

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6509727号
(P6509727)

(45) 発行日 令和1年5月15日 (2019.5.15)

(24) 登録日 平成31年4月12日 (2019.4.12)

(51) Int. Cl.

F I

C 1 2 N 15/10 (2006.01)
 C 1 2 Q 1/6869 (2018.01)
 C 1 2 N 15/63 (2006.01)
 C 1 2 N 9/16 (2006.01)

C 1 2 N 15/10 Z N A Z
 C 1 2 Q 1/6869 Z
 C 1 2 N 15/63 Z
 C 1 2 N 9/16 A

請求項の数 17 (全 33 頁)

(21) 出願番号 特願2015-520370 (P2015-520370)
 (86) (22) 出願日 平成25年6月24日 (2013.6.24)
 (65) 公表番号 特表2015-523866 (P2015-523866A)
 (43) 公表日 平成27年8月20日 (2015.8.20)
 (86) 国際出願番号 PCT/US2013/047370
 (87) 国際公開番号 W02014/004393
 (87) 国際公開日 平成26年1月3日 (2014.1.3)
 審査請求日 平成28年6月23日 (2016.6.23)
 (31) 優先権主張番号 61/731, 627
 (32) 優先日 平成24年11月30日 (2012.11.30)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 61/664, 118
 (32) 優先日 平成24年6月25日 (2012.6.25)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 516286741
 ギンゴー バイオワークス, インコーポ
 レイテッド
 アメリカ合衆国 マサチューセッツ O2
 210, ボストン, ドライドック ア
 ベニュー 27, フロア 8
 (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 憲史
 (74) 代理人 100157956
 弁理士 稲井 史生

最終頁に続く

(54) 【発明の名称】 核酸アセンブリおよび高処理シーケンシングのための方法

(57) 【特許請求の範囲】

【請求項 1】

予め定義された配列を有する核酸を製造する方法であって、

(a) 以下を含む第一の混合物を提供すること；

(i) 第一の標的核酸の 5' 末端と同一の配列を含む第一の複数の末端オリゴヌクレオチド、第一の標的核酸の 3' 末端と同一の配列を含む第二の複数の末端オリゴヌクレオチド、および各々の末端および内部オリゴヌクレオチドが次のオリゴヌクレオチドと重複した領域を有する、第一の標的核酸の異なる部分と同一の配列を各々が含む第一の複数の内部オリゴヌクレオチドを含む、オリゴヌクレオチドの第一のプール、ここで第一のプール中の末端および内部オリゴヌクレオチドはともに第一の標的核酸を含み、各々の末端および内部オリゴヌクレオチドはさらに、共通プライマー認識部位および制限エンドヌクレアーゼ認識部位を含むフランキング配列を含む；

(i i) 制限エンドヌクレアーゼ認識部位を認識する制限酵素、

(b) 第一の混合物をリガーゼに曝して、リガーゼを含む第一の混合物を同時の制限酵素消化およびライゲーションを促進するために適した条件に供し、それにより、その構造が制限酵素の基質ではない、予め定義された配列を有する、第一の標的核酸を生成すること、

を含む方法。

【請求項 2】

第一の標的核酸が：(i) 第二の標的核酸配列の部分と一致する内部配列、(i i) 各

々のフランキンク配列が第二の共通プライマー認識部位および第二の制限酵素認識部位を含む、内部配列の5'末端および3'末端に隣接する5'および3'フランキンク配列、を含む中間核酸である、請求項1に記載の方法。

【請求項3】

制限酵素がIIS型制限酵素であり、IIS型制限酵素での消化が、複数の二本鎖構造オリゴヌクレオチドを生成し、互いに相補的な付着末端をもつ複数の二本鎖構造オリゴヌクレオチドが特有の直鎖配列にライゲーションされて、アセンブリされた第一の標的核酸を生成する、請求項1に記載の方法。

【請求項4】

段階(b)の後に、第一のオリゴヌクレオチドの5'末端および第二のオリゴヌクレオチドの3'末端のプライマー認識部位を認識できるプライマー対を用いて、生成した第一の標的核酸を増幅することをさらに含む、請求項1に記載の方法。

【請求項5】

第一の標的核酸を、その正確性を確認するためにシーケンシングすること、および第一の標的核酸を単離することをさらに含む、請求項1に記載の方法。

【請求項6】

(c)以下を含む第二の混合物を提供すること：

(i)第二の標的核酸の5'末端と同一の配列を含む第三の複数の末端オリゴヌクレオチド、第二の標的核酸の3'末端と同一の配列を含む第四の複数の末端オリゴヌクレオチド、および第二のプール中の各々の末端および内部オリゴヌクレオチドが第二のプール中の次のオリゴヌクレオチドと重複した領域を有する、第二の標的核酸の異なる部分と同一の配列を含む第二の複数の内部オリゴヌクレオチドを含む、オリゴヌクレオチドの第二のプール、ここで第二のプール中の末端および内部オリゴヌクレオチドはともに第二の標的核酸を含み、各々の末端および内部オリゴヌクレオチドはさらに、第二の共通プライマー認識部位および第二の制限エンドヌクレアーゼ認識部位を含む第二のフランキンク配列を含む；

(ii)第二の制限エンドヌクレアーゼ認識部位を認識する、第二の制限酵素、および

(d)第二の混合物をリガーゼに曝して、リガーゼを含む第二の混合物を同時の制限酵素消化およびライゲーションを促進するために適した条件に供し、それによりその構造が第二の制限酵素の基質ではない、予め定義された配列を有する、第二の標的核酸を生成すること、

をさらに含む、請求項2に記載の方法。

【請求項7】

予め定義された配列を有する標的核酸のアセンブリのための組成物であって、

(a)標的核酸の5'末端と同一の配列を含む第一の複数のオリゴヌクレオチド、標的核酸の3'末端と同一の配列を含む第二の複数のオリゴヌクレオチド；および各々のオリゴヌクレオチドが次のオリゴヌクレオチドと重複した領域を有する、標的核酸の異なる部分と同一の配列を含む1つ以上の複数のオリゴヌクレオチド、を含む、少なくとも1つのオリゴヌクレオチドのプール、ここでプール中のオリゴヌクレオチドはともに標的核酸を含み、各々の末端および内部オリゴヌクレオチドはさらに共通プライマー認識部位および制限エンドヌクレアーゼ認識部位を含むフランキンク配列を含む；

(b)(a)の制限エンドヌクレアーゼ認識部位を認識する制限エンドヌクレアーゼ；

(c)リガーゼ；および

(d)オリゴヌクレオチドに隣接する共通配列を除去する際に、第一の複数のオリゴヌクレオチドと親和性のある5'末端および第二の複数のオリゴヌクレオチドと親和性のある3'末端を有する直鎖ベクター、を含む、組成物。

【請求項8】

オリゴヌクレオチドが増幅されるか、エラー修正されるか、または増幅およびエラー修正される、請求項7に記載の組成物。

10

20

30

40

50

【請求項 9】

制限エンドヌクレアーゼまたは制限酵素が I I S 型制限エンドヌクレアーゼである、請求項 1 に記載の方法または請求項 7 に記載の組成物。

【請求項 10】

第一の標的核酸が最終的な標的核酸構築物である、請求項 1 に記載の方法。

【請求項 11】

第二の標的核酸を、その正確性を確認するためにシーケンシングすること、および第二の標的核酸を単離することをさらに含む、請求項 6 に記載の方法。

【請求項 12】

段階 (c) の第二の混合物を提供する前に、共通プライマー認識部位と相補的な共通プライマーを用いて、段階 (c) (i) のオリゴヌクレオチドを増幅し、段階 (c) (i) のオリゴヌクレオチドの第二のプールを生成することをさらに含む、請求項 6 に記載の方法。

10

【請求項 13】

オリゴヌクレオチドの第二のプールが、ミスマッチを含む二本鎖オリゴヌクレオチドと選択的に結合し、切断するミスマッチ結合剤と接触される、請求項 6 に記載の方法。

【請求項 14】

第二の制限酵素が、I I S 型制限酵素である、請求項 6 に記載の方法。

【請求項 15】

第二の標的核酸が、最終的な標的核酸構築物である、請求項 6 に記載の方法。

20

【請求項 16】

第二の標的核酸が中間核酸であり、中間核酸が、(i) 第三の標的核酸配列の部分と一致する内部配列、(ii) 各々のフランキング配列が第三の共通プライマー認識部位および第三の制限エンドヌクレアーゼ認識部位を含む、内部配列の 5' 末端および 3' 末端に隣接する 5' および 3' フランキング配列、を含むものである、請求項 6 に記載の方法。

【請求項 17】

段階 (c) および (d) が、最終的な標的核酸構築物が生成されるまで反復して繰り返される、請求項 6 に記載の方法。

【発明の詳細な説明】

【技術分野】

30

【0001】

関連出願

本願は、2012年6月25日に出願された米国仮出願第61/664,118号、および2012年11月30日に出願された米国仮出願第61/731,627号の優先権およびそれらからの利益を主張し、それらの各々は出典明示によりその全文が本明細書に組み込まれる。

【0002】

技術分野

本明細書において提供される方法および装置は、予め定義された配列を有する核酸および核酸ライブラリの高忠実度の合成およびアセンブリに関する。より具体的に、方法および装置はポリヌクレオチド合成、エラー減少および/または高処理配列照合 (high throughput sequence verification) のために提供される。

40

【背景技術】

【0003】

背景

自然界から複製および増幅され、次いで、成分部分に分解される DNA 配列に、組換え DNA 化学の技術を使用することは、今では一般的なことである。次いで、配列は、成分部分として、新規 DNA 配列に組み換えられるか、または再アセンブリされる。しかし、天然に入手可能な配列に対する確実性のために、研究者によって探索され得る可能性は大

50

きく制限される。今では、短いDNA配列が、個々のヌクレオシドから直接合成されることは可能であるが、一般には、大きなセグメントまたはポリヌクレオチドのアセンブリ、すなわち、約400塩基対より長いポリヌクレオチド配列を直接構築することは非実用的なことであった。

【0004】

オリゴヌクレオチド合成は、マイクロチップでの超並列カスタム合成によって実施できる(Zhou et al. (2004) Nucleic Acids Res. 32:5409; Fodor et al. (1991) Science 251:767)。しかし、現行のマイクロチップは、極めて小さい表面積しかなく、したがって、少量のオリゴヌクレオチドしか製造できない。オリゴヌクレオチドは、溶液中に放出されると、配列あたりピコモル以下の濃度、二分子のプライミング反応を効率的に駆動するのに十分に高くはない濃度で存在する。少数の変異型核酸をアセンブリするための現行法では、多数の特定の变異型を生成するための費用対効果の高い方式までスケールアップすることができない。そのようなものとして、高忠実度の遺伝子アセンブリおよび多数のポリヌクレオチド配列の製造のための、改善された、費用効率が低い方法およびデバイスが依然として必要である。

10

【0005】

さらに、マイクロチップ上のオリゴヌクレオチドは、通常化学反応を経て合成される。誤った化学反応はオリゴヌクレオチド中の無作為の塩基の誤りの原因となる。化学的核酸合成の重大な制限は、エラー率である。化学的に合成されたオリゴヌクレオチドのこのエラー率(例えば、100塩基中に1つの率での欠失および400塩基中に約1のミスマッチおよび挿入)は、存在する核酸を置換する酵素的な手段(例えばPCR)を経て得られるエラー率を超える。したがって、費用対効果の高い方法で、高収率高忠実度のポリヌクレオチドを生成する新たな技術が急ぎ必要とされている。

20

【発明の概要】

【課題を解決するための手段】

【0006】

概要

本発明の態様は、高忠実度ポリマーを調製および/またはアセンブリするための方法、システムおよび構成に関する。また、核酸アセンブリ反応を処理し、核酸をアセンブリするためのデバイスおよび方法も本明細書において提供される。カスタムポリヌクレオチドを合成するための実用的、経済的方法を提供することが本発明の目的である。当分野で既知の方法により製造された合成ヌクレオチドよりも低いエラー率で合成ポリヌクレオチドを製造する方法を提供することも、本願のさらなる目的である。

30

【0007】

いくつかの実施形態によると、本発明はあらかじめ定義された配列を有する標的核酸の製造方法を提供する。いくつかの実施形態において、方法は複数のオリゴヌクレオチド、ここで、各々のオリゴヌクレオチドは(i)標的核酸の配列の異なる部分と同一の内部配列、(ii)内部配列の5'末端に隣接する5'配列および内部配列の3'末端に隣接する3'フランキング配列、ここで各々のフランキング配列は、プライマー対に対応するプライマー認識部位および制限酵素認識部位を含む、を含む。いくつかの実施形態において、方法はさらに、プライマー対を用いてオリゴヌクレオチドの少なくともサブセットを増幅し、複数の増幅オリゴヌクレオチドを生成することを含む。複数の増幅オリゴヌクレオチドは、単一のプール中で制限酵素およびリガーゼに暴露され、ここで、制限酵素は制限酵素認識部位を認識することができ、それにより標的核酸を生成する、ことができる。

40

【0008】

いくつかの実施形態において、方法は、アセンブリされた標的核酸を配列照合(sequence verification)に供することを含む。いくつかの実施形態において、増幅二本鎖オリゴヌクレオチドは配列エラーまたはミスマッチを含み得る。いくつかの実施形態において、方法は、複数の増幅オリゴヌクレオチドをエラー除去に供するこ

50

とを含む。いくつかの実施形態において、複数の増幅オリゴヌクレオチドはミスマッチ結合剤と接触され得る。ミスマッチ結合剤は、ミスマッチを含む二本鎖オリゴヌクレオチドに選択的に関連し、結合および切断作用をもたらすことができる。いくつかの実施形態において、複数の増幅オリゴヌクレオチドはミスマッチ認識剤、例えばリジン、ピペリジンなどの化学剤と接触され得る。

【0009】

いくつかの実施形態において、制限酵素およびリガーゼが増幅オリゴヌクレオチドの単一のプールに切断およびライゲーションの促進に適した条件下で添加され、それによってアセンブリされた標的核酸配列およびそのフランキング領域を含む混合物を生成する。いくつかの実施形態において、各々のフランキング領域は共通のプライマー認識部位を含む。いくつかの実施形態において、制限酵素はIIS型制限酵素である。IIS型制限酵素での消化は、複数の付着末端二本鎖構造オリゴヌクレオチドを作製でき、複数の付着末端二本鎖構造オリゴヌクレオチドは特有の直鎖配列にライゲーションできる。

【0010】

いくつかの実施形態において方法は、標的核酸配列の5'末端および3'末端のプライマー認識部位を認識できるプライマー対を用いて標的核酸を増幅することを含む。いくつかの実施形態において方法は、標的核酸配列を、その正確性を確認するために、例えば高処理シーケンシングにより、シーケンシングすることを含む。いくつかの実施形態において方法は、核酸配列のプールからあらかじめ定義された配列を有する少なくとも一つの標的核酸を単離することを含む。

【0011】

いくつかの実施形態によると、本発明は単離された核酸をさらに処理するための方法を提供する。いくつかの実施形態において、方法は少なくとも二つの標的核酸をアセンブリすることを含む。アセンブリ工程は、階層的アセンブリによってもよい。いくつかの実施形態において、少なくとも二つの標的核酸が制限酵素消化およびライゲーションに供され、それによって長い標的核酸構築物、例えば少なくとも約10キロベースまたは100キロベース長、を形成する。

【0012】

いくつかの実施形態によると、本発明はあらかじめ定義された配列を有する標的核酸をベクター中で作製する方法を提供する。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が(i)標的核酸配列の異なる部分と一致する内部配列、(ii)各々のフランキング配列がプライマー対に対するプライマー認識部位および制限エンドヌクレアーゼに対する制限酵素認識部位を含む、内部配列の5'末端に隣接する5'フランキング配列および内部配列の3'末端に隣接する3'フランキング配列、を含む、複数のオリゴヌクレオチドが提供される。いくつかの実施形態において、オリゴヌクレオチドの少なくともサブセットをプライマー対を用いて増幅でき、それによって複数の増幅オリゴヌクレオチドを生成する。いくつかの実施形態において、複数の増幅オリゴヌクレオチドは、エラー除去および/または修正に供することができる。いくつかの実施形態において、制限エンドヌクレアーゼに対する制限酵素認識部位を有する環状ベクターが提供される。いくつかの実施形態において、複数の増幅オリゴヌクレオチドおよび環状ベクターが単一のプールで制限酵素認識部位を認識できる制限酵素およびリガーゼに曝され、それによって標的核酸をベクター中にアセンブリする。いくつかの実施形態において、方法はさらに、ベクターを宿主細胞の形質転換に用いることおよび標的核酸配列の配列照合を含む。

【0013】

いくつかの実施形態によると、本発明はあらかじめ定義された配列を有する標的核酸のアセンブリのための組成物を提供する。いくつかの実施形態において、組成物は、各々のオリゴヌクレオチドが(i)標的核酸配列の異なる部分と一致する内部配列、(ii)各々のフランキング配列がプライマー対に対するプライマー認識部位および制限エンドヌクレアーゼに対する制限酵素認識部位を含む、内部配列の5'末端に隣接する5'フランキング配列および内部配列の3'末端に隣接する3'フランキング配列、を含む、複数のオ

リゴヌクレオチドを含む。いくつかの実施形態において、組成物はさらに制限酵素および/またはリガーゼを含む。いくつかの実施形態において、組成物はさらに制限エンドヌクレアーゼに対する酵素認識部位の対を含む。いくつかの実施形態において、制限エンドヌクレアーゼはIIS型制限エンドヌクレアーゼである。

【0014】

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が増幅および/またはエラー修正される。

【0015】

本発明のいくつかの態様において、あらかじめ定義された配列を有する標的核酸を作製するための方法は、(i)制限酵素、および(ii)標的核酸の5'末端と同一の配列を含む第一のオリゴヌクレオチド、標的核酸の3'末端と同一の配列を含む第二のオリゴヌクレオチド；および各々のオリゴヌクレオチドが次のオリゴヌクレオチドの配列領域に対応する重複した配列領域を有する、標的核酸の配列の異なる部分と同一の配列を含む複数のオリゴヌクレオチドを含む第一のプール、ここで第一のプール中のオリゴヌクレオチドはともに標的核酸配列を含む；を含む第一の混合物を提供し；混合物をリガーゼに曝して、それにより標的核酸を生成すること、を含む。標的核酸はその後配列照合に供することができる。

【0016】

いくつかの実施形態において、本発明の方法は構築オリゴヌクレオチドのプールの提供を含み、異なる段階でのオリゴヌクレオチドの増幅に関する。「構築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それ自身より長い核酸分子のアセンブリに使用され得る一本鎖オリゴヌクレオチドのことをいう。構築オリゴヌクレオチドは、一本鎖オリゴヌクレオチドまたは二本鎖オリゴヌクレオチドであり得る。いくつかの実施形態において、構築オリゴヌクレオチドは合成オリゴヌクレオチドであり得、基質と並列に合成され得る。

【0017】

いくつかの実施形態において、方法は第一の混合物の提供の前に、さらに各々のオリゴヌクレオチドが(i)標的核酸配列の異なる部分と一致する内部配列、(ii)各々のフランキンク配列がプライマー対に対するプライマー認識部位および制限酵素認識部位を含む、内部配列の5'末端に隣接する5'フランキンク配列および内部配列の3'末端に隣接する3'フランキンク配列、を含む、複数の構築オリゴヌクレオチドを提供する工程を含む。いくつかの実施形態において、各々のフランキンク領域は共通のプライマー認識部位を含み得る。いくつかの実施形態において、複数の構築オリゴヌクレオチドは増幅され得る。いくつかの実施形態において、オリゴヌクレオチドは配列エラーまたはミスマッチを含み得る。いくつかの実施形態において、複数の増幅オリゴヌクレオチドは、エラー除去に供され得る。例えば、複数の増幅オリゴヌクレオチドは、ミスマッチを含む二本鎖オリゴヌクレオチドと選択的に結合し、切断するミスマッチ結合剤と接触され得る。

【0018】

いくつかの実施形態において、制限酵素およびリガーゼが増幅オリゴヌクレオチドの単一のプールに消化およびライゲーションの促進に適した条件下で添加でき、それによってアセンブリされた標的核酸配列およびそのフランキンク領域を含む混合物を生成する。いくつかの実施形態において、制限酵素はIIS型制限酵素であり、IIS型制限酵素での消化は、複数の付着末端二本鎖構造オリゴヌクレオチドを作製でき、複数の付着末端二本鎖構造オリゴヌクレオチドは特有の直鎖配列にライゲーションされる。

【0019】

いくつかの実施形態において、方法は、第一のオリゴヌクレオチドの5'末端および第二のオリゴヌクレオチドの3'末端のプライマー認識部位を認識できるプライマー対を用いて標的核酸を増幅することを含む。

【0020】

いくつかの実施形態において、方法は、標的核酸配列を、その正確性を確認するために

10

20

30

40

50

、例えば高処理シーケンシングにより、シーケンシングすることをさらに含む。

【0021】

いくつかの実施形態において、方法は、核酸配列のプールからあらかじめ定義された配列を有する少なくとも一つの標的核酸を単離することをさらに含む。

【0022】

いくつかの実施形態において、方法は、標的核酸配列を処理することをさらに含む。

【0023】

いくつかの実施形態において、方法は、(i)制限酵素、および(ii)標的核酸の5'末端と同一の配列を含む第一のオリゴヌクレオチド、標的核酸の3'末端と同一の配列を含む第二のオリゴヌクレオチド；および各々のオリゴヌクレオチドが次のオリゴヌクレオチドの配列領域に対応する重複した配列領域を有する、標的核酸の配列の異なる部分と同一の配列を含む複数のオリゴヌクレオチドを含む第二のプール、ここで第二のプール中のオリゴヌクレオチドはともに第二の標的核酸配列を含む；を含む第二の混合物を提供することを含む。いくつかの実施形態において、第二の混合物はリガーゼに曝され、それにより第二の標的核酸を生成する。いくつかの実施形態において、第一のプールの第二のオリゴヌクレオチドは、制限エンドヌクレアーゼについての制限エンドヌクレアーゼ認識部位を含み、第二のプールの第一のオリゴヌクレオチドは、制限エンドヌクレアーゼについての制限エンドヌクレアーゼ認識部位を含む。

10

【0024】

いくつかの実施形態において、方法は、少なくとも2つの標的核酸をアセンブリすることを含む。いくつかの実施形態において、アセンブリの工程は階層的アセンブリによる。いくつかの実施形態において、少なくとも2つの標的核酸は制限エンドヌクレアーゼ消化およびライゲーションに供され、それによって長い標的核酸構築物を形成する。いくつかの実施形態において、長い標的核酸構築物は少なくとも約10キロベース長または少なくとも約100キロベース長である。

20

【0025】

いくつかの態様において、本発明は、標的核酸の5'末端と同一の配列を含む第一のオリゴヌクレオチド、標的核酸の3'末端と同一の配列を含む第二のオリゴヌクレオチド；および各々のオリゴヌクレオチドが次のオリゴヌクレオチドの配列領域に対応する重複した配列領域を有する、標的核酸の配列の異なる部分と同一の配列を含む1つ以上のオリゴヌクレオチド、共に標的核酸を含む複数のオリゴヌクレオチド；プライマー対についてのプライマー認識部位および制限エンドヌクレアーゼ認識部位を含む複数の共通配列、を含む複数のオリゴヌクレオチドを含み、あらかじめ定義された配列を有する標的配列のアセンブリのための組成物に関する。いくつかの実施形態において、組成物は制限エンドヌクレアーゼおよび/またはリガーゼをさらに含む。制限エンドヌクレアーゼは、IIS型制限エンドヌクレアーゼであり得る。

30

【0026】

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増幅および/またはエラー修正され得る。いくつかの実施形態において、組成物は第一のオリゴヌクレオチドと親和性のある5'および第二のオリゴヌクレオチドと親和性のある3'末端を有する直鎖ベクターをさらに含み得る。

40

【図面の簡単な説明】

【0027】

【図1】図1は、本発明の一実施形態による高忠実度核酸アセンブリの例示的な方法を示す。

【0028】

【図2】図2は、あらかじめ定義された配列を有するポリヌクレオチドのアセンブリ方法の非限定的な例を示す。

【0029】

【図3】図3は、あらかじめ定義された配列を有するポリヌクレオチドのベクターへのア

50

センブリ方法の非限定的な例を示す。

【 0 0 3 0 】

【図 4】図 4 は、あらかじめ定義された配列を有するポリヌクレオチドの階層的アセンブリ方法の非限定的な例を示す。

【 0 0 3 1 】

【図 5】図 5 は、制限エンドヌクレアーゼ認識部位（下線）をもつプラスミド p G 9 - 1 の核酸配列を示す。

【 0 0 3 2 】

【図 6】図 6 は、配列照合の非限定的な例示的な方法を示す。

【発明を実施するための形態】

10

【 0 0 3 3 】

発明の詳細な説明

本発明の態様は、核酸アセンブリ反応を最適化して、不正確にアセンブリされた核酸の数を減少させるために有用であり得る。本発明の方法および組成物は、あらかじめ定義された配列を有する標的配列を得る工程を容易にできる。従って、本発明の方法および組成物は、正しくアセンブリされた核酸を得る確率を増加させ、それによってあらかじめ定義された配列を有する核酸の作製に関するコストおよび時間を削減し得る。

【 0 0 3 4 】

本発明の態様は、1つ以上の初期または中間アセンブリ反応の収率を向上させるために使用され得る。いくつかの実施形態において、本発明の方法および組成物は、例えば酵素消化、精製およびライゲーション工程などの多数のアセンブリ工程の分離を要求すること
20
を避けることにより、総合的なアセンブリ手順の効率を高めることができる。従って、本発明のいくつかの態様は、予測できる、および/または、信頼できるアセンブリ法を許容し、遺伝子合成に必要な時間および工程を有意に減少させ、中間産物または最終核酸製品の収率および/または精度を増加させることができる。

【 0 0 3 5 】

本発明のいくつかの態様において、アセンブリ処理は、1つ以上のアセンブリ工程を妨げると知られているか、または予測される配列特性（feature）に適応できる核酸アセンブリ戦略を設計および実行することを含む。例えば、合成された核酸配列は、1つ以上のアセンブリ工程を妨げることができる、繰り返し配列、有意に高いまたは低い GC
30
含量を有する配列、および/または二次構造に関連する他の配列などの、配列特性を分析できる。当業者は、ある配列特性が多重アセンブリ反応（例えば、ポリメラーゼベースの伸長反応）を妨げ得る、および/または望まないアセンブリ産物の形成を促進し得ることによって、正しい核酸産物のアセンブリを減少または防止することを理解するであろう。いくつかの実施形態において、標的核酸配列中に複数の干渉配列特性が同定されるならば、有用な戦略は、アセンブリの間に干渉配列特性を分離することを伴い得る。例えば、目標核酸は少数の干渉配列（例えば、0、1、2、または3）だけを含むように設計された、複数の中間断片または構成要素を伴う工程でアセンブリされ得る。いくつかの実施形態において、各々の中間断片または構成要素は、最大で1つの干渉配列特性を含み得る。従
40
って、各々の中間断片は、効率的にアセンブリされ得る。いくつかの実施形態において、核酸断片または構成要素の設計は、それら5'および/または3'末端から干渉配列特性を除去し得る。従って、干渉配列特性は、アセンブリ反応に使用するために設計された隣接する開始核酸間の相補的に重複する部分から除去され得る。これは、核酸の正しいアセンブリに重要な、配列特異的ハイブリダイゼーション反応の干渉を防止または減少し得る。いくつかの実施形態において、それは構成要素の3'および/または5'末端すぐからの干渉配列特性の除去に十分であり得る。例えば、干渉配列特性が3'末端および/または5'末端からの少なくとも1つのヌクレオチド、好ましくは構成要素の3'末端および/または5'末端から2、3、4、5またはそれ以上のヌクレオチド（例えば、5 - 10、10 - 15、15 - 20、または、それより多くのヌクレオチド）に位置し得る。

【 0 0 3 6 】

50

本発明の態様は、インビトロおよび／またはインビボ核酸アセンブリ手順と併用して用いられ得る。

【0037】

本明細書において提供された方法および組成物の態様は、核酸合成およびアセンブリ反応の正確度、収率、処理能力および／または費用効率の増大のために有用である。本明細書において、用語「核酸」、「ポリヌクレオチド」、「オリゴヌクレオチド」は、同義的に使用され、ヌクレオチドの天然に存在するか、または合成ポリマーの形態を指す。本発明のオリゴヌクレオチドおよび核酸分子は、天然に存在するヌクレオチドから形成され得、例えば、デオキシリボ核酸(DNA)またはリボ核酸(RNA)分子を形成する。あるいは、天然に存在するオリゴヌクレオチドは、例えば、ペプチド核酸(PNA)において、またはロックド核酸(LNA)において、その特性を変更するために構造修飾を含み得る。天然に存在する塩基か、または人工塩基を用いるオリゴヌクレオチドおよび核酸分子の固相合成は、当技術分野で周知である。この用語は、ヌクレオチド類似体から製造されたRNAまたはDNAのいずれかの等価物、類似体を含み、記載されている実施形態、一本鎖または二本鎖ポリヌクレオチドに適応できると理解されるべきである。本発明において有用なヌクレオチドとして、例えば、天然に存在するヌクレオチド(例えば、リボヌクレオチドまたはデオキシリボヌクレオチド)またはヌクレオチドの天然もしくは合成修飾または人工塩基が挙げられる。本明細書において、用語モノマーとは、接続され、接続され得、一緒に2以上のメンバーからなるオリゴマー、ポリマーまたは化合物を形成する一連の小分子のメンバーを指す。ポリマー内のモノマーの特定の順序付けは、本明細書において、ポリマーの「配列」と呼ばれる。モノマーのセットとして、限定されるわけではないが、例えば、一般的なL-アミノ酸のセット、D-アミノ酸のセット、合成および／または天然アミノ酸のセット、ヌクレオチドのセットならびにペントースおよびヘキソースのセットが挙げられる。本発明の態様は、主に、オリゴヌクレオチドの調製に関して本明細書に記載されているが、ペプチドまたはポリペプチド、多糖、リン脂質、ヘテロポリマー、ポリエステル、ポリカーボネート、ポリ尿素、ポリアミド、ポリエチレンイミン、ポリアリーレンスルフィド、ポリシロキサン、ポリイミド、ポリアセテートまたは任意のその他のポリマーなどのその他のポリマーの調製において容易に適用され得る。

【0038】

標的核酸

本明細書において、用語「予め定義された配列」は、ポリマーの配列が公知であり、ポリマーの合成またはアセンブリの前に選択されることを意味する。特に、本発明の態様は、主に核酸分子の調製に関して本明細書に記載されており、核酸の配列は、公知であり、核酸分子の合成またはアセンブリの前に選択される。本明細書において提供された技術のいくつかの実施形態では、固定されたオリゴヌクレオチドまたはポリヌクレオチドは、材料の供給源として使用される。種々の実施形態において、本明細書に記載された方法は、複数のオリゴヌクレオチドを使用し、各々の配列は、合成される最終ポリヌクレオチド構築物の配列に基づいて決定される。一の実施形態において、オリゴヌクレオチドは、短い核酸分子である。例えば、オリゴヌクレオチドは、10～約300ヌクレオチド、20～約400ヌクレオチド、30～約500ヌクレオチド、40～約600ヌクレオチド、または約600超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、種々の長さを有するように設計され得る。いくつかの実施形態では、ポリヌクレオチド構築物の配列は、並列して合成し、単一または複数の所望のポリヌクレオチド構築物にアセンブリできる複数のより短い配列に、本明細書に記載された方法を使用して分割できる。

【0039】

いくつかの実施形態において、標的核酸は天然に存在する遺伝子および／または天然に存在する核酸(例えば天然に存在するコード配列、調節配列、非コード配列、テロメアおよびセントロメアなどの染色体構造配列、それらのいかなる断片またはそれらの2つ以上の組み合わせ)または天然に存在しない配列を有し得る。いくつかの実施形態において、

標的核酸は1以上の位置で天然配列と異なる配列を有するように設計され得る。他の実施形態において、標的核酸は全体として新規な配列を有するように設計され得る。しかしながら、標的核酸は、1以上の天然に存在する配列、天然に存在しない配列、またはそれらの組み合わせを含み得ることが理解されるべきである。

【0040】

いくつかの実施形態において、所定の配列変動を有する核酸を含有するライブラリをアセンブリする方法が、本明細書において提供される。本明細書で提供されるアセンブリ戦略を使用して、多数の異なる対象とする核酸配列を代表する極めて大きなライブラリを作製できる。例えば、本明細書で提供される方法は、10を超える異なる配列変異体を有するライブラリをアセンブリするために用いられ得る。いくつかの実施形態において、核酸のライブラリは、配列変異体のライブラリである。配列変異体は、単一の天然に存在するタンパク質をコードする配列の変異体であり得る。しかしながら、いくつかの実施形態において、配列変異体は、複数の異なるタンパク質をコードする配列の変異体であり得る。したがって、本明細書で提供される技術の一態様は、正確な高密度核酸ライブラリの調製のためのアセンブリ戦略の設計に関する。本明細書で提供される技術の他の態様は、正確な高密度核酸ライブラリのアセンブリに関する。本明細書で提供される技術の態様はまた、正確な高密度核酸ライブラリを提供する。高密度核酸ライブラリは、100種を超える種々の配列変異体（例えば、約 10^2 ～ 10^3 、約 10^3 ～ 10^4 、約 10^4 ～ 10^5 、約 10^5 ～ 10^6 、約 10^6 ～ 10^7 、約 10^7 ～ 10^8 、約 10^8 ～ 10^9 、約 10^9 ～ 10^{10} 、約 10^{10} ～ 10^{11} 、約 10^{11} ～ 10^{12} 、約 10^{12} ～ 10^{13} 、約 10^{13} ～ 10^{14} 、約 10^{14} ～ 10^{15} 種またはより多くの種々の配列）を含むことができ、ここでは、ランダム配列とは対照的に、種々の配列のうちの高いパーセンテージが特定の配列である（例えば、配列の約50%超、約60%超、約70%超、約75%超、約80%超、約85%超、約90%超、約91%、約92%、約93%、約94%、約95%、約96%、約97%、約98%、約99%またはそれ以上が、対象の所定の配列である）。

【0041】

ある実施形態において、標的核酸は機能的配列（例えば、タンパク質結合配列、調節配列、機能性タンパク質をコードする配列など、およびそれらのいかなる組み合わせ）を含み得る。しかしながら、いくつかの実施形態において標的核酸は特定の機能的配列を欠失し得る（例えば、標的核酸はタンパク質結合配列、調節配列、またはタンパク質コード配列の非機能的断片または変異体、または他のいかなる非機能的天然または合成配列、またはそれらのいかなる非機能的組み合わせを含み得る）。ある標的核酸配列は、機能的および非機能的配列の両方を含み得る。標的核酸およびその使用の、これらおよび他の態様は本明細書により詳細に記載される。

【0042】

いくつかの実施形態において、標的核酸はただ一つの多重アセンブリ反応（例えば、ただ一つのオリゴヌクレオチドアセンブリ反応）でアセンブリされ得る。しかしながら、標的核酸はまた、多くの核酸断片からアセンブリされ得る。それは別々のオリゴヌクレオチドアセンブリ反応でそれぞれ生成され得る。多重オリゴヌクレオチドアセンブリで生成された1つ以上の核酸断片が、いくつかの実施形態において、標的核酸を形成するために他の源（例えば、制限断片、核酸増幅産物など）から得られる1つ以上の核酸分子に合成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、第一の反応でアセンブリされた標的核酸は、より大きい標的核酸を作製するための連続アセンブリ反応についての入力核酸断片として使用され得る。本明細書に使用される「多重アセンブリ」および「多重オリゴヌクレオチドアセンブリ反応」なる用語は、通常、より大きい最終核酸を作製するためにアセンブリされる複数の開始核酸（例えば、複数の少なくとも部分的に重なっている核酸）を含むアセンブリ反応をいう。

【0043】

アセンブリ処理

10

20

30

40

50

図1は本発明の一の実施形態により核酸をアセンブリするための工程を示す。初めに、配列情報を得る。配列情報は、アセンブリされるあらかじめ定義された標的核酸の配列であり得る。いくつかの実施形態において、配列は顧客からの注文の形で受領され得る。いくつかの実施形態において、配列は核酸配列（例えば、DNAまたはRNA）として受領され得る。いくつかの実施形態において、配列はタンパク質配列として受領され得る。配列はDNA配列に変換され得る。例えば、得られた配列がRNA配列であれば、対応するDNA配列を得るためにUをTに置換し得る。得られた配列がタンパク質配列であれば、タンパク質配列は、アミノ酸に適切なコドンを使用することでDNA配列に変換され得る。

【0044】

いくつかの実施形態において、配列情報は、アセンブリ戦略を決定するために、標的核酸のあらかじめ定義された配列を生成するようにアセンブリされた断片（構成要素、オリゴヌクレオチドまたは中間的断片としても参照される）の数または配列などが、分析され得る。いくつかの実施形態において、配列分析は、オリゴヌクレオチド合成、増幅またはアセンブリを妨げると知られているか、または予測される1つ以上の干渉配列特性の存在をスキャンすることを含み得る。例えば、干渉配列構造は、少なくとも10ベース長（例えば、10 - 20、20 - 50、50 - 100、または100ベース以上）の長さにわたって少ないGC含量（例えば、30%未満のGC、20%未満のGC、10%未満のGCなど）を有する配列、または二次構造またはステムループ構造を形成し得る配列であり得る。一旦このフィルタを通過すると、核酸配列はオリゴヌクレオチド構成要素などの、より小さい断片に分割できる。

【0045】

いくつかの実施形態において、構造定量化および構文解析工程後にアセンブリのための合成オリゴヌクレオチドが設計され得る（例えば、配列、サイズ、および数）。合成オリゴヌクレオチドは、標準のDNA合成化学（例えば、ホスホロアミダイト法）を使用することで生成できる。合成オリゴヌクレオチドは、本明細書に詳細に説明されるように、当分野で知られているいかなる適切な技術を使用して、例えばマイクロアレイなどの固相担体上で合成され得る。オリゴヌクレオチドは増幅に供される前にマイクロアレイから溶出されるか、またはマイクロアレイ上で増幅できる。異なるオリゴヌクレオチドが異なる長さを持つように設計され得ることが理解されるべきである。

【0046】

いくつかの実施形態において、各々の標的配列のための構成要素オリゴヌクレオチドを増幅できる。例えば、オリゴヌクレオチドは、それらの3'末端および5'末端にプライマー結合配列を有するように設計でき、オリゴヌクレオチドは適切なプライマー対を用いてポリメラーゼ連鎖反応（PCR）によって増幅できる。

【0047】

合成オリゴヌクレオチドには配列エラーがあり得ることが理解されるべきである。従って、オリゴヌクレオチド調製は、本明細書に詳細に説明されるように、エラーを含む分子を除去するために選択またはスクリーニングされ得る。誤りを含むオリゴヌクレオチドは、両方の鎖上にエラーを有する（すなわち、両方の鎖における、不正確な相補的ヌクレオチド、欠失、または追加）二本鎖のホモ二本鎖であり得る。いくつかの実施形態において、配列エラーは二本鎖核酸の変性および再アニーリングを含む技術を使用することで除去され得る。いくつかの実施形態において、各々個別のエラーを含む核酸が、核酸調製において、同じ位置における正しい配列を有する核酸より低頻度で存在しているならば、相補的エラーを含む核酸の一本鎖は共に再アニーリングされないであろう。むしろ、エラーを含む一本鎖は、エラーを全く含まないか、または1つ以上の異なった誤りを含む相補鎖と再アニーリングされ得る。結果として、エラーを含む鎖は、再アニーリング反応産物中にヘテロ二本鎖分子の形態となり得る。エラーのない鎖はエラーを含む鎖または他のエラーのない鎖と再アニーリングされ得る。再アニーリングされたエラーのない鎖は、再アニーリングされた試料でホモ二本鎖を形成する。従って、オリゴヌクレオチドの再アニーリン

10

20

30

40

50

グ調製物からヘテロ二本鎖分子を除去することにより、エラー含有核酸の量または頻度が減少され得る。クロマトグラフィー、電気泳動、ヘテロ二本鎖分子の選択的結合などを含む、当分野で知られているヘテロ二本鎖分子を除去するためのいかなる適切な方法も使用され得る。いくつかの実施形態において、ヘテロ二本鎖核酸分子に選択的に（例えば特異的に）結合するミスマッチ結合タンパク質が使用され得る。いくつかの実施形態において、ミスマッチ結合タンパク質は、二本鎖オリゴヌクレオチドまたはポリヌクレオチド上で、溶液中または支持体上に固定化されて使用され得る。

【0048】

いくつかの実施形態において、エラーを含むオリゴヌクレオチドは、例えば、MutS、MutSホモログ、またはそれらの組み合わせを使用するMutS濾過処理を使用することによって除去される。大腸菌（*E. coli*）では、ホモ二量体として機能するように見えるMutSタンパク質がミスマッチ認識要素として機能する。真核生物では、少なくとも3つのMutSホモログ（MSH）タンパク質、すなわち、MSH2、MSH3、MSH6が同定され、それらはヘテロ二量体を形成する。例えば酵母、サッカロマイセス・セレビシエ（*Saccharomyces cerevisiae*）では、どちらも実質的に冗長な機能を示すが、MSH2-MSH6複合体（MutSアルファとしても知られている）は塩基ミスマッチおよび単一ヌクレオチド挿入/欠失ループを認識し、MSH2-MSH3複合体（MutSベータとしても知られている）は12-16ヌクレオチドの挿入/欠失を認識する。ミスマッチ結合タンパク質は、組換え型または天然の源から得られ得る。ミスマッチ結合タンパク質は熱安定性であり得る。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ミスマッチ結合タンパク質が使用され得る。熱安定性DNAミスマッチ結合タンパク質は：Tth MutS（サーマス・サーモフィラス（*Thermus thermophilus*）からの）、Taq MutS（サーマス・アクアティカス（*Thermus aquaticus*）からの）、Apy MutS（アクイフェクス・ピロフィラス（*Aquifex pyrophilus*）からの）、Tma MutS（サーモトガ・マリティマ（*Thermotoga maritima*）からの）、それらのホモログ、いかなる他の適当なMutSまたはそれらの2以上のいかなる組み合わせ、を含むが、これらに限定されない。

【0049】

異なる種から得られたMutSが、特定のミスマッチまたは異なるミスマッチのための異なる親和性を有することができることが示されている。いくつかの実施形態において、異なるミスマッチのための異なる親和性を有する異なるMutSの組み合わせを使用できる。

【0050】

いくつかの実施形態において、1つ以上の修復タンパク質を使用する酵素複合体が使用できる。修復タンパク質の例は、ミスマッチ認識のためのMutS、標的鎖へのニックの導入のためのMutH、およびMutHおよびMutSの相互作用を仲介するためのMutL、それらのホモログまたはそれらのいかなる組み合わせを含むが、これらに限定されない。いくつかの実施形態において、ミスマッチ結合タンパク質複合体は、MutHLS酵素複合体である。

【0051】

いくつかの実施形態において、スライディングクランプ技術がエラーのない二本鎖オリゴヌクレオチドの濃縮のために使用され得る。いくつかの実施形態において、MutSまたはそれらのホモログはDNAクランプタンパク質と相互作用できる。DNAクランプタンパク質の例は、ホモ二量体として機能できる、dnaN遺伝子によってコードされる細菌性スライディングクランプタンパク質DnaNを含むが、これに限定されない。いくつかの実施形態において、MutSタンパク質またはそれらのホモログのクランプタンパク質との相互作用は、結合ミスマッチのMutSの有効性を増加できる。

【0052】

いくつかの実施形態において、エラーを含むオリゴヌクレオチドは、例えばCEL I、

10

20

30

40

50

C E L I I、または R E S I などのそれらのホモログ、またはそれらの組み合わせなどのタンパク質の S 1 ファミリーからの酵素を使用することで除去できる。タンパク質の S 1 ファミリーからの酵素は塩基ミスマッチ、挿入、および欠失ループを認識できる。いくつかの実施形態において、そのような酵素は、ただ一つのまたは両方の D N A 鎖のいずれかを通して、優先的に認識部位が切断された後のホリデージャンクションに結合できる。いくつかの実施形態において、S 1 タンパク質の熱安定性の同等物が使用され得る。

【 0 0 5 3 】

いくつかの実施形態において、エラーを含むオリゴヌクレオチドは、ミスマッチ塩基部位に結合する小分子、化学的または無機材料を使用することで除去できる。ミスマッチ部位では、ヌクレオチド塩基は、超過螺旋状であり、化学修飾反応に影響されやすい場合がある。過マンガン酸塩、ヒドロキシルアミン、リジン、およびまたはペンタアミルテニウムなどの材料が、それぞれミスマッチしたチミンおよびシトシンを修飾する化学的切断法で使用できる。得られた修飾 D N A は、その後ピペリジンで処理され、脱塩基部位で切断をもたらすことができる。いくつかの実施形態において、切断の特異性は 2 価の塩を用いてモニターできる。

10

【 0 0 5 4 】

いくつかの実施形態において、次の段階で、エラーを修正したオリゴヌクレオチドは共通配列の配列除去およびそれに続くより長い、多重オリゴヌクレオチド構築物へのライゲーションで組み合わせられる。

【 0 0 5 5 】

20

本発明のいくつかの態様において、酵素消化による共通配列の除去段階は、ライゲーション段階と組み合わせられる。当業者は、本発明の処理が、共通配列の並列除去および標的核酸構築物へのライゲーションを許容し、酵素除去、ビーズベースの捕捉およびライゲーション逐次段階の必要性を否定することを理解するであろう。さらに、当業者は本発明の処理が以下のような、標準アセンブリ処理を超える多数の利点を有することを理解するであろう：

(1) 収率の増加。標準的な共通配列の個別の酵素的除去を使用して、反応は、さらなる除去の対象である、反応していない基質または未消化オリゴヌクレオチドがまだ存在する状態で設定時点の後に停止される。当業者は、ライゲーション反応が酵素的除去のための基質でない所望の産物を生成するので、除去およびライゲーション段階の組み合わせに、所望の産物への反応を不可逆的に進行させる効果があることを理解するであろう。

30

(2) 費用効率：共通配列の除去およびライゲーションの間の精製の必要がもはやないため、本発明のいくつかの態様による方法は費用効率が良い。精製を排除するため、本発明の方法の態様はビオチンでラベルされたプライマーの必要性を排除する。そこにはまた、非ビオチン化されたプライマーの、それらのビオチンを含む対応物上での受け取りリードタイムの減少という形もあり得る。

(3) 時間効率：酵素的共通配列除去およびライゲーションとの間の精製段階を除くことにより、遺伝子合成のための時間および段階の数が減少される。

(4) 他の配列を、それらのサイズを考慮せずに容易に加える機会。所望でない配列を除去するための精製法の一部がサイズに基づいているため、精製を排除すると、遺伝子合成のために加えられるいかなる付加配列についてもサイズの規制を除去できる。これはベクターへのワンステップライゲーション、または一般的なフランキング配列の添加を含むことができる。

40

(5) 工程は、遺伝子合成工程自体で使用する遺伝子の制限部位の使用を考慮する。以前の方法論では、切断部位が精製で除去される小さい D N A 断片をもたらすために、これらの制限部位を使用できなかった。これらの制限部位の利用を可能にすることは、再帰的な(階層的な)遺伝子合成が、より長い核酸を構築することを許容できる。

【 0 0 5 6 】

当業者は、オリゴヌクレオチドアセンブリの後にアセンブリ産物(例えば、最終的な標的核酸か中間核酸断片)が所望でない配列を含む配列を含み得ることを理解するであろう

50

。エラーはオリゴヌクレオチド合成またはオリゴヌクレオチドアセンブリの間に導入された配列エラーから、より長い核酸で生じ得る。いくつかの実施形態において、他の核酸配列（予備的インビトロクロニングとしても本明細書で参照される）から正しいあらかじめ定義された配列を有する核酸が単離できる。いくつかの実施形態において、正しい配列は、他の不正確な配列から正しい配列を選択的に単離させることにより、単離され得る。例えば、正しい配列を有する核酸は、支持体の異なる機構（feature）へ、または他のプレートに選択的に移動または輸送することができる。あるいは対象となる核酸を含む機構から不正確な配列を有する核酸を選択的に除去できる（例えば全体が本明細書に参照により組み込まれる P C T / U S 2 0 0 7 / 0 1 1 8 8 6 を参照）。

【 0 0 5 7 】

10

いくつかの実施形態において、オリゴヌクレオチド処理およびライゲーションの後に、クローン単離でアセンブリ構築物またはアセンブリ構築物のコピーを単離できる。アセンブリ構築物は、例えば、高処理シーケンシングを使用することで、配列照合できる。いくつかの実施形態において、標的核酸の配列決定は、単一分子シーケンシング、またはポロニーシーケンシングなどの標的核酸配列の増幅された集団のシーケンシングなどの、個々の分子のシーケンシング配列を使用することで実施できる。ハイブリダイゼーションシーケンシング、ライゲーションによるシーケンシングまたは合成によるシーケンシングなどの、シーケンシングのためのいかなる適切な方法も使用され得る。

【 0 0 5 8 】

本発明のいくつかの態様は、本明細書に記載された方法を使用する遺伝子合成プラットフォームに関連する。いくつかの実施形態において、これらの遺伝子合成プラットフォームは、次世代配列プラットフォーム（例えば、ハイブリダイゼーションシーケンシング、合成によるシーケンシング、ライゲーションによるシーケンシングまたは他のいかなる適当なシーケンシング法も）と組み合わせられる。

20

【 0 0 5 9 】

いくつかの実施形態において、アセンブリ手順は、複数の異なる核酸またはオリゴヌクレオチドがアセンブリされるために合成、または固定化、増幅およびこれらが組み合わせられる、いくつかの並列および/または連続した反応段階を含み得て、さらなるアセンブリ、クローニング、または他の用途（全体が参照として本明細書に組み込まれる P C T 出願 P C T / U S 0 9 / 5 5 2 6 7 を参照）に使用されるために、より長い核酸産物を生成させる（例えば、本明細書に記載される伸長またはライゲーションによる）。

30

【 0 0 6 0 】

オリゴヌクレオチド合成

いくつかの実施形態において、本明細書において提供された方法および装置は、表面または基板上に固定化されているオリゴヌクレオチド（例えば、支持体結合オリゴヌクレオチド）を使用する。本明細書において、用語「支持体」および「基板」は、同義的に使用され、核酸などのポリマーがその上で合成または固定化される多孔性または非多孔性の溶媒不溶性材料を指す。本明細書において「多孔性」とは、材料が、実質的に均一な直径（例えば、nm 範囲の）を有する孔を含有することを意味する。多孔性材料として、紙、合成フィルタなどが挙げられる。このような多孔性材料では、反応は、孔内で起こり得る。支持体は、ピン、ストリップ、プレート、ディスク、ロッド、ベンド、円柱状構造、ビーズ、ナノ粒子などを含めた粒子などのいくつかの形のいずれか 1 種を有し得る。支持体は、可変幅を有し得る。支持体は、親水性であり得るか、または親水性にされ得、これとして、それ自体によってか、またはその他の材料とともに使用される、シリカ、硫酸マグネシウムおよびアルミナなどの無機粉末；天然ポリマー材料、特に、セルロース材料および繊維を含有する紙、例えば、ろ紙、クロマトグラフィー用紙などといったセルロースに由来する材料；ニトロセルロース、酢酸セルロース、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（エチレンテレフタレート）、ナイロン、ポリ（酪酸ビニル）、ポリビニリデンジフルオリド（P V D

40

50

F) メンブレン、ガラス、多孔性ガラス (controlled pore glass)、磁性多孔性ガラス、セラミックス、金属などといった合成または修飾された天然に存在するポリマーが挙げられる。いくつかの実施形態において、オリゴヌクレオチドは、アレイ形式で合成される。例えば、一本鎖オリゴヌクレオチドは、一般的な支持体上でその場で合成され、ここでは、各オリゴヌクレオチドは、基板上の分離した、または別個の機構 (またはスポット) 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、支持体または機構の表面に結合される。本明細書において、用語「アレイ」とは、オリゴヌクレオチドまたはさらなる反応のための相補的オリゴヌクレオチドを保存、増幅および放出するための別個の機構の配置を指す。好ましい実施形態において、支持体またはアレイは、アドレス可能であり、支持体は、支持体上の特定の所定の位置 (すなわち、
10 「アドレス」) に、2 以上の別個のアドレス可能な機構を含む。したがって、アレイ上の各オリゴヌクレオチド分子は、支持体上の公知の、定義された位置に位置づけられる。各オリゴヌクレオチドの配列は、支持体上のその位置から決定できる。

【0061】

いくつかの実施形態において、オリゴヌクレオチドは、表面またはアレイの別個の機構上に、結合され、スポットされ、固定化され、表面に結合され、支持されるか、または合成される。オリゴヌクレオチドは、共有結合によって表面と結合されるか、または表面に沈着され得る。アレイは、構築され、特別注文されるか、または市販の販売会社 (例えば、Agilent、Affymetrix、Nimblegen) から購入され得る。例えば、マスクレス (maskless) アレイシンセサイザー、マスクを利用する光依存性 (light directed) 法、フローチャネル法、スポッティング法など、種々の構築方法が、当技術分野で周知である。いくつかの実施形態において、構築および /
20 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ー (MAS) を使用して固体支持体上で合成され得る。マスクレスアレイシンセサイザーは、例えば、PCT 出願番号 WO 99 / 42813、および対応する米国特許第 6,375,903 号に記載されている。アレイ中の機構の各々が、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り上げることができるマスクレス機器のその他の例が公知である。構築物および / または選択オリゴヌクレオチドを合成するためのその他の方法として、例えば、マスクを利用する光依存性法、フローチャネル法、スポッティング法、ピンベースの (pin-based) 方法および複数の支持体を利用する方法が挙げられる。オリ
30 ゴヌクレオチドの合成のための、マスクを利用する光依存性法 (例えば、VLSIPS (登録商標) 法) が、例えば、米国特許第 5,143,854 号、5,510,270 号および 5,527,681 号に記載されている。これらの方法は、固体支持体の予め定義された領域を活性化すること、次いで、支持体を予め選択したモノマー溶液と接触させることを含む。選択された領域は、集積回路製作において使用されるフォトリソグラフィ技術の方法において多い、マスクを通した光供給源を用いる照射によって活性化され得る。支持体のその他の領域は、照明がマスクによって遮断され、それらは化学的に保護されたままであるので、不活性のままである。したがって、光パターンが、支持体のどの領域が、所与のモノマーと反応するかを定義する。異なるセットの予め定義された領域を反復して活性化し、異なるモノマー溶液を支持体と接触させることによって、支持体上でポリマーの多様なアレイが製造される。場合により、反応していないモノマー溶液を支持体から洗淨するなどのその他の段階を使用してもよい。その他の適用可能な方法として、米国特許第 5,384,261 号に記載されるものなどの機械的技術が挙げられる。単一の支持体上での構築物および / または選択オリゴヌクレオチドの合成に適用可能なさらなる方法は、例えば、米国特許第 5,384,261 号に記載されている。例えば、試薬は、(1) 予め定義された領域上の定義されたチャネル内を流れることまたは (2) 予め定義された領域上に「スポットすること」のいずれかによって支持体に送達され得る。その他のアプローチならびにスポットすることおよび流れることの組合せも同様に使用してよい。各場合において、モノマー溶液が種々の反応部位に送達される場合には、支持体の特定の活性化された領域がその他の領域から機械的に分離される。フローチャネル法は、例えば、
40
50

固体支持体上でのオリゴヌクレオチドの合成を制御するためのマイクロ流体システムを含む。例えば、支持体の表面上に、それによって適当な試薬が流れるか、その中に適当な試薬が入れられるフローチャネルを形成することによって、固体支持体の選択された領域で多様なポリマー配列を合成できる。固体支持体でオリゴヌクレオチドを調製するためのスポットティング法は、選択された領域において反応物を直接沈着させることによってそれらを比較的少量で送達することを含む。いくつかの段階では、そうすることがより効率的である場合には、全支持体表面に溶液を噴霧するか、そうでなければそれをコーティングすることもできる。モノマー溶液の正確に測定されたアリコート、領域から領域へ移動するディスペンサーによって液滴を沈着させてもよい。固体支持体上でのオリゴヌクレオチドを合成するためのピンベースの方法は、例えば、米国特許第5,288,514号に記載されている。ピンベースの方法は、複数のピンまたはその他の拡張を有する支持体を利用する。ピンは各々、トレイ中の個々の試薬容器中に同時に挿入される。96ウェルマイクロタイターディッシュなどの96容器トレイを用いる96ピンのアレイがよく使用される。各トレイは、個々のピンでの特定の化学反応におけるカップリングのための特定の試薬を充填される。したがって、トレイは、異なる試薬を含有することが多い。化学反応は、反応の各々が、比較的同様の反応条件のセット下で実施され得るよう最適化されているので、複数の化学的カップリングステップを同時に実施することが可能となる。

【0062】

他の実施形態において、複数の支持体上に、複数のオリゴヌクレオチドを合成または固定化してもよい。一例として、例えば、米国特許第5,770,358号、5,639,603号および5,541,061号に記載されているビーズベースの合成法がある。ビーズ上でのオリゴヌクレオチドなどの分子の合成のために、多量 (a large plurality) のビーズを、容器中で適した担体 (水など) に懸濁する。ビーズは、場合により、保護基と複合体形成される活性部位を有する任意選択のスペーサー分子とともに提供される。合成の各段階で、カップリングのためにビーズを複数の容器中にわけ、新生オリゴヌクレオチド鎖を脱保護した後、各容器に種々のモノマー溶液を加え、その結果、所与の容器中のすべてのビーズで、同一のヌクレオチド付加反応が生じる。次いで、ビーズを過剰の試薬の洗浄を行い、単一容器中でプールし、合成の次のラウンドに備えて、混合し、別の複数の容器の中に再分配する。最初に使用される多数のビーズのおかげで、同様に、容器中に無作為に分散され、各々、塩基の無作為化付加の多数のラウンド後にその表面上で合成される独特のオリゴヌクレオチド配列を有する多数のビーズがあるということは留意されなければならない。個々のビーズに、その上の二本鎖オリゴヌクレオチドに対して独特である配列を用いてタグをつけ、使用の際に同定を可能にすることもできる。

【0063】

その開示内容が、すべての目的のためにその全文が参照により本明細書に組み込まれる以下の参考文献: 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. 21:10; Microarrays: Making Them and Using Them In Microarray Bioinformatics, Cambridge University Press, 2003; 米国特許出願公開第2003/0068633号および2002/0081582号; 米国特許第6,833,450号、6,830,890号、6,824,866号、6,800,439号、6,375,903号および5,700,637号; ならびにPCT公開番号WO04/031399、WO04/031351、WO04/029586、WO03/100012、WO03/066212、WO03/065038、WO03/064699、WO03/064027、WO03/064026、WO03/046223、WO03/040410およびWO02/24597に示され

る、光依存性法、フローチャネルおよびスポッティング法、インクジェット法、ピンをベースとする方法およびビーズをベースとする方法を使用して、予め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を、支持体と結合させるか、またはその場で合成してもよい。いくつかの実施形態において、あらかじめ合成されたオリゴヌクレオチドは、スポッティング法を使用して、支持体と結合されるか、または合成され、ここでは、モノマー溶液が、領域から領域へ移動するディスペンサー（例えば、インクジェット）によって液滴が沈着される。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、力学的な波によって作動されるディスペンサーを使用して支持体上にスポッティングされる。

【0064】

10

増幅

いくつかの実施形態において、オリゴヌクレオチドは、一のプライマーがオリゴヌクレオチドの各々の末端に対応する（例えば、オリゴヌクレオチドの3'末端と相補的なものとオリゴヌクレオチドの5'末端と同一のもの）適切なプライマー対を使用することで増幅され得る。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'ユニバーサル配列または5'共通増幅配列）および3'増幅配列（例えば、3'ユニバーサル配列または5'共通増幅配列）によって隣接された中央（central）または内部のアセンブリ配列（最終産物に組み込まれるように設計された、標的配列に対応する）を含むように設計され得る。

【0065】

20

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'増幅配列に隣接された中央アセンブリ配列を含み得る。中央アセンブリ配列はアセンブリされた核酸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、アセンブリされた核酸に組み込まれることを意図しない。隣接する増幅配列は、同じ増幅配列を共有するが、異なる中央アセンブリ配列を有する多く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。

【0066】

隣接する増幅配列に対応する増幅プライマー（例えば、10から50ヌクレオチド長の間、15から45ヌクレオチド長の間、約25ヌクレオチド長など）は、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、あるプライマーが3'増幅配列に相補的であり得、あるプライマーが、5'増幅配列と同じ配列であり得る）。いくつかの実施形態において、異なる中央アセンブリ配列がある複数の異なるオリゴヌクレオチド（例えば、5、10、50、約100以上）が5'増幅配列と同一および3'増幅配列と同一であり得る。これらのオリゴヌクレオチドをすべて、同じ反応で同じ増幅プライマーを使用することで増幅できる。増幅配列はその後、増幅されたオリゴヌクレオチドから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のいかなる適当な技術を使用することで除去され得る。いくつかの実施形態において、増幅配列は、本明細書で詳細に説明されるように、制限酵素によって除去される。

30

【0067】

いくつかの実施形態において、オリゴヌクレオチドは支持体に付着したままで増幅され得る。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に支持体から除去または切断され得る。

40

【0068】

いくつかの実施形態において、方法は、テンプレートとして第一の複数の一本鎖オリゴヌクレオチドを使用することにより鎖伸長反応で複数のオリゴヌクレオチドまたはポリヌクレオチドを合成することを含む。上述したように、オリゴヌクレオチドは、まず表面の複数の別個の機構上で合成され得るか、または支持体の複数の機構に沈着され得る。いくつかの実施形態において、オリゴヌクレオチドは支持体に共有結合している。いくつかの

50

実施形態において、第一の複数のオリゴヌクレオチドは固体表面に固定化される。いくつかの実施形態において、固体表面の各機構は異なるあらかじめ定義された配列（例えば、1機構あたり約 $10^6 \sim 10^8$ 個の分子）を有する高密度のオリゴヌクレオチドを含む。支持体は少なくとも100、少なくとも1,000、少なくとも 10^4 、少なくとも 10^5 、少なくとも 10^6 、少なくとも 10^7 、少なくとも 10^8 の機構を含み得る。いくつかの実施形態において、二本鎖オリゴヌクレオチドは、増幅の後に、より長い核酸構築物を形成するために溶液中で溶出され、および/または、エラー減少および/またはアセンブリに供され得る。

【0069】

エラー減少

いくつかの実施形態において、各々の断片は、生成された付着末端を処理する前に、アセンブリおよび忠実度最適化され（例えば、本明細書に記載された1つ以上のアセンブリ後の忠実度最適化技術を用いて）、エラーを含むオリゴヌクレオチド分子を除去する。配列エラーは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エラー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1/10から1/200までの誤りに変化し得る。しかしながら、いくつかの実施形態において、より低いエラー率が達成され得る。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。

【0070】

本発明のいくつかの態様は、合成オリゴヌクレオチドが、ポリヌクレオチドをより長いポリヌクレオチド構築物にアセンブリするために設計され、使用される、ポリヌクレオチドアセンブリ方法に関する。酵素的増幅または鎖伸長反応の間、連続したエラーは忠実に模写される。結果として、この方法で合成されたポリヌクレオチド集団は、エラーのない配列およびエラー傾向のある配列の両方を含む。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オリゴヌクレオチド合成の間に導入されたエラーのため不正確な配列を含み得るため、アセンブリまたは伸長の間に1つ以上の誤りを含むオリゴヌクレオチドを組み込んでいるポリヌクレオチドを除去することは、有用であり得る。いくつかの実施形態において、1個以上のアセンブリされたポリヌクレオチドが、それらがあらかじめ定義された配列を含むかどうか決定するためにシーケンシングされ得る。この手順は、正しい配列をもつ断片を同定することを許容する。他の実施形態において、他の技術がエラーを含む核酸断片を除去するために使用され得る。そのような核酸断片は、新生の合成オリゴヌクレオチドまたはアセンブリされた核酸ポリマーであり得る。アセンブリ手順が1つ以上のラウンドのポリメラーゼ伸長（例えば、アセンブリ産物を増幅するためにアセンブリの間またはアセンブリ後に）を伴い得るので、エラーを含む核酸が両方の鎖（すなわち、両方の鎖上の不正確な相補的ヌクレオチド、欠失、または付加）上にエラーを有する二本鎖のホモ二本鎖であり得ることが理解されるべきである。ポリメラーゼ伸長の間に、エラーを含む入力核酸はテンプレートとして機能し得、それにより相補的エラーを含む相補鎖を生成する。ある実施形態において、二本鎖核酸断片または二本鎖の調製は、アセンブリの間に組み込まれる1つ以上の配列エラーを含む核酸ならびに正確なあらかじめ定義された配列を有する核酸の混合物を含むことが疑われ得る。「二本鎖」なる用語は、少なくとも部分的に二本鎖である核酸分子をいう。「安定した二本鎖」は、特定のハイブリダイゼーション条件のセットのもとで相補的配列にハイブリダイズされたままで比較的残存しそうな二本鎖をいう。例示的な実施形態において、安定した二本鎖は、塩基対ミスマッチ、挿入、または欠失を含まない二本鎖をいう。「不安定な二本鎖」は、ストリンジェントな溶融など特定のハイブリダイゼーション条件のセットのもとで相補的配列にハイブリダイズされたままで比較的残存しそうにない二本鎖をいう。例示的な実施形態において、不

10

20

30

40

50

安定な二本鎖は少なくとも1つの塩基対ミスマッチ、挿入、または欠失を含む二本鎖をいう。本明細書で使用される、「ストリンジェンシー」なる用語は、核酸ハイブリダイゼーションが行われる、温度、イオン強度、および有機溶媒などの他の化合物の存在などの条件に関して使用される。ハイブリダイゼーションのストリンジェンシーは、ハイブリダイゼーション工程の間、温度、および/または、ハイブリダイゼーション溶液中の塩および/またはホルムアミド濃度などの溶液化学特性などにより増加する。「高いストリンジェンシー」条件で、核酸塩基対形成は相補的塩基配列の高い頻度を有する核酸断片の間だけに起こるのである。ストリンジェント条件として、定められたイオン強度およびpHで特定のポリヌクレオチド二本鎖が熱融点(T_m)より約5 低い条件が選択され得る。相補的なポリヌクレオチド鎖の長さおよびGC含量は二本鎖の T_m を決定し、その結果所望の特異的なハイブリダイゼーションを得るために必要なハイブリダイゼーション条件を決定する。 T_m はポリヌクレオチド配列の50%が完全に一致する相補鎖にハイブリダイズされる温度(定められたイオン強度およびpH下)である。ある場合には、ハイブリダイゼーション条件のストリンジェンシーが特定の二本鎖に対して T_m とほぼ等しくなるまで増加することが望ましい場合がある。適切なストリンジェンシー条件は、当業者に知られているか、または当業者が実験的に決定し得る。例えば、Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-12.3.6; Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; S. Agrawal (ed.) Methods in Molecular Biology, volume 20; Tijssen (1993) Laboratory Techniques in biochemistry and molecular biology - hybridization with nucleic acid probes, e.g., part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New Yorkを参照。

【0071】

いくつかの実施形態において、配列エラーは二本鎖核酸変性させ、再アニーリングすることを含む技術を使用することにより除去され得る。いくつかの実施形態において、相補的なエラーを含む核酸の一本鎖は、核酸調製において、各々が個々のエラーを含む核酸が、同じ位置で正しい配列を有する核酸より低頻度で存在しているならば、共に再アニーリングしそうもないであろう。むしろ、エラーを含む一本鎖はエラーのない相補鎖または1つ以上の異なるエラーまたは他の位置にエラーを有する相補鎖と再アニーリングするであろう。結果として、エラーを含む鎖は、最終的に再アニーリング反応産物中のヘテロ二本鎖分子の形となるであろう。エラーのない核酸鎖はエラーを含む鎖または他のエラーのない鎖と再アニーリングし得る。再アニーリングされたエラーのない鎖は、再アニーリングされた試料中でホモ二本鎖を形成する。従って、核酸断片の再アニーリングされた調製物からヘテロ二本鎖分子を除去することにより、エラーを含む核酸の量または誤り率を、減少できる。

【0072】

ヘテロ二本鎖形成はそれゆえ、異なる集団からの核酸鎖が完全な一致およびミスマッチ含有二本鎖が形成できるように他方とハイブリダイズされる、シャッフリングとして理解される工程を経て起こる。ヘテロ二本鎖を除去するための適切な方法は、クロマトグラフィー、電気泳動、二本の鎖の間にミスマッチ配列を有する二本鎖DNAと優先的に結合するヘテロ二本鎖分子との選択的な結合を含む。「ミスマッチ」または「塩基対のミスマッチ」なる用語は、ワトソンおよびクリックの塩基対形成ルールによる核酸を通常形成しない塩基対の組み合わせを示す。例えば、DNAで一般的に見出される塩基、すなわち、アデニン、グアニン、シトシンおよびチミジン、についていうと、塩基対のミスマッチは通

常DNAで見出されるA - TおよびG - Cの組み合わせ以外のそれらの塩基の組み合わせである。本明細書に記載されるように、ミスマッチは、例えばシトシン残基が適切なパートナーであるグアニンとは対照的に、別のシトシンの反対側で見出されることを意味するC / C、として示される。

【0073】

いくつかの実施形態において、オリゴヌクレオチド調製物は、本明細書にさらに詳細に説明されるように、エラーを含む分子を除去するために選択されるか、またはスクリーニングされ得る。いくつかの実施形態において、オリゴヌクレオチドは、本明細書に記載されるように、ミスマッチ結合剤を使用することでエラー修正され得る。

【0074】

一の態様において、本発明は、固体支持体上で高忠実度ポリヌクレオチドを製造する方法に関する。合成ポリヌクレオチドは、少なくとも約1、2、3、4、5、8、10、15、20、25、30、40、50、75または100キロベース(kb)または1メガベース(mb)であるか、またはそれより長い。例示的な実施形態において、合成ポリヌクレオチドの組成物は、エラーのない、少なくとも約1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、50%、60%、70%、80%、90%、95%またはそれ以上のコピーを含有する(例えば、あらかじめ定義された配列から逸脱しない配列を有する)。エラーのないコピーのパーセントは、正しい、例えば、あらかじめ定義されたか、または所定の配列を有するよう意図された組成物中のポリヌクレオチドのコピーの総数と比較した、組成物中のエラーのないコピーの数に基づく。

【0075】

本発明のいくつかの態様は、高忠実度ポリヌクレオチドアセンブリのためのオリゴヌクレオチドの設計に関する。本発明の態様は、核酸アセンブリ手順の処理能力速度を高める、および/または正しくアセンブリされた核酸配列を作製するために使用されるステップ数または試薬の量を低減するために有用であり得る。ある実施形態において、本発明の態様は、自動化された核酸アセンブリの文脈において、各々正しい核酸配列のアセンブリに必要なステップの時間、数、試薬の量およびその他の因子を低減するために有用であり得る。したがって、これらおよびその他の本発明の態様は、1つ以上の核酸アセンブリ手順の費用および時間を低減するために有用であり得る。

【0076】

一本鎖突出

本発明のいくつかの態様において、アセンブリされる核酸断片は、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リング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アセンブリされ、アセンブリさ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、ベクターの第一の付着末端と相補的な第一の付着末端および第二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、第三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、第四の核酸の第一の付着末端と相補的であり得る。最初から二番目の核酸の第二の付着末端に相補的であり得る第一の付着末端を有する最終的な核酸を経るなど。

【0077】

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)のアセンブリを十分に促進できる程度(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。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アセンブリされるとき、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。

【0078】

いくつかの実施形態において、異なる核酸断片間の2つ以上の相補的付着末端の対は、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。このことは、異なる配列アレンジメントをもつ核酸産物のライブラリ、および/または、異なるコピー数の、ある一定の内部配列領域を発生させることに役立ち得る。

【0079】

いくつかの実施形態において、最終的な核酸の第二の付着末端ベクターの第二の付着末端と相補的であり得る。本発明の態様によると、この方法は、あらかじめ定義された線形順序（例えば、第一、第二、第三、第四、、最終）でアセンブリされた核酸断片を含むベクターを生成させるために使用され得る。いくつかの実施形態において、各々2個の終端核酸断片（例えば、アセンブリ産物の各端の末端断片）が、ベクター（例えば、直鎖ベクター）の付着末端と相補的な付着末端を有するように設計され得る。これらの付着末端は直鎖ベクターの同じ相補的末端配列にアニーリングできる同じ付着末端であり得る。しかしながら、いくつかの実施形態において、末端断片上の付着末端は異なっており、ベクターは、一方が直鎖ベクターの各々の末端であり、各々末端断片付着末端の一つと相補的である、異なる2つの付着末端を含む。従って、ベクターは、それらの各々がアセンブリされた核酸断片の一端と相補的である2つの付着末端を有する線状化プラスミドであり得る。

【0080】

本発明のある態様は、一本鎖突出をもつ二本鎖核酸に関する。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アセンブリされた断片）、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、アセンブリ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に消化され得る。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、認識部位の外で単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 ~ 7bp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 - 5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、BstF5I、BtsCI、BsrDI、BtsI、AlwI、BccI、BsmAI、EarI、MlyI（平滑）、PleI、BmrI、BsaI、BsmBI、FauI、MnlI、SapI、BbsI、BciVI、HphI、MboII、BfuAI、BspCNI、BspMI、SfaNI、HgaI、BseRI、BbvI、EciI、FokI、BceAI、BsmFI、BtgZI、BpuEI、BsgI、MmeI、BseGI、Bse3DI、BseMI、AclWI、Alw26I、Bst6I、BstMAI、Eam1104I、Ksp632I、P

10

20

30

40

50

p s I、S c h I (平滑)、B f i I、B s o 3 1 I、B s p T N I、E c o 3 1 I、E s p 3 I、S m u 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 l I、E c o 5 7 I、E c o 5 7 M I、G s u I、およびB e g Iが挙げられるが、これらに限定されるものではない。そのような酵素およびそれらの認識および切断部位に関する情報は、N e w E n g l a n d B i o l a b s (I p s w i c h , M a s s . , U . S . A .)などの供給業者から利用可能である。

【0081】

いくつかの実施形態において、商業的な、または、操作された制限酵素が使用され得る。いくつかの実施形態において、より長い突出長を生成するためにI I S型制限酵素を設計し、操作できる。より長い一本鎖突出を生成するための制限酵素の設計および操作は、多数のオリゴヌクレオチドの接合が、より長い核酸構築物を共に形成することを許容できる。例えば、4個のヌクレオチドの一本鎖突出を生成するB s a Iは、5または6以上の一本鎖突出を生成するために操作できる。そのような操作されたB s a Iによって生成された一本鎖突出長の増加は、接合できる、17個の核酸またはオリゴヌクレオチドの理論的限界を増加させることができる。

【0082】

いくつかの実施形態において、核酸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、および/または適切なベクターを直鎖化するために用いられ得る。しかしながら、異なるI I S型部位も使用され得る。それぞれアセンブリさ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、アセンブリさ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。いくつかの実施形態において、2つの末端核酸断片の各々の一端は、直鎖化ベクター核酸の一本鎖突出に相補的な一本鎖突出（例えば、適切な制限酵素での消化の後）を有するように設計され得る。したがって、得られる核酸断片およびベクターは、宿主細胞に直接的に形質転換され得る。あるいは、核酸断片およびベクターは、宿主細胞の形質転換への利用に先立って、相補的配列のハイブリダイゼーションおよびアニーリングを促進するためにインキュベートされ得る。ベクターが本明細書に記載された技術のいずれか一つまたは末端核酸断片の一つの端に相補的であろう一本鎖突出を生成するいかなる他の適切な技術を使用して作製され得ることが理解されるべきである。

【0083】

I I型または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から6ヌクレオチドの突出を生成する。これらの短い付着末端は相補的な末端を含む2個の核酸断片をライゲーションするために十分であり得る。しかしながら、複数の核酸断片を結合させるとき、より長い相補的な付着末端は、アセンブリを容易にし、特異性を確実にするために好ましいであろう。例えば、付着末端は、類似した付着末端間の誤対合を防ぐか、または減少させるために十分な異なる配列を有する程度の長さであり得る。しかしながら、それらの長さは類似の付着配列間の誤対合を安定できる程度に長くはないことが好ましい。いくつかの実施形態において、約9から約15個の塩基長が使用され得る。しかしながら、いかなる適当な長さが、付着末端を生成するために使用される領域に選択され得る。特異性の重要度は、同時にアセンブリされるであろう異なる断片の数に依存し得る。また、誤対合領域の安定化を避けるために必要な適切な長さは、異なる付着末端をアニーリングする

10

20

30

40

50

ために使用される条件に依存し得る。

【0084】

リガーゼベースのアセンブリ

リガーゼベースのアセンブリ技術は、隣接する3'および5'核酸末端（例えば、3'末端が、5'末端と直接隣接するよう、相補的鋳型核酸上でアニーリングされた核酸（単数または複数）の5'リン酸および3'ヒドロキシル）の共有結合を触媒できる1種または複数の適したリガーゼ酵素を含み得る。したがって、リガーゼは、第一および第二の核酸が鋳型核酸上で隣同士がアニーリングされる場合に、第一の核酸の5'リン酸と第二の核酸の3'ヒドロキシル間のライゲーション反応を触媒し得る。リガーゼは、組換えまたは天然供給源から得ることができる。リガーゼは、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、好熱性生物から得られた熱安定性リガーゼが使用され得る。熱安定性DNAリガーゼの例として、限定されるわけではないが：Tth DNAリガーゼ（例えば、EurogentecおよびGeneCRAFTから入手可能であるサーマス・サーモフィラス（*Thermus thermophilus*）由来の）；Pfu DNAリガーゼ（パイロコッカス・フリオサス（*Pyrococcus furiosus*）由来の超好熱性リガーゼ）；Taq リガーゼ（サーマス・アクアティカス（*Thermus aquaticus*）由来の）、任意のその他の適した熱安定性リガーゼまたはそれらの任意の組合せが挙げられる。いくつかの実施形態において、1種以上の低温リガーゼが使用され得る（例えば、T4 DNAリガーゼ）。低温リガーゼは、高温では安定でない可能性がある、短い突出（例えば、約3、約4、約5または約6塩基の突出）にあって有用であり得る。

10

20

【0085】

いくつかの実施形態において、リガーゼは、形成された所望ではないライゲーション産物を最小化するために、より大きい特異性を有するように設計され、操作され得る。いくつかの実施形態において、リガーゼは、タンパク質と結合して使用され得るか、または核酸分子とリガーゼの相互作用を容易にすること、および/またはライゲーションの特異性を増加させること、ができるタンパク質と融合され得る。

【0086】

非酵素技術を使用して、核酸をライゲーションすることもできる。例えば、酵素を使用することなく（例えば、リガーゼを使用せず）、1種以上の核酸の5'末端（例えば、5'リン酸基）および3'末端（例えば、3'ヒドロキシル）を一緒に共有結合し得る。いくつかの実施形態において、非酵素技術は、酵素ベースのライゲーションを上回る特定の利点を提供し得る。例えば、非酵素技術は、核酸基質中の非天然ヌクレオチド類似体の高い耐性を有し得、短い核酸基質をライゲーションするのに使用でき、RNA基質をライゲーションするのに使用でき、および/またはより安価なおよび/または特定の自動化（例えば、高処理）適用により適したものであり得る。したがって、化学的ライゲーションを使用して、第一の核酸末端の5'末端と第二の核酸末端の3'末端間の共有結合を形成してもよく、ここで、第一および第二の核酸末端は、単一の核酸の末端または別々の核酸の末端であり得る。一態様では、化学的ライゲーションは、連結形成を容易にするか、または促進する1種以上の化学的反応部分を含む、修飾された末端（例えば、修飾された5'および/または3'末端）を有する少なくとも1種の核酸基質を含み得る。いくつかの実施形態において、化学的ライゲーションは、1種以上の核酸末端が密接して一緒になる場合に（例えば、末端が、相補的核酸配列間のアニーリングのために一緒になる場合に）起こる。したがって、相補的3'または5'突出（例えば、二本鎖核酸の制限酵素切断によって生じた突出）間または3'末端が5'末端と密接に接近される相補的核酸の任意の組合せ間のアニーリング（例えば、核酸が相補的鋳型核酸とアニーリングされると、3'および5'末端が互いに隣接する）は、鋳型による（*template-directed*）化学的ライゲーションを促進し得る。化学反応の例は、これらに限定されないが、縮合、還元および/または光化学的ライゲーション反応を含み得る。当然のことではあるが、いくつかの実施形態において、化学的ライゲーションを使用して、天然に存在するホスホ

30

40

50

ジエステルヌクレオチド間連結、天然に存在しないホスファミドピロホスフェートヌクレオチド間連結および/またはその他の天然に存在しないヌクレオチド間連結を生成することができる。

【0087】

同時の、共通オリゴヌクレオチド配列の酵素的除去および処理オリゴヌクレオチドのより長い構造物へのライゲーション

図2は本発明の一の実施形態により核酸をアセンブリするための方法を示す。いくつかの実施形態において、方法は同時の、共通オリゴヌクレオチド配列の酵素的除去および処理オリゴヌクレオチドのより長い構造物へのライゲーションを含む。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、オリゴヌクレオチドはPCRで増幅され、エラー修正される。共通プライマー（増幅）配列（20）、および構築物特異的ペイロードまたは内部配列領域（30）で構成される増幅オリゴヌクレオチド（10）は、適切な制限酵素（40）によって処理される。いくつかの実施形態において、第一および最終のオリゴヌクレオチドは標的構築物の増幅のためのユニークなプライマー配列（25）を含む。制限エンドヌクレアーゼは、すべてのオリゴヌクレオチド（50）により共有された末端共通領域（本明細書において増幅領域またはプライマー認識配列としても参照される）の切断を、末端一本鎖DNA配列（60）と内部領域（本明細書において遊離ペイロードとしても参照される）を残して促進する。いくつかの実施形態において、制限酵素はIIS型制限酵素である。これらの一本鎖配列は、多数のオリゴヌクレオチドのあらかじめ定義された配列（70）への直線配置を許容する、一のオリゴヌクレオチドの他のものとの特異的な相互作用を命令するように設計される。したがって、末端一本鎖DNA配列はオリゴヌクレオチドの適切な相互作用を正しい順番に向けることができ、それによりリガーゼ酵素（80）は、最終的な標的核酸構築物（90）または中間核酸構築物を生成させる、個々のオリゴヌクレオチドの接合を触媒する。

【0088】

当業者は、元の共通配列が共にライゲーションして戻されると（例えば、（60）に相補的な末端配列を用いる（50））、制限エンドヌクレアーゼの存在は、それが再度切断されて自由末端（60）を発生させ得ることを確実にし得る。しかしながら、制限エンドヌクレアーゼの選択のために、適切にライゲーションされた接合点（例えば、1'および2'の間）は、制限部位として認識されず、元に戻されないであろう。反応は自然に所望の産物（90）に向かうように駆動されるべきである

【0089】

いくつかの実施形態において、処理の変形は、共通配列除去に使用される制限部位が、現在は合成される遺伝子の一部であり得ることを認識する。この制限の除去は、遺伝子合成法の再帰的な（階層的な）適用が、より長い核酸配列（図4で示されるような）を構築することを許容する。除去とライゲーションが別々の段階で実施される以前の方法論では、この設計は、部分的にサイズ選択に基づく、除去とライゲーションの間の精製段階の必要性のために許容されなかった。そのような方法論では、所望の標的配列の切断断片は、精製の間に失われる場合があり、所望の標的配列を構築できない結果となった。いくつかの実施形態において、本発明の同時の除去およびライゲーション段階を使用して、それらの切断配列は、連続的に切断および再ライゲーションされ、いくつかの対象となる標的配列の存在をもたらすであろう。所望の配列の量は、いくつかの実施形態において、制限酵素およびリガーゼの相対的な活性の調整に依存され得る。

【0090】

図4で示されるように、オリゴヌクレオチドセット（310）および（311）から遺伝子合成断片（390）および（391）をアセンブリできる。オリゴヌクレオチドセットは、同じ同時の消化およびライゲーション工程（その後の増幅を伴う）を使用することで遺伝子合成断片（390）および（391）を接合できるように、適合する制限酵素部位（340）で設計され得る。いくつかの実施形態において、2番目のラウンドは、第二の制限酵素を使用する、制限酵素部位（340）で設計され得る。しかしながら、これは

処理において複数の酵素を使用する複雑さのため望ましくない場合がある。加えて、同時の消化およびライゲーションがなければ、さらに、合成できる遺伝子を抑制して、2つの制限酵素の使用は標的配列からの2つの制限酵素部位を許容しない結果を生じるであろう。

【0091】

また図4を参照すると、核酸断片(390)はプライマー(325)を使用することで増幅でき、核酸断片(391)はプライマー(326)を使用することで増幅できる。核酸断片はその後、結合した核酸断片(392)を作成するために共に混合され、以前の合成段階と同様に処理され得て、ここで、制限部位(340)は以前のラウンドにおける部位(350)と同様の手順で働く。結合した標的配列(392)は、(325)からの5'プライマーおよび(326)からの3'プライマーを使用することで増幅できる。

10

【0092】

いくつかの実施形態において、本明細書に記載された方法により階層的アセンブリが使用され得る。当業者は、本発明の方法が、最初のラウンドにおける核酸断片の数と、その後のラウンドにおける核酸断片の数が同一にできるように、複数の核酸断片に拡張可能であることを理解するであろう。階層的な組立法は、非常に大きい標的が比較的少数のラウンドで構築されることを許容する、幾何学的であり得る。例えば、1000ベース(1kbp)の標的配列が、プール(310)または(311)の1つから構築できる。(390)または(391)と同様の10個の核酸断片の第二のラウンドは10kbp塩基標的核酸配列をもたらすであろう。10kbp核酸配列を用いる第三のラウンドは、元の1000の源のプール由来の100kbp標的核酸配列をもたらすであろう。

20

【0093】

いくつかの実施形態において、複数のアセンブリ反応が別々のプールで行われ得る。アセンブリ反応からのアセンブリ構築物はさらに長い核酸配列を形成するためにともに混合され得る。いくつかの実施形態において、階層的なアセンブリが、所望の順序で結合し得る付着末端を形成するために制限酵素を使用して実施され得る。構築オリゴヌクレオチドは、特定の順序で結合することを促進する部位での1つ以上の制限酵素のための認識および切断部位を含むように設計および合成され得る。いくつかの実施形態において、1つ以上のIIS型エンドヌクレアーゼ認識部位が、IIS型制限酵素での切断を可能にするために、構築オリゴヌクレオチドの末端に組み込まれ得る。結合の順序は相補的な付着末端のハイブリダイゼーションで決定できる。

30

【0094】

いくつかの実施形態において、オリゴヌクレオチドの第一のプールは3'末端に追加の制限酵素認識部位を有するように設計された3'末端オリゴヌクレオチドを含み、オリゴヌクレオチドの第二のプールは5'末端に追加の制限酵素認識部位を有するように設計された5'末端オリゴヌクレオチドを含む。いくつかの実施形態において、制限酵素は同一である。各々のプールにおけるオリゴヌクレオチドのアセンブリの後に、2つのアセンブリ構築物は本明細書に記載の方法により、制限酵素およびリガーゼに供することができる。

【0095】

40

当業者は、合成の利用可能なアセンブリ空間が本発明の態様により、(幾何学上)抜本的に改良されることを理解するであろう。以前は、2倍の配列サイズ(2n)、2倍の数のオリゴヌクレオチドの構築物を生成することが必要であった。例えば、構築物(390)を生成させるために、2倍の数のオリゴヌクレオチド(310)が必要であり、その結果、2倍の数の、互換性のある一本鎖末端(360)必要であった。図4に示された方法を使用すると、(310)および(311)のための接合点は接合点(340)との互換性が必要であるだけであり、その結果、ただ1つの追加の接合点を使用するだけで、2倍のサイズの核酸のアセンブリを可能にする。したがって、オリゴヌクレオチド(310)および(311)に干渉性または非互換性末端がある場合、それらは本明細書に記載された方法(消化(340)およびライゲーション)で、標的核酸を作成するために結合され

50

得るが、単にオリゴヌクレオチドプール(310)および(311)と一緒に混合することのみによっては可能とならないであろう。

【0096】

オリゴヌクレオチドの同時処理および標的構築物へのアセンブリおよびプラスミドへの同時の挿入の変形例は図3に示される。プラスミド、pG9-1(配列番号:1)の詳細は図5に示される。プラスミドは、制限酵素(この場合BsaI)が2つの位置でプラスミドを切断し、定められた一本鎖配列(図5、反転テキスト)を残すことを許容する制限酵素認識部位(下線テキスト、図5)を含む。図3を参照すると、プラスミド(100)(例えばpG9-1)は本明細書に記載されるようにあらかじめ増幅およびエラー修正されたオリゴヌクレオチドの混合物(110)を含むプールに導入される。いくつかの実施形態において、これらのオリゴヌクレオチド配列(110)は特定の制限エンドヌクレアーゼ(140)によって認識される共通配列(120)を有し得る。いくつかの実施形態において、プラスミド(130)は、同じ制限エンドヌクレアーゼ(140)により認識される配列を有し得る。これらの配列での制限エンドヌクレアーゼ(140)の作用は、オリゴヌクレオチド((310)、(311))、およびプラスミド(150)からの共通配列の除去をもたらし、一本鎖DNA配列(160)を露出する。いくつかの実施形態において、制限酵素はIIS型制限酵素であり得る。いくつかの実施形態において、一本鎖配列は、多数のオリゴヌクレオチドを定められた配列に配置し、この順序付けされたオリゴヌクレオチド配列(170)をプラスミド(100)に導入することを許容する、1つのオリゴヌクレオチドと他のものとの特異的相互作用を行うように設計される。いくつかの実施形態において、リガーゼ(180)酵素は個々のオリゴヌクレオチドの共有結合を触媒する。最終産物は、オリゴヌクレオチド(190)の結合に由来する特定された構築物を含むプラスミド(例えば、pG9-1)である。このプラスミド(190)は、その後細菌の形質転換に用いられ、配列照合され得る。

【0097】

本発明の態様は、本発明の方法によりアセンブリされた構築物の配列照合に関する。構築物の配列照合は図6に示される。この方法では、多くの構築物(200、C1からC4)は、図3に示されるように生成され、細菌の形質転換に用いられる。プラスミドDNAを含む形質転換細胞は、選択に適切な抗生物質抵抗性を使用することで、固相増殖プレート上(210)で選択できる。増殖の後に、それぞれの構築物プレート(220)からの1つの、単一コロニーが選択およびプールされ、各々のプールが各々の構築物の1つのコピーを含む、構築物のプールを生成する。いくつかの実施形態において、プールの数は完全な配列をもつ構築物を同定するためにシークエンシングされた個々の構築物の数に依存し得る。図6に示されるように、4つの構築物の4つのプールが生成され、各々4つの構築物の分析を許容する。プラスミドDNAはその後、プールされた材料(230)から調製できる。プラスミドDNA分子の各々のプールはその後、シークエンシングのために調製できる。この調製は、DNAのより小さい断片への切断、および例えば次世代高処理シークエンシングを用いる、シークエンシングに必要な共通配列の付着をもたらす。生成された4つのプールの各々に特有のDNAの小片は、これらの共通配列の中に含まれている。これらの特有のDNAは、各々シークエンシングされた構築物がどのプールに由来するかを識別することを許容する。正しい配列の構築物は、その後初期細菌増殖プレートに戻し、所望の構築物をもつプラスミドを含む対応するコロニーを再増殖することにより回収できる。

【0098】

ベクターおよび宿主細胞

本発明がそのように限定されないように、いかなる適当なベクターが使用され得る。例えば、ベクターはプラスミド、細菌ベクター、ウイルスベクター、ファージベクター、昆虫ベクター、酵母ベクター、哺乳類ベクター、BAC、YAC、またはいかなる他の適当な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、ただ1つの型の生物の中だけで複製されるベクター(例えば、細菌、酵母、昆虫、哺乳類など)、または、た

だ一つの種だけで複製されるベクターであり得る。いくつかのベクターは、広い宿主範囲を有し得る。いくつかのベクターは、異なる生物中で機能的である、異なる機能的配列（例えば、起源または複製、選択可能なマーカーなど）を有し得る。これらは、2つの異なる型の生物の間（例えば、細菌と哺乳類、酵母と哺乳類などの間など）でベクター（および、ベクターにクローニングされるいかなる核酸断片）を往復させるために使用され得る。いくつかの実施形態において、使用されたベクターの型は、選択される宿主細胞の型によって決定され得る。

【0099】

ベクターが、形質転換細胞が選択的に増殖でき、ベクターが単離でき、いかなる挿入が、それが所望のアセンブリされた核酸を含むかどうか決定するために特徴づけられるように、選択可能なマーカー（例えば、抗生物質耐性など）などの検出可能なマーカーをコードし得ることが理解されるべきである。挿入は、いかなる適当な技術（例えば、サイズ分析、制限断片分析、シーケンシングなど）を使用することで特徴付けられ得る。いくつかの実施形態において、正しくアセンブリされた核酸によってコードされると予測された機能が、宿主細胞の中で発現されるかどうか測定することによって、ベクター中の正しくアセンブリされた核酸の存在がアッセイされ得る。

【0100】

いくつかの実施形態において、核酸挿入を含むベクターが導入された宿主細胞は、正しい（例えば、設計された）末端の核酸断片がベクターにクローニングされたかどうかについての機能的な1つ以上の追加の検出可能なまたは選択可能なマーカーを使用することにより、選択または濃縮され得る。

【0101】

従って、宿主細胞は、ベクターにコードされた1つ以上の薬物耐性マーカーの選択を許容するため（またはベクターにコードされた1つ以上の検出可能なマーカーの検出を許容するため）、適切な表現型を有するべきである。しかしながら、いかなる適当な宿主細胞が使用され得る（例えば、原核、真核、細菌、酵母、昆虫、哺乳類など）。例えば、宿主細胞は細菌細胞（例えば、大腸菌（*Escherichia coli*）、枯草菌（*Bacillus subtilis*）、マイコバクテリウム属（*Mycobacterium* spp）、マイコバクテリウム・ツベルクローシス（*M. tuberculosis*）、または他の適当な細菌細胞）、酵母（例えば、サッカロマイセス属（*Saccharomyces* spp）、ピチア属（*Pichia* spp）、カンディダ属（*Candida* spp）または他の適当な酵母、例えば、サッカロマイセス・セレビシエ（*S. cerevisiae*）、カンディダ・アルビカンス（*C. albicans*）、サッカロマイセス・ポンベ（*S. pombe*）など）、アフリカツメガエル細胞、マウス細胞、サル細胞、ヒト細胞、昆虫細胞（例えば、SF9細胞およびショウジョウバエ（*Drosophila*）細胞）、虫細胞（例えば、カエノラブリディティス属（*Caenorhabditis* spp）、植物細胞、または他の適当な細胞、また例えば、トランスジェニックまたは他の組換え細胞株、であり得る。加えて、チャイニーズハムスター卵巣細胞（CHO）など、多くの異種細胞株が使用され得る。

【0102】

増幅

本発明の態様は、合成核酸の製造および/または使用を含むさまざまな適用にとって有用であり得る。本明細書に記載されるように、本発明は、増加した効率で合成核酸をアセンブリするための方法を提供する。得られたアセンブリされた核酸は、インビトロ（*in vitro*）で（例えば、PCR、LCRまたは任意の適した増幅技術を使用して）増幅され、インビボ（*in vivo*）で増幅され（例えば、適したベクターへのクローニングによって）、単離および/または精製され得る。アセンブリされた核酸（単独またはベクター中にクローニングされた）で、宿主細胞（例えば、原核生物の、真核生物の、昆虫の、哺乳動物の細胞またはその他の宿主細胞）を形質転換してもよい。いくつかの実施形態において、宿主細胞を使用して、核酸を増殖できる。ある実施形態において、核酸は

、宿主細胞のゲノムに組み込まれ得る。いくつかの実施形態において、核酸は、細胞のゲノム上の対応する核酸領域を置換し得る（例えば、相同組換えによって）。したがって、核酸は組換え生物を作製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または一部を置換するために使用される、全ゲノムまたはゲノムの大きな断片であり得る。組換え生物はまた、種々の研究、工業的、農業的および/または医薬的適用のために使用できる。

【0103】

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1つ以上の伸長ベースおよび/またはライゲーションベースのアセンブリ技術の組み合わせを適用することにより、共に使用できる。例えば、協調したアセンブリは、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アセンブリす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、アセンブリに使用され得る。

【0104】

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。ベクターは、複製起点および1つ以上の選択可能なマーカー（例えば、抗生物質耐性マーカー、栄養要求性マーカーなど）を含み得る。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。

【0105】

本明細書に記載された構築物の転写および/または翻訳がインビトロ（すなわち、無細胞系を使用する）またはインビボ（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。

【0106】

本明細書において提供された方法およびデバイスの態様は、本明細書に記載された1つまたは複数の動き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つまたは複数のステップが、1種または複数の自動化サンプル取り扱いデバイス（例えば、1種または複数の自動化液体または流体取り扱いデバイス）を使用して自動化され得る。自動化デバイスおよび手順は、以下：出発核酸、バッファー、酵素（例えば、1種または複数のリガーゼおよび/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任意のその他の適した薬剤、のうち1種または複数を含めた反応試薬を送達するために使用され得る。自動化デバイスおよび手順はまた、反応条件を制御するために使用され得る。例えば、反応温度および使用され得る任意の温度周期を制御するために自動化サーマルサイクラーを使用し得る。いくつかの実施形態において、1つまたは複数の反応温度またはポリヌクレオチドをインキュベートするのに適した温度周期を提供するよう、走査レーザーが自動化され得る。同様に、アセンブリされたポリヌクレオチド産物のその後の分析も自動化され得る。例えば、シーケンシングは、シーケンシングデバイスおよび自動化シーケンシングプロトコルを使用して自動化され得る。追加のステップ（例えば、増幅、クローニングなど）も、1種以上の適当なデバイスお

よび関連プロトコルを使用して自動化され得る。当然のことではあるが、本明細書に記載された1種以上のデバイスまたはデバイス構成成分は、系（例えば、ロボットシステム）において、または微小環境（例えば、微小流体反応チャンバー）において組み合わせられ得る。アセンブリ反応混合物（例えば、液体反応サンプル）は、自動化デバイスおよび手順（例えば、自動化ピペット操作デバイス、マイクロシステムなどを含めた、サンプルおよび/またはサンプル容器のロボット操作および/または移動）を使用して、系のある構成成分から別のものに移動され得る。システムおよびその任意の構成成分は、制御系によって制御できる。

【0107】

したがって、本明細書において提供されたデバイスの方法ステップおよび/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）を使用して自動化され得る。本明細書において提供された技術の態様が実施され得るコンピュータシステムは、任意の種類の処理（例えば、本明細書に記載されたような配列分析および/または自動化デバイス制御）のためのコンピュータを含み得る。しかし、当然のことではあるが、特定の処理ステップは、アセンブリシステムの一部である1つ以上の自動化デバイスによって提供され得る。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つ以上のコンピュータを含み得る。例えば、1つのコンピュータを、ネットワークによって第二のコンピュータとつなぐことができる。1つのコンピュータは、配列分析を実施できる。第二のコンピュータは、系において、1つ以上の自動化合成およびアセンブリデバイスを制御できる。他の態様において、1つ以上の分析または処理動作を制御するために、さらなるコンピュータがネットワーク中に含まれ得る。各コンピュータは、メモリおよび処理装置を含み得る。本明細書において提供された技術の態様が任意の特定のコンピュータプラットフォームで実施されているよう制限されないで、コンピュータは任意の形態をとり得る。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）を含めた任意の形態をとり得る。ディスプレイデバイスが、1種以上のデバイスおよびコンピュータと関連していてもよい。あるいは、またはさらに、ディスプレイデバイスは、本明細書において提供された技術に従って、分析のアウトプットをディスプレイするために、遠隔部位に配置され、接続され得る。システムの異なる構成成分間の接続は、ワイヤ、光ファイバー、ワイヤレス伝送、衛星伝送、任意のその他の適した伝送または上記のうち2種以上の任意の組合せによってであり得る。

【0108】

本明細書において提供された技術の異なる態様、実施形態または動作の各々は、多数の方法のいずれかで、独立に自動化され、実施され得る。例えば、各態様、実施形態または動作は、ハードウェア、ソフトウェアまたはそれらの組合せを使用して、独立に実施され得る。ソフトウェアで実施される場合には、ソフトウェアコードは、単一コンピュータにおいて提供されようと、複数のコンピュータ間に分配されようと、任意の適した処理装置または処理装置の収集物で実行され得る。当然のことではあるが、上記の機能を実施する、任意の構成成分または構成成分の収集物は、一般に、上記で論じられた機能を制御する1種以上の制御装置と考えられ得る。1種以上の制御装置は、専用のハードウェアを用いて、またはマイクロコードを使用してプログラムされている汎用ハードウェア（例えば、1種または複数の処理装置）または上記で列挙された機能を実施するためのソフトウェアを用いてなど、多数の方法で実施され得る。

【0109】

この点において、当然のことではあるが、本明細書において提供された技術の実施形態の一実施は、コンピュータプログラム（すなわち、複数の指示）で符号化された、少なくとも1種のコンピュータによって読み取り可能な媒体（例えば、コンピュータメモリ、フロッピーディスク、コンパクトディスク、テープなど）を含み、これは、処理装置で実行されると、本明細書において提供された技術の1種以上の上記で論じられた機能を実施する。コンピュータによって読み取り可能な媒体は、輸送可能であり得、その結果、それに保存されたプログラムは、本明細書において提供された技術の1種または複数の機能を実

10

20

30

40

50

施するために任意のコンピュータシステムリソースに読み込まれ得る。さらに、当然のことではあるが、実行されると、上記で論じられた機能を実施するコンピュータプログラムへの言及は、ホストコンピュータ上で動くアプリケーションプログラムに制限されない。むしろ、用語コンピュータプログラムは、本明細書において、一般的な意味で、本明細書において提供された技術の上記で論じられた態様を実施するよう処理装置をプログラムするために使用され得る、任意の種類のコンピュータコードに言及するよう使用される（例えば、ソフトウェアまたはマイクロコード）。

【0110】

当然のことではあるが、プロセスが、コンピュータによって読み取り可能な媒体に保存される本明細書において提供された技術のいくつかの実施形態に従って、コンピュータによって実施されるプロセスは、その実行の過程の間、手作業で入力を受け取ることができる（例えば、ユーザーから）。

【0111】

したがって、本明細書に記載されたアセンブリデバイスまたは構成成分のシステム全体レベルの制御は、関連する核酸シンセサイザ、液体取り扱いデバイス、サーマルサイクラー、配列決定デバイス、関連するロボット構成成分ならびに所望のインプット/アウトプットもしくはその他の制御機能を実施するためのその他の適したシステムに制御信号を提供し得るシステム制御装置によって実施され得る。したがって、システム制御装置は、任意のデバイス制御装置と一緒に核酸アセンブリシステムの運転を制御する制御装置を形成する。制御装置は、汎用コンピュータまたは汎用コンピュータのネットワークおよびコミュニケーションデバイス、モデムおよび/または所望のインプット/アウトプットもしくはその他の機能を実施するためのその他の電気回路もしくは構成成分を含めたその他の関連デバイスであり得る、汎用データ処理システムを含み得る。制御装置はまた、少なくとも幾分かは、各々、全体的なシステムレベルの制御のための基本処理装置または中央処理装置セクションと、中央処理装置セクションの制御下で、種々の異なる特定の計算、機能およびその他のプロセスを実施するための専用の別のセクションを有する、単一の特別な目的の集積回路（例えば、ASIC）またはASICのアレイとして実施され得る。制御装置はまた、複数の別の専用のプログラム可能な集積回路またはその他の電子回路またはデバイス、例えば、別個の素子回路またはプログラム可能な論理デバイスなどの、配線で接続された電子回路または論理回路を使用して実施され得る。制御装置はまた、任意のその他の構成成分またはデバイス、例えば、ユーザーインプット/アウトプットデバイス（モニタ、ディスプレイ、プリンタ、キーボード、ユーザーポインティングデバイス、タッチスクリーンまたはその他のユーザーインターフェースなど）、データ保存デバイス、駆動モーター、リンケージ、バルブ制御装置、ロボットデバイス、真空およびその他のポンプ、圧力センサー、検出器、電力供給装置、パルス供給装置、コミュニケーションデバイスまたはその他の電子回路または構成成分などを含み得る。制御装置はまた、当技術分野で公知であるが、本明細書において詳細には記載されていないその他の適した機能を実施するために、自動化されたクライアントオーダー処理、品質管理、パッケージング、輸送、請求書作成などといったシステムのその他の部分の運転を制御し得る。

【0112】

種々の本発明の態様は、単独で使用されても、組み合わせて使用されても、前記で記載された実施形態では具体的に論じられていないさまざまな配置で使用されてもよく、したがって、前記の説明において示されたか、または図面において例示された構成成分のその詳細な適用および配置に制限されない。例えば、一実施形態において記載された態様を、任意の方法で、その他の実施形態において記載された態様と組み合わせてもよい。

【0113】

特許請求の範囲の要素を修飾するための、特許請求の範囲における「第一の」、「第二の」、「第三の」などといった順序を示す用語の使用は、それ自体で、ある特許請求の範囲の要素の、別のものを上回る、任意の優先度、先行もしくは順序または方法の動きが実施される一時的な順序を含意するものではなく、特定の名称を有するある特許請求の範囲

10

20

30

40

50

の要素を、同一の名称を有する（しかし、順序を示す用語の使用のために）別の要素と区別するための標識として単に使用されて、特許請求の範囲の要素を区別する。

【 0 1 1 4 】

また、本明細書において使用された表現および技術用語は、記載の目的のためのものであって、制限と見なされてはならない。「含む (including)」、「含む (comprising) または「有する (having)」、「含有する (containing)」、「包含する (involving)」および本明細書におけるその変形の使用は、以下に列挙された項目およびその等価物ならびにさらなる項目を包含するものとする。

【 0 1 1 5 】

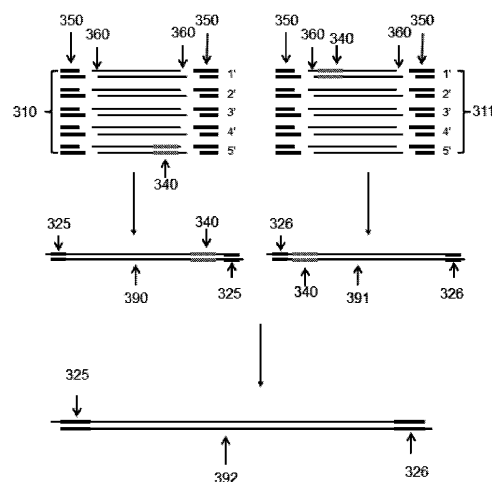
等価物

本発明は、中でも、高忠実度遺伝子アセンブリのための新規方法およびデバイスを提供する。対象発明の特定の実施形態が論じられているが、上記の明細書は例示であって、制限ではない。本発明の多数の等価物は、本明細書を再検討すると当業者には明らかとなろう。本発明の全範囲は、特許請求の範囲を等価物のその全範囲とともに、本明細書をこのような変法と一緒に参照することによって決定されなければならない。

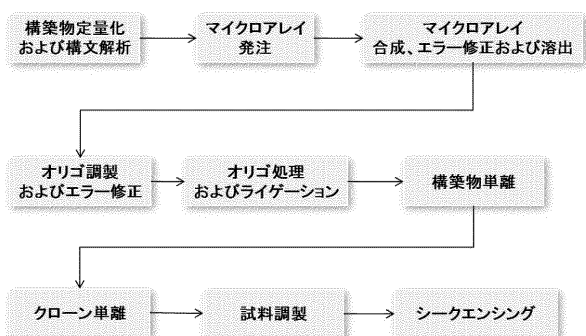
【 0 1 1 6 】

出典明示による組み込み

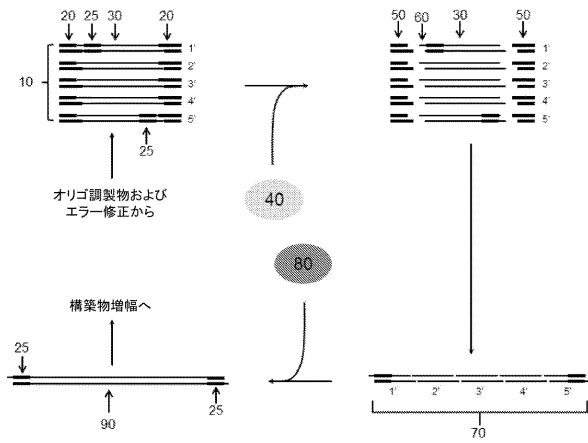
2013年4月24日に出席された米国出席第13/986, 368号および2012年6月15日に出席された米国出席第13/524, 164号、およびPCT公報PCT/US2009/055267が参照される。本明細書において言及されたすべての刊行物、特許および配列データベースの項目は、個々の刊行物または特許が各々、具体的に、また参照により組み込まれるよう個々に指示されるように、出典明示によりその全文が本明細書に組み込まれる。

[illegible]

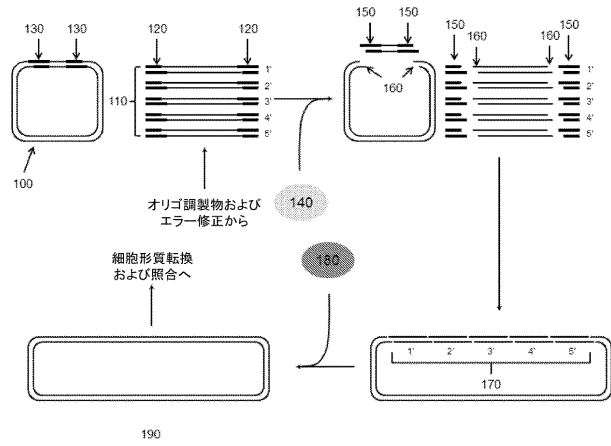
【 义 1 】



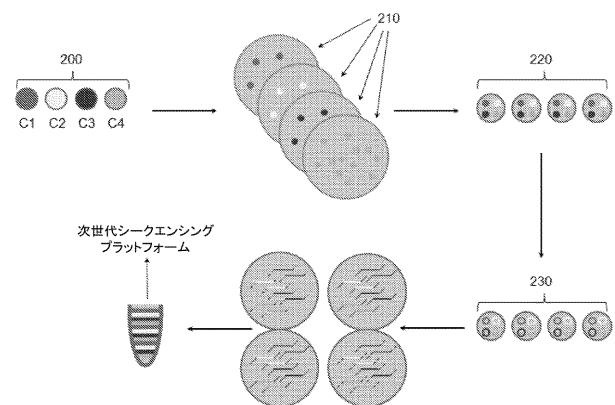
【図 2】



【図 3】



【図 6】



【配列表】

0006509727000001.app

フロントページの続き

- (72)発明者 マイケル・イー・ハドソン
アメリカ合衆国 0 1 7 0 1 マサチューセッツ州フラミンガム、クレストウッド・ドライブ 2 1 番
- (72)発明者 リ・ユン・エイ・クン
アメリカ合衆国 0 2 4 7 4 マサチューセッツ州アーリントン、ノウルズ・ファーム・ロード 4 番
- (72)発明者 ダニエル・シンドラー
アメリカ合衆国 0 2 4 6 4 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 3 1 番
- (72)発明者 スティーブン・アーチャー
アメリカ合衆国 0 2 1 4 1 マサチューセッツ州ケンブリッジ、チャールズ・ストリート 1 2 6 番、
アパートメント 4
- (72)発明者 イシュティアク・サーエム
アメリカ合衆国 0 2 1 5 0 マサチューセッツ州チェルシー、ストックトン・ストリート 1 0 0 番、
アパートメント 2 0 5

審査官 池上 文緒

- (56)参考文献 特表 2 0 0 7 - 5 3 3 3 0 8 (J P , A)
特表 2 0 0 5 - 5 3 8 7 2 5 (J P , A)
Nat. Protoc. (Jan 2012) vol.7, no.1, p.171-192
PLos One (2008) vol.3, issue 11, e3647
PLos ONE (2009) vol.4, issue 5, e5553
BMC Biotechnol. (2011) vol.11, 87
PLos ONE (2011) vol.6, issue 2, e16765

- (58)調査した分野(Int.Cl. , DB名)
C 1 2 N 1 5 / 0 0 - 1 5 / 9 0
J S T P l u s / J M E D P l u s / J S T 7 5 8 0 (J D r e a m I I I)
C A p l u s / M E D L I N E / E M B A S E / B I O S I S (S T N)
W P I D S / W P I X (S T N)