

(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2008-512129

(P2008-512129A)

(43) 公表日 平成20年4月24日(2008.4.24)

(51) Int.Cl.	F I	テーマコード (参考)
C12Q 1/68 (2006.01)	C12Q 1/68 ZNAZ	4B024
C12N 15/09 (2006.01)	C12Q 1/68 A	4B063
GO1N 37/00 (2006.01)	C12N 15/00 A	
GO1N 33/53 (2006.01)	GO1N 37/00 1O2	
	GO1N 33/53 M	
審査請求 未請求 予備審査請求 未請求 (全 93 頁)		

(21) 出願番号 特願2007-531428 (P2007-531428)
 (86) (22) 出願日 平成17年9月8日 (2005.9.8)
 (85) 翻訳文提出日 平成19年5月10日 (2007.5.10)
 (86) 国際出願番号 PCT/US2005/032441
 (87) 国際公開番号 W02006/031745
 (87) 国際公開日 平成18年3月23日 (2006.3.23)
 (31) 優先権主張番号 60/608, 712
 (32) 優先日 平成16年9月10日 (2004.9.10)
 (33) 優先権主張国 米国 (US)

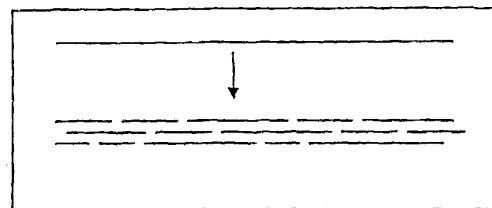
(71) 出願人 505198950
 セクエノム, インコーポレイティド
 アメリカ合衆国, カリフォルニア 921
 21, サンディエゴ, ジョン ホプキンス
 コート 3595
 (74) 代理人 100099759
 弁理士 青木 篤
 (74) 代理人 100077517
 弁理士 石田 敬
 (74) 代理人 100087871
 弁理士 福本 積
 (74) 代理人 100087413
 弁理士 古賀 哲次
 (74) 代理人 100134784
 弁理士 中村 和美

最終頁に続く

(54) 【発明の名称】 核酸の広範囲配列分析法

(57) 【要約】

標的核酸を断片化し、断片を捕獲オリゴヌクレオチドのアレイにハイブリダイズさせ、ハイブリダイズした断片の質量を測定し、および質量測定値から標的核酸のヌクレオチド配列を構築することにより、標的核酸を配列決定するための方法が提供される。



【特許請求の範囲】

【請求項1】

標的核酸を配列決定するための方法であって、以下のステップ：

- a) 標的核酸の重複断片を作製し；
 - b) 捕獲オリゴヌクレオチドへの断片のミスマッチハイブリダイゼーションを排除しない条件下で、捕獲オリゴヌクレオチドのアレイに断片を接触させ；
 - c) 各アレイ位置でハイブリダイズされた断片の質量を質量スペクトル法により測定し；および
 - d) 質量測定値から標的核酸のヌクレオチド配列を構築すること
- を含む、前記方法。

10

【請求項2】

標的核酸を配列決定するための方法であって、以下のステップ：

- a) 標的核酸の重複断片を作製し；
 - b) 捕獲オリゴヌクレオチドのアレイに断片を接触させ、ここで、1つまたはそれ以上の捕獲オリゴヌクレオチドは部分的に縮重している；
 - c) 各アレイ位置で捕獲オリゴヌクレオチドにハイブリダイズした断片の質量を質量スペクトル法により測定し；および
 - d) 質量測定値から標的核酸のヌクレオチド配列を構築すること
- を含む、前記方法。

20

【請求項3】

前記構築ステップd)が、以下のステップ：

- ヌクレオチド位置に仮想ヌクレオチドを含有するヌクレオチド配列を仮に構築し；
 - 仮ヌクレオチド配列の断片化を予測し、どの予測された断片が捕獲オリゴヌクレオチドにハイブリダイズするかを予測し、ハイブリダイズした予測断片の質量を予測し；
 - 断片の予測された質量を観察された質量と比較し；および
 - 予測した質量が観察された質量に一致する場合は、標的核酸分子中のヌクレオチド位置を仮想ヌクレオチドとして同定すること
- を含む、請求項1または2記載の方法。

【請求項4】

- 前記の仮に構築するステップが、ヌクレオチド位置の4つの典型的なヌクレオチドのそれぞれを含有するヌクレオチド配列を仮に構築することをさらに含み、予測し比較するステップがすべての仮のヌクレオチド配列について行われ、予測された質量が観察された質量と最もよく一致する仮のヌクレオチド配列を、標的核酸分子中のヌクレオチド配列として同定することを含む、請求項3記載の方法。

30

【請求項5】

- 前記の仮に構築し、予測し、比較しおよび同定するステップが繰り返され、各繰り返しが、ヌクレオチド位置に仮想ヌクレオチドを含む徐々に長くなるヌクレオチド配列を仮に構築することを含む、請求項3または4記載の方法。

【請求項6】

- 前記構築ステップd)が、以下のステップ：
- 核酸断片化の断片生成物の限界を確立し；
 - 特定の捕獲オリゴヌクレオチドにハイブリダイズできる核酸断片の限界を確立し；
 - 捕獲オリゴヌクレオチドにハイブリダイズしたヌクレオチド断片の質量スペクトルで観察できる可能な質量を予測し；
 - 観察された質量を、観察できる予測された質量と比較して、存在するであろう可能な配列を同定するかおよび/または存在しない配列を同定し；並びに
 - 1つまたはそれ以上の追加の捕獲オリゴヌクレオチドについて比較、確立、予測、および比較ステップを繰り返して、存在し得る可能な配列の数を減少させること
- を含み、

40

それにより標的核酸分子のヌクレオチド配列の少なくとも一部が同定される、請求項1

50

または2記載の方法。

【請求項7】

前記重複断片がランダムに生成される、請求項1~6のいずれか1項記載の方法。

【請求項8】

前記重複断片が非特異的に生成される、請求項1~6のいずれか1項記載の方法。

【請求項9】

前記断片が、酵素的断片化法、物理的断片化法、化学的断片化法、およびこれらの組合せから選択される断片化法を使用して生成される、請求項1~6のいずれか1項記載の方法。

【請求項10】

前記断片が、1つまたはそれ以上の酵素を使用する酵素的断片化法により生成され、酵素的断片化法に使用される1つまたはそれ以上の酵素が、非特異的RNase、非特異的DNase、少なくとも2つの2塩基カッター、選択的切断エンドヌクレアーゼ、制限エンドヌクレアーゼ、1塩基カッター、2塩基カッター、およびこれらの組合せよりなる群から選択される、請求項1~6のいずれか1項記載の方法。

【請求項11】

前記断片が、物理的断片化法により生成され、物理的断片化法が、流体力学的力、攪拌、超音波処理、および噴霧よりなる群から選択される、請求項1~6のいずれか1項記載の方法。

【請求項12】

前記断片が、化学的断片化法により生成され、化学的断片化法が、酸加水分解、塩基加水分解、アルキル化、および放射線照射よりなる群から選択される、請求項1~6のいずれか1項記載の方法。

【請求項13】

前記断片が、統計的に、5~50塩基、10~40塩基、11~35塩基、および12~30塩基からなるサイズ範囲よりなる群から選択される範囲にある、請求項1~12のいずれか1項記載の方法。

【請求項14】

前記断片が、統計的に、20~50塩基、30~60塩基、40~70塩基、および50~80塩基からなるサイズ範囲よりなる群から選択される範囲にある、請求項1~12のいずれか1項記載の方法。

【請求項15】

前記標的核酸が1本鎖である、請求項1~14のいずれか1項記載の方法。

【請求項16】

前記標的核酸が1本鎖RNAである、請求項1~15のいずれか1項記載の方法。

【請求項17】

前記標的核酸が2本鎖である、請求項1~14のいずれか1項記載の方法。

【請求項18】

前記ハイブリダイズステップが、ミスマッチハイブリダイゼーションを排除しない条件下で行われる、請求項2~17のいずれか1項記載の方法。

【請求項19】

前記ハイブリダイズステップが、低緊縮性条件下で行われる、請求項1~18のいずれか1項記載の方法。

【請求項20】

捕獲オリゴヌクレオチド配列のすべての理論的組合せより少ない組合せがアレイ上に存在する、請求項1~19のいずれか1項記載の方法。

【請求項21】

1つまたはそれ以上の捕獲オリゴヌクレオチドが部分的に縮重している、請求項1および3~20のいずれか1項記載の方法。

【請求項22】

10

20

30

40

50

前記捕獲オリゴヌクレオチドのすべてが部分的に縮重している、請求項1~21のいずれか1項記載の方法。

【請求項23】

前記の部分的に縮重したオリゴヌクレオチドが、少なくとも10%、少なくとも20%、少なくとも30%、少なくとも40%、および少なくとも50%よりなる群から選択される割合の縮重位置を含む、請求項2~22のいずれか1項記載の方法。

【請求項24】

前記の部分的に縮重しているオリゴヌクレオチドが、1、2、3、4、5、6、7、8、9、および10よりなる群から選択される数の縮重位置を含む、請求項2~23のいずれか1項記載の方法。

10

【請求項25】

各縮重位置が、ユニバーサル塩基および半ユニバーサル塩基よりなる群から選択される縮重塩基を含む、請求項24記載の方法。

【請求項26】

前記ユニバーサル塩基が、イノシン、キサントシン、3-ニトロピロール、4-ニトロインドール、5-ニトロインドール、6-ニトロインドール、ニトロイミダゾール、4-ニトロピラゾール、5-アミノインドール、4-ニトロベンズイミダゾール、4-アミノベンズイミダゾール、フェニルC-リボヌクレオシド、ベンズイミダゾール、5-フルオロインドール、インドール；非環状糖類似体、ヒポキサンチンの誘導体、イミダゾール4,5-ジカルボキサミド、3-ニトロイミダゾール、5-ニトロイミダゾール；芳香族類似体、ベンゼン、ナフタレン、フェナントレン、ピレン、ピロール、ジフルオロトルエン；イソカルボスチリルヌクレオシド誘導体、MICS、ICS；および水素結合類似体、N8-ピロロピリジンよりなる群から選択される、請求項25記載の方法。

20

【請求項27】

前記半ユニバーサル塩基が、プリンAとGに選択的にハイブリダイズする塩基、ピリミジンCとTに選択的にハイブリダイズする塩基、ピリミジンCとUに選択的にハイブリダイズする塩基、6H,8H-3,4-ジヒドロピリミド[4,5-c][1,2]オキサジン-7-オン、およびN6-メトキシ-2,6-ジアミノプリンよりなる群から選択される、請求項25記載の方法。

【請求項28】

大半の縮重塩基が捕獲オリゴヌクレオチドの3'末端に位置する、請求項25~27のいずれか1項記載の方法。

30

【請求項29】

大半の縮重塩基が捕獲オリゴヌクレオチドの5'末端に位置する、請求項25~27のいずれか1項記載の方法。

【請求項30】

前記アレイが、5,000以下、4096以下、4,000以下、3,000以下、2500以下、2100以下、2000以下、1536以下、1500以下、1400以下、1300以下、1200以下、1100以下、1000以下、900以下、800以下、700以下、600以下、500以下、400以下、384以下、300以下、200以下、100以下、96以下、および64以下よりなる群から選択される数の異なる捕獲オリゴヌクレオチドを含有する、請求項1~29のいずれか1項記載の方法。

40

【請求項31】

前記捕獲オリゴヌクレオチドのアレイが、4096の捕獲オリゴヌクレオチドを含有し、捕獲オリゴヌクレオチドのそれぞれが本質的に12塩基からなる、請求項30記載の方法。

【請求項32】

前記捕獲オリゴヌクレオチドのアレイが、ハイブリダイゼーションチップ、ピンツール、ビーズ、ポリスチレン、ポリカーボネート、ポリプロピレン、ナイロン、ガラス、デキストラン、キチン、砂、軽石、アガロース、多糖、デンドリマー、バッキーボール、ポリアクリルアミド、シリコン、金属、ゴム、マイクロタイターディッシュ、マイクロタイターウェル、ガラススライド、シリコンチップ、ニトロセルロースシート、およびナイロンメッシュよりなる群から選択される固体支持体上に固定化される、請求項1~31のいずれ

50

か1項記載の方法。

【請求項33】

前記捕獲断片のアレイを酵素で処理して、ハイブリダイズした断片の全体の長さを減少させることをさらに含む、請求項1～32のいずれか1項記載の方法。

【請求項34】

前記酵素が、1本鎖特異的RNase、1本鎖特異的DNase、塩基特異的RNase、および塩基特異的DNaseよりなる群から選択される、請求項33記載の方法。

【請求項35】

標的核酸断片の質量スペクトルの複雑さを制御する方法であって、以下のステップ：

(a) 捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸断片の第1の領域中の異なるヌクレオチド配列の数を調節し、これにより各第1の領域中の異なるヌクレオチド配列を含有する2つまたはそれ以上の標的核酸断片は、捕獲オリゴヌクレオチドプローブにハイブリダイズし；および

(b) 捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸断片の質量を質量スペクトル法により測定することを含み、

これにより質量スペクトルの複雑さが制御される、前記方法。

【請求項36】

標的核酸断片の質量を測定する前に標的核酸断片の長さを制御するステップをさらに含む、請求項35記載の方法。

【請求項37】

前記捕獲オリゴヌクレオチドプローブが1つまたはそれ以上の縮重塩基を含有する、請求項35～36のいずれか1項記載の方法。

【請求項38】

前記縮重塩基が、ユニバーサル塩基および半ユニバーサル塩基よりなる群から選択される、請求項37の記載の方法。

【請求項39】

1つまたはそれ以上の標的核酸断片が、捕獲オリゴヌクレオチドプローブにハイブリダイズしない第2の領域をさらに含む、請求項35～38のいずれか1項記載の方法。

【請求項40】

第2の領域を含有する1つまたはそれ以上の標的核酸断片のうちで、少なくとも2つは各第2の領域中に異なるヌクレオチド配列を含有する、請求項39記載の方法。

【請求項41】

前記標的核酸断片が、中程度の緊縮性ハイブリダイゼーション条件と低緊縮性ハイブリダイゼーション条件よりなる群から選択されるハイブリダイゼーション条件下で捕獲オリゴヌクレオチドプローブにハイブリダイズする、請求項35～40のいずれか1項記載の方法。

【請求項42】

1つまたはそれ以上の標的核酸断片の第1の領域が、3'末端および5'末端よりなる群から選択される標的核酸断片の末端を含む、請求項35～41のいずれか1項記載の方法。

【請求項43】

1つまたはそれ以上の標的核酸断片の第2の領域が、3'末端および5'末端よりなる群から選択される標的核酸断片の末端のヌクレオチド位置に1つまたはそれ以上の既知のヌクレオチドを含む、請求項39～42のいずれか1項記載の方法。

【請求項44】

標的核酸断片の長さを制御するステップが、塩基特異的切断をさらに含む、請求項35～43のいずれか1項記載の方法。

【請求項45】

前記標的核酸断片が、捕獲オリゴヌクレオチドプローブのアレイにハイブリダイズし、ここで当該アレイは複数の位置を含み、各アレイ位置の捕獲オリゴヌクレオチドプローブ

10

20

30

40

50

のヌクレオチド配列は、すべての他のアレイ位置で捕獲オリゴヌクレオチドプローブのヌクレオチド配列とは異なる、請求項35～44のいずれか1項記載の方法。

【請求項46】

標的核酸の部分を同定する方法であって、以下のステップ：

(a) 請求項35～45のいずれか1項記載の方法に従って制御された複雑さを有する質量スペクトルを採取し；および

(b) 1つまたはそれ以上の標的核酸断片の質量を、1つまたはそれ以上の参照核酸の1つまたはそれ以上の質量と比較すること
を含み、

ここで、1つまたはそれ以上の標的核酸断片の質量と1つまたはそれ以上の参照質量との相関は、参照核酸に対応するとしてまたは参照核酸の部分に対応するとして標的核酸の部分を同定する、前記方法。

【請求項47】

少なくとも1つの参照核酸の1つまたはそれ以上の参照質量が計算される、請求項46記載の方法。

【請求項48】

少なくとも1つの参照核酸の1つまたはそれ以上の参照質量が実験的に測定される、請求項46～47のいずれか1項記載の方法。

【請求項49】

前記標的核酸断片が、配列特異的断片化および非特異的断片化から選択される方法を用いて生成される、請求項46～48のいずれか1項記載の方法。

【請求項50】

前記同定された標的核酸の部分がSNPを含む、請求項46～49のいずれか1項記載の方法。

【請求項51】

(a) 固体支持体上の2つまたはそれ以上の捕獲オリゴヌクレオチドのアレイ、ここで少なくとも1つの捕獲オリゴヌクレオチドは部分的に縮重している；および

(b) アレイに機能できる形で結合した質量スペクトル計を含む、標的核酸の部分を同定するための組合せ。

【請求項52】

捕獲オリゴヌクレオチドにハイブリダイズする核酸分子から得られる質量シグナルのセットから、標的核酸のヌクレオチド配列を構築するためのコンピュータプログラムをさらに含む、請求項51記載の組合せ。

【請求項53】

1つまたはそれ以上の参照質量ピークのセットをさらに含む、請求項52記載の組合せ。

【発明の詳細な説明】

【技術分野】

【0001】

発明の分野

核酸分析法が提供される。

【背景技術】

【0002】

関連出願

本出願は、米国特許出願第10/412,801号、Linら、2003年4月11日出願、標題「固体支持体上で化学反応を行うための方法と装置」に関連する60/608,712号（2004年9月10日出願）；米国仮特許出願第60/457,847号、Linら、2003年3月24日出願、標題「固体支持体上で化学反応を行うための方法と装置」；米国仮特許出願第60/372,711号、Linら、2002年4月11日出願、標題「固体支持体上で化学反応を行うための方法と装置」；米国特許出願第10/723,365号、van den Boomら、2003年11月27日出願、標題「配列変化の検出と発見のための断片化に基づく方法とシステム」；米国仮特許出願第60/429,895号、van den Boomら、2002年11月27日出願、標題「配列変化の検出と発見のための断片化に基づく方法とシステ

10

20

30

40

50

ム」；米国特許出願第10/830,943号、Boeckerら、2004年4月22日出願、標題「新規配列決定のための断片化に基づく方法とシステム」；および米国仮特許出願第60/466,006号、Boeckerら、2003年4月25日出願、標題「新規配列決定のための断片化に基づく方法とシステム」の利益を請求する。これらの非仮出願と仮出願の主題と内容は、参照することによりその全体が本明細書に組み込まれる。

【0003】

種々のバイオポリマーの構造分析は、医学と研究において非常に重要な領域である。分子遺伝学は、DNAまたはRNA分子のヌクレオチド配列の知識に依存している。タンパク質のアミノ酸配列は、タンパク質の機能と制御を研究するのに有用な情報を提供する。バイオポリマーの配列分析には種々の方策が存在する。核酸配列を決定するのに最も一般的に使用される方法であるジデオキシ法では、4つの塩基のそれぞれで停止するDNA分子の4セットのサブ配列を作製し、ポリアクリルアミドゲル電気泳動（PAGE）により断片を分離し、得られたバンドを読んで配列を決定する。ゲル電気泳動は、時間がかかり、エラーを引き起こしやすい。

【0004】

ゲル電気泳動による配列決定法の欠点を克服するために提唱されている方法はハイブリダイゼーションによる配列決定という方法であり、例えばBainsとSmith, *J. Theoret. Biol.* 135:303-307 (1998)；Lysovら、*Dokl. Acad. Sci. USSR* 303:1508-1511 (1988)；Drmanacら、*Genomics* 4:114-128 (1989)；Pevzner, *J. Biomolec. Struct. Dynamics* 7(1):63-73 (1989)；PevznerとLipschutz, *Nineteenth Symp. on Math. Found. of Comp. Sci.*, LNCS-841:143-258 (1994)；Waterman, 「コンピューター生物学への入門（Introduction to Computational Biology）」、チャップマンアンドホール（Chapman and Hall）、ロンドン、1995年を参照されたい。ハイブリダイゼーションによる配列決定（SBH）は、ヌクレオチド（プローブ）の短い配列のアレイ（SBHチップ）が標的DNA配列の（レプリカの）溶液と接触させられるDNA配列決定法である。生化学的方法が、標的配列（配列のスペクトル）に結合するプローブのサブセットを測定し、コンビナトリアル法を使用してそのスペクトルからDNA配列を再構築する。技術のためにSBHチップ上のプローブの数が限定されるため、コンビナトリアルの課題は、ある長さの任意のランダムなDNA列を配列決定できるプローブの最小セットの設計である。

【0005】

SBHの実施は、「古典的」プローブ結合法[すなわちすべての 4^k の k 量体オリゴヌクレオチド（ギャップの無い「中実の」プローブ）を有するチップ]を使用し、ここで記号は公知のDNA塩基{A、C、G、T}であり、 k は方法に依存する整数のパラメータである。「ハイブリダイゼーションによる配列決定の主要な課題は、完全な2本鎖を信頼性高く検出し、ミスマッチ塩基対を含有する2本鎖から区別することである」と言われている（Chechekinら、*J. of Biomolecular Structure & Dynamics* 18(1):83-101 (2000)）。すなわちハイブリダイゼーションによる配列決定は、偽陽性または偽陰性結果を与え、最終的に配列決定法を失敗させるミスマッチ塩基対合を避けかつ最小にしようとする。

【0006】

SBH法は、偽陽性および/または偽陰性の読み値を排除するためにミスマッチハイブリダイゼーションを避けることに依存する。従って、ミスマッチハイブリダイゼーションを可能にする新規核酸配列情報を得るハイブリダイゼーションに基づく方法に対するニーズがある。すなわち本明細書の目的において、ミスマッチハイブリダイゼーションを可能にする新規核酸配列情報を得る方法を提供することが目的である。

【発明の開示】

【0007】

要約

特に本明細書に記載の方法において、ミスマッチハイブリダイゼーションを可能にする新規核酸配列情報を得るための方法が提供される。本明細書において核酸の配列分析法（新規配列決定を含む）であって、標的核酸の重複断片を作製し；ミスマッチハイブリダイ

10

20

30

40

50

ゼーションを排除しない条件下で、固体支持体上の捕獲オリゴヌクレオチドのアレイに断片をハイブリダイズさせて、捕獲断片のアレイを作製し；アレイ中の各位置の捕獲断片の質量を、例えば質量スペクトル分析によりこれらの質量を測定することにより測定し；そして、各アレイ位置から得られた質量シグナルのセットから標的核酸のヌクレオチド配列もしくはヌクレオチド配列セットを構築することを含んでなる方法が提供される。また、核酸を配列決定する方法であって、標的核酸の重複断片を作製し；固体支持体上の捕獲オリゴヌクレオチドのアレイに断片をハイブリダイズさせて捕獲断片のアレイを作製し（ここで、捕獲オリゴヌクレオチドの少なくとも1つのサブセットは部分的に縮重している）；アレイ中の各位置の捕獲断片の質量を、例えば質量スペクトル分析によりこれらの質量を測定することにより測定し；そして、各アレイ位置から得られた質量シグナルのセットから標的核酸のヌクレオチド配列もしくはヌクレオチド配列セットを構築することを含んでなる方法が提供される。ある実施態様において、重複断片はランダムに生成される。

10

【0008】

本明細書に記載の方法を使用して試料から得られる配列情報は、特に、遺伝子型判定およびハプロタイプ型判定、多重遺伝子型判定およびハプロタイプ型判定、核酸混合物分析、広範囲再配列決定、配列変化と変異の広範囲検出、多重配列決定、広範囲メチル化パターン解析、生物同定、病原体同定および型判定がある。

【0009】

すなわち本明細書に記載の方法は、固相ハイブリダイゼーションに基づく方法をハイブリダイズ生成物のアルゴリズムに基づく組成分析と有利に組合せて、質量スペクトルを使用する固相ハイブリダイゼーションに基づく配列分析を大きく高めている。本明細書に記載の方法の1つの利点は、従来の方法と比較して実現できる標的核酸配列の読みとり長さの量と正確度の顕著な向上である。より大きな（広範囲の）配列読みとり長さは、非特異的切断または部分的な特異的切断された、次にオリゴヌクレオチド（その一部またはすべては部分的に縮重していてもよい）を捕獲するために固相に結合される、標的核酸の質量スペクトル分析を使用して達成される。例えば本明細書に記載の方法は、1つの反応/実験で少なくとも250、500、600、700、800、900、1,000、1,500、2,000、3,000、4,000、5,000、6,000、7,000、8,000、9,000、最大10,000またはそれ以上のヌクレオチドを配列決定することができる。これを達成するために、本明細書に記載の方法による分析のために作製される断片は、より大きな標的核酸の配列を提供するために、最終的に整列される。

20

30

【0010】

別の実施態様において本明細書に記載の方法により、複数のより短い長さの標的核酸断片が配列決定または分析される。これらの多重の短い配列セットは、例えば特定の配列の一部が既知である時、再配列決定法において有用である。これらの多重のより短い配列セットはまた、多重遺伝子型判定、ハプロタイプ型判定、SNP、およびメチル化パターン検出法に有用である。

【0011】

断片は、全体的もしくは部分的非特異的切断および/または部分的特異的切断により作製することができ、典型的には分析のために重複断片が得られる。重複断片は、同じ標的生物学的分子配列の別の重複断片が得られるように、単一の非特異的切断および/または相補的もしくは部分的塩基特異的切断を使用して得ることができる。切断手段は、酵素的、化学的、物理的またはこれらの組合せでもよく、典型的には重複断片が作製される。従って重複断片を作製するのに選択される具体的な方法に依存して、かかる重複断片がランダムに作製されることもされないこともある。

40

【0012】

切断および非切断標的配列断片の質量は、当該分野で公知の方法（特に限定されないが、質量スペクトル法やゲル電気泳動がある）を使用して測定することができる。典型的な実施態様において、断片の質量を測定するのにMALDI-TOFが使用される。高処理能質量スペクトル分析を行うためのチップとキットは、シケノムインク（SEQUENOM, INC.）から商品名マスアレイ7（MassARRAY7）で市販されている。本明細書で使用されるチップの別の

50

例は、関連する米国特許出願第60/372,711号（2002年4月11日出願）、60/457,847号（2003年3月24日出願）、および10/412,801号（2003年4月11日出願）（これらは参照することによりその全体が本明細書に組み込まれる）に記載された「h-チップ」である。

【0013】

従ってある実施態様において本明細書に記載の方法は、高処理能固相ハイブリダイゼーションと、固相上で分類される重複切断産物の質量スペクトル検出および同定とを組合せる。本明細書に記載の方法はまた、非特異的断片化または部分的特異的断片化により産生される断片シグナルの同定の正確性と明瞭性を改良し、また1つの標的核酸または標的核酸セット内の配列を再構築するアルゴリズムを使用してこれらのシグナルの分析速度を増加させる。

10

【0014】

詳細な説明

A. 定義

B. 核酸分子の配列決定法

C. 標的核酸分子

1. 供給源

2. 調製

3. 標的核酸分子のサイズと組成

4. 増幅

D. 断片化

1. ポリヌクレオチドの酵素的断片化

a. ポリヌクレオチドのエンドヌクレアーゼ断片化

b. ヌクレアーゼ断片化

c. 核酸酵素断片化

d. 塩基特異的断片化

2. ポリヌクレオチドの物理的断片化

3. ポリヌクレオチドの化学的断片化

4. 断片化の組合せ

5. ハイブリダイゼーション後の断片化

20

E. 捕獲オリゴヌクレオチド

1. 標的核酸断片の複雑さの制御

a. 複雑さを制御する方法

b. 断片の領域

c. 部分的1本鎖捕獲オリゴヌクレオチド

2. 捕獲オリゴヌクレオチドの組成

a. ヌクレオチドのタイプ

i. ユニバーサル塩基

ii. 半ユニバーサル塩基

b. 他の特徴

c. 捕獲オリゴヌクレオチドの作製

30

40

F. 固体支持体とアレイ

G. 特異的または非特異的ハイブリダイゼーション

H. トリミング

I. 標的核酸断片に関する情報

1. 分子質量

a. 質量スペクトル分析

b. 他の測定法

2. 質量ピーク特性

3. 捕獲オリゴヌクレオチドとハイブリダイゼーション条件

4. 断片化条件

50

- J. ヌクレオチド配列構築
- K. 質量パターンによるヌクレオチド配列の同定
- L. 標的核酸の一部の同定
- M. 応用

1. 広範囲再配列決定
2. 変異 / 配列変化の広範囲検出
3. 多重配列決定
4. 広範囲メチル化パターン解析
5. 生物の同定
6. 病原体同定と型判定
7. 分子育種と有向進化
8. マーカーとしての標的核酸断片
9. 感染を示すウイルスまたは細菌核酸配列の存在の検出
10. 抗体プロファイル化
11. 疾患マーカーの同定
12. ハプロタイプ型判定
13. DNA繰り返し配列
14. アレレ変化の検出
15. アレレ頻度の測定
16. エピジェネティクス (Epigenetics)

10

20

【 0 0 1 5 】

A. 定義

特に明記しない場合は本明細書で使用されるすべての技術的および科学的用語は、本発明が属する分野の当業者により一般的に理解されるものと同じ意味を有するものとする。すべての特許、特許出願、公開された出願と刊行物、ジーンバンク (GenBank) 配列、ウェブサイト、および本明細書の全開示内容を通して参照される他の公表された材料は、参照することによりその全体が本明細書に組み込まれる。本明細書の用語で複数の定義がある場合、このセクションのものが優先する。URLまたは他の識別子もしくはアドレスが参照される場合、かかる識別子は変化し、インターネット上の具体的な情報は出入りがあるが、同等の情報は公知であり、インターネットおよび / または適切なデータベースを検索することにより容易にアクセスできるものと理解される。その参照は、かかる情報の利用可能性や公に普及していることの証拠である。

30

【 0 0 1 6 】

本明細書において「アレイ」は、要素 (例えば核酸) の集合を示す。典型的にはアレイは3つまたはそれ以上のメンバーを含有する。アドレス可能なアレイはアレイのメンバーが、例えば固体支持体上の位置により識別できるものである。従ってアレイのメンバーは、固相または他の識別可能物の表面上の別個の識別可能な位置に、例えばタグ (電子的タグおよび化学的タグを含む) を付けるかまたはタグで標識することにより固定化することができる。アレイには、特に限定されないが、単一の固相表面上の要素の集合、例えばチップ上のオリゴヌクレオチドの集合がある。

40

【 0 0 1 7 】

本明細書において「特異的にハイブリダイズする」とは、プローブまたはプライマーのみが非標的配列ではなく標的配列に選択的に、典型的には高緊縮性ハイブリダイゼーション条件下でハイブリダイズすることを意味する。例えば特異的ハイブリダイゼーションには、プローブに100%相補的な標的配列へのプローブのハイブリダイゼーションがある。当業者は、ハイブリダイゼーションに影響を与えるパラメータ (例えば温度、プローブもしくはプライマーの長さ組成、緩衝液組成と塩濃度) は周知しており、標的配列への核酸の特異的ハイブリダイゼーションを実現するためにこれらのパラメータを容易に調整することができる。

【 0 0 1 8 】

50

本明細書においてハイブリダイゼーションの緊縮性は、標的核酸断片への捕獲オリゴヌクレオチドの非特異結合を除去するための洗浄条件を意味する。ハイブリダイゼーション条件の例には以下がある：

- 1) 高緊縮性：0.1 × SSPE、0.1% SDS、65
- 2) 中緊縮性：0.2 × SSPE、0.1% SDS、50
- 3) 低緊縮性：1.0 × SSPE、0.1% SDS、50

【0019】

当業者は、洗浄工程が安定なハイブリッドを選択すること、およびSSPEの成分を周知している（例えば、Sambrook, E.F. Fritsch, T. Maniatis、「分子クローニング：実験室マニュアル（Molecular Cloning, A Laboratory Manual）」、コールドスプリングハーバークラボラトリープレス（Cold Spring Harbor Laboratory Press）（1989）、第3巻、B.13頁を参照、また一般的に使用される実験室溶液を記載する多くのマニュアルを参照されたい）。SSPEは、pH 7.4のリン酸緩衝液0.18 M NaClである。さらに当業者は、ハイブリッドの安定性が T_m により決定され、これはナトリウムイオン濃度と温度の関数（ $T_m = 81.5 - 16.6(\log_{10}[\text{Na}^+]) + 0.41(\%G + C) - 600/l$ ）であり、従ってハイブリッド安定性に重要な洗浄条件中のパラメータは、SSPE（またはSSC）中のナトリウムイオン濃度と温度であることを知っている。特異的ハイブリダイゼーションは典型的には高緊縮性条件下で起きる。別の緩衝液、塩、および温度を使用して、同等の緊縮性が達成できることは理解される。

【0020】

本明細書において「核酸」または「核酸分子」は、デオキシリボ核酸（DNA）またはリボ核酸（RNA）のようなポリヌクレオチドを意味する。この用語はまた、ヌクレオチド類似体、1本鎖（センスまたはアンチセンス）および2本鎖ポリヌクレオチドから作製されたRNAまたはDNAの同等物、誘導體、変種および類似体を含むと理解されている。デオキシリボヌクレオチドには、デオキシアデノシン、デオキシシチジン、デオキシグアノシン、およびデオキシチミジンがある。RNAではウラシル塩基はウリジンである。

【0021】

本明細書において「質量スペクトル法」は、当業者に公知の任意の適当な質量スペクトルフォーマットを包含する。かかるフォーマットには、特に限定されないが、マトリックス支援レーザー脱離/イオン化、飛行時間型（MALDI-TOF）、電子噴霧（ES）、IR-MALDI（例えば、公開国際PCT出願第99/57318号、および米国特許第5,118,937号明細書）、直交型-TOF（O-TOF）、軸方向-TOF（A-TOF）、リニア/反射（RETOF）、イオンサイクロトロン共鳴（ICR）、フーリエ変換、およびこれらの組合せがある。MALDI、特にUVとIRは当該分野で公知のフォーマットである。また本明細書に記載の方法で使用するのに適した質量スペクトル法の例の総説については、AebersoldとMann、2003年3月13日、Nature, 422:198-207（例えば図2）を参照されたい（これは参照することによりその全体が本明細書に組み込まれる）。MALDI法は、典型的にはUV-MALDIまたはIR-MALDIを含む。

【0022】

本明細書において用語「質量スペクトル分析」は、原子、分子、および分子断片の電荷対質量比の測定を意味する。

【0023】

本明細書において質量スペクトルは、質量スペクトル法により生体高分子またはその断片を分析して得られたデータの、グラフまたは暗号化された数値または他の方法による提示を意味する。

【0024】

本明細書において質量スペクトルまたは質量スペクトル分析についてパターンは、シグナル、ピークの特徴的な分布および数、またはこれらのデジタル表示を意味する。

【0025】

本明細書において質量スペクトルとその分析においてシグナル、ピーク、または測定値は出力データであり、これは、原子、分子、または分子の断片の電荷対質量比を反映する

10

20

30

40

50

ことができ、また存在する原子、分子、または分子の断片の量を反映することができる。電荷対質量比は、原子、分子、または分子の断片の質量を測定するのに使用することができる。量は定量的または半定量的方法で使用することができる。例えばある実施態様においてシグナル、ピーク、または測定値は、特定の電荷対質量比を有する分子の数または相対数を反映することができる。シグナルまたはピークは、出力データの視覚的表示、グラフ表示、およびデジタル表示を含む。

【0026】

本明細書において測定された質量について強度とは、他の試料または組成物成分と比較して試料または組成物中に存在するアナライトの相対量の反映を意味する。例えば第1の質量スペクトルピークまたはシグナルの強度は、第2の質量スペクトルピークに対して報告するか、またはピークのすべての強度の合計に対して報告することができる。当業者は、ピークの相対強度を報告するための種々の方法を認識できるであろう。強度は、ピーク高さ、半分の高さのピーク幅、ピーク下の面積、シグナル対ノイズ比、または当該分野で公知の他の表示法で表示することができる。

10

【0027】

本明細書において、測定された質量または質量ピークの比較は、1つ以上の測定された試料質量ピークを1つ以上の試料もしくは参照質量ピークに対して分析することを意味する。例えば測定された試料質量ピークは、計算された質量ピークパターンと比較して分析ことができ、測定された質量ピークと計算された質量ピークとの重複を使用して、試料質量または分子を同定することができる。参照質量ピークは、参照原子、分子、または分子の断片の質量の表示である。

20

【0028】

本明細書において参照質量は、測定された試料質量が比較される質量である。試料質量と参照質量との比較は、試料質量を参照質量と同じまたは異なるとして同定することができる。かかる参照質量は、計算されるか、データベース中に存在するか、または実験で測定することができる。計算された参照質量は、核酸の予測質量に基づくものでもよい。例えば計算された参照質量は、既知のまたは予測される配列の標的核酸分子の予測された断片化パターンに基づくものでもよい。実験で得られる参照質量は、任意の核酸試料の測定された質量から得られてもよい。例えば実験で得られる質量は、核酸分子を断片化条件下で処理して、断片に捕獲オリゴヌクレオチドを接触させて得られる質量でもよい。参照質量のデータベースは、1つ以上の参照質量（ここで参照質量は計算されるかまたは実験で測定される）を含有することができる；データベースは、標的核酸分子の計算されたかまたは実験で測定された断片化パターンに対応する参照質量を含有することができる；データベースは、2つまたはそれ以上の標的核酸分子の計算されたかまたは実験で測定された断片化パターンに対応する参照質量を含有することができる。

30

【0029】

本明細書において参照核酸分子は、既知のヌクレオチド配列または既知の本体（例えば、配列が不明であるが、DNA配列との相関がわかっている遺伝子座）の核酸分子を意味する。参照核酸は、参照質量を計算するかまたは実験で得るのに使用することができる。参照質量を計算するのに使用される参照核酸は、典型的には既知のヌクレオチド配列を含有する核酸である。参照質量を実験で得るのに使用される参照核酸は、既知の配列を有する（必須ではない）ことができ；参照核酸が既知の配列を持たない時でも、本明細書に開示のまたは当該分野で公知の方法を使用して、参照核酸のヌクレオチド配列を同定することができる。

40

【0030】

本明細書において1つ以上の試料質量（または1つ以上の試料質量ピーク特性）と1つ以上の参照質量（1つ以上の参照質量ピーク特性）およびその標準的变化との相関は、1つ以上の試料質量（または1つ以上の試料質量ピーク特性）と1つ以上の参照質量（1つ以上の参照質量ピーク特性）との比較を意味し、ここで質量の高い類似性は、標的核酸分子またはその断片のヌクレオチド配列が参照核酸のヌクレオチド配列と同じである確率が高いこ

50

とを意味する。

【0031】

本明細書において1つ以上の試料質量ピークと1つ以上の参照質量ピークおよびその標準的变化との相関は、1つ以上の試料質量ピークと1つ以上の参照質量ピークとの比較を意味し、ここで1つ以上の様質量ピークと1つ以上の参照質量ピークの間との1つ以上の質量ピーク特性の高い類似性は、試料標的核酸分子の少なくとも一部が参照核酸の少なくとも一部と同じである確率が高いことを意味するか、または標的核酸の1つ以上のヌクレオチド部分のヌクレオチド配列が参照核酸の1つ以上のヌクレオチド部分のヌクレオチド配列と同じである確率が高いことを意味する。

【0032】

本明細書において標的核酸分子ヌクレオチド配列と参照ヌクレオチド配列との相関は、標的核酸分子のヌクレオチド配列と参照のヌクレオチド配列との類似性または同一性を意味する。

【0033】

本明細書において「分析」は、単一のオリゴヌクレオチドまたはオリゴヌクレオチドの混合物の具体的な性質の測定を意味する。これらの性質には、特に限定されないが、オリゴヌクレオチドまたはオリゴヌクレオチドの混合物のヌクレオチド組成と完全な配列、2つ以上のオリゴヌクレオチド間の単一のヌクレオチド多型性と他の突然変異の存在、オリゴヌクレオチドの質量と長さ、分子または試料中の分子内の配列の存在がある。

【0034】

本明細書において「多重化する」、「多重化された」、「多重化反応」、またはこれらの標準的变化は、単一の反応または単一の質量スペクトル測定もしくは他の配列測定（すなわち、単一の質量スペクトルまたは配列を読む他の方法）中の2つ以上の分子、例えば生物学的分子（例えばオリゴヌクレオチド分子）の同時評価または分析を意味する。

【0035】

本明細書において増幅は、生体高分子（特に核酸）の量を増加させる手段を意味する。選択される5'および3'プライマーに基づき増幅はまた、分析されるゲノムの領域を制限および規定するように作用する。増幅は、当業者に公知の任意の方法により行われ、ポリメラーゼ連鎖反応（PCR）などの使用を含む。多型性の頻度を測定する必要がある場合は、増幅（例えばPCR）は定量的に行わなければならない。

【0036】

本明細書において用語「サイズの統計的範囲」は、断片の一部が特定のサイズ範囲内の大部分の他の断片より実質的に小さいかまたは大きくなるように、部分的切断を使用して作製された大部分の断片のサイズ範囲を意味する。例えば12~30塩基の統計的サイズ範囲は、1ヌクレオチドという小さいオリゴヌクレオチドまたは300ヌクレオチドという大きいオリゴヌクレオチドを含むことができるが、これらの具体的なサイズが起きるのは統計的には比較的まれである。断片の統計的範囲には、60%の断片が所望のサイズ範囲である場合、60%またはそれ以上の断片が所望のサイズ範囲である場合、70%またはそれ以上の断片が所望のサイズ範囲である場合、80%またはそれ以上の断片が所望のサイズ範囲である場合、90%またはそれ以上の断片が所望のサイズ範囲である場合、または95%またはそれ以上の断片が所望のサイズ範囲である場合を含むことができる。

【0037】

本明細書において用語「ハイブリダイズする」またはその標準的变化は、完全または部分的相補鎖への核酸配列の結合を意味する。本明細書において用語「ハイブリダイズする」は、完全に相補的な鎖の結合も、完全に相補的ではない鎖の結合も意味する。すなわち、ハイブリダイズするとは、第1の核酸が第2の核酸に結合する場合を含み、ここで第1の核酸と第2の核酸とは1つ以上のミスマッチ塩基を有する。

【0038】

本明細書において用語「ミスマッチハイブリダイゼーションを排除しない条件下で」は、1つ以上の塩基対ミスマッチを有する捕獲オリゴヌクレオチドの結合を許容するハイブ

10

20

30

40

50

リダイゼーション条件を意味する。ある実施態様において許容されるミスマッチの数は、5つ以下、4つ以下、3つ以下、2つ以下、および1つ以下の塩基対ミスマッチから選択される。

【0039】

本明細書において用語「捕獲された断片」は、捕獲オリゴヌクレオチド（例えば固相上の捕獲オリゴヌクレオチド）に結合した標的核酸断片を意味する。

【0040】

本明細書において「縮重位置」は、4つの典型的に存在する塩基の1つの代わりに、2つ以上のヌクレオチドに結合する置換基を含有するヌクレオチド上の位置を意味する。例えばヌクレオチド上の縮重位置は、ユニバーサル塩基または半ユニバーサル（semi-universal）塩基を含有するヌクレオチド位置でもよい。部分的に縮重しているヌクレオチドは、少なくとも1つの縮重位置と少なくとも1つの非縮重位置とを含有する（例えばユニバーサルまたは半ユニバーサル塩基と、A、G、C、またはT/Uのような非縮重塩基とを含有する）ヌクレオチド、またはあるヌクレオチドに対して他のヌクレオチドに対するより選択的に結合する少なくとも1つの縮重位置を含有する（例えば、少なくとも1つの半ユニバーサル塩基を含有する）ヌクレオチドを意味する。ある実施態様において部分的に縮重しているオリゴヌクレオチドは、少なくとも10%、20%、30%、40%、最大50%までの縮重位置を含有する。例えば20ヌクレオチドの長さを有する捕獲オリゴヌクレオチドについて、これらの部分的に縮重しているオリゴヌクレオチドは、1、2、3、4、5、6、7、8、9、最大10個の縮重位置を含有することができる。別の実施態様において縮重オリゴヌクレオチドは、50%を超える縮重位置（100%の縮重位置を含む）を含有することができる。例えば20ヌクレオチドの長さを有するオリゴヌクレオチドは、20個の半ユニバーサルヌクレオチド、または10個のヌクレオチドと10個の半ユニバーサルヌクレオチドを含有することができる。

【0041】

本明細書において固体支持体粒子は、別個の粒子の形である材料を意味する。これらの粒子は任意の形と大きさを有するが、典型的には100 mmまたはそれ以下、50 mmまたはそれ以下、10 mmまたはそれ以下、1 mmまたはそれ以下、100 μmまたはそれ以下、50 μmまたはそれ以下である少なくとも1つの大きさを有し、かつ典型的には10 0mm³またはそれ以下、50 mm³またはそれ以下、10 mm³またはそれ以下、および1 mm³またはそれ以下、100 μm³またはそれ以下であるサイズを有し、立法ミクロンのオーダーでもよく；典型的には粒子は、約1.5ミクロンより大きくかつ約15ミクロンより小さい直径（例えば約4~6ミクロン）を有する。かかる粒子はまとめて「ビーズ」と呼ばれる。

【0042】

本明細書において「固体支持体」は、その上で反応が行われ、および/または反応生成物が特定可能な位置に保持される表面を与えることができる不溶性支持体を意味する。支持体は、実質的に任意の不溶性または固体材料から作製することができる。例えばシリカゲル、ガラス（例えば、制御孔ガラス（CPG）、ナイロン、ワング（Wang）樹脂、メリフィールド（Merrifield）樹脂、セファデックス、セファロース、セルロース、金属表面（例えば、鋼鉄、金、銀、アルミニウム、および銅）、シリコン、およびプラスチック材料（例えば、ポリエチレン、ポリプロピレン、ポリアミド、ポリエステル、ニフ化ポリビニリデン（PVDF））。固体支持体の例には、特に限定されないが、平面支持体、例えばガラス繊維フィルター、ガラス表面、金属表面（鋼鉄、金、銀、アルミニウム、銅、およびシリコン）、およびプラスチック材料がある。固体支持体は、カートリッジベース上に取り付けるための任意の所望の形であり、例えば特に限定されないが、プレート、膜、ウェーハー、穴のあるウェーハー、多孔性3次元支持体、および当業者に公知の他の形や型がある。支持体の例は、別々の位置の試料を受けるかまたは連結するように設計された平らな表面、例えば試料を受けるか、含有するかまたは結合するための親水性位置の周りに疎水性領域を有する平らな表面である。

【0043】

10
20
30
40
50

本明細書の核酸断片化において用語「非特異的に切断された」または「非特異的断片化」は、異なるサイズとヌクレオチド配列含量の種々の断片がランダムに作製されるように、全体のランダムな位置の標的核酸分子の断片化を意味する。本明細書においてランダムな位置の断片化は絶対的な数学的ランダム性を必要としないが、断片化の強い配列に基づく選択性が欠如しているのみである。例えば放射線照射または剪断力手段による断片化は、ほとんど任意の位置でDNAを切断することができるが、かかる方法はいくつかの位置で他の位置よりわずかに高頻度に断片化を引き起こすことがある。しかし配列選択性がほんのわずかであるほとんどすべての位置での断片化は、本明細書の目的においてランダムであると見なされる。本明細書に記載の方法を使用する非特異的切断は、重複ヌクレオチド断片の生成を引き起こす。

10

【0044】

本明細書において用語、部分的または不完全切断、部分的不完全断片化、またはその標準的变化は、特定の断片化条件について各切断部位のほんの一部のみが実際に切断される反応を意味する。断片化条件は、特に限定されないが、酵素的、化学的または物理的力の存在でもよい。本明細書において部分的断片化を行う1つの方法は、切断反応が完了した時点でさえ、特定の切断部位が切断されないヌクレオチドまたはアミノ酸（これらは標的生物学的分子を部分的に切断させる）を含有するように、標的生物学的分子産生時に切断可能なまたは切断不可能なヌクレオチドまたはアミノ酸の混合物を使用することである。例えば切断されていない標的生物学的分子が4つの可能な切断部位（例えば、核酸の切断される塩基）を有する場合、部分的切断により得られる生成物の混合物は、第1、第2、第3、または第4の切断部位の1つの切断；2つの切断部位の任意の1つまたはそれ以上の組合せの2重切断；または3つの切断部位の任意の1つまたはそれ以上の組合せの2重切断から得られる標的生物学的分子の断片の任意の組合せを有する。部分的切断からの生成物は、全切断からの生成物と同じ混合物中に存在してもよい。

20

【0045】

本明細書において用語「重複断片」は、共通の未変性の標的核酸からの1つ以上のヌクレオチド位置を有する断片を意味する。本明細書において「統計的重複断片」は、規定されたサイズの亜集団が少なくとも1つの他の断片と重複する断片群を意味する。例えば統計的重複断片は、少なくとも50%、少なくとも60%、少なくとも70%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも98%の断片が、少なくとも1つの他の断片と重複する断片群を意味する。

30

【0046】

本明細書において「非特異的RNase」は、切断部位のヌクレオチド配列に無関係にRNA分子を切断する酵素を意味する。非特異的RNaseの例はRNaseIである。

【0047】

本明細書において「非特異的DNase」は、切断部位のヌクレオチド配列に無関係にDNA分子を切断する酵素を意味する。非特異的DNaseの例はDNaseIである。

【0048】

本明細書において用語「単一塩基カッター」は、ある特定の塩基（例えば、DNAについてはA、C、T、もしくはG、またはRNAについてはA、C、U、もしくはG）またはある特定の種類の塩基（例えば、プリンまたはピリミジン）を認識し切断する制限酵素を意味する。

40

本明細書において用語「1-1/4カッター」は、核酸中の2塩基ストレッチを認識し切断する制限酵素（ここで1つの塩基位置の本体は固定され、他の塩基位置の本体は4つの典型的に存在する塩基の任意の3つである）を意味する。

【0049】

本明細書において用語「1-1/2カッター」は、核酸中の2塩基ストレッチを認識し切断する制限酵素（ここで1つの塩基位置の本体は固定され、他の塩基位置の本体は4つの典型的に存在する塩基の任意の2つである）を意味する。

【0050】

本明細書において用語「2塩基カッター」または「2カッター」は、2塩基の長さの特異

50

的核酸部位を認識し切断する制限酵素を意味する。

【0051】

本明細書において用語「質量シグナルのセット」は、2つまたはそれ以上の核酸断片について作製された2つまたはそれ以上の質量測定値を意味する。

【0052】

本明細書においてスコア化またはスコアは、特定の配列変化候補が標的核酸またはタンパク質配列中に実際に存在する確率の計算を意味する。スコアの値は、実際の標的配列に対応する配列変化候補を決定するのに使用される。通常、標的配列の試料セットにおいて最も高いスコアは、標的分子の最も可能性の高い配列変化を示すが、単一の標的配列が存在する時、例えば陽性スコアの検出のような他の選択規則も使用することができる。

10

【0053】

本明細書においてシミュレーション（またはシミュレートする）は、核酸またはタンパク質の配列、および特定の切断試薬についての核酸またはタンパク質配列中の予測される切断部位に基づく断片化パターンの計算を意味する。断片化パターンは、数表（例えば、参照生物学的分子の断片の質量シグナルに対応するピークのリスト）として、質量スペクトルとして、ゲル上のバンドパターンとして、質量分布を測定する技術の表示としてシミュレートすることができる。シミュレーションは、多くの場合コンピュータプログラムにより行われる。

【0054】

本明細書において切断のシミュレートは、標的分子または参照分子が仮想的に切断されるインシリコ（*in silico*）法を意味する。

20

【0055】

本明細書においてインシリコ（*in silico*）は、コンピューターを使用して行われる研究や実験を意味する。インシリコ（*in silico*）法には、特に限定されないが、分子モデリング研究、生物学的分子ドッキング実験、および分子構造および/または方法（例えば分子相互作用）の仮想表示がある。

【0056】

本明細書において用語「ヌクレオチド配列を構築する」は、かかる構築のために設計される種々のアルゴリズムの1つを使用して標的核酸分子のヌクレオチド配列を解明する過程を意味する。

30

【0057】

本明細書において対象は、特に限定されないが、動物、植物、細菌、ウイルス、寄生体、および他の生物または核酸を有する物質がある。対象は特に、哺乳動物、好ましくは（必ずしもではないが）ヒトである。患者は、疾患または障害に罹った対象を意味する。

本明細書において表現型は、生物の見分けが可能な形質を含むパラメータのセットを意味する。表現型は、対象が動物である場合、物理的形質でも、精神的（例えば感情的）形質でもよい。

【0058】

本明細書において「帰属（assignment）」は、核酸またはタンパク質断片の位置が、特定の分子量および特定の末端ヌクレオチドもしくはアミノ酸を示すという決定を意味する。

40

【0059】

本明細書において「a」は1つまたはそれ以上を意味する。

【0060】

本明細書において「複数」は、2つまたはそれ以上を意味する。例えば複数のポリヌクレオチドまたはポリペプチドは、それぞれが異なる配列を有する2つまたはそれ以上のポリヌクレオチドまたはポリペプチドを意味する。そのような差は、配列間の天然に存在する変化（例えば、ヌクレオチドまたはコードされたアミノ酸の対立遺伝子の変化）によるもの、または種々の配列への特定の修飾の導入（例えば複数の各核酸またはタンパク質への質量修飾ヌクレオチドの示差的導入）によるものでもよい。

50

【0061】

本明細書において「明白な」は、標的分子中の特定の配列変化（例えば変異）に対応するピークまたはシグナルのユニークな帰属を意味し、多くの分子または変異が複合している場合、特定の配列変化を示すピークが、各変異または各分子にユニークに帰属される。

【0062】

本明細書においてデータ処理ルーチンは、得られたデータ（すなわちアッセイの最終結果）の生物学的意義を決定する、ソフトウェアで具体化できる方法を意味する。例えばデータ処理ルーチンは、採取されたデータに基づいて遺伝子型決定を行うことができる。本明細書に記載のシステムと方法ではデータ処理ルーチンはまた、得られた結果に基づき、装置および/またはデータ採取ルーチンを制御することができる。データ処理ルーチンおよびデータ採取ルーチンは統合してフィードバックを与えて装置によるデータ採取を自動させることができ、こうして本明細書に記載のアッセイに基づく判定法を提供する。

10

【0063】

本明細書において複数の遺伝子は、少なくとも2、5、10、25、50、100、250、500、1000、2,500、5,000、10,000、100,000、1,000,000、またはそれ以上の遺伝子を含む。複数の遺伝子は、ある生物または複数の生物の完全なまたは部分的なゲノムを含むことができる。生物の種類を選択することは、そこから遺伝子制御領域が選択されるゲノムを決定する。遺伝子スクリーニングのための生物の例には、動物、例えば哺乳動物（ヒトを含む）およびげっ歯類（例えばマウス）、昆虫、酵母、細菌、寄生虫、および植物を含む。

【0064】

本明細書において「試料」は、検出すべき材料を含む組成物を意味する。好適な実施態様において試料は「生物学的試料」である。用語「生物学的試料」は、生きた供給源（例えば、ヒトまたは他の哺乳動物のような動物、植物、細菌、真菌、原生生物、またはウイルス）から得られる任意の材料を意味する。生物学的試料は任意の形でよく、固体材料〔例えば、組織、細胞、細胞ペレット、細胞抽出物、または生検試料〕または生物学的液体〔例えば、尿、血液、血漿、血清、唾液、喀痰、羊水、感染もしくは炎症部位からの浸出液、頬細胞を含む口内洗浄液、髄液、滑液、臓器、精液、眼内液、粘液、分泌液（例えば胃液または乳汁）〕、および病理試料、例えばパラフィン包埋したホルマリン固定試料がある。好ましくは固体材料は液体と混合される。特に本明細書において生物学的試料（例えば核酸）の質量スペクトル分析が行われる時、試料はマトリックスと混合することができる。「から得られる」は、試料を核酸分子の精製または単離および/または増幅などにより処理できることを意味する。

20

30

【0065】

本明細書において組成物は任意の混合物を意味する。これは溶液、懸濁物、液体、粉末、ペースト、これらの水性、非水性または任意の組合せでもよい。

本明細書において組合せは、2つまたはそれ以上の物質の任意の組合せである。

本明細書において用語「アンプリコン」は、複製できるDNAの領域を意味する。

本明細書において用語「完全な切断」または「全切断」は、特定の切断試薬により認識される切断部位が完全に切断される切断反応を意味する。

【0066】

本明細書において用語「擬陽性」は、バックグラウンドノイズより高く、予測される事象の結果としては生成されないシグナルを意味する。例えば標的核酸分子配列を反映しないピークが観察される時、または核酸もしくはタンパク質の特異的な実際ではないもしくはシミュレートされた切断ではない方法により断片が生成される時、擬陽性が発生する。

40

【0067】

本明細書において用語「擬陰性」は、実際の測定からは欠失しているが、予測された実際のシグナルを意味する。例えば実際の質量スペクトルで観察されない質量シグナルが、対応するシミュレートされたスペクトル中に存在することが計算された時、擬陰性が発生する。

【0068】

50

本明細書において断片化または切断は、核酸またはタンパク質分子が小片に分割される方法を意味する。断片化または切断法には、物理的切断、酵素的切断、化学的切断、および核酸の小片が産生される任意の他の方法がある。

【0069】

本明細書において断片化条件または切断条件は、1つまたはそれ以上の断片化試薬、緩衝液、または実際のもしくはシミュレートされた切断反応を行うのに使用される他の化学的もしくは物理的条件のセットを意味する。かかる条件には、例えば時間、温度、pH、または緩衝液の選択のような反応のパラメータがある。

【0070】

本明細書において切断されない切断部位は、切断試薬のための公知の認識部位であるが、反応の条件（例えば時間、温度）下では、または試薬による切断を避けるための切断認識部位での塩基の修飾のために切断試薬により切断されない切断部位を意味する。

10

【0071】

本明細書において相補的切断反応は、同じ標的もしくは参照核酸もしくはタンパク質の交互の切断パターンが生成されるように、異なる切断試薬を使用してまたは切断試薬の切断特異性を変更することにより、同じ標的もしくは参照核酸もしくはタンパク質について行われるかまたはシミュレートされる切断反応を意味する。

【0072】

本明細書において流体は、流ることができる任意の組成物を意味する。従って流体は、半固体、ペースト、溶液、水性混合物、ゲル、ローション、クリーム形の組成物、および他のかかる組成物を包含する。

20

【0073】

本明細書において細胞抽出物は、溶解細胞または破碎細胞から作製される調製物または画分を意味する。

本明細書においてキットは、組合せで使用するための使用説明書および/または試薬および装置とともに成分を随時パッケージングした組合せである。

本明細書においてシステムは、要素とソフトウェア、および本明細書の方法を制御および指令するための任意の他の要素との組合せを意味する。

【0074】

本明細書においてソフトウェアは、コンピューターで実行するとコンピューター操作を実行するコンピューターで読めるプログラム指令を意味する。典型的にはソフトウェアは、コンピューターで読める媒体（特に限定されないが、フロッピー（登録商標）ディスク、ハードディスク、および磁気テープを含む磁気媒体；CD-ROMディスク、DVDディスク、磁気-光ディスクを含む光学的媒体；およびプログラム指令が記録された他の媒体）に記録されたプログラム指令を含むプログラム製品で提供される。

30

【0075】

本明細書において用語「標的核酸または標的核酸分子」は、分析することが目的である核酸分子を意味する。標的核酸分子は、1本鎖分子でも2本鎖分子でもよい。

本明細書において用語「部分的に消化された」は、サブセットの制限部位のみが切断されることを意味する。

40

【0076】

本明細書において「複雑さを制御する」およびその標準的変種は、異なるヌクレオチド配列を有する核酸分子の数、変動、または数と変動を操作するための方法を意味する。例えば捕獲オリゴヌクレオチドにハイブリダイズした標的核酸断片の複雑さを制御することは、特定の捕獲オリゴヌクレオチドプローブ配列にハイブリダイズする異なるヌクレオチド配列を有する標的核酸断片の数、変動、または数と変動を制御するための実験条件を操作することを意味する。捕獲オリゴヌクレオチドプローブにハイブリダイズする異なる標的核酸配列の数は、捕獲オリゴヌクレオチドプローブの特定のヌクレオチド配列の少なくとも一部にハイブリダイズする同一ではない標的核酸または標的核酸断片の量を意味する。例えば互いに異なる配列を有する2つまたはそれ以上の標的核酸断片は、単一のアレイ

50

位置のすべての捕獲オリゴヌクレオチドプローブが同じヌクレオチド配列を有する単一のアレイ位置にハイブリダイズすることができる。ある例では、ハイブリダイゼーションが捕獲オリゴヌクレオチドと標的核酸断片の2つの異なるヌクレオチド配列の間の塩基対合を含む場合、異なる配列を有する2つの標的核酸は捕獲オリゴヌクレオチドにハイブリダイズすることができる。すなわち本明細書に開示の方法のある実施態様において捕獲オリゴヌクレオチドは、2つまたはそれ以上の異なるヌクレオチド配列と塩基対合することができる。捕獲オリゴヌクレオチドプローブにハイブリダイズする異なる標的核酸配列の変動は、長さとヌクレオチド配列の点で、捕獲オリゴヌクレオチドプローブにハイブリダイズする異なる標的核酸配列の配列同一性の程度を意味する。

【0077】

本明細書において捕獲オリゴヌクレオチドプローブにハイブリダイズする配列の数を「調節する」ことは、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸断片の配列の数、変動、および数と変動を設定または修飾するために、条件を設定または修飾することを意味する。設定または修飾可能な条件の例は本明細書で以下に示される。従って捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸断片の複雑さは、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸配列の数を調節することにより制御することができ、これは、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸断片の数、変化、および数と変動に影響を与える条件を設定または修飾することにより達成できる。

【0078】

本明細書において用語「半特異的捕獲」は、部分的に縮重していても縮重ヌクレオチド塩基を含有しなくてもよい単一の捕獲オリゴヌクレオチド配列への2つまたはそれ以上の異なる標的核酸断片の結合を意味する。半特異的捕獲は、すべての標的核酸断片またはランダムに結合する核酸断片への結合は含まないが、少なくとも1つの他の標的核酸断片よりも2つまたはそれ以上の標的核酸断片への結合を意味する。

【0079】

本明細書においてアレイの捕獲オリゴヌクレオチドのヌクレオチド配列を記載するのに用語「ユニークな」および用語「同一の配列」の使用は、厳密な同一性を意味し、従って第1のオリゴヌクレオチドが配列ATCGを有し、第2のオリゴヌクレオチドが配列ATCGAを有する場合、2つのオリゴヌクレオチドはユニークであり、同一の配列は持たない。同様に本明細書において、捕獲オリゴヌクレオチドにハイブリダイズする1つまたはそれ以上の標的核酸または標的核酸断片への言及は、特に明記しない場合は、同一の配列を有する複数の捕獲オリゴヌクレオチドプローブの1つに別々に結合する1つまたはそれ以上の標的核酸または標的核酸断片のそれぞれを意味する。典型的には1つまたはそれ以上の標的核酸または標的核酸断片は、特定のアレイ位置で捕獲オリゴヌクレオチドにハイブリダイズする。

【0080】

本明細書において用語「部分的に縮重した捕獲オリゴヌクレオチド」は、同様の特異性を有する少なくとも2つの異なるヌクレオチド配列にハイブリダイズするが、同様の特異性を有するすべての可能なヌクレオチド配列にはハイブリダイズしないオリゴヌクレオチドを意味する。例えば部分的に縮重した捕獲オリゴヌクレオチドは、ユニバーサル塩基を有するオリゴヌクレオチドでもよい。

【0081】

本明細書において用語「すべての理論的組合せ」は、ある長さのすべての可能なヌクレオチド配列が表示されるように、その長さのオリゴヌクレオチドの完全な群を意味する。

【0082】

本明細書において「縮重塩基」は、「ユニバーサル塩基」または「半ユニバーサル塩基」、または標的核酸もしくは標的核酸断片の2つまたはそれ以上の塩基に、同様の特異性で塩基対合できる他の塩基を意味する。

【0083】

10

20

30

40

50

本明細書において「ユニバーサル塩基」は、実質的に区別することなくゲノムDNA中に存在する4つの任意のヌクレオチドに結合できる塩基を意味する。本明細書においてユニバーサル塩基の例には、イノシン、キサントシン、3-ニトロピロール (Bergstromら、*Abstr. Pap. Am. Chem. Soc.* 206(2): 308 (1993); Nicholsら、*Nature* 369: 492-493; Bergstromら、*J. Am. Chem. Soc.* 117: 1201-1209 (1995))、4-ニトロインドール (Loakesら、*Nucleic Acids Res.* 22: 4039-4043 (1994))、5-ニトロインドール (Loakesら、(1994))、6-ニトロインドール (Loakesら、(1994)) ; ニトロイミダゾール (Bergstromら、*Nucleic Acids Res.* 25: 1935-1942 (1997))、4-ニトロピラゾール (Bergstromら、(1997))、5-アミノインドール (Smithら、*Nucl. Nucl.* 17: 555-564 (1998))、4-ニトロベンズイミダゾール (Seelaら、*Helv. Chim. Acta* 79: 488-498 (1996))、4-アミノベンズイミダゾール (Seelaら、*Helv. Chim. Acta* 78: 833-846 (1995))、フェニルC-リボヌクレオシド (Millicanら、*Nucleic Acids Res.* 12: 7435-7453 (1984)) ; Matulic-Adamicら、*J. Org. Chem.* 61: 3909-3911 (1996))、ベンズイミダゾール (Loakesら、*Nucl. Nucl.* 18: 2685-2695 (1999) ; Papageorgiouら、*Helv. Chim. Acta* 70: 138-141 (1987))、5-フルオロインドール (Loakesら (1999))、インドール (Girgisら、*J. Heterocycle Chem.* 25: 361-366 (1988)) ; 非環状糖類似体 (Van Aerschotら、*Nucl. Nucl.* 14: 1053-1056 (1995) ; Van Aerschotら、*Nucleic Acids Res.* 23: 4363-4370 (1995) ; Loakesら、*Nucl. Nucl.* 15: 1891-1904 (1996)) (ヒポキサントシン、イミダゾール、4,5-ジカルボキサミド、3-ニトロイミダゾール、5-ニトロイミダゾールを含む) ; 芳香族類似体 (Guckianら、*J. Am. Chem. Soc.* 118: 8182-8183 (1996) ; Guckianら、*J. Am. Chem. Soc.* 122: 2213-2222 (2000)) (ベンゼン、ナフタレン、フェナントレン、ピレン、ピロール、ジフルオロトルエンを含む) ; イソカルボスチリルヌクレオシド誘導体 (Bergerら、*Nucleic Acids Res.* 28: 2911-2914 (2000) ; Bergerら、*Angew. Chem. Int. Ed. Engl.*, 39: 2940-2942 (2000)) (MICS、ICSを含む) ; 水素結合類似体 (N8-ピロロピリジンを含む) (Seelaら、*Nucleic Acids Res.* 28: 3224-3232 (2000)) ; およびLNAs、例えばアリアル-C-LNA (Babuら、*Nucleosides, Nucleotides & Nucleic Acids* 22: 1317-1319 (2003) ; WO03/020739) がある。

10

20

30

40

50

【0084】

本明細書において用語「半ユニバーサル塩基」は、2つまたは3つのデオキシリボヌクレオチドに選択的に結合するが、同じまたは同様の特異性で4つすべての典型的に存在するヌクレオチド(すなわち、DNAではA、C、G、およびT、RNAではA、C、G、およびU)には結合しない塩基を意味する。例えば、半ユニバーサル塩基は、少なくとも1つの他の典型的に存在するヌクレオチドに結合するよりはるかに大きなレベルで2つまたは3つの典型的に存在するヌクレオチドに結合する。

【0085】

本明細書において「固体支持体」(不溶性支持体または固体支持体とも呼ばれる)は、目的の分子、典型的には生物学的分子、有機分子、または生体特異的リガンドが結合するかまたは接触する任意の固体もしくは半固体もしくは不溶性支持体を意味する。かかる材料には、化学的および生物学的分子合成や分析のための親和性マトリックスもしくは支持体として使用される任意の材料、例えば特に限定されないが、ポリスチレン、ポリカーボネート、ポリプロピレン、ナイロン、ガラス、デキストラン、キチン、砂、軽石、アガロース、多糖、デンドリマー、パッキーボール、ポリアクリルアミド、シリコン、ゴム、および固相合成、親和性分離と精製、ハイブリダイゼーション反応、イムノアッセイ、および他のかかる応用のための支持体として使用される他の材料がある。

【0086】

本明細書において標的核酸または参照核酸の「部分」は、完全な核酸を含まない核酸のヌクレオチド配列または領域を意味する。例えば部分は、核酸の短いヌクレオチド配列、例えばSNP、メチル化C、またはマイクロサテライト(microsatellite)でもよい。部分はまた例えばヌクレオチド配列が既知または未知の核酸の特定の断片でもよく、ここで断片は、例えば生物、株または種の変動による配列の差から生じ、断片は本明細書に開示の方

法を使用して生成される。部分はまた、他の領域に対して異なって相互作用するか、または異なって処理される核酸の領域でもよい。

【0087】

B. 核酸分子の配列決定法

以下に核酸を配列決定するための方法が提供される：

- a) 標的核酸の重複断片を作製する；
- b) ミスマッチハイブリダイゼーションを排除しない条件下で、固体支持体上の捕獲オリゴヌクレオチドのアレイに断片をハイブリダイズさせて、捕獲断片のアレイを生成する；
- c) 質量スペクトル分析を使用して各アレイ位置で捕獲断片の質量を測定する；および
- d) 各アレイ位置から得られた質量シグナルのセットから標的核酸のヌクレオチド配列を構築する。

10

【0088】

また、以下を含む、核酸を配列決定する方法が提供される：

- a) 標的核酸の重複断片を作製する；
- b) 固体支持体上の捕獲オリゴヌクレオチドのアレイに断片をハイブリダイズさせて、捕獲断片のアレイを生成する（ここで、捕獲オリゴヌクレオチドの少なくとも1つのサブセットは部分的に縮重している）；
- c) 質量スペクトル分析を使用して各アレイ位置で捕獲断片の質量を測定する；および
- d) 各アレイ位置から得られた質量シグナルのセットから標的核酸のヌクレオチド配列を構築する。

20

【0089】

また、以下を含む、核酸を配列決定する方法が提供される：

- a) 標的核酸の重複断片を作製する；
- b) 固体支持体上の捕獲オリゴヌクレオチドのアレイに断片をハイブリダイズさせて、捕獲断片のアレイを生成する（ここで、少なくとも1つの捕獲オリゴヌクレオチドは2つまたはそれ以上の断片にハイブリダイズする）；
- c) 質量スペクトル分析を使用して各アレイ位置で捕獲断片の質量を測定する；および
- d) 各アレイ位置から得られた質量シグナルのセットから標的核酸のヌクレオチド配列を構築する。

30

【0090】

これらの本明細書に記載の各方法の実施態様において、標的核酸の重複断片はランダムに作製される。

【0091】

これらの本明細書に記載の各方法の別の実施態様において、捕獲断片の質量を測定する工程c)の前に、ハイブリダイズした断片は溶液中に再可溶化される。かかる再可溶化は、例えば再可溶化断片を含む溶液中に浸漬されるピンアレイの公知の使用を可能にして、質量スペクトル分析のための適当なチップに断片を移動させる。

【0092】

上記したように本明細書に記載の方法は、固相チップに結合した標的核酸のSBHおよび/または質量スペクトル分析を使用して達成される長さより長い標的核酸配列読み長さを可能にする。別の実施態様において本明細書に記載の方法により、より短い（例えば、200、300、400、500、600、700、800、900、1,000、1,500塩基）複数の標的核酸断片が配列決定または分析できる。本明細書の方法は、5、10、15、20、50、100、200、500またはそれ以上の核酸断片の分析を含む。これらの複数のより短い配列セットはまた、具体的な配列が既知の場合に再配列決定法で有用性である。これらの複数のより短い配列セットはまた、多重遺伝子型判定、ハプロタイプ型判定、SNP、およびメチル化検出法に有用である。

40

【0093】

C. 標的核酸分子

50

標的核酸分子は、1本鎖または2本鎖の核酸分子でもよい。具体的な実施態様ではMALDI-TOF MS分析を使用する時、またはRNA転写に基づくアプローチがチップ上にハイブリダイズした断片の収率を上昇させる時、またはDNA捕獲オリゴにハイブリダイズしたRNAがハイブリダイゼーション後のさらなる修飾を可能にする時は、DNAよりRNAが使用される。別の実施態様ではDNAが使用され、DNA捕獲オリゴにハイブリダイズされ、DNA：DNAハイブリッドのための、ハイブリダイゼーション後のさらなる修飾が行われる。

【0094】

1. 供給源

標的核酸は、1本鎖DNA、2本鎖DNA、cDNA、1本鎖RNA、2本鎖RNA、DNA/RNAハイブリッド、およびDNA/RNAモザイク核酸から選択することができる。標的核酸はまた、メチル化DNAおよび例えばシュードウリジンを含むRNAのような修飾核酸を含むことができる。標的核酸は生物学的試料から直接単離されるか、または生物学的試料から核酸断片を増幅またはクローニングすることにより得ることができる。クローニングまたは増幅の鋳型として機能する標的核酸は、全体の無傷の標的核酸または標的核酸断片でもよく、ここで標的核酸断片はハイブリダイゼーションまたは質量測定のための所望の長さでも、またはまず標的核酸断片が増幅され次に1回またはそれ以上の追加の断片化工程に付される中間の長さでもよい。

【0095】

本明細書に記載の方法で使用される試料は、適用される方法の目的に従って選択することができる。例えば試料は単一の個体由来でもよく、ここで試料を調べて、個体の1つまたはそれ以上の位置でヌクレオチド配列が測定される。当業者は本明細書に記載の方法を使用して、調べるべき所望の試料を測定することができる。

【0096】

試料は、動物、植物、細菌、ウイルス、寄生虫、鳥、は虫類、両生類、真菌、魚、および他の植物や動物を含む任意の対象でもよい。特に対象は哺乳動物、典型的にはヒトである。対象からの試料は任意の形でよく、例えば固体材料[例えば、組織、細胞、細胞ペレット、細胞抽出物、または生検試料]、または生物学的液体[例えば、尿、血液、間質液、腹膜液、血漿、リンパ、腹水、汗、唾液、卵胞液、乳汁、非乳汁性分泌液、血清、髄液、便、精液、肺喀痰、羊水、感染もしくは炎症部位からの浸出液、頬細胞を含む口内洗浄液、髄液、滑液、対象が産生する他の任意の液体試料]がある。さらに試料は、骨髄、上皮、胃、前立腺、腎臓、膀胱、乳房、結腸、肺、膵臓、子宮内膜、ニューロン、および筋肉を含む採取された組織でもよい。試料は、組織、臓器、およびパラフィン包埋されたホルマリン固定試料のような病理試料でもよい。

【0097】

2. 調製

いくつかの試料は本明細書に記載の方法で直接使用できることを、当業者は認識するであろう。例えば所望の細胞または核酸分子の純度を上げるための精製または操作をすることなく、本明細書に記載の方法を使用して試料を調べることができる。

【0098】

所望であれば、例えばManiatisらが記載したような公知の技術(「分子クローニング：実験室マニュアル(Molecular Cloning, A Laboratory Manual)」、コールドスプリングハーバー(Cold Spring Harbor)、ニューヨーク州、280~281頁(1982))を使用して試料を調製することができる。例えば、本明細書に記載の方法を使用して調べられる試料は、試料中の所望の細胞または核酸の純度を上げるために、1回またはそれ以上の精製工程で処理することができる。所望であれば、固体材料に液体を混合することができる。

【0099】

基本的にすべての生物または体内の組織もしくは臓器ならびに培養細胞から試料中の核酸を単離する方法は、公知である。例えば試料は、公知の溶解緩衝液、超音波処理、電気穿孔法、およびこれらの組合せを使用して、臓器、組織または細胞試料をホモジナイズし、細胞を溶解するように処理することができる。さらなる精製は、当業者に公知のように

必要に応じて行うことができる。さらに以後の工程に含むことができる種々の試薬を含むことができる。これらには、塩、緩衝液、タンパク質（例えばアルブミン）、洗浄剤、および最適なハイブリダイゼーション反応もしくは酵素反応を促進し、および/または非特異的もしくはバックグランド相互作用を小さくするのに使用できる試薬がある。また標的核酸分子の試料調製法と純度に依存して、測定法の効率を改善する試薬、例えばプロテアーゼインヒビター、ヌクレアーゼインヒビター、および抗微生物剤を使用することができる。

【0100】

3. 標的核酸分子のサイズと組成

使用できる標的核酸分子の長さは、標的核酸分子の配列、断片化に使用される特定の手法、ハイブリダイゼーションのために使用されるオリゴヌクレオチドを捕獲する特定の手法、ヌクレオチド配列を測定すべき全標的核酸分子の割合、配列決定の正確性の所望のレベル、および配列決定の性質（例えば、新規配列決定対再配列決定）により変化する。例えば、標的核酸分子の長さは、本明細書に開示の方法を使用して、少なくとも約1%、少なくとも約3%、少なくとも約5%、少なくとも約10%、少なくとも約20%、少なくとも約30%、少なくとも約40%、少なくとも約50%、少なくとも約60%、少なくとも約70%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約98%、少なくとも約99%、またはすべての標的核酸分子が測定できる長さ限定することができる。例えば標的核酸分子は、少なくとも約20、25、30、35、40、50、60、70、80、90、100、120、140、160、180、200、225、250、275、300、350、400、450、500、550、600、700、800、900、1000、1200、1400、1600、1800、2000、2500、または3000塩基の長さでもよい。典型的には標的核酸分子は、約10,000、5000、4000、3000、2500、2000、1500、1000、900、800、700、600、500、450、400、350、280、260、240、220、200、190、180、170、160、150、140、130、120、110、または100塩基以下の長さである。

【0101】

4. 増幅

ある実施態様において、以後の工程で処理でき測定できる核酸分子の数を増やすために、そして場合により標的核酸配列を処理するために、標的核酸分子は増幅することができる。増幅は、ポリメラーゼ連鎖反応（PCR）、逆転写ポリメラーゼ連鎖反応（RT-PCR）、ローリングサークル型増幅、全ゲノム増幅、鎖置換増幅（SDA）、および転写に基づく方法により行うことができる。増幅法では、種々の異なる増幅産物を作製することができる種々の異なる増幅法で反応条件および/または反応物を変化させてもよい。

【0102】

a. 反応パラメータ

相補鎖（存在するならば）が分離され、プライマーが鎖にハイブリダイズされ、プライマーがそこにヌクレオチドを加えて新しい相補鎖を生成する増幅工程を行うことができる。鎖の分離は、別の工程としてまたはプライマー伸長産物の合成と同時に行うことができる。この鎖の分離は、物理的、化学的、または酵素的手段を含む種々の適切な変性条件（「変性」という単語はかかる手段のすべてを含む）を使用して行われる。核酸鎖を分離する1つの物理的方法は、標的核酸分子を変性するまで加熱することを含む。典型的な加熱変性は、約80 ~ 105 の範囲の温度を約1~10分間行う。鎖の分離はまた、高塩条件または強い塩基性条件を含む化学的手段により行うことができる。鎖の分離はまた、ヘリカーゼとして知られる酵素群からの酵素または酵素RecA（これはヘリカーゼ活性を有し、リボATPの存在下でDNAを変性させることが知られている）により誘導することができる。ヘリカーゼを用いる核酸の鎖分離に適した反応条件は、Kuhn Hoffmann-Berling、*CSH-Quantitative Biology*, 43: 63 (1978)により記載され、ReaAを使用する技術は、C. Radding, *Ann. Rev. Genetica* 16: 405-437 (1982)の総説がある。

【0103】

各増幅工程後に、増幅産物は典型的には2本鎖であり、各鎖は互いに相補的である。相補鎖は分離することができ、両方の鎖をさらなる核酸鎖の合成用の鋳型として使用するこ

とができる。この合成は、鋳型へのプライマーのハイブリダイゼーションが起きる条件下で行うことができる。一般に合成は、典型的には約7~9のpH（例えば約pH 8）で緩衝化した水溶液で起きる。典型的には、分離された鋳型鎖を含む緩衝液にモル過剰の2つのオリゴヌクレオチドプライマーを加えることができる。ある実施態様において標的核酸の量は不明であり（例えば、本明細書に記載の方法は診断応用で使用される）、従って相補鎖の量に対するプライマーの量を正確に決定することができない。

【0104】

ある方法の例ではデオキシリボヌクレオチド3リン酸であるdATP、dCTP、dGTP、およびdTTPは、別々にまたはプライマーと一緒に合成混合物に加えられ、得られた溶液は約90~100 で約1~10分間、典型的には1~4分間加熱することができる。この加熱時間後、溶液をほぼ室温まで冷却することができる。プライマー伸長反応を実施するための酵素（ここでは「重合のための酵素」）を冷却混合物に加え、当該分野で公知の条件下で反応を行わせる。この合成（または増幅）反応は、室温から重合のために酵素が機能しなくなる温度より高い温度で起きることができる。例えば重合のための酵素は、酵素が熱安定性なら、室温より高い温度で使用することができる。ある実施態様において増幅法は、本明細書に記載され当業者により一般的に行われるPCRである。代替法も記載されており、使用することができる。この目的に適した種々の酵素は当該分野で公知であり、例えば大腸菌（*E. coli*）DNAポリメラーゼI、大腸菌（*E. coli*）DNAポリメラーゼIのクレノウ断片、T4 DNAポリメラーゼ、他の利用可能なDNAポリメラーゼ、ポリメラーゼムテイン、逆転写酵素、および熱安定性酵素（高温で、典型的には増幅すべき核酸の変性を引き起こす温度でプライマー伸長を行う酵素）を含む他の酵素がある。

10

20

【0105】

b. 修飾ヌクレオシド

ある実施態様において標的核酸は、修飾ヌクレオシド（例えば修飾ヌクレシド3リン酸）を使用して増幅される。いくつかの修飾は、各切断法により標的核酸配列の切断特異性を付与するかまたは変化させ得る。他の修飾（例えば質量修飾）は、標的核酸、増幅された核酸、およびこれらの断片の質量を変化させ得る。他のヌクレオシドは、ポリヌクレオチドの機能的性質（特に限定されないが、断片化に対するポリヌクレオチドの感受性の上昇、さらにポリヌクレオチドを伸長させる能力の低下がある）を変化させ得る。修飾ヌクレオチドは必ずしも天然に存在しないものではなく、典型的には特定のポリヌクレオチド中に取り込まれないヌクレオシドである（例えば、DNAが生成される時、A、C、T、およびG以外のヌクレオシド、またはRNAが生成される時、A、C、U、およびG以外のヌクレオシド）。

30

【0106】

ある実施態様において標的核酸は、天然に存在するが標的核酸の通常の前駆体ではないヌクレオシド3リン酸を使用して増幅される。例えば1つのrNTPと3つのdNTPを増幅ポリヌクレオチド中に取り込むことができる（例えば、rCTP、dATP、dTTP、およびdGTP）。別の例では、正常なDNA前駆体ヌクレオチド（例えば、dCTP、dATP、およびdGTP）とdUTPの存在下でDNAを増幅することにより、通常はDNA中に存在しないデオキシウリジン3リン酸を増幅DNA分子中に取り込むことができる。DNAへのかかるウリジンの取り込みは、DNAの塩基特異的切断を促進することができる。例えば増幅されたウリジン含有DNAをウラシル-DNAグリコシラーゼ（UDG）で処理すると、ウラシル残基が切断される。UDG反応からの生成物の以後の化学処理によりリン酸骨格が切断され、ヌクレオ塩基特異的断片が生成される。さらにグリコシラーゼ処理の前に増幅産物の相補鎖を分離すると、断片化の相補的パターンが作製される。すなわちdUTPとウラシルDNAグリコシラーゼの使用は、相補鎖のためのT特異的断片の生成を可能にし、ある配列内のT位置ならびにA位置に関する情報を与える。

40

【0107】

増幅または他のヌクレオチド合成反応（例えば転写）は、伸長を停止させるように機能するヌクレオチド類似体（例えばジデオキシヌクレオチド）を使用して行うことができる

50

。ある実施態様において反応条件は、典型的にはジデオキシヌクレオチド型でオリゴヌクレオチド中に取り込まれる4つのヌクレオチドモノマーのうちの1つを含有する。他の実施態様において反応条件は、ジデオキシヌクレオチド型のヌクレオチドモノマーの4つのうちの2つ、4つのうちの3つ、または4つすべてを含有する。反応条件は、リボヌクレオチド、デオキシヌクレオチド、および/またはジデオキシリボヌクレオチド型の特定のヌクレオチドモノマーの可能な混合物を含有する。例えばアデノシン(A)は、反応混合物中で10%リボヌクレオチド、80%デオキシヌクレオチド、および10%ジデオキシヌクレオチド型で存在することができる。増幅または他の反応(例えば転写)は、完了させる必要はない。例えばPCR中の増幅工程は、すべてのプライマーが完全に伸長される前に停止させて、種々の異なる長さの標的断片核酸を得ることができる。すなわちある実施態様において反応は、伸長中に異なる位置で停止されたオリゴヌクレオチドである標的核酸の不均一プールを与えるように行うことができる。

10

【0108】

ある実施態様において1つまたはそれ以上のヌクレオシド3リン酸を、ヌクレオチド間に選択的非加水分解性結合を作製する類似体で置換することができる。例えばヌクレオチドを-チオ基質で置換し、次にホスホロチオエートヌクレオシド間結合をハロゲン化アルキル(例えば、ヨードアセトアミド、ヨードエタノール)または2,3-エポキシ-1-プロパノールのような試薬を使用してアルキル化して修飾することができる。選択的非加水分解性であるヌクレオシドの他の例には、2'フルオロヌクレオシド、2'デオキシヌクレオシド、および2'アミノヌクレオシドがある。

20

【0109】

質量修飾ヌクレオシドは、質量修飾デオキシヌクレオシド3リン酸、質量修飾ジデオキシヌクレオシド3リン酸、および質量修飾リボヌクレオシド3リン酸から選択することができる。質量修飾ヌクレオシド3リン酸は、塩基、糖、および/またはリン酸残基で修飾ことができ、酵素的工程、化学的工程、またはこれらの組合せにより導入することができる。ある態様において修飾はヒドロキシル基以外の2'置換基を含む。別の態様においてヌクレオシド間の結合はさらに修飾できる(例えばホスホロチオエート結合、またはアルキル化剤でさらに反応させたホスホロチオエート結合)。さらに別の態様において修飾ヌクレオシド3リン酸は、メチル基、例えば5-メチルシトシンまたは5-メチルウリジンで修飾することができる。

30

【0110】

他の公知の質量修飾残基には、ハロゲン、例えばF、Cl、Br、および/またはI、または擬ハロゲン、例えばSCN、NCS、または異なるアルキル、アリールもしくはアラルキル残基、例えばメチル、エチル、プロピル、イソプロピル、t-ブチル、ヘキシル、フェニル、置換フェニル、ベンジル、または官能基、例えば CH_2 、F、 CHF_2 、 CF_3 、 $\text{Si}(\text{CH}_3)_3$ 、 $\text{Si}(\text{CH}_3)_2(\text{C}_2\text{H}_5)$ 、 $\text{Si}(\text{CH}_3)(\text{C}_2\text{H}_5)_2$ 、 $\text{Si}(\text{C}_2\text{H}_5)_3$ の、Hによる置換がある。さらに別の質量修飾は、核酸分子(例えば、検出物質(D))またはヌクレオシド3リン酸を介してホモ-またはヘテロペプチドを結合させることにより得られる。質量増分57を有する質量修飾種を作製するのに有用な1つの例は、オリゴグリシンの結合であり、例えば74($r=1$ 、 $m=0$)、131($r=1$ 、 $m=2$)、188($r=1$ 、 $m=3$)、245($r=1$ 、 $m=4$)の質量修飾が達成される。単純なオリゴアミドもまた使用でき、例えば74($r=1$ 、 $m=0$)、88($r=2$ 、 $m=0$)、102($r=3$ 、 $m=0$)、116($r=4$ 、 $m=0$)などの質量修飾が得られる。

40

【0111】

質量修飾残基は、例えばオリゴヌクレオチドの5'末端に、ヌクレオ塩基(または塩基)に、リン酸骨格に、ヌクレオシドの2'位置に、および/または末端3'位置に結合することができる。質量修飾残基の例には、例えばハロゲン、アジド、またはXR型(ここでXは結合基であり、Rは質量修飾官能基である)がある。質量修飾官能基は、例えば上記したようにオリゴヌクレオチド分子に一定の質量増分を導入するのに使用することができる。アルファ-チオヌクレオシド3リン酸のようなホスホジエステル結合で導入した修飾は、これらの修飾が正確なワトソン-クリック塩基対合を妨害せず、さらに例えばアルキル化反

50

応を介する完全な核酸分子の1工程の合成後部位特異的修飾を可能にするという利点を有する(例えば、Nakamayeら、Nucl. Acids Res. 16: 9947-9959 (1988))。質量修飾官能基の例はホウ素修飾核酸であり、これはポリメラーゼにより核酸中に効率的に取り込まれる(例えば、Porterら、Biochemistry 34: 11963-11969 (1995); Hasanら、Nucl. Acids Res. 24: 2150-2157 (1996); Liら、Nucl. Acids Res. 23: 4495-4501 (1995)を参照)。

【0112】

さらに質量修飾官能基を加えて、ヌクレオシド3リン酸の糖環の3'位置に結合させることにより鎖の停止に影響を与えることができる。本明細書に記載の方法において多くの組合せを使用できることは、当業者に明らかであろう。同じように当業者は、鎖伸長性ヌクレオシド3リン酸もまた、官能基と結合位置に多くの変化と組合せをもって同様に質量修飾されることを認識するであろう。

10

【0113】

種々の異なる核酸断片を同時に検出するために、異なる質量修飾ヌクレオチドを使用することができる。ある実施態様において質量修飾は、増幅プロセス中に取り込まれる。別の実施態様において、異なる標的核酸分子多重性が、1つまたはそれ以上の標的核酸分子を質量修飾することにより行われ、ここで異なる標的核酸分子は、所望であれば示差的に質量修飾することができる。

【0114】

c. 増幅法

増幅法は、所望のアッセイデザインに従って、種々の異なる増幅産物を作製するために使用することができる。

20

【0115】

ある実施態様において増幅反応または他の反応(例えば転写)のヌクレオチド生成物が提供され、生成物ヌクレオチドは、単一の鋳型が提供される時でさえ、サイズが異なる。例えば未変性の標的核酸からの1つまたはそれ以上のヌクレオチド位置が2つまたはそれ以上の生成物ヌクレオチドの間で共通であるように、生成物ヌクレオチドは重複してもよい。かかる重複ヌクレオチドは「ラダー(ladder)」ヌクレオチドを含み、異なるサイズの一連のヌクレオチドが同じコア配列を含み、連続的に大きくなるヌクレオチドが、典型的にはヌクレオチドの3'末端または5'末端のみで1つまたはそれ以上の核酸位置の増分で追加のヌクレオチドを含有する。かかる生成物を生成するのに種々の方法を使用することができ、特に限定されないが、ジデオキシと非ジデオキシヌクレオシドの組合せ中に存在している4つのヌクレオシドの1つとの核酸合成反応がある。

30

【0116】

ある実施態様において増幅反応または他のヌクレオチド合成反応は、鋳型標的核酸または鋳型標的核酸断片中の定常領域と可変領域の両方にハイブリダイズする1つまたはそれ以上のプライマーを使用して実施される。例えば標的核酸分子は本明細書に開示の方法を使用して断片化することができ、かかる標的核酸断片はそこに1つまたはそれ以上のアダプターオリゴヌクレオチドが結合でき、こうして同じ配列を有するアダプターオリゴヌクレオチドが、異なる配列を有する2つまたはそれ以上の標的核酸断片の同じ末端(すなわち3'末端または5'末端)に結合される。各結合生成物は、標的核酸断片とアダプターオリゴヌクレオチドの両方を含有する。標的核酸断片の部分は断片毎に異なるため、プライマーはアダプターオリゴヌクレオチド領域の少なくとも一部およびいくつかの少なくとも一部にハイブリダイズすることにより、一部(すべての結合生成物ではない)にハイブリダイズすることができる。次に、結合した断片の可変領域中のプライマーとハイブリダイズする標的核酸断片のサブセットについて、増幅反応または他のヌクレオチド合成反応が行われる。こうして、1つまたはそれ以上のプライマーのセットを使用してすべての標的核酸断片の亜集団を増幅することができ、これに従って標的核酸断片の可変配列はプライマーとハイブリダイズする。ある実施態様において標的核酸断片の3'末端のみ、5'末端のみ、または3'末端と5'末端の両方に連結するのに1つのプライマー配列のみが使用される。別の実施態様では、標的核酸断片に連結するのに2つのプライマーが使用され、第1のもの

40

50

は3'標的核酸断片末端に、そして第2のものは5'標的核酸断片末端に連結される。別の実施態様では、3'末端または5'末端に連結するのに2つまたはそれ以上のプライマーが使用される。例えば異なる定常領域を認識する複数のプライマーを使用して、第1のセットのプライマーが標的核酸断片の第1の集団にハイブリダイズし、第2のセットのプライマーが標的核酸断片の第2の集団にハイブリダイズするようにすることができ、典型的には標的核酸の第1の集団と第2の集団は重複するメンバーを持たない。

【0117】

選択的ヌクレオチド合成はまた、断片化と組合せて行われる。複数の核酸合成サイクルを介して増幅される標的核酸は、標的核酸分子の2つの異なる領域にハイブリダイズするプライマーを使用する。2つのプライマーハイブリダイゼーション部位の間の中央領域中の標的核酸分子の断片化は、標的核酸分子の増幅を妨害する。従って核酸分子の中央領域の選択的断片化は、核酸合成反応で使用されるプライマーが選択的ではなくてもまたは高度に選択的であっても、標的核酸分子の選択的増幅を与える。

10

【0118】

ある例では試料は、核酸合成条件で処理される前に、断片化条件で処理することができる。そのような例では断片化条件は、特定のヌクレオチド配列を選択的に切断する。例えば試料はそこに制限エンドヌクレアーゼ（例えばEcoRI）が付加されてよい。これにより、EcoRI認識部位を含有する切断された標的核酸分子と、EcoRI認識部位を含有しない無傷の標的核酸分子とを含有する試料を与える。次に試料は、非切断標的核酸分子のみが増幅されるように設計されたプライマーを使用して、核酸合成条件で処理することができる。切断の結果、制限エンドヌクレアーゼ認識部位の存在に従って、増幅は標的核酸分子のサブセットに対して選択的になる。本明細書に記載の方法で使用される断片化条件は、核酸分子を選択的に切断する任意の断片化条件（制限エンドヌクレアーゼを含む）を含む。使用可能な追加の断片化条件には、配列特異性により切断することができる任意の断片化条件がある。

20

【0119】

別の実施態様において、唯一の核酸増幅法としてまたは他の核酸増幅法以外に、転写を行うことができる。RNA分子を生成するのに鋳型DNA分子を使用する転写法は、標的核酸分子を増幅し、標的核酸分子をDNA型からRNA型に修飾するように機能する。鋳型DNAの例には、増幅された生成物標的核酸分子、および処理された非増幅標的核酸分子がある。

30

【0120】

本明細書に記載されるように、処理された標的核酸分子は1回またはそれ以上の核酸合成反応を受ける。核酸合成反応は、処理された標的核酸分子を増幅し、核酸分子の型を修飾するように機能する。ある例では処理された標的核酸分子またはPCR産物は転写される。

【0121】

鋳型DNA（例えば標的核酸分子またはその増幅産物）の転写は、鋳型DNAの片方の鎖または両方の鎖について行うことができる。ある例では転写される核酸分子は、転写を行うことができる酵素が結合する残基を含有し、かかる残基は例えば転写プロモーター配列でもよい。

40

【0122】

転写反応は、当該分野で公知の多くの任意の方法を使用して、当該分野で公知の多くの任意の酵素を使用して行われる。例えばdNTPとrNTPの両方を取り込む能力を有する変異T7 RNAポリメラーゼ（T7 R&DNAポリメラーゼ；エピセンター（Epicentre）、マジソン、ウィスコンシン州）を、転写反応に使用することができる。転写反応は、当該分野で公知の標準的反応条件（例えば40 mM トリス酢酸（pH 7.5）、10 mM NaCl、6 mM MgCl₂、2 mM スペルミジン、10 mM ジチオスレイトール、各1 mMのrNTP、5 mMのdNTP（使用される時）、40 nM DNA鋳型、および5 U/μl T7 R&DNAポリメラーゼ、37 °Cで2時間インキュベート）下で行われる。転写後、エピアルカリホスファターゼ（shrimp alkaline phosphatase、SAP）を切断反応物に加えて、環状1リン酸副産物の量を減少させることができる。T7 R&DN

50

Aポリメラーゼの使用は当該分野で公知であり、例えば米国特許第5,849,546号明細書、同6,107,037号明細書、およびSousaら、EMBO J. 14: 4609-4621 (1995)、Padillaら、Nucl. Acids Res. 27: 1561-1563 (1999)、Huangら、Biochemistry 36: 8231-8242 (1997)、およびStanssensら、Genome Res. 14:1 26-133 (2004)に例示されている。

【0123】

4つの通常のリボヌクレオチド基質 (rCTP、rATP、rGTP、およびrUTP) による転写以外に、1つまたはそれ以上のリボヌクレオチド3リン酸をヌクレオチド類似体 (例えば、本明細書に記載のもので、当該分野で公知のもの) または対応するデオキシリボヌクレオチド3リン酸で置換 (例えば、rCTPをdCTPで、またはrUTPをdUTPもしくはdTTPで置換) して反応を行うことができる。ある例では1つまたはそれ以上のrNTPはヌクレオチドもしくはヌクレオチド類似体で置換され、これは転写核酸に取り込まれると、転写核酸に適用される断片化条件下では切断できない。

10

【0124】

ある例では転写は、1つまたはそれ以上の核酸合成反応後に行われる。例えば増幅産物の転写は、標的核酸分子の増幅後に行われる。別の実施態様において処理された標的核酸分子は、あらかじめ核酸合成工程に付されることなく転写される。

【0125】

ある方法では核酸を含む反応はまた、2本鎖核酸が変性されて1本鎖分子を与える工程を含むことができる。変性は例えば、反応混合物の温度が特定の2本鎖核酸の融解温度を超える条件下で行われる。

20

【0126】

多くの核酸反応 (例えば増幅反応) は温度上昇と低下の繰り返しサイクルを含み、核酸ハイブリッド鎖の変性とアニーリングを与える。出願番号第60/372,711号 (2002年4月11日出願)、第60/457,847号 (2003年3月24日出願)、および第10/412,801号 (2003年4月11日出願) に記載の装置は、比較的低質量の直接的で迅速かつ効率的な加熱と冷却、およびチャンパーの固体支持体である底の高い熱伝導性を介して、および反応物を別のサーマルサイクラー装置に移す工程を避けることにより、チャンパー中の反応混合物の温度の変化を促進する。

【0127】

D. 断片化

公知の方法を使用して充分量の標的核酸がいったん作製されると、標的核酸配列は核酸断片に切断することができる。核酸断片を作製するのに、核酸分子を断片に切断する種々の方法を使用して核酸断片を生成することができる。ある場合には断片化法は、適当な断片サイズ分布を与える。ポリヌクレオチドの断片化は当該分野で公知であり、多くの方法で行われる。例えばDNA、RNA、DNAとRNAの類似体、またはこれらの組合せからなるポリヌクレオチドは、物理的、化学的、または酵素的に断片化することができる。ある例では物理的断片化を使用して、種々のサイズのランダムな標的核酸断片が生成される。別の例では、1つまたはそれ以上の特異的および/または非特異的切断部位での部分的酵素的切断を使用して、本明細書で使用されるランダムな標的核酸断片が生成される。

30

【0128】

具体的な実施態様においてサイズが統計的に5~50塩基、10~40塩基、11~35塩基、および12~30塩基の範囲になるように、標的核酸の断片が調製される。別の実施態様において、例えば質量スペクトル分析の前に捕獲オリゴヌクレオチド：標的断片複合体を「トリム」することが企図される場合は、標的核酸の断片はかなり大きく、サイズが統計的に20~50塩基、30~60塩基、40~70塩基、50~80塩基、60~90塩基、70~100塩基、およびそれ以上の範囲になる。本発明で使用される他のサイズ範囲には、約50~約150塩基、約25~約75塩基、または約12~30塩基がある。ある具体的な実施態様では、約12~約30塩基の断片が使用される。一般に断片のサイズ範囲は、より短い断片が捕獲オリゴヌクレオチドに充分強く結合し、充分な特異性でハイブリダイズし、かつより長い断片が充分効率的にハイブリダイズして不十分にならないように、選択される。またある実施態様においてサ

40

50

イズ範囲は、MALDI-TOF MSの所望の脱着効率を促進するように選択される。

【0129】

断片サイズ長さと断片サイズの範囲は、本明細書に記載の異なる断片化法の任意の方法により行われる。例えば物理的断片化法が使用される時、物理的力/ひずみを負荷するパラメータを調整すると、異なる断片サイズと範囲が得られる。別の例では制限酵素が使用される時、使用される制限酵素の数と種類および選択される具体的な反応条件を使用して、生成される断片の平均長さを制御することができる。断片のサイズが変動してもよく、本発明で使用される適当な断片は典型的には、約500ヌクレオチド未満、約400ヌクレオチド未満、約300ヌクレオチド未満、約200ヌクレオチド未満の長さである。

【0130】

統計的に重複する断片のプールにおいて断片は他の断片と重複する：例えば重複断片は、1個またはそれ以上、2個またはそれ以上、3個またはそれ以上、4個またはそれ以上、5個またはそれ以上、6個またはそれ以上、8個またはそれ以上、10個またはそれ以上、15個またはそれ以上、20個またはそれ以上の他の断片と重複することができ、典型的には少なくとも2個、少なくとも3個、少なくとも4個、少なくとも5個、少なくとも6個、少なくとも8個、少なくとも10個、少なくとも15個、または少なくとも20個の他の断片と重複することができる。

【0131】

重複断片はまた、非断片化標的核酸分子から1つまたはそれ以上のヌクレオチド位置を共通に有する断片である。すなわち重複断片は、第1の断片が第2の断片中に位置するすべてのヌクレオチド位置を含有し、さらに第1の断片は、第1の断片の5'末端、3'末端、または5'末端と3'末端の両方に追加のヌクレオチド位置を含有する断片を含む。重複断片はまた、第1の断片の3'末端が第2の断片の5'末端と重複する断片も含む。重複断片は1つのヌクレオチド位置でのみ重複する必要がある；しかし、統計的に重複する断片のプールは、少なくとも2、少なくとも3、少なくとも4、少なくとも5、少なくとも6、少なくとも8、少なくとも10、少なくとも15、または少なくとも20のヌクレオチド位置で重複することができる。

【0132】

1. ポリヌクレオチドの酵素的断片化

核酸分子断片は、1本鎖のまたは複数鎖の核酸分子の酵素的切断により得られる。複数鎖核酸分子は、核酸分子の2つ以上の鎖（例えば、2本鎖、3本鎖の核酸分子を含む）を含有する核酸分子複合体を含む。使用される酵素に依存して、核酸分子は非特異的に切断されるかまたは特異的なヌクレオチド配列で切断される。核酸分子を切断できる任意の酵素を使用することができ、特に限定されないが、エンドヌクレアーゼ、エキソヌクレアーゼ、1本鎖特異的ヌクレアーゼ、2本鎖特異的ヌクレアーゼ、リボザイム、およびDNAザイムがある。核酸分子を断片化するための種々の酵素が当該分野で公知であり、市販されており、例えばヌクレアーゼBAL-31、大豆ヌクレアーゼ、エキソヌクレアーゼI、エキソヌクレアーゼIII、エキソヌクレアーゼVIII、ラムダエキソヌクレアーゼ、T7エキソヌクレアーゼ、エキソヌクレアーゼT、RecJ、RNase、RNaseIII、RNaseA、RNaseU2、RNaseT1、RNaseHショートカットRNaseIII、AccI、BsaAI、BtgZI、MfeI、SacI、N.BbvCIA、N.BbvCIB、N.BstNBI、I-CeuI、I-SceI、PI-PspI、PI-SceI、McrBC、および他の公知の酵素がある（例えば、ニューイングランドバイオラボズ社（New England Biolabs Inc.）カタログ；Sambrook, J., Russell, D.W., 「分子クローニング：実験室マニュアル（Molecular Cloning, A Laboratory Manual）」、第3版、コールドスプリングハーバーラボラトリープレス（Cold Spring Harbor Laboratory Press）、コールドスプリングハーバー（Cold Spring Harbor）、ニューヨーク、2001を参照）。酵素はまた、大きな核酸分子をより小さい断片に分解するのに使用することができる。本明細書で提供される酵素は、単独でまたは組合せて使用して重複標的核酸断片を生成することができる。重複断片の作製は、多くの異なる方法により行うことができる。例えば非特異的RNase（RNaseI）または非特異的DNase（DNaseI）による限定/部分的消化を使用することができる。

10

20

30

40

50

【 0 1 3 3 】

a. エンドヌクレアーゼ断片化

エンドヌクレアーゼは、核酸分子を断片化するのに有用な酵素群の例である。エンドヌクレアーゼは核酸分子鎖内の結合を切断する。エンドヌクレアーゼは、2本鎖または1本鎖核酸分子のいずれかに特異的であり得る。切断は、核酸分子内でランダムにまたは特異的配列で起きる。2本鎖核酸分子をランダムに切断するエンドヌクレアーゼは、核酸分子の骨格としばしば相互作用する。核酸分子の特異的断片化は、連続反応でまたは同時反応で1つまたはそれ以上の酵素を使用して行われる。均一または不均一な核酸分子が切断される。エンドヌクレアーゼはまた1本鎖核酸を切断することができる：例えば、S1または大豆ヌクレアーゼは1本鎖DNA（大豆）またはDNAもしくはRNA（S1）を分解して平滑末端2本鎖核酸分子を生成する。

10

【 0 1 3 4 】

制限エンドヌクレアーゼは、2本鎖核酸分子内の特異的配列を認識するエンドヌクレアーゼのサブクラスであり、典型的には認識配列内でもまたはその近くで両方の鎖を切断する。DNA分析で一般的に使用される1つの酵素はHaeIIIであり、これは配列5'-GGCC-3'でDNAを切断する。他の制限エンドヌクレアーゼには、AccI、AflIII、AluI、Alw44I、ApaI、AsnI、AvaI、AvaII、BamHI、BanII、BclI、BglI、BglII、BlnI、BsmI、BssHII、BstEII、CfoI、ClaI、DdeI、DpnI、DraI、EclXI、EcoRI、EcoRII、EcoRV、HaeII、HaeIII、HindII、HindIII、HpaI、HpaII、KpnI、KspI、MluI、MluNI、MspI、NciI、NcoI、NdeI、NdeII、NheI、NotI、NruI、NsiI、PstI、PvuI、PvuII、RsaI、SacI、SalI、Sau3AI、Scal、

20

【 0 1 3 5 】

使用される酵素に依存して核酸分子の切断により、「粘着」末端とも呼ばれる他方をオーバーハングする1つの鎖が得られる。例えばBamHIは粘着5'オーバーハング末端を生成し、KpnIは粘着3'オーバーハング末端を生成する。あるいは切断により、オーバーハング末端を持たない「平滑」末端が得られる。例えばDraI切断は平滑末端を生成する。制限酵素は特定のヌクレオチド配列を含有する核酸分子を切断することができ、そのヌクレオチド配列を含有しない核酸分子は切断しない。ある場合には、切断認識部位はメチル化によりマスクすることができる。

30

【 0 1 3 6 】

制限エンドヌクレアーゼは、種々の核酸分子断片サイズを生成するのに使用することができる。例えばCviJIは、2塩基と3塩基DNA配列とを認識する制限エンドヌクレアーゼである。CviJIによる完全な消化により、平均16~64ヌクレオチド長のDNA断片が得られる。従ってCviJIによる部分的消化は、剪断または超音波処理と同様の「準」ランダム法で断片DNAを与える。CviJIは通常、RGCY部位をGとCの間で切断して、容易に切断可能な平滑末端を与える（ここでRは任意のプリンでありYは任意のピリミジンである）。1mM ATPと20%ジメチルスルホキシドの存在下では、切断の特異性が緩和され、CviJIはRGCNとYGCY部位も切断する。これらの「スター（star）」条件下ではCviJI切断は、準ランダム消化物を生成する。消化または剪断されたDNAは、この時点でサイズ選択される。

40

【 0 1 3 7 】

制限エンドヌクレアーゼを使用して核酸分子を断片化する方法は当該分野で公知である。あるプロトコル例では以下を含有する20~50 μ lの反応混合物が調製される：DNA 1~3 μ g；制限酵素緩衝液 1x；および1 μ gのDNAについて制限エンドヌクレアーゼ 2単位。適当な緩衝液はまた当該分野で公知であり、酵素活性の最適条件を提供するために適当なイオン強度、補助因子、および場合によりpH緩衝液を含む。特異的酵素は、一般に酵素の供給業者から入手できる特異的緩衝液を必要とする。緩衝液の例は、グルタミン酸カリウム緩衝液（KGB）である。Hannish, J.とM. McClelland、「KGB、グルタミン酸カリウム緩衝液中のDNA修飾と制限酵素の活性」、Gene Anal. Tech 5: 105 (1988)；McClelland, M.

50

ら、「すべての制限エンドヌクレアーゼのための1つの緩衝液」、Nucl. Acids Res. 16: 364 (1988)。反応混合物は、37 で1時間または所望のサイズまたはサイズ範囲の断片を産生するのに必要な期間インキュベートされる。反応は、混合物を必要に応じて65 または80 で加熱することにより停止することができる。あるいは2価陽イオン（例えば Mg^{2+} ）を例えばEDTAでキレート化させて反応を停止することができる。

【0138】

具体的な実施態様において、核酸分子を断片化するために2つ以上の酵素が使用できる。複数の酵素が同様の条件（例えばイオン強度、温度、またはpH）下で活性なら複数の酵素を同じ反応で使用できるか、または複数の酵素を連続反応で使用できる。典型的には複数の酵素は、KGBのような標準的緩衝液とともに使用される。制限酵素が使用される時、核酸分子は部分的にまたは完全に消化することができる。

10

【0139】

DNaseはまた核酸分子断片を生成するのに使用することができる。Anderson, S.、「クローン化したDNaseI 作製断片を使用する散弾銃DNA配列決定」、Nucl. Acids Res. 9: 3015-3027 (1981)。DNaseI（デオキシリボヌクレアーゼI）は、2本鎖および1本鎖DNAを非特異的に消化してポリ-およびモノ-ヌクレオチドにするエンドヌクレアーゼである。この酵素はまた、1本鎖ならびに2本鎖DNAに、およびクロマチンに作用することができる。

【0140】

デオキシリボヌクレアーゼII型は、核酸研究の多くの用途（酸性pHでのDNA配列決定と消化）で使用される。ブタ脾臓からのデオキシリボヌクレアーゼIIは、未変性DNAおよび変性DNA中のデオキシリボヌクレオチド結合を加水分解して、3'-リン酸塩を有する生成物を与える。これはまた、pH5.6~5.9でp-ニトロフェニルホスホジエステルに作用する。Ehrlich, S.D.ら、「酸性デオキシリボヌクレアーゼの研究。IX. ウシ胸腺デオキシリボ核酸から得られるオリゴヌクレオチドの5'-ヒドロキシ末端および末位から二番目のヌクレオチド」、Biochemistry 10(11): 2000-2009 (1971)。

20

【0141】

エンドヌクレアーゼは、特定のタイプの核酸分子に特異的となり得る。例えばエンドヌクレアーゼは、DNAもしくはRNA、または1本鎖もしくは2本鎖核酸分子に特異的である。エンドヌクレアーゼは配列特異的かまたは非配列特異的である。例えばリボヌクレアーゼHは、RNA-DNAハイブリッド中のRNA鎖を特異的に分解する。リボヌクレアーゼAは、CおよびU残基で1本鎖RNAを特異的に攻撃するエンドヌクレアーゼである。リボヌクレアーゼAは、ヌクレオチドの5'リボースと隣接ピリミジンヌクレオチドの3'リボースに結合したリン酸塩基とのホスホジエステル結合の切断を触媒する。得られる2',3'-環状リン酸塩は加水分解されて、対応する3'ヌクレオシドリン酸塩となる。RNase T1はGリボヌクレオチドでのみRNAを消化して、グアニル残基の3'ヒドロキシル基と隣接するヌクレオチドの5'ヒドロキシル基との間を切断する。RNase U₂はAリボヌクレオチドでのみRNAを消化する。塩基特異的消化の例は、Stanssensら、WO00/66771による刊行物中に見られる。

30

【0142】

ベンゾナーゼJ、ヌクレアーゼP1、およびホスホジエステラーゼIは、200塩基対またはそれ以下の核酸分子断片を作製するのに適した非特異的エンドヌクレアーゼである。ベンゾナーゼJ（ノバゲン（Novagen）、マジソン、ウィスコンシン州）は、すべての型のDNAとRNA（1本鎖、2本鎖、線状、および環状）を分解する遺伝子工学的に作製したエンドヌクレアーゼであり、広範囲の操作条件下で使用することができる。この酵素は、核酸を長さが2~5塩基の5'-1リン酸末端オリゴヌクレオチドに完全に消化する。ベンゾナーゼJのヌクレオチド配列とアミノ酸配列は、米国特許第5,173,418号明細書に提供されている。本明細書に記載の方法の核酸の断片化はまた、ジヌクレオチド（「2カッター」）または緩和ジヌクレオチド（「1-1/2カッター」または「1-1/4カッター」）切断特異性により行われる。ジヌクレオチド特異的切断試薬は当業者に公知である（例えばWO94/21663号；Cannistraroら、Eur. J. Biochem. 181: 363-370 (1989)；Stevensら、J. Bacteriol. 164: 57-62 (1985)；Marottaら、Biochemistry 12: 2901-2904 (1973)を参照）。

40

50

【0143】

制限エンドヌクレアーゼを使用する切断は部分的に行われるか、または制限エンドヌクレアーゼ認識部位にランダムに取り込まれる修飾ヌクレオチドを使用して修飾される。これらの修飾ヌクレオチドは、標準的ヌクレオチドと比較して切断に対する異なる感受性を示す。この異なる感受性は、切断され易さの上昇、および切断され易さの低下（切断に対する完全な抵抗を含む）を含む。例えば、酵素的切断に対して耐性であるデアザヌクレオチドは、部分的にかつランダムに制限エンドヌクレアーゼの認識部位中に取り込まれ、これは、制限エンドヌクレアーゼ反応が完了しても、部分的切断を与える。別の例ではデオキシウリジンはDNAヌクレオチド中に取り込まれ、ウラシル-DNAグリコシラーゼを使用してウラシルを除去し、次にこの位置でDNAを切断することができる：すなわちウリジンのDNAへの取り込みは、切断され易さの上昇を示す。別の例では目的の標的核酸分子の転写体を標準的 - チオ基質の混合物を用いて合成することができ、次にホスホロチオエートヌクレオチド間結合が、ハロゲン化アルキル（例えば、ヨードアセトアミド、ヨードエタノール）または2,3-エポキシ-1-プロパノールのような試薬を使用してアルキル化により修飾される。かかる修飾により生成されるホスホチオエステル結合はRNaseの基質になることは期待できない。RNaseにより切断されない他のヌクレオチドの例には、2'フルオロヌクレオチド、2'デオキシヌクレオチド、および2'アミノヌクレオチドがある。この方法を使用する1つの例では、所望の切断特異性に依存して、CヌクレオチドまたはUヌクレオチドの2'修飾型のような非加水分解ヌクレオチドの取り込みを介して、CpNまたはUpNジヌクレオチドに限定することができる。すなわちある例では、S-dUPT、S-dAPT、S-dCPT、およびGPTヌクレオチドを転写体中に取り込むことにより、転写体（標的分子）が調製できる。有用なジヌクレオチド特異的切断試薬の範囲は、追加のRNase、例えばRNase-U2およびRNase-T1を使用することによりさらに拡張することができる。モノ特異的RNase（例えばRNase-T1）の場合、切断不可能なヌクレオチドの使用は、切断不可能であることが選択されるヌクレオチドに依存して、4つの可能なGpN結合の内の任意の3つ、2つ、または1つに結合するCpN結合の切断を限定する。これらの選択的修飾法はまた、ホモポリマー区画内のヌクレオチドの一部を選択的に修飾することにより、ホモポリマー区画のすべての塩基で切断を妨害して、修飾ヌクレオチドを切断に対して抵抗性を小さくまたは抵抗性を大きくするのに使用することができる。

10

20

30

【0144】

b. エキソヌクレアーゼ断片化

ポリヌクレオチドの末端から種々の長さの塩基を除去するエキソヌクレアーゼ（エキソヌクレアーゼと呼ぶ）を使用して、ポリヌクレオチドを小さいポリヌクレオチドに断片化することができる。エキソヌクレアーゼは2本鎖核酸を断片化するかまたは1本鎖核酸を断片化することができる。1本鎖または2本鎖核酸を断片化できるエキソヌクレアーゼの例は、Ba131ヌクレアーゼである。

【0145】

エキソヌクレアーゼは種々のポリヌクレオチドの末端からヌクレオチドを切断することができる。例えば5'エキソヌクレアーゼ（DNA鎖の5'末端からDNAを切断する）と3'エキソヌクレアーゼ（DNA鎖の3'末端からDNAを切断する）がある。異なるエキソヌクレアーゼは、1本鎖または2本鎖DNAを加水分解することができる。例えばエキソヌクレアーゼIIIは、3'から5'へのエキソヌクレアーゼであり、DNA鎖の3'末端から5'モノヌクレオチドを放出する；これはDNA 3'ホスファターゼであり、3'末端ホスホモノエステルを加水分解する；そしてこれはAPエンドヌクレアーゼであり、アプリン部位またはアピリミジン部位でホスホジエステル結合を切断して、塩基を含まないデオキシリボースである5'ホスファターゼ残基である5'末端を生成する。さらにこの酵素はRNaseH活性を有する；これは、DNA-RNAハイブリッド2本鎖中のRNA鎖を、おそらくエキソヌクレアーゼ的切断により選択的に分解する。S1哺乳動物細胞では、主要なDNA 3'エキソヌクレアーゼはDNaseIII（TREX-1とも呼ばれる）である。すなわちエキソヌクレアーゼを使用してポリヌクレオチドの末端を分解することにより、断片が生成される。

40

50

【0146】

c. 核酸酵素断片化

触媒性DNAおよびRNAは当該分野で公知であり、核酸分子を切断して核酸分子断片を生成するのに使用することができる。Santoro, S.W.とJoyce, G.F.、「多目的RNA切断性DNA酵素」、Proc. Natl. Acad. Sci. USA 94: 4262-4266 (1997)。1本鎖分子としてのDNAはRNAに似た3次元構造に折り畳むことができ、2'ヒドロキシ基は触媒作用のために必須である。リボザイムとしてDNAザイムはまた、選択により補助因子に依存するように作製することができる。これは、RNA加水分解のためのヒスチジン依存性DNAザイムについて証明されている。米国特許第6,326,174号明細書および同6,194,180号明細書は、核酸配列または分子、特にRNA、を切断することができるデオキシリボ核酸酵素、触媒性および酵素性DNA分子を開示する。

10

【0147】

核酸分子を切断するためのリボザイムの使用は当該分野で公知である。リボザイムは、化学反応、例えば共有結合の切断を触媒するRNAである。Uhlenbeckは、触媒性および基質鎖が分離された小さな活性リボザイムであるハンマーヘッドリボザイムを証明した (Uhlenbeck, Nature 328: 596-600 (1987))。かかるリボザイムは塩基対合相互作用により基質RNAに結合し、結合した標的RNAを切断し、切断産物を放出し、リサイクルされて、このプロセスを複数回繰り返すことができる。HaseloffとGerlachは、トランスで作用できる単純なハンマーヘッドリボザイムの一般的設計規則を列挙した (Haseloffら、Nature 334: 585-591 (1988))。高い切断特異性を有する種々の異なるハンマーヘッドリボザイムが開発されており、所望の基質特異性を有するハンマーヘッドリボザイムの設計のための一般的アプローチは、米国特許第5,646,020号明細書および同6,096,715号明細書に例示されるように当該分野で公知である。トランス切断活性を有する他のタイプのリボザイムは、肝炎ウイルスのゲノムから得られたリボザイムである。AnanvoranichとPerraultは、リボザイム切断の基質特異性の要因を記載した (Ananvoranichら、J. Biol. Chem. 273: 13812-13188 (1998))。ヘアピンリボザイムはまた、トランス切断に使用でき、ヘアピンリボザイムの基質特異性の原理もまた公知である (例えばPerez-Ruizら、J. Biol. Chem. 274: 29376-29380 (1999))。当業者は基質特異性の公知の原理を使用してリボザイムを選択しリボザイム配列を設計して、所望の核酸分子切断特異性を達成することができる。

20

30

【0148】

DNAニッカーゼまたはDNaseは、DNA 2本鎖の1つの鎖を認識し切断するのに使用することができる。多くのニッカーゼが公知である。例えば特に、以下の切断部位を有するNY2AニッカーゼとNYS1ニッカーゼ (メガベース (Megabase)) がある：

NY2A: 5'...R AG...3'
 3'...Y TC...5' ここで RはAまたはG、かつYはCまたはT
 NYS1 : 5... CC[A/G/T]...3'
 3'... GG[T/C/A]...5。

【0149】

ニッカーゼ反応の生成物の以後の化学的処理により、リン酸塩骨格が切断され断片が生成される。

40

【0150】

Fen-1断片化法は、Fen-1酵素が関与し、これは「フラップ (flap)」エンドヌクレアーゼとして知られている部位特異的ヌクレアーゼである (米国特許第5,843,669号明細書、同5,874,283号明細書、同6,090,606号明細書)。この酵素は標的DNA鎖にハイブリダイズした2つのオリゴヌクレオチドの重複により作製されたDNA「フラップ (flap)」を認識し切断する。この切断は特異性が高く、単一の塩基の変化を認識することができ、目的のヌクレオチド位置における単一のメチル化塩基の検出を可能にする。Fen-1酵素は、Fen-1様ヌクレアーゼ、例えばヒト、マウス、およびアフリカツメガエル (Xenopus) XPG酵素、および酵母RAD2ヌクレアーゼ、または例えばM. jannaschii、P. furiosus、およびP. woese

50

iのFen-1エンドヌクレアーゼでもよい。

【0151】

使用可能な他の方法は、DNAキメラの切断である。3成分系のDNA-RNA-DNAプローブが、標的核酸分子、例えば結核菌 (*M. tuberculosis*) 特異的配列にハイブリダイズされる。RNaseHを付加するとキメラプローブのRNA部分が分解され、DNA部分を放出する (Yule, *Bio/Technology* 12: 1335 (1994))。

【0152】

d. 塩基特異的断片化

標的核酸分子は、特定の塩基 (例えば、DNAについてはA、C、T、またはG、およびRNAについてはA、C、U、またはG) で選択的に切断するヌクレアーゼを使用して断片化することができる。ある実施態様において、3つのRNAヌクレオチド (例えばU、G、およびA)、2つのRNAヌクレオチド (例えばCとU)、または1つのRNAヌクレオチド (例えばA) を特異的に切断するRNaseを使用して、標的核酸分子の転写体を特異的に切断することができる。例えばRNaseT1はssRNA (1本鎖RNA) をGリボヌクレオチドで切断し、RNaseU2はssRNAをAリボヌクレオチドで切断し、RNaseCL3とクサチビン (cusativin) はssRNAをCリボヌクレオチドで切断し、PhyMはssRNAをUおよびAリボヌクレオチドで切断し、RNaseAはssRNAをピリミジンリボヌクレオチド (CとU) で切断する。単一特異的RNase、例えばRNaseT₁ (G特異的) とRNaseU₂ (A特異的) の使用は当該分野で公知である (Donis-Kellerら、*Nucl. Acids Res.* 4: 2527-2537 (1977); GuptaとRanderath, *Nucl. Acids Res.* 4: 1957-1978 (1977); KuchinoとNishimura, *Methods Enzymol.* 180: 154-163 (1989); およびHahnerら、*Nucl. Acids Res.* 25(10): 1957-1964 (1997))。他の酵素、トリ肝リボヌクレアーゼ (RNase CL3) はシチジンで選択的に切断することが報告されているが、この塩基に対するこの酵素の傾向は反応条件により影響を受けることが報告されている (Boguskiら、*J. Biol. Chem.* 255: 2160-2163 (1980))。報告はまた、キュウリ (*Cucumis sativus* L) の乾燥種子から単離された別々のリボヌクレアーゼクサチビン (cusativin) のシチジン特異性を主張している (Rojoら、*Planta* 194: 328-338 (1994))。あるいはRNase PhyM (AおよびU特異的) (Donis-Keller, H. *Nucl. Acids Res.* 8: 3133-3142 (1980))、およびRNaseA (CおよびU特異的) (Simoncsitsら、*Nature* 269:833-836 (1977); GuptaとRanderath, *Nucleic Acids Res.* 4: 1957-1978 (1977)) の使用によるピリミジン残基の同定が報告されている。かかる切断パターンの例は、StanssensらのW000/66771号に記載されている。

【0153】

さらに、例えば修飾ヌクレオチドを核酸中に取り込み、ヌクレオチドの塩基を切り出すことにより塩基を標的化することができる; 次に適切な条件下または酵素を用いて核酸を処理すると、切り出した塩基の部位で核酸が断片化される。例えばdUPTはDNA中に取り込まれ、UDGを使用してウラシル塩基を除去し、次に公知の切断条件下でDNAを切断することにより、塩基特異的断片化が行われる。別の例ではメチル-シトシンがDNA中に取り込まれ、メチルシトシンドグリコシラーゼを使用してメチルシトシンを除去し、次に公知の条件下で処理してDNA断片化を引き起こすことにより塩基特異的断片化が行われる。塩基特異的断片化は部分的切断反応 (標的核酸がそこに取り込まれた切断不可能なヌクレオチドを含有する時、完了するまで行われる部分的切断反応を含む) と全切断反応で使用することができる。

【0154】

RNaseを使用する塩基特異的反応条件は当該分野で公知であり、例えば4 mM トリス酢酸 (pH 8.0)、4 mM KAc、1 mM スペルミジン、0.5 mM ジチオスレイトール、および1.5 mM MgCl₂がある。

【0155】

ある実施態様において増幅産物は1本鎖RNA分子に転写され、標的核酸分子の転写によりRNA分子が得られ、これは特異的RNAエンドヌクレアーゼを使用して切断することができる。ある実施態様において標的核酸分子の転写はRNA分子を与え、これは特異的RNAエンドヌクレアーゼを使用して切断することができる。例えばRNA分子の塩基特異的切断は、2つの

異なるエンドリボヌクレアーゼ（例えばRNaseT1とRNaseA）を使用して行われる。RNaseT1はGヌクレオチドを特異的に切断し、RNaseAはピリミジンリボヌクレオチド（例えばシトシンおよびウラシル残基）を特異的に切断する。ある実施態様において2つ以上のヌクレオチドを切断する酵素（例えばRNaseA）が切断のために使用される時、非切断可能ヌクレオチド（例えばdNTP）を標的核酸分子または増幅産物の転写中に取り込むことができる。例えば増幅産物の転写中にdCTPを取り込むことができ、得られる転写核酸はUリボヌクレオチドでRNaseAによる切断を受けるが、CデオキシリボヌクレオチドでのRNaseAによる切断に対しては抵抗性である。別の例では標的核酸分子の転写中にdTTPが取り込まれ、得られる転写核酸はCリボヌクレオチドでRNaseAによる切断を受けるが、TデオキシリボヌクレオチドでのRNaseAによる切断に対しては抵抗性である。非切断可能ヌクレオチド（例えばdNTP）を選択的に使用し、RNase（例えばRNaseAおよびRNaseT1）を使用して塩基特異的切断を行うことにより、同じ標的核酸配列の異なる転写体について3つの異なるヌクレオチド塩基に特異的な塩基切断を行うことができる。例えば特定の標的核酸分子の転写体はRNaseT1を使用するG特異的切断を受ける；この転写体は転写反応でdTTPを使用してC特異的切断を受け、次にRNaseAによる消化を受ける；および転写体は転写反応でdCTPを使用するT特異的切断を受け、次にRNaseAによる消化を受ける。

10

20

30

40

【0156】

別の実施態様においてdNTP、異なるRNase、および標的核酸分子の両方の配向の使用は、6つの異なる切断スキームを可能にする。例えば2本鎖標的核酸分子は2つの異なる1本鎖転写産物を与え、これは標的核酸分子のフォワード鎖の転写産物と標的核酸分子のリバース鎖の転写産物と呼ばれる。2つの異なる転写産物のそれぞれは、本明細書に記載のように3つの異なる塩基特異的切断反応（例えばG特異的切断、C特異的切断、およびT特異的切断）を受けて、6つの異なる塩基特異的切断反応を与える。この6つの可能な切断スキームは表1に示す。4つの異なる塩基特異的切断反応の使用は、標的核酸分子の1つの鎖のすべての4つのヌクレオチド塩基についての情報を与えることができる。リバース鎖上の相補的塩基を切断することによりフォワード鎖の切断を模倣することができることを考慮することにより、リバース鎖の切断を参照してフォワード鎖の4つのヌクレオチドのそれぞれについて塩基特異的切断を行うことができる。例えば標的核酸分子フォワード鎖の転写体について3つの塩基特異的切断反応を行うことにより、標的核酸分子フォワード鎖のG-、C-およびT-特異的切断を与えることができる；および4つの塩基特異的切断反応は標的核酸分子リバース鎖の転写体のT特異的切断反応でもよく、結果は標的核酸分子リバース鎖の転写体のA特異的切断に等しい。1つの標的核酸分子のすべての4つのヌクレオチド塩基に関する情報を与える塩基特異的切断は、可能な塩基特異的切断反応の種々の異なる組合せ（RNaseT1とRNaseAについて表1に示した切断反応を含む）を使用して行われ、フォワード鎖またはリバース鎖および/または非加水分解可能ヌクレオチドを使用する追加の切断反応が、当該分野で公知または本明細書に開示の他の塩基特異的RNaseを用いて行うことができることは、当業者が理解している。

【0157】

【表1】

	フォワードプライマー	リバースプライマー
RNase T1	G 特異的開裂	G 特異的開裂
RNase A ; dCTP	T 特異的開裂	T 特異的開裂
RNase A ; dTTP	C 特異的開裂	C 特異的開裂

【0158】

ある例ではRNaseU2を使用して、標的核酸分子転写体を特異的に切断することができる

50

。RNaseU2はAヌクレオチドでRNAを塩基特異的に切断することができる。すなわちRNaseT1、RNaseU2、およびRNaseAを使用し、適切なdNTPを使用（RNaseAの使用と組合せて）することにより、標的核酸分子の1つのみの転写体を塩基特異的に切断して核酸分子のすべての4つの塩基位置を調べることができる。ある実施態様において、4つのリボヌクレオチドの1つのみを塩基特異的に切断するRNaseを使用して塩基特異的切断が行われる時、非切断可能ヌクレオチド3リン酸は必要ではない。例えば塩基特異的切断のためのRNaseT1、RNaseCL3、クサチビン（cusativin）、またはRNaseU2の使用は、標的核酸分子転写体中の非切断可能ヌクレオチドの存在を必要としない。RNase（例えばRNaseT1およびRNaseU2）の使用は、標的核酸分子のすべての4つのヌクレオチド塩基についての情報を与えることができる。例えば標的核酸分子のフォワード鎖およびリバース鎖の両方の転写体または増幅産物を合成することができ、各転写体はRNaseT1とRNaseU2を使用する塩基特異的切断を受けることができる。4つの切断反応の得られる切断パターンは、標的核酸分子の1つの鎖のすべての4つのヌクレオチド塩基についての情報を与える。かかる実施態様において2つの転写反応を行うことができる（フォワード標的核酸分子鎖の第1の転写とリバース標的核酸分子鎖の第2の転写）。

【0159】

種々の塩基特異的切断法もまた本発明の方法での使用が企図される。種々の異なる塩基特異的切断法が当該分野で公知でありかつ本明細書に記載されており、RNAの酵素的塩基特異的切断、修飾DNAの酵素的塩基特異的切断、およびDNAの化学的塩基特異的切断がある。例えばウラシルデグリコシラーゼ（UDG）またはメチルシトシンデグリコシラーゼ（MCDG）を使用する切断のような酵素的塩基特異的切断が当該分野で公知でありかつ本明細書に記載されており、本明細書に記載の酵素的RNase介在塩基特異的切断反応とともに行うことができる。さらに、非加水分解可能塩基を含有するRNAのような核酸を断片化し、こうして部分的に完全な塩基特異的切断反応を与える塩基特異的切断反応の使用が本明細書で企図される。

【0160】

2. ポリヌクレオチドの物理的断片化

核酸分子の断片化は、部分的または機械的力（機械的剪断力および超音波処理を含む）を使用して行われる。核酸分子の物理的断片化は、例えば流体力学的力を使用して行われる。典型的には溶液中の核酸分子は、核酸分子を含有する溶液を針を取り付けたシリンジに繰り返し出し入れすることにより剪断することができる。Thorstenson, Y.R.ら、「制御された不偏DNA剪断のための自動流体力学的方法」、Genome Research 8: 848-855 (1998)；Davison, P.F., Proc. Natl. Acad. Sci. USA 45: 1560-1568 (1959)；Davison, P.F., Nature 185: 918-920 (1960)；Schriefer, L.A.ら、「低圧DNA剪断：ランダムDNA配列分析法」、Nucl. Acids Res. 18: 7455-7456 (1990)。例えば皮下注射針を用いるDNAの剪断は、典型的には1~2kbの範囲の大部分の断片を生成するが、少数の断片は300bpという小さいものもある。

【0161】

核酸分子（例えばゲノムDNAを含む）を剪断するための装置は市販されている。装置の例は、小さな急激な収縮によりDNA試料を押し出すことにより流体力学的剪断力を作製するためのシリンジポンプである。Thorstenson, Y.R.ら、「制御された不偏DNA剪断のための自動流体力学的方法」、Genome Research 8: 848-855 (1998)。剪断のための容量は典型的には100~250 μ lであり、処理時間は15分未満である。試料の剪断は、コンピューター制御により完全に自動化することができる。

【0162】

Oefnerらにより開発された流体力学的点すいこみ（point sink）剪断法は、流体力学的力を利用する核酸分子を剪断する1つの方法である。Oefner, P.J.ら、「再循環点すいこみ（point sink）流れシステムにおいて剪断されるDNAの効率的なランダムサブクローニング」、Nucl. Acids Res. 24(20): 3879-3886 (1996)。「点すいこみ（point sink）」は、このシステムの流体力学的流れの理論的モデルを意味する。ひずみ率テンソルは分子

上の力、従ってその破断を説明する。DNAの破断はこのテンソルの「剪断」項に帰因し、このクラスの断片化法は剪断力と呼ばれた。破断は、剪断項（流体は狭い管すなわちオリフィス中にある）と伸縮ひずみ項（流体がオリフィスに近づく時）の両方により引き起こすことができる。点すいこみ（point sink）剪断法は、核酸分子（例えばDNA）を非常に直径の小さいチューブ中にポンプ（例えばHPLCポンプ）を用いて圧力をかけて押し通すことにより行われる。得られる断片は狭いサイズ範囲を有し、最も大きい断片は最も小さい断片の約2倍である。断片のサイズは流速に反比例する。

【0163】

核酸分子断片はまた、溶液中の大きな核酸分子を攪拌、例えば溶液を混合、ブレンド、攪拌、またはボルテックス混合することにより得ることができる。Hershey, A.D.とBurgi, E. J. *Mol. Biol.* 2: 143-152 (1960); Rosenberg, H.S.とBendich, A. J. *Am. Chem. Soc.* 82: 3198-3201 (1960)。所望のサイズまたはサイズ範囲の断片が得られるまで、溶液は種々の時間攪拌することができる。ビーズまたは粒子の溶液への添加は、核酸分子の断片化を助ける。

10

【0164】

核酸分子を物理的に断片化する1つの適した方法は、核酸分子の超音波処理に基づく。Deininger, P.L.、「迅速なDNA配列分析へのアプローチ」、*Anal. Biochem.* 129:216-223 (1983)。超音波処理による核酸分子断片の作製は、典型的には緩衝化した核酸分子を含有するマイクロ遠心分離管を超音波処理機（例えばカップホーン（cup-horn）超音波処理機）の氷水浴中に入れ、最大出力と連続的出力を使用して短いバーストで異なる回数超音波処理することにより行われる。短いバーストは約10秒間の持続でもよい。例えばBankier, A.T.ら、「M13/ジデオキシヌクレオチド鎖停止法によるランダムクローニングと配列決定」、*Meth. Enzymol.* 155: 51-93 (1987)。ある超音波処理プロトコルの例では、大きな核酸分子の超音波処理により超音波処理条件（例えば持続時間と超音波処理の強度）に依存して300~500bpまたは2~10kbの範囲の断片が得られた。Kawata, Y.ら、「TAベクターを使用するゲノムライブラリーの調製」、*Prep. Biochem & Biotechnol.* 29(1): 91-100 (1999)を参照されたい。

20

【0165】

超音波処理中に温度を上昇させると不均一な分布パターンが得られ、このため浴の温度を注意深く追跡し、必要であれば新鮮な氷水を加える。超音波処理の具体的な条件を決定するための超音波処理プロトコルの例には、100 µgの核酸分子試料を350 µlの適当な緩衝液中に35 µlの10個のアリコートで入れ、このうち5個を10秒間バーストの数を増加させる超音波処理に付す。例えば10秒間バーストの間に少なくとも1分間チューブを氷水浴に入れることにより、核酸分子が冷却される。超音波処理機中の氷水浴は、必要であれば各試料間で交換してもよい。試料を遠心分離して縮合物を回収し、アリコートをアガロースゲルでサイズマーカーに対して電気泳動する。アガロースゲル電気泳動から検出される断片サイズ範囲に基づき、残りの5つのチューブを超音波処理して所望の断片サイズを得ることができる。

30

【0166】

核酸分子の断片化はまたネブライザーを使用して行ってもよい。Bodenteich, A., Chisoe, S., Wang, Y.-F.とRoe, B.A. (1994)。Adams, M.D., Fields, C.とVenter, J.C. (編)「自動DNA配列決定と分析 (Automated DNA Sequencing and Analysis)」、アカデミックプレス (Academic Press)、サンジエゴ、カリフォルニア州。ネブライザーは当該分野で公知であり、市販されている。ネブライザーを使用する核酸分子断片化のプロトコルの例は、25~50%グリセロールを含有する2mlの緩衝化核酸分子溶液（約50 µg）を氷水浴に入れ、溶液に気体流（例えば窒素）を8~10psiで2.5分間送る。任意の気体（特に不活性気体）を使用することができる。ガス圧は、断片サイズの主要な決定因子である。圧力を変えることにより、異なる断片サイズが得られる。均一に分布した断片を作製するのに噴霧化のために氷水浴を使用することができる。同様に高圧噴霧剤噴霧器を使用して断片を作製することができる。Cavaliere, L.F.とRosenberg, B.H., *J. Am. Chem. Soc.* 81

40

50

: 5136-5139 (1959)。

【0167】

核酸分子を断片化するための別の方法は、核酸分子の緩衝化溶液の凍結融解の繰り返しを使用する。核酸分子の試料を必要に応じて凍結融解して所望のサイズまたはサイズ範囲の断片を作製することができる。さらに核酸分子にイオンまたは粒子を衝突させて、種々のサイズの断片を作製することができる。例えば真空下でイオン抽出ビームラインに核酸分子を曝露することができる。イオンは 7kV^+q で電子ビームイオントラップから抽出され、標的核酸分子に向けられる。核酸分子は、例えば 100 イオン/ μm^2 の総流束量が達成されるまで任意の時間(典型的には数時間)照射される。

【0168】

核酸分子断片化はまた、核酸分子に放射線照射することにより行われる。典型的には核酸分子を断片化するのにガンマ線またはX線のような放射線照射が充分である。断片のサイズは、放射線の強度と曝露時間を調整することにより調整することができる。紫外線照射もまた使用できる。核酸分子への放射線の好ましくない影響を最小にするために、曝露強度と時間を調整することができる。

【0169】

核酸分子を沸騰することによっても断片を作製することができる。典型的には核酸分子の溶液は一定攪拌下で数時間沸騰される。約 500 bpの断片が達成される。断片のサイズは沸騰の持続により変化する。

【0170】

3. 核酸分子の化学的断片化

塩基特異性有りまたは塩基特異性無しで、化学的断片化を使用して核酸分子を断片化することができる。核酸分子は、例えば塩基と酸加水分解を含む加水分解反応を含む化学反応により断片化することができる。RNA(または対になっていない塩基)はアルカリ性条件下で不安定なため、アルカリ性条件を使用してニックを含有する核酸分子またはRNAを断片化することができる。Nordhoffら、「赤外線マトリックス支援レーザー脱離/イオン化質量スペクトル法中の核酸のイオン安定性」、Nucl. Acids Res. 21(15): 3347-3357 (1993)を参照されたい。DNAは、酸、典型的には強酸(例えば 6M HCl)の存在下で加水分解してもよい。加水分解を促進するために温度は室温より高くしてもよい。反応条件と反応時間の長さに依存して、核酸分子は種々のサイズ(単一塩基断片を含む)に断片化することができる。加水分解は、厳しい条件下では、両方のホスフェートエステル結合、およびデオキシリボースとプリンおよびピリミジン塩基とのN-グリコシド結合を破断することができる。

【0171】

核酸分子断片を生成するための酸/塩基加水分解プロトコルの例は公知である(例えばSargentら、Meth. Enz. 152: 432 (1988)を参照)。簡単に説明すると 1 gのDNAを 50 mlの 0.1 N NaOHに溶解する。 1.5 mlの濃縮HClを加え、溶液を急速に混合する。直ちにDNAが沈殿し、大きな凝集物の生成を防ぐために数秒間以上攪拌してはならない。試料を室温で 20 分間インキュベートして、DNAを部分的に脱プリン化する。次に、 2 mlの 10 N NaOH(OH濃度は 0.1 Nまで)を加え、DNAが完全に再溶解するまで試料を攪拌する。典型的なサイズは約 250 ~ 1000 ヌクレオチドの範囲であるが、加水分解の条件により短くも長くもなる。

【0172】

化学的切断も特異的となり得る。例えば選択された核酸分子は、アルキル化、特にホスホロチオエート修飾核酸分子により切断することができる(例えばK.A. Browne、「金属イオン触媒核酸アルキル化と断片化」、J. Am. Chem. Soc. 124(27): 7950-7962 (2002)を参照)。ホスホロチオエート修飾のアルキル化は、修飾部位で核酸分子が切断を受けやすくする。I.G. GutとS. Beckは質量スペクトルで検出するためにDNAをアルキル化する方法を記載している。I.G. GutとS. Beck、「選択的DNAアルキル化と質量スペクトルによる検出方法」、Nucl. Acids Res. 22(8): 1367-1373 (1995)。

【0173】

10

20

30

40

50

オリゴヌクレオチドの塩基特異性および塩基非特異的化学的切断のための種々の追加の化学物質と方法は当該分野で公知であり、本明細書に記載の断片化法での使用が企図される。例えば塩基特異性切断は、ギ酸ピペリジン、ピペリジン、硫酸ジメチル、ヒドラジンと塩化ナトリウム、ヒドラジンのような化学物質を使用して行われる。例えばDNAは硫酸ジメチルとピペリジンを使用してGヌクレオチドで塩基特異的に切断できる；DNAは硫酸ジメチル、ピペリジンおよび酸を使用してAおよびGヌクレオチドで塩基特異的に切断できる；DNAはヒドラジンとピペリジンを使用してCおよびTヌクレオチドで塩基特異的に切断できる；DNAはヒドラジン、ピペリジンおよび塩化ナトリウムを使用してCヌクレオチドで塩基特異的に切断できる；およびDNAは強塩基を使用してCヌクレオチドに対してより低い特異性を持ってAヌクレオチドで塩基特異的に切断できる。別の例ではリボヌクレオチドとデオキシリボヌクレオチドは標的核酸分子中に取り込むことができ、標的核酸は、RNAまたはDNAを特異的に切断するための条件に接触させて、標的核酸分子の組成に従って塩基特異的切断（部分的または完全な切断）を与えることができる。

10

20

30

40

50

【0174】

4. 断片化法の組合せ

断片はまた、本明細書に記載の断片化法の任意の組合せ、例えば異なる酵素的断片化法の組合せ、異なる化学的断片化法の組合せ、異なる物理的断片化法の組合せ、または酵素的断片化法と化学的断片化法の組合せ、酵素的断片化法と物理的断片化法の組合せ、化学的断片化法と物理的断片化法の組合せ、または酵素的断片化法と化学的断片化法と物理的断片化法の組合せを使用して、作製することができる。いくつかの具体例には、特に限定されないが、異なる塩基特異的切断法の組合せ、および剪断と配列特異的酵素との組合せがある。特異的断片を作製する方法は、ランダム断片を作製する方法と組合せることができる。さらに、ランダム断片を作製する異なる方法を組合せ、特異的断片を作製する異なる方法を組合せることができる。例えば特定の部位で核酸分子を切断する1つまたはそれ以上の酵素を、異なる部位で核酸分子を特異的に切断する1つまたはそれ以上の酵素と組合せて使用することができる。別の例では、特定の種類の核酸分子を切断する酵素を組合せ、例えばRNaseをDNaseと組合せて使用するか、または1本鎖特異的ヌクレアーゼを2本鎖特異的ヌクレアーゼと組合せて使用するか、またはエキソヌクレアーゼをエンドヌクレアーゼと組合せて使用することができる。さらに別の例では、核酸分子をランダムに切断する酵素を核酸分子を特異的に切断する酵素と組合せて使用することができる。組合せての断片化の使用は、核酸分子について1つまたはそれ以上の方法を続けてまたは同時に行うことを意味する。

【0175】

本明細書に記載のように、組合せの使用は核酸分子試料の第1の画分について第1の断片化法を使用し、核酸分子試料の第2の画分について第2の断片化法を使用することを含む。2つの試料は以後の検出と質量測定法で別々に分析するか、または2つの試料は一緒にプールされて以後の検出と質量測定法で同時に分析することができる。断片化法の組合せには、2つまたはそれ以上の断片化法、3つまたはそれ以上の断片化法、または4つまたはそれ以上の断片化法を含む。

【0176】

5. ハイブリダイゼーション後の断片化

標的核酸はまた、標的核酸が捕獲オリゴヌクレオチドプローブとハイブリダイズした後に断片化することができる。ある実施態様において標的核酸は1つまたはそれ以上の断片化工程を受けた後、捕獲オリゴヌクレオチドプローブとハイブリダイズし、次に捕獲オリゴヌクレオチドプローブとハイブリダイズした後に1回またはそれ以上の追加の断片化を受ける。別の実施態様において標的核酸は捕獲オリゴヌクレオチドプローブとハイブリダイズする前に断片化工程を受けないが、捕獲オリゴヌクレオチドプローブとハイブリダイズした後に1回またはそれ以上の断片化工程を受ける。標的核酸が捕獲オリゴヌクレオチドプローブにハイブリダイズした後に起きる反応の例には、酵素的および化学的断片化がある。ある実施態様においてかかるハイブリダイゼーション後断片化工程は、1本鎖核酸

を選択的に断片化するが、2本鎖核酸は断片化しない。別の実施態様においてハイブリダイゼーション後の断片化には塩基特異的切断を含む。

【0177】

E. 捕獲オリゴヌクレオチド

本明細書に記載の方法と組成物に含まれるのは、標的核酸断片がハイブリダイズできる1つまたはそれ以上の捕獲オリゴヌクレオチドである。本発明の捕獲オリゴヌクレオチドは、典型的にはいくつかの標的核酸断片が捕獲オリゴヌクレオチドにハイブリダイズし、いくつかの標的核酸断片が捕獲オリゴヌクレオチドにハイブリダイズしない条件下で、標的核酸断片と接触させることができる。捕獲オリゴヌクレオチドにハイブリダイズする標的核酸断片は、捕獲オリゴヌクレオチドにハイブリダイズしない標的核酸断片から分離することができる。捕獲オリゴヌクレオチドにハイブリダイズする標的核酸断片と捕獲オリゴヌクレオチドにハイブリダイズしない標的核酸断片は、捕獲オリゴヌクレオチドと接触した後、および/またはハイブリダイズした断片とハイブリダイズしない断片とを分離した後、別の処理工程に付すことができる。標的核酸断片を捕獲オリゴヌクレオチドと接触させた後、標的核酸断片の質量を測定することができる。標的核酸断片を捕獲オリゴヌクレオチドと接触させることは核酸断片の分離を引き起こすため、捕獲オリゴヌクレオチド接触標的核酸断片からの質量スペクトルは捕獲オリゴヌクレオチドと接触していない断片と比較して小さい質量（例えば、異なる質量で小さいピーク）を有する。捕獲オリゴヌクレオチドは1つだけの配列にハイブリダイズするように使用することができるが、例えば縮重塩基または低もしくは中程度の緊縮性条件を使用することにより2つ以上の捕獲オリゴヌクレオチド配列と意図的にハイブリダイズするように捕獲オリゴヌクレオチドを使用することもできる。捕獲オリゴヌクレオチドにハイブリダイズする異なる標的核酸断片の数と多様性は、質量スペクトルにより測定される異なる断片の数と多様性を決定することができる。

【0178】

すなわち本明細書に記載の方法の1つの例は、標的核酸断片の質量を測定する方法であって：

(a) 捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸断片の複雑さを制御する（ここで各標的核酸断片は、捕獲オリゴヌクレオチドプローブにハイブリダイズする少なくとも1つの第1の領域を含む）；そして

(b) 質量スペクトルを使用して捕獲オリゴヌクレオチドにハイブリダイズした標的核酸断片の質量を測定する、ことを含んでなり、

ここで複雑さを制御する工程は、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸断片の第1の領域中の異なる配列の数を調節することを含み、こうして各第1の領域中に異なるヌクレオチド配列を含む2つまたはそれ以上の標的核酸断片が捕獲オリゴヌクレオチドプローブにハイブリダイズするようになる方法である。

【0179】

1. 標的核酸断片の複雑さの制御

本明細書に記載の方法は、別の場所に記載したように標的核酸断片の質量を測定する工程を含む。具体的な測定法（例えばその質量が1回の質量スペクトルで測定される）でその質量が測定される標的核酸断片の数と多様性に依存して、異なる断片の質量が容易に区別でき、特定の質量で表される異なるヌクレオチド配列の数は大きくても小さくても良く、欠如している質量（例えば、考えられるが質量ピークが存在しない）は容易に同定されるかまたは同定されない。断片の複雑さが極めて低い時、質量スペクトルは数個の存在する/欠如した質量を有するのみであり、これは配列決定法により与えられる強固性を限定する（例えば、質量測定により1つのみで断片が存在するかまたは欠如することが測定された時、ハイブリダイゼーション法による従来の配列決定では得られない情報がほとんど得られない）。断片の複雑さが極めて高い時、質量スペクトルは多くの存在する/欠如した質量を有し、各質量は多くの異なるヌクレオチド配列を表し、これが特定の観察（例えば、存在するかまたは後述した質量）を使用して高い確率でヌクレオチド配列を帰属する

ことができる程度を限定する（例えば、存在する／欠如した断片が多すぎる時、捕獲オリゴヌクレオチドのハイブリダイゼーション無しで質量スペクトル法とは異なる複雑さの上昇はほとんど無い）。すなわち標的核酸断片の複雑さを制御することは、質量スペクトルを「調整」するのに有用であり、質量スペクトルが多く分析可能な観察（例えば質量の分析可能な存在または欠如）を与え、場合により観察は、配列決定を可能にする少数の十分な数の異なる配列である。

【0180】

ある実施態様において標的核酸断片の複雑さは、標的核酸断片の質量を測定する前に制御される。別の例において複雑さを制御することは、標的核酸断片のある領域を制御することを含み、ここで少なくとも1つの標的核酸断片は、その複雑さが制御されるかまたは複雑さが異なって制御される第2の領域をさらに含有する。

10

【0181】

a. 複雑さを制御する方法

本明細書に記載のように標的核酸の断片化は、固体支持体に結合した捕獲オリゴヌクレオチドとの標的核酸のハイブリダイゼーションとともに、その質量を分析すべき標的核酸の混合物の複雑さを制御または低下させるように機能することができる。

【0182】

複雑さを制御するある例では、断片化は標的核酸断片の長さを制御し、標的核酸断片中の配列の一部（標的核酸断片の3'末端、5'末端、または3'末端と5'末端の両方の1つまたはそれ以上のヌクレオチド位置の本体を含む）を制御することができる。別の例では標的核酸の捕獲オリゴヌクレオチドへのハイブリダイゼーションは、捕獲オリゴヌクレオチドプローブとハイブリダイズする領域中の標的核酸配列の複雑さを制御することができる。ある実施態様において標的核酸の第1の領域が捕獲オリゴヌクレオチドプローブとハイブリダイズする時、標的核酸の第1の領域の複雑さは、標的核酸の第2の非ハイブリダイズ領域の複雑さとは別々に制御することができる。

20

【0183】

例えば捕獲プローブが5ヌクレオチドの長さで、標的核酸配列が8ヌクレオチドの長さである時、例えばハイブリダイゼーション条件、および2つの異なる標的核酸配列のみが捕獲オリゴヌクレオチドプローブ配列にハイブリダイズすることを可能にする捕獲オリゴヌクレオチドプローブ配列を使用して、複雑さを制御することができ、こうして特定の捕獲プローブオリゴヌクレオチドにハイブリダイズする異なる標的核酸断片の可能な数を512以下に限定する。配列特異的断片化条件を使用して、例えば上記したように配列特異的エンドヌクレアーゼまたは塩基特異的切断を使用することにより、複雑さはさらに限定することができる。

30

【0184】

一般に捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸断片のハイブリダイズ領域および非ハイブリダイズ領域の両方の複雑さは、配列特異的または非特異的断片化法を使用して、標的核酸断片の長さを制御し、標的核酸断片の統計的サイズ範囲の異なる長さの数を制御し、分析される標的核酸の全体の長さを制御することにより、および標的核酸の5'末端または3'末端でヌクレオチド位置にハイブリダイズする捕獲オリゴヌクレオチドプローブの能力を制御することにより、制御することができる。さらにハイブリダイズ領域の複雑さは、標的核酸が捕獲オリゴヌクレオチドプローブに曝露される条件を修飾することにより（例えば、低緊縮性ハイブリダイゼーション条件、中程度の緊縮性ハイブリダイゼーション条件、または高緊縮性ハイブリダイゼーション条件）、および捕獲オリゴヌクレオチドプローブのヌクレオチドの数および／またはヌクレオチドの縮重を修飾することにより（例えばユニバーサルまたは半ユニバーサルヌクレオチドを使用することにより）、さらに制御することができる。例えば捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸断片の複雑さは、配列特異的または非特異的断片化法を使用して、標的核酸断片の5'末端または3'末端でヌクレオチド位置とのハイブリダイズを促進する捕獲オリゴヌクレオチドプローブを使用して、上昇した緊縮性ハイブリダイゼーション

40

50

条件を使用して、および捕獲オリゴヌクレオチド中でより配列特異的なヌクレオチドを含めることにより、標的核酸断片の長さを減少させることにより、標的核酸断片の統計的サイズ範囲の異なる長さの数を減少させることにより、分析される標的核酸の全体の長さを低下させることにより低下させることができる。別の例では、捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸断片のハイブリダイズ領域および非ハイブリダイズ領域の両方の複雑さは、標的核酸特定の領域とのハイブリダイズを促進しない捕獲オリゴヌクレオチドプローブを使用して、低下した緊縮性ハイブリダイゼーション条件を使用して、および捕獲オリゴヌクレオチド中でより少ない配列特異的なヌクレオチド（例えば、ユニバーサルまたは半ユニバーサル塩基）を含めることにより、標的核酸断片の長さを増加させることにより、標的核酸断片の統計的サイズ範囲の異なる長さの数を増加させることにより、分析される標的核酸の全体の長さを増加させることにより増加させることができる。

10

20

30

40

50

【0185】

ある実施態様において捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸断片の複雑さは、標的核酸断片の質量を測定する工程の前に制御される。例えば標的核酸断片の複雑さの制御は、標的核酸断片を捕獲オリゴヌクレオチドプローブにハイブリダイズする前に行われ（例えば断片化工程で）、および/または標的核酸断片の複雑さの制御は、標的核酸断片を捕獲オリゴヌクレオチドプローブにハイブリダイズさせることを含み、および/または標的核酸断片の複雑さの制御は標的核酸断片を捕獲オリゴヌクレオチドプローブにハイブリダイズさせた後であるが標的核酸断片の質量を測定する前（例えば、以後の「トリミング」のような断片化工程で）に行われる。

【0186】

標的核酸断片産物は、種々の方法で固相上に捕獲される。例えば1つまたはそれ以上の断片化生成物と特異的または半特異的にハイブリダイズする捕獲オリゴヌクレオチドは、生成物の特異的または「半特異的」捕獲のために固体支持体に結合することができる。

【0187】

本明細書の教示と当該分野の情報に従って当業者は、特定の捕獲オリゴヌクレオチドに結合した標的核酸断片の予測される複雑さを推定することができる。例えば特定の配列を含有する捕獲オリゴヌクレオチドは、捕獲オリゴヌクレオチドとして同じ長さの最大4つの異なる標的核酸断片までユニバーサルヌクレオチド（例えばイノシン）を含む1つの縮重位置を含有し、同じ配列組成（ユニバーサル塩基と相補的な位置のヌクレオチド以外）がほぼ同等の結合親和性をもって特定の捕獲オリゴヌクレオチドに結合できるであろう。より大きな標的核酸断片が存在し捕獲オリゴヌクレオチドより1~5ヌクレオチド長いなら、最大30,948個の異なる標的核酸断片が1つのオリゴヌクレオチド配列に結合できるであろう（図2参照）。同様に捕獲オリゴヌクレオチドがユニバーサルオリゴヌクレオチドに対応する2つの縮重位置を有する場合、同じ長さで配列組成（ユニバーサル塩基に相補的な位置のヌクレオチドを除く）を有する最大16個の異なる標的核酸断片が、ほぼ同様の結合親和性で特定の捕獲オリゴヌクレオチドに結合できるであろう。

【0188】

ある実施態様において標的核酸断片の非ハイブリダイズ領域は完全に除去される。これは、例えば捕獲オリゴヌクレオチドプローブと同じサイズの標的核酸断片を作製し、または捕獲オリゴヌクレオチドプローブより大きな標的核酸断片を作製し、標的核酸を捕獲オリゴヌクレオチドプローブにハイブリダイズさせ、次に1本鎖特異的ヌクレアーゼを使用してハイブリダイズしなかったヌクレオチドを切断することにより行われる。

【0189】

ある実施態様において、特定の捕獲プローブにハイブリダイズする最小数の異なる配列に関する情報を得ることができる。例えば低緊縮性ハイブリダイゼーション条件または縮重捕獲オリゴヌクレオチドプローブが使用される時、同じ捕獲オリゴヌクレオチドプローブ配列に2つ以上の標的核酸配列が結合できる。そのような場合すべての標的核酸断片が捕獲オリゴヌクレオチドプローブと同じサイズで、すべての標的核酸断片が異なる組成（

すなわち、異なる数のA、C、TおよびG)を有するなら、質量ピークの数、捕獲オリゴヌクレオチドプローブにハイブリダイズした異なる標的核酸配列の数に対応するであろう。異なる配列を有する標的核酸断片が同じ組成(すなわち、同じ数のA、C、TおよびG)を有することは可能であるため、いくつかの異なる配列が同じ質量測定値を有することができ、従って質量ピーク数は、存在する異なる配列の最小数を提供する。

【0190】

非ハイブリダイズ末端(例えば5'末端と3'末端)もまた、例えば1塩基特異的切断のような配列特異的切断により、その塩基組成に基づいて修飾することができる。例えば使用される標的核酸断片がRNAであり、RNAが最初に捕獲プローブにハイブリダイズし、次にRNaseT1(これはGの3'末端で特異的に1本鎖RNAを切断する)に曝露された場合、異なる標的プローブの非ハイブリダイズ末端は標的核酸のハイブリダイズ末端に最も近いGの位置に従って長さが変化するであろう。すなわち非ハイブリダイズ末端の塩基特異的切断のような方法は、塩基特異的切断の前に非ハイブリダイズ末端があらかじめ決められた長さを必要とせずに、非ハイブリダイズ末端の制御を可能にする。

10

【0191】

非ハイブリダイズ末端の塩基特異的切断は、核酸中に典型的に存在する4つの塩基のいずれについても行うことができる。ある実施態様において標的核酸の試料は4つの別の試料に分離され、各分離試料は1つまたは4つの同じチップ上の捕獲プローブにハイブリダイズされる。捕獲プローブにハイブリダイズ後、4つのチップ(または1つのチップ上の4つの異なる位置)の標的核酸はそれぞれ4つの異なる塩基特異的切断反応の1つに付される。最後にハイブリダイズした標的核酸の質量が測定される。この4重塩基特異的切断はまた連続して行うことができ、ここで4つの別の試料は同じチップに連続してハイブリダイズされ、4つの塩基特異的切断反応の1つで処理され質量が測定される。同じ捕獲プローブにハイブリダイズした4つの異なる塩基特異的切断反応物からの標的核酸の質量を測定することにより、1塩基特異的切断後に同じ組成(従って同じ質量)を有するかも知れない非ハイブリダイズ末端の異なる配列は、1回またはそれ以上の異なる塩基特異的切断後に異なる組成(従って異なる質量)を有する。

20

【0192】

当業者に公知のように断片化、ハイブリダイゼーション、および場合によりさらなる断片化の種々の任意の追加的組合せを行って所望の複雑さに達することができる。

30

【0193】

b. 断片の領域

標的核酸断片は、少なくとも1つ、少なくとも2つ、少なくとも3つの領域を含有し得る。例えば1つの領域のみを含有する標的核酸断片は、標的核酸のすべてのヌクレオチドが捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸でもよい;少なくとも2つの領域を含有する標的核酸は、標的核酸のヌクレオチドのサブセットのみが捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸でもよい(例えば、2つの領域を含有する標的核酸は、標的核酸の3'末端が捕獲オリゴヌクレオチドプローブにハイブリダイズし、5'末端はハイブリダイズしないものであり、またその逆もある);少なくとも3つの領域を含有する標的核酸は、標的核酸の中央領域(5'末端でも3'末端でもない)が捕獲オリゴヌクレオチドプローブにハイブリダイズするものであるか、または5'末端と3'末端(中央領域ではない)が捕獲オリゴヌクレオチドプローブにハイブリダイズするものである;4つ以上の領域を有する標的核酸は、捕獲オリゴヌクレオチドプローブにハイブリダイズする2つまたはそれ以上の物理的に分離された領域を有する標的核酸でもよい。

40

【0194】

同様に捕獲オリゴヌクレオチドプローブは1つまたはそれ以上の領域を有することができる。例えば2つの領域を有する捕獲オリゴヌクレオチドは、標的核酸断片にハイブリダイズする第1の領域と少なくとも1つの標的核酸にハイブリダイズしない第2の領域を有することができる。

【0195】

50

c. 部分的1本鎖捕獲オリゴヌクレオチド

別の実施態様において固体支持体上の捕獲オリゴヌクレオチドは、1本鎖オーバーハングを有する部分的2本鎖でもよい。捕獲オリゴヌクレオチドの1本鎖オーバーハングの長さは典型的には5~6ヌクレオチドであり、4~10ヌクレオチドまたはそれ以上の範囲でもよい。捕獲オリゴヌクレオチドが部分的に2本鎖であり、例えば5ヌクレオチド1本鎖オーバーハングを有する時、1024個の別個の場所を有する固体支持体は、すべての可能な標的核酸の5つのヌクレオチドに相補的な捕獲プローブを含有することができる。さらに1本鎖オーバーハングを有する2本鎖捕獲オリゴヌクレオチドの使用は、捕獲オリゴヌクレオチドプローブと標的核酸の一端との塩基積層相互作用を可能にすることにより、捕獲オリゴヌクレオチドへの標的核酸の親和性を上昇させる。標的核酸の一端を捕獲オリゴヌクレオチドプローブと塩基積層させることにより、標的核酸の一端の複雑さは、他端の複雑さから別に制御することができる。

【0196】

例えば捕獲プローブが、1つの鎖の3'末端から伸びる5ヌクレオチド1本鎖オーバーハングを有する時、標的核酸の3'末端の5つのヌクレオチドは捕獲プローブ1本鎖オーバーハングとハイブリダイズすることができる。捕獲プローブが縮重位置を持たない場合は、標的ヌクレオチドの1つの3'末端5塩基配列のみが最も高い相補性でプローブにハイブリダイズする。捕獲プローブが1つのユニバーサルまたは半ユニバーサル塩基を有するなら、標的核酸のそれぞれ4つのみまたは2つのみ3'末端5塩基配列が最も高い相補性でプローブにハイブリダイズする。

【0197】

さらにこの例では、捕獲プローブが1つの鎖の3'末端から伸びる5ヌクレオチド1本鎖オーバーハングを有する時、標的ヌクレオチドはもう5塩基より長くてもよい；この例では単純化するために、標的ヌクレオチドの長さは5~7塩基変動する。すなわち3つの異なる長さ(5塩基、6塩基、および7塩基)のヌクレオチドが、最も高い相補性で非縮重捕獲オリゴヌクレオチドプローブにハイブリダイズする。捕獲オリゴヌクレオチドプローブが非縮重であると仮定すると、標的核酸の各位置は4つの異なる塩基の任意のものを有することができるため、 $21(4^2 + 4^1 + 4^0)$ もの異なる標的核酸が各非縮重捕獲オリゴヌクレオチドプローブにハイブリダイズすることができる。捕獲プローブの1本鎖領域中に5つの塩基の1つがユニバーサル塩基であるなら、 21×4 すなわち84個の標的核酸が各捕獲プローブにハイブリダイズすることができる。ユニバーサル塩基を使用する代わりに、標的ヌクレオチドと捕獲プローブが相互作用する5つの位置の任意の位置で1つのミスマッチを許容するようにハイブリダイゼーション条件が操作されたなら、 $21 \times 4 \times 5$ すなわち420個の標的核酸が各捕獲プローブにハイブリダイズすることができる。標的核酸断片の1つの領域の複雑さまたは全断片の複雑さをモデル化するために、当業者が理解しているように、種々の他のプローブとハイブリダイゼーション緊縮性の任意のものに基づいて同様の計算をすることができる。

【0198】

5'末端の複雑さとは異なる3'末端の複雑さの制御は、上記3つの例で見られる。これらの例では5'末端配列は標的核酸の長さによってのみ制御され、従って5'末端は21個もの異なる配列を有するか、または長さおよび/または長さの変動が増加するとそれ以上の異なる配列を有することができる。この例では3'末端配列は、縮重位置および/またはハイブリダイゼーション条件の使用により制御され、従って3'末端の複雑さは1~20個の異なる配列の間で変化し、またはハイブリダイゼーションの緊縮性がさらに緩和されるかまたは追加の縮重位置が捕獲プローブ中に含まれるなら、それ以上の異なる配列で変化することができる。さらに3'末端の複雑さもまた、捕獲プローブ中に存在する1本鎖オーバーハング塩基の数により制御することができるであろう。

【0199】

2. 捕獲オリゴヌクレオチドの組成

捕獲オリゴヌクレオチドの所望の性質に基づき、捕獲オリゴヌクレオチドは種々の組成

のうちの任意の組成を有することができる。例えば捕獲オリゴヌクレオチドは1本鎖であるか、または1本鎖と2本鎖領域の両方を有し、捕獲オリゴヌクレオチドはユニバーサルおよび/または半ユニバーサル塩基を含有することができ、捕獲オリゴヌクレオチドは種々の長さの任意の長さでもよい。

【0200】

a. ヌクレオチドのタイプ

捕獲オリゴヌクレオチドは種々のヌクレオチド（天然に存在するものおよび天然に存在しないものの両方）のうち任意のヌクレオチドを含有することができる。典型的には捕獲オリゴヌクレオチドは、標的核酸の第2のセットのヌクレオチドより、標的核酸の第1のセットのヌクレオチドによくハイブリダイズする1つまたはそれ以上のヌクレオチドを含有する。例えば捕獲オリゴヌクレオチドは1つまたはそれ以上のA、G、C、またはT/Uを含有することができる。

10

【0201】

ある実施態様において捕獲オリゴヌクレオチドは部分的に縮重であり、1つまたはそれ以上の縮重塩基を含有してもよい。例えば1つまたはそれ以上の縮重塩基は、捕獲オリゴヌクレオチドの「3'末端に位置」することができる。別の実施態様において1つまたはそれ以上の縮重塩基が、捕獲オリゴヌクレオチドの「5'末端に位置」してもよい。例えば1つまたはそれ以上のユニバーサル塩基を捕獲オリゴヌクレオチドの一端に配置することは、捕獲オリゴヌクレオチドの塩基特異性を変更することなく、捕獲オリゴヌクレオチドと標的核酸の間のハイブリダイゼーションを増強するのに有用となり得る；しかしかかる配置は、捕獲オリゴヌクレオチドが選択的に結合する標的核酸の長さを変更するのに使用することができる。

20

【0202】

別の実施態様において1つまたはそれ以上の縮重塩基（例えばユニバーサル塩基および半ユニバーサル塩基）は、捕獲オリゴヌクレオチドプローブ中の特異的非縮重塩基の間に配置される。こうして捕獲オリゴヌクレオチドプローブの認識配列中のヌクレオチド位置の第1の選択されたサブセットは、捕獲オリゴヌクレオチドプローブの認識配列中にヌクレオチド位置の第2のサブセットと比較して、特定のヌクレオチドに対して特異性が上昇している。非縮重塩基間の縮重塩基の分布は、当業者に公知のように種々の型の任意の型を取ることができる。すなわち認識配列中の1つまたはそれ以上の別個の位置（ここで、縮重塩基は非縮重塩基の間に配置される）に1つまたはそれ以上の連続する縮重塩基を分布させてもよい。

30

【0203】

i. ユニバーサル塩基

捕獲オリゴヌクレオチドの縮重はユニバーサル塩基を使用して達成され、これはDNAまたはRNAの4つの典型的に存在する塩基の任意のものに同様の親和性で結合できる。本明細書で使用されるユニバーサル塩基の例には、イノシン、キサントシン、3-ニトロピロール（Bergstromら、*Abstr. Pap. Am. Chem. Soc.* 206(2):3 08 (1993)；Nicholsら、*Nature* 369: 492-493；Bergstromら、*J. Am. Chem. Soc.* 117: 1201-1209 (1995)）、4-ニトロインドール（Loakesら、*Nucleic Acids Res.*, 22: 4039-4043 (1994)）、5-ニトロインドール（Loakesら (1994)）、6-ニトロインドール（Loakesら (1994)）；ニトロイミダゾール（Bergstromら、*Nucleic Acids Res.* 25: 1935-1942 (1997)）、4-ニトロピラゾール（Bergstromら (1997)）、5-アミノインドール（Smithら、*Nucl. Nucl.* 17: 555-564 (1998)）、4-ニトロベンズイミダゾール（Seelaら、*Helv. Chim. Acta* 79: 488-498 (1996)）、4-アミノベンズイミダゾール（Seelaら、*Helv. Chim. Acta* 78: 833-846 (1995)）、フェニルC-リボヌクレオチド（Millicanら、*Nucleic Acids Res.* 12: 7435-7453 (1984)；Matulic-Adamicら、*J. Org. Chem.* 61: 3909-3911 (1996)）、ベンズイミダゾール（Loakesら、*Nucl. Nucl.* 18: 2685-2695 (1999)；Papageorgiouら、*Helv. Chim. Acta* 70: 138-141 (1987)）、5-フルオロインドール（Loakesら (1999)）、インドール（Girgisら、*J. Heterocycl. Chem.* 25: 361-366 (1988)）；非環式糖類似体（Van Aerschotら、*Nucl. Nucl.* 14: 105

40

50

3-1056 (1995); Van Aerschotら、Nucleic Acids Res. 23: 4363-4370 (1995); Loakesら、Nucl. Nucl. 15: 1891-1904 (1996)) (ヒポキサンチン、イミダゾール、4,5-ジカルボキサミド、3-ニトロイミダゾール、5-ニトロインダゾールの誘導体を含む); 芳香族類似体 (Guckianら、J. Am. Chem. Soc. 118: 8182-8183 (1996); Guckianら、J. Am. Chem. Soc. 122: 2213-2222 (2000)) (ベンゼン、ナフタレン、フェナンタレン、ピレン、ピロール、ジフルオロトルエンを含む); イソカルボスチリルヌクレオシド誘導体 (Bergerら、Nucleic Acids Res. 28: 2911-2914 (2000); Bergerら、Angew. Chem. Int. Ed. Engl., 39: 2940-2942 (2000)) (MICS, ICSを含む); 水素結合類似体 (N8-ピロールピロリジンを含む) (Seelaら、Nucleic Acids Res. 28: 3224-3232 (2000)); およびLNA、例えばアシル- β -C-LNA (Babuら、Nucleosides, Nucleotides & Nucleic Acids 22: 1317-1319 (2003); WO 03/020739)がある。

10

【0204】

ii. 半ユニバーサル塩基

半ユニバーサル塩基は、典型的に存在する(すなわち、DNAではA、C、G、およびT、RNAではA、C、G、およびU)ヌクレオチドの2つまたは3つに選択的に結合するが、同じまたは同様の特異性にすべての4つの典型的に存在するヌクレオチドに結合することはない。例えば半ユニバーサル塩基は、2つまたは3つの典型的に存在するヌクレオチドに、少なくとも1つの他の典型的に存在するヌクレオチドに結合するより強い親和性で結合する。本明細書で使用される半ユニバーサル塩基の例は、プリンAとG、またはピリミジンCとTに選択的にハイブリダイズする。例えばピリミジン類似体6H,8H-3,4-ジヒドロピリミド[4,5-c][1,2]オキサジン-7-オンはAまたはGに選択的にハイブリダイズし、プリン類似体N6-メトキシ-2,6-ジアミノプリンはC、TまたはUに選択的にハイブリダイズする(例えば、Bergstromら、Nucleic Acids Res. 25: 1935-1942 (1997)を参照)。

20

【0205】

b. 他の特徴

捕獲オリゴヌクレオチドの配列、長さ、および組成は、当業者に公知の種々の要因(特に限定されないが、標的核酸分子の長さ、断片化法、ハイブリダイゼーション条件、使用される異なる捕獲オリゴヌクレオチドの数と異なるヌクレオチド組成物の所望の数、および/または特定の捕獲オリゴヌクレオチドにハイブリダイズすることが好ましい配列の所望の数)に従って変動する。

30

【0206】

本明細書の具体的な実施形態において捕獲オリゴヌクレオチドのサブセットは部分的に縮重でもよい。例えば少なくとも10%、少なくとも20%、少なくとも30%、少なくとも40%、少なくとも50%、少なくとも60%、少なくとも70%、少なくとも80%、少なくとも90%、少なくとも95%の捕獲オリゴヌクレオチドが部分的に縮重である実施態様が企図される。さらに10%以下、20%以下、30%以下、40%以下、50%以下、60%以下、70%以下、80%以下、90%以下、95%以下の捕獲オリゴヌクレオチドが部分的に縮重である実施態様が企図される。本明細書の他の実施態様においてすべての捕獲オリゴヌクレオチドは部分的に縮重である。他の実施態様において、どの捕獲オリゴヌクレオチドも部分的に縮重ではない。

40

【0207】

部分的に縮重した捕獲オリゴヌクレオチドは、1つまたはそれ以上の非縮重ヌクレオチド(すなわち、DNAではA、C、G、T、およびRNAではA、C、G、U)と1つまたはそれ以上の縮重ヌクレオチドの組合せを含有することができる(例えば、捕獲オリゴヌクレオチドに取り込まれたユニバーサル塩基または半ユニバーサル塩基)。別の実施態様において部分的に縮重したオリゴヌクレオチドは縮重ヌクレオチドのみを含有し、ここで部分的に縮重したオリゴヌクレオチドは、第2のセットのヌクレオチド配列に結合する特異性より高い特異性で第1のセットのヌクレオチド配列に結合する能力を維持している。例えば部分的に縮重したオリゴヌクレオチドは、半ユニバーサル塩基のみまたは半ユニバーサル塩基とユニバーサル塩基との組合せを含有することができ、半ユニバーサル塩基の選択的結合は

50

部分的に縮重したオリゴヌクレオチドに結合特異性を付与する。

【0208】

部分的に縮重した捕獲オリゴヌクレオチドの使用は、2つ以上の特異的標的核酸配列がそれぞれ部分的に縮重した捕獲オリゴヌクレオチドに結合することを可能にし、こうして標的核酸のすべての理論的組合せを捕獲するために、捕獲オリゴヌクレオチド配列のすべての理論的組合せより少ない数がアレイ上に存在することを可能にする。特定の捕獲オリゴヌクレオチド上で使用される縮重位置の数は、単一の捕獲オリゴヌクレオチドが、切断工程で作製される種々の断片からの2つまたはそれ以上の異なる標的核酸断片に選択的にハイブリダイズできるように、選択される。

【0209】

本明細書の別のところに記載されるように、捕獲オリゴヌクレオチド配列のすべての理論的組合せより少ない数の使用で企図されるのは、ミスマッチ結合を可能にするハイブリダイゼーション条件の緊縮性の低下または緩和であり、こうして2つ以上の特異的標的核酸配列が、各部分的に縮重したかもしくは非縮重捕獲オリゴヌクレオチドに結合することが可能になり、標的核酸のすべての理論的組合せを捕獲するために、捕獲オリゴヌクレオチド配列のすべての理論的組合せより少ない数がアレイ上に存在することを可能になる。

【0210】

捕獲オリゴヌクレオチドは各標的核酸断片産物に特異的であるか、または捕獲オリゴヌクレオチドは標的核酸の2つまたはそれ以上の異なる断片の共通領域に相補的でもよい。例えば具体的なハイブリダイゼーション反応測定法において、固相に固定化された捕獲オリゴヌクレオチドは、共通のサブ断片配列を含む異なるサイズの断片化産物にハイブリダイズすることができる。さらに単一の捕獲オリゴヌクレオチドを使用して、緊縮性の低いハイブリダイゼーション条件を使用することにより、および/または捕獲オリゴヌクレオチド内の1つまたはそれ以上の縮重ヌクレオチドを使用することにより、捕獲オリゴヌクレオチドに相補的な領域で互いに1つまたはそれ以上のヌクレオチドだけ異なる配列を有する標的核酸断片を捕獲することができる。すなわち単一の捕獲オリゴヌクレオチド配列が2つ以上の標的核酸断片配列に結合できるようにするために、捕獲ヌクレオチドと緊縮性条件は経験的に選択することができる。また異なる配列を有する異なるヌクレオチド断片の数、または捕獲オリゴヌクレオチドにハイブリダイズする異なる組成を有するヌクレオチド断片の数を制御するために、捕獲ヌクレオチドと緊縮性条件は経験的に選択することができる。

【0211】

従って本明細書で使用される捕獲オリゴヌクレオチドは、接触工程または組合せ工程の条件下で本明細書で調製した標的核酸断片と半特異的にハイブリダイズするのに十分な長さおよび十分な相補性とを有するヌクレオチドの配列を含む。かかるハイブリダイゼーション（ハイブリダイゼーションは溶液中または固相上で起きる）の前、その最中、または後に、捕獲オリゴヌクレオチドは固相上に固定化され、対応する不連続の非重複要素に整列され、従って各要素は異なる捕獲オリゴヌクレオチドを含有する。固体支持体（例えばガラス、シリコン、プラスチック、ナイロン膜、多孔性材料など）の不連続の要素にオリゴヌクレオチドを整列させるために種々の材料と方法が当該分野で公知であり、接触沈着（米国特許第5,807,522号明細書；同5,770,151号明細書など）；写真平版に基づく方法（例えば米国特許第5,861,242号明細書；同5,858,659号明細書；同5,856,174号明細書；同5,856,101号明細書；同5,837,832号明細書など参照）；流路に基づく方法（例えば米国特許第5,384,261号明細書；ディップペンナノリソグラフィーに基づく方法（例えば、Pinerら、Science Jan. 29:661-663 (1999)）がある。具体的な実施態様において捕獲オリゴヌクレオチドは、各固相アレイ（例えばチップ）当たり一般に20,000以下、15,000以下、10,000以下、7,000以下、5,000以下、4,000以下、3,000以下、2,500以下、2,100以下、2,000以下、1,500以下、1,400以下、1,300以下、1,200以下、1,100以下、1,000以下、900以下、800以下、700以下、600以下、500以下、400以下、300以下、200以下、100以下の要素（位置）である対応する不連続の位置（位置）で整列される。

10

20

30

40

50

【0212】

本明細書に記載のように、本発明の方法で使用される固相アレイは、いくつかの縮重ヌクレオチドを有する捕獲オリゴヌクレオチドを含有することができる。これは、元々の標的核酸配列に含まれる情報を捕獲するのに必要なオリゴヌクレオチドの総数を減少させることができる。従って標的核酸の初期切断中に生成される同様の配列の複数の断片は、各位置で同じ捕獲オリゴヌクレオチドにハイブリダイズすることができる。複数の分子種が異なる全体的ヌクレオチド組成を有しても、質量スペクトル分析により分子量からその同定が可能である。

【0213】

本明細書のある実施態様においてユニバーサル塩基または半ユニバーサル塩基の使用は、4096個またはそれ以下の捕獲位置を有するハイブリダイゼーションチップをスクリーニングに使用することを可能にする。具体的な応用は、より少ない数のオリゴヌクレオチドで良いかも知れない。例えば本明細書の実施態様において4096個の捕獲オリゴヌクレオチドは、縮重プリン/ピリミジンにハイブリダイズする塩基について長さ12のすべての捕獲オリゴヌクレオチド（すなわち、12個の半ユニバーサル塩基を含有する12塩基の捕獲オリゴヌクレオチド）、または6個の非縮重塩基（A、C、G、T）と6個のユニバーサル塩基、またはこれらの組合せ（例えば、2個の非縮重塩基、8個の半ユニバーサル塩基、および2個のユニバーサル塩基）を有する捕獲オリゴの作製を可能にする。本実施態様は、すべての捕獲オリゴヌクレオチドを作製するために、アレイの各捕獲オリゴヌクレオチドが、同じ含量の非縮重塩基、半ユニバーサル塩基およびユニバーサル塩基を有することを必要としない。例えばいくつかの捕獲オリゴヌクレオチドは半ユニバーサル塩基のみを含有し、いくつかの捕獲オリゴヌクレオチドは非縮重塩基、ユニバーサル塩基および半ユニバーサル塩基を含有し、さらに別の捕獲オリゴヌクレオチドは非縮重塩基とユニバーサル塩基のみを含有してもよい。種々のタイプの塩基の相対量は、捕獲オリゴヌクレオチドの特異性の所望のレベルに従って当業者が測定することができる。

【0214】

別の実施態様においてハイブリダイゼーション構造体は、例えば1024個という少ない捕獲位置を有してもよい。かかるチップは、複数の試料、例えばそれぞれ異なる塩基を特異的に切断する条件で別々に処理された4つの試料（例えば、試料1をA特異的切断条件で処理し、試料2をC特異的切断条件で処理し、試料3をG特異的切断条件で処理し、試料4をT特異的切断条件で処理する）、をハイブリダイズするのに使用することができる。ある実施態様において4つの異なる切断条件で処理された同じヌクレオチドの4つの試料はハイブリダイゼーション構造体に同時にハイブリダイズされ、標的核酸質量が測定される。別の実施態様において4つの異なる切断条件で処理された同じヌクレオチドの4つの試料が4つの別のハイブリダイゼーション工程でハイブリダイズ構造体にハイブリダイズされ、ここで4つの別のハイブリダイズ工程のそれぞれの後で標的核酸質量が測定される。別の実施態様においてかかる塩基特異的切断は1本鎖核酸について選択的であり、その結果捕獲オリゴヌクレオチドプローブに結合しない標的核酸の部分は塩基特異的に切断されて、標的核酸がハイブリダイズする捕獲オリゴヌクレオチドプローブより長い標的核酸（すなわち捕獲ヌクレオチドプローブをオーバーハングしている）を与え、ここでオーバーハングの長さは、標的核酸のハイブリダイズ部分と比較して最も特異性の近い切断塩基の位置により決定される。

【0215】

c. 捕獲オリゴヌクレオチドの作製

オリゴヌクレオチドは別に合成され、次に固体支持体に結合することができるか、または合成は固体支持体の表面でインサイチュで行われる。オリゴヌクレオチドは多くの会社、例えばインテグレートッドディーエヌエーテクノロジ（Integrated DNA Technology）（IDT）、フィデリティシステムズ（Fidelity Systems）、プロリゴ（Proligo）、エムダブリュー（MWG）、オペロン（Operon）、メタピオン（MetaBion）などから購入することができる。

10

20

30

40

50

【0216】

オリゴヌクレオチドとオリゴヌクレオチド誘導体は、当該分野で公知の標準的方法、例えば自動DNA合成機（例えば、ビオサーチ（Biosearch）（ノバト、カリフォルニア州）；アプライド・バイオシステムズ（Applied Biosystems）（フォスターシティ、カリフォルニア州）などから市販されている）を固体支持体、例えば制御多孔ガラス（CPG）またはポリスチレンおよび他の樹脂などと組合せて、化学的方法、例えばホスホラミダイト法、H-ホスフェート法、またはホスホジエステル法などと組合せて使用して合成することができる。オリゴヌクレオチドはまた、溶液中または可溶性支持体上で合成することができる。例えばホスホロチオエートオリゴヌクレオチドは、Steinらの方法（Nucl. Acids Res. 16: 3209 (1988)）により合成することができ、メチルホスホン酸オリゴヌクレオチドは制御多孔ガラスポリマー支持体（Sarinら、Proc. Natl. Acad. Sci. USA 85: 7448-7451 (1988)）の使用により調製することができる。オリゴヌクレオチドはまた、本明細書に開示され当該分野で公知の増幅のための酵素的な方法、例えばPCRもしくは転写を使用して作製することができる。

10

【0217】

表面結合捕獲オリゴヌクレオチドは、標的核酸断片上の相補的領域にハイブリダイズする核酸である。捕獲オリゴヌクレオチドは一般に、例えば関連出願番号第60/372,711号（2002年4月11日出願）、60/457,847号（2003年3月24日出願）、および10/412,801号（2003年4月11日出願）に開示されたチップのチャンパー中で起きるような、標的核酸断片を作製するように起きる任意の反応に実質的に関与しない。好適なオリゴヌクレオチドは、標的ヌクレオチド配列への特異的または半特異的ハイブリダイゼーションを可能にするのに十分な数のヌクレオチドを有する。

20

【0218】

捕獲オリゴヌクレオチドは種々の長さの任意のものでもよく、標的核酸ヌクレオチド配列に結合するヌクレオチドや標的核酸ヌクレオチド配列に結合しないことが企図されるヌクレオチドを含んでよい。例えば捕獲オリゴヌクレオチドは、捕獲オリゴヌクレオチドを固体支持体に固定するヌクレオチド配列にハイブリダイズする部分、または標的核酸断片のプライマー配列に結合する部分（例えば、標的核酸ヌクレオチド配列の一部ではない転写開始部位）を含むことができる。捕獲オリゴヌクレオチドはまた、標的核酸ヌクレオチド配列に結合できるヌクレオチドを含むことができる。標的核酸配列に結合する捕獲オリゴヌクレオチドの部分は、本明細書に記載でかつ当業者に公知の要因に従って種々の長さの任意のものでもよい。典型的には捕獲オリゴヌクレオチドのこの部分は、5~30塩基の長さを含む。従って本明細書での使用が企図されるオリゴヌクレオチドの具体的な長さは、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、23、24、25、26、27、28、29、30ヌクレオチド、または所望であればそれ以上のヌクレオチドを含む。本明細書に記載のようにオリゴヌクレオチドは、相補配列のハイブリダイゼーション特異性を変化させるために、または生成されるハイブリッドの安定性を変更するために、天然のヌクレオチド、修飾ヌクレオチド、またはヌクレオチド模倣物（例えば、ユニバーサル塩基または半ユニバーサル塩基）からなってもよい。

30

【0219】

捕獲オリゴヌクレオチドの特異性は、縮重塩基または部位を捕獲オリゴヌクレオチド配列中に取り込むことにより制御することができる。例えば配列内で塩基をイノシンで置換すると、標的核酸産物中の多型部位へのユニバーサルハイブリダイゼーションが得られる [例えば、Ohtsukaら、J. Biol. Chem. 260: 2605 (1985)；Takahashiら、Proc. Natl. Acad. Sci. USA 82: 1931 (1985)を参照]。2本鎖核酸ハイブリッドの安定性は、例えばRNA（DNA標的に向けられる場合）、ロックされた核酸（LNA）[Braaschら、Chemistry & Biology 8:1-7 (2001)]、ペプチド核酸（PNA）[Armitageら、Proc. Natl. Acad. Sci. USA 94: 12320-12325 (1997)]、または他の修飾核酸誘導体を捕獲オリゴヌクレオチドの配列内または標的核酸配列内で完全にまたは部分的に使用して、大きく上昇させることができる。安定性はまた、1つまたは数個の非塩基性部位、非ハイブリダイズ性塩基誘導体、

40

50

または融解温度を低下させる核酸修飾（例えばホスホロチオエート）を取り込むことにより低下させることができる。これらの種々の公知のアプローチを使用して、ほとんどすべての配列および長さについて融解温度を所望の融解温度に調整することができる。

【0220】

オリゴヌクレオチド合成

溶液中または固体支持体上のオリゴヌクレオチド合成法は当該分野で公知である [例えば、Beaucageら、Tetrahedron Lett. 22: 1859-1862 (1981); Sasakiら、(1993) Technical Information Bulletin T-1792, Beckman Instrument; Reddyら、米国特許第5,348,868号明細書; Seligerら、DNA and Cell Biol. 9: 691-696 (1990)を参照]。

【0221】

オリゴヌクレオチドのインサイチュ合成

光指向性合成を使用するガラスおよびシリコン表面でのオリゴヌクレオチドのインサイチュ合成は当該分野で公知である [McGallら、J. Am. Chem. Soc. 119: 5081-5090 (1997); Wallraffら、Chemtech 27: 22-32 (1997); McGallら、Proc. Natl. Acad. Sci. USA 93: 13555-13560 (1996); Lipshutzら、Curr. Opin. Structural Biol. 4: 376-380 (1994); およびPeaseら、Proc. Natl. Acad. Sci. USA 91:5022-5026 (1994)を参照]。

【0222】

オリゴヌクレオチドは、化学的に誘導体化された固体支持体または官能基を有するポリマーまたはプラスチックのような固体支持体に結合させることができる。オリゴヌクレオチドは、種々の方法（写真平版法、共有結合、またはイオンの相互作用、ファンデルワールス結合、および水素結合のような非共有結合的相互作用による受動的結合を含む）により固体支持体に結合することができる。オリゴヌクレオチドは5'末端または3'末端修飾により表面に共有結合することができる。リンカーは典型的には、表面からオリゴヌクレオチドを遠ざけるために使用される。例えばオリゴヌクレオチドがその5'末端を介して結合される場合、5'修飾の前にリンカーが5'末端上に存在する。使用される典型的なリンカーには、ヘキシルエチレングリコール（1またはそれ以上の単位）およびオリゴデオキシチミジンdTn (n=5~20)がある。

【0223】

反応性官能基で化学的に誘導体化された表面にオリゴヌクレオチドを結合するために、種々の方法を使用することができる。例えばアミノ修飾オリゴは、エポキシド活性化表面と反応して共有結合を生成する [例えば、Lamtureら、Nuc. Acids Res. 22: 2121-2125 (1994)を参照]。同様にアミノ修飾オリゴヌクレオチドの共有結合は、カルボン酸修飾表面 [Stotherら、J. Am. Chem. Soc. 122: 1205-1209 (2000)]、イソチオシアネート、アミン、チオール [Penchovskyら、Nuc. Acids Res. 28: e98 1-6 (2000); Lenigkら、Langmuir 17: 2497-2501 (2001)]、イソシアネート [Lindroosら、Nuc. Acids Res. 29: e69 1-7 (2001)]、およびアルデヒド修飾表面 [Zammatteoら、Anal. Biochem. 280: 143-150 (2000)]上で行うことができる。

【0224】

典型的にはシリコン表面は、本明細書に記載のオリゴヌクレオチドの固定化後に化学的誘導体化することができる [Bentersら、Nuc. Acids Res. 30: e10 1-7 (2002)を参照]。例えば表面を洗浄後、表面をアミノプロピルトリメトキシシランで処理して表面上にアミノシロキサン層を設けてもよい。表面は2官能性架橋剤である1,4-フェニレンジイソチオシアネートを用いて活性化される。この架橋剤の1つのイソチオシアネート基は表面上のアミノ官能基と反応して、安定なチオ尿素結合を生成する。表面に結合した第2のイソチオシアネート基が、アミノ基を有する他の分子との共有反応のために開放されている。続く工程で、 dendritic ポリアミン、例えばスターバースト (Starburst) (PAMAM) dendritic ポリアミン (64個の末端アミノ基を有する第4世代) は活性化表面と反応して、固体支持体上で密な量の共有結合アミノ基と均一な中間層を形成する。表面上のこれらの官能基は、1,4-フェニレンジイソチオシアネートを用いて活性化される。未反応アミンは4-ニトロ-フェニレンジイソチオシアネートでブロックされる。アミノ修飾オリゴヌク

10

20

30

40

50

レオチドは、同じ種類の反応で共有架橋されて活性化デンドリマー中間層になる。最後の工程で、未反応のイソチオシアネートは、小さい1級アミン（例えばヘキシルアミン）でブロックされる。

【0225】

捕獲オリゴヌクレオチドは複数の不連続の既知の位置またはアレイ位置で固体支持体に結合される。各位置は、同じ配列を有するオリゴヌクレオチドの複数のコピーを含有することができる。例えば捕獲オリゴヌクレオチドプローブのアレイは特定の位置にオリゴヌクレオチドの複数のコピーを有することができ、ここで特定の位置のすべてのオリゴヌクレオチドは同じヌクレオチド配列を有し、特定の位置の捕獲オリゴヌクレオチドのヌクレオチド配列は、アレイ上の他の位置の捕獲オリゴヌクレオチドのヌクレオチド配列と比較してユニークである。すなわち特定のアレイ位置のすべてのオリゴヌクレオチドが同じ配列を有し、異なるアレイ位置のすべてのオリゴヌクレオチド配列がユニークであるように、アレイを構成することができる。

10

【0226】

あるいは各位置は、異なる配列を有するオリゴヌクレオチドを有してもよい。このオリゴヌクレオチドの配置は、例えば多重化反応で使用することができる。同じ位置で異なる配列のオリゴヌクレオチドは混合するか、または同様の配列の群に分離することができる。例えば2つ、3つ、4つ、またはそれ以上の異なるオリゴヌクレオチドが同じ位置にあってもよい。使用される異なるオリゴヌクレオチドの数は、1つの位置内で各異なる配列に結合した生成物を分離する能力によってのみ限定される。

20

【0227】

固体支持体上の異なる位置は典型的には、異なる配列のオリゴヌクレオチドを含有することができる。ある位置のオリゴヌクレオチドは典型的には、 $0.0025 \text{ mm}^2 \sim 1.0 \text{ mm}^2$ の面積を占め、オリゴヌクレオチド量は $10 \text{ amol} \sim 10 \text{ pmol}$ の範囲である。ある実施態様において典型的なフォーマットは、 $20 \times 30 \text{ mm}$ のサイズを有する固体支持体で、96,384または1536位置が 8×12 、 16×24 、または 32×48 のパターン中に、反応プレート上のものと同等の間隔（中心から中心が 2.25 mm 、 1.125 mm 、または 0.5625 mm ）で存在する。他の態様は最大4096の位置を有してもよい。ある実施態様において位置は、質量スペクトル分析のあるタイプで使用されるほぼレーザーの直径であり、例えばいくつかの位置はレーザーの直径より大きくは無い。固体支持体のサイズ、位置とパターン（ここに位置が配置される）の総数は、固体支持体上にアレイを作製するために、液体を扱うために、および/または分析のために使用される設計態様と装置に配置させることができる。例えば間隔とスポットのサイズは、アレイを作製する装置の正確性および/または液滴サイズにより規定される。固体支持体上の横の列または縦の列に置かれるオリゴヌクレオチドの位置の数は、MALDI-TOF質量スペクトル計のレーザーが、同時に2つ以上の位置は含むものではないようにすることができる。

30

【0228】

捕獲オリゴヌクレオチドの群は、固体支持体表面上に任意の配置で置くことができる。例えばオリゴヌクレオチドは、固体支持体中に作製した個々のウェルまたはチャンパー中に配置することができる。固体支持体上に存在するウェルの数は固体支持体のサイズにより変動し、96または384のフォーマットがしばしば使用され、最大4096またはそれ以上が容易に利用できる。典型的にはウェルまたはチャンパーは不連続のままであり、その状態を保持する。ある例ではオリゴヌクレオチドは、共通の重層試薬チャンネルを共有する横の列または縦の列に不連続の既知の位置で固体支持体上に置くことができる。別の例ではオリゴヌクレオチドはまた、完全な平面に不連続の既知の位置および任意の配置で置くことができる。位置はまた、個々のオリゴヌクレオチドまたはオリゴヌクレオチドの混合物を有する小さい領域に分割することができる。固体支持体上に配置された同じかまたは異なる材料からできたマスクを用いて、試薬のチャンネルまたはウェルを作製してもよい。さらに固体支持体上のウェルおよびチャンネルは、そのサイズに従って不連続のおよび分類されたビーズを局在化または平坦化するように設計することができる。この設計ではビーズは

40

50

、反応生成物である核酸断片および誘導体の捕獲のために使用されるオリゴヌクレオチドの担体である。

【0229】

F. 固体支持体とアレイ

本明細書に記載の方法は、配列決定すべき標的核酸の断片の固体支持体上への捕獲を利用する。固体支持体は、化学的および生物学的分子合成と分析のための親和性マトリックスまたは支持体として使用される任意の材料、例えば特に限定されないが、ポリスチレン、ポリカーボネート、ポリプロピレン、ナイロン、ガラス、金属、磁性ビーズ、ラテックス、デキストラン、キチン、砂、軽石、アガロース、多糖、デンドリマー、バックボール、ポリアクリルアミド、シリコン、ゴム、および固相合成、親和性分離と精製、ハイブリダイゼーション反応、イムノアッセイ、および他のかかる応用のための支持体として使用される他の材料がある。ここで固体支持体は粒子状であるか、または連続的表面の型、例えばコーティングされたピンツール、マイクロタイターディッシュもしくはウェル、ガラススライド、金属、プラスチックもしくはシリコンチップ、ニトロセルロースシート、ナイロンメッシュ、多孔性3次元構造体（例えば多孔性3次元ゲル）、またはかかる他の材料でもよい。粒子状である場合は、粒子は典型的には5~10 mmまたはそれ以下の範囲の少なくとも1つの寸法を有する。本明細書においてまとめて「ビーズ」と呼ぶかかる粒子は、しばしば（必ずではないが）球状である。しかしかかる参照が固体支持体の形を制限するものではなく、ランダムな形、針状、繊維状、および楕円形のような任意の形でよい。ほぼ球状の「ビーズ」、特に液相で使用できる微小球もまた企図される。「ビーズ」は、追加の成分が本明細書の方法やその分析を妨害しない限り、磁石を使用して分離するための磁性粒子もしくは常磁性粒子（例えば、ダイナビーズ（Dynabeads）TM（ダイナール（DynaI））、オスロー、ノルウェー））を含有してもよい。

10

20

【0230】

例えば具体的な実施態様において、関連する米国特許出願第60/372,711号（2002年4月1日出願）、60/457,847号（2003年3月24日出願）、および10/412,801号（2003年4月11日出願）に記載されたハイブリダイゼーションチップが、捕獲オリゴヌクレオチドのアレイの固体支持体として使用され、例えばチャンバーの内部底面上の固相固体支持体の表面上（この上で標的核酸断片生成反応が行われる）で捕獲オリゴヌクレオチドにより標的核酸断片が捕獲される。具体的な実施態様において、チャンバー（標的核酸断片化産物と特異的にハイブリダイズすることができる固体支持体を含有するかまたはその底が固体支持体である）から他の分子を除去または洗浄するのに使用される工程中は、標的核酸断片化産物を固体支持体に付着維持するように、断片化反応はチャンバー中で行われる。相互作用は、標的核酸断片化産物と固体支持体（例えば誘導体化されたかまたは官能基化された固体支持体）上に固定化された捕獲オリゴヌクレオチドとの間でもよい。標的核酸断片化産物の特異的捕獲を行う任意のタイプの固体支持体を使用することができる。

30

【0231】

例えば固体支持体は、平らな2次元表面または3次元表面またはビーズでもよい。平らな固体支持体の場合は、本明細書の装置例に記載された「マスク」により提供されるような固体支持体表面から延びる壁により、または不連続な単離されたチャンバーを形成するように固体支持体表面にウェルまたはピラーまたはチャンネルをエッチングすることにより作製される壁により、チャンバーを形成することができる。固体支持体を作製することができる材料には、特に限定されないがシリコン、上に酸化層を有するシリコン、ガラス、金属（例えば白金または金）、ポリマー（例えばポリアクリルアミド）、およびプラスチックがある。具体的な実施態様において固体支持体はシリコンチップまたはウェーハーである。

40

【0232】

平らな固体支持体はまた、チャンバー中の反応混合物の温度調節を容易にするための伝熱性材料を含むように修飾することができる。具体的な実施態様において固体支持体は、金属材料でコートされた平らなシリコンチップである。固体支持体の例は本明細書に記載

50

され、本明細書に記載された装置や方法とともに使用することができる。

【0233】

上記したように捕獲オリゴヌクレオチドは、一般に各固体支持体（例えばチップ）当たり20,000以下、15,000以下、10,000以下、7,000以下、5,000以下、4,000以下、3,000以下、2500以下、2100以下、2000以下、1500以下、1400以下、1300以下、1200以下、1100以下、1000以下、900以下、800以下、700以下、600以下、500以下、400以下、300以下、200以下、100以下の不連続要素である数の位置（場所）で、対応する不連続要素に整列される。さらなる実施態様においてアレイは、捕獲オリゴヌクレオチドを有する4096またはそれ以下、1536またはそれ以下、384またはそれ以下、96またはそれ以下、64またはそれ以下の不連続の位置を含有する。具体的な実施態様において捕獲オリゴヌクレオチドのアレイは4096の捕獲オリゴヌクレオチドを含有する。アレイが4096のオリゴヌクレオチドを含有するある実施態様において、捕獲オリゴヌクレオチドは12塩基の長さでもよい。4096のオリゴヌクレオチドを使用する別の実施態様において捕獲オリゴヌクレオチドは、30塩基の長さ、25塩基の長さ、20塩基の長さ、15塩基の長さ、10塩基の長さ、9塩基の長さ、8塩基の長さ、7塩基の長さ、6塩基の長さでもよい。

10

【0234】

具体的な実施態様において固体支持体上のすべての捕獲オリゴヌクレオチドは完全にまたは部分的に縮重しており、例えばこれらは少なくとも1つのユニバーサル塩基または半ユニバーサル塩基をその中に含有する。別の実施態様において固体支持体は、完全に縮重した、部分的に縮重した、および/または非縮重の捕獲オリゴヌクレオチドをその中に含有してもよい。非縮重捕獲オリゴヌクレオチドは、その中に縮重塩基（ユニバーサル塩基または半ユニバーサル塩基）を含有しないものである。

20

【0235】

捕獲オリゴヌクレオチドのアレイは、捕獲オリゴヌクレオチドの所望の性質に従って種々の方法で設計することができる。アレイを構成する捕獲オリゴヌクレオチドは、長さ、配列、組成、または2本鎖部分の有無、およびこれらの組合せが変化してもよい。例えばアレイは、すべての1本鎖捕獲オリゴヌクレオチドの長さが12塩基で、捕獲オリゴヌクレオチド当たり6つの塩基を含むように設計することができる。あるいはアレイは、種々の異なる長さおよび/または異なる異なる組成（例えば、異なる数のユニバーサル塩基および/または半ユニバーサル塩基）の50%の1本鎖オリゴヌクレオチドと50%の2本鎖オリゴヌクレオチド、または両方を含有するように設計することができる。例えばアレイは、長さが6~18塩基で変化し、さらにもしくはその代わりに、6~12個のユニバーサル塩基または半ユニバーサル塩基を含有する捕獲オリゴヌクレオチドを含有するように設計することができる。

30

【0236】

典型的には捕獲オリゴヌクレオチドプローブのアレイは、4ヌクレオチドもしくはそれ以上の長さ、5ヌクレオチドもしくはそれ以上の長さ、6ヌクレオチドもしくはそれ以上の長さ、7ヌクレオチドもしくはそれ以上の長さ、8ヌクレオチドもしくはそれ以上の長さ、10ヌクレオチドもしくはそれ以上の長さ、12ヌクレオチドもしくはそれ以上の長さ、または15ヌクレオチドもしくはそれ以上の長さの捕獲オリゴヌクレオチドプローブを含有する。さらに捕獲オリゴヌクレオチドプローブの典型的なアレイは、50塩基以下の長さ、40塩基以下の長さ、35塩基以下の長さ、30塩基以下の長さ、25塩基以下の長さ、20塩基以下の長さ、18塩基以下の長さ、16塩基以下の長さ、14塩基以下の長さ、12塩基以下の長さ、10塩基以下の長さ、または8塩基以下の長さの捕獲オリゴヌクレオチドプローブを含有する。さらに捕獲オリゴヌクレオチドプローブは、3'末端、5'末端、または3'末端と5'末端の両方に1つまたはそれ以上の追加の縮重塩基を有することができる。

40

【0237】

設計されたアレイ中の捕獲オリゴヌクレオチドの2本鎖部分のサイズ、組成、および有無は、種々の設計目的により選択することができる。ある実施態様においてアレイは、それぞれが同じ緊縮性条件下でほぼ同じ数の異なる標的核酸配列にハイブリダイズするアレ

50

イを含有するように設計することができる。例えばアレイは、同じハイブリダイゼーション条件下（例えば同じ融解温度を有する）でそれぞれが完全に相補的な配列とハイブリダイズする捕獲オリゴヌクレオチドを含有するように設計することができる。これは、例えば同じ(A+T)/(C+G)比を有するプライマーを設計することにより、A/Tリッチな捕獲オリゴヌクレオチドより短いC/Gリッチな捕獲オリゴヌクレオチドを作製することにより、捕獲オリゴヌクレオチドの長さを変化させることにより、ユニバーサル塩基または半ユニバーサル塩基を含有させることにより、または2本鎖領域を有する捕獲オリゴヌクレオチドを含有させることにより行われる。別の例ではアレイは、異なる融解温度を有するが特定の条件下で同じ数の異なる標的核酸にハイブリダイズする捕獲オリゴヌクレオチドを用いて設計することができる。例えばより高い融解温度を有する捕獲オリゴヌクレオチドは、より低い融解温度を有する捕獲オリゴヌクレオチドと比較して、長さが短いかまたはより多くのユニバーサル塩基または半ユニバーサル塩基を含有することができる。従ってあるハイブリダイゼーション条件下で捕獲オリゴヌクレオチドは、ほぼ同じ数の異なる標的核酸配列にハイブリダイズすることができる。例えば標的核酸断片にハイブリダイズする第1の捕獲オリゴヌクレオチドの部分は数個のみのヌクレオチドを含有するが、このヌクレオチドは主にGとCであって、第1の捕獲オリゴヌクレオチドにハイブリダイズしない標的核酸の部分中の標的核酸配列は制限されないため、結合した種々の異なる標的核酸断片を与える；第2の捕獲オリゴヌクレオチドについて、標的核酸断片とハイブリダイズする部分はより多くのヌクレオチドを含有することができるが、このヌクレオチドは、GやCより弱くハイブリダイズするユニバーサル塩基または半ユニバーサル塩基を含むことができ、捕獲オリゴヌクレオチドに結合する標的核酸配列は捕獲オリゴヌクレオチド中の縮重塩基の数に従って変化することができるため、結合した種々の異なる標的核酸断片を与える；その結果、特定のハイブリダイゼーション条件下で第1および第2の捕獲オリゴヌクレオチドにハイブリダイズする異なる標的核酸配列の総数はほぼ同じである。

10

20

30

40

50

【0238】

あるいは設計されたアレイ中の捕獲オリゴヌクレオチドのサイズと組成は、選択されたハイブリダイゼーション条件下で、異なる捕獲オリゴヌクレオチドが異なる標的核酸の異なる数にハイブリダイズするように選択することもできる。例えば第2の捕獲オリゴヌクレオチドが10個の異なる標的核酸とハイブリダイズするのと同じ条件下で、第1の捕獲オリゴヌクレオチドが20個の異なる標的核酸とハイブリダイズするように設計することができる。例えば第1の捕獲オリゴヌクレオチドが6個の非縮重塩基と6個のユニバーサル塩基を含有し、第2の捕獲オリゴヌクレオチドが第1の捕獲オリゴヌクレオチドと同じ6個の非縮重塩基と2つの追加の非縮重塩基を含有することができる；その結果、第1の捕獲オリゴヌクレオチドに結合する標的核酸のサブセットのみが第2の捕獲オリゴヌクレオチドにも結合する。

【0239】

設計されたアレイ中の捕獲オリゴヌクレオチドのサイズ、組成、およびヌクレオチド配列は、1つまたはそれ以上の以下の基準を満足するように選択することもできる；例えばSNPまたはマイクロサテライトのような特定のタイプの配列を標的とする；ランダム配列または未知配列を標的とする；異なる領域の標的核酸の複雑さを制御する（例えば、いくつかの標的核酸の末端配列部分の複雑さを制御するために、2本鎖となつたいくつかの捕獲オリゴヌクレオチドを有することにより）；および特定の捕獲オリゴヌクレオチドにハイブリダイズする重複断片の数を増加または減少させる（例えば、多くのユニバーサル塩基または半ユニバーサル塩基を使用して減少させるか、または、2本鎖領域を持たず場合により1端または両端以外の位置にユニバーサル塩基を持たないより短い特異的配列を使用して増加させる）。

【0240】

G. 特異的または非特異的ハイブリダイゼーション

本明細書の方法は、典型的には2つまたはそれ以上の核酸分子にハイブリダイズする工程を含む。本方法において捕獲オリゴヌクレオチドは1つまたはそれ以上の標的核酸分子

またはその断片にハイブリダイズして、「捕獲オリゴヌクレオチド：標的断片複合体」または「捕獲オリゴヌクレオチド：標的核酸複合体」を形成することができる。かかる複合体はしばしば2本鎖複合体（すなわち2本鎖）であるが、3本鎖複合体でもよい。

【0241】

ハイブリダイゼーションの程度と特異性は、反応条件、特に温度と塩濃度とともに変化する。ハイブリダイゼーション反応条件は、典型的には緊縮性の程度（例えば低、中、および高緊縮性；これらは、当業者に公知で本明細書に開示された異なる温度と塩濃度下で行われる）で参照される。すなわち例えばある実施態様において、ハイブリダイズする核酸間の不完全一致の量を低下させるために、より高緊縮性条件、例えば高い温度および/または低い塩濃度を使用することができる。逆に、ハイブリダイズする核酸間の不完全一致の量を増加させるために、より低緊縮性条件、例えば低い温度および/または高い塩濃度を使用することができる。

10

【0242】

具体的な実施態様において、標的核酸断片にハイブリダイズするのに使用される捕獲オリゴヌクレオチドは完全な塩基特異性ではハイブリダイズせず、従ってミスマッチハイブリダイゼーションまたはハイブリダイゼーションの縮重を排除しない。これはハイブリダイゼーション緊縮性を低下させることを可能にし、従ってヌクレオチド捕獲配列の必ずしもすべての理論的組合せがチップアレイ上で示される必要は無い。本明細書に記載のように捕獲オリゴヌクレオチドの縮重とハイブリダイゼーション緊縮性条件は、固体支持体上で4096またはそれ以下の少ない捕獲オリゴヌクレオチドを可能にするように経験的に変化させることができる。ミスマッチ断片の組成と配列は、次の質量スペクトル分析で分子質量を得ることにより同定することができる。

20

【0243】

本明細書に記載の方法で有利に使用されるミスマッチハイブリダイゼーションの量は、かかるミスマッチハイブリダイゼーションを排除することを目的とする条件下で、典型的なSBH法で起きるミスマッチハイブリダイゼーションの好ましくない量より有意に多い。例えば本明細書に記載の方法に従って使用される捕獲オリゴヌクレオチドは、そこにハイブリダイズした2つまたはそれ以上の核酸断片を有することができる。ある場合は2つまたはそれ以上の標的核酸断片が完全な相補性で捕獲オリゴヌクレオチドにハイブリダイズすることができる；かかる例は、2つまたはそれ以上の縮重ヌクレオチドを含有する捕獲オリゴヌクレオチドにハイブリダイズした2つまたはそれ以上の標的核酸断片、または捕獲オリゴヌクレオチドより長く、捕獲オリゴヌクレオチドにハイブリダイズしていない断片の部分に従って配列が変化する2つまたはそれ以上の標的核酸断片である。他の例では、2つまたはそれ以上の標的核酸断片が捕獲オリゴヌクレオチドにハイブリダイズできるような、緊縮性が低下したハイブリダイゼーション条件を選択することができる。そのような例では2つまたはそれ以上の標的核酸断片が捕獲オリゴヌクレオチドにハイブリダイズできるような低い緊縮性を有するようにハイブリダイゼーション条件が選択され；かかる例では、1つまたはそれ以上の標的核酸断片が、相補性が完全ではない捕獲オリゴヌクレオチドにハイブリダイズすることが好ましい。捕獲オリゴヌクレオチドにハイブリダイズした標的核酸断片の得られる混合物の例には、どの特定の標的核酸断片も、捕獲オリゴヌクレオチドにハイブリダイズした標的核酸断片の混合物中に、混合物中に95%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、または25%より多い標的核酸断片として存在しない標的核酸断片の混合物がある。別の例では生じる混合物は、少なくとも2つ、少なくとも3つ、少なくとも4つ、または少なくとも5つの標的核酸断片が、捕獲オリゴヌクレオチドにハイブリダイズした標的核酸分子の5%、10%、15%、または20%を超える量で存在する標的核酸断片の混合物がある。別の例では標的核酸断片は、捕獲オリゴヌクレオチドにハイブリダイズした標的核酸断片の混合物中で少なくとも1つの他の標的核酸断片の量より、2倍より多い量、3倍より多い量、4倍より多い量、または5倍より多い量では存在しない（すなわち、最も多い標的核酸断片に対して、最も多い断片の量の少なくとも50%、33%、25%、または20%の量で少なくとも1つの他の断片

30

40

50

が存在する)。

【0244】

具体的な実施態様において捕獲オリゴヌクレオチドは、各チップ位置(典型的には、同じ捕獲オリゴヌクレオチドの複数のコピーを有する)が2つまたはそれ以上の標的核酸断片に結合するように設計される。例えば2~500、2~400、2~300、2~250、2~200、2~150、2~100、2~75、2~50、2~40、2~30、2~25、2~20、2~15、2~10、または2~5の異なる標的核酸断片が、捕獲オリゴヌクレオチドの1つの分子種に結合するような条件が企図される。そのような例では異なる標的核酸断片は、他の断片のサブ断片(例えば断片のラダーを作製する)である断片の結合、ならびに特定のチップ位置および捕獲オリゴヌクレオチドと同じかまたは異なる長さおよび同様のハイブリダイゼーション性を有するが、異なるヌクレオチド組成を有する断片の結合も含む。

10

【0245】

ある実施態様において2つまたはそれ以上の異なるハイブリダイゼーション反応(例えば、標的核酸断片が接触する2つまたはそれ以上の別個の位置を有するアレイ)を含む方法は、2つまたはそれ以上のハイブリダイゼーション反応のすべて(例えばアレイ位置)が、そこに2つまたはそれ以上の標的核酸断片がハイブリダイズした捕獲オリゴヌクレオチドを与えることを必要としない。ある例ではいくつかの反応(例えばアレイ位置)は、そこにハイブリダイズした標的核酸断片を含有しない。他の例ではいくつかの反応(例えばアレイ位置)は、そこにハイブリダイズした1つだけの標的核酸断片を含有する。典型的にはすべての反応の少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、または少なくとも99%が、捕獲オリゴヌクレオチドにハイブリダイズした2つまたはそれ以上のオリゴヌクレオチドを与え、ここで2つまたはそれ以上の捕獲オリゴヌクレオチドの相対量は本明細書に記載したレベルで存在する。

20

【0246】

ハイブリダイゼーション効率を上昇させるために、捕獲オリゴヌクレオチドをユニバーサル塩基により伸長することができる。例えば捕獲オリゴヌクレオチドは2つの領域(ユニバーサル塩基のみを含有する第1の領域、および少なくとも1つの典型的に存在する塩基もしくは半ユニバーサル塩基を含有する第2の領域)を含有することができる。第2の領域は、標的核酸と特異的または半特異的にハイブリダイズするために使用される塩基を含有し、第1の領域のユニバーサル塩基は捕獲オリゴヌクレオチドと標的核酸間のハイブリダイゼーションを安定化させるように機能する。

30

【0247】

さらに、複数の標的核酸が単一の捕獲オリゴヌクレオチドにハイブリダイズすることができるため、捕獲オリゴヌクレオチドは捕獲オリゴヌクレオチドの配列認識部分の縮重塩基を取り込んで縮重捕獲オリゴヌクレオチドを与えることができる。チップアレイ位置の総数を少なく維持すると、縮重捕獲オリゴヌクレオチドの配列認識部分の長さおよび/または特異性は限定される。ある実施態様において標的長さが12ヌクレオチドの捕獲オリゴヌクレオチドは4096個の位置に置くことができるであろう。従って捕獲オリゴヌクレオチドの一端にユニバーサル塩基をさらに付加するとハイブリダイゼーション複合体の安定性が大幅に上昇し、捕獲オリゴヌクレオチドの配列特異性を修飾することなく全体の効率が上昇する。ある実施態様においてさらなる修飾に依存して、これらの追加のユニバーサルヌクレオチドは、捕獲オリゴヌクレオチドの3'末端に置くことができるであろう。別の実施態様においてこれらのユニバーサルヌクレオチドは、捕獲オリゴヌクレオチドの5'末端に置くことができるであろう。別の実施態様において追加のユニバーサルヌクレオチドは、捕獲オリゴヌクレオチドの両端に置くことができる。

40

【0248】

情報量およびシステムの柔軟性と頑強性を向上させるために、または系の組成の複雑さを低下させるために、ハイブリダイズ断片へのさらなる修飾が可能である。例えば固相ア

50

レイ上の捕獲オリゴヌクレオチド：標的断片2本鎖を1本鎖特異的RNaseまたはDNaseで処理すると（「トリミング反応」）、ハイブリダイズ断片の全体的長さがより均一な長さまで低下する。トリミングの使用は、初期断片化条件の選択に影響を与える。例えば初期ランダム断片化法中に負荷される制限を緩和することができ、断片サイズの上限を上昇させることができる。35塩基またはそれ以上サイズのハイブリダイズ断片は、捕獲オリゴの長さまでおよび/またはMALDI-MSにより容易に検出できる長さまで短縮することができる。種々の配列についてのシステムの柔軟性を改良するために、断片化パラメータの緩和が本明細書で企図される。さらに塩基特異的RNaseまたはDNase（「塩基特異的トリミング」）を使用することができ、これは必ずしもハイブリダイズ断片を捕獲オリゴの正確な長さまで短縮するものではなく、標的核酸断片を捕獲オリゴに最も近い標的塩基まで短縮することができる。かかる塩基特異的切断はヌクレオチド中の4つの塩基の任意のものを標的とすることができる。従って特定の塩基特異的切断反応に従って4つの異なる断片の1つに修飾されているものと同じハイブリダイズ断片を与える。

10

20

30

40

50

【0249】

捕獲オリゴヌクレオチドを標的断片とハイブリダイズさせる工程は、対応しない標的核酸断片への捕獲オリゴヌクレオチドの相対的親和性を排除しながら、対応する標的核酸断片への捕獲オリゴヌクレオチドの所望のハイブリダイゼーションレベルを充分与える対応する標的核酸断片についての、捕獲オリゴヌクレオチドの相対的親和性を選択的に制御することを含む。本明細書に記載のようにある実施態様において、緊縮性条件は捕獲オリゴヌクレオチド：標的断片2本鎖中の1つまたはそれ以上のミスマッチを許容するように選択される。すなわち特定の捕獲オリゴヌクレオチドに対応する標的断片は、そこに正確な相補的配列を含有するのみではなくそこに少なくとも1つのまたはそれ以上のヌクレオチドミスマッチを有する標的核酸断片を含むことができる。凝集物中でミスマッチ標的核酸に対する捕獲オリゴヌクレオチドの相対的親和性は、一般に1つまたはそれ以上のミスマッチの標的核酸断片（例えば、捕獲オリゴヌクレオチドと標的核酸間に少なくとも1つの塩基ミスマッチを有する）への捕獲オリゴヌクレオチドの結合の、完全に相補的な標的核酸断片への捕獲オリゴヌクレオチドの結合に対する比として測定される。この比の増加は、完全に一致したオリゴヌクレオチドへの捕獲オリゴヌクレオチドの結合に対する、ミスマッチ標的核酸断片への捕獲オリゴヌクレオチドの結合の増加を意味する。従って本明細書で使用される比は変化可能であり、一般に少なくとも約0.5倍（すなわち捕獲オリゴヌクレオチドプローブは、結合した各2つの完全に相補的な標的核酸について1つのミスマッチ標的核酸に結合する）、少なくとも約1倍、少なくとも約1.5倍、少なくとも約2倍、少なくとも約3倍、少なくとも約5倍、少なくとも約7倍、少なくとも約10倍、少なくとも約15倍、または少なくとも約20倍である。当業者は、種々の因子（試験される標的核酸の長さ、異なる標的核酸断片の長さ数、測定される質量ピークを分解する能力、および標的核酸の核酸配列の測定に測定された質量ピークを使用する能力を含む）に基づいて比率を選択することができる。

【0250】

対応する標的核酸（例えば、特異的または半特異的親和性を有する捕獲オリゴにより結合される標的核酸）に対する各捕獲オリゴヌクレオチドの相対的親和性を調節するために、種々の方法またはアッセイ条件を使用することができる。ある具体的な実施態様において対応する標的核酸に対する各捕獲オリゴヌクレオチドの相対的親和性は、少なくとも一部は、アッセイプローブと形成されるハイブリッドの融解温度を標準化する試薬をハイブリダイゼーション工程に含める工程、特に対応する標的核酸と他の対応しない標的核酸との所望の区別をするのに十分な標的核酸と捕獲オリゴヌクレオチド間で形成されるハイブリッドの融解温度を標準化する工程を含む方法により増加する。界面活性剤（例えばドデシル硫酸ナトリウム、ツイーン）、変性剤（例えば、グアニジン、四級アンモニウム塩）、ポリ陽イオン（例えばポリリジン、スペルミン）、小溝結合剤（例えば、ディスタマイシン、CC-1065、Kutyavinら、1998、米国特許第5,801,155号明細書参照）などを含む種々の適当な標準化試薬およびその使用は、本明細書に記載され、および/または当該分野で

公知である。有効な濃度と適当なアッセイ条件は、経験的に容易に決定される（例えば、以下の実施例を参照）。

【0251】

具体的な実施態様において変性剤は、四級アンモニウム塩、例えば塩化テトラメチルアンモニウム、塩化テトラエチルアンモニウム、フッ化テトラメチルアンモニウム、またはフッ化テトラエチルアンモニウムである。融解温度の標準化は、任意の便利な手段、例えば融解温度の変動係数（CV）または標準偏差の低下により確認することができる。例えば融解温度は、少なくとも20%、少なくとも40%、少なくとも60%、または少なくとも80%のCVまたは標準偏差の低下により標準化することができる。完全な一致と1塩基ミスマッチのシグナル間の比の上昇は、あまり厳しくないCVが必要であることを示す。一致とミスマッチの以下の例の比を与える緊縮性条件が本明細書で企図され、2:1の一致対ミスマッチ、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、15:1、20:1などの一致対ミスマッチの比を含む。5:1の一致対ミスマッチの比の例では、CVは20%またはそれ以下、および10%またはそれ以下が好ましく、50:1の比では、CVは50%またはそれ以下が好ましい。

10

【0252】

特定の捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸配列の数の制御は、ユニバーサル塩基または半ユニバーサル塩基の使用、またはハイブリダイゼーション条件の修飾、またはこれらの両方により行うことができる。ユニバーサル塩基組成とハイブリダイゼーションの使用は、特定のオリゴヌクレオチドプローブにハイブリダイズする標的核酸配列の数を制御するための2つの別個の独立した方法である。当業者は、捕獲オリゴヌクレオチドにハイブリダイズする標的核酸断片の所望の複雑さに基づき、ユニバーサル塩基または半ユニバーサル塩基を使用するか、またはハイブリダイゼーション条件を修飾するか、またはこれらの両方を選択することができる。

20

【0253】

ユニバーサル塩基は、同じかまたは類似の親和性を有する捕獲オリゴヌクレオチドに塩基対合可能な異なる標的核酸配列の理論数を制御するのに使用でき、また配列特異性の無い捕獲オリゴヌクレオチドと塩基対合する標的核酸の部分上の位置を決定するのに有用である。例えば捕獲プローブ中の2つのユニバーサル塩基の使用は、16個の異なる標的核酸配列が、類似の親和性を有する捕獲プローブと塩基対合することを可能にし、非ユニバーサル塩基の捕獲オリゴヌクレオチド上の位置がわかる。すなわち捕獲オリゴヌクレオチドと塩基対合する標的核酸配列の数を制御することができ、ヌクレオチド配列が変化できる標的核酸上のヌクレオチドの位置を知ることができる。

30

【0254】

ハイブリダイゼーション条件の操作は、捕獲オリゴヌクレオチドプローブに実際にハイブリダイズする異なる標的核酸配列の所望の数を得るために、ユーザーがハイブリダイゼーション条件を容易に修飾することを可能にする。例えば特定のハイブリダイズ条件下で捕獲オリゴヌクレオチドプローブにハイブリダイズする異なる標的核酸配列の数は、実験的に決定することができる。かかる実験的測定後に、所望であればハイブリダイゼーション条件を緩和して、オリゴヌクレオチドプローブを捕獲するための種々の異なる標的核酸断片のより多くのハイブリダイゼーションを可能にするか、またはハイブリダイゼーション条件をより緊縮性にして、捕獲オリゴヌクレオチドにハイブリダイズする異なる標的核酸断片の数を減少させることができる。捕獲オリゴヌクレオチドプローブにハイブリダイズする所望の数の異なる標的核酸断片を与えるハイブリダイゼーション条件を選択するために、ハイブリダイゼーション条件を数回変化させることができる。

40

【0255】

標的核酸断片への捕獲オリゴヌクレオチドの非特異結合を排除するための緊縮性条件、および高、中、または低緊縮性に実質的に等しい条件は以下を含む：

- 1) 高緊縮性；0.1×SSPE、0.1% SDS、65
- 2) 中緊縮性；0.2×SSPE、0.1% SDS、50

50

3) 低緊縮性 ; 1.0 × SSPE、0.1% SDS、50 、
ここでSSPEは、約150 mM NaCl、10 mM NaH₂PO₄、1 mM EDTA、pH 7.0、またはこれらと同等の化合物を含有する。

【0256】

別の緩衝液、塩、および温度を使用して同等の緊縮性が達成されることは理解される。具体的な実施態様において1つまたはそれ以上の捕獲オリゴヌクレオチドへの2つ以上の特異的標的核酸断片配列の捕獲を可能にするために、そこに数個~0個の縮重ヌクレオチドを有する捕獲オリゴヌクレオチドのために、ハイブリダイゼーション緊縮性条件を中または低緊縮性に緩和することができる。同様に、捕獲オリゴ内にいくつかの縮重オリゴヌクレオチドが含有される時、ハイブリダイゼーション条件をより緊縮性にすることができ、例えばハイブリダイゼーション条件を高緊縮性条件にすることができる。この条件は、ミスマッチハイブリダイゼーションが完全には排除されず、同時に断片化標的核酸のサブセットのみが特定の捕獲オリゴに結合できるように経験的に選択することができ、緊縮性条件を修飾して、結合する標的核酸断片のサブセットの所望のサイズを達成することができる。

10

【0257】

ある実施態様においてハイブリダイゼーション条件は、初期ハイブリダイゼーション条件から変化させることができる。変化は、ハイブリダイゼーション条件の緊縮性の低下または上昇でもよい。例えば、まず低緊縮性ハイブリダイゼーション条件下で、次にハイブリダイゼーション条件を中または高緊縮性ハイブリダイゼーション条件に上げてハイブリダイゼーションを行うことができる。別の例では、まず高緊縮性ハイブリダイゼーション条件下で、次にハイブリダイゼーション条件を中または低緊縮性ハイブリダイゼーション条件に下げてハイブリダイゼーションを行うことができる。

20

【0258】

ある実施態様においてハイブリダイゼーション条件を変化させて、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸の数を変更することができる。例えばハイブリダイゼーション条件の緊縮性を上げて、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸の数を減少させることができる。あるいはハイブリダイゼーション条件の緊縮性を下げて、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸の数を増加させることができる。すなわち本明細書においてハイブリダイゼーション条件を修飾して、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸の所望の数を得ることができる。

30

【0259】

捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸の数は、オリゴヌクレオチドアレイに結合した核酸を測定するための当該分野で公知の任意の方法により測定することができ、蛍光または吸光度のような光学的測定（これは例えば、オリゴヌクレオチドチップのようなオリゴヌクレオチドアレイについて行われる）；散乱、放射活性、化学発光、熱量測定、または磁性標識の検出；1つまたはそれ以上のアレイ位置の質量スペクトル測定；または米国特許第6,045,996号明細書のような当該分野で公知の他の方法がある。

40

【0260】

1つまたはそれ以上の捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸の数の1回またはそれ以上の測定は、捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸の実際の数、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸の所望の数と比較するために使用することができる。1つまたはそれ以上の捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸の数の測定により、ハイブリダイゼーション条件を修飾して、捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸の数を増加または減少（所望であればいずれでも）させることができる。かかる方法は、1つまたはそれ以上の捕獲オリゴヌクレオチドプローブにハイブリダイズする標的核酸の所望の数が得られるまで繰り返し行うことができる。

50

【0261】

H. トリミング (Trimming)

ある実施態様において、2本鎖の以後の質量スペクトル分析を促進するためにおよび組成の複雑さを低下させるために、捕獲オリゴヌクレオチド：標的断片2本鎖の1本鎖オーバーハング部分をトリミングしてサイズを小さくすることができる。トリミングは、例えば標的核酸断片の平均サイズが比較的大きい時、または広範囲の異なるサイズの標的核酸断片がある時に行われる。トリミングは、質量スペクトル法により測定できるように標的核酸断片のサイズを低下させるために行うことができる。トリミングはまた、質量スペクトル法により測定される標的核酸断片の異なるサイズの範囲を低下させるために、および/または質量スペクトル法により測定される断片の質量を低下させるために行うことができる。

10

【0262】

トリミング法は、種々の公知の任意の方法により行うことができる。例えばトリミングは、捕獲断片のアレイを酵素または化学物質でさらに処理して、ハイブリダイズしていないヌクレオチドを除去するために行うことができる。酵素は例えば、当該分野で公知の任意のエキソヌクレアーゼ、または「1本鎖特異的RNaseまたはDNase」または「塩基特異的RNaseまたはDNase」、または配列特異的ヌクレアーゼである。別の実施態様においてエンドヌクレアーゼ（例えば1本鎖特異的エンドヌクレアーゼ）を使用して、ハイブリダイズしていないヌクレオチドをトリミングすることができる；そのようなトリミング反応においては、必ずしもすべてのハイブリダイズしていないヌクレオチドを除去する必要はない。1本鎖特異的エンドヌクレアーゼは、配列特異的でもまたは配列非特異的でもよい。例えば酵素は塩基特異的RNaseまたはDNaseでもよく、捕獲オリゴヌクレオチドより大きいハイブリダイズ断片は、3'末端または5'末端、またはこれらの両方を、A、C、G、またはT/Uの1つまたはそれ以上の存在の関数としてトリミングすることができる。

20

【0263】

I. 標的核酸断片に関する情報

標的核酸の核酸配列を再構築するための方法および本明細書に記載の他の方法（標的核酸の部分の同定を含む）は、本明細書の方法で標的核酸および標的核酸断片に関する種々の情報を利用して、標的核酸の配列の再構築または部分の同定を行うことができる。かかる情報には、質量測定、質量ピークの特性、標的核酸がハイブリダイズする捕獲オリゴヌクレオチドの配列、ハイブリダイゼーション条件、および使用される断片化法がある。

30

【0264】

1. 分子質量

本明細書に記載のように標的核酸の核酸配列を再構築するための工程、および本明細書に開示の他の方法（標的核酸の部分の同定を含む）は、捕獲核酸にハイブリダイズした標的核酸断片、または捕獲オリゴヌクレオチド：標的断片2本鎖の分子質量の測定を利用して、標的核酸断片の質量を測定することができる。

【0265】

a. 質量スペクトル分析

特定の分子の質量の測定に質量スペクトル分析を使用することができる。かかるフォーマットには、特に限定されないが、マトリックス支援レーザー脱離/イオン化、飛行時間 (MALDI-TOF)、電子噴霧イオン化 (ESI)、IR-MALDI（例えば、国際PCT出願公報第99/57318号および米国特許第5,118,937号明細書を参照）、直交-TOF (O-TOF)、横断-TOF (Axial-TOF; A-TOF)、イオンサイクロトロン共鳴 (ICR)、フーリエ変換、線形/レフレクトロン (RETOF)、およびこれらの組合せがある。また本明細書に記載の方法での使用に適した質量スペクトル法の例の総説については、AebersoldとMann、2003年3月13日、Nature 422:198-207（例えば図2）を参照されたい（これは参照することによりその全体が本明細書に組み込まれる）。MALDI法は典型的には、UV-MALDIまたはIR-MALDIを含む。核酸は、質量スペクトルに依存する検出法およびプロトコールにより分析することができる（例えば、米国特許第5,605,798号明細書、同6,043,031号明細書、同6,197,498号明細書、同6

40

50

,428,955号明細書、同6,268,131号明細書、および国際特許出願第WO96/29431号、国際PCT出願第WO98/20019号を参照)。これらの方法は自動化することができる(例えば米国特許公報第20020009394号、これは自動プロセスラインを記載する)。中解像度装置(特に限定されないが、曲線フィールドレフレクトロンまたは遅延抽出飛行時間MS装置を含む)はまた、配列決定または診断のための改良されたDNA検出を与えることができる。これらはいずれも、30量体鎖中の9 Da ($m(A-T)$)を検出することができる。

【0266】

質量スペクトル(例えばMALDI)を使用して分析が行われる時、ナノリットル量の試料がチップ上に添加される。このような量の使用は、定量的または半定量的質量スペクトル結果を与え得る。例えば得られる質量スペクトル中のピーク下の面積は、試料の成分の相対濃度に比例する。かかるチップを調製し使用する方法は当該分野で公知であり、例えば第6,024,925号明細書、米国特許公報第20010008615号、およびPCT出願第PCT/US97/20195号(WO98/20020号)に例示されている;そのようなチップを調製し使用する方法は、同時係属米国特許出願第08/786,988号、09/364,774号、および09/297,575号に記載されている。これらの分析を行うためのチップとキットは、セクエノム(SEQUENOM)から登録商標マースアレイ7(MassARRAY7)で市販されている。マースアレイ7(MassARRAY7)システムは、結果を迅速に与えるためのMALDI-TOF(マトリックス支援レーザー脱離/イオン化-飛行時間)質量スペクトル法に有用なスペクトロチップ7(SpectroCHIP7)アレイのような小型化アレイを含有する。これは、タグの無い遺伝変種に関するDNA断片のサイズの単一の塩基変化を正確に区別する。

10

20

【0267】

i. 測定された核酸分子の性状解析

ある実施態様において断片化工程で作製されたすべての核酸分子断片の質量が測定される。標的核酸分子断片または増幅産物断片の測定された質量はまた、参照核酸断片から得られる「参照」質量に対して「試料」測定質量と呼ぶことができる。

【0268】

別の実施態様において、その質量が質量スペクトル法を使用して測定される核酸分子断片の長さは、75ヌクレオチド以下の長さ、60ヌクレオチド以下の長さ、50ヌクレオチド以下の長さ、40ヌクレオチド以下の長さ、35ヌクレオチド以下の長さ、30ヌクレオチド以下の長さ、27ヌクレオチド以下の長さ、25ヌクレオチド以下の長さ、23ヌクレオチド以下の長さ、22ヌクレオチド以下の長さ、21ヌクレオチド以下の長さ、20ヌクレオチド以下の長さ、19ヌクレオチド以下の長さ、または18ヌクレオチド以下の長さである。

30

【0269】

別の実施態様においてその質量が質量スペクトル法を使用して測定される核酸分子断片の長さは、3ヌクレオチド以上の長さ、4ヌクレオチド以上の長さ、5ヌクレオチド以上の長さ、6ヌクレオチド以上の長さ、7ヌクレオチド以上の長さ、8ヌクレオチド以上の長さ、9ヌクレオチド以上の長さ、10ヌクレオチド以上の長さ、12ヌクレオチド以上の長さ、15ヌクレオチド以上の長さ、18ヌクレオチド以上の長さ、20ヌクレオチド以上の長さ、25ヌクレオチド以上の長さ、30ヌクレオチド以上の長さ、または35ヌクレオチド以上の長さである。

40

【0270】

ある実施態様においてその質量が測定される核酸分子断片はRNAである。別の実施態様においてその質量が測定される標的核酸分子断片はDNAである。さらに別の実施態様においてその質量が測定される標的核酸分子断片は、1つの修飾または異常ヌクレオチド(すなわち、DNAではデオキシ-C、T、G、またはA、またはRNAではC、U、G、またはA)を含有する。例えば転写反応の核酸分子産物は、リボヌクレオチドとデオキシリボヌクレオチドとの組合せを含有することができる。別々の例では核酸分子は、典型的に存在するヌクレオチドおよび質量が修飾されたヌクレオチドを含有することができ、または典型的に存在するヌクレオチドと天然に存在しないヌクレオチドを含有することができる。

【0271】

50

ii. 調整

質量スペクトル分析の前に、分解能を改良するために核酸分子を処理することができる。そのような方法は、分子の調整と呼ばれる。例えば揮発に必要なレーザーエネルギーを減少させるために、および/または断片化を最小にするために分子は「調整される」。核酸分子調整のための種々の方法が当該分野で公知である。調整の例は、核酸分子のホスホジエステル骨格の修飾（例えば陽イオン交換による）であり、これはヌクレオチド単位当たりに結合した陽イオン中の不均一性によるピークの広がりを排除するのに有用である。他の例では核酸分子にアルキル化剤（例えばヨウ化アルキル）、ヨードアセトアミド、ヨードエタノール、または2,3-エポキシ-1-プロパノールを接触させると、核酸分子のモノチオホスホジエステル結合をホスホトリエステル結合に変換することができる。同様にホスホジエステル結合は、例えば塩化トリアルキルシリルを使用して無変化誘導体に変換することができる。さらなる調整には、脱プリン化に対する感度を低下させるヌクレオチド（MS中の断片化）、例えばプリン類似体（例えばN7-またはN9-デアザプリンヌクレオチド）、RNA構成ブロックを取り込むこと、またはオリゴヌクレオチドトリエステルを使用すること、またはアルキル化されたホスホロチオエート官能基を取り込むこと、またはPNAのようなオリゴヌクレオチド模倣物を使用することがある。

10

【0272】

iii. 多重化 (Multiplexing)

いくつかの応用では、2つ以上の核酸分子断片の同時検出を行うことができる。他の応用では、例えば種々の固体支持体上のオリゴヌクレオチドまたはオリゴヌクレオチド模倣物アレイを使用して平行処理を行うことができる。「多重化」はいくつかの異なる方法により行われる。例えばいくつかの異なる核酸分子からの断片を、同時に質量測定法に付すことができる。典型的には多重質量測定において核酸分子断片は、多重核酸分子断片の同時検出が可能になるように区別できなければならない。核酸分子断片は、使用される質量測定法により断片の質量が区別できることを確実にすることにより、区別できるようになる。これは、配列自体（組成または長さ）により、または1つまたはそれ以上の核酸分子への質量修飾官能基の導入により達成される。

20

【0273】

b. 他の測定法

当該分野で公知の追加の質量測定法が、質量測定の方法で使用でき、電気泳動法（例えばゲル電気泳動および毛细管電気泳動法）、およびクロマトグラフィー法（サイズ排除クロマトグラフィーおよび逆相クロマトグラフィー法を含む）がある。

30

【0274】

2. 質量ピーク特性

本明細書に記載のように質量分析法を使用して、標的核酸分子断片の質量に関する情報を得ることができる。質量測定から得られる質量ピークの追加の情報には、ピークのシグナル対ノイズ比、ピーク面積（例えば、ピーク下の面積または半分の高さのピーク幅で示される）、ピーク高、ピーク幅、1つまたはそれ以上の追加の質量ピークに対するピーク面積、1つまたはそれ以上の追加の質量ピークに対するピーク高さ、および1つまたはそれ以上の追加の質量ピークに対するピーク幅がある。そのようなピーク特性は本配列決定法で使用することができ、例えば増幅断片の少なくとも1つの質量ピーク特性を1つまたはそれ以上の参照核酸の1つまたはそれ以上の質量ピーク特性と比較することにより、標的核酸分子のヌクレオチド配列を同定する方法で使用することができる。

40

【0275】

3. 捕獲オリゴヌクレオチドとハイブリダイゼーション条件

捕獲オリゴヌクレオチドとのハイブリダイゼーションを含む方法では、捕獲オリゴヌクレオチドは典型的には既知のヌクレオチド配列を有する。さらに標的核酸断片が捕獲オリゴヌクレオチドと接触される時使用されるハイブリダイゼーション条件の緊縮性も、典型的には公知である。捕獲オリゴヌクレオチドの配列とハイブリダイゼーション条件の知見は、捕獲オリゴヌクレオチドにハイブリダイズした標的核酸断片のヌクレオチド配列に関

50

する情報を与えるのに使用することができる。

【0276】

標的核酸分子のヌクレオチド配列を構築する方法において捕獲オリゴヌクレオチドプロープの配列を、特定の観察された質量で示される標的核酸配列候補の数を減少させるのに使用することができる。捕獲オリゴヌクレオチドの配列が既知の時、当業者は特定のハイブリダイゼーション条件下で捕獲オリゴヌクレオチドにハイブリダイズする標的核酸断片のヌクレオチド配列を予測することができる。さらに当業者は、特定のハイブリダイゼーション条件下で捕獲オリゴヌクレオチドにハイブリダイズしないと思われる標的核酸断片のヌクレオチド配列を予測することができる。

【0277】

いくつかのヌクレオチド配列の存在の可能性と他のヌクレオチド配列の欠如の可能性は、質量観察結果の解釈を助ける。特定の質量の観察は、その質量で示される標的核酸断片の組成（例えばDNA断片中のC、G、A、およびTの数）を決定するのに使用できるが、典型的にはさらなる情報が無い時、その質量で示される標的核酸断片のヌクレオチド配列を決定するのに使用することができない。すなわち特定の質量観察結果は、種々の異なる標的核酸断片のヌクレオチド配列のいずれかを示す。質量観察結果は、ハイブリダイゼーション情報（捕獲オリゴヌクレオチドとハイブリダイゼーション条件）で補足できるが、これは特定の質量観察結果で示される可能なヌクレオチド配列の数を限定または減少させる。可能なヌクレオチド配列の限定されたかまたは減少した数は、配列構築法でまたは本明細書に記載のように参照と比較するために使用することができる。

【0278】

ある例では4ヌクレオチド捕獲オリゴヌクレオチドはヌクレオチド配列5'ACTG3'を有するが、標的核酸断片は、捕獲オリゴヌクレオチドと完全に相補的な標的核酸断片のみが捕獲オリゴヌクレオチドにハイブリダイズするように、高緊縮性条件下で捕獲オリゴヌクレオチドに接触させることができる。この例でさらに、捕獲オリゴヌクレオチドにハイブリダイズした標的核酸断片の質量が測定され、断片の組成が決定され、組成A₃CTGを有する1つの質量が決定される。質量（従って組成）とハイブリダイゼーション情報が組合わせると、A₃CTG質量はヌクレオチド配列AAACTG、AACTGA、またはACTGAAを有する1つまたはそれ以上の断片を含有することが予測される。すなわち標的核酸分子はヌクレオチド配列AAACTG、AACTGA、またはACTGAAの1つまたはそれ以上を含有することができる。

【0279】

同じ捕獲オリゴヌクレオチドとハイブリダイゼーション条件を有する同様の例では、組成A₃CTGに対応する質量ピークが観察されない。ハイブリダイゼーション情報と組合せるとこの観察結果は、標的核酸分子がAAACTG、AACTGA、またはACTGAAのいずれかも含有しない可能性があることを示す。

【0280】

観察質量特性および参照質量特性を比較することを含む方法では、捕獲オリゴヌクレオチド配列とハイブリダイゼーション条件は、試料パターンと参照パターンとを一致させるための追加の情報源となり得る。例えばアレイ中の複数の捕獲オリゴヌクレオチドについて質量を測定することができる。参照配列は複数の捕獲オリゴヌクレオチドのそれぞれの質量特性の特定のパターンを有することが観察または計算され、これは質量対捕獲オリゴヌクレオチドの2次元パターンを与え得る。1つまたはそれ以上の参照パターンを試料のパターンと比較して、本明細書に記載の方法に従って標的核酸を同定するかまたはヌクレオチド配列を同定することができる。

【0281】

4. 断片化法

標的核酸分子を断片化するのに使用される方法は、ヌクレオチド配列構築または本明細書に記載の他の方法で使用できる情報を提供することができる。ある例では、断片化を行って既知の統計的サイズ範囲を有する標的核酸断片を与えることができる。別の例では、捕獲オリゴヌクレオチドへのハイブリダイゼーション後に断片を「トリミング」して、捕

10

20

30

40

50

獲オリゴヌクレオチドと同じ長さを有するかまたは典型的には捕獲オリゴヌクレオチドよりほんのわずかに大きいようにすることができる（例えば、塩基特異的断片化トリミングが行われる時）。断片化法はまた、断片中の1つまたはそれ以上のヌクレオチド位置のヌクレオチド配列を限定することができる；典型的にはこれは、配列特異的切断（例えば塩基特異的RNaseまたは制限エンドヌクレアーゼを使用して）が行われる時起きる。すなわち産生される断片が既知のサイズ（またはサイズ範囲）、既知のヌクレオチド配列情報、またはその両方を有する断片化法が行われる。

【0282】

使用される断片化法からわかる標的核酸断片についての情報以外に、本明細書に記載のヌクレオチド配列構築法は、断片化法により重複断片が産生される時得られる情報を利用することができる。重複断片の存在は、核酸配列の構築または核酸配列構築の正確さの上昇に使用できる情報の重複を与える。例えば第1と第2の標的核酸断片は、標的核酸中で互いに隣接したヌクレオチド部分から得られる；第3の標的核酸断片は、第1の標的核酸断片のヌクレオチド配列の部分と第2の標的核酸断片のヌクレオチド配列の部分とを含有することができる、隣接ヌクレオチド配列として第1と第2の標的核酸断片を同定するのに使用することができる、従って標的核酸のヌクレオチド配列を構築するのに有用である。

【0283】

J. ヌクレオチド配列構築

標的核酸断片に関する情報、例えば断片化法、質量測定、質量ピーク特性、および標的核酸断片がハイブリダイズした捕獲オリゴヌクレオチド（ハイブリダイゼーション条件）を使用して、標的核酸分子のヌクレオチド配列を構築することができる。例えば配列構築の方法は、成分の質量に従って試料の成分を分離し測定できる質量スペクトル法の能力を利用する。また配列構築の方法は、本明細書に記載のハイブリダイゼーション法を利用して、試料中の核酸断片の複雑さ（例えば、核酸断片の数および/または変動）を低下させ、場合により2つまたはそれ以上の核酸断片を有する試料を与える。また配列構築法は、断片化法により生成する核酸断片のサイズおよび/または配列を利用することができる、重複核酸断片の存在を利用することができる。これらの情報源を利用して、核酸分子の部分的または完全なヌクレオチド配列を決定することができる。ヌクレオチド配列の構築法は：広範囲の新規配列決定、広範囲の再配列決定、広範囲のSNP発見、広範囲の変異発見、より長い配列領域を使用する細菌型判定（例えば、完全な16S rRNA遺伝子に基づく方法を使用する細菌型判定）、多重配列決定（例えば、1つの実験で複数の短いアンプリコン）、広範囲メチル化分析（例えば、さらに少ないチップ位置を用いる特殊メチル化を使用して）、ヒト同定（例えば1つの長い領域または複数の短い領域を使用して）、生物同定（例えば1つの長い領域または複数の短い領域を使用して）、病原体および非病原体混合物の分析、および不均一核酸混合物の定量の方法で使用することができる。

【0284】

1. 標的核酸断片に関する情報の役割

ヌクレオチド配列を構築するための本明細書に記載の方法は、質量スペクトルの質量のヌクレオチド配列についての限界を予測または規定する能力に基づくことができる。例えば質量スペクトルの質量についての予測される配列または配列の限界は、(1)断片化法、(2)捕獲オリゴヌクレオチド、および(3)質量測定のような情報に基づくことができる。

【0285】

本明細書において断片化法は、種々の核酸断片、例えば特定の範囲内のヌクレオチド長さを有する断片（例えば、15~30ヌクレオチドの範囲の長さ）、特定の塩基で切断される断片（例えば、塩基特異的切断）、1つまたはそれ以上の特定のヌクレオチド配列で切断される断片（例えば、配列特異的エンドヌクレアーゼによる消化により生成される断片）、または捕獲オリゴヌクレオチドと同じ長さの断片（例えば「トリミング」した断片）の任意のものを作製するのに使用することができる。生じる断片は、使用された断片化法の関数である低下した複雑さを有する。例えば特定の範囲のヌクレオチド長さ（例えば、15~30ヌクレオチドの長さ）を有する断片のプールは、特定のヌクレオチド長さではない断

10

20

30

40

50

片のプール（任意の長さの断片）に対して低下した複雑さを有する。ヌクレオチド断片の低下した複雑さは、断片のヌクレオチド配列の限界を予測または規定するのに使用することができる。例えば塩基特異的切断ではすべての断片は、一端に単一の特定のヌクレオチド（塩基特異的に切断されたヌクレオチド）を有し、断片の残りは残りの3つのいずれかを有する。ヌクレオチド断片の低下した複雑さはまた、特定の捕獲オリゴヌクレオチドとハイブリダイズする異なるヌクレオチド断片の数を制限するのに使用でき、および/または質量スペクトル法により測定される異なるヌクレオチド断片の数を制限するのに使用することができる。例えばもしすべての断片が捕獲オリゴヌクレオチドと同じ長さを持つなら、捕獲オリゴヌクレオチドにハイブリダイズする断片の数と質量スペクトル法により測定される断片の数は、捕獲オリゴヌクレオチドに相補的なものに限定することができる。

10

【0286】

本明細書において捕獲オリゴヌクレオチドは種々の長さの任意のオリゴヌクレオチドを含有することができ、ユニバーサル塩基および/または半ユニバーサル塩基を含んでよい。各捕獲オリゴヌクレオチドにハイブリダイズした異なるヌクレオチド断片の数は、各捕獲オリゴヌクレオチドの長さや組成に従って制御される。例えば典型的なヌクレオチド（例えば、A、C、G、およびT）のみを含有する長い捕獲オリゴヌクレオチドは、典型的なヌクレオチドのみを含有する短い捕獲オリゴヌクレオチドと比較して、より短い異なるヌクレオチド断片がそこにハイブリダイズしている。別の例では典型的なヌクレオチドのみを含有する捕獲オリゴヌクレオチドは、1つまたはそれ以上のユニバーサル塩基または半ユニバーサル塩基を含有する同じ長さの捕獲オリゴヌクレオチドと比較して、より少ない異なるヌクレオチド断片がそこにハイブリダイズしている。特定の捕獲オリゴヌクレオチドにハイブリダイズした異なるヌクレオチド断片の数についての制限は、断片のヌクレオチド配列の限界を予測または規定するのに使用することができる。特定の捕獲オリゴヌクレオチドにハイブリダイズした異なるヌクレオチド断片の数についての制限はまた、質量スペクトル法により測定される異なるヌクレオチド断片の数を限定するのに使用することができる。

20

【0287】

質量測定は、1つまたはそれ以上のヌクレオチド断片の組成を決定するのに使用することができる。例えば質量測定を使用して、DNA断片中に存在するA、T、G、およびCの数を決定することができる。ヌクレオチド断片の組成は、断片のヌクレオチド配列についての制限を予測または規定するのに使用することができる。

30

【0288】

2. 配列構築の方法

例えば断片化、捕獲オリゴヌクレオチドハイブリダイゼーション、および質量測定により得られる情報は、標的核酸分子のヌクレオチド配列を構築するために本明細書に記載の種々の異なる方法で使用することができる。標的核酸分子のヌクレオチド配列を構築するために、本明細書の教示は、当業者が質量スペクトルによるヌクレオチド配列分析の公知の方法とともにハイブリダイゼーションによる配列決定法によるヌクレオチド配列分析の公知の方法を使用するよう指導している。例えば実験データは、公知の方法によりブリュイジン（Bruijn）グラフのサブグラフに変換することができる；例えばPevzner, J. *Biomol. Struct. Dyn.*, 7: 63-73 (1989)を参照。このグラフのオイラーの路（Eulerian paths）は、当該分野で公知のようにサイクルとパルジをあらかじめ破壊する必要がある場合に検索される；例えばPevznerら、*Proc. Natl. Acad. Sci. USA* 98: 9748-9753 (2001)参照。質量スペクトルは、当該分野で公知の方法により核酸断片のヌクレオチド組成をユニークに同定するのに使用することができる；例えばBoecker, *Lect. Notes Comp. Sci.* 2812: 476-487 (2003)参照。コンポマー（compomers）からヌクレオチド配列を決定するための分岐限界法（branch-and-bound）のような方法は当該分野で公知であり、Boecker, *Lect. Notes Comp. Sci.* 2812: 476-487 (2003)に例示されている。擬陰性ピークが引き起こす分岐限界法（branch-and-bound）の複雑さは当該分野で公知の方法により解決することができ、例えばS. Boecker、「擬陰性ピークの存在中のコンポマー（compomers）からの

40

50

配列決定」、Technical Report 2003-07、ビーレフェルト大学工学部 (Technische Fakultät der Universität Bielefeld)、情報技術部 (Abteilung Informationstechnik)、2003; また http://www.cebitec.uni-bielefeld.de/groups/ims/download/Preprint_2003-07_WeightedSC_SBoecker.pdf に例示されている。

【0289】

ある方法の例では標的核酸の仮定ヌクレオチド配列またはその断片が構築され、断片化 / ハイブリダイゼーション / 断片の質量が予測され、予測される質量を観察された質量と比較して、仮定のヌクレオチド配列が存在する否かを調べることができる。別の例では断片化 / ハイブリダイゼーション法の知見を使用して、観察されるすべての可能な質量を予測し、特定の質量に対応する配列を同定することができ、次にこの情報を観察された質量と比較して、標的核酸分子中に存在できる異なるヌクレオチド配列の数を限定することができる。この情報を使用してヌクレオチド配列を構築する方法の例を以下に示す。

10

【0290】

a. 仮定配列試験

断片化、ハイブリダイゼーション、および質量測定を使用する方法のある例では、標的核酸またはその断片の仮定ヌクレオチド配列を構築し、断片化 / ハイブリダイゼーション / 断片の質量を予測し、予測された質量を観察された質量と比較して、仮定ヌクレオチド配列が存在するか否かを試験することができる。この方法は、標的核酸分子の部分 (例えば1つのヌクレオチド断片) の仮定ヌクレオチド配列を構築し、この部分のヌクレオチド配列を測定し、この部分に1つまたはそれ以上の追加の仮定ヌクレオチドを付加し、追加の仮定ヌクレオチドが存在するか否かを試験することにより行われる。

20

【0291】

ある例では標的核酸分子は、一端または両端 (例えば、3'末端または5'末端、または両端) に既知のヌクレオチド配列を有することができる。例えば標的核酸分子が既知のヌクレオチド配列を有するプライマーで増幅されるような場合である。1つまたはそれ以上の仮定ヌクレオチドが既知の配列に付加され、観察された質量スペクトルを参照して仮定ヌクレオチドの存在を試験することができる。仮定ヌクレオチドと実際のヌクレオチドとのミスマッチは、実験的に観察された質量スペクトルには存在しない仮定質量の存在、および / または実験的に観察された質量スペクトルに存在する仮定質量の欠如を与える。従って実験的に観察された質量に最もよく一致する予測された断片質量を与える仮定ヌクレオチドは、標的核酸分子中の対応する位置に存在するヌクレオチドとして同定することができる。

30

【0292】

複数の質量スペクトルのそれぞれの中の無数の質量の有無を使用して、4つのヌクレオチドのいずれが存在するかを測定し、情報の重複を与え、こうして正確な配列決定の可能性を上昇させることができる。例えば特定のヌクレオチド位置のヌクレオチドの本体は、単一について質量スペクトルの予測された質量と観察された質量を比較することにより決定することができ、かかる決定以外に、1つまたはそれ以上の追加の質量スペクトルを参照することにより、決定を確認または否定するさらなる情報を得ることができる。複数の質量スペクトルを参照することにより、特定のヌクレオチドを同定するのに使用される観察の数を増加させ、従って正確なヌクレオチド同定の可能性を上昇させることができる。

40

【0293】

ヌクレオチド仮定試験に基づく配列構築法の1つの例は以下の通りである：

- (1) 1つまたはそれ以上の特定の位置に仮定ヌクレオチドを帰属する；
- (2) 断片化法に従ってヌクレオチドを含有する断片を予測する；
- (3) 各捕獲オリゴヌクレオチドについて、捕獲オリゴヌクレオチドへの予測される断片のハイブリダイゼーションの有無を予測する；
- (4) 各捕獲オリゴヌクレオチドについてハイブリダイズした断片の質量 / 組成を計算する；および
- (5) 予測された質量を観察された質量と比較する；

50

予測された質量と観察された質量との一致は、標的核酸分子ヌクレオチド配列中の実際のヌクレオチドとして仮想ヌクレオチドを同定することができる。

【0294】

所望であればこの方法は、各ヌクレオチド位置のすべての4つの典型的に存在するヌクレオチド（例えばDNAについてはA、G、C、およびT）について繰り返し、観察された質量と最もよく一致する予測質量を、標的核酸分子中のその位置に存在するヌクレオチドとして選択することができる。単一のまたは複数のヌクレオチド位置がこの方法により同時に試験でき、同時に試験されるヌクレオチド位置の数を、観察の数（例えば、存在する質量の数と欠如する質量の数）、質量スペクトル（例えば、質量スペクトル中に存在し得る異なる配列の数）、および標的核酸分子の長さに従って、本明細書に記載されている当該分野で公知のガイドラインに従って決定することができる。

10

【0295】

ヌクレオチド仮想試験に基づく配列構築の具体的な例では、（未知の）ヌクレオチド配列ACATGAGCTTACAAC（配列番号1）を有する標的オリゴヌクレオチドが断片化されて、5~7ヌクレオチドの長さの断片を与えることができる。次に核酸断片は、4つの半ユニバーサル塩基（例えば、ピリミジン（Y）のみにまたはプリン（R）のみに結合する塩基）のハイブリダイゼーション領域を有する捕獲オリゴヌクレオチドによりハイブリダイズされる。次にハイブリダイズした断片は質量スペクトルにより検出される。この例の目的において、標的オリゴヌクレオチドの最初の7つのヌクレオチドの配列はACATGAGであることがわかっている。8番目のヌクレオチドは、4つの可能な典型的に存在するヌクレオチドの任意のもの、例えば「T」が、仮に帰属される。配列ACATGAGTを含有するオリゴヌクレオチドに基づいて、各異なる捕獲オリゴヌクレオチド配列について測定された各質量スペクトルについて質量を予測することができる。例えば仮にこのヌクレオチド位置に「T」が帰属されると、配列RYYYを有する捕獲オリゴヌクレオチドプローブの質量スペクトルは、組成 $T_2G_2A_2$ 、 $T_2G_2A_2$ 、および $T_2G_2A_2C$ に対応する質量を含有すると予測される。ヌクレオチド配列ACATGAGCTTACAAC（配列番号1）について、この捕獲オリゴヌクレオチドにつき $T_2G_2A_2C$ のみが実験で観察される。同様に「G」の存在は3つの予測される質量を与え、これらのいずれもこの捕獲オリゴヌクレオチドのために実験的に存在しない。各位置が「A」として予測される時、これらの予測される質量の2つは実験的に存在し、8番目の位置が「C」として予測される時、すべての対応する実験的質量が観察される。すなわち「C」が最も近い一致を与える。この位置での「C」の存在をさらに確認するために、1つまたはそれ以上の他の捕獲オリゴヌクレオチドのスペクトルからの質量を比較することができる。例えば「A」が存在する場合、配列YYYYを有する捕獲オリゴヌクレオチドからの質量スペクトルは TG_2A_2 に対応する質量を含む。そのような質量は実験的には観察されない。しかし捕獲オリゴヌクレオチドYYRは組成 TG_2AC に対応する質量を有し、「C」がこの位置にあるかも知れない/あることを示す。

20

30

【0296】

この例では16個の異なる捕獲オリゴヌクレオチドを使用することができ、各捕獲オリゴヌクレオチドは、重複配列を含有するいくつかの核酸断片にハイブリダイズすることができる（例えば、断片が5~7個の長さのヌクレオチドである時、重複配列を有する9個の異なる断片は、同じ4ヌクレオチド長さの捕獲オリゴヌクレオチドにハイブリダイズすることができる）。すなわちこの例では、単一の質量スペクトルの最大9個の異なる質量が特定のヌクレオチド位置のヌクレオチド本体についての情報を与え、16個の異なる質量スペクトルを採取することができる。従ってこの標的オリゴヌクレオチドの各ヌクレオチド位置でヌクレオチドを同定するために大量の情報を使用することができる。

40

【0297】

b. 可能な配列を限定する

ある例では断片化法と捕獲オリゴヌクレオチドの組成を使用して、捕獲オリゴヌクレオチドにハイブリダイズしたヌクレオチド断片の質量スペクトルの特定の質量で示される可能なヌクレオチド配列の数を規定または限定することができ、また捕獲オリゴヌクレオチ

50

ドにハイブリダイズしたヌクレオチド断片の質量スペクトル中に存在できる可能な質量の数を規定または限定することができる。例えばすべての断片を8ヌクレオチドの長さで切断する断片化法は、存在可能な異なるヌクレオチド配列の数を 4^8 に限定し、ある質量スペクトルで可能な異なる質量の数はさらに限定される。ヌクレオチド断片の3'末端の特異的4ヌクレオチド配列にハイブリダイズする捕獲オリゴヌクレオチドは、存在可能な(特定の捕獲オリゴヌクレオチド位置で)ヌクレオチド配列の数を 4^4 にさらに限定し、ある質量スペクトルで可能な異なる質量の数はさらに限定される。

【0298】

これらの制限は実験的に測定される質量スペクトルに応用して、標的核酸分子の可能なヌクレオチド配列に限定を与えることができる。この制限は陽性(例えば、特定のヌクレオチド配列が標的核酸分子中に存在するかまたは存在するかも知れない)かまたは陰性(特定のヌクレオチド配列が標的核酸分子中に存在しない)である。例えば上記断片化例から得られる断片の質量と捕獲オリゴヌクレオチド条件は、24個またはそれ以下の可能なヌクレオチド配列に対応するように限定することができ、こうして標的核酸分子の8ヌクレオチドセグメントを24個またはそれ以下のヌクレオチド配列の1つに限定する。また特定の質量を有する断片が存在しないことは、そのような質量が存在することを示すヌクレオチド配列が標的核酸分子中に存在しないことを示す。さらに、多くの異なる捕獲オリゴヌクレオチドからの質量スペクトルを比較することができ、複数の質量スペクトルからの陰性および陽性制限は、特定の観察された質量に存在できる可能な配列の数を低下させることができる。

【0299】

観察の数(特定の質量の存在または特定の質量の欠如を含む観察)が充分大きく、質量スペクトル(例えば各質量スペクトル中に存在し得る異なる配列の数)が構築されるヌクレオチド配列(本明細書の教示に従って公知の方法により測定できる)に対して十分な単純化されている時、標的核酸分子のヌクレオチド配列を部分的にまたは全体に構築することができる。例えばある場合には、観察されたヌクレオチド断片組成(これは例えば、観察された質量から決定される)は、そこに帰属されたヌクレオチド配列を有し、充分な数のヌクレオチド断片(特に重複断片)が帰属されたヌクレオチド配列を有する時、標的核酸分子の完全なヌクレオチド配列を構築することができる。別の例では、ヌクレオチド断片組成はそこに帰属されたヌクレオチド配列を持たないが、断片の可能なヌクレオチド配列に対する制限を使用して、例えば断片間の重複を決定するための充分な制限を提供し、断片間の重複に基づく断片の配列を決定するための充分な制限を提供することにより、標的核酸分子の配列を決定することができる。さらに別の例では帰属されたヌクレオチド配列を有する断片は、帰属されていないヌクレオチド配列を有するがそのヌクレオチド配列に対する制限を有する断片とともに使用することができる。

【0300】

ヌクレオチド断片および/または標的核酸分子の可能な配列を限定することに基づく配列構築の1つの方法例は、以下の工程に従って行うことができる：

- (1) 核酸断片化の断片生成物の限界を規定または確立する；
- (2) 各特定の捕獲オリゴヌクレオチドにハイブリダイズできる核酸断片の限界を規定または確立する；
- (3) 捕獲オリゴヌクレオチドにハイブリダイズしたヌクレオチド断片の質量スペクトルで観察される可能な質量を予測する；
- (4) 特定の観察された質量中に存在できるであろう可能なヌクレオチド配列についての限界規則セットを作製する；そして
- (5) 観察された質量を規則セットと比較して、存在するであろう可能な配列を同定および/または存在しない配列を同定する。

【0301】

3. 方法の頑強性を測定するためのガイドライン

当業者は本明細書に記載の機能である要因に従って、その配列が構築できる標的核酸分

子の長さおよび/または配列決定が正しい確率の程度を測定することができる。さらに当業者は、その配列が構築できる標的核酸分子の長さおよび/または配列決定が正しい確率の所望の程度に従って、本明細書に記載の方法を設計することができる。例えば本明細書に記載の方法は、配列構築に利用可能な実験情報の量、および実験情報が標的核酸分子中に存在するかまたは欠如するユニークなヌクレオチド配列である程度を管理することができる。

【0302】

例えば本明細書に記載の方法は、ヌクレオチド配列構築で使用可能な異なる質量観察の数を管理することができる。質量観察は、例えば質量スペクトル中に存在する質量、または質量スペクトルからは欠如している質量（例えば可能なヌクレオチド断片の質量でのピークの欠如）でもよい。質量スペクトルの質量観察の数は、使用される断片化法および使用されるハイブリダイゼーション法（例えば、ハイブリダイゼーション条件と捕獲オリゴヌクレオチドの配列）により影響を受ける。例えば長さが10ヌクレオチドの断片のみを与える標的核酸分子の断片化は、長さが5~15ヌクレオチドの断片を与える標的核酸分子の断片化と比較して、質量観察の数を低下させることができる。質量観察の数はまた、異なるハイブリダイゼーション反応（例えば、異なるハイブリダイゼーション条件および/または異なる捕獲オリゴヌクレオチド配列）について採取される質量スペクトルの数により影響を受ける。

10

【0303】

本明細書に記載の方法はまた、同じ質量スペクトルで示される同じ質量を有するヌクレオチド配列の数および/または変化を管理することができる。例えば本明細書に記載の断片化法およびハイブリダイゼーション法は、同じヌクレオチド組成を有し同じ質量スペクトル中に存在する異なるヌクレオチド配列の数に影響を与えることができ、従って質量スペクトルの同じ質量ピークで示される。

20

【0304】

得られる実験情報（例えば観察の数および同じ観察で示される異なるヌクレオチド配列の数）を測定する方法は当該分野で公知である。得られる実験情報が決定されると、当業者は、核酸分子の長さおよび/またはヌクレオチド配列決定の確率の程度を推定することができる。あるいは所望の核酸分子の長さおよび/またはヌクレオチド配列決定の確率の程度に基づいて、当業者は、断片化法および/または所望の結果を得るためのハイブリダイゼーション反応の数と種類を設計することができる。

30

【0305】

K. 質量パターンによるヌクレオチド配列の同定

別の実施態様において、以下を含む標的核酸分子のヌクレオチド配列を同定するための方法が本明細書で提供される：

(a) 標的核酸分子の断片を捕獲オリゴヌクレオチドプローブにハイブリダイズさせる（2つまたはそれ以上の異なる標的核酸断片が捕獲オリゴヌクレオチドプローブにハイブリダイズする）；

(b) 捕獲核酸プローブにハイブリダイズした標的核酸断片の質量を測定する；

(c) 試料の質量を1つまたはそれ以上の参照質量パターンと比較する；

40

(d) 試料質量と一致する参照質量パターンを同定する；

こうして、試料質量と参照質量パターンとの一致により、標的核酸分子中のヌクレオチド配列が参照ヌクレオチド配列に対応するとして同定される。かかる方法において標的核酸中の配列を同定するために質量ピークの2つまたはそれ以上の特性が使用される。かかる同定法において質量ピークの2つまたはそれ以上の特性の集合は「パターン」と呼ばれる。

【0306】

本明細書に記載の方法において具体的なヌクレオチド配列は、そのヌクレオチド配列のユニークな特徴となる質量パターンを与えることができる。例えば具体的なヌクレオチド配列は、核酸がそのヌクレオチド配列を含有する時のみ生成される質量パターンを与える

50

ことができる。かかる状況においてヌクレオチド配列構築はヌクレオチド配列を同定するのに必要ではなく、ヌクレオチド配列は、観察されたパターンを参照パターン（参照パターンは特定のヌクレオチド配列に対応する）に一致させるだけで同定することができる。

【0307】

質量のパターンは、単一の質量スペクトル中に存在するか、または2つまたはそれ以上の異なるハイブリダイゼーション反応の質量スペクトル中に存在してもよい。参照パターンは、計算されたパターンまたは実験的に観察されたパターンでもよい。参照パターンが実験的に観察される場合、ヌクレオチド配列同定は再現し得るエラー（例えば、存在するかまたは欠如することが計算されるピークがそれぞれ再現性よく欠如しているかまたは存在する質量スペクトルのエラー）の影響を受けない。

10

【0308】

ある実施態様においてパターン一致による配列同定は、本明細書に記載のヌクレオチド配列構築法と組合せることができる。例えば標的核酸分子のセクションのヌクレオチド配列は、パターン一致により測定することができ、標的核酸および/または標的核酸分子の残りのヌクレオチド配列中のセクションの位置は、ヌクレオチド配列構築法により測定することができる。別の実施態様においてパターン一致による配列同定は、標的核酸分子の全ヌクレオチド配列を同定するのに使用することができる。

【0309】

ある例（例えば再配列決定とSNP分析）では、標的核酸分子についてすでに既知の配列（例えば公知のデータベース配列）が存在する可能性があるが、目的の具体的な標的核酸の配列は不明である。他の場合には、具体的なヌクレオチド配列について標的核酸断片質量パターンを知ることができる。いずれの場合も、1つまたはそれ以上の捕獲オリゴヌクレオチドにハイブリダイズする標的核酸断片の質量パターンを測定し、このパターンを計算または実験で測定された質量パターンと比較することにより、標的核酸中のヌクレオチド配列を同定することができる。

20

【0310】

同定される質量ピークは、捕獲オリゴヌクレオチドアレイ上の位置（すなわち、標的断片がハイブリダイズする具体的な捕獲オリゴヌクレオチド、および捕獲オリゴヌクレオチドの配列が既知である場合は、標的核酸断片がハイブリダイズする配列）、測定される質量、および質量測定のスIGNAL対ノイズ比を含む、3つまたはそれ以上の同定用の特徴を有することができる。質量ピークの1または2という少ない同定用特徴が、質量パターン一致によるヌクレオチド配列測定法で使用できることが本明細書で企図される。

30

【0311】

既知の配列の分析（例えば、再配列決定法または遺伝子型判定法）において、計算される質量パターンまたは実験的に測定される質量パターンは、標的核酸中のヌクレオチド配列を同定できる1つまたはそれ以上の質量ピーク特性を同定するのに使用することができる。例えばSNP分析は、問題のSNP位置の具体的なヌクレオチドの有無を示す1つまたはそれ以上のピークを測定することにより行われる。すなわち1つまたはそれ以上の指示質量ピークの有無を同定することは、標的核酸分子のすべてまたは任意のヌクレオチド配列を測定するためのヌクレオチド配列構築法を必要とすることなく、問題のSNP位置でヌクレオチドを同定するのに有用である。

40

【0312】

断片化およびハイブリダイゼーションパターンの計算は、質量ピーク特性パターンを予測するのに使用できる質量ピークを同定することができる。そのような方法は、質量ピーク特性（捕獲オリゴヌクレオチドアレイ上の特定の部位の断片の有無、断片の質量、および質量ピークのスIGNAL対ノイズ比）のいずれかまたはすべてを生成することができる。ある例では、問題の同じ位置の異なるヌクレオチド配列についてこれらの計算を繰り返すことにより、標的核酸上の1つまたはそれ以上のヌクレオチド位置の異なるヌクレオチド配列を示す1つまたはそれ以上の質量ピークのいくつかの異なる（および相互に排他性の）集合を作製することができる。

50

【0313】

試料の標的核酸断片の実験分析は、計算された配列指示性質量ピークの1つまたはそれ以上の集合と比較できる質量ピークを生成することができ、理論的に計算された配列指示性質量ピークのこの1つまたはそれ以上の集合は、実験的質量ピークと相関させることができる。次に試料標的核酸の全配列または配列の一部は、実験的質量ピークに最もよく相関する計算された配列指示性質量ピークの集合に対応する参照配列として同定することができるが、ただし場合により相関はユーザーが規定した閾値量より高い。実験的に得られた参照質量パターンと試料標的核酸分子の質量パターンとの同様の相関を作製することができる。

【0314】

試料ピークと参照ピークの相関は、当業者に公知の種々の方法で実施することができる。単純な例では、種々の参照質量ピークパターンの1つにのみ、具体的な捕獲オリゴヌクレオチドについて存在する1つの参照質量が存在してもよい。試料の標的核酸分子について同じ試料質量が検出される場合、標的核酸分子のヌクレオチド配列の少なくとも一部は、参照質量ピークに対応するヌクレオチド配列として同定することができる。試料ピークと参照ピークとの相関はまた、複数のピークを考慮する統計的方法（線形または非線形回帰のような回帰法を含む）を使用して、およびデータ相関について公知の他の方法を使用して行われる。

【0315】

ある実施態様においてユーザーは、十分な確率で標的核酸中のヌクレオチド配列を同定するために、参照核酸について必要な最小の相関を設定する閾値を規定することができる。閾値より高い相関が存在しない時、いずれの参照核酸も、十分な確率で標的核酸中のヌクレオチド配列を同定することができない。

【0316】

ある実施態様においてアレイ中の1つの位置で捕獲プローブにハイブリダイズした標的核酸断片の質量パターンは、標的核酸の1つまたはそれ以上の配列または部分を同定するのに有用である。例えば試料標的核酸が生物の染色体であり、標的核酸が例えば遺伝子発現、遺伝子型、種および変種の測定について特定の遺伝子または配列について試験される時、アレイ中の1つの位置で捕獲プローブにハイブリダイズした標的核酸断片の質量パターン（例えば、すべての標的核酸断片が、すべてが同じヌクレオチド配列を有する捕獲オリゴヌクレオチドプローブにハイブリダイズしている）は、具体的な発現される遺伝子、遺伝子型、種、または変種を示すか、または標的核酸は具体的な発現される遺伝子、遺伝子型、種、または変種に対応しないことを示す。

【0317】

ある実施態様において複数の捕獲プローブアレイ位置にハイブリダイズした標的核酸断片の質量パターンは、標的核酸中のヌクレオチド配列を同定するのに有用であり、ここで標的核酸断片は、アレイ中の500またはそれ以下の位置、アレイ中の250またはそれ以下の位置、アレイ中の100またはそれ以下の位置、アレイ中の75またはそれ以下の位置、アレイ中の50またはそれ以下の位置、アレイ中の25またはそれ以下の位置、アレイ中の20またはそれ以下の位置、アレイ中の15またはそれ以下の位置、アレイ中の10またはそれ以下の位置、アレイ中の8またはそれ以下の位置、アレイ中の6またはそれ以下の位置、アレイ中の5またはそれ以下の位置、アレイ中の4またはそれ以下の位置、アレイ中の3またはそれ以下の位置、またはアレイ中の2またはそれ以下の位置の捕獲プローブにハイブリダイズしている。

【0318】

ヌクレオチド配列構築を必要としない方法では、重複標的核酸断片の作製が使用されるが必須ではない。例えば再配列決定法またはSNPの配列の同定法では、非重複標的核酸断片を作製することができ、ヌクレオチド配列のすべてまたは一部が測定できる。SNP同定のような応用ではわずか1個の標的核酸断片を使用して、標的核酸のヌクレオチド配列をそのSNP位置に示すことができる。

10

20

30

40

50

【0319】

L. 標的核酸の部分の同定

別の実施態様において、以下を含む標的核酸の部分と同定するための方法が本明細書で提供される：

(a) 標的核酸の断片を捕獲オリゴヌクレオチドプローブにハイブリダイズさせる（ここで2つまたはそれ以上の標的核酸断片が捕獲オリゴヌクレオチドプローブにハイブリダイズする）；

(b) 捕獲核酸プローブにハイブリダイズした標的核酸の質量を測定する；および

(c) 質量を参照核酸分子の断片の質量と比較する、

ここで1つまたはそれ以上の試料質量と1つまたはそれ以上の参照質量との相関は、参照核酸分子に対応するとして標的核酸の部分と同定する。かかる同定法において質量ピークの2つまたはそれ以上の特性の集合は「パターン」と呼ばれる。

10

【0320】

ある実施態様において標的核酸の全ヌクレオチド配列を決定する必要無しで、1つまたはそれ以上の捕獲オリゴヌクレオチドにハイブリダイズする標的核酸断片の質量のパターンを使用して、標的核酸の1つまたはそれ以上の部分と同定することができる。別の実施態様において、標的核酸のいずれのヌクレオチド配列も測定することなく、標的核酸の1つまたはそれ以上の部分が同定される。

【0321】

ある場合には、標的核酸の配列が不明であっても、標的核酸分子またはその断片がどこにあるかを証明するための参照核酸質量パターンが公知である。例えば染色体はRFLPまたはAFLP地図と類似の標的核酸断片地図を有することができるが、染色体のすべてまたはサブセットのみが既知のヌクレオチド配列を有してもよい。ヌクレオチド配列が既知であってもなくても、1つまたはそれ以上の捕獲オリゴヌクレオチドにハイブリダイズする標的核酸断片の質量のパターンを測定し、このパターンを計算された（既知の配列の場合）かまたは実験的に測定された質量パターンと比較することにより、標的核酸分子の部分と同定することが可能である。

20

【0322】

問題の領域の配列が未知であっても、標的核酸断片の1つまたはそれ以上の質量ピークを1つまたはそれ以上の参照核酸の1つまたはそれ以上の質量ピークと比較することにより、標的核酸の1つまたはそれ以上の部分と同定することができる。この方法は、未知試料の1つまたはそれ以上のゲル電気泳動バンドが、1つまたはそれ以上の既知試料または参照試料の1つまたはそれ以上のゲル電気泳動バンドと比較される古典的DNAフィンガープリント法に似ている。例えば本発明において試料標的核酸から測定された質量ピークの3つの特性（すなわち、アレイ上の位置、質量、およびシグナル対ノイズ）の1つまたはそれ以上を、1つまたはそれ以上の参照核酸から測定された質量ピークの1つまたはそれ以上の特性と比較することができ、1つまたはそれ以上の参照の質量ピークは質量ピーク中の試料標的核酸と相関させることができる。次に試料標的核酸の部分は、試料標的核酸質量ピークと最もよく相関する1つまたはそれ以上の質量ピークを有する参照核酸の部分に対応するとして同定されるが、ただし場合により相関はユーザーが規定した閾値量より高い。すなわち標的核酸の1つまたはそれ以上の部分の同定は、問題の部分の配列も位置も未知であっても、具体的な参照核酸を同じ質量パターンを有するとして同定することにより行われる。

30

40

【0323】

ある実施態様においてアレイ中の1つの位置で捕獲プローブにハイブリダイズした標的核酸断片の質量パターンは、標的核酸の部分と同定するのに有用である。例えば試料標的核酸が生物の染色体であり、標的核酸が、例えば遺伝子発現、遺伝子型、種および変種について試験される時、アレイ中の1つの位置で捕獲プローブにハイブリダイズした標的核酸断片の質量パターンは、発現される遺伝子、遺伝子型、種、または変種を示すか、または標的核酸が特定の発現される遺伝子、遺伝子型、種、または変種に対応しないことを示

50

す。

【0324】

別の実施態様において複数の捕獲プローブにハイブリダイズした標的核酸断片の質量パターンは、標的核酸の部分と同定するのに有用であり、ここで標的核酸断片は、アレイ中の500またはそれ以下の位置、アレイ中の250またはそれ以下の位置、アレイ中の100またはそれ以下の位置、アレイ中の75またはそれ以下の位置、アレイ中の50またはそれ以下の位置、アレイ中の25またはそれ以下の位置、アレイ中の20またはそれ以下の位置、アレイ中の15またはそれ以下の位置、アレイ中の10またはそれ以下の位置、アレイ中の8またはそれ以下の位置、アレイ中の6またはそれ以下の位置、アレイ中の5またはそれ以下の位置、アレイ中の4またはそれ以下の位置、アレイ中の3またはそれ以下の位置、またはアレイ中の2またはそれ以下の位置の捕獲プローブにハイブリダイズしている。

10

【0325】

ヌクレオチド配列構築を必要としない方法では、重複標的核酸断片の作製が使用されるが必須ではない。例えば標的核酸断片のパターンを使用して生物、株、または種が同定でき、ここでパターンで使用される2つまたはそれ以上の質量ピーク特性のそれぞれは、標的核酸中の非隣接配列である標的核酸断片から得られ、このパターンは1つまたはそれ以上の参照核酸パターンと比較することができ、試料パターンを1つまたはそれ以上の参照パターンと関連させることにより生物、株、または種が同定される。

【0326】

M. 応用

本明細書に開示された方法を使用して、種々の目的のために標的核酸についての情報を得ることができる。以下に開示される応用は、本明細書に開示された方法の使用例である。当業者は、以下に記載した応用が標的核酸のヌクレオチド配列を構築する方法を使用して行われ、また標的核酸の部分と同定する方法（例えば標的核酸質量ピークパターンの分析を含む方法）を使用して行うこともできることを理解するであろう。

20

【0327】

1. 広範囲再配列決定

上記の広範囲新規配列決定法において本明細書に記載された配列決定法はまた、広範囲再配列決定にも使用できる。種々の生物からの劇的に増加している利用可能なゲノム配列情報の量は、配列情報を機能、表現型、または同一性と関連させるための大規模比較配列分析を可能にする技術に対するニーズを上昇させている。比較配列分析のためのそのような技術の応用は広く普及しており、例えばSNP発見や病原体の配列特異的同定がある。従って再配列決定および大量処理突然変異スクリーニング技術は、疾患の原因となっている変異、ならびに薬剤応答の差、および治療法に対する応答の差の原因となっている遺伝的変異性の同定にとって決定的に重要である。

30

【0328】

これらのニーズを満たすためにいくつかのアプローチが開発されている。高処理能DNA配列決定の技術は、電気泳動およびレーザー誘導蛍光検出を使用するDNAシーケンサーを含む。電気泳動に基づく配列決定法は、ヘテロ接合体を検出するために本質的な限界を有し、GC圧縮により犠牲になる。すなわち電気泳動を使用せずにデジタルデータを作り出すDNA配列決定プラットフォームは、これらの問題を克服する。マトリックス支援レーザー脱離/イオン化、飛行時間質量スペクトル法(MALDI-TOF MS)は、デジタルデータ出力でDNA断片を測定する。具体的な切断断片化分析の方法は、本明細書において、参照配列と比較して核酸配列の解明の高処理、高速、および高標準誤差覚醒を可能にした。このアプローチは、正確な配列補正のためならびに変異検出のために、MALDI-TOF MS、ならびに変異検出、例えばBRCA1とBRCA2中の創始者変異のスクリーニングを日常的に使用することを可能にする。

40

【0329】

再配列決定法は、標的核酸分析のために本明細書に開示された種々の方法を使用して行われる。例えば再配列決定は、核酸の大きいセグメントのヌクレオチド配列を決定するの

50

に使用できる配列構築法を使用して行われる。別の例では標的核酸の部分を同定する方法を使用することができる：例えば標的核酸が既知のもしくは参照核酸とはほんのわずかだけ（例えば5%またはそれ以下）異なる場合、質量ピークパターン解析のような方法を使用して、異なるヌクレオチド位置および変種ヌクレオチド位置のヌクレオチドの本体を同定することができる。すなわち、例えば公のデータに基づくヌクレオチド配列がエラーを含有する場合、本明細書に開示された種々の方法を使用して1つまたはそれ以上のエラーを修正することができる。

【0330】

2. 変異 / 配列変化の広範囲検出

本明細書の目的は、疾患およびそのマーカーの遺伝子のベースを同定するのに有用な改良された比較核酸配列決定法を提供することである。本明細書に記載の方法により同定される配列変化候補には、多型である配列変化を含む配列がある。多型には、天然に存在する体性配列変化および変異により生じる変化がある。多型性には特に限定されないが以下がある：SNPを含む配列微小変化（ここで、局所における1つまたはそれ以上のヌクレオチドは個々に異なる）、1ヌクレオチドから数百万の塩基までサイズが異なる挿入および欠失、および繰り返しの数が異なるマイクロサテライトまたはヌクレオチド繰り返し。ヌクレオチド繰り返しには、同じ配列が複数回繰り返される同種繰り返し（例えば、ジヌクレオチド、トリヌクレオチド、テトラヌクレオチド、またはより大きな繰り返し）、および配列モチーフが繰り返される異種繰り返しがある。ある遺伝子座については、ヌクレオチド繰り返しの数は個体により異なる。

【0331】

多型マーカーまたは部位は、分岐が起きる位置である。そのような部位はわずかに1つの塩基対（例えばSNP）でもよい。多型マーカーには、特に限定されないが、制限断片長多型（RFLP）、異なる数のタンDEM繰り返し配列（VNTR）、超可変領域、マイクロサテライト、ジヌクレオチド繰り返し、トリヌクレオチド繰り返し、テトラヌクレオチド繰り返し、および他の繰り返しパターン（例えばサテライト、ミニサテライト、単純な配列繰り返し、およびAluのような挿入要素）がある。多型はまた、ある遺伝子について異なるメンデル対立遺伝子として現れる。多型は、タンパク質の差、タンパク質修飾、RNA発現修飾、エピゲノミックな差、DNAとRNAメチル化、遺伝子発現とDNA複製を変化させる制御因子、およびゲノム核酸または小器官核酸の変化の任意の他の出現がある。

【0332】

さらに、多くの遺伝子が多型性領域を有する。各個体は多型性領域のいくつかの対立遺伝子変種のいずれか1つを有するため、各個体は遺伝子の多型性領域の対立遺伝子変種の種類により同定することができる。これは例えば法医学的応用で使用できる。他の場合には、個体が有する対立遺伝子変種の本体を知ることが重要なことがある。例えばある遺伝子、例えば主要組織適合複合体（MHC）遺伝子の対立遺伝子の差が、骨髄移植におけるような移植片拒絶または移植片対宿主反応に参与している。従って遺伝子または遺伝子病変の多型性領域の対立遺伝子変種の本体を決定するための、迅速で高感度かつ正確な方法を開発することが好ましい。本明細書で提供される方法またはキットは、対象の1つまたはそれ以上の遺伝子または染色体中の1つまたはそれ以上の多型性領域の1つまたはそれ以上の対立遺伝子変種の本体を測定することにより、対象を遺伝子型判定するのに使用できる。本明細書に記載の方法の1つまたはそれ以上を使用して対象を遺伝子型判定することは、法医学的目的または確認試験目的で使用することができ、多型性領域は、例えばミトコンドリア遺伝子に存在するかまたは短いタンDEM繰り返し配列でもよい。

【0333】

単一ヌクレオチド多型（SNP）は一般に2対立遺伝子系であり、すなわち各個体は特定のマーカーについて2つの対立遺伝子を有する。これは、10個以上の対立遺伝子を有するマイクロサテライトマーカーと比較すると、SNPマーカー1個当たりの情報含量は比較的少ないことを意味する。SNPはまた非常に集団特異的である；ある集団で多型性であるマーカーが別の集団ではあまり多型性ではないことがある。ほぼ1キロ塩基毎に存在するSNP（Wa

10

20

30

40

50

ngら、Science 280: 1077-1082 (1998)) は、非常に高密度の遺伝子地図を与える可能性を提供し、これは、目的の遺伝子または領域のハプロタイプ型判定系を開発するのに有用であり、かつSNPの本質のために実際これらは、試験される疾患表現型と関連する多型性である。SNPの低い変異率は、これらを複雑な遺伝形質を研究するのに優れたマーカーとなっている。

【0334】

ゲノミクス (genomics) の焦点の多くはSNPの同定に基づき、これは多くの理由から重要である。これらは間接試験 (ハプロタイプ型の関連) と直接試験 (機能性変種) とを可能にする。これらは最も豊富で安定な遺伝子マーカーである。通常の疾患は一般的な遺伝子改変により最も良く説明でき、ヒト集団の自然の変化は、疾患、治療、および環境相互作用を理解する助けとなる。

【0335】

3. 多重配列決定

また本明細書において、複数の標的核酸配列から核酸配列の高処理能的解明法が企図される。多重化は、2つ以上の標的核酸配列の同時解明を意味する。多重化反応を行う方法 (特に質量スペクトルと組合せて) は公知である (例えば米国特許第6,043,031号明細書、同5,547,835号明細書、国際PCT出願第WO97/37041号を参照)。

【0336】

多重化は、例えばある実験で標的核酸の複数の短いアンプリコンを使用して同じ標的核酸配列の複数の短い領域について行われる。多重化は、各個体の標的核酸配列について別々の質量スペクトル分析を行う必要があることと比較して、わずか1つの質量スペクトルで複数の標的核酸がスクリーニングできるという利点を与える。本明細書に記載の方法は、高速かつ正確に核酸配列を解明するための高処理能の高度に自動化されたプロセスに適用できる。

【0337】

多重化は、複数の異なる標的核酸を含有する試料中の、標的核酸の全配列を決定するために、少なくとも1つのヌクレオチド (標的核酸のすべてのヌクレオチドではない) の配列を決定するために、標的核酸の1つまたはそれ以上の部分を同定するために、または1つまたはそれ以上の特定の標的核酸の存在、または存在と相対的濃度を確定するために使用することができる。ある実施態様において標的核酸は、2つまたはそれ以上のmRNA核酸、または2つまたはそれ以上のmRNA核酸の鋳型を使用して生成される増幅核酸である。かかる方法において1つまたはそれ以上の細胞 (組織試料または血液または骨髄試料を含む) の遺伝子発現プロファイルを調べることができる。例えば2つまたはそれ以上のピークは2つまたはそれ以上のmRNAの発現を示し、2つまたはそれ以上の質量ピークの測定は標的核酸試料中に各mRNAが存在するかどうかおよび標的核酸試料中にmRNAが存在するレベルを明らかにすることができる。かかる方法を使用して、種々のmRNA (例えば癌遺伝子および細胞の新生物性または転移状態を示す他の遺伝子を含む)、細胞表面タンパク質をコードする遺伝子、遺伝性障害に関連する遺伝子、病原体または細胞の他の疾患状態による感染を示すmRNA、および活性化細胞障害剤細胞に関連する遺伝子の発現レベルを調べることができる。かかる方法はまた、種々の異なる試料 (例えば、異なるタイプの細胞、異なるタイプの組織、異なる生物、異なる株、異なる種、新しいタイプの細胞、新しいタイプの組織、新しい生物、新しい株、および新しい種を含む) 中の1つまたはそれ以上の遺伝子の発現レベルを測定するのに使用することができる。異なる試料中の発現レベルの測定は、例えば細胞の転移状態を測定するために、対象 (遺伝子疾患、感染性疾患、自己免疫疾患、または新生物疾患を有する患者を含む) を診断するために、細胞タイプ、組織タイプ、株タイプ、または生物のタイプを区別するために、2つまたはそれ以上の遺伝子の間の発現の関連を調べるために、遺伝子発現と細胞形態 (例えば細胞の有糸分裂または減数分裂状態) との相関を調べるために、使用することができる。

【0338】

任意の2つまたはそれ以上の生物学的分子源からの生物学的試料の混合物は、プールし

10

20

30

40

50

て本明細書での分析のための単一の混合物にすることができる。例えば本明細書に記載の方法を使用して、異なる供給源からの標的核酸またはアミノ酸の複数のコピーを配列決定し、従って生物学的試料中の核酸混合物の標的核酸またはアミノ酸中の配列変化を検出することができる。生物学的試料の混合物はまた、特に限定されないが、個体のプールからの核酸、1つまたはそれ以上の個体からの核酸の異なる領域、単一の組織もしくは細胞タイプから得られる均一な腫瘍試料、2つ以上の組織タイプまたは細胞タイプを含有する不均一な腫瘍試料、または原発性腫瘍から得られる細胞株を含む。また同じ遺伝子中の2つの変異が検出されるハプロタイプ型判定法のような方法が企図される。

【0339】

4. 広範囲メチル化パターン解析

本明細書に記載の方法を使用して、標的配列中のエピジェネティック変化である核酸配列変化（例えば標的配列中のメチル化パターンの変化）を解明することができる。細胞メチル化の解析は新しい研究分野である。シトシンへのメチル基の共有結合的付加は、主にCpGジヌクレオチド（マイクロサテライト）に存在する。プロモーター領域に存在しないCpGアイランドの機能は不明であるが、プロモーター領域中のCpGアイランドは、これらのメチル化状態が関連遺伝子の転写と発現を制御するため、特に重要である。プロモーター領域のメチル化は遺伝子発現のサイレンス化を引き起こす。このサイレンス化は永久的であり、有糸分裂と減数分裂の間継続する。遺伝子発現における重要な役割のためにDNAメチル化は、発生過程、インプリンティング（imprinting）とX染色体不活性化、ならびに腫瘍発生、老化、および寄生性DNAの抑制に影響を与える。メチル化は多くの広範な腫瘍（例えば、肺癌、乳癌、結腸癌）の発癌性および白血病に関連していると考えられている。メチル化とタンパク質機能障害（長いQ-T症状）または代謝性疾患（一過性新生児糖尿病、2型糖尿病）との間にも関係がある。

【0340】

ゲノムDNAの重亜硫酸塩処理は、DNA内のメチル化シトシン残基の位置を分析するのに使用することができる。核酸を重亜硫酸塩で処理するとシトシン残基は脱アミノ化されてウラシル残基になるが、メチル化シトシンは変化しない。すなわち、例えば本明細書に記載の方法で重亜硫酸塩で処理されていない標的核酸の配列を重亜硫酸塩で処理された核酸の配列と比較することにより、核酸中のメチル化の程度およびシトシンがメチル化された位置を推定することができる。処理および未処理標的核酸間のかかる比較は、種々の方法により行われる。例えば未処理標的核酸は、未処理標的核酸から生成された質量ピークが計算されたものであり、実験的に決定されたものではない既知の配列でもよい。さらに未処理標的核酸配列質量ピークは、重亜硫酸塩処理をせずに、断片化と質量ピーク分析とを行って実験的に決定することができる。別の方法では、同じ処理標的核酸の相補鎖はメチル化シトシンを測定するのに有用である。この方法は、シトシンをウラシルに変換するのに重亜硫酸塩が使用される時発生する塩基対ミスマッチに基づく。重亜硫酸塩による処理後、メチル化2本鎖標的核酸は1つまたはそれ以上のG-Uミスマッチを含有する。両方の相補鎖の配列を決定することにより、G-Uミスマッチの存在を使用して、ウラシル位置の非メチル化シトシンの存在を示すことができ、およびG-C一致塩基対の存在を使用してメチル化シトシンの存在を示すことができる。

【0341】

制限エンドヌクレアーゼ反応を介するメチル化分析は、メチル化特異的認識部位を有する制限酵素（例えばHpaIIとMSPI）を使用して可能になる。基本的原理は、認識配列中のメチル化シトシンによりいくつかの酵素が阻止されることである。いったんこの区別が達成されると、生じる断片の以後の分析は本明細書に記載の方法を使用して行うことができる。

【0342】

これらの方法は、共同重亜硫酸塩制限分析（combined bisulfite restriction analysis; COBRA）で一緒に使用することができる。重亜硫酸塩による処理は、増幅PCR産物中のBstUI認識部位中の欠失を引き起こし、これは未処理試料と比較して、分析で新しい検出可

10

20

30

40

50

能な断片を出現させる。本明細書に記載の断片化に基づく配列決定法はメチル化部位の特異的切断と組合せて使用して、標的核酸配列中のメチル化パターンについて迅速で信頼できる情報を与えることができる。

【0343】

5. 生物の同定

本明細書に記載の方法は、生物を同定するかまたはある生物を他の生物から区別するのに使用することができる。ある実施態様においてヒト試料を同定することができる（例えば、1つの長い領域または複数の短い領域）。多型性STR遺伝子座と遺伝子の他の多型性領域は、ヒトの識別、父親および母親試験、遺伝子マッピング、移民および遺伝論争、双子の卵性試験、ヒトの近親についての試験、ヒト培養細胞の品質管理、ヒト遺体の識別、および法医学における精液試料、血痕および他の材料の試験に特に有用なマーカーである配列変化である。かかる遺伝子座はまた、市販の動物家系および血統解析、および市販の植物の品種改良で有用なマーカーである。穀物植物および動物における経済的重要性は、多型性DNAマーカーを使用して系統解析により確定することができる。効率的かつ正確な断片化に基づく核酸配列決定法および標的核酸の部分と同定するための本明細書に記載の方法は、かかる遺伝子座の本体を決定するのに使用することができる。標的核酸（例えばゲノムDNA）は、1つの長い標的核酸領域および/または複数の短い標的核酸領域から得られる。

10

【0344】

別の実施態様において、非ヒト哺乳動物、鳥、植物、真菌、および細菌のような非ヒト生物を同定するための方法を使用することができる。

20

【0345】

6. 病原体同定と型判定

また、本明細書に記載の断片化およびハイブリダイゼーションに基づく方法を使用して、微生物株を同定するプロセスまたは方法が本明細書で企図される。微生物は、特に限定されないが細菌、真菌、原生動物、繊毛虫、およびウイルスを含む種々の生物から選択される。微生物は、特定の属、種、株、または血清型に限定されない。微生物は、1つまたはそれ以上の参照配列と比較して標的微生物配列中の核酸配列および/または配列変化を測定することにより同定される。参照配列は、例えば同じかまたは異なる属、種、株、または血清型からの他の微生物、または宿主原核生物もしくは真核生物から得られる。

30

【0346】

細菌病原体の同定と型判定は、感染性疾患の臨床的管理に決定的に重要である。微生物の正確な同定は、疾患状態を健康な状態から区別するのみでなく、抗生物質または他の抗微生物療法が治療に適しているかおよびどれが最も適しているかを決定するのに重要である。病原体の型判定の伝統的な方法は、細菌を同定するための増殖特性、色、細胞もしくはコロニー形態、抗生物質感受性、染色性、臭い、および特定の抗生物質との反応性を含む種々の表現型の特徴を使用してきた。これらの方法のすべては、疑わしい病原体の培養を必要とし、これは多くの重大な欠点（材料と人件費の高いコスト、作業者が曝露する危険、誤操作による擬陽性、生きた細胞の数が少ないためのもしくは多くの病原体の偏好性培養要求性のための擬陰性を含む）がある。さらに培養法は、診断を下すのに比較的長い時間が必要であり、かかる感染症は生命にかかわるため、抗微生物療法は、結果が得られる前に開始されることが多い。

40

【0347】

多くの場合に病原体は、正常な微生物叢を形成する生物に非常によく似ており、上記の表現型法では無毒な株から区別することが不可能なことがある。こういう場合に病原性株の存在の測定には、本明細書に記載の断片化およびハイブリダイゼーションに基づく方法により与えられる高い分解能を必要とする。例えば標的核酸配列のPCR増幅後に、断片化やマトリックス支援レーザー脱離/イオン化飛行時間質量スペクトル法を使用するハイブリダイゼーションに基づく配列決定を行い、次に本明細書に記載のような配列変化をスクリーニングすることは、1つだけのヌクレオチドが異なる配列の信頼できる区別を可能に

50

し、得られる配列情報の識別力をMALDI-TOF MSの速度と組合せている。同様に1つまたはそれ以上の質量ピークまたは質量ピークパターンを比較することにより標的核酸の部分を同定する方法は、かかる配列変化を検出するのに使用することができる。

【0348】

例えば、本明細書に記載の断片化とハイブリダイゼーションに基づく方法（比較フォーマットにおける断片化に基づく配列決定法を含む）を使用して、より信頼できる長い配列領域（例えば完全長16S rRNA遺伝子）を使用する細菌型判定を行うことができる。例示すると、1つまたはそれ以上の既知タイプの細菌の配列を得て、未知タイプの細菌の配列と比較することができる。

【0349】

10

7. 分子育種と有向進化

ある実施態様において、標的核酸が修飾された核酸、ウイルス、または生物である時、標的核酸の配列または部分を測定するのに本明細書に開示された方法を使用することができる。かかる方法は、生物学的分子の性質または生物もしくはウイルスの表現型を、生物学的分子、生物またはウイルスの遺伝子型と関連させるのに使用することができる。例えば本明細書に開示された方法を使用して、標的核酸の具体的な性質に関連するヌクレオチド配列、質量ピーク、またはピークパターン、標的核酸によりコードされるタンパク質、または標的核酸を含有するウイルスもしくは生物を同定することができる。

【0350】

例えば本明細書の方法を使用して、標的核酸配列、質量ピーク、または質量ピークパターンに関連した特定のタンパク質の性質を同定することができる。この例では、DNAシャuffling（米国特許第6,117,679号明細書および同6,537,746号明細書）、エラーが出やすいPCR（Caldwell, R.C.とJoyce, G.F. (1992) PCR Methods and Applications 2: 28-33）、カセット突然変異誘発（Goldman, E.R.とYouvan D.C. (1992) Bio/Technology 10: 1557-1561；Delagraveら、Protein Engineering 6: 327-331 (1993)）、およびランダムコドン突然変異誘発法（米国特許第5,264,563号明細書および同5,723,323号明細書）を含む、遺伝子修飾のための当該分野で公知の種々の方法を使用して、タンパク質をコードする1つまたはそれ以上の遺伝子を修飾することにより、1つまたはそれ以上のタンパク質を再設計することができる。1つまたはそれ以上の特定の性質を有する再設計されたタンパク質をコードする遺伝子の配列または部分は、本明細書に開示された方法を使用して調べることができ、1つまたはそれ以上の質量ピークが、再設計されたタンパク質の1つまたはそれ以上の具体的な性質に関連するとして同定することができる。タンパク質の性質の例には、結合能力、触媒能力、熱安定性、プロテアーゼに対する感受性、発現レベル、溶解度、膜挿入または会合、翻訳後修飾、光学的性質、電子移動性、小器官ターゲティング、分泌される能力、肝臓での分解に対する感受性、免疫原性、生物学的バリアを通過して輸送される能力（消化管から血流への吸収、および脳血液関門の通過を含む）がある。

20

30

【0351】

再設計されたタンパク質の1つまたはそれ以上の特定の性質に関連するとして1つまたはそれ以上の質量ピークを同定する方法には、1つまたはそれ以上の特定の性質を有する1つまたはそれ以上の再設計されたタンパク質をコードする遺伝子の質量ピークのパターンの分析、およびこれらの特定の性質に関連するヌクレオチド配列または1つまたはそれ以上の質量ピークもしくは質量ピーク特性の同定がある。特定の性質に関連する配列または質量ピークの決定は、特定の性質を有するタンパク質をコードする2つまたはそれ以上の遺伝子に共通の配列または質量ピーク、典型的には特定の性質を有するタンパク質をコードする遺伝子の少なくとも50%、少なくとも70%、少なくとも85%、少なくとも90%、少なくとも95%に共通の配列または質量ピークを測定することにより行われる。特定の性質に関連する配列または質量ピークの決定はまた、そのようなタンパク質の1つだけが特定の性質を有する場合でも、そのタンパク質をコードする遺伝子にユニークな配列または質量ピークを測定することにより行われる。

40

【0352】

50

上記方法に従って他の実施態様は、1つまたはそれ以上の特定の性質を有するタンパク質をコードする1つまたはそれ以上の遺伝子を同定する方法を含み、ここでこの方法は、遺伝子を断片化し、遺伝子断片を1つまたはそれ以上の捕獲オリゴヌクレオチドプローブにハイブリダイズさせ（ここで2つまたはそれ以上の遺伝子断片は、同じヌクレオチド配列を有する捕獲オリゴヌクレオチドプローブにハイブリダイズする異なるヌクレオチド配列を有する）、2つまたはそれ以上の遺伝子断片の質量を測定することを含む。ある実施態様において質量ピークを測定した後、1つまたはそれ以上の測定した質量ピークを1つまたはそれ以上の参照質量ピーク（ここで1つまたはそれ以上の参照質量ピークは、再設計されたタンパク質の1つまたはそれ以上の特定の性質に関連する）と比較することができる。参照質量ピークは、例えば上記方法を使用して実験的に決定されるか、または理論的に決定することができる。別の実施態様において標的核酸のヌクレオチド配列が構築され、1つまたはそれ以上の特定のタンパク質の性質に関連する配列を含有する標的核酸は、かかる性質を有するタンパク質をコードする遺伝子として同定することができる。

10

20

30

40

50

【0353】

さらに本実施態様において再設計されたタンパク質の1つまたはそれ以上の特定の性質に関連する1つまたはそれ以上の質量ピークは、本明細書に開示された方法を使用して分析して、再設計されたタンパク質をコードする標的核酸遺伝子に関するヌクレオチド配列情報を与えることができる。例えば標的核酸配列情報は、1つまたはそれ以上の質量ピーク特性を1つまたはそれ以上の参照質量ピーク特性（1つまたはそれ以上の参照質量ピーク特性は、標的核酸上の1つまたはそれ以上のヌクレオチド位置の特定のヌクレオチド配列に対応する）と比較することにより得ることができる。別の例では1つまたはそれ以上の標的核酸断片のヌクレオチド配列は、測定された質量ピーク特性に従って、または本明細書に記載の配列構築法を使用して決定することができる。さらに別の例では全標的核酸配列またはその部分は、本明細書に記載の配列構築法を使用して決定することができる。

【0354】

別の例ではウイルスゲノムシャフリング（米国特許第6,596,539号明細書）およびウイルス変異と選択法を含む種々の方法を使用してウイルスゲノムを修飾することにより再設計することができる。1つまたはそれ以上の特定の性質を有する1つまたはそれ以上のウイルスを与える修飾されたウイルスゲノムは、本明細書に開示された方法を使用して調べることができ、1つまたはそれ以上の質量ピークが、修飾されたウイルスの1つまたはそれ以上の特定の性質に関連するとして同定される。ウイルスの性質の例には、ウイルスの感染性、複製、宿主範囲、親和性、遺伝子機能、転写制御配列機能、非許容細胞中で複製する能力、宿主範囲、および/または細胞親和性、ウイルス力価（例えば病毒性）、病原性すなわち疾患を生成する能力、感染性、パッケージング能、ウイルス粒子の物理的/化学的安定性、細胞内安定性、1つまたはそれ以上のウイルス遺伝子の発現、染色体組み込み、組織特異性および特定の生物に選択的に感染する能力、宿主（例えばヒト）中のウイルスまたはウイルスタンパク質の免疫原性、生物学的アジュバントとしての機能（例えば、ウイルスにコードされたヒトサイトカインを同時発現する）、および治療薬としての機能（例えば、インターフェロン産生のような一般的抗ウイルス宿主応答を誘導する能力）がある。

【0355】

再設計されたウイルスの1つまたはそれ以上の特定の性質に関連するとして1つまたはそれ以上の質量ピークを同定する方法には、1つまたはそれ以上の特定の性質を有する1つまたはそれ以上の再設計されたウイルスのウイルス配列の質量ピークパターンの分析、およびこれらの特定の性質に関連するヌクレオチド配列または1つもしくはそれ以上の質量ピークまたは質量ピーク特性を同定することを含む。特定の性質に関連する配列または質量ピークを決定することは、特定の性質を有する2つまたはそれ以上のウイルス配列に共通の配列または質量ピークを決定することにより行われ、典型的には配列または質量ピークは、特定の性質を有するウイルス配列の少なくとも50%、少なくとも70%、少なくとも85%、少なくとも90%、または少なくとも95%に共通である。特定の性質に関連して配列

または質量ピークを決定することはまた、そのようなウイルスの1つだけが特定の性質を有する場合でも、そのウイルス配列にユニークな配列または質量ピークを測定することにより行われる。

【0356】

上記方法において別の実施態様は、1つまたはそれ以上の特定の性質を有する1つまたはそれ以上のウイルス配列を同定する方法を含み、ここでこの方法は、ウイルス核酸を断片化し、ウイルス核酸断片を1つまたはそれ以上の捕獲オリゴヌクレオチドプローブにハイブリダイズさせ（ここで2つまたはそれ以上のウイルス核酸断片は、同じヌクレオチド配列を有する捕獲オリゴヌクレオチドプローブにハイブリダイズする異なるヌクレオチド配列を有する）、2つまたはそれ以上のウイルス核酸断片の質量を測定することを含む。ある実施態様において質量ピークを測定した後、1つまたはそれ以上の測定した質量ピークを1つまたはそれ以上の参照質量ピーク（ここで1つまたはそれ以上の参照質量ピークは再設計されたウイルスの1つまたはそれ以上の特定の性質に関連する）と比較することができる。参照質量ピークは、例えば上記方法を使用して実験的に決定されるか、または理論的に決定することができる。別の実施態様においてウイルス核酸のヌクレオチド配列が構築され、1つまたはそれ以上の特定のタンパク質の性質に関連する配列を含有するウイルス核酸は、かかる性質を有するタンパク質をコードするウイルス配列として同定することができる。

10

【0357】

さらに本実施態様において再設計されたウイルスの1つまたはそれ以上の特定の性質に関連する1つまたはそれ以上の質量ピークは、本明細書に開示された方法を使用して分析して、再設計されたウイルスのウイルス核酸に関するヌクレオチド配列情報を与えることができる。例えばウイルス核酸配列情報は、1つまたはそれ以上の質量ピーク特性を1つまたはそれ以上の参照質量ピーク特性（1つまたはそれ以上の参照質量ピーク特性は、ウイルス核酸上の1つまたはそれ以上のヌクレオチド位置の特定のヌクレオチド配列に対応する）と比較することにより得ることができる。別の例では1つまたはそれ以上のウイルス核酸断片のヌクレオチド配列は、測定された質量ピーク特性に従って、または本明細書に記載の配列構築法を使用して決定することができる。さらに別の例では全ウイルス核酸配列またはその部分は、本明細書に記載の配列構築法を使用して決定することができる。

20

【0358】

さらに本明細書において、遺伝子修飾された生物のような生物の1つまたはそれ以上の特定の性質に関連するとして1つまたはそれ以上の質量ピークを同定する方法が企図される。生物の例には、農業植物（トウモロコシ、コメ、小麦、ライ麦、オート麦、大麦、エンドウ豆、インゲン豆、レンズマメ、ピーナツ、ヤムビーン（yam bean）、ササゲ、ハッシュイマメ、大豆、クローバー、アルファルファ、ルピナス、ソラマメ、蓮、シナガワハギ、藤、スイートピー、モロコシ、キビ、ヒマワリ、およびカノーラを含む）のような植物；七面鳥やニワトリを含む鳥；魚；昆虫；線虫；非ヒト哺乳動物（ブタ、ウシ、ウマのような家畜および他の家畜を含む）がある。種々の生物のゲノムを修飾する方法は当該分野で公知であり、DNAシャフリング（米国特許第6,379,964号明細書および同6,500,617号明細書）があり、また有性生殖による伝統的な品種改良がある。生物の性質は生物により変化するが、生存活性、疾患に対する耐性、増殖速度、繁殖能力、栄養要求性、水分要求性、温度感受性、および環境ストレスに対する抵抗性がある。遺伝子修飾された生物のような生物の1つまたはそれ以上の特定の性質に関連するとして1つまたはそれ以上の質量ピークを同定する方法は、ウイルスに関する本明細書で上記した方法を使用して行われる。

30

40

【0359】

8. マーカーとしての標的核酸断片

別の例では標的核酸断片は、大きな標的核酸の配列または部分のマーカーまたは指標として使用することができる。かかる実施態様は、標的核酸の全配列の測定を必要とせず、標的核酸の部分の配列の測定または単に標的核酸断片の質量ピークパターンの測定を含む。これらの実施態様はまた、標的核酸断片が重複することを必要とせず、すなわちこれら

50

の実施態様について標的核酸断片は重複しても重複しなくてもよい。かかる方法には、例えばフィンガープリンティング法およびフィンガープリンティング関連法、および標的核酸の配列または部分の指標として非重複DNA断片の使用を含む他の方法がある。増幅工程を使用するフィンガープリンティング法、例えば増幅リボゾームDNA制限解析 (ARDRA)、ランダム増幅多型DNA解析 (RAPD)、および増幅断片長多型 (AFLP) を、本明細書に開示された方法で使用することができる。

【0360】

ある実施態様において標的核酸の断片が生成され、捕獲核酸のアレイにハイブリダイズされ、断片の質量が測定されて、1、2、3、またはそれ以上の特性 (例えば、標的核酸がハイブリダイズする捕獲オリゴヌクレオチドプローブの位置、質量、および質量ピークのシグナル対ノイズ比) により特徴付けられる質量ピークのパターンを作製することができる。質量ピークのかかるパターンは、標的核酸の配列または部分の指標として使用することができる。

10

【0361】

ある実施態様において具体的に設計されたプライマーと増幅法は、標的核酸断片のサブセットのみが増幅されるように増幅を制御することができ、この断片のサブセットは次に捕獲オリゴヌクレオチドプローブのアレイにハイブリダイズされ、質量を分析することができる。この実施態様は標的核酸として以下を使用することができる：種または株の1つまたはそれ以上の異なる生物からの、遺伝子、染色体断片、酵母人工染色体 (YAC)、細菌人工染色体 (BAC)、全染色体、全ゲノムまたは任意の他の適当な核酸分子；または複数の遺伝子、染色体断片、YAC、BAC、全染色体および全ゲノム。核酸断片のサブセットを増幅する方法は当該分野で公知であり、例えば増幅断片長多型 (AFLP) 法 (例えば米国特許第6,045,994号明細書参照) がある。

20

【0362】

この実施態様において標的核酸の断片を作製するのに1つまたはそれ以上の制限酵素が使用される。典型的には異なるヌクレオチド配列で切断する2つの制限酵素が使用される。例えばレアカッター (rare cutter) (6ヌクレオチドのような長いヌクレオチド配列を認識し、従って核酸上のより少ない部位で切断する制限酵素) およびコモンカッター (common cutter) (4ヌクレオチドのような短いヌクレオチド配列を認識し、核酸上のより多い部位で切断する制限酵素) が使用される。他の例では、2つのレアカッターまたは2つのコモンカッターが使用される。制限酵素の数および酵素の特異性は、標的核酸の長さで標的核酸断片の所望の数と長さにより選択することができる。

30

【0363】

制限断片の末端のヌクレオチド配列が既知であるかどうかにかかわらず、制限断片のPCR増幅を行うことができる。これは、まず既知配列の合成オリゴヌクレオチド (アダプター) を制限断片の両端に連結し、こうしてPCR増幅で使用されるプライマーに相補的となり得る2つの一般的タグの付いた各制限断片が得られる。

【0364】

典型的には請求項は平滑末端 (両方の鎖の末端ヌクレオチドが塩基対合している) を産生するか、または「粘着」末端 (2つの鎖の1つが飛び出て短い1本鎖領域を与える) を産生する。平滑末端を有する制限断片の場合、平滑末端の1本鎖にアダプターが連結される。粘着末端を有する制限断片の場合、アダプターは制限断片の1本鎖領域に相補的な領域を有する。かかるアダプターはまず、アダプター末端が制限断片の1つの鎖の末端に隣接するように、制限断片の1本鎖領域の相補的部分にハイブリダイズされ、次にアダプターは隣接する制限断片末端に連結される。

40

【0365】

従って各タイプの制限切断について、アダプターの一端が特定の対応する制限断片に連結できるように、異なるアダプターを設計することができる。典型的にはアダプターは約10~30ヌクレオチドの長さであり、典型的には12~22ヌクレオチドの長さである。リガーゼ酵素を使用して、アダプターは制限断片の混合物に連結される。制限断片に対して大幅

50

にモル過剰のアダプターを使用する時、ほとんどすべての制限断片は両端でアダプターに連結される。この方法で調製された制限断片は「タグ付き制限断片」と呼ばれる。

【0366】

各タグ付き制限断片は以下の一般的構造を有する：タグ付き制限断片の各末端で定常DNA配列によりフランクされる可変DNA配列。定常DNA配列は、制限エンドヌクレアーゼの認識配列の一部またはすべてを含有し、またタグ付き制限断片の各末端に結合したアダプターの配列も含有する。制限断片の可変配列は定常DNA配列間に位置し、従って制限エンドヌクレアーゼ認識配列を含有しない制限断片の部分を含む。可変配列は既知でも未知でもよく、典型的には制限断片間で変化する。従って定常DNA配列をフランクするヌクレオチド配列は、異なる配列の大きな混合物でもよい。

10

【0367】

ある実施態様においてアダプターはPCRプライマーの正確な相補体でもよい。例えば制限断片は、その両方の末端で同じアダプターを運搬でき、単一のPCRプライマーが、制限断片配列のどこにもハイブリダイズすることなくアダプターにハイブリダイズでき、制限断片を増幅するのに使用することができる。別の例ではDNAを切断するために2つの異なる制限酵素を使用して、2つの異なるアダプターが制限断片の末端に連結される。この場合1つまたは2つの異なるPCRプライマーを使用して、かかる制限断片を増幅することができる。この実施態様のPCRプライマーは、制限断片の可変配列に無関係に、すべてのタグ付き制限断片を増幅するのに使用できる。

20

【0368】

上記工程でタグ付き制限断片が増幅されるか否かに無関係に、タグ付き制限断片は次に、第1のヌクレオチド配列部分と第2の配列部分とを含有する可変配列特異的PCRプライマーを使用して増幅される。第1の配列部分は、タグ付き制限断片の定常DNA配列と完全に塩基対合するように設計される。第2の配列部分は、任意の選択された配列またはランダム配列を含有し、1~約10ヌクレオチドの範囲の長さである。第2の配列部分はタグ付き制限断片のサブセットにのみハイブリダイズして、タグ付き制限断片のハイブリダイズしたサブセットのみが増幅される。ある実施態様においてタグ付き制限断片のより大きなサブセットを増幅するために、その第2の配列部分に異なる配列を有するいくつかの異なる配列特異的PCRプライマーが使用される。

30

【0369】

第2の配列部分を配列特異的プライマーの3'末端に付加すると、PCR工程でどのタグ付き制限断片が増幅されるかが決定される；配列特異的プライマーは、配列特異的PCRプライマーの第2の部分がタグ付き制限断片と塩基対合できるタグ付き制限断片上のDNA合成のみを開始させる。

【0370】

タグ付き制限断片のサブセットの配列特異的増幅後に、制限断片（これはまた標的核酸断片とも呼ばれる）は所望であれば、本明細書に開示された方法に従ってさらに断片化することができる。例えば標的核酸断片（制限断片）は、さらなる配列特異的切断、塩基特異的切断、または非特異的切断に付される。標的核酸断片は次に、捕獲オリゴヌクレオチドプローブのアレイにハイブリダイズされる。ハイブリダイゼーション後、標的核酸断片は所望であれば、本明細書に開示された方法に従ってさらに断片化することができる。例えば標的核酸断片は塩基特異的切断に付される。1つまたはそれ以上の捕獲オリゴヌクレオチドプローブにハイブリダイズした標的核酸断片の複雑さの所望のレベルを達成するために、または例えば質量スペクトル法を使用する質量測定 of 所望の正確性のために標的核酸断片の所望の長さを達成するために、例えばハイブリダイゼーション前またはハイブリダイゼーション後の切断が行われる。

40

【0371】

9. 感染を示すウイルスまたは細菌核酸配列の存在の検出

本明細書に記載の方法を使用して、1つまたはそれ以上の参照配列と比較してウイルスまたは細菌核酸配列中に存在する配列変化を同定することにより、感染を示すウイルスま

50

たは細菌の核酸配列の存在を測定することができる。参照配列には、特に限定されないが、関連する非感染生物から得られる配列または宿主生物の配列がある。

【0372】

ウイルス、細菌、真菌、および他の感染性生物は、宿主細胞中に含まれる配列とは異なる多型を含む明確な核酸配列を含有する。標的DNA配列は、侵入微生物（例えば細菌とファージ、ウイルス、真菌、および原生動物を含む）のゲノムのような外来遺伝子配列の一部でもよい。本明細書で提供される方法は特に、例えば適切な治療介入を選択するために、微生物の異なる変種または株を区別するのに応用される。ヒトおよび動物に感染し、開示された方法で検出される疾患を引き起こすウイルスの例には、特に限定されないが、レトロウイルス科 (Retroviridae)（例えば、HIV-1 (HTLV-III、LAVまたはHTLV-III/LAVとも呼ばれる；Ratnerら、Nature 313: 227-284 (1985)；Wain Hobosonら、Cell 40:9-17 (1985)；HIV-2 (Guyaderら、Nature 328: 662-669 (1987)；ヨーロッパ特許公報第0269520号；Chakrabartiら、Nature 328: 543-547 (1987)；ヨーロッパ特許出願第0655501号)のようなヒト免疫不全症ウイルス、および他の分離株、例えばHIV-LP (国際特許公報第W094/00562号)；ピコルナウイルス科 (Picornaviridae)（例えば、ポリオウイルス、A型肝炎ウイルス (Gustら、Intervirology 20:1-7 (1983))；エンテロウイルス、ヒトコクサッキーウイルス、ライノウイルス、エコーウイルス)；カルシウイルス科 (Calciviridae)（例えば胃腸炎を引き起こす株)；トガウイルス科 (Togaviridae)（例えば、ウマ脳炎ウイルス、風疹ウイルス)；フラビウイルス科 (Flaviridae)（例えば、デング熱ウイルス、脳炎ウイルス、黄熱病ウイルス)；コロナウイルス科 (Coronaviridae)（例えばコロナウイルス)；ラブドウイルス科 (Rhabdoviridae)（例えば水疱性口内炎ウイルス、狂犬病ウイルス)；フィロウイルス科 (Filoviridae)（例えばエボラウイルス)；パラミクソウイルス科 (Paramyxoviridae)（例えばパラインフルエンザウイルス、流行性耳下腺炎ウイルス、RSウイルス)；オルソミクソウイルス科 (Orthomyxoviridae)（例えばインフルエンザウイルス)；ブンガウイルス科 (Bungaviridae)（例えばハンタンウイルス、ブンガウイルス、フレボウイルス、およびナイロウイルス)；アレナウイルス科 (Arenaviridae)（出血熱ウイルス)；レオウイルス科 (Reoviridae)（例えばレオウイルス、オルビウイルス、およびロタウイルス)；ビルナウイルス科 (Birnaviridae)；ヘパドナウイルス科 (Hepadnaviridae) (B型肝炎ウイルス)；パルボウイルス科 (Parvoviridae) (パルボウイルス)；パポバウイルス科 (Papovaviridae) (パピローマウイルス、ポリオーマウイルス)；アデノウイルス科 (Adenoviridae) (ほとんどのアデノウイルス)；ヘルペスウイルス科 (Herpesviridae) (単純ヘルペスウイルス1型 (HSV-1) およびHSV-2、水痘帯状疱疹ウイルス、サイトメガロウイルス、ヘルペスウイルス)；ポックスウイルス科 (Poxviridae) (痘瘡ウイルス、ワクシニアウイルス、ポックスウイルス)；イリドウイルス科 (Iridoviridae) (例えば、アフリカ豚コレラウイルス)；および未分類ウイルス (例えば海綿状脳症の原因菌、デルタ肝炎の原因菌 (B型肝炎ウイルスの欠陥サテライトであると考えられている)、非A非B肝炎の原因菌 (クラス1 = 内部感染；クラス2 = 非経口感染、すなわちC型肝炎)；ノーウォークウイルスおよび関連ウイルス、およびアストロウイルスがある。

10

20

30

40

【0373】

感染性細菌の例には、特に限定されないが、ヘリコバクター・ピロリス (*Helicobacter pylori*)、ボレリア・ブルグドルフェリ (*Borelia burgdorferi*)、レジオネラ・ニューモフィリア (*Legionella pneumophila*)、マイコバクテリア属菌種 (例えば結核菌 (*M. tuberculosis*)、エム・アビウム (*M. avium*)、エム・イントラセルラレ (*M. intracellulare*)、エム・カンサイイ (*M. kansasii*)、エム・ゴルドナエ (*M. goodii*)、黄色ブドウ球菌 (*Staphylococcus aureus*)、淋菌 (*Neisseria gonorrhoeae*)、髄膜炎菌 (*Neisseria meningitidis*)、リステリア菌 (*Listeria monocytogenes*)、ストレプトコッカス・ピオゲネス (*Streptococcus pyogenes*) (A群連鎖球菌)、ストレプトコッカス・アガラクティエ (*Streptococcus agalactiae*) (B群連鎖球菌)、ストレプトコッカス属菌種 (ピリダンス (*viridans*) 群)、大便連鎖球菌 (*Streptococcus faecalis*)、ストレ

50

プトコッカス ボビス (*Streptococcus bovis*)、ストレプトコッカス属菌種 (嫌気性細菌種)、肺炎連鎖球菌 (*Streptococcus pneumoniae*)、病原性カンピロバクター属菌種、腸球菌属菌種、インフルエンザ菌 (*Haemophilus influenzae*)、炭疽菌 (*Bacillus anthracis*)、ジフテリア菌 (*Corynebacterium diphtheriae*)、コリネバクテリウム属菌種、豚丹毒菌 (*Erysipelothrix rhusiopathiae*)、ウェルシュ菌 (*Clostridium perfringens*)、破傷風菌 (*Clostridium tetani*)、エンテロバクター・アエロゲネス (*Enterobacter aerogenes*)、クレブシエラ・ニューモニエ (*Klebsiella pneumoniae*)、パスツレラ・ムルトシダ (*Pasturella multocida*)、バクテロイデス属菌種、フソバクテリウム・ヌクレアツム (*Fusobacterium nucleatum*)、ストレプトバシラス・モニリフォルミス (*Streptobacillus moniliformis*)、梅毒トレポネーマ (*Treponema pallidum*)、トレポネーマ・ペルテニエ (*Treponema pertenuis*)、レプトスピラ (*Leptospira*)、およびイスラエル放線菌 (*Actinomyces israelii*) がある。

10

【0374】

感染性真菌の例には、特に限定されないが、クリプトコックス・ネオフォルマンズ (*Cryptococcus neoformans*)、ヒストプラスマ・カプスラツム (*Histoplasma capsulatum*)、コクシジオイデス・イミチス (*Coccidioides immitis*)、ブラストミセス・デルマテイティディス (*Blastomyces dermatitidis*)、クラミジア・トラコマチス (*Chlamydia trachomatis*)、カンジダ・アルビカンス (*Candida albicans*) がある。他の感染性生物には、原生生物、例えば熱帯熱マラリア原虫 (*Plasmodium falciparum*)、およびトキソプラズマ原虫 (*Toxoplasma gondii*) がある。

20

【0375】

10. 抗生物質プロフィール

本明細書で提供される標的核酸断片の質量分析は、薬剤耐性 (抗生物質耐性を含む) に関するヌクレオチド変化の検出の速度と正確性を改良することができる。イソニアジド、リファンピン、ストレプトマイシン、フルオロキノロン類、およびエチオナミド耐性に関する遺伝子座は同定されている [Heymら、Lancet 344: 293 (1994)、およびMorrisら、J. Infect. Dis. 171: 954 (1995)]。イソニアジド (inh) とリファンピン (rif) をピラジナミドとエタンブトールまたはストレプトマイシンと組合せることは、結核菌 (*M. tuberculosis*) の確認症例に対する最初の攻撃として日常的に使用されている [Banerjeeら、Science 263:227 (1994)]。そのような耐性株の頻度が増加していることが、これらを検出するための、従って非効率的なおそらく有害な治療法を続ける費用と社会の健康への危険を減少させるための、迅速な測定法の開発を必要としている。薬剤耐性に関する遺伝子座のいくつかの同定は、薬剤耐性を引き起こすヌクレオチド変化の迅速なスクリーニングのための変異検出技術の採用を促進した。

30

【0376】

11. 疾患マーカーの同定

疾患の遺伝子マーカーである配列変化の迅速かつ正確な同定のための方法が本明細書で提供され、これは疾患を診断するかまたは予後を測定するのに使用することができる。遺伝子マーカーにより特徴付けられる疾患には、特に限定されないが、アテローム性動脈硬化症、肥満、糖尿病、自己免疫疾患、および癌がある。すべての生物において疾患は、遺伝したものであってもまたは環境ストレス (例えばウイルスおよび毒素) に対する体の応答に起因するものであっても、遺伝的要素を有する。現在進行中のゲノム研究の最終目標は、この情報を使用してこれらの疾患を同定、治療、および治癒させる方法を開発することである。第1段階は疾患組織をスクリーニングし、個々の試料のレベルでゲノム変化を同定することである。これらの「疾患」マーカーの同定は、間違った遺伝子または多型性を同定するためにゲノムマーカーの変化を検出する能力に依存する。ゲノムマーカー (単一ヌクレオチド多型 (SNP)、マイクロサテライト、および他の非コードゲノム領域、タンDEM繰り返し配列、イントロン、およびエキソンを含むすべての遺伝子座) は、ヒトを含むすべての生物の同定のために使用することができる。これらのマーカーは、集団を同定するのみでなく、疾患、薬物治療、環境因子に対する耐性、および他の因子に従って、集

40

50

団を分類することを可能にする。

【0377】

12. ハプロタイプ型判定

本明細書に記載の方法はハプロタイプ型を検出するのに使用することができる。すべての2倍体細胞には、任意の遺伝子、または少なくとも1つの区別できる変化を含む染色体セグメントに、2つのハプロタイプがある。多くのよく研究された遺伝子系ではハプロタイプは、単一ヌクレオチド変化よりよく表現型に相関している。すなわちハプロタイプの測定は、種々の表現型（疾患素因または感受性、治療的加入に対する応答、および医学、動物飼育、および農業目的の他の表現型を含む）の遺伝的基礎を理解するのに有用である。

【0378】

本明細書に記載のハプロタイプ型判定法は、個体の2つの相同的染色体の1つからの配列の部分の選択を可能にし、配列のその部分上の連結SNPを遺伝子型判定することを可能にする。ハプロタイプの直接分析は情報量を増加させ、連結した疾患遺伝子の診断を改良するかまたはこれらの疾患との関連を確定する。

【0379】

13. DNA繰り返し配列

本明細書で提供される断片化に基づく方法は、DNA繰り返し配列中の配列変化の迅速な検出を可能にする。種々のDNA繰り返し配列が疾患に関連している（Thangaveluら、*Prenat. Diagn.* 18: 922-25 (1998)；Bennettら、*J. Autoimmun.* 9: 415-21 (1996)）。DNA繰り返し配列には、サテライト、ミニサテライト、およびマイクロサテライトがある。サテライトは単位サイズが2塩基単位繰り返し～約1000塩基単位繰り返し、またはそれ以上、および典型的には繰り返し単位は約1000繰り返し～約10,000繰り返しの範囲で存在する。ミニサテライトは短いタンDEM繰り返し配列（すなわちSTR）とも呼ばれ、単位サイズが3塩基単位繰り返し～約100塩基単位繰り返し、および典型的には繰り返し単位は約2繰り返し～約100繰り返し、またはそれ以上の範囲で存在し、ミニサテライトの最小長さは典型的には約500塩基になる。マイクロサテライトは単位サイズが1塩基単位繰り返し～約7塩基単位繰り返し、および典型的には繰り返し単位は約5繰り返し～約100繰り返しの範囲で存在する。マイクロサテライトは、染色体上で遺伝子の近くに位置し、遺伝子発現において役割を果たす。サテライト、ミニサテライト、またはマイクロサテライトの変化の検出は、変種または疾患に対する傾向のマーカーとして使用することができる。

【0380】

マイクロサテライト（時に、可変数のタンDEM繰り返し配列、すなわちVNTRと呼ばれる）は、1～7個またはそれ以上塩基の短いタンDEM繰り返しヌクレオチド単位であり、この中で最も顕著なものはジ-、トリ-、およびテトラヌクレオチド繰り返し配列である。マイクロサテライトはゲノムDNA中で100,000bp毎に存在する（J.L. WeberとP.E. Can, *Am. J. Hum. Genet.* 44: 388 (1989)；J. Weissenbachら、*Nature* 359: 794 (1992)）。例えばCAジヌクレオチド繰り返し配列は、ヒトミトコンドリア外ゲノムの約0.5%を構成する；CTとAG繰り返し配列は一緒に約0.2%を構成する。CG繰り返し配列は、おそらくはCpGアイランドの制御機能のためにまれである。マイクロサテライトは長さに関して極めて多型性であり、全ゲノムにわたって分布しており、非コード配列中に最も豊富に有り、ゲノム内のその機能は未知である。

【0381】

集団は、その集団に特徴的であり他の集団（これは異種交配しない）とは区別される種々のマイクロサテライトを維持するため、マイクロサテライトは法医学的応用に重要である。

【0382】

マイクロサテライト内の多くの変化はサイレントであるが、いくつかは遺伝子産物または発現レベルで大きな変化を招く。例えば遺伝子のコード領域中に存在するトリヌクレオチド繰り返し配列は、ある腫瘍中で影響を受け（C.T. Caskeyら、*Science* 256:784 (1992)）、およびマイクロサテライトの変化は癌への素因を引き起こす遺伝的不安定性の原因

10

20

30

40

50

となる (P.J. McKinnen, Hum. Genet. 1(75): 197 (1987); J. Germanら、Clin. Genet. 35: 57 (1989))。

【0383】

また本明細書に記載の方法を使用して、例えばSTR領域を含まないゲノムの参照ゲノム配列と比較して、ゲノムのいくつかの標的配列中のマイクロサテライトまたは短いタンDEM繰り返し配列 (STR) を同定することができる。STR領域は、どの疾患または症状にも関連しない多型性領域である。ヒトゲノム中の多くの遺伝子座は多型性の短いタンDEM繰り返し配列 (STR) 領域を含む。STR遺伝子座は、長さが3~100塩基対の短い繰り返し配列要素を含有する。200,000の予測されるトリマーおよびテトラマーSTRがあると推定され、これはヒトゲノム中で15kb毎に1つという高頻度で存在する (例えば、国際PCT特許出願第 WO9213969A1号、Edwardsら、Nucl. Acids Res. 19: 4791 (1991); Beckmannら、Genomics 12: 627-631 (1992)を参照)。ほとんど半分のこれらのSTR遺伝子座は多型性であり、遺伝マーカーの豊富な供給源となる。特定の遺伝子座中における繰り返し単位の数の変化は、可変ヌクレオチドタンDEM繰り返し (VNTR) 遺伝子座 (Nakamuraら、Science 235: 1616-1622 (1987)); およびミニサテライト遺伝子座 (Jeffreysら、Nature 314: 67-73 (1985)) (これはより長い繰り返し単位を含有する)、およびマイクロサテライトまたはジヌクレオチド繰り返し遺伝子座 (Lutyら、Nucleic Acids Res. 19: 4308 (1991); Littら、Nucleic Acids Res. 18: 4301 (1990); Littら、Nucleic Acids Res. 18: 5921 (1990); Lutyら、Am. J. Hum. Genet. 46: 776-783 (1990); Tautz、Nucl. Acids Res. 17: 6463-6471 (1989); Weberら、Am. J. Hum. Genet. 44: 388-396 (1989); Beckmannら、Genomics 12: 627-631 (1992)) を思わせ次に観察される多型性に関する。

10

20

【0384】

STR遺伝子座の例には、特に限定されないが、ヒトCD4遺伝子座中のペンタヌクレオチド繰り返し配列 (Edwardsら、Nucl. Acids Res. 19: 4791 (1991)); ヒトアロマターゼチトクロームP-450遺伝子中のテトラヌクレオチド繰り返し配列 (CYP19; Polymeropoulosら、Nucl. Acids Res. 19: 195 (1991)); ヒト凝固XIII A因子サブユニット遺伝子中のテトラヌクレオチド繰り返し配列 (F13A1; Polymeropoulosら、Nucl. Acids Res. 19: 4306 (1991)); F13B遺伝子座中のテトラヌクレオチド繰り返し配列 (Nishimuraら、Nucl. Acids Res. 20: 1167 (1992)); ヒトc-les/fps、プロト癌遺伝子中のテトラヌクレオチド繰り返し配列 (FES; Polymeropoulosら、Nucl. Acids Res. 19: 4018 (1991)); LFL遺伝子中のテトラヌクレオチド繰り返し配列 (Zulianiら、Nucl. Acids Res. 18: 4958 (1990)); ヒト膵臓ホスホリパーゼA-2遺伝子におけるトリヌクレオチド繰り返し多型性 (PLA2; Polymeropoulosら、Nucl. Acids Res. 18: 7468 (1990)); VWF遺伝子中のテトラヌクレオチド繰り返し多型性 (Ploosら、Nucl. Acids Res. 18: 4957 (1990)); およびヒト甲状腺ペルオキシダーゼ (hTPO) 遺伝子座中のテトラヌクレオチド繰り返し配列 (Ankerら、Hum. Mol. Genet. 1: 137 (1992)) がある。

30

【0385】

14. 対立遺伝子変種の検出

本明細書に記載の方法は、対立遺伝子変種の高速処理性の迅速かつ正確な検出を可能にする。対立遺伝子変種の研究は、複雑な背景の特異的配列の検出のみでなく、わずかのまたは1つのみのヌクレオチドの差を有する配列の区別を可能にする。PCRによる対立遺伝子特異的変種を検出するための1つの方法は、鋳型鎖とプライマーの3'末端とで mismatches があると、TaqポリメラーゼはDNAを合成することが困難であるという事実に基づく。対立遺伝子特異的変種は、可能な対立遺伝子の1つのみと完全に一致するプライマーの使用により検出でき、他の対立遺伝子に対する mismatches はプライマーの伸長を妨害するように作用し、従ってその配列の増幅を妨害する。この方法は、 mismatches の塩基組成が mismatches を超える伸長を妨害する能力に影響を与えるという点で大きな限界があり、いくつかの mismatches は伸長を妨害しないかまたはわずかな効果しかない (Kwokら、Nucl. Acids Res. 18: 999 [1990])。本明細書に記載の断片化とハイブリダイゼーションに基づく方法は、プライマー伸長法の限界を克服する。

40

50

【0386】

15. 対立遺伝子頻度の測定

本明細書に記載された方法は、その頻度が年齢、民族、性、またはいくつかの他の基準の関数として集団内で変化する1つまたはそれ以上の遺伝的マーカーを同定するのに有用である。例えばApoE遺伝子型の年齢依存性分布は当該分野で公知である（Schaechterら、Nature Genetics 6: 29-32 (1994)参照）。どこかのレベルで疾患と関連していることが知られている多型性の頻度はまた、疾患状態の進行を検出または追跡するのに使用できる。例えばリポタンパク質リパーゼ遺伝子のN291S多型（N291S）（これはアミノ酸コドン291でアスパラギンの代わりにセリンの使用を引き起こす）は、動脈硬化症特に心筋梗塞の男性のリスクの上昇に関連する高密度リポタンパク質コレステロール（HDL-C）のレベルを低下させる（Reymerら、Nature Genetics 10: 28-34 (1995)参照）。さらに対立遺伝子頻度の変化を測定することは、未知の多型の同定および最終的には疾患の発生と進行に關与する遺伝子もしくは経路の同定を可能にする。

10

【0387】

16. エピジェネティクス（Epigenetics）

本明細書に記載の方法を使用して、配列に基づかない、参照核酸と比較した標的核酸またはタンパク質中の変化（核酸の天然に存在するモノマー単位である塩基の本体）を研究することができる。例えば本明細書に記載の方法に使用される特異的切断試薬は、配列非依存性の特徴（例えば、メチル化パターン、修飾塩基の存在、または標的分子と参照分子との高次構造の差）の差を認識して、配列非依存性部位で切断される断片を生成することができる。エピジェネティクス（Epigenetics）は、遺伝子配列の差ではない遺伝子発現の差に基づく情報の遺伝の研究である。エピジェネティクス的变化は、遺伝子機能の有糸分裂的および/または減数分裂的に遺伝可能な変化、または核酸配列の変化では説明できない高次核酸構造の変化を意味する。エピジェネティクス的变化を受ける特徴の例には、特に限定されないが、動物のDMAメチル化パターン、ヒストン修飾、およびポリコーム-トリソックス（Polycomb-trithorax）群（Pc-G/trx）タンパク質複合体がある（例えばBird, A., Genes Dev. 16: 6-21 (2002)参照）。

20

【0388】

エピジェネティクス的变化は通常（必ずではないが）、通常（必ずではないが）遺伝可能な遺伝子発現の変化を引き起こす。例えば上記したようにメチル化パターンの変化は、癌や他の疾患の発生と進行の早期事象である。多くの癌では異常メチル化のために、いくつかの遺伝子は不適切にスイッチオフまたはスイッチオンされる。転写を抑制または活性化するメチル化パターンの能力は遺伝することができる。Pc-G/trxタンパク質複合体はメチル化のように、遺伝可能な形で転写を抑制または活性化する。Pc-G/trxマルチタンパク質アセンブリーはゲノムの特異的領域に標的化され、これは、遺伝子が活性でも不活性でも遺伝子の胚性遺伝子発現状態を有効に凍結し、この状態を発生を通して安定に遺伝していく。Pc-G/trx群のタンパク質がゲノムを標的化し結合する能力は、ゲノムに含有される遺伝子の発現レベルにのみ影響を与え、遺伝子産物の性質には影響を与えない。本明細書に記載の方法は、配列非依存性変化（例えばエピジェネティクス的变化）に基づく参照配列と比較した標的配列の変化を同定する特異的切断試薬とともに使用することができる。

30

40

【実施例】

【0389】

実施例1

基礎となるDNA配列を再構築するために、本例に記載され例示される方法を使用して、ハイブリダイゼーションによる配列決定のヌクレオチド配列分析のための技術、ならびに質量スペクトル法によるヌクレオチド配列分析のための技術を使用することができる。特に実験データをデブリイジン（de Bruijn）グラフのサブグラフに変換することができる（Pevzner, J. Biomol. Struct. Dyn., 7: 63-73 (1989)を参照）。次にこのグラフのオイラーの路（Eulerian paths）を検索することができ、ここでサイクルとバルジをあらかじめ破壊する必要がある：Pevznerら、Proc. Natl. Acad. Sci. USA 98: 9748-9753 (200

50

1) 参照。

【 0 3 9 0 】

例えばACATGAGCTTACAAC (配列番号1) を対象のDNA配列とする。切断反応はこのDNA (またはRNA) 分子を非特異的に5~7ntの断片に切断する。最後に生じた断片を、それぞれの縮重塩基がプリン (文字R、A、またはG) またはピリミジン (文字Y、C、またはT) に結合する4つの縮重塩基を有する16個の位置を含有するハイブリダイゼーションチップに結合させる。この縮重アルファベットでは、対象の配列はRYRYRRRYYYRYRRYとなる。次に以下の結合パターンがチップ上で起きる。

【 0 3 9 1 】

【表 2】

縮重パターン	ハイブリダイゼーション点に結合するフラグメント
RRRR	(フラグメントなし)
RRRY	CATGAGC, ATGAGC, ATGAGCT, TGAGC, TGAGCT, GAGCTT, GAGCT, GAGCTT, GAGCTTA
RRYR	(フラグメントなし)
RRYY	ATGAGCT, TGAGCT, TGAGCTT, GAGCT, GAGCTT, GAGCTTA, AGCTT, AGCTTA, AGCTTAC
RYRR	ACATGA, ACATGAG, CATGA, CATGAG, CATGAGC, ATGAG, ATGAGC, ATGAGCT, CTTACAA, TTACAA, TTACAAC
RYRY	ACATG, ACATGA, ACATGAG
RYYR	(フラグメントなし)
RYYY	TGAGCTT, GAGCTT, GAGCTTA, AGCTT, AGCTTA, AGCTTAC, GCTTA, GCTTAC, GCTTACA
YRRR	ACATGAG, CATGAG, CATGAGC, ATGAG, ATGAGC, ATGAGCT, TGAGC, TGAGCT, TGAGCTT
YRRY	TTACAAC
YRYR	ACATG, ACATGA, ACATGAG, CATGA, CATGAG, CATGAGC, GCTTACA, CTTACA, CTTACAA, TTACA, TTACAA, TTACAAC
YRYY	(フラグメントなし)
YYRR	(フラグメントなし)
YYRY	AGCTTAC, GCTTAC, GCTTACA, CTTAC, CTTACA, CTTACAA, TTACA, TTACAA, TTACAAC
YYYY	GAGCTTA, AGCTTA, AGCTTAC, GCTTA, GCTTAC, GCTTACA, CTTAC, CTTACA, CTTACAA
YYYY	(フラグメントなし)

【 0 3 9 2 】

質量スペクトル分析を使用して断片の組成を測定することができる、例えばBoecker、Lect. Notes Comp. Sci. 2812: 476-487 (2003)を参照。次に以下の成分に対応する質量スペクトルが測定される。

【 0 3 9 3 】

10

20

30

40

【表3】

縮重パターン	ハイブリダイゼーション点上で検出されるコンポマー
RRRR	(ピークなし)
RRRY	A ₂ C ₂ G ₂ T ₁ , A ₂ C ₁ G ₂ T ₁ , A ₂ C ₁ G ₂ T ₂ , A ₁ C ₁ G ₂ T ₁ , A ₁ C ₁ G ₂ T ₂ , A ₁ C ₁ G ₂ T ₃ , A ₁ C ₁ G ₂ T ₁ , A ₁ C ₁ G ₂ T ₂ , A ₂ C ₁ G ₂ T ₁
RRYR	(ピークなし)
RRYY	A ₂ C ₁ G ₂ T ₂ , A ₁ C ₁ G ₂ T ₂ , A ₁ C ₁ G ₂ T ₃ , A ₁ C ₁ G ₂ T ₁ , A ₁ C ₁ G ₂ T ₂ , A ₂ C ₁ G ₂ T ₂ , A ₁ C ₁ G ₁ T ₂ , A ₂ C ₁ G ₁ T ₂ , A ₂ C ₂ G ₁ T ₂
RYRR	A ₃ C ₁ G ₁ T ₁ , A ₃ C ₁ G ₂ T ₁ , A ₂ C ₁ G ₁ T ₁ , A ₂ C ₁ G ₂ T ₁ (2回), A ₂ C ₂ G ₂ T ₁ , A ₂ G ₂ T ₁ , A ₂ C ₁ G ₂ T ₁ , A ₂ C ₁ G ₂ T ₂ , A ₃ C ₂ T ₂ (2回), A ₃ C ₁ T ₂
RYRY	A ₂ C ₁ G ₁ T ₁ , A ₃ C ₁ G ₁ T ₁ , A ₃ C ₁ G ₂ T ₁
RYYR	(ピークなし)
RYYY	A ₁ C ₁ G ₂ T ₃ , A ₁ C ₁ G ₂ T ₂ , A ₂ C ₁ G ₂ T ₂ , A ₁ C ₁ G ₁ T ₂ (2回), A ₂ C ₁ G ₁ T ₂ , A ₂ C ₂ G ₁ T ₂ (2回), A ₁ C ₂ G ₁ T ₂
YRRR	A ₃ C ₁ G ₂ T ₁ , A ₂ C ₁ G ₂ T ₁ (2回), A ₂ C ₂ G ₂ T ₁ , A ₂ G ₂ T ₁ , A ₂ C ₁ G ₂ T ₂ , A ₁ C ₁ G ₂ T ₁ , A ₁ C ₁ G ₂ T ₂ , A ₁ C ₁ G ₂ T ₃
YRRY	A ₃ C ₂ T ₂
YRYR	A ₂ C ₁ G ₁ T ₁ (2回), A ₃ C ₁ G ₁ T ₁ , A ₃ C ₁ G ₂ T ₁ , A ₂ C ₁ G ₂ T ₁ , A ₂ C ₂ G ₂ T ₁ , A ₂ C ₂ G ₁ T ₂ , A ₂ C ₂ T ₂ , A ₃ C ₂ T ₂ (2回), A ₂ C ₁ T ₂ , A ₃ C ₁ T ₂
YRYY	(ピークなし)
YYRR	(ピークなし)
YYRY	A ₂ C ₂ G ₁ T ₂ (2回), A ₁ C ₁ G ₁ T ₂ , A ₁ C ₂ T ₂ , A ₂ C ₂ T ₂ , A ₃ C ₂ T ₂ (2回), A ₂ C ₁ T ₂ , A ₃ C ₁ T ₂
YYYR	A ₂ C ₁ G ₂ T ₂ , A ₂ C ₁ G ₁ T ₂ , A ₂ C ₂ G ₁ T ₂ (2回), A ₁ C ₁ G ₁ T ₂ , A ₁ C ₂ G ₁ T ₂ , A ₁ C ₂ T ₂ , A ₂ C ₂ T ₂ , A ₃ C ₂ T ₂
YYYY	(ピークなし)

10

20

30

【0394】

この情報は以下のように分枝限定 (branch-and-bound) 検索で使用される: ACATGAGが正しい配列の既知のコードであるとする。次の塩基の本体はランダムに帰属することができ、次に1つまたはそれ以上の質量スペクトルと比較される。次の塩基にAを帰属すると、いくつかの質量スペクトルの以下の断片と成分が予測される。

40

【0395】

【表 4】

フラグメント：	コンポマー：	対応するスペクトル：
CATGAGA	A ₃ C ₁ G ₂ T ₁	YRYR, RYRR, YRRR, RRRR
ATGAGA	A ₃ G ₂ T ₁	RYRR, YRRR, RRRR
TGAGA	A ₂ G ₂ T ₁	YRRR, RRRR

10

【0396】

質量スペクトルはこの仮説に矛盾する：ACATGAGAがこの遺伝子座の正しいヌクレオチドなら、ハイブリダイゼーション位置RRRRに対応する質量スペクトルは少なくとも3つのピークを含むであろう。しかしこのスペクトルではピークが1つも検出されない。この決定は、4つの質量スペクトルで9つのピークが観察されるかまたは観察されないことに基づき、従って極めて頑強性である。同様の推論は、GもTもコードACATGAGに結合しないことを示す。

【0397】

これに対して、塩基CをコードACATGAGに付加すると、いくつかの異なる質量スペクトルで以下の断片と組成が生成するであろう。

20

【0398】

【表 5】

フラグメント：	コンポマー：	対応するスペクトル：
CATGAGC	A ₂ C ₂ G ₂ T ₁	YRYR, RYRR, YRRR, RRRY
ATGAGC	A ₂ C ₁ G ₂ T ₁	RYRR, YRRR, RRRY
TGAGC	A ₁ C ₁ G ₂ T ₁	YRRR, RRRY

30

【0399】

すべての9つのピークが4つの異なる質量スペクトルで観察されるため、Cが付加すべき正しい文字である。上記方法によりさらに複雑な切断パターンも分析でき、この方法の頑強性はまた、これらの複雑な状況にも受け継がれる。

【0400】

当業者にはこれらの修飾が明らかであり、従って本発明は請求の範囲によってのみ限定されるものである。

40

【図面の簡単な説明】

【0401】

【図1】図1は、重複断片の作製を示す。

【図2】図2は、固体支持体上の縮重捕獲オリゴヌクレオチドにハイブリダイズする複数の断片を示す。

【図3】図3は、ハイブリダイズした捕獲オリゴヌクレオチド：標的断片2本鎖の「トリミング」を示す。

【 図 1 】

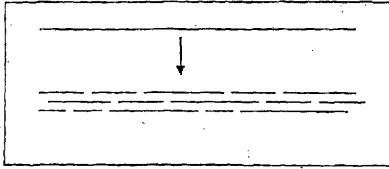


Figure 1

【 図 2 】

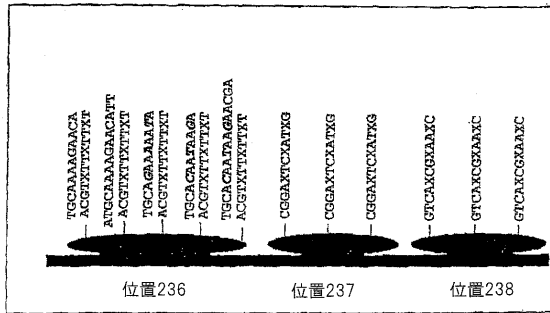


Figure 2

【 図 3 】

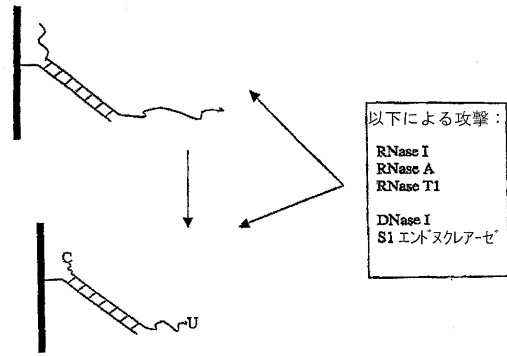


Figure 3

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US05/32441
A. CLASSIFICATION OF SUBJECT MATTER IPC: C12Q 1/68(2006.01);C12P 19/34(2006.01),19/24(2006.01);C07H 21/04(2006.01) USPC: 435/6,91.1,91.4,94;536/24.3 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 435/6,91.1,91.4,94;536/24.3 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EAST--nucleic, fragment, mass spectrometer, reference mass, sequencing		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 6,297,006 B1 (DRMANAC et al) 02 October 2001 (02.10.2001), see entire document.	1-2, 7-27, 29-30, 32, 35-45 ----- 3-6, 28, 31, 33-34, 46-53
Y	US 5,003,059 A (BRENNAN) 26 March 1991 (23.03.1991), see entire document.	3-6
Y	US2002/0120127 A1 (CHURCH et al) 29 August 2002 (29.08.2002), see entire document.	33-34
Y	US 5,453,247 A (BEAVIS et al) 26 September 1995 (26.09.1995), see entire document.	46-53
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents:		
"A"	document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E"	earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P"	document published prior to the international filing date but later than the priority date claimed	
Date of the actual completion of the international search 29 September 2006 (29.09.2006)		Date of mailing of the international search report 08 NOV 2006
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201		Authorized officer <i>Carolina Lawrence Fox</i> Ram Shukla Telephone No. (571) 272.1600

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, LV, 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, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, 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

(72)発明者 ファン デン ブーム, ディルク ヨハネス
アメリカ合衆国, カリフォルニア 92037, ラ ホーヤ, ノーチラス ストリート 385

(72)発明者 ボエッカー, セバスティアン
ドイツ連邦共和国, 33602 ビーレフェルト, ラーフェンスベルガーシュトラッセ 52

Fターム(参考) 4B024 AA11 CA01 CA11 HA08 HA14 HA19
4B063 QA13 QA18 QQ42 QQ52 QR14 QR32 QR84 QS34 QS36 QX01
QX04