

Development of SNP marker set for marker-assisted backcrossing (MABC) in cultivating tomato varieties

GiRim Park^{1†}, Hyun A Jang^{2†}, Sung-Hwan Jo¹, Younghoon Park³, Sang-Keun Oh^{2*}, Moon Nam^{1*}

¹Seeders Inc, Daejeon 34015, Korea

²Department of Applied Biology, College of Agriculture & Life Sciences, Chungnam National University, Daejeon 34134, Korea

³Department of Horticultural Bioscience, Pusan National University, Miryang 50463, Korea

*Corresponding authors: sangkeun@cnu.ac.kr, moonlit51@seeders.co.kr

† These authors equally contributed to this study as first author.

Abstract

Marker-assisted backcrossing (MABC) is useful for selecting offspring with a highly recovered genetic background for a recurrent parent at early generation unlike rice and other field crops. Molecular marker sets applicable to practical MABC are scarce in vegetable crops including tomatoes. In this study, we used the National Center for Biotechnology Information- short read archive (NCBI-SRA) database that provided the whole genome sequences of 234 tomato accessions and selected 27,680 tag-single nucleotide polymorphisms (tag-SNPs) that can identify haplotypes in the tomato genome. From this SNP dataset, a total of 143 tag-SNPs that have a high polymorphism information content (PIC) value (> 0.3) and are physically evenly distributed on each chromosome were selected as a MABC marker set. This marker set was tested for its polymorphism in each pairwise cross combination constructed with 124 of the 234 tomato accessions, and a relatively high number of SNP markers polymorphic for the cross combination was observed. The reliability of the MABC SNP set was assessed by converting 18 SNPs into Luna probe-based high-resolution melting (HRM) markers and genotyping nine tomato accessions. The results show that the SNP information and HRM marker genotype matched in 98.6% of the experiment data points, indicating that our sequence analysis pipeline for SNP mining worked successfully. The tag-SNP set for the MABC developed in this study can be useful for not only a practical backcrossing program but also for cultivar identification and F1 seed purity test in tomatoes.



OPEN ACCESS

Citation: Park GR, Jang HA, Jo SH, Park Y, Oh SK, Nam M. 2018. Development of SNP marker set for marker-assisted backcrossing (MABC) in cultivating tomato varieties. Korean Journal of Agricultural Science. <https://doi.org/10.7744/kjoas.20180061>

DOI: <https://doi.org/10.7744/kjoas.20180061>

Received: July 16, 2018

Revised: July 27, 2018

Accepted: August 6, 2018

Copyright: © 2018 Korean Journal of Agricultural Science

 This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Keywords: Luna probe, marker-assisted backcrossing (MABC), next generation sequencing (NGS), single nucleotide polymorphism (SNP), tomato

Introduction

토마토(*Solanum lycopersicum L.*)는 가지과(Solanaceae)에 속하는 대표적인 채소 작물로서, 국내뿐만 아니라 전세계적으로 생산량 및 소비량이 많아 작물 시장에서 높은 경제적 가치를 가진다. 특히 토마토는 *Agrobacterium*을 통한 형질전환이 가능하고, 비교적 짧은 생활사 및 작은 유전체 크기를 가지고 있어 가지과 작물의 육종 모델로 사용되어 왔다(Foolad, 2007). 최근 토마토의

유전체 해독이 완료되고(The Tomato Genome Consortium, 2012), NGS (Next generation sequencing)를 이용한 대용량 염기서열 생산이 가속화되면서 육종 산업 내 유전체 기반의 분자 육종이 가지는 중요도는 더욱 더 커지고 있다.

분자 육종은 원하는 형질의 유무를 형질에서 유래한 표현형의 관찰 없이 DNA 염기서열의 차이를 보이는 분자 마커(Molecular marker)를 이용하여 판별하는 기법으로(Edwards and Batley, 2010), 전통적 선발 육종 방법에 비해 육종 연한을 단축시키고 우량 개체 선발 효율을 극대화시킬 수 있는 장점을 지니고 있다. 더욱이 분자 마커를 여교잡 육종에 적용하는 분자마커이용여교잡(Marker-assisted backcrossing, MABC)은 분자 마커를 이용하여 유묘기에 여교잡 자손의 염색체 조성을 확인함으로써 신속하고 정확하게 회복친의 염색체 이입개체를 선발할 수 있어 육종 연한과 비용, 노력 등을 효율적으로 감소시킬 수 있다(Servin and Hospital, 2002; Collard and Mackill, 2008). 토마토의 경우 많은 유용 형질 유전자들이 야생종에서 발견되고 있기 때문에(Foolad, 2007), 이러한 MABC 시스템을 활용하면 기존의 육종 기술보다 효율적인 계통의 육성을 기대할 수 있다. 현재 국내에서도 토마토의 내병성 및 주요 형질 연관 마커에 대한 연구와 더불어 MABC 육종 기술을 적용한 사례가 보고되고 있지만(Park et al., 2010a; Park et al., 2010b; Hwang et al., 2012), MABC에 활용할 수 있는 특이적 마커의 선발과 최종적으로 실제 육종 현장에 적용될 수 있는 MABC 기술은 미흡한 실정이다.

최근 MAS 및 MABC에 대표적으로 이용되는 분자 마커는 유전체 내 가장 높은 빈도로 존재하는 단일염기다형성(Single nucleotide polymorphism, SNP)이다. SNP는 국내에서도 토마토의 품종 구분 뿐만 아니라 유용 형질 유전자 선발을 위한 연구에 꼭넓게 활용되고 있다(Bae et al., 2010; Hwang et al., 2012). NGS에 기반한 전장유전체재분석(Whole genome resequencing, WGRS)은 비용이 큰 반면 대량의 SNP와 삽입/결실(Insertion/Deletion, Indel)과 같은 염기서열변이 정보를 확보할 수 있는 큰 장점이 있다(Lee et al., 2015). 최근 국제적으로 다양한 토마토 품종에 대한 WGRS가 수행되고 있는데(Causse et al., 2013; Sahu and Chattopadhyay, 2017), Lin et al. (2014)는 총 360개 토마토 핵심 유전자원(core collection)에 대해 WGRS를 완료하였으며, 염기서열 정보는 NCBI의 SRA (Sequence Read Archive; <https://www.ncbi.nlm.nih.gov/sra>) database에 공시되어 있다.

본 연구는 NCBI-SRA database에 공시된 토마토 유전체 염기서열 정보를 재분석하여 대량 SNP를 탐색하고 이를 이용하여 여교잡 육종 현장에서 활용 가능한 새로운 MABC SNP set를 개발하고자 수행되었다.

Materials and Methods

토마토 유전체 정보의 확보

토마토 유전체 정보는 NCBI의 SRA database로부터 수집하였으며, 본 연구를 위해 수집한 토마토의 계통 별 SRA 정보는 약 800여개, 총 염기서열 정보의 양은 약 7.8 Tb이다(Table 1). 토마토 표준 유전체(reference genome)는 SGN (<https://solgenomics.net/>)에서 수집한 *Solanum lycopersicum* Heinz 1706 (ITAG version 2.4)을 사용하였다. 표준 유전체의 총 길이는 781,666,411 bp이고, 12개 염색체와 12개 염색체 중 하나의 염색체로 자리 잡지 못한 scaffold 염기서열을 묶은 1개의 기타 group으로 구성된 13개의 염색체로 이루어져 있다.

토마토 계통의 선정

토마토 집단을 대변하는 계통을 선발하기 위해 Lin et al. (2014)의 논문을 참고하여 연구 및 분석에 사용 가능한 형태의 염기서열 데이터로서 1차적으로 토마토 234개 계통을 선정하였다(Table 2). 이들은 대과종 완숙용 또는 가공용 재배종으로 많이 이용되고 있는 *Solanum lycopersicum*을 우선으로 하여, 대표적인 방울토마토 재배종인 *S. lycopersicum* var. *cerasiforme*와 방울토마토 육종 과정에서 활용도가 높은 *S. pimpinellifolium*을 포함하였으며, 그 외 일부 야생종

토마토 계통으로 구성되었다. 또한, 선정된 계통을 위주로 교배조합을 작성하여 선발된 SNP 마커의 실효성 검증에 사용되도록 하였다.

토마토 genomic DNA 추출

유전체 분석을 통해 선발한 SNP 마커의 정확도를 검증하기 위해 'L3708' (*Solanum pimpinellifolium*)과 'AV107-4' (*Solanum lycopersicum*) 간 교배집단을 작성하였다. Genomic DNA를 추출하기 위해 토마토 잎 1 mg을 1.5 ml microcentrifuge tube에 넣고 액화질소로 냉동시킨 후, pestle을 이용하여 조직을 완전히 마쇄하였다. DNeasy Plant mini kit (Qiagen, USA)의 매뉴얼에 따라 진행하였으며, 추출한 DNA는 Simplinano™ spectrophotometer (GE Healthcare Life Sciences, UK)를 이용하여 농도와 순도를 측정한 다음 최종 농도가 25 ng· μ l⁻¹ 가 되도록 희석하였다.

토마토 MABC용 SNP 마커 선발

NCBI-SRA database에서 수집한 토마토 234 계통의 유전체 염기서열 정보로부터 haplotype 유전자 분석을 위한 tag-SNP 27,680좌를 선발하였다. 1차적으로 27,680좌의 tag-SNP로부터 group 전체의 missing data가 30% 이내이고 각

Table 1. Tomato genome data from NCBI-SRA.

Scientific name of tomato SRA	No. of SRA	Total bases (bp)
<i>Solanum lycopersicum</i>	557	5,367,216,027,209
<i>Solanum lycopersicum</i> var. <i>cerasiforme</i>	133	818,234,753,829
<i>Solanum pimpinellifolium</i>	65	470,510,067,600
<i>Solanum habrochaites</i>	17	324,511,688,762
<i>Solanum pennellii</i>	14	137,734,433,883
<i>Solanum arcanum</i>	12	165,974,167,372
<i>Solanum huaylasense</i>	3	101,084,736,600
<i>Solanum cheesmaniae</i>	2	68,609,466,800
<i>Solanum chilense</i>	2	64,731,490,900
<i>Solanum chmielewskii</i>	2	67,986,640,700
<i>Solanum neorickii</i>	2	70,608,085,300
<i>Solanum peruvianum</i>	2	69,515,482,100
<i>Solanum corneliomuelleri</i>	1	32,520,818,500
<i>Solanum galapagense</i>	1	30,910,824,200
Total	813	7,790,148,683,755

SRA, Short Read Archive database.

Table 2. SRA information used for analysis.

Scientific name of tomato SRA	No. of used SRA
<i>Solanum lycopersicum</i>	126
<i>Solanum lycopersicum</i> var. <i>cerasiforme</i>	68
<i>Solanum pimpinellifolium</i>	33
<i>Solanum habrochaites</i>	1
<i>Solanum cheesmaniae</i>	2
<i>Solanum chilense</i>	1
<i>Solanum neorickii</i>	1
<i>Solanum peruvianum</i>	2
Total	234

SRA, Short Read Archive database.

group 내에서 PIC (polymorphism information content) value가 0.2 이상인 tag-SNP를 filtering하였다. 각 tag-SNP의 PIC value는 SNP의 다형성 수준을 나타내는 지표로써 Botstein et al. (1980)를 참고하여 산출하였다.

$$PIC_i = 1 - \sum_{j=1}^n P_{ij}^2$$

한 염색체당 최소 10개의 SNP를 선발하고자 각 염색체의 길이를 5 Mb에서 10 Mb 간격으로 영역을 나누었으며, 계산식을 통해 걸러진 tag-SNP의 염색체 영역별 분포 양상을 확인하였다. 최종적으로 해당 영역에서 PIC value가 높고, 교배조합 내에서 다형성을 가장 많이 나타내는 SNP를 선발하였다.

교배조합을 이용한 SNP 마커 검증

각 SNP에 대한 Luna probe(unlabeled probe)를 디자인하여 GBS 결과의 신뢰성을 검증하였다. 해당 SNP의 특이적 증폭을 위해 약 300 bp 길이의 flanking 서열로부터 PCR primer와 Luna probe 세트를 제작하였다. Primer3 프로그램을 이용하여 AT-rich 혹은 GC-rich region을 피하고 GC 비율은 약 50%로 primer를 설계하였으며, 최종 PCR product의 크기는 해당 SNP가 포함된 150 bp 내외로 하였다. Luna probe는 SNP를 포함하여 총 31 mer로 하며, SNP 위치의 염기는 reference의 염기(기준 염기)로 제작하였다(Table 5). PCR 반응액은 genomic DNA (25 ng/μL) 1 μL와 함께 forward primer (1 pmol) 1 μL, reverse primer (10 pmole) 0.5 μL, 5x HOT FIREPol® EvaGreen® qPCR Supermix (Solis Biodyne, Estonia) 4 μL, luna probe (10 pmole) 1 μL, D.W 12.5 μL를 넣어 총 20 μL로 조성하였다. Real time PCR은 CFX96 (Bio-rad, USA) 기기를 이용하여 수행되었으며, PCR 조건은 95°C에서 5분 동안 1차 변성하고, 95°C에서 20초, 60°C에서 20초, 72°C에서 20초간 40 - 45회 반복하였다. 이어서 melting curve 분석을 위해 95°C에서 15초, 40°C에서 30초 반응한 다음, 50°C에서 90°C까지 1초당 0.5°C씩 온도를 상승시켰다. Melting curve 분석 시 기준이 되는 SNP 서열과 perfect match인 template DNA 샘플은 one miss match인 샘플보다 Tm 값이 높기 때문에 melting peak가 늦게 출현하게 되어 결과적으로 두 샘플을 구분할 수 있게 되며, 각 계통에 대한 GBS 분석 결과와 Luna probe 결과를 비교하여 일치율을 백분율로 계산하였다.

Results and Discussion

토마토 유전체 정보를 이용한 계통 선발 및 교배조합의 작성

1차적으로 선발된 234개의 토마토 accession의 경우 재배종으로 이용되고 있는 *Solanum lycopersicum* 이외에도 유전적 배경이 다양한 야생종들이 포함되어 있기 때문에, 선발된 MABC용 마커가 실제 토마토 육종 과정에서 활용 가능한지 여부를 확인할 필요가 있다. 따라서 234개의 계통 중에서 주요 육종 자원으로 이용되고 다양한 계통을 대변할 수 있는 교배조합의 작성이 매우 중요하다.

NCBI-SRA로부터 얻은 토마토 234개 품종 중에서 재배종으로 이용되는 *S. lycopersicum*과 *S. lycopersicum* var. *cerasiforme* 품종을 포함하여 유전체 데이터의 quality가 충분한 178개 품종을 2차로 선발하였다. 이 후 선발된 178개 품종 중 Lin et al. (2014)의 구분에 따른 대과종 그룹(group = BIG lines)과 소과종 그룹(group = CER lines)에 속하는 품종 94종과 68종, 총 162 품종을 재 선발하였으며, 최종적으로 ‘Cocktail tomato’, ‘Vintage cultivar’, ‘Latin American cultivar’ 등 17가지의 다양한 카테고리에 속하는 124종 (group = CER lines 39종, group = BIG lines 85종)을 선정하여 MABC용 SNP 마커 선발에 사용하였다(Table 3). 육종을 통하여 개량된 품종의 경우 품종명을 표기하였다.

MABC용 SNP 마커 선발

토마토 234 계통의 유전체 염기서열 데이터를 이용하여 haplotype 유전자를 분석하기 위해 tag-SNP 27,680개를 선발

Table 3. Information of 124 samples used for selection of MABC markers.(Continued)

No.	SRA	Variety	Group	Categories	Name (cultivar)
1	SRR1572327	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	cerise rose
2	SRR1572335	<i>S.lycopersicum</i> var cerasiforme	CER	Cocktail tomato	
3	SRR1572339	<i>S.lycopersicum</i> var cerasiforme	CER	Cocktail tomato	
4	SRR1572343	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	Ohmiya SunCerise
5	SRR1572344	<i>S.lycopersicum</i> var cerasiforme	CER	Vintage cultivar	Principe Borghese
6	SRR1572345	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	Poire jaune
7	SRR1572346	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	
8	SRR1572348	<i>S.lycopersicum</i> var cerasiforme	CER	Latin American cultivar; Wild species	Trujillo
9	SRR1572362	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	
10	SRR1572364	<i>S.lycopersicum</i> var cerasiforme	CER	Cocktail tomato	
11	SRR1572365	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	N1565
12	SRR1572372	<i>S.lycopersicum</i> var cerasiforme	CER	Latin American cultivar; Wild species	Moyobamba
13	SRR1572373	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	Cuba Plum
14	SRR1572375	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	Malintka 101
15	SRR1572377	<i>S.lycopersicum</i> var cerasiforme	CER	Latin American cultivar	Villa Hermosa
16	SRR1572380	<i>S.lycopersicum</i> var cerasiforme	CER	Latin American cultivar	Veracruz
17	SRR1572383	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	Monplaisir
18	SRR1572384	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	Nagcarlang
19	SRR1572385	<i>S.lycopersicum</i> var cerasiforme	CER	Latin American cultivar	
20	SRR1572386	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	8 bis
21	SRR1572390	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	Phyra
22	SRR1572391	<i>S.lycopersicum</i> var cerasiforme	CER	landrace	Mobalcon
23	SRR1572396	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	Atom
24	SRR1572400	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	L 285
25	SRR1572401	<i>S.lycopersicum</i> var cerasiforme	CER	Latin American cultivar	
26	SRR1572406	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	N 795 Pescio
27	SRR1572409	<i>S.lycopersicum</i> var cerasiforme	CER	cocktail tomato	
28	SRR1572413	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	Minibel
29	SRR1572417	<i>S.lycopersicum</i> var cerasiforme	CER	landrace	Linosa
30	SRR1572424	<i>S.lycopersicum</i> var cerasiforme	CER	landrace	Cerazjnho
31	SRR1572425	<i>S.lycopersicum</i> var cerasiforme	CER	landrace	Allungato piccolo
32	SRR1572426	<i>S.lycopersicum</i> var cerasiforme	CER	Cocktail tomato	
33	SRR1572427	<i>S.lycopersicum</i> var cerasiforme	CER	Latin American cultivar	
34	SRR1572430	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	N 933
35	SRR1572442	<i>S.lycopersicum</i> var cerasiforme	CER	Latin American cultivar	Tarapoto
36	SRR1572446	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	
37	SRR1572449	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	BL 587-S
38	SRR1572450	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	BL 587
39	SRR1572451	<i>S.lycopersicum</i> var cerasiforme	CER	cultivar	BL 359
40	SRR1572452	<i>S.lycopersicum</i>	BIG	Vintage Processing	VF-36
41	SRR1572453	<i>S.lycopersicum</i>	BIG	Vintage Fresh Market	Moneymaker
42	SRR1572455	<i>S.lycopersicum</i>	BIG	Inbreed line	Hawaii 7998
43	SRR1572456	<i>S.lycopersicum</i>	BIG	Vintage Fresh Market	Edkawi
44	SRR1572457	<i>S.lycopersicum</i>	BIG	Vintage Processing	San Marzano
45	SRR1572458	<i>S.lycopersicum</i>	BIG	Modern Fresh Market	Micro-Tom
46	SRR1572460	<i>S.lycopersicum</i>	BIG	Vintage Fresh Market	AilsaCraig
47	SRR1572461	<i>S.lycopersicum</i>	BIG	Vintage Fresh Market	Marglobe
48	SRR1572462	<i>S.lycopersicum</i>	BIG	Modern Fresh Market	KR2

Table 3. Information of 124 samples used for selection of MABC markers.(Continued)

No.	SRA	Variety	Group	Categories	Name (cultivar)
49	SRR1572463	<i>S.lycopersicum</i>	BIG	Modern Fresh Market	yoku improvement
50	SRR1572464	<i>S.lycopersicum</i>	BIG	Cocktail tomato	
51	SRR1572465	<i>S.lycopersicum</i>	BIG	Processing tomato	
52	SRR1572466	<i>S.lycopersicum</i>	BIG	Fresh Market	Moneymaker
53	SRR1572467	<i>S.lycopersicum</i>	BIG	Landrace	A pera abruzzese
54	SRR1572479	<i>S.lycopersicum</i>	BIG	Vintage Fresh Market	New Yorker
55	SRR1572482	<i>S.lycopersicum</i>	BIG	Latin American cultivar	Huachinango
56	SRR1572484	<i>S.lycopersicum</i>	BIG	Fresh Market	N739
57	SRR1572485	<i>S.lycopersicum</i>	BIG	Processing tomato	
58	SRR1572487	<i>S.lycopersicum</i>	BIG	Processing tomato	
59	SRR1572488	<i>S.lycopersicum</i>	BIG	Processing tomato	
60	SRR1572489	<i>S.lycopersicum</i>	BIG	Processing tomato	
61	SRR1572495	<i>S.lycopersicum</i>	BIG	Processing tomato	
62	SRR1572496	<i>S.lycopersicum</i>	BIG	Cocktail tomato	
63	SRR1572497	<i>S.lycopersicum</i>	BIG	Processing tomato	
64	SRR1572498	<i>S.lycopersicum</i>	BIG	Fresh Market	Moneymaker
65	SRR1572499	<i>S.lycopersicum</i>	BIG	Processing	
66	SRR1572514	<i>S.lycopersicum</i>	BIG	Landrace	Scatolone di bolsena
67	SRR1572515	<i>S.lycopersicum</i>	BIG	Landrace/ Latin American cultivar	Hacienda Calera
68	SRR1572516	<i>S.lycopersicum</i>	BIG	Vintage Processing	Pearson
69	SRR1572517	<i>S.lycopersicum</i>	BIG	Modern Fresh Market	Severianin
70	SRR1572518	<i>S.lycopersicum</i>	BIG	Vintage Fresh Market	Primabel
71	SRR1572530	<i>S.lycopersicum</i>	BIG	Landrace/ Latin American cultivar	Hacienda Rosario
72	SRR1572532	<i>S.lycopersicum</i>	BIG	Landrace	Spagnoletta
73	SRR1572533	<i>S.lycopersicum</i>	BIG	Vintage cultivar	Red Setter
74	SRR1572534	<i>S.lycopersicum</i>	BIG	cultivar	149-77
75	SRR1572535	<i>S.lycopersicum</i>	BIG	cultivar	Saladette
76	SRR1572536	<i>S.lycopersicum</i>	BIG	Vintage cultivar	Roma
77	SRR1572537	<i>S.lycopersicum</i>	BIG	Modern Fresh Market	Florida7547
78	SRR1572545	<i>S.lycopersicum</i>	BIG	Landrace/ Latin American cultivar	Santa Cruz B
79	SRR1572548	<i>S.lycopersicum</i>	BIG	Vintage Fresh, Market,Monogenic	Condine Red
80	SRR1572549	<i>S.lycopersicum</i>	BIG	Processing tomato	
81	SRR1572550	<i>S.lycopersicum</i>	BIG	Processing tomato	
82	SRR1572551	<i>S.lycopersicum</i>	BIG	Processing tomato	
83	SRR1572552	<i>S.lycopersicum</i>	BIG	Processing tomato	
84	SRR1572558	<i>S.lycopersicum</i>	BIG	Landrace/ Latin American cultivar	Tegucigalpa
85	SRR1572561	<i>S.lycopersicum</i>	BIG	Landrace	Cuor di bue
86	SRR1572564	<i>S.lycopersicum</i>	BIG	Processing tomato	
87	SRR1572566	<i>S.lycopersicum</i>	BIG	Processing tomato	
88	SRR1572567	<i>S.lycopersicum</i>	BIG	Processing tomato	
89	SRR1572568	<i>S.lycopersicum</i>	BIG	Processing tomato	
90	SRR1572569	<i>S.lycopersicum</i>	BIG	Processing tomato	
91	SRR1572570	<i>S.lycopersicum</i>	BIG	Processing tomato	
92	SRR1572573	<i>S.lycopersicum</i>	BIG	Processing tomato	
93	SRR1572574	<i>S.lycopersicum</i>	BIG	Processing tomato	
94	SRR1572583	<i>S.lycopersicum</i>	BIG	Landrace	Francescano
95	SRR1572584	<i>S.lycopersicum</i>	BIG	Vintage cultivar	Severianin
96	SRR1572585	<i>S.lycopersicum</i>	BIG	Landrace	Pantano dArdea

Table 3. Information of 124 samples used for selection of MABC markers.

No.	SRA	Variety	Group	Categories	Name (cultivar)
97	SRR1572587	<i>S.lycopersicum</i>	BIG	Vintage Fresh Market	Ace
98	SRR1572588	<i>S.lycopersicum</i>	BIG	Processing tomato	N020212
99	SRR1572589	<i>S.lycopersicum</i>	BIG	Landrace	Libanese
100	SRR1572593	<i>S.lycopersicum</i>	BIG	Landrace/ Latin American cultivar	San Salvador
101	SRR1572594	<i>S.lycopersicum</i>	BIG	Landrace	Bell pepper-like
102	SRR1572595	<i>S.lycopersicum</i>	BIG	Modern Fresh Market	Florida7060
103	SRR1572596	<i>S.lycopersicum</i>	BIG	Vintage Fresh Market	PrinceBorghese
104	SRR1572597	<i>S.lycopersicum</i>	BIG	Modern Fresh Market	NC265-1(93)-3-3
105	SRR1572607	<i>S.lycopersicum</i>	BIG	cultivar	Barnaulski Konservnyi
106	SRR1572610	<i>S.lycopersicum</i>	BIG	Processing tomato	
107	SRR1572611	<i>S.lycopersicum</i>	BIG	cultivar	Microtom
108	SRR1572615	<i>S.lycopersicum</i>	BIG	Processing	
109	SRR1572616	<i>S.lycopersicum</i>	BIG	Processing tomato	
110	SRR1572617	<i>S.lycopersicum</i>	BIG	Processing tomato	
111	SRR1572618	<i>S.lycopersicum</i>	BIG	Processing tomato	
112	SRR1572619	<i>S.lycopersicum</i>	BIG	Vintage Fresh Market	Platense
113	SRR1572620	<i>S.lycopersicum</i>	BIG	Modern Fresh Market	NC EBR-5
114	SRR1572623	<i>S.lycopersicum</i>	BIG	Latin American cultivar	Ayacucho
115	SRR1572625	<i>S.lycopersicum</i>	BIG	Processing tomato	
116	SRR1572626	<i>S.lycopersicum</i>	BIG	Landrace/ Latin American cultivar	Merida
117	SRR1572629	<i>S.lycopersicum</i>	BIG	Processing tomato	
118	SRR1572630	<i>S.lycopersicum</i>	BIG	Latin American cultivar	
119	SRR1572631	<i>S.lycopersicum</i>	BIG	Processing tomato	
120	SRR1572654	<i>S.lycopersicum</i>	BIG	Landrace	Ucra2
121	SRR1572655	<i>S.lycopersicum</i>	BIG	Processing tomato	
122	SRR1572656	<i>S.lycopersicum</i>	BIG	Processing tomato	
123	SRR1572658	<i>S.lycopersicum</i>	BIG	Cocktail/ Processing tomato	
124	SRR1572666	<i>S.lycopersicum</i>	BIG	Landrace	

MABC, Marker-assisted Backcrossing; CER, *S. lycopersicum* var. *cerasiforme* accessions; BIG, big-fruited *S. lycopersicum* accessions.

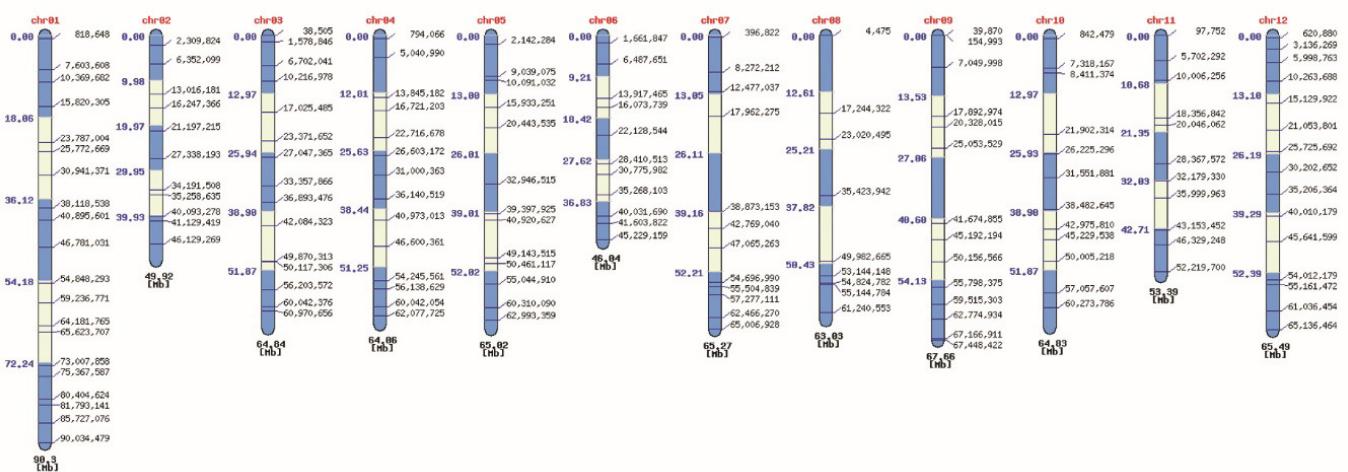


Fig. 1. Mapping the 143 MABC markers on tomato chromosome. Bars represent twelve tomato chromosomes, and the length of the chromosome (Mb) is indicated below each bar. On the left side, chromosomal regions were shown. On the right side, the names and positions of selected MABC markers were shown. MABC, Marker-assisted Backcrossing.

한 후, filtering 과정을 거쳐 총 4,274개를 1차적으로 선발하였다. 그 다음으로 각 염색체를 5 Mb 단위로 영역을 나누었으며, 해당 영역에서 높은 PIC value를 갖는 143개 tag-SNP를 최종적으로 선발하였다(Table 4, Supplemental Table 1). 최종 선발된 143개 마커의 염색체 별 분포 양상을 재확인하기 위해 표준 유전체를 이용하여 mapping 분석을 수행한 결과, 12개의 토마토 염색체 상에 전반적으로 고루 분포하는 것을 확인할 수 있었다(Fig. 1).

효과적인 MABC를 위해서는 먼저 유전체 전체적(genome-wide)으로 분포하고 있는 분자 마커를 확보해야 하며, 교배조합에 따라 적용 가능한 마커를 쉽게 확인할 수 있어야 한다. 또한 genotyping에 필요한 분자 마커만을 선발하여 분석에 소요되는 비용을 최소화하여야 한다(Kim et al., 2013). 따라서 본 연구에서는 단순히 SNP를 임의로 선발하여 사용하던 기존의 방법과 달리, 인간 유전체 분야에서만 사용되던 tag-SNP를 사용함으로써 각 교배집단들의 유전체 조성을 좀 더 효율적으로 대변하고자 하였다. Tag-SNP 기술은 유전자 영역에 존재하는 SNP 중 염색체 블록(block) 단위에서 넓은 범위로 일차 선발하고, 그 중에서 아미노산의 치환으로 단백질 구조 변화를 야기하여 유전자 기능에 영향을 미칠 가능성에 높은 비동의 SNP (non-synonymous SNP)를 이용하는 것을 말한다.

MABC용 SNP 마커 토마토 개체 내 다형성 분석

다형성 지수(PIC)가 높은 마커일수록 개체의 식별뿐만 아니라 유전적 다양성을 평가하기에 매우 유용하다. 최종 선발된 143개 MABC용 tag-SNP 마커의 PIC value는 평균 0.369로 비교적 높고 모든 염색체에 대하여 균일한 값을 나타내었다(Table 4). 또한 선발된 SNP의 MABC용 마커로서의 활용 가능성을 평가하기 위해 124개 품종 간 교배조합을 가상으로 작성(pairwise cross-combination matrix)하고 각 교배조합 당 다형성을 보이는 SNP의 수를 확인하였다(Fig. 2). SNP의 수는 pairwise box의 색도차로 나타내었는데 색이 짙을수록 더 높은 SNP 수를 의미한다. SNP의 수가 많을수록 특정 교배조합의 양친 간 유전적 거리가 크다는 것을 짐작할 수 있는데, *S. lycopersicum*와 *S. lycopersicum* var. *cerasiforme* 품종 간 조합의 경우 pairwise box의 색이 비교적 짙다는 사실에서 증명된다. 전반적으로 선발된 143개의 MABC 마커 대부분이 124개 품종 간에 높은 다형성을 나타내는 것을 확인할 수 있었다(Fig. 2A).

특히 124 품종 중 *S. lycopersicum* (Group = BIG) 14개, *S. lycopersicum* var. *cerasiforme* (Group = CER) 5개를 선발하여 재분석 하였을 경우(Fig. 2B) 동일 그룹 내의 품종 간 교배조합에서도 높은 SNP 수가 존재하였으며, 이는 그룹 간

Table 4. Summary of MABC.

Chr.#	No. of candidate primer	No. of selected SNP marker	Intergenic	Genic (Promoter + Intron + CDS)	Average PIC value
1	3488	19	-	19	0.374
2	2780	10	-	10	0.372
3	2612	13	-	13	0.365
4	2394	13	-	13	0.370
5	2066	12	-	12	0.370
6	2214	10	-	10	0.371
7	1963	11	-	11	0.355
8	2027	8	-	8	0.368
9	1911	11	-	11	0.373
10	1963	11	-	11	0.370
11	1944	11	-	11	0.365
12	1948	14	-	14	0.373
Total	27310	143	-	143	0.369

MABC, Marker-assisted Backcrossing; SNP, Single Nucleotide Polymorphism; CDS, Coding sequence; PIC, Polymorphism Information Content.

교배조합의 경우와 크게 다르지 않았다. 이는 선발된 SNP가 대과종(*S. lycopersicum*)과 소과종(*S. lycopersicum* var. *cerasiforme*) 간 MABC 뿐만 아니라 동일한 아종 내 품종 간 MABC에도 효과적으로 활용될 수 있음을 의미하는 결과이며, 실제 다양한 교배조합에 공통적으로 적용될 가능성이 매우 높은 SNP set을 임을 보여준다. 단 국내 상용 F1 품종 육성에 활용되는 토마토 유전자원의 유전적 다양성이 본 연구의 NCBI-SRA DB로부터 선발한 품종에서보다 낮을 것임을 예측할 때 다양한 국내 F1 품종을 이용한 SNP set의 적용성 검증이 향후 필요하리라 본다.

Luna probe를 이용한 SNP 마커 정확성 검증

SNP 신뢰성 검증을 위해 *Solanum lycopersicum* 과 *S. pimpinellifolium*에 속하는 품종 간의 교배조합에서 다형성을 나타내는 SNP를 염색체 당 1 - 2개 선발하였다. MABC 마커의 경우 PIC value가 0.3 이상인 값을 기준으로 하여 filtering 과정을 수행하였지만, Luna probe 검증에 사용한 SNP의 경우 다양한 PIC value (0.02 - 0.38)를 가지는 것으로 선발하였다(Table 5). 이렇게 선발된 18개의 마커를 이용하여 교배조합의 양친 2 품종과 F1 5 품종에 대하여 genotyping

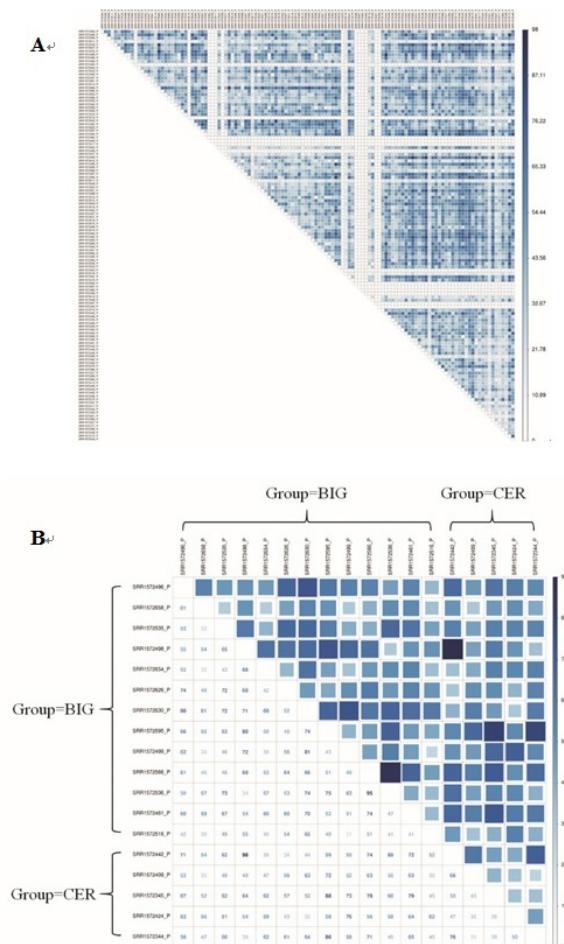


Fig. 2. Analysis of polymorphism in diverse tomato cultivars using 143 MABC markers. (A) Analysis of polymorphism in 124 lines using 143 MABC markers. On the X-axis and Y-axis, 124 individual lines were listed in order. When crossbreeding combinations were made for each individual, the number of markers in the 143 selected markers was shown to indicate the polymorphism. The more the number of markers showing the polymorphism is, the more the color of the box is set to appear dark. (B) Analysis of polymorphism in representative tomato 19 cultivars using markers. A total of 19 representative cultivars were selected from 124 tomato lines to verify the polymorphism of each crossbreeding combination. There are composed of 14 cultivars of *Solanum lycopersicum* (Group = BIG) and 5 cultivars of *S. lycopersicum* var. *cerasiforme* (Group = CER). MABC, Marker-assisted Backcrossing.

을 수행하였고, 그 결과 모든 마커에 대하여 melting peak가 구분되는 것을 확인하였다. 특히, NCBI-SRA의 염기서열 정보를 분석하여 얻은 SNP에 대해 동형접합(homozygosity)의 유전자형을 보인 품종의 경우 정확하게 단일 peak가 나타난 반면, 이형접합(heterozygosity)이었던 품종들의 melting peak는 완만한 곡선을 보이며 두 대립유전자형에 대한 이중성을 나타냈다(data not shown). 결과적으로 7개 품종에 대해 18개의 Luna probe 마커를 검증하여 총 126개의 genotype data point를 얻었고, 그 중 125개의 SNP data point에서 염기서열정보와 Luna probe 검증 결과가 일치하는

Table 5. Information of Luna probe used for SNP genotyping.

No.	Chr	Position	Allele	PIC	Luna probe (unlabeled probe)			Product size (bp)
					Primer sequence (5'-3')		Probe sequence (3'-phosphate)	
					Forward	Reverse		
1	1	4686399	T/G	0.02	ACT CTG CCT CTG CTG GGA TCT TCC TTA TGC GAT CG	CGG ATC	TGC ATG ACT AAT CTG TTA AGG AAG AAG GCT T	156
2	1	80404624	T/C	0.02	CTC CCT TCT AGT GTA ATA TCA G	CCC TGA ATC TGC AAC TTT CTG	TTG CAA ACG GCT TCT TCT GTC TCT GAG AAG C	139
3	2	31384147	C/G	0.02	CCC CTT AAC CAC TTC ACT AGA	AAG GGC GGA GCT GAA GGG A	AGC CCC AAA AAA TCT CTT TTT GCC TTA TTT G	157
4	2	41129419	G/C	0.37	ATG CAG CCA ATC TAA GAT GGT TTA TTA TCC CTA CGA A	TGT TCT GTC	TGG TGT TGG ACA AGA GTT AGT GAA CAA CC CT	151
5	3	1578846	A/C	0.38	AAA GTT GTT GGT CCT GTT GCT G	CTG CAG CAC ATT AAC AGA ATA C	GTG TGG CAC AAC TGC AAT GGT ATG ACT GTG G	154
6	3	60970656	C/T	0.37	GAG AAT CTA AAG AGC CAC TGA A	AAA TTG CAC AGT AAA CTT GAG GA	ATT ATT GAA TTT TAT CCC AAT ACT GAT AGC	145
7	4	59090349	A/G	0.02	CTC AAA TTT TCG CAG TGC CTC A	AGG AAC CAA GAT CTC CTG CTG	TAC ATT GTG TAT GTC ATT GTT GTT ATT CTA T	154
8	4	62077725	T/C	0.29	AGG ACC TTG AAT TGC ACC TTG T	GGA TCT AAA TGG CAG TTT GGG	ATA TAG AAC TAG GTC TTC CGG AGA TGA CTG T	153
9	5	60310090	T/C	0.02	TCA TGG CTT CTG GTA AGT CCT A	CAT CCA CAC AAC AAC CTG TCA TA	TGA TTT TAT GGG GTT TCG TTA GTA TGT TTC T	152
10	6	40031690	G/T	0.02	CTT TTG TAC ATC ATC GTT AGC CAC AAT CTA CTG CAT G	ACA AGG ACA TGA GTT GAA ACA TTA AG	ACA AGG ACA TGA GTT GAA AAA GTA TCT TCA G	144
11	6	44630172	G/A	0.02	CAC CTC TGC ATA TTG CTT TGG	GGT GGT AAC TGA GAA GGG ATT	GTA AAT GCT GCT GGC ATA GAA GAT ACA GCT A	151
12	7	57277111	G/C	0.37	CTT TTG TAC ATC ATC GTT AGC CAC AAT CTA CTG CAT G	AAA CGG AAT ATA AAA GCC ACA TTA AG	TGT TGC TTT TTC C	150
13	8	53144148	G/A	0.02	GTA GTA CTA ATA ATG AAG AGC TC	CAG CAG CAG AAG TAG ATA TTG AA	TGT GCG CCA TTT CCA GCA TTT CCA TTG CTG T	152
14	9	154993	A/G	0.37	ATC CCT TGC ACC TTG ATC TGC ATC CCT AAT AGT ATT T	GTC TTC A	GTA CAA ATA AGC ATT ACG TTT GAT CGT GTG T	152
15	9	55798375	A/G	0.37	AGA CTG TTG CTG CAG CTG GG	GTC ACT GAA AAT ATG TGG TCT TA	GAG CTT GAT ACC TAT AAA AAT GAA GTT GTA T	151
16	9	67448422	T/C	0.29	TGA TTC CCA GTG GCA AAT GTA C	TTC GTA TGA GTT TCC CAT TTG C	ACC AAC GAT TGA AGC TAC CGT CTG TTA ACA G	154
17	10	7318167	A/G	0.02	TAG GTT TAG CAG CGA TTT TGT G	CAA ACC ACC ACC TGA AAC CC	TGG TTG TTT GCA AAA ACT CAC TCA AAA GGG A	156
18	12	3136269	C/T	0.02	GTG GTC AGT CTT ACA CAG AC	GAA GCA GCA ACA GAC AAA GTG	AGA ACA CAC AAT CTG CCT CTA CAT TTC AGT T	151

SNP, Single Nucleotide Polymorphism; PIC, Polymorphism Information Content.

것을 확인할 수 있었다(Table 6). 따라서 본 연구팀의 분석 파이프라인을 통한 NCBI-SRA database의 염기서열정보 분석 및 SNP 탐색 결과가 매우 신뢰성이 있음을 알 수 있었다.

Real-time PCR을 이용한 HMR 방식의 SNP 유전자형 분석법은 CAPS 또는 dCAPS와 같은 gel-based 유전자형 분석법에 비해 자동화가 가능하고 시간과 비용을 절감할 수 있다. 본 연구에서는 유전체 분석을 통해 선발된 SNP 마커의 정확성을 검증하고, 이를 실제 육종에 적용할 수 있을지 판단하기 위해 Luna probe (unlabeled probe)와 Eva green을 이용하는 melting curve 분석에 기반한 Luna probe genotyping 시스템을 구축하였다. Luna probe는 FRET (fluorescence resonance energy transfer) probe 방식을 이용하며, 3'말단에 phosphate를 처리하여 polymerase에 의한 nucleotide 합성을 방지하고, 해당 SNP의 결합 여부에 따라 melting peak에서 나타나는 Tm 값의 차이로 SNP 타입을 구분할 수 있게 된다(Zhou et al., 2004). 특히 NGS에 기반한 Genotyping-by-sequencing (GBS) 분석 기술이 빠르게 확산되면서 토마토를 비롯한 여러 작물에 대하여 폭넓은 연구가 이루어지고 있으며(Elshire et al., 2011; Labate et al., 2014; Jang and Oh, 2017), 이에 따라 다양한 형질과 연관된 SNP가 꾸준히 개발되고 있다. 그럼에도 불구하고 대량의 SNP를 정확하고 효율적으로 검증할 수 있는 시스템은 미진한 상태이다. 현재까지 Luna probe를 SNP genotyping에 적용한 사례가 많지는 않지만(Liew et al., 2007; Costa et al., 2012; Birrer et al., 2014; Zhong et al., 2016), 이용의 편의성과 결과의 신뢰성을 고려할 때 식물에서도 다양한 조건의 SNP와 indel 등의 염기서열변이 판별에 충분히 활용 가능하리라 판단된다.

이와 같이 본 연구에서는 공시된 대량 토마토 유전체 염기서열 정보를 재분석하여 토마토 MABC 육종에 활용 가능한 tag-SNP set를 개발하였다. 또한 여교잡 육종 현장에서 활용될 수 있는 가상의 양친을 대상으로 개발된 tag-SNP set의 적용성을 검증하였다. 본 연구를 통해 개발된 MABC 마커 세트는 토마토 여교배집단에서 반복친의 염색체 회복율이 높은 후대를 조기에 효율적으로 선발할 수 있을 뿐만 아니라 품종 식별과 F1 품종의 순도 검정용 마커로도 활용 가능하리라 본다.

Table 6. Comparison of SNP genotyping results using 18 Luna probe.

No.	Chr	Position	Allele	Comparison of result (Sequencing / Luna probe)							Equal rate (%)
				1	2	3	4	5	6	7	
1	1	4686399	T/G	G/G	T/T	T/T	T/T	T/T	T/T	T/T	100
2	1	80404624	T/C	T/T	C/C	T/T	T/T	T/T	T/T	T/T	100
3	2	31384147	C/G	C/C	G/G	C/C	C/C	C/C	C/C	C/C	100
4	2	41129419	G/C	C/C	G/G	C/C	S/S	S/S	S/S	S/S	100
5	3	1578846	A/C	A/A	C/C	M/M	A/A	A/A	C/C	A/A	100
6	3	60970656	C/T	T/T	C/C	Y/Y	Y/Y	T/T	Y/Y	C/C	100
7	4	59090349	A/G	A/A	G/G	G/G	G/G	G/G	G/G	GG	100
8	4	62077725	T/C	C/C	T/T	Y/Y	T/T	C/C	C/C	Y/Y	100
9	5	60310090	T/C	T/T	C/C	C/C	C/C	C/C	C/C	C/C	100
10	6	40031690	G/T	T/T	G/G	T/T	T/T	T/T	T/T	T/T	100
11	6	44630172	G/A	G/G	A/A	A/A	A/A	A/A	A/A	A/A	100
12	7	57277111	G/C	C/C	G/G	G/G	S/S	G/G	-	C/C	100
13	8	53144148	G/A	A/A	G/G	G/G	G/G	G/G	G/G	G/G	100
14	9	154993	A/G	A/A	G/G	G/G	G/G	R/R	R/R	A/A	100
15	9	55798375	A/G	A/A	G/G	G/G	A/A	G/G	R/R	A/A	100
16	9	67448422	T/C	C/C	T/T	Y/Y	C/C	Y/Y	Y/Y	C/C	100
17	10	7318167	A/G	G/G	A/A	G/G	G/G	G/G	G/G	G/G	100
18	12	3136269	C/T	C/C	T/T	C/C	C/C	C/C	C/C	C/C	100

'-' means no read.

SNP, Single Nucleotide Polymorphism.

Conclusion

분자마커이용여교잡(MABC)은 분자 마커를 이용하여 유묘기에 여교잡 자손의 염색체 전체 조성을 확인함으로써 육종 연한과 비용, 노력 등을 효율적으로 감소시킬 수 있어 목표 형질 유전자의 선발 육종에 매우 적합하다. 그러나 벼를 비롯한 작물을 대상으로 다양하게 연구되고 있는 것과 달리, 실제 육종 산업에 적용될 수 있는 MABC 기술은 매우 미흡하다고 볼 수 있다. 본 연구에서는 토마토 유전자 단위에서 haplotype을 구분하는 tag-SNP 27,680개를 선발하여 마커 선발에 이용하였으며, 실제 토마토의 MABC 육종에 활용될 수 있는 마커 세트를 개발하기 위해 수행되었다. 234개의 SRA 토마토를 이용하여 염색체 내 일정 간격(5 - 10 Mb)으로 영역을 나눈 후, filtering을 거쳐 PIC value 가 높은 SNP로써 총 143개의 MABC용 후보 SNP를 선발하였다. 이렇게 선발된 143개 MABC 마커는 유전적으로 다양한 계통으로 이루어진 교배집단 124 계통 내에서 높은 다형성을 나타내는 것을 확인하였다. 또한, 해당 MABC 마커의 활용 가능성을 검증하기 위해 선발된 18개의 MABC용 SNP 마커와 Luna probe를 이용하여 real-time PCR 분석을 수행하였고, 결과적으로 염기 서열 분석 결과와 실험적 결과가 100% 일치하는 것을 확인할 수 있었다. 따라서 본 연구를 통해 선발된 MABC 마커 세트는 실제 다양한 여교잡 교배조합에 활용될 수 있는 가능성을 높임과 동시에 토마토 품종 육성에 있어서 경쟁력을 크게 향상시킬 수 있으므로 상업적으로 매우 유용하게 이용될 수 있을 것으로 기대된다.

Acknowledgement

본 연구는 충남대학교 학술연구비에 의해 지원되었음.

References

- Bae JH, Han Y, Jeong HJ, Kwon JK, Chae Y, Choi HS, Kang BC. 2010. Development of a SNP marker set for tomato cultivar identification. *Korean Journal of Horticultural Science and Technology* 28:627-637. [in Korean]
- Birrer M, Kolliker R, Manzanares C, Asp T, Studer B. 2014. A DNA marker assay based on high-resolution melting curve analysis for distinguishing species of the *Festuca-Lolium* complex. *Molecular Breeding* 34:421-429.
- Botstein D, White RL, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Society of Human Genetics* 32:314-331.
- Causse M, Desplat N, Pascual L, Paslier MCL, Sauvage C, Bauchet G, Berard A, Bounon R, Tchoumakov M, Brunel D, Bouchet JP. 2013. Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. *BMC Genomics* 14:791.
- Collard BC, Mackill DJ. 2008. Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B* 363:557-572.
- Costa J, Mafra I, Oliverira MBPP. 2012. High resolution melting analysis as a new approach to detect almond DNA encoding for Pru du 5 allergen in foods. *Food Chemistry* 133:1062-1069.
- Edwards D, Batley J. 2010. Plant genome sequencing: Application for crop improvement. *Plant Biotechnology Journal* 8:2-9.
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. 2011. A robust, simple genotyping- by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379.

- Foolad MR. 2007. Genome mapping and molecular breeding of tomato. International Journal of Plant Genomics Article 2007: 64358. doi: 10.1155 / 2007 / 64358.
- Hwang JH, Kim HJ, Chae Y, Choi HS, Kim MK, Park YH. 2012. Evaluation of germplasm and development of SSR markers for marker-assisted backcross in tomato. Korean Journal of Horticultural Science and Technology 30:557-567. [in Korean]
- Jang HA, Oh SK. 2017. Development of an efficient genotyping-by-sequencing (GBS) library construction method for genomic analysis of grapevine. Korean Journal of Agricultural Science 44:495-503. [in Korean]
- Kim JE, Lee BW, Kim SM, Lee BM, Lee JH, Jo SH. 2013. Genome-wide SNP database for marker-assisted background selection in tomato. Korean Journal of Breeding Science 45:232-239. [in Korean]
- Labate JA, Robertson LD, Strickler SR, Mueller LA. 2014. Genetic structure of the four wild tomato species in the *Solanum peruvianum* s.l. species complex. Genome 57:169-180.
- Lee YG, Jeong N, Kim JH, Lee K, Kim KH, Pirani A, Ha BK, Kang ST, Park BS, Moon JK, Kim N, Jeong SC. 2015. Development, validation and genetic analysis of a large soybean SNP genotyping array. The Plant Journal 81:625-636.
- Liew M, Seipp M, Durtschi J, Margraf RL, Dames S, Erali M, Voelkerding K, Wittwer C. 2007. Closed-tube SNP genotyping without labeled probes. American Journal of Clinical Pathology 127:341-348.
- Lin T, Zhu G, Zhang J, Xu X, Yu Q, Zheng Z, Zhang Z, Lun Y, Li S, Wang X, Huang Z, Li J, Zhang C, Wang T, Zhang Y, Wang A, Zhang Y, Lin K, Li C, Xiong G, Xue Y, Mazzucato A, Causse M, Fei Z, Giovannoni JJ, Chetelat RT, Zamir D, Städler T, Li J, Ye Z, Du Y, Huang S. 2014. Genomic analyses provide insights into the history of tomato breeding. Nature Genetics 46:1220-1226.
- Park PH, Chae Y, Kim HR, Chung KH, Oh DG, Kim KT. 2010a. Development of a SCAR marker linked to Ph-3 in *Solanum* ssp. Korean Journal of Breeding Science 42:139-143.
- Park YH, Kim KH, Choi YM, Choi HS, Chae Y, Park KS, Chung SM. 2010b. Evaluation of TYLCV-resistant tomato germplasm using molecular markers. Korean Journal of Horticultural Science and Technology 28:89-97.
- Sahu KK, Chattopadhyay D. 2017. Genome-wide sequence variations between wild and cultivated tomato species revisited by whole genome sequence mapping. BMC Genomics 18:430
- Servin B, Hospital F. 2002. Optimal positioning of markers to control genetic background in marker-assisted backcrossing. Journal of Heredity 93:214-217.
- The Tomato Genome Consortium. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635-641.
- Zhong WL, Wang L, Zhang J, Chen XF, Zhang W, Dou X, Yu B. 2016. Development of unlabeled probe based high-resolution melting analysis for detection of filaggrin gene mutation c.3321delA. Journal of Clinical Laboratory Analysis 30:892-896.
- Zhou L, Myers AN, Vandersteen JG, Wang L, Wittwer CT. 2004. Closed-tube genotyping with unlabeled oligonucleotide probes and a saturating DNA dye. Clinical Chemistry 50:1328-1335.

Supplemental Table 1. Information of MABC 143 markers.(Continued)

No.	SNP ID	Chr	Position	Gene	Allele	PIC value
1	1-1	SL2.40ch01	818648	Solyc01g006180.2	T/C	0.38
2	1-2	SL2.40ch01	7603608	Solyc01g011090.2	A/G	0.38
3	1-3	SL2.40ch01	10369682	Solyc01g013910.1	A/C	0.37
4	1-4	SL2.40ch01	15820305	Solyc01g014910.1	A/G	0.37
5	1-5	SL2.40ch01	23787004	Solyc01g017040.1	T/C	0.38
6	1-6	SL2.40ch01	25772669	Solyc01g017850.1	T/C	0.38
7	1-7	SL2.40ch01	30941371	Solyc01g020460.1	T/C	0.37
8	1-8	SL2.40ch01	38118538	Solyc01g044270.2	T/C	0.38
9	1-9	SL2.40ch01	40895601	Solyc01g049680.2	A/G	0.37
10	1-10	SL2.40ch01	46781031	Solyc01g056370.2	A/C	0.37
11	1-11	SL2.40ch01	54848293	Solyc01g057580.1	G/C	0.38
12	1-12	SL2.40ch01	59236771	Solyc01g058360.1	T/C	0.38
13	1-13	SL2.40ch01	64181765	Solyc01g065520.1	C/T	0.38
14	1-14	SL2.40ch01	65623707	Solyc01g066170.1	A/C	0.37
15	1-15	SL2.40ch01	73007858	Solyc01g081210.2	A/G	0.38
16	1-16	SL2.40ch01	75367587	Solyc01g089900.2	T/G	0.37
17	1-17	SL2.40ch01	81793141	Solyc01g099910.2	G/A	0.38
18	1-18	SL2.40ch01	85727076	Solyc01g105940.2,Solyc01g105950.2	T/C	0.37
19	1-19	SL2.40ch01	90034479	Solyc01g112220.2	C/T	0.35
20	2-1	SL2.40ch02	2309824	Solyc02g005350.2	A/G	0.37
21	2-2	SL2.40ch02	6352099	Solyc02g014130.1	G/A	0.37
22	2-3	SL2.40ch02	13016181	Solyc02g021240.2	T/G	0.37
23	2-4	SL2.40ch02	16247366	Solyc02g030360.1	A/G	0.37
24	2-5	SL2.40ch02	21197215	Solyc02g036290.2, Solyc02g036300.2	G/A	0.37
25	2-6	SL2.40ch02	27338193	Solyc02g055540.1	A/G	0.37
26	2-7	SL2.40ch02	34191508	Solyc02g069780.2	A/T	0.38
27	2-8	SL2.40ch02	35258635	Solyc02g071190.2	A/G	0.38
28	2-9	SL2.40ch02	40093278	Solyc02g081640.2	C/T	0.37
29	2-10	SL2.40ch02	46129269	Solyc02g089840.2	C/T	0.37
30	3-1	SL2.40ch03	38505	Solyc03g005030.2	C/T	0.37
31	3-2	SL2.40ch03	6702041	Solyc03g019770.2	A/G	0.37
32	3-3	SL2.40ch03	10216978	Solyc03g034250.2	A/G	0.37
33	3-4	SL2.40ch03	17025485	Solyc03g046540.1	T/C	0.36
34	3-5	SL2.40ch03	23371652	Solyc03g058810.1	A/G	0.37
35	3-6	SL2.40ch03	27047365	Solyc03g059410.1	A/G	0.35
36	3-7	SL2.40ch03	33357866	Solyc03g063760.2	T/C	0.37
37	3-8	SL2.40ch03	36893476	Solyc03g065080.1	T/G	0.36
38	3-9	SL2.40ch03	42084323	Solyc03g077920.1	A/T	0.34
39	3-10	SL2.40ch03	49870313	Solyc03g095360.2	A/C	0.38
40	3-11	SL2.40ch03	50117306	Solyc03g095530.1	A/G	0.37
41	3-12	SL2.40ch03	56203572	Solyc03g111550.2	C/A	0.37
42	3-13	SL2.40ch03	60042376	Solyc03g116680.2	C/G	0.37
43	4-1	SL2.40ch04	794066	Solyc04g007070.2	G/C	0.38
44	4-2	SL2.40ch04	5040990	Solyc04g014800.2	C/G	0.37
45	4-3	SL2.40ch04	13845182	Solyc04g024670.1	A/T	0.37
46	4-4	SL2.40ch04	16721203	Solyc04g024990.1	C/A	0.37
47	4-5	SL2.40ch04	22716678	Solyc04g025870.2	A/C	0.37
48	4-6	SL2.40ch04	26603172	Solyc04g026350.2	G/C	0.37

Supplemental Table 1. Information of MABC 143 markers.(Continued)

No.	SNP ID	Chr	Position	Gene	Allele	PIC value
49	4-7	SL2.40ch04	31000363	Solyc04g040110.2	C/A	0.37
50	4-8	SL2.40ch04	36140519	Solyc04g047750.2	C/T	0.36
51	4-9	SL2.40ch04	40973013	Solyc04g049400.2	C/T	0.37
52	4-10	SL2.40ch04	46600361	Solyc04g050520.2	C/T	0.37
53	4-11	SL2.40ch04	54245561	Solyc04g057970.2	T/C	0.37
54	4-12	SL2.40ch04	56138629	Solyc04g071570.2	T/C	0.37
55	4-13	SL2.40ch04	60042054	Solyc04g077500.2	T/C	0.37
56	5-1	SL2.40ch05	2142284	Solyc05g007610.2	C/T	0.38
57	5-2	SL2.40ch05	9039075	Solyc05g014830.1	A/G	0.38
58	5-3	SL2.40ch05	10091032	Solyc05g015220.2	A/C	0.37
59	5-4	SL2.40ch05	15933251	Solyc05g016350.1	T/G	0.36
60	5-5	SL2.40ch05	20443535	Solyc05g018320.2	A/G	0.37
61	5-6	SL2.40ch05	32946515	Solyc05g025510.2	A/G	0.36
62	5-7	SL2.40ch05	39397925	Solyc05g026020.1	A/T	0.37
63	5-8	SL2.40ch05	40920627	Solyc05g026350.1	A/T	0.37
64	5-9	SL2.40ch05	49143515	Solyc05g041170.1	T/C	0.36
65	5-10	SL2.40ch05	50461117	Solyc05g041360.1	A/C	0.37
66	5-11	SL2.40ch05	55044910	Solyc05g043300.1	A/G	0.37
67	5-12	SL2.40ch05	62993359	Solyc05g053800.2	A/C	0.38
68	6-1	SL2.40ch06	1661847	Solyc06g007650.1	T/C	0.38
69	6-2	SL2.40ch06	6487651	Solyc06g011330.1	T/A	0.37
70	6-3	SL2.40ch06	13917465	Solyc06g019130.1	A/G	0.36
71	6-4	SL2.40ch06	16073739	Solyc06g030550.1	G/C	0.36
72	6-5	SL2.40ch06	22128544	Solyc06g036060.2	A/G	0.38
73	6-6	SL2.40ch06	28410513	Solyc06g048890.2	G/C	0.37
74	6-7	SL2.40ch06	30775982	Solyc06g051140.2	A/G	0.37
75	6-8	SL2.40ch06	35268103	Solyc06g060860.1	C/G	0.37
76	6-9	SL2.40ch06	41603822	Solyc06g073380.2	A/C	0.38
77	6-10	SL2.40ch06	45229159	Solyc06g083570.2	C/T	0.37
78	7-1	SL2.40ch07	396822	Solyc07g005500.2	G/A	0.37
79	7-2	SL2.40ch07	8272212	Solyc07g017890.1	T/C	0.37
80	7-3	SL2.40ch07	12477037	Solyc07g019630.1	C/T	0.32
81	7-4	SL2.40ch07	17962275	Solyc07g021400.1	A/C	0.34
82	7-5	SL2.40ch07	38873153	Solyc07g032470.1	C/G	0.34
83	7-6	SL2.40ch07	42769040	Solyc07g038150.1	A/T	0.34
84	7-7	SL2.40ch07	47065263	Solyc07g040790.2	C/T	0.33
85	7-8	SL2.40ch07	54696990	Solyc07g043510.2	C/T	0.37
86	7-9	SL2.40ch07	55504839	Solyc07g045070.2	C/T	0.37
87	7-10	SL2.40ch07	62466270	Solyc07g062550.2	T/A	0.38
88	7-11	SL2.40ch07	65006928	Solyc07g066300.2	T/G	0.37
89	8-1	SL2.40ch08	4475	Solyc08g005000.2	G/C	0.37
90	8-2	SL2.40ch08	17244322	Solyc08g023560.2	G/T	0.36
91	8-3	SL2.40ch08	23020495	Solyc08g029200.2	C/T	0.35
92	8-4	SL2.40ch08	35423942	Solyc08g045660.1	G/A	0.35
93	8-5	SL2.40ch08	49982665	Solyc08g063130.2	C/G	0.37
94	8-6	SL2.40ch08	54824782	Solyc08g068440.2	A/G	0.38
95	8-7	SL2.40ch08	55144784	Solyc08g068860.2	A/G	0.38
96	8-8	SL2.40ch08	61240553	Solyc08g080910.2	A/T	0.38

Supplemental Table 1. Information of MABC 143 markers.

No.	SNP ID	Chr	Position	Gene	Allele	PIC value
97	9-1	SL2.40ch09	39870	Solyc09g005060.1	T/A	0.37
98	9-2	SL2.40ch09	7049998	Solyc09g014840.1	T/A	0.37
99	9-3	SL2.40ch09	17892974	Solyc09g019990.1	T/G	0.37
100	9-4	SL2.40ch09	20328015	Solyc09g030450.2	G/A	0.37
101	9-5	SL2.40ch09	25053529	Solyc09g031970.2	A/G	0.37
102	9-6	SL2.40ch09	41674855	Solyc09g055910.2	C/T	0.37
103	9-7	SL2.40ch09	45192194	Solyc09g057540.2	T/G	0.37
104	9-8	SL2.40ch09	50156566	Solyc09g059490.1	A/T	0.37
105	9-9	SL2.40ch09	59515303	Solyc09g065690.2	C/T	0.38
106	9-10	SL2.40ch09	62774934	Solyc09g075590.1	G/C	0.38
107	9-11	SL2.40ch09	67166911	Solyc09g097900.2	G/A	0.38
108	10-1	SL2.40ch10	842479	Solyc10g006100.2	A/C	0.38
109	10-2	SL2.40ch10	8411374	Solyc10g018510.1	A/T	0.38
110	10-3	SL2.40ch10	21902314	Solyc10g044450.1	A/G	0.37
111	10-4	SL2.40ch10	26225296	Solyc10g045100.1	T/C	0.36
112	10-5	SL2.40ch10	31551881	Solyc10g045770.1	A/G	0.34
113	10-6	SL2.40ch10	38482645	Solyc10g047910.1	A/G	0.37
114	10-7	SL2.40ch10	42975810	Solyc10g050020.1	C/T	0.38
115	10-8	SL2.40ch10	45229538	Solyc10g050440.1	G/T	0.38
116	10-9	SL2.40ch10	50005218	Solyc10g054040.1	G/A	0.37
117	10-10	SL2.40ch10	57057607	Solyc10g074390.1	G/T	0.37
118	10-11	SL2.40ch10	60273786	Solyc10g079410.1	C/T	0.37
119	11-1	SL2.40ch11	97752	Solyc11g005110.1	T/C	0.37
120	11-2	SL2.40ch11	5702292	Solyc11g012910.1	A/G	0.37
121	11-3	SL2.40ch11	10006256	Solyc11g020020.1	A/G	0.37
122	11-4	SL2.40ch11	18356842	Solyc11g030490.1	C/T	0.36
123	11-5	SL2.40ch11	20046062	Solyc11g030870.1	C/A	0.36
124	11-6	SL2.40ch11	28367572	Solyc11g040040.1	T/C	0.36
125	11-7	SL2.40ch11	32179330	Solyc11g042500.1	G/A	0.35
126	11-8	SL2.40ch11	35999963	Solyc11g044650.1	C/A	0.36
127	11-9	SL2.40ch11	43153452	Solyc11g056470.1	T/G	0.37
128	11-10	SL2.40ch11	46329248	Solyc11g062310.1	G/A	0.37
129	11-11	SL2.40ch11	52219700	Solyc11g071730.1	A/G	0.38
130	12-1	SL2.40ch12	620880	Solyc12g006010.1	G/T	0.38
131	12-2	SL2.40ch12	5998763	Solyc12g015960.1	T/C	0.38
132	12-3	SL2.40ch12	10263688	Solyc12g019430.1	T/C	0.38
133	12-4	SL2.40ch12	15129922	Solyc12g021280.1	T/C	0.37
134	12-5	SL2.40ch12	21053801	Solyc12g027820.1	A/T	0.37
135	12-6	SL2.40ch12	25725692	Solyc12g035500.1	A/G	0.37
136	12-7	SL2.40ch12	30202652	Solyc12g036260.1	G/A	0.37
137	12-8	SL2.40ch12	35206364	Solyc12g038400.1	T/A	0.37
138	12-9	SL2.40ch12	40010179	Solyc12g040610.1	C/T	0.37
139	12-10	SL2.40ch12	45641599	Solyc12g044840.1	A/T	0.37
140	12-11	SL2.40ch12	54012179	Solyc12g062990.1	G/C	0.38
141	12-12	SL2.40ch12	55161472	Solyc12g070200.1	C/G	0.37
142	12-13	SL2.40ch12	61036454	Solyc12g082740.1	T/C	0.37
143	12-14	SL2.40ch12	65136464	Solyc12g099800.1	C/T	0.37

MABC, Marker-assisted Backcrossing; SNP, Single Nucleotide Polymorphism; PIC, Polymorphism Information Content.