genetic structure and assignment of genotypes to inferred gene pools

Using the model-based clustering method of Pritchard *et al*. (2000) on the SSR data, the highest estimate of the likelihood of the data, conditional on a given number of clusters, was obtained when clustering all genotypes into five gene pools. The assignment of individuals from the different samples to these five gene pools is given in Table 3.5 and described below:

1. **Gene pool 1** comprises the majority of wild individuals, both from Belgian and German origin. Although admixture of genotypes between gene pools is allowed by the method, most wild individuals were fully assigned to this gene pool. Strikingly three individuals (from samples WBVi, WBVo and WBM) are only partially assigned to this wild gene pool (respectively 50%, 74% and 71%) and have a significant proportion of their genetic information assigned to gene pool 2 (respectively 38%, 24% and 23%). These genotypes are indicated as 'inferred hybrids' in figure 3.2.
2. **Gene pool 2** comprises all edible cultivars (ECV), all presumed Belgian hybrids (WBL), three ornamental cultivars (OCV), one Belgian individual sampled in the wild (from sample WBM), and one German individual (from sample WGL, a *M. sylvestris* gene bank). The latter two individuals were those identified as being more related to edible cultivars in the PCO plots (Figure 3.2). All these individuals (even the presumed hybrids (WBL) and the two individuals sampled as wild) were assigned completely to this gene pool, showing no evidence of genetic admixture.
3. **Gene pool 3** comprises most of the ornamental cultivars (OCV), with exception of three individuals assigned to gene pool 2. Four genotypes indicate admixture between the edible and ornamental cultivars with respectively 18, 24, 37 and 38% of their genes assigned to the edible cultivar gene pool.
4. **Gene pool 4** consists of only three closely related wild Belgian trees (from sample WBVo), the three genotypes are assigned completely to this cluster. These trees are indicated as 'inferred gene pool 4' in figure 3.2.
5. **Gene pool 5** consists of only one ornamental cultivar, *M. tschonoskii*, assigned completely to this cluster. This Japanese apple is the only individual of the studied trees that does not belong to the section *Malus* (but to the section *Docyniopsis,* classification according to Phipps *et al*. 1990).

The same clustering method was applied on AFLP data by treating each class of genotypes as being a haploid allele, but no admixture can be taken into account for dominant data. This analysis resulted in very similar clustering of genotypes (results not shown).

Table 3.5: Summary of the results obtained using the assignment procedure based on SSR data. Figures are the proportion of estimated membership to each of five inferred gene pools for genotypes of a given category of origin. Category German wild apples: samples WGL and WGF; Belgian wild apples: samples WBM, WBVo and WBVi.

	Inferred gene pools				
Category of origin	**1**	**2**	**3**	**4**	**5**
Belgian Wild apples	**0.863**	0.048	0.006	**0.077**	0.005
Belgian hybrids (WBL)	0.009	**0.967**	0.005	0.013	0.006
German Wild apples	**0.942**	0.041	0.004	0.008	0.005
Edible cultivars (ECV)	0.013	**0.974**	0.006	0.004	0.004
Ornamental cultivars (OCV)	0.007	**0.168**	**0.748**	0.011	**0.067**

3.4.4 Relationships among samples using nuclear markers

Based on the results shown above, reduced samples were defined. The two 'wild' individuals that were assigned completely to the edible cultivar gene pool were excluded from the dataset. The sample of presumed hybrids (WBL) that was assigned completely to the edible cultivar gene pool was included for studying relationships between samples but was not included in the analysis of genetic diversity within and among wild samples.

3.4.4.1 Estimation of sample differentiation parameters

Results of neighbour-joining trees based on AFLP and SSR data are shown in Figure 3.3. As with the PCO plots, highly concordant results were found based on both marker techniques. Very similar trees were also obtained with both distance measures used to analyse the SSR data (Nei standard genetic distance and $\delta\mu^2$). In all trees, samples from the wild clustered separate from cultivars (both ornamental and edible cultivars) with bootstrap support values ranging from 77 to 100%. As expected from the results presented above, only the formerly presumed hybrid trees from Limburg (sample WBL) clustered together with the edible cultivars (bootstrap support values from 70 to 100%). Both German samples clustered together in the trees based on AFLP data and SSR data (although with low bootstrap values, 55 and 45%, respectively) but only when distances were computed according to Nei. The sample of ornamental cultivars (OCV) was always positioned in between the wild samples and the edible cultivars (ECV).

Divergence between wild samples and edible cultivars was further investigated (Table 3.6). The differentiation among all samples was significant for both marker

systems applied (F_{ST} = 0.14, P < 0.0001 for AFLP data; G_{ST} = 0.097, P < 0.05 for SSR data). Also differentiation among wild samples was small but significant (AFLP: F_{ST} =0.046 P < 0.0001, SSR: G_{ST} = 0.060, P < 0.05). Of the total differentiation among wild samples, AMOVA-analysis based on SSR data attributed 43% (P = 0.016) to the divergence between German and Belgian origins, while 57% (P < 0.0001) was explained by differentiation of samples within origins. Divergence between the wild gene pool and edible cultivars was large based on both marker systems (AFLP: F_{ST} = 0.19, P < 0.0001, SSR: G_{ST} = 0.11, P < 0.05).

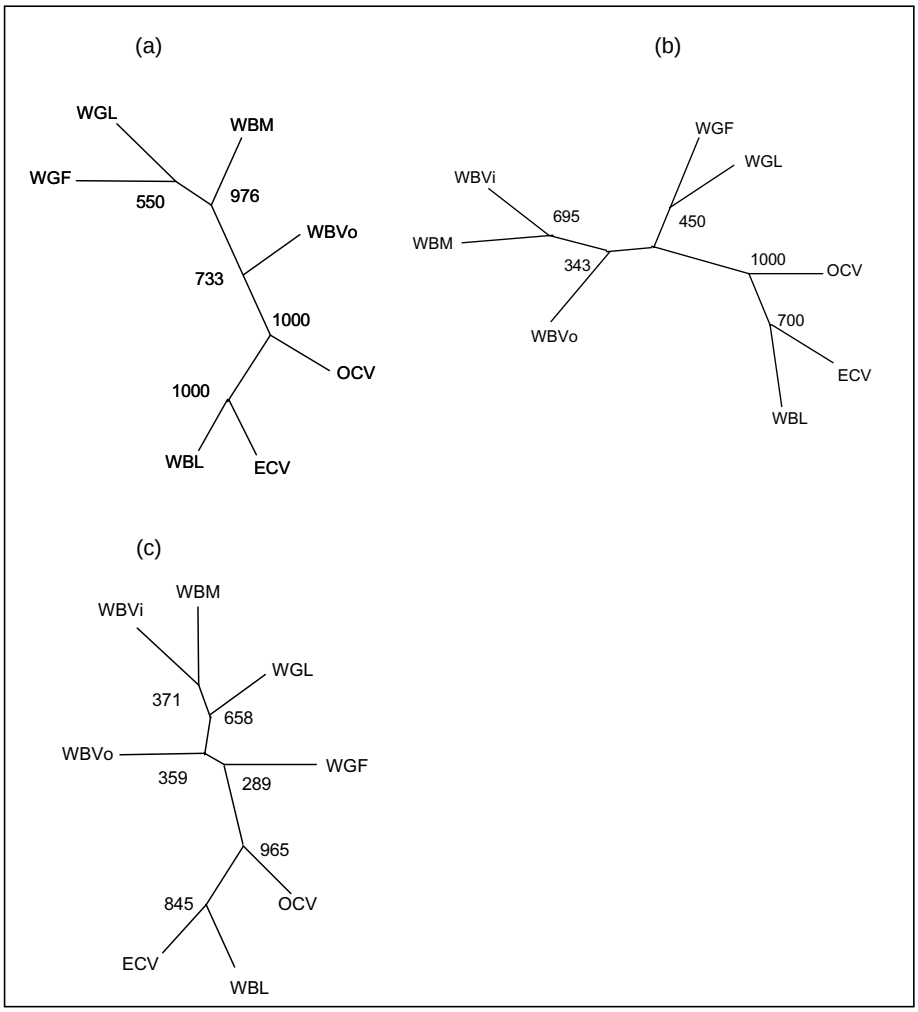


Figure 3.3: Neighbour-joining trees with bootstrap support values at forks (based on 1000 bootstraps). (a) calculated from AFLP data and Nei's genetic distance; (b) from SSR data and Nei's standard genetic distance; (c) from SSR data and $\delta\mu^2$ distance. Sample codes are described in Table 3.2.

Table 3.6: Analysis of genetic differentiation based on 139 AFLP markers and 12 SSR loci (i) among all samples except for the ornamental sample (WBM, WBVo, WBL, WGL, WGF, Wvi (for SSR only), ECV) (ii) among wild samples only (WBM, WBVo, WVi (for SSR only), WGL, WGF) (iii) between wild (WBM, WBVo, Wvi (for SSR only), WGF, WGL) and cultivated (ECV) origins.

	AFLP						**SSR**		
Comparison	N	H_t	H_w	H_b	F_{st}	P-value	N	G_{st}	P-value
(i) Among all samples	6	0.2871	0.2470	0.0401	0.140	<0.0001	7	0.0970	<0.05
(ii) Among all wild samples	4	0.2303	0.2197	0.0106	0.0464	<0.0001	5	0.0602	<0.05
(iii) Between wild and edible cultivated	2	0.2870	0.2325	0.0547	0.1880	<0.0001	2	0.1085	<0.05

N: number of samples; H_t: total diversity; H_w: average diversity within populations; H_b: average diversity between populations; F_{st}: differentiation between defined groups.

A Mantel test on distance matrices between samples (Nei's distance) based on AFLP and SSR data was performed, resulting in a highly significant positive correlation with a much higher r value than that based on relatedness between individual genotypes (r = 0.872, P-value = 0.0035). These computations are based on the 7 samples typed for both marker systems.

3.4.4.2 Nuclear genetic variation within wild samples and comparison with cultivated samples

Diversity statistics are summarised in Table 3.7. High genetic variation at microsatellite loci was observed, with a mean number of alleles per locus equal to 7.2 (range: 4.8 to 9.9) and a mean gene diversity of 0.721 (range: 0.703 to 0.810). Very similar levels of genetic variation were found in all wild samples (CV of gene diversity is only 7.7%). The edible cultivar sample (ECV) and the presumed hybrid sample (WBL) have similar gene diversity (respectively 0.78 and 0.71), whereas higher diversity was detected within the ornamental cultivars (0.84). High variation at AFLP loci was also recorded, with on average 76.6% of polymorphic loci within wild samples. It should be noted however that only loci polymorphic in the overall data set were included in the analysis. Very similar values of gene diversities (mean 0.225, range: 0.204 to 0.253) were found in all wild samples (CV is only 6.5%). The hybrid sample, edible and ornamental cultivar samples have higher gene diversities (respectively 0.30, 0.27 and 0.29).

The average multilocus inbreeding coefficient (F_{is}) for all wild samples is 0.105, showing a significant (P<0.001) overall departure from Hardy Weinberg proportions

with an excess of homozygotes. A significant excess of homozygote genotypes was also detected within each wild sample (Table 3.7).

Table 3.7: Statistics of genetic diversity within samples of *Malus sylvestris* calculated from 139 AFLP and 12 microsatellite loci.

	Microsatellites					AFLP			
Samples	**N**	**A**	H_O	H_E	F_{is} ¥	**N**	**NPL**	**PLP**	H_j
				Wild samples					
WBVo	9	6.7	0.691	0.742	0.065*	9	112	80.6	0.233
WBM	28	9.9	0.706	0.810	0.127***	27	108	77.7	0.204
WBVi	7	6.2	0.708	0.784	0.102*	/	/	/	/
WGL	21	8.3	0.652	0.759	0.140***	21	98	70.5	0.208
WGF	6	4.8	0.636	0.703	0.092**	6	108	77.7	0.253
Mean		7.17	0.679	0.721	0.105***		106.5	76.63	0.225
S.D.		1.96	0.033	0.056	0.030		3.0	4.31	0.015
C.V.		0.27	0.049	0.077	0.29		0.03	0.06	0.065
WBL	6	5.4	0.659	0.709	0.134*	5	107	77.0	0.298
				Cultivated samples					
ECV	11	6.4	0.729	0.775		11	122	87.8	0.265
OCV	21	11.7	0.749	0.841		29	116	83.5	0.287

N: number of individuals typed; *A*: mean number of alleles per locus; H_O: average proportion of heterozygotes; H_E: average gene diversity; F_{is}: average inbreeding coefficient, ¥: Exact test of departure from Hardy-Weinberg genotypic proportions: *P<0.05, **P<0.01, ***P<0.001, other values not significant; NPL: number of polymorphic loci at the 5% level; PLP: proportion of polymorphic loci; *Hj*: same as H_e for AFLP markers.

3.4.5 Relationships among genotypes using the cp-marker matKdupII

3.4.5.1 Sequencing of the matK gene

The *mat*K gene was sequenced for 18 apple genotypes: 9 wild apples (3 from WBM, 3 from WBV, 2 from WGF and 1 from WGL), 2 hybrids (sample WBL), 3 ornamental cultivars (sample OCV) and 4 edible cultivars (3 from sample ECV and cultivar 'Elstar'). In total, 1730 bp could be compared for all genotypes. Figure 3.4 shows the consensus sequence. Within the group of *M. sylvestris* and *M. x domestica* only two point mutations (at positions 338 and 1027) and one insertion of 18bp (duplication II from Robinson *et al.* 2001) were identified. The ornamental cultivar *M. tschonoskii* has two more point mutations (position 1170 and 1553) and one deletion of 7bp (duplication I described by Robinson *et al*. 2001).

Four groups of genotypes, that share identical *mat*K sequences can be defined within the 18 *Malus* genotypes analysed based on this information:

1. **Group A**: duplications I and II present, one point mutation at position 1027 (C → T). This group consists of the modern edible cultivar 'Elstar', the two genotypes from the hybrid sample WBL and five genotypes from wild samples.
2. **Group B**: duplication I present, duplication II absent, one point mutation (at position 338: T → G). This group consists of three old edible cultivars from the collection of CRA-Gembloux (Belgium).
3. **Group C**: duplication I present, duplication II absent, but the point mutation at position 338 (as in group B) is absent. This group consists of three Belgian and two German wild apples and two ornamental varieties (*Malus* 'Red Sentinel' and 'Red Obelisk').
4. **Group D**: Both duplications are absent and two additional point mutations are observed (at position 1170 T → G and 1553 T → A). This group consists only of the Japanese ornamental *Malus tschonoskii* (belonging to the section *Docyniopsis*, Phipps *et al*. 1990) that also made up a separated gene pool based on SSR results.

3.4.5.2 Determination of the presence of duplication II using the marker matKdupII

Sequencing information was used to develop a DNA-marker located in this region that allowed to visualize the presence or absence of duplication II in the *mat*K gene as a length polymorphism. The presence of duplication II was further investigated in all sampled wild apple genotypes and edible cultivars by amplifying the fragment of the *mat*K region containing the location of duplication II. Table 3.8 summarizes the results. Strikingly, duplication II which is considered to be diagnostic for the *M. x domestica* gene pool was found to be present in 37.8% (17 of 45) of the Belgian wild *M. sylvestris* genotypes whereas only one German *M. sylvestris* was found with this duplication. Furthermore, more than half of the cultivated varieties (7 of 12 analysed or 58.3%) did not have duplication II.

```
                    10         20         30         40         50         60
matK ABCD   TTGTTCAAAT AAAAAATGGA AGAATTTCAA GGATATTTAG AACTAGATAG ATATCAGCAA
                    70         80         90        100        110        120
matK ABCD   CATGACTTCC TATACCCACT TATCTTTCGG GAGTATATTT ATGCACTTGC TCATGATCAT
                   130        140        150        160        170        180
matK ABCD   GGTTTAAATA GATCGATTTT GTTGGATAAT GTAGGTTATG ACACTAAATA TAGTTTACTA
                   190        200        210        220        230        240
matK ABCD   ATTATAAAAC GTTTAATTAG TCGAATGTAT CAACAGAATC ATTTGATAAT TTCCGCTAAT
                   250        260        270        280        290        300
matK ABCD   GATTCTAACC AAAATAAATT TTTTGGGTAC AACAAAAATT TGTATTCTCA AATGATGTCG
                   310        320        330        340        350        360
matK ACD    GAGGGATTTG CAGTCATTGT GGAAATTCCG TTTTCCCTAC GATTAGTATC TTCCTTAGAG
matK B      GAGGGATTTG CAGTCATTGT GGAAATTCCG TTTTCCCGAC GATTAGTATC TTCCTTAGAG
                   370        380        390        400        410        420
matK ABCD   GCGACAGAAA TCGTAAAATC TTATAATTTA CGATCAATTC ATTCAATATT TCCTTTTTTA
                   430        440        450        460        470        480
matK ABCD   GAGGACAAAT TCCCACATTT AAATTATGTA TCAGATGTAC TAATACCCTA CCCCATTCAT
                   490        500        510        520        530        540
matK ABCD   CTGGAAATCT TGGTTCAAAC CCTTCGCTAT TGGGTGAAAG ATCCCTCTTC TTTACATTTA
                   550        560        570        580        590        600
matK ABCD   TTACGACTCT TTCTTCACGA GTATTCTAAT TGGAATAGTC TTATTACTCC AAAAAAAATT
                   610        620        630        640        650        660
matK ABCD   ATTTTTTCAA AAAGTAATCC ACGATTATTC TTGCTCCTAT ATAATTCTCA TGTATGTGAA
                   670        680        690        700        710        720
matK ABCD   TACGAATCCA TTTTACTTTT TCTTCGTAAT CAATCTTCTC ATTTACGATT AACCTCTTCT
                   730        740        750        760        770        780
matK ABCD   GGTATCTTTT TTGAGCGAAT ACATTTCTAT GAAAAAAAAA AAGATCCTGT AGAAGAAGTC
                   790        800        810        820        830        840
matK ABCD   TTCGTTAATG ATTTTCCGGC CGCCATCTTA TGGTTCTTCA AGGATCCTTT TATGCATTAT
                   850        860        870        880        890        900
matK ABCD   GTTAGATATC AAGGAAAATC TATTCTGTCT TCGAAGGATA CCCCTCTTCT GATGAATAAG
                   910        920        930        940        950        960
matK ABCD   TGGAAATATT ATCTTGTCAA TTTATGGCAG TGTCATTCTT ATGTGTGGTC TCAACCAGGA
                   970        980        990       1000       1010       1020
matK ABCD   AGGATTTATA TAAACCAATT ATCCAAGCAT TCCCTTGATT TTTTGGGTTA TTTTTCAAGT
                  1030       1040       1050       1060       1070       1080
matK BCD    ATGCGACCAA ACCTTTCGGT GGTACGGGGT CAAATGCTAG AAAATTCATT TATAATGGAT
matK A      ATGCGATCAA ACCTTTCGGT GGTACGGGGT CAAATGCTAG AAAATTCATT TATAATGGAT
                  1090       1100       1110       1120       1130       1140
matK ABCD   AATGCTATGA AGAAGCTTGA TACATTAGTT CCAATTATTC CTTTGATTGG ATCATTGGCT
                  1150       1160       1170       1180       1190       1200
matK ABC    AAAGTGAAAT TTTGTAACGC ATTAGGGCAT CCTATTAGTA AGTCCACCTG GGCAGATTCG
matK D      AAAGTGAAAT TTTGTAACGC ATTAGGGCAG CCTATTAGTA AGTCCACCTG GGCAGATTCG
                  1210       1220       1230       1240       1250       1260
matK ABCD   TCGGATTTTG ATATTATCGA CCGATTTCTG CATATATGCA GAAATCTTTC TCATTATTAC
                  1270       1280       1290       1300       1310       1320
matK ABCD   AGTGGATCCT CAAGAAAAAA GAGTTTGTAT CGAATAAAAT ATATACTTCG ACTTTCTTGT
                  1330       1340       1350       1360       1370       1380
matK ABCD   GTTAAAACTT TGGCTCGTAA ACACAAAAGT ACTGTACGAA CTTTTTTGAA AAGATTAGGT
                  1390       1400       1410       1420       1430       1440
matK ABCD   TATAAAATTA TTGGACGAAT TCTTTACGGA AGAAGAACAG ATTCTTTCTT TAATCTTCCC
                  1450       1460       1470       1480       1490       1500
matK ABCD   AAGAGCTTCT TATACTTTGA AGAAGTTTTA TAGAGGTCGA ATTTGGTATT TGGATATTTT
                  1510       1520       1530       1540       1550       1560
matK ABC    TGCATCAATG ATCTAGTCAA TCATGAATAA TTGGTTATGC GATCGTAGAA ATGGAAATTC
matK D      TGCATCAATG ATCTAGTCAA TCATGAATAA TTGGTTATGC GATCATAGAA ATGGAAATTC
                  1570       1580       1590       1600       1610       1620
matK BC     TATTTAAATA TTAAATAATT AAGAGATAAC AAAA------ ---------- --AATTAATT
matK A      TATTTAAATA TTAAATAATT AAGAGATAAC AAAAAATTAA GAGATAACAA AAAATTAATT
matK D      TATTTAAAT- ------AATT AAGAGATAAC AAAA------ ---------- --AATTAATT
                  1630       1640       1650       1660       1670       1680
matK ABCD   TATTTCTATT ATGAAATGTT CATCCAGTAA GATTAAGGGT TGATCAACTG AGTATTCAAC
                  1690       1700       1710       1720       1730
matK ABCD   TTTCTTAGAG TCGTGTATAG GGAAGGAACT TAATTTTAGA TGTATACATA
```

Figure 3.4: Sequence of the *mat*K region of apples. Positions where one of the defined phylogenetic groups (A, B, C or D, indicated in front of each row and explained p52) has a deviant base composition are marked in grey. ATTT…: Duplication I; AATT…: Duplication II

Table 3.8: Typing of the *mat*KdupII locus for wild apple samples and edible cultivars.

Origin	N	Dup II present	Dup II absent
Wild Belgian samples			
WBM	29	11	18
WBVo	9	6	3
WBVi	7	0	7
Belgian hybrids			
WBL	6	5	1
Wild German samples			
WGL	22	1	21
WGF	6	0	6
Edible cultivars			
ECV	11	4	7
Cultivar 'Elstar'	1	1	0

3.4.6 Assignment of genotypes to gene pools combining morphological, nuclear and cp-information

The purpose of this section is to summarise the results obtained using the information given in the previous sections. The gene pools described in 3.4.3.2 are taken as start point. This information is complemented with information on leaf hairiness (Table 3.3) and presence of duplication II in the *mat*K gene (Table 3.8). The summarised results are given in Table 3.9 and described below. Figure 3.5 gives an overview of characteristics per individual for all wild apples and edible cultivars.

1. **Gene pool 1** comprises the majority of wild individuals, both from Belgian and German origin. It is worth noting that it includes not only individuals without hairy leaves (36 individuals with score 0), but also twelve wild individuals that have hairy inferior leaf surfaces (eleven individuals with score 1, one individual with score 2).
   Following the results of Robinson *et al*. (2001), it was expected that in a group composed mostly by wild genotypes, the majority of the trees (54 individuals) did not have duplication II. However, a significant proportion (14 individuals) was shown to contain duplication II, indicating possibly some hybridisation with cultivated genotypes.
   Of the three individuals displaying admixture between this gene pool and gene pool 2, two (from WBVo and WBM) were scored for leaf hairiness and both had hairless leaves. Two hybrids were scored as duplication II absent, one as duplication II present. The absence of duplication II in admixed

individuals is expected if the hybridisation took place with the cultivar as pollen donor or if the cultivar did not have duplication II. The presence of duplication II in admixed individuals is expected if the hybridisation took place with the cultivar (that has duplication II) as the seed plant.

2. **Gene pool 2** comprises all edible cultivars (ECV), all presumed Belgian hybrids (WBL), three ornamental cultivars (OCV), one Belgian individual sampled in the wild (from sample WBM), and one German individual (from sample WGL, a *M. sylvestris* gene bank). The latter two individuals were those identified as being more related to edible cultivars in the PCO plots (Figure 3.2) and are characterised by moderate hairiness scores (score 2). All the trees with score 3 were assigned to this gene pool, confirming their cultivated origin.
   Duplication II was shown to be present in four edible cultivars (ECV), five presumed Belgian hybrids (WBL) and in the genotype from WGL; duplication II was absent in seven edible cultivars, one presumed Belgian hybrid and in the genotype from WBM.
3. **Gene pool 3** comprises most of the ornamental cultivars (OCV), with exception of the three individuals assigned to gene pool 2. The four genotypes for which indications of admixture with the edible cultivars (gene pool 2) were found were not scored for leaf hairiness or presence of duplication II.
4. **Gene pool 4** consists of only three closely related wild Belgian trees (from sample WBVo), the three genotypes are assigned completely to this cluster. One individual has sparsely hairy leaves (score 1), two trees have hairless leaves (score 0). Duplication II was shown to be present in all three genotypes.
5. **Gene pool 5** consists of only one ornamental cultivar, *M. tschonoskii*, assigned completely to this cluster. This Japanese apple is the only individual of the studied trees that does not belong to the section *Malus* (but to the section *Docyniopsis,* classification according to Phipps *et al*. 1990). This tree was not scored for leaf hairiness. Duplication II was absent.

Table 3.9: Summary of the results obtained using the assignment procedure based on SSR data and relationship with hairiness of leaves and presence of the chloroplast marker.

	Inferred gene pools				
(a) Category of origin	**1**	**2**	**3**	**4**	**5**
Belgian wild apples	**0.863**	0.048	0.006	**0.077**	0.005
Belgian hybrids (WBL)	0.009	**0.967**	0.005	0.013	0.006
German wild apples	**0.942**	0.041	0.004	0.008	0.005
Edible cultivars (ECV)	0.013	**0.974**	0.006	0.004	0.004
Ornamental cultivars (OCV)	0.007	0.168	**0.748**	0.011	**0.067**
(b) Hairiness					
Score 0	36[1]	0	0	2	0
Score 1	11	0	0	1	0
Score 2	1	3	0	0	0
Score 3	0	16	0	0	0
(c) *mat*KDupII locus					
Duplication II present	14[2]	10	0	3	0
Duplication II absent	54[3]	9	0	0	0

(a) Figures are the proportion of estimated membership to each of five inferred gene pools for genotypes of a given category of origin. Category Wild German: samples WGL and WGF, Wild Belgian: samples WBM, WBVo and WBVi;

(b) Figures are the number of trees with a given score for hairiness that were assigned to each inferred gene pool. Only 70 trees (see Table 3.1 for details on their category of origin) could be scored for this trait. [1]: two trees show admixture and are partially assigned to gene pool 1 (74 and 71 %) and partially to gene pool 2 (24 and 23%);

(c) Figures are number of trees that were typed for locus *mat*KdupII (see section 3.4.5). [2]: one tree shows admixture and is partially assigned to gene pool 1 (71 %) and partially to gene pool 2 (23%); [3]: two trees show admixture and are partially assigned to gene pool 1 (50 and 74 %) and partially to gene pool 2 (38 and 24 %).

	Inferred gene pool					Hairiness				Dupl II	
Pop	1	2	3	4	5	0	1	2	3	yes	no
ECV		X							X		X
ECV		X							X		X
ECV		X							X		X
ECV		X							X		X
ECV		X							X		X
ECV		X							X	X	
ECV		X							X		X
ECV		X							X	X	
ECV		X							X	X	
ECV		X							X		X
ECV		X							X	X	
WBL		X							X		X
WBL		X							X	X	
WBL		X							X	X	
WBL		X							X	X	
WBL		X							X	X	
WBL		X						X		X	
WBM	X					X					X
WBM	X					X					X
WBM	X					X					X
WBM	X					X					X
WBM	X					X					X
WBM	X						X			X	
WBM	X					X					X
WBM	X					X				X	
WBM	X					X				X	
WBM	X						X			X	
WBM	X					X					X
WBM	X					X					X
WBM	X					X				X	
WBM	X					X					X
WBM	x	x				X				X	
WBM	X					X					X
WBM	X					X					X
WBM	X					X				X	
WBM	X					X					X
WBM	X					X				X	
WBM	X						X				X
WBM	X						X				X
WBM	X						X				X
WBM	X						X				X
WBM	X					X					X
WBM	X						X			X	
WBM	X						X			X	
WBM		X						X			X
WBVi	x	x				?					X
WBVi	X					?					X
WBVi	X					?					X
WBVi	X					?					X
WBVi	X					?					X
WBVi	X					?					X
WBVi	X					?					X
WBVo				X		X				X	
WBVo				X		X				X	
WBVo				X			X			X	
WBVo	X					X				X	
WBVo	X					X					X
WBVo	X							X			X
WBVo	x	x				X					X
WBVo	X						X			X	
WBVo	X						X			X	

	Inferred gene pool					Hairiness				Dupl II	
Pop	1	2	3	4	5	0	1	2	3	yes	no
WGF	X						X				X
WGF	X					X					X
WGF	X					X					X
WGF	X					X					X
WGF	X					X					X
WGF	X					X					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					?					X
WGL	X					X					X
WGL	X					X					X
WGL	X					X					X
WGL	X					X					X
WGL	X					X					X
WGL	X					X					X
WGL		X						X		X	
WGL	X					X					X
WGL	X					X					X

Figure 3.5: Overview of characteristics for all wild apples and edible cultivars. See Table 3.2 for population codes. x: individual only partially assigned to this gene pool.

3.5 Discussion

3.5.1 Comparison of results from AFLP and microsatellite marker systems

A strong congruence between the results from both marker systems was revealed. Mantel tests on relatedness matrices between genotypes and samples showed highly significant positive correlations. This is in agreement with former observations where similar results were obtained with AFLP and SSR at the highest taxonomic levels (Powell *et al*. 1996, Maguire *et al*. 2002). As in the study of Maguire *et al*. (2002), the correlation between pair-wise genetic distance estimates from AFLP and SSR markers was lower at the intra-group level as compared to the inter-group level, but was still significant. Obviously, because of the nature of markers derived from both systems (binary markers for AFLP, multi-allelic markers for SSR), different values for genetic diversity statistics were obtained (higher for SSR than for AFLP). However, similar estimates of the relative differentiation among samples (F_{st}/G_{st}) were observed (Table 3.6).

Despite the drawbacks of the AFLP technique, data from AFLP markers generated with only three primer combinations led to the same conclusions as data from 12 SSR loci. However, if more apple samples were to be typed and results from different laboratories were to be compared, microsatellites might be the marker system of choice because of their easy exchangeability between different laboratories and amenability for creation and management of databases. Nevertheless, the availability and congruence of the two datasets gives us more confidence on the interpretation of results about differentiation between wild and cultivated gene pools and their level of hybridisation.

3.5.2 Delineation of the Malus sylvestris gene pool

Both nuclear marker systems resulted in concordant groupings of genotypes and a gene pool of *Malus sylvestris*, clearly divergent from cultivated material, could be delineated. The high degrees of genetic similarity between *M. sylvestris* individuals from Germany, Belgium and three European individuals from a private collection further support this observation. Furthermore, the model-based clustering methodology applied, as well as the AFLP- and SSR-PCO plots, allowed us to identify two putative wild apple trees (one tree collected in the wild from the WBM sample and one German tree from an established gene bank, WGL sample) that most likely represent ‘escaped’ edible cultivars. Moreover, the model-based clustering method identified three putative hybrids between the wild and cultivated gene pools, as well as three peculiar wild genotypes that make up a different gene pool and which could not be detected on either of the PCO plots (these genotypes are marked on plots in Figure 3.2).

The results based on molecular markers are to some extent concordant with the morphological trait of hairiness of leaves. Trees with a felted hairy inferior leaf surface (score 3) were assigned to the same gene pool, and represented edible cultivars and the hybrids from sample WBL. Three trees displaying moderate hairiness (score 2) were identified as derived from edible cultivars and the majority of trees with hairless leaves (score 0) were identified as *M. sylvestris* genotypes. However, one tree with hairiness score 2 and eleven trees with sparsely hairy leaves (score 1) were assigned completely to the wild gene pool and two genotypes showing admixture between wild and edible cultivar gene pools displayed hairless leaves. These observations suggest that some genetic variation in the degree of hairiness exists within the wild populations. Thus the degree of hairiness cannot be interpreted as a clear-cut 'degree of wildness' as speculated by Remmy & Grubber (1993) and often used by botanists as a rule of thumb in the field. Our results indicate that the degree of hairiness can be used as a first indication of the origin of an apple tree but the resolution of assignment of individuals to the wild and/or cultivated gene pool reached by molecular markers is much higher. In practice, a tree with felted hairy leaves (score 3) can be regarded as cultivated but the origin of trees with hairless or intermediate hairy leaves (score 0, 1 or 2) cannot be derived from this character only. More *M. sylvestris* genotypes have to be studied in order to establish the genetic variation of this trait.

3.5.3 Chloroplast variation at the matK region: is duplication II diagnostic for M. x domestica?

Robinson *et al*. (2001) concluded from their data that duplication II had arisen only once during the evolution of *Malus*. Moreover, this duplication was shown to be present in eight *M. x domestica* cultivars in the former study (Robinson *et al*. 2001) and was present in the cultivar 'Elstar', four old Belgian cultivars, five of the six presumed hybrids and one German genotype that were all assigned to the edible cultivar group based on their SSR fingerprints. This information suggests that a chloroplast copy that contains duplication II is derived from the edible cultivar gene pool. Hence, the presence of duplication II in *M. sylvestris* trees can be regarded as the result of former hybridisation events with *M. x domestica*. This idea is further supported by the complete absence of duplication II in the German wild samples, only the German individual that was assigned to the cultivar gene pool according to nuclear markers was shown to contain the duplication.

On the other hand, it has been shown that duplication II is also absent in many cultivated varieties. One cultivar in the study of Robinson *et al*. (2001) and seven (of eleven studied) old Belgian varieties lacked the duplication. Also a study of old Dutch varieties revealed that only half of the genotypes had this duplication (29 out

of 60 Dutch genotypes studied, unpublished results E. Coart). Therefore, the absence of duplication II cannot be considered as prove to the complete absence of edible cultivars in the maternal origin of the genotype considered.

These results also point out that *M. x domestica* cultivars do not form one monophyletic group. This is in agreement with the hypothesis that rather than the spread of cultivars, the technique of grafting was distributed during the domestication process of apples and local 'instant' domestication of desirable 'cultivars' took place (Harris *et al*. 2002). The discovery of grafting was undeniably the most important factor in the domestication of this highly heterozygous species with a strong self-incompatible system (Watkins 1995). These results also highlight the importance of the inclusion of local cultivated varieties in studies on the genetic diversity and identity of *M. sylvestris*. Only a thorough characterisation of the cultivated reference group will allow an accurate establishment of the extent of past and present hybridisation processes.

3.5.4 Occurrence of hybrids and cultivated genotypes in the wild

All presumed hybrids sampled in the wild (sample WBL) and the two aberrant genotypes from the putatively wild samples (classified in the edible cultivar gene pool) were not found to form an intermediate (hybrid) group, but instead completely merged into the edible cultivar gene pool. The complete nuclear marker data set therefore detected only three genotypes displaying admixture between wild and edible cultivar gene pools. No genotypes related to ornamental cultivars or hybrids between wild and ornamental gene pools have been detected in the wild. Three putative wild individuals from the same locality (WBVo) show very different fingerprints and make up a distinct gene pool in the clustering method. Based on these data it could be concluded that gene flow between wild and cultivated gene pools is very rare, and that wild genotypes have not lost their genetic identity through hybridisation with cultivated genotypes, despite the presence of cultivated varieties in the landscape (in plantations, in gardens...) and the occurrence of escaped cultivars in the wild. The conclusion that *M. sylvestris* and *M. x domestica* represent different gene pools is further supported by isozyme studies (Wagner & Weeden, 2000).

The information derived from the chloroplast *mat*KdupII marker shows that hybridisation between wild and cultivated gene pools might be present in Belgian forests. More than one third (39%) of the Belgian genotypes that were assigned to the wild gene pool possess the duplication II that has its putative origin in the *M. x domestica* gene pool. The three peculiar genotypes (from the wild Belgian sample WBVo) that made up a separate gene pool were all shown to contain duplication II,

what suggests their hybrid ancestry. The apparently contradicting information derived from nuclear and chloroplast markers suggests the occurrence of historic hybridisation between *M. sylvestris* and *M. x domestica* (at least with *M. x domestica* as female parent) and subsequent backcrosses of the hybrid with wild genotypes until the descendants of the hybrid are indistinguishable from 'genuine' wild genotypes when studied with nuclear markers. Moreover, the introgressive hybridisation as detected at the *mat*KdupII locus still represents an underestimation of the extent of hybridisation because this chloroplast marker is not present in all edible cultivars and obviously the pollen flow from cultivated genotypes to wild genotypes cannot be traced by studying the maternally inherited chloroplast DNA.

Seeds from cultivated trees, growing in nearby gardens or orchards, might have been brought into the forest by mammals that feed on apples. Also human presence (labourers, recreants, …) in the forest may account for the import of cultivated seeds by discarding of apple cores. The imported seeds may germinate to become cultivated trees in nature and the closest compatible trees will most likely be wild genotypes since all taxa in series *Malus* appear to be interfertile (Korban 1986). If the import of cultivated genes into the wild is limited, the wild populations will not lose their genetic identity and will remain distinct from the cultivated gene pool rather than evolve towards a hybrid swarm. The chloroplast DNA, however, maintains the imprint of the cultivated origin.This hypothesis is further supported by the fact that many wild Belgian trees show evidence of past introgression events and no introgression has yet been detected in German apple trees. The small forest sizes in Belgium, especially in the Flemish region, might promote the contact between wild and cultivated genotypes. No introgressed individuals were present in the only Walloon sample included (WBVi, represented by 7 trees) and it could be expected that the introgression in wild apple trees growing in other Walloon forests (particularly the larger forests in the Ardennes) is less important, as was revealed for the German *M. sylvestris* gene pool. However, it is clear that more research is necessary in order to unravel the hybridisation between wild and cultivated gene pools.

Furthermore, in this study we only have investigated the genetic composition of apple trees in the forest and no results are available on the genetic make-up of the fruits originating from these wild trees. But most of the wild apple trees in Belgian forests occur as solitary trees or small groups of old trees and more fragmentation and decline of populations might have occurred since they germinated. This means that the nearest sexually compatible tree may currently be a cultivated tree and hence, the next generation may show a higher level of hybridisation with edible cultivars than the level reported here. However, no spontaneous rejuvenation of wild apple trees was detected during collection of the material in Belgian forests and

fruits were only found on a few trees. Before applying forestry measures to increase fruit set and natural rejuvenation (e.g. opening the dense canopy cover surrounding the apple trees), restocking of the relict locations with wild genotypes from the region would reduce the risk that seeds from hybrid origin are formed. However, a better understanding of the possibilities and barriers for spontaneous hybridisation between wild apples and their cultivated relatives is necessary in order to develop an effective conservation program.

3.5.5 Population genetic structure in Malus sylvestris

As expected for an outcrossing tree species (Hamrick & Godt 1989), a low overall differentiation among wild samples was observed (F_{ST} = 0.046 (AFLP) and 0.06 (SSR)). High diversity values were obtained with both marker systems for all samples. The values of overall genetic diversity obtained agreed with diversity values in other studies of outcrossing tree species from similar geographic regions (H_e = 0.72 in this study, H_e = 0.73 for *Fraxinus excelsior*, Heuertz *et al*. 2001 using SSR data; H_j = 0.225 in this study, H_j = 0.29 for *Quercus petraea* and *Quercus robur* in this thesis, using AFLP data). For all wild samples a positive inbreeding coefficient was observed (mean F_{IS} = 0.11). An often reported cause of positive values of the inbreeding coefficient is the presence of null alleles (Bruford *et al*. 1998). Null alleles can be detected by studying the progeny of one parent or from a controlled cross, or by studying pedigrees. The microsatellite markers used in this study have been applied in only a few controlled crosses as far as they were polymorphic in the parents and could be mapped. CH01e12 had a null allele in cultivar Fiesta (E. van de Weg, Plant Research International, personal communication). In a study of family trees of several apple cultivars, three out of six microsatellite markers tested (all of them different from the ones used here) had a null allele (E. van de Weg *et al*., in preparation). So null alleles do occur and this may partly explain the shortage of heterozygotes. However, in the (multi-population) exact test for departure of Hardy-Weinberg expectations, nine out of twelve microsatellite loci showed a significant deficiency of heterozygotes. It is not so likely that each locus has null alleles at high frequencies. Therefore, a more plausible cause for the shortage of heterozygotes detected is that the remaining individuals of wild apples in the Belgian forests studied here were collected from scattered locations that never formed one random mating population, and that this led to a Wahlund effect, that is a heterozygote defiency due to population subdivision (Wahlund 1928 in Hartl & Clark 1997).

The clustering method grouped Belgian and German wild apples in one gene pool and the majority of differentiation between wild samples (57%) was attributed to divergence among samples from the same country. Nonetheless, some level of

neutral genetic differentiation was observed within and between Belgian and German origins and thus potentially differentiation for adaptive traits may have occurred. Also the discovery of different chloroplast types within Belgian wild genotypes implies the different history of Belgian and German wild apple trees.

3.6 Conclusions

Based on both AFLP and microsatellite markers it was possible to classify the apples analysed into three major gene pools: wild *Malus sylvestris* genotypes, edible cultivars and ornamental cultivars. All individuals sampled as presumed hybrids and two individuals (one Belgian, one German) sampled as *M. sylvestris* were assigned completely to the edible cultivar gene pool, revealing that cultivated genotypes are present in the wild. Furthermore, based on nuclear markers only three genotypes showed evidence of admixture between the wild and edible cultivar gene pools. Assignment of individuals based on the phenotypic trait of hairiness of leaves proved to be a valuable approach especially for trees with felted hairy leaves but did not reach the fine resolution provided by the molecular markers for trees with hairless or intermediate hairy leaves. Wild apples sampled in Belgium and Germany constitute gene pools clearly differentiated from cultivars and although some geographical pattern of genetic differentiation among wild apple populations exists, most variation is concentrated within the locations sampled. Results obtained from the chloroplast *mat*KdupII marker that revealed possible introgressive hybridisation or the involvement of cultivated genotypes in the history of the Belgian wild apple populations. No introgression was established in the gene pool of German wild apples. As the duplication scored is not present in all edible cultivated varieties and its phylogenetic origin is not completely clear, the extent to which hybridisation took place cannot be derived from the present data and might have been underestimated in this study.

It can be concluded that, although historic hybridisation might be present in the Belgian apple samples, the wild apple trees in the Belgian forests do not constitute a hybrid swarm but form a gene pool clearly distinct from both ornamental and edible cultivars. Our results clearly demonstrate that the molecular methods applied provide a valuable approach to ascertain the genetic identity of putative wild apple trees and to discriminate ‘genuine’ wild genotypes from cultivated and hybrid genotypes occurring in the wild. However, more research has to be conducted in order to establish the extent of past hybridisation events. This research ought to include more *M. sylvestris* genotypes and local cultivars and should focus on the phylogenetic origin of both gene pools.

4 Genetic diversity of *Quercus robur* and *Q. petraea*

4.1 Introduction

4.1.1 *Q. petraea* and *Q. robur*

Linnaeus proposed 12 species within the genus *Quercus* but since then the number of oak taxa has increased to several hundred, the exact number depending on the author. Schwarz (1964) considered 320 taxa to be separate species, plus many forms, varieties or subspecies. The three indigenous species in Belgium belong to the white oak section (subgenus *Lepidobalanus*), a species rich section spread over Europe, North America and Asia. *Quercus robur* L. (pedunculate oak) and *Quercus petraea* (Matt.) Liebl. (sessile oak) are indigenous oak species in the Flemish region. *Quercus pubescens* WILLD. (pubescent oak), the third oak indigenous to Belgium is a rather rare species and occurs mainly in the Walloon region. *Q. pubescens* was not included in this study.

Sessile and pedunculate oaks are widespread throughout Europe, the natural occurrence of *Q. petraea* is more limited to the north than *Q. robur*. Both are largely sympatric and generally occupy different but proximal ecological niches (Rushton 1979). *Q. robur* occurs on a wide range of soils, is better adapted to permanent high water levels as alluvial soils and prefers open conditions. *Q. petraea* grows well in more acidic soils, tolerates drier and poorer conditions, is more shade tolerant and prevails on elevated habitats. Both species can be recognised based on their leaf and fruit morphology but morphological intermediates have been reported which started a controversy on the extent of hybridisation between both species that is still ongoing (see further).

Both oak species are monoecious, carrying separate male and female flowers on the same branches. They are highly outcrossing, with selfing rates estimated in mature oak stands varying between 0 and 1% for *Q. petraea* and between 3 and 5% for *Q. robur* (Kremer & Menozzi 2000). Natural vegetative propagation is limited to stump sprouting; recurrent stump sprouting can result in stems separated by several meters belonging to the same genotype (coppice stools). Propagation by root suckers or by natural rooting of cuttings is rare in these species. Both oak species produce regularly seed, on average every seven to eight years in Belgium (called ‘mast’ years). The flowering period depends on the geographic location. In Belgium, oaks flower generally from mid April to the end of May. An important variation of up to thirty days between flowering periods can be observed between the earliest and latest flowering trees of the same species in a given stand. In most years, *Q. robur* and *Q. petraea* flower at the same period, but it has been observed that Q. *petraea* sometimes flushes earlier (Bacilierri *et al.* 1995). Fertilisation takes place eight to ten weeks after pollination, followed by a rapid growth of the acorns, which reach their full size at mid September. *Q. robur* may flower as early as ten years in open

conditions, but as late as thirty years in dense stands. In general, first fruiting of *Q. petraea* occurs later than *Q. robur*.

Oak pollen is of small size (diameter 26 to 29 µm, Rushton 1976) and is dispersed by wind. There exist no size differences between both indigenous species. Observations of pollen dispersion made with pollen traps indicate that oak pollen can be dispersed as far as 7 km from the source (Lahtinen *et al.* 1996). However, morphometric data suggest that pollen grains could be transported by wind at distances exceeding hundreds of kilometres (Stanley & Linskens 1974). Experiments of effective pollen dispersal based on parentage analysis conducted both in a natural stand (Streiff *et al.*1999) and in a seed orchard (Buitevelde *et al.* 2001) do not refute these hypotheses. In these experiments, covering a total area of 5.76 and 4.5 ha repectively, a high percentage (65% for *Q. robur* and 69% *for Q. petraea* in the natural stand; 70% for *Q. robur* in the seed orchard) of the offspring was the result of pollination by pollen from outside the study site.

Acorns can be dispersed by several rodents and by birds. Rodents transport the acorns over small distances and most often eat the fruits. Birds, especially Jays transport the acorns in their oesophagus and hide them to feed later on. Jays can disperse a few thousands of acorns over several kilometres in a given season. Jays prefer acorns of *Q. robur* to acorns of *Q. petraea* because of their shape (Bossema 1979).

4.1.2 Genetic diversity and evolutionary history at the European level

Oaks are diploid organisms comprising 12 pairs of chromosomes (2n=2x=24). In both oak species, the DNA content is 0.9 pg/C (Zoldos *et al.* 1998). The average base composition of the genome, estimated by flow cytometry is about 40% GC, which is typical for higher plants. Estimated genome length in *Q. robur* is 1200 cM (Barrenche *et al.* 1998). Oaks exhibit a level of diversity that is amongst the highest of all woody species (Kremer & Petit 1993), despite their relatively small genome. Biological attributes as highly outcrossing reproduction, large population sizes, large distribution areas, important gene flow, and possibly also hybridisation with related species are likely to be responsible for these levels of diversity.

Studies using molecular markers to describe population genetic structure within both oak species have been conducted in many European countries, on different geographical scales, including studies based on isozymes (Kremer 1991; Zanetto *et al.* 1994; Samuel *et al.* 1995; Zanetto & Kremer 1995; Finkeldey 2001a,b; Gömöry *et al.* 2001; Siegismund & Jensen 2001), RAPD (Moreau *et al.* 1994) or a combination of different marker systems (Bodénès *et al.* 1997a,b using RAPD and

SCAR (a) and SCAR and SSCP (b); Streiff *et al.* 1998 and Degen *et al.* 1999 using isozymes and microsatellites). All studies demonstrated that the vast majority of the genetic variation could be attributed to the within population level (all values higher than 90%), even when populations from distant locations were compared.

The evolutionary history of European oaks since the last ice age is now well understood. Post-glacial migration routes from glacial refuges have recently been reconstructed using cpDNA information and palynological information (see Figure 4.1; Brewer *et al.* 2002; Dumolin-Lapègue *et al.* 1997; Petit *et al*. 1993, 2002a and 2002b). Three areas of southern Europe have been identified as refuges for deciduous oaks: southern Iberian Peninsula, southern Italian Peninsula and southern Balkan Peninsula. Based on the detection of 32 chloroplast haplotypes, spread over Europe, a strong phylogeographic structure has been observed, where related haplotypes have broadly similar geographic distributions. In total, six cpDNA lineages have been identified, which have distinct geographical distributions, mainly along a longitudinal gradient. Most haplotypes found in northern Europe are also present in the south, whereas the converse is not true, suggesting that the majority of mutations observed nowadays in the chloroplast regions investigated were generated prior to post-glacial recolonisation (Dumolin-Lapègue *et al.* 1997; Petit *et al*. 2002a). The spread of the oaks took place in two steps. First, in the late-glacial interstadial *Quercus* spread to the central European mountains from these refuges. Second, with the stabilisation of a climate favourable to deciduous tree species, oak spread into northern Europe, rapidly into the northwest and more slowly into the centre and east, due to physical barriers. Approximately 6,000 years before present oaks reached their maximum extension in Europe. The average colonization speed reached up to 380 metres/year. Rare long distance dispersion events are likely to explain this rapid migration rate (Brewer *et al.* 2002).

Strong funding events during recolonisation and the occurrence of rare long distance dispersion events are probably the cause of the observed phylogeographic structure across Europe. These two factors, together with limited gene flow through seed because of large population sizes, have led to the formation of monotypic stands where all the trees share the same cp-haplotype. Exceptions are expected at the junction between patches characterised by different haplotypes. However, even in regions where different recolonisation routes met, clusters made up of a single haplotype can often be observed. In fact, most of the European oak forests studied to date seem to be completely fixed for a given chloroplast variant. For maternally inherited neutral markers as the described chloroplast haplotypes, the distribution of genetic diversity is therefore the opposite to that of neutral nuclear markers. Inter-population differentiation represents more than 75% of the variation.

These population differentiation studies provide indirect measures of gene dispersal and they support two conclusions of the studies based on direct observations made by parentage analysis: the existence of important pollen flow and the asymmetry between gene flow through pollen and seed dispersal.

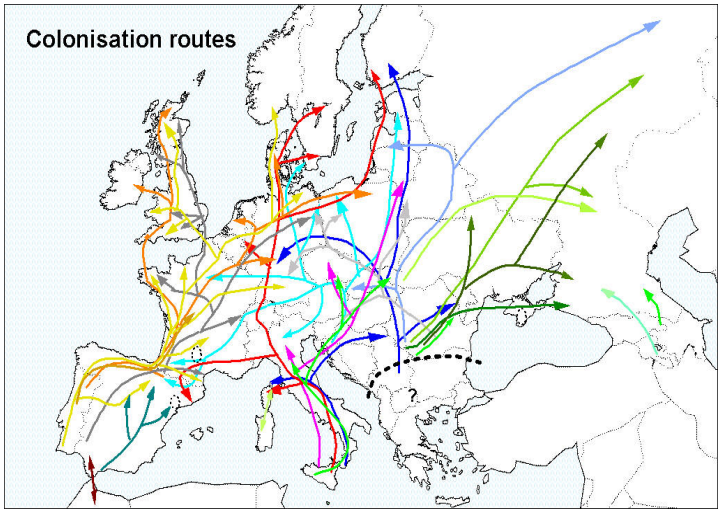


Figure 4.1: Post-glacial migration routes of White oaks, inferred from cpDNA and palynological information (Petit *et al*. 2002b). Each arrow represents the inferred colonisation route of a chloroplast haplotype or group of closely related haplotypes.

4.1.3 Hybridisation between *Q. petraea* and *Q. robur*

The presence of morphological intermediates between both species at first led taxonomists and foresters to believe they were freely interfertile (Wigston 1974). The authors that support the widespread occurrence of hybridisation suggested a revision of their taxonomic status to the status of subspecies (e.g. Olsson 1975; Van Valen 1976). However, crossing experiments (Rushton 1977; Aas 1990; Steinhoff 1993; Kleinschmit & Kleinschmit 1996) demonstrated that the number of successful interspecific crosses was low and that the fertilisation of *Q. robur* with *Q. petraea* pollen was much more successful than the reciprocal crosses. This evidence of fertility barriers led to much speculation about the origin and taxonomic status of the morphological intermediate forms. Several authors postulated that the variation within the two species had been underestimated (e.g. Jones 1959; Gathy 1969; Dupouey 1983; Dupouey & Badeau 1993; Aas 1993) and stated that if hybridisation had been at all common, the close relationship between the two species and their

sympatric distribution, would by now have led to the complete loss of their separate identities.

In surveys based on multivariate approaches for data analysis (e.g. Dupouey & Badeau 1993; Aas 1993; Bacilieri *et al.* 1995), the number of morphological intermediates was low (less than 5%). The phenotypic studies published to date were carried out in various regions of the geographical range of the oaks and different statistical approaches were applied. It is therefore very difficult to attribute the differences found to methodological and/or geographical effects or to state whether the morphological character's variation detected is either the result of modification or of genetically fixed variability (Kleinschmit *et al.* 1995). A recent study however, tried to overcome these difficulties by assessing leaf morphology in nine mixed oak stands located in eight European countries and comparing three multivariate statistical techniques (Kremer *et al.* submitted). The first synthetic variable derived by each multivariate analysis exhibited a clear and sharp bimodal distribution, with overlapping in the central part. The two modes were interpreted as the two species, and the overlapping region was interpreted as an area where the within-species variations were superimposed. No discontinuity was observed in the distribution nor evidence was found of a third mode that would have indicated the existence of a third population composed of trees with intermediate morphologies. These observations were made both over the nine European stands and separately within each stand. As possible explanation for the maintenance of these two modes, maternal effects on morphological characters and fitness of hybrids in comparison with parent species were proposed and the question whether the two modes are composed of either pure species or pure species and introgressed forms remains open.

Although the majority of studies published to date concerning hybridisation are based on morphological evidence, (introgressive) hybridisation is not necessarily indicated by the phenotypic occurrence of the characters of one taxon in another. In their review of plant hybridisation, Rieseberg and Ellstrand (1993) noticed that hybrids were a mosaic of phenotypes with parental and intermediate characters rather than just intermediate ones. Observations of F1 families of controlled crosses between *Q. petraea* and *Q. robur* showed that juvenile F1 hybrids (up to 5 years old) exhibited leaf morphologies that were similar to the female parent rather than intermediate, regardless which species was used as female parent (Kleinschmitt *et al.* 1995). No observations are available on mature F1 families that could have sustained the same conclusions.

With the advent of molecular marker technologies, new tools became available for further investigation of the genetic relationship and interspecific variability of the sessile and pedunculate oaks. However, the first molecular studies were discouraging because markers appeared to be less discriminating than observations of the phenotype. Isozyme-analyses could not reveal any species-specific marker (Zanetto & Kremer 1995; Bacilieri *et al.* 1995), but confirmed directional hybridisation under natural conditions (Bacilieri *et al.* 1996) as previously demonstrated in artificial crosses. RAPD analyses supported the hypothesis that differentiation between both white oak species is likely to be controlled by only a few loci (Bodénès *et al.* 1997a). The overall nucleotide divergence detected between the species was very low (0.5%), however a few 'hot spots' for differentiation with an increased divergence (3%) were observed. Furthermore, these 'hot spots' are located in highly polymorphic genomic regions (Bodénès *et al.* 1997b).

Also chloroplast markers were found to be present with similar frequencies in both oak species over the European continent (e.g. Petit *et al.* 1993; Dumolin-Lapègue *et al.* 1999). According to Muir *et al.* (2000), microsatellites could be the markers of choice for discrimination between related oak species at population level. Microsatellite analysis enabled to group several European oak populations per species, with highly significant bootstrap values, suggesting that sessile and pedunculate oaks represent clearly separate taxonomic units. Furthermore, a recently analysed isozyme locus provided preliminary results on its possible use for differentiation between the two oak species (Gömöry 2000).

4.1.4 Occurrence and management of oak stands in Flanders

Indigenous oaks are major broadleaved species in Flanders, covering ca 8% of the total forest area of 150.000 ha (Jacques & De Cuyper 1998). Oaks are also important species in new forestations. For instance, oaks were planted on 44% of the area forested in the program of forestation of agricultural land (R. De Vreese in Coart *et al*. 2001). Management of public oak forests is 'close to nature'. This silviculture practice is ecologically based, with attention to all components of the forest ecosystem, and pursuing stable healthy forests with a durable economic and ecological value. Although natural rejuvenation is aimed at in public forests and stimulated in private owned forests with a subvention policy, this is frequently hampered by local environmental circumstances and subsequently artificial regeneration is applied.

In general, private owners preferred pedunculate oaks (90% of planted oaks) to sessile oaks, whereas the Forest and Green Areas Division has planted more sessile oaks during recent years (60% of planted oaks) (R. De Vreese in Coart *et al*.

2001). The acknowledged oak provenances in Flanders, selected on phenotypic superiority of their forest characteristics, cannot produce the required amount of oak reproductive material, as they consist of a total area of only 64 ha of forest for *Q. robur* and 19 ha for *Q. petraea*. Therefore, forestations are frequently carried out with foreign provenances whereof it is in most cases still unknown whether these are well adapted to local conditions (Jacques & De Cuyper 1998). As stated in the introduction of this thesis, until present no autochthonous provenances are acknowledged and thus no autochthonous material can be put at the disposal of forest managers.

Although the inventory of autochthonous trees and shrubs is not yet completed for the Flemish region, it has already revealed that the number of oak locations that can be regarded as autochthonous (according to the criteria explained in the general introduction) is very limited and most are only small relics (Maes & Rövekamp 1998; Maes & Rövekamp 2000; Opstaele 2001; Rövekamp & Maes 1999; Rövekamp & Maes 2000; Rövekamp *et al*. 2000). For oak, only coppice wood and trees are assigned an autochthonous status. The historical coppice practice gave oaks a possibility to attain much older ages than non-coppiced trees. The age of oak coppice stools with circumferences of 30 m at soil level is roughly estimated at a few thousand years (pers. comm. B. Maes). Because of their old age, coppice stools may date from before the period of intensive transport of acorns and hence they can be considered autochthonous. These coppiced oak stands are relics within today's sylvicultural practice of high forest, which often survived demolition due to their location on infertile land dunes. The largest stools of both oak species are found in the Northern parts of Flanders, on sandy soils in the Campine region.

4.2 Present oak research and objectives

The general aim of this research was to study the genetic diversity and population genetic structure of a set of representative Flemish oak stands in order to develop conservation guidelines for oak genetic resources in Flanders. Emphasis was laid on (i) the identification and characterisation of putative autochthonous populations and (ii) the comparison of the genetic composition of autochthonous oak populations with selected provenances. In addition, as both historical and present inter-specific hybridisations are known to occur (Petit *et al*. 1997; Dumolin-Lapègue *et al.* 1999), the problem of species identity is also treated in this study on genetic diversity of Flemish oak populations.

The recently established post-glacial migration routes provided a valuable approach to evaluate the autochthonous character of oak populations. In view of the difficulties that often arise from field evaluations, this was highly desirable information. Therefore, the chloroplast haplotypes of Flemish oak populations were determined and the results were confronted with the European data to assess the phylogenetic origin and autochthonous character of the populations studied. Chloroplast haplotypes were identified using the PCR-RFLP protocol that had been developed within the framework of a European collaboration. The results are presented in section 4.4.

All populations where more than ten trees could be sampled (total of twenty-six Belgian and four foreign populations) were typed at six microsatellite loci. First, the diversity and population genetic parameters were estimated for autochthonous populations of both species. Secondly, autochthonous populations and selected provenances were compared for their genetic diversity and population structure present. The results obtained are presented in section 4.5.2 and the usefulness of this set of microsatellite loci for species differentiation is discussed.

Most of the studies on oak genetic diversity published recently were based on the use of microsatellite markers and at the start of this research no study on the genetic variability had been published in which AFLP was applied. AFLPs usually yield high data output per reaction and have good reproducibility, what makes this marker system suited to identify species and detect hybrids (Mueller & Wolfenbarger 1999; Jarvis & Hodgkin 1999). Therefore, the performance of microsatellite and AFLP markers was compared, both to differentiate the two species studied and to unravel levels of genetic diversity at the within species level. This study was based on a sub-set of eight of the populations sampled. The analysis was complemented with a small study on leaf morphology for two populations (one of each species). The results of this comparison are summarised in 4.5.3.

4.3 Sampled oak stands

The putative autochthonous oak populations and selected oak provenances used in this study are listed in Table 4.1. The geographical location of the populations located in Flanders is shown in Figure 4.2 (note that some of the selected provenances sampled are located in Wallony and are not represented in this Figure). Autochthonous populations were chosen from the available inventories of autochthonous trees and shrubs, conducted by Maes and Rövenkamp (Maes & Rövekamp 1998; Rövekamp & Maes 1999). In total 25 Q. *robur* and 11 *Q. petraea* locations were sampled, or 390 and 215 individuals respectively. If present, 30 trees were sampled ad random across the complete surface of the stand. If fewer trees were present, all were sampled. Minimum distances between sampled trees varied from 50 m up to circa 100 m. The populations were coded according to the dominant oak species (as observed by eye during collection; QR for *Q. robur* populations, QP for Q. *petraea* populations). Furthermore, seedlings of four foreign selected provenances that are used for forestations in Belgium were sampled at the tree nursery 'Sylva': two *Q. robur* provenances (originating from The Netherlands and Germany) and two *Q. petraea* provenances (originating from Germany and France). Individual trees (or small groups of trees) that are very old (e.g. the so-called 'thousand-year old oak' in Lummen) were included in the study of occurrence of chloroplast haplotypes in Flanders.

Five trees per stand were analysed for their chloroplast-haplotype (or all trees if less trees were present at the location). For microsatellite analysis, 20 individuals (or all individuals if less than 20 individuals could be sampled) were typed for each stand and locations where less than 10 samples could be collected were not included for SSR analysis. All trees sampled in the eight populations chosen for AFLP analysis were typed in order to compare results of SSR and AFLP marker systems. However, not all samples produced AFLP fingerprints of sufficient quality and sometimes SSR loci were not amplified. These problems are probably due to the insufficient quality of the DNA extracted. The number of individuals that could be typed for each population and for each molecular technique is given in the last columns of Table 4.1. Leaf morphology was described for all sampled trees of two populations (QPAUTKB and QPAUTW), all trees sampled in these populations were described for leaf morphology.

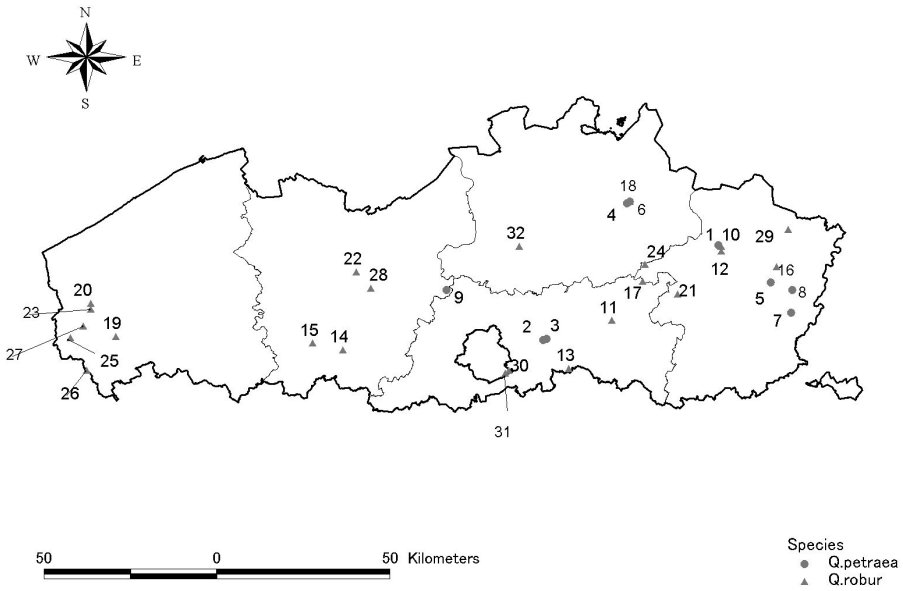


Figure 4.2: Geographical location of oak populations sampled in Flanders. Numbers refer to the first column of Table 4.1.

Table 4. 1: Description of oak populations analysed.

	Nb	Code	Inv.nr	Location	Name	# Sam	# Cp	# AFLP	# SSR	# Morph
	Q. petraea									
Autochthonous	1	QPAUT117	99-117	Hechtel-Eksel	Gemeentebos	30	5	24	20	0
	2	QPAUT153	98-153	Bertem	Bertembos	2	2	0	0	0
	3	QPAUT154	98-154	Bertem	Bertembos	2	2	0	0	0
	4	QPAUTHM	98-102	Kasterlee	Hoge Mouw	6	5	0	0	0
	5	QPAUTKB	99-86	As	Klaverberg	30	5	20	20	30
	6	QPAUTKABB	98-101	Kasterlee	Kabouterkensberg	5	5	0	0	0
	7	QPAUTOG	--	Opgrimbie	Opgrimbie	20	5	17	17	0
	8	QPAUTWS	--	Maasmechelen	Windelsteen	30	5	19	19	0
Selected	9	QPB	--	Buggenhout	Buggenhoutbos	30	5	0	15	0
		QPLP	--	Chimay	La Pointe	30	5	0	13	0
		QPQL	--	Chimay	Queu de L'herse	30	5	0	9	0
		QPD818*	--	Germany		20	5	0	18	0
		QPFR09*	--	France		20	5	0	17	0
					Total *Q. petraea*:	**255**	**59**	**80**	**148**	**30**
	Q. robur									
Autochthonous	10	QRAUT114	99-114	Hechtel-Eksel	Gemeentebos	30	5	27	20	0
	11	QRAUT123	98-123	Tielt-Winge		14	5	0	0	0
	12	QRAUT133	99-133	Hechtel-Eksel	Gemeentebos	30	5	26	20	0
	13	QRAUT166	98-166	St Joris	Meerdaalwoud	20	5	0	18	0
	14	QRAUT208	99-208	St. G. Oudenhove		2	2	0	0	0
	15	QRAUT337	99-337	Oudenaarde	Oosemolen	1	1	0	0	0
	16	QRAUTGR	99-79	Gruitrode	Robertii	30	5	19	19	0
	17	QRAUTH	99-179	Tessenderlo	Heuvelken	8	5	0	0	0
	18	QRAUTKABB	98-101	Kasterlee	Kabouterkensberg	4	4	0	0	0
	19	QRAUTKEM	99-437	Heuvelland	kemmelberg	11	5	0	10	0
	20	QRAUTLR	98-19	Heuvelland	Lo-Reninge	4	4	0	0	0
	21	QRAUTLUM	98-179	Lummen	1000j eik	2	2	0	0	0
	22	QRAUTMK	--	Laarne	Meerskant	30	5	0	19	0
	23	QRAUTOV	98-1	Heuvelland	Oost-Vleteren	4	4	0	0	0
	24	QRAUTPH	99-172	Tessenderlo	Paddenhoek	20	5	0	14	0
	25	QRAUTPR	99-403	Proven	Couthof	1	1	0	0	0
	26	QRAUTVB	99-477	Heuvelland	Vidaigneberg	7	5	0	0	0
	27	QRAUTVL	99-450	Heuvelland	Vletse	2	2	0	0	0
	28	QRAUTW	98-83	Wetteren	Speelbos	30	5	22	20	30
Selected		QRAIS	--	Virton	Aisances	30	5	0	14	0
	29	QRBB	--	Bree	Berkenbroek	15	5	0	10	0
		QRBO	--	Virton	Bochet	30	5	0	10	0
	30	QRDMW	--	Groenendaal	Dronkenmansweg	30	5	0	12	0
	31	QRKW	--	Groenendaal	Kwekerijweg	5	5	0	0	0
	32	QRLI	--	Lint	Kapellekensbos	30	5	0	12	0
		QRD817*	--	Germany		20	5	0	19	0
		QRNL01*	--	The Netherlands		20	5	0	15	0
					Total *Q. robur*:	**430**	**115**	**94**	**232**	**30**
					Total *Quercus*:	**685**	**174**	**174**	**380**	**60**

Nb: number on Figure 4.2. Inv.nr: inventory number for autochthonous populations (Maes & Rövekamp 1998; Rövekamp & Maes 1999); # Sam: number of trees sampled in the population; # Cp, #AFLP, #SSR, #Morph: number of samples used for chloroplast, AFLP, SSR and morphological analysis; *: samples collected from seedlings in a tree nursery. Other samples without Location Nb were collected in Wallony.

4.4 Chloroplast DNA diversity in Flanders and patterns of diversity

This work was carried out in the framework of the European project 'Synthetic maps of gene diversity and provenance performance for utilization and conservation of oak genetic resources in Europe (project FAIROAK, FAIR1 PL95-0297)'. At the Department Plantgenetics and Breeding, Flemish autochthonous populations and Belgian selected provenances were typed. These data were merged with information from other European laboratories and presented together in a special issue of Forest Ecology and Management (vol. 156, 2002). The Belgian data are covered by the regional paper on western-central Europe, together with results from Luxembourg, The Netherlands, Germany, Czech Republic and parts of Austria (König *et al.* 2002). In the following section the results obtained at DvP for the Belgian populations listed in Table 4.1 are presented and interpreted in the context of the findings at European scale.

4.4.1 Results: Occurrence of cpDNA haplotypes in Flanders

In total, five cpDNA-haplotypes were detected in Flanders. These haplotypes represented three different lineages (see Petit *et al.* 2002a for detailed information on lineages and haplotypes). Haplotypes coded as numbers 10, 11 and 12 belong to the lineage B of inferred Iberian origin, the haplotype coded as 1 belongs to lineage C from Italian origin and the haplotype coded as 7 to lineage A from putative Balkan origin. Populations containing only one haplotype are further referred to as monotypic, the remaining as polytypic. Except for haplotype 7 (that was only identified for one pedunculate oak tree originating from a selected provenance), all haplotypes were present in both oak species. Table 4.2 summarizes the frequencies of haplotypes for species, autochthonous populations and selected provenances and the associated diversity statistics. Different trends can be derived about the occurrence of haplotypes:

- Contrast between Iberian, Italian and Balkan haplotypes: Iberian haplotypes are more frequent in Flanders than Italian ones. The three related Iberian haplotypes dominate (66 and 80 % for autochthonous populations and selected provenances respectively), followed by the Italian haplotype 1 (34 and 18% respectively). Figure 4.3 illustrates the geographical distribution of cpDNA haplotypes in Flanders. The most striking characteristic is the occurrence of many monotypic populations of haplotype 1 (of presumed autochthonous origin) in the eastern region and the almost complete absence of type 1 in the rest of the Flemish region (occurs only once in a polytypic population).
- Contrast between *Q. robur* and *Q. petraea*: For haplotypes 1 and 12 big differences were found between species. For both origins, the frequency of the Italian haplotype 1 was higher in *Q. petraea* than in *Q. robur*, while the Iberian haplotype 12 was more frequently found in *Q. robur* trees.

- Contrast between autochthonous populations and selected provenances: Overall frequencies of haplotypes are similar for autochthonous populations and selected provenances. However, autochthonous and allochthonous stands differ significantly for their distribution of genetic diversity. Although the total diversity (h_t) in autochthonous populations and in selected provenances is similar (0.68 and 0.64 respectively), the distribution of this diversity among populations (h_s) is clearly different. This effect is also illustrated in Figure 4.4: for autochthonous populations, the frequency of monotypic populations is much higher than for selected stands, what results in a much higher genetic differentiation between populations (G_{st} = 0.70 and 0.09 for autochthonous and selected stands respectively).

Table 4.2: Frequencies of haplotypes and levels of diversity and differentiation by species and category of origin. N: number of populations analysed; Ntr: number of trees analysed; f10, f11, f12, f1 and f7: frequencies of the respective haplotypes; h_t: total diversity; h_s: within-population diversity; G_{st}: genetic differentiation.

Cat. of origin	N	Ntr	f10	f11	f12	f1	f7	h_t	S.D	h_s	S.D	G_{st}	S.D
Autochthonous populations													
Q. petr + Q. robur	17	85	0.47	0.04	0.15	0.34		0.68	0.05	0.20	0.08	0.70	0.11
Q. petraea	6	30	0.38	0.03	0.05	0.55		0.61	0.08	0.20	0.12	0.71	0.17
Q. robur	11	55	0.56	0.04	0.24	0.16		0.66	0.11	0.22	0.11	0.66	0.17
Selected provenances													
Q. petr + Q. robur	9	45	0.56	0.04	0.20	0.18	0.02	0.64	0.07	0.58	0.12	0.09	0.11
Q. petraea	3	15	0.60		0.07	0.33		0.60	0.13	0.47	0.24	0.22	0.43
Q. robur	6	30	0.53	0.07	0.27	0.10	0.03	0.65	0.08	0.63	0.14	0.03	0.15
Total													
Q. petr + Q. robur	26	130	0.50	0.04	0.17	0.28	0.01	0.65	0.04	0.33	0.07	0.49	0.11
Q. petraea	9	45	0.44	0.02	0.05	0.49		0.60	0.04	0.25	0.11	0.58	0.17
Q. robur	17	85	0.55	0.05	0.25	0.13	0.01	0.64	0.07	0.39	0.10	0.39	0.15

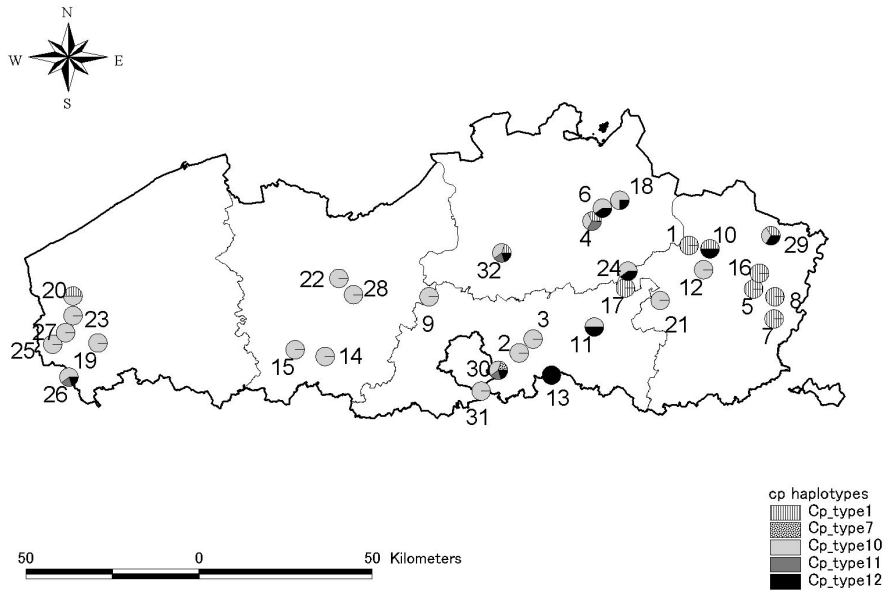


Figure 4.3: Chloroplast DNA haplotypes in Flanders. Numbers: see Table 4.1. Monotypic populations are presented as full circles, polytypic populations by circles with different patterns.

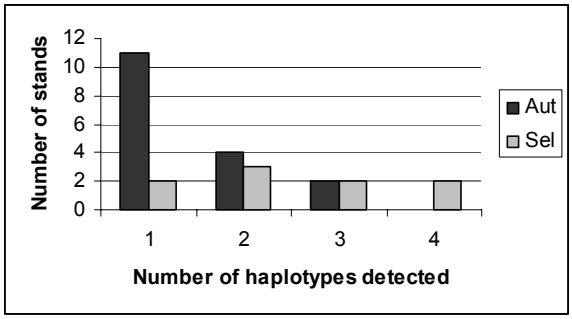


Figure 4.4: Number of haplotypes detected among the trees analysed in each stand for autochthonous populations (Aut) and selected provenances (Sel). Only stands where three or more individuals were typed for their cp-haplotype are included in these calculations.

4.4.2 Discussion

4.4.2.1 Geographical variation of cpDNA variation in relation to inferred post-glacial recolonisation routes

The haplotypes of lineage B migrated from a glacial refuge on the western coast of the Iberian peninsula through Europe (Olalde *et al.* 2002; Petit *et al.* 2002a), mostly

following the Atlantic and North sea coast of the continent (Cottrel *et al.* 2002; Petit *et al.* 2002b; Jensen *et al.* 2002). This early and rapid movement along the Atlantic border resulted in the spread along the Northern European coast into Belgium, The Netherlands, Germany and Denmark. The overall distribution of the three Iberian types occurring in Flanders (10, 11 and 12) is very similar. Haplotype 10 is the most common of the lineage. The autochthonous status of oak populations characterised by these haplotypes in Flanders is supported by the presence of numerous stands fixed for haplotypes of lineage B in Belgium, The Netherlands, northeast Germany and eastern France (König *et al.* 2002).

Haplotype 1 from the Italian lineage C, spread rapidly through the Apennines chain, but its further spread northwards was delayed by the difficult crossing of the Alps (Csaikl *et al.* 2002) that form a formidable barrier for species as oaks (that are currently growing rarely at altitudes higher than 1400 m at these latitudes). More northerly colonisation, into south-western Germany, appears to have proceeded along the lowlands situated between the Jura Mountains and the Swiss Alps. From there the colonisation continued all the way to southern Scandinavia, crossing the centre of Germany (Petit *et al.* 2002b). The presence of many monotypic populations of haplotype 1 in eastern Belgium and eastern parts of The Netherlands shows western movements from the main route (König *et al.* 2002) and thus this haplotype is considered autochthonous to eastern Flanders.

Lineage A, which likely originates from the Balkan, is only detected in Flanders in one individual of haplotype 7. The most likely migration route of this haplotype started from a Balkan refuge (see Petit *et al.* 2002b for a discussion) and spread westwards but also east- and northwards, resulting in a large distribution across Western Europe. The migration may have followed two paths from the Balkan. One path led north to a secondary refuge in the southeast Alps, a second westwards across the exposed Adriatic Sea basin and to the southeast of France (or even the Pyrenees). It migrated from the east of the Alps in a fanlike pattern, passing northwards through Germany and Poland to reach the west and east of the Baltic Sea, respectively. Haplotype 7 also spread west following the northern slopes of the Alpine chain from this secondary refuge in the eastern Alps. It moved from the other secondary refuge in the southeast of France (or the Pyrenees) both to the southwest and the north, thus entouring the Alps in both cases (Matyas & Sperisen 2001). Although the haplotype has a broad distribution in western central Europe, the almost complete absence of monotypic populations of this haplotype in Belgium, The Netherlands and the western most part of Germany, strongly suggests that this haplotype is not native to Flanders. Only one population monotypic for haplotype 7 was recently detected in The Netherlands (pers. comm. J. Buitevelde, Alterra).

4.4.2.2 Inter and intra-specific diversity

In western-central Europe, all major haplotypes are found in both species, with generally similar frequencies (König *et al.* 2002). Past and ongoing hybridisation and introgression can account for this observation (Petit *et al.* 1997). In autochthonous Flemish populations, haplotype 1 is more frequent in *Q. petraea* (55 %) than in *Q. robur* (16 %) populations, what is due to the correlation between the geographic distribution of haplotype 1 and the distribution of *Q. petraea* (both prevailing in the eastern region). Except for the introduced haplotype 7, all haplotypes occur in both species. A higher level of within-population diversity (h_s) and lower G_{ST} in *Q. robur* compared to *Q. petraea* can be seen, which is a general trend that has been found in Europe (Petit *et al.* 2002b; Bordacs *et al*. 2002). This supports the hypothesis that acorns of *Q. robur* have been more frequently transferred by man and planted than those of *Q. petraea*.

4.4.2.3 Consequences of human activities

The differences in genetic composition between autochthonous populations and selected provenances are striking. Whereas 70% of diversity can be attributed to between-population differentiation (G_{st}) for autochthonous oaks, this is only 9% for the selected stands. A similar effect of human impact has been detected in neighbouring countries (König *et al*. 2002). In Dutch oak line plantations along the roads, that are known to have originated from mixture of material from different nurseries, the diversity (h_t = 0.74) is similar to that of a data set of western-central Europe. The mixture between nurseries and introduction of foreign seeds has increased the within-population diversity and decreased G_{ST} substantially (0.28). For these Dutch plantations, the spatial genetic structure as identified for *Q. robur* on the data set of whole western-central Europe, was absent. Similar results were obtained for a lowland region in northern Germany, characterised by long-term human impact. These results illustrate that not only levels of within-population diversity raised with human activities but also resulted in a decrease of levels of spatial genetic structure.

4.4.2.4 Indications for autochthony

The overall spatial genetic structure that was established at the outset of post-glacial re-colonisation can be used to evaluate the autochthony of oak populations (Kremer & Goenaga 2002; König *et al.* 2002). The introduction of non-autochthonous material will be particularly evident where there is a large geographic distance between the stand and the region where the corresponding haplotype occurs at high frequency. However, setting objective limits for the geographical distance within which a stand could be called autochthonous is almost impossible,

as this implies the determination of the maximum distance of a long-distance seed dispersal event during post-glacial recolonisation. When the putative origin of a haplotype in a certain region is questioned, more populations surrounding the population with this haplotype can be analysed to provide more detailed information on its distribution. However, in the Flemish region there is no discussion on putative autochthony of haplotypes: only haplotypes 10, 11, 12 and 1 can be considered autochthonous.

Another indicator for non-autochthony is the heterogeneity of the stand. The existence of polytypic stands, which comprise haplotypes of different lineages or contain haplotypes that do not occur in the surrounding stands, provide strong evidence that the basic material is at least partially introduced. In Flanders, only two populations sampled as autochthonous based on field evidence, contained three different haplotypes, all eighteen other putative autochthonous populations contained one or two haplotypes. Furthermore, the field score for autochthony (according to Maes 1993) in both stands with three haplotypes was low and one contained the western most occurrence of haplotype 1. It can therefore be argued that the presence of three haplotypes must be interpreted as the consequence of human impact. Nevertheless, the mingled haplotypes can still be of autochthonous origin.

On the other hand, stands planted with introduced material may also be fixed for an autochthonous haplotype. This illustrates that this type of analysis is especially useful for the identification of populations (or seed lots in a nursery) that are certainly not of local origin. In all other cases, it only provides additional information on the putative autochthonous status as determined on the basis of field and/or historical evidence. Moreover, the molecular evaluation of the autochthonous character of the oak stands studied was very congruent with the evaluation based on field/historical data.

4.5 Nuclear DNA diversity in Flanders and patterns of diversity

4.5.1 Allelic variation at AFLP and microsatellite loci

In this section, general information on the AFLP and SSR loci typed is presented. Results obtained with these markers are presented and discussed in sections 4.5.2 and 4.5.3.

4.5.1.1 Microsatelite loci

All six SSR loci analysed revealed to be highly polymorphic, displaying many alleles (minimum 17 and maximum 31 alleles per locus) and a wide size range of PCR products (see Table 4.4). More alleles and a wider size range were detected than in the original publications where only 25 to 45 (for loci from Steinkellner *et al*. 1997) and 61 trees (for loci from Dow *et al*. 1995) were analysed.

Table 4.4: Allelic diversity of the nuclear microsatellite loci scored in *Quercus* genotypes. *: Steinkellner *et al*. 1997; **: Dow *et al*. 1995.

			Original publications		This study (4.5.2) 8 populations		This study (4.5.3) 30 populations	
Locus	Repeat motif	Linkage group	Nb of alleles	Range of sizes (bp)	Nb of alleles	Range of sizes (bp)	Nb of alleles	Range of sizes (bp)
*MSQ4***	$(GA)_{17}$	G4	11	203-227	17	194-236	18	194-236
*MSQ13***	$(GA)_{14,11}$	G6	12	222-246	17	202-248	17	202-248
*AG104**	$(AG)_{16}AT(GA)_3$	G2	9	176-196	31	183-245	34	181-249
*AG110**	$(AG)_{15}$	G8	7	206-262	19	191-236	23	189-243
*AG15**	$(AG)_{23}$	G9	11	108-152	18	105-153	22	103-153
*AG9**	$(AG)_{12}$	G7	11	182-210	18	178-258	20	178-258

4.5.1.2 AFLP loci

Twenty-eight primer combinations were tested on 5 samples from different populations. The generated fingerprints were evaluated for overall clearness of the banding pattern and the number of polymorphic markers present was recorded (results not shown). Ten primer combinations were chosen for further screening on 40 oaks. The number of markers scored and the degree of polymorphism found in this sample of 40 oaks are shown in Table 4.3. To evaluate the reproducibility of the fingerprints, three completely independent AFLP fingerprints were generated for 10 genotypes, starting from different DNA extractions. Mean reproducibility values (calculated as the percentage of markers that were identical in the three repeats for the same plant) were very high and ranged from 98.6 to 99.1% for the different primer combinations (Table 4.3). Two primer combinations resulted in not scoreable fingerprints due to the amplification of too many and/or faint bands. Four primer combinations were finally chosen for the diversity screening: *Eco*RI-ACA/*Mse*I-CTC, *Eco*RI-ACC/*Mse*I-CAT, *Eco*RI-ACG/*Mse*I-CTC and *Eco*RI-AGG/*Mse*I-CAA.

Table 4.3: Comparison of the results obtained in the preliminary study using 10 primer combinations.

Primer combination	Number of markers scored (in 40 oaks)	Number of polymorphic markers (in 40 oaks)	Reproducibility value[1]
E-ACA + M-CAA	58	32	98.6%
E-ACA + M-CAC	42	29	98.8%
E-ACA + M-CTC*	**49**	**38**	**99.1%**
E-ACA + M-CTG	40	32	99.0%
E-ACC + M-CAT*	**45**	**36**	**98.6%**
E-ACC + M-CTA	Not scoreable	/	/
E-ACG + M-CTC*	**44**	**37**	**98.8%**
E-ACT + M-CAT	43	36	98.8%
E-AGC + M-CTT	Not scoreable	/	/
E-AGG + M-CAA*	**53**	**35**	**98.8%**

[1]Reproducibility values are given as percentage of bands that were identically scored in three independent repeats, starting from different DNA extractions. *: four primer combinations selected for diversity screening.

The use of four AFLP primer combinations on 174 *Quercus* individuals resulted in 170 scoreable markers, of which 164 (96,5%) were polymorphic (least common state minimum 5%). All trees were characterised by a unique banding pattern. A negative correlation between fragment sizes and frequencies (-0.2525, p=0.0009) was detected, which indicates that some degree of homoplasy might be present in the dataset (Vekemans *et al.* 2002). Therefore, data analysis was repeated with only AFLP fragments larger than 150 bp. For this reduced data set (126 markers), correlation between fragment sizes and frequencies was also negative (-0.2233, p=0.0090) and differentiation values for all groups considered were similar to those obtained with the complete dataset (results not shown), suggesting that the potential presence of size homoplasy of AFLP fragments does not result in underestimating genetic divergence between samples. The low average pair-wise correlation value between markers of 0.074 indicates that only a limited amount of information in the data set is redundant.

4.5.2 Nuclear DNA diversity in Flanders and patterns of diversity studied with microsatellite markers

In this section we present the results of the analysis of a total of 30 oak populations (26 Belgian and 4 foreign), including *Q. robur* and *Q. petraea* samples, using six microsatellite loci. Table 4.1 provides a description of the studied stands. In this section we will characterise autochthonous oak stands for their within-population diversity and population genetic structure and compare them with selected oak provenances. Autochthonous populations and selected provenances will be referred to as different 'categories of origin'.

4.5.2.1 Results

4.5.2.1.1 Relationship among genotypes

The PCO plot (Figure 4.5) based on Moran's I relatedness coefficient between autochthonous genotypes does not show a clear separation of both oak taxa. Furthermore, no grouping of genotypes according to their population of origin is observed (results not shown). The first and second axes account for 8.2 and 7.6% of the variation respectively.

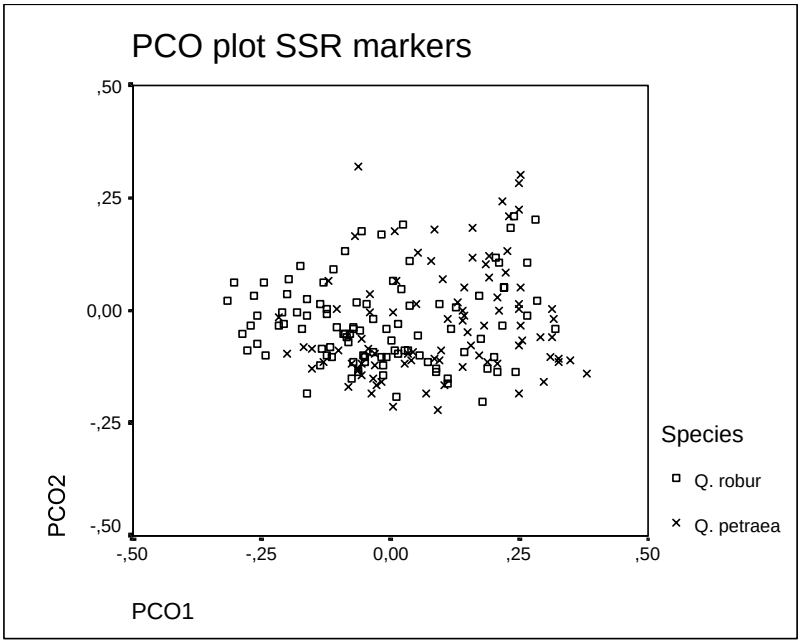


Figure 4.5: Plot of first two principal co-ordinates calculated from the Moran's I relatedness coefficient between genotypes (Hardy & Vekemans 1999) and six polymorphic microsatellite loci. Individuals are coded according to the (dominant) taxonomic species of their population of origin.

The results of the model-based clustering method of Pritchard *et al.* (2000), applied on microsatellite data of autochthonous oaks (selected provenances were not included in the computations) are summarised in Table 4.5. The highest estimate of the likelihood of the data, conditional to a given number of clusters, was obtained when clustering all genotypes into two gene pools. When comparing the inferred gene pools with morphological data (as observed during collection of samples), it was clear that they could be basically interpreted as the two species. However, most genotypes were attributed to both gene pools. The vast majority of genotypes (87.70%) had more than 5% of their alleles attributed to the other gene pool than the dominant one, 32.88% of the genotypes was attributed for more than 25% to the other gene pool. These results can be due to either high levels of admixture between populations and species or to low resolution of the marker data on which the computations were based.

Table 4.5: Summary of the results obtained using the assignment procedure of Pritchard *et al*. (2000) on autochthonous oaks. Figures are proportions of estimated membership to each of two inferred gene pools for genotypes of a given population of origin.

	Population	**Inferred Gene pool 1**	**Inferred Gene pool 2**
Q. petraea	QPAUT117	**0.7988**	0.2012
	QPAUTKB	**0.8145**	0.1855
	QPAUTOG	**0.7281**	0.2719
	QPAUTWS	**0.7385**	0.2615
Q. robur	QRAUTGR	0.2115	**0.7885**
	QRAUT114	0.2307	**0.7693**
	QRAUT123	0.2407	**0.7593**
	QRAUT133	0.2178	**0.7822**
	QRAUT166	0.2593	**0.7407**
	QRAUTHH	0.2800	**0.7205**
	QRAUTKEM	0.3644	**0.6356**
	QRAUTMK	0.2601	**0.7400**
	QRAUTPH	0.2956	**0.7044**
	QRAUTW	0.1864	**0.8136**

4.5.2.1.2 Within-population diversity

Diversity statistics are summarised in Table 4.6. Mean values of diversity statistics are given for both species and both categories of origin (autochthonous populations and selected provenances). Highest genetic diversities were found in autochthonous populations and selected provenances of *Q. petraea*, with mean heterozygosities of

0.86 and 0.85 respectively. For *Q. robur* within-population genetic variation was lower with mean heterozygosities of 0.82 and 0.80 for autochthonous populations and selected provenances respectively. Inbreeding coefficients were significantly different from 0 for all categories of origin, but not for each individual population. A test based on numerical re-sampling was performed to test for differences in the level of within-population diversity between defined groups: between autochthonous populations of both species and between conspecific autochthonous populations and selected provenances. None of the comparisons made resulted significant at the 0.05 level (results not shown).

The analysis made to infer bottlenecks was significant (at $P < 0.05$) for only one out of 14 autochthonous populations (QRAUTPH) under the infinite allele model (IAM). For none of the populations a significant bottleneck effect was detected if the stepwise mutation model (SMM) was assumed. The IAM is more likely to indicate a significant heterozygosity excess based on microsatellite loci and consequently, to be statistically conservative the SMM should be used (Luikart *et al*. 1998).

Table 4.6: Statistics of genetic diversity within *Quercus* populations averaged over six microsatellite loci. Computations were carried out using the software GEN-SURVEY (Vekemans & Lefèbvre 1997).

Population	**N**	***A***	H_o	H_e	F_{is} ¥
		Quercus petraea			
Autochthonous populations					
QPAUT117	26	12.2	0.8358	0.8674	0.0360**
QPAUTKB	23	11.5	0.8409	0.8582	0.0193
QPAUTOG	17	11.5	0.7910	0.8646	0.0942***
QPAUTWS	19	11.5	0.7915	0.8525	0.0712**
	Mean	11.67	0.8148	0.8607	0.0552***
	St.Dev	0.33	0.0273	0.0067	0.0338
Selected provenances					
QPB	15	9.8	0.7906	0.8617	0.0735*
QPLP	13	9.3	0.6724	0.8201	0.1622**
QPQL	9	8.2	0.8519	0.8480	-0.0369
QP09	17	10.0	0.8062	0.8470	0.0484
QP818	18	11.0	0.6847	0.8543	0.1924***
	Mean	9.67	0.7612	0.8462	0.0879***
	St.Dev	1.03	0.0788	0.0157	0.0919
		Quercus robur			
Autochthonous populations					
QRAUTGR	19	9.8	0.8308	0.8176	-0.0149
QRAUT114	27	9.7	0.6998	0.8164	0.1440***
QRAUT123	13	8.8	0.7529	0.8266	0.0821*
QRAUT133	26	10.3	0.7587	0.8105	0.0642*
QRAUT166	18	9.8	0.7658	0.8158	0.0615
QRAUTHH	11	8.5	0.7333	0.8191	0.1065*
QRAUTKEM	10	6.5	0.7567	0.8174	0.0787
QRAUTMK	19	11.0	0.7904	0.8484	0.0694
QRAUTPH	14	8.3	0.7435	0.8388	0.1153**
QRAUTW	22	9.0	0.7869	0.8134	0.0160
	Mean	9.18	0.7619	0.8224	0.0723***
	St.Dev	1.26	0.0355	0.0121	0.04763
Selected provenances					
QRBB	10	7.7	0.6928	0.7703	0.1090*
QRBO	10	6.0	0.7363	0.7799	0.0599
QRDMW	12	7.7	0.7171	0.7922	0.0803*
QRLI	12	8.7	0.7326	0.8361	0.1115*
QRAIS	14	8.0	0.6860	0.8037	0.1329***
QR01	15	9.0	0.7100	0.7913	0.1045**
QR817	19	10.7	0.7422	0.7745	0.0429
	Mean	7.90	0.7146	0.8007	0.1007***
	St.Dev	1.64	0.0210	0.0309	0.0391

N: number of individuals typed; *A*: mean number of alleles per locus; H_o: average proportion of heterozygotes; H_e: average gene diversity; F_{is}: average inbreeding coefficient, ¥: Exact test of departure from Hardy-Weinberg genotypic proportions: *P<0.05, **P<0.01, ***P<0.001, other values not significant.

4.5.2.1.3 Relationship among populations, categories of origin and species

Within each of the four groups (two species each divided in two different categories of origin), a small though significant differentiation at the population level was observed (Table 4.7). Within the selected provenances, a higher differentiation was observed than within the autochthonous origins but re-sampling tests showed that these differences were not significant at the 0.05 level (results not shown).

Table 4.7: Analysis of genetic differentiation based on six SSR loci. Computations carried out with the software GEN-SURVEY (Vekemans & Lefèbvre 1997).

	H_t	H_s	D_{st}	G_{st}	P-value
		Quercus petraea			
Autochthonous populations					
Mean	0.8685	0.8618	0.0067	0.0077	<0.05
Std. Dev.	0.0324	0.0335	0.0068	0.0079	
Selected provenances					
Mean	0.8606	0.8494	0.0113	0.0132	<0.05
Std. Dev.	0.0407	0.0445	0.0120	0.0145	
		Quercus robur			
Autochthonous populations					
Mean	0.8344	0.8285	0.0059	0.0070	<0.05
Std. Dev.	0.0687	0.0670	0.0064	0.0081	
Selected provenances					
Mean	0.8148	0.8051	0.0097	0.0121	<0.05
Std. Dev.	0.0944	0.0957	0.0116	0.0144	

H_t: total diversity; H_s: average diversity within populations; D_{st} average diversity between populations; G_{st}: differentiation of populations within defined groups.

In order to further study the relationship between species and origins, an analysis of hierarchical gene diversity was performed using the software ARLEQUIN (Schneider *et al*. 2000). The distribution of diversity was studied (i) between autochthonous populations of *Q. petraea* and *Q. robur*, (ii) between different categories of origin for *Q. petraea* and (iii) between different categories of origin for *Q. robur*. Of the total differentiation between all autochthonous populations, 78.3% was attributable to the differentiation between species and 21.7% to differentiation of populations within species. If autochthonous populations and selected provenances of the same species were compared, the overall differentiation was lower than in the previous case and more differentiation was attributable to differentiation of populations within a category of origin than between categories. In the case of *Q. petraea* the average gene diversity between populations was 0.010 whereof 63.5% was attributable to differentiation of populations within categories of origin and 36.5% to differentiation between autochthonous and selected stands. For *Q. robur* results were comparable with G_{st} = 0.013, 67.3% of the total differentiation attributable to the diversity between populations of the same origin and 32.7% to differentiation between autochthonous and selected origins.

An additional test based on numerical re-sampling was performed to test for genetic differentiation within and between defined groups. First, autochthonous sessile oak populations were compared with autochthonous pedunculate oak populations and secondly, different categories of origin (autochthonous/selected) were compared within each species. In none of the performed comparisons, the two compared groups differed significantly for their levels of between population differentiation. However, the group of autochthonous populations of *Q. petraea* was significantly differentiated from the group of autochthonous *Q. robur* populations (P-value = 0.0040). For both species, the populations of autochthonous origin were not significantly differentiated from the selected provenances of the same species at the 0.05 level (results not shown).

Finally, neighbour-joining phenograms were constructed for three sub-divisions of the populations: (i) all autochthonous populations, (ii) all *Q. petraea* populations and (iii) all *Q. robur* populations. For each selection of populations, two genetic distances were used as basis for the clustering procedure: Nei's standard genetic distance and $\delta\mu^2$ distance estimates. Results are shown in Figure 4.6 (4.6a to 4.6f). When all autochthonous populations were clustered, the four *Q. petraea* populations grouped together with high bootstrap support values (100% for both distance estimates). Many *Q. robur* populations were clustered closer to *Q. petraea* populations than to conspecific populations. This was observed on both phenograms based on different distance estimates (Figures 4.6a and 4.6d), but the *Q. robur* populations that grouped close to the other species were not the same for both phenograms. On the phenograms including only conspecific populations (Figures 4.6b, 4.6c, 4.6e and 4.6f), no clustering according to origin (autochthonous or selected) could be observed. Different distance measures resulted in very different clustering of populations, mostly with low bootstrap support values.

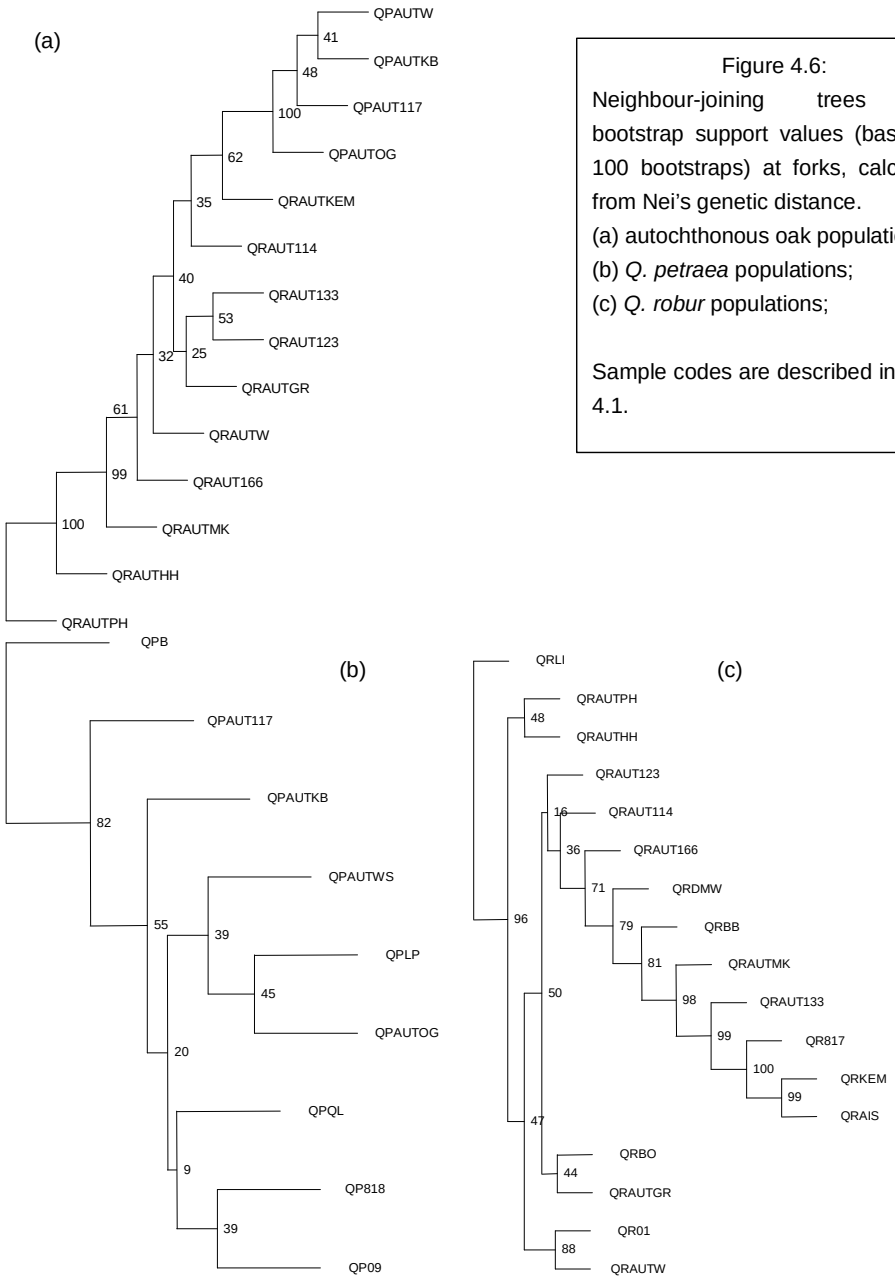


Figure 4.6:
Neighbour-joining trees with bootstrap support values (based on 100 bootstraps) at forks, calculated from Nei's genetic distance.
(a) autochthonous oak populations;
(b) *Q. petraea* populations;
(c) *Q. robur* populations;

Sample codes are described in Table 4.1.

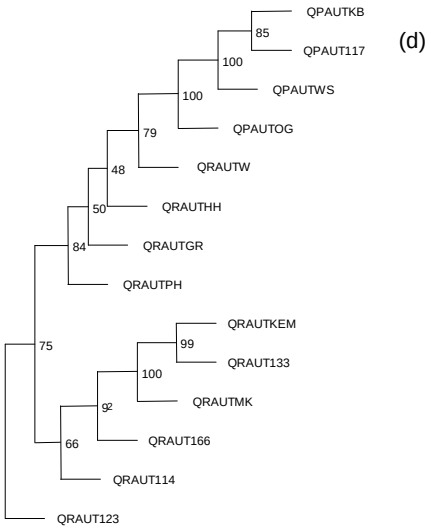
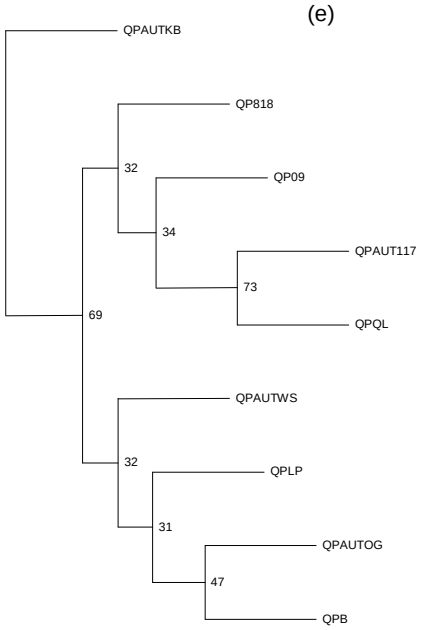
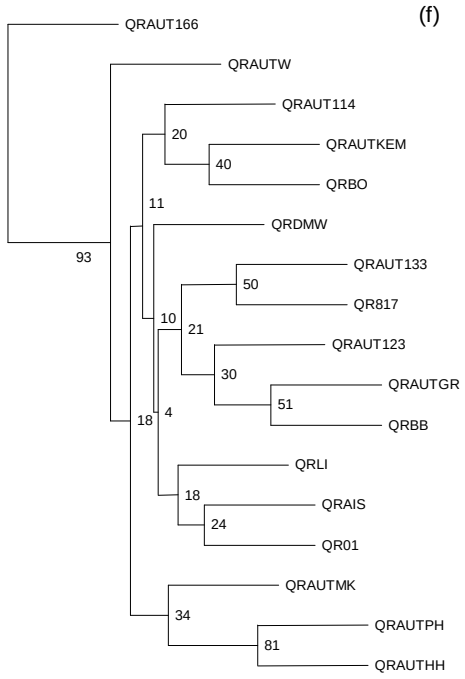


Figure 4.6 (continued): Neighbour-joining trees with bootstrap support values (based on 100 bootstraps) at forks, calculated from $\delta\mu 2$ distance.
(d) autochthonous oak populations;
(e) *Q. petraea* populations;
(f) *Q. robur* populations.

Sample codes are described in Table 4.1.





4.5.2.2 Discussion

4.5.2.2.1 Interspecific variation and species differentiation

On the level of the genotypes, it was not possible to make a clear distinction between both taxa, as demonstrated with a PCO-analysis and using the model-based clustering algorithm of Pritchard *et al.* (2000). However, a low but significant differentiation between autochthonous *Q. petraea* and *Q. robur* populations was observed (F_{st} = 0.0216, P-value = 0.004). This differentiation is higher than the differentiation found between conspecific autochthonous populations (F_{st} = 0.008 for *Q. robur*; F_{st} = 0.007 for *Q. petraea*). Neighbour-joining phenograms cluster all *Q. petraea* populations together with high bootstrap support values (100%). This is observed in the tree based on distance estimates according to Nei (1978, calculated from allele identities) and according to Goldstein (1995, calculated from allele sizes). Some *Q. robur* populations however, are grouped closer to the *Q. petraea* group than to conspecific populations. This is the case for both neighbour-joining trees but different populations are grouped with *Q. petraea* depending on the distance measure applied. These findings are in agreement with results of former studies on genetic divergence between these oak species in Europe, where it was shown that (i) both taxa are closely related and difficulties arise when differentiation is studied at only a few genomic loci and (ii) both taxa are still genetically differentiated despite past and ongoing hybridisation events. More specifically, studies on a comparable geographic scale have been conducted in Switzerland (Finkeldey *et al*. 2001a) and Denmark (Siegismund *et al.* 2001), analysing respectively seventeen and six allozyme loci. Finkeldey found F_{st} values of 0.043 for species differentiation and of 0.017 between conspecific populations. In Denmark, the differentiation between species (F_{st} = 0.235) and conspecific populations (F_{st} = 0.022) was higher. Siegismund *et al.* (2001) mention the high variation of F_{st} values across loci, and the inclusion of one locus with higher differentiation values between species raised the mean value substantially. This example illustrates the caution that should be taken when comparing results of genetic differentiation based on different loci or different marker systems. As in the present study, problems were encountered by Siegismund *et al.* (2001) for the assignment of individual oaks to a species.

A survey of genetic differentiation on a European scale was conducted by Mariette *et al.* (2002b), using six SSR loci. They found significant differentiation between species and conspecific populations but surprisingly, at the European scale, the divergence between species was lower (F_{st} = 0.013) than between populations (F_{st} = 0.023 for *Q. robur* and 0.020 for *Q. petraea*), illustrating the high genetic diversity within both oak species and the importance of hybridisation events. Also at an European scale, Muir *et al.* (2000) showed that by using 20 SSR loci, not only the oak populations, but even the individual oak trees could be correctly assigned to

their taxonomic group. These results illustrate that six microsatellite markers, as applied in the current study, might not be powerful enough to assign oak populations and/or individual trees to the right species. Six loci, however, might be enough to test for significant levels of differentiation between species and among populations. If more SSR markers are applied, the small differences in allele frequencies as present at each locus add up and make the differentiation between species possible, even on the individual level. Soon all 49 published oak microsatellite loci will have been positioned on a genetic linkage map (Steinkellner *et al*. 1997; Kampfer *et al*. 1998; Barrenche *et al*. 1998 and ongoing research within European project 'OAKFLOW'), on which also the regions of species divergence are being mapped (Santaigne *et al*. 2002). At this point a selection of SSR loci located in genomic regions involved in species divergence can be made to increase the power for species determination. However, if assignment of individuals (or populations) to a species is the primary research goal, application of AFLP markers might also be a efficient approach (see section 4.5.3).

4.5.2.2.2 Within-population variation

The within-population diversity of autochthonous *Q. petraea* populations was higher than for autochthonous *Q. robur* populations, but these differences were not significant for any of the diversity statistics computed. Average heterozygosities of 0.86 and 0.82 were found for *Q. petraea* and *Q. robur* populations respectively. Although caution should be taken when comparing results based on different loci or marker systems, it can be stated that similar levels of diversity were detected as in other oak studies. Three groups published recently within-population diversity estimates, calculated from the same set of six SSR loci. Four of these loci were incorporated in the present study. Streiff *et al.* (1998) recorded mean expected heterozygosities of 0.87 for both species in a French forest, Bakker *et al.* (2001b) estimated expected heterozygosities for Dutch autochthonous *Q. robur* populations as 0.86 and Mariette *et al.* (2002b) reported mean H_E values of 0.90 for *Q. petraea* and 0.88 for *Q. robur* averaged over the European populations studied. The observed differences in within-population diversity as observed by the latter study were not significant, what is in accordance with the present oak study. Diversity surveys conducted with isozyme markers found similar values of within-population diversity for both species (Siegismund *et al.* 2001; Finkeldey 2001a) or higher values within *Q. petraea* populations (Kremer *et al.* 1991). Studies with DNA markers (other than SSRs) reported higher within-population diversity for *Q. petraea* (Moreau *et al.* 1994; Bodénès *et al.* 1997a). These observations are in agreement with life-history traits of both oak species (see also 4.5.3.2.4 for further discussion).

Furthermore, for only one population a significant bottleneck effect was found (for population QRAUTPH), but only if the infinite allele model of evolution was assumed for the microsatellite loci studied. It can therefore be concluded that also the comparison of heterozygosity and allelic richness did not show any recent bottleneck effects in the autochthonous oak popualtions.

For many populations of both species, a significant deviation from Hardy-Weinberg proportions was observed with an excess of homozygotes. Mean values of the inbreeding coefficient F_{is} were 0.06 for *Q. petraea* and 0.07 for *Q. robur*. These values were not significantly different. This finding is in agreement with other oak studies. Two papers report on positive inbreeding coefficients based on SSR markers. Streif *et al.* (1998) found a mean F_{is} value of 0.07 in a mixed oak forest; Mariette *et al.* (2002b) report mean inbreeding coefficients of 0.084 and 0.089 for European populations of *Q. petraea* and *Q. robur* respectively. An often reported cause of positive values of the inbreeding coefficient when estimated from allele frequencies at microsatellite loci, is the presence of null alleles (Bruford *et al.* 1998). Although it cannot be ruled out that the presence of null alleles is partially responsible for the present observation of positive F_{is} values, this is presumably not the only reason in the present data set. For the oak SSR loci analysed in this study, null alleles have only been reported for locus *MSQ4* in *Quercus macrocarpa* (Dow & Ashley 1996). No evidence of null alleles has been reported for other loci. Locus *MSQ4* is also the locus with the highest mean F_{is} value over all populations (F_{is} = 0.172, mean of six loci 0.066) in the present data set. Average F_{is} values still deviated significantly from Hardy-Weinberg expections when analyses were repeated without this particular locus. Diversity studies conducted with isozymes also revealed heterozygote deficits with mean values of F_{is} of 0.016 for a mixed oak forest in Switzerland (Finkeldey 2001b) and 0.07 in a mixed oak forest in France (Streiff *et al*. 1998).

The fact that positive inbreeding coefficients are found in other studies using SSR and also based on isozyme markers, suggests the presence of a homozygote excess in many oak populations, wherefore two explanations can be put forward. Autocompatibility is very rare for both species (Steinhoff 1993). Based on the maximum estimates of selfing rates for both species (1% for *Q. petraea* and 5% for *Q. robur*), the inbreeding coefficient as a result of selfing would be only 0.005 for *Q. petraea* and 0.026 for *Q. robur*. However, preferential mating of neighbouring, related trees is a plausible explanation for slightly positive F_{is} values (Bacilieri *et al.* 1994). A weak spatial genetic structure has indeed been observed in a natural population of *Q. petraea* in the distance classes up to 20m (Bacilieri *et al*. 1994; Streiff *et al.* 1998). Also the sympatric occurrence of two differentiated species that can hybridise might be a reason for positive F_{is} values. Genotypic structure of the

offspring generation will be affected by a 'Wahlund effect' if genetically differentiated populations, which are at least partially isolated from each other, contribute to the next generation (Hattemer 1982 in Finkeldey 2001b). This is the case for the two oak taxa that are differentiated for their nuclear genetic information but are shown to hybridise as well. In mixed populations, this effect will inflate F_{is} values. Finkeldey (2001b) found the highest inbreeding coefficients in mixed stands, what supports this theory. In most of the Flemish forests, one oak species dominates but a small mixture of the other oak species is often present, suggesting that part of the observed positive inbreeding coefficients may be due to the partial isolation of gene pools of both oak taxa.

In summary it can be stated that the Flemish autochthonous oak populations have comparable within-population variation as other European oak populations. The fact that most of the Flemish stands are small populations and may have gone through bottlenecks after historical deforestation has not decreased their genetic diversity. These autochthonous relict populations can thus be used as seed sources for the creation of new forests, without jeopardising the genetic variability of future oak populations.

4.5.2.2.3 Comparison of genetic variation within and between autochthonous populations and selected provenances

For both species, a similar trend can be observed when comparing oak stands of different categories of origin. In general, selected provenances exhibit lower within-population diversity levels and have higher heterozygote deficits than autochthonous populations of the same species. However, none of these differences proved to be significant. Levels of population differentiation are similar among autochthonous stands and among selected provenances for both species. Furthermore, the autochthonous populations analysed did not form a gene pool significantly differentiated from conspecific selected provenances, as illustrated in the neighbour joining phenograms and in the re-sampling tests.

The fact that autochthonous populations did not form a distinct gene pool is an expected result. First of all, the exact origin of the trees of the selected provenances is often unknown. Although the chloroplast results indicate that the majority of selected provenances are the result of a mixture of different populations, they might be created with a mixture of autochthonous material. This can be true for all selected stands that do not contain allochthonous chloroplast variants, that is all but one of the studied provenances. Furthermore, high gene flow (through pollen) results in a weak population genetic structure even at large geographical scales, characterised by high within-population diversities and low levels of differentiation

among populations. Therefore, the allelic frequencies of populations imported from nearby regions, will not necessarily be different from the frequencies characteristic for autochthonous populations.

The import of acorns into the Flemish region is known to occur since many centuries. Again through high pollen flow, possible differences in allelic frequencies between autochthonous and introduced allochthonous material will have been homogenised after a few generations. However, some of the oldest oak populations in Flanders consist of coppice stools and their age has been estimated at a few thousand years (pers. comm. B. Maes, Ekologisch Adviesburo Maes). It is unlikely that at the time when these oaks germinated man transported acorns over long distances and it can therefore be assumed that the genetic composition of these old trees was not the result of hybridisation with trees from distant regions. These ancient populations can thus be seen as direct descendants of truly autochthonous oaks that colonised during post-glacial migration. However, even if the autochthonous origin was limited to these extremely old oak populations, numerical re-sampling tests showed that they were not significantly differentiated from selected stands.

Although the trend was not significant, it is rather surprising that the selected provenances have slightly lower diversity values and higher inbreeding coefficients than conspecific autochthonous populations. The knowledge on the mixed origin of most of the selected stands intuitively suggests that diversity values for nuclear markers are higher, despite the low diversity present between populations. If acorns of only a small number of oak trees have been harvested in each location of collection and used for the establishment of a new forest, this could explain the tendency towards lower within-population diversity than in spontaneous populations. Expected heterozygosity levels in half-sib families are lower than in open-pollinated populations of *Q. robur* (Bakker *et al.* 2001b). Thus, if a substantial part of the planted seedlings are half-sibs, this could result in reduced levels of genetic diversity in the newly created oak stand, in comparison to a ‘natural’ forest. Moreover, if subsequently natural regeneration of the oak population takes place, the following generations may exhibit even higher excess of homozygotes than in forests with a ‘natural’ origin due to mating of related trees. These findings are in contradiction with the widespread opinion that selected provenances are to be preferred over autochthonous stands as seed source because bottleneck effects would have decreased the within population diversity of autochthonous populations.

The use of chloroplast information that reflects the evolutionary history of populations has been proposed as basis for the delineation of ESUs in many concepts (e.g. Moritz 1994). However, it can be questioned whether chloroplast

haplotypes (or cp-lineages) are the right bases for the delineation of ESUs in the specific case of oaks. In this study, no significant genetic differentiation was found between autochthonous populations and selected provenances for both oak species. A weak population genetic structure was observed for autochthonous populations of both species from both nuclear marker systems, reflecting the high gene flow (through pollen) among populations. Furthermore, in the specific case of oaks it has been shown on a European scale that local selection pressures, acting on the installed populations, and pollen flow have progressively erased the initial differentiation that existed among the three refuge zones after the last glaciation. As a result, there is no association any more between chloroplastic divergence and phenotypic traits (Kremer *et al*. 2002). Current patterns of differentiation for (adaptive) traits and for nuclear markers have been established which are totally different from those in place immediately following colonisation. As a consequence, the delineation of ESUs on the sole basis of cp-DNA polymorphisms may not reflect the enormous levels of diversity present nowadays in forestry populations and would ignore the possible local adaptations that might have arisen after recolonisation. The delineation of ESUs should also consider present-day levels of nuclear diversity and patterns of differentiation among populations.

4.5.3 Comparison of diversity and differentiation measures as detected with AFLP and microsatellite marker systems

In this section we compare the performance of AFLP and microsatellite markers to distinguish between *Q. robur* and *Q. petraea* and to estimate diversity and population differentiation parameters. Therefore, analysis of four *Q. petraea* and four *Q. robur* populations, all of putative autochthonous origin (see Table 4.1), were carried out and compared with SSR typing of these populations as presented in section 4.5.2. In addition, for one population of each species, the molecular marker data are compared with measurements of leaf morphology parameters. An adapted version of this part has been published in Coart *et al.* (2002).

4.5.3.1 Results

4.5.3.1.1 Relationships among genotypes

PCO plots based on AFLP and microsatellite data give very different results. The PCO analysis based on AFLP markers (Figure 4.7) clearly separated sessile and pedunculate oaks into two differentiated groups, leaving only 9 (out of 174) oaks to be apparently 'misclassified'. These atypical samples represent most likely minor mixtures in a stand dominated by the opposite oak species, as will be confirmed by their morphological evaluation (see below). The two plotted axes accounted for respectively 7.0 and 6.5% of the variation present at the molecular level. As expected, based on the results presented in section 4.6, the PCO plot based on microsatellite data (Figure 4.8) is less informative; no grouping of individuals according to species or population can be seen. The two plotted axes accounted only for respectively 4.9 and 3.6% of the variation present at the molecular level.

The results of the model-based clustering method of Pritchard *et al*. (2000) applied on microsatellite and AFLP data are summarised in Table 4.8. The highest estimate of the likelihood of the data, conditional on a given number of clusters, was for both marker systems obtained when clustering all genotypes into two gene pools. For AFLP data, only 4 genotypes (2.3%) had more than 5% of their alleles attributed to the other gene pool. Nine genotypes were completely assigned to the other gene pool than the dominant one in the location where they were collected; these genotypes were also identified on the PCO plot. Based on SSR data, 59 genotypes (38.0%) had more than 5% of their genetic information attributed to the other gene pool. The 9 genotypes that were identified as outliers based on AFLPs were again assigned for the most part to the non-dominant gene pool. However, 11 other oaks were also assigned for more than 50% to the opposite gene pool.

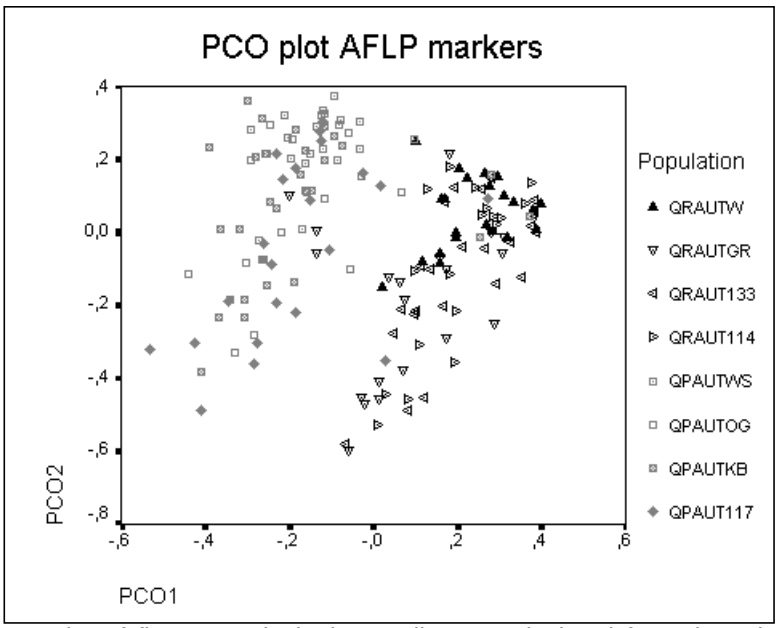


Figure 4.7: Plot of first two principal co-ordinates calculated from the relatedness coefficient r between genotypes (Lynch & Milligan 1994) of presence/absence data of 170 AFLP markers. See Table 4.1 for population names.

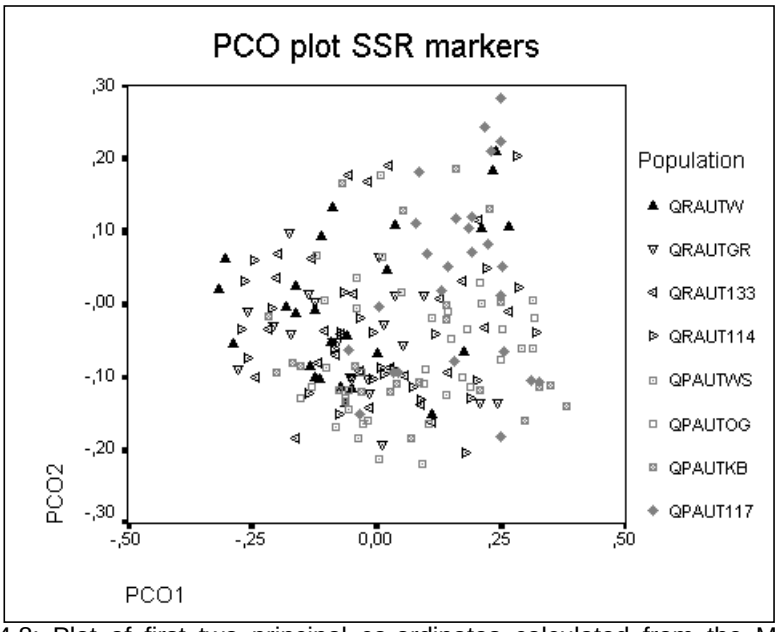


Figure 4.8: Plot of first two principal co-ordinates calculated from the Moran's I relatedness coefficient between genotypes (Hardy & Vekemans, 1999) and six polymorphic microsatellite loci. See Table 4.1 for names of populations.

Table 4.8: Summary of the results obtained using the assignment procedure of Pritchard *et al*. (2002) on eight autochthonous oak populations. Figures are proportions of estimated membership to each of two inferred gene pools for genotypes of a given population of origin.

	SSR		AFLP	
Population	**Gene pool 1**	**Gene pool 2**	**Gene pool 1**	**Gene pool 2**
QPAUT117	**0.7659**	0.2341	**0.9060**	0.0940
QPAUTKB	**0.8068**	0.1932	**0.9030**	0.0970
QPAUTOG	**0.6627**	0.3373	**0.8980**	0.1020
QPAUTWS	**0.7040**	0.2960	**0.9700**	0.0300
QRAUT114	0.1080	**0.8920**	0.0000	**1.0000**
QRAUT133	0.1553	**0.8447**	0.0000	**1.0000**
QRAUTGR	0.1435	**0.8565**	0.1500	**0.8500**
QRAUTW	0.1172	**0.8828**	0.0180	**0.9820**

For two populations, a detailed study of leaf morphology was conducted. Thirty trees from one sessile oak (QPAUTKB) and one pedunculate oak (QRAUTW) population were characterised (Table 4.1). Secondary characters were calculated, based on the measured leaf parameters (see Table 2.6). The discriminatory power of these parameters is shown in Table 4.9. All but one character (LS, lamina shape) were significantly different between sessile and pedunculate oaks. PCO analysis based on leaf morphology data (Figure 4.9) divided the samples into two completely separated groups, without any intermediate positions. The first axis, representing the species divergence, explains 55.0% of the variation and axis 2 stands for 20.1% of the variation present in the morphology data set. As expected, the 3 oaks of stand QPAUTKB that were identified as *Q. robur* based on AFLP genotyping were here again positioned within the *Q. robur* group.

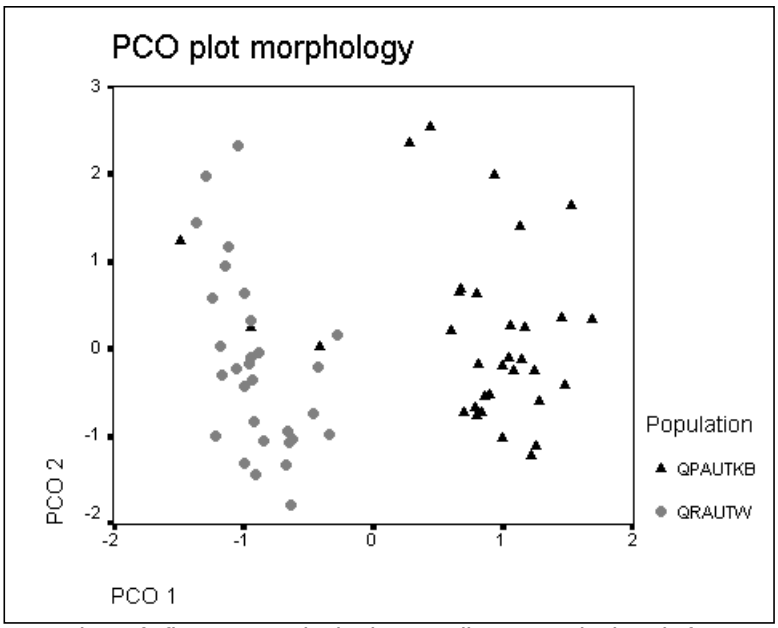


Figure 4.9: Plot of first two principal co-ordinates calculated from Euclidean distances based on 6 morphologic characters of 60 trees (populations QRAUTW and QPAUTKB). Individuals are marked according to their population. Names of stands refer to Table 4.1. The same oaks of QPAUTKB are clustered within the Q. robur group as on the PCO plot based on AFLP markers (Figure 4.6a).

Table 4.9: Descriptive statistics of phenotypic characters and significance values for comparison between species. Five leaves of 30 sessile (QPAUTKB) and 30 pedunculate (QRAUTW) oaks were analysed. Mean values per tree were used for data analysis. See Table 2.6 for abbreviations.

Parameter	Mean	Std. deviation	Minimum	Maximum	Mann-Whitney U	Significance (2-tailed)
PUB	2.63	1.40	1.00	5.00	40.00	0.000
BS	22.71	8.63	8.32	55.16	269.50	0.002
LDR	3.03	0.60	2.13	4.71	118.00	0.000
LS	1.68	0.10	1.45	1.90	474.00	0.773
PR	0.083	0.044	0.01	0.18	27.00	0.000
PV	27.92	24.06	0.00	89.69	57.00	0.000

Based on the AFLP and leaf morphology results presented in section 4.7.1.1, reduced populations were defined. Individuals, identified as belonging to the other oak species than the majority of their population of origin, were excluded from the data set for calculation of within-population diversity statistics and relationships among populations.

4.5.3.1.2 Within population diversity

Diversity statistics are summarised in Table 4.10. High variation at AFLP loci was recorded, with on average 83.7% polymorphic loci within populations. Very similar values of gene diversities (0.2910 for *Q. petraea*, 0.2868 for *Q. robur*) were found in all populations of both species (CV is only 2.2 % for the complete data set). High genetic variation at microsatellite loci was also observed, with a mean number of alleles per locus equal to 10.52 and an average gene diversity of 0.83 (CV 2.9% for the complete data set). Furthermore, based on SSR markers strikingly similar levels of genetic variation were found within conspecific populations (CV of gene diversity is only 0.8% and 0.4% for *Q. petraea* and *Q. robur* respectively), with higher diversity recorded in *Q. petraea* populations (0.86 for *Q. petraea*; 0.81 for *Q. robur*).

Table 4.10: Genetic diversity estimates within *Quercus* populations for AFLP and microsatellite loci.

	Microsatellites					**AFLP**			
Samples	**N**	***A***	H_o	H_e	F_{is} ¥	**N**	***NPL***	***PLP***	H_j
				***Quercus petraea* populations**					
QPAUT117	24	11.7	0.8426	0.8599	0.0195	20	143	84.1	0.3128
QPAUTKB	20	10.7	0.8316	0.8494	0.0197	28	141	82.9	0.2875
QPAUTOG	17	11.5	0.7910	0.8646	0.0942**	14	150	88.2	0.2970
QPAUTWS	19	11.5	0.7915	0.8525	0.0712*	21	143	84.1	0.2667
Mean		11.33	0.8142	0.8566	0.0511***		144.25	84.83	0.2910
S.D.		0.45	0.0268	0.0069	0.0376		3.95	2.32	0.0096
C.V.		0.040	0.0330	0.0080	0.7358		0.0274	0.0273	0.0330
				***Quercus robur* populations**					
QRAUT114	27	9.7	0.6998	0.8164	0.1440***	18	135	79.4	0.2912
QRAUT133	26	10.3	0.7587	0.8105	0.0642*	24	145	85.3	0.2825
QRAUTGR	19	9.8	0.8186	0.8090	-0.0116	17	143	84.1	0.3108
QRAUTW	22	9.0	0.7869	0.8134	0.0160	22	139	81.8	0.2628
Mean		9.71	0.7660	0.8123	0.0532***		140.5	82.65	0.2868
S.D.		0.55	0.0505	0.0033	0.0682		4.44	2.61	0.0100
C.V.		0.0566	0.0659	0.0041	1.2820		0.0316	0.0316	0.0349
				All *Quercus* populations					
Mean		10.52	0.7901	0.8345	0.0522***		142.38	83.74	0.2889
S.D.		0.99	0.0454	0.0242	0.0510		4.37	2.56	0.0065
C.V.		0.0941	0.0575	0.0290	0.0977		0.0307	0.031	0.0224

N: number of individuals typed; *A*: mean number of alleles per locus; H_o: average proportion of heterozygotes; H_e: average gene diversity; F_{is}: average inbreeding coefficient, ¥: Exact test of departure from Hardy-Weinberg genotypic proportions: *P<0.05, **P<0.01, ***P<0.001, other values not significant; NPL: number of polymorphic loci at the 5% level; PLP: proportion of polymorphic loci; H_j: same as H_e for AFLP markers.

The computation of Spearman's correlation coefficient showed that the ranking of diversity estimates calculated from SSR data (*A*, H_o and H_e) were not significantly

correlated with ranking of diversity estimates calculated from AFLP data (*PLP* and H_j) at the 0.05 level. This indicates that the two marker systems target different genomic regions.

4.5.3.1.3 Relationships among populations

Results of neighbour-joining trees based on both marker systems are shown in Figure 4.10. In contrast with the PCO plots, highly concordant results were obtained with AFLP and SSR data. In the phenogram based on AFLP frequency data, all *Q. petraea* populations grouped together with very strong bootstrap support (100%). One *Q. robur* population (QRAUTGR) was clustered close to the *Q. petraea* group. Very similar trees were obtained with both distance measures used to analyse the SSR data (Nei standard genetic distance and $\delta\mu^2$). In both trees, populations of *Q. petraea* clustered together with high bootstrap support values (96% and 100% for Nei standard genetic distance and $\delta\mu^2$, respectively). Similarly to the AFLP phenogram, in both trees derived from SSR data, the population QRAUTGR was clustered close to the *Q. petraea* populations. As for the tree based on AFLP data, grouping of populations (within species) was less stable, although in the tree based on $\delta\mu^2$ distances, high values were obtained for clustering of *Q. petraea* populations. These results fully agree with the results obtained in section 4.5.2 for a bigger set of populations.

Differentiation statistics are summarised in Table 4.11. Analysis of the population genetic structure indicated a significant differentiation between the two species based on AFLP data (F_{ST} = 0.0717, P<0.0001). The differentiation among populations within species was lower though also significant for both species (F_{ST} = 0.0185 (P<0.0001) for *Q. petraea*; F_{ST} = 0.0193 (P=0.003) for *Q. robur*). Based on microsatellite data, differentiation between defined groups was lower. The differentiation between species was much lower (G_{st} = 0.0250, P<0.05) but significant as also shown in section 4.5.2. The differentiation among conspecific populations is higher for *Q. robur* (G_{st} = 0.0146, P<0.05) than for *Q. petraea* (G_{st} = 0.0102, P<0.05).

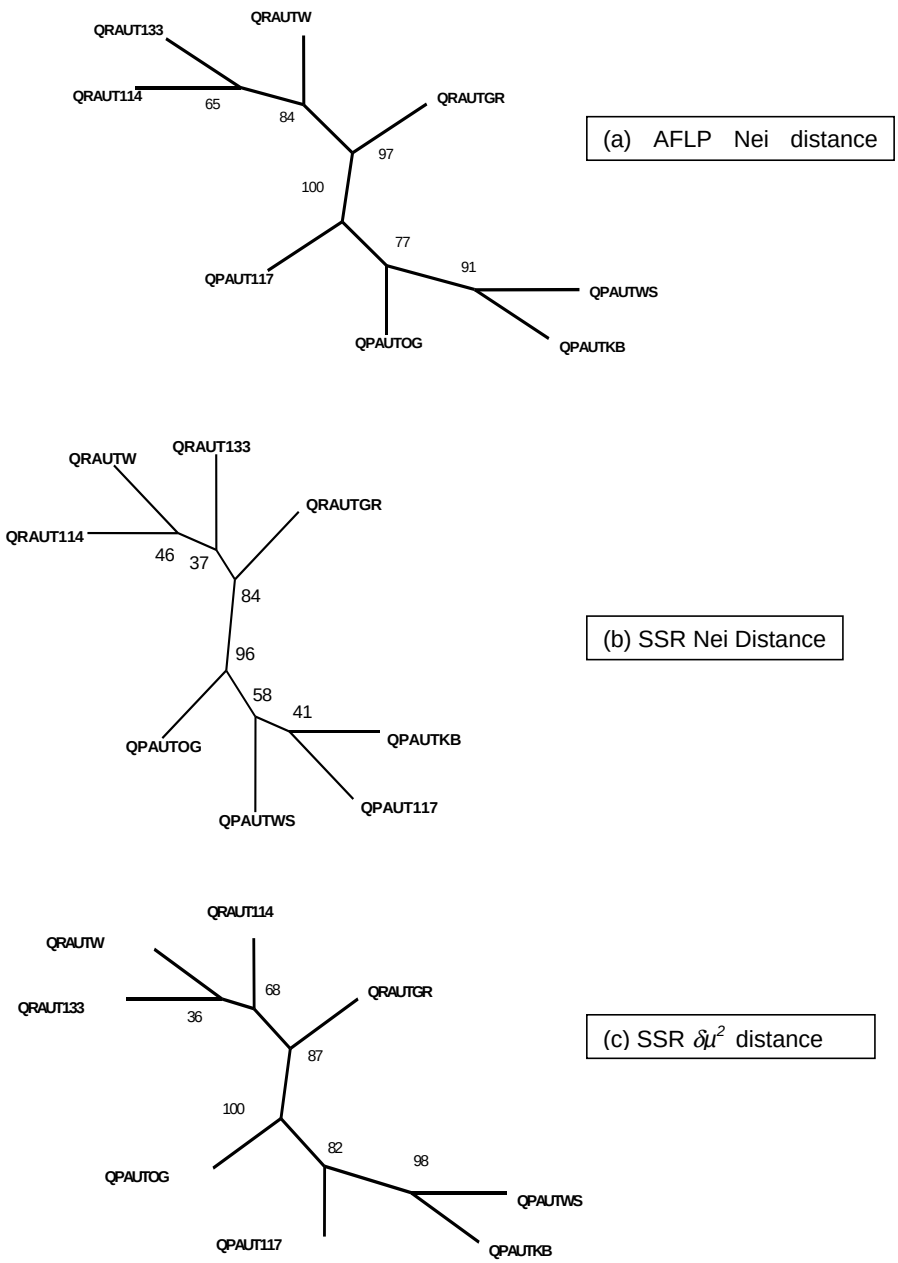


Figure 4.10: Neighbour-joining trees with bootstrap support values at forks (based on 100 bootstraps). (a) Calculated from AFLP data and Nei's genetic distance; (b) from SSR data and Nei's distance; (c) from SSR data and δμ2 distance. Sample codes are described in Table 4.1.

Table 4.11: Analysis of genetic differentiation

(a) based on 170 AFLP markers according to Lynch & Milligan (1994) using AFLP-SURV

(b) based on 6 SSR loci according to Nei & Chesser (1983) using GEN-SURVEY.

(a) AFLP

Comparison	**N**	H_t	H_w	H_b	F_{st}	**P-value**
(i) *Q. robur* vs. *Q. petraea*	2	0.3025	0.2808	0.0217	0.0717	<0.0001
(ii) Among *Q. petraea* populations	4	0.2971	0.2916	0.0055	0.0185	<0.0001
(iii) Among *Q. robur* populations	4	0.2925	0.2868	0.0056	0.0193	=0.003

(b) Microsatellites

Comparison	**N**	H_t	H_s	D_{st}	G_{st}	**P-value**
(i) *Q. robur* vs. *Q. petraea*	2	0.8608	0.8395	0.0213	0.0250	<0.05
(ii) Among *Q. petraea* populations	4	0.8685	0.8618	0.0067	0.0077	<0.05
(iii) Among *Q. robur* populations	4	0.8268	0.8156	0.0112	0.0137	<0.05

N: number of samples; H_t: total diversity; H_w/H_s: average diversity within populations; H_b/D_{st} average diversity between populations; F_{st}/G_{st}: differentiation between defined groups.

4.5.3.2 Discussion

4.5.3.2.1 Use of AFLP markers for population genetic studies of Flemish oak populations

The major drawbacks of using dominant markers such as AFLPs for population genetics is that they only provide information on genotypic frequencies but not on the underlying allelic frequencies. On the other hand, the power of AFLP, generating a huge amount of markers in comparison to other molecular marker systems makes it an important tool for population studies. In order to estimate population genetic parameters one can either assume Hardy-Weinberg equilibrium (for an outcrossing species) or use estimates of the inbreeding coefficient (F_{is}) from other studies performed with co-dominant markers to calculate allelic frequencies (Vekemans *et al.* 2002). In this case it is assumed that the F_{is} values estimated in different populations and using different marker systems are the same. Alternatively, an AMOVA approach (Excoffier *et al.* 1992), assuming input of co-dominant data, could be applied in order to avoid the inaccuracies of estimating F_{is} values. This method can, however, give significantly biased estimates of population genetic parameters when dominant marker data are used, especially for outcrossing species such as oaks. For this study, an estimate for F_{is} as calculated from microsatellite data (mean value F_{is} = 0.05) was used as input value. However, results from this diversity study based on AFLP markers were published before microsatellite data were available for Flemish oak populations (Coart *et al.* 2002). At that time, it was decided to use an average value for F_{is} (F_{is} =0.15), computed in other natural oak populations distributed over the geographic range of the oak species using isozyme and microsatellite markers (Bacilieri *et al.* 1994; Streiff *et al.* 1998). The use of these different values for the inbreeding coefficient resulted in comparable parameters (higher differentiation values for higher inbreeding coefficient) and identical conclusions.

4.5.3.2.2 Interspecific variation and species differentiation

The present AFLP data set revealed a clear differentiation of the Flemish oak gene pool in sessile and pedunculate oaks. PCO analysis and model-based clustering on the AFLP data was able to assign the individual trees to the ‘correct’ species. At first sight, only nine trees were atypical. These nine samples displayed an AFLP fingerprint that did not fit in the stand where they were sampled. Most of these oaks had also been designated as the other species than the dominant one in the stand where they were located during the collection of the material in the field. For the three atypical oaks that were also characterised morphologically, the results of both genotypic and phenotypic classification were identical at the individual tree level. The combination of the AFLP results, the morphological study and additional field information show clearly that these apparently atypical oaks were in fact classified

correctly based on their AFLP fingerprints, and that they represent minor mixtures in a stand dominated by the other oak species.

The assignment of individuals to a species was less straightforward when using microsatellite data. The PCO plot with the first two synthetic variables did not show a clustering of individuals according to their species. The model-based clustering approach divided the trees into two gene pools, roughly representing the two species, but failed to reach a resolution similar to AFLP markers. The clustering based on frequency data of AFLP alleles in populations, grouped the *Q. petraea* populations together, with a bootstrap value of 100%. Very concordant results were obtained based on microsatellite data, both when using distance measures between populations based on allele size ($\delta\mu 2$) and allele identity (Nei's standard genetic distance).

The differences between species proved to be highly significant for estimates based on both marker systems, although the differentiation is much higher when calculated from AFLP data (F_{ST} = 0.0717 and G_{ST} = 0.0185 for AFLP and SSR respectively). The differentiation detected in the present study using AFLP markers was stronger than formerly observed in these oak species using other molecular marker techniques (with isozymes: e.g. Zanetto *et al.* 1994, with RAPD: Moreau *et al.* 1994). AFLP, generating many polymorphic markers, was successfully applied for identifying the two major taxonomic units present in the Flemish oak gene pool. This supports the hypothesis that the formerly often observed discrepancy between the discriminatory power of phenotypic (high) and genetic (low) differentiation between both species can be due to a sampling effect: too few loci were investigated with molecular markers to reveal the same divergence on the molecular level.

An important explanation for the observed discrepancy of power for species discrimination based on both marker systems can be found in the higher number of loci typed with AFLP. Most of the AFLP markers are assumed to be neutral markers (Vos *et al.* 1995), but nevertheless, the likelihood that some markers are linked with genomic regions of species divergence cannot be excluded. The mean difference in marker frequency between both species for AFLP markers was only 0.11 whereas the ten most discriminating markers (ten markers showing the highest difference in frequency between both species) showed marker frequency differences ranging from 39 to 71%. As a result the overall differentiation value is higher for AFLP than for SSR markers. This has been demonstrated in other studies (e.g. Isabel *et al.* 1999) and recently also for oak populations (representing their European distribution range) by Mariette *et al.* (2002b). A similar explanation can be proposed for the higher F_{ST} values for AFLP markers than SSR markers when describing population

differentiation. As oak populations are shown to be highly differentiated for phenological and growth traits (Ducousso *et al.* 1996), some AFLP markers might be associated with these regions of adaptive divergence.

Microsatellite markers are usually considered as evolutionary neutral DNA markers and provided that they do not 'hitchhike' with regions involved in species differentiation, they will fail to distinguish the species. The use of more SSR loci would increase the power for species differentiation of *Q. robur* and *Q. petraea*, adding up the effect of small differences in allele frequencies found for each locus, as has been shown by Muir *et al.* (2000).

4.5.3.2.3 Identification of diagnostic AFLP markers for species assignment

None of the markers screened was species-specific, and none of the (sets of) primer combinations had enough discriminatory power to make the same clear differentiation between species as found using the whole data set (4 primer combinations). However, if only the ten most species-specific markers out of the complete data set (10 markers showing the highest difference in frequency between both species) were used, similar results were obtained as when the complete data set of 170 polymorphic markers was used. When selecting species-diagnostic AFLP markers, a selection is made towards markers associated with genomic regions of species-divergence. Hence, these markers are more likely to be associated with the phenotype, if the observed morphological differences are truly diagnostic characters and not the result of phenotypic plasticity or modifications.

Although many AFLP markers were generated using 4 AFLP primer combinations, only a relatively small portion of the genome was sampled by the 170 polymorphic AFLP-markers scored. Assuming a uniform spread of these markers across the oak genome, the 170 polymorphic markers used correspond to a density of 1 AFLP-marker every 7 cM. For rapidly screening more regions of the genome, 60 different AFLP primer combinations were tested on DNA bulks of each species. In this way, 1 AFLP-marker was screened every 0.22 cM. Six of the primer combinations identified a total of 8 putative species-specific markers. However, none of the 8 markers selected in the bulks was confirmed as species-specific when individual trees were analysed. The markers were present in maximum 8 out of the 10 trees from the same species and were often too faint to score on individual plants (results not shown). These results confirm the previous studies on the divergence of sessile and pedunculate oaks. Up to now, no single marker has been identified which allowed differentiating the two species at the individual tree level (Zanetto & Kremer 1995 using isozymes; Moreau *et al.* 1994 using RAPD; Bodénès *et al.* 1997a using 2800 PCR-amplification products; Muir *et al.* 2000 using SSR markers and Muir *et al.*

2001 using rDNA, Bakker *et al.* 2001a using AFLP). As stated, the overall nucleotide divergence is limited (Bodénès *et al.* 1997a) and past and recent hybridisation events are likely to have occurred, what renders the chance of finding species-specific markers with a random PCR-technique low.

4.5.3.2.4 Within-population variation

Diversity estimates based on AFLP and microsatellite data gave very different results. Based on SSR data, the within-population diversity was strikingly similar for conspecific populations. The average estimated gene diversity (H_E) for *Q. petraea* populations was 0.86 with a coefficient of variance that is only 0.80%. For *Q. robur* the average gene diversity was 0.81 with an even lower coefficient of variation of 0.41%. For this small dataset, the within-population diversity in *Q. petraea* populations was significantly higher than in *Q. robur* populations. The estimates of within-population diversity as calculated from AFLP markers were very similar for both species. Estimates for *Q. petraea* ranged from 0.27 to 0.31 with a mean value of 0.29 and a coefficient of variation of 3.3%. For *Q. robur* populations, estimates of within population diversity also ranged from 0.26 to 0.31 with a mean value of 0.29 and a coefficient of variance of 3.5%. No significant correlation in ranking of populations for both marker systems is present.

It has been shown that estimates of within-population diversity were only seldom correlated when the estimates were made with different marker systems (e.g. Le Corre *et al*. 1997; see Mariette *et al.* 2001 for a cursory review). For outcrossing tree species with high rates of gene flow and large population sizes, within-population diversity may be very similar for different populations, what will result in an almost random ranking of populations for their diversity estimates and depending only on the random variation of diversity observed. However, the higher level of diversity found in *Q. petraea* agrees with life-history traits of this species, in particular a lower selfing rate (Bacilieri *et al.* 1996) and the so-called regeneration of *Q. petraea* from successive unidirectional hybridisation with *Q. robur* (Petit *et al.* 1997). In addition, higher within-population diversity for *Q. petraea* has been shown based on DNA markers (Moreau *et al.* 1994; Bodénès *et al.* 1997a). In a recent study of European oak populations based on AFLP-markers, Mariette *et al.* (2002b) found a slightly higher diversity for *Q. petraea* in comparison *to Q. robur* populations, but this difference was not significant. These results agree quite well with the results reported here. In the study of Mariette *et al.* (2002b), a similar amount of markers was analysed (155 markers) as in the present study (170 markers), but substantially more individuals were sampled per population (mean of 171 individuals per population), what still resulted in a relatively high variance for diversity estimates (C.V. is 6.67 and 6.39 for *Q. petraea* and *Q. robur,* respectively). Also no correlation

between the ranking of populations based on SSR and AFLP markers was found by Mariette *et al*. (2002b).

4.6 Conclusions

Although the landscape has been changed substantially by human impact in Europe, the mapped patterns of chloroplast DNA variation largely reflect the original post-glacial distribution of haplotypes. In Flanders, this resulted in an eastern dominance of the Italian haplotype 1 and the spread of three Iberian haplotypes (10, 11 and 12) over the whole region. Selected provenances contain similar frequencies of haplotypes as autochthonous populations, but are generally more polytypic and often contain two or three haplotypes, whereas autochthonous populations are more often fixed for one haplotype. The applied chloroplast DNA analyses can, to some extent, assist in the identification of the origin of populations and seed lots. Stands not fitting in the general re-colonisation pattern can be classified as having (at least partially) been established with introduced material.

AFLP data classified the oaks into two main groups, according to their taxonomic status. No species-specific AFLP markers were found using 64 primer combinations, but marker frequency differences up to 71% were recorded between both species. AFLP and morphology-based approaches showed a high degree of consistency. The application of six SSR loci could to some extent also differentiate sessile and pedunculate oak populations but more loci should be studied to reach the resolution of AFLP markers. As is expected for outcrossing species and had been previously shown for oaks, only a weak population genetic structure was observed for both oak taxa on a Flemish scale. Within-population diversities were high for autochthonous populations and were not significantly different between the two species. Diversity estimates for the relatively small Flemish forests are comparable to those found for large oak forests in Europe. A small but significant excess of homozygotes has been found in many populations of both species. This agrees with previous observations in other European oak forests and might be due to mating of neighbouring, related individuals or the sympatric presence of partially isolated gene pools of pedunculate and sessile oaks.

No significant genetic differentiation between autochthonous populations and selected provenances was observed for both species. For both taxa, selected stands exhibit slightly lower within-population diversities and have higher heterozygote deficiencies but these differences were not significant. This study revealed that in order to safeguard the genetic variability in future oak forests, there is no reason to choose selected provenances over autochthonous population as seed sources.

5 Genetic diversity of *Carpinus betulus*

5.1 Introduction

5.1.1 Carpinus betulus

Hornbeam (*Carpinus betulus* L.) belongs to the *Corylaceae,* or to the closely related *Betulaceae* family depending on the author (De Langhe *et al*. 1988; Weeda *et al*. 1985). Its natural range includes temperate Europe and Asia Minor. This species cohabits in the south-eastern part of Europe with the related species: *C. orientalis*, a drought-resistant shrub with smaller leaves and fruits, living in more open environments. Hornbeam prefers low lying rich soils and its moisture and warmth demands are high. It is a natural component in the understorey of oak forests. The mature tree is shade tolerant but more light is needed for the germination of the seeds. The tree is monoecious, male and female flowers are grouped in separate catkins. A good seed set requires crosspollination. The spread of seeds and pollen occurs by wind but the dispersal of the fairly heavy seeds may be hampered in the understorey of oaks. On the other hand, the bract is formed to increase potential dispersal distances; distances up to 130m have been observed (Bouman *et al.* 2000). The nuts ripen in September but partially stay on the tree during winter. Nuts are eaten by mammals (e.g. squirells and mice) and birds (e.g. woodpeckers and jays). Seeds are dormant for circa 18 months after ripening (Weeda *et al*. 1985; Bouman *et al.* 2000).

Figure 5.1: Leaves and seeds of *Carpinus betulus* (Photo: E. Evans, NC State University).

The species was of some economic importance in the past owing to its exceptionally hard wood that renders it excellent as firewood and for the production of charcoal and many tools (typical applications were for instance cogwheels and butchers chopping blocks). Despite its excellent wood properties, hornbeam has never been

planted on a large scale for wood production purposes (Grivet & Petit 2002). Because of its ability to withstand repeated coppicing and pollarding however, hornbeam is a usual element in (ornamental) hedges and is cultivated as coppice understorey in middle wood.

5.1.2 Genetic diversity and evolutionary history

Hornbeam colonised Europe very late during the Holocene. Possibly the presence of Neolithic farmers, that opened the European landscape and reactivated the forest dynamics, favoured the colonisation of hornbeam (Ralska-Jasiewiczowa 1964; Küster 1997). These historical processes and some biological and ecological features of the species (delayed reproduction, limited seed dispersal) might explain its late expansion. Recently, the European phylogeography of *C. betulus* and *C. orientalis* was studied using chloroplast-DNA markers (Grivet & Petit 2002). The results of the chloroplast research are illustrated in Figure 5.2. Six haplotypes specific to *C. betulus* were detected, one of them completely fixed in all sampled populations from northern and western Europe, whereas the others were restricted to the eastern European countries or to southern Italy. Two different haplotypes specific to *C. orientalis* were detected, suggesting no ongoing gene flow between both species. A remarkably high value of differentiation was found (G_{ST} = 0.972) among European *C. betulus* populations for these chloroplast-markers. It was the highest differentiation estimate for chloroplast diversity found among the fifteen tree species studied in the framework of a European project (CYTOFOR, Palmé *et al*. 2003). Even in Rumania where several haplotypes occur, only one population out of sixteen was polytypic.

Pollen data (pers. comm. Simon Brewer in Grivet & Petit 2000) are congruent with genetic data: during the last ice age, the hornbeam was restricted to the Balkans, the Iberian Peninsula and to southern Italy. The presence of a single haplotype in western Europe, also found in Croatia, Slovakia and Hungary, suggests that the Iberian populations did not participate in the expansion of the species and became extinct. The divergent Italian populations appear to have remained trapped in this peninsula. Only haplotypes from the refuge in the Balkans participated in the recolonisation in the present interglacial and one haplotype colonised all Western Europe. Similar locations were identified as refuges for both *Carpinus* species in Italy and in the Balkans. The authors further state that conservation priorities should target several populations from these regions that are hotspots for (chloroplast) diversity and that represent a small part of the range of the species. To our knowledge, no other study has investigated the genetic structure of a species in the genus *Carpinus*.

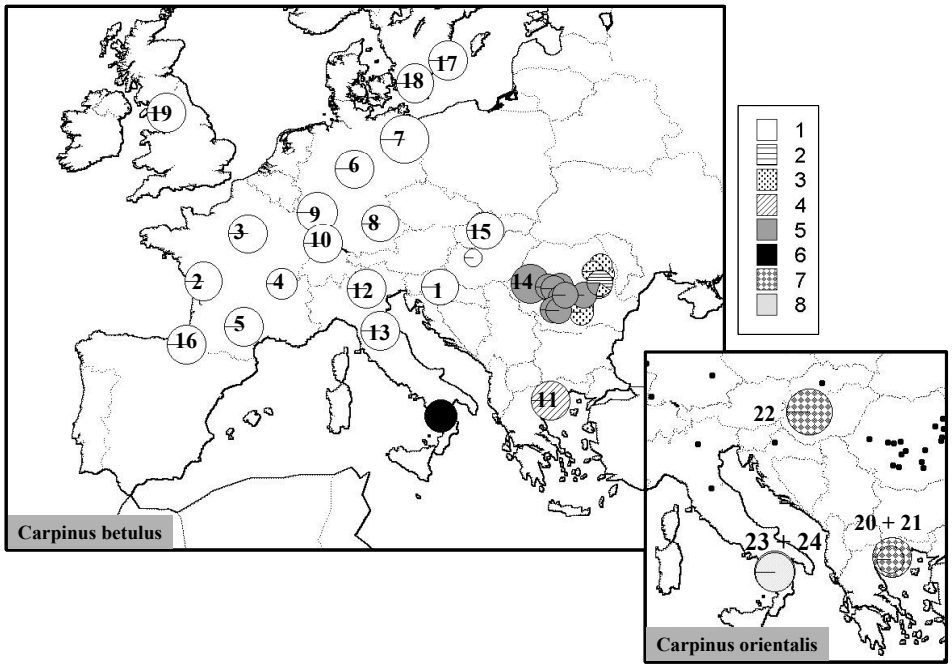


Figure 5.2: Distribution of chloroplast haplotypes in *C. betulus* and *C. orientalis* populations (Grivet & Petit 2002). Symbol sizes are related to the number of genotypes analysed at the location. Numbered populations are analysed in the present study, numbers correspond to Table 5.2.

5.1.3 Occurrence of hornbeam in Flanders

It has become clear from the inventory of autochthonous trees and shrubs in Flanders that natural populations of hornbeam have been conserved in many regions (Maes & Rövekamp 1998, Maes & Rövekamp 2000, Rövekamp & Maes 1999, Rövekamp & Maes 2000, Rövekamp & al. 2000, Opstaele 2001). In accordance with its ecological preferences, it is most abundant in the loam and sandy-loam region. If hornbeam occurs in the sandy or Campine region, both largely characterised by poor sandy soils, it is often on enriched sites such as alluvial soils. Hornbeam is mainly conserved in forests as coppiced understorey in middle wood.

In Flanders, a longstanding tradition of the use of hornbeam in hedges, along roadsides and on the borders between properties exists. It is assumed that hornbeam was grown at nurseries for this purpose from the beginning of the 18th century onwards, where also vegetative propagation was recorded. The further raising success of coppiced hedges might have occurred together with the increasing need for wood and fodder, caused by the growing population and the

increased cultivation of land. For instance, hornbeam leaves were used as fodder for cattle until the beginning of the 20th century. In the Flemish Ardennes, the typical coppiced hedges arose during the second half of the 18th century: rows of low, closely planted, pollarded trees, consisting mainly of ash (*Fraxinus excelsior* L.) and hornbeam (Tack *et al*. 1993). It is assumed that trade in hornbeam reproductive material remained local and hence, most of the old hornbeam individuals are regarded as autochthonous, including the many old line-plantations of pollard and coppiced trees that are still present in the contemporary landscape.

5.2 Occasion of the present research and objectives

This study was started with the general aim of studying the (nuclear) genetic diversity and genetic structure of hornbeam in a European context in order to help devising conservation guidelines for hornbeam in Flanders. This information was especially necessary in view of the new Flemish legislation on the trade in forest reproductive material (see Chapter 1) that has included *C. betulus* in the list of species wherefore certification of reproductive material is compulsory. The recent phylogeographic study of chloroplast haplotypes, conducted in the context of the European project CYTOFOR (see 5.1.2), provided additional information on the evolutionary history of hornbeam. In Western Europe, only one chloroplast haplotype has been detected so far and this raised the question of the impact of the detected bottleneck at the outset of postglacial colonisation on the contemporary genetic diversity of West-European populations. A study of the genetic diversity at nuclear genomic loci would thus provide valuable information.

No information on co-dominant markers for *C. betulus* was available and therefore it was decided to start a survey of genetic diversity with AFLP markers on both European and Flemish hornbeam populations. Since the same haplotype colonised Western Europe, the haplotypes present in Flanders could not be confronted with a European chloroplast structure to evaluate the origin of populations as was done for Flemish oak populations.

5.3 Sampled Carpinus populations

In Flanders, 18 hornbeam locations were sampled. These locations include old coppiced stools and pollarded trees. Most populations were selected from the inventory of autochthonous trees and shrubs in Flanders (Maes & Rövekamp 1998, Rövekamp & Maes 1999) based on a high score for autochthony (score a or b: almost certain or most likely from autochthonous origin). Furthermore, other old populations were sampled for instance in the region ‘Voeren’ and the Flemish Ardennes. All sampled populations are listed in Table 5.1 and their geographical location is shown on Figure 5.3. If possible, leaves were collected from 30 trees per

location, otherwise all trees were sampled. Table 5.1 also provides information on the number of genotypes per population that could be typed successfully with AFLP markers.

Table 5.1: Description of Flemish populations sampled.

Nb	Location	Name	c/p	Lam X	Lam Y	inv.nr	Evaluation*	#	#AFLP
1	As	Kalenhaag	p	236760	191670	99-83	a/b	30	23
2	Bekkevoort	Begijnenbeek	p	194600	181460	98-146	a/b	22	22
3	Bertem	Bertembos	p	168360	175040	98-150	b	22	22
4	Everbeek	Steenbergweide	p	108700	161400	--	--	24	24
5	Heuvelland	Doevebeekdal (Dbd)	p	35750	164440	99-508	a/b	16	16
6	Heuvelland	Rode berg (Rb)	p	36790	164940	99-487	a/b	16	14
7	Heuvelland	Vidaigneberg (Vb)	c	35090	165020	99-477	a/b	10	8
8	Lint	Lachenen	p	159960	200840	98-81	b	16	16
9	Merelbeke	Makkegemsebossen	p	103900	182200	--	--	30	27
10	Michelbeke	Boterhoek	p	108300	170100	99-209	b	30	30
11	Rijkevorsel	Bolk	c	179950	229570	98-110	a/b	16	15
12	Tessenderlo	Achterheide	p	196240	191500	99-182	b/c	9	9
13	Tielt-Winge	Holle weg (Hw)	p	186670	179490	98-123	b	14	14
14	Tielt-Winge	Walembos (Wb)	p	185200	178380	98-131	b	16	16
15	Tongeren	Kolmont	p	225000	166000	--	--	18	15
16	Voeren	Vossenaerde (V)	p	255500	159400	--	--	30	18
17	Voeren	Loods (L)	c	251600	159000	--	--	30	21
18	Voeren	St-Martens-Voeren (SMV)	p	253400	159400	--	--	30	30
		Total						**379**	**340**

Nb: Number on Figure 5.1; c/p: coppice stools (c) or pollarded trees (p); Lam X and Lam Y: coordinates in Lambert X and Lambert Y format; inv.nr: inventory number according to Maes & Rövekamp (1998) and Rövekamp & Maes (1999); *: evaluation of autochthonous character (degree of likelihood) following Maes and Rovekamp (1998): c: likely autochthonous; b: more likely autochthonous, a: most likely autochthonous; #: number of trees sampled; # AFLP: number of individuals that could be typed at AFLP loci.

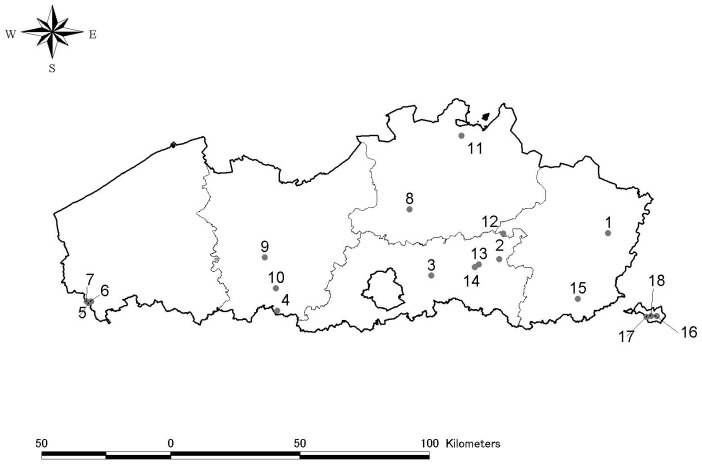


Figure 5.3: Geographical presentation of the Flemish populations sampled. Numbers refer to Table 5.1.

Furthermore, material from 19 European *Carpinus betulus* and 5 *C. orientalis* populations was kindly provided by Dr. R. Petit & Dr. D. Grivet (INRA, Bordeaux, France). These European samples were formerly analysed for their chloroplast haplotype (Grivet & Petit 2002), as shown in the last column of Table 5.2. All *C. betulus* populations, except the Greek and the Rumanian, were shown to be monotypic for the same haplotype (type 1); the Rumanian is monotypic for haplotype 5, the Greek for haplotype 4. The Greek and Hungarian *C. orientalis* populations were shown to be monotypic for haplotype 7, the Italian populations for haplotype 8. All European *Carpinus* populations analysed are listed in Table 5.2 and their geographic location is shown on Figure 5.2. Note that in the Greek location Arnea both species were sampled. Unfortunately, no dried plant material was available for AFLP analysis from the eastern European and southern Italian populations that were shown to contain divergent haplotypes.

The Flemish samples were collected in 1999 and 2000 and AFLPs were performed in 2001; DNA-extractions and AFLP analysis of the European samples were carried out in 2002. For 21 samples of the Flemish population sampled in Voeren (number 17 in Table 5.1), AFLP analysis was repeated together with the European samples in 2002, starting from new DNA extractions of dried leaf material. The comparison of the fingerprints obtained for these samples in 2001 and in 2002 allowed us to check for discrepancies between the two analysis-periods. All AFLP fingerprints were processed jointly in 2002. AFLP markers were defined on a subset of fingerprints that contained fingerprints generated in 2001 and 2002 and genotypes from both *Carpinus* species.

Table 5.2: Description of European populations analysed.

Nb	Country	Location	Long	Lat	# sampled	# AFLP	*Cp** haplotype
***C. betulus* samples**							
1	Croatia	Mt Medvenica	15,95	45,87	9	9	1
2	France	Chizé	-0,40	46,14	10	8	1
3	France	Fontainebleau	2,67	48,42	10	10	1
4	France	Seillon	5,00	46,00	7	5	1
5	France	Senerac	2,36	43,98	10	8	1
6	Germany	Bovenden	10,05	51,57	10	10	1
7	Germany	Grumsiner	13,50	50,00	15	13	1
8	Germany	Kelheim	11,83	48,93	10	10	1
9	Germany	Lemberg	7,46	49,48	11	9	1
10	Germany	Schönsberg	7,83	47,96	10	10	1
11	Greece	Arnea	23,60	40,47	11	11	4
12	Italy	Bresciano	10,88	45,80	10	8	1
13	Italy	Casentinesi	11,80	43,78	10	9	1
14	Rumania	Savarsin	22,23	46,02	11	11	5
15	Slovakia	Boki	19,12	48,57	10	9	1
16	Spain	San Juan Xar	-1,63	43,13	11	10	1
17	Sweeden	Halltorp Hage	16,53	56,75	11	11	1
18	Sweeden	Stenshuvud	14,25	55,65	10	10	1
19	UK	Lake district	-3,00	54,27	10	10	1
				Total:	**196**	**181**	
***C. orientalis* samples**							
20	Greece	Arnea	23,60	40,47	5	5	7
21	Greece	Olymbiada	23,74	40,59	10	8	7
22	Hongary	Csakvar	18,45	47,40	12	9	7
23	Italy	Laino Castello	15,98	39,94	10	10	8
24	Italy	Mormanno	15,99	39,88	10	9	8
				Total:	**47**	**41**	

Nb: number referring to Figure 5.2; # sampled: number of trees sampled; # AFLP: number of individuals that could be typed at AFLP loci; *: from Grivet & Petit (2002).

5.4 Results

5.4.1 Choice of AFLP primer combinations

In total, 36 AFLP primer combinations were tested on 5 samples from different Flemish populations (data not shown). The generated fingerprints were evaluated for overall clearness of the banding pattern and for number of polymorphic markers generated. The three primer combinations finally selected for the diversity study are *Eco*RI-AAC + *Mse*I-CAC, *Eco*RI-ACA + *Mse*I-CAG and *Eco*RI-ACG + *Mse*I-CAC.

A reproducibility test was performed for these primer combinations by repeating the complete AFLP analysis for 5 samples, starting from different DNA extractions. The overall reproducibility was high for the analyses carried out both in 2001 and in 2002: on average 96.9% of the markers were scored identically in the three repeats of each of the five Flemish individuals tested in 2001, on average 96.6% of the markers were scored identically in the two repeats of each of five European individuals tested in 2002.

5.4.2 Reproducibility of AFLP patterns over two years

It was investigated whether AFLP patterns were also reproducible over both years of creation of the dataset. Therefore, the samples from one of the Flemish populations already analysed in 2001 were analysed again in 2002 starting from new DNA-extractions, together with the European samples. In total, AFLP patterns could be compared for repeats of 21 genotypes. Although reactions were performed in a standardised manner and AFLP markers to score were defined on the dataset with both Flemish and European samples, the effect of the year of analysis was shown to cause substantial bias in the dataset. The average reproducibility between repeats produced in different years was 84.8, which is rather low in comparison to the levels of reproducibility obtained between repeats included in the same experiment (96.9 and 96.6 for 2001 and 2002 respectively). Table 5.3 shows the number of markers for which discrepancies between repeats of 2001 and 2002 were observed. Fifty-seven out of 118 markers were scored identically in all the repeats, but more than the half of the markers (61 out of 118) were scored differently in at least one comparison. Strikingly, one marker was scored differently in as much as 17 out of 21 repeats compared. Further inspection of the dataset showed that the deviant markers were spread over the three primer combinations, were both small and large fragments and contained markers with low and high average peak heights.

Table 5.3: Number of comparisons (of 21 repeats) for which markers were scored differently in 2001 and 2002.

# scored differently	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
# markers	**57**	**14**	**11**	**4**	**3**	**6**	**5**	**5**	**2**	**5**	0	0	0	0	0	**2**	**1**	**1**	0	0	0	0

The effect of this lower agreement between fingerprints produced over the two years of laboratory analysis in further data examination is illustrated in the PCO-plot shown in Figure 5.4. In this plot derived from AFLP-marker frequency data, all Flemish populations group very closely together and the Flemish population that was reanalysed with the European samples during 2002 came out very differently from its former analysis (population labelled B* on Figure 5.4). This effect was also apparent in the results of other data analyses such as the model-based clustering method of Pritchard *et al*. (2000) where the dataset of 2001 made up a clearly distinct ‘gene pool’ (results not shown). Hence, the fingerprints from 2001 and 2002 cannot be explored jointly but instead should be treated as separate datasets. These will be further referred to as Flemish and European datasets.

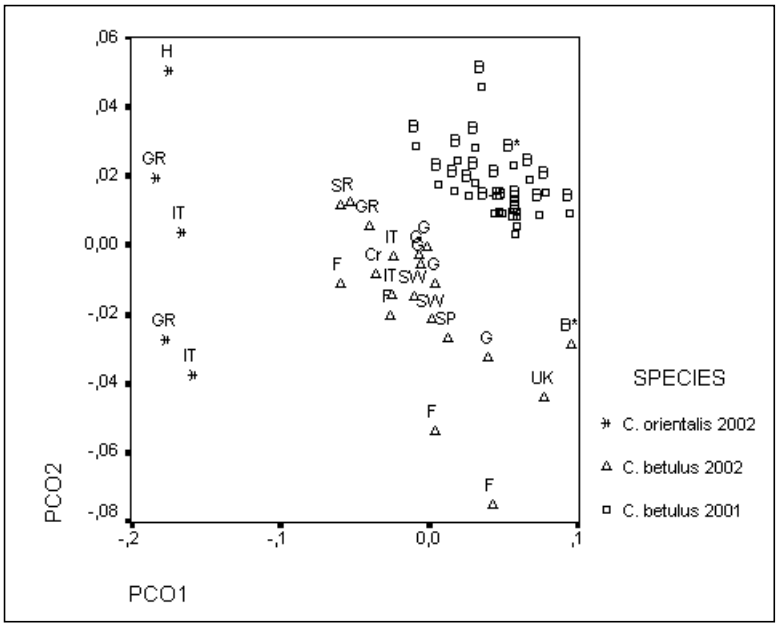


Figure 5.4: Plot of first two principal co-ordinates calculated from Nei's genetic distance between populations (Nei 1987). Populations are labelled according to their country of origin: B(elgium), C(roatia), F(rance), G(ermany), Gr(eece), H(ungary), IT(aly), R(umania), S(lovakia), SP(ain), Sw(eden) and U(nited) K(ingdom). The population B* is the same population analysed in 2001 and 2002.

5.4.3 Allelic variation at AFLP loci

The use of three AFLP primer combinations resulted in 118 scoreable markers, of which 114 and 117 (or 96.6 and 99.2%) were polymorphic in the Flemish and European dataset respectively (with least common state minimum 5%). Each tree was characterised by a unique banding pattern. A negative correlation was detected between fragment sizes and frequencies in both datasets (r = -0.1371 and r = -0.2852 for the Flemish and European dataset respectively) but this correlation was only significant for the European dataset (p values respectively 0.14 and 0.002). Data analysis was repeated for both datasets with only AFLP fragments larger than 150 bp. For this reduced data set (59 markers), the correlation was also negative in both cases (r = -0.2613 and r = -0.1465) but this time only significant for the Flemish data (p values respectively 0.050 and 0.27). This procedure resulted in similar differentiation estimates for both data sets (results not shown), suggesting that the potential presence of size homoplasy of AFLP fragments does not result in underestimation of genetic differentiation between samples.

5.4.4 Relationships among individuals

A PCO analysis was performed for both datasets. No structure according to population of origin or geographic location were apparent on the plot of Flemish samples (results not shown). On the plot of individuals from the European data (Figure 5.5), *C. orientalis* genotypes are separated from *C. betulus* genotypes, with some degree of overlap. The small group of 16 *C. betulus* genotypes that formed a separated cluster (on the bottom of figure 5.5), is composed of some of the Swedish and some of the Rumanian individuals.

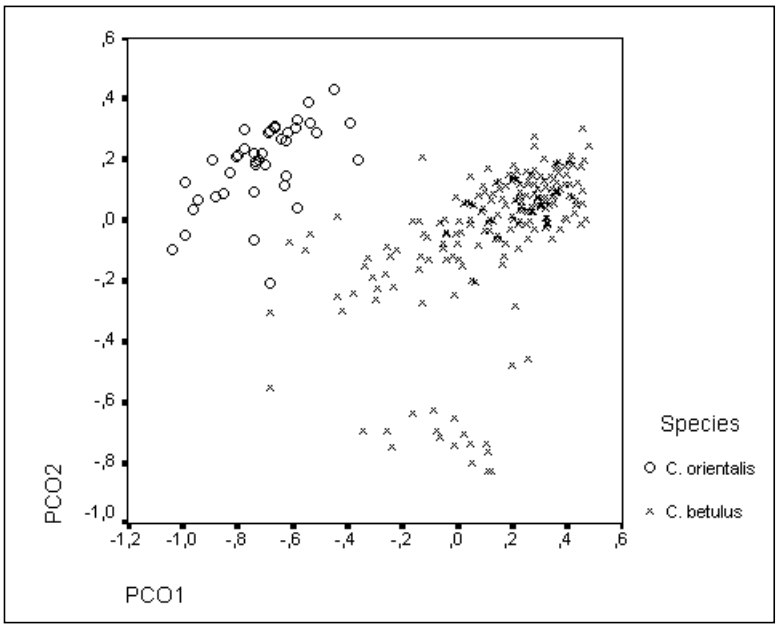


Figure 5.5: Plot of first two principal co-ordinates calculated from the relatedness coefficient r between genotypes (Lynch & Milligan 1994) of 118 AFLP markers for the European dataset.

Additionally, the model-based clustering method of Pritchard (2000) was applied on both data sets. For the **Flemish data set**, the highest estimate of the likelihood of the data, conditional to a given number of clusters, was obtained when clustering all genotypes into five gene pools. No explanation can be given for the division of individuals into these five gene pools. For most populations individuals were assigned to several gene pools and no relation with their geographic location was observed (Table 5.4).

For the **European data set** the highest estimate of the likelihood of the data, conditional to a given number of clusters, was obtained when clustering all genotypes into seven gene pools (Table 5.5). The analysis was repeated with only

C. betulus genotypes and in this case, the highest estimate of the likelihood of the data was obtained when clustering all genotypes into six gene pools, that are almost identical to the six first gene pools described below. In contrast to the Flemish situation, at European scale, the geographic location of the populations can account to some extent for the inference of six gene pools:

- **Gene pool 1** comprises mainly genotypes from the Belgian population (Voeren, number 17) and from the German population Lemberg (nr. 9 on Figure 5.2), located close to the Belgian border;
- **Gene pool 2** consists largely of southern and south-eastern European genotypes (from Italy, Slovakia, Rumania, Greece) and some French genotypes;
- **Gene pool 3** contains a few genotypes from many different regions and half of the Hungarian *C. orientalis* population;
- **Gene pool 4** comprises many German hornbeams, almost all British genotypes and further trees from Sweden, Spain, Italy, Croatia, France and Belgium;
- **Gene pool 5** is made up of French trees and a part of the Spanish population;
- **Gene pool 6** contains many Swedish genotypes and half of the Rumanian individuals. This gene pool corresponds to the group of individuals that are separated from the main *C. betulus* cluster in Figure 5.5;
- **Gene pool 7** contains the vast majority of *C. orientalis* trees.

The Greek trees of *C. betulus* and *C. orientalis* which were collected at the same geographical location (Arnea, number 11 and 20 respectively in Table 5.5) are assigned to different gene pools, indicating a clear genetic separation between the two species, even in locations where they grow sympatric. Table 5.6 gives Nei's genetic distance between inferred gene pools for the European dataset. It is worth noting that the smallest distance in the data set was observed between gene pools 1 and 4 (northern most gene pools), the largest distances are between *C. betulus* and *C. orientalis* gene pools.

Table 5.4: Summary of the results obtained using the assignment procedure based on AFLP data of the Flemish populations. Figures are the proportion of estimated membership to each of five inferred gene pools for genotypes of a given population of origin. Numbers of populations refer to Table 5.1.

			Inferred genepools				
Nb	**Population**	**Province**	**1**	**2**	**3**	**4**	**5**
1	Kalenhaag	Limburg	**0.472**	0.234	0.076	0.13	0.087
2	bekkevoort	Limburg	0.301	**0.559**	0.139		
3	Bertembos	VL Brabant	0.147	**0.751**	0.102		
4	Everbeek	O Vlaanderen	**0.426**	0.209	0.206		0.159
5	Doevebeekdal	W Vlaanderen	**0.332**	0.116	0.177	0.188	0.188
6	Rode berg	W Vlaanderen	**0.409**	0.296	0.08		0.214
7	Vidaigneberg	W Vlaanderen	0.083	**0.766**	0.151		
8	Lint	Antwerpen	0.167	0.324	**0.509**		
9	Makkegem	O Vlaanderen	0.167	0.026	0.063		**0.744**
10	Boterhoek	O Vlaanderen	0.243	**0.295**	0.254	0.032	0.175
11	Bolk	Antwerpen	0.147	0.1	0.176		**0.576**
12	Achterheide	Limburg	0.357	**0.37**	0.262		0.011
13	Tielt-Winge123	VL Brabant	0.087	0.382	**0.531**		
14	Walenbos131	O Vlaanderen	0.149	0.314	**0.537**		
15	Kolmont	Limburg	0.185	0.053	0.028		**0.733**
16	Vossenaerde	Limburg	0.254	0.231	0.146		**0.369**
17	Loods	Limburg	0.222	0.312	**0.336**		0.131
18	SMV	Limburg	**0.421**	0.132	0.221		0.226

Table 5.5: Summary of the results obtained using the assignment procedure based on AFLP data of the European populations. Figures are the proportion of estimated membership to each of seven inferred gene pools for genotypes of a given population of origin. Numbers of populations refer to Table 5.2. *:The Belgian population is number 17 in Table 5.1

		Inferred gene pools						
Nb	**Population**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
17*	B_Loods2002	**0.705**			0.295			
1	CR_MtMedvenica	0.023	**0.444**		0.421		0.111	
2	FR_Chizé				0.319	**0.68**		
3	FR_FBL				0.090	**0.910**		
4	FR_Seillon		0.200	0.200	0,000	**0.600**		
5	FR_Serenac		0.359		0.032	**0.609**		
6	G_Bovenden				**0.997**			
7	G_Grumsiner	0.118	0.091	0.077	**0.713**			
8	G_Kelheim	0.108	0.100	0.100	**0.66**	0.032		
9	G_Lemberg	**0.667**	0.111		0.217			
10	G_Schönberg	0.117	0.328		**0.524**	0.031		
11	GR_Arnea	0.089	**0.609**		0.281	0.021		
12	IT_Bresciano	0.103	0.219	0.125	**0.491**	0.063		
13	IT_Casentinesi	0.041	**0.322**	0.111	0.301	0.225		
14	R_Savarsin		0.274		0.157	0.023	**0.545**	
15	SL_Boki		**0.892**		0.108			
16	SP_SanJuar	0.007	0.097		**0.456**	0.440		
17	SW_Halltorps	0.077			**0.428**	0.200	0.295	
18	SW_Stenshuvud	0.100	0.100		0.100		**0.700**	
19	UK_Lakedistrict				**0.999**			
20	GR_CO_Arnea							**0.999**
21	GR_CO_Olymbiada							**0.999**
22	H_CO_Csakvar			0.456				**0.544**
23	IT_CO_Laino							**1,000**
24	IT_CO_Mormanno			0.128				**0.872**

Table 5.6: Nei's genetic distance between seven inferred gene pools in the European data.

	1	2	3	4	5	6	7
1							
2	0.0906						
3	0.257	0.0815					
4	0.0263	0.0858	0.2609				
5	0.0625	0.0632	0.1905	0.0377			
6	0.1292	0.0708	0.1822	0.1154	0.1197		
7	0.2156	0.0878	0.0757	0.2055	0.1627	0.1752	

5.4.5 Relationships among populations

Results of neighbour-joining trees based on Nei's genetic distance are shown in Figures 5.4 and 5.5. Again, no geographic structure can be observed in the clustering of Flemish populations. Furthermore, the bootstrap support values of forks are mostly low (6 out of 16 values lower than 75%). In the tree calculated from European data, the *C. orientalis* populations are grouped together with a high bootstrap support value (100%). The clustering of conspecific populations on the other hand, has again lower bootstrap support (16 of 22 values lower than 75%). Nevertheless, some structuring according to the geographic location of the populations can be observed. All German populations, the Belgian and the British population are grouped. Together with both Swedish populations they make up a cluster with a bootstrap support of 67.5%. Furthermore, French populations and the Spanish population cluster together with low bootstrap support (49.5%). Close to the *C. orientalis* group, southern and eastern European populations are loosely grouped with bootstrap support values from 34.9 to 65.8%. The highest bootstrap support values are mostly found for clusters of populations of the same country of origin: 97.1% for two French populations, 93.6% for the Swedish populations and 85.9% for both Greek *C. orientalis* samples.

The divergence between populations was tested on both Flemish and European scale. Because no prior information on the within-population level of inbreeding was available, analyses were repeated with different input values for the inbreeding coefficient F_{is} in accordance with the presumed outbreeding nature of the species (F_{is} = 0.0 and F_{is} = 0.1). For both input values the differentiation among populations was highly significant for both data sets (see Table 5.7) and for both input values of F_{is} differentiation was higher on a European than on a Flemish geographical scale. For all further calculations where an input value of F_{is} was required, F_{is} was set to 0.0.

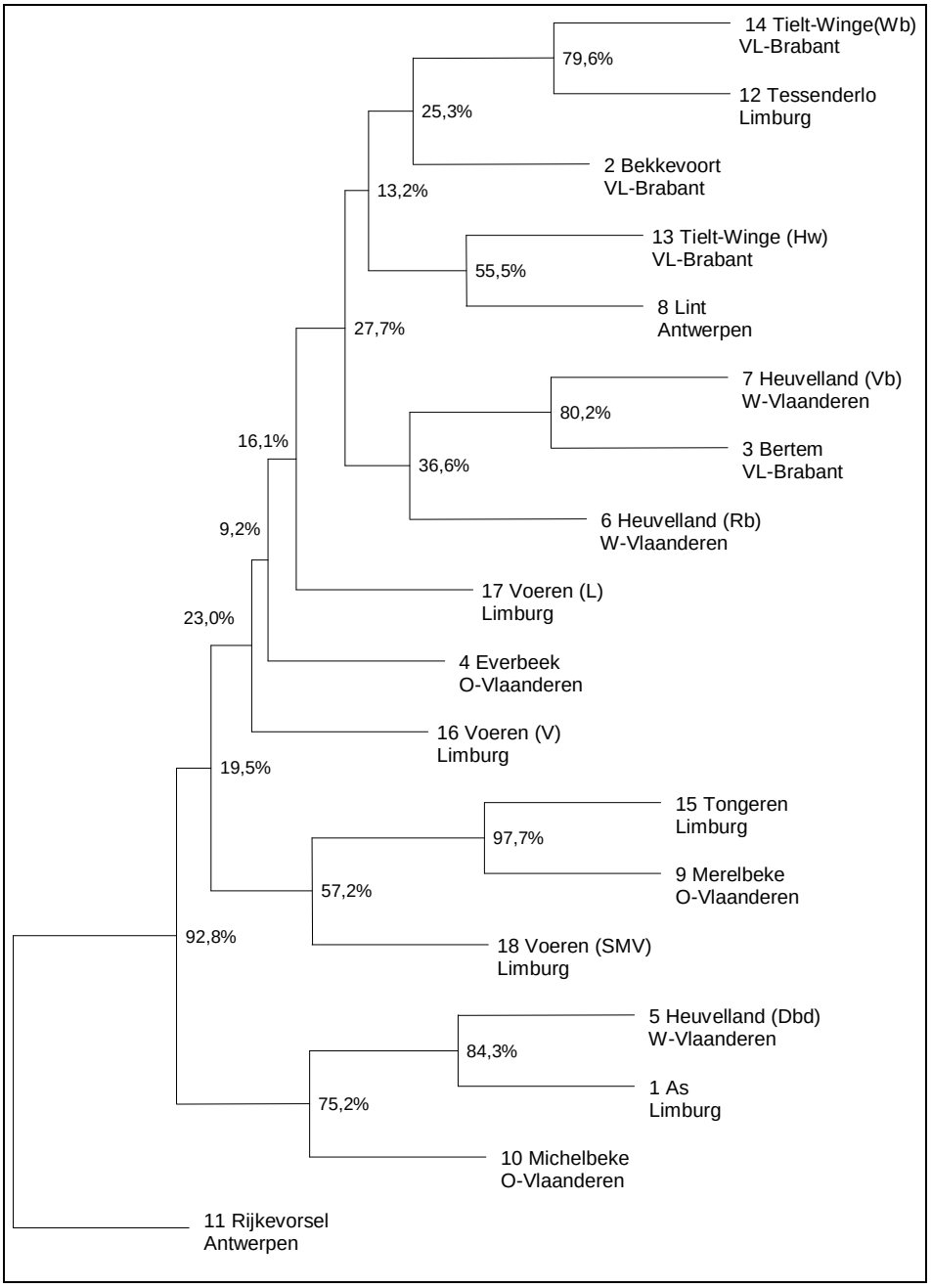


Figure 5.4: Neighbour-joining tree with bootstrap support values (based on 1000 bootstraps) at forks, calculated from Flemish AFLP data and Nei's genetic distance. The province of origin is stated beneath each population. Numbers preceding each population name refer to Table 5.1.

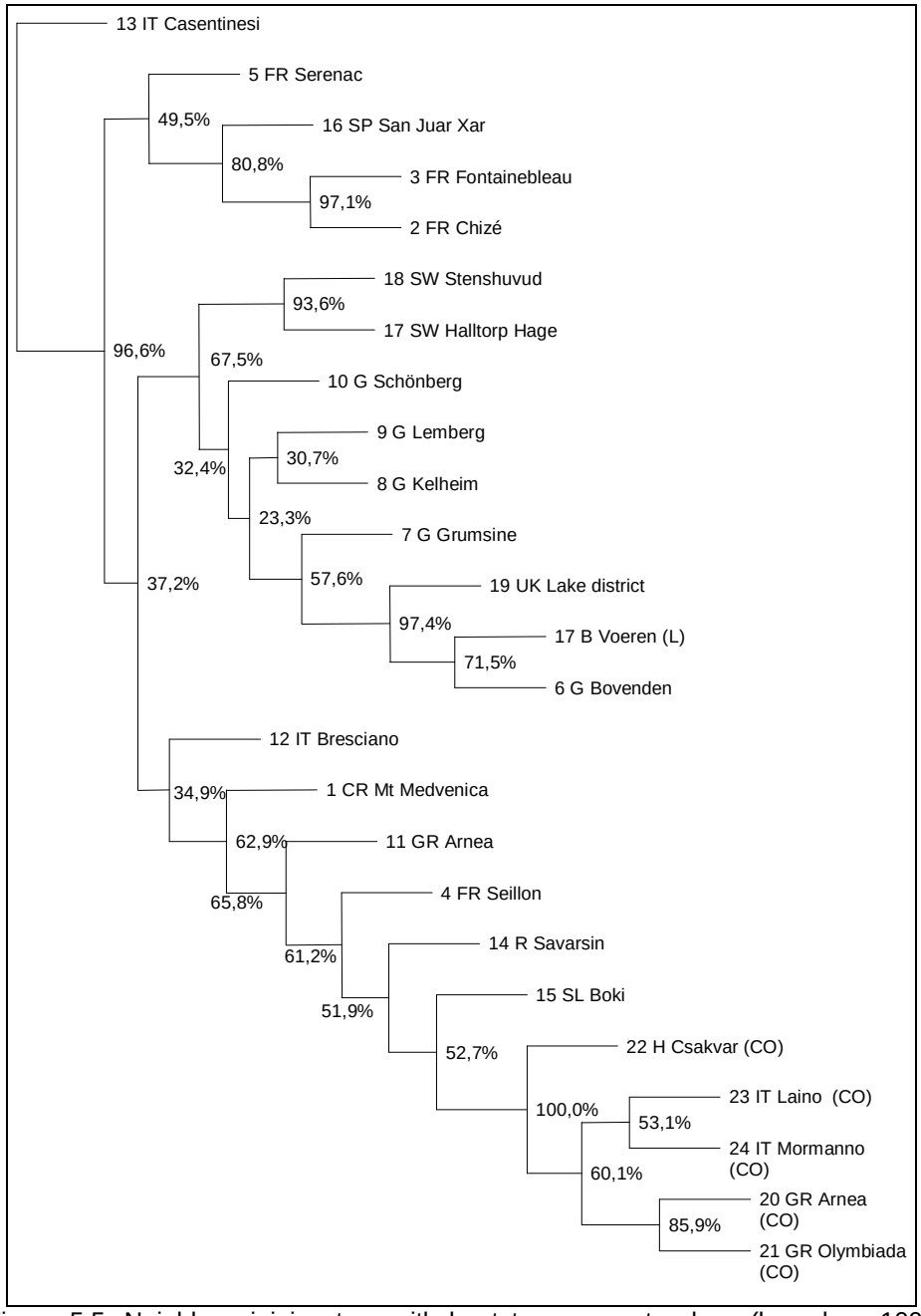


Figure 5.5: Neighbour-joining tree with bootstrap support values (based on 1000 bootstraps) at forks, calculated from European AFLP data and Nei's genetic distance. CO: *C. orientalis*., all other populations *C. betulus*. Each population is preceded with its number (see Table 5.2 (European pop.) and 5.1 (Belgian pop.)) and the abbreviation of the country of origin; see legend of Figure 5.4 for abbreviations used.

Table 5.7: Analysis of population genetic structure based on 118 AFLP markers.

	F_{is} = 0.0					F_{is} = 0.1				
n	H_t	H_w	H_b	F_{st}	**P-value**	H_t	H_w	H_b	F_{st}	**P-value**
Among Flemish *C. betulus* populations										
18	0.3125	0.2979	0.0146	0.0466	<0.0001	0.3044	0.2893	0.0151	0.0496	<0.0001
S.E.		0.0057	0.0017	0.1137			0.0058	0.0017	0.1111	
Var		0.0000	0.0000	0.0129			0.0000	0.0000	0.0123	
Among European *C. betulus* populations										
20	0.3598	0.3332	0.0266	0.074	<0.0001	0.3534	0.324	0.0294	0.0832	<0.0001
S.E.		0.0074	0.0048	0.1768			0.0074	0.0050	0.1658	
Var		0.0000	0.0000	0.0313			0.0000	0.0000	0.0275	
Among *C.betulus* and *C.orientalis*										
2	0.3667	0.2817	0.085	0.2266	<0.0001	0.3693	0.2732	0.0961	0.2544	<0.0001
S.E.		0.0558	0.0000	0.1487		0.0562	0.0000	0.1488		
Var		0.0031	0.0000	0.0221		0.0031	0.0000	0.0221		

N: number of samples; H_t: total diversity; H_w: average diversity within populations; H_b average diversity between populations; F_{st}: differentiation between populations.

Finally, the correlation between pair-wise genetic divergence and geographic distance was further investigated. Therefore, the correlation between the divergence statistic $F_{ST}/(1-F_{ST})$ and the logarithm of the geographical distance was calculated and the significance was tested using a Mantel test. F_{ST} values were calculated using an input value of 0.0 for the level of inbreeding. As expected, no correlation was found between geographical distances and genetic divergence at the Flemish scale (r = -0.0354, P-value = 0.695). For the European *C. betulus* data, a highly significant positive correlation between pair-wise genetic divergence and geographic distance was found: r = 0.395 with P-value = 0.0006.

5.4.6 Within-population diversity

Diversity statistics are summarised in Table 5.8. High variation was recorded at AFLP loci for both data sets, with on average 84 and 87% of polymorphic loci within Flemish and European *C. betulus* samples. On average, only 63% of loci were polymorphic in the *C. orientalis* populations. For Flemish populations, more individuals could be typed and this resulted in a smaller variance component of gene diversity attributed to sampling of individuals (26.4% versus 43.5% for the European populations). The overall variance of H_j, however, is equal for both data sets (Var (H_j) = 0.0002) and did not benefit from the larger amount of individuals typed. The coefficient of variance is larger for the Flemish estimates of H_j (5.2%) than for the European estimate (4.4%). The Flemish populations displayed on average lower gene diversity levels than the European populations and this difference was significant (t = -4.102, p-value <0.001).

The correlation between diversity estimates and geographical variables was calculated for the European *C. betulus* dataset (Table 5.9) in order to determine the

putative influence of the postglacial migration process on the genetic diversity. The coordinates of the Balkan refugium were chosen in the indicated refuge area near the Black sea (on unpublished pollen maps of *C. betulus*, pers. comm. S. Brewer) and set at 28,00 longitude and 45,00 latitude. The distance to this Balkan refugium was calculated for all hornbeam populations. Surprisingly, the correlation between the geographical distances to the refuge and diversity estimates was very high and both are highly significant (Table 5.9). The correlation between latitude and diversity estimates was not significant. The correlation between longitude and PLP was negative and significant (r = -0.372, P = 0.048) whereas the correlation between longitude and H_j was also negative but not significant (r = -0.321, P = 0.078). These correlations were also calculated for the Flemish data set but, as expected, no significant correlations were revealed (results not shown).

Table 5.9: Correlation between geographical variables and diversity estimates

	Log Distance	**Latitude**	**Longitude**
PLP	0,748***	0,258	-0,372*
H_j	0,693***	0,149	-0,321

Log Distance: logarithm of the distance to the refugium; *PLP*: proportion of polymorphic loci; H_j: average gene diversity; *,***: significant correlation coefficients at the 0.05 and 0.001 level respectively.

Table 5.8: Diversity statistics within *Carpinus* populations calculated from 118 AFLP markers.

NB	Location/ Country	Population	N	NPL	PLP	Hj	S.E.(Hj)	Var(Hj)	VarI%	VarL%
Flemish *C. betulus* populations										
1	As	Kalenhaag	23	100	84.7	0.3431	0.0158	0.0003	12.2	87.8
2	Bekkevoort	Begijnenbeek	22	98	83.1	0.2675	0.0157	0.0002	23.6	76.4
3	Bertem	Bertembos	22	98	83.1	0.2736	0.0159	0.0003	22.9	77.1
4	Everbeek	Steenbergweide	26	103	87.3	0.2880	0.0154	0.0002	8.3	91.7
5	Heuvelland	Doevebeekdal	16	98	83.1	0.3382	0.0156	0.0002	18.3	81.7
6	Heuvelland	Rode berg	14	104	88.1	0.3036	0.0152	0.0002	35.4	64.6
7	Heuvelland	Vidaigneberg	8	95	80.5	0.2840	0.0159	0.0003	51.5	48.5
8	Lint	Lachenen	16	94	79.7	0.2802	0.0161	0.0003	28.1	71.9
9	Merelbeke	Makkegem	27	95	80.5	0.2937	0.0161	0.0003	15.8	84.2
10	Michelbeke	Boterhoek	30	103	87.3	0.3253	0.0144	0.0002	15.5	84.5
11	Rijkevorsel	Bolk	15	99	83.9	0.3172	0.0152	0.0002	29.9	70.1
12	Tessenderlo	Achterheide	9	97	82.2	0.2769	0.0157	0.0002	52.0	48.0
13	Tielt-Winge	Holle weg	14	98	83.1	0.3004	0.0165	0.0003	27.2	72.8
14	Tielt-Winge	Walembos	16	94	79.7	0.2699	0.0156	0.0002	31.9	68.1
15	Tongeren	Kolmont	15	105	89	0.3320	0.0151	0.0002	26.8	73.2
16	Voeren	Vossenaerde	18	101	85.6	0.3057	0.0141	0.0002	34.5	65.5
17*	Voeren	Loods	21	101	85.6	0.2746	0.0153	0.0002	23.9	76.1
18	Voeren	SMV	30	101	85.6	0.2887	0.0150	0.0002	17.9	82.1
		Mean	**18**	**99.11**	**84.01**	**0.2979**	**0.0155**	**0.0002**	**26.4**	**73.6**
		CV		**0.034**	**0.034**	**0.052**				
Euopean *C. betulus* populations										
17*	Belgium	Voeren	21	103	87.3	0.2907	0.0151	0.0002	24.3	75.7
1	Croatia	Mt Medvenica	9	108	91.5	0.3692	0.0139	0.0002	39.0	61.0
2	France	Chizé	8	101	85.6	0.2993	0.0151	0.0002	60.8	39.2
3	France	Fontainebleau	10	89	75.4	0.2423	0.0162	0.0003	39.3	60.7
4	France	Seillon	5	87	73.7	0.3305	0.0166	0.0003	41.8	58.2
5	France	Serenac	8	95	80.5	0.3273	0.0163	0.0003	33.9	66.1
6	Germany	Bovenden	10	106	89.8	0.3162	0.0138	0.0002	62.9	37.1
7	Germany	Grumsiner	13	110	93.2	0.3549	0.0135	0.0002	37.3	62.7
8	Germany	Kelheim	10	104	88.1	0.3663	0.0144	0.0002	37.3	62.7
9	Germany	Lemberg	9	105	89	0.3660	0.0141	0.0002	45.8	54.2
10	Germany	Schönberg	10	109	92.4	0.3429	0.0139	0.0002	49.1	50.9
11	Greece	Arnea	11	101	85.6	0.3301	0.0152	0.0002	28.9	71.1
12	Italy	Bresciano	8	104	88.1	0.3726	0.0136	0.0002	46.4	53.6
13	Italy	Casentinesi	9	107	90.7	0.3770	0.0130	0.0002	46.5	53.5
14	Rumania	Savarsin	11	100	84.7	0.3197	0.0150	0.0002	31.7	68.3
15	Slovakia	Boki	9	102	86.4	0.3056	0.0149	0.0002	39.1	60.9
16	Spain	San Juar Xar	10	101	85.6	0.3241	0.0146	0.0002	46.6	53.4
17	Sweden	Halltorps Hage	11	107	90.7	0.3363	0.0141	0.0002	47.0	53.0
18	Sweden	Stenshuvud	10	111	94.1	0.3635	0.0127	0.0002	56.1	43.9
19	U. Kingdom	Lake district	10	103	87.3	0.3299	0.0142	0.0002	56.2	43.8
		Mean	**20**	**102.65**	**86.99**	**0.3332**	**0.0145**	**0.0002**	**43.5**	**56.5**
		CV		**0.061**	**0.062**	**0.044**				
European *C. orientalis* populations										
20	Greece	Arnea	5	69	58.5	0.2277	0.0171	0.0003	46.6	53.4
21	Greece	Olymbiada	8	74	62.7	0.1969	0.0164	0.0003	31.4	68.6
23	Hungary	Csakvar	9	70	59.3	0.2273	0.0183	0.0003	17.1	82.9
24	Italy	Laino	10	79	66.9	0.2006	0.0155	0.0002	37.8	62.2
25	Italy	Mormanno	9	77	65.3	0.2564	0.0181	0.0003	19.3	80.7
		Mean	**5**	**73.8**	**62.5**	**0.2218**	**0.0171**	**0.0003**	**30.4**	**69.6**
		CV		**0.059**	**0.059**	**0.077**				

NB: number in Table 5.1 and 5.2; N: number of individuals typed; *NPL*: number of polymorphic loci at the 5% level; *PLP*: proportion of polymorphic loci; H_j: average gene diversity; VarI%: proportion of variance of H_j due to sampling of individuals; VarL%: proportion of variance of H_j due to sampling of loci; *: Flemish population analysed in 2001 and 2002.

5.5 Discussion

5.5.1 *Use of AFLP for population genetic study of* Carpinus

A remarkable lack of reproducibility of AFLP patterns generated over different years was detected (only 84.8% of markers scored identically in repeats of 2001 and 2002), although AFLP procedures were kept as constant as possible. An important factor for the low reproducibility could be the use of another PCR-machine to carry out the pre-amplifications. The PCR-machine that was used in 2001 (Hybaid Omni Gene Cycler) broke down and pre-amplifications were carried out in another type of PCR-machine in 2002 (Perkin Elmer Geneamp PCR System 9600). Selective amplifications were always carried out in the latter PCR system. However, it cannot be determined at this point whether the effect was (partially) caused by other factors such as other fabrication batches of products used or the storage of leaf material. Nonetheless, it is clear that caution is required when constructing databases of AFLP markers over a longer period of time and reproducibility should be monitored carefully.

As for the AFLP data sets for apple and oak (Chapters 3 and 4), allelic frequencies where calculated assuming a non-uniform prior distribution (Zhyvotovski 1999) that requires the input of the within-population inbreeding level. As far as we know, no population genetic study of *Carpinus* sp. has been performed with nuclear markers and therefore, we could only rely on the putative outcrossing breeding system of the species to propose input values for F_{is}. Calculations were carried out assuming Hardy-Weinberg equilibrium (F_{is} = 0.0) and assuming a low level of within-population inbreeding (F_{is} = 0.1). In both situations, among-population genetic differentiation was significant in both data sets with a higher divergence among European than among Flemish populations.

5.5.2 *Relationship between* C. betulus *and* C. orientalis

Similar refugia in Italy and the Balkans have been identified for *C. betulus* and *C. orientalis* (Grivet & Petit 2002) and both species have largely overlapping ranges. Nonetheless, two different haplotypes specific to *C. orientalis* were detected, suggesting no ongoing gene flow between these species. One haplotype detected in *C.betulus* however (type 5, found in the Rumanian population studied in the present research), was found to be more related to the *C. orientalis* haplotypes than other haplotypes occurring in *C. betulus* populations (Grivet & Petit 2002).

Although the present AFLP-data set includes only 41 fingerprints of *C. orientalis* individuals and the study of divergence between the two *Carpinus* species was not a main objective, some information on the relationship between both species can be derived. First of all, AFLP-data clustered the genotypes largely according to their

taxonomic status. This could be seen on the PCO plot as well as in the model-based assignment of individuals to inferred gene pools, where *C. orientalis* makes up a distinct gene pool (number 7 in Table 5.5). No evidence is present at the nuclear genome level that the Rumanian trees with haplotype 5 are more closely related to *C. orientalis* than the other hornbeams with haplotype 1. This would suggest that gene flow from conspecific *C. betulus* populations and no (or only limited) gene flow from populations of C. *orientalis* has altered the putative former nuclear relatedness between *C. betulus* populations with haplotype 5 and *C. orientalis* populations. Furthermore, genetic differentiation between both species was much larger (0.23 for input F_{is} = 0.0) than differentiation between *C. betulus* populations on a Flemish and a European scale (respectively 0.05 and 0.07 for input F_{is} = 0.0). On the neighbour-joining tree, all *C. orientalis* populations were clustered together at the maximum bootstrap support value (100%). Strikingly, this group is clustered more closely with southeastern European *C. betulus* populations that occur in the range where both species are sympatric. Together with the fact that both species are almost completely separated on both genotypic and population level this suggests that gene flow between species might be present at a low level. However, this is most likely the reflection at the nuclear level of past hybridisation events that were also inferred from the relationship of chloroplast haplotypes in both species (Grivet & Petit 2002) and it can be stated that no evidence for recent hybridisation was detected.

5.5.3 Structure of nuclear genetic diversity

In accordance with the wind-pollinated and outcrossing breeding system of the hornbeam, little differentiation between populations was recorded in the Flemish data set (F_{ST} = 0.048) and moderate differentiation between European populations (F_{ST} = 0.074), suggesting high gene flow between populations. Both divergence estimates were highly significant (P < 0.0001).

5.5.3.1 Structure at European scale

The moderate divergence between populations on a European scale (F_{ST} = 0.074) is somewhat lower than was found when studying oak populations with AFLP markers on a similar geographic scale (F_{ST} = 0.111 for *Q. robur* and *Q. petraea*, Mariette *et al*. 2002b). However, despite the moderate differentiation, a geographical pattern of genetic diversity was apparent on the genotypic as well as on the population level. The model-based clustering divided the *C. betulus* genotypes in 6 gene pools. Two closely related gene pools (number 1 and 4, genetic distance only 0.026, see Table 5.6) are made up mainly by Belgian, German, British and Swedish genotypes, representing northern and western Europe. A second gene pool

(number 2) represents the southern and southeastern genotypes by grouping Italian, Slovakian, Rumanian and Greek hornbeams. The French and Spanish trees form another gene pool (number 5). A rather distinct gene pool (number 6) contains Swedish and Rumanian trees; these individuals also formed a different cluster on the PCO plot (Figure 5.5). No explanation can be given for this observation. The geographic structuring was confirmed when calculating the correlation between geographic distance and genetic divergence between populations ($r = 0.40$; $P < 0.001$) and when populations were clustered in a neighbour-joining tree.

Two populations were previously shown to be monotypic for a different haplotype than all other investigated populations (Greek population: haplotype 4 and Rumanian population haplotype 5; Grivet & Petit 2002). Despite their distinct chloroplast haplotype, individuals from these populations clustered following the geographic pattern and on the population level, they did not form an outgroup but were also grouped with other southeastern populations. These data suggest that if a nuclear divergence between populations with different haplotypes existed at the verge of recolonisation, local selection pressure and gene flow has erased this pattern and installed a genetic structure following a different pattern. This process was previously revealed in a comprehensive study for white oaks on a European level (Kremer *et al*. 2002).

5.5.3.2 Structuring at Flemish scale

The divergence among Flemish hornbeam populations ($F_{ST} = 0.0466$) was larger than detected for Flemish oak population also based on AFLP markers (0.0185 for *Q. petraea*, 0.0193 for *Q. robur*). However, similarly as for the oak dataset, no geographical pattern of genetic diversity could be revealed. Neither the genotypes nor the populations clustered according to their location of origin and no correlation was established between the geographical and genetic distance between populations. In the light of the established geographic pattern at the European level, different hypotheses can be put forward to explain the lack of structure on a Flemish scale.

First of all, if the influence of historical founding events has already faded and the situation of equilibrium between drift and gene flow can be assumed, the expected geographical pattern might be absent because the spatial scale is too small in relation to the dispersing abilities of the species. This would mean that the gene flow (mainly by pollen) between populations is so high that a geographic pattern can only be revealed at a larger geographic scale. Unfortunately, the discrepancy between fingerprints generated in 2001 (Flemish data) and 2002 (European data) hampers the possibility of positioning the Flemish populations in a European context in order

to further explore this hypothesis. However, the average population divergence between Flemish populations is lower than between European populations so a certain effect of the geographic location can be assumed.

Secondly, if historical factors still have an impact on the genetic structure, then the genetic pattern will reflect the historical relationship between founding populations, which does not necessarily follow an obvious geographic pattern. It can be derived from the European results that although hornbeam has recolonised the Flemish region fairly recently, the pattern of recolonisation is not the major factor shaping present-day nuclear genetic diversity. On the other hand, almost all Flemish hornbeam populations sampled are old pollard trees along fields and roads and are therefore certainly planted. Even if fairly local material was used for these plantations as is assumed in Flanders (Tack *et al.* 1993), the translocation of individuals over a limited distance might have obscured the weak geographic pattern that was present on this small geographic scale. The information on the positive correlation between the genetic and geographical distance among European populations suggests that if hornbeam reproductive material from other European regions was introduced, this would have been apparent in the Flemish dataset provided that the distance to the Flemish region was substantial.

5.5.4 Within population variation

Estimated levels of within-population diversity for *C. betulus* populations are of the same magnitude as was detected for oak populations on a Flemish (mean H_j 0.29 for both *Q. petraea* and *Q. robur*, Coart *et al*. 2002) and European scale (mean H_j 0.23 and 0.22 Mariette *et al*. 2002b). The average gene diversity is lower for Flemish (0.30) than for the European populations (0.33) and this difference was shown to be significant. On the other hand, the diversity estimates of the Flemish population that was analysed in 2001 and 2002 (H_j = 0.27 in 2001; H_j = 0.29 in 2002) showed that, although these estimates are not significantly different, caution is required when making pair-wise comparisons of the level of within-population gene diversities.

Gene diversity estimates were significantly lower for *C. orientalis* populations (mean H_j was only 0.22) but this must be attributed to the small proportion of *C. orientalis* genotypes in the overall dataset. Many markers, typical for *C. betulus* were scored as absent in the vast majority of *C. orientalis* trees, resulting in a high proportion of monotypic markers in *C. orientalis* populations and lower estimates of gene diversity. The fact that AFLP markers were defined on a subset of genotypes that contained both species could not compensate for this.

Surprisingly, the significant correlation between the geographic distance to the glacial refuge and the diversity parameters gene diversity and proportion of polymorphic loci was positive (respectively r = 0.69 and = 0.75). The correlation between longitude and diversity parameters was negative (though only significant for PLP) and indicates the increase of diversity from east to west across Europe. These results suggest that gene diversity has increased during recolonisation despite the bottleneck at the outset of postglacial colonisation. Similar results were found in a comprehensive study of beech (*Fagus sylvatica)* across Europe using isozymes (Comps *et al*. 2001). In this study it was shown that although maximum allelic richness was found in the southeastern part of the range, gene diversity was lower near these refuges than in recently colonised regions. Clearly, no measure equivalent to allelic richness can be computed based on binary AFLP markers and thus no information is available on the putative patterns of allelic richness across European hornbeam populations.

Several factors may have limited the initial loss of gene diversity and then possibly increased it above the initial values found in the refuges. Wind-pollinated trees typically have a delayed reproduction and chances could be high that numerous juvenile migrants arrive at a newly colonised site before reproduction begins and therefore reduce the intensity of the founder event. Furthermore, adult trees may serve as pollen traps, and low density of adult trees in recently founded populations could allow long-distance gene-flow through pollen (Comps *et al*. 2001). For beech it was shown that pollen flow is more widespread where trees are far apart than within dense populations (Merzeau *et al*. 1994). Also the selection pressure during the establishment of populations (acting predominant indirectly on AFLP loci) might promote an increase in heterozygosity, partially caused by the removal of the more inbred individuals (Müller-Starck & Starke 1993).

5.6 Conclusions

A clear separation between *C. betulus* and *C. orientalis* was revealed both on the genotypic and population level. The *C. betulus* genotypes with the haplotype that was shown to be derived from *C. orientalis* (type 5) were not more related to *C. orientalis* than other *C. betulus* individuals. However, the *C. orientalis* group was clustered more closely with *C. betulus* populations of southeastern Europe where both species occur sympatric. This is most likely the reflection at the nuclear level of past hybridisation events that were also inferred from the relationship of chloroplast haplotypes in both species.

In accordance with the wind-pollinated and outcrossing breeding system of the hornbeam, high within-population diversity and little (but significant) genetic

differentiation between populations was detected at Flemish and European scales. At European scale, a significant correlation between genetic and geographic distance was revealed and to some extent gene pools with a distinct geographic location could be delineated. Furthermore, results on the within-population diversity suggest that gene diversity has increased during postglacial recolonisation despite the bottleneck at the outset of recolonisation. At Flemish scale, no apparent geographic pattern underlying the genetic structure was discovered. The weak pattern that possibly was present on this small geographic scale might have been obscured by the translocation of individuals over a limited distance to construct the many line plantations. Alternatively, this scale of analysis might be too detailed in relation to the dispersion capacity of the species.

The current study provides further information on the vastly similar evolutionary history of hornbeam and beech. They are both late successional species (Pott 2000), might have benefited from anthropogenic influence for their colonisation (Küster 1997), refugia were identified in similar regions (unpublished information S. Brewer in Comps *et al*. 2001 and Grivet & Petit 2002) and mainly one Balkan haplotype has colonised western Europe (Demesure *et al*. 1996; Grivet & Petit 2002). This study suggests that the process of recolonisation might have led to similar patterns of genetic diversity in both species. However, the latter should be confirmed by studying more populations preferably with co-dominant markers.

6 General Discussion

6.1 AFLP versus microsatellite markers

Both nuclear marker techniques applied here are generally assumed to generate evolutionary neutral DNA markers. However, it should be pointed out that some AFLP markers will most likely be associated with adaptive traits (see also discussion on differentiation estimates calculated from AFLP data in section 6.1.3) and the functional significance of some microsatellites has been proven (see Li *et al*. 2002 for a comprehensive review).

6.1.1 Use of dominant markers in population genetic studies

For apple and oak, AFLP and microsatellites were applied to study the amount and structure of nuclear diversity. The most important differences between both techniques applied is the fact that AFLP is a multi-locus dominant marker system (visualises multiple loci at a time and is not able to identify allelic composition) whereas SSRs are single locus co-dominant markers (visualise both alleles at a single locus at a time). The estimation of allele frequencies within populations and the corresponding variances is straightforward from co-dominant genotypic data, but estimating allele frequencies from dominant marker data requires previous knowledge of the inbreeding coefficient. For outcrossing species as studied in the present thesis, one can either assume Hardy-Weinberg equilibrium or use estimates of the inbreeding coefficient obtained in the same or in other populations of the species, using a co-dominant marker system. Approaches relying on a known inbreeding coefficient were developed by Lynch & Milligan (1994) and Zhivotovsky (1999). Alternatively, the multilocus phenotype can be treated as a haplotype and distances among haplotypes can subsequently be described in an analysis of molecular variance (AMOVA: Excoffier *et al*. 1992) but this might cause biased results for outcrossing species. In this study, population genetic parameters for AFLP-data were estimated using the software AFLP-SURV 1.0 (Vekemans *et al*. 2002), assuming a non-uniform prior distribution of allelic frequencies with its parameters derived from the observed distribution of fragment frequencies among loci, according to Zhivotovsky (1999). For oak and hornbeam it was shown in this study that the effect of different input values for the inbreeding coefficient affected the computed diversity estimates only slightly (as far as the inbreeding coefficient oscillated between 0.0 and 0.15), resulting in similar estimates and identical conclusions.

Recently, Holsinger *et al*. (2002) proposed a Bayesian approach to infer population structure from dominant markers that does not assume previous information on the within-population level of inbreeding but incorporates uncertainty about the magnitude of inbreeding. For hornbeam, the largest AFLP data set considered in this thesis, the Bayesian method of Holsinger was compared with the method

implemented in AFLP-SURV (results not shown). In this case, the estimates of the inbreeding coefficient did not appear reliable in view of the breeding system of the species and therefore this method only confirmed that in both datasets the population differentiation was significantly different from zero and that the differentiation was higher on the European than on the Flemish level.

6.1.2 Estimates of genetic diversity

Obviously, the different nature of the markers derived from both systems (bi-allelic markers for AFLP, multi-allelic markers for SSR) resulted in different estimates of **heterozygosity values** (summarised in Table 6.1).

Table 6.1: Mean values for nuclear within-population diversity (mean within-population heterozygosity, H_j/H_E) and among-population diversity (mean between conspecific populations genetic differentiation, F_{ST}/G_{ST}).

	AFLP	**SSR**
Within-population diversity		
*Q. petraea**	0.2910	0.8566
*Q. robur**	0.2868	0.8123
M. sylvestris	0.2250	0.7210
C. betulus Flanders	0.2979	-
C. betulus Europe	0.3332	-
Among-population diversity		
*Q. petraea**	0.0185	0.0077
*Q. robur**	0.0193	0.0137
M. sylvestris	0.0464	0.0602
C. betulus Flanders	0.0466	-
C. betulus Europe	0.0740	-

*: For oak, data are taken from the eight populations studied with both marker systems (section 4.5.3).

However, the more important question is whether both techniques resulted in a similar ranking of populations according to the within-population levels of diversity. In the case of oaks, no congruence was found between the ranking of populations for their levels of diversity for the different marker systems. This is in accordance with results from similar studies that also compared within-population diversity levels as estimated from different marker systems (Isabel *et al*. 1995; Le Corre *et al*. 1997; Mariette *et al*. 2001, 2002a and 2002b). The simulation studies of Mariette *et al*. (2002a) indicated that low correlations between diversity estimates might be expected in three situations:

(i) Low heterogeneity of populations for their diversity. For large populations and high levels of gene flow, both life-history traits of trees, all populations tended to be display similar levels of diversity;

(ii) High heterogeneity within the genome and low sampling efforts of markers within the genome. Also in the present study, the sampling variance was relatively large for both marker systems, what might have resulted in inaccurate estimates;

(iii) In recent populations, where within-population diversity did not reach an equilibrium between drift, migration and mutation, a substantial discrepancy between diversity estimated with markers and diversity over the whole genome was observed. Fewer than 500 generations of tree populations have elapsed since the last glaciations and thus the genetic equilibrium may not have been reached (Kremer 1994).

In consequence, it is clear that caution is required not only when comparing diversity estimates based on different marker systems but also when interpreting small differences in diversity estimates derived from one marker system. Especially for tree species, where within-population diversities are high and may be very similar among populations, the ranking observed could be generated by the noise of the estimation. We can thus conclude that in species for which no co-dominant marker systems are available, dominant AFLP-markers provide nice estimations of average genetic diversities in population surveys (as shown in Table 6.1), even though the ranking of the individual populations according to diversity parameters might not fully agree with that estimated from co-dominant marker data.

The most important advantage of using co-dominant markers is that they allow the calculation of other diversity statistics that are affected differently by evolutionary processes. For example, allelic richness is a diversity statistic that is much more affected by historical processes such as bottlenecks and founder events than the expected heterozygosities because rare alleles (that do not affect heterozygosity significantly) are easily lost (Widmer & Lexer 2001; Comps *et al*. 2001). The inbreeding coefficient (F_{IS}), also termed heterozygote deficit, is another important population statistic as its different, more rapid dynamics than heterozygosity reflects better the ongoing evolutionary processes. However, the estimation of F_{IS} from marker data remains problematic, as it can differ greatly between loci and might be greatly affected by the number and nature of loci analysed. The comparison of these different diversity statistics can be exploited in order to unravel the current and historical evolutionary processes that shaped population genetic structure such as bottlenecks and population admixture (Luikart *et al*. 1998; Sunnucks 2000) what was shown remarkably in the study on *Fagus sylvatica* (Comps *et al*. 2001). This is, however, not possible when only dominant marker data are available.

6.1.3 Estimates of inter-specific and inter-population genetic differentiation

The differentiation estimates between taxa were in most cases higher when calculated from AFLP markers, for the oak and apple study (summarised in Table 6.1). However, SSR markers were also able to identify significant differentiation levels and both marker techniques led to essentially the same conclusions. The higher differentiation calculated from AFLP data is in accordance with the conclusions reached in similar studies (e.g. Mariette *et al*. 2001), and might be explained by the large amount of AFLP markers which are usually used (in comparison to co-dominant markers such as SSRs), what increases the chance that some may be localised in the chloroplast or mitochondrial genomes that show a higher population differentiation for most tree species. For instance, it has been shown that some RAPD markers were derived from the mtDNA in Douglas fir (*Pseudotsuga menziesii*, Ennos 1994); Limber pine (*Pinus flexilis*, Latta & Mitton 1997) and sessile oak (*Quercus petraea*, Le Corre *et al*. 1997). Occasional linkage between organelle DNA and AFLP markers is equally possible. Also, the possibility exists that some AFLP markers are linked with adaptive traits or regions of high species/population divergence. This was illustrated in AFLP studies on oaks (this study and Mariette *et al*. 2002), where a high heterogeneity of differentiation values was obtained for different AFLP fragments. Furthermore, it has been shown by different authors (e.g. Lynch & Milligan 1994, Ribeiro *et al*. 2002) that loci with a low frequency of the null allele for a dominant marker system such as AFLP will introduce bias in the analysis which tends to overestimate the genetic differentiation among populations. If the differentiation values were recalculated for the oak data set after pruning of loci according to the Lynch and Milligan correction (in this case all loci with fewer than 3 phenotypes with the null phenotype in any population) the differentiation between species was significantly lower (F_{ST} = 0.0437 in comparison to 0.0717 for the complete data set) but still larger than the differentiation value calculated from SSR markers (G_{ST} = 0.0250). For the genetic differentiation between conspecific populations this correction did not result in very different estimates (for *Q. petraea* populations F_{ST} = 0.0199, for *Q. robur* populations F_{ST} = 0.0177) in comparison to the complete data set (F_{ST} = 0.0185 and 0.0193 respectively). On the other hand, mutation rates are higher in microsatellites and cannot be ignored when compared to migration rates (Slatkin 1995), what could result in lower differentiation estimates. Also homoplasy due to the high mutation rate might increase the underestimation of differentiation, but this effect is often largely compensated by the large amount of variability present at SSR loci (Estoup *et al*. 2002). Homoplasy may also be present among AFLP markers (Vekemans *et al*. 2002), but it was shown in this study (by comparing results from of a reduced data set including a subset of larger AFLP markers that are less sensitive to size homoplasy) that the effect on differentiation estimates is limited.

The high power of AFLPs to distinguish between closely related taxa was illustrated in the study of both *Quercus* species and both *Carpinus* species. In the case of oaks, AFLP markers were able to differentiate between the species (that are shown to be different morphologically and in ecological preferences) whereas six microsatellites were not able to assign an individual oak to the right species. More microsatellites could have the same power (as was shown by Muir *et al*. 2000 analysing 20 SSR loci) but it is apparent that AFLP makes up a very (cost) effective approach to distinguish between closely related taxa.

6.2 Nuclear and cytoplasmic genetic diversity in tree species

High within-population diversity was found in this study for all the species studied, based on both AFLP and microsatellite markers and this in contrast to differentiation between conspecific populations, which was low (all values are summarised in Table 6.1). In oaks, the chloroplast markers analysed revealed low within population diversities and high levels of differentiation for autochthonous populations (H_S = 0.20 and G_{ST} = 0.70). For hornbeam, no chloroplast diversity was found in western Europe (although the authors stress that analysing more cpDNA fragments into more detail might reveal some polymorphisms). At European scale, values even more extreme than for oak were found: H_S = 0.02 and G_{ST} = 0.97 (Grivet & Petit 2002). These clear differences between nuclear and cytoplasmic diversity patterns explain why the decision on conservation and management units should be supported by information on both nuclear and organelle loci (Newton *et al*. 1999).

The present results are typical for the peculiar distribution of genetic variability in forest tree species. Isozyme data collected on a large number of species indicated that trees maintain a significantly higher level of genetic diversity within species and within populations than annual plants for nuclear genes and also show a lower level of genetic differentiation among populations (Hamrick *et al*. 1992, Hamrick & Godt 1996). These results have been confirmed by molecular markers for several species (e.g. for *Q. petraea* by Le Corre *et al*. 1997 and Mariette *et al*. 2002). In contrast, a clear geographic structure and high population divergence are often observed with cytoplasmic markers of the chloroplast DNA (e.g. Ennos 1994, Demesure *et al*. 1996, Palmé *et al*. 2003). The detected patterns of nuclear diversity were at first unexpected in tree species, since the known successive foundation events that occur during colonization yield a strong genetic differentiation and low within-population diversity, especially in populations far from refuges (Austerlitz *et al*. 1997). The usual explanation given for the limited differentiation of nuclear genes in trees and high population differentiation is the asymmetry between high pollen flow and limited seed flow (Ennos 1994, Le Corre *et al*. 1997). Nevertheless, high pollen

flow alone cannot account for this observation since animal-pollinated species were shown not to maintain much higher population differentiation than wind-pollinated species although their pollen dispersal ability is much lower (Hamrick *et al*. 1992, Mariette *et al*. 1997; Raspé & Jacquemart 1998). Recent research pointed out that life-history traits of trees, that do not depend on the level of pollen flow, have played a very important role in shaping the genetic diversity currently present in tree species (Austerlitz *et al*. 2000; Comps *et al*. 2001; Widmer & Lexer 2001). Especially delayed reproduction is highlighted as the key factor in avoiding severe founder effects of trees species, since it allows a significant increase in the number of initial founders of a given population before reproduction begins (Austerlitz *et al*. 2000).

6.3 The relationship between neutral markers, adaptive potential and genetic diversity

A major objective in the present study of genetic diversity of tree species in the Flemish region was the identification of ESUs (Evolutionary Significant Units), populations that are sufficiently distinct to merit conservation status. ESUs are thought to preserve evolutionary potential that can recreate lost biodiversity, provided that evolutionary processes are able to operate (different ESU concepts are comprehensively reviewed in Fraser *et al*. 2001). It is often presumed that statistical significance between groups for neutral markers indicates the presence of biologically important differences or that populations have been separated long enough for biologically important differences to accumulate, what would justify the use of genetic differentiation as observed by neutral markers as basis for the delineation of ESUs. The current study showed expected results for tree species when studied with neutral markers: nuclear diversity was high and mainly attributed to the within-population level, whereas differentiation between conspecific populations was low. However, the observed low differentiation values between populations in this study were often significant and it can be questioned whether the detected differences have a biological meaning. Differentiation was for instance not larger between Belgian and German apple samples than among Belgian samples and similarly, differentiation values between autochthonous and non-autochthonous oak populations were not larger than among autochthonous populations. On the other hand, in the European hornbeam study the genetic divergence was correlated with the geographical distance, whereas this correlation was not present at Flemish scale.

It is clear that with the discovery of highly variable loci (e.g. microsatellites) and the use of large numbers of independent markers (e.g. AFLPs) the association between statistical significance based on molecular markers and biological meaningfulness of

comparisons of groups has become more complicated (Hedrick 2001). The power of AFLP and SSR markers can be extremely high, so that equally small molecular genetic differences between groups become statistically significant. On the other hand, in a study of *Pinus sylvestris* from Finland (Karhu *et al*. 1996) the nuclear markers revealed low differentiation between populations, reflecting (correctly) high gene flow whereas many important quantitative traits showed high differentiation in provenance trials under different environmental conditions.

The overall picture from empirical studies is of adaptive divergence for specific traits taking place in the face of gene flow, with little relationship to patterns exhibited by molecular markers (McKay & Latta 2002). The lack of relationship between neutral markers and quantitative traits can be intuitively deduced from their different evolutionary processes. The variation in quantitative traits might reflect the past influences of selection, which can be different for each gene, superimposed on the pattern of variation as a result of history, migration and drift that is expected to affect all loci in similar ways (van Tienderen *et al*. 2002). Furthermore, the phenotypic variance for polygenic traits includes the covariance of allelic effects of different loci, what gives the possibility of substantial trait differentiation with only minor differentiation of allele frequencies at the underlying loci (or vice versa). Hence, if the QTLs, the targets for selection at the genetic level, themselves only differentiate slightly, there can be no reason to expect neutral molecular markers to reflect the adaptive differentiation of populations. Therefore, the interpretation of genetic variation must distinguish among (neutral) genetic markers, quantitative genetic (polygenic) traits and the genes (QTLs) underlying quantitative traits since each type of variation is likely to have its own pattern of geographic distribution which is likely to be poorly correlated across the three different types (Booy *et al*. 2000; McKay & Latta 2002). The significant differentiation between populations based on SSR or AFLP markers as detected in the present study can therefore only be used to infer the nuclear genetic structure and are by no means a replacement for the provenance trials under different ecological conditions, where adaptive traits can be observed.

Adaptation implies the genetic or phenotypic response of a population or an individual to an environmental change in order to increase or maintain fitness (Booy *et al*. 2000). Therefore, it is stressed in several concepts of ESUs that the unit should contain a significant amount of genetic diversity, because this is often assumed to be the basis for adaptation and therefore determines species or population survival. It was shown in this thesis that small oak relics contain as much diversity as larger forests and it was therefore concluded that these populations can be used as seed source for forestations without jeopardising the genetic diversity in future oak forests. However, the relationship between genetic diversity and adaptive

potential is not straightforward. For instance, in a study on the effect of air pollution on gene pools of Norway spruce, silver fir and beech it was shown that for all species damaged trees exhibited higher levels of both allelic richness and heterozygosity at isozyme loci (Longuaur *et al*. 2001). The authors explained this by the possible deleterious effect associated with rare alleles under the conditions of air pollution. Nevertheless, the latter example might also illustrate that a reason for the preservation of polygenic variation in natural populations lies in the inconsistent relative fitness of genotypes in different environments. If genetic variation is a prerequisite for adaptation to future environmental changes, this unfortunately implies the current presence of suboptimally adapted genotypes, i.e. a genetic load (Booy *et al*. 2000).

To conclude, it can be stated, that although neutral molecular markers provide essential information on the current patterns of genetic diversity and help to understand the evolutionary history and population dynamics of species, as has been shown in this thesis, this information is not sufficient for the development of conservation programs. Furthermore, the greatest understanding of molecular markers information will come when it is used in conjunction with ecological, demographic or physical data collected in the field (Haig 1998; Cruzan 2000). However, the recently developed approaches to assess genetic diversity using markers directly targeted at specific genes or gene families could overcome the problems related to neutral markers (van Tienderen *et al*. 2002).

6.4 Dealing with hybrids in conservation issues

Although hybridisation has long been recognized as an important factor in the evolution of plant species, the harmful effects of hybridisation have also led to the extinction of many populations and species (Barton 2001; Allendorf *et al*. 2001). In this study two hybrid complexes have been studied: hybridisation between the two sympatric oak species and between the wild and domesticated apple. Two main differences exist between both hybrid complexes. First of all, the apple complex is the result of anthropogenic hybridisation after the introduction and cultivation of a domesticated related species whereas the oaks are both indigenous species that are largely sympatric over their range of distribution. In the second place, the wild apple is a rare species in Flanders, in comparison to the cultivated individuals whereas both oak species are fairly common.

M. sylvestris is a very rare species in Flanders (as in most of its distribution range). Although its limited competitive ability suggests that it has always been a rather rare species, the destruction of suitable habitats and habitat fragmentation has further decreased the number and size of populations. It is in this situation, when a rare

species comes in contact with a more abundant one that the effects of hybridisation can be the most problematic (Rhymer & Simberloff 1996). The current study provided first insights into the occurring hybridisation of wild apple with the domesticated apple. The majority of nuclear fingerprints are typical for either wild or cultivated gene pools and only a small percentage of hybrids (4,1%) was detected although cultivated apple trees are present in nature. Study of a chloroplast marker revealed that introgression had taken place in 39% of Belgian wild apples but was absent among the German wild trees. These figures might represent an underestimation of actual hybridisation rates, as this marker is not present in all cultivated genotypes. In addition, as a cytoplasmatic marker was used, only hybridisation in one direction was detected. However, based on our current knowledge, pure individuals (trees with nuclear fingerprints typical of *M. sylvestris*) are still present and the continued existence of hybridised populations might pose a threat to remaining pure populations, what renders the conservation value of hybridised individuals low (Allendorf *et al*. 2001). It is clear that in the case of the wild apple, efforts should focus on maintaining and expanding the remaining 'pure' genotypes and populations.

The two indigenous oak species clearly present a natural hybridisation complex. Many experimental results have supported the hypothesis of the existence of interspecific gene flow and introgression between *Q. petraea* and *Q. robur*, the most important being: (1) the possibility of interspecific crossings (Steinhoff, 1993), (2) the occurrence of natural hybrids and the so called 'regeneration' of *Q. petraea* from successive unidirectional hybridisation with *Q. robur* (Bacilieri *et al.* 1996; Petit *et al*. 1997) and (3) the similar geographic structure of chloroplast haplotypes (Dumolin-Lapègue *et al.* 1999). Moreover, analyses of chloroplast and mitochondrial DNA variation revealed that hybridisations took place during the stay of the oaks in glacial refuges but that also recent hybridisation and introgression events are most likely to occur (Petit *et al.* 1997; Dumolin-Lapègue *et al.* 1999). On the other hand, although the observed differentiation between species is low in comparison to other related species, both nuclear markers and morphological analysis have shown that both species are differentiated on a local scale (this study for Flanders; e.g. Bakker *et al*. 2001a for The Netherlands; Finkeldey 2001a & 2001b for Switzerland) and throughout their natural geographical range (e.g. Zanetto *et al.* 1994; Muir *et al*. 2000; Mariette *et al*. 2002b). Even in populations where both species coexist only very few putative hybrid forms were detected based on both molecular and morphological data (this study; Finkeldey 2001b; Kremer *et al*. submitted). This type of hybridisation and subsequent introgression must be seen as a part of the evolutionary process of the species and should not preclude protection of individuals and populations that are shown to be the result of hybridisation. The occasional formation of species hybrids may even add to the adaptive potential of progenies

from mixed populations (Finkeldey 2001b). Furthermore, the apparent partial reproductive isolation of *Q. petraea* and *Q. robur* also implies that gene exchange and hybridisation among species may be limited even in mixed forests, as the majority of oak populations in Flanders. Indeed, in Flanders, the hybrid *Q. x rosaceae* often occurs sympatric with *Q. petraea.* Taking into account that the hybridisation between these two oak species is a 'natural' phenomenon, there is no reason to treat populations that include hybrid individuals differently in conservation issues.

7 General conclusions and future prospects

7.1 Malus sylvestris

It is clear that the key issue to be solved in order to organise the conservation of *M. sylvestris* is the discrimination between 'genuine' wild genotypes, cultivated genotypes and hybrids. The molecular markers applied in this thesis provided a valuable approach to ascertain the identity of a putative wild genotype. Both AFLPs and SSRs revealed a very clear differentiation between the wild gene pool, edible cultivars and ornamental cultivars, despite the fact that individuals derived from edible cultivars are present in the landscape. A very low contemporary admixture (revealed by nuclear markers) between the wild and edible cultivar group was detected, with only three 'wild' genotypes (or 4.1%) being identified as hybrids. If apple trees were assigned to the wild or cultivated gene pool based on the hairiness of leaves, the results were only congruent with nuclear marker assignment for trees with felted hairy leaves: they were all cultivars. Most wild apples had hairless leaves, but some variation in the degree of hairiness was recorded within *M. sylvestris.* Furthermore, also two of the detected hybrids had hairless leaves. The chloroplast *mat*K-marker resulted in a contrasting picture and showed possible evidence of introgression with edible cultivars for up to 39% of Belgian wild apples. As most of the trees showing putative traces of admixture with the cultivated gene pool at the chloroplast genome, displayed a typical 'wild' nuclear fingerprint, it can be concluded that most of the hybridisation detected nowadays using the chloroplast *mat*K-marker reflects historical hybridisation events between the wild and the cultivated gene pools. No evidence for introgression was detected in the German wild apples. However, hybridisation might have been underestimated since the *mat*K-marker typed was not present in all edible cultivars.

Based on the results of this thesis, further research priorities can be established. First of all, more individuals should be typed in order to confirm the conclusions presented here. Because it is now established that the edible cultivars do not form a monophyletic group, also this gene pool should be further characterised before the extent of past and current hybridisation can be established. For the moment, a larger data set containing ca 150 Dutch wild apples and ca 60 Dutch old cultivated varieties is being analysed in collaboration with Plant Research International (Wageningen, The Netherlands). First results are very concordant with the findings presented in this thesis: SSR markers delineate clearly a wild and cultivated gene pool with only a very low percentage of admixture, whereas hybridisation as detected with the *mat*K-markers seems to be present in the Dutch material but at a lower level than in Belgium (unpublished results PRI and DvP). However, in order to clearly establish the importance of hybridisation of the wild and cultivated gene pools, a chloroplast marker should be used that is present in all cultivars and that unambiguously differentiates between wild and cultivated origins. Ideally, also

nuclear markers typical of cultivars should be used, in order to monitor gene flow between the cultivated gene pool and the wild gene pool through pollen.

This study revealed the status of the adult wild apple trees found nowadays in forests, but no information is yet available on the genetic make-up of their progeny. Since their germination, the forests were further fragmentised and the nearest compatible tree might currently be a cultivated individual. Furthermore, the lack of hybridisation in the German genotypes indicates that it would be useful to study the factors that influence hybridisation between gene pools. Studying the progeny of wild apple trees in different situations might help to identify these factors: e.g. population size, forest size, recreation pressure and management practices.

It was shown that the degree of leaf hairiness, although considered one of the best discriminators between wild and cultivated apples, could not be interpreted as a 'degree of wildness'. The genetic variability for this trait should be further studied in the wild gene pool in order to facilitate determination in the field. Also the correlation between other phenotypic parameters, easy to evaluate in the field, and molecular markers has yet to be studied, including fruit characteristics (form, colour, taste).

Most genetic variation was present within the locations sampled but a significant differentiation was detected within and between Belgian and German origins. Furthermore, the discovery of introgressive hybridisation only in Belgian populations implies the different history of Belgian and German wild apple trees. At this stage, it is therefore advisable to create a regional gene bank including all Belgian trees with nuclear fingerprints typical of *M. sylvestris*. Further decisions on the choice of genotypes to include in a future seed orchard can be made when results on more Belgian genotypes become available. This information will be on hand at the end of 2005, as a result of a OSTC[1]-funded project concerning apple biodiversity. This project brings together different Belgian partners involved in the conservation of apple biodiversity and apple breeding. It is the purpose of this collaboration to study *M. sylvestris*, modern and ancient *M x domestica* cultivars and the putative progenitor species of cultivated apple. Furthermore, also the usefulness of genetic resources of *M. sylvestris* for future apple breeding programs will be evaluated.

7.2 Quercus robur and Q. petraea

As stated in the introduction, emphasis has been laid recently on the preservation of autochthonous genetic resources. Autochthonous populations are seen as a

[1] Belgian federal Office for Scientific, Technical and Cultural affairs; project funded under the second multi-annual scientific support plan for a sustainable development policy – SPSD II.

valuable genetic heritage that has to be preserved for future generations. However, the correct identification of autochthonous oak populations based on field observations can be complex. In this study it was attempted to evaluate the autochthonous character of Flemish oak populations based on their cpDNA diversity and taking as reference the expected pattern of cpDNA haplotypes in the region as a result of postglacial migration. Our conclusions demonstrate that the detected patterns of chloroplast DNA variation in Flemish autochthonous populations still largely reflect the original post-glacial distribution of haplotypes and this method proved thus helpful to determine the origin of oak populations. However, the chloroplast-haplotype method is unable to detect translocation of material along the migration routes, and is therefore unable to provide an airtight evidence of autochthonous origin. As a consequence this methodology should be used in combination with other criteria for the determination of the autochthonous character of an oak stand. In most cases the results were concordant with the field evaluation method developed by Maes (1993) that was applied during the inventory of autochthonous trees and shrubs in Flanders. The congruence between the field assessment and the chloroplast data emphasize the applicability of both evaluation methods.

Furthermore, an important concern of conservationists is the potential effect that past bottlenecks might have had on the genetic diversity and adaptive potential of the populations. This study of nuclear genetic diversity of Flemish autochthonous oak populations revealed that these contained as much genetic diversity as larger natural European forests and no recent bottleneck effects were observed. We can therefore conclude that the comparatively small population sizes and the putative historic bottlenecks that these populations have endured did not result in a significant loss of genetic diversity. At this point, there is no reason to assume that autochthonous populations are under more threat than other oak stands. Moreover, most old coppiced stands are owned by public institutions that are aware of their value. Coppiced oak stools can attain high ages but their rejuvenation can be prepared by stimulating natural regeneration or planting seedlings from a local origin.

Once autochthonous populations are identified and characterised, their conservation can be organised and ESUs need to be delineated. In this study, no significant genetic differentiation was found between autochthonous populations and selected provenances for both oak species. Furthermore, on a European level it was shown that no association exists any more between chloroplastic divergence and phenotypic traits. As a consequence, the delineation of ESUs on the sole basis of cp-DNA polymorphisms does not reflect the enormous levels of diversity present nowadays in forestry populations and would ignore the possible local adaptations

that might have arisen after recolonisation. The delineation of ESUs should also consider present-day levels of nuclear diversity and patterns of differentiation among populations.

In forestry practice, ESUs for forestry trees are often defined as provenances. These provenances can group different seed sources that are assumed to be adapted to similar ecological conditions and are therefore applicable under the same conditions. It could be considered to create one autochthonous provenance for each oak species that would include all conspecific autochthonous populations. In this way, a more representative part of present oak diversity would be used as basis for the creation of future oak forests. If this approach is followed, difficulties might be encountered due to the fact that most autochthonous populations occur on poor soils (that were not deforested and taken into cultivation) and they might not be suitable for forestation of richer habitats. The establishment of provenance trials including autochthonous origins and selected stands and planted under different ecological conditions, would reveal very valuable information on adaptive traits that could hardly be obtained otherwise.

This study also provided information on the differentiation between *Q. robur* and *Q. petraea*. AFLP markers demonstrated a clear differentiation between both oak taxa and were able to assign an individual oak tree to a species, but no species-specific markers were found. Assignment based on AFLP markers and morphological parameters were congruent. The application of six SSR loci could to some extent also differentiate sessile and pedunculate oak populations but more loci should be studied to reach the resolution of AFLP markers. These results are in accordance with the knowledge that both oak species are only differentiated for a few loci. If required, AFLP markers could be applied to a population or seed lot to determine the taxonomic oak species present. However, taking into account that the hybridisation between these two oak species is an occasional and natural occurring phenomenon, there is no reason to treat populations that include *Q. x rosaceae* individuals differently in conservation issues.

The technique applied to analyse the chloroplast-DNA variation can to some extent be used to control the origin of oak reproductive material. It will be especially useful to control the material of autochthonous origin since their progenies should be (in most cases) completely monotypic for one haplotype. Before this technique would be installed as an official control, more trees of the source populations should be typed in order to avoid future problems. The high gene flow between populations and high within-population diversity levels hamper the use of AFLP or SSR markers for the control on reproductive material.

7.3 Carpinus betulus

In accordance with the wind-pollinated and outcrossing breeding system, high within-population diversity and little (but significant) genetic differentiation was detected at Flemish and European scales. However, only at European scale the genetic structure was correlated with the geographic structure. The translocation of individuals over limited distances to create the many line plantations present nowadays in Flanders, might have contributed to the weak geographic pattern that is found at Flemish scale. Within-population genetic diversities were shown to be significantly correlated with the distance from the glacial refuge, suggesting that gene diversity has increased during postglacial recolonisation across Europe, despite the bottleneck at the outset of recolonisation. However, more research including co-dominant markers would be necessary to study the current patterns of genetic diversity into greater detail.

The fact that only one chloroplast haplotype has so far been detected in western Europe obviously implies that this methodology is, in its current state, not informative in order to evaluate the origin of Flemish hornbeam populations. If the cp-DNA of hornbeam would be studied into more detail, more variants might well be detected. However, the chance that this information would become useful on the Flemish level is limited.

The detection at European scale of geographic structure of nuclear diversity patterns, with clear clusters of populations originating from nearby geographic regions, suggests that adaptive changes might exist between hornbeam populations from distant locations. Based on our current knowledge, it is therefore advisable to use reproductive material from own or nearby provenances. Similar as described for oak species, one Flemish hornbeam provenance could be created that groups all hornbeam populations. At present, there is indeed no reason to create different autochthonous provenances for hornbeam in Flanders. Flemish populations were shown to contain lower within-population diversities than European populations; it can be advisable to group the reproductive material of different populations for future forestations. By combining different populations in a single provenance, increased levels of within-population diversity could be obtained in new forestations.

7.4 Importance of a European context

For all the species studied, information on other conspecific European populations has proven to be extremely valuable. The comparison of Belgian and German wild apples allowed us to delineate a *M. sylvestris* gene pool with more certainty, the European chloroplast structure in oak populations was the basis for the evaluation of the origin of Flemish oak populations and, based on the European hornbeam samples, a geographic structure of genetic diversity was inferred. It is clear that when studying genetic diversity of forest tree species and the factors shaping it, a European or range-wide context is crucial, even if the interest is to optimise conservation programmes at regional scale.

8 References

Aas G (1990) Kreuzbarkeit und Unterscheidung von Stiel- und Traubeneiche. Allgemeine Forst und Jagdzeitung, 9-10, 219-221.

Aas G (1993) Taxonomical impact of morphological variation in *Quercus robur* and *Quercus petraea*: a contribution to the hybrid controversy. Annals of Forest Science, 50 (suppl.1), 107-113.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. Trends in Ecology and Evolution, 16, 613-622.

Austerlitz F, Jung-Muller B, Godelle B, Gouyon P-H (1997) Evolution of coamlescence times, genetic diversity and structure during colonization. Theoretical Population Biology, 51, 148-164.

Austerlitz F, Mariette S, Machon N, Gouyon P-H, Godelle B (2000) Effects of colonization processes on genetic diversity: differences between annual plants and tree species. Genetics, 154, 1309-1321.

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

Bacilieri R, Ducousso A, 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. Sylva Genetica, 44, 1-10.

Bacilieri R, Ducousso A, Petit RJ, Kremer A (1996) Mating system and asymmetric hybridization in a mixed stand of European oaks. Evolution, 50, 900-908.

Bakker EG, van Dam BC, Van Eck HJ, Jacobsen E (2001a) A discrimination between *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. based on species-indicative AFLP markers. Forest Genetics, 8, 315-322.

Bakker EG, van Dam BC, Van Eeuwijk FA, Jacobsen E (2001b) Genetic variation and structure of indigenous *Quercus robur* L. populations. In: Towards molecular tools for management of oak forests, Genetic studies on indigenous *Q. robur* L. and *Q. petraea* (Matt.) Liebl. Populations, PhD thesis of the University of Wageningen.

Barrenche T, Bodènes C, Lexer C, Trontin JF, Fluch S, Streiff R, Plomion C, Rousel G, Steinkellner H, Burg K, Favre JM, Glossl J, Kremer A (1998) A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellites, minisatellite, issozyme and 5s rDNA markers. Theoretical and Applied Genetics, 97, 1090-1103.

Barton NH (2001) The role of hybridisation in evolution. Molecular Ecology, 10, 551-568.

Bodénès C, Joandet S, Laigret F, Kremer A (1997a) Detection of genomic regions differentiating two closely related oak species *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. Heredity, 78, 433-441.

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

Booy G, Hendriks RJJ, Smulders MJM, Van Groenendael JM, Vosman B (2000) Genetic diversity and the survival of populations. Plant biology, 2, 379-395.

Bordacs S, Popescu F, Slade D, Csaikl UM, Lesur I, Borovics A, Kézdy P, König AO, Gömöry D, Brewer S, Burg K, Petit RJ (2002) Chloroplast DNA variation of white oaks in northern Balkans and in the Carpatian basin. Forest Ecology Management, 156, 197-209.

Bossema I (1979) Jays and oaks: an eco-ethological study of symbiosis. Behaviour, 70, 1-117.

Bouman F, Boesewinkel D, Bregman R, Devente N, Oostermeijer G (2000) Verspreiding van zaden. KNNV, Utrecht.

Brewer S, Cheddadi R, de Beaulieu JL, Reille M (2002) The spread of deciduous *Quercus* throughout Europe since the last glacial period. Forest Ecology Management, 156, 27-48

Bruford MW, Ciofi C, Funk SM (1998) Characteristics of microsatellites. In: Molecular tools for screening biodiversity (eds Karp A, Isaac PG, Ingram DS), pp. 202-205. Chapman & Hall, London.

Buiteveld, J Bakker EG, Bovenschen J, de Vries SMG (2001) Paternity analysis in a seed orchard of *Quercus robur* L. and estimation of the amount of background pollination using microsatellite markers. Forest genetics, 8, 331-337.

Coart E, Vanden Broeck A, Van Slycken J (1998) Behoud van autochtone bomen en struiken: een verkenning. Mededelingen van het Instituut voor Bosbouw en Wildbeheer 1998-1.

Coart E., De Vreese R. & De Cuyper B. (2001). Studie van Vlaamse relictbestanden van inlandse eik: Genetische diversiteit en situering in Europese migratiestromen. VLINA 00/09, Ministry of the Flemish Community, Brussels.

Coart E, Lamote V, De Loose M, Van Bockstaele E, Lootens P, Roldan-Ruiz I (2002) AFLP markers demonstrate local genetic differentiation between two indigenous oak species (*Quercus robur* L. and *Quercus petraea* (Matt.) Liebl) in Flemish populations. Theoretical and Applied Genetics,105, 431-439.

Comps B, Gömöry D, Letouzey J, Thiebaut B, Petit RJ (2001) Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. Genetics, 157, 389-397.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics, 144, 2001-2014.

Cottrell JE, Munro RC, Tabbener HE, Gillies ACM, Forrest GI, Deans JD, Lowe AJ (2002) Distribution of chloroplast DNA variation in British oaks (Quercus robur and Quercus petraea): the influence of postglacial colonization and human management. Forest Ecology and Management, 156, 181-195.

Coulibaly S, Pasquet RS, Papa R, Gepts P (2002) AFLP analysis of the phenetic organization and genetic diversity of *Vigna unguiculata* L. Walp. reveals extensive gene flow between wild and domesticated types. Theoretical and Applied Genetics, 104, 358-366.

Cruzan MB, Templeton AR (2000) Paleoecology and coalescence: phylogeographic analysis of hypotheses from the fossil record. Trends in Ecology and Evolution, 15, 491-496.

Csaikl UM, Burg K, Fineschi S, König AO, Matyas G, Petit RJ (2002) Chloroplast DNA variation of white oaks in the Alpine region. Forest Ecology Management, 156, 131-195.

De Langhe JE, Delvosalle L, Duvigneaud J, Lambinon J, Vanden Berghen C (1988) Flora van België, het Groothertogdom Luxemburg, Noord-Frankrijk en aangrenzende gebieden, tweede druk. Nationale Plantentuin van België, Meise.

De Riek J, Dendauw J, Mertens M, De Loose M, Heursel J, Van Bockstaele E (1999) Validation of criteria for the selection of AFLP markers to assess the genetic variation of a breeders' collection of evergreen azaleas. Theoretical and Applied Genetics, 99, 1155-1165.

Degen B, Streiff R, 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.

Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. Molecular Ecology. 4: 129-131.

Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. Evolution, 50, 2515-2520.

Dendauw J, De Riek J, De Loose M, Van Bockstaele E. (2002) Identification of 33 Chinese Rhodondendron species using matK sequences and AFLP data.

Proceedings of the 20th EUCARPIA symposium 3-6 July 2001, Melle Belgium.

Djè Y, Heuertz M, Lefèbvre C, Vekemans X (2000) Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. Theoretical and Applied Genetics, 100, 918-925.

Dow BD, Ashley MV, Howe HF (1995) Characterisation of highly variable (GA/CT)n microsatellites in the bur oak, *Quercus macrocarpa*. Theoretical and Applied Genetics, 91, 137-141.

Dow BD, Ashley MV (1996) Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. Molecular Ecology, 5, 615-627.

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus, 12, 13-15.

Ducousso A, Guyon JP, Kremer A (1996) Latitudinal and altitudinal variation of bud burst in western populations of sessile oak (*Quercus petraea* (Matt.) Liebl.). Annals of Forest sciences, 53,775-782.

Dumolin S, Demesure B, Petit J (1995) Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. Theoretical and Applied Genetics, 91, 1253-1256.

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

Dumolin-Lapègue S, Kremer A, Petit RJ (1999) Are chloroplast and mitochondrial DNA variation species independent in oaks? Evolution, 53, 1406-1413.

Dupouey JL (1983) Analyse multivariable de quelques charactères morphologiques de populations de chènes. Annals of Forest Science, 40, 265-282.

Dupouey JL, Badeau V (1993) Morphological variability of oaks (*Quercus robur* L., *Quercus petraea* (Matt.) Liebl., *Quercus pubescens* (Willd.)) in northeastern France: preliminary results. Ann Sci For, 50 (suppl 1), 35s-40s.

Ennos RA (1994) Estimating the relative rates of pollen and seed migration among plant populations. Heredity, 72, 250-259.

Estoup A, Jarne PJ, Cornuet J-M (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Molecular Ecology, 11, 1591-1604.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics, 131, 479-491.

Felsenstein J (1993) PHYLIP, Phylogenetic inference package, version 3.5.7. A computer program distributed by the author, http://evolution.genetics.washington.edu/phylip.html. Department of genetics, University of Washington, Seattle, USA.

Finkeldey R (2001a) Genetic variation of oaks (*Quercus spp*.) in Switzerland. 1. Allelic diversity and differentiation at isozyme gene loci. Forest genetics, 8(3), 185-195.

Finkeldey R (2001b) Genetic variation of oaks (*Quercus spp*.) in Switzerland, 2. Genetic structure in “pure” and “mixed” forests of pedunculate (*Q.robur* L.) and sessile oak (*Q. petraea* (Matt.) Liebl.). Sylvae Genetica, 50, 22-30.

Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Molecular Ecology, 10, 2741-2752.

Gathy P (1969) Contribution a l’étude de la génétique des chènes. FAO2: 979-987.

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998) Simple sequence repeats for the genetic analysis of apple. Theoretical and Applied Genetics, 96, 1069-1076.

Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. Genetics, 139, 463-471.

Gömöry D (2000) A gene coding for a non-specific NAD-dependent dehydrogenase shows a strong differentiation between *Quercus robur* and *Quercus petraea*. Forest Genetics, 7, 167-170.

Gömöry D, Yakovlev I, Zhelev P, Jedinakova J, Paule L (2001) Genetic differentiation of oak populations within the *Quercus robur*/*Quercus petraea* complex in Central and eastern Europe. Heredity, 86, 557-563.

Gregorius HR & Kleinschmit JRG (1999). The environmental dichotomy of adaptation and the role of genetic diversity. Silvae Genetica, 48, 193-199.

Grivet D & Petit R (2002) Chloroplast DNA phylogeography of the hornbeam in Europe: Evidence for a bottleneck at the outset of postglacial colonization. Conservation Genetics, 4, 47-53.

Guilford P, Prakash S, Zhu JM, Rikkerink E, Gardiner S, Bassett H, Forster R (1997) Microsatelittes in *Malus x domestica*: abundance, polymorphism and cultivar identification. Theoretical and Applied Genetics, 94, 249-254.

Haig SM (1998) Molecular contributions to conservation. Ecology, 79, 413-425.

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. New Forest, 6, 95-124.

Hamrick JL, Godt MJW (1996) Effect of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society of London series B-Biological Sciences, 351, 1291-1298.

Hardy OJ, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. Heredity, 83, 145-154.

Hardy OJ, Vekemans X (2002) SPAGeDI: a versatile computer program to analyse spatial structure at the individual or population levels. Molecular Ecology Notes, 2, 618-620.

Harris SA, Robinson JP, Juniper BE (2002) Genetic clues to the origin of the apple. Trends in Genetics, 18, 426-430.

Hartl DL, Clark AG (1997). Principles of Population Genetics, 3nd edn. Sinauer Associates, Inc, Sunderland, MA.

Hedrick PW (1999) Highly variable loci and their interpretation in evolution and conservation. Evolution, 53, 313-318.

Hedrick PW (2001) Conservation genetics: where are we now? Trends in Ecology and Evolution 16, 629-636.

Heuertz M, Hausman J.-F., Tsvetkov I, Frascaria-La Coste N, Vekemans X (2001) Assessment of genetic structure within and among Bulgarian populations of the common ash (*Fraxinus excelsior* L.). Molecular Ecology, 10, 1615-1623.

Heybroek HM (1992). Behoud en ontwikkeling van het genetisch potentiaal van onze bomen en struiken. Dorschkamprapport nr. 684, IKC-NBLF/IBN-DLO, Wageningen.

Hokanson SC, Swezc-McFadden AK, Lamboy WF, McFerson JR (1998) Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus x domestica* borkh. core subset collection. Theoretical and Applied Genetics, 97, 671-683.

Hokanson SC, Lamboy WF, Szewc-McFadden AK, McFerson JR (2001) Microsatellite (SSR) variation in a collection of *Malus* (apple) species and hybrids. Euphytica, 118, 281-294.

Holsinger KE, Lewis PO, Dey DK (2002) A Baysesian approach to inferring population genetic structure from dominant markers. Molecular Ecology, 2002, 1157-1164.

Isabel N, Beaulieu J, Bousequet J (1995) Complete congruence between gene diversity estimates derived from genotypic data at enzyme and random amplified polymorphic DNA loci in black spruce. Proceedings of the National Academy of Sciences USA, 92, 6369-6373.

Isabel N, Beaulieu J, Theriault P, Bousquet J (1999) Direct evidence for biased gene diversity estimates from dominant random amplified polymorphic DNA (RAPD) fingerprints. Molecular Ecology, 8, 477-483.

Jacques D, De Cuyper B (1998) Genetic conservation strategy for social broadleaves in Belgium. In: report of the first EUFORGEN meeting on social broadleaves, 23th-25th October 1997, Bordeaux (France). IPGRI, Rome, Italy.

Janssens GA, Goderis IJ, Broekaert WF, Broothaerts W (1995) A molecular method for S-allele identification in apple based on allele-specific PCR. Theoretical and Applied Genetics, 91, 691-698.

Jarvis DI, Hodgkin T (1999) Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. Molecular Ecology, 8, S159-S173.

Jensen JS, Gillies A, Csaikl U, Munro R, Madsen SF, Rouland H, Lowe A (2002) Chloroplast DNA variation within the Nordic countries. Forest Ecology and Management, 156, 167-180.

Jones EW (1959) Biological flora of the British islands, *Quercus* L. Journal of Ecology, 17, 169-222.

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

Karhu A, Hurme P, Karjalainen M, Karvonen P, Karkkainen K, Neale D, Savolainen O (1996) Do molecular markers reflect patterns of differentiation in adaptive traits of conifers? Theoretical and Applied Genetics, 93, 215-221.

King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. Molecular Ecology, 7, 1151-1163.

King RA, Ferris C (2000) Chloroplast DNA and nuclear DNA variation in the sympatric alder species, *Alnus cordata* (Lois.) Duby and *A. glutinosa* (L.) Gaertn. Biological journal of the Linnean society, 70, 147-160.

King TL, Burke T (1999) Special issue on gene conservation: identification and management of genetic diversity. Molecular Ecology, 8, S1-S3.

Kirby GC (1975) Heterozygote frequencies in small subpopulations. Theoretical Population Biology, 8, 31-48.

Kleinschmit J, Bacilieri R, Kremer A, Roloff A (1995) Comparison of morphological and genetic traits of pedunculate oak (*Q .robur* L.) and sessile oak (*Q. petraea* (Matt.)Liebl.). Silvae Genetica, 5-6, 256-267.

Kleinschmit JRG, Kleinschmit J (1996) Artificial hybridization between *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. In: Inter- and intra-specific variation in European oaks: evolutionary implications and practical

consequences (eds: Kremer A, Mühs H). European commission, EUR 1617EN, Luxembourg: pp 69-87.

Kleinschmit J, Stephan R, Wagner I (1998) Wild fruit trees (*Prunus avium*, *Malus sylvestris* and *Pyrus pyraster*). In: Noble hardwoods network; IPGRI/EUFORGEN Report of the second meeting, 22-25 March 1997, Lourizan, Spain. Compilers Turok J, Collin E, Demesure B, Erikson G, Kleinschmit J, Rusanen M, Stephan R. p51-60. (http://www.ipgri.cgiar.org/networks/euforgen/Networks/Noble_hardwoods/Strategies/nhwildstrat.htm)

König AO, Ziegenhagen B, van Dam BC, Csaikl UM, Coart E, Degen B, Burg K, de Vries SMG, Petit RJ (2002) Chloroplast DNA variation of oaks in western central Europe and genetic consequences of human influences. Forest Ecology and Management, 156, 147-166.

Korban S (1986) Interspecific hybridization in *Malus*. Hortscience, 21, 41-48.

Kremer A, Petit RJ, Zanetto A, Fougère V, Ducousso A, Wagner D, Chauvin C (1991) Nuclear and organelle gene diversity in *Quercus robur* and *Quercus petraea*. In: Genetic variation in European populations of forest trees (eds.: Müller-Starck G, Ziehe M.), pp 141-166. Sauerländers Verlag, Frankfurt am Main.

Kremer A, Petit RJ (1993) Gene diversity in natural populations of oak species. Annals of Forest Science, 50 (suppl 1): 186s-202s.

Kremer A (1994) Genetic diversity and phenotypic variability of forest trees. Genetics Selection Evolution, 26 (suppl 1), 105s-123s.

Kremer A and Menozzi P (2000) Genetic prosesses promoting genetic diversity in oak and beech forest. Proceedings of the 3rd meeting of the EUFORGEN network “Social hardwoods”, Borovets, June 2000.

Kremer A, Goenaga X (2002) Preface. Forest Ecology Management, 156, 1-3.

Kremer A, Kleinschmit J, Cottrell J, Cundall EP, Deans JD, Ducousso A, König AO, Lowe AJ, Munro RC, petit RJ, Stephan RB (2002) Is there a correlation between chloroplastic and nuclear divergence, or what are the roles of history in the selection on genetic diversity in European oaks? Forest Ecology and Management, 156, 75-87.

Kremer A, Dupouey J-L, Deans JD, Cottrell J, Csaikl U, Finkeldey R, Espinel S, Jensen J, Kleinschmit J, Van Dam B, Ducousso A, Forrest I, de Heredia UL, Lowe AJ, Tutkova M, Munro RC, Steinhoff S, Badeau V (submitted) Morphological variation in mixed oak stands (*Quercus robur* and *Quercus petraea*) is stable across western European populations. Annals of Forest Science.

Küster H (1997) The role of farming in the postglacial expansion of beech and hornbeam in the oak woodlands of central Europe. The Holocene, 7, 239-242.

Lahtinen MJ, Pulkkinen P, Helander ML (1996) Potential gene flow by pollen between English oak (*Quercus robur* L.) stands in Finland. Forestry Studies, 28, 47-50.

Latta RG & Mitton JB (1997) A comparison of population differentiation across four classes of gene markers in limber pine (*Pinus flexiblis* L.). Genetics, 146, 1153-1163.

Le Corre V, Dumolin-Lapegue S, Kremer A (1997) Genetic variation at allozyme and RAPD loci in sessile oak (*Quercus petraea*): the role of history and geography. Molecular Ecology, 6, 519-529.

Lefèvre F (2001) Managing the dynamic conservation networks. In: Forest Genetic Resources Management and Conservation. France as a case study. Editor: Teissier du Cros E. Ministry of Agriculture and Fisheries, Bureau of Genetic Resources, Commision of Forest Genetic Resources, INRA DIC, Paris.

Lefèvre F & Collin E. (2001) Static conservation of collections. In: Forest Genetic Resources Management and Conservation. France as a case study. Editor: Teissier du Cros E. Ministry of Agriculture and Fisheries, Bureau of Genetic Resources, Commision of Forest Genetic Resources, INRA DIC, Paris.

Lefort F, Douglas GC (1999) An efficient micro-method of DNA isolation from mature leaves of four hardwood tree species *Acer*, *Fraxinus*, *Prunus* and *Quercus*. Annals of Science, 56, 259-263.

Li Y-C, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Molecular Ecology, 11, 2453-2465.

Liebhard R, Gianfrancschi L, Koller B, Ryder CD, Tarchini R, Van De Weg E, Gessler C. (2002) Development of microsatellites covering the entire apple genome. Molecular Breeding, 10, 217-241.

Linder CR, Taha I, Seiler GS, Snow A.A., Rieseberg LH (1998) Long-term introgression of crop genes into wild sunflower populations. Theoretical and Applied Genetics, 96, 339-347.

Longauer R, Gömöry D, Paule L, Karnosky DF, Mankovská B, Müller-Starck G, Percy K, Szaro R (2001) Selection effects of air pollution on gene pools of Norway spruce, European silver fir and European beech. Environmental pollution, 115, 405-411.

Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. Molecular Ecology, 7, 963-974.

Lynch M (1988) Design and analysis of experimentals on random drift and inbreeding depression. Genetics, 120, 791-807.

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

Maes N (1993). Genetische kwaliteit inheemse bomen en struiken. Deelproject: Randvoorwaarden en knelpunten bij behoud en toepassing van inheems genenmateriaal. IBN-rapport 20, IKC-NBLF, IBN-DLO, Wageningen.

Maes N & Rövekamp C (1998). Oorspronkelijk inheemse bomen en struiken in Vlaanderen. Een onderzoek naar autochtone genenbronnen in de Ecologische Impulsgebieden. Brussel, Ministerie van de Vlaamse Gemeenschap, Afdeling Bos & Groen.

Maes N, Rövekamp C (2000) Onderzoek naar autochtone genenbronnen in de Regionale Landschappen Vlaamse Ardennen en West-Vlaamse heuvels. Vertrouwelijke rapporten, AMINAL, afdeling Bos en Groen, Brussel.

Maguire TL, Peakall R, Saenger P (2002) Comparative analysis of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae) detected by AFLPs and SSRs. Theoretical and Applied Genetics, 104, 388-398.

Maliepaard C, Alston FH, Van Arkel G, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill) using multi-allelic markers. Theoretical and Applied Genetics, 97, 60-73.

Mantel (1967) The detection of disease clustering and a generalized regression approach. Cancer Research, 27, 209-220.

Mariette S, Lefranc M, Legrand P, Taneyhill D, Frascaria-Lacoste N (1997) Genetic variability in wild cherry populations in France. Effects of colonization processes. Theoretical and Applied Genetics, 94, 904-908.

Mariette S, Chagné D, Lézier C et al. (2001) Genetic diversity within and among *Pinus pinaster* populations:comparison between AFLP, microsatellite markers. Heredity, 86, 469-479.

Mariette S, Le Corre V, Austerlitz F, Kremer A (2002a) Sampling within the genome for measuring within-poulation diversity: trade-offs between markers. Molecular Ecology, 11, 1145-1156.

Mariette S, Cottrell J, Csaikl UM, Goikoechea P, König A, Lowe AJ, Van Dam BC, Barrenche T, Bodénès, Streiff R, Burg K, Groppe K, Munro RC, Tabbener H, Kremer A (2002b) Comparison of levels of genetic diversity detected with AFLP and microsatellite markers within and among mixed *Q. petraea* (Matt.) Liebl. and *Q. robur* L. stands. Silvae Genetica, 51, 72-79.

Marquard RD, Chan CR (1995) Identifying crabapple cultivars by isozymes. Journal of the American Society for Horticultural Science, 120, 706-709.

Matyas G and Sperisen C (2001) Chloroplast DNA polymorphisms provide evidence for postglacial recolonization of oaks (*Quercus spp*.) across the Swiss Alps. Theoretical Applied Genetics, 102, 12-20.

McKay JK & Latta RG (2002) Adaptive population divergence: markers, QTL and traits. Trends in Ecology and Evolution, 17, 285-291.

Merzeau D, Comps B, Thiébaut B, Letouzey (1994) Estimation of *Fagus sylvatica* L. mating system parameters in natural populations. Annals of Forest Science, 51, 163-173.

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

Morgan DR, Soltis DE, Robertson KR (1994) Systematic and evolutionary implications of *rbc*L sequence variation in *Rosaceae*. American Journal of Botany, 81, 890-903.

Moritz C (1994) Applications of mitochondrial DNA analysis in conservation: a critical review. Molecular Ecology, 3, 401-411.

Moritz C (1999) A molecular perspective of biodiversity. In: The biology of biodiversity (Editor Kato S), Springer Verlag, Tokyo.

Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. Trends in Ecology and Evolution, 14, 389-394.

Müller-Starck G & Starke R (1993) Inheritance of iso-enzymes in European beech (*Fagus sylvatica* L.). Journal of Heredity, 84, 291-296.

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

Muir G, Fleming CC, Schlötterer C (2001) Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. Molecular Biology and Evolution, 18, 112-119.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, 89, 583-590.

Nei M (1987) Molecular evolutionary genetics. Colombia university press, 512 pp. New York.

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. Ann. Hum. Gent., 47, 253-259.

Newton AC, Allnut TR, Gillies ACM, Lowe AJ, Ennos RA (1999) Molecular phylogeography, intraspecific variation and the conservation of tree species. Trends in Ecology and Evolution, 14, 140-145.

Nicholas KB, Nicholas HBJ, Hugh BJ (1997) GeneDoc: a tool fro editing and annotating multiple sequence alignments. Distributed by the author: www.psc.edu/biomed/genedoc.

Olalde M, Herrán A, Espinel S, Goicoechea PG (2002) White oaks phylogeography in the Iberian Peninsula. Forest Ecology and Management, 156, 89-102.

Olsson V (1975) Morphological analysis of phenotype in populations of *Quercus* species in Sweden. Botaniska Notiser, 128, 55-69.

Opstaele B (2001). Autochtone bomen en struiken in de houtvesterijen Leuven en Hasselt. Brussel, Ministerie van de Vlaamse Gemeenschap, afdeling Bos & Groen, Brussel.

Palmé AE, Su Q, Rautenberg A, Manni F, Lascoux M (2003) Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. Molecular Ecology, 12, 201-212

Petit RJ, Kremer A, Wagner DB (1993) Geographical structure of chloroplast DNA polymorphisms in European oaks. Theoretical and Applied Genetics, 87, 122-128.

Petit RJ, Pineau E, Demesure B, Bacilieri R, Ducousso A, Kremer A (1997) Chloroplast DNA footprints of postglacial recolonisation by oaks. Proceedings of the National Acadadamy of Science of USA, 94, 9996-10001.

Petit JP, Csaikl UM, Bordács S, Burg K, Coart E, Cottrell J, van Dam B, Deans JD, Dumolin-Lapègue S, Fineshi S, Finkeldey R, Gillies A, Glaz I, Goicoechea PG, Jensen JS, König K, Lowe AJ, Madsen SF, Mátyás G, Munro RC, Olalde M, Pemonge M-H, Popoescu F, Slade D, Tabbener H, Taurchini D, de Vries SGM, Ziegenhagen B, Kremer A (2002a) Chloroplast DNA variation in European white oaks: Phylogeography and patterns of diversity based on data from over 2600 populations. Forest Ecology and Management,156, 5-26.

Petit JP, Brewer S, Bordács S, Burg K, Cheddadi R, Coart E, Cottrell J, Csaikl UM,van Dam B, Deans JD, espinel S, Fineshi S, Finkeldey R, Glaz I, Goicoechea PG, Jensen JS, König K, Lowe AJ, Madsen SF, Mátyás G, Munro RC, Olalde M, Pemonge M-H, Popoescu F, Slade D, Tabbener H, de Vries SGM, Ziegenhagen B, de Beaulieu J-L, Kremer A (2002b) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. Forest Ecology and Management, 156, 49-74.

Phipps JB, Robertson KR, Smith PG, Rohrer JR (1990) A checklist of the subfamily *Maloideae* (*Rosaceae*). Canadian Journal of Botany, 68, 2209-2269.

Pons O, Petit RJ (1995) Estimation, variance and optimal sampling of gene diversity: 1. haploid locus. Theoretical and Applied Genetics, 90, 462-470.

Pott R (2000) Paleoclimate and vegetation- long term vegetation dynamics in central Europe with particular reference to beech. Phytocoenologica, 30, 285-333.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Molecular Breeding, 2, 225-238.

Pritchard JK, Stephens M,Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics, 155, 945-959.

Ralska-Javiewiczowa M (1964) Correlation between the Holocene history of *Carpinus betulus* and prehistoric settlement in North Poland. Acta societatis Botanicorum Poloniae, 33, 461-468.

Raspé O, Jacquemart A-L (1998) Allozyme diversity and genetic structure of European populations of *Sorbus aucuparia* L. (*Rosaceae: Maloideae*). Heredity, 81, 537-545.

Raymond M, Rousset F (1995) GENEPOP (version12) Population genetics software for exact tests and ecumenicism. Journal of Heredity,86, 248-249.

Remmy K, Gruber F (1993) Untersuchungen zur Verbreitung und Morphologie des Wild-Apfels (*Malus sylvestris* (L.) Mill.). Mitteilungen der Deutschen Dendrologischen Gesellschaft, 81, 71-94.

Rhymer JM, Simberloff D (1996) Extinction by hybridisation and introgression. Annual Review of Ecology and Systematics, 27, 83-109

Ribeiro MM, Mariette S, Vendramin GG, Szmidt AE, Plomion C, Kremer A (2002) Comparison of genetic diversity estimates within and among populations of maritime pine using chloroplast simple-sequence repeat and amplified fragment length polymorphism data. Molecular Ecology, 11, 869-877.

Rieseberg LH and Ellstrand NC (1993) What can molecular and morphological markers tell us about plant hybridisation. Critical reviews in Plant Science, 12, 213-241

Robinson JP, Harris SA, Juniper BE (2001) Taxonomy of the genus *Malus* Mill. (*Rosaceae*) with emphasis on the cultivated apple, *Malus x domestica* Borkh. Plant Systematics and Evolution, 226, 35-58.

Rohlf FJ (2000) NTSys: numerical taxonomy and multivariate analysis system, version 2.1. State University of New York

Roldán-Ruiz I, Dendauw J, Van Bockstaele E, Depicker A, De Loose M (2000) AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). Molecular Breeding, 6, 125-134.

Roldán-Ruiz I, van Eeuwijk FA, Gilliland TJ, Dubreuil P, Dillmann C, Lallemand J, De Loose M, Baril P (2001) A comparative study of molecular and morphological methods describing relationships between perennial ryegrass (*Lolium perenne* L.) varieties. Theoretical and Applied Genetics, 103, 1138-1150.

Rövekamp C & Maes N (1999). Oorspronkelijk inheemse bomen en struiken in Vlaanderen. Een onderzoek naar autochtone genenbronnen in de Regionale Landschappen West-Vlaamse Heuvels, Vlaamse Ardennen en de houtvesterijen Hechtel en Bree. Brussel, Ministerie van de Vlaamse Gemeenschap, Afdeling Bos & Groen, Brussel.

Rövekamp C & Maes N (2000). Oorspronkelijk inheemse bomen en struiken in het Regionaal Landschap West-Vlaamse Heuvels. Een onderzoek naar autochtone genenbronnen. Brussel, Ministerie van de Vlaamse Gemeenschap, Afdeling Bos & Groen, Brussel.

Rövekamp C, Maes N & Zwaenepoel A (2000). Oorspronkelijk inheemse bomen en struiken en cultuurwilgen in de Vlaamse Vallei. Een onderzoek naar autochtone genenbronnen in Gent en omgeving. Brussel, Ministerie van de Vlaamse Gemeenschap, Afdeling Bos & Groen, Brussel.

Rushton BS (1976) Pollen grain size in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. Watsonia, 11, 137-140.

Rushton BS (1977) Artificial hybridization between *Q. robur* and *Q. petraea*. Watsonia, 11, 229-236.

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

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. Trends in Ecology and Evolution, 1, 9-10.

Samuel R, Pinsker W, Ehrendorfer F (1995) Electrophoretic analysis of genetic variation within and between populations of *Quercus cerris*, *Q. pubescens*, *Q. petraea* and *Q. robur* (*Fagaceae*) from eastern Austria. Botanica Acta, 108, 290-299.

Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA phylogeny, reticulate evolution and biography of *Paeonia* (*Paeoniaceae*). American Journal of Botany, 84, 120-136.

Santaigne C, Barrenche T, Bodénès C, Bertocchi E, Roussel G, Plomion C, Kremer A (2002) Genomic organisation of molecular differentiation between closely

related white oak species (*Q. petraea* (Matt.) Liebl. and *Q. robur* L.). Proceedings of the Plant, Animal & Microbe Genomes Conference 12th-16th January 2002, San Diego USA

Savolainen V, Corbaz R, Moncousin C, Spichiger R, Manen JF (1995) Chloroplast DNA variation and parentage analysis in 55 apples. Theoretical and Applied Genetics, 90, 1138-1141.

Schaal BA, Hayworth DA, Olsen KM. (1998) Phylogeographic studies in plants: problems and prospects. Molecular Ecology, 7, 465-474.

Schlosser S, Reichhoff L, Hanelt P (1991) Wildpflanzen Mitteleuropas, Nutzung und Schutz. DLV Berlin.

Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2000: A software program for population genetics data analysis. Distributed by the authors, http://anthro.unige.ch/arlequin. Genetics and biometry laboratory, University of Geneva, Switzerland.

Schwarz O (1964) *Quercus* L.; In: Tutin TG, Heywood VH, Burges NA, Valentine DH, Walters SM, Webb DA (editors) Flora Europeae, vol 1: *Lycopodiaceae* to *Platanaceae*, pp 61-64. Cambridge University Press.

Siegismund HR and Jensen JS (2001) Intrapopulation and interpopulation genetic variation of *Quercus* in Denmark. Scandinavian Journal of Forest Research, 16 103-116.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics, 139, 457-462.

Smulders MJM, Van der Schoot J, Geerts RHEM, Antonisse-de Jong AG, Korevaar H, van der Werf A, and Vosman B (2000) Genetic diversity and the reintroduction of meadow species. Plant Biology, 2, 447-454.

Stanley RG and Linskens HF (1974) Pollen: biology, biochemistry and management. Springer verlag.

Steinhoff S (1993) Results of species hybridization with *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. Annals of Forest Science, 53, 173-180.

Steinkellner H, Fluch S, Turetschek E, Lexer C, Streiff R, Kremer A, Burg K, Glossl J (1997) Identification and characterisation of (GA/CT)n microsatellite loci from *Quercus petraea*. Plant Molecular Biology, 33, 1093-1096.

Stephan BR, Wagner I, Kleinschmit, J (in press) Wild apple (*Malus sylvestris* [L.] Mill.) and pear (*Pyrus pyraster* [L.] Burgsd.) Technical guidelines EUFORGEN: www.ipgri.cgiar.org/networks/euforgen/euf_home.htm.

Streiff R, Labbe T, Bacilieri R, Steinkellner H, Glossl J, Kremer A (1998) Within-population genetic structure in *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. Assessed with isozymes and microsatellites. Molecular Ecology, 7, 317-328.

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

Sunnucks P (2000) Efficient genetic markers for population biology. Trends in Ecology and Evolution, 15, 199-203.

Taberlet P, Gielly L, patou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology, 17, 1105-1109.

Tack G, van den Bremt P, Hermy M (1993). Bossen van Vlaanderen, een historische ecologie. Davidsfonds, Leuven.

Thompson JD, Higgins DG and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22,4673-4680.

Vander Mijnsbrugge K, Coart E, Van Slycken J Conservation measures for autochthonous oaks in Flanders. Submitted to Forest Genetics.

van Tienderen PH, de Haan AA, van der Linden G, Vosman B (2002) Biodiversity assessment using markers for ecologically important traits. Trends in Ecology and Evolution, 17, 577-582

Van Valen L (1976) Ecological species, multispecies and oaks. Taxon, 25, 233-239

Vekemans X & Lefèbvre C (1997) On the evolution of the heavy-metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. Journal of Evolutionary Biology, 10, 175-191.

Vekemans X, Beauwens T, Lemaire M, Roldan-Ruiz I (2002) Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and a relationship between degree of homoplasy and fragment size. Molecular Ecology, 11, 139-151.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research, 23, 4407-4414.

Wagner I (1996) Zusammenstellung morphologischer Merkmale und ihrer Ausprägungen zur Unterscheidung von Wild- und Kulturformen des Apfel- (*Malus*) und des Birnbaumes (*Pyrus*). Mitteilungen der Deutschen Dendrologischen Gesellschaft, 82, 87-108.

Wagner I (1998) Artenschutz bei Wildapfel – Die Blattbehaarung von 116 Apfelklonen auf zwei Samenplantagen. Forst und Holz, 53 (2), 40-43.

Wagner I, Weeden NF (2000) Isozymes in *Malus sylvestris*, *Malus domestica* and in related *Malus* species. Acta Horticulturae, 538, 51-56

Watkins R (1995) apples and pears. In: Smartt J, Simmonds NW (Editors) Evolution of crop plants. Longman, London, 418-422.

Weeda EJ, Westra R, Westra CH, Westra T (1985) Nederlandse oekologische flora, wilde planten en hun relaties, deel I. IVN.

Widmer A & Lexer C (2001) Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. Trends in Ecology and Evolution, 16, 267-269.

Wigston DL (1974) Cytology and genetics of oaks. In: The British oak (eds.: :Morris MG, Perring FH), London.

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

Zanetto A, Kremer A (1995) Geographical structure of gene diversity in *Quercus petraea* monolocus patterns of variation. Heredity, 75, 506-517.

Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. Molecular Ecology, 8, 907-913.

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

List of publications

Publications – peer reviewed

- E. Coart, V. Lamote, M. De Loose, E. Van Bockstaele, P. Lootens, I. Roldán-Ruiz (2002) AFLP markers demonstrate genetic differentiation between two indigenous oak species (*Quercus robur* L. and *Quercus petraea* (Matt.) Liebl.) in Flemish populations. Theoretical and Applied Genetics, 105, 431-439.
- E. Coart, X. Vekemans, M.J.M. Smulders, I. Wagner, J. Van Huylenbroeck, E. Van Bockstaele, I. Roldán-Ruiz (2003) 'Genetic variation in the endangered Wild apple (*Malus sylvestris* (L.) Mill.) in Belgium as revealed by AFLP and microsatellite markers. Molecular Ecology, 12, 845-857.
- L. Triest, B. De Greef, R. De Bondt, D. Vanden Bossche, M. D'Haeseleer, J. Van Slycken, E. Coart (1997) Use of RAPD markers to estimate hybridization in *Salix alba* and *Salix fragilis*. Belgian Journal of Botany, 129, 140-148.
- L. Triest, B. De Greef, M. D'Haeseleer, G. Echchgadda, J. Van Slycken, E. Coart (1997). Variation and inheritance of isozyme loci in controlled crosses of *Salix alba* and *Salix fragilis*. Silvae Genetica, 47, 88-94.
- L. Triest, B. De Greef, S. vermeersch, J. Van Slycken, E. Coart (1999) Genetic variation and putative hybridization in Salix alba and S. fragilis (Salicaceae): evidence from allozyme data. Plant Systematics and Evolution, 215, 169-187.
- A.O. König, B. Ziegenhagen, B.C. van Dam, U.M. Csaikl, E. Coart, B. Degen, K. Burg, S.M.G. de Vries, R.J. Petit (2002) Chloroplast DNA variation of oaks in western central Europe and genetic consequences of human influences. Forest Ecology and Management, 156, 147-166.
- R.J Petit, U.M. Csaikl, S. Bordács, K. Burg, E. Coart, J Cottrell, B.C. van Dam, J.D. Deans, S. Dumolin-Lapègue, S. Fineshi, R. Finkeldey, A. Gillies, I. Glaz, P.G. Goicoechea, J.S. Jensen, A. König , A.J. Lowe, S.F. Madsen, G. Mátyás, R.C. Munro, M. Olalde, M.-H. Pemonge, F. Popoescu, D. Slade, H. Tabbener, D. Taurchini, S.G.M. de Vries, B. Ziegenhagen, A. Kremer (2002) Chloroplast DNA variation in European white oaks: Phylogeography and patterns of diversity based on data from over 2600 populations. Forest Ecology and Management, 156, 5-26.
- R.J. Petit, S. Brewer, S. Bordács, K. Burg, R. Cheddadi, E. Coart, J. Cottrell, U.M. Csaikl, B.C. van Dam, J.D. Deans, S. Espinel, S. Fineshi, R. Finkeldey, I. Glaz, P.G. Goicoechea, J.S. Jensen, A. König, A.J. Lowe, S.F. Madsen, S. Mátyás, R.C. Munro, M. Olalde, M.-H. Pemonge, F. Popoescu, D. Slade, H. Tabbener, S.G.M. de Vries, B. Ziegenhagen, J.-L. de Beaulieu, A. Kremer (2002) Identification of refugia and post-glacial colonisation routes of

European white oaks based on chloroplast DNA and fossil pollen evidence. Forest Ecology and Management, 156, 49-74.

- V. Lamote, I. Roldán-Ruiz, E. Coart, M. De Loose, E. Van Bockstaele (2002) A study of genetic variation in *Iris pseudacorus* populations using Amplified Fragment Length Polymorphisms (AFLP). Aquatic botany, 73, 19-31.

Publications – not peer reviewed

- E. Coart, V. Lamote, E. Van Bockstaele, M. De Loose (1999) Estimating the genetic diversity in autochthonous and introduced populations. Case studies on *Quercus* sp. and *Phragmites australis*. Med. Fac. Landbouww. Univ. Gent, 64/5b.
- V. Lamote, E. Coart, E. Van Bockstaele, M. De Loose (2000) Estimating genetic diversity in autochthonous populations. Case studies on *Quercus sp*. and *Iris pseudacorus* L. Med. Fac. Landbouww. Univ. Gent, 65/3b.
- V. Lamote, E. Coart, M. De Loose, E. Van Bockstaele, I. Roldán-Ruiz (2001) Genetic variation within and among patches of the clonal species, Phragmites australis. Med. Fac. Landbouww. Univ. Gent, 66/4.

Reports

- E. Coart (1999) Genetische diversiteit van Inlandse eiken (*Quercus robur* L. en *Quercus petraea* (Matt.) Liebl.), Wilde appel (*Malus sylvestris* (L.) Mill.) en Haagbeuk (*Carpinus betulus* L.) in Vlaanderen. Verslag betreffende de periode 01/1999 – 08/1999, in opdracht van afdeling Bos & Groen.
- E. Coart (2000) Genetische diversiteit van Inlandse eiken (*Quercus robur* L. en *Quercus petraea* (Matt.) Liebl.), Wilde appel (*Malus sylvestris* (L.) Mill.) en Haagbeuk (*Carpinus betulus* L.) in Vlaanderen. Verslag betreffende de periode 09/1999 – 02/2000, in opdracht van afdeling Bos & Groen.
- E. Coart (2000) Genetische diversiteit van Inlandse eiken (*Quercus robur* L. en *Quercus petraea* (Matt.) Liebl.), Wilde appel (*Malus sylvestris* (L.) Mill.) en Haagbeuk (*Carpinus betulus* L.) in Vlaanderen. Verslag betreffende de periode 03/2000 – 10/2000, in opdracht van afdeling Bos & Groen.
- E. Coart (2001) Genetische diversiteit van Inlandse eiken (*Quercus robur* L. en *Quercus petraea* (Matt.) Liebl.), Wilde appel (*Malus sylvestris* (L.) Mill.) en Haagbeuk (*Carpinus betulus* L.) in Vlaanderen. Verslag betreffende de periode 11/2000 – 05/2001, in opdracht van afdeling Bos & Groen.
- E. Coart (2001) Studie van de relatie tussen genetische parameters van eik en morfologische en standplaatskundige parameters. Verslag betreffende de periode 07/2001 – 12/2001. Verslag in opdracht van het Instituut voor Bosbouw en Wildbeheer.
- E. Coart, R. De Vreese, B. De Cuyper (2001) Studie van Vlaamse relictbestanden van inlandse eik: Genetische diversiteit en situering in

Europes migratiestromen. Verslag betreffende de periode 07/2000-12/2000, in opdracht van afdeling Natuur (VLINA-programma)

- E. Coart, R. De Vreese, B. De Cuyper (2001) Studie van Vlaamse relictbestanden van inlandse eik: Genetische diversiteit en situering in Europes migratiestromen. Verslag betreffende de periode 01/2001-06/2001, in opdracht van afdeling Natuur (VLINA-programma)
- E. Coart (2002) Genetische diversiteit van Inlandse eiken (*Quercus robur* L. en *Quercus petraea* (Matt.) Liebl.), Wilde appel (*Malus sylvestris* (L.) Mill.) en Haagbeuk (*Carpinus betulus* L.) in Vlaanderen. Verslag betreffende de periode 06/2001 – 02/2002, in opdracht van afdeling Bos & Groen.
- K. De Cock, E. Coart (2002) Populatiebiologie van autochtone rozen (*Rosa spp*.) en meidoornen (*Crataegus spp*.) in Vlaanderen. Verslag betreffende de periode 01/2002 – 10/2002, in opdracht van afdeling Bos & Groen.
- E. Coart (2003) Genetische diversiteit van Inlandse eiken (*Quercus robur* L. en *Quercus petraea* (Matt.) Liebl.), Wilde appel (*Malus sylvestris* (L.) Mill.) en Haagbeuk (*Carpinus betulus* L.) in Vlaanderen. Verslag betreffende de periode 02/2002 – 12/2002, in opdracht van afdeling Bos & Groen.
- E. Coart (2003) Onderzoek naar de plaats van de Vlaamse haagbeuk (*Carpinus betulus* L.) in Europa: opstellen van aanbevelingen voor het gebruik van autochtone en buitenlandse herkomsten. Verslag betreffende de periode 02/2002 – 12/2002, in opdracht van afdeling Bos & Groen.

Curriculum Vitae

Coart Elisabeth (Els)

B.A. Verhaegenlaan 52
9820 Merelbeke
09/231.83.40
Els.Diego@pandora.be

geboren te Lier, 10 augustus 1972
Belg
Ongehuwd

Opleiding

1985-1990: Humaniora ASO, Latijn-Wiskunde, St-Ursula-lyceum Lier.

1990-1995: Bio-ingenieur in het land- en bosbeheer, minor fytotechnie, K.U.Leuven. Afgestudeerd met onderscheiding.
Thesis: Study of the ancient oak woodlands and dynamics of the natural regeneration in Coedydd Aber National Nature Reserve (North Wales). University of Bangor, UK.

Werkervaring

01/09/1995-31/12/1995: Stage houtvesterij Turnhout voor het opstellen van bosbeheersplannen.

01/01/1996-31/12/1998: Wetenschappelijk medewerkster aan het Instituut voor Bosbouw en Wildbeheer (Ministerie van de Vlaamse Gemeenschap), afdeling veredeling en genetica.

Sinds 01/01/1999: Tewerkgesteld op het CLO, departement Plantengenetica en –veredeling als doctoraatstudent. Onderzoeksthema: Genetische diversiteit van Inlandse eik, Wilde appel en Haagbeuk in Vlaanderen.

Bijkomende opleidingen

- Workshop “Molecular tools for screening biodiversity”, EU DGXII Biotechnology research programme, Long-Ashton, UK. 1-11 juli 1997.
- Cursus "Autochtone bomen en struiken", Educatief Bosbouwcentrum Groenendaal, voorjaar 1998.
- Plantenveredeling, FLTBW, UG. Academiejaar 1997-1998.
- Toegepaste moleculaire biologie en genetica. FLTBW, UG. Academiejaar 1998-1999.
- Classification, identification, typing and phylogeny of bacteria. Faculteit wetenschappen, UG. Academiejaar 1999-2000.
- Praktijkgerichtte statistiek: basismodule en module multivariate statistiek. IVPV, Gent, België, voorjaar 2000 en voorjaar 2001.
- Biostatistics. IVPV, Gent, België, oktober 2001.

Deelname aan nationale en internationale congressen, symposia en studiedagen

- Biodiversiteit en biometrie. CPRO-DLO, 's Graveland, Nederland, 11 februari 1999.
- L' approvisionnement en espèces végétales locales dans les aménagements: Quel(s) enjeu(x) pour la diversité végétale? Versailles, Frankrijk, 16 maart 1999.
- 9th European congress on biotechnology, Brussel, België, 11 – 15 juli 1999.
- 13th Forum for Applied Biotechnology, Gent, België, 22 – 23 september 1999. **Poster**: E. Coart, V. Lamote, E. Van Bockstaele, M. De Loose. Estimating the genetic diversity in autochthonous and introduced populations. Case studies on *Quercus sp*. and *Phragmites australis*.
- Colloquium "De uitvoering door België van het verdrag inzake biologische diversiteit", 17 november 1999, Brussel.
- Molecular markers for characterizing genotypes and identifying cultivars in horticulture, ISHS, Montpellier-France, March 6-8, 2000. **Poster**: E. Coart, V. Lamote, J. De Riek, E. Van Bockstaele, M. De Loose. The use of AFLP markers for characterisation of biodiversity in indigenous trees and data analysis.
- Special symposium British Ecological Society. Plants stand still but their genes don't: Integrating ecological and evolutionary processes in a spatial context. London, UK, 29 – 31 augustus 2000. **Poster**: V. Lamote, E. Coart, E. Van Bockstaele, M. De Loose. The use of AFLP-markers for characterization of biodiversity in yellow flag and oak species.
- 14th Forum for Applied Biotechnology, Brugge, België, 27 – 28 september 2000. **Poster**: V. Lamote, E. Coart, E. Van Bockstaele, M. De Loose. Estimating genetic diversity in autochthonous populations. Case studies on *Quercus sp*. and *Iris pseudacorus* L.
- Symposium on biodiversity: From genes to landscapes. Louvain La Neuve, België, 13 – 15 december 2000.
- EUCARPIA, 20th international symposium section ornamentals. Strategies for new ornamentals. Melle, België, 3 – 6 juli 2001. **Poster**: J. Van Huylenbroeck, E. Coart, F. Janneteau, J. De Riek. Identification of woody ornamentals by means of AFLP fingerprinting.
- Forth EUFORGEN Social Broadleaves Network Meeting, Bergen-Norway, 14-16 june 2001.
- Moleculaire technieken in de milieumicrobiologie: een nieuwe kijk op de enorme diversiteit. KaHo Sint Lieven, Gent, België, 7 mei 2001.
- IUFRO symposium on 'Population and Evolutionary Genetics of Forest Trees', Stara Lesna, Slovakije, August 25-29, 2002. **Poster**: E. Coart, X. Vekemans, M.J.M. Smulders, I. Wagner, J. Van Huylenbroeck, E. Van Bockstaele, I. Roldán-

Ruiz. Genetic ariation in the endangered wild apple (*Malus sylvestris* (L.) Mill.) in Belgium as revealed by AFLP and microsatellite markers.

- DYGEN conference, Dynamics and conservation of genetic diversity in forest ecosystems. Straatsburg, Frankrijk, 2 – 5 december 2002. **Poster**: E. Coart, X. Vekemans, M.J.M. Smulders, I. Wagner, J. Van Huylenbroeck, E. Van Bockstaele, I. Roldán-Ruiz. Genetic variation in the endangered wild apple (*Malus sylvestris* (L.) Mill.) in Belgium as revealed by AFLP and microsatellite markers.
- Third international symposium on 'Ecological Genetics', Leuven, Belgium, 5 – 7 februari 2003. **Voordracht**: Genetic variation in the endangered wild apple (*Malus sylvestris* (L.) Mill.) in Belgium as revealed by AFLP and microsatellite markers.
- Studiedag 'Starters in het bosonderzoek', Brussel, 25 februari 2003. **Voordracht**: Hoe de eik Europa veroverde: onderzoek naar de herkomst van eikenbestanden.

Begeleiding scripties

- Reinhilde Vergucht (1999-2000) Veredeling en identificatie van sierappels. Hogeschool Gent, dep. Biotechnologische wetenschappen, Melle.
- Eveline Neyrinck (2000-2001). Studie van de genetische diversiteit van inheemse eiken aan de hand van microsatellietanalyses. KAHO Sint-Lieven, Gent.
- Sabine Van Glabeke (2001-2002). Diversiteitstudie van Wilde appel (*Malus sylvestris* (L.) Mill.) aan de hand van microsatellietmerkers en sequentie-analyse van chloroplastgenen. BME-CTL, Gent.
- Sophie Deschaumes (2002-2003). Studie van de genetische diversiteit binnen het genus Crataegus (Meidoorn): Voorkomen van soorten en hybriden in Vlaanderen. BME-CTL, Gent.