



## 차세대염기서열분석법을 이용한 잔대의 SSR 마커 개발

박기찬\* · 김영국\* · 황보경\*\* · 길진수\*\* · 정 희\*\* · 박신기\*\*\* · 홍창표\*\*\* · 이 이\*\*†

\*농촌진흥청 국립원예특작과학원 인삼특작부, \*\*충북대학교 특용식물학과, \*\*\*㈜테라젠이텍스 바이오연구소

### Development of Simple Sequence Repeat Markers from *Adenophora triphylla* var. *japonica* (Regel) H. Hara using Next Generation Sequencing

Ki Chan Park\*, Young Guk Kim\*, Kyeong Hwangbo\*\*, Jinsu Gil\*\*,  
Hee Chung\*\*, Sin Gi Park\*\*\*, Chang Pyo Hong\*\*\* and Yi Lee\*\*†

\*Department of Herbal Crop Research, NIHHS, RDA, Eumseong 27709, Korea.

\*\*Department of Industrial Plant Science and Technology, Chungbuk National University, Cheongju 28644, Korea.

\*\*\*Theragen Etex Bio Institute, Suwon 16229, Korea.

#### ABSTRACT

**Background:** *Adenophora triphylla* var. *japonica* (Regel) H. Hara shows vegetative growth with radical leaves during the first year and shows reproductive growth with cauline leaves and bolting during the second year. In addition, the shape of the plant varies within the same species. For this reason, there are limitations to classifying the species by visual examination. However, there is not sufficient genetic information or molecular tools to analyze the genetic diversity of the plant.

**Methods and Results:** Approximately 34.59 Gbp of raw data containing 342,487,502 reads was obtained from next generation sequencing (NGS) and these reads were assembled into 357,211 scaffolds. A total of 84,106 simple sequence repeat (SSR) regions were identified and 14,133 primer sets were designed. From the designed primer sets, 95 were randomly selected and were applied to the genomic DNA which was extracted from five plants and pooled. Thirty-nine primer sets showing more than two bands were finally selected as SSR markers, and were used for the genetic relationship analysis.

**Conclusions:** The 39 novel SSR markers developed in this study could be used for the genetic diversity analysis, variety identification, new variety development and molecular breeding of *A. triphylla*.

**Key Words:** *Adenophora triphylla*, Genetic Diversity, Molecular Breeding, Next Generation Sequencing, SSR Marker

#### 서 언

잔대 [*Adenophora triphylla* var. *japonica* (Regel) H. Hara]는 초롱꽃과에 속하는 타식성 식물이며 전 세계적으로 약 60 여 종이 분포하지만 주로 아시아권에서 자생하며 그 중 약 30 여 종이 우리나라에 자생하고 있는 것으로 알려져 있다 (Lee *et al.*, 2006). 잔대는 같은 종이라 하더라도 개체 간에 다양한 형태적 특성을 나타내고 그 유전형질 또한 다양하다고 보고되었다 (Lee *et al.*, 1988; Lee, 1989). 예로부터 알려진

약용식물인 잔대 (사삼)의 효능은 주로 폐 기능에 도움을 주며, 강한 해독능력이 있다고 전해지고 있다 (Shin and Seo, 2000).

현재까지 약용식물에 대한 연구는 물질분석을 통한 이차대 사산물의 이용과 같은 성분내 대한 연구와 DNA barcode기법을 통한 종판별이나 randomly amplified polymorphic DNA (RAPD)를 이용한 유연관계 분석 등의 연구가 주로 이루어졌다. 최근 들어 한약재의 기원이 되는 식물에 대해 분자수준에서의 분류학적 연구가 활발해지고 있으며 유전정보의 중요

†Corresponding author: (Phone) +82-43-261-3373 (E-mail) leeyi22@hanmail.net

Received 2017 September 5 / 1st Revised 2017 September 29 / 2nd Revised 2017 October 20 / 3rd Revised 2017 October 30 / Accepted 2017 October 31

This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

성이 대두되면서 next generation sequencing (NGS)을 통한 유전체 분석 및 유전적 다양성분석과 이를 기초로 하는 분자 육종의 필요성이 대두되고 있다 (Eu *et al.*, 2009; Lee *et al.*, 2010).

분자마커는 생물마다 가지고 있는 DNA의 염기서열 정보에 기반하여 중합효소연쇄반응 (polymerase chain reaction, PCR) 결과물의 길이 차이를 비교하거나 염기서열의 차이를 이용하여 종 혹은 집단을 비교 분석하는 방법이다. 좋은 마커의 요건으로는 다형성을 가지고 있으며 공우성적으로 검출할 수 있고, 대상으로 하는 식물의 유전체 전체에 고르게 분포하며 변이를 탐색하는데 있어 저렴한 비용으로 짧은 시간에 분석이 가능할 뿐 아니라 재현성이 높게 나타나는 것이 좋다.

분자마커의 대표적인 종류로는 RAPD, restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), single nucleotide polymorphism (SNP), simple sequence repeat (SSR) 등이 있다. 이 중에서 SSR마커는 식물체의 유전체에 존재하는 단순반복 염기서열의 반복횟수의 차이를 이용하는 마커이며 다른 마커보다 다형성 정도가 높고 공우성을 나타내기 때문에 검정 계통의 이형접합도 여부를 쉽게 확인할 수 있는 등의 장점이 있다 (McCouch *et al.*, 1997; Coburn *et al.*, 2002). 실제로 국제식물신품종보호 연맹 (International Union for the Protection of New Varieties of Plants, UPOV) 내의 분자생물학실 무작업반 (Working Group on Biochemical and Molecular Techniques and DNA-Profiling in Particular, BMT) 내에서도 품종식별에 SSR마커가 가장 효과적임을 제안한 바 있다 (UPOV-BMT, 2002).

최근 들어 한약재의 기원이 되는 식물 종을 대상으로 분자 수준에서 식물 분류학적 측면의 연구가 많이 이루어지고 있으며 (Eu *et al.*, 2009; Lee *et al.*, 2010), 종 다양성 및 종간 계통 분류학적 유연관계 분석을 위해 특정한 유전자 부위의 염기서열이나 유전체를 이용한 DNA fingerprinting을 이용한 연구가 다수 보고되었다 (Kress *et al.*, 2005; Weising *et al.*, 2005).

최근 NGS 기술의 발전은 대량의 마커 발굴을 가능하게 하였으며 (Mardis, 2008; Wang *et al.*, 2009), 그 결과 옥수수, 밀, 벼 등 주요 식량작물이나 채소, 화훼에서는 육종기술의 발달과 마커의 분자 표지 개발에 관한 연구가 활발히 이루어졌다 (Nybom, 2004; Henry, 2012; Kilian and Graner, 2012). 그러나, 약용작물의 종내 다양성에 대한 연구는 더덕 (Kim *et al.*, 2016)과 같은 주요 약용작물을 제외하면 분자수준의 연구가 매우 미미한 실정이다.

본 연구에서는 국내의 여러 지역에 걸쳐 자생하며 재배, 유통되고 있는 잔대를 대상으로 NGS를 통한 SSR마커를 개발하고 이를 이용하여 잔대의 유전적 다양성을 분석하여 향후 육종에 필요한 연구 기반을 마련하고자 수행하였다.

## 재료 및 방법

### 1. SSR 구간 탐색 및 프라이머 제작

본 실험의 NGS 분석에는 강원도 평창에서 수집된 유전자원 PC001에 속하는 한 개체를 임의로 선택하였으며, SSR마커의 효율성 검정과 유전다양성 분석을 위한 재료는 평창, 보은, 임실, 진안, 나주, 곡성, 화순 등 7 개 지역에서 뿌리를 수집 후 농촌진흥청 국립원예특작과학원 인삼특작부에서 증식한 총 23 유전자원을 사용하였다 (Table 1).

NGS는 Illumina Hiseq 2500 플랫폼 (Illumina, San Diego, CA, USA)을 이용하여 수행하였으며, 이를 통해 얻은 핵산조각을 Gil 등 (2017)의 방법에 따라 SOAPdenovo2 version 2.04의 SOAPec을 이용하여 error를 보정한 후 SOAPdenovo2를 이용하여 조합하였다. Microsatellites, transposable elements, rDNAs 등과 같은 반복서열은 RepeatMasker version 4.0.5과 RepeatModeler version 1.0.8 (<http://www.repeatmasker.org/>)을 이용하여 분석하였다. 또한 SSR Finder (<ftp://ftp.gramene.org/pub/gramene/arch-ives/software/scripts/ssr.pl>)를 사용하여 SSR 구간을 탐색하였다. SSR 구간을 탐색하기 위해 2 가지의 변수를

**Table 1.** List of *Adenophora triphylla* genetic resources used in this study.

No.	Collection ID	Collection region	Collection year
1	PC001 [NGS <sup>1)</sup>	Gangwon-do, Pyeongchang-gun	2015
2	PC002	Gangwon-do, Pyeongchang-gun	2015
3	PC003	Gangwon-do, Pyeongchang-gun	2015
4	PC004	Gangwon-do, Pyeongchang-gun	2015
5	BE001	Chungcheongbuk-do, Boeun-gun	2015
6	BE002	Chungcheongbuk-do, Boeun-gun	2015
7	BE003	Chungcheongbuk-do, Boeun-gun	2015
8	IS006	Jeollabuk-do, Imsil-gun	2016
9	IS009	Jeollabuk-do, Imsil-gun	2016
10	IS012	Jeollabuk-do, Imsil-gun	2016
11	IS016	Jeollabuk-do, Imsil-gun	2016
12	JA001	Jeollabuk-do, Jinan-gun	2015
13	JA002	Jeollabuk-do, Jinan-gun	2015
14	JA003	Jeollabuk-do, Jinan-gun	2015
15	NJ001	Jeollanam-do, Naju-si	2016
16	NJ002	Jeollanam-do, Naju-si	2016
17	NJ003	Jeollanam-do, Naju-si	2016
18	NJ004	Jeollanam-do, Naju-si	2016
19	GS002	Jeollanam-do, Gokseong-gun	2015
20	GS003	Jeollanam-do, Gokseong-gun	2015
21	HS001	Jeollanam-do, Hwasun-gun	2016
22	HS007	Jeollanam-do, Hwasun-gun	2016
23	HS008	Jeollanam-do, Hwasun-gun	2016

<sup>1)</sup>NGS; Used for next generation sequencing.

고려하였다. 첫 번째는 2 개에서 6 개의 염기쌍 (base pairs) 이 직렬로 최소 5 회 이상 반복되는 염기서열을 우선 선택하였고, 두 번째는 반복되는 motif에 변이가 있는 것은 제외하였다. Primer3 program (Untergasser *et al.*, 2012)을 이용하여 SSR 프라이머 서열을 다음과 같은 조건에서 디자인하였다. 길이는 18 bp에서 26 bp (최적 길이 23 bp)이고 GC함량은 50%에서 70% (최적 60%)로 하였으며, Tm값은 55°C에서 62°C (최적 58°C)로 설정하였다. PCR product의 길이는 150 bp에서 250 bp 사이가 되게 설계하되, 프라이머쌍 사이에 서로 다른 motif의 반복구간이 들어간 것들은 모두 제외하였다.

### 2. 다형성 SSR 프라이머 선발

본 연구에서 수집된 유전자원 중 5 개 지역 (평창, 보은, 나주, 임실, 화순)에서 각각 1 개체씩을 임의로 선별하고 genomic DNA를 추출한 후, 이를 동량으로 혼합하는 genomic DNA pooling 방법을 이용하여 혼합된 genomic DNA를 95 개의 프라이머에 적용하였다. 그 후, Fragment Analyzer Automated CE system (Advanced Analytical Technologies, Ankeny, IA, USA) 기기를 이용하여 Quant-iT PicoGreen dsDNA reagent kit, 1-500 bp (Invitrogen Co., Carlsbad, CA, USA)로 전기영동을 수행하였다. 이후, 증폭된 밴드는 PROSize version 2.0 software (Advanced Analytical Technologies, Ankeny, IA, USA) 프로그램을 이용하여 분석하였다. 이를 바탕으로, 결과 값이 뚜렷하고, PCR product의 증폭 길이가 예상되었던 수치와 비슷하며, 또한 2 개 이상의 밴드를 보이는 프라이머를 선별하였다.

### 3. SSR 마커의 효율성 검증

선발된 마커를 수집된 자원에 적용하여 개발된 마커의 효율성을 검증하였다. 이를 위해서 선발된 프라이머를 총 23 개체에 각각 적용하였고, 증폭된 PCR 산물을 DNA Fragment Analyzer Automated CE System (Advanced Analytical Technologies, Ankeny, IA, USA)을 이용하여 분석하였다. 그 결과를 마커 분석 프로그램인 Power Marker software (Version 3.23) (Botstein *et al.*, 1980; Liu and Muse 2005)를 이용하여 선발된 SSR 마커의 효율성을 분석하였다.

## 결과 및 고찰

### 1. 염기서열 분석 및 SSR 구간 탐색과 프라이머 제작

본 연구에서는 Illumina Hiseq 2500 platform을 이용하여 염기서열 분석을 수행하였다. Paired-end sequencing 방법을 이용하여 총 342,487,502 개의 read를 포함하는 총 34.59 Gbp의 염기서열 정보를 생산하였고, GC의 비율은 41.37%였다. 생산된 염기서열로부터 총 357,211 개의 scaffold를 얻

**Table 2.** Details of the repeat search from the *Adenophora triphylla* PC001 whole genome assembly.

Type	Number of elements	Length (bp)	In genome (%)
Retrotransposons	271,852	51,826,041	11.38
SINEs	9,922	1,139,203	0.25
LINEs	28,230	5,370,052	1.18
LTR elements	233,700	45,320,139	9.95
DNA transposons	123,770	16,890,297	3.71
DNA	120,710	16,345,389	3.59
RC	3,060	561,985	0.12
Unclassified	772,971	100,313,631	22.03
Simple sequence repeats	84,106	3,316,165	0.73
Low complexity	15,057	720,565	0.16
Total		170,626,675	37.47

SINEs; Short interspersed nuclear elements, LINEs; Long interspersed nuclear elements, LTR; Long terminal repeat, RC; Rolling-circle.

**Table 3.** SSR motif types of *Adenophora triphylla*.

Repeat type	No. of searched SSR	In SSR (%)	No. of primer designed SSR	In SSR (%)	No. of tested SSR
Di	70,497	84.58	11,785	14.14	NT
Tri	9,998	11.99	1,899	2.28	31
Tetra	1,263	1.52	222	0.27	40
Penta	752	0.90	117	0.14	24
Hexa	757	0.91	102	0.12	NT
Hepta	87	0.10	8	0.01	NT
Total	83,354	100	14,133	16.96	95

었고, N50은 1,459 bp였다.

이렇게 얻어진 scaffold로부터 SSR 구간을 검색하고 primer를 제작하였다. 검색된 SSR구간은 총 84,106 개로 assembly된 전체 염기서열의 0.73%를 차지하였으며, 그 중 di-nucleotide가 70,497 개, tri-nucleotide가 9,998 개, tetra-nucleotide가 1,263 개, penta-nucleotide가 752 개, hexa-nucleotide가 757 개, hepta-nucleotide가 87 개를 차지하였다 (Table 2, Table 3).

탐색된 SSR 구간에서 서로 다른 SSR motif가 없도록 14,133 개의 SSR 프라이머를 설계하였으며, 이 중에서 임의로 tri-nucleotide 31 개, tetra-nucleotide 40 개, penta-nucleotide 24 개를 포함한 총 95 개의 프라이머 세트를 제작하였다.

### 2. 다형성을 보이는 SSR 프라이머 선발

제작된 95 개의 프라이머 세트 중 5 개체를 이용하여 다형성을 조사한 결과 총 39 개의 프라이머 세트가 선발되었다 (Table 4). 선발된 39 개의 프라이머 세트를 이용하여 23 개

**Table 4.** Characteristics of the SSR markers developed from *Adenophora triphylla*.

No.	Primer ID	SSR motif	T <sub>m</sub> (°C)	Allele size range (bp)	Primer sequences (5'-3')	GenBank accession No.
1	Atri00004	AAAC <sub>(6)</sub>	59	149-229	F :AGCAGAGTTCGTACAGGTTCC R :GGGGCTCAAATACCACTTGT	KY906103
2	Atri00007	AAAG <sub>(5)</sub>	59	181-193	F :CGTCCGCCGATACTAGTTACA R :GACATGACATCAGTCCACG	KY906104
3	Atri00011	AAAG <sub>(8)</sub>	59	179-199	F :CTCACATCTTACGAGCCTCTCT R :GGTTGCTTTCTCCTGGAATG	KY906105
4	Atri00021	AAAT <sub>(5)</sub>	58	200-212	F :GGGCATAGCTCTCTGCTTAGT R :GACTGAAGATCTGAAGGTCACC	KY906106
5	Atri00048	ACAG <sub>(5)</sub>	58	168-204	F :GCATACTAGGCCACTGATTCC R :GCAGTCTCCACTTCTTATAGGC	KY906107
6	Atri00051	ACCT <sub>(5)</sub>	59	168-184	F :GTACCCTCTGCCAAGATAACG R :ATGAGTCCAATAGCCCTAGCTC	KY906108
7	Atri00079	ATTT <sub>(5)</sub>	58	146-162	F :GGCAGAGTATTAGGTGAAGTGG R :TGTGAGGAGAGACAGAGGTAGTG	KY906109
8	Atri00095	CCAG <sub>(5)</sub>	57	132-164	F :AACCCTGAGCTTCTTCGAG R :CTTCTCAGCTAGGTTCTCAGC	KY906110
9	Atri00108	CTCA <sub>(5)</sub>	58	200-232	F :CATTGATGAGCTCTCGAACC R :CAATGCCTGTAGTTTGGTCC	KY906111
10	Atri00111	CTCC <sub>(5)</sub>	59	167-199	F :GCTAAGGCGCAGGTATTATGAC R :GATGGTGCATGTGACTCAAG	KY906112
11	Atri00146	GCTA <sub>(5)</sub>	59	179-203	F :CATAGCTGAGCTGGTGAACAGT R :ATTCATCTCCCCACCTTCAG	KY906113
12	Atri00171	TATG <sub>(7)</sub>	58	166-190	F :CCATCCATTGCGTAGCTG R :CCATTGGTACGTACATCGTG	KY906114
13	Atri00172	TATG <sub>(5)</sub>	58	161-177	F :CCAGAGAACGCAAAGAAGTCTG R :CTCCTTACGATGCTCTTCT	KY906115
14	Atri00174	TATT <sub>(5)</sub>	58	185-213	F :GGCCATAATGATCTGGGAG R :CGACAGTCGTGAAAAGAGGA	KY906116
15	Atri00186	TCTT <sub>(6)</sub>	59	166-186	F :CCACCAAGGGACTAAATCCT R :GAACAGAAGCTCTGTAGCGAGA	KY906117
16	Atri00200	TTCC <sub>(5)</sub>	59	151-171	F :GGTGGCCTTGCAAATATCTC R :AATATGATCGGGGGCAAC	KY906118
17	Atri00206	TTCT <sub>(5)</sub>	58	169-173	F :ATTGAGGCCTCTGAGACAGAC R :CGCAAGACTACCCAGAAACTAC	KY906119
18	Atri00223	AAAAG <sub>(5)</sub>	58	192-208	F :CGATCACTGGCTTAACCTTG R :GGCTGCCACATAAATCGTAG	KY906120
19	Atri00232	AAAGA <sub>(5)</sub>	58	153-208	F :TAAGTATGCACCAGGTTGCC R :ACCCTCTTCTGCATCACTTG	KY906121
20	Atri00237	AAATT <sub>(5)</sub>	58	183-203	F :CCAACCTCCTTCTTCTTCTCCT R :GGTTTCGAGGAGAGGGAAT	KY906122
21	Atri00245	ACCAG <sub>(5)</sub>	59	190-230	F :CTTTTCTTTGAGTGGCGG R :GCCATGAGATAAGATCTCGACC	KY906123
22	Atri00246	ACTCA <sub>(7)</sub>	59	190-210	F :GTGAACACGAAGGGAGGTTT R :TGCCTAGTGTATGAGGCTTCC	KY906124
23	Atri00254	AGGTG <sub>(7)</sub>	58	164-204	F :GGCCCTTGACTAAAGGTATAGG R :AGAGGATCAAGGCACATGG	KY906125
24	Atri00276	CCTCA <sub>(5)</sub>	60	185-215	F :CCCTCTTCTAATTCCCTCAGGT R :CCAAGAAATAACAGCCAGGG	KY906126

Table 4. continued.

No.	Primer ID	SSR motif	Tm (°C)	Allele size range (bp)	Primer sequences (5'-3')	GenBank accession No.
25	Atri00285	CTTGT <sub>(5)</sub>	57	163-183	F :GGAGTGACTAGGATACCTGTTG R :CAGGGACTCTGGGATTATACAG	KY906127
26	Atri00287	CTTTT <sub>(5)</sub>	59	176-206	F :GATGGCGCTACAGCTTATCTAC R :ACAAAAAGGGACACAGGCAC	KY906128
27	Atri00326	TGTAC <sub>(6)</sub>	59	191-211	F :TGGAGGTTGATGATTGCC R :GACATAGTGAGCAACAACCTGGC	KY906129
28	Atri00338	TTTGG <sub>(6)</sub>	57	199-219	F :GCTGCAGCAATGCAGTAT R :AGTTAGAGAAGAGCGCCAGTC	KY906130
29	Atri00456	AAC <sub>(8)</sub>	59	152-161	F :CTAAGCACTCGCTCAAGTACCTC R :GTGAATCCAGAAAGAGTCCG	KY906131
30	Atri00543	AAT <sub>(9)</sub>	59	181-193	F :CAGTGCTTCAATAGACCGTCAC R :GACCATGCCAAGGTTGAA	KY906132
31	Atri00855	ATC <sub>(7)</sub>	58	182-197	F :GAGATAAGTCTTGATGCCGC R :CAATCCACCATTCTCAGCAG	KY906133
32	Atri00959	CAA <sub>(7)</sub>	58	176-251	F :ACTATAGAAAGGCCCGCT R :GGCAAGTGATTGTGCAGAC	KY906134
33	Atri01120	CCA <sub>(7)</sub>	58	148-175	F :AACTCGTCCTCACTCAAGGTAG R :CTTTAGGTTGAGGCTAAGAGGG	KY906135
34	Atri01347	CTT <sub>(8)</sub>	59	197-248	F :CCCTAAGAACCCTCCCTTCTACA R :CTATCGTGACTCAGCCTTTGAC	KY906136
35	Atri01349	CTT <sub>(7)</sub>	58	180-195	F :GTCTGGTTACCTCCACCATCT R :CCTGGCAGGTATCCTGTACATA	KY906137
36	Atri01490	GAG <sub>(7)</sub>	59	180-198	F :GAAACCTGTGAGTTGACGACTG R :CATTGAGACGCGTACGAAG	KY906138
37	Atri01559	GCA <sub>(7)</sub>	59	188-194	F :CTCATTGAGGAGGTCAGTC R :TCGAGTTGGACTATTGCCTG	KY906139
38	Atri01585	GCT <sub>(8)</sub>	58	180-201	F :GAGTAGCACCTGAAGAGGAAGA R :AAGTAGTAGCAGGCAACAGGTG	KY906140
39	Atri01627	GGC <sub>(7)</sub>	58	171-186	F :GATTTCCGCTCCCAATTC R :ACTCAAACCCACGTTTCG	KY906141

의 자원에서 유전적 다양성을 분석한 결과 allele의 수가 2 개에서 12 개로 나타나 명확한 다형성을 보였으며 (Table 5), 이 SSR 프라이머 세트는 잔대 [*Adenophora triphylla* var. *japonica* (Regel) H. Hara]의 SSR 마커로 활용하기에 매우 적합하다고 판단되었다. 이들 마커의 motif 종류는 tri motif 11 개, tetra motif 17 개, penta motif 11 개였으며, 결과적으로 motif 별로 고르게 마커가 개발되었음을 알 수 있었다.

### 3. SSR 마커의 효율성 검증

개발된 마커의 효율성을 검증하기 위하여 23 개체에 적용하여 얻은 총 263 개의 allele을 분석하였다. Frequency of major alleles ( $M_{AF}$ )은 0.17 - 0.87 (평균 0.46) 이었고, genotype number ( $G_N$ )는 의 값은 2 - 18 (평균 9.9)였으며, heterozygosity ( $H_E$ )는 0.23 - 0.88 (평균 0.68)이었다. Observed heterozygosity

( $H_O$ )의 값은 0.24 - 1.00 (평균 0.59)였으며, polymorphism information content (PIC)의 값은 0.20 - 0.87 (평균 0.64)이었다. PIC 값이 가장 컸던 마커는 Atri00959였으며, 이 마커는 PIC값과  $H_E$  뿐만 아니라  $G_N$ 도 높게 나타났다 (Table 5). 위 결과를 바탕으로 적용된 39 개의 마커는 잔대의 유전자원 분석에 적합하다고 사료되었다.

39 개의 SSR마커를 이용하여 얻은 결과를 바탕으로 국내 각 지역에서 수집된 잔대 유전자원 23 개에서 유전자형을 분석하여 계통학적 분석을 한 결과 수집지역별로 그룹이 만들어지지 않고 섞여서 만들어지는 것을 알 수 있었다 (Fig. 1). 또한, 한 지역에서 수집된 자원이라도 다시 세부 그룹으로 나뉘어져, 잔대는 유전적으로 다양하다는 것을 잔대 유전체 정보를 바탕으로 개발된 SSR 마커를 이용하여 확인할 수 있었다. 또한, 개발된 마커의 평균 PIC 값이 0.64인 것으로 볼 때 본

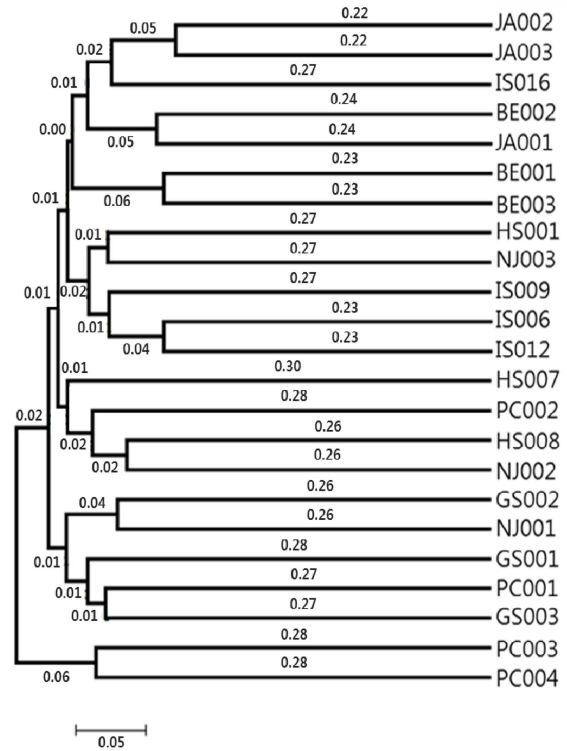
**Table 5.** Genetic diversity measures for 39 polymorphic SSR loci using 23 *Adenophora triphylla* accessions.

No.	Marker	$M_{AF}$	$G_N$	$A_N$	$H_E$	$H_O$	PIC
1	Atri00004	0.65	7	6	0.53	0.65	0.50
2	Atri00007	0.65	4	4	0.51	0.65	0.45
3	Atri00011	0.47	9	7	0.69	0.53	0.65
4	Atri00021	0.24	17	9	0.85	0.91	0.84
5	Atri00048	0.59	6	4	0.59	0.41	0.54
6	Atri00051	0.61	5	5	0.57	0.74	0.51
7	Atri00079	0.50	6	4	0.63	0.56	0.57
8	Atri00095	0.33	14	9	0.81	0.87	0.78
9	Atri00108	0.33	15	8	0.79	0.65	0.76
10	Atri00111	0.39	11	9	0.72	0.83	0.67
11	Atri00146	0.28	12	7	0.80	1.00	0.77
12	Atri00171	0.33	11	7	0.79	0.86	0.76
13	Atri00172	0.35	9	5	0.76	0.35	0.72
14	Atri00174	0.48	9	7	0.68	0.45	0.64
15	Atri00186	0.61	6	5	0.57	0.26	0.53
16	Atri00200	0.59	9	6	0.61	0.52	0.58
17	Atri00206	0.87	2	2	0.23	0.26	0.20
18	Atri00223	0.35	10	7	0.76	0.25	0.73
19	Atri00232	0.41	13	10	0.73	0.74	0.70
20	Atri00237	0.52	8	5	0.67	0.74	0.63
21	Atri00245	0.35	14	9	0.75	0.61	0.71
22	Atri00246	0.57	10	5	0.62	0.44	0.58
23	Atri00254	0.35	10	7	0.74	0.65	0.69
24	Atri00276	0.52	12	7	0.69	0.61	0.66
25	Atri00285	0.35	10	5	0.72	0.78	0.67
26	Atri00287	0.64	9	6	0.56	0.36	0.53
27	Atri00326	0.65	9	6	0.54	0.44	0.51
28	Atri00338	0.44	8	5	0.67	0.26	0.61
29	Atri00456	0.52	6	4	0.58	0.57	0.50
30	Atri00543	0.39	8	5	0.71	0.52	0.65
31	Atri00855	0.50	10	7	0.68	0.65	0.65
32	Atri00959	0.17	18	12	0.88	0.74	0.87
33	Atri01120	0.23	18	10	0.85	0.77	0.84
34	Atri01347	0.37	10	8	0.78	0.52	0.75
35	Atri01349	0.39	11	8	0.75	0.52	0.71
36	Atri01490	0.30	12	8	0.80	0.78	0.77
37	Atri01559	0.67	5	3	0.49	0.48	0.43
38	Atri01585	0.48	9	7	0.67	0.24	0.62
39	Atri01627	0.60	8	6	0.60	0.57	0.57
	Mean	0.46	9.9	6.6	0.68	0.59	0.64

$M_{AF}$ ; Major allele frequency,  $G_N$ ; Genotype number,  $A_N$ ; Number of alleles,  $H_E$ ; Expected heterozygosity,  $H_O$ ; Observed heterozygosity, PIC; Polymorphic information content.

연구에서 개발된 잔대의 SSR 마커는 유전자형의 다양성을 확인하는데 우수한 마커임을 알 수 있었다.

지금까지 잔대 (사삼)에서는 항산화와 비만 등에 관련된 성



**Fig. 1.** Phylogenetic tree for 23 *Adenophora triphylla* collections using 39 SSR markers. PC; Gangwon-do Pyeongchang-gun 2015, BE; Chungcheongbuk-do Boeun-gun 2015, IS; Jeollabuk-do Imsil-gun 2016, JA; Jeollabuk-do Jinan-gun 2015, NJ; Jeollanam-do Naju-si 2016, GS; Jeollanam-do Gokseong-gun 2015, HS; Jeollanam-do Hwasung-gun 2016. The genetic distances between strains were calculated based on the shared allelic methods (Jin and Chakraborty, 1994) using PowerMarker v3.25 software. The values on each branch indicate the length of the branch. The scale bar indicates the branch length.

분물질 분석, 약리활성에 대한 연구가 꾸준히 이루어져 왔지만 (Lee *et al.*, 2015; Kim, 2016), 잔대의 유전적 다양성이나 신품종 육성을 위한 유전적 연구는 RAPD나 DNA barcoding 등을 통해 이루어진 속간 유연관계의 포괄적인 규명에 국한되었고, 더구나 잔대속의 경우 속간 유전적 분자표지의 연구나 종내 다양성에 대한 연구는 매우 한정적이어서 (Eddie *et al.*, 2003; Cheon and Yoo, 2013), 확보한 잔대 유전자원의 다양성 확인 및 체계적 관리, 계통 분류 등에 많은 어려움이 있었다.

본 연구에서 개발된 SSR 마커는 잔대의 유전체 정보를 기반으로 한 것으로써, 잔대의 유전적 다양성을 염기서열 수준에서 밝히는데 매우 유용하다는 것을 알 수 있었다. 또한, 이를 이용하여 국내 잔대 유전자원의 다양성을 성공적으로 파악함으로써, 잔대의 체계적인 계통 분류 및 분자 육종에 필요한

기본적 자료를 제공하였다. 본 연구에서 개발된 마커는 앞으로 국내외 잔대 유전자원의 확보 및 다양성 분석, 우수 형질의 잔대 품종 개발에도 도움을 줄 것으로 사료된다.

## 감사의 글

본 연구는 농촌진흥청 연구사업(과제번호: PJ0102882017, PJ01102202)의 지원에 의해 이루어진 것으로 이에 감사드립니다.

## REFERENCE

- Botstein D, White RL, Skolnick M and Davis RW.** (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*. 32:314-331.
- Cheon KS and Yoo KO.** (2013). Phylogeny of *Hanabusaya* (Campanulaceae), a Korean endemic, based on ITS sequences of nuclear ribosomal DNA. *Journal of Systematics and Evolution*. 51:704-714.
- Coburn JR, Temnykh SV, Paul EM and McCouch SR.** (2002). Design and application of microsatellite marker panels for semiautomated genotyping of rice (*Oryza sativa* L.). *Crop Science*. 42:2092-2099.
- Eddie WMM, Shulkina T, Gaskin J, Haberle RC and Jansen RK.** (2003). Phylogeny of Campanulaceae S. Str. inferred from ITS sequences of nuclear ribosomal DNA. *Annals of the Missouri Botanical Garden*. 90:554-575.
- Eu GS, Park MR and Yun SJ.** (2009). Internal transcribed spacer(ITS) regions reveals phylogenetic relationships of *Rubus* species cultivated in Korea. *Korean Journal of Medicinal Crop Science*. 17:165-172.
- Gil J, Um Y, Kim S, Kim OT, Koo SC, Reddy CS, Kim SC, Hong CP, Park SG, Kim HB, Lee DH, Jeong BH, Chung JW and Lee Y.** (2017). Development of genome-wide SSR markers from *Angelica gigas* Nakai using next generation sequencing. *Genes*. 8: 238. <http://www.mdpi.com/2073-4425/8/10/238/htm> (cited by 2017 Sep 18).
- Henry RJ.** (2012). Next-generation sequencing for understanding and accelerating crop domestication. *Briefings in Functional Genomics*. 11:51-56.
- Jin L and Chakraborty R.** (1994). Estimation of genetic distance and coefficient of gene diversity from single-probe multilocus DNA fingerprinting data. *Molecular Biology and Evolution*. 11:120-127.
- Kilian B and Graner A.** (2012). NGS technologies for analyzing germplasm diversity in genebanks. *Briefings in Functional Genomics*. 11:38-50.
- Kim HK.** (2016). *Adenophora remotiflora* protects human skin keratinocytes against UVB-induced photo-damage by regulating antioxidative activity and MMP-1 expression. *Nutrition Research and Practice*. 10:371-376.
- Kim S, Jeong JH, Chung H, Kim JH, Gil J, Yoo J, Um Y, Kim OT, Kim TD, Kim YY, Lee DH, Kim HB and Lee Y.** (2016). Simple sequence repeat marker development from *Codonopsis lanceolata* and genetic relation analysis. *Journal of Plant Biotechnology*. 43:181-188.
- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA and Janzen DH.** (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*. 102:8369-8374.
- Lee DR, Lee YS, Choi BK, Lee HJ, Park SB, Kim TM, Oh HJ, Yang HY and Suh JW.** (2015). Roots extracts of *Adenophora triphylla* var. *japonica* improve obesity in 3T3-L1 adipocytes and high-fat diet-induced obese mice. *Asian Pacific Journal of Tropical Medicine*. 8:898-906.
- Lee JH, Jo IH, Lee JW, Park CG, Bang KH, Kim HS and Park CB.** (2010). Molecular authentication of *Scrophularia* herbs by PCR-RFLP based on *rpl-5* region of mitochondrial DNA. *Korean Journal of Medicinal Crop Science*. 18:173-179.
- Lee JK.** (1989). Taxonomy of the genus *Adenophora* (Adenophorae) in Korea. Master Thesis. Sungkyunkwan University. p.75.
- Lee ST, Chung YJ and Lee JK.** (1988). A palynotaxonomic study of the Korean Campanulaceae. *Korean Journal of Plant Taxonomy*. 18:115-131.
- Lee SY, Lee SK, Lee JK, Kim KH and Lee SK.** (2006). Cultural characteristics and genetic diversity of *Rhizina undulata* isolates by random amplified polymorphic DNA(RAPD). *Journal of Korean Forestry Society*. 95:388-392.
- Liu KJ and Muse SV.** (2005). PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics*. 21:2128-2129.
- Mardis ER.** (2008). Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics*. 9:387-402.
- McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T and Blair M.** (1997). Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Molecular Biology*. 35:89-99.
- Nybohm H.** (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*. 13:1143-1155.
- Shin DH and Seo YB.** (2000). Bibliographical studies on the shashen(*Adenophora triphylla* var. *japonica* Hara.). *The Journal of Daejeon Oriental Medicine*. 8:107-122.
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M and Rozen SG.** (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Research*. 40:e115. <https://doi.org/10.1093/nar/gks596> (cited by 2017 Jan 25).
- UPOV-BMT.** (2002). C/36/10 Progress report of the 36th session of the technical committee, the technical working parties and working group on biochemical and molecular techniques and DNA-profiling in particular. UPOV-BMT. Geneva, Switzerland. p.21-22.
- Wang Z, Gerstein M and Snyder M.** (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics*. 10:57-63.
- Weising K, Nybohm H, Wolff K and Kahl G.** (2005). DNA fingerprinting in plants: Principles, methods, and applications, second edition. CRC Press. Boca Raton. FL, USA. p.21-146.