

(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2007-534320  
(P2007-534320A)

(43) 公表日 平成19年11月29日(2007.11.29)

(51) Int. Cl.	F I	テーマコード (参考)
<b>C 1 2 N 15/09 (2006.01)</b>	C 1 2 N 15/00 Z N A A	4 B O 2 4
<b>C 1 2 M 1/00 (2006.01)</b>	C 1 2 N 15/00 F	4 B O 2 9
	C 1 2 M 1/00 A	

審査請求 未請求 予備審査請求 未請求 (全 87 頁)

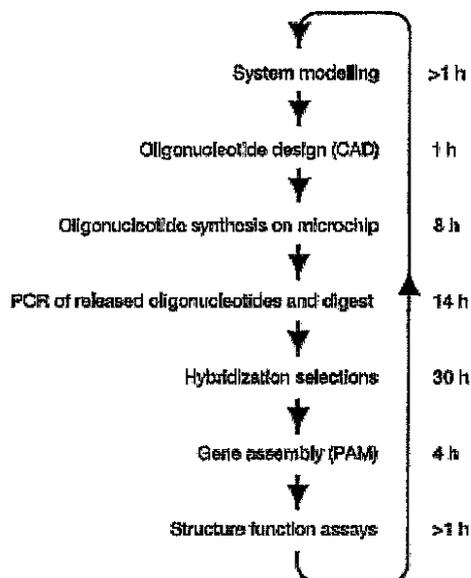
(21) 出願番号	特願2007-500808 (P2007-500808)	(71) 出願人	592257310 プレジデント・アンド・フェロウズ・オブ ・ハーバード・カレッジ アメリカ合衆国02138マサチューセツ ツ州ケンブリッジ、クウィンシー・ストリ ート17
(86) (22) 出願日	平成17年2月28日 (2005. 2. 28)	(74) 代理人	100102978 弁理士 清水 初志
(85) 翻訳文提出日	平成18年10月4日 (2006. 10. 4)	(74) 代理人	100128048 弁理士 新見 浩一
(86) 国際出願番号	PCT/US2005/006429	(72) 発明者	チャーチ ジョージ エム. アメリカ合衆国 マサチューセツ州 ブ ルックリン ケント ストリート 218
(87) 国際公開番号	W02005/089110		
(87) 国際公開日	平成17年9月29日 (2005. 9. 29)		
(31) 優先権主張番号	60/548, 637		
(32) 優先日	平成16年2月27日 (2004. 2. 27)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/600, 957		
(32) 優先日	平成16年8月12日 (2004. 8. 12)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/636, 672		
(32) 優先日	平成16年12月16日 (2004. 12. 16)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 ポリヌクレオチド合成法

(57) 【要約】

反応物質が低濃度で存在する場合の二分子相互作用の速度式を向上させる方法が提供される。高濃度のユニバーサルプライマーを用いて1種または複数種のオリゴヌクレオチドを予め増幅する方法が提供される。オリゴヌクレオチドおよび/またはポリヌクレオチド合成中のエラー率を改善する方法が同様に提供される。配列最適化およびオリゴヌクレオチド設計の方法がさらに提供される。



## 【特許請求の範囲】

## 【請求項1】

以下の段階を含む、所定の配列を有するポリヌクレオチド構築体を調製する方法：

- a) 以下の段階を含む、構築用オリゴヌクレオチドのプールを提供する段階
  - i) 該ポリヌクレオチド構築体の配列を規定 (define) する部分的に重複する配列、
  - ii) 該構築用オリゴヌクレオチドの少なくとも一部分で隣接するおよび該構築用オリゴヌクレオチドの少なくともサブセットに共通する少なくとも1組のプライマーハイブリダイゼーション部位、ならびに
  - iii) プライマーハイブリダイゼーション部位と構築用オリゴヌクレオチドとの間の切断部位；
- b) プライマーハイブリダイゼーション部位に結合する少なくとも1つのプライマーを用いて、該構築用オリゴヌクレオチドのプールを増幅させる段階；
- c) 該プライマーハイブリダイゼーション部位を該切断部位で該構築用オリゴヌクレオチドから除去する段階；
- d) 該構築用オリゴヌクレオチドの相補鎖を分離する段階；ならびに
- e) 該構築用オリゴヌクレオチドのプールをハイブリダイゼーション条件ならびに(i) ライゲーション条件、(ii) 鎖伸長条件、または(iii) 鎖伸長およびライゲーション条件に曝露し、ポリヌクレオチド構築体を形成させる段階。

10

## 【請求項2】

構築用オリゴヌクレオチドが固体支持体上で合成される、請求項1記載の方法。

20

## 【請求項3】

構築用オリゴヌクレオチドが増幅の前に支持体から切断される、請求項2記載の方法。

## 【請求項4】

合成が光誘導による反応を支持体上の別々の位置で利用する、請求項2記載の方法。

## 【請求項5】

マスクを用いて光が別々の位置に向けられる、請求項4記載の方法。

## 【請求項6】

光案内マスクレスオプティクス(light directing maskless optics)を用いて光が別々の位置に向けられる、請求項4記載の方法。

## 【請求項7】

PCRを用いて構築用オリゴヌクレオチドのプールが増幅される、請求項1記載の方法。

30

## 【請求項8】

構築用オリゴヌクレオチドをアッセンブリの前にエラー低減過程に供する段階をさらに含む、請求項1記載の方法。

## 【請求項9】

構築用オリゴヌクレオチドをアッセンブリの前に少なくとも2ラウンドの増幅およびエラー低減に供する段階をさらに含む、請求項8記載の方法。

## 【請求項10】

エラー低減過程がハイブリダイゼーション条件の下で構築用オリゴヌクレオチドのプールを選択用オリゴヌクレオチドのプールに曝露する段階および選択用オリゴヌクレオチドにハイブリダイズされる際に mismatches を含む構築用オリゴヌクレオチドのコピーを除去する段階によるエラーの過である、請求項8記載の方法。

40

## 【請求項11】

ポリヌクレオチド構築体をさらなるアッセンブリに供して、それによりさらに長いポリヌクレオチド構築体を形成させる段階をさらに含む、請求項1記載の方法。

## 【請求項12】

ポリヌクレオチド構築体をエラー低減過程に供する段階をさらに含む、請求項1または11記載の方法。

## 【請求項13】

ポリヌクレオチド構築体を増幅に供する段階をさらに含む、請求項1または12記載の方

50

法。

【請求項14】

ポリヌクレオチド構築体が少なくとも約1キロベースである、請求項1記載の方法。

【請求項15】

ポリヌクレオチド構築体が少なくとも約10キロベースである、請求項14記載の方法。

【請求項16】

ポリヌクレオチド構築体が少なくとも約100キロベースである、請求項15記載の方法。

【請求項17】

ポリヌクレオチド構築体が少なくとも約1メガベースである、請求項16記載の方法。

【請求項18】

ポリヌクレオチド構築体が少なくとも約1ギガベースである、請求項16記載の方法。

10

【請求項19】

ポリヌクレオチド構築体をベクターに挿入する段階をさらに含む、請求項1記載の方法

【請求項20】

ポリヌクレオチド構築体を宿主細胞に導入する段階をさらに含む、請求項1または19記載の方法。

【請求項21】

ポリヌクレオチド構築体が少なくとも1種のポリペプチド配列をコードする、請求項1または20記載の方法。

20

【請求項22】

ポリヌクレオチド構築体から少なくとも1種のポリペプチドを発現させる段階をさらに含む、請求項1または20記載の方法。

【請求項23】

構築用オリゴヌクレオチドの全てが共通する少なくとも1組のプライマーハイブリダイゼーション部位を含む、請求項1記載の方法。

【請求項24】

構築用オリゴヌクレオチドが光切断可能なリンカーによって固体支持体に結合される、請求項2記載の方法。

【請求項25】

制限エンドヌクレアーゼを用いてプライマーハイブリダイゼーション部位が構築用オリゴヌクレオチドから除去される、請求項1記載の方法。

30

【請求項26】

制限エンドヌクレアーゼがIIS型制限エンドヌクレアーゼである、請求項25記載の方法

【請求項27】

以下の段階を含む、精製済みの構築用オリゴヌクレオチドのプールを調製する方法：

a) 構築用オリゴヌクレオチドのプールを提供する段階；

b) ハイブリダイゼーション条件の下で構築用オリゴヌクレオチドのプールを選択用オリゴヌクレオチドのプールと接触させ、それにより、少なくとも、二重鎖の一部分が相補的な領域中に mismatches を含まない構築用オリゴヌクレオチドのコピーと選択用オリゴヌクレオチドのコピーとを含んだ安定な二重鎖でありおよび二重鎖の一部分が相補的な領域中に1つまたは複数の mismatches を含む構築用オリゴヌクレオチドのコピーと選択用オリゴヌクレオチドのコピーとを含んだ不安定な二重鎖である、二重鎖を形成させる段階；ならびに

40

c) 不安定な二重鎖を形成した構築用オリゴヌクレオチドのコピーを除去し、それにより精製済みの構築用オリゴヌクレオチドのプールを形成させる段階。

【請求項28】

以下の段階を含む、所定の配列を有するポリヌクレオチド構築体を調製する方法：

d) 該ポリヌクレオチド構築体の配列を規定する部分的に重複する配列を含んだ構築用

50

オリゴヌクレオチドのプールを提供する段階；

e) ハイブリダイゼーション条件の下で構築用オリゴヌクレオチドのプールを選択用オリゴヌクレオチドのプールと接触させ、それにより、少なくとも、二重鎖の一部分が相補的な領域中にミスマッチを含まない構築用オリゴヌクレオチドのコピーと選択用オリゴヌクレオチドのコピーとを含んだ安定な二重鎖でありおよび二重鎖の一部分が相補的な領域中に1つまたは複数のミスマッチを含む構築用オリゴヌクレオチドのコピーと選択用オリゴヌクレオチドのコピーとを含んだ不安定な二重鎖である、二重鎖を形成させる段階；および

f) 不安定な二重鎖を形成した構築用オリゴヌクレオチドのコピーを除去する段階；

g) 残存する複合体を変性させ、それにより精製済みの構築用オリゴヌクレオチドのプールを形成させる段階；ならびに

h) 精製済みの構築用オリゴヌクレオチドをハイブリダイゼーション条件ならびに(i) ライゲーション条件、(ii) 鎖伸長条件、または(iii) 鎖伸長およびライゲーション条件に曝露して、それによりポリヌクレオチド構築体を形成させる段階。

10

【請求項29】

ポリヌクレオチド構築体を形成させる段階の前に構築用オリゴヌクレオチドを増幅させる段階をさらに含む、請求項28記載の方法。

【請求項30】

精製済みの構築用オリゴヌクレオチドのプールを選択用オリゴヌクレオチドのプールと少なくとも二度接触させる段階および不安定な二重鎖を形成した構築用オリゴヌクレオチドのコピーを除去する段階をさらに含む、請求項28または29記載の方法。

20

【請求項31】

選択用オリゴヌクレオチドが固体支持体上に固定化される、請求項27または28記載の方法。

【請求項32】

選択用オリゴヌクレオチドがカラム中に含まれる、請求項27または28記載の方法。

【請求項33】

構築用オリゴヌクレオチドのプールへの曝露の前に選択用オリゴヌクレオチドを増幅させる段階をさらに含む、請求項27または28記載の方法。

【請求項34】

選択用オリゴヌクレオチドが、該選択用オリゴヌクレオチドの少なくとも一部分で隣接するおよび構築用オリゴヌクレオチドの少なくともサブセットに共通する少なくとも1組のプライマーハイブリダイゼーション部位を含む、請求項27、28または33のいずれか一項記載の方法。

30

【請求項35】

選択用オリゴヌクレオチドの全てが共通する少なくとも1組のプライマーハイブリダイゼーション部位を含む、請求項34記載の方法。

【請求項36】

構築用オリゴヌクレオチドが、該構築用オリゴヌクレオチドの少なくとも一部分で隣接するおよび該構築用オリゴヌクレオチドの少なくともサブセットに共通する少なくとも1組のプライマーハイブリダイゼーション部位を含む、請求項27または28記載の方法。

40

【請求項37】

少なくとも大量の構築用オリゴヌクレオチドが共通する少なくとも1組のプライマーハイブリダイゼーション部位を含む、請求項36記載の方法。

【請求項38】

選択用オリゴヌクレオチドが固体支持体上で合成される、請求項27または28記載の方法。

【請求項39】

構築用オリゴヌクレオチドが増幅の前に支持体から切断される、請求項38記載の方法。

【請求項40】

50

合成が光誘導による反応を支持体上の別々の位置で利用する、請求項38記載の方法。

【請求項41】

マスクを用いて光が別々の位置に向けられる、請求項40記載の方法。

【請求項42】

光案内マスクレス光学を用いて光が別々の位置に向けられる、請求項30記載の方法。

【請求項43】

選択用オリゴヌクレオチドが光切断可能なリンカーによって固体支持体に結合される、請求項38記載の方法。

【請求項44】

精製済みの構築用オリゴヌクレオチドが500塩基中約1エラー未満の塩基エラー率を有する、請求項27、28、または30のいずれか一項記載の方法。

【請求項45】

精製済みの構築用オリゴヌクレオチドが1,000塩基中約1エラー未満の塩基エラー率を有する、請求項44記載の方法。

【請求項46】

精製済みの構築用オリゴヌクレオチドが10,000塩基中約1エラー未満の塩基エラー率を有する、請求項45記載の方法。

【請求項47】

構築用オリゴヌクレオチドのプールが重複領域において相補的なプラス鎖およびマイナス鎖を含む、請求項27記載の方法。

【請求項48】

以下の段階を含む、単一のプール中で異なる所定の配列を有する複数のポリヌクレオチド構築体を調製する方法：

a) 該複数のポリヌクレオチド構築体の各々の配列を規定する部分的に重複する配列を含む構築用オリゴヌクレオチドのプールを提供する段階；および

b) 該構築用オリゴヌクレオチドのプールをハイブリダイゼーション条件ならびに以下の条件：(i) ライゲーション条件、(ii) 鎖伸長条件、または(iii) 鎖伸長およびライゲーション条件のうち少なくとも1つの下でインキュベートし、それにより単一のプール中で該複数のポリヌクレオチド構築体を形成させる段階。

【請求項49】

構築用オリゴヌクレオチドのプールが重複領域において相補的なプラス鎖およびマイナス鎖を含む、請求項48記載の方法。

【請求項50】

少なくとも4種のポリヌクレオチド構築体が単一のプール中で調製される、請求項48記載の方法。

【請求項51】

少なくとも10種のポリヌクレオチド構築体が単一のプール中で調製される、請求項48記載の方法。

【請求項52】

少なくとも約100種のポリヌクレオチド構築体が単一のプール中で調製される、請求項48記載の方法。

【請求項53】

複数のポリヌクレオチド構築体をさらなるアッセンブリに供して、それにより少なくとも1種のさらに長いポリヌクレオチド構築体を形成させる段階をさらに含む、請求項48記載の方法。

【請求項54】

さらなるアッセンブリが以下の段階を含む、請求項53記載の方法：

a) ポリヌクレオチド構築体を融解させる段階；および

b) ポリヌクレオチド構築体をハイブリダイゼーション条件ならびに(i) ライゲーション

10

20

30

40

50

ン条件、(ii) 鎖伸長条件、または(iii) 鎖伸長およびライゲーション条件に曝露して、それによりさらに長いポリヌクレオチド構築体を形成させる段階。

【請求項55】

さらなるアッセムブリが以下の段階を含む、請求項51記載の方法：

- a) ポリヌクレオチド構築体を制限酵素と接触させる段階；および
- b) ポリヌクレオチド構築体をハイブリダイゼーション条件ならびに(i) ライゲーション条件、(ii) 鎖伸長条件、または(iii) 鎖伸長およびライゲーション条件に曝露して、それによりさらに長いポリヌクレオチド構築体を形成させる段階。

【請求項56】

構築用オリゴヌクレオチドを少なくとも1ラウンドの(i) 増幅、(ii) エラー低減、または(iii) いずれかの順序の増幅およびエラー低減に供する段階をさらに含む、請求項48記載の方法。

【請求項57】

ポリヌクレオチド構築体を少なくとも1ラウンドの(i) 増幅、(ii) エラー低減、または(iii) いずれかの順序の増幅およびエラー低減に供する段階をさらに含む、請求項53記載の方法。

【請求項58】

以下の段階を含む、異なる所定の配列を有する複数のポリヌクレオチド構築体を調製する方法：

- a) 各ポリヌクレオチド構築体の配列を部分的に重複する配列セグメントにコンピュータにより分割する段階；
- b) 部分的に重複する配列セグメントのセットに応答する配列を含む構築用オリゴヌクレオチドを合成する段階；および
- c) 該構築用オリゴヌクレオチドをハイブリダイゼーション条件ならびに以下の条件：(i) ライゲーション条件、(ii) 鎖伸長条件、または(iii) 鎖伸長およびライゲーション条件のうち少なくとも1つの下でインキュベートし、それにより複数のポリヌクレオチド構築体を形成させる段階。

【請求項59】

構築用オリゴヌクレオチドが重複領域において相補的なプラス鎖およびマイナス鎖を含む、請求項58記載の方法。

【請求項60】

以下の段階をさらに含む、請求項58記載の方法：

- a) 構築用オリゴヌクレオチドの少なくとも一部分の末端に、該構築用オリゴヌクレオチドの少なくともサブセットに共通するおよびプライマーハイブリダイゼーション部位と構築用オリゴヌクレオチドとの間の切断部位を規定する1組または複数組のプライマーハイブリダイゼーション部位をコンピュータにより付加する段階；
- b) 該プライマーハイブリダイゼーション部位に結合する少なくとも1つのプライマーを用いて該構築用オリゴヌクレオチドを増幅する段階；および
- c) 該プライマーハイブリダイゼーション部位を該切断部位で該構築用オリゴヌクレオチドから除去する段階。

【請求項61】

ポリヌクレオチド構築体の少なくとも一部分の末端を規定する構築用オリゴヌクレオチドに、該ポリヌクレオチド構築体の少なくともサブセットに共通する1組または複数組のプライマーハイブリダイゼーション部位をコンピュータにより付加する段階をさらに含む、請求項60記載の方法。

【請求項62】

プライマーハイブリダイゼーション部位と末端を規定する構築用オリゴヌクレオチドとの間の切断部位を規定する配列をコンピュータにより設計する段階をさらに含む、請求項61記載の方法。

【請求項63】

10

20

30

40

50

ポリヌクレオチド構築体を増幅する段階をさらに含む、請求項61または62記載の方法。

【請求項64】

プライマーハイブリダイゼーション部位をポリヌクレオチド構築体から除去する段階をさらに含む、請求項63記載の方法。

【請求項65】

ポリヌクレオチド構築体をさらに長いポリヌクレオチド構築体にアッセンブルする段階をさらに含む、請求項58、63、または64のいずれか一項記載の方法。

【請求項66】

構築用オリゴヌクレオチドを少なくとも1ラウンドのエラー低減に供する段階をさらに含む、請求項58記載の方法。

【請求項67】

エラー低減過程が以下の段階を含む、請求項66記載の方法：

a) 構築用オリゴヌクレオチドの少なくとも一部分に相補的である配列を含む選択用オリゴヌクレオチドの少なくとも1つのプールをコンピュータにより設計する段階；

b) 該選択用オリゴヌクレオチドを合成する段階；

c) ハイブリダイゼーション条件の下で構築用オリゴヌクレオチドのプールを選択用オリゴヌクレオチドのプールと接触させ、それにより、少なくとも、二重鎖の一部分が相補的な領域中に mismatches を含まない構築用オリゴヌクレオチドのコピーと選択用オリゴヌクレオチドのコピーとを含んだ安定な二重鎖でありおよび二重鎖の一部分が相補的な領域中に1つまたは複数の mismatches を含む構築用オリゴヌクレオチドのコピーと選択用オリ

ゴヌクレオチドのコピーとを含んだ不安定な二重鎖である、二重鎖を形成させる段階；

d) 不安定な二重鎖を形成した構築用オリゴヌクレオチドのコピーを除去する段階；および

e) 残存する二重鎖を変性させ、それにより精製済みの構築用オリゴヌクレオチドのプールを形成させる段階。

【請求項68】

二重鎖の融解温度を最適化するため、構築用オリゴヌクレオチド、選択用オリゴヌクレオチド、またはその両方の長さ、コドン使用頻度、または長さおよびコドン使用頻度をコンピュータにより最適化する段階をさらに含む、請求項67記載の方法。

【請求項69】

選択用オリゴヌクレオチドが固体支持体上に固定化される、請求項67記載の方法。

【請求項70】

選択用オリゴヌクレオチドのプール中の1種または複数種の選択用オリゴヌクレオチドの配列が構築用オリゴヌクレオチドのプール中の構築用オリゴヌクレオチドの少なくとも一部分の配列全体に相補的である、請求項67記載の方法。

【請求項71】

構築用オリゴヌクレオチド、選択用オリゴヌクレオチド、またはその両方が固体支持体上で合成される、請求項58または67記載の方法。

【請求項72】

ポリヌクレオチド構築体のコピーの配列を決定する段階をさらに含む、請求項58記載の方法。

【請求項73】

ポリヌクレオチド構築体を細胞に導入する段階をさらに含む、請求項58記載の方法。

【請求項74】

ポリヌクレオチド構築体が少なくとも1種のポリペプチドをコードする、請求項58記載の方法。

【請求項75】

ポリヌクレオチド構築体が特定の宿主細胞で発現させるためにコドン最適化された、請求項74記載の方法。

【請求項76】

10

20

30

40

50

以下の段階を含む、単一のプール中で複数の異なるポリヌクレオチド配列をアッセンブルする方法：

- a) 相補的な末端領域と該異なるポリヌクレオチド配列の末端を含んだオリゴヌクレオチドに隣接するプライマー部位とを有する合成オリゴヌクレオチドの群を提供する段階；
- b) 該合成オリゴヌクレオチドをdNTPsおよびポリメラーゼとともに混合する段階；
- c) 以下をもたらすよう混合物をサイクリングにかける段階：

該相補的な末端領域のハイブリダイゼーション；

重複するオリゴヌクレオチドを伸長させるおよび完全長の異なるポリヌクレオチド配列のコピーを産生させるポリメラーゼを介した塩基の取込み；および  
複数の該完全長配列の増幅。

10

【請求項 77】

産生されたポリヌクレオチド配列をさらにアッセンブルしてより大きなポリヌクレオチドを産生させる段階を含む請求項76記載の方法であって、以下のさらなる段階を含む方法：

d) 複数の別個のプール中で請求項76記載の方法を行い、異なる合成ポリヌクレオチド配列の少なくともいくつかをそれにより、相補的な末端領域と該より大きなポリヌクレオチドの末端を含んだ異なるポリヌクレオチド配列に隣接するプライマー部位とを有するポリヌクレオチドを含む各プール中で産生させる段階；

e) 該複数のプールの少なくともいくつかをdNTPsおよびポリメラーゼとともに混合する段階；ならびに

20

f) 以下をもたらすよう混合物をサイクリングにかける段階：

該異なるポリヌクレオチド配列の該相補的な末端領域のハイブリダイゼーション；

重複するオリゴヌクレオチド配列を伸長させるおよび完全長のより大きなポリヌクレオチドのコピーを産生させるポリメラーゼを介した塩基の取込み；および

複数の該完全長のより大きなポリヌクレオチドの増幅。

【請求項 78】

合成オリゴヌクレオチドが複数の塩基配列の連続自動並行アッセンブリにより並行して合成され、精製されて、配列エラーを包含するオリゴヌクレオチドコピーの濃度を減らす、請求項76記載の方法。

【請求項 79】

精製がハイブリダイゼーションにより行われる、請求項78記載の方法。

30

【請求項 80】

合成オリゴヌクレオチドが表面上で合成される、請求項76記載の方法。

【請求項 81】

相補的な末端領域の複数のペアがほぼ同じ融解温度を有するように設計される、請求項76または77記載の方法。

【請求項 82】

プールがウェルまたはマイクロチャンネルである、請求項76記載の方法。

【請求項 83】

混合する段階e)が微小流体システムの中に混合物の成分をともに流すことで行われる、請求項77記載の方法。

40

【請求項 84】

ポリメラーゼが熱安定性ポリメラーゼである、請求項76記載の方法。

【請求項 85】

多数の異なる回収可能なポリヌクレオチドを含む製造物品であって、以下を含む物品：

容器由来の異なるポリヌクレオチドの部分集団の増幅を可能にするプライマー配列の異なるペアを含んだ異なるポリヌクレオチド混合物を含むポリヌクレオチド用容器；および  
複数のプライマー用容器であって、それぞれが構築用容器中のポリヌクレオチドのプライマー配列のペアに相補的なオリゴヌクレオチドプライマーのペアを含んだ容器。

【請求項 86】

50

ポリヌクレオチド用容器中のポリヌクレオチドのプライマー配列ペアが互いに異なる、請求項85記載の物品。

【請求項87】

ポリヌクレオチドが合成DNAを含む、請求項85記載の物品。

【請求項88】

ポリヌクレオチドが遺伝子を含む、請求項85記載の物品。

【請求項89】

ポリヌクレオチドが野生型配列の複数の変異体を含む、請求項85記載の物品。

【請求項90】

ポリヌクレオチドがベクターを含む、請求項85記載の物品。

10

【請求項91】

ポリヌクレオチドの少なくとも一部分が少なくとも1 Kb長である、請求項85記載の物品。

【請求項92】

ポリヌクレオチドの少なくとも一部分が少なくとも2 Kb長である、請求項85記載の物品。

【請求項93】

ポリヌクレオチドの少なくとも一部分が少なくとも10 Kb長である、請求項85記載の物品。

【請求項94】

ポリヌクレオチドの少なくとも一部分が環状化される、請求項85記載の物品。

20

【請求項95】

ポリヌクレオチドが、ポリヌクレオチド配列の操作を容易にするアダプター配列に隣接しているポリヌクレオチド配列を含む、請求項85記載の物品。

【請求項96】

アダプター配列がベクターへの挿入、固定化、および配列の機能の同定の1つまたは複数を容易にする、請求項95記載の物品。

【請求項97】

ポリヌクレオチドの混合物が哺乳類の配列、酵母の配列、原核生物の配列、植物の配列、キイロショウジョウバエ (*D. melanogaster*) の配列、線虫 (*C. elegans*) の配列、およびアフリカツメガエル (*Xenopus*) の配列からなる群より選択される1つまたは複数の配列を含む、請求項85記載の物品。

30

【請求項98】

異なる回収可能なポリヌクレオチド構築体の混合物が独立して回収可能である、請求項85記載の物品。

【請求項99】

複数の異なるポリヌクレオチドを含有する複数のポリヌクレオチド用容器、プライマー配列の同一のペアを含んだ異なる容器中のポリヌクレオチドを含む、請求項85記載の物品。

【請求項100】

複数のプライマー用容器の1つまたは複数が相補的なオリゴヌクレオチドプライマーのペアを含む、請求項85または99記載の物品。

40

【請求項101】

ポリヌクレオチド用容器がD個の異なる独立して回収可能なポリヌクレオチドであって、それぞれがN個のネスティッドプライマーペアを含むポリヌクレオチド、少なくとも  $N/2 \times D^{1/N}$  であるプライマー用容器の数を含む、請求項85または99記載の物品。

【請求項102】

ポリヌクレオチド用容器がD個の異なるポリヌクレオチドおよびプライマーのペアを含有するD個のプライマー用容器を含む、請求項85または99記載の物品。

【請求項103】

50

ポリヌクレオチド用容器が、プライマー配列の複数のネスティッドペアを含む異なるポリヌクレオチドであって、該複数のネスティッドペアのそれぞれが該容器中の選択のポリヌクレオチド群のまたはその容器中の該異なるポリヌクレオチドの個々のものの増幅を可能にするポリヌクレオチドを含む、請求項85または99記載の物品。

【請求項104】

$10^2$ 個の異なるポリヌクレオチドを含む、請求項85または99記載の物品。

【請求項105】

$10^3$ 個の異なるポリヌクレオチドを含む、請求項85または99記載の物品。

【請求項106】

$10^4$ 個の異なるポリヌクレオチドを含む、請求項85または99記載の物品。

10

【請求項107】

$10^5$ 個の異なるポリヌクレオチドを含む、請求項99記載の物品。

【請求項108】

$10^6$ 個の異なるポリヌクレオチドを含む、請求項99記載の物品。

【請求項109】

多数の異なる回収可能なポリヌクレオチドを含有する包装を含む製造物品であって、以下を含む物品：

異なるポリヌクレオチドの少なくともいくつかはプライマー配列の複数のネスティッドペアを含み、該複数のネスティッドペアのそれぞれが容器中の選択のポリヌクレオチド群のまたはその容器中の該異なるポリヌクレオチドの個々のものの増幅を可能にする異なるポリヌクレオチドの混合物を含んだポリヌクレオチド用容器；および

20

複数のプライマー用容器であって、それぞれが構築用容器中のポリヌクレオチドのプライマー配列のペアに相補的なオリゴヌクレオチドプライマーのペアを含んだ容器。

【請求項110】

容器中の各ポリヌクレオチド上のネスティッドペアの組合せが、該容器中のその他全てのポリヌクレオチドのネスティッドペアの組合せと異なる、請求項109記載の物品。

【請求項111】

それぞれが複数の異なるポリヌクレオチドを含む複数の構築用容器、所与のプライマーペアが異なる容器中の異なるポリヌクレオチドとアニーリングするようにプライマー配列の同一のペアを含んだ異なる容器中のポリヌクレオチドを含む、請求項85または109記載の物品。

30

【請求項112】

ポリヌクレオチド構築体の選択の1つまたは選択の群に富む溶液を供給する装置であって、以下を含む装置：

容器由来の異なるポリヌクレオチドのうち選択のものの増幅を可能にするおよび該容器中のその他のポリヌクレオチドのプライマー配列のその他のペアとは異なる、プライマー配列の少なくとも1ペアを含む同定済みのポリヌクレオチドの混合物を含んだポリヌクレオチド用容器；

複数のプライマー用容器であって、それぞれが構築用容器中の異なるポリヌクレオチドのプライマー配列のペアに相補的なオリゴヌクレオチドプライマーのペアを含んだ容器；

40

同定済みのポリヌクレオチドと各同定済みのポリヌクレオチドに相補的なプライマーのペアまたは複数ペアを含む1つまたは複数の容器の位置とを収載するデータ保存庫；

ポリヌクレオチドまたはポリヌクレオチドの群を利用者が指定するのを可能にするインターフェース；

該インターフェースで入力された仕様に応答する自動的手段、および該指定されたポリヌクレオチドまたはポリヌクレオチドの群を選択的に増幅するのに必要とされる試薬を調製するよう該構築用容器からポリヌクレオチドおよび選択のプライマー用容器からプライマーの一定分量を抽出するために該データ保存庫からアクセスされる命令。

【請求項113】

異なる同定済みのポリヌクレオチドを含んだ複数のポリヌクレオチド用容器を含む、請

50

求項112記載の装置。

【請求項114】

異なる容器中の同定済みのポリヌクレオチドが、プライマー配列の同一ペアを含む、請求項113記載の装置。

【請求項115】

異なる容器中の同定済みのポリヌクレオチドが、プライマー配列の複数のネスティッドペアを含む、請求項113記載の装置。

【請求項116】

少なくとも10個のポリヌクレオチド用容器を含む、請求項112記載の装置。

【請求項117】

異なる容器中の同定済みのポリヌクレオチドが、プライマー配列の固有のネスティッドペアを含む、請求項112記載の装置。

10

【請求項118】

構築用容器から回収される選択の同定済みポリヌクレオチドを選択のプライマーペアによる指定のとおり増幅するよう適合された増幅用チャンバをさらに含む、請求項112記載の装置。

【請求項119】

増幅用チャンバから回収される同定済みのポリヌクレオチドの1つまたは部分集団を選択のプライマーペアによる指定のとおり増幅するよう適合された第2の増幅用チャンバをさらに含む、請求項118記載の装置。

20

【請求項120】

以下の段階を含む、選択のポリヌクレオチドを得る方法：

容器由来のポリヌクレオチドのうち選択のものの増幅を可能にするプライマー配列の複数のネスティッドペアを含み、該容器中のあるポリヌクレオチドのプライマーペアの組合せが該容器中のその他のポリヌクレオチドのプライマー配列のその他のペアとは異なる同定済みの合成ポリヌクレオチドの混合物を含有する複数の構築用容器を提供する段階；

複数のプライマー用容器であって、それぞれが構築用容器中のポリヌクレオチドのプライマー配列のペアに相補的なオリゴヌクレオチドプライマーのペアを含む容器を提供する段階；

第1の増幅手順を、選択の該構築用容器から回収されるポリヌクレオチドの該混合物と1つまたは複数のプライマー用容器から回収される、プライマー配列の外側ネスティッドペアに相補的なプライマーのペアの一定分量を含む第1の増幅混合物で行う段階；および

30

第2の増幅手順を、該第1の増幅混合物から回収される単位複製配列と1つまたは複数のプライマー用容器から回収される、プライマー配列の内側ネスティッドペアに相補的なプライマーのペアの一定分量を含む第2の増幅混合物で行う段階。

【請求項121】

多数のポリヌクレオチド種であって、少なくともいくつかライブラリーから回収される種のうちの選択群の増幅を可能とするのに十分な長さのプライマー配列の外側ペアを有する種、および該外側ペアを用いた増幅によって産生された単位複製配列の混合物から回収される種のうちの1つまたは選択群の増幅を可能とするのに十分な長さを有するプライマー配列の内側ペアを含む、ライブラリーを形成する混合物中の多数の合成ポリヌクレオチド。

40

【請求項122】

ライブラリー中の個々の種の濃度が、該ライブラリーから直接的にその選択的増幅を可能とするには十分ではないが、外側のプライマー配列ペアを用いた増幅後にその選択的増幅を可能とするには十分である、請求項121記載のライブラリー。

【請求項123】

合成ポリヌクレオチドがプライマー配列の3組のネスティッドペアを含む、請求項121記載のライブラリー。

【請求項124】

50

合成ポリヌクレオチドが各々、ライブラリー中のプライマー配列のその他全てのネステッドペアとは異なる核酸配列を有するプライマー配列のネステッドペアを含む、請求項121記載のライブラリー。

【発明の詳細な説明】

【技術分野】

【0001】

発明の分野

本発明は合成ポリヌクレオチドを作製する方法に関する。

【0002】

関連する米国出願

本願は、あらゆる目的でその全体が参照により本明細書に組み入れられる2004年2月27日付で出願された米国仮特許出願第60/548,637号、2004年8月12日付で出願された同第60/600,957号および2004年12月16日付で出願された同第60/636,672号の優先権を主張する。

【0003】

政府の利益についての記述

本発明は国防総省国防高等研究事業局(DARPA)から授与された助成番号F30602-01-2-0586の下で政府の支援によりなされた。政府は本発明において一定の権利を有する。

【背景技術】

【0004】

発明の背景

配列決定、マイクロアレイおよびプロテオミクスなどの大規模生化学分析の進歩によって膨大なデータが生み出されており、これを計算生物学者らは多数の仮説に利用している。しかしながら、新しい遺伝要素、遺伝経路および遺伝子操作細胞を構築する際の障害が克服されなければならない。ダーウィン淘汰を用いる複雑な生物過程を最適化するには、コンビナトリアル・オリゴヌクレオチド合成で利用できる有限の多様性(無作為化された約25塩基対(bp)または等価物)では、DNA配列の大きなストレッチを通じて(メガ塩基レベルで)注意深く指図されなければならない。これらは合成生物学の新興分野にとって大きな課題と潜在的な利益である。

【0005】

カスタム遺伝子およびゲノムを十分に供与する有用な種々の分子、細胞のおよび無細胞の系を作製するため、方法が当技術分野において利用可能である。しかしながら、実に単純なオリゴヌクレオチドを作製する現行の方法は高価であり(1ヌクレオチド当たり0.11米ドル)、非常に高いレベルのエラー(100塩基中1塩基の割合で欠失ならびに400塩基中およそ1塩基の割合でミスマッチおよび挿入)を有する。結果的に、オリゴヌクレオチドからの遺伝子またはゲノム合成は、高価でもありエラーを起こしやすくもある。クローン配列決定および突然変異誘発法によるエラーの補正は、労働量と総費用をさらに(1塩基対当たり少なくとも2米ドルまで)増大させる。

【0006】

オリゴヌクレオチド合成の費用は、マイクロチップ上で大規模並列カスタム合成を行うことにより削減することができる(Zhou et al. (2004) *Nucleic Acids Res.* 32:5409 (非特許文献1); Fodor et al. (1991) *Science* 251:767 (非特許文献2))。これは、標準的な試薬を用いたインクジェット印刷(Agilent; 例えば、米国特許第6,323,043号(特許文献1)を参照のこと)、光に不安定な5'保護基(Nimbelgen/Affymetrix; 例えば、米国特許第5,405,783号(特許文献2); およびPCT公報番号W0 03/065038(特許文献3); W0 03/064699(特許文献4); W0 03/064026(特許文献5); W0 02/04597(特許文献6)を参照のこと)、光生成酸による脱保護(例えば、AtacticおよびXeotron技術、例えば、X. Gao et al., *Nucleic Acids Res.* 29: 4744-50 (2001) (非特許文献3); X. Gao et al., *J. Am. Chem. Soc.* 120: 12698-12699 (1998) (非特許文献4); O. Srivannavit et al., *Sensors and Actuators A.* 116: 150-160 (2004) (非特許文献5); および米国特許第6,426,184号(特許文献7)を参照のこと)および電氣的酸/塩基アレイ(Oxamer/Combimatrix; 例えば

10

20

30

40

50

、米国特許第2003/0054344号(特許文献8);米国特許第6,093,302号(特許文献9);米国特許第6,444,111号(特許文献10);米国特許第6,280,595号(特許文献11)を参照のこと)を含めて、様々な方法を用いて達成することができる。しかしながら、現行のマイクロチップは表面積が非常に小さく、故にごく少量のオリゴヌクレオチドしか産生できない。溶液中に放出された場合、オリゴヌクレオチドは1配列当たりピコモル(picomolar)またはそれより低い濃度、つまり二分子プライミング反応を効率的に推進するのに十分に高くはない濃度で存在することになる。

【0007】

正確なDNA構築体の製造は、化学合成技術に付いて回るエラー率によって大きな影響を受ける。図1が示すように、例証として、1000塩基中1塩基のエラー率を有する方法によって合成された、3000塩基対を含む読み取り枠を包含するDNAでは、合成されたDNAのコピーの5%未満は正しいと考えられる。

10

【0008】

ホスホアミダイト化学合成法を利用する最先端のオリゴヌクレオチド合成機は、200塩基中およそ1塩基の割合でエラーを起こす。光解離性合成技術を用いてチップ上で合成されたDNAは、伝えられるところでは、約1/50のエラー率を有し、場合によっては約1/100まで改善されることができる。高忠実度PCRは約1/10<sup>5</sup>のエラー率を有する。このような高忠実度の複製でさえも、長さが3000 bpの遺伝子の場合、エキスピボで作用するポリメラーゼは、エラーをその回の約3%含んだコピーを産生する。現行の最良の商業的DNA合成プロトコルは数十年の開発の頂点に当たるので、ポリヌクレオチドの化学合成における大規模のさらなる改善が近い将来にやって来るといった可能性は低いように思われる。

20

【0009】

遺伝子およびゲノム合成技術の広範な利用は、原価高および高エラー率、ならびに自動化の不足などの制限によって阻まれている。カスタムポリヌクレオチドを合成する実用的、経済的な方法、大規模遺伝子システム、ならびに当該技術分野において知られる方法により作製された合成ポリヌクレオチドよりも低いエラー率を有する合成ポリヌクレオチドの産生方法が必要とされる。

【0010】

【特許文献1】米国特許第6,323,043号

【特許文献2】米国特許第5,405,783号

【特許文献3】WO 03/065038

【特許文献4】WO 03/064699

【特許文献5】WO 03/064026

【特許文献6】WO 02/04597

【特許文献7】米国特許第6,426,184号

【特許文献8】米国特許第2003/0054344号

【特許文献9】米国特許第6,093,302号

【特許文献10】米国特許第6,444,111号

【特許文献11】米国特許第6,280,595号

【非特許文献1】Zhou et al. (2004) *Nucleic Acids Res.* 32:5409

【非特許文献2】Fodor et al. (1991) *Science* 251:767

【非特許文献3】X. Gao et al., *Nucleic Acids Res.* 29: 4744-50 (2001)

【非特許文献4】X. Gao et al., *J. Am. Chem. Soc.* 120: 12698-12699 (1998)

【非特許文献5】O. Srivannavit et al., *Sensors and Actuators A.* 116: 150-160 (2004)

30

40

【発明の開示】

【0011】

概要

広くは、本発明は、個別にまたは一緒に利用できるMullis (Mullis et al. (1986) *Col d Spring Harb. Symp. Quant. Biol.* 51 Pt 1:263)およびStemmer (Stemmer et al. (199

50

5) Gene 164:49)のDNAアッセムブリ法に対し一群の改良を行うことにより、有用な高忠実度の合成DNA構築体の費用効果的な産生を可能にする。この改良には、アッセムブリに用いられるオリゴヌクレオチドのコンピュータによる設計の、すなわち「構築用オリゴヌクレオチド」および精製の場合、すなわち「選択用オリゴヌクレオチド」の設計の進歩、「構築用オリゴヌクレオチドアッセムブリの多重化」、すなわち同一プール中での多数の異なるアッセムブリの作製、構築用オリゴヌクレオチド増幅技術、ならびに構築用オリゴヌクレオチドのエラー低減技術が含まれる。

#### 【0012】

1つの態様では、本発明は、様々な段階でのオリゴヌクレオチドの増幅を含む所定の配列を有するポリヌクレオチド構築体の調製方法を提供する。この方法は、(i) ポリヌクレオチド構築体の配列を規定する (define) 部分的に重複する配列、(ii) 構築用オリゴヌクレオチドの少なくとも一部分で隣接するおよび構築用オリゴヌクレオチドの少なくともサブセットに共通する少なくとも1組のプライマーハイブリダイゼーション部位、ならびに(iii) プライマーハイブリダイゼーション部位と構築用オリゴヌクレオチドとの間の切断部位を有する構築用オリゴヌクレオチドのプールを提供する段階を含む。次に、プライマーハイブリダイゼーション部位に結合する少なくとも1つのプライマーを用いて、構築用オリゴヌクレオチドのプールを増幅することができる。任意で、プライマーハイブリダイゼーション部位はその後、構築用オリゴヌクレオチドから切断部位で(例えば、制限エンドヌクレアーゼ、化学的切断などを用いて)除去されてもよい。増幅後、例えば、オリゴヌクレオチドを変性させて相補鎖を分離することその後ハイブリダイゼーション条件ならびにライゲーションおよび/または鎖伸長条件に構築用オリゴヌクレオチドのプールを曝すことにより、構築用オリゴヌクレオチドをアッセムブリに供することができる。

#### 【0013】

別の態様では、本発明は構築用オリゴヌクレオチドの精製済みのプールを調製する方法を提供する。この方法は、ハイブリダイゼーション条件の下で構築用オリゴヌクレオチドのプールを選択用オリゴヌクレオチドのプールと接触させて二重鎖を形成させる段階を含む。この反応は安定な二重鎖(例えば、相補的な領域中に mismatches を含まない構築用オリゴヌクレオチドのコピーと選択用オリゴヌクレオチドのコピーとを含んだ二重鎖)と不安定な二重鎖(例えば、相補的な領域中に1つまたは複数の mismatches、例えば、塩基の mismatches、挿入もしくは欠失を含む構築用オリゴヌクレオチドのコピーと選択用オリゴヌクレオチドのコピーとを含んだ二重鎖)の両方を形成するはずである。その後、不安定な二重鎖を形成した構築用オリゴヌクレオチドのコピーをプールから除去して(例えば、カラムなどの分離技術により)、精製済みの構築用オリゴヌクレオチドのプールを形成させることができる。任意で、精製過程(例えば、構築および選択用オリゴヌクレオチドの混合)は、構築用オリゴヌクレオチドの使用の前に少なくとも1度繰り返されてもよい。さらに、構築用オリゴヌクレオチドのプールは、選択による各種精製ラウンドの前におよび/または後に増幅されてもよい。精製済みの構築用オリゴヌクレオチドのプールを形成させた後、それらプールをアッセムブリ条件に供することができる。例えば、構築用オリゴヌクレオチドのプールをハイブリダイゼーション条件ならびにライゲーションおよび/または鎖伸長条件に曝すことができる。

#### 【0014】

別の態様では、本発明は、単一のプール中で異なる所定の配列を有する複数のポリヌクレオチド構築体を調製する方法を提供する。この方法は、(i) 複数のポリヌクレオチド構築体の各々の配列を規定する部分的に重複する配列を含む構築用オリゴヌクレオチドのプールを提供する段階および(ii) 前記構築用オリゴヌクレオチドのプールをハイブリダイゼーション条件ならびにライゲーションおよび/または鎖伸長条件の下でインキュベートする段階を含む。任意で、オリゴヌクレオチドおよび/またはポリヌクレオチド構築体は必要に応じ、1ラウンドまたは複数ラウンドの増幅および/またはエラー低減に供されてもよい。さらに、ポリヌクレオチド構築体をさらなるラウンドのアッセムブリに供して、さらに長いポリヌクレオチド構築体を産生してもよい。少なくとも約2、4、5、10、50、100

、1,000またはそれ以上のポリヌクレオチド構築体を単一のプール中でアッセンブルすることができる。

【0015】

別の態様では、本発明は、構築および/または選択用オリゴヌクレオチドを設計する方法ならびに1つまたは複数のポリヌクレオチド構築体を産生するアッセンブリストラテジーを提供する。この方法は、例えば、(i) 各ポリヌクレオチド構築体の配列を部分的に重複する配列セグメントにコンピュータにより分割する段階；(ii) 部分的に重複する配列セグメントのセットに対応する配列を含む構築用オリゴヌクレオチドを合成する段階；および(iii) 前記構築用オリゴヌクレオチドをハイブリダイゼーション条件ならびにライゲーションおよび/または鎖伸長条件の下でインキュベートする段階を含むことができる。

任意で、この方法は(i) 構築用オリゴヌクレオチドの少なくとも一部分の末端に、前記構築用オリゴヌクレオチドの少なくともサブセットに共通するおよびプライマーハイブリダイゼーション部位と構築用オリゴヌクレオチドとの間の切断部位を規定する1組または複数組のプライマーハイブリダイゼーション部位をコンピュータにより付加する段階；(ii)

前記プライマーハイブリダイゼーション部位に結合する少なくとも1つのプライマーを用いて前記構築用オリゴヌクレオチドを増幅する段階；および(iii) 前記プライマーハイブリダイゼーション部位を前記切断部位で前記構築用オリゴヌクレオチドから除去する段階をさらに含んでもよい。好ましくは、そのようなプライマー部位はプール中の構築用オリゴヌクレオチドの少なくとも一部分に共通とすることができる。この方法は、構築用オリゴヌクレオチドの少なくとも一部分に相補的である配列を含む選択用オリゴヌクレオチドの少なくとも1つのプールをコンピュータにより設計する段階、前記選択用オリゴヌクレオチドを合成する段階、および選択用オリゴヌクレオチドのプールとの構築用オリゴヌクレオチドのプールのハイブリダイゼーションによってエラー低減過程を行う段階をさらに含むことができる。

【0016】

本発明の態様は同様に、単一のプール中で複数の異なるポリヌクレオチド配列をアッセンブルする方法に向けられる。これらの方法は、相補的な末端領域と異なるポリヌクレオチド配列の末端を含んだオリゴヌクレオチドに隣接するプライマー部位とを有する合成オリゴヌクレオチドの群を提供する段階、合成オリゴヌクレオチドをdNTPsおよびポリメラーゼとともに混合する段階、ならびに相補的な末端領域のハイブリダイゼーション、ポリメラーゼを介した塩基の取込みを誘導して重複するオリゴヌクレオチドを伸長させるようおよび完全長の異なるポリヌクレオチド配列のコピーや、複数の前記完全長配列の増幅をもたらすよう混合物をサイクリングにかける段階を含む。

【0017】

ある種の局面では、そのような方法は同様に、複数の別個のプールを利用し、異なる合成ポリヌクレオチド配列の少なくともいくつかをそれにより、相補的な末端領域とより大きなポリヌクレオチドの末端を含んだ異なるポリヌクレオチド配列に隣接するプライマー部位とを有するポリヌクレオチドを含む各プール中で産生させることを含む。複数のプールの少なくともいくつかをdNTPsおよびポリメラーゼとともに混合し、この混合物を異なるポリヌクレオチド配列の相補的な末端領域のハイブリダイゼーションを誘導するようサイクリングにかける。ポリメラーゼを介した塩基の取込みを利用して、重複するポリヌクレオチド配列を伸長し、完全長のより大きなポリヌクレオチドのコピーや、複数の前記完全長のより大きなポリヌクレオチドの増幅をもたらす。

【0018】

ある種の局面では、合成オリゴヌクレオチドを複数の塩基配列の連続自動並行アッセンブリにより並行して合成し精製(例えば、ハイブリダイゼーションによる精製)して、配列エラーを包含するオリゴヌクレオチドコピーの濃度を減らす。他の局面では、合成オリゴヌクレオチドを表面上で合成する。他の局面では、ほぼ同じ融解温度を有するように相補的な末端領域の複数のペアを設計する。他の局面では、プールはウェルまたはマイクロチャンネルである。他の局面では、混合する段階は微小流体システムの中に、ポリメラーゼが

10

20

30

40

50

熱安定性ポリマーゼである混合物の成分をとともに流すことで行われる。

【0019】

本発明の態様は、多数の異なる回収可能なポリヌクレオチドを含む製造物品に向けられる。この物品は、容器由来の異なるポリヌクレオチドの部分集団の増幅を可能にするプライマー配列の異なるペアを含んだ異なるポリヌクレオチド混合物を含むポリヌクレオチド用容器、および複数のプライマー用容器であって、それぞれが構築用容器中のポリヌクレオチドのプライマー配列のペアに相補的なオリゴヌクレオチドプライマーのペアを含んだ容器を含む。ポリヌクレオチド用容器中のポリヌクレオチドのプライマー配列ペアは、互いに異なることができる。ポリヌクレオチドは合成DNA、遺伝子、野生型配列の複数の変異体、ベクターおよび同様のものを含むことができる。ポリヌクレオチドの少なくとも一部分は、少なくとも1キロベース長である。ある種の局面では、ポリヌクレオチドの少なくとも一部分は、少なくとも2キロベース長、少なくとも5キロベース長、少なくとも10キロベース長、またはそれ以上の長さである。

10

【0020】

ある種の局面では、ポリヌクレオチドを環状化させることができる。ポリヌクレオチドを任意でアダプター配列に隣接させて、ベクターへの挿入、固定化または配列の機能の同定などの、ポリヌクレオチド配列の操作を容易にしてもよい。ポリヌクレオチドは、哺乳類の配列、酵母の配列、原核生物の配列、植物の配列、キイロショウジョウバエ (*D. melanogaster*) の配列、線虫 (*C. elegans*) の配列およびアフリカツメガエル (*Xenopus*) の配列からなる群より選択される1つまたは複数の配列を含むことができる。

20

【0021】

他の局面では、異なる回収可能なポリヌクレオチド構築体の混合物は、独立して回収可能である。例えば、製造物品は、複数の異なるポリヌクレオチドを含有する複数のポリヌクレオチド用容器、プライマー配列の同一のペアを含んだ異なる容器中のポリヌクレオチドを含んでもよく、その際に複数のプライマー用容器の1つまたは複数が相補的なオリゴヌクレオチドプライマーのペアを含む。ポリヌクレオチド用容器は、D個の異なる独立して回収可能なポリヌクレオチドであって、それぞれがN個のネスティッドプライマーペアを含むポリヌクレオチド、少なくとも  $N/2 \times D^{1/N}$  であるプライマー用容器の数を含むことができ、またはD個の異なるポリヌクレオチドおよびプライマーのペアを含有するD個のプライマー用容器を含むことができる。ポリヌクレオチド用容器は、プライマー配列の複数のネスティッドペアを含む異なるポリヌクレオチドであって、前記複数のネスティッドペアのそれぞれが前記容器中の選択のポリヌクレオチド群のまたはその容器中の前記異なるポリヌクレオチドの個々のものの増幅を可能にするポリヌクレオチドを含むことができる。製造物品は  $10^2$  個の異なるポリヌクレオチド、 $10^3$  個の異なるポリヌクレオチド、 $10^4$  個の異なるポリヌクレオチド、 $10^5$  個の異なるポリヌクレオチド、 $10^6$  個の異なるポリヌクレオチドまたはそれ以上を含むことができる。

30

【0022】

本発明の態様は、多数の異なる回収可能なポリヌクレオチドを含有する包装を含む製造物品にさらに向けられる。この物品は、異なるポリヌクレオチドの少なくともいくつかのプライマー配列の複数のネスティッドペアを含み、複数のネスティッドペアのそれぞれが容器中の選択のポリヌクレオチド群のまたはその容器中の前記異なるポリヌクレオチドの個々のものの増幅を可能にする異なるポリヌクレオチドの混合物を含んだポリヌクレオチド用容器を含む。物品は同様に、複数のプライマー用容器であって、それぞれが構築用容器中のポリヌクレオチドのプライマー配列のペアに相補的なオリゴヌクレオチドプライマーのペアを含んだ容器を含む。容器中の各ポリヌクレオチド上のネスティッドペアの組合せは、容器中のその他全てのポリヌクレオチドのネスティッドペアの組合せと異なってもよい。物品は、それぞれが複数の異なるポリヌクレオチドを含む複数の構築用容器、所与のプライマーペアが異なる容器中の異なるポリヌクレオチドとアニーリングするようにプライマー配列の同一のペアを含んだ異なる容器中のポリヌクレオチドを含むことができる。

40

50

## 【0023】

本発明の態様は同様に、ポリヌクレオチド構築体の選択の1つまたは選択の群に富む溶液を供給する装置に向けられる。この装置は、容器由来の異なるポリヌクレオチドのうち選択のものの増幅を可能にするおよび前記容器中のその他のポリヌクレオチドのプライマー配列のその他のペアとは異なる、プライマー配列の少なくとも1ペアを含む同定済みのポリヌクレオチドの混合物を含んだポリヌクレオチド用容器、ならびに複数のプライマー用容器であって、それぞれが構築用容器中の異なるポリヌクレオチドのプライマー配列のペアに相補的なオリゴヌクレオチドプライマーのペアを含んだ容器を含む。装置は同様に、同定済みのポリヌクレオチドと各同定済みのポリヌクレオチドに相補的なプライマーのペアまたは複数ペアを含む1つまたは複数の容器の位置とを収載するデータ保存庫、およびポリヌクレオチドまたはポリヌクレオチドの群を利用者が指定するのを可能にするインターフェースを含む。装置は、インターフェースで入力された仕様に応答する自動的手段、および指定されたポリヌクレオチドまたはポリヌクレオチドの群を選択的に増幅するのに必要とされる試薬を調製するよう構築用容器からポリヌクレオチドおよび選択のプライマー用容器からプライマーの一定分量を抽出するためにデータ保存庫からアクセスされる命令をさらに含む。

10

## 【0024】

ある種の局面では、装置は、異なる同定済みのポリヌクレオチドを含んだ複数のポリヌクレオチド用容器を含む。他の局面では、異なる容器中のポリヌクレオチドは、プライマー配列の同一ペアを含む。他の局面では、異なる容器中のポリヌクレオチドは、少なくとも10個のポリヌクレオチド用容器のものを含んだプライマー配列の複数のネスティッドペアを含む。他の局面では、異なる容器中のポリヌクレオチドは、プライマー配列の固有のネスティッドペアを含む。

20

## 【0025】

装置は、構築用容器から回収される選択の同定済みポリヌクレオチドを選択のプライマーペアによる指定のとおり増幅するよう適合された増幅用チャンバを含むことができる。他の局面では、装置は同様に、増幅用チャンバから回収される同定済みのポリヌクレオチドの1つまたは部分集団を選択のプライマーペアによる指定のとおり増幅するよう適合された第2の増幅用チャンバを含む。

## 【0026】

本発明の態様は同様に、選択のポリヌクレオチドを得る方法に向けられる。この方法は、容器由来のポリヌクレオチドのうち選択のものの増幅を可能にするプライマー配列の複数のネスティッドペアを含み、前記容器中のあるポリヌクレオチドのプライマーペアの組合せが前記容器中のその他のポリヌクレオチドのプライマー配列のその他のペアとは異なる同定済みの合成ポリヌクレオチドの混合物を含有する複数の構築用容器を提供する段階を含む。次に、複数のプライマー用容器であって、それぞれが構築用容器中のポリヌクレオチドのプライマー配列のペアに相補的なオリゴヌクレオチドプライマーのペアを含む容器を提供する。第1の増幅手順は、選択の構築用容器から回収されるポリヌクレオチドの混合物と1つまたは複数のプライマー用容器から回収される、プライマー配列の外側ネスティッドペアに相補的なプライマーのペアの一定分量を含む第1の増幅混合物で行う。第2の増幅手順は、第1の増幅混合物から回収される単位複製配列と1つまたは複数のプライマー用容器から回収される、プライマー配列の内側ネスティッドペアに相補的なプライマーのペアの一定分量を含む第2の増幅混合物で行う。

30

40

## 【0027】

本発明の態様は同様に、ライブラリーを形成する混合物中の多数の合成ポリヌクレオチドに向けられる。このライブラリーは、多数のポリヌクレオチド種であって、少なくともいくつかはライブラリーから回収される種のうちの選択群の増加を可能とするのに十分な長さのプライマー配列の外側ペアを有する種を含む。ライブラリーは同様に、外側ペアを用いた増幅によって産生された単位複製配列の混合物から回収される種のうちの1つまたは選択群の増幅を可能とするのに十分な長さを有するプライマー配列の内側ペアを含む。

50

ある種の局面では、ライブラリー中の個々の種の濃度は、ライブラリーから直接的にその選択的増幅を可能とするには十分ではないが、外側のプライマー配列ペアを用いた増幅後にその選択的増幅を可能とするには十分である。別の局面では、合成ポリヌクレオチドは、プライマー配列の3組のネスティッドペアを含む。別の局面では、合成ポリヌクレオチドは各々、ライブラリー中のプライマー配列のその他全てのネスティッドペアとは異なる核酸配列を有するプライマー配列のネスティッドペアを含む。

【0028】

本明細書に記述される方法は同様に、機能的なスクリーニングおよび選択に向けた変異体配列のライブラリーを作製するのに有用である。

【0029】

本発明の前述のおよびその他の特徴および利点は、添付の図面とともに以下の例示的態様の詳細な説明からもっと十分に理解されると思われる。

【0030】

詳細な説明

本発明は、カスタムポリヌクレオチドを合成する経済的な方法、ならびに当技術分野において知られる方法により作製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よりも低いミスマッチエラー率を有する合成オリゴヌクレオチドおよび/またはポリヌクレオチドを作製する方法を提供する。

【0031】

当技術分野において知られる方法と比べての本明細書に記述される方法の大きな進歩の1つは、表面オリゴヌクレオチドアレイ合成から得られるわずかな分子を利用する能力である。本明細書において提供される方法では、反応物質が低濃度で存在する場合の二分子相互作用の速度式を向上させるため、さらに2通りのストラテジーを利用する。1つの態様では、本発明は、高濃度の「ユニバーサル」プライマーを用いて1種または複数種のオリゴヌクレオチドを予め増幅する方法を提供する。別の態様では、本発明は、合成時に初めのうち高濃度のオリゴヌクレオチドを利用する方法を提供する。

【0032】

本明細書で用いられる、以下の用語および語句は下記の意味を有するものとする。他に特に規定がなければ、本明細書で用いられる全ての技術用語および科学用語は、当業者に共通して理解されるのと同じ意味を有する。

【0033】

単数形「a(1つの)」、「an(1つの)」および「the(その)」は、文脈により他に明記されていないならば複数対象を含む。

【0034】

「増幅」という用語は核酸断片のコピー数が増やされることを意味する。

【0035】

「塩基対合」という用語は、例えば、アデニン(A)とチミン(T)、グアニン(G)とシトシン(C)、(A)とウラシル(U)、およびグアニン(G)とシトシン(C)、ならびにそれらの相補体を含む、二本鎖核酸中のプリンとピリミジンとの間の特定の水素結合のことをいう。塩基対合は2本の相補的な一本鎖からの核酸二重らせんの形成をもたらす。

【0036】

本明細書で用いられる「切断」という用語は、ホスホジエステル結合などの、2本のヌクレオチド間の結合の切断のことをいう。

【0037】

「comprise(含む)」および「comprising(含む)」という用語は、さらなる要素が含まれるという包含的な広い意味で使われる。

【0038】

「構築用オリゴヌクレオチド」という用語は、オリゴヌクレオチドそれ自体よりも長い核酸分子をアッセンブルするのに使用できる一本鎖オリゴヌクレオチドのことをいう。典型的な態様では、構築用オリゴヌクレオチドは、構築用オリゴヌクレオチドよりも少なく

10

20

30

40

50

とも約3倍、4倍、5倍、10倍、20倍、50倍、100倍、またはそれ以上長い核酸分子をアッセンブルするのに使用できる。通常、所定の配列を有する異なった構築用オリゴヌクレオチドのセットが所望の配列を有するいっそう長い核酸分子へのアッセンブリに使われると考えられる。典型的な態様では、構築用オリゴヌクレオチドは、長さが約25から約200、約50から約150、約50から約100、または約50から約75ヌクレオチドとすることができる。構築用オリゴヌクレオチドのアッセンブリは、例えば、PAM、PCRアッセンブリ、ライゲーション連鎖反応、ライゲーション/融合PCR、二重非対称PCR、重複伸長PCR、およびそれらの組合せを含む、さまざまな方法により行うことができる。構築用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とすることができる。典型的な態様では、構築用オリゴヌクレオチドは、基板上に並行して合成された合成オリゴヌクレオチドである。構築用オリゴヌクレオチドの配列設計は、例えば、DNAWorks (Hoover and Lubkowski, *Nucleic Acids Res.* 30: e43 (2002)、Gene2Oligo (Rouillard et al., *Nucleic Acids Res.* 32: W176-180 (2004)およびberry.engin.umich.edu/gene2oligoのワールドワイドウェブ)などのコンピュータプログラム、または以下でさらに論じられる実装システムや方法を活用して行うことができる。

10

**【0039】**

「dam」という用語は、DNA複製開始、DNAミスマッチ修復およびいくつかの遺伝子の発現の調節を調整する役割を果たすアデニンメチル基転移酵素のことをいう。この用語は原核生物のdamタンパク質ならびにそのホモログ、オルソログ、パラログ、変異体、または断片を包含するよう意図される。典型的なdamタンパク質としては、例えば、以下のGenBankアクセッション番号AF091142 (髄膜炎菌(*Neisseria meningitidis*)株BF13)、AF006263 (梅毒トレポネーマ(*Treponema pallidum*))、U76993 (ネズミチフス菌(*Salmonella typhimurium*))およびM22342 (バクテリオファージT2)を有する核酸によってコードされるポリペプチドが挙げられる。

20

**【0040】**

「変性する(させる)」または「融解する(させる)」という用語は、二重核酸分子の鎖が一本鎖分子に分離される過程のことをいう。変性の方法としては、例えば、熱変性およびアルカリ変性が挙げられる。

**【0041】**

「検出可能なマーカー」という用語は、ポリヌクレオチド配列であって、この配列を持つ細胞の同定を容易にするポリヌクレオチド配列のことをいう。ある種の態様では、検出可能なマーカーは、例えば、緑色蛍光タンパク質(GFP)、強化緑色蛍光タンパク質(EGFP)、レニラ・レニフォルミス(*Renilla Reniformis*)由来の緑色蛍光タンパク質、GFPmut2、GFPuv4、強化黄色蛍光タンパク質(EYFP)、強化シアン蛍光タンパク質(ECFP)、強化青色蛍光タンパク質(EBFP)、シトリンおよびイソギンチャク(*discosoma*)由来の赤色蛍光タンパク質(dsRED)などの、化学発光または蛍光タンパク質をコードする。その他の態様では、検出可能なマーカーは、例えば、ポリHisタグ、myc、HA、GST、プロテインA、プロテインG、カルモジュリン結合ペプチド、チオレドキシン、マルトース結合タンパク質、ポリアルギニン、ポリHis-Asp、FLAG、および同様のものなどの抗原性またはアフィニティータグであってもよい。

30

40

**【0042】**

「二重鎖」という用語は、少なくとも部分的に二本鎖である核酸分子のことをいう。「安定な二重鎖」とは、一定のハイブリダイゼーション条件の下で相補配列にハイブリダイズされたままである可能性が相対的により高い二重鎖のことをいう。典型的な態様では、安定な二重鎖とは、塩基対のミスマッチ、挿入、または欠失を含まない二重鎖のことをいう。「不安定な二重鎖」とは、一定のハイブリダイゼーション条件の下で相補配列にハイブリダイズされたままである可能性が相対的により低い二重鎖のことをいう。典型的な態様では、不安定な二重鎖とは、少なくとも1つの塩基対のミスマッチ、挿入、または欠失を含む二重鎖のことをいう。

**【0043】**

50

「エラー低減」という用語は、核酸分子、または核酸分子のプール中での配列エラー数を減らし、それによって核酸分子の組成物中でのエラーなしのコピー数を増やすよう利用できる過程のことをいう。エラー低減はエラーろ過、エラー中和およびエラー補正過程を含む。「エラーろ過」とは、核酸分子のプールから配列エラーを含む核酸分子が除去される過程である。エラーろ過を行う方法としては、例えば、選択用オリゴヌクレオチドとのハイブリダイゼーション、またはミスマッチ結合剤との結合、その後分離が挙げられる。「エラー中和」とは、配列エラーを含む核酸は増幅することおよび/またはアッセムブルすることを制限されるが、核酸のプールから除去はされない過程である。エラー中和の方法としては、例えば、ミスマッチ結合剤との結合および任意でDNA二重鎖とのミスマッチ結合剤の共有結合が挙げられる。「エラー補正」とは、核酸分子中の配列エラーが補正される(例えば、特定位置の不正確なヌクレオチドが既定の配列に基づき存在するはずの核酸に変えられる)過程である。エラー補正の方法としては、例えば、相同組換えまたはDNA修復タンパク質を使った配列補正が挙げられる。

10

## 【0044】

「遺伝子」という用語は、エキソン配列および任意でイントロン配列を有するポリペプチドをコードする読み取り枠を含んだ核酸のことをいう。「イントロン」という用語は、ある遺伝子に存在するDNA配列であって、タンパク質に翻訳されず、一般にエキソン間に見出されるDNA配列のことをいう。

## 【0045】

「ハイブリダイズする」または「ハイブリダイゼーション」という用語は、2本の相補的な核酸鎖間の特異的結合のことをいう。種々の態様では、ハイブリダイゼーションとは、2本の核酸鎖のうち完全に適合する相補領域間の会合および相補領域中に1つまたは複数のミスマッチ(ミスマッチ、挿入、または欠失を含め)を含む2本の核酸鎖間の結合のことをいう。ハイブリダイゼーションは、例えば、1つ、2つ、3つ、4つ、5つまたはそれ以上のミスマッチを含む2本の相補核酸鎖の間で起こりうる。種々の態様では、ハイブリダイゼーションは、例えば、部分的に重複するオリゴヌクレオチドと相補的な構築用オリゴヌクレオチドとの間で、部分的に重複するオリゴヌクレオチドと相補的な構築および選択用オリゴヌクレオチドとの間で、プライマーとプライマー結合部位との間などで起こりうる。2本の核酸鎖間のハイブリダイゼーションの安定性は、例えば温度および/または塩濃度を含め、ハイブリダイゼーション条件および/または洗浄条件を変化させることで制御

20

30

## 【0046】

「を含む(including)」という用語は「を含むがこれ(ら)に限定され(ることは)ない(including but not limited to)」を意味するように使われる。「を含む(including)」および「を含むがこれ(ら)に限定され(ることは)ない(including but not limited to)」は同義的に使われる。

## 【0047】

「リガーゼ」という用語は、同じオリゴヌクレオチドにアニーリングしている隣接のオリゴヌクレオチドにホスホジエステル結合を形成する際の酵素類およびその機能のことをいう。1つのオリゴヌクレオチドの末端リン酸基ともう1つの隣接オリゴヌクレオチドの末端ヒドロキシル基が二重らせん内でその相補配列の向かいに共にアニーリングされる場合に、すなわち、ライゲーション過程が連結可能なニック部位の「ニック」を連結し、相補的な二重鎖を作製する場合に、特に効率的なライゲーションが行われる(Blackburn, M. and Gait, M. (1996) *Nucleic Acids in Chemistry and Biology*, Oxford University Press, Oxford, pp. 132-33, 481-2中)。隣接するオリゴヌクレオチド間の部位は「連結可能なニック部位」、「ニック部位」または「ニック」と称されており、それらによってホスホジエステル結合は存在していないか、または切断されている。

40

50

## 【0048】

「連結する」という用語は、ヌクレオチド間結合の形成を通じて隣接するオリゴヌクレオチドを共有結合的に結合する反応のことをいう。

## 【0049】

「選択可能なマーカー」という用語は、ポリヌクレオチド配列であって、選択可能なマーカーを欠いた類似細胞に比べて、そのポリヌクレオチド配列を持った細胞の所与の増殖環境での増殖能または生存能を変化させる遺伝子産物をコードするポリヌクレオチド配列のことをいう。そのようなマーカーは、陽性または陰性の選択可能なマーカーとすることができる。例えば、陽性の選択可能なマーカー(例えば、抗生物質耐性または栄養要求性増殖遺伝子)は、選択培地(例えば、抗生物質を含むまたは必須栄養素を欠く)中での増殖能または生存能を与える産物をコードする。陰性の選択可能なマーカーは、対照的に、ポリヌクレオチドを持っていない細胞に比べて、ポリヌクレオチドを持った細胞が陰性の選択培地中で増殖するのを妨げる。選択可能なマーカーは細胞を増殖させるのに使われる培地に応じて、陽性および陰性の両選択可能性を与えてもよい。選択可能なマーカーを原核細胞および真核細胞で使用することは、当業者によってよく知られている。適当な陽性選択マーカーとしては、例えば、ネオマイシン、カナマイシン、hyg、hisD、gpt、プレオマイシン、テトラサイクリン、hprt、SacB、 $\beta$ -ラクタマーゼ、ura3、アンピシリン、カルベニシリン、クロラムフェニコール、ストレプトマイシン、ゲンタマイシン、フレオマイシン、およびナリジクス酸が挙げられる。適当な陰性選択マーカーとしては、例えば、hs v-tk、hprt、gpt、およびシトシンデアミナーゼが挙げられる。

10

20

## 【0050】

「選択用オリゴヌクレオチド」という用語は、構築用オリゴヌクレオチド(または構築用オリゴヌクレオチドの相補体)の少なくとも一部分に相補的である一本鎖オリゴヌクレオチドのことをいう。選択用オリゴヌクレオチドは、配列決定のエラー(例えば、所望の配列からのずれ)を含む構築用オリゴヌクレオチドのコピーを構築用オリゴヌクレオチドのプールから除去する方法で使われてもよい。典型的な態様では、選択用オリゴヌクレオチドは基板上に末端で固定化することができる。1つの態様では、選択用オリゴヌクレオチドは、基板上に並行して合成された合成オリゴヌクレオチドである。選択用オリゴヌクレオチドは構築用オリゴヌクレオチド(または構築用オリゴヌクレオチドの相補体)の全長の少なくとも約20%、25%、30%、50%、60%、70%、80%、90%、または100%に相補的とすることができる。典型的な態様では、選択用オリゴヌクレオチドのプールは複数の構築/選択用オリゴヌクレオチド対の融解温度( $T_m$ )が実質的に同様であるように設計される。1つの態様では、選択用オリゴヌクレオチドのプールは実質的に全ての構築/選択用オリゴヌクレオチド対の融解温度が実質的に同様であるように設計される。例えば、構築/選択用オリゴヌクレオチド対の少なくとも約50%、60%、70%、75%、80%、90%、95%、97%、98%、99%、またはそれ以上の融解温度が相互の約10、7、5、4、3、2、1、またはそれ未満の範囲内である。選択用オリゴヌクレオチドの配列設計は、例えば、DNAWorks (Hoover and Lubkowski, *Nucleic Acids Res.* 30: e43 (2002)、Gene2oligo (Rouillard et al., *Nucleic Acids Res.* 32: W176-180 (2004)およびberry.engin.umich.edu/gene2oligoのワールドワイドウェブ)などのコンピュータプログラム、または以下でさらに論じられる実装システムや方法を活用して行うことができる。

30

40

## 【0051】

「ストリンジェントな条件」または「ストリンジェントなハイブリダイゼーション条件」という用語は、二重鎖を形成するように2本の相補的なポリヌクレオチド鎖間の特異的ハイブリダイゼーションを促進する条件のことをいう。ストリンジェントな条件は、規定のイオン強度およびpHでの所与のポリヌクレオチド二重鎖に対する熱融解温度( $T_m$ )よりも約5%低いように選択することができる。相補的なポリヌクレオチド鎖の長さおよびそのGC含量によって、二重鎖の $T_m$ 、したがって所望のハイブリダイゼーション特異性を得るために必要なハイブリダイゼーション条件を判断できるはずである。 $T_m$ とは、完全に適合する相補鎖にポリヌクレオチド配列の50%がハイブリダイズする温度(規定のイオン強度およ

50

びpHの下)である。ある種の場合には、特定の二重鎖の $T_m$ にはほぼ等しくなるようにハイブリダイゼーション条件のストリンジェンシーを高くすることが望ましいかもしれない。

#### 【0052】

$T_m$ を推定するための様々な技術を利用することができる。通常、約80~100の理論上の最大値まで、二重鎖中のG-C塩基対は $T_m$ に約3寄与すると推定され、その一方でA-T塩基対は約2寄与するものと推定される。しかしながら、G-Cスタッキング相互作用、溶媒効果、所望のアッセイ温度などを考慮に入れるもっと精緻な $T_m$ モデルを利用することができる。例えば、以下の式を用いて、およそ60の解離温度( $T_d$ )を有するようにプローブを設計することができる： $T_d = (((((3 \times \#GC) + (2 \times \#AT)) \times 37) - 562) / \#bp) - 5$ ; 式中で $\#GC$ 、 $\#AT$ 、および $\#bp$ は、それぞれ、二重鎖の形成に関与する、グアニン-シトシン塩基対の数、アデニン-チミン塩基対の数および全塩基対の数である。 $C_T$ は全モル鎖濃度であり、 $R$ は気体定数1.9872 cal/K-molであり、および $x$ は非自己相補的な二重鎖の場合には4に等しく、自己相補的な二重鎖の場合には1に等しい、式 $T_m = H^0 \times 1000 / (S^0 + R \times \ln(C_T/x)) - 273.15$ を用いる、 $T_m$ を算出するその他の方法がSantaLucia and Hicks, *Ann. Rev. Biomol. Struct.* 33: 415-40 (2004)に記述されている。

10

20

30

#### 【0053】

ハイブリダイゼーションは5×SSC、4×SSC、3×SSC、2×SSC、1×SSCまたは0.2×SSC中で少なくとも約1時間、2時間、5時間、12時間、または24時間行うことができる。ハイブリダイゼーションの温度を、例えば、約25(室温)から約45、50、55、60、または65まで高くして反応のストリンジェンシーを調整してもよい。ハイブリダイゼーション反応には同様に、ストリンジェンシーに影響を与える別の作用物質が含まれてもよく、例えば、50%ホルムアミドの存在下で行われるハイブリダイゼーションでは、規定の温度でのハイブリダイゼーションのストリンジェンシーが高くなる。典型的な態様では、ベタイン、例えば、約5 Mベタインをハイブリダイゼーション反応に加えて、DNA熱融解転移の塩基対組成依存性を最小限に抑えてもまたは取り除いてもよい(例えば、Rees et al., *Biochemistry* 32: 137-144 (1993)を参照のこと)。別の態様では、低分子量アミドまたは低分子量スルホン(例えば、DMSO、テトラメチレンスルホキシド、メチルsec-ブチルスルホキシドなどのような)をハイブリダイゼーション反応に加えて、GC含量が豊富な配列の融解温度を低下させてもよい(例えば、Chakarbarti and Schutt, *BioTechniques* 32: 866-874 (2002)を参照のこと)。

#### 【0054】

ハイブリダイゼーション反応に続けて単回の洗浄ステップ、または2回もしくはそれ以上の回数の洗浄ステップが行われてもよく、このステップは同じまたは異なる塩度および温度であってもよい。例えば、洗浄の温度を約25(室温)から約45、50、55、60、65またはそれ以上まで高くしてストリンジェンシーを調整してもよい。洗浄ステップは界面活性剤、例えば、0.1%または0.2% SDSの存在下で行われてもよい。例えば、ハイブリダイゼーションに続けて各2×SSC、0.1% SDS中で約20分間の65の洗浄ステップが2回、および任意で各0.2×SSC、0.1% SDS中で約20分間の65の洗浄ステップがさらに2回行われてもよい。

#### 【0055】

典型的なストリンジェントなハイブリダイゼーション条件としては50%ホルムアミド、10×デンハルト(0.2% Ficoll、0.2%ポリビニルピロリドン、0.2%ウシ血清アルブミン)および200 μg/mlの変性済みキャリアDNA、例えば、剪断サケ精子DNAを含む、またはそれらからなる溶液中65で一晩のハイブリダイゼーション、続けて各2×SSC、0.1% SDS中で約20分間の65の洗浄ステップ2回、および各0.2×SSC、0.1% SDS中で約20分間の65の洗浄ステップ2回を行うことが挙げられる。

40

#### 【0056】

ハイブリダイゼーションは溶液中の核酸2つをハイブリダイズさせるか、または溶液中の核酸1つを固体支持体、例えば、フィルタに付着している核酸1つにハイブリダイズさせることからなってもよい。一方の核酸が固体支持体上にある場合、プレハイブリダイゼー

50

ションステップがハイブリダイゼーションの前に行われてもよい。プレハイブリダイゼーションはハイブリダイゼーション溶液と同じ溶液中でおよび同じ温度で少なくとも約1時間、3時間または10時間行われてもよい(相補的なポリヌクレオチド鎖がなければ)。

【0057】

適切なストリンジェンシー条件は当業者に知られており、または当業者によって実験的に決められてもよい。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 (編) Methods in Molecular Biology, volume 20; Tijssen (1993) Laboratory Techniques in biochemistry and molecular biology-hybridization with nucleic acid probes、例えば、part I chapter 2「Overview of principles of hybridization and the strategy of nucleic acid probe assays」、Elsevier, New York; Tibanyenda, N. et al., Eur. J. Biochem. 139: 19(1984)およびEbel, S. et al., Biochem. 31:12083 (1992); Rees et al., Biochemistry 32: 137-144 (1993); Chakrabarti and Schutt, BioTechniques 32:866-874 (2002); ならびにSantaLucia and Hicks, Annu. Rev. Biomol. Struct. 33: 415-40 (2004)を参照されたい。

10

【0058】

タンパク質に適用される場合、「実質的な同一性」という用語は、2つの配列が、例えば、デフォルトのギャップ重み付けを用いるGAPまたはBESTFITプログラムによって最適に整列された際に、典型的には少なくとも約70パーセントの配列同一性、あるいは少なくとも約80、85、90、95パーセントの配列同一性またはそれ以上を共有することを意味する。アミノ酸配列の場合、同一ではないアミノ酸残基は、上述されている同類アミノ酸置換によって異なってもよい。

20

【0059】

「サブアッセムブリ」という用語は、構築用オリゴヌクレオチドのセットからアッセムブルされた核酸分子のことをいう。サブアッセムブリは構築用オリゴヌクレオチドよりも少なくとも約3倍、4倍、5倍、10倍、20倍、50倍、100倍、またはそれ以上長いこと、例えば、300~600塩基長であることが好ましい。

【0060】

核酸分子に関連して本明細書で用いられる「合成の」という用語は、インビトロの化学的および/または酵素的合成による作製のことをいう。

30

【0061】

「転写調節配列」とは、これが作動可能に連結されるタンパク質コード配列の転写を誘導するまたは制御する、開始シグナル、エンハンサーおよびプロモーターなどの、DNA配列のことを指して本明細書で用いられる総称である。好ましい態様では、組換え遺伝子の1つの転写は、発現が意図される細胞種において組換え遺伝子の発現を制御するプロモーター配列(またはその他の転写調節配列)の制御下にある。組換え遺伝子は、本明細書に記述される天然型の遺伝子の転写を制御する配列と同じであるかまたはその配列と異なる転写調節配列の制御下とできることも理解されると思われる。

【0062】

本明細書で用いられる「トランスフェクション」という用語は、受容細胞への核酸、例えば、発現ベクターの導入を意味し、ウイルスまたはウイルスベクターについて「感染させる」などのよく使われる用語を含むよう意図される。「形質導入」という用語は、核酸のトランスフェクションが核酸のウイルス送達によるものである場合に本明細書で一般に使用される。「形質転換」という用語は、DNAなどの外来分子を細胞に導入する任意の方法のことをいう。リポフェクション、DEAE-デキストランを介したトランスフェクション、マイクロインジェクション、プロトプラスト融合、リン酸カルシウム沈殿、レトロウイルス送達、エレクトロポレーション、自然形質転換、およびバイオリスティック形質転換は、当業者に公知の、利用できるほんの一握りの方法である。

40

【0063】

50

「ユニバーサルプライマー」という用語は、複数のポリヌクレオチドの鎖伸長/増幅に使用できるプライマーのセット(例えば、フォワードおよびリバースプライマー)のことをいい、例えば、このプライマーは、複数のポリヌクレオチドに共通している部位にハイブリダイズする。例えば、ユニバーサルプライマーは、例えば、構築用オリゴヌクレオチドのプール、選択用オリゴヌクレオチドのプール、サブアッセムブリのプール、および/またはポリヌクレオチド構築体のプールなどのような、単一のプール中の全ての、または本質的に全てのポリヌクレオチドの増幅に使われてもよい。1つの態様では、単一のプライマーを使用して、単一のプール中において複数のポリヌクレオチドのフォワードおよびリバース両鎖を増幅してもよい。ある種の態様では、ユニバーサルプライマーは、酵素的または化学的切断を介して増幅後に除去できる一過性プライマーであってもよい。その他の態様では、ユニバーサルプライマーは、鎖伸長によってポリヌクレオチド分子に組み入れられるようになる修飾を含んでもよい。典型的な修飾としては、例えば、3'もしくは5'末端キャップ、標識(例えば、フルオレセイン)、またはタグ(例えば、ビオチンなどのような、ポリヌクレオチドの固定化または単離を容易にするタグ)が挙げられる。

10

**【0064】**

「ベクター」とは、挿入された核酸分子を宿主細胞中におよび/または宿主細胞間に移入する自己複製核酸分子である。この用語には、細胞への核酸分子の挿入で主に機能するベクター、核酸の複製で主に機能する複製ベクター、ならびにDNAまたはRNAの転写および/または翻訳で機能する発現ベクターが含まれる。前記の機能のうち2つ以上を供与するベクターも含まれる。本明細書で用いられる「発現ベクター」とは、適切な宿主細胞中に導入されると、転写されポリペプチドに翻訳されるポリヌクレオチドと同定される。「発現系」とは、通常、所望の発現産物を産生するように機能できる発現ベクターからなる適当な宿主細胞のことを意味する。

20

**【0065】**

本発明の態様は、構築用オリゴヌクレオチドおよび選択用オリゴヌクレオチドなどの合成オリゴヌクレオチド配列を作製し増幅する方法に向けられる。本明細書で用いられる「オリゴヌクレオチド」という用語は、合成手段によって通常調製される一本鎖DNAまたはRNA分子を含むよう意図されるが、これらに限定されることはない。本発明のヌクレオチドは、通常、アデノシン、グアノシン、ウリジン、シチジンおよびチミジン由来のヌクレオチドなどの天然に存在するヌクレオチドであると考えられる。オリゴヌクレオチドが「二本鎖(の)」と称される場合、オリゴヌクレオチドのペアが、例えば、DNAと通常結合される水素結合らせん状アレイに存在することが当業者によって理解される二本鎖オリゴヌクレオチドの100%相補的な形態に加えて、本明細書で用いられる「二本鎖(の)」という用語は同様に、バルジおよびループのような構造的特徴を含んだ形態を含むよう意図される(あらゆる目的でその全体が参照により本明細書に組み入れられるStryer, Biochemistry, Third Ed. (1988)を参照のこと)。本明細書で用いられる「ポリヌクレオチド」という用語は、ともに結合した(例えば、ハイブリダイゼーション、ライゲーション、重合および同様のものにより)2つまたはそれ以上のオリゴヌクレオチドを含むよう意図されるが、これに限定されることはない。

30

**【0066】**

「作動可能に連結される」という用語は、2つの核酸領域間の関係を記述する場合、それらの領域がその所期の通りそれらを機能させる関係にある並置のことをいう。例えば、コード配列に「作動可能に連結される」制御配列は、適切な分子(例えば、インデューサーおよびポリメラーゼ)が制御または調節配列に結合される場合のような、制御配列に適合する条件の下でコード配列の発現が達成されるように連結される。

40

**【0067】**

「同一性の割合」という用語は、2つのアミノ酸配列間のまたは2つのヌクレオチド配列間の配列同一性のことをいう。比較の目的で並べられうる各配列中の位置を比較することで、同一性をそれぞれ判定することができる。比較配列中の等価の位置が同じ塩基またはアミノ酸によって占有されているなら、その位置で分子は同一であり; 等価の部位が同じ

50

または類似のアミノ酸残基(例えば、立体的および/または電氣的性質が類似)によって占有されているなら、その位置で分子は相同(類似)ということができる。相同性、類似性または同一性の割合と表現することは、比較配列によって共有される位置での同一または類似のアミノ酸の数の関数のことをいう。FASTA、BLASTまたはENTREZを含めて、さまざまなアライメント・アルゴリズムおよび/またはプログラムが使われてもよい。FASTAおよびBLASTはGCG配列解析パッケージ(ウィスコンシン大学、Madison, Wis.)の一部として利用可能であり、例えば、デフォルト設定で使用することができる。ENTREZは全米バイオテクノロジー情報センター(National Center for Biotechnology Information)、米国立医学図書館(National Library of Medicine)、米国立衛生研究所(National Institutes of Health)、Bethesda, MDを通じて利用可能である。1つの態様では、2つの配列の同一性の割合は、1のギャップ重み付けを用いてGCGプログラムにより決定することができ、例えば、各アミノ酸のギャップは、それが2つの配列間の単一のアミノ酸またはヌクレオチドのミスマッチであるかのように重み付けされる。

10

## 【0068】

アライメントのための他の技術はMethods in Enzymology, vol. 266: Computer Methods for Macromolecular Sequence Analysis (1996), ed. Doolittle, Academic Press, Inc., a division of Harcourt Brace & Co., San Diego, California, USAに記述されている。配列中のギャップを許容するアライメントプログラムを利用して、配列を整列させることが好ましい。Smith-Watermanは配列アライメント中のギャップを許容するアルゴリズムの一種である。Meth. Mol. Biol. 70: 173-187 (1997)を参照されたい。同様に、NeedlemanおよびWunschアライメント法を使ったGAPプログラムを利用して、配列を整列させることができる。別の検索方法では、MASPARコンピュータで作動するMPSRCHソフトウェアを利用する。MPSRCHはSmith-Watermanアルゴリズムを使って、大規模並列処理コンピュータで配列をスコア化する。この手法では遠縁の適合を拾い上げる能力が向上しており、とりわけ小さなギャップやヌクレオチド配列のエラーが許容される。核酸にコードされたアミノ酸配列を利用して、タンパク質およびDNAの両データベースを検索することができる。

20

## 【0069】

「ポリヌクレオチド構築体」という用語は、所定の配列を有する長い核酸分子のことをいう。ポリヌクレオチド構築体は構築用オリゴヌクレオチドのセットおよび/またはサブアッセムブリのセットからアッセムブルされてもよい。

30

## 【0070】

「制限エンドヌクレアーゼ認識部位」という用語は、1つまたは複数の制限エンドヌクレアーゼを結合できる核酸配列のことをいう。「制限エンドヌクレアーゼ切断部位」という用語は、1つまたは複数の制限エンドヌクレアーゼによって切断される核酸配列のことをいう。ある酵素に対し、制限エンドヌクレアーゼ認識部位および切断部位は同じであってもまたは異なってもよい。制限酵素はI型酵素、II型酵素、IIS型酵素、III型酵素およびIV型酵素を含むが、これらに限定されることはない。

## 【0071】

本発明のある種の局面では、分子の塩基部分が糖部分のいずれかに保護基を有する、あるいは付着されたもしくは組み入れられた標識、または人工環境か生理環境のいずれかで親単量体と同じようにふるまう単量体をもたらす等配電子置換を有するヌクレオシドまたはヌクレオチドなどの、ヌクレオチド類似体または誘導体が使われると考えられる。このヌクレオチドは、ヌクレオチドの反応基に結合されて、その基をマスクする保護基を有することができる。さまざまな保護基が本発明で有用であり、利用される合成法に応じて選択されてもよく、以下でさらに論じられる。ヌクレオチドを支持体または成長中の核酸に付着させた後に、保護基を取り除くことができる。

40

## 【0072】

本明細書で用いられる「構築用オリゴヌクレオチド」という用語は、標的核酸配列(例えば、遺伝子)またはその一部分に同一であるかまたは相補的であるオリゴヌクレオチド配列を含むよう意図されるが、これに限定されることはない。

50

## 【0073】

本明細書で用いられる「選択用オリゴヌクレオチド」という用語は、構築用オリゴヌクレオチドの少なくとも一部分に相補的であり、配列特異的にその部分にハイブリダイズできるオリゴヌクレオチド配列を含むよう意図されるが、これに限定されることはない。

## 【0074】

オリゴヌクレオチドまたはその断片は、天然供給源から単離されてもよく、または商業的供給源から購入されてもよい。オリゴヌクレオチド配列は、任意の適当な方法、例えば、どちらもあらゆる目的でその全体が参照により本明細書に組み入れられるBeaucageおよびCarruthers ((1981) *Tetrahedron Lett.* 22: 1859)によって報告されているホスホラミダイト法またはMatteucciら ((1981) *J. Am. Chem. Soc.* 103: 3185)によるトリエステル法により、あるいは本明細書に記述されており当技術分野において知られている高処理・高密度アレイ法または市販の自動オリゴヌクレオチド合成機法のいずれかを用いたその他の化学的方法(あらゆる目的でその全体が参照により本明細書に組み入れられる米国特許第5,602,244号、同第5,574,146号、同第5,554,744号、同第5,428,148号、同第5,264,566号、同第5,141,813号、同第5,959,463号、同第4,861,571号および同第4,659,774号を参照のこと)により調製されてもよい。予め合成されたオリゴヌクレオチドおよびオリゴヌクレオチドを含むチップを様々な業者から商業的に得ることもできる。

## 【0075】

種々の態様では、本明細書に記述される方法は構築および/または選択用オリゴヌクレオチドを利用する。構築および/または選択用オリゴヌクレオチドの配列は、合成されることが望まれる最終のポリヌクレオチド構築体の配列に基づいて決定されると思われる。本質的にポリヌクレオチド構築体の配列は複数の重複するいっそう短い配列に分けられてもよく、これを本明細書に記述される方法により並行して合成し、最終の望ましいポリヌクレオチド構築体にアッセンブルすることができる。構築および/または選択用オリゴヌクレオチドの設計は、例えば、DNAWorks (Hoover and Lubkowski (2002) *Nuc. Acids Res.* 30:e43、Gene2Oligo (Rouillard et al., *Nucleic Acids Res.* 32:W176-180 (2004)およびberry.engin.umich.edu/gene2oligoのワールドワイドウェブ)、または以下でさらに記述されるCAD-PAMソフトウェアなどのコンピュータプログラムを用いて容易にされてもよい。ある種の態様では、単一のプール中での複数のオリゴヌクレオチドの操作を容易にするため、実質的にほぼ同じ融解温度を有するように複数の構築用オリゴヌクレオチド/選択用オリゴヌクレオチドのペアを設計することが望ましいかもしれない。この過程は上記のコンピュータプログラムにより容易にされてもよい。さまざまなオリゴヌクレオチド配列間の融解温度の規準化は、オリゴヌクレオチドの長さを変えることによりおよび/または配列のコドン再マッピング(例えば、最終的にポリヌクレオチドによりコードされるポリヌクレオチドの配列を変化させることなく1つまたは複数のオリゴヌクレオチド中のA/T対G/C含量を変えること)により行うことができる(例えば、WO 99/58721を参照のこと)。

## 【0076】

ある種の態様では、構築用オリゴヌクレオチドは、本質的に所望のポリヌクレオチド構築体のセンスおよびアンチセンス鎖の完全な相補体を供与するよう設計される。例えば、構築用オリゴヌクレオチドは完全なポリヌクレオチド構築体を形成するため、ともにハイブリダイズされライゲーションに供される必要があるだけである。その他の態様では、構築用オリゴヌクレオチドの相補体は全配列を網羅するが、ライゲーション前の鎖伸長によって埋められうる一本鎖のギャップを残すように設計されてもよい。この態様はより少ないおよび/またはより短い構築用オリゴヌクレオチドおよび/または選択用オリゴヌクレオチドの合成を必要とするので、ポリヌクレオチド構築体の産生を容易にするはずである。

## 【0077】

典型的な態様では、構築および/または選択用オリゴヌクレオチドは、1セット、または数セットのプライマーによる核酸プールの増幅に利用できるユニバーサルプライマーに対する結合部位の1つまたは複数のセットを含んでもよい。ユニバーサルプライマー結合部

10

20

30

40

50

位の配列は、効率的なプライマーハイブリダイゼーションと鎖伸長を可能とするのに適した長さで配列を有するよう選択されてもよい。さらに、ユニバーサルプライマー結合部位の配列は、プール中の核酸の望ましくない領域との非特異的結合を最小限に抑えるように最適化されてもよい。ユニバーサルプライマーおよびユニバーサルプライマーに対する結合部位の設計は、例えば、DNAWorks (前記)、Gene20ligo (前記)などのコンピュータプログラム、または以下でさらに論じられる実装システムや方法を利用して容易にされてもよい。ある種の態様では、ポリヌクレオチド構築の異なる段階で核酸の増幅を可能にできるユニバーサルプライマー/プライマー結合部位のいくつかのセットを設計することが望ましいかもしれない(図6)。例えば、1セットのユニバーサルプライマーを利用して、構築および/または選択用オリゴヌクレオチドのセットを増幅してもよい。サブアッセムブリへの構築用オリゴヌクレオチドのセットのアッセムブリ後、このサブアッセムブリは同じまたは異なるユニバーサルプライマーのセットを用いて増幅されてもよい。例えば、サブアッセムブリに組み入れられる最も3'および5'末端の構築用オリゴヌクレオチドは、ユニバーサルプライマー結合部位の2つまたはそれ以上のネスティッドセット、つまり構築用オリゴの最初の増幅に利用できる最も外側のセットおよびサブアッセムブリを増幅するのに利用できる第2のセットを含んでもよい。アッセムブリ(例えば、構築および/もしくは選択用オリゴヌクレオチド、サブアッセムブリならびに/またはポリヌクレオチド構築体)の各段階に増幅用の複数セットのユニバーサルプライマーを組み入れることが可能である。

10

**【0078】**

典型的な態様では、ユニバーサルプライマーは一過性プライマー、例えば、化学的または酵素的切断によって核酸分子から除去できるプライマーとして設計されてもよい。核酸の化学的、熱的、光に基づく、または酵素的切断の方法は以下に詳述される。典型的な態様では、ユニバーサルプライマーはIIS型制限エンドヌクレアーゼを用いて除去することができる。

20

**【0079】**

構築および/または選択用オリゴヌクレオチドは、所望の配列を有するオリゴヌクレオチドの調製で当技術分野において知られている任意の方法によって調製することができる。例えば、オリゴヌクレオチドは天然供給源から単離されても、商業的供給源から購入されても、または第一原理から設計されてもよい。好ましくは、オリゴヌクレオチドは、費用および生産時間を減らし柔軟性を高めるため、高処理・並行合成を可能にする方法により合成することができる。典型的な態様では、構築および/または選択用オリゴヌクレオチドは固体支持体上で、各オリゴヌクレオチドが基板上で個別のフィーチャまたは位置に合成されるアレイ形式に、例えば、共通の基板上でインサイチュー合成される一本鎖DNAセグメントのマイクロアレイに合成することができる。アレイは構築されても、特別注文されても、または商業者から購入されてもよい。アレイを構築する様々な方法は当技術分野において周知である。例えば、固体支持体上での、例えば、アレイ形式の構築および/または選択用オリゴヌクレオチドの合成に適用できる方法および技術は、例えば、WO 00/58516、米国特許第5,143,854号、同第5,242,974号、同第5,252,743号、同第5,324,633号、同第5,384,261号、同第5,405,783号、同第5,424,186号、同第5,451,683号、同第5,482,867号、同第5,491,074号、同第5,527,681号、同第5,550,215号、同第5,571,639号、同第5,578,832号、同第5,593,839号、同第5,599,695号、同第5,624,711号、同第5,631,734号、同第5,795,716号、同第5,831,070号、同第5,837,832号、同第5,856,101号、同第5,858,659号、同第5,936,324号、同第5,968,740号、同第5,974,164号、同第5,981,185号、同第5,981,956号、同第6,025,601号、同第6,033,860号、同第6,040,193号、同第6,090,555号、同第6,136,269号、同第6,269,846号、および同第6,428,752号ならびにZhou et al., *Nucleic Acids Res.* 32: 5409-5417 (2004)に記述されている。

30

40

**【0080】**

典型的な態様では、構築および/または選択用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)を用いて固体支持体上で合成することができる。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におよび同時係属中の米国特

50

許第6,375,903号に記述されている。その他の例では、アレイ中のフィーチャのそれぞれが所望の配列の一本鎖DNA分子を有するカスタムDNAマイクロアレイを製造できるマスクレス装置が知られている。好ましいタイプの装置は反射光学系の利用に基づいた、米国特許第6,375,903号の図5に示されるタイプである。このタイプのマスクレス・アレイ・シンセサイザは、ソフトウェア制御下にあることが望ましい。マイクロアレイ合成の全過程をたった数時間のうちに行うことができるので、および適当なソフトウェアによって所望のDNA配列を自由自在に変えられるので、このクラスの装置は異なる配列のDNAセグメントを含むマイクロアレイを毎日、場合によっては1つの装置で1日当たり複数回製造することを可能にする。マイクロアレイ中のDNAセグメントのDNA配列の相違は、僅かでもまたは劇的にもよく、これによって過程が異なることはない。MAS装置は、ハイブリダイゼーション実験用のマイクロアレイを作製するためにこれが通常使われる形式で使用されてもよいが、これは本明細書に記述される組成物、方法およびシステムに特別に適した特徴を有するように適合されてもよい。例えば、上述の米国特許第6,375,903号の図5に示される光源の代わりにコヒーレント光源、すなわちレーザーを使うことが望ましいかもしれない。レーザーが光源として使用される場合、マスクレス・アレイ・シンセサイザで使われるマイクロミラーアレイを照射するため、ビーム拡大および散乱板をレーザーの後に使用して、狭いレーザー光ビームをより広い光源に変換することができる。マイクロアレイが合成されるフローセルに変化を加えられることも想定される。具体的には、フローセルは、直線列のアレイ要素が共通の流体チャンネルによって互いと連通しているが、各チャンネルが隣接列のアレイ要素に付随する隣接のチャンネルからは分離された状態で、区画化できると想定される。マイクロアレイ合成の間に、チャンネルは全て同じ流体を同時に受ける。DNAセグメントを基板から分離した後、チャンネルはアレイ要素の列からのDNAセグメントを互いと集合させるのを可能にし、ハイブリダイゼーションによって自己アッセンブルし始めるのを可能にする働きをする。

10

20

**【0081】**

構築および/または選択用オリゴヌクレオチドを合成するその他の方法としては、例えば、マスクを利用した光による方法、流体チャンネル法、フローチャンネル法、スポットティング法、ピンに基づく方法、および多数の支持体を利用した方法が挙げられる。

**【0082】**

オリゴヌクレオチドの合成のためのマスクを利用した光による方法(例えば、VLSIPS(商標)法)は、例えば、米国特許第5,143,854号、同第5,510,270号および同第5,527,681号に記述されている。これらの方法は、固体支持体の所定の領域を活性化する段階、その後支持体を事前に選択した単量体溶液と接触させる段階を含む。選択の領域は、集積回路製作で使われる光リソグラフィ技術のように多量のマスクを通じての光源を使った照射により活性化することができる。支持体のその他の領域は不活性なままである。その理由は照射がマスクにより遮断され、その領域が化学的に保護されたままであるためである。すなわち、光のパターンが、支持体のどの領域が所与の単量体と反応するかを規定する。繰り返し異なる所定領域のセットを活性化し、異なる単量体溶液を支持体と接触させることにより、多様な重合体アレイが支持体上で作製される。未反応の単量体溶液を支持体から洗浄するなどの、その他のステップが必要に応じて使われてもよい。その他の適用可能な方法

30

40

**【0083】**

単一の支持体上での構築および/または選択用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号に記述されている。例えば、試薬を(1) 所定の領域に規定されたチャンネル内にフローイングさせるかまたは(2) 所定の領域に「スポットティング」するかのいずれかにより支持体に送達することができる。その他の手法、およびスポットティングとフローイングさせる組合せを同様に利用してもよい。どの場合にも、単量体溶液を種々の反応部位に送達する際に、支持体の特定の活性化領域をその他の領域から機械的に分離する。

**【0084】**

50

フローチャンネル法には、例えば、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微小流体システムが含まれる。例えば、多様な重合体配列は、支持体の表面上に、適切な試薬が流れるまたは適切な試薬が入れられるフローチャンネルを形成させることにより固体支持体の選択の領域に合成することができる。当業者であれば、チャンネルを形成するまたは支持体表面の一部分を別の方法で保護する別法が存在することを認識すると考えられる。例えば、親水性または疎水性コーティング(溶媒の性質に応じ)などの保護コーティングを、場合によってはその他の領域中での反応物質溶液による湿潤を促進する材料と組み合わせて、保護する支持体の部分一面に利用する。このように、流通溶液がその指定流路の外側を通ることをさらに阻止する。

#### 【0085】

固体支持体上でのオリゴヌクレオチド調製のためのスポットティング法には、反応物質を選択の領域に直接的に堆積させることで、反応物質を比較的少量で送達する段階が含まれる。ステップによっては、そうすることがより効率的であるならば、支持体表面全体に溶液を吹き付けてもまたは別の方法でコーティングしてもよい。領域間を移動する分注器によって、正確に測定された一定分量の単量体溶液を液滴状に堆積させることができる。典型的な分注器としては、単量体溶液を支持体に送達するマイクロピペットおよびマイクロピペットの位置を支持体に対して制御するロボットシステム、またはインクジェットプリンタが挙げられる。その他の態様では、分注器は、各種の試薬を反応領域に同時に送達できるように一連の試験管、マニホールド、ずらりと並んだピペットまたは同様のものを含む。

10

20

#### 【0086】

固体支持体上でのオリゴヌクレオチド合成のためのピンに基づく方法は、例えば、米国特許第5,288,514号に記述されている。ピンに基づく方法では、複数のピンまたはその他の伸長部を有する支持体を利用する。これらのピンはそれぞれがトレイ内の個々の試薬用容器中に同時に挿入される。96ピンのアレイが96ウェルマイクロタイターディッシュなどの、96容器トレイとともに広く用いられる。各トレイは個々のピン上での特定の化学反応における特定のカップリング試薬で満たされている。したがって、それらのトレイは多くの場合、異なる試薬を含むはずである。それらの化学反応は、比較的似通った反応条件のセットの下で、反応の各々が行われうるように最適化されているので、多数の化学的カップリングステップを同時に行うことが可能になる。

30

#### 【0087】

別の態様では、多数の構築および/または選択用オリゴヌクレオチドを複数の支持体上で合成することができる。例には、例えば、米国特許第5,770,358号、同第5,639,603号および同第5,541,061号に記述されているビーズに基づく合成法がある。ビーズ上でのオリゴヌクレオチドなどの分子の合成の場合、かなり多数のビーズを容器中の適当な担体(水などの)に懸濁する。これらのビーズには、任意で保護基を複合体形成させた活性部位を有する任意のスペーサー分子が付与されている。合成の各ステップで、カップリングのため、ビーズを多数の容器に分ける。新生オリゴヌクレオチド鎖を脱保護した後、異なる単量体溶液を各容器に加え、したがって所与の容器中の全てのビーズ上で、同じヌクレオチド付加反応が行われる。その後、ビーズは過剰の試薬を洗い流され、単一の容器中にプールされ、混合され、次の合成ラウンドに向けて別の多数の容器の中に再分配される。最初に利用されるビーズが多数であることによって、幾多のラウンドの塩基無作為付加の後にそれぞれ固有のオリゴヌクレオチド配列がその表面上で合成されている、多数のビーズが同じように容器中に無作為に分散していることに留意されたい。個々のビーズは使用中の識別を可能とするため、ビーズ上の二本鎖オリゴヌクレオチドに固有の配列でタグ付けされてもよい。

40

#### 【0088】

固体支持体上でのオリゴヌクレオチドの合成に有用な種々の典型的な保護基は、例えば、Atherton et al., 1989, Solid Phase Peptide Synthesis, IRL Pressに記述されている。

50

## 【0089】

種々の態様では、本明細書に記述される方法は核酸の固定化のため固体支持体を利用する。例えば、オリゴヌクレオチドを1つまたは複数の固体支持体上で合成することができる。さらに、選択用オリゴヌクレオチドを固体支持体上に固定化して、配列エラーを含む構築用オリゴヌクレオチドの除去を容易にすることができる。典型的な固体支持体としては、例えば、スライド、ビーズ、チップ、粒子、鎖、ゲル、シート、管、球体、容器、毛细管、パッド、薄片、フィルム、またはプレートが挙げられる。種々の態様では、固体支持体は生物学的でも、非生物学的でも、有機的でも、無機的でも、またはそれらの組合せでもよい。実質的に平面である支持体を使用する場合、支持体は、例えば、細長いくぼみ、溝、ウェル、または化学的障壁(例えば、疎水性コーティングなど)で各領域に物理的に分離されてもよい。光を通す支持体は、アッセイ法が光学的検出を含む場合に有用である(例えば、米国特許第5,545,531号を参照のこと)。固体支持体の表面はカルボキシル、アミノおよびヒドロキシルなどの反応基を通常含むはずであり、または官能化シリコン化合物でコーティングされてもよい(例えば、米国特許第5,919,523号を参照のこと)。

10

## 【0090】

1つの態様では、固体支持体上で合成されたオリゴヌクレオチドをさらに長いポリヌクレオチド構築体へのアッセムブリに向けた構築用オリゴヌクレオチドおよび/または選択用オリゴヌクレオチドの産生用の鑄型として使用することができる。例えば、支持体に結合したオリゴヌクレオチドを、プライマーの鎖伸長を可能にする条件の下で、オリゴヌクレオチドにハイブリダイズするプライマーと接触させることができる。次いで、支持体に結合した二重鎖を変性させ、さらなる増幅ラウンドに供することができる。

20

## 【0091】

別の態様では、支持体に結合したオリゴヌクレオチドをポリヌクレオチド構築体へのアッセムブリの前に固体支持体から除去することができる。オリゴヌクレオチドは、例えば、酸、塩基、酸化、還元、熱、光、金属イオン触媒反応、置換もしくは脱離化学反応などの条件に曝すことにより、または酵素的切断により固体支持体から除去することができる。

## 【0092】

1つの態様では、オリゴヌクレオチドは切断可能な結合部分を通じて固体支持体に結合させることができる。例えば、固体支持体を官能化して、オリゴヌクレオチドとの共有結合用の切断可能なリンカーを付与することができる。このリンカー部分は長さが6またはそれ以上の原子のものですることができる。あるいは、切断可能な部分はオリゴヌクレオチド内であってもよく、インサイチュウ合成の間に導入されてもよい。幅広い種類の切断可能な部分が固相およびマイクロアレイ・オリゴヌクレオチド合成の分野で利用可能である(例えば、Pon, R., *Methods Mol. Biol.* 20:465-496 (1993); Verma et al., *Ann. Rev. Biochem.* 67:99-134 (1998); 米国特許第5,739,386号、同第5,700,642号および同第5,830,655号; ならびに米国特許第2003/0186226号および同第2004/0106728号を参照のこと)。適した切断可能な部分は、とりわけ、ヌクレオシド塩基の保護基の性質、固体支持体の選択および/または試薬送達の方法と適合するように選択することができる。典型的な態様では、固体支持体から切断されたオリゴヌクレオチドは、遊離3'-OH末端を含む。あるいは、遊離3'-OH末端をオリゴヌクレオチドの切断後、化学的または酵素的処理によって得ることもできる。切断可能な部分は、オリゴヌクレオチドを分解しない条件の下で除去することができる。好ましくは、リンカーは2通りの手法により、つまり(a) 脱保護ステップと同じ条件の下で同時にまたは(b) 脱保護ステップの完了後にリンカー切断用の異なる条件もしくは試薬を用いて逐次的に切断することができる。

30

40

## 【0093】

共有結合固定化部位はオリゴヌクレオチドの5'末端にあってもまたはオリゴヌクレオチドの3'末端にあってもよい。場合によっては、固定化部位はオリゴヌクレオチド内に(すなわちオリゴヌクレオチドの5'または3'末端以外の部位に)あってもよい。切断可能な部位はオリゴヌクレオチド主鎖、例えば、リボース、ジアルコキシシラン、ホスホロチオエ

50

ートおよびホスホルアミデート・ヌクレオチド間結合などの、ホスホジエステル基の1つに代わる修飾3'-5'ヌクレオチド間結合に沿って位置してもよい。切断可能なオリゴヌクレオチド類似体は同様に、7-デアザグアノシン、5-メチルシトシン、イノシン、ウリジンおよび同様のものなどの、塩基または糖の1つへの置換、またはそれらの1つの交換を含んでもよい。

#### 【0094】

1つの態様では、修飾オリゴヌクレオチドの内部に含まれる切断可能な部位は、ジアルコキシシラン、3'-(S)-ホスホロチオエート、5'-(S)-ホスホロチオエート、3'-(N)-ホスホルアミデート、5'-(N)ホスホルアミデートおよびリボースなどの、化学的に切断可能な基を含むことができる。化学的に切断可能なオリゴヌクレオチドの合成および切断条件は、米国特許第5,700,642号および同第5,830,655号に記述されている。例えば、導入される切断可能な部位の選択に応じて、官能化ヌクレオシドまたは修飾ヌクレオシド二量体のいずれかを最初に調製し、その後オリゴヌクレオチド合成の間に成長中のオリゴヌクレオチド断片に選択的に導入してもよい。ジアルコキシシランの選択的切断は、フッ化物イオンを用いた処理により達成することができる。ホスホロチオエート・ヌクレオチド間結合は、穏やかな酸化条件の下で選択的に切断することができる。ホスホルアミデート結合の選択的切断は、80%酢酸などの、穏やかな酸性条件の下で行うことができる。リボソームの選択的切断は、希水酸化アンモニウムを用いた処理により達成することができる。

#### 【0095】

別の態様では、米国特許出願第2003/0186226号に記述されているように、ホスホルアミダイトまたはH-ホスホネート・オリゴヌクレオチド合成の前に特殊なホスホルアミダイトをヒドロキシル基にカップリングさせることにより、非切断可能なヒドロキシルリンカーを切断可能なリンカーに変換することができる。オリゴヌクレオチド合成の完了時の化学的リン酸化剤での切断により、リン酸基を3'末端に持ったオリゴヌクレオチドが得られる。3'-リン酸末端はアルカリホスファターゼなどの、酵素または化学物質を用いた処理により3'ヒドロキシル末端に変換されてもよく、これは当業者により日常的に行われている。

#### 【0096】

別の態様では、切断可能な連結部分はTOPS(1合成当たり2オリゴヌクレオチド)リンカー(例えば、PCT公開WO 93/20092を参照のこと)であってもよい。例えば、TOPSホスホルアミダイトを使用して、固体支持体上の非切断可能なヒドロキシル基を切断可能なリンカーに変換することができる。TOPS試薬の好ましい態様は、Universal TOPS(商標)ホスホルアミダイトである。Universal TOPS(商標)ホスホルアミダイト調製の条件、カップリングおよび切断は、例えば、Hardy et al, Nucleic Acids Research 22(15):2998-3004 (1994)に詳述されている。Universal TOPS(商標)ホスホルアミダイトは、長時間のアンモニアおよび/またはアンモニア/メチルアミン処理などの、塩基性条件の下で除去できる環状3'ホスフェートをもたらす、結果的に天然の3'ヒドロキシオリゴヌクレオチドを生ずる。

#### 【0097】

別の態様では、切断可能な連結部分はアミノリンカーであってもよい。結果的に得られる、ホスホルアミダイト結合を介してリンカーに結合したオリゴヌクレオチドを80%酢酸で切断し、3'-リン酸化オリゴヌクレオチドをもたらすことができる。

#### 【0098】

別の態様では、切断可能な連結部分は光切断可能なオルト-ニトロベンジルリンカーなどの、光切断可能なリンカーであってもよい。固体支持体上での光解離性オリゴヌクレオチドの合成および切断条件は、例えば、Venkatesan et al. J. of Org. Chem. 61:525-529 (1996), Kahl et al., J. of Org. Chem. 64:507-510 (1999), Kahl et al., J. of Org. Chem. 63:4870-4871(1998), Greenberg et al., J. of Org. Chem. 59:746-753 (1994), Holmes et al., J. of Org. Chem. 62:2370-2380 (1997)、および米国特許第5,739,386号に記述されている。ヒドロキシメチル、ヒドロキシエチルおよびFmoc-アミノエチルカルボン酸リンカーなどの、オルト-ニトロベンジルに基づくリンカーは、商業的に入手す

10

20

30

40

50

ることできる。

【0099】

別の態様では、より短い構築用オリゴヌクレオチドはより長い構築用オリゴヌクレオチドよりも純粋であり、配列エラーが少ないはずなので、より短い構築用オリゴヌクレオチドを合成し、構築に使用してもよい。例えば、構築用オリゴヌクレオチドは、約30から約100ヌクレオチド、約30から約75ヌクレオチドまたは約30から約50オリゴヌクレオチドとすることができる。その他の態様では、構築用オリゴヌクレオチドは、合成ポリヌクレオチドの配列全体を本質的に網羅するのに十分である(例えば、ポリメラーゼによって埋められる必要のあるギャップがオリゴヌクレオチド間に存在していない)。オリゴヌクレオチドそれら自体がチェック機構として働くことができる。何故ならば、不適合のオリゴヌクレオチドは完全適合のオリゴヌクレオチドよりも優先的にアニーリングすることが少ないはずであり、したがってハイブリダイゼーション条件を注意深く制御することにより、エラーを含む配列を減らすことができるからである。

10

20

30

【0100】

別の態様では、オリゴヌクレオチドはヌクレアーゼなどの酵素により固体支持体から除去することができる。例えば、オリゴヌクレオチドは、例えば、IIs型制限酵素を含めて、1種または複数種の制限エンドヌクレアーゼへの曝露により固体支持体から除去することができる。制限エンドヌクレアーゼ認識配列を固定化オリゴヌクレオチドに導入することができ、オリゴヌクレオチドを1種または複数種の制限エンドヌクレアーゼと接触させて、オリゴヌクレオチドを支持体から除去することができる。種々の態様では、酵素的切断を利用してオリゴヌクレオチドを支持体から除去する場合、一本鎖固定化オリゴヌクレオチドをプライマー、ポリメラーゼおよびdNTPsと接触させて、固定化二重鎖を形成させることが望ましいかもしれない。二重鎖を次いで酵素(例えば、制限エンドヌクレアーゼ)と接触させて、二重鎖を支持体の表面から除去することができる。支持体に結合したオリゴヌクレオチド上で第2鎖を合成する方法および支持体に結合した二重鎖の酵素的除去の方法は、例えば、米国特許第6,326,489号に記述されている。あるいは、制限エンドヌクレアーゼ認識および/または切断部位に相補的である(例えば、しかし支持体に結合したオリゴヌクレオチド全体には相補的でない)短いオリゴヌクレオチドをハイブリダイゼーション条件の下で、支持体に結合したオリゴヌクレオチドに加えて、制限エンドヌクレアーゼによる切断を容易にすることができる(例えば、PCT公報番号W0 04/024886を参照のこと)。

30

【0101】

種々の態様では、本明細書に開示される方法は、例えば、構築用オリゴヌクレオチド、選択用オリゴヌクレオチド、サブアッセンブルおよび/またはポリヌクレオチド構築体を含む、核酸の増幅を含む。増幅はアッセンブリスキームの間の1つまたは複数の段階で行われてもよく、および/またはアッセンブリの間のある段階で1回または複数回行われてもよい。増幅方法は、ハイブリダイゼーションおよび鎖伸長を促進する条件の下で核酸に特異的にハイブリダイズする1つまたは複数のプライマーと核酸を接触させる段階を含むことができる。核酸を増幅させる典型的な方法としては、ポリメラーゼ連鎖反応(PCR)法(例えば、Mullis et al. (1986) Cold Spring Harb. Symp. Quant. Biol. 51 Pt 1:263およびCleary et al. (2004) Nature Methods 1:241; ならびに米国特許第4,683,195号および同第4,683,202号を参照のこと)、アンカーPCR法、RACE PCR法、ライゲーション連鎖反応(LCR)法(例えば、Landegran et al. (1988) Science 241:1077-1080; およびNakazawa et al. (1994) Proc. Natl. Acad. Sci. U.S.A. 91:360-364を参照のこと)、自己持続性配列複製法(Guatelli et al. (1990) Proc. Natl. Acad. Sci. U.S.A. 87:1874)、転写増幅系(Kwoh et al. (1989) Proc. Natl. Acad. Sci. U.S.A. 86:1173)、Q-レプリカーゼ増幅(Lizardi et al. (1988) BioTechnology 6:1197)、反復PCR法(Jaffe et al. (2000) J. Biol. Chem. 275:2619; およびWilliams et al. (2002) J. Biol. Chem. 277:7790)、米国特許第6,391,544号、同第6,365,375号、同第6,294,323号、同第6,261,797号、同第6,124,090号および同第5,612,199号に記述されている増幅方法、または当業者に周知の技術を用

40

50

いたその他任意の核酸増幅方法が挙げられる。典型的な態様では、本明細書に開示される方法はPCR増幅を利用する。

#### 【0102】

ある種の態様では、核酸配列に特異的なプライマーセットを利用して、単離される特異的核酸配列を増幅するまたは核酸配列のプールの一部である特異的核酸配列を増幅することができる。別の態様では、複数のプライマーセットを利用して、任意で単一の反応混合物の中にもプールされてもよい複数の特異的核酸配列を増幅することができる。典型的な態様では、ユニバーサルプライマーのセットを利用して、単一のプール中にあるまたは複数のプールに分けられてもよい複数の核酸配列を増幅することができる(図32)。アッセムブリの間の異なる段階で核酸を増幅する場合、増幅が望まれる各段階用に異なるユニバーサルプライマーのセットを利用することが望ましいかもしれない(図33)。例えば、ユニバーサルプライマーの第1セットを利用して、構築および/または選択用オリゴヌクレオチドを増幅することができ、ユニバーサルプライマーの第2セットを利用して、サブアッセムブリまたはポリヌクレオチド構築体を増幅することができる(図33)。上述のように、ユニバーサルプライマーの1つまたは複数のセットに向けたプライマー結合部位を考慮して、構築用オリゴヌクレオチドおよび/または選択用オリゴヌクレオチドは設計されてもよい。あるいは、プライマー結合部位は、標的核酸に相補的な領域と増幅過程の間に組み入れられるようになる非相補的な領域とを含むキメラプライマーの使用により、合成後に核酸に付加されてもよい(例えば、WO 99/58721を参照のこと)。

10

#### 【0103】

典型的な態様では、プライマー/プライマー結合部位は一過性であるように、例えば、アッセムブリの間の所望の段階でプライマー/プライマー結合部位の除去を可能とするように設計することができる。一過性プライマーは、化学的、熱的、光に基づく、または酵素的切断により除去可能であるように設計することができる。切断は外部要因(例えば、酵素、化学物質、熱、光など)の付加によって行われてもよく、またはある時間周期の後に(例えば、*n*ラウンドの増幅の後に)自動的に行われてもよい。1つの態様では、一過性プライマーは化学的切断により除去することができる。例えば、酸に不安定なまたは塩基に不安定な部位を有するプライマーが増幅に使われてもよい。その後、増幅されたプールを酸または塩基に曝露してプライマー/プライマー結合部位を所望の位置で除去することができる。あるいは、一過性プライマーは熱および/または光への曝露により除去することができる。例えば、熱に不安定なまたは光に不安定な部位を有するプライマーが増幅に使われてもよい。その後、増幅されたプールを熱および/または光に曝露してプライマー/プライマー結合部位を所望の位置で除去することができる。別の態様では、RNAプライマーを増幅に使用し、それによって核酸分子の末端にRNA/DNAハイブリッドの短いストレッチを形成させることができる。その後、プライマー部位をRNase(例えば、RNase H)への曝露により除去することができる。種々の態様では、プライマーを除去する方法は、増幅された二重鎖の一本鎖を切断するのみであり、それによって3'または5'突出部を残すことがある。そのような突出部はエキソヌクレアーゼにより除去して、平滑末端二本鎖の二重鎖を形成させることができる。例えば、RecJ<sub>I</sub>を使用して一本鎖5'突出部を除去することができ、エキソヌクレアーゼIまたはエキソヌクレアーゼTを使用して一本鎖3'突出部を除去

20

30

40

#### 【0104】

1つの態様では、一過性プライマーは化学的、熱的または光に基づく切断により核酸が

50

ら除去することができる。本明細書に記述される方法で用いる典型的な化学的に切断可能なヌクレオチド間結合としては、例えば、 $\beta$ -シアノエーテル、5'-デオキシ-5'-アミノカルバメート、3'-デオキシ-3'-アミノカルバメート、尿素、2'-シアノ-3',5'-ホスホジエステル、3'-(S)-ホスホロチオエート、5'-(S)-ホスホロチオエート、3'-(N)-ホスホルアミデート、5'-(N)-ホスホルアミデート、 $\alpha$ -アミノアミド、ピシナルジオール、リボヌクレオシド挿入、2'-アミノ-3',5'-ホスホジエステル、アリルスルホキシド、エステル、シリルエーテル、ジチオアセタール、5'-チオ-フォーマル(formal)、 $\beta$ -ヒドロキシ-メチル-ホスホニックビスアミド、アセタール、3'-チオ-フォーマル、メチルホスホネートおよびホスホトリエステルが挙げられる。トリアルキルシリルエーテルおよびジアルコキシシランなどのヌクレオシド間シリル結合は、フッ化物イオンを用いた処理により切断される。塩基切断可能な部位としては $\beta$ -シアノエーテル、5'-デオキシ-5'-アミノカルバメート、3'-デオキシ-3'-アミノカルバメート、尿素、2'-シアノ-3',5'-ホスホジエステル、2'-アミノ-3',5'-ホスホジエステル、エステルおよびリボースが挙げられる。3'-(S)-ホスホロチオエートおよび5'-(S)-ホスホロチオエートなどの硫黄含有ヌクレオチド間結合は、硝酸銀または塩化第二水銀を用いた処理により切断される。酸切断可能な部位としては3'-(N)-ホスホルアミデート、5'-(N)-ホスホルアミデート、ジチオアセタール、アセタールおよびホスホニックビスアミドが挙げられる。 $\alpha$ -アミノアミドヌクレオシド間結合はイソチオシアネートを用いた処理により切断可能であり、チタンを利用して2'-アミノ-3',5'-ホスホジエステル-0-オルト-ベンジルヌクレオシド間結合を切断することができる。ピシナルジオール結合は過ヨウ素酸塩を用いた処理により切断可能である。熱的に切断可能な基のなかにはアリルスルホキシドおよびシクロヘキセンがあり、その一方で光に不安定な結合のなかにはニトロベンジルエーテルおよびチミジン二量体がある。化学的に切断可能な、熱的に切断可能な、および光に不安定な基を含む核酸を合成し切断する方法は、例えば、米国特許第5,700,642号に記述されている。

10

20

#### 【0105】

その他の態様では、一過性プライマー/プライマー結合部位は、酵素的切断により除去することができる。例えば、プライマー/プライマー結合部位は制限エンドヌクレアーゼ切断部位を含むように設計することができる。増幅後、核酸のプールを1種または複数種のエンドヌクレアーゼと接触させて、二本鎖切断をもたらし、それによってプライマー/プライマー結合部位を除去することができる。ある種の態様では、フォワードおよびリバースプライマーは、同じまたは異なる制限エンドヌクレアーゼにより除去することができる。任意の種類のリミットエンドヌクレアーゼを利用して、プライマー/プライマー結合部位を核酸配列から除去することができる。特異的な結合および/または切断部位を有する多種多様な制限エンドヌクレアーゼは、例えば、New England Biolabs (Beverly, MA)から市販されている。種々の態様では、3'突出部、5'突出部または平滑末端を作製する制限エンドヌクレアーゼを利用することができる。突出部を作製する制限エンドヌクレアーゼを用いる場合、エキソヌクレアーゼ(例えば、RecJ<sub>1</sub>、エキソヌクレアーゼI、エキソヌクレアーゼT、S<sub>1</sub>ヌクレアーゼ、P<sub>1</sub>ヌクレアーゼ、マングビーンヌクレアーゼ、CEL Iヌクレアーゼなど)を利用して平滑末端を作製することができる。あるいは、特定の制限エンドヌクレアーゼによって形成された付着端を利用して、所望の配列でのサブアッセムブリのアクセスを円滑にすることができる(例えば、図31Aを参照のこと)。典型的な態様では、HIS型制限エンドヌクレアーゼの結合および/または切断部位を含むプライマー/プライマー結合部位を利用して、一過性プライマーを除去することができる。

30

40

#### 【0106】

本明細書に開示される増幅方法で用いるのに適したプライマーは、例えば、DNAWorks (前記)、Gene2Oligo (前記)、または本明細書に記述されるCAD-PAMソフトウェアなどの、コンピュータプログラムを用いて設計することができる。通常、プライマーは、長さが約5から約500、約10から約100、約10から約50、または約10から約30ヌクレオチドである。典型的な態様では、複雑な反応混合物の操作を容易にするため、実質的にほぼ同じ融解温度を有するようにプライマーのセットまたはプライマーの複数のセットを設計することが

50

できる。融解温度は例えば、プライマー長およびヌクレオチド組成により影響を受ける可能性がある。

【0107】

ある種の態様では、キャップ(例えば、エキソヌクラーゼ切断を阻止するため)、連結部分(基板上へのオリゴヌクレオチドの固定化を容易にする前述のものなどの)、または標識(例えば、核酸構築体の検出、単離および/または固定化を容易にするため)などの1つまたは複数の修飾を含むプライマーを利用することが望ましいかもしれない。適当な修飾としては、例えば、さまざまな酵素、補欠分子団、発光マーカ、生物発光マーカ、蛍光マーカ(例えば、フルオレセイン)、放射能標識(例えば、<sup>32</sup>P、<sup>35</sup>Sなど)、ピオチン、ポリペプチドエピトープなどが挙げられる。本明細書の開示に基づいて、当業者は所与の用途に適したプライマー修飾を選択することができる。

10

【0108】

1つの態様では、本発明は配列最適化およびオリゴヌクレオチド設計の方法を提供する。1つの局面では、本発明は各遺伝子に対しマイナスとプラスの両鎖で互い違いとなった末端重複オリゴヌクレオチドのセットを設計する方法を提供する。別の局面では、オリゴヌクレオチドはともに、合成される配列全体を網羅する。別の局面では、オリゴヌクレオチド設計はコンピュータプログラムにより支援される。別の局面では、タンパク質をコードする配列はコンピュータプログラム、すなわち、本明細書に記述されるCAD-PAMプログラムにより最適化される。

【0109】

本発明の態様は、1つまたは複数の増幅配列または増幅部位を有するオリゴヌクレオチド配列(すなわち、構築用オリゴヌクレオチド配列および選択用オリゴヌクレオチド配列)に向けられる。本明細書で用いられる「増幅部位」という用語は、相補的な核酸配列をハイブリダイズする本発明のオリゴヌクレオチド配列の5'および/または3'末端に位置する核酸配列を含むよう意図されるが、これらに限定されることはない。本発明の1つの局面では、増幅部位は増幅後にオリゴヌクレオチドから除去される。本発明の別の局面では、増幅部位は、1種または複数種の制限酵素により認識される1種または複数種の制限エンドヌクラーゼ認識配列を含む。別の局面では、増幅部位は熱に不安定および/または光に不安定であり、それぞれ熱または光により切断可能である。別の局面では、増幅部位はRNaseにより切断可能なり核酸配列である。

20

30

【0110】

本明細書で用いられる「制限エンドヌクラーゼ認識部位」という用語は、1種または複数種の制限酵素が結合し、結果的に制限エンドヌクラーゼ認識配列それ自体でまたは制限エンドヌクラーゼ認識配列より遠位の配列でDNA分子の切断を引き起こす特定の核酸配列を含むよう意図されるが、これに限定されることはない。制限酵素はI型酵素、II型酵素、IIS型酵素、III型酵素およびIV型酵素を含むが、これらに限定されることはない。REBASEデータベースは、制限修飾に関する制限酵素、DNAメチル基転移酵素および関連タンパク質に関する情報の包括的データベースを供与する。これは、制限エンドヌクラーゼ認識部位および制限エンドヌクラーゼ切断部位、イソ制限酵素、商業的入手性、結晶および配列データに関する情報を扱った既刊および未刊の両研究を含んでいる(あらゆる目的でその全体が参照により本明細書に組み入れられるRoberts et al. (2005) Nuc. Acids Res. 33:D230を参照のこと)。

40

【0111】

ある種の局面では、本発明のプライマーは、制限エンドヌクラーゼ認識配列の3'側にIIS型酵素が核酸塩基対を切断できる1種または複数種の制限エンドヌクラーゼ認識部位を含む。本明細書で用いられる「IIS型」という用語は、制限酵素の認識配列から遠く離れた部位を切断する制限酵素のことをいう。IIS型酵素はその認識部位より0から20塩基対までの距離を切断することが知られている。例えば、IIS型エンドヌクラーゼの例としては、例えば、Bsr I、Bsm I、BstF5 I、BsrD I、Bts I、Mnl I、BciV I、Hph I、Mbo II、Eci I、Acu I、Bpm I、Mme I、BsaX I、Bcg I、Bae I、Bfi I、TspDT I、TspGW I、T

50

aq I I、Eco57 I、Eco57M I、Gsu I、Ppi IおよびPsr Iなどの3'突出部を作製する酵素；例えば、BsmA I、Ple I、Fau I、Sap I、BspM I、SfaN I、Hga I、Bvb I、Fok I、BceA I、BsmF I、Ksp632 I、Eco31 I、Esp3 I、Aar Iなどの5'突出部を作製する酵素；ならびに例えば、Mly IおよびBtr Iなどの平滑末端を作製する酵素が挙げられる。IIs型エンドヌクレアーゼは市販されており、当技術分野において周知である(New England Biolabs, Beverly, MA)。IIs型エンドヌクレアーゼを用いた認識部位、切断部位および消化の条件に関する情報は、例えば、[neb.com/nebecomm/enzymefindersearch\\_bytypeIIs.asp](http://neb.com/nebecomm/enzymefindersearch_bytypeIIs.asp)のワールドワイドウェブで見出すことができる。制限エンドヌクレアーゼ配列および制限酵素は当技術分野において周知であり、制限酵素は市販されている(New England Biolabs, Beverly, MA)。

10

#### 【0112】

ある種の態様では、検出可能な標識を有するプライマーが提供される。検出可能な標識としては、さまざまな酵素、補欠分子団、発光マーカー、生物発光マーカー、蛍光マーカーおよび同様のものが挙げられるが、これらに限定されることはない。適当な発光および生物発光マーカーの例としては、ビオチン、ルシフェラーゼ(例えば、細菌由来、ホタル由来、コメツキムシ由来など)、ルシフェリン、エクオリンおよび同様のものが挙げられるが、これらに限定されることはない。適当な蛍光タンパク質の例としては、黄色蛍光タンパク質(YFP)、緑色蛍光タンパク質(GFP)、シアン蛍光タンパク質(CFP)、ウンベリフェロン、フルオレセイン、フルオレセインイソチオシアネート、ローダミン、ジクロロトリアジニルアミンフルオレセイン、塩化ダンシル、フィコエリスリンおよび同様のものが挙げられるが、これらに限定されることはない。視覚的に検出可能なシグナルを有する適当な酵素系の例としては、ガラクトシダーゼ、グルコリニダーゼ、ホスファターゼ、ペルオキシダーゼ、コリンエステラーゼおよび同様のものが挙げられるが、これらに限定されることはない。検出可能な標識としては同様に、直接的にまたは間接的に放射能標識された、例えば、<sup>32</sup>P、<sup>35</sup>Sなどで標識された核酸が挙げられるが、これらに限定されることはない。あるいは、例えば、西洋ワサギペルオキシダーゼ、アルカリホスファターゼまたはルシフェラーゼを用いて化合物を酵素的に標識することができ、適切な基質の生成物への変換を測定することにより酵素標識を検出することができる。

20

#### 【0113】

本発明のある種の態様は核酸配列および非常に長い配列(例えば、遺伝子、遺伝子セット、ゲノムおよび同様のもの)を合成する方法であって、重複オリゴヌクレオチドおよび/または増幅プライマーのセットを配列特異的なハイブリダイゼーションに有利に働く条件の下で混合し、ハイブリダイズしている鎖を鋳型として利用して1種または複数種のポリメラーゼによりオリゴヌクレオチドを伸長する方法(すなわち、あらゆる目的でその全体が参照により本明細書に組み入れられるTian et al. (2004) Nature 432:1050に記述されているポリメラーゼアッセンブリ多重化(PAM)法)に向けられる。多重アッセンブリは図29~33に図解されている。1つの局面では、完全長の二本鎖DNA分子が合成され増幅されるまで、二本鎖伸長産物はさらなるラウンドの上記過程のために変性される。多重遺伝子合成は、本明細書に記述されるように溶液中でまたは支持体上で(アレイの一部として)行われてもよい。本明細書に記述される方法の利用の成功が、最近になってZhou et al. (2004) Nucleic Acids Res. 32:5409およびRichmond et al. (2004) Nucleic Acids Res. 32:5011(あらゆる目的でその全体が参照により本明細書に組み入れられる)により追認されている。

30

40

#### 【0114】

ポリメラーゼアッセンブリ多重化法に加えて、本明細書に記述される本発明のオリゴヌクレオチドおよび方法を用いて大きな二本鎖核酸配列を得るのに様々な方法が適している。例えば、PCRに基づくアッセンブリ法(PAM法またはポリメラーゼアッセンブリ多重化法を含む)およびライゲーションに基づくアッセンブリ法(例えば、付着端を有するポリヌクレオチドセグメントの連結)である。典型的な態様では、複数のポリヌクレオチド構築体を単一の反応混合物中でアッセンブルすることができる。その他の態様では、例えば、多

50

数のポリヌクレオチド構築体を合成する場合、内部の相同性領域を含むポリヌクレオチド構築体を合成する場合、または相同性の高いもしくは相同性領域を含む2つまたはそれ以上のポリヌクレオチド構築体を合成する場合、階層に基づくアッセンプリ法を利用することができる。

**【0115】**

1つの態様では、アッセンプリPCRを本明細書に記述される方法によって利用することができる。アッセンプリPCRでは、ポリヌクレオチドの少なくとも1つがポリメラーゼ(例えば、熱安定性ポリメラーゼ(例えば、Taqポリメラーゼ、VENT(商標)ポリメラーゼ(New England Biolabs)、TthIポリメラーゼ(Perkin-Elmer)および同様のもの)によってポリヌクレオチド鎖伸長可能な遊離3'-ヒドロキシルを有するような、アニーリングできる相補末端を有する少なくとも2つのポリヌクレオチドと組み合わせてポリメラーゼを介した鎖伸長を利用する。dNTPs、ポリメラーゼおよび緩衝液を含有する標準的なPCR反応液の中に、重複オリゴヌクレオチドを混合することができる。オリゴヌクレオチドの重複末端は、アニーリングにより、PCR反応においてポリメラーゼによる伸長のためのプライマーとして働く二本鎖核酸配列の領域を作製する。伸長反応の産物はもっと長い二本鎖核酸配列の形成のための基質として働き、最終的には完全長の標的配列の合成をもたらす(例えば、図3Bを参照のこと)。PCR条件を最適化して、標的の長いDNA配列の収量を増加させることができる。

10

**【0116】**

ある種の態様では、関心対象のポリヌクレオチド構築体を形成させるのに必要な重複オリゴヌクレオチドの全てをとともに混合することにより、標的配列を単一のステップで得ることができる。あるいは、産物を混合し、第2ラウンドのPCRに供する一連の別個のPCR反応からさらに長いポリヌクレオチド構築体をアッセンプルできるような、一連のPCR反応が並行してまたは逐次的に行われてもよい。さらに、自己プライミングPCRが単一の反応から完全なサイズの産物を産生できない場合、重複オリゴヌクレオチドのペア、もしくはは標的核酸配列のもっと小さな部分を別々にPCR増幅することによって、または従来のフィルインおよびライゲーション法によってアッセンプリを救済することができる。

20

**【0117】**

アッセンプリPCRを行う方法は、例えば、Kodumal et al. (2004) Proc. Natl. Acad. Sci. U.S.A. 101:15573; Stemmer et al. (1995) Gene 164:49; Dillon et al. (1990) BioTechniques 9:298; Hayashi et al. (1994) BioTechniques 17:310; Chen et al. (1994) J. Am. Chem. Soc. 116:8799; Prodromou et al. (1992) Protein Eng. 5:827; 米国特許第5,928,905号および同第5,834,252号; ならびに米国特許出願第2003/0068643号および同第2003/0186226号に記述されている。

30

**【0118】**

典型的な態様では、ポリメラーゼアッセンプリ多重化(PAM)法を利用して、本明細書に記述される方法によりポリヌクレオチド構築体をアッセンプルすることができる(例えば、Tian et al. (2004) Nature 432:1050; Zhou et al. (2004) Nucleic Acids Res. 32:5409; およびRichmond et al. (2004) Nucleic Acids Res. 32:5011を参照のこと)。ポリメラーゼアッセンプリ多重化法は、配列特異的なハイブリダイゼーションに有利に働く条件の下での重複オリゴヌクレオチドおよび/または増幅プライマーのセットの混合ならびにハイブリダイズしている鎖を鋳型として用いたポリメラーゼによる鎖伸長を含む。所望のポリヌクレオチド構築体の合成が完了するまで、二本鎖伸長産物を任意で変性し、さらなるラウンドのアッセンプリに使用してもよい。

40

**【0119】**

種々の態様では、本明細書に記述される方法によってポリヌクレオチド構築体をアッセンプルする方法としては、例えば、予め形成された二重鎖のライゲーション(例えば、Sca rpulla et al., Anal. Biochem. 121: 356-365 (1982); Gupta et al., Proc. Natl. Acad. Sci. USA 60: 1338-1344 (1968)を参照のこと)、Fok I法(例えば、Mandecki and Bolling, Gene 68: 101-107 (1988)を参照のこと)、二重非対称PCR (DA-PCR) (例えば、Stemm

50

er et al., Gene 164: 49-53 (1995); Sandhu et al., Biotechniques 12: 14-16 (1992); Smith et al., Proc. Natl. Acad. Sci. USA 100: 15440-15445 (2003)を参照のこと)、重複伸長PCR (OE-PCR) (例えば、Mehta and Singh, Biotechniques 26: 1082-1086 (1999)を参照のこと)、DA-PCR/OE-PCRの組合せ(例えば、Young and Dong, Nucleic Acids Res. 32: e59 (2004)を参照のこと)が挙げられる。

#### 【0120】

別の態様では、コンビナトリアル・アッセンプリストラテジーをポリヌクレオチドのアッセンプリに利用することができる(例えば、米国特許第6,670,127号、同第6,521,427号および同第6,521,427号を参照のこと)。手短に言えば、温度に基づく緩徐なアニーリングによってオリゴヌクレオチドと一緒に同時アニーリングし、その後ステップごとに新しいオリゴヌクレオチドの付加によるライゲーション鎖反応ステップを行うことができる。鎖中の第1のオリゴヌクレオチドを支持体に結合させる。逆鎖由来の第2の重複オリゴヌクレオチドを添加し、アニーリングさせ連結させる。第3の重複オリゴヌクレオチドを添加し、アニーリングさせ連結させるなど。関心対象の全てのオリゴヌクレオチドがアニーリングされ連結されるまで、この手順を反復する。自動化装置を利用し、この手順を長い配列に向け行うことができる。その後、二本鎖核酸配列を固体支持体から除去する。

#### 【0121】

ある種の態様では、階層的なアッセンプリストラテジーを本明細書に開示される方法によって利用することができる。階層的なアッセンプリストラテジーには、アッセンプリを段階的にまたは逐次的に制御するため反応混合物の各種成分を混合制御する種々の方法がある(例えば、米国特許第6,586,211号; 米国特許出願第2004/0166567号; PCT公報番号WO 02/095073; Zhou et al. (2004) Nucleic Acids Res. 32:5409を参照のこと)。例えば、複数のアッセンプリ反応を別々のプールで行うことができる。その後、これらのアッセンプリからの産物とともに混合して、さらに長くアッセンプルされた産物を形成させることができるなど。あるいは、階層的なアッセンプリストラテジーには、混合物中の反応種を変化させることによって外部制御を可能にする単一の反応混合物が含まれてもよい。例えば、光に不安定なリンカーを介して固体支持体に結合されているオリゴヌクレオチドを、定序アッセンプリを促進するのに利用できる極めて特異的且つ制御された方法で支持体から放出することができる(例えば、オリゴヌクレオチドを、制御された方法で固体支持体上の単一のアドレス可能な位置から除去することができる)。構築用オリゴヌクレオチドの第1セットを支持体から放出し、アッセンプリに供することができる。その次に、構築用オリゴヌクレオチドの第2セットを支持体から放出し、アッセンプルすることができるなど。1つの態様では、構築用オリゴヌクレオチドのプラス鎖およびマイナス鎖を異なる位置でまたは異なる支持体で合成することができる。その後、プラス鎖およびマイナス鎖をチップから別個のプールの中に放出し、制御された方法で混合することができる。別の態様では、階層的なアッセンプリを固体支持体上での構築用オリゴヌクレオチドの近接性によって制御することができる。例えば、相補的な領域を有する2種の構築用オリゴヌクレオチドを相互と近接近で合成することができる。固体支持体からの放出により、相互と近接近に位置するオリゴヌクレオチドは、オリゴヌクレオチドの局所濃度がいっそう高いので、有利に相互作用すると考えられる。典型的な態様では、2種またはそれ以上の種の構築用オリゴヌクレオチドを固体支持体の同じ位置で合成し、それによってその相互作用を促進させることができる(例えば、米国特許第2004/0101894号を参照のこと)。別の態様では、微小流体システムを利用して、反応混合を制御し、アッセンプリ過程を促進することができる。例えば、互いに物理的に分離されている直線の列に、すなわち、流体が流れる別個の直線チャンネルにアレイのフィーチャが並べられているようなチャンネルを含むフローセルの中で、オリゴヌクレオチドを合成することができる。所与のチャンネル中のオリゴヌクレオチドは、同じチャンネル中のその他のオリゴヌクレオチドとハイブリダイズし相互作用できるが、その他のチャンネル由来のオリゴヌクレオチドには曝されないはずである。隣接のオリゴヌクレオチド配列を同じチャンネル中で合成すれば、それらはアレイからの切断後に互いにハイブリダイズして「サブアッセンプリ」を形成することができる。次い

10

20

30

40

50

で、さらに大きな核酸配列をハイブリダイズさせるため、各種のサブアッセンプリをその他のサブアッセンプリと接触させることができる。リガーゼおよび/またはポリメラーゼを必要に応じ添加して、核酸配列中のギャップをフィルインし連結することができる。

#### 【0122】

さらに別の態様では、階層的なアッセンプリを制限エンドヌクレアーゼにより行って、所望の順序でともに連結されうる付着端を形成させることができる。構築用オリゴヌクレオチドは、指定の順序で連結を促進しうる部位に1種または複数種の制限エンドヌクレアーゼに対する認識および切断部位を含むよう設計し合成することができる。DNA二重鎖を形成させた後、オリゴヌクレオチドのプールを1種または複数種の制限エンドヌクレアーゼと接触させて、付着端を形成させることができる。次いで、このプールをハイブリダイゼーションおよびライゲーション条件に曝して、二重鎖をともに連結させる。連結の順序は相補的な付着端のハイブリダイゼーションによって決められるはずである。制限エンドヌクレアーゼは、1度に付着端の1サブセットだけを形成させるよう段階的に加えられてもよい。その後、これらの末端をともに連結し、引き続き別ラウンドのエンドヌクレアーゼ消化、ハイブリダイゼーション、ライゲーションなどを行うことができる。典型的な態様では、IIS型エンドヌクレアーゼ認識部位を構築用オリゴヌクレオチドの末端に組み入れて、IIS型エンドヌクレアーゼによる切断を可能にすることができる。

10

#### 【0123】

オリゴヌクレオチド合成の間に被る変異は、アッセンプルされたDNA分子中のエラーの主因であり、取り除くのは費用がかかり且つ困難である(あらゆる目的でその全体が参照により本明細書に組み入れられるCello et al. (2002) Science 297:1016; Smith et al. (2003) Proc. Natl. Acad. Sci. USA 100:15440)。したがって、種々の態様では、さまざまなエラー低減法を利用して、構築用オリゴヌクレオチド、サブアッセンプリおよび/またはポリヌクレオチド構築体中のエラーを除去することができる。エラー補正法としては、例えば、後述のエラー過法、エラー中和法およびエラー補正法を挙げることができる。

20

#### 【0124】

ミスマッチ結合タンパク質などの、ミスマッチ修復に関与するタンパク質を利用して、正しいヌクレオチド配列を有する合成オリゴヌクレオチドを選択することができる(図34~36)。ミスマッチ修復タンパク質は、種々のDNAミスマッチ、欠失および挿入に結合する(Carr et al. (2004) Nucleic Acids Res. 32:e162)。したがって、ミスマッチ結合タンパク質を利用して、エラーを有する合成オリゴヌクレオチド配列に結合させることができる。エラーがない二本鎖オリゴヌクレオチド配列(例えば、ハイブリダイズした構築用オリゴヌクレオチド、ハイブリダイズした選択用オリゴヌクレオチドおよび/または選択用オリゴヌクレオチドにハイブリダイズした構築用オリゴヌクレオチド)をその後、ミスマッチ結合タンパク質に結合した二本鎖オリゴヌクレオチド配列から分離することができる。このように、エラーを含むオリゴヌクレオチド配列からエラーのないオリゴヌクレオチド配列を効果的に分離することができる。

30

#### 【0125】

「DNA修復」という用語は、損傷したまたは変異した領域を核酸から切り取るヌクレアーゼが核酸(DNA:DNA二重鎖、DNA:RNAおよび、本明細書の目的では、同様にRNA:RNA二重鎖)中の配列エラーを認識し、その後さらに酵素または酵素活性が鎖の交換部分を合成して、正しい配列を作製する過程のことをいう。

40

#### 【0126】

「DNA修復酵素」という用語は、核酸構造および配列中のエラーを補正する1種または複数種の酵素のことをいい、すなわち、この酵素は核酸二重鎖中の異常な塩基対合を認識し、結合して補正する。DNA修復酵素の例としては、MutH、mutL、mutM、mutS、mutY、dam、チミジンDNAグリコシラーゼ(TDG)、ウラシルDNAグリコシラーゼ、AlkA、MLH1、MSH2、MSH3、MSH6、エキソヌクレアーゼI、T4エンドヌクレアーゼV、エキソヌクレアーゼV、RecJエキソヌクレアーゼ、FEN1 (RAD27)、dnaQ (mutD)、polC (dnaE)、またはそれらの組合せな

50

どのタンパク質、およびこれらのホモログ、オルソログ、パラログ、変異体、または断片が挙げられるが、これらに限定されることはない。DNAらせん内部の塩基対合エラーの認識および補正ができる酵素系は、細菌、真菌および哺乳類細胞などで証明されている。

#### 【0127】

本明細書で用いられる「ミスマッチ結合剤」または「MMBA」という用語は、ミスマッチを含む二本鎖核酸分子に結合する薬剤のことをいう。この薬剤は化学性でもまたはタンパク性でもよい。ある種の態様では、MMBAは、例えば、Fok I、MutS、T7エンドヌクレアーゼ、本明細書に記載のDNA修復酵素、米国特許第2004/0014083号に記載の変異DNA修復酵素、またはそれらの断片もしくは融合体などの、ミスマッチ結合タンパク質(MMBP)である。MMBAにより認識されうるミスマッチとしては、例えば、1つまたは複数のヌクレオチド挿入もしくは欠失、またはA:A、A:C、A:G、C:C、C:T、G:G、G:T、T:T、C:U、G:U、T:U、U:U、5-ホルミルウラシル(fU):G、7,8-ジヒドロ-8-オキソ-グアニン(8-オキソG):C、8-オキソG:Aもしくはそれらの相補体などの不適当な塩基対合が挙げられる。

10

#### 【0128】

本明細書で用いられる「MLH1」および「PMS1」(ヒトではPMS2)という用語は、誤対合塩基に結合したMSH2を含む複合体と相互作用する真核生物mutL関連タンパク質複合体、例えば、MLH1-PMS1の構成成分のことをいう。典型的なMLH1タンパク質としては、例えば、以下のGenBankアクセッション番号A1389544(キイロショウジョウバエ)、A1387992(キイロショウジョウバエ)、AF068257(キイロショウジョウバエ)、U80054(ドブネズミ(*Rattus norvegicus*))およびU07187(出芽酵母(*S. cerevisiae*))を有する核酸によりコードされるポリペプチド、ならびにそのホモログ、オルソログ、パラログ、変異体、または断片が挙げられる。

20

#### 【0129】

本明細書で用いられる「MSH2」という用語は、塩基ミスマッチおよび12塩基までの挿入または欠失を認識する真核生物DNA修復複合体の構成成分のことをいう。MSH2はMSH3またはMSH6とヘテロ二量体を形成する。MSH2タンパク質としては、例えば、以下のGenBankアクセッション番号AF109243(シロイヌナズナ(*A. thaliana*))、AF030634(アカパンカビ(*N eurospora crassa*))、AF002706(シロイヌナズナ)、AF026549(シロイヌナズナ)、L47582(ヒト(*H. sapiens*))、L47583(ヒト)、L47581(ヒト)およびM84170(出芽酵母)を有する核酸によりコードされるポリペプチドならびにそのホモログ、オルソログ、パラログ、変異体、または断片が挙げられる。MSH3タンパク質としては、例えば、GenBankアクセッション番号J04810(ヒト)およびM96250(出芽酵母)を有する核酸によりコードされるポリペプチドならびにそのホモログ、オルソログ、パラログ、変異体、または断片が挙げられる。MSH6タンパク質としては、例えば、以下のGenBankアクセッション番号U54777(ヒト)およびAF031087(ハツカネズミ(*M. musculus*))を有する核酸によりコードされるポリペプチドならびにそのホモログ、オルソログ、パラログ、変異体、または断片が挙げられる。

30

#### 【0130】

本明細書で用いられる「mutH」という用語は、ヘミメチル化DNAの非メチル化鎖に切り込みを入れる、またはd(GATC)配列のGの5'側、非メチル化DNAで二本鎖切断を行う潜在性エンドヌクレアーゼのことをいう。この用語は原核生物mutH(例えば、Welsh et al., 26 2 J. Biol. Chem. 15624 (1987))およびそのホモログ、オルソログ、パラログ、変異体、または断片を含むよう意図される。

40

#### 【0131】

本明細書で用いられる「mutHLS」という用語は、mutH、mutLおよびmutSタンパク質(またはそのホモログ、オルソログ、パラログ、変異体、もしくは断片)の間の複合体のことをいう。

#### 【0132】

本明細書で用いられる「mutL」という用語は、ATP依存的に5'-GATC-3'配列でのmutH切断とmutSによる異常な塩基対合認識を共役させるタンパク質のことをいう。この用語は原核生物mutLタンパク質およびそのホモログ、オルソログ、パラログ、変異体、または断片

50

を包含するよう意図される。MutLタンパク質としては、例えば、以下のGenBankアクセッション番号AF170912 (*C. クレセントス* (*C. crescentus*)), AI518690 (キイロシヨウジョウバエ)、AI456947 (キイロシヨウジョウバエ)、AI389544 (キイロシヨウジョウバエ)、AI387992 (キイロシヨウジョウバエ)、AI292490 (キイロシヨウジョウバエ)、AF068271 (キイロシヨウジョウバエ)、AF068257 (キイロシヨウジョウバエ)、U50453 (*T. アクアチクス* (*T. aquaticus*)), U27343 (枯草菌 (*B. subtilis*)), U71053 (U71053 (*T. マリチマ* (*T. maritima*))), U71052 (*A. ピロフィラス* (*A. pyrophilus*)), U13696 (ヒト)、U13695 (ヒト)、M29687 (*S. ティフィムリウム* (*S. typhimurium*)), M63655 (大腸菌 (*E. coli*))およびL19346 (大腸菌)を有する核酸によりコードされるポリペプチドが挙げられる。MutLホモログとしては、例えば、真核生物MLH1、MLH2、PMS1、およびPMS2タンパク質(例えば、その全体が参照により本明細書に組み入れられる米国特許第5,858,754号および同第6,333,153号を参照のこと)が挙げられる。

10

#### 【0133】

本明細書で用いられる「mutS」という用語は、さまざまな誤対合塩基や小さな(1~5塩基)一本鎖ループを認識し、これらに結合するDNAミスマッチ結合タンパク質のことをいう。この用語は原核生物mutSタンパク質およびそのホモログ、オルソログ、パラログ、変異体、または断片を包含するよう意図される。この用語は同様に、さまざまなmutSタンパク質のホモおよびヘテロ二量体ならびに多量体を包含する。MutSタンパク質としては、例えば、以下のGenBankアクセッション番号AF146227 (ハツカネズミ)、AF193018 (シロイヌナズナ)、AF144608 (腸炎ビブリオ (*V. parahaemolyticus*)), AF034759 (ヒト)、AF104243 (ヒト)、AF007553 (*T. アクアチクス・カルドフィラス* (*T. aquaticus caldophilus*)), AF109905 (ハツカネズミ)、AF070079 (ヒト)、AF070071 (ヒト)、AH006902 (ヒト)、AF048991 (ヒト)、AF048986 (ヒト)、U33117 (*T. アクアチクス*)、U16152 (腸炎エルシニア (*Y. enterocolitica*)), AF000945 (コレラ菌 (*V. cholerae*)), U698873 (大腸菌)、AF003252 (*H. インフルエンザ菌* (*H. influenzae*))株b型(Eagan)、AF003005 (シロイヌナズナ)、AF002706 (シロイヌナズナ)、L10319 (ハツカネズミ)、D63810 (*T. サーモフィルス* (*T. thermophilus*)), U27343 (枯草菌)、U71155 (*T. マリチマ*)、U71154 (*A. ピロフィラス*)、U16303 (ネズミチフス菌)、U21011 (ハツカネズミ)、M84170 (出芽酵母)、M84169 (出芽酵母)、M18965 (ネズミチフス菌)およびM63007 (窒素固定細菌 (*A. vinelandii*))を有する核酸によりコードされるポリペプチドが挙げられる。MutSホモログとしては、例えば、真核生物MSH2、MSH3、MSH4、MSH5、およびMSH6タンパク質(例えば、米国特許第5,858,754号および同第6,333,153号を参照のこと)が挙げられる。

20

30

#### 【0134】

1つの局面では、本発明は、エラーを含むポリヌクレオチドのコピーを1種または複数種の選択用オリゴヌクレオチドとのハイブリダイゼーションを通じ除去することで、ポリヌクレオチドプールの忠実度を高める方法を提供する。この種のエラーを過工程をアッセンプリの任意段階のオリゴヌクレオチドで、例えば、構築用オリゴヌクレオチドで、サブアッセンプリで、および場合によってはさらに大きなポリヌクレオチド構築体で行うことができる。選択用オリゴヌクレオチドを用いたエラーを過は、ポリヌクレオチドプールの増幅の前におよび/または後に行うことができる。典型的な態様では、選択用オリゴヌクレオチドを用いたエラーを過を利用して、増幅の前におよび/または後に構築用オリゴヌクレオチドのプールの忠実度を高める。選択用オリゴヌクレオチドとのハイブリダイゼーションを通じたエラーを過の例示的態様を図32に示す。構築用オリゴヌクレオチドのプールをユニバーサルプライマーによって増幅した。一部の構築用オリゴヌクレオチドは、鎖の中で隆起により表されるエラーを含む。これらのエラーは構築用オリゴヌクレオチドの初期合成から生じていることがありまたは増幅過程の間に導入されていることがある。その後、構築用オリゴヌクレオチドのプールを変性させて一本鎖を作製し、ハイブリダイゼーション条件の下で選択用オリゴヌクレオチドの少なくとも1プールと接触させる。選択用オリゴヌクレオチドのプールは、プール中の構築用オリゴヌクレオチドの各々に相補的な1種または複数種の選択用オリゴヌクレオチドを含む(例えば、選択用オリゴのプールは構

40

50

築用オリゴヌクレオチドのプールと少なくとも同じ規模であり、および場合によっては、例えば、構築用オリゴヌクレオチドのプールと比べて2倍多くの異なるオリゴヌクレオチドを含むことができる)。選択用オリゴヌクレオチドと完全には対合しない構築用オリゴヌクレオチドのコピー(例えば、ミスマッチが存在する)は、完全に適合するコピーほど強固にはハイブリダイズしないはずであり、ハイブリダイゼーション条件のストリンジェンシーを制御することによりプールから除去することができる。ミスマッチを含むオリゴヌクレオチドの除去後、完全に適合する構築用オリゴヌクレオチドコピーは、ストリンジェンシー条件を増大させて、それらを選択用オリゴヌクレオチドから流出させることにより除去することができる。典型的な態様では、選択用オリゴヌクレオチドを末端固定化して(例えば、化学結合、ビオチン/ストレプトアビジンなどを介し)、エラーを含むオリゴヌクレオチドコピーの除去を容易にすることができる。例えば、構築用オリゴヌクレオチドのプールとのハイブリダイゼーションの前にまたは後に、選択用オリゴヌクレオチドをビーズに固定化することができる。その後、ビーズをペレットにし、またはカラムに負荷し、異なるストリンジェンシー条件に曝して、選択用オリゴヌクレオチドを使いミスマッチを含む構築用オリゴヌクレオチドのコピーを除去することができる。ある種の態様では、オリゴヌクレオチドを反復ラウンドの増幅および選択用オリゴヌクレオチドのプールとのハイブリダイゼーションを通じたエラーを過に掛けて、それによりプールの忠実度を維持しながら、または好ましくは増大させながらプール中のオリゴヌクレオチドのコピー数を増加させる(例えば、プール中のエラーのないコピー数を増加させる)ことが望ましいかもしれない。

10

20

## 【0135】

場合によっては、構築および選択用オリゴヌクレオチド間のミスマッチは選択用オリゴヌクレオチド中の配列エラーに起因することがあり、それによってエラーのない構築用オリゴヌクレオチドをプールから除去することに留意されるべきである。しかしながら、正味の効果はやはり、構築用オリゴヌクレオチドプールの忠実度の増大と考えられる。

## 【0136】

図34は、二本鎖構築用オリゴヌクレオチド、サブアッセムブリおよび/またはポリヌクレオチド構築体のプールの忠実度を増大させるのに利用できるエラーを過の典型的な別法を図解している。一本鎖DNA中のエラーがDNA二重鎖中のミスマッチを引き起こす。MutSの二量体などの、ミスマッチ結合タンパク質(MMBP)はDNAのこの部位に結合する。図34Aに示されるように、DNA二重鎖のプールは、ミスマッチが有る二重鎖(左)とエラーなしのもの(右)とを含む。各DNA鎖の3'末端が矢印により示されている。エラーが引き起こすミスマッチは、左上鎖にて隆起した三角形のバンプとして示されている。図34Bに示されるように、ミスマッチの部位に選択的に結合するMMBPを加えることができる。その後、MMBPに結合されたDNA二重鎖を除去し、エラーなしの二重鎖が劇的に濃縮されているプールを残すことができる(図34C)。1つの態様では、DNAに結合されたタンパク質は、エラーを含むDNAをエラーなしのコピーから分離する手段となる(図34D)。タンパク質-DNA複合体は、例えば、特異抗体、固定化ニッケルイオン(タンパク質はhis-タグ融合体として産生させる)、ストレプトアビジン(タンパク質はビオチンの共有結合付加によって修飾してある)またはタンパク質精製の分野に普及しているようなその他の機構で官能化された固体支持体にタンパク質の親和性によって捕捉することができる。あるいは、タンパク質-DNA複合体は、例えば、サイズ排除カラムクロマトグラフィーを用いてまたは電気泳動によって移動度の違いでエラーなしのDNA配列のプールから分離される(図34E)。この例では、ゲル中の電気泳動移動度がMMBP結合により変化する、つまりMMBPの非存在下では全ての二重鎖が一緒に移動するが、MMBPの存在下では、ミスマッチ二重鎖は遅れる(上側のバンド)。その後、ミスマッチのないバンド(下側)を切り出し抽出する。

30

40

## 【0137】

図35はミスマッチ結合剤を用いて配列エラーを中和する典型的な方法を図解している。この種のエラー低減法は、二本鎖構築用オリゴヌクレオチド、サブアッセムブリおよび/またはポリヌクレオチド構築体のプールの忠実度を増大させるのに有用とすることができ

50

る。この態様では、エラーを含むDNA配列はDNA産物のプールから除去されない。むしろ、その配列は化学架橋剤(例えば、スベルイミノ酸ジメチル、DMS)の、または別のタンパク質(MutLなど)の作用によりミスマッチ認識タンパク質と不可逆的に複合体をなすようになる。その後、DNA配列のプールを増幅させる(例えば、ポリメラーゼ連鎖反応、PCRにより)が、エラーを含むものは増幅を妨害され、エラーなしの配列の増大によって急速に数で劣るようになる。図35Aは、ミスマッチが有る二重鎖(左)とエラーなしのもの(右)とを含むDNA二重鎖の典型的なプールを図解している。MMBPを利用して、ミスマッチを含むDNA二重鎖に選択的に結合させることができる(図35B)。MMBPは架橋剤の適用によってミスマッチの部位に不可逆的に結合させることができる(図35C)。共有結合しているMMBPの存在下において、DNA二重鎖のプールの増幅は、エラーのない二重鎖のコピーをより多くもたらす(図35D)。結合タンパク質は二重鎖の2本の鎖が解離するのを妨害するので、MMBP-ミスマッチDNA複合体は増幅に関与することができない。長いDNA二重鎖の場合、MMBP結合部位の外側の領域は部分的に解離し、その(エラーのない)領域の部分的増幅に関与することができるかもしれない。

#### 【0138】

ますます長いDNA配列が作製されるにつれて、完全にエラーがない配列の画分は減少する。ある長さで、完全に正しい配列を含んだ分子が全プール中に存在しなくなるという可能性が高くなる。すなわち、極端に長いDNAセグメントの作製の場合、上記のエラー制御法にかけられるもっと短い単位を最初に産出することが有用になりうる。その後、これらのセグメントを結合させて、より長い完全長の産物を得ることができる。しかしながら、長いDNA二重鎖全体を除去するまたは中和することなく、これらの極端に長い配列中のエラーを局所的に補正できれば、より複雑な段階的アッセムブリの過程を回避することができる。

#### 【0139】

多くの生物学的DNA修復機構は、変異(エラー)部位を認識することおよび正しくない配列を交換するために鋳型鎖(たいがいエラーがない)を利用することに依る。DNA配列の新規生成では、どちらの鎖がエラーを含むかおよびどちらが鋳型として使われるべきかを判定する難しさから、この過程は複雑である。この問題の解決方法は、混合物の中に、補正用の鋳型を供与する他の配列プールを用いることに依る。これらの方法は非常に強力でありうる、つまり、たとえすべてのDNA鎖が1つまたは複数のエラーを含んでいても、鎖の大部分が各位置に正しい配列を有する限り(エラーの位置には一般に鎖間で相関性がないと予想されるので)、所与のエラーは正しい配列と置き換えられるという高い可能性がある。

#### 【0140】

図36は、鎖特異的なエラー補正を行う典型的な方法を図解している。複製生物では、酵素を介したDNAメチル化は、鋳型(親)DNA鎖を同定するために使われることが多い。新しく合成された(娘)鎖が最初に非メチル化される。ミスマッチが検出される場合には、二重鎖DNAのヘミメチル化状態を利用して、ミスマッチ修復系に娘鎖だけを補正するよう指令する。しかしながら、相補的DNA鎖のペアの新規合成では、両鎖が非メチル化されてしまい、修復系にはどちらの鎖を補正すべきかを選択するのに本来備わっている基準がない。本発明のこの局面では、メチル化および部位特異的な脱メチル化を利用して、選択的にヘミメチル化されたDNA鎖を作製する。大腸菌のDamメチル化酵素などの、メチル化酵素を利用して、各鎖の潜在的な標的部位の全てを一様にメチル化する。その後、DNA鎖を解離させ、新たなパートナー鎖と再アニーリングさせる。新たなタンパク質、つまりミスマッチ結合タンパク質(MMBP)の脱メチル化酵素との融合体を加える。この融合タンパク質はミスマッチにのみ結合し、脱メチル化酵素の近接性により両鎖から、しかしミスマッチ部位の近傍だけでメチル基が除去される。その後の解離とアニーリングのサイクルによって、(脱メチル化された)エラー含有鎖をその配列のこの領域にエラーがない(メチル化された)鎖と会合させる。(相補鎖のエラー位置には相関性がないので、これは鎖の大部分に当てはまるはずである。)ヘミメチル化されたDNA二重鎖はその時点で、エラーの修復を指令する

のに必要な情報の全てを含んでおり、MutS、MutL、MutHおよびこの目的のDNAポリメラーゼタンパク質を利用する、大腸菌のものなどの、DNAミスマッチ修復系の成分を利用することができる。この過程を複数回繰り返して、全てのエラーが補正されるのを確実にすることができる。

#### 【0141】

図36Aは、ミスマッチを引き起こす、左上鎖中の1塩基エラーを除いては同一である2つのDNA二重鎖を示す。右側二重鎖の鎖は極太線で示されている。次に、メチル化酵素(M)を利用して、各DNA鎖の可能な部位を全て一様にメチル化することができる(図36B)。その後、メチル化酵素を除去し、ミスマッチ結合タンパク質(MMBP)および脱メチル化酵素(D)の両方を含む融合タンパク質を加える(図36C)。融合タンパク質のMMBP部分はミスマッチ部位に結合し、融合タンパク質をミスマッチ部位に局在化させる。次いで、融合タンパク質の脱メチル化酵素部分はミスマッチの近くで両鎖からメチル基を特異的に除去するように作用することができる(図36D)。その後、MMBP-Dタンパク質融合体を除去させることができ、DNA二重鎖を解離させ、新たなパートナー鎖と再会合させることができる(図36E)。エラー含有鎖は、a) その部位に相補的なエラーを含んでいない。およびb) ミスマッチ部位の近傍でメチル化されている、相補鎖と再会合する可能性が最も高いはずである。この新しい二重鎖はその時点で、DNAミスマッチ修復系の天然基質を模倣している。次に、ミスマッチ修復系の成分(大腸菌MutS、MutL、MutHおよびDNAポリメラーゼなどの)を利用して、エラー含有(エラーを含む)鎖の中の塩基を除去することができ、代わりに合成するための鋳型として反対(エラーなし)の鎖を利用し、補正された鎖を残すことができる(図36F)。

10

20

#### 【0142】

1つの態様では、検出されるおよび補正されるエラーの数は、エラー低減の前にDNA二重鎖のプールを融解することおよび再アニーリングすることにより増やすことができる。例えば、問題のDNA二重鎖がポリメラーゼ連鎖反応(PCR)などの技術によって増幅されたなら、新しい(完全に)相補的な鎖の合成は、これらのエラーが直ちにDNAミスマッチとして検出可能とされるわけではないことを意味すると考えられる。しかしながら、これらの二重鎖を融解し、この鎖を新しい(および無作為の)相補的パートナーと再会合させることで、大部分のエラーがミスマッチとして明らかであるような二重鎖を作製することができよう(図37)。エラー制御の各サイクルは一部のエラーのない配列も除去してしまうことがあるので(プールをエラーのない配列に比例的に濃縮しながらも)、エラー制御とDNA増幅の交互サイクルを利用して、大きな分子プールを維持することができる。

30

#### 【0143】

ミスマッチ結合タンパク質が結合したオリゴヌクレオチド配列は、ゲル電気泳動、アフィニティーカラム、免疫学的方法および同様のものを含むが、これらに限定されない当該技術分野において知られている様々な方法を用いて、未結合のオリゴヌクレオチド配列から分離することができる。

#### 【0144】

ゲル電気泳動は、電界の影響のもとゲル媒体中での移動に基づいてDNA-タンパク質複合体を複合体未形成のDNAから分離できる別の方法である。DNA-タンパク質複合体は、複合体未形成のDNAよりも遅い移動速度を示し、このように複合体未形成のDNAから分離することができる。複合体未形成のDNAは、当該技術分野において知られている様々な方法を用いてゲルから取り出すことができる(あらゆる目的でその全体が参照により本明細書に組み入れられる Ausubel et al(編)., 1992, current protocols in Molecular Biology, John Wiley & Sons, New York)。

40

#### 【0145】

本発明は同様に、エラーを含むオリゴヌクレオチド配列のアフィニティー分画による試料内でのエラーなしのオリゴヌクレオチド配列の選択的濃縮を提供する。ミスマッチ結合タンパク質が結合したオリゴヌクレオチド配列は、ミスマッチ結合タンパク質を結合する固体支持体を利用したアフィニティー分画により未結合のオリゴヌクレオチドから分離す

50

ることができる。オリゴヌクレオチド配列-ミスマッチ結合タンパク質複合体は、複合体を結合できる任意の成分、例えば、結合タンパク質特異的なまたは複合体特異的な抗体がカップリングされている充填剤によって選択的に保持される。この過程を繰り返して、エラーがほとんどまたは全くないオリゴヌクレオチド配列を溶出液中でさらに濃縮することができる。

#### 【0146】

抗体がミスマッチ結合タンパク質またはオリゴヌクレオチド配列-ミスマッチ結合タンパク質複合体に直接結合する抗体支持体に加えて、その他のアフィニティー支持体が使われてもよい。例えば、1つには、固定化金属アフィニティークロマトグラフィーによってポリペプチド中のヒスチジン残基に結合する金属、例えば、ニッケル・カラムの能力を利用することができる。ヒスチジン尾部、例えば、ヒスチジン6残基をHochuliら(あらゆる目的でその全体が参照により本明細書に組み入れられる(1988) *Biotechnology* 6:1321)により記述されているように、ミスマッチ結合タンパク質のアミノ末端に共有結合させることができる。オリゴヌクレオチド配列-ミスマッチ結合タンパク質複合体をニッケル・カラムに加えると、結合タンパク質のヒスチジン部分がカラムに結合されるはずである。

#### 【0147】

アフィニティー支持体の別の例は、flag配列、すなわち、抗体が特異的に結合する任意のアミノ酸配列(例えば、10残基)を認識し、その配列に結合する抗体が結合されている支持体である。flag配列はミスマッチ結合タンパク質のアミノ末端に遺伝子工学的に作製することができる。オリゴヌクレオチド配列-結合タンパク質複合体を抗体カラムに加えると、抗体は結合タンパク質中のflag配列に結合し、したがって複合体を保持できる。The Flag Biosystemとして知られるこの技術の1つの態様は、International Biotechnologies社(New Haven, Conn.)から市販されている。さらに大きなflag配列、例えば、マルトース結合タンパク質を使用することもできる(あらゆる目的でその全体が参照により本明細書に組み入れられる Ausubel et al(編)., 1992, *current protocols in Molecular Biology*, John Wiley & Sons, New York)。

#### 【0148】

本発明で有用な固体支持体は、多種多様な支持体のいずれか1つとすることができ、合成高分子支持体、例えば、ポリスチレン、ポリプロピレン、置換ポリスチレン(例えば、アミノ化またはカルボキシル化ポリスチレン)、ポリアクリルアミド、ポリアミド、ポリ塩化ビニルおよび同様のもの、ガラスビーズ、高分子ビーズ、セファロース、アガロース、セルロース、またはアフィニティークロマトグラフィーに有用な任意の材料を含むことができるが、これらに限定されることはない。支持体を反応基、例えば、カルボキシル基、アミノ基などと共に供与して、支持体とのタンパク質の直接結合を可能にしてもよい。ミスマッチ結合タンパク質が支持体に直接架橋されてもよく、またはミスマッチの結合タンパク質もしくは核酸/結合タンパク質複合体を結合できるタンパク質(例えば、抗体)が支持体にカップリングされてもよい。

#### 【0149】

例えば、支持体はセファロースビーズを含み、ミスマッチ結合タンパク質はビーズにカップリングされるならば、結合タンパク質が結合したビーズをカラムの中に詰め込み、平衡化し、このカラムを核酸試料にさらす。適切な結合条件の下で、カラム中のビーズにカップリングされているタンパク質は、これが認識する核酸断片またはタンパク質/核酸複合体を保持する。

#### 【0150】

吸着、共有結合を含め様々な技術により、例えば、支持体の活性化により、または適当なカップリング剤の使用もしくは支持体上での反応基の使用により、タンパク質を支持体に結合させることができる。そのような手順は一般的に当技術分野において公知であり、さらなる詳細が本発明の完全な理解に必要とは考えられない。適当なカップリング剤の代表例は、ジアルデヒド、例えば、グルタルアルデヒド、スクシナルデヒドまたはマロンアルデヒド、不飽和アルデヒド、例えば、アクロレイン、メタクロレインまたはクロトン

10

20

30

40

50

アルデヒド、カルボジイミド、ジイソシアン酸エステル、アジピン酸ジメチルおよび塩化シアヌルである。適当なカップリング剤の選択は、本明細書の教示から当業者には明らか  
なはずである。

【0151】

オリゴヌクレオチド配列-ミスマッチ結合タンパク質複合体のアフィニティー精製の別の形態には、Ausubel (1992、前記、あらゆる目的でその全体が参照により本明細書に組み入れられる)に記述されている、タンパク質を結合するがタンパク質のない核酸を結合しないニトロセルロースフィルターの使用が含まれる。

【0152】

エラーを有する合成オリゴヌクレオチドを検出する別の適当な方法は、ミスマッチ結合タンパク質に対するモノクローナルまたはポリクローナル抗体などの抗体を用いた免疫学的方法を介するものである。抗ミスマッチ結合タンパク質抗体を利用して、アフィニティークロマトグラフィー(前記)または免疫沈降などの標準的な技術により、ミスマッチ結合タンパク質-オリゴヌクレオチド配列複合体を複合体未形成のオリゴヌクレオチド配列から分離することができる。

10

【0153】

免疫沈降の場合、抗原(すなわち、ミスマッチ結合タンパク質)、1次抗体およびプロテインA-、G-もしくはL-基材複合物または2次抗体-基材複合物を含む免疫複合体によって、ミスマッチ結合タンパク質を沈降させる。基材はアガロース、(例えば、磁気、ガラス、高分子)ビーズ、細胞(例えば、黄色ブドウ球菌(*S. aureus*))および同様のものを含むが、これらに限定されることはない。アガロース複合物の選択は種起源や1次抗体のアイソタイプに依る。免疫沈降の試薬およびプロトコルは市販されている(例えば、Sigma-Aldrich社)。

20

【0154】

本明細書で用いられる「抗体」という用語は、免疫グロブリン分子および免疫グロブリン分子の免疫学的に活性な部分、すなわち、特異的にミスマッチ結合タンパク質などの抗原を結合する(抗原と免疫反応する)抗原結合部位を含む分子のことをいう。免疫グロブリン分子の免疫学的に活性な部分の例には、抗体をペプシンなどの酵素で処理することによって作製できるF(ab)およびF(ab')<sub>2</sub>断片が含まれる。本発明は、ミスマッチ結合タンパク質を結合するポリクローナルおよびモノクローナル抗体を提供する。本明細書で用いられる「モノクローナル抗体」という用語は、ミスマッチ結合タンパク質の特定の抗原決定基と免疫反応できる抗原結合部位を1種しか含まない抗体分子の集団のことをいう。

30

【0155】

ポリクローナル抗体は、適当な被験体にミスマッチ結合タンパク質免疫原を免疫することによって調製することができる。免疫された被験体において抗ミスマッチ結合タンパク質抗体の力価を標準的な技術により、例えば、固定化されたミスマッチ結合タンパク質を利用する酵素結合免疫測定法(ELISA)を用いて長期にわたりモニターすることができる。必要に応じて、ミスマッチ結合タンパク質に対して作製された抗体分子を哺乳類から(例えば、血液から)分離し、プロテインAクロマトグラフィーなどの周知の技術によりさらに精製して、IgG分画を得ることができる。

40

【0156】

免疫化後の適当な時点で、例えば、抗ミスマッチ結合タンパク質抗体の力価が最も高い場合に、抗体産生細胞を被験体から入手し使用して、もとはKohlerおよびMilstein ((1975) *Nature* 256:495-497) (同様にBrown et al. (1981) *J. Immunol.* 127:539-46; Brown et al. (1980) *J. Biol. Chem.* 255:4980-83; Yeh et al. (1976) *Proc. Natl. Acad. Sci. U.S.A.* 76:2927-31; およびYeh et al. (1982) *Int. J. Cancer* 29:269-75を参照のこと)によって報告されたハイブリドーマ技術、もっと最近のヒトB細胞ハイブリドーマ技術(Kozbor et al. (1983) *Immunol. Today* 4:72)、EBV-ハイブリドーマ技術(Cole et al. (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96)またはトリオーマ技術などの標準的な技術によりモノクローナル抗体を調製することができ

50

る。モノクローナル抗体ハイブリドーマを産生する技術は周知である(一般的にはR. H. Kenneth, *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, N.Y. (1980); Lerner (1981) *Yale J. Biol. Med.* 54:387-402; M. L. Gefter et al. (1977) *Somatic Cell Genet.* 3:231-36を参照のこと)。手短に言えば、上記のようにミスマッチ結合タンパク質免疫原を免疫した哺乳類からのリンパ球(通常、脾細胞)に不死化細胞系(通常、骨髄腫)を融合し、得られたハイブリドーマ細胞の培養上清をスクリーニングして、ミスマッチ結合タンパク質を結合するモノクローナル抗体を産生するハイブリドーマを同定する。上記の各参考文献はあらゆる目的でその全体が参照により本明細書に組み入れられる。

#### 【0157】

ある種の態様では、DNA配列決定法、ハイブリダイゼーションに基づく診断方法、分子生物学技術、例えば、制限消化、選択マーカーアッセイ法、インビボでの機能的選択、またはその他の適当な方法により、サブアセンブリおよび/または合成ポリヌクレオチド構築体のアセンブリ成功を評価することが望ましいかもしれない。例えば、ポリヌクレオチド構築体を細胞に導入し、構築体上の1つまたは複数のポリヌクレオチドの発現をアッセイすることにより、機能的選択を行うことができる。検出可能なマーカー、選択可能なマーカー、所与のサイズのポリペプチドをアッセイすることにより(例えば、サイズ排除クロマトグラフィー、ゲル電気泳動などにより)、またはポリヌクレオチド構築体によりコードされる1つまたは複数のポリペプチドの酵素機能をアッセイすることにより、アセンブリ成功を判定することができる。DNA操作および酵素処理は、当技術分野において確立されたプロトコルおよび製造元の推奨手順にしたがって行う。適当な技術はSambrookら(第2版), *Cold Spring Harbor Laboratory, Cold Spring Harbor* (1982, 1989); *Methods in Enzymol.* (Vols. 68, 100, 101, 118および152-155) (1979, 1983, 1986および1987); およびDNA Cloning, D.M. Glover(編), IRL Press, Oxford (1985)に記述されている。

#### 【0158】

ある種の態様では、ポリヌクレオチド構築体を発現ベクターに導入し、宿主細胞にトランスフェクションすることができる。宿主細胞は任意の原核または真核細胞とすることができる。例えば、本発明のポリペプチドを細菌細胞、例えば、大腸菌、昆虫細胞(バキュロウイルス)、酵母、植物または哺乳類細胞で発現させることができる。ポリペプチドの発現を最適化するため、宿主には通常見られないtRNA分子を宿主細胞に補充してもよい。ポリヌクレオチド構築体を発現ベクターに連結すること、および真核生物(酵母、鳥類、昆虫または哺乳類)または原核生物(細菌細胞)のいずれかの宿主に形質転換することまたはトランスフェクションすることは、標準的な手順である。大腸菌などの原核細胞での発現に適した発現ベクターの例としては、例えば、次のタイプのプラスミド: pBR322-由来プラスミド、pEMBL-由来プラスミド、pEX-由来プラスミド、pBTac-由来プラスミドおよびpUC-由来プラスミドが挙げられ; 酵母での発現に適した発現ベクターとしては、例えば、YEP24、YIP5、YEP51、YEP52、pYES2およびYRP17が挙げられ; ならびに哺乳類細胞での発現に適した発現ベクターとしては、例えば、pcDNA1/amp、pcDNA1/neo、pRc/CMV、pSV2gpt、pSV2neo、pSV2-dhfr、pTk2、pRSVneo、pMSG、pSVT7、pko-neoおよびpHyg由来ベクターが挙げられる。

#### 【0159】

本発明の態様は、異なるプライマー配列を有する複数の異なるポリヌクレオチドを含有する少なくとも1つの容器(すなわち、構築用容器)、およびプライマーを含有する容器(すなわち、プライマー用容器)を提供する製造物品(例えば、キット、自動システム)にさらに向けられる。ある種の局面では、製造物品は複数の異なるポリヌクレオチドを含有する少なくとも1つの容器を含み、プライマーは利用者によって供与される。さまざまなプライマーの組合せを選択して、特定のポリヌクレオチド配列を増幅することができる。各ポリヌクレオチドは増幅プライマーの固有のセットを含むので、各種の異なるポリヌクレオチドを単一の容器から取り出すことができる。ある種の局面では、複数の異なるポリヌク

10

20

30

40

50

レオチドはネスティッドプライマー配列を含む。ポリヌクレオチド用容器は $10^2$ 、 $10^3$ 、 $10^4$ 、 $10^5$ 、 $10^6$ 、 $10^7$ 、 $10^8$ 、 $10^9$ 、 $10^{10}$ またはそれ以上の異なるポリヌクレオチド配列を含むことができる。

【0160】

容器を供与する製造物品の部分は、各種のプラスチック、ポリマー、ガラスおよびその組合せを含むがこれらに限定されない、当技術分野において知られる様々な材料から製造されてもよく、例えば、マイクロタイタープレート(384ウェルプレート)、マイクロチップ、管(例えば、PCR管、微量遠心管、試験管、組織培養プレートなど)および同様のものの形態であってもよい。

【0161】

ある種の局面では、複数の異なるポリヌクレオチドおよび/またはプライマーは1つまたは複数の容器に共有結合される。したがって、本明細書において提供される製造物品は、ポリメラーゼおよびヌクレオチドと共に増幅されることを望むポリヌクレオチド配列に特異的なさらなるプライマーペアを付加するだけで、1つまたは複数のポリヌクレオチド配列および/またはプライマーセットを繰り返し増幅できるという点で再利用可能である。適当な増幅の方法は本明細書にさらに記述されている。本明細書に記述される製造物品は、遺伝子、遺伝子セット、ゲノム、ベクターおよび同様のものに対応するポリヌクレオチドを増幅するのに有用である。

【0162】

本明細書に記述される合成ポリヌクレオチドを作製する方法のどれも自動増幅系を用いて行うことができる。ある種の局面では、本明細書に記述される製造物品の少なくとも一部分は自動部品を含む。したがって、製造物品はデータ保存庫(例えば、供与されるポリヌクレオチドおよび/またはプライマーペアを収蔵する)、増幅されるポリヌクレオチドまたはポリヌクレオチドの群を利用者が指定するのを可能にするインターフェース、ならびにインターフェースで入力された仕様に応答する自動的手段を含むことができる。1つまたは複数の構築用容器からおよび1つまたは複数のプライマー用容器からポリヌクレオチドの一定分量を抽出するため、命令をデータ保存庫からアクセスされ、1つまたは複数の増幅ポリヌクレオチド配列を調製することができる。

【0163】

本発明の態様は、遺伝子およびオリゴヌクレオチド配列の設計を自動化するためのコンピュータソフトウェアの使用を含む。そのようなソフトウェアは、ポリヌクレオチド合成を手動で行う個人と一緒にもしくは半自動的に使用されてもよくまたは自動合成システムと併用されてもよい。少なくともいくつかの態様では、遺伝子/オリゴヌクレオチド設計ソフトウェアは、JAVAプログラミング言語で書かれたプログラムにより実行される。このプログラムは、WINDOWS XPオペレーティング・システムのコマンド・プロンプトから実行できる実行ファイルにコンパイルすることができる。このソフトウェア(コンピュータ支援設計-ポリメラーゼアッセンプリ多重化を表す「CAD-PAM」と命名)の操作は、本章および図7~27に記述されている。しかしながら、CAD-PAMは本発明の様々な局面の単なる1態様にすぎない。特許請求の範囲に具体的に記述されていない限り、本発明はCAD-PAMの全特徴を含む実装にまたは同じアルゴリズム、組織構成もしくはCAD-PAMのその他の特定の40 特徴を利用する実装に限定されることはない。本発明は同様に、特定のプログラミング言語、オペレーティング・システム環境またはハードウェア・プラットフォームを利用する実装に限定されることはない。

【0164】

図7はCAD-PAMプログラムの操作を示すフローチャートである。このプログラムは入力を2回受ける。第1(ブロック10)は、選択および構築用オリゴヌクレオチドを設計するための、FASTAフォーマットの、1つまたは複数のヌクレオチド配列(例えば、遺伝子配列)を含んだファイル(「配列.txt」)である。図8は入力配列ファイルの例を示す。配列rs-1中に示される長方形は、後述される配列の部分の部分を指し示すために含まれている。図8に示されるファイルは2つの配列(rs-1およびrs-2)を含むが、単一の配列(または3つ以上の配列)しか 50

10

20

30

40

50

入力することができない。CAD-PAMへの第2の入力は、オリゴヌクレオチドの設計を制御するパラメータを含んだファイル(「cadpam.特性」、ブロック12)ある。

#### 【0165】

図9Aおよび9Bはcadpam.特性の入力パラメータファイルの例を示す。図9Aから始めて、括弧102に示されるとおり、第1パラメータ(「最適化する=」)は、入力配列の1つまたは複数のヌクレオチド配列を発現できる生物により最も頻繁に使われるコドンに基づいて、入力配列(図8、配列.txtファイル中の)を修飾するかどうかを指定する。図9Aの例では、このパラメータは「最適化する=オフ」に設定されている。したがって、入力配列は修飾されないはずである。このパラメータが「オン」(「最適化する=オン」)に設定された場合、以下により詳細に記述されるとおり、発現生物によって使われるコドンに基づいて、入力配列は修飾されると考えられる。発現生物に関する情報は、利用者により別のファイルで供給される。ファイルが指定されていない場合、デフォルトの生物(例えば、大腸菌K12)に関する情報が使われる。デフォルトの生物ではない場合のファイル名は、次のパラメータ(括弧104に示される「コドンファイル=」)として付与される。そのようなファイルの内容は以下にさらに論じられる。

10

#### 【0166】

図9A中の次の入力パラメータは「配列を除去する」(括弧106)である。このパラメータは、入力配列から除去されることになるヌクレオチド配列を指定する。このパラメータの操作に関するさらなる詳細は以下に示される。配列を除去するというパラメータの次に来るのが「GCトレードオフ値(GCTradeOffValue)」(括弧108)というパラメータである。このパラメータは、最適化配列のGC含量を調整することでヌクレオチド配列の生物特異的な最適化をさらに制御する。このパラメータの操作のさらなる詳細は同様に、以下に示される。

20

#### 【0167】

図9A中の入力パラメータの次のセット(「オリゴ設計」以下)は、所望の遺伝子配列(すなわち、任意の生物特異的な修飾を含む、配列.txtで指定された配列)を作製するのに利用される構築および選択用オリゴヌクレオチドの設計を制御する。括弧110に示されたパラメータ(「次のもので配列を選択する(pickSequenceBy)」)は、オリゴヌクレオチドを所望のオリゴヌクレオチドの重複末端の $T_m$ にのみ基づいて(pickSequenceBy= $T_m$ )またはオリゴヌクレオチドの長さに基づいて(pickSequenceBy=長さ)設計するかどうかを指定する。pickSequenceBy=長さの場合、長さ(ヌクレオチド数の)は「chipSeqLen」パラメータ(括弧112)として指定される。pickSequenceBy=長さで、長さが指定されない場合、デフォルト値(例えば、40ヌクレオチド)が使われる。

30

#### 【0168】

chipSeqLenパラメータの次に来るのが括弧114の「chipExtraSeqLen」および「フィルアップを終わらせる(endFillUp)」パラメータである。chipExtraSeqLenパラメータは、制限酵素(RE)切断の結果として残存しうる構築用オリゴヌクレオチドの付着端の長さを指定する。endFillUpパラメータは、余分の配列を付加して等しい長さのオリゴヌクレオチドを作製するかどうかを指定する。構築用オリゴヌクレオチドおよび選択用オリゴヌクレオチドの長さは、一定または可変とすることができる。余分の配列をオリゴヌクレオチドの一端または両端に付加することができる。付加配列は、構築用オリゴヌクレオチドに隣接する遺伝子中の天然核酸配列から選択される。

40

#### 【0169】

括弧116に示されるのは、パラメータ「oligo $T_m$ 」である。このパラメータは、設計されたオリゴヌクレオチドの重複部分に対する $T_m$ の指定を可能にする。括弧118に示されるのは、パラメータ「DNA濃度」および「塩濃度」である。これらのパラメータは、オリゴヌクレオチドの配列特異的ハイブリダイゼーションの間のDNA鎖および塩の溶液濃度に対する特定値の入力を可能にする。以下でさらに詳細に論じられるように、これらの値は、重複するオリゴヌクレオチドセグメントの $T_m$ を算出する際に使われる。

#### 【0170】

50

パラメータ入力ファイルは図9Bに引き継がれる。図9Bの第1部(「オリゴ・チップ-合成」以下)のなかには、パラメータ「センス5endAddOn」および「センス3endAddOn」(括弧120)である。以下でさらに十分に論じられるこれらのパラメータは、各構築用オリゴヌクレオチドの5'および3'末端に付加される配列を指定する。これらの配列は、例えば、制限酵素認識部位とすることができる。パラメータ「選択5endAddOn」および「選択3EndAddOn」(括弧122)は同様に、以下に論じられており、選択用オリゴヌクレオチドの5'および3'末端に付加される配列を指定する。括弧124には、所望のオリゴヌクレオチド長に到達するため選択用オリゴヌクレオチドに付加されうるアデニン塩基の数に制限を指定するパラメータ「選択FillUpLen」がある。パラメータ「選択ChipTM」(括弧126)は、構築用オリゴヌクレオチドの部分に重複する選択用オリゴヌクレオチドの部分に対する $T_m$ である。

10

## 【0171】

図9Bの最終部分はパラメータ「re部位」および「プールサイズ」(それぞれ括弧128および130)を含む。re部位というパラメータは、配列がさらに小さい配列に切断されうる制限酵素(RE)部位を特定する。これらの部位は「配列を除去する」というパラメータによって先に特定された配列と同じであってもよい(しかし同じである必要はない)。少なくともいくつかの態様では、複数のRE部位がフォーマット<5'-3'方向のRE部位1>; <3'-5'方向のRE部位1>; <5'-3'方向のRE部位2>; <3'-5'方向のRE部位2>;などで示される。プールサイズというパラメータは、入力配列を切断して構築用オリゴヌクレオチドを作製できる断片の数に制限を設定する。プールサイズパラメータの操作は同様に、以下で論じられる。

## 【0172】

配列.txtおよびcadpam.特性の入力を受けた後に、プログラムはブロック20に進む。判断ブロック20で、プログラムは、発現生物のコドン使用頻度に基づく最適化が望まれるかどうか(すなわち、図9Aの「最適化する」というパラメータが「オン」または「オフ」であるかどうか)を判断する。最適化が望まれない場合、プログラムは、ブロック20からブロック26の「いいえ」の分岐に進む。ブロック26は以下で論じられる。最適化が望まれる場合、プログラムは、ブロック20からブロック22の「はい」の分岐に進む。ブロック22で、利用者指定のまたはデフォルトの生物に対するコドン表がロードされる。図10Aおよび10Bは、デフォルトの生物、大腸菌K12に対するコドン使用頻度表を示す。標準的なGCC-通常フォーマットである、図10Aおよび10Bの表は、多数の生物に使用可能なコドン使用頻度表に類似している。そのような表の情報源の1つを<<http://www.kazusa.or.jp/codon/>>でオンライン確認できる。縦列140は20種のアミノ酸に対する略語を記載しており、縦列142はそれら20種のアミノ酸の各々をコードするのに使われるコドンを記載している。縦列148は特定の生物に対する各コドンの使用割合を記載している。例えば、図10A中の最初の4行はグリシン(「Gly」)に相当する。グリシンをコードする4つのヌクレオチドトリプレットのうち、GGGは大腸菌K12によってグリシンをコードする機会のうち15%(すなわち、0.15)使用される。GGA、GGTおよびGGCはそれぞれ、11%、34%および40%使用される。縦列144および146は、本発明の少なくともいくつかの態様によっては使用されないが、それらは標準的なGCC-通常フォーマットの一部であるため留置されている。別の生物に対するコドン使用頻度表は、同じフォーマットになるが、その他の生物に応じて縦列144~148の値は異なるはずである。

20

30

40

## 【0173】

ブロック22でコドン使用頻度表をロードする一環として、プログラムは、各コドンのGC含量に基づいて表中のコドン使用割合を調整する。配列中の特定のコドンをその同じアミノ酸に対し発現生物が最も頻繁に使用する別のコドンと置換することが望ましいかもしれないが、その生物による全体的な発現を向上させるため、配列のGC含量を最小限に抑えることも望ましいかもしれない。これらは相反する目的であることが多い(すなわち、使用割合が最も高いコドンは同様に、GC含量が最も高いコドンになることが多い)ので、これらの2つの判定基準間のトレードオフをGCTradeOffValueパラメータで指定することができる(図9A)。2つまたは3つのGまたはC塩基を有する使用頻度表中の各コドンの場合には、GCTradeOffValueをそのコドンの使用割合から差し引く。例えば、GCTradeOffValue=0.12の

50

場合、図10AのGGGおよびGGAコドンはそれぞれ、その使用割合が-0.21 (0.15-0.12-0.12-0.12)および0.0 (0.11-0.12、負の値が端数を0に丸めている)にまで減った。0のまたは1つのGまたはC塩基を有する使用頻度表中の各コドンの場合には、GCTradeOffValueをそのコドンの使用割合に加える。GCTradeOffValue=0.12の本願の例では、スレオニンに対する2つのコドン(ACAおよびACT)は、その使用割合が増えた(それぞれ0.25および0.29にまで)。

#### 【0174】

ブロック22でコドン表をロードした後、プログラムはブロック24に進む。ブロック24で、プログラムは次いで、ロードされたコドン表に基づいておよびcadpam.特性ファイルで指定されたその他のパラメータに基づいて入力ファイル(配列.txtファイル由来の)を最適化する。図11に示されるのは、最適化手順を記述するフローチャートである。ブロック24-1から始めて、プログラムは入力配列中の最初の3塩基を調べる。複数の配列が配列.txtファイルに含まれる場合、図11の最適化手順は各配列に対し連続的に行われる(すなわち、手順は最初の配列に対し、それから次の配列に対し実行される、など)。ブロック24-3において、プログラムは、検査する塩基をブロック22でロードされたコドン使用頻度表と比較し(図11)、(GCTradeOffValueがゼロに等しくはなく、最適化する=オンならば調整の後に)最も高い使用割合を有する同一アミノ酸に対するコドンを同定する。プログラムは次に、ブロック24-5でもとのコドンに代えて最も高い使用コドンを用いる。場合によっては(例えば、もとのコドンが最も使用されるコドンであり、低いGC含量を有する)、プログラムは事実上、コドンをその同じコドンに置換していることになる。

#### 【0175】

ブロック24-5から、プログラムはブロック24-7に進み、配列中にさらに良いコドンが存在するかを判断する。もし存在するなら、プログラムは「はい」の分岐をブロック24-9に進み、配列中の次の3塩基を調べる。ブロック24-9から、プログラムは次に、ブロック24-3に戻り、その次の3塩基についてブロック24-3から24-7を繰り返す。ブロック24-7で、プログラムが配列の末端に到達した場合、プログラムは「いいえ」の分岐をブロック24-11に進む。

#### 【0176】

ブロック24-11で、プログラムは配列中の二次構造を探し、その二次構造を代替コドンに置換する。具体的には、プログラムは配列全体に沿って、ループ、ヘアピンなどを形成しうる塩基の組合せを検索する。少なくともいくつかの態様では、プログラムは所与の領域内の自己相補的な配列を探すことでこの検索を行う。二次構造の発見と同時に、プログラムは次いで二次構造のコドンを、同一アミノ酸をコードする代替コドンに置換する。いくつかの態様では、置換コドンはブロック24-11で、使用頻度表から最も高い使用割合を有する代替コドンを選択することにより選択される。

#### 【0177】

いくつかの態様では、ブロック24-11のステップは、二次構造を特定することなく配列全体を横断できるまで、または他の停止条件に到達する(例えば、配列をある回数通過する)まで繰り返される。例えば、1つまたは複数のコドンを置換してある領域中の二次構造を排除することで、配列の別領域の中に二次構造をうっかり導入してしまう可能性がある。これが起こる場合、うっかり作製された二次構造は、配列の次回通過時に補正される。簡単にするため、ブロック24-11が繰り返される別の態様は、破線矢印で示されている。

#### 【0178】

ブロック24-11を完了した(またはブロック24-11の繰り返しを全て完了した)後、プログラムはブロック24-13に進む。ブロック24-19で、プログラムはcadpam.特性ファイルの配列を除去するというパラメータ(図9A)で特定された塩基の組合せがないか配列を検索する。そのような塩基の組合せの発見と同時に、プログラムはその塩基を、同一アミノ酸をコードするコドンに置換する。いくつかの態様では、置換コドンはブロック24-13で、使用頻度表から最も高い使用割合を有する代替コドンを選択することにより選択される。いくつかの態様では、およびブロック24-11について記述されているものと類似の理由で、ブロック24-13は、配列を除去するの塩基の組合せを発見することなく配列全体を横断できる

まで、または他の停止条件に到達するまで繰り返される。

【0179】

ブロック24-13後、プログラムは図7のメインプログラムフローに戻り、ブロック26に進む。ブロック26で、プログラムは、最適化された入力配列(またはブロック20からブロック26に直接到達した場合にはもとの入力配列)をre部位パラメータによって特定されたRE部位(図9B)がないか調べる。ブロック28で、およびそれらのRE部位のいずれかが見られる場合、プログラムはその発見部位で配列を分割する。プログラムは、その後に設計される構築用オリゴヌクレオチドがそのような部位を望ましくない位置に(例えば、構築用オリゴヌクレオチド配列の中央に)持たないように、入力配列をRE部位で分割する。

【0180】

ブロック28での配列の分割は、図12を図8と比較することによって見定められる。図12は、さらに短い4つの配列rs1-f1、rs1-f2、rs1-f3およびrs1-f4に分割された入力配列rs1を示す。配列rs2は指定のRE部位をどれも含まれていなかったため、配列rs2は分割されなかった。指定のRE部位のrs1内での位置が図8中で囲みにより示されている。それらの部位の各々で、RE部位は中央で分割される。すなわち、例えば、rs1-f1とrs1-f2との間の分割は、図8の最初の囲みで示されるRE部位acctgcの真ん中で起こる。さらに短い配列rs1-f1からrs1-f4(図12)の末端近くの部分的な囲みは、図12の囲みの半分を表す。

【0181】

その後、プログラムは判断ブロック30に進む。ブロック30で、プログラムは、オリゴヌクレオチドを $T_m$ に基づいてまたはオリゴヌクレオチドの長さに基づいて設計するかどうかを判断する。次のもので配列を選択するという入力パラメータ(括弧110、図9A)が「tm」に相当する場合、プログラムはブロック34に進み、設計される構築用オリゴヌクレオチドの重複部分の $T_m$ に基づいて構築および選択用オリゴヌクレオチドを設計する。

【0182】

ブロック34中のプログラムの操作は、図13Aから18にさらに詳細に示されている。図13Aおよび13Bは、図7のブロック34中の続く、少なくともいくつかの態様による、アルゴリズムのステップを示すフローチャートである。ブロック34-1(図13A)から始めて、プログラムは、構築および選択用オリゴヌクレオチドを作製することになる最初の配列を読み出す。例として図8の入力を用いて、および上述のとおり配列rs1のさらに短い配列への分割の後で、図13A~Bのアルゴリズムで分析される配列はrs1-f1、rs1-f2、rs1-f3、rs1-f1およびrs-2である。したがって、プログラムは、ブロック34-1での分析にこれらのうち最初のもの(rs1-f1)を選択する。

【0183】

プログラムは次いでブロック34-3に進み、ブロック34-1中で選択された配列の3'末端に開始点を置く。これは図14に図示されており、ここでは開始点が配列rs1-f1の3'末端に位置する三角形として示されている。その後、プログラムはブロック34-5に進む。ブロック34-5で、プログラムは所定数(W)の塩基をrs1-f1の開始点から5'末端の方に伸ばす検索ウィンドウを特定する。少なくともいくつかの態様では、検索ウィンドウの長さは、Wが一番近い整数に四捨五入された $T_m$ (図9A、括弧116)に等しくなるように設定される。本願の例では、 $W=50$ 塩基である。プログラムはその後ブロック34-7に進み、そこでプログラムは、検索ウィンドウが現行の配列の5'末端を超過するかどうかを判断する。別の言い方をすれば、プログラムは、W塩基が開始点から現行の配列の5'末端を越えて伸びるかどうかを判断する。もし伸びるなら、プログラムは、以下で論じられる「はい」の分岐をブロック34-21に進む。もし伸びないなら、プログラムは「いいえ」の分岐をブロック34-9に進む。

【0184】

ブロック34-9で、プログラムは次いで検索ウィンドウ中の重複領域を特定する。以下で説明されるように、プログラムによって分析される配列はさらに重複断片の一群に分割される。検索ウィンドウ内の重複領域を特定するため、プログラムは、入力パラメータ(括弧116、図9A)で指定された所望の $T_m$ 値に最も近い融解点 $T_m$ を有する領域を検索する。図13

10

20

30

40

50

Bはブロック34-9中のプログラムの操作をさらに詳細に示す。ブロック34-9-1で、プログラムは、開始点が分析される配列の3'末端に現在あるかどうかを判断する。もしあるなら、プログラムは「はい」の分岐をブロック34-9-3に進む。ブロック34-9-3で、プログラムは次に、オフセット距離を検索ウィンドウ内の5'末端の方に移動する。これは図14にも図示されている。プログラムは、重複領域が配列の3'末端の位置で始まらないようにオフセット距離を5'方向に移動する。これが行われる場合、重複領域は検索ウィンドウ全体を消費すると考えられる。以下で見られるように、これにより、別の構築用オリゴヌクレオチドによって完全に重複される構築用オリゴヌクレオチドがもたらされると考えられる。

【0185】

オフセット距離を5'末端の方に移動した後、プログラムはブロック34-9-5に進む。ブロック34-9-5において、および同様に図14に示されるとおり、プログラムは、入力パラメータで指定された $T_m$ 値に最も近い融解点を有する検索ウィンドウ内の領域を検索する。少なくともいくつかの態様では、融解点は、入力パラメータ(括弧118、図9A)によって指定されたDNA濃度および塩濃度に対する値を考慮しながら、最近接法を用いて算出される。融解点算出の最近接法は当技術分野において公知であり、Breslauer et al. (1986) Proc. Natl. Acad. Sci. U.S.A. 83:3746 (前記)に記述されている。最近接法を実装するコンピュータアルゴリズムは、当技術分野において公知であり、したがって本明細書ではさらに記述されない。

【0186】

図15は、入力 $T_m$ 値に最も近い融解点を有する重複領域(下線)がブロック34-9-5で見つられた後のrs1-f1の3'末端を図示している。図15に見られるとおり、重複領域は最初のオリゴヌクレオチド断片(rs1-f1-1)を規定する。重複領域は断片の左側(rs1-f1-1L)であり、重複領域の3'末端と断片の3'末端との間の部分が断片の右側(rs1-f1-1R)である。ブロック34-11(図13A)で、rs1-f1-1、rs1-f1-1Lおよびrs1-f1-1R中の塩基が保存され、プログラムはブロック34-13に進む。ブロック34-13で、プログラムはこれが、分析される配列の末端に到達したかどうかを判断する。もし到達していないなら、プログラムは「いいえ」の分岐をブロック34-15に進む。ブロック34-15で、開始点は先に特定された重複領域の3'末端(図15に示される)に移動し、プログラムはブロック34-5に戻る。

【0187】

ブロック34-5に戻った後、プログラムはブロック34-7および(ブロック34-7で「いいえ」と判断されると仮定すれば)ブロック34-9を繰り返す。しかしながら、この場合、開始点はもはや配列の初めの位置にはなく、プログラムはしたがって「いいえ」の分岐をブロック34-9-1(図13B)からブロック34-9-7に進む。ブロック34-9-7で、プログラムは次いで、図16に示されるように次の重複領域を判断する。先に見つられた重複領域(rs1-f1-1L)の5'側の最初の塩基位置で始めて、プログラムは検索ウィンドウの5'末端の方に移動し、所望の $T_m$ に最も近い融解点を有するrs1-f1-1Lに隣接の塩基を判断する。いったんこれらの塩基が見つけれたら(図16中で二重下線により示されている)、プログラムはブロック34-11(図13A)に進み、最新のおよび先の最新の重複領域によって規定される配列の部分を次のオリゴヌクレオチド断片(rs1-f1-2)として保存する。最新の重複領域はrs1-f1-2Lになり、先の重複領域(rs1-f1-1L)は同様にrs1-f1-2Rである。その後、プログラムはブ

【0188】

図17は、配列の末端に到達した場合のプログラムの操作を図示している。これはブロック34-7(図13A)およびブロック34-21からの「はい」の分岐に相当する。図17に示されるように、プログラムは所望の $T_m$ を達成するため必要に応じて塩基を付加する。これらの付加塩基を有する断片に相当する構築用オリゴヌクレオチドの部分は後に、構築中の遺伝子または配列から除外することができる。最後の断片(rs1-f1-n、または例では、rs1-f1-38)は先の重複領域、調査中の断片の残存5'末端、および付加塩基によって規定される。この情報が保存されて、プログラムはブロック34-13に進む。ブロック34-13で、配列の末端に到達したら、プログラムは「はい」の分岐をブロック34-17に進む。ブロック34-17で、

10

20

30

40

50

プログラムは、分析されるさらなる配列が存在するかどうかを判断する。もし存在するならば、プログラムは「はい」の分岐をブロック34-19に進み、次の配列(例えば、図12中のrs1-f2)に行く。もし存在しないならば、プログラムは「いいえ」の分岐をブロック36に進む(図7)。

**【0189】**

図18は、図13A~13Bに示されるステップの間にプログラムによって作製されたデータを含む出力ファイル(例では、「info.out」と題した)の一部を示す。図18に示されるデータの一部は、下記のように、引き続きステップのプログラムによって作製される。

**【0190】**

次のもので配列を選択するという入力パラメータ(括弧110、図9A)が「tm」の代わりに「長さ」に設定された場合、プログラムはブロック30(図7)からブロック32に進むと考えられる。ブロック32中のプログラムの操作は、図19から22にさらに詳細に示されている。図19は、図7のブロック32中の続く、少なくともいくつかの態様による、アルゴリズムのステップを示すフローチャートである。ブロック32-1で始めて、プログラムは、構築および選択用オリゴヌクレオチドを設計することになる最初の配列または配列を読み出す。この場合も先と同様に例として図8の入力を用いて、プログラムはブロック32-1での分析にrs1-f1を最初に選択する。

10

**【0191】**

プログラムは次いでブロック32-3に進み、ブロック32-1中で選択された配列の3'末端に開始点を置く。これは図20に図示されており、ここでは開始点が配列rs1-f1の3'末端に位置する三角形として示されている。その後、プログラムはブロック32-5に進む。ブロック32-5で、プログラムは入力の「chipSeqLen」パラメータ(括弧112、図9A)に対応する、現行の配列の開始点から5'末端の方向に伸びる、塩基の数を特定しようと試みる。図19~22の例では、chipSeqLen=40塩基とされる。プログラムはブロック32-7に進み、そこでプログラムは、これがrs1-f1の5'末端を超過したかどうかを判断する。別の言い方をすれば、プログラムは、chipSeqLen塩基が開始点からrs1-f1の5'末端を越えて伸びるかどうかを判断する。もし伸びるならば、プログラムは、以下で論じられる「はい」の分岐をブロック32-21に進む。もし伸びないならば、プログラムは「いいえ」の分岐をブロック32-9に進む。

20

**【0192】**

ブロック32-9で、長さに基づいてステップ32-5で特定された断片がrs1-f1-1(図20)になる。rs1-f1-1の5'末端の位置で開始し、所望の値(括弧110の入力パラメータ「tm」、図9A)に最も近い融解温度を有するrs1-f1-1の5'末端の塩基を特定することにより、プログラムはrs1-f1-1に対する重複領域を判断する。オリゴヌクレオチド断片が今回は必要な長さに基づいて選択されているので、重複領域に対するもっと広い範囲の $T_m$ 値が必要とされるかもしれない。いったん重複領域が特定されたら、プログラムはブロック32-11に進む。ブロック32-11で、プログラムはrs1-f1-1、rs1-f1-1L(ブロック32-9中で見つけられた重複領域)、およびrs1-f1-1R(所望の $T_m$ に最も近い $T_m$ を有する3'末端のrs1-f1-1の部分)中の塩基に関するデータを保存する。

30

**【0193】**

その後、プログラムはブロック32-13に進み、現行の配列の末端に到達したかどうかを判断する。もし到達していないならば、プログラムは「いいえ」の分岐をブロック32-15に進み、特定されたばかりの重複領域の3'末端に開始点を置く。これが図21に示される。その後、プログラムはブロック32-5に戻り、ブロック32-5、32-7および(現行の配列の末端が超過されていないと仮定して)32-9から32-13のステップを繰り返す。図21は、第2の長さに基づくオリゴヌクレオチド断片(rs1-f1-2)ならびにその左および右の部分の決定を図示する。第2のおよびその後の長さに基づく断片の場合、その右側が前の断片の左側に相当するように設定される(例えば、rs1-f1-2Rはrs1-f1-1Lと同じである)。

40

**【0194】**

図22は、配列の末端に到達した場合のプログラムの操作を図示している。これはブロック32-7(図19)およびブロック34-21からの「はい」の分岐に相当する。図22に示されるよ

50

うに、プログラムは、指定の長さを達成するためおよび所望の $T_m$ に可能な限り近い融解点を有する左末端を得るため必要に応じて塩基を付加する。これらの付加塩基に相当する構築用オリゴヌクレオチドの部分は後に、構築中の遺伝子または配列から除外することができる。最後の断片(rs1-f1-n、または例では、rs1-f1-23)は、その左および右末端と共に、図22に示されている。この情報が保存されて、プログラムはブロック32-13に進む。ブロック32-13で、配列の末端に到達したら、プログラムは「はい」の分岐をブロック32-17に進む。ブロック32-17で、プログラムは、分析されるさらなる配列が存在するかどうかを判断する。もし存在するなら、プログラムは、はいの分岐をブロック32-19に進み、次の配列(例えば、図12中のrs1-f2)に行く。もし存在しないなら、プログラムは「いいえ」の分岐をブロック36に進む(図7)。

10

## 【0195】

図23は、図19~22に示されるステップの間にプログラムによって作製されたデータを含む出力ファイル(例では、「info.out」と題した)の一部分を示す。図23に示されるデータの一部は、下記のように、引き続くステップのプログラムによって作製される。

## 【0196】

ブロック36で、構築および選択用オリゴヌクレオチドは、ブロック32またはブロック34で決定された断片(例えば、rs1-f1-1、rs1-f1-2など)に基づいて作製される。図24は、どのようにして構築用オリゴヌクレオチドが作製されるかを図示しており、ならびに図18のinfo.outファイル、図9Bのcadpam.特性ファイル、および作製された構築用オリゴヌクレオチドを含んだ第3のファイル(「chipProduction.out」と名付けた)の一部分を図示している。第1の構築用オリゴヌクレオチド(rs1-f1-1c)は、rs1-f1-1(info.out)の相補体を利用し、「センス5endAdd0n」および「センス3endAdd0n」入力パラメータ(cadpam.特性由来の)によって特定された配列を付加して作製される。rs1-f1に対する残りの構築用オリゴヌクレオチド(例えば、rs1-f1-2c)(および処理中のその他の配列)は、同じように作製される。

20

## 【0197】

図25では選択用オリゴヌクレオチドの作製を図示しており、例として構築用オリゴヌクレオチドrs1-f1-1c(図24)を利用している。各構築用オリゴヌクレオチドに対し、2つの選択用オリゴヌクレオチド(「a」および「b」)が作製される。図25では、センス5endAdd0nおよびセンス3endAdd0n配列を除くrs1-f1-1cの部分がステップ(1)でより大きな書体を用いて強調されている。プログラムは、選択Chip $T_m$ (括弧126、図9B)の指定値に基づいて「a」および「b」部分を決定する。具体的には、プログラムは、指定の選択Chip $T_m$ 値に最も近い $T_m$ を有する構築用オリゴヌクレオチドの左および右側の部分を特定する。その後、「a」選択用オリゴヌクレオチド(rs1-f1-1s-a)は、「a」部分の相補体を利用し(ステップ(2))、「選択3endAdd0n」パラメータ(図9B)によって指定された配列を相補体の3'末端に付加し(ステップ(3))、「選択5endAdd0n」パラメータ(図9B)によって指定された配列を付加した時点でrs1-f1-1-s-aが60塩基(選択Chip $T_m$ パラメータに基づいて決められた塩基の数)を有するように十分なアデニン塩基を付加し(ステップ(4))、それから選択5endAdd0n配列を付加して(ステップ(5))作製される。この手順をステップ(6)から(9)で続けて、選択用オリゴヌクレオチドrs1-f1-1-s-bを得る。その後、同様のステップを続けて、全ての構築用オリゴヌクレオチドに対する「a」および「b」選択用オリゴヌクレオチドを得る。

30

40

## 【0198】

ブロック38(図7)では、プログラムは次いで遺伝子断片および末端プライマーを設計する。具体的には、プログラムは構築用オリゴヌクレオチドに応じて合成される遺伝子断片の長さを決める。入力パラメータ「プールサイズ」(図9B)を用いて、プログラムは、どれくらいの構築用オリゴヌクレオチドを各断片に使用できるかを判断する。例えば、プールサイズ=50なら、最大で50個までの構築用オリゴヌクレオチドを各断片に使用することができる。プールサイズが配列に対し設計された構築用オリゴヌクレオチドの数以上である場合、配列を単一の遺伝子断片として合成することができ、左側および右側プライマーの単一セットをその断片に対し設計することができる。プールサイズが配列に対し設計され

50

た構築用オリゴヌクレオチドの数に満たない場合、配列を複数の遺伝子断片として合成しなければならず、各断片は左側および右側プライマーのそれぞれのセットを持たなければならない。

#### 【0199】

図18は、プールサイズ=50および次のもので配列を選択する=tmでの、rs1-f1に対するinfo.outファイルの一部分を示す。これは結果的にrs1-f1に対する38種の構築用オリゴヌクレオチドをもたらす(すなわち、構築用オリゴヌクレオチドはrs1-f1-1からrs1-f1-38の各々に相当する)ので、rs1-f1を単一の配列として合成することができる。次いで、5'および3'プライマーが所定の範囲のオリゴTm内で融解点を有するように十分な塩基を遺伝子断片の各末端の位置に選択することにより、末端プライマーをrs1-f1に対して設計する。図26は、プールサイズ=5および次のもので配列を選択する=tmでの、rs1-f1に対するinfo.outファイルの一部分を示す。この場合には、rs1-f1に対する38種の構築用オリゴヌクレオチドは8種の「プール」に分割されており、rs1-f1は8種の遺伝子断片として合成される。次いで、5'および3'プライマーが所定の範囲のオリゴTm内で融解点を有するように十分な塩基を遺伝子断片の各末端の位置に選択することにより、末端プライマーをそれら8種の断片の各々に対して設計する。

10

#### 【0200】

図27は、プールサイズ=50、次のもので配列を選択する=tmおよびchipExtraSeqLen=7での、rs1-f1に対するinfo.outファイルの例である。この場合には、断片の7塩基長の付着端が「余分の5末端」および「余分の3末端」と見なされている。

20

#### 【0201】

ブロック38(図7)から、プログラムはブロック40に進み、設計された構築および選択用オリゴヌクレオチドに対するデータを含んだファイルを出力する。先に論じられた「info.out」および「chipProduction」出力ファイルに加えて、プログラムは選択用オリゴヌクレオチド(「chipSelectionA」および「chipSelectionB」、示されていない)を収載する2つのファイル、ブロック28で分割の入力配列(図12に示される「full\_sequences.out」)を含んだファイル、および構築用オリゴヌクレオチドに対し逆相補性を有するオリゴヌクレオチド配列を含んだファイルを出力する。

#### 【0202】

限定するものと解釈されるべきではない以下の例により、本発明をさらに説明する。本出願を通じて引用される全ての参考文献、特許および公開された特許出願の内容は、あらゆる目的でその全体が参照により本明細書に組み入れられる。

30

#### 【0203】

実施例I

ブレ増幅

1つまたは複数のオリゴヌクレオチドには、5から30塩基長とできるおよび/またはその3'末端のタグに相補的なさらに長いプライマーを持つことで増幅サイクルの間に伸長される「一過性タグ」または「増幅部位」(例えば、ユニバーサルな一過性タグ、またはユニバーサルな増幅部位)を隣接させることができる。プライマーは3'末端の易動性ヌクレオチド、例えば、N7位でアルキル化されたプリン(N7me-dGTP)を有することができる。これらは熱に不安定および/または光に不安定とすることができ、ほんの数ラウンドのPCRしか続けることができない。放出されるとまたは損傷を受けると、次のラウンドのポリメラーゼ作用は、オリゴまたは伸長されたオリゴ(所望の最終配列に隣接する)上でプライミングするのに適した「長いプライマー」が作製されるようなその位置でまたはその位置の近傍で終結することができる。理論により束縛されることを意図するわけではないが、これは、たとえ選択の鋳型に依然として一過性タグが隣接していたとしても機能するはずである。この一番末端のタグは不安定ではなく、故に最終ラウンドで優位を占める。所望のプライマーを合成する方法の1つは、5'末端に相補的な余分のヌクレオチドを有する鋳型上で、硫酸ジメチル処理されたdATPまたはdGTP(および精製された)を用いて伸長することである。魅力的な代替手段は、1つまたは複数のrNTPsを鋳型プライマーの3'末端に用いる

40

50

ことである。これらは熱およびMg<sup>++</sup>によりまたはRNaseにより不安定化されると考えられる。RnaseHは保留物ではなく伸長されたプライマーを選択的に攻撃するので、特に適しており；これはある最高の程度まで、正確に切断された鋳型を作製しながらもとのプライマーを再生することができる。

## 【0204】

一過性タグのその他の変化形としては、一過性タグのIIS型制限切断(以下に示される)、およびまたは増幅の間の反応への接近が必要になる化学的切断が挙げられる。

## 【0205】

## 実施例II

IIS型制限部位を用いた反復PCRアッセムブリ

14から28塩基対重複の、Xeotron型チップ上の70mer516種のプールから選択のプレ増幅された40mer38種の反復PCRアッセムブリ。選択した2種のIIS酵素は以下であった。

5' ... GGTCTC(N)<sup>1</sup> ... 3' BsaI  
 3' ... CCAGAG(N)<sup>5</sup> ... 5'  
 5' ... ACCTGC(N)<sup>4</sup> ... 3' BfuAI  
 3' ... TGGACG(N)<sup>8</sup> ... 5'

ストラテジーは実施例IXに示されている。

## 【0206】

## 実施例III

44mer(放出後30mer)の両端での同じ7merタグの使用

ユニバーサルな一過性プライマー: 5' tagtaga 3' (3'側の下線塩基は容易に除去可能である)

rs1-1由来の一過性PCR産物は以下である。

5' tagtagaTAAACAGGAAGATGCAAATTTTAGTAATAtctatcta 3' (SEQ ID NO:64)  
 3' atcatctATTTGTCCTTCTACGTTTAAAATCATTATagatagat 5' (SEQ ID NO:65)

## 【0207】

下側の鎖の特異的塩基での切断および7merによる伸長の後、本発明者らは以下のssの37merを得る。

5' tagtagaTAAACAGGAAGATGCAAATTTTAGTAATAA 3' (SEQ ID NO:66) |||||  
 3' atcatctCATTATTAACGTTACCGTCTTCGTAAATTTcagatagat 5' (SEQ ID NO:67)

これは上記の重複する43mer(タグなし30塩基)と対合することができる。

## 【0208】

2回の伸長の後、以下のdsの68mer(タグなし54塩基)が作製される。

5' tagtagaTAAACAGGAAGATGCAAATTTTAGTAATAATGCAATGGCAGAAGCA  
 TTTAAAGtctatcta (SEQ ID NO:68)

3' atcatctATTTGTCCTTCTACGTTTAAAATCATTATTACGTTACCGTCTTCGT  
 AAATTTcagatagat (SEQ ID NO:69)

## 【0209】

## 実施例IV

隣接オリゴの添加の順序にバイアスをかける固定化合成パターンの利用

遺伝子が2Dレイアウトのオリゴヌクレオチドのクラスターとして合成されるならば、そ

10

20

30

40

50

れらは「インサイチュー」ポロニー(すなわち、ポリメラーゼコロニー)に類似の方法でアッセンブルすることができる。鑄型は固定の70merとすることができ、移動相(例えば、ゲルまたは高分子媒体中の)はユニバーサルプライマーおよびその伸長産物とすることができる。遺伝子のもっと大きな染色体へのアッセンブリまたはインサイチューを目的に、部位特異的な組換えタンパク質を遺伝子工学で作り変えることができる。理論により束縛されることを意図するわけではないが、このパターン化アッセンブリは、選択の数が各ステップで非常に小さいので、ミスプライミング/ミスアッセンブリの問題を大幅に減らすことができる。別の恩典は、混合物全体が通常のPCR反応容量中に放出された場合よりも局所濃度が高い(例えば、フェムトリットルのポロニースケール反応 対 マイクロリットルスケール)ということである。例えば、現行のXeotronアレイでは、20 nI容量中で8000個のオリゴヌクレオチドを合成する。これらが通常のPCR容量(10マイクロリットル)に希釈されたなら、その濃度は各オリゴ1 pM (= 6 M分子)である。PCRプライマーは1000 nMで通常使用されており、したがって未希釈の1 nM濃度でさえも初め(二分子のうち的一方が通常よりも希薄な二分子反応)は1000倍ゆっくと進むものと予想される。

10

## 【0210】

限定するものではない例は、以下の2Dアレイレイアウトであり、この場合に、4組のプライマーペア(例えば、70merのペアabとbc)が最初に互いに伸長するはずであり(ダッシュを参照のこと、abc、cde、efgおよびghiの産生)、その後、伸長と拡散のため、これらの産物のうち2ペアが共に伸長して(縦線に沿って)、abcdeとefghiを作り出すはずである。最後に、これらが融合して、所望のabcdefghiを作り出す。各オリジナルペアに対する斑点中央の間の距離は40ミクロンおよび最も近い点の間は5ミクロンでありえ、その一方で最初のペアの次のペアからの重心は次のものに対し100ミクロンと200などでありえる。

20

ab-bc ef-fg

|-----|

cd-de gh-hi

30

## 【0211】

実施例V

増幅後の鎖選択ストラテジー

代替案は合成された鎖を交互に入れ替えることである(例えば、実施例IXのrs3-2およびその他偶数のオリゴは、上記実施例IIに例示されているものの逆鎖になっているはずである)。二つのPCR反応物をオリジナルのチッププールから作り出す。一方は、BfuAIのみで切断されるピオチン化L-プライマーに使用する。このプールをストレプトアビジン-ビーズに結合させ、ピオチン化されていない鎖を遊離させて、一本鎖の55mer(ss-55-mers)を残すことができる。もう一方の反応物は、両酵素で切断される2本の非ピオチン化プライマーに使用して、二本鎖の40merを遊離させる。40merのうち片鎖だけが、40塩基対が完全に適合した状態で55merのビーズに結合するはずである。14から28 bpの重複部分は顕著に結合しないはずである。不完全な適合物は融解温度( $T_m$ )よりも少し低い温度で洗い流すことができ、完全な適合物は $T_m$ を少し超える温度で溶出することができる。

40

## 【0212】

ソフトウェアを利用して、40merの位置を変化させることによりほぼ同じの $T_m$ 点をもたらすことができる(またはサイズ選択が緩ければ、「40mer」の39、41などまでの長さ変化により、 $T_m$ 均等化をさらに良好にすることができる)。

## 【0213】

実施例VI

ブレ増幅に加えてのライゲーションストラテジー

50

ライゲーションは、通常、1 nMまたはそれ以上の濃度で行われる。より小さな(故により高価な、例えば、Xeoチップ 8000 \* 40/\$2000 = 160塩基/\$ 対 Illumina 6塩基/\$)アレイエレメントを使用するにつれて、各オリゴマーの量は減少する(1チップ当たり70merの配列4000個で、これは約1 fmolであり、キャピラリー電気泳動精製で10%まで低下する) = ライゲーション反応10マイクロリットル中で各オリゴ0.1 fmole = 0.01 nM。すなわち、二分子反応速度は少なくとも希釈係数の二乗単位で遅く((1/.01)<sup>2</sup> = 10,000倍遅く)なるものと予想される。各チップ・オリゴマー(例えば、70mer)の末端に共有のタグプライマーを含めることで、PCR増幅が可能になる。初期の伸長反応は同じ二分子(二乗)相互作用に依存するので、これは同様に、反復PCRに役立つはずである。PCRがもたらすこのことからの通常の逸脱は適切ではない。これには、稀有な中間反応が起こるまでは起き得ない、過剰の両末端プライマーを用いた反応を推進させることが必要になるからである。ライゲーションと反復PCRの組合せは、原理上は、PCRサイクル数を(例えば、2<sup>6</sup> > 38なので、上記実施例1では少なくとも6サイクルだけ)減らすのに役立つが、実施にはそれらの余分のサイクルは必要なDNA量を得るため、いずれにせよ行われる必要があった。ライゲーションは同様に、5'および3'末端でミスマッチに対する負の選択をすることができるが、反復PCRは同じことをすると考えられる。たとえライゲーションの理論上の利点が明白ではないにしても、実験的組合せが場合によっては成功することもある。

10

## 【0214】

## 実施例VII

統合的多重化サイズ、ミスマッチおよび読み取り枠選択ストラテジー  
サイズ選択

20

チップ・オリゴマーの全てが同じ(またはほぼ同じサイズ)を有するならば、プール(またはサブセット)全体を、例えば、キャピラリー電気泳動またはHPLCにより多重化サイズ選択することができる(増幅の前および/または後に)。同様に、ライゲーションまたは反復PCR産物がほぼ同じサイズを有するならば、多重化サイズ選択を適用することができる。各遺伝子(または断片)に対する末端オリゴへのユニバーサルな遺伝子隣接PCRプライマーの設計は、望ましいことが多く、同様に遺伝子特異的プライマーの利用を妨げないはずである。DNAが異なるサイズを有するなら、これらの特性を利用して、任意の段階で脱多重化(分離)を始めることができる。

## 【0215】

## ミスマッチ選択

30

## 方法1:

上記実施例Vの鎖選択を同様に利用して、プールのT<sub>m</sub>の直下で事前溶出させることによりミスマッチに対する負の選択をすることができる。ソフトウェアプログラムを利用して、T<sub>m</sub>がかなり均質となるようにプールを設計してもよく、必要に応じて、2つまたはそれ以上のT<sub>m</sub>のプール用に別のチップを作製し、T<sub>m</sub>選択の後にプールをプールしてもよい。ミスマッチ識別を最大にするためおよびサイズ均一性とT<sub>m</sub>均一性との間の矛盾を減少させるため、1つまたは複数の「選択用オリゴ・セット」を上記のとおり、しかし主要プールとの重複部分をもっと短くして合成し増幅することができる(例えば、40mer-プラス-タグを1つではなく固定化された24mer(プラス・タグ)を2つ使った連続選択)。

40

## 【0216】

連続ラウンドのハイブリダイゼーション選択は、化学的合成エラーを相乗的に減少させることが分かった。エラー率は選択なしのアッセムブリで1/160から約50 bpの「構築用オリゴ」を網羅する約26mer重複の「選択用」オリゴを用いた2連続ステップによるアッセムブリで1/1400 bpに低下したことが認められた。選択用および構築用の長さはさまざまとできるが、T<sub>m</sub>は両端の選択用オリゴの長さを変化させることで望まれる均一性に近づけることができる。

## 【0217】

## 方法2:

MutSタンパク質に基づく選択。

50

## 【0218】

方法3:

二本鎖および/または一本鎖断片間のインビボでのまたはインビトロでの相同組換え。

## 【0219】

方法4:

無作為にニックが入れられ、再アニーリングされたプールは、3'末端が相補的な鋳型に適合する場合、選択的にDNAポリメラーゼによって伸長される。

## 【0220】

ORF選択

アッセンブルされた遺伝子(または中間にある断片)をインビボで(Lutz et al. (2002) Protein Eng. 15:1025)またはインビトロで(あらゆる目的でその全体が参照により本明細書に組み入れられるJermutus et al. (2001) Proc. Natl. Acad. Sci. U.S.A. 98:75)選択して、読み枠を維持する(例えば、読み枠の移動やナンセンス突然変異を克服する)ことができる。

## 【0221】

上記の選択方法のいずれの場合にも、最適な数の複数ラウンドの選択を利用して、最終産物の忠実度を増大させることができる。

## 【0222】

実施例VIII

ウェルプレートDNAプールセット

本発明のこの態様によれば、ユニバーサル384ウェルプレートなどの、標準的なウェルプレートを本明細書に記述されるプール合成方法および多数の高忠実度(または相違制御)のDNAを合成する他の方法と併せて、DNAの分配と利用のためのプラットフォームを有利に提供することができる。

## 【0223】

合成遺伝子に向けられる1つの態様では、データベース中のRNAおよびタンパク質コード遺伝子の数がますます増加しており、それらを単独でおよび種々の組合せで使いたいという願望が増えてはいるが、保存、複製および分配の費用はひどく高くなりうることを認識している。本発明によれば、1つの標準化された384ウェルプレートを利用して、総計何百万にも容易に達する、例えば、ヒト全遺伝子の一群、植物、微生物およびウイルス由来の数多くの遺伝子、多くの観測されているおよび理論上のスプライス変異体、よく見られる突然変異体、コドン最適化形などを含むDNA試料を収集するおよび手に入れることができるようにする。

## 【0224】

例として、前述のとおり遺伝子884,736 (=96<sup>\*</sup> 96<sup>\*</sup> 96)個を50チップ当たりわずか\$35,000で作製することができる(50mer/Kbp/遺伝子700,000個 = 1チップ当たり遺伝子17,500個)。一度マスター・プレートが作製されると、さらなる384ウェルプレートを各およそ\$300で(PCR、プライマー、労働力およびインフラ償還を含め)複製することができる。これらの遺伝子のそれぞれに3組のプライマーペアのネスティッドセットを隣接させることができる。本発明によれば、288組のユニバーサル・プライマーペアを利用して、任意遺伝子の任意量を手に入れることができる。これにより、一連の広範な利用者がcDNAクローニングまたは個々の貯蔵費用なくしてさまざまな遺伝子または遺伝子セグメントを利用することができる。

## 【0225】

例示のみを目的として、各遺伝子は次のような代表的構造を有する。

CCCCBBBBAAAAGGGGaaaabbbccccpppp

## 【0226】

上記において、aaaaおよびAAAAは内側プライマーペアである。aaaaの配列はPCRプライミングに適した任意の配列(例えば、その他のプライマーから遠く離れているように選択

された25mer)であってよく、AAAAと無関係であってよい。BBBBおよびbbbbは第2のペアであり、CCCCおよびccccは最も外側のペアであり、GGGGは所望の遺伝子である。

【0227】

384ウェルプレートなどの標準的なウェルプレートを、それぞれが9216 (=96<sup>\*</sup> 96)個の各遺伝子(最も外側のプライマーCCCC/ccccによって既に増幅されている)を含有する96サブウェルを含んだ左上部を有する四半部に分割する。左下の四半部は、前記96プールの任意/全部を再増幅するのに十分な量でその96組のプライマーペアを含む。右上の四半部は第2のプライマー(BBBBおよびbbbbタイプ)の全てを含み、右下部は最も内側のプライマー(AAAA/aaaa)を含む。左上の四半部から適切なウェルを選び、これを右上部中の適切なプライマーペアと混合し、PCRを行うことで、任意の遺伝子を増幅させることができる。その後、右下部から正しいウェルを利用して第2のPCR(任意でPCR間の精製ステップ)を行う。最終産物には、下流の適用での便宜を目的にその後の切断、発現、ライゲーション、アニーリング、または結合のためのシグナルを含みうるAAAA/aaaaペアの1つが隣接すると考えられる。

10

【0228】

代替の態様によれば、ヒトゲノム中において、遺伝的関連研究(例えば、罹患症例での遺伝的変異を対照での同一部位と比較する場合)で配列決定の「標的とされる」のに妥当なセグメントであると思われるヒトエクソンやその他の重要な保存要素は300,000をはるかに上回ると推定される。たとえ非常に安価なDNA配列決定法があるにしても、ゲノムのサブセットに対するアッセイ法(例えば、よくあるがんゲノム調査やプロファイリング)を開発し利用する要求が存在している。

20

【0229】

本発明のこの態様により、本実施例VIIIの protokol を行うが、ただし遺伝子をプライマーと置き換える。288組のユニバーサル・プライマーペアを利用して、任意のプライマーペアの任意量入手することができる。この結果が症例/対照の配列決定に向けた多数のプライマーセットの複合検査と分配の方法であり、1つの具体的な態様によれば、それを1枚の384ウェルプレートで行うことができる。

【0230】

参考文献

Prodromou and Pearl (1992) *Protein Eng.* 5:827

30

Dillon, P.J. and Rosen, C.A. (1993) In White, B.A. (ed.), *PCR Protocols: Current Methods and Applications*. Humana Press, Totowa, NJ, Vol. 15, pp. 263–267.

Sardana et al. (1996) *Plant Cell Rep.* 15:677

Stemmer (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91:10747

Ho et al. (1989) *Gene* 77:51

40

各参考文献は全ての目的でその全体が参照により本明細書に組み入れられる。

【0231】

実施例IX

大腸菌小リボソームサブユニット

以下の3つの遺伝子(rs1、rs3およびrs14)を大腸菌での発現用に最適化した。

【0232】

大腸菌抽出物での発現用に最適化した遺伝子rs1

ATGACCGAATCATTTCGCACAGTTATTCGAGGAAAGTTTAAAAGAAATTGAAACC  
 CGTCCGGGCTCAATCGTGCCTGGCGTAGTTGTTGCTATAGACAAAGATGTTGTTT  
 TAGTTGATGCAGGTTTAAAAGTGAAAGTGCAATTCCGGCAGAACAGTTTAAAA  
 ATGCACAGGGTGAATTAGAAATTCAGGTAGGCGATGAGGTAGATGTAGCTTTAG  
 ATGCAGTAGAGGATGGCTTCGGTGAAACCTTATTAAGTCGTGAAAAAGCAAAAC  
 GTCATGAAGCATGGATTACCTTAGAAAAAGCATATGAAGATGCAGAAACTGTAA  
 CCGGTGTAATCAACGGCAAGGTAAAAGGCGGCTTTACTGTTGAGTTAAATGGTAT  
 TCGTGCATTTTTACCAGGCAGTTTAGTTGATGTTTCGTCCGGTTCGTGATACCTTAC  
 ATTTAGAAGGTAAAGAATTAGAATTTAAAGTAATCAAATTAGATCAGAAACGTA  
 ACAACGTAGTAGTTAGTCGTGTCGAGTAATCGAAAGTGAAAACCTCAGCAGAAC  
 GTGATCAGTTATTAGAAAATCTGCAAGAAGGTATGGAAGTAAAGGGTATTGTAA  
 AGAATTTAACCGATTATGGTGCATTTGTCGACTTAGGGCGGCGTTGATGGTTTATT  
 ACACATCACCGACATGGCATGGAAACGTGTTAAACATCCGAGTGAAATCGTAAA  
 TGTTGGCGACGAGATAACCGTAAAGGTTTTAAAATTTGATCGTGAACGTACCCGT  
 GTTAGTTTAGGATTGAAACAGTTAGGTGAAGATCCGTGGGTTGCAATTGCAAAAC  
 GTTATCCGGAAGGTACCAAATTAACCGGCAGAGTTACCAATTTAACCGATTATGG  
 TTGCTTCGTAGAGATCGAGGAAGGTGTAGAGGGCCTTGTTACAGTTAGTGAATG  
 GACTGGACCAATAAAAACATCCATCCGAGTAAAGTAGTAAACGTAGGTGACGTA  
 GTGGAGGTAATGGTTTTAGATATCGACGAAGAACGTCGTCGTATTAGTTTAGGTT  
 TAAAACAGTGCAAGGCTAACCCGTGGCAGCAGTTCGCTGAAACCCATAATAAAG  
 GCGACCGTGTAGAGGGTAAGATTA AAAAGCATTACTGACTTTGGCATCTTTATCGG  
 CCTTGACGGTGGCATCGATGGTCTTGTCCATTTAAGTGACATCAGTTGGAATGTT  
 GCAGGTGAAGAAGCTGTACGTGAATATAAAAAGGAGACGAAATTGCAGCAGTT  
 GTTTTACAGGTAGACGCAGAACGTGAACGTATTAGTCTGGGCGTAAAGCAACTG  
 GCAGAAGACCCGTTTAAACAATTGGGTAGCTTTAAATAAAAAGGTGCAATTGTT  
 ACCGGTAAAGTTACCGCAGTAGACGCAAAAGGTGCAACTGTAGA ACTGGCTGAC  
 GCGGTTGAAGGCTACTTACGTGCAAGTGAAGCAAGTCGTGATCGTGTTGAAGAT  
 GCAACCCTTGTCTTAAGTGTAGGCGATGAAGTTGAAGCAAAATTTACCGGTGTAG  
 ACCGTAAAATCGTGCAATTAGTTTAAAGTGTTTCGTGCAAAAGATGAAGCAGATG  
 AAAAAGATGCAATTGCAACCGTTAATAAACAGGAAGATGCAAATTTTAGTAATA  
 ATGCAATGGCAGAAGCATTTAAGCAGCAAAAGGTGAATAA (SEQ ID NO:7)

10

20

30

## 【 0 2 3 3 】

大腸菌抽出物での発現用に最適化した遺伝子 rs3

ATGGGACAGAAAGTTCATCCGAACGGCATTTCGTCTGGGCATCGTAAAGCCTTGG  
 AATAGTACCTGGTTCGCTAATACCAAAGAATTTGCAGATAATCTGGACAGTGACT  
 TCAAAGTTCGTGAGTATTTAACCAAAGA ACTGGCTAAAGCAAGTGTTAGTCGTAT  
 TGTTATTGAACGTCCGGCAAAAAGTATTCGTGTTACCATTCATACCGCACGTCCG  
 GGAATAGTTATTGGTAAAAAAGGTGAAGACGTAGAAAAATTACGTAAAGTTGTT  
 GCAGACATAGCAGGCGTACCGGCACAGATTAATATTGCAGAAGTTCGTAAACCG  
 GAATTAGATGCAAACTTGTGCGCAGATAGTATTACCAGTCAGTTAGAAAGAAGA  
 GTTATGTTCCGTGCAATGAAGAGAGCAGTTCAGAACGCTATGCGTTTAGGTG  
 CAAAAGGTATTAAGTTGAAGTTAGTGGTTCGTTTAGGTGGTGCAGAAATTGCAC  
 GTACCGAATGGTATCGTGAAGGTCGTGTTCCGTTACATACCTTACGTGCAGATAT  
 TGATTATAACACAAGTGAAGCACACACTACCTATGGCGTAATTGGTGTAAAGGTA  
 TGGATTTTCAAGGGTGAAATTTTAGGTGGTATGGCAGCAGTTGAACAGCCGGAA  
 AAACCGGCAGCACAGCCGAAAAACAGCAGCGTAAAGGTTCGTAAATAA (SEQ ID  
 NO:70)

40

## 【 0 2 3 4 】

大腸菌抽出物での発現用に最適化した遺伝子 rs14

ATGGCAAAACAGTCAATGAAAGCTAGAGAAGTTAAACGTGTTGCATTAGCAGAT  
 AAATATTTTCGCTAAACGTGCAGAATTAAGCAATCATCTCAGACGTTAATGCAT  
 CAGACGAAGATCGTTGGAACGCAGTTTTAAAATTACAGACCTTACCGCGTGACTC  
 AAGTCCGAGTCGTCAGCGTAACAGATGTCGTCAGACCGGCAGACCGCATGGCTT  
 CTTACGTAAATTCGGCTTAAGTAGAATCAAAGTTCGTGAAGCAGCAATGCGTGGT  
 GAAATTCCGGGTTTTAAAAAAGCAAGTTGGTAA (SEQ ID NO:71)

【 0 2 3 5 】

配列 rs1、rs3、および rs14由来のオリゴヌクレオチド

- rs1-1: TAAACAGGAAGATGCAAATTTTAGTAATAATGCAATGGCAGAAGCATT  
 TAAAGCAGCAAAGGTGAATAA (SEQ ID NO:72) 10
- rs1-2: AGATGAAGCAGATGAAAAAGATGCAATTGCAACCGTTAATAAACAG  
 GAAGATGCAAATTTTAGTAATAAT (SEQ ID NO:73)
- rs1-3: GGTGTAGACCGTAAAAATCGTGCAATTAGTTTAAGTGTTTCGTGCAA  
 AGATGAAGCAGATGAAAAAGATG (SEQ ID NO:74)
- rs1-4: GCAACCCTTGTCTTAAGTGTAGGCGATGAAGTTGAAGCAAATTTACC  
 GGTGTAGACCGTAAAAATCGTG (SEQ ID NO:75) 20
- rs1-5: AAGGCTACTTACGTGCAAGTGAAGCAAGTCGTGATCGTGTTGAAGATG  
 CAACCCTTGTCTTAAGTGTAGG (SEQ ID NO:76)
- rs1-6: CGCAGTAGACGCAAAGGTGCAACTGTAGAAGTGGCTGACGGCGTTG  
 AAGGCTACTTACGTGCAAGTGAA (SEQ ID NO:77)
- rs1-7: CAATTGGGTAGCTTTAAATAAAAAAGGTGCAATTGTTACCGGTAAAGT  
 TACCGCAGTAGACGCAAAGGT (SEQ ID NO:78)
- rs1-8: TATTAGTCTGGGCGTAAAGCAACTGGCAGAAGACCCGTTTAAACAATTG  
 GGTAGCTTTAAATAAAAAAGGT (SEQ ID NO:79) 30
- rs1-9: ACGAAATTGCAGCAGTTGTTTTACAGGTAGACGCAGAACGTGAACGTA  
 TTAGTCTGGGCGTAAAGCAACT (SEQ ID NO:80)
- rs1-10: AGTTGGAATGTTGCAGGTGAAGAAGCTGTACGTGAATATAAAAAAGG  
 AGACGAAATTGCAGCAGTTGTTT (SEQ ID NO:81)
- rs1-11: TTATCGGCCTTGACGGTGGCATCGATGGTCTTGTCCATTTAAGTGACA  
 TCAGTTGGAATGTTGCAGGTGA (SEQ ID NO:82) 40
- rs1-12: TAAAGGCGACCGTGTAGAGGGTAAGATTAAGCATTACTGACTTTG  
 GCATCTTTATCGGCCTTGACGGT (SEQ ID NO:83)

- rs1-13: TTAAAACAGTGCAAGGCTAACCCGTGGCAGCAGTTCGCTGAAACCCA  
TAATAAAGGCGACCGTGTAGAGG (SEQ ID NO:84)
- rs1-14: GTAATGGTTTTAGATATCGACGAAGAACGTCGTCGTATTAGTTAGGT  
TTAAAACAGTGCAAGGCTAACC (SEQ ID NO:85)
- rs1-15: ATCCATCCGAGTAAAGTAGTAAACGTAGGTGACGTAGTGGAGGTAAT  
GGTTTTAGATATCGACGAAGAAC (SEQ ID NO:86)
- rs1-16: GGCCTTGTTACGTTAGTGAAATGGACTGGACCAATAAAAACATCCA  
TCCGAGTAAAGTAGTAAACGTAG (SEQ ID NO:87) 10
- rs1-17: CAATTTAACCGATTATGGTTGCTTCGTAGAGATCGAGGAAGGTGTAG  
AGGGCCTTGTTACGTTAGTGAA (SEQ ID NO:88)
- rs1-18: TTGCAAACGTTATCCGGAAGGTACCAAATTAACCGGCAGAGTTACC  
AATTTAACCGATTATGGTTGCTT (SEQ ID NO:89)
- rs1-19: CCCGTGTTAGTTTAGGATTGAAACAGTTAGGTGAAGATCCGTGGGTT  
GCAATTGCAAACGTTATCCGGA (SEQ ID NO:90) 20
- rs1-20: GGCGACGAGATAACCGTAAAGGTTTTAAAATTTGATCGTGAACGTAC  
CCGTGTTAGTTTAGGATTGAAAC (SEQ ID NO:91)
- rs1-21: CCGACATGGCATGGAAACGTGTTAAACATCCGAGTGAAATCGTAAAT  
GTTGGCGACGAGATAACCGTAAA (SEQ ID NO:92)
- rs1-22: CCGATTATGGTGCATTTGTCGACTTAGGCGGCGTTGATGGTTTATTAC  
ACATCACCGACATGGCATGGAA (SEQ ID NO:93)
- rs1-23: AGAAAATCTGCAAGAAGGTATGGAAGTAAAGGGTATTGTAAAGAATT  
TAACCGATTATGGTGCATTTGTC (SEQ ID NO:94) 30
- rs1-24: GTGCAGTAATCGAAAGTGAAAACCTCAGCAGAACGTGATCAGTTATTA  
GAAAATCTGCAAGAAGGTATGGA (SEQ ID NO:95)
- rs1-25: TAATCAAATTAGATCAGAAACGTAACAACGTAGTAGTTAGTCGTCGT  
GCAGTAATCGAAAGTGAAAACCTC (SEQ ID NO:96)
- rs1-26: TGATACCTTACATTTAGAAGGTAAAGAATTAGAATTTAAAGTAATCAA  
ATTAGATCAGAAACGTAACAAC (SEQ ID NO:97) 40
- rs1-27: TTACCAGGCAGTTTAGTTGATGTTTCGTCCGGTTCGTGATACCTTACAT  
TTAGAAGGTAAAGAATTAGAAT (SEQ ID NO:98)
- rs1-28: GTAAAAGGCGGCTTTACTGTTGAGTTAAATGGTATTCGTGCATTTTAA  
CCAGGCAGTTTAGTTGATGTTT (SEQ ID NO:99)

rs1-29: AAAGCATATGAAGATGCAGAAACTGTAACCGGTGTAATCAACGGCAA  
GGTAAAAGGCGGCTTTACTGTTG (SEQ ID NO:100)

rs1-30: AAGTCGTGAAAAAGCAAAACGTCATGAAGCATGGATTACCTTAGAAA  
AAGCATATGAAGATGCAGAAACT (SEQ ID NO:101)

rs1-31: AGATGTAGCTTTAGATGCAGTAGAGGATGGCTTCGGTGAAACCTTAT  
TAAGTCGTGAAAAAGCAAAACGT (SEQ ID NO:102)

rs1-32: AAATGCACAGGGTGAATTAGAAATTCAGGTAGGCGATGAGGTAGATG  
TAGCTTTAGATGCAGTAGAGGAT (SEQ ID NO:103)

10

rs1-33: ATGCAGGTTTAAAAAGTGAAAGTGCAATTCGGCAGAACAGTTTAAA  
AATGCACAGGGTGAATTAGAAAT (SEQ ID NO:104)

rs1-34: GTGCGTGGCGTAGTTGTTGCTATAGACAAAGATGTTGTTTTAGTTGAT  
GCAGGTTTAAAAAGTGAAAGTG (SEQ ID NO:105)

rs1-35: ACAGTTATTCGAGGAAAGTTTAAAAGAAATTGAAACCCGTCCGGGCT  
CAATCGTGCCTGGCGTAGTTGTT (SEQ ID NO:106)

20

rs1-36: ATGACCGAATCATTTCGCACAGTTATTCGAGGAAAGTTTAAAAGAAAT  
(SEQ ID NO:107)

rs3-1: GGCAGCAGTTGAACAGCCGGAAAAACCGGCAGCACAGCCGAAAAAAC  
AGCAGCGTAAAGGTCGTAAATAA (SEQ ID NO:108)

rs3-2: GGCGTAATTGGTGTAAAGGTATGGATTTTCAAGGGTGAAATTTTAGGT  
GGTATGGCAGCAGTTGAACAGC (SEQ ID NO:109)

rs3-3: TACGTGCAGATATTGATTATAACACAAGTGAAGCACACACTACCTATG  
GCGTAATTGGTGTAAAGGTATG (SEQ ID NO:110)

30

rs3-4: TACCGAATGGTATCGTGAAGGTCGTGTTCCGTTACATACCTTACGTGCA  
GATATTGATTATAACACAAGT (SEQ ID NO:111)

rs3-5: GTATTAAAGTTGAAGTTAGTGGTCGTTTAGGTGGTGCAGAAATTGCAC  
GTACCGAATGGTATCGTGAAGG (SEQ ID NO:112)

rs3-6: AAGAGAGCAGTTCAGAACGCTATGCGTTTAGGTGCAAAGGTATTAAA  
GTTGAAGTTAGTGGTCGTTTAG (SEQ ID NO:113)

40

rs3-7: GTATTACCAGTCAGTTAGAAAGAAGAGTTATGTTCCGTCGTGCAATGA  
AGAGAGCAGTTCAGAACGCTAT (SEQ ID NO:114)

rs3-8: GTAAACCGGAATTAGATGCAAAACTTGTCGCAGATAGTATTACCAGTC  
AGTTAGAAAGAAGAGTTATGTT (SEQ ID NO:115)

- rs3-9: CAGACATAGCAGGCGTACCGGCACAGATTAATATTGCAGAAGTTCGTA  
AACCGGAATTAGATGCAAAACT (SEQ ID NO:116)
- rs3-10: TAGTTATTGGTAAAAAAGGTGAAGACGTAGAAAAATTACGTAAAGTT  
GTTGCAGACATAGCAGGCGTACC (SEQ ID NO:117)
- rs3-11: AAAAAGTATTCGTGTTACCATTTCATACCGCACGTCCGGGAATAGTTAT  
TGGTAAAAAAGGTGAAGACGTA (SEQ ID NO:118)
- rs3-12: GGCTAAAGCAAGTGTTAGTCGTATTGTTATTGAACGTCCGGCAAAAA  
GTATTCGTGTTACCATTTCATACC (SEQ ID NO:119) 10
- rs3-13: TCTGGACAGTGACTTCAAAGTTCGTTCAGTATTTAACCAAAGAAGTGGC  
TAAAGCAAGTGTTAGTCGTATT (SEQ ID NO:120)
- rs3-14: CCTTGGAAATAGTACCTGGTTCGCTAATACCAAAGAATTTGCAGATAAT  
CTGGACAGTGACTTCAAAGTTC (SEQ ID NO:121)
- rs3-15: ATGGGACAGAAAGTTCATCCGAACGGCATTTCGTCTGGGCATCGTAAA  
GCCTTGGAATAGTACCTGGTTCG (SEQ ID NO:122) 20
- rs3-16: ATGGGACAGAAAGTTCATCC (SEQ ID NO:123)
- rs14-1: AAGTAGAATCAAAGTTCGTGAAGCAGCAATGCGTGGTGAAATTCCGG  
GTTTAAAAAAGCAAGTTGGTAA (SEQ ID NO:124)
- rs14-2: TCGTCAGACCGGCAGACCGCATGGCTTCTTACGTAAATTCGGCTTAAG  
TAGAATCAAAGTTCGTGAAGCA (SEQ ID NO:125)
- rs14-3: AAAATTACAGACCTTACCGCGTGACTCAAGTCCGAGTCGTCAGCGTAA  
CAGATGTCGTCAGACCGGCAGA (SEQ ID NO:126) 30
- rs14-4: CATCTCAGACGTTAATGCATCAGACGAAGATCGTTGGAACGCAGTTTT  
AAAATTACAGACCTTACCGCGT (SEQ ID NO:127)
- rs14-5: GCATTAGCAGATAAATATTTGCTAAACGTGCAGAATTTAAAGCAATC  
ATCTCAGACGTTAATGCATCAG (SEQ ID NO:128)
- rs14-6: ATGGCAAAACAGTCAATGAAAGCTAGAGAAGTTAAACGTGTTGCATT  
AGCAGATAAATATTTGCTAAAC (SEQ ID NO:129)
- rs14-7: ATGGCAAAACAGTCAATGAAAG (SEQ ID NO:130) 40

## 【 0 2 3 6 】

rs3遺伝子に対し作製された70merの完全なリスト。15merのタグは下線が引かれており、6merのIIS部位は太字になっている。

- rs3-1: CACTCCAGGGTCTCGTTATTTACGACCTTTACGCTGCTGTTTTTTTCGG  
CTGTGCTCGTCGCAGGTGTCAC (SEQ ID NO:131)
- rs3-2: CACTCCAGGGTCTCGCTGTTTTTTTCGGCTGTGCTGCCGGTTTTTCCGG  
CTGTTACGTCGCAGGTGTCAC (SEQ ID NO:132)
- rs3-3: CACTCCAGGGTCTCGGTTTTTCCGGCTGTTCAACTGCTGCCATACCAC  
CTAAAATCGTCGCAGGTGTCAC (SEQ ID NO:133)
- rs3-4: CACTCCAGGGTCTCGCTGCCATACCACCTAAAATTTACCCTTGAAA  
ATCCATACCGTCGCAGGTGTCAC (SEQ ID NO:134) 10
- rs3-5: CACTCCAGGGTCTCGTCACCCTTGAAAATCCATACCTTAACACCAAT  
TACGCCATCGTCGCAGGTGTCAC (SEQ ID NO:135)
- rs3-6: CACTCCAGGGTCTCGATACCTTAACACCAATTACGCCATAGGTAGTG  
TGTGCTTCCGTCGCAGGTGTCAC (SEQ ID NO:136)
- rs3-7: CACTCCAGGGTCTCGGCCATAGGTAGTGTGTGCTTCACTTGTGTTATA  
ATCAATACGTCGCAGGTGTCAC (SEQ ID NO:137) 20
- rs3-8: CACTCCAGGGTCTCGTGCTTCACTTGTGTTATAATCAATATCTGCACG  
TAAGGTACGTCGCAGGTGTCAC (SEQ ID NO:138)
- rs3-9: CACTCCAGGGTCTCGTATAATCAATATCTGCACGTAAGGTATGTAAC  
GGAACACGCGTCGCAGGTGTCAC (SEQ ID NO:139)
- rs3-10: CACTCCAGGGTCTCGGTAAGGTATGTAACGGAACACGACCTTCACGA  
TACCATTCCGTCGCAGGTGTCAC (SEQ ID NO:140)
- rs3-11: CACTCCAGGGTCTCGCGACCTTCACGATACCATTCCGGTACGTGCAAT  
TTCTGCACCGTCGCAGGTGTCAC (SEQ ID NO:141) 30
- rs3-12: CACTCCAGGGTCTCGGGTACGTGCAATTTCTGCACCACCTAAACGAC  
CACTAACTCGTCGCAGGTGTCAC (SEQ ID NO:142)
- rs3-13: CACTCCAGGGTCTCGCACCTAAACGACCACTAACTTCAACTTTAATAC  
CTTTTGCCGTCGCAGGTGTCAC (SEQ ID NO:143)
- rs3-14: CACTCCAGGGTCTCGCACTAACTTCAACTTTAATACCTTTTGCACCTA  
AACGCATCGTCGCAGGTGTCAC (SEQ ID NO:144) 40
- rs3-15: CACTCCAGGGTCTCGAATACCTTTTGCACCTAAACGCATAGCGTTCT  
GAACTGCTCGTCGCAGGTGTCAC (SEQ ID NO:145)
- rs3-16: CACTCCAGGGTCTCGGCATAGCGTTCTGAACTGCTCTCTTCATTGCA  
CGACGGAACGTCGCAGGTGTCAC (SEQ ID NO:146)

- rs3-17: CACTCCAGGGTCTCGCTCTTCATTGCACGACGGAACATAACTCTTCTT  
TCTAACTCGTCGCAGGTGTCAC (SEQ ID NO:147)
- rs3-18: CACTCCAGGGTCTCGGGAACATAACTCTTCTTTCTAACTGACTGGTA  
ATACTATCCGTCGCAGGTGTCAC (SEQ ID NO:148)
- rs3-19: CACTCCAGGGTCTCGTCTTTCTAACTGACTGGTAATACTATCTGCGAC  
AAGTTTTCGTCGCAGGTGTCAC (SEQ ID NO:149)
- rs3-20: CACTCCAGGGTCTCGGTAATACTATCTGCGACAAGTTTTGCATCTAA  
TTCCGGTTCGTCGCAGGTGTCAC (SEQ ID NO:150) 10
- rs3-21: CACTCCAGGGTCTCGAAGTTTTGCATCTAATTCCGGTTTACGAACTTC  
TGCAATACGTCGCAGGTGTCAC (SEQ ID NO:151)
- rs3-22: CACTCCAGGGTCTCGGGTTTACGAACTTCTGCAATATTAATCTGTGC  
CGGTACGCCGTCGCAGGTGTCAC (SEQ ID NO:152)
- rs3-23: CACTCCAGGGTCTCGATATTAATCTGTGCCGGTACGCCTGCTATGTC  
TGCAACAACGTCGCAGGTGTCAC (SEQ ID NO:153) 20
- rs3-24: CACTCCAGGGTCTCGCCTGCTATGTCTGCAACAACCTTTACGTAATTTT  
TCTACGTCGTCGCAGGTGTCAC (SEQ ID NO:154)
- rs3-25: CACTCCAGGGTCTCGAACAACCTTTACGTAATTTTTCTACGTCTTCACC  
TTTTTTACGTCGCAGGTGTCAC (SEQ ID NO:155)
- rs3-26: CACTCCAGGGTCTCGTTTCTACGTCTTCACCTTTTTTACCAATAACTA  
TTCCCGGCGTCGCAGGTGTCAC (SEQ ID NO:156)
- rs3-27: CACTCCAGGGTCTCGACCTTTTTTACCAATAACTATTCCCGGACGTGC  
GGTATGACGTCGCAGGTGTCAC (SEQ ID NO:157) 30
- rs3-28: CACTCCAGGGTCTCGGGACGTGCGGTATGAATGGTAACACGAATACT  
TTTTGCCGCGTCGCAGGTGTCAC (SEQ ID NO:158)
- rs3-29: CACTCCAGGGTCTCGATGGTAACACGAATACTTTTTGCCGGACGTTT  
AATAACAACGTCGCAGGTGTCAC (SEQ ID NO:159)
- rs3-30: CACTCCAGGGTCTCGCCGGACGTTCAATAACAATACGACTAACACTT  
GCTTTAGCCGTCGCAGGTGTCAC (SEQ ID NO:160) 40
- rs3-31: CACTCCAGGGTCTCGAATACGACTAACACTTGCTTTAGCCAGTTCTTT  
GGTTAAACGTCGCAGGTGTCAC (SEQ ID NO:161)
- rs3-32: CACTCCAGGGTCTCGTTTAGCCAGTTCTTTGGTTAAATACTGACGAAC  
TTTGAAGCGTCGCAGGTGTCAC (SEQ ID NO:162)

rs3-33: CACTCCAGGGTCTCGGGTTAAATACTGACGAACTTTGAAGTCACTGT  
CCAGATTACGTCGCAGGTGTCAC (SEQ ID NO:163)

rs3-34: CACTCCAGGGTCTCGTTTGAAGTCACTGTCCAGATTATCTGCAAATTC  
TTTGGTACGTCGCAGGTGTCAC (SEQ ID NO:164)

rs3-35: CACTCCAGGGTCTCGAGATTATCTGCAAATTCTTTGGTATTAGCGAA  
CCAGGTACCGTCGCAGGTGTCAC (SEQ ID NO:165)

rs3-36: CACTCCAGGGTCTCGTGGTATTAGCGAACCAGGTACTATTCCAAGGC  
TTTACGATCGTCGCAGGTGTCAC (SEQ ID NO:166)

10

rs3-37: CACTCCAGGGTCTCGGTAATAATTCCAAGGCTTTACGATGCCAGACG  
AATGCCGTCGTCGCAGGTGTCAC (SEQ ID NO:167)

rs3-38: CACTCCAGGGTCTCGGCCAGACGAATGCCGTTCCGATGAACTTTCT  
GTCCCATACGTCGCAGGTGTCAC (SEQ ID NO:168)

【 0 2 3 7 】

2本の15mer「タグ」プレプライマーをPCRに使用する。

L = 5' CACTCCAGGGTCTCG (SEQ ID NO:169)

R = 5' GTGACACCTGCGACG (SEQ ID NO:170)

20

【 0 2 3 8 】

ヌクレアーゼ切断がギャップによって示された第1のオリゴヌクレオチド配列(rs3-1)に対する二本鎖の70mer。

CACTCCAGGGTCTCG TTATTTACGACCTTTACGCTGCTGTTTTTTTCGGCTG TGCTCGTCGCAGGTGTCAC  
 (SEQ ID NO:171)

gtgagggtcccagagcaata aatgctggaaatgcgacgacaaaaaagccgacacga gcagcgtcccacagag  
 (SEQ ID NO:172)

【 0 2 3 9 】

タグおよび重複部分および逆相補鎖を除去することで、以下の708merが得られる。

TATGGGACAGAAAGTTCATCCGAACGGCATTTCGTCTGGGCATCGTAAAGCCTTGG  
 AATAGTACCTGGTTCGCTAATAACCAAAGAATTTGCAGATAATCTGGACAGTGACT  
 TCAAAGTTCGTCAAGTATTTAACCAAAGAAGTGGCTAAAGCAAGTGTTAGTCGTAT  
 TGTTATTGAACGTCCGGCAAAAAGTATTCGTGTTACCATTTCATACCGCACGTCCG  
 GGAATAGTTATTGGTAAAAAAGGTGAAGACGTAGAAAAATTACGTAAAGTTGTT  
 GCAGACATAGCAGGCGTACCGGCACAGATTAATATTGCAGAAGTTCGTAAACCG  
 GAATTAGATGCAAACTTGTTCGCAGATAGTATTACCAGTCAGTTAGAAAGAAGA  
 GTTATGTTCCGTCGTGCAATGAAGAGAGCAGTTCAGAACGCTATGCGTTTAGGTG  
 CAAAAGGTATTAAGTGAAGTTAGTGGTTCGTTTAGGTGGTGCAGAAATTGCAC  
 GTACCGAATGGTATCGTGAAGGTCGTGTTCCGTTACATACCTTACGTGCAGATAT  
 TGATTATAACACAAGTGAAGCACACACTACCTATGGCGTAATTGGTGTTAAGGTA  
 TGGATTTTCAAGGGTGAATTTTAGGTGGTATGGCAGCAGTTGAACAGCCGGAA  
 AAACCGGCAGCACAGCACAGCCGAAAAAACAGCAGCGTAAAGGTCGTAAATAA  
 (SEQ ID NO:63)

40

【 0 2 4 0 】

最後のPCRに使用したフランキングプライマーは以下である。

rs3-L = 5' ATGGGACAGAAAGTTCATC (SEQ ID NO:36)

rs3-R = 5' TTATTTACGACCTTTACGCT (SEQ ID NO:37)

#### 【 0 2 4 1 】

実施例 X

配列の設計

遺伝子およびオリゴヌクレオチド配列は、JavaプログラムのCAD-PAMを用いて設計した。基本的に、CAD-PAMはn mer(典型的には50mer)の構築用オリゴマーといっそう短い選択用オリゴマー(典型的には26mer)のほぼ最適な重複セットを作製するために、アミノ酸配列、コドン使用頻度、伝令RNA二次構造および構築用オリゴヌクレオチドを放出させるのに使われる制限酵素に対する制約を利用する。隣接する遺伝子構築用オリゴヌクレオチド間のまたは構築用と選択用オリゴヌクレオチドとの間の重複領域の融解温度( $T_m$ )を均等化した。選択用オリゴヌクレオチドに余分のアデニン残基を挿入して、最適なサイズ選択(通常のパムには使われない)のためにオリゴマー長を一定(70mer)にした。 $T_m$ 値は最近接法を用いて算出した(あらゆる目的でその全体が参照により本明細書に組み入れられる Breslauer et al. (1986) Proc. Natl. Acad. Sci. U.S.A. 83:3746)。コドンを固定化させてまたは変化させて、発現向上を可能にした。

10

#### 【 0 2 4 2 】

実施例 XI

合成オリゴヌクレオチドの増幅

現行のマイクロチップは表面積が非常に小さく、故にごく少量のオリゴヌクレオチドしか産生しない。溶液中に放出された場合、オリゴヌクレオチドは1配列当たりピコモルまたはそれより低い濃度、つまり、例えば、PCRアッセムブリ、ライゲーションアッセムブリなどに関与するものなどの二分子プライミング反応を効率的に推進するのに十分に高くない濃度で存在することになる。

20

#### 【 0 2 4 3 】

この規模の問題に取り組むため、マイクロチップから得られたオリゴヌクレオチドを各配列のわずかおよそ $10^5$  (低密度アレイの場合には $10^9$ )から最大で $10^9$  (または $10^{12}$ )分子まで増幅させ、それによりその後の選択およびアッセムブリステップを可能にした。統合過程の概要が図6に示されている。

30

#### 【 0 2 4 4 】

この増幅方法の場合、ユニバーサルプライマー配列が隣接したオリゴヌクレオチドをプログラムブル・マイクロチップ上で合成した。これによって、化学的または酵素的処理によりマイクロチップから放出可能な、 $10^2 \sim 10^5$ 個の異なるオリゴヌクレオチドのプールが作製される。IIS型制限酵素認識部位を含んだプライマーを用いたポリメラーゼ連鎖反応(PCR)によって、放出オリゴヌクレオチドを増幅させた。対応する制限酵素でPCR産物を消化することで、遺伝子またはゲノムアッセムブリに使われる十分な量の純粋なオリゴヌクレオチド配列が得られた。

#### 【 0 2 4 5 】

この手法の実現可能性を初めにAtactic/Xeotron 4K (すなわち、3,968個の合成チャンバ)光プログラマブル微小流体マイクロアレイ(Zhou, X. et al., Nucleic Acids Res. 32 : 5409-5417 (2004))で証明した。オリゴヌクレオチドの合成およびマイクロチップからの切断をモニターするため、オリゴヌクレオチドの5'末端にフルオレセインを結合させた。切断の前後に、マイクロチップをマイクロアレイ・スキャナーでスキャンした。オリゴヌクレオチドの切断部分を、相補的なオリゴヌクレオチド配列を用いて合成された「品質評価(QA)-チップ」上にハイブリダイズさせた。これらの結果から、個々のオリゴヌクレオチドが合成され、QA-チップハイブリダイゼーション過程によって測定可能とされる量でマイクロチップからほぼ完全に放出されていることが証明された。4Kマイクロチップの各チャンバから放出されたオリゴヌクレオチドの典型的収量は、定量的PCR (Zhou, X. et

40

50

al. *Nucleic Acids Res.* 32: 5409-5417 (2004))によって測定したところ、約5 fmoleであった。オリゴヌクレオチド配列に隣接するユニバーサルプライマーに特異的にアニーリングしたプライマーを用い、PCR反応を行って、百万倍を超えるオリゴヌクレオチドを増幅させた。

#### 【0246】

##### 実施例XII

##### 合成および/または増幅オリゴヌクレオチドのエラー低減

オリゴヌクレオチド合成の間に被る変異は、アッセンブルされたDNA分子中のエラーの主因であり、取り除くのは費用がかかり且つ困難である(Cello et al. (2002) *Science* 297:1016; Smith et al. (2003) *Proc. Natl. Acad. Sci. USA* 100:15440)。本実施例は、そのような変異を有するオリゴヌクレオチドを除去するストリンジェントなハイブリダイゼーションに基づく簡便な方法について記述する。構築用オリゴヌクレオチド中の変異に対する負の選択をするため、合すると構築用オリゴヌクレオチドの全長に及ぶ、ビーズに固定化された短い相補的な選択用オリゴヌクレオチド(図5)の2つのプールに、これらのオリゴヌクレオチドを連続してハイブリダイズさせた。選択用オリゴヌクレオチドは全て、その長さを変化させることでほぼ同一の融解温度を有するように設計された。適切なハイブリダイゼーション条件の下で、塩基ミスマッチまたは欠失による選択用と構築用オリゴヌクレオチドとの間の不完全なペアはいっそう低い融解温度を有し、不安定である。ハイブリダイゼーション、洗浄および溶出のサイクル後、選択用オリゴヌクレオチドに完全に適合する配列を有するオリゴヌクレオチドが選択的に保持され濃縮された。IIS型制限酵素を用いたPCR産物の消化によって、オリゴヌクレオチドの両端から汎用プライマー配列が除去された。これらの実験では、増幅用タグは選択の直前に除去された。しかしながら、消化を後回しにすれば、オリゴヌクレオチドをPCRにより再増幅し、さらなるラウンドのハイブリダイゼーション選択に供することができる。理論により束縛されることを意図するわけではないが、構築および選択用オリゴヌクレオチド上の適合位置で起こる相補的な変異の確立は非常に小さいので、原理上は、変異を有するほとんどのオリゴヌクレオチドはこの選択手順によって除去することができる。

10

20

#### 【0247】

構築用オリゴヌクレオチドと同様に、選択用オリゴヌクレオチドを合成しプログラマブル・マイクロアレイから放出した。アームを有する選択用オリゴヌクレオチドをPCRにより増幅し、遺伝子構築用オリゴヌクレオチドに相補的な鎖を5'末端でビオチンを用いて標識し、ストレプトアビジンビーズ上に選択的に固定化した。標識されていない鎖を変性し除去した。固定化された選択用オリゴヌクレオチドは、正しい50塩基対の構築用オリゴヌクレオチドを選択的に保持した。

30

#### 【0248】

エラー低減された構築用オリゴヌクレオチドは、遺伝子アッセンブリに適している。自動化を促進するため、一段階ポリメラーゼアッセンブリ多重化(PAM)反応がオリゴヌクレオチドの単一プールからの複数の遺伝子合成向けに開発された。単一断片アッセンブリ法では、伝統的に2つまたは3つの段階(ライゲーション、アッセンブリおよびPCR)が使われてきた(Cello, J., et al., *Science* 297: 1016-1018 (2002); Smith, H. O. et al., *Proc. Natl. Acad. Sci. USA* 100: 15440-15445 (2003); Stemmer, W. P. et al., *Gene* 164: 49-53 (1995))。PAMの場合、遺伝子に隣接するプライマーペアを熱安定性ポリメラーゼおよびdNTPsとともに、遺伝子構築用オリゴヌクレオチドのプールに(プライマーペアをオリゴヌクレオチドよりも高い濃度で)加えた。重複するオリゴヌクレオチドの伸長と引き続く複数の完全長遺伝子の増幅は、サーマルサイクラーを用いて密閉型試験管中、一段階反応で行った。異なる汎用アダプター配列を各遺伝子または遺伝子セットの末端に組み入れることができたので、相補的なアダプタープライマーペアのセット(例えば、標準的なマルチウェルプレートに合わせた96または394種の汎用アダプター)を予め合成して、遺伝子特異的なPAMプライマーペアを合成する費用負担を回避することや自動化を促進することができる。

40

50

## 【0249】

ミスマッチ変異を除去するハイブリダイゼーション選択法の効率を測定するため(Eason, R. G. et al. Proc. Natl. Acad. Sci. USA 101: 11046-11051 (2004))、3通りの異なる方法で精製された、つまり精製されていない、ポリアクリルアミドゲル電気泳動(PAGE)精製されたまたはハイブリダイゼーション精製されたマイクロチップ合成によるオリゴヌクレオチドの同一プールを用いて、遺伝子を構築した。これらの遺伝子をクローニングし、各分類由来の無作為のクローンを両方向から配列決定して、各分類に対するエラーの種類と割合を測定した。(図38)に示されるように、精製されていないオリゴヌクレオチドを用いて合成された遺伝子は最も高いエラー率(160中1 bp)を有し; (ライゲーションまたはPAMを用いた)遺伝子アッセムブリの方法では相違を生じなかった。オリゴヌクレオチドのPAGE精製は、主に欠失変異の除去を通じてエラー率を450中1 bpに減少させた。この割合はPAGE精製を利用した他のグループによって報告されている数量(Cello, J., et al., Science 297: 1016-1018 (2002); Smith, H. O. et al., Proc. Natl. Acad. Sci. USA 100: 15440-15445 (2003))に匹敵する。ハイブリダイゼーション選択では、エラー率がおよそ1,394bp中1 bpにさらに低下した。

10

## 【0250】

## 実施例XII

## 単一プール中での複数遺伝子の並行アッセムブリ

マイクロチップを利用して、大腸菌の小リボソームサブユニットを構成する21種のタンパク質コード遺伝子のコドン改変形を再設計し合成した。これら21種のタンパク質の天然形の翻訳効率は、インビトロにおいてこれらのタンパク質が高い発現レベルを有するにもかかわらず、インビトロにおいて非常に低い(Culver, G. M. & Noller, H. F. RNA 5: 832-843 (1999))。コドン使用頻度を再設計することは、タンパク質翻訳効率を増加させる方法であるが、ほとんど理想的なコドンから始める場合にはこれはいっそう達成しづらい。その他多くのタンパク質はこのインビトロ系で十分に発現されるので、一部の問題は二次構造が原因であった(場合によってはT7ポリメラーゼを介した転写の速度が翻訳よりも8倍高いという事実によって悪影響を受けたかもしれない)(Iost, I., et al., J. Bacteriol. 174: 619-622 (1992); Iost, I. & Dreyfus, M. Nature 372: 193-196 (1994))と仮定された。コドンを、二次構造をあまり持たない可能性が高い配列と置き換えた(例えば、G + C含量を低下させることにより)。CAD-PAMソフトウェア(図7)は、21種のリボソーム遺伝子に対し重複する50-bpのオリゴヌクレオチド配列(70mer中に包含される)を設計し、それら全てを4K Xeo chip上で合成した。これらのオリゴヌクレオチドを処理し、選択用オリゴヌクレオチドを用いてハイブリダイゼーション選択し、その後、複数のPAM反応で21種のリボソーム遺伝子を構築するために使用した。インビトロ転写翻訳共役反応を利用し、エラーなしのクローンを大腸菌で試験した。合成遺伝子の翻訳プロファイルを測定した。いくつかのコドン改変遺伝子は、その各野生型遺伝子と比べて、大腸菌抽出物中での翻訳レベルが高かった。遺伝子単位間に固有の約30merの重複リンカーを導入し、連続PAM反応を行うことで、これらの21種の遺伝子を連続PAM反応により結合して、約14.6 kbのアッセムブリのプールを得た。高忠実度PCR反応によって作製された、ともに構築体全体を網羅するすべての重複DNAセグメント由来の平均して4つの個々のクローンを配列決定することにより、正しいアッセムブリを確認した。正しい入力遺伝子配列から始めることにより、および高忠実度のポリメラーゼに基づく伸長反応の繰返しを通じ、アッセムブリ過程は、図38に示される方法(これは全て合成エラーを含むオリゴヌクレオチドから始まった)のいずれかよりも低いエラー率(約7,300 bp中1 bp)をもたらした。このことから、遺伝子アッセムブリのエラーの主因が、ポリメラーゼ校正活性ではなくオリゴヌクレオチド化学合成によってもたらされることが明示された。PCR産物の長さを増やすことは後のアッセムブリの収率を低下させるものと予想されうるが、反応成分の数が減少し、それで効率は高いままである。PAMによる長さが制限になるならば、相同組換えを利用して、アッセムブリをメガベース範囲で可能にすることができる。

20

30

40

## 【0251】

50

成功した幾つかのアッセンブリ反応は、本明細書に記述される方法を用いて行われた。例えば、21種のリボソーム遺伝子の14-kbのオペロンは、本明細書に記述されるポリメラーゼアッセンブリ多重化法を用いてアッセンブルされた。完全長の断片の産生はゲル電気泳動により確認した。さらに、s19遺伝子をオリゴ(6.7メガベース) 95,376種のNimbelgenカスタムアレイ由来のオリゴ混合物からアッセンブルすることに成功した。この結果はゲル電気泳動により確認した。

#### 【0252】

実施例XIV

実施例XI~XIIの方法

#### 配列の設計

遺伝子およびオリゴヌクレオチド配列は、本明細書にさらに記述されるとおりJavaプログラムのCAD-PAMを用いて設計した。基本的に、CAD-PAMは $n$  mer(典型的には50mer)の構築用オリゴマーといっそう短い選択用オリゴマー(典型的には26mer)のほぼ最適な重複セットを作製するために、アミノ酸配列、コドン使用頻度、伝令RNA二次構造および構築用オリゴヌクレオチドを放出させるのに使われる制限酵素に対する制約を利用する。隣接する遺伝子構築用オリゴヌクレオチド間のまたは構築用と選択用オリゴヌクレオチドとの間の重複領域の融解温度( $T_m$ )を均等化した。選択用オリゴヌクレオチドに余分のアデニン残基を挿入して、最適なサイズ選択(通常のパAMには使われない)のためにオリゴマー長を一定(70mer)にした。 $T_m$ 値は最近接法を用いて算出した(Breslauer, K. J., et al., Proc. Natl. Acad. Sci. USA 83: 3746-3750 (1986))。コドンを固定化させてまたは変化させて、

10

20

#### 【0253】

#### オリゴヌクレオチドのマイクロチップ合成、増幅および選択

オリゴヌクレオチドは5'末端および3'末端のリン酸基をウラシル残基の3'-ヒドロキシ末端にカップリングさせて、光プログラマブル微小流体マイクロチップ上で合成した。合成後、RNase Aを用いてまたは水酸化アンモニウム処理(標準的なオリゴヌクレオチド合成のように脱保護に使われる)により、オリゴヌクレオチドを切断し、続けて沈殿を行った。20mer(初めは末端の10塩基に相補的な)でPCR増幅された遺伝子構築用オリゴヌクレオチドをIIS型制限酵素BsaIおよびBseRIで消化した(「PAGE」対照を除いてはゲル精製なしで)。ストレプトアビジン磁気ビーズ上でのビオチン標識済みの選択用オリゴヌクレオチドの固定化(Dynal Biotech, Brown Deer, WI)およびビオチン化されていない鎖の除去は、報告のように行った(Espelund, M., et al., Nucleic Acids Res. 18: 6157-6158 (1990))。構築用オリゴヌクレオチドを95 で3分間変性させ、ローター上42 で14~16時間ハイブリダイゼーション緩衝液(5×SSPET緩衝液、50%ホルムアミド、0.2 mg ml<sup>-1</sup> BSA)中で選択用オリゴヌクレオチドにハイブリダイズさせた。0.5×SSPETを用いて3回および洗浄緩衝液(20 mM Tris-HCl pH 7.0、5 mM EDTA、4 mM NaCl)を用いて3回室温でビーズを洗浄した。0.1 M NaOH中で15分間の変性とその後の中和により、構築用オリゴヌクレオチドを回収した。

30

40

#### 【0254】

#### ポリメラーゼアッセンブリ多重化反応

PAM反応は、オリゴヌクレオチド混合物2 μl、0.4 μMの各遺伝子末端プライマーペア、1×dNTP混合物および1×緩衝液中のAdvantage 2ポリメラーゼ混合物(Clontech ADVANTAGE 2(商標)PCRキット) 0.5 μlを含有する反応液25 μl中で行った。試料を95 で3分間変性させ、その後95 で30秒間、49 で1分間および68 で1 kb当たり1分間の40~45回の熱サイクルにかけ、その後68 で10分間終結させた。連続PAM反応を利用して、複数の遺伝子を結合させた。初めに、RTS大腸菌直鎖状鑄型生成キット(Roche)を用いたPCRにより、21種のリボソームタンパク質遺伝子の正しい配列のHis6タグ付き直鎖状発現構築体を事前構築した。これらの構築体をその後、同一の $T_m$ を有する固有の約30merのリンカー(各0.4 μM、Integrated DNA Technologies社)を導入して第2のPAM反応に十分な遺伝子間重複領域を作製させる別のPCR反応で鑄型として利用した。これらの場合、次の3つの大きな断

50

片RS1-5 (1-5,513)、RS6-13 (5,483-10,526)およびRS14-21 (10,497-14,593)を別のRoche Expand long template PCR反応で作製した。これらの断片をゲル精製し、RS1-21 (1-14,593)を用いた最後のアッセンプリ反応で完全な14,593 bpのオペロンにアッセンプルした。後の2つのアッセンプリの場合、試料を92 °Cで2分間変性させ、その後92 °Cで30秒間、65 °Cで1分間および68 °Cで1 kb当たり1分間の熱サイクル10回にかけ、その後92 °Cで30秒間、65 °Cで1分間および68 °Cで1 kb当たり1分間に加えて1サイクル当たり10秒のさらなるサイクル25回にかけ、68 °Cで10分間終結させた。

【0255】

#### インビトロ転写翻訳共役

アッセンプルされた遺伝子をクローニングし、エラーなしのクローンを配列決定により選択した。インビトロでのタンパク質発現に用いる直鎖状構築体は、Roche RTS大腸菌直鎖状鑄型生成セット、His-タグを用いて作製した。インビトロ転写翻訳共役は、Roche Rapid Translation System RTS 100大腸菌HYキットを用いて行った。タンパク質は標準的な手順を利用して、抗His6-ペルオキシダーゼ抗体(Roche)を用いたウエスタンブロッティングにより検出した。

【0256】

#### 等価物

その他の態様は当業者には明らかであると思われる。前記述は明確にするためだけに与えられており単なる例示にすぎないことが理解されるべきである。本発明の趣旨および範囲は上記の例に限定されないが、以下の特許請求の範囲により包含される。上記に引用した全ての出版物および特許出願は、個々の出版物または特許出願のそれぞれが参照により確かに組み入れられることが具体的に示されているかのごとく、全ての目的でその全体が参照により本明細書に組み入れられる。

【図面の簡単な説明】

【0257】

【図1】カスタムマイクロアレイからの遊離オリゴヌクレオチドの調製を表す。(A)はマイクロチップ表面からのPCR増幅可能なオリゴヌクレオチドの合成と切断の図を表す。遺伝子構築に使われたオリゴヌクレオチドの部分は黒色で描かれている。PCRプライマーアダプターは灰色で示されている。(B)はXeotron/Atactic 4K光プログラマブル微小流体マイクロチップからのオリゴヌクレオチドの合成と切断を表す。左側：切断前のオリゴヌクレオチドアレイの蛍光スキャン顕微鏡写真。挿入部分：微小流体チャンバおよび接続路(チャンネル)の細部。右側：切断後のアレイ。(C)は品質評価(QA)-チップとの遊離フルオレセイン(FAM)-標識オリゴヌクレオチドのハイブリダイゼーションを表す。左側：ハイブリダイゼーション前；中央：ハイブリダイゼーション後；右側；ハイブリダイズしたヌクレオチドのストリッピング後。

【図2】新規のRS3 対 本来の大腸菌K12のアミノ酸配列を表す。2AをSEQ ID NO:1として示し；2B上段をSEQ ID NO:2として示し；2B中段をSEQ ID NO:3として示し；2B下段をSEQ ID NO:4として示す。

【図3】新規のRS3 対 本来の大腸菌K12の核酸配列を表す。スコア(Score) = 212ビット(107)、期待値(Expect) = 6e-52、同一残基率(Identities) = 557/707 (78%)、ギャップ(Gaps) = 5/707 (0%)。上側の配列をSEQ ID NO:5として示し；下側の配列をSEQ ID NO:6として示す。

【図4】21個の合成rs遺伝子T7発現構築体を示すアガロースゲルを表す。

【図5】マイクロチップ合成されたオリゴヌクレオチドのハイブリダイゼーション選択に関するハイブリダイゼーションストラテジーの図を表す。90merのオリゴヌクレオチド(上側鎖黒色、下側鎖灰色)をIIS型制限酵素で切断して、一部のものは正しくない配列(2番目の90merのオリゴヌクレオチドの上側鎖中バルジで示される)を有する50merと相補的44merとのハイブリッドを放出させる。正しい上側50merの鎖だけが左側(L)その後右側(R)の選択用オリゴヌクレオチド(灰色のビーズ上に固定化された)とうまくハイブリダイズする。

【図6】プール中の複数の遺伝子の設計、合成および分析に関するフローチャートを表す

。現在の作業タイミングの推定値(必ずしも可能な最小時間とは限らない)が記載されている。

【図7】本発明のある種の態様によってオリゴヌクレオチドを設計するためのプログラムの操作を示すフローチャートを表す。

【図8】図7のプログラムに対する典型的な入力配列ファイルを表す。Rs1をSEQ ID NO:7として示し; rs2をSEQ ID NO:8として示す。

【図9】図7のプログラムに対する典型的なパラメータ入力ファイルを表す。

【図10】図7のプログラムに対する典型的なコドン使用頻度表を表す。

【図11】本発明のある種の態様による入力配列の最適化を示すフローチャートを表す。

【図12】制限酵素切断後の図8由来の配列の1つを表す。Rs1-f1をSEQ ID NO:9として示し; rs1-f2をSEQ ID NO:10として示し; rs1-f3をSEQ ID NO:11として示し; rs1-f4をSEQ ID NO:12として示し; rs2-f1をSEQ ID NO:13として示す。

【図13】本発明のある種の態様による融解点( $T_m$ )に基づくオリゴヌクレオチド断片の選択を示すフローチャートを表す。

【図14】図13A~13Bの選択アルゴリズムを示した図を表す。配列をSEQ ID NO:9として示す。

【図15】図13A~13Bの選択アルゴリズムを示した図を表す。配列をSEQ ID NO:14として示す。

【図16】図13A~13Bの選択アルゴリズムを示した図を表す。配列をSEQ ID NO:14として示す。

【図17】図13A~13Bの選択アルゴリズムを示した図を表す。配列をSEQ ID NO:15として示す。

【図18】図13A~13Bのアルゴリズムに対するデータ出力の例を表す。Rs1-f1-1をSEQ ID NO:16として示し; rs1-f1-1LをSEQ ID NO:17として示し; rs1-f1-1RをSEQ ID NO:18として示し; rs1-f1-38をSEQ ID NO:19として示し; rs1-f1-38LをSEQ ID NO:20として示し; rs1-f1-38RをSEQ ID NO:21として示し; rs1-f1-LをSEQ ID NO:22として示し; rs1-f1-RをSEQ ID NO:23として示し; 左側プライマーをSEQ ID NO:24として示し; 右側プライマーをSEQ ID NO:25として示す。

【図19】本発明のある種の態様による長さに基づくオリゴヌクレオチド断片の選択を示すフローチャートを表す。

【図20】図19の選択アルゴリズムを示した図を表す。配列をSEQ ID NO:14として示す。

【図21】図19の選択アルゴリズムを示した図を表す。配列をSEQ ID NO:26として示す。

【図22】図19の選択アルゴリズムを示した図を表す。配列をSEQ ID NO:27として示す。

【図23】図19のアルゴリズムに対するデータ出力の例を表す。Rs1-f1-1をSEQ ID NO:28として示し; rs1-f1-1LをSEQ ID NO:29として示し; rs1-f1-1RをSEQ ID NO:30として示し; rs1-f1-23をSEQ ID NO:31として示し; rs1-f1-23LをSEQ ID NO:32として示し; rs1-f1-23RをSEQ ID NO:33として示し; rs1-f1-LをSEQ ID NO:22として示し; rs1-f1-RをSEQ ID NO:23として示し; 左側プライマーをSEQ ID NO:24として示し; 右側プライマーをSEQ ID NO:28として示す。

【図24】本発明のある種の態様によってどのように構築用オリゴヌクレオチドを設計するかを図式的に表す。Rs1-f1-1をSEQ ID NO:16として示し; rs1-f1-1LをSEQ ID NO:17として示し; rs1-f1-1RをSEQ ID NO:18として示し; rs1-f1-1cをSEQ ID NO:38として示し; センス5末端付加(sense5endAddOn)をSEQ ID NO:39として示し; センス3末端付加(sense3endAddOn)をSEQ ID NO:40として示す。

【図25】本発明のある種の態様によってどのように選択用オリゴヌクレオチドを設計するかを図式的に表す。配列(1)をSEQ ID NO:38として示し; 配列(2)をSEQ ID NO:37として示し; 配列(3)をSEQ ID NO:41として示し; 配列(4)をSEQ ID NO:42として示し; 配列(5)をSEQ ID NO:43として示し; 配列(6)をSEQ ID NO:36として示し; 配列(7)をSEQ ID NO:44として示し; 配列(8)をSEQ ID NO:45として示し; 配列(9)をSEQ ID NO:46として示す。

【図26】異なるプールサイズパラメータを指定する場合の典型的なプログラム出力を表

10

20

30

40

50

す。Rs1-f1-1をSEQ ID NO:35として示し；rs1-f1-1LをSEQ ID NO:36として示し；rs1-a1-1RをSEQ ID NO:37として示し；プール-1左側プライマーをSEQ ID NO:47として示し；プール-1右側プライマーをSEQ ID NO:23として示し；プール-2左側プライマーをSEQ ID NO:49として示し；プール-2右側プライマーをSEQ ID NO:50として示し；プール-3左側プライマーをSEQ ID NO:51として示し；プール-3右側プライマーをSEQ ID NO:52として示し；プール-4左側プライマーをSEQ ID NO:53として示し；プール-4右側プライマーをSEQ ID NO:54として示し；プール-5左側プライマーをSEQ ID NO:55として示し；プール-5右側プライマーをSEQ ID NO:56として示し；プール-6左側プライマーをSEQ ID NO:57として示し；プール-6右側プライマーをSEQ ID NO:58として示し；プール-7左側プライマーをSEQ ID NO:59として示し；プール-7右側プライマーをSEQ ID NO:60として示し；プール-8左側プライマーをSEQ ID NO:24として示し；プール-8右側プライマーをSEQ ID NO:48として示す。

10

【図27】異なるチップExtraSeqLenパラメータを指定する場合の典型的なプログラム出力を表す。Rs1-f1-1をSEQ ID NO:35として示し；rs1-f1-1LをSEQ ID NO:36として示し；rs1-f1-1RをSEQ ID NO:37として示し；rs1-f1-38をSEQ ID NO:61として示し；rs1-f1-38LをSEQ ID NO:62として示し；rs1-f1-38RをSEQ ID NO:21として示し；rs1-f1-LをSEQ ID NO:22として示し；rs1-f1-RをSEQ ID NO:23として示し；左側プライマーをSEQ ID NO:24として示し；右側プライマーをSEQ ID NO:25として示す。

【図28】ポリヌクレオチド忠実度に対するエラー率の影響を表す。

【図29】オリゴヌクレオチドの設計から所定の配列を有する複数のポリヌクレオチド構築体の産生までの、複数のポリヌクレオチド構築体の多重アッセムブリ方法の1つの態様の図式的概観を表す。

20

【図30】(A) ライゲーション、(B) 鎖伸長ならびに(C) 鎖伸長およびライゲーションを含む、サブアッセムブリおよび/またはポリヌクレオチド構築体への構築用オリゴヌクレオチドの3通りの典型的なアッセムブリ方法の図式的概観を表す。破線は、ポリメラーゼによって伸長された鎖を示す。

【図31】複数ラウンドのアッセムブリを含むポリヌクレオチドアッセムブリ方法の1つの態様の図式的概観を表す。

【図32】オリゴヌクレオチドプールを増幅するためのユニバーサルプライマーを利用するポリヌクレオチドアッセムブリ方法の1つの態様の図式的概観を表す。

【図33】構築用オリゴヌクレオチドのプールを増幅するためのユニバーサルプライマーの1セットとサブアッセムブリ(例えば、abc)を増幅するためのユニバーサルプライマーの1セットとを利用するポリヌクレオチドアッセムブリ方法の1つの態様を示す図式的概観を表す。

30

【図34】ミスマッチ結合タンパク質を用いたエラー配列の除去方法の1つを表す。

【図35】ミスマッチ認識タンパク質を用いたエラー配列の中和を表す。

【図36】鎖特異的なエラー補正方法の1つを表す。

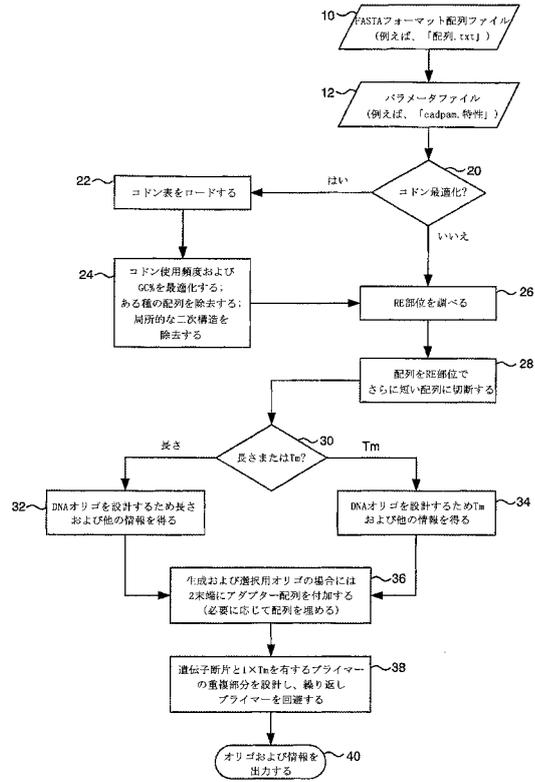
【図37】オリゴヌクレオチドプールをエラー低減の前に変性/再生のラウンドに供することでエラー低減過程の効率を増加させる方法の1つを示す図式的概観を表す。Xは配列エラー(例えば、挿入、欠失または正しくない塩基の形での所望の配列からの逸脱)を表す。

【図38】各種の方法によって作製された配列エラーの比較を表す。<sup>2</sup>試験をハイブリダイゼーション選択 対 PAGE選択( $P = 2 \times 10^{-5}$ )、およびハイブリダイゼーション選択 対 選択なし( $P = 2 \times 10^{-21}$ )に対し行った。「PAGE選択」と表示された横列の構築体のみにゲル精製が含まれた。

40



【 図 7 】



【 図 8 】

配列.txt

```

>rs1
atgaactgaatttttgcacacttttgaagagctctfaanaaganaagaacccggggtctctctgcttggggtt
ggttgcctatcgcaanaagactgtagctggtgacgctgctgaatctgagctccatccggctgagcagctca
aanaaccgacggcgagcfcggaatccagfaggggaaggtgagctgctctggacgagfagaagcggctt
cggfnaacfcctgctcccgfagaagctnaacgctcaagagcctggatcagctggaanaagcttcaagaatg
ctgaaactgttaccggttttaacagcnaagfnaagggcgcttcactggtgagctgacggfcttctgctctc
ccaggttctctgtagcgtctccggctgctgacactctgacccggagcaagagctgaaatfnaagfnaatca
agctgatacaagagcnaacacggttcttctctgctgctgctgacaaaccgaaacagcggagcggatc
agctcgganaacccggaaggaagctggaagfnaaggtatcgttaagaacctcagctacggtgcaattctgtg
atctggggcggttgcagcgctgctgacatcaactgacatggcctggaaacagcttaagctcagagcaaatctgca
acgtggcgagcaaatcactgfaaagctgaggttcaacccgcaacctccgctatccggcgtgaaacagc
fggggaagatcctgggtagctatgctaaactgctcggaaaggaacnaacfcgctgctgctgcaacactgca
cagactagctgcttctggaatcgaagagggctggaagcctggfagcaactctccgnaatggactggaccaca
naaacatccacccgtccaaaggttgaactgtggagtagtggaaggtatggtctggatgacgaagaagcctg
ctgtatctccctggctgaaacagctgcaaacctaacccgtggcagcagctggcgaacccaacaagggcgacc
ggtgtagggfanaatcaagctatcaactgctgctgctgctgctgctgctgctgctgctgctgctgctgctg
ctgctgcaatcctggaaactgtaggggaagagcaggttctgtaatacnaaaaggcgaanaatcctgctgctg
gttctgcaaggtgagcgaacgtgaactatcctccggcgtgaaacagctcagagaagctcctcaacaactg
gggtctgcaagaagaagggcctctgtaacggtaaacctgctgctgctgctgctgctgctgctgctgctgctg
ctgctgcaagctgcaagaagatgcaatcgaactgtaacaaacaggaagatgcaaaccttcccaacaacgcaatg
gctgaagcttcaagcagfanaagggcgagtaa
  
```

```

>rs2
atggcaactgttccatgcgcgaatgctcaagcctgggttcaactcggctacacagccctgactgaaacccaana
tgaagcctgctatctctggcgttaacaaggttcaatcaactgagaaactgtaaccgatgctcaacgaagc
ctctggtaactgaaacagatgctctgcaaggtgaaactcttctggttaaacgcgctcaaggaagcgg
ggaanaagcctgctgctgagcggcagcaggttctctggaacactcctgctggcctgctgctgctgctgctg
accgttctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg
aanaagcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg
ctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg
atgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg
gacctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg
ctctgtagcctgagtaa
  
```

【 図 9 A 】

cadpam.特性

```

# CAD-PAMプログラムの特性ファイル #
#----- 配列修飾 -----#
102 { # 配列最適化: オン/オフ
      # 最適化する=オフ
      #
      # コドンファイル名(デフォルト=ecoli-k12)を付与する
      # コドンファイル=
      #
      # 配列を除去する(最適化する=オンの場合に作動する)
      # 配列=accctgc; gcagggt; agcttc; gagaccを除去する
      #
      # %GC調整 (推奨値: 0.08, 0.1, 0.12)
      # (注: 値が高くなるほど、GC含量が低くなる)
      # GCトレードオフ値(GCTradeOffValue)=0.12
      #----- オリゴ設計 -----#
110 { # 「長さ」または「tm」によってオリゴ配列を設計する?
      # 次のもので配列を選択する(pickSequenceBy)=tm
      #
      # 長さによる場合は、オリゴ長(デフォルト=40 nt)を指定する
      # chipSeqLen=40
      #
      # 現在4塩基対長に設定されている
      # chipExtraSeqLen=0
      # フィルアップを終わらせる(endFillUp)=はい
      #
      # オリゴ重複領域のTm (デフォルト=66C)
      # オリゴTm=50
      # Tm算出の場合 (ほとんどの条件下でデフォルトを使用する)
      # DNA濃度=
      # 塩濃度=
  
```

【 図 9 B 】

cadpam.特性

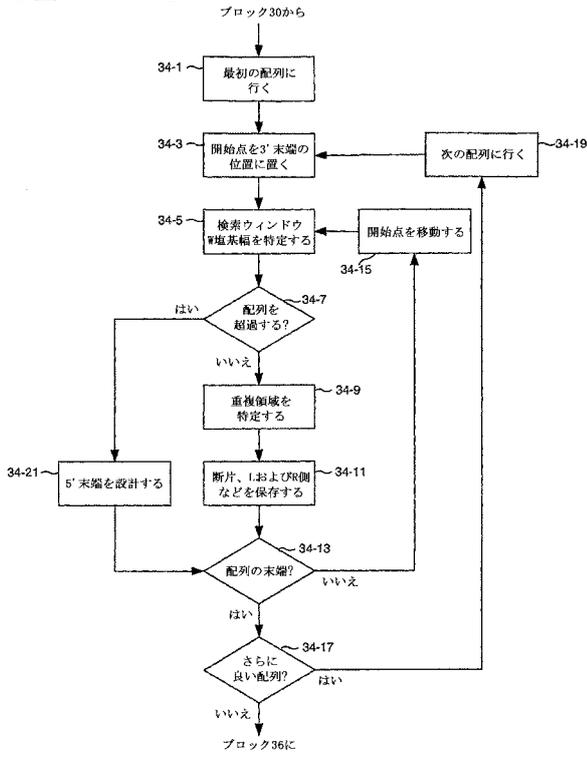
(対照)

```

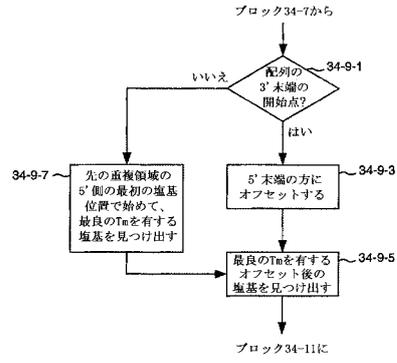
#----- オリゴ・チップ合成 -----#
#
# 生成チップ用オリゴ
120 { センス 5endAadOn=ACCTGcgttc
      # センス 3endAadOn=gGAGACCcta
      # 選択 5endAadOn=cggcggtGCTCTCa
      # 選択 3endAadOn=gctcgGAGCCTgag
      #
      # 選択用チップオリゴ
      # 選択FillUpLen=30
      #
      # 選択用オリゴTm (デフォルト=60C)
      # 選択ChipTm=50
      #----- DNA配列アセンブリ -----#
128 { # 以下の115型R2部位配列において長いDNAを切断する
      # R2部位=センス; アンチセンス
      # R2部位 =accctgc; gcagggt; ggtctc; gagacc
      #
      # アセンブリのオリゴ・プールサイズ(偶数)
      # プールサイズ=50
  
```



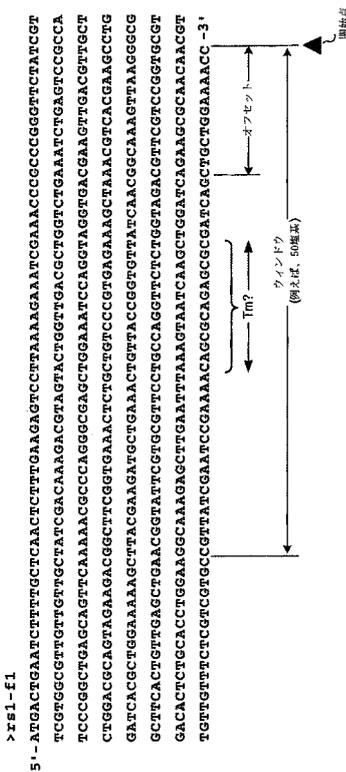
【 図 1 3 A 】



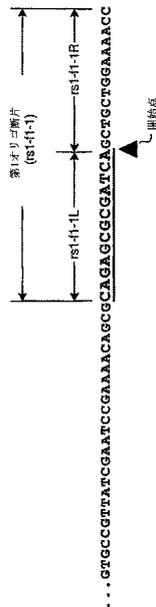
【 図 1 3 B 】



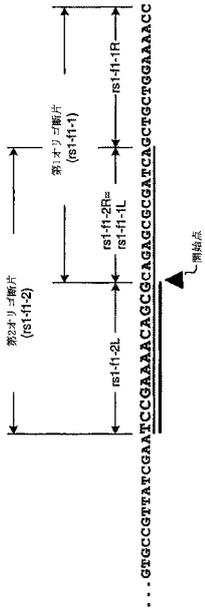
【 図 1 4 】



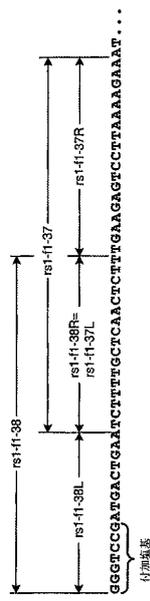
【 図 1 5 】



【 図 16 】



【 図 17 】



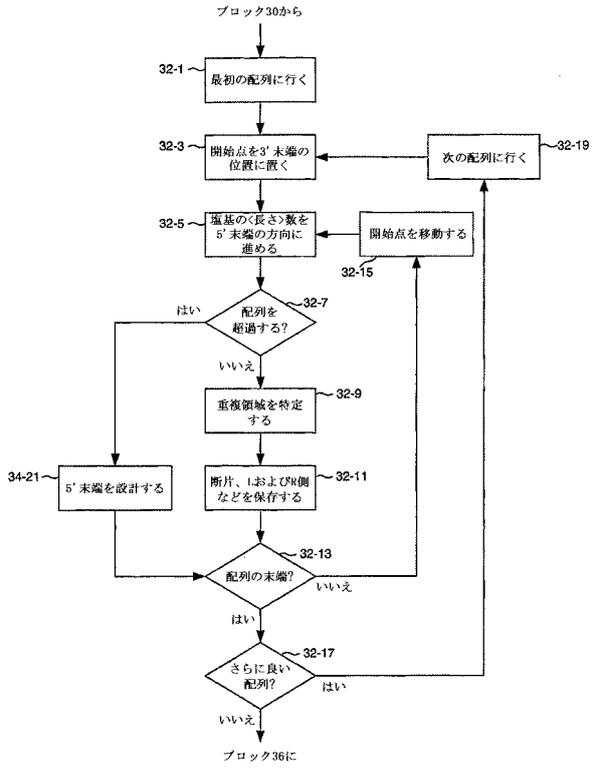
【 図 18 】

info.out

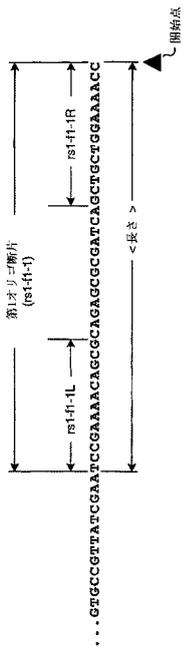
```

>rs1-f1
GC%: 50.61946902654867
rs1-f1-1(538-565),CAGAGCGCGATCAGCTGCTGGAAAACC
rs1-f1-1L(53.437313043640536),CAGAGCGCGATCA
rs1-f1-1R(52.3888909406698),GCTGCTGGAAAACC
rs1-f1-1 余分の5末端 : 余分の3末端 :
.
.
.
rs1-f1-38(0-26),GGTCCGATGACTGAACTCTTTGGCTCAACTCTT
rs1-f1-38L(52.92028903441997),GGTCCGATGACTGAA
rs1-f1-38R(50.028873309323735),GCTTTCCTCAACTCTT
rs1-f1-38 余分の5末端 : 余分の3末端 :
rs1-f1-L(ATGACTGAACTCTTTGC),ATGACTGAACTCTTTGC
rs1-f1-R(GGTTTCCAGCAGC),GGTTTCCAGCAGC
Pool:
pool-1:1-38(0-565)
  左側プライマー : ATGACTGAACTCTTTGC
  右側プライマー : CAGAGCGCGATCAGCTGCTGGAAAACC
  サイズ: 565
  
```

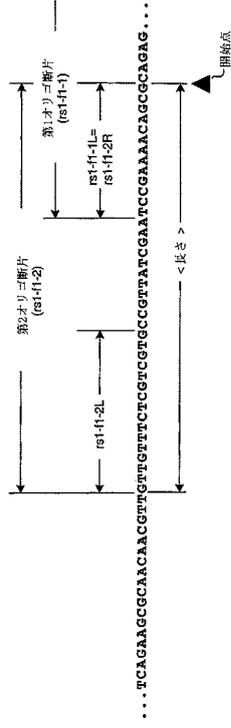
【 図 19 】



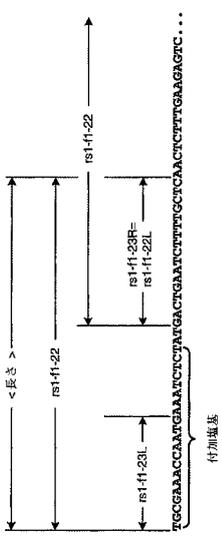
【 図 2 0 】



【 図 2 1 】



【 図 2 2 】



【 図 2 3 】

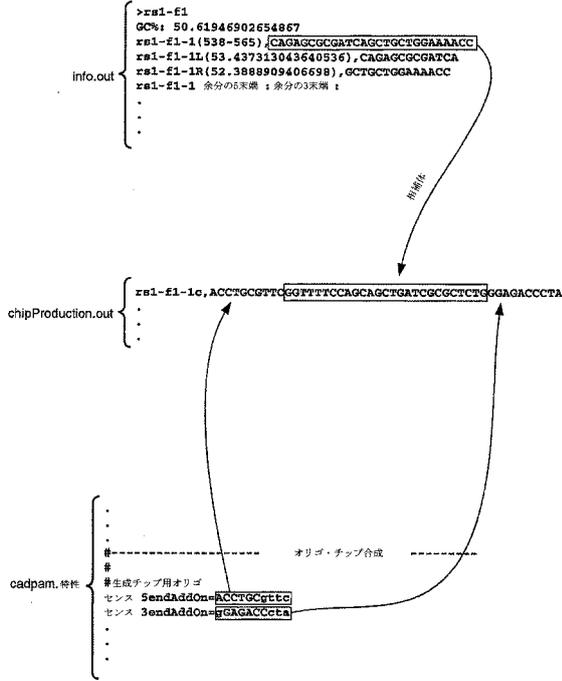
```

info.out

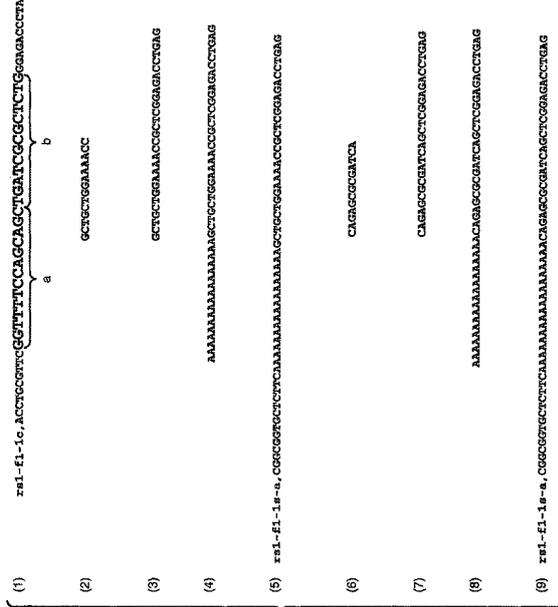
rs1-f1のGC%
>rs1-f1
GC%: 50.61946902654867
rs1-f1-1(565-605), TCCGAAAACAGCCGACAGCCGATCAGCTGCTGGAAAACC
rs1-f1-1L(52.10256306291746), TCCGAAAACAGCG
rs1-f1-1R(52.3888939406698), GCTGCTGGAAAACC
rs1-f1-1 余分の5末端 | 余分の3末端 | rs1-f1-1RのTm
.
.
rs1-f1-23(0-19), TCCGAAACCAATGAAATCTCTATGACTGAAATCTTTTCTC
rs1-f1-23L(49.336517787243515), TCCGAAACCAATG
rs1-f1-23R(49.60228808783075), GACTGAAATCTTTTCTC
rs1-f1-23 余分の5末端 | 余分の3末端 |
rs1-f1-L(ATGACTGAAATCTTTTCTC), ATGACTGAAATCTTTTCTC
rs1-f1-R(GGTTTCCAGCAGC), GGTTTCCAGCAGC
Pool:
pool-i1-1-23(0-605)
左側プライマー : ATGACTGAAATCTTTTCTC
右側プライマー : TCCGAAAACAGCCGACAGCCGATCAGCTGCTGGAAAACC
サイズ: 605

```

【 2 4 】



【 2 5 】



【 2 6 】

```

>rs1-f1
GC%: 50.61946902654867
rs1-f1-1(538-565), CAGAGCCGATCAGCTGCTGAAAACC
rs1-f1-1L(53.437313043640536), CAGAGCCGATCA
rs1-f1-1R(52.3888909406698), GCTGCTGAAAACC
rs1-f1-1 余分の5末端 ; 余分の3末端 ;
.
.
.
Pools:
pool-1:1-5(485-565)
  左側プライマー : GCGCAACACGTTGTTGTTTTCTGTCG
  右側プライマー : GGTTTTCCAGCAGC
  サイズ: 80
pool-2:16-10(410-497)
  左側プライマー : CGTTCGTCGGTCCGTGACACTCTGCA
  右側プライマー : CAAGCTGGATCAGAAGCCCAACACGT
  サイズ: 87
pool-3:11-15(338-423)
  左側プライマー : CGCCAAAGTAAAGGCGCTTCACCTG
  右側プライマー : CCAGGTTCTCTGATAGACGTTGCTCCGCTG
  サイズ: 85
pool-4:16-20(258-352)
  左側プライマー : GAGAAAGCTAAACGTCAGAAAGCTGGATCA
  右側プライマー : TTACCGGTGTTATCAACGGCAAGGTTAAGG
  サイズ: 94
pool-5:21-25(185-275)
  左側プライマー : CCAGGTAGGTGACGAAAGTTGACGTTGCTCT
  右側プライマー : TCTGCTGTCCTCCGTTGAGAAAGCTAAACGTC
  サイズ: 90
pool-6:26-30(118-199)
  左側プライマー : CTGCTCTGAAATCTGATCCCGCAACCC
  右側プライマー : GCGAGCTGAAATCCAGGTAAGGTGACG
  サイズ: 81
pool-7:31-35(45-135)
  左側プライマー : ATCGAAACCCCGCGGTTCTATCGTT
  右側プライマー : CGTAGTACTGGTTGACGCTGCTGTAATCTGAG
  サイズ: 90
pool-8:36-38(0-57)
  左側プライマー : ATGACTGAATCTTTTC
  右側プライマー : TGAAGAGTCTTAAAAGAAATCGAAACCCGC
  サイズ: 57

```

【 2 7 】

```

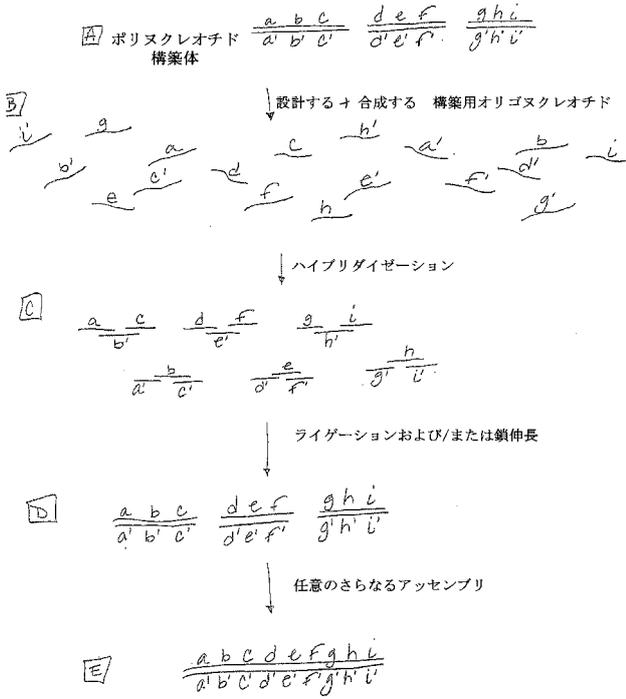
>rs1-f1
GC%: 50.61946902654867
rs1-f1-1(538-565), CAGAGCCGATCAGCTGCTGAAAACC
rs1-f1-1L(53.437313043640536), CAGAGCCGATCA
rs1-f1-1R(52.3888909406698), GCTGCTGAAAACC
rs1-f1-1 余分の5末端 ; AACAGCC 余分の3末端 ;
.
.
.
rs1-f1-38(0-26), CTTTGGAAATATGACTGAATCTTTTGTCTCAACTCTT
rs1-f1-38L(49.50509389465078), CTTTGGAAATATGACTGAA
rs1-f1-38R(50.028873309323735), TCTTTTGTCTCAACTCTT
rs1-f1-38 余分の5末端 ; TAAGGGT 余分の3末端 ; TGAAGAG
rs1-f1-1L(ATGACTGAATCTTTTC), ATGACTGAATCTTTTC
rs1-f1-1R(GGTTTTCCAGCAGC), GGTTTTCCAGCAGC
Pools:
pool-1:1-38(0-565)
  左側プライマー : ATGACTGAATCTTTTC
  右側プライマー : CAGAGCCGATCAGCTGCTGAAAACC
  サイズ: 565

```

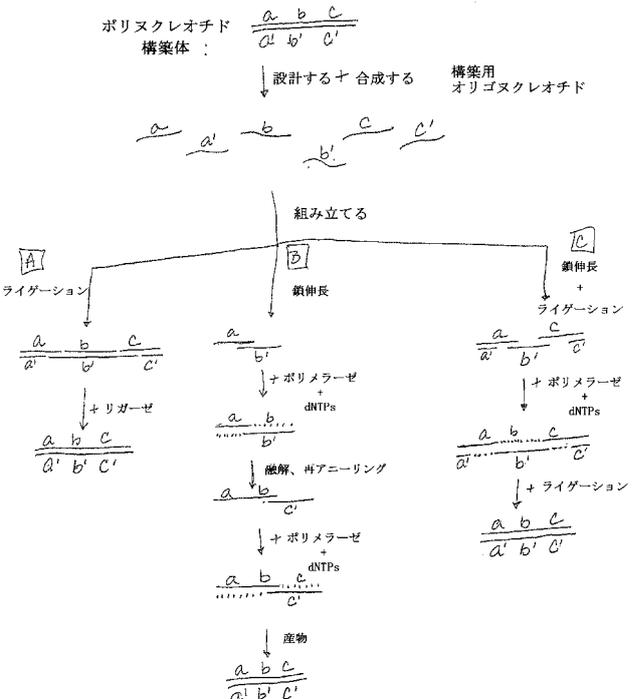
【 2 8 】

ポリヌクレオチド長	塩基エラー率の逆数	正しいコピー分率
1000	20	5.2918E-23
	50	1.683E-09
	100	4.3171E-05
	200	0.00665397
	1000	0.36769542
	10,000	0.90483289
2000	100,000	0.99004978
	200	4.4275E-05
	1000	0.13519993
3000	10,000	0.81872257
	100,000	0.98019858
	200	2.9461E-07
10,000	1000	0.04971239
	10,000	0.74080711
	100,000	0.97044539
10,000	1000	4.5173E-05
	10,000	0.36786105
	100,000	0.90483697

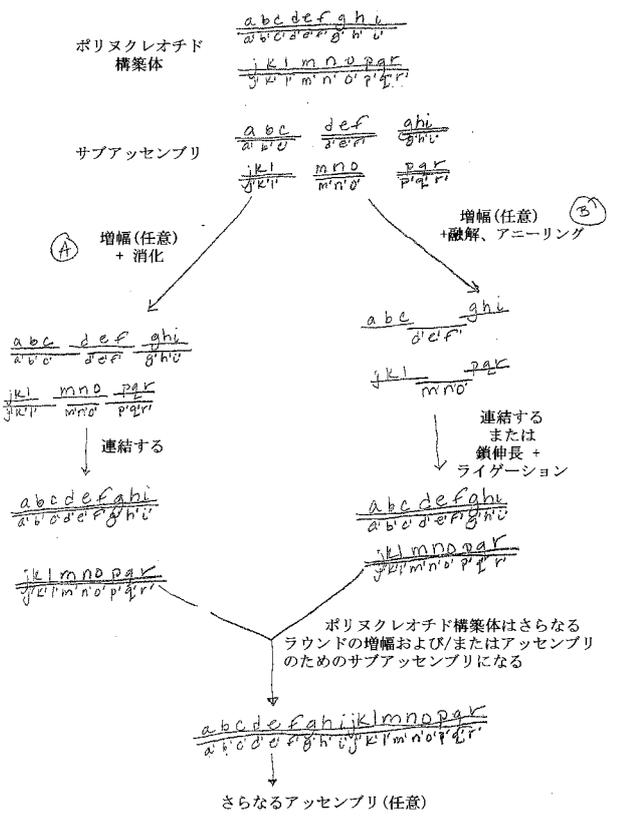
【図 29】



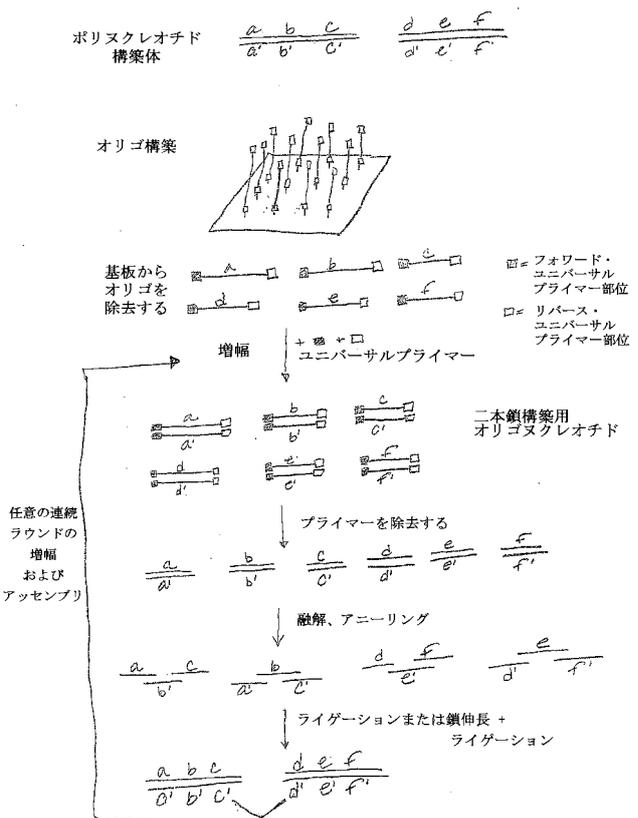
【図 30】



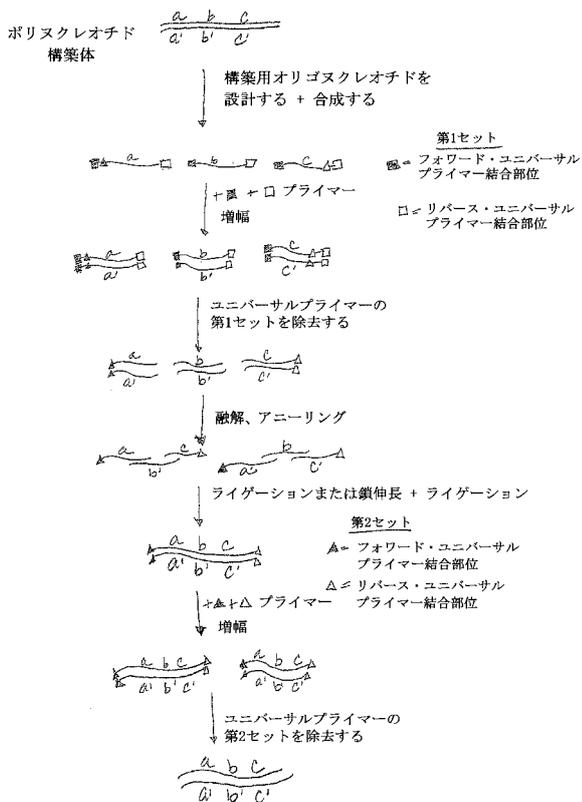
【図 31】



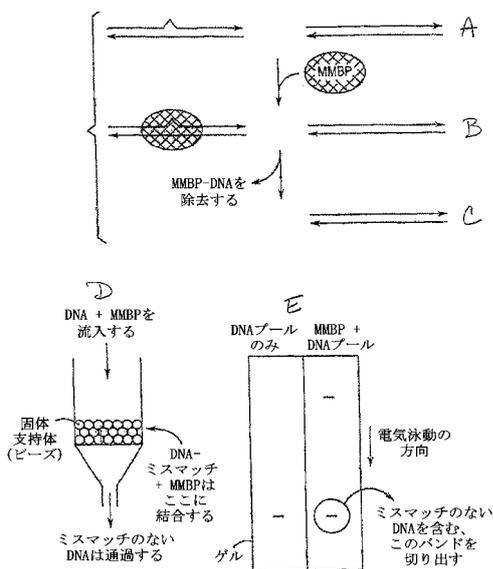
【図 32】



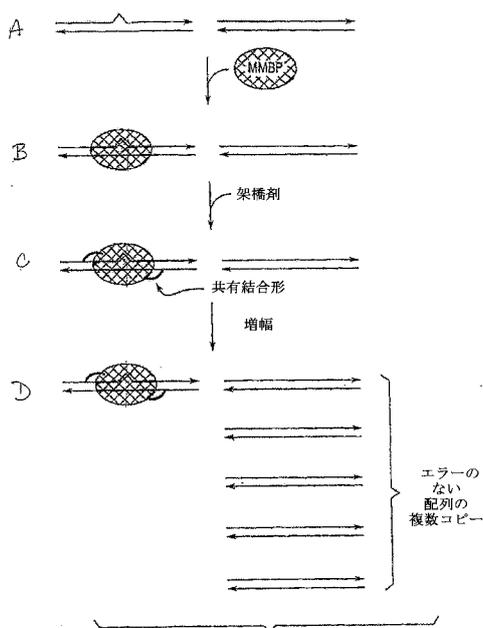
【 図 3 3 】



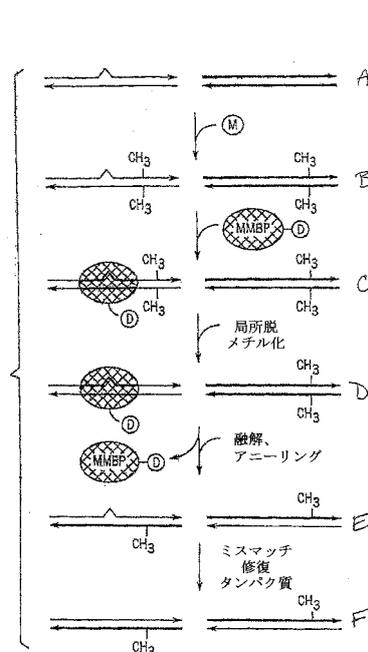
【 図 3 4 】



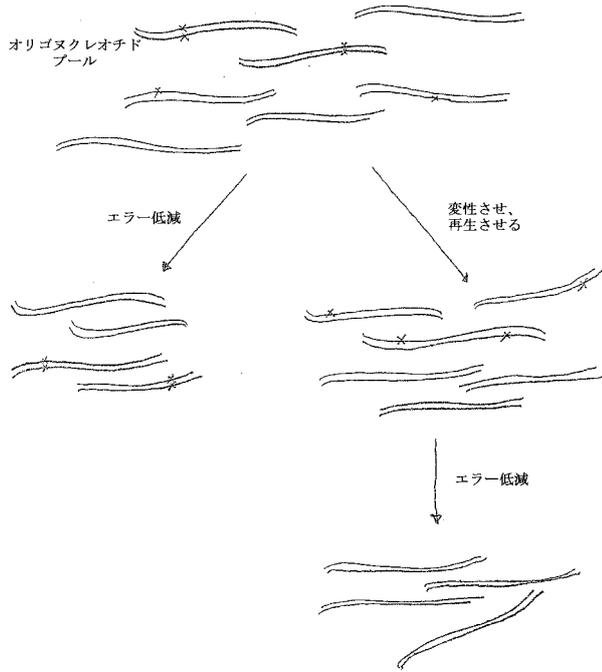
【 図 3 5 】



【 図 3 6 】



【 図 3 7 】



【 図 3 8 】

方法	総BP	転移	塩基転換	欠失	付加	エラー 当たりのBP
ハイブリダイ ゼーション選択 (PAM)	23,641	7	3	5	2	1,394
PAGE選択 (PAM)	24,546	28	12	11	3	455
選択なし (PAM)	9,243	25	13	19	1	159
選択なし (ライゲーション)	6,093	6	6	22	4	160

【 配 列 表 】

2007534320000001.app

## フロントページの続き

(81) 指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(特許庁注：以下のものは登録商標)

1. J A V A
2. W I N D O W S

(72) 発明者 ティアン ジンドン

アメリカ合衆国 マサチューセッツ州 アーリントン ウィグワム サークル 10

Fターム(参考) 4B024 AA11 AA20 CA04 CA09 HA12 HA14 HA19 HA20

4B029 AA23 BB20 CC03

## 【要約の続き】

