

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7459010号

(P7459010)

(45)発行日 令和6年4月1日(2024.4.1)

(24)登録日 令和6年3月22日(2024.3.22)

(51)国際特許分類

F I

C 1 2 N 15/11 (2006.01)

C 1 2 N 15/11

Z Z N A

C 1 2 Q 1/6813(2018.01)

C 1 2 Q 1/6813

Z

C 1 2 Q 1/6851(2018.01)

C 1 2 Q 1/6851

Z

C 1 2 Q 1/686(2018.01)

C 1 2 Q 1/686

Z

C 1 2 Q 1/6869(2018.01)

C 1 2 Q 1/6869

Z

請求項の数 16 (全134頁)

(21)出願番号 特願2021-52809(P2021-52809)

(22)出願日 令和3年3月26日(2021.3.26)

(62)分割の表示 特願2018-208639(P2018-208639)
の分割

原出願日 平成25年7月3日(2013.7.3)

(65)公開番号 特開2021-104031(P2021-104031
A)

(43)公開日 令和3年7月26日(2021.7.26)

審査請求日 令和3年4月22日(2021.4.22)

(31)優先権主張番号 61/667,919

(32)優先日 平成24年7月3日(2012.7.3)

(33)優先権主張国・地域又は機関
米国(US)

(31)優先権主張番号 61/745,435

(32)優先日 平成24年12月21日(2012.12.21)

最終頁に続く

(73)特許権者 510288529

インテグレイテッド・デー・エヌ・エ
イ・テクノロジーズ・インコーポレイテ
ッドアメリカ合衆国、アイオワ・5 2 2 4 1
、コーラルビル、コマーシャル・パーク
・1 7 1 0

(73)特許権者 515002735

ファンデーション・メディスン・インコ
ーポレイテッド
アメリカ合衆国、マサチューセッツ・0
2 1 3 9、ケンブリッジ、ワン・ケンド
ール・スクエア、スイート・ビー・3 5
0 1

(74)代理人 110001173

最終頁に続く

(54)【発明の名称】 Tm増強ブロッキングオリゴヌクレオチド、ならびに標的濃縮の改善およびオフターゲット
選択の低減のためのペイト

(57)【特許請求の範囲】

【請求項1】

末端アダプターを有する所望の鑄型核酸の選択方法における使用のためのオリゴヌクレ
オチドであって、以下の一つから選択されるオリゴヌクレオチド、

配列番号 2、3、4、5、6、7、8、10、11、12、13、14、15、16、1
8、19、21、22、25、27、28、32、34 および 36

またはその組み合わせ。

【請求項2】

末端アダプターが、15個のヌクレオチドから75個のヌクレオチドのサイズ範囲を有
する、請求項1に記載のオリゴヌクレオチド。

【請求項3】

末端アダプターがバーコードメインを含む、請求項1に記載のオリゴヌクレオチド。

【請求項4】

請求項1～3のいずれか一項に記載のオリゴヌクレオチドを含む、末端アダプターを有
する所望の鑄型核酸の選択方法における使用のためのキット。

【請求項5】

請求項1に記載のオリゴヌクレオチドであって、

該オリゴヌクレオチドは少なくとも1つのTm増強基を含み、少なくとも1つのTm増強
基が、ロックド核酸基、二環式核酸基、C5修飾ピリミジン、ペプチド核酸基およびこれ
らの組合せからなる群から選択される少なくとも1つのメンバーであって、少なくとも1

つのT_m増強基が1、4℃から25℃を含む範囲の増強T_m値を提供する、請求項1に記載のオリゴヌクレオチド。

【請求項6】

オリゴヌクレオチドが、所望の鑄型の少なくとも1つの配列に相補的である、請求項5に記載のオリゴヌクレオチド。

【請求項7】

オリゴヌクレオチドが、ブロッカーまたはベイトから選択される少なくとも1つのメンバーを含む、請求項5に記載のオリゴヌクレオチド。

【請求項8】

ロックド核酸基または二環式核酸基が、シトシン、アデニンおよびチミンからなる群から選択される核酸塩基を含む、請求項5に記載のオリゴヌクレオチド。

10

【請求項9】

オリゴヌクレオチドがブロッカーを含む、請求項5に記載のオリゴヌクレオチド。

【請求項10】

ブロッカーが、所望の鑄型核酸の末端における少なくとも1つの配列に対する配列相補性を有する、請求項9に記載のオリゴヌクレオチド。

【請求項11】

ブロッカーが、複数のヌクレオチドを有するバーコードドメインをさらに含む、請求項9に記載のオリゴヌクレオチド。

【請求項12】

複数のヌクレオチドが、連続して位置する5個から12個のヌクレオチドを含む、請求項11に記載のオリゴヌクレオチド。

20

【請求項13】

バーコードドメインが、アデニン、チミン、シトシン、グアニン、イノシン、3-ニトロピロール、5-ニトロインドール、およびこれらの組合せから選択される群から選択される少なくとも1つのメンバーを核酸塩基として有するヌクレオチドを含む、請求項11に記載のオリゴヌクレオチド。

【請求項14】

オリゴヌクレオチドからのポリメラーゼ指向性合成を防止する、3'末端修飾をさらに含む、請求項5に記載のオリゴヌクレオチド。

30

【請求項15】

3'末端修飾が、3'-デオキシヌクレオチド、2', 3'-ジデオキシヌクレオチドまたは3'-スパーサーC3基を含む、請求項4に記載のオリゴヌクレオチド。

【請求項16】

3'-デオキシヌクレオチド、2', 3'-ジデオキシヌクレオチドまたは3'-スパーサーC3基から選択される3'末端修飾をさらに含む、請求項5に記載のオリゴヌクレオチド。

【発明の詳細な説明】

【技術分野】

【0001】

40

この出願は、35 U. S. C. 119により、2012年7月3日に出願され、「METHODS AND COMPOSITIONS FOR REDUCING OFF-TARGET SELECTION」という名称である米国仮出願第61/667,919号、および2012年12月21日に出願され、「T_M-ENHANCED BLOCKING OLIGONUCLEOTIDES AND BAITS FOR IMPROVED TARGET ENRICHMENT IN MASSIVELY PARALLEL SEQUENCING EXPERIMENTS」という名称である米国仮出願第61/745,435号からの優先権の利益を請求し、これら両方の内容は参照によりこれらの全体を本明細書に組み込む。

【0002】

50

この発明は、修飾オリゴヌクレオチド組成物、ならびに核酸選択および配列決定のための方法におけるこれらの使用に関する。特に、本発明は、ブロッカーおよびベイトとしての T_m 増強オリゴヌクレオチド、ならびに標的濃縮の改善およびオフターゲット選択の低減のための他の試薬に属する。オリゴヌクレオチド組成物および試薬は、次世代配列決定用途のための核酸鋳型を調製するための強力な用途を見出す。

【背景技術】

【0003】

核酸ハイブリダイゼーションは、核酸の同定、選択および配列決定に属するバイオテクノロジー応用において重要な役割を有する。標的材料としてゲノム核酸を用いる配列決定用途は、高複合混合物から目的の核酸標的を選択することを要求する。配列決定努力の質は、選択プロセスの効率に依存し、これはさらに、核酸標的が非標的配列に対して、どのくらい良好に濃縮され得るかにかかっている。

【0004】

ゲノムDNAまたはcDNAなど核酸の複合プールから所望の配列を濃縮するために、様々な方法が使用されてきた。これらの方法としては、ポリメラーゼ連鎖反応（PCR）、分子反転プローブ（MIP）、またはハイブリッド形成による配列捕捉（「ハイブリッド捕捉」；例えば、Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., Howard, E., Shendure, J. および Turner, D. J. (2010) 「Target-enrichment strategies for next-generation sequencing」、Nat. Methods 7:111-118 を参照されたい。）が挙げられる。ハイブリッド捕捉は、この方法が、他の方法と比較して、より少ない標的核酸の酵素増幅または操作手順を必要とすることにおいて他の方法を超える利点を提供する。ハイブリッド捕捉方法は、結果として最終の配列決定ライブラリーに、より少ないエラーを導入する。この理由のため、ハイブリッド捕捉方法は、核酸の複合プールから所望の配列を濃縮するための好ましい方法であり、次世代配列決定（NGS）用途において鋳型を調製するのに理想的である。

【0005】

NGS用途は、通常、使用されるNGSプラットフォームに依存して、長いゲノムDNAまたはcDNAを、100-3,000bp長のサイズ分布を有するより小さい断片サイズに無作為に切断することに関与する。DNA末端は、酵素的に処理されてライゲーションを容易にし、ユニバーサルDNAアダプターは、末端にライゲートされて、その結果得られるNGS鋳型を提供する。末端アダプター配列がプライマーハイブリダイゼーションのためのユニバーサル部位を提供して、所望のDNA標的のクローン拡大が達成され、NGS用途において使用される自動化配列決定プロセスに導入され得る。ハイブリッド捕捉方法は、目的の特異的配列が濃縮されている例えば 3×10^9 塩基（ヒトゲノム）から 10^3 から 10^8 塩基のずっと小さいサブセットまでのランダムDNA断片のプールの複雑性を低減すると意図される。このプロセスの効率は、出発複合プールからの所望のDNA配列のために達成される捕捉および濃縮の質に直接的に関係する。

【0006】

NGS用途は、典型的に、以下の方式における濃縮のハイブリッド捕捉方法を使用する。NGS鋳型の調製プールは熱変性され、捕捉プローブオリゴヌクレオチド（「ベイト」）のプールと混合される。ベイトは、標的ゲノム内で目的の領域にハイブリダイズするように設計されており、通常、60-200塩基長であり、さらに、これらのプローブの後続の捕捉を可能にするリガンドを含有するように修飾される。1つの共通の捕捉方法は、ベイト上にビオチン基（単数または複数）を組み込む。ハイブリダイゼーションが完了してDNA鋳型：ベイトハイブリッドを形成した後に、ベイトに対してだけ親和性を有する構成成分を用いて捕捉が行われる。例えば、ストレプトアビジン-磁性ビーズが使用されて、ハイブリダイズされているビオチン化ベイトのビオチン部分を、NGS鋳型のプールからの所望のDNA標的に結合することができる。洗浄して非結合核酸を除去し、保持さ

10

20

30

40

50

れた材料の複雑性を低減する。保持された材料は、次いで、磁性ビーズから溶出され、自動化配列決定プロセスに導入される。

【0007】

ベイトとのDNAハイブリダイゼーションは極めて特異的であり得るが、望まれない配列が、ハイブリッド捕捉方法の完了に続く濃縮プール中に残る。これらの望まれない配列の最も大きい部分は、ベイトに対する相補性を有さないNGS鑄型とそれを有するNGS鑄型との間の所望されないハイブリダイゼーション事象により存在する。ハイブリッド捕捉方法において起こる所望されないハイブリダイゼーションの2つの型としては、以下の配列が挙げられる：(1) 内在性ゲノムDNAにおいて見出される高反復性DNA要素；および(2) プールのNGS鑄型の各々に操作される末端アダプター配列。

10

【0008】

複合プールにおける1つのDNA断片中に存在する、A l u配列またはL I N E配列などの反復性の内在性DNA要素は、別の非関連性DNA断片中に存在する別の同様の要素にハイブリダイズすることができる。本来はゲノム内の非常に異なる位置に由来し得るこれらの断片は、ハイブリッド捕捉方法のハイブリダイゼーションプロセス中に連結される。これらのDNA断片の1つが、ベイトのための結合部位を含有する所望の断片を表すならば、望まれない断片は所望の断片と一緒に捕捉される。このクラスの望まれないNGS鑄型は、過剰な反復要素をハイブリダイゼーション反応に付加することによって低減され得る。最も共通して、ヒトのC o t - 1 DNAがハイブリダイゼーション反応に付加され、反応が、標的中のA l u、L I N Eおよび他の反復部位を結合させ、これに基づいて互いに相互作用するNGS鑄型の能力をブロックする。

20

【0009】

ハイブリッド捕捉中に回復される望まれないNGS鑄型の、より問題のあるクラスは、プールのNGS鑄型の各々の上で操作される末端アダプター配列間の相互作用から起こる。NGS鑄型のプールは、典型的に、あらゆるDNA断片上に同一の末端アダプター配列を含有するので、アダプター配列は、ハイブリダイゼーション溶液中に非常に高い有効濃度(単数または複数)で存在する。その結果として、非関連性NGS鑄型は、これらの末端を介して互いにアニールし、これによって、一緒に連結される別段の非関連性DNA断片の「デージー鎖」をもたらすことがある。それで、これらの連結断片の1つがベイトのための結合部位を含有するならば、全デージー鎖が捕捉される。このやり方で、単一の所望の断片の捕捉が、多数の所望されない断片をもたらす恐れがあり、このことが、所望の断片の濃縮の全体的効率を低減する。このクラスの望まれない捕捉事象は、過剰な一本鎖アダプター配列をハイブリダイゼーション反応に付加することによって低減され得る。なお、過剰なアダプター配列を用いるいわゆるデージー鎖捕捉事象を有効に低減するための能力は、所望の断片を捕捉するための効率が約50% - 60%に制限される。

30

【0010】

ハイブリダイゼーション反応におけるC o t - 1 DNAおよびアダプターブロッキングオリゴヌクレオチドの使用にもかかわらず、混入する望まれないDNA断片の相当量がハイブリッド捕捉ステップ後の配列決定プールに残るが、なぜならば主に、ブロッキング方法が完全に成功していないからである。したがって、目的の標的のより大きな部分および興味対象でないより少ない標的を配列決定することに供給源を充てることができるように、捕捉効率を改善するとともに所望されない配列からの混入を低減する必要がある。

40

【0011】

したがって、オフターゲット核酸相互作用は、捕捉プローブ、例えばオリゴヌクレオチドベイトに対するハイブリダイゼーション(例えば、溶液ハイブリダイゼーション)によって、標的核酸の選択の効率を制限する恐れがある。オフターゲット選択は、例えば、ハイブリダイゼーション捕捉および／またはアーチファクト的ハイブリッド捕捉の収率減少の1つ以上をもたらすことができ、これが順じて、後続のステップ、例えば配列決定における非効率に至る。

【0012】

50

オフターゲット選択は、ハイブリッド選択のストリンジェンシー条件が低減される場合、例えば、より低い核酸溶融温度を有する標的：捕捉物二重鎖（例えば、RNA：DNA二重鎖と比較してのDNA：DNA二重鎖）のために選択する場合に典型的に増加される。したがって、オフターゲット配列の捕捉は、DNA：DNAハイブリダイゼーションにおいてよりも問題であり得る。

【0013】

典型的に、ライブラリーメンバーとしては、ライブラリー挿入物、しばしば、目的の遺伝子からの配列のセグメント、例えば、配列決定するためのセグメントが挙げられる。メンバーがオンターゲットであるならば、ライブラリー挿入物は、捕捉プローブと二重鎖を形成する。典型的に、ライブラリーメンバーとしては、1つ以上の非標的配列も挙げられる。これらは典型的に、目的の遺伝子の部分ではなく、むしろアダプター配列、増幅プライマーもしくは増幅タグ、またはバーコードタグである。捕捉プローブハイブリダイズ化ライブラリーメンバーの非標的配列は、反応混合物における他の配列との二重鎖形成によって、所望されない配列、例えばオフターゲットライブラリーメンバーの選択に至る恐れがある。理論によって束縛されることを望まないが、オンターゲットライブラリーメンバーとオフターゲット配列との間の鎖状体形成は、オフターゲット配列の選択をもたらす恐れがある。

【先行技術文献】

【非特許文献】

【0014】

【文献】Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., Howard, E., Shendure, J. および Turner, D. J. (2010) 「Target-enrichment strategies for next-generation sequencing」、Nat. Methods 7:111-118

【発明の概要】

【発明が解決しようとする課題】

【0015】

オフターゲット核酸の選択を最小化するため、例えば、捕捉プローブで二重鎖を形成しないライブラリーメンバーの選択を最小化するための方法および組成物が、本明細書において開示されている。非標的配列、例えばアダプター媒介選択を低減する方法および組成物が本明細書において開示されている。

【課題を解決するための手段】

【0016】

（発明の要旨）

一態様において、本発明は、少なくとも1個の T_m 増強基を有するオリゴヌクレオチドを含む、所望の鋳型核酸の選択方法におけるオリゴヌクレオチドに関する。第1の観点において、オリゴヌクレオチドは、ハイブリッド捕捉方法などの選択方法において有用である。第2の観点において、オリゴヌクレオチドは、所望の鋳型核酸として、鋳型の集団から選択される少なくとも1つのメンバーを含む。第3の観点において、オリゴヌクレオチドは、所望の鋳型の少なくとも1つの配列に対して実質的に相補的である。第4の観点において、オリゴヌクレオチドは、ブロッカーまたはベイトから選択される少なくとも1つのメンバーを含む。第5の観点において、オリゴヌクレオチドは、少なくとも1個の T_m 増強基として、ロックド核酸基、二環式核酸基、C5修飾ピリミジン、ペプチド核酸基、およびこれらの組合せからなる群から選択される少なくとも1つのメンバーを含む。第6の観点において、オリゴヌクレオチドは、少なくとも1個の T_m 増強基として、ロックド核酸基、二環式核酸基、またはこれらの組合せの1つを含む。第7の観点において、オリゴヌクレオチドは、少なくとも1個の T_m 増強基として、ロックド核酸基または二環式核酸基を含む。第7の観点の好ましい実施形態として、オリゴヌクレオチドは、ロックド核酸基または二環式核酸基として、シトシンおよびアデニンの混合物ならびにシトシンおよ

10

20

30

40

50

びチミンの混合物を含めて、シトシン、アデニンおよびチミンからなる群から選択される核酸塩基を有する。第9の観点において、オリゴヌクレオチドは、少なくとも1個の T_m 増強基として、約1.4°Cから約25°Cを含む範囲の最適な増強 T_m 値を提供するものを含む。第10の観点において、オリゴヌクレオチドは、配列番号2、3、4、5、6、7、8、10、11、12、13、14、15、16、18、19、21、22、24、25、27、28、30、32、34および36からなる群から選択される少なくとも1つのメンバーを含む。第11の観点において、オリゴヌクレオチドは、ブロッカーを含む。これらの観点の一つの好ましい実施形態において、ブロッカーは、所望の鋳型核酸の末端における少なくとも1つの配列に対する実質的な配列相補性を有する。この好ましい実施形態のさらなる精巧化において、ブロッカーは、複数のヌクレオチドを有するバーコード（またはインデックス）ドメインを含む。この観点のさらなる実施形態において、複数のヌクレオチドとしては、実質的に連続して位置する約5個から約12個のヌクレオチドが挙げられる。別の実施形態において、バーコードドメインは、核酸塩基として、アデニン、チミン、シトシン、グアニン、イノシン、3-ニトロピロール、5-ニトロインドールおよびこれらの組合せから選択される群から選択される少なくとも1つのメンバーを有するヌクレオチドを含む。第12の観点において、オリゴヌクレオチドは、所望の鋳型核酸の選択方法における改善を提供する。この観点の一つの好ましい実施形態において、改善は、所望されない鋳型核酸に対して所望の鋳型核酸の濃縮の改善からなる。なお別の実施形態において、濃縮の改善は、少なくとも65%の濃縮を含む。第13の観点において、オリゴヌクレオチドは、3'末端修飾をさらに含む。この観点において、3'末端修飾の好ましい実施形態は、オリゴヌクレオチドからのポリメラーゼ指向性合成を防止する。別の観点において、3'末端修飾としては、とりわけ2'，3'-ジデオキシヌクレオチド、3'-スパーサーC3基が挙げられる。

【0017】

第2の態様において、本発明は、鋳型核酸の集団から所望の鋳型核酸を選択する方法に関する。方法は、2つのステップを含む。第1のステップは、鋳型核酸の集団と、 T_m 増強オリゴヌクレオチドを含む第1のオリゴヌクレオチドとを接触させて、混合物を形成することである。第2のステップは、混合物から所望の鋳型核酸を単離することを含む。第1の観点において、方法は、接触ステップの一部として、 T_m 増強オリゴヌクレオチドのほとんど最適な増強 T_m 値の温度で混合物をインキュベートするサブステップを提供する。この観点の第1の好ましい実施形態において、 T_m 増強オリゴヌクレオチドは、複数の T_m 増強基を含む。これに関して、複数の T_m 増強基は、約2個から約25個の T_m 増強基を含む。さらなる実施形態は、複数の T_m 増強基がロックド核酸基または二環式核酸基を含むことを提供する。これらの実施形態の好ましい態様としては、シトシン、アデニンおよびチミンからなる群から選択される核酸塩基を有するロックド核酸基または二環式核酸基の特色が挙げられる。第2の観点において、方法は、 T_m 増強オリゴヌクレオチドとして、配列番号2、3、4、5、6、7、8、10、11、12、13、14、15、16、18、19、21、22、24、25、27、28、30、32、34および36からなる群から選択される少なくとも1つのメンバーを含む。第3の観点において、方法は、ブロッカーを含む T_m 増強オリゴヌクレオチドを提供する。この観点の第1の好ましい実施形態において、ブロッカーは、鋳型核酸の集団の各メンバーの末端における少なくとも1つの配列に対する実質的な配列相補性を有する。なお別の好ましい実施形態において、ブロッカーは、複数のヌクレオチドを有するバーコードドメインをさらに含む。一部の実施形態において、複数のヌクレオチドとしては、実質的に連続して位置する約5個から約12個のヌクレオチドが挙げられる。他の実施形態において、バーコードドメインとしては、核酸塩基としてアデニン、チミン、シトシン、グアニン、またはユニバーサル塩基、例えばイノシン、3-ニトロピロール、5-ニトロインドール、およびこれらの組合せから選択される群から選択される少なくとも1つのメンバーを有するヌクレオチドが挙げられる。第3の観点において、この方法は、接触ステップとして、所望の鋳型核酸と所望されない鋳型核酸との間の複合体形成の実質的阻害をもたらすという目標を有する。第4の

観点において、方法は、所望の鋳型核酸を単離するステップとして、2つの追加のステップを含む。第1のステップは、所望の核酸と第2のオリゴヌクレオチドとの間にハイブリッド複合体を形成することである。第2のステップは、混合物からハイブリッド複合体を分離することである。この第4の観点に関して、第2のオリゴヌクレオチドは、ベイトを含む。特定の実施形態において、ベイトは、所望の鋳型核酸内の配列に対する実質的な配列相補性を有する配列を含む。他の実施形態において、ベイトは、複数の T_m 増強基を含む。なお他の実施形態において、ベイトは、ハイブリッド複合体の選択をできるようにするための共有結合修飾を含む。これらの後者実施形態の一部として、共有結合修飾は、ビオチン化された基である。なお他の実施形態は、アビジンまたはストレプトアビジンで固定化された固体支持体と接触されているハイブリッド複合体を提供する。

10

【0018】

第3の態様において、本発明は、大規模平行配列決定を行う方法に関する。方法は、4つのステップを含む。第1のステップは、鋳型核酸のライブラリー集団を調製することである。第2のステップは、鋳型核酸のライブラリー集団と、ブロッカーとしての少なくとも1つの T_m 増強オリゴヌクレオチド、ベイトとしての複数のオリゴヌクレオチド、および $C_{\theta-1}$ DNAとを接触させて、混合物を形成することである。第3のステップは、混合物から複数の所望の鋳型核酸を単離することである。第4のステップは、複数の所望の鋳型核酸を配列決定することである。ベイトとしての複数のオリゴヌクレオチドの少なくとも1つのメンバーは、複数の所望の鋳型核酸の少なくとも1つのメンバー内の配列に対する実質的な相補性を有する。第1の観点において、方法は、鋳型核酸のライブラリー集団のメンバーを含み、各々は、約15個のヌクレオチドから約75個のヌクレオチドのサイズ範囲を有する少なくとも1つの同一の末端アダプター配列を含む。第2の観点において、方法は、鋳型核酸のライブラリー集団の少なくとも1つの同一の末端アダプター配列に対する実質的な配列相補性を有するブロッカーを含む。第3の観点において、方法は、少なくとも1つの同一の末端アダプター配列として、バーコードドメインを含む。第4の観点において、方法は、少なくとも1つの同一の末端アダプター配列に対する実質的な配列相補性を有するブロッカーを提供する。第5の観点において、方法は、接触ステップとして、少なくとも1つの T_m 増強オリゴヌクレオチドのほとんど最適な増強 T_m 値の温度で混合物をインキュベートするステップを含む。第6の観点において、方法は、少なくとも1つの T_m 増強オリゴヌクレオチドが、ブロッカーとして、配列番号2、3、4、5、6、7、8、10、11、12、13、14、15、16、18、19、21、22、24、25、27、28、30、32、34および36からなる群から選択される少なくとも1つのメンバーを含むことを提供する。第7の観点において、方法は、混合物から複数の所望の鋳型核酸を単離するステップが、2つのステップを含むことを提供する。第1のステップは、複数の所望の鋳型核酸と、ベイトとしての複数のオリゴヌクレオチドとの間に、複数のハイブリッド複合体を形成することである。第2のステップは、混合物から複数のハイブリッド複合体を分離することである。第8の観点において、方法は、ベイトとしての複数のオリゴヌクレオチドが、複数の T_m 増強基を含むことを提供する。この観点の実施形態において、各ベイトは、ベイトを含むハイブリッド複合体の選択を可能にするための共有結合修飾を含む。この観点のさらなる実施形態において、共有結合修飾は、ビオチン化された基である。この観点の別の実施形態として、複数のハイブリッド複合体は、アビジンまたはストレプトアビジンで固定化されている固体支持体と接触される。

20

30

40

【0019】

別の態様において、本発明は、核酸を選択するまたはハイブリダイゼーション反応におけるオフターゲット核酸選択を低減する方法を特色とする。ハイブリダイゼーション反応は、固相または溶液相ハイブリダイゼーションであり得る。この方法は、後続のプロセッシング、例えば配列決定のためのライブラリーメンバーの選択において使用され得る。

【0020】

この方法は、

(a) 場合によって、複数の標的メンバー、例えば標的核酸（例えば、DNAまたはR

50

NA) メンバーを含むライブラリーを獲得し、ここで、1つ以上の標的メンバーは、挿入配列（例えば、目的の遺伝子のセグメント）および非標的核酸配列（例えば、アダプター配列）を含む、こと；ならびに

(b) ライブラリーと、捕捉プローブ、例えばベイトセットまたは複数のベイトセット、およびブロッキングオリゴヌクレオチドとを接触させることを含む、

(i) ブロッキングオリゴヌクレオチドは、ライブラリーメンバーの非標的核酸配列（例えば、アダプター配列）に相補的である、またはライブラリーメンバーの非標的核酸配列（例えば、アダプター配列）で二重鎖を形成し、

(ii) ブロッキングオリゴヌクレオチドとライブラリーメンバーの非標的核酸配列との間の結合相互作用に関連したパラメータについての値は、バックグラウンド核酸に対する非標的核酸配列、例えば他の相補的非標的核酸配列についての値より高く、これによってオフターゲット選択を最小化する。

【0021】

一実施形態において、方法は、選択ライブラリーメンバー（本明細書において「ライブラリーキャッチ」とも称される）を提供することをさらに含む。

【0022】

一実施形態において、方法は、捕捉プローブから選択ライブラリーメンバーを分離することをさらに含む。

【0023】

一実施形態において、方法は、選択ライブラリーメンバーの挿入物を配列決定すること、例えば、少なくとも2個、5個、10個、15個、20個、30個または50個の遺伝子または核酸改変、例えば、本明細書に記載されている遺伝子または核酸改変からの挿入物を配列決定することをさらに含む。

【0024】

一実施形態において、結合相互作用に関連したパラメータについての値は、親和性、会合率、解離率の反比例、または核酸溶融温度（例えば、 T_m 、DNA鎖の半分が二重ヘリックス状態であり半分がランダムコイル状態である温度）についての値であってよい。

【0025】

一実施形態において、方法は、第1の非標的核酸配列、例えば第1のアダプター配列と二重鎖を形成する第1のブロッキングオリゴヌクレオチド、および場合によって、第2の非標的核酸配列、例えば第2のアダプター配列と二重鎖を形成する第2のブロッキングオリゴヌクレオチドの使用を含む。オリゴヌクレオチドブロッカーのセットは、複数の異なるオリゴヌクレオチドブロッカーを含む。

【0026】

一実施形態において、ブロッキングオリゴヌクレオチドは、反応における配列と、捕捉プローブに対して二重鎖化されるライブラリーメンバーの非標的配列との間の二重鎖の形成を阻害する（例えば、ブロッキングオリゴヌクレオチドは、ライブラリーメンバーの連鎖鎖の形成を阻害する。）。

【0027】

一実施形態において、ライブラリーメンバーは、挿入、例えばサブゲノム間隔、および非標的配列、例えば複数のライブラリーメンバーに共通の配列を含む。一実施形態において、挿入は、例えば腫瘍試料からの核酸からのサブゲノム配列であり、非標的配列は、非ゲノムの発生配列またはサブゲノム配列中に存在しない配列、例えば増幅タグまたはバーコーディングタグである。

【0028】

一実施形態において、ライブラリーメンバーまたは選択ライブラリーメンバーは、少なくとも2個、5個、10個、15個、20個、30個または50個の遺伝子または核酸改変、例えば本明細書に記載されている遺伝子または核酸改変からのサブゲノム間隔を含む。

【0029】

10

20

30

40

50

一実施形態において、複数のライブラリーメンバー、または選択ライブラリーメンバー、例えば、X個の（ここで、Xは、2、5、10、20、50、100、200、またはこれ以上に等しい。）ライブラリーメンバー、または選択ライブラリーメンバーは、挿入の5'末端に第1の非標的配列を有し、挿入の3'末端に第2の非標的配列を有する。

【0030】

一実施形態において、非標的配列としては、複数の非標的配列中に存在する非標的配列、例えば増幅のための配列、および独特である非標的配列、例えばバーコードが挙げられる。ライブラリーのメンバーの典型的に一部、最も実質的に全て、または全ては、共通の非標的配列を含む。実施形態において、ライブラリーまたは選択ライブラリーメンバーは、共通の非標的配列を有する少なくともX個のメンバー（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）を含む。

10

【0031】

一実施形態において、ブロッキングオリゴヌクレオチドは、少なくともX個のライブラリーメンバー（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）の非標的核酸配列と二重鎖を形成し、この二重鎖は、バックグラウンド核酸に対する非標的核酸配列、例えば非標的配列の相補体によって形成される二重鎖の T_m より高い T_m を有する。一実施形態において、修飾ブロッキングオリゴヌクレオチド二重鎖のより高い核酸溶解温度は、非修飾ブロッキングオリゴヌクレオチド二重鎖の核酸溶解温度より約5℃から約25℃、またはこれ以上（例えば、2℃、5℃、10℃、15℃、20℃、25℃、またはこれ以上）高い。一実施形態において、ブロッキングオリゴヌクレオチドとライブラリーメンバーの非標的核酸配列との間の二重鎖についての T_m は、非標的核酸配列およびこの正確な相補体の二重鎖についての T_m より高い。

20

【0032】

他の実施形態において、ブロッキングオリゴヌクレオチドは、バックグラウンド核酸に対する非標的核酸配列、例えば非標的配列の相補体の会合率より高い、少なくともX個のライブラリーメンバー（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）の非標的核酸配列に対する会合率を有する。一実施形態において、より高い会合率は、バックグラウンド核酸に対する非標的核酸配列の会合率の約2倍から10倍超（例えば、2倍、4倍、6倍、8倍、10倍、またはこれ以上）である。

30

【0033】

なお他の実施形態において、ブロッキングオリゴヌクレオチドは、バックグラウンド核酸に対する非標的核酸配列、例えば非標的配列の相補体の解離率より低い、少なくともX個のライブラリーメンバー（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）の非標的核酸配列に対する解離率を有する。一実施形態において、より低い解離率は、バックグラウンド核酸に対する非標的核酸配列の解離率の約2倍から10倍超（例えば、2倍、4倍、6倍、8倍、10倍、またはこれ以上）である。

【0034】

一実施形態において、ブロッキングオリゴヌクレオチドの長さは、バックグラウンド核酸に対して、ライブラリーメンバー（例えば、アダプター配列）の非標的核酸配列のためのブロッキングオリゴヌクレオチドの結合相互作用における増加をもたらす。

40

【0035】

一実施形態において、ブロッキングオリゴヌクレオチドと少なくともX個のライブラリーメンバー（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）の非標的核酸配列との間に形成される二重鎖は、非標的配列とこの相補体との間、例えば二本鎖化アダプターのワトソン鎖とクリック鎖との間に形成される二重鎖より長い。実施形態において、ブロッキングオリゴヌクレオチドと非標的核酸配列との間の二重鎖は、非標的配列とこの相補体との間、例えば二本鎖化アダプターのワトソン鎖とクリック鎖との間に形成される二重鎖より長い、少なくとも1個、2個、3個、4個、5個、6個、7個、8個、9個、10個、15個または20個のヌクレオチドである。

50

【0036】

一実施形態において、ブロッキングオリゴは、1個以上の非天然ヌクレオチドを含む。実施形態において、非天然ヌクレオチドを有するブロッキングオリゴヌクレオチドと、少なくともX個のライブラリーメンバー（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）の非標的核酸配列との間に形成される二重鎖は、バックグラウンド核酸に対する非標的核酸配列、例えば他の相補的非標的核酸配列についての値より高い、結合相互作用に関連したパラメータ（例えば、親和性、会合率、解離率の反比例、または T_m ）についての値を有する。例示的非天然オリゴヌクレオチドとしては、修飾DNAまたはRNAヌクレオチドが挙げられる。例示的修飾ヌクレオチド（例えば、修飾RNAまたはDNAヌクレオチド）としては、これらに限定されないが、LNAヌクレオチドのリボース部分が2' 酸素および4' 炭素を接続する余分な架橋で修飾されているロックド核酸（LNA）；ペプチド核酸（PNA）、例えば、ペプチド結合によって連結されている反復N-（2-アミノエチル）-グリシン単位に構成されているPNA；低GC領域を捕捉するように修飾されているDNAまたはRNAオリゴヌクレオチド；二環式核酸（BNA）；架橋オリゴヌクレオチド；修飾5-メチルデオキシシチジン；および2, 6-ジアミノプリンが挙げられる。他の修飾DNAおよびRNAヌクレオチドは、当技術分野において知られている。

10

【0037】

一実施形態において、ブロッキングオリゴヌクレオチドは、RNAであるまたはRNAを含み、非標的核酸配列、例えばアダプターは、DNAであるまたはDNAを含む。一実施形態において、非標的核酸配列は、複数のライブラリーメンバー、例えば少なくともX個のライブラリーメンバー（ここで、Xは、2、5、10、20、50、100、または200に等しい。）に共通の配列であり、例えば増幅のため、例えばPCR、架橋PCR、増幅のために使用され得る配列である。

20

【0038】

一実施形態において、非標的核酸配列は、増幅のため、例えばPCR、架橋PCR、増幅のために使用され得る配列であり、バックグラウンド核酸は、第2の非標的配列である。

【0039】

一実施形態において、捕捉プローブは、DNAである（例えばRNAに対立するものとして）。実施形態において、捕捉プローブとしては、1種以上のDNAオリゴヌクレオチド（例えば、天然または非天然のDNAオリゴヌクレオチド）が挙げられる。

30

【0040】

一実施形態において、捕捉プローブはRNAである。実施形態において、捕捉プローブとしては、1種以上のRNAオリゴヌクレオチド（例えば、天然または非天然のRNAオリゴヌクレオチド）が挙げられる。

【0041】

一実施形態において、ブロッキングオリゴヌクレオチドは、20-80、30-80、40-80、50-80、70-80、30-75、30-65、30-55、30-45、40-70、40-60、40-50、50-60、50-70、60-70ヌクレオチド長である。一実施形態において、ライブラリー挿入は、本明細書において他所で記載されている通り、50-10、000、50-1,000、50-500、50-200、77-150または100-150ヌクレオチド長である。

40

【0042】

別の態様において、本発明は、例えば本明細書において記載されている通りの複数のブロッキングオリゴヌクレオチドを含む調製物を特色とする。一実施形態において、調製物は：例えば本明細書において記載されている通りの複数のライブラリーメンバー；および例えば本明細書において記載されている通りの捕捉プローブの一方または両方をさらに含む。

【0043】

別の態様において、本発明は、例えば本明細書において記載されている通りの複数のブ

50

ロッキングオリゴヌクレオチドを含むキットを特色とする。一実施形態において、キットは：例えば本明細書において記載されている通りの複数のライブラリーメンバー；および例えば本明細書において記載されている通りの捕捉プローブの一方または両方をさらに含む。実施形態において、構成成分は、別々の容器中に提供され、例えば、ブロックングオリゴヌクレオチドは容器中に提供され、別の構成成分、例えば緩衝剤、または例えば本明細書において記載されている通りの複数のライブラリーメンバー、または例えば本明細書において記載されている通りの捕捉プローブは、異なる容器（単数または複数）中で提供される。

【0044】

別の態様において、本発明は、本明細書に記載されている別の方法、例えば、本明細書に記載されている配列決定方法、本明細書に記載されているアラインメント方法、本明細書に記載されている変異呼び出し方法、または本明細書に記載されているベイトを使用する方法と組み合わされている、本明細書に記載されているオフターゲット核酸選択を低減する方法を特色とする。

【0045】

オフターゲット選択は、非標的配列の二重鎖がライブラリーメンバーの挿入配列および捕捉プローブの二重鎖より安定でないように、十分に短い非標的配列の使用によって最小化することもできる。したがって、別の態様において、本発明は、例えば固相または溶液ハイブリダイゼーションにおけるオフターゲット核酸選択を低減する方法を特色とする。この方法は、後続の配列決定のためのライブラリーメンバーの選択において使用され得る。

【0046】

この方法は、

(a) 場合によって、複数の標的メンバー、例えば標的核酸（例えば、DNAまたはRNA）メンバーを含むライブラリーを獲得し、ここで、標的メンバーの1つ以上は、挿入配列（例えば、目的の遺伝子のセグメント）および非標的核酸配列（例えば、アダプター配列）を含む、こと；ならびに

(b) ライブラリーと、捕捉プローブ、例えばベイトセットまたは複数のベイトセットとを接触させること
を含み、

非標的配列は十分に短いので、挿入配列と捕捉プローブとの間の結合相互作用に関連したパラメータについての値は、非標的核酸配列およびこの相補体についてのその値より高く、これによってオフターゲット選択を最小化する。

【0047】

一実施形態において、この方法は、選択ライブラリーメンバー（本明細書において「ライブラリーキャッチ」とも称される）を提供することをさらに含む。

【0048】

一実施形態において、方法は、捕捉プローブから選択ライブラリーメンバーを分離することをさらに含む。

【0049】

一実施形態において、方法は、選択ライブラリーメンバーの挿入を配列決定すること、例えば、少なくとも2個、5個、10個、15個、20個、30個または50個の遺伝子または核酸改変、例えば本明細書に記載されている遺伝子または核酸改変からの挿入を配列決定することをさらに含む。

【0050】

一実施形態において、結合相互作用に関連したパラメータについての値は、親和性、会合率、解離率の反比例、または核酸溶融温度（例えば、 T_m 、DNA鎖の半分が二重ヘリックス状態であり、半分がランダムコイル状態である温度）についての値であり得る。

【0051】

一実施形態において、ライブラリーメンバーは、挿入、例えばサブゲノム間隔、および非標的配列、例えば複数のライブラリーメンバーに共通の配列を含む。一実施形態におい

10

20

30

40

50

て、挿入は、例えば腫瘍試料からの核酸からのサブゲノム配列であり、非標的配列は、非天然配列またはサブゲノム配列中に存在しない配列、例えば増幅タグまたはバーコーディングタグである。

【0052】

一実施形態において、ライブラリーメンバーまたは選択ライブラリーメンバーは、少なくとも2個、5個、10個、15個、20個、30個または50個の遺伝子または核酸改変、例えば本明細書に記載されている遺伝子または核酸改変からのサブゲノム間隔を含む。

【0053】

一実施形態において、複数のライブラリーメンバーまたは選択ライブラリーメンバー、例えば、X個の（ここで、Xは、2、5、10、20、50、100、200、またはこれ以上に等しい。）ライブラリーメンバーまたは選択ライブラリーメンバーは、挿入の5'末端に第1の非標的配列を有し、挿入の3'末端に第2の非標的配列を有する。

【0054】

一実施形態において、非標的配列としては、複数の非標的配列中に存在する非標的配列、例えば増幅のための配列、および独特である非標的配列、例えばバーコードが挙げられる。ライブラリーのメンバーの典型的に一部、最も実質的に全て、または全ては、共通の非標的配列を含む。実施形態において、ライブラリーまたは選択ライブラリーメンバーは、共通の非標的配列を有する少なくともX個のメンバー、（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）を含む。

【0055】

一実施形態において、挿入配列は、少なくともX個のライブラリーメンバー（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）のための捕捉プローブと二重鎖を形成し、この二重鎖は、バックグラウンド核酸に対する非標的核酸配列、例えば非標的配列の相補体によって形成される二重鎖の T_m より高い T_m を有する。一実施形態において、挿入配列／捕捉プローブ二重鎖のより高い核酸溶融温度は、約5℃から25℃、またはこれ以上（例えば、5℃、10℃、15℃、20℃、25℃、またはこれ以上）である。一実施形態において、挿入配列／捕捉プローブ間の二重鎖についての T_m は、非標的核酸配列およびこの正確な相補体の二重鎖についての T_m より高い。

【0056】

他の実施形態において、挿入配列は、バックグラウンド核酸に対する非標的核酸配列、例えば非標的配列の相補体の会合率より高い、少なくともX個のライブラリーメンバー（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）のためのプローブに対する会合率を有する。一実施形態において、より高い会合率は、バックグラウンド核酸に対する非標的核酸配列の会合率の約2倍から10倍超（例えば、2倍、4倍、6倍、8倍、10倍、またはこれ以上）である。

【0057】

なお他の実施形態において、挿入配列は、バックグラウンド核酸に対する非標的核酸配列、例えば非標的配列の相補体の解離率より低い、少なくともX個のライブラリーメンバー（ここで、Xは、1、2、5、10、20、50、100、200、またはこれ以上に等しい。）のための捕捉プローブに対する解離率を有する。一実施形態において、より低い解離率は、バックグラウンド核酸に対する非標的核酸配列の解離率の約2倍から10倍超（例えば、2倍、4倍、6倍、8倍、10倍、またはこれ以上）である。

【0058】

一実施形態において、非標的核酸配列は、増幅のため、例えばPCR、架橋PCR、増幅のために使用され得る配列であり、バックグラウンド核酸は、第2の非標的配列である。

【0059】

一実施形態において、捕捉プローブは、DNA（例えば、RNAに対立するものとして）である。一実施形態において、捕捉プローブはRNAである。

【0060】

一実施形態において、ライブラリー挿入は、本明細書において他所で記載されている通り、50-10, 000, 50-1, 000, 50-500, 50-200, 77-150または100-150ヌクレオチド長である。

【0061】

一実施形態において、方法は、本明細書において記載されている通り、ブロッキングオリゴヌクレオチドの使用をさらに含む。

【0062】

本発明の追加の特色および実施形態は、本明細書に記載されている。

【0063】

一実施形態において、この方法は、

(c) 例えば配列決定によって、例えば次世代配列決定方法を用いて、前記ライブラリーまたはライブラリーキャッチからの腫瘍メンバーからのサブゲノム間隔のための読み取りを獲得すること；

(d) 前記読み取りをアラインメントすること；および

(e) 予備選択ヌクレオチド位置、例えば、複数のサブゲノム間隔の各々、例えば複数の遺伝子の各々における予備選択ヌクレオチド位置に関する前記読み取りからのヌクレオチド値を割り当て（例えば、ベイズ方法を用いて、変異を呼び出す）、これによって前記試料を分析すること
をさらに含む。

【0064】

一実施形態において、

(i) Xヌクレオチド位置の各々は、ステップ(b)、(c)、(d)もしくは(e)の1つまたはこれらの組合せのための独特のセットの条件下で分析される（ここで、独特とは、他のX-1セットの条件と異なることを意味し、Xは、少なくとも2、5、10、20、30、40、50、100、200、300または500である。）。例えば、第1セットの条件、例えば本明細書に記載されている条件のセットは、例えば第1のサブゲノム間隔または遺伝子における第1のヌクレオチド位置のために使用され、第2セットの条件、例えば本明細書に記載されている条件の第2セットは、例えば第2のサブゲノム間隔または遺伝子における第2のヌクレオチド位置のために使用される；

(ii) X個のヌクレオチド位置の各々について、ヌクレオチド位置で発生し得る予備選択改変、例えば変異の特徴、例えば本明細書に記載されている特徴に応答して、ヌクレオチド位置は、独特のセットの条件下で分析される（ここで、独特とは、他のX-1セットの条件と異なることを意味し、Xは、少なくとも2、5、10、20、30、40、50、100、200、300または500である。）。例えば、第1のサブゲノム間隔におけるヌクレオチド位置で発生し得る予備選択改変、例えば変異の特徴、例えば本明細書に記載されている特徴に応答して、ヌクレオチド位置は第1セットの条件下で分析され、第2のサブゲノム間隔におけるヌクレオチド位置で発生し得る予備選択改変、例えば変異の特徴、例えば本明細書に記載されている特徴に応答して、ヌクレオチド位置は第2セットの条件下で分析される；

(iii) ここで、前記方法は、試料、例えば保存腫瘍試料上にて、少なくとも2、5、10、20、50または100個のサブゲノム間隔、例えば遺伝子におけるヌクレオチド位置に対する95%、98%もしくは99%の感受性または特異性を可能にする条件下で行われる；または

(iv) ここで、方法は、

a) 第1のサブゲノム間隔を配列決定して、約500×またはこれより高い配列決定深度を提供すること、例えば、試料からの細胞の5%以下に存在する変異を配列決定すること；

b) 第2のサブゲノム間隔を配列決定して、約200×以上、例えば約200×-約500×の配列決定深度を提供するステップこと、例えば、試料からの細胞の10%以下に存在する変異を配列決定すること；

10

20

30

40

50

c) 第3のサブゲノム間隔を配列決定して、約10-100×の配列決定深度を提供すること、例えば、以下から選択される1つ以上のサブゲノム間隔（例えば、エクソン）を配列決定すること：a) 異なる薬物を代謝するための患者の能力を説明し得る薬理ゲノミクスの（PGx）単一ヌクレオチド多型（SNP）またはb) 患者を独特に同定する（例えば、指紋をとる。）ために使用され得るゲノムSNP；

d) 第4のサブゲノム間隔を配列決定して、約5-50×の配列決定深度を提供すること、例えば、ゲノム転座またはインデルなどの構造ブレイクポイントを検出することの1つ以上または全てを含む。例えば、イントロンブレイクポイントの検出は、高い検出信頼性を確保するための5-50×の配列対スパンニング深度を必要とする。こうしたバイトセットは、例えば転座／インデル傾向の癌遺伝子を検出するために使用され得る；または

e) 第5のサブゲノム間隔を配列決定して、約0.1-300×の配列決定深度を提供すること、例えばコピー数変化を検出すること。一実施形態において、配列決定深度は、コピー数変化を検出するために約0.1-10×の配列決定深度を範囲とする。他の実施形態において、配列決定深度は、ゲノムDNAのコピー数獲得／消失、またはヘテロ接合性の消失（LOH）を判定するために使用されるゲノムSNP／座位を検出するために約100-300×を範囲とする。

【0065】

例示的第1および第2のセットの条件は、

第1のバイトセットは、第1のサブゲノム間隔のために使用され、第2のバイトセットは、第2のサブゲノム間隔のために使用されること；

第1のアラインメント方法は、第1のサブゲノム間隔のための読み取りに適用され、第2のアラインメント方法は、第2のサブゲノム間隔のための読み取りに適用されること；

第1の変異呼び出し方法は、第1のサブゲノム間隔のヌクレオチド位置に適用され、第2の変異呼び出し方法は、第2のサブゲノム間隔のヌクレオチド位置に適用されることを含む。

【0066】

一実施形態において、

第1のヌクレオチド位置は、第1のセットのバイト条件、第1のアラインメント方法、および第1の変異呼び出し方法を用いて分析され；

第2のヌクレオチド位置は、前記第1のセットのバイト条件、第2のアラインメント方法、および前記第1の変異呼び出し方法を用いて分析され；

第3のヌクレオチド位置は、前記第1のセットのバイト条件、前記第1のアラインメント方法、および第2の変異呼び出し方法を用いて分析されて、各々が他の2つと比較して独特の条件下で分析される3つのヌクレオチド位置を提供する。

【0067】

一実施形態において、条件は、

第1のバイトセットは、第1のサブゲノム間隔のために使用され、第2のバイトセットは、第2のサブゲノム間隔のために使用されること；

第1のアラインメント方法は、第1のサブゲノム間隔のための読み取りに適用され、第2のアラインメント方法は、第2のサブゲノム間隔のための読み取りに適用されること；または

第1の変異呼び出し方法は、第1のサブゲノム間隔のヌクレオチド位置に適用され、第2の変異呼び出し方法は、第2のサブゲノム間隔のヌクレオチド位置に適用されることを含む。

【0068】

例示的特徴としては、

(i) 改変が位置されている遺伝子または遺伝子の型、例えばオンコジーンもしくは腫瘍抑制剤、予備選択バリエーションもしくはバリエーションの型、例えば変異体によってもしくは予備選択頻度の変異体によって特徴づけられる遺伝子もしくは遺伝子の型、または本明細書に記載されている他の遺伝子もしくは遺伝子の型；

(i i) 改変の型、例えば置換、挿入、欠失または転座；

(i i i) 試料の型、例えば改変のために分析される F F P E 試料；

(i v) 評価される改変の前記ヌクレオチド位置またはこの付近における配列、例えば、サブゲノム間隔のためのミスアラインメント、例えばヌクレオチド位置またはこの付近における反復配列の存在の予想傾向に影響し得る配列；

(v) 例えば予備選択型の腫瘍における改変、例えば変異を示す読み取りを観察する事前（例えば、文献）予想；

(v i) 単独での塩基呼び出しエラーによる改変を示す読み取りを観察する確率；または

(v i i) 改変を検出するのに所望される配列決定の予備選択深度が挙げられる。

10

【 0 0 6 9 】

一実施形態において、特徴は、配列決定されるヌクレオチドの同一性以外であり、すなわち、特徴は、配列が a または t であるかどうかではない。

【 0 0 7 0 】

一実施形態において、少なくとも X 個の遺伝子、例えば表 1 および 1 A からの少なくとも X 個の遺伝子、例えば表 1 および 1 A において優先 1 のアノテーションを有する遺伝子からのサブゲノム間隔は、異なる条件下で分析され、X は、2、3、4、5、10、15、20 または 30 に等しい。

【 0 0 7 1 】

一実施形態において、方法は、以下の 1 つ以上を含む：

20

(i) 方法、例えば上記の方法の (b) は、本明細書に記載されているバイトセットの使用を含む；

(i i) 方法、例えば上記の方法の (c) は、サブゲノム間隔のセットまたは群のための読み取り、または本明細書に記載されている遺伝子のセットまたは群からの読み取りを獲得するステップを含む；

(i i i) 方法、例えば上記の方法の (d) は、本明細書に記載されている複数のアラインメント方法の使用を含む；

(i v) 方法、例えば上記の方法の (e) は、本明細書に記載されている予備選択ヌクレオチド位置にヌクレオチド値を割り当てるための複数の方法の使用を含む；または

(v) 方法は、本明細書に記載されているサブゲノム間隔のセットにヌクレオチド値を割り当てることを含む。

30

【 0 0 7 2 】

一実施形態において、方法は： (i) および (i i) - (v) の 1 つ、2 つ、3 つまたは全てを含む。一実施形態において、方法は： (i i) ならびに (i) および (i i i) - (v) の 1 つ、2 つ、3 つまたは全てを含む。一実施形態において、方法は： (i i i) ならびに (i) 、 (i i) 、 (i v) および (v) の 1 つ、2 つ、3 つまたは全てを含む。一実施形態において、方法は： (i v) ならびに (i) - (i i i) および (v) の 1 つ、2 つ、3 つまたは全てを含む。一実施形態において、方法は： (v) ならびに (i) - (i v) の 1 つ、2 つ、3 つまたは全てを含む。

バイト

40

本明細書に記載されている方法は、配列決定される標的核酸の選択のために、バイト、例えば溶液ハイブリダイゼーションにおける使用のためのバイトの適切な選択によって、1 つ以上の対象からの試料、例えば腫瘍試料からの多数の遺伝子および遺伝子産物の選択および／または配列決定を提供する。様々なサブゲノム間隔またはこれらのクラスのための選択の効率は、選択の予備選択効率を有するバイトセットに従って整合される。この項において使用される場合、「選択の効率」は、それが標的サブゲノム間隔（単数または複数）に従って調整される場合の配列カバレッジのレベルまたは深度を指す。

【 0 0 7 3 】

したがって、方法（例えば、上記に列挙されている方法の要素 (b) ）は、ライブラリーと複数のバイトとを接触させて、選択メンバー（例えば、ライブラリーキャッチ）を提

50

供することを含む。特定の実施形態において、方法は、ライブラリーとベイトまたはベイトセットの複数、例えば少なくとも2つ、3つ、4つまたは5つとを接触させることを含み、ここで、前記複数の各ベイトまたはベイトセットは、選択（複数の他のベイトセットに対立するものとして。）の独特の予備選択効率を有する。例えば、各独特のベイトまたはベイトセットは、配列決定の独特の深度を提供する。「ベイトセット」という用語は、本明細書で使用される場合、1つのベイトまたは複数のベイト分子を集合的に指す。

【0074】

一実施形態において、複数における第1のベイトセットの選択の効率は、複数における第2のベイトセットの効率と少なくとも2倍異なる。一実施形態において、第1および第2のベイトセットは、少なくとも2倍異なる配列決定の深度を提供する。

【0075】

別の実施形態において、方法は、以下のベイトセットの1つまたは複数とライブラリーとを接触させることを含む：

a) サブゲノム間隔を含む十分なメンバーを選択して、約500×以上の配列決定深度を提供する、例えば、試料からの細胞の5%以下に存在する変異を配列決定する、ベイトセット；

b) サブゲノム間隔を含む十分なメンバーを選択して、約200×以上、例えば約200×から約500×の配列決定深度を提供する、例えば、試料からの細胞の10%以下に存在する変異を配列決定する、ベイトセット；

c) サブゲノム間隔を含む十分なメンバーを選択して、約10-100×の配列決定深度を提供する、例えば、以下から選択される1つ以上のサブゲノム間隔（例えば、エクソン）を配列決定する、ベイトセット：a) 異なる薬物を代謝するための患者の能力を説明することができる薬理ゲノミクスの（PGx）単一ヌクレオチド多型（SNP）、またはb) 患者を独特に同定する（例えば、指紋をとる。）ために使用され得るゲノムSNP；

d) サブゲノム間隔を含む十分なメンバーを選択して、約5-50×の配列決定深度を提供する、例えば、ゲノム転座またはインデルなどの構造ブレイクポイントを検出するベイトセット。例えば、イントロンブレイクポイントの検出は、高い検出信頼性を確保するために5-50×の配列対スパン深度を必要とする。こうしたベイトセットは、例えば転座／インデル傾向の癌遺伝子を検出するために使用され得る；または

e) サブゲノム間隔を含む十分なメンバーを選択して、約0.1-300×の配列決定深度を提供する、例えば、コピー数変化を検出する、ベイトセット。

【0076】

一実施形態において、配列決定深度は、コピー数変化を検出するために約0.1-10×の配列決定深度の範囲である。他の実施形態において、配列決定深度は、ゲノムDNAのコピー数獲得／消失またはヘテロ接合性の消失（LOH）を判定するために使用されるゲノムSNP／座位を検出するために約100-300×の範囲である。こうしたベイトセットは、例えば増幅／欠失傾向の癌遺伝子を検出するために使用され得る。配列決定深度のレベル（例えば、配列決定深度のX倍レベル）は、本明細書で使用される場合、二重読み取り、例えばPCR二重読み取りの検出および除去後の読み取り（例えば、独特の読み取り）のカバレッジのレベルを指す。

【0077】

一実施形態において、ベイトセットは、1つ以上の再編成を含有するサブゲノム間隔、例えばゲノム再編成を含有するイントロンを選択する。こうした実施形態において、ベイトセットは、選択効率を増加させるために反復性配列がマスクされるように設計されている。再編成が公知の接合配列を有するような実施形態において、相補的ベイトセットは、選択効率を増加させるために接合配列に設計することができる。

【0078】

実施形態において、方法は、2つ以上の異なる標的カテゴリーを捕捉するように設計されているベイトの使用を含み、各カテゴリーは、異なるベイト設計戦略を有する。実施形態において、本明細書において開示されているハイブリッド捕捉方法および組成物は、標

10

20

30

40

50

的配列（例えば、標的メンバー）の定義されたサブセットを捕捉し、標的配列の相同カバレッジを提供し、他方でこのサブセットの外側でカバレッジを最小化する。一実施形態において、標的配列は、ゲノムDNAからの全エキソーム、またはこの選択サブセットを含む。本明細書において開示されている方法および組成物は、複合標的核酸配列（例えば、核酸ライブラリー）についてのカバレッジの異なる深度およびパターンを達成するため異なるベイトセットを提供する。

【0079】

一実施形態において、この方法は、核酸ライブラリーの選択メンバー（例えば、ライブラリーキャッチ）を提供することを含む。この方法は、

複数のメンバー、例えば標的核酸メンバー（例えば、複数の腫瘍メンバー、参照メンバー、および／またはPGxメンバーを含む）を含むライブラリー（例えば、核酸ライブラリー）を提供すること；

例えば溶液ベースの反応におけるライブラリーと、複数のベイト（例えば、オリゴヌクレオチドベイト）とを接触させて、複数のベイト／メンバーハイブリッドを含むハイブリダイゼーション混合物を形成すること；

前記ハイブリダイゼーション混合物から複数のベイト／メンバーハイブリッドを、例えば前記ハイブリダイゼーション混合物と前記複数のベイト／メンバーハイブリッドの分離を可能にする結合実体とを接触させることによって分離し、

これによって、ライブラリーキャッチ（例えば、ライブラリーからの核酸分子の選択されたまたは濃縮されたサブグループ）を提供すること

を含む、複数のベイトは、以下の2つ以上を含む：

a) 低い頻度、例えば約5%以下で出現する改変（例えば、1つ以上の変異）（すなわち、試料からの細胞の5%は、これらのゲノムにおいて改変を持つ。）に対する高いレベルの感受性を可能にするために最も深いカバレッジが必要とされる高レベルの標的（例えば、遺伝子、エクソンまたは塩基などのサブゲノム間隔を含む1つ以上の腫瘍メンバー）を選択する第1のベイトセット。一実施形態において；第1のベイトセットは、約500×以上の配列決定深度を必要とする改変（例えば、点変異）を含む腫瘍メンバーを選択する（例えば、それに相補的である。）；

b) a)における高レベルの標的より高い頻度、例えば約10%の頻度で出現する改変（例えば、1つ以上の変異）（すなわち、試料からの細胞10%は、これらのゲノムにおける改変を持つ。）に対する高いレベルの感受性を可能にするために高いカバレッジが必要とされる中レベルの標的（例えば、遺伝子、エクソンまたは塩基などのサブゲノム間隔を含む1つ以上の腫瘍メンバー）を選択する第2のベイトセット。一実施形態において；第2のベイトセットは、約200×以上の配列決定深度を必要とする改変（例えば、点変異）を含む腫瘍メンバーを選択する（例えば、それに相補的である。）；

c) 高いレベルの感受性を可能にするために、例えばヘテロ接合性対立遺伝子を検出するために低中程度のカバレッジが必要とされる低レベルの標的（例えば、遺伝子、エクソンまたは塩基などのサブゲノム間隔を含む1つ以上のPGxメンバー）を選択する第3のベイトセット。例えば、ヘテロ接合性対立遺伝子の検出は、高い検出信頼性を確保するために10-100×の配列決定深度を必要とする。一実施形態において、第3のベイトセットは、以下から選択される1つ以上のサブゲノム間隔（例えば、エクソン）を選択する：a) 異なる薬物を代謝するための患者の能力を説明することができる薬理ゲノミクスの（PGx）単一ヌクレオチド多型（SNP）、またはb) 患者を独特に同定する（例えば、指紋をとる。）ために使用され得るゲノムSNP；

d) 例えばゲノム転座またはインデルなどの構造ブレイクポイントを検出するために低中程度のカバレッジが必要とされる第1のイントロン標的（例えば、イントロン配列を含むメンバー）を選択する第4のベイトセット。例えば、イントロンブレイクポイントの検出は、高い検出信頼性を確保するために5-50×の配列対スパン深度を必要とする。前記第4のベイトセットは、例えば転座／インデル傾向の癌遺伝子を検出するために使用され得る；または

10

20

30

40

50

e) コピー数変化を検出するための能力を改善するために低密度のカバレッジが必要とされる第2のイントロン標的(例えば、イントロンメンバー)を選択する第5のバイトセット。例えば、いくつかの末端のエクソンの1コピー欠失の検出は、高い検出信頼性を確保するために0.1-300×のカバレッジを必要とする。一実施形態において、カバレッジ深度は、コピー数変化を検出するために約0.1-10×からを範囲とする。他の実施形態において、カバレッジ深度は、ゲノムDNAのコピー数獲得/消失またはヘテロ接合性の消失(LOH)を判定するために使用されるゲノムSNP/座位を検出するため、約100-300×からを範囲とする。前記第5のバイトセットは、例えば、増幅/欠失傾向の癌遺伝子を検出するために使用され得る。

【0080】

前述のバイトセットの2つ、3つ、4つ、またはこれ以上の任意の組合せ、例えば、第1および第2のバイトセット;第1および第3のバイトセット;第1および第4のバイトセット;第1および第5のバイトセット;第2および第3のバイトセット;第2および第4のバイトセット;第2および第5のバイトセット;第3および第4のバイトセット;第3および第5のバイトセット;第4および第5のバイトセット;第1、第2および第3のバイトセット;第1、第2および第4のバイトセット;第1、第2および第5のバイトセット;第1、第2、第3、第4のバイトセット;第1、第2、第3、第4および第5のバイトセットなどの組合せが使用され得る。

【0081】

一実施形態において、第1、第2、第3、第4または第5のバイトセットの各々は、選択(例えば、捕捉)の予備選択効率を有する。一実施形態において、選択の効率についての値は、a)-e)による全て5つのバイトの少なくとも2つ、3つ、4つについて同じである。他の実施形態において、選択の効率についての値は、a)-e)による全て5つのバイトの少なくとも2つ、3つ、4つについて異なる。一部の実施形態において、少なくとも2つ、3つ、4つ、または全て5つのバイトセットは、異なる予備選択効率値を有する。

【0082】

例えば、選択の効率についての値は、以下の1つ以上から選択される:

(i) 第1の予備選択効率は、少なくとも約500×以上の配列決定深度である選択の第1の効率についての値を有し、例えば、選択の第2、第3、第4または第5の予備選択効率を超える、選択の効率についての値を有する(例えば、選択の第2の効率についての値より約2-3倍超;選択の第3の効率についての値より約5-6倍超;選択の第4の効率についての値より約10倍超;選択の第5の効率についての値より約50から5000倍超);

(ii) 第2の予備選択効率は、少なくとも約200×以上の配列決定深度である選択の第2の効率についての値を有し、例えば、選択の第3、第4または第5の予備選択効率を超える、選択の効率についての値を有する(例えば、選択の第3の効率についての値より約2倍超;選択の第4の効率についての値より約4倍超;選択の第5の効率についての値より約20から2000倍超);

(iii) 第3の予備選択効率は、少なくとも約100×以上の配列決定深度である、選択の第3の効率についての値を有し、例えば、選択の第4または第5の予備選択効率を超える、選択の効率についての値を有する(例えば、選択の第4の効率についての値より約2倍超;選択の第5の効率についての値より約10から1000倍超);

(iv) 第4の予備選択効率は、少なくとも約50×以上の配列決定深度である、選択の第4の効率についての値を有し、例えば、選択の第5の予備選択効率を超える、選択の効率についての値を有する(例えば、選択の第5の効率についての値より約50から500倍超);または

(v) 第5の予備選択効率は、少なくとも約10×から0.1×の配列決定深度である、選択の第5の効率についての値を有する。

【0083】

特定の実施形態において、選択の効率についての値は、以下の1つ以上によって修正されている：異なるバイトセットの示差表示、バイトサブセットの示差重なり、示差バイトパラメータ、異なるバイトセットの混合、および／または異なる型のバイトセットを使用すること。

【0084】

例えば、選択の効率（例えば、各バイトセット／標的カテゴリーの相対配列カバレッジ）における変動は、以下の1つ以上を改変することによって調整され得る：

(i) 異なるバイトセットの示差表示—所定の標的（例えば、標的メンバー）を捕捉するためのバイトセット設計は、より多い／より少ない数のコピーに含まれることにより、相対的な標的カバレッジ深度を増強／低減することができる；

(ii) バイトサブセットの示差重なり—所定の標的（例えば、標的メンバー）を捕捉するためのバイトセット設計に、隣接しているバイト間のより長いまたはより短い重なりが含まれることにより、相対的な標的カバレッジ深度を増強／低減することができる；

(iii) 示差バイトパラメータ—所定の標的（例えば、標的メンバー）を捕捉するためのバイトセット設計に、配列修飾／より短い長さが含まれることにより、捕捉効率を低減し、相対的な標的カバレッジ深度を低下させることができる；

(iv) 異なるバイトセットの混合—異なる標的セットを捕捉するように設計されているバイトセットは、異なるモル比で混合されることにより、相対的な標的カバレッジ深度を増強／低減することができる；

(v) 異なる型のオリゴヌクレオチドバイトセットを使用すること—特定の実施形態において、バイトセットは、以下を含むことができる：

(a) 1つ以上の、化学的に（例えば、非酵素的に）合成された（例えば、個別に合成された）バイト、

(b) アレイに合成された1つ以上のバイト、

(c) 1つ以上の、酵素的に調製された、例えばインビトロで転写されたバイト；

(d) (a)、(b) および／もしくは (c) の任意の組合せ、

(e) 1つ以上のDNAオリゴヌクレオチド（例えば、天然または非天然のDNAオリゴヌクレオチド）、

(f) 1つ以上のRNAオリゴヌクレオチド（例えば、天然または非天然のRNAオリゴヌクレオチド）、

(g) (e) および (f) の組合せ、または

(h) 上記の任意のものの組合せ。

【0085】

異なるオリゴヌクレオチド組合せは、異なる比、例えば1:1、1:2、1:3、1:4、1:5、1:10、1:20、1:50；1:100または1:1000などから選択される比で混合され得る。一実施形態において、化学的に合成されたバイト対アレイで発生されたバイトの比は、1:5、1:10または1:20から選択される。DNAまたはRNAオリゴヌクレオチドは、天然または非天然であり得る。特定の実施形態において、バイトは、1つ以上の非天然ヌクレオチドを含むことで、例えば溶融温度を増加させる。例示的非天然オリゴヌクレオチドとしては、修飾されたDNAまたはRNAヌクレオチドが挙げられる。例示的修飾ヌクレオチド（例えば、修飾されたRNAまたはDNAヌクレオチド）としては、これらに限定されないが、LNAヌクレオチドのリボース部分が、2' 酸素および4' 炭素を接続する余分な架橋で修飾されている、ロックド核酸（LNA）；ペプチド核酸（PNA）、例えば、ペプチド結合によって連結されている反復N-（2-アミノエチル）-グリシン単位から構成されているPNA；低GC領域を捕捉するように修飾されているDNAまたはRNAオリゴヌクレオチド；二環式核酸（BNA）；架橋オリゴヌクレオチド；修飾5-メチルデオキシシチジン；および2, 6-ジアミノプリンが挙げられる。他の修飾DNAおよびRNAヌクレオチドは、当技術分野において知られている。

【0086】

特定の実施形態において、標的配列（例えば、標的メンバー）の実質的に均一または均質のカバレッジが得られる。例えば、各バイトセット／標的カテゴリー内で、カバレッジの均一性は、バイトパラメータを修飾することによって、例えば以下の1つ以上によって最適化され得る：

(i) バイトの表示または重なりが増加／減少が使用されて、同じカテゴリーにおける他の標的に対して、不足して／過度にカバーされる標的（例えば、標的メンバー）のカバレッジを増強／低減することができる；

(ii) 標的配列（例えば、高GC含有量配列）を捕捉するのが困難な低いカバレッジは、バイトセットで標的化されている領域を拡大して、例えば隣接配列（例えば、GCが豊富でない隣接配列）をカバーする；

(iii) バイト配列を修飾して、バイトの二次構造を低減し、選択のその効率を増強することができる；

(iv) バイトの長さの修飾が使用されて、同じカテゴリー内の異なるバイトの融解ハイブリダイゼーション動態を等しくすることができる。バイトの長さは、直接的に（変動する長さを有するバイトを生成することによって）または間接的に（一貫した長さのバイトを生成すること、およびバイト末端を任意の配列と置き換えることによって）修飾され得る；

(v) 同じ標的領域（すなわち、前方向鎖および逆方向鎖）に対して異なる配向のバイトの修飾は、異なる結合効率を有することができる。各標的に対して最適なカバレッジを提供するいずれかの配向を有するバイトセットが選択され得る；

(vi) 結合実体、例えば各バイト上に存在する捕捉タグ（例えば、ビオチン）の量の修飾は、その結合効率に影響することができる。特異的標的を標的にするバイトのタグレベルの増加／減少が使用されて、相対的な標的カバレッジを増強／低減することができる；

(vii) 異なるバイトのために使用されるヌクレオチドの型の修飾が変化されて、標的に対する結合親和性に影響し、相対的な標的カバレッジを増強／低減することができる；または

(viii) 例えばより安定な塩基対形成を有する修飾オリゴヌクレオチドバイトが使用されて、高GC含有量に対して低いまたは正常のGC含有量の区域間の融解ハイブリダイゼーション動態を等しくすることができる。

【0087】

例えば、異なる型のオリゴヌクレオチドバイトセットが使用され得る。一実施形態において、選択の効率についての値は、異なる型のバイトオリゴヌクレオチドを使用することによって修飾されて、予備選択された標的領域を包含する。例えば、第1のバイトセット（例えば、10,000-50,000のRNAまたはDNAバイトを含むアレイベースのバイトセット）が使用されて、大きな標的領域（例えば、1-2MBの合計標的領域）をカバーすることができる。第1のバイトセットは、第2のバイトセット（例えば、5,000未満のバイトを含む、個別に合成されたRNAまたはDNAバイトセット）でスパイクされて、予備選択された標的領域（例えば、標的領域の目的のスパンニングの選択サブゲノム間隔、例えば250kb以下）および／またはより高い二次構造、例えばより高GC含有量の領域をカバーすることができる。目的の選択サブゲノム間隔は、本明細書に記載されている遺伝子もしくは遺伝子産物、またはこれらの断片の1つ以上に対応することができる。第2のバイトセットは、所望のバイト重なり依存して、約1-5,000、2-5,000、3-5,000、10-5,000、100-5,000、500-5,000、1000-5,000、2,000-5,000のバイトを含むことができる。他の実施形態において、第2のバイトセットは、第1のバイトセット中にスパイクされた選択オリゴバイト（例えば、400、200、100、50、40、30、20、10、5、4、3、2または1未満のバイト）を含むことができる。第2のバイトセットは、個別のオリゴバイトの任意の比で混合することができる。例えば、第2のバイトセットは、1:1の等モル比として存在する個別のバイトを含むことができ

10

20

30

40

50

る。代替として、第2のベイトセットは、異なる比（例えば、1：5、1：10、1：20）で存在する個別のベイトを含むことで、例えば、特定の標的の捕捉を最適化することができる（例えば、特定の標的は、他の標的と比較して、5－10×の第2のベイトを有することができる。）。

【0088】

他の実施形態において、選択の効率は、等モルミックスのベイトを使用する場合に観察される示差配列捕捉効率を参照にしてベイトの相対的豊富度または結合実体の密度（例えば、ハプテンまたは親和性タグの密度）を調整することおよび次いで内部平準化群2に対して示差過剰の内部平準化群1を全体的ベイトミックスに導入することにより、群内（例えば、第1、第2または第3の複数のベイト）の個別のベイトの効率を平準化することによって調整される。

【0089】

一実施形態において、方法は、腫瘍メンバー、例えば腫瘍細胞からのサブゲノム間隔を含む核酸分子を選択するベイトセット（本明細書において、「腫瘍ベイトセット」とも称される。）を含む複数のベイトセットの使用を含む。腫瘍メンバーは、腫瘍細胞または癌細胞中に存在する、本明細書において記載されている通りの、腫瘍細胞、例えば変異型、野生型、PGx、参照またはイントロンヌクレオチド配列中に存在する任意のヌクレオチド配列であってよい。一実施形態において、腫瘍メンバーは、低い頻度で出現する改変（例えば、1つ以上の変異）を含め、例えば、腫瘍試料からの細胞の約5%以下が、これらのゲノムにおける改変を持つ。他の実施形態において、腫瘍メンバーは、腫瘍試料からの細胞の約10%の頻度で出現する改変（例えば、1つ以上の変異）を含む。他の実施形態において、腫瘍メンバーは、PGx遺伝子または遺伝子産物からのサブゲノム間隔、イントロン配列、例えば本明細書において記載されている通りのイントロン配列、腫瘍細胞中に存在する参照配列を含む。

【0090】

別の態様において、本発明は、本明細書に記載されているベイトセット、本明細書に記載されている個別のベイトセットの組合せ、例えば本明細書に記載されている組合せを特色とする。ベイトセット（単数または複数）は、指示書、標準物、緩衝剤もしくは酵素、または他の試薬を場合によって含むことができるキットの一部であってよい。

【0091】

遺伝子選択

分析のための予備選択されたサブゲノム間隔、例えば、遺伝子および他の領域のセットまたは群のためのサブゲノム間隔の群またはセットは、本明細書に記載されている。

【0092】

したがって、実施形態において、方法は、獲得された核酸試料からの少なくとも5個、6個、7個、8個、9個、10個、15個、20個、25個、30個、またはこれ以上の遺伝子または遺伝子産物からのサブゲノム間隔を含むライブラリーメンバーの選択および／または配列決定を含み、ここで、遺伝子または遺伝子産物は、ABL1、AKT1、AKT2、AKT3、ALK、APC、AR、BRAF、CCND1、CDK4、CDKN2A、CEBPA、CTNNB1、EGFR、ERBB2、ESR1、FGFR1、FGFR2、FGFR3、FLT3、HRAS、JAK2、KIT、KRAS、MAP2K1、MAP2K2、MET、MLL、MYC、NF1、NOTCH1、NPM1、NRAS、NTRK3、PDGFRA、PIK3CA、PIK3CG、PIK3R1、PTCH1、PTCH2、PTEN、RB1、RET、SMO、STK11、SUFU、またはTP53から選択され、これによって、腫瘍試料を分析する。

【0093】

他の実施形態において、方法は、試料からの少なくとも5個、6個、7個、8個、9個、10個、15個、20個、25個、30個、またはこれ以上の遺伝子または遺伝子産物からのサブゲノム間隔を含むライブラリーメンバーの選択および／または配列決定を含み、ここで、遺伝子または遺伝子産物は、以下から選択される：ABL1、AKT1、AK

T2、AKT3、ALK、APC、AR、BRAF、CCND1、CDK4、CDKN2A、CEBPA、CTNNB1、EGFR、ERBB2、ESR1、FGFR1、FGFR2、FGFR3、FLT3、HRAS、JAK2、KIT、KRAS、MAP2K1、MAP2K2、MET、MLL、MYC、NF1、NOTCH1、NPM1、NRAS、NTRK3、PDGFRA、PIK3CA、PIK3CG、PIK3R1、PTCH1、PTCH2、PTEN、RB1、RET、SMO、STK11、SUFU、またはTP53。

【0094】

別の実施形態において、以下のセットまたは群の1つのサブゲノム間隔が分析される。例えば、腫瘍または癌遺伝子もしくは遺伝子産物、参照（例えば、野生型）遺伝子または遺伝子産物、およびPGx遺伝子または遺伝子産物に関連したサブゲノム間隔は、腫瘍試料からのサブゲノム間隔の群またはセットを提供することができる。

10

【0095】

一実施形態において、方法は、腫瘍試料からのサブゲノム間隔のセットのライブラリーメンバーの選択および／または配列決定を含み、ここで、サブゲノム間隔は、以下の少なくとも1、2、3、4、5、6、7、8、9、10、11、12、13個または全てから選択される：

A) 以下の少なくとも5つ以上から選択される変異もしくは野生型遺伝子または遺伝子産物からの少なくとも5、6、7、8、9、10、15、20、25、30個、またはこれ以上のサブゲノム間隔：ABL1、AKT1、AKT2、AKT3、ALK、APC、AR、BRAF、CCND1、CDK4、CDKN2A、CEBPA、CTNNB1、EGFR、ERBB2、ESR1、FGFR1、FGFR2、FGFR3、FLT3、HRAS、JAK2、KIT、KRAS、MAP2K1、MAP2K2、MET、MLL、MYC、NF1、NOTCH1、NPM1、NRAS、NTRK3、PDGFRA、PIK3CA、PIK3CG、PIK3R1、PTCH1、PTCH2、PTEN、RB1、RET、SMO、STK11、SUFU、またはTP53；

20

B) 以下の少なくとも5個以上から選択される変異もしくは野生型遺伝子または遺伝子産物からのサブゲノム間隔の少なくとも5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、105、110、115、120個、またはこれ以上：ABL2、ARAF、ARFRP1、ARID1A、ATM、ATR、AURKA、AURKB、BAP1、BCL2、BCL2A1、BCL2L1、BCL2L2、BCL6、BRCA1、BRCA2、CBL、CARD11、CBL、CCND2、CCND3、CCNE1、CD79A、CD79B、CDH1、CDH2、CDH20、CDH5、CDK6、CDK8、CDKN2B、CDKN2C、CHEK1、CHEK2、CRKL、CRLF2、DNMT3A、DOT1L、EPHA3、EPHA5、EPHA6、EPHA7、EPHB1、EPHB4、EPHB6、ERBB3、ERBB4、ERG、ETV1、ETV4、ETV5、ETV6、EWSR1、EZH2、FANCA、FBXW7、FGFR4、FLT1、FLT4、FOXP4、GATA1、GNA11、GNAQ、GNAS、GPR124、GUCY1A2、HOXA3、HSP90AA1、IDH1、IDH2、IGF1R、IGF2R、IKBKE、IKZF1、INHBA、IRS2、JAK1、JAK3、JUN、KDM6A、KDR、LRP1B、LRP6、LTK、MAP2K4、MCL1、MDM2、MDM4、MEN1、MITF、MLH1、MPL、MRE11A、MSH2、MSH6、MTOR、MUTYH、MYCL1、MYCN、NF2、NKX2-1、NTRK1、NTRK2、PAK3、PAX5、PDGFRB、PKHD1、PLCG1、PRKDC、PTPN11、PTPRD、RAF1、RARA、RICTOR、RPTOR、RUNX1、SMAD2、SMAD3、SMAD4、SMARCA4、SMARCB1、SOX10、SOX2、SRC、TBX22、TET2、TGFB2、TMPRSS2、TNFAIP3、TNK、TNKS2、TOP1、TSC1、TSC2、USP9X、VHL、またはWT1；

30

40

50

C) 表1、1A、2、3または4による遺伝子または遺伝子産物からの少なくとも5、6、7、8、9、10、15、20個、またはこれ以上のサブゲノム間隔；

D) 腫瘍または癌に関連した、例えば正または負の処置応答予測因子である、正または負の予測因子である、または腫瘍または癌の鑑別診断を可能にする遺伝子または遺伝子産物、例えば以下の1つ以上から選択される遺伝子または遺伝子産物からの少なくとも5、6、7、8、9、10、15、20個、またはこれ以上のサブゲノム間隔：ABL1、AKT1、ALK、AR、BRAF、BRCA1、BRCA2、CEBPA、EGFR、ERBB2、FLT3、JAK2、KIT、KRAS、MET、NPM1、PDGFRA、PIK3CA、RARA、AKT2、AKT3、MAP2K4、NOTCH1、およびTP53；

E) 以下の1つ以上から選択される変異または野生型コドンを含む、少なくとも5、6、7、8、9、10個、またはこれ以上のサブゲノム間隔：ABL1遺伝子のコドン15；APCのコドン1114、1338、1450または1556；BRAFのコドン600；CTNNB1のコドン32、33、34、37、41または45；EGFRのコドン719、746-750、768、790、858または861；FLT3のコドン835；HRASのコドン12、13または61；JAK2のコドン617；KITのコドン816；KRASのコドン12、13または61；PIK3CAのコドン88、542、545、546、1047、または1049；PTENのコドン130、173、233または267；RETのコドン918；TP53のコドン175、245、248、273または306（例えば、表1に示されているコドンの1つ以上を含む、少なくとも5、10、15、20個、またはこれ以上サブゲノム間隔）。

【0096】

F) 以下から選択される薬物代謝、薬物応答性、または毒性（この中で「PGx」遺伝子とも称される。）の1つ以上に関連した遺伝子または遺伝子産物中に存在するサブゲノム間隔の変異もしくは野生型遺伝子または遺伝子産物（例えば、単一ヌクレオチド多型（SNP））からのサブゲノム間隔の少なくとも5、6、7、8、9、10、15、20、25、30個、またはこれ以上：ABCB1、BCC2、ABCC4、ABCG2、C1orf144、CYP1B1、CYP2C19、CYP2C8、CYP2D6、CYP3A4、CYP3A5、DPYD、ERCC2、ESR2、FCGR3A、GSTP1、ITPA、LRP2、MAN1B1、MTHFR、NQO1、NRP2、SLC19A1、SLC22A2、SLCO1B3、SOD2、SULT1A1、TPMT、TYMS、UGT1A1、またはUMPS；

G) 以下の1つ以上に関連した遺伝子または遺伝子産物中に存在するサブゲノム間隔の変異もしくは野生型PGx遺伝子または遺伝子産物（例えば、単一ヌクレオチド多型（SNP））からのサブゲノム間隔の少なくとも5、6、7、8、9、10、15、20、25、30個、またはこれ以上：(i) 薬物で処置された癌患者のより良好な生存（例えば、パクリタキセル（例えば、ABCB1遺伝子）で処置された乳癌患者のより良好な生存）；(ii) パクリタキセル代謝（例えば、表2に示されている、異なる座位および変異でのCYP2C8遺伝子；CYP3A4遺伝子）；(iii) 薬物に対する毒性（例えば、ABCC4遺伝子（表2）で見られる通りの6-MP毒性；DPYD遺伝子、TYMS遺伝子またはUMPS遺伝子（表2）で見られる通りの5-FU毒性；TPMT遺伝子（表2）で見られる通りのプリン毒性；NRP2遺伝子で見られる通りのダウノルビシン毒性；C1orf144遺伝子、CYP1B1遺伝子（表2））；または(iv) 薬物に対する副作用（例えば、ABCG2遺伝子、TYMS遺伝子、UGT1A1遺伝子、ESR1遺伝子およびESR2遺伝子（表2））；

H) 表3による少なくとも5個、10個、15個、20個、25個、30個、35個、40個、45個、50個、75個、110個、またはこれ以上の遺伝子または遺伝子産物の転座改変；

I) ここに明記されている癌型からの固形腫瘍試料における、表3による少なくとも5個、10個、15個、20個、25個、30個、35個、40個、45個、50個、7

10

20

30

40

50

5 個、110 個、またはこれ以上の遺伝子または遺伝子産物の転座改変；

J) 表 4 による少なくとも 5 個、10 個、15 個、20 個、25 個、30 個、35 個、40 個、45 個、50 個、75 個、100 個、150 個、200 個、またはこれ以上の遺伝子または遺伝子産物の転座改変；

K) ここに明記されている癌型からのヘム腫瘍試料における、表 4 による少なくとも 5 個、10 個、15 個、20 個、25 個、30 個、35 個、40 個、45 個、50 個、75 個、100 個、150 個、200 個、またはこれ以上の遺伝子または遺伝子産物の転座改変；

L) 表 1-4 から選択される少なくとも 5 個の遺伝子または遺伝子産物、ここで、例えば予備選択位置での対立遺伝子の変動は、腫瘍の予備選択型に関連し、前記対立遺伝子の変動は、前記腫瘍型における細胞の 5%未満に存在する；

M) GC が豊富な領域に埋め込まれている、表 1、1A-4 から選択される少なくとも 5 個の遺伝子または遺伝子産物；または

N) 癌を発生させる遺伝子（例えば、生殖系列リスク）因子を示す少なくとも 5 個の遺伝子または遺伝子産物（例えば、遺伝子または遺伝子産物は、BRCA1、BRCA2、EGFR、HRAS、KIT、MPL、ALK、PTEN、RET、APC、CDKN2A、MLH1、MSH2、MSH6、NF1、NF2、RB1、TP53、VHLまたはWT1の1つ以上から選択される。）。

【0097】

なお別の実施形態において、方法は、腫瘍試料からのサブゲノム間隔のセットを含むライブラリーメンバーの選択および／または配列決定を含み、ここで、サブゲノム間隔は、表 1B に記載されている改変の 1、2、3、4、5、10、15 個、または全てから選択される。

【0098】

一実施形態において、サブゲノム間隔は、カテゴリー A、B、C、D または E の 1 つ以上に分類される改変を含む。

【0099】

他の実施形態において、サブゲノム間隔は、腫瘍試料、例えば大腸、肺または乳房腫瘍の試料におけるKRAS G13Dの改変を含む。

【0100】

他の実施形態において、サブゲノム間隔は、腫瘍試料、例えばメラノーマまたは大腸腫瘍の試料におけるNRAS Q61Kの改変を含む。

【0101】

なお他の実施形態において、サブゲノム間隔は、腫瘍試料、例えばメラノーマ、大腸または肺腫瘍の試料におけるBRAF V600Eの改変を含む。

【0102】

他の実施形態において、サブゲノム間隔は、腫瘍試料、例えば肺腫瘍試料におけるBRAF D594Gの改変を含む。

【0103】

他の実施形態において、サブゲノム間隔は、腫瘍試料、例えば乳房または大腸腫瘍の試料におけるPIK3CA H1047Rの改変を含む。

【0104】

なお他の実施形態において、サブゲノム間隔は、腫瘍試料、例えば肺腫瘍試料におけるEGFR L858RまたはT790Mの改変を含む。

【0105】

他の実施形態において、サブゲノム間隔は、腫瘍試料におけるERBB2の改変、例えば乳房腫瘍試料におけるERBB2増幅を含む。

【0106】

他の実施形態において、サブゲノム間隔は、腫瘍試料におけるBRCA1の改変、例えば乳房腫瘍試料におけるBRCA1二対立遺伝子の不活性化を含む。

10

20

30

40

50

【0107】

他の実施形態において、サブゲノム間隔は、腫瘍試料におけるBRCA2の改変、例えば膀胱腫瘍試料におけるBRCA2二対立遺伝子の不活性化を含む。

【0108】

他の実施形態において、サブゲノム間隔は、腫瘍試料におけるATMの改変、例えば乳房腫瘍試料におけるATM二対立遺伝子の不活性化を含む。

【0109】

他の実施形態において、サブゲノム間隔は、腫瘍試料におけるTSCの改変、例えば大腸腫瘍試料におけるTSC二対立遺伝子の不活性化を含む。

【0110】

他の実施形態において、サブゲノム間隔は、腫瘍試料におけるPTENの改変、例えば乳房または大腸腫瘍の試料におけるPTEN二対立遺伝子の不活性化を含む。

【0111】

なお他の実施形態において、サブゲノム間隔は、腫瘍試料におけるVHLの改変、例えば腎腫瘍試料におけるVHL二対立遺伝子の不活性化を含む。

【0112】

他の実施形態において、サブゲノム間隔は、腫瘍試料におけるATRの改変、例えば乳房腫瘍試料におけるATR二対立遺伝子の不活性化を含む。

【0113】

他の実施形態において、サブゲノム間隔は、腫瘍試料におけるMYCの改変、例えば乳房腫瘍試料におけるMYC二対立遺伝子の不活性化を含む。

【0114】

サブゲノム間隔のこれらのおよび他のセットおよび群は、本明細書において他所でより詳細に考察されている。

【図面の簡単な説明】

【0115】

【図1】ハイブリッド捕捉方法を用いる所望の鋳型の選択につながる鋳型ライブラリーの典型的な設定を描写している図である。

【図2A】102：102の二重鎖形成に好都合である温度条件下で鋳型203において見出される対応アダプター102配列にハイブリダイズするための、ブロッカーとしてオリゴヌクレオチド102を使用する従来のオリゴヌクレオチドブロッキング戦略を描写している図である。複数の鋳型203は、1つのオリゴヌクレオチドベイト204の結合および固定化支持体205上の捕捉試薬に対するその相互作用を介して捕捉されることに注目されたい。

【図2B】NGS鋳型の複合プールからの望まれないDNA配列の同時選択がない、所望のDNA標的の濃縮のためのT_m増強オリゴヌクレオチドブロッキング戦略を描写する図である。ブロッカーとしてオリゴヌクレオチド102を使用するよりむしろ、戦略は、ブロッカーとしてT_m増強オリゴヌクレオチド202を使用して、202：102の二重鎖形成に好都合である温度条件下で鋳型203において見出される対応アダプター102配列にハイブリダイズする。202：102の二重鎖は、最適な増強T_m値に近い温度で、102：102の二重鎖より好都合であるので、より少ない所望されない鋳型203は、1つのオリゴヌクレオチドベイト204の結合および固定化支持体205上の捕捉試薬に対するその相互作用を介して捕捉される。

【図3A】腫瘍試料の多重遺伝子分析のための方法の実施形態の工程表描写を示す図である。試料受容、品質管理およびDNA単離のための実施形態の工程表描写を描写している。

【図3B】腫瘍試料の多重遺伝子分析のための方法の実施形態の工程表描写を示す図である。DNA品質管理およびライブラリー生成のための実施形態の工程表描写を描写している。

【図3C】腫瘍試料の多重遺伝子分析のための方法の実施形態の工程表描写を示す図である。ハイブリッド捕捉および配列決定のための実施形態の工程表描写を描写している。

10

20

30

40

50

【図3D】腫瘍試料の多重遺伝子分析のための方法の実施形態の工程表描写を示す図である。配列データ品質管理および変異呼び出しのための実施形態の工程表描写を描写している。

【図3E】腫瘍試料の多重遺伝子分析のための方法の実施形態の工程表描写を示す図である。レポート生成のための実施形態の工程表描写を描写している。

【図3F】腫瘍試料の多重遺伝子分析のための方法の実施形態の工程表描写を示す図である。レポート生成の追加の詳細のための実施形態の工程表描写を描写している。

【図4】変異検出における事前予想および読み取り深度の影響を描写するグラフである。

【図5】100個超の臨床的癌試料における変異頻度を描写する図である。

【図6】カバレッジヒストグラムの線状表示を示すグラフである。標的の数（y軸）は、カバレッジ（x軸）の関数として描写されている。線番号1は、個別合成のデオキシリボ核酸オリゴヌクレオチドベイトでスパイクされたアレイ誘導のデオキシリボ核酸オリゴヌクレオチドベイトを含むベイトセット（本明細書において「ベイトセット番号1」と称される。）を使用するカバレッジを表す。線番号2は、アレイ誘導のデオキシリボ核酸オリゴヌクレオチドベイトのみを含むベイトセット（本明細書において「ベイトセット番号2」と称される。）を使用して得られたカバレッジを表す。ベイトセット番号2を使用する全体平均カバレッジは924であったが、一方、ベイトセット番号2を使用する高GC含有量（約68%）の区域におけるカバレッジは73であった。対照的に、ベイトセット番号1が使用された場合、全体カバレッジは約918であったが、カバレッジは、高GC含有量の区域における183に対して改善された。

【図7】アレイ誘導のデオキシリボ核酸オリゴヌクレオチドベイトのみ（「ベイトセット番号3」）を含むベイトセットと比較して、個別合成のデオキシリボ核酸オリゴヌクレオチドベイトのみからなるベイトセット（ベイトセット番号1）および個別合成のデオキシリボ核酸オリゴヌクレオチドベイトでスパイクされたアレイ誘導のデオキシリボ核酸オリゴヌクレオチドベイトを含むベイトセット（「ベイトセット番号2」）を用いて検出されたカバレッジにおける均一性を比較しているカバレッジヒストグラムを示す図である。ベイトセットは、図7における番号1、2および3として示されている。カバレッジにおけるいくつかのギャップは、図7に描写されている通り、ベイトセット番号3を使用して検出されたが、ベイトセット番号1-2を使用して検出されなかった。

【図8】ライブラリーメンバーの非標的鎖状体の例示的立体配置を形成しているダイアグラムを例示する図である。非標的領域（例えば、「P5」および「P7」として描写されているアダプター）は、これらの相補的非標的鎖（それぞれ「rcP5」および「rcP7」として描写されている。）にハイブリダイズしていると示されている。ライブラリーメンバー標的挿入の相補的領域にハイブリダイズしているデオキシリボ核酸タグ付けベイトが示されている。

【図9】標準的および延長されたブロッキングオリゴを使用する標的選択の百分率を描写している棒グラフを示す図である。

【図10】標準的または延長されたブロッカーを使用する捕捉結果を示すエクソンカバレッジヒストグラムを描写している図である。

【図11】1試料当たりの配列読み取りの数を、実施形態のハイブリッド捕捉相において使用されたブロッカーの異なる型の関数として描写している図である。図において左から右へ、非濃縮対照（すなわち、ブロッキングオリゴヌクレオチドを用いない。）；「P7/P5」オリゴ（配列番号それぞれ81および23）；「P7 Comp 6xI/P5」オリゴ（それぞれ配列番号82および23）；「P7 Comp Med BNA 6xI/P5 Med BNA」オリゴ（それぞれ配列番号84および85）；「P7 Comp High BNA 6xI/P5 High BNA」オリゴ（それぞれ配列番号86および87）；「P7 6xI/P5 Comp」オリゴ（それぞれ配列番号88および89）；「P7 Med BNA 6xI/P5 Comp Med BNA」オリゴ（それぞれ配列番号90および91）；および「P7 High BNA 6xI/P5 Comp High BNA」オリゴ（それぞれ配列番号92および93）。

【図12】 $>1\times$ （すなわち、1倍超）でカバーされた標的領域の百分率を、実施形態のハイブリッド捕捉相において使用されたブロッカーの異なる型の関数として描写している図である。図において左から右へ、「P7/P5」オリゴ（それぞれ配列番号81および23）；「P7 Comp 6xI/P5」オリゴ（それぞれ配列番号82および23）；「P7 Comp Med BNA 6xI/P5 Med BNA」オリゴ（それぞれ配列番号84および85）；「P7 Comp High BNA 6xI/P5 High BNA」オリゴ（それぞれ配列番号86および87）；「P7 6xI/P5 Comp」オリゴ（それぞれ配列番号88および89）；「P7 Med BNA 6xI/P5 Comp Med BNA」オリゴ（それぞれ配列番号90および91）；および「P7 High BNA 6xI/P5 Comp High BNA」オリゴ（それぞれ配列番号92および93）；百分率は、 $2\times$ 、 $10\times$ 、 $20\times$ および $30\times$ でカバーされた標的領域の関数として表現されている（描写されている各々について左から右へ、試験されたブロッカー対合型。）。 10

【図13】標的上でアラインメントする配列読み取りの百分率を、実施形態のハイブリッド捕捉相において使用されたブロッカーの異なる型の関数として描写している図である。図において左から右へ、「P7/P5」オリゴ（それぞれ配列番号81および23）；「P7 Comp 6xI/P5」オリゴ（それぞれ配列番号82および23）；「P7 Comp Med BNA 6xI/P5 Med BNA」オリゴ（それぞれ配列番号84および85）；「P7 Comp High BNA 6xI/P5 High BNA」オリゴ（それぞれ配列番号86および87）；「P7 6xI/P5 Comp」オリゴ（それぞれ配列番号88および89）；「P7 Med BNA 6xI/P5 Comp Med BNA」オリゴ（それぞれ配列番号90および91）；および「P7 High BNA 6xI/P5 Comp High BNA」オリゴ（それぞれ配列番号92および93）。 20

【図14】オンターゲット配列の倍数濃縮を、実施形態のハイブリッド捕捉相において使用されたブロッカーの異なる型の関数として描写している図である。図において左から右へ、「P7/P5」オリゴ（それぞれ配列番号81および23）；「P7 Comp 6xI/P5」オリゴ（それぞれ配列番号82および23）；「P7 Comp Med BNA 6xI/P5 Med BNA」オリゴ（それぞれ配列番号84および85）；「P7 Comp High BNA 6xI/P5 High BNA」オリゴ（それぞれ配列番号86および87）；「P7 6xI/P5 Comp」オリゴ（それぞれ配列番号88および89）；「P7 Med BNA 6xI/P5 Comp Med BNA」オリゴ（それぞれ配列番号90および91）；および「P7 High BNA 6xI/P5 Comp High BNA」オリゴ（それぞれ配列番号92および93）。 30

【図15】標的上での配列読み取りの百分率を実施形態のハイブリッド捕捉相において使用されたブロッカーの異なる型の関数として描写している図である。図において左から右へ、「0」オリゴは、BNA修飾 T_m 増強基を含有しない（それぞれ配列番号94および97）；「8」オリゴは、各ブロッカーオリゴにおいて8個のBNA修飾 T_m 増強基を含有し（それぞれ配列番号95および98）；「22」オリゴは、ブロッカーオリゴに依存して各ブロッカーオリゴにおいて17個または22個のBNA修飾 T_m 増強基を含有し（それぞれ配列番号96および97）；ブロッカー群の各型についての個別の棒グラフは、非依存性複製アッセイを表す。 40

【発明を実施するための形態】

【0116】

特定の用語が最初に定義されている。追加の用語は明細書の全体にわたって定義されている。

【0117】

本明細書において使用される用語は、「オープン」用語として意図される（例えば、「含む」という用語は「含むが限定されない」と解釈されるべきであり、「有する」という用語は「少なくとも有する」と解釈されるべきであり、「挙げられる」という用語は「挙げられるが限定されない」と解釈されるべきである、など。）。 50

【0118】

さらに、「A、BおよびCの少なくとも1つ、など」に類似した慣例が使用される例において、一般にこうした構文は、当業者が慣例を理解するという意味において意図される（例えば、「A、BおよびCの少なくとも1つを有するシステム」は、これらに限定されないが、Aを単独で、Bを単独で、Cを単独で、AおよびBと一緒に、AおよびCと一緒に、BおよびCと一緒に、ならびに／またはA、BおよびCと一緒に有するシステムを含む。）。2つ以上の代替の用語を表す実質的に任意の離接語および／または成句は、記載中または図中であろうと、用語の1つ、用語のいずれかまたは両方の用語を含むという可能性を企図すると理解されるべきであることが、当業者によってさらに理解される。例えば、「AまたはB」という成句は、「AまたはB」あるいは「AおよびB」の可能性を含むと理解される。

10

【0119】

「から」「まで」「最大～まで」、「少なくとも」、「超」および「未満」などの全ての言語は、列挙されている数を含め、上記で考察されている通り下位範囲に引き続いて分けられ得る範囲を指す。

【0120】

範囲は、各個別メンバーを含む。したがって、例えば、1－3つのメンバーを有する群は、1つ、2つまたは3つのメンバーを有する群を指す。同様に、6つのメンバーを有する群は、1つ、2つ、3つ、4つまたは6つのメンバーを有する群、その他を指す。

【0121】

「あり得る」という法動詞は、1つ以上の選択肢の好ましい使用もしくは選択、または同一内に含有されているいくつかの記載の実施形態もしくは特色の中での選択を指す。同一に含有されている特別な実施形態または特色に関して、選択肢または選択が開示されていない場合、「あり得る」という法動詞は、同一に含有されている記載の実施形態もしくは特色の態様をどのように作製もしくは使用するのかに関する肯定的行為、または同一に含有されている記載の実施形態もしくは特色に関する特異的技術を使用するという確定的決定を指す。この後者の文脈において、「あり得る」という法動詞は、「できる」という助動詞と同じ意味および含意を有する。

20

【0122】

本明細書で使用される場合、「a」および「an」という冠詞は、冠詞の文法上の目的語の1つまたは1つ超まで（例えば、少なくとも1つまで）を指す。

30

【0123】

「約」および「およそ」は、一般に、測定の性質または精度を前提として、測定された定量についてのエラーの許容程度を意味する。エラーの例示的程度は、所定の値または値の範囲の20－25パーセント（%）内、典型的に10%内、より典型的に5%内である。

【0124】

「獲得する」または「獲得すること」は、用語が本明細書において使用される場合、物理的実体または値を「直接的に獲得すること」または「間接的に獲得すること」によって、物理的実体、または値、例えば数値の所有を得ることを指す。「直接的に獲得すること」は、物理的実体または値を得るためのプロセス（例えば、合成または分析の方法を行うこと）を行うことを意味する。「間接的に獲得すること」は、別の団体または供給源（例えば、物理的実体または値を直接的に獲得した第三者の実験室）から物理的実体または値を受容することを指す。物理的実体を直接的に獲得することは、物理的物質、例えば出発材料における物理的変化を含むプロセスを行うことを含む。例示的变化は、2つまたは1つの出発材料から物理的実体を作製すること、物質を剪断または断片化すること、物質を分離または精製すること、2つ以上の別々の実体を混合物に組み合わせること、共有結合または非共有結合を切断または形成することを含む化学反応を行うことを含む。値を直接的に獲得することは、試料または別の物質における物理的変化を含むプロセスを行うこと、例えば、物質、例えば試料、分析物、または試薬における物理的変化を含む分析的プロセス（本明細書において「物理的分析」と称されることがある。）を行うこと、分析方法

40

50

、例えば、以下の1つ以上を含む方法を行うことを含む：物質、例えば分析物、または別の物質からの断片もしくはこの他の誘導体を分離または精製すること；分析物、またはこの断片もしくは他の誘導体と、別の物質、例えば緩衝剤、溶媒または反応物とを組み合わせること；または分析物またはこの断片もしくは他の誘導体の構造を、例えば分析物の第1原子と第2原子との間で共有結合または非共有結合を切断もしくは形成することによって変化させること；または試薬またはこの断片もしくは他の誘導体の構造を、例えば試薬の第1原子と第2原子との間で共有結合または非共有結合を切断もしくは形成することによって変化させること。

【0125】

「配列を獲得すること」または「読み取りを獲得すること」は、用語が本明細書において使用される場合、ヌクレオチド配列またはアミノ酸配列の所有を得ることを指し、配列または読み取りを「直接的に獲得すること」または「間接的に獲得すること」によって得ることを指す。配列または読み取りを「直接的に獲得すること」は、配列決定方法（例えば、次世代配列決定（NGS）方法）を行うことなど、配列を得るためのプロセス（例えば、合成または分析の方法を行うこと）を行うことを意味する。配列または読み取りを「間接的に獲得すること」は、別の団体または供給源（例えば、配列を直接的に獲得した第三者の実験室）から配列の情報または知識を受容すること、または配列を受容することを指す。獲得された配列または読み取りは、全配列である必要がなく、例えば、少なくとも1つのヌクレオチドを配列決定すること、または対象に存在する通りの本明細書において開示されている改変の1つ以上を同定する情報もしくは知識を得ることは、配列を獲得することと相当する。

【0126】

配列または読み取りを直接的に獲得することは、物理的物質、例えば出発材料、例えば組織もしくは細胞の試料、例えば生検、または単離された核酸（例えば、DNAまたはRNA）試料における物理的変化を含むプロセスを行うことを含む。例示的変化は、2つ以上の出発材料から物理的実体を作製すること、ゲノムDNA断片などの物質を剪断または断片化すること；物質を分離または精製すること（例えば、組織から核酸試料を単離すること）；2つ以上の別々の実体を混合物に組み合わせること、共有結合または非共有結合を切断または形成することを含む化学反応を行うことを含む。値を直接的に獲得することは、上に記載されている通りの試料または別の物質における物理的変化を含むプロセスを行うことを含む。

【0127】

「試料を獲得すること」は、用語が本明細書において使用される場合、試料を「直接的に獲得すること」または「間接的に獲得すること」によって、試料、例えば、組織試料または核酸試料の所有を得ることを指す。「試料を直接的に獲得すること」は、試料を得るためのプロセス（例えば、外科手術または抽出などの物理的方法を行うこと）を行うことを意味する。「試料を間接的に獲得すること」は、別の団体または供給源（例えば、試料を直接的に獲得した第三者の実験室）から試料を受容することを指す。試料を直接的に獲得することは、物理的物質、例えば出発材料、例えば組織、例えばヒト患者における組織または患者から事前に単離された組織における物理的変化を含むプロセスを行うことを含む。例示的変化としては、出発材料から物理的実体を作製すること、組織を解剖または断片化すること；物質（例えば、試料組織または核酸試料）を分離または精製すること；2つ以上の別々の実体を混合物に組み合わせること；共有結合または非共有結合を切断または形成することを含む化学反応を行うことが挙げられる。試料を直接的に獲得することには例えば上に記載されている通りの、試料または別の物質における物理的変化を含むプロセスを行うことを含む。

【0128】

遺伝子または遺伝子産物（例えば、マーカー遺伝子または遺伝子産物）の、本明細書で使用される場合の「改変」または「改変された構造」は、正常または野生型の遺伝子と比較して、変異または遺伝子もしくは遺伝子産物内の変異、例えば遺伝子もしくは遺伝子産

10

20

30

40

50

物の量または活性に影響する変異の存在を指す。改変は、正常または健常の組織または細胞（例えば、対照）における量、構造および／または活性と比較して、癌組織または癌細胞における量、構造および／または活性における改変であってよく、癌などの疾患状態を伴う。例えば、癌を伴うまたは抗癌治療に対する応答性の前兆となる改変は、正常の健常組織または細胞と比較して、癌組織または癌細胞における改変されたヌクレオチド配列（例えば、変異）、アミノ酸配列、染色体転座、染色体内逆位、コピー数、発現レベル、タンパク質レベル、タンパク質活性、またはメチル化状態を有することができる。例示的変異としては、これらに限定されないが、点変異（例えば、サイレント、ミスセンスまたはナンセンス）、欠失、挿入、逆位、連結変異、重複、転座、染色体間および染色体内の再編成が挙げられる。変異は、遺伝子のコード領域または非コード領域中に存在することができる。特定の実施形態において、改変（単数または複数）は、再編成、例えば、1つ以上のイントロンまたはこの断片を含むゲノム再編成（例えば、5' - および／または 3' - UTR における1つ以上の再編成）として検出される。特定の実施形態において、改変は、表現型、例えば癌表現型（例えば、癌リスク、癌進行、癌処置、または癌処置に対する抵抗性の1つ以上）と関連する（または関連しない）。一実施形態において、改変は、癌に対する遺伝子危険因子、陽性処置応答予測因子、陰性処置応答予測因子、陽性予後因子、陰性予後因子または診断因子の1つ以上と関連する。

【0129】

「ベイト」は、本明細書で使用される場合、ハイブリッド捕捉試薬の型である。ベイトは、標的核酸にハイブリダイズする（例えば、相補的である。）ことができ、これによっての標的核酸の捕捉を可能にする核酸分子、例えばDNAまたはRNA分子であってよい。一実施形態において、ベイトは、RNA分子（例えば、天然のまたは修飾されたRNA分子）；DNA分子（例えば、天然のまたは修飾されたDNA分子）、またはこれらの組合せである。他の実施形態において、ベイトとしては、例えば、ベイトおよびベイトにハイブリダイズされた核酸によって形成されたハイブリッドを結合実体に結合することによって捕捉および分離を可能にする結合実体、例えば親和性タグが挙げられる。一実施形態において、ベイトは、溶液相ハイブリダイゼーションに適當である。

【0130】

「ベイトセット」は、本明細書で使用される場合、1つ以上のベイト分子を指す。

【0131】

「結合実体」は、分析物に特異的に結合できる、分子タグが直接的または間接的に付着され得る任意の分子を意味する。結合実体は、各ベイト配列上の親和性タグであり得る。特定の実施形態において、結合実体は、アビジン分子またはハプテンもしくはこの抗原結合断片に結合する抗体などのパートナーに結合することによって、ハイブリダイゼーション混合物からのベイト／メンバーハイブリッドの分離を可能にする。例示的結合実体としては、これらに限定されないが、ビオチン分子、ハプテン、抗体、抗体結合断片、ペプチドおよびタンパク質が挙げられる。

【0132】

「相補的」は、2つの核酸鎖の領域間、または同じ核酸鎖の2つの領域間の配列相補性を指す。第1の核酸領域のアデニン残基は、残基がチミンまたはウラシルであるならば第1の領域に逆平行である第2の核酸領域の残基との特異的水素結合（「塩基対合」）を形成できることが知られている。同様に、第1の核酸鎖のシトシン残基は、残基がグアニンであるならば第1の鎖に逆平行である第2の核酸鎖の残基と塩基対合できることが知られている。核酸の第1の領域は、2つの領域が逆平行様式で編成されている場合に第1の領域の少なくとも1つのヌクレオチド残基が第2の領域の残基と塩基対合できるならば、同じまたは異なる核酸の第2の領域に相補的である。特定の実施形態において、第1の領域は第1の部分を含み、第2の領域は第2の部分を含み、これによって、第1および第2の部分が逆平行様式で編成される場合、第1の部分のヌクレオチド残基の少なくとも約50%、少なくとも約75%、少なくとも約90%、または少なくとも約95%は、第2の部分におけるヌクレオチド残基と塩基対合できる。他の実施形態において、第1の部分の全

でのヌクレオチド残基は、第2の部分におけるヌクレオチド残基と塩基対合できる。

【0133】

「癌」または「腫瘍」という用語は、本明細書において相互交換可能に使用される。これらの用語は、非制御増殖、不死性、転移能、急速な成長率および増殖率、ならびに特定の特徴的形態学的特色など、癌原因細胞に典型的な特徴を所有する細胞の存在を指す。癌細胞は、しばしば腫瘍の形態であるが、こうした細胞は、動物内に単独で存在することができる、または白血病細胞などの非腫瘍原性癌細胞であり得る。これらの用語は、固形腫瘍、軟組織腫瘍、または転移巣を含む。本明細書で使用される場合、「癌」という用語は、前悪性癌、ならびに悪性癌を含む。

【0134】

「しそうである」または「見込みの増加」は、本明細書で使用される場合、物品、物体、物または人間が生じる確率の増加を指す。したがって、1つの例において、処置に応答しそうである対象は、参照対象または対象の群に対して、処置に応答する確率の増加を有する。

【0135】

「しそうでない」は、参照に関して事象、物品、物体、物または人間が生じる確率の減少を指す。したがって、処置に応答しそうでない対象は、参照対象または対象の群に対して、処置に応答する確率の減少を有する。

【0136】

「対照メンバー」は、非腫瘍細胞からの配列を有するメンバーを指す。

【0137】

「インデルアラインメント配列セレクター」は、本明細書で使用される場合、予備選択インデルの場合において読み取りがアラインメントされるべき配列の選択を可能にするまたは指向するパラメータを指す。こうした配列の使用は、インデルを含む予備選択サブゲノム間隔の配列決定を最適化することができる。インデルアラインメント配列セレクターについての値は、予備選択インデルの関数、例えばインデルについての識別子である。一実施形態において、値は、インデルの同一性である。

【0138】

本明細書で使用される場合、「ライブラリー」という用語は、メンバーの収集物を指す。一実施形態において、ライブラリーとしては、核酸メンバーの収集物、例えば、全ゲノム、サブゲノム断片、cDNA、cDNA断片、RNA、RNA断片、またはこれらの組合せの収集物が挙げられる。一実施形態において、ライブラリーメンバーの一部分または全ては、非標的アダプター配列を含む。アダプター配列は、一方または両方の末端に位置され得る。アダプター配列は、例えば、配列決定方法（例えば、NGS方法）に、増幅に、逆転写に、またはベクターへのクローニングに有用であり得る。

【0139】

ライブラリーは、メンバー、例えば標的メンバー（例えば、腫瘍メンバー、参照メンバー、PGxメンバー、またはこれらの組合せ）の収集物を含むことができる。ライブラリーのメンバーは、単一の個体からであり得る。実施形態において、ライブラリーは、1個超の対象（例えば、2、3、4、5、6、7、8、9、10、20、30個、またはこれ以上の対象）からのメンバーを含むことができ、例えば、異なる対象からの2つ以上のライブラリーは、組み合わされて、1つ超の対象からのメンバーを有するライブラリーを形成することができる。一実施形態において、対象は、癌または腫瘍を有するまたは有するリスクがあるヒトである。

【0140】

「ライブラリーキャッチ」は、ライブラリーのサブセット、例えば、予備選択サブゲノム間隔、例えば予備選択ベイトとのハイブリダイゼーションによって捕捉された生成物が濃縮されたサブセットを指す。

【0141】

「メンバー」もしくは「ライブラリーメンバー」または他の同様の用語は、本明細書で

10

20

30

40

50

使用される場合、ライブラリーのメンバーである核酸分子、例えばDNA、RNA、またはこれらの組合せを指す。典型的に、メンバーは、DNA分子、例えばゲノムDNAまたはcDNAである。メンバーは、断片化された、例えば剪断された、または酵素調製されたゲノムDNAであり得る。メンバーは、対象からの配列を含み、対象から誘導されたのではない配列、例えば、非標的配列、例えばアダプター配列、プライマー配列、または同定を可能にする他の配列、例えば「バーコード」または「インデックス」配列を含むこともできる。

【0142】

「次世代配列決定またはNGSもしくはNG配列決定」は、本明細書で使用される場合、高スループット様式において、個別核酸分子（例えば、単一分子配列決定における。）または個別核酸分子のためのクローン的に拡大されたプロキシのいずれかのヌクレオチド配列を決定する任意の配列決定方法を指す（例えば、 10^3 つ、 10^4 つ、 10^5 つ超、またはこれ以上の分子が同時に配列決定される。）。一実施形態において、ライブラリーにおける核酸種の相対的豊富度は、配列決定実験によって生成されたデータにおけるそれらの同族配列の出現の相対的数をカウントすることによって概算され得る。次世代配列決定方法は、当技術分野において知られており、例えば、参照により本明細書に組み込む Metzker, M. (2010) Nature Biotechnology Reviews 11: 31-46に記載されている。次世代配列決定は、試料における核酸の5%未満に存在するバリエーションを検出することができる。

【0143】

本明細書において称される「ヌクレオチド値」は、予備選択ヌクレオチド位置を占有するまたは割り当てられるヌクレオチド（単数または複数）の同一性を表す。典型的なヌクレオチド値は：欠損（例えば、欠失）；追加（例えば、同一性を含むことができるまたは含むことができない1つ以上のヌクレオチドの挿入）；または存在（占有）；A；T；C；またはGを含む。他の値は、例えば、YがA、T、GまたはCである場合にはYでなくてよく；XがT、GまたはCの1つまたは2つである場合にはAまたはX；XがA、GまたはCの1つまたは2つである場合にはTまたはX；XがT、AまたはCの1つまたは2つである場合にはGまたはX；XがT、GまたはAの1つまたは2つである場合にはCまたはX；ピリミジンヌクレオチド；またはプリンヌクレオチドであり得る。ヌクレオチド値は、ヌクレオチド位置における1個以上、例えば2個、3個または4個の塩基の頻度（または本明細書に記載されている他の値、例えば欠損または追加）であり得る。例えば、ヌクレオチド値は、ヌクレオチド位置におけるAの頻度およびGの頻度を含むことができる。

【0144】

「または」は、別段に文脈が明確に示していないかぎり、本明細書において「および／または」という用語を意味するために使用されるとともに相互交換可能に使用される。本明細書において一部の箇所における「および／または」という用語の使用は、別段に文脈が明確に示していないかぎり、「または」という用語の使用が「および／または」という用語と相互交換可能でないことを意味しない。

【0145】

「一次対照」は、腫瘍試料におけるNAT組織以外の非腫瘍組織を指す。血液は典型的な一次対照である。

【0146】

「再編成アラインメント配列セレクター」は、本明細書で使用される場合、予備選択再編成の場合において読み取りがアラインメントされるべき配列の選択を可能にするまたは指向するパラメータを指す。こうした配列の使用は、再編成を含む予備選択サブゲノム間隔の配列決定を最適化することができる。再編成アラインメント配列セレクターについての値は、予備選択再編成の関数、例えば再編成の識別子である。一実施形態において、値は再編成の同一性である。「インデルアラインメント配列セレクター」（本明細書において他所でも定義している。）は、再編成アラインメント配列セレクターの一例である。

【0147】

「試料」、「組織試料」、「患者試料」、「患者細胞または組織試料」または「標本」は、各々、対象または患者の組織または循環細胞から得られる同様の細胞の収集物を指す。組織試料の供給源は、新たな凍結および／または保存された臓器、組織試料、生検もしくはは吸引物からの固体組織；血液もしくはは任意の血液構成要素；脳脊髄液、羊水、腹水もしくはは間質液などの体液；または対象の妊娠もしくはは発育における任意の時点からの細胞であってよい。組織試料は、保存料、抗凝固剤、緩衝剤、固定液、栄養素または抗生物質などの組織と本来天然に混合されていない化合物を含有することができる。一実施形態において、試料は、凍結試料としてまたはホルムアルデヒドもしくははパラホルムアルデヒド固定パラフィン包埋（FFPE）組織調製物として保存される。例えば、試料は、マトリックス、例えばFFPEブロックまたは凍結試料中に包埋され得る。

10

【0148】

一実施形態において、試料は腫瘍試料であり、例えば、1つ以上の前悪性細胞または悪性細胞を含む。特定の実施形態において、試料、例えば腫瘍試料は、固形腫瘍、軟組織腫瘍または転移巣から獲得される。他の実施形態において、試料、例えば腫瘍試料は、外科的周縁部からの組織または細胞を含む。別の実施形態において、試料、例えば腫瘍試料は、1つ以上の循環腫瘍細胞（CTC）（例えば、血液試料から獲得されたCTC）を含む。

【0149】

「感受性」は、本明細書で使用される場合、配列の異質の集団における予備選択配列バリエーションを検出するための方法の能力の尺度である。予備選択配列バリエーションが試料における配列の少なくともF%として存在する試料を前提として、方法がC%の予備選択信頼度、S%の確率で予備選択配列を検出することができるならば、方法はF%のバリエーションに対してS%の感受性を有する。例として、予備選択バリエーション配列が試料における配列の少なくとも5%として存在する試料を前提として、方法が99%の予備選択信頼度、10のうち9の確率で予備選択配列を検出することができるならば（F=5%；C=99%；S=90%）、方法は5%のバリエーションに対して90%の感受性を有する。例示的感受性としては、C=90%、95%、99%および99.9%の信頼レベルにて、F=1%、5%、10%、20%、50%、100%で、配列バリエーションについてS=90%、95%、99%の感受性が挙げられる。

20

【0150】

「特異性」は、本明細書で使用される場合、本当に生じる予備選択配列バリエーションを配列決定アーチファクトまたは他の密接に関連した配列と区別するための方法の能力の尺度である。それは、誤った陽性検出を回避するための能力である。誤った陽性検出は、試料調製中に目的の配列に導入されたエラー、配列決定エラー、または偽遺伝子もしくはは遺伝子ファミリーのメンバーのような密接に関連した配列の不注意な配列決定から起こり得る。X_{True}配列が本当にバリエーションであるとともにX_{Not true}が本当にはバリエーションでないN_{Total}配列の試料セットに適用される場合に、方法が、本当ではないバリエーションの少なくともX%をバリエーションでないとして選択するならば、方法は、X%の特異性を有する。例えば、500個の配列が本当にバリエーションであるとともに500個が本当ではないバリエーションである1,000個の配列の試料セットに適用される場合に、方法が、500個の本当ではないバリエーション配列の90%をバリエーションでないとして選択するならば、方法は、90%の特異性を有する。例示的特異性としては、90%、95%、98%および99%が挙げられる。

30

40

【0151】

「腫瘍核酸試料」は、本明細書で使用される場合、腫瘍または癌の試料からの核酸分子を指す。典型的に、それは、DNA、例えば、ゲノムDNA、または腫瘍もしくはは癌の試料からのRNAから誘導されたcDNAである。特定の実施形態において、腫瘍核酸試料は、精製または単離される（例えば、それは、その天然状態から除去される。）。

【0152】

「対照」または「参照」の「核酸試料」は、本明細書で使用される場合、対照または参

50

照の試料からの核酸分子を指す。典型的に、それは、DNA、例えばゲノムDNA、またはRNAから誘導されたcDNAであり、遺伝子または遺伝子産物において改変または変動を含有しない。特定の実施形態において、参照または対照の核酸試料は、野生型または非変異型の配列である。特定の実施形態において、参照核酸試料は、精製または単離される（例えば、それは、その天然状態から除去される。）。他の実施形態において、参照核酸試料は、同じまたは異なる対象からの非腫瘍試料、例えば血液対照、正常の隣接腫瘍（NAT）、または任意の他の非癌性試料からである。

【0153】

核酸分子を「配列決定すること」は、分子における少なくとも1個のヌクレオチドの同一性を決定することを必要とする。実施形態において、分子におけるヌクレオチドの全て未満の同一性が決定される。他の実施形態において、分子におけるヌクレオチドの大多数または全ての同一性が決定される。

【0154】

本明細書において称されている「サブゲノム間隔」は、ゲノム配列の一部を指す。一実施形態において、サブゲノム間隔は、例えばヌクレオチド位置バリエーションが腫瘍表現型と（正または負に）関連する単一のヌクレオチド位置であってよい。一実施形態において、サブゲノム間隔は、1つ超のヌクレオチド位置を含む。こうした実施形態は、少なくとも2、5、10、50、100、150または250のヌクレオチド位置長の配列を含む。サブゲノム間隔は、全遺伝子、またはこの予備選択部分、例えばコード領域（またはこの一部）、予備選択イントロン（またはこの一部）またはエクソン（またはこの一部）を含むことができる。サブゲノム間隔は、天然核酸、例えばゲノム核酸の断片の全てまたは一部を含むことができる。例えば、サブゲノム間隔は、配列決定反応にかけられるゲノムDNAの断片に対応し得る。実施形態において、サブゲノム間隔は、ゲノム供給源からの連続的配列である。実施形態において、サブゲノム間隔は、ゲノムにおいて近接していない配列を含め、例えば、それは、cDNAにおいてエクソン-エクソン接合部が見出される形成された接合部を含む。

【0155】

一実施形態において、サブゲノム間隔は、以下を含むまたは以下からなる：単一のヌクレオチド位置；遺伝子内領域または遺伝子間領域；エクソンもしくはイントロンまたはこれらの断片、典型的にエクソン配列またはこの断片；コード領域または非コード領域、例えば、プロモーター、エンハンサー、5'非翻訳領域（5'UTR）、または3'非翻訳領域（3'UTR）、またはこれらの断片；cDNAまたはこの断片；SNP；体細胞変異、生殖細胞系列変異または両方；改変、例えば点変異または単一変異；欠失変異（例えばインフレーム欠失、遺伝子内欠失、全遺伝子欠失）；挿入変異（例えば、遺伝子内挿入）；逆位変異（例えば、染色体逆位）；連結変異；連結されている挿入変異；逆位された重複変異；タンDEM重複（例えば、染色体タンDEM重複）；転座（例えば、染色体転座、非相互転座）；再編成、例えば、ゲノム再編成、例えば、1つ以上のイントロンの再編成、またはこの断片；再編成イントロンは、5'-および/または3'-UTRを含むことができる。）；遺伝子コピー数における変化；遺伝子発現における変化；RNAレベルにおける変化、またはこの組合せ。「遺伝子のコピー数」は、特別な遺伝子産物をコードする細胞におけるDNA配列の数を指す。一般に、所定の遺伝子について、哺乳動物は、各遺伝子の2つのコピーを有する。コピー数は、例えば遺伝子の増幅もしくは重複によって増加され、または欠失によって低減され得る。

【0156】

「閾値」は、本明細書で使用される場合、サブゲノム間隔にヌクレオチド値を割り当てるために存在するのに必要とされる読み取りの数の関数である値である。例えば、それは、ヌクレオチド位置における特異的ヌクレオチド値、例えばAを有する読み取りの数の関数であって、サブゲノム間隔においてそのヌクレオチド値をそのヌクレオチド位置に割り当てるのに必要とされる関数である。閾値は、例えば、読み取りの数、例えば整数として（または、その関数として）、または予備選択値を有する読み取りの割合として表現され

得る。例として、閾値が X であり、「A」のヌクレオチド値を有する $X+1$ つの読み取りが存在するならば、「A」の値は、サブゲノム間隔における予備選択位置に割り当てられる。閾値は、変異もしくはバリエーションの予想、変異頻度、またはベイズ事前の関数としても表現され得る。一実施形態において、予備選択変異頻度は、そのヌクレオチド値を呼び出すため、予備選択位置におけるヌクレオチド値、例えばAまたはGを有する読み取りの予備選択される数または割合を必要とする。実施形態において、閾値は、変異予想、例えば変異頻度、および腫瘍型の関数であり得る。例えば、予備選択ヌクレオチド位置での予備選択バリエーションは、患者が第1の腫瘍型を有するならば第1の閾値を有し、患者が第2の腫瘍型を有するならば第2の閾値を有し得る。

【0157】

本明細書で使用される場合、「標的メンバー」は、核酸ライブラリーから単離することを所望する核酸分子を指す。一実施形態において、標的メンバーは、本明細書において記載されている通りの腫瘍メンバー、参照メンバー、対照メンバーまたはPGxメンバーであり得る。

【0158】

「腫瘍メンバー」または他の同様の用語（例えば、「腫瘍または癌関連メンバー」）は、本明細書で使用される場合、腫瘍細胞からの配列を有するメンバーを指す。一実施形態において、腫瘍メンバーは、癌表現型と関連する改変（例えば、変異）を有する配列（例えば、ヌクレオチド配列）を有するサブゲノム間隔を含む。他の実施形態において、腫瘍メンバーは、野生型配列（例えば、野生型ヌクレオチド配列）を有するサブゲノム間隔を含む。例えば、癌細胞中に存在するヘテロ接合性またはホモ接合性の野生型対立遺伝子からのサブゲノム間隔。腫瘍メンバーは、参照メンバーまたはPGxメンバーを含むことができる。

【0159】

「参照メンバー」または他の同様の用語（例えば、「対照メンバー」）は、本明細書で使用される場合、癌表現型と関連する配列（例えば、ヌクレオチド配列）を有するサブゲノム間隔を含むメンバーを指す。一実施形態において、参照メンバーは、変異される場合に癌表現型と関連する遺伝子または遺伝子産物の野生型または非変異のヌクレオチド配列を含む。参照メンバーは、癌細胞または非癌細胞中に存在し得る。

【0160】

「PGxメンバー」または他の同様の用語は、本明細書で使用される場合、遺伝子の薬理遺伝学的または薬理ゲノミクスのプロファイルと関連するサブゲノム間隔を含むメンバーを指す。一実施形態において、PGxメンバーは、SNP（例えば、本明細書において記載されている通りのSNP）を含む。他の実施形態において、PGxメンバーは、表1または表2によるサブゲノム間隔を含む。

【0161】

本明細書で使用される場合、「ユニバーサル核酸塩基」は、隣接している塩基対の相互作用を著しく不安定化することなく4つの正常な核酸塩基の任意のものを置き換える能力を呈する核酸塩基を指す。ユニバーサル核酸塩基組成物を含めて、こうした混合核酸塩基組成物がブロッカー中に存在する場合、それらは、好ましくは約5から約12ヌクレオチド長を範囲とする、複数の実質的に連続するヌクレオチド位置を占有する。

【0162】

「バリエーション」は、本明細書で使用される場合、1つ超の構造、例えば対立遺伝子を多型部位に有することができるサブゲノム間隔中に存在することができる構造を指す。

【0163】

見出し、例えば(a)、(b)、(i)などは、明細書および請求項を読み取ることの簡便さのために存在する。明細書または請求項における見出しの使用は、ステップまたは要素が、アルファベット順もしくは番号順またはそれらが表される順で行われることを必要としない。

【0164】

10

20

30

40

50

本発明は、標的濃縮を改善するとともにオフターゲット選択を低減するためのブロッカーおよびベイトとしての新規な T_m 増強オリゴヌクレオチドに関する。オリゴヌクレオチド組成物は、次世代配列決定用途のための核酸鋳型を調製する強力な用途を有する。オリゴヌクレオチドは、 T_m 増強基で修飾されて、ハイブリダイゼーション／捕捉反応が非修飾オリゴヌクレオチドより高い温度およびストリンジェントな洗浄条件下で実行されることを可能にするそれらのそれぞれの標的に対するオリゴヌクレオチドの結合親和性を増加させる。NGS鋳型の末端アダプターと同一の配列を有するオリゴヌクレオチドブロッカーについて、ハイブリッド捕捉方法におけるブロッカーとしての T_m 増強オリゴヌクレオチドの含有は、NGS鋳型中のアダプター媒介ハイブリッド形成に由来する望まれない混入配列のレベルを低減し（「デイズ鎖効果」）、これによって、所望のNGS鋳型のための濃縮プロセスの全体的効率を増加させる。ブロッカーおよびベイトとしての新規な T_m 増強オリゴヌクレオチドの組成物、ならびに大規模平行配列決定実験などの用途におけるそれらの使用を含めて標的濃縮の改善およびオフターゲット選択の低減のためのそれらの特異的使用が、下記においてさらに詳細に開示されている。

【0165】

図1を参照すると、インプットDNA100は、適切なサイズ範囲を提供するために断片化される。その結果得られるDNA断片101の好ましいサイズ範囲は、特別な用途および／またはNGSプラットフォームに依存するが、典型的には200–500bp長の範囲である。DNA100を断片化する好ましい方法は、超音波処理手順を使用してDNAを剪断することによる。市販のソニファイアーおよび他の超音波処理器は、DNA100を適切なサイズ範囲に断片化するために使用され得る。剪断することによってDNA100を断片化することが好ましい断片化手段である一方で、エンドヌクレアーゼ（例えば、DNaseまたは制限エンドヌクレアーゼ）を使用するDNA100の部分的消化など、他の断片化手順が使用され得る。

【0166】

その結果得られるDNA断片101が酵素処理されて、NGS鋳型103を生成するために少なくとも1つの平滑末端を有するオリゴヌクレオチドアダプター102がライゲートされる平滑末端化された末端を調製する。典型的に、剪断DNAは、様々な末端、こうした平滑末端、5'-オーバーハング末端、および3'-オーバーハング末端を含むことができる。5'-オーバーハング末端を含むDNA断片は、適当なポリメラーゼ（例えば、とりわけT4 DNAポリメラーゼ、DNAポリメラーゼIの大きな（クレノウ）断片、Vent DNAポリメラーゼ、Deep Vent DNAポリメラーゼ）を使用して陥凹3'-末端に充填することによって平滑末端化され得る。3'-オーバーハングを含むDNA断片は、好ましくはdNTP（例えば、とりわけT4 DNAポリメラーゼ、DNAポリメラーゼIの大きな（クレノウ）断片、Pfuポリメラーゼ）の存在下でDNAポリメラーゼの3'→5'エキソヌクレアーゼ活性を使用することによって、平滑末端化され得る。5'-オーバーハング末端または3'-オーバーハング末端を有するDNA断片は、一本鎖ヌクレアーゼ（例えば、とりわけマングベーンヌクレアーゼ、P1ヌクレアーゼ、S1ヌクレアーゼ）を使用して平滑末端化することもできる。DNAポリメラーゼの使用は、断片101のための平滑末端化された末端を調製するための使用に好ましい。

【0167】

場合によって、その結果得られる断片101は、酵素操作されて、相補的単一ヌクレオチドオーバーハング（上記の例において、3'-dTオーバーハング）を持つ少なくとも1つの末端を有するアダプター102とのライゲーションを容易にすることができる単一ヌクレオチドオーバーハング（例えば、3'-dAオーバーハング）を含むことができる。こうした断片101は、典型的に、上に記載されている通りの平滑末端化された末端で作製され、次いで引き続いて、3'-ポリメラーゼ（「テーリング」）活性（例えば、とりわけTth DNAポリメラーゼ、Bst DNAポリメラーゼ、Taq DNAポリメラーゼ、クレノウDNAポリメラーゼ（エキソ-））を有する酵素で処理される。

【0168】

さらに、剪断DNAは、二本鎖DNA構造の完全な破損をもたらさない2つの相補鎖の1つ内の内部切断（例えば、ニック）を含むことができる。こうした内部切断は、dNTP（例えば、とりわけT4 DNAポリメラーゼまたはDNAポリメラーゼIの大きな（クレノウ）断片）の存在下、またはATP（例えば、T4 DNAリガーゼ）の存在下での適当なリガーゼの存在下でニック翻訳活性を有するDNAポリメラーゼを使用して、修復され得る。断片101の剪断DNA内の任意の一本鎖切断を修復するのは好ましく、というのは、最終の鑄型103が、好ましくは、2つの連続鎖の各末端上にライゲートされる2つのアダプター102を含むからである。

【0169】

アダプター102は、末端の異なる型を含むように設計されている。この好ましい設計が選択されて、その結果得られる鑄型103の各末端のために二本鎖化アダプター102の単一コピーを提供する。平滑末端化された末端を含むように酵素処理された断片101のため、アダプター102は、平滑末端を有する第1の末端およびオーバーハング末端を有する第2の末端を含むように設計されている。こうしたアダプター102のため、第2の末端は、他のアダプター102（例えば、とりわけリガーゼコンピテント基質、例えば5'-ホスフェート基、3'-ヒドロキシル基、および／または配列相補性を欠如する。）へのライゲーションを妨げる1つ以上の特色を含むように設計されている。単一ヌクレオチド末端を含むように酵素処理された断片101のため、アダプター102は、相補的単一ヌクレオチドオーバーハングを有する第1の末端および異なる型の末端を有する第2の末端を含むように設計されている。上に記載されていることのように、後者アダプター102の第2の末端は、好ましくは、他のアダプター102へのライゲーションを妨げる1つ以上の特色を含むように設計されている。

【0170】

アダプター102のオリゴヌクレオチド組成物は、好ましくは、従来の核酸塩基を含め、ここで、ヌクレオチド間の連結は、従来のホスホジエステル部分である。アダプター102は、好ましくは、下記でさらに説明されている通りにT_m増強特性を示す化学基を除く。オリゴヌクレオチドアダプター102の好ましい長さは、約15ヌクレオチドから約75ヌクレオチドを範囲とする。

【0171】

特定のNGS用途のため、大規模平行配列決定実験において多重配列決定できるようにするために「バーコード」配列を含むことが望ましい。この目的のため、アダプター102は、バーコード配列タグを表す「ユニバーサル」核酸塩基組成物（例えば、とりわけイノシン、3-ニトロピロール、5-ニトロインドール）を含め、混合核酸塩基組成物（例えば、特別な位置（単数または複数）での2つ以上の正準核酸塩基の混合物）を有する複数のヌクレオチド位置を含むことができる。本明細書で使用される場合、「ユニバーサル核酸塩基」は、隣接している塩基対の相互作用を著しく不安定化することなく、4つの正常の核酸塩基の任意のものを置き換える能力を呈する核酸塩基を指す。ユニバーサル核酸塩基組成物を含めて、こうした混合核酸塩基組成物がアダプター102に存在する場合、それらは、好ましくは約5から約12ヌクレオチド長を範囲とする複数の実質的に連続するヌクレオチド位置を占有する。好ましくは、これらの核酸塩基を含む複数の実質的に連続するヌクレオチド位置は、末端から離れた中心位置におけるオリゴヌクレオチド内に位置されている。

【0172】

アダプター102の一次配列組成は、多くの考慮に依存し得る。1つの考慮は、大規模平行配列決定実験のために使用されるNGSプラットフォームである。例えば、NGS用途のために使用される市販の自動計器は、異なるアダプター102を含有する鑄型103の異なるライブラリーを有するので、任意の所定の市販のNGS計器プラットフォームのための一次配列組成の選択は、その判断基準に依存する。別の考慮は、ブロッカーとしての相補的T_m増強オリゴヌクレオチドの一次配列組成設計である。下記で明らかになる通り、相補的アダプター102の一次配列組成に関する設計決定に影響を及ぼすことができ

るブロッカーのための特定の一次配列組成が好ましい。

【0173】

図2A-Bを参照すると、ブロッカーおよびベイトとしての T_m 増強オリゴヌクレオチドの原理が、典型的なNGS用途について例示されている。二本鎖鑄型203、 T_m 増強オリゴヌクレオチドブロッカー202、ビオチン化オリゴヌクレオチドベイト204およびC₆-1 DNA（登録商標）（示されていない。）は、一緒に混合され、5×の生理食塩水クエン酸ナトリウム緩衝剤（SSC）（または、当業者によく知られている通りの同様のハイブリダイゼーション緩衝剤）の最終濃度を含むように調整された緩衝剤混合物中にて95℃で熱変性され、ベイト：標的ハイブリッドのための予測平均 T_m 値より低いハイブリダイゼーション温度で2時間から3日の間維持される。ハイブリダイゼーション混合物が、95℃変性ステップからハイブリダイゼーションステップに冷却すると、ベイト：標的ハイブリッドが形成する。ライブラリーにおける全ての標的核酸について同じである異なる標的核酸間またはアダプタードメイン間に共通であり得る反復ドメイン間の相互作用を介して互いに結合する相補的配列の領域を有する標的：標的ハイブリッドも形成する。ブロッキング核酸（反復ドメインを結合するためのC₆-1 DNA、およびアダプタードメインを結合するためのオリゴヌクレオチドブロッカー）は、所望されない標的：標的反応と競合するために添加される。しかしながら、非修飾DNAブロッキングオリゴヌクレオチドが用いられるならば、ブロッカーおよびアダプター-アダプター二重鎖の T_m は同一であり、標的：標的二重鎖が形成するのを防止するための唯一のやり方は、大過剰の非修飾ブロッキングオリゴヌクレオチドをハイブリダイゼーション反応に添加することによる質量作用を介する。 T_m 増強修飾ブロッキングオリゴヌクレオチドを使用することは、非修飾ブロッカーの使用より良好に、標的：標的形成を防止する。修飾ブロッカーの T_m は非修飾アダプターより高いので、ブロッカー：アダプター-ハイブリッドは、アダプター：アダプター-ハイブリッドが形成する前に形成し、これによって、「デジ-鎖」の形成を防止する。例えば、非修飾アダプターの T_m が65℃であり、修飾ブロッカーの T_m が75℃であるならば、修飾ブロッカー：標的二重鎖は75℃で形成し、全てのアダプターは、標的：標的二重鎖がアダプター相互作用を介して形成し得る温度である65℃にハイブリダイゼーション混合物が冷却する前にブロックされる。混合物は次いで、203：204ハイブリッドの捕捉を可能にするためのストレプトアビジンを含む固体支持媒体205に添加される。支持媒体／混合物は、連続的によりストリンジェントな条件下（例えば、1×SSC、続いて0.1×SSC）にて、概算されたベイト：標的 T_m 値のものより下、好ましくは、非修飾アダプターの T_m 値のものより上の温度で洗浄される。アダプターが通常ベイトオリゴマーよりずっと短いと仮定して、ベイト T_m は通常、アダプター T_m のものより非常に上である。ブロッカー202は、鑄型上に見出される非修飾アダプターと比較して増強された T_m 値を有するので、鑄型203は、増加したハイブリダイゼーション温度下でブロッカー202に優先的にハイブリダイズし、これによって、異なる鑄型203がこれらのそれぞれのアダプター配列を介してデジ-鎖化凝集体を形成することを最小化する。上昇させたハイブリダイゼーション温度でのストリンジェントな洗浄に続いて、1つの最終のストリンジェントな洗浄が室温で行われ、所望の鑄型が固定化支持体205から回復される。

【0174】

アダプター配列に対応する典型的オリゴヌクレオチドブロッカーは、ハイブリッド捕捉から得られる所望の標的配列の約60%濃縮を提供することができる。対照的に、ブロッカーとしての T_m 増強オリゴヌクレオチドは、ハイブリッド捕捉から得られる所望の標的配列の約80%を超える濃縮を提供することができる。ブロッカーとして T_m 増強オリゴヌクレオチドを用いるハイブリッド捕捉実験からの標的濃縮におけるその結果得られる改善は、ブロッカーとして非修飾オリゴヌクレオチドを用いて得られる収率に対して、所望の鑄型標的の収率の増加が約30%を超える範囲である。

【0175】

ブロッカーおよびベイトとしての T_m 増強オリゴヌクレオチドの設計の各種実施形態が

、ここで記載される。本明細書で使用される場合、「 T_m 増強オリゴヌクレオチド」は、ハイブリダイゼーションパートナーとして同一の核酸塩基組成物を有するオリゴヌクレオチドおよび非修飾基を含む二重鎖核酸に対して、ハイブリダイゼーションパートナーとしてオリゴヌクレオチドを含む二重鎖核酸について熱熔融温度値の増加（「 T_m 値の増強」）を提供する少なくとも1つの修飾基（「 T_m 増強基」）を含むオリゴヌクレオチドである。典型的に、 T_m を増加させる修飾の使用は、標的に対するブロッカーの結合親和性も増加させて、 K_a または会合定数を増加させ、 K_d 、または逆反応についての解離定数を減少させることもできる。

【0176】

多数の T_m 増強基が、 T_m 増強オリゴヌクレオチドの設計において使用され得る。この目的のための適当な T_m 増強基の例としては、例えばロックド核酸（LNA）、二環式核酸（BNA、例えば Isis Pharmaceuticalsからの拘束エチル核酸）、C5修飾ピリミジン塩基（例えば、とりわけ5-メチル-dC、プロピニルピリミジン）を含めて、核酸塩基またはリボース部分への修飾物が挙げられる。とりわけペプチド核酸（PNA）、モルホリノなど、代替りの骨格化学も用いられ得る。マイナーグロブバインダー（MGB）、スペルミン、G-クランプ、またはUaqアントラキノンキャップなど、非塩基モディファイヤーが用いられて、 T_m （または結合親和性）を増加させることができる。結合親和性を増加させるための多くの戦略が当業者に知られており、全てのこうした修飾物の使用は、本発明の範囲内と考えられる。

【0177】

好ましくは、 T_m 増強オリゴヌクレオチドは、複数の T_m 増強基を含む。 T_m 増強基の好ましい数は、1つの相補鎖として T_m 増強オリゴヌクレオチドを含有する二重鎖DNAのために少なくとも約1.4°Cのストリンジェントな条件（ $0.1 \times SSC$ ）（「最適な T_m 値増強」）下で最適な T_m 値における増加を提供するような数である。 T_m 増強オリゴヌクレオチドにおける好ましい数の T_m 増強基は、約2°Cから約25°Cを範囲とする最適な T_m 値増強を提供する。BNAヌクレオチドと末端MGB基との組合せなど、1つ超の型の T_m 増強修飾が、単一修飾ブロッカーにおいて用いられ得る。

【0178】

ハイブリッド捕捉方法における鋳型濃縮の改善のための T_m 増強オリゴヌクレオチドを設計するための好ましい手法は、オリゴヌクレオチドにおいて使用される T_m 増強基に依存する。前述の T_m 増強基の任意のものを含有する T_m 増強オリゴヌクレオチドの T_m 値は、ルーチンの経験的方法を使用して決定され得る。最小のまたは低減された経験的評価を必要とするこれらの後者 T_m 増強基を含有する T_m 増強オリゴヌクレオチドの T_m 値を正確に予測するための信頼できる方法が利用可能であるため、LNAまたはBNAの T_m 増強基の使用が好ましい。この目的のための1つのこうした方法の例は、Behlkeへの、METHODS FOR PREDICTING STABILITY AND MELTING TEMPERATURES OF NUCLEIC ACID DUPLEXESという表題で2012年2月2日に公開された米国特許出願公開第2012/002989号に提供されており、参照により全体を本明細書に組み込む。

【0179】

特定の好ましい実施形態のため、 T_m 増強オリゴヌクレオチドは、バーコード配列タグを含む。バーコード要素は、しばしば、ライブラリー構築中に標的核酸に付着される2つのアダプターオリゴヌクレオチドの1つに含まれる。バーコード要素は典型的に6個の塩基の長さであり；8個の塩基以上の長さなど、より長い要素も用いられ得る。典型的にバーコードアダプターは、NGSライブラリー調製において用いられる2つのアダプターの1つのみを含み、1つのアダプターは「独特であり、コードされており」、1つのアダプターは「ユニバーサル」である。両アダプター上にバーコードを置くことも可能である。バーコードされたアダプターの使用は、多重試料が、単一の多重配列決定実行において一緒に混合および加工されるのを可能にして、著しいコスト節約およびスループットの増加を提供する。配列は、配列決定した後の分析によってデコンボリューションされる。多重

実験は、2、3、4、または最大100個以上までのバーコード修飾アダプター配列の使用に与与することがある。各異なるバーコードアダプターが独特の配列を有すると、最も有効なブロッキングオリゴヌクレオチド（単数または複数）は、セットで存在する各独特のバーコードアダプターに対して完璧に相補的なマッチである配列である。アダプターとブロッカーとの間のバーコードドメイン内ミスマッチが T_m を低下させるため、この手法は、ブロッカーについての最も高い可能な T_m を確保する。そのため、例えば、4重反応における4つのバーコードアダプターの使用は、4つのバーコードアダプターのための4個の独特の配列および共通のユニバーサルアダプターのための1個の独特の配列を含む5つの明確に異なるブロッキングオリゴヌクレオチドの使用を必要とする。しかしながら、多くの明確に異なるバーコードアダプターが用いられるならば、この手法は、高レベル多重実験のための100以上もの多くの独特のブロッキングオリゴヌクレオチドの使用を必要とすることがあり、この手法は費用効果が高くない。さらに、アダプター「B」へのブロッカー「A」のミスハイブリダイゼーションが発生しそうであり、ブロッキングオリゴヌクレオチドの結合親和性を低下させ、ブロッキングステップの有効性を減少する。1つの解決法は、アダプターにおけるバーコードドメインをスパンするためにアダプターオリゴヌクレオチド内の適当な位置にランダムN-merドメイン（例えば、NNNNNN混合塩基6量体配列）を含むブロッキングオリゴヌクレオチドに「ユニバーサル」ドメインを組み込むことである。この手法を用いて、単一のブロッキングオリゴヌクレオチドは、多数のバーコード化アダプターとともに使用され得る。6塩基N-merドメインを使用して、4096個の異なる配列は、ブロッキングオリゴヌクレオチドプール中に存在する。存在するこの多数のバーコードを有することは、大部分のブロッカー：アダプター対をもたらして、バーコードドメインにおけるミスマッチを含む。代替として、「ユニバーサル塩基」が、N塩基の代わりに用いられ得る。ユニバーサル塩基は、一部または全ての天然塩基にハイブリダイズする修飾核酸塩基であり、G：A対またはT：T対など本当の塩基ミスマッチよりミスマッチの熱力学的コストが少ない。当業者によく知られているイノシン（「I」）、5-ニトロインドール（「5-NI」）など、多くのユニバーサル塩基が存在する。イノシンドメイン（IIIIII）とバーコードとの対合は、平均して、完全にミスマッチのN-merドメイン（NNNNNN）より高い T_m を有する。そのため、以下の3つの手法が使用されて、バーコードアダプターのためのブロッキングオリゴヌクレオチドを作製することができる：1）各アダプターに完璧にマッチする一連のブロッカーを合成する、2）アダプターのバーコードドメインと対合するためのN-merドメインを有する単一のブロッカーを合成する、または3）アダプターのバーコードドメインと対合するためのユニバーサル塩基ドメインを有する単一のブロッカーを合成する。前述の方法を用いて、混合されているまたはユニバーサルな核酸塩基の配列に由来する配列寄与を省略することによって、バーコードアダプターを有するLNA基またはBNA基を含有する特別な T_m 増強ブロッキングオリゴヌクレオチドについての T_m 値の十分に正確な概算を算出することができる。こうしたオリゴヌクレオチドについての正確な T_m 値が次いで、ルーチンの経験的方法を使用して、より高精度に決定され得る。

【0180】

前に記述されている通り、アダプター102は、鋳型103上の2つの相補鎖として存在する。ハイブリッド捕捉の二本鎖化鋳型103の集団の変性に続いて、各一本鎖鋳型103は、アダプター102の対応する一本鎖コピーを含む。多くの非関連の鋳型103のデイズ鎖凝集をもたらす異なる一本鎖鋳型103の間での相互作用を防止するため、2つのアダプター102鎖の1つのみが、別の相補鎖アダプター102とのハイブリダイゼーションのためにブロックされる必要がある。この理由のため、および好ましい実施形態において、ブロッカーとしてただ1つの T_m 増強オリゴヌクレオチド鎖が、NGS鋳型103を用いるハイブリッド捕捉方法において鋳型濃縮の改善を達成するために含められる必要がある。

【0181】

ブロッカーとしての T_m 増強オリゴヌクレオチドの一次配列の設計は、オリゴヌクレオ

チドアダプター102の2つの相補鎖の1つの一次配列に基づく。T_m増強基として、利用可能な核酸塩基のいずれかまたは全てをT_m増強オリゴヌクレオチドに含めることができるが、修飾核酸塩基のただ1つの単一型または修飾核酸塩基の2つの異なる型を含むことが好ましい。

【0182】

T_m増強核酸塩基で修飾されたオリゴヌクレオチドは、ヘアピンまたは自己二量体形成のリスクが増加する。オリゴヌクレオチド設計アルゴリズムまたはカルキュレーターが使用されて、配列のヘアピンおよび二量体の潜在性をモデル化することができ、使用されてスクリーン修飾パターンに役立つはずである。例えば、IDTウェブサイト：<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/>上で公的に利用可能であるOligoAnalyzerを参照されたい。この課題は、LNA修飾またはBNA修飾が用いられるならば特に重要である。LNA：DNA塩基対は、DNA：DNA対に対してT_m増加を示す。LNA：LNA塩基対は、LNA：DNA対に対してT_m増加を示す。ヘアピン事象または自己二量体事象において発生する任意のLNA：LNA対は特にこのましい（LNA：DNA対だけがブロッカーと標的との間に形成することができることに留意されたい。）。そのため、T_m増強オリゴヌクレオチドの設計に注意して、LNA：LNA対合事象を介する自己二量体またはヘアピンの形成を促進するパターンを回避しなければならない。これはBNA修飾にも同等に当てはまる。

【0183】

この問題を防止するための1つの好ましい手法は、修飾核酸塩基の単一型だけを用いることである。例えば、T_m増強ブロッキングオリゴヌクレオチドは、LNA-CまたはBNA-Cを使用することだけで作製され得る。塩基組成に依存して、単一塩基型の完全な置換は、最適な性能を提供するのに十分高いT_m増加を達成しない恐れがある。この場合において、LNA-CとLNA-A、またはBNA-CとBNA-Aなど、2つの異なる修飾核酸塩基が用いられ得る。一般に、修飾されたCは、修飾されたAまたは修飾されたTとともに使用され得るが、修飾されたGとともに使用され得ない。同様に、修飾されたAは、修飾されたCまたは修飾されたGとともに使用され得るが、修飾されたTとともに使用され得ない。修飾されたGとの修飾されたCの使用または修飾されたTとの修飾されたAの使用は、回避されるべきである。この戦略は、修飾された塩基間の潜在的相互作用を制限することによって、ヘアピン／二量体形成増加のリスクを制限する。プロピニルピリミジン修飾は、pdU塩基およびpdC塩基としてだけ利用可能である。この場合において、修飾ブロッキングオリゴヌクレオチドは、1個または多くのpdC塩基を含むことができる。代替として、修飾ブロッキングオリゴヌクレオチドは、pdC塩基およびpdU塩基の混合物を含み、前に樹立された設計基準に応じ得る。

【0184】

T_m増強オリゴヌクレオチドのため、T_m増強基の好ましい数は、オリゴヌクレオチドの組成の約2%から約50%まで変動することができる。一般に、ブロッカーとして役立つオリゴヌクレオチドは、アダプター102において使用されるオリゴヌクレオチドの2つの相補鎖の1つと同じ長さ（例えば、約15から約75のヌクレオチド長）を有する。例えば、T_m増強基の好ましい数は、50個のヌクレオチドを有するブロッカーとしてのT_m増強オリゴヌクレオチドには1から約25の範囲であり得る。修飾された残基のより高い部分の使用は、ますますT_mを増加させ、その試薬の「ブロッキング力」に増加改善を加える。しかしながら、修飾された残基の付加は、合成オリゴヌクレオチドのコストを増加させ、自己二量体およびヘアピンの形成のリスクを増加させるので、こうした基の公正な使用が推奨される。NGS用途の大多数において、ブロッカーとしてT_m増強オリゴヌクレオチドだけが使用されて、大規模平行配列決定実験において標的濃縮における所望の改善を達成する。

【0185】

ベイトとしてのT_m増強オリゴヌクレオチドのため、T_m増強基の好ましい数は、プロ

ッカーとしてのオリゴヌクレオチドの記載されている通りの百分率と同じ範囲内に入る。非修飾ベイトとして役立つオリゴヌクレオチドは、約60から約200のヌクレオチド長のサイズにおける範囲であり、ここで、最も共通して使用されるベイトの長さは、約120ヌクレオチド長である。しかしながら、ベイトとしてのオリゴヌクレオチドにT_m増強基を含めることによって、約20から約100のヌクレオチド長の範囲である、より短いベイトを使用することができる。NGS用途における特定の大規模平行配列決定実験のため、何百ものオリゴヌクレオチドの集団がベイトとして使用される。特定の用途において必要とされるベイトの数に非常に依存して、その集団内の各ベイト候補のための、より短いT_m増強オリゴヌクレオチドの使用は、ベイトとして非修飾オリゴヌクレオチドを使用することに対して経済的利点を提供することができる。

10

【0186】

さらなる利点は、品質管理手順を用いる標準化生成物規格に従って構造および活性が立証されたベイトの使用によって与えられる。他の手順は、ベイトを調製するのに利用可能であるが、複数のベイトを含む組成物（すなわち、別個のベイトオリゴヌクレオチドのセット）を捕捉試薬として調製することが好ましく、ここで、複数のベイトの各メンバーは、個別に調製される。

【0187】

この文脈において使用される場合、複数のベイトオリゴヌクレオチドのメンバーの数としては、約10から1,000,000の範囲が挙げられる。用途に依存して、この範囲が好ましい。例えば、「スパイクイン」は、既存のアレイ合成セットを再平衡する必要があり、約10-100個のベイトオリゴヌクレオチドほども低いことがあり；完全な焦点化オリゴセットは、500個から25,000個のベイトオリゴヌクレオチドを範囲としてよい；全体のエキソームセットは、約600,000個のベイトオリゴヌクレオチドを含むことができる。

20

【0188】

より好ましくは、複数のベイトの各メンバーは、化学的プロセスによって個別に合成され、ここで、生成物の品質は、合成中および精製後にモニタリングされ得る。いっそう好ましくは、複数の各メンバーは、合成化学プロセスによって調製され、精製され、ここで、合成および精製の両品質は、独立して判定され得る。最も好ましくは、複数のベイトの各メンバーは、複数のベイトが得ることができるように複数のベイトの他のメンバーからの非依存性生成物規格を有し、ここで、各メンバーの構造および活性は、複数のベイト内の他のメンバーに対して標準化される。標準化された活性を有する複数のベイトの使用は、特に高GC含有量領域を有する標的のために、目的の所定の標的の、より完全および均一なカバレッジを可能にする。これらの利点は、全ての型のオリゴヌクレオチドベイト、すなわち、非修飾オリゴヌクレオチドベイトおよびT_m増強オリゴヌクレオチドベイトのために実現され得る。

30

【0189】

T_m増強オリゴヌクレオチドは、内部または末端の修飾など、追加の特色を含むことができる。ブロッカーとして役立つT_m増強オリゴヌクレオチドのため、ハイブリッド捕捉に続く所望のNGS鑄型の回復は、典型的に、ブロッカーの同時精製をもたらし得る。ブロッカーは、PCR増幅の後続のステップとして鑄型の集団から実質的に希釈され、配列決定が進行する。なお、これらの後続のステップ中におけるプライマーとしてのブロッカーの参加を制限することが望ましい。この理由のため、T_m増強オリゴヌクレオチドは、DNA合成のプライマーとして役立つためのブロッカーの利用能を妨げる3'末端基（例えば、とりわけ3'-dC；2'，3'-ddC；逆位dT；3'-スパーサーC3、）を含むことができる。

40

【0190】

ベイトとして役立つオリゴヌクレオチドは、ハイブリッド捕捉中において鑄型103の集団からの所望の鑄型：ベイトハイブリッドの選択を可能にする少なくとも1つの修飾を含む。好ましい修飾の1つの例としては、化学合成中にオリゴヌクレオチドベイトに組み

50

込まれるとともにハイブリッド選択のためのアビジンまたはストレプトアビジンを含有する固体支持媒体とともに使用され得るビオチンが挙げられる。当業者によく知られている通りのジゴキシゲニンまたは他の基など、他の捕捉リガンドが用いられ得る。

【0191】

ブロッカーとしての T_m 増強オリゴヌクレオチドの好ましい例としては、配列番号2、3、4、5、6、7、8、10、11、12、13、14、15、16、18、19、21、22、24、25、27、28、30、32、34および36が挙げられる。これらの特別な配列、これらの組成物、および大規模平行配列決定用途における使用の方法は、より詳細に実施例に記載されている。

【0192】

遺伝子または遺伝子産物の選択

選択された遺伝子または遺伝子産物（本明細書において「標的遺伝子または遺伝子産物」とも称される。）は、遺伝子内領域または遺伝子間領域を含むサブゲノム間隔を含むことができる。例えば、サブゲノム間隔は、エクソンまたはイントロン、またはこれらの断片、典型的にエクソン配列またはこの断片を含むことができる。サブゲノム間隔は、コード領域または非コード領域、例えばプロモーター、エンハンサー、5'非翻訳領域（5' UTR）、もしくは3'非翻訳領域（3' UTR）、またはこれらの断片を含むことができる。他の実施形態において、サブゲノム間隔は、cDNAまたはこの断片を含む。他の実施形態において、サブゲノム間隔は、例えば本明細書において記載されている通りのSNPを含む。

【0193】

他の実施形態において、サブゲノム間隔、例えば、本明細書において記載されている通りのサブゲノム間隔の1つ以上は、ゲノムにおける実質的に全てのエクソン（例えば、目的の選択された遺伝子または遺伝子産物（例えば、本明細書において記載されている通りの癌表現型と関連する遺伝子または遺伝子産物）からのエクソン）を含む。一実施形態において、サブゲノム間隔は、体細胞変異、生殖細胞系列変異、または両方を含む。一実施形態において、サブゲノム間隔は、改変、例えば点変異もしくは単一変異、欠失変異（例えば、インフレーム欠失、遺伝子内欠失、全遺伝子欠失）、挿入変異（例えば、遺伝子内挿入）、逆位変異（例えば、染色体内逆位）、連結変異、連結挿入変異、逆位重複変異、タンDEM重複（例えば、染色体内タンDEM重複）、転座（例えば、染色体転座、非相互転座）、再編成、遺伝子コピー数における変化、またはこれらの組合せを含む。特定の実施形態において、サブゲノム間隔は、試料における腫瘍細胞のゲノムのコード領域の5%、1%、0.5%、0.1%、0.01%、0.001%未満を構成する。他の実施形態において、サブゲノム間隔は疾患に関与せず、例えば、本明細書において記載されている通りの癌表現型を伴わない。

【0194】

一実施形態において、標的の遺伝子または遺伝子産物は、バイオマーカーである。本明細書で使用される場合、「バイオマーカー」または「マーカー」は、遺伝子、mRNA、または改変され得るタンパク質であり、ここで、前記改変は癌を伴う。改変は、正常または健常の組織または細胞（例えば、対照）におけるその量、構造および／または活性と比較して、癌組織または癌細胞における量、構造および／または活性における改変であってよく、癌などの疾患状態を伴う。例えば、癌を伴うまたは抗癌治療に対する応答性の前兆となるマーカーは、正常の健常組織または細胞と比較して、癌組織または癌細胞において、改変されたヌクレオチド配列、アミノ酸配列、染色体転座、染色体内逆位、コピー数、発現レベル、タンパク質レベル、タンパク質活性、またはメチル化状態を有し得る。さらに、「マーカー」としては、癌などの疾患状態を伴う組織または細胞中に存在する場合に、例えば置換、欠失または挿入によって、構造が改変されている、例えば変異されている（変異を含有する。）、例えばヌクレオチドまたはアミノ酸のレベルで野生型配列と異なる分子が挙げられる。

【0195】

一実施形態において、標的の遺伝子または遺伝子産物としては、単一ヌクレオチド多型（SNP）が挙げられる。別の実施形態において、遺伝子または遺伝子産物は、小さい欠失、例えば小さい遺伝子内欠失（例えば、インフレームまたはフレームシフトの欠失）を有する。なお別の実施形態において、標的配列は、全遺伝子の欠失に由来する。また別の実施形態において、標的配列は、小さい挿入、例えば小さい遺伝子内挿入を有する。一実施形態において、標的配列は、逆位、例えば染色体内逆位に由来する。別の実施形態において、標的配列は、染色体間転座に由来する。なお別の実施形態において、標的配列は、タンDEM重複を有する。一実施形態において、標的配列は、望ましくない特色（例えば、高GC含有量または反復要素）を有する。別の実施形態において、標的配列は、例えばその反復性質のためにそれ自体首尾よく標的化され得ないヌクレオチド配列の一部を有する。一実施形態において、標的配列は、代替のスプライシングに由来する。別の実施形態において、標的配列は、表1、1A、2、3または4による遺伝子または遺伝子産物、またはこれらの断片から選択される。

【0196】

癌としては、これらに限定されないが、B細胞癌、例えば、多発性骨髄腫、メラノーマ、乳癌、肺癌（非小細胞肺癌腫またはNSCLCなど）、気管支癌、直腸結腸癌、前立腺癌、膀胱癌、胃癌、卵巣癌、尿管癌、脳癌または中枢神経系癌、末梢神経系癌、食道癌、子宮頸癌、子宮癌または子宮内膜癌、口腔または咽頭の癌、肝臓癌、腎臓癌、精巣癌、胆管癌、小腸癌または虫垂癌、唾液腺癌、甲状腺癌、副腎癌、骨肉腫、軟骨肉腫、血液組織の癌、腺癌腫、炎症性筋線維芽細胞腫瘍、消化管ストローマ腫瘍（GIST）、大腸癌、多発性骨髄腫（MM）、骨髄異形成症候群（MDS）、骨髄増殖性疾患（MPD）、急性リンパ性白血病（ALL）、急性骨髄球性白血病（AML）、慢性骨髄球性白血病（CLL）、慢性リンパ性白血病（CLL）、真性多血症、ホジキンリンパ腫、非ホジキンリンパ腫（NHL）、軟組織肉腫、線維肉腫、粘液肉腫、脂質肉腫、骨原性肉腫、脊索腫、血管肉腫、内皮肉腫、リンパ管肉腫、リンパ管内皮肉腫、滑膜腫、中皮腫、ユーイング腫瘍、平滑筋肉腫、横紋筋肉腫、扁平細胞癌腫、基底細胞癌腫、腺癌腫、汗腺癌腫、皮脂腺癌腫、乳頭状癌腫、乳頭状腺癌腫、髄様癌腫、気管支原性肺癌腫、腎細胞癌腫、肝細胞癌腫、胆管癌腫、絨毛癌腫、精上皮腫、胚性癌腫、ウィルムス腫瘍、膀胱癌腫、上皮癌腫、神経膠腫、星状細胞腫、髄芽腫、頭蓋咽頭腫、上衣細胞腫、松果体腫、血管細胞芽腫、聴神経腫瘍、乏突起神経膠腫、髄膜腫、神経芽細胞腫、網膜芽細胞腫、濾胞性リンパ腫、びまん性大細胞型B細胞リンパ腫、マントル細胞リンパ腫、肝細胞癌腫、甲状腺癌、胃癌、頭頸部癌、小細胞癌、本態性血小板血症、特発性骨髄化生、過好酸性症候群、全身性肥満細胞症、家族性過好酸球増加症、慢性好酸球性白血病、神経内分泌癌、およびカルチノイド腫瘍などが挙げられる。

【0197】

一実施形態において、標的の遺伝子または遺伝子産物は、ABCB1、ABCC2、ABCC4、ABCG2、ABL1、ABL2、AKT1、AKT2、AKT3、ALK、APC、AR、ARAF、ARFRP1、ARID1A、ATM、ATR、AURKA、AURKB、BCL2、BCL2A1、BCL2L1、BCL2L2、BCL6、BRCA1、BRCA2、C1orf144、CARD11、CBL、CCND1、CCND2、CCND3、CCNE1、CDH1、CDH2、CDH20、CDH5、CDK4、CDK6、CDK8、CDKN2A、CDKN2B、CDKN2C、CEBPA、CHEK1、CHEK2、CRKL、CRLF2、CTNNB1、CYP1B1、CYP2C19、CYP2C8、CYP2D6、CYP3A4、CYP3A5、DNMT3A、DOT1L、DPYD、EGFR、EPHA3、EPHA5、EPHA6、EPHA7、EPHB1、EPHB4、EPHB6、ERBB2、ERBB3、ERBB4、ERCC2、ERG、ESR1、ESR2、ETV1、ETV4、ETV5、ETV6、EWSR1、EZH2、FANCA、FBXW7、FCGR3A、FGFR1、FGFR2、FGFR3、FGFR4、FLT1、FLT3、FLT4、FOXP4、GATA1、GNA11、GNAQ、GNAS、GPR124、GSTP1、GUCY1A2、HOXA

3、HRAS、HSP90AA1、IDH1、IDH2、IGF1R、IGF2R、IKBKE、IKZF1、INHBA、IRS2、ITPA、JAK1、JAK2、JAK3、JUN、KDR、KIT、KRAS、LRP1B、LRP2、LTK、MAN1B1、MAP2K1、MAP2K2、MAP2K4、MCL1、MDM2、MDM4、MEN1、MET、MITF、MLH1、MLL、MPL、MRE11A、MSH2、MSH6、MTHFR、MTOR、MUTYH、MYC、MYCL1、MYCN、NF1、NF2、NKX2-1、NOTCH1、NPM1、NQO1、NRAS、NRP2、NTRK1、NTRK3、PAK3、PAX5、PDGFRA、PDGFRB、PIK3CA、PIK3R1、PKHD1、PLCG1、PRKDC、PTCH1、PTEN、PTPN11、PTPRD、RAF1、RARA、RB1、RET、RICTOR、RPTOR、RUNX1、SLC19A1、SLC22A2、SLCO1B3、SMAD2、SMAD3、SMAD4、SMARCA4、SMARCB1、SMO、SOD2、SOX10、SOX2、SRC、STK11、SULT1A1、TBX22、TET2、TGFB2、TMPRSS2、TOP1、TP53、TPMT、TSC1、TSC2、TYMS、UGT1A1、UMPS、USP9X、VHL、およびWT1からなる群から選択される全長またはこの断片である。

【0198】

一実施形態において、標的の遺伝子もしくは遺伝子産物、またはこれらの断片は、薬理遺伝学および薬理ゲノミクス（PGx）、例えば薬物代謝および毒性に関連している1つ以上のSNPを有する。例示的遺伝子または遺伝子産物としては、これらに限定されないが、ABCB1、ABCC2、ABCC4、ABCG2、C1orf144、CYP1B1、CYP2C19、CYP2C8、CYP2D6、CYP3A4、CYP3A5、DPYD、ERCC2、ESR2、FCGR3A、GSTP1、ITPA、LRP2、MAN1B1、MTHFR、NQO1、NRP2、SLC19A1、SLC22A2、SLCO1B3、SOD2、SULT1A1、TPMT、TYMS、UGT1A1、およびUMPSが挙げられる。

【0199】

別の実施形態において、標的の遺伝子もしくは遺伝子産物、またはこれらの断片は、癌と関連する1つ以上のコドンを含む。例示的遺伝子または遺伝子産物としては、これらに限定されないが、ABL1（例えば、コドン315）、AKT1、ALK、APC（例えば、コドン1114、1338、1450および1556）、AR、BRAF（例えば、コドン600）、CDKN2A、CEBPA、CTNNB1（例えば、コドン32、33、34、37、41および45）、EGFR（例えば、719、746-750、768、790、858および861）、ERBB2、ESR1、FGFR1、FGFR2、FGFR3、FLT3（例えば、コドン835）、HRAS（例えば、コドン12、13および61）、JAK2（例えば、コドン617）、KIT（例えば、コドン816）、KRAS（例えば、コドン12、13および61）、MET、MLL、MYC、NF1、ノッチ1、NPM1、NRAS、PDGFRA、PIK3CA（例えば、コドン88、542、545、546、1047および1049）、PTEN（例えば、コドン130、173、233および267）、RB1、RET（例えば、コドン918）、TP53（例えば、175、245、248、273および306）が挙げられる。

【0200】

なお別の実施形態において、標的の遺伝子もしくは遺伝子産物またはこれらの断片は、癌と関連する。例示的遺伝子または遺伝子産物としては、これらに限定されないが、ABL2、AKT2、AKT3、ARAF、ARFRP1、ARID1A、ATM、ATR、AURKA、AURKB、BCL2、BCL2A1、BCL2L1、BCL2L2、BCL6、BRCA1、BRCA2、CARD11、CBL、CCND1、CCND2、CCND3、CCNE1、CDH1、CDH2、CDH20、CDH5、CDK4、CDK6、CDK8、CDKN2B、CDKN2C、CHEK1、CHEK2、CRKL、CRLF2、DNMT3A、DOT1L、EPHA3、EPHA5、EPHA6、EPHA7、

10

20

30

40

50

EPHB1、EPHB4、EPHB6、ERBB3、ERBB4、ERG、ETV1、ETV4、ETV5、ETV6、EWSR1、EZH2、FANCA、FBXW7、FGFR4、FLT1、FLT4、FOXP4、GATA1、GNA11、GNAQ、GNAS、GPR124、GUCY1A2、HOXA3、HSP90AA1、IDH1、IDH2、IGF1R、IGF2R、IKBKE、IKZF1、INHBA、IRS2、JAK1、JAK3、JUN、KDR、LRP1B、LTK、MAP2K1、MAP2K2、MAP2K4、MCL1、MDM2、MDM4、MEN1、MITF、MLH1、MPL、MRE11A、MSH2、MSH6、MTOR、MUTYH、MYCL1、MYCN、NF2、NKX2-1、NTRK1、NTRK3、PAK3、PAX5、PDGFRB、PIK3R1、PKHD1、PLCG1、PRKDC、PTCH1、PTPN11、PTPRD、RAF1、RARA、RICTOR、RPTOR、RUNX1、SMAD2、SMAD3、SMAD4、SMARCA4、SMARCB1、SMO、SOX10、SOX2、SRC、STK11、TBX22、TET2、TGFB2、TMPRSS2、TOP1、TSC1、TSC2、USP9X、VHL、およびWT1が挙げられる。

10

【0201】

前述の方法の用途は、医学標本において配列決定するための特別な遺伝子または遺伝子の全ての公知配列バリエーション（またはこのサブセット）を含有するオリゴヌクレオチドのライブラリーを使用することを含む。

【0202】

核酸試料

20

様々な組織試料が、本方法において使用される核酸試料の供給源であり得る。ゲノムまたはサブゲノムの核酸（例えば、DNAまたはRNA）は、対象の試料（例えば、腫瘍試料、正常の隣接組織（NAT）、血液試料、循環腫瘍細胞（CTC）、または任意の正常の対照を含有する試料）から単離され得る。特定の実施形態において、組織試料は、凍結試料としてまたはホルムアルデヒドもしくはパラホルムアルデヒド固定パラフィン包埋（FFPE）組織調製物として保存される。例えば、試料は、マトリックス、例えばFFPEブロックまたは凍結試料中に包埋され得る。単離ステップは、個別の染色体をフロー選別すること；および／または対象の試料（例えば、腫瘍試料、NAT、血液試料）を顕微解剖することを含むことができる。

【0203】

30

「単離」核酸分子は、核酸分子の天然供給源中に存在する他の核酸分子から分離されるものである。特定の実施形態において、「単離」核酸分子は、核酸が誘導される生物体のゲノムDNAにおける核酸（すなわち、核酸の5'および3'末端に位置されている配列）に天然にフランクする配列（タンパク質エンコード配列など）がない。例えば、各種実施形態において、単離核酸分子は、核酸が誘導される細胞のゲノムDNAにおける核酸分子に天然にフランクするヌクレオチド配列を約5kB未満、約4kB未満、約3kB未満、約2kB未満、約1kB未満、約0.5kB未満または約0.1kB未満含有することができる。さらに、cDNA分子などの「単離」核酸分子は、組換え技法によって生成される場合に他の細胞性物質または培養培地が実質的にないことがあり、または化学合成される場合に化学的前駆体または他の化学薬品が実質的にないことがある。

40

【0204】

「他の細胞性物質または培養培地が実質的にない」という言語は、分子が単離または組換え生成される細胞の細胞構成成分から分子が分離される核酸分子の調製物を含む。したがって、細胞性物質が実質的にない核酸分子は、他の細胞性物質または培養培地を約30%未満、約20%未満、約10%未満または約5%未満（乾燥重量による）有する核酸分子の調製物を含む。

【0205】

特定の実施形態において、核酸は、熟成試料、例えば熟成FFPE試料から単離される。熟成試料は、例えば、年齢、例えば、1年、2年、3年、4年、5年、10年、15年、20年、25年、50年、75年、または100年齢、またはこれより古くてよい。

50

【0206】

核酸試料は、様々なサイズの組織試料（例えば、生検試料またはFFPE試料）から得られてよい。例えば、核酸は、5から200 μ m、またはこれ以上の組織試料から単離され得る。例えば、組織試料は、5 μ m、10 μ m、20 μ m、30 μ m、40 μ m、50 μ m、70 μ m、100 μ m、110 μ m、120 μ m、150 μ mもしくは200 μ m、またはこれ以上の大きさであり得る。

【0207】

組織試料からのDNA単離のためのプロトコールは、実施例1において提供されている。ホルムアルデヒド固定またはパラホルムアルデヒド固定のパラフィン包埋（FFPE）組織から核酸（例えば、DNA）を単離するための追加の方法は、例えばCronin M.ら、(2004) Am J Pathol. 164 (1) : 35-42 ; Masuda N.ら、(1999) Nucleic Acids Res. 27 (22) : 4436-4443 ; Specht K.ら、(2001) Am J Pathol. 158 (2) : 419-429、Ambion RecoverAll (商標) Total Nucleic Acid Isolation Protocol (Ambion、カタログ番号AM1975、2008年9月)、Maxwell (登録商標) 16 FFPE Plus LEV DNA Purification Kit Technical Manual (Promega Literature 番号TM349、2011年2月)、E. Z. N. A. (登録商標) FFPE DNA Kit Handbook (OMEGA bio-tek、Norcross、GA、生成物番号D3399-00、D3399-01、およびD3399-02 ; 2009年6月)、およびQIAamp (登録商標) DNA FFPE Tissue Handbook (Qiagen、カタログ番号37625、2007年10月) に開示されている。RecoverAll (商標) Total Nucleic Acid Isolation Kitは、パラフィン包埋試料を可溶化するために昇温でキシレンを使用し、核酸を捕捉するためにガラス繊維フィルターを使用する。Maxwell (登録商標) 16 FFPE Plus LEV DNA Purification Kitは、FFPE組織の1 μ mから10 μ mのセクションのゲノムDNAの精製のためにMaxwell (登録商標) 16 Instrumentとともに使用される。DNAは、シリカクラッド常磁性粒子（PMP）を使用して精製され、低溶出体積で溶出される。E. Z. N. A. (登録商標) FFPE DNA Kitは、ゲノムDNAの単離のためにスピンカラムおよび緩衝剤系を使用する。QIAamp (登録商標) DNA FFPE Tissue Kitは、ゲノムおよびミトコンドリアのDNAの精製のためにQIAamp (登録商標) DNA Micro技術を使用する。血液からのDNA単離のためのプロトコールは、例えばMaxwell (登録商標) 16 LEV Blood DNA Kit およびMaxwell 16 Buccal Swab LEV DNA Purification Kit Technical Manual (Promega Literature 番号TM333、2011年1月1日) に開示されている。

【0208】

RNA単離のためのプロトコールは、例えばMaxwell (登録商標) 16 Total RNA Purification Kit Technical Bulletin (Promega Literature 番号TB351、2009年8月) に開示されている。

【0209】

単離核酸試料（例えば、ゲノムDNA試料）は、ルーチン的な技法を実践することによって、断片化または剪断される。例えば、ゲノムDNAは、物理的剪断方法、酵素的開裂方法、化学的開裂方法、および当業者によく知られている他の方法によって断片化され得る。核酸ライブラリーは、ゲノムの複雑性の全てまたは実質的に全てを含有することができる。この文脈における「実質的に全て」という用語は、手順の初期ステップ中においてゲノム複雑性の一部の望まれない消失が実際にあり得るという可能性を指す。本明細書に記載されている方法は、核酸ライブラリーがゲノムの一部である場合、すなわち、ゲノム

10

20

30

40

50

の複雑性が設計によって低減される場合においても有用である。一部の実施形態において、ゲノムの任意の選択部分は、本明細書に記載されている方法で使用され得る。特定の実施形態において、全エキソームまたはこのサブセットは単離される。

【0210】

本発明において特色とされている方法は、ライブラリー（例えば、本明細書において記載されている通りの核酸ライブラリー）を提供するために核酸試料を単離することをさらに含むことができる。特定の実施形態において、核酸試料は、全ゲノム断片、サブゲノム断片、または両方を含む。単離核酸試料は、核酸ライブラリーを調製するために使用され得る。したがって、一実施形態において、本発明において特色とされている方法は、ライブラリー（例えば、本明細書において記載されている通りの核酸ライブラリー）を提供するために核酸試料を単離することをさらに含む。全ゲノムまたはサブゲノムの断片からライブラリーを単離および調製するためのプロトコールは、当技術分野において知られている（例えば、IlluminaのゲノムDNA試料調製キット）。特定の実施形態において、ゲノムまたはサブゲノムのDNA断片は、対象の試料（例えば、腫瘍試料、正常の隣接組織（NAT）、血液試料または任意の正常の対照）から単離される。一実施形態において、試料（例えば、腫瘍またはNATの試料）は、保存標本である。例えば、試料は、マトリックス、例えばFFPEブロックまたは凍結試料中に包埋されている。特定の実施形態において、単離ステップは、個別の染色体をフロー選別すること；および／または対象の試料（例えば、腫瘍試料、NAT、血液試料）を顕微解剖することを含む。特定の実施形態において、核酸ライブラリーを生成するために使用される核酸試料は、5マイクログラム未満、1マイクログラム未満、または500ng未満、200ng未満、100ng未満、50ng未満、10ng未満、5ng未満または1ng未満である。

【0211】

また他の実施形態において、ライブラリーを生成するために使用される核酸試料は、RNAから誘導されるRNAまたはcDNAを含む。一部の実施形態において、RNAとしては、全細胞RNAが挙げられる。他の実施形態において、特定の豊富なRNA配列（例えば、リボソームRNA）は枯渇されていた。一部の実施形態において、全RNA調製物におけるポリ（A）尾部mRNA画分は濃縮されていた。一部の実施形態において、cDNAは、ランダムプライム化cDNA合成方法によって生成される。他の実施形態において、cDNA合成は、オリゴ（dT）含有オリゴヌクレオチドによるプライミングによって、成熟したmRNAのポリ（A）尾部で惹起される。枯渇、ポリ（A）濃縮、およびcDNA合成のための方法は、当業者によく知られている。

【0212】

方法は、当業者によく知られている特異的または非特異的核酸増幅方法によって核酸試料を増幅することをさらに含むことができる。一部の実施形態、特定の実施形態において、核酸試料は、例えば、ランダムプライム化鎖置換増幅などの全ゲノム増幅方法によって増幅される。

【0213】

他の実施形態において、核酸試料は、物理的または酵素的方法によって断片化または剪断され、サイズ選択（例えば、調製用ゲル電気泳動法による。）および増幅（例えば、PCRによる。）された合成アダプターにライゲートされる。他の実施形態において、核酸の断片化およびアダプターライゲートされた基は、ハイブリッド選択より前の顕在的サイズ選択または増幅なく使用される。

【0214】

他の実施形態において、単離DNA（例えば、ゲノムDNA）は、断片化または剪断される。一部の実施形態において、ライブラリーは、ゲノムの簡約表示または定義部分であるゲノムDNAのサブ部分、例えば他の手段によってサブ部分にされたゲノムDNAのサブ部分などのゲノムDNAを50%未満含む。他の実施形態において、ライブラリーは、全てのゲノムDNAまたは実質的に全てのゲノムDNAを含む。

【0215】

一部の実施形態において、ライブラリーは、ゲノムの簡約表示または定義部分であるゲノムDNAのサブ部分、例えば他の手段によってサブ部分にされたゲノムDNAのサブ部分などのゲノムDNAの50%未満を含む。他の実施形態において、ライブラリーは、全てまたは実質的に全てのゲノムDNAを含む。全ゲノムまたはサブゲノムの断片からライブラリーを単離および調製するためのプロトコールは、当技術分野において知られており（例えば、IlluminaのゲノムDNA試料調製キット）、実施例2A、2Bおよび3として本明細書に記載されている。DNA剪断のための代替法は、例2Bとして本明細書に記載されている。例えば、代替のDNA剪断方法は、より自動化可能および／またはより効果的であり得る（例えば、分解FFPE試料を用いる。）。DNA剪断方法の代替が使用されて、ライブラリー調製中にライゲーションステップを回避することもできる。

10

【0216】

本明細書に記載されている方法は、例えば供給源DNAの量が限定的である場合（例えば、全ゲノム増幅後でさえも）、少量の核酸を使用して行われ得る。一実施形態において、核酸は、約5 μg、4 μg、3 μg、2 μg、1 μg、0.8 μg、0.7 μg、0.6 μg、0.5 μg、または400 ng、300 ng、200 ng、100 ng、50 ng、10 ng、5 ng、1 ng未満、またはこれ以下の核酸試料を含む。例えば、典型的に、50–100 ngのゲノムDNAで開始することができる。しかしながら、ハイブリダイゼーションステップ、例えば溶液ハイブリダイゼーションの前にゲノムDNAを（例えば、PCRを使用して）増幅するならば、それより少ない量で始めることができる。したがって、ハイブリダイゼーション、例えば溶液ハイブリダイゼーションの前にゲノムDNAを増幅することは可能であるが、必要不可欠ではない。

20

【0217】

ライブラリーを生成するために使用される核酸試料は、RNAから誘導されるRNAまたはcDNAも含むことができる。一部の実施形態において、RNAは、全細胞RNAを含む。他の実施形態において、特定の豊富なRNA配列（例えば、リボソームRNA）は、枯渇されていた。他の実施形態において、全RNA調製におけるポリ（A）尾部化mRNA画分は、濃縮されていた。一部の実施形態において、cDNAは、ランダムプライム化cDNA合成方法によって生成される。他の実施形態において、cDNA合成は、オリゴ（dT）含有オリゴヌクレオチドによるプライミングによって、成熟したmRNAsのポリ（A）尾部で惹起される。枯渇、ポリ（A）濃縮、およびcDNA合成のための方法は、当業者によく知られている。

30

【0218】

方法は、当業者に知られている特異的または非特異的核酸増幅方法によって核酸試料を増幅することをさらに含むことができる。核酸試料は、例えば、ランダムプライム化鎖置換増幅などの全ゲノム増幅方法によって、増幅され得る。

【0219】

核酸試料は、本明細書において記載されている通りの物理的または酵素的方法によって断片化または剪断され、サイズ選択（例えば、調製用ゲル電気泳動法による。）および増幅（例えば、PCRによる。）された合成アダプターにライゲートされ得る。核酸の断片化およびアダプターライゲートされた基は、ハイブリッド選択より前の顕在的サイズ選択または増幅なく使用される。

40

【0220】

ライブラリーメンバー

「メンバー」もしくは「ライブラリーメンバー」または他の同様の用語は、本明細書で使用される場合、ライブラリーのメンバー（または「ライブラリーキャッチ」）である核酸分子、例えばDNAまたはRNAを指す。ライブラリーメンバーは、本明細書において記載されている通りの腫瘍メンバー、参照メンバー、またはPGxメンバーの1つ以上であり得る。典型的に、メンバーは、DNA分子、例えばゲノムDNAまたはcDNA分子である。メンバーは、例えば酵素的に断片化されたまたは剪断することによって断片化されたゲノムDNAであり得る。メンバーは、対象からのヌクレオチド配列を含むことがで

50

き、対象から誘導されたのではないヌクレオチド配列、例えばプライマーもしくはアダプター（例えば、PCR増幅用または配列決定用）、または試料の同定を可能にする配列、例えば「バーコード」配列を含むこともできる。

【0221】

本明細書で使用される場合、「標的メンバー」は、核酸ライブラリーから単離するのを所望する核酸分子を指す。一実施形態において、標的メンバーは、本明細書において記載されている通りの腫瘍メンバー、参照メンバー、またはPGxメンバーであり得る。核酸ライブラリーから実際に選択されたメンバーは、本明細書において「ライブラリーキャッチ」と称される。一実施形態において、ライブラリーキャッチは、ライブラリーのメンバーの選択または濃縮、例えば、本明細書において記載されている通りのハイブリッド捕捉の1回以上のラウンド後のライブラリーの濃縮または選択されたアウトプットを含む。

10

【0222】

標的メンバーは、ライブラリーのサブグループであってよく、すなわち、ライブラリーメンバーの全てが、本明細書に記載されているプロセスの任意の特別な使用によって選択されるというわけではない。他の実施形態において、標的メンバーは、所望の標的領域内にある。例えば、標的メンバーは、一部の実施形態において、10%ほども低いまたは95%–98%ほども高い、またはこれより高いライブラリーメンバーの百分率であり得る。一実施形態において、ライブラリーキャッチは、標的メンバーの少なくとも約20%、30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、98%、99%、99.9%、またはこれ以上を含む。別の実施形態において、ライブラリーは、標的メンバーの100%を含有する。一実施形態において、ライブラリーキャッチの純度（アラインメント標的に対する読み取りの百分率）は、少なくとも約20%、30%、40%、50%、60%、70%、75%、80%、85%、90%、95%、98%、99%、99.9%、またはこれ以上である。

20

【0223】

ゲノムDNAから得られる標的メンバー（またはライブラリーキャッチ）は、ゲノムDNAの約0.0001%未満、少なくとも約0.0001%、少なくとも約0.001%、少なくとも約0.01%もしくは少なくとも約0.1%を含むように、全ゲノムDNAの小部分を含むことができ、またはゲノムDNAの少なくとも約1%、2%、3%、4%、5%、6%、7%、8%、9%もしくは10%、もしくはゲノムDNAの10%超を含むように、全ゲノムDNAのより大きな部分を含むことができる。

30

【0224】

一実施形態において、標的メンバー（またはライブラリーキャッチ）は、ゲノムの複合混合物から選択される。例えば、1つの細胞型（例えば、癌細胞）からのDNAの選択は、他の細胞型（例えば、正常細胞）からのDNAを含有する試料からである。こうした用途において、標的メンバーは、複合試料中に存在する核酸配列の全複雑性の0.0001%未満、少なくとも0.0001%、少なくとも約0.001%、少なくとも約0.01%、または少なくとも約0.1%を含むことができ、または錯体試料中に存在する核酸配列の全複雑性の少なくとも約1%、2%、5%、10%、または10%超を含むように、より大きな部分を含むことができる。

40

【0225】

一実施形態において、本明細書に記載されている方法（例えば、溶液ハイブリダイゼーション選択方法）によって選択される標的メンバー（またはライブラリーキャッチ）は、ゲノムエクソンの約0.1%、1%、2%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、または95%超など、ゲノムにおけるエクソンの全てまたは一部を含む。別の実施形態において、標的メンバー（またはライブラリーキャッチ）は、エクソン、例えば少なくとも約100個、200個、300個、400個、500個、600個、700個、800個、900個または1000個の特別なエクソン、例えば、癌などの特別な疾患と関連するエクソンの特異的群であり得る。なお別の実施形態において、標的メンバー（またはライブラリーキャッチ）は、目的の選択遺伝子のエ

50

クソンまたは他の対を含有する。特異的ベイト配列の使用は、実務者が、特別な選択のための核酸の群からの多くのエクソンまたはわずかなエクソン（または他の配列）を含有する標的配列（選択された配列の理想的セット）および核酸のサブグループ（選択された配列の実際のセット）を選択するのを可能にする。

【0226】

一実施形態において、標的メンバー（またはライブラリーキャッチ）は、cDNAのセットを含む。cDNAを捕捉することは、例えばスプライスバリエーションを見出すため、および融合転写物（例えば、ゲノムDNA転座から。）を同定するために使用され得る。別の実施形態において、標的メンバー（およびライブラリーキャッチ）は、例えば腫瘍において、細胞、組織または臓器のRNA画分中に発現される単一塩基の変化および他の配列変化を見出すために使用される。

10

【0227】

標的メンバー（またはライブラリーキャッチ）（例えば、エクソン、cDNAおよび他の配列）は、所望通りに関連または非関連であってよい。例えば、選択標的メンバー（およびライブラリーキャッチ）は、疾患に関与される遺伝子である核酸の群、例えば、癌など1種以上の疾患に関係づけられる遺伝子の群、特異的SNPを含有する核酸の群から得ることができる。

【0228】

一実施形態において、ライブラリーメンバーの一部または全ては、非標的アダプター配列を含む。アダプター配列は、例えば配列決定方法（例えば、NGS方法）に、増幅に、逆転写に、またはベクターへのクローニングに有用であり得る。アダプター配列は、一方または両方の末端に位置され得る。アダプターは、例えば付属の実施例に記載されている通り、ライブラリー挿入の5' - または3' - 3'末端でライゲートされ得る。アダプターは、DNAオリゴについてはNimbleGen (Roche)、Integrated DNA Technologies (IDT)、またはAgilent Technologiesなどの市販供給元から得ることができる。

20

【0229】

アダプターに相補的なブロックングオリゴヌクレオチドは、当技術分野において知られている方法、例えばオリゴ合成の方法によって、設計および調製され得る。ブロックングオリゴヌクレオチドは、DNAオリゴについてはNimbleGen (Roche)、Integrated DNA Technologies (IDT)、またはAgilent Technologiesなどの市販供給元から得ることもできる。これらのアダプターの長さおよび組成は、例えば、当技術分野において知られている方法に従って相補的アダプターとの結合相互作用（例えば、本明細書において記載されている通りの T_m ）を修飾するために調整され得る。

30

【0230】

ブロックングオリゴヌクレオチドは、DNA、RNA、または両方の組合せを含むことができる。DNAまたはRNAオリゴヌクレオチドは、天然または非天然であってよい。特定の実施形態において、ブロックングオリゴヌクレオチドは、例えば溶融温度を増加させるため、1つ以上の非天然ヌクレオチドを含む。例示的非天然オリゴヌクレオチドとしては、修飾されたDNAまたはRNAヌクレオチドが挙げられる。例示的修飾RNAヌクレオチドは、ロックド核酸 (LNA) であり、ここで、LNAヌクレオチドのリボース部分は、2' 酸素および4' 炭素を接続する余分な架橋で修飾されている (Kaur, H; Arora, A; Wengel, J; Maiti, S; Arora, A.; Wengel, J.; Maiti, S. (2006). 「Thermodynamic, Counterion, and Hydration Effects for the Incorporation of Locked Nucleic Acid Nucleotides into DNA Duplexes」。Biochemistry 45 (23): 7347-55)。他の修飾された例示的DNAおよびRNAヌクレオチドとしては、これらに限定されないが、ペプチド結合によって連結されている反復N- (2-アミノエチル)

40

50

ーグリシン単位から構成されているペプチド核酸（PNA）（Egholm, M. ら（1993）Nature 365（6446）：566－8）；低GC領域を捕捉するために修飾されたDNAまたはRNAオリゴヌクレオチド；二環式核酸（BNA）または架橋オリゴヌクレオチド；修飾5－メチルデオキシシチジン；および2，6－ジアミノプリンが挙げられる。他の修飾されたDNAおよびRNAヌクレオチドは、当技術分野において知られている。

【0231】

ベイトの設計および構築

ベイトは、標的核酸にハイブリダイズし（例えば、相補的である。）、これによって、標的核酸の捕捉を可能にすることができる核酸分子、例えばDNAまたはRNA分子であり得る。一実施形態において、ベイトはRNA分子である。他の実施形態において、ベイトは、例えば結合実体に結合することによって、ベイトおよびベイトにハイブリダイズされる核酸により形成されるハイブリッドの捕捉および分離を可能にする結合実体、例えば親和性タグを含む。一実施形態において、ベイトは、溶液相ハイブリダイゼーションに相当である。

【0232】

典型的に、RNA分子は、ベイト配列として使用される。RNA－DNA二重鎖は、DNA－DNA二重鎖より安定であり、そのため、核酸の潜在的により良好な捕捉を提供する。

【0233】

RNAベイトは、これらに限定されないが、DNA依存性RNAポリメラーゼを使用するDNA分子のデノボ化学合成および転写を含めて当技術分野において知られている方法を使用して、本明細書において他所で記載されている通りに作製され得る。一実施形態において、ベイト配列は、PCRなどの公知の核酸増幅方法を使用して、例えば、鑄型としてヒトDNAまたはプールされたヒトDNA試料を使用して生成される。オリゴヌクレオチドは、次いで、RNAベイトに変換され得る。一実施形態において、インビトロ転写は、例えば、オリゴヌクレオチドの1つの末端にRNAポリメラーゼプロモーター配列を付加することに基づいて使用される。一実施形態において、RNAポリメラーゼプロモーター配列は、例えばPCRまたは他の核酸増幅方法を使用してベイト配列を増幅または再増幅することによって、例えばRNAプロモーター配列を用いる各標的特異的プライマー対の1つのプライマーをテーリングすることによって、ベイトの末端に付加される。一実施形態において、RNAポリメラーゼは、T7ポリメラーゼ、SP6ポリメラーゼまたはT3ポリメラーゼである。一実施形態において、RNAベイトは、タグ、例えば親和性タグで標識される。一実施形態において、RNAベイトは、例えばビオチン化UTPを使用するインビトロ転写によって作製される。別の実施形態において、RNAベイトは、ビオチンを用いることなく生成され、次いでビオチンは、ソラレン架橋など当技術分野においてよく知られている方法を使用して、RNA分子に架橋される。一実施形態において、RNAベイトは、RNAse分解に抵抗するRNA分子を生成するために例えば転写中に修飾ヌクレオチドを使用することによって作製され得るRNAse抵抗性RNA分子である。一実施形態において、RNAベイトは、二本鎖化DNA標的の1つの鎖だけに対応する。典型的に、こうしたRNAベイトは、自己相補的でなく、ハイブリダイゼーションドライバースとして、より有効である。

【0234】

ベイトセットは、ベイトが参照配列の標的を選択するのに最適であるように参照配列から設計され得る。一部の実施形態において、ベイト配列は、混合塩基（例えば、縮重）を使用して設計される。例えば、混合塩基（単数または複数）は、共通のSNPまたは変異の位置（単数または複数）でベイト配列に含められることで、対立遺伝子の両方（例えば、SNPおよび非SNP；変異体および非変異体）をキャッチするためのベイト配列を最適化することができる。一部の実施形態において、全ての公知配列変動（またはこのサブセット）は、混合縮重オリゴヌクレオチドを使用することによるよりむしろ、複数のオリ

ゴヌクレオチドベイトを用いて標的化され得る。

【0235】

特定の実施形態において、ベイトセットは、約100ヌクレオチド長から300ヌクレオチド長の間のオリゴヌクレオチド（または複数のオリゴヌクレオチド）を含む。典型的に、ベイトセットは、約130ヌクレオチド長から230ヌクレオチド長、または約150ヌクレオチド長から200ヌクレオチド長の間のオリゴヌクレオチド（または複数のオリゴヌクレオチド）を含む。他の実施形態において、ベイトセットは、約300ヌクレオチド長から1000ヌクレオチド長の間のオリゴヌクレオチド（または複数のオリゴヌクレオチド）を含む。

【0236】

一部の実施形態において、オリゴヌクレオチドにおける標的メンバー特異的配列は、約40ヌクレオチド長から1000ヌクレオチド長、約70ヌクレオチド長から300ヌクレオチド長、約100ヌクレオチド長から200ヌクレオチド長の間、典型的には、約120ヌクレオチド長から170ヌクレオチド長の間である。

【0237】

一部の実施形態において、ベイトセットは、結合実体を含む。結合実体は、各ベイト配列上の親和性タグであり得る。一部の実施形態において、親和性タグは、ビオチン分子またはハプテンである。特定の実施形態において、結合実体は、アビジン分子などのパートナー、またはハプテンに結合する抗体もしくはこの抗原結合断片に結合することによって、ハイブリダイゼーション混合物からベイト／メンバーハイブリッドの分離を可能にする。

【0238】

他の実施形態において、ベイトセットにおけるオリゴヌクレオチドは、同じ標的メンバー配列に対して順方向および逆の相補配列を含有し、これによって、逆相補体メンバー特異的配列を有するオリゴヌクレオチドは、逆相補体ユニバーサル尾部も保有する。これは、同じ鎖であるRNA転写物、すなわち互いに相補的でないRNA転写物に至ることができる。

【0239】

他の実施形態において、ベイトセットは、1つ以上の位置に縮重塩基または混合塩基を含有するオリゴヌクレオチドを含む。また他の実施形態において、ベイトセットは、生物体の単一の種または群集の集団中に存在する複数のまたは実質的に全ての公知配列バリエーションを含む。一実施形態において、ベイトセットは、ヒト集団中に存在する複数のまたは実質的に全ての公知配列バリエーションを含む。

【0240】

他の実施形態において、ベイトセットは、cDNA配列を含む、またはcDNA配列から誘導される。他の実施形態において、ベイトセットは、ゲノムDNA、cDNAまたはクローニングDNAから増幅される増幅生成物（例えば、PCR生成物）を含む。

【0241】

他の実施形態において、ベイトセットは、RNA分子を含む。一部の実施形態において、セットは、これらに限定されないが、より安定であり、RNaseに抵抗性であるものを含めて、化学的酵素的に修飾されたまたはインビトロ転写されたRNA分子を含む。

【0242】

なお他の実施形態において、ベイトは、参照により本明細書に組み込むUS 2010/0029498 およびGnirke, A.ら(2009) Nat Biotechnol. 27(2):182-189に記載されている方法によって生成される。例えば、ビオチン化RNAベイトは、元々マイクロアレイ上で合成された合成の長いオリゴヌクレオチドのプールを得ること、およびオリゴヌクレオチドを増幅してベイト配列を提供することによって生成される。一部の実施形態において、ベイトは、ベイト配列の1つの末端でRNAポリメラーゼプロモーター配列を付加すること、およびRNAポリメラーゼを使用してRNA配列を合成することによって生成される。一実施形態において、合成のオリゴデオキシヌクレオチドのライブラリーは、Agilent Technologies,

10

20

30

40

50

Inc.などの市販供給元から得られ、公知の核酸増幅方法を使用して増幅させることができる。

【0243】

したがって、上述のバイトセットを作製する方法が提供される。方法は、1つ以上の標的特異的バイトオリゴヌクレオチド配列（例えば、本明細書において記載されている通りの1つ以上の変異捕捉、参照または対照オリゴヌクレオチド配列）を選択すること；標的特異的バイトオリゴヌクレオチド配列のプールを得ること（例えば、マイクロアレイ合成によって、標的特異的バイトオリゴヌクレオチド配列のプールを合成すること）；および場合によって、オリゴヌクレオチドを増幅してバイトセットを生成することを含む。

【0244】

他の実施形態において、方法は、1つ以上のビオチン化プライマーを使用するオリゴヌクレオチドを増幅すること（例えば、PCRによる。）をさらに含む。一部の実施形態において、オリゴヌクレオチドは、マイクロアレイに付着された各オリゴヌクレオチドの末端にユニバーサル配列を含む。方法は、オリゴヌクレオチドからユニバーサル配列を除去することをさらに含む。こうした方法は、オリゴヌクレオチドの相補鎖を除去すること、オリゴヌクレオチドをアニーリングすること、およびオリゴヌクレオチドを延長することを含むことができる。これらの実施形態の一部において、オリゴヌクレオチドを増幅する（例えば、PCRによる。）ための方法は、1つ以上のビオチン化プライマーを使用する。一部の実施形態において、方法は、増幅オリゴヌクレオチドをサイズ選択することをさらに含む。

【0245】

一実施形態において、RNAバイトセットが作製される。方法は、本明細書に記載されている方法に従ってバイト配列のセットを生成すること、バイト配列の1つの末端でRNAポリメラーゼプロモーター配列を付加すること、およびRNAポリメラーゼを使用してRNA配列を合成することを含む。RNAポリメラーゼは、T7 RNAポリメラーゼ、SP6 RNAポリメラーゼまたはT3 RNAポリメラーゼから選択され得る。他の実施形態において、RNAポリメラーゼプロモーター配列は、バイト配列を増幅すること（例えば、PCRによる。）によって、バイト配列の末端で付加される。ゲノムまたはcDNAからの特異的プライマー対を用いるPCRによってバイト配列が増幅される実施形態において、各対における2つの特異的プライマーの1つの5'末端にRNAプロモーター配列を付加することは、標準的方法を使用してRNAバイトに転写され得るPCR生成物に至る。

【0246】

他の実施形態において、バイトセットは、鋳型としてヒトDNA試料またはプールされているヒトDNA試料を使用して生成され得る。こうした実施形態において、オリゴヌクレオチドは、ポリメラーゼ連鎖反応（PCR）によって増幅される。他の実施形態において、増幅されたオリゴヌクレオチドは、ローリングサークル増幅または超分岐ローリングサークル増幅によって再増幅される。同方法は、鋳型としてヒトDNA試料またはプールされているヒトDNA試料を使用してバイト配列を生成することにも使用され得る。同方法は、これらに限定されないが、制限消化、パルスフィールドゲル電気泳動、フロー選別、CsCl密度勾配遠心分離、選択的動態学的再会合、染色体調製物の顕微解剖、および当業者に知られている他の部分化方法を含めて、他の方法によって得られるゲノムのサブ部分を使用して、バイト配列を生成するためにも使用され得る。

【0247】

特定の実施形態において、バイトセットにおけるバイトの数は、1,000未満である。他の実施形態において、バイトセットにおけるバイトの数は、1,000超、5,000超、10,000超、20,000超、50,000超、100,000超または500,000超である。

【0248】

バイト配列の長さは、約70ヌクレオチドから1000ヌクレオチドの間であり得る。

一実施形態において、ベイトの長さは、約100から300ヌクレオチド長、110から200ヌクレオチド長、または120から170ヌクレオチド長の間である。上に記述されているものに加えて、約70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、300、400、500、600、700、800および900ヌクレオチド長の中間体オリゴヌクレオチド長さが、本明細書に記載されている方法において使用され得る。一部の実施形態において、約70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220または230塩基長のオリゴヌクレオチドが使用され得る。

【0249】

各ベイト配列は、一方または両方の末端上に標的特異的（例えば、メンバー特異的）ベイト配列およびユニバーサル尾部を含むことができる。本明細書で使用される場合、「ベイト配列」という用語は、標的特異的「ベイト配列」およびオリゴヌクレオチドの他のヌクレオチドを含めて、標的特異的ベイト配列または全オリゴヌクレオチドを指すことができる。ベイトにおける標的特異的配列は、約40ヌクレオチド長から1000ヌクレオチド長の間である。一実施形態において、標的特異的配列は、約70ヌクレオチド長から300ヌクレオチド長の間である。別の実施形態において、標的特異的配列は、約100ヌクレオチド長から200ヌクレオチド長の間である。なお別の実施形態において、標的特異的配列は、約120ヌクレオチド長から170ヌクレオチド長の間、典型的には120ヌクレオチド長である。約40、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、300、400、500、600、700、800および900ヌクレオチド長の標的特異的配列、ならびに上に記述されている長さの間の長さの標的特異的配列など、上に記述されているものに加えて中間の長さも、本明細書に記載されている方法において使用され得る。

【0250】

一実施形態において、ベイトは、約50から200ヌクレオチド長（例えば、約50、60、80、90、100、110、120、130、140、150、160、170、190または200ヌクレオチド長）のオリゴマー（例えば、RNAオリゴマー、DNAオリゴマー、またはこれらの組合せで構成されている。）である。一実施形態において、各ベイトオリゴマーは、標的特異的ベイト配列である、約120から170、または典型的には約120ヌクレオチドを含む。ベイトは、一方または両方の末端に追加の非標的特異的ヌクレオチド配列を含むことができる。追加のヌクレオチド配列は、例えば、PCR増幅のためまたはベイト識別子として使用され得る。特定の実施形態において、ベイトは、本明細書において記載されている通りの結合実体（例えば、ビオチン分子などの捕捉タグ）を追加として含む。結合実体、例えばビオチン分子は、ベイトに、例えばベイトの5'、3'末端または内部に（例えば、ビオチン化ヌクレオチドを組み込むことにより）付着され得る。一実施形態において、ビオチン分子は、ベイトの5'末端で付着される。

【0251】

1つの例示的实施形態において、ベイトは、120ヌクレオチドが標的特異的「ベイト配列」である約150ヌクレオチド長のオリゴヌクレオチドである。この他の30ヌクレオチド（例えば、各末端上の15ヌクレオチド）は、PCR増幅のために使用されるユニバーサルな任意の尾部である。尾部は、ユーザーによって選択される任意の配列であってよい。例えば、合成のオリゴヌクレオチドのプールは、 $N_{120}CACTGCGGCTCCTCA-3'$ （配列番号100）の配列のオリゴヌクレオチドを含むことができる。

【0252】

本明細書に記載されているベイト配列は、エクソンおよび短い標的配列の選択のために使用され得る。一実施形態において、ベイトは、約100ヌクレオチド長から300ヌクレオチド長の間である。別の実施形態において、ベイトは、約130ヌクレオチド長から

230ヌクレオチド長の間である。なお別の実施形態において、ベイトは、約150ヌクレオチド長から200ヌクレオチド長の間である。例えばエクソンおよび短い標的配列の選択のための、ベイトにおける標的特異的配列は、約40ヌクレオチド長から1000ヌクレオチド長の間である。一実施形態において、標的特異的配列は、約70ヌクレオチド長から300ヌクレオチド長の間である。別の実施形態において、標的特異的配列は、約100ヌクレオチド長から200ヌクレオチド長の間である。なお別の実施形態において、標的特異的配列は、約120ヌクレオチド長から170ヌクレオチド長の間である。

【0253】

一部の実施形態において、長いオリゴヌクレオチドは、標的配列を捕捉するのに必要なオリゴヌクレオチドの数を最小化することができる。例えば、エクソン1個当たり1個のオリゴヌクレオチドが使用され得る。ヒトゲノムにおけるタンパク質コードエクソンの平均長さおよび中央長さはそれぞれ約164および120塩基対であることが、当技術分野において知られている。より長いベイトは、より特異的であり、より短いものより良好に捕捉することができる。結果として、オリゴヌクレオチドベイト配列1個当たりの成功率は、短いオリゴヌクレオチドを用いるより高い。一実施形態において、最小ベイトにカバーされる配列は、例えばエクソンサイズ標的を捕捉するため、1つのベイトのサイズ（例えば、120-170塩基）である。ベイト配列の長さを決定する際に、不必要に長いベイトが、標的に直接隣接する望まれないDNAをより多くキャッチすることも考慮に入れることができる。より長いオリゴヌクレオチドベイトは、DNA試料中の標的化領域における多型に対して、より短いものより耐性であることもあり得る。典型的に、ベイト配列は、参照ゲノム配列から誘導される。実際のDNA試料における標的配列が参照配列から外れるならば、例えばそれが単一ヌクレオチド多型（SNP）を含有するならば、ベイトに対して効果的にハイブリダイズできず、そのため、ベイト配列にハイブリダイズされた配列において表示不足または完全に存在しないことがある。例えば120個から170個の塩基における単一の不対合が、それぞれ多重増幅およびマイクロアレイ捕捉における典型的なベイトまたはプライマーの長さである20個または70個の塩基における単一のミスマッチよりもハイブリッド安定性に対する効果がないことがあるという理由のため、SNPによる対立遺伝子の脱落は、より長い合成のベイト分子にはありそうにない。

【0254】

ゲノム領域など、捕捉ベイトの長さと比較して長い標的の選択のため、ベイト配列長さは、典型的に、上に記述されている短い標的のためのベイトと同じサイズ範囲であるが、但し、隣接配列の標的化を最小化するという唯一の目的のためにベイト配列の最大サイズを限定する必要がない場合を除く。代替として、オリゴヌクレオチドは、よりずっと広いウィンドウ（典型的に600個の塩基）にわたってタイトル付けされ得る。この方法は、典型的エクソンよりずっと大きい（例えば、約500個の塩基）DNA断片を捕捉するために使用され得る。結果として、より多くの望まれないフランキング非標的配列が選択される。

【0255】

ベイト合成

ベイトは、任意の型のオリゴヌクレオチド、例えばDNAまたはRNAであり得る。DNAまたはRNAベイト（「オリゴベイト」）は、個別に合成され得る、またはDNAまたはRNAベイトセットとしてのアレイ（「アレイベイト」）において合成され得る。オリゴベイトは、アレイフォーマットで提供されるかまたは単離オリゴとして提供されるかにかかわらず、典型的に一本鎖である。ベイトは、本明細書において記載されている通りの結合実体（例えば、ビオチン分子などの捕捉タグ）を追加として含み得る。結合実体、例えばビオチン分子は、ベイトに、例えばベイトの5'または3'末端で、典型的にはベイトの5'末端で付着され得る。

【0256】

一部の実施形態において、個別のオリゴベイトは、アレイベイトセットに付加され得る。これらの場合において、オリゴベイトは、アレイベイトによって標的化されるものと同

じ区域を標的にするように設計することができ、追加のオリゴペイトは、ゲノムの特定の区域において、増強されたまたはより徹底的なカバレッジを達成するように設計され、標準的アレイペイトに付加され得る。例えば、追加のオリゴペイトは、標準的アレイペイトセットを用いる初期の配列決定ラウンドに続いて不十分な配列決定カバレッジの区域を標的にするように設計され得る。一部の実施形態において、オリゴペイトは、アレイペイトセットのためのカバレッジの区域に対するタイル張り効果、または他のオリゴペイトのためのカバレッジの区域に対するタイル張り効果を有するように設計されている。

【0257】

一実施形態において、個別のオリゴペイトは、RNAもしくはDNAオリゴアレイペイトセットまたはこれらの組合せ（例えば、市販のアレイペイトセット）を補充するために使用されるDNAオリゴである。他の実施形態において、個別のオリゴペイトは、個別に設計および合成されているオリゴの収集物であるRNAもしくはDNAオリゴペイトセットまたはこれらの組合せを補充するために使用されるDNAオリゴである。一実施形態において、個別のオリゴペイトは、RNAもしくはDNAオリゴアレイペイトセットまたはこれらの組合せ（例えば、市販のアレイペイトセット）を補充するために使用されるRNAオリゴである。他の実施形態において、個別のオリゴペイトは、個別に設計および合成されているオリゴの収集物である、RNAもしくはDNAオリゴペイトセットまたはこれらの組合せを補充するために使用されるRNAオリゴである。

10

【0258】

なお別の実施形態において、個別のオリゴペイトは、DNAオリゴアレイペイトセット（例えば、市販のアレイペイトセット）を補充するために使用されるDNAオリゴであり、他の実施形態において、個別のオリゴペイトは、個別に設計および合成されているオリゴの収集物であるDNAオリゴペイトセットを補充するために使用されるDNAオリゴである。

20

【0259】

なお別の実施形態において、個別のオリゴペイトは、RNAオリゴアレイペイトセット（例えば、市販のアレイペイトセット）を補充するために使用されるDNAオリゴであり、他の実施形態において、個別のオリゴペイトは、個別に設計および合成されているオリゴの収集物であるRNAオリゴペイトセットを補充するために使用されるDNAオリゴである。

30

【0260】

なお別の実施形態において、個別のオリゴペイトは、RNAオリゴアレイペイトセット（例えば、市販のアレイペイトセット）を補充するために使用されるRNAオリゴであり、他の実施形態において、個別のオリゴペイトは、個別に設計および合成されているオリゴの収集物であるRNAオリゴペイトセットを補充するために使用されるRNAオリゴである。

【0261】

なお別の実施形態において、個別のオリゴペイトは、DNAオリゴアレイペイトセット（例えば、市販のアレイペイトセット）を補充するために使用されるRNAオリゴであり、他の実施形態において、個別のオリゴペイトは、個別に設計および合成されているオリゴの収集物であるDNAオリゴペイトセットを補充するために使用されるRNAオリゴである。

40

【0262】

一実施形態において、オリゴペイトは、特に目的の遺伝子における配列を標的にするように、例えば、拡大された遺伝子セットの配列決定カバレッジの増加を達成するように設計される。

【0263】

別の実施形態において、オリゴペイトは、ゲノムのサブセットを表す配列を標的にするように設計され、アレイペイトの代わりにまたはアレイペイトに加えて、プールとして混合および使用される。

50

【0264】

一実施形態において、オリゴベイトの第1のセットは、不十分な配列決定カバレッジの区域を標的にするように設計され、オリゴベイトの第2のセットは、特に目的の遺伝子を標的にするように設計される。次いで、オリゴベイトの両セットは組み合わせられ、場合によって、配列決定のために使用されるべき標準的アレイベイトセットと混合される。

【0265】

一実施形態において、オリゴベイトミックスは、例えば、標的化遺伝子パネルを同時に配列決定するため、およびゲノム再編成およびコピー数の改変（アレイ化CGH（比較用ゲノムハイブリダイゼーション）と同等）を探するという目的のために創出される単一ヌクレオチド多型（SNP）のパネルをスクリーンするために使用される。例えば、最初にSNPのパネルが、アレイベイトとしてアレイ方法によって創出され、次いで、追加のDNAオリゴヌクレオチドベイトが、遺伝子の標的化セットに対して不十分な配列決定カバレッジの区域を標的にするように設計され得る。SNPの収集物の配列決定は、次いで、元のアレイベイトセットプラス追加のオリゴベイトで反復されて、意図される全配列決定カバレッジを達成することができる。

【0266】

一部の実施形態において、オリゴベイトは、標準的アレイベイトセットに付加されて、より徹底的な配列決定カバレッジを達成する。一実施形態において、オリゴベイトは、標準的アレイベイトセットを用いる初期の配列決定ラウンドに続いて、不十分な配列決定カバレッジの区域を標的にするように設計されている。

【0267】

別の実施形態において、オリゴベイトは、特に目的の遺伝子における配列を標的にするように設計されている。これらのオリゴベイトは、標準的アレイベイトセットに、または現存のオリゴ／アレイハイブリッドベイトセットに付加されて、例えば、全アレイベイトプール再設計サイクルを経ることなく拡大された遺伝子セットの配列決定カバレッジの増加を達成することができる。

【0268】

オリゴベイトは、NimbleGen (Roche) またはDNAオリゴについてはIntegrated DNA Technologies (IDT) などの市販供給源から得ることができる。オリゴは、Agilent Technologiesからも得ることができる。濃縮のためのプロトコールは、公的に入手可能であり、例えばSureSelect Targetである。

【0269】

濃縮系

ベイトは、参照により本明細書に組み込むUS 2010/0029498およびGnirke, A. ら (2009) Nat Biotechnol. 27 (2) : 182-189に記載されている方法によって生成され得る。例えば、ビオチン化RNAベイトは、マイクロアレイ上で元々合成される合成の長いオリゴヌクレオチドのプールを得ること、およびオリゴヌクレオチドを増幅してベイト配列を生成することによって生成され得る。一部の実施形態において、ベイトは、ベイト配列の1つの末端でRNAポリメラーゼプロモーター配列を付加すること、およびRNAポリメラーゼを使用してRNA配列を合成することによって生成される。一実施形態において、合成のオリゴデオキシヌクレオチドのライブラリーは、Agilent Technologies, Inc. などの市販供給元から得られ、公知の核酸増幅方法を使用して増幅され得る。

【0270】

例えば、ベイトの大きな収集物は、オリゴヌクレオチドアレイ、例えばAgilent プログラム制御可能なDNAマイクロアレイ上で元々合成された合成のオリゴヌクレオチドのカスタムプールから生成され得る。したがって、少なくとも約2,500、5,000、10,000、20,000、30,000、40,000、50,000、または60,000個の独特のオリゴヌクレオチドは、同時に合成され得る。

【0271】

一実施形態において、独特のオリゴヌクレオチドの最小セットが選択され、追加のコピー（例えば、逆相補体と元の順方向鎖との間で交互する。）が、例えば標的の予備選択セット（例えば、エクソンの予備選択セット）を捕捉するように設計されているベイトのために合成のオリゴヌクレオチドアレイの最大容量が達されるまで付加される。別の実施形態において、標的は、例えば順方向および逆相補体オリゴヌクレオチドの両方を合成することによって、少なくとも2回表される。所定の標的に対する順方向および逆相補体オリゴヌクレオチドを合成することは、非常に同じ配列を2回合成することより良好なりダンダンシーを合成ステップで提供することができる。なお別の実施形態において、PCR生成物またはベイトは、順方向および逆相補体オリゴヌクレオチドに関して同じである。

10

【0272】

チップからのオリゴヌクレオチドは、1回合成され、次いで増幅されて、何回も使用され得るオリゴヌクレオチドのセットを創出することができる。この手法は、多数の選択実験のためのベイトとして使用され得るユニバーサル試薬を生成し、これによって、配列決定コストのごく一部であるべきチップのコストを償却する。代替として、ベイト配列は、PCRなどの公知の核酸増幅方法を使用し、鑄型としてヒトDNA試料またはプールされているヒトDNA試料を使用して生成され得る。

【0273】

合成に続いて、オリゴヌクレオチドは、化学的開裂、続いて保護基の除去、およびユニバーサルプライマーを使用して二本鎖化DNAに増幅されたPCRによって、アレイから放出され得る（例えば、剥離される。）。PCRの第2のラウンドは、DNAを一本鎖RNAに転写するために使用されるアンプリコンに、プロモーター（例えば、T7、SP6またはT3プロモーター）部位を組み込むために使用され得る。

20

【0274】

一実施形態において、ベイトは、ギャップまたは重なりなく配列（例えば、エクソン）に沿ってタイル張りされている。例えば、ベイトはUCSCゲノムブラウザにおいて示されている参照ゲノム配列の鎖中の最も「左」のコード塩基で開始することができ（例えば、遺伝子の配向に依存してコード配列に沿って5'から3'または3'から5'）、追加のベイトは、全てのコード塩基がカバーされるまで付加される。別の実施形態において、各標的に対して少なくとも2つ、3つ、4つまたは5つのベイトが、少なくとも約15個、30個、45個または60個の塩基が重なって設計される。オリゴヌクレオチド合成、およびユニバーサルプライマーを使用するPCR増幅の後、二本鎖化DNAの尾部の1つが酵素分解され、続いて、鎖の1つが分解され得る。一本鎖生成物は、ハイブリダイズされ、充填することによって完全二本鎖化にされ、PCRによって増幅され得る。この方式において、化学合成され得るよりも多い、少なくとも約300個、400個、500個または600個の連続する標的特異的塩基を含有するベイトを生成することが可能である。こうした長いベイトは、高い特異性および感受性を必要とする用途に、またはベイトの長さを限定することから必ずしも利益を得ない用途（例えば、長い近接するゲノム領域の捕捉）に有用であり得る。

30

【0275】

一実施形態において、各標的のカバレッジは判定することができ、同様のカバレッジを得る標的はグループ化することができる。ベイト配列の明確に異なるセットは、標的の各群のために創出されて、表示をさらに改善することができる。別の実施形態において、マイクロアレイチップからのオリゴヌクレオチドは、ハイブリダイゼーションの効力について試験され、オリゴヌクレオチドがそれらの捕捉効力によってグループ化されるマイクロアレイチップの生成ラウンドが要求され、したがって、ベイト効力における変動を相殺する。なお別の実施形態において、相対的に少数の複合体プールを形成するためにオリゴヌクレオチドプールが凝集されて、これらの間の捕捉効力における変動はほとんどない。

40

【0276】

本明細書に記載されているベイトは、タグ、例えば親和性タグで標識され得る。例示的

50

親和性タグとしては、これらに限定されないが、ビオチン分子、磁性粒子、ハプテン、またはタグ分子でタグ付けされたベイトの単離を可能にする他のタグ分子が挙げられる。こうした分子および核酸にこれらを付着させる方法（例えば、本明細書において開示されている方法において使用されるベイト）は、当技術分野においてよく知られている。ビオチン化ベイトを作製するための例示的方法は、例えば、参照により全体が本明細書に組み込むGnirk et al.、Nat. Biotechnol. 2009; 27(2): 182-9に記載されている。

【0277】

その上、当技術分野において知られているのは、タグ付けされたベイトのセットに結合するまたはハイブリダイゼーション混合物からそれを分離させられる分子、粒子または装置である。一実施形態において、分子、粒子または装置は、タグ（例えば、親和性タグ）に結合する。一実施形態において、分子、粒子または装置は、アビジン分子、磁石、または抗体もしくははこの抗原結合断片である。一実施形態において、タグ付けされたベイトは、ストレプトアビジン分子でコーティングされている磁性ビーズを使用して分離される。

【0278】

オリゴヌクレオチドライブラリーを調製するための例示的方法は、例えば、参照により全体を本明細書に組み込むGnirk et al.、Nat. Biotechnol. 2009; 27(2): 182-9、およびBlumenstiel B.、Curr. Protoc. Hum. Genet. 2010; 18章: 18.4ユニットに記載されている。

【0279】

本発明において特色とされている方法および組成物は、各ベイトセット／標的カテゴリの相対配列カバレッジを調整することを伴う。ベイト設計における相対配列カバレッジの差異を与えるための方法は、以下の1つ以上を含む：

(i) 異なるベイトセットの示差表示—所定の標的（例えば、標的メンバー）を捕捉するためのベイトセット設計は、相対的な標的カバレッジ深度を増強／低減するため、より多い／より少ない数のコピーに含められ得る；

(ii) ベイトサブセットの示差重なり—所定の標的（例えば、標的メンバー）を捕捉するためのベイトセット設計は、隣接しているベイト間により長いまたはより短い重なりを含むことで、相対的な標的カバレッジ深度を増強／低減することができる；

(iii) 示差ベイトパラメータ—所定の標的（例えば、標的メンバー）を捕捉するためのベイトセット設計は、配列修飾／より短い長さを含むことで、捕捉効率を低減するとともに相対的な標的カバレッジ深を低下させることができる；

(iv) 異なるベイトセットの混合—異なる標的セットを捕捉するように設計されているベイトセットは、異なるモル比で混合されて、相対的な標的カバレッジ深度を増強／低減することができる；

(v) オリゴヌクレオチドベイトセットの異なる型を使用すること—特定の実施形態において、ベイトセットは、以下を含むことができる：

(a) 1つ以上の化学的に（例えば、非酵素的に）合成された（例えば、個別に合成された）ベイト、

(b) アレイにおいて合成された1つ以上のベイト、

(c) 1つ以上の酵素的に調製された、例えばインビトロ転写されたベイト；

(d) (a)、(b) および／もしくは (c) の任意の組合せ、

(e) 1つ以上のDNAオリゴヌクレオチド（例えば、天然または非天然のDNAオリゴヌクレオチド）、

(f) 1つ以上のRNAオリゴヌクレオチド（例えば、天然または非天然のRNAオリゴヌクレオチド）、

(g) (e) および (f) の組合せ、または

(h) 上記の任意のものの組合せ。

【0280】

10

20

30

40

50

異なるオリゴヌクレオチド組合せは、異なる比、例えば 1 : 1、1 : 2、1 : 3、1 : 4、1 : 5、1 : 10、1 : 20、1 : 50 ; 1 : 100 または 1 : 1000 などから選択される比で混合され得る。一実施形態において、化学合成ベイト対アレイ生成ベイトの比は、1 : 5、1 : 10 または 1 : 20 から選択される。DNA または RNA オリゴヌクレオチドは、天然または非天然であり得る。特定の実施形態において、ベイトは、例えば溶融温度を増加させるため、1 つ以上の非天然ヌクレオチドを含む。例示的非天然オリゴヌクレオチドとしては、修飾されている DNA または RNA ヌクレオチドが挙げられる。例示的修飾 RNA ヌクレオチドはロックド核酸 (LNA) であり、ここで、LNA ヌクレオチドのリボース部分は、2' 酸素および 4' 炭素を接続する余分な架橋で修飾される (Kaur, H; Arora, A; Wengel, J; Maiti, S; Arora, A.; Wengel, J.; Maiti, S. (2006). 「Thermodynamic, Counterion, and Hydration Effects for the Incorporation of Locked Nucleic Acid Nucleotides into DNA Duplexes」. *Biochemistry* 45 (23) : 7347-55)。他の例示的修飾 DNA および RNA ヌクレオチドとしてはこれらに限定されないが、ペプチド結合によって連結されている反復 N- (2-アミノエチル) -グリシン単位から構成されているペプチド核酸 (PNA) (Egholm, M. ら (1993) *Nature* 365 (6446) : 566-8) ; 低 GC 領域を捕捉するために修飾された DNA または RNA オリゴヌクレオチド ; 二環式核酸 (BNA) または架橋オリゴヌクレオチド ; 修飾 5-メチルデオキシシチジン ; および 2, 6-ジアミノプリンが挙げられる。他の修飾 DNA および RNA ヌクレオチドは、当技術分野において知られている。

【0281】

特定の実施形態において、標的配列 (例えば、標的メンバー) の実質的に均一なまたは均質のカバレッジが得られる。例えば、各ベイトセット / 標的カテゴリー内で、カバレッジの均一性は、例えば以下の 1 つ以上によってベイトパラメータを修飾することによって最適化され得る :

(i) ベイト表示または重なりを増加 / 減少することが使用されて、同じカテゴリーにおける他の標的に対して、不足して / 過度にカバーされる標的 (例えば、標的メンバー) のカバレッジを増強 / 低減することができる ;

(ii) 低いカバレッジのため、標的配列 (例えば、高 GC 含有量配列) を捕捉するのが難しいのであれば、カバーするためにベイトセット、例えば隣接配列 (例えば、GC が豊富でない隣接配列) で標的化されている領域を拡大する ;

(iii) ベイト配列の修飾がなされて、ベイトの二次構造を低減するとともに選択のその効率を増強することができる ;

(iv) ベイトの長さを修飾することが使用されて、同じカテゴリー内の異なるベイトの融解ハイブリダイゼーション動態を等しくすることができる。ベイトの長さは、直接的に (変動する長さでベイトを生成することによって) または間接的に (一貫性の長さのベイトを生成すること、およびベイト末端を任意の配列で置換することによって) 修飾され得る ;

(v) 同じ標的領域のための異なる配向のベイト (すなわち、順方向および逆の鎖) を修飾することは、異なる結合効率を有することがある。各標的のための最適なカバレッジを提供するいずれかの配向を有するベイトセットが選択され得る ;

(vi) 各ベイト上に存在する結合実体、例えば捕捉タグ (例えば、ビオチン) の量を修飾することは、その結合効率に影響することがある。特異的標的を標的にするベイトのタグレベルを増加 / 減少することが使用されて、相対的な標的カバレッジを増強 / 低減することができる ;

(vii) 異なるベイトのために使用されるヌクレオチドの型を修飾することが改変されて、標的に対する結合親和性に影響するとともに相対的な標的カバレッジを増強 / 低減することができる ; または

10

20

30

40

50

(v i i i) 修飾オリゴヌクレオチドベイトを使用し、例えば、より安定な塩基対合を有することが使用されて、高GC含有量に対して低いまたは正常のGC含有量の区域間の融解ハイブリダイゼーション動態を等しくすることができる。

【0282】

例えば、オリゴヌクレオチドベイトセットの異なる型が使用され得る。

【0283】

一実施形態において、選択の効率についての値は、予備選択標的領域を包含するためのベイトオリゴヌクレオチドの異なる型を使用することによって修飾される。例えば、第1のベイトセット（例えば、10,000-50,000個のRNAまたはDNAベイトを含むアレイベースのベイトセット）は、大きな標的領域（例えば、1-2MB全標的領域）をカバーするために使用され得る。第1のベイトセットは、第2のベイトセット（例えば、5,000個未満のベイトを含む、個別に合成されたRNAまたはDNAベイトセット）でスパイクされて、予備選択標的領域（例えば、標的領域の目的のスパニングの選択サブゲノム間隔、例えば250kb以下）および／またはより高い二次構造、例えばより高GC含有量の領域をカバーすることができる。目的の選択サブゲノム間隔は、本明細書に記載されている遺伝子もしくは遺伝子産物、またはこの断片の1つ以上に対応することができる。第2のベイトセットは、所望されるベイト重なりに依存して、約2,000-5,000個のベイトを含むことができる。なお他の実施形態において、第2のベイトセットは、第1のベイトセットにスパイクされた選択オリゴベイト（例えば、400、200、100、50、40、30、20、10個未満のベイト）を含むことができる。第2のベイトセットは、個別のオリゴベイトの任意の比で混合され得る。例えば、第2のベイトセットは、1:1の等モル比で存在する個別のベイトを含むことができる。代替として、第2のベイトセットは、異なる比（例えば、1:5、1:10、1:20）で存在する個別のベイトを含むことで、例えば、特定の標的の捕捉を最適化することができる（例えば、特定の標的は、他の標的と比較して、5-10×の第2のベイトを有することができる。）。

【0284】

ハイブリダイゼーション条件

本発明において特色とされる方法は、ライブラリー（例えば、核酸ライブラリー）と複数のベイトとを接触させて、選択ライブラリーキャッチを提供するステップを含む。接触ステップは、溶液ハイブリダイゼーションにおいて達成され得る。特定の実施形態において、方法には、溶液ハイブリダイゼーションの1つ以上の追加のラウンドによってハイブリダイゼーションステップを反復することが含まれる。一部の実施形態において、方法は、ベイトの同じまたは異なる収集物を用いる溶液ハイブリダイゼーションの1つ以上の追加のラウンドに、ライブラリーキャッチをかけることをさらに含む。

【0285】

他の実施形態において、本発明において特色とされる方法は、ライブラリーキャッチを増幅すること（例えば、PCRによる）をさらに含む。他の実施形態において、ライブラリーキャッチは増幅されない。

【0286】

なお他の実施形態において、方法は、ライブラリーキャッチを遺伝子型同定にかけ、これによって、選択核酸の遺伝子型を同定するステップをさらに含む。

【0287】

より具体的に、数千個のベイト配列の混合物は、核酸の群における相補的核酸に有効にハイブリダイズすることができ、こうしたハイブリダイズ化核酸（核酸のサブグループ）は、有効に分離および回復される。一実施形態において、本明細書に記載されている方法は、約1,000個超のベイト配列、約2,000個超のベイト配列、約3,000個超のベイト配列、約4,000個超のベイト配列、約5,000個超のベイト配列、約6,000個超のベイト配列、約7,000個超のベイト配列、約8,000個超のベイト配列、約9,000個超のベイト配列、約10,000個超のベイト配列、約15,000

個超のベイト配列、約 20,000 個超のベイト配列、約 30,000 個超のベイト配列、約 40,000 個超のベイト配列、または約 50,000 個超のベイト配列を含有するベイト配列のセットを使用する。

【0288】

一部の実施形態において、選択プロセスは、例えば選択核酸の濃縮を増大するために、核酸の選択サブグループ上で反復される。例えば、ハイブリダイゼーションの 1 つのラウンド後、核酸の数千倍の濃縮が観察され得る。第 2 のラウンド後、濃縮は、単一のシーケンサー実行において標的の何百倍のカバレッジを提供できる例えば約 15,000 倍の平均濃縮率まで上昇することができる。したがって、ハイブリッド選択の単一のラウンドにおいて達成不可能な濃縮因子を必要とする実験のため、方法は、典型的に、ベイト配列のセットを用いる溶液ハイブリダイゼーションの 1 つ以上の追加のラウンドに核酸の単離サブグループ（すなわち、標的配列の一部または全て）をかけることを含む。

10

【0289】

2 つの異なるベイト配列（ベイト 1、ベイト 2）を用いる順次のハイブリッド選択は、「交差点」、すなわち、例えばこれらに限定されないが染色体間を濃縮することが挙げられる用途のために使用されるベイト 1 およびベイト 2 に結合する DNA 配列のサブグループを、単離および配列決定するために使用され得る。例えば、第 1 染色体上の配列に対して特異的なベイトを用いる腫瘍試料からの DNA の選択、続いて、第 2 染色体に対して特異的なベイトにハイブリダイズする配列の第 1 の選択の生成物からの選択は、両染色体からの配列を含有する染色体転座接合部で配列を濃縮することができる。

20

【0290】

核酸の選択サブグループの体積モル濃度は、任意の特別な核酸の体積モル濃度が、核酸のサブグループにおける全ての選択核酸の平均体積モル濃度の小さい変動内であるように制御され得る。標的表示の均等性を制御および最適化するための方法として、これらに限定されないが、当技術分野においてよく知られているプローブ設計の物理化学的および経験的規則に基づくベイト配列の合理的設計、ならびに働きが不十分であると知られているまたは疑われている配列がそれらの固有の弱点を補うために過剰表示されているベイトのプールが挙げられる。一部の実施形態において、核酸の単離サブグループの少なくとも約 50%、60%、65%、70%、75%、80%、85%、90% または 95% は、平均体積モル濃度の約 20 倍、15 倍、10 倍、5 倍、3 倍または 2 倍内である。一実施形態において、核酸の単離サブグループの少なくとも約 50% は、平均体積モル濃度の約 3 倍内である。別の実施形態において、核酸の単離サブグループの少なくとも約 90% は、平均体積モル濃度の約 10 倍内である。

30

【0291】

選択の効率における変動は、ベイトの濃度を改変することによってさらに調整され得る。一実施形態において、選択の効率は、ベイトの等モルミックスを使用する場合に観察される示差配列捕捉効率を参照してベイトの相対的豊富度または結合実体（例えば、ハプテンまたは親和性タグの密度）の密度を調整することによって群（例えば、第 1、第 2 または第 3 の複数のベイト）内の個別のベイトの効率を平準化すること、および次いで内部平準化群 2 に対して、ちょうど等しい内部平準化群 1 の示差的過剰を全体的ベイトミックスに導入することによって調整される。

40

【0292】

特定の実施形態において、本明細書に記載されている方法は、標的配列の均等カバレッジを達成することができる。一実施形態において、予想されるカバレッジの少なくとも約 50% を有する標的塩基のパーセントは、例えば、タンパク質コードエクソンなどの短い標的について少なくとも約 60%、70%、80% または 90% である。別の実施形態において、予想されるカバレッジの少なくとも約 50% を有する標的塩基のパーセントは、例えば、ゲノム領域など捕捉ベイトの長さと比較して長い標的について少なくとも約 80%、90% または 95% である。

【0293】

50

ハイブリダイゼーションより前に、ベイトは、当技術分野においてよく知られている方法に従って変性され得る。一般に、ハイブリダイゼーションステップは、過剰なブロッキングDNAを標識化ベイト組成物に付加すること、ブロックされたベイト組成物を、ハイブリダイズ条件下で、検出されるべき標的配列と接触させること、未ハイブリダイズ化ベイトを洗い流すこと、および標的へのベイト組成物の結合を検出することを含む。

【0294】

ベイトは、ハイブリダイズ条件下で標的配列にハイブリダイズまたはアニールされる。「ハイブリダイズ条件」は、ベイトと標的核酸との間のアニーリングを容易にする条件である。異なるベイトのアニーリングは、プローブ長さおよび塩基濃度などに依存して変動するので、アニーリングは、ベイト濃度、ハイブリダイゼーション温度、塩濃度、および当技術分野においてよく知られている他の因子を変動することによって容易にされる。

【0295】

ハイブリダイゼーション条件は、ベイトの濃度、塩基組成、複雑性および長さ、ならびにインキュベーションの塩濃度、温度および長さを変動することによって容易にされる。例えば、ハイブリダイゼーションは、5×のSSPE、5×のデンハート液、5mMのEDTAおよび0.1%のSDS、ならびに非特異的ハイブリダイゼーションを抑制するためのブロッキングDNAを含有するハイブリダイゼーション緩衝剤中で行われ得る。RNAase阻害剤は、ベイトがRNAであるならば使用され得る。一般に、ハイブリダイゼーション条件は、上に記載されている通り、約25℃から約65℃、典型的に約65℃の温度、および約0.5時間から約96時間、典型的に約66時間のインキュベーション長さを含む。追加の例示的ハイブリダイゼーション条件は、本明細書における実施例12A-12Cおよび表14にある。

【0296】

本明細書に記載されている方法は、標準的液体取扱い方法および装置に適応可能である。一部の実施形態において、方法は、マルチウェルプレートを取り扱う装置など、当技術分野において知られている通りの自動化液体取扱い技術を使用して実施される（例えば、Gnirke, A.ら(2009) Nat Biotechnol. 27(2):182-189を参照されたい。）。これは、これらに限定されないが自動化ライブラリーの構築、ならびに設定を含む溶液ハイブリダイゼーションおよび溶液ハイブリダイゼーション後の洗浄のステップを含むことができる。例えば、こうした自動化方法を実施するために、溶液ハイブリダイゼーション反応後のビーズの捕捉および洗浄のステップのための器具が使用され得る。例示的器具は、これらに限定されないが、以下の位置を含むことができる：ストレプトアビジンがコーティングされた磁性ビーズを含有するマルチウェルプレートのための位置、溶液ハイブリッド選択反応を含有するマルチウェルプレート、試薬を予備加熱するとともにユーザーに定義された温度で洗浄ステップを実施するためのI/O制御の熱ブロックのための位置、ピペット先端のラックのための位置、磁石で固定化されたビーズからの上澄みの分離を容易にする特定の立体配置でレイアウトされている磁石、ピペット先端を洗浄するとともに廃棄物を処分する洗浄ステーションを有する位置、ならびに低いおよび高ストリンジェンシーの洗浄緩衝剤または最終キャッチのアルカリ性溶出のための溶液など他の溶液および試薬のための位置。一実施形態において、器具は、平行してキャッチ中和ステップを介するビーズ捕捉ステップからの最大96個までのハイブリッド選択物を加工するように設計されている。別の実施形態において、1つ以上の位置は、二重機能を有する。なお別の実施形態において、ユーザーは、1つのプレートを別のプレートと交換するためのプロトコールによって指示される。

【0297】

直接的に選択された核酸は、連鎖状化および剪断することができ、これは、短い配列決定読み取りの限界を乗り越えるのに用いられる。一実施形態において、各エクソンサイズの配列決定標的は、標的とほぼ同じサイズであるとともに標的の末端点に近い末端点を有する単一のベイト分子で捕捉される。およそ100個以上の連続する塩基対を有する二重鎖分子を形成するハイブリッドだけが、ストリンジェントなハイブリダイゼーション後の洗

10

20

30

40

50

浄を生き残る。結果として、核酸の選択サブグループ（すなわち、「キャッチ」）は、末端がベイト分子の末端に近い無作為に切断されたゲノムDNA断片が濃縮される。非常に短い配列決定読み取りを用いる「キャッチ」の、より末端の配列決定は、標的の末端に（または外側さえに）近くでより高いカバレッジ、および中間近くでより低いカバレッジを与えることができる。

【0298】

ライゲーションによって「キャッチ」分子を連鎖状化すること、ならびに続いて、ランダム切断することおよびショットガン配列決定することは、標的配列の全長に沿って配列カバレッジを得るための1つの方法である。この方法は、非常に短い読み取りでの末端配列決定より高い百分率の、標的（近くの標的に相対する。）上にある配列決定塩基を生成する。同時ライゲーションによって分子を連鎖状化するための方法は、当技術分野においてよく知られている。鎖状体形成は、単純なプラント末端ライゲーションによって行われ得る。効果的ライゲーションのための「粘性」末端は、それらの5'末端近くに制限部位を有するPCRプライマーを用いる「キャッチ」のPCR増幅、続いて、対応する制限酵素（例えば、Not I）を用いる消化を含む様々な方法によって、またはT4 DNAポリメラーゼ（Aslanidisおよびde Jong、Nucleic Acids Res. 18:6069-6074、1990）による部分的「チューバック」もしくはUDGグリコシラーゼおよびリアーゼエンドVII（例えば、New England BiolabsカタログE5500S）を用いるウラシル含有PCR生成物の処理など、PCR生成物のライゲーション非依存性クローニングのために共通して使用されるものと同様の戦略によって生成され得る。

【0299】

別の実施形態において、ベイト分子のスタガーセットが、領域を標的にするために使用されて、標的領域の全体にわたって頻出のベイト末端を得る。一部の実施形態において、単に末端配列決定された「キャッチ」（すなわち、鎖状体形成および切断することがない。）は、実際の配列決定標的（例えば、エクソン）を含むベイトによってカバーされる全領域に沿ってかなり均一な配列カバレッジを提供する。ベイト分子をスタガーすることは、ベイトによってカバーされるセグメントを広くするので、配列決定塩基は、より広い区域にわたって分布される。結果として、標的上の配列対近くの標的の比は、標的1つ当たり単一のベイトだけをしばしば必要とする重ならないベイトを用いる選択よりも低い。

【0300】

別の実施形態において、若干より長い読み取り（例えば、76個の塩基）を用いる末端配列決定は、短い選択標的（例えば、エクソン）を配列決定するための典型的な方法である。非常に短い読み取りを用いる末端配列決定とは異なり、この方法は、中間でのカバレッジにおいて低下することなく、単様式のカバレッジプロファイルに至る。この方法は、上に記載されている連鎖状化および切断方法より行うのが簡便であり、標的に沿って相対的に均等なカバレッジをもたらす、ベイトおよび適正な標的に向かう高い百分率の配列決定塩基を生成する。

【0301】

一実施形態において、核酸の選択サブグループは、増幅される（例えば、PCRによる。）後に、配列決定することまたは遺伝子型を同定することによって分析される。別の実施形態において、サブグループは、例えば単一の分子を読み取ることができる高感度分析方法によって選択サブグループが分析される場合、増幅ステップがなくて分析される。

【0302】

配列決定

本発明は、核酸を配列決定する方法も含む。これらの方法において、核酸ライブラリーメンバーは、本明細書に記載されている方法を使用すること、例えば溶液ハイブリダイゼーションを使用することによって単離され、これによって、ライブラリーキャッチを提供する。ライブラリーキャッチまたはこのサブグループは、配列決定され得る。したがって、本発明において特色とされる方法は、ライブラリーキャッチを分析することをさらに含

む。一実施形態において、ライブラリーキャッチは、配列決定方法、例えば本明細書において記載されている通りの次世代配列決定方法によって分析される。方法は、溶液ハイブリダイゼーションによってライブラリーキャッチを単離すること、およびライブラリーキャッチを核酸配列決定にかけることを含む。特定の実施形態において、ライブラリーキャッチは、再配列決定され得る。

【0303】

当技術分野において知られている配列決定の任意の方法が、使用され得る。選択方法によって単離された核酸の配列決定は、典型的に、次世代配列決定（NGS）を使用して実施される。次世代配列決定は、高度に平行な様式（例えば、 10^5 個超の分子が同時に配列決定される。）にて、個別の核酸分子または個別の核酸分子のためのクローン的に拡大されたプロキシのいずれかのヌクレオチド配列を決定する任意の配列決定方法を含む。一実施形態において、ライブラリーにおける核酸種の相対的豊富度は、配列決定実験によって生成されるデータにおけるそれらの同族配列の出現の相対的な数をカウントすることによって概算され得る。次世代配列決定方法は、当技術分野において知られており、例えば、参照により本明細書に組み込む Metzker, M. (2010) Nature Biotechnology Reviews 11:31-46に記載されている。

【0304】

一実施形態において、次世代配列決定は、個別の核酸分子のヌクレオチド配列の決定を可能にする（例えば、Helicos BioSciencesの Heliscope Gene Sequencingシステム、およびPacific Biosciencesの PacBio RSシステム）。他の実施形態において、配列決定方法は、個別の核酸分子のためのクローン的に拡大されたプロキシのヌクレオチド配列を決定する（例えば、Solexaシーケンサー、Illumina Inc.、San Diego、Calif.; 454 Life Sciences (Branford, Conn.)、および Ion Torrent）。例えば、大規模に平行な短い読み取り配列決定（例えば、Solexaシーケンサー、Illumina Inc.、San Diego、Calif.）は、より少ないがより長い読み取りを生成する他の配列決定方法より、配列単位1個当たりの配列の塩基を多く生成する。次世代配列決定するための他の方法または機械としては、これらに限定されないが、454 Life Sciences (Branford, Conn.)、Applied Biosystems (Foster City, Calif.; SOLiD sequencer)、Helicos BioScience Corporation (Cambridge, Mass.) によって提供されるシーケンサー、ならびにエマルジョンおよびマイクロ流体配列決定技術ナノ液滴（例えば、GnuBi液滴）が挙げられる。

【0305】

次世代配列決定するためのプラットフォームとしては、これらに限定されないが、Roche/454の Genome Sequencer (GS) FLX System、Illumina/Solexaの Genome Analyzer (GA)、Life/AMPGの Support Oligonucleotide Ligation Detection (SOLiD) システム、Polonatorの G.007 システム、Helicos BioSciencesの Heliscope Gene Sequencing システム、および Pacific Biosciencesの PacBio RS システムが挙げられる。

【0306】

NGS技術は、1つ以上のステップ、例えば鋳型調製、配列決定および撮像、ならびにデータ分析を含むことができる。

【0307】

追加の例示的配列決定方法論は、当技術分野において知られており、例えば、これらの一部は、ともに2011年12月29日に出願された共同所有のUS SN第13/339,986およびPCT/US11/67725に記載されており、これらの内容は参照に

10

20

30

40

50

より組み込まれる。

【0308】

アラインメント

アラインメントは、読み取りを位置、例えばゲノムの位置に関して整合するプロセスである。ミスアラインメント（例えば、ゲノムにおける不正確な位置上の短い読み取りからの塩基対の配置）、例えば、実際の癌変異の周辺での読み取りの配列状況（例えば、反復性配列の存在）によるミスアラインメントは、代わりの対立遺伝子の読み取りが代わりの対立遺伝子読み取りの主な重層から回避され得るので、変異検出の感受性における低減に至る恐れがある。実際の変異が存在しないという問題のある配列状況が発生するならば、ミスアラインメントは、参照ゲノム塩基の実際の読み取りを間違った位置に設置することによって、「変異された」対立遺伝子のアーチファクツ的読み取りを導入する恐れがある。増殖多重遺伝子分析のための変異－呼び出しアルゴリズムは、低豊富度変異にも高感度であるべきなので、これらのミスアラインメントは、偽陽性発見率を増加／特異性を低減する恐れがある。

【0309】

本明細書において考察されている通り、実際の変異に対する感受性の低減は、分析されている遺伝子における予想変異部位の周辺でのアラインメントの質を（手動で、または自動化様式で）評価することによって取り組まれ得る。評価されるべき部位は、癌変異のデータベース（例えば、COSMIC）から得ることができる。問題ありと同定される領域は、例えば、Smith-Watermanアラインメントなど、より緩徐であるがより正確なアラインメントアルゴリズムを使用するアラインメントの最適化（または再アラインメント）によって、関連の配列状況におけるより良好な性能を与えるために選択されたアルゴリズムの使用で修正され得る。一般のアラインメントアルゴリズムが問題を修正することができない場合において、カスタム化されたアラインメント手法が、例えば以下によって創出され得る：置換を含有するという高い見込みを有する遺伝子のための最大差異ミスマッチペナルティーパラメータの調整；特定の腫瘍型（例えば、メラノーマにおけるC→T）において共通である特異的変異型に基づく特異的ミスマッチペナルティーパラメータを調整すること；または特定の試料型において共通である特異的変異型（例えば、FFPEにおいて共通である置換）に基づく特異的ミスマッチペナルティーパラメータを調整すること。

【0310】

ミスアラインメントによる遺伝子領域の評価における特異性の低減（偽陽性率の増加）は、配列決定された試料における全ての変異呼び出しの手動検査または自動検査によって判定され得る。ミスアラインメントによる擬似変異呼び出しの傾向があると見出されたそれらの領域は、上記と同じアラインメント修正にかけられ得る。アルゴリズムの修正が可能だと見出されない場合において、問題の領域からの「変異」は、試験パネルから分類またはスクリーンして除かれ得る。

【0311】

挿入／欠失（インデル）

一般に、インデル変異の正確な検出は、本明細書において無効にされた配列決定プラットフォーム上での擬似インデル率が相対的に低いので、アラインメントにおける運動である（したがって、正しくアラインメントされたインデルのわずかな観察でも変異の強い証拠であり得る。）。しかしながらインデルの存在下での正確なアラインメントは難しいことがある（殊に、インデル長さが増加する場合。）。アラインメントに伴う一般の、例えば置換の課題に加えて、インデルこれ自体が、アラインメントの問題を引き起こす恐れがある。（例えば、ジヌクレオチド反復の2bpの欠失は、容易に断定的に配置することはできない。）感受性および特異性の両方は、より短い（<15bp）見かけのインデル含有読み取りの不正確な配置によって低減され得る。より大きいインデル（本発明者らの本プロセスにおける個別の読み取りの長さ－36bpと規模が近づいている。）は、読み取りを全くアラインメントし損ねる原因となり、アラインメントされた読み取りの標準的セ

ットにおけるインデルの検出を不可能にする恐れがある。

【0312】

癌変異のデータベースは、これらの問題に取り組むとともに性能を改善するために使用され得る。偽陽性インデル発見を低減する（特異性を改善する。）ため、共通して予想されるインデルの周辺の領域は、配列状況による問題のあるアラインメントについて検査され、上記の置換と同様に取り組まれ得る。インデル検出の感受性を改善するため、癌において予想されるインデル上の情報を使用するいくつかの異なる手法が使用され得る。例えば、短い読み取りが含有されている予想されるインデルがシミュレートされ、アラインメントが試みられ得る。アラインメントは研究することができ、問題のあるインデル領域は、例えばギャップ開口／延長ペナルティを低減することによって、または部分的な読み取り（例えば、読み取りの前半または後半）をアラインメントすることによって、アラインメントパラメータを調整させることができる。

10

【0313】

代替として、初期のアラインメントは、正常の参照ゲノムを用いるだけでなく、公知のまたは見込みのある癌インデル変異の各々を含有するゲノムの代わりのバージョンも用いて試みることができる。この手法において、初期にアラインメントし損ねたまたは不正確にアラインメントしたインデルの読み取りは、ゲノムの代わりの（変異された）バージョン上に首尾よく配置される。

【0314】

追加の例示的アラインメント方法論は当技術分野において知られており、例えば、これらの一部は、ともに2011年12月29日に出願された共同所有のUS SN 第13／339,986およびPCT／US 11／67725に記載されており、これらの内容は参照により組み込まれる。

20

【0315】

変異呼び出し

塩基呼び出しは、配列決定装置の生アウトプットを指す。変異呼び出しは、配列決定されているヌクレオチド位置について、ヌクレオチド値、例えばA、G、TまたはCを選択するプロセスを指す。典型的に、位置についての配列決定読み取り（または塩基呼び出し）は、1つ超の値を提供し、例えば、一部の読み取りはTを与え、一部はGを与える。変異呼び出しは、ヌクレオチド値、例えば配列に対してそれらの値の1つを割り当てるプロセスである。それは「変異」呼び出しと称されるが、それは、任意のヌクレオチド位置、例えば変異体対立遺伝子、野生型対立遺伝子、変異体または野生型のいずれかと特徴付けられなかった対立遺伝子に対応する位置にまたは変動によって特徴付けされていない位置にヌクレオチド値を割り当てるために適用される。変異呼び出しのための方法は、以下の1つ以上を含むことができる：参照配列における各位置での情報に基づき非依存性呼び出しをすること（例えば、配列読み取りを検査すること；塩基呼び出しおよび品質スコアを検査すること；潜在的遺伝子型を前提として、観察された塩基および品質スコアの確率を算出すること；および遺伝子型を割り当てること（例えば、ベイズ規則を使用する。））；偽陽性を除去すること（例えば、予想されるのよりずっと低いまたは高い読み取り深度を有するSNPを拒絶するための深度閾値を使用する；小さいインデルによる偽陽性を除去するための局所再アラインメント）；および連鎖不平衡（LD）／インピュテーションに基づく分析を行うことで呼び出しを精密にすること。

30

40

【0316】

特異的遺伝子型および位置に関連する遺伝子型見込みを算出するための反応式は、例えばLi H. およびDurbin R. Bioinformatics、2010；26（5）：589－95に記載されている。特定の癌型における特別な変異のための事前予想は、その癌型からの試料を評価する場合に使用され得る。こうした見込みは、癌変異の公開データベース、例えば、Catalogue of Somatic Mutation in Cancer（COSMIC）、HGMD（Human Gene Mutation Database）、The SNP Consortium、Breast Can

50

cer Mutation Data Base (BIC)、およびBreast Cancer Gene Database (BCGD) から誘導され得る。

【0317】

LD／インプテーションに基づく分析の例は、例えばBrowning B. L. およびYu Z. Am. J. Hum. Genet. 2009、85(6):847-61に記載されている。低カバレッジSNP呼び出し方法の例は、例えばLi Y. ら、Ann u. Rev. Genomics Hum. Genet. 2009、10:387-406に記載されている。

【0318】

変異呼び出し：置換

アラインメントの後、置換の検出は、呼び出し方法、例えば、ベイズ変異呼び出し方法を使用して行うことができ；これは、サブゲノム間隔の各々における各塩基、例えば評価されるべき遺伝子のエクソンに適用され、ここで、代わりの対立遺伝子の存在が観察される。この方法は、変異の存在下での読み取りデータを観察する確率を、塩基呼び出しエラー単独の存在下での読み取りデータを観察する確率と比較する。この比較が変異の存在を十分に強く支持しているならば、変異が呼び出され得る。

【0319】

癌DNAの分析について50%または100%の頻度からの偏差の制限に取り組む方法が開発された。(例えば、SNVMix - Bioinformatics. 2010年3月15日; 26(6):730-736。) 本明細書において開示されている方法は、しかしながら、試料DNAの1%から100%の間のどこかで、殊に50%より低いレベルで変異体対立遺伝子の存在の可能性の考慮を可能にするこの手法は、天然型(多クローン性)腫瘍DNAの低純度FFPE試料における変異の検出にとって特に重要である。

【0320】

ベイズ変異検出手法の利点は、変異の存在の確率と塩基呼び出しエラー単独での確率との比較が、部位での変異の存在の事前予想によって重み付けされ得ることである。代わりの対立遺伝子の一部の読み取りが、所定の癌型に関する頻繁に変異される部位で観察されるならば、変異の存在は、変異の証拠の量が通常の閾値に合わなくても確信的に呼び出しされ得る。この柔軟性は、なお希少な変異／なお低い純度試料に対する検出感受性を増加させるため、または読み取りカバレッジにおける減少に対して試験をより頑強にするために使用され得る。癌において変異されているゲノムにおけるランダム塩基対の見込みは、約 $1e-6$ である。典型的多重遺伝子癌ゲノムパネルにおける多くの部位での特異的変異の見込みは、より高い桁であり得る。これらの見込みは、癌変異の公開データベース(例えば、COSMIC)から誘導され得る。

【0321】

変異呼び出し：インデル

インデル呼び出しは、挿入または欠失が参照配列と異なる配列決定データにおける塩基を見出すプロセスであり、典型的には、関連の信頼スコアまたは統計的証拠メトリックを含む。

【0322】

インデル呼び出しの方法は、候補インデルを同定するステップ、局所再アラインメントを介する遺伝子型見込みを算出するステップ、ならびにLDベースの遺伝子型推測および呼び出しを行うステップを含むことができる。典型的に、潜在的インデル候補を得るためにベイズ手法が使用され、次いで、これらの候補は、ベイズフレームワークにおいて参照配列と一緒に試験される。

【0323】

候補インデルを生成するためのアルゴリズムは、例えば、McKenna A. ら、Genome Res. 2010; 20(9):1297-303; Ye K. ら、Bioinformatics, 2009; 25(21):2865-71; Lunter G. およびGoodson M. Genome Res. 2010、印刷に先立って電子出版；

10

20

30

40

50

Li H. ら、Bioinformatics 2009、Bioinformatics 25 (16) : 2078-9 に記載されている。

【0324】

インデル呼び出しおよび個別レベルの遺伝子型見込みを生成するための方法としては、例えば、Dindel アルゴリズム (Albers C. A. ら、Genome Res. 2010 年 10 月 27 日。[印刷に先立って電子出版]) が挙げられる。例えば、ベイズ EM アルゴリズムは、読み取りを分析し、初期インデル呼び出しをし、各候補インデルについて遺伝子型見込みを生成するために使用することができ、その後、例えば QCALL (Le S. Q. および Durbin R. Genome Res. 2010 年 10 月 27 日。[印刷に先立って電子出版]) を使用する遺伝子型のインビュテーションが続く。インデルを観察するという事前予想などのパラメータは、インデルのサイズまたは位置に基づいて調整され得る (例えば、増加または減少される。) 。

10

【0325】

追加の例示的変異呼び出し方法論は、当技術分野において知られており、例えば、これらの一部は、ともに 2011 年 12 月 29 日に出願された共同所有の US SN 13/339,986 および PCT/US 11/67725 に記載されており、これらの内容は参照により組み込まれる。

【実施例】

【0326】

本発明は、追加として、例示の目的で提供されるとともに任意の方式で本発明を限定すると意図されない以下の実施例を参照することにより記載される。当技術分野においてよく知られている標準的技術または下に具体的に記載されている技法が利用され得る。

20

【0327】

【実施例 1】 捕捉生成物への DNA プローブのハイブリダイゼーション

下記の手順は、捕捉生成物を用いる DNA プローブのハイブリダイゼーションに必要なステップを要約している。

【0328】

A. ハイブリダイゼーション

プールされたビオチン化バイト 100 ナノグラム、適応 DNA ライブラリー 500 ng、Cot-1 DNA 2 μ g、2.0 μ L 中のオリゴヌクレオチドブロッカー 2 ナノモルを、10 μ L の体積に合わせ、予備加温された Genisphere Buffer 6 (2 $\times$ の SDS 系ハイブリダイゼーション緩衝剤: 0.50 M の NaPO₄、1% の SDS、2 mM の EDTA、2 $\times$ の SSC、4 $\times$ のデンハート液) 10 μ L と混合する。混合物のボルテックス混合に続いて、40 μ L の鉱物油のオーバーレイを適用し、混合物を熱サイクラー内にて 95 $^{\circ}$ C で 5 分間変性し、71 $^{\circ}$ C にゆっくり減少する。混合物を 71 $^{\circ}$ C で 48 時間インキュベートする。

30

【0329】

B. ストレプトアビジンビーズに結合

ストレプトアビジンビーズを、ハイブリダイゼーション混合物への付加の前に以下の方式で調製する。ストレプトアビジンビーズを室温で 30 分間静置する。各ハイブリダイゼーション反応のため、Invitrogen M270 Streptavidin ビーズ (磁性) 50 μ L を、2 $\times$ の Bind、および Wash Buffer (10 mM の Tris-HCl (pH 7.5)、2 M の NaCl、1 mM の EDTA) で 2 回洗浄する。50 μ L の Bind および Wash Buffer ならびに水 30 μ L を含む 80 μ L 中に、ビーズを再懸濁する。

40

【0330】

48 時間のハイブリダイゼーション期間の最後に、ハイブリダイゼーション液体 20 μ L を鉱物油の下から除去し、ビーズ 80 μ L に添加して、100 μ L の総体積を提供する。混合物をチューブローター上で 30 分間回転させて、ハイブリダイズ化鑄型: バイト複合体上のビオチンとビーズ上のストレプトアビジンとの間に結合が生成するのを可能

50

にする。

【0331】

C. ストレプトアビジンビーズを洗浄

回転期間に続いて、試料を磁性分離ラック上へ配置する。ビーズを上澄みから分離させておき、捕捉プローブに結合していないDNAを含有する上澄みを除去し、捨てる。ビーズに結合されたプローブを以下の溶液で順次に洗浄する。各洗浄のため、所定の温度に予備平衡化された洗浄溶液を添加し、示された時間ローテーター上に配置し、短時間でスピンドウンし（磁石）、上澄みを回収し、捨てる。第1の洗浄は、回転しながら71℃で5分間、1000μLの1×のSSC/0.1%SDSを用いる。第2の洗浄は、回転しながら71℃で5分間、1000μLの0.1×SSC/0.1%SDSを用いる。第3の洗浄は、回転しながら71℃で5分間、1000μLの0.1×のSSC/0.1%SDSを用いる。第4の洗浄は、回転しながらRTで5分間、1000μLの0.1×のSSC/0.1%SDSを用いる。第5の洗浄は、なお磁石上のチューブを用いてRTで30秒間、1000μLの0.2×のSSCを用いる。最終の洗浄溶液は、下記に説明されている通り、事前のさらなる加工より前に完全に除去する。

【0332】

最終洗浄の後、50μLの0.125N NaOHを添加し、混合物をRTで10分間インキュベートし、2分毎にボルテックス処理して、溶液中にビーズを保持する。ビーズを有するチューブを磁石上に戻して1分間配置する。ビーズが磁石上にある間、1M Tris-HCl (pH 8.8) の50μLのアリコットを、新たな1.5mLのRNAse/DNAseフリーのPCRチューブに添加する。磁石上のチューブからの上澄み（0.125NのNaOH）を、1M Tris-HCl (pH 8.8) を含有するチューブに添加して、溶液を中和する。回復された鋳型断片を、1.5×の体積を使用するとともに20μLのEB緩衝剤（10mMのTris-Cl, pH 8.5）中で溶出して、AMPureビーズで精製する。

【0333】

【実施例2】一本鎖鋳型材料を用いるPCR反応

A. 最終のPCR濃縮

回復された一本鎖鋳型（16μL）を、以下の反応ミックス構成成分（KAPA HiFi マスターミックス（25μL）；10μMのプライマー1（2.5μL）、10μMのプライマー2（2.5μL）、水（4μL））を有する50μLの総体積に調製した。DNAを短時間でボルテックスし、簡潔な遠心分離に続いて溶液として再回収した。以下のプログラムを用いる熱サイクラー内に、反応物を配置した：5つ以上のサイクルに関して、98℃（5.0分）；98℃（20秒）；60℃（15秒）；72℃（20秒）；77℃（5.0分）。増幅された生成物を、1.5×の体積を使用するとともに20μLのEB Buffer (Qiagen)（10mMのTris-HCl, pH 8.5）中で溶出して、AMPureビーズで精製した。その結果得られるDNA濃度を、Quibit Fluorometerで測定し、適切なNGS配列決定プラットフォームでの使用のために希釈した。

【0334】

5つのサイクルの増幅を、捕捉後のIon Torrentライブラリーのために使用し、典型的に18以下のサイクルの増幅を、捕捉後のIlluminaライブラリーに使用する。標準的Illuminaプロトコルを、以下のPCR手順を使用して最適化する。回復された一本鎖DNA鋳型（2mL）を、25μLのSyberGreen Master Mix、8pmolのプライマー1および8pmolのプライマー2を含む50μLの最終体積中で合わせる。反応物を96ウェルqPCRプレート中に設置して、最終のPCR濃縮を模倣し、以下のプログラムを実行する：95℃（5分）、続いて、95℃（30秒）および60℃（45秒）の30サイクル。曲線の中点（増幅開始とプラトーとの中間）を見出すために、閾値を手動で調整し、この値から3サイクルを引き、最終のPCR濃縮に対して実行するためのサイクルの数を決定する。最適化するDNAの量は、最

終の濃縮反応に入れるものより8×少ないので、3サイクルを引き；中和された捕捉生成物の2μLはPCR最適化反応に入り、16μLは最終のPCR濃縮に入る。

【0335】

【実施例3】 バーコードドメイン対してイノシン塩基を用いるIllumina配列決定プラットフォームにおける使用のためのT_m増強オリゴヌクレオチド

表Iにおいて、Illumina配列決定プラットフォームを用いるDNA鋳型ライブラリーに対するハイブリッド捕捉実験において使用するため、以下のブロッキングオリゴヌクレオチドを設計した。T_m増強基としてLNA（「+C」または「+A」）を使用して、T_m増強オリゴヌクレオチドを調製した。ホスホラミダイト化学法を使用して、全てのオリゴヌクレオチドを調製した。参照によりその全体が組み込まれるOwczarzy (Biochemistry 2011 50:9352-9367) の方法を使用して、750mMのNaCl緩衝剤（5×のSSCと同様のイオン強度）中のLNA塩基について、および15mMのNaCl緩衝剤（0.1×のSSCと同様のイオン強度）についてT_m値を概算する。BNA修飾はLNA修飾と同様の熱力学的な効果を有するので、本明細書に示される予測は両クラスの修飾ブロッカーに当てはまり、LNA/BNA修飾は全ての実施例において置換することができる。例えば、LNA誘導の最も近い隣接するパラメータを使用して、下記の実施例における熱力学的モデル化を行い、他方で、BNA塩基を使用して、オリゴヌクレオチド合成を行った。

【0336】

【表1】

表I. T_m増強オリゴヌクレオチドブロッカー

配列 番号	配列	LNA数	T _m (°C) 750 mM (Na)	T _m (°C) 15 mM (Na)
1	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT	0	90.3	62.0
2	AGATCGGAAGAGCGT+CGTGTAGGGAAAGAGTGTAGATCT+CGGTGGTCGCCGTATCATT	2	92.0	63.4
3	AGAT+CGGAAGAGCGT+CGTGTAGGGAAAGAGTGTAGATCT+CGGTGGTCGC+CGTATCATT	4	93.9	65.0
4	AGAT+CGGAAGAGCGT+CGTGTAGGGAAAGAGTGTAGATCT+CGGTGGT+CGC+CGTAT+CATT	6	>95	66.6
5	AGAT+CGGAAGAG+CGT+CGTGTAGGGAAAGAGTGTAGAT+CT+CGGTGGT+CGC+CGTAT+CATT	8	>95	68.3
6	AGAT+CGGA+AGAG+CGT+CGTGT+AGGG+AAAGAGTGTAGAT+CT+CGGTGGT+CG+C+CGTAT+CATT	12	>95	71.1
7	AG+AT+CGGA+AGAG+CGT+CGTGT+AGGG+A+A+AGAGTGTAG+AT+CT+CGGTGGT+CG+C+CGTAT+CATT	16	>95	73.3
8	AG+AT+CGG+A+AG+AG+CGT+CGTGT+AGGG+A+A+AG+AGTGT+AG+AT+CT+CGGTGGT+CG+C+CGT+AT+C+ATT	22	>95	77.4

【0337】

表Iにおいて、全てのC位（すなわち、Cを有する9つの位置）が徹底的に置換されるまで、LNA-C T_m増強基が最初にT_m増強オリゴヌクレオチドに含められ、続いて、LNA-A T_m増強基がA位でその後に含められる。

【0338】

下記の表IIにおいて、バーコード配列（8-イノシン）を含有するアダプターに反するブロッカーとしての使用のための一連のT_m増強オリゴヌクレオチドの設計が示されて

いる。詳細な記述で説明すると、可変塩基と対合されたイノシン（表ⅠⅠの配列において「I」および「/ i d e o x y I /」と定義されている。）で評価された増強 T_m をモデル化するやり方はない。したがって、イノシン塩基を T_m 分析に含めなかったが、最終配列中に存在する。実際の配列についての正確な増強 T_m 値は、しかしながら、ルーチンの経験的方法によって容易に決定される。 T_m 増強基として LNA（「+C」または「+A」）を使用して、 T_m 増強オリゴヌクレオチドを調製した。ホスホラミダイト化学法を使用して、全てのオリゴヌクレオチドを調製した。

【0339】

【表2】

表Ⅱ. バーコード配列を有する T_m 増強オリゴヌクレオチドブロッカー

配列 番号	配列	LNA数	T_m (°C) 750 mM (Na)	T_m (°C) 15 mM (Na)
9	GATCGGAAGAGCACACGTCTGAACTCCAGTCAC (IIIIII II) ATCTCGTATGCCGTCTTCTGCTTG	0	89.8	61.7
10	GATCGGAAGAGCACACGTCTGAACTCCAGT+CAC (IIIIII III) ATCT+CGTATGCCGTCTTCTGCTTG	2	91.5	63.1
11	GATCGGAAGAGCACACGTCTGAA+CTCCAGT+CAC (IIIIIIII) ATCT+CGTATGC+CGTCTTCTGCTTG	4	93.9	65.2
12	GATCGGAAGAG+CACACGTCTGAA+CTCCAGT+CAC (IIIIIIII) ATCT+CGTATGC+CGTCTTCTG+CTTG	6	>95	66.8
13	GAT+CGGAAGAG+CACACGTCTGAA+CTCCAGT+CA+C (I IIIIIIII) ATCT+CGTATGC+CGTCTTCTG+CTTG	8	>95	68.7
14	GAT+CGGAAGAG+CACA+CGT+CTGAA+CTC+CAGT+CAC (IIIIIIII) ATCT+CGTATGC+CGT+CTT+CTG+CTTG	12	>95	72.7
15	GAT+CGGAAGAG+CA+CA+CGT+CTGAA+CT+C+CAGT+C A+C (IIIIIIII) AT+CT+CGTATG+C+CGT+CTT+CTG+ CTTG	17	>95	77.4
16	GAT+CGG+A+AGAG+CA+CA+CGT+CTG+AA+CT+C+CAG T+CA+C (IIIIIIII) +AT+CT+CGT+ATG+C+CGT+CTT +CTG+CTTG	22	>95	80.4

【0340】

表ⅠⅠにおいて、全てのC位（すなわち、Cを有する17つの位置）が徹底的に置換されるまで、LNA-C T_m 増強基が初期に T_m 増強オリゴヌクレオチドに含められ、続いて、LNA-A T_m 増強基がA位でその後に含められる。この実施例において、イノシン塩基を組み込んでバーコードドメインを広げた。全ての4核酸塩基または他のユニバーサル塩基の「N」塩基ランダムミックス、例えば5-ニトロインドールは、本明細書においてすでに記載されている通りに用いることができる。

【0341】

【実施例4】 バーコードドメインのための混合塩基（「N」塩基）を用いる Illumina 配列決定プラットフォームにおける使用のための T_m 増強ブロッキングオリゴヌクレオチド

表ⅠⅠⅠにおいて、Illumina 配列決定プラットフォームを用いるDNA鋳型ライブラリーに対するハイブリッド捕捉実験における使用のために、以下のオリゴヌクレオチドを設計した。 T_m 増強基として BNA（「/ i BNA-me C /」、「/ i BNA-T /」、または「/ i BNA-A /」）を使用して、 T_m 増強オリゴヌクレオチドを調製した。ホスホラミダイト化学法を使用して、全てのオリゴヌクレオチドを調製した。

【0342】

【表 3】

表 III. バーコード配列(「N」)を有する T_m 増強オリゴヌクレオチドブロッカー

配列番号	配列	BNA 数	T _m (°C) 750 mM (Na)	T _m (°C) 15 mM (Na)
17	インデックスブロック: CAAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTTCAG ACGTGTGCTCTTCCGATCT	0	約 90	61.8
18	インデックスブロック +10-BNA: CAAG/iBNA-meC/AGAAGA/iBNA-meC/GG/iBNA-meC/ATA /iBNA-meC/GAGATNNNNNNGTGA/iBNA-meC/TGGAGTT/iB NA-meC/AGA/iBNA-meC/GTGTG/iBNA-meC/TCTT/iBNA- meC//iBNA-meC/GATCT	10	>90	70.9
19	インデックスブロック +20-BNA: CAAG/iBNA-meC/AGAAGA/iBNA-meC/GG/iBNA-meC/A/i BNA-T/A/iBNA-meC/GAGA/iBNA-T/NNNNNG/iBNA-T/G A/iBNA-meC//iBNA-T/GGAG/iBNA-T//iBNA-T//iBNA- meC/AGA/iBNA-meC/G/iBNA-T/G/iBNA-T/G/iBNA-meC /TC/iBNA-T/T/iBNA-meC//iBNA-meC/GA/iBNA-T/CT	20	>90	76.8
20	インデックスブロック_RevComp: AGATCGGAAGAGCACACGTCTGAACTCCAGTCACNNNNNATCTC GTATGCCGTCTTCTGCTTG	0	約 90	61.8
21	インデックスブロック_RevComp +10-BNA: AGAT/iBNA-meC/GGAAGAG/iBNA-meC/ACA/iBNA-meC/G TCTGAA/iBNA-meC/TCCAGT/iBNA-meC/A/iBNA-meC/NN NNNATCT/iBNA-meC/GTATGC/iBNA-meC/GT/iBNA-meC /TTCTG/iBNA-meC/TTG	10	>90	70.9
22	インデックスブロック_RevComp +20-BNA: AGA/iBNA-T//iBNA-meC/GGAAGAG/iBNA-meC/A/iBNA- meC/A/iBNA-meC/GT/iBNA-meC/TGAA/iBNA-meC/T/iB NA-meC//iBNA-meC/AGT/iBNA-meC/A/iBNA-meC/NNNN NNAT/iBNA-meC/T/iBNA-meC/G/iBNA-T/A/iBNA-T/G/ iBNA-meC//iBNA-meC/GT/iBNA-meC/TT/iBNA-meC/TG /iBNA-meC/TTG	20	>90	80.0
23	PE1.0: AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGAC GCTCTTCCGATCT	0	約 90	62.0
24	PE1.0 +10-BNA: AATGATA/iBNA-meC/GG/iBNA-meC/GA/iBNA-meC/CA/i BNA-meC/CGAGAT/iBNA-meC/TACA/iBNA-meC/TCTTTC/ iBNA-meC/CTACACGACGCT/iBNA-meC/TTC/iBNA-meC/G AT/iBNA-meC/T	10	>90	72.6
25	PE1.0 +17-BNA: AATGATA/iBNA-meC/GG/iBNA-meC/GA/iBNA-meC/CA/i BNA-meC/CGAGAT/iBNA-meC/TA/iBNA-meC/A/iBNA-me C/T/iBNA-meC/TTT/iBNA-meC/C/iBNA-meC/TA/iBNA- meC/A/iBNA-meC/GA/iBNA-meC/G/iBNA-meC/T/iBNA- meC/TTC/iBNA-meC/GAT/iBNA-meC/T	17	>90	81.0
26	PE1.0_Rev Comp: AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGG TCGCCGTATCATT	0	約 90	62.0
27	PE1.0_Rev Comp +9-BNA: AGAT/iBNA-meC/GGAAGAG/iBNA-meC/GT/iBNA-meC/GT GTAGGAAAGAGTGTAGAT/iBNA-meC/T/iBNA-meC/GGTGG T/iBNA-meC/G/iBNA-meC//iBNA-meC/GTAT/iBNA-meC /ATT	9	>90	69.0
28	PE1.0_Rev Comp +20-BNA: AG/iBNA-A/T/iBNA-meC/GG/iBNA-A/AG/iBNA-A/G/iB NA-meC/GT/iBNA-meC/GTGT/iBNA-A/GGG/iBNA-A/A/i BNA-A/G/iBNA-A/GTGT/iBNA-A/G/iBNA-A/T/iBNA-me C/T/iBNA-meC/GGTGGT/iBNA-meC/G/iBNA-meC//iBNA -meC/GT/iBNA-A/T/iBNA-meC//iBNA-A/TT	20	>90	77.3

【0343】

表 III は、ブロッカーとしてアダプター配列のいずれかの鎖を使用して T_m 増強オリゴヌクレオチドを設計した実施例を提供している。ブロッカーとしての T_m 増強オリゴヌクレオチドとしての使用のための好ましい鎖は、最も少ない T_m 増強基の含有で、修飾基 1 個当たり最大「ブロッキング力」（すなわち、最も大きい最適な増強 T_m 値）を提供するものである。例えば、配列番号 19 および 22（配列番号 22 が好ましい。）および配列番号 25 および 28（配列番号 25 が好ましい。）を比較されたい。

【0344】

【実施例 5】 Ion Torrent PGM 配列決定プラットフォームにおける使用の

10

20

30

40

50

ための T_m 増強オリゴヌクレオチド

表 I V において、Ion Torrent PGM 配列決定プラットフォームを用いる DNA 鋳型ライブラリーに対するハイブリッド捕捉実験における使用のために、以下のオリゴヌクレオチドを設計した。T_m 増強基として LNA (「+C」または「+A」) を使用して、T_m 増強オリゴヌクレオチドを調製した。ホスホラミダイト化学法を使用して、全てのオリゴヌクレオチドを調製した。

【0345】

【表 4】

表 IV. Ion Torrent アダプターのための T_m 増強オリゴヌクレオチドブロッカー

配列番号	配列	LNA 数	T _m (°C) 750 mM (Na)	T _m (°C) 15 mM (Na)
29	Ion P1 top: CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGA T	0	89.2	62.3
30	Ion P1 top 11C: C+CA+CTA+CGC+CTC+CG+CTTTC+CT+CT+CTATGGG+ CAGT+CGGTGAT	11	>90	77.9
31	Ion P1 bot: ATCACC GACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTG GTT	0	89.6	62.1
32	Ion P1 bot 6C8A: AT+CAC+CGA+CTGC+CC+AT+AG+AG+AGGA+A+AG+CG G+AGG+CGT+AGTGGTT	14	>90	77.8
33	Ion A top: CCATCTCATCCCTGCGTGTCTCCGACTCAG	0	83.2	57.4
34	Ion A top 11C: C+CAT+CT+CAT+C+CCTG+CGTGT+CT+C+CGA+CT+CA G	11	>90	76.1
35	Ion A bot: CTGAGTCGGAGACACGCAGGGATGAGATGGTT	0	84.0	57.4
36	Ion A bot 5C5A: +CTG+AGT+CGG+AGA+CA+CG+CAGGG+ATG+AG+ATGG TT	10	>90	74.3

【0346】

表 I V は、ブロッカーとしてアダプター配列のいずれかの鎖を使用して T_m 増強オリゴヌクレオチドが設計され得る追加実施例を提供している。ブロッカーとしての T_m 増強オリゴヌクレオチドとしての使用のための好ましい鎖は、最も少ない T_m 増強基の含有で、修飾基 1 個当たり最大「ブロッキング力」（すなわち、最も大きい最適な増強 T_m 値）を提供するものである。この実施例は、T_m 増強オリゴヌクレオチドとして選択された鎖に依存して、LNA-C が、LNA-A と比較して T_m 増強基 1 個当たりに基づく優れた「ブロッキング力」を有することも示す。例えば、配列番号 30 および 32（配列番号 30 が好ましい。）および配列番号 34 および 36（配列番号 34 が好ましい。）を比較されたい。

【0347】

【実施例 6】 Illumina に基づく配列決定を用いるブロッキングオリゴヌクレオチドとしての T_m 増強オリゴヌクレオチドの使用。

【0348】

T_m 増強ブロッカーの有効性を検証するために、8 つの個別にバーコードされた（イン

デックスされた) ライブラリーを、同じ gDNA 供給源 NA07034 (Coriell Institute for Medical Research, Camden, NJ) から、Illumina TruSeq アダプター (01、03、05、06、08、09、10、11) (Illumina, San Diego, CA) を使用し、Illumina のライブラリー調製プロトコルに従って作製した。S220 焦点化ウルトラソニケーター (Covaris, Woburn, MA) を使用し、175W ピークパワー、デューティファクタ 2.0 および 200 サイクル/バーストで、1 マイクログラムのゲノム DNA を切断した。Agencourt AMPure XP システム (Beckman Coulter) を使用して、切断された DNA を精製にかけ、Illumina TruSeq DNA Sample Prep Kit を使用して、末端を修復し、A を尾部にした。

10

【0349】

表 V にリストされている様々なブロッカーを使用し、実施例 1 に概説されているプロトコルを使用し、各ライブラリー 500 ng および各ブロッキングオリゴヌクレオチド 1 ナノモルを用いて、ライブラリーの 7 つをハイブリダイゼーション捕捉に基づく濃縮にかけた。捕捉オリゴヌクレオチドプールは、265 個の遺伝子を標的にするおよそ 11,000 の 120mer 5' ビオチン化オリゴヌクレオチド (Lockdown (商標) プロープ、Integrated DNA Technologies, Coralville Iowa) からなっていた。捕捉/濃縮された配列を増幅し、標準化し、非濃縮ライブラリーを含めてプールした。Illumina MiSeq ベンチトップシーケンサーを使用して、プールされた材料を配列決定した。Galaxy 計算生物学プラットフォームを使用して、配列決定の結果を分析した。図 11 は、各インデックス化ライブラリーについて得られた独特の配列の数を示す。配列の数は 4.3 百万から 3.1 百万を範囲とし、任意の所定のブロッキングオリゴ配列に有意な低減は見られなかった。

20

【0350】

30

40

50

【表 5】

表V. Illumina TruSeqアダプターとともに使用されたオリゴヌクレオチド

配列番号	名称	配列
81	TruSeq P7 インデックス 5	GATCGGAAGAGCACACGTCTGAACTCCAGTCACACAGTGATCTCGTATGCCGTCTTCTGCTTG
23	TruSeq P5 ユニバーサル	AATGATACGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
82	TruSeq P7 相補体 インデックス 6 x I	CAAGCAGAAGACGGCATACGAGAT/ideoxyI//ideoxyI//ideoxyI//ideoxyI//ideoxyI//GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT/3SpC3/
83	TruSeq P5 ユニバーサル	AATGATACGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT/3SpC3/
84	TruSeq P7 相補体 中BNA 6-I	/5BNA-mC/AAG/iBNA-mC/AGAAGA/iBNA-mC/GG/iBNA-mC/ATA/iBNA-mC/GAGAT/ideoxyI//ideoxyI//ideoxyI//ideoxyI//ideoxyI//GTGA/iBNA-mC/TGGAGTT/iBNA-mC/AGA/iBNA-mC/GTGTG/iBNA-mC/T/iBNA-mC/TT/iBNA-mC//iBNA-mC/GAT/iBNA-mC/T/3SpC3/
85	TruSeq P5 ユニバーサル 中BNA	AATGATA/iBNA-mC/GG/iBNA-mC/GA/iBNA-mC/CAC/iBNA-mC/GAGAT/iBNA-mC/TA/iBNA-mC/ACT/iBNA-mC/TTT/iBNA-mC/CCTA/iBNA-mC/ACGA/iBNA-mC/GCT/iBNA-mC/TTC/iBNA-mC/GAT/iBNA-mC/T/3SpC3/
86	TruSeq P7 相補体 高BNA 6-I	/5BNA-mC/AAG/iBNA-mC/AGA/iBNA-A/GA/iBNA-mC/GG/iBNA-mC/ATA/iBNA-mC/G/iBNA-A/G/iBNA-A/T/ideoxyI//ideoxyI//ideoxyI//ideoxyI//ideoxyI//GTG/iBNA-A/iBNA-mC/TGG/iBNA-A/GTT/iBNA-mC/AG/iBNA-A/iBNA-mC/GTGTG/iBNA-mC/T/iBNA-mC/TT/iBNA-mC//iBNA-mC/G/iBNA-A/T/iBNA-mC/T/3SpC3/
87	TruSeq P5 ユニバーサル 高BNA	AATGATA/iBNA-mC/GG/iBNA-mC/GA/iBNA-mC/CAC/iBNA-mC/GAGAT/iBNA-mC/TA/iBNA-mC/A/iBNA-mC/T/iBNA-mC/TTT/iBNA-mC/C/iBNA-mC/TA/iBNA-mC/A/iBNA-mC/GA/iBNA-mC/G/iBNA-mC/T/iBNA-mC/TT/iBNA-mC//iBNA-mC/GAT/iBNA-mC/T/3SpC3/
88	P7 インデックス 6-I	GATCGGAAGAGCACACGTCTGAACTCCAGTCAC/ideoxyI//ideoxyI//ideoxyI//ideoxyI//ATCTCGTATGCCGTCTTCTGCTTG/3SpC3/
89	P5 相補体	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT/3SpC3/
90	P7 中BNA インデックス 6-I	GAT/iBNA-mC/GGAAGAG/iBNA-mC/ACACGT/iBNA-mC/TGAA/iBNA-mC/TCCAGT/iBNA-mC/A/iBNA-mC//ideoxyI//ideoxyI//ideoxyI//ideoxyI//AT/iBNA-mC/T/iBNA-mC/GTATGC/iBNA-mC/GTCTTCTG/iBNA-mC/TTG/3SpC3/
91	P5 相補体 中BNA	AGAT/iBNA-mC/GGAAGAG/iBNA-mC/GT/iBNA-mC/GTG TAGGGAAAGAGTGTAGAT/iBNA-mC/T/iBNA-mC/GGTGGT/iBNA-mC/G/iBNA-mC//iBNA-mC/GTAT/iBNA-mC/ATT/3SpC3/

10

20

30

40

50

配列番号	名称	配列
92	P7 高BNA インデックス 6-I	GAT/iBNA-mcC/GGAAGAG/iBNA-mcC/A/iBNA-mcC/A/iBNA-mcC/GT/iBNA-mcC/TGAA/iBNA-mcC/T/iBNA-mcC//iBNA-mcC/AGT/iBNA-mcC/A/iBNA-mcC/ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl/AT/iBNA-mcC/T/iBNA-mcC/GTATG/iBNA-mcC//iBNA-mcC/GT/iBNA-mcC/TT/iBNA-mcC/TG/iBNA-mcC/TTG
93	P5 相補体 高BNA	AG/iBNA-A/T/iBNA-mcC/GG/iBNA-A/iBNA-A/G/iBNA-A/G/iBNA-mcC/GT/iBNA-mcC/GTGT/iBNA-A/GGG/iBNA-A//iBNA-A//iBNA-A/G/iBNA-A/GTGT/iBNA-A/G/iBNA-A/T/iBNA-mcC/T/iBNA-mcC/GGTGGT/iBNA-mcC/G/iBNA-mcC/iBNA-mcC/GT/iBNA-A/T/iBNA-mcC//iBNA-A/TT/3SpC3/
94	PEblock_0-BNAs	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCCGTCGTCGTCATT
95	PEblock_8-BNAs	AGAT/iBNA-mcC/GGAAGAG/iBNA-mcC/GT/iBNA-mcC/GTGTAGGGAAAGAGTGTAGAT/iBNA-mcC/T/iBNA-mcC/GGTGGT/iBNA-mcC/GC/iBNA-mcC/GTAT/iBNA-mcC/ATT
96	PEblock_22-BNAs	AG/iBNA-A/T/iBNA-mcC/GG/iBNA-A/iBNA-A/G/iBNA-A/G/iBNA-mcC/GT/iBNA-mcC/GTGT/iBNA-A/GGG/iBNA-A//iBNA-A//iBNA-A/G/iBNA-A/GTG T/iBNA-A/G/iBNA-A/T/iBNA-mcC/T/iBNA-mcC/GGTGGT/iBNA-mcC/G/iBNA-mcC//iBNA-mcC/GT/iBNA-A/T/iBNA-mcC//iBNA-A/TT
97	InBlock_0-BNAs	GATCGGAAGAGCACACGTCTGAACTCCAGTCAC/ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl/ATCTCGTATGCCGTCTTCTGCTTG
98	Inblock_8-BNAs	GAT/iBNA-mcC/GGAAGAG/iBNA-mcC/ACACGTCTGAA/iBNA-mcC/TC CAGT/iBNA-mcC/A/iBNA-mcC//ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl/ATCT/iBNA-mcC/GTATGC/iBNA-mcC/GTCTT CTG/iBNA-mcC/TTG
99	Inblock_17-BNAs	GAT/iBNA-mcC/GGAAGAG/iBNA-mcC/A/iBNA-mcC/A/iBNA-mcC/GT/iBNA-mcC/TGAA/iBNA-mcC/T/iBNA-mcC//iBNA-mcC/AGT/iBNA-mcC/A/iBNA-mcC//ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl//ideoxyl/AT/iBNA-mcC/T/iBNA-mcC/GTATG/iBNA-mcC//iBNA-mcC/GT/iBNA-mcC/TT/iBNA-mcC/TG/iBNA-mcC/TTG

10

20

30

【0351】

1つの性能メトリックは、20-30倍（「20×-30×」）のカバレッジの深度で所望の標的配列にアラインメントする（マップする）配列の百分率である。図12は、非修飾DNAブロッキングオリゴヌクレオチドならびにT_m増強ブロッキングオリゴヌクレオチドについてのカバレッジの深度を示す。インデックス配列における分散について説明するため、一連の縮重塩基位置（全ての4個のヌクレオシド塩基）、またはデオキシイノシンなどのユニバーサル塩基のいずれかを使用するのが共通技術である。カバレッジ値の深度は表VIに示されている。T_m増強ブロッキングオリゴヌクレオチド（ユニバーサル塩基を含有する。）の性能は、同じユニバーサル塩基を含有する非T_m増強ブロッキングオリゴヌクレオチドと比較した場合、カバレッジの深度における全体的増加を20×および30×で示す。二重化配列決定読み取りを、分析より前に除いた。

40

【0352】

50

【表 6】

表VI. ブロッキングオリゴヌクレオチドを用いて得られたカバレッジ値の深度

バーコード:	インデックス 05	インデックス 06	インデックス 01	インデックス 03	インデックス 10	インデックス 08	インデックス 09
使用された ブロッキングオリ ゴ対:	P7 +	P7Comp 6xI +	P7Comp 中BNA 6xI +	P7Comp 高BNA 6xI +	P7 6xI +	P7 中BNA 6xI +	P7 高BNA 6xI +
	P5	P5	P5 中BNA	P5 高BNA	P5Comp	P5Comp 中BNA	P5Comp 高BNA
カバーされた 標的%							
> 1×							
2×	99.08%	98.99%	99.10%	99.11%	99.04%	99.12%	99.04%
10×	98.69%	97.78%	98.71%	98.68%	97.65%	98.84%	98.62%
20×	98.05%	92.12%	98.36%	98.13%	90.76%	98.46%	97.83%
30×	96.74%	81.09%	97.57%	97.10%	78.14%	97.63%	96.06%

【 0 3 5 3 】

別の性能メトリックは、目的の配列にアラインメントする配列の百分率（標的に対するパーセント）である。図 1 3 は、様々なブロッキングオリゴヌクレオチドを用いて得られたオン標的百分率を要約している。デオキシイノシンユニバーサル塩基を有する非修飾 D N A ブロッカーと比較して、デオキシイノシンユニバーサル塩基を有する T_m 増強ブロッカーで、有意な改善（1 6 – 1 9 %）が観察される。特異的インデックス配列を有するブロッキングオリゴヌクレオチドを含有する非 B N A 対デオキシイノシンを含有するブロッキングオリゴヌクレオチドを含有する B N A の間に、限定された改善（9 %）が観察された。さらに、ブロッキングオリゴヌクレオチドを含有する 1 3 個対 2 0 個の B N A の間に、有意な追加の改善が観察されなかった（3 9 . 8 6 % 対 3 8 . 0 9 %）。

【 0 3 5 4 】

第 3 の性能メトリックは、非標的配列を超える標的配列の濃縮の平均レベルであり、図 1 4 は、倍数濃縮レベルを要約している。非増強ブロッキングオリゴヌクレオチドと比較して、T_m 増強ブロッキングオリゴヌクレオチドでの倍数濃縮における有意な増加が観察される。濃縮の範囲は、デオキシイノシンを含有する非増強ブロッキングオリゴヌクレオチドについて 1 9 7 の低さから、T_m 増強ブロッキングオリゴヌクレオチドについて 5 8 0 倍の高さまでであった。

【 0 3 5 5 】

配列番号 9 5 および 9 8 ならびに配列番号 9 6 および 9 9 に対応する T_m 増強ブロッカーオリゴヌクレオチド対を持つ 8 – ヌクレオチドバーコードインデックス配列を有する I l l u m i n a ライブラリーを使用するハイブリッド捕捉実験において、匹敵できる傾向が得られた。配列番号 9 4 および 9 7 に対応する非修飾ブロッカーオリゴヌクレオチド対は、標的上の配列決定読み取りの約 5 0 % をもたらした。対照的に、配列番号 9 5 および 9 8 ならびに配列番号 9 6 および 9 9 に対応する T_m 増強ブロッカーオリゴヌクレオチド

対は、標的上の配列決定読み取りのそれぞれ約70%および約75%をもたらした。したがって、8個の T_m 増強基を有する T_m 増強ブロッカーは、約40% ($([70\% - 50\%] / 50\%) \times 100\%$) の、非修飾ブロッカーを超えるおよその相対増加を生み出した。そして、17-22個の T_m 増強基を有する T_m 増強ブロッカーは、約50% ($([75\% - 50\%] / 50\%) \times 100\%$) の非修飾ブロッカーを超えるおよその相対増加をもたらした。(図15を参照されたい。)

【0356】

下記に開示されている実施例A-Nは、図3に提供されている流れ図を介して描写されている、腫瘍試料の多重遺伝子分析のための方法についての実施形態の特色を示している。

【0357】

【実施例A】 腫瘍試料からの核酸単離

パラフィンブロックから切断された $3 \times 20 \mu\text{m}$ のセクションを $400 \mu\text{L}$ のBuffer FTLと、ボルテックスすることによって混合し、 90°C で15分間、 1.5 mL の遠心管内でインキュベートした。 $88-92^\circ\text{C}$ の範囲が、インキュベーションのための許容範囲であった。次いで、 $20 \mu\text{L}$ のプロテイナーゼKを用いて 55°C で6時間、および $10 \mu\text{L}$ のRNase (1 mg/mL)を用いて室温で5分間、試料をインキュベートした。次に、 $460 \mu\text{L}$ のBuffer BLおよび $500 \mu\text{L}$ の無水エタノールを試料に添加した。結果として得られた試料溶液を、さらに使用するまで室温で保持した。

【0358】

DNA結合のためのカラムを調製するため、 $100 \mu\text{L}$ の平衡緩衝剤をMicro Eluteカラムに添加し、カラムを $10,000 \times g$ で30秒間遠心分離した。上に記載されている試料溶液 $700 \mu\text{L}$ をMicro Eluteカラムに移し、カラムを $10,000 \times g$ で1分間遠心分離した。流体がMicro Eluteカラムを完全に通過しなかった場合には遠心分離ステップを反復した。残留している試料溶液を、上に記載されているのと同じやり方でMicro Eluteカラムに適用した。次いで、Micro Eluteカラムを $500 \mu\text{L}$ のBuffer HBで処理し、 $10,000 \times g$ で1分間遠心分離した。次に、エタノールで希釈された $700 \mu\text{L}$ のDNA Wash BufferをMicro Eluteカラム中に添加し、カラムを $10,000 \times g$ で1分間遠心分離した。エタノールで希釈された $700 \mu\text{L}$ のDNA Wash Bufferを再び使用してMicro Eluteカラムを洗浄し、 $10,000 \times g$ で1分間遠心分離し、 $>13,000 \times g$ で3分間遠心分離して、カラムを乾燥した。Micro Eluteカラムを、上部が除去された標準的 1.5 mL 遠心管中に配置した。 70°C に予備加熱された $50-75 \mu\text{L}$ のElution Bufferをカラム中に添加し、室温で3分間インキュベートした。カラムをコレクションチューブ内で $>13,000 \times g$ で1分間遠心分離した。 70°C に予備加熱した別の $50-75 \mu\text{L}$ のElution BufferをMicro Eluteカラム中に添加し、室温で3分間インキュベートした。カラムを再びコレクションチューブ中にて $>13,000 \times g$ で1分間遠心分離した。全溶液を新たな 1.5 mL の遠心管に移し、 -20°C で貯蔵した。

【0359】

FTL緩衝剤、プロテイナーゼK、BL Buffer、Equilibration Buffer、Micro Eluteカラム、Buffer HB、DNA Wash BufferおよびElution Bufferは、E. Z. N. A. (商標) FFPE DNA Kit (OMEGA bio-tek、NORCROSS、GA; カタログ番号D3399-00、D3399-01およびD3399-02)において提供された。

【0360】

ホルムアルデヒド固定またはパラホルムアルデヒド固定のパラフィン包埋(FFPE)組織から核酸(例えば、DNA)を単離するための追加の方法は、例えば、Cronin M.ら、(2004) Am J Pathol. 164 (1): 35-42; Masuda N.ら、(1999) Nucleic Acids Res. 27 (22): 4436-4443; Specht K.ら、(2001) Am J Pathol. 158 (2): 4

10

20

30

40

50

19-429、Ambion RecoverAll (商標) Total Nucleic Acid Isolation Protocol (Ambion、カタログ番号AM1975、2008年9月)、Maxwell (登録商標) 16 FFPE Plus LEV DNA Purification Kit Technical Manual (Promega Literature番号TM349、2011年2月)、およびQIAamp (登録商標) DNA FFPE Tissue Handbook (Qiagen、カタログ番号37625、2007年10月)に開示されている。RecoverAll (商標) Total Nucleic Acid Isolation Kitは、キシレンを昇温で使用して、パラフィン包埋試料およびガラス繊維フィルターを可溶化して核酸を捕捉する。Maxwell (登録商標) 16 FFPE Plus LEV DNA Purification Kitを、FFPE組織の1 μ mから10 μ mのセクションからのゲノムDNAの精製のために、Maxwell (登録商標) 16 Instrumentとともに使用する。DNAは、シリカクラッドの常磁性粒子(PMP)を使用して精製し、低溶出体積で溶出する。QIAamp (登録商標) DNA FFPE Tissue Kitは、ゲノムおよびミトコンドリアのDNAの精製のために、QIAamp (登録商標) DNA Micro技術を使用する。

【0361】

【実施例B.1】 DNAの剪断

循環チラー付きのCovaris (商標) E210機器を4℃に設定した。機器の水タンクに蒸留／脱イオン水を充填ラインのレベル「6」まで充填した。SonoLab (商標) ソフトウェアを起動し、指示された時にシステムがホーミング配列を実行するのを可能にした。試料を剪断する前に少なくとも45分間、機器のタンク中の水を脱気した。

【0362】

剪断するためのゲノムDNA試料を調製するため、試料を最初に、マイクロプレートリーダー(Spectramax M2、Molecular Devices)上のPicogreen (登録商標) アッセイ(Invitrogen)を使用して定量化した。濃度に基づいて、低いTE(10mMのTris、0.2mMのEDTA、pH8.0)を有する120 μ lの所望のインプットDNA(2ng/ μ l)を実験に使用した。100 μ lの個別の試料を、チューブの蓋の中の隔壁を介してCovaris MicroTUBE (Covaris カatalog番号520045)中にピペットでゆっくり取った。Covaris MicroTUBEを次いで、Covaris Eシリーズのチューブラック内に配置した。200bp剪断のため、設定は以下の通りであった: 10%のデューティサイクル、5の強度、200のサイクル／バースト、時間180秒、およびFrequency Sweepingモード。剪断した後、ミニ遠心機における適切なアダプターを使用して、Covaris MicroTUBEを短時間でスピンドウンし、剪断された試料を清潔な1.5ml微量遠心機チューブに移した。QIAGEN MinElute (登録商標) カラムを使用して、各剪断DNA試料を精製した。短時間で、5 $\times$ のQIAGEN PBI緩衝剤を、1.5ml微量遠心機チューブ中の試料に添加した(例えば、PBI緩衝剤500 μ lを試料100 μ lに添加した。)。各試料をボルテックスし、短時間でスピンドウンし、MinEluteスピнкаラムに移した。MinEluteスピнкаラムを13,000rpmで1分間遠心分離し、フロースルーを捨てた。750 μ lのQIAGEN PE緩衝剤をカラムに添加し、13,000rpmで1分間遠心分離し、フロースルーを捨てた。スピнкаラムを再び13,000rpmで1分間遠心分離し、清潔な1.5ml微量遠心機チューブに移した。カラムを2-3分間風乾した。第1の溶出のため、18 μ lのQIAGEN Elution Bufferを各カラムに添加し、2-3分間インキュベートし、次いで13,000rpmで1分間遠心分離した。第2の溶出のため、15 μ lのQIAGEN Elution Bufferを添加し、1分間インキュベートし、次いで13,000rpmで1分間遠心分離した。溶出液を回収し、スピнкаラムを捨てた。

【0363】

典型的には、200 ngがDNA剪断のために使用されるが、DNAの量は、20 ngから200 ngまたはこれより高い範囲であってよい。

【0364】

〔実施例B. 2〕 DNA剪断の代替

この実施例は、実施例2AからのDNA剪断のための代替方法を記載する。

【0365】

二本鎖ゲノムDNAを最初に一本鎖DNAに変性し、次いで、プライマー、DNAポリメラーゼ（例えば、Exo-DNAポリメラーゼ）、dNTP、および少量のddNTPと混合する。プライマー配列は、ランダム6量体、またはアダプター配列と5'末端でタグ付けされたランダム6量体であってよい。タグ付けされたランダム6量体増幅を使用して微量のDNAクローンおよび配列決定する方法は、例えばWong K. K.ら、Nucleic Acids Res. 1996; 24(19): 3778-83に記載されている。プライマー鋳型アニーリングおよびDNA合成を可能にする条件下で、反応物をインキュベートする。新たに合成された第1の鎖にddNTPが組み込まれると、DNA合成は終了する。合成された第1の鎖DNAの長さは、dNTP対ddNTPの比によって制御することができる。例えば、dNTP対ddNTPのモル比は、少なくとも約1000:1、約5000:1、または約10000:1である。第1の鎖合成後、短い断片（短い長さおよびddNTPを有するプライマーおよび合成された第1の鎖DNAなど）を、サイズ選択によって（例えば、サイズ選択スピンカラムを使用して）除去することができる。結果として得られた第1の鎖DNAを、プライマー（例えば、ランダム6量体またはアダプター配列でタグ付けされたランダム6量体）、DNAポリメラーゼ（例えば、Exo+DNAポリメラーゼ）、およびdNTPと混合する。Exo+DNAポリメラーゼを使用して、第1の鎖DNAから末端3'-ddNTPを除去することができ、または第2のプライミング部位上でプラント末端を生成することさえできる。反応物を次いで、プライマー鋳型アニーリングおよびDNA合成を可能にする条件下でインキュベートする。第2の鎖の合成後、結果として得られた二本鎖DNA断片を精製し、ライブラリー構築において直接的に使用することができる。代替として、二本鎖DNA断片は、アダプター配列を含有するプライマーを使用してPCR増幅することができるが、これらのアダプター配列が第1および第2の鎖合成のためのプライマーに含められていた場合である。PCR増幅のためのプライマーは、全配列および／またはバーコード配列も含むことができる。

【0366】

〔実施例C〕 ライブラリー調製

末端修復反応

末端修復試薬（NEB番号E6050L）を解凍し、末端修復マスターミックスを氷上で調製した。1試料当たり70 µlのマスターミックス調製するため、55 µlのヌクレアーゼフリー水を10 µlの10×のEnd Repair反応緩衝剤および5 µlのEnd Repair酵素ミックスと混合した。次いでマスターミックス70 µlを各剪断DNA試料30 µlに、氷上の96ウェルPCRプレート内で添加した。反応物を熱サイクラー内にて20℃で30分間インキュベートした。QIAGEN MinElute（登録商標）カラムを使用して、各試料を精製した。短時間で、5×のQIAGEN PBI緩衝剤を試料に（例えば、PBI緩衝剤500 µlを試料100 µlに添加した。）、1.5 ml微量遠心機チューブ内で添加した。各試料をボルテックスし、短時間でスピンドウンし、MinEluteスピンカラムに移した。MinEluteスピンカラムを13,000 rpmで1分間遠心分離し、フロースルーを捨てた。QIAGEN PE緩衝剤750 µlをカラムに添加し、13,000 rpmで1分間遠心分離し、フロースルーを捨てた。スピンカラムを再び13,000 rpmで1分間遠心分離し、清潔な1.5 ml微量遠心機チューブに移した。カラムを2-3分間風乾した。第1の溶出のため、22 µlのQIAGEN Elution Buffer（10 mMのTris、pH 8.5）を各カラムに添加し、2-3分間インキュベートし、次いで、13,000 rpmで1分間遠心分離した。第2の溶出のため、22 µlのQIAGEN Elution Buffer

を添加し、1分間インキュベートし、次いで、13,000rpmで1分間遠心分離した。溶出液を回収し、スピncラムを捨てた。

【0367】

3' A塩基付加

A塩基付加試薬（NEB番号E6053L）を氷上で解凍し、A塩基付加マスターミックスを氷上で調製した。1試料当たり10μlのマスターミックスを調製するため、2μlのヌクレアーゼフリー水を5μlの10×のdA-Tailing反応緩衝剤および3μlのKlenow Fragment（3'→5' exo-）と混合した。マスターミックス10μlを各精製末端修復DNA試料40μlに、氷上の96ウェルPCRプレート内で添加した。反応物を熱サイクラー内にて37℃で30分間インキュベートした。QIAGEN MinElute（登録商標）カラムを使用して、各試料を精製した。短時間で、5×のQIAGEN PBI緩衝剤を、1.5ml微量遠心機チューブ内の試料に添加した（例えば、PBI緩衝剤250μlを試料50μlに添加した。）。各試料をボルテックスし、短時間でスピncダウンし、MinEluteスピncラムに移した。MinEluteスピncラムを13,000rpmで1分間遠心分離し、フロースルーを捨てた。QIAGEN PE緩衝剤750μlをカラムに添加し、13,000rpmで1分間遠心分離し、フロースルーを捨てた。スピncラムを再び13,000rpmで1分間遠心分離し、清潔な1.5ml微量遠心機チューブに移した。カラムを2-3分間風乾した。第1の溶出のため、13μlのQIAGEN Elution Buffer（10mMのTris、pH8.5）を各カラムに添加し、2-3分間インキュベートし、次いで、13,000rpmで1分間遠心分離した。第2の溶出のため、13μlのQIAGEN Elution Bufferを添加し、1分間インキュベートし、次いで、13,000rpmで1分間遠心分離した。溶出液を回収し、スピncラムを捨てた。

【0368】

マルチプレックスアダプターのライゲーション

ライゲーション試薬（NEB番号E6056L）を解凍し、ライゲーションマスターミックスを氷上で調製した。1試料当たり36μlのマスターミックスを調製するため、12μlの5×のQuick Ligation反応緩衝剤を3.3μlのIllumina Multiplex Adaptor（15μM、Illuminaカタログ番号PE-400-1001中に含まれる。）に添加した（3.3μlのアダプター／1μgの出発インプットDNAを使用した。）。例えば、500ngのインプットDNAの1つの試料のため、アダプターを最初に水中で希釈し（2μlのアダプタープラス2μlのH₂O）、次いで、この希釈アダプターミックス3.3μl、ヌクレアーゼフリー水15.7μlおよびQuick T4 DNAリガーゼ5μlを、ライゲーション反応に添加した。>1μgの出発材料のため、アダプター>3.3μlを使用した。したがって、より少ない水を添加して、希釈アダプターミックスおよびヌクレアーゼフリー水の総体積を19μlに保持した。

【0369】

マスターミックス36μlおよび各dA尾部DNA試料24μlを氷上の96ウェルPCRプレートのウェルに添加した。反応物を熱サイクラー内にて25℃で30分間インキュベートした。QIAGEN MinElute（登録商標）カラムを使用して、各試料を精製した。短時間で、5×のQIAGEN PBI緩衝剤を、1.5ml微量遠心機チューブ内の試料に添加した（例えば、PBI緩衝剤300μlを試料60μlに添加した。）。各試料をボルテックスし、短時間でスピncダウンし、MinEluteスピncラムに移した。MinEluteスピncラムを13,000rpmで1分間遠心分離し、フロースルーを捨てた。QIAGEN PE緩衝剤750μlをカラムに添加し、13,000rpmで1分間遠心分離し、フロースルーを捨てた。スピncラムを再び13,000rpmで1分間遠心分離し、清潔な1.5ml微量遠心機チューブに移した。カラムを2-3分間風乾した。第1の溶出のため、20μlのQIAGEN Elution Buffer（10mMのTris、pH8.5）を各カラムに添加し、2-3分間インキ

ュベートし、次いで、13,000rpmで1分間遠心分離した。第2の溶出のため、20μlのQIAGEN Elution Bufferを添加し、1分間インキュベートし、次いで、13,000rpmで1分間遠心分離した。溶出液を回収し、スピンカラムを捨てた。

【0370】

PCR濃縮

PCR試薬を解凍し、PCRマスターミックスを氷上で調製した。1試料当たり62μlのマスターミックスのため、HF Buffer (Finnzyme、NEB カタログ番号F-531S)を有する50μlの2×のPhusion High Fidelityマスターミックス、8μlのヌクレアーゼフリー水、2μlのIllumina Primer 1.0 (25μM)、および2μlのIllumina Primer 2.0 (0.5μM)を使用した。次いでマスターミックス62μlを、適切なバーコードを有する2μlのIllumina Index Primer (25μM、Illumina カタログ番号PE-400-1001に含まれる。)および36μlのライゲートDNA試料と、96ウェルPCRプレート内で混合した。

【0371】

反応物を熱サイクラー内にて以下の通りにインキュベートした：

1サイクル	98℃	30秒
18サイクル	98℃	10秒
65℃	30秒	
72℃	30秒	
1サイクル	72℃	5分
4℃	保持	

【0372】

各PCR反応物を1.8×体積のAMPure XPビーズ (Agencourt ; Beckman Coulter Genomics カタログ番号A6388) でサイズ選択した。短時間で、1.8×のAMPure XPビーズを、1.5ml微量遠心機チューブ内の試料に添加し（例えば、180μlのビーズを100μlの試料に添加した。）、ボルテックスし、転倒回転混合しながら5分間インキュベートした。溶液が清澄になるまでチューブを磁石台上に配置した（2分）。磁石上に捕捉されたビーズを妨げることなく上澄みを捨てた。600μlの新たに作製された70%エタノールをビーズに添加し、1分間インキュベートし、続いてエタノールを除去した。600μlの新たに作製された70%エタノールの第2のアリコットをビーズに添加し、1分間インキュベートし、エタノールを除去した。チューブを磁石台上に戻して1-2分間置くことで、ビーズを再捕捉した。任意の残留しているエタノールを除去し、ビーズを室温で5-10分間風乾した。30μlのQIAGEN Elution Bufferをビーズに添加し、ボルテックスし、2分間インキュベートした。溶液が清澄になるまでチューブを磁石台上に戻して配置した（2分）。上澄みを新たな1.5mlチューブに移し、ビーズを捨てた。Q-PCRアッセイを使用して溶出されたDNA試料を定量化した。これらの定量化は、等モルにプールのを可能にして、プールされたハイブリッド捕捉選択内の各ライブラリーの等しい表示を確保する。

【0373】

【実施例D】 ハイブリッド選択

プールインデックス試料ライブラリー

インデックスされ、精製され、Q-PCRによって定量化されたライブラリーのプール（最大12plexまで）を氷上で作製した。等モルのプールを1.5ml微量遠心機チューブ内で調製して、ハイブリッド選択プロセスにおいて同等に各試料を表すことを確保した。これらのプールの各々に関するDNAの全インプットは、2000ngから500ngの範囲であってよい。典型的に、全インプットDNAは2000ngである。したがって、12個の試料がプールされるなら、各々166.67ngがプールされて、合計2

000 ngを達成することができる。2000 ngのライブラリープールの最終体積は、4 μ lであるはずである。インデックス化ライブラリーの濃度を変動することにより、プールはより大きい任意の体積で作製することができるが、次いでプールはspeed vacによって（低熱を使用して）乾燥し、4 μ lのヌクレアーゼフリー水中で再構成することができる。

【0374】

ライブラリー構築における収率が大きければ大きいほど、ライブラリーの複雑性が大きくなる。

ビオチン化RNAベイトへのプールDNAライブラリーのハイブリダイゼーション

Agilent SureSelect Target Enrichment Paired Endキット（番号G3360A-J）を、この実験に使用した。Hybridization Buffer番号3、SureSelect Block番号1、SureSelect Block番号2、Paired End Primer 1.0ブロック、Index Primer 1-12ブロック、RNaseブロック、およびビオチン化RNAベイトを氷上で解凍した。

【0375】

以下のマスターミックスを調製した：

a. ハイブリダイゼーション緩衝剤ミックス（1反応当たり13 μ l）：

- i. Hybridization Buffer番号1（Agilent）-25 μ l
- ii. Hybridization Buffer番号2（Agilent）-1 μ l
- iii. Hybridization Buffer番号3（Agilent）-10 μ l
- iv. Hybridization Buffer番号4（Agilent）-13 μ l

b. ブロッキングミックス（1反応当たり8 μ l）：

- i. SureSelect Block番号1（Agilent）-2.5 μ l
- ii. SureSelect Block番号2（Agilent）-2.5 μ l
- iii. Paired Endプライマー1.0ブロック（IDT、H₂Oを用いて200 μ Mに再懸濁される。）-1.5 μ l
- iv. Index Primer 1-12ブロック（IDT、H₂Oを用いて200 μ Mに再懸濁される。）-1.5 μ l

c. RNaseブロックの希釈

- i. テリトリ<3 Mbを有するカスタムビオチン化RNAベイト用：1 μ lのRNase Block（Agilent）を9 μ lの水中に希釈した。
- ii. ベイトテリトリ>3 Mbを有するカスタムベイト用：1 μ lのRNaseブロックを3 μ lの水中に希釈した（捕捉反応7 μ l当たり、まだ0.5 μ lのRNaseブロック）

d. ベイトミックス：（1反応当たり7 μ l）

- i. RNAベイト-2 μ l（ベイトテリトリ>3 Mbを有するベイトには、5 μ lベイトを使用した。）

- ii. 希釈RNaseブロック-5 μ l（ベイトテリトリ>3 Mbを有するベイトには、上記に示されている通りに希釈された2 μ lのRNaseブロックを使用した。）

【0376】

ハイブリダイゼーション緩衝剤ミックス、ブロッキングミックス、およびベイトミックス（単数または複数）を調製すると、ハイブリダイゼーション緩衝剤ミックスをボルテックスし、スピンドウンし、熱ブロック内で65℃に加熱した。ハイブリッド選択すべき各プール試料ライブラリー4 μ lを、ブロッキングミックス8 μ lと、96ウェルPCRプレート内で混合した。反応物を熱サイクラー内にて95℃で5分間インキュベートし、次いで65℃で保持した。プール試料ライブラリー／ブロッキングミックスを95℃で5分間および次いで65℃で2.5分間インキュベートした時、ベイトミックス（=ベイト

10

20

30

40

50

／RNAseブロックミックス)を熱ブロック内に65℃で2.5分間置いた。ハイブリダイゼーション緩衝剤含有チューブを迅速にスピンドウンし、次いで直ちに65℃の熱ブロックに戻した。加熱されたハイブリダイゼーション緩衝剤ミックス13μlを各試料ライブラリー／ブロックミックス中にピペットで取り、一方で96ウェルプレート熱サイクラー内に65℃で残した。ベイトミックスを2.5分間65℃でインキュベートするとすぐに、ベイトミックス7μlを各試料ライブラリー／ブロック／ハイブリダイゼーション緩衝剤ミックスに添加し、一方96ウェルプレート熱サイクラー内に65℃で残した。反応物(総体積は32μlであった。)を65℃で24時間熱サイクラー内にてインキュベートした。

【0377】

磁性ビーズの調製

SureSelect Wash Buffer 番号2を65℃で熱ブロック内にて予備加温した。Dynal MyOne Streptavidin T1ビーズ(Invitrogen)をボルテックスし、再懸濁した。Dynalビーズ(例えば、1200μlのSureSelect Binding Bufferを使用して、300μlのDynalビーズを調製した。)50μl当たり200μlのSureSelect Binding Bufferを添加することによって、ビーズを洗浄した。ビーズを5秒間ボルテックスし、短時間でスピンドウンする。ビーズを磁石台上に約15秒間または全てのビーズを捕捉するまで配置した。上澄みを除去し、捨てた。SureSelect Binding Bufferを用いてさらに2回洗浄を反復し、合計3回洗浄した。洗浄した後、Dynalビーズ50μl当たり200μlのSureSelect Binding Buffer中にビーズを再懸濁した(例えば、1200μlのSureSelect Binding Bufferを使用して、300μlのDynalビーズを調製した。)。再懸濁ビーズをボルテックスし、短時間でスピンドウンした。再懸濁ビーズ200μlを個別の1.5ml微量遠心機チューブ内にアリコットした。

【0378】

ハイブリッド捕捉DNAの選択

インキュベーションの24時間後に、65℃での熱サイクラー内のPCRプレートから、調製ビーズ200μlを含有するチューブ中に室温で、各ハイブリダイズ化試料を迅速にピペットで取った。試料およびビーズの混合物を5秒間ボルテックスし、ローテーター上に室温で30分間インキュベートして、適正な混合を確保した。次いでチューブを迅速にスピンドウンした。ビーズを磁石上に捕捉し(2分間)、上澄みを除去し、捨てた。低ストリンジェンシー洗浄のために、ビーズを500μlのSureSelect Wash Buffer 番号1中に再懸濁した。試料を5秒間ボルテックスし、15分間室温でインキュベートし、磁石から外した。試料を3-5分毎に5秒間ボルテックスした。チューブを迅速にスピンドウンした。ビーズを次いで磁石台上で2分間捕捉し、上澄みを除去し、捨てた。オフターゲット材料を除去するための高ストリンジェンシー洗浄のため、65℃に予備加熱されたSureSelect Wash Buffer 番号2でビーズを洗浄した。短時間で、ビーズを500μlの予備加温SureSelect Wash Buffer 番号2中に再懸濁し、ボルテックス上で5秒間混合して、ビーズを再懸濁した。ビーズを遠心機内にて短時間でスピンドウンし、5秒間室温で時折ボルテックス混合しながら65℃で10分間熱ブロック内にてインキュベートした。次いでビーズを遠心機内にて短時間でスピンドウンし、磁石上で2分間捕捉した。予備加温SureSelect Wash Buffer 番号2を用いて65℃でさらに2回洗浄を反復し、合計3回洗浄した。次いで洗浄緩衝剤を完全に除去し、50μlのSureSelect Elution Bufferをビーズに添加し、続いて、5秒間ボルテックスしてビーズを混合した。5秒間時折ボルテックス混合しながら、試料を10分間室温でインキュベートした。ビーズを遠心機内にて短時間でスピンドウンし、磁石台上で捕捉した。捕捉DNAを含有する上澄みを、新たな1.5ml微量遠心機チューブにピペットで取った。50μlのSureSelect Neutralization Bufferを捕捉DNAに添

加した。試料を5秒間ボルテックスし、遠心機内にて短時間でスピンドウンし、1.8×体積のAMPure XPビーズを使用して精製した。DNAを40 μlのヌクレアーゼフリー水中に溶出した。

【0379】

捕捉DNAのPCR濃縮

PCR試薬を解凍し、PCRマスターミックスを氷上で調製した。1試料当たりのマスターミックス60 μlのため、HF緩衝剤（NEB番号F-531S）を有する50 μlの2×のPhusion High Fidelityマスターミックスを、8 μlのヌクレアーゼフリー水、1 μlのQPCRプライマー1.1（H₂O中100 μM）、および1 μlのQPCRプライマー2.1（H₂O中100 μM）と混合した。Q-PCRのためのプライマー配列は：

QPCRプライマー1.1（IDTからHPLC精製された。）：

5' AATGATACGGCGACCAACCGAGAT3'（配列番号79）

QPCRプライマー2.1（IDTからHPLC精製された。）：

5' CAAGCAGAAGACGGCATACGA3'（配列番号80）

マスターミックス60 μlを各精製捕捉DNA試料40 μlに、96ウェルPCRプレート中で添加した。反応物を熱サイクラー内にて以下の通りにインキュベートした：

1 サイクル 98℃ 30秒

12 サイクル 98℃ 10秒

65℃ 30秒

72℃ 30秒

1 サイクル 72℃ 5分

4℃ 保持

【0380】

1.8×体積のAMPure XPビーズを用いて、PCR反応物各100 μlを精製し、溶出緩衝剤（10mMのTris、pH8.5）35 μl中に溶出した。Q-PCRアッセイを使用して、ハイブリッド選択／捕捉されたDNA試料を定量化した。Q-PCRアッセイは末端アダプターを検出し、読み取りは、適切なクラスター密度を得るためにどれだけの各試料を配列決定フローセル上に装填するべきかを示した。

【0381】

【実施例E】 方法

以下は、実施例に従って改変を同定するために使用された方法および実験条件の特定の実施形態を例証する。追加の転座スクリーニングは、例えば予備選択腫瘍試料から調製されたcDNAのいずれかのqRT-PCR分析を使用して行うことができる。

【0382】

アーカイブされた固定パラフィン包埋組織から単離されたDNAを使用してハイブリダイゼーション捕捉されたアダプターライゲーションベースのライブラリー上で、大規模平行DNA配列決定を行った。分析ツールの組合せを使用してデータを分析し、DNA改変呼び出しを割り当てた。凍結腫瘍から調製されたcDNAのqRT-PCR分析またはアーカイブされたFFPE標本のIHC判定のいずれかを使用して、追加の転座スクリーニングを行った。FFPE組織から単離されたRNAを使用する両方の新規な転座の発現を確認するため、大規模平行cDNA配列決定を行った。再編成の体細胞起源を確認するためインデックスNSCLC患者について、血液からの整合された正常の参照ゲノムDNAを配列決定した。

【0383】

ゲノムDNA配列決定

アーカイブされたホルマリン固定パラフィン包埋（FFPE）腫瘍標本からのDNA；NSCLC患者から24個を使用して、145個の癌遺伝子の2574個のエクソンの配列決定を行った。ゲノムDNAを使用するアダプターライゲーション方法、続いて、最適化されたRNAハイブリダイゼーション捕捉プローブを用いるハイブリダイゼーション選

10

20

30

40

50

択によって、配列決定ライブラリーを構築した (Agilent SureSelect カスタムキット)。253×の平均深度に対して36×36つの対合読み取りを使用して、HiSeq 2000機器 (Illumina) 上での配列決定を行った。腫瘍組織からの変異呼び出しのために最適化されたツールの組合せを使用して、塩基置換、インデル、コピー数改変およびゲノム再編成のデータ加工および変異割り当てを行った。

【0384】

cDNA配列決定

Roche High Pureキットを使用して単一の5-10μmのFFPE組織セクションから抽出された全RNAからcDNAを生成し、SuperScript (登録商標) III First-Strand Synthesis System (Invitrogen) によってランダム6量体プライマーを有するcDNAに逆転写した。二本鎖cDNAをNEBNext (登録商標) mRNA Second Strand Synthesis Module (New England Biolabs) で作製し、FFPE DNA試料に関してライブラリー構築、ハイブリッド捕捉および配列決定へのインプットとして使用した。分析ツールの組合せを用いて、発現レベルの分析を行った。

【0385】

【実施例F】 マルチプレックス分析のための例示的選択遺伝子およびバリエーション

この実施例は、マルチプレックス分析のための遺伝子、バリエーションおよび癌型の選択を要約する4つの例示的表を提供する。

【0386】

10

20

30

40

50

【表 7】

表1: マルチプレックス分析のための例示的選択遺伝子およびバリエーション、関連癌型、および優先コドンのリスト。

Hugo遺伝子	遺伝子 カテゴリー	癌型	優先コドン
ABL1	優先1	白血病(例えば、慢性骨髄性白血病(CML)、急性骨髄球性白血病(AML)、急性リンパ芽球性白血病(ALL))	315
AKT1	優先1	乳癌、直腸結腸癌、卵巣癌	
ALK	優先1	リンパ腫(例えば、非ホジキンリンパ腫、未分化大細胞リンパ腫(ALCL))、炎症性筋線維芽細胞腫瘍	
APC	優先1	直腸結腸癌、髄芽腫、ミスマッチ修復癌症候群	1114, 1338, 1450, 1556
AR	優先1	前立腺癌	
BRAF	優先1	肺癌、非ホジキンリンパ腫、直腸結腸癌、甲状腺癌、メラノーマ	600
CDKN2A	優先1	メラノーマ、膀胱癌、リー・フラウメニ症候群、肺癌(例えば、非小細胞肺癌(NSCLC))、扁平細胞癌腫、網膜芽細胞腫、星状細胞腫	
CEBPA	優先1	白血病(例えば、急性骨髄球性白血病(AML)、急性骨髄性白血病(AML)、単芽球性白血病)、網膜芽細胞腫	
CTNNB1	優先1	直腸結腸癌、卵巣癌、前立腺癌、肝臓癌(例えば、肝芽腫(HB)、肝細胞癌腫(HCC))、石灰化上皮腫、髄芽腫、唾液腺多形腺腫	32, 33, 34, 37, 41, 45
EGFR	優先1	肺癌、扁平細胞癌腫、神経膠芽腫、神経膠腫、直腸結腸癌	719, 746-750, 768, 790, 858, 861
ERBB2	優先1	胃癌、神経膠腫、卵巣癌、肺癌	
ESR1	優先1	乳癌、子宮内膜癌、子宮内膜腺癌腫、平滑筋腫、乳管癌腫	
FGFR1	優先1	白血病、リンパ腫	
FGFR2	優先1	乳癌、前立腺癌	
FGFR3	優先1	膀胱癌、子宮頸癌、多発性骨髄腫	
FLT3	優先1	白血病(例えば、急性骨髄球性白血病(AML)、急性前骨髄球性白血病、急性リンパ芽球性白血病)	835
HRAS	優先1	ハースル細胞甲状腺癌腫、膀胱癌、メラノーマ、直腸結腸癌	12, 13, 61
JAK2	優先1	白血病(例えば、慢性リンパ芽球性白血病(CLL)、急性リンパ芽球性白血病(ALL)、慢性骨髄性白血病(CML)、急性骨髄性白血病(AML))	617
KIT	優先1	消化管ストローマ腫瘍(GIST)、精巣腫瘍、白血病(例えば、急性骨髄球性白	816

10

20

30

40

50

Hugo遺伝子	遺伝子 カテゴリー	癌型	優先コドン
		血病(AML)), 肥満細胞腫、間葉腫瘍、腺様嚢胞癌腫、肺癌(例えば、小細胞肺癌)、リンパ腫(例えば、バーキットリンパ腫)	
KRAS	優先1	白血病(例えば、急性骨髄性白血病(AML)、若年性骨髄単球性白血病(JMML)), 直腸結腸癌、肺癌	12, 13, 61
MET	優先1	胃癌、肝細胞癌腫(HCC)、遺伝性乳頭状腎癌腫(HPRC)、肺癌(例えば、非小細胞肺癌)、乳頭状甲状腺癌腫、神経膠腫、食道腺癌腫、骨肉腫、子宮内膜癌、扁平細胞癌腫、メラノーマ、乳癌	
MLL	優先1	白血病(例えば、急性リンパ芽球性白血病(ALL)、急性骨髄球性白血病(AML))	
MYC	優先1	慢性リンパ性白血病(CLL)、バーキットリンパ腫、プラズマ細胞腫	
NF1	優先1	白血病(例えば、若年性骨髄単球性白血病(JMML)), 神経線維腫	
NOTCH1	優先1	扁平細胞癌腫、白血病(例えば、急性リンパ芽球性白血病(ALL)), 延髄性甲状腺癌腫、リンパ腫(例えば、胸腺リンパ腫、T細胞リンパ腫)	1575, 1601
NPM1	優先1	リンパ腫(例えば、非ホジキンリンパ腫、未分化大細胞リンパ腫、未分化リンパ腫)、白血病(例えば、急性前骨髄球性白血病、急性骨髄性白血病(AML))	
NRAS	優先1	白血病(例えば、若年性骨髄単球性白血病(JMML)、急性骨髄球性白血病(AML)、急性リンパ芽球性白血病)、メラノーマ	12, 13, 61
PDGFRA	優先1	消化管ストローマ腫瘍(GIST)、白血病(例えば、慢性好酸球性白血病(CEL)、急性リンパ性白血病(ALL)), 間葉腫瘍	
PIK3CA	優先1	直腸結腸癌、乳癌、卵巣癌、肝細胞癌腫、頭頸部扁平細胞癌腫(HNSCC)、未分化甲状腺癌腫、子宮内膜癌、胆嚢腺癌腫、神経膠芽腫	88, 542, 545, 546, 1047, 1049
PTEN	優先1	頭頸部扁平細胞癌腫(HNSCC)、子宮内膜癌、神経膠腫、前立腺癌、神経膠芽腫	130, 173, 233, 267
RB1	優先1	網膜芽細胞腫、膀胱癌、骨肉腫、肺癌(例えば、小細胞肺癌、非小細胞肺癌)、白血病(例えば、急性リンパ芽球性白血病(ALL))	
RET	優先1	直腸結腸癌、延髄甲状腺癌腫、多発性	918

10

20

30

40

50

Hugo遺伝子	遺伝子 カテゴリー	癌型	優先コドン
		腫瘍症2B型、褐色細胞腫、多発性腫瘍症2A型、甲状腺乳頭状癌腫、甲状腺癌、網膜芽細胞腫	
TP53	優先1	TP53は、癌の約60%で頻繁に変異または不活性化されている、例えば、食道扁平細胞癌腫、リー・フラウメニ症候群、頭頸部扁平細胞癌腫(HNSCC)、肺癌、遺伝性副腎皮質癌腫、星状細胞腫、扁平細胞癌腫、膀胱癌、直腸結腸癌、神経膠芽腫、網膜芽細胞腫	175, 245, 248, 273, 306
ABL2	癌遺伝子	急性骨髄球性白血病(AML)	
AKT2	癌遺伝子	卵巣癌、膀胱癌	
AKT3	癌遺伝子	メラノーマ、神経膠腫、子宮癌、前立腺癌、口腔癌、卵巣癌	
ARAF	癌遺伝子	血管免疫芽細胞性T細胞リンパ腫、エールリッヒ腹水腫瘍	
ARFRP1	癌遺伝子	乳癌	
ARID1A	癌遺伝子	神経芽細胞腫、急性リンパ芽球性白血病(ALL)、神経内分泌腫瘍	
ATM	癌遺伝子	白血病(例えば、T細胞前リンパ球性白血病(T-PLL))、リンパ腫、髄芽腫、神経膠腫	
ATR	癌遺伝子	臍胸関連リンパ腫、T細胞リンパ腫	
AURKA	癌遺伝子	喉頭扁平細胞癌腫、卵巣癌、膀胱癌、頭頸部扁平細胞癌腫(HNSCC)、喉頭癌腫、食道扁平細胞癌腫(ESCC)、膀胱癌	
AURKB	癌遺伝子	直腸結腸癌、星状細胞腫、上衣腫瘍、神経膠腫、食道扁平細胞癌腫(ESCC)、急性骨髄球性白血病(AML)	
BCL2	癌遺伝子	リンパ腫、結腸直腸腺癌腫、食道扁平細胞癌腫(ESCC)、滑膜肉腫、白血病	
BCL2A1	癌遺伝子	肺肉芽腫、胃腺腫、バーキットリンパ腫、耳下腺腫、カポジ肉腫、胃癌、大腸癌	
BCL2L1	癌遺伝子	頭頸部扁平細胞癌腫、神経膠芽腫、中皮腫、膀胱癌、腺癌腫肺	
BCL2L2	癌遺伝子	脳癌、白血病、リンパ腫、結腸直腸腺癌腫、直腸結腸癌、腺腫、子宮頸部扁平細胞癌腫	
BCL6	癌遺伝子	リンパ腫、白血病	
BRCA1	癌遺伝子	乳癌、卵巣癌	
BRCA2	癌遺伝子	乳癌、卵巣癌、膀胱癌	
CARD11	癌遺伝子	リンパ腫	
CBL	癌遺伝子	リンパ腫、白血病	
CCND1	癌遺伝子	慢性リンパ芽球性白血病(CLL)、B細胞急性リンパ芽球性白血病(B-ALL)、乳癌	

10

20

30

40

50

Hugo遺伝子	遺伝子 カテゴリー	癌型	優先コドン
CCND2	癌遺伝子	網膜芽細胞腫、マントル細胞リンパ腫、T細胞急性リンパ芽球性白血病(T-ALL)、バーキットリンパ腫、精巢生殖細胞腫瘍、卵巣顆粒膜細胞腫瘍、多発性骨髄腫	
CCND3	癌遺伝子	網膜芽細胞腫、マントル細胞リンパ腫、未分化大細胞リンパ腫、リンパ腫(非ホジキン)、B細胞リンパ腫、喉頭扁平細胞癌腫、無痛性リンパ腫、マル細胞腺腫	
CCNE1	癌遺伝子	乳癌、卵巣癌、膀胱癌、網膜芽細胞腫	
CDH1	癌遺伝子	胃癌、小葉癌腫、扁平細胞癌腫、浸潤性腺管癌腫、浸潤性小葉癌腫	
CDH2	癌遺伝子	メラノーマ、悪性中皮腫、胸膜中皮腫、線維形成性メラノーマ、肺腺癌腫、類内膜腫瘍、中皮腫、膀胱癌、食道扁平細胞癌腫(ESCC)	
CDH20	癌遺伝子	乳癌	
CDH5	癌遺伝子	肉芽腫、類上皮肉腫	
CDK4	癌遺伝子	メラノーマ	
CDK6	癌遺伝子	急性リンパ芽球性白血病(ALL)	
CDK8	癌遺伝子	大腸癌、肺癌、直腸癌、急性リンパ芽球性白血病(ALL)	
CDKN2B	癌遺伝子	白血病、網膜芽細胞腫、喉頭扁平細胞癌腫	
CDKN2C	癌遺伝子	甲状腺癌腫、下垂体腺腫、乏突起神経膠腫、脚臓内分泌腫瘍、多発性骨髄腫、肝芽腫、リンパ腫瘍、多発性内分泌腫瘍症型1、未分化乏突起神経膠腫	
CHEK1	癌遺伝子	白血病、大腸癌	
CHEK2	癌遺伝子	乳癌	
CRKL	癌遺伝子	白血病、リンパ腫	
CRLF2	癌遺伝子	白血病	
DNMT3A	癌遺伝子	精巢生殖細胞腫瘍、リンパ肉腫、肝細胞癌腫、唾液腺腫瘍	
DOT1L	癌遺伝子	白血病	
EPHA3	癌遺伝子	横紋筋肉腫、リンパ腫、前立腺癌、肝細胞癌腫、白血病、メラノーマ	
EPHA5	癌遺伝子	神経膠芽腫、乳癌、星状細胞腫、ウィルムス腫瘍、神経膠腫	
EPHA6	癌遺伝子	乳癌	
EPHA7	癌遺伝子	多形神経膠芽腫(GBM)、大腸癌、十二指腸癌、副甲状腺腫瘍、前立腺癌	
EPHB1	癌遺伝子	直腸結腸癌、胚性癌腫、胃癌、奇形癌腫、粘液性癌腫	
EPHB4	癌遺伝子	頭頸部扁平細胞癌腫(HNSCC)、脳癌、子宮内膜癌、卵巣癌	

10

20

30

40

50

Hugo遺伝子	遺伝子 カテゴリー	癌型	優先コドン
EPHB6	癌遺伝子	神経芽細胞腫、メラノーマ、非小細胞肺癌(NSCLL)	
ERBB3	癌遺伝子	乳癌、非小細胞肺癌(NSCLC)、膀胱癌、浸潤性腺管癌腫、肺腺癌腫、類内膜癌腫、毛様細胞性星状細胞腫	
ERBB4	癌遺伝子	乳癌、髄芽腫、子宮頸部扁平細胞癌腫、前立腺癌、白血病	
ERG	癌遺伝子	前立腺癌、ユーイング肉腫、白血病、前立腺癌	
ETV1	癌遺伝子	前立腺癌、乳癌、ユーイング肉腫、線維形成性小円形細胞腫瘍、粘液様脂質肉腫、明細胞肉腫	
ETV4	癌遺伝子	乳癌、卵巣癌、扁平細胞癌腫舌、ユーイング肉腫	
ETV5	癌遺伝子	神経節膠腫、脳腫瘍	
ETV6	癌遺伝子	白血病、先天性線維肉腫、分泌性癌腫、骨髄異形成症候群	
EWSR1	癌遺伝子	ユーイング肉腫、明細胞肉腫、線維形成性小円形細胞腫瘍、骨外性粘液様軟骨肉腫、粘液様脂質肉腫、血管腫様繊維状組織球腫	
EZH2	癌遺伝子	前立腺癌、胆嚢腺癌腫、乳癌、膀胱癌、胃癌、ユーイング肉腫	
FANCA	癌遺伝子	白血病	
FBXW7	癌遺伝子	直腸結腸癌、子宮内膜癌、T細胞急性リンパ芽球性白血病(T-ALL)	
FGFR4	癌遺伝子	下垂体腫瘍、前立腺癌、肺癌、星状細胞腫、横紋筋肉腫、下垂体腺腫、線維腺腫	
FLT1	癌遺伝子	乳癌、前立腺癌	
FLT4	癌遺伝子	肺癌、カポシ肉腫、胃癌、リンパ管腫、扁平細胞癌腫	
FOXP4	癌遺伝子	リンパ腫、脳腫瘍	
GATA1	癌遺伝子	ダウン症候群の巨核芽球白血病	
GNA11	癌遺伝子	乳癌	
GNAQ	癌遺伝子	ぶどう膜メラノーマ	
GNAS	癌遺伝子	下垂体腺腫	
GPR124	癌遺伝子	大腸癌	
GUCY1A2	癌遺伝子	乳癌	
HOXA3	癌遺伝子	乳癌	
HSP90AA1	癌遺伝子	リンパ腫、骨髄腫	
IDH1	癌遺伝子	多形神経膠芽腫(GBM)	
IDH2	癌遺伝子	多形神経膠芽腫(GBM)	
IGF1R	癌遺伝子	ユーイング肉腫、乳癌、ぶどう膜メラノーマ、副腎皮質癌腫、膀胱癌	
IGF2R	癌遺伝子	胃腸腫瘍、肝臓癌	
IKBKE	癌遺伝子	乳癌	

10

20

30

40

50

Hugo遺伝子	遺伝子 カテゴリー	癌型	優先コドン
IKZF1	癌遺伝子	リンパ腫、白血病	
INHBA	癌遺伝子	赤白血病、バレット化生、食道腺癌腫、顆粒膜細胞腫瘍、性索間質性腫瘍、肺腺癌腫、褐色細胞腫、クルケンベルグ腫瘍、卵巣癌	
IRS2	癌遺伝子	高インスリン血症、子宮平滑筋肉腫	
JAK1	癌遺伝子	白血病、卵巣癌、乳癌	
JAK3	癌遺伝子	急性リンパ芽球性白血病(ALL)	
JUN	癌遺伝子	皮膚癌、白血病	
KDR	癌遺伝子	非小細胞肺癌(NSCLC)、血管肉腫	
LRP1B	癌遺伝子	肺癌、胃癌、食道癌	
LTK	癌遺伝子	リンパ腫、乳癌	
MAP2K1	癌遺伝子	前立腺癌、胃癌	
MAP2K2	癌遺伝子	脾癌、腸腫瘍	
MAP2K4	癌遺伝子	脾癌、乳癌、直腸結腸癌	
MCL1	癌遺伝子	多発性骨髄腫、白血病、リンパ腫	
MDM2	癌遺伝子	肉腫、神経膠腫、直腸結腸癌	
MDM4	癌遺伝子	多形神経膠芽腫(GBM)、膀胱癌、網膜芽細胞腫	
MEN1	癌遺伝子	副甲状腺腫瘍	
MITF	癌遺伝子	メラノーマ	
MLH1	癌遺伝子	直腸結腸癌、子宮内膜癌、卵巣癌、CNS癌	
MPL	癌遺伝子	骨髄増殖性疾患(MPD)	
MRE11A	癌遺伝子	乳癌、リンパ腫	
MSH2	癌遺伝子	直腸結腸癌、子宮内膜癌、卵巣癌	
MSH6	癌遺伝子	直腸結腸癌	
MTOR	癌遺伝子	リンパ腫肺癌、腎癌、腎明細胞癌腫、神経膠腫	
MUTYH	癌遺伝子	直腸結腸癌	
MYCL1	癌遺伝子	小細胞肺癌(SCLC)	
MYCN	癌遺伝子	神経芽細胞腫	
NF2	癌遺伝子	髄膜腫、聴神経腫瘍、腎癌	
NKX2-1	癌遺伝子	肺癌、甲状腺癌、腺癌腫	
NTRK1	癌遺伝子	乳頭甲状腺癌	
NTRK3	癌遺伝子	先天性線維肉腫、分泌性乳癌	
PAK3	癌遺伝子	肺癌	
PAX5	癌遺伝子	非ホジキンリンパ腫(NHL)、急性リンパ芽球性白血病(ALL、例えば、B細胞ALL)	
PDGFRB	癌遺伝子	骨髄増殖性疾患(MPD)、急性骨髄球性白血病(AML)、慢性骨髄性白血病(CML)、慢性骨髄単球性白血病(CMML)	
PIK3R1	癌遺伝子	神経膠芽腫、卵巣癌、直腸結腸癌	
PKHD1	癌遺伝子	脾癌	
PLCG1	癌遺伝子	頭頸部癌、白血病	
PRKDC	癌遺伝子	神経膠腫、神経膠芽腫、胃癌、卵巣癌	

10

20

30

40

50

Hugo遺伝子	遺伝子 カテゴリー	癌型	優先コドン
PTCH1	癌遺伝子	皮膚基底細胞、髄芽腫	
PTPN11	癌遺伝子	若年性骨髄単球性白血病(JMML)、急性骨髄球性白血病(AML)、骨髄異形成症候群(MDS)	
PTPRD	癌遺伝子	肺癌、皮膚扁平細胞癌腫、神経膠芽腫、神経芽細胞腫	
RAF1	癌遺伝子	毛様細胞性星状細胞腫	
RARA	癌遺伝子	白血病	
RICTOR	癌遺伝子	大腸癌、リンパ腫、神経膠腫、乳癌	
RPTOR	癌遺伝子	乳癌、前立腺癌	
RUNX1	癌遺伝子	急性骨髄球性白血病(AML)、前駆B細胞急性リンパ芽球性白血病(プレB-ALL)、T細胞急性リンパ芽球性白血病(T-ALL)	
SMAD2	癌遺伝子	食道扁平細胞癌腫(ESCC)	
SMAD3	癌遺伝子	皮膚癌、絨毛癌腫	
SMAD4	癌遺伝子	膵癌、大腸癌	
SMARCA4	癌遺伝子	非小細胞肺癌(NSCLC)	
SMARCB1	癌遺伝子	悪性ラブドイド	
SMO	癌遺伝子	皮膚基底細胞癌	
SOX10	癌遺伝子	乏突起神経膠腫	
SOX2	癌遺伝子	胚性癌腫、生殖細胞腫瘍	
SRC	癌遺伝子	肉腫、大腸癌、乳癌	
STK11	癌遺伝子	非小細胞肺癌(NSCLC)、膵癌	
TBX22	癌遺伝子	乳癌	
TET2	癌遺伝子	骨髄異形成症候群(MDS)	
TGFBR2	癌遺伝子	肺癌、胃癌、大腸癌	
TMPRSS2	癌遺伝子	前立腺癌	
TOP1	癌遺伝子	急性骨髄球性白血病(AML)	
TSC1	癌遺伝子	過誤腫、腎細胞癌	
TSC2	癌遺伝子	過誤腫、腎細胞癌	
USP9X	癌遺伝子	白血病	
VHL	癌遺伝子	腎癌、血管腫、褐色細胞腫	
WT1	癌遺伝子	ウィルムス腫瘍、線維形成性小円形細胞腫瘍	
ABCB1	PGx遺伝子		
ABCC2	PGx遺伝子		
ABCC4	PGx遺伝子		
ABCG2	PGx遺伝子		
C1orf144	PGx遺伝子		
CYP1B1	PGx遺伝子		
CYP2C19	PGx遺伝子		
CYP2C8	PGx遺伝子		
CYP2D6	PGx遺伝子		
CYP3A4	PGx遺伝子		
CYP3A5	PGx遺伝子		
DPYD	PGx遺伝子		

10

20

30

40

50

Hugo遺伝子	遺伝子 カテゴリー	癌型	優先コドン
ERCC2	PGx遺伝子		
ESR2	PGx遺伝子		
FCGR3A	PGx遺伝子		
GSTP1	PGx遺伝子		
ITPA	PGx遺伝子		
LRP2	PGx遺伝子		
MAN1B1	PGx遺伝子		
MTHFR	PGx遺伝子		
NQO1	PGx遺伝子		
NRP2	PGx遺伝子		
SLC19A1	PGx遺伝子		
SLC22A2	PGx遺伝子		
SLCO1B3	PGx遺伝子		
SOD2	PGx遺伝子		
SULT1A1	PGx遺伝子		
TPMT	PGx遺伝子		
TYMS	PGx遺伝子		
UGT1A1	PGx遺伝子		
UMPS	PGx遺伝子		

10

20

【 0 3 8 7 】

「優先 1」は、選択遺伝子または遺伝子産物の最も高い優先を指す。

【 0 3 8 8 】

「癌遺伝子」は、優先 1 に対して優先でない癌関連遺伝子または遺伝子産物を指す。

【 0 3 8 9 】

「PGx 遺伝子」は、薬理遺伝学および薬理ゲノミクス（PGx）に重要である遺伝子を指す。

【 0 3 9 0 】

30

40

50

【表 8】

表1A: 追加の例示的選択遺伝子およびバリエーション、関連癌型、優先コドン、作用性カテゴリーおよび潜在的治療。

Hugo 遺伝子	遺伝子 カテゴリー	癌型	優先 コドン	作用性 カテゴリー	理由
ASXL1	優先1	多発性骨髄腫(MM)		D	予後(陰性MDS)
BACH1	優先1	乳房		C	PARP阻害剤
BAP1	優先1	ぶどう膜メラノーマ、乳房、NSCLC		C	PARP阻害剤
BARD1	優先1	乳房		C	PARP阻害剤
BLM	優先1	白血病、リンパ腫、皮膚扁平細胞、他の癌		C	
BRIP1	優先1	急性骨髄球性白血病(AML)、白血病、乳房		C	PARP阻害剤
CDKN1B	優先1	乳房		D	
CREBBP	優先1	急性リンパ芽球性白血病(ALL)、AML、DLBCL、B細胞非ホジキンリンパ腫(B-NHL)		D	
DDR2	優先1	NSCLC		C	ダサチニブ
EMSY	優先1	乳房		C	PARP阻害剤
FANCC	優先1	AML、白血病		C	PARP阻害剤
FANCD2	優先1	AML、白血病		C	PARP阻害剤
FANCE	優先1	AML、白血病		C	PARP阻害剤
FANCF	優先1	AML、白血病		C	PARP阻害剤
FANCG	優先1	AML、白血病		C	PARP阻害剤
FANCL	優先1	AML、白血病		C	PARP阻害剤
HGF	優先1	MM		C	抵抗性
NFKB1	優先1	乳房		D	可能な生命予後不良
NOTCH2	優先1	辺縁帯リンパ腫、DLBCL		D	-
PALB2	優先1	ウィルムス腫瘍、髄芽腫、AML、乳房		C	PARP阻害剤
PBRM1	優先1	明細胞腎癌腫、乳房		E	HDAC阻害剤?
PDK1	優先1	NSCLC		C	PDK1阻害剤
PIK3R2	優先1	NSCLC		C	PI3K-PATHWAY阻害剤
RAD50	優先1	乳房		C	PARP阻害剤
RAD51	優先1	乳房		C	PARP阻害剤
ROS1	優先1	神経膠芽腫、NSCLC		C	
SF3B1	優先1	MDS、CML、ALL、膵臓、乳房		E	
SPOP	優先1	悪性黒色腫		E	

10

20

30

40

50

Hugo 遺伝子	遺伝子 カテゴリー	癌型	優先 コード	作用性 カテゴリー	理由
ACVR1B	癌遺伝子	膵臓、乳房		E	
ALOX12B	癌遺伝子	多発性骨髄腫(MM)		E	
ATRX	癌遺伝子	膵臓神経内分泌腫瘍		E	
AXL	癌遺伝子	非小細胞肺癌(NSCLC)、MM		E	
BCOR	癌遺伝子	乳房		E	
BCORL1	癌遺伝子	乳房		E	
C17orf39	癌遺伝子	乳房		E	
CASP8	癌遺伝子	乳房		E	
CBFB	癌遺伝子	AML		E	
CD22	癌遺伝子	NSCLC、乳房		E	
CD79A	癌遺伝子	びまん性大細胞型B細胞リンパ腫(DLBCL)		E	
CD79B	癌遺伝子	DLBCL		E	
CDC73	癌遺伝子	副甲状腺		E	
CDK12	癌遺伝子	卵巣		E	
CHUK	癌遺伝子	結腸直腸		E	
CRBN	癌遺伝子	上気道消化管		E	
CSF1R	癌遺伝子	NSCLC		E	
CTCF	癌遺伝子	乳房		E	
CTNNA1	癌遺伝子	乳房		E	
CUL4A	癌遺伝子	白血病		E	
CUL4B	癌遺伝子	白血病		E	
CYP17A1	癌遺伝子	乳房		E	
DAXX	癌遺伝子	膵臓神経内分泌腫瘍		E	
DIS3	癌遺伝子	MM		E	
EP300	癌遺伝子	結腸直腸、乳房、膵臓、AML、ALL、DLBCL		E	
ERCC2	癌遺伝子	皮膚基底細胞、皮膚扁平細胞、メラノーマ		E	
FAM46C	癌遺伝子	MM		E	
FGF1	癌遺伝子	乳房		E	
FGF10	癌遺伝子	乳房		E	
FGF12	癌遺伝子	乳房		E	
FGF14	癌遺伝子	乳房		E	
FGF19	癌遺伝子	乳房		E	
FGF23	癌遺伝子	乳房		E	
FGF3	癌遺伝子	乳房		E	
FGF4	癌遺伝子	乳房		E	
FGF6	癌遺伝子	乳房		E	
FGF7	癌遺伝子	乳房		E	

10

20

30

40

50

Hugo 遺伝子	遺伝子 カテゴリー	癌型	優先 コード	作用性 カテゴリー	理由
FOXO2	癌遺伝子	卵巣の顆粒膜細胞腫瘍	134	E	
GATA2	癌遺伝子	AML、慢性骨髄性白血病(CML、芽球形 質転換)		E	
GATA3	癌遺伝子	乳房		E	
GRAF	癌遺伝子	AML、骨髄異形成症候群(MDS)		E	
GRIN2A	癌遺伝子	悪性黒色腫		E	
GSK3B	癌遺伝子	NSCLC		E	
HLA-A	癌遺伝子	MM		E	
IGF1	癌遺伝子	乳房		E	
IGF2	癌遺伝子	乳房		E	
IL7R	癌遺伝子	T細胞急性リンパ芽球性白血病(T-ALL)		E	
INSR	癌遺伝子	NSCLC、神経膠芽腫、胃		E	
IRF4	癌遺伝子	多発性骨髄腫(MM)		E	
KDM4C	癌遺伝子	卵巣、乳房		E	
KDM5A	癌遺伝子	AML		E	
KDM6A	癌遺伝子	腎性、食道扁平細胞癌腫(SCC)、MM		E	
KEAP1	癌遺伝子	NSCLC		E	
KLHL6	癌遺伝子	慢性リンパ球白血病(CLL)		E	
LMO1	癌遺伝子	T細胞急性リンパ芽球性白血病(T-ALL)、神経芽細胞腫		E	
LRP6	癌遺伝子	NSCLC、悪性黒色腫		E	
LRRK2	癌遺伝子	卵巣、NSCLC		E	
MAGED1	癌遺伝子	MM		E	
MAP3K1	癌遺伝子	乳房		E	
MAP3K13	癌遺伝子	乳房		E	
MLL2	癌遺伝子	髄芽腫、腎性		E	
MLST8	癌遺伝子	乳房		E	
MYD88	癌遺伝子	活性化B細胞様DLBCL(ABC-DLBCL)		E	
MYST3	癌遺伝子	乳房		E	
NCOR1	癌遺伝子	乳房		E	
NFE2L2	癌遺伝子	NSCLC、頭頸部扁平細胞癌腫(HNSCC)		E	
NFKB1A	癌遺伝子	乳房		E	
NOTCH3	癌遺伝子	NSCLC、乳房		E	
NOTCH4	癌遺伝子	NSCLC、乳房		E	
NSD1	癌遺伝子	AML		E	
NTRK2	癌遺伝子	腎性、NSCLC		E	
NUP93	癌遺伝子	乳房		E	
PAK7	癌遺伝子	NSCLC、悪性黒色腫		E	

10

20

30

40

50

Hugo 遺伝子	遺伝子 カテゴリー	癌型	優先 コード	作用性 カテゴリー	理由
PHLPP2	癌遺伝子	卵巣、神経膠芽腫、NSCLC		E	
PHOX2B	癌遺伝子	神経芽細胞腫		E	
PIK3C2G	癌遺伝子	NSCLC		E	
PIK3C3	癌遺伝子	NSCLC		E	
PIK3CG	癌遺伝子	NSCLC		E	
PNRC1	癌遺伝子	MM		E	
PRDM1	癌遺伝子	DLBCL		E	
PRKAR1A	癌遺伝子	副腎、甲状腺		E	
PRSS8	癌遺伝子	乳房		E	
PTCH2	癌遺伝子	悪性黒色腫		E	
PTK2	癌遺伝子	NSCLC、神経膠芽腫		E	
PTK2B	癌遺伝子	NSCLC、乳房		E	
REL	癌遺伝子	ホジキンリンパ腫		E	
RHEB	癌遺伝子	NSCLC、結腸直腸		E	
ROCK1	癌遺伝子	乳房		E	
RUNXT1	癌遺伝子	NSCLC、結腸直腸		E	
SETD2	癌遺伝子	明細胞腎癌腫		E	
SH2B3	癌遺伝子	骨髄異形成症候群(MDS)		E	
SOCS1	癌遺伝子	DLBCL		E	
SPEN	癌遺伝子	腺様嚢胞癌腫		E	
STAG2	癌遺伝子	神経膠芽腫		E	
STAT3	癌遺伝子	乳房		E	
STAT4	癌遺伝子	乳房		E	
STK12	癌遺伝子	PNET、NSCLC		E	
SUFU	癌遺伝子	髄芽腫		E	
TBX23	癌遺伝子	乳房		E	
TBX3	癌遺伝子	乳房		E	
TNFAIP3	癌遺伝子	辺縁帯B細胞リンパ腫、ホジキンリンパ腫、原発性縦隔B細胞リンパ腫		E	
TNFRSF14	癌遺伝子	濾胞性リンパ腫		E	
TNFRSF17	癌遺伝子	腸T細胞リンパ腫		E	
TNKS	癌遺伝子	NSCLC		E	
TNKS2	癌遺伝子	メラノーマ、乳房		E	
TRRAP	癌遺伝子	結腸直腸、神経膠芽腫		E	
TYK2	癌遺伝子	NSCLC、乳房		E	
XBP1	癌遺伝子	MM		E	
XP01	癌遺伝子	慢性リンパ球白血病(CLL)		E	
ZNF217	癌遺伝子	乳房		E	
ZNF703	癌遺伝子	乳房		E	

【0391】

作用性カテゴリーは、下に記載されている通りに分類されている。表1Bは、異なる癌型における例示的改変に対する異なるカテゴリーの用途の要約を提供する。

【0392】

カテゴリーA：承認された／標準的な治療に対する感受性または抵抗性を予測する承認された／標準的な改変

転移性大腸癌におけるKRAS G13D

乳癌におけるERBB2増幅

非小細胞肺癌におけるEGFR L858R

10

20

30

40

50

【0393】

カテゴリーB：特異的実験治療のための包含または除外の基準である改変

大腸癌、肺癌または乳癌におけるKRAS G13D

メラノーマ、大腸癌または肺癌におけるBRAF V600E

メラノーマにおけるNRAS Q61K

乳癌におけるPIK3CA H1047R

乳癌におけるFGFR1増幅

乳癌におけるPTEN二対立遺伝子不活性化

乳癌または膵癌におけるBRCA1二対立遺伝子不活性化

【0394】

カテゴリーC：標準治療または実験治療に対する感受性または抵抗性を予測する限定された証拠（早期臨床的データ、矛盾する臨床的データ、前臨床データ、理論）を有する改変

大腸癌におけるKRAS Q61H（早期臨床）

乳癌におけるPIK3CA H1047R（矛盾する臨床）

大腸癌におけるBRAF V600E（矛盾する臨床）

肺癌におけるERBB2変異または増幅（事例レポート）

肺癌におけるBRAF D594G（前臨床）

乳癌におけるFGFR1増幅（前臨床）

乳癌におけるATM二対立遺伝子不活性化（前臨床）

大腸癌におけるTSC1二対立遺伝子不活性化（前臨床）

乳癌におけるATR二対立遺伝子不活性化（理論）

肉腫におけるBRAF V600E変異（理論）

【0395】

カテゴリーD：癌の特別なサブタイプにおける予後または診断の有用性を有する改変

大腸癌におけるMSH2二対立遺伝子不活性化（強い臨床的証拠）

大腸癌におけるBRAF V600E（強い臨床的証拠）

肺癌におけるKRAS G13D（強い臨床的証拠）

乳癌におけるBRCA1不活性化（強い臨床的証拠）

【0396】

カテゴリーE：明確な臨床的意味合いのない癌（すなわち、ドライバー変異）における明確な生物学的有意性を有する改変

大腸癌におけるAPC二対立遺伝子不活性化

乳癌におけるTP53二対立遺伝子不活性化

メラノーマにおけるMITF増幅

卵巣癌におけるARID1A

【0397】

カテゴリーF：癌における公知の生物学的有意性のない改変

公知癌遺伝子における新規な改変

治療の標的

公知癌遺伝子のオルソログ

【0398】

10

20

30

40

50

【表 9】

表1B: 異なる癌型における改変の例示的分類

		A	B	C	D	E
KRAS G13D	大腸癌	×	×		×	×
KRAS G13D	肺癌		×		×	×
KRAS G13D	乳癌		×			×
NRAS Q61K	メラノーマ		×	×		×
KRAS Q61H	大腸癌		×	×		×
BRAF V600E	メラノーマ		×			×
BRAF V600E	大腸癌		×	×	×	×
BRAF V600E	肺癌		×			×
BRAF D594G	肺癌			×		×
PIK3CA H1047R	乳癌		×	×		×
PIK3CA H1047R	大腸癌		×	×		×
EGFR L858R	肺癌	×				×
EGFR T790M	肺癌	×	×			×
ERBB2増幅	乳癌	×				×
BRCA1 二対立遺伝子不活性化	乳癌		×	×	×	×
BRCA2 二対立遺伝子不活性化	膀胱癌		×	×	×	×
ATM 二対立遺伝子不活性化	乳癌			×		×
TSC 二対立遺伝子不活性化	大腸癌			×		×
PTEN 二対立遺伝子不活性化	大腸癌			×		×
PTEN 二対立遺伝子不活性化	乳癌		×	×		×
VHL 二対立遺伝子不活性化	腎臓癌				×	×
MSH2 二対立遺伝子不活性化	大腸癌				×	×
ATR 二対立遺伝子不活性化	乳癌			×		×
MYC増幅	乳癌			×		×

10

20

【 0 3 9 9 】

30

40

50

【表 10】

表2: 薬理遺伝学および薬理ゲノミクス(PGx)に関連の例示的選択遺伝子。

遺伝子	部位	変異	効果
ABCB1	chr7:86976581	3853C>T	Ida/AraCで処理されたAsian AMLにおけるより良好な生存;パクリタキセルで処理された乳癌患者における生存
ABCB1	chr7:86998554	2677G>T/A	タキサン、白金およびGI毒性への応答;Ida/AraCで処理されたAsian AMLにおけるより良好な生存
ABCC2	chr10:101610761		ドセタキセル誘発の白血球減少
ABCC4	chr13:94613416		6MP毒性
ABCG2	chr4:89252551		MTX
ABCG2	chr4:89271347	q141K	ゲフィチニブ後の下痢症
ABCG2	chr4:89274403		MTX
C1orf144	chr1:16578662		ダウノルビシンからの毒性
CYP1B1	chr2:38151707	CYP1B1*3	ダウノルビシンからの毒性;パクリタキセルで処置された乳癌患者における生存
CYP2C19	chr10:96509051	CYP2C19*17	タモキシフェンからの改善利益
CYP2C19	chr10:96511647	CYP2C19*17	タモキシフェンからの改善利益
CYP2C8	chr10:96786964	461delV	パクリタキセル代謝
CYP2C8	chr10:96788739	K399R	パクリタキセル代謝
CYP2C8	chr10:96808096		パクリタキセル代謝
CYP2C8	chr10:96808109		パクリタキセル代謝
CYP2C8	chr10:96817020		パクリタキセル代謝
CYP2D6	chr22:40853554	CYP2D6: 3183 G>A	CYP2D6*29、タンザニア人において存在
CYP2D6	chr22:40853749	CYP2D6: 2988 G>A	CYP2D6*41 (IM)
CYP2D6	chr22:40853887	CYP2D6: 2850 C>T	CYP2D6*2 (EM)
CYP2D6	chr22:40854122	CYP2D6: 2613-2615 del AGA	CYP2D6*9 (不明な機能?)
CYP2D6	chr22:40854188	CYP2D6: 2549 del A	CYP2D6*3
CYP2D6	chr22:40854891	CYP2D6: 1846 G>A	CYP2D6*4
CYP2D6	chr22:40855030	CYP2D6: 1707 del T	CYP2D6*6
CYP2D6	chr22:40855078	CYP2D6: 1659G>A	CYP2D6*29、タンザニア人において存在
CYP2D6	chr22:40855716	CYP2D6: 1023 C>T	CYP2D6*17において存在
CYP2D6	chr22:40856638	CYP2D6: 100C>T	CYP2D6*10(原因)および*4(関連)において存在
CYP3A4	chr7:99196395		
CYP3A4	chr7:99196460		
CYP3A4	chr7:99197606		
CYP3A4	chr7:99204017		
CYP3A4	chr7:99204029	CYP3A4*16B	日本人におけるパクリタキセル代謝
CYP3A4	chr7:99205328		

10

20

30

40

50

遺伝子	部位	変異	効果
CYP3A4	chr7:99205363		
CYP3A4	chr7:99219597		
CYP3A4	chr7:99220032	CYP3A4*1B	ドセタキセルの、より大きいクリアランス
CYP3A5	chr7:99088330		
CYP3A5	chr7:99100771		
CYP3A5	chr7:99108475		
DPYD	chr1:97688202	DPYD*2A	5FUに対する毒性
DPYD	chr1:97753983	DPYD*5	5FUに対する毒性
DPYD	chr1:97937679	496A>G	5FU、Xeloda毒性
DPYD	chr1:98121473	DPYD*9A	5FUに対する毒性
ERCC2	chr19:50546759	2251A>C	アジア人における5FU後の再発
ESR1	chr6:152205074		タモキシフェン誘発高コレステロール血症
ESR2	chr14:63769569		タモキシフェン誘発高コレステロール血症
FCGR3A	chr1:159781166	V158F	セツキシマブに対する応答
FGFR4	chr5:176452849	GLY388ARG	
GSTP1	chr11:67109265	I105V	多重化学療法に対する抵抗性
GSTP1	chr11:67110155	A114V	I105Vとの不明な連鎖不均衡
ITPA	chr20:3141842		6MP毒性
LRP2	chr2:169719231		シスプラチンからの聴器毒性と関連
MAN1B1	chr9:139102689		ダウノルビンからの毒性
MTHFR	chr1:11777044		MTX
MTHFR	chr1:11777063		MTX
MTHFR	chr1:11778965	677C>T	MTX
NQO1	chr16:68302646	NQO1*2	急速な分解(シスプラチン、ドキソルビン); アントラサイクリンで処置された乳癌における生存不良
NRP2	chr2:206360545		ダウノルビンからの毒性
SLC19A1	chr21:45782222		MTX
SLC22A2	chr6:160590272	Ala270Ser	低減されたシスプラチン腎毒性
SLCO1B3	chr12:20936961		ドセタキセル誘発白血球減少
SOD2	chr6:160033862	V16A	シクロホスファミドで処置された乳癌における低い生存
SULT1A1	chr16:28524986		
SULT1A1	chr16:28525015		
SULT1A1	chr16:28528073		
SULT1A1	chr16:28528301		
TPMT	chr6:18247207	TPMT*3B	プリン毒性
TPMT	chr6:18238897		6MP毒性
TPMT	chr6:18238991		6MP毒性
TPMT	chr6:18251934		6MP毒性
TYMS	chr18:647646	28bp タンデム重複	5FUに対する毒性
TYMS	chr18:663451	6bp欠失	5FUに対する毒性
UGT1A1	chr2:234255266		イリノテカンからの貧血
UGT1A1	chr2:234255709		イリノテカンからの血小板減少症
UGT1A1	chr2:234330398		UGT1A1*60
UGT1A1	chr2:234330521		UGT1A1*93
UGT1A1	chr2:234333620		UGT1A1*28
UGT1A1	chr2:234333883		UGT1A1*6
UGT1A1	chr2:234334358		UGT1A1*27
UMPS	chr3:125939432	Gly213Ala	5FUに対する毒性

10

20

30

40

【 0 4 0 0 】

50

【表 1 1】

表3: 固形腫瘍における転座変異と関連の例示的選択遺伝子

Hugo遺伝子	遺伝子カテゴリー	転座パートナー	癌型
ACSL3	優先1	ETV1	前立腺
ALK	優先1	NPM1, TPM3, TFG, TPM4, ATIC, C LTC, MSN, ALO17, CARS, EML4	ALCL、NSCLC、神経芽細胞腫
BRAF	優先1	AKAP9, KIAA1549	メラノーマ、結腸直腸、乳頭状甲状腺、境界性ov、非小細胞肺癌(NSCLC)、胆管腺癌腫、毛様細胞性星状細胞腫
C15orf21	優先1	ETV1	前立腺
CANT1	優先1	ETV4	前立腺
CCND1	優先1	IGH, FSTL3	CLL、B-ALL、乳房
DDX5	優先1	ETV4	前立腺
ELK4	優先1	SLC45A3	前立腺
EML4	優先1	ALK	NSCLC
EP300	優先1	MLL, RUNXBP2	結腸直腸、乳房、脾臓、AML
ERG	優先1	EWSR1, TMPRSS2, ELF4, FUS, HE RPUD1	ユーイング肉腫、前立腺、AML
ETV1	優先1	EWSR1, TMPRSS2, SLC45A3, C15orf21, HNRNPA2B1, ACSL3	ユーイング肉腫、前立腺
ETV4	優先1	EWSR1, TMPRSS2, DDX5, KLK2, CANT1	ユーイング肉腫、前立腺癌腫
ETV5	優先1	TMPRSS2, SCL45A3	前立腺
FGFR3	優先1	IGH@, ETV6	膀胱、MM、T細胞リンパ腫
HERPUD1	優先1	ERG	前立腺
HNRNPA2B1	優先1	ETV1	前立腺
KLK2	優先1	ETV4	前立腺
RET	優先1	H4, PRKAR1A, NCOA4, PCM1, GOLGA5, TRIM33, KTN1, TRIM27, HOOK3	延髄甲状腺、乳頭状甲状腺、褐色細胞腫
ROS1	優先1	GOPC, ROS1	神経膠芽腫、NSCLC
SLC45A3	優先1	ETV1, ETV5, ELK4, ERG	前立腺
TMPRSS2	優先1	ERG, ETV1, ETV4, ETV5	前立腺
AKAP9		BRAF	乳頭状甲状腺
ASPSR1		TFE3	肺軟部肉腫
ATF1		EWSR1, FUS	軟部の悪性黒色腫、血管腫様組織球腫
BRD3		NUT	若い人々の致死的中線癌腫
BRD4		NUT	若い人々の致死的中線癌腫
C12orf9		LPP	脂肪腫
CD74		ROS1	NSCLC
CDH11		USP6	動脈瘤性骨嚢胞
CHCHD7		PLAG1	唾液腺腫
CHN1		TAF15	骨外性粘液様軟骨肉腫
CIC		DUX4	軟組織肉腫
CMKOR1		HMGA2	脂肪腫
COL1A1		PDGFB, USP6	隆起性皮膚線維肉腫、動脈瘤性骨嚢胞
COX6C		HMGA2	子宮平滑筋腫

10

20

30

40

50

Hugo遺伝子	遺伝子カテゴリー	転座パートナー	癌型
CREB1		EWSR1	明細胞肉腫、血管腫様繊維状組織球腫
CREB3L2		FUS	線維粘液様肉腫
CRTC3		MAML2	唾液腺粘膜表皮
CTNNB1		PLAG1	結腸直腸、卵巣、肝芽腫、他、多形性唾液腺腫
D10S170		RET, PDGFRB	乳頭状甲状腺、CML
DDIT3		FUS	脂質肉腫
DUX4		CIC	軟組織肉腫
ELKS		RET	乳頭状甲状腺
ETV6		NTRK3, RUNX1, PDGFRB, ABL1, MN1, ABL2, FACL6, CHIC2, ARNT, JAK2, EVI1, CDX2, STL, HLXB9, MDS2, PER1, SYK, TTL, FGFR3, PAX5	先天性線維肉腫、多発性白血病およびリンパ腫、分泌性乳房、MDS、ALL
EWSR1		FLI1, ERG, ZNF278, NR4A3, FEV, ATF1, ETV1, ETV4, WT1, ZNF384, CREB1, POU5F1, PBX1	ユーイング肉腫、線維形成性小円形細胞腫瘍、ALL、明細胞肉腫、肉腫、筋上皮腫
FEV		EWSR1, FUS	ユーイング肉腫
FLI1		EWSR1	ユーイング肉腫
FOXO1A		PAX3	肺胞横紋筋肉腫
FUS		DDIT3, ERG, FEV, ATF1, CREB3L2	脂質肉腫、AML、ユーイング肉腫、血管腫様繊維状組織球腫、線維粘液様肉腫
GOLGA5		RET	乳頭状甲状腺
HEI10		HMGA2	子宮平滑筋腫
HMGA1		?	微小濾胞状甲状腺腺腫、様々な良性間葉腫瘍
HMGA2		LHFP, RAD51L1, LPP, HEI10, COX6C, CMKOR1, NFIB	脂肪腫
HOOK3		RET	乳頭状甲状腺
JAZF1		SUZ12	子宮内膜間質性腫瘍
KTN1		RET	乳頭状甲状腺
LHFP		HMGA2	脂肪腫
LIFR		PLAG1	唾液腺腫
LPP		HMGA2, MLL, C12orf9	脂肪腫、白血病
MAML2		MECT1, CRTC3	唾液腺粘膜表皮
MECT1		MAML2	唾液腺粘膜表皮
MN1		ETV6	AML、髄膜腫
MYB		NFIB	腺様嚢胞癌腫
MYC		IGK, BCL5, BCL7A, BTG1, TRA, IGH	他の癌において増幅されたパーキットリンパ腫、B-CLL
NCOA1		PAX3	肺胞性横紋筋肉腫
NCOA4		RET	乳頭状甲状腺
NFIB		MYB, HGMA2	腺様嚢胞癌腫、脂肪腫
NONO		TFE3	乳頭状腎癌
NR4A3		EWSR1	骨外性粘液様軟骨肉腫
NTRK1		TPM3, TPR, TFG	乳頭状甲状腺

10

20

30

40

50

Hugo遺伝子	遺伝子カテゴリー	転座パートナー	癌型
NTRK3		ETV6	先天性線維肉腫、分泌性乳房
NUT		BRD4, BRD3	若い人々の致死的中線癌腫
OMD		USP6	動脈瘤性骨嚢胞
PAX3		FOXO1A, NCOA1	肺胞横紋筋肉腫
PAX7		FOXO1A	肺胞横紋筋肉腫
PAX8		PPARG	濾胞性甲状腺
PBX1		TCF3, EWSR1	プレB-ALL、筋上皮腫
PCM1		RET, JAK2	乳頭状甲状腺、CML、MPD
PDGFB		COL1A1	DFSP
PDGFRA		FIP1L1	GIST、特発性過好酸性症候群
PLAG1		TCEA1, LIFR, CTNNB1, CHCHD7	唾液腺腫
POU5F1		EWSR1	肉腫
PPARG		PAX8	濾胞性甲状腺
PRCC		TFE3	乳頭状腎
PRKAR1A		RET	乳頭状甲状腺
PRO1073		TFEB	腎細胞癌腫(幼児期類上皮)
RAD51L1		HMGA2	脂肪腫、子宮平滑筋腫
RAF1		SRGAP3	毛様細胞性星状細胞腫
SFPQ		TFE3	乳頭状腎細胞
SRGAP3		RAF1	毛様細胞性星状細胞腫
SS18		SSX1, SSX2	滑膜肉腫
SS18L1		SSX1	滑膜肉腫
SSX1		SS18	滑膜肉腫
SSX2		SS18	滑膜肉腫
SSX4		SS18	滑膜肉腫
SUZ12		JAZF1	子宮内膜間質性腫瘍
TAF15		TEC, CHN1, ZNF384	骨外性粘液様軟骨肉腫、ALL
TCEA1		PLAG1	唾液腺腫
TCF12		TEC	骨外性粘液様軟骨肉腫
TFE3		SFPQ, ASPSCR1, PRCC, NONO, CLTC	乳頭状腎性の、肺胞軟部肉腫、腎臓
TFEB		ALPHA	腎臓(幼児期類上皮)
TFG		NTRK1, ALK	乳頭状甲状腺、ALCL、NSCLC
THRAP3		USP6	動脈瘤性骨嚢胞
TPM3		NTRK1, ALK	乳頭状甲状腺、ALCL
TPR		NTRK1	乳頭状甲状腺
TRIM27		RET	乳頭状甲状腺
TRIM33		RET	乳頭状甲状腺
USP6		COL1A1, CDH11, ZNF9, OMD	動脈瘤性骨嚢胞
ZNF278		EWSR1	ユーイング肉腫
ZNF331		?	濾胞性甲状腺腺腫
ZNF9		USP6	動脈瘤性骨嚢胞

10

20

30

40

【 0 4 0 1 】

50

【表 1 2】

表4: 血液悪性腫瘍における転座変異と関連の例示的選択遺伝子。

Hugo遺伝子	遺伝子カテゴリー	転座パートナー	癌型
ABL1	優先1	BCR, ETV6, NUP214	CML、ALL、T-ALL
ALK	優先1	NPM1, TPM3, TFG, TPM4, ATIC, C LTC, MSN, ALO17, CARS, EML4	ALCL、NSCLC、神経芽細胞腫
BCL2	優先1	IGH	NHL、CLL
BCL6	優先1	IG座位, ZNFN1A1, LCP1, PIM1, TFR, MHC2TA, NACA, HSPCB, HSPCA, HIST1H4I, IL21R, POU2AF1, ARHH, EIF4A2, SFRS3	NHL、CLL
CCND1	優先1	IGH, FSTL3	CLL、B-ALL、乳房
CREBBP	優先1	MLL, MORF, RUNXBP2	AL、AML
FGFR1	優先1	BCR, FOP, ZNF198, CEP1	MPD、NHL
FGFR3	優先1	IGH, ETV6	膀胱、MM、T細胞リンパ腫
JAK2	優先1	ETV6, PCMI, BCR	ALL、AML、MPD、CML
MLL	優先1	MLL, MLLT1, MLLT2, MLLT3, MLLT4, MLLT7, MLLT10, MLLT6, ELL, EPS15, AF1Q, CREBBP, SH3GL1, FNBP1, PNUTL1, MSF, GPHN, GMP, SSH3BP1, ARHGEF12, GAS7, FOXO3A, LAF4, LCX, SEPT6, LPP, CBFA2T1, GRAF, EP300, PICALM, HEAB	AML、ALL
PDGFRA	優先1	FIP1L1	GIST、特発性過好酸性症候群
RARA	優先1	PML, ZNF145, TIF1, NUMA1, NPM1	APL
SEPT6		MLL	AML
ABL2		ETV6	AML
AF15Q14		MLL	AML
AF1Q		MLL	ALL
AF3p21		MLL	ALL
AF5q31		MLL	ALL
ALO17		ALK	ALCL
ARHGEF12		MLL	AML
ARHH		BCL6	NHL
ARNT		ETV6	AML
ATIC		ALK	ALCL
BCL10		IGH	MALT
BCL11A		IGH	B-CLL
BCL11B		TLX3	T-ALL
BCL3		IGH	CLL
BCL5		MYC	CLL
BCL7A		MYC	BNHL
BCL9		IGH, IGL	B-ALL
BCR		ABL1, FGFR1, JAK2	CML、ALL、AML
BIRC3		MALT1	MALT
BTG1		MYC	BCLL
CARS		ALK	ALCL
CBFA2T1		MLL, RUNX1	AML
CBFA2T3		RUNX1	AML

10

20

30

40

50

Hugo遺伝子	遺伝子カテゴリー	転座パートナー	癌型
CBFB		MYH11	AML
CBL		MLL	AML、JMML、MDS
CCND2		IGL	NHL、CLL
CCND3		IGH	MM
CDK6		MLLT10	ALL
CDX2		ETV6	AML
CEP1		FGFR1	MPD、NHL
CHIC2		ETV6	AML
CLTC		ALK, TFE3	ALCL、腎臓
CLTCL1		?	ALCL
DDX10		NUP98	AML*
DDX6		IGH	B-NHL
DEK		NUP214	AML
EIF4A2		BCL6	NHL
ELF4		ERG	AML
ELL		MLL	AL
ELN		PAX5	B-ALL
EP300		MLL, RUNXBP2	結腸直腸、乳房、膵臓、AML
EPS15		MLL	ALL
ERG		EWSR1, TMPRSS2, ELF4, FUS, HEPUD1	ユーイング肉腫、前立腺、AML
ETV6		NTRK3, RUNX1, PDGFRB, ABL1, MN1, ABL2, FACL6, CHIC2, ARNT, JAK2, EVI1, CDX2, STL, HLXB9, MDS2, PER1, SYK, TTL, FGFR3, PAX5	先天性線維肉腫、多発性白血病およびリンパ腫、分泌性乳房、MDS、ALL
EVI1		RUNX1, ETV6, PRDM16, RPN1	AML、CML
EWSR1		FLI1, ERG, ZNF278, NR4A3, FEV, ATF1, ETV1, ETV4, WT1, ZNF384, CREB1, POU5F1, PBX1	ユーイング肉腫、線維形成性小円形細胞腫瘍、ALL、明細胞肉腫、肉腫、筋上皮腫
FACL6		ETV6	AML、AEL
FCGR2B		?	ALL
FGFR1OP		FGFR1	MPD、NHL
FIP1L1		PDGFRA	特発性過好酸性症候群
FNBP1		MLL	AML
FOXO3A		MLL	AL
FOXP1		PAX5	ALL
FSTL3		CCND1	B-CLL
FUS		DDIT3, ERG, FEV, ATF1, CREB3L2	脂質肉腫、AML、ユーイング肉腫、血管腫様組織球腫、線維粘液様肉腫
FVT1		IGK	B-NHL
GAS7		MLL	AML*
GMPS		MLL	AML
GPHN		MLL	AL
GRAF		MLL	AML、MDS
HCMOGT-1		PDGFRB	JMML
HEAB		MLL	AML
HIP1		PDGFRB	CMML
HIST1H4I		BCL6	NHL

10

20

30

40

50

Hugo遺伝子	遺伝子カテゴリー	転座パートナー	癌型
HLF		TCF3	ALL
HLXB9		ETV6	AML
HOXA11		NUP98	CML
HOXA13		NUP98	AML
HOXA9		NUP98, MSI2	AML*
HOXC11		NUP98	AML
HOXC13		NUP98	AML
HOXD11		NUP98	AML
HOXD13		NUP98	AML*
HSPCA		BCL6	NHL
HSPCB		BCL6	NHL
IGH		MYC, FGFR3,PAX5, IRTA1, IRF4, C CND1, BCL9, BCL8, BCL6, BCL2, BCL3, BCL10, BCL11A, LHX4, DD X6, NFKB2, PAFAH1B2, PCSK7	MM、バーキットリンパ腫、NHL、C LL、B-ALL、麦芽、MLCLS
IGK		MYC, FVT1	バーキットリンパ腫、B-NHL
IGL		BCL9, MYC, CCND2	バーキットリンパ腫
IL2		TNFRSF17	腸T細胞リンパ腫
IL21R		BCL6	NHL
IRF4		IGH	MM
IRTA1		IGH	B-NHL
ITK		SYK	末梢T細胞リンパ腫
KDM5A		NUP98	AML
LAF4		MLL, RUNX1	ALL、T-ALL
LASP1		MLL	AML
LCK		TRB	T-ALL
LCP1		BCL6	NHL
LCX		MLL	AML
LMO1		TRD	T-ALL
LMO2		TRD	T-ALL
LPP		HMGA2, MLL, C12orf9	脂肪腫、白血病
LYL1		TRB	T-ALL
MAF		IGH	MM
MAFB		IGH	MM
MALT1		BIRC3	MALT
MDS1		RUNX1	MDS、AML
MDS2		ETV6	MDS
MHC2TA		BCL6	NHL
MKL1		RBM15	急性巨核球性白血病
MLF1		NPM1	AML
MLLT1		MLL	AL
MLLT10		MLL, PICALM, CDK6	AL
MLLT2		MLL	AL
MLLT3		MLL	ALL
MLLT4		MLL	AL
MLLT6		MLL	AL
MLLT7		MLL	AL
MN1		ETV6	AML、髄膜腫
MSF		MLL	AML*
MSI2		HOXA9	CML
MSN		ALK	ALCL

10

20

30

40

50

Hugo遺伝子	遺伝子カテゴリー	転座パートナー	癌型
MTCPI		TRA	T細胞前リンパ球性白血病
MUC1		IGH	B-NHL
MYC		IGK, BCL5, BCL7A , BTG1, TRA, IGH	他の癌において増幅されたバーキットリンパ腫、B-CLL
MYH11		CBFB	AML
MYH9		ALK	ALCL
MYST4		CREBBP	AML
NACA		BCL6	NHL
NCOA2		RUNXBP2	AML
NFKB2		IGH	B-NHL
NIN		PDGFRB	MPD
NOTCH1		TRB	T-ALL
NPM1		ALK, RARA, MLF1	NHL、APL、AML
NSD1		NUP98	AML
NUMA1		RARA	APL
NUP214		DEK, SET, ABL1	AML、T-ALL
NUP98		HOXA9, NSD1, WHSC1L1, DDX10, TOP1, HOXD13, PMX1, HOXA13, HOXD11, HOXA11, RAPIGDS1, HOXC11	AML
OLIG2		TRA	T-ALL
PAFAH1B2		IGH	MLCLS
PAX5		IGH, ETV6, PML, FOXP1, ZNF521, ELN	NHL、ALL、B-ALL
PBX1		TCF3, EWSR1	プレB-ALL、筋上皮腫
PCM1		RET, JAK2	乳頭状甲状腺、CML、MPD
PCSK7		IGH	MLCLS
PDE4DIP		PDGFRB	MPD
PDGFRB		ETV6, TRIP11, HIP1, RAB5EP, H4, NIN, HCMOGT-1, PDE4DIP	MPD、AML、CMML、CML
PER1		ETV6	AML、CMML
PICALM		MLLT10, MLL	TALL、AML、
PIM1		BCL6	NHL
PML		RARA, PAX5	APL、ALL
PMX1		NUP98	AML
PNUTL1		MLL	AML
POU2AF1		BCL6	NHL
PRDM16		EVII	MDS、AML
PSIP2		NUP98	AML
RAB5EP		PDGFRB	CMML
RANBP17		TRD	ALL
RAPIGDS1		NUP98	T-ALL
RBM15		MKL1	急性巨核球性白血病
RPL22		RUNX1	AML、CML
RPN1		EVII	AML
RUNX1		RPL22, MDS1, EVII, CBFA2T3, CBFA2T1, ETV6, LAF4	AML、プレB-ALL、T-ALL
RUNXBP2		CREBBP, NCOA2, EP300	AML
SET		NUP214	AML
SFRS3		BCL6	濾胞性リンパ腫
SH3GL1		MLL	AL

10

20

30

40

50

Hugo遺伝子	遺伝子カテゴリー	転座パートナー	癌型
SIL		TAL1	T-ALL
SSH3BP1		MLL	AML
STL		ETV6	B-ALL
SYK		ETV6, ITK	MDS、末梢T細胞リンパ腫
TAF15		TEC, CHN1, ZNF384	骨外性粘液様軟骨肉腫、ALL
TAL1		TRD, SIL	リンパ芽球性白血病/二相性
TAL2		TRB	T-ALL
TCF3		PBX1, HLF, TFPT	プレB-ALL
TCL1A		TRA	T-CLL
TCL6		TRA	T-ALL
TFG		NTRK1, ALK	乳頭状甲状腺、ALCL、NSCLC
TFPT		TCF3	プレB ALL
TFRC		BCL6	NHL
TIF1		RARA	APL
TLX1		TRB, TRD	T-ALL
TLX3		BCL11B	T-ALL
TNFRSF17		IL2	腸T細胞リンパ腫
TOP1		NUP98	AML*
TPM3		NTRK1, ALK	乳頭状甲状腺、ALCL
TPM4		ALK	ALCL
TRA		ATL, OLIG2, MYC, TCL1A, TCL6, MTCPI, TCL6	T-ALL
TRB		HOX11, LCK, NOTCH1, TAL2, LY L1	T-ALL
TRD		TAL1, HOX11, TLX1, LMO1, LMO 2, RANBP17	T細胞白血病
TRIP11		PDGFRB	AML
TTL		ETV6	ALL
WHSC1		IGH	MM
WHSC1L1		NUP98	AML
ZNF145		RARA	APL
ZNF198		FGFR1	MPD、NHL
ZNF384		EWSR1, TAF15	ALL
ZNF521		PAX5	ALL
ZNFN1A1		BCL6	ALL、DLBL

10

20

30

【0402】

[実施例G] ハイブリッド捕捉のための例示的ベイト配列

表7は、3つの標的：SMAD3__標的__10、SMAD3__標的__11、SMAD3__標的__12のための例示的ベイトを提供している。

【0403】

40

50

【表 1 3】

表7: 例示のベイト

1.遺伝子	標的	ベイトゲノム位置	
SMAD3	SMAD3_標的_10	chr15:67477013-67477132	
CCATTGTGTGTGAGCAAAGGCACCCTGTCCAGTCTAACCTGAATCTCTGTAGGAAGAGG CGTGCGGCTCTACTACATCGGAGGGGAGGTCTTCGCAGAGTGCCTCAGTGACAGCGCTA TT (配列番号37)			
(ベイトID: SMAD3_標的_10.2)			
2.遺伝子	標的	ベイトゲノム位置	10
SMAD3	SMAD3_標的_10	chr15:67477037-67477156	
CTGTCCAGTCTAACCTGAATCTCTGTAGGAAGAGGCGTGCGGCTCTACTACATCGGAGG GGAGGTCTTCGCAGAGTGCCTCAGTGACAGCGCTATTTTGTCCAGTCTCCCAACTGTAA C (配列番号38)			
(ベイトID: SMAD3_標的_10.4)			
3.遺伝子	標的	ベイトゲノム位置	
SMAD3	SMAD3_標的_10	chr15:67477061-67477180	
GTAGGAAGAGGCGTGCGGCTCTACTACATCGGAGGGGAGGTCTTCGCAGAGTGCCTCAG TGACAGCGCTATTTTGTCCAGTCTCCCAACTGTAACCAGCGCTATGGCTGGCACCCGGC C (配列番号39)			20
(ベイトID: SMAD3_標的_10.6)			
4.遺伝子	標的	ベイトゲノム位置	
SMAD3	SMAD3_標的_10	chr15:67477085-67477204	
TACATCGGAGGGGAGGTCTTCGCAGAGTGCCTCAGTGACAGCGCTATTTTGTCCAGTCT CCCAACTGTAACCAGCGCTATGGCTGGCACCCGGCCACCGTCTGCAAGATCCCACCAGG T (配列番号40)			
(ベイトID: SMAD3_標的_10.1)			
5.遺伝子	標的	ベイトゲノム位置	30
SMAD3	SMAD3_標的_10	chr15:67477109-67477228	
GAGTGCCTCAGTGACAGCGCTATTTTGTCCAGTCTCCCAACTGTAACCAGCGCTATGGC TGGCACCCGGCCACCGTCTGCAAGATCCCACCAGGTAAACGAGCCGCACAGGCACCCCT G (配列番号41)			
(ベイトID: SMAD3_標的_10.5)			
6.遺伝子	標的	ベイトゲノム位置	
SMAD3	SMAD3_標的_10	chr15:67477133-67477252	40

TTTGTCCAGTCTCCCAACTGTAACCAGCGCTATGGCTGGCACCCGGCCACCGTCTGCAAG
ATCCCACCAGGTAAACGAGCCGCACAGGCACCCCTGCCTTGAGGTCCCTCTCCGAGTGC
A (配列番号42)

(ベイトID: SMAD3_標的_10.3)

7.遺伝子	標的	ベイトゲノム位置
SMAD3	SMAD3_標的_11	chr15:67479655-67479774

GACCTGGCCACTTCCATCCCCACAGCCCTGTTTCTGTGTTTTTGGCAGGATGCAACCTGA
AGATCTTCAACAACCAGGAGTTCGCTGCCCTCCTGGCCCAGTCGGTCAACCAGGGCTTTG
(配列番号43)

10

(ベイトID: SMAD3_標的_11.1)

8.遺伝子	標的	ベイトゲノム位置
SMAD3	SMAD3_標的_11	chr15:67479679-67479798

GCCCTGTTTCTGTGTTTTTGGCAGGATGCAACCTGAAGATCTTCAACAACCAGGAGTTCG
CTGCCCTCCTGGCCCAGTCGGTCAACCAGGGCTTTGAGGCTGTCTACCAGTTGACCCGAA
(配列番号44)

(ベイトID: SMAD3_標的_11.5)

9.遺伝子	標的	ベイトゲノム位置
SMAD3	SMAD3_標的_11	chr15:67479703-67479822

20

GATGCAACCTGAAGATCTTCAACAACCAGGAGTTCGCTGCCCTCCTGGCCCAGTCGGTCA
ACCAGGGCTTTGAGGCTGTCTACCAGTTGACCCGAATGTGCACCATCCGCATGAGCTTCG
(配列番号45)

(ベイトID: SMAD3_標的_11.3)

10.遺伝子	標的	ベイトゲノム位置
SMAD3	SMAD3_標的_11	chr15:67479727-67479846

ACCAGGAGTTCGCTGCCCTCCTGGCCCAGTCGGTCAACCAGGGCTTTGAGGCTGTCTACC
AGTTGACCCGAATGTGCACCATCCGCATGAGCTTCGTCAAAGGCTGGGGAGCGGAGTAC
A (配列番号46)

30

(ベイトID: SMAD3_標的_11.4)

11.遺伝子	標的	ベイトゲノム位置
SMAD3	SMAD3_標的_11	chr15:67479751-67479870

CCCAGTCGGTCAACCAGGGCTTTGAGGCTGTCTACCAGTTGACCCGAATGTGCACCATCC
GCATGAGCTTCGTCAAAGGCTGGGGAGCGGAGTACAGGTCAGTTATGGGTGCTGCCTAC
A (配列番号46)

(ベイトID: SMAD3_標的_11.2)

40

12.遺伝子 標的 バイトゲノム位置
SMAD3 SMAD3_標的_11 chr15:67479775-67479894
AGGCTGTCTACCAAGTTGACCCGAATGTGCACCATCCGCATGAGCTTCGTCAAAGGCTGG
GGAGCGGAGTACAGGTCAGTTATGGGTGCTGCCTACATCAGGGGACCCAACTCCAGGTG
AC (配列番号48)

(バイトID: SMAD3_標的_11.6)

13.遺伝子 標的 バイトゲノム位置
SMAD3 SMAD3_標的_12 chr15:67482692-67482811
TGTAACCCCCTGGAGATTTTTTAAGTCCCCACCCCACCCCTTTCCCTATTTCTTACAGGA
GACAGACTGTGACCAGTACCCCTGCTGGATTGAGCTGCACCTGAATGGGCCTTTGCAG (配列番号49)

(バイトID: SMAD3_標的_12.5)

14.遺伝子 標的 バイトゲノム位置
SMAD3 SMAD3_標的_12 chr15:67482716-67482835
GTCCCCACCCCACCCCTTTCCCTATTTCTTACAGGAGACAGACTGTGACCAGTACCCCC
TGCTGGATTGAGCTGCACCTGAATGGGCCTTTGCAGTGGCTTGACAAGGTCCTCACCCAG (配列番号50)

(バイトID: SMAD3_標的_12.3)

15.遺伝子 標的 バイトゲノム位置
SMAD3 SMAD3_標的_12 chr15:67482740-67482859
ATTTCTTACAGGAGACAGACTGTGACCAGTACCCCCTGCTGGATTGAGCTGCACCTGAAT
GGGCCTTTGCAGTGGCTTGACAAGGTCCTCACCCAGATGGGCTCCCCAAGCATCCGCTGT (配列番号51)

(バイトID: SMAD3_標的_12.2)

16.遺伝子 標的 バイトゲノム位置
SMAD3 SMAD3_標的_12 chr15:67482764-67482883
ACCAGTACCCCCTGCTGGATTGAGCTGCACCTGAATGGGCCTTTGCAGTGGCTTGACAAG
GTCCTCACCCAGATGGGCTCCCCAAGCATCCGCTGTTCCAGTGTGTCTTAGAGACATCAA (配列番号52)

(バイトID: SMAD3_標的_12.4)

17.遺伝子 標的 バイトゲノム位置
SMAD3 SMAD3_標的_12 chr15:67482788-67482907

10

20

30

40

50

CTGCACCTGAATGGGCCTTTGCAGTGGCTTGACAAGGTCCTCACCCAGATGGGCTCCCCA
AGCATCCGCTGTTCCAGTGTGTCTTAGAGACATCAAGTATGGTAGGGGAGGGCAGGCTT
G (配列番号53)

(ベイトID: SMAD3_標的_12.6)

18.遺伝子 標的 ベイトゲノム位置

SMAD3 SMAD3_標的_12 chr15:67482812-67482931

TGGCTTGACAAGGTCCTCACCCAGATGGGCTCCCCAAGCATCCGCTGTTCCAGTGTGTCT
TAGAGACATCAAGTATGGTAGGGGAGGGCAGGCTTGGGGAAAATGGCCATGCAGGAGG
TG (配列番号54)

10

(ベイトID: SMAD3_標的_12.1)

【0404】

は、2つの標的：二次構造を低減するために修飾されたFLT3__標的__2のための配列を有するベイトを提供する。FLT4__標的__31は、より短いベイトと効果的に同様であるベイトの両末端に一部の任意配列を有する。両方が、カバレッジを約4×改善する（カバレッジにおける約4×の改善）。

【0405】

20

30

40

50

【表 1 4】

表8: 例示的ベイト

1.遺伝子	標的	ベイトゲノム位置
FLT3	FLT3_標的_24	chr13:28674626-28674745

元の配列

CGTCGCGCGCCAACGCCGGCATGGCCTCCGGAGCCCCGGGGTCCCCAGGCCGCGCCGGCC
CAGCCCTGCGATGCCGCTGGAGCGGCGCGCCTCGCGCTGCAGGTGGCTCTCTTAAGGA
TG (配列番号55)

修飾された配列

CGTCTCACGCCAACGCAAGCATGTCCTCCGGAGCCCCGGGGTCCCCAGGCCGCGCCGGCC
CAGCCCTGCGATGCCGCTGGAGCGGCGCGCCTCGCACTGCAGATGGCTCTCTTAAGGA
TG (配列番号56)

(ベイトID: FLT3_標的_24.1)

2.遺伝子	標的	ベイトゲノム位置
FLT3	FLT3_標的_24	chr13:28674602-28674721

元の配列

TACCGAGCAGCGGCAGCTGGCCGCCGTCGCGCGCCAACGCCGGCATGGCCTCCGGAGCC
CGGGGTCCCCAGGCCGCGCCGGCCCAGCCCTGCGATGCCGCTGGAGCGGCGCGCCTCG
CG (配列番号57)

修飾された配列

TACCGAGCAGCGGCAGCTGGCCGCCGTCGCGCGCCAACGCCGGCATGGCCTCCGGAGCC
CGGGGTCCCCAGGCCGCGCATGCCAGCCCTGCGATGCCGCTTGAGCAACGCGCCTCA
CG (配列番号58)

(ベイトID: FLT3_標的_24.2)

3.遺伝子	標的	ベイトゲノム位置
FLT3	FLT3_標的_24	chr13:28674578-28674697

元の配列

GCTGCGAGCGAGCGAGCGGGGCCTTACCGAGCAGCGGCAGCTGGCCGCCGTCGCGCGCC
AACGCCGGCATGGCCTCCGGAGCCCCGGGGTCCCCAGGCCGCGCCGGCCCAGCCCTGCGA
TG (配列番号59)

修飾された配列

GCTTCGAGAGAGCGAGCGGGGCCTTACCGAGCAGCAGCAGCTGGCCGCCGTCGCGCGCC
AACGCCGGCATGGCCTCCGGAGCCCCGGGGTCCCCAGGCCGCGCCAGCCCAGCCCTGAGA
TG (配列番号60)

10

20

30

40

50

(バイトID: FLT3_標的_24.3)

4.遺伝子	標的	バイトゲノム位置
FLT3	FLT3_標的_24	chr13:28674554-28674673

元の配列

GTGGGGGCTGAGGGACCGCGAGGGGCTGCGAGCGAGCGAGCGGGGCCTTACCGAGCAG
CGGCAGCTGGCCGCCGTCGCGCGCCAACGCCGGCATGGCCTCCGGAGCCCGGGTCCCC
AGG (配列番号61)

修飾された配列

GAGGTGGCTGAGAGACCGCGAGGAGCTGCGAGCGAGCGAGCGGGGCCTTACCGAGCAG
CGGCAGCTGGCCGCCGTCGCGCGCCAACGCAGGCATGGCCTCCGGAGCCCAGGGTCCCC
AGG (配列番号62)

(バイトID: FLT3_標的_24.4)

5.遺伝子	標的	バイトゲノム位置
FLT3	FLT3_標的_24	chr13:28674506-28674625

元の配列

CGAGGCGGCTGGGCCGGAGGAGGCGCGCGCCCGGGTCCACACTGCGGGGTGGGGGCTG
AGGGACCGCGAGGGGCTGCGAGCGAGCGAGCGGGGCCTTACCGAGCAGCGGCAGCTGG
CCGC (配列番号63)

修飾された配列

CGAGGCGGCTGGGCCGGAGGAGGCGCGCGCCCGGATCCACACTGCGGGGTGGGGGCTG
AGGGACCGCGAGGGGCTGCGAGCGAGCGAGCGGGGACTTACCGAGCAGCGGCAACTGG
ACGC (配列番号64)

(バイトID: FLT3_標的_24.5)

6.遺伝子	標的	バイトゲノム位置
FLT3	FLT3_標的_24	chr13:28674530-28674649

元の配列

GCGCGCCCGGGTCCACACTGCGGGGTGGGGGCTGAGGGACCGCGAGGGGCTGCGAGCG
AGCGAGCGGGGCCTTACCGAGCAGCGGCAGCTGGCCGCCGTCGCGCGCCAACGCCGGC
ATGG (配列番号65)

修飾された配列

GCACGCACGGATCCACACTGCGGGGTGGGGGCTGAGGGACCGCGAGGAGCTGCGAGCG
AGCGAGCGGGGCCTTACCGAGCAGCGGCAGCTGGCAGCCGTCGCGCGCCAACGCCGGC
ATGG (配列番号66)

(バイトID: FLT3_標的_24.6)

10

20

30

40

50

7.遺伝子 標的 ベイトゲノム位置
FLT4 FLT4_標的_31 chr5:180076516-180076635

元の配列

TCGCAGGCACAGCGCGGCGCCCCGCTGCATCTCCGGCCGCTGCGCGTGGGTCCGACCCG
AGCGGCCGCGGCTCGGGGCTGAAAGTGTCCGCGCGGGCGCCGGCTGGCCTGGGGCGGG
GCG (配列番号67)

修飾された配列

CACACACACAAGCGCGGCGCCCCGCTGCATCTCCGGCCGCTGCGCGTGGGTCCGACCCG
AGCGGCCGCGGCTCGGGGCTGAAAGTGTCCGCGCGGGCGCCGGCTGGCCTGCACACACA
CA (配列番号68)

(ベイトID: FLT4_標的_31.1)

8.遺伝子 標的 ベイトゲノム位置
FLT4 FLT4_標的_31 chr5:180076396-180076515

元の配列

GGCGGAGCGGTCTCAGCGCCCCGCCCCAGGTGCGCGGTACCCCCTCCCCGGCCAGCCCCA
CGCTCGGGCGGGTGGCCCCGTTCCGCCGCTCACCGTCCAGGAGTCCCAGGCAGAGCCAC
AG (配列番号69)

修飾された配列

CACACACACATCTCAGCGCCCCGCCCCAGGTGCGCGGTACCCCCTCCCCGGCCAGCCCCA
CGCTCGGGCGGGTGGCCCCGTTCCGCCGCTCACCGTCCAGGAGTCCCAGGCCACACACA
CA (配列番号70)

(ベイトID: FLT4_標的_31.2)

9.遺伝子 標的 ベイトゲノム位置
FLT4 FLT4_標的_31 chr5:180076420-180076539

元の配列

CCAGGTGCGCGGTACCCCCTCCCCGGCCAGCCCCACGCTCGGGCGGGTGGCCCCGTTCCG
CGCGCTACCGTCCAGGAGTCCCAGGCAGAGCCAGTCGCAGGCACAGCGCGGCGCCCCG
(配列番号71)

修飾された配列

CACACACACAGGTACCCCCTCCCCGGCCAGCCCCACGCTCGGGCGGGTGGCCCCGTTCCG
CGCGCTACCGTCCAGGAGTCCCAGGCAGAGCCACAGTCGCAGGCACAGCGCACACACA
CA (配列番号72)

(ベイトID: FLT4_標的_31.3)

10.遺伝子 標的 ベイトゲノム位置

10

20

30

40

50

FLT4 FLT4_標的_31 chr5:180076468-180076587

元の配列

GGCCCGTTTCGCCGCGCTCACCGTCCAGGAGTCCCAGGCAGAGCCACAGTCGCAGGCACA
GCGCGGCGCCCCGCTGCATCTCCGGCCGCTGCGCGTGGGTCCGACCCGAGCGGCCGCGG
CT (配列番号73)

修飾された配列

CACACACACACCGCGCTCACCGTCCAGGAGTCCCAGGCAGAGCCACAGTCGCAGGCACA
GCGCGGCGCCCCGCTGCATCTCCGGCCGCTGCGCGTGGGTCCGACCCGAGCCACACACA
CA (配列番号74)

10

(ベイトID: FLT4_標的_31.4)

11.遺伝子 標的 ベイトゲノム位置

FLT4 FLT4_標的_31 chr5:180076444-180076563

元の配列

GGCCAGCCCCACGCTCGGGCGGGTGGCCCGTTTCGCCGCGCTCACCGTCCAGGAGTCCCA
GGCAGAGCCACAGTCGCAGGCACAGCGCGGCGCCCCGCTGCATCTCCGGCCGCTGCGCG
TG (配列番号75)

20

修飾された配列

CACACACACAACGCTCGGGCGGGTGGCCCGTTTCGCCGCGCTCACCGTCCAGGAGTCCCA
GGCAGAGCCACAGTCGCAGGCACAGCGCGGCGCCCCGCTGCATCTCCGGCCACACACA
CA (配列番号76)

(ベイトID: FLT4_標的_31.5)

12.遺伝子 標的 ベイトゲノム位置

FLT4 FLT4_標的_31 chr5:180076492-180076611

元の配列

CAGGAGTCCCAGGCAGAGCCACAGTCGCAGGCACAGCGCGGCGCCCCGCTGCATCTCCG
GCCGCTGCGCGTGGGTCCGACCCGAGCGGCCGCGGCTCGGGGCTGAAAGTGTCCGCGCG
GG (配列番号77)

30

修飾された配列

CACACACACAAGGCAGAGCCACAGTCGCAGGCACAGCGCGGCGCCCCGCTGCATCTCCG
GCCGCTGCGCGTGGGTCCGACCCGAGCGGCCGCGGCTCGGGGCTGAAAGTGCACACACA
CA (配列番号78)

(ベイトID: FLT4_標的_31.6)

40

【0406】

【実施例H】 臨床的癌標本の次世代配列決定からの体細胞ゲノムの改変の高感度検出のためのベイズ手法

本明細書に記載されているベイズ手法を、以下の実施例において実施した。

【0407】

この手法の有用性は電力算出によって例示され、臨床状況における関連の変異頻度のより低い範囲における置換検出に対するデータ駆動の事前影響を記載する。図4に示されている通り、事前予想（例えば、 $1e-6$ または10%の事前）および変異頻度（例えば、1%、5%または15%の変異）の値は、それぞれ「A Bayesian Approa

50

ch for Sensitive Detection of Somatic Genomic Alterations from Next-generation Sequencing of Clinical Cancer Specimens」の(i)および(ii)に記載されている値に対応する。図4は、事前予想を組み込むことが、例えば、変異部位での必要とされるカバレッジ深度を低減すること、または変異を検出するための概算された力(感受性)を増加させることによって、より希少な変異に対する検出力を改善することができることを示す。

【0408】

【実施例I】 バイズ手法：構築された低純度の多クローン性試料への適用

本明細書において開示されているバイズ手法のこれらの利益をさらに実証するため、人工的な低純度の多クローン性「腫瘍」試料を、1000ゲノムプロジェクトにおける10人の参加者からのDNAの等しい付加混合物によって構築し、これによって、全DNA(非公式のヘテロ接合性SNPから生じる。)の約5%または10%に存在する多数の配列バリエーションを含有するDNAプールを創出する。ミックスを、182個の癌関連遺伝子エクソンのハイブリッド選択にかけ、Illumina HiSeq 2000プラットフォーム上で配列決定して、遺伝子パネルにわたっておよそ350×の平均カバレッジを得た。各構成要素試料を同様に個別に加工して、全てのSNP部位で遺伝子型を決定した。プール中に存在するおよそ260個の約5%「変異」のうち、 $1e-6$ の事前を使用して高い信頼で89%を検出したが、一方で、それぞれ1%および10%の事前を使用して94%および95%が検出可能であり(見逃された部位の平均カバレッジ約125×)、上記の理論的結論を支持した。プール中に存在する102個の10%「変異」のうち、 $1e-6$ の事前を使用して高い信頼で98%を検出したが、一方で、1%および10%の事前を使用して99%および99%が検出可能であった(見逃した部位のカバレッジ13×)。

【0409】

【実施例J】 バイズ手法：肺および大腸腫瘍試料への適用。

COSMICデータベース(ワールドワイドウェブ上でsanger.ac.uk/genetics/CGP/cosmic)からのいくつかの癌型における関連の変異の頻度の事前予想を誘導し、ルーチンの臨床標本から抽出された80超の肺癌および大腸癌の試料を分析した。この癌型におけるこの変異についての3%事前の組込みによってのみ検出することができる大腸癌における1%のPIK3CA変異p.H1047Rを含めて、20個超の異なる遺伝子における公知の変異を観察した。これらの結果は、腫瘍型特異的な変異スペクトルの周辺での事前予想の賢明な組込みが、臨床状況にとってNGSベースの腫瘍ゲノム分析の翻訳に有益であり得ることを示している。

【0410】

【実施例K】 バイズ手法：乳癌試料への適用

FFPE乳癌試料について約260×に配列決定された182個の癌関連遺伝子のエクソンにおける置換変異呼び出しを行った。代わりの対立遺伝子の>2コピーを有する部位の数は1,793である。変異の存在における>99%の事後信頼を有する部位の数は402である。フィルター後に残留している部位の数は188であり、これはおよそ、バリエーション部位の予想数である。dbSNPにない部位の数は14であり、これは、およそ、dbSNPが>90%の変動を捕捉する場合にdbSNPにない部位の予想数である。非同義部位の数は5である。COSMICにおける部位の数は2である(PIK3CA p.H1047RおよびP53 p.F113S)。

【0411】

【実施例L】 バイズ手法：稀な変異の検出

多くのルーチンの臨床標本は、関連の希少変異を含有する。図5は、100個超の臨床的癌試料における変異頻度を示す。試料は、主に大腸癌および肺癌のFFPE生検物、外科的切除物または細針吸引物であった。一連の臨床試料において見出された公知の変異の頻度スペクトルは、表12に示されている。

【0412】

【表 15】

表12: 一連の臨床試料において見出だされた公知の変異の頻度スペクトル

一連の臨床試料において見出だされた公知の変異の頻度スペクトル				
変異の割合	変異の割合	変異の割合	変異の割合	変異の割合
<5%	<10%	<25%	<50%	<100%
7%*	17%	50%	85%	100%

*おそらく過小評価されている

10

【0413】

【実施例M. 1】 個別に合成されたオリゴヌクレオチド捕捉プローブを使用する、高性能溶液ベースの標的選択

溶液ベースのゲノム標的選択技法の利用能は、標的化配列決定用途の急速な開発を可能にし、この一部は、臨床的配列決定試験の導入に至った。市販化されたハイブリダイゼーション捕捉試薬は、ビオチン化DNAまたはRNAプローブ（「ベイト」）に変換されているアレイ合成オリゴヌクレオチドに基づく。しかしながら、プローブのこれらの複合プールを生成する方法は、性能課題、例えば高GC含有量標的の捕捉に直面する。

【0414】

20

57個の臨床的に関連のおよび作用性の癌関連遺伝子を表す約130kbの標的領域を捕捉するための個別に合成された、5'-ビオチン化オリゴヌクレオチド（「オリゴベイト」）を使用する代替の手法は、本明細書に記載されている。24時間ハイブリダイゼーション手順を用いてこれらのオリゴベイトを使用して選択されたインデックス化配列決定ライブラリーで、5,000倍の標的濃縮を得た。50Mの49×49対合末端読み取りで、568×の標準偏差を有する2100×の平均標的カバレッジを生成した（27%）。全ての標的は首尾よくカバーされ、標的化塩基の99.95%が>500×でカバーされた。さらに、標的カバレッジは、GCバイアスを実質的に有していなかった。GC含有量>70%を有する標的は、平均すると1,975×カバレッジになり、GC含有量<35%を有する標的は、平均すると1,996×のカバレッジになった。

30

【0415】

いっそう短いハイブリダイゼーション時間を使用して、高性能を保持し：2.5時間のハイブリダイゼーション後に、標的化塩基の99.3%が>500×でカバーされた。

【0416】

SSPE（Salmon Sperm、PE）／デンハート液の使用は、TEACL、TMACLおよび／または硫酸デキストランを含有するhyb／洗浄緩衝剤より性能が優れていた。

【0417】

オリゴベイトをアレイ誘導ベイトプールにスパイクして、そうでなければ捕捉するのが困難な（例えば、高い%GC）領域のカバレッジを増加させる、または新たな遺伝子含有量を急速に増大することができる。この手法は、高性能の標的化臨床的配列決定試験を開発するための非常に有効なおよびスケーラブルな方法を提供する。

40

【0418】

【実施例M. 2】 捕捉ベイトを最適化する方法

3つのベイトセットを試験した。結果は図7に要約されている。ベイトセットは以下の通りであった：

ベイトセット番号1は、5'-ビオチン化の個別合成DNAオリゴヌクレオチドベイトのみからなる。

【0419】

ベイトセット番号2は、5'-ビオチン化の個別合成DNAオリゴヌクレオチドベイトで

50

スパイクされた、ビオチン化されたアレイ誘導RNAベイトを含む。

【0420】

ベイトセット番号3は、ビオチン化されたアレイ誘導RNAベイトのみからなる。

【0421】

全ての5'-ビオチン化の個別合成DNAオリゴヌクレオチドは、5'ビオチンを有する120個の塩基であった。

【0422】

図7は、ベイトセット番号3と比較した、ベイトセット番号1およびベイトセット番号2で検出されたカバレッジにおける均一性を比較するカバレッジヒストグラムである。ベイトセットは、図7において番号1、2および3として示されている。高い%GCに対応するベイトセット番号3を使用するカバレッジにおけるいくつかのギャップが存在したが、一方、対応する領域は、図7において描写されている通り、ベイトセット番号1および番号2を使用して深くカバーされた。図7において、「GC__密度__標的…」と標識されている左手のパネルは、標的内の局所GC含有量を示している。線は65%GC含有量を表し、ここで、線を超える任意の値は、より高いGC含有量を表す。ヒストグラムに示されている通り、カバレッジは、高GC含有量の区域におけるベイトセット番号3が最も低い。「IDT__ベイト…」と標識されている図7における最下のパネルは、示されている標的をカバーするオリゴの配置を示している。

【0423】

標的の数、および単独でアレイ誘導ベイトセットを使用するまたは個別に合成されたベイトでスパイクされたアレイ誘導ベイトセットを使用するカバレッジにおける変化のグラフィック表示が、図6に描写されている。より具体的に、図6は、カバレッジヒストグラムの線状表示である。標的の数(y軸)は、カバレッジ(x軸)の関数として描写されている。線番号1は、5'-ビオチン化の個別合成DNAオリゴヌクレオチドベイトでスパイクされた5'-ビオチン化されたアレイ誘導RNAオリゴヌクレオチドベイトを含むベイトセット(図6において「ベイトセット番号1」と称される。)を使用するカバレッジを表す。線番号2は、ビオチン化されたアレイ誘導RNAオリゴヌクレオチドベイトのみを含むベイトセット(図6において「ベイトセット番号2」と称される。)を使用して得られたカバレッジを表す。ベイトセット番号2を使用する全体の平均カバレッジは924であったが、一方、ベイトセット番号2を使用する高GC含有量(約68%)の区域におけるカバレッジは73であった。対照的に、ベイトセット番号1を使用した場合、全体カバレッジはベイトセット番号1と同様で約918であったが、高GC含有量の区域において、カバレッジは183に改善された。

【0424】

【実施例M.3】 ベイトセットを評価するための例示的実験条件

ベイトセットAは、5'-ビオチン化の個別合成DNAオリゴヌクレオチドベイトのみからなる。元のセットは、標的テリトリーの133kbをカバーする1000個のオリゴであった(本明細書において「大きなセット」、「ベイトセットA」または「DNAオリゴベイト」と称される。)

【0425】

「スパイクイン」実験のため、元の1000のDNAオリゴセット(「大きなセット」)を、ビオチン化されたアレイ誘導RNAオリゴヌクレオチドベイトからなるベイトセットに添加した(この例において「ベイトセットB」または「RNAベイト」と称される。)。ベイトセットAからの異なる比のDNAオリゴベイトを、ベイトセットBからのRNAベイトと混合した。特に、1:10のDNAオリゴベイト:RNAベイト比を使用した(10ngの全DNAオリゴベイト対100ngの全RNAベイト)。ハイブリダイゼーションおよび洗浄の条件を、RNAベイトに最も理想的である条件に整合した。

【0426】

ベイト配置に対応したDNAオリゴベイトを使用した場合、低いタイリング密度で、カバレッジにおける強い周期性を検出した。加えて、低いタイリング密度は、インデルを有

する対立遺伝子の捕捉をより困難にする恐れがある。そのため、表13に描写されている異なるタイリング密度でMAP3K1のために、ベイトセットを設計した。下記のミックスにおいて、6つの癌関連遺伝子(DAXX、TRRAP、CREBBP、GRIN2A、SPOP、GNA11)のエクソンを捕捉するように設計された5'-ビオチン化の個別合成DNAオリゴベイトを含有するミックス1を、アレイ誘導RNAオリゴヌクレオチドベイトのみにスパイクした(ベイトセットB)。DAXX、TRRAP、CREBBP、GRIN2AおよびSPOPは、RNAベイトセット中に存在しなかった。ミックス2-4をベイトセットA(DNAオリゴベイトの大きなセット)にスパイクして、MAP3K1のエクソンのための捕捉ベイトの異なるタイリング密度(最も高密度であるミックス2を用いる。)を試験した。RNAベイトセットは単独で、約1MBの配列をカバーした。

10

【0427】

【表16】

表13. 捕捉プローブを使用する方法のためのミックス

カテゴリー	数
ミックス1	メラノーマ遺伝子に対して369個のオリゴ
ミックス2	MAP3K1に対して60のタイリング密度の91個のオリゴ
ミックス3	MAP3K1に対して100のタイリング密度の57個のオリゴ
ミックス4	MAP3K1に対して150のタイリング密度の40個のオリゴ
ミックス5	STK11エクソン3に対する3個のオリゴ

20

【0428】

捕捉物にインプットしたのは、2μgのプールされた細胞株DNAライブラリーであった。2μgのライブラリーをブロックミックス(表14)と混合し、乾燥させ、9μLの水中に再懸濁した。この混合物を次いでプレート中に置き、サイクラーに移し、98℃で5分間、続いて68℃で2分間実行した。プレートを次いで開封し、68℃の11μLのDNAベイト/hyb緩衝剤混合物を添加した。68℃でのDNAベイト/hyb混合物=10μLのhyb緩衝剤+1μLのベイト(10ng、50ngまたは100ngのベイトを含有する。))。

30

【0429】

DNAベイト単独(例えば、ベイトセットA)を用いた捕捉物について、ハイブリダイゼーションを68℃で行い、洗浄を行った。ベイトを5ng、10ng、100ng、1000ngおよび2000ng(インプットライブラリー2μg当たり)で試験した。24時間のhybについて、5-10ngの条件、および最大100ngまでの条件を試験した。

【0430】

不十分な性能/高GCの領域を救出するためにRNA-アレイベイトセット(B)にスパイクされた大きなDNAベイトセット(100kb)を用いた捕捉物について、ハイブリダイゼーションを68℃で行い、洗浄を70℃で行った。ベイトセットを、1:10のDNAオリゴ:RNAベイト(すなわち、10ngの総質量のオリゴベイトおよび100ngの総質量のRNAベイト)で試験した。

40

【0431】

RNAベイトセットにスパイクされた小さい遺伝子焦点化DNAベイトセットを用いた捕捉物について、ハイブリダイゼーションを68℃で行い、広範な洗浄温度を試験した(62℃、64℃、66℃、68℃、70℃および72℃)。

【0432】

ミックス1(6個の新たな遺伝子を添加する。)を、以下の比:1:5、1:10および1:20の全オリゴDNAベイト質量:RNAベイト質量(すなわち、20ng:100ng、10ng:100ngおよび5ng:100ng)で試験した。

50

【0433】

ミックス5（経路低カバレッジに対するSTK11のエクソン3を表す3個のオリゴ）を、1：500、1：1000および1：2000のDNAオリゴ：RNAオリゴで試験した。全RNAベイト100ngを使用した。STK11は、RNAベイト単独を用いて捕捉された場合にそれが不十分な検出性能で重要な癌標的を表す時に試験した。STK11のエクソン3のDNAオリゴスパイキングは、カバレッジを70×の平均から300×に押し上げた。

【0434】

【表17】

表14. 捕捉プローブを使用する方法のための緩衝剤

ベイト(プールされたIDTオリゴ) 39600 100 nmol = 0.0039600グラム =
(g/mol) 3,960,000ナノグラム

低TE中に再懸濁された 25 mL 250 μ L Tris
5 μ L EDTA

ブロッキングミックス	[ストック]	[作業]	14.5 μ L/rxn
Cot1	1 μ g/ μ l	1 μ g/ μ l	10
サケの精子	10 μ g/ μ l	10.0 μ g/ μ l	1
PE 1.0	800 μ M	800 μ M	1.75
ユニバーサルインデックス	800 μ M	800 μ M	1.75

2×のHyb緩衝剤	[ストック]	[最終]	10ml(10 μ L/rxn) 中
SSPE	20×	10×	5 ml
デンハート液	50×	10×	2 ml
EDTA	0.5M	0.01 M	200 μ l
SDS	10%	0.20%	200 μ l
水			2.6 ml

ビーズ洗浄	[ストック]	[最終]	50ml(200 μ L/洗浄) 中
NaCl	5 M	1 M	10 ml
Tris	1 M	10 mM	500 μ l
EDTA	0.5 M	1 mM	100 μ l
水			39.4 ml

洗浄緩衝剤1	[ストック]	[最終]	50ml(150 μ L/洗浄) 中
SSC	20×	1×	2.5 ml
SDS	10%	0.10%	500 μ l
水			47 ml

洗浄緩衝剤2	[ストック]	[最終]	50ml(150 μ L/洗浄) 中
SSC	20×	0.1×	250 μ l
SDS	10%	0.10%	500 μ l
水			49.25 ml

【0435】

10

20

30

40

50

【実施例 N】 ライブラリーメンバーのオフターゲット核酸結合を低減

オフターゲット核酸相互作用は、捕捉プローブ、例えばオリゴヌクレオチドベイトへのハイブリダイゼーション（例えば、溶液または固相ハイブリダイゼーション）によって標的核酸の選択の効率を制限する恐れがある。オフターゲット選択は、典型的に、ハイブリッド選択のストリンジェンシー条件が低減される場合、例えばより低い核酸溶解温度（例えば、RNA：DNA二重鎖と比較した場合のDNA：DNA二重鎖の T_m ）を有する標的：捕捉二重鎖のために選択する場合に増加される。したがって、オフターゲット配列の捕捉は、DNA：DNAハイブリダイゼーションにおいて、より問題であり得る。オフターゲット選択は、例えば、ハイブリダイゼーション捕捉および／またはアーチファクト的ハイブリッド捕捉の減少収率の1つ以上をもたらす恐れがあり、これが順じて、後続のステップ、例えば配列決定における非効率に至る。

10

【0436】

ライブラリーメンバーは、ライブラリー挿入（オンターゲットならば、捕捉プローブ、例えばベイトで二重鎖を形成する。）および1つ以上の非標的配列（例えば、アダプター配列、増幅プライマーまたはタグの1つ以上、およびバーコードタグ）を含むことができる。典型的に、ベイトは、ライブラリー挿入、例えば標的DNAにハイブリダイズする。しかしながら、ライブラリー挿入は、ライブラリーにおいてあらゆる断片上に典型的に存在するユニバーサルアダプターを有することができる。反応混合物における他の配列を用いる（例えば、アダプター配列への結合を介する。）二重鎖形成による、捕捉プローブハイブリダイズ化ライブラリーメンバーの非標的配列は、所望されない配列、例えばオフターゲットライブラリーメンバーの選択に至る恐れがある。

20

【0437】

理論に束縛されることを望まないが、捕捉プローブで二重鎖を形成したオンターゲットライブラリーメンバーとオフターゲット配列との間の鎖状体形成は、オフターゲット配列の選択をもたらすことがある。図6は、ダイアグラム形態において、ライブラリーメンバーの非標的鎖状体の例示的立体配置を例示している。非標的領域（例えば、「P5」および「P7」として描写されているアダプター）は、それらの相補的非標的鎖（それぞれ「rcP5」および「rcP7」として描写されている。）にハイブリダイズするとして示されている。ビオチンタグ付けされたベイトは、ライブラリーメンバーの標的挿入の相補的領域にハイブリダイズして示されている。オフターゲット結合は、ライブラリーメンバーの鎖状体形成に至り、したがって、標的結合特異性における低減（本明細書においてオフターゲット選択の増加とも称される。）に至る恐れがある。

30

【0438】

DNA（ライブラリーメンバー）：RNA（ベイト）二重鎖を伴う標的：捕捉二重鎖において、捕捉プローブで二重鎖を形成したオンターゲットライブラリーメンバーとオフターゲット配列との間の鎖状体形成は、65–70℃で典型的に行われる高ストリンジェンシー洗浄中に崩壊されることがある。典型的に、DNA：DNA二重鎖のより低い溶解を伴う洗浄は、RNA：DNA二重鎖に対して、より低い温度で行われる。鎖状体形成を崩壊することができないことで、DNAベイトを使用する場合に、標的捕捉の百分率を相対的に低く保持してきた（45–50%）。アダプターに相補的な市販のブロッキングオリゴは、鎖状体形成を最小化するために添加されるが、それらは典型的に、特にDNA：DNAハイブリダイゼーションにおいて鎖形成を適切に阻害することはない。

40

【0439】

非標的配列（例えば、アダプター）媒介選択を低減する方法および組成物は、本明細書において開示されている。特定の実施形態において、ライブラリーメンバーの非標的核酸配列（例えば、アダプター配列）に相補的であるまたはライブラリーメンバーの非標的核酸配列（例えば、アダプター配列）で二重鎖を形成することができるとともに、バックグラウンド核酸に対する非標的核酸配列、例えば他の相補的非標的核酸配列についての値より高い、ブロッキングオリゴヌクレオチドとライブラリーメンバーの非標的核酸配列との間の結合相互作用に関連のパラメータについての値を有するブロッキングオリゴヌクレオチ

50

ドが、開示されている。増加された結合相互作用を有する例示的ブロッキングオリゴヌクレオチドとしては、延長されたブロッカー長さ、例えば非標的核酸に対する延長された相補性を有するオリゴヌクレオチド；1つ以上の非天然ヌクレオチドを有するブロッキングオリゴヌクレオチド；およびデオキシリボヌクレオチドの代わりにオリゴリボヌクレオチドを含む（または、これから実質的に構成されている。）ブロッキングオリゴヌクレオチドが挙げられる。

【0440】

【実施例0】 延長されたブロッカー長さ

この実施例は、パーセントオンターゲット選択が、ブロッキングオリゴヌクレオチドの長さを延長することによって改善され得ることを実証する。

【0441】

アダプター特異的ブロッキングオリゴヌクレオチドが、本明細書において記載されている通りに行われるハイブリダイゼーション反応に添加されて、実施例14に記載されている通りのオフターゲット核酸結合のキャリーオーバーを防止する。実施例4に記載されている実験条件において、オフターゲット結合を変性しそうである高ストリンジェンシー洗浄が行われる。しかしながら、DNA：DNA相互作用のための最適なハイブリダイゼーションおよび洗浄の条件は、実施例13A-13Cに記載されている通り、洗浄の温度を低下させ、したがって、オフターゲット結合を増加させる。

【0442】

ブロッキングオリゴは、アダプター、例えば実施例Cに記載されているIlluminaマルチプレックスアダプターに相補的に設計されて、アダプターとブロッキングオリゴとの間の相補性の程度を増加させることができる。例えば、P5ブロッキングオリゴは58bp塩基長であるが、ブロッカーは46塩基だけである。P5ブロッキングオリゴの長さを19塩基延長した。ブロッキングオリゴの長さを19塩基延長して、選択効率がおおよそ5%増加した（図9に示されている。）。図9は、標準的および延長されたブロッキングオリゴを使用する標的選択の百分率を描写する棒グラフである。4つの代表的実験からのデータが示されている。図10は、標準的または延長されたブロッカーを使用する捕捉結果を示すエクソンカバレージヒストグラムを描写している。

【0443】

改善されたブロッキングは、アダプターとブロッキングオリゴとの間の相補性領域の長さを延長することによって達成され、したがって、溶融温度を増加させることができる。

【0444】

参照による組み込み

本明細書において記述されている全ての公報、特許および特許出願は、各個別の公報、特許または特許出願が参照により組み込まれると具体的および個別に示されているかのように、本明細書によって参照によりそれらの全体を組み込まれる。矛盾の場合において、本明細書における任意の定義を含めて、本出願が優先される。

【0445】

その上、参照によりそれらの全体を組み込まれるのは、ワールドワイドウェブ上のtigr.orgにおけるゲノム研究所（TIGR）および／またはワールドワイドウェブ上のncbi.nlm.nih.govにおける国立バイオテクノロジー情報センター（NCBI）によって維持されているものなど、公開データベースにおけるエントリーと関連する受託番号を参照する任意のポリヌクレオチドおよびポリペプチドの配列である。

【0446】

本明細書において使用される用語は、特定の実施形態を記載する目的のためだけであり、限定的なものではない。当業者は、本明細書に記載の複数および／または単数の用語の実質的な使用に関して、文脈および／または用途に適切であるように複数形に言い換えることができる。様々な単数／複数の置き換えは、明確さのために本明細書においてははっきりと述べられていることがある。

【0447】

10

20

30

40

50

本発明は特定の実施形態を参照して記載されてきたが、本発明は様々に変更することができるとともに本発明の範囲から逸脱することなく均等物に置き換え得ることは、当業者によって理解されよう。加えて、特別な状況または材料を、本発明の範囲から逸脱することなく本発明の教示に適応させるために、多くの改変が採用される。そのため、本発明は開示されている特別な実施形態または実施例に限定されるべきでないが、本発明は添付の請求項の範囲内にある全ての実施形態を含むものとみなす。

【図面】

【図 1】

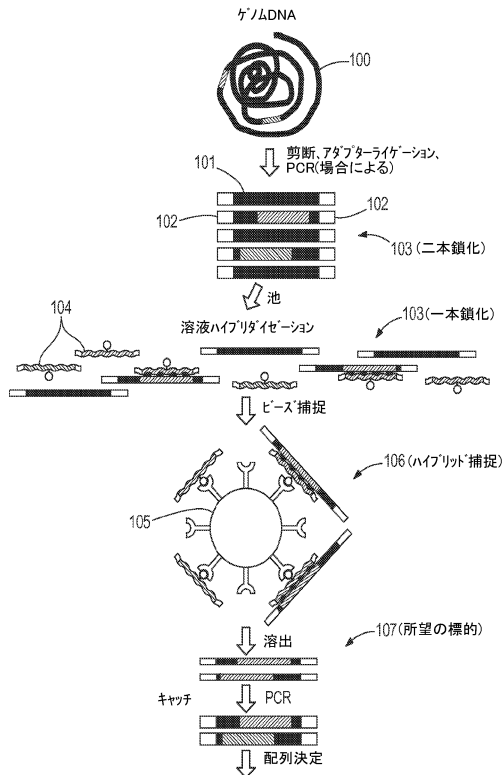


FIG. 1

【図 2 A】

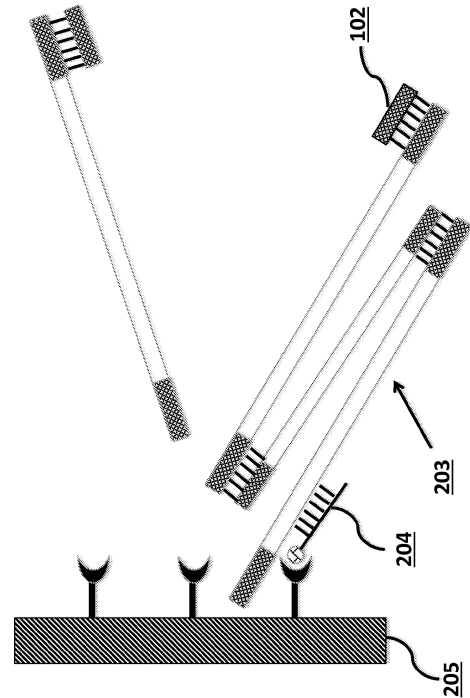


FIG. 2A

【図 2 B】

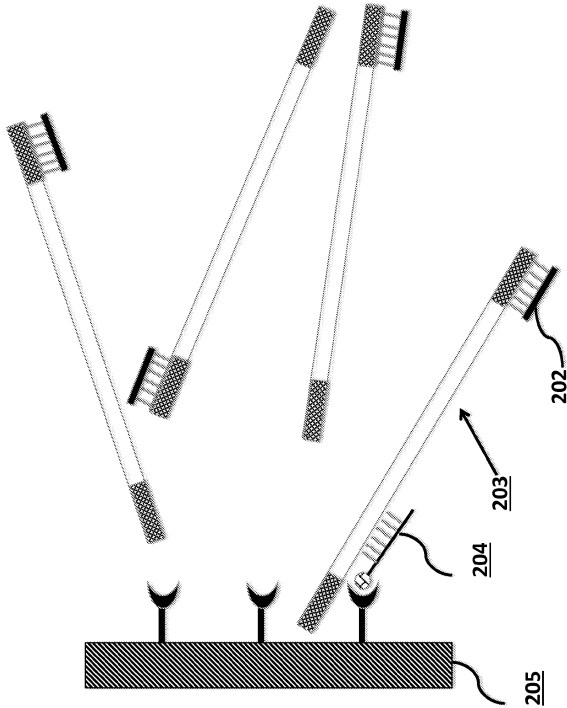


FIG. 2B

【図 3 A】

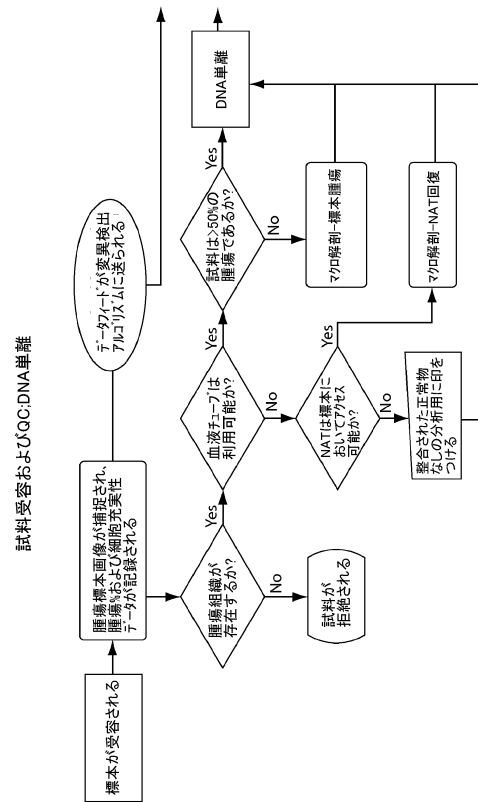


FIG. 3A

【図 3 B】

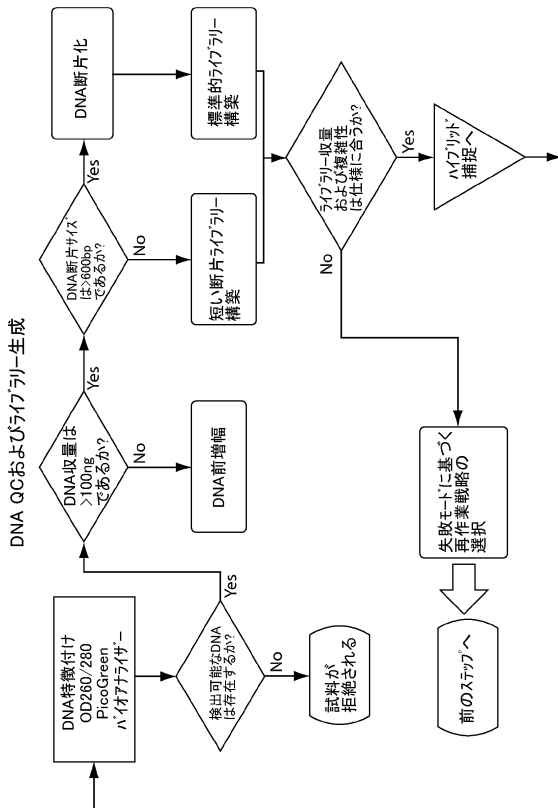


FIG. 3B

【図 3 C】

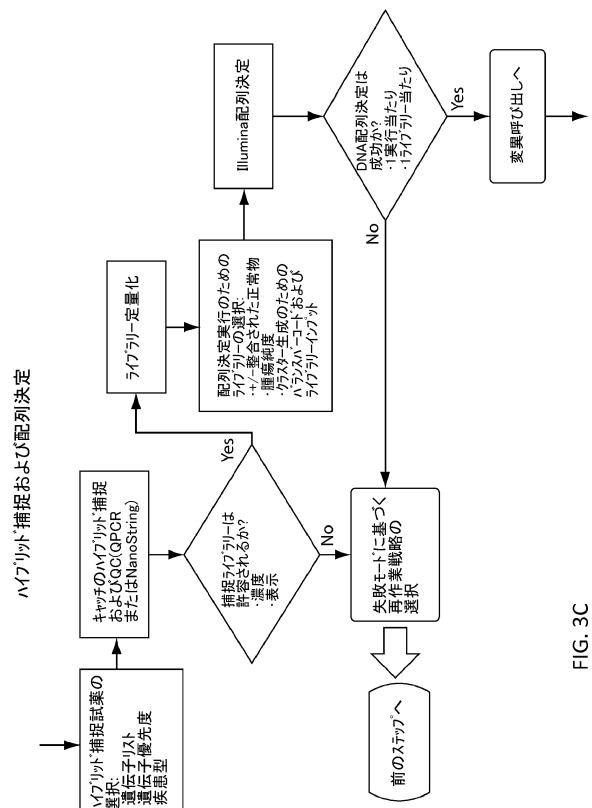
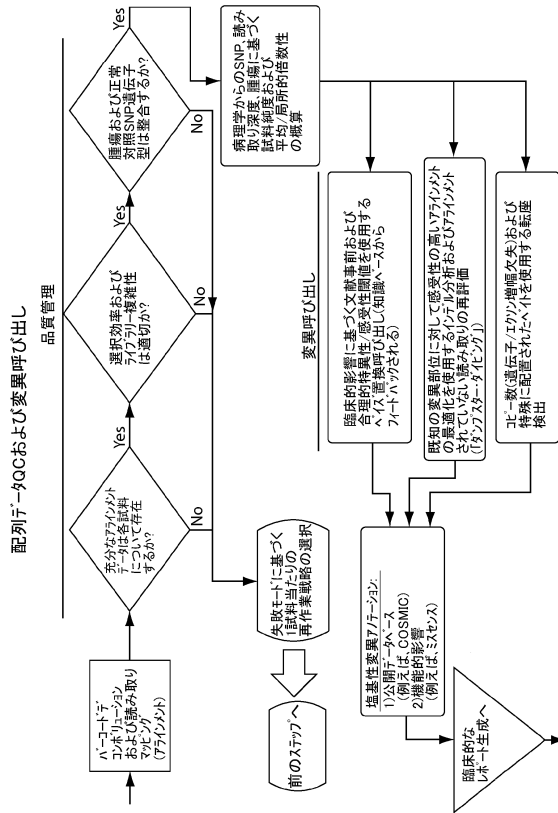


FIG. 3C

【図 3 D】



【図 3 E】

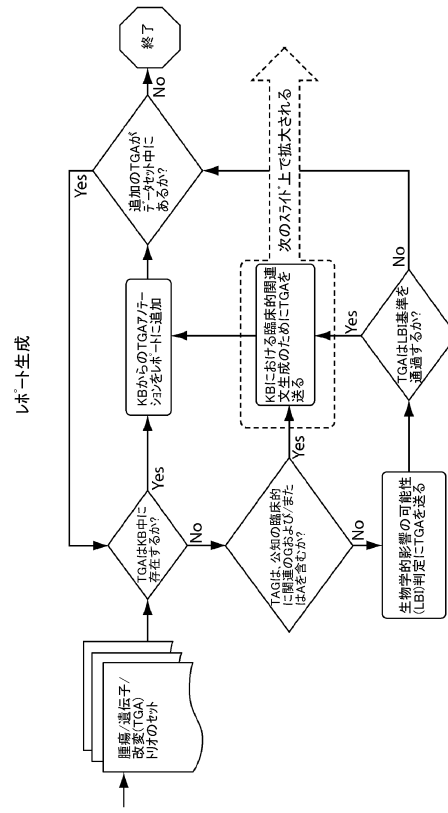


FIG. 3E

【図 3 F】

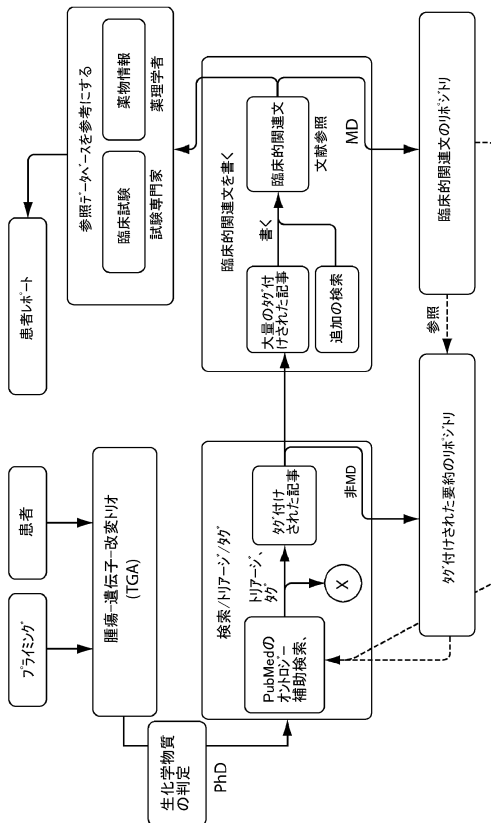


FIG. 3F

【図 4】

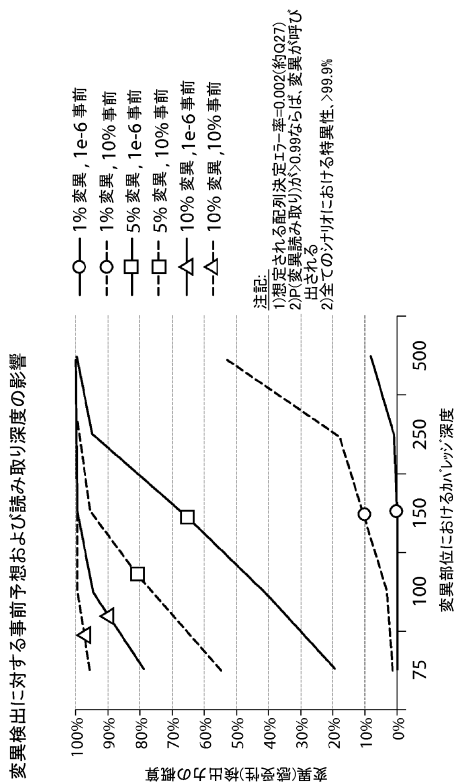


FIG. 4

【図 5】

>100個の臨床的癌試料における変異頻度

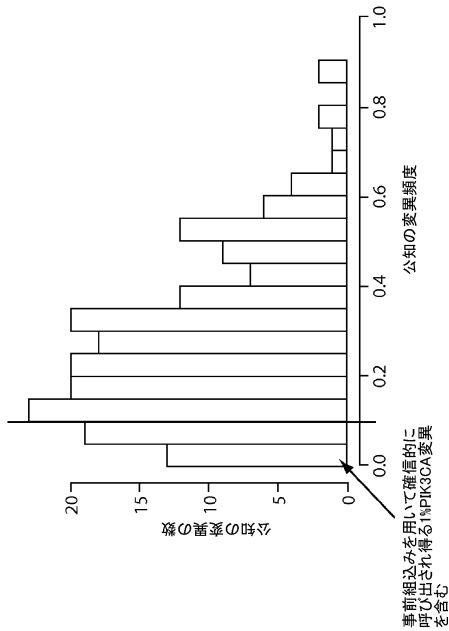


FIG. 5

【図 6】

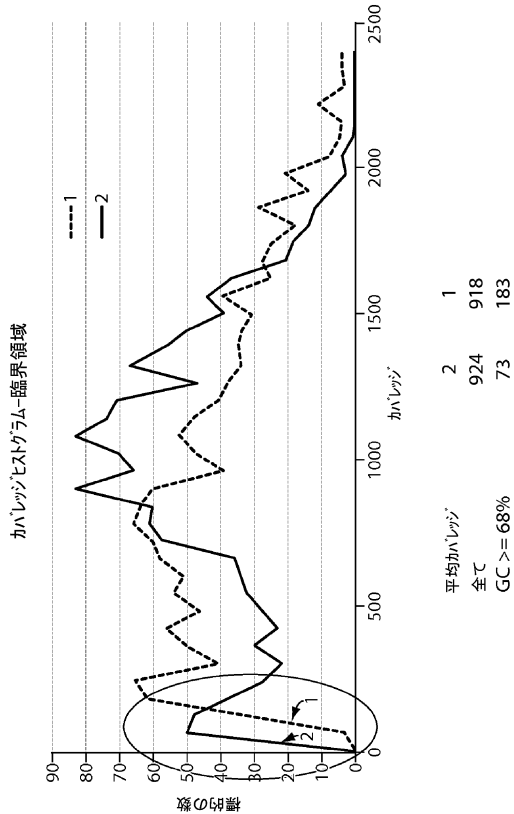


FIG. 6

【図 7】

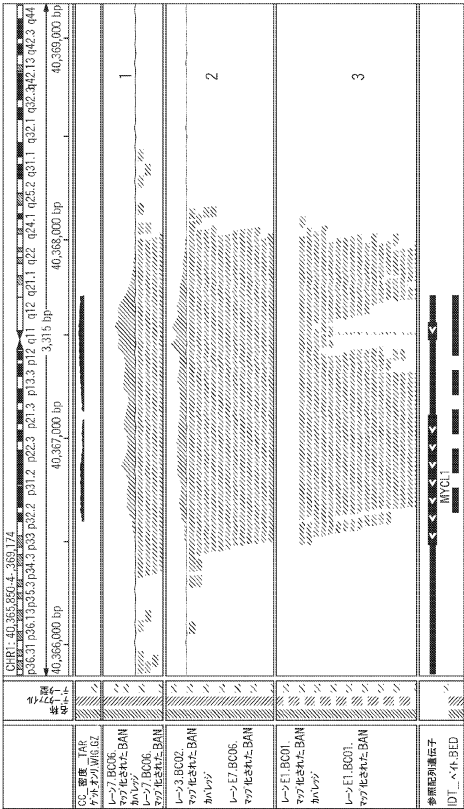


FIG. 7

【図 8】

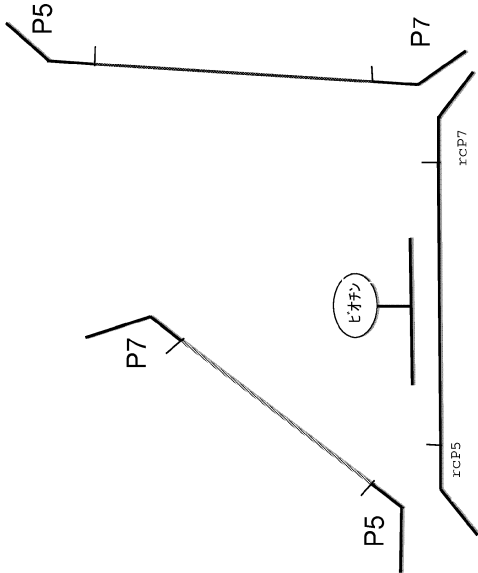


FIG. 8

【図 9】

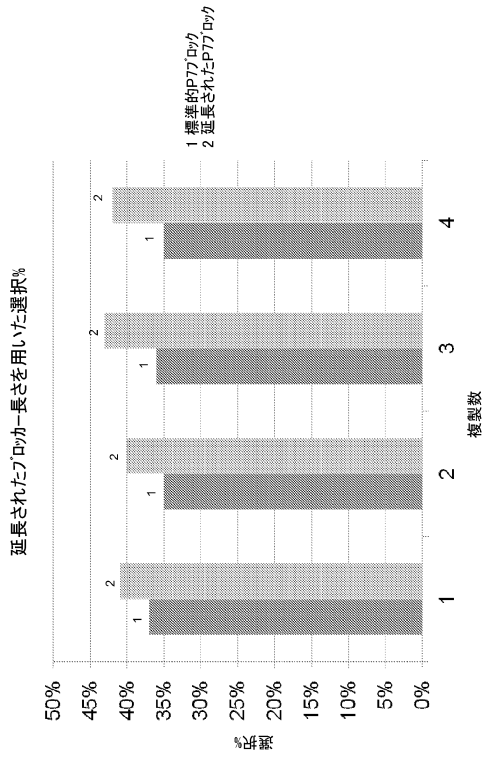


FIG. 9

【図 10】

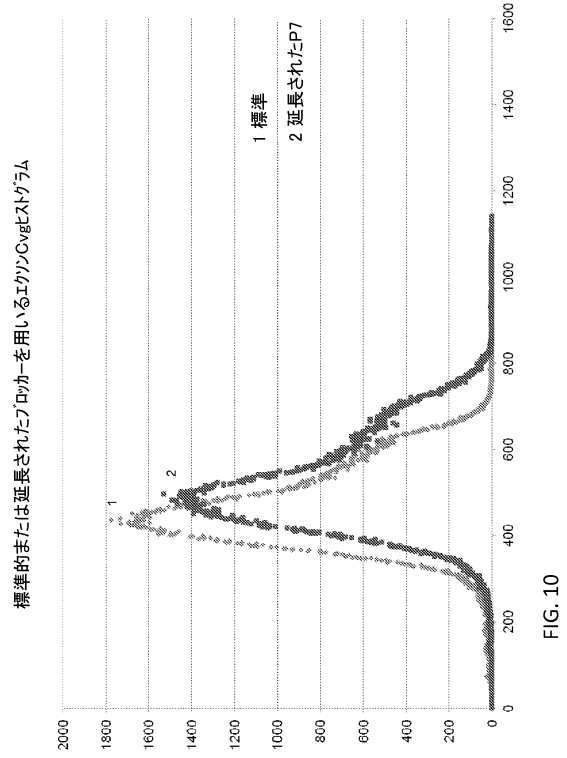


FIG. 10

【図 11】

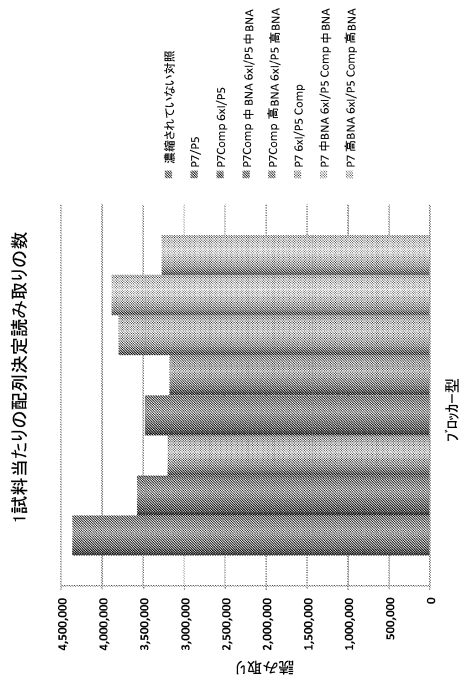


FIG. 11

【図 12】

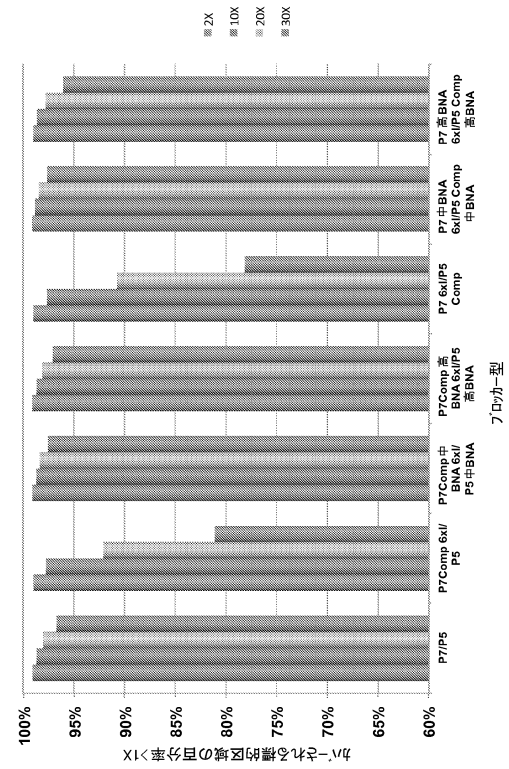


FIG. 12

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

弁理士法人川口国際特許事務所

(72)発明者 マーク・アーロン・ベールケ

アメリカ合衆国、アイオワ・５２２４１、コーラルビル、ティンバー・リッジ・コート・９６１

(72)発明者 ジョン・ロバート・ヘブンス

アメリカ合衆国、コロラド・８０４３９、エバーグリーン、ウィリアム・コーディー・ドライブ・８７４９

(72)発明者 スコット・ダニエル・ローズ

アメリカ合衆国、アイオワ・５２２４１、コーラルビル、メスキート・ドライブ・１１８０

(72)発明者 ミルナ・ジャロス

アメリカ合衆国、カリフォルニア・９４３０４、パロ・アルト、クラーク・ウェイ・９６２

(72)発明者 ザカリー・ズビルコ

アメリカ合衆国、マサチューセッツ・０２１３５、ブライトン、コモンウェルス・コート・１１、アパートメント・ナンバー・５

(72)発明者 ドロン・リブソン

アメリカ合衆国、マサチューセッツ・０２４５７、チェスナット・ヒル、ミドルセックス・ロード・１４２、ナンバー・１

(72)発明者 フランク・スー・ジュン

アメリカ合衆国、マサチューセッツ・０２４４６、ブルックライン、ビルス・ストリート・８５

審査官 長谷川 強

(56)参考文献 米国特許出願公開第２０１１／０２６３４３４（US，A１）

米国特許出願公開第２０１２／００１００９１（US，A１）

(58)調査した分野 (Int.Cl.，DB名)

C 1 2 N 1 5／1 1

C 1 2 Q 1／6 8 1 3

C 1 2 Q 1／6 8 5 1

C 1 2 Q 1／6 8 6

C 1 2 Q 1／6 8 6 9

CAplus／REGISTRY (STN)