行政院國家科學委員會專題研究計畫　期中進度報告

核、葉綠體及粒線體微衛星 DNA 應用於青剛櫟（Cyclobalanopsis glauca）分類、親緣及生物地理之研究 (2/3)

計畫類別：個別型計畫
計畫編號：NSC93-2621-B-002-011-
執行期間：93 年 08 月 01 日至 94 年 07 月 31 日
執行單位：國立臺灣大學生命科學系

計畫主持人：何國傑

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 94 年 6 月 1 日

行政院國家科學委員會多年期專題研究計畫 93 年度進度報告書

核、葉綠體及粒線體微衛星 DNA 應用於青剛 （*Cyclobalanops glauca*）分類、親緣及生物地理之研究（2/3）

計畫編號：NSC 93-2621-B-002-011

執行期限：93 年 8 月 1 日至 94 年 7 月 31 日

主持人：何國傑 教授　國立台灣大學生命科學系

一. 摘要:

In this work, we attempted to study the genetic differentiation between populations of *Quercus glauca* in Taiwan using microsatellite markers to infer the evolutionary history in the last glaciation stage. We found that *Q. glauca* has relatively high within-population diversity (H_E = 0.741) and low population differentiation (F_{ST} = 0.042), but shows isolation by distance. The most-divergent populations, according to the average F_{ST} for individual populations in comparison with every other population, were found in populations Chinshuiying, Shanping and Hungyeh in southern Taiwan, and Paling in north-central Taiwan. Moreover, populations Chinshuiying, Shanping and Paling were recognized as being the source populations for gene recolonization after the last glaciation stage. In addition, the three sites of Wushe, Yangmingshan, and Chinshuiying exhibited the highest gene diversities which coincided with populations with the highest chloroplast DNA variations. This may have resulted from an admixture of colonization routes.

二. 緣起與目的:

Quercus glauca is an evergreen broadleaf tree found not only in Taiwan, but also in China, Korea, and Japan. *Q. glauca* is the most-commonly occurring plant among the 50 native species of the family Fagaceae in Taiwan. *Q. glauca* is a medium-sized evergreen tree growing up to 20 m high, and 0.6 m in breast height diameter. *Machilus-Castanopsis* is one of the two major types of evergreen broadleaf forests in Taiwan occurring at 400~1500 m. A large number of tree species appear in these forests, the most prominent among them belonging

to the families Fagaceae and Lauraceae. *Q. glauca* is found from sea level up to 1700 m, and is one of the dominant species of *Machilus-Castanopsis* forests (Liao, 1991; Hsieh et al., 1994). No pure stand has been found, but populations may form aggregates on steep hillsides as one of the pioneer species. With respect to the usage of *Q. glauca,* aboriginal people have used it in construction, in mushroom culture, and for making household items, such as wheels, various poles, etc. The only man-made plantations were established in mixed forests in selected sites in order to build fire-deterrent belts in areas with pine forests and a long dry season. *Q. glauca* is selected as material because the methodology and primers for the exploitation of *Quercus* spp are available and have been tested in numerous European species.

However, the population genetics of *Q. glauca* have never been studied in Taiwan. Using chloroplast DNA, Huang et al. (2002) examined the phylogeography of *Q. glauca*. They suggested the possibility that the southeastern part of Taiwan could have been a refugium during the last glaciation. A star-like genealogy was characterized in *Q. glauca*, as a general outcome of population expansion from a bottleneck.

Microsatellites are highly polymorphic, abundant, and fairly evenly distributed throughout the euchromatic part of genomes (Jeffreys et al., 1985; Winberg et al., 1993; Kumar and Rostad, 1998; Richard et al., 1999), and these properties have made them the most-popular marker among others in the past 10 years (Brown et al., 1996; Hahn and Grifo, 1996; Liu et al., 1996; Xiao et al., 1996; van de Ven and McNicol, 1996; Ayreset al., 1997Chen et al., 1997; Bowers et al., 1999; Ramsayet al., 1999; Li et al., 2000a, 2000b;Schlotterer, 2004). Because of the great variability in the number of repeat units at most microsatellite loci, simple sequence repeat (SSR) data provide an important source of molecular markers for many areas of genetic research. Using microsatellites to study the population genetic structure of *Q. glauca* allowed us to reappraise the findings described above. The four microsatellites using in this study were in Table 1.

The aims of this work were: (i) to estimate the genetic diversity within and between populations of *Q. glauca* in Taiwan and to test, in particular, whether central Taiwan is

the area with the highest genetic diversity; and (ii) to examine patterns of genetic structure differentiation between populations in order to infer the possible evolutionary history since the last glaciation.

三. 結果與討論:

Microsatellite diversity within populations — Averaged over the four loci, the number of alleles per locus (A) was 5.75 to 7.25 at the population level and 7.25 to 8.75 at the group level (Table 2). Allelic richness (A_R) was calculated for gene copies; there was a maximum value of 6.42 in the Wu population and a minimum value of 5.08 in the Wl population. Coincidently, the Wu population harbored the highest H value (0.765), while population Wl exhibited the lowest value (0.682). The fixation index, a measure of heterozygotic deficiency, statistically deviated from zero in all populations.

If the population was grouped using geographical approximations, it could be divided into group A (1, 2, and 4), group B (3, 5, and 8), group C (6 and 7), and group D (9 and 10) (Table 2). Group C (populations 6 and 7) of southern Taiwan harbored the highest genetic diversity but not the greatest number of alleles nor allelic richness. The fixation indices statistically deviated from zero at the group level, due to Wahlunds' effect.

The four SSR loci investigated in the present study were polymorphic in all populations. The number of alleles observed per locus in samples of 181 individuals ranged from 8 and 17, with an overall total of 45 alleles scored over the four loci (Table 3). Among the four loci, Mic67 was the most polymorphic. The size ranges of PCR products corresponding to these alleles were roughly between 145 and 268 nucleotides (Table 3). We analyzed the flanking region and found substitutions and indels. Four of 144 alleles (in 72 individuals, about 2.8% of total alleles) showed variations in length. Without knowing the flanking sequence, the same length of sequence could be wrongly identified as the same allele. In the case of variation in length of a flanking sequence, the allelic identity was determined by the length of the core sequence. Providing for 2.8% length variations in flanking regions of 72 individuals, less than 2% of the total alleles may have been wrongly identified after

subtracting the four alleles mentioned above.

Overall genetic diversities (H_E) were similar for each locus and ranged from 0.643 to 0.820. The average genetic diversity (H_E, 0.741) was higher than H_O (0.612). This was also reflected in the positive F_{IS} values indicating the existence of more homozygotes in the population than expected under Hardy-Weinberg equilibrium (Table 3). A low variation in F_{ST} and only 4.2% of the total variation existed among populations.

Genetic differentiation and isolation by distance — Genetic differentiation between populations was low as revealed by SSR markers using AMOVA, and F_{ST} varied from 0.011 to 0.041 (data not shown). The unexpectedly high level of gene flow as revealed by the low F_{ST} value was found between Pa and Wu, Wl and Ym, Yu and Ch, and Yu and Ta. Interestingly, through the calculation of the average F_{ST} for individual populations in comparison with every other population, we found that Cy was genetically the most-distinct population, followed by Sa and Hy (Fig. 1), all situated in the south. The Paling population in the north-central area had a value comparable to Hy.

A Mantel test for SSRs also showed a significant correlation between population differentiations measured as $F_{ST}/(1 - F_{ST})$, and this significantly increased with the natural logarithm of the geographical distance between populations ($r = 0.5652$; $p = 0.007$ with 1000 permutations). To test the correlation between geographical distances and genetic distances for a specific population against to the remaining populations, the r value (correlation coefficient) was used in a study to find the source population of the Northern Fulmar (*Fulmarus glacialis*) (Burg et al., 2003). We have obtained R^2 values from calculations of the simple Mantel test generated by zt, a software tool for performing simple and partial Mantel tests (Bonnet and Van de Peer, 2002). According to the R^2 value for each population, all of the regressions were significant. We found that four populations, i.e., Pa, Cy, Sa, and Wl, had values of ≥ 0.8 (Table 4) and can be considered source populations for gene recolonization after the last glaciation stage. Populations Pa and Wl are located in the northern portion of the Central Mountain Range (CMR), and Sa and Cy are in the south.

Genetic relationships of populations — The population Neighbor-joining trees according to Nei's D (Tamura and Nei, 1983) and Goldstein's Dmμ (1995) are shown in Figure 2A and B, respectively. The first major cluster included populations Hy, Sa, and Cy. The second cluster included populations Ch and Yu. The third included Wl, Ym, Wu, Pa, and Ta. These two dendrograms, however, are not the same in terms of population grouping, and no high bootstrap values exist.

Same alleles present in both eastern and western populations of the CMR can be observed in Table 5. For example, in locus Mic 57, allele 223 is present in populations 3, 9, and 10; allele 226 is present in populations 6 and 8; and allele 232 is present in populations 2, 5, and 7. Every population harbored few unique alleles, because there are no geographically isolated populations. Similar conditions could be observed for the three other loci.

In conclusion, we found the southern part of Taiwan showed the highest population genetic divergence of *Quercus glauca*, compared to all the others which suggests the existence of a potential major refuge. This is in agreement with a previous report according to chloroplastic markers (Huang et al., 2002). Even though there has been gradual erosion of nuclear differentiation among populations, this study implies that a correlation still exists between maternal lineage (cpDNA) and nuclear marker variations of *Quercus glauca* in Taiwan.

Table 1. Core sequence for each microsatellite locus and primer sequence. The annealing temperatures for all were 50 °C.

Locus	Core sequence	Primer sequence
Mic57	(CCA)n	5`-GCTAAGATTTATCGCAGCCATAGG-3`
		5`-TGAGGAGGTTGGTGGAGAAAA-3`
Mic67	(CCA)n+(CCG)n+(CCT)n	5`-TGGCTTATCCAATGTTTGTGATT-3`
		5`-CGGCTTAGAGATTGGTGTCAAAG-3`
Mic51	(CA)n+ (TA)n	5`-CAAAAACCTAAACCTACAAACGCTAAA-3`

		5`-AATAGCAAGAGAGAAGATGTTGCAAC-3`
Mic69	(TGG)n+(TTG)(TGG)$_2$CGG(TGG)$_2$	5`-CACAATCTGCCCACATCATC -3` 5`- GGATGGACGAAGAGAAAAGAT-3`

Table 2. Genetic variation in individual *Quercus glauca* populations or in each group of populations analyzed using microsatellite loci. *A*r, allele richness; *H*, genetic diversity; F_{IS}, Wright's inbreeding coefficient. The Geo-group refers to population groupings according to geographic location.

Population	Location	Sample size	No. of alleles/ locus	*A*r	*H*	F_{IS}
1. Wushe (Wu)	121.08°E, 24.03°N	20	7.25	6.42	0.765	0.215**
2. Paling (Pa)	121.23°E, 24.38°N	20	6.75	5.82	0.721	0.299**
3. Yuli (Yu)	121.15°E, 23.19°N	19	6	5.45	0.747	0.236**
4. Chilan (Ch)	121.33°E, 24.38°N	20	6.75	5.58	0.703	0.287**
5. Taroko (Ta)	121.37°E, 24.11°N	20	6.75	5.70	0.729	0.282**
6. Shanping[?] (Sa)	120.40°E, 23.00°N	20	7	6.24	0.739	0.443**
7. Chinshuiying (Cy)	120.43°E, 22.25°N	20	6	5.54	0.758	0.492**
8. Hungyeh (Hy)	121.03°E, 22.53°N	20	6.5	5.77	0.733	0.422**
9. Wulai (Wl)	121.33°E, 24.53°N	20	5.75	5.08	0.682	0.392**

10. Yangmingshan (Ym)	121.33°E, 25.08°N	12	6	6.00	0.756	0.367**
Geo-group						
A (1, 2, 4)	North-central	60	8.75	7.78	0.731	0.122**
B (5, 3, 8)	Eastern	59	8.25	7.68	0.741	0.132**
C (6, 7)	Southern	40	8	7.65	0.753	0.286**
D (9, 10)	Northern	32	7.25	7.25	0.704	0.179**

* Significant at the 5% level; ** Significant at the 1% level.

Table 3. Allelic diversity of microsatellite loci scored in *Quercus glauca*. Size distr. (nt), range of sizes of PCR products; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, Wright's inbreeding coefficient; and F_{ST}, relative differentiation according to allele identity.

Locus	Size distr. (nt)	No. of alleles	No. of heterozygotes	H_O	H_E	F_{IS}	F_{ST}
Mic67	220~262	17	115	0.604	0.643	0.028*	0.033*
Mic57	145~166	9	109	0.572	0.767	0.221**	0.042*
Mic51	250~268	8	88	0.461	0.737	0.332**	0.063**
Mic69	221~251	11	154	0.811	0.820	-0.019	0.030*
Mean	--	--	116.5	0.612	0.741	0.140	0.042

* Significant at the 5% level; ** Significant at the 1% level.

[1] The number of plants showing heterozygotic alleles in a total number of 191 individuals.

Table 4. Isolation by distance correlations for 10 populations, calculated separately based on the genetic distance (F_{ST}) and geographical distances between a given population and all other populations. The R^2 values were derived from calculations by simple Mantel test generated by zt, a software tool for simple and partial Mantel tests (Bonnet and Van de Peer, 2002).

Population	R^2 value	p value
Wu	0.7263	< 0.001
Pa	0.8603	< 0.001
Yu	0.3879	0.002
Ch	0.7662	< 0.001
Ta	0.7104	< 0.001
Sa	0.8374	< 0.001
Cy	0.7965	0.002
Hy	0.6092	< 0.001
Wl	0.9444	< 0.001
Ym	0.7712	< 0.001

TABLE 5. Representatives of allele-size distributions for locus *Mic57* in each population of *Quercus glauca*

Locus *Mic57*	Population									
Allele size	Wu	Pa	Yu	Ch	Ta	Sp	Cy	Hy	Wl	Ym
220	0	0	0	0	0	0.025	0	0	0	0
223	0	0	0.053	0	0	0	0	0	0.05	0.083
226	0	0	0	0	0	0.075	0	0.05	0	0
229	0.15	0.05	0.026	0.125	0.075	0.05	0.3	0.15	0.025	0.083
231	0.025	0.05	0	0	0	0	0	0.1	0	0
232	0	0.05	0	0	0.025	0	0.025	0	0	0
235	0	0.025	0.026	0	0.025	0.025	0.05	0	0.025	0.042
238	0.325	0.25	0.421	0.55	0.55	0.5	0.35	0.325	0.525	0.375
241	0.2	0.175	0.105	0.05	0.025	0.075	0.025	0.075	0.075	0
244	0.075	0.275	0.289	0.2	0.2	0.125	0.1	0.175	0.15	0.292
247	0.1	0.05	0.053	0.025	0.025	0.1	0.1	0.025	0.1	0
250	0.025	0.025	0.026	0.025	0.075	0	0.05	0.05	0	0.042
253	0.1	0	0	0	0	0.025	0	0	0	0
256	0	0.05	0	0	0	0	0	0	0	0.042
259	0	0	0	0.025	0	0	0	0	0.05	0
261	0	0	0	0	0	0	0	0.05	0	0
262	0	0	0	0	0	0	0	0	0	0.042

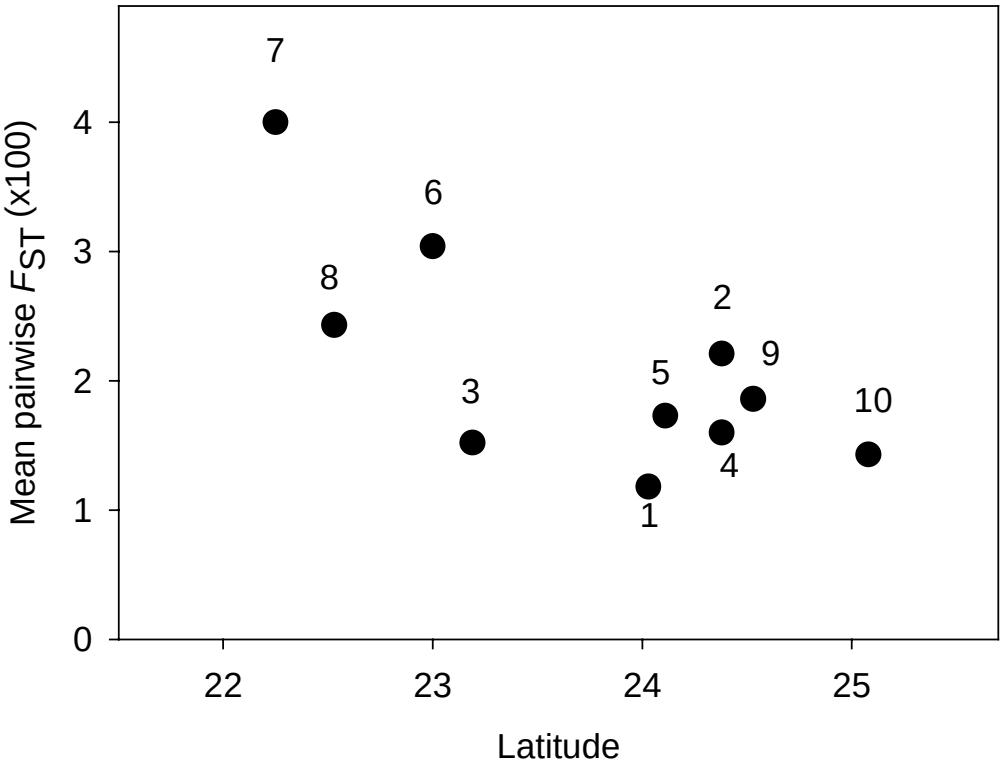
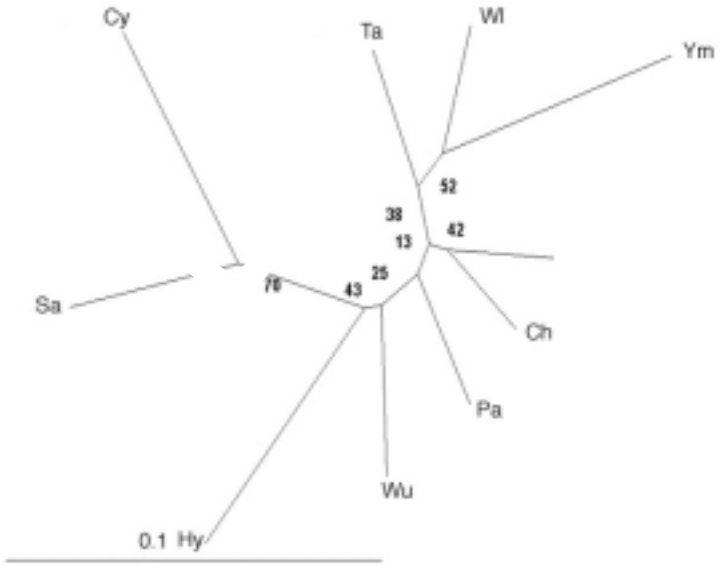


Fig 1. Plot of the average pairwise F_{ST} values for each population in comparison with the remaining populations against latitude in *Quercus glauca*. Chinshuiying (22.25°N, population no. 7) was genetically the most-distinct population, followed by Shanping (23.00°N, no. 6) and Hungyeh (22.53°N, no. 8), all located in southern Taiwan. For population numbers see Table 1.

(a) Nei's D, Neighbor-joining



(b) Goldstein's Dmμ, Neighbor-joining

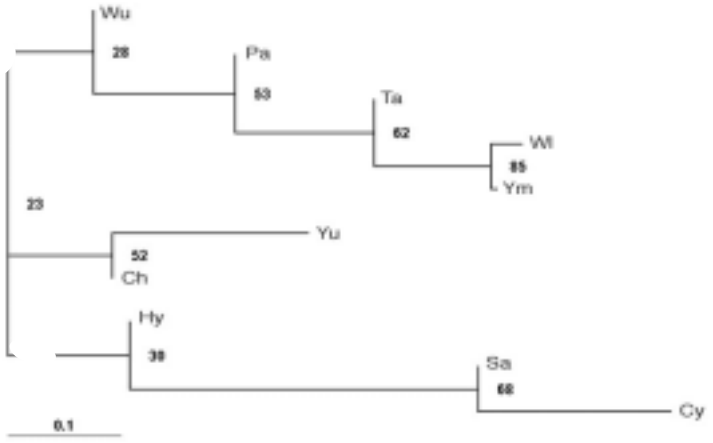


Fig 2. Unrooted Neighbor-joining tree of the 10 *Quercus glauca* populations drawn using Nei's D (Nei, 1983) and Goldstein's $D_{m\mu}$ (Goldstein et al., 1995) by (a) and (b) Neighbor-joining methods. The numbers in the figures are percentage values over 2000 bootstrap replicates. Only bootstrap values over 50% are presented.

四、參考資料

. Ayres, N. M., McClung, A. M., Larkin, P. D., Bligh, H. F. J., Jones, C. A. and Park, W. D. (1997) Microsatellite and a single nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germplasm. Theor. Appl. Genet. 94,773-781.

Bonnet, E. and van de Peer, Y. 2002. At: a software tool for simple and partial mantel tests. J. Statistical software 7, 1-12.

Bowers, J., Boursiquot, J., This, P., Chu, K., Johansson, H. and Meredith, C. (1999) Historical genetics: the parentage of Chardonnay, Gamay, and other wine grapes of northeastern France. Science 285,1562-1565.

Brown, S. M., Hopkins, M. S., Mitchell, S. E., Senior, M. L., Wang, T. Y., Duncan, R. R., Gonzalez-Candelas, F.and Kresovich, S. (1996) Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* Moench (L.)]. Theor. Appl. Genet. 93,190-198.

Burg, T.M., Lomax, J., Almond, R., Brooke, M, and Amos, W. 2003. Unravelling dispersal patterns in an expanding population of a highly mobile seabird, the northern fulmar (*Fulmarus glacialis*). Proc. Roy. Soc. London. 270, 979-984.

Chen, X., Temnykh, S., Xu, Y., Cho, Y. G. and McCouch, S. R. (1997) Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.) Theor. Appl. Genet. 95, 553-567.

Goldstein, D.B., Linares1, A.R. and Cavalli-Sforza, L.L. 1995. An evaluation of genetic distances for use with microsatellite loci. Genetics 139, 463-471.

Hahn, W. J. and Grifo, F. T. (1996) Molecular markers in plant conservation genetics. In *The impact of plant molecular genetics*, Sobral, B.W.S. (ed). Birkhauser, Boston, pp 113-116.

Huang, S.F., Hwang, S.Y. and Lin, T.P. 2002. Spatial pattern of chloroplast DNA

variation of *Quercus glauca* in Taiwan and eastern Asia. Mole. Ecol. 11, 2349-2458.

Hsieh, C.F., Shen, C.F. and Yang, K.C. 1994. Introduction to the flora of Taiwan, 3: Floristics, phytogeography, and vegetation. *In* Editorial Committee of the Flora of Taiwan [ed], *Flora of Taiwan*, 2nd ed, vol. 1, 7–18.Editorial Committee of the Flora of Taiwan, Taipei.

Jeffreys, A. J., Wilson, V. and Thein, S. L. (1985) Hypervariable "minisatellite" regions in human DNA. Nature 314,67-73.

Kumar, A. and Rogstad, S. H. (1998) A hierarchical analysis of minisatellite DNA diversity in Gambel oak (Quercus gambelii Nutt.; Fagaceae). Mol. Ecology 7,859-869.

Li, C.-D., Rossnagel, B. G. and Scoles, G. J. (2000a) The development of oat microsatellite markers and their use in identifying relationships among Avena species and oat cultivars. Theor. Appl. Genet. 101,1259-1268.

Li, C.-D., Rossnagel, B. G. and Scoles, G. J. (2000b) Tracing the phylogeny of the hexaploid Oat *Avena sativa* with satellite DNAs. Crop Sci. 40,1755-1763.

Liao, J. C. (1991) The taxonomic revisions of the family Fagaceae in Taiwan. Department of Forestry. College of Agriculture. Natl. Taiwan Univ., Taipei, Taiwan.

Liu, Z. W., Biyashev, R. M., Saghai Maroof, M. A. (1996) Development of simple sequence repeat markers and their integration into a barley linkage map. Theor. Appl. Genet. 93, 869-876.

Rajora, O. P., Rahman, M. H. and Dayanandan, S. (2001) Isolation, characterization, inheritance and linkate of microsatellite DNA markers in white spruce (*Picea glaueca*) and their usefulness in other spruce species. Mol. Gen. Genet. 264,871-882.

Ramsay, L., Macaulay, M., Cardle, L., Morgante, M., Ivanissevich, S. D., Maestri, E., Powell, W. and Waugh, R. (1999) Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plant J. 17, 415-425.

Richard, G-F., Henneguin, C., Thierry, A. and Dujon, B. (1999) Trinucleotide repeats and other microsatellites in yeasts. Res. Microbiol. 150,589-602.

Schlotter, C. 2004. The evolution of molecular markers-just a matter of fashion? Nature Rev. Genet. 5: 63-69.

Simth, S. and Helentjaris, T. (1996) DNA fingerprinting and plant variety protection. In *Genome mapping in plants*, Paterson, A.H., ed, New York, pp 95-110.

Tamura, K. and Nei, N. 1983. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mole. Biol. Evol. 10, 512-526.

van de Ven, W.T.G. and McNicol, R. J. (1996) Microdatellites as DNA markers in *Sikta spruce.* Theor. Appl. Genet. 93, 613-617.

Winberg, B. C., Zhou, Z., Dallas, J. F., Mclntype, C. L. and Gustafson, J. P. (1993) Characterization of microsatellite sequences from *Oryza sativa.* Genome 36,978-983.

Xiao, J., Li, J., Yuan, L., McCouch, S. R. and Tanksley, S. D. (1996) Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. Theor. Appl. Genet. 92,637-643.