



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년04월20일
(11) 등록번호 10-2102107
(24) 등록일자 2020년04월13일

(51) 국제특허분류(Int. Cl.)
C12N 15/10 (2017.01) C12Q 1/6855 (2018.01)
C12Q 1/6869 (2018.01)
(52) CPC특허분류
C12N 15/1093 (2013.01)
C12Q 1/6855 (2018.05)
(21) 출원번호 10-2019-0063834
(22) 출원일자 2019년05월30일
심사청구일자 2019년05월30일
(56) 선행기술조사문헌
KR1020180055905 A
JP2017512491 A

(73) 특허권자
단국대학교 천안캠퍼스 산학협력단
충청남도 천안시 동남구 단대로 119, 단국대학교천안캠퍼스내(안서동)
(72) 발명자
한규동
충청남도 천안시 동남구 안서1길 5-16, 102동 304호 (안서동, 고운여의주아파트)
문세영
충청남도 천안시 서북구 3공단6로 33, 103동 803호 (차암동, 한화꿈에그린스마일시티)
신원석
경기도 안산시 단원구 다리간2길 8, 203호 (고잔동)
(74) 대리인
특허법인태백

전체 청구항 수 : 총 7 항

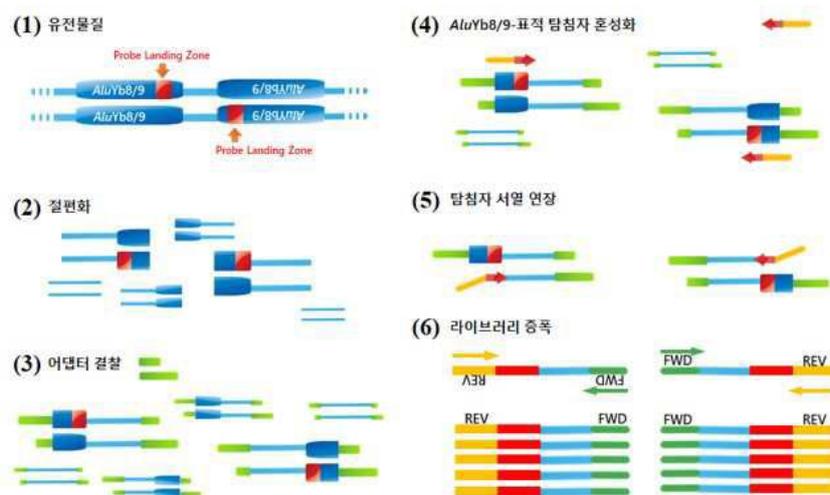
심사관 : 한지혜

(54) 발명의 명칭 이동성 유전인자 SINE 선택적 발굴을 위한 HiSeq 시퀀서 기반의 DNA 라이브러리 제작 방법

(57) 요약

본 발명은 이동성 유전인자 SINE 선택적 발굴을 위한 HiSeq 시퀀서 기반의 DNA 라이브러리 제작 방법에 관한 것으로서, 이동성 유전인자의 임상 및 학술적 가치와 차세대 염기서열해독 기술의 발달 그리고 비용의 감소 및 분석 시스템 발전을 기반으로 인간의 특이적인 *AluYb8/9*를 동정할 수 있는 *AluYb8/9*-표적 염기서열 분석 방법을 개발함으로써 간단하고 빠르게 *AluYb8/9*에 의한 개개인의 유전적 변이와 그에 따른 유전적 영향연구에 한 방법이 될 수 있다. 또한 본 발명을 바탕으로 *AluYb8/9*를 포함한 인간 유전체내 이동성을 보이고 유전적 변이를 일으킬 수 있는 활동적인 이동성 유전인자만을 표적화하는 추가 기술 개발이 가능하다. 이동성 유전인자 표적화 염기서열 분석뿐만 아니라 표적화 서열 데이터 분석 방법이 단순화 보편화됨으로써 임상의학적인 진단 마커 개발에 이용될 수 있을 것이라 기대할 수 있다.

대표도 - 도2



(52) CPC특허분류

- C12Q 1/6869 (2018.05)
- C12Q 2523/301 (2013.01)
- C12Q 2525/191 (2013.01)
- C12Q 2535/122 (2019.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	2017R1D1A1A02019421
부처명	교육부
연구관리전문기관	한국연구재단
연구사업명	이공분야기초연구사업/기본연구지원사업(SGER)
연구과제명	이동성 유전인자 표적 NGS 기법을 통해 한국인 대장암 환자의 암세포 내에서 활동적인 이
동성 유전인자 발굴	
기여율	1/1
주관기관	단국대학교 천안캠퍼스 산학협력단
연구기간	2018.06.01 ~ 2019.05.31

명세서

청구범위

청구항 1

서열번호 3으로 표시되며, 이동성 유전인자 SINE(Short INterspersed Element)의 아종(Subfamily) 중 *AluYb8* 또는 *AluYb9*를 표적하는 탐침자(probe)를 유효성분으로 포함하는 *AluYb8* 또는 *AluYb9* 표적용 DNA 라이브러리(library) 제작용 조성물.

청구항 2

삭제

청구항 3

제1항에 있어서, 상기 DNA 라이브러리는 차세대 염기서열 해독(Next generation sequencing, NGS) 분석을 위한 것임을 특징으로 하는 *AluYb8* 또는 *AluYb9* 표적용 DNA 라이브러리 제작용 조성물.

청구항 4

제3항에 있어서, 상기 NGS 분석은 일루미나(Illumina) 플랫폼 HiSeq 시퀀서를 이용하는 것을 특징으로 하는 *AluYb8* 또는 *AluYb9* 표적용 DNA 라이브러리 제작용 조성물.

청구항 5

제1항, 제3항 및 제4항 중 어느 한 항의 조성물을 포함하는 *AluYb8* 또는 *AluYb9* 표적용 DNA 라이브러리 제작용 키트.

청구항 6

게노믹(genomic) DNA를 절편화하는 단계;

상기 절편화된 DNA를 어댑터 접합시켜 양 말단을 평활화(end repair)시키는 단계;

상기 평활화된 DNA와, 서열번호 3으로 표시되는 이동성 유전인자 SINE의 아종(Subfamily) 중 *AluYb8* 또는 *AluYb9*를 표적하는 탐침자(probe)를 혼성화시키는 단계;

상기 혼성화된 DNA를 연장화시키는 단계; 및

상기 연장화된 DNA 단편 양 말단의 어댑터 서열을 이용하여 PCR 증폭시키는 단계를 포함하는 *AluYb8* 또는 *AluYb9* 표적용 DNA 라이브러리 제작 방법.

청구항 7

삭제

청구항 8

제6항에 있어서, 상기 DNA 라이브러리는 NGS 분석을 위한 것임을 특징으로 하는 *AluYb8* 또는 *AluYb9* 표적용 DNA 라이브러리 제작 방법.

청구항 9

제8항에 있어서, 상기 NGS 분석은 일루미나(Illumina) 플랫폼 HiSeq 시퀀서를 이용하는 것을 특징으로 하는 *AluYb8* 또는 *AluYb9* 표적용 DNA 라이브러리 제작 방법.

발명의 설명

기술분야

[0001] 본 발명은 이동성 유전인자 SINE 선택적 발굴을 위한 HiSeq 시퀀서 기반의 DNA 라이브러리 제작 방법에 대한 것으로, 상세하게는 차세대 염기서열 해독 장비인 HiSeq 시퀀서를 이용하여 인간 유전체에서 이동성 유전인자 SINE (*AluYb8* 그리고 *AluYb9*)들을 선택적으로 동정하기 위한 탐침자(Probe)의 제작, 실험법 및 생명정보학적 분석 파이프 라인 구축에 관한 것이다.

배경기술

[0002] 차세대 염기서열해독 기술(Next-Generation Sequencing Technology, NGS technology)은 유전물질(DNA)로부터 다수 그리고 대량의 염기서열 해독을 동시에 진행할 수 있고, 단기간에 비교적 적은 비용으로 높은 처리량의 유전체 데이터를 얻을 수 있는 기술이다. 이 기술이 상용화된 2007년 이후 현재까지 매년 혁신적인 기술개발이 진행되었고, 이에 맞추어 어플리케이션 기술 역시 다양하게 개량 및 개발이 되고 있다.

[0003] 2003년 인간 유전체 프로젝트(Human Genome Project)는 미국을 주도로 선진 6개국의 18개 기관이 참여하였고 13년의 시간과 약 30억 달러의 비용이 사용되어 약 30억개의 인간 전체 유전체 해독 및 해석을 완료한 전세계적 프로젝트였다. 하지만, 현재 차세대 염기서열해독 기술을 활용하면 인간 전체 유전체 해독 및 해석이 약 1주일이라는 짧은 소요시간과 상대적으로 대단히 저렴한 비용으로 가능하게 되었다. 차세대 염기서열해독 및 분석 기술의 발전함에 따라, 의생명과학 전분야에 적용되어 다양한 지적재산권이 생산되고 관련산업이 성장하는 큰 동력이 되었다. 또한, 차세대 염기서열 해독 및 분석 기술의 발전으로 새로운 어플리케이션 기술(예, 진단 마커)과 원천기술의 발전 또한 요구되는 시점이다.

[0004] 미국 Illumina사의 플랫폼 중 HiSeq 차세대 염기서열해독장비는 다양한 유전체 데이터 생성 장비 종류들 중 가장 보편적으로 사용되는 장비이다. Illumina사의 플랫폼은 전세계 유전체 데이터 생산량의 70% 이상을 차지하고 있고 높은 재연성과 정확도를 보이고 있다. 또한, 다양한 논문과 특허에서 HiSeq 장비를 이용해 얻어진 데이터를 분석 및 보고하고 있다. HiSeq은 Illumina사의 Sequencing By Synthesis (SBS) 기술과 Cluster generation 기술을 이용한 장비로 유리 Chip인 Flow Cell 표면에 존재하는 Adaptor oligo와 라이브러리를 혼성화 및 증폭한 후 하나씩 합성되는 핵산 종류에 따른 형광을 한 개씩 측정함으로써 염기서열을 해독하는 원리이다.

[0005] 현재 높은 정확성과 대량의 데이터를 필요로 하는 임상 및 연구 분야에서 사용되고 있고 다양한 목표 염기서열을 해독할 수 있는 응용기법들이 있다. 특히 인간을 포함하는 대형동물의 전체 유전체, 전사체, 후성유전체, 진 유전체 그리고 목표서열 증폭법을 이용한 특정 부위에 대한 집단연구에 특화된 플랫폼이다. 차세대 염기서열해독 기술의 발달은 전체 유전체 분석에 막대한 영향을 주었으며 라이브러리 구축 방법이 개발됨에 따라 연구자가 원하는 표적영역에 초점을 맞추어 집단 염기서열 분석이 가능하여 특정 집단(예, 환자군)에서 유전학적 차이를 설명할 수 있는 좋은 기반이 된다.

[0006] 하지만, 인간 유전체에 유전자 발현의 변화, 유전자 구조 변형 등을 야기하여 암과 같은 유전적 질병을 유발시키는 이동성 유전인자(Transposable Element, TE)가 약 44%를 차지하고 있다. 이들은 염기서열의 유사성이 매우 높아서 차세대 염기서열 해독기술을 직접 적용할 수 없으며, 목표 서열을 동정 방법과 분석법에 대한 연구가 미흡하다. 이동성 유전자는 삽입 메커니즘에 따라 레트로트랜스포손(Retrotransposon, Class I)과 DNA 트랜스포손(DNA transposon, Class II) 두 클래스로 분류된다. 이중 레트로트랜스포손은 스스로 리보핵산(Ribo nucleic acid, RNA)을 전사하여 이를 매개로 하는 삽입기작을 통해 새로운 유전체 영역에 삽입함으로써 유전적 변이 및 다양성에 관여하고 있다.

[0007] 레트로트랜스포손 중, Short Interspersed Element (SINE) 인자에 속하는 *Alu*과(科)는 인간 유전체의 약 10% 이상을 차지하는 유전인자이다. 인간 유전체에 약 100만 카피(copy)가 있다고 알려져 있으며 영장류 유전체에 특이적으로 삽입되어져 있다. *Alu*는 약 140 bp에 존재하는 Mid A-stretch를 기준으로 Left Arm과 Right Arm으로 구성되며 약 300 bp의 이동성 유전인자이다. 대부분의 *Alu*는 긴 영장류 진화 역사 동안 인자 이내에 유전적 구조적 변이와 단일염기 다형성(single nucleotide polymorphism)의 축적에 의해 *Alu*의 유전적 이동성을 상실한 것으로 알려져 있다. 하지만 최근 삽입된 몇몇의 *Alu*과와 *AluYb8/9*아과(亞科)는 여전히 이동성을 보인다고 보고되어 있으며, 이들은 인간의 질병에 관련된 유전자 조절 및 구조적 변화에 다양한 영향을 끼친다고 보고되었다. 따라서 참조 유전체에 보고 되지 않으며 인간 유전체에 새롭게 삽입되어 있는 *AluYb8/9*의 유전적 부위를 동정하고 *AluYb8/9* 삽입에 의한 유전적 변이 연구가 필요하다. 보고되지 않은 *AluYb8/9*의 삽입을 추적할 수 있는 기술을 통하여 특정 집단에서 공통변이 연구에 이용한다면 정확도 높은 진단마커가 개발될 수 있을 것이다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 미국등록특허 US 8663924 (2014.03.04 등록)

발명의 내용

해결하려는 과제

[0009] 본 발명의 목적은 이동성 유전인자 SINE(Short INterspersed Element)을 표적하는 탐침자(probe)를 유효성분으로 포함하는 SINE 표적용 DNA 라이브러리(library) 제작용 조성물 또는 이를 포함하는 SINE 표적용 DNA 라이브러리 제작용 키트를 제공하는데 있다.

[0010] 또한, 본 발명의 다른 목적은 이동성 유전인자 SINE(Short INterspersed Element)을 표적하는 탐침자(probe)를 이용한 SINE 표적용 DNA 라이브러리 제작 방법을 제공하는데 있다.

과제의 해결 수단

[0011] 상기 목적을 달성하기 위하여, 본 발명은 서열번호 3으로 표시되는 이동성 유전인자 SINE(Short INterspersed Element)을 표적하는 탐침자(probe)를 유효성분으로 포함하는 SINE 표적용 DNA 라이브러리(library) 제작용 조성물을 제공한다.

[0012] 또한, 본 발명은 조성물을 포함하는 SINE 표적용 DNA 라이브러리 제작용 키트를 제공한다.

[0013] 또한, 본 발명은 게노믹(genomic) DNA를 절편화하는 단계; 상기 절편화된 DNA를 어댑터 접합시켜 양 말단을 평활화(end repair)시키는 단계; 상기 평활화된 DNA와, 서열번호 3으로 표시되는 이동성 유전인자 SINE을 표적하는 탐침자(probe)를 혼성화시키는 단계; 상기 혼성화된 DNA를 연장화시키는 단계; 및 상기 연장화된 DNA 단편 양 말단의 어댑터 서열을 이용하여 PCR 증폭시키는 단계를 포함하는 SINE 표적용 DNA 라이브러리 제작 방법을 제공한다.

발명의 효과

[0014] 본 발명은 이동성 유전인자 SINE 선택적 발굴을 위한 HiSeq 시퀀서 기반의 DNA 라이브러리 제작 방법에 관한 것으로서, 이동성 유전인자의 임상 및 학술적 가치와 차세대 염기서열해독 기술의 발달 그리고 비용의 감소 및 분석 시스템 발전을 기반으로 인간의 특이적인 *AluYb8/9*를 동정할 수 있는 *AluYb8/9*-표적 염기서열 분석 방법을 개발함으로써 간단하고 빠르게 *AluYb8/9*에 의한 개개인의 유전적 변이와 그에 따른 유전적 영향연구에 한 방법이 될 수 있다. 또한 본 발명을 바탕으로 *AluYb8/9*를 포함한 인간 유전체내 이동성을 보이고 유전적 변이를 일으킬 수 있는 활동적인 이동성 유전인자만을 표적화하는 추가 기술 개발이 가능하다. 이동성 유전인자 표적화 염기서열 분석뿐만 아니라 표적화 서열 데이터 분석 방법이 단순화 보편화됨으로써 임상의학적 진단 마커 개발에 이용될 수 있을 것이라 기대할 수 있다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 일 실시예에 따른 차세대 염기서열 해독 및 해석을 위한 유전물질 가공부터 *AluYb8/9* 표적화 서열 데이터 분석 방법까지의 연구 절차를 나타낸다.

도 2는 본 발명의 일 실시예에 따른 차세대 염기서열분석을 위한 *AluYb8/9* 표적 라이브러리 구축 과정을 나타낸다.

도 3은 본 발명의 일 실시예에 따른 차세대 염기서열해독 데이터의 교정/가공 및 생명정보학적 분석 과정을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0016] 본 발명은 서열번호 3으로 표시되는 이동성 유전인자 SINE(Short INterspersed Element)을 표적하는 탐침자

(probe)를 유효성분으로 포함하는 SINE 표적용 DNA 라이브러리(library) 제작용 조성물을 제공한다.

- [0017] 바람직하게는, 상기 SINE은 *AluYb8* 또는 *AluYb9*일 수 있다.
- [0018] 바람직하게는, 상기 DNA 라이브러리는 차세대 염기서열해독(Next generation sequencing, NGS) 분석을 위한 것이고, 보다 바람직하게는, 상기 NGS 분석은 일루미나(Illumina) 플랫폼 HiSeq 시퀀서를 이용하는 것일 수 있으나, 이에 제한되는 것은 아니다.
- [0019] 또한, 본 발명은 상기 조성물을 포함하는 SINE 표적용 DNA 라이브러리 제작용 키트를 제공한다.
- [0020] 한편, 상기 키트는 SINE 표적용 DNA 라이브러리 제작에 일반적으로 사용되는 다른 기구, 용액 등을 추가로 포함할 수 있으나, 이에 제한되지 않는다.
- [0021] 또한, 본 발명은 게노믹(genomic) DNA를 절편화하는 단계; 상기 절편화된 DNA를 어댑터 접합시켜 양 말단을 평활화(end repair)시키는 단계; 상기 평활화된 DNA와, 서열번호 3으로 표시되는 이동성 유전인자 SINE을 표적하는 탐침자(probe)를 혼성화시키는 단계; 상기 혼성화된 DNA를 연장화시키는 단계; 및 상기 연장화된 DNA 단편 양 말단의 어댑터 서열을 이용하여 PCR 증폭시키는 단계를 포함하는 SINE 표적용 DNA 라이브러리 제작 방법을 제공한다.
- [0022] 바람직하게는, 상기 SINE은 *AluYb8* 또는 *AluYb9*일 수 있다.
- [0023] 바람직하게는, 상기 DNA 라이브러리는 차세대 염기서열해독(Next generation sequencing, NGS) 분석을 위한 것이고, 보다 바람직하게는, 상기 NGS 분석은 일루미나(Illumina) 플랫폼 HiSeq 시퀀서를 이용하는 것일 수 있으나, 이에 제한되는 것은 아니다.
- [0024] 본 발명에서 사용한 서열번호 3으로 표시되는 이동성 유전인자 SINE을 표적하는 탐침자(probe)는 서열번호 1로 표시되는 Illumina HiSeq system-based NuGen adaptor sequence(P7 어댑터 서열)과, 서열번호 2로 표시되는 *AluYb8/9*-표적 서열로 이루어져 있다.
- [0025] 본 발명의 SINE 표적용 DNA 라이브러리는 일루미나(Illumina) 플랫폼 장비인 HiSeq 시퀀서에 적용될 수 있으며, HiSeq에 사용되는 어댑터 서열은 P5 어댑터 서열과 P7 어댑터 서열 두 가지로 나누어지는데, P5와 P7 어댑터 서열은 방향성에 따라서 나뉜다. 보통 P7 어댑터가 주형가닥으로 부착되었을 때 첫 번째 염기분석을 시작하는 위치의 어댑터이고 P5 어댑터는 주형가닥 전환을 통해 상보적인 반대편 서열을 읽기 시작할 수 있는 어댑터이다. 이 두 어댑터는 모두 HiSeq 장비에 사용되는 플로우셀(Flow Cell) 표면에 심어져 있는 염기서열과 상보적이므로 플로우셀 내에서 혼성화 하여 증폭함으로써 라이브러리가 플로우셀에 가득차게 된다. P5 어댑터 서열과 P7 어댑터 서열은 상업적으로 공개된 서열을 사용할 수 있다.
- [0026] 이하에서는, 본 발명을 한정하지 않는 실시예에 따라 본 발명을 상세히 설명한다. 본 발명의 하기 실시예는 본 발명을 구체화하기 위한 것일 뿐 본 발명의 권리범위를 제한하거나 한정하는 것이 아님은 물론이다. 따라서, 본 발명의 상세한 설명 및 실시예로부터 본 발명이 속하는 기술분야의 전문가가 용이하게 유추할 수 있는 것은 본 발명의 권리범위에 속하는 것으로 해석된다.
- [0027] <실시예>
- [0028] 본 발명은 인간 유전체에 특이적으로 존재한다고 알려진 *AluYb8/9*를 표적하는 탐침자(Probe)를 제작하여 개인 유전체 내에 개별적으로 가지고 있는 이동성 유전인자 *AluYb8/9*에 의한 구조적 변이를 동정하는 것이다. 이는 특정 집단연구를 통하여 정확도 높은 진단마커를 개발할 수 있다.
- [0029] - 단계 1 -
- [0030] 이동성 유전인자 *AluYb8/9*의 염기서열을 인간 참조 유전체 데이터베이스인 UCSC genome 데이터베이스로부터 모두 검출한 뒤 다른 아종(Subfamily)의 *Alu*들과 다중염기서열정렬법을 통하여 *AluYb8/9*에 특이적으로 가지고 있는 염기 서열 부위를 확인하여 목표 서열 탐침자를 제작할 위치를 선정하였다.
- [0031] *AluYb8/9* 표적 탐침자는 *AluYb8/9* 인자의 3' 비번역영역에 위치한 서열을 포함하며 HiSeq (Illumina사 플랫폼 장비) 차세대 염기서열해독 장비에 적용할 수 있는 해독 어댑터 서열, 본 2가지 서열로 구성하였다.
- [0032] 본 발명의 실시예에 따른 이동성 유전인자인 *AluYb8/9* 표적 염기서열을 추적하기 위해 제작된 탐침자(Probe)의 표적서열은 표 1에 나타냈다.

표 1

[0033]	Usage	염기서열 (5'-Sequence-3')
	Illumina HiSeq system-based NuGen adaptor sequence	5'-CAAGCAGAAGACGGCATACGAGATGGAGTTCAGACGTGTGCTCTTCCGATCT GCAAGGATCGCTCTC-3' (서열번호 1)
	AluYb8/9-targeted sequence	5'-TGCAGTGAGCCGAGATTGCGCCACTGCAGTCCGAGTCCG-3' (서열번호 2)
	제작된 AluYb8/9 표적 탐침자 염기서열 (AluYb8/9-targeted Probe)	5'-CAAGCAGAAGACGGCATACGAGATGGAGTTCAGACGTGTGCTCTTCCGATCT GCAAGGATCGCTCTC TGCAGTGAGCCGAGATTGCGCCACTGCAGTCCGAGTCCG-3' (서열번호 3)

[0034] - 단계 2 -

[0035] 차세대 염기서열해독에 적합한 형태로 유전물질을 절편화 하기 위해 초음파 DNA Sonication 시스템을 사용하여 약 550 bp 크기로 절단한다. HiSeq 염기서열 해독장치에 사용할 수 있는 어댑터를 접합하기 위해 점착성 말단 (Sticky end) 형태의 단편화된 유전물질을 End Repair 과정을 통해 평활 말단(Blunt end) 형태로 복원해 준다 (25°C (30 분) 1 사이클 그리고 70°C (10 분)). HiSeq에 사용되는 어댑터 서열은 P5 어댑터 서열과 P7 어댑터 서열 두 가지로 나누어지는데, P5와 P7 어댑터 서열은 방향성에 따라서 나뉜다. 보통 P7 어댑터가 주형가닥으로 부착되었을 때 첫 번째 염기분석을 시작하는 위치의 어댑터이고 P5 어댑터는 주형가닥 전환을 통해 상보적인 반대편 서열을 읽기 시작할 수 있는 어댑터이다. 이 두 어댑터는 모두 HiSeq 장비에 사용되는 플로우셀(Flow Cell) 표면에 심어져 있는 염기서열과 상보적이므로 플로우셀 내에서 혼성화 하여 증폭함으로써 라이브러리가 플로우셀에 가득차게 된다. 차세대 염기서열해독 장비인 HiSeq을 사용하기 위해 어댑터 서열을 절편화된 유전물질 양 말단에 접합한다(25°C (30 분) 1 사이클 그리고 70°C (10 분)).

[0036] - 단계 3 -

[0037] P5 어댑터 서열을 절편화된 유전물질 양 말단에 붙이는 과정 이후 AluYb8/9 표적 탐침자 서열(Illumina P7 어댑터 서열과 함께 구성됨)을 사용해 유전물질 단편 안에 존재하는 AluYb8/9 서열과 유전물질간 혼성화(DNA hybridization) 과정을 거쳐 두 서열을 붙여준다. 이 과정은 P5 어댑터가 접합된 유전물질 절편 500ng과 AluYb8/9 표적 탐침자 100 pM을 사용하여 증합효소 연쇄반응(PCR, Polymerase chain reaction) 기기에서 95°C (5 분) 1사이클, 80°C(10 초, 매 사이클 -0.1°C 차감) 200 사이클, 60°C(16 시간) 동안 유전물질간 혼성화 반응을 진행한다.

[0038] - 단계 4 -

[0039] 유전물질 혼성화 반응이 끝난 후 혼성화된 유전물질 단편들을 연장화 작업(72°C (10분)) 이후 증합효소 연쇄반응(Polymerase Chain Reaction, PCR)을 통해 증폭한다. 이 과정에서 P5와 P7 어댑터에 상응하는 프라이머를 사용하여 증합효소 연쇄반응을 진행한다. 탐침자가 표적하는 AluYb8/9를 포함하는 유전물질 절편서열이 접합되어 있는 경우에만 선택적으로 증폭반응이 일어나게 된다(37°C (10분), 95°C(3 분) 1 사이클, [95°C(30 초), 62°C (15 초), 72°C(20 초) 총 15 사이클], 72°C(3 분) 1 사이클, 그리고 4°C (보관)). Illumina사 P5와 P7 프라이머를 이용하여 증폭된 유전물질 절편 시료를 라이브러리라고 하며, 이 라이브러리를 사용해 차세대 염기서열해독을 진행한다.

[0040] 본 발명의 실시예에 따른 AluYb8/9 표적 염기서열해독을 위한 라이브러리 구축을 위한 실험적 단계에 적용된 유전물질 절편화 준비 및 표적서열 증폭조건은 표 2에 나타냈다.

표 2

DNA Fragmentation(Covaris S2 system (Covaris, Woburn, MA, USA))		
Ultrasonication	Duty cycle	10%
	Intensity	2
	Cycles per Burst	200
	Duration (seconds)	45
	Water bath temperature	4°C
DNA end repair (PCR and End Repair Master Mix (NuGen, San Carlos, USA))		

End repair	25°C	30 minutes	
Enzyme deactivation	70°C	10 minutes	
Adaptor ligation (PCR and Ligation Master Mix (NuGen, San Carlos, USA))			
Adaptor ligation	25°C	30 minutes	
Enzyme deactivation	70°C	10 minutes	
<i>AluYb8/9</i> -targeted probe hybridization (PCR)			
Denaturation	95°C	5 minutes	
Probe hybridization	80°C (-0.1°C) 200 cycle	10 seconds in each cycle	
	60°C	16 hours	
Library amplification (PCR and Target Enrichment Mix (NuGen, San Carlos, USA))			
Initialization	37°C	10 minutes	
Pre-denaturation	95°C	3 minutes	
Denaturation	95°C	30 seconds	15 cycles
Annealing	62°C	15 seconds	
Extension	72°C	20 seconds	
Final extension	72°C	3 min	
Storage	4°C	Hold	

[0042] - 단계 5 -

[0043] *AluYb8/9* 표적서열이 농축된 라이브러리를 플로우셀에서 증폭시키고 HiSeq 장비를 이용하여 염기서열해독을 진행한다. 시퀀싱 프라이머는 개시영역에 붙어서 표적서열의 반대쪽 말단부터 한 염기서열씩 합성해 나가간다. 이 때, 각 A, T, G, C 염기서열은 고유의 파장을 갖는 올리고 뉴클레오타이드(Oligo Nucleotide)로 합성되어 Sequencing By Synthesis 방법으로 진행된다. 각 파장은 HiSeq 내부에 존재하는 검출카메라를 이용하여 파장을 측정하여 염기서열해독을 결정한다.

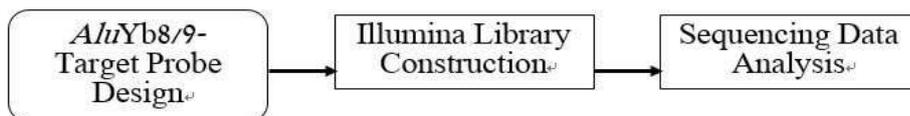
[0044] - 단계 6 -

[0045] 차세대 염기서열해독이 종료된 후 염기서열을 결정해주고 bcl파일로 CASAVA라는 2차 분석 장비로 넘겨준다. 이 때 Demultiplexing이라는 작업을 통해 각 샘플 별로 분류가 가능하며 fastq 파일을 생성해주며 Illumina에서 정의된 quality filters를 통해 Raw 데이터를 생성한다. Illumina 어댑터는 *AluYb8/9* 표적 서열이 아니기 때문에 cutadapt 1.1 version (<https://cutadapt.readthedocs.io/en/stable>) 생명정보학 분석 도구를 이용하여 제거한다. *AluYb8/9* 표적 탐침자 서열을 제거한 게놈서열(genomic sequence)을 Bowtie (<http://bowtie-bio.sourceforge.net>) aligner 프로그램을 이용하여 인간 참조 유전체 서열에 맵핑(Mapping)을 진행하여 유사한 유전체 서열 영역에 정렬하는 단계를 진행한다. 정렬된 유전체 서열 내에 유의한 맵핑 정도를 보이는 표적 영역을 표시하기 위해 개발된 HOMER와 MACS2 두 개의 서로 다른 프로그램을 이용하여 공통적으로 검출한 표적 영역을 *AluYb8/9* 삽입영역으로 선별한다.

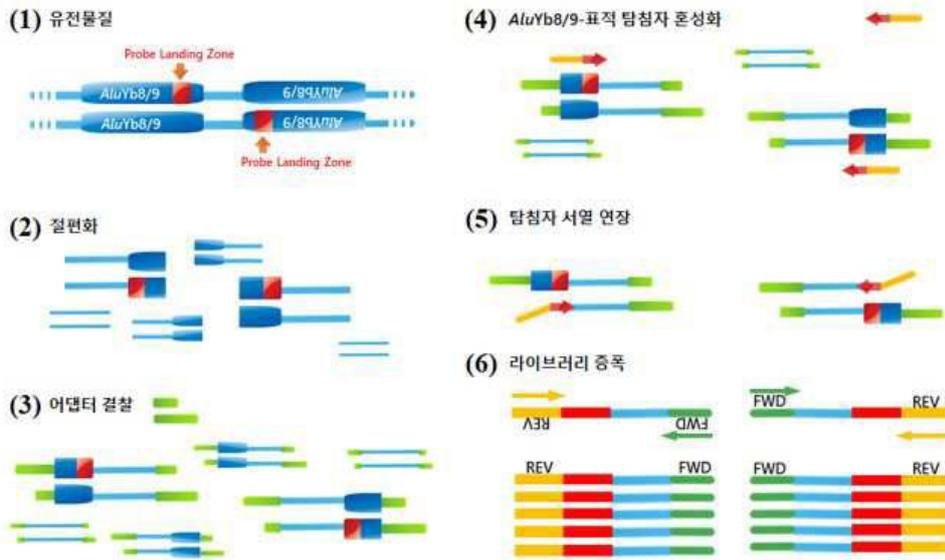
[0046] 이 발명을 통해 개인 유전체 내 새롭게 삽입된 인간에 특이적인 *AluYb8/9* 유전인자 영역을 쉽고 빠르게 스크리닝 가능하며, 저렴한 가격으로 개인 유전체 내 *AluYb8/9*에 의한 유전학적 변이를 확인할 수 있다. 자동화 분석 방법을 이용해 임상의학 및 기초과학 연구자가 손쉽게 분석이 가능하며 *AluYb8/9* 유전인자에 국한된 변이 분석이 아닌 다른 목적 인자를 위한 탐침자 제작에 다양하게 이용 가능하다.

도면

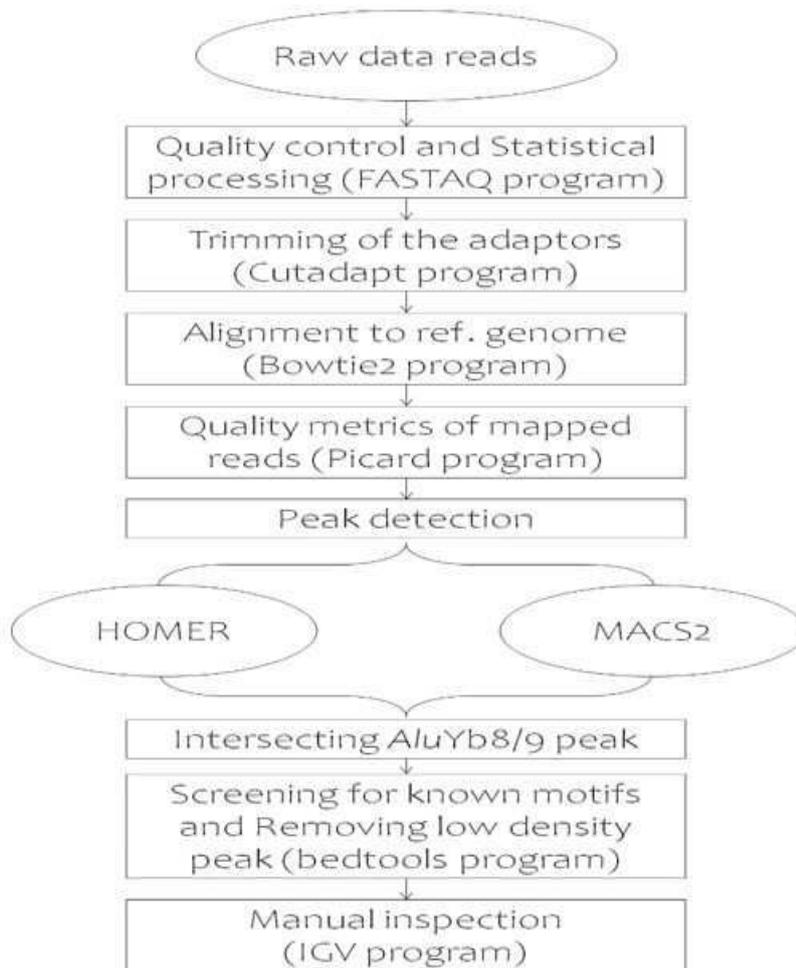
도면1



도면2



도면3



서열 목록

<110> Dankook University Cheonan Campus Industry Academic Cooperation Foundation

<120> Method for preparing DNA library for SINE targeted-probe enrichment using HiSeq sequencer

<130> ADP-2019-0123

<160> 3

<170> KopatentIn 2.0

<210> 1

<211> 67

<212> DNA

<213> Artificial Sequence

<220><223> Illumina HiSeq system-based NuGen adaptor sequence

<400> 1

caagcagaag acggcatacg agatggagtt cagacgtgtg ctcttccgat ctgcaaggat 60

cgctctc 67

<210> 2

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> AluYb8/9-targeted sequence

<400> 2

tgcagtgagc cgagattgcg cactgcagtt ccgcagtcgg 40

<210> 3

<211> 107

<212> DNA

<213> Artificial Sequence

<220><223> AluYb8/9-targeted Probe

<400> 3

caagcagaag acggcatacg agatggagtt cagacgtgtg ctcttccgat ctgcaaggat 60

cgctctctgc agtgagccga gattgcgcca ctgcagtcgg cagtcgg 107