

## 국내 소나무 집단에 있어서 cpSSR 표지자 변이체의 분포양상

洪鎔杓\* · 權海燕 · 金龍律

국립산림과학원, 산림유전자원부

### Distribution Pattern of cpSSR Variants in Korean Populations of Japanese Red Pine

Yong-Pyo Hong\*, Hae-Yun Kwon and Yong-Yul Kim

Department of Forest Genetic Resources, Korea Forest Research Institute, Suwon 441-350, Korea

**요 약:** 국내 소나무 19개 집단을 대상으로 cpSSR 표지자 분석을 통해서 관찰된 28개 변이체로부터 총 167개의 독특한 haplotype이 확인되었고, 동일한 haplotype을 보인 13개체가 10집단에 고르게 분포하였으며, 각 집단에서 관찰된 유효 haplotype의 수는 평균 13.37개로 나타났다. cpSSR haplotype의 집단내 다양도( $H_e$ )는 0.987로 계산되어 기존의 임목을 대상으로 한 연구에서 보고된 수치와 유사하거나 약간 높은 수치를 보였다. 각 haplotype을 구성하고 있는 cpSSR 변이체를 대상으로 각 집단에서의 다양성( $S.I.$ )을 계산한 결과 강원도 영월집단이 1.109로 계산되어 가장 높은 수치를 보였으며, 경북 문경집단이 0.411로 가장 낮은 수치를 보였다(평균 0.887). 관찰된 cpSSR 변이체들의 대부분이 19개 집단에 공통적으로 존재하는 것으로 나타났으며(97.62%), 집단간에 cpSSR 변이체 분화는 미약한 것으로 나타났다( $\Phi_{ST} = 0.024$ ) cpSSR 표지자의 높은 돌연변이 발생빈도가 주요 원인인 것으로 추정된다. 반면에 비교 가능한 173개 집단 쌍 간에 동일한 haplotype이 전혀 존재하지 않는 집단 쌍이 39쌍으로 나타나 집단간의 유전적 유연관계에 대한 직접적인 비교가 불가능했으며, 따라서 분석된 19개 집단간에 유전적 교류가 자유롭게 일어나지 않는 것으로 나타났다. cpSSR 표지자 변이체의 분포양상과 기존의 I-SSR 표지자 변이체의 분포양상을 비교 고찰해 볼 때 국내 소나무 유전자원의 효율적인 관리를 위해서는 분석된 소나무 집단의 현재 위치 정보와 유전정보가 함께 고려되어야 할 것으로 생각된다.

**Abstract:** A total of 167 peculiar haplotypes confirmed from 28 cpSSR variants that were observed in 19 populations of Japanese red pine in Korea through cpSSR marker analysis. Thirteen individuals that showed identical haplotype dispersed evenly in 10 populations, and the average number of effective haplotype within population was 13.37. Estimate of genetic diversity ( $H_e$ ) was 0.987 on the basis of cpSSR haplotype variants that was equivalent to or higher than the estimates reported in other studies on some forest tree species. Estimation of genetic diversity ( $S.I.$ ) on the basis of cpSSR variants composing each haplotype revealed the highest estimate of 1.109 for the population of Gangwon-Yeongwol and the lowest estimate of 0.411 for the population of Gyeongbuk Mungyeong with the average of 0.887. Most of observed cpSSR variants appeared to exist commonly in 19 populations (97.62%), and genetic differentiation of cpSSR variants among populations was turned out to be weak ( $\Phi_{ST} = 0.024$ ). Relatively fast rate of mutation of cpSSR marker might be a major cause for such weak population differentiation. There was no identical haplotype shared between 39 population pairs of 173 pair-wise population pairs. Estimation of genetic distance among 19 populations on the basis of cpSSR haplotype frequencies was impossible. Therefore, reconstruction of genetic relationships among pair-wise population pairs was also impossible, that might be resulted from restricted migration among 19 populations. Considering the observed distribution patterns of cpSSR variants in addition to the previous studies on I-SSR variants, informations on the present geographic location and genetic status of populations should be considered together for effective sustainable management of the genetic resources of Japanese red pine in Korea.

**Key words :** cpSSR variants, Japanese red pine, genetic diversity, population differentiation, genetic resources

\*Corresponding author  
E-mail: yphong@foa.go.kr

## 서 론

소나무(*Pinus densiflora*)는 우리나라의 대표적인 침엽수 종으로 예로부터 건축재, 연료재, 선박재, 관곽재, 구황식 등으로 인간에 의해 빈번히 이용되어 왔으며, 이와 관련하여 고려사나 조선왕조실록, 비변사등록 등에 소나무림 보호 및 국가적 차원에서의 소나무 조림 등에 대한 다수의 기록이 남아 있다. 최근 Hong 등(2004, 2006)은 I-SSR 표지자를 사용하여 전국 19개 소나무 집단에 대한 유전다양성 분석을 수행하였는데, 우리나라의 소나무 집단이 타수종에 비해 비교적 높은 유전다양성을 나타내었으며( $S.I.=0.453$ ) 전반적으로 집단간 유전적 거리와 지리적 거리가 일치하지 않는 것으로 나타나 소나무의 유전자원 보존 전략 수립을 위한 유전구역 설정 시 이를 참고해야 할 것이라고 하였다.

일반적으로 엽록체 DNA의 염기서열은 핵 DNA(nuclear DNA)에 비해 느리게 진화되며(낮은 돌연변이율), 유성생식 시 감수분열과정에서 수반될 가능성이 있는 일련의 재조합 과정을 거치지 않고 부모 중 한 쪽에 의해서만 자식에게 일방적으로 전달되기 때문에(uniparental inheritance, 소나무는 부계유전) 수종 및 개체 간 변이가 핵에 존재하는 DNA의 경우보다 훨씬 적다. 그러나 엽록체 DNA의 특정 부위에는 핵 DNA처럼 SSR(simple sequence repeats, 單純鹽基反復)이라는 지역이 존재하는데, 이는 소수의 염기로 구성된 主要素(motif)가 반복 배열되어 이루어지는 genome 상의 특정부위로 엽록체 genome 상에 무작위적으로 분포하는 것으로 알려져 있다(Tautz, 1993). 노 등(2005, GenBank Accession NC 004677)에 의해서 잣나무의 엽록체 DNA의 염기서열이 결정됨에 따라 잣나무 전체 엽록체 DNA의 배열상에 한 개의 염기가 반복적으로 나타나는 부위가 35개(T: 19개, A: 14개, G: 1개, C: 1개), 두 개의 반복염기(TA)가 반복적으로 나타나는 부위가 4개 존재하는 것이 확인되었다. 이러한 SSR 부위들은 해당 SSR을 구성하는 主要素의 반복수에서 동일 종내 개체별로 많은 차이를 나타내기도 하는데, 이는 주로 DNA 복제 과정 중 DNA 중합효소(DNA polymerase)의 불완전한 반응으로 인해서 나타나는 것으로 추정되는 DNA polymerase slip에 의한 single repeat unit의 첨가(addition) 또는 결실(deletion)에 의해 생겨나는 frameshift 돌연변이의 생성 과정을 통해 만들어진다고 알려져 있다(Gragg 등, 2002). 이들 단순반복염기에 존재하는 DNA들을 그 양쪽에 위치한 DNA 정보를 토대로 제작된 primer를 이용해서 PCR(polymerase chain reaction)법에 의해서 대량 증폭할 수 있으며, Provan 등(1998)은 *Pinus torreyana*의 엽록체 SSR DNA(이하 cpSSR) 분석을 통해 cpSSR의 돌연변이율을 계산한 결과 cpSSR 부위 염기서열의 돌연변이율이 세대당 약  $3.2-7.9 \times 10^{-5}$  mutation/allele/generation으로 식물체내

다른 genome 부위에 비해 상대적으로 높다고 하였는데, 이는 cpSSR이 개체식별이나 집단 유전학적 연구에 있어 기존의 표지자들에 비해 보다 풍부한 정보를 제공할 수 있다는 것을 의미한다. 이러한 다양한 主要素 반복수에 따른 변이에도 불구하고 전반적인 엽록체 DNA의 진화적 안정성으로 인해서 동일 속내의 종 간에 단순반복염기와 이웃하고 있는 DNA들이 동일한 구성을 하고 있는 경우가 있다. 따라서 분석 대상종에 대한 특이적인 cpSSR primer를 개발하지 않더라도 동일 속에 속한 별개의 종으로부터 개발된 cpSSR primer를 이용할 수 있는 장점이 있다. 또한 소수의 반복염기로 구성된 cpSSR 표지자는 대부분 특정 유전자를 coding하거나 그 발현에 영향을 미치지 않기 때문에 자연선택의 직접적인 영향을 받지 않으며 표현형 상 공우성적 특성을 나타내므로, 최근 종 및 개체 식별과 집단 분석 등에 유용한 분자적 표지자로서 널리 이용되고 있다(Anzidei 등, 1999). 임목에서는 1994년 Wakasugi 등에 의해 해송(*Pinus thunbergii*)의 엽록체 DNA 염기서열이 완전히 알려지면서 cpSSR의 이용이 가능하게 되었는데, 또한 대부분의 침엽수에서 엽록체 DNA는 부계로만 유전되는 반수체 계놈이기 때문에 생식세포 분열시 계놈 간 재조합이 거의 일어나지 않으며 대부분의 변이체가 DNA 염기상에 발생하는 돌연변이를 대변하는 관계로 집단의 이주 및 분화 역사를 재구성하는 연구에 nuclear SSR보다 유용한 정보를 제공할 것으로 기대된다.

본 연구는 cpSSR 표지자를 분석하여 organelle genome 수준에서 국내 소나무 집단의 유전변이량 및 분포를 추정함으로써 전국 소나무 집단의 유전구역 설정과 이를 통한 유전자원 보존 전략에 유용한 정보를 얻고자 수행하였다.

## 재료 및 방법

### 1. 시료 및 cpSSR 분석

국내 19개 집단을 대상으로 299개체의 total DNA를 cpSSR 분석에 사용하였다(Figure 1). cpSSR haplotype의 확인을 위해서 해송의 엽록체 DNA 염기서열을 기초로 개발된 20개의 cpSSR primer 중(Vendramin 등, 1996) 소나무에서 다형성을 나타내는 것으로 판명된 6개 primer를 사용하였으며(Pt1254, Pt30204, Pt15169, Pt71936, Pt100783, Pt45002), 각 primer의 forward sequence의 5' 말단에는 19 base pair의 oligonucleotide를 붙여 합성함으로써 PCR 반응시 첨가된 M13(-29) IRD 800 dye-labeled primer와 반응하여 형광표지가 감지될 수 있도록 하였다. PCR 증폭은 김 등(2002)의 방법을 사용하였는데, 반응용액 11  $\mu$ l 당 template DNA 5ng, M13(-29) labeled primer 0.02  $\mu$ M, Forward/Reverse SSR primer 각 0.1  $\mu$ M, 0.1 mM dNTP, 1unit의 *Taq* polymerase가 포함되도록 하였다.

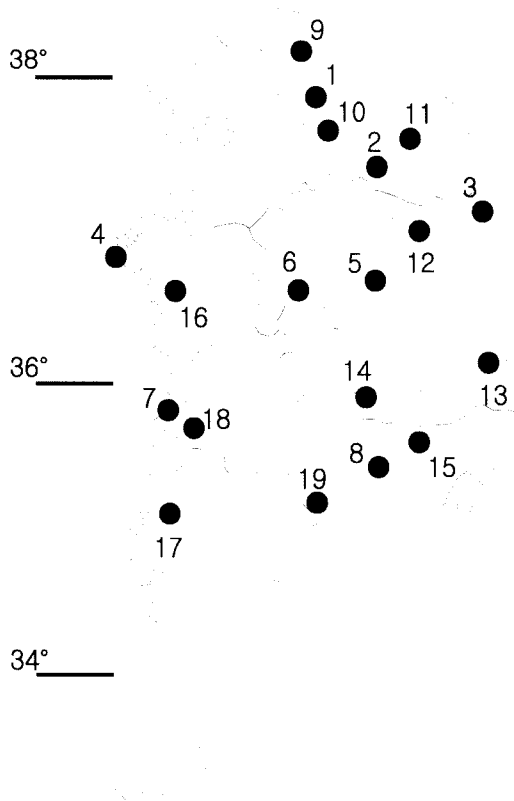


Figure 1. Location of sampled populations. Population names refer to those corresponding to the numbers in Table 1.

PCR 반응은 94°C에서 1분간 열변성 후 57°C에서 1분간 annealing, 72°C에서 2분간 증폭이 이루어지는 과정을 6회 반복한 후, 다시 94°C에서 30초간 열변성, 52°C에서 30초간 annealing, 72°C에서 1분간 증폭의 과정을 30회 반복하고 마지막으로 72°C에서 5분간 추가로 증폭되도록 하였다. PCR 증폭산물의 분획은 7% long range polyacrylamide gel(FMC)을 사용, Li-cor automatic DNA 4000 sequencer로 1시간 동안 수행하였으며, 여기서 얻어진 gel 영상을 관찰하여 증폭산물의 크기를 기준으로 개체별 haplotype을 결정하였다. 증폭산물의 크기를 결정하기 위해서 DAN size standard(50-350bp concentrated sizing standard, LI-COR inc., USA) 동시에 전기영동하였으며, 관찰된 각 primer별 염록체 DNA의 haplotype은 증폭산물의 크기에 준해서 명명되었다.

## 2. Data 분석

분석된 집단에서 cpSSR 표지자 변이체의 다양성을 추정하기 위해 POPGENE 1.31 program(Yeh 등, 1999)을 이용하여 관찰된 각각의 변이체에 대한 집단내 다양성은 Shannon's index(S.I.; Shannon, 1948)를 구하였다. 또한 관찰된 cpSSR 변이체들이 감수분열 시에 재조합이 일어나지 않는 염록체 게놈 상에 놓여있기 때문에 각 개체별

로 관찰된 변이들의 조합에 의해서 각 haplotype을 결정하고 다음과 같은 통계량을 계산하였다. 각 집단에서 관찰된 전체 시료의 haplotype 결정에 기여한 유효 haplotype의 수( $N_e$ )를 다음과 같은 공식을 이용해서 계산했다.

$$N_e = \frac{1}{\sum_{i=1}^m p_i^2}$$

위의 수식 중,  $p_i$ 는  $i$  번째 haplotype의 상대빈도,  $m$ 은 집단내에서 관찰된 haplotype의 수를 의미한다.

또한 분석된 집단에서 cpSSR 표지자 변이체 다양성을 추정하기 위해 POPGENE 1.31 program(Yeh 등, 1999)을 이용하여 Shannon's index(S.I.; Shannon, 1948)를 구하였으며, Arlequin 2.0 program(Schneider 등, 2000)에 의해 동일 집단내 시료 중에서 선정된 임의의 haplotype들이 동일한 haplotype일 확률( $H_e$ : haplotype 다양도, Nei, 1987)을 계산했고, cpSSR 표지자 변이체의 분포양상을 파악하기 위해서 Euclidean distance의 유전적 거리를 기초로 AMOVA 분석을 실시하여 집단간 유전적 분화의 정도를 계산하였다(Excoffier 등, 1992).

## 결과 및 고찰

6개 cpSSR primer를 이용한 cpSSR 표지자 분석결과 총 28개 변이체가 관찰되어[Pt1254(3개), Pt30204(3개), Pt15169(7개), Pt71936(9개), Pt100783(2개), Pt45002(4개), 평균 4.7개/primer], 분석된 개체별로 관찰된 모든 변이체들을 취합한 결과 총 167개의 독특한 haplotype이 확인되었으며(Figure 2), 이 중 50개의 haplotype은 2개 이상의 개체목들에서 공통으로 관찰되었는데 가장 빈도가 높게 관찰된 haplotype은 13개체가 동일한 haplotype을 지닌 것으로 나타났으며 10집단에 고르게 분포하였다. 각 primer들에 의해서 생성된 haplotype 변이체들의 전체 집단에 있어서의 분포정도를 분석한 결과(Table 1), 이 중 가장 많은 변이체가 확인된 Pt15169 primer에 의해서 증폭된 haplotype 변이체의 경우 이들 변이체가 다수의 집단에 고르게 분포되어 있었으며, Pt45002 primer에 의해서 확인된 변이체들은 상대적으로 일부 집단에 다소 편중되어 출현하는 것을 알 수 있었다. 각 집단에서 관찰된 유효 haplotype의 수는 모든 개체(19 개체)가 서로 다른 haplotype을 지닌 충남 태안집단에서 19개로 가장 높았으며, 경남 함안집단에서 7.54개로 가장 낮았는데 이 집단에서 분석된 개체의 수는 총 11개로 나타났다. 전체 19개 집단에서 분석된 집단별 평균 개체 수는 14.32개이었으며, 평균 유효 haplotype의 수는 13.37개로 나타나 각 집단에서의 cpSSR 표지자 분석에 의한 개체 인식력(각 유효

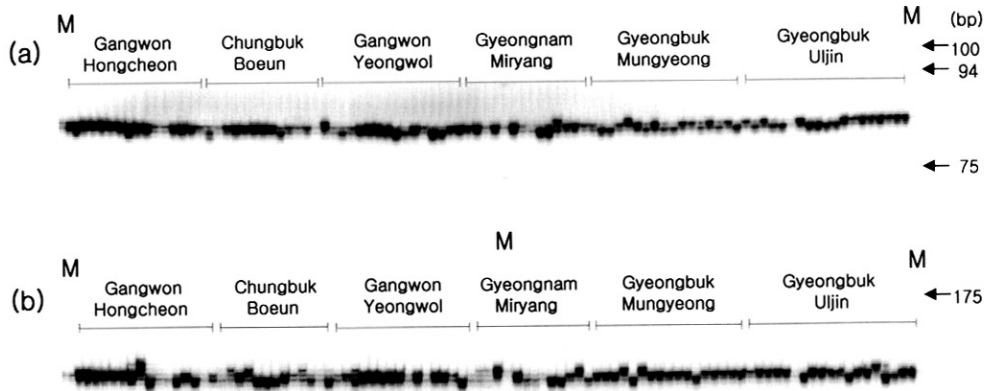


Figure 2. Example of cpSSR amplicon profiles. (a) Pt1254 primer. (b) Pt30204 primer. ‘M’ denotes sizing standard IRDye800™.

Table 1. Diversity of I-SSR variants within population.

	Population	Individual No.	No	Ne	He	S.I.
1	Gangwon-Hongcheon	16	15	14.22	0.99	0.836
2	Gangwon-Yeongwol	16	16	16	1.00	1.109
3	Gyeongbuk-Uljin	18	16	14.73	0.99	0.882
4	Chungnam-Taean	19	19	19	1.00	0.990
5	Gyeongbuk-Mungyeong	17	14	11.56	0.97	0.411
6	Chungbuk-Boeun	13	12	11.27	0.99	0.813
7	Jeonbuk-Buan	17	17	17	1.00	0.973
8	Gyeongnam-Haman	14	11	7.54	0.93	0.844
9	Gangwon-Hwacheon	17	17	17	1.00	0.953
10	Gangwon-Hoengseong	17	15	12.57	0.98	0.897
11	Gangwon-Jeongseon	16	16	16	1.00	0.989
12	Gyeongbuk-Bonghwa	12	12	12	1.00	0.803
13	Gyeongbuk-Pohang	17	16	15.21	0.99	0.879
14	Gyeongbuk-Goryeong	14	13	12.25	0.99	0.931
15	Gyeongnam-Miryang	12	12	12	1.00	0.991
16	Chungnam-Cheongyang	17	14	11.56	0.97	0.838
17	Jeonnam-Naju	13	12	11.27	0.99	0.887
18	Jeonbuk-Jeongeup	12	10	10.29	1.00	0.913
19	Gyeongnam-Hadong	17	15	12.57	0.98	0.911
	Mean	15.47	14.32	13.37	0.987	0.887

No: total number of haplotypes, Ne: effective number of haplotypes  
 He: unbiased haplotype diversity (Nei 1987), S.I.: Shannon's index

haplotype 당 개체 수 = 1.07개)이 매우 높은 것으로 나타났다(Parducci, 2001).

cpSSR haplotype의 집단내 다양도(He)를 계산한 결과 모든 집단에서 0.93 이상의 높은 수치(평균 0.987)를 보여 각 집단내에 매우 다양한 haplotype이 존재함이 확인되었다. 높은 수준의 haplotype 다양도는 기존의 연구에서도 관찰된 일반적인 현상으로 스코트랜드와 유럽 전역에 자생하는 *Pinus sylvestris*를 대상으로한 연구에서 0.977의 높은 값이 보고된 바 있다(Provan 등, 1998). 미국 북서부에 자생하는 *Pinus albicaulis*를 대상으로한 연구에서 0.907의 haplotype 다양도가 보고되었으며(Richardson 등, 2002),

포르투갈의 해안성 소나무류인 *Pinus pinaster*에서 haplotype 다양도가 0.87로 상대적으로 낮은 다양도가 보고된 바 있으나(Ribeiro 등, 2001), 유럽에 자생하는 *Abies* 속 4개 수종에 대한 분석결과 haplotype 다양도가 평균 0.964로 높게 나타났으며, 그 중 북부 이탈리아에 유일하게 잔존하고 있는 *Abies nebrodensis* 단일 집단에서도 0.846의 비교적 높은 haplotype 다양도가 보고된 바 있다(Parducci, 2001). 본 연구에서 계산된 haplotype 다양도(0.987)는 기존의 연구에서 보고된 수치와 유사하거나 약간 높은 수치로서 이를 토대로 볼 때 각 국내 소나무 집단 내에 존재하는 개체들에는 본 연구에서 분석된 6개 primer

**Table 2. Distribution pattern of cpDNA haplotype variants among 19 populations.**

	Pt1254	Pt30204	Pt15169	Pt71936	Pt100783	Pt45002	Over all loci
No. of variants	3	3	7	9	2	4	28
$\theta_{ST}^*$	0.029	0.014	-0.007	0.029	0.014	0.066	0.024

\*Weir and Cockerham (1984)

에 의해서 분석된 엽록체 게놈상에 매우 다양한 cpSSR 표지자 변이체의 조합이 존재하는 것이 확인되었다. 그러나 본 연구에서 각 primer별로 관찰된 변이체의 수(Table 2)를 기준으로 볼 때 이들 변이체들이 상호간에 연관되어있지 않다는 전제하에서 총 31,752개의 cpSSR 표지자 변이체의 조합이 가능함에도 불구하고 본 연구에서 단지 167개의 조합만이 확인된 것은 이들 표지자들(loci)이 게놈 상에 밀접하게 연관되어 있기 때문인 것으로 생각되는데, 이는 엽록체 게놈이 원형의 plasmid 형태로 존재하며 생식 세포 분열시 게놈 간에 유전자를 교환하는 교차가 거의 일어나지 않는 특성에서 기인하는 것에서 그 원인을 찾을 수 있다.

비록 각각의 haplotype에 존재하는 6개 표지자들이 상호간에 연관되어 있을지라도 각각의 표지자에서의 돌연변이 발생은 상호간에 독립적으로 일어날 것이라는 가정하에 각 집단내에 존재하는 haplotype 변이체가 지니고 있는 cpSSR 표지자 변이체의 종류에 근거해서 유전다양성(S.I.)을 계산한 결과 강원도 영월집단이 유전다양성이 1.109로 계산되어 가장 높은 수치를 보였으며, 문경집단이 0.411로 가장 낮은 수치를 보였다(Table 1). 전체 집단에 대한 평균은 0.887로 나타나 비교적 다양한 cpSSR 표지자 변이체가 집단내에 존재함을 알 수 있었다. 이는 기존의 소나무 집단을 대상으로 I-SSR 표지자 분석을 통해서 보고된 유전다양성(0.453; Hong 등, 2005) 보다 2배 정도 높은 수치인데, 그 원인은 분석된 표지자의 진화적 특성에 기인할 수도 있겠으나 간접적으로 Shannon's index의 계산 원리에서 유추해낼 수 있다. 비록 동일한 개체를 대상으로 DNA 분석을 수행하더라도 I-SSR 표지자의 경우에는 각각의 유전자좌에서 관찰될 수 있는 변이체의 최대치가 2개(중복산물의 유무)에 불과하나, cpSSR 표지자의 경우에는 변이체를 구성하는 simple sequence repeat DNA의 가감에 의해서 생성될 수 있는 가능한 변이체의 수에 제한이 없다. 따라서 분석 집단내에서 관찰된 변이체의 다양성을 계산하는 통계치인 Shannon's index는 관찰된 각 표지자 변이체 빈도를 근거로 계산된 수치를 합산한 값이기 때문에 변이체의 종류가 많으면 많을수록 그 수치가 증가하는 특성이 있다. 예를 들면, 변이체 수가 최대 2개인 I-SSR 표지자의 경우 그 최대 값은 0.6931(각 변이체의 빈도가 0.5일 때)로 한정되어져 있다. 반면에 이론상 그 수가 제한되어 있지 않은 cpSSR 표지자의 경우 각

**Table 3. Results of AMOVA for 19 populations.**

Source of variance	d.f.	Component variance
Among populations	18	2.38%
Within populations	275	97.62%

집단에서 관찰된 변이체의 종류가 증가함에 따라 다양성 정도가 증가하게 된다. 따라서 동일한 수종에 대한 DNA 분석 결과일지라도 관찰되는 변이체의 특성에 차이가 나는 표지자로부터 계산된 유전다양성 정도를 직접적으로 비교하는 것은 큰 의미가 없다고 할 수 있다.

전국 19개 집단에 있어서 cpSSR 표지자의 분포양상을 분석한 결과 관찰된 유전변이의 대부분이 19개 집단에 존재하는 집단내 개체간 차이에 의해서 기인하는 것으로 나타났다(97.62%), 따라서 낮은 수준의 유전분화가 집단 간에 일어난 것으로 관찰되어( $\Phi_{ST} = 0.024$ , Table 3) 소나무속 3개 수종을 대상으로 수행된 cpSSR 표지자 분석결과와 유사한 수준이었다(*Pinus sylvestris* : 0.0324, Provan 등, 1998; *Pinus albicaulis* : 0.046, Richardson 등, 2002; *Pinus pinaster* : 0.023, Ribeiro 등, 2001). 이는 기존의 소나무를 대상으로 수행된 핵 게놈에 존재하는 I-SSR 표지자의 분석결과에서 계산된 분화정도( $\Phi_{ST} = 0.092$ , Hong 등, 2006)의 약 1/4에 불과한 낮은 수준으로 나타났다. 엽록체 게놈과 핵 게놈으로부터 계산된 집단간 유전분화에 있어서 이와 같은 차이를 초래한 진화적 요인이 확실히 구분되어질 수는 없겠으나 게놈 집단의 크기 차이에 의한 유전자 부동의 영향과 돌연변이율 차이 등의 복합적인 요인이 작용했을 것으로 추정된다. 이들 연구에서 분석된 cpSSR 표지자와 I-SSR 표지자의 경우 이들이 유전자를 coding하고 있는 부분이 아니기 때문에 자연선택의 영향을 직접적으로 받지 않는 것으로 알려져 있으며(selectively neutral DNA markers), 부계유전양식을 보이는 소나무 엽록체 게놈의 경우 유전자 이주의 경로가 핵 게놈의 이주와 동시에 일어나며(화분 및 종자) 분석된 국내 소나무 집단의 분포형태를 볼 때 지리적 연속성이 단절된 상태이며 거리상 충분히 떨어져 있으므로 유전자 이동에 따른 영향은 최근 수 세대 동안에는 자유롭게 이루어지지 못했을 것으로 생각된다. 따라서 두 게놈이 돌연변이율이 동일하고 자연선택의 대상이 아니라면, 엽록체 게놈이 반수체(1N)이기 때문에 자웅동주인 소나무의 경우 2배체(2N)인 핵 게놈에 비해서 게놈집단의 크기가 절반에 불과할 것으

로 기대된다. 또한 핵 게놈과 부계유전되는 엽록체 게놈은 화분이나 종자에 의해서 이주가 일어나더라도 동등한 비율로 일어날 것이다. 이론적으로 동일한 과정의 유전자 부동(random genetic drift)을 겪게 되면 집단의 크기 차이로 인해서 평형상태에서 핵 게놈에 비해서 엽록체 게놈에는 집단간 분화현상이 더 크게 초래될 것으로 추정되어진다 [ $F_{ST} = 1/(1 + 4Nm)$ ,  $N =$  집단의 크기,  $m =$  이주율 : Slatkin, 1987]. 그러나 본 연구에서 cpSSR 표지자 분석결과 나타난 집단간 분화정도는 위의 공식에 근거해서 예상되었듯이 집단간 분화현상이 증가된 게 아니라 오히려 I-SSR 표지자의 경우에 비해서 1/4 정도에도 미치지 못하는 것으로 나타났다. 따라서 단순히 유전양식의 차이(부계 유전 대 양친유전)로 인한 게놈집단의 크기가 1/2에 불과한 엽록체 게놈에서 예상되는 집단간 유전분화의 심화정도를 상회하고도 남는 다른 진화적 요인이 작용했을 것으로 추정된다.

일반적으로 식물체에 있어서 DNA의 돌연변이율은 엽록체 DNA와 핵 DNA 간에 차이가 큰 것으로 알려져 있으며, cpSSR 표지자의 돌연변이율( $3.2-7.9 \times 10^{-5}$  mutation/allele/year)이 핵 게놈을 조성하고 있는 핵 DNA( $5-30 \times 10^{-9}$  mutation/allele/year) 보다 높은 것으로 알려져 있다 (Provan 등 1998). 핵 DNA에 존재하는 변이를 토대로 분석하는 I-SSR 표지자의 경우에도 정확한 돌연변이율이 계산되어진 보고는 없으나 primer 결합부위에 생겨난 핵 DNA 돌연변이( $5-30 \times 10^{-9}$  mutation/allele/year)에 의해서 새로운 I-SSR 변이체가 생성될 것이기 때문에 단순염기의 추가 또는 삭제에 의해서 새로운 변이체가 생성되는 cpSSR 표지자( $3.2-7.9 \times 10^{-5}$  mutation/allele/year) 보다는 여하한 경우에도 높지 않을 것으로 추정된다. Birky 등 (1989)이 세포소기관(엽록체와 미토콘드리아) 게놈을 대상으로 돌연변이율( $u$ ), 이주율( $m_e$ ), 유효집단의 크기( $N_{em}$ ), 집단의 수( $L$ ) 등을 고려하여 집단간 분화정도( $G_{ST}$ )를 계산하는 공식을 유도한 바 있다 :  $G_{ST} \approx 1/[1+2N_{em}\{L/(L-1)\}(m_e + u)]$ . 이 공식에서 집단의 수를 19개로 고정하고 부계유전하는 엽록체 게놈의 이주율은 핵 게놈과 동일하고 현재 집단들이 지리적으로 충분히 격리되어 있기 때문에 현존하는 당대의 집단 생성에 기여한 이주율은 0으로 가정하고, 2배체인 핵 DNA의 유효집단의 크기를 2배로 수정할 경우 핵 DNA에 대한 유전분화의 계산에 그대로 적용될 수 있다 :  $G_{ST} \approx 1/[1+4N_{em}\{19/(19-1)\}u]$ . 이 공식을 이용해서 소나무 19개 집단에 있어서 집단의 크기와 돌연변이율의 차이에 따른 유전분화 정도의 변화를 추정할 수 있다. 즉, 유효집단 크기가 2배 증가하면 집단간 유전분화 정도는 감소하는 경향이 있으나, 돌연변이율의 감소에 의해서 초래되는 유전분화의 심화정도는 유효집단의 크기가 2배 증가됨에 따라 야기되는 집단간 분화의 변화를

현 우도는 양상을 보인다. 따라서 비록 핵 DNA로부터 분석된 I-SSR 표지자를 대상으로 계산된 돌연변이율에 대한 정보가 없으나 일반적으로 핵 DNA 보다 돌연변이율이 훨씬 높은 것으로 알려진 cpSSR 표지자의 경우 I-SSR 표지자 분석을 통해서 계산된 집단간 유전분화 보다 낮은 수준으로 계산된 것으로 생각된다. 본 연구에서 관찰된 집단간 유전분화 정도( $\Phi_{ST} = 0.024$ )와 cpSSR 표지자의 돌연변이율( $3.2-7.9 \times 10^{-5}$  mutation/allele/year)의 최소치를 위의 공식에 대입해서 계산했을 때 전체 19개 집단에 대한 유효집단의 크기는 600,000개 정도로 추산되었다. 핵 DNA 인 I-SSR 표지자의 경우 전체 유효집단의 크기는 엽록체 게놈의 2배인 1,200,000이 되며 이때 돌연변이율이 1/10 정도만 감소되어도 그 분화정도는 4.5배 이상 증가함을 알 수 있었다. 비록 모의계산을 통해서 확인된 수치가 비평형상태의 결과로서 과대치를 초래할 수도 있으나 최소한 cpSSR 표지자의 높은 돌연변이율이 집단간 분화현상에 미치는 영향은 I-SSR 표지자에 비해서 절반에 불과한 집단의 크기로 인해서 야기될 수 있는 집단간 유전분화의 심화현상을 크게 상회한다는 것을 설명해 준다고 할 수 있다. 본 연구결과에 따르면 I-SSR 표지자 보다 높은 cpSSR 표지자의 돌연변이 발생빈도로 인해서 집단을 구성하고 있는 개체들이 대부분 서로 다른 haplotype으로 나타났다(평균 15.47개 개체 = 평균 14.32개 haplotype, Table 1), 각각의 haplotype을 조성하고 있는 6개 유전자좌에 다양한 cpSSR 변이체들이 포함되어 있었으며 별개의 haplotype의 경우에도 각각의 유전자좌에서 공유하고 있는 변이체들의 종류가 다양한 것으로 나타났다. 또한 비교 가능한 집단 쌍 중에서 77.5%의 집단간에는 동일한 haplotype이 존재하였기 때문에 비교 집단간에 존재하는 개체간 유전적 거리의 분산 보다는 이질적인 집단내 개체간에 계산되는 유전적 거리의 분산이 훨씬 크기 때문인 것으로 추정되며, 따라서 전체 개체간 유전적 거리의 분산 중 집단내 개체간의 유전적 거리의 분산이 차지하는 비율이 상대적으로 훨씬 크게 나타나기 때문에 집단크기가 절반에 불과한 cpSSR 표지자에서 예상과는 달리 상대적으로 낮은 수준의 집단간 유전분화 현상이 관찰된 것으로 생각된다.

이론적으로 현존하는 집단들이 안정적인 집단의 확산(이주) 및 정착 과정을 겪었다면 지리적으로 인접한 집단들이 지리적으로 먼 집단들 보다 유전적으로 동일 내지는 유사한 선조 집단으로부터 유래되었을 것이기에 이들 간에 유전적으로 가까운 현상을 보이게 될 것이다. 이를 검증하기 위해서는 비교 집단간에 계산된 유전적 거리를 토대로 UPGMA법 등의 유집분석법에 의해서 집단간 유전적 유연관계를 분석하고 이를 지리적 거리와의 비교를 통해서 확인해야한다. 그런데 소나무 19개 집단 총 299개체

**Table 4. Population pairs of sharing haplotypes. “0” denotes pairs without common haplotypes and “+” denotes pairs with common haplotypes.**

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 Hongcheon																			
2 Yeongwol	+																		
3 Uljin	+	+																	
4 Taean	+	0	+																
5 Mungyeon	0	+	+	+															
6 Boeun	+	0	+	+	0														
7 Buan	+	+	+	+	+	+													
8 Haman	+	0	+	0	0	+	+												
9 Hwacheon	+	+	+	+	+	+	+	+											
10 Hoengseong	+	+	-	0	0	0	+	+	+										
11 Jeongseon	+	+	+	+	+	+	+	+	+	+									
12 Bonghwa	+	0	+	+	+	+	+	+	+	+	+								
13 Pohang	+	+	+	0	+	+	+	+	+	0	+	+							
14 Goryeong	+	+	+	0	+	0	0	+	+	+	+	+	0						
15 Miryang	+	+	+	+	+	0	+	+	+	+	+	+	0	0	0				
16 Cheongyang	+	+	+	+	0	+	+	0	+	+	+	+	0	0	+				
17 Naju	+	+	0	+	+	+	+	0	+	+	+	+	+	0	+	+			
18 Jeongeup	+	0	-	0	0	+	+	+	+	+	+	+	+	0	+	+	+		
19 Hadong	0	0	-	0	+	+	+	+	+	+	+	+	+	+	0	+	0	+	

에 대한 cpSSR 표지자 분석 결과 167개의 haplotype이 확인되었다. 즉 각각의 haplotype당 1.79개의 개체만이 동일한 haplotype을 지닐 가능성이 있기 때문에 전체 19개 집단에 동일한 haplotype들이 고르게 분포될 가능성이 적다. 실제로 전체 비교 가능한 173개 집단 쌍 간에 동일한 haplotype이 전혀 존재하지 않는 집단 쌍이 39쌍으로 나타나(Table 4) 관찰된 haplotype 빈도에 근거해서 이들 39쌍 간의 비교가 가능한 유전적 거리를 계산해 낼 수 없을 뿐만 아니라, 이들로 인하여 집단간에 최소한 1개 이상의 haplotype을 공유하고 있는 나머지 134쌍을 포함한 19개 집단 상호간의 유전적 유연관계를 계산하는데 오류를 유발시키게 되기 때문에 집단간의 유집분석을 수행할 수 없었다. 이와 같이 현존하는 집단의 분포양상은 이론상 과거 공동의 임의교배 조상집단(common panmictic ancestral population)으로부터 분화되었을 것임에도 불구하고 각 집단별로 독특한 haplotype이 다수 존재하는 이유는 simple sequence repeats을 이루고 있는 主要素(motif)가 주로 단일염기인 cpSSR 표지자의 특성상 돌연변이 발생빈도가 높고 자연선택의 직접적인 대상이 아니기 때문에 돌연변이 발생의 방향성(주요소가 일방적으로 추가 또는 삭제되는 일관성)이 무작위로 6개 유전자좌에서 발생하기 때문일 것으로 추정된다. 또한 분석된 19개 집단간에 유전적 교류가 빈번하게 일어나지 못해서 야기된 cpSSR haplotype 수준에서의 유전적 격리현상이 초래한 결과일 가능성을 배제할 수 없다. 특이한 점은 울진집단의 경우 나주집단

을 제외한 모든 집단 쌍들과 최소한 1개 이상의 haplotype을 공유하고 있는 것으로 나타났으며, 봉화집단은 밀양과 영월집단을 제외한 모든 집단 쌍들과 최소한 1개 이상의 haplotype을 공유하고 있는 것으로 나타나 이들이 비교적 원조 집단에 가까운 집단일 가능성이 있는 것으로 추정된다. 한편 고령집단의 경우 비교 가능한 18개 집단 쌍 중에 8쌍 간에 동일한 haplotype이 전혀 존재하지 않는 것으로 나타나 19개 집단 중에서 가장 유전적으로 이질적인 분화가 많이 된 집단인 것으로 추정된다.

본 연구에서 분석된 cpSSR 표지자의 경우 과도한 유전 변이를 보이기 때문에 이들을 이용해서 집단간 분화정도를 평가할 때에는 시료선정 과정에서 발생할 수 있는 오류를 미연에 방지하기 위해서 분석시료의 수를 충분히 확보하여 각 집단에 존재하는 모든 변이체가 고르게 시료에 포함 되도록 해야 할 것으로 생각된다. 또한 본 연구를 통해서 관찰된 결과에 비추어 볼 때 향후 육종목적으로 국내 소나무 집단에 대한 유전구역을 설정하는 기준으로 지리적 위치와 더불어 각 집단간 유전적 유연관계도 함께 고려되어야 할 것으로 생각된다.

**인용문헌**

1. 김영중, 송정호, 조경진, 김용울, 구영분. 2002. 準人工交配에 의한 리기다×테다 소나무 잡종종자 大量生産과 雌花芽의 生長特性. 한국육종학회지 34(3): 228-235.

2. 노은운, 이재순, 최영입, 한무석, 한상억, 이효신, 이용섭. 2005. 잣나무의 엽록체 유전체. 국립산림과학원 연구자료 251호. pp. 74-139.
3. Anzidei, M., A. Madaghie, C. Sperisen, B. Ziegenhagen, and G.G. Vendramin. 1999. Chloroplast microsatellites for analysis of the geographic distribution of diversity in conifer species. In: Gillet, E.M. (ed.). *Which DNA Marker for Which Purpose?* Final Compendium of the Research Project Development, optimization and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity. URL <http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm>
4. Birky, C.W.Jr., F. Paul, and M. Takeo. 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.
5. Black, W. C., IV. 1996. RAPDDIST 1.0. Department of Microbiology, Colorado State University, Fort Collins, CO. USA.
6. Excoffier, L., P. Smouse, and J. Quattro. 1992. Analysis of molecular variance inferred from metric distance among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.
7. Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics. University of Washington. Seattle, WA. USA.
8. Hong, Y.-P., H.-Y. Kwon, K.-S. Kim, K.-N. Hong, and Y.-Y. Kim. 2004. Discordance between geographical distribution and genetic relationship among populations of Japanese red pine in Korea revealed by analysis of I-SSR markers. *Silvae Genetica* 53(3): 89-92.
9. Hong, Y.-P., H.-Y. Kwon, and I.-S. Kim. 2006. I-SSR markers revealed inconsistent phylogeographic patterns among populations of Japanese red pine in Korea. *Silvae Genetica* (인쇄중).
10. Nei, M. 1987. *Molecular evolutionary genetics*. Columbia University Press, New York.
11. Parducci, L., A.E. Schmidts, A. Madaghi, M. Anzidei, and G.G. Vendramin. 2001. Genetic variation at chloroplast microsatellite (cpSSRs) in *Abies nebrodensis* (Lojac.) Mattei and three neighboring *Abies* species *Theoretical Applies Genetics* 102: 733-740.
12. Provan, J., N. Soranzo, N.J. Wilson, J.W.McNicol, G.I. Forrest, J. Cottrell, and W. Powell. 1998. Gene pool variation in Caledonian and European Scots pine (*Pinus sylvestris* L.) revealed by chloroplast simple-sequence repeats. *Proceedings of the Royal Society of London. Series B.* 265: 1697-1705.
13. Provan, J., N. Soranzo, N.J. Wilson, D.B. Goldstein, and W. Powell. 1999. A low mutation rate for chloroplast microsatellites. *Genetics* 153: 943-947.
14. Richardson, B.A., S.J. Brunfeld, and N.B. Klopfenstein. 2002. DNA from birds-dispersed seeds and wind-disseminated pollen provides insights into postglacial colonization and population genetic structure of whitebark pine (*Pinus albicaulis*). *Molecular Ecology* 11: 215-227.
15. Ribeiro, M.M., C. Plemion R. Petit, G.G. Vendramin, and A.E. Schmidts. 2001. Variation in chloroplast single-sequence repeats in Portuguese maritime pine (*Pinus pinaster* Ait.). *Theoretical Applies Genetics* 102: 97-103.
16. Shannon, C.E. 1948. A mathematical theory of communication. *Bell System Tech. J.* 27: 379-423.
17. Schneider, S., D. Roessli and L. Excoffier. 2000. Arlequin V2.000. A software for population genetics data analysis. Dept. of Anthropology and Ecology, University of Geneva, Geneva, Switzerland.
18. Slatkin, M. 1987. Gene flow and the Geographic structure of natural populations. *Science* 236: 787-792.
19. Tautz, D. 1993. Notes on the definition and nomenclature of tandemly repetitive DNA sequences. In: Pena S.D.J., R. Chakraborty, J.T. Epplen, and A.J. Jeffreys (eds.). *DNA Fingerprinting: State of the Science*. Birkhauser Verlag, Basel, Switzerland, pp. 21-28.
20. Vendramin G.G., L. Lelli, P. Rossi, and M. Morgante. 1996. A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*. *Mol. Ecol.* 5: 111-114.
21. Wakasugi T., J. Tsudzuki, S. Ito, M. Shibata, and M. Sugiura. 1994. A physical map and clone bank of the black pine (*Pinus thunbergii*) chloroplast genome. *Plant Mol. Biol. Rep.* 12: 227-241.
22. Yeh, F.C., R.C. Yang and T. Boyle. 1999. POPGENE v 1.31. Microsoft window-based freeware for population genetic analysis. Dept. of Renewable Resources. Univ. of Chicago Press, Chicago, USA.

---

(2006년 5월 9일 접수; 2006년 7월 10일 채택)