Population genetic studies of two *Quercus* species in the Japanese Archipelago

（日本列島に分布するコナラ属 2 種の集団遺伝学的研究）

The United Graduate School of Agricultural Sciences,

Ehime University

Huan-Zhen Liu

劉 煥 臻

2013

Table of Contents

Abstract

In order to investigate the historical expansion of common oaks (*Quercus* sect. *Prinus*) in the northeastern part of Japan, the relation between the chloroplast haplotypes formerly defined by Okaura *et al*. (2007) in *Quercus mongolica* var. *crispula* and the chloroplast type defined by Kanno *et al*. (2004) were examined. Complete linkage between haplotype II and chloroplast C-type was found. The chloroplasts examined in the oak species collected from Sakhalin and Primorski Krai, Russia, and Harbin, China were all T-type. This suggests that the T to C mutation had occurred in haplotype II in Japan. From these findings we could show estimated southern limit of the distribution range of haplotype I about 100 km north of the southern limit of haplotype II in central Honshu. Neither of the haplotype I nor haplotype II was found outside Japan suggesting both occurred in Japan independently from the ancestral haplotype VI.

Quercus phillyraeoides is an evergreen broad-leaved shrub tree that grows in the warm-temperate regions of Japan and China. In order to elucidate the phylogeographical relationships and to find the past colonization routes of evergreen oak (*Quercus phillyraeoides*) we collected samples from 19 populations from the central to western part of Japan and examined chloroplast DNA sequence variation. I analyzed the genetic diversity and differentiation of the populations for four non-coding and one coding region and examined a total of 3,665 bp. Total nucleotide diversity (π) was calculated to be 0.00017 ± 0.00001, and the scaled mutation rate (θ_w) was 0.00021 ± 0.00011. These values were relatively smaller than those formerly reported for Japanese deciduous oak species (*Q. mongolica* var. *crispula*). A total of five haplotypes, two of which included two subtypes (A1, A2, B, C, D1, D2, E), were identified, with types A1 and D2 the most frequent. A large degree of genetic differentiation had occurred across the populations ($G_{ST} = 0.833$ for haplotype data and $F_{ST} = 0.858$ for sequence data). The data suggested several possible refugia populations on the Pacific Ocean coast and the East China Sea coast, nearly identical

to those previously suggested for the lucidophyllus forest refugia. There was no evidence showing that any part of the Seto Inland Sea populations could be possible refugia. Largest genetic differentiation across the populations in the Seto Inland Sea region suggested that these areas had been recently colonized following the last glacial maximum from the refugia in the Pacific coast or in southern Kyushu.

Genetic structure of *Q. phillyraeoides* was determined by using nine microsatellite markers in 423 individuals of 19 populations covering the distribution in Japan. Results showed that alleles number ranged from 1.421 (locus bcqm07) to 4.105 (locus bcqm76) at the population level. Expected heterozygosity ranged from 0.312 (population MR) to 0.457 (population TB) with the average of 0.383. Genetic differentiation among populations was 0.131 showing significant, but moderate genetic differentiation. Genetic variations mainly allocated within population (89.83%) according to AMOVA. STRUCTURE analysis showed that *Q. phillyraeoides* populations could be defined into four clusters. The northernmost population (IW) and the southernmost population (IZ) were most genetically differentiated. These populations are suggested to be the important candidates for conservation as well as refugia populations previously defined. The results of BOTTLENECK showed that *Q. phillyraeoides* in Japan had not experienced a recent bottleneck.

Acknowledgement

I would like to express my gratitude to my supervisor, professor Ko Harada, College of Agriculture, Ehime University, Japan for his guidance, advice, and encouragement during the entire course of this work. My thanks also go to professors Ikuo Ninomiya and Masaharu Kyo, my co-supervisors. I would like to thank my colleagues in the Laboratory of Forest Genetics, Department of Forest Resources Biology, College of Agriculture, Ehime University for their advices, supports, and friendships.

I would like to thank Mr. Shigehiro Yamamoto and Mr. Hideyuki Katai in the Forest and Forestry Research Center of Shizuoka Prefecture for collecting leaf samples of *Quercus phillyraeoides* in Ogasayama, Shizuoka Prefecture. I also thank Mr. Yoshinori Takeichi for helping sample collection of *Q. phillyraeoides* and providing a part of cpDNA sequence data.

I am grateful to many people who assisted me in my Ph. D. Course and in the preparation of my dissertation. I sincerely thank Dr. Koichi Kamiya for his advising on laboratory experiments and supporting on my paper writing. The Japanese government was thanked for their generosity of offering me a Monbukagakushio scholarship.

Finally, I would like to thank my parents, my husband and my son, and other close relatives for their continuous supports and encouragements.

1. **General Introduction**

1.1 Population genetics

The concept of population genetics is that the study of allele frequency distribution and change under the influence of the four main evolutionary processes: natural selection, genetic drift, mutation and gene flow. In a broad sense, population genetics refers to the study of naturally occurring genetic differences among organisms. Genetic differences that are common among individuals of the same species are called genetic polymorphisms, whereas genetic differences that accumulate between species constitute genetic divergence. Therefore population genetics also can be defined as the study of polymorphism and divergence (Hartl, 2000). Population genetics was a vital ingredient in the emergence of the modern evolutionary synthesis. Its primary founders were Sewall Wright, J. B. S. Haldane and R. A. Fisher, who also laid the foundations for the related discipline of quantitative genetics. Population genetics seeks to understand the causes of genetic differences within and among species, and molecular biology provides a rich repertoire of techniques for identifying these differences since the last 40 years.

1.2 Molecular phylogenetics

Molecular phylogenetics is the analysis of hereditary molecular differences, mainly in DNA sequences, to gain information on an organism’s evolutionary relationships. The

result of a molecular phylogenetic analysis is expressed in a phylogenetic tree. Molecular phylogenetics is one aspect of molecular systematics, in a broader sence, it also includes the use of molecular data in taxonomy and biogeography.

Molecular evolution includes two area of study, the evolution of macromolecules and the reconstruction of evolutionary history of genes and organisms (Li, 1997). Molecular phylognetics belongs to the second area of molecular evolution. It can deal with the evolutionary relationships among organisms or genes based on molecular biology and statistical techniques. When the relationships among organisms are concerned, it is particularly called molecular systematics, as narrow sense of molecular phylogenetics. In modern time, protein and nucleotide sequences have been extensively used to reconstruct the phylogenetic relationship. The using of molecular sequences has more powerful than morphological and physiological data, because they generally evolve much more regular manner and are much more easy to treat quantitatively (Li, 1997).

1.3 Two *Quercus* species studied

The Fagaceae was formally divided into three subfamilies, *Fagoideae* (*Fagus* and *Nothofagus*), *Castanoideae* (*Castanea*, *Castanopsis*, *Chryisolepis*, and *Lithocarpus*), and *Quercoideae* (*Colombobalanus*, *Formanodendron*, *Quercus* and *Trigonobalanus*) (Forman, 1964). *Quercus* is without doubt one of the most important woody genus of the northern hemisphere, and includes deciduous and evergreen species extending from cold latitudes to tropical Asia and the America.

Quercus mongolica var. *crispula* is a deciduous, diploid, long-living, broadleaved, and mostly outcrossing wind-pollinated oak species that is commonly distributed throughout the cool-temperate deciduous forests. This species distribute throughout Japanese archipelago, and in the northeastern part of Japan they distribute continuously from the sea level to higher mountains up to 1,000m, however, in the southwestern part of Japan they show patchy distribution from the altitude of 900m to 1,500m. *Quercus mongolica*, the believed ancestor species, broadly distribute in northwestern part of Eurasian contintent. *Quercus mongolica* var. *crispula* is one of the members of the section Prinus, which includes *Q. dentata*, *Q. serata* and *Q. aliena* in Japan.

Quercus phillyraeoides is a kind of evergreen tree with 4.7- 6 m high and with wind-pollinated flower. Geographically they distribute in subtropical climate zone in China, and moderate climate zone in Japan. These oaks commonly grow on sunny, dry, windy hills near the coast. Its leaves are small and thick with shallow sinuses on the edge. They grow along the ocean coast from Boso Peninsula at their northern limit to the southwest. The southern-most population in the Japanese Archipelago is seen in Okinawa-jima. They are especially abundant around the Seto Inland Sea.

1.4 The objectives of this study

This study focuses on molecular phylogeny and population genetics of two *Quercus* species. The thesis is divided into three parts.

In the first part, geographic distribution and chloroplast T/C-type in *Quercus*

mongolica var. crispula in the northeastern Japan were evaluated by using chloroplast DNA variation.

In the second part, genetic diversities and genetic differentiation were evaluated and phylogeography of *Quercus phillyraeoides* (*Fagaceae*) in Japan were studied using chloroplast DNA variation.

In the third part, genetic variations and genetic differentiation of *Q. phillyraeoides* (*Fagaceae*) in Japan were studied using microsatellite markers.

2. Geographic distribution and origin of the chloroplast T/C-type in *Quercus mongolica* var. *crispula* in northeastern Japan

2.1 Introduction

Quercus mongolica var. *crispula* and related species in section *Prinus* (*Q. serrata*, *Q. dentata*, and *Q. aliena*) are common oak species that are widely distributed in the deciduous forests of Japan. Palynological studies have examined historical changes in the distribution range of temperate deciduous trees including these oak species during the last glacial age (reviews in Yasuda and Mihoshi 1998). More recently, molecular markers such as chloroplast DNA variants and microsatellite markers have been used to study migration routes and the colonization process for these oak species (Kanno *et al*. 2004; Okaura *et al*. 2007; Ohsawa *et al*. 2011). Kanno *et al*. (2004) examined five cpDNA non-coding regions (*trn*D-*trn*T, *trn*T-*trn*L, *rps*14-*psa*B, *trn*S-*trn*T, and *trn*Q-*trn*S intergenic spacers) in a total of 3396 bp in the four oak species and found a single nucleotide (T/C) substitution in the *trn*Q-*trn*S region that could distinguish two chloroplast types, the T-type and the C-type. Here we refer to the T/C-type (Kanno *et al*. 2004) as the "chloroplast type" in order to differentiate it from the "haplotype" defined in *Q. mongolica* var. *crispula* by Okaura *et al*. (2007). They showed by phylogenetic analysis of the four above-mentioned oak species and related species that the C-type is a derived and is possibly of monophyletic origin, while the T-type is ancestral (Kanno *et al*. 2004). The T-type spans from South Korea to whole Japan, whereas the C-type is

restricted to northeastern Japan. They suggested that the T to C mutation occurred in an individual of one of the four oak species and then was transferred among these species by hybridization in northeastern Japan (Kanno *et al*. 2004). Further, they suggested that this likely occurred in the Kanto region, which is thought to contain glacial age refugia of deciduous oaks (Tsukada 1982; Uchiyama 1998), then spread northwards during the subsequent warm period (Kanno *et al*. 2004). Later, Okaura *et al*. (2007) examined six cpDNA regions (*trn*T-*trn*L, *trn*L-*trn*F, *atp*B-*rbc*L, the *trn*H-*psb*A intergenic spacers, the *trn*L intron, and the *mat*K gene) in a total of 4253 bp in the same four oak species in Japan and found nine haplotypes (haplotypes I to IX). The distribution of these haplotypes clearly showed that they were divided into two non-overlapping groups in northeastern Japan (haplotypes I and II) and southwestern Japan (haplotypes III to IX). The boundary between these non-overlapping groups is located in central Honshu, coincident with the Itoigawa-Shizuoka tectonic line (Okaura *et al*. 2007). Because the same haplotypes VI, VII, VIII in southwestern Japan were also found in Korea and Liaoning Province, China, they hypothesized that the southwestern oak populations had migrated from the Eurasian Continent across the southern land bridge via the Korean peninsula during the last few glacial ages (Okaura *et al*. 2007). On the other hand, northeastern oak populations had probably migrated from the Eurasian Continent via the northern land bridge, retracted once to refugia in central Japan, and subsequently expanded northward after the last glacial maximum (LGM) (Okaura *et al.* 2007). According to these results, we suspected that the T-type described by Kanno *et al*. includes several haplotypes, as defined by Okaura *et al*. (2007). In order to test this

hypothesis, we sequenced the *trn*Q-*trn*S region in individuals of *Q. mongolica* var. *crispula* from northeastern Japanese populations whose haplotypes are already known in Okaura et al. (2007). Because of the intensive sampling of the oak species by Kanno *et al*. (2004), especially in northeastern Japan, by relating the T/C-type to the haplotypes of Okaura *et al*. (2007), we expected that we could more precisely describe the distribution of the haplotypes in northeastern Japan. This will enable us to estimate the location of refugia in the LGM and will shed light on the hypothesized northward migration routes of deciduous oaks. We also examined the *trn*Q-*trn*S region and defined the haplotypes in samples of *Quercus* species collected from surrounding areas of Sakhalin and Primorski Krai, Russia, and Harbin, China, and discuss the possible origin of the T/C-type in northeastern Japan.

2.2 Materials and methods

We selected 12 populations of *Q. mongolica* var. *crispula*, including 79 individuals from northeastern Japan, whose haplotypes had already been determined and reported by Okaura *et al*. (2007), and examined the DNA sequence of the *trn*Q-*trn*S intergenic spacer region using the same primer set (QSf and QSr) as described in Kanno *et al*. (2004). We also included 10 newly collected samples from Wakkanai, Hokkaido and determined the sequence of all six cpDNA regions (*trn*T-*trn*L, *trn*L-*trn*F, *atp*B-*rbc*L, the *trn*H-*psb*A intergenic spacers, the *trn*L intron and the *mat*K gene), in addition to the *trn*Q-*trn*S spacer, to determine their haplotypes.

Ten samples of *Quercus mongolica* var. *mongolica* collected form Harbin, Notheast China in 2004, 13 samples of the same species from Primosrski Kari (also known as the Maritime Territory), Russia collected in 2002, and 66 samples of *Q. mongolica* var. *crispula* collected from Sakhalin, Russia in 2005 were also examined for the six cpDNA regions to determine their haplotypes and the *trn*Q-*trn*S region to determine the chloroplast type. The sequences of the *trn*Q-*trn*S spacer were adjusted to 974 bp and aligned by ClustalW on DDBJ (DNA Data Bank of Japan). Sampling sites for each locations are listed in Table 2.1 and also shown in Fig. 2.1 (a). Detailed PCR and sequencing methodologies for determining the haplotypes are described by Okaura *et al*. (2007). In order to reconstruct phylogenetic relationship among haplotyopes, a haplotype network was constructed by using TCS 1.21 (Clement *et al*., 2000). All newly sequenced data were deposited in DDBJ and assigned accession numbers AB725917 through AB725986 for the *trn*Q-*trn*S spacer, AB727873-AB727882 for the *mat*K gene, AB72883-AB72892 for the *trn*T-*trn*L spacer, AB727893-AB727902 for the *trn*L-*trn*F spacer, AB727903-AB727912 for the *atp*B-*rbc*L spacer, and AB727912-AB727922 for the *trn*H-*psb*A spacer.

2.3 Results and discussion

Among the five Hokkaido populations, one population (Eiura) was monomorphic, containing only haplotype I, and all of its individual members were shown to be T-type. Two populations (Shikaoi and Hidaka) were polymorphic, including haplotypes I and II,

and within these, all samples of haplotype I were T-type, and all of the samples of haplotype II were C-type. Two other populations (Wakkanai and Ohnuma) were monomorphic for haplotype II and all members of these populations were shown to be C-type. Among the eight populations in Honshu, only one population (Mt. Hayachine) exhibited haplotype I and all of its members were T-type. The remaining seven populations (Shirakami, Tazawako, Izumigadake, Iizaka, Myoko, Takaharayama, and Chichibu) were monomorphic for haplotype II, and all samples were C-type. These results indicate that the haplotypes I and haplotype II are completely linked to T-type and C-type, respectively. Neither individuals of haplotype I with C-type nor individuals of haplotype II with T-type were found. We also examined the six cpDNA regions to determine their haplotypes and *trn*Q-*trn*S region to determine their chloroplast types in the populations from Sakhalin and Primorski Krai in Russia and Harbin in China. The Russian populations had either haplotype VI or XI, but all were T-type. All individuals of *Q. mongolica* var. *mongolica* from Harbin were haplotype XII with a 30 bp newly found insertion at position 770 of *atp*B-*rbc*L spacer (we defined this haplotype XII'), and all were T-type. Because the T-type was shown to be the ancestral type and the C-type was a derived (Kanno *et al.* 2004), these findings strongly suggest that the T to C mutation occurred in an individual of haplotype II in Japan. These results were summarized in Table 2.1. The haplotypes detected in this work were redefined including the T/C mutation in the *trn*Q-*trn*S spacer and presented in Table 2.2.

Formally, the haplotypes in the populations in Sakhalin was reported to be the haplotype I (Okaura *et al*., 2007), but reexamination of the samples collected in

Sakhalin indicated that they were either the haplotype VI or haplotype XI. Similarly, the haplotype I (and also the haplotype II) was not found in other populations outside Japan. This means that we have no evidence for northeastern haplotypes (i. e. the haplotype I and the haplotype II) to have once migrated into Japan via northern land bridge as Okaura *et al*. (2007) suggested. It is more probable that both occurred in Japan derived from the ancestral haplotype VI. It is noticeable that the Russian populations in Sakhalin and in Primorski Krai shared common haplotypes VI and XI, and both has been detected in Liaoning Province, China by Okaura *et al*. (2007). Preceding haplotype found in Harbin (i. e., haplotype XII) also has been detected in Korea by Okaura *et al*. (2007). These findings suggest diversification of haplotype had occurred in the Eurasian Continent. We reconstructed haplotype network for the redefined haplotypes in Fig. 2.2. Haplotype I could have occurred from the haplotype VI by one mutation (T/C at position 745 in the *atp*B-*rbc*L spacer). Haplotype II could have occurred independently from the haplotype VI by two mutations (T/G at position 1321 in *mat*K gene and T/C at position 276 in *trn*Q-*trn*S spacer), although it is not possible to tell which of the mutations occurred first.

The complete linkage discovered here between haplotypes II and the chloroplast C-type provides a clue to allow us to draw more precise picture for the distribution of the two haplotypes in northeastern Japan. The boundary of haplotype II can be approximated by imposing the distribution of the C-type shown by Kanno *et al*. (2004) on the populations in central Japan, as indicated by the solid lines on Fig. 2.1 (b). The southern limit of haplotype I could be estimated by connecting the populations carrying

the T-type in the northeastern part of Japan, as indicated by double lines in Fig. 2.1 (b), more than 100 km north of the southern limit of haplotype II, and at least 250 km south of Mt. Hayachine, the only population with haplotype I detected by Okaura *et al*. (2007). It is notable that the estimated southern limit of haplotype I extends to the south along the Pacific coast to the Boso Peninsula in Chiba Prefecture (Fig. 2.1 (b)). Although palynological studies have provided no evidence of deciduous oak refugia in the Tohoku district in Japan during the LGM, a study of the Kanto region showed that a mixed forest of conifers and deciduous broad-leaved trees including oaks did exist during the LGM (Uchiyama 1998). Probably, two groups of refugia, one containing haplotype II along the Itoigawa-Shizuoka tectonic line and one containing haplotype I in the Kanto region, existed at one time and then expanded northward after the LGM.

Recently, Ohsawa *et al*. (2011) studied the genetic structure of 16 *Quercus mongolica* var. *crispula* populations along a latitudinal gradient in northeastern Japan using seven nuclear microsatellite markers. They detected two geographically differentiated clusters: one dominated most of the populations in Hokkaido (cluster I) and the other dominated all Honshu populations and the southernmost population in Hokkaido in Hakodate (cluster II). They suggested that cluster I had remained in Hokkaido during the LGM, while another had expanded northward from the refugia that existed in the south (probably in central Japan), based on the finding of significant negative correlation between allelic richness (although not heterozygosity) and latitude (Ohsawa *et al*. 2011). The distribution of cluster I is nearly identical to that of haplotype I in Hokkaido. The cluster II may be a mixture of lineage with haplotype II and lineages with haplotype I

remaining in Boso peninsula and probably in Mt. Hayachine. This possibility could be tested by investigating the relationship between microsatellite genotypes and chloroplast haplotypes in Hokkaido and Honshu.

2.4 Conclusions

In order to investigate the historical expansion of common oaks (*Quercus* sect. *Prinus*) in the northeastern part of Japan, the relation between the chloroplast haplotypes formerly defined by Okaura *et al*. (2007) in *Quercus mongolica* var. *crispula* and the chloroplast type defined by Kanno *et al*. (2004) were examined. Complete linkage between haplotype II and chloroplast C-type was found. The chloroplasts examined in the oak species collected from Sakhalin and Primorski Krai, Russia, and Harbin, China were all T-type. This suggests that the T to C mutation had occurred in haplotype II in Japan. From these findings we could show estimated southern limit of the distribution range of haplotype I about 100 km north of the southern limit of haplotype II in central Honshu. Neither of the haplotype I nor haplotype II was found outside Japan suggesting both occurred in Japan independently from the ancestral haplotype VI.

3. Phylogeography of *Quercus phillyraeoides* (Fagaceae) in Japan as revealed by chloroplast DNA variation

3.1 Introduction

The lowlands of the southwestern part of Japan are covered with forests dominated by *Castanopsis*, *Quercus* subgenus *Cyclobalanopsis*, *Cinnamomum*, *Machilus* and *Camellia* with several other evergreen shrubs and herbs. These forests are socio-botanically defined as *Camellietea japonicae* Miyawaki et Chiba 1963 and often called lucidophyllus forests or forests with shiny leaves (Kira 1977). The geographic distribution of lucidophyllus forests corresponds to zones that follow the contours of the warmth index, which suggests that their distribution is primarily determined by temperature (Tsukada 1974; Hattori 1985). Although some tree species such as *Castanopsis* and evergreen oaks (*Quercus* subgenus *Cyclobalanopsis*) commonly appear in the lucidophyllus forests, because of contrasting winter precipitation and snowfall, species composition largely differs from the Japan Sea side to the Pacific Ocean side. In addition to climatic conditions, the geographical distribution and species composition of lucidophyllus forests is thought to be related to postglacial migration routes from different refugia at the last glacial maximum (LGM) (Hattori 2002).

Quercus phillyraeoides is an evergreen oak and typical member of the lucidophyllus forests on the Pacific and on the East China Sea side, and does not appear on the Japan Sea side. These oaks commonly grow on sunny, dry, windy hills near the coast and can

grow to a height of 15 m, but typically reach 3-5 m. Their leaves are small and thick with shallow sinuses on the edge. They grow along the ocean coast from Boso Peninsula at their northern limit to the southwest. The southern-most population in the Japanese Archipelago is seen in Okinawa-jima, the main island of the Ryukyu Islands. They are especially abundant around the Seto Inland Sea, which is known to have one of the lowest precipitation rates in Japan throughout the year. While the other evergreen oak species belong to subgenus *Cyclobalanopsis*, this evergreen species is unique in belonging to the subgenus *Quercus*, which is mainly deciduous. Taxonomically, this species belongs to sect. *Ilex*, but is closely related to sect. *Cerris*, which includes deciduous oaks *Q. variabilis* and *Q. actissima*, as shown by molecular phylogenetic analysis (Harada *et al.* 2003).

Lucidophyllus forests develop in subtropical to warm temperate monsoon zones and range from the southern foothills of the Himalayas to Yunnan, South China, East China, Taiwan, Ryukyu, and the western part of Japan. They are most often found at altitudes below 1,000 m and many have been converted to secondary forests because they have often occurred in areas of human activity since the pre-agricultural age (Nakao 1966). Like other tree species in lucidophyllous forests, *Q. phillyraeoides* has been utilized for producing daily supplies. *Quercus phillyraeoides* is characterized by its slow growth and hard wood known to be the heaviest among Japanese trees. It is often used for hedges in western Japan on the Pacific coast side because of its dense leaves, but is most famously used for making high-quality “white charcoal” known as Bincho-tan. The largest quantities of Bincho-tan are produced in Wakayama Prefecture in the Kinki

district, but it is also produced in Ehime and Kochi Prefectures in Shikoku, and in Miyazaki and Oita Prefectures in Kyushu. Although the total production of charcoal in Japan dramatically decreased in the 1960s upon the shift to oil as a primary energy source, the production of white charcoal has remained at a constant level of about 3,000 t per year, mainly for use in traditional Japanese restaurants (Consortium of Japan Bincho-tan Meeting 2011). The oak populations used for making charcoal have maintained themselves by sprouting, but many of these forests are isolated and occur in small patches often threatened with extinction. The easternmost populations exist in Chiba and Kanagawa prefectures and are assigned to critically endangered (Chiba Biodiversity Center 2009) and endangered (Kanagawa Prefecture 2006) species. The species is also assigned as near threatened in Nagasaki Prefecture (2011) and Miyazaki Prefecture (2010), and vulnerable in Okinawa Prefecture (2006).

To protect and preserve the oak tree populations as genetic resources for production forests and to conserve coastal ecosystems, it is necessary to elucidate their genetic structure and to estimate genetic variation within and among these populations. The genetic information present in plant chloroplast (cpDNA) is of great interest in phylogeny and population genetics, largely because of the non-Mendelian mode of inheritance of its genome. Due to predominantly maternal chloroplast inheritance in angiosperms and the absence of recombination, the maternal lineage of a population can be traced at the level of the whole chloroplast genome (Corriveau and Coleman 1988). It is useful for determining the route of seed migration and for identifying lineages that are expected to have been derived from refugia during the last glacial period (McCauley

1995; Newton *et al*. 1999; Abbott *et al*. 2000). Sequence variation in cpDNA has been widely used to investigate intraspecific phylogeographical relationships both in deciduous trees (Dumolin-Lapegue *et al.* 1997; Fujii *et al.* 1997, 2002; Johnk and Siegismund 1997; Ferris *et al*. 1998; King and Ferris 1998; Okaura and Harada 2002; Okaura *et al*. 2007) and evergreen trees (Huang *et al.* 2002, 2004; Aoki *et al.* 2004, 2006; Wu *et al.* 2006).

The temperature at the LGM, which occurred about 20,000 years ago, was estimated to be 7 ℃ below the present temperature and the oceanic water line was 80-140 m lower than its present level (Minato and Ijiri 1976). The main islands of the Japanese archipelago were once connected in a contiguous land mass. Widespread lucidophyllus forests during the inter-glacial period retreated southward and to the lowlands, and survived in a restricted number of small populations; in other words, in refugia. Present vegetation is thought to be the result of expansion from a restricted number of refugia. Palynological studies suggested these lucidophyllus forest refugia to have been located at the south end of Kyushu (Tsukada 1974), but several other studies including climatological and biogeographical studies have indicated other possible refugia along the Pacific coast, such as on the Muroto and Ashizuri Peninsulas of Shikoku, on the Kii Peninsula in the Kinki district, and on the Izu, Miura, and Boso Peninsulas in the Kanto district (Maeda 1980; Hattori 2002). The islands of Goto and Amakusa in Kyushu on the East China Sea side have also been proposed as refugia (Maeda 1980; Hattori 2002). On the other hand, palynological studies have rarely been undertaken around the Seto Inland Sea, where *Q. phillyraeoides* grows abundantly (Miyoshi 1998). To date, only a

few studies have focused on this species, and even fewer have done so at the molecular genetic level. This chapter is intended to reveal the pattern of genetic variation and intraspecific phylogeographical relationships among *Q. phillyraeoides* populations in Japan and estimate refugial populations in the LGM. This could be utilized to draw outlines for assessing the groups of populations to be conserved for protecting costal ecosystems, and could also be useful for future breeding studies such as to use the detected haplotypes for provenance tests and selective markers.

3.2 Materials and methods

3.2.1 Plant materials

Samples were collected from 19 populations covering the distribution range of this species in Japan: 1 population in the Kanto district, 2 populations in the Chubu district, 2 populations in the Kinki district, 7 populations in Shikoku, 2 populations in the Chugoku district, 4 populations in Kyushu, and 1 population in Okinawa. Total populations were divided into three areal groups considering geological accommodation: one on the Pacific coast (Group 1), one on the Seto Inland Sea coast (Group 2), and the other in the western and southern Kyushu and Okinawa (Group 3), which are located on the edge of the East China Sea. 7 populations were included in Group 1, 8 were included in Group 2, and 4 were included in Group 3 (Table 3.1; Fig. 3.1). Fresh leaves were collected from a total of 440 trees from 19 populations (about 20 samples per

population). Five individuals were randomly chosen from each population for a total of 95 samples analyzed in this study.

3.2.2 DNA extraction and PCR amplification

DNA was extracted from fresh leaves using a modified CTAB method (Doyle and Doyle 1990). The quality and concentration of DNA were confirmed by electrophoresis on 0.8 % agarose gels along with a *λ-Hin*dIII digest as DNA size standard. The extracted DNA was stored at 4 ℃ until required.

Five cpDNA regions were amplified by PCR using universal primers for two non-coding regions of the *trn*L intron and the *trn*L-*trn*F spacer reported by Taberlet *et al.* (1991), one non-coding region of the *atp*B-*rbc*L spacer reported by Terachi (1993), one non-coding region of the *trn*H-*psb*A spacer reported by Okaura *et al.* (2007), and one coding region of *mat*K reported by Kamiya *et al*. (2002). Each reaction mixture contained 2×GoTaq Hot Start Colorless Master Mix (Promega) with 400 μM dNTPs and 4 mM $MgCl_2$, 10 ng of DNA, and 5 pmol of primers in a total volume of 25 μL. PCR conditions were initial denaturation at 94 ℃ for 2 min, followed by 35 cycles at 94 ℃ for 1 min, 50-55 ℃ for 1 min (annealing temperature depending on the regions to be amplified), and 72 ℃ for 2 min, with a final extension at 72 ℃ for 7 min using the GeneAMP PCR System 9700 Thermal Cycler (Applied Biosystems) and Veriti 96-Well Thermal Cycler (Applied Biosystems), and subsequently held at 4 ℃. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) and electrophoresed

on a 0.8 % agarose gel to determine the concentration of the amplified DNA.

3.2.3 Sequencing of the cpDNA regions

Sequencing reactions were performed directly on the PCR products using the BigDye Terminator (version 1.1) Cycle sequencing Kit (Applied Biosystems). Template DNA (50 ng) was subjected to sequencing in a reaction cycle consisting of preheating for 2 min at 96 ℃ and 25 cycles of denaturation at 90 ℃ for 30 s, annealing at 50 ℃ for 30 s, and extension at 60 ℃ for 4 min. The five cpDNA regions (*trn*L intron, *trn*L-*trn*F spacer, *atp*B-*rbc*L spacer, *mat*K gene, and *trn*H-*psb*A spacer) were sequenced using the abovementioned universal primers. Labeled products were purified by ethanol precipitation and subsequently applied to ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Sequence data were deposited to DDBJ (accession nos. AB650016-AB650465).

3.2.4 Data analysis

ATGC version 6.0 software (GENETYX) was used to assemble sequence data. Sequence data were aligned using ClustalW on DDBJ (DNA Data Bank of Japan). Nucleotide diversity (π) (Nei and Li 1979) and scaled mutation rate (θ_w) (Watterson 1975) were calculated for the five cpDNA regions using DnaSP ver.5.10 (Librado and Rozas 2009). Gaps were excluded from the calculation. Neutrality tests of Tajiam's *D*

(Tajima 1989) and Fu and Li's (1993) D^* and F^* were also performed by using DnaSP software, and the statistical significance was determined by coalescent simulation implemented into the software. Sequence variations were manually assembled in haplotypes primarily based on nucleotide substitutions. Variations in mononucleotide repeat number accounted for subtypes because of their higher mutation rate (Okaura and Harada 2002). Haplotypes in the frequencies less than 5 % in the total sample are defined as rare haplotypes following Aoki et al. (2004). Haplotypes in the frequencies more than 5 % are considered as common haplotypes. To clarify the genetic relationships among the haplotypes, a haplotype network was constructed for *Q. phillyraeoides* using TCS 1.21 (Clement *et al*. 2000) using the corresponding sequence of *Quercus mongolica* var. *crispula* as an outgroup. Fixation indices for haplotype data (G_{ST}, Nei 1973) and for sequence data (F_{ST}, Hudson *et al.* 1992) were calculated by using DnaSP. Analysis of molecular variance (AMOVA, Excoffier *et al*. 1992) was applied in order to estimate the variance components for whole cpDNA regions examined using Arlequin version 3.5.1 (Excoffier *et al*. 2005). Fixation indices in hierarchical population structure, Φ_{SC}, Φ_{ST}, and Φ_{CT} were estimated using Arlequin version 3.5.1. Φ_{SC}, Φ_{ST}, and Φ_{CT} were the original notations used by Excoffier *et al.* (1992) for F_{SC}, F_{ST}, and F_{CT}, respectively. We used these notations to avoid confusion with Hudson *et al.*'s F_{ST}.

3.3 Results

3.3.1 cpDNA variation

A total of 3,665 bp excluding alignment gaps was sequenced. The lengths of the *trn*L intron, *trn*L-*trn*F spacer, *atp*B-*rbc*L spacer, *mat*K gene, and *trn*H-*psb*A spacer were 532 bp, 444 bp, 790 bp, 1,329 bp, and 570 bp, respectively. Four nucleotide substitutions, two mononucleotide repeat variations and one single nucleotide deletion were identified in 95 individuals. π and θ_w were calculated for each locus with Tajima's *D* and Fu and Li's *D** and *F** (Table 3.2). Among the five cpDNA regions, the *mat*K gene showed the highest nucleotide diversity. No variation was found in the *trn*L intron and the *trn*L-*trn*F spacer. Both the two substitutions in the *mat*K gene are non-synonymous (T > C transition at position 177 corresponds to a change of Phe to Leu, and the T > C transition at position 1,098 corresponds to a change of Leu to Ser). Tajima's *D* and Fu and Li's *D** and *F** were negative for all three groups, but not significant for the entire sequence. *D** and *F** were significant at the 5 % level when the *mat*K gene was excluded.

3.3.2 Geographical distribution of haplotypes

Five cpDNA haplotypes (haplotypes A, B, C, D, and E) were identified based on nucleotide substitutions (Table 3.3). Both haplotype A and haplotype D were divided

into two subtypes: A1 and A2, and D1 and D2, respectively. Haplotype frequencies were 51.6 % in type A1, 1.1 % in type A2, 1.1 % in type B, 6.3 % in type C, 6.3 % in type D1, 32.6 % in type D2, and 1.1 % in type E. Types A1, C, D1, and D2 were considered common haplotypes (with frequencies greater than 5 %). Among them, A1 and D2 were the most frequent haplotypes. Types A2, B, and E were rare haplotypes (with frequencies less than 5 %). Type A1 was identified in samples from all the regions investigated except Okinawa and was fixed in six populations of IW, OG, HZ, SY, TM, and AS, whereas type A2 was found only in the population NG and was polymorphic with type A1. Type D2 was found in Chugoku, Shikoku, Kyushu and Okinawa, but not in Kinki, Chubu and Kanto. Type D2 was fixed in five populations of TK, SG, TS, US and IZ. Type D1 was found in two populations in SD and ST. Both populations were polymorphic with type D2. Types B and E were found uniquely in the populations MR and RY, respectively. Both were polymorphic with type A1. Type C was found only in the population KS and the nearby population TB in Kii Peninsula. The two major haplotypes A1 and D2 were polymorphic in the population NM in Kyushu. Haplotype frequencies in the populations are summarized in Table 4 and their geographical distribution is shown in Fig. 3.1.

3.3.3 Phylogenetic relationships among haplotypes

Haplotype network shows type A1 is ancestral to the other haplotypes (Fig. 3.2). Type A2 was derived from A1 by a change of mononucleotide repeat number. Types B, C,

and D1 were derived from type A1 by single-nucleotide substitutions. Type D2 was derived from type D1 by a change of mononucleotide repeat number. Type E was derived from type A1 with one nucleotide substitution and one single nucleotide deletion (Table 3.3).

3.3.4 Genetic differentiation of *Q. phillyraeoides*

G_{ST} for haplotype data was calculated to be 0.833 for the total populations, while F_{ST} for sequence data was 0.858. Both values indicated that very large genetic differentiation had occurred among the populations. G_{ST} for Groups 1, 2, and 3 were 0.657, 0.809, and 0.526, respectively. F_{ST} for Groups 1, 2, and 3 were 0.700, 0.873, and 0.643, respectively. AMOVA was performed in order to test the genetic differentiation among the groups and among the populations (Table 3.5). The difference among the groups was significant at 5 % level (Φ_{CT} = 0.289, p = 0.027). Largest variation (54.1 %) occurred among populations within groups. Fixation indices were 0.761 for Φ_{SC}, and 0.830 for Φ_{ST}. Both the differentiation among populations within groups (Φ_{SC}) and the differentiation among populations in the whole area (Φ_{ST}) were highly significant (p < 0.001) by 1,023 permutations.

3.4 Discussion

Relatively small levels of genetic variation were revealed in cpDNA in *Quercus*

phillyraeoides populations. If we exclude the *mat*K region, π and θ_w for non-coding regions become 0.00002 ±0.00001 and 0.00017 ±0.00012, respectively (Table 3.2). The average nucleotide diversity for the same four non-coding regions in the Japanese oak, *Q. mongolica* var. *crispula* in deciduous forest was 0.00074 ±0.00004 (Okaura *et al*. 2007). The values in non-coding regions of cpDNA for three tree species in lucidophyllus forest in Japan, including *Prunus zippeliana*, *Daphne kiusiana*, and *Elaeocarpus sylvestris* var. *ellipticus* were 0.00084, 0.00027, and 0.00031, respectively (Aoki *et al.* 2004). The values obtained for typical members of lucidophyllus forests in Taiwan and surrounding area were 0.00039 for *Castanopsis carlesii* (Cheng *et al.* 2005), 0.00065 for *Quercus glauca* (Huang *et al.* 2002), and 0.00031 and 0.00051 for *Machilus thunbergii* and *M. kusanoi*, respectively (Wu *et al.* 2006). Although the selected regions are not the same, these values are much larger than that of *Quercus phillyraeoides*. Tajima's *D* and Fu and Li's *D** and *F** for the combined dataset showed negative values, but none were significant. Although Fu and Li's *D** and *F** were marginally significant if we exclude *mat*K gene, this does not clearly indicate the occurrence of a bottleneck event in the past. This probably means that many populations have been maintained in small, isolated populations for long time.

Refugial populations are sometimes defined as regions rich with rare haplotypes, or as regions rich in number of common haplotypes (Aoki *et al.* 2004). The populations MR on the Pacific coast and RY on the East China Sea are polymorphic with common and rare haplotypes, respectively. Muroto Peninsula, where the population MR locates, and Amakusa-jima, where the population RY locates, have been proposed to be refugia

of lucidophyllus forests in the LGM (Hattori 1985). The population NG was polymorphic with common type A1 and rare haplotype A2. Type A2 was probably derived from A1 by a *de novo* mutation in this population. The population KS in Kii Peninsula is polymorphic with two common haplotypes A1 and C. Type C was also found in TB, which were 148 km from KS, but not found in population HZ, which is only 30 km from TB separated by the water of Ise Bay. Kii Peninsula has also been mentioned as one of refugia of lucidophyllus forests in the LGM (Hattori 2002). This suggests that KS is a refugium and type C migrated from KS northeastward to form the population TB and fixed. Both the populations SD and ST are polymorphic with prevailing type D2 and related type D1. TCS network (Fig. 3.1) shows D2 is a derived type from D1 by parsimony. However, type D2 is a prevailing haplotype dominating in Kyushu and on the coast of Seto Inland Sea. This probably means that D2 is derived from A1 first by one nucleotide substitution, and a change in mononucleotide repeat number. D1 could be derived from D2 by restoring the same repeat number of A1. One possible scenario to explain the same status in the populations SD and ST is that either one of the populations was a refugium and the other one was colonized by a population migrated from that refugium. However, this is not likely because the direct distance between SD and ST is 290 km, which is nearly the estimated maximum range of vegetation expansion (300 km) of lucidophyllus forest trees proposed by Hattori (2002). Another possibility is that D1 occurred in each population independently, because any two-adenine additions at the *trn*H-*psb*A spacer mononucleotide repeat region of D2 could be accounted as D1 (see Table 3.3). The study of ancient climates suggested that

southern Kyushu have been the only possible refugia of lucidophyllus forests in the LGM (Tsukada 1974). Existence of refugia in Kyushu is also suggested by fossilized pollen data from the last glacial period (Matsuoka and Miyoshi 1998). This may support the status of ST as a refugium, but this should be confirmed by examining intermediate populations between SD and ST. Population NM is polymorphic with the prevailing types A1 and D2. This could most probably be an admixture of two lineages with A1 and D2, because these haplotypes were found as major haplotypes in the neighboring populations of RY and ST, respectively. IZ population is fixed with type D2, which is commonly found in Kyushu and on the Seto Inland Sea coast. The Ryukyu Islands had been connected with the southwestern Japan and Taiwan to the mainland China via land bridge from the Pliocene to the early Pleistocene (2.6 million years ago). The land bridge was disconnected and the Ryukyu Islands were isolated in accordance with submergence in the middle of Pleistocene (Kizaki 1982). Several waves of migration from the dissemination center in the subtropical southeastern coast of China might have occurred through the land bridge to Japan during the Pliocene (Xie *et al*. 2011). Probably type D2 was the prevailing type at the southeastern coast of China and had migrated to the Ryukyu Islands and Kyushu at that time. The IZ population might have persisted in Izena-jima since the isolation of Ryukyu Islands.

The study of the intraspecific phylogeographic pattern of cpDNA variation in lucidophyllus forest species suggested the existence of possible important refugia in Muroto and Kii Peninsula beside southern Kyushu (Aoki *et al.* 2004). Our study suggests that RY in Amakusa-jima, MR in Muroto Peninsula, and KS in Kii Peninsula,

and possibly ST in Sata Peninsula, are the refugial populations for *Q. phillyraeoides*. This means *Q. phillyraeoides* survived in these refugia during the LGM together with other members of lucidophyllus forest trees. On the other hand, our study does not show any evidence that any of the populations in the Seto Inland Sea were the possible refugia in the LGM.

Genetic variations for each areal group were summarized in Table 3.6. Highest genetic variation is shown in Group 3. The result of AMOVA showed significant differentiation among the three areal groups at the 5 % level (Table 3.5). This suggests that the grouping by sea areas reflects genetic differentiation caused by the difference in composition of source populations or refugia. Although it is not statistically significant, consistent negative values in Tajima's *D*, and Fu and Li's *D** and *F** were obtained in Group 1. This means that there is a trend of population expansion in Group 1. The populations on the Pacific Ocean side may have experienced bottleneck events during the LGM or earlier glaciations. An opposite trend is apparent in Group 2. The Seto Inland Sea used to be terrestrial at the LGM, so that the climatic conditions may not have been suitable for *Q. phillyraeoides* considering the conditions in the present habitat. *Quercus phillyraeoides* populations on the Seto Inland Sea are thought to have migrated and colonized from refugia in southern Kyushu or on the Pacific coast after the LGM. The Seto Inland Sea is open to the Pacific Ocean through two straits, the Bungo Strait on the west and the Kii Strait on the east. Type D2 probably migrated from refugia in southern Kyushu along the opening and extension of the Bungo Strait to the north, and type A1 probably migrated from the refugia on the Muroto Peninsula or on the Kii

Peninsula following the opening of the Kii Strait. Largest values for G_{ST} and F_{ST} in Group 2 indicate extensive isolations have proceeded in the populations on the Seto Inland Sea coast. They have often fixed by random genetic drift as shown by alternative pattern of the haplotype distribution. Six of the eight populations in Group 2 are monomorphic, and four are fixed with type D2 (others are A1). Positive values of Tajima's *D*, and Fu and Li's *F** and *D** indicates founder effect according to migration from the Pacific coast or from southern Kyushu. Furthermore, five of the seven populations in Group 1 were monomorphic, and four were fixed with type A1. Although the sample size in this study is too small to be conclusive, fixation of haplotypes must have occurred frequently because of the smaller effective size of the chloroplast genome compared to the nuclear genome. This must be confirmed in future studies using highly polymorphic markers, such as microsatellites, with a larger number of samples, which are now underway. Our study estimated several refugial populations for *Q. phillyraeoides*, and most of them were consistent with the refugia identified for other lucidophyllus species (Aoki *et al*. 2004, 2006). Conservation effort should be concentrated on these populations.

3.5 Conclusions

19 *Q. Phillyraeoides* populations from the central to western part of Japan were collected and cpDNA sequence variations were examined. I analyzed the genetic diversity and differentiation of the populations for four non-coding and one coding

region and examined a total of 3,665 bp. Total nucleotide diversity (π) was calculated to be 0.00017 $\pm$ 0.00001, and the scaled mutation rate (θ_w) was 0.00021 $\pm$ 0.00011. These values were relatively smaller than those formerly reported for Japanese deciduous oak species (*Q. mongolica* var. *crispula*). When only non-coding regions were considered, nucleotide diversity was much lower than that reported for other evergreen lucidophyllus forest tree species. A total of five haplotypes, two of which included two subtypes (A1, A2, B, C, D1, D2, E), were identified, with types A1 and D2 the most frequent. A large degree of genetic differentiation had occurred across the populations (G_{ST} = 0.833 for haplotype data and F_{ST} = 0.858 for sequence data). The data suggested several possible refugia populations on the Pacific Ocean coast and the East China Sea coast, nearly identical to those previously suggested for the lucidophyllus forest refugia. There was no evidence showing that any part of the Seto Inland Sea populations could be possible refugia. Largest genetic differentiation across the populations in the Seto Inland Sea area suggested that these areas had been recently colonized following the last glacial maximum from the refugia in the Pacific coast or in southern Kyushu.

4. Genetic variations and genetic differentiation of *Quercus phillyraeoides* (Fagaceae) in Japan revealed by microsatellite markers

4.1 Introduction

Geographically *Q. phillyraeodies* distributes in subtropical climate zone in China at 23 °- 24 N, 100 °- 120 °E, and moderate climate zone in Japan at 26 °- 37 °N, 128 °- 140 ° E. Vertical distribution is almost altitude over 500m in China, while it becomes lower in Japan. The distribution in modern times can be divided into five areas including eastern part of China (distribution center), southern part of China, western part of China, southwest part of China, and western part of Japan (Xie *et al*. 2011). The distribution in Japan consists of southwestern Honshu, Shikoku, Kyushu, and Ryukyuan Okinawa. The northeastern distribution limit is Kyonan-machi (Chiba prefecture) along the Pacific side of Japan (Hara *et al*., 2000.). *Q. phillyraeoides* is often used as gardening (ornamental tree) or bonsai, and especially for producing high-grade white charcoal called Binchotan. Nowadays *Q. phillyraeoides* populations in many places are considered to be critically endangered, endangered, near threatened, or vulnerable (Chiba Biodiversity Center 2009; Kanagawa Prefecture 2006; Nagasaki Prefecture 2011; Miyazaki Prefecture 2010; Okinawa Prefecture 2006).

Previously, we elucidated the phylogeographical relationships and found the past colonization routes of *Q. phillyraeoides* by examining chloroplast DNA (cpDNA) sequence variation. CpDNA sequence data suggested possible refugia population on the

Pacific Ocean (MR) and the East China Sea coast (RY), and KS population in Kii Peninsula. These nearly identical to those previously suggested for the lucidophyllus forest refugia which existed in Muroto and Kii Peninsula beside southern Kyushu (Aoki *et al*. 2004). But we could not find any evidence showing that any part of the Seto Inland Sea populations could be possible refugia. *Q. phillyraeoides* species may survive in these refugia above mentioned during the LGM together with other members of lucidophyllus forest trees. Although the sample size used in cpDNA variation was too small (five individuals for each population) to be conclusive, fixation of haplotypes must have occurred frequently because of the smaller effective size of the chloroplast genome compared to nuclear genome. This must be confirmed in this study using highly polymorphic markers (microsatellites) with large number of samples.

Microsatellites are defined as simple sequence repeats (SSRs) with a repeats length of up to 13 bases, whereas longer repeats give rise to minisatellites (Greg Gibson and Spencer V. Muse 2009). These repetitive DNA sequences are randomly distributed throughout eukaryotic genomes (K. S. Kim *et al*. 2001). They are usually highly polymorphic and exhibit a large number of alleles. The high degree of polymorphism and the codominant mode of inheritance make them extremely informative and suitable for analysis of gene flow and for characterization of the mating system in many plant species. Microsatellites display higher levels of variation, and consequently enable clearer clustering of the populations (D. B. Goldstein, *et al*. 1995). Differences in the number of repeats at microsatellite loci are commonly recognized by differences in sizes of the amplified fragments (Tautz, 1989; Weber and May, 1989). Microsatellite

markers provide a powerful tool for estimating the amount of intraspecific and interspecific gene flow (Dow and Ashley, 1998) and within-population genetic structure (Streiff et al., 1998). They have been widely used to study population genetics, evolutionary processes and conservation or management of biological resources (Jarne and Lagoda, 1996).

The aims of this chapter are to provide information on distribution of genetic diversity within and among *Q. phillyraeoides* populations in Japan and also to elucidate genetic structure of this species. The information may provide us the past migration and expansion process after LGM and also testify the status of the refugia populations suggested in the previous study in chapter 3.

4.2 Materials and methods

4.2.1 Sample collection and DNA extraction

Q. phillyraeoides is a common oak in western Japan. We collected 19 populations covering the distribution of this species in Japan: 1 population in Kanto, 2 populations in Chubu; 2 populations in Kinki; 7 populations in Shikoku; 2 populations in Chugoku; 4 populations in Kyushu; and 1 population in Okinawa. Iwai (IW) population in Chiba Prefecture and Izenajima (IZ) population in Okinawa Prefecture are the northernmost and the southernmost populations, respectively. Fresh leaves were collected from a total of 423 trees in 19 populations (Table 3.1 and Fig. 3.1). The 19 populations were divided

into three groups (designed 1-3), based on their geographical locations. Group 1 includes 7 populations on the Pacific coast. Group 2 includes 8 populations in the Seto Inland Sea. Group 3 includes 4 populations in Kyushu and Okinawa (Table 3.1).

DNA was extracted from fresh leaves using a modified CTAB method (Doyle and Doyle, 1990). The quality and concentration of DNA were confirmed by electrophoresis on 0.8% agarose gels along with a *λ-Hin*dIII digest as DNA size standard. The extracted DNA was stored at 4 ℃ until required.

4.2.2 Microsatellite markers and genotyping

Nine microsatellite markers were confirmed to be amplified in *Q. phillyraeoides*: bcqm07, bcqm42, bcqm74, bcqm76, and bcqm94 were developed for *Q. mongolica* var. *crispula* (Mishima K. *et al*., 2006), QM58TGT and QM69-2M1 were developed for *Q. myrsinifolia* (Isagi Y. and Suhandono S., 1997), QpZAG36 was developed for *Q. petraea* / *Q. robur* (Steinkellner H. *et al*., 1997), and MSQ13 was developed for *Q. macrocarpa* (Dow BD *et al*., 1995). Some of aforementioned primers were modified for *Q. phillyraeoides* (Table 4.1). The sequences of reverse primer for the locus bcqm07 was CCCCACATCAAAGAACTATTG, forward primer for the locus QM58TGT was GGTCAGTGTATTTTGTTGCT, reverse primer for the locus QM69-2M1 was CAATCTGCCCACATCAGC, respectively (Table 4.1).

PCR amplification was performed with using Type-it Microsatellite PCR Kit (Qiagen). Nine primer pairs, which corresponded to nine loci were divided into three

sets: A, B. and C. Set A included primer pairs bcqm07 (FAM), bcqm42 (HEX), and bcqm74 (NED). Set B included primer pairs bcqm76 (FAM), bcqm94 (HEX), and QM58TGT (NED). Set C included primer pairs QM69-2M1 (FAM), QpZAG36 (HEX), and MSQ13 (NED). The forward primer of each pair was labeled with a fluorescent dyes (FAM, HEX, or NED) to enable the detection of PCR products in ABI PRISM 310 Genetic Analyzer (Applied Biosystems). For each primer set mixture, it contained 0.2 μM of each primer in a total volume of 50 μL. Multiple PCR contained 3 μL of RNase free water, 5 μL of 2× Type-it Multiple PCR Maseter Mix, 1 μL of primer set mixture, and 1 μL of template DNA (200 ng for each PCR reaction). The PCR conditions included an initial denaturation at 95℃ for 5 minutes, followed by 28 (for set A) or 31 (for set B and set C) cycles at 95℃ for 30 seconds, 49℃ (for set B and set C) to 52℃ (for set A) for 90 seconds, and 72℃ for 30 seconds, with a final extension at 60℃ for 30 minutes, keep in 4℃ forever. PCR was conducted with a GeneAMP® PCR System 9700 Thermal Cycler (Applied Biosystems) and Veriti 96 Well Thermal Cycler (Applied Biosystems). Then electrophoresis was performed on a 0.8 % agarose gel to determine the concentration of the amplified DNA. PCR products were diluted into 200 times and applied on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Fragment size was determined by GeneScan software (Applied Biosystems).

4.2.3 Statistical data analysis

4.2.3.1 Genetic diversity within populations

The statistical analysis of the microsatellite data was performed with Genetic analysis in Excel (GenAlEx) version 6.4.1 software (Peakall R and Smouse PE, 2006). The total number of detected alleles (N_a), number of effective alleles (N_e), the average observed and expected heterozygosity within populations (H_O and H_E), unbiased expected heterozygosity (UH_E), Shannon's information index (I), fixation index (F) were estimated. Alleles were characterized as private (Ap) if they presented a frequency more than 5% in one population and did not occur in any other population (Payne et al. 2008). The allelic richness was determined for 17 diploid individuals as the minimum population size, and. Allelic richness were detected by FSTAT V2.9.3.2 software (Weir and Cockerham 1984).

MICRO-CHECKER software (van Oosterhout et al. 2004) was used to check for genotyping errors and to estimate the existence of null alleles. The program constructs random genotypes by randomizing the observed alleles for each locus within samples and then compares the observed genotypes with the distribution of randomized genotypes.

4.2.3.2 Population genetic structure

Wright's inbreeding coefficient in the total populations (F_{IT}) and within populations (F_{IS}), together with genetic differentiation based on allele identity (F_{ST}), number of migrants (Nm) were determined. G_{ST} defined by Nei (1973) was also determined. R_{ST}

for over all samples were estimated following Rousset (1996) and Goodman (1997). All above testes were done by using FSTAT V2.9.3.2 software (Weir and Cockerham 1984). Estimation of pairwise F_{ST} between populations and its significance test was also done by using FSTAT V2.9.3.2 software. A genetic distance matrix of pairwise F_{ST} values was also used to perform a hierarchical analysis of molecular variance (AMOVA) (Excoffier *et al*. 1992) with Arlequin software version 3.1 (Excoffier *et al*. 2005). AMOVA was also used to partition the total molecular variances into among areas, among populations within areas, and within populations.

Structure software version 2.3 (Rritchard *et al*., 2009) was used to detect population genetic structure. The program Structure implements a model-based clustering method for inferring population structure using genotype data consisting of unlinked markers. To estimate the number of clusters (K), ΔK, which is based on the rate of change in the "log probability of data" (Ln P(D)) between successive K value, was calculated (Evanno *et al*. 2005). There are two approaches to determine the best K. One is according to Ln P(D) and the other is according to ΔK. In the first approach, when K is approaching a true value, Ln P(D) plateaus (or continues increasing slightly) and has high variance between runs (Rosenberg *et al*. 2001). In the second approach, the calculation is based on the second order rate of change of the likelihood (ΔK) (Evanno *et al*. 2005). The ΔK shows a clear peak at the true value of K. As the number of total populations was 19, we chose K equaled from 1 to 20. Structure Harvester software was used to calculate ΔK and determine best K (Dent A. Earl and Bridgett M. vonHoldt, 2012).

Populations were tested for a recent reduction of their effective population size

(Cornuet and Luikart 1997) using the BOTTLENECK software (Piry et al. 1999). The software assumes the mutation-drift equilibrium. The software calculated the deviation in observed gene diversity from the expected equilibrium value and computed from the observed number of alleles. If a population suffers from a recent bottleneck, a larger reduction of alleles number at neutral loci is observed compared to the gene diversity. The two-phased model of mutation (TPM) was used with 95% stepwise mutation model (SMM) and 5% multistep mutations. The TPM is intermediate to the SMM and IAM (infinite alleles model). The Wilcoxon signed-rank test was used for the estimation of the significance of heterozygote excess.

4.3 Results

4.3.1 Genetic diversity within populations

Eight of the nine SSR loci were polymorphic in a total of 423 individuals over 19 populations *Q. phillyraeoides* in Japan. Average number of alleles over all populations for eight polymorphic loci was 2.993 ranging from 1.421 (locus bcqm07) to 4.105 (locus bcqm76) (Table 4.2). The results of the genetic variation using the eight polymorphic microsatellite loci are summarized in Table 4.2.

The average expected heterozygosity (H_E) for over all loci ranged from 0.312 (population MR) to 0.457 (population TB) with the average of 0.383 (Table 4.3). These are rather smaller compared to the previous study of deciduous oak, *Q. mongolica* var.

crispula (0.724, Ohsawa et al., 2011). Inbreeding coefficient, *F*, indicates heterozygosity excess. Over 19 populations it ranged from -0.042 (AS) to 0.210 (RY) with the average of 0.074 (Table 4.3).

The allelic richness, *As*, per locus and over all populations was estimated based on the minimum population size of 17 individuals. The value ranged from 1.931 (locus MSQ13) to 4.623 (locus bcqm76) and averaged 3.466 (Table 4.2). RY population in Kyushu showed maximum allelic richness (3.771), and MR population in Shikoku showed minimum (2.236) (Table 4.2). No evidence of null alleles was found as estimated using MICRO-CHECKER (Table 4.8).

In a total eight private alleles were detected, two in population IZ, one in each of OG, HZ, TB, TS, RY, and ST populations (Table 4.4).

4.3.2 Genetic differentiation among populations

Total F_{ST} was 0.131* (Table 4.2) for over all loci. According to the Wright's standard, *Q. phillyraeoides* populations in Japan had moderate to large genetic differentiation, which is significantly larger than zero at the level of 5%. R_{ST}, which was obtained by analysis of variance of alleles, was 0.081 (Table 4.2). The result of AMOVA was listed in Table 4.5. Genetic variance was divided into three components: among groups, among populations within groups, and within populations with respective genetic variance components of 1.4%, 8.8% and 89.8%. All are shown to be significantly larger than zero. Largest genetic variation occurred at the level of within population among

individuals. Pairwise F_{ST} was calculated in all combinations of populations (Table 4.6). IW, ST and IZ populations showed highly significant differentiation in any pairs of the other populations. Whereas, KS, AS, SY and NM populations showed more than half of the combinations (9 of 18) were not significantly different from zero.

4.3.3 Population genetic structure

The best K, the number of clusters was determined by using LnP(D). When K is approaching a true value, LnP(D) plateaus (or continues increasing slightly) and has high variance between runs (Rosenberg et al. 2001). We chose K=4 as the best K following this criteria. Structure Harvester software also showed that K to be 4 as the best K (Fig. 4.1). The bar plot for K = 4 for over all populations was shown in Fig. 4.2a. The northeastern most population Iwai (IW) was specifically dominated by cluster I (87.5%), while the southernmost population Izena-jima (IZ) was specifically dominated by cluster III (95.8%). Other populations were admixed with the four clusters at various degrees. Cluster I was dominated (more than 50%) in SG and TK beside IW. Cluster II dominated in KS, MR, SY, TM, TS and NG. Cluster III dominated in ST beside IZ. Cluster IV dominated in HZ. In order to make the clustering clearer in each geographic regions we did STRUCTURE analysis separately in each group. Group 1 includes populations on the Pacific coast. Structure Harvester indicated the best K to be three. The bar plot was shown in Fig. 4.2b. Three clusters corresponding cluster I, II and IV from the total bar plot in Fig. 4.2a (we defined these to be clusters I', II' and IV'). One

cluster (cluster I') was dominate in IW, and cluster IV' dominated in HZ. Other populations were admixture of cluster I', II' and IV' at various degrees, but the cluster I' was the least frequent. Group 2 includes populations on the Seto Inland Sea. Structure Harvester indicated that K equaled to 2 was the best and these are corresponding cluster I and II in the total bar plot in Fig. 4.2a (we defined these to be cluster I' and II'). In the populations SG, SD and TK, cluster I' was dominant while in populations SY, TM and NG, cluster II' was dominant. Populations TS and US were the admixture of the two. Group 3 includes populations on the East China Sea and the best K was 3 according to the result of Structure Harvester. These clusters were assigned as cluster IV', III' and III" corresponding to the total bar plot in Fig. 4.2a. Population ST was dominated with cluster III' and population IZ was dominated with cluster III". This shows the genetic differentiation exist between population ST, in the southernmost Kyushu and IZ in Okinawa. This was supported by highly significant pairwise F_{ST} between IZ and ST.

4.3.4 Detection of recent bottleneck

BOTTLENECK analysis was done for all the populations (Table 4.10). All populations did not show significant heterozygosity excess, and suggested that *Q. phillyraeoides* populations had not experienced a recent bottleneck. The Wilcoxon signed-rank test revealed a normal L-shaped distribution for most of populations. But some populations revealed a skewed distribution, i. e. AS and MR populations in Shikoku, NG population in Chugoku, IZ population from Okinawa. These skewness is

trivial and all investigated populations are conclude to have not experienced a recent bottleneck.

4.4 Discussion

In a previous study (chapter 3), several populations were proposed as the LGM refugia based on the criteria that the population was polymorphic with rare haplotypes (here the frequency <5%, Aoki et al., 2004). KS and MR on the Pacific coast and RY on the East China Sea coast were polymorphic with common and rare haplotypes and considered as possible refugia. These populations were formerly proposed as the refugia of tree species in lucidophyllous forest (Tsukada 1974; Hattori 1985).

The microsatellite markers examined here had low variability and a small number of alleles (ranged from 1.421 in locus bcqm07, to 4.105 in locus bcqm76), including one monomorphic locus of bcqm74. The level of polymorphism detected (observed heterozygosity was 0.346) was smaller than the value observed in other *Quercus* species (*Q. petraea* (Streiff *et al*., 1998; Cottrell *et al*., 2003; Jiri Dostalek *et al*., 2011), *Q. lobata* (Dutech et al., 2005), *Q. aquifolioides* (Xuejiang Zhang et al., 2006)) based on microsatellite loci. The observed heterozygosity being slightly lower than expected heterozygosity (F_{IS}=0.086*) in *Q. phillyraeoides* populations. It was similar to that detected in other oak species (Guttman and Weigt, 1989; Schnabel and Hamrick, 1990; Berg and Hamrick, 1993, 1994; Chung *et al*., 2002). But in this study F_{IS} was significant at the 0.1% level, which means inbreeding occurred. Heterozygosity

deficiency were found at all the polymorphic loci except for locus QM58-TGT. Positive value of F_{IS}, which occurred across the loci was indicative of inbreeding. Selfing may be the possible reason.

High levels of genetic differentiation were detected for the cpDNA sequencing (F_{ST}= 0.858) and showed sharp contrast to the nuclear microsatellite loci (F_{ST}= 0.131*) (Table 4.2). Theoretically, the level of genetic differentiation among populations is expected to be higher for maternally inherited cpDNA markers than for biparentally inherited nuclear genes (Birky et al., 1989; Petit *et al*., 1993; Ennos, 1994; Hu and Ennos, 1997). Three main factors contribute to the higher genetic differentiation for cpDNA markers: firstly, effective gene flow is limited to seed dispersal for maternally inherited DNA; secondly, drift is expected to be twice as strong for a haploid genome compared to a diploid one; and thirdly, it has been shown in some hermaphrodite tree species that flowering, seed dispersal and recruitment patterns resulted in an effective number of trees contributing to the next generation as females that is much reduced compared to the effective number of trees acting as male parents (Dow and Ashley, 1996). Moreover, differences in species's pollination ecology or seed dispersal mechanisms might have important consequences on the balance between female and male reproductive success of individual trees (Parvin *et al*., 2010).

The nine microsatellites loci we used showed low polymorphism including one monomorphic locus (bcqm74). One reason may be that these markers were not specifically developed for *Q. phillyraeoides*, and they could not well represent the genetic background of this species.

Nuclear microsatellites (nSSRs) are multiallelic within and among populations and inherited in a codominant fashion (Morgante and Olivieri 1993). Population structure will be developed in subdivided populations. This process decreases in heterozygosity and generates genetic differentiation through natural selection and genetic drift. F_{ST} was calculated to be 0.131*. According to Wright's F standard, *Q. phillyraeoides* populations in Japan had moderate genetic differentiation.

The results of AMOVA (Table 4.5) showed significant differentiation among three defined areal groups at the 5% level. This suggested that the grouping by sea areas reflects genetic differentiation caused by the difference in composition of source populations or refugia. These results were similar with analyzing by using cpDNA variation. cpDNA analysis in previous paper showed largest genetic differentiation across the populations in the Seto Inland Sea are. It suggested that these areas had been recently colonized following the last glacial maximum from the refugia in the Pacific coast or in southern Kyushu. The Seto Inland Sea used to be terrestrial at the LGM. So that climatic conditions may not have been suitable for *Q. phillyraeoides* considering the conditions in the present habitat. *Q. phillyraeoides* populations on the Seto Inland Sea are thought to have migrated and colonized from refugia in southern Kyushu or on the Pacific coast after LGM. The Seto Inland Sea is open to the Pacific Ocean through two straits, the Bungo Strait on the west and Kii Strait on the east.

Most of populations contain individuals made up of all four clusters according to Structure software (Fig. 4.2 and Fig. 4.3). IW population did not contain cluster II. SG and NG populations did not contain cluster III. TM population did not contain cluster I.

TK population did not contain cluster IV. ST and IZ populations contained two shared clusters (III and IV). ST population was nearest to Okinawa population IZ. And these two populations both showed that cluster III was the dominant cluster.

4.5 Conclusions

Genetic structure of *Q. phillyraeoides* was determined by using nine microsatellite markers in 423 individuals of 19 populations covering the distribution in Japan. Results showed that alleles number ranged from 1.421 (locus bcqm07) to 4.105 (locus bcqm76) at the population level. Expected heterozygosity ranged from 0.312 (population MR) to 0.457 (population TB) with the average of 0.383. Genetic differentiation among populations was 0.131 showing significant, but moderate genetic differentiation. Genetic variations mainly allocated within population (89.83%) according to AMOVA. STRUCTURE analysis showed that *Q. phillyraeoides* populations could be defined into four clusters. The northernmost population (IW) and the southernmost population (IZ) were most genetically differentiated. These populations are suggested to be the important candidates for conservation as well as refugia populations previously defined. The results of BOTTLENECK showed that *Q. phillyraeoides* in Japan had not experienced a recent bottleneck.

Table 2.1 Location of the populations, numbers of samples, haplotypes, and chloroplast types detected in *Quercus mongolica* var. *crispula*

District	Population‡	Location	No. sample	Combination of haplotype and chloroplast type §				
				I-T	II-C	VI-T	XI-T	XII'-T
Harbin	Mt. Mao'er (Ma)	45˚20'N, 127˚18'E	10	0	0	0	0	10
Primorski Krai	Spassk-Dalny (Sd)	44˚41'N, 132˚38'E	5	0	0	5	0	0
	Roshino (Ro)	45˚54'N, 134˚51'E	5	0	0	0	5	0
	Okeanskaya (Ok)	43˚15'N, 132˚00'E	3	0	0	3	0	0
Sakhalin	Storodubskoe (St)	47˚25'N, 142˚48'E	25	0	0	23	2	0
	Kostromoskoe (Ko)	47˚17'N, 142˚01'E	34	0	0	19	15	0
	Ohotsukoe (Oh)	46˚51'N, 143˚10'E	7	0	0	7	0	0

Hokkaido	Wakkanai (Wk)	45˚20’N, 141˚45’E	10	0	10	0	0	0
	Eiura (Ew)	44˚07’N, 143˚58’E	7	7	0	0	0	0
	Shikaoi (Sk)	43˚06’N, 142˚59’E	8	7	1	0	0	0
	Hidaka (Hd)	42˚52’N, 142˚26’E	8	7	1	0	0	0
	Ohnuma (On)	41˚58’N, 140˚40’E	8	0	8	0	0	0
Tohoku	Mt. Shirakami (Sr)	40˚31’N, 140˚15’E	10	0	10	0	0	0
	Tazawako (Tz)	39˚44’N, 140˚43’E	8	0	8	0	0	0
	Mt. Hayachine (Hy)	39˚32’N, 141˚29’E	8	8	0	0	0	0
	Izumigadake (Iz)	38˚24’N, 140˚42’E	8	0	8	0	0	0
	Iizaka (Ik)	37˚48’N, 140˚27’E	2	0	2	0	0	0

Chubu	Myoko (My)	36˚52’N, 138˚13’E	2	0	2	0	0	0
Kanto	Mt. Takahara (Tk)	36˚54’N, 139˚46’E	8	0	8	0	0	0
	Chichibu (Cc)	35˚59’N, 139˚04’E	2	0	2	0	0	0

Note: †The species sampled in China and Russia are *Quercus mongolica* var. *mongolica*. ‡The numbers in parentheses are abbreviations for the location. §Combination of haplotypes defined by Okaura *et al*. (2007) and chloroplast types defined by Kanno *et al*. (2004).

Table 2.2 Nucleotide variations of newly defined cpDNA haplotypes in

Quercus mongolica var. *crispula* and *Q. mongolica* var. *mongolica*

	Nucleotide(s) of variable sites‡												
	*trn*T-*trn*L		*trn*L-*trn*F		*atp*B-*rbc*L		*mat*K			*trn*H-*psb*A			*trn*Q-*trn*S
	1	2	5	1	7	7	1	4	6	1	1	2	2
	2	1	4	2	4	7	3	5	7	5	7	7	7
	7	3	6	1	5	0	2			3	4	4	6
Haplotype†							1						
I	T	A	A	T_{11}	C	-	T	A	-	C	-	A_9	T
II	•	•	•	T_{11}	T	-	G	•	-	•	-	A_9	C
VI	•	•	•	T_{10}	T	-	•	•	-	•	-	A_{10}	•
XI	•	C	C	T_{12}	T	-	•	C	-	-	G	A_9	•
XII’	G	•	•	T_{10}	T	I	•	•	I	•	-	A_9	•

Note: † Haplotpes were newly defined after Okaura *et al*. 2007 including *trn*Q-*trn*S intergenic spacer region. ‡Dots represent the same

nucleotide as shown in haplotype I. Mononucleotide repeat variations are indicated as nucleotide followed by the repeat number. Minus sign indicates absence of the nucleotide(s). “I” at position 770 of *atp*B-*rbc*L spacer indicates the 30 bp insertion of “TGAGTTGTAGGGAGGGACTTATGTCACCAC”, and at position 67 of *trn*H-*psb*A intron indicates the 27 bp insertion of “ATTTTAATTTCTACCCATTCCTCTTGTT”.

Table 3.1 Collection sites of *Q. phillyraeoides* in western Japan

Districts		Location	Latitude (N)	Longitude (E)	Code	Pop Size	Group
Kanto		Iwai	35 °06'07"	139 °49'55"	IW	32	1
Chubu		Ogasayama	34 °44'44"	137 °58'16"	OG	26	1
		Hazu	34 °41'51"	136 °58'13"	HZ	29	1
Kinki		Toba	34 °27'17"	136 °52'16"	TB	18	1
		Kushimoto	33 °26'10"	135 °45'44"	KS	20	1
Shikoku	the Pacific coast	Muroto	33 °17'24"	134 °9'7"	MR	19	1
		Ashidsuri	32 °46'54"	132 °57'18"	AS	20	1
	the Seto Inland Sea	Shyodoshima	34 °30'35'	134 °21'08"	SY	36	2
		Toomiyama	34 °23'18"	134 °8'33"	TM	20	2
		Suga	34 °20'44"	133 °54'13"	SG	23	2

	Sada	33 °22'35"	132 °6'58"	SD	20	2
	Toshima	33 °12'23"	132 °22'9"	TS	20	2
Chugoku	Takarahachimangu	34 °43'23"	134 °14'19"	TK	19	2
	Nagahama	34 °20'00"	132 °58'38'	NG	20	2
Kyushu	Usuki	33 °08'22"	131 °48'39"	US	19	2
	Ryugadake	32 °24'23"	130 °22'51"	RY	19	3
	Noma	31 °23'58"	130 °13'40"	NM	20	3
	Sata	31 °02'46"	130 °40'39"	ST	19	3
Okinawa	Izenajima	26 °54'53"	127 °56'25"	IZ	24	3
Total				19	423	

Note: Group 1 includes populations on the Pacific coast (7 populations) (IW, OG, HZ, TB, KS, MR, and AS). Group 2 includes populations in the Seto Inland Sea (8 populations) (SY, TM, SG, SD, TS, TK, NG, and US). Group 3 includes populations in Kyushu and Okinawa (4 populations) (RY, NM, ST, and IZ).

Table 3.2 Summary of nucleotide variation in five cpDNA regions in *Q. phillyraeoides*

Region	Length (bp)	S^a	π^b	$\theta_W{}^c$	Tajima's *D*	Fu and Li's *D**	Fu and Li's *F**	$H_d{}^d$
*trn*L intron	532	0	0	0	0	0	0	0
*trn*L-*trn* F spacer	444	0	0	0	0	0	0	0
*atp*B-*rbc*L spacer	790	1	0.00003 ± 0.00003	0.00025 ± 0.00025	-1.033 (p > 0.10)	-2.018 (p > 0.05)	-2.006 (p > 0.05)	0.021 ± 0.022
*mat*K gene	1,329	2	0.00045 ± 0.00003	0.00029 ± 0.00021	0.836 (p > 0.10)	0.689 (p > 0.10)	0.858 (p > 0.10)	0.551 ± 0.025
*trn*H-*psb*A spacer	570	1	0.00004 ± 0.00004	0.00034 ± 0.00034	-1.034 (p > 0.10)	-2.018 (p > 0.05)	-2.006 (p > 0.05)	0.021 ± 0.022
Total	3,665	4	0.00017 ± 0.00001	0.00021 ± 0.00011	-0.361 (p > 0.10)	-1.467 (p > 0.10)	-1.307 (p > 0.10)	0.568 ± 0.027

Total without *mat*K	2,336	2	0.00002 ± 0.00001	0.00017 ± 0.00012	-1.385 (p > 0.10)	-2.807* (p < 0.05)	-2.770* (p < 0.05)	0.042 ± 0.030

[a] *S*: number of polymorphic sites (segregating sites).

[b] π: nucleotide diversity. The numbers after plus-minus sign are standard deviations.

[c] θ_W : scaled mutation rate ($\theta_W = 4N\mu$, *N*: population size, μ: mutation rate per site per generation). The numbers after plus-minus sign are standard deviations.

[d] H_d: Haplotype diversity. The numbers after plus-minus sign are standard deviations.

Table 3.3 Nucleotide variation in cpDNA haplotypes in *Q. phillyraeoides*

Haplotype[c]	Frequency (%)	Nucleotides of variable sites[a]						
		*atp*B-*rbc*L spacer	*mat*K		*trn*H-*psb*A spacer[b]			
				1				
		4	1	0		1	2	4
		2	7	9	9	0	9	9
		1	7	8	7	9	8	8
A1	48.9	G	T	T	T	C	A_{11}	T_2
A2	1.1	-	-	-	-	-	-	T_3
B	1.1	A	-	-	-	-	-	-
C	6.7	-	-	C	-	-	-	-
D1	6.7	-	C	-	-	-	-	-
D2	34.4	-	C	-	-	-	A_9	-

E	1.1	-	-	-	A	D[d]	-	-

[a]Minus sign indicates the same nucleotide as shown in type A1

[b]Suffix numbers in nucleotides are the number of repeats in mononucleotide repeats

[c]Haplotypes with numbers 1 and 2 are subtypes with variable numbers of mononucleotide repeats

[d]One nucleotide deletion of C (cytosine)

Table 3.4 Haplotype distribution of *Q. phillyraeoides* populations

Region	Population		Haplotype				
	Location	Code	A (A1, A2)	B	C	D (D1, D2)	E
Kanto	Iwai	IW	5 (5, 0)	-	-	-	-
Chubu	Ogasayama	OG	5 (5, 0)	-	-	-	-
	Hazu	HZ	5 (5, 0)	-	-	-	-
Kinki	Toba	TB	-	-	5	-	-
	Kushimoto	KS	4 (4, 0)	-	1	-	-
Chugoku	Takara-hachiman	TK	-	-	-	5 (0, 5)	-
	Nagahama	NG	5 (4, 1)	-	-	-	-
Shikoku	Shyodoshima	SY	5 (5, 0)	-	-	-	-
	Muroto	MR	4 (4, 0)	1	-	-	-
	Toomiyama	TM	5 (5, 0)	-	-	-	-
	Suga	SG	-	-	-	5 (0, 5)	-
	Ashizuri	AS	5 (5, 0)	-	-	-	-
	Toshima	TS	-	-	-	5 (0, 5)	-
	Sada	SD	-	-	-	5 (4, 1)	-
Kyushu	Usuki	US	-	-	-	5 (0, 5)	-
	Sata	ST	-	-	-	5 (2, 3)	-
	Ryugadake	RY	4 (4, 0)	-	-	-	1

	Noma	NM	3 (3, 0)	-	-	2 (0, 2)	-
Okinawa	Izenajima	IZ	-	-	-	5 (0, 5)	-
Total			50 (49, 1)	1	6	37 (6, 31)	1

Table 3.5 Summary of AMOVA for *Q. phillyraeoides* in Japan

Source of variation	d.f.	Sum of squares	Variance components[a]	Percentage of variation (%)
Among groups	2	20.99	0.259*	28.86
Among populations Within group	16	41.34	0.486***	54.14
Within populations	76	11.60	0.153***	16.99
Total	94	73.94	0.898	100.00

[a]***: P-value by permutation test less than 0.001. *: P-value less than 0.05.

Table 3.6 Nucleotide variation in the three groups of *Q. phillyraeoides*

Group[a]	Length(bp)	S[b]	π[c]	θ_W[d]	Tajima's D	Fu and Li's		H_d[e]
						D*	F*	
1	3,665	2	0.00010 ± 0.00003	0.00013 ± 0.00010	-0.549 (p>0.10)	-0.787 (p>0.10)	-0.831 (p>0.10)	0.339 ± 0.0079
2	3,665	1	0.00013 ± 0.00001	0.00006 ± 0.00006	1.492 (p>0.10)	0.564 (p>0.10)	0.954 (p>0.10)	0.481 ± 0.0018
3	3,665	2	0.00017 ± 0.00003	0.00015 ± 0.00011	0.173 (p>0.10)	-0.593 (p>0.10)	-0.444 (p>0.10)	0.542 ± 0.0058

[a]Group 1 includes populations on the Pacific Ocean side. Group 2 includes populations in the Seto Inland Sea. Group 3 includes populations in western and southern Kyushu and Okinawa (See Table 1).

[b]S: number of polymorphic sites (segregating sites)

[c]π: nucleotide diversity. The numbers after plus-minus sign are standard deviations.

[d]θ_W : scaled mutation rate ($\theta_W=4N\mu$, N: population size, μ: mutation rate per site per generation). The numbers after plus-minus sign are

standard deviations.

[e]H_d: Haplotype diversity. The numbers after plus-minus sign are standard deviations.

Table 4.1 The description of nine microsatellite primer pairs for *Q. Phillyraeoides*

Locus	Primer sequence (5'-3')	T_a (℃)	Primer set	Repeat motif	Fluorescent Dye	Allele size (bp)	
						Min size	Max size
bcqm07	F: TTCCTTTTCCTCAGTTTGGG R: CCCCACATCAAAGAACTATTG	52	A	$(GT)_8$, $(GA)_{18}$	FAM-tailed	199	225
bcqm42	F: CAGTGGGACCTTCTTATGCC R: GCTTTGGAACAACTGCTACATC	52	A	$(GT)_{11}$	HEX-tailed	129	140
bcqm74	F: GAATGGATCTTCATTTATCGTTG R: TCTGCATATTTTCAACATACATTTAG	52	A	$(TG)_{10}$	NED-tailed	146	161
bcqm76	F: ATTAGTTTGCCTAGCTCTACCATG R: AGTCGCAGTCCCAATAGTAAAG	49	B	$(GT)_7$	FAM-tailed	180	223
bcqm94	F: CCAACGATTTGCTTTACCTG R: AGCACATTCACACATGGATG	49	B	$(AC)_9$	HEX-tailed	143	152
QM58TGT	F: GGTCAGTGTATTTTGTTGCT	49	B	$(CAA)_1$	NED-	212	

	R: AAATGTATTTTGCTTGCTCA			$_1$	tailed		
QM69-2M1	F: GGATGGACGAAGAGAAAGAT R: CAATCTGCCCACATCAGC	49	C	$(TGG)_9$	FAM-tailed	217	
QpZAG36	F: GATCAAAATTTGGAATATTAAGAGAG R: ACTGTGGTGGTGAGTCTAACATGTAG	49	C	$(AG)_{19}$	HEX-tailed	210	236
MSQ13	F: TGGCTGCACCTATGGCTCTTAG R: ACACTCAGACCCACCATTTTTCC	49	C	$(TC)_n$	NED-tailed	222	246

Note: T_a: annealing temperature.

Table 4.2 Summary of the genetic variation of *Q. phillyraeoides* for the studied loci

Locus	*N*	N_a	N_e	*I*	H_O	H_E	H_T	*F*	F_{IT}	F_{IS}	F_{ST}	R_{ST}	*Nm*	*As*
bcqm07	22.211	1.421	1.142	0.098	0.045	0.050	0.077	0.003	0.418	0.101	0.353	0.024	0.458	2.087
bcqm42	22.263	3.579	1.887	0.784	0.427	0.443	0.508	0.046	0.159	0.035	0.129	0.101	1.690	3.950
bcqm76	22.158	4.105	2.960	1.189	0.416	0.649	0.725	0.365	0.426	0.359	0.105	0.094	2.138	4.623
bcqm94	22.263	2.632	1.265	0.368	0.178	0.185	0.198	0.016	0.100	0.039	0.064	0.030	3.642	3.159
QM58-TGT	22.105	3.789	2.572	1.052	0.637	0.604	0.665	-0.054	0.043	-0.054	0.091	0.052	2.485	4.338
QM69-2M1	22.053	3.737	2.235	0.910	0.466	0.530	0.596	0.120	0.218	0.120	0.112	0.202	1.982	4.331
QpZAG36	22.053	2.947	1.955	0.777	0.483	0.470	0.531	-0.025	0.091	0.028	0.116	0.092	1.902	3.307
MSQ13	22.000	1.737	1.175	0.228	0.116	0.131	0.142	0.061	0.182	0.114	0.077	0.056	3.002	1.931
Mean	22.138	2.993	1.899	0.676	0.346	0.383	0.431	0.074*	0.205^{NS}	0.086*	0.131*	0.081	2.162	3.466

Note: *N*: No. of detected individuals in each population; N_a: No. of different alleles; N_e: No. of effective alleles; *I*: Shannon's information index; H_O: observed heterozygosity; H_E: expected heterozygosity within populations; *F*: fixation index; F_{IT}: Wright's F-statistics in the total populations; F_{IS}: Wright's F-statistics within populations; F_{ST}: Wright's F-statistics among populations; *Nm*: number of migrants; *As*:

allelic richness per locus and overall populations.

H_O, H_S, H_T, G_{ST} were estimated according to Nei's formula (Nei, 1973); F statistics (F_{IT}, F_{ST}, and F_{IS}) were estimated by Weir and Cockerham (1984)'s methods; R_{ST} Fixation index for over all samples estimated following Rousset (1996) and Goodman (1997); NS: not significant, P>0.05; *: P<0.05; **: P<0.01; ***: P<0.001.

Table 4.3 Summary of the genetic variation of *Q. phillyraeoides* for the studied populations

District	Code	N	N_a	N_e	I	H_O	H_E	UH_E	F	Ap	A_S
Kanto	IW	32.000	2.750	1.952	0.651	0.383	0.376	0.382	-0.025*	0	2.592
Chubu	OG	25.875	3.250	2.112	0.788	0.352	0.447	0.455	0.202^{NS}	1	3.100
	HZ	29.000	3.500	1.976	0.740	0.371	0.398	0.405	0.021^{NS}	1	3.187
Kinki	TB	18.000	3.500	2.148	0.836	0.375	0.457	0.470	0.148^{NS}	1	3.479
	KS	20.000	3.250	1.871	0.715	0.338	0.403	0.413	0.156^{NS}	0	3.147
Shikoku	MR	19.000	2.250	1.676	0.521	0.283	0.312	0.320	0.100^{NS}	0	2.236
	AS	20.000	2.875	1.949	0.707	0.413	0.404	0.415	-0.042*	0	2.835
	SY	36.000	3.125	1.925	0.698	0.340	0.398	0.404	0.101^{NS}	0	2.867
	TM	20.000	2.875	1.841	0.686	0.375	0.409	0.420	0.066^{NS}	0	2.799
	SG	23.000	2.625	1.901	0.616	0.375	0.373	0.381	-0.021^{NS}	0	2.486
	SD	20.000	2.875	1.878	0.649	0.300	0.365	0.374	0.143^{NS}	0	2.814

	TS	20.000	3.375	1.751	0.642	0.344	0.346	0.355	0.001^{NS}	1	3.220
Chugoku	TK	17.750	2.500	1.807	0.569	0.295	0.349	0.359	0.125^{NS}	0	2.474
	NG	19.625	2.625	1.987	0.728	0.471	0.441	0.453	-0.017^{NS}	0	2.622
Kyushu	US	19.000	2.875	1.830	0.604	0.322	0.329	0.337	-0.016^{NS}	0	2.820
	RY	18.500	3.875	1.828	0.734	0.293	0.380	0.391	0.210^{NS}	1	3.771
	NM	20.000	3.250	1.933	0.702	0.306	0.382	0.392	0.148^{NS}	0	3.198
	ST	19.000	2.875	1.871	0.627	0.362	0.351	0.360	-0.041^{NS}	1	2.819
Okinawa	IZ	23.875	2.625	1.845	0.627	0.279	0.355	0.362	0.126^{NS}	2	2.609
Mean		22.138	2.993	1.899	0.676	0.346	0.383	0.392	0.074*		3.466

Note: Code: code of each population (see Table 3.1); *N*: No. of detected individuals in each population; N_a: No. of different alleles; N_e: No. of effective alleles = 1 / (Sum π^2); *I*: Shannon's information index = -1* Sum (π * Ln (pi)); H_O: observed heterozygosity; H_E: expected heterozygosity; UH_E: unbiased expected heterozygosity = (2*N* / (2*N*-1)) * H_E; *F*: fixation index = (H_E - H_O) / H_E = 1 - (H_O / H_E); *Ap*: number of private alleles.

Table 4.4 Summary of private alleles for studied loci

Locus	Private allele	Population	Frequency
bcqm76	174	OG	0.038
bcqm42	132	RY	0.053
QpZAG36	213	ST	0.132
QpZAG36	223	TB	0.056
bcqm07	201	IZ	0.130
bcqm07	207	HZ	0.017
bcqm07	211	IZ	0.174
QM69-2M1	229	TS	0.025

Table 4.5 Summary of AMOVA for *Q. phillyraeoides* in Japan

Source of variation	d.f.	Sum of squares	Variance components[a]	Percentage of variation (%)
Among groups	2	29.57	0.023*	1.36
Among populations	16	132.78	0.152***	8.81
Within groups				
Within populations	827	1283.52	1.552***	89.83
Total	845	1445.85	1.728	100.00

[a]*: *P*-value by permutation test more than 0.01 and less than 0.05. ***: *P*-value by permutation test less than 0.001.

Table 4.6 Pairwise F_{ST} and significance level of *Q. phillyraeoides*

	IW	OG	HZ	TB	KS	MR	AS	SY	TM	SG	SD	TS	TK	NG	US	RY	NM	ST	IZ
IW		***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
OG	0.053		NS	NS	NS	**	NS	NS	NS	**	**	**	**	NS	*	NS	NS	**	**
HZ	0.069	0.025		NS	**	**	*	**	**	**	**	**	**	**	**	*	**	**	**
TB	0.076	0.020	0.025		NS	**	NS	*	*	**	**	*	**	NS	**	NS	NS	**	**
KS	0.067	0.024	0.030	0.021		NS	NS	NS	NS	**	**	NS	**	*	NS	NS	NS	**	**
MR	0.089	0.070	0.073	0.066	0.035		**	NS	NS	**	**	**	**	**	*	**	*	**	**
AS	0.051	0.014	0.024	0.024	0.018	0.052		NS	NS	*	NS	NS	**	NS	NS	NS	NS	**	**
SY	0.061	0.020	0.032	0.028	0.020	0.027	0.016		NS	**	**	*	**	NS	NS	NS	NS	**	**
TM	0.076	0.017	0.034	0.029	0.015	0.040	0.016	0.013		**	**	**	**	NS	NS	NS	NS	**	**
SG	0.045	0.047	0.058	0.055	0.045	0.062	0.029	0.046	0.059		NS	NS	NS	**	**	**	NS	**	**
SD	0.032	0.045	0.063	0.062	0.050	0.077	0.027	0.040	0.058	0.019		NS	NS	**	**	**	NS	**	**
TS	0.050	0.045	0.059	0.056	0.037	0.039	0.025	0.027	0.041	0.026	0.019		**	NS	NS	*	NS	**	**
TK	0.062	0.065	0.071	0.065	0.045	0.038	0.048	0.048	0.066	0.017	0.045	0.039		**	**	**	*	**	**
NG	0.078	0.021	0.045	0.024	0.031	0.051	0.024	0.015	0.017	0.056	0.050	0.038	0.070		**	*	NS	**	**
US	0.071	0.048	0.043	0.066	0.029	0.037	0.028	0.030	0.037	0.053	0.055	0.031	0.048	0.062		NS	NS	**	**

RY	0.079	0.028	0.024	0.022	0.010	0.055	0.023	0.026	0.025	0.063	0.061	0.054	0.068	0.041	0.043		NS	**	**
NM	0.042	0.023	0.036	0.036	0.019	0.041	0.011	0.014	0.027	0.028	0.025	0.016	0.037	0.032	0.020	0.030		**	**
ST	0.106	0.065	0.049	0.056	0.054	0.060	0.054	0.042	0.061	0.075	0.095	0.073	0.074	0.068	0.045	0.054	0.053		**
IZ	0.166	0.123	0.114	0.093	0.133	0.170	0.132	0.114	0.143	0.143	0.172	0.184	0.148	0.125	0.171	0.119	0.140	0.120	

Note: *P*-values were obtained after 171,000 permutations. Indicative adjusted nominal level (5%) for multiple comparisons is 0.000292; F_{ST} value: below diagonal; significant of F_{ST}: above diagonal; NS: not significant; *: $P<0.05$; **: $P<0.01$.

Table 4.7 Allelic richness per locus and over all populations

District	Code	bcqm07	bcqm42	bcqm76	bcqm94	QM58-TGT	QM69-2M1	QpZAG36	MSQ13	Mean
Kanto	IW	1.000	1.999	3.992	2.783	3.903	3.531	2.000	1.531	2.592
Chubu	OG	1.000	3.643	5.862	3.000	2.999	4.296	2.000	2.000	3.100
	HZ	1.586	4.758	4.586	2.561	3.419	3.586	2.997	1.999	3.187
Kinki	TB	1.000	4.997	4.000	3.000	3.944	4.889	3.998	2.000	3.479
	KS	1.850	3.831	3.981	2.831	3.000	4.698	2.981	2.000	3.147
Shikoku	MR	1.000	2.991	4.000	1.895	3.000	2.000	2.000	1.000	2.236
	AS	1.000	3.000	4.000	2.979	3.850	2.850	3.000	2.000	2.835
	SY	1.000	3.197	4.436	2.960	3.000	3.892	2.725	1.725	2.867
	TM	1.000	2.850	3.000	3.846	3.000	3.848	2.850	2.000	2.799
	SG	1.000	2.739	3.739	1.739	3.000	2.739	2.997	1.936	2.486
	SD	1.000	2.850	3.981	2.848	3.850	3.981	3.000	1.000	2.814
	TS	1.850	3.850	4.848	1.981	3.831	4.550	3.850	1.000	3.220
Chugoku	TK	1.000	3.789	3.000	1.000	4.000	2.000	3.000	2.000	2.474
	NG	1.000	2.981	3.000	3.000	2.998	3.000	3.000	2.000	2.622
Kyushu	US	1.000	2.894	4.000	1.895	3.991	2.895	3.895	1.991	2.820

	RY	1.895	4.781	3.991	3.781	4.889	5.887	2.944	2.000	3.771
	NM	1.000	3.831	4.981	2.979	4.829	3.000	2.981	1.981	3.198
	ST	1.895	3.991	3.991	1.000	3.895	3.789	2.991	1.000	2.819
Okinawa	IZ	4.000	2.000	3.000	1.919	4.953	2.000	2.000	1.000	2.609
Mean		2.087	3.950	4.623	3.159	4.338	4.331	3.307	1.931	3.466

Table 4.8 Detection of null alleles by using MICRO-CHECKER

Locus	Repeat of motif	Results
bcqm07	dinucleotide	No evidence for scoring error due to stuttering. No evidence for large allele dropout. No evidence for null alleles.
bcqm42	dinucleotide	No evidence for scoring error due to stuttering. No evidence for large allele dropout. No evidence for null alleles.
bcqm74	dinucleotide	No evidence for scoring error due to stuttering. No evidence for large allele dropout. No evidence for null alleles.
bcqm76	dinucleotide	No evidence for scoring error due to stuttering. No evidence for large allele dropout. No evidence for null alleles.
bcqm94	dinucleotide	No evidence for scoring error due to stuttering. No evidence for large allele dropout. No evidence for null alleles.
QM58TGT	trincleotide	No evidence for scoring error due to stuttering. No evidence for large allele dropout. No evidence for null alleles.
QM69-2M1	trincleotide	No evidence for scoring error due to stuttering. No evidence for large allele dropout. No evidence for null alleles.
QpZAG36	dinucleotide	No evidence for scoring error due to stuttering. No evidence for large allele dropout. No evidence for null alleles.
MSQ13	dinucleotide	No evidence for scoring error due to stuttering. No evidence for large allele dropout. No evidence for null alleles.

Table 4.9 Results of Bayesian clustering analysis for *Q. phillyraeoides* in Japan

Population	ΔK	Best K
The Pacific coast	28.373271	3
The Set Inland Sea	78.241316	2
Kyushu	81.055206	2
Kyushu and Okinawa	267.793464	3
East China Sea	309.444687	3
Total	31.726480	4

Table 4.10 Provabilities of each cluster in each population and results of bottleneck detection of *Q. phillyraeoides*

			Cluster						Recent bottleneck detection	
District	Group	Code	I	II	III	IV	Dominant	cpDNA haplotype	Probability (one tail for H excess)	Model-shift
Kanto	1	IW	0.875	0	0.031	0.094	1	A1	0.1875	normal L-shaped distribution
Chubu	1	OG	0.231	0.346	0.038	0.385	2 and 4	A1	0.14844	normal L-shaped distribution
	1	HZ	0.241	0.069	0.172	0.517	4	A1	0.84375	normal L-shaped distribution
Kinki	1	TB	0.056	0.389	0.167	0.389	2 and 4	C	0.46875	normal L-shaped distribution
	1	KS	0.1	0.5	0.2	0.2	2	4/5A1, 1/5C	0.72656	normal L-shaped distribution
	1	MR	0.053	0.684	0.105	0.158	2	4/5A1, 1/5B	0.07813	shifted mode
	1	AS	0.15	0.3	0.15	0.4	2 and 4	A1	0.34375	shifted mode
Shikoku	2	SY	0.083	0.556	0.111	0.25	2	A1	0.34375	normal L-shaped distribution
	2	TM	0	0.6	0.05	0.35	2	A1	0.14844	normal L-shaped distribution
	2	SG	0.696	0.261	0	0.043	1	D2	0.28906	normal L-shaped distribution
	2	SD	0.45	0.35	0.1	0.1	1 and 2	4/5D1, 1/5D2	0.57813	normal L-shaped distribution
	2	TS	0.25	0.5	0.1	0.15	2	D2	0.98047	normal L-shaped distribution
Chugoku	2	TK	0.789	0.158	0.053	0	1	D2	0.28125	normal L-shaped distribution

	2	NG	0.05	0.65	0	0.3	2	4/5A1, 1/5A2	0.01172	shifted mode
	2	US	0.368	0.211	0	0.421	1 and 4	D2	0.76563	normal L-shaped distribution
Kyushu	3	RY	0.105	0.211	0.211	0.474	4	4/5A1, 1/5E	0.99414	normal L-shaped distribution
	3	NM	0.35	0.3	0.15	0.2	1 and 2	3/5A1, 2/5D2	0.8125	normal L-shaped distribution
	3	ST	0	0	0.737	0.263	3	2/5D1, 3/5D2	0.57813	normal L-shaped distribution
Okinawa	3	IZ	0	0	0.958	0.042	3	D2	0.03906	shifted mode

Note: For detecting cpDNA haplotype, we chose 5 individuals for each population. Haplotype frequencies of each haplotype were also showed. Probablility (one tail for H excess) using Wilcoxon's signed-rank test as determined by Bottleneck software. Shifted mode means certain population had experienced a recent bottleneck.

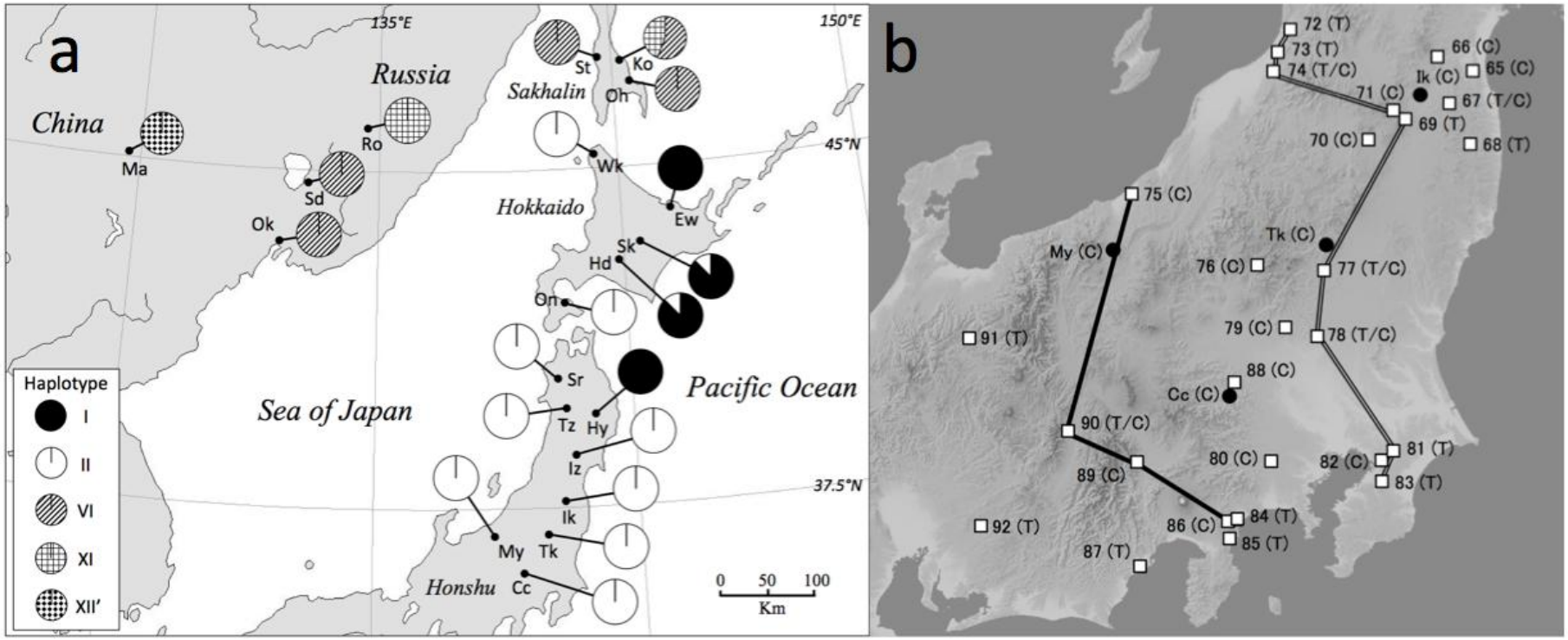

Fig. 2.1 (a) Geographic distribution of cpDNA haplotypes in *Quercus* species in the northeastern part of Japan and surrounding areas. The species in Japan and Sakhalin was *Q. mongolica* var. *crispula* and the species in Primorski Krai and Harbin was *Q. mongolica* var. *mongolica*. Abbreviation for each population follows the notation in Table 1. Haplotype frequencies are shown by pie chart.

(b) Geographic distribution of T/C-types in central Japan. The part below the latitude 38˚N was enlarged. Solid circles indicate the location of populations used in this study. Open squares with locality numbers show the collection sites of *Quercus* sect. *Prinus* by Kanno *et al.* (2004). Chloroplast types are indicated in the parentheses. The solid line connects the populations with chloroplast C-types in the southernmost boundary. A double line connects the populations with chloroplast T-types at the southernmost boundary.

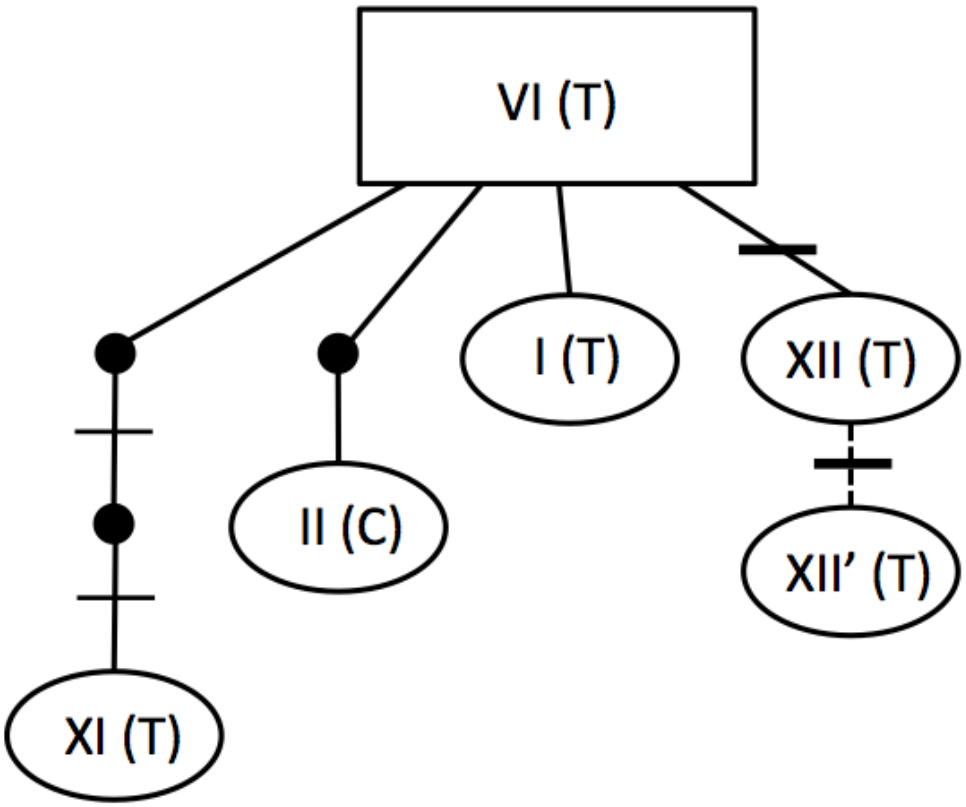

Fig. 2.2 Haplotype network constructed for the newly defined cpDNA haplotypes of *Quercus mongolica* var. *crispula* and *Q. mongolica* var. *mongolica* using TCS software (Clement *et al*., 2000). Haplotype VI in square is ancestral to the other haplotypes. Closed circles represent hypothetical haplotypes. Thick horizontal bars indicate position of large insertions. Thin horizontal bars indicate position of one nucleotide indels whose positions on the branch are arbitrary. The branch connecting haplotype XII and XII' was dotted because of no nucleotide substitution between them. Mononucleotide repeats variation was not considered.

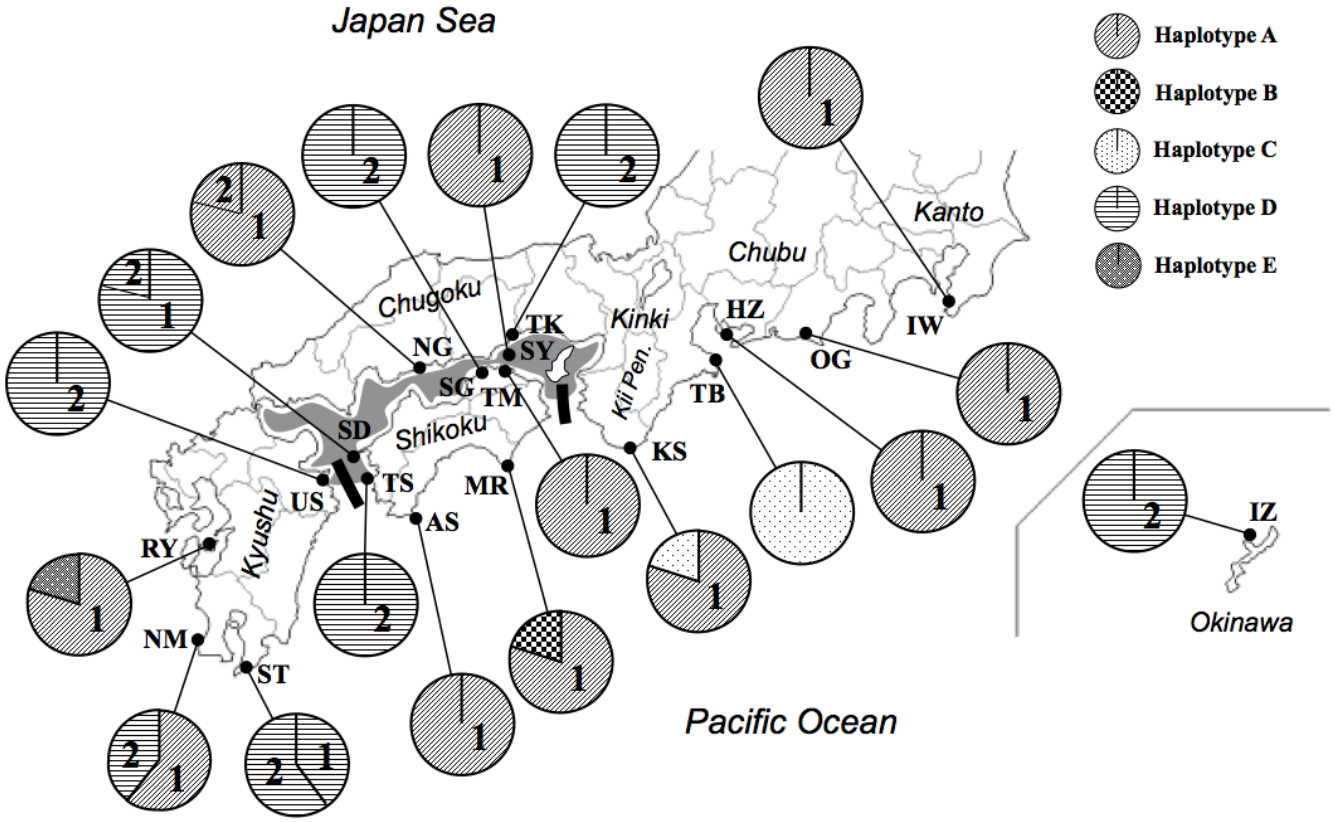

Fig. 3.1 Distribution of haplotypes. Numbers in the circles indicate subtypes. The shaded area is the Seto Inland Sea. The thick line between Kyushu and Shikoku shows Bungo Strait. Another thick line between Shikoku and Kii Peninsula shows Kii Strait.

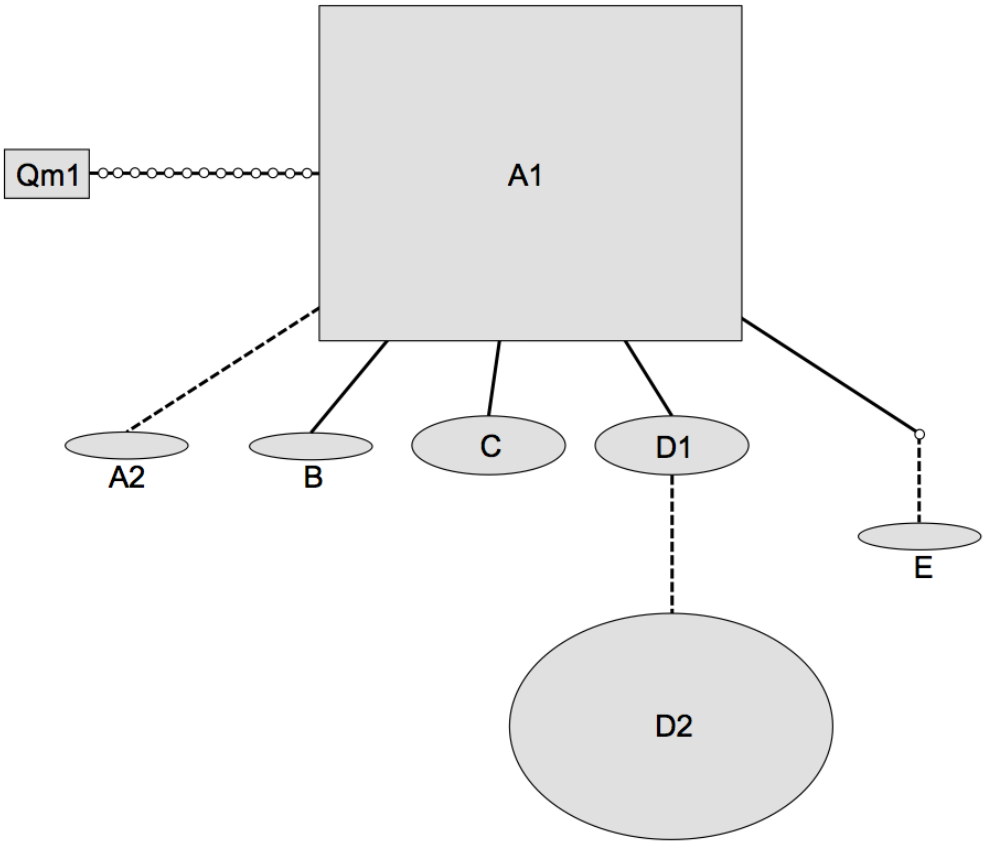

Fig. 3.2 Haplotype network constructed for the seven haplotypes of *Q. phillyraeoides* in Japan using TCS software (Clement *et al*., 2000). One haplotype sequence of *Quercus mongolica* var. *crispula* (*Qm*) was used as an outgroup. Type A1 (shown in *square*) is ancestral to the other haplotypes. Circled or squared areas are proportional to the frequencies of the haplotypes. Nodes between A1 and E, and A1 and *Qm* represent hypothetical haplotypes. Branches shown by broken lines indicate mononucleotide number changes or indels.

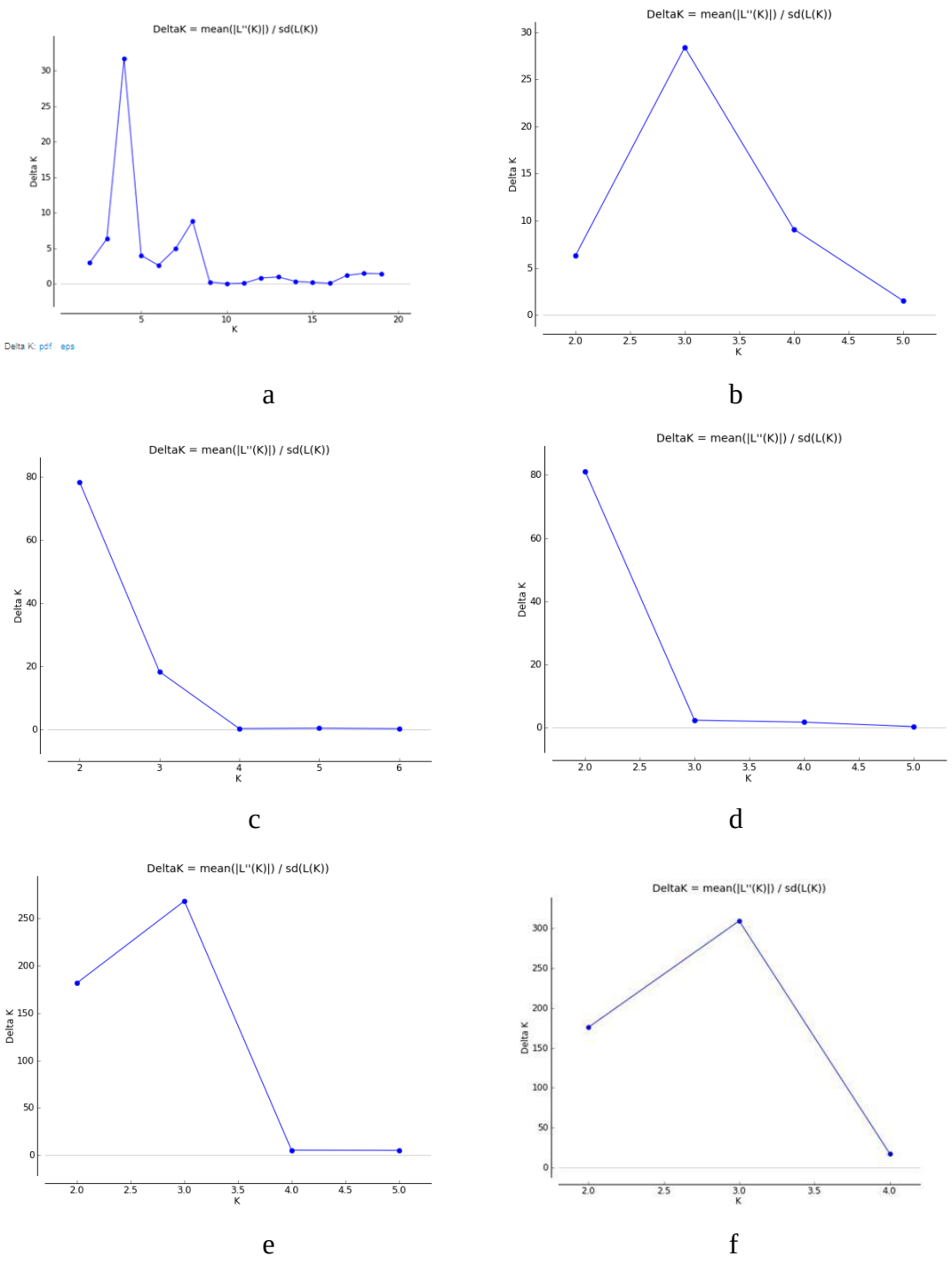

Fig. 4.1 Delta K value of investigated populations according to results of Structure Harvester software.

a: Delta K value of total investigated populations;

b: Delta K value of the Pacific coast populations;

c: Delta K value of the Seto Inland Sea populations;

d: Delta K value of Kyushu populations;

e: Delta K value of Kyushu and Okinawa populations.

f. Delta K value of East China Sea populations.

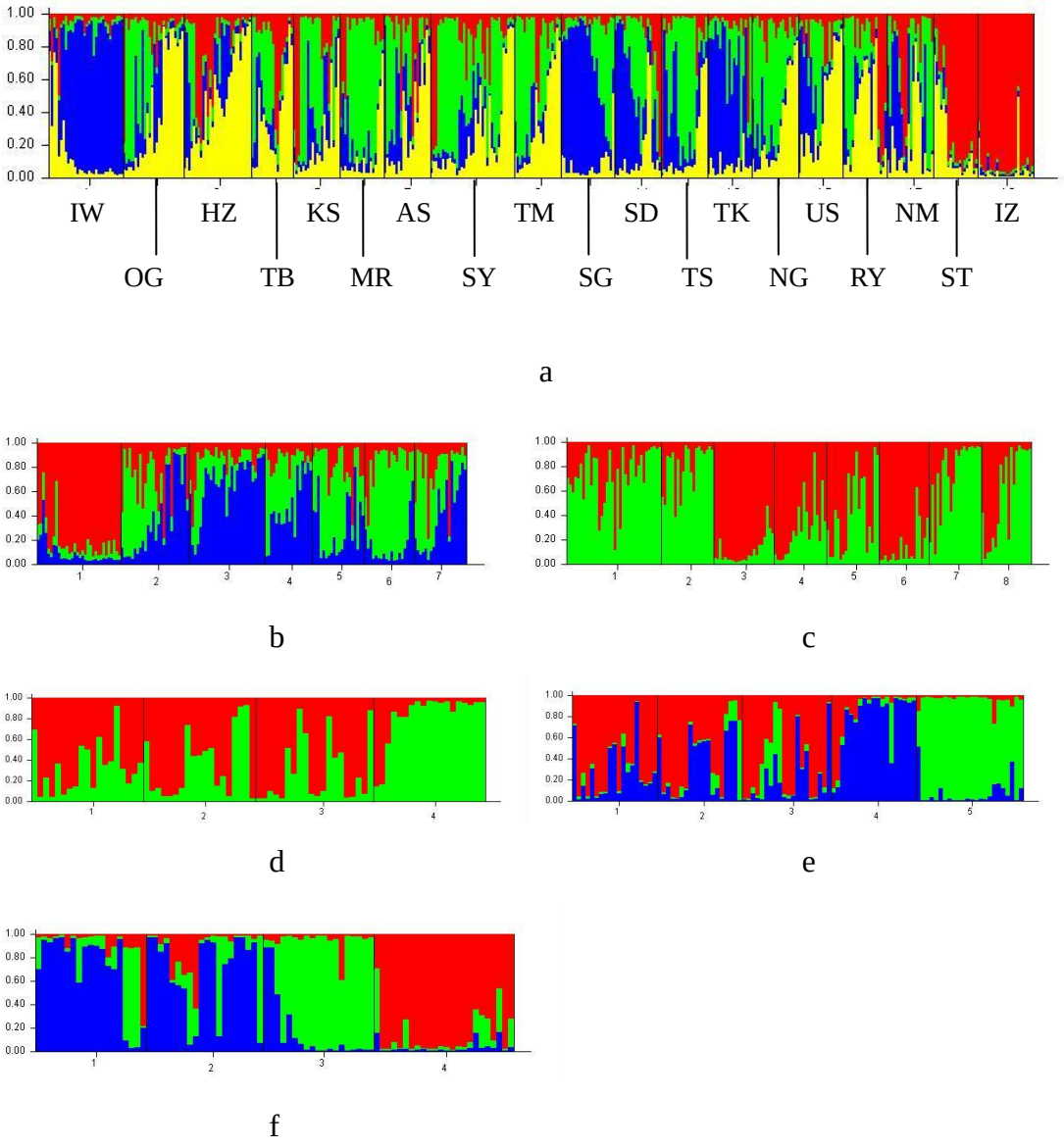

Fig. 4.2 Bar plot group by population code of best K.

a: Bar plot of total investigated populations.

b: Bar plot of the Pacific coast populations. Code from 1 to 7 are IW, OG, HZ, TB, KS, MR, and AS.

c: Bar plot of the Seto Inland Sea populations. Code from 1 to 8 are SY, TM, SG, SD, TS, TK, NG, and US.

d: Bar plot of Kyushu populations. Code from 1 to 4 are US, RY, NM, and ST.

e: Bar plot of Kyushu and Okinawa populations. Code from 1 to 5 are US, RY, NM, ST, and IZ.

f: Bar plot of East China Sea populations. Code from 1 to 4 are RY, NM, ST, and IZ.

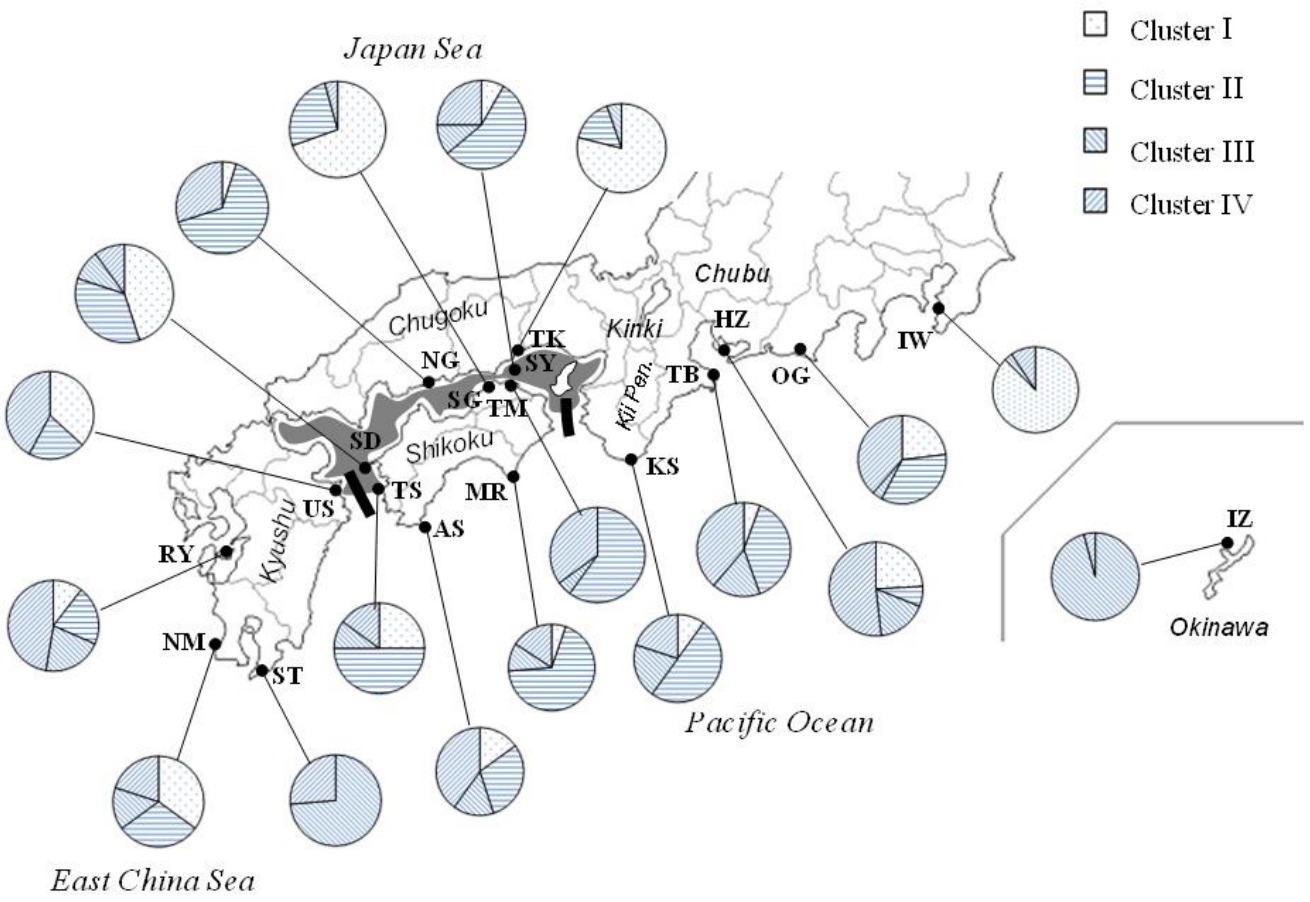

Fig. 4.3 Structure clusters distribution.

References

Abbott RJ, Smith LC, Milne RI, Crawford RMM, Wolff K, Balfor J (2000) Molecular analysis of plant migration and refugia in Arctic. *Science* 289: 1343-1346.

Aoki K, Matsumura T, Hattori T, Murakami N (2006) Chloroplast DNA phylogeography of *Photinia glabra* (Rosaceae) in Japan. *Am J Bot* 93: 1852-1858.

Aoki K, Suzuki T, Hsu T-W, Murakami N (2004) Phylogeography of the component species of broad-leaved evergreen forests in Japan, based on chloroplast DNA variation. *J Plant Res* 117: 77-94.

Berg, E.E. and Hamrick, J.L. (1993) Regional genetic variation in turkey oak, *Quercus laevis*. *Canadian Journal of Forest Research* 23: 1270-1274.

Berg, E.E. and Hamrick, J.L. (1994) Spatial and genetic structure of two sand-hills oaks: *Quercus laevis* and *Quercus margaretta* (Fagaceae). *American Journal of Botany* 81: 7-14.

Birky CW, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics* 121: 613-627.

Cheng Y-P, Hwang S-Y, Lin T-P (2005) Potential refugia in Taiwan revealed by the phylogeographical study of *Castanopsis carlesii* Hayata (Fagaceae). *Mol Ecol* 14:2075-2085.

Chiba Biodiversity Center (2009) The red data book of Chiba Prefecture. http://www.bdcchiba.jp/index.html

Chung, M.Y., Nason, J., Chung, M.G., Kim, K.J., Park, C.W., Sun, B.Y., and Pak, J.H. (2002) Landscape-level spatial genetic structure in *Quercus acutissima* (Fagaceae). *American Journal of Botany* 89: 1229-1236.

Clement, M., Posada, D., and Crandall, K. A. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657-1660.

Consortium of Japan Bincho-tan Meeting (2011) Bulletin (in Japanese).

Cornuet J.M. and Luikart G. (1997) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*. 144:2001-2014.

Corriveau JL, Coleman AW (1988) Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. *Am J Bot* 75: 1443-1458.

Cottrell, J.E., Munro, R.C., Tabbener, H.E., Milner, A.D., Forrest, G.I. And Lowe, A.J. (2003) Comparison of fine-scale genetic structure using nuclear microsatellites within two British oak woods differing in population history. Forest Ecology and Management 176: 287-303.

D. B. Goldstein, A. R. Linares, L. L. Cavalli-Sforza, and M. W. Feldman (1995) An Evaluation of Genetic Distances for Use With Microsatellite Loci. Genetics 139 (1): 463–471.

Daniel L. Hartl (2000) "*A Primer of Population Genetics*." 3rd edition. Sinauer Associates, Inc., Publishers, Sunderland, Massachusetts, U.S.A.

Dent A. Earl and Bridgett M. vonHoldt (2012) Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. Conservation Genet Resour 4: 359- 361.

Dow BD and Ashley MV (1998) High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. J. Hered. 89, 62–70.

Dow BD and Ashley MV (1996). Microsatellite analysis of seed dispersal and parentage of sampling in the bur oak, *Quercus macrocarpa*. *Mol Ecol* 5: 615-627.

Dow BD, Ashley MV, and Howe HF (1995). Characterization of highly variable (GA/CT)$_n$ microsatellites in the bur oak, *Quercus macrocarpa*. Theor Appl Genet 91: 137-141.

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

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

Dutech Cyril, Victoria L. Sork, Andrew J. Irwin, Peter E. Smouse, and Frank W. Davis. (2005) Gene flow and fine-scale genetic structure in a wild-pollinated tree species, *Quercus Lobata* (Fagaceaee). *American Journal of Botany*. 92 (2): 252-261.

Ennos RA (1994) Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* 80: 548-593.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinformatics* 1: 47-50.

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

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

Ferris C, King R, Vainola R, Hewitt GM (1998) Chloroplast DNA recognizes three refugial sources of European oaks and suggests independent eastern and western immigrations to Finland. *Heredity* 80: 584-593.

Forman, L. L. (1964) Trigonobalanus, a new genus of Fagaceae, with notes on the classification og the family. *Kew Bull.* 17: 381-396.

Fu Y-X, Li W-H (1993) Statistical tests of neutrality mutations. *Genetics* 133: 693-709.

Fujii N, Tomaru N, Okuyama K, Koike T, Mikami T, Ueda K (2002) Chloroplast DNA phylogeography of *Fagus crenata* (Fagaceae) in Japan. *Plant Syst Evol* 232: 21-33.

Fujii N, Ueda K, Watano Y, Shimizu T (1997) Intraspecific sequence variation of chloroplast DNA in *Pedicularis chamissonis* Steven (Scrophulariaceae) and geographic structuring of the Japanese “alpine” plants. *J Plant Res* 110: 195-207.

Goodman SJ (1997) Rst Calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and a determining their significance. *Molecular Ecology*. 6: 881-885.

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *J Hered* 86:485–486

Greg Gibson and Spencer V. Muse (2009) Book: A Primer of Genome Science (Third Edition). Sinauer Associates, Inc. Publishers. Sunderland, Massachuestts USA. Page 110.

Guttman, S.I. And Weigt, L.A. (1989) Electrophoretic evidence of relationships among *Quercus* (oaks) of eastern North America. *Canadian Journal of Botany* 67: 339-351.

Hara Y., Yokoyama A. and Kim J.H. (2000) Systematic relationship of unicellular red algae from phenetics to the phylogeny. Algae. 15: 7-11.

Harada K, Tamaki K, Kamiya K, Takechi Y (2003) Pollen morphology observed by scanning electron microscopy on Japanese Fagaceae species and molecular phylogeny. *Bulletin of the Ehime University Forest* 42: 1-19 (in Japanese).

Hattori T (1985) On the distribution limits of the lucidophyllous forest in the Japanese Archipelago. *Bot Mag Tokyo* 98: 317-333.

Hattori T (2002) Ch. 7: Forest conservation from the view point of biogeography in Lucidophyllus forests. In: Yahara T, Kawakubo N (eds) Biology of Conservation and Restoration. Bun-ichi Sogoshuppan, Tokyo, pp 203-222 (in Japanese).

Huang SF, Hwang SY, Wang JC, Lin TP (2004) Phylogeography of *Trochodendron aralioides* (Trochodendraceae) in Taiwan and its adjacent areas. *J Biogeography* 31: 1291-1239.

Huang SF, Hwang SY, Kin TP (2002) Spatial pattern of chloroplast DNA variation of

Cyclobalanopsis glauca in Taiwan and East Asia. *Mol Ecol* 11: 2349-2358.

Hudson RR, Slatkin M, Madison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics* 132: 583-589.

Hu XS, Ennos R (1997) On estimation of the ratio of pollen to seed flow among plant populations. *Heredity* 79: 541-552.

Isagi, Y. and Suhandono, S (1997) PCR primers amplifying microsatellite loci of *Quercus myrsinifolia* Blume and their conservation between oak species. Mol. Ecol. 6(9), 897-899.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. Trends Ecol Evol 11:424–429.

Jiri Dostalek, Tomas Frantik, Miroslava Lukasova (2011) Genetic differences within natural and planted stands of *Quercus petraea*. Central European Journal of Biology 6 (4): 597-605.

Johnk N, Siegismund HR (1997) Population structure and postglacial migration routes of *Quercus robur* and *Quercus petraea* in Denmark, based on chloroplast DNA analysis. *Scandinavian J. Forest Res*. 12: 130-137.

K. S. Kim, Y. Tanabe, C. K. Park, and J. H. Ha. (2001) Genetic Variability in East Asian Dogs Using Microsatellite Loci Analysis. The Journal of Heredity: 92 (5): 398-403.

Kamiya K, Harada K, Mahani MC, Latiff AM (2002) Genetic variation of *Trigonobalanus verticillata*, a primitive species of Fagaceae, in Malaysia revealed by chloroplast sequences and AFLP markers. *Genes Genet Syst* 77: 177-186.

Kanagawa Prefecture (2006) The red data book of Kanagawa Prefecture. http://www.e-tanzawa.jp/rdb06/

Kanno M., Yokoyama J., Suyama Y., Ohyama M., Itoh T. & Suzuki M. (2004) Geographic distribution of two haplotypes of chloroplast DNA in four oak species (*Quercus*) in Japan. *Journal of Plant Research* 117: 311-317.

King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Mol Ecol* 7: 1151-1161.

Kira T (1977) A climatological interpretation of Japanese vegetation zones. In: Miyawaki A, Tuxen T (eds) Vegetation science and environmental protection. Maruzen, Tokyo, pp 21-30.

Kizaki, K (1982) Establishment of Ryukyu Arc - special lecture. *Soils and Foundation* 30: 3-5, (in Japanese).

Li, W. H. (1997) "*Molecular Evolution*." Sinauer Associates, Inc., Publishers, Sunderland, Massachusetts, U.S.A.

Librado P and Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.

Maeda Y (1980) Sea and forest at the Jomon Period. Aoki Shobo, Tokyo, pp149-216 (in Japanese).

Matsuoka K, Miyoshi N (1998) Ch. III-4: Transition of Lucidophyllous forests after the last glacial maximum. In: Yasuda Y, Miyoshi N (eds.) Illustration: vegetation history of the Japanese Archipelago. Asakura Shoten, Tokyo, pp224-236 (in

Japanese).

McCauley DE (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plant. *Trends Ecol. Evol.* 10: 198-202.

Minato M, Ijiri S (1976) The Japanese Archipelago. Iwanami Shoten, Tokyo, pp 91-98 (in Japanese).

Mishima, K., Watanabe, A., Isoda, K., Ubukata, M. and Takata, K. (2006) Isolation and characterization of microsatellite loci from *Quercus mongolica* var. *crispula*. Mol. Ecol. Notes 6, 695-697.

Miyazaki Prefecture (2010) The red data book of Miyazaki Prefecture. http://www.pref.miyazaki.lg.jp/contents/org/kankyo/shizen/reddatabook/index.html

Miyoshi N (1998) Ch. II-7: Vegetation history of Chugoku and Shikoku districts. In: Yasuda Y, Miyoshi N (eds.) Illustration: vegetation history of the Japanese Archipelago. Asakura Shoten, Tokyo, pp 224-236 (in Japanese).

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182.

Nagasaki Prefecuter (2011) The red data book of Nagasaki Prefecture. www.pref.nagasaki.jp/sizen/2reddata/index.html

Nakagawa S, Ito M (2009) Development and characterization of microsatellite loci in *Ixeridium dentatum* (Asteraceae, Lactuceae). J Plant Res 122:581–584.

Nakao S (1966) The origin of cultivated plants and agriculture. Iwanami Shoten, Tokyo, pp 59-76 (in Japanese).

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70: 3321-3323.

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

Newton AC, Allnutt TR, Gollies ACM, Lowe AJ, Ennos RA (1999) Molecular phylogeography, intraspecific variation and the conservation of tree species. *Trends Ecol Evol* 14: 140-145.

Ohsawa T, Tsuda Y., Saito Y. & Ide Y. (2011) The genetic structure of *Quercus crispula* in northeastern Japan as revealed by nuclear simple sequence repeat loci. *Journal of Plant Research* 124: 645-654.

Okaura T, Harada K (2002) Phylogeographical structure revealed by chloroplast DNA variation in Japanese Beech (*Fagus crenata* Blume). *Heredity* 88: 322-329.

Okaura T, Quang ND, Ubukata M, Harada K (2007) Phylogeographic structure and late Quaternary population history of the Japanese oak *Quercus mongolica var. crispula*. *Genes Genet Syst* 82: 465-477.

Okinawa Prefecture (2006) The red data book of Okinawa Prefecture. http://www3.pref.okinawa.jp/site/view/contview.jsp?cateid=70&id=9962&page=1

Parvin Salehi Shanjani, G. G. Vendramin, and M. Calagari (2010) Genetic diversity and differentiation of *Fagus orientalis* Lipsky in Hyrcanian forests revealed by nuclear and chloroplast microsatellite markers. *Conserv Genet* 11: 2321–2331.

Payne RW, Murray DA, Harding SA, Baird DB, Soutar DM. (2008) GenStat for

Windows Introduction. 11. Hemel Hempstead, UK: VSN International.

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

Petit RJ, Kremer A, Wagner DB (1993) Finite island model for organelle and nuclear genes in plants. *Heredity* 71: 630-641.

Piry S, Luikart G and Cornuet J-M (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered.* 90: 502-503.

Rosenberg Noah A., Terry Burke, Kari Elo, Marcus W. Feldman, Paul J. Freidlin, Martien A. M. Groenen, Jossi Hillel, Asko Maki-Tanila, Michele Tixier-Boichard, Alain Vignal, Klaus Wimmers and Steffen Weigend (2001) Empirical Evaluation of Genetic Clustering Methods Using Multilocus Genotypes From 20 Chicken Breeds. *Genetics* 159:699-713.

Rousset F. 1996 Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics.* 142: 1357-1362.

Rritchard Jonathan K, Xiaoquan Wen, and Daniel Falush (2009) Documentation for structure software: version 2.3.

Schnabel, A. And Hamrick, J.L. (1990) Comparative analysis of population genetic structure in *Quercus macrocarpa* Michx. and *Q. Gambeli* Numm. *Systematic Botany* 15: 240-251.

Steinkellner, H., Fluch, S., Turetschek, E., Lexel, C., Streiff, R., Kremer, A., Burgand, K. and Gloess, J. (1997) Identification and characterization of (GA/CT)n microsatellite loci from *Quercus petraea*. *Plant Mol. Biol.* 33(6), 1093-1096.

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

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

Tajima F (1989) Statistical method for testing neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.

Tautz D (1989) Hypervariablity of simple sequences as a general source of polymorphic DNA markers. *Nucleic Acids Res* 17: 6463-6471.

Terachi T (1993) Structural alterations of chloroplast genome and their significance to higher plant evolution. *Bull Inst Natl Land Util Devel, Kyoto Sangyo University* 14: 138-148.

Tsukada M (1974) Paleoecology II. Kyoritsu Shuppan, Tokyo (in Japanese).

Tsukada M. (1982) Cryptomeria japonica: glacial refugia and late-glacial and postglacial migration. *Ecology* 63: 1091–1105.

Uchiyama T. (1998) Vegetation history of the Kanto district. In: Yasuda Y. & Miyoshi N. (eds.) *The Illustrated Vegetation History of the Japanese Archipelago.* (in Japanese). Asakura-Shoten, Tokyo, pp 73-88.

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

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theor Pop Biol* 7: 256-276.

Weber JL and May PE (1989) Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44: 388-396.

Wu SH, Hwang CY, Lin TP, Chung JD, Cheng YP, Hwang SY (2006) Contrasting phylogeographical patterns of two closely related species, *Machilus thunbergii* and *Machilus kasanoi* (Lauraceae), in Taiwan. *J Biogeography* 33: 936-947.

Xie CP, Yan F, Fang YM. (2011) Geographical distribution of *Quercus phillyraeoides* A. Gray. *Tropical Geography* 31: 8-13, (in Chinese).

Xuejiang Zhang, Helena Korpelainen and Chunyang Li (2006) Microsatellite Variation of *Quercus aquifolioides* Populations at Varying Altitudes in the Wolong Natural Reserve of China. *Silva Fennica* 40 (3): 407-415.