Theor Appl Genet (1999) 99: 11–15

K. Tanaka · Y. Tsumura · T. Nakamura

Development and polymorphism of microsatellite markers for *Fagus crenata* and the closely related species, *F. japonica*

Received: 5 December 1998 / Accepted: 28 December 1998

Abstract We have developed microsatellite markers (SSRs) applicable to *Fagus crenata* using the RAHM method and investigated their polymorphisms. We also applied the SSRs in an analysis of a closely related species, *F. japonica*. Here we describe the isolation and characterization of nine polymorphic microsatellite markers, of which eight are applicable to both species. Among 30 individuals of each of *F. crenata* and *F. japonica* we detected a total of 79 and 77 alleles, respectively, with an average of 9.9 and 8.6 alleles per locus. The mean expected heterozygosity (*He*) was 0.615 (range: 0.216–0.925) in *F. crenata* and 0.660 in *F. japonica* (range: 0.259–0.827). The *He* values were considerably higher than those previously found for isozymes. Paternity exclusion probabilities for multiple loci, calculated over all loci, were extremely high (0.999 and 0.998 in *F. crenata* and *F. japonica*, respectively): sufficiently high to study pollen flow in both species.

Key words *Fagus crenata* · *Fagus japonica* · Microsatellite · RAPD · RAHM · SSR

Introduction

The effectiveness of microsatellite markers (simple-sequence repeats, SSRs) has been utilized in recent studies to quantify gene flows among plant populations and determine their genetic structure. While SSR markers allow us to estimate gene flows mediated by both seed dispersal and pollen flow, they have been particularly useful for analyzing the latter, which has proven difficult to study by field analysis of tree populations (Dow Ashley 1996; Chase et al. 1996).

Fagus species, beech, are dominant in many deciduous forests of the temperate zone of the northern hemisphere, and their ecology and population dynamics have been intensively studied in many such areas (Nakashizuka 1987; Yamamoto 1989). Their seed dispersal has been roughly estimated by seed-trap methods and direct observation (Maeda and Miyagawa 1971), however, the pollen flow of *F. crenata* has not been reported until now. Development of the microsatellite markers reported here (in combination with the large demographic data sets pertaining to these species) will provide us with the opportunity to study in greater detail their genetic structure, the effects of pollen and seed flow, and their ecological adaptive processes.

Microsatellites often have to be developed individually for each species, though some common sites have been found for related species (Echt et al. 1996). Therefore, we have developed microsatellite markers and investigated their polymorphism in *Fagus crenata* and, subsequently, their applicability to the closely related species, *F. japonica*.

Communicated by P. M. A. Tigerstedt

K. Tanaka[1] (✉) · T. Nakamura
Institute of Agriculture and Forestry, University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan

Y. Tsumura
Bio-resources Technology Division, Forestry and Forest Products Research Institute, Kukizaki, Ibaraki 305-8687, Japan

Present address:
[1]Center for Ecological Research, Kyoto University, Kamitanakami Hiranocho 509-3, Otsu 520-2113, Japan
E-mail: kenta@ecology.kyoto-u.ac.jp

Materials and methods

Plant material

Total DNA for developing microsatellite markers was extracted from leaves of *Fagus crenata* individuals using the modified CTAB method (Murray and Thompson 1980). We also extracted total DNA using the procedure of Tsumura et al. (1995) from each of 30 *F. crenata* and *F. japonica* individuals at the Mt. Takahara research plot in Tochigi Prefecture, which was established for a long-term ecological survey (Ohkubo 1996). These DNAs were used to survey

the polymorphism of SSRs in both species. This was anticipated to be a valid approach, since *F. japonica* is closely related to *F. crenata* with respect to its chloroplast DNA sequences (T. Kawahara, personal communication).

Screening of microsatellites

We employed the random amplified hybridization microsatellites (RAHM) method developed by Cifarelli et al. (1995) to identify the SSR region in the genome of *F. crenata*. This method is convenient for detecting SSR regions because it does not require a genomic library, and we merely selected the random amplified polymorphic DNA (RAPD) fragments containing the SSR region. RAPD fragments were amplified in a PTC-100 thermocycler (MJ Research) in a reaction mixture (25 ml) containing PCR buffer (10 m*M* TRIS-HCl, pH 8.3, 50 m*M* KCl, 100 m*M* of each dNTP, 0.02% Triton X-100, and 0.01% gelatin), 2 m*M* $MgCl_2$, 2 U of *Taq* DNA polymerase (GibcoBRL), 500 n*M* random primer (Operon), and 50 ng template DNA. The amplification cycle consisted of an initial 5-min denaturation at 94°C, followed by 40 cycles of a 1-min denaturation at 94°C, 1-min annealing at 36°C, and 2-min extension at 72°C, with a final 5-min extension at 72°C. The PCR products were separated on 2% agarose gels and transferred to nylon membranes. Southern hybridization was carried out using a DIG detection kit (Boehringer Mannheim), whereby $(CT)_{20}$ oligonucleotides were used as probes to detect CT/GA-repeat SSRs. The hybridization procedure followed was that of Tsumura and Suyama (1998) except for the hybridization and washing temperature (here, 40°C).

Cloning and sequencing of positive fragments and primer design

The fragments found to be positive in the above screening were purified by band-stab polymerase chain reaction (PCR) (Bjourson and Cooper 1992) and then cloned into pCR vectors (Invitrogen). The cloned vectors were transformed into Hot-1 *E. coli* (Invitrogen), and white colonies were selected. The plasmid DNA was extracted using the Alkarin-SDS method, and the insert DNA fragment was sequenced from both ends in an ABI 377 DNA sequencer using a Dye Terminator Cycle Sequencing method. The following universal primers were used for sequencing, 5′CCCAGTCACGACGTTGT3′ and 5′GGAAACAGCTATGACCATG3′.

PCR primers for microsatellites were designed based on the sequence data using the computer program OLIGO ver 4.0 (National Bioscience).

PCR amplification of SSRs and their polymorphism

PCR amplification of SSRs was carried out using the designed SSR primers. The reaction mixtures (25 ml), consisting of PCR buffer, 1.5 m*M* $MgCl_2$, 200 n*M* forward fluorescently labeled primer and reverse primer, 1 U DNA polymerase (*Taq* polymerase, GibcoBRL or Ampli-Taq Gold, Perkin Elmer), and 10 ng template DNA, were amplified in a Perkin Elmer 9600 thermal cycler. Amplification cycles consisted of an initial 5 min at 94°C (or 10 min at 95°C for Ampli-Taq Gold), followed by 30 cycles (or 40 cycles) of 15 s at 94°C and 30 s at an optimized temperature, with a final 5 min at 72°C. The PCR products were checked electrophoretically by 2% agarose gel fractionation to identify whether or not they were controlled by a single locus. The polymorphism of the SSRs was then investigated on a Genotyper ABI 310 system using the GeneScan program (Perkin Elmer).

Statistical analysis

The number of alleles per locus, the expected and observed heterozygosities (*He* and *Ho*, respectively), where *He* was an unbiased estimate (Nei 1978; Nei and Roychoudhury 1974), the fixation index (F_{IS}, Wright 1965), and paternity exclusion probability for single loci (*Pe*) were calculated for each locus and averaged over all loci (Weir 1996). Paternity exclusion probabilities for multiple loci (*Q*) were also calculated using all the loci and several selected loci. Deviations from Hardy-Weinberg expectations were tested using chi-square tests (Li and Horvitz 1953).

Results and discussion

Isolation and characterization of microsatellites

RAPD-amplified fragments were screened with a $(CT)_{20}$ oligonucleotide repeat. We used a total of 360

Table 1 Primer sequences of microsatellites loci in *F. crenata* and their PCR conditions

Locus	Core sequences	PCR primer(5′ → 3′)	Annealing temperature (°C)	Fragment size (bp) (sequenced allele)
mfc2	$(CT)_{27}$	ACAACTAATCCCACCAGTCC GGCCTAAAGCGACAACTCTA	62[a]	188
mfc3	$(ATT)_9$	TAATGGGATGGAGGGTGGAT GGAAGACAAACAATGGAAAG	55[b]	214
mfc4	$(CT)_6$	ATAAAGCAACTGAATAAAGA GTCAAGAGAGGGAAGAGAGT	52[b]	193
mfc5	$(AG)_{10}$	ACTGGGACAAAAAAACAAAA GAAGGACCAAGGCACATAAA	58[b]	276
mfc7	$(GA)_9$	AAAATACACTGCCCCCAAAA CAGGTTTTGGTTTCTTACAC	60[a]	112
mfc9–2	$(GA)_{19}$	TTCCCTCCTCTTCTCTAAAT TTATACTTCCTCTCTCATCCC	55[a]	175
mfc11	$(AG)_{10}$	ACAGATAAAAACAGAAGCCA TTTGGTTTTGTTGAGTTTAG	58[b]	324
mfc12	$(AGG)(AG)_6(AGG)(AG)_7(AGG)$	ACACCTCACAATCCACGAAA CCCAATAACTAAGAATACCA	56[b]	299
mfc13	$(CT)_4(GT)_2(CA)_2$	GAGAGCAGGATGGGATGAAA ACAGCACCTCTCCTTCTCTT	55[b]	317

[a] Ampli-Taq Gold (Perkin Elmer)
[b] *Taq* polymerase (GibcoBRL)

random primers for SSR screening. Sixty positive fragments were obtained, amplified by 38 primers. Thus, 10.6% of the primers gave positive fragments. Two or three positive fragments were also detected by 19 primers. Finally, sequencing analysis confirmed that 15 fragments contained repeated sequences with more than four repetitions. A microsatellite consisting of $(AAT)_9$ was obtained. Primer pairs were designed for these 15 SSRs. Eight primer pairs were used successfully to amplify single-locus products in *F. crenata*, all of which showed polymorphisms (Tables 1 and 2).

The screening efficiency of the RAHM technique (Cifarelli et al. 1995) is very high compared to the colony hybridization method, because 1 random primer can amplify many fragments, which can be screened simultaneously. We detected 60 positive fragments using 38 random primers, so 10.6% of the primers detected positive fragments.

Some plant species contain high contents of secondary metabolites in their cells. If we employ the colony hybridization method, these components have to be excluded from the genomic library during the DNA extraction. The RAHM procedure, being based on the PCR, is very convenient because we do not need to prepare high-quality genomic DNA to make a genomic DNA library (Cifarelli et al. 1995).

Applications of the *F. crenata* SSRs to *F. japonica*

We also applied these SSR primer pairs to *F. japonica*. In addition to eight single-locus, polymorphic SSR primer pairs developed for *F. crenata*, *mfc7* (which was a multiple-locus product in *F. crenata*) was found to be a single-locus product in *F. japonica*. Consequently, the nine primer pairs developed in the *F. crenata* study can be used to analyze *F. japonica* (Table 2).

Microsatellite primers developed in *F. crenata* could be used in analysis of *F. japonica*, indicating that they may be applicable to the study of other *Fagus* species, especially since species within the genus *Fagus* are very closely related (T. Kawahara, personal communication). Ujino et al. (1998) showed that SSR loci developed for *Shorea curtisii* are conserved among closely related species and genera and that there is high potential for the application of SSR markers among closed related taxa, as also found in our study. Conservation of SSRs for closely related plant taxa was also reported by Dayanandan et al. (1997), Kajis et al. (1995) and Wu and Tanksley (1993).

The *mfc7* sequence, which was a multiple-locus product in *F. crenata*, was a single-locus product in *F. japonica*. One possible explanation is that *F. crenata* may have several copies of this gene in its genome, and it may have become a single copy during speciation.

Table 2 Genetic parameters of microsatellite loci and paternity exclution probabilities in *F. crenata* and *F. japonica*

Locus	Fragment size (bp)	Number of alleles	*He*[a]	*Ho*[b]	F_{IS}[c]	*Pe*[d]
F. crenata						
mfc2	134–219	21	0.925	0.867	0.063	0.850
mfc3	201–211	6	0.521	0.500	0.041	0.316
mfc4	190–198	5	0.216	0.067	0.690**	0.117
mfc5	256–320	16	0.901	0.833	0.075	0.802
mfc9–2	164–186	4	0.494	0.467	0.056	0.284
mfc11	313–337	10	0.830	0.633	0.237	0.665
mfc12	280–327	14	0.799	0.567	0.291	0.627
mfc13	315–337	3	0.238	0.267	− 0.121	0.118
Mean		9.88	0.615	0.525	0.092	0.472
Q[e]						0.999
F. japonica						
mfc2	129–177	13	0.827	0.933	− 0.128	0.675
mfc3	202–253	6	0.580	0.300	0.482*	0.350
mfc4	191–195	3	0.259	0.333	− 0.285	0.122
mfc5	268–312	18	0.887	0.867	0.023	0.781
mfc7	103–135	11	0.817	0.967	− 0.183	0.645
mfc9–2	186–237	9	0.694	0.400	0.424	0.506
mfc11	314–349	4	0.611	0.800	− 0.310	0.344
mfc12	228–284	10	0.764	0.600	0.215	0.560
mfc13	313–319	3	0.499	0.233	0.532**	0.204
Mean		8.56	0.660	0.604	0.086	0.465
Q						0.998

* $P < 0.01$, ** $P < 0.001$
[a] *He*, Expected heterozygosity
[b] *Ho*, Observed heterozygosity
[c] F_{IS}, Fixation index
[d] *Pe*, Paternity exclusion probabilities for single loci
[e] *Q*, Multiple loci

The regeneration systems of the two *Fagus* species are different; *F. crenata* is regenerated predominantly by seedlings, but *F. japonica* has some clonal structure (Ohkubo 1992; Kitamura et al. 1992). It is likely that the SSR markers will be useful for comparing the genetic diversity and structure of the two species.

Level of the microsatellite polymorphisms

We have investigated the polymorphism of SSR loci in two *Fagus* species, *F. crenata* and *F. japonica*. Among 30 individuals of each of *F. crenata* and *F. japonica*, we detected a total of 79 and 77 alleles, respectively. The number of alleles per locus, the observed (*Ho*) and expected (*He*) heterozygosities, the fixation indices (F_{IS}), the paternity exclusion probabilities for both single loci (*Pe*) and multiple loci (*Q*) are shown in Table 2. When we estimated these parameters, a few data, which showed more than 3 fragments, were excluded.

The average number of alleles per locus was 9.9 (range 3–21) and 8.6 (range 3–18) in *F. crenata* and *F. japonica*, respectively. The mean expected heterozygosity (*He*) was 0.615 (range: 0.216–0.925) in *F. crenata* and 0.660 in *F. japonica* (range: 0.259–0.827). The microsatellite *mfc2* showed the highest variability, with *He* values of 0.925 in *F. crenata* and 0.827 in *F. japonica*. This microsatellite has the longest repeat $(CT)_{27}$, and generated a total of 21 and 13 alleles in *F. crenata* and *F. japonica*, respectively. The microsatellite *mfc5*, consisting of a $(CT)_{10}$ repeat, also had extremely large *He* values (0.901 and 0.88, respectively) and numbers of alleles (16 and 18, respectively). There were no significant deviations from Hardy-Weinberg expectations for the mean F_{IS} values (Wright 1965) in either *F. crenata* or *F. japonica*. However, in some individual loci, chi-square tests with the procedure of Li and Horvitz (1953) indicated significant deviations from Hardy-Weinberg expectations (Table 2).

We have developed nine polymorphic microsatellite markers, of which eight are applicable to both *F. crenata* and *F. japonica*. The polymorphic level was extremely high in both species. Tomaru et al. (1997) investigated genetic diversity and population differentiation of *F. crenata* using allozymes. According to their results, the average observed heterozygosity of 11 polymorphic allozyme loci in total populations was 0.213 (range 0.002–0.478). The average *Ho* value of eight microsatellite loci in *F. crenata* was 2.46 times higher than this allozyme figure.

We can apply these microsatellite markers to assess the range and exact distances of pollen flow for the investigated species. The ability of genetic makers to study pollen flow is represented by the paternity exclusion probability, which is the parameter used to describe the chance of correctly identifying pollen donors. The *Pe* values of *mfc2* and *mfc5* (the most variable loci in both species) were very high, 0.850 and 0.675 at *mfc2* and 0.901 and 0.781 at *mfc5* in *F. crenata* and *F. japonica*, respectively. Paternity exclusion probabilities for multiple loci (*Q*) calculated for all loci were extremely high: 0.999 and 0.998 in *F. crenata* and *F. japonica* respectively. These values correspond to the ability to find the true pollen donor among approximately 1000 mature candidate trees.

Significantly high F_{IS} values were observed at one and two loci of *F. crenata* and *F. japonica*, respectively, due to homozygote excess. If excessively inbreeding occurred in the populations, we may expect deviations from zero for most of the loci. Another explanation for deviations from Hardy-Weinberg expectations with an excess of homozygotes is the presence of null alleles. Null alleles of microsatellite regions, which occasionally fail to yield an amplification product (Weber and May 1989; Kooley et al. 1993; Pemberton et al. 1995), may arise through a point mutation at one or other priming site. If we include loci with null alleles, parentage analysis is more difficult. However, even if we exclude these SSR loci with high F_{IS} terms to calculate *Q* values, the paternity exclusion probabilities for the remaining loci are still high: 0.998 in *F. crenata* and 0.997 in *F. japonica*. Thus, we can conclude that the microsatellite markers we have developed are very useful for studying pollen flow in both *F. crenata* and *F. japonica*.

Acknowledgements The authors are particularly grateful to T. Ohkubo for allowing the use of his research plot, to H. Yoshimaru, T. Ujino, N. Tani, and K. Maruyama for their help in collecting samples, and to T. Nakashizuka for his critical advice while preparing this manuscript. We also thank T. Kawahara and T. Ujino for their advice on molecular techniques. This study was supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences.

References

Bjourson AJ, Cooper JE (1992) Band-stab PCR: a simple technique for the purification of individual PCR products. Nucleic Acids Res 20:4675

Chase MR, Moller R, Kessell R, Bawa KS (1996) Distant gene flow in tropical trees. Nature: 383:398–399

Cifarelli RA, Gallitelli M, Cellini F (1995) Random amplified hybridization microsatellites (RAHM): isolation of a new class of microsatellite-containing DNA cones. Nucleic Acids Res 23:3802–3803

Dayanandan S, Bawa KS, Kesseli R (1997) Conservation of microsatellites among tropical trees (Leguminosae). Am J Bot 84:1658–1663

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

Echt CS, MayMarquardt P, Hseih M, Zahorchak R (1996) Characterization of microsatellite markers in eastern white pine. Genome 39:1102–1108

Kajis J MH, Fowler JCS, Thomas M (1995) An evaluation of sequence tagged microsatellite site markers for genetic analysis *Citrus* and related species. Genome 38:349–355

Kitamura K, Okuizumi H, Seki T, Niiyama K, Shiraishi S (1992) Isozyme analysis of the mating system in natural populations of *Fagus crenata* and *F. japonica*. Jpn J Ecol 42:61–69

Kooley DJ, Bishop GA, McCaughan GW (1993) Allele non-amplification a source of confusion in linkage studies employing microsatellite polymorphism. Hum Mol Genet 44 : 397–401

Li CC, Horvitz DG (1953) Some methods estimating the inbreeding coefficient. Am J Hum Genet 5 : 107–117

Maeda T, Miyagawa K (1971) New sylvicultual techniques for natural regeneration of beech forests. In: Yanagisawa F, Yamaya K, Nakano M, Maeda T, Miyakawa K, Kato R, Ogata N (eds) New sylvicultural techniques for natural regeneration (in Japanese). Sobun, Tokyo, pp 179–252

Murray MG, Thompson W F (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8 : 4321–4325

Nakashizuka T (1987) Regeneration dynamics of beech forests in Japan. Vegetatio 69 : 169–175

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

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

Ohkubo T (1992) Structure and dynamics of Japanese beech (*Fagus japonica* Maxim.) stools and sprouts in the regeneration of the natural forests. Vegetatio 101 : 65–80

Ohkubo T (1996) Response of Japanese beech (*Fagus japonica* Maxim.) sprouts to canopy gaps. Vegetatio 124 : 1–8

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Non-amplifying alleles at microsatellite loci: a caution for parentage studies. Mol Ecol 4 : 242–252

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

Tsumura Y, Suyama Y (1998) Differentiation of mitochondrial DNA polymorphisms in populations of five Japanese *Abies*. Evolution 52 : 1031–1042

Tsumura Y, Yoshimura K, Tomaru N, Ohba K (1995) Molecular phylogeny of conifers using PCR-RFLP analysis of chloroplast genes. Theor Appl Genet 91 : 1222–1236

Ujino T, Kawahara T, Tsumura T, Yoshimaru H (1998) Development and polymorphism of simple repeat DNA markers for *Shorea curtisii* and other Dipterocarpaceae species. Heredity 81 : 422–428

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be tried using the polymerase chain reaction. Am J Hum Genet 22 : 24–49

Weir BS (1996) Genetic Data Analysis II. Sinauer Assoc, Sunderland, Mass.

Wright S (1965) The interpretation of population structure by *F*-statics with special regard to systems of mating. Evolution 19 : 395–420

Wu K, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241 : 225–235

Yamamoto S (1989) Gap dynamics in climax *Fagus crenata* forests. Bot Mag Tokyo 102 : 93–114