(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년11월19일
(11) 등록번호 10-1919431
(24) 등록일자 2018년11월12일

(51) 국제특허분류(Int. Cl.)
C12Q 1/68 (2018.01) *C12N 15/10* (2017.01)
C40B 50/06 (2017.01)
(52) CPC특허분류
C12Q 1/6855 (2018.05)
C12N 15/1093 (2013.01)
(21) 출원번호 10-2017-0081625
(22) 출원일자 2017년06월28일
심사청구일자 2017년06월28일
(56) 선행기술조사문헌
PLoS ONE, 6(5):e19379 (2011)
Korean Journal of Agricultural Science, 44(4):495-503 (2017.12.)
PLoS ONE, 10(8):e0134880 (2015)
PLoS ONE, 12(1):e0170580 (2017.01.26.)

(73) 특허권자
충남대학교 산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
(72) 발명자
오상근
대전광역시 서구 관저로 83, 203동 1303호 (관저동, 신선마을아파트)
장현아
대전광역시 서구 월평북로 11, 207동 402호
(74) 대리인
최규환

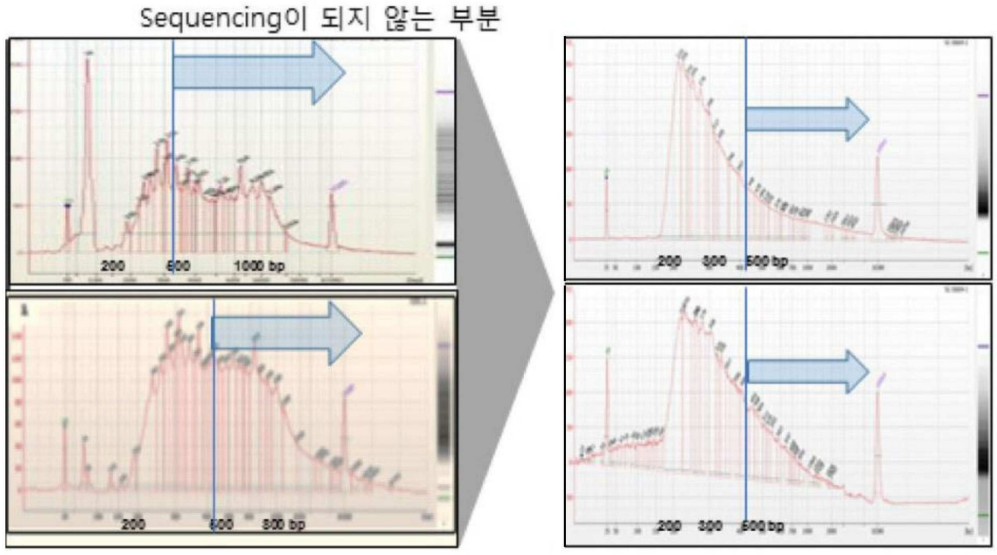
전체 청구항 수 : 총 2 항 심사관 : 김승범

(54) 발명의 명칭 **균일한 크기의 GBS 라이브러리 구축 방법**

(57) 요 약

본 발명은 균일한 크기의 GBS 라이브러리 구축 방법에 관한 것으로, 게노믹 DNA 절단을 위한 제한효소의 처리시간, 및 라이브러리 제작을 위한 증폭반응의 조건 등을 조절하여 GBS 분석을 위한 최적의 크기로 GBS 라이브러리를 구축하는 방법에 관한 것으로, 본 발명의 방법을 이용하면 GBS 분석의 효율을 증대시킬 수 있을 것이다.

대 표 도 - 도4



(52) CPC특허분류
C40B 50/06 (2013.01)
C12Q 2525/191 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 PJ011187022017
부처명 농촌진흥청
연구관리전문기관 농생물게놈활용사업단
연구사업명 차세대바이오그린21 사업
연구과제명 포도 핵심집단 기반 GWAS 및 유용유전자 선발
기 여 율 1/1
주관기관 국립원예특작과학원
연구기간 2015.01.15 ~ 2017.12.31

명 세 서

청구범위

청구항 1

(1) 포도로부터 게노믹 DNA를 추출하는 단계;

(2) 상기 (1) 단계의 추출한 게노믹 DNA에 *ApeK*I 제한효소를 처리하고 70~80℃에서 4~6시간 동안 반응시켜 게노믹 DNA를 절단하는 단계;

(3) 서열번호 1 및 서열번호 2의 염기서열로 이루어진 바코드 어댑터 세트와 서열번호 3 및 서열번호 4의 염기서열로 이루어진 공통 어댑터 세트를 각각 듀플렉스(duplex) 형태로 준비하는 단계;

(4) 상기 (2) 단계의 절단된 DNA와 상기 (3) 단계에서 준비된 각각의 바코드 및 공통 어댑터 듀플렉스를 혼합하고 라이게이즈(ligase)를 상기 바코드 및 공통 어댑터 듀플렉스 각각 2μM에 대해 300~400U 처리하여 바코드 어댑터-절단된 DNA-공통 어댑터 형태로 결합시키는 단계; 및

(5) 상기 (4) 단계의 결합된 바코드 어댑터-절단된 DNA-공통 어댑터를 주형으로 변성(denaturation), 어닐링(annealing), 및 연장(extension)의 사이클을 15~17회 반복하고, 상기 연장 시간은 18~22초로 하여 증폭반응을 수행하는 단계;를 포함하는 것을 특징으로 하는, 200 내지 400bp 크기의 포도 GBS(genotyping by sequencing) 라이브러리 구축 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

제1항의 방법으로 구축된 포도 GBS 라이브러리.

발명의 설명

기 술 분 야

[0001] 본 발명은 균일한 크기의 GBS(genotyping by sequencing) 라이브러리 구축 방법에 관한 것이다.

배 경 기 술

[0002] NGS(next generation sequencing)를 통한 대량 마커의 개발이 중요해짐에 따라 이를 쉽고 빠르며 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. 국내에서도 벼, 고추, 들깨, 수박 등 다양한 작물에서 GBS 분석이 시도되고 있으며, GBS 기술은 신품종 육성과 순계 선발, 유연관계 분석 등을 위한 다양한 시료를 대상으로 한 효율적인 게놈 수준의 진단 시스템으로 매우 유용하게 사용될 수 있다. 이에 따라 많은 연구실 및 기업에서 GBS 기술을 적용하고자 하였으나, 안정적인 GBS 라이브러리 제작 기술 확보에 어려움이 있다.

[0003] 포도(*Vitis vinifera* L.)는 3대 과수작물 중 하나로서 레스베라트롤, 플라보노이드, 폴리페놀계 화합물이 풍부하여 건강식품의 소재로 주목받고 있음에도 불구하고, 현재 국내에서 약 70%의 재배면적을 차지하는 캠벨얼리 품종은 외국 도입품종이다. 이에 따라 국내 재배종 포도를 비롯해 내재해성(disaster tolerance)을 갖는 새로운 포도 육성이 필요한 실정이다.

[0004] 본 발명자들은 포도 자원을 이용하여 기존에 보고된 GBS 제작 방법을 참고하되, 자체적으로 GBS 라이브러리를 효율적으로 개선 및 제작하는 기술을 개발하였다.

[0005] 한편, 한국등록특허 제1632881호에는 유전체 DNA 말단 염기서열에 특이적인 프라이머를 사용하여 유전체 DNA의 말단 염기서열만을 대량 증폭시킨 후 NGS를 이용하여 유전체 DNA의 말단 염기서열을 분석하는 방법에 관한 '차세대 염기서열분석법을 이용한 유전체 DNA 말단 서열의 대량 분석방법'이 개시되어 있고, 미국공개특허 제2015-0344947호에는 'Genotyping by next-generation sequencing'이 개시되어 있으나, 본 발명의 균일한 크기의 GBS 라이브러리 구축 방법에 대해서는 기재된 바가 없다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 상기와 같은 요구에 의해 도출된 것으로서, 본 발명자들은 국내 포도속(*Vitis* spp.) 유전자원에 대한 GBS 라이브러리 구축을 위해, 국내 포도속 유전자원 약 350여점을 이용하여, 게노믹 DNA 절단을 위한 제한효소(*ApeK*Ⅰ)의 처리시간, 및 라이브러리 제작을 위한 증폭반응의 조건 등을 조절하여, 기존의 방법 대비 개선된 방법에서 균일한 크기의 고품질 GBS 라이브러리가 구축되는 것을 확인함으로써, 본 발명을 완성하였다.

과제의 해결 수단

[0007] 상기 과제를 해결하기 위해, 본 발명은 (1) 생물체로부터 게노믹 DNA를 추출하는 단계; (2) 상기 (1) 단계의 추출한 게노믹 DNA에 제한효소를 처리하고 70~80℃에서 4~6시간 동안 반응시켜 게노믹 DNA를 절단하는 단계; (3) 서열번호 1 및 서열번호 2의 염기서열로 이루어진 바코드 어댑터 세트와 서열번호 3 및 서열번호 4의 염기서열로 이루어진 공통 어댑터 세트를 각각 듀플렉스(duplex) 형태로 준비하는 단계; (4) 상기 (2) 단계의 절단된 DNA와 상기 (3) 단계에서 준비된 각각의 바코드 및 공통 어댑터 듀플렉스를 혼합하고 라이게이즈(ligase)를 처리하여 바코드 어댑터-절단된 DNA-공통 어댑터 형태로 결합시키는 단계; 및 (5) 상기 (4) 단계의 결합된 바코드 어댑터-절단된 DNA-공통 어댑터를 주형으로 증폭반응을 수행하는 단계;를 포함하는 것을 특징으로 하는, 균일한 크기의 GBS(genotyping by sequencing) 라이브러리 구축 방법을 제공한다.

[0008] 또한, 본 발명은 상기 방법으로 구축된 GBS 라이브러리를 제공한다.

발명의 효과

[0009] 본 발명의 방법은 GBS 라이브러리 제작 기술의 최적화를 통해 기존 방식을 보완하여, 최적의 균일한 DNA 단편 크기로 고품질의 GBS 라이브러리를 구축할 수 있으므로, 본 발명의 방법은 GBS를 통한 genome-wide SNP 분석과 분석 시료의 유전적 정보를 결합하여 형질 연관 마커 개발 및 신품종 육성을 위한 분자 마커 개발에 유용하게 활용될 수 있을 것이다.

도면의 간단한 설명

[0010] 도 1은 포도 유전자원에 대한 GBS 분석 방법의 모식도이다.

도 2는 제한효소(*ApeK*Ⅰ) 처리 시간에 따른 포도 GBS 라이브러리의 겔 전기영동 결과이다.

도 3은 PCR 증폭조건에 따른 포도 GBS 라이브러리의 겔 전기영동 결과이다.

도 4는 기존방법(좌)과 개선방법(우)으로 구축된 포도 GBS 라이브러리의 QC(quality control) 분석 결과이다.

발명을 실시하기 위한 구체적인 내용

[0011] 본 발명의 목적을 달성하기 위하여, 본 발명은

[0012] (1) 생물체로부터 게노믹 DNA를 추출하는 단계;

[0013] (2) 상기 (1) 단계의 추출한 게노믹 DNA에 제한효소를 처리하고 70~80℃에서 4~6시간 동안 반응시켜 게노믹 DNA를 절단하는 단계;

[0014] (3) 서열번호 1 및 서열번호 2의 염기서열로 이루어진 바코드 어댑터 세트와 서열번호 3 및 서열번호 4의 염기서열로 이루어진 공통 어댑터 세트를 각각 듀플렉스(duplex) 형태로 준비하는 단계;

[0015] (4) 상기 (2) 단계의 절단된 DNA와 상기 (3) 단계에서 준비된 각각의 바코드 및 공통 어댑터 듀플렉스를 혼합하고 라이게이즈(ligase)를 처리하여 바코드 어댑터-절단된 DNA-공통 어댑터 형태로 결합시키는 단계; 및

[0016] (5) 상기 (4) 단계의 결합된 바코드 어댑터-절단된 DNA-공통 어댑터를 주형으로 증폭반응을 수행하는 단계;를 포함하는 것을 특징으로 하는, 균일한 크기의 GBS(genotyping by sequencing) 라이브러리 구축 방법을 제공한다.

[0017] 본 명세서에서 용어 '라이브러리(library)'는 유전형 분석 대상 생물체의 게노믹 DNA 시료의 단편들을 증폭한 산물의 통합(pooling)물을 의미하는 것으로, 본 발명에서는 제한효소를 처리하여 절단된 DNA 단편과 상기 제한효소의 인식 자리를 포함하고 있는 바코드 어댑터 세트 및 공통 어댑터 세트가 라이게이션된 형태를 가지며 PCR을 통해 증폭된 산물의 통합물을 의미한다. 따라서, 본 발명의 '균일한 크기의 GBS 라이브러리'의 의미는, 균일한 크기를 갖는 다양한 '바코드 어댑터-절단된 DNA-공통 어댑터 형태의 폴리뉴클레오티드'가 통합되어 구축된 것을 의미한다.

[0018] GBS 기술은 유전체 전체의 염기서열 분석(whole genome sequencing) 또는 재염기서열 분석(resequencing) 방식과 같은 NGS 기반 기술이지만, 바코드를 이용하여 대량의 시료를 한 번에 분석 가능한 기술로, 제한효소로 게노믹 DNA를 절단하여 절단된 주변 서열만 부분적으로 해독하며, 분석 비용 및 시간 절약이 가능한 장점이 있는 유전형 분석 기술이다. 다른 유전형 분석 기술과는 달리, GBS는 저렴한 비용으로 높은 수준의 SNP 마커들을 참조 유전체에 맵핑할 수 있는 장점이 있다. 또한, GBS는 RRL(reduced representation library), RAD-seq(restriction site associated DNA sequencing) 등과 같이 제한 효소를 이용하는 기본원리는 동일하지만 제한효소로 자른 후 사이즈 크기를 상관하지 않는 점에서 라이브러리 제작이 더 간단한 장점이 있다.

[0019] 본 발명의 일 구현 예에 따른 GBS 라이브러리 구축 방법에 있어서, 상기 제한효소는 *ApeK*I 일 수 있으나, 이에 제한되지 않는다. GBS 라이브러리 구축을 위한 방법에 있어서 제한효소의 선택시 고려사항은, 4 내지 6bp의 크기를 자르며, 메틸화에 따른 활성 차이를 보이고, 제한 효소 처리 후 DNA 단편의 길이가 200~700bp 정도 되는 비율이 높으며, 시퀀싱 되는 영역에서 약 10x 커버리지 될 수 있도록 하는 제한효소를 선택할 수 있다.

[0020] 본 발명의 일 구현 예에 따른 GBS 라이브러리 구축 방법에 있어서, 서열번호 1 및 서열번호 2의 염기서열로 이루어진 바코드 어댑터 세트는, 시퀀싱 프라이머와 상보적인 서열(어댑터 서열)의 뒷부분에 개체 식별을 위한 4~6bp의 바코드 서열이 연결된 정방향 프라이머(서열번호 1)와, 제한효소 인식자리, 바코드 서열 및 시퀀싱 프라이머와 상보적인 서열이 순차적으로 연결된 역방향 프라이머(서열번호 2)로 이루어진 프라이머 세트일 수 있으나, 이에 제한되지 않는다.

[0021] 본 발명의 GBS 라이브러리 구축 방법에 있어서, 상기 생물체는 과수작물 또는 미생물일 수 있으나, 이에 제한되지 않는다.

[0022] 본 발명의 일 구현 예에 따른 GBS 라이브러리 구축 방법에 있어서, 상기 생물체는 과수작물일 수 있고, 바람직하게는 포도, 사과, 배 또는 귤 등일 수 있으며, 가장 바람직하게는 포도일 수 있으나, 이에 제한되지 않는다.

[0023] 본 발명의 GBS 라이브러리 구축 방법에 있어서, 상기 GBS 라이브러리는 균일한 크기를 가지는 다양한 '바코드 어댑터-절단된 DNA-공통 어댑터 형태의 폴리뉴클레오티드'의 라이브러리로, 상기 폴리뉴클레오티드가 200 내지 400bp 크기인 라이브러리일 수 있으나, 이에 제한되지 않는다.

[0024] 본 발명의 일 구현 예에 따른 GBS 라이브러리 구축 방법에 있어서, 상기 (4) 단계의 라이게이즈 처리는 바코드

및 공통 어댑터 듀플렉스 각각 2μM에 대해 300~400U 처리하는 것일 수 있으며, 바람직하게는 360U 처리하는 것일 수 있으나, 이에 제한되지 않는다.

[0025] 본 발명의 일 구현 예에 따른 GBS 라이브러리 구축 방법에 있어서, 상기 (5) 단계의 증폭반응은 변성(denaturation), 어닐링(annealing), 및 연장(extension)의 사이클을 15~17회 반복하고, 상기 연장 시간은 18~22초로 수행하는 것일 수 있고, 바람직하게는 변성, 어닐링, 및 연장의 사이클을 16회 반복하고, 상기 연장 시간은 20초로 수행하는 것일 수 있으나, 이에 제한되지 않는다. 상기 반복 횟수 및 연장(extension) 단계의 시간은 비효율적인 큰 단편의 증폭을 억제하며, 비특이적인 구역의 증폭을 방지하기 위해 최적화된 것이다.

[0026] 본 발명은 또한, 상기 방법으로 구축된 GBS 라이브러리를 제공한다. 본 발명의 GBS 라이브러리는 제한효소의 처리 시간 및 증폭반응의 조건이 개선된 방법에 따라 구축된 것으로, GBS 분석을 위해 최적화된 균일한 크기를 갖는 라이브러리이다.

[0027] 이하, 본 발명을 실시예에 의해 상세히 설명한다. 단, 하기 실시예는 본 발명을 예시하는 것일 뿐, 본 발명의 내용이 하기 실시예에 한정되는 것은 아니다.

[0028] **재료 및 방법**

[0029] **1. 고순도 게노믹 DNA의 분리**

[0030] 포도의 잎은 섬유질을 다량 함유하고 있어서 컬럼을 이용하여 DNA를 분리해야 효율적이다. 국립원예특작과학원으로부터 제공받은 350개의 포도 유전자원의 잎을 액체질소에 얼린 다음 막자를 이용하여 완전히 갈아준 후, DNeasy Plant mini kit(Qiagen, 독일)를 사용하여 제조사의 매뉴얼에 따라 게노믹 DNA를 추출하였다. 또한 RNA가 남아있지 않도록 RNase를 65℃에서 20분 동안 처리해주었다.

[0031] **2. 바코드 & 공통 어댑터**

[0032] 대량의 샘플을 한꺼번에 분석하기 위해 96개의 서로 다른 바코드 어댑터 쌍(PLoS One. 2011. 6(5)e19379)을 사용하였으며, 바코드 어댑터는 F: 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNN-3'(서열번호 1)과 R: 5'-CWGnnnnAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-3'(서열번호 2)의 두 가닥으로 구성하였다. 바코드 어댑터는 상보적인 33bp 염기와 서로 다른 염기(NNNN, nnnn)로 구성되어 있으며, CWG는 *ApeK*I 인식자리로써 공통 어댑터 F와 상보적이다. 공통 어댑터는 F: 5'-CWGAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAG-3'(서열번호 3)와 R: 5'-CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT-3'(서열번호 4)로 구성하였다. 1X 용출 버퍼(10mM Tris-Cl, pH 8.0-8.5)를 이용하여 단일가닥의 어댑터를 50μM로 희석하고, 희석한 각 어댑터 F와 R은 10X 어댑터 버퍼(500mM NaCl, 100mM Tris-Cl)를 넣어 10μM 농도로 맞춘 후, 95℃ 2분 → ramp down to 25℃ by 0.1℃/초(s) → 25℃ 30분의 반응을 통해 어닐링시켰다. 그 후, 멸균수를 넣어 최종 2μM의 농도로 희석하여 준비하였다.

[0033] **3. 제한효소의 처리**

[0034] 모든 샘플의 균일한 라이브러리 분포를 위해 DNA 샘플 당 100ng을 정확히 정량하여 사용하였다. 제한효소(*ApeK*I) 3.6U와 10X 버퍼를 넣어 총 20㎕에 맞춘 후 75℃에서 5시간동안 중합효소연쇄반응(polymerase chain reaction, PCR)을 수행하였다. 처리시간이 2시간일 경우 잘리지 않은 DNA가 상당수 남아있는 것을 확인, 이를 해결하기 위해 제한효소 처리 시간을 조정하였다.

[0035] **4. 어댑터 라이게이션**

[0036] 제한효소를 처리하여 잘린, 절단된 DNA 샘플 20㎕와 어닐링시킨 바코드 어댑터, 및 공통 어댑터를 라이게이션하여 GBS 라이브러리를 증폭을 위한 주형으로 준비하였다. *ApeK*I으로 절단된 DNA의 양 끝은 바코드 어댑터의 R-CWG(*ApeK*I 인식자리)와 공통 어댑터의 F-CWG가 각각 붙게 된다. *ApeK*I으로 절단된 DNA 20㎕와 바코드 어댑터, 공통 어댑터 각각을 2μM씩 넣고 T4 ligase(Takara, 일본) 360U와 혼합하여 총 30㎕에 맞춘 후 22℃에서 2시간

반응시킨 후, 65℃에서 30분 동안 비활성화시켜 반응을 종료시켰다.

[0037] **5. 각 샘플의 통합(pooling) 및 정제**

[0038] 하나의 튜브에 어댑터 라이게이션이 끝난 샘플(바코드 어댑터-절단된 DNA-공통 어댑터 형태)을 각 5㎕씩 옮기고, Qiaquick PCR purification kit(Qiagen)를 사용하여 정제하였다. 따뜻한 멸균수 50㎕로 용출하여 라이브러리 증폭을 위한 주형가닥으로 사용하였다.

[0039] **6. 라이브러리 제작**

[0040] PCR 증폭에 사용된 프라이머는 F: 5'-AATGATACGG CGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'(서열번호 5)와 R: 5'-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT-3'(서열번호 6)이다. 상기 프라이머 F(서열번호 5)의 3' 말단은 바코드 어댑터와 상보적으로 결합할 33개 염기서열을 포함하고, 프라이머 R(서열번호 6)의 3'말단은 공통 어댑터와 상보적으로 결합할 33개 염기서열을 포함한다. 500bp가 넘는 샘플의 경우 NGS 시퀀싱 품질이 매우 떨어지므로, 절단된 DNA 단편의 길이가 200-450bp 범위에서 메인 피크가 유지되도록 하였다. 이와 같이 500bp가 넘는 큰 단편의 증폭을 억제하기 위해 기존의 GBS 라이브러리 증폭 조건에서 연장(extension) 시간을 조절하였으며, 비특이적인 구역의 증폭을 방지하기 위해 반복 횟수를 최대 16회로 제한하였다. 라이브러리 제작을 위한 증폭반응의 조건은 다음과 같다: 72℃ 5분 → 98℃ 30초 → [98℃ 30초 → 65℃ 30초 → 72℃ 20초](16회) → 72℃ 5분. 완료된 라이브러리는 Qiaquick PCR purification kit(Qiagen)으로 한 번 더 정제하였으며, 정제 후 샘플은 겔 전기영동과 QC(quality control) 분석을 통해 GBS 분석을 위한 적합 여부를 판단하였다.

[0041] **실시예 1. 제한효소의 처리 조건에 따른 결과 비교**

[0042] 제한효소(*ApeK*Ⅰ)는 G/CWGC(W=A 또는 T) 염기서열을 인식하여 절단하며, 제한효소 처리는 게놈 전체가 아닌 일부 염기서열만을 폭넓게 분석하기 위한 GBS 라이브러리 제작과정 중 가장 중요한 단계라고 할 수 있다. 350개의 포도 시료 중 96개 시료의 게노믹 DNA를 100ng씩 정량하여 2개의 96 웰 플레이트에 각각 넣고 *ApeK*Ⅰ(New england biolabs, 미국)을 75℃에서 2시간(기존 방법) 또는 5시간(개선 방법)동안 처리하여 최종 구축된 GBS 라이브러리의 결과를 비교하였다.

[0043] 그 결과, 기존의 방법인 2시간 처리 조건에서는 DNA가 균일하게 잘리지 않아 PCR 산물이 부분적으로 모여 특정 크기의 밴드처럼 확인되는 반면, 제한효소를 5시간 처리한 조건에서는 균일한 크기의 PCR 산물이 증폭되는 것으로 확인되었다(도 2). 결과적으로 포도 시료에서는 제한효소(*ApeK*Ⅰ)의 처리시간이 2시간으로는 불충분하며, 5시간으로 늘렸을 때 균일한 크기의 GBS 라이브러리가 만들어지는 것을 알 수 있었다. 또한, 절단된 DNA와 어댑터 라이게이션 단계에서 사용된 T4 ligase의 사용량도 기존 방법인 640U보다 적은 360U만으로도 충분히 라이게이션됨을 알 수 있었다. 본 결과는 반복과정을 거쳐 도출된 결과이다.

[0044] **실시예 2. PCR 증폭 최적 조건에 따른 결과 비교**

[0045] 각기 다른 96개의 바코드와 라이게이션시킨 샘플을 한 튜브에 모아 정제하여 PCR 증폭을 위한 최종 주형가닥으로 사용하였다. 증폭반응은 Accupower PCR premix(Bioneer, 한국)를 이용하여 통합된(pooled) DNA 5㎕, 서열번호 5 및 서열번호 6의 각 프라이머(10pmol) 1㎕, 증류수 43㎕를 넣어 최종 50㎕로 맞춘 후 PCR을 수행하였다. 기존의 방법(72℃ 5분 → 98℃ 30초 → [98℃ 30초 → 65℃ 30초 → 72℃ 30초](18-20회) → 72℃ 5분 → 4℃ 유지)에 따라 PCR을 수행한 결과, 최종 구축된 GBS 라이브러리에서 1kb 이상의 크기까지 증폭된 산물이 있는 것으로 확인되었으며, 정제 후에도 약 400-600bp의 크기에 PCR 산물이 밀집되어 나타났다(도 3A). 이에 따라 분석에 비효율적인 큰 단편의 증폭을 억제하기 위해 기존의 증폭 조건에서 연장 시간을 10초 줄이고, 비특이적인 구역의 증폭을 방지하기 위해 반복 횟수를 16회로 제한하여 PCR을 재수행하였다: 72℃ 5분 → 98℃ 30초 → [98℃ 30초 → 65℃ 30초 → 72℃ **20초**](**16회**) → 72℃ 5분 → 4℃ 유지. 그 결과, 500bp 이하 크기의 PCR 산물이 확인되었고, 정제를 통해 다이머(dimer)를 제거함과 동시에 200-400bp의 크기에 밀집된 라이브러리를 확인할 수 있었다(도 3B).

[0046] **실시예 3. QC 분석**

[0047] QC(quality control) 분석을 수행하여 포도로부터 제작한 라이브러리가 GBS 분석에 적합한지 분석하였다. 도 4와 같이 기존의 방법을 이용하여 제작한 라이브러리는 200 내지 1,000bp 범위 내에 분포하며, 특히 400~600bp 사이에 밀집되어 나타났다(도 4 좌측). 반면, 본 발명의 개선된 방법을 이용하여 동일한 포도 DNA로부터 제작한 라이브러리의 경우, 200~500bp 범위 내에 분포하였으며 특히 200~400bp에 밀집되어 나타남을 확인하였다(도 4 우측). 상기의 결과를 통해, 본 발명의 개선된 방법을 이용하여 제작한 라이브러리가 GBS 분석에 더 적합하다고 판단되었다.

[0048] **실시예 4. GBS 분석 통계치에 따른 결과 비교**

[0049] 기존방법과 개선된 방법에 따라 제작된 각 라이브러리는 Illumina HiSeq 2000 플랫폼을 이용하여 시퀀싱하였고, *Vitis vinifera*(data source: Phytozome v10.1)를 참조 서열로 하여 GBS 분석을 수행하였다.

[0050] 그 결과, 하기 표 1에서 보여지는 것처럼, GBS 분석 통계치에 따르면 기존 방법에 따라 제작된 1, 2차 라이브러리는 질(quality)이 떨어져 2 lane을 읽은 것으로써, 이는 1, 2차 라이브러리의 효율이 2배가 되었을 때 비로소 1 lane만 읽은 3차 라이브러리의 초기 원시 리드(raw reads) 수와 비슷해진 것으로 해석할 수 있었다. 또한, 원시 리드를 양질의 리드로 정제하는 트리밍(trimming) 과정을 통해 1, 2차 라이브러리에서 모두 95% 이상의 비율로 리드가 걸러져 3, 4차 라이브러리에 비해 높았으나, 최종적으로 게놈 내에 매핑되어 분석에 사용되는 리드의 수 비율은 1, 2차 평균 71.7%, 3, 4차 평균 81.87%를 나타냄으로써 개선된 방법을 통해 제작한 GBS 라이브러리가 매핑 효율을 10.17% 더 증가시킬 수 있음을 알 수 있었다. 또한, 4차 라이브러리는 96개 샘플을 한 번에 통합(pooling) 했음에도 불구하고 질이 좋아 1 lane에 분석이 가능했고, 3차(48개 샘플 통합) 라이브러리를 분석한 것보다 리드 수는 적었으나 매핑 효율은 유사하게 나타나, 개선된 방법으로 제작된 GBS 라이브러리의 경우, 샘플의 수가 많은 경우에도 그 효율이 유지되는 것으로 예측되었다.

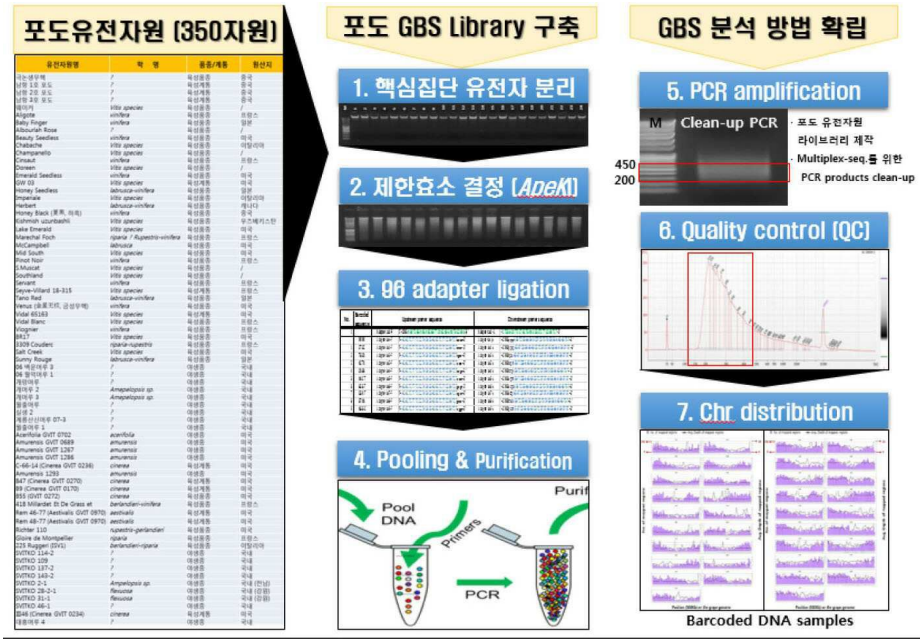
표 1

[0051] GBS 분석 결과

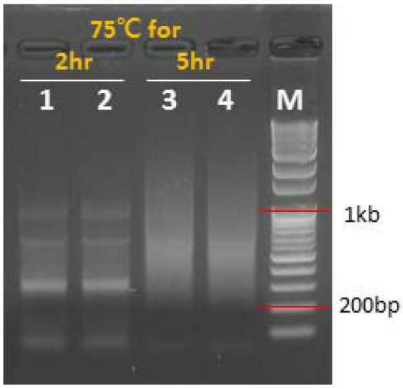
		Sum of raw reads(bp)	Sum of trimmed reads(bp)	Avg. length of trimmed reads(bp)	Trimmed/ Raw(%)	No. of mapped reads(bp)	Percent of mapped reads(%)
기존 방법 (샘플 수)	1차 (66)	6,427,547	6,144,051	85.83	95.58	4,014,664	65.11
	2차 (48)	7,806,595	7,473,294	89.15	95.56	5,936,001	78.29
개선 방법 (샘플 수)	3차 (48)	**6,350,266**	5,855,557	80.37	92.21	4,788,251	**81.51**
	4차 (96)	**3,285,859**	2,904,955	79.82	88.64	2,380,631	**82.23**

도면

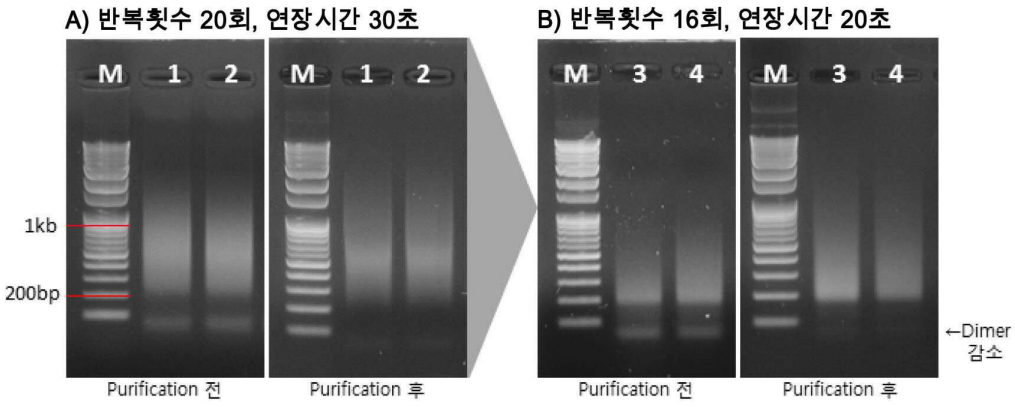
도면1



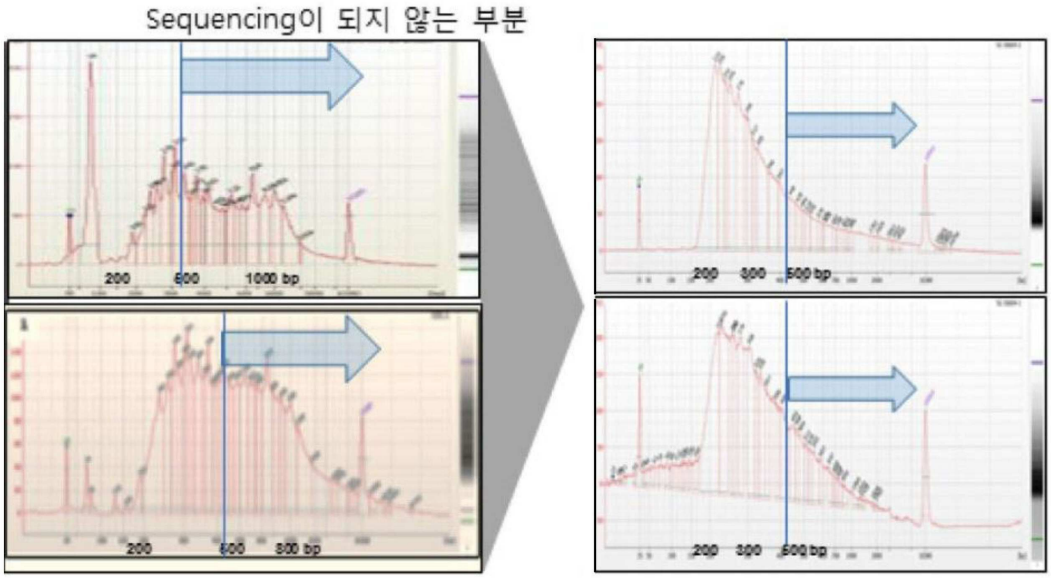
도면2



도면3



도면4



서 열 목 록

<110> The Industry & Academic Cooperation in Chungnam National University (IAC)

<120> Method for constructing GBS library having uniform size

<130> PN17199

<160> 6

<170> KopatentIn 2.0

<210> 1

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> barcode adapter F

<400> 1

acactctttc cctacacgac gctcttccga tctnnnn 37

<210> 2

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> barcode adapter R

<400> 2

cwgnnnnaga tcggaagagc gtcgtgtagg gaaagagtgt 40

<210> 3

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> common adapter F

<400> 3

cwgagatcgg aagagcggtt cagcaggaat gccgag 36

<210> 4

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> common adapter R

<400> 4

ctcggcattc ctgctgaacc gctcttccga tct 33

<210> 5

<211> 58

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 5

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatct 58

<210> 6

<211> 61

<212> DNA

<213> Artificial Sequence

<220><223> primer

<400> 6

caagcagaag acggcatacg agatcggtct cggcattcct gctgaaccgc tcttccgatc 60

t 61