(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2018 년 11 월 15 일 (15.11.2018)

WIPO | PCT

(10) 국제공개번호
WO 2018/208133 A1

(51) 국제특허분류:
C12Q 1/6806 (2018.01) *C12N 15/10* (2006.01)

(21) 국제출원번호: PCT/KR2018/005493

(22) 국제출원일: 2018 년 5 월 14 일 (14.05.2018)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:
10-2017-0059558 2017 년 5 월 12 일 (12.05.2017) KR

(71) 출원인: **서울대학교산학협력단 (SEOUL NATIONAL UNIVERSITY R&DB FOUNDATION)** [KR/KR]; 08826 서울시 관악구 관악로 1, Seoul (KR).

(72) 발명자: **권성훈 (KWON, Sung Hoon)**; 06291 서울시 강남구 삼성로 151, 3동 502호, Seoul (KR). **김성식 (KIM, Sung Sik)**; 08826 서울시 관악구 관악로1 서울대학교 902동 408호, Seoul (KR). **염희란 (YEOM, Hui Ran)**; 24045 강원도 철원군 갈말읍 명성로139번길 13-22, 101동 406호, Gangwon-do (KR). **이용희 (LEE, Yong Hee)**; 06603 서울시 서초구 서초중앙로24길 57, 102동 1401호, Seoul (KR).

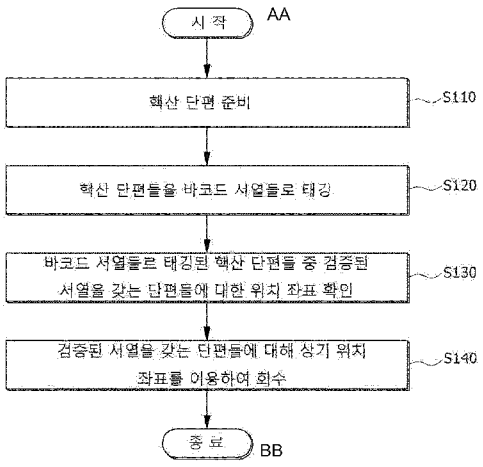
(74) 대리인: **곽현규 (KWAK, Hyun Kyu)**; 06236 서울시 강남구 테헤란로 22길 15 A&C빌딩 4층, Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) **Title:** METHOD AND APPARATUS FOR ACQUIRING HIGH-PURITY NUCLEOTIDES

(54) **발명의 명칭:** 고순도 뉴클레오타이드 획득 방법 및 장치



S110 ... Prepare nucleic acid fragments
S120 ... Tag nucleic acid fragments with barcode sequences
S130 ... Confirm location coordinates of fragments having verified sequences among nucleic acid fragments tagged with barcode sequences
S140 ... Collect fragments having verified sequences using location coordinates
AA ... Start
BB ... End

(57) **Abstract:** The present invention relates to a method and apparatus for acquiring high-purity nucleotides that can be favorably used in a genome/gene synthesis field, as DNA synthesis techniques capable of synthesizing nucleotide sequences with high purity at a low cost.

(57) **요약서:** 본 발명은 저비용 고순도로 염기서열을 합성할 수 있는 DNA 합성기술로서, 유전체/유전자 합성 분야에 유용하게 사용되어질 수 있는 고순도 뉴클레오타이드 획득 방법 및 장치에 관한 것이다.

WO 2018/208133 A1

[다음 쪽 계속]

공개:
— 국제조사보고서와 함께 (조약 제21조(3))
— 명세서의 서열목록 부분과 함께 (규칙 5.2(a))

명세서

발명의 명칭: 고순도 뉴클레오타이드 획득 방법 및 장치

기술분야

[1] 본 발명은 저비용 고순도로 염기서열을 합성할 수 있는 DNA 합성기술로서, 유전체/유전자 합성 분야에 유용하게 사용되어질 수 있는 고순도 뉴클레오타이드 획득 방법 및 장치에 관한 것이다.

배경기술

[2] 마이크로어레이 DNA 합성 방법이 개발되면서 한 번에 매우 많은 종류의 DNA를 저렴하게 합성할 수 있게 되었다. 그러나 마이크로어레이 합성 방법은 화학 합성을 기반으로 하므로 DNA 길이에 제한이 있으며 오류가 많았다. 무엇보다도 섞여있는 형태로 합성되므로 오류가 없는 염기서열을 분리하고자 할 때에는 염기서열을 일일이 확인한 후 분리해야 하며, 이는 합성보다 더 비용이 소요되어 사용상 제한으로 작용할 수밖에 없다. 비록 염기서열을 한꺼번에 병렬적으로 확인할 수 있는 차세대 염기서열 분석기법(NGS)이 적용되면서 염기서열정보를 확인하는 비용은 현저히 줄어들었더라도 정제 기술에 있어 사람이 일일이 손으로 분리해야 했기에 여전히 많은 비용과 시간이 요구되었다.

[3] 즉, 기존의 정제 기술은 오류가 없을 만한 DAN를 무작위로 분리한 다음 분리된 DNA들의 시퀀스를 확인하여 오류 없는 DNA를 찾아내는 방식으로, 무작위로 사람이 손으로 분리한다는 점에서 시간이 많이 소요되며, 시퀀스 확인도 개별적으로 생어 염기서열 분석기법을 사용하고 있기 때문에 비용이 많이 소요된다. 이에 초-저비용(ultra-low cast)로 오류율(error rate)을 개선할 수 있는 기술에 대한 요구가 여전히 필요한 실정이다.

[4] 한편 고속 DNA 시퀀싱 과정이 부분적으로 어드레스할 수 있는 스팟(Roche-454, Illumina 사의 클론성 스팟(clonal spots 등)에 적용될 수 있을지라도, 고속 시퀀싱 플레이트로부터 소망하는 DNA를 수득하는 것과 관련된 어려움으로 인하여 타겟 DNA 분리는 쉽지 않다. 관련하여, Matzas 등에 따르면, 칩-절단된 올리고뉴클레오타이드들이 454 시퀀싱 기술로 시퀀싱되고, 비드 추출 피펫을 이용하여 454 시퀀싱 플레이트로부터 직접 분리된 다음 서열-확인된 올리고뉴클레오타이드들을 프로세싱하여 200bp의 타겟 DNA 단편들의 어셈블리에 이용하였다(Matzas et al. 2010). 상기 연구는 DNA 합성 비용의 절감에 있어 NGS 시퀀싱 기술 및 마이크로칩 올리고뉴클레오타이드의 융합 가능성을 증명한 것이나, 어셈블리된 DNA 단편들이 아닌 칩 올리고뉴클레오타이드들의 고속 시퀀싱을 실시한 기술에 해당한다.

[5] 이에 더 큰 서열로의 DNA 어셈블리와 관련한 기술 개발이 필요하며, 또한 단일세포의 유전체를 분석하고 유전자를 합성하기 위한 기술인 DNA 분자

바코딩 기술을 기반으로 하여 오류 없는 DNA 가닥만을 확인하는 기술의 정확도를 높이는 기술이 요구된다.

발명의 상세한 설명

기술적 과제

[6] 이에 본 발명은 저비용 고순도로 염기서열을 합성하고 검증된 서열만을 확인하여 고 효율 저비용으로 정제하는 핵산분자의 획득 방법 및 장치를 제공한다.

과제 해결 수단

[7] 본 발명의 일 구현예에 따르면, (a) 핵산 단편들을 준비하는 단계; (b) 상기 핵산 단편들을 바코드 서열들로 태깅하는 단계; (c) 상기 바코드 서열들로 태깅된 핵산 단편들 중에서 검증된 서열을 갖는 단편들에 대한 위치 좌표를 확인하는 단계; 및 (d) 상기 검증된 서열을 갖는 단편들을 상기 위치 좌표를 이용하여 회수하는 단계를 포함하는 고순도 핵산분자의 획득방법이 제공된다.

[8] 본 발명의 다른 구현예에 따르면, 서로 다른 종류의 상기 바코드 서열들로 태깅된 핵산 단편들이 개별 스폿 단위별로 구분되어 부착되어 있으며, 상기 개별 스폿들이 이격하여 배열된 지지체가 장착된 스테이지; 상기 바코드 서열들로 태깅된 핵산 단편들 중에서 검증된 서열을 갖는 단편들의 분리를 위한 병렬적 시퀀싱 장치; 상기 병렬적 시퀀싱 장치에서 검증된 서열을 갖는 단편들에 부착된 바코드 서열정보를 지지체 상에서 관찰하기 위한 이미징 장치; 및 상기 검증된 서열을 갖는 단편들이 위치한 좌표에 상기 지지체로부터 분리하기 위해 접촉식 또는 비접촉식으로 에너지를 인가하는 추출 장치를 포함하는 고순도 핵산분자의 획득장치가 제공된다.

[9] 본 발명의 또 다른 구현예에 따르면, 전술한 방법에 의하여 만들어지며, 음의 전하를 띤 단일가닥 DNA 분자들을 갖는 것인 비드 프리(bead free) 핵산 마이크로스피어가 제공된다.

발명의 효과

[10] 본 기술에 따르면, 상업적으로 입수가능한 지지체와 기존의 합성 방법을 이용하되, 1개의 유전자를 구입할 비용으로 500개의 유전자를 합성할 정도로 처리량(throughput)과 오류율(error rate)을 현저하게 개선시킬 뿐 아니라 제조 공정의 단축으로 생산 시간 또한 10배 정도 단축시킨 고 효율 DNA 정제 기술을 제공하는 효과가 있다.

도면의 간단한 설명

[11] 도 1과 도 2는 본 발명의 일 구현예에 따른 검증된 서열을 합성하는 고순도 DNA 회수 공정을 나타낸 개략적인 공정흐름도와 구체적인 공정모식도이다.

[12] 도 3은 본 발명의 일 구현예에 따른 바코드 서열들로 태깅된 핵산 단편들 중 RCA 반응을 수행하도록 디자인된 프라이머(서열목록 9,10)를 이용하여 27 cycle PCR 후 증폭된 DNA 젤 사진이다.

[13] 도 4는 본 발명의 일 구현예에 따른 RCP(RCA product)와 형광 분자가 달린 프로브와의 결합을 통한 RCP(RCA product) 형성을 확인한 사진이다.

[14] 도 5는 본 발명의 일 구현예에 따른 반응에 사용된 칩 구조의 측면도(a)와 상부도(b)이다.

[15] 도 6은 본 발명의 일 구현예에 따른 기판 상에 목적 핵산 분자 캡쳐 후 증폭을 통해 분자 클론을 고정하는 방식을 개략적으로 나타내는 모식도(a)와 지지체에 증폭된 목적 핵산 분자 클론을 고정하는 방식을 개략적으로 나타내는 모식도(b)이다.

[16] 도 7은 본 발명의 일 구현예에 따른 기판 상에서 RCP(RCA product)의 바코드 시퀀스를 확인한 사진으로, Cy7 형광은 RCP 유무 확인용, Cy3: 구아닌(Guanine), Cy5: 아데닌(Adenine), TAM: 티민(Thymine), Texas Red: 시토신(Cytosine) 확인용이다.

[17] 도 8은 레이저로 RCP(RCA product)를 분리한 다음 PCR 증폭을 통해 확인한 사진으로, #1-13: RCP, PCR product(1/10^6): 양성 대조군(positive control), NTC: 음성 대조군(negative control)이다.

[18] 도 9는 본 발명의 일 구현예에 따른 폴로니 시퀀싱(polony sequencing)을 통한 바코드 염기서열 확인 및 검증 그래프로서, square box는 바코드 시퀀스 영역을 나타낸다.

[19] 도 10은 본 발명의 일 구현예에 따른 양성 대조군을 폴로니 시퀀싱을 통해 바코드의 염기서열을 읽은 결과 사진이다.

[20] 도 11은 본 발명의 일 구현예에 따른 임의의 바코드 샘플을 폴로니 시퀀싱을 통해 바코드의 염기서열을 읽은 결과 사진이다.

[21] 도 12는 이하의 도 13에 따른 1개의 RCP(RCA product)가 결합한 SU-8 필라를 폴로니 시퀀싱으로 읽은 결과와, 이들 필라를 레이저로 분리하여 Sanger 시퀀싱으로 바코드의 염기서열을 확인한 결과를 대비한 그래프이다.

[22] 도 13은 본 발명의 다른 구현예에 따른 기판 상에 정전기력에 의한 RCP(RCA product)의 고정을 나타낸 단면도이다.

[23] 도 14는 도 13의 방식으로 기판 상에 고정된 RCP(RCA product)를 나타낸 사진이다.

[24] 도 15는 DNA가 캡쳐된 bead 회수 후, 염기서열 확인을 위해 DNA를 증폭한 뒤 전기영동한 젤 사진이다.

[25] 도 16은 DNA분자 상에서 오류를 확인하면 절단을 일으키는 효소반응 후 전기영동한 젤 사진이다.

[26] 도 17은 본 발명의 일 구현예에 따른 비드 프리(bead free) 핵산 마이크로스피어가 음의 전하를 띠는 롤링된 형태로 제공됨에 따라 비드패킹 등 비드 부착(좌측 도면에 해당) 없이 정전기적 인력등에 의해 지지체의 표면에 직접 고정되는 것(우측 도면에 해당)을 모식화한 도면이다.

발명의 실시를 위한 최선의 형태

[27] 이하, 본 발명을 첨부 도면을 참고하여 보다 상세하게 설명한다. 이하에서는 핵산 분자 중에서 DNA를 예로 들어 설명한 것으로, 본 발명에서 획득 및 분리 대상을 DNA에 한정하는 것은 아니다.

[28] 도 1과 도 2는 각각 본 발명의 일 구현예에 따른 검증된 서열을 합성하는 고순도 DNA 회수 공정을 나타낸 개략적인 공정흐름도와 구체적인 공정모식도이다. 도 1 및 도 2를 참조하면, 단계 S110에서 핵산 단편들을 준비한다. 상기 핵산 단편들은 일례로 핵산 샘플을 단편화하여 얻어진 목적 핵산 분자의 전체 서열의 적어도 일부를 구성하는 단편들을 조립하여 병렬적 시퀀싱 기술로 서열 확인 가능한 크기의 핵산 단편들을 제조한다. 상기 핵산 샘플은 일례로 DNA(deoxyribonucleic acid), RNA(ribonucleic acid), PNA(peptide nucleic acid), LNA(locked nucleic acid), GNA(glycol nucleic acid), TNA(threose nucleic acid), XNA(xeno nucleic acid), 합성 핵산(synthetic nucleic acid), 변형 핵산(modified nucleic acid) 혹은 그 조합을 포함할 수 있으나, 바람직하게는 DNA가 사용되며 이는 시험관 내 혹은 세포 내 증폭 용이성을 고려한 것이다. 상기 DNA는 마이크로어레이에서 화학적인 합성, 생물체로부터 유래된 DNA 등을 통해 수득할 수 있다.

[29] 상기 목적 핵산 분자는 생물체로부터 분리된 것, 핵산 라이브러리로부터 분리된 것, 또는 분리된 핵산 단편을 유전공학적 방법에 의해 변형하거나 조합하여 수득된 것, 화학적으로 합성된 것, 또는 이들의 조합일 수 있다. 상기 목적 핵산 분자는 단일 가닥 또는 이중 가닥일 수 있다.

[30] 상기 목적 핵산 분자는 마이크로어레이 기반 합성을 통해 생성된 것일 수 있다. 마이크로어레이 기반 합성은 고체 기판 상에 ㎛ 단위 내지 ㎝ 단위 수준의 합성 스폿이 간격을 두고 고정되어 있는, 동일, 유사, 또는 상이한 종류의 여러 생화학 분자들을 동시에 병렬적으로 합성하는 기술을 의미한다.

[31] 바람직하게는 상기 핵산 단편들은 낮은 가격으로 수백만 종 이상의 염기서열을 제공하는 DNA 마이크로칩으로부터 유래하거나 합성 올리고뉴클레오타이드들의 풀로부터 유래한 것일 수 있다. 상기 합성 올리고뉴클레오타이드들의 풀은 당업계에 잘 알려진 방법에 따라 제조될 수 있으며, 예를 들어 레진-기반 올리고뉴클레오타디드들로 제조될 수 있으나, 이에 국한되는 것은 아니다. 상기 핵산 단편들은 DNA 마이크로칩으로부터 유래한 것이 바람직하다.

[32] 상기 검증된 서열을 갖는 단편들의 활용을 위해서는 상기 핵산 단편들은 삽입(insertion), 결식(deletion), 전이(transtion), 역전(transversion) 등과 같은 서열 오류들이 없는 것일 수 있다.

[33] 단계 S110의 핵산 단편들은 올리고뉴클레오타이드들의 풀로부터 직접 추출한 것들이거나 일정 길이 이상을 갖도록 하기 위해 올리고뉴클레오타이드들을

증폭 또는 조립한 것들일 수 있다.

[34] 상기 핵산 단편들은 목적 핵산분자의 전체 서열을 포함할 수 있다. 오류없는 검증된 DNA 합성을 위해 상기 핵산 단편들은 병렬적 시퀀싱 기기들로 서열 확인될 수 있다. 상기 병렬적 시퀀싱 기기로 서열확인 가능한 핵산 단편의 길이를 고려할 때 바람직한 상기 핵산 단편의 길이는 20~3,000bp이고, 보다 바람직하게는 200~1,000bp이며, 가장 바람직하게는 300~500bp이다. 이러한 수치 범위에도 불구하고 병렬적 시퀀싱 기기의 성능 향상에 의해 수천 bp가 넘는 DNA를 분석할 수 있을 경우, 상기 핵산 단편들의 크기는 수천 bp 이상의 크기를 가지는 DNA로 확대 적용될 수 있다.

[35] 본 명세서에서 용어 "뉴클레오타이드"는 단일가닥(single strand) 또는 이중가닥(double strand) 형태로 존재하는 디옥시리보뉴클레오타이드 또는 리보뉴클레오타이드이며, 달리 특정하지 않는 한 자연의 뉴클레오타이드의 유사체를 포함한다.

[36] 본 명세서에서 용어 "올리고뉴클레오타이드"는 뉴클레오타이드의 올리고머, 폴리머 또는 이의 모방체를 의미한다. 본 발명의 일 구현예에 따르면, 본 발명의 유전자 증폭은 PCR 등을 이용하여 실시되며, 본 발명의 프라이머가 이러한 유전자 증폭 반응(amplification reactions)에 이용된다.

[37] 본 명세서에 기재된 용어 "증폭반응(amplification reaction)"은 목적 핵산서열을 증폭하는 반응을 의미한다. 다양한 증폭 반응들이 보고되고 있으며, 일례로 중합효소 연쇄반응(PCR), 역전사-중합효소 연쇄반응(RT-PCR), RCA(Rolling cicle amplification), LAMP(Loop-mediated isothermal amplification), SR(Sequence replication), SDA(strand displacement amplification), LCR(Ligase chain reaction), NASBA(Nucleic acid sequence based amplification), MDA(Multiple displacement amplification) 등을 포함하지만 이에 한정되지는 않는다.

[38] 본 명세서에 기재된 용어 "병렬적 시퀀싱 방법"은 달리 한정하지 않는 한, 초병렬 시퀀싱, 생어(Sanger) 시퀀싱, 질량 분석, 전기영동, 혼성화, 디지털 PCR, 대립유전자-특이 PCR, 정량적 PCR, 형광 기반 분류, 또는 그 조합을 포함하는 방법을 지칭한다.

[39] 상기 초병렬 시퀀싱은 합성에 의한 시퀀싱, 이온 토렌트 시퀀싱, 파이로시퀀싱, 라이게이션에 의한 시퀀싱, 나노포어 시퀀싱 및 단일-분자 실시간 시퀀싱 또는 그 조합을 포함하는 방법으로부터 선택되는 것일 수 있다. 예컨대 차세대 염기서열 분석(NGS)를 들 수 있다.

[40] 본 발명의 일 구현예에서, 상기 증폭반응은 가장 잘 알려진 PCR을 사용하여 수행할 수 있다.

[41] 본 발명에 이용되는 프라이머는 주형의 한 부위에 혼성화 또는 어닐링되어 이중쇄 구조를 형성한다. 이러한 이중쇄 구조를 형성하는데 적합한 핵산 혼성화의 조건은 Joseph Sambrook 등, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.(2001) Haymes, B. D.,

등, Nucleic Acid Hybridization, A Practial Approach, IRL Press, Washington, D.C.(1985)에 개시되어 있다.

[42] 다양한 DNA 중합효소(DNA polymerase)가 본 발명의 증폭에 이용될 수 있으며, E.Coli DNA 중합효소 I의 "클레나우" 단편, 열안정성 DNA 중합효소, 박테리오파지 T7 DNA 중합효소를 포함한다. 바람직하게는 중합효소는 다양한 박테리아 종으로부터 얻을 수 있는 열안정성 DNA 중합효소이고, 이는 Thermus aquaticus(Taq), Thermus thermophilus(Tth), Thermus filiformis, Thermis flavus, Thermococcus literalis, Pyrococcus furiosus(Pfu), Thermus antranikianii, Thermus caldophilus, Thermus chliarophilus, Thermus flavus, Theremus igniterrae, Theremus lacteus, Thermus oshimai, Thermus ruber, Thermus rubes, Thermus scotoductus, Thermus silvanus, Thermus species Z05, Thermus species sps 17, Thermus thermophilus, Thermotoga maritima, Thermotoga neapolitana 및 Thermosipho africanus의 DNA 중합효소 및 Phusion 중합효소를 포함하며, 가장 바람직하게는 Pyrococcus furiosus(Pfu) 또는 Phusion High-Fidelity DNA 중합효소를 이용한다.

[43] 중합 반응을 실시할 때, 반응 용기에 반응에 필요한 성분들을 과량으로 제공하는 것이 바람직하다. 증폭 반응에 필요한 성분들의 과량은, 증폭반응이 성분의 농도에 실질적으로 제한되지 않는 정도의 양을 의미한다. Mg^{2+}와 같은 조인자, dATP, dCTP, dGTP 및 dTTP를 소망하는 증폭 정도가 달성될 수 있을 정도로 반응 혼합물에 제공하는 것이 소망된다. 증폭 반응에 이용되는 모든 효소들은 동일한 반응 조건에서 활성 상태일 수 있다. 실제로 완충액은 모든 효소들이 최적의 반응 조건에 근접하도록 한다. 따라서, 본 발명의 증폭 과정은 반응물의 첨가와 같은 조건의 변화 없이 단일 반응물에서 실시될 수 있다.

[44] 본 발명에 있어 어닐링은 타겟 뉴클레오타이드 서열과 프라이머 사이에 특이적 결합을 가능하게 하는 엄격한 조건 하에 실시된다. 어닐링을 위한 엄격한 조건은 서열-의존적이며 주위 환경적 변수에 따라 다양하다. 이렇게 증폭된 올리고뉴클레오타이드 풀을 이용하여 제1차 증폭산물들을 생성시키고 이를 이용하여 제2차 증폭산물들을 제조하여 보다 더 큰 목적 핵산분자(예를 들어 10kb 이상의 핵산분자)를 어셈블리시킬 수 있다.

[45] 본 발명에서 사용하는 용어 "프라이머"는 올리고뉴클레오타이드를 의미하는 것으로, 핵산쇄(주형)에 상보적인 프라이머 연장 산물의 합성이 유도되는 조건, 즉 뉴클레오타이드와 DNA 중합효소와 같은 중합체의 존재, 그리고 적합한 온도와 pH 조건에서 합성의 개시점으로 작용할 수 있다. 바람직하게는 프라이머는 디옥시리보뉴클레오타이드이며 단일쇄이다. 본 발명에서 이용되는 프라이머는 자연 dNMP(dAMP, dGMP, dCMP, dTMP에 해당), 변형 뉴클레오타이드 또는 비-자연 뉴클레오타이드를 포함할 수 있다. 또한 상기 프라이머는 리보뉴클레오타이드도 포함할 수 있다.

[46] 상기 프라이머는 중합제(DNA 중합효소에 해당)의 존재 하에 연장 산물의 합성을 프라이밍시킬 수 있을 정도로 충분히 길어야 한다. 프라이머의 적합한

길이는 다수의 요소, 예컨대, 온도, 응용분야 및 프라이머의 공급원(source)에 따라 변화가 있지만 전형적으로 15~30 뉴클레오타이드이다. 짧은 프라이머 분자는 주형과 충분히 안정된 혼성 복합체를 형성하기 위하여 일반적으로 보다 낮은 온도를 요구한다. 상기 목적 핵산 분자를 증폭하기 위한 프라이머 세트는 상기 목적 핵산 분자의 말단 서열 영역에 결합할 수 있다.

[47] 본 명세서에서 사용되는 용어 "어닐링" 또는 "프라이밍"은 주형 핵산에 올리고디옥시뉴클레오타이드 또는 핵산이 병치(apposition)되는 것을 의미하며, 상기 병치는 중합효소가 뉴클레오타이드를 중합시켜 주형 핵산 또는 그 일부에 상보 핵산 분자를 형성하게 한다. 본 명세서에서 사용되는 용어 "혼성화(hybridization)"는 2개의 단일 가닥 핵산이 상보적인 염기 서열들의 페어링(pairing)에 의하여 이합체 구조(duplex structure)를 형성하는 것을 의미한다. 혼성화는 단일 가닥 핵산 서열 간의 상보성이 완전할 경우(perfect match) 일어나거나 일부 미스매치(mismatch) 여기가 존재하여도 일어날 수 있다. 혼성화에 필요한 상보성의 정도는 혼성화 반응 조건에 따라 달라질 수 있으며, 특히 온도에 의하여 조절될 수 있다.

[48] 본 명세서에서 언급되는 용어 "상보적(complementary)"은 어떤 특정한 혼성화(hybridization) 또는 어닐링 조건 하에서 상술한 뉴클레오타이드 서열에 선택적으로 혼성화할 수 있을 정도의 상보성을 갖는 것을 의미하며, 실질적으로 상보적 및 완전히 상보적인 것을 모두 포괄하는 의미로, 이중에서 바람직하게는 완전히 상보적인 것이다.

[49] 도 1을 재참조하면, 단계 S120에서 상기 핵산 단편들을 바코드 서열들로 태깅한다. 상기 바코드 서열들은 전 단계에서 제공되는 핵산 단편들 중에서 검증된 단편들 또는 소망하는 단편들을 회수하거나 이들을 선택적으로 증폭하고 어셈블리 하여 목적 핵산 분자를 합성하기 위해 상기 핵산 단편들에 도입될 뿐 아니라 후술하는 단계 S130에서 보듯이, 병렬적 시퀀싱을 통해 전체 염기서열을 확인하고, 이중에서 검증된 염기서열을 갖는 바코드 정보를 활용하여 지지체 상에서 해당 바코드 영역을 분리하면 염기서열 확인과 오류없는 목적 핵산 분자의 분리를 간편하면서도 효과적으로 제공할 수 있어 바람직하다.

[50] 상기 바코드 서열들은 상기 핵산 단편들에 부가되어 서로 다른 핵산 단편들을 구분할 수 있도록 하는 물질이라면 그 종류가 한정되는 것은 아니다. 각 핵산 단편들간의 구분을 위해 상기 바코드 서열들의 종류의 수는 상기 핵산 단편들의 종류의 수와 대응되거나 이보다 많은 것이 바람직하다. 예를 들어, 상기 바코드 서열들은 임의 또는 계획적으로 설계된 2종 이상의 올리고뉴클레오타이드들의 혼합물일 수 있다. 상기 바코드 서열의 길이는 5 내지 300 bp일 수 있다.

[51] 본 발명의 일 구현예에 따르면, 상기 핵산 단편들을 바코드 서열들로 태깅하는 단계는 고정된 핵산(nucleic acid) 시퀀스와 목적 핵산분자에 태깅될 수 있는 핵산 시퀀스가 같이 포함된 프라이머를 준비하는 단계; 상기 프라이머를 목적

핵산분자와 함께 섞어 반응용액을 준비하는 단계; 상기 반응용액으로부터 목적 핵산 분자를 얻는 단계를 포함하는 핵산 증폭 방법을 통해 수행될 수 있다.

[52] 상기 프라이머의 일부 또는 전체가 DNA, RNA 중 하나일 수 있다.

[53] 상기 시퀀싱 태깅 핵산 시퀀스의 길이가 4nt이상 35nt이하일 수 있다. 상기 시퀀싱 태깅 핵산 시퀀스가 두개 이상 포함된 프라이머를 이용한 것일 수 있다. 상기 목적 핵산 분자에 태깅될 수 있는 핵산 시퀀스가 4nt이상 10nt이하일 수 있다. 상기 목적 핵산 분자에 태깅될 수 있는 핵산 시퀀스가 랜덤 시퀀스일 수 있다.

[54] 상기 바코드 서열들의 태깅은 특별히 제한되지 않으나, PCR, 에멀젼 PCR, 라이게이션(ligation) 또는 그 조합을 포함하는 방법 중에서 선택되는 어느 하나의 방법을 사용하여 수행될 수 있다. 구체적으로는, 상기 바코드 서열들에 상응하는 프라이머들로 목적 단편들을 선택적으로 증폭하여 회수하는 방식으로 수행될 수 있다.

[55] 도 1 및 도 2를 재참조하면, 단계 S130에서 상기 바코드 서열들로 태깅된 핵산 단편들 중에서 검증된 서열을 갖는 단편들에 대한 위치 좌표를 확인한다. 상기 태깅된 핵산 단편들의 서열의 확인은 바람직하게는 병렬적 시퀀싱 방법에 의해 수행될 수 있다. 그 결과 상기 태깅된 핵산 다편들의 서열을 태깅된 바코드 서열과 함께 확인할 수 있다.

[56] 본 발명의 일 구현예에 따르면, 본 발명의 상기 병렬적 시퀀싱 또는 고속 시퀀싱은 당업계에 잘 알려진 고속 시퀀싱 방법에 의해 실시되며, 예를 들어 Illumina 또는 기타 판독 길이가 100 bp 이상인 고속 시퀀싱 기기에 의해 실시될 수 있다(도 1의 단계 S130, 도 2 참조). 이로부터 오류없이 검증된 DNA와 바코드 서열정보를 획득할 수 있다.

[57] 본 명세서에서 핵산 단편에 부가되는 바코드 서열을 포함하는 서열을 일명“바코드 프라이머”라고 명명한다. 본 명세서의 용어 “어댑터 서열”은 상기 해당 단편의 고속 시퀀싱 분석을 가능하게 하는 서열로, 예를 들어 본 발명에 이용된 Illumina 시퀀싱을 위해 상업적으로 이용가능한 모든 서열을 포함한다. 바람직하게는, 본 발명의 어댑터 서열은 Illumina 시퀀싱 플랫폼의 어댑터 서열 및 다른 종류의 차세대 시퀀싱 기술의 어댑터 서열을 포함하지만, 이에 한정되는 것은 아니다.

[58] 본 발명의 일 구현예에 따르면 서열 확인을 위해 상기 태깅된 핵산 단편들을 템플레이트로 하여 상기 어댑터 서열에 결합하는 프라이머를 이용한 증폭 과정이 수행될 수 있다. 상기 바코드 서열의 길이는 특별히 제한되지 않으며, 핵산 단편을 포함한 전체 서열에 대한 시퀀싱 성능을 고려하여 예를 들어 5~300bp, 바람직하게는 10~100bp, 보다 바람직하게는 12~40bp일 수 있다. 시퀀싱 기기에 따라 상기 수치 범위는 변경될 수 있다.

[59] 상기 프라이머의 일부 또는 전체가 DNA, RNA 중 하나일 수 있다.

[60] 상기 프라이머에 선별적으로 캡쳐 가능한 분자가 결합되어있을 수 있다. 상기

선별적으로 캡쳐 가능한 분자가 비오틴, 비오틴 dT, 비오틴-TEG, Dual 비오틴, PC-비오틴, 아크릴기(acrylic group), 싸이올기(thiol group), 다이싸이올기(dithiol group), 아민기(amine group), NHS 에스터, 아자이드, 헥시닐(Hexynyl), 옥타디닐(Octadiynyl) dU, his 태그(his tag), 프로테인 태그(protein tag), polyA (poly adenine), NGS용 시퀀싱 어댑터의 일부, 10nt이상의 고정된 DNA시퀀스로 이루어진 군중에서 선택되는 것일 수 있다.

[61] 상기 캡쳐 가능한 분자에 알맞은 캡쳐 방법을 이용하여 상기 프라이머로부터 원하는 핵산 분자만을 선별적으로 얻는 단계를 더 포함할 수 있다. 상기 캡쳐 방법은 아비딘, 뉴트라비딘, 스트렙타비딘, 카복시기(carboxylic group), 싸이올(thiol), 알킨(alkyne), Digoxigenin, 아자이드(azide), 헥시닐(Hexynyl)로 이루어진 군중에서 선택되는 물질로 이루어지는 방법이거나 PCR(Polymerase Chain reaction)일 수 있다.

[62] 상기 캡쳐 이전에 상기 목적 핵산 분자로 증폭된 핵산을 지지체에 고정하는 과정을 더 포함할 수 있다. 상기 캡쳐 이후에 상기 목적 핵산 분자로 증폭된 핵산을 지지체에 고정하는 과정을 더 포함할 수 있다. 상기 캡쳐 이전에 상기 목적 핵산 분자로 증폭된 핵산을 조각내는(fragmentation) 과정을 더 포함할 수 있다. 상기 캡쳐 이후에 상기 목적 핵산 분자로 증폭된 핵산을 조각내는 과정을 더 포함할 수 있다. 상기 캡처를 통해 선별적으로 얻어낸 분자가 DNA이중가닥(dsDNA) 또는 ssDNA일 수 있고, 반응 공정상 ssDNA가 풍부한 것이 바람직하다. 상기 프라이머가 입체장애(steric hindrance)를 통해 단백질의 접근을 방해하는 물질이 결합되어있는 프라이머일 수 있다.

[63] 상기 캡쳐 이전에 상기 목적 핵산 분자로 증폭된 핵산을 지지체에 직접 고정하는 과정을 포함하고, 상기 캡쳐 단계는 생략할 수 있다. 상기 캡쳐를 통해 선별적으로 얻어낸 분자, 또는 캡쳐 이전에 직접 고정시킨 분자에 혼성화(hybridization)될 수 있는 프로브를 혼성화시키고 라이게이션시켜 원형화(circularization) 처리하고 DNA 중합 효소를 이용하여 증폭 처리하는 방식으로 RCP(RCA product)를 제조할 수 있다. 상기 DNA 중합효소는 일례로 Bst DNA 폴리머라아제, 엑소뉴클레아제 마이너스, Tth DNA 폴리머라아제, 피로파지 (PyroPhage TM) 3173 DNA 폴리머라아제, BcaBEST DNA 폴리머라아제 및 파이(Phi) 29 DNA 폴리머라아제 등을 사용할 수 있으며, 특히 RCA 반응효율이 하나의 염기서열 당 10^(-8)의 확률의 오류를 가지는 파이(Phi) 29 DNA 폴리머라아제를 사용하는 것이 바람직하다. 상기 DNA 중합효소는 일례로 30℃에서 분당 1nmole의 dNTP를 중합할 수 있는 분량으로 포함될 수 있다.

[64] 상기 RCP(RCA product)는 지지체 상에 고정된 형태로 제공되는 것이 바코드 서열정보 확인 후 해당 위치에서 소망하는 RCP(RCA product)를 고순도로 분리해낼 수 있어 바람직하다. 게다가 상기 RCP(RCA product)는 지름이 5㎛ 이하로서 이들을 지지체 상에 고정하여 얻어진 서열들로 태깅된 단편들은 그

지름이 30 ㎛ 이하의 분자 클론 형태로 형성될 수 있으므로 종전 마이크로웰 지지체를 비롯한 특정 형상을 제공하는 지지체 대비 지지체 표면의 전체 면적을 충분히 활용할 수 있어 분자클론의 처리량을 현저하게 개선할 수 있다.

[65] 상기 서열 확인을 통해 상기 핵산 단편들 중 검증된 핵산 단편들 및 이에 부가된 바코드 서열을 동정할 수 있다. 본 발명에서는 이렇게 동정된 바코드 서열을 갖는 핵산 단편을 손쉽게 분리하도록 확인된 바코드 서열을 갖는 단편들이 고정된 지지체를 준비하고, 지지체 상에서 해당 바코드 서열을 나타내는 위치 좌표에 근거하여 검증된 핵산 단편을 신속하게 분리할 수 있다.

[66] 일례로, 도 1 및 도 2를 재참조하면, 앞서 S120 단계에서 바코드 서열들로 태깅된 단편들의 전체 혹은 일부 염기서열을 병렬적 시퀀싱 방법에 의해 확인하여 검증된 서열을 갖는 단편들의 바코드 서열을 확인하는 병렬적 시퀀싱 단계와 상기 검증된 서열을 갖는 단편들로서 확인된 바코드 서열을 갖는 단편들이 고정된 지지체를 준비하고, 해당 지지체 상에서 해당 바코드 서열을 나타내는 위치 좌표를 확인하는 시퀀싱 단계가 동시 혹은 순차적으로 수행될 수 있다(도 1의 단계 S130 및 도 2 참조).

[67] 상기 시퀀싱 단계는 예컨대, i) 상기 S120 단계에서 바코드 서열들로 태깅된 단편들을 지지체 상에 고정하는 단계; 및 ii) 상기 지지체 상에 고정된 태깅된 단편들의 바코드 서열을 확인하여 상기 병렬적 시퀀싱 단계에서 검증된 서열을 갖는 단편들로 확인된 바코드 서열에 대한 위치 좌표를 확인하는 단계로 구성될 수 있다(도 1, 2 참조).

[68] 구체적으로, 상기 i) 단계는, 상기 바코드 서열들로 태깅된 단편들을 RCA(Rolling cicle amplification), LAMP(Loop-mediated isothermal amplification), SR(Sequence replication), SDA(strand displacement amplification), LCR(Ligase chain reaction), NASBA(Nucleic acid sequence based amplification) 또는 MDA(Multiple displacement amplification)로 증폭된 다음 지지체 표면과의 상호작용에 의해 지지체 상에 고정된 것일 수 있다.

[69] 여기서 상기 RCA 또는 MDA는 phi29 중합 효소를 비롯한 DNA 중합 효소들을 사용하여 수행될 수 있다. 상기 DNA 중합 효소들은 하나의 염기서열 당 10^(-8)의 오류율을 가지는 것으로 고순도 DNA 합성이 가능 하다.

[70] 상기 바코드 서열들로 태깅된 단편들, 상기 증폭물 또는 그 조합을 고정하는 지지체 표면층은 일례로 평평한 형상, 우물 형상, 기둥 형상, 반복 무늬 혹은 격자 형상이 패터닝된 구조, 또는 그 조합을 포함하는 구조를 가질 수 있다.

[71] 상기 지지체 표면층은 하나의 DNA cluster가 하나의 패턴에 포함할 수 있도록 구조가 조절된 것일 수 있다. 예를 들어 미세입자의 높이나 넓이, 간격을 작게 조절하여 이를 성취할 수 있다. 상기 지지체 표면층은 리소그래피를 이용하여 패터닝을 형성할 수 있다. 상기 지지체 표면층의 패터닝된 구조는 광중합(photopolymerization)에 의하여 형성될 수 있다. 상기 지지체 표면층의 패터닝된 구조는 빛의 회절이나 간섭 현상을 이용하여 100 ㎛ 내지 5 ㎛의 작은

크기를 형성할 수 있다.

[72] 상기 바코드 서열들로 태깅된 단편들, 상기 증폭물 또는 그 조합을 고정하는 지지체 표면은 일례로 폴리머, 실리카, 하이드로겔, 유리 또는 그 조합을 포함하는 물질로 이루어질 수 있다.

[73] 상기 지지체 표면과의 상호작용에 의해 지지체 상에 고정하는 단계는 전술한 지지체 상에 전자기 인력, 흡착, 단백질 결합, 자기 조립, 화학작용기의 결합, 또는 그 조합을 포함하는 상호작용에 의해 고정되는 것일 수 있다.

[74] 본 발명에서 상기 바코드 서열들로 태깅된 단편들, 상기 증폭물 또는 그 조합을 지지체 상에 고정하는 단계는 필요에 따라 기판의 표면 혹은 내부, 셀의 표면 혹은 내부, 기판 상에 패터닝된 표면 혹은 내부 또는 그 조합을 포함하는 위치에 고정되는 것일 수 있다. 특히 고정의 용이성을 고려할 때 상기 지지체 혹은 기판은 평판형 플레이트 기판 혹은 역방향의 나노/마이크로 우물 구조를 갖는 기판인 것이 바람직하다.

[75] 상기 바코드 서열들로 태깅된 단편들, 상기 증폭물 또는 그 조합을 지지체 상에 고정하는 단계는 또한 상기 바코드 서열들로 태깅된 단편들의 목적 핵산 분자와 상보적 시퀀스를 갖는 프로브를 지지체에 고정한 다음 상기 프로브에 상기 바코드 서열들로 태깅된 단편들을 캡쳐하고 RCA(Rolling cicle amplification), LAMP(Loop-mediated isothermal amplification), SR(Sequence replication), SDA(strand displacement amplification), LCR(Ligase chain reaction), NASBA(Nucleic acid sequence based amplification) 또는 MDA(Multiple displacement amplification)로 증폭시켜 수행된 것일 수 있다. 여기서 상기 RCA 또는 MDA는 phi29 중합 효소를 비롯한 DNA 중합 효소들을 사용하여 수행될 수 있다. 상기 DNA 중합효소는 일례로 하나의 염기서열 당 10^(-8)의 합성 오류율을 가지는 것으로 고순도 DNA 합성이 가능하다.

[76] 본 발명에서 상기 위치 좌표 확인은 일례로 상기 바코드에 형광 분자를 붙인 후 직접 형광을 관찰하는 방식으로 수행될 수 있다. 여기서 형광 분자는 플루오레세인 이소티오시아네이트(fluorescein isothiocyanate), 5,6-카복시메틸 플루오레신(5,6-carboxymethyl fluorescein), 텍사스 레드(Texas red), 니트로벤즈-2-옥사-1,3-디아졸-4-일(nitrobenz-2-oxa-1,3-diazol-4-yl), 쿠마린(coumarin), dansyl chloride, 로다민(rhodamine), 아미노메틸 쿠마린(aminomethyl coumarin), 에오신(Eosin), 에리스로신(Erythrosin), BODIPY®, Cascade Blue®, Oregon Green®, 피렌(pyrene), 리사민(lissamine), 크산텐(xanthene), 아크리딘(acridine), 옥사진(oxazine), 피코에리쓰린(phycoerythrin), Cy3, Cy3.5, Cy5, Cy5.5, Cy7 중에서 선택되는 1종 이상을 태깅하도록 바코드 서열을 개질할 수 있고, 이를 통해 증폭 반응산물 중 바코딩이 된 증폭산물의 양을 정량할 수 있다.

[77] 구체적으로, 상기 바코드 시퀀싱은 일례로, SBS(sequencing by synthesis), SBL(sequencing by ligation), ELISA 또는 이들의 조합으로 수행될 수 있다.

[78] 따라서 본 발명에서 전술한 위치 좌표 확인 단계는 전술한 시퀀싱을 통해 소망하는 핵산 단편들과 이에 태깅된 바코드 서열들의 서열이 확인되었으므로 상기 바코드 서열들에 상응하는 프라이머들로 소망하는 상기 핵산 단편들을 선택적으로 증폭하여 회수, 상기 바코드 서열들을 갖는 핵산 단편들을 지지체 상에 고정한 다음 미리 확인된 바코드 서열들을 이용하여 원하는 핵산 단편들을 해당 지지체 상의 위치 좌표를 바탕으로 회수하는 방식 또는 그 조합으로 신속하게 회수할 수 있다. 예를 들어 상기 검증된 핵산 단편들은 오류없는 핵산 단편들일 수 있다.

[79] 상기 신속한 회수 처리에 이미징을 제공하는 컴퓨터 프로그램이 이용될 수 있다. 구체적으로, 상기 핵산 단편들의 서열을 이미징을 제공하는 컴퓨터 프로그램을 이용하여 가상적으로 어셈블리시켜 소망하는 목적 핵산분자의 전체 서열과 비교한다. 이후 가장 최적화된 DNA 단편이 갖는 바코드 서열정보를 기반으로 별도로 제작하여 놓은 지지체 상에서 대응하는 최적화된 핵산 단편들을 효과적으로 회수하는 단계를 수행할 수 있다(도 1의 단계 S140 및 도 2 참조).

[80] 구체적으로, 상기 전체 염기서열 시퀀싱 결과와 바코드 정보의 시퀀싱 결과를 매칭하고 원하는 시퀀싱을 가지는 핵산 분자의 위치를 지지체 상에서 확인할 수 있다. 상기 원하는 핵산 분자로 확인된 지지체의 위치를 기반으로 추출식 또는 비추출식으로 해당 핵산 분자를 분리할 수 있다. 일례로, 레이저를 사용하는 것이 손상 없이 해당 핵산 분자를 분리할 수 있고 자동화 공정을 제공할 수 있어 더욱 바람직하다.

[81] 상기 위치 좌표를 이용하여 회수하는 방식은 초음파, 공기압, 레이저 또는 그 조합을 포함하는 방법에 의해 접촉식 또는 비접촉식으로 에너지를 인가하는 방식으로 수행되는 것이 정제 효율을 고려할 때 바람직하나 이에 한정하는 것은 아니다.

[82] 구체적으로, 원하는 시퀀스의 핵산분자로 확인된 영역에 레이저를 영역의 외곽부를 포함하여 영역 전체에 인가하고 원하는 시퀀스를 가진 핵산 분자를 분리하거나, 케미컬로 부분 작용기를 제거하거나, 전기력 혹은 열 등을 사용하여 분리할 수 있다. 상기 비접촉식으로 에너지를 인가하는 방식은 초음파, 공기압 및 레이저로 이루어진 군중에서 선택되는 1종 이상의 방식일 수 있다. 일례로 상기 비접촉식으로 에너지를 인가하는 방식은 펄스 에너지의 입사에 의해 펄스 레이저 삭마(pulse laser ablation) 혹은 복사압 배출(radiation pressure ejection)이 일어나는 방식일 수 있다.

[83] 상기 펄스 레이저는 10 내지 10,000nm, 바람직하게는 20 내지 5,000nm, 보다 바람직하게는 100 내지 2,000nm의 파장을 가질 수 있다. 가시광선 영역 또는 적외선 영역을 포함하는 상기 범위에서는 전자기장이 광학 부품에 큰 영향을 주지 않으며 기판 혹은 원하는 바코드 서열을 갖는 핵산 분자에 충분한 에너지를 전달 가능하다. 또한 대부분의 상용 펄스 레이저는 상기 범위에서 동작하기

때문에 시스템의 구현에 용이하며, 희생 층을 이용한 기판을 사용하는 경우에도 시스템의 큰 변경 없이 본 발명에 개시된 기술을 실시할 수 있다.

[84] 상기 펄스 레이저는 1as 내지 1ms, 바람직하게는 1fs 내지 100ns의 펄스 조사시간(pulse duration)을 가질 수 있다. 상기 펄스 조사시간은 상기 펄스 레이저로 인한 펄스 레이저 삭마가 발생할 시 상기 분리된 기판들과 바코드 서열을 갖는 핵산 분자의 분리 진행 경로가 보다 일정하여 회수가 용이하다는 장점이 있다. 한편, 상기 펄스 레이저의 출력은 펄스 당 10 내지 1 kJ/㎠, 바람직하게는 펄스 당 100 내지 300 J/㎠일 수 있다. 상기 펄스 조사시간 및 상기 출력에서 상기 펄스 레이저로 인한 펄스 레이저 삭마가 발생할 시에 원하는 바코드 서열을 갖는 핵산 분자 분리시 덜 손상을 받을 수 있고 이는 분리된 핵산의 후처리 과정을 수행할 때 높은 효율을 가져줄 수 있다.

[85] 상기 원하는 바코드 서열을 갖는 핵산 분자를 상기 기판으로부터 분리시키는 단계는 원하는 바코드 서열을 갖는 핵산 분자를 저장용기(reservoir)로 이전시키는 과정을 포함한다. 상기 저장용기로의 이전은 분리한 핵산 분자를 보관하고 기타 반응물과의 반응이 필요할 시 사용하기 위해 필요한 과정이다. 상기 저장용기는 물리적 혹은 화학적 반응을 일으키거나 이를 관측할 목적으로 제작된 용기를 포함할 수 있다. 또한 상기 저장용기는 생화학 분자의 보관을 위해 제작된 통상의 용기를 포함할 수 있다. 상기 저장 용기는 예를 들어, 각 웰의 부피가 1pL 내지 1uL인 마이크로웰의 어레이 구조일 수 있다.

[86] 상기 핵산 분자의 추출을 위해 사용되는 추출 도구는 위치 이동을 위한 장치와 결합되어 있을 수 있다. 상기 위치 이동을 위한 장치는 전동장치일 수 있으며, 바람직하게는 1mm 이하, 바람직하게는 100 ㎛ 이하, 보다 바람직하게는 5㎛ 이하의 미세 범위에서 조작가능한 정밀도를 가질 수 있으며, 이러한 범위에서 원하는 핵산 분자의 정확한 분리를 가능하게 한다.

[87]

[88] 본 발명의 다른 구현예에 따르면, 타겟 핵산 분자를 PCR로 증폭하여 바코드 시퀀스를 갖는 타겟 핵산 분자가 최소 2 이상이 되도록 준비하는 단계; 및 최소 2 이상의 바코드가 부착된 복제 핵산 분자를 랜덤으로 분리하여, 두 개 이상의 시퀀싱 단계를 수행하고 바코드-시퀀싱 정보 매칭을 수행하는 단계를 포함하되, 상기 두 개 이상의 시퀀싱 단계는 바코드 분자가 부착된 복제 DNA로 RCP(Rolling circulation amplification product)를 포함하는 핵산분자 클러스터를 형성하고 상기 핵산분자 클러스터는 지지체에 고정되어 있으며, 상기 지지체에 고정된 핵산분자 클러스터 중에서 바코드 영역을 시퀀싱하는 단계를 포함하는 고순도 핵산분자 획득 방법을 제공한다.

[89] 구체적으로, 서로 다른 바코딩 핵산 시퀀스를 가지는 두 종류 이상의 타겟 핵산 분자를 랜덤으로 분리하고, 해당 분리군에 대하여 지지체 상에 타겟 핵산 분자를 고정하고 RCP(RCA product)에 대한 바코드 영역 시퀀싱을 수행하며, 동시에 혹은 순차적으로, 잔류 분리군에 대해 차세대 염기서열 분석으로 전체 시퀀스

정보를 얻어 상기 지지체의 어느 위치에 원하는 시퀀스가 있는지 정보를 제공하는 시퀀스 정보 매칭을 수행하는 단계를 포함하는 고순도 핵산분자 획득 방법을 제공한다.

[90] 또한, 상기 지지체 상에 타겟 핵산 분자 고정은 지지체 상에 패터닝된 미세입자에 RCP(RCA 반응으로 이루어진 product)를 고정시켜 수행되며, 예컨대 3-아미노프로필트리에톡시실란(APTES)과 같은 다분지쇄 실란 화합물을 코팅을 통해 지지체 표면에 아민기를 부착하거나, 이와 같은 방식으로 히드록실기, 에폭시기 또는 아크릴기를 부착시킨 기판에 용액 상에서 RCA 반응을 진행하여 형성된 RCP(RCA product)를 적재하여 고정시킬 수 있으며, 바코딩된 DNA 분자를 RCA 반응을 진행하여 생성된 RCP(RCA product)를 지지체에 분사하여 고정시킬 수도 있다.

[91] 특히, 상기 S130 단계에서 지지체 상에 RCA 반응을 수행하여 RCP(RCA product)를 고정한 경우, 상기 지지체 상에 대응하는 최적화된 핵산 단편들은 바코드 서열정보를 갖는 복제 핵산 분자로, 폴로니(polony), RCP(RCA product)를 포함하는 핵산 분자 클러스터를 형성하는 것일 수 있다. 여기서 복제 핵산 분자로, 폴로니(polony), RCP(RCA product)를 포함하는 핵산 분자 클러스터는 시퀀스 분석이 가능한 어떠한 형태의 핵산 분자 산물로도 대체 가능하다.

[92] 상기 지지체는 해당 지지체에 핵산 분자를 패킹한 다음 상기 핵산 분자 상에 핵산 분자 클러스터를 고정하거나, 혹은 해당 지지체 상에 직접 핵산 분자 클러스터를 고정하는 것일 수 있다. 관련하여 본 발명의 일 구현예에 따른 기판 상에 목적 핵산 분자 캡쳐 후 증폭을 통해 분자 클론을 고정하는 방식을 개략적으로 나타내는 모식도(a)와 지지체에 증폭된 목적 핵산 분자 클론을 고정하는 방식을 개략적으로 나타내는 모식도(b)를 도 6 등에 나타낸 것으로, 도 5는 패터닝 기판에 해당한다. 도 5는 본 발명의 일 구현예에 따른 반응에 사용된 칩 구조의 측면도(a)와 상부도(b)에 해당한다. 여기서 패터닝은 투명한 기판(ex. 유리 등) 혹은 희생층 및 폴리머 층을 포함하는 투명 기판에 RCP(RCA product)(핵산 분자 클러스터에 해당)를 필요에 따라서는 격자 형상으로 패킹한 것(미도시)일 수 있다. 상기 RCP(RCA product)의 패킹 방법으로는 이에 한정하는 것은 아니나, 리소그래피, 자기 조립, 등을 사용할 수 있다. 기타 단백질 결합, 화학작용기의 결합, 전자기 인력과 흡착 등의 상호작용이 가능하다. 패킹이 아닌 직접 고정 방법으로도 어레이를 형성할 수 있다.

[93] 상기 지지체는 SU-8 포토레지스트(epoxy-based negative photoresist)를 비롯한 지지체 상에 RCP(RCA product)의 규칙적인 배열을 인위로 보조하는 수단(이하 필라(pillar)라고도 함)을 필요로 할 수 있다. 상기 필라는 원하는 배열의 규칙성을 제공하기 위한 것으로, 상기 필라들의 간격은 1㎛ 내지 5mm, 바람직하게는 100nm 내지 1mm, 보다 바람직하게는 1㎛ 내지 500 ㎛의 거리를 가질 수 있다.

[94] 도 6은 본 발명의 일 구현예에 따른 기판 상에 DNA 캡쳐를 통한 DNA 분자 고정 후 RCA 반응을 수행하여 고정하는 방식(상부 도면)과 RCA 증폭산물을

고정하는 방식(하부 도면)을 나타낸 단면도이다. 상기 패터닝 기판 상에서 RCP(RCA product)는 하나의 핵산 분자 클러스터만 포함하게끔 생성할 수 있고, RCP(RCA product)의 넓이, 간격을 작게 조절하여 기판 위에 고정된 개수를 자유로이 조절할 수 있다. 도 6의 상부 도면 및 하부 도면을 참조하면, 지지체에 고정된 DNA 클러스터가 RCP(RCA product)의 형태로 말려(rolling) 양 말단 바코드가 연결되거나 혹은 하나 이상의 바코드가 연결된 형태를 띠면서 작은 간격으로 빼곡히 지지체 상에 형성할 수 있을 것으로, 결과 처리율 개선에 특히 바람직하다.

[95] 다음 합성된 바코드 올리고를 이용한 증폭(PCR) 또는 혼성화(hybridization)을 통해 핵산 단편들(오류를 가진 단편들과 오류없는 단편들의 혼합물) 중, 목적 DNA의 합성을 최적화할 수 있는 오류없는 단편들만을 회수한다. 예를 들어, 상기 합성 바코드 서열을 이용한 오류없는 단편들을 획득하기 위한 방법은 PCR 이외에도, 소망하는 바코드 서열을 바이오틴-비드(biotynylated bead), 또는 마그네틱 비드에 붙여서 소망하는 오류없는 단편들을 획득하는 방법 등과 같은 혼성화 방법을 포함하지만, 이에 한정되는 것은 아니다. 또한, 상기 (d) 단계에서 회수된 검증된 서열들을 갖는 단편들로부터 필요에 따라 특정 염기서열의 증폭 또는 절단 등의 공지된 방법을 통해 추가로 원하는 염기서열을 가진 핵산 분자를 획득할 수 있다.

[96]

[97] 본 발명의 또 다른 구현예에 따르면, 본 발명의 오류없는 것으로 검증된 바코딩된 핵산 단편들의 길이는 200 bp 이상일 수 있다. 또한 차세대 시퀀싱의 길이가 1,000 bp가 없는 DNA를 분석할 수 있는 기기를 사용한다면 오류없는 바코딩된 핵산 단편의 길이는 1,000 bp 이상일 수 있다. 보다 바람직하게는 본 발명의 오류없는 바코딩된 핵산 단편들의 길이는 200 bp 내지 약 10kb 또는 그 이상일 수 있다.

[98]

[99] 본 발명의 또 다른 구현예에 따르면, 서로 다른 종류의 상기 바코드 서열들로 태깅된 핵산 단편들이 개별 스폿 단위별로 구분되어 부착되어 있으며, 상기 개별 스폿들이 이격하여 배열된 지지체가 장착된 스테이지; 상기 바코드 서열들로 태깅된 핵산 단편들 중에서 검증된 서열을 갖는 단편들의 분리를 위해 병렬적 시퀀싱 장치; 상기 병렬적 시퀀싱 장치에서 검증된 서열을 갖는 단편들에 부착된 바코드 서열정보를 지지체 상에서 관찰하기 위한 이미징 장치; 및 상기 검증된 서열을 갖는 단편들이 위치한 좌표에 상기 지지체로부터 분리하기 위해 접촉식 또는 비접촉식으로 에너지를 인가하는 추출 장치를 포함하는 고순도 핵산분자의 획득장치를 제공한다. 상기 추출 장치는 일례로 펄스 레이저 광원 및 집광 장치(condenser)를 포함할 수 있다. 상기 집광 장치는 광학 렌즈, 광원, 이미지 센서(image sensor)일 수 있으며, 이는 펄스 레이저 에너지의 집중을 목적으로 사용됨과 동시에 상기 핵산분자와 기판의 관찰에 사용될 수도 있기

때문이다. 상기 펄스 레이저는 원하는 시점에 이미 확인해놓은 xy 위치좌표를 이용하여 상기 기판 상의 분리를 원하는 핵산 분자의 위치에 정확하게 조사할 수 있으며, 이는 일반적으로 컴퓨터와 같은 제어장치로 펄스레이저를 제어함에 의해 자동화 수행될 수 있다. 상기 광학 렌즈는 2X 내지 100X, 바람직하게는 10X 내지 40X 배율의 광학 렌즈일 수 있다. 상기 배율은 상기 기판 상에 분리 가능한 적당한 에너지를 전달할 수 있음과 동시에 렌즈와 기판 상의 거리가 너무 가까워서 서로 접촉하거나 초점 거리를 벗어나는 일을 막아줄 수 있다. 상기 광원의 파장은 10nm 내지 10,000nm, 바람직하게는 50nm 내지 2,000nm의 파장을 가질 수 있으며, 해당 파장 영역에서 형광이나 가시광선을 이용한 기판의 관찰, 계측이 용이할 수 있다. 상기 광원으로는 예를 들어 할로겐 램프 등이 사용될 수 있다. 상기 이미지 센서는 일반적으로 CCD(charge-coupled device)가 사용되지만 이에 한정되는 것인 아니다.

[100] 일 실시예에 따르면, 상기 원하는 분자클론의 회수를 위한 저장 용기가 장착된 별도의 스테이지를 더 포함할 수 있다. 상기 스테이지는 1mm 이하, 바람직하게는 100㎛ 이하의 미세 범위에서 조작가능한 정밀도를 가질 수 있다. 상기 저장 용기가 장착된 별도의 스테이지는 상기 분리된 생화학 분자의 활용을 더욱 용이하게 할 수 있다.

[101] 예컨대, 상기 추출 장치는 크게 상부 시스템과 하부 시스템으로 구분된다. 상기 상부 시스템은 컴퓨터로 제어되며 기판이 아래쪽 방향으로 부착된 상부 스테이지(Motorized XY stage)와 에너지 인가 장치 및 상부 이미징 장치로 구성된다. 즉, 집광 장치를 통해 집광된 펄스 레이저 빔은 기판에 입사되어 펄스 레이저 삭마에 의한 팽창 압력, 혹은 복사압에 의해 기판 상의 핵산 분자를 추출하여 하부의 저장 용기 역할을 하는 PCR 플레이트로 밀어낸다. 하부 시스템은 저장 용기의 역할을 하는 PCR 튜브 랙 혹은 PCR 플레이트가 위쪽 방향으로 부착되어 있음, Z축으로 이동가능한 하부 스테이지(Motorized XYZ stage)와 하부 이미징 장치로 구성된다. 상기 하부 이미징 장치에는 상기 기판에서 분리된 분자클론을 수집할 경우 광학적으로 확인하도록 투과 이미징이 가능한 저장 용기가 사용될 수 있다. 상기 저장 용기의 바닥이 평면인 것이 상기 하부 이미징 장치의 광원으로부터 나오는 빛의 경로에 미치는 영향을 줄여 이미징을 용이하게 할 수 있다.

[102] 상기 장치는 일례로 상기 검증된 서열을 갖는 단편들의 분리를 위하여 상기 지지체의 특정 영역이 상기 추출 장치와 대응되도록 위치시키는 제어 장치를 더 포함할 수 있다. 상기 이미징 장치는 일례로 광학 렌즈, 파장이 10~10,000nm인 광원, 이미지 센서, 형광 분자 중 하나 이상을 포함할 수 있다. 구체적으로, 본 발명에 사용되는 장치는 이하 도면으로 표시한 장치와 전술한 장치들을 사용하여 다양하게 조합하여 사용할 수 있다.

[103]

[104] 본 발명의 또 다른 구현예에 따르면, 전술한 방법에 의하여 만들어지며, 음의

전하를 띤 단일가닥 DNA 분자들을 갖는 것인 비드 프리(bead free) 핵산 마이크로스피어가 제공된다. 상기 단일가닥 DNA 분자들은 이에 한정하는 것은 아니나, 50kbp 내지 100Gbp 정도로 긴 길이의 단일가닥 DNA 분자들일 수 있다.

[105] 상기 긴 길이의 단일가닥 핵산 분자들이 핵산분자간 서로 엉기거나 반복 서열의 단일가닥이 서로 이어진 입체물의 형태로 제공되며, 상기 입체물은 1nm ~ 30μm의 직경을 갖는 것일 수 있다. 참고로, 상기 입체물의 직경은 달리 언급되지 않는 한 해당 입체물에서 측정된 가장 긴 지름을 제공하는 부분의 직경을 지칭한다.

[106] 본 발명에 따른 비드 프리(bead free) 핵산 마이크로스피어는 통상적으로 비드에 부착하여 제공하던 것(도 17의 좌측 도면에 해당)과 달리 핵산분자를 지지체에 따로 고정할 필요가 없고 지지체에 바로 직접적인 고정이 가능하기 때문에(도 17의 우측 도면에 해당) 고정 단계 과정에 있어서 간소화가 가능하다. 또한, 비드패킹을 고려하지 않아도 되는 장점으로 지지체 기판 디자인 및 제작 단순화가 가능한 이점이 있다.

[107] 상기 고순도 비드 프리(bead free) 핵산 마이크로스피어는 단일 가닥의 핵산분자가 엉겨있는 형태를 이루고 있으며, 전체적으로 음전하를 띄고 있으므로 이후 지지체 기판 위 고정에서 전자기적 결합 및 흡착이 가능하다. 혹은, 반복적인 염기서열을 구성할 수 있으며, 이에 상보적인 염기서열을 가지는 프로브로 캡쳐가 가능하다.

발명의 실시를 위한 형태

[108] 이하, 본 발명의 실시예를 제시하나 예시적인 기재로서 이에 한정하는 것은 아니다.

[109] **실시예 1**

[110] 단계1: 병렬적 시퀀싱 기술로 서열확인 가능한 크기의 핵산 단편 제조

[111] 타겟 DNA 분자 준비에 있어서, 길이 147bp의 DNA 클론들을 마이크로어레이를 이용한 화학적 합성을 통해 준비하였다. 합성된 DNA 염기서열 일부는 아래와 같다.

[112] DNA template A(서열목록 1):

[113] TGCCTTGGCAGTCTCAaGATCAgctcttcGTTGAACGCACTCACAGCGAcgtctcA TGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAG CTGGACGGCGACgaagagcACACTGCGGCTCCTCATCCACG

[114] DNA template B(서열목록 2):

[115] TGCCTTGGCAGTCTCAaGATCAgctcttcGGCGAGGAGCTGTTCACCGGGGTG GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAG CGTGCGCGGCGAGGGgaagagcACACTGCGGCTCCTCATCCACG

[116] DNA template C(서열목록 3):

[117] TGCCTTGGCAGTCTCAaGATCAgctcttcGGCGAGGAGCTGTTCACCGGGGTG

GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAG CGTGCGCGGCGAGGGgaagagcACACTGCGGCTCCTCATCCACG

[118] DNA template D(서열목록 4):

[119] TGCCTTGGCAGTCTCAaGATCAgctcttcGGGGTGGTGCCCATCCTGGTCGAG CTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGCGCGGCGAGGGCGA GGGCGATGCCACCAAgaagagcACACTGCGGCTCCTCATCCACG

[120]

[121] 단계2: 핵산 단편들의 바코드 서열 태깅

[122] 단계1의 타겟 DNA분자에 각기 다른 바코드 염기서열을 가진 DNA를 추가하여 개별 분류가 가능하도록 하였다. 구체적으로는, 타겟 DNA분자 염기서열과 일부 겹치는 염기서열과 바코드 염기서열을 모두 포함하는 프라이머를 디자인하고, 본 프라이머로 PCR을 통해서 각 타겟 DNA분자에 서로 다른 바코드 염기서열을 가지도록 하였다. 필요에 따라서는, Ligation에 의해 바코드 DNA를 이어붙였다. 타겟 DNA 분자 염기서열과 일부 겹치는 염기서열 정보는 다음과 같다:

[123] a. 순방향 (sense: 서열목록 5):

[124] TGCCTTGGCAGTCTCA

[125] b. 역방향 (antisense: 서열목록 6):

[126] CGTGGATGAGGAGCCGCAGTGT

[127] 타겟 DNA 분자의 일부 염기서열과 바코드 염기서열을 모두 포함하는 정보는 다음과 같다.

[128] a. 순방향 (sense: 서열목록 7)

[129] AATGATACGGCGACCACCGAGATCTACACCTTCGCCTACACTCTTTCCCT ACACGACGCTCTTCCGATCTTGCGTCTATTTAGTGGAGCCNNNNCTATCTTC TTTTGCCTTGGCAGTCTCAAGAT

[130] b. 역방향 (antisense: 서열목록 8)

[131] CAAGCAGAAGACGGCATACGAGATATTCAGAAGTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTCGTGGATGAGGAGCCGCAGTGT

[132] 그런 다음 DNA 증폭방법을 통해 하나의 바코드 염기서열을 가지는 DNA분자가 최소 2개 이상이 되도록 함으로써 추후 같은 바코드를 가지는 분자클론이 각각 RCP(RCA product) 형성과 전체 염기서열 분석을 위한 준비과정으로 나누어질 수 있도록 준비하였다. 구체적으로는, 10cycle의 PCR 증폭방법을 통해 분자클론 갯수를 늘렸다.

[133]

[134] 단계3: 바코드 서열 태깅된 핵산 단편들 중 검증된 서열을 갖는 단편들에 대한 위치 좌표 확인

[135] 기판상에 DNA 분자클론 고정방법에 있어서, 직경 5㎛ 이하 크기의 RCP의 DNA 분자클론을 직접적으로 기판에 고정시켰다. 구체적으로는 하기 단계3-1과 단계3-2를 수행하여 기판 상에 RCP(RCA product)를 고정하였다:

[136]

[137] 단계3-1: 전체 염기서열 확인에 의한 검증된 염기서열을 갖는 DNA의 바코드 서열정보

[138] 상기 단계2에서 형성한 바코드 시퀀스가 추가된 DNA샘플을 증폭 후, 전체 염기서열을 분석하여 원하는 시퀀스를 선별하기 위하여 NGS를 이용하였으며, 해당 분석에서는 같은 바코드를 가지는 시퀀싱 리드(read)끼리의 패밀리 리드(family read)를 형성하여 무작위로 생기는 NGS의 시퀀싱 에러를 필터링함으로써 분석 오류율을 최소한으로 줄일 수 있도록 하였다. 구체적으로는, 바코드와 전체 시퀀스를 포함하는 DNA 시퀀싱을 진행하기 위하여 300bp paired end Illumina 시퀀싱을 이용하였으며, 한 바코드 당 100x의 이상의 심도(depth)로 90% 이상의 시퀀스 결과에서 같은 시퀀스를 가지는 결과만을 선별하여, 원하는 시퀀스를 0.001%/base이하의 오류율로 정확하게 얻을 수 있었다.

[139]

[140] 단계 3-2: 단계 3-1에서 검증된 염기서열을 갖는 DNA의 바코드 서열정보를 나타내는 바코드 서열을 지지체 상에서 확인

[141] 상기 단계 2에서 바코딩된 DNA 분자를 용액 상에서 RCA 반응을 통해 RCP(RCA product)를 생성한 뒤에 기판에 해당 RCP(RCA product)를 뿌려서 고정시켰다.

[142] A. RCA 반응 진행

[143] RCA 반응을 위해서 상기 단계 3-1에서 바코딩한 DNA분자 (sense) 스트랜드의 5'을 인산화 시켜주기 위하여 인산화된 프라이머(서열목록 9)를 이용하여 PCR 증폭을 통해 DNA 말단을 인산화 시켰다. 또한, 뒷단계에서 (sense)스트랜드의 ssDNA만 풍부하기 위하여 (antisense)스트랜드에는 비오틴분자를 추가함으로써(서열목록 10) 비오틴-아비딘결합을 통해 분리되도록 하고자 하였다.

[144] 구체적으로, DNA(sense) 스트랜드 5'에는 phosphorylation을, (antisense)스트랜드의 5'에는 비오틴 분자를 달아주기 위한 PCR을 진행하였다. 사용된 프라이머 디자인은 다음과 같다:

[145] a. 순방향(sense) 서열목록 9

[146] 5' - phos - AATGATACGGCGACCAC - 3'

[147] b. 역방향(antisense) 서열목록 10

[148] 5' - 비오틴- CAAGCAGAAGACGGCATA - 3'

[149] 상기 프라이머를 이용하여 27 cycle PCR 후 증폭된 DNA 젤 사진을 도 3에 나타내었다. 도 3에서 보듯이, 왼쪽의 DNA 래더에 비교하여 보았을 때, 약 300bp 근처에 DNA 증폭물이 보이는 것을 확인 할 수 있었으며, 인산화와 비오틴이 각각 추가된 프라이머 사용으로 인해 증폭물에도 인산화와 비오틴이 추가 되었음을 추측해 볼 수 있다.

[150] 그런 다음 인산화된 (sense)스트랜드만 풍부하기 위하여 dsDNA가 ssDNA가 되도록 변성 후, 상기 단계에서 추가한 (antisense) 스트랜드의 비오틴 분자를 이용하여 아비딘이 코팅된 자성 비드와의 결합 및 수세를 통해 (antisense) 스트랜드는 제거하였다. 또한 풍부한 (sense)스트랜드의 ssDNA를 padlock 프로브와 혼성화한 다음 라이게이션 시켜 원형화하였고(circularization), 이어 Phi29 중합효소를 이용하여 RCA 반응을 진행하였다. 그 결과, RCP(RCA product)와 형광 분자가 달린 프로브간 결합을 통해 RCP(RCA product)를 형성하는 것을 확인하였다(도 4 참조).

[151]

[152] B. 기판 설계: RCA(Rolling Cycle Amplification) product 고정 및 레이저 분리를 위한 칩 설계

[153] 저비용, 초고속 생산을 위해서는 칩 하나 당 약 10만개의 오류없는 DNA 생산이 필요하며, 이를 위하여 한 칩에서 약 100만개의 서로 다른 종류의 DNA 고정이 필요하였다. 이를 구현하도록 마이크로 크기의 필라 구조물을 디자인하여 그 위에 병렬적으로 서로 다른 종류의 DNA를 고정시켰다.

[154] 본 실험은 오류 없는 DNA를 회수하기 위해 레이저를 이용하는데, 레이저 분리를 통해 원하는 부분만 분리하기 위해 ITO가 코팅된 글래스를 사용하였다(도 5 참조). 구체적으로 100만개의 다른 종류 DNA 클러스터를 RCP(RCA product)(RCP)형태로 만들어 고정시킨 다음 레이저로 분리하기 위한 목적으로 필라 어레이를 디자인하였다. 도 5에서 보듯이, ITO 글래스 위에 일정한 패턴으로 SU-8 필라 구조를 제작하였다. 레이저 영역의 크기와 각 필라에 1개의 RCP가 형성되어야 하는 것을 고려하여 지름 10㎛, 높이 5㎛, 간격 10㎛로 디자인하였다.

[155] 또한 해당 필라에 RCP(RCA product)를 고정하기 위해, 뉴트라비딘을 필라에 코팅하고 비오틴이 달린 프로브를 로딩한 다음, 상기 프로브에 합성 유전자가 원형으로 결합하여 RCA 반응을 진행하였다(도 6a 참조)

[156]

[157] C. 바코드 영역 시퀀싱

[158] 상기 기판 위에 고정된 RCP로부터 SBS(sequencing by synthesis) 또는 SBL(sequencing by ligation)을 통해 바코드 시퀀스를 읽을 수 있었다. 구체적으로는, RCP의 반복되는 서열에 염기서열을 읽고싶은 시작점이 되는 위치에 앵커 프라이머(anchor primer)를 붙이고, 바코드 염기서열을 읽기 위한 시퀀싱 프로브(시퀀싱 probe)를 상기 앵커 프라이머와 라이게이션하였다. 네 종류의 시퀀싱 프로브에는 각각 Cy3(Guanine), Cy5(Adenine), TAM(Thymine), Texas Red(Cytosine) 형광 물질이 달려있도록 하여, 형광 현미경을 통해 각 시퀀스 구분이 가능하도록 구현하였다. 상기 기판 위에서 바코드 시퀀스를 읽기 위하여 SBL 시퀀싱 방법을 수행한 결과를 나타내는 도 7에서 보듯이,'GCA'의 바코드를 가지는 것을 확인하였다. 참고로, 도 7에서 Cy7 형광은 RCP 유무 확인 용도로,

Cy3: 구아닌(Guanine), Cy5: 아데닌(Adenine, TAM: Thymine, Texas Red: Cytosine 확인용이다.

[159]

[160] 단계4: 검증된 서열을 갖는 단편의 위치 좌표를 활용한 회수

[161] 단계 3-1과 단계 3-2의 매칭을 통한 검증된 염기서열을 갖는 DNA의 바코드 서열 정보를 나타내는 위치 좌표를 확인하였다. 구체적으로는, Illumina 분석결과에서 선별된 시퀀스와 매치되는 바코드를 가지는 RCP를 레이저로 분리하여 Sanger 시퀀싱으로 해당 바코드의 염기서열을 최종 확인하였다. 실제로 도 8을 통해 레이저로 RCP 분리 후 PCR 증폭을 통한 확인을 수행하였다. 참고로, #1-13: RCP, PCR product(1/10^6): 양성 대조군, NTC: 음성 대조군이었다.

[162]

[163] 단계 5: 폴로니 시퀀싱을 통한 바코드 염기서열 확인 및 검증

[164] RCP가 앞서 도 4, 도 5를 통해 디자인된 칩 상에 고정된 상태에서 폴로니 시퀀싱(Polony sequencing)을 통해 정확히 바코드 시퀀스를 읽을 수 있는지 확인하였으며 최종적으로 총 10base의 바코드 시퀀스를 읽을 수 있음을 확인하였다. 상기 폴로니 시퀀싱은 도 9와 같이 SBL(sequencing by ligation) 방식을 이용하여 구현한 것으로, RCP 반복되는 서열에 앵커 프라이머(anchor primer)를 붙이고, 바코드 염기서열을 읽기 위한 시퀀싱 프로브를 앵커 프라이머와 라이게이션하였다. 네 종류의 시퀀싱 프로브에는 각각 Cy3(Guanine), Cy5(Adenine), TAM(Thymine), Texas Red(Cytosine) 형광 물질이 달려있다.

[165] 양성 대조군 실험을 위해 생어 시퀀싱으로 먼저 바코드의 염기서열을 파악한 샘플을 가지고 폴로니 시퀀싱을 진행하였다. 도 10은 리버스 프라이머로 수행한 생어 시퀀싱 결과이며, 바코드의 염기서열이 'GCAT'임을 확인하였다.

[166] 또한 검증된 염기서열을 갖는 DNA의 바코드 서열 정보를 나타내는 위치 좌표를 확인한 다음 적절한 펄스 레이저를 인가하여 회수하였다(도 2의 최우측 단계 참조). 분리된 RCP의 바코드 시퀀스 생어 시퀀싱 결과를 확인한 결과, 도 12를 통해 정확히 일치한 결과를 확인하였다. 참고로, 도 12에서 square box는 바코드 시퀀스 영역을 나타낸다.

[167] 도 10은 폴로니 시퀀싱을 통해 바코드의 염기서열을 읽은 결과로서, 도면에서 보듯이, 각 필라의 표면에 많은 수의 RCP가 붙어있으며 형광 신호가 순서대로 Cy3, Texas Red, Cy5, FAM으로 나타났다. 즉, 바코드의 염기서열을 'GCAT'로 정확하게 읽어냈음을 확인할 수 있었다. 이후에 임의의 바코드 염기서열을 가지는 샘플을 가지고 폴로니 시퀀싱으로 읽은 후 레이저로 분리하여 생어 시퀀싱을 통해 바코드 염기서열을 재확인하는 실험을 진행하였다(도 11 참조). 이번에는 먼지와 실제 RCP를 구분하기 위해 앞서 사용한 앵커 프라이머에 Cy7 형광물질을 달아서 구분할 수 있도록 하였다.

[168] 그 결과, 도 13과 같이 1개의 RCP가 결합한 SU-8 필라를 폴로니 시퀀싱으로

읽은 결과, 각각 'CAG', 'GCA'로 확인되었다. 이들 필라를 레이저로 분리하여 생어 시퀀싱으로 바코드의 염기서열을 확인한 결과, 폴로니 시퀀싱으로 읽은 결과와 마찬가지로 'CAG', 'GCA'임을 확인할 수 있었다(도 12 참조).

[169]

[170] 3. 오류 없는 합성 DNA의 회수 및 오류율 계산

[171] - 오류율 계산을 위한 실험 과정

[172] a) Illumina 시퀀싱을 통해 오류가 없는 합성 DNA를 파악하고 이들의 바코드 염기서열을 추출하였다. b) 칩 상에서 합성 DNA로 형성된 RCP를 고정하고 인시츄(*in situ*) 시퀀싱으로 바코드 염기서열을 읽었다. c) 앞서 추출한 바코드 염기서열과 매칭되는 RCP를 확인하고 레이저를 통해 합성 DNA를 회수하였다. d) 회수된 DNA의 시퀀스를 확인하여 오류염기의 개수를 파악하였다.

[173] 바코드의 길이가 총 10 bases(5 bases × 2)인 합성 DNA를 가지고 실험하였으며, 합성 DNA의 길이는 (바코드 영역을 제외하고) 263 bases이었다. Illumina 시퀀싱에서 오류없는(error-free) 것으로 검증된 샘플 중 바코드가 매칭되는 1,504개의 RCP를 추출하였으며, 이 샘플들의 염기서열을 재확인해본 결과 2개의 RCP에서 오류가 있음을 확인하였다. 본 데이터를 통해 오류율을 계산해보면 약 0.0005%/base($\frac{1}{1,504} \times \frac{2}{263(base)} \times 100(\%)$)으로, 목표했던 0.001%/base 수준을 만족함을 확인하였다.

[174]

[175] 4. 오류 없는 합성 DNA의 회수 처리량, 처리시간 정보 및 제조단가 계산

[176] 처리량 및 제조단가 정보:

[177] NGS(Illumina)를 통한 전체 염기서열 분석과, 한 칩에 약 10만개의 RCP 형성이 가능함에 따라 하나의 원하는 DNA 분자클론을 얻기 위한 가격은 약 $0.05 수준으로 계산된다.

[178]

[179] 5. 오류 없는 합성 DNA의 직경 및 지지체 고정 형태:

[180] 수득된 오류없는 합성 DNA를 투과전자현미경으로 찍어 측정한 결과, 해당 입체물에서 측정된 가장 긴 지름을 갖는 부분의 직경을 확인한 결과, 1 ㎛ 내지 300 ㎛ 범위 내로서 평균적으로 대략 3 ㎛인 것을 확인하였다. 해당 입체물이 차지하는 부피는 약 30 μm^3 이고 해당 합성에 사용된 바코드 서열 길이가 총 20bp 이고, DNA template 길이는 각각 50bp~ 5kbp 인 점을 고려할 때 적어도 50kbp 내지 100Gbp 정도의 긴 길이를 갖는 DNA 단일분자가 롤링된 형태로 음의 전하를 띠고 있는 것으로 확인되었다. 이와 같은 음의 전하로 긴 길이를 갖는 DNA 단일분자가 롤링된 형태로 제공됨에 따라 핵산분자를 지지체에 따로 고정할 필요가 없고 지지체에 바로 직접적인 고정이 가능하기 때문에 고정 단계 과정에 있어서 간소화가 가능하다. 해당 고정과정을 모식화한 도 17에 따르면, 비드패킹 등 비드 부착을 필요로 하지 않고(좌측 모식도 참조), 정전기적 인력

등에 의해 지지체의 표면에 직접 고정되는 것을(우측 모식도 참조) 확인할 수 있다.

[181]

[182] **실시예 2**

[183] 상기 실시예 1의 단계 B. 기판 설계: RCA(Rolling Cycle Amplification) product 고정 및 레이저 분리를 위한 칩 설계를 아래의 단계 B'로 대체한 것을 제외하고는 상기 실시예 1과 동일한 공정을 반복하였다. 구체적으로는 기판 상에 DNA 캡쳐를 통한 고정을 수행한 다음 RCA 반응을 진행하였다.

[184] 단계 B': 기판 설계: RCA(Rolling Cycle Amplification) product 고정 및 레이저 분리를 위한 칩 설계:

[185] 기판 상에 3-아미노프로필트리에톡시실란(APTES)를 코팅하고 표면에 아민기가 드러나도록 하고, 아민기가 양전하를 띠게 되어 음전하를 띠는 DNA와 정전기적 결합을 하게 하였다. 이 때, 상기 실시예 1에서 사용한 비오틴-아비딘 방법과는 달리 미리 용액 상에서 RCA를 진행하여 형성된 RCP를 아민기가 드러난 칩에 로딩하였다(도 6b 및 도 13 참조). 결과 기판 상에 고정된 RCP(RCA product)를 도 14를 통해 확인하였다.

[186]

[187] 3. 오류 없는 합성 DNA의 회수 및 오류율 계산

[188] 그런 다음 상기 실시예 1과 동일한 방법으로 오류율을 계산하였다. 그 결과, 419개의 RCP를 회수하였을 때에 1개의 RCP에서 1개의 base에 오류가 있음을 확인하였고, 본 결과를 통해 오류율을 계산해보면 약 0.0009%/base($\frac{1}{419} \times \frac{1}{263(base)} \times 100(\%)$)이었다.

[189]

[190] 4. 오류 없는 합성 DNA의 회수 처리량, 처리시간 계산

[191] 본 실험을 통해 한 기판에 100,000개의 RCP를 고정할 수 있는 가능성을 확인하였으며, 최종적으로는 염기서열이 검증된 10,000개의 RCP를 회수 가능함을 확인하였다. 10,000개의 검증된 염기서열을 가지는 DNA분자 회수 처리시간은, DNA 분자 화학적 합성에서 2일, Illumina 시퀀싱과 기판 위 RCP(RCA product)의 위치정보 획득(*in situ* 시퀀싱)의 병렬적 과정에서 3일, 최종 RCP(RCA product) 분자 획득에서 2일로, 총 1주일이 소요되었다.

[192]

[193] **비교예**

[194] 비드와 이멀젼 PCR를 사용하여 원하는 염기서열의 DNA를 회수한 방법에 따른 실험 결과:

[195] 150bp 길이의 DNA를 화학적으로 합성한 후 30 ㎛의 직경을 갖는 비드에 DNA를 캡쳐하여 이멀젼 PCR을 진행한 뒤에 해당 비드에서 염기서열을 확인하였다. 염기서열이 확인된 총 100개의 DNA가 캡쳐된 비드를 회수하여

오류율을 계산하여 본 결과 25012개의 염기 중 18개의 염기에 오류가 있음을 염기서열 분석을 통해 확인하였다. 간접적으로는, DNA 분자 상에서 오류를 확인하면 절단을 일으키는 효소와의 반응 후에 전기영동으로 확인한 결과, 젤 사진에서 원하는 위치의 밴드 형성 이외에 형성된 끌리는 듯한 밴드 이미지 관찰을 통해 DNA 가닥 상에서 절단이 무작위로 일어났음을 확인할 수 있었다(도 15 및 도 16 참조). 도 15는 DNA가 캡쳐된 비드의 회수 후, 염기서열 확인을 위해 DNA를 증폭한 뒤 전기영동한 젤 사진이고, 도 16은 DNA분자 상에서 오류를 확인하면 절단을 일으키는 효소반응 후 전기영동한 젤 사진이다.

[196] 이와 같이 비드를 사용하고, 100bp 길이의 DNA, 총 10,000개의 염기서열이 검증된 DNA 회수를 가정할 때 DNA 회수 처리량은 7일이 소요될 뿐 아니라, 비용 $ 1,000과 오류율이 한 염기서열 당 0.072%를 보였다.

산업상 이용가능성

[197] 이상 살펴본 바와 같이, 본 기술에 따르면 상업적으로 입수가능한 지지체와 기존의 합성 방법을 이용하되 1개의 유전자를 구입할 비용으로 500개의 유전자를 합성할 정도로 처리량(throughput)과 오류율(error rate)을 현저하게 개선시킬 뿐 아니라 제조 공정의 단축으로 생산 시간 또한 10배 정도 단축시킨 고 효율 DNA 정제 기술을 제공할 수 있다.

[198] 이러한 본 기술 특유의 효과는 10배 이하 작은 DNA 클론 분자를 비드 없이 직접적으로 기판에 고정하기 때문에 동일한 사이즈의 기판이라 할지라도 10배 이상 많은 DNA 클론 분자를 고정할 수 있고, 이러한 공간적 고정 개수 효율 개선을 통해 한 칩당 회수가능한 DNA 개수를 늘릴 수 있게 하며, 이를 통해 최종적으로는 하나의 DNA 회수 당 제조단가를 약 10배 이상 낮출 수 있고, 제조 공정의 단축으로 생산 시간 또한 개선할 수 있으며, 특히 기존 비드를 사용한 이멀젼 PCR 방법에서 벗어나 증폭과정에서 생기는 오류율이 낮은 phi29의 DNA 중합효소를 사용함으로써 오류율도 약 70배 낮출 수 있었다.

[199] 이러한 본 기술 특유의 장점은 실시예와 비교예를 대비해보면 명백히 확인되는 것으로, 실제로 동일한 DNA 회수 처리량(7일) 기준으로, 비용은 $ 100로 현저히 저감될 뿐 아니라 오류율이 하나의 염기서열 당 0.001%로 현저히 저감된 결과를 입증하였다.

서열목록 Free Text

[200] <110> Seoul National University R&DB Foundation

[201]

[202] <120> Method and apparatus for obtaining high purity nucleotides

[203]

[204] <130> SPC2018-1004

[205]

[206] <150> KR 10/2017/0059558

[207] <151> 2017-05-12
[208]
[209] <160> 10
[210]
[211] <170> KoPatentIn 3.0
[212]
[213] <210> 1
[214] <211> 147
[215] <212> DNA
[216] <213> Artificial Sequence
[217]
[218] <220>
[219] <223> DNA Template A
[220]
[221]
[222] <400> 1
[223] tgccttggca gtctcaagat cagctcttcg ttgaacgcac tcacagcgac gtctcatgag 60
[224]
[225] caagggcgag gagctgttca ccggggtggt gcccatcctg gtcgagctgg acggcgacga 120
[226]
[227] agagcacact gcggctcctc atccacg 147
[228]
[229]
[230] <210> 2
[231] <211> 147
[232] <212> DNA
[233] <213> Artificial Sequence
[234]
[235] <220>
[236] <223> DNA Template B
[237]
[238]
[239] <400> 2
[240] tgccttggca gtctcaagat cagctcttcg gcgaggagct gttcaccggg gtggtgccca 60
[241]
[242] tcctggtcga gctggacggc gacgtaaacg gccacaagtt cagcgtgcgc ggcgagggga 120
[243]
[244] agagcacact gcggctcctc atccacg 147

[245]
[246]
[247] <210> 3
[248] <211> 147
[249] <212> DNA
[250] <213> Artificial Sequence
[251]
[252] <220>
[253] <223> DNA Template C
[254]
[255]
[256] <400> 3
[257] tgccttggca gtctcaagat cagctcttcg gcgaggagct gttcaccggg gtggtgccca 60
[258]
[259] tcctggtcga gctggacggc gacgtaaacg gccacaagtt cagcgtgcgc ggcgagggga 120
[260]
[261] agagcacact gcggctcctc atccacg 147
[262]
[263]
[264] <210> 4
[265] <211> 147
[266] <212> DNA
[267] <213> Artificial Sequence
[268]
[269] <220>
[270] <223> DNA Template D
[271]
[272]
[273] <400> 4
[274] tgccttggca gtctcaagat cagctcttcg gggtggtgcc catcctggtc gagctggacg 60
[275]
[276] gcgacgtaaa cggccacaag ttcagcgtgc gcggcgaggg cgagggcgat gccaccaaga 120
[277]
[278] agagcacact gcggctcctc atccacg 147
[279]
[280]
[281] <210> 5
[282] <211> 16

[283] <212> DNA
[284] <213> Artificial Sequence
[285]
[286] <220>
[287] <223> Sequence information partially overlapping the target DNA
[288] molecule base sequence-sense
[289]
[290]
[291] <400> 5
[292] tgccttggca gtctca 16
[293]
[294]
[295] <210> 6
[296] <211> 22
[297] <212> DNA
[298] <213> Artificial Sequence
[299]
[300] <220>
[301] <223> Sequence information partially overlapping the target DNA
[302] molecule base sequence-antisense
[303]
[304]
[305] <400> 6
[306] cgtggatgag gagccgcagt gt 22
[307]
[308]
[309] <210> 7
[310] <211> 125
[311] <212> DNA
[312] <213> Artificial Sequence
[313]
[314] <220>
[315] <223> Information including both a partial nucleotide sequence of a
[316] target DNA molecule and a barcode nucleotide sequence-sense
[317]
[318]
[319] <400> 7
[320] aatgatacgg cgaccaccga gatctacacc ttcgcctaca ctctttccct acacgacgct 60

[321]
[322] cttccgatct tgcgtctatt tagtggagcc nnnnctatct tcttttgcct tggcagtctc 120
[323]
[324] aagat 125
[325]
[326]
[327] <210> 8
[328] <211> 88
[329] <212> DNA
[330] <213> Artificial Sequence
[331]
[332] <220>
[333] <223> Information including both a partial nucleotide sequence of a
[334] target DNA molecule and a barcode nucleotide sequence-antisense
[335]
[336]
[337] <400> 8
[338] caagcagaag acggcatacg agatattcag aagtgactgg agttcagacg tgtgctcttc 60
[339]
[340] cgatctcgtg gatgaggagc cgcagtgt 88
[341]
[342]
[343] <210> 9
[344] <211> 17
[345] <212> DNA
[346] <213> Artificial Sequence
[347]
[348] <220>
[349] <223> barcode primer-sense
[350]
[351]
[352] <400> 9
[353] aatgatacgg cgaccac 17
[354]
[355]
[356] <210> 10
[357] <211> 18
[358] <212> DNA

[359] <213> Artificial Sequence

[360]

[361] <220>

[362] <223> barcode primer-antisense

[363]

[364]

[365] <400> 10

[366] caagcagaag acggcata 18

청구범위

[청구항 1] (a) 핵산 단편들을 준비하는 단계;
(b) 상기 핵산 단편들을 바코드 서열들로 태깅하는 단계;
(c) 상기 바코드 서열들로 태깅된 핵산 단편들 중에서 검증된 서열을 갖는 단편들에 대한 위치 좌표를 확인하는 단계; 및
(d) 상기 검증된 서열을 갖는 단편들을 상기 위치 좌표를 이용하여 회수하는 단계를 포함하는 고순도 핵산분자의 획득방법.

[청구항 2] 제1 항에 있어서,
상기 (c) 단계에서 상기 바코드 서열들로 태깅된 단편들은 PCR, RCA(Rolling cicle amplification), LAMP(Loop-mediated isothermal amplification), SR(Sequence replication), SDA(strand displacement amplification), LCR(Ligase chain reaction), NASBA(Nucleic acid sequence based amplification), MDA(Multiple displacement amplification) 또는 그 조합을 포함하는 방법들로 증폭된 것인 고순도 핵산분자의 획득방법.

[청구항 3] 제1 항에 있어서,
상기 (c) 단계는, (c-1) 상기 (b) 단계에서 바코드 서열들로 태깅된 단편들의 전체 혹은 일부 염기서열을 병렬적 시퀀싱 방법에 의해 확인하여 검증된 서열을 갖는 단편들의 바코드 서열을 확인하는 병렬적 시퀀싱 단계; 및 (c-2) 상기 검증된 서열을 갖는 단편들로서 확인된 바코드 서열을 갖는 단편들이 고정된 지지체를 준비하고, 해당 지지체 상에서 해당 바코드 서열을 나타내는 위치 좌표를 확인하는 시퀀싱 단계가 동시 혹은 순차적으로 수행되는 것인 고순도 핵산분자의 획득방법.

[청구항 4] 제3 항에 있어서,
상기 (c-2) 단계는, i) 상기 (b) 단계에서 바코드 서열들로 태깅된 단편들을 지지체 상에 고정하는 단계; 및 ii) 상기 지지체 상에 고정된 태깅된 단편들의 바코드 서열을 확인하여 상기 (c-1)의 병렬적 시퀀싱 단계에서 검증된 서열을 갖는 단편들로 확인된 바코드 서열에 대한 위치 좌표를 확인하는 단계를 포함하는 것인 고순도 핵산분자의 획득방법.

[청구항 5] 제4 항에 있어서,
상기 i) 단계는, 상기 바코드 서열들로 태깅된 단편들을 RCA(Rolling cicle amplification), LAMP(Loop-mediated isothermal amplification), SR(Sequence replication), SDA(strand displacement amplification), LCR(Ligase chain reaction), NASBA(Nucleic acid sequence based amplification) 또는 MDA(Multiple displacement amplification)로 증폭된 다음 지지체 표면과의 상호작용에 의해 지지체 상에 고정된 것인 고순도 핵산분자의 획득방법.

[청구항 6] 제4 항에 있어서,
상기 i) 단계는, 상기 바코드 서열들로 태깅된 단편들을 고정하는 지지체

표면층은 평평한 구조, 우물 구조, 기둥 구조, 반복 무늬 혹은 격자 형상이 패터닝된 구조, 또는 그 조합을 포함하는 구조인 고순도 핵산분자의 획득방법.

[청구항 7] 제4 항에 있어서,
상기 i) 단계는, 상기 바코드 서열들로 태깅된 단편들을 고정하는 지지체 표면은 폴리머, 실리카, 하이드로겔, 유리 또는 그 조합을 포함하는 물질인 고순도 핵산분자의 획득방법.

[청구항 8] 제4 항에 있어서,
상기 i) 단계는, 상기 지지체 표면과의 상호작용에 의해 지지체 상에 고정은, 지지체 상에 전자기 인력, 흡착, 단백질 결합, 자기 조립, 화학작용기의 결합, 또는 그 조합을 포함하는 상호작용에 의해 고정되는 것인 고순도 핵산분자 획득 방법.

[청구항 9] 제4 항에 있어서,
상기 i) 단계는, 상기 바코드 서열들로 태깅된 단편들을 지지체 상에 고정은 기판의 표면 혹은 내부, 셀의 표면 혹은 내부, 기판 상에 패터닝된 표면 혹은 내부 또는 그 조합을 포함하는 위치에 고정되는 것인 고순도 핵산분자 획득 방법.

[청구항 10] 제4 항에 있어서,
상기 i) 단계는, 상기 바코드 서열들로 태깅된 단편들의 목적 핵산 분자와 상보적 시퀀스를 갖는 프로브를 지지체에 고정한 다음 상기 프로브에 상기 바코드 서열들로 태깅된 단편들을 캡쳐하고 RCA(Rolling cicle amplification), LAMP(Loop-mediated isothermal amplification), SR(Sequence replication), SDA(strand displacement amplification), LCR(Ligase chain reaction), NASBA(Nucleic acid sequence based amplification) 또는 MDA(Multiple displacement amplification)로 증폭시켜 수행된 것인 고순도 핵산분자의 획득방법.

[청구항 11] 제3 항에 있어서,
상기 (c-2) 해당 지지체 상에서 해당 바코드 서열을 나타내는 위치 좌표를 확인하는 시퀀싱 단계는 SBS(sequencing by synthesis), SBL(sequencing by ligation), ELISA 방식, 또는 그 조합을 포함하는 방법으로 수행되는 것인 고순도 핵산분자의 획득방법.

[청구항 12] 제5 항 또는 제10 항에 있어서,
상기 RCA 또는 MDA는 파이(Phi) 29 DNA 폴리머라아제를 사용하여 수행되는 것인 고순도 핵산 분자의 획득방법.

[청구항 13] 제5 항 또는 제7 항에 있어서,
상기 지지체 상에 고정된 서열들로 태깅된 단편들은 지름 30㎛ 이하의 분자 클론 형태로 형성된 것인 고순도 핵산 분자의 획득방법.

[청구항 14] 제1 항에 있어서,

상기 (c) 단계에서 위치 좌표 확인은 상기 바코드에 형광 분자를 붙인 후 직접 형광을 관찰하는 방식으로 수행하는 것인 고순도 핵산분자의 획득방법.

[청구항 15] 제1 항에 있어서,
상기 (a) 단계의 핵산 단편들의 크기는 20 내지 3,000 bp인 고순도 핵산분자의 획득방법.

[청구항 16] 제3 항에 있어서,
상기 병렬적 시퀀싱 방법은 초병렬 시퀀싱, 생어(Sanger) 시퀀싱, 질량 분석, 전기영동, 혼성화, 디지털 PCR, 대립유전자-특이 PCR, 정량적 PCR, 형광 기반 분류, 또는 그 조합을 포함하는 것인 방법 중에서 선택된 것인 고순도 핵산분자의 획득방법.

[청구항 17] 제16 항에 있어서,
상기 초병렬 시퀀싱은 합성에 의한 시퀀싱, 이온 토렌트 시퀀싱, 파이로시퀀싱, 라이게이션에 의한 시퀀싱, 나노포어 시퀀싱 및 단일-분자 실시간 시퀀싱 또는 그 조합을 포함하는 방법으로부터 선택되는 것인 고순도 핵산분자의 획득방법.

[청구항 18] 제1 항에 있어서, 상기 바코드 서열들은 임의 또는 계획적으로 설계된 2종 이상의 올리고뉴클레오타이드들의 혼합물인 고순도 핵산분자의 획득방법.

[청구항 19] 제1 항에 있어서,
상기 바코드 서열의 길이는 5 내지 300 bp인 고순도 핵산분자의 획득방법.

[청구항 20] 제1 항에 있어서,
상기 바코드 서열들의 태깅은 PCR, 에멀젼 PCR, 라이게이션(ligation) 또는 그 조합을 포함하는 방법 중에서 선택되는 어느 하나의 방법에 의해 수행되는 것인 고순도 핵산분자의 획득방법.

[청구항 21] 제1 항에 있어서,
상기 (b) 단계는 상기 바코드 서열들에 상응하는 프라이머들로 목적 단편들을 선택적으로 증폭하여 회수하는 방식으로 수행되는 것인 고순도 핵산분자의 획득방법.

[청구항 22] 제1 항에 있어서,
상기 (d) 단계에서 상기 위치 좌표를 이용하여 회수하는 방식은 초음파, 공기압, 레이저 또는 그 조합을 포함하는 방법에 의해 접촉식 또는 비접촉식으로 에너지를 인가하는 방식인 것인 고순도 핵산분자의 획득방법.

[청구항 23] 제1 항에 있어서,
상기 (d) 단계에서 회수된 검증된 서열들을 갖는 단편들로부터 원하는 염기서열을 가진 핵산 분자를 획득하도록 특정 염기서열의 증폭 또는 절단 단계를 더 포함하는 것인 고순도 핵산분자의 획득방법.

[청구항 24] 서로 다른 종류의 상기 바코드 서열들로 태깅된 핵산 단편들이 개별 스폿 단위별로 구분되어 부착되어 있으며, 상기 개별 스폿들이 이격하여 배열된 지지체가 장착된 스테이지; 상기 바코드 서열들로 태깅된 핵산 단편들 중에서 검증된 서열을 갖는 단편들의 분리를 위해 병렬적 시퀀싱 장치; 상기 병렬적 시퀀싱 장치에서 검증된 서열을 갖는 단편들에 부착된 바코드 서열정보를 지지체 상에서 관찰하기 위한 이미징 장치; 및 상기 검증된 서열을 갖는 단편들이 위치한 좌표에 상기 지지체로부터 분리하기 위해 접촉식 또는 비접촉식으로 에너지를 인가하는 추출 장치를 포함하는 고순도 핵산분자의 획득장치.

[청구항 25] 제24 항에 있어서,
상기 장치는 상기 검증된 서열을 갖는 단편들의 분리를 위하여 상기 지지체의 특정 영역이 상기 추출 장치와 대응되도록 위치시키는 제어 장치를 더 포함하는 것인 고순도 핵산분자의 획득장치.

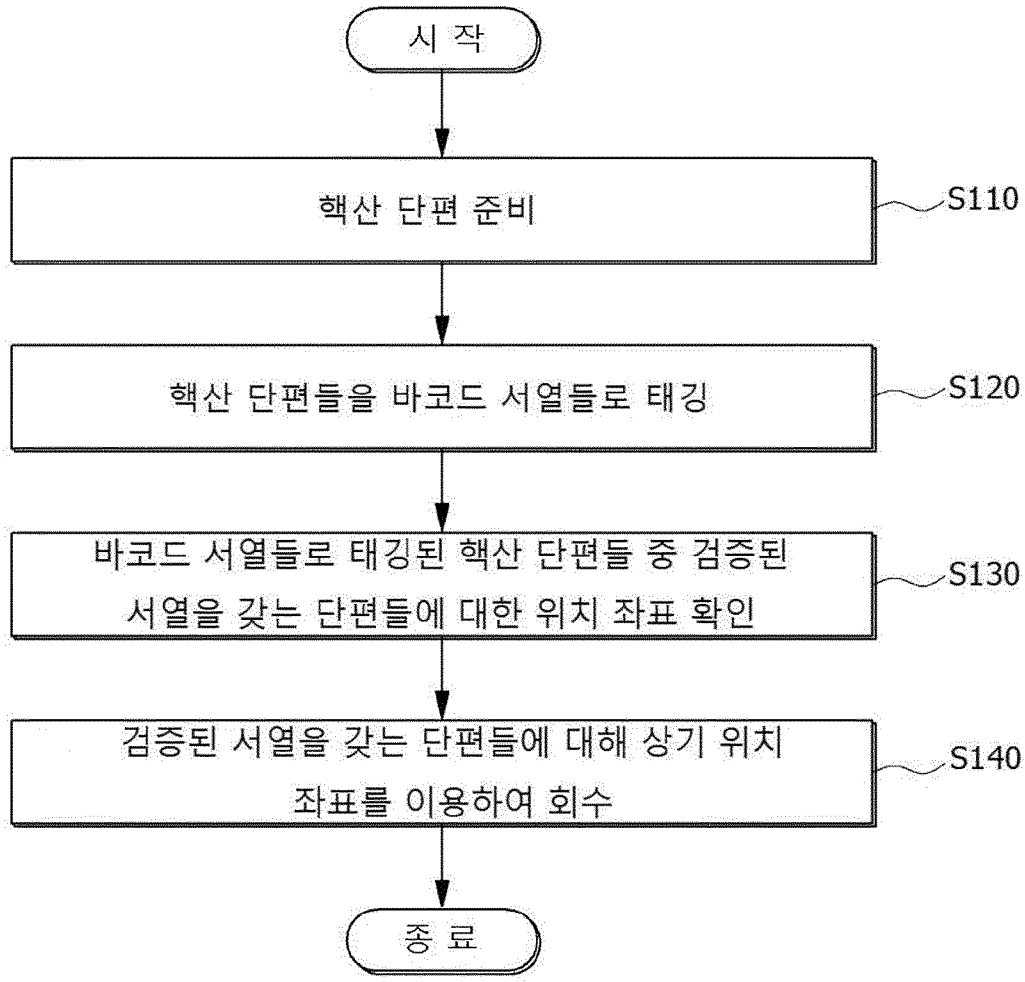
[청구항 26] 제24 항에 있어서,
상기 이미징 장치는 광학 렌즈, 파장이 10~10,000nm인 광원, 이미지 센서, 형광 분자 중 하나 이상을 포함하는 고순도 핵산분자의 획득장치.

[청구항 27] 제1 항 내지 제23 항 중 어느 한 항의 방법에 의하여 만들어지며, 음의 전하를 띤 단일가닥 DNA 분자들을 갖는 것인 비드 프리(bead free) 핵산 마이크로스피어.

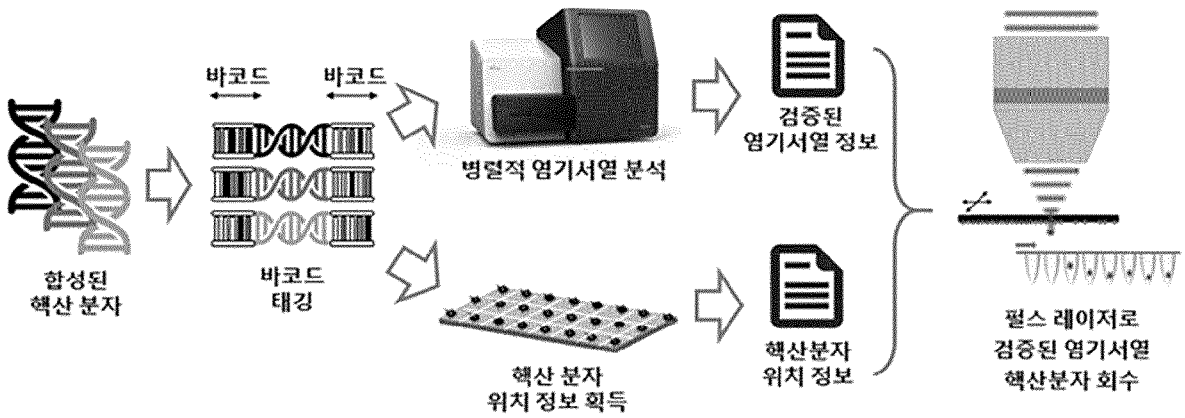
[청구항 28] 제27 항에 있어서,
상기 단일가닥 DNA 분자들은 50kbp 내지 100Gbp 정도로 긴 길이의 단일가닥 DNA 분자들인 비드 프리 핵산 마이크로스피어.

[청구항 29] 제28 항에 있어서,
상기 긴 길이의 단일가닥 핵산 분자들이 핵산분자간 서로 엉기거나 반복 서열의 단일가닥이 서로 이어진 입체물의 형태로 제공되며, 상기 입체물은 1nm ~ 30μm의 직경을 갖는 것인 비드 프리 핵산 마이크로스피어.

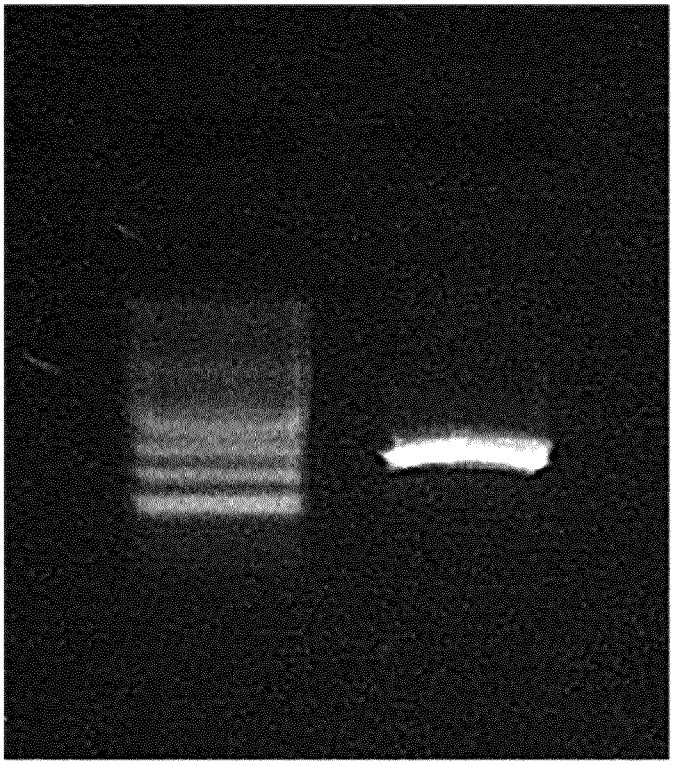
[도1]


시 작
핵산 단편 준비
S110
핵산 단편들을 바코드 서열들로 태깅
S120
바코드 서열들로 태깅된 핵산 단편들 중 검증된 서열을 갖는 단편들에 대한 위치 좌표 확인
S130
검증된 서열을 갖는 단편들에 대해 상기 위치 좌표를 이용하여 회수
S140
종 료

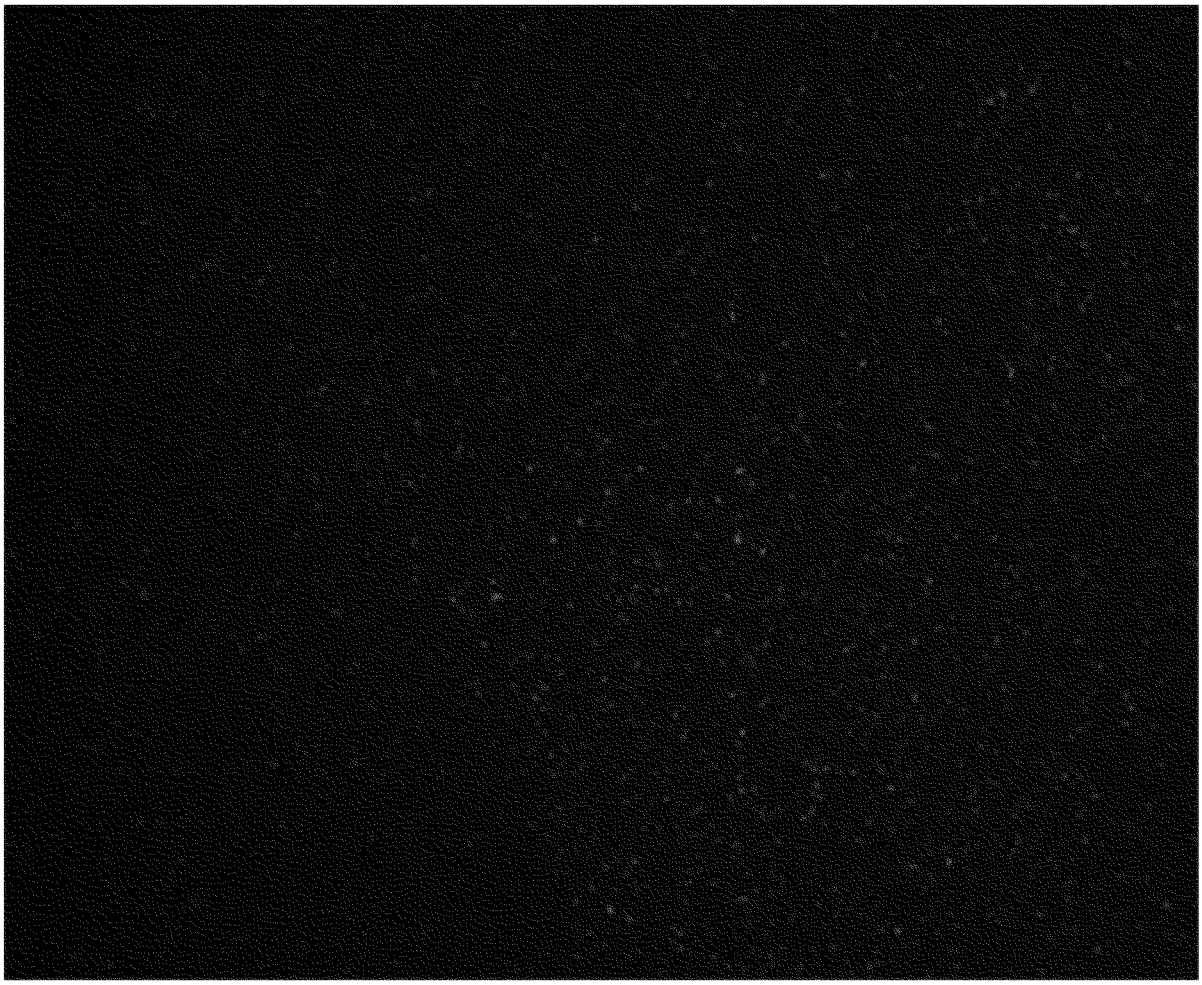
[도2]


바코드
바코드
합성된 핵산 분자
바코드 태깅
병렬적 염기서열 분석
검증된 염기서열 정보
핵산 분자 위치 정보 획득
핵산분자 위치 정보
펄스 레이저로 검증된 염기서열 핵산분자 회수

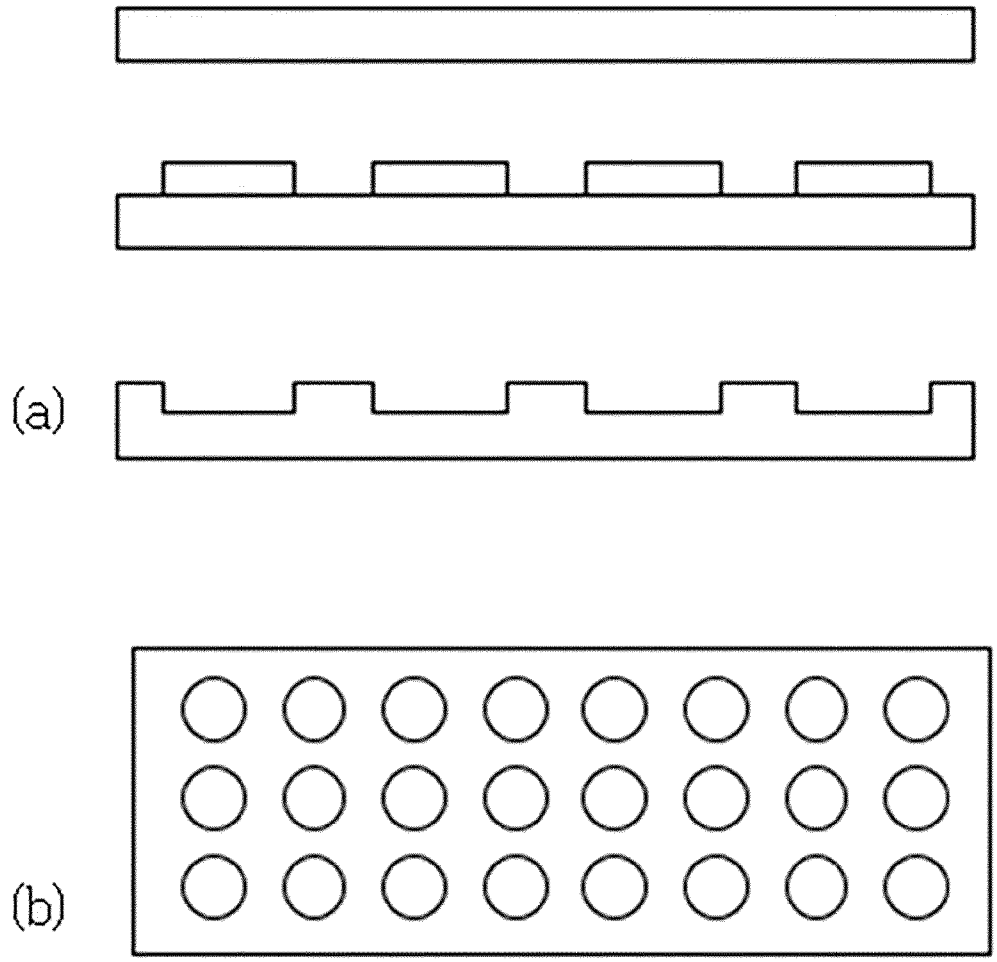
[도3]



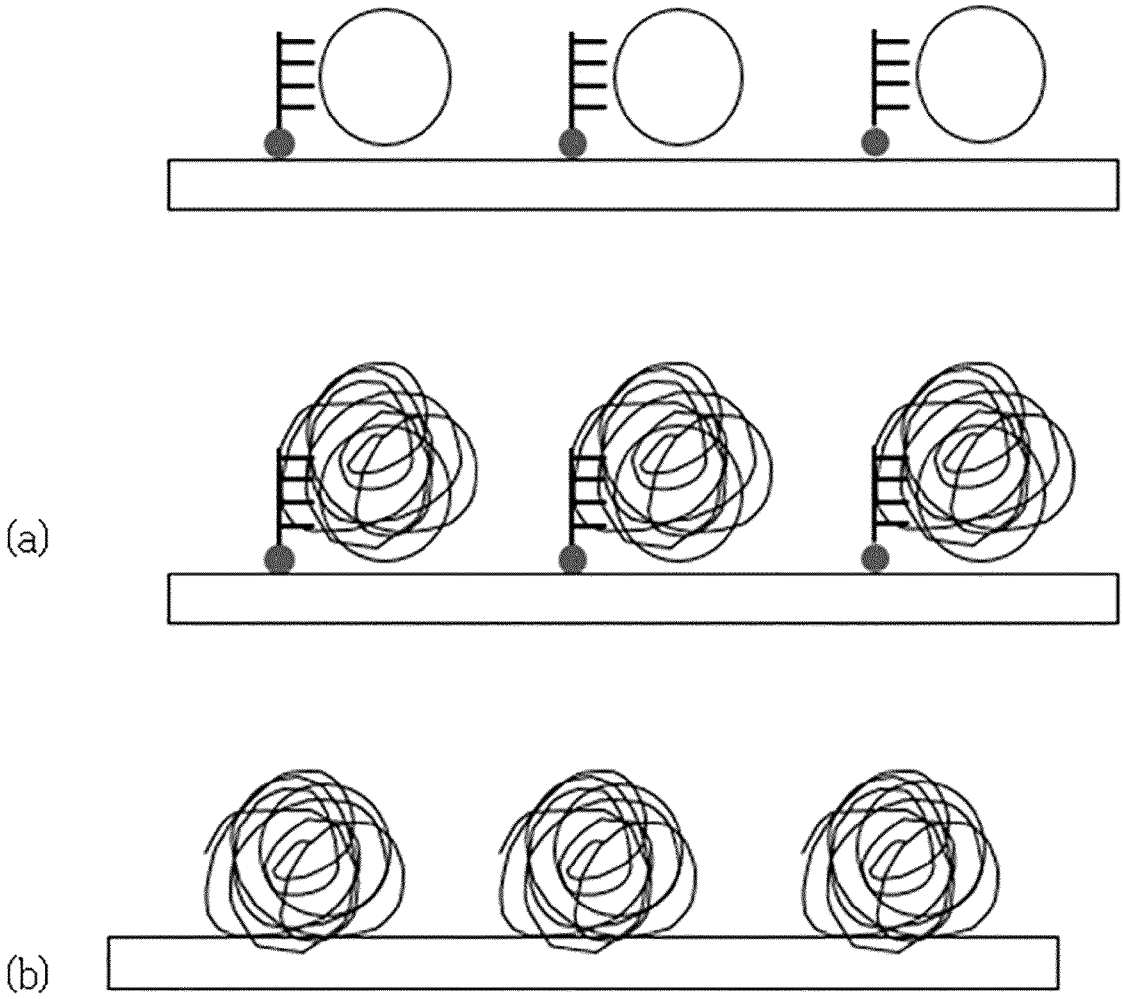
[도4]



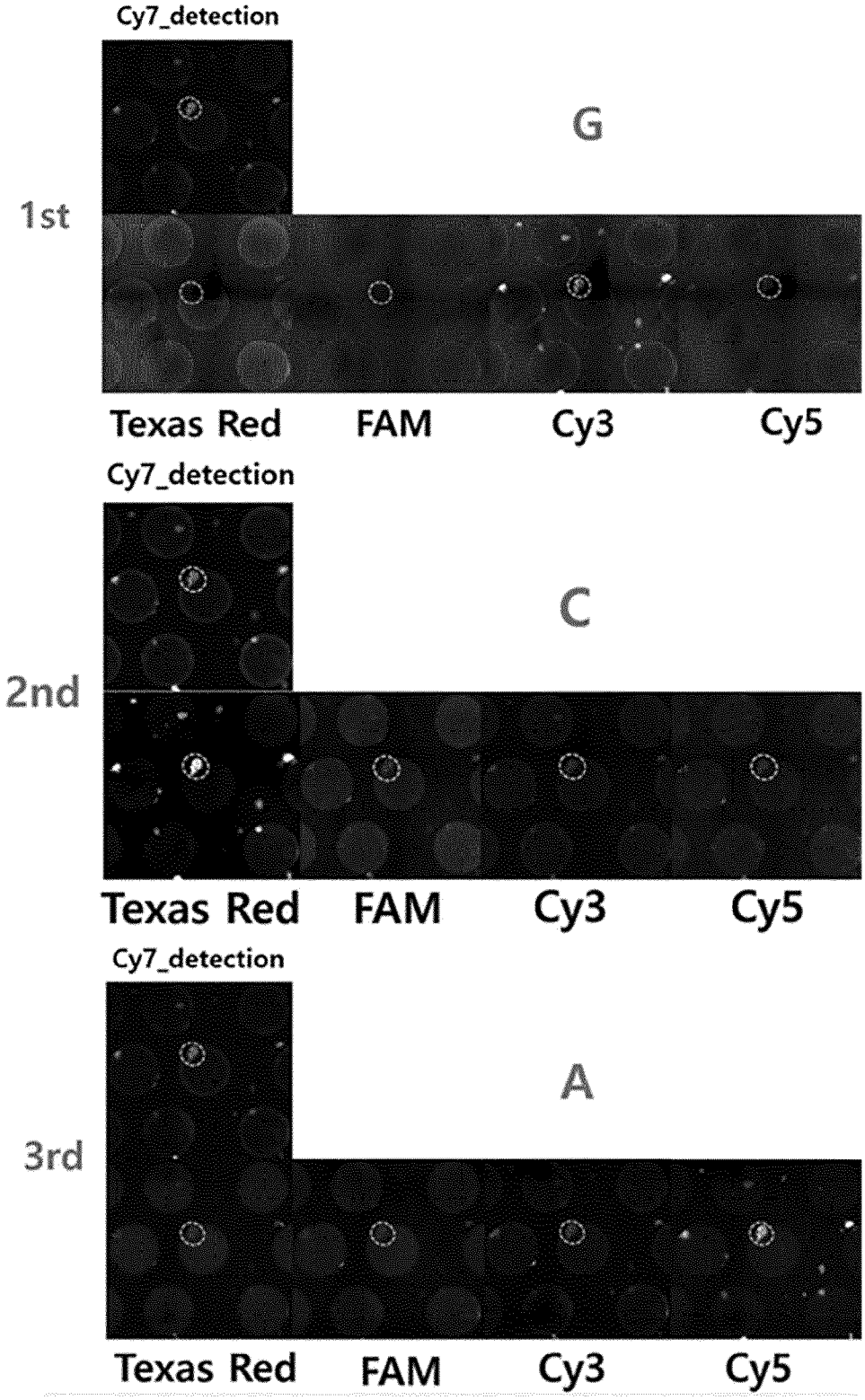
[도5]



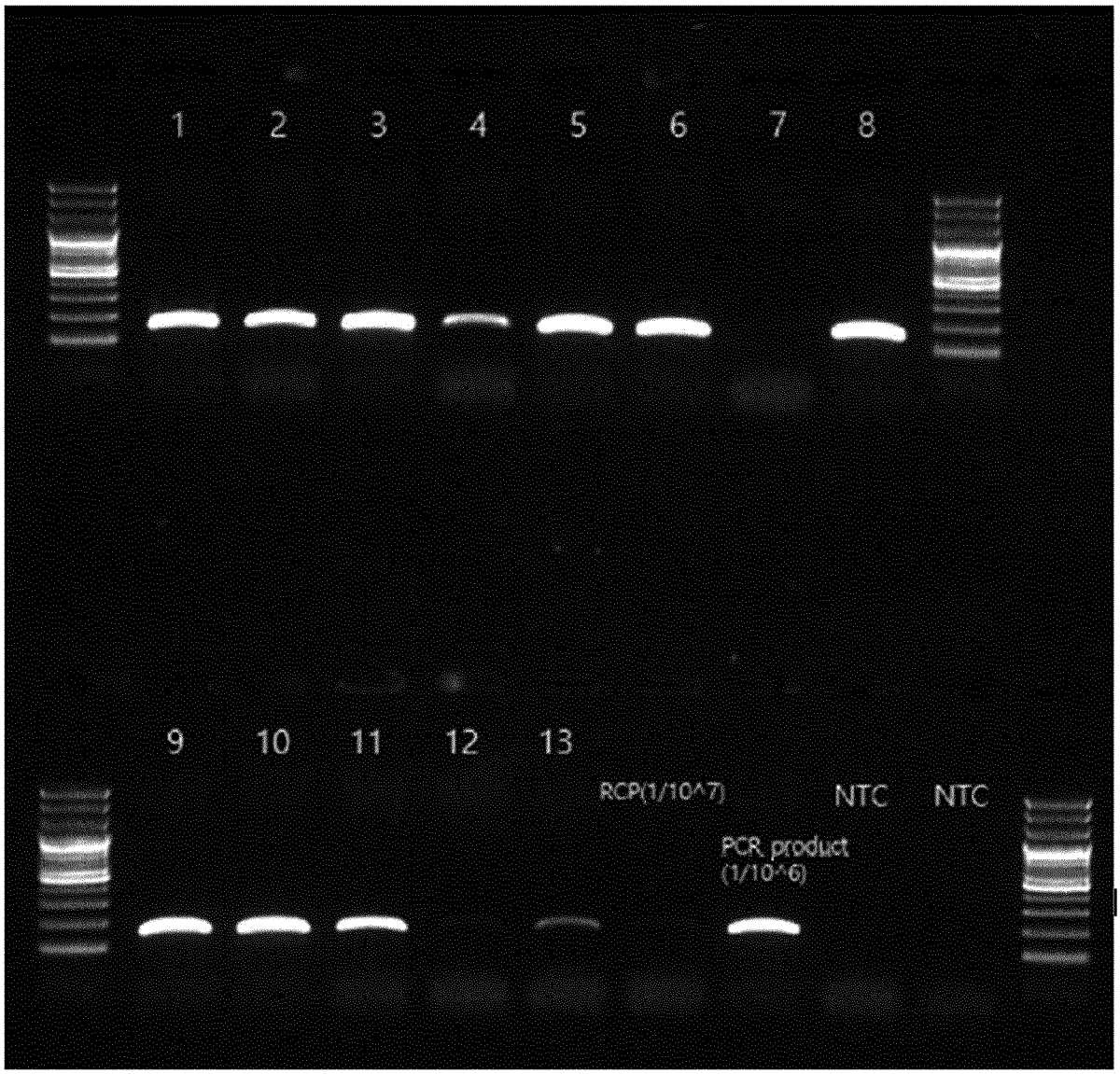
[도6]



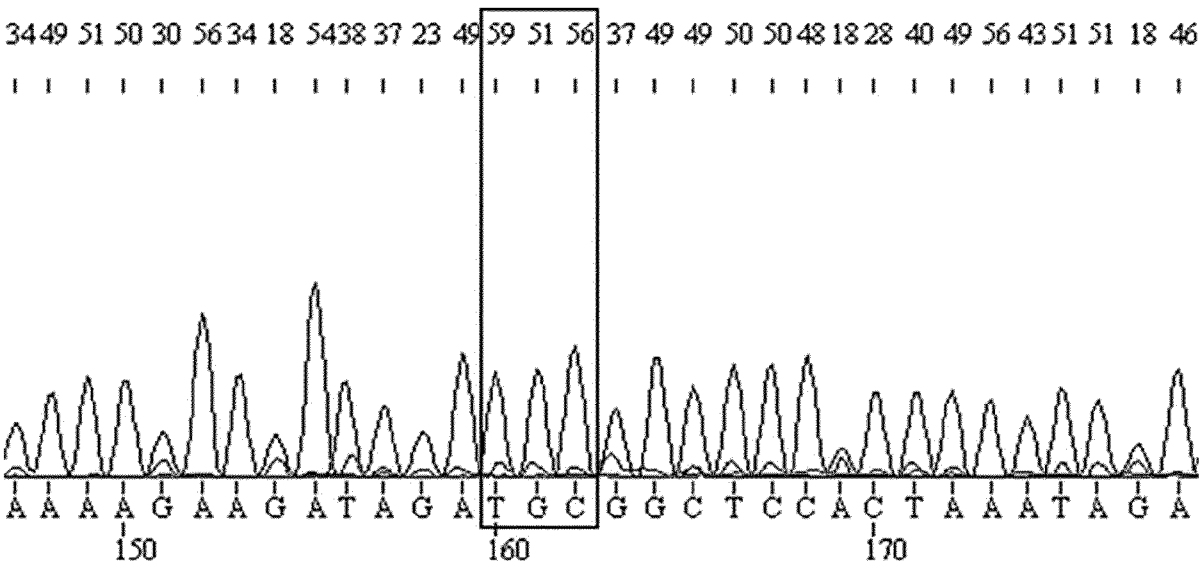
[도7]


Cy7_detection
G
1st
Texas Red
FAM
Cy3
Cy5
Cy7_detection
C
2nd
Texas Red
FAM
Cy3
Cy5
Cy7_detection
A
3rd
Texas Red
FAM
Cy3
Cy5

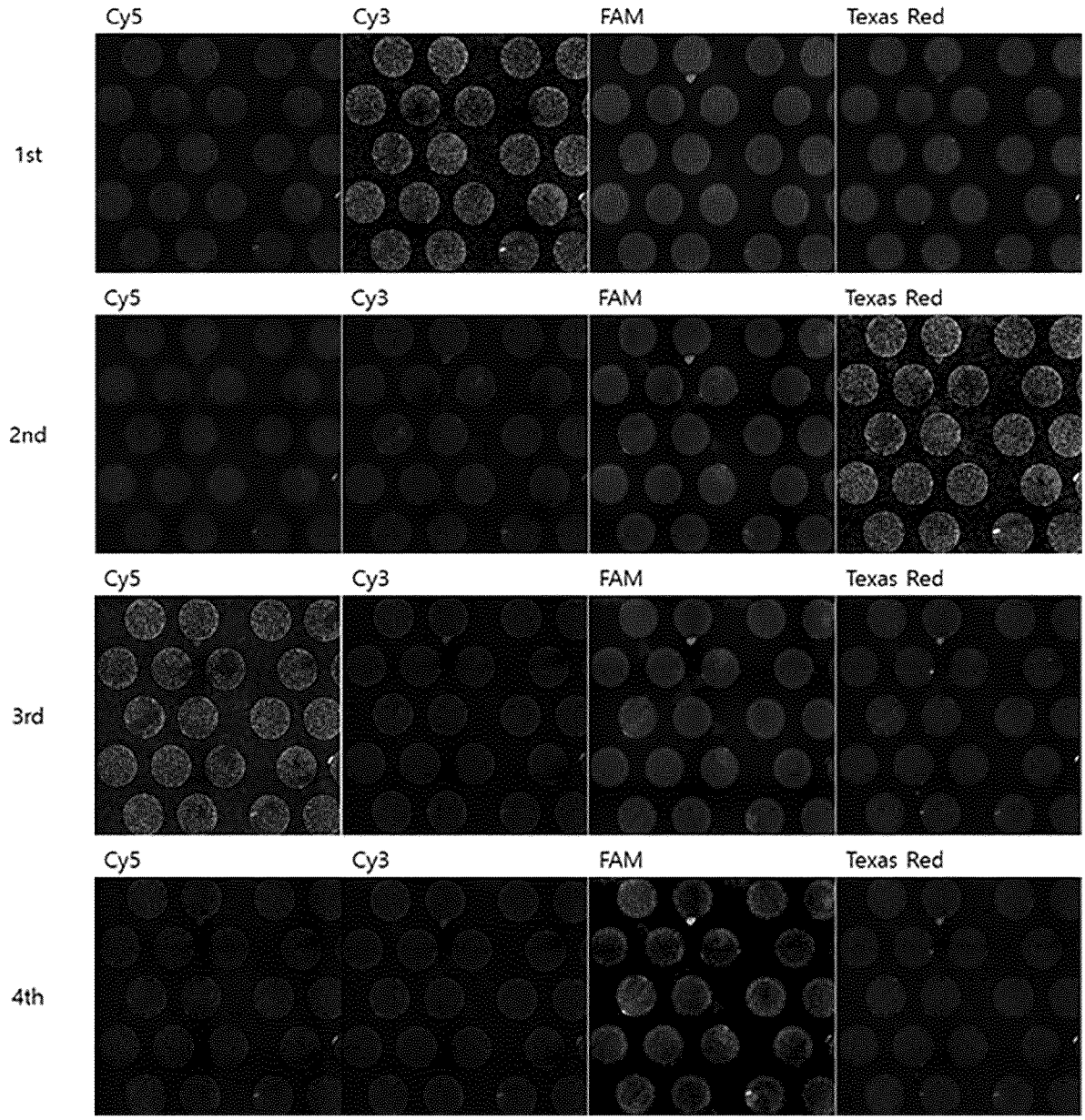
[도8]


1
2
3
4
5
6
7
8
9
10
11
12
13
RCP(1/10^7)
NTC
NTC
PCR product
(1/10^6)

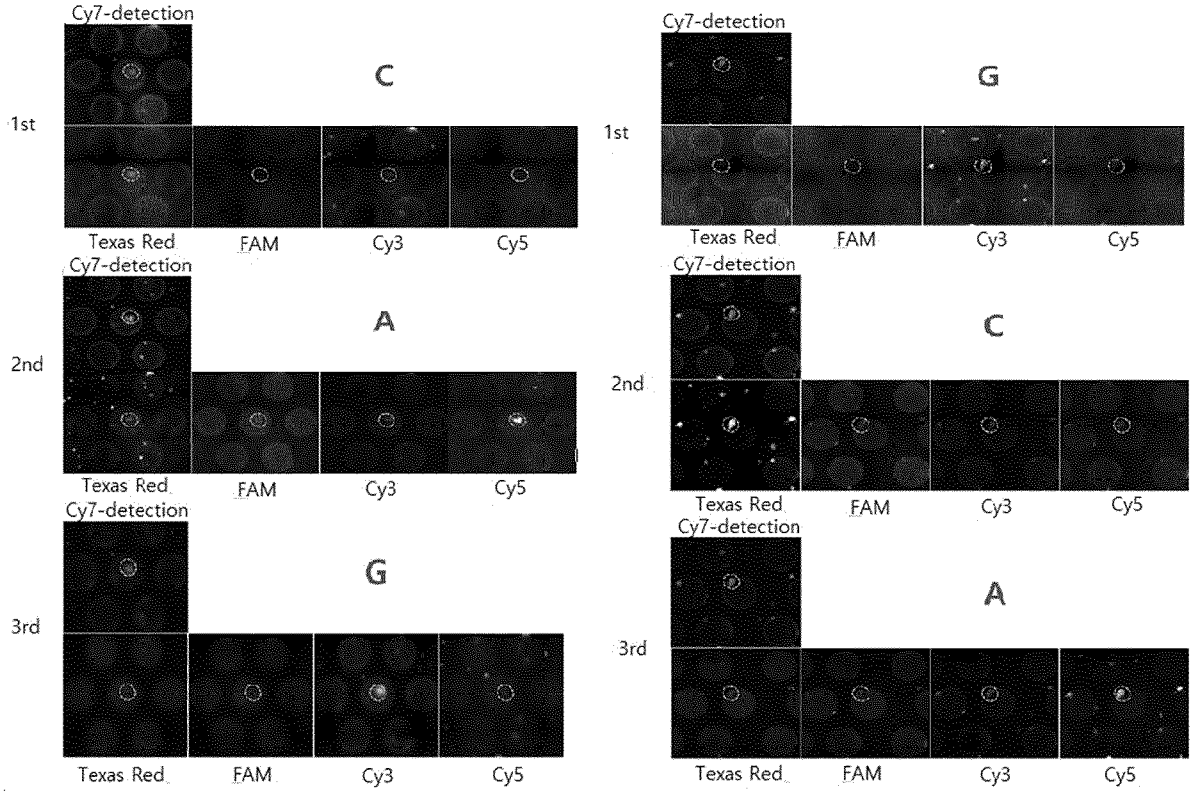
[도9]


34 49 51 50 30 56 34 18 54 38 37 23 49 59 51 56 37 49 49 50 50 48 18 28 40 49 56 43 51 51 18 46
A A A A G A A G A T A G A T G C G G C T C C A C T A A A T A G A
150
160
170

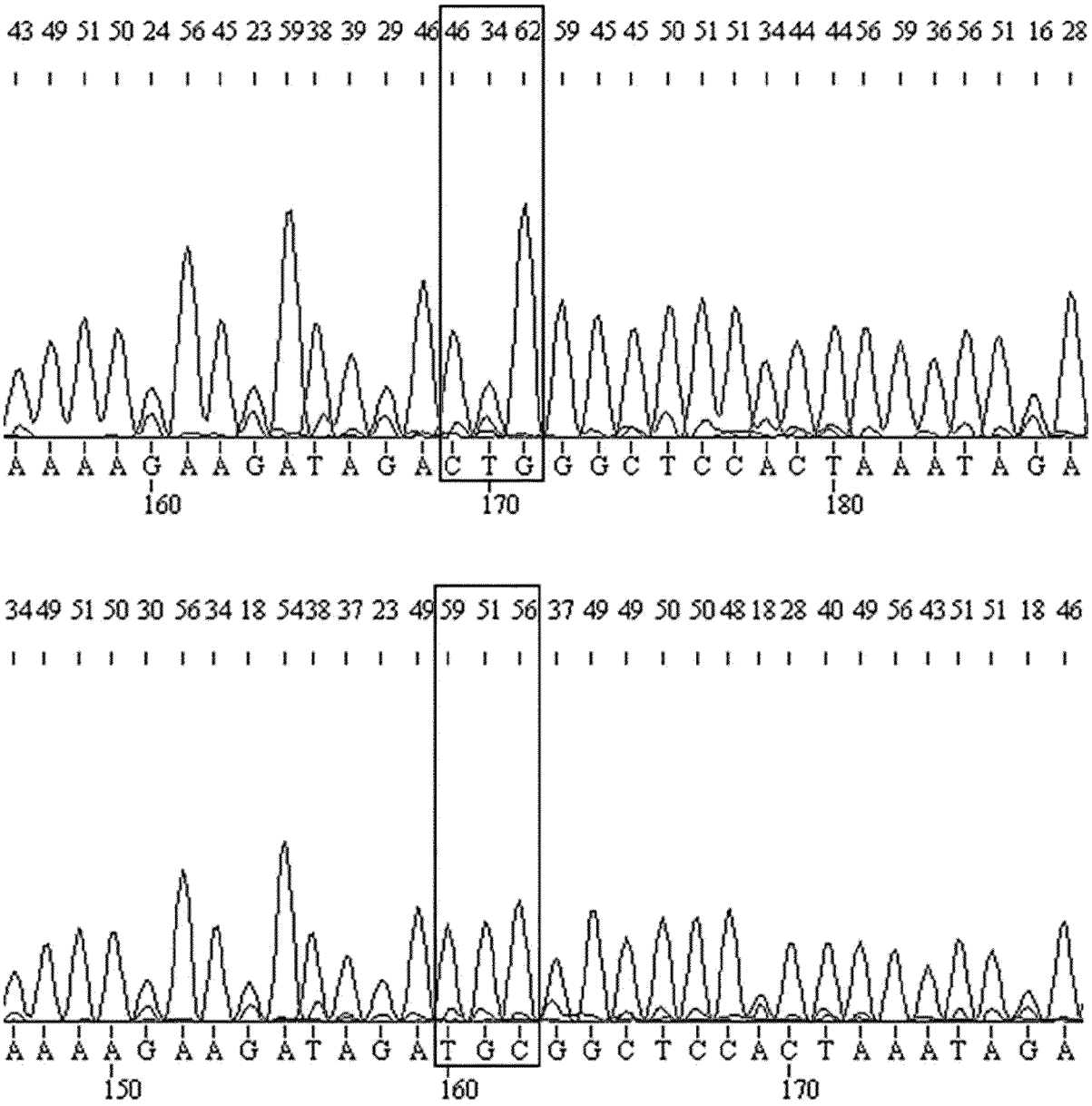
[도10]



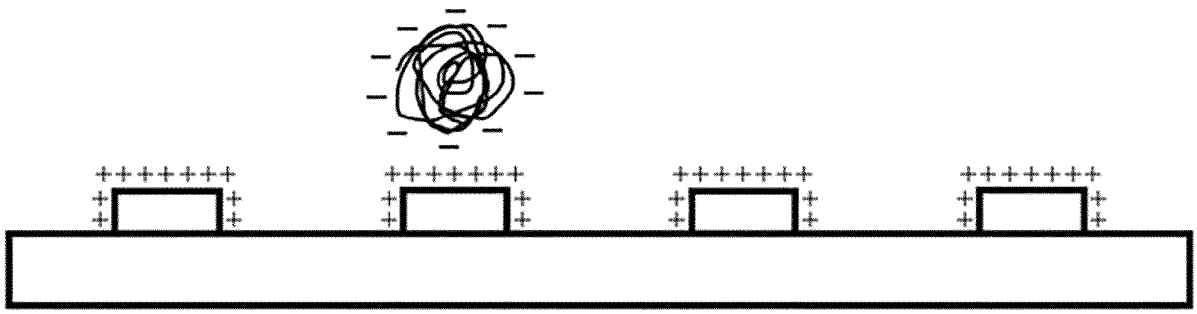
[도11]


Cy7-detection
1st
C
Texas Red
FAM
Cy3
Cy5
Cy7-detection
2nd
A
Texas Red
FAM
Cy3
Cy5
Cy7-detection
3rd
G
Texas Red
FAM
Cy3
Cy5
Cy7-detection
1st
G
Texas Red
FAM
Cy3
Cy5
Cy7-detection
2nd
C
Texas Red
FAM
Cy3
Cy5
Cy7-detection
3rd
A
Texas Red
FAM
Cy3
Cy5

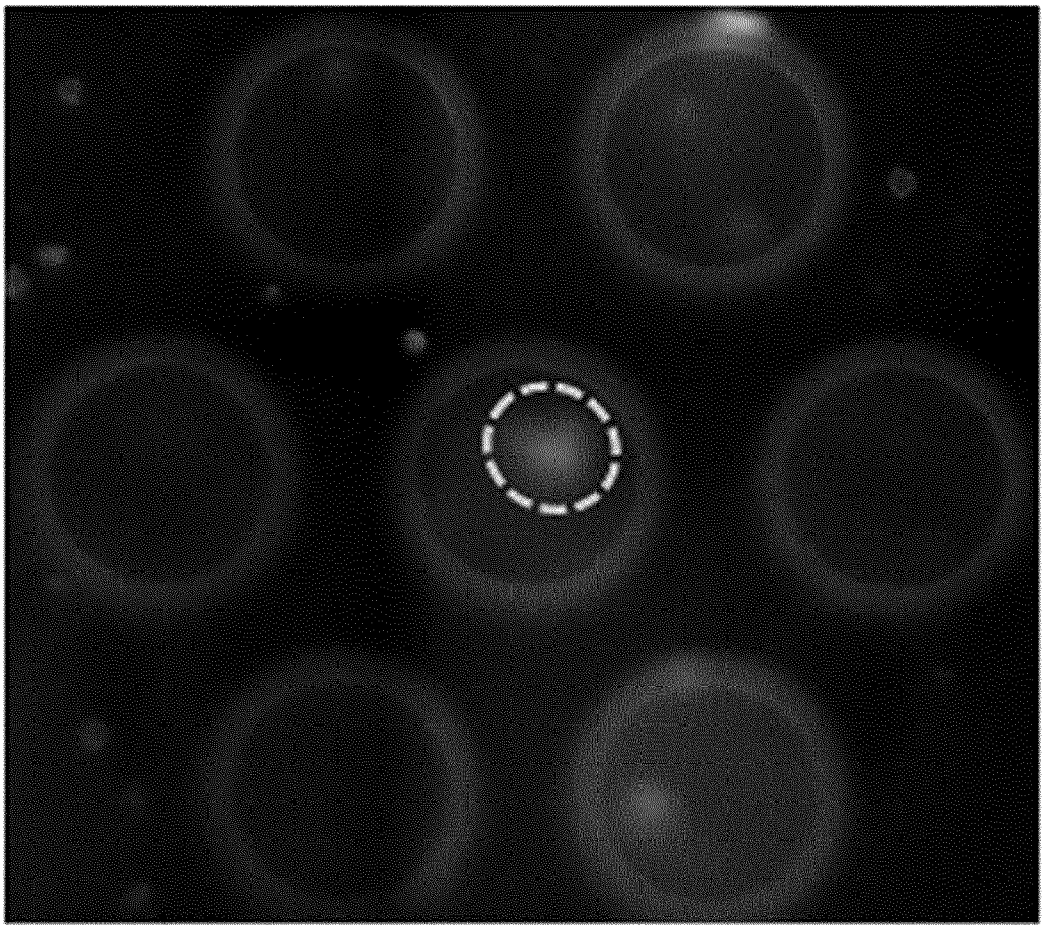
[도12]


43 49 51 50 24 56 45 23 59 38 39 29 46 46 34 62 59 45 45 50 51 51 34 44 44 56 59 36 56 51 16 28
A A A A G A A G A T A G A C T G G G C T C C A C T A A A T A G A
160
170
180
34 49 51 50 30 56 34 18 54 38 37 23 49 59 51 56 37 49 49 50 50 48 18 28 40 49 56 43 51 51 18 46
A A A A G A A G A T A G A T G C G G C T C C A C T A A A T A G A
150
160
170

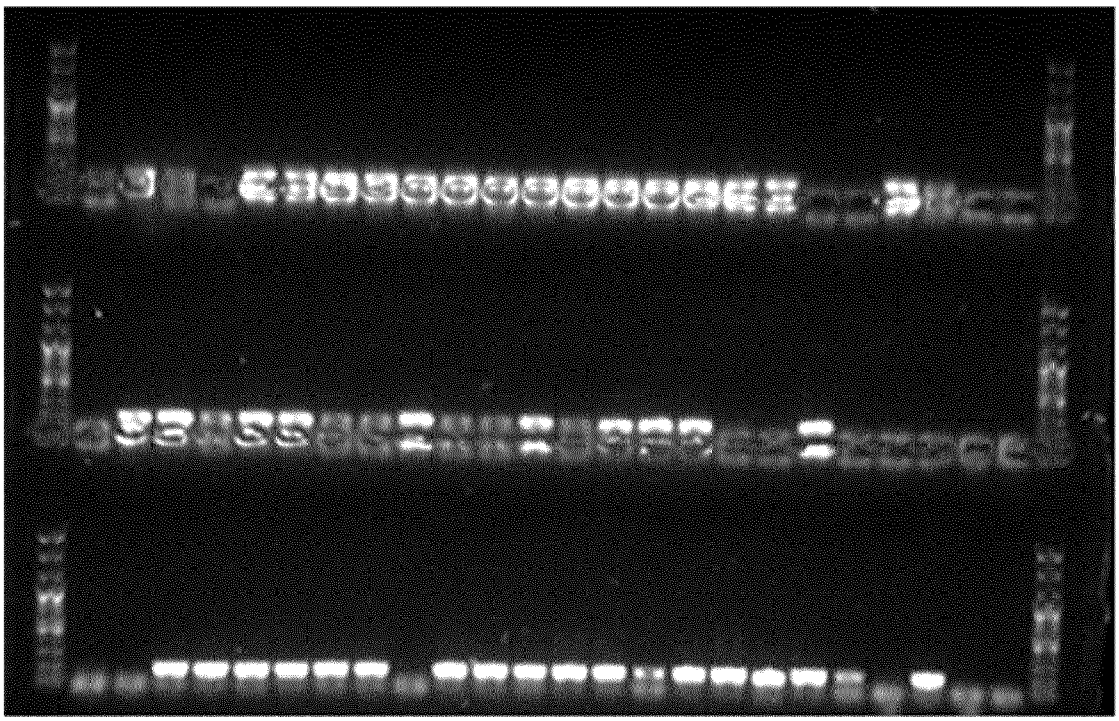
[도13]



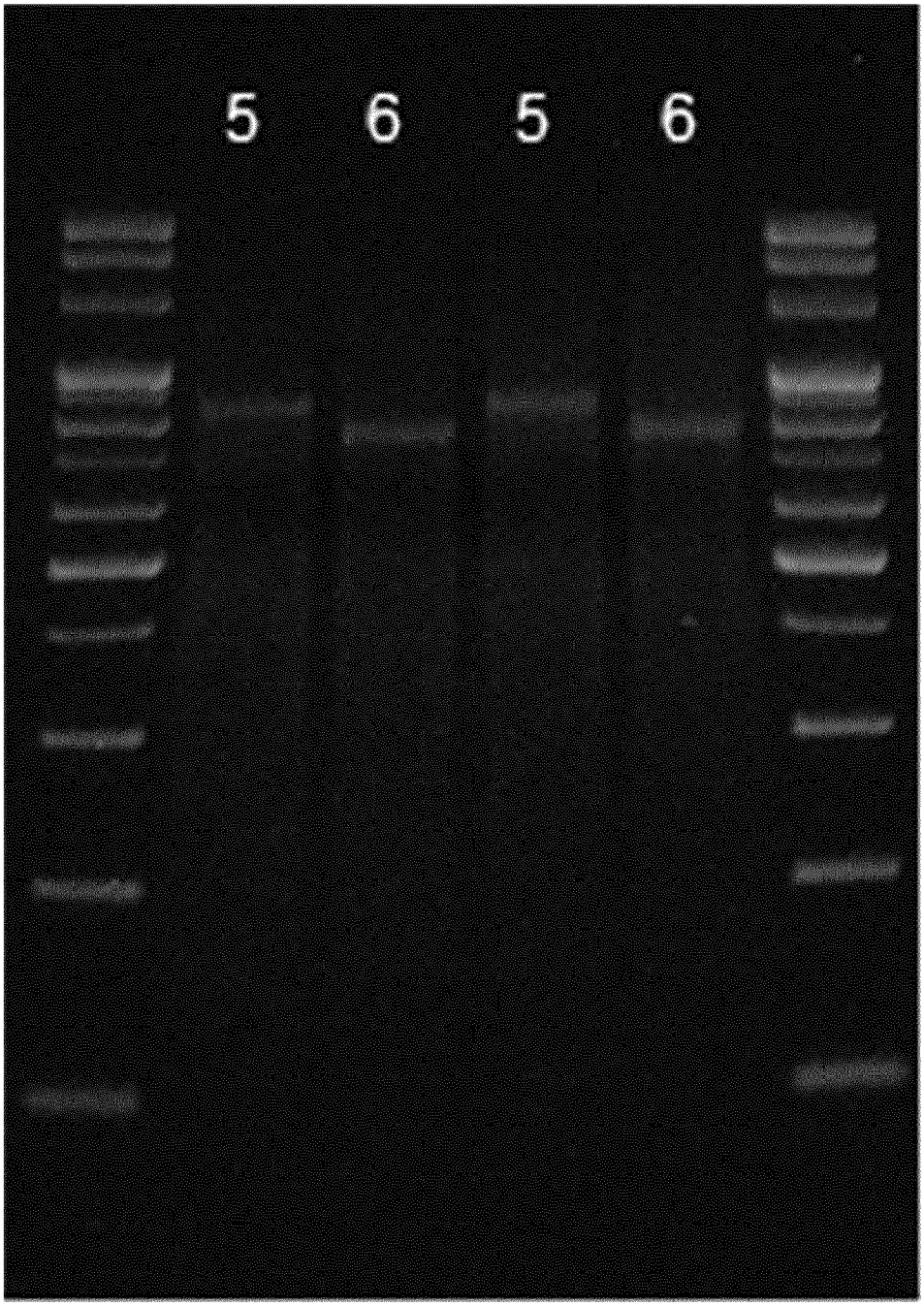
[도14]



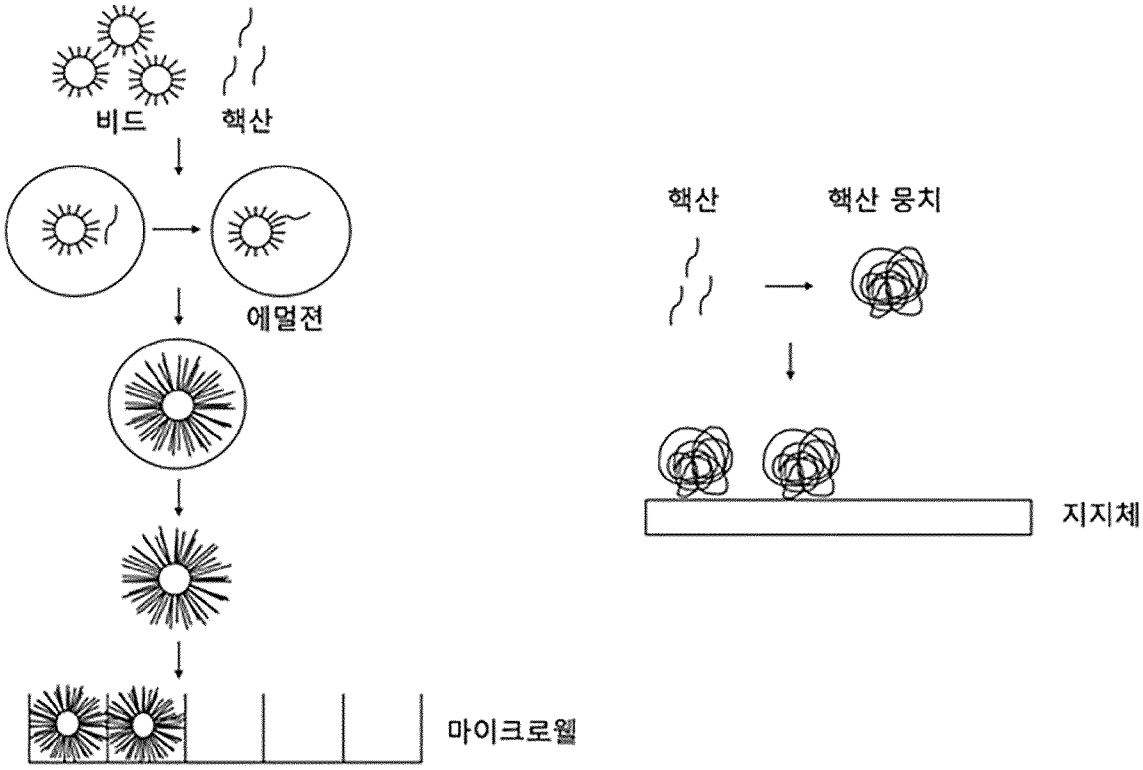
[도15]



[도16]



[도17]



INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2018/005493

A. CLASSIFICATION OF SUBJECT MATTER

C12Q 1/6806(2018.01)i, C12N 15/10(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12Q 1/6806; G06F 19/22; C12N 15/10; G01N 33/50; C12Q 1/68; C12Q 1/24

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean Utility models and applications for Utility models: IPC as above
Japanese Utility models and applications for Utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: nucleic acid, bar code, location coordinate, parallel sequence, support body, phi 29 DNA polymerase

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2013-0018575 A (INDUSTRY-ACADEMIC COOPERATION FOUNDATION, YONSEI UNIVERSITY) 25 February 2013 See paragraphs [0033], [0073]; claims 1-13.	1-26
Y	KR 10-2014-0075611 A (SEOUL NATIONAL UNIVERSITY R&DB FOUNDATION) 19 June 2014 See paragraphs [0013]-[0042], [0062]-[0064]; claims 1-21.	1-26
A	US 2015-0072873 A1 (CAMBRIAN GENOMICS, INC.) 12 March 2015 See the entire document.	1-26
A	KR 10-2014-0111224 A (SEOUL NATIONAL UNIVERSITY R&DB FOUNDATION) 18 September 2014 See the entire document.	1-26
A	US 2016-0244825 A1 (ABVITRO, INC.) 25 August 2016 See the entire document.	1-26

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 AUGUST 2018 (23.08.2018)	**23 AUGUST 2018 (23.08.2018)**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office Government Complex Daejeon Building 4, 189, Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea Facsimile No. +82-42-481-8578	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2018/005493

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: **28, 29**
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
Claims 28 and 29 refer to a claim not meeting the requirement of PCT Rule 6.4(a).

3. ☒ Claims Nos.: **27**
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2018/005493

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2013-0018575 A	25/02/2013	KR 10-1454886 B1	03/11/2014
		KR 10-1467969 B1	02/12/2014
		KR 10-2014-0004053 A	10/01/2014
		US 2014-0309118 A1	16/10/2014
		US 2016-0222380 A1	04/08/2016
		US 9340826 B2	17/05/2016
		WO 2013-019075 A2	07/02/2013
		WO 2013-019075 A3	04/07/2013
KR 10-2014-0075611 A	19/06/2014	CN 104884957 A	02/09/2015
		CN 104884957 B	26/12/2017
		EP 2919008 A1	16/09/2015
		KR 10-1595159 B1	18/02/2016
		KR 10-2016-0019913 A	22/02/2016
		US 2015-0322485 A1	12/11/2015
		WO 2014-088380 A1	12/06/2014
US 2015-0072873 A1	12/03/2015	WO 2014-179735 A1	06/11/2014
KR 10-2014-0111224 A	18/09/2014	KR 10-1575457 B1	11/12/2015
		US 2016-0010085 A1	14/01/2016
		US 9850482 B2	26/12/2017
		WO 2014-137193 A1	12/09/2014
US 2016-0244825 A1	25/08/2016	AU 2015-318011 A1	06/04/2017
		CA 2961210 A1	24/03/2016
		CN 107002076 A	01/08/2017
		EP 3194593 A1	26/07/2017
		JP 2017-527313 A	21/09/2017
		KR 10-2017-0063726 A	08/06/2017
		WO 2016-044227 A1	24/03/2016

국 제 조 사 보 고 서

국제출원번호
PCT/KR2018/005493

A. 발명이 속하는 기술분류(국제특허분류(IPC))

C12Q 1/6806(2018.01)i, C12N 15/10(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

C12Q 1/6806; G06F 19/22; C12N 15/10; G01N 33/50; C12Q 1/68; C12Q 1/24

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 핵산, 바코드, 위치 좌표, 병렬적 시퀀싱, 지지체, 파이 29 DNA 폴리머라아제

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2013-0018575 A (연세대학교 산학협력단) 2013.02.25 단락 [0033], [0073]; 청구항 1-13 참조.	1-26
Y	KR 10-2014-0075611 A (서울대학교산학협력단) 2014.06.19 단락 [0013]-[0042], [0062]-[0064]; 청구항 1-21 참조.	1-26
A	US 2015-0072873 A1 (CAMBRIAN GENOMICS, INC.) 2015.03.12 전체 문헌 참조.	1-26
A	KR 10-2014-0111224 A (서울대학교산학협력단) 2014.09.18 전체 문헌 참조.	1-26
A	US 2016-0244825 A1 (ABVITRO, INC.) 2016.08.25 전체 문헌 참조.	1-26

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.

☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

"A" 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

"E" 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

"L" 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

"O" 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

"P" 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

"T" 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

"X" 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

"Y" 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

"&" 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일	국제조사보고서 발송일
2018년 08월 23일 (23.08.2018)	2018년 08월 23일 (23.08.2018)

ISA/KR의 명칭 및 우편주소
대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)
팩스 번호 +82-42-481-8578

심사관
감유림
전화번호 +82-42-481-3516

국 제 조 사 보 고 서

국제출원번호
PCT/KR2018/005493

제2기재란 일부 청구항을 조사할 수 없는 경우의 의견(첫 번째 용지의 2의 계속)

PCT 제17조(2)(a)의 규정에 따라 다음과 같은 이유로 일부 청구항에 대하여 본 국제조사보고서가 작성되지 아니하였습니다.

1. ☐ 청구항:
이 청구항은 본 기관이 조사할 필요가 없는 대상에 관련됩니다. 즉,

2. ☒ 청구항: 28, 29
이 청구항은 유효한 국제조사를 수행할 수 없을 정도로 소정의 요건을 충족하지 아니하는 국제출원의 부분과 관련됩니다. 구체적으로는,
청구항 28-29는 PCT 규칙 6.4(a)의 규정을 충족시키지 않는 청구항을 인용하고 있습니다.

3. ☒ 청구항: 27
이 청구항은 종속청구항이나 PCT규칙 6.4(a)의 두 번째 및 세 번째 문장의 규정에 따라 작성되어 있지 않습니다.

제3기재란 발명의 단일성이 결여된 경우의 의견(첫 번째 용지의 3의 계속)

본 국제조사기관은 본 국제출원에 다음과 같이 다수의 발명이 있다고 봅니다.

1. ☐ 출원인이 모든 추가수수료를 기간 내에 납부하였으므로, 본 국제조사보고서는 모든 조사 가능한 청구항을 대상으로 합니다.

2. ☐ 추가수수료 납부를 요구하지 않고도 모든 조사 가능한 청구항을 조사할 수 있었으므로, 본 기관은 추가수수료 납부를 요구하지 아니하였습니다.

3. ☐ 출원인이 추가수수료의 일부만을 기간 내에 납부하였으므로, 본 국제조사보고서는 수수료가 납부된 청구항만을 대상으로 합니다. 구체적인 청구항은 아래와 같습니다.

4. ☐ 출원인이 기간 내에 추가수수료를 납부하지 아니하였습니다. 따라서 본 국제조사보고서는 청구범위에 처음 기재된 발명에 한정되어 있으며, 해당 청구항은 아래와 같습니다.

이의신청에 관한 기재

☐ 출원인의 이의신청 및 이의신청료 납부(해당하는 경우)와 함께 추가수수료가 납부되었습니다.

☐ 출원인의 이의신청과 함께 추가수수료가 납부되었으나 이의신청료가 보정요구서에 명시된 기간 내에 납부되지 아니하였습니다.

☐ 이의신청 없이 추가수수료가 납부되었습니다.

국제조사보고서
대응특허에 관한 정보

국제출원번호
PCT/KR2018/005493

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2013-0018575 A	2013/02/25	KR 10-1454886 B1	2014/11/03
		KR 10-1467969 B1	2014/12/02
		KR 10-2014-0004053 A	2014/01/10
		US 2014-0309118 A1	2014/10/16
		US 2016-0222380 A1	2016/08/04
		US 9340826 B2	2016/05/17
		WO 2013-019075 A2	2013/02/07
		WO 2013-019075 A3	2013/07/04
KR 10-2014-0075611 A	2014/06/19	CN 104884957 A	2015/09/02
		CN 104884957 B	2017/12/26
		EP 2919008 A1	2015/09/16
		KR 10-1595159 B1	2016/02/18
		KR 10-2016-0019913 A	2016/02/22
		US 2015-0322485 A1	2015/11/12
		WO 2014-088380 A1	2014/06/12
US 2015-0072873 A1	2015/03/12	WO 2014-179735 A1	2014/11/06
KR 10-2014-0111224 A	2014/09/18	KR 10-1575457 B1	2015/12/11
		US 2016-0010085 A1	2016/01/14
		US 9850482 B2	2017/12/26
		WO 2014-137193 A1	2014/09/12
US 2016-0244825 A1	2016/08/25	AU 2015-318011 A1	2017/04/06
		CA 2961210 A1	2016/03/24
		CN 107002076 A	2017/08/01
		EP 3194593 A1	2017/07/26
		JP 2017-527313 A	2017/09/21
		KR 10-2017-0063726 A	2017/06/08
		WO 2016-044227 A1	2016/03/24