

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0025430

(43) 공개일자 2020년03월10일

(51) 국제특허분류(Int. Cl.)

G16B 20/00 (2019.01) G16B 5/00 (2019.01)

G16B 50/00 (2019.01)

(52) CPC특허분류

G16B 20/00 (2019.02)

G16B 5/00 (2019.02)

(21) 출원번호 10-2018-0102771

(22) 출원일자 2018년08월30일

심사청구일자 2018년08월30일

(71) 출원인

한동대학교 산학협력단

경상북도 포항시 북구 흥해읍 한동로 558

(72) 발명자

안태진

서울특별시 서초구 서초중앙로29길 28, 306동 210호(반포동, 반포미도아파트)

반하민

경상북도 포항시 북구 흥해읍 한동로 558

(74) 대리인

김정훈

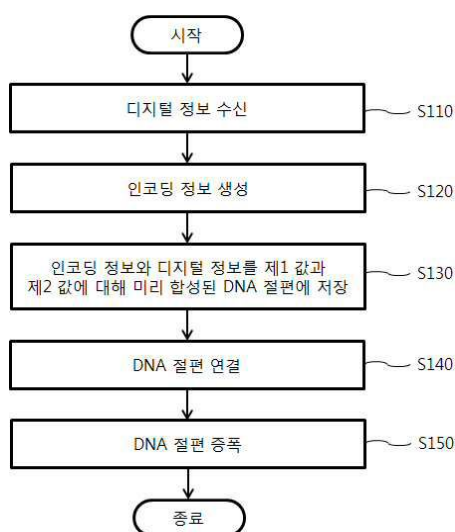
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 디지털 정보를 DNA 분자에 저장하는 방법 및 그 장치

(57) 요약

디지털 정보를 DNA 분자에 저장하는 방법 및 그 장치가 개시된다. 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 방법은 디지털 정보를 수신하는 단계; 및 디지털 데이터의 제1 값과 제2 값에 대응하는 미리 합성된 일정 길이의 DNA 절편(fragment)을 이용하여 상기 수신된 디지털 정보를 저장하는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G16B 50/00 (2019.02)

이 발명을 지원한 국가연구개발사업

과제고유번호 2017R1E1A2A0178518

부처명 과학기술정보통신부

연구관리전문기관 한국연구재단

연구사업명 개인기초연구(미래부)

연구과제명 하나의 DNA 분자에 정보를 기록하고 비파괴적으로 읽을 수 있는 방법

기 여 율 1/1

주관기관 한동대학교

연구기간 2017.11.01 ~ 2018.10.31

명세서

청구범위

청구항 1

디지털 정보를 수신하는 단계; 및

디지털 데이터의 제1 값과 제2 값에 대응하는 미리 합성된 일정 길이의 DNA 절편(fragment)을 이용하여 상기 수신된 디지털 정보를 저장하는 단계

를 포함하는 DNA 분자 기반 정보 저장 방법.

청구항 2

제1항에 있어서,

상기 제1 값과 제2 값에 대응하는 DNA 절편을 합성하는 단계

를 더 포함하고,

상기 DNA 절편은

상기 DNA 절편의 전기적 속성이 달라지도록 상기 제1 값과 제2 값 각각에 대응하는 미리 설정된 서열 패턴을 포함하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 방법.

청구항 3

제1항에 있어서,

상기 저장하는 단계는

루프 구조를 포함하는 프라이머(primer)를 사용하여 상기 DNA 절편을 연결함으로써, 상기 DNA 절편을 이용하여 상기 디지털 정보를 저장하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 방법.

청구항 4

제1항에 있어서,

상기 저장하는 단계는

상기 DNA 절편의 DNA 서열이 가지는 전기적 특성을 이용하여 상기 디지털 정보를 저장하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 방법.

청구항 5

제1항에 있어서,

상기 DNA 절편은

상기 DNA 절편의 시작을 나타내는 스타트(start) 서열 패턴, 상기 제1 값과 상기 제2 값 중 어느 하나의 값에 대응하는 값 서열 패턴, 상기 DNA 절편의 끝을 나타내는 엔드(end) 서열 패턴 및 리게이션 어세이(ligation assay)를 위한 링커 서열 패턴을 포함하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 방법.

청구항 6

제1항에 있어서,

상기 저장하는 단계는

디코딩을 위한 인코딩 정보를 상기 DNA 절편에 추가적으로 저장하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 방법.

청구항 7

제1항에 있어서,

상기 디지털 정보가 저장된 DNA 절편을 증폭 프라이머를 이용하여 증폭하는 단계

를 더 포함하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 방법.

청구항 8

디지털 정보를 수신하는 수신부; 및

디지털 데이터의 제1 값과 제2 값에 대응하는 미리 합성된 일정 길이의 DNA 절편(fragment)을 이용하여 상기 수신된 디지털 정보를 저장하는 저장부

를 포함하는 DNA 분자 기반 정보 저장 장치.

청구항 9

제8항에 있어서,

상기 제1 값과 제2 값에 대응하는 DNA 절편을 합성하는 합성부

를 더 포함하고,

상기 DNA 절편은

상기 DNA 절편의 전기적 속성이 달라지도록 상기 제1 값과 제2 값 각각에 대응하는 미리 설정된 서열 패턴을 포함하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 장치.

청구항 10

제8항에 있어서,

상기 저장부는

루프 구조를 포함하는 프라이머(primer)를 사용하여 상기 DNA 절편을 연결함으로써, 상기 DNA 절편을 이용하여 상기 디지털 정보를 저장하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 장치.

청구항 11

제8항에 있어서,

상기 저장부는

상기 DNA 절편의 DNA 서열이 가지는 전기적 특성을 이용하여 상기 디지털 정보를 저장하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 장치.

청구항 12

제8항에 있어서,

상기 DNA 절편은

상기 DNA 절편의 시작을 나타내는 스타트(start) 서열 패턴, 상기 제1 값과 상기 제2 값 중 어느 하나의 값에 대응하는 값 서열 패턴, 상기 DNA 절편의 끝을 나타내는 엔드(end) 서열 패턴 및 리게이션 어세이(ligation assay)를 위한 링커 서열 패턴을 포함하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 장치.

청구항 13

제8항에 있어서,

상기 저장부는

디코딩을 위한 인코딩 정보를 상기 DNA 절편에 추가적으로 저장하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 장치.

청구항 14

제8항에 있어서,

상기 디지털 정보가 저장된 DNA 절편을 증폭 프라이머를 이용하여 증폭하는 증폭부

를 더 포함하는 것을 특징으로 하는 DNA 분자 기반 정보 저장 장치.

발명의 설명

기술 분야

[0001]

본 발명은 DNA 분자 기반 정보 저장 기술에 관한 것으로서, 보다 상세하게는 디지털 데이터 값에 대응하여 미리 설정된 서열 패턴을 포함하는 DNA 절편(fragment)을 이용하여 디지털 정보를 DNA 분자에 저장할 수 있는 DNA 분자 기반 정보 저장 방법 및 그 장치에 관한 것이다.

배경 기술

[0003]

2020년이면 전세계 데이터량이 40 ZB에 이를 것으로 예측되며, 이미 2007년부터 전세계에서 생산되는 데이터량이 저장장치를 초과하여 새로운 방식의 저장장치가 필요하게 되었다. 장기 기록매체로서 널리 사용되고 있는 자기 테이프의 경우 데이터 저장 수명이 10년 정도로 제한되어 유지 및 관리비용이 지속적으로 요구 되고 있다. 반도체 저장장치의 경우 HDD와 SSD가 대표적인데, HDD의 경우 충격에 매우 취약하다는 점과 수명이 25만 시간, 그리고 최대 용량치에 한계가 있으며, SSD의 경우 충격에는 강하지만 수명이 상대적으로 HDD보다 짧다고 알려져 있다.

[0004]

기존 정보 저장 매체가 직면한 데이터 저장밀도 한계를 해결 가능한 대체 기술로서 DNA 염기서열이 초대용량 정보의 장기 보존 가능한 스토리지 매체로 거론되었고, DNA 구조로 초대용량 데이터를 저장하는 새로운 기술의 가능성을 발견하였다. DNA는 잘 알려져 있듯이, 생물체의 가장 작은 단위인 세포 안에 들어있으며, 모든 유전정보를 담고 있다. DNA가 가지고 있는 정보에 따라 마치 모든 생물체는 프로그램 되어있는 대로 움직인다. 인간의 경우, 1개의 단일 세포에 들어있는 DNA는 30억쌍의 염기 서열로 구성되어 있고, 이를 모두 해독한 유전정보의 크기를 환산한다면 약 1TB정도의 용량이다. 그리고 1개의 단일세포에는 폭이 2nm, 길이가 3m나 되는 두 가닥의 DNA가 들어있다. 따라서 이론적으로 $EB(10^{18})$ 이상 저장할 수 있는 차세대 바이오스토리지인 DNA는 초집약적으로 정보를 저장하기 위한 바이오소재로서 매우 적합하다. 저장 수명도 1,000년 이상이며, 저비용으로 저장이 용이

할 것으로 보인다.

- [0005] 최근에는 폭발적으로 생산되는 데이터량이 저장매체의 용량을 초과하여 과부하를 일으키고 있는 실정이고, 이로 인해 새로운 저장매체를 개발하기 위한 시도가 계속해서 이루어지고 있다. 그 중 DNA를 이용하여 새로운 저장매체 개발의 시도는 매우 흥미롭다. DNA를 저장매체로 이용할 경우, 기존의 저장매체의 단점인 데이터 저장 밀도를 뛰어 넘을 수 있고, 물리적인 충격에도 안정적으로 정보를 저장할 수 있다.
- [0006] 종래 DNA에 정보를 저장하는 일 실시예 기술은 DNA 분자 중 A, T, G, C 염기 부분에 정보를 저장하는 방법을 사용하는데, 종래 기술은 원하는 정보를 저장하기 위해서 정보를 표현하기 위한 특정한 서열을 정확하게 합성해야 하는 조건이 요구되며, 합성의 정확도를 담보할 수 없기 때문에 동일한 정보를 많은 수의 DNA 분자에 합성하고, 또한 긴 정보는 작게 나누어 서로 중첩되도록 합성하여 정보를 기록한다.
- [0007] 이러한 종래 기술은 실질적으로 많은 수의 분자들을 사용하여 정보를 저장하고, 결국은 많은 수의 분자들에서 투표를 하듯이 정보를 복원해야 하기 때문에 하나의 분자를 사용한 정확한 정보의 저장 방법이라고 할 수 없다. 이러한 접근을 하게 되는 근본적인 이유는 올리고뉴클레오타이드(oligonucleotide) 합성 정확도에 한계가 있기 때문이다.

발명의 내용

해결하려는 과제

- [0009] 본 발명의 실시예들은, 디지털 데이터 값에 대응하여 미리 설정된 서열 패턴을 포함하는 DNA 절편(fragment)을 이용하여 디지털 정보를 DNA 분자에 저장할 수 있는 DNA 분자 기반 정보 저장 방법 및 그 장치를 제공한다.

과제의 해결 수단

- [0011] 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 방법은 디지털 정보를 수신하는 단계; 및 디지털 데이터의 제1 값과 제2 값에 대응하는 미리 합성된 일정 길이의 DNA 절편(fragment)을 이용하여 상기 수신된 디지털 정보를 저장하는 단계를 포함한다.
- [0012] 나아가, 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 방법은 상기 제1 값과 제2 값에 대응하는 DNA 절편을 합성하는 단계를 더 포함하고, 상기 DNA 절편은 상기 DNA 절편의 전기적 속성이 달라지도록 상기 제1 값과 제2 값 각각에 대응하는 미리 설정된 서열 패턴을 포함할 수 있다.
- [0013] 상기 저장하는 단계는 루프 구조를 포함하는 프라이머(primer)를 사용하여 상기 DNA 절편을 연결함으로써, 상기 DNA 절편을 이용하여 상기 디지털 정보를 저장할 수 있다.
- [0014] 상기 저장하는 단계는 상기 DNA 절편의 DNA 서열이 가지는 전기적 특성을 이용하여 상기 디지털 정보를 저장할 수 있다.
- [0015] 상기 DNA 절편은 상기 DNA 절편의 시작을 나타내는 스타트(start) 서열 패턴, 상기 제1 값과 상기 제2 값 중 어느 하나의 값에 대응하는 값 서열 패턴, 상기 DNA 절편의 끝을 나타내는 엔드(end) 서열 패턴 및 리게이션 어레이(ligation assay)를 위한 링커 서열 패턴을 포함할 수 있다.
- [0016] 상기 저장하는 단계는 디코딩을 위한 인코딩 정보를 상기 DNA 절편에 추가적으로 저장할 수 있다.
- [0017] 더 나아가, 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 방법은 상기 디지털 정보가 저장된 DNA 절편을 증폭 프라이머를 이용하여 증폭하는 단계를 더 포함할 수 있다.
- [0019] 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 장치는 디지털 정보를 수신하는 수신부; 및 디지털 데이터의 제1 값과 제2 값에 대응하는 미리 합성된 일정 길이의 DNA 절편(fragment)을 이용하여 상기 수신된 디지털 정보를 저장하는 저장부를 포함한다.
- [0020] 나아가, 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 장치는 상기 제1 값과 제2 값에 대응하는 DNA 절편을 합성하는 합성부를 더 포함하고, 상기 DNA 절편은 상기 DNA 절편의 전기적 속성이 달라지도록 상기 제1 값

과 제2 값 각각에 대응하는 미리 설정된 서열 패턴을 포함할 수 있다.

- [0021] 상기 저장부는 루프 구조를 포함하는 프라이머(primer)를 사용하여 상기 DNA 절편을 연결함으로써, 상기 DNA 절편을 이용하여 상기 디지털 정보를 저장할 수 있다.
- [0022] 상기 저장부는 상기 DNA 절편의 DNA 서열이 가지는 전기적 특성을 이용하여 상기 디지털 정보를 저장할 수 있다.
- [0023] 상기 DNA 절편은 상기 DNA 절편의 시작을 나타내는 스타트(start) 서열 패턴, 상기 제1 값과 상기 제2 값 중 어느 하나의 값에 대응하는 값 서열 패턴, 상기 DNA 절편의 끝을 나타내는 엔드(end) 서열 패턴 및 리게이션 어레이(ligation assay)를 위한 링커 서열 패턴을 포함할 수 있다.
- [0024] 상기 저장부는 디코딩을 위한 인코딩 정보를 상기 DNA 절편에 추가적으로 저장할 수 있다.
- [0025] 더 나아가, 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 장치는 상기 디지털 정보가 저장된 DNA 절편을 증폭 프라이머를 이용하여 증폭하는 증폭부를 더 포함할 수 있다.

발명의 효과

- [0027] 본 발명의 실시예들에 따르면, 디지털 데이터 값에 대응하여 미리 설정된 서열 패턴을 포함하는 DNA 절편을 이용하여 디지털 정보를 DNA 분자에 저장할 수 있다.
- [0028] 본 발명의 실시예에 따르면, DNA 절편을 이용하여 디지털 정보를 저장할 수 있기 때문에 자기 디스크를 보완하거나 대체할 정보 저장 매체로 활용할 수 있으며, 화재나 재해에 대비하여 중요 정보를 전기소모 없이 적은 공간에 안전하게 저장할 수 있다.
- [0029] 본 발명의 실시예에 따르면, 디지털 정보를 DNA 절편을 이용하여 저장하기 때문에 합성 시 발생할 수 있는 에러에 강건하고, 시퀀싱 장비 예를 들어, 나노포어(nanopore) 시퀀싱 장비를 이용하여 DNA 분자에 저장된 디지털 정보를 읽을 수 있다.
- [0030] 이러한 본 발명은 자기를 복제하는 DNA 분자의 특성을 이용하여 다양한 경우의 수를 계산하는 분자 컴퓨팅 분야, 정보를 저장하도록 합성한 DNA를 증폭하여 보관하거나 박테리아 등에 넣어서 보관 하는 등의 바이오인포매틱스(Bioinformatics) 분야, 개량된 유전체를 보유한 바이오 균주 또는 바이러스 식별을 위한 디지털 정보 삽입, 정보 백업용 스토리지, 대용량 자료의 장기간 보관 서비스 등의 다양한 분야에 적용할 수 있다.

도면의 간단한 설명

- [0032] 도 1은 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 방법에 대한 동작 흐름도를 나타낸 것이다.
- 도 2는 기존 방법과 본 발명에 따른 방법을 비교하기 일 예시도를 나타낸 것이다.
- 도 3은 DNA 절편을 연결하는 과정을 설명하기 위한 일 예시도를 나타낸 것이다.
- 도 4는 허프만 코드를 이용한 디지털 정보를 DNA 분자에 저장하는 일 예시도를 나타낸 것이다.
- 도 5는 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 장치에 대한 구성을 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

- [0033] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나, 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다.
- [0034] 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며, 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성요소, 단계, 동작 및/또는 소자는 하나 이상 의 다

른 구성요소, 단계, 동작 및/또는 소자의 존재 또는 추가를 배제하지 않는다.

- [0035] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또한, 일반적으로 사용되는 사 전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0036] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예들을 보다 상세하게 설명하고자 한다. 도면 상의 동일한 구성요소에 대해서는 동일한 참조 부호를 사용하고 동일한 구성요소에 대해서 중복된 설명은 생략한다.
- [0038] DNA를 정보 저장 매체로 사용하는 연구는 계속 시도되어 왔으며, 자연계에 존재하는 원리를 재사용 한다는 측면에서 상징적인 의미가 있다. 화재나 재해에 대비하여 중요 정보를 전기소모 없이 적은 공간에 안전하게 저장한다는 개념으로 우선적으로 활용될 수 있으며, 나아가 통상적으로 활용되는 자기 디스크를 보완하거나 대체할 정보 저장 매체로 활용할 수 있다. 또한 자기를 복제하는 DNA 분자의 특성을 이용하여 다양한 경우의 수를 계산하는 분자 컴퓨팅 분야에도 나아가 활용할 수 있다
- [0039] 이러한 DNA 기반 정보 저장 분야는 바이오인포매틱스(Bioinformatics)의 한 분야로, 정보를 저장하도록 합성한 DNA를 증폭하여 보관하거나, 박테리아 등에 넣어서 보관 하는 등의 생물학적 시도를 했던 나라에서 주도하고 있다.
- [0040] 또한, 현존하는 시퀀싱 기술은 반도체 기술이 활용되는 실리콘 산업에 기반하여 제작되므로, 실리콘 기판 위에 제작되는 나노포어를 활용한 정보 저장 및 관독이 반도체 산업과도 연계될 수도 있다.
- [0042] 본 발명의 실시예들은, 디지털 데이터 값에 대응하여 미리 설정된 서열 패턴을 포함하는 DNA 절편(fragment)을 이용하여 디지털 정보를 DNA 분자에 저장하는 것으로, 자기 디스크를 보완하거나 대체할 정보 저장 매체로 활용할 수 있으며, 화재나 재해에 대비하여 중요 정보를 전기소모 없이 적은 공간에 안전하게 저장할 수 있으며, 디지털 정보를 DNA 절편을 이용하여 저장하기 때문에 합성 시 발생할 수 있는 에러에 강건하고, 시퀀싱 장비 예를 들어, 나노포어(nanopore) 시퀀싱 장비를 이용하여 DNA 분자에 저장된 디지털 정보를 읽을 수 있다.
- [0043] 여기서, 본 발명은 디지털 데이터의 제1 값 예를 들어, 1과 제2 값 예를 들어, 0에 대하여 미리 합성된 일정 길이의 DNA 절편을 이용하여 디지털 정보를 저장할 수 있으며, DNA 절편은 제1 값 또는 제2 값에 대응하는 값 서열 패턴을 포함하고, 나아가 DNA 절편의 시작을 나타내는 스타트(start) 서열 패턴, DNA 절편의 끝을 나타내는 엔드(end) 서열 패턴 및 리게이션 어세이(ligation assay)를 위한 링커 서열 패턴을 더 포함할 수 있다.
- [0044] 링커 서열 패턴은 값 서열 패턴의 좌우측 또는 앞뒤에 배치될 수 있으며, 시작 서열 패턴과 엔드 서열 패턴은 DNA 절편의 시작과 끝에 배치될 수 있다.
- [0045] 나아가, DNA 절편들은 루프 구조를 포함하는 프라이머(primer)에 의해 연결될 수 있으며, 이렇게 연결된 DNA 절편들에 디지털 정보가 저장될 수 있다.
- [0046] 또한, 본 발명은 DNA 절편에 저장되는 디지털 정보에 대하여, 허프만 코드에 대응하는 디지털 정보를 저장할 수 있으며, 이러한 인코딩에 대한 정보 즉, 인코딩 정보를 DNA 절편에 추가로 함께 저장할 수 있다.
- [0047] 여기서, 인코딩 정보는 DNA 분자에서 디지털 정보가 시작된다는 시그널(DNA서열 패턴), 보관할 디지털 정보의 총량, 디지털 정보를 디코딩할 때 필요한 정보, DNA 분자에서 디지털 정보 기록한 부분이 끝난다는 시그널(DNA 서열 패턴), 정보 기록 날짜, 정보 기록자 등 부차적인 정보를 포함할 수 있는데, 인코딩 정보에 포함되는 정보는 상술한 정보로 한정되지 않으며, 본 발명에서 디지털 정보를 인코딩하고 디코딩 시 필요로 하는 정보가 모두 포함될 수 있다.
- [0048] 더 나아가, 본 발명은 디지털 정보를 DNA 절편들을 포함하는 DNA 분자에 저장한 후 디지털 정보가 저장된 DNA 절편을 증폭 프라이머를 이용하여 증폭할 수도 있다.
- [0049] 이러한 본 발명은 제1 값과 제2 값에 대응하는 값 서열 패턴을 포함하는 DNA 절편을 미리 합성하고, DNA 절편에 포함된 값 서열 패턴은 전기적 속성이 상이할 수 있다. 즉, 본 발명은 DNA 절편의 DNA 서열이 가지는 전기적 특성을 이용하여 제1 값 또는 제2 값의 디지털 정보를 DNA 분자에 저장할 수 있다.

- [0050] 이러한 본 발명에 대하여, 도 1 내지 도 5를 참조하여 설명한다.
- [0052] 도 1은 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 방법에 대한 동작 흐름도를 나타낸 것이다.
- [0053] 도 1을 참조하면, 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 방법은 디지털 정보를 수신하고, 수신된 디지털 정보를 디코딩할 때 사용하기 위한 인코딩 정보를 생성한다(S110, S120).
- [0054] 여기서, 단계 S110은 저장하고자 하는 디지털 정보를 인코딩한 방식에 의하여 다양한 형태의 디지털 데이터로 표현될 수 있다. 예를 들어, 단계 S110은 UTF8 인코딩 방식에 의해 인코딩되는 경우 한글 한글자당 24비트로 변환될 수 있고, 허프만(Huffman) 코드에 의해 인코딩되는 경우 한글 한글자당 최소 3비트에서 최대 7비트로 변환될 수 있다.
- [0055] 예를 들어, 단계 S110은 도 4a에 도시된 바와 같이 "나라의말이중국"이 호프만 코드에 의해 인코딩되면, "나"는 "0111100"의 7비트로 변환되고, "라"는 "100110"의 6비트로 변환되며, "의"는 "1110010"의 7비트로 변환되고, "말"은 "101100"의 6비트로 변환되며, "이"는 "0101"의 4비트로 변환되고, "중"은 "1110111"의 7비트로 변환되며, "국"은 "0111000"의 7비트로 변환됨으로써, 저장하고자 하는 정보 "나라의말이중국"에 대한 44비트의 디지털 정보를 수신할 수 있다.
- [0056] 단계 S120에서 생성되는 인코딩 정보는 DNA 분자에서 디지털 정보가 시작된다는 시그널(DNA서열 패턴) 예를 들어, 스타트 서열 패턴, 보관할 디지털 정보의 총량, 디지털 정보를 디코딩할 때 필요한 정보, DNA 분자에서 디지털 정보 기록한 부분이 끝난다는 시그널(DNA서열 패턴) 예를 들어, 엔드 서열 패턴, 정보 기록 날짜 및 정보 기록자 등의 부차적인 정보를 포함할 수 있으며, 허프만 코딩에 대한 정보를 포함할 수도 있다. 즉, 단계 S120은 DNA 분자에 디지털 정보를 저장할 때, 해당 분자에 기록할 정보 외에 복원 및 보관 무결성에 필요한 부차적 정보를 생성하는 단계일 수 있다.
- [0057] 단계 S110과 S120에 의해 저장하고자 하는 디지털 정보가 수신되고 인코딩 정보가 수신되면, 디지털 정보와 인코딩 정보를 제1 값과 제2 값에 대해 미리 합성된 DNA 절편에 저장하고, 이러한 DNA 절편을 연결한다(S130, S140).
- [0058] 여기서, 디지털 데이터의 제1 값과 제2 값은 DNA 절편의 전기적 속성이 달라지도록 미리 설정된 상이한 값 서열 패턴을 가질 수 있다.
- [0059] 예를 들어, 디지털 데이터 값 0은 "TTTTTTTTTTTT"의 서열 패턴을 가질 수도 있고, "TATT"의 서열 패턴을 가질 수도 있으며, 디지털 데이터 값 1은 "CCCGGGGCCCC"의 서열 패턴을 가질 수도 있고, "ACCC"의 서열 패턴을 가질 수도 있다. 물론, 이러한 디지털 데이터 값 각각의 서열 패턴은 해당 기술을 제공하는 사업자 또는 개인에 의해 결정될 수 있으며, DNA 분자에 저장된 디지털 정보를 감지 또는 읽기 위한 상이한 전기적 특성을 가지도록 미리 합성될 수 있다. 즉, 디지털 데이터의 제1 값과 제2 값에 대한 서열 패턴은 미리 합성된 상태로, 디지털 정보가 수신되면 수신된 디지털 정보와 인코딩 정보에 대한 디지털 데이터 비트 값에 대응하는 서열 패턴을 그대로 사용할 수 있다.
- [0060] 디지털 정보가 저장되는 DNA 절편은 DNA 절편의 시작을 나타내는 스타트(start) 서열 패턴, 제1 값과 상기 제2 값 중 어느 하나의 값에 대응하는 값 서열 패턴, DNA 절편의 끝을 나타내는 엔드(end) 서열 패턴 및 리게이션 어세이(ligation assay)를 위한 링커 서열 패턴을 포함할 수 있다. 예를 들어, 디지털 데이터 값 0의 서열 패턴이 "TTTTTTTTTTTT"인 경우 0의 값 서열 패턴과 링커들로 구성된 DNA 서열 절편은 "GAGGTGAGGTGAGGTGAGGTTTTTTTTTTAAGTGAAGTGAAGTGAAGT"로 표현될 수 있으며, 디지털 데이터 값 1의 서열 패턴이 "CCCGGGGCCCC"인 경우 1의 값 서열 패턴과 링커들로 구성된 DNA 서열 절편은 "GAGGTGAGGTGAGGTGAGGTCCCGGGGCCCCAAGTGAAGTGAAGTGAAGT"로 표현될 수 있다. 여기서, 링커들에 대한 서열 패턴은 랜덤 서열을 가질 수 있으며, 값 서열 패턴의 앞뒤 또는 좌측과 우측에 연결되는 서열일 수 있다. 물론, 값 서열 패턴과 링커 서열 패턴은 상황에 따라 달라질 수 있고, 그 길이 또한 상황에 따라 달라질 수 있다.
- [0061] 스타트 서열 패턴과 엔드 서열 패턴 또한 미리 설정될 수 있으며, 예를 들어, 스타트 서열 패턴은 35mer의 "CAGTCGCTCCACAAGTACCAGCCTCGTCTCCACAT"(start-5)일 수도 있고, 39mer의 "TCAGATGTGGAGACGAGGCTGGTACTGTGGAGCGACTG"(start-3)일 수도 있으며, 엔드 서열 패턴은 39mer의 "ATCGGGATATGATTGAGCAAGCAATGGCAAGATTGACGG"(end-5)일 수도 있고, 35mer의 "CCGTCATCTTGCCATTGCTTGCTCAATCATATCC"일 수도 있다. 물론, 본 발명에서의 스타트 서열 패턴과 엔드 서열 패턴

턴은 상황에 따라 달라질 수 있으며, 그 길이 또한 달라질 수 있다.

- [0062] 단계 S140에서 DNA 절편을 연결하는 과정은 DNA 절편을 연결하는 과정을 반복적으로 수행하여 원하는 양의 디지털 정보를 DNA 분자에 저장할 수 있는데, 도 3a에 도시된 바와 같이, 두 개의 DNA 절편을 순서대로 이어 붙여 제1 DNA 절편 그룹을 생성하고, 도 3b에 도시된 바와 같이 두 개 이상의 제1 절편 그룹을 원하는 순서대로 이어 붙여 더 큰 절편 그룹을 생성하며, 도 3c에 도시된 바와 같이 더 많은 절편을 그룹화하기 위하여 도 3a와 도 3b의 과정을 반복할 수 있는 속성을 부여한다.
- [0063] 여기서, 도 3a에서 사용되는 DNA 절편은 디지털 정보 0, 1을 표현하기 위한 서열과 리게이션 어레이가 가능하도록 하는 링커, 그리고 DNA 절편들을 순서대로 연결할 수 있는 스타트 어댑터(Start adapter) 및 엔드 어댑터(End adapter) 서열을 가질 수 있으며, 도 3b에서 절편 연결을 위한 리게이션 시퀀스는 루프 구조를 포함하는 프라이머를 사용할 수 있다.
- [0064] 추가적으로, 본 발명에 따른 방법은 디지털 정보가 저장되어 연결된 DNA 절편을 증폭 프라이머를 이용하여 증폭함으로써, 많은 수의 DNA 분자를 보관할 수 있도록 한다(S150).
- [0065] 여기서, 단계 S150은 증폭 프라이머를 이용하여 DNA 분자를 증폭하는 것으로 설명하였지만, 이에 한정되지 않으며 플라즈마 벡터에 넣을 미생물을 배양함으로써, DNA 분자를 증폭할 수도 있다.
- [0066] 이 때, 증폭 프라이머 서열은 미리 설정될 수 있으며, 예를 들어, 증폭 프라이머 서열은 "5' CAGTCGCTCCACAAGTACCAGCCTCGTCTCCACAT 3'"일 수도 있고, "5' CCGTCAATCTTGCCATTGCTTGCTCAATCATATCC 3'"일 수도 있다.
- [0067] 물론, 본 발명은 DNA 분자 증폭이 증폭 프라이머와 플라즈마 벡터 등에 의해 이루어지는 것으로 한정되지 않으며 DNA 분자를 증폭할 수 있는 모든 방법을 이용할 수 있다.
- [0068] 이러한 과정을 통해 본 발명에 따른 방법은 디지털 데이터 값에 대응하여 미리 설정된 서열 패턴을 포함하는 DNA 절편을 이용하여 디지털 정보를 DNA 분자에 저장할 수 있다.
- [0069] 또한, 본 발명에 따른 방법은 DNA 절편을 이용하여 디지털 정보를 저장할 수 있기 때문에 자기 디스크를 보완하거나 대체할 정보 저장 매체로 활용할 수 있으며, 화재나 재해에 대비하여 중요 정보를 전기소모 없이 적은 공간에 안전하게 저장할 수 있다.
- [0070] 본 발명의 실시예에 따르면, 디지털 정보를 DNA 절편을 이용하여 저장하기 때문에 합성 시 발생할 수 있는 에러에 강건하고, 시퀀싱 장비 예를 들어, 나노포어(nanopore) 시퀀싱 장비를 이용하여 DNA 분자에 저장된 디지털 정보를 읽을 수 있다. 즉, 본 발명의 방법에 의해 DNA 분자에 저장된 디지털 정보는 나노포어 시퀀싱 장비 등을 이용하여 DNA 서열의 컨텍스트(context)가 가지는 전기적 속성의 차이를 통하여 특정 길이의 절편(fragment)의 차이를 인지할 수 있으며, 따라서 합성 시 에러가 일부 있다 하여도 디지털 데이터 값을 정확하게 인지할 수 있다.
- [0071] 예컨대, 기존 방법은 도 2a에 도시된 바와 같이, 00 = A(아데닌), 01 = T(티민), 10 = G(구아닌), 11 = C(시토신)처럼, 2비트 디지털 데이터를 DNA의 염기에 정보를 표현하는 방법으로, 저장하고자 하는 디지털 정보가 "0100101101"인 경우 DNA의 염기 "TAGCT"로 저장될 수 있다. 하지만, 이런 방법은 합성 에러가 발생될 수 있고, 시퀀싱 에러로 하나의 분자에 저장에 불가능하며, 시퀀싱 수행 시 분자를 파괴할 수 있다. 반면, 본 발명에 따른 방법은 도 2b에 도시된 바와 같이 디지털 데이터 값 0과 1에 대응하는 서열 패턴을 포함하는 DNA 절편을 이용하여 디지털 정보를 저장하는 방법으로, 저장하고자 하는 디지털 정보가 "0100101101"인 경우 0에 대한 DNA 절편과 1에 대한 DNA 절편을 이용하여 DNA 절편을 연결함으로써, DNA 절편을 이용하여 디지털 정보를 저장할 수 있다.
- [0072] 본 발명에서 디지털 데이터 값 0과 1에 대하여, 상이한 값 서열 패턴을 가지도록 설명하였지만, 이에 한정되지 않으며, 모디파이드(modified) 염기의 포함 여부로 미리 합성할 수도 있다. 예를 들어, 디지털 데이터 값 0의 경우 모디파이드되지 않은 염기로 구성되고 20~40bp 길이의 서열을 가질 수 있으며, 디지털 데이터 값 1의 경우 모디파이드된 염기를 포함하고 20~40bp 길이의 서열을 가질 수 있다. 여기서, 모디파이드된 염기는 이노신(inosine), 5-수산화메틸시토신(5-hydroxymethylcytosine) 등을 포함할 수 있다.
- [0073] 이와 같이, 본 발명에 따른 방법은 DNA 염기 서열이 아닌, DNA 서열이 가지는 전기적 특성을 특징으로 정보를 저장하기 때문에 현재의 합성, 어레이 기술을 이용하여 하나의 분자에 디지털 정보를 저장할 수 있으며, 시퀀싱 수행 시 분자를 파괴하지 않을 수 있다. 즉, 본 발명에 따른 방법은 DNA 염기 서열이 아닌, DNA 서열이 가지는

전기적 특성을 특징으로 정보를 저장하기 때문에 합성 시 발생할 수 있는 에러에 강건하고, 저장된 디지털 정보를 읽는 경우 약속한 길이의 DNA 절편 전체의 전기적 특성을 판별하기 때문에 콜링(calling) 에러에 강건할 수 있다. 예컨대, DNA 분자에 저장된 디지털 정보를 읽기 위하여, 약속한 길이의 DNA 절편 전체의 전기 신호를 검출함으로써, 디지털 데이터 값 0과 1을 판별할 수 있다. 물론, 값 서열 패턴이 DNA 절편에 저장되는 경우 나노포어 시퀀싱 장비에 의해 검출되는 전기 신호의 차이를 이용하여 서열 패턴을 읽고 이를 통해 디지털 데이터 값 1과 0을 판별할 수도 있고, 디지털 정보 저장 시 모디파이된 염기의 포함 여부로 디지털 데이터 값 1과 0을 저장하는 경우 모디파이된 염기를 포함하는 DNA 절편의 전기 신호를 검출함으로써, 디지털 데이터 값을 판별할 수 있는데, 0과 1을 판별하기 위한 컷 오프를 설정할 수 있으며, 이에 대한 컷 오프는 해당 기술을 제공하는 사업자 또는 개인에 의해 결정될 수 있고, 이러한 컷 오프는 실험 등을 통해 결정될 수도 있다.

[0075] 도 4는 허프만 코드를 이용한 디지털 정보를 DNA 분자에 저장하는 일 예시도를 나타낸 것으로, 도 4에 도시된 바와 같이 DNA 분자에 허프만 코드에 의해 인코딩된 "나라의말이중국"을 저장하는 경우에 대한 것이다.

[0076] 도 4a에 도시된 바와 같이, 저장하고자 하는 데이터 "나라의말이중국"에 대하여 허프만 코드에 의해 인코딩된 디지털 정보는 44비트의 "0111100100110111001010110001011101110111000"이고, 이러한 44비트의 디지털 정보는 0에 대한 값 서열 패턴 "TATT"와 1에 대한 값 서열 패턴 "ACCC"를 이용하여 DNA 절편에 저장하며, 각각의 디지털 데이터 값이 저장된 DNA 절편을 연결함으로써, 도 4b에 도시된 바와 같은 817mer의 DNA 서열로 이루어질 수 있다.

[0077] 여기서, 도 4b에 도시된 값 서열 패턴 "TATT"와 "ACCC" 앞뒤 또는 좌우에 배치된 일정 길이의 서열 패턴은 링커 서열 패턴이며, 처음의 일정 길이에 대한 서열 패턴은 스타트 서열 패턴이고, 마지막의 일정 길이에 대한 서열 패턴을 엔드 서열 패턴일 수 있다. 링커 서열 패턴은 정보가 없는 랜덤 패턴이므로, DNA 분자에 저장된 이러한 디지털 정보를 나노포어 시퀀싱 장비와 같은 장비를 이용하여 읽는 경우 스타트 서열 패턴을 통해 디지털 정보에 대한 시작을 알 수 있고, 값 서열 패턴을 통해 저장된 디지털 데이터 값을 알 수 있으며, 엔드 서열 패턴을 통해 디지털 정보의 끝을 알 수 있기 때문에 DNA 분자에 저장된 디지털 정보를 알 수 있다. 물론, 허프만 코드와 같은 인코딩 정보도 함께 저장될 수 있으며, 이렇게 저장된 인코딩 정보를 통해 DNA 분자에 저장된 디지털 정보를 디코딩함으로써, 도 4b와 같이 DNA 분자에 저장된 정보가 "나라의말이중국"이라는 것을 알 수 있다.

[0079] 상술한 바와 같이, 본 발명에 따른 방법은 올리고뉴클레오타이드(oligonucleotide)의 합성 에러가 있는 상태에서도 구현 가능한 단일 DNA 분자 기반 정보 저장 방법으로, DNA 분자 중 염기가 아닌 특정 길이의 DNA 절편에 정보를 저장할 수 있고, DNA 절편의 전기적 속성이 달라지도록 서열을 배치하거나 모디파이된 염기를 사용하여 디지털 데이터 값을 DNA 절편에 저장할 수 있기 때문에 합성 시 에러가 일부 있다 하여도, DNA 서열의 컨텍스트가 가지는 전기적 속성의 차이를 통하여 특정 길이의 절편의 차이를 인지할 수 있다.

[0080] 즉, 기존 방법은 단일 염기에 정보를 기록하기 때문에 한 두 개의 염기가 손상을 입어도 저장한 정보에 손상을 입을 수 있는 반면, 본 발명은 염기보다 긴 영역인 DNA 절편에 정보를 기록하기 때문에 한 두 개의 염기가 손상을 입어도 정보 자체는 잘 변하지 않는 강건한 속성이 있다. 또한, 기존 방법은 정확도를 위하여 비트당 최소 수백회 중첩되도록 정보를 중첩하여 사용하는데, 동일한 개수의 분자를 사용하는 경우 본 발명은 기존 방법보다 정보 정확도가 향상될 수 있다.

[0081] 그리고, 본 발명은 저장하고자 하는 디지털 정보가 달라질 때마다 DNA 서열을 새로 합성할 필요가 없으며, 미리 합성해둔 0과 1을 표현하는 DNA 절편을 원하는 순서대로 연결함으로써, 디지털 정보를 DNA 분자에 저장할 수 있다.

[0082] 또한, 본 발명의 방법에 의해 디지털 정보를 저장하는 경우 DNA B O

[0084] 도 5는 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 장치에 대한 구성을 나타낸 것으로, 상기 도 1 내지 도 4에서 설명한 방법을 수행하는 장치에 대한 구성을 나타낸 것이다.

[0085] 도 5를 참조하면, 본 발명의 일 실시예에 따른 DNA 분자 기반 정보 저장 장치(500)는 수신부(510), 합성부(520), 저장부(530) 및 증폭부(540)를 포함한다.

- [0086] 수신부(510)는 저장하고자 하는 디지털 정보를 수신한다.
- [0087] 여기서, 수신부(510)는 저장하고자 하는 정보를 미리 설정된 인코딩 방식 예를 들어, 허프만 코드에 의해 인코딩되고, 이렇게 인코딩된 디지털 정보를 수신할 수 있다.
- [0088] 합성부(520)는 디지털 데이터 값인 제1 값과 제2 값에 대응하는 DNA 절편을 합성한다.
- [0089] 여기서, 제1 값과 제2 값에 대응하는 DNA 절편은 합성부(520)에 의해 미리 합성된 상태이기 때문에 디지털 데이터 값 각각에 대응하는 DNA 절편을 이용하여 원하는 디지털 정보를 저장할 수 있다.
- [0090] 나아가, 합성부(520)에 의해 합성되는 DNA 절편은 DNA 절편의 시작을 나타내는 스타트 서열 패턴, 제1 값과 상기 제2 값 중 어느 하나의 값에 대응하는 값 서열 패턴, DNA 절편의 끝을 나타내는 엔드 서열 패턴 및 리게이션 어세이(ligation assay)를 위한 링커 서열 패턴을 포함할 수 있다.
- [0091] 저장부(530)는 수신부(510)에 의해 수신된 디지털 정보를 제1 값과 제2 값에 대응하는 미리 합성된 일정 길이의 DNA 절편을 이용하여 저장한다.
- [0092] 여기서, 저장부(530)는 디지털 정보와 인코딩 정보를 제1 값과 제2 값에 대해 미리 합성된 DNA 절편에 저장하고, 이러한 DNA 절편을 연결할 수도 있다.
- [0093] 저장부(530)는 디지털 정보의 디지털 데이터 값과 인코딩 정보의 디지털 데이터 값 각각에 대응하는 DNA 절편을 루프 구조를 포함하는 프라이머를 이용하여 연결하고, 이러한 연결 과정을 반복 수행함으로써, 디지털 정보와 인코딩 정보를 DNA 분자에 저장할 수 있다.
- [0094] 예컨대, 저장부(530)는 두 개의 DNA 절편을 순서대로 이어 붙여 제1 DNA 절편 그룹을 생성하고, 두 개 이상의 제1 절편 그룹을 원하는 순서대로 이어 붙여 더 큰 절편 그룹을 생성하며, 이러한 과정을 반복 수행함으로써, 원하는 양의 디지털 정보를 DNA 분자에 저장할 수 있다.
- [0095] 증폭부(540)는 디지털 정보가 저장된 DNA 절편 또는 DNA 분자를 증폭 프라이머 또는 플라스마 벡터에 넣을 미생물 배양 등을 통해 증폭한다.
- [0097] 비록, 도 5의 장치에서 그 설명이 생략되었더라도, 본 발명에 따른 장치는 상기 도 1 내지 도 4에서 설명한 모든 내용을 포함할 수 있다는 것은 이 기술 분야에 종사하는 당업자에게 있어서 자명하다.
- [0099] 또한, 본 발명은 상술한 방법에 의해 DNA 분자에 저장된 디지털 정보를 검출할 수도 있으며, DNA 분자의 서열에 대한 전기 신호의 차이를 통해 디지털 데이터 값에 대한 서열 패턴을 확인함으로써, DNA 분자에 저장된 디지털 데이터 값을 인지하고, 이렇게 인지된 디지털 데이터를 디코딩하여 DNA 분자에 저장된 정보를 확인할 수 있다. 즉, 본 발명은 디지털 데이터를 DNA 분자에 저장하는 방식에 한정되지 않으며, 이러한 방식으로 DNA 분자에 저장된 디지털 정보를 읽는 방법 또한 포함할 수 있다. 구체적으로, 본 발명은 나노포어 시퀀싱 방법과 같은 방식을 통해 DNA 분자의 서열 패턴에 대한 전기 신호를 검출하고, 이렇게 검출된 전기 신호 차를 이용하여 서열 패턴을 검출함으로써, 저장된 디지털 데이터를 검출할 수 있다. 물론, 검출 과정은 DNA 분자에서 스타트 서열 패턴을 검출하여 디지털 정보가 저장된 시작 부분을 인식하고, 해당 부분 이후의 DNA 서열에 대한 전기 신호 차이를 통해 값 서열 패턴을 확인함으로써, DNA 분자에 저장된 디지털 데이터 값을 인지하며, 이러한 과정을 엔드 서열 패턴이 검출될 때까지 수행함으로써, DNA 분자에 저장된 디지털 데이터를 검출할 수 있다. 이렇게 검출된 디지털 데이터는 미리 설정된 디코딩 방식을 이용하여 해당 정보를 확인할 수 있다.
- [0101] 이상에서 설명된 시스템 또는 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 시스템, 장치 및 구성요소는, 예를 들어, 프로세서, 컨트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하

여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0102] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

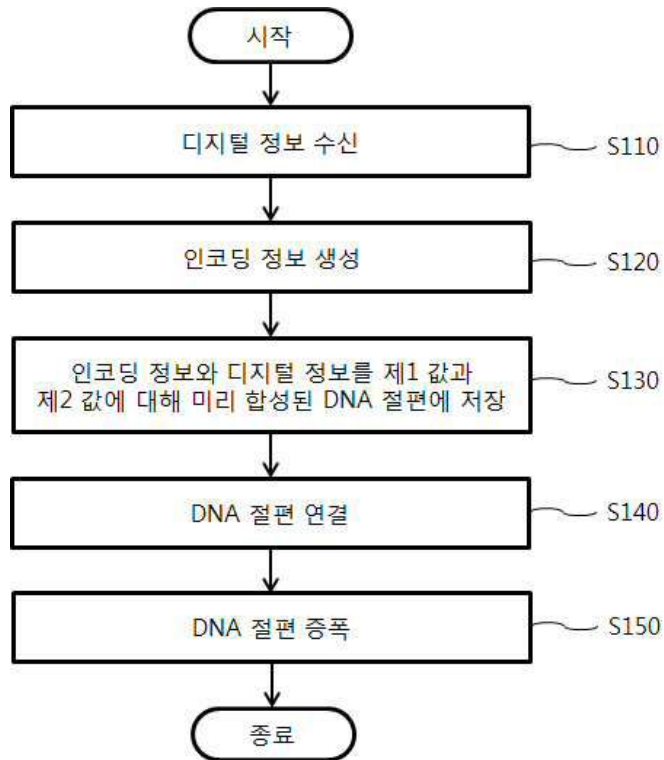
[0103] 실시예들에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0104] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

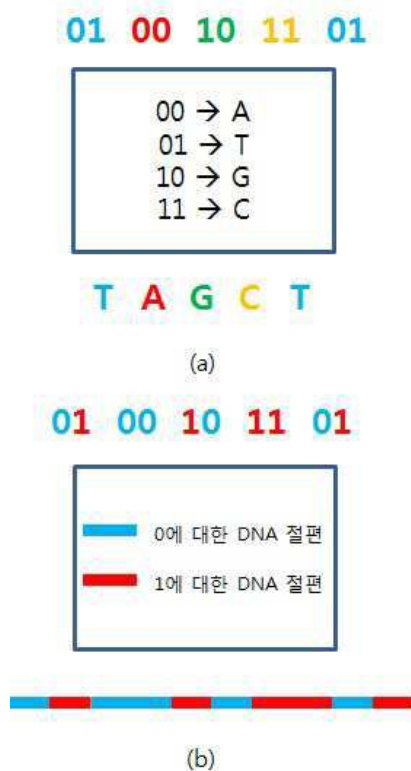
[0105] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

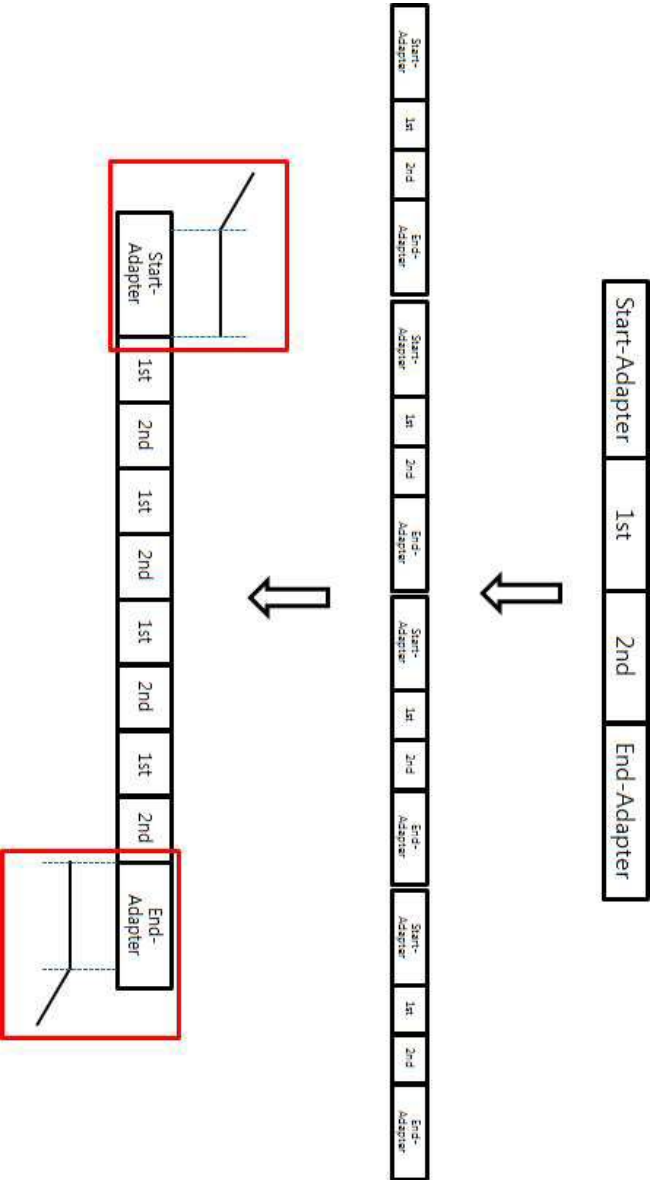
도면1



도면2



도면3



(a) 나라의말이중국
01111001001101110010101100010111101110111000

나라의말이중국(817mer)

GAATTGCGGGCGCTTAGAGCAGTCGCTCCACAAGTACCAAGCCTCGTCTCCACATgtgc
cgTATTacaactgatccaACCCttctctacggaACCCtcgctcatgagACCCagacagcgctgACCCg
ctgatcaacgtTATTgtcgtatctcgTATTctctaggtactACCCgttcggaagcgTATTgtcacagtact
TATTaagcactcgacaACCCtagctgacagaACCCggaagaactgTATTccaaggaactaACCCac
gagagccggaACCCatgatgacatcaACCCtgagacagtgagTATTccaagtgaagagTATTgcccgtac
aagAACCCaagatgtcgtactTATTgttagcgtctcAACCCggaagacggtccgTATTgtcttcgttctAACCC
ggtcagttcttaACCCCaagtgtacgagTATTcagcacagacgtTATTgacaacgacgacTATTgcgagagt
gcacAACCCCaagtcgcccgaagTATTcaagcgagactgAACCCagaagccttagtAACCCtacacagctgtgAC
CCgacacgaggttaACCCcttagagacggtTATTcagtlacacgagAACCCttcacacacggaACCCatcg
gcttagtgACCCttgtagcttgctTATTgtgagttcactAACCCatcagaagcggaACCCgagcagcgct
ACCCttcttaggtcactTATTcgtcttacaactTATTgttcgcccgaactTATTgttagtgCCGTCAATCTTGCCA
TTGCTTGCTCAATCAATCTACTAGTAGGCGGGCGGCTGCAG

(b)

도면4

도면5

500

