

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2014-526899
(P2014-526899A)

(43) 公表日 平成26年10月9日(2014.10.9)

(51) Int.Cl.

C 12 N 15/09 (2006.01)
C 12 M 1/00 (2006.01)

F 1

C 12 N 15/00 Z N A A
C 12 M 1/00 A

テーマコード(参考)

4 B 0 2 4
4 B 0 2 9

審査請求 未請求 予備審査請求 未請求 (全 42 頁)

(21) 出願番号	特願2014-527291 (P2014-527291)
(86) (22) 出願日	平成24年8月23日 (2012.8.23)
(85) 翻訳文提出日	平成26年4月24日 (2014.4.24)
(86) 国際出願番号	PCT/US2012/052036
(87) 国際公開番号	W02013/032850
(87) 国際公開日	平成25年3月7日 (2013.3.7)
(31) 優先権主張番号	61/527,922
(32) 優先日	平成23年8月26日 (2011.8.26)
(33) 優先権主張国	米国(US)
(31) 優先権主張番号	61/532,825
(32) 優先日	平成23年9月9日 (2011.9.9)
(33) 優先権主張国	米国(US)
(31) 優先権主張番号	13/592,827
(32) 優先日	平成24年8月23日 (2012.8.23)
(33) 優先権主張国	米国(US)

(71) 出願人	513117136 ジェン9・インコーポレイテッド Gen9, INC. アメリカ合衆国O 2 1 3 9マサチューセッ ツ州ケンブリッジ、テクノロジー・スクエ ア500番、スウィート130
(74) 代理人	100081422 弁理士 田中 光雄
(74) 代理人	100084146 弁理士 山崎 宏
(74) 代理人	100122301 弁理士 富田 慎史
(74) 代理人	100157956 弁理士 稲井 史生

最終頁に続く

(54) 【発明の名称】核酸の高忠実度アセンブリのための組成物および方法

(57) 【要約】

本発明の態様は、標的核酸を設計して製造するための方法、組成物、およびアルゴリズムに関する。方法は、(1)各々の両端に制限酵素認識配列を有する、複数の平滑末端二本鎖核酸断片を提供する；(2)制限酵素認識配列の近接における複数の平滑末端二本鎖核酸断片の酵素消化を経て、各々が2つの異なる非相補性突出を有する、複数の付着末端二本鎖核酸断片を生成する；(3)第一の付着末端二本鎖核酸断片の第一の突出が、第二の付着末端二本鎖核酸断片の第二の突出に一意的に相補であるところの、複数の付着末端二本鎖核酸断片のリガーゼでのライゲーション；および(4)一意的な配置が標的核酸を含む、複数の付着末端二本鎖核酸断片の直線配置の形成、を含む。ある実施形態において、複数の平滑末端二本鎖核酸断片は、固相担体上で合成された複数のオリゴヌクレオチドを放出し、ポリメラーゼを基にした反応を用いる、複数のオリゴヌクレオチドの相補鎖を合成することにより提供できる。

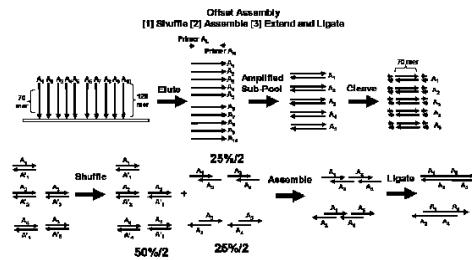


Figure 10A

【特許請求の範囲】**【請求項 1】**

所定の配列を有する標的核酸を作製する方法であって、複数の平滑末端二本鎖核酸断片の各々の両端に制限酵素認識配列を有する、複数の平滑末端二本鎖核酸断片を提供すること；複数の平滑末端二本鎖核酸断片の酵素消化を経て、標的核酸配列とともに含む複数の付着末端二本鎖核酸断片を生成すること、ここで、複数の付着末端二本鎖核酸配列の各々が2つの異なる非相補性突出を有する；

第一の付着末端二本鎖核酸断片の第一の突出が、第二の付着末端二本鎖核酸断片の第二の突出に一意的に相補であるところの、複数の付着末端二本鎖核酸断片のリガーゼでのライゲーション；および

一意的な配置が所定の配列を有する標的核酸を含む、複数の付着末端二本鎖核酸断片の直線配置の形成、
を含む、方法。

【請求項 2】

複数の平滑末端二本鎖核酸断片が、固相担体上に固定化された複数の一一本鎖オリゴヌクレオチドから生成される、請求項1に記載の方法。

【請求項 3】

複数の平滑末端二本鎖核酸断片が：
固相担体上で合成された複数のオリゴヌクレオチドを放出すること；および
ポリメラーゼを基にした反応を用いて、複数のオリゴヌクレオチドの相補鎖を合成すること、
を含む、請求項1に記載の方法。

【請求項 4】

複数のオリゴヌクレオチドの各々がユニバーサルプライマー結合部位を含み、該ユニバーサルプライマー結合部位に相補的なユニバーサルプライマーが、該ポリメラーゼを基にした反応で使用される、請求項3に記載の方法。

【請求項 5】

複数のオリゴヌクレオチドの各々が制限酵素認識配列を含む、請求項4に記載の方法。
【請求項 6】

該制限酵素認識配列がユニバーサルプライマー結合部位の一部であり、ユニバーサルプライマー結合部位の5'または3'末端に位置するか、または制限酵素認識配列がユニバーサルプライマー結合部位の上流または下流に位置する、請求項5に記載の方法。

【請求項 7】

ユニバーサルプライマーが望ましくない酵素消化産物の親和性除去を促進するために親和性タグを有する、請求項4に記載の方法。

【請求項 8】

親和性タグがビオチンである、請求項7に記載の方法。

【請求項 9】

複数の平滑末端二本鎖核酸が、少なくとも3、4、5、6、7、8、10、15または20の異なる平滑末端二本鎖核酸断片を含む、請求項1に記載の方法。

【請求項 10】

複数の平滑末端二本鎖核酸断片が、少なくとも50、100、200または300塩基長である、請求項1に記載の方法。

【請求項 11】

制限酵素認識配列がすべての平滑末端二本鎖核酸断片に共通する、請求項1に記載の方法。

【請求項 12】

複数の平滑末端二本鎖核酸断片が、同じ数の塩基を有する突出を生成するために選択された2つの異なる制限酵素によって認識できる少なくとも2つの異なる制限酵素認識配列

を含む、請求項 1 に記載の方法。

【請求項 1 3】

制限酵素認識配列が I I s 型制限酵素で認識されることができる、請求項 1 に記載の方法。

【請求項 1 4】

I I s 型制限酵素が B s a I 、 B s m B I 、 B s p Q I 、 B t g Z I 、 B s m F I 、 F o k I 、 B b v I 、 それらのいかなる変異体、またそれらのいかなる組み合わせである、請求項 1 3 に記載の方法。

【請求項 1 5】

複数の付着末端二本鎖核酸断片が、付着末端二本鎖核酸断片の付着末端が、隣接する付着末端二本鎖核酸断片における次の付着末端に一意的に相補であるように設計される、請求項 1 に記載の方法。 10

【請求項 1 6】

突出が少なくとも 3 、 4 、 5 、 6 、 7 、または 8 塩基長である、請求項 1 に記載の方法。

【請求項 1 7】

突出が互いに少なくとも 1 、 2 、 3 または 4 塩基異なる、請求項 1 に記載の方法。

【請求項 1 8】

突出が、 5 ' または 3 ' 突出である、請求項 1 に記載の方法。

【請求項 1 9】

ライゲーション段階の前に、複数の付着末端二本鎖核酸断片を精製して望ましくない酵素消化産物を除去することをさらに含む、請求項 1 に記載の方法。 20

【請求項 2 0】

望ましくない酵素消化産物が、約 40 、約 35 、約 30 、約 25 、約 20 または約 15 塩基長未満の断片を含む、請求項 1 9 に記載の方法。

【請求項 2 1】

該精製が、シリカとの異なる親和性、サイズろ過、ポリエチレングリコールまたは臭化セチルトリメチルアンモニウムでの沈殿、またはそれらのいかなる組み合わせを含む、請求項 1 9 に記載の方法。 30

【請求項 2 2】

リガーゼが T 3 DNA リガーゼ、 T 4 DNA リガーゼ、 T 7 DNA リガーゼ、大腸菌 (E . c o l i) DNA リガーゼ、それらのいかなる変異体、またはそれらのいかなる組み合わせである、請求項 1 に記載の方法。

【請求項 2 3】

標的核酸が非天然核酸である、請求項 1 に記載の方法。

【請求項 2 4】

標的核酸が少なくとも 500 、 800 、 1000 、 1500 、 2000 、または 3000 塩基長である、請求項 1 に記載の方法。

【請求項 2 5】

標的核酸に特異的なプライマー対およびポリメラーゼを使用する標的核酸の増幅をさらに含む、請求項 1 に記載の方法。 40

【請求項 2 6】

標的核酸の配列の確認をさらに含む、請求項 1 に記載の方法。

【請求項 2 7】

複数の平滑末端二本鎖核酸断片が合成オリゴヌクレオチドから階層的に組み立てられる、請求項 1 に記載の方法。

【請求項 2 8】

複数の核酸断片が单一のプールでライゲーションされる、請求項 1 に記載の方法。

【請求項 2 9】

複数の核酸断片が少なくとも 2 つのプール中にあり、第一のプール中の各々の核酸断片

が第二のプールの核酸断片に相補的な末端を有する、請求項 1 に記載の方法。

【請求項 3 0】

複数の核酸断片がオリゴヌクレオチド二量体である、請求項 2 9 に記載の方法。

【請求項 3 1】

標的核酸に組み立てられる複数の出発核酸を設計するための方法であって、

(a) 標的核酸の標的配列を得る；

(b) あらゆる 2 個の隣接する部分配列が、お互い N 塩基で重なり合うように複数の部分配列を選択する；

(c) 得られる重なり合った N - 塩基配列をメモリに格納する；

(d) 重なり合う N - 塩基配列をお互いに比較し、少なくとも 1 個の塩基によりお互いが異なることを確実にする；および 10

(e) いかなる 2 個の隣接する出発核酸が、お互いが N 塩基で一意的に重なり合う、複数の満足できる出発核酸を得るまで、段階 (b) から (d) を繰り返す、ことを含む、方法。

【請求項 3 2】

ランキング配列をその 5' 末端および 3' 末端で設計することをさらに含み、ランキング配列が I I s 型制限酵素で認識されることができる制限酵素認識部位を含む、請求項 3 1 に記載の方法。

【請求項 3 3】

制限酵素認識部位が、 I I s 型認識部位である、請求項 3 2 に記載の方法。 20

【請求項 3 4】

ランキング配列が、いかなる 2 つの隣接出発核酸断片が、制限酵素での切断の後に一意的な付着末端を有するように、核酸の伸長をさらに含む、請求項 3 2 に記載の方法。

【請求項 3 5】

ランキング配列がプライマー結合部位をさらに含む、請求項 3 2 に記載の方法。

【請求項 3 6】

標的核酸配列が、非天然核酸である、請求項 3 1 に記載の方法。

【請求項 3 7】

標的核酸が、少なくとも 500、800、1000、1500、2000 または 3000 塩基長である、請求項 3 1 に記載の方法。 30

【請求項 3 8】

各々の部分配列が、約 50、60、70、80、90、100、120、140、160、180、200、220、240、260、280、300 またはそれ以上の塩基長である、請求項 3 1 に記載の方法。

【請求項 3 9】

N が整数であり、3、4、5、6、7、8 またはそれ以上である、請求項 3 1 に記載の方法。

【請求項 4 0】

請求項 3 1 に記載の方法で設計された、標的核酸に組み立てられる複数の出発核酸。

【請求項 4 1】

各々がさらに、そこから複数の出発核酸を增幅するための設計されたユニバーサルプライマー結合部位を含む、請求項 4 0 に記載の複数の出発核酸。 40

【請求項 4 2】

各々がさらに設計された制限酵素認識配列を有する、請求項 4 0 に記載の複数の出発核酸。

【請求項 4 3】

標的核酸配列を組み立てるためのシステムであって、

各々の出発核酸がさらに設計されたユニバーサルプライマー結合部位および設計された制限酵素認識配列を含む、請求項 4 0 に記載の複数の出発核酸配列の合成のための固相担体； 50

ユニバーサルプライマー結合部位に相補的なユニバーサルプライマーを用いるポリメラーゼを基にした反応により複数の出発核酸配列の相補鎖を合成し、それにより複数の平滑末端二本鎖核酸断片を生成するためのポリメラーゼ反応ユニット；

複数の平滑末端二本鎖核酸断片の酵素消化を経て、各々が2つの異なる非相補的突出を有する複数の付着末端二本鎖核酸フラグメントを生成するための消化ユニット；および

第一の付着末端二本鎖核酸断片の第一の突出が、第二の付着末端二本鎖核酸断片の第二の突出に一意的に相補であるところの、複数の付着末端二本鎖核酸断片のリガーゼでのライゲーションのための、ライゲーションユニット、

を含む、システム。

【請求項 4 4】

標的核酸に組み立てられる複数の出発核酸を設計するための、コンピュータプログラム製品であって、該プログラムはハードウェアコンピュータ記憶媒体に存在し、プロセッサにより実行されると、プロセッサに、

(a) 標的核酸の標的配列を得る；

(b) あらゆる2個の隣接する部分配列が、お互いN塩基で重なり合うように複数の部分配列を選択する；

(c) 得られる重なり合ったN - 塩基配列をメモリに格納する；

(d) 重なり合うN - 塩基配列をお互いに比較し、少なくとも1個の塩基によりお互いが異なることを確実にする；および

(e) いかなる2個の隣接する出発核酸が、お互いがN塩基で一意的に重なり合う、複数の満足できる出発核酸を得るまで、段階(b)から(d)を繰り返す、
を含む操作を実施させる、コンピュータプログラム製品。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

関連出願

本願は、2012年8月23日に出願された米国特許出願第13/592,827号、
2011年8月26日に出願された米国仮出願第61/527,922号、および2011年9月9日に出願された米国仮出願61/532,825号の優先権および利益を主張し、これらの各々の全体が参考により本明細書に組み込まれる。

【0 0 0 2】

発明の分野

本発明の方法および組成物は核酸アセンブリ、および特に高忠実度の、多重核酸アセンブリ反応に関する。

【背景技術】

【0 0 0 3】

背景

組換え体および合成核酸は、研究、工業、農業および医薬において多くの用途がある。組換え体および合成核酸は、種々の医薬、工業または農業の目的に用いられ得る酵素、抗体、成長因子、受容体および他のポリペプチドを含む、多量のポリペプチドを発現および得るために用いることができる。組換え体および合成核酸はまた、変異細菌、酵母、哺乳動物、植物および他の生物を含む遺伝的に変更された生物に使用できる。遺伝的に変更された生物は研究（例えば、疾患の動物モデルとして、生物学的プロセスの解明のためのツールとして、等）、工業（例えば、タンパク質発現の宿主生物として、工業産物の生成のためのバイオリアクターとして、環境修復のためのツールとして、天然化合物の工業用途での単離または変更のため、等）、農業（例えば、収率の上昇した、または疾患または環境ストレスへの耐性が増加した変更作物、等）、および他の用途において使用され得る。組換え体および合成核酸はまた、治療組成物（例えば、遺伝子発現の変更のため、遺伝子治療のため、等）または診断ツール（例えば、症状のためのプローブとして、等）として使用され得る。

【0004】

存在する核酸（例えば、天然に生じる核酸）を改変して組換え核酸を生成するために、多数の技術が開発されている。例えば、核酸増幅、変異誘発、ヌクレアーゼ消化、ライゲーション、クローニング、および他の技術の組み合わせは、多くの異なった組換え核酸を作出するために使用され得る。化学的に合成されたポリヌクレオチドは核酸増幅、変異誘発、およびクローニングのためにプライマーまたはアダプターとしてしばしば使用される。

【0005】

核酸が作成され（例えば、化学的に合成され）、組み立てられて、対象となるより長い標的核酸を製造する *d e n o v o* 核酸アセンブリのための技術もまた開発され得る。例えば、異なる多重アセンブリ技術は、オリゴヌクレオチドを、研究、工業、農業、および / または医薬において使用できるより大きい合成核酸を組み立てるために開発されている。しかしながら、現在利用可能なアセンブリ技術の一つの制限は、比較的高い誤り率である。それゆえ、高忠実度で、低コストのアセンブリ方法が必要とされている。

10

【発明の概要】**【課題を解決するための手段】****【0006】****発明の概要**

本発明の態様は、標的核酸を製造する方法に関する。いくつかの実施形態によると、方法は、（1）複数の平滑末端二本鎖核酸断片の各々の両端に制限酵素認識配列を有する、複数の平滑末端二本鎖核酸断片を提供する；（2）制限酵素認識配列の近接における複数の平滑末端二本鎖核酸断片の酵素消化を経て、各々が2つの異なる非相補性突出を有する、複数の付着末端二本鎖核酸断片を生成する；（3）第一の付着末端二本鎖核酸断片の第一の突出が、第二の付着末端二本鎖核酸断片の第二の突出に一意的に相補であるところの、複数の付着末端二本鎖核酸断片のリガーゼでのライゲーション；および（4）一意的な配置が標的核酸を含む、複数の付着末端二本鎖核酸断片の直線配置の形成、を含む。ある実施形態において、複数の平滑末端二本鎖核酸断片は、固相担体上で合成された複数のオリゴヌクレオチドを放出し、ポリメラーゼを基にした反応を用いる、複数のオリゴヌクレオチドの相補鎖を合成することにより提供できる。

20

【0007】

本発明の他の態様において、標的核酸に組み立てられる複数の出発核酸を設計するための方法が提供される。いくつかの実施形態によると、方法は、（1）標的核酸の標的配列を得る；（2）あらゆる2個の隣接する部分配列（*s u b s e q u e n c e*）が、お互いN塩基で重なり合うように複数の部分配列を選択する；（3）得られる重なり合ったN-塩基配列をメモリに格納する；（4）重なり合うN-塩基配列をお互いに比較し、少なくとも1個の塩基によりお互いが異なることを確実にする；および（5）いかなる2個の隣接する出発核酸が、お互いがN塩基で一意的に重なり合う、複数の満足できる出発核酸を得るまで、段階（2）から（4）を繰り返す、を含む。

30

【0008】

本発明のまた他の態様は、本明細書に記載された方法により設計された、標的核酸に組み立てられる複数の出発核酸に関する。ある実施形態において、複数の出発核酸は、それぞれさらにそこから複数の出発核酸を増幅するために設計されたユニバーサルプライマー結合部位を含むことができる。複数の出発核酸配列はまた、それぞれさらに設計された制限酵素認識配列を含むことができる。

40

【0009】

また他の態様において、標的核酸配列を組み立てるためのシステムが提供される。システムは、（1）各々の出発核酸がさらに設計されたユニバーサルプライマー結合部位および設計された制限酵素認識配列を含む、本明細書に記載された複数の出発核酸配列の合成のための固相担体；（2）ユニバーサルプライマー結合部位に相補的なユニバーサルプライマーを用いるポリメラーゼを基にした反応により複数の出発核酸配列の相補鎖を合成し

50

、それにより複数の平滑末端二本鎖核酸断片を生成するためのポリメラーゼ反応ユニット；(3)制限酵素認識配列の近接における複数の平滑末端二本鎖核酸断片の酵素消化を経て、各々が2つの異なる非相補的突出を有する複数の付着末端二本鎖核酸フラグメントを生成するための消化ユニット；および(4)第一の付着末端二本鎖核酸断片の第一の突出が、第二の付着末端二本鎖核酸断片の第二の突出に一意的に相補であるところの、複数の付着末端二本鎖核酸断片のリガーゼでのライゲーションのための、ライゲーションユニット、を含む。

【0010】

本発明の他のさらなる態様は、標的核酸に組み立てられる複数の出発核酸を設計するための、コンピュータプログラム製品を提供し、該プログラムはハードウェアコンピュータ記憶媒体に存在し、プロセッサにより実行されると、プロセッサに、(1)標的核酸の標的配列を得る；(2)あらゆる2個の隣接する部分配列が、お互いN塩基で重なり合うように複数の部分配列を選択する；(3)得られる重なり合ったN-塩基配列をメモリに格納する；(4)重なり合うN-塩基配列をお互いに比較し、少なくとも1個の塩基によりお互いが異なることを確実にする；および(5)いかなる2個の隣接する出発核酸が、お互いがN塩基で一意的に重なり合う、複数の満足できる出発核酸を得るまで、段階(2)から(4)を繰り返す、を含む操作を実施させる。

10

【図面の簡単な説明】

【0011】

【図1】図1は多重オリゴヌクレオチドアセンブリ反応のためのオリゴヌクレオチドの例示的設計を説明する。

20

【0012】

【図2】図2は多重オリゴヌクレオチドアセンブリ反応からの試験製品で使用されるプライマーの相対的な位置を説明する。

【0013】

【図3】図3は対オリゴヌクレオチドアセンブリ反応の実施形態を説明する。

【0014】

【図4】図4は多重オリゴヌクレオチドアセンブリ反応の実施形態を説明する。

【0015】

【図5】図5は、図4の多重オリゴヌクレオチドアセンブリ反応の産物のPCRを基にした試験を説明する。

30

【0016】

【図6】図6は、図4の多重オリゴヌクレオチドアセンブリ反応の産物の配列確認を説明する。

【0017】

【図7】図7Aおよび7Bは、対ミスマッチライゲーションアッセイの実施形態を説明する。

【0018】

【図8】図8は図1の設計に基づく代替のアセンブリ産物を説明する。

40

【0019】

【図9】図9Aおよび9Bは、アセンブリ断片に隣接する配列のための2つの設計戦略を説明する。

【0020】

【図10】図10Aおよび10Bは、2つのオフセットアセンブリ戦略を説明する。

【発明を実施するための形態】

【0021】

発明の詳細な説明

本発明の態様は、単一のアセンブリ段階においてより長い核酸産物を生成するために複数の核酸断片を共有結合するための方法および組成物に関する。本発明の態様は、アセンブリ誤り率を減少させながら、複数の核酸断片を効率的に組み立てるため、および／また

50

は大きい核酸産物を生成するために必要な複数の段階を減少させるために使用できる。本発明の態様は、アセンブリ忠実度、処理量および／または効率を増加させ、費用を下げ、および／またはアセンブリ時間を減少させるために核酸アセンブリ手順に組み込むことができる。いくつかの実施形態において、本発明の態様は多くの異なる標的核酸産物の並行生産を容易にするために自動化および／または高処理アセンブリにおいて実行され得る。

【0022】

多重オリゴヌクレオチドアセンブリ

所定の核酸断片が、多重アセンブリ反応（例えば、多重酵素仲介反応、多重化学的アセンブリ反応、またはそれらの組み合わせ）において、複数の異なる出発核酸（例えば、オリゴヌクレオチド）から組み立てられ得る。多重核酸アセンブリ反応のある態様は、以下に記載される多重オリゴヌクレオチドアセンブリ反応により説明される。アセンブリ反応の記載が、オリゴヌクレオチドの文脈において制限することを意図しないことが理解されるべきである。本明細書に記載されたアセンブリ反応は、1つ以上のさまざまな源（例えば、合成または天然ポリヌクレオチド、核酸增幅産物、核酸分解産物、オリゴヌクレオチド等）から得られた出発核酸を使用することで実行され得る。出発核酸はアセンブリ核酸（例えば、アセンブリオリゴヌクレオチド）と称され得る。本明細書で用いられるように、アセンブリ核酸はアセンブリ工程の間に生成される核酸産物に組み込まれるように設計された配列を有する。しかしながら、アセンブリ反応の記載が二本鎖核酸の文脈において制限されることを意図しないことが理解されるべきである。いくつかの実施形態において、図面で説明され、本明細書に記載された1つ以上の出発核酸は、一本鎖核酸として提供され得る。従って、図面および明細書が付着末端二本鎖核酸のアセンブリを説明する場合に1つ以上の一一本鎖核酸の存在が予期されることが理解されるべきである。

10

20

30

【0023】

本明細書で用いられるように、オリゴヌクレオチドは少なくとも2つの共有結合された核酸残基を含む核酸分子であり得る。いくつかの実施形態において、オリゴヌクレオチドは10から1,000の間のヌクレオチド長であり得る。例えば、オリゴヌクレオチドは500から1,000の間のヌクレオチド長、または10から500の間のヌクレオチド長であり得る。いくつかの実施形態において、オリゴヌクレオチドは約20から約300の間のヌクレオチド長（例えば、約30から250、40から220、50から200、60から180、または約65または約150のヌクレオチド長）、約100から約200の間、約200から約300ヌクレオチドの間、約300から約400の間、約400から約500の間のヌクレオチド長であり得る。しかしながら、より短いか、またはより長いオリゴヌクレオチドも使用され得る。オリゴヌクレオチドは一本鎖核酸であり得る。しかしながら、いくつかの実施形態において、本明細書に記載されるように二本鎖オリゴヌクレオチドが使用され得る。ある実施形態において、オリゴヌクレオチドは、以下でより詳細に記載されるように、化学的に合成され得る。いくつかの実施形態において、入力核酸（input nucleic acid）（例えば、合成オリゴヌクレオチド）は使用の前に増幅され得る。得られた産物は二本鎖であり得る。

【0024】

ある実施形態において、各々のオリゴヌクレオチドは、組み立てられるべき所定の標的核酸配列の異なる部分と同じ配列を有するように設計され得る。従って、いくつかの実施形態において、各々のオリゴヌクレオチドは二本鎖標的核酸の2つの鎖の1つの部分と同じ配列を有し得る。明確化のため、二本鎖核酸の2個の相補鎖が本明細書においてポジティブ（P）およびネガティブ（N）鎖と称される。この名称は、鎖がコード配列のセンス鎖とアンチセンス鎖であると暗示することを意図するものではない。それらは、核酸の配列や機能にかかわらず、核酸（例えば、標的核酸、中間体核酸断片等）の2つの相補鎖についてのみ言及する。従って、いくつかの実施形態において、P鎖はコード配列のセンス鎖であり得るが、一方で他の実施形態において、P鎖はコード配列のアンチセンス鎖であり得る。本明細書における相補的核酸または相補的核酸領域への参照が天然DNAに典型的な逆並行の型（antiparallel fashion）でハイブリダイズできる

40

50

ようにお互いに逆相補である配列を有する核酸またはその領域を意味することが理解されるべきである。

【0025】

本発明の一の態様によると、標的核酸はP鎖、N鎖またはP鎖およびN鎖の両方を含む二本鎖核酸のいずれかであり得る。異なるオリゴヌクレオチドが、異なる長さを有するように設計され得ることが理解されるべきである。いくつかの実施形態において、1つ以上の異なるオリゴヌクレオチドが重なり合う配列領域（例えば、重なり合う5'領域および/または重なり合う3'領域）を有し得る。重なり合う配列領域は、同一（すなわち、核酸断片の同一鎖に対応する）、または相補的（すなわち、核酸断片の相補鎖に対応する）であり得る。複数のオリゴヌクレオチドは、同一配列領域で重なり合う1つ以上のオリゴヌクレオチド対、相補配列領域で重なり合う1つ以上のオリゴヌクレオチド対、またはそれらの組み合わせを含み得る。重なり合う配列は、いかなる適当な長さであり得る。例えば、重なり合う配列は、アセンブリ反応で使用される1つ以上の核酸の全長を含み得る。重なり合う配列は、約2から約50の間（例えば、3から20の間、3から10の間、3から8の間、または4、5、6、7、8、9などヌクレオチド長）であり得る。しかしながら、より短い、より長い、または中間的な重なり合う長さが使用され得る。アセンブリ反応で使用される異なる入力核酸の間の重なりは異なる長さ、および/または、配列があり得ることが理解されるべきである。例えば、重なり合う配列は、少なくとも1つのヌクレオチド、2ヌクレオチド、3ヌクレオチド、またはそれ以上で互いに異なり得る。重なり合う配列が×ヌクレオチドでお互いに異なると仮定する場合、一回の反応で、(4^x+1)までの異なる入力核酸部品が共に組み立てられる。10

【0026】

所定の核酸断片を生成するよう設計された多重オリゴヌクレオチドアセンブリ反応において、反応における、異なるオリゴヌクレオチドの組み合わせ配列はポジティブ鎖、ネガティブ鎖、両方の鎖、またはポジティブ鎖の一部およびネガティブ鎖の一部の組み合わせ上の全体の核酸断片の配列に及び得る。複数の異なるオリゴヌクレオチドは、組み立てられるべき核酸断片の全配列に対応する、ポジティブ配列、ネガティブ配列、またはポジティブおよびネガティブ配列の両方のいずれかを提供し得る。いくつかの実施形態において、複数のオリゴヌクレオチドは、核酸断片の、ポジティブ配列の1つ以上の部分と同一の配列を有する1つ以上のオリゴヌクレオチドおよびネガティブ配列の1つ以上の部分と同一の配列を有する1つ以上のオリゴヌクレオチドを含み得る。異なるオリゴヌクレオチドの1つ以上の対は本明細書に記載される所定の核酸断片配列の重なり合う部分（例えば、核酸断片の同一鎖または相補鎖からの重なり合う配列部分）と同一の配列を含み得る。いくつかの実施形態において、複数のオリゴヌクレオチドは所定の核酸断片のポジティブ配列全体に及ぶように結合する配列を有するオリゴヌクレオチドの組および所定の核酸断片のネガティブ配列全体に及ぶように結合する配列を有するオリゴヌクレオチドの組を含む。しかしながら、ある実施形態において、複数のオリゴヌクレオチドは、核酸断片の1つの鎖（ポジティブまたはネガティブのいずれか）の配列部分と、相補的な配列をもつオリゴヌクレオチドは全く含まず、同一の配列をもつ1つ以上のオリゴヌクレオチドを含み得る。一の実施形態において、複数のオリゴヌクレオチドは、所定の核酸断片のポジティブ配列の部分と同一の配列を有するオリゴヌクレオチドのみを含む。一の実施形態において、複数のオリゴヌクレオチドは、所定の核酸断片のネガティブ配列の部分と同一の配列を有するオリゴヌクレオチドのみを含む。これらのオリゴヌクレオチドは連続ライゲーションまたは伸長を基にした反応（例えば、1つ以上の複数のオリゴヌクレオチドと相補的な3'領域を有するオリゴヌクレオチドが反応に加えられる場合）で組み立てられ得る。30

【0027】

一の態様において、核酸断片は、リガーゼ仲介ライゲーションの1以上の周回において結合およびライゲーションされる複数のオリゴヌクレオチドからのリガーゼ仲介アセンブリ反応において組み立てられ得る。リガーゼを基にしたアセンブリ技術は、隣接する3'および5'核酸末端（例えば、3'末端が5'末端のすぐに隣接するように相補的テンプ40

レート核酸上にアニールされた 5' リン酸基および 3' ヒドロキシル基の核酸) の共有結合を触媒できる適切な 1 つ以上のリガーゼ酵素を伴い得る。従って、リガーゼは第一の核酸の 5' リン酸基と第二の核酸の 3' ヒドロキシル基の間のライゲーション反応を、第一および第二の核酸がテンプレート核酸上でお互いに続いてアニールされると、触媒し得る。リガーゼは、組換え型または天然源から得られ得る。いくつかの実施形態において、1 つ以上の低温(例えば、室温以下)リガーゼが使用され得る(例えば、T3 DNA リガーゼ、T4 DNA リガーゼ、T7 DNA リガーゼ、および / または大腸菌(E. coli) DNA リガーゼ)。より低温のリガーゼは高温で安定ではないであろうより短い突出(例えば、約 3、約 4、約 5 または約 6 塩基の突出)に有用であり得る。リガーゼはまた、熱安定性リガーゼであり得る。いくつかの実施形態において、好熱性生物からの熱安定性リガーゼが使用され得る。熱安定性リガーゼの例は、Tth DNA リガーゼ(例えば、Eurogentec および Genecraft から利用可能な、Thermus thermophilic 由来); Pfu DNA リガーゼ(Pyrococcus furiosus からの超好熱性リガーゼ); Taq リガーゼ(Thermus aquaticus 由来)他の好適な熱安定性リガーゼ、またはこれらのいかなる組み合わせを含むが、これらに限定されるものではない。

10

20

30

40

50

【0028】

本発明の態様は、異なる型の核酸アセンブリ反応(例えば、多重核酸アセンブリ反応)を増幅するために使用され得る。本発明の態様は、例えば、その開示が参照により本明細書に組み込まれる、Carr et al., 2004, Nucleic Acid Research, Vol. 32, No 20, e162 (9 頁); Richmond et al., 2004, Nucleic Acids Research, Vol. 32, No 17, pp. 5011-5018; Caruthers et al., 1972, J. Mol. Biol. 72, 475-492; Hecker et al., 1998, Biotechniques 24: 256-260; Kodumal et al., 2004, PNAS Vol. 101, No. 44, pp. 15573-15578; Tian et al., 2004, Nature, Vol. 432, pp. 1050-1054; および米国特許番号 6,008,031 および 5,922,539 に記載される 1 以上のアセンブリ反応の組み合わせで使用され得る。所定の核酸断片を生成する多重核酸アセンブリ反応のある実施形態は、図 1 ~ 10 を参照することによって説明される。本明細書に記載の合成およびアセンブリ反応(例えば、オリゴヌクレオチド合成、ステップワイズアセンブリ、多重核酸アセンブリ、核酸断片の階層的なアセンブリ、またはそれらのいかなる組み合わせも含む)は、反応チューブ中、マルチウェルプレート中、表面上、カラム上、マイクロ流体デバイス(例えば、マイクロ流体チューブ)中、キャピラリーチューブ、等を含むいかなる適切な形式でも実施され得ることを理解すべきである。例えば、いくつかの実施形態では、「再帰的アセンブリ」または「階層的アセンブリ」により、標的核酸を組み立てることができる。この実施形態において、標的核酸は、まず 2 つ以上の重なり合う核酸断片(または部分アセンブリ断片(subassembly fragment))に分割される。各々の核酸断片はその後、2 つ以上の重なり合うより小さい核酸断片に細分される。

【0029】

合成オリゴヌクレオチド

オリゴヌクレオチドは、いかなる適切な技術を使用することで合成され得る。例えば、オリゴヌクレオチドはカラムまたは他の担体(例えば、チップ)上で合成され得る。チップを基にした合成技術の例は CombiMatrix、Agilent、Affymetrix、または他の源から利用可能な合成デバイスまたは方法で使用される技術を含む。合成オリゴヌクレオチドはいかなる適当なサイズ、例えば 10 と 1,000 ヌクレオチド長の間(例えば、10 と 200、200 と 500、500 と 1,000 ヌクレオチド長の間、またはそれらのいかなる組み合わせ)であり得る。アセンブリ反応は、各々が独立し

て10と300の間のヌクレオチド長（例えば、20と250の間、30と200の間、50と150の間、50と100の間、またはいかなる中間数のヌクレオチド）であり得る、複数のオリゴヌクレオチドを含み得る。しかしながら、1つ以上のより短い、またはより長いオリゴヌクレオチドがある実施形態において使用され得る。

【0030】

本明細書で使用されるように、「担体」および「基板」なる用語は互換的に使用され、核酸などの高分子が合成または固定化される多孔性または無孔性の溶剤不溶性物質をいう。本明細書で使用されるように、「多孔性」は実質的に均一な直径を有する（例えば、n mの範囲）を有する孔を含む物質を意味する。多孔性物質は、紙、合成フィルター等を含むことができるが、これらに限定されるものではない。そのような多孔性物質において、反応は孔内で起こり得る。担体は、多数の形状、例えばピン、片、板、円盤、棒、曲板（bend）、筒状構造、ビーズおよびナノ粒子を含む粒子等、のどれかを有することができる。担体は可変幅を有することができる。

10

【0031】

担体は、吸水できるか、または吸水できるようにされている。担体は、シリカ、硫酸マグネシウム、アルミナなどの無機粉；紙を含む纖維、例えばろ紙、クロマトグラフ紙などの天然高分子材料、特にセルロース系材料およびセルロース由来の材料；ニトロセルロース、セルロースアセテート、ポリ（塩化ビニル）、ポリアクリルアミド、架橋デキストラン、アガロース、ポリアクリレート、ポリエチレン、ポリプロピレン、ポリ（4-メチルブテン）、ポリスチレン、ポリメタクリレート、ポリ（エチレンテレフタラート）、ナイロン、ポリ（酪酸ブチル）、ポリビニリデンジフルオライド（P V D F）膜、ガラス、気孔ガラス、磁気制御気孔ガラス、セラミック、金属などの合成または改変天然由来高分子；を含むことができ、それらをそれら自身、または他の材料と結合して使用できる。

20

【0032】

いくつかの実施形態において、オリゴヌクレオチドはアレイ形式で合成される。例えば、一本鎖のオリゴヌクレオチドは、各オリゴヌクレオチドが基板上で別々に、または、別個の機構（*discrete feature*）（または、スポット）で合成される、一般的な担体上で*in situ*で合成される。好ましい実施形態において、一本鎖オリゴヌクレオチドは担体または機構の表面に結合する。本明細書で使用されるように、「アレイ」なる用語は更なる反応のためにオリゴヌクレオチドまたは相補的オリゴヌクレオチドを格納し、送達し、増幅し、放出するための離散的な機構の配置のことをいう。好ましい実施形態において、担体またはアレイがアドレス可能であり：担体は担体上の特定の所定位置（すなわち、「アドレス」）の2つ以上の離散的なアドレス可能な機構を含む。それゆえ、アレイの各オリゴヌクレオチド分子は、担体上の既知および定義された位置に局所化される。各々のオリゴヌクレオチド配列は担体上のその位置から決定できる。さらに、アドレス可能な担体またはアレイは、液滴などの個別に分離された容量の直接制御を可能にする。機構上の、各小滴がお互いからの分離を保たれている、マイクロ容量の小滴の形成を許容するために、定義された機構のサイズを選ぶことができる。本明細書に記載されるように、必須というわけではないが、機構は典型的に、2つの隣接する機構の間の小滴が合わされることを確実にするために、機構間（*interfere*）の空間によって分離されている。機構間は、典型的に、それらの表面にいかなるオリゴヌクレオチドも保持せず、不活性な空間に相当するであろう。いくつかの実施形態において、機構と機構間はそれらの親水性または疎水性の特性において異なり得る。いくつかの実施形態において、機構と機構間は、本明細書に記載されるように、修飾因子を含み得る。

30

【0033】

アレイは、構築されるか、特注されるか、または業者（例えば、CombiMatrix, Agilent, Affymetrix, Nimblegen）から購入され得る。オリゴヌクレオチドは、アレイ表面の別個の機構上に付着されるか、スポットされるか、固定化されるか、表面結合されるか、支持されるかまたは合成される。オリゴヌクレオチドは、表面に共有結合されるか、または表面に置かれ（deposit）得る。構築の様

40

50

々な方法は当分野で周知であり、例えば、マスクレスアレイシンセサイザ、マスク利用光指向性法 (light directed method)、流路法 (flow channel method)、スポット法などである。

【0034】

いくつかの実施形態において、構築および／または、選択オリゴヌクレオチドは、マスクレスアレイシンセサイザ (MAS) を用いて固相担体上で合成され得る。例えば、マスクレスアレイシンセサイザは、例えば PCT 出願番号 WO 99 / 42813 および対応する米国特許番号 6,375,903 に記載される。他の例で、アレイの各々の機構が所望の配列の一本鎖 DNA 分子を有するカスタム DNA マイクロアレイを作ることができるマスクレス機器が知られている。

10

【0035】

構築および／または選択オリゴヌクレオチドを合成する他の方法は、例えば、マスク利用光指向性法、流路法、スポット法、ピンを基にした方法 (pin-based method)、多重担体利用法 (method utilizing multiple support) などを含む。

【0036】

オリゴヌクレオチド合成のためのマスク利用光指向性法（例えば、VLSPS（登録商標）法）は、例えば、米国特許番号 5,143,854、5,510,270 および 5,527,681 に記載される。これらの方法は、固相担体の所望の領域を活性化し、ついで予備選択された単量体溶液で担体に接触させることを伴う。選択された領域は、集積回路製作に使用されるフォトリソグラフィー技術の手法で、マスクを通した光源での照射により活性化できる。担体の他の領域は、照射がマスクによって妨げられるため、不活性なままで残り、それらは化学的に保護されたままで残る。したがって、光のパターンは、サポートのどの範囲が特定の単量体と反応するかを定義する。所定の領域の異なるセットを繰り返し活性化し、担体を異なる単量体溶液に接触することにより、さまざまの一連の重合体が担体上に生成される。担体からの未反応単量体溶液の洗浄などの、他の段階が任意に使用できる。他の適用可能な方法は、米国特許番号 5,384,261 に記載されるような機械技術を含む。

20

【0037】

单一担体上での構築および／または選択オリゴヌクレオチドの合成に適切な追加の方法は、例えば、米国特許番号 5,384,261 に記載される。例えば、試薬は、(1) 所定の領域上に定義されたチャンネルないでの流れ、または(2) 所定の領域上の「スポットティング」のいずれかによって担体に送達され得る。また、スポットティングと流れの組み合わせだけでなく他のアプローチも用いられ得る。どの場合でも、単量体溶液が様々な反応部位に送達されるとき、担体のある一定の活性化領域は、他の領域と機械的に切り離される。流路法は、固相担体上のオリゴヌクレオチドの合成を制御するために、例えば、マイクロ流体システムを伴う。例えば、適切な試薬の流れを通じた、または適切な試薬が置かれた担体の表面上で流路を形成することにより、固相担体の選択された領域でさまざまな重合体配列が合成され得る。固相担体上のオリゴヌクレオチド生成のためのスポットティング方法は、比較的少量の反応剤を、それらを直接選択された領域に置くことにより送達することを伴う。いくつかの段階において、より効率的であるならば、溶液で担体表面全体を、スプレーするかまたは他の方法でコーティングできる。正確に測定された単量体溶液のアリコートは、領域から領域まで移動するディスペンサーによって滴下され得る。

30

【0038】

固相担体上でのオリゴヌクレオチド合成のためのピンを基にした方法は、例えば米国特許番号 5,288,514 に記載される。ピンを基にした方法は、多くのピンまたは他の伸長を有する担体を利用する。ピンはトレイ内の個々の試薬容器の中に、各々同時に挿入される。一般に、96 ピンのアレイが、96 ウェルのマイクロタイタープレートなどの 96 ウェルのトレイと共に利用される。各々のトレイは、個々のピン上で特定の化学反応で結合するための特定の試薬で満たされる。したがって、トレイはしばしば異なった試薬を

40

50

含むであろう。化学反応は、比較的同様の反応条件下でそれぞれの反応を実行できるよう に最適化されるため、同時に複数の化学的結合段階を行うことが可能となる。

【0039】

オリゴヌクレオチドを合成するための他の適当なマイクロアレイおよび方法は、その全体が参照により本明細書に組み込まれる、米国特許番号7,323,320および7,563,600に記載されたものを全体として含む。例では、そこから合成されたオリゴヌクレオチドは、化学的、酵素的または物理的に切断されるかまたは他の方法で、さらなる増幅、制限酵素での消化および／またはアセンブリのためにマイクロアレイから放出される。

【0040】

他の実施形態において、複数のオリゴヌクレオチドが、ビーズなどの多重担体に合成されるか、または固定化（例えば、付着）され得る。一例は、例えば米国特許番号5,770,358、5,639,603および5,541,061に記載されるビーズを基にした合成法である。ビーズ上のオリゴヌクレオチドなどの分子の合成のため、大きな多数のビーズは容器内の適当な担体（水など）で懸濁される。ビーズは、任意に保護基と複合された活性部位を有する、任意のスペーサー分子とともに提供される。合成の各々の段階で、ビーズは、多数の容器への結合のために分割される。初期のオリゴヌクレオチド鎖が脱保護された後に、同じヌクレオチド添加反応が特定の容器の中のすべてのビーズ上に起こるように、異なる単量体溶液が各容器に加えられる。ビーズはその後、合成の次のラウンドに備えて、過剰の試薬を洗浄され、单一の容器にプールされ、混合され、他の多数の容器に再分配される。最初に利用された多くのビーズによって、塩基の無作為化された付加の無数の周回の後に、各々がその表面上に合成された一意的なオリゴヌクレオチド配列を有する、容器内に無作為で分散された多くのビーズが同様にあることに注意すべきである。個々のビーズは使用の間、識別を考慮するためにその上に二本鎖オリゴヌクレオチドに特有の配列でタグ付けされ得る。

10

20

30

40

【0041】

さらに他の実施形態において、複数のオリゴヌクレオチドが、ナノ粒子上に付着されるか、合成され得る。ナノ粒子は金属（例えば、金、銀、銅、およびプラチナ）、半導体（例えばCdSe、CdS、およびZnSでコーティングされたCdS）、および磁性（例えば、強磁性体）コロイド物質を含むが、これらに限られない。オリゴヌクレオチドをナノ粒子に付着させるための方法は当分野で知られている。他の実施形態において、ナノ粒子は基板に付着される。固定化オリゴヌクレオチドの有無にかかわらず、ナノ粒子は例えば、Grabar et al., Analyt. Chem., 67, 73-743 (1995); Bethell et al., J. Electroanal. Chem., 409, 137 (1996); Bar et al., Langmuir, 12, 1172 (1996); Colvin et al., J. Am. Chem. Soc., 114, 5221 (1992)に記載されるように基板に付着できる。そのままのナノ粒子がまず基板に結合し得、その後、固定されたナノ粒子にオリゴヌクレオチドが付着できる。

【0042】

予備合成されたオリゴヌクレオチドおよび／またはポリヌクレオチド配列は、その全体が参照により全ての目的において本明細書に組み込まれる、以下の文献に示される光指向性法、流路法およびスポット法、インクジェット法、ピンを基にした方法およびビーズを基にした方法を用いて、in situで基板に付着または合成され得る：McGall et al. (1996) Proc. Natl. Acad. Sci. U.S.A. 93:13555; Synthetic DNA Arrays In Genetic Engineering, Vol. 20:111, Plenum Press (1998); Duggan et al. (1999) Nat. Genet. S21:10; Microarrays: Making Them and Using Them In Microarray Bioinforma

50

t i c s , C a m b r i d g e U n i v e r s i t y P r e s s , 2 0 0 3 ; 米国出願公開番号 2 0 0 3 / 0 0 6 8 6 3 3 および 2 0 0 2 / 0 0 8 1 5 8 2 ; 米国特許番号 6 , 8 3 3 , 4 5 0 , 6 , 8 3 0 , 8 9 0 , 6 , 8 2 4 , 8 6 6 , 6 , 8 0 0 , 4 3 9 , 6 , 3 7 5 , 9 0 3 および 5 , 7 0 0 , 6 3 7 ; および P C T 公開番号 W O 0 4 / 0 3 1 3 9 9 , W O 0 4 / 0 3 1 3 5 1 , W O 0 4 / 0 2 9 5 8 6 , W O 0 3 / 1 0 0 0 1 2 , W O 0 3 / 0 6 6 2 1 2 , W O 0 3 / 0 6 5 0 3 8 , W O 0 3 / 0 6 4 6 9 9 , W O 0 3 / 0 6 4 0 2 7 , W O 0 3 / 0 6 4 0 2 6 , W O 0 3 / 0 4 6 2 2 3 , W O 0 3 / 0 4 0 4 1 0 および W O 0 2 / 2 4 5 9 7 。いくつかの実施形態において、予備結合されたオリゴヌクレオチドは、担体に付着されるか、または単量体溶液が領域から領域に移動するディスペンサーによって滴下されるスポット方法（例えば、インクジェット）を使用することで結合される。いくつかの実施形態において、オリゴヌクレオチドは、例えば力学的な波が作動するディスペンサーを用いて担体上にスポットされる。

10

【 0 0 4 3 】

ある配列を有するように設計されたオリゴヌクレオチドの調製は、誤り（例えば、それは少なくとも 1 つの位置において設計された配列と異なる）を含むオリゴヌクレオチド分子に加えて設計された配列を有するオリゴヌクレオチド分子を含み得る。配列誤りは 1 つ以上のヌクレオチド欠失、付加、置換（例えば、塩基転換または転移）、反転、重複、またはこれらのいかなる 2 つ以上の組み合わせを含み得る。オリゴヌクレオチド誤りはオリゴヌクレオチド合成の間に発生し得る。異なる合成技術は異なる誤りプロフィールと頻度の傾向を有し得る。いくつかの実施形態において、誤り率は使用された合成プロトコルによって、塩基あたり 1 / 1 0 から 1 / 2 0 0 までの誤りに変化し得る。しかしながら、いくつかの具体化において、より低い誤り率が達成され得る。また、誤りの型は、使用された合成技術に依存し得る。例えば、いくつかの実施形態において、チップを基にしたオリゴヌクレオチド合成は、カラムを基にした合成技術より比較的多くの欠失をもたらし得る。

20

【 0 0 4 4 】

いくつかの実施形態において、1 つ以上のオリゴヌクレオチド調製品が、誤りを含むオリゴヌクレオチドを除去する（またはその頻度を減少させる）誤り減少または誤りろ過の工程に供され得る。そのような工程は、オリゴヌクレオチド調製品における、誤りのないオリゴヌクレオチドの数を増加させるために使用できる。誤り減少または誤りろ過を行うための方法は、例えば選択オリゴヌクレオチドのハイブリダイゼーション、ミスマッチ結合剤、ミスマッチ結合蛋白質またはそれらの組み合わせとの結合を含むことができる。

30

【 0 0 4 5 】

いくつかの実施形態において、ハイブリダイゼーション技術は、オリゴヌクレオチド調製品（すなわち、構築オリゴヌクレオチド）がストリンジエントな条件で、相補的配列を有するように設計された固定化オリゴヌクレオチド調製品（すなわち、選択オリゴヌクレオチド）とハイブリダイズされるところで用いられ得る。本明細書で使用される、「選択オリゴヌクレオチド」なる用語は、構築オリゴヌクレオチド（または、構築オリゴヌクレオチドの相補体）の少なくとも一部と相補である一本鎖オリゴヌクレオチドを意味する。選択オリゴヌクレオチドは、構築オリゴヌクレオチドのプールから配列誤り（例えば、所望の配列からの逸脱）を含む構築オリゴヌクレオチドのコピーを除去するために使用され得る。いくつかの実施形態において、選択オリゴヌクレオチドは、基板上に末端固定され得る。また、他の実施形態において、選択オリゴヌクレオチドは溶液内となることができる。一の実施形態において、選択オリゴヌクレオチドは、本明細書に開示されるように、基板と平行に合成された合成オリゴヌクレオチドであり得る。

40

【 0 0 4 6 】

結合しないまたは不安定な二本鎖を形成する構築オリゴヌクレオチドは、使用条件下ハイブリダイゼーションを不安定化させるであろう誤りを含むオリゴヌクレオチドを選択的にまたは特異的に除去するために、除去され得る。いくつかの誤りを含むオリゴヌクレオチドがこの選択工程を通して十分な親和力固定化選択オリゴヌクレオチドに結合し得る

50

ので、この工程すべての誤りを含むオリゴヌクレオチドを除去し得るわけではないことが理解されるべきである。例えば、誤りを含むオリゴヌクレオチドは、1または2個の塩基によって選択オリゴヌクレオチドと異なり得て、選択工程反応条件下で選択オリゴヌクレオチドとまだ結合し得る。

【0047】

いくつかの実施形態において、核酸結合タンパク質またはリコンビナーゼ（例えば、R e c A）が、誤りのないオリゴヌクレオチドの選択を改良するためにオリゴヌクレオチド処理工程の1つ以上に含まれ得る。例えば、固定化オリゴヌクレオチドと完全に相補的なオリゴヌクレオチドのハイブリダイゼーションを選択的に促進することにより、結合した誤りを含むオリゴヌクレオチドの量を減少させ得る。結果として、本明細書に記載されたオリゴヌクレオチド処理手順は、誤りを含むオリゴヌクレオチドをより除去し、低い誤り頻度（例えば、塩基当たりの誤りが1/50未満、1/100未満、1/200未満、1/300未満、1/400未満、1/500未満、1/1,000未満、または1/2,000未満の誤り率）を有するオリゴヌクレオチド調製物を生成し得る。

10

【0048】

いくつかの実施形態において、誤り修正は、所望の配列からの逸脱なしに合成ポリヌクレオチドの相対的な数を増加させるように各々の工程の反復および合成工程の終わりの間に含まれ得る。そのような誤り修正は、直接配列決定および／または誤り修正ヌクレアーゼ（例えば、C E L I）などの修正酵素を基にした誤り修正、M u t S またはM u t S ホモログ結合または他のミスマッチ結合タンパク質を基にした誤り修正（例えば、国際出願番号P C T / U S 2 0 1 0 / 0 5 7 4 0 5）、当分野で知られる他の誤り修正手段、またはそれらのいかなる組み合わせを含み得る。例示的な実施形態において、C E L Iは液状媒体中のオリゴヌクレオチド二本鎖に添加され得る。C E L Iは、単一塩基多型、小規模の挿入または欠失などのすべての型のミスマッチを切断する、ミスマッチ特異的エンドヌクレアーゼである。エンドヌクレアーゼの添加は、ミスマッチの部位または領域で二本鎖オリゴヌクレオチドの切断をもたらす。

20

【0049】

最適化手順が、ヘテロ二本鎖の生産および除去を経て誤り含有核酸を除去することを伴うため、1つ以上の核酸結合タンパク質またはリコンビナーゼが、合成後忠実性最適化技術（例えば、M u t S またはM u t S ホモログを用いるスクリーニング技術）に好ましくは含まれないことが理解されるべきである。したがって、好ましくは、合成後、忠実性最適化の前に合成段階に含まれるいかなる核酸結合タンパク質またはリコンビナーゼ（例えば、R e c A）が、好ましくは除去される（例えば、不活性化、カラム精製または他の適当な技術により）。

30

【0050】

ある実施形態において、1つ以上の修飾オリゴヌクレオチドを含むことは、有用であり得る。オリゴヌクレオチドは、合成の間に改変された塩基（例えば、ヌクレオチドアナログ）を組み込むことにより、合成後にオリゴヌクレオチドを改変することにより、またはそれらのいかなる組み合わせにより、修飾され得る。修飾の例は：ニトロインドール、d P およびd K、イノシン、ウラシルなどのユニバーサル塩基；B r d Uなどのハロゲン化塩基；蛍光ラベル塩基；ビオチン（d T 誘導体として）およびジゴキシゲニン（D I G）などの非放射性標識；2,4-ジニトロフェニル（D N P）；放射性ヌクレオチド；d R - N H 2（デオキシリボース-N E b）などのカップリング後修飾；アクリジン（6-クロロ-2-メトキシアクリジン）；およびC 3、C 8（オクタンジオール）、C 9、C 12、H E G（ヘキサエチレングリコール）およびC 18などの配列にスペ-サー「アーム」を加えるために合成の間に使用されるスペ-サーホスホアミド、などの1つ以上を含むが、これらに限定されるものではない。

40

【0051】

オリゴヌクレオチド増幅

オリゴヌクレオチドは、一本鎖合成品として、提供または合成され得る。いくつかの実

50

施形態において、オリゴヌクレオチドは、アニーリングされた相補鎖を含む二本鎖調製品として提供または合成され得る。オリゴヌクレオチドは、DNA、RNA、PNAまたはそれらのいかなる組み合わせの分子であり得る。二本鎖オリゴヌクレオチドは、一本鎖合成オリゴヌクレオチドまたは他の適当なテンプレート（例えば、核酸ベクターまたはゲノム核酸などの核酸調製品における配列）を増幅することによって、生成され得る。従って、本明細書に記載される配列特性を有するように設計された複数のオリゴヌクレオチドは、これらの特性を有する複数の一本鎖オリゴヌクレオチドとして提供されるか、または相補的オリゴヌクレオチドと共に提供され得る。いくつかの実施形態において、オリゴヌクレオチドはリン酸化（例えば、5' リン酸）され得る。いくつかの実施形態において、オリゴヌクレオチドは非リン酸化であり得る。

10

【0052】

いくつかの実施形態において、オリゴヌクレオチドは、オリゴヌクレオチドの各々の末端に対応する一のプライマーをもつ適切なプライマー対（例えば、1つがオリゴヌクレオチドの3' 末端に相補的であり、1つがオリゴヌクレオチドの5' 末端と同じもの）を用いて増幅され得る。いくつかの実施形態において、オリゴヌクレオチドは、5' 増幅配列（例えば、5' ユニバーサル配列）および／または3' 增幅配列（例えば、3' ユニバーサル配列）によって隣接された（f l a n k e d）主要なアセンブリ配列（標的核酸に組み込まれるように設計される）を含むように設計され得る。フランкиング増幅配列に対応する増幅プライマー（例えば、10から50の間のヌクレオチド長、15から45の間のヌクレオチド長、約25ヌクレオチド長など）が、オリゴヌクレオチドを増幅するために使用され得る（例えば、1つのプライマーが3' 増幅配列に相補的であり得、1つのプライマーが5' 増幅配列に相補的であり得る）。増幅配列はその後、アセンブリ配列だけを含むオリゴヌクレオチドを生産するために、いかなる適当な技術を使用して増幅オリゴヌクレオチドから除去され得る。

20

【0053】

いくつかの実施形態において、異なる主要なアセンブリ配列をもつ複数の異なるオリゴヌクレオチド（例えば、約5、10、50、100、それ以上）が、5' 増幅配列と同一、および／または3' 增幅配列と同一であり得る。これらのオリゴヌクレオチドは、同じ増幅プライマーを使用する同じ反応において増幅できる。

30

【0054】

アセンブリ反応で使用される複数のオリゴヌクレオチドが、合成オリゴヌクレオチド、一本鎖オリゴヌクレオチド、二本鎖オリゴヌクレオチド、増幅産物、誤りを含む変異体を除去（または頻度を減少させる）するために処理されるオリゴヌクレオチド等の調製物、またはそれらの2つ以上のいかなる組み合わせを含み得る。いくつかの態様において、二本鎖増幅産物は、本明細書に記載のように、アセンブリオリゴヌクレオチドとして使用され、アセンブリ反応に加えられ得る。いくつかの実施形態において、オリゴヌクレオチドはそれがまだ指示体に付着しながら、増幅され得る。いくつかの実施形態において、オリゴヌクレオチドは増幅の前または増幅の後に支持体から除去されるか、切断され得る。

【0055】

いくつかの実施形態において、合成オリゴヌクレオチドは、5' および3' 増幅配列に隣接した主要なアセンブリ配列を含み得る。主要なアセンブリ配列は、組み立てられた標的核酸または標的アセンブリへの組み込みのために設計される。フランキング配列は、増幅のために設計され、組み立てられた核酸に組み込まれることを意図しない。フランキング増幅配列は、同じ増幅配列を共有するが、異なる主要なアセンブリ配列を有する複数の異なるアセンブリオリゴヌクレオチドを増幅するためにユニバーサルプライマー配列として使用され得る。いくつかの実施形態において、増幅の後にアセンブリ配列だけを含むオリゴヌクレオチドを生成するためにフランキング配列を除去する。

40

【0056】

ある実施形態において、二本鎖増幅産物は、フランキング配列を除去するために制限酵素での消化に供され得る。そのために、フランキング配列は、1つ以上の制限部位または

50

制限酵素認識部位を含むように設計される場合がある。除去されるフランキング配列と主要なアセンブリ配列の間に切断部位がある限り、制限部位は增幅配列の5'または3'末端に存在し得る。制限部位は增幅配列（すなわち、プライマー結合部位）に含まれ得る。制限部位はまた、增幅配列の外側にあり得る。

【0057】

制限酵素での消化の後に、いかなる適当な技術を使用することにより、切断されたフランキング配列が分離され、除去され得る。いくつかの実施形態において、切断されたフランキング配列は、約40、約35、約30、約25、約20、または約15個塩基長未満であり得る。そういうものとして、フランキング配列よりも長いサイズとなるように設計できる主要なアセンブリ配列から切断されたフランキング配列を分離するために、当分野で知られているサイズ依存分離技術、例えば、シリカとの異なる親和性、サイズろ過、PEG（ポリエチレングリコール）またはCTAB（臭化セチルトリメチルアンモニウム）での沈殿、またはそれらのいかなる組み合わせが、使用され得る。

10

【0058】

いくつかの実施形態において、增幅プライマーはビオチン化され得る。その結果、得られた增幅産物もまた両端でビオチン化されることとなる。制限酵素での消化の際に、ビオチン化プライマーを有する切断されたフランキング配列は、主要なアセンブリ配列が非ビオチン化されている一方で、ビオチントグを保有する。それゆえ、切断されたフランキング配列は、ストレプトアビジンを使用して親和性精製および除去できる（例えば、ビーズ、カラム、または他の表面に結合する）。いくつかの実施形態において、増幅プライマーもまた、増幅の後にフランキング増幅配列のないアセンブリ配列を含む二本鎖アセンブリ断片を生成するために、増幅後のプライマー領域を除去するために使用できるある配列特性（すなわち、制限部位）を含むように設計され得る。

20

【0059】

一本鎖突出

本発明のある態様は、一本鎖突出をもつ二本鎖核酸に関連する。突出は、いかなる適当な技術を使用することにより生成し得る。いくつかの実施形態において、二本鎖核酸断片（例えば多重アセンブリで組み立てられた断片）、は適切な制限酵素で消化され、末端の一本鎖突出を生成し得る。いくつかの実施形態において、組み立てられた産物内でお互いに隣接するように設計されている断片は、同じ酵素によって、相補的な突出を露出するよう消化され得る。相補的な突出を生成する異なる酵素もまた使用され得る。

30

【0060】

いくつかの実施形態において、突出は、IIS型制限酵素を使用して生成され得る。IIS型制限酵素は、認識部位と呼ばれる1つの部位の二本鎖核酸に結合する酵素であり、認識部位の外で单一の二本鎖切断を作成する。切断部位と呼ばれる二本鎖切断は、通常、認識部位から0-20塩基離れて位置する。認識部位は通常、約4~8bp長である。すべてのIIS型制限酵素は少なくとも部分的に非対称な認識を示す。非対称な認識は、核酸の各々の鎖において、5'、3'認識配列が異なることを意味する。酵素活性もまた、切断部位が認識部位の片面だけに位置することを意味する極性を示す。それゆえ、一般に、それぞれの認識部位に対応するただ1つの二本鎖切断がある。いくつかの酵素は平滑末端を生成するが、一般に、切断は5'または3'末端で1-6ヌクレオチドの一本鎖の突出を生成する。いくつかの場合に一本鎖の突出を生成するものが生産されるが、いずれの切断も本発明の文脈において有用である。これまで、約80のIIS型酵素が同定されている。好適な例として、BstF5 I、BtsC I、BsrD I、Bts I、Alw I、Bcc I、BsmA I、Ear I、Mly I（平滑）、Ple I、Bmr I、Bsa I、BsmB I、BspQ I、Fau I、MnI I、Sap I、Bbs I、BciV I、Hph I、Mbo II、BfuA I、Bsp CN I、BspM I、SfaN I、Hga I、BseR I、Bbv I、Eci I、Fok I、BceA I、BsmF I、BtgZ I、BpuE I、Bsg I、Mme I、BseG I、Bse3D I、BseM I、AclW I、A

40

50

l w 2 6 1、B s t 6 1、B s t M A I、E a m l 1 0 4 1、K s p 6 3 2 I
 、P p s I₅ S c h Iは(平滑)、B f i I、B s o 3 1 1、B s p T N I
 、E c o 3 1 I、E s p 3 I、S mu I、B f u I、B p i I、B p u A I、
 B s t V 2 I、A s u H P I、A c c 3 6 I、L w e I、A a r I、B s e M
 I I、T s p D T I、T s p G W I、B s e X I、B s t V 1 I、E c o 5 7
 1₅ E c o 5 7 M I₅ G s u I₅ およびB e g I Iが挙げられるが、これらに
 限定されるものではない。いくつかの実施形態において、B s a I、B s m B I、B
 s p Q I、B t g Z I、B s m F I、F o k I、B b v I、それらのいかなる
 变異体、またはそれらのいかなる組み合わせも使用できる。そのような酵素およびそれら
 の認識および切断部位は、N e w E n g l a n d B i o l a b sなどの供給業者から
 利用可能である。

10

【0061】

いくつかの実施形態において、アセンブリのために設計された複数の核酸断片の各々が、各端にI I S型制限部位を有し得る。I I S型制限部位は、切断部位が認識配列に対して内部となるように配置され得る。結果として、酵素消化は、内部配列(例えば、内部配列内の突出)を露出して、末端から認識配列を除去する。したがって、同じI I S型部位は、アセンブリのために調製される核酸断片すべての両端に使用され得る。しかしながら、異なるI I S型部位も使用され得る。それぞれ組み立てられた産物中で隣接するように設計された2つの断片が、同一の重複末端配列および、制限酵素消化の際に重複配列内で相補的な突出を露出するために適切に配置されるフランキングI I S型部位を含み得る。したがって、複数の核酸断片が異なる相補的な突出と共に生成され得る。核酸断片の各端の制限部位は、適切なI I S型酵素の消化が制限部位を除去し、組み立てられた核酸産物内で隣接するように設計された核酸断片の一本鎖領域に相補的な一本鎖領域を露出するように位置し得る。ある実施形態において、制限酵素はアセンブリ核酸断片に対応する制限部位がないように選択され得る。

20

【0062】

上述のとおり、制限部位は、增幅配列の内側または外側、5'または3'へ配置できる。図9Aに示すとおり、制限部位(太字で示す)を、増幅配列(イタリック体で示す)内および主要なアセンブリ断片(黒)の遠位に含むことができる。例として、B t g Z IおよびB s m F I部位は二本鎖アセンブリ断片のいずれの末端でも使用され、それらの各々の切断部位は矢印で示される。B t g Z IおよびB s m F Iの両方が、それらの認識部位から10ヌクレオチド/14ヌクレオチド離れて切断する。認識部位から短距離(例えば、5~25、10~20、または約15ヌクレオチド)を切断する他の制限酵素もまた使用できる。あるいはまた、図9Bに示すように、制限部位(太字で示す)は、増幅配列(イタリック体で示す)の外側および主要なアセンブリ断片(通常のフォント)の近位であることができる。B s a I部位は例えば、その部位が矢印によって示される切断部位として、二本鎖アセンブリ断片の両端で使用される。図の9Aおよび9Bに示すように、制限部位が主要なアセンブリ断片の遠位に位置し、増幅配列内に含まれるとき、出発核酸の全長は、制限部位が主要なアセンブリ断片の近位に位置し、増幅配列に含まれていない場合より短い。したがって、第一の戦略(図9A)は、より短い出発核酸を合成する場合に(例えば、チップで)、より費用効率が良く、より少ない誤りとすることができます。第一の戦略は、また、より短いユニバーサルプライマー(断片を増幅するための)を使用し、その結果、さらにコストを削減する。制限酵素での消化の後に、主要なアセンブリ断片から除去されるべき末端部はまた、より短く、その結果、第二よりも、第一の戦略においてより簡単に、安く、速く除去される。

30

【0063】

I I S型または他の部位特異的制限酵素でのD N Aの酵素消化は、典型的に4~6個のヌクレオチドの突出を生成する。本発明で示される、これらの短い付着末端が、強敵核酸に相補的な末端を含む複数の核酸断片をライゲーションするために十分であることは、予期できないものである。3個以上の断片でライゲーション効率が有意に減少するとして、

40

50

慣習上、効率を確実にするために、ライゲーション反応は典型的に2個の断片を含む。加えて、より長い付着末端は、ミスマッチがしばしば起こるのに応じて特異性を改良するために従来の方法において必要とされる。その上、正しいライゲーション産物、労働集約的で時間のかかるクローニング、および選別のために過程を選択することが必要とされる。

【0064】

本発明は、とりわけ(1)单一反応(例えば、単一のプール)における、複数の断片(例えば、少なくとも4、少なくとも5、少なくとも6、少なくとも7、少なくとも8以上)の成功したライゲーション、(2)迅速で安価なライゲーション反応(例えば、室温30分)、(3)ミスマッチを区別する高特異性、および(4)クローニングおよびスクリーニングを必要としないクイックPCR段階、を提供する。本発明の他の利点は、いかなる配列および/またはいかなる長さ(例えば、少なくとも500bp、少なくとも1kb、少なくとも2kb、少なくとも5kb、少なくとも10kb、またはより長いもの)となることができる、市販のチップまたはマイクロアレイの合成オリゴヌクレオチドを直接利用できることである。そのような合成オリゴヌクレオチドは、実質的に同じサイズ(例えば、約50塩基、約100塩基、約200塩基、300以上の塩基)であって、その結果、扱いの容易さを提供できる。

【0065】

一例では、チップ上の各オリゴヌクレオチドまたは断片が100ヌクレオチドのペイロード(pay load)を有し、断片が4'-塩基の突出を有すると仮定すると、断片の数をnとして、ライゲーション産物長=(n*100)-(4*(n-1))、(n-1)ライゲーション接合部分(ligation junction)、である。ライゲーションの特異性を確実にするために、突出が各々のライゲーション部位に特異的になるように選択または設計することができる、すなわち、組み立てられた産物内で隣接するように設計された2つの断片の相補的な突出の各々の対が、少なくとも1つのヌクレオチドにより、相補的な突出のいかなる他の対とも異なっているはずである、ことに注意すべきである。

【0066】

付着末端を露出させるための他の戦略(オフセットアセンブリ)は図10Aに示される。チップから始まり、複数のオリゴ(例えば、A₁-A₁₀)が合成できる。オリゴは、適切に組み立てられたときに、主要なアセンブリ配列を有するように設計でき、標的核酸5'-A₁-A₃-A₅-A₇-A₉-3'(逆行鎖(reverse strand))は3'-A₂-A₄-A₆-A₈-A₁₀-5')を形成する。すなわち、2つの隣接オリゴヌクレオチドA_nおよびA_{n+1}が、重なるように設計できる。本明細書で使用されるように、隣接オリゴヌクレオチドは、直線核酸配列に沿って第一のオリゴヌクレオチドが第二のオリゴヌクレオチドの5'末端または3'末端にあるオリゴヌクレオチドを意味する。いくつかの実施形態において、隣接オリゴヌクレオチドは連続(contiguous)となる。本明細書で使用されるように、連続オリゴヌクレオチドは、直線核酸配列に沿って、第一のオリゴヌクレオチドが任意に-1に設定される位置で終わり、第二の断片が任意に0に設定される位置で開始する2つのオリゴヌクレオチドを意味する。主要なアセンブリ配列は、約50~500ヌクレオチド、約60~300ヌクレオチド、約70~200ヌクレオチドなどの、またはそれより短いまたは長い、いかなる所望の長さとすることができます。複数のオリゴは扱いの容易さのために均一長を有することができます。一例として、合成オリゴはまた、いずれかの末端にも増幅配列を含むことができ、それは、制限部位を組み込むことができる。増幅配列は、約10~30ヌクレオチド、約15~25ヌクレオチド、またはそれより短いまたは長くできる。図10Aは70-マーの主要なアセンブリ配列および120-マーの全長のオリゴを示す。合成オリゴは、チップから溶出、切断または他の方法で放出され、プライマー対ALおよびARを用いてPCR増幅に供することができる。増副産物は、増幅配列(矢じり)を除去する切断(例えば、制限酵素で)され、主要な70-マーの二本鎖のアセンブリ配列をそこから精製できる。これらの二本鎖アセンブリ配列は次いで、単一のシャッフリング段階で溶解され(例えば、9

10

20

30

40

50

5 で)、再アニーリングできる(例えば、65 で)。一本鎖オリゴヌクレオチドのシャッフルリングの後に、製品の25%は付着末端を有するオフセットアセンブリ産物(例えば、A₁/A₂、A₂/A₃、A₃/A₄、A₄/A₅など)になるであろう。これらの付着末端をリガーゼを用いて一緒に組み立てる(徐々にまたは単一反応で階層的に)ことができ、その結果、標的核酸5'-A₁-A₃-A₅-A₇-A₉-3'(逆行鎖は3'-A₂-A₄-A₆-A₈-A₁₀-5')を形成できる。オリゴはまた、標的核酸が5'-A₁...A₃...A₅...A₇...A₉-3'(すなわち、テンプレートとしてA_{n+1}配列を使用することで満たすことができるギャップがA_nおよびA_{n+2}の間に許容される)となるように設計できる。そのために、ポリメラーゼおよびdNTPをライゲーションの前にギャップをひろげ、埋めるために使用できる。

10

【0067】

第二のオフセットアセンブリ戦略は図10Bで示され、別々の2段階(すなわち、アセンブリ段階およびライゲーション段階)とは対照的に、単一の組み合わせられたアセンブリ-(拡張)-ライゲーション段階が使用され得る。例えば、シャッフルリング段階(例えば、95 で溶解し、65 で再アニーリングする)の後、ギャップのない解析オリゴヌクレオチド(gapless parse oligonucleotide)が、完全長産物または部分組立品を形成するためにライゲーションできる。ギャップが解析中に存在するならば、オリゴヌクレオチドはポリメラーゼおよびdNTPの存在下で培養され、ライゲーションの前の鎖延長によりギャップを埋めることができる。いくつかの実施形態において、ギャップ化された解析をポリメラーゼ鎖延長およびライゲーションに同時に供することができる。本明細書で使用されるように、用語「部分アセンブリ(subassembly)」は構築オリゴヌクレオチドの組から組み立てられた核酸分子をいう。好ましくは、部分アセンブリ構築オリゴヌクレオチドよりも少なくとも約2倍、3倍、4倍、5倍、10倍、20倍、50倍、100倍、またはそれ以上長い。

20

【0068】

付着末端を精製する他の方法もまた使用できる。例えば、所望の付着末端を合成するために、ポリメラーゼを基にした方法(例えば、T4 DNAポリメラーゼ)を使用できる。特異的な突出(例えば、組み立てられた核酸産物中で隣接するように設計された核酸に相補的な突出)を生成する方法にかかわらず、異なる長さの突出が、設計および/または生成され得る。いくつかの実施形態において、長い一本鎖の突出(3'または5')は、特異的および/または効率的なアセンブリを促進するために使用され得る。例えば、3'または5'の一本鎖の突出は8塩基より長い、例えば、8~14、14~20、20~25、25~50、50~100、100~500、または、より長い塩基長であり得る。

30

【0069】

高忠実度アセンブリ

本発明の態様によると、複数の核酸断片が、複数の断片が特定のより長い核酸を生成する、断片の共有結合アセンブリを促進する条件下で複数の断片が共に混合される、单一の手順で組み立てられ得る。本発明の態様によると、複数の核酸断片が、リガーゼを使用することにより、in vitroで共有結合を介して組み立てられ得る。いくつかの実施形態において、5以上(例えば、10以上、15以上、15~20、20~25、25~30、30~35、35~40、40~45、45~50、50以上など)の異なる核酸断片が組み立てられ得る。しかしながら、いかなる数の核酸(例えば、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20など)が適当なアセンブリ技術を使用することで組み立てられ得ることが理解されるべきである。組み立てられる各々の核酸断片は、約100ヌクレオチド長および約1,000ヌクレオチド長の間(例えば、約200、約300、約400、約500、約600、約700、約800、約900)あり得る。しかしながら、より長い(例えば、約2,500以上のヌクレオチド長、約5,000以上のヌクレオチド長、約7,500以上のヌクレオチド長、約10,000以上のヌクレオチド長など)または、より短い核酸断片が、アセンブリ技術(例えば、プラスミドベクターへのショットガンアセンブリ)を使用する

40

50

ことにより組み立てられ得る。各々の核酸断片のサイズがアセンブリに追加された他の核酸断片のサイズから独立し得ることが理解されるべきである。しかしながら、いくつかの実施形態において、各々の核酸断片は、ほぼ同じサイズまたは長さ（例えば、約100のヌクレオチド長および約400のヌクレオチド長の間）であり得る。例えば、オリゴヌクレオチドの長さは、約+/-1ヌクレオチド、+/-4ヌクレオチド、+/-10ヌクレオチドで変化する約100ヌクレオチド長および約400ヌクレオチド長の間の中央値を有し得る。二本鎖核酸断片の長さが塩基対の数によって示され得ることが理解されるべきである。本明細書で使用されるように、二本鎖核酸断片の文脈で使用されるとき、「x」ヌクレオチド長と称される核酸断片は、長さにおいて「x」塩基対に対応する。いくつかの実施形態において、1回の反応で組み立てられた1つ以上の核酸（例えば、1~5、5~10、10~15、15~20など）は、コドン最適化され得、および/または、非天然に生じ得る。いくつかの実施形態において、1回の反応で組み立てられる核酸のすべてが、コドン最適化され、および/または、非天然に生じる。

10

20

30

40

50

【0070】

本発明のいくつかの態様において、組み立てられる核酸断片は、重複する相補的配列を有するように設計される。いくつかの実施形態において、核酸断片は、3'および/または5'一本鎖突出をもつ二本鎖核酸断片である。これらの突出は、異なる核酸断片の相補的な付着末端にアニールできる付着末端であり得る。本発明の態様によると、2個の核酸断片の相補的配列（および、特に相補的付着性末端）の存在はそれらの共有結合アセンブリを促進する。いくつかの実施形態において、異なる重複した相補的一本鎖付着末端をもつ複数の核酸断片が組み立てられ、組み立てられた核酸産物における順番は、各々の断片の付着末端の同一性により決定される。例えば、核酸断片は、第一の核酸が、第二の核酸の第一の付着末端と相補的な第一の付着末端および第三の核酸の第一の付着末端と相補的な第二の付着末端を有するように、設計され得る。第二の核酸の第二の付着末端は、第四の核酸の第一の付着末端と相補的であり得る。第三の核酸の第二の付着末端は、第五の核酸の第一の付着末端と相補的であり得る。最終的な核酸をへるなど。本発明の態様によると、この技術は、所定の線形順序で組み立てられた核酸断片を含む直線的配置を生成するために使用され得る（例えば、第一、第二、第三、第四、...最終）。

【0071】

ある実施形態において、隣接している核酸断片の間で重複する相補的な領域は、核酸断片の一意的なアラインメント（例えば、断片の選択または設計されたアラインメント）のアセンブリを促進できる程度（例えば、熱力学的な支持）に異なるように設計（または、選択）される。驚くべきことに、適切なライゲーション条件下で、ただ1つのヌクレオチドによる違いは、完全な適合（100%相補的な付着末端）およびミスマッチ（100%未満の相補的な付着末端）の間を十分に区別する力を提供する。そして、4-塩基の突出は、 $(4^4 + 1) = 257$ までの異なる断片が高特異性および高忠実性でライゲーションされることを許容できる。

【0072】

異なる長さの重複領域が使用され得ることが理解されるべきである。いくつかの実施形態において、より大きい数の核酸断片が組み立てられとき、より長い付着末端が使用され得る。より長い付着末端は、正しい付着末端アニーリング（例えば、お互いにアニーリングされるように設計された付着末端に関する）および不正確な付着末端アニーリング（例えば、非相補的な付着末端の間）を区別する、十分に区別可能な配列をより柔軟に設計または選択することを提供し得る。

【0073】

このような高忠実性アセンブリを達成するために、1つ以上の好適なリガーゼが使用され得る。リガーゼは組換えまたは天然資源から得られ得る。いくつかの実施形態において、T3 DNAリガーゼ、T4 DNAリガーゼ、T7 DNAリガーゼおよび/またはE. coli DNAリガーゼが使用され得る。これらのリガーゼは比較的低温（例えば、室温）で使用され得、比較的短い突出（例えば、約3、約4、約5または約6塩基突出

)に特に有用であり得る。あるライゲーション反応(例えば、室温で30分のインキュベーション)において、T7 DNAリガーゼは他のリガーゼと比較して多重ライゲーション(multi-way ligation)に、より有効となる。Tth DNAリガーゼ; Pfu DNAリガーゼ; Taq リガーゼ、他のいかなる好適な熱安定リガーゼ、またはそれらのいかなる組み合わせなどの熱安定リガーゼもまた使用され得る。

【0074】

いくつかの実施形態において、2つ以上の相補的付着末端と異なる核酸断片の対は、類似または同一の付着末端を有する断片の比較的無作為な配列(および/または数)を含む産物のアセンブリを促進するために、同一または類似の配列を有するように設計または選択され得る。

10

【0075】

組み立てられるべき個々の断片の濃度の変化が不完全な中間構築物のアセンブリをもたらし得ることを理解すべきである。例えば、各々が適切な付着性突出を有するオリゴヌクレオチドA、B、C、D、E、Fを用いる標的核酸配列(A B C D E F)のアセンブリにおいて、個々の断片の濃度が等モル(例えば、A、B、およびCの濃度がD、E、およびFの濃度より大きい)でないならば、ライゲーションされていない中間産物の混合をもたらす、終端種(A BおよびB Cなど)を形成できる。不完全な中間構築物の形成を避けるために、個々の断片の少なくとも2つのプール(例えば、プール1:A、C、E、およびプール2:B、D、F)から標的核酸を組み立てることができる。いくつかの実施形態において、各々2つのプールが、第一のプールの各々の核酸断片が第二のプールの核酸配列の末端と相補的な末端を有する、複数の核酸断片を含む。いくつかの実施形態において、オリゴヌクレオチドの集団を少なくとも2つのプールに分けて、別々に各プールの中でオリゴヌクレオチドを増幅することによって、少なくとも2つのプールを形成できる。他の実施形態において、オリゴヌクレオチドを第一のオリゴヌクレオチドアレイから第一のプールに放出して(例えば、溶出、切断または増幅によって)、第二のオリゴヌクレオチドアレイのオリゴヌクレオチドを第二のプールに放出することによって、少なくとも2つのプールを形成できる。また他の実施形態において、本明細書に記載されるように、少なくとも2組の異なる増幅タグを使用することでオリゴヌクレオチド配列を増幅することにより、少なくとも2つの異なるプールを形成できる。例えば、オリゴヌクレオチドB、DおよびFを含む第二のプールは、第二のプール中に存在するオリゴヌクレオチドB、DおよびFのモル濃度が第一のプール中に存在するオリゴヌクレオチドA、CおよびEのモル濃度より低くなるように希釈できる。例えば、第二のプール中のオリゴヌクレオチドのモル濃度は、第一のプール中のオリゴヌクレオチドのモル濃度より約2倍、10倍、20倍、50倍、100倍、またはそれ以上低くし得る。2つのプールを混合し、ライゲーションした後に、得られた産物は所定の配列を有する標的核酸を含み、第一のプールで形成された過剰なオリゴヌクレオチドから分離できる。ある実施形態において、異なるモル濃度を有するオリゴヌクレオチド二量体のプールを形成することが好ましい場合もある。例えば、標的核酸配列A B C D E F G Hのアセンブリは、オリゴヌクレオチドA、B、E、Fを含む第一のプールおよびオリゴヌクレオチドC、D、G、Hを含む第二のプールといった、少なくとも2つの異なるプールを使用することにより実施できる。第二のプールは、オリゴヌクレオチドC、D、G、Hのモル濃度がオリゴヌクレオチドA、B、E、Fのモル濃度より低くなるように(例えば10倍または100倍)希釈され得る。適切な付着性突出末端を有するオリゴヌクレオチドは、第一のプール中の中間産物A BおよびE F、および第二のプール中の中間産物C DおよびG Hを形成するためにライゲーションできる。中間産物A B、C D、E F、G Hをライゲーション条件下で混合した後に、所定の配列を有する標的核酸配列を含む最終産物が過剰な二量体A BおよびE Fから分離できる。

20

20

30

40

【0076】

いくつかの実施形態において、核酸断片はリガーゼと混合され、インキュベートされる。付着末端の特定のアニーリングを促進する状態でのインキュベーションが、アセンブリ(例えば、正しいアセンブリ)の頻度を増加させ得ることが理解されるべきである。いく

50

つかの実施形態において、異なる付着末端は、同様の条件下で全ての断片の正しいアニーリングが促進されるように、類似の融点（例えば、互いに約5以内）を有するよう設計される。正しいアニーリングは、異なる温度において、使用される付着末端の長さに応じて促進され得る。いくつかの実施形態において、約4から約30の間のヌクレオチド長（例えば、約5、約10、約15、約20、約25または約30ヌクレオチド長の付着末端）の付着末端が使用され得る。インキュベーション温度は約20から約50の範囲（例えば、室温を含む）であり得る。しかしながら、より高いまたはより低い温度が使用され得る。インキュベーションの長さは、突出の長さ、突出の複雑さ、および共に混合される異なる核酸の数（それゆえ、異なった突出の数）に基づいて最適化され得る。インキュベーション時間もまた、アニーリング温度および混合物中の他の剤の有無に依存し得る。例えば、核酸結合タンパク質、および／または、リコンビナーゼが添加され得る（例えば、RecA、例えば、熱安定RecAタンパク質）。

10

【0077】

得られた核酸複合体は、1組の標的配列特異的プライマーの存在下、正しいライゲーション産物（すなわち、目標核酸）を増幅および選択するために、ポリメラーゼ連鎖反応に供され得る。あるいは、更なるコロニー選別のために得られた核酸複合体を適當なベクターにライゲーションして、宿主細胞を形質転換することができる。

【0078】

配列分析および断片設計および選択

本発明の態様は、標的核酸の配列を分析して、適切な付着末端（例えば、一本鎖の突出）を生成するために使用できる、標的核酸配列内での、領域の同定に基づいたアセンブリ戦略を設計することを含み得る。これらの領域は、標的核酸を生成させるように（例えば、1回の反応で）組み立てることができる核酸断片の両端を定義するために使用され得る。核酸断片は、その後提供するかまたは作成することができる（例えば、多重アセンブリ反応で）。核酸断片は、扱いが容易になるように、それらが比較的均一なサイズになるように選択できる（例えば、精製）。

20

【0079】

いくつかの実施形態において、核酸配列は、解析された二本鎖または一本鎖オリゴヌクレオチドの組を生成するコンピュータ利用方法で設計および／または分析できる。本明細書で使用されるように、用語「解析（parse）」は、標的核酸の配列が、たとえば一連の隣接オリゴヌクレオチド配列を同定するような、コンピュータ利用方法で描写されることを意味する。隣接するオリゴヌクレオチドまたは核酸断片は、好ましくは本発明の方法によるアセンブリを容易にするために適切な数のヌクレオチドで重複する。オリゴヌクレオチド配列は、本発明の方法を用いて個別に合成および組立てられる。

30

【0080】

いくつかの実施形態において、標的核酸配列は、標的核酸の一本鎖上の少なくとも1つの異なるヌクレオチドを含む領域を同定するために分析され得る。これらの領域は、付着末端を生成するために使用され得る。好ましくは、付着末端の長さが特異性を提供するために十分であることが理解されるべきである。例えば、付着末端は、類似の付着末端との間の対誤り（mispair）を防ぐか、または減少させるために十分異なった配列（例えば、少なくとも1-塩基の違い）を持つくらいの長さであり得る。しかしながら、好ましくは、それらの長さは類似の付着性配列の間の対誤り形成を安定させることができるようにには長くない。いくつかの実施形態において、約3～約10塩基の長さが使用され得る。しかしながら、付着性突出を生成するために用いられる領域のために、いかなる適當な長さも選択され得る。特異性の重要性は同時に組み立てられる予定である異なる断片の数に依存し得る。また、対誤り領域を安定させるのを避けるのに必要である適當な長さは、付着末端をアニーリングするために使用される条件に依存し得る。

40

【0081】

いくつかの実施形態において、代わりの領域が、それらがアセンブリ設計のための適當な長さをもつ断片を定義する距離によって分離されるなら、選択され得る。いくつかの実

50

施形態において、交互の領域は約 100 から約 500 個の塩基によって分離され得る。しかしながら、いかなる適当な、より短いまたはより長い距離も選択され得る。例えば、その付着領域は約 200 から約 1,000 個の塩基によって切り離され得る。いくつかの要素（例えば、標的核酸の配列、付着末端の選択された長さ、および所望の断片の長さに依存する）によって、代わりの領域の異なるパターンが利用可能であり得ることが理解されるべきである。いくつかの実施形態において、いくつかのオプションが利用可能であるなら、その領域が異なった付着末端の配列差を最大にするように選択され得る。

【0082】

その付着領域の選択は、標的核酸を生成させるように組み立てられる断片を定義する。したがって、断片サイズは、約 100 から約 500 塩基対長の間、約 200 から約 1,000 塩基長の間、または標的核酸により、より短いかまたはより長くなり得る。いかなる適当な技術も使用することにより、断片を生成または入手し得る。いくつかの実施形態において、各々の断片が、付着性一本鎖領域を生成するために使用できる二本鎖領域によって隣接されるように、組み立てられ得る（例えば、多重重複のアセンブリ反応で）。

10

【0083】

いくつかの実施形態において、標的核酸の配列の情報に基づく標的ポリヌクレオチドのアセンブリを可能にするための方法である。いくつかの実施形態において、標的配列を解析するためにコンピュータ・ソフトウェアを使用でき（例えば、 $A_1 \sim A_n$ ）、特定の長さの重複オリゴヌクレオチドの組（ $A_1, A_2, A_3, \dots, A_n$ ）にそれを分解する。オリゴ $A_1, A_2, A_3 \dots A_n$ は、チップまたはマイクロアレイから合成できる。いくつかの実施形態において、オリゴヌクレオチド配列は：增幅プライマー配列、ITS型制限酵素などの制限酵素のための認識部位、パディング（padding）、ペイロード、パディング制限酵素のための認識部位の逆相補（同じであるか異なる）、異なる增幅プライマー配列の逆相補を含むように設計され得る。ペイロードは標的遺伝子（または、いかなる任意の核酸配列）の重複するサブセットであり得る。ペイロードを、所望により、制限酵素での切断の一意的に相補な付着末端の生成を許与する m ヌクレオチド M (M_m) でパディングできる。プライマーは増幅を許す。制限酵素のための認識部位は、プライマーがペイロードから離れて切断されることを許容する。

20

【0084】

ある実施形態において、複数の標的配列にわたる同じ認識部位を使用することは有利である。しかしながら、標的配列が既に認識部位を含むとその認識部位（右から左、または右から左の解析）を含むオリゴが切断され、正しいアセンブリが妨げられることに注意すべきである。いくつかの実施形態において、標的配列が認識部位の単一の発生を含むだけであるならば、部位内の解析を開始し、左からオリゴの組を、さらに認識部位の右側のもう片方の組を解析することによって問題を解決できる。部位が 2 オリゴの間で分けられるので、その結果、完全な配列として存在せず、認識また切断されることはないであろう。所望のオリゴ長または長さの範囲がある場合、解析の各々の側の最後のオリゴは、適切な数の m ヌクレオチド M (M_m) でパディングできる。

30

【0085】

このアプローチは、それらの制限部位がペイロードで許容される長さの範囲の整数倍で現れるならば、認識部位の 1 回を超える発生まで広げることができる。最も簡単な場合（この例の目的のためにいかなる所望の重複も無視する）の例として、何か 2 つの制限部位のいかなる部分が所望の 100 bp ペイロードサイズのためにちょうど 100 bp 離れており、次に、いずれか一方から解析され、他方が自動的に分割されるであろう。ペイロードが 90 ~ 110 bp と変化することができるならば、この距離内の 1 組の制限部位の対に対応できる。また、この同じペイロード範囲で、より長い間の距離で 1 組を分割することもできるであろう： 180 ~ 220 bp, 270 ~ 330 bp など。

40

【0086】

オリゴへ標的配列を解析する場合、最後のオリゴ（または、内部から解析する場合は各々の方向における最後）の長さは、所望の範囲のオリゴ長の外であり得る。最後のオリゴ

50

は所望の長さにパディングできる。これは、しかしながら、特に複数の標的配列を組み立てる場合に、有用でない追加の塩基対を作成するコストを生じ得る。いくつかの実施形態において、この問題の解決法は、単一の長い疑似標的（実際の標的配列の間に任意のプライマーがある）に各標的配列を連結し、次いで、所望の長さの、より小さい、重複する断片へ分割する（例えば、P C Rによる切断または増幅）ことである。断片の長さの計算を以下に示す：

$$\text{長さ} = (\text{片}(\text{p i e c e}) * \text{最大_オリゴ_長}) - (\text{接合} * \text{重複})$$

ここで、接合 = 片 - 1

例えば：

$$\text{長さ } 4\ 8\ 4 = (\text{片 } 5 * \text{最大_オリゴ_長 } 1\ 0\ 0) - (\text{接合 } 4 * \text{重複 } 4)$$

$$\text{長さ } 5\ 0\ 4 = (\text{片 } 5 * \text{最大_オリゴ_長 } 1\ 0\ 4) - (\text{接合 } 4 * \text{重複 } 4)$$

10

【0087】

標的配列のいくつかが制限部位を含むならば、いくつかの場合、接合部位において制限部位（および所望のオリゴ長の範囲内で）を有するように標的配列が連結される順番を選択できる。一般的な場合、標的配列の大部分でパディングを除去する完全な利益を得ながら、制限部位を含む標的配列のサブセットに追加パディングを追加できる。

【0088】

本発明の例は、同じ最後の塩基を有し、異なる最後から2番目の塩基を有する突出をもつ2つのオリゴを正確に区別するある条件におけるあるリガーゼ酵素を示す。いくつかの実施形態において、各々の突出における最後の塩基が一意的になるように、オリゴを設計することが望ましい場合がある。末端（4つの接合）の一意的なA、C、G、Tは、組み立てのために商業的に有用な数である、5片までのライゲーションを許容する。より多くのライゲーション片もまた、以下で示されるように、本発明で考慮され得る：

20

- 最後の2塩基で一意的： $4^2 = 1\ 6$ 接合、17片まで
- 最後の3塩基で一意的： $4^3 = 6\ 4$ 接合、65片まで
- 最後の4塩基で一意的： $4^4 = 2\ 5\ 6$ 接合、257片まで

【0089】

本発明の態様は、入力標的核酸配列を解析するためのアルゴリズムに関する。いくつかの実施形態において、アルゴリズムを、複数のオリゴの最後の塩基（または、最後の2、3または4塩基）が確実に一意的になるようにするために使用できる。例えば、本発明のアルゴリズムは、一意的な標的配列（天然に生じるかまたは非天然で生じる、またはいかなる任意の核酸配列、およそ同じ長さを有するオリゴヌクレオチド、および最後の塩基（または最後の2、3または4塩基）4塩基の重複を有する）を共に含む、複数の解析されたオリゴヌクレオチドを定義するために使用できる。またいくつかの実施形態において、オリゴヌクレオチドを、最後から2番目、または最後から3番目など、またはそれらの組み合わせが一意的であるように定義できる。

30

【0090】

いくつかの実施形態において、第一のアルゴリズムは以下の設計または分解段階を含む：

40

段階1：目標量、例えば、100bpまで移動する、

段階2：組中（例えばメモリ中）に関連する1～4塩基を格納

段階3：重複（4bp）によるバックアップ

段階4：再度の移動、この第二のおよびそれに続く100bpの移動のため、関連する1～4塩基が組中に既に存在するならば、その後に組中にまだない1～4塩基の配列に遭遇するまで1回に1塩基移動させる、

段階5：組に、新たな1～4塩基の配列を加える。

段階6：その後、繰り返す。DNA配列の末端に到達する前に所望の数の片に達しているならば、断片のアセンブリのための適切な重複をバックアップしながら（断片へのオリゴの組立と異なる方法であり得る）、新しい組でやり直す。

【0091】

50

当業者は、1塩基の移動が方向で変化できる、例えば、正常な長さが最大の所望の長さである場合に常に左(より短く)、正常な長さが最小の所望の長さである場合に常に右(より長く)またはそれらのいくつかの組み合わせ、ことに気付くであろう。仮の長さを中心置くために、移動を変える、例えば、チェック位置を以下の順番とする：-1、+1、-2、+2、など、ことができるであろう。移動はまた、好ましい方に重きを置く、例えば、より短いが、より長いものも許容される、例えば、-1、-2、+1、-3、-4、+2など、とすることもできるであろう。

【0092】

このアルゴリズムは、各々のその後の追加に伴って自由度が減少するため、必要な移動が大きくなり得るとして、ある標的配列の設計に制限され得る。例えば、第一の末端は「A」であり得るが、最後の末端はいくつかの塩基の中に「A」を有しない場合があり、その結果、最後のオリゴが非常に短いか非常に長くなり、望ましくない場合がある。この問題の1つの解決法は、各接合点のためのデータのアレイを格納し、次いで、オリゴの最も小さい数のいずれかを選択して移動させるか、またはすべてのオリゴまたはそれのいくつかの組み合わせの中で最も少ない全体の移動距離を選択することである。

10

【0093】

いかなる生じた短い配列(例えば制限部位のための)が、どの程度の頻度で無作為の1,000bp配列において現れるかの統計を以下に示す。例えば、標的配列の中央から解析されない6-bp制限部位が使用されているならば、その制限部位で22%の配列が構築できないであろう。同じ6-bp部位での中間からの解析で、2つの部位を含む3%の配列のみが構築できない(または、追加解析を必要とする)であろう。

20

より具体的に：

- 制限部位による構築の妨害が1回発生する場合：

5bpの長さの1つの量で、62%が少なくとも1部位を有するであろう

6bpの長さの1つの量で、22%が少なくとも1部位を有するであろう

7bpの長さの1つの量で、6%が少なくとも1部位を有するであろう

- 内部からの解析が2回の発生を許容する場合：

5bpの長さの1つの量で、25%が少なくとも2部位を有するであろう

6bpの長さの1つの量で、3%が少なくとも2部位を有するであろう

7bpの長さの1つの量で、<1%が少なくとも2部位を有するであろう(約0.2%)

30

- 1を超える制限酵素(および対応する部位)が使用されていて、1回の発生を許容する場合：

5bpの長さの2つの量で、38%が少なくとも1部位を有するであろう

6bpの長さの2つの量で、5%が少なくとも1部位を有するであろう

7bpおよび6bpの長さで、1%が少なくとも1部位を有するであろう

5bpの長さの3つの量で、24%が少なくとも1部位を有するであろう

6bpの長さの3つの量で、1%が少なくとも1部位を有するであろう

- 2回の発生を許容する1を超える制限酵素の場合：

5bpの長さの2つの量で、6%が少なくとも2部位を有するであろう

6bpの長さの2つの量で、<1%が少なくとも2部位を有するであろう(約0.06%)

40

【0094】

適用

本発明の態様は合成核酸の製造および/または使用に関するさまざまな用途に有用であり得る。本明細書に記載のように、本発明は増加した効率で合成核酸を組み立てるための方法を提供する。得られた組み立てられた核酸は、in vitroで増幅され(例えば、PCR、LCR、またはいかなる適当な増幅法を使用して)、in vivoで増幅され(例えば、適当なベクターでのクローニングを経て)、単離され、および/または、精製され得る。組み立てられた核酸(単独またはベクターにクローニングされる)は宿主細

50

胞（例えば、原核生物、真核生物、昆虫、哺乳類または他の宿主細胞）に導入され得る。いくつかの実施形態において、宿主細胞は、核酸を伝播するために使用され得る。ある実施形態において、核酸は宿主細胞のゲノムに取り込まれ得る。いくつかの実施形態において、核酸は細胞のゲノムの対応する核酸領域に置き換わり得る（例えば、相同組換えを通して）。したがって、核酸は、組換え生物を生産するために使用され得る。いくつかの実施形態において、標的核酸は、宿主生物のゲノムのすべてか一部を置き換えるために使用されるゲノムの全ゲノムまたは大きい断片であり得る。組換え生物もさまざまな研究、工業、農業、および／または、医療の用途に使用され得る。

【0095】

本明細書に記載された技術の多くは、長い核酸配列を生産するために1つ以上のポイントで適當な組み立て技術を適用することにより、共に使用できる。例えば、リガーゼを基にするアセンブリは、100未満から10,000を超える塩基対（例えば、100マーから500マー、500マーから1,000マー、1,000マーから5,000マー、5,000マーから10,000マー、25,000マー、50,000マー、75,000マー、100,000マーなど）の長さの二重オリゴヌクレオチドおよび核酸断片を組み立てるために使用され得る。例示的な実施形態において、本明細書に記載の方法は、生物（例えば、ウイルス性、細菌、酵母、または他の原核生物、または真核生物）の全ゲノム（または、その大きい断片、例えば、約10%、20%、30%、40%、50%、60%、70%、80%、90%以上）の、任意で1以上の所望の位置に特異的修飾を組み込む、組み立てに使用され得る。

10

20

【0096】

いかなる核酸産物（例えば、増幅、クローニング、精製、単離等された核酸を含む）は、保管および／または流通（例えば、配送センターまたは顧客に流通するため）のためのいかなる適當な形式（例えば、安定バッファー中、凍結乾燥など）でパッケージされ得る。同様に、いかなる宿主細胞（例えば、ベクターで形質転換されたか、または修飾ゲノムを有する細胞）が保存およびまたは輸送（例えば、顧客に配送するため）のために適當なバッファーで調製され得る。いくつかの実施形態において、細胞は凍結され得る。しかしながら、他の安定した細胞調製もまた使用され得る。

30

【0097】

宿主細胞は、培地中で増殖および拡大され得る。宿主細胞は、対象となる1つ以上のRNAまたはポリペプチド（例えば、医療、工業、農業、および／または、医薬のタンパク質）を発現するために使用され得る。発現されたポリペプチドは、天然ポリペプチドまたは非天然ポリペプチドであり得る。ポリペプチドは、その後の使用のために単離または精製され得る。

40

【0098】

したがって、本発明の方法を使用して生成された核酸分子はベクターに組み込むことができる。ベクターは、クローニングベクターまたは発現ベクターであり得る。いくつかの実施形態において、ベクターはウイルスベクターであり得る。ウイルスベクターは標的細胞を感染させることができる核酸配列を含み得る。同様に、いくつかの実施形態において、標的細胞を改変するのに適切なプロモーター系に作動可能に結合された原核生物の発現ベクターが使用できる。他の実施形態において、適切なプロモーター系に作動可能に結合された真核ベクターはトランスフェクト標的細胞または組織に使用できる。

【0099】

本明細書に記載された構築物の転写および／または翻訳が *in vitro*（すなわち、無細胞系を使用する）または *in vivo*（すなわち、細胞内で発現される）で行われ得る。いくつかの実施形態において、細胞溶解物が調製され得る。ある実施形態において、発現されたRNAまたはポリペプチドが、単離または精製され得る。本発明の核酸も、発現されたポリペプチドまたはその断片に検出および／または生成タグを追加するために使用され得る。ポリペプチドを基にした融合／タグはヘキサヒスチジン（His⁶）MycおよびHA、および他のポリペプチドと、GFP5、GST、MBP、キチンな

50

どの多用途品 (utility) を含むが、これらに限られるというものではない。いくつかの実施形態において、ポリペプチドは1つ以上の非天然アミノ酸残基を含み得る。

【0100】

いくつかの実施形態において、1個以上の合成核酸によってコードされたポリペプチドまたはその断片に対して抗体を作成することができる。ある実施形態において、研究開発（例えば潜在的治療用タンパク質またはペプチドを同定して、創薬などのための潜在的タンパク質標的を同定する）における選別のためのライプラリとして合成核酸を提供し得る。いくつかの実施形態において、合成核酸は医療用として（例えば、遺伝子療法または遺伝子調節のため）使用され得る。例えば、合成核酸は、治療上の量のタンパク質を発現できる程度の量で患者に投与され得る。他の実施形態において、合成核酸は遺伝子の発現を調節する（例えば、下方制御する）ために十分な量で患者に投与され得る。

10

【0101】

異なる行為または本明細書に記載された実施形態が、別個に実施され得、また合衆国内、または合衆国外と異なる場所で実施され得ることが理解されるべきである。例えば、標的核酸の注文を受ける、標的核酸を分析する、1つ以上の出発核酸（例えば、オリゴヌクレオチド）を設計する、出発核酸を合成する、出発核酸を精製する、出発核酸を組み立てる、組み立てられた核酸を単離する、組み立てられた核酸を配列を確認する、組み立てられた核酸を操作する（例えば、增幅、クローニング、宿主ゲノムへの挿入など）などの行為、およびいかなる他の行為、さらにこれらの行為のいかなる部分も、合衆国内または合衆国外の1つの位置でまたは異なる場所で独自に実行され得る。いくつかの実施形態において、アセンブリ手順は1つの場所（合衆国内または合衆国外）で実施される行為および1つ以上の遠隔地（合衆国内または合衆国外）で実施される行為の組み合わせを伴い得る。

20

【0102】

自動化適用

本明細書で提供された方法および機器の態様は、本明細書に説明された1つ以上の行為を自動化することを含み得る。いくつかの実施形態において、増幅および／またはアセンブリ反応の1以上の段階は、1台以上の自動試料処理機器（例えば、1台以上の自動化された液体または流体操作機器）を使用することで自動化され得る。自動化機器および手順は、以下の1つ以上を含む反応試薬を提供するために用いられ得る：出発核酸、バッファー、酵素（例えば、1つ以上のリガーゼおよび／またはポリメラーゼ）、ヌクレオチド、塩、および安定剤などのいかなる他の適当な試薬。自動化機器および手順はまた、反応条件を制御するために用いられ得る。例えば、自動化された熱循環装置は、使用され得る反応温度およびいかなる温度サイクルを制御するために使用され得る。いくつかの実施形態において、スキャンレーザーは、ポリヌクレオチドをインキュベートするために適当な1つ以上の反応温度または温度サイクルを提供するために自動化され得る。同様に、組み立てられたポリヌクレオチド製品のその後の分析が自動化され得る。例えば、配列決定は、配列決定機器および自動配列決定プロトコルを使用することで自動化され得る。付加段階（例えば、増幅、クローニングなど）もまた、1つ以上の適切な機器および関連するプロトコルを使用することで自動化され得る。本明細書に記載される1つ以上の機器または機器の構成要素が、システム（例えば、ロボットシステム）または、マイクロ環境（例えば、マイクロ流体反応チャンバー）で組み合わせられ得ることが理解されるべきである。アセンブリ反応混合物（例えば、液体反応試料）は、自動化機器および手順（例えば、自動化されたピペット操作機器、マイクロシステムなどを含む、ロボット操作、および／または試料輸送および／または試料容器）を用いることで、システムの1つの構成要素から他の構成要素に輸送され得るシステムとそのいかなる構成要素も制御システムによって制御され得る。

30

【0103】

したがって、本明細書で提供された機器の方法の段階および／または態様は、例えば、コンピュータシステム（例えば、コンピュータ制御システム）を用いて自動化され得る。

40

50

本明細書で提供された技術の態様を実装することができるコンピュータシステムはいかなる型の処理（例えば、本明細書に記載される配列分析および／または自動化された機器制御）のためのコンピュータを含み得る。しかしながら、ある処理段階がアセンブリシステムの一部である1つ以上の自動化機器によって提供され得ることが理解されるべきである。いくつかの実施形態において、コンピュータシステムは2台以上のコンピュータを含み得る。例えば、1台のコンピュータがネットワークを通じて第二のコンピュータと連動され得る。1台のコンピュータが配列分析を実行し得る。第二のコンピュータはシステムの自動合成およびアセンブリ機器の1つ以上を制御し得る。他の態様において、追加のコンピュータは、分析または処理実行の1つ以上を制御するためにネットワーク内に含まれ得る。各コンピュータはメモリおよびプロセッサを含み得る。コンピュータはいかなる形態をとることができ、本明細書で提供される技術の態様として、いかなる特定のコンピュータプラットホームで制限なしに実装される。同様に、ネットワークはプライベートネットワークまたはパブリックネットワーク（例えば、インターネット）を含むいかなる形態をとることができる。ディスプレイ装置を、機器とコンピュータの1つ以上に関連させることができます。代わりに、または追加で、ディスプレイ装置は、本明細書に提供された技術に従って、リモートサイトに位置し、分析の出力を表示するために接続され得る。システムの異なる構成要素の間の接続は、ワイヤー、光ファイバー、ワイヤレス通信、衛星通信、他のいかなる適当な転送、またはそれらのいかなる組み合わせであり得る。

【0104】

本明細書で提供される各々の異なる態様、実施形態または技術の実施は、独立して自動化でき、いかなる複数の方法で実装できる。例えば、各々の態様、実施形態、または実施が、ハードウェア、ソフトウェアまたはその組み合わせを使用して、個別に実装され得る。ソフトウェアで実装されるとき、ソフトウェアコードは、1台のコンピュータまたは複数のコンピュータに分配されるかにかかわらず、いかなる適切なプロセッサまたはプロセッサの集合で実行できる。上述した機能を実行するいかなる構成要素または構成要素の集合が、上述された機能を制御する1台以上のコントローラであると一般的にみなすことができることが理解されるべきである。1台以上のコントローラを専用ハードウェアで、または上述した機能を実行するためにマイクロコードかソフトウェアを使用することでプログラムされる汎用のハードウェア（例えば、1台以上のプロセッサ）で、などの複数の方法で実装できる。

【0105】

これについて、本明細書で提供される技術の実施形態の1つの実装は、プロセッサ上で実行される時に、本明細書で提供される技術の1つ以上の上述の機能を実施するコンピュータプログラム（すなわち、複数の指示）をコードされた、少なくとも1つ以上のコンピュータ読み取り可能媒体（例えばコンピュータメモリ、フロッピー（登録商標）ディスク、コンパクトディスク（登録商標）、テープなど）を含むことが理解されるべきである。コンピュータ読み取り可能な媒体は、それらに格納されたプログラムが、本明細書で提供された技術の1つ以上の機能を実装するいかなるコンピュータシステム資源上でロードできるように、輸送可能とすることができます。さらに、実行されると、上述の機能を実施するコンピュータプログラムの参照は、ホストコンピュータ上で実行されるアプリケーションプログラムに制限されないことが理解されるべきである。むしろ、本明細書における用語、コンピュータプログラムは、本明細書で提供された技術の上述の態様を実装するプロセッサをプログラムするためのコンピュータ・コード（例えば、ソフトウェアまたはマイクロコード）のいかなる型を参照する、一般的な意味で本明細書で用いられる。

【0106】

プロセスがコンピュータ読み取り可能な媒体に保存される、本明細書で提供される技術のいくつかの実施形態に従って、プロセスが実装されたコンピュータは、それらの実行の進行する間、手動で入力を受け取り得る（例えば、ユーザーから）ことが理解されるべきである。

【0107】

10

20

30

40

50

したがって、本明細書に記載されたアセンブリ機器または構成要素の総合的なシステムレベルでの制御は、関連する核酸シンセサイザ、液体処理機器、熱循環器、配列機器、関連するロボット構成要素、ならびに希望の入力／出力または他の制御機能を実行する他の適当なシステムの制御信号を提供し得るシステムコントローラによって実行され得る。したがって、いかなる機器コントローラを伴うシステムコントローラが、共に核酸アセンブリシステムの操作を制御するコントローラを形成する。コントローラは汎用目的コンピュータとなる汎用目的のデータ処理システム、または汎用目的コンピュータおよび、モデムなどのコミュニケーション機器および／または所望の入力／出力または他の機能を実施する他の回路または構成要素を含む、他の関連デバイスのネットワークを含み得る。コントローラもまた少なくとも部分的に、単一の専用集積回路（例えば、A S I C）または各々が、全体の、システムレベルでの制御のための主または中央プロセッサ部分、および中央プロセッサ部分の制御下で様々な異なった特定の計算、機能および他のプロセスを実行するため設けられた分離された部分を有する、A S I Cのアレイとして実装できる。また、コントローラが離散的な要素回路かプログラム可能論理回路などの電子で配線されて、例えば統合するか他の多くの別々の専用プログラム可能な電子回路かデバイスを使用するか、論理回路で実装できる。コントローラはまた、複数の、別個に専用プログラム可能に統合されたかまたは他の電子回路または機器、例えば、配線された電子回路または別個の基本回路またはプログラム可能な論理デバイスのような論理回路を用いて実装できる。また、コントローラは、ユーザー入力／出力装置（モニター、ディスプレイ、プリンタ、キーボード、ユーザーインティングデバイス、タッチスクリーン、または他のユーザーインターフェースなど）、データ保存機器、駆動モーター、リンク機構、弁制御器、ロボット機器、真空および他のポンプ、圧力センサ、検出器、電源、パルス源、通信装置または他の電子回路または構成要素などのいかなる他の構成要素または機器も含めることができる。コントローラはまた、当分野で知られているが、本明細書に詳細に記載されなかつた他の適当な機能を実行するために自動化された、顧客注文処理、品質管理、パッケージング、出荷、支払いなどのシステムの他の部分の操作を制御し得る。

【0108】

本発明の様々な態様は、単独、組み合わせ、または上記された実施形態で特に議論しなかつたさまざまなアレンジメントで使用され得、したがって、その用途は上記で詳しく説明されたか、または図面で示された構成要素の詳細およびアレンジメントに制限されない。例えば、一の実施形態で説明された態様は、いかなる方法で他の実施形態で説明される態様と組み合わせられ得る。

【0109】

請求項において、請求項の要素を修飾する、「第一」、「第二」、「第三」などの序数の用語の使用は、それ自体によって、いかなる優先権、先行、他からの一の請求項の要素の順序または方法の行為が実施される時間的な順序を暗示するものではなく、（序数の用語としての使用でなく）単に同じ名称を有する他の要素から一つの請求項の要素を区別するための標識として用いられる。

【0110】

また、本明細書で使用された言い回しおよび用語は、記述の目的のためのものであり、制限を意図すると見なすべきではない。「含む」、「包含する」または「有する」、「包括する」、「持つ」およびそれらの変形の使用は、追加の項目と同様に、その後記載された項目とその同等物を含むことを意味する。

【実施例】

【0111】

図1は、任意に選択された約836bp長の二本鎖配列を示す。60-bp断片が選択され、1から28まで標識された（断片1～14は陽性鎖、断片15～28は陰性鎖）。これらの60-bp断片は下記のフランギング配列とともに、I D T (Integrated DNA Technologies, Coralville, Iowa) (「I D T」オリゴ)に注文した：

10

20

30

40

50

【化1】

GTCACTACCGCTATCATGGCGGTCTC.....GAGACCAGGAGACAGGACCGACCAAA
CAGTGATGGCGATAGTACCGCCAGAG.....CTCTGGTCCTCTGTCCTGGCTGGTTT

下線は4-塩基突出を生成するBsa I-HF認識部位である。

【化2】

5'...GGTCTC(N)....3'
3'...CCAGAG(N)....5'

Bsa I-HF認識部位は、これらの断片を増幅するために有用なユニバーサルプライマーによって隣接される。 10

【0112】

PCRプライマーA~Eは、正しいライゲーション産物を増幅するために設計された(図1の破線矢印)。図2は対応するPCR産物の予想されるサイズならびに矢じりで示されるプライマー(「オリゴA」から「オリゴE」)の相対的な位置を示す。

【0113】

二本鎖IDTオリゴは、下記の条件でBsa I-HF消化に供された:

- 1 X NEバッファー4
- 100 μg/mlのウシ血清アルブミンで補足
- 37度でインキュベート。

【0114】

付着末端(オリゴ1~28)を有する消化された二本鎖オリゴは、4%ゲル上の電気泳動で精製された。いくつかの異なるリガーゼ、温度およびインキュベーション時間が、最適なライゲーション条件のために試験された。試験したリガーゼは以下を含む: 20

T4 DNAリガーゼ

T4 DNAリガーゼ + 300 mM塩(活性の減少、より高い特異性のため)

T3 DNAリガーゼ

T7 DNAリガーゼ

Pfu DNAリガーゼ

Taq DNAリガーゼ

E.coli DNAリガーゼ

【0115】

室温で30分間行われた典型的な結果を、図3~5に示す。図3は(2オリゴの)対ライゲーションの電気泳動結果を示す。ゲルの左から右に:ラダー、リガーゼなし、T4DNAリガーゼ、T4 DNAリガーゼ+塩、T3 DNAリガーゼ、T7 DNAリガーゼ。ゲルの下部から上部までのバンドは以下に対応する:遊離オリゴ、正しいライゲーション産物、1.5ライゲーション産物、ライゲーション産物二量体。他は同程度であったが、T7 DNAリガーゼは、最も正しいライゲーション産物を生産し、その結果、この実験条件下で最も効率的であるように見えた。 30

【0116】

図4はオリゴ1~10(レーン1~6)およびオリゴ11~14(レーン7~10)の、ゲルの上端に示された異なるリガーゼでのライゲーション結果を示す。異なるライゲーション産物の存在を示す多重バンドが観察された。しかしながら、プライマーとしてオリゴAおよびBを使用するPCR増幅の際に、約300 bpの強いバンドが観測された。オリゴAおよびBから予測されたPCR産物は337 bp(図2参照)であるので、このバンドはオリゴ1~6を含む正しいライゲーション産物に対応する(図1参照)。バンドは、ゲルから切り出され、精製され、配列決定された。配列決定結果は図6に示され、予想された配列と比べて、ライゲーション産物の100%の忠実度を確認した。高い温度(45~65)でのみ活性であるTaq DNAリガーゼは、低反応温度(室温)のためか、いかなるライゲーション産物も生産しなかった。 40

【0117】

10

20

30

40

50

対ミスマッチアッセイは、様々な合成酵素の特異性を試験するために開発された。1組のオリゴは4塩基の突出で設計され、完全な適合（「P」）配列はGGTGであり、ミスマッチ（「M」）配列は1個のヌクレオチドが正しい配列と異なるGCTGである。図7Aおよび7Bに示されるように、ライゲーションされていないオリゴに対応する下側のバンド（リガーゼ無しの対照で示される）、およびライゲーション産物に対応する上側のバンドの、2つの主要なバンドを観測できる。T4 DNAリガーゼ+塩、T3 DNAリガーゼ、T7 DNAリガーゼおよびE.coli DNAリガーゼは全て、完全な適合の突出を用いたライゲーション産物に対応する強いバンドを生成した。対照的に、ミスマッチ突出が使用されたとき、産物の大部分がライゲーションされていないオリゴであった。これらの実験は、これらの反応条件において、T4 DNAリガーゼ+塩、T3 DNAリガーゼ、T7 DNAリガーゼおよびE.coli DNAリガーゼの全てが、高い特異性を有し、ミスマッチの1個のヌクレオチドの違いと区別されることを示した。
10

【0118】

上述したオリゴ1～6を有するライゲーション産物に加えて、より長い産物を含む他のライゲーション産物も生産された。1つの産物はオリゴ14にライゲーションされたオリゴ1～6とみられた。これはオリゴ7およびオリゴ14が同じ付着末端（GTTTC、図8の箱）を有するという事実による。

【0119】

本発明は高忠実度の遺伝子アセンブリのための特に新規な方法および機器を提供する。本発明の要素の特定の実施形態について議論したが、上記の詳細は説明であり制限ではない。本発明の多くの変化がこの明細書を参照した当業者に明らかになるであろう。本発明の完全な範囲はそれらの同等物の完全な範囲、および明細書、そのような変化に沿って、特許請求の範囲を参照して決定されるべきである。
20

【0120】

本明細書で引用したすべての刊行物、特許および配列データベースは、それらがあたかも個別に示されるかのように、参照によりその全体が本明細書に組み込まれる。

【図1】

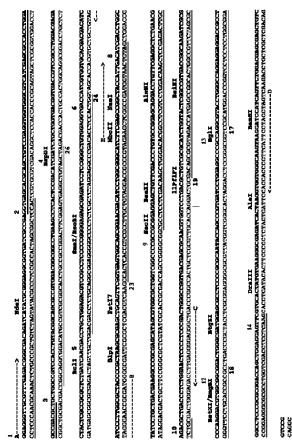


Figure 1

【 図 2 】

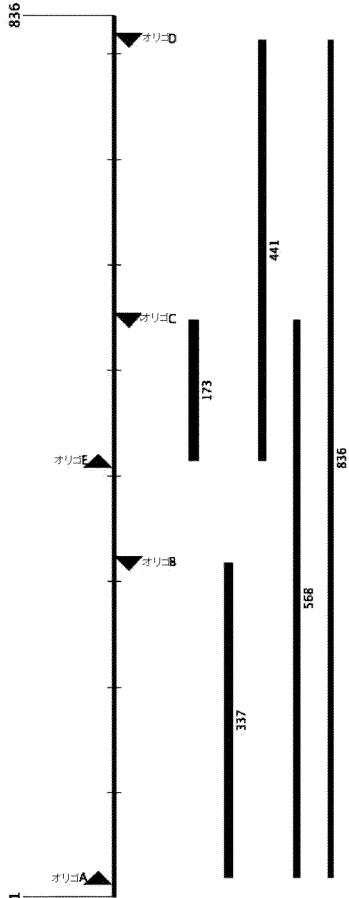
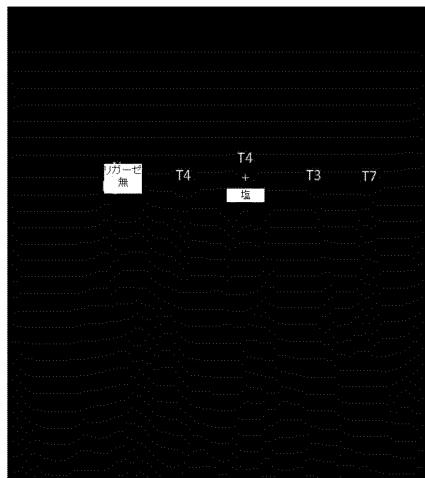


Figure 2

【図3】

ライゲーション



二量体 一. 五 ライゲーション オリゴ

【 図 4 】

オリゴ 1-10

オリゴ 11-14

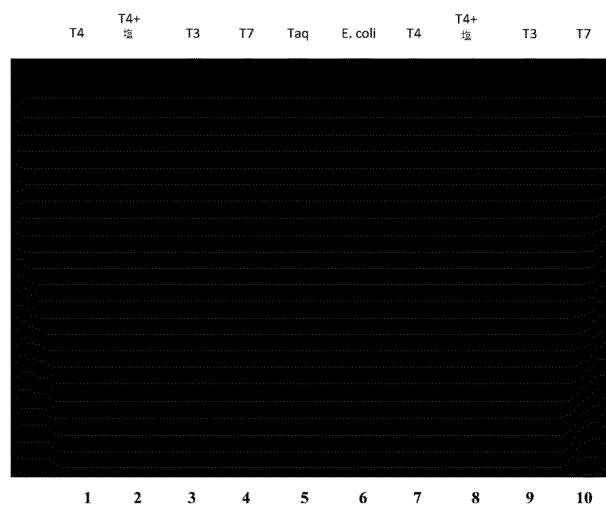


Figure 4

Figure 3

【図5】

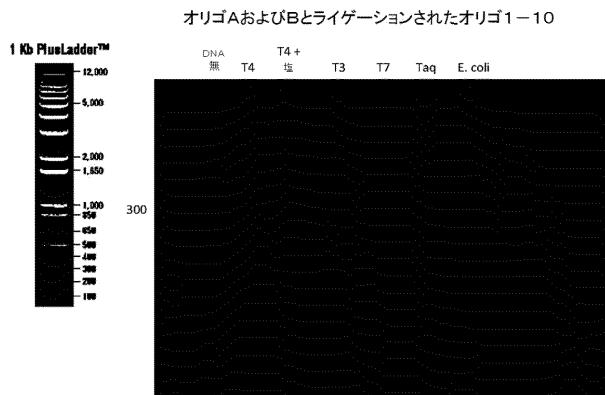


Figure 5

【図6】

予想と比較したライゲーション産物の配列

```
>lcl142826 2
長さ = 550
スコア = 623 bits (33%), 予想 = 0.0
同一性 = 337/337 (100%), キャップ = 0/337 (0%)
鎖 = プラス/プラス
クエリー 1 GGAGGGTTGCGTTTGAGACGGCGGNCAAGATCATATGCCGGAGGCCGGTGTATGCCGGAGGTG 60
対象 102 GGAGGGTTGCGTTTGAGACGGCGNCAAGATCATATGCCGGAGGCCGGTGTATGCCGGAGGTG 161
クエリー 61 AGCACGAGCTGTCAGCTGGTGGCGGTGATCGAGGCCACCTTGAGGCCACCTGGAGCCGACCTGGCTG 120
対象 162 AGCACGAGCTGTCAGCTGGTGGCGGTGATCGAGGCCACCTTGAGGCCACCTGGAGCCGACCTGGCTG 221
クエリー 121 GCGCTCCACCTGTACGGCAGCGCCGTGAGCGCGGCGCTGAAGCCTCACTCCGACATCGAT 180
対象 222 GCGCTCCACCTGTACGGCAGCGCCGTGAGCGCGGCGCTGAAGCCTCACTCCGACATCGAT 281
クエリー 181 CTGCTGGTACGGTGAACGGTCCCTCCCCGGAGAGCGAGACTACTCGCCGGCGCTCTGATCAACGAC 240
対象 282 CTGCTGGTACGGTGAACGGTCCCTCCCCGGAGAGCGAGACTACTCGCCGGCGCTCTGATCAACGAC 341
クエリー 241 CTGCTGGAGACGTCGCTCCCTCCCCGGAGAGCGAGACTCTGGAGGGTGAAC 300
対象 342 CTGCTGGAGACGTCGCTCCCTCCCCGGAGAGCGAGACTCTGGAGGGTGAAC 401
クエリー 301 ATCGTGGTGCACGACATCATCCCTGGCGCTACC 337
対象 402 ATCGTGGTGCACGACATCATCCCTGGCGCTACC 438
```

Figure 6

【図7】

対ミスマッチライゲーションアッセイ

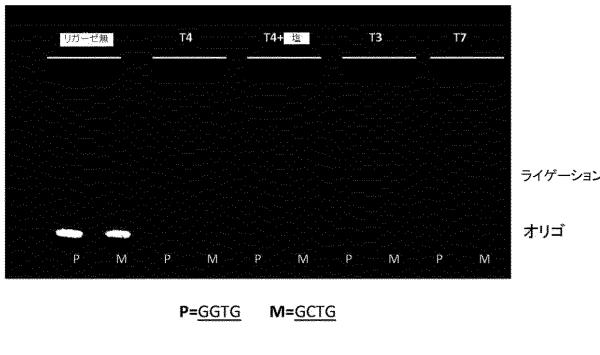


Figure 7A

【図8】

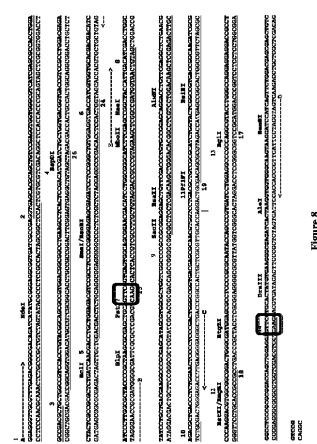
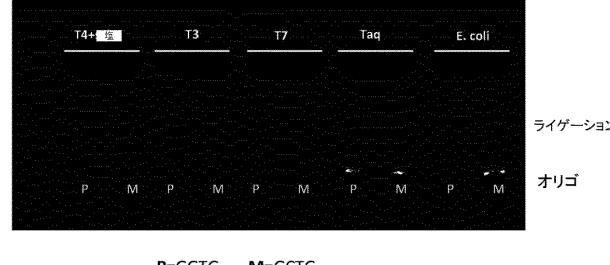


Figure 8



P=GGTG M=GCTG

Figure 7B

【図9】

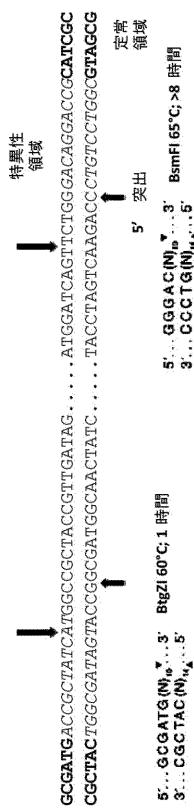


Figure 9A

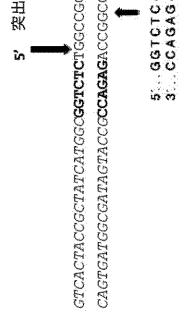


Figure 9B

【図10】

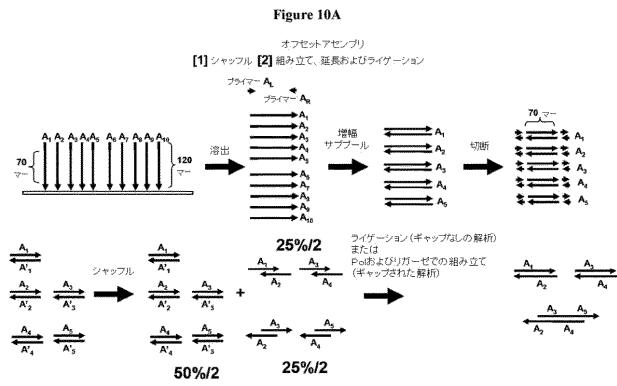
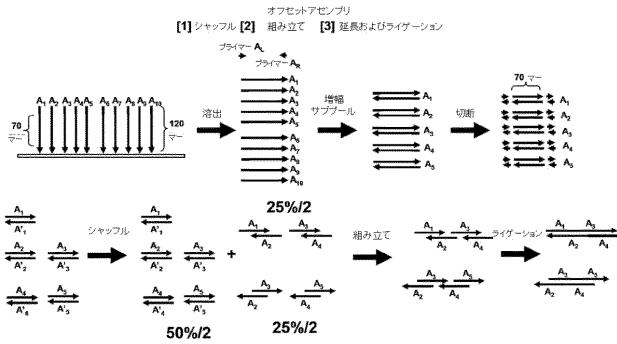


Figure 10B

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No PCT/US2012/052036									
A. CLASSIFICATION OF SUBJECT MATTER INV. C12N15/10 C12N15/66 C12P19/34 C40B50/06 G06F19/10 ADD.											
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) C12N C12P G06F											
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched											
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, BIOSIS, CAB Data, Sequence Search, WPI Data											
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left; padding: 2px;">Category*</th> <th style="text-align: left; padding: 2px;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="text-align: left; padding: 2px;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td style="padding: 2px;">X</td> <td style="padding: 2px;">WO 01/81568 A1 (GENENCOR INT [US]) 1 November 2001 (2001-11-01) claims 1-17 -----</td> <td style="padding: 2px;">1-3,5, 9-30</td> </tr> <tr> <td style="padding: 2px;">Y</td> <td style="padding: 2px;">RICHMOND K E ET AL: "Amplification and assembly of chip-eluted DNA (AACED): a method for high-throughput gene synthesis", NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 32, no. 17, 1 January 2004 (2004-01-01), pages 5011-5018, XP002344586, ISSN: 0305-1048, DOI: 10.1093/NAR/GKH793 cited in the application the whole document ----- -/-</td> <td style="padding: 2px;">2,3,5</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	X	WO 01/81568 A1 (GENENCOR INT [US]) 1 November 2001 (2001-11-01) claims 1-17 -----	1-3,5, 9-30	Y	RICHMOND K E ET AL: "Amplification and assembly of chip-eluted DNA (AACED): a method for high-throughput gene synthesis", NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 32, no. 17, 1 January 2004 (2004-01-01), pages 5011-5018, XP002344586, ISSN: 0305-1048, DOI: 10.1093/NAR/GKH793 cited in the application the whole document ----- -/-	2,3,5
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.									
X	WO 01/81568 A1 (GENENCOR INT [US]) 1 November 2001 (2001-11-01) claims 1-17 -----	1-3,5, 9-30									
Y	RICHMOND K E ET AL: "Amplification and assembly of chip-eluted DNA (AACED): a method for high-throughput gene synthesis", NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 32, no. 17, 1 January 2004 (2004-01-01), pages 5011-5018, XP002344586, ISSN: 0305-1048, DOI: 10.1093/NAR/GKH793 cited in the application the whole document ----- -/-	2,3,5									
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		<input checked="" type="checkbox"/> 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											
Date of the actual completion of the international search 10 December 2012		Date of mailing of the international search report 22/02/2013									
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Hornig, Horst									

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2012/052036

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.: 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:

3. Claims Nos.: 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:

see additional sheet

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 an 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.:

see additional sheet(s)

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

International application No PCT/US2012/052036

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2011/085075 A2 (GEN9 INC [US]; JACOBSON JOSEPH [US]; CHU LARRY LI-YANG [US]) 14 July 2011 (2011-07-14) the whole document -----	2,3,5
Y	WO 2011/066186 A1 (GEN9 INC [US]; RAMU SENTHIL [US]; JACOBSON JOSEPH [US]; CHU LARRY LI-Y) 3 June 2011 (2011-06-03) the whole document -----	2,3,5
A	WO 00/49142 A1 (FEBIT FERRARIUS BIOTECHNOLOGY [DE]; STAehler PEER F [DE]; STAehler COR) 24 August 2000 (2000-08-24) the whole document -----	1-30
A	WO 03/010311 A2 (DSM NV [NL]; BOVENBERG ROELOF ARY LANS [NL]; KERKMAN RICHARD [NL] DSM) 6 February 2003 (2003-02-06) the whole document -----	1-30
A	WO 00/40715 A2 (UNIV BOSTON [US]; JARRELL KEVIN A [US]; COLJEE VINCENT W [US]; DONAHUE) 13 July 2000 (2000-07-13) the whole document -----	1-30
A	WO 98/05765 A1 (NOVO NORDISK AS [DK]; MIYOTA YOSHIAKI [JP]; FUKUYAMA SHIRO [JP]) 12 February 1998 (1998-02-12) the whole document -----	1-30
A	WO 2007/113688 A2 (LIBRAGEN [FR]; LEFEVRE FABRICE [FR]; NALIN RENAUD [FR]; JARRIN CYRILLE) 11 October 2007 (2007-10-11) the whole document -----	1-30

INTERNATIONAL SEARCH REPORT Information on patent family members				International application No PCT/US2012/052036	
Patent document cited in search report	Publication date	Patent family member(s)		Publication date	
WO 0181568	A1 01-11-2001	AU CA EP WO	5300101 A 2406466 A1 1276858 A1 0181568 A1	07-11-2001 01-11-2001 22-01-2003 01-11-2001	
WO 2011085075	A2 14-07-2011		NONE		
WO 2011066186	A1 03-06-2011	EP US WO	2504449 A1 2012028843 A1 2011066186 A1	03-10-2012 02-02-2012 03-06-2011	
WO 0049142	A1 24-08-2000	AT AT AU CA EP EP EP JP US US US US WO	334197 T 456652 T 767606 B2 2362939 A1 1153127 A1 1728860 A1 2175021 A2 2002536977 A 6586211 B1 2003198948 A1 2007196854 A1 2009170802 A1 0049142 A1	15-08-2006 15-02-2010 20-11-2003 24-08-2000 14-11-2001 06-12-2006 14-04-2010 05-11-2002 01-07-2003 23-10-2003 23-08-2007 02-07-2009 24-08-2000	
WO 03010311	A2 06-02-2003	AT DE US WO	321850 T 60210298 T2 2005130140 A1 03010311 A2	15-04-2006 02-11-2006 16-06-2005 06-02-2003	
WO 0040715	A2 13-07-2000	AT AU AU CA EP JP JP JP JP US US US US WO	439437 T 774306 B2 3692100 A 2360011 A1 1141275 A2 2002537762 A 2007075128 A 2011200257 A 2005191623 A1 2006105358 A1 2007269858 A1 2009317875 A1 0040715 A2	15-08-2009 24-06-2004 24-07-2000 13-07-2000 10-10-2001 12-11-2002 29-03-2007 13-10-2011 01-09-2005 18-05-2006 22-11-2007 24-12-2009 13-07-2000	
WO 9805765	A1 12-02-1998	AU JP WO	3691797 A 10066576 A 9805765 A1	25-02-1998 10-03-1998 12-02-1998	
WO 2007113688	A2 11-10-2007	CA EP EP US WO	2647498 A1 1842915 A1 2002002 A2 2010015667 A1 2007113688 A2	11-10-2007 10-10-2007 17-12-2008 21-01-2010 11-10-2007	

International Application No. PCT/ US2012/ 052036

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-30

A method of producing a target nucleic acid having a predefined sequence, the method comprising: providing a plurality of blunt-end double-stranded nucleic acid fragments having a restriction enzyme recognition sequence at both ends of each of the plurality of blunt-end double-stranded nucleic acid fragments; producing a plurality of cohesive-end double-stranded nucleic acid fragments that together comprises the target nucleic acid sequence via enzymatic digestion of the plurality of blunt-end double-stranded nucleic acid fragments, wherein the plurality of cohesive-end double-stranded nucleic acid fragments each have two different and non-complementary overhangs; ligating the plurality of cohesive-end double-stranded nucleic acid fragments with a ligase, wherein a first overhang of a first cohesive-end double-stranded nucleic acid fragment is uniquely complementary to a second overhang of a second cohesive-end double-stranded nucleic acid fragment; and forming a linear arrangement of the plurality of cohesive-end double-stranded nucleic acid fragments, wherein the unique arrangement comprises the target nucleic acid having a predefined sequence;

2. claims: 31-42, 44

A method for designing a plurality of starting nucleic acids to be assembled into a target nucleic acid, the method comprising, (a) obtaining an input target sequence of a target nucleic acid; (b) selecting a plurality of subsequences therein such that every two adjacent subsequences overlap with each other by N bases; (c) storing the resulting overlapping N-bases in a memory; (d) comparing the overlapping N-base sequences to one another to ensure that they differ from one another by at least one base; and (e) repeating steps (b) to (d) until a plurality of satisfactory nucleic acid fragments are obtained wherein any two adjacent starting nucleic acid fragments uniquely overlap with each other by N bases; a plurality of starting nucleic acids to be assembled into a target nucleic acid, designed by said method; a computer program product for designing a plurality of starting nucleic acids to be assembled into a target nucleic acid, said program residing on a hardware computer readable storage medium and having a plurality of instructions which, when executed by a processor, cause the processor to perform operations comprising: (a) obtaining a target sequence of a target nucleic acid; (b) selecting a plurality of subsequences therein such that every two adjacent subsequences overlap with each other by N bases; (c) storing the resulting

International Application No. PCT/US2012/052036

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

overlapping N-base sequences in a memory; (d) comparing the overlapping N-base sequences to one another to ensure that they differ from one another by at least one base; and (e) repeating steps (b) to (d) until a plurality of satisfactory starting nucleic acids are obtained wherein any two adjacent starting nucleic acids uniquely overlap with each other by N bases;

3. claim: 43

A system for assembling a target nucleic acid, the system comprising: a solid support for synthesizing the plurality of starting nucleic acids, wherein each starting nucleic acid further comprises an engineered universal primer binding site and an engineered restriction enzyme recognition sequence; a polymerase reaction unit for synthesizing complementary strands of the plurality of starting nucleic acids a polymerase-based reaction using a universal primer complementary to the universal primer binding site, thereby producing a plurality of blunt-end double-stranded nucleic acid fragments; a digestion unit for producing a plurality of cohesive-end double-stranded nucleic acid fragments via enzymatic digestion of the plurality of blunt-end double-stranded nucleic acid fragments , wherein the plurality of cohesive-end double-stranded nucleic acid fragments each have two different and non-complementary overhangs; and a ligation unit for ligating the plurality of cohesive-end double-stranded nucleic acid fragments with a ligase, wherein a first overhang of a first cohesive-end double-stranded nucleic acid fragment is uniquely complementary to a second overhang of a second cohesive-end double-stranded nucleic acid fragment;

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(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,R,S,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),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,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,H,U,ID,IL,IN,IS,JP,KE,KG,KM,KN,KP,KR,KZ,LA,LK,LR,LS,LT,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PE,PG,PH,PL,PT,QA,RO,RS,RW,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN

(72)発明者 ジョセフ・ジェイコブソン

アメリカ合衆国02459マサチューセッツ州ニュートン、グラント・アベニュー223番

(72)発明者 ダニエル・シンドラー

アメリカ合衆国02464マサチューセッツ州ニュートン、ウィリアムズ・ストリート31番

(72)発明者 スコット・エス・ロートン

アメリカ合衆国01824マサチューセッツ州チャーチルムズフォード、コロニアル・ドライブ24番

Fターム(参考) 4B024 AA19 AA20 CA05 HA09

4B029 AA23 BB20 CC08