



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0048451
(43) 공개일자 2019년05월09일

(51) 국제특허분류(Int. Cl.)

C12Q 1/68 (2018.01)

(52) CPC특허분류

C12Q 1/6895 (2018.05)

C12Q 2600/13 (2013.01)

(21) 출원번호 10-2017-0143417

(22) 출원일자 2017년10월31일

심사청구일자 2017년10월31일

(71) 출원인

주식회사 씨더스 농업회사법인

대전광역시 유성구 테크노로 11-3, N218(관평동, 배재대학교 대덕산학협력관)

(72) 발명자

이정희

대전광역시 동구 대전로542번길 121, 306동 502호(천동, 위드힐아파트)

조성환

대전광역시 동구 대전로542번길 121, 306동 502호(천동, 위드힐아파트)

(뒷면에 계속)

(74) 대리인

최규환

전체 청구항 수 : 총 5 항

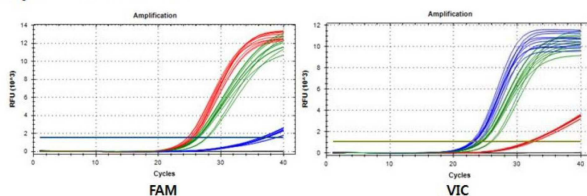
(54) 발명의 명칭 들깨 잎의 보라색 판별용 분자마커 및 이의 용도

(57) 요약

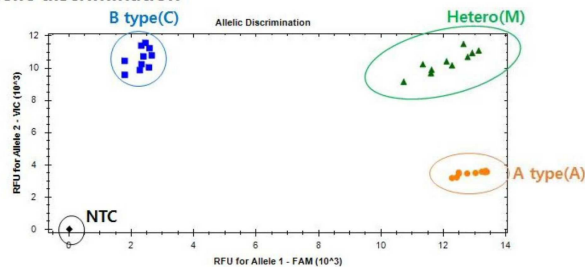
본 발명은 서열번호 1, 서열번호 2 및 서열번호 3의 올리고뉴클레오타이드 프라이머를 포함하는 들깨 잎의 보라색 판별용 프라이머 세트, 상기 프라이머 세트를 이용한 들깨 잎의 보라색 판별용 키트 및 들깨 잎의 보라색 판별 방법에 관한 것으로, 본 발명의 프라이머 세트는 들깨잎의 뒷면의 보라색 여부를 재배 초기 단계에 간단하고 신속, 정확하게 판별할 수 있으므로, 들깨 신품종 육종 등에 유용하게 활용될 수 있을 것이다.

대표도 - 도4

(A) Amplification results



(B) Allelic discrimination



(C)

Genotyping result

Sample id	Leaf color	scaffold8_13707141	
		분석결과	실험결과
31	purple	A	A
6	purple	A	A
49	purple	A	A
15	purple	A	A
16	purple	A	A
23	light purple	M	M
75	light purple	M	M
60	light purple	M	M
22	light purple	M	M
13	light purple	M	M
35	green	C	C
9	green	C	C
76	green	C	C
55	green	C	C
3	green	C	C

(52) CPC특허분류

C12Q 2600/156 (2013.01)

(72) 발명자

남문

대전광역시 서구 도안동로 123, 1708동 1803호(도안동, 도안리슈빌)

임보라

경상북도 김천시 개령면 서부3길 23-2

이보미

충청북도 청주시 흥덕구 가로수로1380번길 71, 104호(복대동)

강윤주

세종특별자치시 금남면 금강변길 1216

이 발명을 지원한 국가연구개발사업

과제고유번호 PJ01040803

부처명 농촌진흥청

연구관리전문기관 포스트게놈 다부처유전체사업단

연구사업명 포스트게놈 다부처유전체사업

연구과제명 들깨 고밀도 유전지도 작성 및 유전체 정보 해독

기 여 율 1/1

주관기관 (주)씨더스

연구기간 2014.04.01 ~ 2017.12.31

명세서

청구범위

청구항 1

서열번호 1, 서열번호 2 및 서열번호 3의 올리고뉴클레오타이드 프라이머를 포함하는 들깨 잎의 보라색 판별용 프라이머 세트.

청구항 2

제1항에 따른 프라이머 세트; 및 증폭 반응을 수행하기 위한 시약을 포함하는 들깨 잎의 보라색 판별용 키트.

청구항 3

제2항에 있어서, 상기 증폭 반응을 수행하기 위한 시약은 RNase H, DNA 중합효소, dNTPs 및 버퍼를 포함하는 것인 키트.

청구항 4

들깨 잎 시료에서 게놈 DNA를 분리하는 단계;

상기 분리된 게놈 DNA를 주형으로 하고, 제1항에 따른 프라이머 세트를 이용하여 증폭 반응을 수행하여 표적서열을 증폭하는 단계; 및

상기 증폭 단계의 산물을 검출하는 단계를 포함하는, 들깨 잎의 보라색을 판별하는 방법.

청구항 5

제4항에 있어서, 상기 증폭 단계의 산물의 검출은 DNA 칩, 겔 전기영동, 방사성 측정, 형광 측정 또는 인광 측정을 통해 수행되는 것인 방법.

발명의 설명

기술 분야

[0001] 본 발명은 들깨 잎의 보라색 판별용 분자마커 및 이의 용도에 관한 것이다.

배경 기술

[0002] NGS(next generation sequencing)를 통한 대량 마커의 개발이 중요해짐에 따라 이를 쉽고 빠르게 경제적으로 고 효율적인 마커개발을 위한 염기서열 결정기술과 고밀도 유전자지도 작성 등의 중요성이 대두되고 있다. 기존의 NGS 방식은 유전체 모든 영역의 염기서열을 세밀하게 해독할 수 있는 반면, 비용 부담이 크고 필요 이상의 정보를 제공하여 분석 효율이 떨어지는 점이 있다. 최근 급격히 발달한 GBS(genotyping by sequencing) 시스템은 2011년 미국 코넬 대학의 연구팀이 옥수수를 소재로 처음 보고되었으며, 이후 보리, 밀, 콩 등을 소재로 국외에서 활발히 연구되고 있다. GBS 기술은 신품종 육성과 순계 선발, 유연관계 분석 등을 위한 다양한 시료를 대상으로 한 효율적인 게놈 수준의 진단 시스템으로 매우 유용하게 사용될 수 있다.

[0003] SNP(Single nucleotide polymorphism)는 개체 간의 DNA에 나타나는 하나의 염기 차이로, 식물체 게놈에 나타나는 빈번한 SNP는 유전자 지도제작(mapping), 분자표지 연관 육종(marker assisted breeding) 및 유전자지도 기반 클로닝(map-based cloning)을 가능하게 한다. SNP는 게놈 유전자 표지 중 가장 많으며, 이러한 SNP를 검출하기 위해서 실험의 용이성 및 비용을 고려한 다양한 SNP 검출 방법과 실험장비가 개발되고 있다.

[0004] 들깨(*Perilla frutescens* L. Britt)는 꿀풀과(Labiatae)의 일년생 초본이며, 원산지는 한반도를 포함한 동부아시아 고원지대이다. 들깨는 들기름을 추출하는 유지작물(oilseed crop)이며, 생육하는 동안에 잎(깻잎)을 수확하여 식용으로도 사용할 수 있다. 들깨 잎은 나물 반찬이나 장아찌, 깻잎 김치 등의 밑반찬으로 먹기도 하고, 무침이나 탕 등에 향신료처럼 사용되기도 한다. 과거에는 주로 종실유를 채취할 목적으로 들깨가 재배되어 왔으나 최근 육류의 소비증가, 외식문화의 발달 및 웰빙 열풍에 의한 쌈채소 소비 시장의 급성장으로 잎용 들깨 품

종이 개발되어 연중 들깨잎의 생산이 가능해졌다. 따라서 들깨잎의 의학적 기능성을 밝히고 이를 식품, 의약 부문에 다양하게 응용할 경우 산업상 이용 가치가 클 것으로 예상된다. 들깨 잎의 기능성을 나타내는 물질로는 페릴라 알데하이드(perilla aldehyde), 로즈마리산, 안토시아닌, 카페인산 등이 보고된 바 있다. 이들 성분은 높은 수용성을 갖기 때문에 섭취하였을 경우 높은 생체이용률을 기대할 수 있다.

[0005] 특히, 들깨 잎의 뒷면이 보라색일 경우 안토시아닌의 함량이 높은 것으로 알려져 있다. 안토시아닌(anthocyanin)은 플라보노이드(flavonoid) 계열의 수용성 색소로 식물의 잎, 꽃 및 과실의 푸른색, 자주색 및 붉은색을 결정하며, 식물의 생장과 발달에도 중요한 역할을 한다. 안토시아닌을 포함한 플라보노이드 화합물은 곤충 및 외부 스트레스에 대한 식물의 손상을 보호하는 기능을 하고(즉 생물학적 및 비생물학적 스트레스 저항성 부여), 특히 활성산소 등에 의한 손상을 막기 위해 항산화제로서의 기능을 한다. 또한, 안토시아닌은 사람에게도 항산화제 역할을 하여 세포사멸을 감소시키고, 심혈관 질병, 암, 당뇨, 신경퇴화, 염증, 바이러스 감염 및 비만을 예방할 수 있어, 건강식품으로 각광을 받고 있다.

[0006] 한편, 한국등록특허 제1036609호에는 '들깨잎과 들깨를 이용한 오메가-3 들깨차 및 이의 제조방법'이 개시되어 있고, 한국등록특허 제0326841호에는 '들깨잎으로부터 고순도 로즈마리닉산의 분리, 정제방법'이 개시되어 있으나, 본 발명의 들깨 잎의 보라색 판별용 분자마커 및 이의 용도에 대해서는 기재된 바가 없다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 상기와 같은 요구에 의해 도출된 것으로서, 본 발명자들은 들깨 잎의 뒷면이 보라색인 레몬들깨(*Perilla citriodora*)와 녹색인 범꼬리소엽(*P. hirtella*)의 중간교잡을 통해 확보된 F2 교배집단의 게놈 서열을 분석하여 SNP 마커를 발굴하였고, 상기 SNP 마커를 증폭할 수 있는 프라이머를 이용하여 F2 교배집단 간의 유전적 다양성을 분석한 결과, 본 발명의 SNP 마커 및 이의 프라이머 세트가 들깨 잎 뒷면의 보라색 여부를 효과적으로 구별할 수 있음을 확인함으로써, 본 발명을 완성하였다.

과제의 해결 수단

[0008] 상기 과제를 해결하기 위해, 본 발명은 서열번호 1, 서열번호 2 및 서열번호 3의 올리고뉴클레오타이드 프라이머를 포함하는 들깨 잎의 보라색 판별용 프라이머 세트를 제공한다.

[0009] 또한, 본 발명은 상기 프라이머 세트; 및 증폭 반응을 수행하기 위한 시약을 포함하는 들깨 잎의 보라색 판별용 키트를 제공한다.

[0010] 또한, 본 발명은 상기 프라이머 세트를 이용하여 들깨 잎의 보라색을 판별하는 방법을 제공한다.

발명의 효과

[0011] 본 발명에 따른 들깨 잎의 보라색 판별용 분자마커는 보라색 형질이 표현되지 않은 어린 개체들에서 잎의 보라색 여부를 신속하고 정확하게 파악할 수 있으므로, 고부가가치의 들깨 신품종 육종 등에 유용하게 이용될 수 있을 것으로 기대된다.

도면의 간단한 설명

[0012] 도 1은 차세대 염기서열 분석법으로 들깨 F2 교배집단의 서열을 분석하여 얻어진 리드 깊이를 보여주는 결과이다.

도 2는 모본인 레몬들깨와 부분인 범꼬리소엽의 F2 교배집단의 잎 뒷면의 색을 분석한 결과이다.

도 3은 들깨 잎의 보라색 형질과 연관된 양적형질유전자좌(LG3)와, 안토시아닌 생합성 관련 WD40, bHLH DNA-결합 단백질 및 Myb 도메인 단백질 코딩 유전자와 SNP 마커 후보의 위치를 표시한 모식도이다. 역삼각형은 SNP를 의미한다.

도 4는 최종 선발된 bHLH DNA-결합 단백질과 WD40 사이의 SNP(PLCM_2)를 이용하여 rhAmp SNP 유전형 분석을 수행한 결과이다. A; 형광물질(FAM, VIC)을 사용하여 보라색 잎 형질(FAM, A type) 및 녹색 잎 형질(VIC, B type)의 검출 결과를 나타내는 그래프, B; 보라색 잎(A type, A), 연보라색 잎(Hetero, M) 및 녹색 잎(B type, C)과 관련된 대립유전자를 구별하여 나타낸 모식도, C; 15개의 F2 교배집단을 대상으로 하여 SNP 종류에 따른

들깨 잎의 색을 구분한 결과를 나타낸 표.

발명을 실시하기 위한 구체적인 내용

- [0013] 본 발명의 목적을 달성하기 위해, 본 발명은 서열번호 1, 서열번호 2 및 서열번호 3의 올리고뉴클레오타이드 프라이머를 포함하는 들깨 잎의 보라색 판별용 프라이머 세트를 제공한다.
- [0014] 본 발명의 상기 프라이머 세트는 rhPCR(RNase H-dependent PCR) 반응을 수행하기 위해 이용될 수 있다. rhPCR은 DNA/RNA 혼성화 이중 가닥의 RNA만 특이적으로 분해하는 리보핵산가수분해효소 H2(ribonuclease H2, RNase H2)를 이용한 PCR 방법으로, rhPCR 반응용 프라이머 세트는 5'에서 3' 방향으로, SNP 위치 염기에 인접한 서열을 표적하는 DNA 염기서열, SNP 위치 염기 및 상기 SNP 위치 염기에 바로 인접한 표적 서열에 상보적인 RNA 염기의 형태로 제작되는 대립 형질 특이적(allele-specific) 프라이머 1 및 2와, 표적 서열에 상보적인 DNA 염기와 RNA 염기서열로 이루어진 좌 특이적(locus-specific) 프라이머로 이루어질 수 있다. 상기 대립 형질 특이적 프라이머 1 및 2는 SNP 위치 염기의 서열만 상이한 프라이머이다.
- [0015] 본 발명에 따른 프라이머 세트에 있어서, 상기 서열번호 1 및 서열번호 2의 올리고뉴클레오타이드 프라이머는 각각 대립 형질 특이적인 프라이머 1 및 2이고, 서열번호 3의 올리고뉴클레오타이드 프라이머는 좌 특이적 프라이머이다. 상기 서열번호 1 및 서열번호 2의 올리고뉴클레오타이드 프라이머는 서열 내 24번째 염기가 SNP 위치 염기이며, 25~29번째 염기는 RNA 염기이다(표 7). 또한, 서열번호 3의 올리고뉴클레오타이드 프라이머는 5' 말단에 서열 내의 RNA 염기를 안정화하기 위해 GC의 염기를 추가로 포함할 수 있으나, 이에 제한되지 않는다.
- [0016] rhPCR 기법은, RNase H2가 증폭단계의 프라이머와 주형 DNA 간의 결합(annealing) 과정에서 RNA 염기를 포함하고 있는 대립 형질 특이적인 프라이머와 주형 DNA가 결합된 RNA/DNA 이중가닥을 인식하여 프라이머 상의 RNA 가닥만을 특이적으로 분해하고, RNase H2의 RNA 분해(절단)를 통해 생성된 3' 말단의 하이드록시기(-OH)를 고도로 인식할 수 있는 변형된 중합효소에 의해서 중합반응이 개시되게 된다.
- [0017] 본 발명의 일 구현 예에 따른 프라이머 세트에 있어서, 상기 서열번호 1 및 서열번호 3의 프라이머 세트는 들깨 잎의 뒷면이 보라색인 품종을 판별할 수 있고, 서열번호 2와 서열번호 3의 프라이머 세트는 들깨잎의 뒷면이 녹색인 품종을 판별할 수 있다.
- [0018] 본 발명의 프라이머는 상기 서열번호 1, 2 및 3의 염기서열의 부가, 결실 또는 치환된 서열도 포함할 수 있다.
- [0019] 본 발명에 있어서, "프라이머"는 카피하려는 핵산 가닥에 상보적인 단일 가닥 올리고뉴클레오타이드 서열을 말하며, 프라이머 연장 산물의 합성을 위한 개시점으로서 작용할 수 있다. 상기 프라이머의 길이 및 서열은 연장 산물의 합성을 시작하도록 허용해야 한다. 프라이머의 구체적인 길이 및 서열은 요구되는 DNA 또는 RNA 표적의 복잡도(complexity)뿐만 아니라 온도 및 이온 강도와 같은 프라이머 이용 조건에 의존할 것이다.
- [0020] 본 발명에 있어서, 프라이머로서 이용된 올리고뉴클레오타이드는 또한 뉴클레오타이드 유사체(analogue), 예를 들면, 포스포로티오에이트(phosphorothioate), 알킬포스포로티오에이트 또는 펩티드 핵산(peptide nucleic acid)을 포함할 수 있거나 또는 삽입 물질(intercalating agent)을 포함할 수 있다.
- [0021] 본 발명은 또한, 본 발명에 따른 프라이머 세트; 및 증폭 반응을 수행하기 위한 시약을 포함하는 들깨 잎의 보라색 판별용 키트를 제공한다.
- [0022] 본 발명의 키트에서, 상기 증폭 반응을 수행하기 위한 시약은 RNase H, DNA 폴리머라제, dNTPs, 버퍼 등을 포함할 수 있다. 또한, 본 발명의 키트는 증폭 산물에 결합할 수 있는 프로브를 추가로 포함할 수 있다.
- [0023] 본 발명에서 용어 "프로브"는 혼성화 프로브로서, 핵산의 상보성 가닥에 서열 특이적으로 결합할 수 있는 더욱 시리보뉴클레오타이드 및 리보뉴클레오타이드를 포함하는 자연적인 또는 변형되는 모노머 또는 결합을 갖는 선형의 올리고머를 의미하며, 5' 또는 3' 말단에 형광물질이 결합된 형태일 수 있으나, 이에 제한되지 않는다.
- [0024] 또한, 본 발명의 키트는 최적의 반응 수행 조건을 기재한 사용자 안내서를 추가로 포함할 수 있다. 안내서는 키트 사용법, 예를 들면, PCR 완충액 제조 방법, 제시되는 반응 조건 등을 설명하는 인쇄물이다. 안내서는 팜플렛 또는 전단지 형태의 안내 책자, 키트에 부착된 라벨 및 키트를 포함하는 패키지의 표면에 설명을 포함한다. 또한, 안내서는 인터넷과 같이 전기 매체를 통해 공개되거나 제공되는 정보를 포함한다.
- [0025] 또한, 본 발명은
- [0026] 들깨 잎 시료에서 게놈 DNA를 분리하는 단계;

- [0027] 상기 분리된 게놈 DNA를 주형으로 하고, 상기 프라이머 세트를 이용하여 증폭 반응을 수행하여 표적서열을 증폭하는 단계; 및
- [0028] 상기 증폭 단계의 산물을 검출하는 단계를 포함하는, 들깨 잎의 보라색을 판별하는 방법을 제공한다.
- [0029] 본 발명의 방법은 들깨 잎 시료로부터 게놈 DNA를 분리하는 단계를 포함한다. 상기 들깨 잎 시료에서 게놈 DNA를 분리하는 방법은 당업계에 공지된 방법을 이용할 수 있으며, 예를 들면, CTAB 방법을 이용할 수도 있고, Wizard prep 키트(Promega 사)를 이용할 수도 있다. 상기 분리된 게놈 DNA를 주형으로 하고, 본 발명의 일 실시예에 따른 프라이머 세트를 이용하여 증폭 반응을 수행하여 표적 서열을 증폭할 수 있다.
- [0030] 본 발명의 방법에 있어서, 상기 증폭된 표적 서열은 검출 가능한 표지 물질로 표지될 수 있다. 일 구현 예에서, 상기 표지 물질은 형광, 인광 또는 방사성을 발하는 물질일 수 있으나, 이에 제한되지 않는다. 바람직하게는, 상기 표지 물질은 6-FAM (6-Carboxyfluorescein), NED, VIC, PET 또는 ROX이다. 표적 서열의 증폭시 프라이머의 5'-말단에 6-FAM, NED, VIC, PET 또는 ROX를 표지하여 PCR을 수행하면 표적 서열이 검출가능한 형광 표지 물질로 표지될 수 있다. 또는, 증폭 서열에 혼성화될 수 있는 프로브의 5'-말단에 상기 표지 물질이 결합된 경우, PCR을 수행하면 증폭된 산물에 상기 프로브가 혼성화되어, 증폭 산물의 양에 비례하게 형광 물질이 표지될 수 있다. 또한 상기 표지 물질에는 Cy-5 또는 Cy-3가 포함될 수 있다. 또한, 방사성 물질을 이용한 표지는 PCR 수행시 ^{32}P 또는 ^{35}S 등과 같은 방사성 동위원소를 PCR 반응액에 첨가하면 증폭 산물이 합성되면서 방사성이 증폭 산물에 혼입되어 증폭 산물이 방사성으로 표지될 수 있다.
- [0031] 본 발명의 일 구현 예에 있어서, 상기 들깨 잎의 보라색을 판별하는 방법은 상기 증폭 산물을 검출하는 단계를 포함하며, 상기 증폭 산물의 검출은 DNA 칩, 겔 전기영동, 모세관 전기영동, 방사성 측정, 형광 측정 또는 인광 측정을 통해 수행될 수 있으나, 이에 제한되지 않는다.
- [0032] 이하, 본 발명을 실시예에 의해 상세히 설명한다. 단, 하기 실시예는 본 발명을 예시하는 것일 뿐, 본 발명의 내용이 하기 실시예에 한정되는 것은 아니다.
- [0033] **실시예 1. 시료 준비 및 게놈 DNA 추출**
- [0034] 들깨(*Perilla frutescens* L. Britt.) 유전체 정보 해독을 위해 잎 뒷면이 보라색인 레몬들깨(*P. citriodora*)를 모본으로, 잎 뒷면이 녹색인 범꼬리소엽(*P. hirtella*)을 부본으로 하여 F1 세대를 얻었고, 이들의 중간교잡(interspecific hybridization)을 통해 96개체의 F2 교배집단을 확보하였다. 상기 F2 교배집단의 잎을 이용하여 CTAB(Cetyl Trimethyl Ammonium Bromide) 방법으로 게놈 DNA를 추출하였다.
- [0035]
- [0036] **실시예 2. 레몬들깨 및 범꼬리소엽의 F2 집단 96계통에 *MspI*와 *PstI* 제한효소 처리 및 GBS(genotyping by sequencing) 라이브러리 제작**
- [0037] 상기에서 추출한 DNA에 제한효소 *MspI* 및 *PstI*를 처리하여 절단(digestion)시킨 후, 절단된 DNA 단편에 바코드(barcode) 및 어댑터(adaptor) 시퀀스를 라이게이션(ligation)하여 96개체의 복수 샘플을 동시에 분석하여 GBS 라이브러리를 제작하였고, 상기 제작된 GBS 라이브러리를 이용하여 Illumina Hiseq 2000(Illumina, 덴마크)로 차세대 염기서열(Next Generation Sequencing, NGS) 분석을 실시한 결과, 약 51 Gbp 길이의 유전체 염기서열을 확보하였다(표 1).

표 1

- [0038] *MspI* & *PstI* 제한효소를 이용한 GBS 결과

파일 명칭	로우 리드(raw reads) 수	평균 길이	총 길이(bp)
TN1604L0104_1.fastq	257,280,053	101	25,985,285,353
TN1604L0104_2.fastq	257,280,053	101	25,985,285,353
합계	514,560,106		51,970,570,706

- [0039] **실시예 3. 역다중화(demultiplexing) 및 대량 변이 정보(genome-wide SNP) 분석**

- [0040] **3-1. GBS 데이터의 역다중화 및 전처리**

- [0041] 96개체의 F2 교배집단을 바코드 서열에 따라 역다중화(demultiplexing)한 후 바코드 및 어댑터 시퀀스를 제거하

였고, 프레드 스코어(phred score)를 20으로 적용한 SolexaQA package(version 1.13)를 이용하여 서열 품질을 분석하였다. 염기서열을 이루는 염기 하나의 품질 값이 20보다 낮은 것은 제거하였고, 나머지 염기서열 중 길이가 25 bp 이상인 염기서열만 선별하였다.

[0042]

그 결과, 459,068,228개의 제거된 짧은 리드(trimmed short reads)를 확보하였으며, 이는 5,241,830개 로우 리드(raw reads)의 91.22%에 해당하는 개수인 것으로 확인되었다(표 2, 3).

표 2

들깨 F2 자손개체의 96개통 GBS 데이터의 역다중화(Demultiplexing) 및 서열 품질 분석 결과

BarCode	Sample name	Sum of raw reads	Total length of raw reads	Sum of trimmed reads	Total length of trimmed reads (bp)	Avg. length of trimmed reads (bp)	Trimmed/ Raw (%)
CTCC	1	1,066,966	107,763,566	963,852	80,781,253	83.31	90.34%
TGCA	2	3,118,844	315,003,244	2,830,766	233,929,379	82.64	90.76%
ACTA	3	3,951,776	399,129,376	3,579,444	295,396,712	82.53	90.58%
CAGA	4	1,453,102	146,763,302	1,324,064	110,773,079	83.66	91.12%
AACT	5	3,815,038	385,318,838	3,469,434	284,589,722	82.03	90.94%
GCGT	6	4,227,636	426,991,236	3,848,296	318,873,742	82.86	91.03%
CGAT	7	1,503,880	151,891,880	1,361,276	114,148,641	83.85	90.52%
GTAA	8	6,675,586	674,234,186	6,091,394	502,945,049	82.57	91.25%
AGGC	9	5,692,442	574,936,642	5,231,272	433,914,215	82.95	91.90%
GATC	10	5,118,516	516,970,116	4,688,852	387,388,600	82.62	91.61%
TCAC	11	4,770,332	481,803,532	4,327,282	357,251,653	82.56	90.71%
TGCGA	12	5,560,296	561,589,896	5,086,182	418,453,025	82.27	91.47%
CGCTT	13	1,850,784	186,929,184	1,670,900	139,458,651	83.46	90.28%
TCAAC	14	5,069,448	512,014,248	4,598,274	378,167,037	82.24	90.71%
CTAGC	15	1,926,198	194,545,998	1,750,580	146,058,441	83.43	90.88%
ACAAA	16	4,435,354	447,970,754	4,046,738	332,451,740	82.15	91.24%
TTCTC	17	3,508,430	354,351,430	3,167,602	258,438,152	81.59	90.29%
AGCCC	18	7,163,822	723,546,022	6,546,754	537,339,957	82.08	91.39%
GTATT	19	4,179,320	422,111,320	3,820,774	312,516,872	81.79	91.42%
CTGTA	20	3,871,958	391,067,758	3,518,288	293,051,497	83.29	90.87%
ACCGT	21	6,239,054	630,144,454	5,674,312	467,198,822	82.34	90.95%
CGTTA	22	7,813,252	789,138,452	7,106,578	582,564,321	81.98	90.96%
GGTGT	23	9,174,072	926,581,272	8,404,172	691,926,262	82.33	91.61%
AGGAT	24	2,477,706	250,248,306	2,279,286	186,919,022	82.01	91.99%
ATTGA	25	9,340,770	943,417,770	8,587,396	702,797,439	81.84	91.93%
CATCT	26	1,177,068	118,883,868	1,068,348	88,781,842	83.10	90.76%
CCTAC	27	3,193,878	322,581,678	2,880,770	239,629,179	83.18	90.20%
GAGGA	28	5,031,510	508,182,510	4,649,474	381,348,283	82.02	92.41%
GGAAC	29	7,756,074	783,363,474	7,106,304	581,743,322	81.86	91.62%
GTCAA	30	7,912,250	799,137,250	7,262,450	596,025,924	82.07	91.79%
TAATA	31	5,027,626	507,790,226	4,582,746	374,378,664	81.69	91.15%
TACAT	32	6,049,394	610,988,794	5,490,310	449,698,630	81.91	90.76%
TCGTT	33	7,996,372	807,633,572	7,243,094	597,441,343	82.48	90.58%
GGTTGT	34	5,530,124	558,542,524	5,058,226	411,700,654	81.39	91.47%
CCAGCT	35	1,609,666	162,576,266	1,461,260	120,824,967	82.69	90.78%
TTCAGA	36	4,488,670	453,355,670	4,075,874	330,396,389	81.06	90.80%
TAGGAA	37	5,173,362	522,509,562	4,757,550	387,367,717	81.42	91.96%
GCTCTA	38	4,734,086	478,142,686	4,310,992	352,045,578	81.66	91.06%
CCACAA	39	3,956,786	399,635,386	3,587,976	297,254,541	82.85	90.68%
CTTCCA	40	2,751,850	277,936,850	2,480,128	204,985,094	82.65	90.13%
GAGATA	41	2,680,656	270,746,256	2,467,288	201,796,581	81.79	92.04%
ATGOCT	42	7,129,762	720,105,962	6,509,370	530,880,492	81.56	91.30%
AGTUGA	43	6,705,854	677,291,254	6,167,594	502,126,210	81.41	91.97%
ACCTAA	44	5,753,016	581,054,616	5,232,630	426,888,393	81.58	90.95%
ATATGT	45	4,951,566	500,108,166	4,550,576	368,964,027	81.08	91.90%
ATCGTA	46	5,180,672	523,247,872	4,730,298	386,336,718	81.67	91.31%
CATCGT	47	2,004,508	202,455,308	1,825,022	151,139,705	82.82	91.05%
CGCGGT	48	653,320	65,985,320	589,654	49,055,054	83.19	90.26%
CTATTA	49	3,237,466	326,984,066	2,944,914	243,301,972	82.62	90.96%
GCCAGT	50	9,710,874	980,798,274	8,859,334	725,219,663	81.86	91.23%
GGAAGA	51	2,764,908	279,255,708	2,543,240	206,534,497	81.21	91.98%

[0043]

표 3

들개 F2 자손개체의 96계통 GBS 데이터의 역다중화(Demultiplexing) 및 서열 품질 분석 결과 (continued)

BarCode	Sample name	Sum of raw reads	Total length of raw reads	Sum of trimmed reads	Total length of trimmed reads (bp)	Avg. length of trimmed reads (bp)	Trimmed/ Raw (%)
GTAAGT	52	8,741,552	882,896,752	7,986,688	651,218,781	81.54	91.36%
GTTGAA	53	12,808,616	1,293,670,216	11,739,370	955,440,656	81.39	91.65%
TAACGA	54	11,185,422	1,129,727,622	10,199,128	832,459,682	81.62	91.18%
TGGCTA	55	7,181,702	725,351,902	6,548,374	537,067,799	82.02	91.18%
TATTTTT	56	8,830,768	891,907,568	7,837,882	621,632,798	79.31	88.76%
CTTGCTT	57	3,210,462	324,256,662	2,897,902	238,457,769	82.29	90.26%
ATGAAAC	58	8,119,716	820,091,316	7,437,446	601,800,401	80.91	91.60%
AAAAGTT	59	6,555,700	662,125,700	6,037,982	488,475,078	80.90	92.10%
GAATTCA	60	11,432,504	1,154,682,904	10,403,642	841,866,267	80.92	91.00%
GAAGTTC	61	5,659,774	571,637,174	5,188,460	420,534,335	81.05	91.67%
GGACCTA	62	7,395,404	746,935,804	6,793,920	552,575,120	81.33	91.87%
GTGATT	63	2,652,438	267,896,238	2,447,996	199,715,393	81.58	92.29%
AAGGCTT	64	5,586,940	564,280,940	5,089,368	413,854,653	81.32	91.09%
AATATGC	65	7,801,726	787,974,326	7,119,304	576,878,094	81.03	91.25%
ACGTGTT	66	8,928,894	901,818,294	8,176,484	667,714,043	81.66	91.57%
ATTAATT	67	8,290,764	837,367,164	7,549,310	605,271,772	80.18	91.06%
ATTGGAT	68	6,845,258	691,371,058	6,283,944	505,084,139	80.38	91.80%
GATAAGT	69	4,558,044	460,362,444	4,158,112	341,574,029	82.15	91.23%
CGCTGAT	70	1,643,846	166,028,446	1,489,018	122,879,572	82.52	90.58%
CGGTAGA	71	517,638	52,281,438	468,370	38,634,209	82.49	90.48%
CTACGGA	72	3,573,394	360,912,794	3,242,228	267,473,981	82.50	90.73%
GCGGAAT	73	2,687,462	271,433,662	2,451,834	201,157,500	82.04	91.23%
TAGCGGA	74	5,547,072	560,254,272	5,071,104	412,732,327	81.39	91.42%
TGGAAGA	75	6,501,248	656,626,048	5,910,312	481,318,935	81.44	90.91%
TCTGTGA	76	3,971,528	401,124,328	3,601,040	292,608,913	81.26	90.67%
TGCTGGA	77	7,756,830	783,439,830	7,049,460	572,612,657	81.23	90.88%
ACGACTAC	78	10,137,312	1,023,868,512	9,237,112	750,139,895	81.21	91.12%
TAGCATGC	79	9,695,354	979,230,754	8,849,038	715,643,861	80.87	91.27%
TAGGCCAT	80	6,647,260	671,373,260	6,074,468	492,338,454	81.05	91.38%
TGCAAGGA	81	7,653,400	772,993,400	6,961,702	563,624,490	80.96	90.96%
TGGTACGT	82	3,653,292	368,982,492	3,331,740	269,847,277	80.99	91.20%
TCTCAGTC	83	9,124,846	921,609,446	8,289,842	671,099,484	80.95	90.85%
CCGGATAT	84	3,553,470	358,900,470	3,236,760	265,919,407	82.16	91.09%
CGCCTTAT	85	1,473,110	148,784,110	1,333,490	109,725,534	82.28	90.52%
AACCGAGA	86	2,915,956	294,511,556	2,660,106	215,156,685	80.88	91.23%
ACAGGGAA	87	5,935,414	599,476,814	5,427,106	440,400,104	81.15	91.44%
ACGTGGTA	88	3,174,562	320,630,762	2,910,272	237,466,430	81.60	91.67%
CCATGGGT	89	4,061,708	410,232,508	3,707,012	305,272,614	82.35	91.27%
CGCGGAGA	90	1,509,754	152,485,154	1,362,694	112,083,799	82.25	90.26%
CGTGTGGT	91	1,456,472	147,103,672	1,318,856	108,575,321	82.33	90.55%
GCTGTGGA	92	8,835,068	892,341,868	8,059,364	654,626,794	81.23	91.22%
GGATTGGT	93	6,000,426	606,043,026	5,549,552	450,516,841	81.18	92.49%
GTGAGGGT	94	3,887,898	392,677,698	3,581,732	290,248,740	81.04	92.13%
TATCGGGA	95	3,729,374	376,666,774	3,406,328	274,736,514	80.65	91.34%
TTCCTGGA	96	4,620,430	466,663,430	4,183,186	338,340,572	80.88	90.54%
Average	-	5,241,830	529,424,859	4,781,961	390,416,671	82	91.15%
Sum	96	503,215,708	50,824,786,508	459,068,228	37,480,000,441	-	-

[0044]

[0045]

3-2. 표준 유전체에 대한 전처리 염기서열의 정렬(alignment)

[0046]

상기 선발된 염기서열들에서 SNP를 추출하기 위해, 정렬 프로그램인 BWA(Burrows-Wheeler Aligner, version 0.6.1)를 이용하여 들개 유전체 초안(표준유전체, 김태호 박사님(농업과학원)으로부터 제공받음)에 정렬(alignment)하였다. 프로그램 옵션으로는 미스매치(mismatch) 최대 2개, 씨드 랭스(seed length) 27 설정 외에 기본값(default)을 적용하였다.

[0047]

그 결과, 제거된 짧은 리드(trimmed short reads) 459,068,228개 중에서 4,104,955개의 리드(85.84%)가 매핑(mapping) 되었으며, 리드 매핑 부위를 확인한 결과, 리드 깊이(read depth)가 평균 67.17X인 것을 확인하였다(도 1). 계통별 1,000,000 리드 이상을 확보하는 것을 목표로 했을 때, 대부분의 계통에서 목표치의 리드를 확보할 수 있었으며, *ApkI* 제한효소를 이용한 GBS 라이브러리에 비해 *MspI* & *PstI* 제한효소를 이용한 GBS 데이터에서 보다 높은 리드 깊이를 확보할 수 있어, SNP의 정확도가 높을 것으로 예측되었다.

[0048]

3-3. SNP 탐색 및 필터링을 통한 유전자 지도 작성

[0049] 공개된 공통 염기서열 추출 프로그램인 SAMtools(version 0.1.16)를 이용하여 상기 정렬된 염기서열로부터 단일 염기 다형성(single nucleotide polymorphism, SNP) 마커 후보를 탐색하였다. 프로그램 옵션으로는 대부분 기본값을 적용하였고, 정확도 높은 SNP를 분석하기 위해 정렬 품질(alignment quality) 값을 기본 값인 25보다 높은 30으로 적용하고, 적어도 3개 이상의 로우 리드가 정렬된 위치에서 추출된 SNP만을 선별하였다.

[0050] 교배양친 간 존재하는 SNP를 확보하기 위해, 표준 유전체와 잎 뒷면이 녹색인 범꼬리소엽 간의 SNP 및 표준 유전체와 잎 뒷면이 보라색인 레몬들깨 간의 SNP를 각각 발굴하여 부분과 모본 사이 즉, 교배 양친 간 차이를 보이는 SNP를 확인하였다. 그 후, 상기 범꼬리소엽과 레몬들깨 간의 SNP에서 최종 마커를 선별한 뒤 ICIM(inclusive composite interval mapping) 프로그램을 이용하여 유전자 지도를 작성하였다.

[0051] *Msp1* 및 *Pst1* 제한효소를 이용하여 제작된 GBS 라이브러리는 총 7단계의 SNP 필터링(filtering)를 실시하였고(표 4), 이를 통해 확보된 총 2,518개 SNP를 이용하여 10개의 연관군이 형성된 것을 확인하였으며, 총 유전거리는 1,309.39 cM, 마커 간의 유전거리는 평균 0.56 cM인 것으로 분석되었다(표 5).

표 4

[0052] *Msp1* 및 *Pst1* 제한효소를 이용한 GBS 라이브러리 SNP 필터링 통계치

번호	분석항목	SNP 수(SNP loci)
1	Total SNP loci	91,132
2	Missing data	38,668
3	MAF (minor allele frequency)	14,223
4	Polymorphic SNP loci between parents	9,607
5	SNP remark (genotype code 변환)	9,445
6	Identical genotype filtering	7,669
7	joinmap 이용 가능한 수준의 마커선발	2,518

표 5

[0053] 들깨의 F2 교배집단을 이용하여 작성된 유전자 지도의 통계치

연관군	마커 수	길이(cM)	마커 간의 평균 길이(cM)
LG 1	661	271.86	0.41
LG 2	182	154.68	0.85
LG 3	254	154.54	0.61
LG 4	233	100.84	0.43
LG 5	279	164.90	0.59
LG 6	195	57.26	0.30
LG 7	247	158.66	0.64
LG 8	204	97.36	0.48
LG 9	203	103.49	0.51
LG 10	60	45.8	0.78
합계	2518	1309.39	0.56

[0054] 실시예 4. QTL 매핑 및 마커 후보 선별

[0055] 잎 뒷면이 보라색인 레몬들깨와 녹색인 범꼬리소엽의 F2 교배집단 96개체의 잎 뒷면의 색을 육안으로 분석한 결과, 보라색(purple)인 개체는 20개, 녹색(green)인 개체는 14개, 연한 보라색(light purple)인 개체는 62개인으로 확인되었다(도 2).

[0056] 또한, CIM(Composite Interval mapping) 방법을 사용하여 Windows QTL Cartographer(version 2.5)(WinQTLCart) 프로그램으로 잎 뒷면이 보라색을 나타내는 형질에 관한, 양적 형질 유전자좌(qualitative trait loci, QTL) 분석을 수행하였으며, 95% 유의수준으로 500회 반복실험을 실시하여 유의성 있는 LOD(logarithm of the odds) 값을 확인하였다.

[0057] 그 결과, 들깨잎 뒷면의 보라색과 관련된 QTL은 12개의 마커로 구성된 3번 연관군에서만 확인되었고, QTL에 대한 LOD값이 14.3 내지 18.29의 범위를 나타내었으며 표현형 변이에 대한 QTL의 설명 비율이 41~48%로, 보라색

잎의 형질이 우성임을 알 수 있었다(표 6).

표 6

보라색의 들깨 잎 형질의 양적 형질 유전자좌(QTL)

Leaf Color	LG ^a	마커	Position (cM) ^b	LOD ^c	PVE(%) ^d	Add ^e	Dom ^f	d/a ^g
LC	3	scaffold8_17172553	94.38	14.48	41	-0.51	0.05	A
		scaffold8_17004842	94.38	14.94	42	-0.53	0.05	A
		scaffold8_17030380	95.59	15.53	43	-0.55	0.06	A
		scaffold8_16101189	96.12	16.45	45	-0.57	0.08	A
		scaffold8_14696338	97.19	15.72	44	-0.55	0.05	A
		scaffold8_14806662	97.87	15.23	43	-0.57	0.12	PD
		scaffold8_14219959	98.52	16.41	45	-0.59	0.15	PD
		scaffold8_12737907	102.39	17.39	47	-0.64	0.18	PD
		scaffold8_12548782	103.67	15.79	44	-0.61	0.18	PD
		scaffold8_12511056	104.3	16.43	45	-0.60	0.14	PD
		scaffold8_12611865	104.94	17.26	46	-0.63	0.13	PD
		scaffold8_12737895	105.54	18.29	48	-0.64	0.12	A

a; 연관군(Linkage group)

b; 마커로부터 유전적 거리

c; QTL 분석에 따른 LOD값 (LOD 역치값 = 3.0)

d; 테스트 위치에서 H0:H3 하에서 QTL에 의해 변이가 설명되는 비율 (마커가 표현형(Leaf color)에 미치는 영향의 정도)

e; H1 하에서 상가적 효과 (Leaf Color 형질에 대한 마커의 상가적 효과)

f; H2 하에서 우성효과 (Leaf Color 형질에 대한 마커의 우성 효과)

g; 유전자 기능의 추정 (A: 상가적 효과 0-0.2, PD: 부분적 우성 0.21-0.80, D: 우성 0.81-1.2, OD: 초우성 >1.20)

또한, 잎 뒷면이 보라색인 들깨 유전자 내 SNP 마커를 선별하기 위해 기존에 보고된 레몬들깨 전사체 어셈블리 서열을 분석하여, 보라색 잎 QTL에서 360개의 들깨 유전자들을 확인하였다. 또한, 모델식물인 애기장대에서 보라색과 관련된 안토시아닌 성분의 생합성 조절인자로 알려진 WD40(WD-repeat), bHLH(basic helix-loop-helix) DNA-결합 단백질 및 Myb 도메인 단백질(Myb domain protein)과 상기 들깨 유전자들을 비교하여 서열유사성이 있는, 들깨 전사체 서열 3개를 확인하였다(도 3).

상기 3개 유전자 내 SNP를 확인한 결과, WD40에서는 부모 양친의 다형성을 보이는 SNP가 확인되지 않았고, Myb에서는 부모 양친의 다형성을 보이는 2개의 SNP가 확인되었으나, F2 교배집단에 적용하였을 때 잎 뒷면의 보라색 형질과 관련성을 확인할 수 없었다. 반면, bHLH에서는 부모 양친의 다형성을 보이는 2개의 SNP가 확인되었을 뿐만 아니라 F2 교배집단에 적용하였을 때에도 잎 뒷면의 보라색 형질과 관련된 1개의 SNP가 확인되었으며, 이를 PLCM_1(G/A)로 명명하였다. 또한, 상기 3개 유전자 주변 영역에서 SNP를 분석한 결과, bHLH와 WD40 사이에서 잎 뒷면의 보라색 형질과 연관성이 있는 2개의 SNP를 확인하였으며, 이를 각각 PLCM_2(A/C)와 PLCM_3(C/T)으로 명명하였다.

실시예 5. rhAmp SNP 유전형 분석(genotyping)을 통한 SNP 검증

상기 선별된 들깨 잎 뒷면의 보라색 형질의 후보 SNP 3개에 대해 보라색 판별능을 확인하기 위하여 rhAmp SNP 유전형 분석을 실시하였고, 이를 위해 3개의 SNP 종류에 따라 5개씩 선정하여 총 15개의 F2 교배집단을 사용하였다.

SNP 종류에 따라 대립 형질 특이적 프라이머를 각각 디자인하였고, 상기 프라이머의 3' 말단에는 RNA 염기

(base)가 연결되어 있어, RNase H2가 반응할 수 있도록 디자인하였다. 각 대립형질 특이적 프라이머는 SNP 종류에 따라 다른 상용 시퀀스(Universal sequence)를 포함하여 각 상용 시퀀스에 따라서 FAM 또는 VIC이 결합된 Probe가 반응하도록 하였다. 또한, 좌 특이적 프라이머(PLCM Specific primer)는 서열 내의 RNA 염기를 안정화하기 위해 서열의 끝에 GC를 추가하여 제작하였다(표 7). 표 7의 각 대립 형질 특이적 프라이머에서 PLCM specific primer는 정방향 프라이머이고, primer 1 및 primer 2는 역방향 프라이머이다. PLCM_3의 경우 시퀀스의 방향성이 정방향인 것과 역방향인 것 2종류로 디자인하여 실험을 진행하였다.

[0072] 상기 대립 형질 특이적 프라이머와 좌 특이적 프라이머를 사용한 PCR 분석은 rhAmp™ SNP Assays(Integrated DNA Technologies, 미국)를 이용하여 제조사의 방법에 따라 실시하였다. 구체적으로는, 게노믹 DNA(10ng/ μ l) 1 μ l, rhAmp Genotyping Master Mix(2x) 2.5 μ l, rhAmp Reporter Mix(40x) 0.15 μ l와 rhAmp SNP Assay(20x) 0.25 μ l의 조성에 증류수 1.1 μ l를 넣어, 총 부피가 5 μ l가 되도록하여 수행하였다. 증폭반응은 CFX96 Touch™ Real-Time PCR Detection System(Bio-Rad, California, USA) 장비를 사용하였다. 증폭 조건은 초기 변성 단계 95℃, 10분; 변성 단계 95℃, 10초 → 결합 단계 60℃, 30초 → 신장 단계 68℃, 20초로 하여 40회 수행하였으며; 68℃에서 20초 동안 FAM 및 VIC 형광을 감지하여 분석하였다. 실험은 각 2회 반복하였으며 최종적으로 확인된 형광의 종류에 따라 SNP 종류를 구분하였다.

[0073] 그 결과, PLCM_1과 PLCM_3에서는 PCR 증폭 반응이 일어나지 않았으나, PLCM_2에서는 서열번호 1의 대립 형질 특이적 프라이머에서 앞 뒷면이 보라색인 5개의 F2 교배집단이 확인되었으며, 이는 모본인 레몬들깨와 같은 A 유형임을 알 수 있었다. 또한, PLCM_2의 서열번호 1의 대립 형질 특이적 프라이머는 앞 뒷면이 녹색인 5개의 F2 교배집단을 확인하였고 이를 통해 부분인 범꼬리소엽과 같은 B 유형임을 알 수 있었으며, 연한 보라색을 보이는 5개의 F2 교배집단은 헤테로(hetero) 유형임을 확인하였다(도 4).

표 7

SNP 검출에 이용된 PLCM 프라이머 서열 정보

SNP			rhAmp SNP genotyping assay		
번호	위치	대립유전자	프라이머 명칭	프라이머 서열	검증
1	Scffold8_13762404	G/A	PLCM_1 primer 1	GGTGTTTAAGGCAGCAGAGrUrGrUrC (서열번호 4)	X
			PLCM_1 primer 2	GGTGTTTAAGGCAGCAGAArUrGrUrC (서열번호 5)	
			PLCM_1 Specific primer	GCCCATGCTCGACTTTTATTCCTTATrCrUrCrGrC (서열번호 6)	
2	Scffold8_13707141	A/C	PLCM_2 primer 1	ACATCAATTCACATGAAAGTGTTrGrArArGrG (서열번호 1)	O
			PLCM_2 primer 2	ACATCAATTCACATGAAAGTGTGrGrArArGrG (서열번호 2)	
			PLCM_2 Specific primer	GCGACCTCATCTCGAGTATGTTTrGrUrGrUrU (서열번호 3)	
3	Scffold8_13706957	C/T	PLCM_3F primer 1	CCAACTGTTTTTCAGCTTTGATTGrUrArCrArC (서열번호 7)	X
			PLCM_3F primer 2	CCAACTGTTTTTCAGCTTTGATTArUrArCrArC (서열번호 8)	
			PLCM_3F Specific primer_1	GCTAGCAGTGGACTATGGGTTTTTrGrUrGrUrU (서열번호 9)	
	Scffold8_13706957	C/T	PLCM_3R primer 1	CCCATCTTTAGCAGCGTGTACrArArUrCrA (서열번호 10)	X
			PLCM_3R primer 2	CCCATCTTTAGCAGCGTGTATrArArUrCrA (서열번호 11)	
			PLCM_3R Specific primer_2	GCCATACTGCAGAGATGAGGTACrUrGrArArA (서열번호 12)	

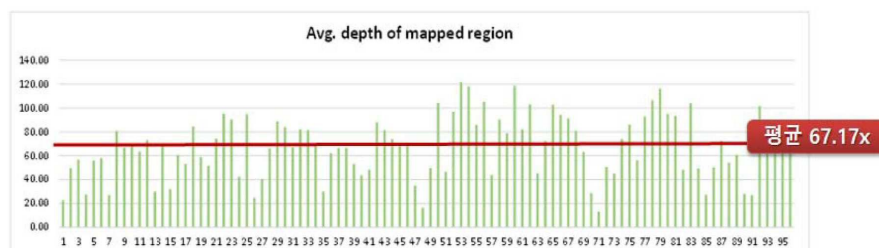
(상기 rA, rU, rG, rC는 RNA 염기)

[0074]

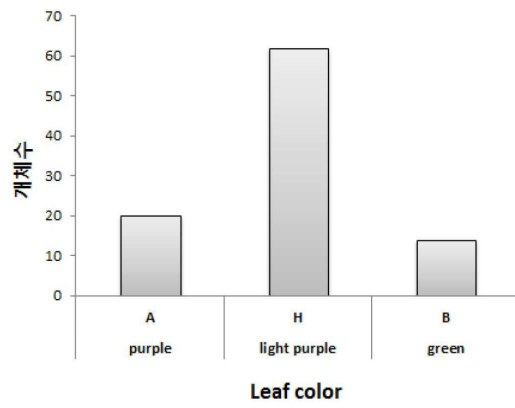
[0075] 상기 결과를 통해, 앞의 색이 보라색인 들깨에서는 'A'의 염기유전형이 확인되었고, 앞의 색이 녹색인 들깨에서는 'C'의 염기유전형이 확인됨으로써, 본 발명의 마커 및 상기 마커를 증폭할 수 있는 프라이머를 통해 들깨 앞의 색을 구분할 수 있음을 확인하였다.

도면

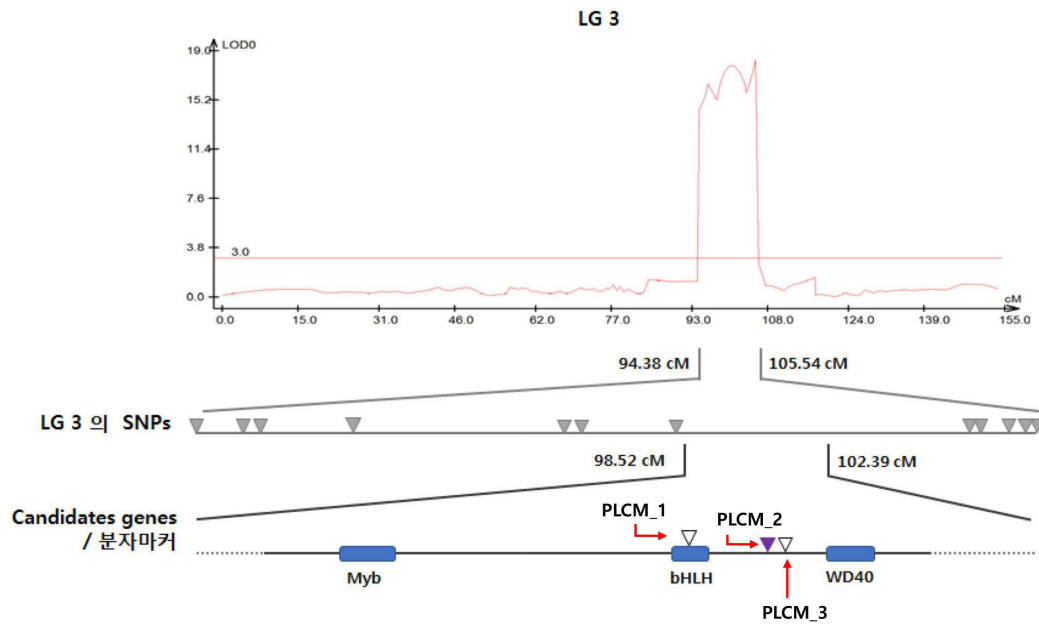
도면1



도면2

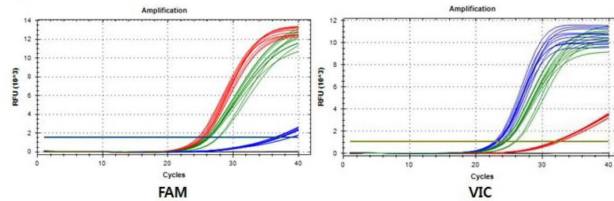


도면3

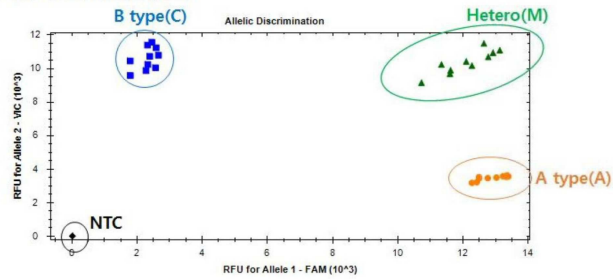


도면4

(A) Amplification results



(B) Allelic discrimination



(C)

Genotyping result

Sample id	Leaf color	scaffold8_13707141	
		분석결과	실험결과
31	purple	A	A
6	purple	A	A
49	purple	A	A
15	purple	A	A
16	purple	A	A
23	light purple	M	M
75	light purple	M	M
60	light purple	M	M
22	light purple	M	M
13	light purple	M	M
35	green	C	C
9	green	C	C
76	green	C	C
55	green	C	C
3	green	C	C

서열 목록

<110> SEEDERS

<120> Molecular marker for discriminating purple color of Perilla leaf
and uses thereof

<130> PN17420

<160> 12

<170> Kopatent In 2.0

<210> 1

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 1

acatcaattc cacatgaaag tgttgaagg

<210> 2

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 2

29

acatcaattc cacatgaaag tgtggaagg

29

<210> 3

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 3

gcgacctcat ctctgcagta tgtttguguu

30

<210> 4

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 4

ggtgtttaag gcagcagagu gguc

24

<210> 5

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 5

ggtgtttaag gcagcagaau gguc

24

<210> 6

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 6

gccccatgctc gacttttatt cttatcucg c

31

<210> 7

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> primer
 <400> 7
 ccaactgttt ttcagctttg attguacac 29
 <210> 8
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 8
 ccaactgttt ttcagctttg attauacac 29
 <210> 9
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 9
 gctagcagtg gactatgggt ttgugau 28
 <210> 10
 <211> 26
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 10
 cccatcttta gcagcgtgta caauca 26
 <210> 11
 <211> 26
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer
 <400> 11
 cccatcttta gcagcgtgta taauc 26
 <210> 12
 <211> 29

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 12

gccatactgc agagatgagg tcacugaaa

29