국내 유통국화의 품종판별체계 및 유연관계 연구를 위한 EST-SSR마커 개발

한정윤¹ · 김정분¹ · 이호진¹ · 홍창표² · 박하승³ · 김태성^{1*}

¹한국방송통신대학교 농학과, ²(주)테라젠이텍스 테라젠바이오연구소, ³충남농업기술원 화훼연구소

Development of EST-SSR Markers for Cultivar Determination and Genetic Diversity Studies of Commercial Chrysanthemums in the Korean Floral Market

Jeong-Yun Han¹, Jung-Bun Kim¹, Ho-Jin Lee¹, Chang Pyo Hong², Ha-Seung Pak³, and Tae-Sung Kim^{1*}

¹Department of Agriculture and Life Sciences Korea National Open University, Seoul 03087, Republic of Korea

²Theragen Bio Institute, Theragen Etex, Suwon 16229, Korea,

³Flower Research Institute, Chungcheongnam-do ARES, Yesan 32427, Republic of Korea

Abstract *Chrysanthemum morifolium* Ramat. is one of the major flowering crop plants worldwide. However, domestic chrysanthemum markets have recently faced a downturn. To stimulate related industries, breeding technologies and efficient protection systems using molecular markers must be established. However, high cost and intensive efforts are required to develop useful molecular markers for the chrysanthemum as it is a polyploid crop with highly complex genome organization. Thus, the aim of this research was to develop expressed sequence tag-based simple sequence repeat (EST-SSR) markers, which are applicable to the chrysanthemum breeding program and cultivar protection, based on next-generation sequencing technology. From the RNAseq data of the standard chrysanthemum cultivars 'Jungwoon' and 'Seinoisei,' we identified 31,121 SSR loci and further retrieved 1,846 polymorphic SSRs. To test the marker efficiency of the 1,846 SSRs, we first chose 50 of the SSRs and designed primers by using the flanking sequences. It is noted that the nine EST-SSR markers show a single band-like amplicon, which can be exploited in various genetic studies. We proceeded to polymorphism tests for those SSRs with 56 chrysanthemum cultivars, confirming that the average polymorphism index content (PIC) was 0.69 ± 0.058 . Among those, we found that six SSRs were sufficient to specify the genetic identities of 55 chrysanthemum cultivars, which may be useful for protections of the related cultivars, as well as breeding programs, in the future.

Keywords chrysanthemum, EST-SSR, next-generation sequencing, RNAseq, molecular marker, polymorphism

Received on July 18, 2019. Revised on July 25, 2019. Accepted on July 30, 2019.

* Corresponding Author (E-mail: kts117@knou.ac.kr, Tel: +82-2-3668-4638)

서 언

국화(*Chrysanthemum morifolium* Ramat.)는 전 세계적으로 널리 재배되는 주요 화훼작물로, 우리나라에서는 2017년 재배면적 341 ha, 생산액 49,213백만 원으로 장미, 백합과 함께 3대 화훼류(Hwang et al. 2012)의 하나에 속한다. 이는 국내 전체 절화류 재배면적과 생산액의 각각 26%와 27%를 차지하고 있다(MIFAFF 2018). 하지만 최근 국화 재배면적이 2009년 대비 약 46%로 크게 감소하면서 국내산 국화의 경쟁력을 높이기 위한 방안 모색이 필요하다. 생산규모 감소의 원인은 국내에서 유통되는 국화가 주로 축하용 절화로 동절기에 소비되거나 지속적인 인건비 및 연료비 상승 등으로 농가의 부담이 늘었기 때문으로 보인다(MIFAFF 2018).

한편, 국화는 장미, 백합과 더불어 수출 주력 품목으로서 대부분 일본으로 수출되는데, 이것은 선도유지가 중요한 화훼류의 특성으로 인하여 인접한 일본으로 신속한 수송이 가능할 뿐만 아니라 거리가 가까워 물류비가 절감되는 이점 때문이다(Lim et al. 2013). 하지만 국화 내수시장의 부진과 더불어 주 수출국인 일본에 중국 및 말레이시아 등 경쟁국의 국화 수출비율이 최근 비약적으로 늘어나면서, 국내 국화의 총 수출이 최전성기인 2010년 약 1천3백 8십만 달러에서 2017년에는 약 2백만 달러로 크게 감소하였다(MIFAFF 2018). 말레이시아는 고랭지에서 국화가 재배되어 품질이 우수할 뿐만 아니라 가격도 낮아 수요가 많으며, 중국 또한 대규모로 국화를 생산하여 저가로 일본시장을 공략하고 있는 실정이다(Lim et al. 2013). 따라서 국내 국화산업에 활력을 불어넣기 위해서는 혁신적인 품종을 개발할 수 있는 육종기술 개발과 국내

에서 육성된 품종을 효율적으로 보호할 수 있는 체계가 절실하다 (Shim et al. 2015).

DNA 염기서열을 기초로 한 분자표지 마커는 유전적 다양성 평가, 유전자원의 도입과 선발, 계통선발 및 품종구별 등의 육종 분야 및 품종보호에 이용되고 있다(Hong et al. 2013). 특히 Microsatellite라고도 불리는 SSR (Simple sequence repeat)은 DNA에서 6개 미만의 염기서열이 단순 반복되는 motif를 가지고 있으며, 식물체 게놈 전체에 널리 분포한다. SSR은 염기서열 motif 개수의 변화가 종 내에서도 비교적 높게 일어나며 특정 종의 집단에서의 유전적 다양성과 유연관계를 평가하는데 매우 유용하다(Hong et al. 2013). 국화의 유전체는 배수성 정도가 높아 유전체계가 매우 복잡하여, 재현성이 높고 효과적인 SSR 분자표지 마커 개발에 있어 상당한 비용과 노력이 소요된다(Hong et al. 2013, Shim et al. 2015). 따라서 국내에서 SSR마커가 다양한 작물에서 활용되고 있지만 아직 국화에서는 충분히 활용되지 못하고 있다(Hong et al. 2013, Shim et al. 2015). 그러나 최근 NGS (Next Generation Sequencing) 기술의 진보로 염기서열 생산 비용이 크게 저렴해져 향후 국화에서도 DNA 염기서열에 기초한 분자표지 마커 개발이 더욱 활발하게 진행될 전망이다(Won et al. 2016). 특히 국화는 표준유전체 정보가 없지만 transcriptome 정보는 재배되는 국화품종에 대한 SSR마커 등 유용 분자표지 마커를 대량으로 발굴할 수 있는 기초를 제공할 수 있다(Wang et al. 2013, Won et al. 2016, Han et al. 2018).

본 연구는 국화품종 육성과 품종판별에 대한 활용도를 높이고자, 국내에서 유통되는 스탠다드 국화 ‘정운’과 ‘세이노이세이’ 품종을 기반으로 EST_SSR 서열을 transcriptome분석을 통해 대량 발굴하였고, 그 중 50개를 임의로 선택하여 분자표지 마커로서의 효율을 검정하였다.

재료 및 방법

국화 전사체 데이터 기반 EST_SSR 대량 추출

공시재료는 예산 화훼연구소에서 재배 중인 국내 유통되는 품종인 스탠다드 국화 ‘정운’과 ‘세이노이세이’의 두 개의 품종에 대해 각각 액아 및 여러 조직을 혼합 채취하여 -80℃에 보관하였고, 이들 국화 sample을 HiGene™ Total RNA Prep Kit (BIOFACT, Korea)으로 RNA를 추출 하였다. 형광분광광도계(NanoDrop® ND-1000, Thermo Scientific, USA)와 Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA)를 통해 추출한 total RNA의 순도와 농도를 확인하였고(Cox et al. 2010), Illumina

TruSeq RNA 샘플 키트 v2 (RS-122-2001, Illumina, San Diego, CA)를 사용하여 ‘정운’과 ‘세이노이세이’의 RNA-seq 라이브러리를 구축하였다. 구축된 라이브러리는 Illumina Hiseq2500 플랫폼으로 시퀀싱하여 short read를 생산하였고, 생산된 read를 pooling하여 ‘Trinity’로 De novo assembly를 수행한(Grabherr et al. 2011) 후, TGILC과 CAP3 software를 기반으로 중복을 최소화한 unigene을 구축하였다(Pertea et al. 2003). 각 샘플의 read를 구축된 unigene에 mapping하여 ‘정운’과 ‘세이노이세이’ 각각의 transcriptome 데이터셋을 구성하였다(Mortazavi et al. 2008). EST_SSR sequence는 최초 De novo assembly를 수행한 unigene을 대상으로 Tandem Repeats Finder를 통해 발굴하였고(Benson 1999), 이들 SSR을 ‘정운’과 ‘세이노이세이’ 각 transcriptome 내에서 다시 검색한 후, polymorphic EST_SSR sequence만을 선발하였다.

SSR마커개발을 위한 Primer 제작

선발된 polymorphic EST_SSR의 flanking sequence를 대상으로 Polymerase Chain Reaction (PCR)의 증폭산물이 200~300base pair (bp)가 될 수 있게 forward 및 reverse primer쌍을 Primer3 software를 이용하여 default setting으로 일괄 design 하였다 (Untergasser et al. 2012).

국화 품종판별을 위한 공시재료 선택 및 DNA 추출

공시재료는 충청남도농업기술원 예산화훼연구소에서 재배되고 있는 품종 중 국내에서 주로 유통되는 스프레이 타입 52품종, 스탠다드 타입 4품종 등 총 56품종을 선발하여 분석에 사용하였다. 그 중에 마커 분류를 위한 예비실험에는 스프레이 국화 Argus, Borami, Mujigae, Yes Ruby 4품종과 스탠다드 국화 Baekseol, Baekson, Zinba, Sumi 4품종 등 총 8품종을 이용하였으며(Supplementary Table 2), 국내 국화 DNA profiling을 위한 유연관계분석을 위한 실험에서는 전체 56품종을 이용하였다(Supplementary Table 2). 또한 공시재료의 DNA 추출은 Pak et al. (2019)에 기술되어 있는 대로 수행되었다.

PCR 증폭 및 증폭산물의 다형성 분석

PCR 반응은 Pak et al. (2019) 기술되어있는 대로 PCR Mixture를 조성하였고, 반응조건은 95℃에서 3분간 Initial denaturation를 시행한 후, 95℃에서 1분간 Denaturation, 55℃에서 45초 Annealing, 72℃에서 1분간 Extension으로 이 세 단계를 35번 반복한 뒤에 72℃에서 10분간 Final Extension을 수행하였다. PCR 완료 후

증폭여부는 Agarose gel 상에서 전기영동하여 Gel doc™ system (Biorad, USA)에서 확인하였다. 국화 품종의 각 SSR locus의 다형성 분석을 위해 증폭산물을 모세관전기영동기인 ZAG DNA Analyzer (Agilent, USA)에서 분리한 후, PROSize software (v.2.0) 상에서 육안으로 보이는 SSR의 major band의 형태를 고려하였다. 최종 증폭 산물의 size는 증폭산물의 가장 높은 peak를 기준으로 결정하였다.

국내 국화 품종의 유연관계 및 다양성 분석

국내 국화 품종의 유연관계 및 다양성은 PROSize 2.0 프로그램을 통해 도출된 SSR의 size variation을 기반으로 PowerMaker (v.3.25) 프로그램을 사용하여 분석하였다. 국화 품종의 유연관계는 UPGMA (unweighted pair-group method with arithmetical average) 방법으로 clustering되었으며, 각 SSR locus에 대한 집단 내 다양성은 대립유전자 빈도(MAF, Major Allele Frequency), 대립유전자 수(NA, Number of alleles), 유전적 다형성 정보 함유량(PIC, Polymorphic Information Content), 이형접합(H, Heterozygosity) 등으로 제시하였다.

결과 및 고찰

Short read assembly, Polymorphic EST_SSR sequence 발굴

‘정운’과 ‘세이노이세이’와의 잎 조직을 포함한 혼합샘플에서 total RNA를 추출한 후, Hiseq2500 기반 RNA sequencing을 진행한 결과 약 13,554 Mbp (megabasepair)와 13,694 Mbp에 해당하는 135,549,954개와 136,943,528개의 read를 각각 생산하였고, 이 중 약 5%의 염기서열이 clean-up과정에서 제외되었다 (Table 1). De novo assembly의 정확도를 높이기 위해 filtering된 read dataset을 pooling하여 assembly를 진행하여 240,434개의 unigene을 생산하였고, 평균 unigene의 길이는 851 bp였다 (Fig. 1). ‘정운’과 ‘세이노이세이’ 각각의 transcriptome data set은 초기 pooling dataset으로 구축된 unigene에 mapping하여 재구성되었다.

Pooling dataset의 unigene sequence를 기반으로 총 31,121개의 EST_SSR을 발굴하였고, 그 SSR sequence들을 ‘정운’과 ‘세이노이세이’ 각각의 unigene set에 blast를 수행하여 1,846개의 polymorphic SSR을 탐색하였다. 국화 unigene에 존재하는 polymorphic EST_SSR 중 di nucleotide SSR (di-nt SSR)에서는 AT/TA type이, tri-nt SSR은 ATG/TGA/GAT가 가장 많이 분포하였다 (Fig. 2). Tri-nt SSR은 frame-shift를 최소화 시킬 수 있기 때문에 coding region에 주로 관찰된다 (Kim et al. 2008). Tri-nt

Fig. 1. The biplot result between the number of unigenes and their corresponding sizes as a result of de novo assembly using standard type chrysanthemum short read sequences of ‘Jung Woon’ and ‘Seino Isei’.

Fig. 2. Frequency of polymorphic SSRs by SSR types originating from chrysanthemum unigene sequences.

Table 1. RNA sequencing results of standard type chrysanthemum ‘Jung Woon’ and ‘Seino Isei’ generated by Hiseq2500.

No.	Varieties	Raw		Clean reads	
		Reads	Basepair	Reads	Basepair
1	SI ^z	135,549,954	13,554,995,400	128,967,106	12,854,378,198
2	JW ^y	136,943,528	13,694,352,800	129,961,130	12,954,593,060
	Mean	136,246,741	13,624,674,100	129,464,118	12,904,485,629

^{z, y}SI and JW stand for the standard chrysanthemum cultivar ‘Seino Isei’ and ‘Jung Woon’ respectively.

SSR 중 size variation율이 높은 type은 exon내 글루타민(glutamine) 반복을 일으킬 수 있는 CAA와 CAG 서열이 포함된 SSR type—즉 AAC/ACA/CAA나 AGC/GCA/CAG 등—이 보편적이다. 하지만, start codon이 될 수 있는 ATG/TGA/GAT SSR type은 exon에서 상대적으로 분포가 낮은 편이기 때문에 국화 transcriptome data에서 관찰된 높은 polymorphic ATG/TGA/GAT type의 비율은 이례적인 사례이며, 향후 이에 대한 면밀할 검토가 필요하다 (Kim et al. 2008, Kim et al. 2015).

유용 SSR 마커 검정

1,846개의 Polymorphic EST_SSR 중 size 차가 크고 200-300 bp 정도의 증폭산물을 생성할 수 있는 SSR loci 50개를 무작위적으로 선택 후 flanking sequence를 이용하여 primer를 design 하였고 (Supplementary Table 1), Yes Ruby 등 8개의 품종을 대상으로 PCR을 수행하였다. PCR 후 증폭산물을 Agarose 상에서 1차 확인한 뒤, 동일 증폭산물을 모세관 전기영동 시스템에서 분리하여 ProSize software 상에서 증폭산물의 size를 판독하였다(Fig. 3, Supplementary Fig. 1). 총 31개 SSR loci에서 증폭산물을 안정적으로 관찰할 수 있었고, 31개중 22개는 major band를 구분할 수 있었으나 multiple band와 같은 양상을 보였고, 9개 SSR은 비교적 scoring이 용이한 single band 형태를 보였다(Fig. 3, Supplementary Fig. 1).

국화 집단 내 SSR polymorphism 정도를 예측하기 위해 9개의 single band형 EST_SSR을 대상으로 다형성 정도와 연관된

parameter들—major allele frequency (MAF), allele number (AN), genotype number (GD), Heterozygosity (H), Polymorphism index content (PIC)—를 모세관 전기영동 시스템에서 확인된 size를 정보를 바탕으로 산출하였다. MAF값은 평균 0.42, AN값은 평균 0.43, GN값은 평균 4.13, H값은 평균 0.01, PIC값은 평균 0.61으로, SSR locus별로 다형성 정도가 다양한 차이를 보였고, KNOU_CrySSR40, KNOU_CrySSR45 및 KNOU_CrySSR50은 다른 single 밴드형 SSR보다 상대적으로 다형성이 낮은 것으로 판명되었다(Table 2). 종합적으로 본 연구에서 확인된 국화 EST_SSR 마커의 다형성은 국화에서 관찰되었던 다른 연구와 비슷한 정도였다(Shim et al. 2015). 한 개의 마커 당 9개의 품종 중 약 4.13개의 유전자형을 구분할 수 있어 다수의 품종도 안정적으로 구분해 낼 수 있다고 판단되어 국내에서 유통되는 국화 56품종에 대해 다형성 분석을 확장하였다(Table 2). 확장된 국화 품종에 대해 single band 형 EST_SSR 9개 중 8개에 대한 다형성 분석 결과, MAF값은 평균 0.44, AN값은 평균 14.00, GN값은 평균 15.75, H값은 평균 0.27, PIC값은 평균 0.69으로 8개 품종 결과와 유사하였다(Table 2). 하지만 PIC 기준 약 58%의 데이터 셋만이 선형회귀로 설명되는 반면, 2차함수로 설정된 곡선회귀가 적용이 되었을 때 약 81%의 데이터 셋이 관련 회귀분석으로 설명이 되는 것을 알 수 있었다(Fig. 4A). 특히 다형성이 높은 KNOU_CrySSR4 등과 같은 SSR loci군은 샘플의 숫자가 늘어남에 따라 다형성 정도가 크게 변화가 없지만, 상대적으로 다형성 정도가 낮게 평가된 KNOU_CrySSR40 등과 같은 SSR

Fig. 3. Exemplary images of PCR amplification for 22 chrysanthemum SSR loci, which produce multiple bands (A, C, E) and single band like pattern (B, D, F), which include KNOU_CRYSSR1 (A), KNOU_CRYSSR7 (B), KNOU_CRYSSR8 (C), KNOU_CRYSSR31 (D), KNOU_CRYSSR23 (E) and KNOU_CRYSSR42 (F). The DNA template of each PCR reaction for lane 1 is Argus; lane 2, Borami; lane 3, Mujigae; lane 4, Yes Ruby; lane 5, Baekseol; lane 6, Baekson; lane 7, Zinba; lane 8, Sumi. PCR amplification results of Six SSR loci out of 22 are shown as examples. The PCR amplicon of each SSR locus is separated by capillary electrophoresis system and visualized as a virtual gel image, obtained from PROSize software (v.2.0).

loci군들은 중간 정도의 SSR군들은 다형성 정도가 뚜렷이 증가되었다(Fig. 4A, Table 2). 특정 마커에 대해 집단 내 다형성을 예측할 때 비교적 작은 샘플로 예측을 시도하게 되면 다형성에 관계된 각 지표들을 과대평가 혹은 과소평가를 할 수 있음이 보고되었다(Chung et al. 2014, Kim et al. 2015). 본 연구에서도 KNOU_SSR40 등과 같이 PIC 값 기준 0.75이하인 loci인 경우

집단의 규모가 8품종에 불과할 때, 관련 마커의 다형성 예측이 평가절하될 수 있음을 확인하였다(Fig. 4A). KNOU_CrySSR40에 대해 샘플 규모를 8, 24, 40 품종으로 확대한 결과, PIC 예측에 대한 오차범위가 점차 줄어들었고, 특히 집단 규모가 24품종 이상이 되면 평가 절하 효과가 완화가 되는 것을 관찰할 수 있었다(Fig. 4B). 따라서 향후 국화 SSR마커 다형성 예측 시

Table 2. The polymorphism tests using single band like chrysanthemum EST_SSR markers.

No.	Name	Product size (bp)	8 cultivar							56 cultivars					
			A ^z	MAF ^y	AN ^x	GN ^w	H ^v	PIC ^u	A	MAF	AN	GN	H	PIC	
1	KNOU_CRYSSR4	248-285	7	0.43	5	5	0	0.70	52	0.34	20	23	0.25	0.83	
2	KNOU_CRYSSR7	240-267	7	0.43	4	4	0	0.64	47	0.30	14	15	0.15	0.80	
3	KNOU_CRYSSR9	371-492	6	0.16	6	6	0	0.81	-	-	-	-	-	-	
4	KNOU_CRYSSR16	185-190	7	0.29	6	6	0	0.79	44	0.43	18	20	0.45	0.74	
5	KNOU_CRYSSR31	140-157	8	0.31	6	2	0.12	0.76	54	0.44	11	13	0.15	0.71	
6	KNOU_CRYSSR40	146-152	7	0.57	2	6	0	0.37	42	0.44	15	15	0.36	0.71	
7	KNOU_CRYSSR42	141-170	8	0.25	6	2	0	0.79	54	0.31	17	21	0.43	0.82	
8	KNOU_CRYSSR45	159-188	8	0.63	2	2	0	0.36	56	0.46	8	10	0.29	0.57	
9	KNOU_CRYSSR50	124-129	8	0.75	2	5	0	0.30	55	0.80	9	9	0.11	0.34	
Mean			7.33	0.42	4.13	0.67	0.01	0.61	50.50	0.44	14.00	15.75	0.27	0.69	
SEM ^t			0.236	0.065	0.666	0.058	0.013	0.070	1.909	0.056	1.535	1.820	0.047	0.058	

^zA: stands for availability of dataset.

^yMAF: stands for major allele frequency.

^xAN: stands for allele number.

^wGN: stands for genotype number.

^vH: stands for heterozygosity.

^uPIC: stands for polymorphism information content.

^tSEM: stands for standard error mean.

Fig. 4. The PIC estimation of the chrysanthemum SSR marker of our study. (A) The correlation of PIC estimation between 8 and 56 chrysanthemum cultivars among the 8 KNOU_SSR markers. The dataset is fitted to linear or polynomial regression model. The equation and R-square value of the linear and polynomial model are shown lower and upper part of the figure respectively. (B) The PIC estimation of KNOU_CrySSR40 with randomly selected 8, 24 and 40 cultivars from the 56 chrysanthemums. In the given number of cultivars which are randomly selected from the 56 chrysanthemums, the PIC values are estimated independently.

정확성을 확보하기 위해 적어도 24품종 혹은 자원을 사용해야 할 것으로 보인다.

국화 56품종 유연관계 분석

56 품종에 대해 8개 SSR 마커의 증폭된 DNA size 패턴을 기반으로 UPGMA 방법으로 클러스터링을 수행하였다(Fig. 5). 그 결과, 비유사도 지수(dissimilarity index) 약 0.05에서 3개의 집단으로 나눌 수 있었다. 56개 품종 중 55개 품종이 약 비유사도

지수 0.41 범위안에서 구분할 수 있었지만, 이 범위 내에서 유럽종 스프레이 타입 ‘Leopard’와 국내종 스프레이타입 ‘Loving You’는 구분해낼 수 없었다. 국내종 스프레이타입 ‘Loving You’는 녹색인 ‘Furore’를 모본, 흰녹병에 강한 녹색 스파이더 화형의 ‘Green Joy’를 부분으로 인공교배하여 획득한 교배종이고, 스프레이국 ‘Leopard’와는 형태적으로 구분이 비교적 뚜렷하다. 본 연구에서 개발된 SSR마커로 ‘Loving You’와 ‘Furore’를 유전적으로 구분을 하지 못한 데에는 개발된 SSR마커의 resolution이

Fig. 5. A UPGMA dendrogram based on dissimilarity indices from 8 SSR markers as to 56 commercial chrysanthemum cultivars. Numbers below each branches refer to dissimilarity indices from UPGMA algorithm and the tree length is equivalent to dissimilarity index value. The scale bar is indicated in the left bottom. The dotted line on the left is the point where the populations are specified as denoted by different alphabets in the right. More info for the cultivars are listed in the Supplementary Table 1.

충분치 않기 때문일 수 있지만, 두 품종에 대한 각 마커의 missing 데이터의 영향도 배제할 수 없다(Supplementary Table 3). 다형성이 상대적으로 낮은 KNOU_CrySSR45와 KNOU_CrSSR50 없이도 국화 55품종을 충분히 유전적으로 구분해 낼 수 있었지만, 6개 이하의 마커에서는 55품종을 성공적으로 구분하기 어려웠다.

현재 국내 국화품종 육성은 대부분 인공교배 후 실생 선발에 의해 이루어지고 있다(Kim et al. 2015, Pak et al. 2019). 특히 국화는 후대에서 표현형의 변이 폭이 매우 크기 때문에 재배안정성 및 농업형태적 특성과 더불어 다양한 화색, 화형을 모두 갖춘 우량품종을 육성하는데 많은 자본과 노력이 요구되나(Shim et al. 2016), 시장성이 높은 품종들의 품종연한은 비교적 긴 편이다(Lim et al. 2003). 하지만 영양번식작물이기 때문에 어디서나 쉽게 대량생산이 가능하여 효과적인 품종보호체계가 절실히 요구된다(Hong et al. 2013, Shim et al. 2015). SSR은 다형성 및 재현성이 높고 이용하기 쉬운 공유성 마커이기 때문에 국내외에서 육성 품종의 보호에 많이 활용되어 왔으나(Nguyen et al. 2019), 국화의 경우에는 유전체의 배수성 정도가 높고 유전적 잡박성과 중복성이 심하기 때문에 비교적 보편적으로 사용되지 못하는 실정이다.

SSR (Simple sequence repeat)은 식물체 게놈 전체에 널리 분포하여 유전체 정보가 있는 경우에는 대량으로 추출이 가능하지만 국화유전체 정보는 아직 가용하지 않기 때문에, 최근 next generation sequencing에 의한 transcriptome 정보를 통해 다양한 국화 근연종에서 EST-SSR마커를 대량으로 발굴하는 노력이 시도되고 있으나, 이를 국화 재배종에 적용할 때 마커로서의 효율이 떨어짐을 관찰할 수 있었다(Wang et al. 2013, Won et al. 2016; Han et al. 2018). 본 연구에서는 국내에서 재배되는 화훼국화품종(*Chrysanthemum morifolium* Ramat.)의 SSR마커를 transcriptome 정보를 이용하여 대량발굴 및 마커 개발 가능성을 목적으로 수행하였다. ‘정운’과 ‘세이노이세이’ RNAseq정보를 통해 31,121개 중 1,846 polymorphic SSR을 추출하였으며, 그 중 50개를 테스트 한 결과 약 16%에서 평균 PIC 값이 0.69 ± 0.058 되며, single-band 형의 마커를 개발할 수 있음을 확인하였다. 이는 국화 근연종인 *Chrysanthemum nankingense*의 transcriptome data를 기반으로 SSR마커를 개발한 경우보다 약 5배 이상의 효율에 해당되기 때문에(Shim et al. 2015), 향후 본 연구는 국화의 SSR마커 활용도를 높이는데 기여할 것이다.

적 요

국화(*Chrysanthemum morifolium* Ramat.)는 전세계적으로 주요 화훼작물로, 국내 국화 품종 개발을 위한 육종 기술과 육성 품종의 효율적인 보호체계가 필요하다. 본 연구는 기존에 활용되지 못하였던 국화의 염기서열을 활용하여 국내 국화 품종의 다형성 평가와 품종 구별 등 육종 분야와 품종 보호에 적용 가능한 SSR마커를 개발하는 것에 목적이 있다. 이를 위해 국내에서 유통되는 스탠다드 국화 ‘정운’과 ‘세이노이세이’ 품종을 기반으로 EST-SSR 서열을 transcriptome 분석을 통해 대량 발굴하였고, 그중 50개를 임의로 선택하여 분자표지 마커로 개발을 시도하였다. Transcriptome 분석으로 예측된 flanking region의 염기서열을 대상으로 프라이머를 디자인을 하여 PCR (Polymerase chain reaction)을 시도한 결과 50개중 31개가 증폭되었다. 31개의 후보 SSR마커중 9개는 single band를 보였고, 22개는 multiple band를 보였다. 56개의 국화 품종을 대상으로 single band형 SSR마커 중 8개에 대해 다형성 정보 함유량(Polymorphic Information Content)을 측정한 결과 평균 0.69 ± 0.058 이었다. 또한 8개 마커셋 중 6개로도 국내에서 유통되는 국화 55품종을 profiling할 수 있었다. 본 연구는 향후 국화 품종육성과 유전분석 및 품종판별에 대한 SSR마커의 활용도를 높이는 기초자료를 제공하며, 기존에 개발되었던 SSR마커의 결과를 보완하여, 국내 유통 국화의 품종판별체계의 정확성을 높이고 국내 국화분자육종의 인프라 확충에 기여할 것으로 기대된다.

보충자료

본문의 Supplementary Table 1과 Supplementary Table 2, Supplementary Table 3, Supplementary Fig. 1은 한국육종학회지 홈페이지에서 확인할 수 있습니다.

사 사

이 논문은 2016년도 한국방송통신대학교 학술연구비 지원을 받아 작성된 것임.

REFERENCES

- Benson G. 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* 27: 573-580.

2. Chung JW, Kim TS, Sundan S, Lee G, An JH, Park GT, Cho HS, Lee JY, Lee MC, Lee HJ, Lee SY. 2014. New cDNA-SSR markers in the narrow-leaved vetch (*Vicia sativa* subsp. *nigra*) using 454 pyrosequencing. *Mol Breeding* 33: 749-754.
3. Cox MP, Peterson DA, Biggs PJ. 2010. At-a-glance quality assessment of illumina second-generation sequencing data. *BMC Bioinformatics* 11: 485.
4. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29: 644-652.
5. Han Z, Ma X, Wei M, Zhao T, Zhan R, Chen W. 2018. SSR marker development and intraspecific genetic divergence exploration of chrysanthemum indicum based on transcriptome analysis. *BMC Genomics* 19: 291.
6. Hong WJ, Khaing AA, Park YJ. 2013. Cultivar identification of chrysanthemum (*Dendranthema grandiflorum*. Ramat.) using SSR markers. *Kor J Intl Agri* 25: 385-394.
7. Hwang JC, Chin YD, Chung YM, Kim SK. 2012. A new spray chrysanthemum cultivar, 'Green Hof' with pompon type and green color petals for cut flower. *Flower Res J* 20: 134-138.
8. In BC, Sim SC, Lim JH. 2015. Analysis of genetic diversity and evaluation of phenotypic traits in *Chrysanthemums morifolium*. *Flower Res J* 23: 261-269.
9. Kim DC, Choi HG, Park HS, Lee YH, Won M, Jung YK, Lee JS. 2015. Breeding of garden chrysanthemum cultivar nuri ball (*Dendranthema grandiflorum* Ramat.) with white color petals and semi-decorative type characteristics. *Kor J of Hort Sci Technol* 33: 789-795.
10. Kim TS, Booth JG, Gauch HG, Sun Q, Park J, Lee YH, Lee K. 2008. Simple sequence repeats in *neurospora crassa*: Distribution, polymorphism and evolutionary inference. *BMC Genet* 9: 31.
11. Kim TS, Raveendar S, Suresh S, Lee GA, Lee JR, Cho JH, Lee SY, Ma KH, Cho GT, Chung JW. 2015. Transcriptome analysis of two *vicia sativa* subspecies: Mining molecular markers to enhance genomic resources for vetch improvement. *Genes* 6: 1164-1182.
12. Lim JH, Seo JY, Shim MS. 2013. Characteristics that affect Japanese consumer preferences for chrysanthemum. *Kor J Hort Sci Technol* 31: 640-647.
13. Ministry for Food, Agriculture, Forestry and Fisheries (MFAFF). 2018. Statistics for floricultural industry in 2017. Sejong, Korea. pp. 334-336.
14. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5: 621-628.
15. Nguyen NN, Kwon YS, Park JR, Sim SC. 2019. Development of a core set of SSR markers for cultivar identification and seed purity tests in oriental melon (*Cucumis melo* L. var. *makuwa*). *Kor J of Hort Sci Technol* 37: 119-129.
16. Pak HS, Choi HG, Kim D, Son KS, Kim TS, Choi JJ, Kim JH, Kim JO. 2019. Breeding of cut spray chrysanthemum 'Yes Ruby' with thermotolerance. *Flower Res J* 27: 73-79.
17. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S. 2003. TIGR Gene Indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. *Bioinformatics* 19: 651-652.
18. Shim EJ, Heo EJ, Yoon MK, Soh EH, Hong JH. 2015. Construction of SSR marker database of chrysanthemum varieties collected in Korea. *Kor J Breed Sci* 47: 366-375.
19. Shim SI, Lim KC, Kim CK, Chung MY, Kim KM, Chung JD. 2016. Morphological characteristics, and coefficient of variation, heritability and genetic advance of major cultivars of spray chrysanthemum. *Kor J of Hort Sci Technol* 34: 269-281.
20. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3-new capabilities and interfaces. *Nucleic Acids Res* 40: e115.
21. Wang H, Jiang J, Chen S, Qi X, Peng H, Li P, Song A, Guan Z, Fang W, Liao Y, Chen F. 2013. Next-generation sequencing of the chrysanthemum *nankingense* (Asteraceae) transcriptome permits large-scale unigene assembly and SSR marker discovery. *PLoS One* 8: e62293.
22. Won SY, Kim JS, Kang SH, Sohn SH. 2016. Current status and prospects of chrysanthemum genomics. *J Plant Bio Technol* 43: 272-280.