Patterns of genetic diversity and their underlying processes in a dominant subtropical tree *Castanopsis eyrei* at multiple scales

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der

Naturwissenschaftlichen Fakultät I
Biowissenschaften
der Martin-Luther-Universität Halle-Wittenberg

von

Miaomiao Shi
geboren am 30.09.1983 in Henan, China

Gutachter:

1. Dr. W. Durka
2. Prof. Dr. H. Bruelheide
3. Prof. Dr. R. Finkeldey

Halle (Saale), 21.12.2012

Table of contents

Summary

As a fundamental component of biodiversity, genetic diversity can represent the abilities of organisms to evolve and adapt to changing environmental conditions. An increasing number of studies have focused on revealing spatial and temporal genetic structures of populations and identifying the underling ecological and evolutionary processes. In this thesis, I focused on the evergreen tree species *Castanopsis eyrei*, which is dominant in subtropical forests in China, analysed its patterns of genetic diversity and structure at multiple scales and revealed underlying processes.

The first study I investigated small scale spatial genetic structures (SGS) of *C. eyrei* in adults and seedlings which were collected from six plots varying in individual density, elevation and species richness. No significant differences in genetic diversity were found between seedlings and adults. Significant SGS was detected in both adults and seedlings up to a maximum of 20 m, likely due to limited seed dispersal. The *Sp* statistic, quantifying the intensity of SGS, showed similar values between adults and seedlings, and no correlation with habitat conditions. The consistent SGS in seedlings and adults may be attributed to overlapping seed shadows caused by secondary dispersal by rodents and high density of adults and suggests that extensive pollen and seed dispersal prevents the build up of increased SGS in the seeding stage.

In the second study, I assessed the effects of elevation and successional stage on genetic diversity of *C. eyrei* controlling for neutrality of the microsatellite loci used. Diversity and differentiation within and among 24 populations from different elevations and successional stages were analysed by eight microsatellite loci. One locus was found to strongly deviate from a neutral model due to either divergent selection or hitchhiking with an unknown selected locus. Genetic variation increased with elevation for both the putatively selected locus and the neutral loci. Additionally, significant isolation by elevation was found for both neutral loci and the putatively selected locus. The results indicate higher gene flow among similar elevational levels than across different elevational levels and suggest a selective influence of elevation on the distribution of genetic diversity in *C. eyrei*.

Lastly, in order to reveal the effect of glacial ages on the current species distribution and genetic structure, I carried out a phylogeographic study. Climate modelling revealed a

potential distribution of *C. eyrei* in a narrow belt along the southern coastline during the Last Glacial Maximum (LGM). Nuclear microsatellites revealed two clusters corresponding to a split between western and eastern range, and a south-north decline of genetic variation. The eastern cluster harboured significantly higher nuclear genetic diversity. Populations were strongly differentiated at cpDNA with many fixations for different haplotypes, but mostly lacked a phylogeographic structure. Both data sets indicated higher genetic differentiation in the western than in the eastern cluster. The results provide evidence for at least two putative refugial regions during the LGM and a postglacial re-colonization from the South.

With this PhD thesis I'm able to evaluate the responses of genetic structure of *C. eyrei* to some biotic and abiotic factors and its historic dynamics. It provides insights into the dynamics of genetic structure over life stages, underlines the fact that both neutral and adaptive processes interact in determining population genetic structure, and likely represents a template for evolutionary history and phylogeography for wide-spread subtropical species in subtropical China since the LGM.

Zusammenfassung

Als ein wesentlicher Bestandteil der biologischen Vielfalt repräsentiert die genetische Diversität eines Organismus dessen Möglichkeit zur Evolution und zur Adaptation an sich ändernde Umweltbedingungen. Viele Studien haben die räumlichen und zeitlichen Muster genetischer Diversität und deren zugrundeliegenden ökologischen und evolutionären Prozesse untersucht. Hier bearbeite ich *Castanopsis eyrei*, eine in den subtropischen Wäldern Chinas dominante, immergrüne Baumart, analysiere genetische Diversität und Struktur auf unterschiedlichen Skalen und suche nach möglichen Einflussfaktoren.

Im der ersten Studie untersuche ich die räumliche-genetische Struktur innerhalb Populationen (SGS). Dazu wurden Proben von Adultpflanzen und Sämlingen aus sechs Plots gesammelt, welche sich in Individuendichte, Höhenlage und Artenreichtum unterscheiden. Genetische Diversität und SGS waren ähnlich zwischen den Generationen. Eine signifikante SGS wurde bis zu einer Maximaldistanz von 20 m detektiert, was vermutlich auf eine limitierte Samenausbreitung zurückzuführen ist. Die Intensität der SGS, quantifiziert durch die *sp*-Statistik, zeigte keine Korrelation mit den verschiedenen Habitatcharakteristika.
Die unerwartete Ähnlichkeit in der SGS zwischen Adulten und Sämlingen kann möglicherweise durch eine räumlich stark überlappende Samenausbreitung der Individuen durch zoochore Verschleppung der Samen sowie durch eine hohe Dichte der adulten Individuen erklärt werden.

In der zweiten Studie untersuche ich die Effekte von Höhenlage und Suksessionsstadium auf die genetische Diversität von *C. eyrei*. Unter Nutzung von acht Mikrosatellitenloci quantifiziere ich Diversität innerhalb von und Differenzierung zwischen 24 Populationen. Ein Lokus wich dabei von stark von neutralen Erwartungen ab, was durch divergente Selektion und/ oder genetisches ‚hitchhiking' mit einem unbekanntem, selektiertem Lokus erklärt werden kann. Für diesen Lokus, aber auch die neutralen Loci stieg die genetische Diversität der Populationen mit der Höhenlage an. Ebenso fand ich ein ‚Isolation-by-elevation'-Muster für neutrale und nicht-neutrale Variation, welches einerseits nahelegt, dass Genfluss entlang gleicher Höhenlagen stärker ist als zwischen verschiedenen Lagen, andererseits einen selektiven Einfluss der Höhenlage vermuten lässt.

In der letzten Studie präsentiere ich eine phylogeographische Arbeit, in welcher ich versuche die heutige, großskalige Verteilung genetischer Diversität durch nacheiszeitliche Prozesse zu erklären. Eine Modellierung der klimatischen Verhältnisse in der letzten Eiszeit (LGM) zeigt, dass zu der Zeit *C. eyrei* in einem schmalen Bereich entlang der südlichen Ostasiatischen Küste eine mögliche Verbreitung gefunden hat. In rezenten Proben, lässt sich die gefundene genetische Variation an nukleäre Mikrosatelliten am besten mit zwei genetischen Gruppen erklären, welche einerseits die westlichen und andererseits die östlichen Populationen umfassen. Dabei zeigte die östliche Gruppe eine signifikant höhere genetische Diversität als die westliche. Genetische Variation auf Chloroplastenebene zeigte eine starke Fixierung für verscheiden Haplotypen und damit eine hohe Differenzierung zwischen den Populationen. Beide Marker weisen jedoch auf eine stärkere genetische Differenzierung der westlichen Gruppe hin. Meine Ergebnisse lassen vermuten, dass die post-glaziale Wiederbesiedlung durch genetische Linien aus wenigstens zwei, südlichen Refugialgebieten erfolgte.

Mit meinen hier präsentierten Arbeiten untersuche ich den Einfluss historischer sowie einiger biotischer und abiotischer Faktoren sowie auf die genetische Strukture von *C. eyrei*. Sie ermöglichen Einblick in die Dynamik genetischer Struktur auf verschiedenen Skalen, unterstreichen die Bedeutung der Interaktion neutraler und adaptiver Prozesse für die Bildung genetischer Diversitätsmuster und sind möglicherweise exemplarisch für die nacheizeitliche Evolutionsgeschichte und Phylogeographie von weitverbreiteten Arten im subtropischen China.

Chapter 1 - General introduction

Genetic diversity

Genetic diversity, as a fundamental source of biodiversity, is defined as a measure to quantify the magnitude of genetic variability within a population (Hughes *et al.* 2008). It is the raw material for evolution and can represent the abilities of organisms to adapt to changing environmental conditions. Thus, genetic diversity is often adopted in studies of evolutionary and conservation biology (Petit *et al.* 1998; Yakimowski & Eckert 2007). Genetic variation within and among populations is shaped by the interplay of genetic drift, gene flow, mutation and natural selection, which may vary over life spans of populations, hence natural populations are genetically structured in space and time. Molecular markers have helped to reveal spatial and temporal genetic structures of populations during the last four decades. Studies have been increasingly focusing on the identification of ecological (Booth & Grime 2003; Wang *et al.* 2012; Yoshida *et al.* 2003) and evolutionary processes (Alleaume-Benharira *et al.* 2006; Garcia-Ramos & Kirkpatrick 1997) responsible for such genetic structure. However, it is well documented that population genetic structure depends on either biotic factors, e.g. life-history traits (Duminil *et al.* 2007), or abiotic factors, e.g. environmental conditions (Ohsawa & Ide 2008) and their historic dynamics, e.g. during glacial periods (Hewitt 2000).

Factors influencing genetic structure

Biotic factors

Generally, life-history traits of species, including mating system, breeding system, life form or mode of seed and pollen dispersal, can largely affect the distribution pattern of genetic diversity. For instance, mating system was indicated as the only factor significantly related with genetic structure for nuclear markers (Duminil *et al.* 2007). Indeed, outcrossing promotes the exchange of genes within and among populations, thus maintaining high genetic diversity and weakening genetic structure (Muir *et al.* 2004; Nettel *et al.* 2009). In contrast, selfing species are apt to fix alleles in one population, where genetic drift dominates the patterns, leading to low genetic diversity and strong genetic structure (Michalski & Durka 2007). Increasing the outcrossing rate is expected to increase heterozygosity and spatial mixing of genes, thus decreasing spatial genetic structure (Doligez *et al.* 1998). However, it can not be neglected that other traits, like breeding system, perenniality and pollination mode, are strongly correlated with mating system (Duminil *et al.* 2007; Michalski & Durka 2009),

thus, somewhat affecting the patterns of genetic structure.

In addition, since gene flow is one of the most important processes dominating genetic variation of populations (Austerlitz & Garnier-Géré 2003), the movement of genes within and between populations determines their spatial and temporal genetic variation and structure, which in turn influences their evolutionary potential. Theoretical studies indicated that frequent long distance dispersal can considerably increase genetic diversity within populations and decrease differentiation among populations (Bialozyt *et al.* 2006; Fayard *et al.* 2009). In plants, seed dispersal involves colonization of new areas and plays an important role in migration and expanding distribution, especially in the context of climate change (Shaw & Etterson 2012). As indicated, seed dispersal mode was significantly related with genetic structure typically for organelle markers (Duminil *et al.* 2007). Seeds primarily dispersed by birds, bats or monkeys caused a relatively weaker spatial genetic structure than did seed primarily dispersed by gravity or rodents (Dick *et al.* 2008). Besides seed dispersal, pollen dispersal also contributes to the gene pool of colonizing populations (Chen *et al.* 2008; Wang *et al.* 2011). Especially in outcrossing species, e.g. pollinated by wind or insects, extensive pollen flow can be expected, thus resulting in homogeneity of genes, i.e. weak genetic structure (Petit *et al.* 2005; Sutherland *et al.* 2010). Therefore, detecting dispersal capacity of seed and pollen of plant species is helpful for understanding population genetic structure and evaluating their evolutionary potential.

Besides, population density can indirectly affect genetic structure by its effect on outcrossing rates and patterns of gene dispersal (Born *et al.* 2008; Hanson *et al.* 2008). In wind-pollinated species, pollen flow among individuals in low-density populations is likely to be reduced, hence leading to more pronounced genetic structure than in high-density populations (Friedman & Barrett 2008), while in insect-pollinated species, populations with lower individual density and a higher number of co-flowering species are expected to result in less effective movement of pollinators among individuals and then stronger genetic structure (Zeng *et al.* 2012). However, still, inconsistent patterns indicated that distances covered by pollinating insects during foraging activities increased when local density of flowering trees decreased, facilitating the maintenance of genetic diversity (Born *et al.* 2008).

Furthermore, spatial genetic structure (SGS) within plant populations is also influenced by variation in demographic processes through space and time, e.g. life stages (Chung *et al.* 2003a; Ueno *et al.* 2002) and successional status (Chung *et al.* 2007). For example, some studies have found significant fine-scale genetic structure in seedlings because of limited seed dispersal which is greatly reduced or absent in adults due to thinning processes

(Chung *et al.* 2003a; Qi *et al.* 2011). Generally, in outcrosssing tree species, seeds fall around maternal trees. Although probably experiencing low rate of germination, seedlings are characterized by high density and a high level of relatedness of adjacent seedling individuals. Neverthesis, from seedlings to adults, a thinning process is underlying due to individual competition and predation of herbivore (Chybicki & Burczyk 2010b; Oddou-Muratorio *et al.* 2011), thus, leading to a low density of sibs. Therefore, decreasing SGS from seedlings to adults can be expected. However, increasing SGS was also indicated which may be correlated with founder effects and microenvironmental selection (Kalisz *et al.* 2001), whereas consistent patterns of SGS across different life stage was most likely the result of limited gene flow (Yao *et al.* 2011). Besides of life stages, population succession involves a number of potential demographic processes including the changes of environmental factors, such as illumination intensity, humidity and competition, which are recognized to influence not only levels of genetic variation but the spatial distribution of alleles and genotypes.

Abiotic factors

Among abiotic factors, environmental conditions, such as soil type, topology or elevation, play an important role in genetic structuring because they may affect phenology, population size or density and thus gene flow or genetic drift (Byars *et al.* 2009). Elevation is of particular importance, because elevations clines encompass a suite of environmental factors that are either physically linked with elevation (e.g. temperature) or instead correlated with it like land use. These factors can affect not only the neutral processes of gene flow and genetic drift (Kraj & Sztorc 2009), but also selective processes (Jump *et al.* 2006). Different patterns of genetic variation within populations varying along elevational gradients have been identified (Fig. 1.1) (Ohsawa & Ide 2008). First, mid-elevation populations may hold higher levels of diversity compared with both low and high elevation populations due to the optimal mid-elevation habitats following the central-marginal hypothesis (Herrera & Bazaga 2008). Second, low elevation populations may have highest diversity which decreases with elevation as a result of bottlenecks occurring throughout upward range expansion (Quiroga & Premoli 2007). Third, highest genetic diversity was found at high elevations which was attributed to various reasons like decreased human disturbance and/or historical downward range shifts due to climate change, and adaptation (Gämperle & Schneller 2002; Ohsawa & Ide 2008). Lastly, genetic variation also has been found to stay rather constant along a given elevational gradient due to extensive gene flow (Truong *et al.* 2007a). Overall, these inconsistent patterns

highlight the predominant role of biogeographic history in determing patterns of genetic variation along elevational gradients.

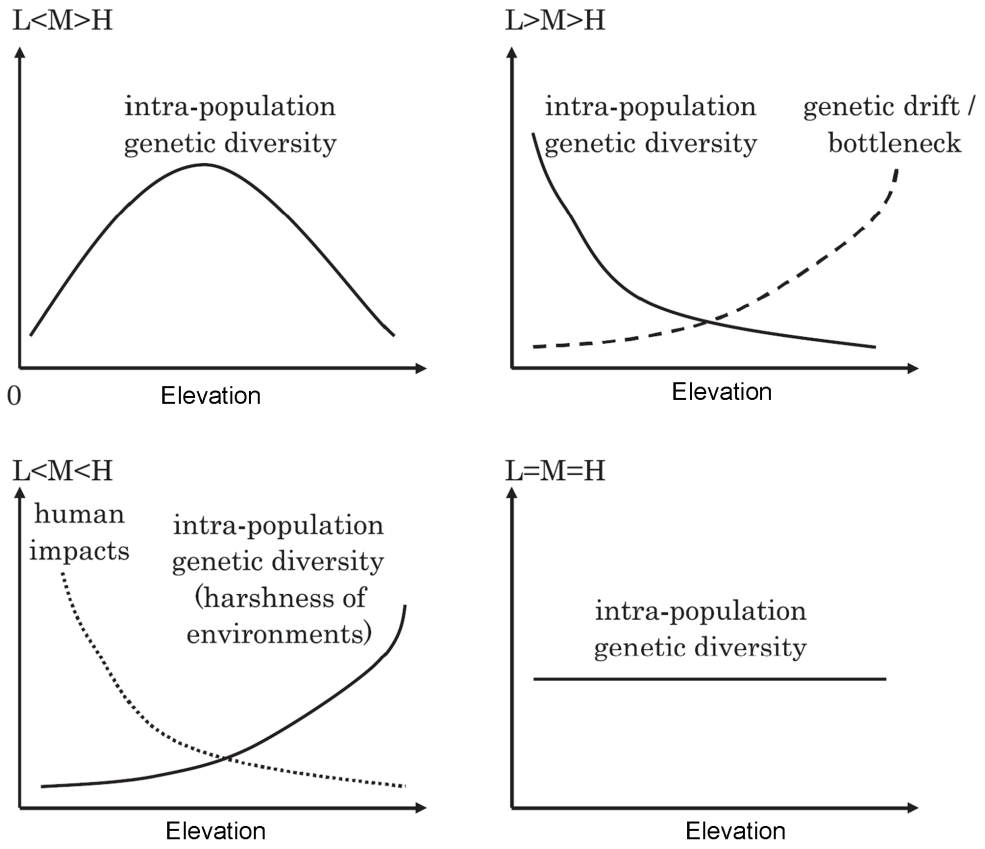

Fig. 1.1 Four patterns of genetic variation along changing elevations in plant species. L, M and H represent low, middle and high elevations, respectively. (Modified after Ohsawa and Ide 2008)

Phylogeographic history

The present distribution of plant populations is determined not only by the current environment, but also by past events like Pleistocene climate cycles (Avise 2000). Repeated drastic climate changes of Quaternary have resulted in repeated contraction-expansion processes of many plants and profoundly shaped their current genetic structure and phylogeographic patterns (Hewitt 2004). During glacial expansions, most species would have retreated into more southerly or warmer lowland areas, namely refugia. When the ice retreated, a reverse course occurred with species recolonizing from south to north, or from lowland to higher elevations. Generally, populations located in refugia are old and may maintain high genetic diversity, which will be gradually lost due to genetic drift and environmental stresses (Hampe & Petit 2005), while the recolonizing populations are characterized by low genetic diversity and high differentiation because of founder effect and

rare long-distance dispersal events (Hampe & Petit 2005). Therefore, glacial periods have been imprinted in population genetic structure which, in return, can draw out the glacial history. A number of phylogeographic studies have documented complex historic processes and the effects of climate changes on species distribution and population genetic structure in the last 20 years, especially in Europe and North America (Avise 2000; Heuertz *et al.* 2004a; Taberlet *et al.* 1998).

The various factors may have a complex influence on spatial and temporal distribution of genetic diversity in plants. Thus, it is clear that elucidating the responses of populations to those factors requires a multi-scale study, such as at local, ecological scales as well as large, species-range scales which coincide with time scales during which patterns of genetic variation emerge.

Molecular markers

Molecular markers have been extensively used to identify the effect of life history traits, environmental factors and phylogeographic history on the genetic structure of plant populations (Hamrick & Godt 1996; Nybom 2004). Among the various markers, because of their high mutation rates and high levels of polymorphism, biparentally inherited nuclear microsatellites are highly adopted to evaluate the level of genetic variation in natural populations, reveal the effects of genetic drift and recent bottlenecks (Selkoe & Toonen 2006), and even be involved to phylogeographic studies by identifying the gene pools. However, for such markers with high levels of allelic variability, genetic differentiations (F_{ST}) are always underestimated (Hedrick 2005). Thus, the standardized parameter e.g. F'_{ST} is increasingly used to evaluate the level of population differentiation (Hedrick 2005). Generally, microsatellites are assumed to represent neutral markers as they are usually located in non-coding regions. Consequently, patterns of differentiation among populations revealed by microsatellites are almost exclusively interpreted as genetic drift and gene flow. However, increasing studies indicated the presence of non-neutral microsatellite loci (Casa *et al.* 2005; Nielsen *et al.* 2006). Besides, chloroplast DNA (cpDNA), which is inherited maternally, is also commonly used especially in phylogeographic studies. cpDNA does not experience recombination and represents only a single gene genealogy, thus reflecting exclusive gene flow by seeds. More and more studies used both nuclear and organelle markers to detect a relatively complete description of population genetic structure (Bai *et al.* 2010; Hu *et al.* 2010; Wang & Ge 2006).

Subtropical China

Subtropical China (ca. 22°N-30/33°N), as one of the most prominent hotspots of terrestrial biodiversity, represents an exceptional high level of biodiversity, higher than any other region at the similar latitude (Fig. 1.2), especially for tree and shrub species. In this region, evergreen broadleaved forests are the representative vegetation and dominated mainly by species of Fagaceae and Lauraceae. While large parts of Eurasia have been glaciated, subtropical China was never covered by large ice sheets during the Last Glacial Maximum (LGM) (Fig. 1.3) (Hewitt 2000). Therefore subtropical China is considered to be one of the most important refugial regions for lineages that evolved prior to the late Tertiary and Quaternary glaciations (Axelrod *et al.* 1996). However, contrary to its importance in evolution and high level of biodiversity, phylogeographic studies in subtropical China are surprisingly limited (Qiu *et al.* 2011). Most of the studies available have focused on endangered species with narrow distribution range (Qiu *et al.* 2009; Wang & Ge 2006), while common species with broad distribution range, which may reflect the general pattern of effect of glacial periods, however, are less studied.

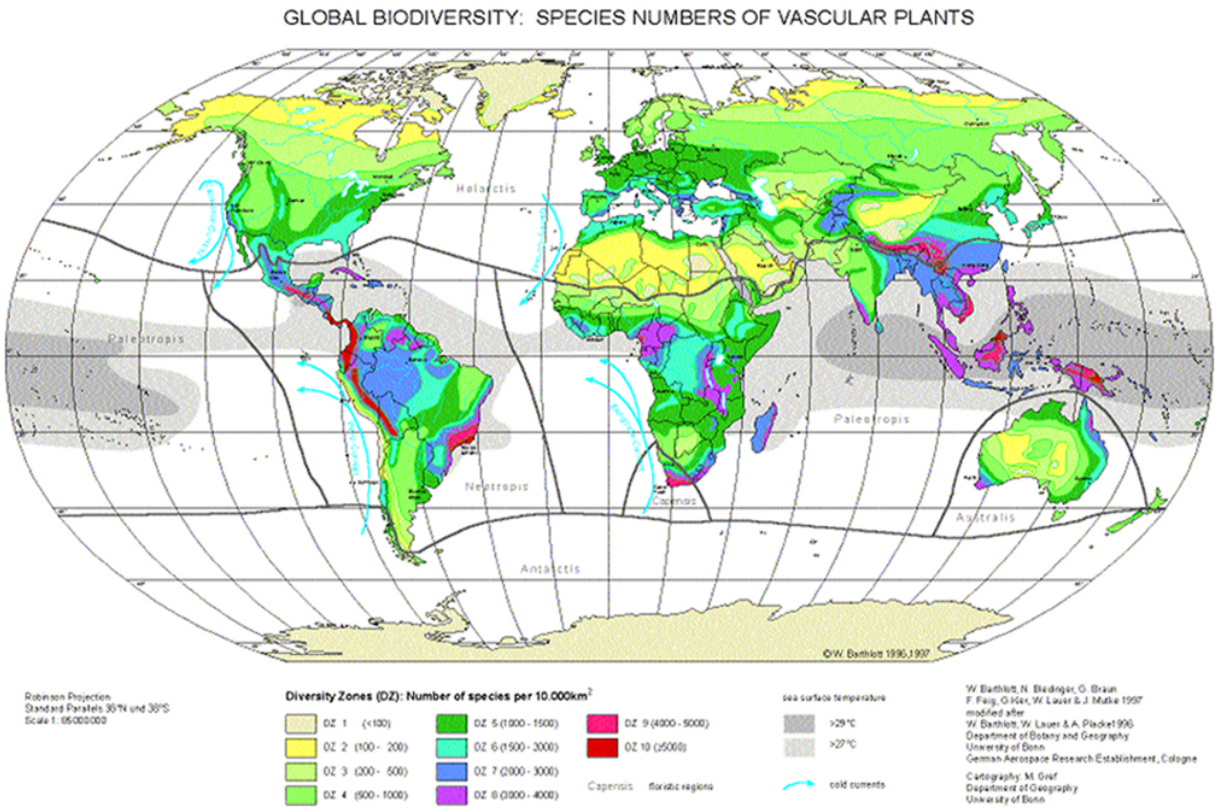

Fig. 1.2 Distribution of global biodiversity in vascular plant species (Barthlott *et al.* 2005).

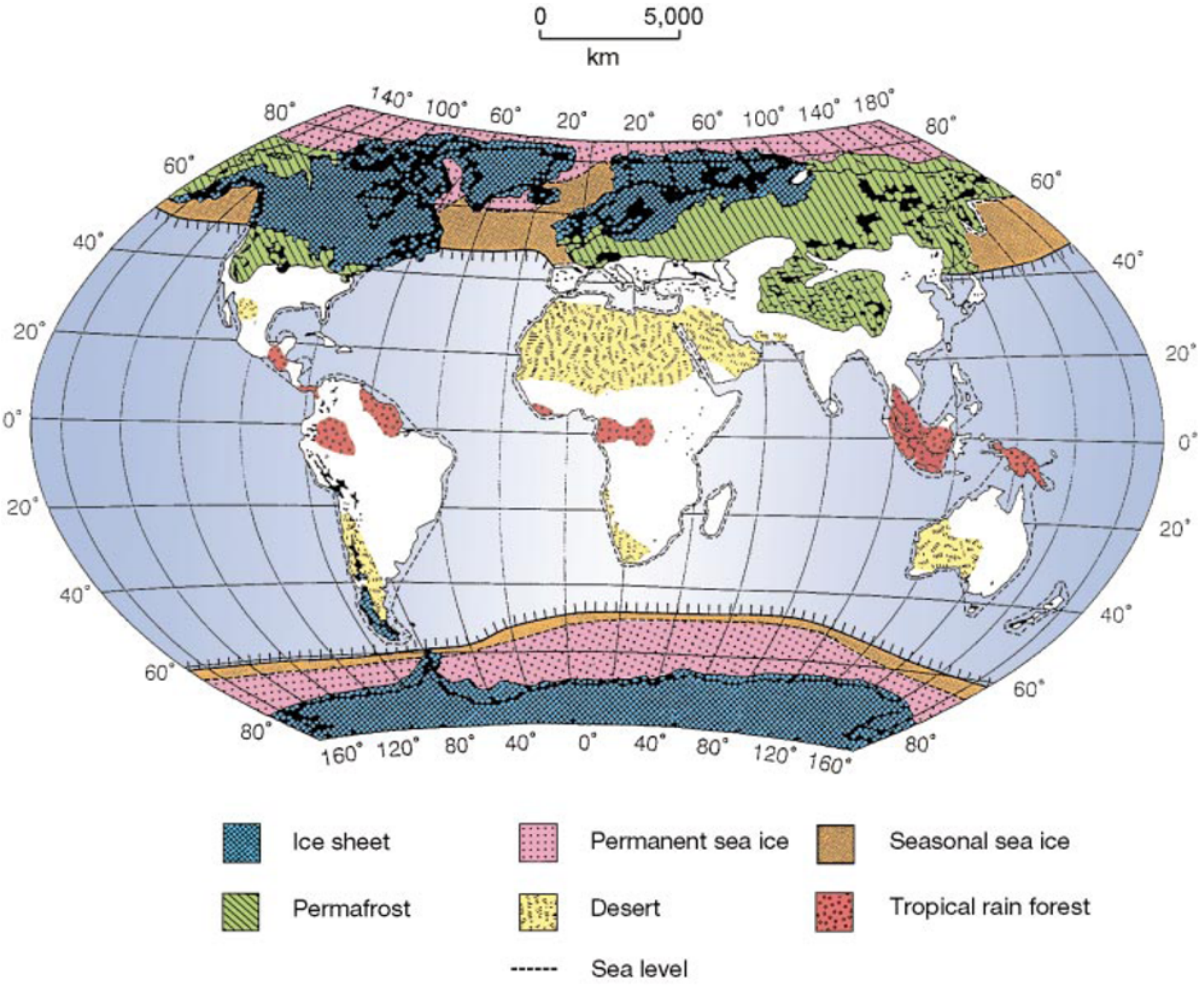

Fig. 1.3 The maximum extent of ice and permafrost at the end of the last ice age 20,000 years BP (Hewitt 2000).

Study species

In this thesis, the focal species is *Castanopsis eyrei* (Champion ex Bentham) Tutcher (Fagaceae), one of the dominant tree species in late successional forests of evergreen broadleaved forests in subtropical China (Fig. 1.4). As a common species, it represents a template to reveal the general patterns of genetic structure and underlying processes at multiple scales. *Castanopsis eyrei* is monoecious, and wind- and insect- pollinated, although the extent to which these two types of pollination contribute is not clear, but suggesting that pollen dispersal is likely to be extensive. Acorn seeds are principally dispersed by gravity, thus limited seed dispersal is expected, although small rodents can transport the acorn seeds and serve as secondary seed dispersers (Li & Jin 2006). The geographical distribution of *C. eyrei* is mainly restricted to the area south of Yangtze River in mainland China and Taiwan, where humid subtropical climate without dry seasons dominates, typically characterized by hot, humid summers and mild winters. It occurs from 300 m to 1700 m a.s.l. (Huang *et al.*

1999), frequently in Southeast China and more scattered in the Southwest due to unsuitable karst habitat (Fig. 1.4).

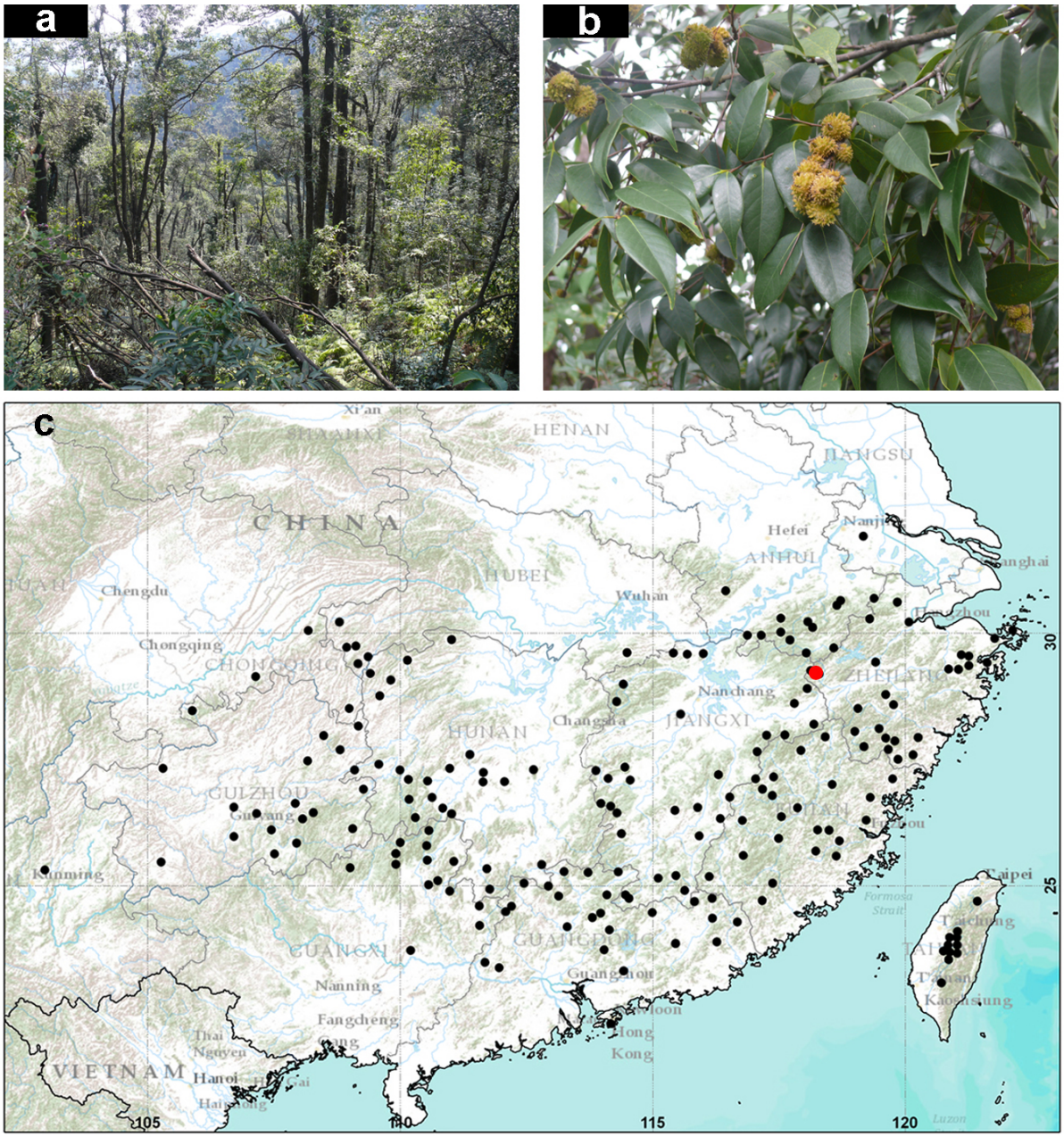

Fig. 1.4 (a) Forests dominated by *Castanopsis eyrei*, in Gutianshan Nature Reserve, Zhengjiang, China; (b) The fruits of *C. eyrei*. (photos from W. Durka); (c) Distribution of *C. eyrei* (map from E. Welk). The red dot indicates the location of Gutianshan Nature Reserve where the Comparative Study Plots (CSPs) of BEF-China project were set up.

Study region

This study was carried out in different scales, i.e. fine-, regional-, and large-scale. Large scale covered the whole distribution range of *C. eyrei*, while both fine- and regional-scales were studied in Gutianshan National Nature Reserve (GNNR) in Zhejiang Province, southeast China (29º8'18" - 29º17'29" N, 118º2'14" - 118º11'12" E). There, in order to focus on biodiversity and ecosystem functioning (BEF), as an observational experiment, BEF-China project set up 27 comparative study plots (CSPs) with a size of 30 m × 30 m per plot in 2008 (Fig. 1.5), varying in elevation, age of the tree layer and species richness (Bruelheide *et al.* 2011).

The Reserve was initially established as a National Forest Reserve in 1975 and became a National Nature Reserve in 2001. It consists of species-rich broad-leaved forests including old growth forest and successional stages that developed after cease of human use in 1975. Because of the steep terrain of the entire reserve with slopes frequently exceeding 30°, the GNNR was only marginally usable for agricultural activities, and thus, an exceptionally intact forest cover has been preserved especially on the steep slopes. Totally, 1426 seed-plant species have been recorded from 149 families and 648 genera with approximately similar proportions of deciduous and evergreen species in species numbers (Lou & Jin 2000). Within the tree layer, evergreen species dominate in abundance of individuals (Yu *et al.* 2001).

The GNNR has an area of approximately 81 km^2 with elevations ranging from 250 to 1258 m a.s.l.. The climate in GNNR is warm-temperate with an annual average temperature of 15.1 °C. The mean minimal and maximal temperatures are -6.8 °C (January) and 38.1 °C (July), respectively. The mean annual precipitation sums up to approximately 2000 mm with a wet season occurring from March to September and a short dry season in November and December.

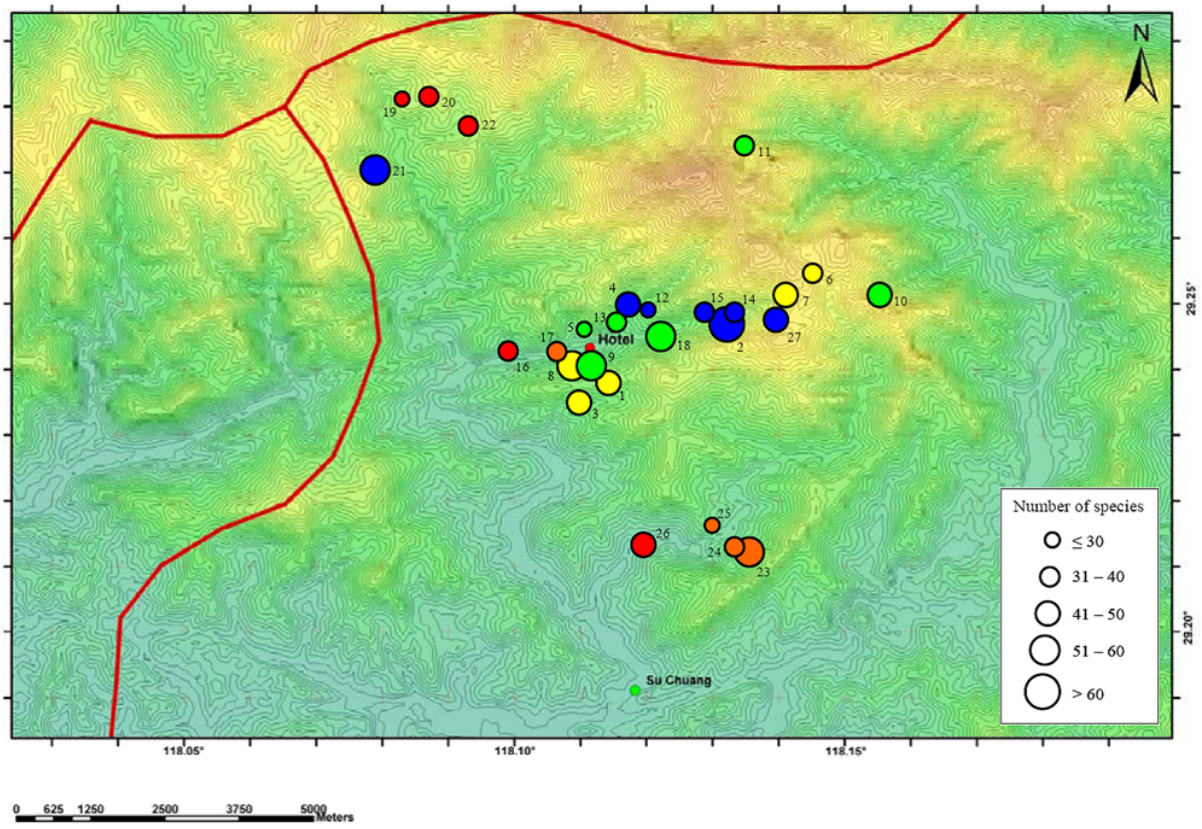

Fig. 1.5 Locations of the 27 comparative study plots (CSPs) established by BEF-China project in Gutianshan Nature Reserve (Böhnke 2011). Size of the circles corresponds to the number of species in the plots (see the legend in the corner) and filled colours indicates the successional stages referring to the average age of the tree layer: red, < 20 yrs; orange, < 40 yrs; yellow, < 60 yrs; green, < 80 yrs; blue, > 80 yrs.

Objectives and aims of this thesis

In order to detect the complex effects of several important biotic and abiotic factors on the genetic structure of a common species *C. eyrei*, I evaluated patterns of genetic diversity and underlying processes at multiple scales.

Chapter 2 investigates the fine-spatial genetic structure of *C. eyrei*. In this study, two generations, i.e. seedlings and adults, were sampled from six plots with different environmental (e.g. elevation) and ecological (density, species diversity) conditions. In total, more than 2500 individuals were genotyped by nine highly polymorphic microsatellite markers. Based on multi-loci genotypes and coordinates of seedlings and adults, we analysed spatial genetic structure (SGS) indirectly by spatial autocorrelation and directly by spatially explicit mating models, and compared genetic variation, SGS and patterns of gene flow among plots and between the two generations. Considering that fact that seeds tend to

aggregate around the mother trees, a stronger small scale genetic structure in seedlings is expected than in adults.

Chapter 3 looks at the influence of elevation or successional stage on genetic variation of *C. eyrei* at a regional scale. I sampled 24 representative sites of 30 × 30 m which were spread across all successional stages and the local elevational range of the species (251 - 920 m) and evaluated the level of genetic variation in each site controlling for neutrality of the microsatellite loci used. I correlated genetic diversity (allelic richness: A_R) with elevation and successional stage and tested for a pattern of isolation by elevation by comparing pairwise genetic differentiation and pairwise elevational distances.

Chapter 4 concerns the phylogeography of *C. eyrei* and in particular the impact of the Last Glacial Maximum on the distribution of genetic variation across the whole distribution range. First, climatic envelope modelling was used combined with downscaled high-resolution estimates of LGM climate parameters to predict the distribution of *C. eyrei* during LGM. In addition, microsatellite and cpDNA data were obtained from 31 populations throughout its range. For cpDNA sequence data, relationships among haplotypes were depicted in a statistical parsimony network. Based on microsatellites, I analysed genetic clustering, and patterns of variation within and among populations. I compared the genetic variation between clusters and detected a correlation between genetic diversity and latitude. I hypothesize that there are multiple refugia located in South China and genetic variation decreases with increasing latitude because of northward recolonizing route.

Chapter 5 summarizes the main results of the preceding chapters, and presents an overall discussion and comprehensive conclusions.

Chapter 2 - Seedlings and adults do not differ in small scale genetic structure in *Castanopsis eyrei*

Miaomiao Shi, Stefan Michalski, Xueqin Zeng and Walter Durka

Manuscript

Abstract

Castanopsis eyrei is one of the dominant tree species in evergreen broadleaved forests in subtropical China. Comparison of fine-scale spatial genetic structure (SGS) over life stages can provide insights into the dynamics of natural populations and ecological and evolutionary processes within and among populations. In this study, 910 adults and 1594 seedlings of *C. eyrei* were collected from six plots varing in individual density and elevation in Gutianshan Nature Reserve. All the individuals were mapped and genotyped using nine microsatellite loci to investigate genetic diversity and SGS. No significant differnences in genetic diversity were found between seedlings and adults. However, significant SGS was detected in both adults and seedlings, although the extent was weak, as expected for outcrossing tree species, up to a maximum of 20 m, suggesting limited seed dispersal. The *Sp* statistic, quantifying the intensity of SGS, showed similar values between adults (0.0083) and seedlings (0.0109), and no correlations with environmental conditions, especially density and elevation, indicating its independence. The consistent SGS in seedlings and adults may be attributed to overlapping seed shadows caused by secondary dispersal by rodents and high density of adults. In addition, seed and pollen dispersal kernel was estimated using a spatially explicit mating model, indicating a limited seed dispersal distance (mean = 21.0 m) in line with that deduced with SGS. Pollen dispersal distance reached a maximum of 132.3 m, however, large proportions of pollen came from outside the study sites (46% < m_p < 95%), indicating extensive pollen flow.

Keywords：Dispersal kernel, life stages, microsatellites, spatial genetic structure, tree.

Introduction

One of the major objectives in plant biology is to understand the effects of demographic and genetic processes on the amount and spatial distribution of genetic variation within and among natural populations. Especially, fine-scale spatial genetic structure (SGS) is drawing more and more attention in population ecology and evolutionary biology (Dick 2008; Hardy *et al.* 2006; Vekemans & Hardy 2004). The extent of SGS allows to estimate the extent of gene dispersal within populations (Rousset 2000) and indirectly quantify the dispersal distance of the species (Oddou-Muratorio *et al.* 2010). Therefore, SGS is important for the conservation and management in natural populations, and for understanding the capacities of local adaptation and colonization into new areas in the context of rapid climate change (Oddou-Muratorio *et al.* 2010).

SGS is dependant on a number of biotic and abiotic factors. Since gene flow largely influences the distribution of genetic variation, self-compatible species or those with poorly dispersed seeds and pollen tend to exhibit strong SGS (Vekemans & Hardy 2004). With increasing outcrossing rate an increase of heterozygosity and spatial mixing of genes is expected, thus decreasing SGS (Doligez *et al.* 1998). In addition, population density can indirectly affect SGS via effects on dispersal distance of seed or pollen (Born *et al.* 2008; Vekemans & Hardy 2004), although inconsistent patterns exist (Born *et al.* 2008; Hanson *et al.* 2008). For example, pronounced SGS was found at lower individual density probably due to less effective movement of pollinators among individuals (Zeng *et al.* 2012). Furthermore, abiotic conditions, e.g. elevation, may influence the distribution of genetic variation. Along an elevational gradient, environmental conditions, population density and successional stages can rapidly change which likely affects genetic structure (Ohsawa & Ide 2008; Shi *et al.* 2011a). Also, anthropogenic processes such as habitat fragmentation can have a great influence on SGS. The isolation of small populations can result in restricted seed dispersal resulting in increasing relatedness of adjacent individuals and consequently pronounced SGS (Hanson *et al.* 2008; Sebbenn *et al.* 2011; Wang *et al.* 2011).

Generally, tree species display weaker SGS than herbaceous species (Vekemans & Hardy 2004), although those with restricted seed and pollen dispersal still show pronounced SGS (Cavers *et al.* 2005; Hardy *et al.* 2006). Especially, in outcrossing tree species, high gene flow within populations is expected to result in relatively low levels of SGS (Leonardi & Menozzi 1996; Streiff *et al.* 1998). For example, significant SGS is rarely detectable beyond 30-40 m in the common ash *Fraxinus excelsior* (Heuertz *et al.* 2003), oaks *Quercus petraea* and *Q. robur* (Streiff *et al.* 1998), and beech *Fagus sylvatica* and *F. crenata* (Chybicki *et al.*

2009; Oddou-Muratorio *et al.* 2010; Oddou-Muratorio *et al.* 2011), as well as in *Castanopsis sclerophylla* (Wang *et al.* 2011). For a given tree species with seed dispersal primarily by gravity, most of the seeds are followed by a short-distance dispersal and colonize close to the mother tree (Asuka *et al.* 2005; Nakanishi *et al.* 2009; Sebbenn *et al.* 2011). Thus, seedlings tend to grow around their mother tree and exist in much higher density leading to a high level of inbreeding rate and relatedness of adjacent seedling individuals, and then a low level of genetic diversity and a significant SGS. In addition, during an individual lifetime the step from seedling to an adult is characterized in particular by a massive mortality due to competition, predation and selection during eatablishment (Chybicki & Burczyk 2010a; Oddou-Muratorio *et al.* 2011), typically with only one seed in a million surviving as a reproductive tree (Petit & Hampe 2006). Such a thinning process reduces the density of half-sibs and weakens the degree of allele aggregation (Chung *et al.* 2002; Chung *et al.* 2003b). Thus, weaker or even absent genetic structure is expected to be found in adults than seedlings. Overall, studying established seedlings can provide important insights into the dynamics of natural populations (Burczyk & Chybicki 2004) and allow comparisons of dispersal distances and patterns of seeds and pollen (Chybicki & Burczyk 2010b; Oddou-Muratorio *et al.* 2010; Oddou-Muratorio *et al.* 2011; Sebbenn *et al.* 2011).

Historical gene flow can be indirectly studied based on population genetic structure or spatial genetic autocorrelation analysis, both targetting the extant population (Hardy *et al.* 2006; Rousset 1997; Vekemans & Hardy 2004). In contrast, direct estimation of current gene flow necessitates the analysis of parents and offspring, such as parentage analysis, which largely depends on the polymorphism of genetic markers and assigns each offspring to specific parent trees. In fact, because of insufficient polymorphism and presence of null alleles or missing data, precise parantage assignment is often impossible (Chybicki & Burczyk 2010a). An alternative way for detecting dispersal patterns uses spatially explicit mating models, which can estimate, among others, immigration rate of pollen and seed from outside, selfing rate and dispersal distance of pollen or seed, without complete parentage assignment (Burczyk *et al.* 2006). The use of both indirect and direct methods can reveal historical and contemporary pattern of pollen and seed dispersal.

In subtropical China, evergreen broadleaved forests are the zonal vegetation which is dominated mainly by Fagaceae and Lauraceae. Among them, *Castanopsis eyrei* (Champion ex Bentham) Tutcher (Fagaceae) is locally dominant in late successional forests. The species is monoecious, and wind- and insect- pollinated, although the extent to which these two types of pollination contribute is not clear, but suggesting that pollen dispersal is likely to be

extensive. Fruits are principally dispersed by gravity, thus limited seed dispersal is expected, although small rodents can transport the acorn seeds and serve as secondary seed dispersers (Li & Jin 2006). In a previous study low levels of genetic differentiation among populations were detected (Shi *et al.* 2011a), indicating high gene flow most probably mediated by pollen dispersal. However, detailed patterns of seed and pollen dispersal on a small scale are missing.

In this paper, we study of fine-scale spatial genetic structure of *C. eyrei* in a continuous subtropical forest in China. Two generations, i.e. seedlings and adults, were sampled from six plots with different environmental (e.g. elevation) and ecological (density, species diversity) conditions. Based on genotypes of microsatellite markers we analysed SGS indirectly by spatial autocorrelation and directly by spatially explicit mating models, and compared genetic variation, SGS and patterns of gene flow among plots and between the two generations. Specificly, the following questions were investiged: (i) Do adults and seedlings differ in genetic variation or SGS? (ii) Do habitat conditions influence the patterns of SGS? (iii) What is the gene dispersal distance of pollen and seeds?

Materials and methods

Study area and populations

Our study was carried out in Gutianshan National Nature Reserve (NNR) located in the west of Zhejiang province of China (29º08′18′′-29º17′29′′N, 118º02′14′′-118º11′12′′E). The NNR mainly consists of species-rich evergreen broad-leaved subtropical forests including old growth forests and successional stages that developed after cease of human use in 1975. *Castanopsis eyrei* is distributed continuously in this area (Xu *et al.* 2005) and is the dominant tree species in late successional stages.

We established six plots of 30 × 30 m at different elevations (Table 2.1), corresponding to comparative study plots reported in Bruelheide et al. (2011), where elevation, successional stage, species richness and species diversity of all woody species for each plot were recorded. In each plot, we sampled two generation cohorts: adults, i.e. trees >1cm diameter at breast height (DBH) and seedlings (<1cm DBH), forming 12 subpopulations (Fig. 2.S1). For seedlings, we collected all present in the plots varying from 110 to 422 individuals (total N = 1594). Their heights ranged from 1 to 416 cm (median = 14 cm). For adults, beside all trees within plots, additional samples were collected outside of the plots in a 40 m buffer in order to collect more than 100 individuals which can provide a close approximation of real levels of SGS (Cavers *et al.* 2005), thus resulting in 113 to 191 trees

(total N = 910). The diameter at breast height (DBH) of the trees ranged from 1.31 to 109.18 cm (mean = 18.65 cm). We mapped all the individuals and determined the number of individuals in the plot as the density of each cohort. All the leaf samples were dried and preserved in silica gel for further analysis.

Table 2.1 Habitat parameters of study plots and samples of *Castanopsis eyrei*. Plot numbers according to (Bruelheide *et al.* 2011).

Plot	Latitude (°N)	Longitude (°E)	No. of samples		Elevation	Successional Stage [b]	Inclination	Species richness	Species diversity
			Adults (in plot [a])	Seedlings	(m a.s.l)				
CSP6	29.25497	118.14747	113 (38)	110	880	3	24	39	22.3
CSP13	29.24630	118.12190	173 (42)	288	402	5	25	32	24.8
CSP14	29.24944	118.13518	120 (36)	422	639	5	28	38	24.2
CSP15	29.24917	118.13106	191 (35)	260	618	5	44	39	26.3
CSP21	29.27059	118.08084	165 (23)	363	566	5	40	56	32.7
CSP27	29.24709	118.13605	148 (55)	151	665	5	33	46	25.5

[a] number of individuals in the CSP plot 30 × 30 m.

[b] successional stage: 3 = intermediate forest (< 60 years); 5 = old forest (≥ 80 years) (Bruelheide *et al.* 2011)

Genotyping

Total genomic DNA was extracted from leaf samples using the DNeasy 96 plant extraction kit (QIAGEN) following manufacturers instructions. Nine polymorphic microsatellite loci were used for genotyping. Among them, five loci: Ccu102F36, Ccu16H15 (redesigned reverse primer: gaaattgagttgggttagttccac), Ccu62F15 (redesigned forward primer: catcagcccaattgatac agg), Ccu90T17 and Ccu97H18 were developed for *C. cuspidata* var. *sieboldii* (Ueno *et al.* 2000, 2003), two (Cch13 and Cch15) for *C. chinensis* Hance (Huang *et al.* 2009), while the other two: ca-sc2 (F: ccagctaatgtggactatc; R: gagctaagacaatgcagaag) and cy-gl 7 (F: atggttgaa gaggccggtg; R: ttcgccagcgcgttatttg) were newly designed based on the sequences of *C. sclerophylla* (GenBank accession no. HQ265410, (Shi *et al.* 2011b)) and *Quercus myrsinifolia* (accession no. Y10723, (Isagi & Suhandono 1997)). Multiplex polymerase chain reactions (PCRs) were performed following the programme reported in Shi et al. (2011a). PCR products were separated on an ABI 3130 genetic analyzer (Applied Biosystems, Foster City, California, USA) with internal size standard GeneScanTM 500 LIZ. Individuals were genotyped using GeneMapper version 3.7 (Applied Biosystems).

Population genetic analysis

To prove the neutrality of loci, we tested the nine loci for outliers, i.e. markers potentially under selection, using the program FDIST (Beaumont & Nichols 1996). A null distribution of F_{ST} values expected from a neutral model is generated and quantile limits are calculated. Loci outside a 99% confidence interval are regarded as potentially under selection. Because most of wind pollinated trees in Fagaceae are characterized by outcrossing, we assumed the populations of *C. eyrei* to be in Hardy-Weinberg equilibrium (HWE). Under this assumption, we checked the data for the presence of null alleles using the software MICRO-CHECKER v.2.2.3 (Van Oosterhout *et al.* 2004), since trans-specific amplification of microsatellites often results in null alleles. Within each subpopulation, we calculated the number of alleles per locus (A), allelic richness (A_R, correcting for sample size by rarefaction for a minimum sample size of 110), expected heterozygosity (H_e) and the fixation index (F_{IS}). We tested for significant differences in these descriptors between adults and seedlings by permutation tests (n = 1000). Genetic differentiation among subpopulations within cohorts and between cohorts within plots were estimated as F_{ST} (Weir & Cockerham 1984) and F'_{ST} (Hedrick 2005), a standardized parameter of genetic differentiation, as F_{ST} is likely to underestimate genetic differentiation between populations for markers with high levels of allelic variability. The necessary parameter $F_{ST\,max}$ for standardization was calculated after recoding the data using RECODEDATA (Meirmans 2006). In order to test for an isolation by distance pattern among subpopulations in each cohort, matrices of pairwise log transformed geographical distances and genetic distances ($F_{ST}/(1-F_{ST})$) were correlated and the significance was calculated by a Mantel test with 999 randomizations using Package "ecodist" (Goslee & Urban 2007) in R. If not mentioned otherwise, all above parameters were analysed with FSTAT v. 2.9.3.2 (Goudet 1995).

Fine-scale SGS and estimation of gene dispersal

Fine-scale spatial genetic structure (SGS) within subpopulations was assessed by a spatial autocorrelation approach using SPAGeDi v.1.3 (Hardy & Vekemans 2002). The kinship coefficients (F_{ij}) (Loiselle *et al.* 1995) were computed in a set of distance classes: 5, 10, 15, 20, 30, 40 and 80 m. Standard errors were estimated by Jackknifing across loci. Significance of mean F_{ij} per class was tested by random permutation with 1000 repetitions. We used the *Sp* statistic (Vekemans & Hardy 2004) to quantify the intensity of SGS for each subpopulation following the formula $Sp = -blog/(1-F_{(1)})$, where *blog* is the slope of the linear regression for kinship coefficient on the logarithm of geographical distance, and $F_{(1)}$ is the mean kinship

coefficient between individuals in the first distance class. The significance of *blog* was also test with 1000 permutations. We estimated the maximum extent of SGS extent as the maximum distance at which significantly positive kinship coefficients were detected before they turn into negative values (Jump *et al.* 2012).

To compare the SGS patterns of adult and seedling cohorts we performed a heterogeneity test (Smouse *et al.* 2008) using GENALEX v.6.3 (Peakall & Smouse 2006), which compares the observed differences among spatial autocorrelation analyses with permutated ones from pooled data. The criteria ω and t^2 were used to assess the divergence in pairs of spatial autocorrelation analyses as a whole and in each distance class, respectively (Smouse *et al.* 2008). The heterogeneity test was restricted to the distance range covered by the seedlings sampled. We use 999 repetitions for permutation tests and bootstrapping.

Under the assumption that observed SGS is representative of an isolation by distance pattern at dispersal-drift equilibrium, effective gene dispersal distance (σ) can be estimated from the *Sp* statistic, which is expected to equal $1/4\pi D_e\sigma^2$ in two-dimensional space in a restricted range of σ to 20σ (Vekemans & Hardy 2004), where D_e is the effective population density. We used $D_e = D/4$, as it is the best approximation for tree species (Hardy *et al.* 2006), where D is the census population density. We computed the gene dispersal distance for adults and seedlings separately in each plot and also for a combined data set with SPAGeDi v.1.3. The analysis iteratively determines σ, however, the algorithm does not converge necessarily, depending on the strength of SGS, in which case σ can not be determined. When no convergence was achieved, we also set D_e to $D/2$ and $D/10$, which, however, did not improve the results.

Effect of habitat conditions on SGS

To test the effects of habitat conditions on fine-scale within-subpopulation SGS as quantified by *Sp* for each cohort, we performed a multiple linear regression between *Sp* value and six predictors: density, elevation, successional stage, inclination, species richness and woody species diversity. The analysis was done using the `step` function (Venables & Ripley 2002) with R 2.8.1.

Estimation of the seed and pollen dispersal kernels

We used a spatially explicit mating model (SEMM) developed by (Burczyk *et al.* 2006) to estimate the range of seed and pollen dispersal. The SEMM considers that each individual can be mothered either by a tree located outside the study site due to seed immigration (with

probability m_s) or by a tree within the site. In the latter case, there are three possibilities for paternity of an offspring: self-pollination with probability s, pollination by a male tree outside the site with probability m_p, and pollination by a parent within the study site with probability $1 - s - m_p$. A maximum likelihood approach is used to simultaneously estimate parameters of the dispersal kernel, selfing rate and dispersal distances by the software NM+ (Chybicki & Burczyk 2010a), which accounts for null alleles and missing data.. We used an exponential-power dispersal kernel characterized by the parameters a and b (Clark 1998):

$$P(r) = \frac{b}{2\pi a^2 \Gamma(\frac{2}{b})} \exp\left\{-(\frac{r}{a})^b\right\}$$

where r is pollination distance, Γ is the classically defined gamma function, a is a scale parameter, and b is the shape parameter affecting the tail of dispersal function (Austerlitz *et al.* 2004). When $b = 1$, the above equation degenerates to the exponential distribution. When $b < 1$, the dispersal kernel is fat-tailed, i.e. the long-range decrease is slow, indicating more long-distance dispersal events. Conversely, when $b > 1$, the dispersal is thin tailed, with a rapid decrease of the dispersal function, implying few long-distance dispersal events (Austerlitz *et al.* 2004). We adjusted the null alleles and estimated the best initial values of all the parameters individually by setting an initial value of one parameter and leaving the others not estimated (a constant). If a given value leads to the highest likelihood, this implies that such value is a better initial point.

Results

Genetic diversity and differentiation

The neutrality test performed with FDIST revealed that all observed mean F_{ST} values calculated from all loci remained within the 99% confidence interval ($P > 0.26$), indicating their neutrality. Except for ca-sc2 and Ccu16H15, all loci showed signs of null alleles with average frequency between 0.38% and 17.24% (mean = 7.65%). A high level of polymorphism was present. The number of alleles ranged from 9 to 37 among loci (mean = 20). The average fixation index (F_{IS}) differed significantly from Hardy-Weinberg expectations in all subpopulations (Table 2.2), likely due to the effect of Null-alleles. However, no significant difference in F_{IS} was found between adults and seedlings ($P = 0.23$). All the six seedling subpopulations harboured similar levels of genetic variation to the adults, showing no significant difference in allelic richness (A_R, $P = 0.98$) and expected heterozygosity (H_e, $P = 0.48$) (Table 2.2).

Values of overall genetic differentiation for both seedling (F_{ST} = 0.043) and adult (F_{ST} = 0.030) subpopulations were low, but significant ($P < 0.05$). The standardized F'_{ST} values were 0.140 and 0.100, respectively, for seedlings and adults. There was no significant difference in F_{ST} values when comparing adults and seedlings (P = 0.293). However, in four of the six sites, significant genetic differentiation between adults and seedlings was found, ranging from 0.0025 to 0.0302 (Table 2.2).

Table 2.2 Genetic variation and fine-scale spatial genetic structure at nine microsatellite loci in 12 subpopulations of *Catanopsis eyrei*.

Plot	Subpopulation	*A*	A_R	H_e	F_{IS}	F_{ST}	$F_{(1)}$	*blog*	*Sp*	σ (m)
CSP6	Adults	13.56	11.88	0.74	0.198**	0.0015	0.0275	-0.0069**	0.0071	NC
	Seedlings	12.78	11.48	0.74	0.202**		0.0166	-0.0121**	0.0123	9.8
CSP13	Adults	11.78	9.73	0.65	0.226**	0.0093*	0.0431	-0.0122**	0.0127	NC
	Seedlings	13.89	10.35	0.67	0.217**		0.0141	-0.0090**	0.0091	13.6
CSP14	Adults	9.89	8.94	0.64	0.311**	0.0110*	-0.0003	-0.0052*	0.0052	NC
	Seedlings	14.78	10.74	0.71	0.176**		0.0119	-0.0076**	0.0077	9.7
CSP15	Adults	12.22	10.17	0.69	0.219**	0.0302*	0.0238	-0.0071**	0.0073	NC
	Seedlings	13.56	10.47	0.75	0.146**		0.0111	-0.0082**	0.0083	12.1
CSP21	Adults	11.56	9.16	0.68	0.186**	0.0025*	0.0306	-0.0052**	0.0053	NC
	Seedlings	12.44	8.26	0.67	0.182**		0.0123	-0.0096**	0.0097	8.0
CSP27	Adults	14.22	11.85	0.74	0.227**	-0.0003	0.0266	-0.0059**	0.0061	NC
	Seedlings	14.22	11.72	0.74	0.224**		0.0114	-0.0082**	0.0083	NC
Overall										
Adult			10.52	0.69	0.210		0.0515	-0.0079**	0.0083	25.4
Seedling			10.50	0.70	0.185		0.0478	-0.0103**	0.0109	9.1
P (A *v.* S)			0.98	0.48	0.23					

A, number of alleles per locus; A_R, allelic richness; H_e, expected heterozygosity; F_{IS}, fixation index; F_{ST}, genetic differentiation between adult and seedling subpopulations in each plot; $F_{(1)}$, kinship coefficient in the first distance class; *blog*, regression slope for kinship coefficient on the logarithm of geographical distance; *Sp*, statistic reflecting the intensity of SGS; σ, estimate of effective gene dispersal distance.

NC, no convergence; *$P < 0.05$, **$P < 0.01$.

Fine-scale SGS and effective gene dispersal

Significant values of regression slope (*blog*) were found in all subpopulations and in the combined analysis, indicating significant SGS (Fig. 2.1, Table 2.2). Significant positive autocorrelations existed in all seedling and adult subpopulations within 10 m, except for the adult cohort of CSP14. The maximum extent of SGS was 20 m appearing in adult cohorts of CSP13. In the first distance class (5 m), the kinship coefficients of adults were higher than those of seedlings in all plots except for CSP14.

The *Sp* statistic showed weak intensity of SGS in all subpopulations with values ranging from 0.0052 to 0.0127 (Table 2.2). Adult trees showed a lower *Sp* value than seedlings in all but one plot and for the combined data set (adults: *Sp* = 0.0083; seedlings: *Sp* = 0.0109). This result was corroborated by heterogeneity tests indicating that patterns of SGS differ between adults and seedlings (Table 2.3). However, for single plots overall significance was only found in three plots and the distance classes that differed were plot-specific. Only in the largest distance class (40m), no heterogeneity was found between adults and seedlings. When pooling all individuals to a single dataset, the adults showed higher autocorrelation than seedlings at 5 and 30 m, while at 10 m the value was significantly lower in adults (Fig. 2.2).

The iteration analysis of effective gene dispersal distances (σ) converged in five of the seedling subpopulations but in none of the adult subpopulations (Table 2.2). The effective gene dispersal distance of seedlings varied from 8.0 to 13.6 m. For the combined data sets, effective gene dispersal distance converged and equalled 25.4 m in adults and 9.1 m in seedlings.

Table 2.3 Heterogeneity tests of SGS between adult and seedling subpopulations using GENALEX. Statistics t^2 and ω represent the degree of differentiation of SGS between two pairwise subpopulations in each class and as a whole, respectively.

plot	t^2						ω
	5m	10m	15m	20m	30m	40m	
CSP6	0.8	0.1	1.1	3.7	3.4	0.0	7.4
CSP13	14.3**	10.2**	5.3*	9.6**	0.8	0.8	25.4**
CSP14	0.4	0.0	0.8	0.1	0.1	1.4	5.0
CSP15	2.2	4.3*	2.2	8.1**	8.4**	0.0	42.7**
CSP21	5.5*	0.5	0.5	3.0	13.6**	0.9	17.3**
CSP27	1.5	0.5	1.4	2.3	0.6	0.0	7.6
All plots combined	6.0*	4.6*	0.1	1.6	22.0**	0.3	19.1**

$^{*}P < 0.05$, $^{**}P < 0.01$.

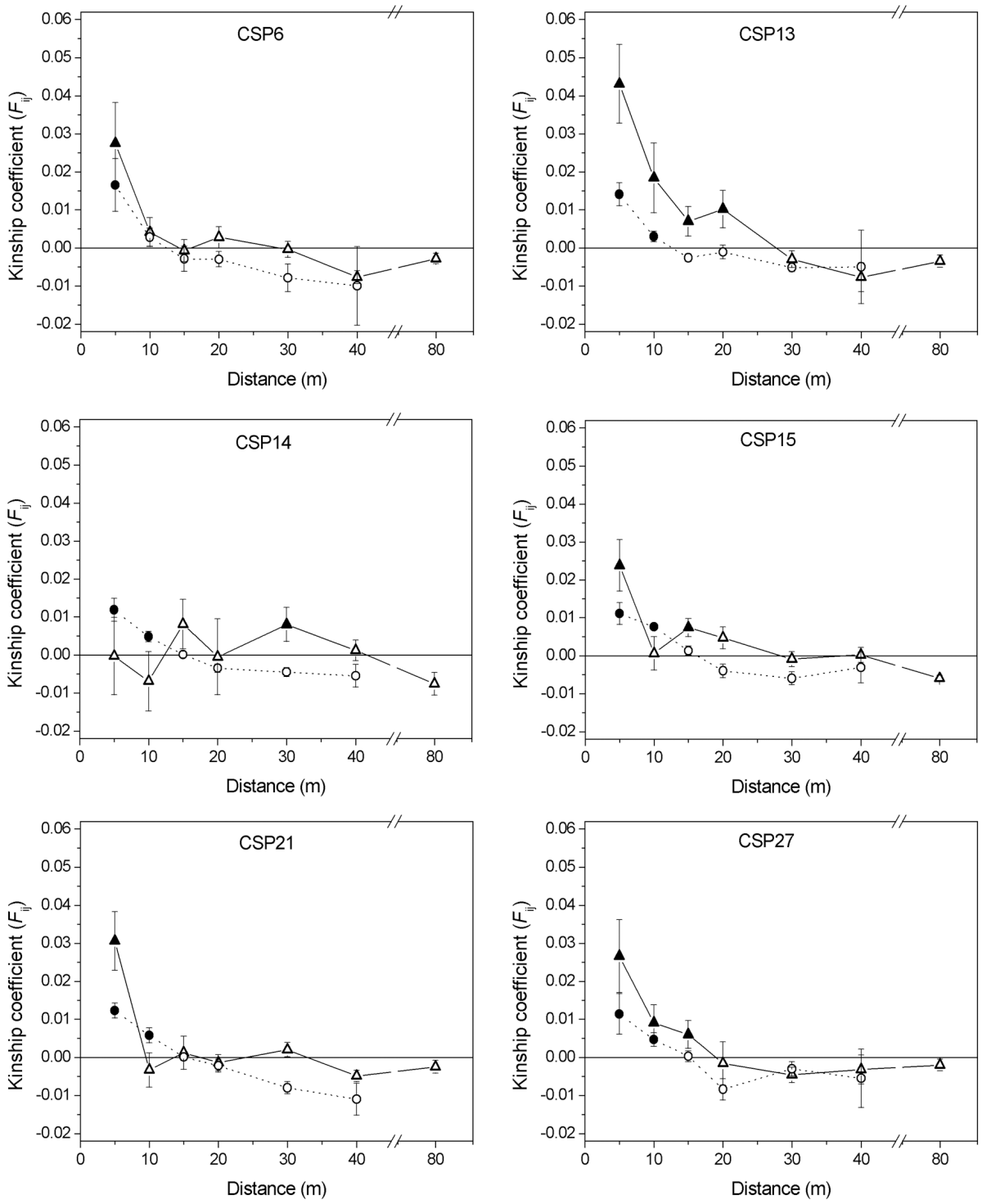

Fig. 2.1 Spatial genetic autocorrelation of adults (triangles) and seedlings (circles) within six plots using SPAGeDi. Error bars indicate standard errors. Filled symbols represent values significantly different from the expected value under a random distribution of genotypes (95% confidence level).

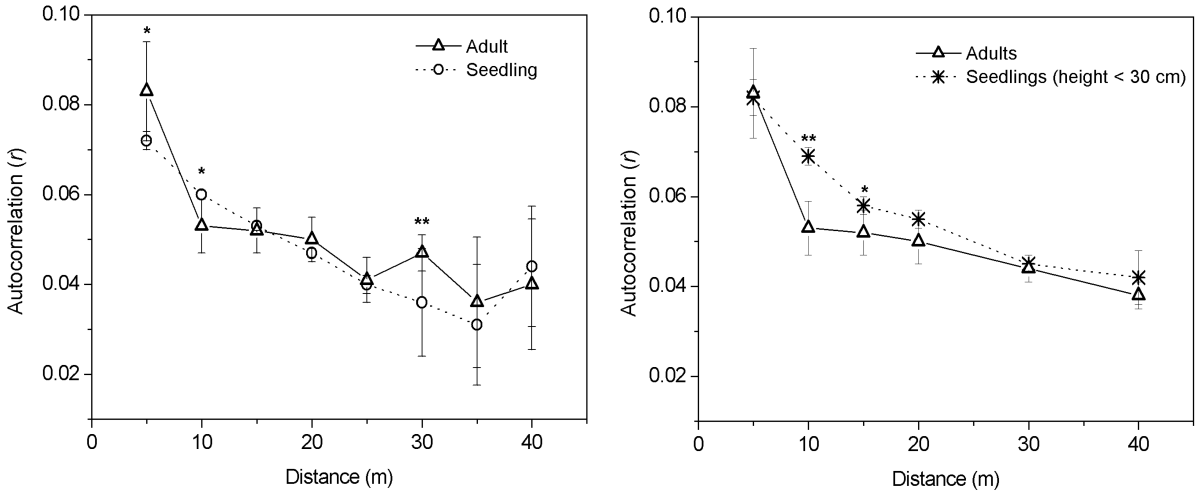

Fig. 2.2 Spatial genetic autocorrelation of the combined dataset of adult and seedling subpopulations (Left: all seedlings; right: seedlings with hight < 30 cm) with results of a heterogeneity test using GENALEX. ω-test indicates overall significance (ω = 19.050, P = 0.002), $^{*}P < 0.05$ and $^{**}P < 0.01$ indicate significant differences at single distance classes.

Effect of habitat conditions on SGS

For adults and seedlings multiple linear regressions were performed between within-population SGS, quantified by the *Sp* value, and the six habitat predictors. The results showed that no significant factors left in the model ($P > 0.727$), indicating SGS was independent of density, elevation, successional stage, inclination, species richness and woody species diversity in this study.

Direct estimation of the seed and pollen dispersal

Maximum likelihood estimates of seed and pollen dispersal as well as mating system parameters are shown in Table 2.4. The average distance of seed dispersal ranged from 9.7 (CSP13) to 43.2 m (CSP6) among plots with a mean value of 21.0 m, whereas the average pollen dispersal distance was higher with a mean value of 55.4 m (range 11.0 - 132.3 m). Both high proportions of seed and pollen immigration were detected. A range from 46% to 89% seeds (mean m_s = 0.69) came from outside of the plots while pollen immigration rate varied from 46% to 95% (mean m_p = 0.75). As expected, the average selfing rate was close to zero.

Table 2.4 Estimation of seed and pollen dispersal kernel based on seedling neighbourhood model.

Plot	Parameters						
	m_s	m_p	s	δ_s (m)	δ_p (m)	b_s	b_p
CSP6	0.78	0.95	0.00	43.2	85.4	1.00[a]	1.00[a]
CSP13	0.46	0.70	0.01	9.7	25.3	0.59	0.32
CSP14	0.89	0.77	0.06	16.4	NC	1.11	1.00[a]
CSP15	0.59	0.89	0.00	10.3	23.1	1.08	2.05
CSP21	0.65	0.46	0.00	22.7	132.3	0.85	1.00[a]
CSP27	0.81	0.76	0.00	23.9	11.0	0.90	0.90
Mean	0.69	0.75	0.01	21.0	55.4		

m_s, seed immigration rate; m_p, pollen immigration rate; s, selfing rate; δ_s, average distance of seed dispersal; δ_p, average distance of pollen dispersal; b_s and b_p, shape parameters of exponential-power kernel for seeds and pollen, respectively. NC, no convergence. [a] Fixed value for the parameter.

Discussion

Genetic diversity of adults and seedlings

In species with limited seed dispersal, seedlings are suggested to aggregate around the maternal trees and display high relatedness. This may lead to lower genetic diversity in the seedling cohort as shown for tropical tree *Copaifera langsdorffii* (Sebbenn *et al.* 2011). However, still many other studies have demonstrated that high genetic diversity was maintained across different life stages. For example, similar levels of expexted heterozygosity were detected in the three life stages of *Sinojackia rehderiana* (Styracaceae) (adult : juvenile : seedling = 0.778 : 0.782 : 0.785) (Yao *et al.* 2011). In our study, genetic variation did not differ between seedlings and adults. In contrast to the study of *C. langsdorffii*, which targeted a small, isolated and fragmented population subject to genetic drift, our study refers to a common species in continuously distributed forest conforming to random mating. In fact, significant patterns of isolation by distance (Table 2.2, indicated by significant *blog* values) were detected in both cohorts, indicating the equilibrium between genetic drift and gene dispersal (Born *et al.* 2008; Hutchison & Templeton 1999).

In addition, outcrossing species are expected to be able to maintain the level of genetic diversity (Duminil *et al.* 2007). Although *C. eyrei* is a monoecious tree, highly outcrossing was suggested by only 1% selfing rate on average (Table 2.4). Significant fixation index (F_{IS}) was found for both cohorts in any plot. However, it can not always indicate the high inbreeding depression. First, most importantly, null alleles increase the estimation of F_{IS} values when heterozygous individuals were misclassified as homozygous ones because of

lacking amplification of null alleles (Shi & Chen 2012). Indeed, most of loci showed signs of null alleles with a men frequency of 7.65%. The similar values of F_{IS} in adults and seedlings seem to prove the effect of null alleles on estimation of F_{IS}. Additionally, if inbreeding happens, the fixation index should be equal or similar to the kinship coefficient (F_{ij}) among individuals (Sebbenn *et al.* 2011). Nevertheless, all the F_{IS} values were far larger than the largest values of F_{ij}, indicating inbreeding can not, at least fully, explain the high values of fixation index.

SGS in adults and seedlings

In woody species, strong SGS present in seedlings is expected to largely disappear in adults due to thinning processes. Such a hypothesis has been confirmed by several studies (Chung *et al.* 2003a; Qi *et al.* 2011). For example, within populations of *Camellia japonica* (Theaceae) (Chung *et al.* 2003a), weak but significant SGS was found in seedlings, which disappeared in older age classes with a random spatial distribution of genetic variation. However, this is not the case in this study of *C. eyrei*. Significant SGS was detected in both adults and seedlings, although the extent was weak, as expected for outcrossing tree species (Vekemans & Hardy 2004), up to a maximum of 20 m, which is consistent with the estimation on other Fagaceae species (Oddou-Muratorio *et al.* 2010; Streiff *et al.* 1998; Wang *et al.* 2011). The *Sp* statistic, quantifying the intensity of SGS, showed similar values between adults (0.0083) and seedlings (0.0109) although SGS differed at some distance classes between adults and seedlings rather idiosyncratically in some plots and in the combined analysis. Extremely high seedlings may result in a bias of SGS, however, when restricting the seedlings < 30 cm height, very similar patterns were revealed, excluding the effect of the size of seedlings (Fig. 2.2).

Although SGS can be influenced by various factors, gene flow should be a key determinant. (Vekemans & Hardy 2004; Wright 1943). The homogeneity of SGS in adults and seedlings suggests that overall patterns of gene flow at population level are similar over time, which could be demonstrated by the similar level of their effective gene dispersal distance (seedlings vs. adults: 9.1 m vs. 25.4 m). Similarly, the consistent SGS revealed in different life stages in *S. rehderiana* was attributed to limited gene flow (Yao *et al.* 2011). In woody species, the initial SGS in established seedlings largely depends on seed dispersal and adult density. Although seeds dispersed by gravity often aggregate around the maternal trees, secondary dispersal by small rodents could blend the seeds from neighbouring trees and promote overlapping distribution of half-sibs from different families. In addition, high density

of adult trees should be another explanation of weak SGS in seedlings, since a high density results in overlapping seed shadows, leading to mixing of different half-sib families, thus a reduction in relatedness of adjacent individuals. On the other hand, high density of adults reflects a relatively low mortality, thus permitting the initial fine-scale SGS within maternal seed shadows to persist into the adult stage (Hamrick *et al.* 1993).

Genetic structure within populations is dependent on the processes of pollen and seed dispersal which can be influenced by environmental conditions. However, none of the factors tested significantly affected SGS of adults or seedlings, indicating its independence from density, elevation, inclination, successional stage and species richness in this study. Thus, although the plots covered the whole range of local environmental conditions, particularly for tree density and elevation, SGS is maintained at a similar level.

Gene flow, pollen and seed dispersal

Both adults and seedlings revealed weak but significant SGS. Assuming that effective density equals ¼ of census density, the historical gene dispersal distance was estimated between 9.14 m and 25.36 m, which is surprisingly low compared to other tree species (Hardy *et al.* 2006; Oddou-Muratorio & Klein 2008). For example, a range of gene dispersal distance from 150 m to 1200 m was found in 10 tropical tree species (Hardy *et al.* 2006). Since *Sp* is expected to equal $1/4\pi D_e\sigma^2$ in two-dimensional space in a restricted range of σ to 20σ (Vekemans & Hardy 2004), it implies that the intensity of genetic structure decreases both with increasing dispersal and increasing individual density. On the other hand, when *Sp* values are similar, the estimation of gene dispersal will be affected by effective density. Indeed, in the study of *Sorbus torminalis* (Oddou-Muratorio & Klein 2008), similar *Sp* values were reported (0.0187) comparable to *C. eyrei* (0.0083), however, much higher dispersal distance (472 m) was computed, as almost 20 times of that of *C. eyrei*, on the condition that the adult density of *S. torminalis* (0.39 trees/ha) was only one-thousandth of that of *C. eyrei* (424 trees/ha). Nevertheless, it is suggested that long-distance dispersal can counterbalance the decrease of population density, thus maintaining the level of genetic variation (Born *et al.* 2008; Dick *et al.* 2003).

Direct estimation of seed and pollen dispersal using the spatial explicit mating model revealed relatively restricted seed dispersal abilities (mean δ_s = 21.04 m) in *C. eyrei* which were in line with gene dispersal distance as deduced from SGS. Similar values of seed dispersal were obtained in other Fagaceae species, e.g. *Q. robur* and *Q. petraea* (8.8 m and 15.6 m, respectively) (Chybicki & Burczyk 2010b), *F. sylvatica and F. crenata* (10.5 m and

12.4 m, respectively) (Oddou-Muratorio *et al.* 2010). However, relatively high seed immigration rates (46% < m_s < 89%) were detected, indicating that long distance seed dispersal events can not be neglected in our study sites which should have resulted from secondary dispersal by small rodents. The unusual values, compared to other studies, might be attributed to the relatively smaller sampling area and smaller number of adults. In our study (appr. 0.5 ha) (Supporting information), on average 151 adults per plot were collected and in the core of each plot, seedlings were sampled in an area of 900 m^2, whereas in the study of Chybicki & Burczyk (2010b), they collected 421 adult trees within an area of 5 ha and seedlings in the core with an area of 350 m^2. Large coverage of sampling for adults will increase the possibilities of involving the maternal tree of seedlings. In addition, seed immigration rate may be overestimated because genotyping error and the presence of null alleles could result in false parentage exclusion (Oddou-Muratorio *et al.* 2010).

Regarding pollen dispersal, greater dispersal abilities of pollen versus seeds are suggested, as expected. The maximal distance was determined as 132.3 m, although in some plots, the distance of pollen dispersal (δ_p) seems quite close (e.g CSP27, δ_p = 11.0). However, pollen dispersal distance should be considered with caution. As discussed above, when compared to other studies, the area investigated should be considered, because it directly influences the possibilities to involve both parents. Additionally, as discussed in Oddou-Muratorio et al. (2010), δ_p is estimated based on seedlings whose parent pair was found within the study sites. However, few pairs were matched, since large proportions of pollen came from outside the study sites (46% < m_p < 95%), indicating extensive pollen flow.

Acknowledgments

This study was financially supported by the German Research Foundation (DFG FOR 891/1 and 891/2) and the China Scholarship Council (2008 3039). We thank Christoph Hahn for primer development and Juliane Horn for DNA extraction.

Supporting information:

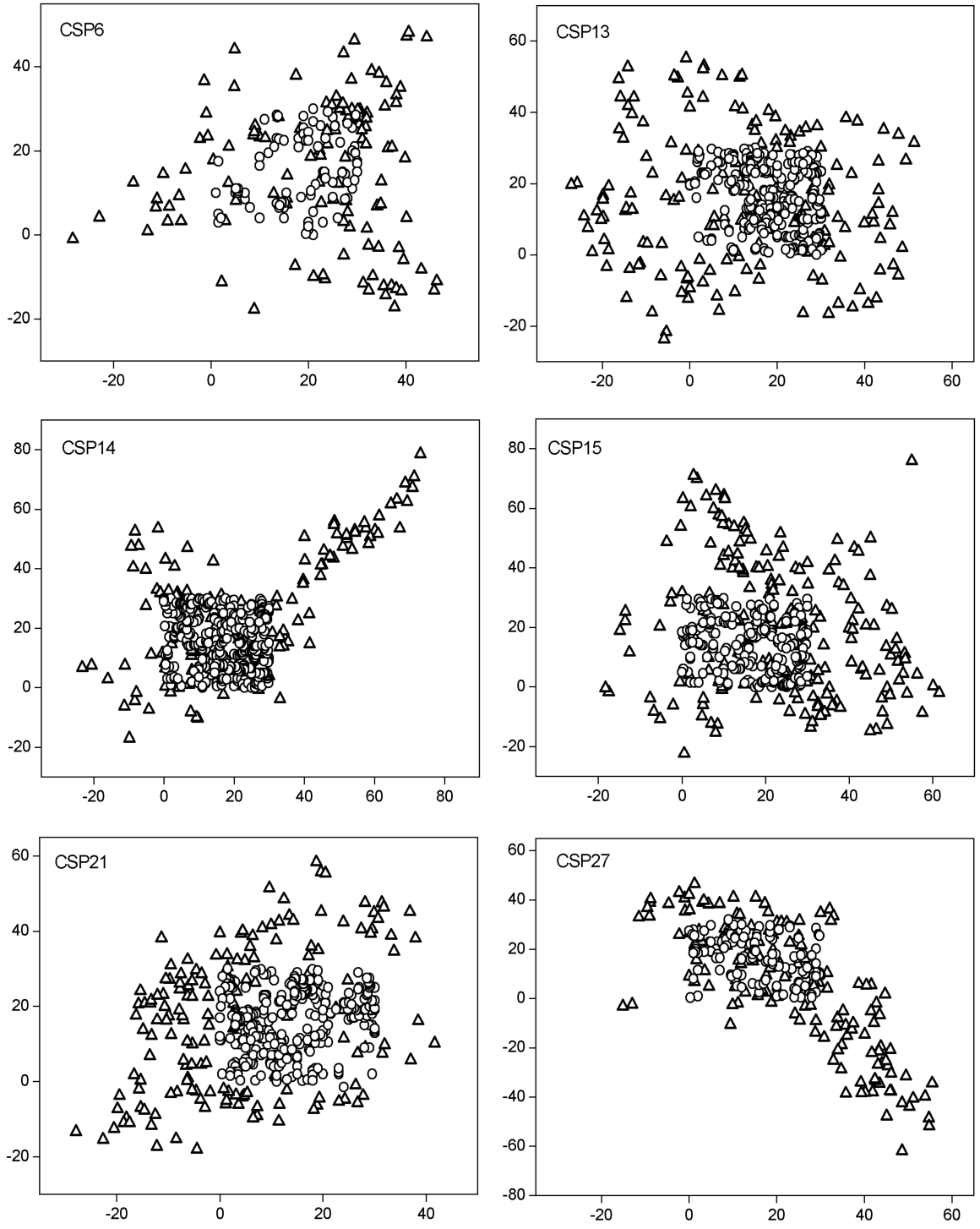

Fig. 2.S1 Distribution of adult trees (triangles) and seedlings (circles) of *Castanopsis eyrei* in the six plots. X and Y scale in meters.

Chapter 3 - Isolation by elevation: genetic structure at neutral and putatively non-neutral loci in a dominant tree of subtropical forests, *Castanopsis eyrei*

Miao-Miao Shi, Stefan G. Michalski, Xiao-Yong Chen and Walter Durka

PLoS ONE 2011, 6(6):e21302

Abstract

Background: The distribution of genetic diversity among plant populations growing along elevational gradients can be affected by neutral as well as selective processes. Molecular markers used to study these patterns usually target neutral processes only, but may also be affected by selection. In this study, the effects of elevation and successional stage on genetic diversity of a dominant tree species were investigated controlling for neutrality of the microsatellite loci used.

Methodology/Principal Findings: Diversity and differentiation among 24 populations of *Castanopsis eyrei* from different elevations (251-920m) and successional stages were analysed by eight microsatellite loci. We found that one of the loci (Ccu97H18) strongly deviated from a neutral model of differentiation among populations due to either divergent selection or hitchhiking with an unknown selected locus. The analysis showed that *C. eyrei* populations had a high level of genetic diversity within populations (A_R = 7.6, H_E = 0.82). Genetic variation increased with elevation for both the putatively selected locus Ccu97H18 and the neutral loci. At locus Ccu97H18 one allele was dominant at low elevations, which was replaced at higher elevations by an increasing number of other alleles. The level of genetic differentiation at neutral loci was similar to that of other Fagaceae species (F_{ST} = 0.032, F'_{ST} = 0.15). Population differentiation followed a model of isolation by distance but additionally, strongly significant isolation by elevation was found, both for neutral loci and the putatively selected locus.

Conclusions/Significance: The results indicate higher gene flow among similar elevational levels than across different elevational levels and suggest a selective influence of elevation on the distribution of genetic diversity in *C. eyrei*. The study underlines the importance to check the selective neutrality of marker loci in analyses of population structure.

Introduction

Genetic composition within and among populations is shaped by the interplay of genetic drift, gene flow, mutation and natural selection. Molecular markers have helped to identify the effect of life history traits, phylogeographic history and environmental factors on the genetic structure of plant populations (Hamrick & Godt 1996; Nybom 2004). Among environmental factors, abiotic factors, such as soil type, topology or elevation, play an important role in genetic structuring because they may affect phenology, population size or density and thus gene flow or genetic drift (Byars *et al.* 2009). Elevation is of particular importance, and many studies focused on its relationships with plant performance and phenotype (Körner 2007), but also on genetic variation of molecular markers (Byars *et al.* 2009; Gämperle & Schneller 2002; Goto *et al.* 2009).

Genetic variation within populations often varies along elevational gradients and among species different patterns have been identified (Ohsawa & Ide 2008). First, mid-elevation populations may hold higher levels of diversity compared with both low and high elevation populations due to the optimal mid-elevation habitats following the central-marginal hypothesis (e.g (Herrera & Bazaga 2008)). Second, low elevation populations may have highest diversity which decreases with elevation as a result of bottlenecks occurring throughout upward range expansion (e.g (Quiroga & Premoli 2007)). Third, highest genetic diversity was found at high elevations which was attributed to various reasons like decreased human disturbance and/or historical downward range shifts due to climate change, and adaptation (Gämperle & Schneller 2002; Ohsawa & Ide 2008). Lastly, genetic variation also has been found to stay rather constant along a given elevational gradient due to extensive gene flow (e.g. (Truong *et al.* 2007b)). Overall, these inconsistent patterns support a predominant role of life history traits and of biogeographic history in determining patterns of genetic variation along elevational gradients. The processes underlying these patterns are either neutral, like genetic drift and bottleneck effects as a result of the demographic history, or are selective due to the climatic clines related to elevation.

Elevational clines encompass a suite of environmental factors that are either physically linked with elevation like temperature (Zhu *et al.* 2010) or that are instead correlated with it, like land use (Körner 2007). Depending on the ability of these factors to exert selection or to affect the neutral processes of gene flow and drift, molecular markers may display elevational patterns. Of the various molecular markers, which have been used to study genetic variation, microsatellites are assumed to represent neutral markers because microsatellites are generally found in non-coding regions (Nielsen *et al.* 2006) and are characterized by high levels of

variability. Consequently, patterns of differentiation among populations at microsatellite loci are almost exclusively interpreted as genetic drift and gene flow. However, some empirical studies indicated the presence of non-neutral microsatellite loci (Casa *et al.* 2005; Lazrek *et al.* 2009; Nielsen *et al.* 2006). Thus, in order to study neutral processes the neutrality of loci should be confirmed before performing other genetic analyses (Beaumont & Nichols 1996; Ohsawa & Ide 2008). Due to the steep clines in environmental conditions with increasing elevation accompanied by changes in selective conditions, non-neutral behaviour of individual molecular markers is likely, e.g. due to physical linkage to specific genes under selection (e.g. (Jump *et al.* 2006)).

In mixed and evergreen broad-leaved subtropical forests of Southeast Asia, *Castanopsis eyrei* (Fagaceae) is often the dominant tree species in late successional forests. The long lived evergreen species is native to southeastern China and Taiwan and occurs along a large elevational gradient from <300 m to 1700 m a.s.l. (http://www.efloras.org/florataxon.aspx?flora_id=620&taxon_id=200006236). It is monoecious and wind-pollinated and the acorn seeds are predominantly dispersed by gravity and small rodents (Li & Jin 2006). Due to these life history traits, *C. eyrei* populations are expected to have high genetic diversity and efficient gene flow mediated by pollen dispersal should result in low levels of genetic differentiation.

In this study we examined the distribution of genetic variation in *C. eyrei* populations within a nature reserve of continuous mixed broad-leaved forest across a mountain range. Specifically, we ask (1) whether individual loci are more strongly differentiated than expected from a neutral model, and (2) whether spatial structure, elevation or successional stage affect the patterns of neutral and of putatively adaptive genetic variation, respectively.

Results

Identification of loci under selection

Outlier tests performed using FDIST detected a significant departure of the F_{ST} value from neutral expectations for locus Ccu97H18 (F_{ST} = 0.316, Fig. 3.1), while for other loci F_{ST} values ranged from F_{ST} = 0.029 to 0.055. However, four of them with lower F_{ST} values were also situated out of the simulated distribution, which was probably due to the extremely high value of Ccu97H18. When we excluded this locus and reanalysed the other seven loci, the result confirmed their neutrality as all of them were situated within the 0.99 quantile.

Analysing only locus Ccu97H18, we found an increase in the number of alleles with elevation from an average of 2.2±1.2 below 400 m a.s.l. to 16.8±3.9 above 800 m a.s.l (Fig.

3.2). This was due to one allele in particular (145 bp) which was most common with frequency close to 1.0 at lower elevations (< ca. 700 m), whereas its frequency decreased drastically at higher elevations.

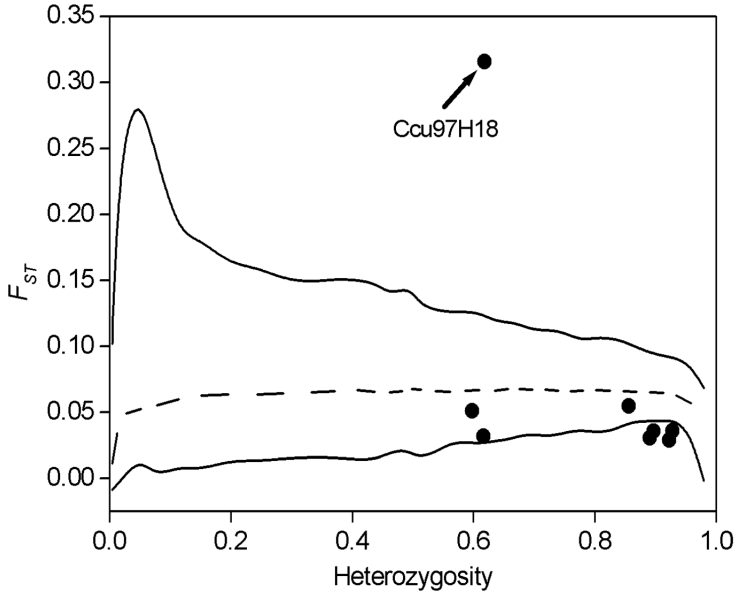

Fig. 3.1 F_{ST} values of eight microsatellite loci in *Castanopsis eyrei* populations plotted against heterozygosity. The lines represent the median (broken line) and 99% quantiles of expected F_{ST} values estimated from a neutral model (Beaumont & Nichols 1996).

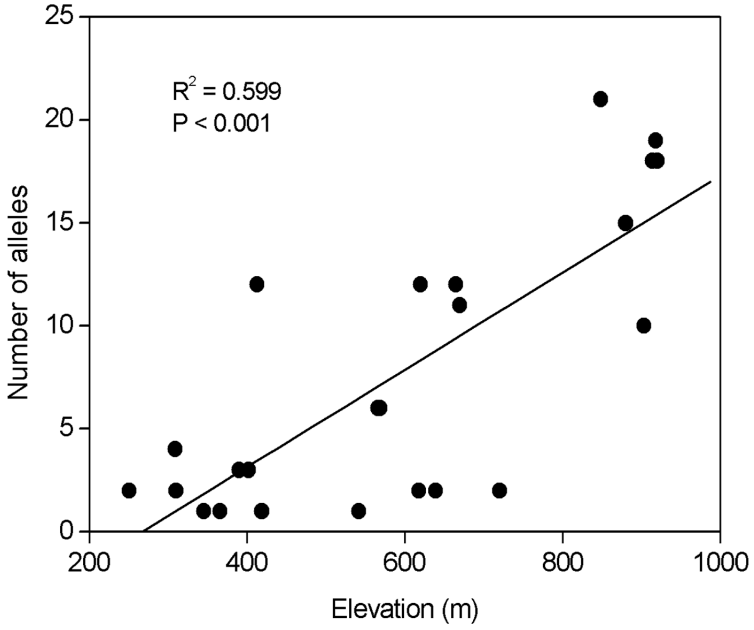

Fig. 3.2 Number of alleles per population of *Castanopsis eyrei* as a function of elevation at microsatellite locus Ccu97H18.

Genetic diversity at species and population level

Genetic parameters at species and population levels for both the putatively neutral loci and the putatively selected locus Ccu97H18 are displayed in Table 3.1. In a total of 583 individuals and at the seven putatively neutral loci, we identified 129 alleles with a number of 10 to 25 alleles per locus. At the population level, the mean number of alleles per locus ranged from 6.1 to 12.1 (mean = 9.4) and allelic richness (A_R) varied from 5.4 to 7.7 (mean = 6.7). The expected heterozygosity (H_E) ranged from 0.68 to 0.86 among populations (mean = 0.78). At the species level, *C. eyrei* had a H_E value of 0.82. The bottleneck analyses indicated recent reduction in population size in five sites (Table 3.1), which were located at low, medium and high elevations.

Effects of environmental factors

Successional stage was significantly interrelated with elevation (r = 0.567, P = 0.004 Spearman correlation). Over all neutral loci, the multiple regression of allelic richness (A_R) against elevation and successional stage showed that A_R increased significantly with elevation but succession had no significant contribution (r = 0.586, P = 0.005). For the putatively selected locus Ccu97H18, similarly only elevation had a significantly positive strong effect on A_R in the mutliple regression analysis (r = 0.708, P < 0.001).

Population differentiation

Populations were significantly structured as revealed by overall F_{ST} over the seven neutral loci of 0.032 (P < 0.01). However, the standardized F'_{ST} value was 0.15, indicating considerable differentiation. When only the putatively selected locus Ccu97H18 was analysed, we detected much higher values of F_{ST} = 0.316 and F'_{ST} = 0.571. Significant patterns of isolation by geographic distance were found at neutral loci both for pairwise F_{ST} and F'_{ST} values (Fig. 3.3). For the putatively selected locus Ccu97H18, a significant pattern of isolation by distance was detected for pairwise F_{ST} (r = 0.193, P = 0.016), but the pattern did not exist for pairwise F'_{ST} (r = 0.069, P = 0.194). When we checked for a pattern of isolation by elevational distance, both pairwise F_{ST} and pairwise F'_{ST} revealed much more strongly significant correlations, indicating isolation by elevation, both in the neutral loci (Fig. 3.3) and the putatively selected locus (F_{ST}: r = 0.251, P = 0.002; F'_{ST}: r = 0.210, P = 0.003). Since elevational distance was correlated with geographic distance (r = 0.329, P = 0.002), we performed partial Mantel tests to test whether elevational distance was related to genetic differentiation after accounting for

geographic distance. For the neutral loci, elevational distance remained significant for pairwise F'_{ST} (r = 0.129, P = 0.010) but not for pairwise F_{ST} (r = 0.060, P = 0.123). For the putatively selected locus elevational distance remained significantly related to differentiation after accounting for geographic distance in both F_{ST} and F'_{ST} (F_{ST}: r = 0.188, P = 0.013; F'_{ST}**:** r = 0.201, P = 0.001).

Table 3.1 Sample sites and genetic diversity estimates for *Castanopsis eyrei* samples.

					Neutral loci				Ccu97H18		
Population	location	Elevation (m)	Successional stage	*N*	*A*	A_R	H_E	H_{Eeq}	*A*	A_R	H_E
CSP 02	E 118.13484, N 29.24926	390	5	20	10.1	7.2	0.78	0.656	3	2.3	0.11
CSP 03	E 118.12402, N 29.23885	720	3	18	9.1	6.8	0.76	0.852	2	1.8	0.06
CSP 04	E 118.12015, N 29.24963	542	5	22	10.0	7.3	0.86	0.004**	1	1.0	0.00
CSP 06	E 118.14747, N 29.25497	880	3	49	12.7	7.4	0.81	0.406	15	8.0	0.75
CSP 07	E 118.14373, N 29.25184	903	4	21	10.0	7.1	0.81	0.656	10	9.3	0.91
CSP 08	E 118.11019, N 29.24106	413	3	17	8.4	6.5	0.78	0.055	12	10.9	0.91
CSP 10	E 118.15791, N 29.25188	670	4	18	8.9	6.7	0.79	0.289	11	9.7	0.90
CSP 12	E 118.12190, N 29.24939	620	4	18	8.4	6.5	0.78	0.344	12	10.7	0.90
CSP 13	E 118.11621, N 29.24630	402	5	34	10.6	6.3	0.74	0.852	3	1.9	0.09
CSP 14	E 118.13518, N 29.24944	639	5	24	9.0	6.4	0.78	0.289	2	1.8	0.10
CSP 15	E 118.13106, N 29.24917	618	5	17	6.1	5.4	0.69	0.188	2	2.0	0.08
CSP 16	E 118.09966, N 29.24253	309	1	15	8.1	6.8	0.78	0.531	4	3.6	0.21
CSP 17	E 118.10828, N 29.24342	310	2	17	6.6	5.7	0.76	0.012*	2	1.7	0.06
CSP 18	E 118.12461. N 29.24516	569	3	18	7.3	6.1	0.79	0.039*	6	4.6	0.31
CSP 21	E 118.08084, N 29.27059	566	5	45	10.7	6.5	0.77	0.289	6	3.1	0.20
CSP 23	E 118.13723, N 29.21450	419	2	19	8.0	6.2	0.77	0.594	1	1.0	0.00
CSP 24	E 118.13469, N 29.21483	366	1	20	8.1	6.3	0.78	0.469	1	1.0	0.00
CSP 25	E 118.13155, N 29.21713	345	2	12	6.1	5.4	0.68	0.813	1	1.0	0.00
CSP 26	E 118.12155, N 29.21489	251	1	17	8.7	6.7	0.79	0.711	2	1.7	0.06
CSP 27	E 118.13605, N 29.24709	665	5	47	12.1	6.9	0.79	0.813	12	7.4	0.76
A	E 118.14136, N 29.24811	848	5	28	11.0	7.6	0.85	0.004**	21	14.8	0.96
B	E 118.14381, N 29.25244	914	5	27	11.9	7.4	0.83	0.406	18	13.1	0.94
C	E 118.14345, N 29.25249	918	5	30	11.4	7.5	0.83	0.148	19	14.1	0.96
D	E 118.14306, N 29.25265	920	5	30	11.9	7.7	0.85	0.004**	18	13.2	0.94
Mean				24	9.4	6.7	0.78		7.7	5.8	0.43
All				583	18.4	7.6	0.82		29	8.6	0.43

N, number of individuals sampled per population; *A*, mean number of alleles; A_R, allelic richness; H_E, expected heterozygosity;. H_{Eeq}, expected heterozygosity under drift-migration equilibrium and TPM with significant departure indicated by asterisks. *P<0.05, **P<0.01

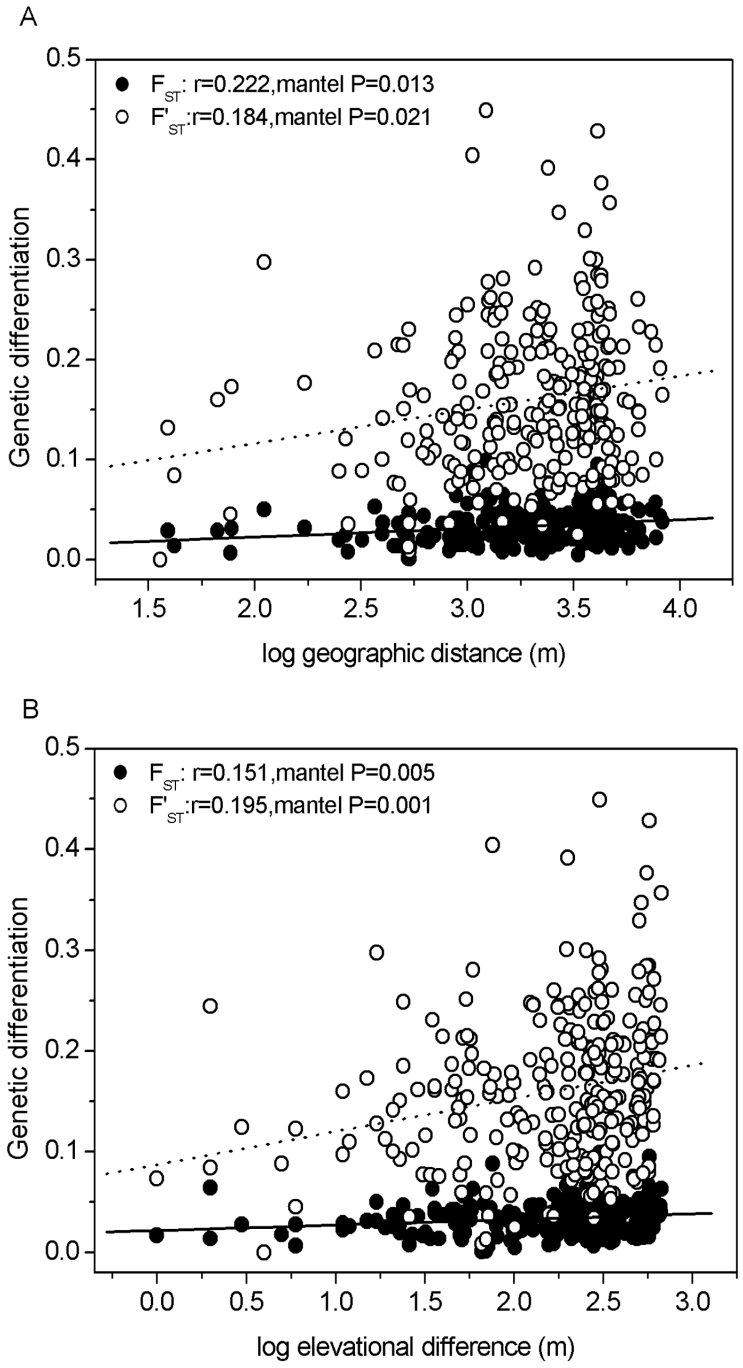

Fig. 3.3 Patterns of isolation by distance and isolation by elevation in *Castanopsis eyrei*. Population differentiation (closed circle and solid line: pairwise F_{ST}; open circle and dotted line: standardized F'_{ST}) at neutral loci as a function of (A) geographic distance and (B) as a function of elevational difference. Correlation coefficient and Mantel-*P*.

Discussion

Neutrality of microsatellite loci

Microsatellites are assumed to represent ideal neutral markers, so that only gene flow and genetic drift rather than selection should affect their genetic structure. However, an increasing number of studies indicated the presence of non-neutral loci (Acheré *et al.* 2005; Lazrek *et al.* 2009; Scascitelli *et al.* 2010). In the present study one out of eight loci that were originally developed for *C. cuspidata* var. *sieboldii* (Ueno *et al.* 2000, 2003) showed non-neutral behaviour. However, no information on the genomic position and putatively linked genes of this locus is available (Ueno pers. comm.). Based on the analysis of expressed genes of *C. sieboldii*, Ueno *et al*. (Ueno *et al.* 2009) showed that microsatellites are widespread with 314 microsatellites in 2417 potential unigenes. Consequently, microsatellite markers may be linked to expressed genes and, hence, tests of neutrality should precede population genetic analyses. Since only a limited number of microsatellite loci are routinely analysed in such studies and given that average linkage disequilibrium is expected to be low in outcrossing species, the likelihood of finding a marker linked to an adaptively important gene may be low (Jump *et al.* 2006). However, based on studies that used the method of Beaumont and Nichols (1996) to identify non-neutral microsatellite loci in plants, between 4% (one out of twenty six for *Fucus serratus*, (Coyer *et al.* 2008)) and 33% (three out of nine for *Astronium urundeuva*, (Caetano et al. 2008)) of loci were found to behave non-neutrally. However, these seemingly high levels of non neutral loci may be overestimated as the identification of outlier loci with non-neutral behaviour also produces false-positives (Pérez-Figueroa *et al.* 2010), which should be controlled e.g. by correlating allele frequencies with potentially selective site conditions (e.g. (Michalski *et al.* 2010)).

Genetic diversity of Castanopsis eyrei

At the seven neutral microsatellite loci employed in this study a total of 129 alleles were detected with 10 to 25 alleles found per locus. Ueno *et al*. (2000, 2003) detected a total of 78 alleles in *C. cuspidata* with these same loci in a limited number of individuals. In our study, *C. eyrei* showed many more alleles than *C. cuspidata* in the original work, possibly due to the larger sample size ($N = 583$ and $N = 24$ for *C. eyrei* and *C. cuspidata*, respectively).

Genetic variation at the species level in *C. eyrei* was high ($H_E = 0.82$) and similar to that of other congeneric species like *C. cuspidata* ($H_E = 0.83$, (Vidhanaarachchi *et al.* 2005)), and *C. acuminatissima* ($H_E = 0.72$, (Blakesley *et al.* 2004)). These species share common characteristics like an outcrossing mating system, wind pollination and a long life span.

Furthermore, they are all climax species with a broad current distribution and thus may also have similar demographic histories. Species exhibiting these traits are generally expected to show high levels of genetic variation (Hamrick & Godt 1996).

Population structure

Focusing on neutral genetic variation and thus, excluding the putatively selected locus, overall population differentiation was low (F_{ST} = 0.032) indicating only little differentiation (Wright 1978). However, the adjusted F'_{ST} (Hedrick 2005) equaled 0.15 for the neutral loci. Such values would indicate substantial differentiation, quite unexpected for perennial, wind pollinated and outcrossing species. However, other tree species of the Fagaceae that were analysed with microsatellite markers show similar values with a mean of F'_{ST} = 0.145 (Table 3.2). Levels of differentiation derived from dominant markers are somewhat lower (F'_{ST} = 0.124 for AFLP or RAPD markers, Table 3.2) and drastically lower when estimated from allozyme markers (mean F'_{ST} = 0.054, Table 3.2). This discrepancy indicates that the absolute level of F'_{ST} values has to be interpreted with caution, e.g. marker specific mutation rates have to be taken into account. In fact it seems unlikely that across the scale of a few kilometres populations of these tree species are strongly differentiated in neutral markers because of extensive pollen flow and seed dispersal by animals.

The absolute levels of standardized pairwise population differentiation, F'_{ST}, approached unity in several cases at the putatively selected locus Ccu97H18. This demonstrates that these population pairs are almost fixed for different alleles, a fact that is not obvious with traditional F_{ST}. However, the relationship between population differentiation and spatial or elevational distance was almost the same for the traditional and standardized F_{ST} values. Thus a more comprehensive understanding of differentiation patterns is possible using standardized differentiation measures (Heller & Siegismund 2009; Lange *et al.* 2010).

Table 3.2 Genetic differentiation among populations of tree species in the family Fagaceae as assessed with different molecular markers.

species	*N*	H_S	F_{ST}	$F_{ST\,max}$	F'_{ST}	Reference
microsatellites						
Castanea crenata	6	0.780	0.034	0.190	0.179	(Tanaka *et al.* 2005)
Castanopsis acuminatissima	3	0.716	0.006	0.209	0.029	(Blakesley *et al.* 2004)
Cyclobalanopsis myrsinaefolia	5	0.553	0.061	0.393	0.155	(Liu *et al.* 2008)
Fagus crenata	23	0.839	0.027	0.155	0.174	(Hiraoka & Tomaru 2009a)
Fagus japonica	16	0.659	0.023	0.327	0.070	(Hiraoka & Tomaru 2009b)
Lithocarpus densiflorus	19	0.540	0.090	0.447	0.202	(Nettel *et al.* 2009)
Quercus garryana	22	0.596	0.063	0.393	0.160	(Marsico *et al.* 2009)
Quercus glauca	10	0.741	0.042	0.239	0.176	(Lee *et al.* 2006)
Quercus robur	7	0.868	0.018	0.115	0.156	(Mariette *et al.* 2002)
allozymes						
Cyclobalanopsis championii	5	0.151	0.092	0.818	0.112	(Cheng *et al.* 2001)
Cyclobalanopsis glauca	6	0.222	0.065	0.745	0.087	(Chen *et al.* 1997)
Fagus crenata	23	0.187	0.038	0.806	0.047	(Tomaru *et al.* 1997)
Fagus sylvatica	38	0.250	0.030	0.745	0.040	(Hazler *et al.* 1997)
Quercus acutissima	3	0.152	0.010	0.788	0.013	(Chung *et al.* 2002)
Quercus chrysolepis	6	0.443	0.018	0.511	0.035	(Montalvo *et al.* 1997)
Quercus petraea	21	0.381	0.027	0.607	0.044	(Le Corre *et al.* 1997)
AFLP/RAPD						
Castanopsis fargesii	5	0.283	0.043	0.670	0.064	(Zhu *et al.* 2002)
Cyclobalanopsis glauca	4	0.315	0.104	0.620	0.167	(Zhang *et al.* 2006)
Lithocarpus harlandii	3	0.232	0.227	0.688	0.330	(Li *et al.* 2008)
Quercus petraea	4	0.289	0.021	0.649	0.032	(Coart *et al.* 2002)
Quercus petraea	21	0.233	0.024	0.758	0.032	(Le Corre *et al.* 1997)
Quercus robur	4	0.284	0.021	0.654	0.033	(Coart *et al.* 2002)
Quercus robur	7	0.190	0.110	0.785	0.140	(Mariette *et al.* 2002)
Trigonobalanus verticillata	3	0.153	0.153	0.787	0.194	(Kamiya *et al.* 2002)
mean SSR		**0.699**	**0.040**	**0.274**	**0.145**	
mean allozyme		**0.255**	**0.040**	**0.717**	**0.054**	
mean AFLP/RAPD		**0.247**	**0.088**	**0.701**	**0.124**	

Isolation by elevation

Significant isolation by distance was found for the neutral loci and locus Ccu97H18 (only for pairwise F_{ST}). However, additionally significant isolation by elevation was detected in both the potentially adaptive locus and the non-adaptive loci after accounting for the effect geographic distance. This pattern of isolation by elevation suggests higher rates of gene flow among sites at similar elevations than along elevational clines (Byars *et al.* 2009). Elevation can result in reproductive isolation due to phenological shifts, e.g. delayed budding (Rusch 1993) or shift of flowering time or prolonged floral longevity and stigma receptivity (Blionis *et al.* 2001) resulting in temporal separation of the timing of flowering (Borchert 1983). Phenological differences in flowering time in turn will lead to partial reproductive isolation which both may facilitate adaptation to elevation and lead to neutral genetic differentiation as has been shown for other forest trees (Kraj & Sztorc 2009).

Populations of *C. eyrei* at the top of the mountains harboured the largest amount of genetic variation whereas populations at lower elevation had reduced levels of variation. Although not often observed among trees (Ohsawa & Ide 2008) a similar pattern was found in other tree species (Jump *et al.* 2006; Maghuly *et al.* 2006; Peng *et al.* 2003). As both the non-selected loci and the putatively selected locus displayed the same pattern, a number of non-mutually exclusive processes may have contributed. First, mutation rate may be higher at higher elevations due to increased ultraviolet-B radiation (Ohsawa & Ide 2008). If effective, this process should apply to all loci in a similar manner and may have contributed to the general trend across all loci. However, microsatellites are polymorphic due to slippage mutation of the DNA polymerase and UV radiation seems not necessarily to affect this process (Clarke *et al.* 2008; Jackson *et al.* 1998). Second, human disturbance is much stronger at lower elevations. Charcoal has been detected in many local soil profiles (Bruelheide *et al.* 2011) indicating past fire clearance. Populations at higher elevations are more rarely influenced by human activities and, thus, are able to preserve genetic diversity. We found a significant positive correlation between elevation and successional stage, indicating that older, less disturbed forests are often located in higher elevations. Hence, it is likely that undisturbed upland forests served as sources for colonization after logging at low elevations. Recent and older bottleneck and founder effects may thus have contributed to reduced variation at lower elevations. However, bottleneck tests did not support the hypothesis of recent reductions of population size at lower elevation. However, in wind pollinated trees, large gamete pools may be involved in colonization, maintaining high levels of diversity in colonized areas (e.g. (Poitti *et al.* 2009)). Third, selection may be a significant force. Locus Ccu97H18 showed a

strong cline as the most common allele at low elevations almost went extinct at higher elevations and many other alleles appeared instead. These patterns are unlikely due to random genetic drift or restricted gene flow, but most likely due to selection. Since Ccu97H18 is a short microsatellite, genetic hitchhiking is the most probable reason for the observed patterns, assuming that the locus is linked with loci under selection, as has been shown for other microsatellite loci in trees (Edelist *et al.* 2006; Ingvarsson 2010; Stefenon *et al.* 2008). We do not have evidence on the genes potentially involved. Thus, both the target of selection and the potential contribution of diversifying selection producing the cline and/or balancing selection bearing high allelic diversity remain obscure. Overall, the study suggests that elevation can be an efficient driver for genetic differentiation at both neutral and adaptive loci at the landscape scale.

Materials and methods

Ethics Statement

Field work and the collecion of leaves were performed in cooperation with and under approval by the Gutianshan National Nature Reserve in China.

Study area and populations

Our study was carried out in Gutianshan National Nature Reserve (NNR) located in the western part of Zhejiang Province, China (29º8' - 29º17' N, 118º2' - 118º11' E). *C. eyrei* is the dominant tree species in the area and continuously distributed throughout (Xu *et al.* 2005). The Gutianshan NNR has an area of approximately 81 km^2 with elevations ranging from 250 to 1250 m a.s.l.. It mainly consists of species-rich evergreen broad-leaved forests including old growth forest and successional stages that developed after cease of human use in 1975 (Bruelheide *et al.* 2011). In 2008, we sampled 24 representative sites of 30×30 m which were spread across all successional stages and the local elevational range of the species (251-920 m). We did not sample at >1000 m a.s.l. because the species was too rare. Five successional stages were distinguished according to the average age of the tree layer ((Bruelheide *et al.* 2011), 1: < 20 yrs, 2: < 40 yrs, 3: < 60 yrs, 4: < 80 yrs, 5: ≥ 80 yrs). Additional details of site selection and conditions for 20 of the sites are reported in Bruelheide *et al*. (2011). We sampled all mature individuals of *C. eyrei* inside the sites and some additional individuals outside of sites CSP 6 and CSP 21 because there were too few inside. In each of the 24 populations, 12 to 49 individuals (mean = 24) were collected, totalling 583 individuals (Table 3.1).

Population genetic analysis

Total genomic DNA was isolated from ca. 10 mg dried leaf material using the DNeasy 96 plant extraction kit (QIAGEN) following manufacturers instructions. Samples were genotyped at eight microsatellite loci previously developed for *C. cuspidata* var. *sieboldii* (Ueno *et al.* 2000, 2003). Multiplex polymerase chain reactions (PCR) were performed in a total volume of 10 µl. Ccu16H15 (Label: PET, redesigned reverse primer: GAAATTGAGT TGGGTTAGTTCCAC), Ccu28H18 (FAM), Ccu62F15 (NED), Ccu33H25 (FAM), Ccu90T-17 (PET), Ccu102F36 (VIC) and Ccu87F23 (FAM) were in one PCR amplification. Another single PCR amplification was performed for Ccu97H18 (VIC). Thermocycler protocol was one cycle of 95 °C for 15 min, followed by 35 cycles of 30 s at 94 °C, 90 s at 58 °C and 1 min at 72 °C, with a final extension of 20 min at 72 °C. PCR products from the two amplifications were mixed and separated on an ABI 3130 genetic analyzer (Applied Biosystems) with internal size standard GeneScan^{TM} 600 LIZ. Individuals were genotyped using GeneMapper version 3.7 (Applied Biosystems). *C. eyrei* is diploid and none of the samples displayed more than two alleles.

Because the study species is a wind pollinated perennial tree of the Fagaceae, many of which exhibit populations in Hardy-Weinberg equilibrium (HWE, e.g. (Cheng *et al.* 2006; Hiraoka & Tomaru 2009a)), we assumed that *C. eyrei* microsatellite loci should conform to HWE. Because trans-species amplification of microsatellites often results in null alleles we checked the data for the presence of null alleles under the assumption of HWE using MICRO-CHECKER (Van Oosterhout *et al.* 2004). Except for Ccu16H15, all loci showed signs of null-alleles, the frequency of which ranged from 1.3 % to 20 % (mean = 6.99 %). We took three approaches to deal with null-alleles. First, we adjusted homozygous single locus genotypes, if necessary, by adding an additional allele in the frequency of the null-allele. This approach assumes that there is one single null allele, which is treated as an additional allele. Second, we used the null allele correction procedure of the FreeNA software (Chapuis & Estoup 2007) to calculate pairwise F_{ST} values. This approach corrects for null alleles but restricts the analysis to the visible alleles. Third, we left data unchanged. However, all subsequent analyses showed similar results irrespective of the type of null allele treatment. Therefore, we only present the results of the MICRO-CHECKER approach as it allows the calculation of standard diversity descriptors.

We tested the eight microsatellite loci for outliers, i.e. markers potentially under selection, using the program FDIST (Beaumont & Nichols 1996). A null distribution of target F_{ST} values expected from a neutral model is generated and quantile limits are calculated. Loci

outside a 99% confidence interval are regarded as potentially under selection. Following Acheré *et al*. (2005), the neutral expectation was first based on the observed overall mean F_{ST} calculated from all markers. In a second step, the overall mean F_{ST} was recalculated without the putatively selected locus and used as target value for a new null distribution to test the remaining loci. As our analyses suggested that locus Ccu97H18 was potentially under directional selection, we performed all the following analyses both for the seven loci conforming to a neutral model (" neutral loci") and only for locus Ccu97H18.

Genetic diversity at population level was characterized by number of alleles (A), allelic richness (A_R, correcting for sample size by rarefaction for a minimum sample size of 12) and expected heterozygosity (H_E) using FSTAT version 2.9.3.2 (Goudet 1995). Because genotypes were adjusted for null-alleles, we did not calculate inbreeding coefficients. In the dataset of neutral loci we tested for recent bottlenecks (reductions of effective population size) by testing for an excess of heterozygosity relative to expectations of a mutation–drift equilibrium (Cornuet & Luikart 1996). We used the software BOTTLENECK (Piry *et al.* 1999) and applied the recommended two-phase mutation model (TPM) with 70% stepwise and 30% multistep mutations, a variance of 12, 1000 iterations in the coalescent simulations and one-tailed Wilcoxon´s signed-rank tests. To assess population differentiation, pairwise F_{ST} values based on Weir and Cockerham's (Weir & Cockerham 1984) estimator θ were calculated using FSTAT. As F_{ST} is likely to underestimate genetic differentiation between populations for markers which show high levels of allelic variability, we calculated F'_{ST}, a standardized parameter of genetic differentiation as $F'_{ST} = F_{ST} / F_{ST\max}$ (Hedrick 2005). $F_{ST\max}$ was calculated after recoding the data using RECODEDATA (Meirmans 2006). To test for isolation by distance (Wright 1943), the association between pairwise genetic differentiation (F_{ST}) and pairwise geographic distances (log transformed) was tested using the Mantel test implemented in R 2.8.1 (R Development Core Team 2008). We also performed a Mantel test between F_{ST} and pairwise elevational differences (log transformed) to test for isolation by elevational distance. Since pairwise elevational difference was correlated to pairwise geographic distance, we performed partial Mantel tests to test for effects of elevation after accounting for geographic distance. Because allelic diversity differed between populations and was correlated with elevation, this may have biased the estimates of pairwise differentiation using F_{ST}. We therefore calculated standardized pairwise F_{ST} values (pairwise F'_{ST}, (Hedrick 2005), eqn. 4b) and repeated the tests for isolation by distance and isolation by elevation. In order to compare the genetic differentiation of *C. eyrei* with other species of the

Fagaceae, we reviewed empirical studies and calculated F'_{ST} following Hedrick ((Hedrick 2005), eqn. 4b).

Statistical analysis

To test the effects of environmental factors on genetic variation, we analysed the relationship between allelic richness (A_R) and the two predictors elevation and successional stage in a multiple regression. We used A_R, which corrects for sample size, rather than H_E, because sample size varied among populations; however, the results were qualitatively the same for H_E. Collinearity of elevation and successional stage was checked by Spearman correlation. All analyses were performed with R 2.8.1 (R Development Core Team 2008).

Acknowledgments

We thank Xueqin Zeng, Xiaoyan Wang and Teng Fang for sample collecting and the BEF-China team for logistic help.

Chapter 4 - Phylogeography of a widespread Asian subtropical tree: East-west differentiation and climate envelope modelling suggest multiple glacial refugia

Miao-Miao Shi, Stefan Michalski, Erik Welk, Xiao-Yong Chen and Walter Durka

Journal of Biogeography (Under review)

ABSTRACT

Aim Climate models predict that during the Last Glacial Maximum (LGM), the subtropical zone of Eastern Asia was reduced to a narrow southern belt. In contrast, previous phylogeographic studies of subtropcial plant species, many of which were rare species, indicated different glacial refugial areas north of this predicted area. Here, we aim to elucidate the phylogeographic structure and putative refugial areas of *Castanopsis eyrei*, a widely distributed tree dominant in subtropical evergreen broadleaved forests.

Location Subtropical China.

Methods We compiled distribution data and employed climate envelope model projections for the LGM. Microsatellite data and chloroplast DNA (cpDNA) sequence data were obtained for 31 populations sampled throughout the species range. Microsatellites were analysed with Bayesian clustering. For cpDNA sequence data relationships among haplotypes were depicted in a statistical parsimony network. We analysed patterns of variation within and among populations and clusters and along latitudinal clines.

Results Modelling revealed a potential distribution of *C. eyrei* mainly in a narrow belt along the southern coastline. Nuclear microsatellites revealed two clusters corresponding to a split between the western and eastern range, and a south-north decline of genetic variation. The eastern cluster harboured significantly higher nuclear genetic diversity. Sixteen, closely related cpDNA, haplotypes formed a "star-like" network. Populations were strongly differentiated at cpDNA, but lacked phylogeographic structure. Both data sets revealed higher genetic differentiation in the western than the eastern cluster.

Main conclusions Our results provide evidence for at least two putative refugial regions during the LGM and a postglacial re-colonization from the South. Inconsistency between nuclear and chloroplast data might be attributed to ancestral polymorphism of cpDNA and chloroplast capture, but does not contradict the hypothesis of multiple refugia. Our results are

likely to represent a template for evolutionary history and phylogeography since the LGM for wide spread subtropical species in this region.

Keywords

Castanopsis eyrei, chloroplast capture, genetic structure, glacial refugia, microsatellites, phylogeography, subtropical China, widespread species.

INTRODUCTION

The present-day distribution of plant populations is determined not only by the current environment, but also by past events like Quaternary climate cycles (Avise 2000). The repeated drastic climate changes of Quaternary have resulted in multiple contraction-expansion processes of many species and profoundly shaped their phylogeographic patterns and current genetic structure (Hewitt 2004). Phylogeographic studies can shed light on the effects of climate changes on species distributions and help to unravel complex historic processes (Avise 2000; Hickerson *et al.* 2010). Many phylogeographic studies have been performed in species of the temperate zone, in particular in Europe and North America (Avise 2000; Heuertz *et al.* 2004a; Taberlet *et al.* 1998), as well as in Asia (Cheng *et al.* 2005; Hiraoka & Tomaru 2009a; Hu *et al.* 2010; Huang *et al.* 2002; López-Pujol *et al.* 2011). However, in contrast to their highly diverse flora, subtropical areas have not been intensively studied. The Asian subtropics are one of the global biodiversity hotspots (Barthlott *et al.* 2005) and are considered to be one of the most important refugial regions for lineages that evolved prior to the late Tertiary and Quaternary glaciations (Axelrod *et al.* 1996). However, the evidence for effects of Quaternary climate changes on the phylogeography of organisms in the Asian subtropics is surprisingly limited (Qiu *et al.* 2011).

In China, the zone ranging between the Qinling Mountains and Huai River (34°N) and the tropical South (22°N) is characterized by a subtropical climate (Qiu *et al.* 2011) and a vegetation of evergreen broadleaved forests. This region has never been covered by large ice sheets during the Last Glacial Maximum (LGM) (Hewitt 2000), but nevertheless underwent particularly complex changes of climate and vegetation distributions throughout the last ice-age cycles (Qiu *et al.* 2011). At the LGM, the climate was cooler by 4-6 °C than today (Zheng 2000), and especially dryer by ca. 400-600 mm/year (Qiu *et al.* 2011). According to climate models and derived LGM biome maps, evergreen broadleaved forests were forced to retreat southwards into the current tropical zone (Harrison *et al.* 2001; Qiu *et al.* 2011; Sun *et al.* 1999). Thus the potential refugia for evergreen broadleaved forests are expected to be located

in a narrow range in the most southern mainland of China (< ca. 24˚N, Qui *et al*., 2011). After the LGM, species should have moved northward again from southern refugia and should reveal signs of northward expansion. However, large scale climate models can not resolve smaller scale climatic discontinuities and local climatic abnormalities e.g. in mountain ranges. The use of species distribution models (SDMs) in combination with downscaled paleoenvironmental reconstructions allows to locate and characterise potentially refugial areas of particular species. In general, the development of SDM assumes species' current large scale geographical distribution to be largely in equilibrium with the environment. A further precondition to the application of ecological niche models (ENMs) for predicting past distribution is a certain degree of niche conservatism (Peterson 2003). SDMs relate current presence data of species to spatial data of environmental variables to infer models of environmental requirements. Successfully parameterized models that are projected onto paleoclimate reconstructions might allow identifying species specific past potential distributions.

A number of empirical studies revealed multiple isolated refugial areas within subtropical China, that are located outside of the regions of predicted subtropical vegetation, mainly in the mountain regions of the Yunnan-Guizhou (Yungui) Plateau (Shen *et al.* 2005) and the Nanling Mountains and the far east of subtropical China, e.g. the Tianmu Mountains (Gao *et al.* 2007; Gong *et al.* 2008; Wang *et al.* 2009; Yan *et al.* 2007; Zhou *et al.* 2010). However, most of these studies have either focused on endangered species with narrow distribution ranges, or on coniferous species. For example, Wang and Ge (2006) suggested at least four separate glacial refugia for the endangered *Cathaya argyrophylla* (Pinaceae). For *Pinus kwangtungensis*, Tian *et al.* (2010) suggested three major refugia, located in the vicinity of the Yungui Plateau and the Nanling Mountains. In order to reveal a more complete and general pattern of the phylogeography of the subtropical flora, common and widespread keystone species of the subtropical biomes should be analysed.

Maternally inherited chloroplast DNA (cpDNA) is one of the most commonly used markers to study the phylogeography of angiosperms, because of lack of recombination, smaller effective population size and exclusive gene flow by seeds (Avise 2000; Hewitt 2004; Petit *et al.* 2003a; Taberlet *et al.* 1998). However, cpDNA represents only a single gene genealogy, therefore can hardly capture all major events in a species history (Heuertz *et al.* 2004b). Alternatively, biparentally inherited nuclear DNA markers, particularly nuclear microsatellites, allow for recombination, represent high levels of polymorphism, and thus are increasingly used to reveal the effects of climate changes on genetic pattern of organisms (Bai

et al. 2010; Hu *et al.* 2010). Heuertz *et al.* (2004b) detected complex geographical patterns and cryptic refugia using microsatellites in European populations of *Fraxinus excelsior*, which were not detected by cpDNA (Heuertz *et al.* 2004a). Thus, the use of molecular markers derived from both nuclear and organelle DNA can provide a more complete description of population structure and insights into population history and dynamics (Petit *et al.* 2005).

Castanopsis eyrei (Champion ex Bentham) Tutcher (Fagaceae) is one of the dominant tree species in late successional evergreen broadleaved forests in subtropical China. It is monoecious and pollinated by wind and insects. The acorn seeds are predominantly dispersed by gravity and small rodents (Li & Jin 2006) presumably leading to spatially more restricted gene flow by seed than by pollen. The geographical distribution of *C. eyrei* is mainly restricted to the area south of Yangtze River in mainland China and Taiwan, occurring frequently in Southeast China and more scattered in the Southwest due to unsuitable karst habitat (Fig. 4.1). It occurs from 300 m to 1700 m a.s.l (Flora of China 1999). Within its distribution range, a few scattered mountain ranges with altitudes over 2000 m a.s.l. are located in the Wuyi and Nanling Mountains (Liu *et al.* 2003). The western part of the distribution range is characterized by a more complex topography with numerous mountains, whereas the eastern part is much flatter with fewer high mountains. The distribution of *C. eyrei* in the southwest is less continuous, and populations are spatially more isolated. Thus, gene flow among populations, especially seed dispersal, is likely to be more obstructed in the West, and higher genetic differentiation is expected.

In this study, we used climatic envelope modelling combined with downscaled high-resolution estimates of LGM climate parameters, maternally inherited cpDNA as well as biparentally inherited nuclear microsatellite loci, to investigate the phylogeography of *C. eyrei*. In particular we tested whether a single southern or several northern glacial refugia can be proposed for *C. eyrei*. Therefore we ask (1) Which areas were climatically suitable for *C. eyrei* during the LGM? (2) Is there evidence for differentiated gene pools indicative for multiple refugia? (3) Are there clines of genetic diversity indicative of refugia and of postglacial colonization?

MATERIALS AND METHODS

LGM distribution model

To identify potential glacial refugia and their spatial arrangement as well as to evaluate and complement the phylogeographic analyses, we calibrated a climate based distribution model

on the present native distribution range of the species that describes the present climatic niche of *C. eyrei* and projected it onto a scenario of the Pleistocene's last glacial maximum (LGM) about 21,000 years ago.

The climatic niche model of *C. eyrei* was parametrized in East Asia using county based presence records for the species and environmental data for the present and the LGM managed as geographic information system (GIS) map layers. Occurrence information of *C. eyrei* was compiled using complementary data sources. Specimen location data was obtained from the Global Biodiversity Information Facility (GBIF 2010) and the Chinese Virtual Herbarium (CVH 2010). Additionally, GIS datasets of county occurrence data for China were obtained from Fang *et al.* (2010). For the modeling, we ended up with 197 unique records with an average nearest neighbor distance of ca. 50 km. These localities sample the entire distribution range of the species (Fig. 4.1).

Past and present climatic attributes at these locations (available for 19 bioclimatic variables in the WorldClim database (version 1.4; www.worldclim.org; Hijmans *et al.* 2005) were obtained using the 'Extract Values by Points' function in DIVA-GIS 7.3.0 (Hijmans *et al.* 2001). The bioclimatic datasets were represented by climate layers of a spatial resolution of 2.5 arc minutes (0.25' x 0.25' grid cells).

To develop climate-based distribution models, a simple envelope algorithm was applied in a software environment developed in Visual Basic and C^{++} by E.W. For each environmental variable, the algorithm finds lower and upper limits over all occurrence sites. During model projection, the climatic suitability is computed as the ratio of the number of layers within an optimized minimum–maximum threshold to the total number of layers. Interval threshold limits were estimated by inferentially optimizing the trade-off between omission and commission errors to maximize model accuracy to the highest possible AUC value (area under the receiver operating characteristics curve) (Jiménez-Valverde 2012). The maximized model fit is based on 100-fold cross-validation with a random re-sampling of 50% of the input training data set aside to serve as test data.

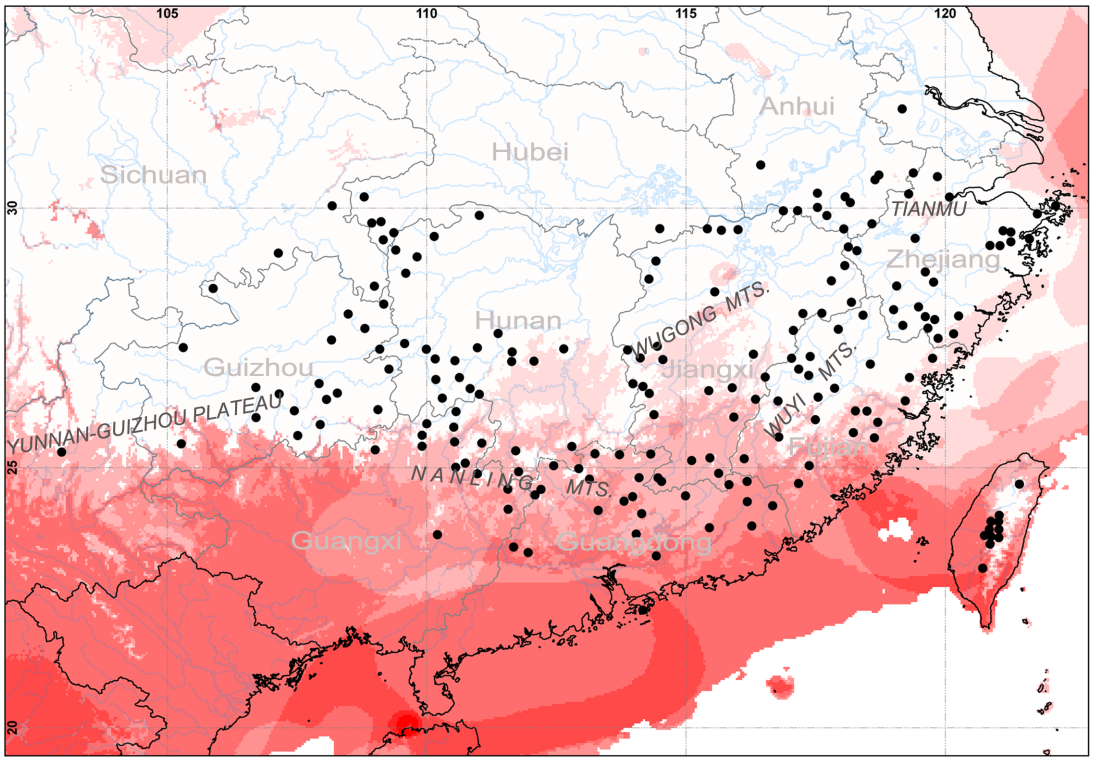

Fig. 4.1 Current distribution of *Catstanopis eyrei* and predicted distribution during the Last Glacial Maximum (LGM, 21,000 BP). The dots indicate specimen locations from the Global Biodiversity Information Facility (GBIF 2010) and the Chinese Virtual Herbarium (CVH 2010) representing the current distribution. Red shading represents the predicted distribution, dark corresponding to high scores.

Population sampling

We selected 31 natural populations of *C. eyrei* across the species range (Table 4.1; Fig. 4.2a). The populations covered a wide latitudinal (22˚ - 30˚ N) and longitudinal range (108˚ - 121˚ E) and were separated by distances between 24 and 1328 km (average nearest neighbour distance: 104.9 km). In each population, leaves from between 2 and 31 individuals were collected and dried with silica gel. As outgroups for the cpDNA analysis, we sampled two individuals each of the sympatric congeners *C. fargesii* Franchet, *C. carlesii* (Hemsley) Hayata, *C. tibetana* Hance and *C. sclerophylla* (Lindley & Paxton) Schottky in Gutianshan nature reserve (29º10´19´´N, 118º03´50´´E).

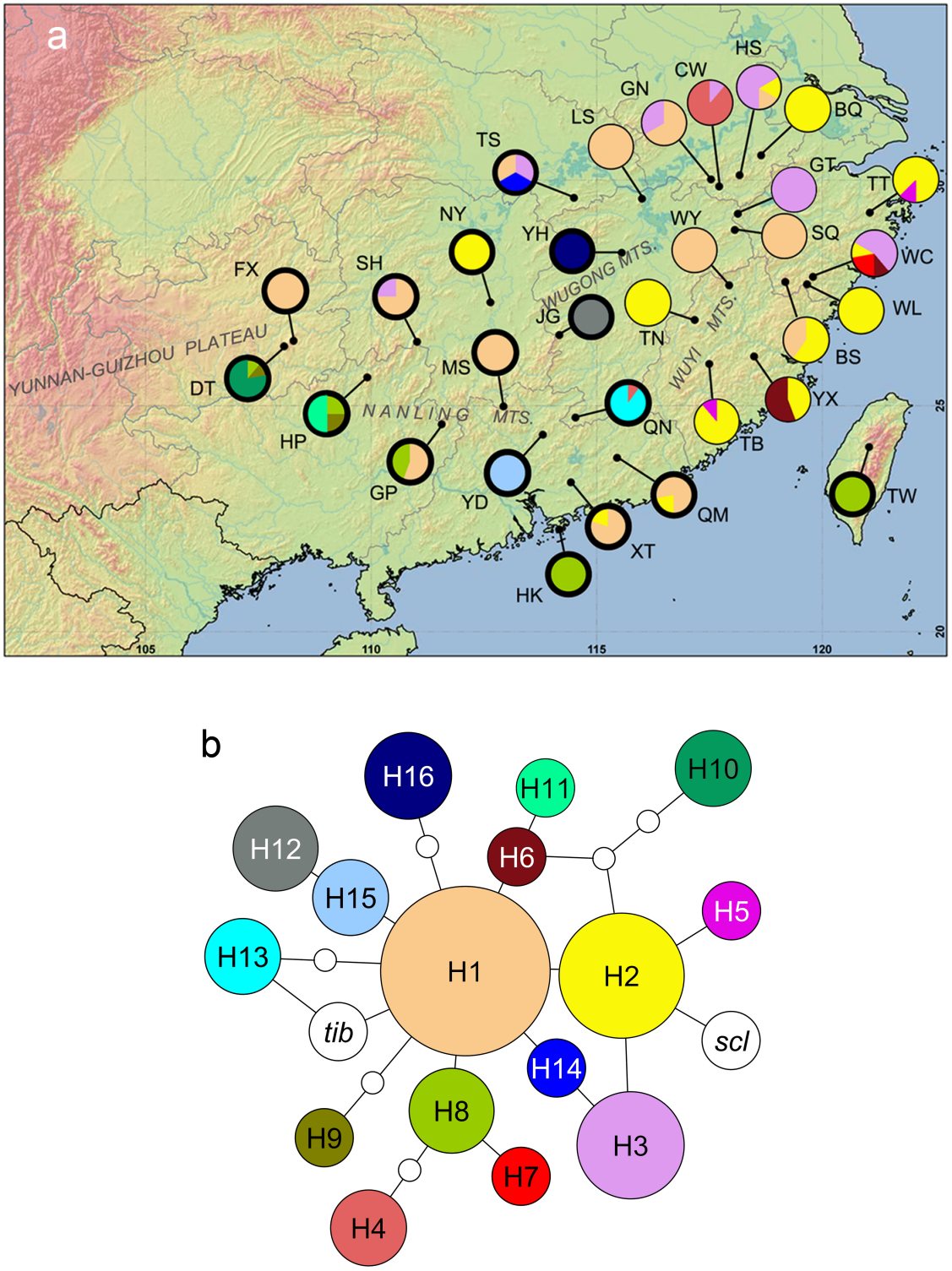

Fig. 4.2 (a) Locations of *Castanopsis eyrei* populations sampled in this study and the geographical distribution and frequency of chloroplast DNA (cpDNA) haplotypes H1 to H16. Colours of haplotypes correspond to those in (b). Bold and thin outlines indicate the two genetic clusters, western and eastern, respectively, resulting from STRUCTURE analysis of microsatellites (see Fig. 4.3). (b) Statistical parsimony network of cpDNA haplotypes. Size of

the circle corresponds to haplotype frequency. Open circles indicate one hyplotype. *tib* and *scl* correspond to haplotypes of *C. tibetana* and *C. sclerophylla*, respectively.

DNA extraction and genotyping

DNA was extracted with the DNeasy 96 plant extraction kit (QIAGEN) from 10 mg dried leaf material. Samples were genotyped at 8 microsatellite loci following Shi *et al.* (2011a). A first screening for polymorphisms at 15 cpDNA intergenic spacer regions was done using 8 individuals from 4 geographically distant populations. Finally, we chose two regions for further analysis, which could be sequenced successfully in both directions and showed variation, *trn*T - *trn*L (Taberlet *et al.* 1991) (redesigned reverse primer: 5'-TCGAAGATCCAGAGTTGATCC -3') and *pet*G - *trn*P (Wang *et al.* 2003). A subsample of 263 individuals from 31 *C. eyrei* populations was sequenced, as well as the 8 outgroup samples. Around 10 individuals were analysed in most populations, whereas only 2 - 3 individuals were available in 4 populations (NY, HK, TS, and TW) (Table 4.1). Polymerase chain reactions (PCR) were performed in a total volume of 10 µl. The amplification conditions were initial denaturing of 15 min at 95 °C followed by 35 cycles of 30 s at 94 °C, 90 s of annealing at 63 °C for *trn*T - *trn*L and 62 °C for *pet*G - *trn*P, 1 min of elongation at 72 °C, ending with a 10 min extension at 72 °C. The PCR products were checked on 1.5% agarose gels. DNA was sequenced in both directions using BigDye Terminator v 3.1 cycle sequencing (95 °C, 3 min; 30 cycles of 20 s at 95 °C, 15 s at 50 °C, 4 min at 60 °C; 60 °C, 15 min) and separation on an ABI 3130 sequencer (Applied Biosystems, Foster City, California, USA).

Table 4.1 Details of population locations, sample sizes and parameters of genetic diversity at each population of *Castanopsis eyrei* in China.

Sampling sites	ID	Location coordinates		SSR				cpDNA			
		Longitude (°E)	Latitude (°N)	*n*	*A*	A_R	H_E	*n*	Haplotypes	*h*	$\pi\times10^3$
Datang Town	DT	108.06860	26.31821	22	9.9	7.1	0.78	9	H8 (1), H9 (1), H10 (7)	0.417	2.233
Fangxiang Town	FX	108.27875	26.43920	15	8.8	7.4	0.81	8	H1 (8)	0	0
Huaping Nature Reserve	HP	109.91148	25.62730	30	10.6	6.5	0.71	8	H8 (2), H9 (2), H11 (4)	0.714	2.441
Shunhuangshan	SH	111.01667	26.41667	19	4.0	3.5	0.57	8	H1 (6), H3 (2)	0.429	0.862
Guposhan	GP	111.56500	24.59111	25	11.9	7.6	0.75	9	H1 (5), H8 (4)	0.556	0.559
Nanyue	NY*	112.64250	27.29167					3	H2 (3)	0	0
Mangshan Nature Reserve	MS	112.93277	24.98517	18	11.0	8.4	0.84	9	H1 (9)	0	0
Yingde	YD	113.80696	24.36038	24	11.3	7.4	0.79	9	H15 (9)	0	0
Hongkong	HK*	114.16708	22.26600					2	H8 (2)	0	0
Jinggangshan	JG	114.16976	26.56771	30	14.8	9.0	0.84	10	H12 (10)	0	0
Xiangtoushan	XT	114.42689	23.30767	23	13.9	9.3	0.86	10	H1 (8), H2 (2)	0.356	0.358
Tongshan County	TS*	114.49433	29.60207					3	H1 (1), H3 (1), H14 (1)	1.000	1.341
Quannan Maoshan Forestry	QN	114.53013	24.74236	21	5.8	4.8	0.70	10	H4 (1), H13 (9)	0.200	1.005
Qimuzhang Nature Reserve	QM	115.45178	23.84994	29	9.1	6.5	0.79	9	H1 (7), H2 (2)	0.389	0.391
Yuhuashan	YH	115.55707	28.38932	13	5.8	5.3	0.70	13	H16 (13)	0	0
Lushan	LS	116.00653	29.58381	31	11.6	7.8	0.80	13	H1 (13)	0	0
Taining	TN*	117.17575	26.90018					7	H2 (7)	0	0
Tianbaoyan Nature Reserve	TB	117.49645	25.92735	15	11.0	8.9	0.89	9	H2 (8), H5 (1)	0.222	0.224
Guniujiang Nature Reserve	GN	117.53612	30.01086	19	9.6	7.8	0.84	9	H1 (6), H3 (3)	0.5	1.006
Chawan Nature Reserve	CW	117.71739	29.85405	22	10.0	7.3	0.81	9	H3 (1), H4 (8)	0.222	1.118
Wuyishan	WY	117.93962	27.66782	29	11.5	7.5	0.80	10	H1 (10)	0	0
Sanqingshan	SQ	118.06308	28.89072	29	13.9	8.6	0.86	10	H1 (10)	0	0
Gutianshan Nature Reserve	GT	118.13333	29.25000	30	12.4	8.1	0.83	8	H3 (8)	0	0
Huangshan	HS	118.16889	30.10503	22	10.8	7.6	0.83	12	H1 (2), H2 (2), H3 (8)	0.546	0.793
Youxi County	YX	118.49098	26.09329	29	15.0	9.2	0.87	9	H2 (4), H6 (5)	0.556	1.118
Banqiao Nature Reserve	BQ	118.64900	30.54232	19	8.8	6.6	0.71	8	H2 (8)	0	0
Baishanzu Nature Reserve	BS	119.18281	27.74117	31	13.6	8.3	0.83	10	H1 (6), H2 (4)	0.533	0.538
Wuyanling Nature Reserve	WL	119.66487	27.68931	25	12.3	8.4	0.85	9	H2 (9)	0	0
Wencheng County	WC	119.79669	27.85350	20	11.9	8.5	0.84	9	H2 (1), H3 (5), H6 (1), H7 (2)	0.694	2.068
Taiwan	TW*	121.03903	24.08920					3	H8 (3)	0	0
Tiantai	TT	121.05611	29.27500	24	10.3	7.3	0.79	8	H2 (7), H5 (1)	0.250	0.252
Mean				23.6	10.7	7.5	0.79	8.5		0.245	0.526
overall				614	28.1	10.1	0.80	263		0.814	1.750

n: sample size; *A*: number of alleles per locus; A_R: allelic richness based on 13 samples; H_E: expected heterozygosity; *h*: haplotype diversity; π: nucleotide diversity; * populations with small sample size ($n < 10$) were excluded when analyzing microsatellite data. Values in parentheses indicate the frequency of each haplotype.

Data analysis

Genetic diversity at microsatellite loci

Populations with small sample size ($n < 10$) were excluded from the microsatellite analysis to minimize sample size effects, restricting genetic analyses to 26 populations (Table 4.1). We calculated the total number of alleles detected per locus (A), genetic diversity in the total population (H_T) and average gene diversity within populations (H_S) (Nei 1987) across all populations at each locus and over all loci to assess the level of genetic diversity over all populations. In each population and genetic cluster (see below), genetic diversity was characterized by number of alleles per locus (A), allelic richness (A_R, correcting for sample size by rarefaction for a minimum sample size of 13) and expected heterozygosity (H_E). These calculations were performed using FSTAT v. 2.9.3.2 (Goudet 1995). We compared genetic diversity parameters (A_R and H_E) between genetic clusters by a randomization procedure implemented in FSTAT. The two-sided P-values were obtained after 1,000 permutations. To determine whether genetic variation within populations was correlated with geographical gradients, we performed a multiple regression with backward elimination of geographic factors (latitude, longitude and cluster factor: western or eastern) in R 2.8.1 (R Development Core Team 2008).

Genetic differentiation at microsatellites

We determined levels of genetic differentiation among populations using θ as an estimate of F_{ST} (Weir & Cockerham 1984) in FSTAT and the standardized genetic differentiation $F'_{ST} = F_{ST}/F_{ST\,max}$ (Hedrick 2005). $F_{ST\,max}$ was calculated with FSTAT after recoding the data using RECODEDATA v. 0.1 (Meirmans 2006). The significance of differences of genetic differentiation (F_{ST} and F'_{ST}) between clusters were assessed by permuting populations (1,000 repeats) among clusters. To evaluate isolation by distance (Wright 1943), the associations between pairwise genetic differentiation (F'_{ST}) and pairwise log geographical distances were tested using a Mantel test. These analyses were performed in R.

Identification of genetic clusters at microsatellites

Samples from all populations were involved in this analysis (total N = 629), except the three individuals from NY, because they failed to be amplified successfully at microsatellites. To determine the phylogeographical genetic structure, Bayesian clustering was employed using STRUCTURE v.2.3.3 (Pritchard *et al.* 2000). This program assigns individual genotypes to K gene pools based on their allelic frequencies and estimates the posterior probability of the data

given *K* assuming Hardy-Weinberg equilibrium. An admixture model was run with correlated allele frequencies. Each run was pursued for 50,000 MCMC cycles, with an initial burn-in of 10,000. To estimate the most probable *K* number of genetic clusters and the ancestry membership coefficients of each individual in these clusters, the algorithm was run 10 times for each *K* value from 1 to 10. The mean log-likelihood at each value of *K,* [ln Pr(*X*|*K*)], was plotted, and an *ad-hoc* statistics Δ*K*, as suggested by Evanno *et al.* (2005), was used to estimate the most likely number of clusters. From the whole dataset, two clusters were identified. To detect further genetic structure, we repeated the STRUCTURE analysis for separate data sets of these two clusters. For further analyses populations were assigned to clusters based on the population level membership coefficient.

Diversity and differentiation of cpDNA

All DNA sequences were deposited in Gen Bank under the accession numbers JX215141 - JX215196 (*trn*T - *trn*L) and JX215197 - JX215239 (*pet*G - *trn*P). They were combined and aligned in BioEdit v.7.0.5.0 (Hall 1999). Three mononucleotide repeats were removed from the alignment because they could not be aligned unequivocally. Insertions and deletions (indels) were treated as single mutation events. We calculated haplotype diversity (*h*) and nucleotide diversity (π) for each population using ARLEQUIN v.3.5 (Excoffier & Lischer 2010), in which also pairwise F_{ST} values were computed. Patterns of isolation by distance were evaluated by associating pairwise F_{ST} values with pairwise log geographical distances in Mantel tests. We also computed the mean within-population gene diversity (H_S) and the total gene diversity (H_T) based on unordered alleles, and also the equivalent parameters (V_S and V_T) based on ordered alleles, as well as differentiation G_{ST} for unordered and N_{ST} for ordered alleles by PERMUT v.1.0. A higher N_{ST} than G_{ST} usually indicates the presence of phylogeographic structure (Pons & Petit 1996). The comparison of G_{ST} versus N_{ST} was conducted based on 2,000 random permutations. In order to visualize relationships among haplotypes a statistical parsimony network was computed with TCS (Clement *et al.* 2000).

RESULTS

LGM projection of current climate distribution model

For the current climate distribution model of *C. eyrei* the internal model validation revealed a reasonable high model fit with AUC = 0.88. The resulting current potential distribution range indicates that there are a number of lowland and hill counties in China (provinces Guizhou, Hunan, Jiangxi, Zhejiang, Fujian, Guangdong, and Guangxi) that are climatically suitable, yet

from where the species is currently not reported (result not shown). Projecting the optimized distribution model to the climatic conditions during LGM resulted in a prediction of potentially suitable habitats during the last ice age cold cycles (Fig. 4.1). The LGM potential distribution range is located at distinctly lower latitudes than most of the extant populations. Only the southernmost current occurrences are clearly located within the potential LGM range. Two regions that expand further northwards and, thus potentially supported a closer contact of current distribution and suitable climatic conditions during the LGM are NW-Guangxi/SW-Guizhou and SE-Fujian-Taiwan. Large parts of the potential LGM distribution range are currently below sea level due to postglacial transgression.

Genetic population structure at nuclear microsatellites

The Bayesian analysis of population structure with STRUCTURE revealed a most likely number of $K = 2$ clusters, each of which were subdivided further (Fig. 4.S1, Supporting Information). The two clusters largely corresponded to a split between the western/southern and the eastern/northern part of the distribution range, which will be referred to western and eastern cluster, respectively (Fig. 4.3). Some populations in the contact area of the two clusters showed comparatively high levels of mixture between gene pools (JG, QN). Both main clusters were further subdivided into two subclusters (Fig. 4.3). In the centre of the western cluster, in the western Nanling Mts., populations HP, SH and GP formed a clearly separated subcluster without much mixture or admixture. In contrast, the two gene pools of the eastern cluster were mixed in many populations (Fig. 4.3). Here, populations TB, SQ, YX and BS consisted mainly of one gene pool, however, without forming a contiguous geographic group (Fig. 4.3).

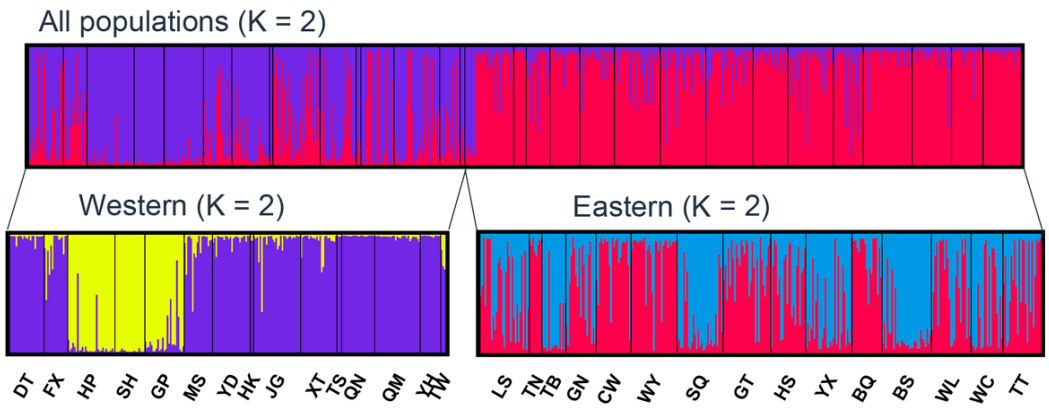

Fig. 4.3 Results of Bayesian cluster analysis with STRUCTURE based on microsatellite data of *C. eyrei*. Thin bars represent the cluster membership of 629 individuals in 30 populations at

K = 2 and further analyses of the western and eastern cluster at K = 2 (see Fig. 4.S1 for analysis of the most probable K).

Nuclear genetic diversity and differentiation

Across all microsatellite loci, high levels of gene diversity in the total population (H_T = 0.88) and within populations (mean H_S = 0.80) were observed (Table 4.1). Population differentiation was significant at each locus ($P < 0.05$), with the overall F_{ST} value for the multilocus estimate equal to 0.097 (range: 0.057 - 0.149 across loci). The standardized genetic differentiation (F'_{ST}) equalled 0.443 and thus was much higher than F_{ST}. Genetic diversity at population level (A_R and H_E) was significantly higher in the eastern than in the western cluster (Table 4.2). Accordingly, populations within the western cluster (F_{ST} = 0.122, F'_{ST} = 0.542) showed a significantly higher level of genetic differentiation than in the eastern cluster (F_{ST} = 0.052, F'_{ST} = 0.309; $P < 0.001$). The subclusters of the western cluster did not differ in genetic diversity (A_R: P = 0.29) and differentiation (F_{ST}: P = 0.784). In the eastern cluster, the subcluster including populations TB, SQ, YX and BS harboured significantly higher genetic diversity (A_R: P = 0.003) than the other subcluster, but did not differ in differentiation (F_{ST}: P = 0.262). In the multiple regression model of genetic variation, latitude and cluster factor were retained indicating that genetic variation was significantly different in two clusters (P = 0.001) and affected by latitude (P = 0.018). In the eastern cluster, allelic richness decreased significantly with latitude (r = -0.802, $P < 0.001$), while in the western cluster, the decreasing trend was not significant (r = -0.368, P = 0.238) (Fig. 4.4).

Table 4.2 Comparison of microsatellite genetic diversity within and differentiation among populations of *Castanopsis eyrei* in the two main genetic clusters identified in China. Significance after permutation test with 1,000 permutations.

Parameters	Western	Eastern	P
A_R	6.890	8.000	0.036
H_E	0.764	0.823	0.019
F_{ST}	0.122	0.052	< 0.001
F'_{ST}	0.542	0.309	< 0.001

A_R: allelic richness; H_E: expected heterozygosity; F_{ST}: Weir and Cockerham's genetic differentiation; F'_{ST}: standardized genetic differentiation; P value indicates the significance of difference between western and eastern group.

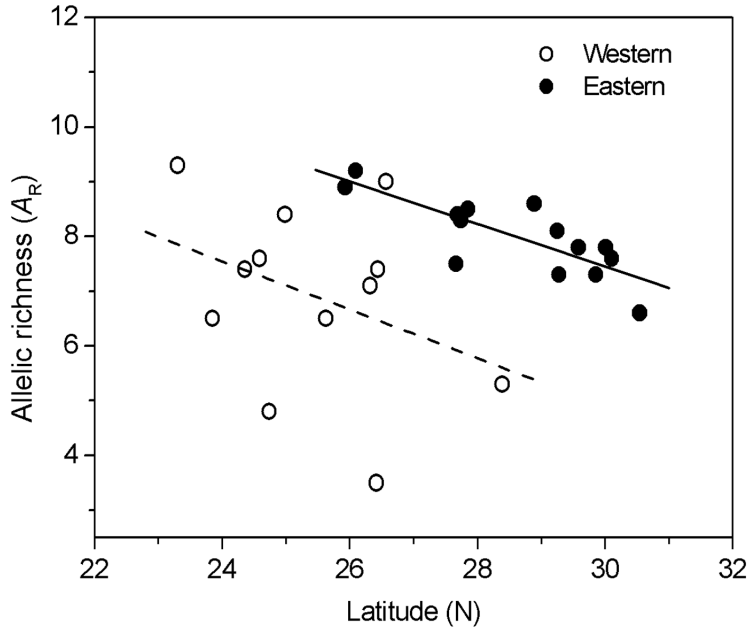

Fig. 4.4 The relationship between allelic richness (A_R) and latitude in the western (open circles, dashed line; r = -0.368, P = 0.238) and eastern cluster (filled circles, solid line; r = -0.802, $P < 0.001$).

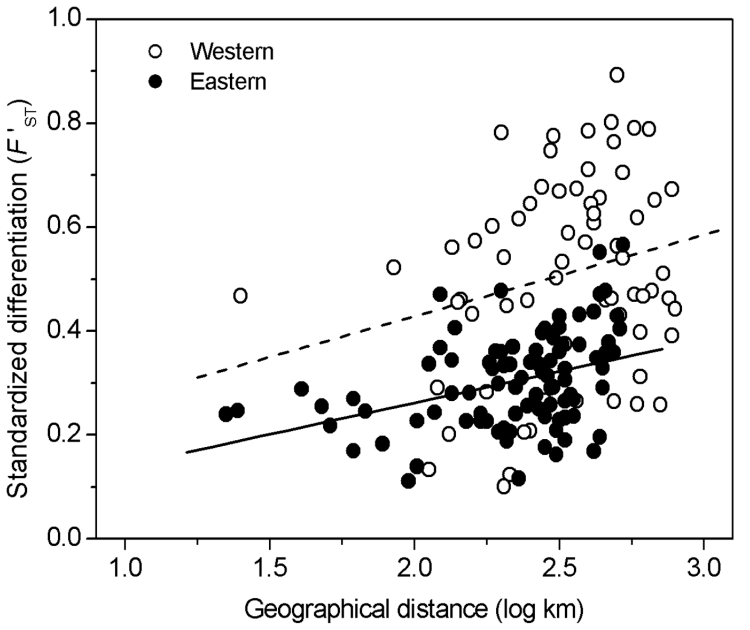

Fig. 4.5 Patterns of isolation by distance in the western (open circle, dashed line; r = 0.232, Mantel P = 0.049) and eastern cluster (filled circle, solid line; r = 0.376, Mantel P = 0.004). Population differentiation (standardized F'_{ST} values) as a function of log geographical distance.

To evaluate the spatial genetic structure, patterns of isolation by distance were tested. Significant isolation by distance was found when analysing all populations (r = 0.407, P = 0.001) and in both the eastern (r = 0.376, P = 0.004) and the western cluster (r = 0.232, P = 0.049) (Fig. 4.5).

Genetic variation at chloroplast DNA sequences

Alignment lengths in *C. eyrei* were 568 bp for *trn*T - *trn*L and 474 bp for *pet*G – *trn*P. In total, 19 polymorphic informative sites including 12 point mutations and 7 indels were detected within these two intergenic spacers, defining 16 haplotypes (Table 4.S1, Supporting Information). The geographic distribution of haplotypes is shown in Fig. 4.2a. Among 16 haplotypes, H1 was the most common one, followed by H2. Nine haplotypes (56.3%, H8 - H16) occurred only in the western cluster, while three (18.7%, H5 - H7) appeared only in the eastern cluster. Eight haplotypes were restricted to a single population, seven of which (87.5%) were found in the western cluster, particularly in the southern part of the Nanling Mts. and the Wugong Mts (Fig. 4.2a).

A parsimony network revealed the genealogical relationship of the haplotypes showing many closely related haplotypes (Fig. 4.2b) with the most abundant haplotype H1 representing the ancestral type of numerous tip haplotypes. Five haplotypes (H2, H6, H8, H14, H15) are separated by one step from H1, eight haplotypes are separated by two steps and only two haplotypes (H4, H10) are separated by three steps. Of these, H4 was confined to populations CW and QN, while the most divergent haplotype H10 was confined to one population (DT) in the western cluster. In the outgroups, *C. fargesii* and *C. carlesii* shared haplotype H2 with *C. eyrei*, while *C. tibetana* and *C. sclerophylla* harboured new haplotypes differing from H1 or H2 by only one mutation step (Fig. 4.2b).

Over all populations, *C. eyrei* had high haplotype diversity (h_T = 0.814), while haplotype diversity within populations varied from 0 to 1 with an average of h_S = 0.245 (Table 4.1). In the West, the largest number of haplotypes was found in populations HP, DT and TS and in the East in populations WC and HS. At the species level, nucleotide diversity was estimated as $\pi_T = 1.750 \times 10^{-3}$, and at population level it ranged from 0 to 2.441×10^{-3} with an average of $\pi_S = 0.526 \times 10^{-3}$ (Table 4.1). Overall, diversity of haplotypes and nucleotides was found to be slightly higher in the western populations compared to eastern populations (Table 4.3).

Table 4.3 Population diversity and differentiation in *Castanopsis eyrei* at chloroplast DNA.

Parameters	Total	Western	Eastern
H_T	0.842 (0.038)	0.876 (0.053)	0.758 (0.050)
V_T	0.554 (0.085)	0.636 (0.121)	0.487 (0.121)
H_S	0.245 (0.051)	0.254 (0.078)	0.235 (0.068)
V_S	0.150 (0.037)	0.158 (0.056)	0.171 (0.058)
G_{ST}	0.709 (0.061)	0.710 (0.093)	0.690 (0.085)
N_{ST}	0.729 (0.061)	0.751 (0.079)	0.649 (0.098)
N_{ST} - G_{ST}	0.020	0.041	-0.041

H_T and V_T, total gene diversity based on unordered and ordered alleles; H_S and V_S, mean within-population gene diversity based on unordered and ordered alleles; G_{ST} and N_{ST}, genetic differentiation based on unordered and ordered alleles. Values in parentheses are standard error.

Population differentiation at cpDNA was substantial as revealed by high values of both G_{ST} (0.709) and N_{ST} (0.729), which did not differ significantly ($P = 0.594$), indicating that related haplotypes were not clustered. When populations were divided into two geographic groups defined by the STRUCTURE analysis of microsatellites, western populations ($G_{ST} = 0.710$, $H_S = 0.254$, $H_T = 0.876$) exhibited slightly stronger differentiation and higher diversity than eastern populations ($G_{ST} = 0.690$, $H_S = 0.235$, $H_T = 0.758$) (Table 4.3). Neither western nor eastern populations showed significant differences between N_{ST} and G_{ST} ($P > 0.103$). In addition, pairwise F_{ST} values indicated most population pairs differed significantly ($P < 0.05$). We did not find significant patterns of isolation by distance either in all populations or in the two subgroups ($P > 0.64$) indicating a predominant role of genetic drift for shaping cpDNA structure.

DISCUSSION

Contrasting patterns at plastid and nuclear genomes

Before the phylogeography of *C. eyrei* can be discussed, the incongruent patterns of nuclear and plastid genomes have to be considered. A clear geographical clustering was revealed by nuclear genes, showing a split between western and eastern populations, however, such a pattern was not found in cpDNA. Similar inconsistencies between gene trees from cytoplasmic (e.g. chloroplast) and nuclear genes have been repeatedly reported (Rieseberg & Soltis 1991; Tsitrone *et al.* 2003). They can result from a variety of factors, such as evolutionary rate heterogeneity, differential lineage sorting of ancestral polymorphisms (Comes & Abbott 2001) or evolutionary convergence (Davis *et al.* 1998). Moreover, seeds and pollen typically differ in their dispersal distance in Fagaceae leading to much stronger

patterns of differentiation at maternally inherited chloroplast than in biparentally inherited nuclear genes (Petit *et al.* 2005). However, most probably, two non-exclusive factors may be responsible: ancestral polymorphism or chloroplast capture. A number of haplotypes were shared between the congeners *C. eyrei, C. fargesii* and *C. carlesii* and haplotypes from *C. tibetana* and *C. sclerophylla* were only one or two steps apart from the ancestral one. This suggests that the haplotypes and their distribution did not originate recently but predate the most recent glaciation (Zhou *et al.* 2010) or even speciation within *Castanopsis*. In addition, chloroplast capture may have occurred, i.e. introgression of the chloroplasts from one species into another, which appears to be the most common source of phylogenetic incongruence in plants (Rieseberg & Soltis 1991). Shared haplotypes occur most commonly among congeneric species, and have been observed in many long-lived tree species of the Fagaceae (Dumolin-Lapègue *et al.* 1997; Petit *et al.* 2003b). Dumolin-Lapègue *et al.* (1997) found haplotypes shared widely among eight white oaks which was attributed to hybridization and introgression. Petit *et al.*(2003b) explained the fact of haplotype congruence between two widespread and largely sympatric European oak species (*Quercus petraea* and *Q. robur*) by a complex interaction of pollen swamping and different rates of seed dispersal between the species. In our study, two congeneric species (*C. fargesii* and *C. carlesii*) shared one haplotype with *C. eyrei*, although only two individuals per species were tested. They are largely sympatric with *C. eyrei* allowing for interspecific gene exchange. Hence, a more comprehensive sampling of both, congeners and the study species itself, is necessary to quantify and minimize the noise introduced by putative interspecific chloroplast sharing. However, although a phylogeographic signal in our cpDNA data is low (see below), it complements genetic variation and differentiation descriptors obtained with nuclear markers.

At the species level, both markers revealed high levels of genetic diversity in *C. eyrei*. Nuclear diversity (H_E = 0.80) was comparable to other Fagaceae (Shi *et al.* 2011a) and typical for wind pollinated, outcrossing trees (Duminil *et al.* 2007). Variation at cpDNA also represented a considerably high level (H_T = 0.842) among Fagaceae species (Magni *et al.* 2005), higher than *C. hystrix* (H_T = 0.686) (Li *et al.* 2007) and *C. carlesii* (H_T = 0.761) (Cheng *et al.* 2005). However, across the species range, different patterns were shown by the two markers. Microsatellites revealed significantly higher genetic diversity in eastern than western populations (discussed below). In contrast, the chloroplast markers did not reveal such differences, which is likely due to the overall lower level of variation at cpDNA, small effective sample size and noise due to chloroplast capture.

Phylogeographic history

During the Quaternary, glacial cycles are considered to have strongly affected the distribution of plant species and the structure of genetic variation within species. Although large ice sheets have never covered southeast China, most subtropical species are supposed to have retreated to tropical or warmer lowland areas during the LGM. Our climate based distribution model revealed the potential glacial distribution of *C. eyrei* as a narrow belt located in the south along the coastlines similar to the predicted area of subtropical vegetation (Harrison *et al.* 2001), as well as two regions in NW-Guangxi/SW-Guizhou and SE-Fujian-Taiwan expanding somewhat further northwards (Fig. 4.1). Additionally, some regions located more northward, like the Wugong Mts., also showed some degree of climatic suitability. Taiwan Island was connected with the mainland because of ocean regressions and was part of the glacial distribution area similar to other species (Huang *et al.* 2002). Thus, overall, distribution modelling did not give strong support of multiple isolated LGM refugia. However, since large scale distribution modelling can not integrate small scale local climatic variation in varied topographies, it is not the ultimate tool to identify precise locations of cryptic refugia. On the other hand, *C. eyrei* probably did not occupy all areas predicted to have been climatically suitable during the LGM. Thus it would be desirable to evaluate the estimated LGM range against palaeoecological records providing clear evidence of species' occurrences within a given timeframe in the past. However, this type of evidence is not available for *C. eyrei* in China.

Bayesian analysis identified two nuclear gene pools in the West and East of the range of *C. eyrei*. Thus we propose that during the LGM the species had retreated into at least two independent southern refugia, or that the recolonization took place from two genetically differentiated source regions. No such phylogeographic structure was evident from the cpDNA data, as the haplotypes formed a star-like genealogy which is usually interpreted as the outcome of a population expansion (Page & Holmes 1998). However, as discussed above, ancestral polymorphism and chloroplast capture together with genetic drift may have strongly influenced the distribution of haplotypes. The fact that cpDNA haplotypes were strongly differentiated between populations, does not contradict the existence of two glacial refugial areas in the West and East as indicated by microsatellites.

It is a basic phylogeographic hypothesis that because of the role of the rugged topography in putative refugial regions of the southern temperate up to the tropical zone, species may have survived in multiple separate refugia (Hewitt 2000). Thus the mountain ranges in southern China, i.e. the Nanling and Wuyi Mts., are putative refugial areas for the

western and eastern cluster, respectively, of *C. eyrei*. These two regions have been considered a key refugial area during glacial periods (Shen *et al.* 2002), since their complex topography could include a relatively mild Pleistocene climate zone supporting various habitats including subtropical forests. This hypothesis is supported by several studies, e.g. on *Pinus kwangtungensis* (Tian *et al.* 2010), *Fagus longipetiolata* (Liu 2008) and *Ficus pumila* (Chen *et al.* 2012). Furthermore, northward post-glacial recolonizations are characterized by a loss of genetic variation because of founder events or bottleneck effects (Hampe & Petit 2005). Indeed, a cline of decreasing genetic variation towards north was evident, indicating the presence of southern refugia.

Actually, as a common wide-distributed species, it can be hypothesised that *C. eyrei* might have survived in a relatively wider range or more cryptic refugia, compared to more rare species (Wang & Ge 2006). As our climate model revealed, beside the southern area along the coastline, some regions located more nothwards still showed some climatic suitability. A high level of population differentiation for cpDNA with many cases of fixation of different haplotypes may also indicate that multiple glacial refugia have existed during the LGM (Zhou *et al.* 2010), especially in the western cluster. This could also partially explain the lack of geographic structuring of haplotypes. Still, further genetic differentiation can be a useful criterion for inferring the existence of glacial refugia (Petit *et al.* 2003a). The significant substructure in the western cluster (yellow cluster in Fig. 4.3) may imply the presence of such further cryptic refugia or even 'refugia within refugia' (Gómez & Lunt 2007).

Strong population differentiation in the western range

Over the whole distribution range, *C. eyrei* presented a considerably high level of nuclear differentiation (F'_{ST}=0.443) compared to results of a previous smaller scale study located in the eastern part of the range (F'_{ST}=0.15) (Shi *et al.* 2011a). Indeed, western populations exhibited significantly higher differentiation than eastern populations (Table 4.2), indicating stronger genetic drift in the western range. STRUCTURE results supplied strong evidence for extensive gene flow with genetic homogeneity among eastern populations, although substructures were still evident. In addition, population differentiation was less variable and patterns of isolation by distance were more stringent in the east suggesting a higher level of gene flow. On the contrary, the low levels of genetic diversity observed in western populations, particularly in the yellow subcluster, together with the higher level and larger scatter of population differentiation (Fig. 4.5) suggest that severe genetic drift occurred in the

west. Lack of gene flow and a major role of genetic drift in the west were also corroborated by cpDNA haplotypes (Fig. 4.2). Seven haplotypes were found in eastern populations, most of which were shared among populations indicating high levels of gene flow. In contrast, more than half of the haplotypes in the western part were restricted to a single population, showing considerable genetic drift.

A number of facts may have contributed to the observed pattern. First, topographical isolation may be the most important factor. In the western distribution range of *C. eyrei*, large mountain areas prevail and the Nanling Mts., which extend from east to west, may have acted as a geographical barrier obstructing migration between southern and northern populations (Wang *et al.* 2004). As a result of restricted northward gene flow, some haplotypes may be confined to the south, a pattern shown by haplotype H8, which was found in four populations, all south of the Nanling Mts. Second, fragmentation always results in small and isolated populations, where low genetic diversity and high differentiation arise (Young *et al.* 1996). Actually, the distribution range of *C. eyrei* is more fragmented in the west, especially in the southwest, where the calcareous karst habitats are not suitable for *C. eyrei* (Guo *et al.* 2011). Third, the western cluster included a number of populations which represent the current southern distribution margin of *C. eyrei* (e.g. DT, HP, GP). Small size and prolonged isolation in marginal populations presumably have resulted in reduced within-population genetic diversity and pronounced genetic differentiation (Eckert *et al.* 2008).

CONCLUSION

In order to reveal more complete and general phylogeographical patterns of subtropical floras, studies on common and widespread species are still surprisingly limited. In our study, based on nuclear and chloroplast DNA sampled from the common and widespread *C. eyrei* in subtropical China, a western and an eastern phylogeographical lineage were identified, and at least in the western lineage multiple refugial areas are probable. Additionally, a south-north cline of genetic variation indicates postglacial re-colonization from the south. Range wide analysis of closely related species at high spatial resolution will be needed to resolve incongruent patterns of nuclear and chloroplast markers that possibly arouse due to chloroplast capture. However, in contrast to previous analyses of rare species, our results are likely to represent a template in this region for evolutionary history and phylogeography since the LGM for wide spread species.

ACKNOWLEDGEMENTS

We acknowledge Chiang-Her Yen, Min Liu, Ming Jin, Nana Xu, Shipin Chen, Shuo Yu, Wenhui Guo, Xiaoyan Wang, Xinghua Hu, Xiushan Li and Yingliang Liu for helping with sample collections, local forestry stations in Anhui, Fujian, Guangdong, Guangxi, Hubei, Jiangxi and Zhengjiang provinces for assistance in sampling, and Ina Geier and Martina Hermann for their kind help in the lab. This study was financially supported by the German Research Foundation (DFG FOR 891/1 and 891/2) and the China Scholarship Council (2008 3039). Various travel grants for project preparation financed by DFG, NSFC and the Sino-German Centre for Research Promotion in Beijing (GZ 524, 592, 698 and 699) are gratefully acknowledged.

SUPPORTING INFORMATION

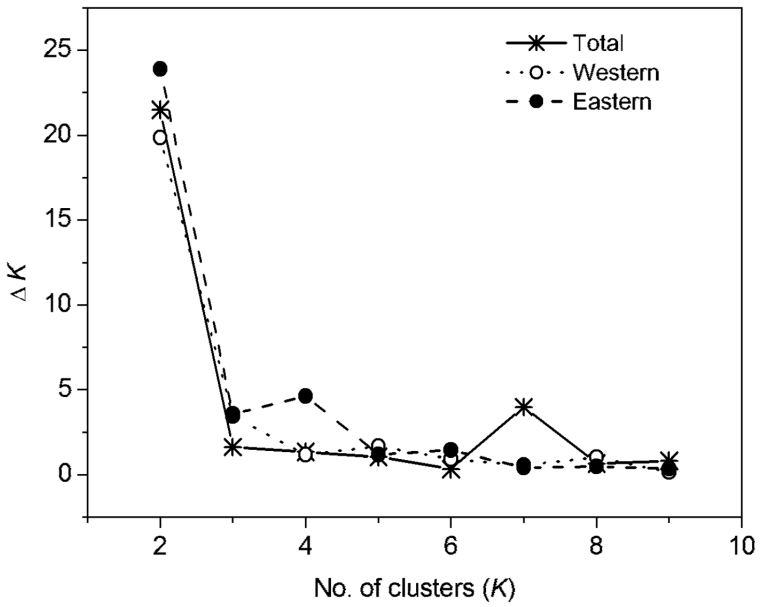

Fig. 4.S1 Values of ΔK based on the rate of the change of ln Pr($X|K$) as a function of the number of clusters K for the whole data set and the two genetic clusters identified.

Table 4.S1 Haplotypes of *Castanopsis eyrei* (H1 - H16) and outgroups (*tib* and *scl*) defined by two chloroplast intergenic spacers and their frequencies (n).

		trnT-trnL											*petG-trnP*														
			1	1	1	2	2	3	3	3	4	4		1	1	1	1	1	1	1	1	1	1	1	2	3	4
		6	0	1	3	3	8	0	2	3	0	7	8	0	1	2	2	2	2	2	2	2	2	2	9	2	2
Hap	n	6	1	7	2	4	1	5	9	0	3	0	5	8	4	0	1	2	3	4	5	6	7	8	7	7	6
H1	91	C	C	-	I2	A	G	-	T	C	G	C	-	A	C	G	A	A	A	T	T	G	C	T	-	C	G
H2	57	.	.	-	.	.	.	-	.	A	.	.	-	.	.	.	.	.	.	.	.	.	.	.	-	.	.
H3	28	.	.	-	.	.	.	-	.	A	.	.	-	.	.	.	.	.	.	.	.	.	.	.	-	T	.
H4	9	.	A	-	.	.	.	-	.	.	.	.	-	G	.	.	.	.	.	.	.	.	.	.	-	.	A
H5	2	.	.	-	.	.	.	-	A	A	.	.	-	.	.	.	.	.	.	.	.	.	.	.	-	.	.
H6	6	.	.	-	.	.	.	-	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.	.	-	.	.
H7	2	.	.	-	.	.	.	-	.	.	.	.	-	G	A	.	.	.	.	.	.	.	.	.	-	.	.
H8	12	.	.	-	.	.	.	-	.	.	.	.	-	G	.	.	.	.	.	.	.	.	.	.	-	.	.
H9	3	.	.	I1	.	G	.	-	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	-	.	.
H10	7	.	.	-	.	.	.	-	.	A	T	-	-	.	.	T	.	.	.	.	.	.	.	.	-	.	.
H11	4	T	.	-	.	.	.	-	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.	.	-	.	.
H12	10	.	.	-	.	.	.	I3	.	.	.	.	-	.	.	.	.	.	-	-	-	-	-	-	-	.	.
H13	9	.	.	-	.	.	.	-	.	.	.	.	-	.	.	.	.	.	.	.	G	.	.	.	I5	.	.
H14	1	.	.	-	.	.	.	-	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	-	T	.
H15	9	.	.	-	.	.	.	-	.	.	.	.	-	.	.	.	.	.	-	-	-	-	-	-	-	.	.
H16	13	.	.	-	-	.	.	-	.	.	.	.	I4	.	.	.	.	.	.	.	.	.	.	.	-	.	.
tib	2	.	.	-	.	.	C	-	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	-	.	.
scl	2	.	.	-	.	.	.	-	.	A	.	.	-	.	.	.	-	-	-	-	-	-	.	.	-	.	.

Insertions: I1, TAAAT; I2, ATATA; I3, ATAA; I4, TTATTAATTTTAATACAT; I5, AGTAACCCTAATA. *tib* is the haplotype from *C. tibetana* and *scl* is from *C. sclerophylla*. The positions of polymorphic sites in the *trn*T-*trn*L and *pet*G-*trn*P spacers and deletions (-) are indicated.

Chapter 5 - Synthesis

Main results

Genetic diversity is the raw material for organisms to evolve and adapt to new conditions. Detecting the level of genetic diversity and underling processes is essential for management and conservation. This thesis provides insights into genetic responses of a common evergreen broad-leaved tree species *Castanopsis eyrei* to different factors at multiple scales. The studies presented in the preceding chapters include the following main results:

Spatial genetic structure (SGS) in adults and seedlings

The first study (Chapter 2) investigated and compared SGS of adults and seedlings from six plots. I tested the correlations of *Sp* statistic, quantifying the intensity of SGS, with several habitat conditions, such as individual density, species richness and elevation. Weak but significant SGS was found within 20 m in both adults and seedlings. *Sp* values were similar in adults (0.0083) and seedlings (0.0109), and showed no correlations with any index of habitat conditions ($P > 0.727$). Heterogeneity tests did not show significant difference in SGS between adults and seedlings in all plots. In addition, similar levels of genetic diversity (Allelic richness and expected heterozygosity) were detected in both adults and seedlings. Thus overall seedlings do not differ from adults in small scale genetic structure suggesting that pollen and seed dispersals are far enough to prevent the build up of increased SGS in the seeding stage.

Isolation by elevation

A second study examined the effect of elevation on distribution of genetic diversity in a regional scale. Diversity and differentiation among 24 populations from different elevation and successional stages were analysed by eight microsatellite loci. One of the loci was found strongly deviated from a neutral model. High level of genetic variation was revealed in *C. eyrei*. A significantly positive correlation was detected between genetic variation and elevation for both putatively selected locus and the neutral loci. Genetic differentiation was similar to that of other Fagaceae species, and followed a model of isolation by distance, but additionally significant isolation by elevation was found both in neutral loci and putatively selected locus, suggesting higher gene flow among similar elevational levels than across different elevational levels and a selective influence of elevation on the distribution of genetic diversity in *C. eyrei*.

Phylogeography

The last study researched the phylogeographic structure and putative refugia of *C. eyrei* on species-range scale. Climate modelling revealed a potential distribution of *C. eyrei* in a narrow belt along the southern coastline during the Last Glacial Maximum. Nuclear microsatellites revealed two clusters corresponding to a split between western and eastern range, and a south-north decline of genetic variation. The eastern cluster harboured significantly higher nuclear genetic diversity. Populations were strongly differentiated at cpDNA with many of them fixed for different haplotypes, but lacked phylogeographic structure. Both data sets indicated higher genetic differentiation in the western than eastern cluster. Overall, the results suggest that multiple refugial areas are probable especially in the western lineage.

Overall discussion

Genetic variation within and among populations is the result of interplay among gene flow, genetic drift, mutation and selection which can be affected by different factors, leading to varying population structure in space and time. Such spatial and temporal genetic structures of populations can be revealed by various molecular markers. In this thesis, genetic diversity and structure of *C. eyrei* were detected at multiple scales.

High level of genetic diversity in Castanopsis eyrei

Regardless of focused scale, *C. eyrei* shows a high level of genetic variation as expected for such a climax tree species with broad distribution range. Life-history traits may determine genetic structure among populations (Duminil *et al.* 2007; Hamrick & Godt 1996; Pannell & Dorken 2006), particularly mating system (Duminil *et al.* 2007). The species with outcrossing mating system are considered capable to harbour high genetic diversity (Van Rossum *et al.* 1997), which has been evidenced by many studies, as well as other congeneric species, such as *C. cuspidata* (Vidhanaarachchi *et al.* 2005), and *C. acuminatissima* (Blakesley *et al.* 2004). In fact, a high outcrossing rate in *C. eyrei* was suggested close to unity (Chapter 2), allowing the maintainance of high genetic diversity.

No difference in SGS between adults and seedlings

In general, tree species display weaker SGS than herbaceous species (Vekemans & Hardy 2004). Especially, in outcrossing tree species, high gene flow within populations is expected

to result in relatively low level of SGS (Leonardi & Menozzi 1996; Streiff *et al.* 1998). Indeed, significant SGS is rarely detectable beyond 30 - 40 m , such as in the common ash *Fraxinus excelsior* (Heuertz *et al.* 2003), oaks *Quercus petraea* and *Q. robur* (Streiff *et al.* 1998), and beech *Fagus sylvatica* and *F. crenata* (Chybicki *et al.* 2009; Oddou-Muratorio *et al.* 2010; Oddou-Muratorio *et al.* 2011), as well as in *C. sclerophylla* (Wang *et al.* 2011). Similarly, in *C. eyrei*, significant SGS was only found within 20 m (Fig. 2.1), which is attributed to limited seed dispersal (mean = 21.04 m, Chapter 2), whereas in larger distance, extensive pollen flow will contribute to the absence of SGS.

SGS is expected to vary across the life stages of species. For a given tree species with seed dispersal primarily by gravity, seeds tend to be aggregated around maternal trees (Asuka *et al.* 2005; Sebbenn *et al.* 2011), thus, seedlings often exsit in much higher density with a high level of relatedness of adjacent seedling individuals. However, during the life time from a seedling to an adult, the process is characterized particularlly by a massive mortality due to competition, predation and selection during eatablishment (Oddou-Muratorio *et al.* 2011). Such a thinning process leads to low density of adults and low level of relatedness. Hence, strong SGS present in seedlings is expected to largely disappear in adults, which has been demonstrated in many studies (Chung *et al.* 2003a; Qi *et al.* 2011), however, was not found in *C. eyrei*. Instead, I found significant SGS and similar *Sp* values in both adults and seedlings. Heterogeneity test did not show consistent patterns in all plots although in combined analysis, significant difference in SGS between adults and seedlings was detected at some distances. All present results indicate no difference between adults and seedlings in SGS, suggesting that pollen dispersal and seed dispersal are far enough to prevent the build up of increased SGS in the seeding stage. Consistent SGS could be resulted from limited seed dispersal, e.g. in *Sinojackia rehderiana* (Styracaceae) (Yao *et al.* 2011). In addition, secondary dispersal by small rodents could blend the seeds from neighbouring trees and promote overlapping distribution of half sibs. Furthermore, high density of adults should be another explanation of weak SGS in seedlings, as overlapping seed shadows will be resulted in, leading to mixing of different half-sib families and a reduction of relatedness of adjacent individuals.

Isolation by elevation at neutral and putatively non-neutral loci

Microsatellites are assumed to represent ideal neutral markers, so that only gene flow and genetic drift rather than selection should affect their genetic structure. However, an increasing number of studies indicated the presence of non-neutral loci (Acheré *et al.* 2005; Lazrek *et al.*

2009; Scascitelli *et al.* 2010). In Chapter 3, one out of eight loci (Ccu97H18) showed non-neutral behaviour (Fig. 3.1), representing a putatively selected locus, although, no information on the genomic position and putatively linked genes of this locus is available.

Elevation showed a significant effect on distribution of genetic variation both at neutral and non-neutral loci. Populations at the top of the mountains harboured the largest amount of genetic variation which was reduced at low elevations. A number of processes may have contributed. First, human disturbance is much stronger at lower elevations. Charcoal has been detected in many local soil profiles (Bruelheide *et al.* 2011) indicating past fire clearance. Populations at higher elevations are more rarely influenced by human activities and, thus, are able to preserve genetic diversity. Second, significantly positive correlation was found between elevation and successional stage, indicating that older, less disturbed forests are often located in higher elevations. Hence, it is likely that upland forests served as sources for colonization after logging at low elevations. Last but most important factor may be selection. The most common allele at locus Ccu97H18 present at low elevations almost went extinct at higher elevations. Instead, many other alleles appeared. These patterns are most likely due to selection. Microstellite markers linked with loci under selection have been shown in other trees (Edelist *et al.* 2006; Ingvarsson 2010; Stefenon *et al.* 2008).

Additionally, significant isolation by elevation was detected at both the potentially adaptive locus and the non-adaptive loci after accounting for the effect of geographic distance, suggesting effective gene flow is not restricted by horizontal distance, but restricted across elevations (Byars *et al.* 2009). Elevation is indicated to influence population genetic structure by the effect on neutral and selective processes. On one side, elevation can result in reproductive isolation due to phenological shifts in neutral markers, e.g. delayed budding (Rusch 1993) or shift of flowering time or prolonged floral longevity and stigma receptivity (Blionis *et al.* 2001) resulting in temporal separation of the timing of flowering (Borchert 1983). Phenological differences in flowering time in turn will lead to partial reproductive isolation which both may facilitate adaptation to elevation and lead to neutral genetic differentiation (Kraj & Sztorc 2009). On the other side, numerous factors associated with elevation may act selectively at many loci, e.g. temperature and irradiance. For example, gene frequency in *Fagus sylvatica* followed temperature-linked spatial and temporal trends (Jump *et al.* 2006). Thus, elevation could affect the selective process via temperature- or other factor-linked genes.

Phylogeographic history

Although subtropical China was never covered by large ice sheets during the Last Glacial Maximum (LGM) (Hewitt 2000), complex climate changes have profoundly affected vegetation distributions and shaped current population structures. Climatic models and derived LGM biome maps have shown that evergreen-broadleaved forests were forced to retreat southwards into the current tropical zone (Harrison *et al.* 2001; Sun *et al.* 1999). Similarly, in this thesis, the climate-based distribution model revealed the potential glacial distribution of *C. eyrei* as a narrow belt located in the south along the coastlines, as well as two regions in NW-Guangxi/SW-Guizhou and SE-Fujian-Taiwan expanding somewhat further northwards (Fig. 4.1). However, climate modelling is not be the ultimate tool to identify precise locations of cryptic refugia, since large scale modelling can not integrate small scale local climate variation, especially in varied topographies. More information can be drawn from molecular analyses.

Based on nuclear microsatellites, two gene pools in the west and east of the range can be identified (Fig. 4.3). Additionally, a cline of decreasing genetic variation towards north suggests the recolonization route from south to north. Thus, I propose that during the LGM *C. eyrei* had retreated into at least two independent southern refugia, or that the recolonization took place from two genetically differentiated source regions. Although cpDNA did not find such a phylogeographic structure (Fig. 4.2) which could be attributed to ancestral polymorphism and chloroplast capture, as the haplotypes formed a star-like genealogy, the fact that cpDNA haplotypes were strongly differentiated between populations, does not contradict the existence of two glacial refugial areas in the west and east as indicated by microsatellites. Considering the role of rugged topography in putative refugia regions, the mountain ranges in southern China, i.e. the Nanling and Wuyi Mts., are considered as putative refugial areas of *C. eyrei* for the western and eastern cluster, respectively. These two regions have been considered a key refugial area during glacial periods (Shen *et al.* 2002), since their complex topography could include a relatively mild Pleistocene climate zone supporting various habitats including subtropical forests. Several studies have confirmed this hypothesis e.g. on *Pinus kwangtungensis* (Tian *et al.* 2010), *Fagus longipetiolata* (Liu 2008) and *Ficus pumila* (Chen *et al.* 2012).

Actually, as a common wide-distributed species, it can be hypothesised that *C. eyrei* might have survived in a relatively wider range or more cryptic refugia, compared to more rare species (Wang & Ge 2006). In my study, some clues can be found. For example, in the western cluster, a high level of population differentiation for cpDNA with many cases of

fixation of different haplotypes may indicate multiple glacial refugia. Additionally, the climate modelling also revealed some regions located more northwards with some climatic suitability. Furthermore, the significant substructure in the western cluster (Fig. 4.3) may imply the presence of cryptic refugia or even 'refugia within refugia' as indicated in other studies (Gómez & Lunt 2007; Heuertz *et al.* 2004b; Wang *et al.* 2009).

Conclusions

This thesis revealed the patterns of genetic diversity and their underling processes in a wide distributed subtropical tree species *C. eyrei*, focusing multiple scales, from small to large scale, i.e species-range scale. Following conclusions can be highlighted.

Spatial genetic structure (SGS) is indicated independent of elevation, density and species richness. Significant SGS was found within 20 m which may result from limited seed dispersal. No difference was found in SGS between adults and seedlings which could be likely attributed to overlapping of seed shadows by secondary dispersal and/or high density of adults. My study contributes to the population dynamics in genetic structure of a common species in continuous forests.

Populations at the top of the mountains harboured the largest amount of genetic variation in *C. eyrei*. Significant isolation by elevation indicates that gene flow is higher among smilar elevational levels than across different elevational levels. My study underlines the importance to check the selective neutrality of marker loci in analyses of population structure and highlights the fact that both neutral and adaptive processes interact in determining population genetic structure.

Based on nuclear and chloroplast DNA data, a western and an eastern phylogeographical lineage were identified in *C. eyrei*, and at least in the western lineage multiple refugial areas are probable. Additionally, a south-north cline of genetic variation indicates postglacial re-colonization from the south. In contrast to previous analyses of rare species, my results are likely to represent a template in this region for evolutionary history and phylogeography since the LGM for wide spread species.

In the time of genomics, with the development of genomic resources in the Fagaceae (Kremer *et al.* 2012), my results not only contribute to the chloroplast genomes (Chapter 4), but also

shed light on the nuclear genomes since putatively selected locus was found (Chapter 3), which will promote examining relevant expressed sequence tag (EST), e.g. by developing genic microsatellite markers (Ueno *et al.* 2009). *Castanopsis* is one of the major genera of Fagaceae including 120 species and covers a wide distribution in Asia. Thus, it can be as a model taxon for genetic evolutionary researches, which my studies on *C. eyre*i contribute to.

References

Acheré V, Favre JM, Besnard G, Jeandroz S (2005) Genomic organization of molecular differentiation in Norway spruce (*Picea abies*). *Molecular Ecology* **14**, 3191-3201.

Alleaume-Benharira M, Pen IR, Ronce O (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *Journal of Evolutionary Biology* **19**, 203-215.

Asuka Y, Tomaru N, Munehara Y*, et al.* (2005) Half-sib family structure of *Fagus crenata* saplings in an old-growth beech–dwarf bamboo forest. *Molecular Ecology* **14**, 2565-2575.

Austerlitz F, Dick CW, Dutech C*, et al.* (2004) Using genetic markers to estimate the pollen dispersal curve. *Molecular Ecology* **13**, 937-954.

Austerlitz F, Garnier-Géré PH (2003) Modelling the impact of colonisation on genetic diversity and differentiation of forest trees: interaction of life cycle, pollen flow and seed long-distance dispersal. *Heredity* **90**, 282-290.

Avise JC (2000) *Phylogeography: the history and formation of species* Harvard University Press, Cambridge.

Axelrod DI, Al-Shehbaz I, Raven PH (1996) Floristic characteristics and diversity of East Asian plants. In: *History of the modern flora of China* (eds. Zhang A, Wu S), pp. 43-55. Springer, New York.

Bai WN, Liao WJ, Zhang DY (2010) Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. *New Phytologist* **188**, 892-901.

Barthlott W, Mutke J, Rafiqpoor D, Kier G, Kreft H (2005) Global centers of vascular plant diversity. *Nova Acta Leopoldina* **92**, 61-83.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceeding of the Royal Society B: Biological Sciences* **263**, 1619-1626.

Bialozyt R, Ziegenhagen B, Petit RJ (2006) Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology* **19**, 12-20.

Blakesley D, Pakkad G, James C, Torre F, Elliott S (2004) Genetic diversity of *Castanopsis acuminatissima* (BI.) A. DC. in northern Thailand and the selection of seed trees for forest restoration. *New Forests* **27**, 89-100.

Blionis GJ, Halley JM, Vokon D (2001) Flowering phenology of *Campanula* on Mt Olympos, Greece. *Ecography* **24**, 696-706.

Böhnke M (2011) *Functional diversity in a subtropical forest based on anatomical and morphological species traits* PhD dissertation, Martin-Luther-University Halle-Wittenberg.

Booth RE, Grime JP (2003) Effects of genetic impoverishment on plant community diversity. *Journal of Ecology* **91**, 721-730.

Borchert R (1983) Phenology and control of flowering in tropical trees. *Biotropica* **15**, 81-89.

Born C, Hardy OJ, Chevallier M-H*, et al.* (2008) Small-scale spatial genetic structure in the Central African rainforest tree species *Aucoumea klaineana*: a stepwise approach to infer the impact of limited gene dispersal, population history and habitat fragmentation. *Molecular Ecology* **17**, 2041-2050.

Bruelheide H, Böhnke M, Both S*, et al.* (2011) Community assembly during secondary forest succession in a Chinese subtropical forest. *Ecological Monographs* **81**, 25-41.

Burczyk J, Adams WT, Birkes DS, Chybicki IJ (2006) Using genetic markers to directly estimate gene flow and reproductive success parameters in plants on the basis of naturally regenerated seedlings. *Genetics* **173**, 363-372.

Burczyk J, Chybicki IJ (2004) Cautions on direct gene flow estimation in plant populations. *Evolution* **58**, 956-963.

Byars SG, Parsons Y, Hoffmann AA (2009) Effect of altitude on the genetic structure of an Alpine grass, *Poa hiemata*. *Annals of Botany* **103**, 885-899.

Caetano S, Prado D, Pennington RT*, et al.* (2008) The history of seasonally dry tropical forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Molecular Ecology* **17**, 3147-3159.

Casa AM, Mitchell SE, Hamblin MT*, et al.* (2005) Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats. *Theoretical and Applied Genetics* **111**, 23-30.

Cavers S, Degen B, Caron H*, et al.* (2005) Optimal sampling strategy for estimation of spatial genetic structure in tree populations. *Heredity* **95**, 281-289.

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24**, 621-631.

Chen X, Wang X, Song Y (1997) Genetic diversity and differentiation of *Cyclobalanopsis glauca* populations in East China. *Acta Botanica Sinica* **39**, 149-155.

Chen XY, Fan XX, Hu XS (2008) Roles of seed and pollen dispersal in natural regeneration of *Castanopsis fargesii* (Fagaceae): Implication for forest management. *Forest Ecology and Management* **256**, 1143-1150.

Chen Y, Compton SG, Liu M, Chen X-Y (2012) Fig trees at the northern limit of their range: the distributions of cryptic pollinators indicate multiple glacial refugia. *Molecular Ecology* **21**, 1687-1701.

Cheng YP, Chien CT, Chen HW, Lln TP (2001) Allozyme variation of *Cyclobalanopsis championii* (Fagaceae), a narrowly distributed species in sourthern Taiwan. *Journal of Heredity* **92**, 65-70.

Cheng YP, Hwang SY, Chiou WL, Lin TP (2006) Allozyme variation of populations of *Castanopsis carlesii* (Fagaceae) revealing the diversity centres and areas of the greatest divergence in Taiwan. *Annals of Botany* **98**, 601-608.

Cheng YP, Hwang SY, Lin TP (2005) Potential refugia in Taiwan revealed by the phylogeographical study of *Castanopsis carlesii* Hayata (Fagaceae). *Molecular Ecology* **14**, 2075-2085.

Chung MY, Epperson BK, Chung MG (2003a) Genetic structure of age classes in *Camellia japonica* (Theaceae). *Evolution* **57**, 62-73.

Chung MY, Nason J, Chung MG*, et al.* (2002) Landscape-level spatial genetic structure in *Quercus acutissima* (Fagaceae). *American Journal of Botany* **89**, 1229-1236.

Chung MY, Nason JD, Chung MG (2007) Effects of population succession on demographic and genetic processes: predictions and tests in the daylily *Hemerocallis thunbergii* (Liliaceae). *Molecular Ecology* **16**, 2816-2829.

Chung MY, Nason JD, Epperson BK, Chung MG (2003b) Temporal aspects of the fine-scale genetic structure in a population of *Cinnamomum insularimontanum* (Lauraceae). *Heredity* **90**, 98-106.

Chybicki IJ, Burczyk J (2010a) NM+: software implementing parentage-based models for estimating gene dispersal and mating patterns in plants. *Molecular Ecology Resources* **10**, 1071-1075.

Chybicki IJ, Burczyk J (2010b) Realized gene flow within mixed stands of *Quercus robur* L. and *Q. petraea* (Matt.) L. revealed at the stage of naturally established seedling. *Molecular Ecology* **19**, 2137-2151.

Chybicki IJ, Trojankiewicz M, Oleksa A, Dzialuk A, Burczyk J (2009) Isolation-by-distance within naturally established populations of European beech (*Fagus sylvatica*). *Botany* **87**, 791-798.

Clark JS (1998) Why trees migrate so fast: confronting theory with dispersal biology and the paleorecord. *American Naturalist* **152**, 204-224.

Clarke LJ, Ayre DJ, Robinson SA (2008) Somatic mutation and the Antarctic ozone hole. *Journal of Ecology* **96**, 378-385.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.

Coart E, Lamote V, De Loose M*, et al.* (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.

Comes HP, Abbott RJ (2001) Molecular phylogeography, reticulation, and lineage sorting in Mediterranean *Senecio* Sect. *Senecio* (Asteraceae). *Evolution* **55**, 1943-1962.

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.

Coyer JA, Hoarau G, Sjøtun K, Olsen JL (2008) Being abundant is not enough: a decrease in effective population size over eight generations in a Norwegian population of the seaweed, *Fucus serratus*. *Biology Letters* **4**, 755-757.

Davis JI, Simmons MP, Stevenson DW, Wendel JF (1998) Data decisiveness, data quality, and incongruence in phylogenetic analysis: an example from the monocotyledons using mitochondrial atp A sequences. *Systematic Biology* **47**, 282-310.

Dick C, Hardy O, Jones F, Petit R (2008) Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropical Plant Biology* **1**, 20-33.

Dick CW (2008) New interpretations of fine-scale spatial genetic structure. *Molecular Ecology* **17**, 1873-1874.

Dick CW, Etchelecu G, Austerlitz F (2003) Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. *Molecular Ecology* **12**, 753-764.

Doligez A, Baril C, Joly HI (1998) Fine-scale spatial genetic structure with nonuniform distribution of individuals. *Genetics* **148**, 905-919.

Duminil J, Fineschi S, Hampe A*, et al.* (2007) Can population genetic structure be predicted from life-history traits? *The American Naturalist* **169**, 662-672.

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

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central - marginal hypothesis and beyond. *Molecular Ecology* **17**, 1170-1188.

Edelist C, Lexer C, Dillmann C, Sicard D, Rieseberg LH (2006) Microsatellite signature of ecological selection for salt tolerance in a wild sunflower hybrid species, *Helianthus paradoxus*. *Molecular Ecology* **15**, 4623-4634.

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

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

Fang J, Wang Z, Tang Z (2010) *Atlas of woody plants in China: distribution and climate* Springer.

Fayard J, Klein EK, Lefèvre F (2009) Long distance dispersal and the fate of a gene from the colonization front. *Journal of Evolutionary Biology* **22**, 2171-2182.

Flora of China (1999) **4**, 314-400.

Friedman J, Barrett SCH (2008) High Outcrossing in the Annual Colonizing Species *Ambrosia artemisiifolia* (Asteraceae). *Annals of Botany* **101**, 1303-1309.

Gämperle E, Schneller JJ (2002) Phenotypic and isozyme variation in *Cystopteris fragilis* (Pteridophyta) along an altitudinal gradient in Switzerland. *Flora* **197**, 203-213.

Gao LM, Möller M, Zhang XM*, et al.* (2007) High variation and strong phylogeographic pattern among cpDNA haplotypes in *Taxus wallichiana* (Taxaceae) in China and North Vietnam. *Molecular Ecology* **16**, 4684-4698.

Garcia-Ramos G, Kirkpatrick M (1997) Genetic models of adaptation and gene flow in peripheral populations. *Evolution* **51**, 21-28.

Gómez A, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography of southern European refugia* (eds. Weiss S, Ferrand N), pp. 155-188. Springer, New York.

Gong W, Chen C, Dobeš C, Fu CX, Koch MA (2008) Phylogeography of a living fossil: Pleistocene glaciations forced *Ginkgo biloba* L. (Ginkgoaceae) into two refuge areas in China with limited subsequent postglacial expansion. *Molecular Phylogenetics and Evolution* **48**, 1094-1105.

Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software* **22**, 1-19.

Goto S, Tsuda Y, Koike Y, Chunlan L, Ide Y (2009) Effects of landscape and demographic history on genetic variation in *Picea glehnii* at the regional scale. *Ecological Research* **24**, 1267-1277.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Guo K, Liu CC, Dong M (2011) Ecological adaptation of plants and control of rocky-desertification on karst region of Southwest China. *Chinese Journal of Plant Ecology* **35**, 991-999.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95-98.

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters* **8**, 461-467.

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

Hamrick JL, Murawski DA, Nason JD (1993) The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. *Vegetatio* **107-108**, 281-297.

Hanson TR, Brunsfeld SJ, Finegan B, Waits LP (2008) Pollen dispersal and genetic structure of the tropical tree *Dipteryx panamensis* in a fragmented Costa Rican landscape. *Molecular Ecology* **17**, 2060-2073.

Hardy OJ, Maggia L, Bandou E*, et al.* (2006) Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. *Molecular Ecology* **15**, 559-571.

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

Harrison SP, Yu G, Takahara H, Prentice IC (2001) Palaeovegetation (Communications arising): Diversity of temperate plants in east Asia. *Nature* **413**, 129-130.

Hazler K, Comps B, Šugar I*, et al.* (1997) Genetic structure of *Fagus sylvatica* L. populations in southeastern Europe. *Silvae Genetica* **46**, 229-236.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* **59**, 1633-1638.

Heller R, Siegismund HR (2009) Relationship between three measures of genetic differentiation G_{ST}, D_{EST} and G'_{ST}: how wrong have we been? *Molecular Ecology* **18**, 2080-2083.

Herrera CM, Bazaga P (2008) Adding a third dimension to the edge of a species' range: altitude and genetic structuring in mountainous landscapes. *Heredity* **100**, 275-285.

Heuertz M, Fineschi S, Anzidei M*, et al.* (2004a) Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. *Molecular Ecology* **13**, 3437-3452.

Heuertz M, Hausman J-F, Hardy OJ*, et al.* (2004b) Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior* L.). *Evolution* **58**, 976-988.

Heuertz M, Vekemans X, Hausman JF, Palada M, Hardy OJ (2003) Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology* **12**, 2483-2495.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907-913.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **359**, 183-195.

Hickerson MJ, Carstens BC, Cavender-Bares J*, et al.* (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution* **54**, 291-301.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965-1978.

Hijmans RJ, Guarino L, Cruz M, Rojas E (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter* **127**, 15-19.

Hiraoka K, Tomaru N (2009a) Genetic divergence in nuclear genomes between populations of *Fagus crenata* along the Japan Sea and Pacific sides of Japan. *Journal of Plant research* **122**, 269-282.

Hiraoka K, Tomaru N (2009b) Population genetic structure of *Fagus japonica* revealed by nuclear microsatellite markers. *International Journal of Plant Sciences* **170**, 748-758.

Hu LJ, Uchiyama K, Saito Y, Ide Y (2010) Contrasting patterns of nuclear microsatellite genetic structure of *Fraxinus mandshurica* var. *japonica* between northern and southern populations in Japan. *Journal of Biogeography* **37**, 1131-1143.

Huang CJ, Zhang YT, Bartholomew B (1999) Fagaceae. In: *Flora of China* (eds. Wu ZY, Raven PH), pp. 314-400. Sciences Press, Beijing.

Huang G, Hong L, Ye W, *et al.* (2009) Isolation and characterization of polymorphic microsatellite loci in *Castanopsis chinensis* Hance (Fagaceae). *Conservation Genetics* **10**, 1069-1071.

Huang SSF, Hwang SY, Lin TP (2002) Spatial pattern of chloroplast DNA variation of *Cyclobalanopsis glauca* in Taiwan and East Asia. *Molecular Ecology* **11**, 2349-2358.

Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M (2008) Ecological consequences of genetic diversity. *Ecology Letters* **11**, 609-623.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53**, 1898-1914.

Ingvarsson PK (2010) Natural selection on synonymous and nonsynonymous mutations shapes patterns of polymorphism in *Populus tremula*. *Molecular Biology and Evolution* **27**, 650-660.

Isagi Y, Suhandono S (1997) PCR primers amplifying microsatellite loci of *Quercus myrsinifolia* Blume and their conservation between oak species. *Molecular Ecology* **6**, 897-899.

Jackson A, Chen R, Loeb L (1998) Induction of microsatellite instability by oxidative DNA damage. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 12468-12473.

Jiménez-Valverde A (2012) Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography* **21**, 498-507.

Jump AS, Hunt JM, Martínez-Izquierdo JA, Peñuelas J (2006) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Molecular Ecology* **15**, 3469-3480.

Jump AS, Rico L, Coll M, Penuelas J (2012) Wide variation in spatial genetic structure between natural populations of the European beech (*Fagus sylvatica*) and its implications for SGS comparability. *Heredity* **108**, 633-639.

Kalisz S, Nason JD, Hanzawa FM, Tonsor SJ (2001) Spatial population genetic structure in *Trillium grandiflorum*: the role of dispersal, mating, history, and selection. *Evolution* **55**, 1560-1568.

Kamiya K, Harada K, Clyde MM, Mohamed AL (2002) Genetic variation of *Trigonobalanus verticillata*, a primitive species of Fagaceae, in Malaysia revealed by chloroplast sequences and AFLP markers. *Genes & Genetic systems* **77**, 177-186.

Körner C (2007) The use of 'altitude' in ecological research. *TRENDS in Ecology and Evolution* **22**, 569-574.

Kraj W, Sztorc A (2009) Genetic structure and variability of phenological forms in the European beech (*Fagus sylvatica* L.). *Annals of Forest Science* **66**, 203-209.

Kremer A, Abbott A, Carlson J*, et al.* (2012) Genomics of Fagaceae. *Tree Genetics & Genomes* **8**, 583-610.

Lange R, Durka W, Holzhauer SIJ, Wolters V, Diekötter TIM (2010) Differential threshold effects of habitat fragmentation on gene flow in two widespread species of bush crickets. *Molecular Ecology* **19**, 4936-4948.

Lazrek F, Roussel V, Ronfort J*, et al.* (2009) The use of neutral and non-neutral SSRs to analyse the genetic structure of a Tunisian collection of *Medicago truncatula* lines and to reveal associations with eco-environmental variables. *Genetica* **135**, 391-402.

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

Lee YJ, Hwang SY, Ho KC, Lin TP (2006) Source populations of *Quercus glauca* in the last glacial age in Taiwan revealed by nuclear microsatellite markers. *Journal of Heredity* **97**, 261-269.

Leonardi S, Menozzi P (1996) Spatial structure of genetic variability in natural stands of *Fagus sylvatica* L. (beech) in Italy. *Heredity* **77**, 359-368.

Li J, Ge XJ, Gao HL, Ye WH (2007) Chloroplast DNA diversity in *Castanopsis hystrix* populations in south China. *Forest Ecology and Management* **243**, 94-101.

Li J, Jin Z (2006) Genetic diversity of *Castanopsis eyrei* populations in three forest communities with different succession stage. *Journal of Zhejiang University (Agricultural & Life Science)* **32**, 232-236.

Li J, Jin Z, Lou W, Li J (2008) Genetic diversity of *Lithocarpus harlandii* populations in three forest communities with different succession stages. *Frontiers of Forestry in China* **3**, 106-111.

Liu HY, Xing QR, Ji ZK, XU LH, Tian YH (2003) An outline of Quaternary development of *Fagus* forest in China: palynological and ecological perspectives. *Flora* **198**, 249-259.

Liu MH (2008) *Phylogeography of Fagus longipetiolata: insights from nuclear DNA microsatellites and chloroplast DNA variation* East China Normal University, Shanghai, China.

Liu MH, Chen XY, Zhang X, Shen DW (2008) A population genetic evaluation of ecological restoration with the case study on *Cyclobalanopsis myrsinaefolia* (Fagaceae). *Plant Ecol* **197**, 31-41.

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany* **82**, 1420-1425.

López-Pujol J, Zhang F-M, Sun H-Q, Ying T-S, Ge S (2011) Centres of plant endemism in China: places for survival or for speciation? *Journal of Biogeography* **38**, 1267-1280.

Lou L, Jin S (2000) Spermatophyta flora of Gutianshan Nature Reserve in Zhejiang. *Journal of Beijing Forestry University* **22**, 33-39.

Maghuly F, Pinsker W, Praznik W, Fluch S (2006) Genetic diversity in managed subpopulations of Norway spruce [*Picea abies* (L.) Karst.]. *Forest Ecology and Management* **222**, 266-271.

Magni CR, Ducousso A, Caron H, Petit RJ, Kremer A (2005) Chloroplast DNA variation of *Quercus rubra* L. in North America and comparison with other Fagaceae. *Molecular Ecology* **14**, 513-524.

Mariette S, Cottrell J, Csaikl U, Goikoechea P, König A (2002) 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.

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

Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* **60**, 2399-2402.

Michalski S, Durka W, Jentsch A*, et al.* (2010) Evidence for genetic differentiation and divergent selection in an autotetraploid forage grass (*Arrhenatherum elatius*). *TAG Theoretical and Applied Genetics* **120**, 1151-1162.

Michalski SG, Durka W (2007) High selfing and high inbreeding depression in peripheral populations of *Juncus atratus*. *Molecular Ecology* **16**, 4715-4727.

Michalski SG, Durka W (2009) Pollination mode and life form strongly affect the relation between mating system and pollen to ovule ratios. *New Phytologist* **183**, 470-479.

Montalvo AM, Conard SG, Conkle MT, Hodgskiss PD (1997) Population structure, genetic diversity, and clone formation in *Quercus chrysolepis* (Fagaceae). *American Journal of Botany* **84**, 1553-1564.

Muir G, Lowe AJ, Fleming CC, Vogl C (2004) High nuclear genetic diversity, high levels of outcrossing and low differentiation among remnant populations of Quercus petraea at the margin of its range in Ireland. *Annals of Botany* **93**, 691-697.

Nakanishi A, Tomaru N, Yoshimaru H, Manabe T, Yamamoto S (2009) Effects of seed- and pollen-mediated gene dispersal on genetic structure among *Quercus salicina* saplings. *Heredity* **102**, 182-189.

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

Nettel A, Dodd RS, Afzal-Rafii Z (2009) Genetic diversity, structure, and demographic change in tanoak, *Lithocarpus Densiflorus* (Fagaceae), the most susceptible species to sudden oak death in California. *American Journal of Botany* **96**, 2224-2233.

Nielsen EE, Hansen MM, Meldrup D (2006) Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. *Molecular Ecology* **15**, 3219-3229.

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**, 1143-1155.

Oddou-Muratorio S, Bontemps A, Klein EK*, et al.* (2010) Comparison of direct and indirect genetic methods for estimating seed and pollen dispersal in *Fagus sylvatica* and *Fagus crenata*. *Forest Ecology and Management* **259**, 2151-2159.

Oddou-Muratorio S, Klein EK (2008) Comparing direct vs. indirect estimates of gene flow within a population of a scattered tree species. *Molecular Ecology* **17**, 2743-2754.

Oddou-Muratorio S, Klein EK, Vendramin GG, Fady B (2011) Spatial vs. temporal effects on demographic and genetic structures: the roles of dispersal, masting and differential mortality on patterns of recruitment in *Fagus sylvatica*. *Molecular Ecology* **20**, 1997-2010.

Ohsawa T, Ide Y (2008) Global patterns of genetic variation in plant species along vertical and horizontal gradients on mountains. *Global Ecology and Biogeography* **17**, 152-163.

Page R, Holmes E (1998) *Molecular evolution, a phylogenetic approach* Blackwell Science Ltd, Oxford.

Pannell JR, Dorken ME (2006) Colonisation as a common denominator in plant metapopulations and range expansions: effects on genetic diversity and sexual systems. *Landscape Ecology* **21**, 837-848.

Peakall R, Smouse PE (2006) GENALEX6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Peng SL, Li QF, Li D, Wang ZF, Wang DP (2003) Genetic diversity of *Pinus massoniana* revealed by RAPD markers. *Silvae Genetica* **52**, 60-63.

Pérez-Figueroa A, García-Pereira MJ, Saura M, Rolán-Alvarez E, Caballero A (2010) Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology* **23**, 2267-2276.

Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. *The Quarterly Review of Biology* **78**, 419-433.

Petit RJ, Aguinagalde I, de Beaulieu JL*, et al.* (2003a) Glatical refugia: hotspots but not melting pots of genetic diversity. *Science* **300**, 1563-1565.

Petit RJ, Bodénès C, Ducousso A, Roussel G, Kremer A (2003b) Hybridization as a mechanism of invasion in oaks. *New Phytologist* **161**, 151-164.

Petit RJ, Duminil J, Fineschi S*, et al.* (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* **14**, 689-701.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology* **12**, 844-855.

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annual Review Ecology, Evolution, and Systematics*.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *The Journal of Heredity* **90**, 502-503.

Poitti A, Leonardi S, Piovani P, Scalfi M, Menozzi P (2009) Spruce colonization at treeline: where do those seeds come from? *Heredity* **103**, 136-145.

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**, 1237-1245.

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

Qi C, Jin Z, Li J (2011) Smallscale spatial patterns of genetic structure in *Castanopsis eyrei* populations based on autocorrelation analysis in the Tiantai Mountain of Zhejiang Province. *Acta Botanica Sinica* **31**, 5130-5137.

Qiu Y-X, Fu C-X, Comes HP (2011) Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. *Molecular Phylogenetics and Evolution* **59**, 225-244.

Qiu Y-X, Guan B-C, Fu C-X, Comes HP (2009) Did glacials and/or interglacials promote allopatric incipient speciation in East Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Dysosma versipellis*. *Molecular Phylogenetics and Evolution* **51**, 281-293.

Quiroga MP, Premoli AC (2007) Genetic patterns in *Podocarpus parlatorei* reveal the long-term persistence of cold-tolerant elements in the southern Yungas. *Journal of Biogeography* **34**, 447-455.

R Development Core Team (2008) *R: A language and environment for statistical computing* R Foundation for Statistical Computing, Vienna, Austria.

Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* **5**, 65-84.

Rousset (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13**, 58-62.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**, 1219-1228.

Rusch VE (1993) Altitudinal variation in the phenology of *Nothofagus pumilio* in Argentina. *Revista Chilena de Historia Natural* **66**, 131-141.

Scascitelli M, Whitney KD, Randell RA*, et al.* (2010) Genome scan of hybridizing sunflowers from Texas (*Helianthus annuus* and *H. debilis*) reveals asymmetric patterns of introgression and small islands of genomic differentiation. *Molecular Ecology* **19**, 521-541.

Sebbenn AM, Carvalho ACM, Freitas MLM*, et al.* (2011) Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree *Copaifera langsdorffii* Desf. *Heredity* **106**, 134-145.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9**, 615-629.

Shaw RG, Etterson JR (2012) Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. *New Phytologist* **195**, 752-765.

Shen L, Chen XY, Li YY (2002) Glacial refugia and postglacial recolonization patterns of organisms. *Acta Ecologica Sinica* **22**, 1983-1990.

Shen L, Chen XY, Zhang X*, et al.* (2005) Genetic variation of *Ginkgo biloba* L. (Ginkgoaceae) based on cpDNA PCR-RFLPs inference of glacial refugia. *Heredity* **94**, 396-401.

Shi MM, Chen XY (2012) Leading-edge populations do not show low genetic diversity or high differentiation in a wind-pollinated tree. *Population Ecology*, in press.

Shi MM, Michalski SG, Chen XY, Durka W (2011a) Isolation by elevation: genetic structure at neutral and putatively non-neutral loci in a dominant tree of subtropical forests, *Castanopsis eyrei*. *Plos One* **6**, e21302

Shi Y-S, Zhang J, Jiang K, Cui M-Y, Li Y-Y (2011b) Development and characterization of polymorphic microsatellite markers in *Castanopsis sclerophylla* (Fagaceae). *American Journal of Botany* **98**, e19-e21.

Smouse PE, Peakall ROD, Gonzales EVA (2008) A heterogeneity test for fine-scale genetic structure. *Molecular Ecology* **17**, 3389-3400.

Stefenon VM, Gailing O, Finkeldey R (2008) Genetic structure of plantations and the conservation of genetic resources of Brazilian pine (*Araucaria angustifolia*). *Forest Ecology and Management* **255**, 2718-2725.

Streiff R, Labbe T, Bacilieri R*, et al.* (1998) Within-population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Molecular Ecology* **7**, 317-328.

Sun XJ, Song CQ, Chen XD (1999) China quaternary pollen database (CPD) and " Biome 6000" project. *Advance in Earth Sciences* **14**, 407-411.

Sutherland BG, Belaj A, Nier S*, et al.* (2010) Molecular biodiversity and population structure in common ash (*Fraxinus excelsior* L.) in Britain: implications for conservation. *Molecular Ecology* **19**, 2196-2211.

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

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

Tanaka T, Yamamoto T, Suzuki M (2005) Genetic diversity of *Castanea crenata* in northern Japan assessed by SSR markers. *Breeding Science* **55**, 271-277.

Tian S, López-Pujol J, Wang HW, Ge S, Zhang ZY (2010) Molecular evidence for glacial expansion and interglacial retreat during Quaternary climatic changes in a montane temperate pine (*Pinus kwangtungensis* Chun ex Tsiang) in southern China. *Plant Systematics and Evolution* **284**, 219-229.

Tomaru N, Mitsutsuji T, Takahashi M*, et al.* (1997) Genetic diversity in *Fagus crenata* (Japanese beech): influence of the distributional shift during the late-Quaternary. *Heredity* **78**, 241-251.

Truong C, PalmÉ AE, Felber F (2007a) Recent invasion of the mountain birch *Betula pubescens* ssp. *tortuosa* above the treeline due to climate change: genetic and ecological study in northern Sweden. *Journal of Evolutionary Biology* **20**, 369-380.

Truong C, Palmé AE, Felber F (2007b) Recent invasion of the mountain birch *Betula pubescens* ssp. *tortuosa* above the treeline due to climate change: genetic and ecological study in northern Sweden. *Journal of Evolutionary Biology* **20**, 369-380.

Tsitrone A, Kirkpatrick M, Levin DA (2003) A model for chloroplast capture. *Evolution* **57**, 1776-1782.

Ueno S, Aoki K, Tsumura Y (2009) Generation of expressed sequence tags and development of microsatellite markers for *Castanopsis sieboldii* var. *sieboldii* (Fagaceae). *Annals of Forest Science* **66**, 509-519.

Ueno S, Tomaru N, Yoshimaru H, Manabe T, Yamamoto S (2002) Size-class differences in genetic structure and individual distribution of *Camellia japonica* L. in a Japanese old-growth evergreen forest. *Heredity* **89**, 120-126.

Ueno S, Yoshimaru H, Kawahara T, Yamamoto S (2000) Isolation of microsatellite markers in *Castanopsis cuspidata* var. *sieboldii* Nakai from an enriched library. *Molecular Ecology* **9**, 1188-1190.

Ueno S, Yoshimaru H, Kawahara T, Yamamoto S (2003) A further six microsatellite markers for *Castanopsis cuspidata* var. *sieboldii* Nakai. *Conservation Genetics* **4**, 813-815.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

Van Rossum F, Vekemans X, Meerts P, Gratia E, Lefebvre C (1997) Allozyme variation in relation to ecotypic differentiation and population size in marginal populations of *Silene nutans*. *Heredity* **78**, 552-560.

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* **13**, 921-935.

Venables WN, Ripley BD (2002) *Modern applied statistics with S*, 4th edn. Springer, New York.

Vidhanaarachchi VRM, Baba S, Murayama S (2005) Genetic diversity of *Castanopsis cuspidata* (Thunb.) Schottky var. *sieboldii* (Makino) Nakai distributed in Ryukyu islands, Japan. *Japanese Journal of Tropical Agriculture* **49**, 53-60.

Wang HW, Ge S (2006) Phylogeography of the endangered *Cathaya argyrophylla* (Pinaceae) inferred from sequence variation of mitochondrial and nuclear DNA. *Molecular Ecology* **15**, 4109-4122.

Wang J, Gao PX, Kang M, Lowe AJ, Huang HW (2009) Refugia within refugia: the case study of a canopy tree (*Eurycorymbus cavaleriei*) in subtropical China. *Journal of Biogeography* **36**, 2156-2164.

Wang L, Guo J, Zhao GF (2006) Genetic diversity of the endangered and endemic species *Psathyrostachys huashanica* natural populations using simple sequence repeats (SSRs) markers. *Biochemical Systematics and Ecology* **34**, 310-318.

Wang R, Compton SG, Chen X-Y (2011) Fragmentation can increase spatial genetic structure without decreasing pollen-mediated gene flow in a wind-pollinated tree. *Molecular Ecology* **20**, 4421-4432.

Wang WP, Hwang CY, Lin TP, Hwang SY (2003) Historical biogeography and phylogenetic relationships of the genus *Chamaecyparis* (Cupressaceae) inferred from chloroplast DNA polymorphism. *Plant Systematics and Evolution* **241**, 13-28.

Wang XP, Wang ZH, Fang JY (2004) Mountain ranges and peaks in China. *Biodiversity Science* **12**, 206-212.

Wang XY, Shen DW, Jiao J*, et al.* (2012) Genotypic diversity enhances invasive ability of *Spartina alterniflora*. *Molecular Ecology* **21**, 2542-2551.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114-138.

Wright S (1978) Variability within and among natural populations In: *Evolution and the genetics of populations*, p. 590. University of Chicago Press, Chicago.

Xu XH, Yu MJ, Hu ZH, Li MH, Zhang FG (2005) The structure and dynamics of *Castanopsis eyrei* populations in Gutian Mountain Natural Reserve in Zhejiang, East China. *Acta Ecologica Sinica (In Chinese)* **25**, 645-653.

Yakimowski SB, Eckert CG (2007) Threatened peripheral populations in context: geographical variation in population frequency and size and sexual reproduction in a clonal woody shrub. *Conservation Biology Volume* **21**, 811-822.

Yan HF, Peng CI, Hu CM, Hao G (2007) Phylogeographic structure of *Primula obconica* (Primulaceae) inferred from chloroplast microsatellites (cpSSRs) markers. *Acta Phytotaxonomica Sinica* **45**, 488-496.

Yao X, Zhang J, Ye Q, Huang H (2011) Fine-scale spatial genetic structure and gene flow in a small, fragmented population of Sinojackia rehderiana (Styracaceae), an endangered tree species endemic to China. *Plant Biology* **13**, 401-410.

Yoshida T, Jones LE, Ellner SP, Fussmann GF, Hairston Jr NG (2003) Rapid evolution drives ecological dynamics in a predator-prey system. *Nature* **424**, 303-306.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *TRENDS in Ecology and Evolution* **11**, 413-418.

Yu M, Hu ZH, Yu J, Ding B, Fang T (2001) Forest vegetation types in Gutianshan Natural Reserve in Zhejiang *Journal of Zhejiang University (Agriculture & Life Science Edition)* **27**, 375-380.

Zeng X, Michalski SG, Fischer M, Durka W (2012) Species diversity and population density affect genetic structure and gene dispersal in a subtropical understory shrub. *Journal of Plant Ecology* **5**, 270-278.

Zhang X, Chen X, Zhang D (2006) Effect of regeneration method on RAPD-based genetic variation of *Cyclobalanopsis glauca* (Fagaceae). *New Forests* **32**, 347-356.

Zheng Z (2000) Late Quaternary vegetational and climatic changes in the tropical and subtropical area of China. *Acta Micropalaeontologica Sinica* **17**, 125-146.

Zhou YF, Abbott RJ, Jiang ZY*, et al.* (2010) Gene flow and species delimitation: a case study of two pine species with overlapping distributions in southeast China. *Evolution* **64**, 2342-2352.

Zhu QH, Pan HX, Zhuge Q*, et al.* (2002) Analysis of genetic structure of natural populations of *Castanopsis fargesii* by RAPDs. *Acta Botanica Sinica* **44**, 1321-1326.

Zhu Y, Siegwolf RTW, Durka W, Körner C (2010) Phylogenetically balanced evidence for structural and carbon isotope responses in plants along elevational gradients. *Oecologia* **162**, 853-863.

Acknowledgements

Many people have contributed to this thesis. It is a great pleasure to express my acknowledgements to them.

First of all, I would like to give my deepest gratitude to my supervisor Dr. Walter Durka. Unforgotten are his patient instructions in analyzing data and his constructive comments and suggestions on the early versions of manuscripts. He is always available when I need any help or suggestion. I am very grateful for what he has supported during the last four years!

I would like to acknowledge Prof. Dr. Helge Bruelheid to be my uni-supervisor. I still clearly remember what he described to me about the university and the city of Halle in the campus of East China Normal University. His kindness and encouragement are what I can not forget!

I am also especially grateful to Prof. Dr. Chen. Since eight years ago, his help and encouragement never leave me. It is he that recommended me to Dr. Walter Durka and Prof. Dr. Helge Bruelheid. It is he that gave his constructive suggestions and comments on the manuscripts. It is he that helped to collect samples, especially in Taiwan on his birthday. Still it is he that recommends me to a new position now. No word can express my gratitude to him!

The premise of my thesis is extensive sampling. Without the people who helped to collect samples, I could not complete it. They are Chiang-Her Yen, Min Liu, Ming Jin, Nana Xu, Shipin Chen, Shuo Yu, Wenhui Guo, Xiaoyan Wang, Xinghua Hu, Xiushan Li, Yingliang Liu. I am very grateful to them, as well as the local forestry stations and Nature Reserves in Anhui, Fujian, Guangdong, Guangxi, Hubei, Jiangxi and Zhengjiang provinces for assistance in sampling.

I am indebted to Dr. Erik Welk for contributing maps and climate modelling.

Many thanks also to the members of BEF-China project who built a great platform to carry out researches as well as great friendship. Especially, I feel happy to be friends with Xueqin Zeng, Yuting Wu, Xuefei Yang, Lin Chen, Sabine Both, Anne Lang, Stefan Trogisch and Martin Böhnke.

The UFZ provided excellent working facilities and I enjoyed my time at the Department of Community Ecology. Thanks to all colleagues, especially to Ina Geier and Martina Hermann for kind help in the lab, and to my working group members Stefan Michalski, Gitte Hornemann, Lena Kloss and Christoph Hahn for their helpful comments, discussions and literature sharing.

I am so lucky to meet my good friends Yun Huang, Jun Li, Qi Zhang and Xueqin Zeng. Because of them, I did not feel so lonely in the foreign country. Because of them, my foreign life became much easier and more colourful. I deeply appreciate their company, help and friendship!

Finally, I am forever grateful to my family! I appreciate them for always being there and giving me warmness. Especially, I am deeply thankful to my husband for his understanding, tolerance, encouragement and endless love.

Appendix

Curriculum vitae

Personal data

Name	Miaomiao Shi
Date and place of birth	30.09.1983, Henan, China
Nationality	Chinese

Education

2008 - 2012	Helmholtz Centre for Environmental Research - UFZ, Germany Martin Luther University Halle-Wittenberg PhD study in community ecology Topic: Patterns of genetic diversity and their underlying processes in a dominant subtropical tree *Castanopsis eyrei* at multiple scales Supervisors: Dr. Walter Durka, Prof. Dr. Helge Bruelheide
2005 - 2008	East China Normal University Master study in molecular ecology Topic: A comparative study of genetic structure between the central and peripheral populations of *Castanopsis sclerophylla* Supervisor: Prof. Dr. Xiaoyong Chen
2001 - 2005	East China Normal University Bachelor study in environmental science Topic: Effect of habitat fragmentation on genetic structure of *Castanopsis sclerophylla* Supervisor: Prof. Dr. Xiaoyong Chen

List of publications

Publications of this dissertation

Shi MM, Michalski SG, Chen XY, Durka W (2011) Isolation by elevation: genetic structure at neutral and putatively non-neutral loci in a dominant tree of subtropical forests, *Castanopsis eyrei*. *PLoS ONE* 6(6): e21302.

Shi MM, Michalski SG, Welk E, Chen XY, Durka W. Phylogeography of a widespread Asian subtropical tree: East-west differentiation and climate envelope modelling suggest multiple glacial refugia. *Journal of Biogeography* (under review).

Shi MM, Michalski SG, Zeng X, Durka W. Seedlings and adults do not differ in small scale genetic structure in *Castanopsis eyrei* (manuscript)

Additional publications

Shi MM, Chen XY (2012) Leading-edge populations do not show low genetic diversity or high differentiation in a wind-pollinated tree. *Population Ecology* (on line). DOI10.1007/ s10144-012-0332-7.

Zhang X*, **Shi MM***, Shen DW, Chen XY (2012) Habitat loss other than fragmentation decreased nuclear and chloroplast genetic variation in a monoecious tree. *PLoS ONE* 7(6): e39146. (* co-first author)

Wang XY, Shen DW, Jiao J, Xu NN, Yu S, Zhou XF, **Shi MM**, Chen XY (2012) Genotypic diversity enhances invasive ability of *Spartina alterniflora*. *Molecular Ecology* 21: 2542-2551.

Liu M, **Shi MM**, Liu MH, Chen XY (2009) Isolation and characteri- zation of microsatellite loci in *Fagus longipetiolata* Seem. (Fagaceae).*Conservation Genetics* 10: 1981-1983

Chen Y, **Shi MM**, Ai B, Gu JM, and Chen XY. (2008) Genetic variation in island and mainland populations of *Ficus pumila* (Moraceae) in eastern Zhejiang of China. *Symbiosis* 45: 37-44.

Contributions to conferences

Shi MM, Michalski SG, Chen XY, Durka W (2011) Phylogeography of a late successional subtropical tree indicatesmultiple refugia.
Verhandlungen der Gesellschaft für Ökologie 41, Oldenburg, Germany. *Oral presentation*

Shi MM, Michalski SG, Chen XY, Durka W (2010) Isolation by elevation in a subtropical dominant tree.
Verhandlungen der Gesellschaft für Ökologie 40, Giessen, Germany. *Oral presentation*

Eigenständigkeitserklärung

Hiermit erkläre ich, dass die Arbeit mit dem Titel "Patterns of genetic diversity and their underlying processes in a dominant subtropical tree *Castanopsis eyrei* at multiple scales" bisher weder der Naturwissenschaftlichen Fakultät I Biowissenschaften der Martin-Luther-Universität Halle-Wittenberg noch einer anderen wissenschaftlichen Einrichtung zum Zweck der Promotion vorgelegt wurde.

Ferner erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst sowie keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen wurden als solche von mir kenntlich gemacht.

Ich erkläre weiterhin, dass ich mich bisher noch nie um einen Doktorgrad beworben habe.

Halle, den 1. September 2012

Miaomiao Shi