

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6410726号
(P6410726)

(45) 発行日 平成30年10月24日 (2018.10.24)

(24) 登録日 平成30年10月5日 (2018.10.5)

(51) Int. Cl.	F I
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/09 Z N A Z
C 1 2 Q 1/6806 (2018.01)	C 1 2 Q 1/6806 Z
C 1 2 Q 1/686 (2018.01)	C 1 2 Q 1/686 Z

請求項の数 18 (全 193 頁)

(21) 出願番号	特願2015-545927 (P2015-545927)	(73) 特許権者	515138791
(86) (22) 出願日	平成25年12月10日 (2013.12.10)		レゾリューション バイオサイエンス、
(65) 公表番号	特表2015-536156 (P2015-536156A)		インコーポレイテッド
(43) 公表日	平成27年12月21日 (2015.12.21)		アメリカ合衆国 ワシントン 98005
(86) 国際出願番号	PCT/US2013/074102		, ベルビュー, 120ティーエイチ
(87) 国際公開番号	W02014/093330		アベニュー ノースイースト 2023,
(87) 国際公開日	平成26年6月19日 (2014.6.19)		スイート 100
審査請求日	平成28年11月16日 (2016.11.16)	(74) 代理人	100078282
(31) 優先権主張番号	61/794,049		弁理士 山本 秀策
(32) 優先日	平成25年3月15日 (2013.3.15)	(74) 代理人	100113413
(33) 優先権主張国	米国 (US)		弁理士 森下 夏樹
(31) 優先権主張番号	61/735,417	(74) 代理人	100181674
(32) 優先日	平成24年12月10日 (2012.12.10)		弁理士 飯田 貴敏
(33) 優先権主張国	米国 (US)	(74) 代理人	100181641
			弁理士 石川 大輔

最終頁に続く

(54) 【発明の名称】 標的化ゲノム解析のための方法

(57) 【特許請求の範囲】

【請求項 1】

標的化遺伝学的解析のための方法であって、
前記方法は、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュールとハイブリダイズさせるステップであって、前記タグ付けされたゲノムライブラリーが、多機能性アダプター分子にライゲーションされたゲノムDNAを含み、前記多機能性捕捉プローブモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズすることができる領域を含み、前記多機能性アダプター分子がさらに、1種または複数の核酸タグと、1種のPCRプライマー配列とを含む、ステップと、

(b) (a) から得られる前記タグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エクソヌクレアーゼ活性を有する酵素を使用して、(b) から得られる前記単離されたタグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体に対して3' - 5' エクソヌクレアーゼ酵素プロセッシングを実行して、一本鎖3'末端を除去するステップと、

(d) (c) から得られる前記酵素によりプロセッシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分がコピーされ、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能

10

20

性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

(e) (d) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップと

を含む、方法。

【請求項 2】

少なくとも 2 種のハイブリッド核酸分子を作製するために、ステップ (a) ~ (d) が少なくとも 2 回反復され、前記少なくとも 2 回の (a) ステップにおいて、少なくとも 2 種の異なる多機能性捕捉プローブモジュールが使用され、前記少なくとも 2 回の (a) ステップが、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる、請求項 1 に記載の方法。

10

【請求項 3】

少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする、請求項 2 に記載の方法。

【請求項 4】

標的化遺伝学的解析のための方法であって、

前記方法は、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュールとハイブリダイズさせるステップであって、前記タグ付けされたゲノムライブラリーが、多機能性アダプター分子にライゲーションされたゲノム DNA を含み、前記多機能性捕捉プローブモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズすることができる領域を含み、前記多機能性アダプター分子がさらに、1 種または複数の核酸タグと、1 種の PCR プライマー配列とを含む、ステップと、

20

(b) (a) から得られる前記タグ付けされたゲノムライブラリー多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するため (b) から得られる前記複合体に対して前記多機能性捕捉プローブの 5' から 3' への DNA ポリメラーゼ伸長を実行して、前記多機能性捕捉プローブの 3' にある前記捕捉されたタグ付けされたゲノム標的領域の一領域を複製するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記ゲノム標的領域とハイブリダイズする位置から 3' 方向に位置する前記タグ付けされたゲノム標的領域の一領域の相補体とを含むステップと、

30

(d) (c) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップと

を含む、方法。

【請求項 5】

少なくとも 2 種のハイブリッド核酸分子を作製するために、ステップ (a) ~ (c) が少なくとも 2 回反復され、前記少なくとも 2 回の (a) ステップにおいて、少なくとも 2 種の異なる多機能性捕捉プローブモジュールが使用され、前記少なくとも 2 回の (a) ステップが、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる、請求項 4 に記載の方法。

40

【請求項 6】

少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする、請求項 5 に記載の方法。

【請求項 7】

前記標的化遺伝学的解析が、前記ハイブリッド核酸分子のシーケンシングと、前記ハイブリッド核酸分子配列の配列アラインメントとを含む、請求項 1 ~ 6 のいずれか一項に記載の方法。

【請求項 8】

50

前記標的化遺伝学的解析がさらに、ゲノム配列変異を同定するため、または、コピー数変異を決定するための配列解析を含む、請求項 7 に記載の方法。

【請求項 9】

前記多機能性アダプター分子がさらに、試料コードタグ配列を含む、請求項 1 ～ 8 のいずれか一項に記載の方法。

【請求項 10】

特異的ゲノム標的領域のコピー数を決定するための方法であって、
前記方法は、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュールとハイブリダイズさせるステップであって、前記タグ付けされたゲノムライブラリーが、多機能性アダプター分子にライゲーションされたゲノム DNA を含み、前記多機能性捕捉プローブモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズすることができる領域を含み、前記多機能性アダプター分子がさらに、1 種または複数の核酸タグと、1 種の PCR プライマー配列とを含む、ステップと、

(b) (a) から得られる前記タグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、(b) から得られる前記単離されたタグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体に対して 3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖 3' 末端を除去するステップと、

(d) (c) から得られる前記酵素によりプロセッシングされた複合体に対して PCR 反応を実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分が複製され、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

(e) (d) における前記ハイブリッド核酸分子の PCR 増幅を実行するステップと、

(f) (e) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップと

を含む、方法。

【請求項 11】

ステップ (a) ～ (e) が少なくとも 2 回反復され、(f) の前記標的化遺伝学的解析が、前記少なくとも 2 回の (e) ステップから得られる前記ハイブリッド核酸分子配列の配列アラインメントを実行することを含む、請求項 10 に記載の方法。

【請求項 12】

前記少なくとも 2 回の (a) ステップにおいて、少なくとも 2 種の異なる多機能性捕捉プローブモジュールが使用され、前記少なくとも 2 回の (a) ステップが、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる、請求項 11 に記載の方法。

【請求項 13】

少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする、請求項 12 に記載の方法。

【請求項 14】

特異的ゲノム標的領域のコピー数を決定するための方法であって、

前記方法は、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュールとハイブリダイズさせるステップであって、前記タグ付けされたゲノムライブラリーが、多機能性アダプター分子にライゲーションされたゲノム DNA を含み、前記多機能性捕捉プローブモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズすることができる領域を含み、前記多機能性アダプター分子がさらに、1 種または複数の核酸タグと、1 種の PCR プライマー配列とを含む、ステップと、

(b) (a) から得られる前記タグ付けされたゲノムライブラリー多機能性捕捉プローブハイブリッド分子複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、(b) から得られる前記複合体に対して前記多機能性捕捉プローブの 5' から 3' への DNA ポリメラーゼ伸長を実行して、前記多機能性捕捉プローブの 3' にある前記捕捉されたタグ付けされたゲノム標的領域の一領域を複製するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記ゲノム標的領域とハイブリダイズする位置から 3' 方向に位置する前記タグ付けされたゲノム標的領域の一領域の相補体とを含むステップと、

(d) (c) における前記ハイブリッド核酸分子の PCR 増幅を実行するステップと、
(e) (d) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップと

を含む、方法。

【請求項 15】

ステップ (a) ~ (d) が少なくとも 2 回反復され、(e) の前記標的化遺伝学的解析が、前記少なくとも 2 回の (d) ステップから得られる前記ハイブリッド核酸分子配列の配列アラインメントを実行することを含む、請求項 14 に記載の方法。

【請求項 16】

前記少なくとも 2 回の (a) ステップにおいて、少なくとも 2 種の異なる多機能性捕捉プローブモジュールが使用され、前記少なくとも 2 回の (a) ステップが、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる、請求項 15 に記載の方法。

【請求項 17】

少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする、請求項 16 に記載の方法。

【請求項 18】

前記多機能性アダプター分子がさらに、試料コードタグ配列を含む、請求項 10 ~ 17 のいずれか一項に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の引用

本願は、2013 年 3 月 15 日に出願した米国仮出願第 61/794,049 号および 2012 年 12 月 10 日に提出した米国仮出願第 61/735,417 号の米国特許法第 119 条 (e) 項の下での利益を主張する。これらの出願は、その全体が参考として援用される。

【0002】

配列表に関する記載

本願に関連する配列表は、書面による複写物の代わりにテキスト形式で提供され、これにより、本明細書に参照として援用する。配列表を含有するテキストファイルの名称は、CLFK__001__02WO__ST25.txt である。このテキストファイルは、188 KB であり、2013 年 12 月 10 日に作成され、EFS-Web 経由で電子提出されている。

【0003】

背景

技術分野

本発明は、全般的には、単一アッセイにおいて標的化された特異的なゲノム遺伝子座の遺伝的配列および染色体コピー数の両方を明らかにする、個体における遺伝学的解析のための方法に関する。特に、本発明は、標的遺伝子配列または遺伝子転写物の高感度かつ特異的な検出をもたらす方法と、単一アッセイにおいてバリエーション配列および全体的な遺伝

10

20

30

40

50

子コピー数の両方を明らかにする方法に関する。

【背景技術】

【0004】

関連出願の説明

個々のヒト対象の完全ヒトゲノム配列および部分的ゲノムリシーケンシング試験の両方により、あらゆるヒトが、完全とは言えないゲノムを保有すると考えられるという基本テーマが明らかになった。特に、正常な健康ヒト対象は、そのゲノム配列内に数千種まではいかないとしても数百種の遺伝的病変を持つことが判明した。このような病変の多くは、病変が存する遺伝子の機能を排除することが公知であるまたは予測される。正常二倍体のヒトは、大部分の遺伝子の機能的コピーを2個保有するが、あらゆるヒトにおいて、機能的遺伝子コピーが僅か1（またはゼロ）個存在するという事例が多くあることが暗示される。同様に、遺伝子が、遺伝子重複／増幅事象により過剰出現するという事例にも有意な頻度で遭遇する。

10

【0005】

生体ネットワークにおける重要な特色の1つは、機能的冗長性である。正常な健康個体は、1コピーの損失の重要度が低くなるように、全遺伝子を平均して2コピー保有するため、遺伝的病変の平均負荷に耐容性を示すことができる。さらに、特異的遺伝子機能における軽微な攪乱が、機能的要素のより大きなネットワーク内で全般的に補償されるように、遺伝子のセットは、多くの場合、同様の機能を実行する。生体システムにおける機能的補償は、一般テーマであるが、特異的遺伝子損失が、急性の破壊的な事象を誘発し得るという事例が多くある。例として、がんは、複数の個々の病変の複合効果が、制御されない細胞増殖である遺伝的疾患の結果であると考えられる。同様に、処方薬は、多くの場合、非常に特異的な遺伝子により輸送、代謝および／または排除される特異的な化学物質である。このような遺伝子における攪乱は、正常環境下では一般に重要度が低い、化学治療法において有害事象（例えば、副作用）として顕在化し得る。

20

【0006】

「精密医学」と称されることがますます多くなった「個別化医学」の中心的な目標は、患者に特異的な遺伝情報を、個体の遺伝的プロファイルと適合する処置オプションと統合することである。しかし、個別化医学の莫大な潜在力は、まだ現実化されていない。この目標を現実化するためには、関連性のある遺伝子の遺伝的状态を確実に決定することができる、臨床的に許容される頑強な遺伝的診断検査が存在する必要がある。

30

【発明の概要】

【課題を解決するための手段】

【0007】

簡単な概要

本明細書中で企図される特定の実施形態は、タグ付けされたDNAライブラリーを作製するための方法であって、断片化されたDNAを末端修復酵素で処理して、断片化され末端修復されたDNAを作製するステップと、ランダム核酸タグ配列ならびに必要な応じて試料コード配列および／またはPCRプライマー配列を、前記断片化され末端修復されたDNAにライゲーションして、前記タグ付けされたDNAライブラリーを作製するステップとを含む方法を提供する。

40

【0008】

特定の実施形態では、前記ランダム核酸タグ配列は、約2～約100ヌクレオチドである。一部の実施形態では、本発明は、前記ランダム核酸タグ配列が、約2～約8ヌクレオチドであることを提供する。

【0009】

ある特定の実施形態では、前記断片化され末端修復されたDNAは、平滑末端を含有する。一部の実施形態では、前記平滑末端は、単一塩基対オーバーハングを含有するようさらに修飾される。

【0010】

50

ある特定の実施形態では、前記ライゲーションステップは、多機能性アダプターモジュールを前記断片化され末端修復されたDNAにライゲーションして、前記タグ付けされたDNAライブラリーを作製することを含み、前記多機能性アダプター分子は、

- i) ランダム核酸タグ配列を含む第1の領域と、
- i i) 試料コード配列を含む第2の領域と、
- i i i) PCRプライマー配列を含む第3の領域と

を含む。

【0011】

さらなる実施形態では、この方法は、タグ付けされたDNAライブラリーを少なくとも1種の多機能性捕捉プローブモジュールとハイブリダイズさせて、複合体を形成するステップをさらに含み、前記多機能性捕捉プローブモジュールが、前記DNAライブラリーにおける特異的標的領域とハイブリダイズする。

10

【0012】

さらなる実施形態では、この方法は、前記タグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップをさらに含む。

【0013】

一部の実施形態では、この方法は、前記単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を3'-5'エキソヌクレアーゼ酵素プロセッシングして、一本鎖3'末端を除去するステップをさらに含む。一部の実施形態では、前記3'-5'エキソヌクレアーゼ酵素プロセッシングにおいて使用される酵素は、T4ポリメラーゼである。

20

【0014】

特定の実施形態において、方法は、単離されたタグ付けされたDNAライブラリー断片を鋳型として利用して、多機能性捕捉プローブの3'末端からの、単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体の5'-3'DNAポリメラーゼ伸長をさらに含む。

【0015】

ある特定の実施形態において、方法は、5'FLAPエンドヌクレアーゼ、DNA重合およびDNAリガーゼによるニック閉鎖の協奏的作用により、多機能性捕捉プローブおよび単離されたタグ付けされたDNAライブラリー断片を連結するステップをさらに含む。

30

【0016】

さらなる実施形態では、この方法は、前記3'-5'エキソヌクレアーゼ酵素によりプロセッシングされた複合体に対してPCRを実行するステップをさらに含み、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分がコピーされ、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含む。

【0017】

種々の実施形態では、標的化遺伝学的解析のための方法が提供され、この方法は、

a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュールが、前記DNAライブラリーにおける特異的DNA標的領域と選択的にハイブリダイズするステップと、

40

b) a) から得られる前記タグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、

c) 3'-5'エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体に対して3'-5'エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖3'末端を除去するステップと、

d) c) から得られる前記酵素によりプロセッシングされた複合体に対してPCRを実行

50

するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分がコピーされ、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

e) d) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む。

【0018】

様々な特定の実施形態において、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、c) この単離されたタグ付けされたDNAライブラリー断片を鋳型として利用して、多機能性捕捉プローブの5' - 3' DNAポリメラーゼ伸長を実行するステップと、d) c) から得られる酵素によりプロセシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために単離された標的領域の相補体がコピーされ、このハイブリッド核酸分子が、DNA標的領域の相補体、多機能性捕捉プローブの標的特異的領域および多機能性捕捉プローブモジュールテイル配列を含むステップと、e) d) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む、標的化遺伝学的解析のための方法が提供される。

【0019】

様々なある特定の実施形態において、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、c) 5' FLAPエンドヌクレアーゼ、DNA重合およびDNAリガーゼによるニック閉鎖の協奏的作用により、ハイブリッド多機能性捕捉プローブによって単離されたタグ付けされたDNA標的分子の作製を実行するステップと、d) c) から得られる酵素によりプロセシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために、この多機能性捕捉プローブ分子が、単離されたタグ付けされたDNA標的クローンに連結され、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができるゲノム標的領域と、多機能性捕捉プローブモジュールの相補体とを含むステップと、e) d) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む、標的化遺伝学的解析のための方法が提供される。

【0020】

特定の実施形態において、特異的標的領域のコピー数を決定するための方法であって、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体に対して3' - 5' エキソヌクレアーゼ酵素プロセシングを実行して、一本鎖3' 末端を除去するステップと、d) c) から得られる酵素によりプロセシングされた複合体に対してPCR反応を実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分が複製され、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができる標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、e)

d) におけるハイブリッド核酸のPCR増幅を実行するステップと、e) d) におけるPCR反応を定量化するステップであって、定量化が、特異的標的領域のコピー数の決定を可能にするステップとを含む方法が提供される。

【0021】

ある特定の実施形態において、特異的標的領域のコピー数を決定するための方法であって、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、c) この単離されたタグ付けされたDNAライブラリー断片を鋳型として利用して、多機能性捕捉プローブの5' - 3' DNAポリメラーゼ伸長を実行するステップと、d) c) から得られる酵素によりプロセシングされた複合体に対してPCR反応を実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分が複製され、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができる標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、e) d) におけるハイブリッド核酸のPCR増幅を実行するステップと、e) d) におけるPCR反応を定量化するステップであって、定量化が、特異的標的領域のコピー数の決定を可能にするステップとを含む方法が提供される。

10

【0022】

さらなる実施形態において、特異的標的領域のコピー数を決定するための方法であって、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、c) 5' FLAPエンドヌクレアーゼ、DNA重合およびDNAリガーゼによるニック閉鎖の協奏的作用により、ハイブリッド多機能性捕捉プローブによって単離されたタグ付けされたDNA標的分子の作製を実行するステップと、d) c) から得られる酵素によりプロセシングされた複合体に対してPCR反応を実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分が複製され、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができる標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、e) d) におけるハイブリッド核酸のPCR増幅を実行するステップと、e) d) におけるPCR反応を定量化するステップであって、定量化が、特異的標的領域のコピー数の決定を可能にするステップとを含む方法が提供される。

20

30

【0023】

追加的な実施形態において、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブハイブリッドモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、c) ハイブリッド核酸分子を作製するために、b) から得られる複合体に対してPCRを実行して、多機能性捕捉プローブの配列と比べて3'の領域を複製するステップであって、このハイブリッド核酸分子が、多機能性捕捉プローブハイブリッドモジュールと、多機能性捕捉プローブと比べて3'に位置するタグ付けされたDNAライブラリー配列の領域の相補体とを含むステップと、d) c) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む、標的化遺伝学的解析のための方法が提供される。

40

【0024】

特定の実施形態において、a) タグ付けされたDNAライブラリーを多機能性捕捉プロ

50

ープハイブリッドモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブハイブリッドモジュールが、ゲノムライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、c) 単離されたタグ付けされたDNAライブラリー断片を鋳型として利用して、多機能性捕捉プローブの5' - 3' DNAポリメラーゼ伸長を実行するステップであって、このハイブリッド核酸分子が、多機能性捕捉プローブハイブリッドモジュールと、多機能性捕捉プローブと比べて3' に位置するタグ付けされたDNAライブラリー配列の領域の相補体とを含むステップと、d) c) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む、標的化遺伝学的解析のための方法が提供される。

10

【0025】

ある特定の実施形態において、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブハイブリッドモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、c) 5' FLAPエンドヌクレアーゼ、DNA重合およびDNAリガーゼによるニック閉鎖の協奏的作用により、ハイブリッド多機能性捕捉プローブによって単離されたタグ付けされたDNA標的分子の作製を実行するステップであって、このハイブリッド核酸分子が、多機能性捕捉プローブハイブリッドモジュールの相補体と、多機能性捕捉プローブと比べて5' に位置するタグ付けされたDNAライブラリー配列の領域を含むステップと、d) c) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む、標的化遺伝学的解析のための方法が提供される。

20

【0026】

特定の実施形態において、特異的標的領域のコピー数を決定するための方法であって、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブハイブリッドモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、c) ハイブリッド核酸分子を作製するために、b) から得られる複合体に対してPCRを実行して、多機能性捕捉プローブの配列と比べて3' である領域を複製するステップであって、このハイブリッド核酸分子が、多機能性捕捉プローブハイブリッドモジュールと、多機能性捕捉プローブと比べて3' に位置するタグ付けされたDNAライブラリー配列の領域の相補体とを含むステップと、d) c) におけるハイブリッド核酸のPCR増幅を実行するステップと、e) d) におけるPCR反応を定量化するステップであって、定量化が、特異的標的領域のコピー数の決定を可能にするステップとを含む方法が提供される。

30

【0027】

様々な実施形態において、標的化遺伝学的解析は、配列解析である。

40

【0028】

特定の実施形態において、タグ付けされたDNAライブラリーは、PCRにより増幅されて、増幅されたタグ付けされたDNAライブラリーを作製する。

【0029】

ある特定の実施形態において、DNAは、血液、皮膚、毛、毛包、唾液、口腔粘膜、膣粘膜、汗、涙、上皮組織、尿、精液、精子液 (seminal fluid)、精漿、前立腺液、尿道球腺液 (カウパー氏腺液)、排泄物、生検、腹水、脳脊髄液、リンパ液および組織抽出物試料または生検試料からなる群から選択される生体試料に由来する。

【0030】

さらなる実施形態において、タグ付けされたDNAライブラリーは、タグ付けされたD

50

N A配列を含み、各タグ付けされたD N A配列は、i) 断片化され末端修復されたD N Aと、i i) ランダムヌクレオチドタグ配列と、i i i) 試料コード配列と、i v) P C Rプライマー配列とを含む。

【0031】

追加的な実施形態において、ハイブリッドタグ付けされたD N Aライブラリーは、標的化遺伝学的解析において使用されるハイブリッドタグ付けされたD N A配列を含み、各ハイブリッドタグ付けされたD N A配列は、i) 断片化され末端修復されたD N Aと、i i) ランダムヌクレオチドタグ配列と、i i i) 試料コード配列と、i v) P C Rプライマー配列と、v) 多機能性捕捉プローブモジュールテイル配列とを含む。

【0032】

さらなる実施形態において、多機能性アダプターモジュールは、i) ランダムヌクレオチドタグ配列を含む第1の領域と、i i) 試料コード配列を含む第2の領域と、i i i) P C Rプライマー配列を含む第3の領域とを含む。

【0033】

特定の実施形態において、多機能性捕捉プローブモジュールは、i) パートナーオリゴヌクレオチドとハイブリダイズすることができる第1の領域と、i i) 特異的標的領域とハイブリダイズすることができる第2の領域と、i i i) テイル配列を含む第3の領域とを含む。一部の実施形態において、捕捉プローブモジュールの第1の領域は、パートナーオリゴヌクレオチドに結合している。一部の実施形態において、パートナーオリゴヌクレオチドは、化学修飾されている。

【0034】

一実施形態において、組成物は、タグ付けされたD N Aライブラリーと、多機能性アダプターモジュールと、多機能性捕捉プローブモジュールとを含む。

【0035】

特定の実施形態において、組成物は、本発明の方法によるハイブリッドタグ付けされたゲノムライブラリーを含む。

【0036】

ある特定の実施形態において、組成物は、本明細書において企図される方法を実行するための反応混合物を含む。

【0037】

特定の実施形態において、タグ付けされたD N Aライブラリーを作製することができる反応混合物は、a) 断片化されたD N Aと、b) 断片化され末端修復されたD N Aを作製するためのD N A末端修復酵素とを含む。

【0038】

ある特定の実施形態において、反応混合物は、多機能性アダプターモジュールをさらに含む。

【0039】

追加的な実施形態において、反応混合物は、多機能性捕捉プローブモジュールをさらに含む。

【0040】

一部の実施形態において、反応混合物は、3' - 5' エキソヌクレアーゼ活性およびP C R増幅活性を有する酵素をさらに含む。

【0041】

一実施形態において、反応混合物は、F L A Pエンドヌクレアーゼと、D N Aポリメラーゼと、D N Aリガーゼとを含む。

【0042】

前述の実施形態のいずれかにおいて、D N Aは、単離されたゲノムD N Aまたはc D N Aとなり得る。

【0043】

様々な実施形態において、タグ付けされたゲノムライブラリーを作製するための方法で

10

20

30

40

50

あって、断片化されたゲノムDNAを末端修復酵素で処理して、断片化され末端修復されたゲノムDNAを作製するステップと、ランダム核酸タグ配列ならびに必要なに応じて試料コード配列および／またはPCRプライマー配列を、前記断片化され末端修復されたゲノムDNAにライゲーションして、タグ付けされたゲノムライブラリーを作製するステップとを含む方法が提供される。

【0044】

特定の実施形態において、ランダム核酸タグ配列は、約2～約100ヌクレオチドである。

【0045】

ある特定の実施形態において、ランダム核酸タグ配列は、約2～約8ヌクレオチドである。

10

【0046】

追加的な実施形態において、断片化され末端修復されたゲノムDNAは、平滑末端を含有する。

【0047】

さらなる実施形態において、平滑末端は、単一塩基対オーバーハングを含有するように修飾される。

【0048】

一部の実施形態において、ライゲーションステップは、断片化され末端修復されたゲノムDNAに多機能性アダプターモジュールをライゲーションして、タグ付けされたゲノムライブラリーを作製することを含み、多機能性アダプター分子は、ランダム核酸タグ配列を含む第1の領域と、試料コード配列を含む第2の領域と、PCRプライマー配列を含む第3の領域とを含む。

20

【0049】

特定の実施形態において、本明細書において企図される方法は、タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュールとハイブリダイズさせて、複合体を形成するステップを含み、この多機能性捕捉プローブモジュールは、ゲノムライブラリーにおける特異的ゲノム標的領域とハイブリダイズする。

【0050】

ある特定の実施形態において、本明細書において企図される方法は、タグ付けされたゲノムライブラリー多機能性捕捉プローブモジュール複合体を単離するステップを含む。

30

【0051】

追加的な特定の実施形態において、本明細書において企図される方法は、単離されたタグ付けされたゲノムライブラリー多機能性捕捉プローブモジュール複合体を3'－5'エキソヌクレアーゼ酵素プロセッシングして、一本鎖3'末端を除去するステップを含む。

【0052】

さらなる特定の実施形態において、3'－5'エキソヌクレアーゼ酵素プロセッシングにおいて使用される酵素は、T4 DNAポリメラーゼである。

【0053】

一部の特定の実施形態において、本明細書において企図される方法は、先行する請求項から得られる3'－5'エキソヌクレアーゼ酵素によりプロセッシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分がコピーされ、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができるゲノム標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含む。

40

【0054】

様々な実施形態において、(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、ゲノムライブラリーにおける特異的ゲノム標的領域に選択的とハイブリダイズするステップと、(b) a) から得られるタグ付けされたゲノムライブラリー

50

ー多機能性捕捉プローブモジュール複合体を単離するステップと、(c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる単離されたタグ付けされたゲノムライブラリーー多機能性捕捉プローブモジュール複合体に対して3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖3' 末端を除去するステップと、(d) c) から得られる酵素によりプロセッシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分がコピーされ、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができるゲノム標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、(e) d) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む、標的化遺伝学的解析のための方法が提供される。

10

【0055】

特定の実施形態では、ステップa) ~ d) は、少なくとも約2回反復され、e) の前記標的化遺伝学的解析は、前記少なくとも2回のd) ステップから得られるハイブリッド核酸分子配列の配列アラインメントを含む。

【0056】

さらなる実施形態では、少なくとも2種の異なる多機能性捕捉プローブモジュールは、少なくとも2回のa) ステップにおいて使用され、前記少なくとも2回のa) ステップは、それぞれ1種の多機能性捕捉プローブモジュールを用いる。

【0057】

20

一部の実施形態では、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする。

【0058】

種々の実施形態では、特異的ゲノム標的領域のコピー数を決定するための方法が提供され、この方法は、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュール複合体が、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズするステップと、

30

(b) a) から得られる前記タグ付けされたゲノムライブラリーー多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされたゲノムライブラリーー多機能性捕捉プローブモジュール複合体に対して3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖3' 末端を除去するステップと、

(d) c) から得られる前記酵素によりプロセッシングされた複合体に対してPCR反応を実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分が複製され、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

40

(e) d) における前記ハイブリッド核酸分子のPCR増幅を実行するステップと、

(f) e) におけるPCR反応を定量化するステップであって、前記定量化が、前記特異的ゲノム標的領域のコピー数の決定を可能にするステップとを含む。

【0059】

一部の実施形態では、本明細書中で企図される方法は、ステップe) から得られる前記ハイブリッド核酸分子の配列を得るステップを含む。

【0060】

さらなる実施形態では、ステップa) ~ e) は、少なくとも約2回反復され、前記少な

50

くとも2回のe)ステップから得られる前記ハイブリッド核酸分子配列を使用して、配列アラインメントが実行される。

【0061】

追加的な実施形態では、少なくとも2種の異なる多機能性捕捉プローブモジュールは、前記少なくとも2回のa)ステップにおいて使用され、前記少なくとも2回のa)ステップは、それぞれ1種の多機能性捕捉プローブモジュールを用いる。

【0062】

ある特定の実施形態では、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の上流とハイブリダイズする。

10

【0063】

種々の実施形態では、特異的ゲノム標的領域のコピー数を決定するための方法が提供され、この方法は、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュール複合体が、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体に対して3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖3' 末端を除去するステップと、

20

(d) c) から得られる前記酵素によりプロセッシングされた複合体に対してPCR反応を実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分が複製され、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

(e) d) における前記ハイブリッド核酸分子のPCR増幅を実行するステップと、を含む。

30

【0064】

ある特定の実施形態では、本明細書中で企図される方法は、ステップe) から得られる前記ハイブリッド核酸分子の配列を得るステップを含む。

【0065】

特定の実施形態では、ステップa) ~ e) は、少なくとも約2回反復され、前記少なくとも2回のe)ステップから得られる前記ハイブリッド核酸分子配列を使用して、配列アラインメントが実行される。

【0066】

一部の実施形態では、少なくとも2種の異なる多機能性捕捉プローブモジュールは、前記少なくとも2回のa)ステップにおいて使用され、前記少なくとも2回のa)ステップは、それぞれ1種の多機能性捕捉プローブモジュールを用いる。

40

【0067】

追加的な実施形態では、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の上流とハイブリダイズする。

【0068】

種々の実施形態では、特異的ゲノム標的領域のコピー数を決定するための方法が提供され、この方法は、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュール複合体が

50

、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体に対して3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖3' 末端を除去するステップと、

(d) c) から得られる前記酵素によりプロセッシングされた複合体に対してPCR反応を実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分が複製され、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

(e) d) における前記ハイブリッド核酸分子のPCR増幅を実行するステップと、

(f) e) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む。

【0069】

特定の実施形態では、ステップa) ~ e) は、少なくとも約2回反復され、f) の前記標的化遺伝学的解析は、前記少なくとも2回のe) ステップから得られる前記ハイブリッド核酸分子配列の配列アラインメントの実行を含む。

【0070】

ある特定の実施形態では、少なくとも2種の異なる多機能性捕捉プローブモジュールは、前記少なくとも2回のa) ステップにおいて使用され、前記少なくとも2回のa) ステップは、それぞれ1種の多機能性捕捉プローブモジュールを用いる。

【0071】

追加的な実施形態では、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の上流とハイブリダイズする。

【0072】

種々の実施形態では、標的化遺伝学的解析のための方法が提供され、この方法は、(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、b) から得られる前記複合体に対して前記多機能性捕捉プローブの5' から3' へのDNAポリメラーゼ伸長を実行して、前記多機能性捕捉プローブの3' にある前記捕捉されたタグ付けされたゲノム標的領域の一領域を複製するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記ゲノム標的領域とハイブリダイズする位置から3' 方向に位置する前記タグ付けされたゲノム標的領域の一領域の相補体とを含むステップと、

(d) c) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む。

【0073】

さらなる実施形態では、ステップa) ~ c) は、少なくとも約2回反復され、d) の前記標的化遺伝学的解析は、前記少なくとも2回のd) ステップから得られる前記ハイブリ

10

20

30

40

50

ッド核酸分子配列の配列アラインメントを含む。

【0074】

一部の実施形態では、少なくとも2種の異なる多機能性捕捉プローブモジュールは、前記少なくとも2回のa)ステップにおいて使用され、前記少なくとも2回のa)ステップは、それぞれ1種の多機能性捕捉プローブモジュールを用いる。

【0075】

特定の実施形態では、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の上流とハイブリダイズする。

【0076】

種々の実施形態では、特異的ゲノム標的領域のコピー数を決定するための方法が提供され、この方法は、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、b) から得られる前記複合体に対して前記多機能性捕捉プローブの5'から3'へのDNAポリメラーゼ伸長を実行して、前記多機能性捕捉プローブの3'にある前記捕捉されたタグ付けされたゲノム標的領域の一領域を複製するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記ゲノム標的領域とハイブリダイズする位置から3'方向に位置する前記タグ付けされたゲノム標的領域の一領域の相補体とを含むステップと、

(d) c) における前記ハイブリッド核酸分子のPCR増幅を実行するステップと、

(e) d) におけるPCR反応を定量化するステップであって、前記定量化が、前記特異的ゲノム標的領域のコピー数の決定を可能にするステップとを含む。

【0077】

特定の実施形態では、本明細書中で企図される方法は、ステップd) から得られる前記ハイブリッド核酸分子の配列を得るステップを含む。

【0078】

ある特定の実施形態では、ステップa)～d)は、少なくとも約2回反復され、前記少なくとも2回のd)ステップから得られる前記ハイブリッド核酸分子の配列アラインメントが行われる。

【0079】

追加的な実施形態では、少なくとも2種の異なる多機能性捕捉プローブモジュールは、前記少なくとも2回のa)ステップにおいて使用され、前記少なくとも2回のa)ステップは、それぞれ1種の多機能性捕捉プローブモジュールを用いる。

【0080】

さらなる実施形態では、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の上流とハイブリダイズする。

【0081】

一部の実施形態では、前記標的化遺伝学的解析は、配列解析である。

【0082】

特定の実施形態では、前記タグ付けされたゲノムライブラリーがPCRにより増幅されて、増幅されたタグ付けされたゲノムライブラリーを作製する。

【0083】

関連の特定の実施形態では、前記ゲノムDNAは、血液、皮膚、毛、毛包、唾液、口腔粘膜、腔粘膜、汗、涙、上皮組織、尿、精液、精子液、精漿、前立腺液、尿道球腺液（カウパー氏腺液）、排泄物、生検、腹水、脳脊髄液、リンパ液および組織抽出物試料または生検試料からなる群から選択される生体試料に由来する。

【0084】

種々の実施形態では、タグ付けされたゲノム配列を含むタグ付けされたゲノムライブラリーが提供され、ここで、各タグ付けされたゲノム配列は、断片化され末端修復されたゲノムDNAと、ランダムヌクレオチドタグ配列と、試料コード配列と、PCRプライマー配列とを含む。

【0085】

種々の関連の実施形態では、タグ付けされたcDNA配列を含むタグ付けされたcDNAライブラリーが提供され、ここで、各タグ付けされたcDNA配列は、断片化され末端修復されたcDNAと、ランダムヌクレオチドタグ配列と、試料コード配列と、PCRプライマー配列とを含む。

【0086】

種々の特定の実施形態では、標的化遺伝学的解析において使用するためのハイブリッドタグ付けされたゲノム配列を含むハイブリッドタグ付けされたゲノムライブラリーが提供され、ここで各ハイブリッドタグ付けされたゲノム配列は、断片化され末端修復されたゲノムDNAと、ランダムヌクレオチドタグ配列と、試料コード配列と、PCRプライマー配列と、ゲノム標的領域と、多機能性捕捉プローブモジュールテイル配列とを含む。

【0087】

種々のある特定の実施形態では、標的化遺伝学的解析において使用するためのハイブリッドタグ付けされたcDNA配列を含むハイブリッドタグ付けされたcDNAライブラリーが提供され、ここで各ハイブリッドタグ付けされたcDNA配列は、断片化され末端修復されたcDNAと、ランダムヌクレオチドタグ配列と、試料コード配列と、PCRプライマー配列と、cDNA標的領域と、多機能性捕捉プローブモジュールテイル配列とを含む。

【0088】

種々の特定の実施形態では、多機能性アダプターモジュールが提供され、この多機能性アダプターモジュールは、ランダムヌクレオチドタグ配列を含む第1の領域と、試料コード配列を含む第2の領域と、PCRプライマー配列を含む第3の領域とを含む。

【0089】

種々の追加的な実施形態では、多機能性捕捉プローブモジュールが提供され、この多機能性捕捉プローブモジュールは、パートナーオリゴヌクレオチドとハイブリダイズすることができる第1の領域と、特異的ゲノム標的領域とハイブリダイズすることができる第2の領域と、テイル配列を含む第3の領域とを含む。

【0090】

特定の実施形態では、前記第1の領域は、パートナーオリゴヌクレオチドに結合している。

【0091】

特定の実施形態では、多機能性アダプタープローブハイブリッドモジュールが提供され、この多機能性アダプタープローブハイブリッドモジュールは、パートナーオリゴヌクレオチドとハイブリダイズすることができ、PCRプライマーとして機能することができる第1の領域と、特異的ゲノム標的領域とハイブリダイズすることができる第2の領域とを含む。

【0092】

ある特定の実施形態では、前記第1の領域は、パートナーオリゴヌクレオチドに結合している。

【0093】

一部の実施形態では、前記パートナーオリゴヌクレオチドは、化学修飾されている。

10

20

30

40

50

【0094】

さらなる実施形態では、タグ付けされたゲノムライブラリーと、多機能性アダプターモジュールと、多機能性捕捉プローブモジュールとを含む組成物が提供される。

【0095】

追加的な実施形態では、先行する実施形態のいずれかに記載のハイブリッドタグ付けされたゲノムライブラリーまたはcDNAライブラリーを含む組成物が提供される。

【0096】

種々の実施形態では、先行する実施形態のいずれか一つに記載の方法を実行するための反応混合物が提供される。

【0097】

特定の実施形態では、タグ付けされたゲノムライブラリーを作製することができる反応混合物が提供され、この反応混合物は、断片化されたゲノムDNAと、断片化され末端修復されたゲノムDNAを作製するためのDNA末端修復酵素とを含む。

【0098】

特定の実施形態では、タグ付けされたゲノムライブラリーを作製することができる反応混合物が提供され、この反応混合物は、断片化されたcDNAと、断片化され末端修復されたcDNAを作製するためのDNA末端修復酵素とを含む。

【0099】

特定の実施形態では、多機能性アダプターモジュールを含む、反応混合物が提供される。

【0100】

一部の実施形態では、反応混合物は、多機能性捕捉プローブモジュールを含む。

【0101】

ある特定の実施形態では、反応混合物は、3' - 5' エキソヌクレアーゼ活性およびPCR増幅活性を有する酵素を含む。

【0102】

種々の実施形態では、DNA配列解析のための方法が提供され、この方法は、各クローンが第1のDNA配列および第2のDNA配列を含む1種または複数のクローンを得るステップであって、前記第1のDNA配列が、標的化ゲノムDNA配列を含み、前記第2のDNA配列が、捕捉プローブ配列を含むステップと、前記1種または複数のクローンに対してペアエンドシーケンシング反応を実行して1種または複数のシーケンシングリードを得るステップと、前記シーケンシングリードの前記プローブ配列に従って、前記1種または複数のクローンの前記シーケンシングリードを順序付けまたはクラスター形成するステップとを含む。

【0103】

特定の実施形態では、DNA配列解析のための方法が提供され、この方法は、各クローンが第1のDNA配列および第2のDNA配列を含む1種または複数のクローンを得るステップであって、前記第1のDNA配列が、標的化ゲノムDNA配列を含み、前記第2のDNA配列が、捕捉プローブ配列を含むステップと、前記1種または複数のクローンに対してシーケンシング反応を実行するステップであって、約100ヌクレオチドを超える単一の長いシーケンシングリードが得られ、前記リードが、前記第1のDNA配列および前記第2のDNA配列の両方の同定に十分であるステップと、前記シーケンシングリードの前記プローブ配列に従って、前記1種または複数のクローンのシーケンシングリードを順序付けまたはクラスター形成するステップとを含む。

【0104】

ある特定の実施形態では、前記1種または複数のクローンの配列は、1種または複数のヒト参照DNA配列と比較される。

【0105】

追加的な実施形態では、前記1種または複数のヒト参照DNA配列にマッチしない配列が同定される。

10

20

30

40

50

【0106】

さらなる実施形態では、非マッチ配列が使用されて、前記非マッチ配列データからデノボアセンブリーを作成する。

【0107】

一部の実施形態では、前記デノボアセンブリーが使用されて、前記捕捉プローブに関連する新規配列再編成を同定する。

【0108】

種々の実施形態では、ゲノムコピー数決定解析のための方法が提供され、この方法は、各クローンが第1のDNA配列および第2のDNA配列を含む1種または複数のクローンを得るステップであって、前記第1のDNA配列が、ランダムヌクレオチドタグ配列および標的化ゲノムDNA配列を含み、前記第2のDNA配列が、捕捉プローブ配列を含むステップと、前記1種または複数のクローンに対してペアエンドシーケンシング反応を実行して1種または複数のシーケンシングリードを得るステップと、前記シーケンシングリードの前記プローブ配列に従って、前記1種または複数のクローンの前記シーケンシングリードを順序付けまたはクラスター形成するステップとを含む。

10

【0109】

一部の実施形態では、ゲノムコピー数決定解析のための方法が提供され、この方法は、各クローンが第1のDNA配列および第2のDNA配列を含む1種または複数のクローンを得るステップであって、前記第1のDNA配列が、ランダムヌクレオチドタグ配列および標的化ゲノムDNA配列を含み、前記第2のDNA配列が、捕捉プローブ配列を含むステップと、前記1種または複数のクローンに対してシーケンシング反応を実行するステップであって、約100ヌクレオチドを超える単一の長いシーケンシングリードが得られ、前記リードが、前記第1のDNA配列および前記第2のDNA配列の両方の同定に十分であるステップと、前記シーケンシングリードの前記プローブ配列に従って、前記1種または複数のクローンのシーケンシングリードを順序付けまたはクラスター形成するステップとを含む。

20

【0110】

ある特定の実施形態では、前記ランダムヌクレオチドタグ配列は、約2～約50ヌクレオチドの長さである。

【0111】

さらなる実施形態では、本明細書中で企図される方法は、独特のシーケンシングリードおよび重複したシーケンシングリード(unique and redundant sequencing reads)の分布を決定し、独特のリードが遭遇する回数を計数し、前記独特のリードの度数分布を統計分布に適合させ、独特のリードの総数を推量し、推量された独特のリードの前記総数を、大部分のヒト遺伝子座が一般に二倍体であるという仮定に対して正規化することにより、第2のリード配列に関連するあらゆるシーケンシングリードを解析するステップを含む。

30

【0112】

追加的な実施形態では、1種または複数の標的化遺伝子座の推量されたコピー数が決定される。

40

【0113】

一部の実施形態では、予想されるコピー数値から逸脱する前記1種または複数の標的遺伝子座が決定される。

【0114】

さらなる実施形態では、遺伝子の前記1種または複数の標的化遺伝子座は、遺伝子座のコレクションにおいて共にグループ化され、標的化遺伝子座のコレクションから得られる前記コピー数測定値は、平均化および正規化される。

【0115】

追加的な実施形態では、遺伝子の前記推量されたコピー数は、この遺伝子を表す全標的遺伝子座の前記正規化された平均によって表される。

50

【0116】

ある特定の実施形態では、タグ付けされたRNA発現ライブラリーを作製するための方法が提供され、この方法は、cDNAライブラリーを断片化するステップと、前記断片化されたcDNAライブラリーを末端修復酵素で処理して、断片化され末端修復されたcDNAを作製するステップと、多機能性アダプター分子を前記断片化され末端修復されたcDNAにライゲーションして、タグ付けされたRNA発現ライブラリーを作製するステップとを含む。

【0117】

特定の実施形態では、タグ付けされたRNA発現ライブラリーを作製するための方法が提供され、この方法は、1個または複数の細胞の全RNAからcDNAライブラリーを調製するステップと、前記cDNAライブラリーを断片化するステップと、前記断片化されたcDNAを末端修復酵素で処理して、断片化され末端修復されたcDNAを作製するステップと、多機能性アダプター分子を前記断片化され末端修復されたcDNAにライゲーションして、タグ付けされたRNA発現ライブラリーを作製するステップとを含む。

10

【0118】

種々の実施形態では、前記cDNAライブラリーは、オリゴdTプライミングcDNAライブラリーである。

【0119】

特定の実施形態では、前記cDNAライブラリーは、約6～約20ランダムヌクレオチドを含むランダムオリゴヌクレオチドによってプライミングされる。

20

【0120】

ある特定の実施形態では、前記cDNAライブラリーは、ランダムヘキサマーまたはランダムオクタマーによってプライミングされる。

【0121】

追加的な実施形態では、前記cDNAライブラリーは、約250bp～約750bpのサイズに断片化される。

【0122】

さらなる実施形態では、前記cDNAライブラリーは、約500bpのサイズに断片化される。

【0123】

一部の実施形態では、前記多機能性アダプターモジュールは、ランダム核酸タグ配列を含む第1の領域と、必要に応じて、試料コード配列を含む第2の領域と、必要に応じて、PCRプライマー配列を含む第3の領域とを含む。

30

【0124】

関連の実施形態では、前記多機能性アダプターモジュールは、ランダム核酸タグ配列を含む第1の領域と、試料コード配列を含む第2の領域と、PCRプライマー配列を含む第3の領域とを含む。

【0125】

種々の実施形態では、本明細書中で企図される方法は、タグ付けされたcDNAライブラリーを多機能性捕捉プローブモジュールとハイブリダイズさせて、複合体を形成するステップであって、前記多機能性捕捉プローブモジュールが、前記cDNAライブラリーにおける特異的標的領域とハイブリダイズするステップを含む。

40

【0126】

一部の実施形態では、本明細書中で企図される方法は、前記タグ付けされたcDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップを含む。

【0127】

特定の実施形態では、本明細書中で企図される方法は、前記単離されたタグ付けされたcDNAライブラリー-多機能性捕捉プローブモジュール複合体を3'-5'エキソヌクレアーゼ酵素プロセッシングして、一本鎖3'末端を除去するステップを含む。

【0128】

50

一部の実施形態では、前記 3' - 5' エキソヌクレアーゼ酵素プロセッシングにおいて使用される酵素は、T4 DNAポリメラーゼである。

【0129】

ある特定の実施形態では、本明細書中で企図される方法は、前記 3' - 5' エキソヌクレアーゼ酵素によりプロセッシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分がコピーされ、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記cDNA標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップを含む。

【0130】

さらなる実施形態では、標的化遺伝子発現解析のための方法が提供され、この方法は、

(a) タグ付けされたRNA発現ライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュールが、前記タグ付けされたRNA発現ライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたRNA発現ライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされたRNA発現ライブラリー-多機能性捕捉プローブモジュール複合体に対して3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖3' 末端を除去するステップと、

(d) c) から得られる前記酵素によりプロセッシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分がコピーされ、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

(e) d) から得られる前記ハイブリッド核酸分子に対して標的化遺伝子発現解析を実行するステップとを含む。

【0131】

追加的な実施形態では、標的化遺伝子発現解析のための方法が提供され、この方法は、

(a) タグ付けされたRNA発現ライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記RNA発現ライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたRNA発現ライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、b) から得られる前記複合体における前記多機能性捕捉プローブの5' から3' へのDNAポリメラーゼ伸長を実行して、前記多機能性捕捉プローブの3' にある前記捕捉されたタグ付けされた標的領域の一領域を複製するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記標的領域とハイブリダイズする位置の3' 方向に位置する前記タグ付けされた標的領域の相補体とを含むステップと、

(d) c) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む。

【0132】

種々の実施形態では、標的化遺伝子発現解析のための方法が提供され、この方法は、(

a) タグ付けされたcDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール

10

20

30

40

50

ル複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記 cDNA ライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされた cDNA ライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、b) から得られる前記複合体における前記多機能性捕捉プローブの 5' から 3' への DNA ポリメラーゼ伸長を実行して、前記多機能性捕捉プローブの 3' にある前記 cDNA ライブラリーにおける前記捕捉されたタグ付けされた標的領域の一領域を複製するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記標的領域とハイブリダイズする位置の 3' 方向に位置する前記 cDNA ライブラリーにおける前記タグ付けされた標的領域の相補体とを含むステップと、

10

(d) c) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む。

【0133】

特定の実施形態では、少なくとも 2 種の異なる多機能性捕捉プローブモジュールは、前記少なくとも 2 回の a) ステップにおいて使用され、前記少なくとも 2 回の a) ステップは、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる。

20

【0134】

ある特定の実施形態では、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記標的領域の上流とハイブリダイズする。

【0135】

追加的な実施形態では、cDNA 配列解析のための方法が提供され、この方法は、(a) 各クローンが第 1 の cDNA 配列および第 2 の cDNA 配列を含む 1 種または複数のクローンを得るステップであって、前記第 1 の cDNA 配列が、標的化ゲノム cDNA 配列を含み、前記第 2 の cDNA 配列が、捕捉プローブ配列を含むステップと、

(b) 前記 1 種または複数のクローンに対してペアエンドシーケンシング反応を実行して 1 種または複数のシーケンシングリードを得るステップと、

30

(c) 前記シーケンシングリードの前記プローブ配列に従って、前記 1 種または複数のクローンの前記シーケンシングリードを順序付けまたはクラスター形成するステップとを含む。

【0136】

種々の実施形態では、cDNA 配列解析のための方法が提供され、この方法は、(a) 各クローンが第 1 の cDNA 配列および第 2 の cDNA 配列を含む 1 種または複数のクローンを得るステップであって、前記第 1 の cDNA 配列が、標的化ゲノム DNA 配列を含み、前記第 2 の cDNA 配列が、捕捉プローブ配列を含むステップと、

40

(b) 前記 1 種または複数のクローンに対してシーケンシング反応を実行するステップであって、約 100 ヌクレオチドを超える単一の長いシーケンシングリードが得られ、前記リードが、前記第 1 の cDNA 配列および前記第 2 の cDNA 配列の両方の同定に十分であるステップと、

(c) 前記シーケンシングリードの前記プローブ配列に従って、前記 1 種または複数のクローンのシーケンシングリードを順序付けまたはクラスター形成するステップとを含む。

【0137】

特定の実施形態では、本明細書中で企図される方法は、独特のシーケンシングリードおよび重複したシーケンシングリードの分布を決定し、独特のリードが遭遇する回数を

50

計数し、前記独特のリードの度数分布を統計分布に適合させ、独特のリードの総数を推量し、各cDNAライブラリー試料内で収集される前記総リードに対する正規化を使用して、独特のリード計数を転写物存在量に変換することにより、第2のリード配列に関連するあらゆるシーケンシングリードを解析するステップを含む。

【0138】

ある特定の実施形態では、標的化遺伝学的解析のための方法が提供され、この方法は、
(a) タグ付けされたDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、

10

(b) a) から得られる前記タグ付けされたDNAライブラリー多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、5'FLAPエンドヌクレアーゼ活性、5'から3'へのDNAポリメラーゼ伸長およびDNAリガーゼによるニック閉鎖を含む、b) から得られる前記タグ付けされたDNAライブラリー多機能性捕捉プローブハイブリッドモジュール複合体の協奏的酵素プロセッシングを実行して、前記多機能性捕捉プローブ結合部位の5'にある前記標的領域に前記多機能性捕捉プローブの相補体を連結するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールの相補体と、前記多機能性捕捉プローブハイブリッドモジュールが前記ゲノム標的領域とハイブリダイズする位置の5'に位置する前記タグ付けされた標的領域の一領域を含むステップと、

20

(d) c) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む。

【0139】

種々の実施形態では、ステップa)～c)は、少なくとも約2回反復され、d)の前記標的化遺伝学的解析は、前記少なくとも2回のd)ステップから得られる前記ハイブリッド核酸分子配列の配列アラインメントを含む。

【0140】

ある特定の実施形態では、少なくとも2種の異なる多機能性捕捉プローブモジュールは、前記少なくとも2回のa)ステップにおいて使用され、前記少なくとも2回のa)ステップは、それぞれ1種の多機能性捕捉プローブモジュールを用いる。

30

【0141】

特定の実施形態では、少なくとも1種の多機能性捕捉プローブモジュールは、前記標的領域の下流とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールは、前記標的領域の上流とハイブリダイズする。

【0142】

追加的な実施形態では、特異的標的領域のコピー数を決定するための方法が提供され、この方法は、

(a) タグ付けされたDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記ゲノムライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、

40

(b) a) から得られる前記タグ付けされたDNAライブラリー多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、5'FLAPエンドヌクレアーゼ活性、5'から3'へのDNAポリメラーゼ伸長およびDNAリガーゼによるニック閉鎖を含む、b) から得られる前記タグ付けされたDNAライブラリー多機能性捕捉プローブハイブリッドモジュール複合体の協奏的酵素プロセッシングを実行して、前記多機能性捕捉プローブ結合部位の5'にある前記標的領域に前記多機能性捕捉プローブの相補体を連結す

50

るステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールの相補体と、前記多機能性捕捉プローブハイブリッドモジュールが前記標的領域とハイブリダイズする位置の 5' に位置する前記タグ付けされた標的領域の一領域を含むステップと、

(d) c) における前記ハイブリッド核酸分子の PCR 増幅を実行するステップと、

(e) d) における PCR 反応を定量化するステップであって、前記定量化が、前記特異的標的領域のコピー数の決定を可能にするステップとを含む。

【0143】

種々の実施形態では、本明細書中で企図される方法は、ステップ d) から得られる前記ハイブリッド核酸分子の配列を得るステップを含む。

【0144】

特定の実施形態では、ステップ a) ~ d) は、少なくとも約 2 回反復され、少なくとも 2 回の d) ステップから得られる前記ハイブリッド核酸分子の配列アラインメントが行われる。

【0145】

特定の実施形態では、少なくとも 2 種の異なる多機能性捕捉プローブモジュールは、少なくとも 2 回の a) ステップにおいて使用され、前記少なくとも 2 回の a) ステップは、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる。

【0146】

ある特定の実施形態では、少なくとも 1 種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールは、前記ゲノム標的領域の上流とハイブリダイズする。

【0147】

追加的な実施形態では、前記標的化遺伝学的解析は、配列解析である。

【0148】

さらなる実施形態では、前記標的領域は、ゲノム標的領域であり、前記 DNA ライブラリーは、ゲノム DNA ライブラリーである。

【0149】

一部の実施形態では、前記標的領域は、cDNA 標的領域であり、前記 DNA ライブラリーは、cDNA ライブラリーである。

特定の実施形態では、例えば以下が提供される：

(項目 1)

タグ付けされたゲノムライブラリーを作製するための方法であって、

(a) 断片化されたゲノム DNA を末端修復酵素で処理して、断片化され末端修復されたゲノム DNA を作製するステップと、

(b) ランダム核酸タグ配列ならびに必要なに応じて試料コード配列および／または PCR プライマー配列を、前記断片化され末端修復されたゲノム DNA にライゲーションして、前記タグ付けされたゲノムライブラリーを作製するステップと

を含む方法。

(項目 2)

前記ランダム核酸タグ配列が、約 2 ~ 約 100 ヌクレオチドである、先行する項目のいずれかに記載の方法。

(項目 3)

前記ランダム核酸タグ配列が、約 2 ~ 約 6 ヌクレオチドである、先行する項目のいずれかに記載の方法。

(項目 4)

前記断片化され末端修復されたゲノム DNA が、平滑末端を含有する、先行する項目のいずれかに記載の方法。

(項目 5)

10

20

30

40

50

前記平滑末端が、単一塩基対オーバーハングを含有するようさらに修飾される、先行する項目のいずれかに記載の方法。

(項目 6)

前記ライゲーションステップが、多機能性アダプターモジュールを前記断片化され末端修復されたゲノム DNA にライゲーションして、前記タグ付けされたゲノムライブラリーを作製することを含み、前記多機能性アダプター分子が、

(i) ランダム核酸タグ配列を含む第 1 の領域と、

(i i) 試料コード配列を含む第 2 の領域と、

(i i i) P C R プライマー配列を含む第 3 の領域とを含む、先行する項目のいずれかに記載の方法。

10

(項目 7)

タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュールとハイブリダイズさせて、複合体を形成するステップをさらに含み、前記多機能性捕捉プローブモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域とハイブリダイズする、先行する項目のいずれかに記載の方法。

(項目 8)

前記タグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップをさらに含み、先行する項目のいずれかに記載の方法。

(項目 9)

前記単離されたタグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体を 3' - 5' エクソヌクレアーゼ酵素プロセッシングして、一本鎖 3' 末端を除去するステップをさらに含み、先行する項目のいずれかに記載の方法。

20

(項目 10)

前記 3' - 5' エクソヌクレアーゼ酵素プロセッシングにおいて使用される酵素が、T 4 DNA ポリメラーゼである、先行する項目のいずれかに記載の方法。

(項目 11)

先行する項目から得られる前記 3' - 5' エクソヌクレアーゼ酵素によりプロセッシングされた複合体に対して P C R を実行するステップをさらに含み、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分がコピーされ、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含む、先行する項目のいずれかに記載の方法。

30

(項目 12)

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エクソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体に対して 3' - 5' エクソヌクレアーゼ酵素プロセッシングを実行して、一本鎖 3' 末端を除去するステップと、

40

(d) c) から得られる前記酵素によりプロセッシングされた複合体に対して P C R を実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分がコピーされ、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

(e) d) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップと

50

を含む、標的化遺伝学的解析のための方法。

(項目 1 3)

ステップ a) ~ d) が、少なくとも約 2 回反復され、e) の前記標的化遺伝学的解析が、前記少なくとも 2 回の d) ステップから得られるハイブリッド核酸分子配列の配列アラインメントを含む、項目 1 2 に記載の方法。

(項目 1 4)

少なくとも 2 種の異なる多機能性捕捉プローブモジュールが、少なくとも 2 回の a) ステップにおいて使用され、前記少なくとも 2 回の a) ステップが、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる、項目 1 3 に記載の方法。

(項目 1 5)

少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする、項目 1 4 に記載の方法。

(項目 1 6)

特異的ゲノム標的領域のコピー数を決定するための方法であって、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュール複合体が、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされたゲノムライブラリー-多機能性捕捉プローブモジュール複合体に対して 3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖 3' 末端を除去するステップと、

(d) c) から得られる前記酵素によりプロセッシングされた複合体に対して PCR 反応を実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分が複製され、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

(e) d) における前記ハイブリッド核酸分子の PCR 増幅を実行するステップと、

(f) e) における PCR 反応を定量化するステップであって、前記定量化が、前記特異的ゲノム標的領域のコピー数の決定を可能にするステップとを含む方法。

(項目 1 7)

ステップ e) から得られる前記ハイブリッド核酸分子の配列を得るステップをさらに含む、項目 1 6 に記載の方法。

(項目 1 8)

ステップ a) ~ e) が、少なくとも約 2 回反復され、前記少なくとも 2 回の e) ステップから得られる前記ハイブリッド核酸分子配列を使用して、配列アラインメントが実行される、項目 1 7 に記載の方法。

(項目 1 9)

少なくとも 2 種の異なる多機能性捕捉プローブモジュールが、前記少なくとも 2 回の a) ステップにおいて使用され、前記少なくとも 2 回の a) ステップが、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる、項目 1 8 に記載の方法。

(項目 2 0)

少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする、項目 1 9 に記載の方法。

(項目 2 1)

10

20

30

40

50

特異的ゲノム標的領域のコピー数を決定するための方法であって、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュール複合体が、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされたゲノムライブラリー多機能性捕捉プローブモジュール複合体に対して 3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖 3' 末端を除去するステップと、

(d) c) から得られる前記酵素によりプロセッシングされた複合体に対して P C R 反応を実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分が複製され、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

(e) d) における前記ハイブリッド核酸分子の P C R 増幅を実行するステップと、を含む方法。

(項目 2 2)

ステップ e) から得られる前記ハイブリッド核酸分子の配列を得るステップをさらに含む、項目 2 1 に記載の方法。

(項目 2 3)

ステップ a) ~ e) が、少なくとも約 2 回反復され、前記少なくとも 2 回の e) ステップから得られる前記ハイブリッド核酸分子配列を使用して、配列アラインメントが実行される、項目 2 2 に記載の方法。

(項目 2 4)

少なくとも 2 種の異なる多機能性捕捉プローブモジュールが、前記少なくとも 2 回の a) ステップにおいて使用され、前記少なくとも 2 回の a) ステップが、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる、項目 2 3 に記載の方法。

(項目 2 5)

少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする、項目 2 4 に記載の方法。

(項目 2 6)

特異的ゲノム標的領域のコピー数を決定するための方法であって、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュール複合体が、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされたゲノムライブラリー多機能性捕捉プローブモジュール複合体に対して 3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖 3' 末端を除去するステップと、

(d) c) から得られる前記酵素によりプロセッシングされた複合体に対して P C R 反応を実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分が複製され、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記ゲノム標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

10

20

30

40

50

(e) d) における前記ハイブリッド核酸分子のPCR増幅を実行するステップと、
(f) e) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行
するステップと
を含む方法。

(項目27)

ステップa)～e) が、少なくとも約2回反復され、f) の前記標的化遺伝学的解析が
、前記少なくとも2回のe) ステップから得られる前記ハイブリッド核酸分子配列の配列
アラインメントの実行を含む、項目26に記載の方法。

(項目28)

少なくとも2種の異なる多機能性捕捉プローブモジュールが、前記少なくとも2回のa
) ステップにおいて使用され、前記少なくとも2回のa) ステップが、それぞれ1種の多
機能性捕捉プローブモジュールを用いる、項目27に記載の方法。

(項目29)

少なくとも1種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイ
ブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールが、前記ゲノム標的
領域の上流とハイブリダイズする、項目28に記載の方法。

(項目30)

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブハイブリッドモジュ
ール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブ
リッドモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的に
ハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー多機能性捕捉プロー
ブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、b) から得られる前記複合体に対して
前記多機能性捕捉プローブの5'から3'へのDNAポリメラーゼ伸長を実行して、前記
多機能性捕捉プローブの3'にある前記捕捉されたタグ付けされたゲノム標的領域の一領
域を複製するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プロー
ブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記
ゲノム標的領域とハイブリダイズする位置から3'方向に位置する前記タグ付けされたゲ
ノム標的領域の一領域の相補体とを含むステップと、

(d) c) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行
するステップと

を含む、標的化遺伝学的解析のための方法。

(項目31)

ステップa)～c) が、少なくとも約2回反復され、d) の前記標的化遺伝学的解析が
、前記少なくとも2回のd) ステップから得られる前記ハイブリッド核酸分子配列の配列
アラインメントを含む、項目30に記載の方法。

(項目32)

少なくとも2種の異なる多機能性捕捉プローブモジュールが、前記少なくとも2回のa
) ステップにおいて使用され、前記少なくとも2回のa) ステップが、それぞれ1種の多
機能性捕捉プローブモジュールを用いる、項目31に記載の方法。

(項目33)

少なくとも1種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイ
ブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールが、前記ゲノム標的
領域の上流とハイブリダイズする、項目32に記載の方法。

(項目34)

特異的ゲノム標的領域のコピー数を決定するための方法であって、

(a) タグ付けされたゲノムライブラリーを多機能性捕捉プローブハイブリッドモジュ
ール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブ
リッドモジュールが、前記ゲノムライブラリーにおける特異的ゲノム標的領域と選択的に

10

20

30

40

50

ハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたゲノムライブラリー—多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、b) から得られる前記複合体に対して前記多機能性捕捉プローブの 5' から 3' への DNA ポリメラーゼ伸長を実行して、前記多機能性捕捉プローブの 3' にある前記捕捉されたタグ付けされたゲノム標的領域の一領域を複製するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記ゲノム標的領域とハイブリダイズする位置から 3' 方向に位置する前記タグ付けされたゲノム標的領域の一領域の相補体とを含むステップと、

10

(d) c) における前記ハイブリッド核酸分子の PCR 増幅を実行するステップと、

(e) d) における PCR 反応を定量化するステップであって、前記定量化が、前記特異的ゲノム標的領域のコピー数の決定を可能にするステップとを含む方法。

(項目 35)

ステップ d) から得られる前記ハイブリッド核酸分子の配列を得るステップをさらに含む、項目 34 に記載の方法。

(項目 36)

ステップ a) ~ d) が、少なくとも約 2 回反復され、前記少なくとも 2 回の d) ステップから得られる前記ハイブリッド核酸分子の配列アラインメントが行われる、項目 35 に記載の方法。

20

(項目 37)

少なくとも 2 種の異なる多機能性捕捉プローブモジュールが、前記少なくとも 2 回の a) ステップにおいて使用され、前記少なくとも 2 回の a) ステップが、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる、項目 36 に記載の方法。

(項目 38)

少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする、項目 37 に記載の方法。

(項目 39)

前記標的化遺伝学的解析が、配列解析である、先行する項目のいずれかに記載の方法。

30

(項目 40)

前記タグ付けされたゲノムライブラリーが PCR により増幅されて、増幅されたタグ付けされたゲノムライブラリーを作製する、先行する項目のいずれかに記載の方法。

(項目 41)

前記ゲノム DNA が、血液、皮膚、毛、毛包、唾液、口腔粘膜、膣粘膜、汗、涙、上皮組織、尿、精液、精子液、精漿、前立腺液、尿道球腺液（カウパー氏腺液）、排泄物、生検、腹水、脳脊髄液、リンパ液および組織抽出物試料または生検試料からなる群から選択される生体試料に由来する、先行する項目のいずれかに記載の方法。

40

(項目 42)

タグ付けされたゲノム配列を含むタグ付けされたゲノムライブラリーであって、各タグ付けされたゲノム配列が、

(a) 断片化され末端修復されたゲノム DNA と、

(b) ランダムヌクレオチドタグ配列と、

(c) 試料コード配列と、

(d) PCR プライマー配列と

を含む、ゲノムライブラリー。

(項目 43)

標的化遺伝学的解析において使用するためのハイブリッドタグ付けされたゲノム配列を

50

含むハイブリッドタグ付けされたゲノムライブラリーであって、各ハイブリッドタグ付けされたゲノム配列が、

(a) 断片化され末端修復されたゲノムDNAと、

(b) ランダムヌクレオチドタグ配列と、

(c) 試料コード配列と、

(d) PCRプライマー配列と、

(e) ゲノム標的領域と、

(f) 多機能性捕捉プローブモジュールテイル配列とを含む、ゲノムライブラリー。

(項目44)

(a) ランダムヌクレオチドタグ配列を含む第1の領域と、

(b) 試料コード配列を含む第2の領域と、

(c) PCRプライマー配列を含む第3の領域とを含む多機能性アダプターモジュール。

(項目45)

(a) パートナーオリゴヌクレオチドとハイブリダイズすることができる第1の領域と、

(b) 特異的ゲノム標的領域とハイブリダイズすることができる第2の領域と、

(c) テイル配列を含む第3の領域とを含む多機能性捕捉プローブモジュール。

(項目46)

前記第1の領域が、パートナーオリゴヌクレオチドに結合している、先行する項目のいずれかに記載の多機能性捕捉プローブモジュール。

(項目47)

(a) パートナーオリゴヌクレオチドとハイブリダイズすることができ、PCRプライマーとして機能することができる第1の領域と、

(b) 特異的ゲノム標的領域とハイブリダイズすることができる第2の領域とを含む多機能性アダプタープローブハイブリッドモジュール。

(項目48)

前記第1の領域が、パートナーオリゴヌクレオチドに結合している、先行する項目のいずれかに記載の多機能性捕捉プローブハイブリッドモジュール。

(項目49)

前記パートナーオリゴヌクレオチドが、化学修飾されている、先行する項目のいずれかに記載の方法。

(項目50)

タグ付けされたゲノムライブラリーと、多機能性アダプターモジュールと、多機能性捕捉プローブモジュールとを含む組成物。

(項目51)

先行する項目のいずれかに記載のハイブリッドタグ付けされたゲノムライブラリーを含む組成物。

(項目52)

先行する項目のいずれか一項に記載の方法を実行するための反応混合物。

(項目53)

タグ付けされたゲノムライブラリーを作製することができる反応混合物であって、

(a) 断片化されたゲノムDNAと、

(b) 断片化され末端修復されたゲノムDNAを作製するためのDNA末端修復酵素とを含む反応混合物。

(項目54)

多機能性アダプターモジュールをさらに含む、先行する項目のいずれかに記載の反応混合物。

10

20

30

40

50

(項目 5 5)

多機能性捕捉プローブモジュールをさらに含む、先行する項目のいずれかに記載の反応混合物。

(項目 5 6)

3' - 5' エキソヌクレアーゼ活性および PCR 増幅活性を有する酵素をさらに含む、先行する項目のいずれかに記載の反応混合物。

(項目 5 7)

(a) 各クローンが第 1 の DNA 配列および第 2 の DNA 配列を含む 1 種または複数のクローンを得るステップであって、前記第 1 の DNA 配列が、標的化ゲノム DNA 配列を含み、前記第 2 の DNA 配列が、捕捉プローブ配列を含むステップと、

(b) 前記 1 種または複数のクローンに対してペアエンドシーケンシング反応を実行して 1 種または複数のシーケンシングリードを得るステップと、

(c) 前記シーケンシングリードの前記プローブ配列に従って、前記 1 種または複数のクローンの前記シーケンシングリードを順序付けまたはクラスター形成するステップとを含む、DNA 配列解析のための方法。

(項目 5 8)

(a) 各クローンが第 1 の DNA 配列および第 2 の DNA 配列を含む 1 種または複数のクローンを得るステップであって、前記第 1 の DNA 配列が、標的化ゲノム DNA 配列を含み、前記第 2 の DNA 配列が、捕捉プローブ配列を含むステップと、

(b) 前記 1 種または複数のクローンに対してシーケンシング反応を実行するステップであって、約 100 ヌクレオチドを超える単一の長いシーケンシングリードが得られ、前記リードが、前記第 1 の DNA 配列および前記第 2 の DNA 配列の両方の同定に十分であるステップと、

(c) 前記シーケンシングリードの前記プローブ配列に従って、前記 1 種または複数のクローンのシーケンシングリードを順序付けまたはクラスター形成するステップとを含む、DNA 配列解析のための方法。

(項目 5 9)

前記 1 種または複数のクローンの配列が、1 種または複数のヒト参照 DNA 配列と比較される、項目 5 7 または項目 5 8 に記載の方法。

(項目 6 0)

前記 1 種または複数のヒト参照 DNA 配列にマッチしない配列が同定される、項目 5 9 に記載の方法。

(項目 6 1)

非マッチ配列が使用されて、前記非マッチ配列データからデノボアセンブリーを作成する、項目 6 0 に記載の方法。

(項目 6 2)

前記デノボアセンブリーが使用されて、前記捕捉プローブに関連する新規配列再編成を同定する、項目 6 1 に記載の方法。

(項目 6 3)

(a) 各クローンが第 1 の DNA 配列および第 2 の DNA 配列を含む 1 種または複数のクローンを得るステップであって、前記第 1 の DNA 配列が、ランダムヌクレオチドタグ配列および標的化ゲノム DNA 配列を含み、前記第 2 の DNA 配列が、捕捉プローブ配列を含むステップと、

(b) 前記 1 種または複数のクローンに対してペアエンドシーケンシング反応を実行して 1 種または複数のシーケンシングリードを得るステップと、

(c) 前記シーケンシングリードの前記プローブ配列に従って、前記 1 種または複数のクローンの前記シーケンシングリードを順序付けまたはクラスター形成するステップとを含む、ゲノムコピー数決定解析のための方法。

10

20

30

40

50

(項目 6 4)

(a) 各クローンが第 1 の DNA 配列および第 2 の DNA 配列を含む 1 種または複数のクローンを得るステップであって、前記第 1 の DNA 配列が、ランダムヌクレオチドタグ配列および標的化ゲノム DNA 配列を含み、前記第 2 の DNA 配列が、捕捉プローブ配列を含むステップと、

(b) 前記 1 種または複数のクローンに対してシーケンシング反応を実行するステップであって、約 100 ヌクレオチドを超える単一の長いシーケンシングリードが得られ、前記リードが、前記第 1 の DNA 配列および前記第 2 の DNA 配列の両方の同定に十分であるステップと、

(c) 前記シーケンシングリードの前記プローブ配列に従って、前記 1 種または複数のクローンのシーケンシングリードを順序付けまたはクラスター形成するステップとを含む、ゲノムコピー数決定解析のための方法。

10

(項目 6 5)

前記ランダムヌクレオチドタグ配列が、約 2 ～約 50 ヌクレオチドの長さである、項目 6 3 または項目 6 4 に記載の方法。

(項目 6 6)

(a) 独特のシーケンシングリードおよび重複したシーケンシングリードの分布を決定し、

(b) 独特のリードが遭遇する回数を計数し、

(c) 前記独特のリードの度数分布を統計分布に適合させ、

20

(d) 独特のリードの総数を推量し、

(e) 推量された独特のリードの前記総数を、大部分のヒト遺伝子座が一般に二倍体であるという仮定に対して正規化すること

により、第 2 のリード配列に関連するあらゆるシーケンシングリードを解析するステップをさらに含む、項目 6 3 または項目 6 4 に記載の方法。

(項目 6 7)

1 種または複数の標的化遺伝子座の推量されたコピー数が決定される、項目 6 6 に記載の方法。

(項目 6 8)

予想されるコピー数値から逸脱する前記 1 種または複数の標的遺伝子座が決定される、項目 6 7 に記載の方法。

30

(項目 6 9)

遺伝子の前記 1 種または複数の標的化遺伝子座が、遺伝子座のコレクションにおいて共にグループ化され、標的化遺伝子座のコレクションから得られる前記コピー数測定値が、平均化および正規化される、項目 6 7 に記載の方法。

(項目 7 0)

遺伝子の前記推量されたコピー数が、この遺伝子を表す全標的遺伝子座の前記正規化された平均によって表される、項目 6 7 に記載の方法。

(項目 7 1)

タグ付けされた RNA 発現ライブラリーを作製するための方法であって、

40

(a) cDNA ライブラリーを断片化するステップと、

(b) 前記断片化された cDNA ライブラリーを末端修復酵素で処理して、断片化され末端修復された cDNA を作製するステップと、

(c) 多機能性アダプター分子を前記断片化され末端修復された cDNA にライゲーションして、タグ付けされた RNA 発現ライブラリーを作製するステップとを含む方法。

(項目 7 2)

タグ付けされた RNA 発現ライブラリーを作製するための方法であって、

(a) 1 個または複数の細胞の全 RNA から cDNA ライブラリーを調製するステップと、

50

(b) 前記 c D N A ライブラリーを断片化するステップと、
(c) 前記断片化された c D N A を末端修復酵素で処理して、断片化され末端修復された c D N A を作製するステップと、

(d) 多機能性アダプター分子を前記断片化され末端修復された c D N A にライゲーションして、タグ付けされた R N A 発現ライブラリーを作製するステップとを含む方法。

(項目 7 3)

前記 c D N A ライブラリーが、オリゴ d T プライミング c D N A ライブラリーである、項目 7 1 または項目 7 2 に記載の方法。

(項目 7 4)

前記 c D N A ライブラリーが、約 6 ～ 約 2 0 ランダムヌクレオチドを含むランダムオリゴヌクレオチドによってプライミングされる、項目 7 1 または項目 7 2 に記載の方法。

(項目 7 5)

前記 c D N A ライブラリーが、ランダムヘキサマーまたはランダムオクタマーによってプライミングされる、項目 7 1 または項目 7 2 に記載の方法。

(項目 7 6)

前記 c D N A ライブラリーが、約 2 5 0 b p ～ 約 7 5 0 b p のサイズに断片化される、項目 7 1 または項目 7 2 に記載の方法。

(項目 7 7)

前記 c D N A ライブラリーが、約 5 0 0 b p のサイズに断片化される、項目 7 1 または項目 7 2 に記載の方法。

(項目 7 8)

前記多機能性アダプターモジュールが、

(i) ランダム核酸タグ配列を含む第 1 の領域と、必要に応じて、

(i i) 試料コード配列を含む第 2 の領域と、必要に応じて、

(i i i) P C R プライマー配列を含む第 3 の領域とを含む、項目 7 1 ～ 7 7 のいずれか一項に記載の方法。

(項目 7 9)

前記多機能性アダプターモジュールが、ランダム核酸タグ配列を含む第 1 の領域と、試料コード配列を含む第 2 の領域と、P C R プライマー配列を含む第 3 の領域とを含む、請求項 7 1 ～ 7 8 のいずれか一項に記載の方法。

(項目 8 0)

タグ付けされた c D N A ライブラリーを多機能性捕捉プローブモジュールとハイブリダイズさせて、複合体を形成するステップであって、前記多機能性捕捉プローブモジュールが、前記 c D N A ライブラリーにおける特異的標的領域とハイブリダイズするステップをさらに含む、項目 7 1 ～ 7 8 のいずれか一項に記載の方法。

(項目 8 1)

前記タグ付けされた c D N A ライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップをさらに含む、項目 7 1 ～ 7 8 のいずれか一項に記載の方法。

(項目 8 2)

前記単離されたタグ付けされた c D N A ライブラリー-多機能性捕捉プローブモジュール複合体を 3' - 5' エクソヌクレアーゼ酵素プロセッシングして、一本鎖 3' 末端を除去するステップをさらに含む、項目 7 1 ～ 7 8 のいずれか一項に記載の方法。

(項目 8 3)

前記 3' - 5' エクソヌクレアーゼ酵素プロセッシングにおいて使用される酵素が、T 4 D N A ポリメラーゼである、項目 8 2 に記載の方法。

(項目 8 4)

前記 3' - 5' エクソヌクレアーゼ酵素によりプロセッシングされた複合体に対して P C R を実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分がコピーされ、前記ハイブリッド核酸分子が、前記多機能

10

20

30

40

50

性捕捉プローブモジュールとハイブリダイズすることができる前記 cDNA 標的領域と、
前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップをさらに含む、
項目 82 または項目 83 に記載の方法。

(項目 85)

(a) タグ付けされた RNA 発現ライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブモジュールが、
前記タグ付けされた RNA 発現ライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされた RNA 発現ライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、

(c) 3' - 5' エクソヌクレアーゼ活性を有する酵素を使用して、b) から得られる前記単離されたタグ付けされた RNA 発現ライブラリー-多機能性捕捉プローブモジュール複合体に対して 3' - 5' エクソヌクレアーゼ酵素プロセッシングを実行して、一本鎖 3' 末端を除去するステップと、

(d) c) から得られる前記酵素によりプロセッシングされた複合体に対して PCR を実行するステップであって、ハイブリッド核酸分子を作製するために、前記多機能性捕捉プローブ分子のテイル部分がコピーされ、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブモジュールとハイブリダイズすることができる前記標的領域と、前記多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、

(e) d) から得られる前記ハイブリッド核酸分子に対して標的化遺伝子発現解析を実行するステップと
を含む、標的化遺伝子発現解析のための方法。

(項目 86)

(a) タグ付けされた RNA 発現ライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記 RNA 発現ライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされた RNA 発現ライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、b) から得られる前記複合体における前記多機能性捕捉プローブの 5' から 3' への DNA ポリメラーゼ伸長を実行して、前記多機能性捕捉プローブの 3' にある前記捕捉されたタグ付けされた標的領域の一領域を複製するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記標的領域とハイブリダイズする位置の 3' 方向に位置する前記タグ付けされた標的領域の相補体とを含むステップと、

(d) c) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップと
を含む、標的化遺伝子発現解析のための方法。

(項目 87)

(a) タグ付けされた cDNA ライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記 cDNA ライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされた cDNA ライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、b) から得られる前記複合体における前記多機能性捕捉プローブの 5' から 3' への DNA ポリメラーゼ伸長を実行して、前記多機能性捕捉プローブの 3' にある前記 cDNA ライブラリーにおける前記捕捉されたタグ付けされた標的領域の一領域を複製するステップであって、前記ハイブリッド核酸分子

10

20

30

40

50

が、前記多機能性捕捉プローブハイブリッドモジュールと、前記多機能性捕捉プローブハイブリッドモジュールが前記標的領域とハイブリダイズする位置の3'方向に位置する前記cDNAライブラリーにおける前記タグ付けされた標的領域の相補体とを含むステップと、

(d) c) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップと

を含む、標的化遺伝子発現解析のための方法。

(項目88)

少なくとも2種の異なる多機能性捕捉プローブモジュールが、前記少なくとも2回のa) ステップにおいて使用され、前記少なくとも2回のa) ステップが、それぞれ1種の多機能性捕捉プローブモジュールを用いる、項目85～87のいずれか一項に記載の方法。

(項目89)

少なくとも1種の多機能性捕捉プローブモジュールが、前記標的領域の下流とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールが、前記標的領域の上流とハイブリダイズする、項目88に記載の方法。

(項目90)

(a) 各クローンが第1のcDNA配列および第2のcDNA配列を含む1種または複数のクローンを得るステップであって、前記第1のcDNA配列が、標的化ゲノムcDNA配列を含み、前記第2のcDNA配列が、捕捉プローブ配列を含むステップと、

(b) 前記1種または複数のクローンに対してペアエンドシーケンシング反応を実行して1種または複数のシーケンシングリードを得るステップと、

(c) 前記シーケンシングリードの前記プローブ配列に従って、前記1種または複数のクローンの前記シーケンシングリードを順序付けまたはクラスター形成するステップと

を含む、cDNA配列解析のための方法。

(項目91)

(a) 各クローンが第1のcDNA配列および第2のcDNA配列を含む1種または複数のクローンを得るステップであって、前記第1のcDNA配列が、標的化ゲノムDNA配列を含み、前記第2のcDNA配列が、捕捉プローブ配列を含むステップと、

(b) 前記1種または複数のクローンに対してシーケンシング反応を実行するステップであって、約100ヌクレオチドを超える単一の長いシーケンシングリードが得られ、前記リードが、前記第1のcDNA配列および前記第2のcDNA配列の両方の同定に十分であるステップと、

(c) 前記シーケンシングリードの前記プローブ配列に従って、前記1種または複数のクローンのシーケンシングリードを順序付けまたはクラスター形成するステップとを含む、cDNA配列解析のための方法。

(項目92)

(a) 独特のシーケンシングリードおよび重複したシーケンシングリードの分布を決定し、

(b) 独特のリードが遭遇する回数を計数し、

(c) 前記独特のリードの度数分布を統計分布に適合させ、

(d) 独特のリードの総数を推量し、

(e) 各cDNAライブラリー試料内で収集される前記総リードに対する正規化を使用して、独特のリード計数を転写物存在量に変換すること

により、第2のリード配列に関連するあらゆるシーケンシングリードを解析するステップをさらに含む、項目90または項目91に記載の方法。

(項目93)

(a) タグ付けされたDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記DNAライブラリーにおける特異的標的領域と選択的にハイブ

10

20

30

40

50

リダイズするステップと、

(b) a) から得られる前記タグ付けされたDNAライブラリー多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、5'FLAPエンドヌクレアーゼ活性、5'から3'へのDNAポリメラーゼ伸長およびDNAリガーゼによるニック閉鎖を含む、b) から得られる前記タグ付けされたDNAライブラリー多機能性捕捉プローブハイブリッドモジュール複合体の協奏的酵素プロセッシングを実行して、前記多機能性捕捉プローブ結合部位の5'にある前記標的領域に前記多機能性捕捉プローブの相補体を連結するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールの相補体と、前記多機能性捕捉プローブハイブリッドモジュールが前記ゲノム標的領域とハイブリダイズする位置の5'に位置する前記タグ付けされた標的領域の一領域を含むステップと、

(d) c) から得られる前記ハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップと

を含む、標的化遺伝学的解析のための方法。

(項目94)

ステップa)～c)が、少なくとも約2回反復され、d)の前記標的化遺伝学的解析が、前記少なくとも2回のd)ステップから得られる前記ハイブリッド核酸分子配列の配列アラインメントを含む、項目93に記載の方法。

(項目95)

少なくとも2種の異なる多機能性捕捉プローブモジュールが、前記少なくとも2回のa)ステップにおいて使用され、前記少なくとも2回のa)ステップが、それぞれ1種の多機能性捕捉プローブモジュールを用いる、項目94に記載の方法。

(項目96)

少なくとも1種の多機能性捕捉プローブモジュールが、前記標的領域の下流とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールが、前記標的領域の上流とハイブリダイズする、項目95に記載の方法。

(項目97)

特異的標的領域のコピー数を決定するための方法であって、

(a) タグ付けされたDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、前記多機能性捕捉プローブハイブリッドモジュールが、前記ゲノムライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、

(b) a) から得られる前記タグ付けされたDNAライブラリー多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、

(c) ハイブリッド核酸分子を作製するために、5'FLAPエンドヌクレアーゼ活性、5'から3'へのDNAポリメラーゼ伸長およびDNAリガーゼによるニック閉鎖を含む、b) から得られる前記タグ付けされたDNAライブラリー多機能性捕捉プローブハイブリッドモジュール複合体の協奏的酵素プロセッシングを実行して、前記多機能性捕捉プローブ結合部位の5'にある前記標的領域に前記多機能性捕捉プローブの相補体を連結するステップであって、前記ハイブリッド核酸分子が、前記多機能性捕捉プローブハイブリッドモジュールの相補体と、前記多機能性捕捉プローブハイブリッドモジュールが前記標的領域とハイブリダイズする位置の5'に位置する前記タグ付けされた標的領域の一領域を含むステップと、

(d) c) における前記ハイブリッド核酸分子のPCR増幅を実行するステップと、

(e) d) におけるPCR反応を定量化するステップであって、前記定量化が、前記特異的標的領域のコピー数の決定を可能にするステップと
を含む方法。

(項目98)

ステップd) から得られる前記ハイブリッド核酸分子の配列を得るステップをさらに含

10

20

30

40

50

む、項目 97 に記載の方法。

(項目 99)

ステップ a) ~ d) が、少なくとも約 2 回反復され、少なくとも 2 回の d) ステップから得られる前記ハイブリッド核酸分子の配列アラインメントが行われる、項目 98 に記載の方法。

(項目 100)

少なくとも 2 種の異なる多機能性捕捉プローブモジュールが、少なくとも 2 回の a) ステップにおいて使用され、前記少なくとも 2 回の a) ステップが、それぞれ 1 種の多機能性捕捉プローブモジュールを用いる、項目 99 に記載の方法。

(項目 101)

少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の下流とハイブリダイズし、少なくとも 1 種の多機能性捕捉プローブモジュールが、前記ゲノム標的領域の上流とハイブリダイズする、項目 100 に記載の方法。

(項目 102)

前記標的化遺伝学的解析が、配列解析である、項目 93 ~ 101 のいずれか一項に記載の方法。

(項目 103)

前記標的領域が、ゲノム標的領域であり、前記 DNA ライブラリーが、ゲノム DNA ライブラリーである、項目 93 ~ 102 のいずれか一項に記載の方法。

(項目 104)

前記標的領域が、cDNA 標的領域であり、前記 DNA ライブラリーが、cDNA ライブラリーである、項目 93 ~ 103 のいずれか一項に記載の方法。

【図面の簡単な説明】

【0150】

【図 1】増幅可能で、試料コード処理され、タグ付けされた、ゲノム DNA ライブラリーの構築。精製ゲノム DNA は、全血液または口腔内頬スワブなどの供給源から単離した。DNA を断片化（例えば、機械的手段、酵素的手段、または化学的手段により）し、DNA の末端を、この例では、平滑末端に修復した。修復された DNA を、ユニバーサル増幅配列、ランダムヌクレオチドタグ配列、および試料コード配列を含有する、多機能性アダプターモジュールにライゲーションした。例を目的として、典型的なアダプター二重鎖分子の具体例を示す。

【図 2】ゲノム捕捉プローブのデザイン。（A）原型的な 114 ヌクレオチドのプローブの構造。領域 1 は、高度に修飾された、34 ヌクレオチドのパートナーオリゴヌクレオチドとの完全な相補性を共有する、34 ヌクレオチドの配列である。パートナーオリゴヌクレオチドを、その 5' 端において、ストレプトアビジンでコーティングされた磁気ビーズ上の捕捉を可能とする、ビオチン-TEG 化学実体で修飾した。B は、「Bio-TEG」修飾を表す。領域 2 は、ゲノム DNA 標的と相互作用するようにデザインされた、60 ヌクレオチドのプローブ領域である。領域 3 は、PCR 増幅配列を、捕捉されたゲノム断片に導入する、20 ヌクレオチドのテイルである。（B）各個別のプローブの領域 1 と相補的な、高度に修飾されたパートナー鎖の例。

【図 3】従来の、ハイブリダイゼーションベースの捕捉実験における、配列の「スプレッディング」。（A）フランキング断片は、プローブ断片-フランキング断片間相互作用を伴う、正規のハイブリダイゼーション三重鎖を介して、捕捉ライブラリーに「ヒッチハイク」しうる。（B）交差断片ハイブリダイゼーションの正味の結果は、標的領域（破線）の近傍内にあるが、所望の標的の境界から外れる配列を意味する、配列の「スプレッディング」である。

【図 4】断片：プローブハイブリダイズ複合体の酵素プロセッシング。（A）プロセッシング前の、断片（明るいグレー）とプローブ（黒色）との精製複合体。B：ビオチンアフィニティー修飾である。（B）DNA ポリメラーゼ（例えば、T4 DNA ポリメラーゼ）は、捕捉された断片の 3' セグメントを除去する、3' → 5' エキソヌクレアーゼ活性をコ

10

20

30

40

50

ードした。(C)プローブ：断片二重鎖領域に遭遇すると、ポリメラーゼは、プローブのテイルセグメントを、ハイブリダイズしたゲノム断片にコピーする。(D)多機能性捕捉プローブハイブリッドモジュールのテイルセグメントが、標的ゲノム断片にコピーされた、最終的な修飾断片。

【図5】捕捉複合体の酵素プロセッシングは、標的領域上のシーケンシングリードに「フォーカス」する。(A)捕捉プローブとの関連で示される、酵素プロセッシングされた断片。シーケンシングリードの配向性は、プローブにより指定され、薄い黒色の矢印により描示される。(B)「スプレディング」を最小化した、フォーカスされたリードの仮説的トレース。

【図6】有向プローブの「フォーカシング効果」。(A)典型的なエキソンの平均長は、100~150bpである。有向捕捉プローブは、標的セグメントを挟むイントロン領域内に配置される。(B)各個別のプローブ配列のリード分布を破線で示す。相加的なカバレッジを実線で示す。この例において示される通り、カバレッジの有向的性格を使用して、クエリー領域に対する鋭利なフォーカスを得ることができる。

【図7】アダプター二量体非含有断片ライブラリーが、「チューニング可能なオン/オフ」増幅特性を伴うことの実証。全く同じゲル画像を、4つの異なる色スキームおよびコントラストスキームで示す。試料は、(1)ACA2 20により増幅される、インサートを伴わない、アダプターだけのライゲーション；(2)ACA2 (通常の25ヌクレオチドのPCRプライマー)により増幅される、インサートを伴わない、アダプターだけのライゲーション；(3)ACA2 FLFP (全長フォワードプライマー：full length forward primer)により増幅される、インサートを伴わない、アダプターだけのライゲーション；(4)ACA2 20により増幅される、約200bpのhgDNAインサート+アダプターによるライゲーションミックス20ng；(5)ACA2 (通常の25ヌクレオチドのPCRプライマー)により増幅される、約200bpのhgDNAインサート+アダプターによるライゲーションミックス20ng；および(6)ACA2 FLFP (全長フォワードプライマー)により増幅される、約200bpのhgDNAインサート+アダプターによるライゲーションミックス20ngであった。アダプター単独のライゲーション→PCR産物では、増幅された材料は、目視可能ではなかった(レーン1~3)。短鎖の20ヌクレオチドのACA2プライマーは、「通常の」25ヌクレオチドのACA2プライマー(レーン5)と比べて非効率的な増幅(レーン4)を示した。58ヌクレオチドのACA2 FLFPプライマー(レーン6)では、材料のごくかすかな痕跡が目視可能であるに過ぎなかった。

【図8】CovarisによるgDNAの断片化の後における、平均断片サイズの均等な分布。男性(M)および女性(F)ヒトgDNA(Promega Corporation、Madison、WI、USAから受領した)を、Covaris条件によりシアリングし、2μl(約120ng)または5μl(約300ng)の断片化前(U)試料または断片化後(C)試料を、2%のアガロースゲル上にロードした。平均断片サイズは、200bp近傍を中心とする均等分布であった。

【図9】男性試料と女性試料との間の予測コピー数の差違を実証する、プロテオリピドタンパク質1(PLP1：proteolipid protein 1)qPCRアッセイ6についての増幅トレース。男性ゲノムDNA鋳型または女性ゲノムDNA鋳型を、PLP1 qPCRアッセイ6を使用する、Illumina Eco測定器上のリアルタイムPCRにより3連で増幅した。増幅トレースは、女性試料と男性試料との間のコピー数の差違を明確に実証した。

【図10】PLP1 qPCRアッセイプライマーによる従来のPCRの後において予測される、アンプリコンのサイズおよび独特のものであること(uniqueness)の実証。男性ゲノムDNA鋳型または女性ゲノムDNA鋳型を、qPCRアッセイプライマーセット1~8を使用する従来のPCRにより増幅し(実施例3)、精製されていないPCR反応物を、2%のアガロースゲル上に直接ロードした。各二重項の上側バンドは、アッセイPCR産物の推定移動度と符合した。下側の「ファジー」材料は、使用されないPCRプライ

10

20

30

40

50

マーであった可能性が高い。

【図11】2倍濃度のABI SYBRミックスおよび複数の条件を使用する、PLP1 qPCRアッセイの性能についての解析。ゲノムライブラリーI（実施例4で構築される）から得られるDNA断片を、鋳型として使用して、2倍濃度のABI SYBRマスターミックスを2ステップのPCR反応で使用して、室温で準備された場合のPLP1 qPCRアッセイの性能を測定した。非鋳型対照トレース（A）および+gDNAトレース（B）は、アッセイの性能についての定性的描像を提示することが示される。

【図12】T4-DNAポリメラーゼによる捕捉後プロセシングの後におけるインサートサイズの低減。アダプターをライゲーションされた、ゲノムライブラリーI（実施例4で構築される）から得られるgDNA断片の試料4例を、記載される通りに（実施例6：PLP1エクソン2）捕捉した。これらの試料のうちの2例では、ユニバーサル結合性オリゴC1を活用したが、他の2例の試料は、オリゴC10と結合させた。次いで、試料を、T4-DNAポリメラーゼで処理する（T4プロセシング）か、またはT4ポリメラーゼを欠く反応溶液中で同様に加工した（非処理）。T4ポリメラーゼによる捕捉後プロセシングは、試料のサイズ分布の全体的な低減を誘導したことから、インサートの平均サイズの低減が示唆される。加えて、T4プロセシングは、2つの弱いバンド（約250bpおよび約175bp）の出現を結果としてもたらした。

【図13】捕捉後プロセシング感度の直接的な測定。まず、PLP1エクソン2特異的ゲノムDNA断片を、単一のPLP1捕捉プローブを独立の反応で使用する女性gDNAライブラリー（実施例1）からのプルダウン／プルアウトにより単離した。捕捉された材料は、（A）で例示される通り、隣接するPLP1 qPCRアッセイプライマー対を使用して定量化した。酵素プロセシングの後、プロセシングされた複合体の量を、（B）で示される通り、1種のPLP1特異的プライマーおよび1種のプローブ特異的プライマーを使用するqPCRにより再度測定した。[B/A×100%]における測定値の比により、プロセシング効率の推定値がもたらされるであろう。リアルタイム反応から得られるPCR産物を抽出し、ゲル解析にかけて、予測される長さのアンプリコンが作製されたことを検証した（C）。これは、いずれのPCR反応も個別の開始点および停止点を有したので可能であった。プロセシング効率は、A+B+Cから解釈可能なデータをもたらしプルアウトから推量した。

【図14】プロセシング前およびプロセシング後における、PLP1エクソン2で捕捉されたDNA断片の、リアルタイムの定量化から得られるqPCR産物についてのゲル解析。6種の独立の捕捉反応物（プローブ1による2種の反応物、プローブ4による2種の反応物、プローブ2による1種の反応物、およびプローブ3による1種の反応物）を、図16において記載される通りにプロセシングした。プローブは、B10ユニバーサルオリゴセット（実施例4）に由来し、ユニバーサルオリゴおよびUltramerプローブを含んだ。これらの条件下で、アッセイセット3（プローブ4）、5（プローブ2）、および6（プローブ3）は、アッセイのアンプリコン（上のゲル）またはアダプターアンプリコンに照らしたプロセシング後におけるPLP1（下のゲル）と符合するPCR産物をもたらししたが、他のアッセイセットでは、検出可能な産物が観察されなかった。

【図15】代替的な断片：プローブハイブリダイズ複合体の酵素プロセシング。図4で概括した方法とは対照的に、この代替法は、クローンにプローブをコピーさせる手法から、プローブにクローンをコピーさせる手法にシフトした。この極性の逆転は、プローブの5'端を、プルダウン配列およびリバースPCR配列のいずれとしても使用したことを意味する。プローブの3'端は、非修飾のまま放置し、これにより、DNAポリメラーゼを使用してクローンをコピーすることが可能となり、タグ付けされた単離DNAライブラリー断片を鋳型として活用して、5'-3'DNAポリメラーゼにより、多機能性捕捉プローブを伸長させた。

【図16】代替的な酵素プロセシングの概念について調べるために援用された実験デザイン。挙動が良好な4種のqPCRアッセイ（10、14、15、および16）は、これらのアッセイを「指さす」プローブとマッチした。プローブの標的配列と、qPCRアッセ

10

20

30

40

50

イとは、互いの近傍内にある領域を指向したが、重複しなかったことは重要である。したがって、これらの十分に確立されたアッセイセットを使用すれば、プロセシングの効果を直接調べることができよう。

【図17】代替のプロセシング法により誘導される、ライブラリーの平均インサートサイズの減殺。DNA断片を、捕捉プローブ/Ultramerとハイブリダイズさせ、前出で記載した(図12)通りに、ストレプトアビジンビーズ上で捕捉した。捕捉後プロセシングは、図17において記載した代替法およびを使用して実行し、投入試料ならびにプロセシング前における試料およびプロセシング後における試料を、2%のアガロースゲル電気泳動により解析した。予測される通り、プロセシング試料中で、ライブラリーの平均インサートサイズは減殺され、これにより、プロセシングが働くという結論が裏付けられた。また、ライブラリーの、プロセシング試料の下方の見かけのバンドへの分解も観察されたことから、プローブのある程度のプライミングオフが生じたことが指し示される。

10

【図18】代替のプロセシング法による標的配列のフォーカシングの増強。実施例13において得られた配列リードを、UCSC Genome Browserに表示して、特異的標的部位内で捕捉された断片のカバレッジおよび分布を評価した。一方はPLP1遺伝子のエキソンに対応し(AおよびB)、他方はZNF630遺伝子のイントロンセグメントに対応する(CおよびD)、X染色体上の2つの標的領域のための、「捕捉だけ」のライブラリーおよび「プロセシング」ライブラリーに対応する配列リードの密度(黒色)を示す。代替のプロセシング法により作製されたライブラリー(BおよびD)から得られるリードは、捕捉単独により構築されるライブラリー(AおよびC)より、標的部位内で高度に濃縮される。捕捉プローブ結合性部位を赤色で示す。各トラックは、ゲノム座標(x軸)の所与の連なりについて観察される、最大リード密度値(y軸)に照らしてスケールリングする。

20

【図19】パーティカルアラインメントについての概略図。全ての次世代配列(NGS: next generation sequence)解析は、参照ゲノムに照らしたアラインメントから始まる。(A)初期リードアラインメントは、一塩基変異体(SNV: single nucleotide variant)に対応し、限定された程度で、挿入/欠失にも対応しうる、コンフィギュラブルワードストリング検索に基づく。(B)整列されたリードの集合セットを、SNVについて解析する。示される例では、候補SNVを2回にわたり観察したが、これらの2つのリードのためのリード座標は同一であった。パーティカルアラインメントパラダイムは、直交的な検証を必要とする、多数のSNV仮説および/または挿入/欠失仮説をもたらす。

30

【図20】データ解析スキームの概略図。ステップ1は、リードをプローブとマッチさせることである。ステップ2は、各プローブに「水平に」接続された配列情報を解析することとなる。

【図21】リードの、プローブ1およびプローブ2との会合を「強い」水平アラインメントについての概略図。アセンブリは、1つは野生型のエキソン構造を伴い、1つは挿入構造を伴う、2種のコンティグをもたらすであろう。2つの重要な原理が立ち現れる: 1) 隣接するプローブから得られる重複リードは、捕捉されたエキソンのインデル含有対立遺伝子についての仮説に対する実証または反証となり、2) 捕捉プローブの外側のマイクロCNV対立遺伝子は、水平法により容易に検出可能である。

40

【図22】「信頼性の低い」SNV判定についての概略図。候補ヌクレオチド変異体は、試料を回収およびプロセシングされる個体において保有される実際のイベントでありうるが、また、試料のプロセシングおよびシーケンシングにおいて導入されるアーチファクトでもありうる。本明細書で記載される方法は、実際の「信頼性の高い」変異体判定と、アーチファクト的な「信頼性の低い」変異体判定とを差別化するようにデザインした。標的領域をカバーするシーケンシングリードを、複数の異なるクローンから、可能なシーケンシング配向性の両方において回収し、各リードをタグ情報で注釈する。タグは、同じクローニングイベントから得られるリードの同定および群分けを可能とする。全てが同じクローニングイベントから得られるクローンのうちの1種のセット内だけで発生するS

50

N Vおよびインデルは、さらなる解析から棄却される信頼性の低い判定である。

【図23】「信頼性の高い」S N V判定についての概略図。候補ヌクレオチド変異体は、試料を回収およびプロセシングされる個体において保有される実際のイベントでありうるが、また、試料のプロセシングおよびシーケンシングにおいて導入されるアーチファクトでもありうる。本明細書で記載される方法は、実際の「信頼性の高い」変異体判定と、アーチファクト的な「信頼性の低い」変異体判定とを差別化するようにデザインした。標的領域をカバーするシーケンシングリードを、複数の異なるクローンから、可能なシーケンシング配向性の両方において回収し、各リードをタグ情報で注釈する。タグは、異なるクローニングイベントから得られるリードの同定を可能とする。示される例は、(A) 開始点は同じであるが、配列標識が異なるリード、(B) 異なる開始点および異なる標識を有する、配向性が同じリード、ならびに(C) 配向性が逆向きのリードである。これらの全ての場合では、独立のクローニングイベントにおける変異体の発生および検出により、その変異体を高い信頼性でマークし、このような変異体を、さらなる直交的な検証法で追跡する。

10

【図24】分子注釈されたシーケンシングリード。(A) フォワードフローセル (Illuminaによる化学反応) のグラフト配列およびシーケンシングプライマー結合性部位。(B) リバースフローセルのグラフト配列およびリバースシーケンシングプライマーのアニリング部位。(1) 配列標識。(2) 試料標識。(3) フォワードリード開始部位。(4) ゲノム断片の配列。(5) ゲノムインデックス (プローブ配列)。(1) + (3) の組合せは、変異体の判定およびコピー数の決定のいずれにもきわめて重要な独特のリードタグを構成する。

20

【図25】DNA配列変異体 (挿入、欠失、点突然変異、および/または転座の発生) のうちの最も重要なクラスはまた、アラインメントベースの方法で検出するのが最も困難でもある。

【図26】二重プローブによる標的領域 (例えば、エキソン) の精査。(A) 典型的なエキソンの平均長は、100~150bpである。捕捉プローブは、標的セグメントを挟むイントロン領域内に配置される。これらのプローブは、配列極性が逆である (一方は「+」鎖をクエリーし、他方は「-」鎖クエリーする)。(B) 各個別のプローブ配列のリード分布は、影を付したエリアで指し示し、リードの配向性は、矢印で指定する。鍵となる側面は、標的領域を、いずれの配向性にもある、複数のリードでシーケンシングすることである。さらに、各プローブは、隣接するプローブ結合性部位をシーケンシングするリードも捕捉する。この配置は、変異体判定の信頼性を増大させる1つの要素である。

30

【図27】変異体の判定における配列タグの役割。配列「タグ」は、ヌクレオチドコード (卵型; Clearforkによる場合、3つのヌクレオチドによる16の可能な配列のコレクション) および任意のクローン断片の粘着末端配列を含む。(A) 偽陽性の変異体判定とは、全てが同一な配列タグを保有する同種配列のコレクションの中で変異体が同定される判定である。(B) 信頼性の高い変異体判定は、異なる配列タグを有する配列のコレクションの中で見出される。

【図28】リード観察統計を使用するコピーの決定。

【図29】分子注釈されたシーケンシングリード。(A) フォワードフローセル (IlluminaによるSBS化学反応) のグラフト配列およびシーケンシングプライマー結合性部位。(B) リバースフローセルのグラフト配列およびリバースシーケンシングプライマーのアニリング部位。(1) 配列標識。(2) 試料標識。(3) フォワードリード開始部位。(4) ゲノム断片の配列。(5) ゲノムインデックス (プローブ配列)。(6) 捕捉標識。(1) + (3) の組合せは、コピー数の決定にきわめて重要な独特のリードタグを構成する。(5) + (6) の組合せは、捕捉イベントをモニタリングし、定量化するのに使用されうる、ゲノムインデックスタグを構成する。注釈エレメント1、2、3、および4の配列を決定するフォワードシーケンシングリード1、ならびに注釈エレメント5および6の配列を決定するペアエンドリバースリード2が指し示される。

40

【図30】プローブ (例えば、多機能性捕捉プローブ) は一般に有向であり、プローブは

50

、配列をそれらの位置の一方の側（一般に 3' 側）において捕捉することを意味する。60 マーの標的化コアに加えて、さらなる機能性（例えば、PCR プライマー結合性部位、ビオチンによるプルアウトなどを可能とするパートナーオリゴのための結合性部位）を付加するテイル配列も付加する。60 ヌクレオチドの標的化配列は、以下の制約および基準：（１）プローブは、標的配列の開始点に照らして -100 ~ +50 ヌクレオチドに配置すること。目下の例示では、標的配列の「開始点」とは、イントロン：エキソン接合部であること；（２）プローブは、例示される通り、プローブ対から得られる配列が、逆向きの配向性で重複しているように、冗長性を伴ってデザインすること；（３）プローブは、33 % 以上の GC 含量（60 マー当たり > 20 の G または C）および 67 % 以下の（60 マー当たり < 40 の G または C）を保有するように選択する（可能な場合）こと；（４）10
プローブは、可能な場合は常にリピートを回避するように選択すること。これは、いずれもが U C S C ゲノムブラウザー上で閲覧しうる、R E P E A T M A S K E R および／または独特の整列可能性基準を一助として行うこと；（５）万一、位置要件、GC 要件、および独特のものであるという要件を満たすことができない場合は、選択規則を、以下の順位（GC > 位置 > 独特のものであること）で緩和することにより選択する。言い換えると、GC 基準および位置どり基準は厳密でないが、独特のものであるという基準は厳密である。

【図 3 1】標的化されたゲノムシーケンシングライブラリーを創出するためのプロセッシング。（A）初期捕捉複合体は、タグ付けされた「基準」ゲノムライブラリー断片、プローブに対して 5' のゲノム「標的領域」を標的化する、テイル処理された捕捉プローブ、20
および全てのプローブに共通な、ビオチニル化パートナーオリゴヌクレオチドを含む。（B）複合体の、配列準備されたクローンへのプロセッシングは、3つのステップ：（１）DNA ポリメラーゼホロ酵素（例えば、全長 B s t ポリメラーゼ）の 5' F L A P エンドヌクレアーゼにより、ゲノムクロンの 5' テイルをクリッピングするステップ；（２）ポリメラーゼにより、重合化を介して、パートナーオリゴ配列を伸長させるステップ（ステップ 1 と同時に生じうる）；および（３）T a q リガーゼにより、パートナーオリゴとゲノム断片との間のニックを修復するステップを含む。これらの協奏的ステップにより、配列準備されたクローンを創出する。

【図 3 2】図 3 2 は、捕捉後 PCR 産物／プロセッシング PCR 産物を示す。レーン 1 は、A C A 2 単一プライマーにより増幅された、プロセッシングされていない捕捉複合体である 30
。レーン 2 ~ 4 は、A F + C R 二重 PCR プライマーにより増幅した。

【図 3 3】図 3 3 ~ 3 5 は、タグ付けされたゲノム DNA を作製するためのライブラリーフリー法、ならびに関連する捕捉法、プロセッシング法、および解析法を示す。

【図 3 4】図 3 3 ~ 3 5 は、タグ付けされたゲノム DNA を作製するためのライブラリーフリー法、ならびに関連する捕捉法、プロセッシング法、および解析法を示す。

【図 3 5】図 3 3 ~ 3 5 は、タグ付けされたゲノム DNA を作製するためのライブラリーフリー法、ならびに関連する捕捉法、プロセッシング法、および解析法を示す。

【図 3 6】図 3 6 は、タグ付けされたゲノム DNA を作製するためのライブラリーフリー法におけるプライマー二量体によるアーチファクトを回避するための抑制 PCR 戦略を示す。 40

【図 3 7】図 3 7 は、タグ付けされたゲノム DNA を作製するためのライブラリーフリー法で使用される、生の g DNA および超音波処理した g DNA についてのゲル電気泳動結果を示す。

【図 3 8】図 3 8 は、ライブラリーフリー法により調製される 4 つの g DNA 試料についての、q PCR 増幅プロットを示す。

【図 3 9】図 3 9 は、ライブラリーフリー法により調製される試料から増幅される、生の PCR 産物によるゲル電気泳動結果を示す。

【図 4 0】図 4 0 は、ライブラリーフリー法により調製される試料から増幅される、ビーズで洗浄された PCR 産物によるゲル電気泳動結果を示す。

【図 4 1】図 4 1 は、異なる酵素の組合せ：T 4 DNA ポリメラーゼ（P）、T 4 D 50

NAリガーゼ (L)、およびT4遺伝子32タンパク質 (32)、または無酵素対照を使用する、ライブラリーフリー法により調製される試料についてのqPCR増幅プロットを示す。

【図42】図42は、異なる酵素の組合せ：T4 DNAポリメラーゼ (P)、T4 DNAリガーゼ (L)、およびT4遺伝子32タンパク質 (32)、または無酵素対照を使用する、ライブラリーフリー法により調製される、PCR増幅 (10サイクルまたは16サイクル) 試料についてのゲル電気泳動結果を示す。

【図43】図43は、プーリングの前にライブラリーフリー法により調製される個々の試料についてのゲル電気泳動結果を示す。

【図44】図44は、PLP1のCNVを、X染色体のコピー量を可変とする試料を通じた、常染色体の遺伝子座であるKRASおよびMYCの標準化との関連で示す。試料は、ライブラリーフリー法を使用して調製した。

10

【図45】図45は、XXXX (4コピー量) 試料中の、X染色体領域15のDNA配列の開始点を、捕捉プローブ配列と比べて示す。リードは、左から右に進み、試料は、ライブラリーフリー法を使用して調製した。

【図46】図46は、本明細書で想定されるRNA-seq法、捕捉されたcDNA、およびPippin自動式DNAサイズセクターでサイズ処理されたcDNA調製物を使用してRNA試料から調製される、cDNAについてのゲル電気泳動結果を示す。

【図47】図47は、ライブラリーを調製するための標的化された発現戦略と対比した、全RNAを使用して調製されるライブラリーにおける、心臓内の多様な転写物の遺伝子発現の相関を、肝臓と対比して示す。

20

【図48】図48は、全RNA-seqにおいて測定される多様な転写物の絶対発現レベルの相関を、標的化されたRNA-seqと比較して示す。

【発明を実施するための形態】

【0151】

詳細な説明

A. 概要

本発明は、少なくとも一部には、標的化遺伝学的解析の実行において、数種類の重要な分子モジュールの協調的な利用を用いることができるという発見に基づく。

【0152】

30

本発明の実施において、それと反対のことが特に示されていなければ、化学、生化学、有機化学、分子生物学、微生物学、組換えDNA技法、遺伝学、免疫学および細胞生物学の従来方法を当業者の技能範囲内で利用することができ、例証目的のために、その多くについて後述する。斯かる技法は、文献において十分に説明されている。例えば、Sambrook, et al., Molecular Cloning: A Laboratory Manual (3rd Edition, 2001); Sambrook, et al., Molecular Cloning: A Laboratory Manual (2nd Edition, 1989); Maniatis et al., Molecular Cloning: A Laboratory Manual (1982); Ausubel et al., Current Protocols in Molecular Biology (John Wiley and Sons, updated July 2008); Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology, Greene Pub. Associates and Wiley-Interscience; Glover, DNA Cloning: A Practical Approach, vol. I & II (IRL Press, Oxford, 1985); Anand, Techniques for the Analysis of Complex Genomes, (Academic Press, New York, 1992); Transcription and Translation (B. Hames & S. Higgins, Eds., 1984); Perbal, A Practical Guide to Molecular Cloning (1984);ならびにHarlow and Lane, Antibodies, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1998)を参照されたい。

40

【0153】

本明細書に引用されているあらゆる刊行物、特許および特許出願は、これにより参照としてその内容全体を援用する。

50

【0154】

B. 定義

他に定義されていなければ、本明細書に使用されているあらゆる技術および科学用語は、本発明が属する技術分野における当業者により一般に理解されているものと同じ意義を有する。本発明の実施または検査において、本明細書に記載されているものと同様または均等ないかなる方法および材料を使用することもできるが、組成物、方法および材料の好ましい実施形態が本明細書に記載されている。本発明の目的のために、次の用語を下に定義する。

【0155】

冠詞「a」、「an」および「the」は、この冠詞の文法上の目的語の1個または2個以上（即ち、少なくとも1個）を指すよう本明細書において使用されている。例として、「要素（an element）」は、1個の要素または2個以上の要素を意味する。

10

【0156】

選択肢（例えば、「または」）の使用は、選択肢のいずれか一方、両方またはこれらのいずれかの組合せを意味するものと理解されたい。

【0157】

用語「および／または」は、選択肢のいずれか一方または両方を意味するものと理解されたい。

【0158】

本明細書において、用語「約」または「およそ」は、参照含量、レベル、値、数、頻度、パーセンテージ、寸法、サイズ、量、重量または長さに対して15%、10%、9%、8%、7%、6%、5%、4%、3%、2%または1%ほど変動する含量、レベル、値、数、頻度、パーセンテージ、寸法、サイズ、量、重量または長さを指す。一実施形態において、用語「約」または「およそ」は、参照含量、レベル、値、数、頻度、パーセンテージ、寸法、サイズ、量、重量または長さに関して±15%、±10%、±9%、±8%、±7%、±6%、±5%、±4%、±3%、±2%または±1%の含量、レベル、値、数、頻度、パーセンテージ、寸法、サイズ、量、重量または長さの範囲を指す。

20

【0159】

本明細書を通して、文脈がそれ以外を要求しない限り、単語「を含む（comprise、comprisesおよびcomprising）」は、記述されているステップもしくはは要素またはステップもしくはは要素の群の包接を暗示するが、他のいかなるステップもしくはは要素またはステップもしくはは要素の群の排除も暗示しないものと理解されよう。特定の実施形態において、用語「含む（include）」、「有する（has）」、「含有する（contains）」および「を含む（comprise）」は、同義的に使用されている。

30

【0160】

「からなる（consisting of）」とは、いかなるものであれ語句「からなる」に続くものを含み、これに限定されることを意味する。よって、語句「からなる」は、列挙されている要素が必要または必須であり、他のいかなる要素も存在しなくてよいことを示す。

【0161】

「から本質的になる（consisting essentially of）」とは、この語句の後に列挙されているいかなる要素も含み、列挙されている要素の開示に指定されている活性または作用に干渉または寄与しない他の要素に限定されることを意味する。よって、語句「から本質的になる」は、列挙されている要素が、必要または必須であるが、他の要素が必要に応じてあり、列挙されている要素の活性または作用に影響するか否かに応じて、存在してもしなくてもよいことを示す。

40

【0162】

本明細書を通して、「一実施形態」、「実施形態」、「特定の実施形態」、「関連する実施形態」、「ある特定の実施形態」、「追加的な実施形態」もしくは「さらなる実施形態」またはこれらの組合せの言及は、実施形態に関連して記載されている特定の特色、構造または特徴が、本発明の少なくとも一実施形態において含まれることを意味する。よっ

50

て、本明細書を通した様々な箇所における前述の語句の出現は、必ずしも全てが同じ実施形態を指す訳ではない。さらに、特定の特色、構造または特徴は、1種または複数の実施形態においていずれかの適した様式で組み合わせることができる。

【0163】

本明細書において使用される場合、用語「単離される」は、その天然状態では通常それに付随している構成成分を実質的にまたは本質的に含まない材料を意味する。特定の実施形態において、用語「得られる」または「由来する」は、単離と同義的に使用される。

【0164】

本明細書において使用される場合、用語「DNA」は、デオキシリボ核酸を指す。様々な実施形態において、用語、DNAは、ゲノムDNA、組換えDNA、合成DNAまたはcDNAを指す。一実施形態において、DNAは、ゲノムDNAまたはcDNAを指す。特定の実施形態において、DNAは、「標的領域」を含む。本明細書において企図されるDNAライブラリーは、ゲノムDNAライブラリーおよびRNAから構築されるcDNAライブラリー、例えば、RNA発現ライブラリーを含む。様々な実施形態において、DNAライブラリーは、1種または複数の追加的なDNA配列および／またはタグを含む。

【0165】

「標的領域」は、DNA配列内の対象とする領域を指す。様々な実施形態において、標的化遺伝学的解析は、標的領域において実行される。特定の実施形態において、標的領域がシーケンシングされる、あるいは標的領域のコピー数が決定される。

【0166】

C. 例示的な実施形態

本発明は、一部には、タグ付けされたゲノムライブラリーを作製するための方法を企図する。特定の実施形態において、方法は、断片化されたDNA、例えば、ゲノムDNAまたはcDNAを末端修復酵素で処理して、断片化され末端修復されたDNAを作製し、続いてランダム核酸タグ配列をライゲーションして、タグ付けされたゲノムライブラリーを作製するステップを含む。一部の実施形態において、試料コード配列および／またはPCRプライマー配列が、必要に応じて、断片化され末端修復されたDNAにライゲーションされる。

【0167】

本発明は、一部には、タグ付けされたDNAライブラリーを作製するための方法を企図する。特定の実施形態において、方法は、断片化されたDNAを末端修復酵素で処理して、断片化され末端修復されたDNAを作製し、続いてランダム核酸タグ配列をライゲーションして、タグ付けされたDNAライブラリーを作製するステップを含む。一部の実施形態において、試料コード配列および／またはPCRプライマー配列が、必要に応じて、断片化され末端修復されたDNAにライゲーションされる。

【0168】

DNAを断片化するための例証的方法として、剪断、超音波処理、制限消化を含む酵素消化および他の方法が挙げられるがこれらに限定されない。特定の実施形態において、DNAを断片化するための当技術分野において公知のいずれかの方法を、本発明と共に用いることができる。

【0169】

一部の実施形態において、断片化されたDNAは、末端修復酵素によってプロセッシングされて、末端修復されたDNAを作製する。一部の実施形態において、末端修復酵素は、例えば、平滑末端、5'-オーバーハングおよび3'-オーバーハングを生じることができる。一部の実施形態において、末端修復されたDNAは、平滑末端を含有する。一部の実施形態において、末端修復されたDNAは、平滑末端を含有するようプロセッシングされる。一部の実施形態において、末端修復されたDNAの平滑末端は、単一塩基対オーバーハングを含有するようさらに修飾される。一部の実施形態において、平滑末端を含有する末端修復されたDNAは、アデニン(A)／チミン(T)オーバーハングを含有するようさらにプロセッシングすることができる。一部の実施形態において、平滑末端を含有する末

端修復されたDNAは、単一塩基対オーバーハングとしてアデニン（A）／チミン（T）オーバーハングを含有するようさらにプロセッシングすることができる。一部の実施形態において、末端修復されたDNAは、鋳型によらない（non-templated）3'オーバーハングを有する。一部の実施形態において、末端修復されたDNAは、3'オーバーハングを含有するようプロセッシングされる。一部の実施形態において、末端修復されたDNAは、ターミナルトランスフェラーゼ（TdT）により、3'オーバーハングを含有するようプロセッシングされる。一部の実施形態において、TdTによりG-テイルを付加することができる。一部の実施形態において、末端修復されたDNAは、いずれか公知の制限酵素（例えば、酵素Sau3Aその他）による部分的消化を使用して、オーバーハング末端を含有するようプロセッシングされる。

10

【0170】

特定の実施形態において、DNA断片は、1種または複数の「ランダムヌクレオチドタグ」または「ランダム核酸タグ」を使用してタグ付けされる。本明細書において、用語「ランダムヌクレオチドタグ」または「ランダム核酸タグ」は、個別の長さのポリヌクレオチドを指し、ヌクレオチド配列は、ランダムに作製または選択された。特定の例証的实施形態において、ランダム核酸タグの長さは、約2～約100ヌクレオチド、約2～約75ヌクレオチド、約2～約50ヌクレオチド、約2～約25ヌクレオチド、約2～約20ヌクレオチド、約2～約15ヌクレオチド、約2～約10ヌクレオチド、約2～約8ヌクレオチドまたは約2～約6ヌクレオチドである。ある特定の実施形態において、ランダムヌクレオチドタグの長さは、約2～約6ヌクレオチドである（例えば、図1を参照）。一実施形態において、ランダムヌクレオチドタグ配列は、約2、約3、約4、約5、約6、約7、約8、約9または約10ヌクレオチドである。

20

【0171】

特定の実施形態において、当技術分野において公知の方法を用いて、断片化されたDNAに本発明のランダムヌクレオチドタグを付加することができる。一部の実施形態において、「タグメンテーション（tagmentation）」を用いることができる。タグメンテーションは、市販のNextera技術であり（Illumina and Epicenter、米国）、トランスポゾンタンパク質複合体に本発明のランダムヌクレオチドタグおよび／または多機能性アダプターモジュールをロードするために使用することができる。次に、ロードされたトランスポゾン複合体を、記載されている方法に従ってタグ付けされた

30

【0172】

本方法において使用されるDNAは、当業者に公知のいかなる供給源に由来してもよい。DNAは、いずれかの供給源から収集し、コピーDNA（cDNA）としてRNAから合成し、本方法において使用される純粋または実質的に純粋DNAにプロセッシングすることができる。一部の実施形態において、断片化されたDNAのサイズは、約2～約500塩基対、約2～約400塩基対、約2～約300塩基対、約2～約250塩基対、約2～約200塩基対、約2～約100塩基対または約2～約50塩基対の範囲内である。

【0173】

DNA断片末端配列と導入された「ランダム核酸タグ」（単数または複数）の組合せは、以後「ゲノムタグ」または「cDNAタグ」と称される2要素の組合せを構成する。一部の実施形態において、「ゲノムタグ」または「cDNAタグ」が独特のものであることは、DNA断片末端配列プールの多様性を乗じた、付着されたランダムヌクレオチドタグプール内の多様性の組合せの積により決定することができる。

40

【0174】

本発明は、一部には、多機能性アダプターモジュールも企図する。本明細書において、用語「多機能性アダプターモジュール」は、（i）ランダムヌクレオチドタグ配列を含む第1の領域と、必要に応じて（ii）試料コード配列を含む第2の領域と、必要に応じて（iii）PCRプライマー配列を含む第3の領域とを含むポリヌクレオチドを指す。特定の実施形態において、多機能性アダプターモジュールは、PCRプライマー配列と、ラ

50

ンダムヌクレオチドタグと、試料コード配列とを含む。ある特定の実施形態において、多機能性アダプターモジュールは、PCRプライマー配列およびランダムヌクレオチドタグまたは試料コード配列を含む。一部の実施形態において、試料コードを含む第2の領域は必要に応じてである。一部の実施形態において、多機能性アダプターモジュールは、第2の領域を含まないが、代わりに、第1および第3の領域のみを含む。本発明の多機能性アダプターモジュールは、本明細書の他の箇所に開示されている末端と共に、断片化されたDNAに多機能性アダプターモジュールをライゲーションするための当業者に公知のその他の末端を含む、用いられるライゲーション方法に適切な平滑または相補的末端を含むことができる。

【0175】

10

様々な実施形態において、第1の領域は、ランダムヌクレオチドタグ配列を含む。特定の実施形態において、第1の領域は、約2～約100ヌクレオチド、約2～約75ヌクレオチド、約2～約50ヌクレオチド、約2～約25ヌクレオチド、約2～約20ヌクレオチド、約2～約15ヌクレオチド、約2～約10ヌクレオチド、約2～約8ヌクレオチドもしくは約2～約6ヌクレオチドまたはいずれかの介在する数のヌクレオチドのランダムヌクレオチドタグ配列を含む。

【0176】

特定の実施形態において、第2の領域は、必要に応じて存在する場合、試料コード配列を含む。本明細書において、用語「試料コード配列」は、試料の同定に使用されるポリヌクレオチドを指す。特定の実施形態において、第2の領域は、約1～約100ヌクレオチド、約2～約75ヌクレオチド、約2～約50ヌクレオチド、約2～約25ヌクレオチド、約2～約20ヌクレオチド、約2～約15ヌクレオチド、約2～約10ヌクレオチド、約2～約8ヌクレオチドもしくは約2～約6ヌクレオチドまたはいずれかの介在する数のヌクレオチドの試料コード配列を含む。

20

【0177】

ある特定の実施形態において、第3の領域は、必要に応じて存在する場合、PCRプライマー配列を含む。特定の実施形態において、第3の領域は、約5～約200ヌクレオチド、約5～約150ヌクレオチド、約10～約100ヌクレオチド、約10～約75ヌクレオチド、約10～約50ヌクレオチド、約10～約40ヌクレオチド、約20～約40ヌクレオチドもしくは約20～約30ヌクレオチドまたはいずれかの介在する数のヌクレオチドのPCRプライマー配列を含む。

30

【0178】

特定の実施形態において、ライゲーションステップは、多機能性アダプターモジュールを断片化され末端修復されたDNAにライゲーションすることを含む。このライゲーション反応を使用して、多機能性アダプター分子および／またはランダムヌクレオチドタグにライゲーションされた末端修復されたDNAを含む、タグ付けされたDNAライブラリーを作製することができる。一部の実施形態において、単一の多機能性アダプターモジュールが用いられる。一部の実施形態において、2種以上の多機能性アダプターモジュールが用いられる。一部の実施形態において、同一配列の単一の多機能性アダプターモジュールが、断片化され末端修復されたDNAの各末端にライゲーションされる。

40

【0179】

本発明は、多機能性捕捉プローブモジュールも提供する。本明細書において、用語「多機能性捕捉プローブモジュール」は、(i) パートナーオリゴヌクレオチドとハイブリダイズすることができる第1の領域と、(ii) 特異的標的領域とハイブリダイズすることができる第2の領域と、必要に応じて(iii) テイル配列を含む第3の領域とを含むポリヌクレオチドを指す。

【0180】

一実施形態において、多機能性捕捉プローブモジュールは、パートナーオリゴヌクレオチドとハイブリダイズすることができる領域と、DNA標的配列とハイブリダイズすることができる領域と、テイル配列とを含む。

50

【0181】

一実施形態において、多機能性捕捉プローブモジュールは、パートナーオリゴヌクレオチドとハイブリダイズすることができる領域と、ゲノム標的配列とハイブリダイズすることができる領域とを含む。

【0182】

特定の実施形態において、多機能性捕捉プローブモジュールは、必要に応じて、ランダムヌクレオチドタグ配列を含む。

【0183】

様々な実施形態において、第1の領域は、パートナーオリゴヌクレオチドとハイブリダイズすることができる領域を含む。本明細書において、用語「パートナーオリゴヌクレオチド」は、多機能性捕捉プローブモジュールのヌクレオチド配列に相補的なオリゴヌクレオチドを指す。特定の実施形態において、パートナーオリゴヌクレオチドとハイブリダイズすることができる第1の領域は、約20～約200ヌクレオチド、約20～約150ヌクレオチド、約30～約100ヌクレオチド、約30～約75ヌクレオチド、約20～約50ヌクレオチド、約30～約45ヌクレオチドまたは約35～約45ヌクレオチドの配列である。ある特定の実施形態において、領域は、約30～約50ヌクレオチド、約30～約40ヌクレオチド、約30～約35ヌクレオチドもしくは約34ヌクレオチドまたはいずれかの介在する数のヌクレオチドである。

【0184】

特定の実施形態において、第2の領域は、必要に応じて存在する場合、特異的DNA標的領域とハイブリダイズすることができる領域を含む。本明細書において、用語「DNA標的領域」は、本明細書において企図される組成物および方法を使用した解析に選択されるゲノムまたはcDNAの領域を指す。特定の実施形態において、特異的標的領域とハイブリダイズすることができる領域を含む第2の領域は、約20～約200ヌクレオチド、約30～約150ヌクレオチド、約50～約150ヌクレオチド、約30～約100ヌクレオチド、約50～約100ヌクレオチド、約50～約90ヌクレオチド、約50～約80ヌクレオチド、約50～約70ヌクレオチドまたは約50～約60ヌクレオチドの配列である。ある特定の実施形態において、第2の領域は、約60ヌクレオチドまたはいずれかの介在する数のヌクレオチドである。

【0185】

ある特定の実施形態において、第3の領域は、必要に応じて存在する場合、テイル配列を含む。本明細書において、用語「テイル配列」は、特定の実施形態において、PCRプライマー結合部位として役立つことができる、多機能性捕捉プローブモジュールの5'末端におけるポリヌクレオチドを指す。特定の実施形態において、第3の領域は、約5～約100ヌクレオチド、約10～約100ヌクレオチド、約5～約75ヌクレオチド、約5～約50ヌクレオチド、約5～約25ヌクレオチドまたは約5～約20ヌクレオチドのテイル配列を含む。ある特定の実施形態において、第3の領域は、約10～約50ヌクレオチド、約15～約40ヌクレオチド、約20～約30ヌクレオチドもしくは約20ヌクレオチドまたはいずれかの介在する数のヌクレオチドである。

【0186】

一実施形態において、多機能性捕捉プローブモジュールは、パートナーオリゴヌクレオチドとハイブリダイズすることができる領域と、ゲノム標的配列とハイブリダイズすることができる領域とを含む。多機能性捕捉プローブモジュールが、パートナーオリゴヌクレオチドとハイブリダイズすることができる領域と、ゲノム標的配列とハイブリダイズすることができる領域を含む特定の実施形態において、パートナーオリゴは、テイル配列またはプライマー結合部位として機能することもできる。

【0187】

一実施形態において、多機能性捕捉プローブモジュールは、テイル領域と、ゲノム標的配列とハイブリダイズすることができる領域とを含む。

【0188】

様々な実施形態において、多機能性捕捉プローブは、結合対の特異的メンバーを含み、多機能性捕捉プローブとハイブリダイズする、タグ付けされたDNAライブラリーの1種または複数の捕捉された断片の単離および／または精製を可能にする。特定の実施形態において、多機能性捕捉プローブは、ビオチンまたは別の適したハプテン、例えば、ジニトロフェノール、ジゴキシゲニンにコンジュゲートされる。

【0189】

本発明は、一部には、タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュールとハイブリダイズさせて、複合体を形成するステップをさらに企図する。一部の実施形態において、多機能性捕捉プローブモジュールは、DNAライブラリーにおける特異的ゲノム標的領域と実質的にハイブリダイズする。

10

【0190】

ハイブリダイゼーションまたはハイブリダイズ条件は、2種のヌクレオチド配列が、安定的複合体を形成するいずれかの反応条件を含むことができる；例えば、タグ付けされたDNAライブラリーおよび多機能性捕捉プローブモジュールが、安定的なタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を形成する条件。斯かる反応条件は、当技術分野において周知のものであり、当業者であれば、斯かる条件を適切にかつ本発明の範囲内で修正することができることを認められよう。実質的ハイブリダイゼーションは、多機能性捕捉プローブ複合体の第2の領域が、タグ付けされたDNAライブラリーの領域に対し、100%、99%、98%、97%、96%、95%、94%、93%、92%、91%、90%、89%、88%、85%、80%、75%または70%配列同一性、相同性または相補性を提示する場合に生じ得る。

20

【0191】

特定の実施形態において、多機能性捕捉プローブモジュールの第1の領域は、第2の領域が実質的にハイブリダイズするタグ付けされたDNAライブラリーの領域と実質的にハイブリダイズしない。一部の実施形態において、多機能性捕捉プローブモジュールの第3の領域は、多機能性捕捉プローブモジュールの第2の領域が実質的にハイブリダイズするタグ付けされたDNAライブラリーの領域と実質的にハイブリダイズしない。一部の実施形態において、多機能性捕捉プローブモジュールの第1および第3の領域は、多機能性捕捉プローブモジュールの第2の領域が実質的にハイブリダイズするタグ付けされたDNAライブラリーの領域と実質的にハイブリダイズしない。

30

【0192】

ある特定の実施形態において、本明細書において企図される方法は、タグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップを含む。特定の実施形態において、DNA複合体を単離するための方法は、当業者に周知のものであり、当業者により適切と考慮されるいかなる方法を、本発明の方法と共に用いてよい（Ausubelら、Current Protocols in Molecular Biology、2007～2012頁）。特定の実施形態において、複合体は、ビオチン-ストレプトアビジン単離技法を使用して単離される。一部の実施形態において、多機能性捕捉プローブモジュールの第1の領域とハイブリダイズすることができるパートナーオリゴヌクレオチドは、DNA複合体単離方法において使用されるカラム、ビーズまたは他の基板に連結されたストレプトアビジンと相互作用することができるビオチンを5'末端または3'末端に含有するよう修飾される。

40

【0193】

特定の実施形態において、多機能性捕捉プローブモジュールの第1の領域は、パートナーオリゴヌクレオチドに結合している。一部の実施形態において、多機能性捕捉プローブモジュールは、タグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体の形成前に、パートナーオリゴヌクレオチドに結合している。一部の実施形態において、多機能性捕捉プローブモジュールは、タグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体の形成後に、パートナーオリゴヌクレオチドに結合している。一部の実施形態において、多機能性捕捉プローブモジュールは、タグ付けされたDNA

50

Aライブラリー-多機能性捕捉プローブモジュール複合体の形成と同時に、パートナーオリゴヌクレオチドに結合している。一部の実施形態において、パートナーオリゴヌクレオチドは、化学修飾されている。

【0194】

特定の実施形態において、単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体からの一本鎖3'末端の除去が企図される。ある特定の実施形態において、方法は、単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を3'-5'エキソヌクレアーゼ酵素プロセッシングして、一本鎖3'末端を除去するステップを含む。

【0195】

ある特定の他の実施形態において、方法は、単離されたタグ付けされたDNAライブラリー断片を鋳型として利用した、多機能性捕捉プローブの5'-3'DNAポリメラーゼ伸長を実行するステップを含む。

【0196】

ある特定の他の実施形態において、方法は、5'FLAPエンドヌクレアーゼ、DNA重合およびDNAリガーゼによるニック閉鎖の協奏的作用により、ハイブリッド多機能性捕捉プローブによって単離されたタグ付けされたDNA標的分子を作製するステップを含む。

【0197】

単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体の3'-5'エキソヌクレアーゼ酵素プロセッシングのために、種々の酵素を用いることができる。特定の実施形態において用いることのできる、3'-5'エキソヌクレアーゼ酵素活性を提示する適した酵素の例証のための例として、T4またはエキソヌクレアーゼI、III、Vが挙げられるがこれらに限定されない (Shevelev IV, Hubscher U., 「The 3' 5' exonucleases」、Nat Rev Mol Cell Biol. 3巻 (5号) : 364~76頁 (2002年) も参照)。特定の実施形態において、3'-5'エキソヌクレアーゼ活性を含む酵素は、T4ポリメラーゼである。特定の実施形態において、例えば、T4またはエキソヌクレアーゼI、III、Vを含む、3'-5'エキソヌクレアーゼ酵素活性を提示し、プライマー鋳型伸長することができる酵素を用いることができる。Id. 3' 5'

【0198】

一部の実施形態において、本明細書において企図される方法は、上記および本明細書の他の箇所に記述される3'-5'エキソヌクレアーゼ酵素によりプロセッシングされた複合体に対してPCRを実行するステップを含む。特定の実施形態において、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分がコピーされる。一実施形態において、作製されるハイブリッド核酸分子は、多機能性捕捉プローブモジュールとハイブリダイズすることができる標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含む。

【0199】

様々な実施形態において、標的化遺伝学的解析のための方法も企図される。ある特定の実施形態において、標的化遺伝学的解析のための方法は、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、ゲノムライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、c) 3'-5'エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体に対して3'-5'エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖3'末端を除去するステップと、d) c) から得られる酵素によりプロセッシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉

10

20

30

40

50

ローブ分子のテイル部分がコピーされ、ハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができる標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、e) d) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む。

【0200】

様々な実施形態において、特異的標的領域のコピー数を決定するための方法が企図される。特定の実施形態において、特異的標的領域のコピー数を決定するための方法は、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、c) 3' - 5' エキソヌクレアーゼ活性を有する酵素を使用して、b) から得られる単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体に対して3' - 5' エキソヌクレアーゼ酵素プロセッシングを実行して、一本鎖3' 末端を除去するステップと、d) c) から得られる酵素によりプロセッシングされた複合体に対してPCR反応を実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分が複製され、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができる標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、e) d) におけるハイブリッド核酸のPCR増幅を実行するステップと、f) e) におけるPCR反応を定量化するステップであって、定量化が、特異的標的領域のコピー数の決定を可能にするステップとを含む。

【0201】

様々な実施形態において、標的化遺伝学的解析のための方法も企図される。ある特定の実施形態において、標的化遺伝学的解析のための方法は、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、ゲノムライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、c) 単離されたタグ付けされたDNAライブラリー断片を鋳型として利用して、多機能性捕捉プローブの5' - 3' DNAポリメラーゼ伸長を実行するステップと、d) c) から得られる酵素によりプロセッシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分がコピーされ、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができる標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、e) d) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む。

【0202】

様々な実施形態において、特異的標的領域のコピー数を決定するための方法が企図される。特定の実施形態において、特異的標的領域のコピー数を決定するための方法は、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、c) 単離されたタグ付けされたDNAライブラリー断片を鋳型として利用して、多機能性捕捉プローブの5' - 3' DNAポリメラーゼ伸長を実行するステップと、d) c) から得られる酵素によりプロセッシングされた複合体に対してPCR反応を実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分が複製され、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができる標的領域と、多機能性捕捉プローブモジ

ュールテイル配列の相補体とを含むステップと、e) d) におけるハイブリッド核酸のPCR増幅を実行するステップと、f) e) におけるPCR反応を定量化するステップであって、定量化が、特異的標的領域のコピー数の決定を可能にするステップとを含む。

【0203】

様々な実施形態において、標的化遺伝学的解析のための方法も企図される。ある特定の
実施形態において、標的化遺伝学的解析のための方法は、a) タグ付けされたDNAライ
ブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであ
って、この多機能性捕捉プローブモジュールが、ゲノムライブラリーにおける特異的標的
領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたD
NAライブラリー—多機能性捕捉プローブモジュール複合体を単離するステップと、(c
10) 5'FLAPエンドヌクレアーゼ、DNA重合およびDNAリガーゼによるニック閉鎖
の協奏的作用により、ハイブリッド多機能性捕捉プローブによって単離されたタグ付けさ
れたDNA標的分子を作製するステップと、d) c) から得られる酵素によりプロセシン
グされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作
製するために、多機能性捕捉プローブ分子のテイル部分がコピーされ、このハイブリッド
核酸分子が、多機能性捕捉プローブモジュールとハイブリダイズすることができる標的領
域と、多機能性捕捉プローブモジュールテイル配列の相補体とを含むステップと、e) d
20) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップと
を含む。

【0204】

様々な実施形態において、特異的標的領域のコピー数を決定するための方法が企図され
る。特定の実施形態において、特異的標的領域のコピー数を決定するための方法は、a)
タグ付けされたDNAライブラリーを多機能性捕捉プローブモジュール複合体とハイブリ
ダイズさせるステップであって、この多機能性捕捉プローブモジュールが、DNAライ
ブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から
得られるタグ付けされたDNAライブラリー—多機能性捕捉プローブモジュール複合体を
単離するステップと、c) 5'FLAPエンドヌクレアーゼ、DNA重合およびDNAリ
30 ガーゼによるニック閉鎖の協奏的作用により、ハイブリッド多機能性捕捉プローブによ
って単離されたタグ付けされたDNA標的分子を作製するステップと、d) c) から得られ
る酵素によりプロセッシングされた複合体に対してPCR反応を実行するステップであって
、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分が複
製され、このハイブリッド核酸分子が、多機能性捕捉プローブモジュールとハイブリダイ
ズすることができる標的領域と、多機能性捕捉プローブモジュールテイル配列の相補体と
を含むステップと、e) d) におけるハイブリッド核酸のPCR増幅を実行するステップ
と、f) e) におけるPCR反応を定量化するステップであって、定量化が、特異的標的
領域のコピー数の決定を可能にするステップとを含む。

【0205】

特定の実施形態において、当業者に周知のいずれかの標準PCR反応条件を使用して、
PCRを実行することができる。ある特定の実施形態において、e) におけるPCR反応
は、2種のPCRプライマーを用いる。一実施形態において、e) におけるPCR反応は
40 、標的領域とハイブリダイズする第1のPCRプライマーを用いる。特定の実施形態にお
いて、e) におけるPCR反応は、標的領域／テイル接合部におけるハイブリッド分子と
ハイブリダイズする第2のPCRプライマーを用いる。ある特定の実施形態において、e
) におけるPCR反応は、標的領域とハイブリダイズする第1のPCRプライマーと、標
的ゲノム領域／テイル接合部におけるハイブリッド分子とハイブリダイズする第2のPC
Rプライマーとを用いる。特定の実施形態において、第2のプライマーは、プライマーの
少なくとも1個または複数のヌクレオチドが、標的領域とハイブリダイズし、プライマー
の少なくとも1個または複数のヌクレオチドが、テイル配列とハイブリダイズするよう
に、標的領域／テイル接合部とハイブリダイズする。ある特定の実施形態において、ステ
50 ュップe) から得られるハイブリッド核酸分子は、シーケンシングされ、配列は、ホリゾン

タルに整列される、即ち、互いに整列されるが、参照配列とは整列されない。特定の実施形態において、ステップa)～e)は、1種または複数の多機能性捕捉プローブモジュール複合体により、1回または複数回反復される。多機能性捕捉プローブ複合体は、同じであっても異なっている場合、標的配列のいずれかのDNA鎖を標的化するように設計される。一部の実施形態において、多機能性捕捉プローブ複合体が異なっている場合、これらは、タグ付けされたDNAライブラリー内の同じ標的領域の近くとハイブリダイズする。一実施形態において、1種または複数の多機能性捕捉プローブは、標的領域からのあらゆる介在する距離を含む、タグ付けされたDNAライブラリーにおける標的領域の約5、10、15、20、25、30、35、40、45、50、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000bp以上の内でハイブリダイズする。

10

【0206】

一部の実施形態において、方法は、標的領域当たり2種の多機能性捕捉プローブモジュールを使用して実行することができ、一方は、標的領域上流の「ワトソン」鎖（非コードまたは鋳型鎖）とハイブリダイズし、一方は、標的領域下流の「クリック」鎖（コードまたは非鋳型鎖）とハイブリダイズする。

【0207】

特定の実施形態において、本明細書において企図される方法は、いずれかの数の多機能性プローブモジュールにより、複数回さらに実行することができ、例えば、標的領域当たり2、3、4、5、6、7、8、9または10種以上の多機能性捕捉プローブモジュールを用い、そのうちいずれかの数が、いずれかの組合せでワトソンまたはクリック鎖とハイブリダイズする。一部の実施形態において、差異の数のいずれかを同定するために、得られる配列は、互いに整列することができる。

20

【0208】

ある特定の実施形態において、1種または複数の多機能性プローブモジュールを使用して、単一の反応において複数の標的領域が照合され、例えば、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、50000、100000、500000種以上が照合される。

30

【0209】

コピー数は、独特のリードおよび重複したリードに関する有用な情報を提供することができると共に、公知リードのバリエーションの探索を補助することができる。本明細書において、用語「リード」、「リード配列」または「シーケンシングリード」は、同義的に使用されており、ポリヌクレオチドのシーケンシングにより得られるポリヌクレオチド配列を指す。特定の実施形態において、DNAタグ、例えば、ランダムヌクレオチドタグを使用して、解析されている核酸配列のコピー数を決定することができる。

【0210】

一実施形態において、多機能性捕捉プローブハイブリッドモジュールは、(i) パートナーオリゴヌクレオチドとハイブリダイズすることができ、PCRプライマーとして機能することができる第1の領域と、(ii) 特異的ゲノム標的領域とハイブリダイズすることができる第2の領域とを含む。

40

【0211】

様々な実施形態において、多機能性捕捉プローブハイブリッドモジュールの第1の領域は、PCRプライマー配列を含む。特定の実施形態において、この第1の領域は、いずれかの介在する数のヌクレオチドを含む、約5～約200ヌクレオチド、約5～約150ヌクレオチド、約10～約100ヌクレオチド、約10～約75ヌクレオチド、約10～約50ヌクレオチド、約10～約40ヌクレオチド、約20～約40ヌクレオチドまたは約20～約30ヌクレオチドのPCRプライマー配列を含む。

【0212】

50

特定の実施形態において、多機能性捕捉プローブハイブリッドモジュールの第1の領域は、パートナーオリゴヌクレオチドに結合している。ある特定の実施形態において、多機能性捕捉ハイブリッドプローブモジュールは、タグ付けされたDNAライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体の形成前に、パートナーオリゴヌクレオチドに結合している。特定の実施形態において、多機能性捕捉プローブハイブリッドモジュールは、タグ付けされたDNAライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体の形成後に、パートナーオリゴヌクレオチドに結合している。一部の実施形態において、多機能性捕捉プローブハイブリッドモジュールは、タグ付けされたDNAライブラリー-多機能性捕捉ハイブリッドプローブモジュール複合体の形成と同時に、パートナーオリゴヌクレオチドに結合している。一部の実施形態において、パートナーオリゴヌクレオチドは、化学修飾されている。

10

【0213】

様々な実施形態において、本明細書において企図される方法は、捕捉されたタグ付けされたDNAライブラリー配列をコピーして、多機能性捕捉プローブハイブリッドモジュール複合体と、ハイブリッドモジュールがゲノム標的とハイブリダイズする位置に対して多機能性捕捉プローブ配列の3'または5'に位置する捕捉されたタグ付けされたDNAライブラリー配列の領域に相補的な配列を含むハイブリッド核酸分子を作製することができるように、タグ付けされたDNAライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体に対してPCRを実行するステップを含む。特定の実施形態において、コピーされた標的領域は、多機能性捕捉プローブハイブリッドモジュールがゲノム標的にハイブリダイズする位置の配列の3'または5'末端から1~5000ntのいずれかの位置である。ある特定の実施形態において、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブハイブリッドモジュールがハイブリダイズする位置に対して3'の領域の相補的配列がコピーされる。作製されるハイブリッド核酸分子は、多機能性捕捉プローブハイブリッドモジュールと、多機能性捕捉プローブハイブリッドモジュールが標的領域とハイブリダイズする位置から3'または5'に位置する捕捉されたタグ付けされたDNAライブラリー配列の領域の相補体とを含む。

20

【0214】

様々な実施形態において、本明細書において企図される方法は、タグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体をプロセッシングして、ハイブリッド核酸分子（即ち、ハイブリッド多機能性捕捉プローブによって単離されたタグ付けされたDNA標的分子）を作製するステップを含む。特定の実施形態において、ハイブリッド核酸分子は、多機能性捕捉プローブハイブリッドモジュールと、多機能性捕捉プローブハイブリッドモジュールが標的領域とハイブリダイズする位置に対して3'に位置するタグ付けされたDNAライブラリー配列の領域の相補体とを含む。非限定的な一実施形態において、ハイブリッド核酸分子は、単離されたタグ付けされたDNAライブラリー-多機能性捕捉プローブモジュール複合体から一本鎖3'末端を除去する3'-5'エキソヌクレアーゼ酵素プロセッシングおよび/または多機能性捕捉プローブの5'-3'DNAポリメラーゼ伸長により作製される。

30

【0215】

他の特定の実施形態において、ハイブリッド核酸分子は、多機能性捕捉プローブハイブリッドモジュールと、多機能性捕捉プローブハイブリッドモジュールが標的領域とハイブリダイズする位置に対して5'に位置するタグ付けされたDNAライブラリー配列の領域の相補体とを含む。非限定的な一実施形態において、ハイブリッド核酸分子は、5'FLAPエンドヌクレアーゼ、DNA重合およびDNAリガーゼによるニック閉鎖の協奏的作用により作製される。

40

【0216】

様々な実施形態において、標的化遺伝学的解析のための方法が提供される。一実施形態において、標的化遺伝学的解析のための方法は、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップ

50

であって、この多機能性捕捉プローブハイブリッドモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、c) b) から得られる複合体に対してPCRを実行して、ハイブリッド核酸分子を形成するステップと、d) c) から得られるハイブリッド核酸分子に対して標的化遺伝学的解析を実行するステップとを含む。特定の実施形態において、ステップc) から得られるハイブリッド核酸分子は、シーケンシングされ、配列は、ホリゾンタルに整列される、即ち、互いに整列されるが、参照配列とは整列されない。ある特定の実施形態において、ステップa) ~ c) は、1種または複数の多機能性捕捉プローブモジュールにより1回または複数回反復される。

10

【0217】

多機能性捕捉プローブモジュールは、同じであっても異なってもよく、ゲノムのいずれかの鎖に対しハイブリダイズするよう設計される。一部の実施形態において、多機能性捕捉プローブモジュールが異なっている場合、これらは、タグ付けされたDNAライブラリーにおける同じ標的領域の1~5000ntからのいずれかの位置とハイブリダイズする。

【0218】

特定の実施形態において、方法は、2種の多機能性捕捉プローブモジュールを使用して2回実行することができ、一方は、ゲノム標的領域の上流とハイブリダイズし（即ち、5'末端に；即ち、フォワード多機能性捕捉プローブモジュールまたは複合体）、一方は、反対側のゲノム鎖におけるゲノム標的領域の下流とハイブリダイズする（即ち、3'末端に；即ち、リバース多機能性捕捉プローブモジュールまたは複合体）。

20

【0219】

一実施形態において、1種または複数の多機能性捕捉プローブは、標的領域からのあらゆる介在する距離を含む、タグ付けされたDNAライブラリーにおける標的領域の約5、10、15、20、25、30、35、40、45、50、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000bp以上の内でハイブリダイズする。

【0220】

一部の実施形態において、方法は、いずれかの数の多機能性プローブモジュールにより、複数回さらに実行することができ、例えば、標的領域当たり2、3、4、5、6、7、8、9、10種以上の多機能性捕捉プローブモジュールを用いて、そのうちいずれかの数がいずれかの組合せでワトソンまたはクリック鎖とハイブリダイズする。

30

【0221】

ある特定の実施形態において、1種または複数の多機能性プローブモジュールを使用して、単一の反応において複数の標的領域が照合され、例えば、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、50000、100000、500000種以上が照合される。

40

【0222】

特定の実施形態において、突然変異を同定するために、本方法により得られる配列は、参照配列と整列することなく、互いに整列することができる。ある特定の実施形態において、得られる配列は、必要に応じて、参照配列と整列することができる。

【0223】

様々な実施形態において、特異的標的領域のコピー数を決定するための方法が企図される。特定の実施形態において、特異的標的領域のコピー数を決定するための方法は、a) タグ付けされたDNAライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブハイブリッドモジュールが、DNAライブラリーにおける特異的標的領域と選択的にハイブリダイズする

50

ステップと、b) a) から得られるタグ付けされたDNAライブラリー-多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、c) b) から得られる複合体に対してPCRを実行して、ハイブリッド核酸分子を形成するステップと、d) c) におけるハイブリッド核酸のPCR増幅を実行するステップと、e) d) におけるPCR反応を定量化するステップであって、定量化が、特異的標的領域のコピー数の決定を可能にするステップとを含む。特定の実施形態において、PCRは、当業者に周知のいずれかの標準PCR反応条件を使用して実行することができる。ある特定の実施形態において、d) におけるPCR反応は、2種のPCRプライマーを用いる。特定の実施形態において、d) におけるPCR反応は、2種のPCRプライマーを用い、そのそれぞれは、多機能性捕捉プローブハイブリッドモジュールがタグ付けされたDNAライブラリーとハイブリダイズする位置に対して下流の領域とハイブリダイズする。さらなる実施形態において、PCRプライマーがハイブリダイズする領域は、ステップc) において増幅される領域に位置する。様々な実施形態において、ステップc) から得られるハイブリッド核酸分子は、シーケンシングされ、配列は、ホリゾンタルに整列され、即ち、互いに整列されるが、参照配列とは整列されない。特定の実施形態において、ステップa) ~ c) は、1種または複数の多機能性捕捉プローブモジュールにより1回または複数回反復される。多機能性捕捉プローブモジュールは、同じであっても異なってもよく、ゲノムのいずれかの鎖とハイブリダイズするよう設計される。

10

【0224】

一実施形態において、1種または複数の多機能性捕捉プローブは、標的領域からのあらゆる介在する距離を含む、タグ付けされたDNAライブラリーにおける標的領域の約5、10、15、20、25、30、35、40、45、50、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000bp以上の内でハイブリダイズする。

20

【0225】

一部の実施形態において、方法は、いずれかの数の多機能性プローブモジュールにより、複数回さらに実行することができ、例えば、標的領域当たり2、3、4、5、6、7、8、9、10種以上の多機能性捕捉プローブモジュールを用いて、そのうちいずれかの数がいずれかの組合せでワトソンまたはクリック鎖とハイブリダイズする。

30

【0226】

ある特定の実施形態において、1種または複数の多機能性プローブモジュールを使用して、単一の反応において複数の標的領域が照合され、例えば、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、50000、100000、500000種以上が照合される。

【0227】

特定の例証的实施形態において、例えばPCRにより、タグ付けされたDNAライブラリーが増幅されて、増幅されたタグ付けされたDNAライブラリーを作製する。

【0228】

あらゆるゲノム標的領域は、5'末端および3'末端を有するであろう。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、それぞれ5'および3'方向の両方からの標的化ゲノム領域の増幅をもたらす2種の多機能性捕捉プローブ複合体により実行することができる。一実施形態において、1種または複数の多機能性捕捉プローブは、標的領域からのあらゆる介在する距離を含む、タグ付けされたDNAライブラリーにおける標的領域の約5、10、15、20、25、30、35、40、45、50、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000bp以上の内でハイブリダイズする。

40

【0229】

50

一部の実施形態において、方法は、いずれかの数の多機能性プローブモジュールにより、複数回さらに実行することができ、例えば、標的領域当たり2、3、4、5、6、7、8、9、10種以上の多機能性捕捉プローブモジュールを用いて、そのうちいずれかの数がいずれかの組合せでワトソンまたはクリック鎖とハイブリダイズする。

【0230】

ある特定の実施形態において、1種または複数の多機能性プローブモジュールを使用して、単一の反応において複数の標的領域が照合され、例えば、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、50000、100000種以上が照合される。

10

【0231】

特定の実施形態において、標的化遺伝学的解析は、配列解析である。特定の実施形態において、配列解析は、一配列が第2の配列から区別されるいずれかの解析を含む。様々な実施形態において、配列解析は、シーケンシングのための組成物または方法の非存在下で実行されるいずれかの純粋にメンタルな(mental)配列解析を排除する。ある特定の実施形態において、配列解析として、シーケンシング、一塩基多型(SNP)解析、遺伝子コピー数解析、ハプロタイプ解析、突然変異解析、メチル化状態解析(限定されない例として、非メチル化シトシン残基の亜硫酸水素塩変換により決定される)、クロマチン免疫沈降実験において得られるDNA配列の標的化リシーケンシング(CHIP-seq)、妊娠中の母体血漿DNAから収集された捕捉された胎児DNAの配列における父子鑑定、微生物特異的捕捉プローブにより捕捉された試料における微生物存在および集団評価ならびに胎児遺伝的配列解析(例えば、母体試料における胎児細胞または細胞外胎児DNAを使用)が挙げられる。

20

【0232】

コピー数解析として、特定の遺伝子または所定のゲノムDNA試料において生じる突然変異のコピー数を試験する解析が挙げられるがこれに限定されず、所定の遺伝子のコピー数または所定の試料における配列の差の定量的決定をさらに含むことができる。

【0233】

参照配列とアラインメントする必要なく実行することができる、配列アラインメント解析のための方法も本明細書において企図され、これは、本明細書において、水平配列解析と称される(例えば、図20に例証)。斯かる解析は、本明細書において企図される方法または他のいずれかの方法によって作製されるいずれかの配列において実行することができる。特定の実施形態において、配列解析は、本明細書において企図される方法により得られるハイブリッド核酸分子において配列アラインメントを実行するステップを含む。一実施形態において、1種または複数の多機能性捕捉プローブは、標的領域からのあらゆる介在する距離を含む、タグ付けされたDNAライブラリーにおける標的領域の約5、10、15、20、25、30、35、40、45、50、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000bp以上の内でハイブリダイズする。

30

40

【0234】

一部の実施形態において、方法は、いずれかの数の多機能性プローブモジュールにより、複数回さらに実行することができ、例えば、標的領域当たり2、3、4、5、6、7、8、9、10種以上の多機能性捕捉プローブモジュールを用いて、そのうちいずれかの数がいずれかの組合せでワトソンまたはクリック鎖とハイブリダイズする。

【0235】

ある特定の実施形態において、1種または複数の多機能性プローブモジュールを使用して、単一の反応において複数の標的領域が照合され、例えば、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、50000

50

、100000種以上が照合される。

【0236】

特定の実施形態において、DNAは、いずれかの生物学的供給源から単離することができる。DNAの例証的な供給源として、血液、皮膚、毛、毛包、唾液、口腔粘膜、腔粘膜、汗、涙、上皮組織、尿、精液、精子液、精漿、前立腺液、尿道球腺液（カウパー氏腺液）、排泄物、生検、腹水、脳脊髄液、リンパ液または組織抽出物試料または生検試料が挙げられるがこれらに限定されない。

【0237】

一実施形態において、本明細書において企図される方法により使用するためのタグ付けされたDNAライブラリーが提供される。一部の実施形態において、タグ付けされたDNAライブラリーは、タグ付けされたゲノム配列を含む。特定の実施形態において、各タグ付けされたDNA配列は、i) 断片化され末端修復されたDNAと、ii) 1種または複数のランダムヌクレオチドタグ配列と、iii) 1種または複数の試料コード配列と、iv) 1種または複数のPCRプライマー配列とを含む。

10

【0238】

一実施形態において、ハイブリッドタグ付けされたDNAライブラリーが企図される。特定の実施形態において、ハイブリッドタグ付けされたDNAライブラリーは、ハイブリッドタグ付けされたDNA配列を含む。ある特定の実施形態において、各ハイブリッドタグ付けされたDNA配列は、i) 標的領域を含む断片化され末端修復されたDNAと、ii) 1種または複数のランダムヌクレオチドタグ配列と、iii) 1種または複数の試料コード配列と、iv) 1種または複数のPCRプライマー配列と、v) 多機能性捕捉プローブモジュールテイル配列とを含む。

20

【0239】

様々な実施形態において、本明細書において企図される方法において使用される試薬のキットおよび組成物が提供される。一部の実施形態において、組成物は、タグ付けされたDNAライブラリーと、多機能性アダプターモジュールと、多機能性捕捉プローブモジュールとを含む。特定の実施形態において、組成物は、タグ付けされたゲノムライブラリーを含む。ある特定の実施形態において、組成物は、ハイブリッドタグ付けされたゲノムライブラリーを含む。

【0240】

30

様々な実施形態において、本明細書において企図される方法を行うための反応混合物が提供される。特定の実施形態において、反応混合物は、本明細書において企図される方法のいずれかを実行するための反応混合物である。ある特定の実施形態において、反応混合物は、タグ付けされたDNAライブラリーを作製することができる。一部の実施形態において、タグ付けされたDNAライブラリーを作製することができる反応混合物は、a) 断片化されたDNAと、b) 断片化され末端修復されたDNAを作製するためのDNA末端修復酵素とを含む。特定の実施形態において、反応混合物は、多機能性アダプターモジュールをさらに含む。様々な実施形態において、反応混合物は、多機能性捕捉プローブモジュールをさらに含む。ある特定の実施形態において、反応混合物は、3' - 5' エキソヌクレアーゼ活性およびPCR増幅活性を有する酵素をさらに含む。

40

【0241】

様々な実施形態において、本明細書において企図される1種または複数のクローンの配列のために、DNA配列解析のための方法が提供される。一実施形態において、方法は、各クローンが第1のDNA配列および第2のDNA配列を含む、1種もしくは複数のまたは複数のタグ付けされたDNAライブラリークローンを得るステップであって、この第1のDNA配列が、標的化DNA配列を含み、この第2のDNA配列が、捕捉プローブ配列を含むステップと、1種または複数のクローンに対してペアエンドシーケンシング反応を実行して1種または複数のシーケンシングリードを得るステップと、あるいは1種または複数のクローンに対してシーケンシング反応を実行するステップであって、約100、200、300、400、500以上のヌクレオチドを超える単一の長いシーケン

50

シングリードが得られ、このリードが、第1のDNA配列および第2のDNA配列の両方の同定に十分であるステップと、シーケンシングリードのプロープ配列に従って、1種または複数のクローンのシーケンシングリードを順序付けまたはクラスター形成するステップとを含む。

【0242】

配列リードは、1種または複数のヒト参照DNA配列と比較することができる。参照配列とマッチしない配列リードを同定し、非マッチ配列データからのデノボアセンブリの作成に使用することができる。特定の実施形態において、デノボアセンブリを使用して、捕捉プローブに関連する新規配列再編成を同定する。

【0243】

様々な実施形態において、各クローンが第1のDNA配列および第2のDNA配列を含む、1種もしくは複数のまたは複数のクローンを得るステップであって、この第1のDNA配列が、ランダムヌクレオチドタグ配列および標的化DNA配列を含み、この第2のDNA配列が、捕捉プローブ配列を含むステップを含む、コピー数決定解析のための方法が提供される。関連する実施形態において、1種または複数のクローンにおけるペアエンドシーケンシング反応が実行され、1種または複数のシーケンシングリードが得られる。別の実施形態において、1種または複数のクローンにおけるシーケンシング反応が実行され、約100ヌクレオチドを超える単一の長いシーケンシングリードが得られ、このリードが、第1のDNA配列および第2のDNA配列の両方の同定に十分である。1種または複数のクローンのシーケンシングリードは、シーケンシングリードのプロープ配列に従って、順序付けまたはクラスター形成することができる。

【0244】

特定の実施形態において、コピー数を決定するための方法が提供される。特定の実施形態において、方法は、各クローンが第1のDNA配列および第2のDNA配列を含む、1種もしくは複数のまたは複数のクローンを得るステップであって、この第1のDNA配列が、ランダムヌクレオチドタグ配列および標的化DNA配列を含み、この第2のDNA配列が、捕捉プローブ配列を含むステップと、シーケンシングリードのプロープ配列に従って、1種または複数のクローンのシーケンシングリードを順序付けまたはクラスター形成するステップとを含む。特定の実施形態において、ランダムヌクレオチドタグは、約2～約50ヌクレオチドの長さである。

【0245】

方法は、独特のシーケンシングリードおよび重複したシーケンシングリードの分布を決定することにより、第2のリード配列に関連するあらゆるシーケンシングリードを解析するステップと、独特のリードが遭遇する回数を計数するステップと、独特のリードの度数分布を統計分布に適合させるステップと、独特のリードの総数を推量するステップと、推量された独特のリードの総数を、ヒトが一般に二倍体であるという仮定に対して正規化するステップとをさらに含むことができる。

【0246】

特定の実施形態において、本明細書において企図される方法を使用して、1種または複数の標的化遺伝子座の推量されるコピー数を計算し、あるとすれば、予想されるコピー数値からのこの計算の逸脱を計算することができる。ある特定の実施形態において、遺伝子の1種または複数の標的化遺伝子座は、遺伝子座のコレクションにおいて共にグループ化され、標的化遺伝子座のコレクションのコピー数測定値は、平均化および正規化される。一実施形態において、遺伝子の推量されるコピー数は、この遺伝子を表す全標的遺伝子座の正規化された平均により表すことができる。

【0247】

様々な実施形態において、本明細書において企図される組成物および方法は、RNA発現の作製および解析にも適用可能である。いかなる特定の理論に制約されることも望まないが、タグ付けされたgDNAライブラリーの作製に使用される方法および組成物のいずれかを、タグ付けされたcDNAライブラリーの作製に使用することもでき、配列解析を

限定することなく含む、その後のRNA発現解析のためにcDNAにおいて具体化されるRNA配列に相当する標的領域を捕捉およびプロセッシングできることが企図される。

【0248】

様々な実施形態において、タグ付けされたRNA発現ライブラリーを作製するための方法は、まず、cDNAライブラリーを得るまたは調製するステップを含む。cDNAライブラリー合成の方法は、当技術分野において公知のものであり、様々な実施形態に適用可能となり得る。cDNAライブラリーは、適用に応じて、1種または複数の同じまたは異なる細胞型から調製することができる。一実施形態において、方法は、cDNAライブラリーを断片化するステップと、この断片化されたcDNAライブラリーを末端修復酵素で処理して、断片化され末端修復されたcDNAを作製するステップと、多機能性アダプター分子を、断片化され末端修復されたcDNAにライゲーションして、タグ付けされたRNA発現ライブラリーを作製するステップとを含む。

10

【0249】

特定の実施形態において、タグ付けされたRNA発現ライブラリー（cDNAライブラリー）は、1個または複数の細胞の全RNAからcDNAライブラリーを得るまたは調製し、このcDNAライブラリーを断片化し、断片化されたcDNAを末端修復酵素で処理して、断片化され末端修復されたcDNAを作製し、多機能性アダプター分子を、断片化され末端修復されたcDNAにライゲーションして、タグ付けされたRNA発現ライブラリーを作製することにより調製される。

20

【0250】

ある特定の実施形態において、cDNAライブラリーは、オリゴdTプライミングcDNAライブラリーである。

【0251】

ある特定の実施形態において、cDNAライブラリーは、約6〜約20ランダムヌクレオチドを含むランダムオリゴヌクレオチドによってプライミング作製される。特定の好ましい実施形態において、cDNAライブラリーは、ランダムヘキサマーまたはランダムオクタマーによってプライミングされる。

【0252】

所望の平均ライブラリー断片サイズを達成するために、cDNAライブラリーは、公知の方法を使用して剪断または断片化することができる。一実施形態において、cDNAライブラリーは、約250bp〜約750bpの平均サイズに断片化される。ある特定の実施形態において、cDNAライブラリーは、約500bpの平均サイズに断片化される。

30

【0253】

様々な実施形態において、本明細書において企図されるRNA発現ライブラリーは、軽微な変化ありまたはなしで、タグ付けされたゲノムDNAライブラリーを捕捉、プロセッシングおよびシーケンシングするための本明細書において企図される方法のいずれかを使用して、捕捉、プロセッシング、増幅およびシーケンシング等することができる。

【0254】

一実施形態において、タグ付けされたRNA発現ライブラリーを多機能性捕捉プローブモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブモジュールが、タグ付けされたRNA発現ライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、タグ付けされたRNA発現ライブラリー-多機能性捕捉プローブモジュール複合体を単離するステップと、単離されたタグ付けされたRNA発現ライブラリー-多機能性捕捉プローブモジュール複合体に対して3'〜5'エキソヌクレアーゼ酵素プロセッシングおよび／または5'〜3'DNAポリメラーゼ伸長を実行するステップと、酵素によりプロセッシングされた複合体に対してPCRを実行するステップであって、ハイブリッド核酸分子を作製するために、多機能性捕捉プローブ分子のテイル部分（例えば、PCRプライマー結合部位）がコピーされ、このハイブリッド核酸分子が、標的領域の相補体、特異的多機能性捕捉プローブ配列および捕捉モジュールテイル配列を含むステップと、ハイブリッド核酸分子に対して標的化遺伝子発現解析を実行するステッ

40

50

プとを含む、標的化遺伝子発現解析のための方法が提供される。

【0255】

一実施形態において、標的化遺伝子発現解析のための方法は、タグ付けされたRNA発現ライブラリーを多機能性捕捉プローブハイブリッドモジュール複合体とハイブリダイズさせるステップであって、この多機能性捕捉プローブハイブリッドモジュールが、RNA発現ライブラリーにおける特異的標的領域と選択的にハイブリダイズするステップと、タグ付けされたRNA発現ライブラリー多機能性捕捉プローブハイブリッドモジュール複合体を単離するステップと、複合体に対してPCRを実行して、ハイブリッド核酸分子を形成するステップとを含む。

【0256】

10

特定の実施形態において、少なくとも2回のハイブリダイゼーションステップにおいて、少なくとも2種の異なる多機能性捕捉プローブモジュールが使用され、この少なくとも2回のハイブリダイゼーションステップは、それぞれ1種の多機能性捕捉プローブモジュールを用いる。ある特定の実施形態において、少なくとも1種の多機能性捕捉プローブモジュールは、標的領域の5'とハイブリダイズし、少なくとも1種の多機能性捕捉プローブモジュールは、標的領域の3'とハイブリダイズする。

【0257】

一実施形態において、1種または複数の多機能性捕捉プローブは、標的領域からのあらゆる介在する距離を含む、タグ付けされたRNA発現またはcDNAライブラリーにおける標的領域の約5、10、15、20、25、30、35、40、45、50、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000bp以上の内にハイブリダイズする。

20

【0258】

一部の実施形態において、方法は、いずれかの数の多機能性プローブモジュールにより、複数回さらに実行することができ、例えば、標的領域当たり2、3、4、5、6、7、8、9、10種以上の多機能性捕捉プローブモジュールを用いて、そのうちいずれかの数がいずれかの組合せでワトソンまたはクリック鎖とハイブリダイズする。

【0259】

ある特定の実施形態において、1種または複数の多機能性プローブモジュールを使用して、単一の反応において複数の標的領域が照合され、例えば、100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、50000、100000種以上が照合される。

30

【0260】

さらなる実施形態において、当業者にcDNAライブラリーからの遺伝子発現解析を実行させることができる、cDNA配列解析のための方法が提供される。特定の実施形態において、本明細書において企図されるシーケンシング方法のいずれかは、タグ付けされたゲノムクロンのシーケンシングへのその適用から殆どまたは全く逸脱することなく、cDNAライブラリーのシーケンシングに適応させることができる。上述の通り、本明細書において企図されるRNA発現解析におけるcDNAの標的領域のタグ付けされたcDNAシーケンシングリードの統計分布は、cDNAライブラリーを調製したまたは得た細胞における標的領域の遺伝子発現のレベルと相関する。

40

【0261】

本明細書に引用されているあらゆる刊行物、特許出願および交付された特許は、あたかも個々の刊行物、特許出願または交付された特許のそれぞれが、参照として援用すると特にかつ個々に示されているかのように、ここに参照として援用する。

【0262】

理解を明確にするために、例証および具体例として、前述の発明を詳細に記載してきたが、当業者であれば、本発明の教示するところを鑑み、添付の特許請求の範囲の精神また

50

は範囲から逸脱することなく、これにある一定の変更および修正を行ってよいことが容易に明らかとなるであろう。次の実施例は、限定としてではなく、単なる例証として示されている。当業者であれば、本質的に同様の結果を得るために変更または修正することができる、種々の重大でないパラメータを容易に認識するであろう。

【実施例】

【0263】

(実施例1)

遺伝子解析のための標的ゲノム領域の調製
総括

特定の実施形態では、本明細書で想定される方法は、いくつかの鍵となる分子モジュールの組織化された活用を含む。以下の節では、各モジュールについて個別に記載する。本節の末尾では、モジュールの相互関連について記載する。

【0264】

1節：ゲノムDNA断片へのタグ付け

個体から得られるゲノムDNAは、回収し、純粋なDNAにプロセッシングすることができ、1ヌクレオチド以上の、一部の実施形態では、2～100ヌクレオチドの範囲の、または2～6ヌクレオチドの範囲の、断片化されたランダムなヌクレオチド配列を、ゲノムDNA断片のランダム末端に結合させる（図1）。導入されたランダムヌクレオチドタグ配列の、ゲノム断片末端配列との組合せは、本明細書の以下では、多機能性アダプターモジュールの第1の領域と称する、2つのエレメントの独特の組合せを構成する。多機能性アダプターモジュールの第1の領域が独特のものであることは、結合させた多機能性アダプターモジュールの第1の領域のプール内の多様性に、ゲノム断片末端配列の多様性を乗じた組合せの積により決定する。

【0265】

2節：試料特異的コードおよびユニバーサル増幅配列の付加

多機能性アダプター分子は、試料特異的コード（本明細書では、多機能性アダプターモジュールの第2の領域と称する）およびユニバーサル増幅配列（本明細書では、PCRプライマー配列または多機能性アダプターモジュールの第3の領域と称する）もさらに含む。多機能性アダプターモジュールの第1の領域から得られる、導入されたランダムヌクレオチドに加えて、断片化されたゲノムDNAに結合させた各セグメントは、この領域のDNA配列を使用して、複数の試料が一体に組み合わせられている配列のセット内の所与の試料配列を独特に同定（言い換えると、試料のバーコード化）しうるように、各試料に共通であるが、試料間で異なるヌクレオチドのさらなるセットも含みうる。加えて、結合させたヌクレオチド配列は、ポリヌクレオチドを増幅する（例えば、PCRにより）のに使用されうる、ユニバーサル配列も含有しうる。ランダムヌクレオチドタグ配列、試料コード、およびユニバーサル増幅配列のエレメントの組合せは、最も一般的には、ヌクレオチドライゲーションにより、断片化されたゲノムDNAに結合させる、「アダプター」（また、多機能性アダプターモジュールとも称する）を構成する。

【0266】

断片化されたgDNAにライゲーションされた多機能性アダプターモジュールの例示的な（illustrative）例を、図1に例示し、（例により限定されることを望まないが）このような配列の例示的な（exemplary）セットを、表1に示す。表1内では、アダプター配列のセットを、4種のアダプター配列のセットにクラスター形成する。各欄内では、同じ2種の塩基コードを共有する全てのアダプターおよび可能な16のランダムタグの全てを表す。可能な16のアダプターは、断片とのライゲーションの前に混合する。各アダプターのトップ鎖である「ライゲーション鎖」だけを示すが、これは、末端修復されたDNA断片に共有結合させる鎖である。最終的に失われるボトム鎖であるパートナー鎖は、図1には示すが、表1には組み入れない。

【0267】

【表 1 - 1】

表1: 64のタグ付け試料およびIDアダプターのセットの5'-3'DNA配列

試料コードAC	配列番号	試料コードGA	配列番号
GACCAGAGAATTCGAATACAAAAC	1	GACCAGAGAATTCGAATACAAAGA	17
GACCAGAGAATTCGAATACAACAC	2	GACCAGAGAATTCGAATACAACGA	18
GACCAGAGAATTCGAATACAAGAC	3	GACCAGAGAATTCGAATACAAGGA	19
GACCAGAGAATTCGAATACAATAC	4	GACCAGAGAATTCGAATACAATGA	20
GACCAGAGAATTCGAATACACAAC	5	GACCAGAGAATTCGAATACACAGA	21
GACCAGAGAATTCGAATACACCAC	6	GACCAGAGAATTCGAATACACCGA	22
GACCAGAGAATTCGAATACACGAC	7	GACCAGAGAATTCGAATACACGGA	23
GACCAGAGAATTCGAATACACTAC	8	GACCAGAGAATTCGAATACACTGA	24
GACCAGAGAATTCGAATACAGAAC	9	GACCAGAGAATTCGAATACAGAGA	25
GACCAGAGAATTCGAATACAGCAC	10	GACCAGAGAATTCGAATACAGCGA	26
GACCAGAGAATTCGAATACAGGAC	11	GACCAGAGAATTCGAATACAGGGA	27
GACCAGAGAATTCGAATACAGTAC	12	GACCAGAGAATTCGAATACAGTGA	28
GACCAGAGAATTCGAATACATAAC	13	GACCAGAGAATTCGAATACATAGA	29
GACCAGAGAATTCGAATACATCAC	14	GACCAGAGAATTCGAATACATCGA	30
GACCAGAGAATTCGAATACATGAC	15	GACCAGAGAATTCGAATACATGGA	31
GACCAGAGAATTCGAATACATTAC	16	GACCAGAGAATTCGAATACATTGA	32
試料コードCT	配列番号	試料コードTG	配列番号
GACCAGAGAATTCGAATACAAACT	33	GACCAGAGAATTCGAATACAAATG	49
GACCAGAGAATTCGAATACAACCT	34	GACCAGAGAATTCGAATACAACCTG	50
GACCAGAGAATTCGAATACAAGCT	35	GACCAGAGAATTCGAATACAAGTG	51
GACCAGAGAATTCGAATACAATCT	36	GACCAGAGAATTCGAATACAATTG	52
GACCAGAGAATTCGAATACACACT	37	GACCAGAGAATTCGAATACACATG	53
GACCAGAGAATTCGAATACACCCT	38	GACCAGAGAATTCGAATACACCTG	54
GACCAGAGAATTCGAATACACGCT	39	GACCAGAGAATTCGAATACACGTG	55
GACCAGAGAATTCGAATACACTCT	40	GACCAGAGAATTCGAATACACTTG	56
GACCAGAGAATTCGAATACAGACT	41	GACCAGAGAATTCGAATACAGATG	57
GACCAGAGAATTCGAATACAGCCT	42	GACCAGAGAATTCGAATACAGCTG	58

10

20

30

40

【表 1 - 2】

GACCAGAGAATTTCGAATACAGGCT	43	GACCAGAGAATTTCGAATACAGGTG	59
GACCAGAGAATTTCGAATACAGTCT	44	GACCAGAGAATTTCGAATACAGTTG	60
GACCAGAGAATTTCGAATACATACT	45	GACCAGAGAATTTCGAATACATATG	61
GACCAGAGAATTTCGAATACATCCT	46	GACCAGAGAATTTCGAATACATCTG	62
GACCAGAGAATTTCGAATACATGCT	47	GACCAGAGAATTTCGAATACATGTG	63
GACCAGAGAATTTCGAATACATTCT	48	GACCAGAGAATTTCGAATACATTTG	64

10

【0268】

単一の増幅配列を伴う単一のアダプター群（すなわち、ユニバーサル増幅配列、試料特異的コード、およびランダムタグのセット；また、多機能性アダプターモジュールとも称する）の、ゲノム断片の両方の末端への適用は、その2つの末端において、同じゲノム断片に独立にタグ付けするという事実を含む、いくつかの顕著な利点を有する。以下のいくつかの節において記載する通り、所与の任意の断片による2種の鎖は、最終的には互いから分離され、本発明では独立の分子として挙動するであろう。したがって、同じ断片の2つの末端における2種の異なるタグの存在は、本発明の欠点ではなく、利点となる。加えて、アダプター間のライゲーション事象は、初期目標が異種末端を伴うアンプリコンを創出することである、次世代ライブラリーの構築において大きな問題であるという事実もある。本発明の方法を使用する場合、方法でこの非対称性を導入するのは、工程の後半であり、したがって、本発明では、同一な末端も許容可能である。本方法の未見で驚くべき利益は、本発明のライブラリー構築法では、アダプター二量体が観察されないことである。理論により束縛されずに、本発明者らは、これが、急速に形成された、希少なアダプター二量体分子種は、PCR増幅に必要な変性ステップおよびアニーリングステップにおいて、緊密なヘアピン構造を形成するためでありうることを想定するが、これらのヘアピン構造が、プライマーに方向付けられたさらなる増幅に対して完全に耐性であることもさらに想定される。アダプター二量体非含有ライブラリーを作製する可能性は、単一細胞～少数細胞によるゲノム解析、循環DNA解析（胎児診断法、組織移植拒絶の監視、またはがんスクリーニングへの適用）、または単一細胞のトランスクリプトーム解析など、極低投入量の適用において重要な技術的特色である。本方法はそれ自体、このような適用において、顕著な有用性を提供する。単一プライマーによるアンプリコンのさらに顕著な特色は、25ヌクレオチドのPCRプライマーによる増幅を「オン」にし、より長い58ヌクレオチドのプライマーによる増幅を「オフ」にすることが可能なことである。これについては、下記の6-5節でより詳細に記載し、本発明に対する重要性についても強調する。

20

30

【0269】

概説

単一のユニバーサル増幅配列を、標的断片の両方の末端において使用するアダプター戦略により、アダプター二量体に関する問題が解消される。これは、一例として、実施例3：単一アダプターによるゲノムライブラリーの構築において明確に実証される。

40

【0270】

3節：ライブラリーによる定量化

ゲノム解析戦略のための本方法のさらなる側面は、「カバレッジ深度」が既知である、すなわち、ライブラリー内に存在する平均ゲノムコピー数が既知であるか、またはこれを決定しうることである。カバー深度は、後続のステップに必要なライブラリーのバルク増幅の前に、精製されたライゲーション反応物を使用して測定する。例示を目的として述べると、50ゲノム相当のDNAを、本発明の実施形態のライブラリースキームに投入し、アダプターの、断片の両方の末端へのライゲーションが100%の効率でなされる場合、各アダプター末端が他のアダプター末端とは独立に作用し、よって、2末端×50ゲノム

50

=100カバレッジであるため、カバレッジ深度は100である。本明細書で想定されるユニバーサルPCRプライマーでは、アダプター二量体は増幅されず、両方の末端においてアダプター処理された断片が増幅されるという単純な事実は、ライブラリーによる定量化が、単に、ユニバーサルプライマーを使用する定量的PCR (qPCR: quantitative PCR) によりライブラリーの複雑性を測定し、結果をカバレッジ深度が既知である基準物質に対して較正する問題となることを意味する。本明細書では、「ゲノムコピー」および「カバレッジ深度」という語句は、同じ事柄を意味し、互換的に使用することができる。本方法は、4~1000、好ましくは20~100倍のカバレッジ深度を、次の段階 (phase) である、本発明に従う試料プロセッシングに送り込む。

【0271】

10

4節：ライブラリーによる増幅

特定の実施形態では、アダプターをライゲーションされた、20~100倍のカバレッジ深度に相当するゲノム断片ライブラリーの一部を、増幅を駆動する単一のユニバーサルプライマー配列による、標準的なPCR法を使用して増幅することになる。特定の実施形態では、この段階 (stage) において、初期ライブラリー内の数ピコグラムの材料を、10,000倍の増幅を含意する、数マイクログラムの増幅された材料に転換することが有利である。

【0272】

5節：標的ライブラリー断片の、捕捉プローブとのハイブリダイゼーション

オリゴヌクレオチド合成化学における進歩により、洗練されたゲノム捕捉戦略のための新たな機会が創出されている。特に、現在では、1塩基当たりの合成費用が妥当であり、収率が比較的高く、塩基精度が優れた、長いオリゴヌクレオチド (100~200ヌクレオチドの長さ) が、様々な販売元から市販されている。この能力により、本発明者は、多機能性捕捉プローブを創出する (図2)。多機能性捕捉プローブの例示的な例の要素は、以下を含む。

20

領域1は、全てのプローブに共通な34ヌクレオチドの領域であって、修飾された相補的なオリゴヌクレオチド (また、パートナーオリゴヌクレオチドとも称する) とハイブリダイズする領域を含む。この修飾オリゴヌクレオチドは、5'端において、ストレプトアビジンタンパク質に緊密に結合することが可能なビオチンであって、ビオチン結合に対する立体障害を緩和する、長い親水性のスペーサーアームである、ビオチン-T E G修飾もさらに含む。3'端において、オリゴヌクレオチドは、このパートナーオリゴヌクレオチドをプライマーの伸長に対して不活性とする、ジデオキシシトシン残基で終結する。プローブデザインのこの要素は、前記プローブを直接修飾することなく、非限定的な数のプローブを、ビオチンによる捕捉機能性でアダプター処理することを可能とする。

30

【0273】

領域2は、標的特異的であり、gDNA断片分子と相互作用する、特注の60ヌクレオチドの領域を含む。ゲノム内の配列が独特のものであること、結合効率を損なう共通なSNPの存在、および二次構造についての検討を促すこの領域は、コンピュータによる方法を介してデザインする。

【0274】

40

領域3は、後続の断片増幅において、PCRプライマー結合性部位として用いられる、20ヌクレオチドのセグメントを含む。この特色については、以下の段落においてさらに詳細に記載する。

プローブの多重性を使用して、対象のゲノム領域を捕捉することができる (プローブのマルチプレックス化)。少なくとも2種のプローブを援用して、典型的な、100~150bpの長さのコードエキソンを完全にクエリーすることができる。一例として、これにより、20種のプローブを使用して、典型的な10種のエキソン遺伝子を捕捉し、合計2000種のプローブを使用して、100種の遺伝子パネルを精査することになることが指し示される。ゲノムライブラリー断片の、プローブとのハイブリダイゼーションは、熱変性に続く再アニーリングにより実行することができる。一実施形態では、ステップは、

50

以下を含む。

【0275】

1. ゲノムライブラリー断片を、プールされたプローブ配列（この場合、「プローブ配列」とは、個々のプローブの、等モル量の、高度に修飾されたパートナーオリゴヌクレオチドとの組合せを指す）と、プローブ1種に対する標的1種～プローブ1,000,000種に対する標的1種の範囲にわたる具体的な標的対プローブ比で組み合わせるステップ。一実施形態では、最適の比は、プローブ10,000種に対する断片約1種である。

【0276】

2. 組み合わせた断片+プローブを、1MのNaCl、10mMのトリス、pH8.0、1mMのEDTA、および0.1%のTween 20（非イオン性の洗浄剤）を含有する溶液中で、>30秒間にわたり95℃まで加熱して、全ての二本鎖DNA構造を変性させるステップ。

【0277】

3. 組み合わせたプローブおよび断片を、段階的に制御して、例えば、2分ごとに温度1℃ずつ、<60℃までの低下により冷却するステップ。この緩徐な冷却により、標的ゲノム断片とプローブ配列との二重鎖が結果としてもたらされるであろう。

【0278】

4. プローブ：断片複合体を、カルボキシルでコーティングされ、ストレプトアビジンで修飾された常磁性ビーズに結合させ、これらのビーズを、強力な磁石を使用して「プルアウト」するステップ。

【0279】

5. 結合させた複合体を、25（v/v）%のホルムアミド、10mMのトリス、pH8.0、0.1mMのEDTA、および0.05%のTween 20を含有する溶液で洗浄するステップ。特定の実施形態では、洗浄ステップを、少なくとも2回にわたり実行する。

【0280】

6. 洗浄されたビーズを、後続の酵素プロセッシングステップに適する溶液中に再懸濁させるステップ。

【0281】

捕捉反応

捕捉反応の実施形態は、実施例3（単一アダプターによるゲノムライブラリーの構築）において提示し、実施例5（PLP1-qPCRアッセイの検証）において展開され、さらに記載される、qPCRアッセイを援用する。

【0282】

6節：ハイブリダイズさせたプローブ：標的複体の酵素プロセッシング

当技術分野で現在実行されている通り、ハイブリダイゼーションベースの配列捕捉法は一般に、最適未満の標的配列の濃縮を結果としてもたらず。文献および市販の刊行物からは、リードのうちで、それらの意図される標的配列にマップされるのは、せいぜい約5%～10%であると推定することができる。残りのリードは、意図される標的の近傍にマップされることが多く、市販品の販売元は、「的中」を、意図される遺伝子座から約100塩基以内に収まるリードと定義し直している。この「スプレッディング」効果の理由は、完全に理解されてはならず、正規の配列ハイブリダイゼーションイベントの結果である可能性が高い（例えば、図3を参照されたい）。

【0283】

本明細書で想定される複体の酵素プロセッシングは、捕捉された配列を、正確な対象領域に、より鋭利にフォーカスする。このステップでは、3'-5'エキソヌクレアーゼ活性もまた保有するDNAポリメラーゼを援用する。このような酵素の例示的な例は、T4 DNAポリメラーゼである。この酵素は、ダングリングテイル配列を、プローブと標的配列との間で形成された二重鎖領域から「齧り取る」であろう。次いで、T4 DNAポリメラーゼは、プローブ上のテイルセグメントをコピーするであろう。例えば、図4を参

照されたい。このステップによりもたらされる利益は、以下を含むがこれらに限定されない。

【0284】

1. この種の酵素プロセッシングを援用することにより、プローブと直接にハイブリダイズして二重鎖化された断片だけを先に進める。最終的なシーケンシングライブラリーは、断片およびプローブの両方から得られる分子のキメラ（ハイブリッド）セットである。

【0285】

2. プローブは、鎖特異的であり、したがって、捕捉された標的は、プローブに照らした独特の有向性を有する（図5で例示する）。これは、単一の断片から作製された2種の鎖のうち的一方だけが、プローブと相互作用し、プロセッシングは、リードを、プローブ配列の5'領域に「フォーカス」することを意味する。この点で、断片の相補鎖は、完全に独立の分子種となる。有向プローブを、標的領域（例えば、エキソン）の片側に配置することにより、技術は、標的領域上のシーケンシングリードの高度に特異的なフォーカシングを可能とする（図6）。

【0286】

3. 標的断片と正規に交差ハイブリダイズした（しかし、プローブとは交差ハイブリダイズしなかった；図3）標的分子は、不可欠のプローブ配列を取得せず、したがって、後続の増幅ステップでは失われる。

【0287】

4. プローブの実際の「テイル」配列は、増幅配列の一部として、標的断片にコピーされる。全ての市販の実行可能なシーケンシングプラットフォーム（例えば、Illumina製の、可逆性ターミネーター化学反応によるシーケンシングプラットフォーム）は、標的断片が非対称末端を有するシーケンシングライブラリーを必要とし、これは、「フォワード」アダプター配列および「リバース」アダプター配列と称するか、またはシーケンシングラボの略称では、「P1」および「P2」と称することが多い。特定の実施形態では、この時点までに、本明細書で想定される断片ライブラリーは、末端において単一の分子種を有し、「P1」と判定される。酵素プロセッシングステップは、2つの事柄を達成する。第1に、酵素プロセッシングステップは、これらのP1末端のうちの1つを「消失させる」（3'-5'エキソヌクレアーゼ活性により）。第2に、酵素プロセッシングステップは、P1とは異種のP2末端の基部を「付加する」（DNAポリメラーゼによるプローブテイル配列のコピーを介して）。

【0288】

5. 正規のP1-P2末端で酵素的に修飾された標的分子は、プロセッシングに後続するPCR増幅ステップにおいて、選択的に濃縮することができる。これは、長いPCRプライマーの使用により達成する。特に、長いプライマーは、次世代シーケンシングに必要とされる十分な機能性を付加するために必要であり、また、増幅に選択性も付与する。増幅の第1ラウンドから得られる「夾雑物」である、残留P1-P1ライブラリー断片は、長いP1プライマーでは増幅されない。これは、本方法の顕著な利点である。初期P1-P1ライブラリーは、単一の、25ヌクレオチドのPCRプライマーで効果的に増幅される。このプライマーの長さが、57ヌクレオチドまで伸長し（シーケンシングの機能性が付加され）たら、これらの同じP1-P1分子は、いかなる程度でも増幅されない。したがって、初期ライブラリーの増幅は、25ヌクレオチドのプライマーで「オン」にし、57ヌクレオチドのプライマーで「オフ」にすることができる。

【0289】

概説

P1-インサート-P1ライブラリーが増幅されないことは、実施例3（単一アダプターによるゲノムライブラリーの構築）で実証される。P1-インサート-P2のプロセッシングされたDNA断片の優先的な増幅を、実施例3（単一アダプターによるゲノムライブラリーの構築）に示す。実施例3では、プロセッシングに伴う標的特異度の実質的な改善についてもさらに実証する。最後に、プロセッシングされた初期複合体のパーセントを意味す

るプロセシングの「感度」は、捕捉された全ての複合体のうちの10%のオーダーにあることを、実施例9（捕捉後プロセシングの直接的な測定）において実証する。

【0290】

7節：増幅およびシーケンシング

初期概念実証実験に適用されるコアアダプター配列およびプライマー配列を表2に示す。ステップ6から得られた、酵素プロセシングされた複合体を、全長フォワードPCRプライマーおよび全長リバースPCRプライマーを含有する、PCR増幅反応物に直接付加する。増幅の後、ライブラリーを、精製し、定量化し、ハイスループットの次世代シーケンサー上にロードすることができ（この実施形態では、ライブラリーを、Illuminaによる可逆性ターミネーターベースのプラットフォーム用に構成する）、約数百万種の断片の配列を決定する。この段階では、>36ヌクレオチド、好ましくは72または100+ヌクレオチドの長さの単一のリードを観察することができる。

【0291】

【表2-1】

表2: プライマー配列の出自

説明	配列(5'→3')	配列番号
全長ライブラリーのためのフォワード増幅プライマー。	AATGATACGGCGACCACCGA	65
Illumina製のペアエンド(PE)フローセル上に存在するフォワードプライマー配列。	AATGATACGGCGACCACCGAGATCTACAC	66
4つの塩基全てによる非回文状4マー。これは、リード1シーケンシングプライマーを、Illumina基準配列と同じ長さおよび位置となるように付加する。	GTCA	
ACA2。PCRプライマーであるACA2およびCAC3は、NuGENに先立って存在する。	TGCAGGACCAGAGAATTCGAATACA	67
ACA2フォワードシーケンシングプライマー(FSP):リード1	ACACGTCATGCAGGACCAGAGAATTCGAA TACA	68
ACA2全長フォワード増幅	AATGATACGGCGACCACCGAGATCTACAC	69

【表 2 - 2】

プライマー(FLFP)。最終長=58nt。	GTCATGCAGGACCAGAGAATTTCGAATACA	
全長ライブラリーのためのリバース増幅プライマー。	CAAGCAGAAGACGGCATAACG	70
PEフローセルリバースプライマー	CAAGCAGAAGACGGCATAACGAGAT	71
これらは、リード2プライマー部位を創出するように挿入されたTruSeqリバースアダプターから得られる9つの5'塩基である。	GTGACTGG	
CAC3リバースプライマー	CACGGGAGTTGATCCTGGTTTTTCAC	72
CAC3リバースシーケンシングプライマー(RSP):リード2	GTGACTGGCACGGGAGTTGATCCTGGTTTT CAC	73
CAC3全長リバースプライマー(FLRP)。最終長=57nt。	CAAGCAGAAGACGGCATAACGAGATGTGAC TGGCACGGGAGTTGATCCTGGTTTTTCAC	74

【0292】

8 節：データ解析

シーケンシング後データ解析には、少なくとも2つの主要な側面がある。第1の側面は、配列変異体（参照配列の確立されたセットに照らした、一塩基変異体、微細挿入および／または微細欠失）の同定である。複雑ではあるが、当技術分野では、これらの方法が十分に記載されており、当業者であれば、このような方法を理解するであろう。第2の側面は、標的化シーケンシングデータから得られるコピー数変異の決定である。

【0293】

(実施例2)

コピー数の決定

コピー数の決定は、DNAシーケンシングの分野において様々に使用される。非限定的な例として、大量パラレルDNAシーケンシング技術は、生物学的試料を精査および解析する、少なくとも2つの機会をもたらす。十分に確立された1つの側面は、試料中に存在するデノボ配列を意味するDNA配列の決定（例えば、新たに単離された微生物のシーケンシング）または変異体のための、既知の領域のリシーケンシング（例えば、既知の遺伝子内の変異体の検索）を意味するDNA配列の決定である。大量パラレルシーケンシングの第2の側面は、定量的生物学および特定の配列に遭遇する回数をカウントする可能性である。これは、カウンティングを使用して、それぞれ、遺伝子発現または特定のタンパク質のゲノムDNAとの関連を推量する、「RNA-seq」および「CHIP-seq」などの技術の根本的な側面であろう。本実施例は、DNAシーケンシングの定量的でカウンティングベースの側面に関する。

【0294】

DNA断片は、高度な類似性を共有する配列の配座としてカウントされる（すなわち、DNA断片は、公知のゲノム配列の特異的領域と整列する）ことが極めて多い。これらのクラスター内の配列は、同一であることが多い。a) 異なる開始DNA配列および終結DNA配列のリード、またはb) セット内の他のリードから得られる高品質の配列差を伴うDNA配列はしばしば、「独特のリード」と考えられることに注目されたい。したがって、異なる開始配列位置および配列変異は、クローンから得られる独特のイベントを差別化するのに使用される「タグ付け」の一形態である。本実施例ではまた、ライブラリー構築

のコースにおいて、ゲノム断片に、ランダムヌクレオチドタグ（例えば、ランダム6ヌクレオチド配列）も導入する。1）ランダムヌクレオチドタグ配列を、2）DNAシーケンシングリードの開始点、および3）リードの実際の配列と組み合わせることにより、タグを集合的に構成する。このタグにより、同じ断片が2回にわたりクロニングされた収束イベント（このような断片は、ライブラリー構築において導入された、異なるランダムヌクレオチドタグ配列を有するであろう）と、ライブラリーによる増幅において複製された同じ由来の断片（これらの「クローン」は、同じランダムヌクレオチドセグメント、および同じクローン開始点を有するであろう）とを差別化することが可能となる。この種のタグ付けはとりわけ、ゲノムDNAについての定量的解析もさらに可能とするが、より一般には、DNA分子（例えば、RNA-seqライブラリー）についての定量的解析もさらに可能とする。

10

【0295】

ランダムヌクレオチドタグ（DNAクローン末端と組み合わせたランダムNマー）を、DNAシーケンシングライブラリーに導入することにより、理論的には、ライブラリー内の各独特のクローンを、その独特のタグ配列により同定することが可能となる。「理論的には」と明記することにより、シーケンシングにおけるエラー、ライブラリーによる増幅において導入されるエラー、他のライブラリーからの夾雑クローンの導入など、生じる通常の実験データセットの交絡特徴を認識する。これらの交絡源の全ては、本明細書で提起される理論的検討を交絡させる。配列捕捉および標的化リシーケンシングの文脈では、ライブラリーをタグ付けすることにより、捕捉されたライブラリー内の遺伝子座コピー数についての定量的解析を可能とすることができる。

20

【0296】

非限定的な例として、男性対象から創出された、100二倍体ゲノム相当の投入量から構築されるライブラリーについて検討しよう。予測では、各常染色体の遺伝子座において、約200種のライブラリークローンが存在し、各X染色体遺伝子座において、100種のクローンが存在するであろう。常染色体領域を捕捉し、2000回にわたりシーケンシングすれば、200種のタグの全てに、99%の確実性を超える信頼区間で遭遇するであろう。X染色体領域では、理論的には、2000種のリードは、合計100種のタグを明らかにするであろう。例示として、本実施例では、DNAシーケンシングライブラリー内でDNAタグを創出することにより、コピー数の差違を保存するという一般的概念が裏付けられる。この一般的枠組みを、本明細書で記載される方法に適用することができる。経験的証拠は、遺伝子座ごとのベースのクロニング効率の差違について、実験エラーなどからのアーチファクトタグの散在的な導入について調整を行う必要があることを示唆する。この概念の実際的な実装は、異なる文脈において異なる可能性があり、事例ごとの配列解析法を伴いうるが、本明細書で概括される一般的原理は、このような適用の全てに通底するであろう。

30

【0297】

現時点まで、タグ付けされたDNAライブラリーの創出は、ゲノムDNA解析の文脈において検討されているが、この概念は、全てのカウンティングベースのDNAシーケンシング適用に当てはまることを強調しなければならない。特定の実施形態では、タグ付けは、mRNA試料から作製されるcDNA分子を、タグを創出する方法によりクロニングする、RNA-seqに適用することができる。このような手法は、配列ベースの遺伝子発現解析の忠実度を実質的に増大させる。ある特定の実施形態では、タグ付けにより、クロマチン免疫沈降（CHIP（chromatin immunoprecipitation）-seq）実験の分解能を増大させることが想定される。多様な実施形態では、タグ付けにより、マイクロビーム区画内および環境試料内の微生物の存在および存在量を決定するのに使用される、配列カウンティングの定量的側面が増強されるであろう。

40

【0298】

（実施例3）

50

単一アダプターによるゲノムライブラリーの構築

目的

本実施例の目標は、音響処理により断片化された P r o M e g a による女性 h g D N A (約 2 0 0 b p) から、ゲノム D N A ライブラリーを創出することであった。

【0299】

概要

結果は、アダプターをデザインするための本方法の顕著な特色を明確に実証した。特に、アダプター単独のライゲーション反応は、検出可能なアダプター二量体分子種を存在させなかった。本方法と同様、投入量の限界は、アダプター二量体のバックグラウンドレベルにより不変的に決定されるので、極低投入量のシーケンシングライブラリー調製技術の文脈では、これは、極めて重要であった。アダプター二量体の夾雑に対する点検を維持しようとする試みでは、高度に特化された技術が適用されている。これらは、カラムまたはゲル精製などのサイズ除外法、アダプターの自己ライゲーションイベントを最小化するようにデザインされた、高価な特注のオリゴヌクレオチド修飾、およびライブラリー構築の後における、制限消化によるアダプター二量体の破壊を可能とする、アダプター配列修飾を含む。

【0300】

本明細書で想定される、単純な、単一アダプター、単一プライマーの概念により、D N A 構造原理の基本原則を喚起する単純な溶液に伴うアダプター二量体問題に取り組む。この極低投入量の技術は、ゲノム解析のためのゲノムライブラリーを構築するのに有用であり、例えば、1種または数種の特異的な細胞に対する R N A - s e q 適用における、クローニングされた二本鎖 c D N A についてのトランスクリプトーム解析にも有用であり、高度に修飾され、保存不良の、ホルマリン固定され、パラフィン包埋された (F F P E) 核酸試料中に存在するいくつかのインタクトな断片をレスキューするのに有用であろう。

【0301】

本発明のアダプターデザインの別の不可欠の特色は、異なる P C R プライマーの長さを使用することにより、標的アンプリコンライブラリーの P C R 増幅を「オン」および「オフ」にする能力である。明確に実証された通り、ライブラリーによる増幅に最適のプライマーの長さは、推定 T m (標準的なイオン強度条件下における) を $\geq 55^{\circ}\text{C}$ とする、25ヌクレオチドのプライマー分子種であった。短く、T m の低いプライマーは、アンプリコンの低効率の増幅を提示し、平均インサートサイズが小さなアンプリコンに好都合であると考えられた。異種配列の逆向きのプライマーと対合させる場合は、このサイズクラスのプライマーが良好に働くという先例は多くみられる。

【0302】

まとめると、これらのデータは、本発明のアダプターおよび P C R 増幅法により、「チューニング可能な、オン/オフ」の増幅特性を伴う、アダプター二量体非含有断片ライブラリーが作製されることを実証した。

【0303】

方法

I D T から受領したプライマーを、T E z e r o (10 mM の トリス、p H 8. 0、0. 1 mM の E D T A) 中に 100 μM まで水和させた。

【0304】

断片の修復：

- 14 μl の水
- 5 μl の h g D N A
- 10 倍濃度の末端修復緩衝液 2. 5 μl
- 1 mM の d N T P 2. 5 μl
- 1 μl の末端修復酵素と、0. 5 μl の P r e C R 酵素修復ミックスとを、混合および添加したもの

を組み合わせることにより、融解させた g D N A および 500 n g の g D N A を末端修復

10

20

30

40

50

した。

【0305】

混合物を、20℃で30分間にわたり、かつ、70℃で10分間にわたりインキュベートし、10℃で保持した。

【0306】

アダプターのアニーリング：68μlのTE zeroと、5MのNaCl 2μlと、20μlのオリゴ11と、10μlのオリゴ12とを組み合わせた。10秒間にわたり95℃まで加熱し、65℃で5分間にわたり加熱し、RTまで冷却した。

【0307】

【表3】

10

表3: アダプター配列であるAdap_P23およびAdapt_L23

配列番号	オリゴID	名称	配列	説明
75	11	Adap_P23	TCCGTGTATTCGAAT	64の初期アダプターセットの番号23のパートナー鎖
23	12	Adap_L23	GACCAGAGAATTCGA ATACACGGA	64の初期アダプターセットの番号23のライゲーション鎖

20

【0308】

ライゲーション：20μlの総容量：

- 13μlまたは8μlの水
- 0または5μlの末端修復された断片=100ng。
- 10倍濃度のT4リガーゼ緩衝液2μl
- 50%のPEG8000 3μl
- 10μMのACA2アダプター23の二重鎖1μl
- 1μlのT4 DNAリガーゼを混合および添加したもの

中に、1=インサートを伴わないhgDNA、2=100ngの末端修復されたhgDNAを組み合わせた。

30

【0309】

23℃で30分間にわたり、かつ、65℃で10分間にわたりインキュベートした。反応1回当たり80μlのTEzおよび120μlのビーズを添加した。混合し、RTで10分間にわたりインキュベートした。70%のEtOH：水（v/v）のアリコート200μlで、2回にわたり洗浄し、50μlのTEz中に再懸濁させた。

【0310】

PCR増幅：ライゲーションミックスのアリコート各10μl=20ngずつのライブラリー。18サイクルにわたり増幅するように計画した。

【0311】

40

【表 4】

表4: 反応混合物

	100 μ l	600 μ l
水	50	300
10X STD Taq 緩衝液	10	60
25 mM MgCl ₂	10	60
10 μ M PCR プライマー	10	60
鋳型	10	60
DMSO	5	30
10 mM dNTPs	5	30
Taq ポリメラーゼ	1	6

10

【0 3 1 2】

プライマーおよび鋳型を除く全ての成分を含有する600 μ lのミックスを作製した。80 μ lのアリコート6つを作製した。10 μ lのインサートを伴わないライゲーションミックスを、セット1に添加し、10 μ lのhg DNAインサートを、セット2に添加した。下記に示される対合したプライマーのうちの10 μ Mのプライマーを、インサートを伴わないライゲーションミックスおよびhg DNAインサートによるライゲーションミックスに添加した。混合した。94 $^{\circ}$ Cで30秒間、60 $^{\circ}$ Cで30秒間、および72 $^{\circ}$ Cで60秒間の18サイクルにわたりサーマルサイクリングさせ、最後に72 $^{\circ}$ Cで2分間にわたりサーマルサイクリングさせ、10 $^{\circ}$ Cで保持した。

20

【表 5】

表5: プライマー配列

配列 番号	オリゴ ID	名称	配列
76	3	ACA2_20	GACCAGAGAATTCGAATACA
77	5	ACA2	TGCAGGACCAGAGAATTCGAATACA
78	8	ACA2_FLFP	AATGATACGGCGACCACCGAGATCACACGTCATG CAGGACCAGAGAATTCGAATACA

30

【0 3 1 3】

PCR産物を、120 μ lのビーズで精製した。70%のEtOH 200 μ lで、2回にわたり洗浄した。ビーズを乾燥させ、DNAを、50 μ lのTE_zで溶出させる。5 μ lの各試料を、2%のアガロースゲル上で解析した。

40

【0 3 1 4】

結果

全く同じゲル画像を、4つの異なる色およびコントラストスキームで、図7に示す。ゲル上にロードした試料は、

1. ACA2_20により増幅される、インサートを伴わない、アダプターだけのライゲーションミックス
2. ACA2 (通常の25ヌクレオチドのPCRプライマー)により増幅される、インサートを伴わない、アダプターだけのライゲーションミックス
3. ACA2_FLFP (全長フォワードプライマー)により増幅される、インサート

50

を伴わない、アダプターだけのライゲーションミックス

4. A C A 2 20により増幅される、約200bpのhgDNAインサート+アダプターによるライゲーションミックス20ng

5. A C A 2 (通常の25ヌクレオチドのPCRプライマー)により増幅される、約200bpのhgDNAインサート+アダプターによるライゲーションミックス20ng

6. A C A 2 F L F P (全長フォワードプライマー)により増幅される、約200bpのhgDNAインサート+アダプターによるライゲーションミックス20ngであった。

【0315】

アダプター単独のライゲーション→PCR産物中(レーン1~3)では、材料が増幅されないことが明らかであった。短い、20ヌクレオチドのA C A 2プライマー(レーン4)は、「通常の」、25ヌクレオチドのA C A 2プライマー(レーン5)と比べて非効率的な増幅を示した。58ヌクレオチドのA C A 2 F L F Pプライマー(レーン6)では、材料のごくかすかな痕跡が目視可能であるに過ぎなかった。

【0316】

さらなる実施形態では、A C A 2プライマーの量で滴定し、収率をモニタリングすることが有用でありうる。通常の高収率のPCRプライマーは、合計2μMのプライマー(100μlのPCR反応物当たり)につき、フォワードプライマーおよびリバースプライマーの両方を1μMずつ保有する。したがって、A C A 2を2μMまで(フォワードプライマーおよびリバースプライマーの両方であるので)添加することにより、収率を増大させることができる。同様に、特定の実施形態では、60℃より低いプライマーアニーリング温度において、ライブラリーの増幅特徴をモニタリングすることが有用でありうる。

【0317】

(実施例4)

gDNAの断片化

目的

初期の原理実証実験のために、男性および女性から得られた、シアリングされたヒトgDNAを必要とした。本例では、Promega製のヒト女性およびヒト男性のgDNAを援用する。チューブ上に示される量に基づき、これらを、100ng/μlのDNA 1000μlまで希釈し、200bpの範囲の断片を作製することが意図される、Covaris条件下に置いた。

【0318】

概要

実験室の研究基盤には、少なくとも2つの構成要素が存在する。一方は、DNAを定量化する能力であり、他方は、ゲル上のDNAのサイズ分布を可視化する能力である。本実施例では、Life Technologies製のQubit 2.0測定器を援用して、DNA濃度を測定した。記録された読取り値は、Nanodropによる本発明者らのかつての実験より一般に小さいことが分かった。Qubitによる読取り値は、dsDNA特異的な色素結合および蛍光に基づいた。Qubitの1つの主要な利点は、あらかじめのクリーンアップを伴わずに、DNA増幅反応(例えば、PCR)を定量化するのに使用しうることである。これらの実験では、100ng/μlであると考えられたPromega gDNAが、Qubitでは、約60ng/μlと測定されることが分かった。ゲルおよびサイズ分布についての定性的評価に関しては、電気泳動が施され、システムは効果的に働いたという記録がなされている。本実施例では、断片化されたgDNAは、所望の約200bpを中心とする平均サイズ分布を有することが分かった。

【0319】

方法および結果

Covaris処理の後、Qubit測定器を使用して、DNA濃度を測定した。gDNAを10倍に希釈し、2μlを、200μlの最終容量のアッセイ溶液に添加した。女性試料および男性試料のいずれについての読取り値も、約60ng/mLと記録されたが

、これは、開始溶液が、 $60\text{ ng}/\mu\text{ l}$ であることを意味する。これは、当初の予想を下回ったが、十分に、特定の実施形態に適切な範囲内にはあった。次いで、本発明者らは、断片化前および断片化後両方の材料 $2\mu\text{ l}$ (120 ng) および $5\mu\text{ l}$ (300 ng) を、2%のアガロースゲル上にロードした(図8)。上の行の記号は、M: 男性のgDNA およびF: 女性のgDNAを表す。下の行の記号は、U: 断片化されていない、およびC: Covarisにより断片化されたである。1つの重要な観察は、平均断片サイズが、 200 bp 近傍を中心とする均等分布であることであった。

【0320】

(実施例5)

PLP1 qPCRアッセイの検証

10

目的

X染色体上のプロテオリピドタンパク質1 (PLP1) 遺伝子を、初期概念実証捕捉研究のために検討した。この遺伝子は、がんに関与性であり、X染色体上に存在し、これは、男性と女性との間で天然のコピー変異を有することを意味するために選択した。PLP1のRefSeq転写物であるNM000533. 3の、187ヌクレオチドのエキソン2領域を、標的領域として使用した。原理実証研究のために、PLP1エキソン2内およびPLP1エキソン2近傍の領域を、qPCRによりモニタリングする能力が必要とされた。本実施例は、8つのこのようなアッセイのデザインおよびバリデーションについての記載を提示する。

【0321】

20

概要

PLP1エキソン2の捕捉をモニタリングするように、8種のqPCRアッセイ(この場合、単純なプライマー対を意味する)をデザインした。5種のアッセイが的中したことは、それらが、捕捉プローブにより標的化される領域内にあることを意味する。2種が「標的の近傍」にあることは、1種のアッセイが、標的領域から 200 bp のゲノム座標に位置し、1種のアッセイが、逆向きの鎖上の標的領域から 1000 bp のゲノム座標に位置することを意味する。これらの2種のアッセイは、捕捉実験において、標的遺伝子座の近傍の領域が、「ヒッチハイカー」として引き回される現象である「スプレディング」を定量化するようにデザインした。最後に、9番染色体領域に対しても、1種のアッセイをデザインしたが、これは、ヒトgDNAの任意の非類縁セグメントをモニタリングするようにデザインしている。本明細書では、例は、8種のアッセイ全てにより、予測されたサイズのアンプリコンと符合するPCR断片が作製されることを示す。例は、X染色体上に位置するPLP1アッセイにより、投入されたgDNA 1 ng あたりの比活性は、男性より女性において大きいことが適切に示された。これらのデータにより、これらのアッセイの、gDNAの捕捉をモニタリングするさらなる実験における使用の妥当性が検証された。

30

【0322】

方法、結果、および考察

PLP1エキソン2の近傍を中心とする 400 bp の領域を、 $80\sim100\text{ bp}$ の長さであるアンプリコンを作製するために、平均24ヌクレオチドの長さであり、 T_m が $60\sim65\text{ }^{\circ}\text{C}$ であるプライマーのためのプライマー3に切り出した。検索領域を操作して、エキソン2の5'イントロン-エキソン間境界部から、CDSを通して、3'エキソン-イントロン間境界部に「ウォーキングする」プライマー対(qPCRのためのアンプリコン)を得た。この近くでまた、エキソン2に対して遠位であり、エキソン3に向かい、エキソン2から約200ヌクレオチド~約1000ヌクレオチドに配置された、近傍捕捉アッセイもデザインした。これらは、副次的ハイブリダイゼーションイベントにおいて捕捉される「ヒッチハイカー」ゲノム断片をモニタリングするのに使用されるであろう。最後に、9番染色体上で、実験におけるバルクのゲノムDNAレベルをモニタリングする1種のアッセイも創出した。これらのアッセイのプライマー配列を下記に示し、詳細は本実施例の末尾の付録とする。

40

50

【 0 3 2 3 】

【表 6 - 1】

表6: PLP1エキソン2の5'イントロン-エキソン間境界部からCDSを
通って「ウォーキング」するqPCRプライマー配列

アッセイ	配列 番号	名称	配列	説明
アッセイ1	79	F1	TTAGAGTGCTGTGCAAGATGTCTG	PLP1エキソン2内に位置する フォワードqPCRプライマー
	80	R1	ACCCCAAAGAAACACAATCCAGT	PLP1エキソン2内に位置する リバースqPCRプライマー
アッセイ2	81	F2	GCCACTGGATTGTGTTTCTTTG	PLP1エキソン2内に位置する フォワードqPCRプライマー
	82	R2	TCAATTAGCTTTTCTGTGCCAGTG	PLP1エキソン2内に位置する リバースqPCRプライマー
アッセイ3	83	F3	ACTGTTCTGTGGCTGTGGACAT	PLP1エキソン2内に位置する フォワードqPCRプライマー
	84	R3	TCTTGGTAGTTTTTGGAGAAATAGGTC	PLP1エキソン2内に位置する リバースqPCRプライマー
アッセイ4	85	F4	ttttttccccagGCTTGTTAGA	PLP1エキソン2の 上流領域内に位置し、これを挟む フォワードqPCRプライマー(小文字)
	86	R4	CACCCCAAAGAAACACAATCCAGT	PLP1エキソン2の 上流領域内に位置し、これを挟む リバースqPCRプライマー(小文字)

10

20

30

【表 6 - 2】

アッセイ5	87	F5	CCCTCACTGGCACAGAAAAGCTAA	PLP1エキソン2の下流領域内に位置し、これを挟むフォワードqPCRプライマー(小文字)
	88	R5	tgaggagggcaggtacttacACATT	PLP1エキソン2の下流領域内に位置し、これを挟むリバースqPCRプライマー(小文字)
アッセイ6	89	F6	CCCCTTGTTTTCTTACACGTGTTCT	PLP1エキソン2の200bpの下流のフォワードqPCRプライマー
	90	R6	CCTCCCTTGGCTTCTCCATACCTA	PLP1エキソン2の200bpの下流のリバースqPCRプライマー
アッセイ7	91	F7	GTGTGTCATTGTTTGGGAAAATGG	PLP1エキソン2の1000bpの下流(実際にはエキソン3内)のフォワードqPCRプライマー
	92	R7	CACCCCTTGTTATTGCCACAAAAT	PLP1エキソン2の1000bpの下流(実際にはエキソン3内)のリバースqPCRプライマー
アッセイ8	93	F8	TACCAGCCAAGCCCATACTAGAGG	9番染色体内に位置するX染色体のPLP1遺伝子座に連結されないフォワードqPCRプライマー
	94	R8	GGGATCAACAGTGGCATAATTGAA	9番染色体内に位置するX染色体のPLP1遺伝子座に連結されないリバースqPCRプライマー

【0324】

プライマー対の性能を検証するため、男性ゲノムDNAまたは女性ゲノムDNAを鋳型として含有するPCR反応物を準備した。次いで、これらを、Illumina Eco測定器上のリアルタイムPCRにより、または従来のPCRにより増幅した。qPCRにより、女性は、男性より若干強いPLP1(X染色体)シグナルを示すと推論した。従来のPCRにより、本発明者らは、アンプリコンのサイズおよび独特のものであることを点検することが可能であった。いずれの試験も、8種のアッセイ全ての性能が良好であると

10

20

30

40

50

いう解釈と符合するデータをもたらした。

【0325】

PCR反応物の準備：各女性のPCR反応物または各男性のPCR反応物について、

- － 100 μ lの水
- － 25 μ lの10× STD Taq緩衝液
- － 25mMのMgCl₂ 25 μ l
- － 60 ng/ μ lのシアリングされたgDNA 25 μ l (Qubitにより、女性および男性ともに同じ濃度とした)

- － 12.5 μ lのDMSO
 - － 10mMのdNTP 12.5 μ l
 - － 6.25 μ lのEvaGreen色素 (Biotum)
 - － 5 μ lのROX色素 (Invitrogen)
 - － 2.5 μ lのTaq DNAポリメラーゼをよく混合し、添加したもの
- を含有する、250 μ lのマスターミックスを、氷上で作製した。

【0326】

【表7】

表7:PLP1 qPCRアッセイバリデーションによるCq値

	アッセイ1	アッセイ2	アッセイ3	アッセイ4	アッセイ5	アッセイ6	アッセイ7	アッセイ8
	chrX	chrX	chrX	chrX	chrX	chrX	chrX	chr9
F	23.27371	24.01229		24.19554	22.3413	22.10594	22.65759	22.31407
	23.33794	22.49233	23.87286	22.53665	22.49127	22.19686	22.73586	22.0446
	22.33889	22.40355	23.68538	22.50178	22.42961	22.23099	22.83415	22.32891
M	24.00953	23.56869		23.2648	23.51487	23.11367	23.65077	22.15477
	23.63462	23.48547	23.95081	23.18794	23.25046	22.93675	23.66965	21.23208
	23.77216	23.16852	24.16562	23.21774	22.93373	22.95635	23.61168	22.16391
F	22.98351	22.96939	23.77912	22.51921	22.42073	22.17793	22.74253	22.2292
M	23.80544	23.40756	24.05821	23.22349	23.23302	23.00226	23.64403	21.85026
M-F	0.8	0.4	0.3	0.7	0.8	0.8	0.9	-0.4

【0327】

実験のために、24 μ lのミックスを、8本のストリップチューブ2セット（女性または男性）にアリコート分割し、各アッセイから得られた10 μ Mのフォワードプライマーおよびリバースプライマーを含有する、6 μ lのプライマーミックスを添加した。混合した後、3つの同一な5 μ lずつの量を、48ウェルEco PCRプレートの列（列内の上側における3連の女性試料、列内の下側における3連の男性試料）にアリコート分割した。SYBRおよびROXをモニタリングして、95℃まで30秒間、60℃で30秒間、および72℃で30秒間の40サイクルにわたりサイクリングさせるように測定器を設定した。アッセイ6についての増幅トレースのJPG画像を、図9に示す。女性試料と男性試料との間のコピー差は、明確であった。女性試料および男性試料について、全ての「Cq」値（蛍光曲線が、自動で規定されたあるベースラインを超えるときの値）を集積し、次いで、3連の測定値の平均値の間の差を計算した。これを、上記の表7に示す（下線＝男性－女性）が、ここでは、9番染色体アッセイを除き、全ての値が正である。全体的なデータにより、8種のアッセイ全ての性能は同様であり（22～24のCq値）、X染色体によるアッセイは一般に、女性におけるシグナルが大きいことが指し示される。

【0328】

従来のPCR反応は、94℃で30秒間、60℃で30秒間、および72℃で30秒間、72℃で2分間にわたる休止、10℃での保持の30サイクルにわたりサイクリングさせた。合計5 μ lの産物を、精製せずに、2%のアガロースゲル上に直接ロードしたが、これを図10に示す。各二重項の上側バンドは、アッセイPCR産物の推定移動度と符合した。下側の「ファジー」材料は、使用されないPCRプライマーであった可能性が高い

。

【0329】

リアルタイムPCRおよび従来のPCRに続くゲル解析の結果から、これらの8種のアッセイは、もっぱらそれらの意図された領域を増幅し、断片の濃縮をモニタリングするのに適すると結論付けることができる。

【0330】

実施例5の付録：アッセイデザインについての詳細

【0331】

PLP1遺伝子：転写物ID：NM_000533.3；エキソン2：187ヌクレオチド；UCSCブラウザーから得られるCDS2

CDSは下線を付した太字の大文字で示し、プライマー配列には影を付す。

ランキング配列は小文字で示す。

【化1-1】

配列番号	名称	配列
95	フォワード鎖	agtgcctactatctccgagcctgtgagcacagggcctggcagaggggtttgagtgccatgagct <u>acctactggatgacctgactgtttcccttcttcttcccccag</u> GCTTGTTAGAGTGC TGTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTG <u>GTGGCCACTGGATTGTGTTTCTTTGGGGT</u> GGCACTGTTC TGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAG <u>CTAATTGAGACCTATTTCTCCAAAACTACCAAGACTAT</u> GAGTATCTCATCAATGTgtaagtacctgccctcccacacagaccatcttttttcc ctctctccatccctggagatagagaactctcaglaccttagtaactagcaggggactggggggga
96	リバーズ相補鎖	tccaccccagccccctgctagtactaaggtactgaagagtctctatctccaggatggagagagggg aaaaaaagatgggtctgtgtggagggcagggtactacACATTGATGAGATACT CATAGTCTTGGTAGTTTTTGGAGAAATAGGTCTCAATTAGC

10

20

30

【化 1 - 2】

		TTTTCTGTGCCAGTGAGGGCTTCATGTCCACAGCCACAGAA CAGTGCCACCCCAAAGAAACACAATCCAGTGGCCACCAGG GAAGCAAAGGGGGCCCCCTACCAGACATCTTGCACAGCACT CTAACAAGCctggggaagaagaaggggaaacagtcaggcacatccagtaggtagetca tgccactcaaaccctctgccaggccctgtgtcacaggctcggagatagtgggcac
97	77ヌクレオチドの qPCRアッセイ1:	agtgccactatctccgagcctgtgagcacaggccctggcagaggggtttgagtgccatgagct <u>acctactggatgtgacctgactgtttcccttcttcttcccccag</u> GCTTGTTAGAGTGC <u>TGTGCAAGATGCTGGTAGGGGGCCCCCTTTGCTTCCCTG</u> <u>GTGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTT</u> TGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAG <u>CTAATTGAGACCTATTTCTCCAAAACTACCAAGACTAT</u> GAGTATCTCATCAATGTgtaagtacctgcccctccacacagaccatcttttttcc ctctctccatcctggagatagagaactcttcagtaaccttagtaactagcaggggactgggggtgga
98	83ヌクレオチドの qPCRアッセイ2:	agtgccactatctccgagcctgtgagcacaggccctggcagaggggtttgagtgccatgagct <u>acctactggatgtgacctgactgtttcccttcttcttcccccag</u> GCTTGTTAGAGTGC TGTGCAAGATGCTGGTAGGGGGCCCCCTTTGCTTCCCTG <u>GTGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTT</u> TGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAG <u>CTAATTGAGACCTATTTCTCCAAAACTACCAAGACTAT</u> GAGTATCTCATCAATGTgtaagtacctgcccctccacacagaccatcttttttcc ctctctccatcctggagatagagaactcttcagtaaccttagtaactagcaggggactgggggtgga
99	81ヌクレオチドの qPCRアッセイ3:	agtgccactatctccgagcctgtgagcacaggccctggcagaggggtttgagtgccatgagct <u>acctactggatgtgacctgactgtttcccttcttcttcccccag</u> GCTTGTTAGAGTGC TGTGCAAGATGCTGGTAGGGGGCCCCCTTTGCTTCCCTG <u>GTGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTT</u> TGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAG <u>CTAATTGAGACCTATTTCTCCAAAACTACCAAGACTAT</u> GAGTATCTCATCAATGTgtaagtacctgcccctccacacagaccatcttttttcc ctctctccatcctggagatagagaactcttcagtaaccttagtaactagcaggggactgggggtgga
100	97ヌクレオチドの qPCRアッセイ4:	agtgccactatctccgagcctgtgagcacaggccctggcagaggggtttgagtgccatgagct <u>acctactggatgtgacctgactgtttcccttcttcttcccccag</u> GCTTGTTAGAGTGC

10

20

30

40

【化 1 - 3】

		TGTGCAAG<u>ATGTCTGGTAGGGGCCCCCTTTGCTTCCCTG</u> <u>GTGGCCACTGGATTGTGTTTCTTTGGGGT</u>GGCACTGTTC TGTGGCTGTGGAC<u>ATGAAGCCCTCACTGGCACAGAAAAG</u> <u>CTAATTGAGACCTATTTCTCCAAAACTACCAAG</u>ACTAT GAGTATCTCATCAATGTgtaagtacctgccctcccacagacccatcttttttcc ctctctccatcctggagatagagaactcttcaglaccttagtaactagcaggggactgggggtgga	10
101	95ヌクレオチドの qPCRアッセイ5:	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggttttgaatgcataagct <u>acctactgcatgtgctgactgttttcccccttcttcttccccag</u>GCTTGTTAGAGTGC TGTGCAAG<u>ATGTCTGGTAGGGGCCCCCTTTGCTTCCCTG</u> <u>GTGGCCACTGGATTGTGTTTCTTTGGGGT</u>GGCACTGTTC TGTGGCTGTGGAC<u>ATGAAGCCCTCACTGGCACAGAAAAG</u> <u>CTAATTGAGACCTATTTCTCCAAAACTACCAAG</u>ACTAT GAGTATCTCATCAATGTgtaagtacctgccctcccacagacccatcttttttcc ctctctccatcctggagatagagaactcttcaglaccttagtaactagcaggggactgggggtgga	20
102	エキソン2に対して200bp 遠位にある100bpの産 物である、qPCRアッセイ 6 (>chrX:103040918+ 103041017 100bp)	CCCCCTGTTTTCTTACACGTGTTCTgacttctgtaggtgtggttcatat tgcccaagltggagcctccagcgtagTAGGTATGGAGAAGCCAAGGGAGG	
103	エキソン2に対して 1000bp遠位(エキソン3 内)にある81bpの産物で ある、qPCRアッセイ7 (> chrX:103041614+ 103041694 100bp)	GTGTGTCATTGTTTGGGAAAATGGcttaggacatcccagacaggtgatcat cctcaggATTTTGTGGCAATAACAAGGGGTG	30
104	9番染色体上の非連結 アッセイであり、96bpの産 物である、qPCRアッセイ 8 (>chr9:103041653+ 103041748 96bp)	TACCAGCCAAGCCCATACTAGAGGctgtccccagatgctgaacccatct gattgaataaccatctgtatcaTTCAATTATGCCACTGTTGATCCC	40

【0332】

(実施例6)

PLP1エキソン2の捕捉
目的

一実施形態では、Clearfork Bioscience v1.0によるDNA
捕捉戦略は、特異的ゲノム標的領域に標的化された、多機能性プローブの使用を伴う。目

標は、PLP1エキソン2を標的とする、Ultramer（商標）（Integrated DNA Technologies（IDT）、Coralville、IA；Ultramerとは、45～200ヌクレオチドの範囲にわたる長さの特化合成オリゴヌクレオチドに与えられた商標名である）を使用する手法を検証することであった。

【0333】

概要

本実施例では、捕捉反応を実証した。IDT-DNA製のUltramerは、捕捉のために良好に働き、基本プロトコルは、捕捉ステップを通じて、試薬の化学量論比に関して妥当であり、分子密集剤であるPEGは、効果的な捕捉に干渉した。捕捉後、続いて、酵素プロセッシングに取り組んだ。

10

【0334】

簡単な記載

多機能性プローブを、図2に概略図化する。この実験データセットの目標は、これらのプローブの3つの特色全てについて調べることであった。領域1は、34ヌクレオチドの、5'ビオチン-TEG修飾され、かつ、3'ジデオキシシトシン修飾された、ユニバーサル「プルダウン」オリゴのための結合性部位であった。これらのユニバーサル領域のうちの2種は、同等の（望ましくは）性能を検証する（validate/verify）ようにデザインした。

【0335】

これらの2種のユニバーサルオリゴの配列を、下記の表8に示す。

20

【0336】

【表8】

表8: PLP1エキソン2を標的化するプルダウンオリゴ配列

配列番号	名称	配列	説明
105	Univ_seq_1	/5BioTEG/CACTGGACTATGTAGTACCTCACTCAGCAATAC/3ddC/	2つの5'ビオチン-TEGおよび3'ジデオキシシトシンを修飾したユニバーサル捕捉オリゴのうちの1つ
106	Univ_seq_10	/5BioTEG/GGCAACGAACGGACTGGAATCTACGGTCACCAA/3ddC/	2つの5'ビオチン-TEGおよび3'ジデオキシシトシンを修飾したユニバーサル捕捉オリゴのうちの1つ

30

【0337】

以下は、これらの配列をどのようにして選択したのかについての簡単な記載である。これらのオリゴの機能的役割は、捕捉プローブとハイブリダイズし、これにより、ストレプトアビジンで修飾された磁気ビーズ上で捕捉するために使用されうる、安定的に結合させたビオチン突出をもたらすことであった。

40

【0338】

10種のランダム配列を、ランダムDNA配列作製セットにより、塩基組成が約50%のGCとなるように作製した。使用したウェブサイトは、www.faculty.ucr.edu/mmaduro/random.htmであった。次いで、10種の配列を、BLATにより、ヒトゲノムのhg19 buildに照らしてスクリーニングした。顕著なアラインメントを示したのは、配列3だけであった。「C」で終結する2種の配列は、ddCによりブロックされうるので選択した。両方の配列を、IDT Oligo Analyzerにより解析した。配列1は、GCが47%であり、1MのNaCl中の融解温度が76℃である。配列2のGC含量は、57%であり、高濃度の塩中の融解温度は86℃である。選択された配列である1お

50

よび10は、実際の「ユニバーサル」5'ビオチンTEG-ddCと相補的なプローブ配列である。これらのうちのリバース相補鎖を、捕捉プローブ上のテイルとして使用した。続いて、4種の塩基、A、G、C、およびTを付加することにより、これらの配列を変化させて、長さを34塩基まで延長した。この長さは、SBCについて良好に働き、変化させる確固たる理由は見いだされなかった。第2に、CGCG型のモチーフのうちのいくつかを破壊して、自己二量体の形成を低下させた。

【0339】

領域2は、試料であるゲノムライブラリー内のゲノム配列に接触するようにデザインされたプローブの部分を含括した。この実験では、標的領域は、PLP1のエキソン2であった。PLP1エキソン2のDNA配列を下記に示す。CDSであるエキソン2を、下線を付した太字の大文字体で強調する。等間隔の捕捉プローブ配列には影を付す。

【0340】

【化2-1】

配列番号	名称	配列
107	フォワード鎖	agtgcceactatctccgagccctgtgagcacagggcctggcagagggg <u>tttgagtgcatgagctacctactggatgtgcc</u> <u>tgactgtttcccttcttcttcccccag</u> GCTTGTTAGAGTGCTGTGCAAGATGTCTGGTAGG <u>GGCCCCCTTTGCTTCCCTGGTGGCCACTGGATTGTGTTTCTTTGGGGTG</u> GCACTGTTCTGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCT <u>AATTGAGACCTATTTCTCCAAAACTACCAAG</u> ACTATGAGTATCTCATCA

【化2-2】

	ATGT <u>gtaagtacctgccctccacacagaccatcttttttccctctctccatcctggagatagagaactcttcagtacctta</u> gtaactagcaggggactggggggga
--	--

【0341】

領域3は、CAC3と呼ばれる、検証されたPCRプライマーと相補的であった。CAC3 PCRプライマーの配列は、C A C G G G A G T T G A T C C T G G T T T T C A C (配列番号72)である。

【0342】

これらのプローブ領域を含むUltramerの配列を表9に示す。

【0343】

【表 9 - 1】

表9: PLP1エキソン2捕捉で使用したUltramerの配列

配列番号	名称	配列	説明
108	B1_PLP1ex2_p1	GAGTTGATCCTGGTTFITCACTT TGAGTGGCATGAGCTACCTAC TGGATGTGCTGACTGTTTCCC CTTCTTCTCCCCAGGGTATTG CTGAGTGAGGTACTACATAGT CCAGTG	PLP1エキソン2を標的化し、 ユニバーサル配列1と適合的 なUltramer
109	B1_PLP1ex2_p2	GAGTTGATCCTGGTTFITCACAT GTCTGGTAGGGGCCCCCTTTG CTTCCCTGGTGGCCACTGGATT GTGTTTCTTTGGGGTGGTATTG CTGAGTGAGGTACTACATAGT CCAGTG	PLP1エキソン2を標的化し、 ユニバーサル配列1と適合的 なUltramer
110	B1_PLP1ex2_p3	GAGTTGATCCTGGTTFITCACCT TGGTAGTTTTTGGAGAAATAG GTCTCAATTAGCTTTTCTGTGC CAGTGAGGGCTTCATGGTATT GCTGAGTGAGGTACTACATAG TCCAGTG	PLP1エキソン2を標的化し、 ユニバーサル配列1と適合的 なUltramer
111	B1_PLP1ex2_p4	GAGTTGATCCTGGTTFITCACAT CTCCAGGATGGAGAGAGGGAA AAAAAAGATGGGTCTGTGTGG GAGGGCAGGTACTTACGGTAT TGCTGAGTGAGGTACTACATA GTCCAGTG	PLP1エキソン2を標的化し、 ユニバーサル配列1と適合的 なUltramer
112	B10_PLP1ex2_p1	GAGTTGATCCTGGTTFITCACTT TGAGTGGCATGAGCTACCTAC TGGATGTGCTGACTGTTTCCC CTTCTTCTCCCCAGGTTGGTG ACCGTAGATTCCAGTCCGTTT GTTGCC	PLP1エキソン2を標的化し、 ユニバーサル配列10と適合的 なUltramer
113	B10_PLP1ex2_p2	GAGTTGATCCTGGTTFITCACAT GTCTGGTAGGGGCCCCCTTTG CTTCCCTGGTGGCCACTGGATT GTGTTTCTTTGGGGTGGTGGTG ACCGTAGATTCCAGTCCGTTT GTTGCC	PLP1エキソン2を標的化し、 ユニバーサル配列10と適合的 なUltramer

【表 9 - 2】

114	B10_PLP1ex2_p3	GAGTTGATCCTGGTTFITCACCT TGGTAGTTTTTGGAGAAATAG GTCTCAATTAGCTTTTCTGTGC CAGTGAGGGCTTCATGTTGGT GACCGTAGATTCCAGTCCGTT CGTTGCC	PLP1エキソン2を標的化し、 ユニバーサル配列10と適合的 なUltramer
115	B10_PLP1ex2_p4	GAGTTGATCCTGGTTFITCACAT CTCCAGGATGGAGAGAGGGAA AAAAAAGATGGGTCTGTGTGG GAGGGCAGGTACTTACGTTGG TGACCGTAGATTCCAGTCCGTT CGTTGCC	PLP1エキソン2を標的化し、 ユニバーサル配列10と適合的 なUltramer

【0344】

モル、マイクログラム、および分子についての検討：実施例3で構築したゲノムライブラリー（Promega製の女性hgDNAライブラリー）を使用した。このライブラリーの大スケール（800 μl）の増幅は、投入物としての20 μlのライゲーションミックスで開始して実行した。精製ライブラリー（400 μl）の最終濃度は、22 ng/μ

10

20

30

40

50

1であった。本明細書で記載される実験1回当たり1マイクログラムを使用した。さらに、50bpである全アダプターおよび150～200bpであるインサートに基づき、ライブラリー質量のうちの75%は、ゲノムDNAであると仮定した。次いで、この仮定と、1つのヒトゲノムの質量は3pgであるという事実とに基づき、約250,000 ($750 \times 10^{-9} / 3 \times 10^{-12} = 250,000$) コピーの所与の任意のゲノム領域が存在した。かつての実験および文献は、10,000倍のモル過剰量のプローブが、妥当な出発点であることを示唆した。これは、2,500,000,000個のプローブ分子を含意する。分子 2.5×10^9 個/分子 6.02×10^{23} 個/モル = 4.15×10^{-15} モル = 4アトモルのプローブである。これを、原液の容量に変換すると、4nM (の各プローブ) 1μl = 4アトモルのプローブである。最後に、Invitrogen製のMyOne strepでコーティングされたC1ビーズは、1μlのビーズ当たり約1ピコモルの500bpビオチニル化dsDNAに結合する。この実験では、合計4アトモル×4プローブ=16アトモルのプローブを添加した。1μlのビーズは、1000アトモルに結合し、1μlは、ビーズが働くための実際のな量であり、1μlのビーズは、添加されるプローブの60倍過剰量の結合能を有する。したがって、本実施例では、以下のパラメータ：

- － 単位質量のライブラリー内の標的分子の数 (1μgのライブラリー当たり、250,000コピーの独特の二倍体遺伝子座)；
- － 10,000倍のモル過剰量のプローブで標的遺伝子座に取り組むのに必要なプローブのモル濃度 (4アトモルの各プローブ、16アトモルの全プローブ (4種のプローブ)、4nMのプローブ溶液1μl)；および
- － 添加される全てのプローブを定量的に捕捉するのに必要なビーズの量 (1μlは、1000アトモルのdsDNAおよび/または結合していないプローブに結合する) を計算した。

【0345】

緩衝液および作業溶液

溶液1：結合性プローブ：ユニバーサル結合性パートナーおよびPLP1のプローブを、100μMまで水和させた。2つの個別のチューブ内で、92μlのTEz+0.05%のTween-20緩衝液と、4μlのユニバーサルオリゴと、各々1μlずつの4種のコグネイト (ユニバーサルオリゴと) プローブとを組み合わせた。これにより、1μMのプローブ原液のうちの2種を作製した。これらの各々4μlずつを、1000μlのTEz+Tweenに希釈して、4nMのプローブ作業溶液をもたらした。

【0346】

4倍濃度の結合緩衝液 = 4MのNaCl、40mMのトリス、pH8.0、0.4mMのEDTA、および0.4%のTween 20。5MのNaCl 40mlと、1Mのトリス、pH8.0 2mlと、10%のTween 20 2mlと、0.5MのEDTA 40μlと、6mlの水とを組み合わせることにより、50mlを作製した。

【0347】

洗浄緩衝液 = 25%のホルムアミド、10mMのトリス、pH8.0、0.1mMのEDTA、および0.05%のTween 20。37mlの水と、12.5mlのホルムアミドと、1Mのトリス、pH8.0 500μlと、0.5MのEDTA 10μlと、10%のTween 20 250μlとを組み合わせることにより、50mlを作製した。

【0348】

ビーズ：4倍濃度の結合緩衝液250μlと、750μlの水とを組み合わせ、1倍濃度の結合緩衝液を作製した。10μlのビーズを、1倍濃度の結合緩衝液90μlに添加し、磁石で引き寄せ、ビーズを、1倍濃度の結合緩衝液100μlで2回にわたり洗浄し、洗浄されたビーズを、1倍濃度の結合緩衝液100μl中に再懸濁させた。10マイクロリットルの洗浄されたビーズは、製造元のチューブから取り出されたときの1μlのビーズに相当する。

【0349】

方法

以下の3つのパラメータについて調べた。

1. オリゴ10と対比したユニバーサルピオチンオリゴ1；
2. 7.5%のPEG8000（アニーリングの速度を増強しうる分子密集剤）を加えた1倍濃度の結合緩衝液中の結合と対比した、1倍濃度の結合緩衝液中の結合；
3. プロセッシングを伴わない結合後におけるPLP1領域の濃縮倍数、および酵素プロセッシングを加えた結合後におけるPLP1領域の濃縮倍数

【0350】

これらのパラメータについて調べるために、8例の試料（ $2 \times 2 \times 2$ ）を作製した。これらの試料は、 $50 \mu\text{l}$ の $20 \text{ ng} / \mu\text{l}$ のゲノムDNA、4倍濃度の結合緩衝液 $25 \mu\text{l}$ 、 $1 \mu\text{l}$ の結合性プローブ、および $24 \mu\text{l}$ の水または50%のPEG8000 $20 \mu\text{l} + 4 \mu\text{l}$ の水（PEGを伴う4例の試料およびPEGを伴わない4例の試料）を含有した。OligoAnalyzerについてのIDT DNAウェブサイトによれば、高濃度の塩（例えば、1MのNaCl）中では、オリゴの T_m は、劇的に高い温度にシフトすることが記載されていた。したがって、試料を 95°C で融解させ、次いで、 1°C および2分間の減分で、 60°C まで温度を下げた（本発明者らによるABI2720サーマルサイクラー上のAutoXによる35サイクルであり、各サイクルでは、 1°C ずつ低下させ、各サイクルは、2分間にわたり持続させる）。試料を室温（RT）まで冷却した後、試料1例当たり $10 \mu\text{l}$ の洗浄されたビーズを添加し、20分間にわたりインキュベートした。ビーズは、強力な磁石でブルアウトし、溶液を吸引し、廃棄した。ビーズを、洗浄緩衝液による $200 \mu\text{l}$ の洗浄液で4回にわたり洗浄し、ビーズを再懸濁させるたびごとに、RTで5分間にわたりインキュベートした。最終回の洗浄の後、残りの洗浄液の大部分を、チューブから十分に吸引した。

【0351】

4本のチューブのセットを、T4 DNAポリメラーゼで処理した。 $10 \mu\text{l}$ のNew England Biolab $10 \times$ Quick blunting bufferと、同じキットから得られる1mMのdNTP $10 \mu\text{l}$ と、 $10 \mu\text{l}$ の水と、 $1 \mu\text{l}$ のT4 DNAポリメラーゼとを組み合わせることにより、カクテルを作製した。 $20 \mu\text{l}$ を、4本のチューブのセットに添加し、反応物を、 20°C で15分間にわたりインキュベートした。

【0352】

捕捉後におけるPCR増幅のため、非T4処理試料（捕捉だけ施された）を、単一プライマー反応において、ACA2-25（TGCAAGGACCAAGAGAAATTCGAATACA；配列番号67）で増幅した。T4処理試料は、ACA2FLプライマーおよびCAC3FLプライマー（それぞれ、AATGATACGGCGACCACCGAGATCTACACGTCATGCAGGACCAGAGAATTCGAATACA（配列番号69）およびCAAGCAGAAGACGGCATACGAGATGTGACTGGCACGGGAGTTGATCCTGGTTTTCAC（配列番号74））により増幅した。コア反応ミックスは、反応物 $400 \mu\text{l}$ 当たり、 $120 \mu\text{l}$ の水、 $40 \mu\text{l}$ の $10 \times$ STD Taq緩衝液（NEB）、 25 mM のMgCl₂ $40 \mu\text{l}$ 、 $10 \mu\text{M}$ の単一プライマー $80 \mu\text{l}$ 、または $40 \mu\text{l} + 40 \mu\text{l}$ のFプライマーおよびRプライマー、 $20 \mu\text{l}$ のDMSO、 10 mM のdNTP $20 \mu\text{l}$ 、ならびに $4 \mu\text{l}$ のTaqポリメラーゼを含有した。 $80 \mu\text{l}$ のアリコートを、 $20 \mu\text{l}$ のTEz中に再懸濁させたビーズ（結合だけ）または $20 \mu\text{l}$ のT4ミックスに添加した。最終的な容量は、 $100 \mu\text{l}$ であった。これらの試料を、 94°C で30秒間、 60°C で30秒間、および 72°C で60秒間の30サイクルにわたるPCRにより増幅した。ゲル解析（レーン1つ当たり $5 \mu\text{l}$ のPCR後材料をロードした）を、結果節に示す。Qubitによる読取り値により、各PCR反応物の濃度は、約 $20 \sim 25 \text{ ng} / \mu\text{l}$ であることが指し示された。

【0353】

増幅後解析のために、 $200 \mu\text{l}$ の水と、10倍濃度のTaq緩衝液 $50 \mu\text{l}$ と、25

mMのMgCl₂ 50 μlと、25 μlのDMSOと、10 mMのdNTP 25 μlと、12.5 μlのEvaGreen (Biotum)と、5 μlのTaqポリメラーゼ (NEB)とを組み合わせることにより、従来のPCRミックスによる、500 μl (最終容量)のマスターミックスを作製した。42 μlのアリコートをし、8本のチューブに分配し、10 μMのF+R PLP1プライマーミックス12 μl (アッセイについては、実施例5: PLP1 qPCRアッセイの検証において記載されている)を添加した。9 μlのミックスを、8列の各アッセイに分配した。ウェル1つ当たりの試料1 μlで合計6例の試料をアッセイした。これらの試料は、

- 行1: gDNAライブラリー出発材料
- 行2: ビオチンオリゴ1 捕捉材料
- 行3: ビオチンオリゴ1 + PEG 捕捉材料
- 行4: ビオチンオリゴ10 捕捉材料
- 行5: ビオチンオリゴ10 + PEG 捕捉材料
- 行6: TEZ NTC 対照

であった。

【0354】

T4 処理試料は、PCR増幅で処理されたのが異常な材料だけであることがゲル解析により示されたため、アッセイしなかった。

【0355】

結果

捕捉だけのライブラリーは、予測される通り、投入されたゲノムライブラリーと同様のスミアをもたらした。試料は、左から右へ、(1) オリゴ1、(2) オリゴ1 + PEG、(3) オリゴ10、および(4) オリゴ10 + PEGであった。T4 処理試料には、残留T4 ポリメラーゼが夾雑した(5~8)。特定の実施形態では、T4 ポリメラーゼを熱不活化した。

【0356】

Qubitで測定した4つの捕捉だけのライブラリーの収率を下記の表10に示す。

【0357】

【表10】

表10: Qubitで測定した4つの捕捉だけのライブラリーの収率

C1ユニバーサルオリゴ	23.2	ng/ul
C1+PEG	27.2	ng/ul
C10ユニバーサルオリゴ	24.8	ng/ul
C10+PEG	25.6	ng/ul

【0358】

qPCRのために、8つの検証されたPLP1アッセイ (実施例5) の全てを、列において使用し、試料を行において使用した。試料のアレイは、

- 行1: 1 μlの25 ng/μlのgDNAライブラリー
- 行2: 1 μlの約25 ng/μlのC1 捕捉試料
- 行3: 1 μlの約25 ng/μlのC1 + P 捕捉試料
- 行4: 1 μlの約25 ng/μlのC10 捕捉試料
- 行5: 1 μlの約25 ng/μlのC10 + P 捕捉試料
- 行6: 1 μlのTEZ (NTC)

であった。

【0359】

ウェル1つ当たり1例の試料という、この構成では、データは、厳密な定量的測定ではなく、定性的概観であることを意図した。データを下記の表に示す。上の表は、生のCq

値である。次の表は、全ての試料およびアッセイは、同じ2倍の検量線に適合するという仮定に基づいて絶対値に変換されたCq値である。下の表は、捕捉された試料のCq値を、gDNAライブラリーのCq値で除した商を示す。これにより、捕捉後における濃縮倍数の意味が分かる。

【0360】

【表11】

表11: 標的配列の捕捉誘導型濃縮の成功を実証する、8種のPLP1 qPCRアッセイの結果

Cq値	アッセイ1	アッセイ2	アッセイ3	アッセイ4	アッセイ5	アッセイ6	アッセイ7	アッセイ8
gDNA lib	22	25	29	25	25	26	25	25
C1	14	11	11	11	9	15	23	22
C1+P	18	15	18	14	13	18	29	26
C10	12	12	13	12	10	15	22	25
C10+P	16	13	15	13	12	18	27	25
NTC	29	25	34	28	30	28	30	31

Cq値	アッセイ1	アッセイ2	アッセイ3	アッセイ4	アッセイ5	アッセイ6	アッセイ7	アッセイ8
gDNA lib	33	3	0	4	3	2	3	2
C1	4782	56660	50773	43156	138885	2950	10	28
C1+P	293	3196	407	6187	16082	450	0	1
C10	29864	34343	11961	28507	123383	2894	31	3
C10+P	1234	11632	3976	10060	25772	413	1	2
NTC	0	3	0	0	0	1	0	0

cap/gDNA	アッセイ1	アッセイ2	アッセイ3	アッセイ4	アッセイ5	アッセイ6	アッセイ7	アッセイ8
C1	147	19303	330609	11735	48601	1720	3	11
C1+P	9	1089	2649	1682	5628	262	0	1
C10	917	11700	77886	7752	43176	1688	10	1
C10+P	38	3963	25890	2736	9019	241	0	1

【0361】

データからいくつかの結論が引き出された。(1) 捕捉は働いた。C1について、的中アッセイ1～5を通じた捕捉濃縮の平均は、82,000倍であった。C10についての平均は、28,000倍であった。アッセイ部位では、約数百～数万倍の濃縮が観察された。これは、UltraMerが働き、基本プローブのデザインが効果的であることを含意する。これは、gDNA対プローブ対ピーズの基本的な化学量論比が適正であることを意味した；(2) 2つのビオチンデザインは、ほぼ同様に働いた；(3) PEGは、捕捉効率を増強せず、阻害した；および(4) アッセイ6では、標的から200bpの、顕著な「混獲」が観察された。1000bp離れた領域について認められる逸脱活性は小さかった。

【0362】

特定の実施形態では、捕捉された複合体の酵素プロセッシングが、このスキームにおける感度（濃縮倍数）および特異度（「混獲」の程度）に寄与するのかどうかを決定することが重要でありうる。

【0363】

(実施例7)

SYBR空間におけるPLP1 qPCRアッセイ
目的

場合によっては、リアルタイムの条件が、非リアルタイムの増幅条件を正確に模倣することが有用である。本実施例では、これは、氷上における準備と、3段階の比較的緩徐なPCR反応とを意味した。代替的に、いくつかのアッセイは、増幅条件のセットの複製を必要とせず、定量的測定値を厳密に求めることが意図される。例えば、PLP1 qPCRアッセイは、断片を作製するためには使用せず、遺伝子座の濃縮を測定するためだけに

使用することが好ましい。この種の状況では、室温で準備された q P C R 反応物および迅速なサイクリングが有利である。この実験では、A B I 2 × S Y B R ミックスによる 8 種の P L P 1 アッセイについて調べた。これらは、実施例 5 (P L P 1 q P C R アッセイの検証) で記載したアッセイと同じプライマーアッセイである。

【0364】

概要

これらのデータは、8 種の P L P 1 q P C R アッセイのうちの少なくとも 6 種を、S Y B R G r e e n q P C R ミックスおよび条件と共に使用しうることを示唆した。

【0365】

方法

女性 g D N A ライブラリー (実施例 3 : P r o m e g a 製の女性 h g D N A ライブラリー) に対する P L P 1 アッセイの性能を測定した。10 μ l のウェル 1 つ当たり、5 μ l の A B I 2 × S Y B R マスターミックスと、10 μ M の F + R プライマー原液 0.2 μ l と、1 μ l の g D N A ライブラリー (20 ng / μ l の) と、3.8 μ l の水とを組み合わせた (大容量のマスターミックスを作製し、アリコート分割した)。非鋳型対照についての 3 連の測定および g D N A ライブラリーについての 3 連の測定を、各アッセイを通じて行った。R O X パッシブ基準色素による標準化を伴う、標準的な 2 ステップの P C R (95 °C で 15 秒間、60 °C で 45 秒間) を使用して、I l l u m i n a E c o リアルタイム P C R 上で、40 サイクルにわたりサイクリングさせた。

【0366】

結果

各ウェルについて判定された C q 値を、下記の表 12 に示す。N T C は、極めて清浄であり、g D N A の C q は可変的であり、これは、ピペティングに起因する可能性が高い。一般的な主題は、アッセイ 1 および 7 の性能が低かったのに対し、残りのアッセイは、S Y B R 空間において妥当な程度に良好に働いたことである。図 11 では、N T C トレース (A) および + g D N A トレース (B) は、アッセイの性能についての定性的描像を提示するようにコピーされた。

【0367】

【表 12】

表 12: SYBR green ベースの PLP1 迅速アッセイにより判定された Cq 値

アッセイ	1	2	3	4	5	6	7	8
NTC	38	40	40	40	40	40	40	40
	40	40	40	40	40	40	40	40
	35	40	40	40	40	40	40	40
gDNA	35	25	25	24	27	22	32	26
	35	25	25	24	26	23	32	27
	32	24	22	22	23	21	29	24

【0368】

(実施例 8)

複合体の酵素プロセシングの前および後における、P L P 1 エキソン 2 濃縮の測定
目的

本実施例では、P L P 1 エキソン 2 D N A の「比活性」を、プロセシング前およびプロセシング後の捕捉複合体において測定することにより、複合体の酵素プロセシングを、収率について直接調べた。U l t r a m e r は、優れた捕捉効率を裏付け、コア捕捉プロトコールの性能も良好であった。

【0369】

概要

この実験は、T4-DNAポリメラーゼによる捕捉後プロセシングが、捕捉反応の特異度を劇的に改善することを実証した。

【0370】

背景

実施例6 (PLP1エキソン2の捕捉)では、捕捉の成功について記載したが、PCRの前にT4ポリメラーゼを除去しない捕捉後プロセシングステップでは、アーチファクトライブラリーがもたらされた。本実施例では、PCRの前にT4を95℃で1分間にわたり熱不活化したことを除き、同じ基本実験を繰り返す。

【0371】

方法、結果、考察

この実験では、酵素プロセシングの前および後における複合体の捕捉効率を評価するために、2つのユニバーサルビオチン捕捉プローブを含む4例の試料を作製した。各試料は、20ng/μlのゲノムDNA 50μl、4倍濃度の結合緩衝液20μl、1μlの結合性プローブ、および最終容量を80μlとするための9μlの水を含有した。試料は、95℃で1分間にわたり融解させ、1℃、2分間の減分で60℃まで温度を低下させることによりアニーリングさせる(本発明者らによるABI2720サーマルサイクラー上のAutoXによる35サイクル)のに続いて、RTまで冷却した。次いで、試料1例当たり合計10μlの洗浄されたビーズ(1μlのMyOneビーズ溶液(ストレプトアビジンでコーティングされたC1; Invitrogen)に相当する)を添加し、20分間にわたりインキュベートした。ビーズは、強力な磁石でプルアウトし、溶液を吸引し、廃棄した。ビーズを、洗浄緩衝液による200μlの洗浄液で4回にわたり洗浄し、ビーズを再懸濁させるたびごとに、RTで5分間にわたりインキュベートした。最終回の洗浄の後、残りの洗浄液の大部分を、チューブから十分に吸引し、捕捉複合体でコーティングされたビーズを静置した。

【0372】

2例の試料のT4プロセシングのために、本発明者らは、40μlの水、5μlの10× quick blunt buffer (New England Biolabs)、1mMのdNTP 5μl、および0.5μlのT4 DNAポリメラーゼを含有する、50μlの酵素プロセシングミックスを調製した。複合体の2つのアリコート、20μl(ずつ)のT4ミックス中に再懸濁させ、20℃で15分間、95℃で1分間にわたりインキュベートし、RTまで冷却した。「非処理」対照を、T4ポリメラーゼを欠く、20μlの同じ緩衝液(40μlの水、5μlの10× quick blunt buffer (New England Biolabs)、1mMのdNTP 5μl)中に懸濁させた。

【0373】

比活性を測定するため、捕捉単独試料および捕捉+プロセシング試料の両方を、30サイクルのPCRにより増幅した。次いで、DNAを定量化し、特殊量および既知量の増幅されたDNAによるPLP1アッセイシグナルを測定した。本実施例では、2つの増幅反応物を準備した。捕捉単独では、これらのライブラリーは、この単一プライマーだけで増幅可能であるので、増幅をACA2-25 (TGCAAGGACCAAGAGAAATTCGAATACA; 配列番号67)で実行した。酵素プロセシングされた複合体では、増幅を、ACA2FLプライマーおよびCAC3FLプライマー(それぞれ、AATGATACGGCGACCACCGAGATCTACACGTCATGCAGGACCAGAGAATTCGAATACA(配列番号69)およびCAAGCAGAAGACGGCATACGAGATGTGACTGGCACGAGGTTGATCCTGGTTCAC(配列番号74))で実行した。100μlのPCRミックスは、10μlの10× STD Taq緩衝液(別段に指定されない限り、全ての試薬は、NEBから得た)、25mMのMgCl2 10μl、10μMの単一プライマー20μl、または10μMの二重プライマー10μl+10μl、20μlの鋳型(非処理対照またはT4プロセシング、ビーズおよび全ての鋳型)、5μlのDMSO、10

mMのdNTP 5 μ l、および1 μ lのTaq DNAポリメラーゼ（全ては、増幅の前に氷上で準備した）を含有した。試料は、95℃で30秒間、60℃で30秒間、72℃で60秒間の3ステッププロトコールに続き、72℃で2分間を施し、10℃で休止させる、30サイクルのPCRにより増幅した。

【0374】

増幅の後、DNA収率を測定し、PCR増幅された材料を、DNAゲル電気泳動で検討した。Qubit (Invitrogen) により測定された (DNA HSキット) 収率を、下記の表13に示す。これらのデータは、二重プライマーによる増幅は、単一プライマーによる増幅より大きな全体的収率を裏付けるという基本特色を強調する。

【0375】

10

【表13】

表13: 増幅されたDNAのQubitによる定量化

試料	収率 (ng/ul)
C1ビオチン非プロセシング複合体	23
C10ビオチン非プロセシング複合体	24
C1ビオチンT4プロセシング複合体	38
C10ビオチンT4プロセシング複合体	34

20

【0376】

ゲル画像（2%のアガロース、ロードされた材料100 ng）を、図12に示す。プロセシングは、2つの注目すべき効果を及ぼした。第1に、プロセシングは、約250 bp（上矢印）および約175 bp（下矢印）の2つの弱いバンドに加えて、予測されるスミアをもたらした。下側のバンドは、プローブの不測のクローニング（115 bpのアダプター+60 bpのプローブ=175 bp）と符合した。第2に、プロセシングは、全体的な試料のサイズ分布を低減した。これは、50 bpの単一のアダプターを、プロセシングされた材料の、上方への65 bpの全体的なシフトをもたらすことが予測される、115 bpの全長アダプターで置きかえたので、注目に値した。プロセシングにより、ライブラリーの平均インサートサイズが、顕著に低減されると解釈された。

30

【0377】

qPCRによる濃縮効率を測定するために、2種の試みを行った。第1のより定性的な試みでは、8種のPLP1アッセイ全て（実施例5：PLP1 qPCRアッセイの検証において詳細に記載した）を使用して、6例の試料：

1. アッセイ1種当たり25 ngの出発gDNAライブラリー
2. アッセイ1種当たり0.25 ngの非処理C1
3. アッセイ1種当たり0.25 ngの非処理C10
4. アッセイ1種当たり0.25 ngのT4処理C1
5. アッセイ1種当たり0.25 ngのT4処理C10
6. 非鋳型対照

40

を測定した。

【0378】

これらの単回の測定から得られるCq値を、下記の表14に示す。gDNAおよびNTC対照の性能は良好であり（上および下；最も明るい影）、さらには査定しなかった。

【0379】

【表 1 4】

表 14: 濃縮効率についての定性的PLP1アッセイ解析によるCq値

	アッセイ 1	アッセイ 2	アッセイ 3	アッセイ 4	アッセイ 5	アッセイ 6	アッセイ 7	アッセイ 8
gDNA	N/A	27	29	40	27	30	23	32
un-C1	12	8	8	9	5	11	24	29
un-C10	16	14	14	15	9	10	30	31
T4-C1	7	6	3	3	4	19	25	N/A
T4-C10	4	3	4	6	3	17	26	28
NTC	32	32	31	23	24	28	24	30

10

【0380】

T4 処理試料（暗い影）のシグナルは、定量的解析がそれほどインフォーマティブでなくなる程度に強かった（10未満のCq）。しかし、定性的レベルでは、非処理捕捉複合体（中程度の影）と比較して、2つの傾向が明確であった。一方は、アッセイ1～5から得られる的中シグナルが劇的に増大する（低値のCq）ことであった。他方は、プロセシングされると、アッセイ6から得られる、標的領域から200bp離れた標的外シグナルが、顕著に減殺されることであった。データには、いくつかの凹凸が見られたが、中心的なメッセージは、プロセシングにより、PLP1エキソン2シグナルの特異度が大きく増強されることであった。

20

【0381】

この実験のより定量的側面を捉えるため、qPCRの前に、非処理C10捕捉アンプリコンを1000倍に希釈し、プロセシングされたC10アンプリコンを15,000倍に希釈した。これは、Cq値を測定可能な範囲内に収めるために行った。次いで、出発gDNAライブラリーを検討し、qPCRプレートの4連ウェル内のこれらの希釈された試料を、2種の的中アッセイ（アッセイ2および5）および2種の標的外アッセイ（アッセイ6および7）を通じて検討した。4連ウェルのCq値を平均し、これらの値を下記の表15に示す。ここでもまた、gDNAシグナルは弱かったが、これらの実験の目標は、プロセシングされていない捕捉複合体におけるPLP1エキソン2シグナルを、T4ポリメラーゼ処理された捕捉複合体と比較することであるため、弱いシグナルの、データ解釈に対する影響は、それほど顕著ではなかった。Cq値は、あらゆるPCRサイクルによる2倍の増幅を仮定する「ユニバーサル」検量線を使用して、絶対値に変換した。表の第3の区画は、希釈率についての調整を示す。第4の区画である、プロセシングのされていない複合体およびT4処理された複合体の、gDNAに対する比は、それほど有用ではないが、表の下欄の、T4処理された複合体と対比した、プロセシングされていない複合体の定量比は有用である。実施例6では、C1についての82,000倍の非処理捕捉濃縮およびC10についての28,000倍の非処理捕捉濃縮が観察された（これらの実験全てと同様、gDNAの分母は、極めて低度のシグナルから導出されたので、倍数範囲は、これに定性的側面をもたらした）ので、捕捉単独は、300bpのPLP1エキソン2領域の50,000倍の濃縮をもたらすと推定するのが妥当であった。プロセシングは、この濃縮を、さらに50倍（表15から得られる、83倍と24倍との平均値）増大させ、濃縮を、250万倍～1千万倍（ゲノム1つ当たり30億塩基／300bpの標的）に押し上げる。したがって、qPCR測定値のレベルでは、捕捉＋プロセシングは、濃縮に関する最良の状況に近づくと考えられた。アッセイ6によりモニタリングされた標的から200bp離れた標的外シグナルが、捕捉単独では大幅に濃縮される（ヒッチハイカー、交差ハイブリダイゼーション効果）が、プロセシングにより顕著に減殺されることは注目に値した。

30

40

【0382】

【表 15】

表15:濃縮効率についての定量的PLP1アッセイ解析によるCq値

生のCq値				
Cq	アッセイ2	アッセイ5	アッセイ6	アッセイ7
gDNA	27	33	35	27
捕捉	15	12	17	32
cap+enz	13	11	24	31
絶対値に変換したCq				
abs val	アッセイ2	アッセイ5	アッセイ6	アッセイ7
gDNA	81	1	0	95
捕捉	253464	2538822	89450	3
cap+enz	1408479	4050620	506	4
試料の希釈率について調整した絶対値				
abs val	アッセイ2	アッセイ5	アッセイ6	アッセイ7
gDNA	81	1	0	95
捕捉	253463648	2538821701	89449912	2573
cap+enz	21127189405	60759299218	7584546	66931
捕捉された試料のgDNAに対する比				
/gDNA	アッセイ2	アッセイ5	アッセイ6	アッセイ7
捕捉	3126652	2394183716	239697693	
cap+enz	260618748	57297762664	20324204	
非プロセッシング試料のプロセッシング試料に対する比				
enz/cap	アッセイ2	アッセイ5	アッセイ6	アッセイ7
cap+enz	83	24	0.08	

【0383】

この実験は、捕捉＋プロセッシングの特異度（非標的 qPCR シグナル）に取り組んだ。増幅された DNA（PLP1 エキソン 2 から得られる）の 1 ng 当たりの比活性は、捕捉後プロセッシングにより大幅に増強された。この実験は、感度、すなわち、酵素により転換される捕捉複合体のパーセントには取り組まなかった。特定の実施形態ではまた、本方法の特異度および感度の両方についての定量的理解も重要である。

【0384】

（実施例 9）

捕捉後プロセッシングの直接的な測定

【0385】

目的

実施例 8 では、捕捉後プロセッシングが、標的捕捉の特異度を実質的に増大させる所望の目的を達成することを決定した。検討される他のきわめて重要なパラメータは、感度、すなわち、初期に捕捉された複合体であって、最終的なシーケンシングライブラリー内に回収される複合体の百分率である。本実施例では、本発明者らは、感度の直接的な測定により、酵素プロセッシングが、初期に捕捉された配列のうちの > 10 % について効果的であることを実証した。

【0386】

概要

この実験から得られるデータにより、的中捕捉複合体のうちの 10 % が、T4 ポリメラーゼにより、捕捉後シーケンシングライブラリー断片にプロセッシングされることが指示された。

【0387】

検討

参照として、捕捉後プロセッシングについての図式的例示を、図 4 に示す。この図では、

プロセシングの感度を、図13の右下で例示される、3ステップの手順で測定した。まず、単一のPLP1捕捉プローブを、女性gDNAライブラリー（実施例3：Promega製の女性hgDNAライブラリー）から、PLP1エキソン2特異的ゲノムDNA断片をプルダウン／プルアウトする独立の反応において使用した。4種のプローブが存在するので、4種のプルダウンを実行した。図13（A）で例示する通り、捕捉された材料の量は、隣接するPLP1-qPCRアッセイプライマー対を使用して測定した。図13（B）に示す通り、複合体の酵素プロセシングの後、1種のPLP1特異的プライマーおよび1種のプローブ特異的プライマーを使用することにより、qPCRを介して、プロセシングされた複合体の量を再度測定した。[B/A×100%]における測定値の比により、プロセシング効率の推定値がもたらされた。リアルタイム反応から得られるPCR産物を抽出し、ゲル解析により、予測される長さのアンプリコンが作製されたことを検証した（図13（C））ことは、実験結果の適正な解釈に極めて重要である。これは、いずれのPCR反応も個別の開始点および停止点を有したので可能であった。プロセシング効率は、A+B+Cから解釈可能なデータをもたらすプルアウトを使用して、プロセシング効率を決定した。

10

【0388】**アッセイ**

個々のプローブは、qPCRアッセイにマッチさせることが必要であった。qPCRアッセイの前工程および後工程にマッチさせたプローブの6種の組合せを選び出した。これらを、プローブ配列を斜字体とし、PLP1エキソン2特異的プライマーに影を付して、下記に示す。暗い影を付したプライマーは、プロセシング後にCAC3プライマーと対合させたプライマーである。また、各アッセイセットについて、PCRアンプリコンの予測される産物サイズも示す。

20

【0389】

【化 3 - 1】

配列番号	名称	配列
116	プローブ1およびアッセイ1; qPCRアッセイ1=77bp; CAC3⇄R1=167bp	agtgcctactatctccgagcctgtgagcacagggcctggcagaggggtttgagtgcatgagct <u>acctactggatgtgcctgactgtttcccttcttcttccccag</u> GCTTGTTAGAGTGCTG TGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGT <u>GGCCACTGGATTGTGTTTCTTTGGGGT</u> GGCACTGTTCTGT GGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCTA <u>ATTGAGACCTATTTCTCCAAAACTACCAAG</u> ACTATGAG TATCTCATCAATGTgtaagtacctgccctccccacagaccatcttttttccctctct ccatcctggagatagagaactcttcagtaccttagtaactagcaggggactgggtgga
117	プローブ1およびアッセイ2; qPCRアッセイ2=83bp; CAC3⇄R2=224bp:	agtgcctactatctccgagcctgtgagcacagggcctggcagaggggtttgagtgcatgagct <u>acctactggatgtgcctgactgtttcccttcttcttccccag</u> GCTTGTTAGAGTGCTG TGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGT <u>GGCCACTGGATTGTGTTTCTTTGGGGT</u> GGCACTGTTCTGT GGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCTA <u>ATTGAGACCTATTTCTCCAAAACTACCAAG</u> ACTATGAG TATCTCATCAATGTgtaagtacctgccctccccacagaccatcttttttccctctct ccatcctggagatagagaactcttcagtaccttagtaactagcaggggactgggtgga
118	アッセイ3およびプローブ4; qPCRアッセイ3=81bp; F3⇄CAC3=187bp:	agtgcctactatctccgagcctgtgagcacagggcctggcagaggggtttgagtgcatgagct <u>acctactggatgtgcctgactgtttcccttcttcttccccag</u> GCTTGTTAGAGTGC TGTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTG <u>GTGGCCACTGGATTGTGTTTCTTTGGGGT</u> GGCACTGTTC TGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAG <u>CTAATTGAGACCTATTTCTCCAAAACTACCAAG</u> ACTAT GAGTATCTCATCAATGTgtaagtacctgccctccccacagaccatcttttttcc cctcttccatcctggagatagagaactcttcagtaccttagtaactagcaggggactgggtgga
119	アッセイ2およびプローブ4; qPCRアッセイ2=83bp; F2⇄CAC3=216bp:	agtgcctactatctccgagcctgtgagcacagggcctggcagaggggtttgagtgcatgagct <u>acctactggatgtgcctgactgtttcccttcttcttccccag</u> GCTTGTTAGAGTGC TGTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTG <u>GTGGCCACTGGATTGTGTTTCTTTGGGGT</u> GGCACTGTTC TGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAG

10

20

30

40

【化 3 - 2】

		<u>CTAATTGAGACCTATTTCTCCAAAACTACCAAGACTAT</u> <u>GAGTATCTCATCAATGT</u><u>gtaaglacctgccctccacacagacccatcttttttc</u> <u>ccctctccatcctggagatagagaactcttcagtaccttagtaactagcaggggactgggtgga</u>
120	プローブ2およびアッセイ5; qPCRアッセイ5=95bp; CAC3⇄R5=209bp:	agtgccactatctccgagcctgtgagcacagggcctggcagagggg<u>tttgagtggcatgagct</u> <u>acctactggatgtgcctgactgtttccctcttcttctccag</u>GCTTGTTAGAGTGC TGTGCAAG<u>ATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGT</u> <u>GGCCACTGGATTGTGTTTCTTTGGGGT</u>GGCACGTGTTCTGTGG CTGTGGAC<u>ATGAAGCCCTCACTGGCACAGAAAAGCTAAT</u> <u>TGAGACCTATTTCTCCAAAACTACCAAGACTAT</u>GAGTA TCTCATCAATGT<u>gtaaglacctgccctccacacagacccatcttttttccctctccat</u> <u>cctggagatagagaactcttcagtaccttagtaactagcaggggactgggtgga</u>
121	アッセイ4およびプローブ3; qPCRアッセイ4=97bp; F4⇄CAC3=204bp:	agtgccactatctccgagcctgtgagcacagggcctggcagagggg<u>tttgagtggcatgagct</u> <u>acctactggatgtgcctgactgtttccctcttcttctccag</u>GCTTGTTAGAGTGC TGTGCAAG<u>ATGTCTGGTAGGGGCCCCCTTTGCTTCCCTG</u> <u>GTGGCCACTGGATTGTGTTTCTTTGGGGT</u>GGCACGTGTTCT TGTGGCTGTGGAC<u>ATGAAGCCCTCACTGGCACAGAAAAGCT</u> <u>AATTGAGACCTATTTCTCCAAAACTACCAAGACTAT</u>GAGTA TCTCATCAATGT<u>gtaaglacctgccctccacacagacccatcttttttccctctccat</u> <u>cctggagatagagaactcttcagtaccttagtaactagcaggggactgggtgga</u>

10

20

30

【0390】

方法

プローブ：これらのアッセイでは、プローブのB10ユニバーサルオリゴセットを選択した（2012年8月24日の実験4：PLP1エキソン2の捕捉）。個々の捕捉プローブを作製するために、1μlのユニバーサルオリゴ10（100μM）を、100μMのUltramerプローブ1μlおよび98μlのTEz+0.05%のTween20と組み合わせた。これを、4μlから996μlのTEz+Tweenにさらに希釈して、4nMの作業溶液をもたらした。

【0391】

捕捉：捕捉のために、22ng/μlのgDNAライブラリー50μlと、4倍濃度の結合緩衝液20μlと、1μlのプローブと、9μlの水とを組み合わせた。6種の独立の捕捉反応物（プローブ1による2種の反応物、プローブ4による2種の反応物、プローブ2による1種の反応物、およびプローブ3による1種の反応物）が存在した。これらを、1分間にわたり95℃まで加熱し、次いで、前出で記載した通り、-1℃および2分間の35「サイクル」で60℃まで冷却した。アニーリングの後、10μlの洗浄されたビーズ（=1μlのビーズ原液）を添加し、RTで20分間にわたり結合をインキュベートした。次いで、ビーズを引き寄せ、洗浄緩衝液の200μlアリコートで、5分間ずつ4回にわたり洗浄した。最後の洗浄の後、全ての残りの接触可能な流体をビーズから吸引した。

40

【0392】

50

プロセッシング：ビーズを、 $10\mu\text{l}$ のquick blunt 溶液 ($200\mu\text{l} = 20\mu\text{l}$ の $10\times$ quick blunting buffer、 1mM のdNTP $20\mu\text{l}$ 、および $160\mu\text{l}$ の水)中に再懸濁させた。6つのビーズのアリコートの各々を、 $5\mu\text{l}$ ずつの2つのアリコートに分割した。酵素を伴わない $5\mu\text{l}$ のQB緩衝液を、チューブの1種のセット (これらは、捕捉だけのアリコートである)に添加した。他の $5\mu\text{l}$ のアリコートへは、 $0.025\mu\text{l}$ のT4ポリメラーゼを含有する $5\mu\text{l}$ のQB緩衝液 (これは、 $100\mu\text{l}$ のQB緩衝液を、 $0.5\mu\text{l}$ のT4ポリメラーゼと組み合わせ、 $5\mu\text{l}$ ずつのアリコートに分配することにより作製した)を添加した。捕捉だけのチューブおよび捕捉+プロセッシングのチューブの両方を、 20°C で15分間、 98°C で1分間にわたりインキュベートし、RTまで冷却し、速やかに磁石上に置いた。約 $10\mu\text{l}$ の上清を、捕捉だけの複合体およびT4-プロセッシングされた複合体の6種の対から取り出した (これで、合計12本のチューブ)。これらの上清を、下記に記載されるqPCRにおいて直接使用した。

10

【0393】

qPCR：これらのアッセイのために、標準的なTaq 反応ミックスおよび3ステップのサーマルサイクリングを選択した。各々が、

- $14\mu\text{l}$ の水
- $4\mu\text{l}$ の $10\times$ STD Taq 緩衝液
- 25mM のMgCl₂ $4\mu\text{l}$
- 各々 $10\mu\text{M}$ ずつのFプライマーとRプライマーとのブレンド $4\mu\text{l}$
- $8\mu\text{l}$ の鋳型 (上記から得られた上清)
- $2\mu\text{l}$ のDMSO
- 10mM のdNTP $2\mu\text{l}$
- $1\mu\text{l}$ のEvaGreen
- $0.8\mu\text{l}$ のROX
- $0.4\mu\text{l}$ のTaqポリメラーゼ

20

を含有する、 $40\mu\text{l}$ のqPCRミックス12種を構築した。

【0394】

反応物を4幅対に分配し、 94°C で30秒間、 55°C で30秒間、 72°C で60秒間の40サイクルにわたりサイクリングさせた。PCRの後、反応ミックスを、4幅対の4つのウェルの各々からプールし、 $5\mu\text{l}$ を2%のアガロースゲル上で解析した。

30

【0395】

結果

実験データを解釈するために、図14に示されるアガロースゲルを検討した。使用されるプライマー (など)と共に使用されるサイクリング条件下で、アッセイセット3、5、および6は、アッセイのアンプリコン (上のゲル) またはアダプターアンプリコンに照らしたプロセッシング後におけるPLP1 (下のゲル)と符合するPCR産物をもたらすことが観察された。より良好なアッセイセットは、

- アッセイ3を伴うプローブ4
- アッセイ5を伴うプローブ2
- アッセイ4を伴うプローブ3

40

に対応した。

【0396】

qPCRのCq値を、下記の表16に示す。アッセイ1および2は、ゲル解析が不良であった。良好なアッセイを、アッセイ3、5、および4に示す。プロセッシング%値を導出するため、Cqを、絶対値 (「Excel言語」では、絶対値=power (10, log10 (1/2) × Cq + 10))に変換した。次いで、プロセッシング後Cq/捕捉だけによるCqの商を、百分率として表した。この測定は、全てのアンプリコンの増幅効率が同じであり、理想化された検量線 (おそらく妥当な程度に正確な)に適合することを仮定した。次いで、これが適正であると仮定すると、捕捉された材料のうちの約10%がプロセ

50

シングされると考えられた。

【0397】

【表16】

表16: 捕捉後プロセシング感度についてのqPCR解析

	プローブ1	プローブ1	プローブ4	プローブ4	プローブ2	プローブ3
	アッセイ1	アッセイ2	アッセイ3	アッセイ2	アッセイ5	アッセイ4
捕捉			25		25	26
プロセシング後			29		28	29
プロセシング%			7%		17%	10%

10

【0398】

(実施例10)

拡張コード処理された男性gDNAライブラリーおよび女性gDNAライブラリーの構築目的

単回のMiSeqシーケンシングランで複数の捕捉パラメータについて調べるのに使用される、16種のコード処理された男性gDNAライブラリーおよび女性gDNAライブラリーのセットを構築する。

【0399】

方法

ステップ1: gDNA: 修復されたgDNAを調製した。

20

ステップ2: 全16種の可能なアダプターコードを作製した。これらのコードは、4種の塩基構造である。-4および-3位（インサートに照らして）の塩基位置は、ランダム塩基であり、-2および-1位の塩基位置は、試料コードである。4種の「クラスター」試料コードが存在する。これらは、

- クラスター1: AC、GA、CT、TG
- クラスター2: AA、GC、CG、TT
- クラスター3: AG、GT、CA、TC
- クラスター4: AT、GG、CC、TA

である。

【0400】

30

クラスター2~4は、プレート内の100μMのオリゴとして並べられた。プレートのうちの1種のセットは、ライゲーション鎖を有し、プレートのうちの1種のセットは、パートナー鎖を有した。プレートアレイは、A1-H1、A2-H2などであった。96ウェルPCRプレートの2種のセット内でアダプターをアニーリングさせるために、ウェル1つ当たり70μlの、68μlのTEzおよび2μlの5MのNaClを含有する「アニーリング溶液」を、テープで覆われた、20μlのパートナー鎖オリゴおよび10μlのライゲーション鎖オリゴに添加し、95℃で10秒間、65℃で5分間にわたりアニーリングさせ、RTまで冷却した。16種のコードのセット（同じ試料コードを有するランダムコード）を、4種のコードのセットにプールした。赤色=セットAA、GC、CG、およびTT。紫色=セットAG、GT、CA、およびTC。青色=セットAT、GG、CC、およびTA（この順序で並べる）。

40

【0401】

ステップ3: いずれの種類も、16種の独特のアダプター型の同じセットを受容する、女性DNAのための16種のライゲーション鎖および男性DNAのための16種のライゲーション鎖を創出することが最も容易である。これにより、後で、どの試料の組合せを創出しようと欲するかを、最大限に柔軟に決定することが可能となる。これを行うために、実験から得られる、末端修復されたgDNAを使用した。これにより、以下の通り、反応1回当たり20μl中の、必須の32回にわたるライゲーションが実行されるであろう。

1種を女性とし、1種を男性とする、2種のgDNAカクテルであって、

50

- － 144 μ l の水
- － 10 倍濃度のライゲーション緩衝液 32 μ l
- － 50 %のPEG8000 48 μ l
- － 64 μ l のgDNA

を含有するgDNAカクテルを作製した。

【0402】

カクテルを混合し、18 μ l ずつを伴う16本のチューブにアリコート分割した。2 μ l のアダプターおよび0.5 μ l のHC T4リガーゼを添加し、結果として得られる反応物を、22℃で60分間、65℃で10分間にわたりインキュベートし、RTまで冷却した。80 μ l のTEz、次いでまた、120 μ l のAmpureビーズも、反応物に添加し、混合し、RTで10分間にわたりインキュベートした。反応物を、70%のEtOH／水（v／v）200 μ l で、2回にわたり洗浄し、通気乾燥させ、100 μ l のTEz中に再懸濁させた。

10

【0403】

ステップ4：qPCR：

- － 175 μ l の水
- － 50 μ l の10×STD Taq緩衝液
- － 25mMのMgCl₂ 50 μ l
- － 100 μ l のACA2プライマー（10 μ M）
- － （50 μ l の鋳型：後で添加される）
- － 25 μ l のDMSO
- － 10mMのdNTP 25 μ l
- － 12.5 μ l のEva Green
- － 10 μ l のROX
- － 5 μ l のTaq DNAポリメラーゼ

20

を含有するqPCRマスターミックスを作製した。

【0404】

9 μ l を、Illumina Eco qPCRプレートの48のウェルに分配した。ライブラリー校正基準物質の、10 pg／ μ l および1 pg／ μ l の2種の系列希釈液を作製した。プレートの残りには、下記の表に示されるライブラリーをロードした。

30

【0405】

【表17】

表17:48ウェルアッセイプレート1についての試料凡例

	1	2	3	4	5	6	7	8
A	10 pg/ul STD	1 pg/ul STD	F1	F2	F3	F4	F5	F6
B	10 pg/ul STD	1 pg/ul STD	F1	F2	F3	F4	F5	F6
C	F7	F8	F9	F10	F11	F12	F13	F14
D	F7	F8	F9	F10	F11	F12	F13	F14
E	F15	F16	M1	M2	M3	M4	M5	M6
F	F15	F16	M1	M2	M3	M4	M5	M6

40

【0406】

第2のプレートは、下記の表18に示されるレイアウトを有した。

【0407】

【表 18】

表18: 48ウェルアッセイプレート2についての試料凡例

	1	2	3	4	5	6	7	8
A	100 pg/ul STD	10 pg/ul STD	1 pg/ul STD	NTC	F15	M1	M2	M3
B	100 pg/ul STD	10 pg/ul STD	1 pg/ul STD	NTC	F15	M1	M2	M3
C	M1	M2	M3	M4	M5	M6	M7	M8
D	M1	M2	M3	M4	M5	M6	M7	M8
E	M9	M10	M11	M12	M13	M14	M15	M16
F	M9	M10	M11	M12	M13	M14	M15	M16

【0408】

10

ライゲーション効率は、以下のサイクリングプログラム：

ー 72℃で2分間

ー 94℃で30秒間、60℃で30秒間、および72℃で60秒間の40サイクルにわたる

により測定した。

【0409】

結果

下記の表19は、基準物質および試料のCq値（(i)プレート2上で繰り返された実験、ならびに(ii)M1、M2、およびM3を2連のセット3種において測定した（3つの測定値の平均値を取った）ことを除き、2連の測定値の平均値）を示す。

20

【0410】

【表 19】

表19

生Ct平均値								
プレート1	9	13	9	10	9	10	9	9
	10	9	9	9	9	9	9	9
	14	9	9	9	9	9	9	9
プレート2	#DIV/0!	9	13	#DIV/0!	9	9	9	9
	9	9	9	9	9	9	10	9
	9	9	10	9	9	9	9	9
絶対値に変換したCt								
プレート1	1583	164	1984	1338	1739	970	2225	1756
	869	1729	1981	2142	1555	1751	2002	1688
		1729	2346	1947	1817	1545	2337	1697
プレート2	#DIV/0!	1469	151	#DIV/0!	1843	2218	1495	1489
	2308	1622	1885	1439	1834	1500	852	1733
	1822	2061	1340	1812	2100	1689	1857	2168
pg/ul単位に標準化したCt								
プレート1	10	1	13	8	11	6	14	11
	5	11	13	14	10	11	13	11
		11	15	12	11	10	15	11
プレート2	#DIV/0!	10	1	#DIV/0!	13	15	10	10
	16	11	13	10	12	10	6	12
	12	14	9	12	14	11	13	15
ライゲーションミックス1ul当たりのゲノム								
プレート1			4	2	3	2	4	3
	2	3	4	4	3	3	4	3
		3	4	4	3	3	4	3
プレート2					4	4	3	3
	5	3	4	3	4	3	2	3
	4	4	3	4	4	3	4	4

【0411】

これらを、Excel（青色の影）の式量=power(10, log10(1/2) × Cq + 8) により、任意の絶対値に変換した。次いで、絶対値に、10/1583（プレート1）または10/1469（プレート2）を乗じることにより、値を、公知の基準（赤色の影）に照らして標準化した。1μl当たりのゲノムは、7/8（アダプター質量に対応する）を乗じ、次いで、ゲノム1つ当たり3pgで除することにより計算した。ライゲーション効率を計算し（ライゲーション1回当たり20ngで、1/100が測定される=200pgがライゲーションされる）、計算された効率により、ライブラリーへの

【0412】

10

20

30

40

【表 20】

表20:ライブラリー濃度(1ul当たりのゲノム)

	1	2	3	4	5	6	7	8
A	10 pg/ul STD	1 pg/ul STD	F1	F2	F3	F4	F5	F6
B	10 pg/ul STD	1 pg/ul STD	4	2	3	2	4	3
C	F7	F8	F9	F10	F11	F12	F13	F14
D	2	3	4	4	3	3	4	3
E	F15	F16	M1	M2	M3	M4	M5	M6
F	4	3	4	4	3	3	4	3

	1	2	3	4	5	6	7	8
A	100 pg/ul STD	10 pg/ul STD	1 pg/ul STD	NTC	F15	M1	M2	M3
B	100 pg/ul STD	10 pg/ul STD	1 pg/ul STD	NTC	F15	4	3	3
C	M1	M2	M3	M4	M5	M6	M7	M8
D	5	3	4	3	4	3	2	3
E	M9	M10	M11	M12	M13	M14	M15	M16
F	4	4	3	4	4	3	4	4

【0413】

この実験の目標は、g DNAライブラリーを含有するライゲーションミックスを創出し、測定された数のゲノムを、マイクログラム量のライブラリー材料に増幅しうるように、ライゲーションミックス1 μ l 当たりのゲノム相当量を定量化することであった。上記の表20は、作製された各ライブラリーについて、1 μ l 当たりのゲノムを示す。下記に示される表21の目標は、表示される試料（ランダムな抜き取りにより採取された）を、後段の捕捉試験のための、10コピー、20コピー、40コピー、80コピーなどのライブラリーに転換することであった。表は、1 μ l 当たりのゲノム数を、表示のカバレッジ深度を達成するための、PCR反応1回当たりの μ l 数に転置する。表では、試料1例当たり200 μ l のPCRおよび40 μ l の鋳型投入量が仮定される。これらの実験は、実際のライブラリーを作製および精製する場合の指針として使用することができる。

【0414】

【表 21】

表21: 所望のゲノム表示に必要とされる各試料ライブラリーの容量の計算

試料	ゲノム/ul	# 必要とされるゲノム	得られるul	H2O	試料	ゲノム/ul	# 必要とされるゲノム	得られるul	H2O
F1	4	10.0	2.7	37.3	M1	ave			
F2	2	40.0	16.2	23.8	M2	ave			
F3	3				M3	ave	40.0	11.9	28.1
F4	2	80.0	44.8	4.8	M4	3			
F5	4	20.0	4.9	35.1	M5	4			
F6	3	20.0	6.2	33.8	M6	3			
F7	2	40.0	25.0	15.0	M7	2			
F8	3				M8	3	80.0	23.3	16.7
F9	4				M9	4			
F10	4	40.0	10.1	29.9	M10	4			
F11	3				M11	3	40.0	15.0	25.0
F12	3				M12	4	10.0	2.8	37.2
F13	4				M13	4	20.0	4.8	35.2
F14	3				M14	3			
F15	4				M15	4	40.0	10.9	29.1
F16	3				M16	4	20.0	4.6	35.4

【0415】

(実施例 11)

8 種の新規捕捉 q PCR アッセイの検証

目的

拡張プローブコレクションの捕捉効率を追求するようにデザインされた、8 種の新たな q PCR プライマーセットの性能を検証する。

【0416】

概要

8 種のアッセイ全ては、ヒト g D N A を増幅するのに使用される場合に予測されるサイズのアンプリコンをもたらした。X 染色体：1 5 4 3 7 6 0 5 1 領域（女性において 4 コピー、男性において 2 コピー）についての定量的解析は、観察されたコピー数と予測されるコピー数との驚くほど緊密な相関を示した。

【0 4 1 7】

方法

4 9 種のプローブ標的領域のサンプリングを表す、アッセイデザインのための 8 種のセグメントを選択した。アッセイをデザインするために、プローブの 5 ' 端から 2 0 0 b p 以内の D N A セグメントを同定した。標的領域の多少ともランダムな選択となるように、下記の表 2 2 に示される 8 種の領域を選択した。2 0 0 b p のセグメントを、本発明者らが、プライマーの T m を 6 5 °C（最適の T m）とし、プライマーの長さを 2 4 ヌクレオチド（最適の長さ）とする、5 0 ~ 1 0 0 b p のアンプリコンを指定する、P C R プライマー採取のプライマー 3 に切り出した。下記の表 2 2 は、領域および独特のゲノム属性、フォワード（F）プライマー配列およびリバース（R）プライマー配列、予測されるアンプリコンの長さ、ならびにゲノム配列の文脈における実際のアンプリコンを示す。

【0 4 1 8】

【表 2 2】

表22: ゲノム捕捉標的、およびそれらを精査するためのプライマー

配列 番号	アッセイ	領域	特徴	名称	配列	産物
122	9	CYP2D6	男性および女性 において4コピー	CYP2D 6_F	GGCTTCGACTGAACGTCTCCA	53 bp
123				CYP2D 6_R	AGTGCTCCATGGCTGCTCAGTT	
124	10	chrX- 1543760 51	女性において 4コピー、 男性において 2コピー	chrX- 15_1_F	GAACCAGAGGAACGCTGTGGT AACT	52 bp
125				chrX- 15_1_R	GACGTGTGCCTAGATGCGTTTTC	
126	11	chrX- 1543760 51	女性において 4コピー、 男性において 2コピー	chrX- 15_2_F	TGGCTGAAAAGTCTCCTTGAA ACTG	53 bp
127				chrX- 15_2_R	CTCAGTGGGTCTCCTTGAGAGAGGT	
128	12	chrX- 6929264	女性において 6コピー、 男性において 5コピー (19番染色体 において2コ ピー、X染色体 において1回)	chrX- 69_F	CTTTCCTCCAGTCACAAGCCA TCTA	78 bp
129				chrX- 69_R	CATGCTGATAGAAAGTCCCCTGGA	
130	13	KRAS 領域1	女性および男性 において2コピー	KRAS_ r1_F	TTTTCACACAGCCAGGAGTCT TTTC	66 bp
131				KRAS_ r1_R	GGGAGATCCGACAATACAGATTGAA	
132	14	KRAS 領域2	女性および男性 において2コピー	KRAS_ r2_F	CCTGTCTTGTCTTTGCTGATGT TTC	73 bp
133				KRAS_ r2_R	ACAAAACAGGCTCAGGACTTAGCAA	
134	15	MYC 領域2	女性および男性 において2コピー	MYC_r 2_1_F	CAACGTTAGCTTCACCAACAG GAAC	92 bp
135				MYC_r 2_1_R	GCTGGTAGAAGTTCTCCTCCTCGTC	
136	16	MYC 領域2	女性および男性 において2コピー	MYC_r 2_2_F	GTCTGCTCCACCTCCAGCTTGT A	93 bp
137				MYC_r 2_2_R	GTTGAGAGGGTAGGGGAAGACCAC	

【0419】

200 ng (2 ng/μl) の女性ゲノムDNAを含有する100 μlのPCR反応を実行することにより、各プライマー対の性能を探索した。反応物ミックスは、100 μl当たり、50 μlの水、10 μlの10× STD Taq緩衝液、25 mMのMgCl₂ 10 μl、各プライマーが10 μMで存在する10 μlのF+Rプライマーブレンド、10 μlの20 ng/μlのgDNA、5 μlのDMSO、10 mMのdNTP 5 μl、および1 μlのTaqポリメラーゼを含有した。反応物は、氷上で準備した。増幅は、94℃で30秒間、60℃で30秒間、および72℃で30秒間の30サイクルにわた

10

20

30

40

50

り実行し、これに続き、72℃で2分間にわたるインキュベーションを施し、10℃で保持した。5μlのPCR産物を、2%のアガロースゲル上で検討した。

【0420】

PCR産物は、残りのPCR産物95μlを、500μlのPBと組み合わせることにより、Qiagen PCR精製カラム上で精製した。材料は、カラムを通して、6KRPMで30秒間にわたりスピンし、750μlのPEで洗浄し、13.2KRPMでスピンのした。産物は、50μlのEBでカラムから溶出させ、Qubitにより定量化した。

【0421】

qPCR解析のために、chrX-154376051領域（アッセイ10および11）を、より詳細に検討した。精製PCR産物を、100fg/μl、10fg/μl、および1fg/μlに希釈した。ゲノムDNAを、10ng/μlに希釈した。2マイクロリットルの基準物質またはgDNAを、48ウェルEco qPCRプレートのウェル1つ当たり8μlのPCRマスターミックスと組み合わせた。マスターミックスは、500μlの最終反応容量（鋳型の添加に対応する）当たり、175μlの水、50μlの10×STD Taq緩衝液、25mMのMgCl₂ 50μl、10μMのF+Rプライマーブレンド50μl、25μlのDMSO、10mMのdNTP 25μl、12.5μlのEvaGreen、10μlのROX、および5μlのTaqポリメラーゼを含有した。32μlのミックスを、16のウェルに分配し、8μlの鋳型を添加した。次いで、これらを、4幅対でqPCRプレートに分配した。プレートのレイアウトを、下記の表23に示す。

【0422】

【表23】

表23: chrX-154376051領域についてqPCR精査するためのプレートレイアウト

	1	2	3	4	5	6	7	8
A	100 fg PCR prod		10 fg PCR prod		1 fg PCR prod		NTC	
B								
C	100 fg PCR prod		10 fg PCR prod		1 fg PCR prod		NTC	
D								
E	F-gDNA		M-gDNA		F-gDNA		M-gDNA	
F								
		アッセイ 10						
		アッセイ 11						

【0423】

結果および考察

ゲノムDNAから増幅されるPCR産物についてのゲル解析は、8種のPCR反応物全てが、予測されたサイズの独特の産物をもたらすことを示した（データは示さない）。アンプリコンは、十分に清浄であり（剰余バンドを伴わず、余剰プライマーを伴わない）、定量的解析のための検量線を作成するのに有用であった。アンプリコンを、Qiagen

PCRスピнкаラムを使用して精製し、産物を50μl中に溶出させた。産物収率は、アッセイ9-18.4ng/μl；アッセイ10-26.1ng/μl；アッセイ11-13.9ng/μl；アッセイ12-26.6ng/μl；アッセイ13-7.9ng/μl；アッセイ14-19.2ng/μl；アッセイ15-23.1ng/μl；およびアッセイ16-20.4ng/μlであった。

【0424】

定量的解析は、女性が4コピーを有し、男性が2コピーを有するように、X染色体上の潜在性セグメント重複に対応する、アッセイ10および11について実行した。

【0425】

平均C_q値を、下記の表24に示す。これらを使用して、示される検量線を作成した。

2つの反応は、基本的に重ね合せ可能であった。これらの曲線を使用して、本発明者らは、検量線ウェルおよびゲノム投入ウェルの両方について、フェムトグラム単位の絶対量を計算した。データは、表24の検量線データの下に示す。

【0426】

【表24】

表24: アッセイ10および11により解析された基準試料のC_q値

検量線 10	13	17	20	31
検量線 11	13	16	20	33
女性試料／男性試料	21	22	21	22

10

【0427】

本実施例の1つの要点は、定量的分子生物学の効力を強調することであった。この実験で、2 μlの基準物質を添加し、サンプリングしたことは、1 fg/μlの基準物質が、qPCR反応物中に、実際に2 fg存在することを意味する。これは、アッセイ10の53 bpの断片による分子17,500個に対応する。20 ngのゲノムDNAを、反応物に投入した。これは、6667ゲノム相当のDNAに対応する。ゲノムDNAを、200 bpの平均サイズに断片化したことは、標的領域のうちの75%だけが、インタクトを維持することを意味する。よって、gDNAは、約5000の「qPCR実行可能」ゲノムコピーを有した。最後に、男性では、予測されるゲノム1つ当たりのX染色体の重複領域は平均1コピーであり、女性では、予測される平均は2コピーであった。観察される分子の数だけ発生した観察値と対比した予測値は以下：男性予測値=5000コピー；男性観察値=3500コピー；女性予測値=10000コピー；および女性観察値=7000コピーの通りであることが分かった。

20

【0428】

【表25】

表25: 予測値対観察値

検量線 10	102.8004	9.421966	1.02996
女性試料／男性試料	0.447179	0.204926	
検量線 11	98.27451	10.35326	0.983644
女性試料／男性試料	0.388559	0.174757	

30

【0429】

(実施例12)

さらなる捕捉後プロセッシング戦略

目的

捕捉後プロセッシングを達成するための代替法（図15を参照されたい）を開発した。

【0430】

概要

再デザインされたプローブにより実行された捕捉後プロセッシングステップは、既に頑健な捕捉を、さらに5～9倍増強すると考えられた。全般的に、試験は極めて成功した。

【0431】

背景

アッセイデザインの他の実施形態では、PCRプライミング部位を付加する、プローブ

40

50

テイル配列をコピーする前に、クローンの3'端において、エキソヌクレアーゼステップを使用することが想定された。特定の実施形態では、クローンにプローブをコピーさせることから、プローブにクローンをコピーさせることにシフトすることがさらに想定された。この極性の逆転は、本発明者らが、プローブの5'端を、プルダウン配列およびリバーサPCRプライマー配列のいずれとしても使用することを意味する。プローブの3'端は、非修飾のまま放置し、ついで、DNAポリメラーゼを使用してクローンをコピーすることができる。構想として、この手法にはいくつかの利点がある。まず、エキソヌクレアーゼ活性および重合化の両方を必要とするステップから、単純な重合化ステップへのシフトがなされたため、このステップを、PCRと協奏させて施すことができる。さらに、このステップは、熱安定性のポリメラーゼ酵素により、72℃で施しうることから、一本鎖クローンの潜在的な二次構造がそれほど問題とならないことも意味する。最後に、プローブは、114ヌクレオチドから95ヌクレオチドに短縮され、これにより、費用節減の利点もたらされることが含意された。

【0432】

4種の良好に挙動したqPCRアッセイ（実施例11：8種の新たな捕捉qPCRアッセイの検証）である、アッセイ10、14、15、および16は、これらのアッセイを「指さす」プローブとマッチした。プローブと、qPCRアッセイとは、互いの近傍内にあったが、それらのDNA配列が互いと重複しなかったことは重要である（図16を参照されたい）。プローブ配列および対応するアッセイを、下記の表26および27に示す。

【0433】

【表26】

表26: ゲノム捕捉標的およびそれらを精査するためのプライマー

配列番号	アッセイ	領域	特徴	名称	配列	産物
138	10	chrX-154376051	女性において4コピー、男性において2コピー	chrX-15_1_F	GAACCAGAGGAACGCTGTGGTAACT	52 bp
139				chrX-15_1_R	GACGTGTGCCTAGATGCGTTTTC	
140	14	KRAS領域2	女性および男性において2コピー	KRAS_r2_F	CCTGTCTTGTCTTTGCTGATGTTTC	73 bp
141				KRAS_r2_R	ACAAAACAGGCTCAGGACTTAGCAA	
142	15	MYC領域2	女性および男性において2コピー	MYC_r2_1_F	CAACGTTAGCTTCACCAACAGGAAC	92 bp
143				MYC_r2_1_R	GCTGGTAGAAGTTCTCCTCCTCGTC	
144	16	MYC領域2	女性および男性において2コピー	MYC_r2_2_F	GTCTGCTCCACCTCCAGCTTGTA	93 bp
145				MYC_r2_2_R	GTTGAGAGGGTAGGGGAAGACCAC	

【0434】

【表 27】

表27: 捕捉プローブ配列およびゲノム標的

配列番号	アッセイ	領域	60マーのプローブ配列
146	10	chrX- 154376051_2x_chrX:154376979; 領域 _1:1011nt:929:988:-	AGAATTCATTGCCAGCTATAAAATCTG TGGAAACGCTGCCACACAATCTTAGC ACACAAGA
147	14	KRAS_NM_033360_chr12:2537 8485: 領域 _2:161nt:-62:-3:+	TAAATGACATAACAGTTATGATTTTG CAGAAAACAGATCTGTATTTATTCA GTGTTACT
148	15	MYC_NM_002467_chr8:128750 431: 領域 _2:773nt:-62:-3:+	AGGTTTCCGCACCAAGACCCCTTTAA CTCAAGACTGCCTCCCGCTTTGTGTG CCCCGCTC
149	16	MYC_NM_002467_chr8:128751 122: 領域 _2:773nt:630:689:-	AGAGCAGAGAATCCGAGGACGGAGA GAAGGCGCTGGAGTCTTGCGAGGCG CAGGACTTGG

10

【0435】

方法

20 μ l ずつのライゲーションミックスを、合計 80 μ l に組み合わせ、合計 800 μ l 中で増幅することにより、gDNAライブラリーを、試料 F13～F16（実施例10）から作製し直した。ビーズを洗浄して400 μ l とし、Qubitにより、32 ng/ μ l のプール濃度を測定した。

20

【0436】

下記に列挙されるIDT製のオリゴを、100 μ Mに再懸濁させた。Ultramerは、4ナノモルで得られるので、これらを40 μ l のTEzero中に懸濁させた。4種の試験プローブの、2 μ l ずつ4つのアリコート、8 μ l の100 μ Mのユニバーサルトイル配列（全長リバースプライマー9の最初の35塩基から得られる）と組み合わせて、50 μ Mの二重鎖のチューブをもたらした。この二重鎖を、10 μ l から、990 μ l のTEzero+Tweenに希釈して、500 nMとし、10 μ l から、990 μ l に再度希釈して、5 nMとした。

【0437】

組み合わせられた40 μ l のgDNAを、4倍濃度の結合緩衝液15 μ l および5 μ l の捕捉二重鎖と組み合わせた。反応物ミックスをアニーリングさせ、2 μ l の洗浄されたMyOneストレプトアビジンコーティングビーズ上で捕捉した。反応物を、洗浄緩衝液で4回にわたり洗浄し、洗浄緩衝液をビーズペレットから吸引した。捕捉単独試料を測定するため、1種のビーズペレットを、単一のPCRプライマーであるACA2を含有する、100 μ l のPCRミックス中に再懸濁させた。捕捉+プロセッシング試料を測定するため、別種のビーズペレットを、全長ACA2フォワードプライマー（オリゴ8）および全長CAC3リバースプライマー（オリゴ9）を含有する、100 μ l のPCRミックス中に再懸濁させた。後者の試料を、72 $^{\circ}$ Cで2分間にわたりインキュベートした。両方の試料を、94 $^{\circ}$ Cで30秒間、60 $^{\circ}$ Cで30秒間、および72 $^{\circ}$ Cで60秒間の25サイクルにわたり増幅した。72 $^{\circ}$ Cで2分間にわたり保持し、RTまで冷却した後、ビーズ上でPCRアンプリコンを精製し、50 μ l のTEzero中に再懸濁させた。

30

40

【0438】

qPCRのために、レポーター色素としてのEvaGreen、基準色素としてのROX、ならびに94 $^{\circ}$ Cで30秒間、60 $^{\circ}$ Cで30秒間、および72 $^{\circ}$ Cで60秒間の40サイクルにわたる3ステップPCRを使用して、試料を、アッセイ9～16（アッセイ10、14、15、および16は標的である）でアッセイした。元のgDNAライブラリーは、2 ng/ μ l の最終濃度で存在した。捕捉された試料および捕捉+プロセッシング試料は、2 pg/ μ l の最終濃度（TEzero+0.05%のTween20中で希釈された）で存在した。

50

【0439】

結果および考察

捕捉だけによるPCR収率は、 $27.8 \text{ ng}/\mu\text{l}$ であり、捕捉+プロセッシングによるPCR収率は、 $40.4 \text{ ng}/\mu\text{l}$ であった。これらの頑健な収率により、増幅は完全であることが指し示された。2%のアガロースゲル画像により、出発投入ライブラリー、捕捉されたライブラリー、および捕捉+プロセッシングライブラリーを示す（図17）。プロセッシングが働いた場合、ライブラリーの平均インサートサイズは減少するはずであり、実際に減少した。ライブラリーの下端が多少とも「バンド」をなすという事実により、プローブのある程度のプライミングオフが生じうることが指し示される。このフォーマットでは、本発明者らのプローブの3'端が露出されているため、残留する、結合していないブ

10

【0440】

この実験における重要な計量は、qPCRによる、捕捉感度および捕捉特異度の測定であった。qPCRデータを、下記の表28に示す。

【0441】

【表28】

表28: 捕捉感度および捕捉特異度についてのqPCR解析

生のCq	アッセイ 9	アッセイ 10	アッセイ 11	アッセイ 12	アッセイ 13	アッセイ 14	アッセイ 15	アッセイ 16
gDNA lib	20	19	20	19	20	19	20	20
1st cap	20	15	23	21	20	15	16	15
cap + proc	23	13	27	20	20	12	13	13
Abs 値	アッセイ 9	アッセイ 10	アッセイ 11	アッセイ 12	アッセイ 13	アッセイ 14	アッセイ 15	アッセイ 16
gDNA lib	122	188	114	211	76	184	85	79
1st cap	72	3240	15	59	85	2305	1231	2197
cap + proc	11	16518	1	69	97	19504	11189	13275
希釈率につい ての調整	アッセイ 9	アッセイ 10	アッセイ 11	アッセイ 12	アッセイ 13	アッセイ 14	アッセイ 15	アッセイ 16
gDNA lib	122	188	114	211	76	184	85	79
1st cap	71606	3239810	14835	58862	84616	2304816	1230995	2196798
cap + proc	10793	16517508	713	69047	97429	19504350	11189209	13274986
濃縮倍数	アッセイ 9	アッセイ 10	アッセイ 11	アッセイ 12	アッセイ 13	アッセイ 14	アッセイ 15	アッセイ 16
1st cap	585	17261	131	279	1108	12530	14513	27753
cap + proc	88	88000	6	328	1276	106034	131921	167709
c+p/c	0	5	0	1	1	8	9	6

20

30

【0442】

特異度に関しては、標的化された領域（明るいグレーの強調）だけが、著明な濃縮を呈示した。さらに、プロセッシングされたライブラリーは、全ての標的領域について、比活性の、捕捉単独と比べて著明な増大を示した。これらのデータにより、このさらなるプローブデザインの実施形態ならば、効率的な捕捉後プロセッシングに使用しうることが指し示された。

40

【0443】

(実施例13)

捕捉後プロセッシング戦略についての配列解析

目的

この実験の目的は、シーケンシングライブラリー内の標的領域の濃縮およびカバー

50

ジを評価することであった。

【0444】

概要

標的配列の濃縮およびフォーカシングのレベルは、ハイブリダイゼーションベースの捕捉を、酵素プロセッシングとカップリングすることにより、捕捉単独と比較して劇的に改善された。

【0445】

背景

本明細書で開示されるかつての実験により、捕捉後プロセッシングは、標的含有および qPCR により測定される濃縮ライブラリーの比活性を増大させることが実証されている。この実験では、次世代 DNA シークエンシングを使用して、捕捉単独または代替のプロセッシング法により作製されたライブラリー内の標的配列の表示および分布を比較した。

【0446】

方法

特異的遺伝子 (KRAS、MYC、PLP1、CYP2D6 および AMY1) 内の部位および X 染色体上の重複領域を標的化する、49 種の捕捉プローブのセットを使用して、2 種の濃縮ライブラリープールを、男性ヒトゲノム DNA と女性ヒトゲノム DNA との均等なミックスから構築した。プローブ配列を、下記の表 29 に示す。

【0447】

【表 29-1】

表29: 捕捉感度および捕捉特異度についてのqPCR解析

配列番号	番号	プローブ名	プローブ配列
150	1	CYP2D6_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC AAGCACCTAGCCCCATTCTGCTGAGCAGGAGGTG GCAGGTACCCACAGACTGGGAGGTAA
151	2	CYP2D6_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC AGTCGGTGGGGCCAGGATGAGGCCAGTCTGTTCAC CACATGGCTGCTGCCTCTCAGCTCT
152	3	AMY1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC ACCTGAGTAGCATCATTGTAGTTCTCGATATCTCCA CTTCCAGTTTTACATTACCATCA
153	4	chrX_15_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC CCTGGCCCTCAGCCAGTACAGAAAGTCATTTGTCA AGGCCTTCAGTTGGCAGACGTGCTC
154	5	chrX_15_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC AGAATTCATTGCCAGCTATAAATCTGTGGAAACGC TGCCACACAATCTTAGCACACAAGA
155	6	chrX_477_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GACTTCAAAGAAATTACAAGTTGACATCTTGGACT CTACCCCTCGTACTTTATCTCCTAT
156	7	chrX_477_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TCTCTTTGGGGTCAAGAAAGAATCCCTAGTGGATT TGGGATTCTAGAGGAGGTGTTATAA
157	8	chrX_478_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TGCGATACCATGCTGAAGATGAGCTAACCCAACCA GCCAAGCAGGCAGGGCTGCGAAGGA

【表 29-2】

158	9	chrX_478_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GGGGTAGGTGGAAAACCCAAGTAATGTGATTTTGT AACATCCACTGCTGCATTTGTTTGC
159	10	chrX_69_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TTACTTCCCTCCAGTTTGTGCTTGCAAAACAACA GAATCTTCTCTCCATGAAATCATG
160	11	chrX_69_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC CAGGGGTATCTATTATCCCCATTTTCTCACAAAGG AAACCAAGATAAAAGGTTTAAATGG
161	12	PLP1_ex1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GAAATTCTCTTGTGAATTCCTGTGTCCTCTTGAATC TTCAATGCTAAAGTTTTTGAAGT
162	13	PLP1_ex2_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GGGTTTGAGTGGCATGAGCTACCTACTGGATGTGC CTGACTGTTTCCCTTCTTCTTCCC
163	14	PLP1_ex2_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC CTATCTCCAGGATGGAGAGAGGGAAAAAAAAGAT GGGTCTGTGTGGGAGGGCAGGTACTT
164	15	PLP1_ex3_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GAAAGAAGCCAGGTCTTCAATTAATAAGATTCCT GGTCTCGTTTGTCTACCTGTTAATG
165	16	PLP1_ex3_M	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC CAGACTCGCGCCCAATTTTCCCCCACCCTTGTTAT TGCCACAAAATCCTGAGGATGATC
166	17	PLP1_ex3_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TCTTTCTTCTTCCTTTATGGGGCCCTCCTGCTGGCT GAGGGCTTCTACACCACCGGCGCA
167	18	PLP1_ex4_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GTTTGTGTTTCTACATCTGCAGGCTGATGCTGATTT CTAACCACCCCATGTCAATCATTT
168	19	PLP1_ex4_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC AACCAAATATATAGTGCTTCCATAGTGGGTAGGAG AGCCAAAGCACCCGTACCCTAACTC
169	20	PLP1_ex5_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC AGTCTCCATGTGGCCCCGTAAGTCCATAAAGCTTA CCCTGCTTGCTTTTTGTGTCTTACT
170	21	PLP1_ex5_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC CCATGGGTGTAAATTTGTATGGTATTAGCTACTCCCT TGTAATAATAACCCAAATAACCCAC
171	22	PLP1_ex6_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TTTACAGTGGAGCATATTACTGCTGTTGCAAGAAA CAGTTCTTCCCTCTTTCATTTTCTG
172	23	PLP1_ex6_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC ATAGCTGTACCCACACTATCTCAGGCCTATTTACTT GCCAAGATCATTCAAAGTCAACTC

10

20

30

40

【表 29 - 3】

173	24	PLP1_ex7_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GATTTGAGGAGGGAGTGCTTTCTTTTCTACTCTCAT TCACATTCTCTCTTCTGTTCCCTA
174	25	PLP1_ex7_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC CAGCATTGTAGGCTGTGTGGTTAGAGCCTCGCTAT TAGAGAAAGGGGGATTCTACGGGG
175	26	KRAS_ex1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TGTTACCTTTAAAAGACATCTGCTTTCTGCCAAAAT TAATGTGCTGAACTTAAACTTACC
176	27	KRAS_ex1_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TTCCCAGTAAATTACTCTTACCAATGCAACAGACT TTAAAGAAGTTGTGTTTTACAATGC
177	28	KRAS_ex2_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TAAATGACATAACAGTTATGATTTTGCAGAAAACA GATCTGTATTTATTTTCACTGTTACT
178	29	KRAS_ex2_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GACAGGTTTTGAAAGATATTTGTGTACTAATGAC TGTGCTATAACTTTTTTTTCTTTCC
179	30	KRAS_ex3_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC ACTCAAAAAATAAAAACTATAATTACTCCTTAATG TCAGCTTATTATATTCAATTTAAAC
180	31	KRAS_ex3_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC AACACCTTTTTTGAAGTAAAAGGTGCACTGTAATA ATCCAGACTGTGTTTCTCCCTTCTC
181	32	KRAS_ex4_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GAAACCTTTATCTGTATCAAAGAATGGTCCTGCAC CAGTAATATGCATATTAAAACAAGA
182	33	KRAS_ex4_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GTGTATTAACTTATGTGTGACATGTTCTAATATAG TCACATTTTCATTATTTTTATTAT
183	34	MYC_r1_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC CCCCAGCCAGCGGTCCGCAACCCTTGCCGCATCCA CGAAACTTTGCCCATAGCAGCGGGC
184	35	MYC_r1_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC CGACTCATCTCAGCATTAAAGTGATAAAAAAATAA ATTAAAAGGCAAGTGGACTTCGGTG
185	36	MYC_r2_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC CTGTGGCGCGCACTGCGCGCTGCGCCAGGTTTCCG CACCAAGACCCCTTTAACTCAAGAC
186	37	MYC_r2_F2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TTCTACTGCGACGAGGAGGAGAACTTCTACCAGCA GCAGCAGCAGAGCGAGCTGCAGCCC
187	38	MYC_r2_F3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC ACCGAGCTGCTGGGAGGAGACATGGTGAACCAGA GTTTCATCTGCGACCCGGACGACGAG

10

20

30

40

【表 2 9 - 4】

188	39	MYC_r2_F4	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GCCGCCGCTCAGAGTGCATCGACCCCTCGGTGGT CTCCCCCTACCCTCTCAACGACAGC
189	40	MYC_r2_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GGCGGCTAGGGGACAGGGGCGGGGTGGGCAGCAG CTCGAATTTCTTCCAGATATCCTCGC
190	41	MYC_r2_R2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC AGACGAGCTTGGCGGCGGCCGAGAAGCCGCTCCA CATACAGTCCTGGATGATGATGTTTT
191	42	MYC_r2_R3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC AGGAGAGCAGAGAATCCGAGGACGGAGAGAAGGC GCTGGAGTCTTGCGAGGCGCAGGACT
192	43	MYC_r2_R4	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TAAGAGTGGCCCGTTAAATAAGCTGCCAATGAAAA TGGGAAAGGTATCCAGCCGCCCACT
193	44	MYC_r3_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TTGTATTTGTACAGCATTAACTGGTAATTGATTAT TTAATGTAACCTTGCTAAAGGAG
194	45	MYC_r3_F2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GAGGCCACAGCAAACCTCCTCACAGCCCACTGGTC CTCAAGAGGTGCCACGTCTCCACAC
195	46	MYC_r3_F3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC AGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCC CTGCGTGACCAGATCCCGGAGTTGG
196	47	MYC_r3_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC TCCAACCTTGACCTCTTGGCAGCAGGATAGTCCTT CCGAGTGGAGGGAGGCGCTGCGTAG
197	48	MYC_r3_R2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GCTTGGACGGACAGGATGTATGCTGTGGCTTTTTT AAGGATAACTACCTTGGGGGCTTTT
198	49	MYC_r3_R3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCAC GCATTTGATCATGCATTTGAAACAAGTTCATAGGT GATTGCTCAGGACATTTCTGTTAGA

10

20

30

【0 4 4 8】

第1のライブラリープールは、実施例12において「捕捉+プロセッシング」ライブラリーについて記載した通りに作製した。第2のプールは、以下の改変を除き、実施例12において「捕捉だけ」ライブラリーについて記載した通りに作製した。捕捉PCRの後、第2ラウンドのPCRを実行して、単一のプライマーであるACA2で増幅されたライブラリーを、二重プライマーによる、Illuminaシーケンシングに適する異種末端ライブラリーに転換した。これを行うため、ライブラリーを希釈し、以下のプライマー：

40

プライマー55

AATGATACGGCGACCACCGAGATCTACACGTCATGCAGGAC
CAGAG (配列番号199) および

プライマー56

CAAGCAGAAGACGGCATACGAGATGTGACTGGCACGGGAG
TTGAGAAATTCGAATACA (配列番号200)

により再増幅した。

【0 4 4 9】

50

100 μ l の反応ミックスは、40 ng のライブラリー、10 μ l の10 $\times$ STD Taq 緩衝液、25 mM のMgCl₂ 10 μ l、いずれも10 μ Mである10 μ l の5プライマーおよび10 μ l の56プライマー、5 μ l のDMSO、5 μ l のdNTP、ならびに1 μ l のTaq DNAポリメラーゼを含有した。試料は、94℃で30秒間、50℃で30秒間、52.5℃で30秒間、55℃で30秒間、57.5℃で30秒間、60℃で30秒間、72℃-1分間の2サイクルにわたり増幅した。次いで、試料を、94℃で30秒間、60℃で30秒間、および72℃で60秒間の8サイクルにわたり増幅するの続き、72℃で2分間を施した。PCRミックスをビーズで精製し、50 μ l ずつに再懸濁させた。

【0450】

10

結果および考察

Illumina MiSeq Personal Sequencerを使用して、両方のプールを解析した。各ライブラリープールから得られる50ヌクレオチドの配列リードをトリミングして、4塩基のバーコード配列を除去し、Bowtie配列アラインメントプログラムを使用して、ヒトゲノムの参照配列(hg19変化形)にマップした。両方のライブラリー内のリードのうちの約80%が、参照配列と明確に整列した。整列されたリードのさらなる特徴付けは、ハイブリダイゼーションベースの捕捉を、酵素プロセシングとカップリングすることにより、4.9キロ塩基の標的領域の、投入されたゲノムDNAと比べた979,592倍の濃縮が結果としてもたらされることを明らかにした。これは、ライブラリー含量の、プロセシングされない「捕捉だけ」の手法と比較した3倍の改善を表した。全般的に、この代替的プロセシング法により得られる5種の配列のうちの約4種が、捕捉プローブにより特異的に標的化されるゲノム部位にマップされた。

20

【0451】

各ライブラリープールについてのアラインメント統計の概要を、下記の表30に示す。

【0452】

【表30】

表30: シークエンシング結果

濃縮プロトコール	ゲノム(hg19)と整列するリードの数	標的領域と整列するリードの数	標的領域内で整列されたリードの%
捕捉だけ	4,122,637	1,031,370	25.0%
捕捉+プロセシング	6,949,695	5,433,442	78.2%

30

【0453】

各ライブラリープールから得られるリードはまた、UCSC Genome Browserによっても表示して、標的部位近傍の局所的配列カバレッジおよび配列分布を評価した。X染色体の2つのセグメントの拡大画像は、プロセシングされたライブラリーにより、「捕捉だけ」ライブラリーより標的化部位内で高度に濃縮された配列カバレッジがもたらされることを示す(図18)。さらに、標的領域への配列マッピングは、プロセシングされたライブラリーにおいて、プロセシングされていない対照より均一に分配された。まとめると、これらのデータにより、代替的プロセシング法により、濃縮されるライブラリー内に存在する標的配列の量および品質が劇的に改善されることが指し示された。

40

【0454】

(実施例14)

バイオインフォマティクス

総括

従来の次世代シーケンシング(NGS)解析は、「パーティカル」である。本明細書

50

で想定される本発明の分子の独特のデザインは、臨床リシーケンシング手法を革新する、「ホリゾンタル」法を可能とする。

【0455】

配列アラインメントに関連して本明細書で用いられる「バーティカル」とは、図19により例示される手法を指す。インフォマティクス解析へのかつての手法は、短いリードを参照ゲノムと整列させる、第1のステップを伴う。アラインメントの後、重複するリードを、SNV（一塩基変異体）を指し示しうる塩基変化について解析する。本明細書では、リードの垂直的な積み重ねとして描示されることが多いアラインメントに依拠するため、手法を「バーティカル」という愛称と呼ぶ。多様なプログラムが、SNVおよびインデル（挿入／欠失）を許容するが、コアとなる手法は、アラインメント認識ベースである。

10

【0456】

これに対し、本明細書で想定される方法により得られるペアエンドリードデータは、リード1内のDNAタグ付けされた配列情報およびリード2内のプローブID情報を有するであろう。データ解析における第1のステップは、リードをプローブにマッチさせるステップである。ステップ2は、各プローブと「ホリゾンタルに」関連する配列情報を解析するステップである。例えば、図20を参照されたい。

【0457】

リード深度が十分であるとき、ホリゾンタルな、プローブベースの配列の会合は、アラインメントに依拠しない。そうではなくて、リードは、デノボのコンティグにアセンブルすることができる。方法の利点は、いずれも、従来の、アラインメントベースの方法が、検出するのに難航し、最も多くの困難を経る状況である、短い配列の連なりにおける挿入／欠失および複数の配列変化に対して極度に頑健なことである。さらに、ホリゾンタルな会合を、プローブおよびタグ付けと組み合わせることにより、より正確な仮説の提起（すなわち、観察される配列変異体が、真の配列変異体である可能性が高いのか、偽の配列変異体である可能性が高いのかの決定）が容易となる。

20

【0458】

CNVおよび構造的変異I

大スケールのコピー数変異（CNV：copy number variation）解析では、方法は、捕捉された配列領域と関連する独特のリードの数の決定を含む。観察されるCNVの大多数は、2～100bpの長さのオーダーにある塩基の挿入および欠失を伴う、「マイクロCNV」である。バーティカルアラインメント法は、それらが、多数の偽陽性仮説を促進する、アラインメントの厳密性緩和を必要とするため、マイクロ挿入／欠失（インデル）で難航する。ホリゾンタル法およびデノボコンティグアセンブリーは、アラインメントパラメータのこのような緩和を必要とせず、構造的変異の把握を要求する。

30

【0459】

図21に例示される、エキソンの1種の対立遺伝子内の小規模な挿入という単純な場合について検討しよう。本実施例では、ホリゾンタルアラインメントにより、リードの、プローブ1およびプローブ2との会合を「強いる」。アセンブリーは、1種は野生型のエキソン構造を伴い、1種は挿入構造を伴う、2種のコンティグを作製するであろう。この解析から、2つの原理が立ち現れる：1）隣接するプローブから得られる重複リードは、捕捉されたエキソンのインデル含有対立遺伝子についての仮説に対する実証または反証となり、2）捕捉プローブの外側のマイクロCNV対立遺伝子は、ホリゾンタル法によりたやすく検出可能である。

40

【0460】

CNVおよび構造的変異II

CNVの検証は、バーティカルアラインメント法を伴うことが多い。これらの研究では、参照配列に照らして完全であることが典型的なアラインメントが要求される。基準と異なる、リードを横切るSNVを棄却するこのような方法は、SNV（一般的なSNPなど）に対して脆弱である。正味の結果は、コピー数の常習的な過小評価となろう。先に進む

50

には、本発明の方法により可能なホリゾンタル法を使用すべきである。

【0461】

S N Vについてのホリゾンタルな仮説の検定 I

S N Vを検出するための、パーティカルな、アラインメントベースの方法は、解析が困難である。単一の塩基に関与するホモ接合性変異体の対立遺伝子は、同定するのがごく単純であるが、これらの変化はまれである。より一般には、S N Vは、ヘテロ接合性であり、変異体は、いくつかの連続的な位置で生じる場合もあり、近接した間隔を置いた位置で生じる場合もある（エラープロード修復は、複数の塩基がコンセンサスでない位置を追跡する傾向がある）。49種のリードがS N Vを保有し、47種のリードが野生型の基準塩基を保有する場合（厳密に仮説的な例としての）、ヘテロ接合性のS N V仮説は、真の高度なカバレッジ検出から逸脱する。リード深度が浅く、WTのリードと対比したS N Vのリードの数が、50/50から著明に解離する（例えば、8種がWTであり、2種が変異体である、合計10種のリード）場合、判定は、さらにより思弁的となる。直交的な検証にかけられる仮説は、不変的に任意のカットオフを受ける。

10

【0462】

ホリゾンタルプローブベースの会合をタグと組み合わせる特定の実施形態では、S N V仮説におけるはるかに大きな精度が達成される。単一のタグ（タグ＝コード＋端点）上に存在するS N Vは、とりわけ、同じタグ内のリードがWTである場合には無視される。例えば、図22を参照されたい。

20

【0463】

S N Vについてのホリゾンタルな仮説の検定 I I

リードの開始部位が同一であってもなお、2つの異なるタグ上で生じるS N V仮説（A）、または同じプローブとホリゾンタルに会合する、異なるリード上で生じるS N V仮説（B）、または同じエクソンにおける異なるプローブの会合から生じるS N V仮説（C）は、必然的に、真剣に検討しなければならない仮説である。例えば、図23を参照されたい。

【0464】

（実施例15）

分子注釈

総括

30

本実施例は、シーケンシングライブラリーについての「分子注釈」（図24）と、結果として得られるシーケンシング情報を査定するのに後続のステップで使用されるインフォマティクスとの相互作用について記載する。プローブから得られるリバースリードは有用である。DNAプローブ配列の領域を決定するリバースリード2は、全ての後段における解析検討において、著明に有用である。例えば、有用性は、変異体の判定において見出すことができ、ここから得られる成果は、コピー数の決定において有用である。データ解析のこれらの2つの側面について、下記で記載する。

【0465】

リード__2のプローブ配列

プローブセットとは、1種もしくは2種のプローブ、なおまたは数万種のプローブを含みうる、独特の、公知の配列コレクションである。これは、リード__2を使用して、ある実験における任意の全てのプローブを同定しうることを意味する。これは当然ながら、リード2の長さが十分であり、プローブは、リード__2により精査される領域が、独特の識別子を構成するようにデザインされることを仮定する。表31は、192種のプローブのコレクションおよび各プローブについての独特の識別子として用いられる10ヌクレオチドのリード__2配列について記載する。2種のプローブ（CYP2C19__r5__FおよびCYP2C9__r5__F）は当然ながら、同一な10ヌクレオチドの5' DNA配列を共有し、それらの間を識別する、「AG」または「CT」の2ヌクレオチドのコード（影を付した）を付加されていることに注目されたい。

40

【0466】

50

【表 3 1 - 1】

表31: 192のプローブのコレクション

名称	配列	配列 番号	10nt リード_2	配列 番号
CYP2D6_F	AAGCACCTAGCCCCATTCTGCTGAGC AGGAGGTGGCAGGTACCCAGACTGG GAGGTAA	201	AAGCACCTAG	393
CYP2D6_R	AGTCGGTGGGGCCAGGATGAGGCCCA GTCTGTTACACATGGCTGCTGCCTCT CAGCTCT	202	AGTCGGTGGG	394
AMY1_F	ACCTGAGTAGCATCATTGTAGTTCTCG ATATCTCCACTTCCAGTTTTACATTTA CCATCA	203	ACCTGAGTAG	395
chrX_15_F	CCTGGCCCTCAGCCAGTACAGAAAGT CATTTGTCAAGGCCTTCAGTTGGCAGA CGTGCTC	204	CCTGGCCCTC	396
chrX_15_R	AGAATTCATTGCCAGCTATAAATCTGT GGAAACGCTGCCACACAATCTTAGCA CACAAGA	205	AGAATTCATT	397
chrX_477_F	GACTTCAAAGAAATTACAAGTTGACA	206	GACTTCAAAG	398

10

20

【表 3 1 - 2】

名称	配列	配列 番号	10nt リード_2	配列 番号
	TCTTGGACTCTACCCCTCGTACTTTAT CTCCTAT			
chrX_477_R	TCTCTTTGGGGTCAAGAAAGAATCCCT AGTGGATTTGGGATTCTAGAGGAGGT GTTATAA	207	TCTCTTTGGG	399
chrX_478_F	TGCGATACCATGCTGAAGATGAGCTA ACCCAACCAGCCAAGCAGGCAGGGCT GCGAAGGA	208	TGCGATACCA	400
chrX_478_R	GGGGTAGGTGGAACCAAGTAATG TGATTTTGTAACATCCACTGCTGCATT TGTTTGC	209	GGGGTAGGTG	401
chrX_69_F	TTACTTCCCTCCAGTTTGTGCTTGCA AAACAACAGAATCTTCTCTCCATGAA ATCATG	210	TTACTTCCCT	402
chrX_69_R	CAGGGGTATCTATTATCCCCATTTTCT CACAAAGGAAACCAAGATAAAAGGTT TAAATGG	211	CAGGGGTATC	403
PLP1_ex1_F	GAAATTCTCTTGGAATTCCTGTGTCC TCTTGAATCTTCAATGCTAAAGTTTTT GAAACT	212	GAAATTCTCT	404
PLP1_ex2_F	GGGTTTGAGTGGCATGAGCTACCTACT GGATGTGCCTGACTGTTTCCCCTTCTT CTTCCC	213	GGGTTTGAGT	405
PLP1_ex2_R	CTATCTCCAGGATGGAGAGAGGGA AAAAAGATGGGTCTGTGTGGGAGGGC AGGTACTT	214	CTATCTCCAG	406
PLP1_ex3_F	GAAAGAAGCCAGGTCTTCAATTAATA AGATCCCTGGTCTCGTTTGTCTACCT GTTAATG	215	GAAAGAAGCC	407

10

20

30

40

【表 3 1 - 3】

名称	配列	配列 番号	10nt リード 2	配列 番号
PLP1_ex3_M	CAGACTCGCGCCCAATTTTCCCCCACC CCTTGTTATTGCCACAAAATCCTGAGG ATGATC	216	CAGACTCGCG	408
PLP1_ex3_R	TCTTTCTTCTTCCTTTATGGGGCCCTCC TGCTGGCTGAGGGCTTCTACACCACCG GCGCA	217	TCTTTCTTCT	409
PLP1_ex4_F	GTTTGTGTTTCTACATCTGCAGGCTGA TGCTGATTTCTAACCACCCCATGTCAA TCATTT	218	GTTTGTGTTT	410
PLP1_ex4_R	AACCAAATATATAGTGCTTCCATAGTG GGTAGGAGAGCCAAAGCACCCGTACC CTAACTC	219	AACCAAATAT	411
PLP1_ex5_F	AGTCTCCATGTGGCCCCGTAACCTCAT AAAGCTTACCCTGCTTGCTTTTGTGT CTTACT	220	AGTCTCCATG	412
PLP1_ex5_R	CCATGGGTGTAATTTGTATGGTATTAG CTACTCCCTTGTAATAAACCCTAATA ACCCAC	221	CCATGGGTGT	413
PLP1_ex6_F	TTTACAGTGGAGCATATTACTGCTGTT GCAAGAAACAGTTCTTCTCTTTCATT TTCCTG	222	TTTACAGTGG	414
PLP1_ex6_R	ATAGCTGTACCCACACTATCTCAGGCC TATTTACTTGCCAAGATCATTCAAAGT CAACTC	223	ATAGCTGTAC	415
PLP1_ex7_F	GATTTGAGGAGGGAGTGCTTTCTTTTC TACTCTCATTACATTCTCTTCTGTGTT CCCTA	224	GATTTGAGGA	416
PLP1_ex7_R	CAGCATTGTAGGCTGTGTGGTTAGAG CCTCGCTATTAGAGAAAGGGGGATTT	225	CAGCATTGTA	417

10

20

30

40

【表 3 1 - 4】

名称	配列	配列 番号	10nt リード 2	配列 番号
	CTACGGGG			
KRAS_ex1_F	TGTTACCTTTAAAAAGACATCTGCTTTC TGCCAAAATTAATGTGCTGAACTTAA ACTTACC	226	TGTTACCTTT	418
KRAS_ex1_R	TTCCCAGTAAATTACTCTTACCAATGC AACAGACTTTTAAAGAAGTTGTGTTTAA CAATGC	227	TTCCCAGTAA	419
KRAS_ex2_F	TAAATGACATAACAGTTATGATTTTGC AGAAAAACAGATCTGTATTTATTTCAGT GTTACT	228	TAAATGACAT	420
KRAS_ex2_R	GACAGGTTTTGAAAGATATTTGTGTTA CTAATGACTGTGCTATAACTTTTTTTT CTTTCC	229	GACAGGTTTT	421
KRAS_ex3_F	ACTCAAAAAATAAAAACTATAATTAC TCCTTAATGTCAGCTTATTATATTCAA TTTAAAC	230	ACTCAAAAAA	422
KRAS_ex3_R	AACACCTTTTTTGAAGTAAAAGGTGC ACTGTAATAATCCAGACTGTGTTTCTC CCTTCTC	231	AACACCTTTT	423
KRAS_ex4_F	GAAACCTTTATCTGTATCAAAGAATG GTCCTGCACCAGTAATATGCATATTAA AACAAGA	232	GAAACCTTTA	424
KRAS_ex4_R	GTGTATTAACCTTATGTGTGACATGTT CTAATATAGTCACATTTTCATTATTTT ATTAT	233	GTGTATTAAC	425
MYC_r1_F1	CCCCAGCCAGCGGTCCGCAACCCCTTG CCGCATCCACGAAACTTTGCCCATAGC AGCGGGC	234	CCCCAGCCAG	426
MYC_r1_R1	CGACTCATCTCAGCATTAAAGTGATA	235	CGACTCATCT	427

10

20

30

40

【表 3 1 - 5】

名称	配列	配列 番号	10nt リード 2	配列 番号
	AAAAAATAAATTAAAAAGGCAAGTGGA CTTCGGTG			
MYC_r2_F1	CTGTGGCGCGCACTGCGCGCTGCGCC AGGTTTCCGCACCAAGACCCCTTTAAC TCAAGAC	236	CTGTGGCGCG	428
MYC_r2_F2	TTCTACTGCGACGAGGAGGAGAACTT CTACCAGCAGCAGCAGCAGAGCGAGC TGCAGCCC	237	TTCTACTGCG	429
MYC_r2_F3	ACCGAGCTGCTGGGAGGAGACATGGT GAACCAGAGTTTCATCTGCGACCCGG ACGACGAG	238	ACCGAGCTGC	430
MYC_r2_F4	GCCGCCGCCTCAGAGTGCATCGACCC CTCGGTGGTCTTCCCTACCTCTCAA CGACAGC	239	GCCGCCGCCT	431
MYC_r2_R1	GGCGGCTAGGGGACAGGGGCGGGTG GGCAGCAGCTCGAATTTCTTCCAGATA TCCTCGC	240	GGCGGCTAGG	432
MYC_r2_R2	AGACGAGCTTGGCGGCGCCGAGAAG CCGCTCCACATACAGTCCTGGATGATG ATGTTTT	241	AGACGAGCTT	433
MYC_r2_R3	AGGAGAGCAGAGAATCCGAGGACGG AGAGAAGGCGCTGGAGTCTTGCGAGG CGCAGGACT	242	AGGAGAGCAG	434
MYC_r2_R4	TAAGAGTGGCCCGTTAAATAAGCTGC CAATGAAAATGGGAAAAGGTATCCAGC CGCCCACT	243	TAAGAGTGGC	435
MYC_r3_F1	TTGTATTTGTACAGCATTAATCTGGTA ATTGATTATTTAATGTAACCTTGCTA AAGGAG	244	TTGTATTTGT	436

10

20

30

40

【表 3 1 - 6】

名称	配列	配列 番号	10nt リード_2	配列 番号
MYC_r3_F2	GAGGCCACAGCAAACCTCCTCACAGC CCACTGGTCCTCAAGAGGTGCCACGT CTCCACAC	245	GAGGCCACAG	437
MYC_r3_F3	AGAGGAGGAACGAGCTAAAACGGAG CTTTTTTGCCTGCGTGACCAGATCCC GGAGTTGG	246	AGAGGAGGAA	438
MYC_r3_R1	TCCAACCTTGACCCTCTTGGCAGCAGGA TAGTCCTTCCGAGTGAGGGAGGCGC TGCGTAG	247	TCCAACCTGA	439
MYC_r3_R2	GCTTGGACGGACAGGATGTATGCTGT GGCTTTTTTAAGGATAACTACCTTGGG GGCCTTT	248	GCTTGGACGG	440
MYC_r3_R3	GCATTTGATCATGCATTTGAAACAAGT TCATAGGTGATTGCTCAGGACATTTCT GTTAGA	249	GCATTTGATC	441
AMY1A_r_9	CATCATTGTTGAAAACAATGAATCCTC TGTTTCCTCTCCCAAAGCCACTTGGT TGCTCC	250	CATCATTGTT	442
AMY1A_r_10	TTTATCTCCAGAAATGACATCACAGTA TGTGCCAGCAGGAAGACCAGTTTGCA AAGTTAA	251	TTTATCTCCA	443
AMY1A_r_8	CTATTAGAGGACATGTCTAAATACAT ATTCTCACCTTATTTGGCGCCATCGAT GTTACA	252	CTATTAGAGG	444
AR_r1_F	TCTGAGCAAGAGAAGGGGAGGCGGG GTAAGGGAAGTAGGTGGAAGATTCAG CCAAGCTCA	253	TCTGAGCAAG	445
AR_r1_R	CCAGAACACAGAGTGA CTCTGCCCTG GGCCGAAAGGCGACATTTCTGGAAGG	254	CCAGAACACA	446

10

20

30

40

【表 3 1 - 7】

名称	配列	配列 番号	10nt リード 2	配列 番号
	AAAAACTT			
AR_r2_F	CTTCACTTGCCTATTTCTGCCATTTCAG TGACATGTGTTGCATTGGTTTTTTGTG TCTTTC	255	CTTCACTTGC	447
AR_r2_R	GTGTCTCTCTCTGGAAGGTAAAGGAG AAAGGGAAAGAGAGAGTGCATGTGCAA GACCCTTT	256	GTGTCTCTCT	448
AR_r3_F	CCGAAGAAAGAGACTCTGGAAACTCA TTATCAGGTCTATCAACTCTTGTATTT GTTCTCC	257	CCGAAGAAAG	449
AR_r3_R	ACTAGAAAATGAGGGAGAAGGGGGA GAGAGGAAGGAGGAGGAAGAGAAAG AAAAGTATCT	258	ACTAGAAAAT	450
AR_r4_F	GTAGTTGCATTGTGTGTTTTTGACCAC TGATGATAAATTCAAGTCTCTTTCCT TCCCAA	259	GTAGTTGCAT	451
AR_r4_R	CATAGGAGCGTTCCTAAATATGATC CCCCTTATCTCATGCTCCCACTTCCCTT TTCCTT	260	CATAGGAGCG	452
AR_r5_F	CTCAGACTTAGCTCAACCCGTCAGTAC CCAGACTGACCACTGCCTCTGCCTCTT CTTCTC	261	CTCAGACTTA	453
AR_r5_R	CACCAACCAGGTCTGGCCAAGCTGCT GTATTTTAGTGAGGTCTGGGCCCCAGG AGCACTT	262	CACCAACCAG	454
ATR_r44_F	GGAAGATACAGTTGTTGAGAAAAGGA AATTGAGAGAAAACACAATTAGTAAG AGTAACTC	263	GGAAGATACA	455
ATR_r44_R	TTTtagatttATATTGGCCTCATATGT	264	TTTtagattt	456

10

20

30

40

【表 3 1 - 8】

名称	配列	配列 番号	10nt リード 2	配列 番号
	ATATGGATATTTTCATAGGCATTGTGTT TCTTTT			
ATR_r45_F	TGTAGGGGCCAATAATTATATTCGAG GTTACTGTAAATTATTTACAAAGTAT AGGTGAT	265	TGTAGGGGCC	457
ATR_r45_R	TTTGAGTAAAGATTTTAAATTCTAAC ATTGTTAGTTTGTAATAAAATGTATTG TTTCTA	266	TTTGAGTAAA	458
ATR_r46_F	CATATCAAGTTCATTTGTAGAGATGAG GACTACAGCCCATATCAAGCTATACCT TCTACT	267	CATATCAAGT	459
ATR_r46_R	ATGTCAGCAAGATTTCTTCTTGCAAAG ATAACATCATACCATAATATTTGTTTC AATTTT	268	ATGTCAGCAA	460
ATR_r47_F	CAACCACAGATTCATACCAAATGCAT TACTTTTAGATTATTAACATATTCTTTT ACATAA	269	CAACCACAGA	461
ATR_r47_R	AAGAAAGGTAGTAATTCCAAATTATT AACATCTGTTTTTGTTTTATGTTTCTT CTTTTT	270	AAGAAAGGTA	462
C4A_r20_F	GGCCTATGTGTGGCCACCCAGTCCA GCTCCGGGTGTTCCGCGAGTTCCACCT GCACCTC	271	GGCCTATGTG	463
C4A_r24_R	AGGCGTGGCCTCCCTCTTGAGGCTTCC TCGAGGCTGTGGGGAGCAAACCATGA TCTACTT	272	AGGCGTGGCC	464
C4A_r1_F	CCTAGCTTGGCCAGAAGGTAGCAGAC AGACAGACGGATCTAACCTCTCTTGG ATCCTCCA	273	CCTAGCTTGG	465

10

20

30

40

【表 3 1 - 9】

名称	配列	配列 番号	10nt リード 2	配列 番号
C4A_r1_R	TGAATCGGGTCCCCGATGCCAGCCCTG CCCCAATCCAAGCACCCAGCATCCCCG CCTCCAGG	274	TGAATCGGGT	466
C4A_r2_F	AGGGAGAGCAGGGGTGGAGGTGTCAG AGCGAAGTCTGACTGCTGATCCTGTCT GTTCTCC	275	AGGGAGAGCA	467
C4A_r2_R	TGACACTTACAAGACAGATGGGAACA GGGCAGGAGGCCCCCACAAGCAGCAG GAGGGCAT	276	TGACACTTAC	468
C4A_r3_F	CCTTTTGGTCAGCTGTCTCTTGCTCTG TGACCCGCTCCCTCTCCCTCTCCCTCT CCTGA	277	CCTTTTGGT	469
C4A_r3_R	CAAACTCATCTGAGAGGGCTCGGAG GGGGTTAAAGGTTGAGGCCCTGGGGC TGAGACTC	278	CAAACTCATC	470
C4A_r4_F	AGTTTGACCCACCCTCCCCTTGCACAT GGACCCCTGCTCACCTCTCTCCTCCTC CACTCC	279	AGTTTGACCC	471
C4A_r4_R	AAGGGGAGAAGTGCTCACAGGCAGGA GGTCACATCAGTGGCCAGGATCAGGA AGGCCAGA	280	AAGGGGAGAA	472
CCL4_r1_F	GTTCTGAAGCTTCTGAGTTCTGCAGCC TCACCTCTGAGAAAACCTCTTTGCCAC CAATAC	281	GTTCTGAAGC	473
CCL4_r2_F	GGCAGTGTTGATCTCACCTTGGCCTTT CCTTTCAGTGGGCTCAGACCCTCCCAC CGCCTG	282	GGCAGTGTTG	474
CCL4_r2_R	ACCACAGCTGGCTGGGAGCAGAGGCT GCTGGTCTCATAGTAATCTACCACAAA	283	ACCACAGCTG	475

10

20

30

40

【表 3 1 - 1 0】

名称	配列	配列 番号	10nt リード 2	配列 番号
	GTTGCGA			
CCL4_r3_F	TCAGGTGACCTTCCCTGAAGACTTCCT GTCTCTGAGCAGCTCAGTTCAGTTCCA GGTCAT	284	TCAGGTGACC	476
CYP2C19_r1_ F	GGAGCATATAGTGGGCCTAGGTGATT GGCCACTTTATCCATCAAAGAGGCAC ACACACTT	285	GGAGCATATA	477
CYP2C19_r1_ R	CCTTTCAAAGTATTTTACTTTACAATG ATCTCTTGTAACATTGTACCTCTAGGG ATATAC	286	CCTTTCAAAG	478
CYP2C19_r3_ F	ATGGGGAGGATGGAAAACAGACTAGC AGAGCTTCTCGGGCAGAGCTTGGCCC ATCCACAT	287	ATGGGGAGGA	479
CYP2C19_r3_ R	GATCTGGCCACCCCTGAAATGTTTCCA AGAATGTCAGTAGAGAGGAGAGCAGT CCAGAAA	288	GATCTGGCCA	480
CYP2C19_r5_ F	AGCAACCAGAGCTTGGCATATTGTAT CTATACCTTTTATTAAATGCTTTTAATTT AATAAATT	289	AGCAACCAGA	481
CYP2C19_r5_ R	CAAACTAGTCAATGAATCACAAATA CGCAAGCAGTCACATAACTAAGCTTTT GTTAACA	290	CAAACTAGT	482
CYP2C9_r1_F	AAGGAGCATATAGTGGACCTAGGTGA TTGGTCAATTTATCCATCAAAGAGGCA CACACCG	291	AAGGAGCATA	483
CYP2C9_r1_R	AGCCTTTCAAAGTATTTTACTTTACCA TTACCTCTTGTAACATGTACCTCTAGG GATACA	292	AGCCTTTCAA	484
CYP2C9_r3_F	GGGGAGGATGGAAAACAGAGACTTAC	293	GGGGAGGATG	485

10

20

30

40

【表 3 1 - 1 1】

名称	配列	配列 番号	10nt リード 2	配列 番号
	AGAGCTCCTCGGGCAGAGCTTGGCCC ATCCACAT			
CYP2C9_r3_R	GATATGGCCACCCCTGAAATGTTTCCA AGAATGTCAGTAGAGAAGATAGTAGT CCAGTAA	294	GATATGGCCA	486
CYP2C9_r5_F	CTCAACCAGAGCTTGGTATATGGTATG TATGCTTTTATTAAAATCTTTTAATTTA ATAAATT	295	CTCAACCAGA	487
CYP2C9_r5_R	CAGAACTAGTCAACAAATCACAAATT CACAAGCAGTCACATAACTAAGCTTTT GTTTACA	296	CAGAACTAGT	488
CYP2D6_r1_F	GGAAGTCCCCCAAACCTGCTTCCCCTT CTCAGCCTGGCTTCTGGTCCAGCCTGT GGTTTC	297	GGAAGTCCCC	489
CYP2D6_r1_R	GCCTGGTGGGGTGGGGGTGCCAGGTG TGTCCAGAGGAGCCCATTGGTAGTG AGGCAGGT	298	GCCTGGTGGG	490
CYP2D6_r2_F	GCCCAGCTCGGACTACGGTCATCACC CACCCGGGTCCACGGAATCTGTCTC TGTCCCC	299	GCCCAGCTCG	491
CYP2D6_r2_R	AATAGGACTAGGACCTGTAGTCTGGG GGGATCCTGGCTTGACAAGAGGCCCT GACCCCTCC	300	AATAGGACTA	492
CYP2D6_r5_F	AAGTTCATGGGCCCCCGCCTGTACCCT TCCTCCCTCGGCCCCCTGCACTGTTTCC CAGATG	301	AAGTTCATGG	493
CYP2D6_r5_R	GGGCTGACAGGTGCAGAATTGGAGGT CATTTGGGGGCTACCCCGTTCTGTCCC GAGTATG	302	GGGCTGACAG	494

10

20

30

40

【表 3 1 - 1 2】

名称	配列	配列 番号	10nt リード 2	配列 番号
CYP2D6_r8_F	CATTGCTTTATTGTACATTAGAGCCTC TGGCTAGGGAGCAGGCTGGGGACTAG GTACCCC	303	CATTGCTTTA	495
CYP2D6_r8_R	TGGAGTCTTGCAGGGGTATCACCCAG GAGCCAGGCTCACTGACGCCCTCCC CTCCCCAC	304	TGGAGTCTTG	496
DCC_r1_F	AGTGCATGTGTGTGAGTGCTGCCGCTG CCCGCGACCCCTGGCCCCGAAGGTGT TGGCTGA	305	AGTGCATGTG	497
DCC_r1_R	AATGAGAAGGGAAGTGGGGTACGGA AGGGGGTGCACGAGAAGAAAGGAA AGAGCCACTT	306	AATGAGAAGG	498
DCC_r2_F	GGAATCTAAGCCTGAGATTTATTTGAA TACATGAACATATTTCCCTGTGCTCTC TTGTTC	307	GGAATCTAAG	499
DCC_r2_R	AAGATGGCATTCTGCGCTTGGGA ATGGATGAAAGGAAGCAGCAACTTTC AAATGGGT	308	AAGATGGCAT	500
DCC_r2	GGTAAAGTCCCTCATCTGGCTTGTGGT GTCTGGAATGAAGTATGTTTTGTATCA GCAGAG	309	GGTAAAGTCC	501
DCC_r3_F	TTTATTGGCGATTATTGTGCTTTATTG GAAGACTTATTCTTCTTCTTTGTTTT CTCC	310	TTTATTGGCG	502
DCC_r3_R	GAAAAAAATTGTTTTCAACTTATTCC AGAGAATATCATTCTGAAGGCAACAA AGAGCAT	311	GAAAAAAATT	503
DCC_r4_F	ATATATCATTATCTTTGCAATGTTTT CATATATCATATGATACTGTGTTTTCC	312	ATATATCATT	504

10

20

30

40

【表 3 1 - 1 3】

名称	配列	配列 番号	10nt リード 2	配列 番号
	CCTCA			
DCC_r4_R	ATTAGAGAAATATAAATAGCAGCATA TACCATACAAAAGTACACTTTACAAA AAAGTCTT	313	ATTAGAGAAA	505
EP300_r18_F	ATACTCCATCTCCCGTAAAAATAGTGA GACTTGAGTAATGTTTGATGTCACCTG TCTTTC	314	ATACTCCATC	506
EP300_r18_R	CAGTCACCACTATATTATTCTAGGTAT CCCAGAAAAGTTAAAGTCAAATCTGA AACACAT	315	CAGTCACCAC	507
EP300_r19_F	TTCTTACTGTTCTAGCTTGTCTTAAG GCCTCTGTGCTTTTTAACAAATGGTTT CTTTTG	316	TTCTTACTGT	508
EP300_r19_R	TCCGCATGCACTCCCTGGACATGTGGA CACATGGACCATGGTCCACACCTGGC CAAGCTT	317	TCCGCATGCA	509
EP300_r20_F	CTTGGCTTGGGCTGTGTTGTGTGAACG GAACAGTTCACCCAGTATGGCCTTCT TGCCGA	318	CTTGGCTTGG	510
EP300_r20_R	GCTGTGCATAATCACTGGACAACAAA CTAATTAGCACTTTTCAAAATAATGCA GTACTT	319	GCTGTGCATA	511
EP300_r21_F	GAACAGCAGTCAGATTGCTCATCTCTA TCACTTTTTCTCATTGTGTCCCTTTTCT CTCCT	320	GAACAGCAGT	512
EP300_r21_R	AGAGAATGAAAGGGAAAAAGAAAAG CCAAAGCGTACTGACTATTCAAGGGG ATCGTACTT	321	AGAGAATGAA	513
EP300_r22_F	TATTGCAAGTTTTCAATTTGGTTAAGGT	322	TATTGCAAGT	514

10

20

30

40

【表 3 1 - 1 4】

名称	配列	配列 番号	10nt リード 2	配列 番号
	TTGGGGTTAATTTTGAATTGGCTCTG CTCTTC			
EP300_r22_R	AAAACCTTGTTAATATTCACGATAAAG AAAAATTCCAGAGAAAGTAACAACGT TAAGACTT	323	AAAACCTTGTT	515
PTEN_r1_F	GCAGAAGAAGCCCCGCCACCAGCAGC TTCTGCCATCTCTCTCCTCTTTTCTT CAGCCA	324	GCAGAAGAAG	516
PTEN_r1_R	CTACTCCCACGTTCTAAGAGAGTGAC AGAAAGGTAAAGAGGAGCAGCCGCA GAAATGGAT	325	CTACTCCCAC	517
PTEN_r2_F	AGTATTCTTTTAGTTTGATTGCTGCAT ATTCAGATATTTCTTTCCTTAATAA AGTACT	326	AGTATTCTTT	518
PTEN_r2_R	AATGAAAACACAACATGAATATAAAC ATCAATATTTGAAATAGAAAATCAAA GCATTCTT	327	AATGAAAACA	519
PTEN_r3_F	GTAATTTCAAATGTTAGCTCATTTTTG TTAATGGTGGCTTTTTGTTTGTTT TGTTT	328	GTAATTTCAA	520
PTEN_r3_R	CTTTCACTTAATAGTTGTTTTAGAAGA TATTTGCAAGCATACAAATAAGAAAA CATACTT	329	CTTTCACTTA	521
PTEN_r4_F	GTACTTTTTTTTCTTCCTAAGTGCAAA AGATAACTTTATATCACTTTTAACTT TTCTTT	330	GTACTTTTTT	522
PTEN_r4_R	ACAGTAAGATACAGTCTATCGGGTTT AAGTTATACAACATAGTACAGTACAT TCATACCT	331	ACAGTAAGAT	523

10

20

30

40

【表 3 1 - 1 5】

名称	配列	配列 番号	10nt リード 2	配列 番号
PTEN_r5_F	TACTTGTTAATTAAAAATTCAAGAGTT TTTTTTTCTTATTCTGAGGTTATCTTTT TACCA	332	TACTTGTTAA	524
PTEN_r5_R	CCAAAATCTGTTTTCCAATAAATTCTC AGATCCAGGAAGAGGAAAGGAAAAA CATCAAAA	333	CCAAAATCTG	525
RHD_r1_F	TCAAGTAGGTGTTGGAGAGAGGGGTG ATGCCTGGTGCTGGTGGAACCCCTGC ACAGAGAC	334	TCAAGTAGGT	526
RHD_r2_F	CCTGTCCTTTCGGGGTCCATTCCCTCT ATGACCCAGAAGTGATCCAGCCACCA TCCCAAT	335	CCTGTCCTTT	527
RHD_r5_F	AACCCCTCGAGGCTCAGACCTTTGGA GCAGGAGTGTGATTCTGGCCAACCAC CCTCTCTG	336	AACCCCTCGA	528
RHD_r5_R	CATAAATATGTGTGCTAGTCCTGTTAG ACCCAAGTGCTGCCAAGGGCAGCGC CCTGCTC	337	CATAAATATG	529
RHD_r6_F	TTGCAGCAAGATGGTGTCTCTCTCTA CCTTGCTTCCTTTACCCACACGCTATT TCTTTG	338	TTGCAGCAAG	530
RHD_r7_F	GAGATCAAGCCAAAATCAGTATGTGG GTTCATCTGCAATAAAAAATGTTTGTTT TGCTTTT	339	GAGATCAAGC	531
RHD_r7_R	GCAACAGTGAGAGGAAGTTGTCTTGT TTTTGAACAGGCCTTGTTTTCTTGGA TGCTTTT	340	GCAACAGTGA	532
RUNX1_r1_F	CTGCCATTTTCATTACAGGCAAAGCTGA GCAAAAGTAGATATTACAAGACCAGC	341	CTGCCATTTT	533

10

20

30

40

【表 3 1 - 1 6】

名称	配列	配列 番号	10nt リード 2	配列 番号
	ATGTACT			
RUNX1_r1_R	AAGGTAAAAGAAATCATTGAGTCCCC CGCCTTCAGAAGAGGGTGCATTTCA GGAGGAAG	342	AAGGTAAAAG	534
RUNX1_r3_F	GCGGATCTCCCCCGGCCTCGCCGGCCT CCGCCTGTCTCCACCACCCTCTCCG GGCCAG	343	GCGGATCTCC	535
RUNX1_r3_R	CTGGTAGGAGCTGTTTGCAGGGTCCTA ACTCAATCGGCTTGTGTGATGCGTAT CCCCGT	344	CTGGTAGGAG	536
RUNX1_r4_F	TTTTGAAATGTGGGTTTGTGCCATGA AACGTGTTTCAAGCATAGTTTTGACAG ATAACG	345	TTTTGAAATG	537
RUNX1_r4_R	TGCCCTAAAAGTGATGTATAACATCC CTGATGTCTGCATTTGTCCTTTGACTG GTGTTT	346	TGCCCTAAAA	538
RUNX1_r5_F	GTATACCAGCCTGGAGGGTGTACCAG CCTGGAGGGTGTACCAGCCCCAAGTG GATGCACT	347	GTATACCAGC	539
RUNX1_r5_R	TTCAACAGATATGTTTCAGGCCACCAA CCTCATTCTGTTTTGTTCTCTATCGTGT CCCCAC	348	TTCAACAGAT	540
SKI_r4_F	AGGATGTGTCTGGGTGGTGTCTGGGG ACAGAGGCACCTTCCCGACACCCGCC TGCCCCCTC	349	AGGATGTGTC	541
SKI_r4_R	CCTCAGCCAGTGCCACCCCCACAGCC CACAGGGAGGAGGCACAGAAAGCGA CTCACACGT	350	CCTCAGCCAG	542
SKI_r5_F	CTGGTGTGGAGCTGCCGGGCACTCC	351	CTGGTGTGGA	543

10

20

30

40

【表 3 1 - 1 7】

名称	配列	配列 番号	10nt リード 2	配列 番号
	ATGACTTTGTTTCTGTCTCTGCTTCCTC CTCAGT			
SKI_r5_R	CTGGTGCAGGCTGTGCTCACTGCCCCGT GCCCTGGACCTCCCAGCACCACTCGCC CCGCTC	352	CTGGTGCAGG	544
SKI_r6_F	GTCATGGTGAGGGGTGTGCTGGGACC GGCTGGGCAGTGACCCCAGCCGCCT CCGGCCCC	353	GTCATGGTGA	545
SKI_r6_R	CCGGGGCCCCACGGCGGCTCCACGCCC ACCGTGCTGCGTGCCTCAGTCTCCCCA CCCGCAT	354	CCGGGGCCCCA	546
SRY_r1_F	CTGTAAGTTATCGTAAAAAGGAGCAT CTAGGTAGGTCTTTGTAGCCAATGTTA CCCGATT	355	CTGTAAGTTA	547
SRY_r1_M1	AGTAAAATAAGTTTCGAACTCTGGCA CCTTTCAATTTTGTGCGCACTCTCCTTGT TTTGA	356	AGTAAAATAA	548
SRY_r1_M2	GAAGCAAAGTCAATTCTTCGGCAGC ATCTTCGCCTTCCGACGAGGTCGATAC TTATAAT	357	GAAGCAAAGT	549
SRY_r1_M3	AATGGCCATTCTTCAGGAGGCACAG AAATTACAGGCCATGCACAGAGAGAA ATACCCGA	358	AATGGCCATT	550
SRY_r1_M4	CTTGCGCCTCTGATCGCGAGACCACAC GATGAATGCGTTCATGGGTCGCTTCAC TCTATC	359	CTTGCGCCTC	551
SRY_r1_R	GAAAGCTGTAAGTCTAAGTATCAGTG TGAAACGGGAGAAAAACAGTAAAGGC AACGTCCAG	360	GAAAGCTGTA	552

10

20

30

40

【表 3 1 - 1 8】

名称	配列	配列 番号	10nt リード 2	配列 番号
TNFRSF14_r1 _F	AGTTCCTCTGCTGGAGTTCATCCTGCT AGCTGGGTTCCTGAGCTGCCGGTCTG AGCCTGA	361	AGTTCCTCTG	553
TNFRSF14_r1 _R	TCTGTGGGAGGCCCTGGGGTCAGAAC TGGGATCTGCGAGCAGACGGAGAGGA GGCTCGGG	362	TCTGTGGGAG	554
TNFRSF14_r2 _F	CAAGCCTGGCAGAGCCACAGGGCAG CCAGGGCATCTCCCAATGCCTGTCCTG ACCCCT	363	CAAGCCTGGC	555
TNFRSF14_r2 _R	TGTCTGGGGCAGAAGGGGGCAAGAGT GTCTGCCCTCGGCCACAGAGCTGGC CCGCCAAA	364	TGTCTGGGGC	556
TNFRSF14_r3 _F	TGATGGGTGGGCTCCCGAAGGGGCCT CCCGCAGACTTGCGAAGTTCCCACTCT CTGGGCG	365	TGATGGGTGG	557
TNFRSF14_r3 _R	CAGGGTGCGGGGCATCCAGGCTGCC CAAGCGGAGGCTGGGCCGCTGTGCT GGCCTCTT	366	CAGGGTGCGG	558
UGT1A1_rP_ P	CTCCACCTTCTTTATCTCTGAAAGTGA ACTCCCTGCTACCTTTGTGGACTGACA GCTTTT	367	CTCCACCTTC	559
UGT1A1_r1_ F	TCGATCCAAAGTAATACATCTGAAAG AATATACGCATGTAAAAGTCCCACTC CAATACAC	368	TCGATCCAAA	560
UGT1A1_r1_ M1	CTGCAGCAGAGGGGACATGAAATAGT TGTCCTAGCACCTGACGCCTCGTTGTA CATCAGA	369	CTGCAGCAGA	561
UGT1A1_r1_ M2	ACTCTTTCACATCCTCCCTTTGGAATG GCACAGGGTACGTCTTCAAGGTGTAA	370	ACTCTTTCAC	562

10

20

30

40

【表 3 1 - 1 9】

名称	配列	配列 番号	10nt リード_2	配列 番号
	AATGCTC			
UGT1A1_r1_ M3	GCGTGTGATCAAAACATACAAGAAAA TAAAAAAGGACTCTGCTATGCTTTTGT CTGGCTG	371	GCGTGTGATC	563
UGT1A1_r1_ M4	ACATCAAAGCTGCTTTCTGCCAGGGA GGCCATGAGCTCCTTGTTGTGCAGTAA GTGGGAA	372	ACATCAAAGC	564
UGT1A1_r1_ M5	CTCTGCCCCACTGTATTCTTCTTGCATG CACTGCCATGCAGCCTGGAATTTGAG GCTACCC	373	CTCTGCCAC	565
UGT1A1_r1_ M6	TCATGTGATCTGAATGAGAGGAGAGA GGCCTGGGCACGTAGGAGAATGGGTT GGGGCACT	374	TCATGTGATC	566
UGT1A1_r1_ M7	TGCGACGTGGTTTATTCCCCGTATGCA ACCCTTGCCCTCAGAATTCCTTCAGAGA GAGGTG	375	TGCGACGTGG	567
UGT1A1_r1_ R	CTTCACAAAGTCACTTCTAAACAGCCA GACAGATGCAGAGCTCAATAGGTCCT GGACAGT	376	CTTCACAAAG	568
UGT1A1_r2_ F	ACTGTATGTAGTCATCAAAGAATATG AGAAAAAATTAAGTAAAAATTTTCTT CTGGCTC	377	ACTGTATGTA	569
UGT1A1_r2_ R	ATTAATCTGGAAGCTGGAAGTCTGGG ATTAGCGCTCCTGTGAAAATAGATAT GAGGCCAT	378	ATTAATCTGG	570
UGT1A1_r3_ F	CAGATTTGTTTTCTAATCATATTATGT TCTTTCTTTACGTTCTGCTCTTTTGCC CCTCC	379	CAGATTTGTT	571
UGT1A1_r3_ F	GCTTAAGCCATTTTCTTAATTTGACCC	380	GCTTAAGCCA	572

10

20

30

40

【表 3 1 - 2 0】

名称	配列	配列 番号	10nt リード 2	配列 番号
R	TGGTTTGACCTATACATCCAATCCGCC CAACAT			
UGT1A1_r4_ F	TGTGTCCAGCTGTGAAACTCAGAGAT GTAAGTGTGACATCCTCCCTATTTTG CATCTCA	381	TGTGTCCAGC	573
UGT1A1_r4_ R	ATTTGAAACAATTTTATCATGAATGCC ATGACCAAAGTATTCTTCTGTATCTTC TTTCTT	382	ATTTGAAACA	574
UGT1A1_r5_ F	ATAAAGAGAGGATTGTTTCATACCACA GGTGTTCAGGCATAACGAAACTGTC TTTGTGTT	383	ATAAAGAGAG	575
UGT1A1_r5_ R	TTTTCAAGTTTGGAATGACTAGGGA ATGGTTCAAAATTTTACCTTATTTCCC ACCCACT	384	TTTTCAAGTT	576
VHL_r1_F	CGCCCCGCGTCCGACCCGCGGATCCC GCGGCGTCCGGCCCCGGGTGGTCTGGA TCGCGGAG	385	CGCCCCGCGT	577
VHL_r1_M1	TAGAGGGGCTTCAGACCGTGCTATCG TCCCTGCTGGGTGCGGCCTAAGCGCC GGGCCCCGT	386	TAGAGGGGCT	578
VHL_r1_M2	GGCGCCGAGGAGGAGATGGAGGCCG GGCGGCCGCGGCCCGTGCTGCGCTCG GTGAACTCG	387	GGCGCCGAGG	579
VHL_r1_R	CCATACGGGCAGCACGACGCGCGGAC TGCGATTGCAGAAGATGACCTGGGAG GGCTCGCG	388	CCATACGGGC	580
VHL_r2_F	GGTGTGGGCCACCGTGCCAGCCACC GGTGTGGCTCTTTAACAACCTTTGCTT GTCCCGA	389	GGTGTGGGCC	581

10

20

30

40

【表 3 1 - 2 1】

名称	配列	配列 番号	10nt リード 2	配列 番号
VHL_r2_R	AAGTGGTCTATCCTGTACTTACCACAA CAACCTTATCTTTTAAAAAGTAAAC GTCAGT	390	AAGTGGTCTA	582
VHL_r3_F	CTTGTTTCGTTCTTGTACTGAGACCCT AGTCTGCCACTGAGGATTGGTTTTTG CCCTTC	391	CTTGTTTCGTT	583
VHL_r3_R	ATCAAGACTCATCAGTACCATCAAAA GCTGAGATGAAACAGTGTAAGTTTCA ACAGAAAT	392	ATCAAGACTC	584

【0 4 6 7】

ペアエンドシーケンシング実験では、リード__1 およびリード__2 を、同じ DNA クローンから導出する。これは、リード__1 のゲノム配列（図 2 4 の部分図（3）および（4））が、特定のプローブ（図 2 4 の部分図（5））と会合したために存在することを含意する。まとめると、このデータにより、次世代配列のコレクション内に存在する各 DNA 配列は、それを標的化したプローブ配列と会合しうることが指し示される。特定のプローブと会合する全ての DNA 配列を回収することができる。

【0 4 6 8】

次世代リシーケンシング解析（標的化解析または他の形の解析）のための本パラダイムは、リードを再び参照ゲノムと整列させることである。プローブ会合を標的化する知見は、リードをプローブによりまず分取し、次いで、アラインメントベースの方法、デノボアセンブリー法、またはこれらの両方により解析する新規のワークフローをもたらす。実施例 1 4 において記載する通り、PARSAR（probe-associated-read-scaffold-assembly）は、変異体の発見における、より複雑で困難な問題のうちの 1 つであって、最も興味深い変異体は、参照配列から最も著明に逸脱する変異体であるが、これらはまさに、従来の配列ベースのアラインメントに対して最も屈折的な配列（図 2 5）であるという問題を解決する。プローブ会合に続いて、デノボの局所的アセンブリーを使用すると、このような変異体は、容易に同定される。

【0 4 6 9】

プローブベースのリードの群分けを、変異体を高度な初回通過の信頼性で同定する、分子デザインの他の側面と共に使用する。図 2 6 において示される通り、プローブは一般に、標的領域を一括するようにデザインする。リードの重複的側面により、潜在的な変異体部位を、双方向的な独立のリードによりクエリーすることが可能となる。加えて、この二重プローブのデザインは、隣接するプローブ結合性部位自体をシーケンシングすることも確保する。これは、プローブの捕捉性能が問題となりうる場合に重要な特色である。例として述べると、この分子デザインでは、一塩基変異体が捕捉プローブ配列のうちの 1 種の基底をなす変異体の対立遺伝子が同定され、後段のインフォマティクス解析において把握される。

【0 4 7 0】

分子注釈から後段の変異体解析への情報の流れのさらなる側面は、3 つの塩基配列による標識と、粘着末端の配列開始部位との組合せとして規定される配列「タグ」（図 2 4 の（1）+（3））を伴う。配列タグは、各シーケンシングクローンが独特のものであることを規定する。図 2 7 で例示される通り、同種クローンのコレクション内で生じる変異

10

20

30

40

50

体であって、同一な配列タグを共有する変異体は、偽陽性である可能性が高い。これに対し、異なるタグ（低頻度で生じる場合であってもなお）を伴う配列の間で共有される変異体は、真陽性の変異体である可能性が高い。配列をタグ付けし、タグを使用して、信頼性の予測を変異体の判定に割り当てるこのシステムは、後段における変異体の検証（費用がかさみ、時間がかかる可能性がある）の負担を実質的に軽減する見通しがある。分子注釈は、分子技術によるシーケンシングプラットフォームについて記載する、実施例 16 の文献においてより詳細に記載されている。

【0471】

まとめると、本明細書で想定される技術プラットフォームの顕著な特色のうちの 1 つは、全ての「アニーリングプローブ」イベントが、さらなる分子注釈もまた保有する DNA クローンにコピーされるということである。配列は、プローブおよび試料標識により、特異的投入試料の特異的標的領域に属するコレクションに分別される。次いで、アラインメントとデノボアセンブリーとの組合せを、変異体を検出するために使用することができる。最後に、候補変異体の出現の冗長性を使用して、変異体の判定における信頼性を割り当てることができる。変異体の解析に加えて、コピー数を決定するための方法もまた提供される。これら 2 つのエレメントは、緊密にカップリングされており、コピー数の決定は、信頼性の高いシーケンシングリードに依存するために特異的である。コピー数を配列情報から決定するための全体的な図式を、図 28 に示す。

【0472】

（実施例 16）

分子技術シーケンシングプラットフォーム
総括

本明細書で想定されるゲノムシーケンシングプラットフォームは、（1）複数の個体から得られるゲノム試料に、単一のシーケンシングランで取り組み；（2）一（および／または複数）塩基変異体（SNV）ならびに一（および／または複数）塩基挿入および欠失（SNID）を、高い信頼性で検出し；（3）クエリーされる全ての遺伝子環境における、大スケールおよび小スケールのコピー数変異（CNV）を検出し；（4）クエリーされる遺伝子環境において、マイクロスケールの転座、反転、および挿入／欠失イベントを検出し；（5） $\geq$ エキソームスケールの探索（ヒトゲノム配列全体のうちの $\geq 1 \sim 2\%$ ）から、 $\leq$ 単一遺伝子スケールの検証までスケーラブルな技術システムを開発し；（6）ゲノム変異検定における高特異度（低偽陰性率）および高感度（低偽陽性率）を達成し；（7）単純で、携帯可能で、その実行においてスケーラブルな、分子技術およびバイオインフォマティクス技術を創出し；（8）品質管理測定にたやすく適する分子法をもたらす方法を提供する。

【0473】

ゲノムシーケンシングリードの全体的な図式を、図 29 に示す。各要素についての記載は、以下の通りである。

【0474】

（1）「配列標識」とは、リード開始位置（3）と共に使用されて、各シーケンシングリードが独特のものであることを確立する、（連続*）ヌクレオチドのセット（すなわち、独特の 3 マーのセット）である。先駆的文献では、この標識とリード開始点との組合せを、「独特の配列タグ」と称した。配列標識は、遭遇する最初の塩基セットであり、一塩基合成（SBS: sequencing by synthesis）化学反応では、4 つの DNA 塩基全てが、各リード位置において同等に提示されなければならないため、シーケンシング標識に対する制約は、独特のものであることだけでなく、また、配列標識の集合セットで使用される塩基のコレクションは、4 つの塩基全てを、シーケンシングされる全ての位置に存在させなければならないことでもある。局所的 CNV を決定する独特の配列タグの使用については、この文献のバイオインフォマティクス節に記載されている。

【0475】

(2) 「試料標識」とは、マルチプレックス化された試料のセット内の特定の試料を独特に同定する、(連続*)ヌクレオチドコードのセットである。シーケンシング標識と同様に、試料標識のコレクションもまた、SBSによる配列塩基判定の要件を満たす、4つの塩基全てを含有する。試料コードは、ゲノムDNA断片に隣接して意図的に配置される。このデザインの駆動因子は、ライゲーションバイアスであることは、ライゲーション接合部の約2塩基上流および1~2塩基下流において、DNAライゲーション効率への塩基優先性が存在することを意味する。試料コードを、ライゲーション接合部に配置することにより、特異的試料内の全ての断片は、ライゲーションの影響/バイアスを受ける。

【0476】

* いかなる特定の理論にも束縛されることを望まずに述べると、配列標識および試料標識は、楕円形ヌクレオチド配列として創出しうることが想定される。

【0477】

(3) ゲノム断片内の「リード開始点」とは、「独特の配列リード」を規定する、2種の鍵となるエレメントのうちの1種である。上記の節(1)で論じた通り、各リードに独特の同定「タグ」は、シーケンシング標識およびリード開始点を含む。下記でより詳細に検討する通り、独特の[(1)+(3)]配列タグのコレクションは、大スケールのCNVを決定するのに不可欠である。本明細書では、「大スケールのCNV」を、何らかの隣接配列を加えた、少なくとも1種のプローブ結合性領域の全体を伴う任意のCNVとして規定する。大スケールのCNVは、染色体全体の獲得または喪失と同程度に大規模のCNVでありうる。

【0478】

[(1)+(2)]配列標識および試料標識は、ゲノムライブラリーの全体が創出される、ライブラリー構築工程の初期段階において、末端修復されたゲノム断片にライゲーションされたアダプター配列内に埋め込まれる。

【0479】

(4) シーケンシングリード：ゲノム断片から得られる配列情報は当然ながら、ゲノムアッセイの中心的な焦点である。各リードは、同じアッセイ内でもたらされる、複数の、重複するリードの文脈において検討される。

【0480】

(5) プローブレベル(「ゲノムインデックス処理」)：全体的なゲノムアッセイ戦略は、複数の配列標識を、各シーケンシングリードをゲノム解析のより大きな枠組みに位置づける、複合的「分子注釈」と組み合わせることである。この操作パラダイム内では、リード1は、各注釈されたクローンのエレメント(1~4)を明らかにする。リード2は、ハイブリダイゼーションベースの捕捉および後続の酵素プロセッシングにより各クローンを回収したプローブ配列を明らかにする。全てのリードは、それらを捕捉したプローブに従い、まずクラスター形成されるため、プローブ配列情報は、ゲノム戦略に中心的である。各リードは、解析の前に、ゲノムプローブに照らしてインデックス処理されるため、このプローブごとベースの情報のクラスター形成を、「ゲノムインデックス処理」と称する。

【0481】

プローブ標識の興味深い特色のうちの1つは、捕捉反応物中の全てのプローブ配列の配座が、十分に規定されている(どのプローブが捕捉反応にかけられるのかが既知である)ことである。これは、リード2が、60ヌクレオチドのプローブ配列の全体を、必ずしも対象とする必要がないことを含意する。そうではなく、リード2は、特異的反応物中の全てのプローブの明確な同定を可能とする程度に十分な長さであるだけでよい。非限定的な一例として述べると、実施例15において論じたプローブセットは、5'プローブ配列のうちの7ヌクレオチドのだけに基づいて差別化しうる(同一な7ヌクレオチドの5'末端を伴う2種のプローブを、ジヌクレオチドコードでタグ付けしたので、情報論的に差別化することができるであろう)、192のプローブからなる。

【0482】

(6) 捕捉標識：ライブラリーの組成は、プローブと標的配列との緊密な分子相互作用により決定する。各独特のプローブ配列の性能は、いくつかの(4~6つの)ランダム塩基の連なりほどに単純でありうる、捕捉標識を使用してモニタリングすることができる。シーケンシングにおいて検出される捕捉標識の多様性および統計学的分布は、プローブ性能の直接的な尺度である。一例として、特定のプローブと会合する配列が極めて少数の配列である場合を想像されたい。この配列の欠損をプローブ性能の不良に帰し、これにより、プローブ再デザインの反復サイクルを開始してみたいかもしれない。しかし、配列の提示不足はまた、アダプターに良好にライゲーションされない配列、および／または使用される特定のPCRレジメンで良好に増幅されない配列の帰結でもありうる。捕捉標識の使用は、これらの欠陥モードの差別化を可能とする。プローブの性能が不良であると、実際に生じる極めて少数の捕捉イベントは、多数回にわたり現れる極めて少数の捕捉標識として顕示されるであろう。これに対し、実際の捕捉反応の前段における理由(ライゲーション、PCR、末端修復など)で提示が不良である場合は、該して、独特に提示される捕捉標識の大規模な配座が結果としてもたらされるであろう。特定の実施形態では、数千種のプローブの自動式デザインに移行する場合、プローブの性能を情報論的に品質管理する能力が、ますます重要となるであろう。

10

【0483】

(実施例17)

プローブの選択および実装

【0484】

20

概要

プローブ配列の選択とそれらを使用する方法とは、必ずしも協奏的に開発されてはいない。本実施例は、I節でプローブの選択基準について記載し、II節でそれらを最も効果的とする実験室法について記載する。例えば、図30を参照されたい。

【0485】

I節：標的化プローブの選択

最も一般的に述べると、標的濃縮プローブは、60ヌクレオチドの長さである。プローブが一般に有向であることは、それらの位置の一方の側(一般に3'側)において、配列を捕捉することを意味する。60マーの標的化コアに加えて、さらなる機能性(例えば、PCRプライマー結合性部位、ビオチンによるプルアウトなどを可能とする相補的オリゴのための結合性部位)を付加するテイル配列も付加する。60ヌクレオチドの標的化配列は、以下の制約および基準：(1)プローブは、標的配列の開始点に照らして-100~+50ヌクレオチドに配置すること。図30では、標的配列の「開始点」とは、イントロン：エキソン接合部であること；(2)プローブは、例示される通り、プローブ対から得られる配列が、逆向きの配向性で重複しているように、冗長性を伴ってデザインすること；(3)プローブは、33%以上のGC含量(60マー当たり>20のGまたはC)および67%以下の(60マー当たり<40のGまたはC)を保有するように選択する(可能な場合)こと；(4)プローブは、可能な場合は常にリピートを回避するように選択すること。これは、いずれもがUCSCゲノムブラウザー上で閲覧しうる、REPEATMASKERおよび／または独特の整列可能性基準を一助として行うこと；および(5)万

30

40

【0486】

II節：実験室法

標的濃縮への投入物は、本明細書の他の箇所で記載した、プローブ、gDNAライブラリー、および緩衝液である。標的化濃縮における第1のステップは、二本鎖PCR断片としての形態で始まる、gDNAライブラリーの融解ステップである。これは、gDNAを、好ましくは、10μlの総容量中100ng/μlの濃度で、98℃で2分間にわたり変性させるのに続いて、速やかに氷に移すことにより達成する。gDNAライブラリーは

50

、10 mMのトリス、pH 8.0および0.1 mMのEDTAを含有する低濃度の塩緩衝液中に懸濁させる。第2のステップは、5 μ lの濃縮結合緩衝液（4 MのNaCl、40 mMのトリス、pH 8.0、0.4 mMのEDTA、および0.4%のTween 20）を添加することである。これらの条件は特殊であるが、全体の構想は、塩濃度を、2 Nのオスモル濃度まで増大させて、相補的なDNA鎖の、急速な反応速度による会合を達成しなければならないということである。プローブの最終濃度が、各プローブ250 pMずつとなるように、5マイクロリットルのプローブもまた添加する。gDNAライブラリー、緩衝液、およびプローブの混合物を、2分間にわたり98℃まで加熱し、4分間ごとに1℃ずつの減分で68℃まで下げて冷却した。第3のステップでは、プローブ：gDNA複合体（プローブには、ビオチンを会合させている）を、ストレプトアビジンでコーティングされた磁気ビーズに結合させた。第4のステップでは、厳密な洗浄を使用して、例えば、プローブとgDNAとの間のヌクレオチド配列の短いマッチのために生じうる、プローブと非標的配列との望ましくない会合を除去する。厳密性は、塩濃度が低く、ホルムアミド濃度が高い洗浄緩衝液、例えば、30%～35%（v/v）のホルムアミド、10 mMのトリス、pH 8.0、0.1 mMのEDTA、および0.5%のTween 20を含有する緩衝液を使用することにより達成する。複数回（例えば、4回）にわたるビーズの洗浄を使用して、標的配列の所望の純度を達成する。洗浄されたビーズは、プロセッシングされ、増幅され、シーケンシングされたプローブに結合した標的配列を保有する。まとめると、低塩濃度によるgDNAライブラリーの融解、高塩濃度によるプローブのアニリング、および高いホルムアミド濃度による洗浄を、プローブのデザインと協奏的に使用して、高レベルの標的配列濃縮を達成する。

【0487】

（実施例18）

例示的な配列

【0488】

総括

例示的なゲノムタグ、試料コード、およびライブラリー情報を、下記の表32～34に示す。

【0489】

【表32】

表32: 例示的なゲノムタグ

ゲノムタグ			
AAC	CAA	GAT	TAG
ACA	CCC	GCG	TCT
AGT	CGG	GGC	TGA 5
ATG	CTT	GTA	TTC

【0490】

10

20

30

40

【表 3 3】

表33: 例示的な試料コード

試料コード			
GAA	TCA	CGA	ATA
TAC	GCC	AGC	CTC
AAG	CCG	TGG	GTG
CAT	ACT	GGT	TTT

10

【 0 4 9 1】

【表 3 4】

表34: 例示的なライブラリー

ゲノムタグ	試料コード	試料ID	入力
AAC	ATA	NA18917	50 ゲノム
ACA	CCG	NA18917	100 ゲノム
AGT	GGT	NA18917	200 ゲノム
ATG	TAC	NA18917	400 ゲノム
CAA	AGC	NA12878	50 ゲノム
CCC	CAT	NA12878	100 ゲノム
CGG	GTG	NA12878	200 ゲノム
CTT	TCA	NA12878	400 ゲノム
GAT			
GCG			
GGC			
GTA			
TAG			
TCT			
TGA			
TTC			

20

30

40

【 0 4 9 2】

(実施例 1 9)

タグ付けされ、標的化された、ゲノムライブラリーの構築概要

本明細書では、タグ付けされ、標的化された、ゲノムシーケンシングライブラリーへの複数の方法が想定されている。本実施形態では、DNA修復を使用して、プローブ会合配列を、捕捉されたゲノム断片に接合させる。この手法は、配列準備された標的化ゲノムライブラリーを創出するのに良好に働いた。

【 0 4 9 3】

構想

50

ライブラリー構築の重要な原理は、配列準備されたクローンが、ゲノム断片および捕捉プローブの両方から得られるDNA配列を含むことである。この部分の「組換え」により、プローブと直接接触するゲノム断片が大幅に濃縮され、シーケンシングリードを、プローブ配列の一方の側にフォーカスすることが可能となる。このデザインでは、標的ゲノムライブラリー断片、捕捉プローブ、および共通なパートナーオリゴの間で形成される三幅対の複合体は、DNA複製フォークを想起させる構造を保有する。このようなフォークは、通常のDNA複製において生じるが、また、DNA修復工程においても生じる。後者の場合、5'置換鎖をトリミングして、新たに重合化した鎖の、隣接する3'配列への接合を可能とすることが必要であることが多い。この修復工程は、2種の酵素および3種の酵素活性を必要とする。E. coli DNAポリメラーゼまたはBst DNAポリメ

10

【0494】

特定の実施形態では、Bstポリメラーゼはまた、DNAポリメラーゼホロ酵素と関連することが多い、3'-5'ヌクレアーゼ活性も欠くために好ましい。例えば、図31を参照されたい。この特色は、標的ゲノムクローンの一本鎖3'DNA突出は、保護を必要としないことを示唆するために有用である。必要とされる他の酵素および活性は、NAD⁺要求性Taq DNAリガーゼなどの、ニック閉鎖型DNAリガーゼである。プロセシングの後、プロセシングされた断片を、PCRにより増幅して、シーケンシングの前の

20

【0495】

原理実証オリゴヌクレオチド

この実験のために、本発明者らがqPCRアッセイを施す、8種のゲノム領域に対応する、8種の標的領域を選択した。これらの8種の領域のためのフォワードプライマーおよびリバースプライマーを表35に示す。捕捉プローブは、本明細書の他の箇所で使用および検証された捕捉プローブの、正確なリバース相補鎖である。これらのプローブは、表37で言及される通り、22%~73%のGC%の範囲にわたる。

【0496】

【表 3 5】

表35: 原理実証オリゴヌクレオチド

表1: 原理実証オリゴヌクレオチド			
配列番号	qPCR アッセイ	名称	配列
585	17	SRY_F	CTGGTGTCTCCATTCTTGAGTGTGT
586		SRY_R	GGACAACAGGTTGTACAGGGATGA
587	18	VHL_F	TACAGGAGACTGGACATCGTCAGG
588		VHL_R	CTTTCTGCACATTTGGGTGGTCTT
589	19	UGT_F	GGTGATCAGATGGACAATGCAAAG
590		UGT_R	TCATTTCCAGAACATTTCAGGGTCA
591	20	TNF_F	ACCTCAATGGCCTAAGCAAGTGTC
592		TNF_R	GCCTCTTACCTGGGTCACACATTT
593	21	RUNX_F	ATCTTGCAACCTGGTTCTTCATGG
594		RUNX_R	GCTGGCAATGATGAAAACACTACTCG
595	22	RHD_F	CCAAGTTTCAACTCTGCTCTGCTG
596		RHD_R	GTTGAACACGGCATTCTTCCTTTC
597	23	PTEN_F	GGAAAGGGACGAACTGGTGTAATG
598		PTEN_R	AAATCTAGGGCCTCTTGTGCCTTT
599	24	EP300_F	AGCCCCATGGATCTTTCTACCATT
600		EP300_R	ATACTGCCAGGGCTCCTGATACTG
	捕捉プローブ	名称	配列
601	17	SRY_rlf_V3	AATCGGGTAACATTGGCTACAAAGACCTACCTAGATGCTCCTTTTTACGATAACTTACAGGTGAAAACCAGGATCAACTCCCGTGCCAGTCACATCTCGT
602		SRY_rlr_V3	ATTATAAGTATCGACCTCGTCGGAAGGCGAAGATGCTGCCGAAGAATTGCAGTTTGCTTCGTGAAAACCAGGATCAACTCCCGTGCCAGTCACATCTCGT

【0497】

プローブは、IDTによりUltramerとして合成されたものであり、100 uMまで再水和させ、プールしたが、プール内の各プローブは、6.25 uMで存在する。各プローブが100 nMで存在する、100倍濃度の原液を創出するために、10 uMのプールと100 uMの共通なビオチニル化パートナーオリゴ10 µlとを、605 µlのTE zero+0.05%のTween 20 (TT) 中で組み合わせた。100倍濃度の原液を、さらに100倍に希釈し(10 µlを、990 µlのTTに)て、各プローブが1 nMの濃度で存在する作業溶液をもたらした。

【0498】

捕捉／プロセッシングプロトコール

原理実証研究の1つの目的は、プローブの性能を検証し、配列準備されたライブラリー収率に対するプロセッシングの効率について調べることであった。ゲノムライブラリープールは、16例の試料セットライブラリーから導出した。プローブをアニーリングさせるた

め、個別のPCRストリップチューブ内に10 μ lずつのライブラリーアリコート4つを、2分間にわたり98 $^{\circ}$ Cまで加熱し、氷上で冷却した。4倍濃度の結合緩衝液5 μ lおよび5 μ lのプロブを、各チューブに添加し、98 $^{\circ}$ Cから69 $^{\circ}$ Cへの、4分間にわたり1 $^{\circ}$ Cずつのステップによるサーマルサイクラープログラムを使用して、溶液をアニーリングさせた。アニーリングされた複合体を、ストレプトアビジンでコーティングされた磁気ビーズに結合させ、25%のホルムアミド含有洗浄緩衝液で4回にわたり洗浄し、TE zeroで1回洗浄した。最終的な複合体を、2 μ lのTE zero中に懸濁させた。

【0499】

4種の複合体に対する4種の処理：1) 生の捕捉効率を決定する、プロセッシングを伴わない、ACA2プライマー単独による増幅；(2) プロセッシングを伴わない増幅および捕捉効率を決定する、プロセッシングを伴わない、AFおよびCRによる増幅；(3) 低容量のプロセッシングについて探索する、AFおよびCRによる増幅の前の、10 μ l中のPreCRによるプロセッシング；ならびに(4) 高容量のプロセッシング効果を確立する、50 μ lのAFおよびCRによる増幅の前の、PreCRによるプロセッシングについて探索した。

10

【0500】

PreCRによるプロセッシングは、100 μ l当たり

- － 82 μ lの水
- － 10 μ lのThermopol緩衝液
- － 1 μ lの100倍濃度のNAD⁺
- － 10mMのdNTP 1 μ l
- － 2 μ lのPreCR酵素ミックス

20

を含有する、製造元の推奨溶液を添加することにより達成した。

【0501】

10 μ lのPreCRカクテルを、チューブ3に添加し、50 μ lを、チューブ4に添加した。これらを、37 $^{\circ}$ Cで20分間にわたりインキュベートした。

【0502】

PreCR処理の後、TE zeroを添加することにより、4例の試料全てを、50 μ lに再懸濁させ、適切なPCRプライマーを伴うQ5 PCRカクテルを、250 μ lの最終容量に添加した。PCRカクテルの各アリコートは、

30

- － 125 μ lの水
- － 5倍濃度のQ5反応緩衝液50 μ l
- － 10 μ Mのプライマー25 μ l (ACA2またはAFとCRとの1:1ブレンド)
- － 10mMのdNTP 5 μ l
- － 2.5 μ lのQ5ホットスタート酵素

を含有した。

【0503】

50 μ lずつの各PCR反応ミックスを、1.25 μ lのEvaGreenおよび1 μ lのROX色素を含有するチューブにアリコート分割し、混合し、4連の10 μ lアリコートを、qPCR用光学PCRプレートに添加した。残りの200 μ lを、100 μ lのアリコートに分割した。qPCR反応および従来のPCR反応の両方を、

40

- － 98 $^{\circ}$ Cで30秒間
- － 98 $^{\circ}$ Cで10秒間、69 $^{\circ}$ Cで10秒間、72 $^{\circ}$ Cで10秒間の40サイクル (qPCR) およびプラトーサイクル (従来のPCR)

の通りにサイクリングさせた。

【0504】

リアルタイムPCR反応をモニタリングして、従来のPCR反応に最適の停止点を決定した。ACA2反応では、停止点は、21サイクル目であった。残りの反応では、停止点は、28サイクル目であった。これらのqPCR反応については、以下の結果節においてさらに記載する。

50

【0505】

10 μl の生のPCRを、ゲル解析のために回収し、残りの100 μl のアリコートをし、ビーズと1:1で精製した。精製PCR産物を、50 μl のTE zeroで溶出させ、Qubitにより定量化した。DNA収率は、(1) 7.44 ng/ μl ; (2) 10.6 ng/ μl ; (3) 12.1 ng/ μl ; および (4) 15.7 ng/ μl であった。

【0506】

捕捉／プロセッシングについてのqPCR解析

6例ずつの試料による8種のアッセイアッセイ17～24(表37)のアレイを含有する、単一のEco qPCRプレートを使用して、捕捉効率を評価した。6例の試料は、

1. 元のgDNAライブラリー10 ng/ μl
2. NTC
3. 0.01 ng/ μl の試料1
4. 0.01 ng/ μl の試料2
5. 0.01 ng/ μl の試料3
6. 0.01 ng/ μl の試料4

であった。

【0507】

Q5ホットスタートアッセイ混合物は、

- 237.5 μl のH₂O
- 5倍濃度のQ5反応緩衝液100 μl
- 10 μl のdNTP
- 12.5 μl のEvaGreen
- 10 μl のROX
- 5 μl のQ5ホットスタート酵素

を含有した。

【0508】

このカクテルを、48 μl のアリコートに分配し、3 μl のアッセイプライマー(FプライマーおよびRプライマーのいずれも10 uM)を添加した。これを列に分配した。2 μl の試料を行に添加し、プレートを上記で記載した通りにサイクリングさせた。

【0509】

結果

複合体の増幅：増幅用複合体の蛍光プロファイルを、増幅プラトー(単一のプライマーによるアンプリコンについては、二重プライマーによるアンプリコンはるかに早く生じる)を同定するのに主に使用する一方で、Cq値を使用して、アンプリコンの含量を、試料間で注視することができる。この実験では、観察されたCq値は、

【化4】

試料	条件	Cq
1	PreCRなし、ACA2単一プライマー	15
2	PreCRなし、AF + CR	21
3	10 μl PreCR, AF + CR	20
4	50 μl PreCR, AF + CR	19

であった。

【0510】

これらのデータは、PreCR処理により、P1 + P2(AF + CR)によるアンプリコンの存在量が増大することを実証した。

【0511】

図33に示される、プロセッシング後PCR産物のゲル画像は、PreCR処理により、サイズ分布の大きなクローンの増幅が支援されたことを示す。非処理試料2のアンプリコンは主に、サイズの小さな断片のクラスターである。試料3と、より大きな程度において、試料4とは、より広く分布するスメアである。

【0512】

標的濃縮を示すqPCRの結果を、下記の表36に示す。試料1における生の配列捕捉は、驚くべき程度に多かった。既往のデータセットに対する、このような予測外の改善には、少なくとも2つの因子：(1) コアのアニーリング工程（融解前の、高温で低塩濃度によるストレプトアビジンビーズへの結合）を最適化したこと；および(2) 長いパートナーオリゴ（35ヌクレオチドと対比した40ヌクレオチド）を使用したことが寄与した可能性がある。

10

【0513】

PreCR処理を伴わないP1+P2（AF+CR）による増幅用材料（試料2）も作製したところ、gDNA（および／またはNTC）に対する標的シグナルの実質的な濃縮が観察された。

【0514】

PreCRで処理された複合体はまた、プロセッシングされていない対照（試料1）と同等な濃縮レベルももたらした。これは、PreCRによるプロセッシングが、プローブベースのパートナーオリゴの、ゲノムライブラリーベースの標的クローンによる組換えを刺激するという事実の目覚ましい実証である。濃縮レベルが顕著ではない場合、クローン材料の大部分は少量であり、qPCRアッセイの範囲から外れた。本明細書の他の箇所で見られる通り、ビーズによる注意深い濃縮は、qPCR部位をカバーするライブラリーの比率を劇的に増大させる。

20

【0515】

加えて、結果により、PreCR処理が大量であるほど、必ずしも良好であるわけではないことも指し示された。濃縮比活性に関する、8種のアッセイのうちの6種において、試料3（10μlのPreCR処理）の成果は、試料4（50μlのPreCR処理）を上回った。

【0516】

【表 3 6 - 1】

表36: qPCR濃縮データ

表36: V3実験試料についてのqPCR濃縮データ								
Cqs	アッセイ 17	アッセイ 18	アッセイ 19	アッセイ 20	アッセイ 21	アッセイ 22	アッセイ 23	アッセイ 24
gDNA lib	24	21	21	22	22	22	22	22
NTC	35	31	34	26	33	N/A	30	29
試料 1	16	13	13	12	14	13	15	14
試料 2	21	18	19	16	18	18	22	20
試料 3	16	14	15	13	14	15	17	15
試料 4	16	14	15	13	15	15	17	15
Abs 値	アッセイ 17	アッセイ 18	アッセイ 19	アッセイ 20	アッセイ 21	アッセイ 22	アッセイ 23	アッセイ 24
gDNA lib	5	40	36	32	20	22	34	29
NTC	0	0	0	1	0		0	0
試料 1	1918	14374	9372	20243	8444	10189	2820	8133
試料 2	59	382	172	1526	295	403	34	87
試料 3	1085	5090	3051	10275	5213	4144	1068	2822
試料 4	1492	5381	2751	8770	3866	2777	777	2233
Adj 値	アッセイ 17	アッセイ 18	アッセイ 19	アッセイ 20	アッセイ 21	アッセイ 22	アッセイ 23	アッセイ 24
gDNA lib	5	40	36	32	20	22	34	29
NTC	0	0	0	1	0	0	0	0

【表 3 6 - 2】

試料 1	1918016	14374082	9372333	20243011	8443797	10188837	2820166	8132920
試料 2	59292	381694	171590	1526000	295038	403467	33575	87344
試料 3	1085442	5089739	3050781	10275412	5212989	4143563	1068184	2821950
試料 4	1491501	5380882	2751029	8770152	3865764	2777330	777035	2232604
濃縮倍数	アッセイ 17	アッセイ 18	アッセイ 19	アッセイ 20	アッセイ 21	アッセイ 22	アッセイ 23	アッセイ 24
試料 1	378673	356798	262412	641805	415341	453405	84090	281573
試料 2	11706	9475	4804	48382	14513	17954	1001	3024
試料 3	214298	126339	85418	325782	256421	184389	31850	97700
試料 4	294466	133566	77025	278058	190152	123592	23169	77296

【0 5 1 7】

考察

本実施例で開示される捕捉法およびプロセッシング法の性能は、非処理複合体を使用しても良好であった。いかなる特定の理論にも束縛されることを望まずに述べると、非処理複合体の成果が良好である 1 つの理由は、捕捉プローブおよびゲノム断片のいずれもが、プライマー結合性部位を保有するためであったことが想定される。

【0 5 1 8】

実施例 1 9 への付録

本実施例のためのプライマーおよびアンプリコンのデザインを、下記の表 3 7 に示す。

【0 5 1 9】

【表 3 7 - 1】

表37: 実施例19を実行するのに使用された配列

配列 番号	標的	名称	配列	%GC
603	SRY_r1, アッセイ 17	F qPCR プライマー	CTGGTGCTCCATTCTTGAGTGTGT	
604		R qPCR プライマー	GGACAACAGGTTGTACAGGGATGA	
605		qPCR アンプリコン	CTGGTGCTCCATTCTTGAGTGTGTGGCTTTCGTA CAGTCATCCCTGTACAACCTGTTGTCC	
606		F-プローブ	CTGTAAGTTATCGTAAAAAGGAGCATCTAGGTA GGTCTTTGTAGCCAATGTTACCCGATT	40%
607		R-プローブ	GAAGCAAACCTGCAATTCTTCGGCAGCATCTTCG CCTTCCGACGAGGTCGATACTTATAAT	47%

10

20

【表 3 7 - 2】

608		F- プローブ リバー 相補鎖	AATCGGGTAACATTGGCTACAAAGACCTACCTA GATGCTCCTTTTTACGATAACTTACAG	
609		R-プローブ リバー 相補鎖	ATTATAAGTATCGACCTCGTCGGAAGGCGAAGA TGCTGCCGAAGAATTGCAGTTTGCTTC	
610	VHL_r3 アッセイ 18	F qPCR プライマー	TACAGGAGACTGGACATCGTCAGG	
611		R qPCR プライマー	CTTCTGCACATTGCGGTGGTCTT	
612		qPCR アンプリコン	TACAGGAGACTGGACATCGTCAGGTCGCTCTAC GAAGATCTGGAAGACCACCCAAATGTGCAGAA AG	
613		F- プローブ	CTTGTTTCGTTTCCTTGTAAGTACGAGACCCTAGTCTGC CACTGAGGATTTGGTTTTTGGCCCTTC	48%
614		R- プローブ	ATCAAGACTCATCAGTACCATCAAAAGCTGAGA TGAAACAGTGTAAGTTTCAACAGAAAT	35%
615		F- プローブ リバー 相補鎖	GAAGGGCAAAAACCAAATCCTCAGTGGCAGAC TAGGGTCTCAGTACAAGGAACGAACAAG	
616		R- プローブ リバー 相補鎖	ATTTCTGTTGAAACTTACACTGTTTCATCTCAGC TTTTGATGGTACTGATGAGTCTTGAT	
617	UGT1A 1_r4, アッセイ 19	F qPCR プライマー	GGTGATCAGATGGACAATGCAAAG	
618		R qPCR プライマー	TCATTTCCAGAACATTTCAGGGTCA	

10

20

30

40

【表 3 7 - 3】

619		qPCR アンプリコン	GGTGATCAGATGGACAATGCAAAGCGCATGGA GACTAAGGGAGCTGGAGTGACCCTGAATGTTCT GGAAATGA	
620		F-プローブ	TGTGTCCAGCTGTGAAACTCAGAGATGTAACTG CTGACATCCTCCCTATTTTGCATCTCA	45%
621		R-プローブ	ATTTGAAACAATTTTATCATGAATGCCATGACC AAAGTATTCTTCTGTATCTTCTTTCTT	28%
622		F-プローブ リバー ス相補鎖	TGAGATGCAAAATAGGGAGGATGTCAGCAGTTA CATCTCTGAGTTTCACAGCTGGACACA	
623		R-プローブ リバー ス相補鎖	AAGAAAGAAGATACAGAAGAATACTTTGGTCAT GGCATTTCATGATAAAATTGTTTCAAAT	
624	TNFRS F14_r3, アッセイ 20	F qPCR プライマー	ACCTCAATGGCCTAAGCAAGTGTC	
625		R qPCR プライマー	GCCTCTTACCTGGGTCACACATTT	
626		qPCR アンプリコン	ACCTCAATGGCCTAAGCAAGTGTCTGCAGTGCC AAATGTGTGACCCAGGTAAGAGGC	
627		F-プローブ	TGATGGGTGGGCTCCCGAAGGGGCCTCCCGCAG ACTTGCGAAGTTCCCACTCTCTGGGCG	67%
628		R-プローブ	CAGGGTGCGGGGGCATCCAGGCTGCCCAAGCG GAGGCTGGGCCGGCTGTGCTGGCCTCTT	73%
629		F-プローブ リバー ス相補鎖	CGCCCAGAGAGTGGGAACTTCGCAAGTCTGCGG GAGGCCCTTCGGGAGCCCAACCATCA	
630		R-プローブ リバー ス相補鎖	AAGAGGCCAGCACAGCCGGCCAGCCTCCGCTT GGGCAGCCTGGATGCCCCCGCACCCCTG	

10

20

30

40

【表 3 7 - 4】

631	RUNX_ r4, アッセイ 21	F qPCR プライマー	ATCTTGCAACCTGGTTCTTCATGG	
632		R qPCR プライマー	GCTGGCAATGATGAAAACACTCG	
633		qPCR アンプリコン	AATCTTGCAACCTGGTTCTTCATGGCTGCGGTAG CATTTCTCAGCTCAGCCGAGTAGTTTTCATCATT GCCAGC	
634		F-プローブ	TTTGTAAATGTGGGTTTGTGTCATGAAACGTGT TTCAAGCATAGTTTGTGACAGATAACG	37%
635		R-プローブ	TGCCCTAAAAGTGTATGTATAACATCCCTGATG TCTGCATTTGTCCTTTGACTGGTGTTT	40%
636		F-プローブ リバー 相補鎖	CGTTATCTGTCAAACTATGCTTGAAACACGTTT CATGGCAACAAACCCACATTTCAAAA	
637		R-プローブ リバー 相補鎖	AAACACCAGTCAAAGGACAAATGCAGACATCA GGGATGTTATACATACACTTTTAGGGCA	
638	RHD_r5 アッセイ 22	F qPCR プライマー	CCAAGTTTCAACTCTGCTCTGCTG	
639		R qPCR プライマー	GTTGAACACGGCATTCTTCCTTTC	
640		qPCR アンプリコン	CCAAGTTTCAACTCTGCTCTGCTGAGAAGTCCA ATCGAAAGGAAGAATGCCGTGTTCAAC	
641		F-プローブ	AACCCCTCGAGGCTCAGACCTTTGGAGCAGGAG TGTGATTCTGGCCAACCACCCTCTCTG	58%
642		R-プローブ	CATAAATATGTGTGCTAGTCCTGTTAGACCCAA GTGCTGCCCAAGGGCAGCGCCCTGCTC	55%

10

20

30

40

【表 3 7 - 5】

643		F-プローブ リバー 相補鎖	CAGAGAGGGTGGTTGGCCAGAATCACACTCCTG CTCCAAAGGTCTGAGCCTCGAGGGGTT	
644		R-プローブ リバー 相補鎖	GAGCAGGGCGCTGCCCTTGGGCAGCACTTGGGT CTAACAGGACTAGCACACATATTTATG	
645	PTEN_r 5, アッセイ 23	F qPCR プライマー	GGAAAGGGACGAACTGGTGTAATG	
646		R qPCR プライマー	AAATCTAGGGCCTCTTGTGCCTTT	
647		qPCR アンプリコン	GGAAAGGGACGAACTGGTGTAATGATATGTGCA TATTTATTACATCGGGGCAAATTTTAAAGGCA CAAGAGGCCCTAGATTT	
648		F-プローブ	TACTTGTTAATTAAAAATTCAAGAGTTTTTTTTT CTTATTCTGAGGTTATCTTTTTACCA	22%
649		R-プローブ	CCAAAATCTGTTTTCCAATAAATTCTCAGATCCA GGAAGAGGAAAGGAAAAACATCAAAA	33%
650		F-プローブ リバー 相補鎖	TGGTAAAAAGATAACCTCAGAATAAGAAAAAA AAACTCTTGAATTTTAAATTAACAAGTA	
651		R-プローブ リバー 相補鎖	TTTTGATGTTTTTCCTTTCCTCTTCCTGGATCTGA GAATTTATTGGAAAACAGATTTTGG	
652	EP300_r 18, アッセイ 24	F qPCR プライマー	AGCCCCATGGATCTTCTACCATT	
653		R qPCR プライマー	ATACTGCCAGGGCTCCTGATACTG	

10

20

30

40

【表 3 7 - 6】

654		qPCR アンプリコン	AGCCCCATGGATCTTTCTACCATTAAGAGGAAG TTAGACACTGGACAGTATCAGGAGCCCTGGCAG TAT	
655		F-プローブ	ATACTCCATCTCCCGTAAAAATAGTGAGACTTG AGTAATGTTTGATGTCACTTGTCTTTC	37%
656		R-プローブ	CAGTCACCACTATATTATTCTAGGTATCCCAGA AAAGTTAAAGTCAAATCTGAAACACAT	33%
657		F-プローブ リバース 相補鎖	GAAAGACAAGTGACATCAAACATTACTCAAGTC TCACTATTTTTACGGGAGATGGAGTAT	
658		R-プローブ リバース 相補鎖	ATGTGTTTCAGATTTGACTTTAACTTTTCTGGGA TACCTAGAATAATATAGTGGTGACTG	
659	追加の オリゴ	捕捉プローブ のための テイル配列	GTGAAAACCAGGATCAACTCCCGTGCCAGTCAC ATCTCGT	
660		共通のパー トナーオリゴ 配列	/5BioTEG/ACGAGATGTGACTGGCACGGGAGTTG ATCCTGGTTTTTCAC	
661		フォワード 増幅プライ マーAF	AATGATACGGCGACCACCGAGATCTACACGTCA TGCAGGACCAGAGAATTCGAATACA	
662		リバース 増幅プライ マーCR	CAAGCAGAAGACGGCATAACGAGATGTGACTGG CACGGGAGTTGATCCTGGTTTTTCAC	

【 0 5 2 0 】

(実施例 2 0)

ライブラリーフリーの標的化ゲノム解析
概要

本実施例は、ライブラリーフリーゲノム解析について実証する。目標は、このような方法を、信頼でき、再現可能で、廉価で、ハイスループットのフォーマットで実装するのに最も有用なパラメータを同定することであった。特に、T4ポリメラーゼは、PEG 8000 (分子密集剤) の存在下で T4 遺伝子 32 タンパク質を補充すると仮定した場合において、多くの多様なゲノム配列をコピーしうることが発見された。加えて、配列ライブラリー構築のすぐ前段における抑制 PCR が、長いインサートのシーケンシングクローンを濃縮する強力な方法であることも分かった。

【0521】

背景

ライブラリーフリー法の背景にある分子概念は、

(1) gDNAを、約400bpに断片化すること、またはddNTPの存在下で、ランダムな15マーによる第1鎖のcDNA合成を実行すること(図33)；

(2) gDNAまたはcDNAを、標識された捕捉プローブと共に融解させ、末端修復されたgDNA/cDNAを精製すること。gDNAについては、ゲノム配列を、テイル部分内に含有されるランダムヘキサマー配列を含む配列タグで修復すること(図33)；

(3) 20℃の単一の反応においてDNA複合体をプロセッシングすること。使用される緩衝液は、NEB CutSmart (NEB#4およびBSA)、ATP、dNTP、およびPEG8000である。複合体は、T4 DNAポリメラーゼ、T4 遺伝子32タンパク質(SB)、およびT4 DNAリガーゼでプロセッシングする。アダプターライゲーション鎖は、5'リン酸化されており、パートナー鎖は、3'ddCを含む。アダプターの逆側の末端は、粘着末端であり、保護することができる。平滑末端構成は、自己二量体をなさず、極めて効果的であり、ライゲーション鎖を含有するP1を、標的を含有するP2に接合させること(図34)；

(4) フローセルに適合的な配列を添加し、試料特異的インデックス配列を各反応物に導入するPCR増幅(図35)；ならびに

(5) DNAシーケンシング(図35)を含む。

【0522】

特定の実施形態において生じうる、1種の潜在的なアーチファクトは、占有されていないプローブの存在量と関連する。T4 DNAポリメラーゼの3'-5'エキソヌクレアーゼ活性により、これらの分子において、平滑末端を作製することが可能であり、次いで、これを、P1アダプター配列へのライゲーションの基質とする(図36)。これらの短い「オリゴ二量体」産物は、介入しなければ、後続のPCR反応物を圧倒するであろう。潜在的なアーチファクトを回避するために、25ヌクレオチドのP2セグメントをP1アダプター内に組み入れる抑制PCRデザインを使用した。このセグメントを伴う抑制PCR増幅の後、P1またはP2に、特異的な突出を伴うフォワードプライマーおよびリバースプライマーを使用して、インデックス配列およびフローセル適合的突出を付加する。

【0523】

プロセッシング後抑制PCR、全長増幅、およびシーケンシングを可能とするオリゴヌクレオチドを、下記の表38に示す。

【0524】

【表 3 8】

表38

配列番号	名称	配列	説明
663	LF_part 鎖	AGTTGATCCTGGTTATACA/3ddC/	アダプターパートナー鎖
664	LF_lig 鎖	/5Phos/GTGTATAACCAGGATCAACTCCC GTGCCAGT	アダプターライゲーション鎖
665	LF_I1P	GTGAAAACCAGGATCAACTCCCGTGCCA GTCAC	インデックス1シーク エンシングプライマー
666	LF_FSP	GTCATGCAGGAGTTGATCCTGGTTATAC AC	ライブラリーフリーフォワード シークエンシングプライマー
667	LF_単一 プライマー 配列	ACTGGCACGGGAGTTGATCCTGGTT	プロセッシング後増幅 プライマー
668	LF_FLFP	AATGATACGGCGACCACCGAGATCTAC ACGTCATGCAGGAGTTGATCCTGGTTAT ACAC	ライブラリーフリーフォ ワード増幅プライマー
669	LF_FLRP_ N701	CAAGCAGAAGACGGCATAACGAGATTCG CCTTAGTGACTGGCACGGGAGTTGATCC TGGTTTTCAC	インデックスN701 リバースプライマー
670	LF_FLRP_ N702	CAAGCAGAAGACGGCATAACGAGATCTA GTACGGTGACTGGCACGGGAGTTGATCC TGGTTTTCAC	インデックスN702 リバースプライマー
671	LF_FLRP_ N703	CAAGCAGAAGACGGCATAACGAGATTTC TGCTGTGACTGGCACGGGAGTTGATCC TGGTTTTCAC	インデックスN703 リバースプライマー
672	LF_FLRP_ N704	CAAGCAGAAGACGGCATAACGAGATGCT CAGGAGTGACTGGCACGGGAGTTGATC CTGGTTTTCAC	インデックスN704 リバースプライマー

【 0 5 2 5】

材料

ゲノムDNA試料を、4例の対象から回収し、Oragene唾液回収キットを使用して精製した。この研究でシークエンシングした試料は、以下の通りであった。

【化 5】

この研究で使ったゲノムDNA試料およびIlluminaインデックス		
Coriel ID #	試料の説明	IlluminaインデックスID
NA19240	ヨルバ人女性	インデックスN701 リバースプライマー
GM03623	異数体XXXおよび18番染色体トリソミー	インデックスN702 リバースプライマー
GM11226	XXXX	インデックスN703 リバースプライマー
GM18917	ヨルバ人/ナイジェリア人:欠失変異体1、2、4、9を保有する	インデックスN704 リバースプライマー

【 0 5 2 6】

これらの実験で使ったプローブを、下記の表39に提示する。捕捉後増幅において発生する同種クローンから得られる、同じシークエンシング開始部位を伴う、独立の捕捉イベントを確立するのに、ヘキサマータグが必要とされる。

【表 3 9 - 1】

表39

配列番号	名称	配列
673	CYP2D6_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NAAGCACCTAGCCCCATTCTGCTGAGCAGGAGGTGGCA GGTACCCCAGACTGGGAGGTAA
674	CYP2D6_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NAGTCGGTGGGGCCAGGATGAGGCCAGTCTGTTCACAC ATGGCTGCTGCCTCTCAGCTCT
675	AMY1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NACCTGAGTAGCATCATTGTAGTTCTCGATATCTCCACTT CCAGTTTTACATTTACCATCA
676	chrX_15_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NCCTGGCCCTCAGCCAGTACAGAAAGTCATTTGTCAAGG CCTTCAGTTGGCAGACGTGCTC
677	chrX_15_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NAGAATTCATTGCCAGCTATAAATCTGTGGAAACGCTGCC ACACAATCTTAGCACACAAGA
678	chrX_477_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGACTTCAAAGAAATTACAAGTTGACATCTTGGACTCTAC CCCTCGTACTTTATCTCCTAT
679	chrX_477_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NTCTCTTTGGGGTCAAGAAAGAATCCCTAGTGGATTGGG ATTCTAGAGGAGGTGTTATAA
680	chrX_478_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NTGCGATACCATGCTGAAGATGAGCTAACCCAACCAGCC AAGCAGGCAGGGCTGCGAAGGA

10

20

【表 3 9 - 2】

681	chrX_478_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGGGGTAGGTGGAACCCAAAGTAATGTGATTTTGTAAC ATCCACTGCTGCATTTGTTTGC
682	chrX_69_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NTTACTTCCCTCCAGTTTGTGCTTGCAAAACAACAGAA TCTTCTCTCCATGAAATCATG
683	chrX_69_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NCAGGGGTATCTATTATCCCCATTTTCTCACAAAGGAAAC CAAGATAAAAGGTTTAAATGG
684	PLP1_ex1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGAAATTCTCTTGTAATTCCTGTGTCCTCTTGAATCTTCA ATGCTAAAGTTTTTGAACT
685	PLP1_ex2_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGGGTTTGAGTGGCATGAGCTACCTACTGGATGTGCCTGA CTGTTTCCCCTTCTTCTTCCC
686	PLP1_ex2_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NCTATCTCCAGGATGGAGAGAGGGAAAAAAAAGATGGGT CTGTGTGGGAGGGCAGGTACTT
687	PLP1_ex3_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGAAAGAAGCCAGGTCTTCAATTAATAAGATTCCCTGGT CTCGTTTGTCTACCTGTTAATG
688	PLP1_ex3_M	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NCAGACTCGCGCCCAATTTTCCCCACCCCTGTTATTGC CACAAAATCCTGAGGATGATC
689	PLP1_ex3_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NTCTTTCTTCTTCTTTATGGGGCCCTCCTGCTGGCTGAGG GCTTCTACACCACCGGCGCA
690	PLP1_ex4_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGTTTGTGTTTCTACATCTGCAGGCTGATGCTGATTTCTA ACCACCCCATGTCAATCATTT
691	PLP1_ex4_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NAACCAAATATATAGTGCTTCCATAGTGGGTAGGAGAGC CAAAGCACCCGTACCCTAACTC
692	PLP1_ex5_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NAGTCTCCATGTGGCCCCGTAAGTCCATAAAGCTTACCCT GCTTGCTTTTTGTGTCTTACT
693	PLP1_ex5_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NCCATGGGTGTAATTTGTATGGTATTAGCTACTCCCTTGT AAAATAACCCAAATAACCCAC
694	PLP1_ex6_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NTTACAGTGGAGCATATTACTGCTGTTGCAAGAAACAGT TCTTCTCTTTTCAATTTTCTG
695	PLP1_ex6_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NATAGCTGTACCCACACTATCTCAGGCCTATTTACTTGCC AAGATCATTCAAAGTCAACTC

10

20

30

40

【表 3 9 - 3】

696	PLP1_ex7_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NGATTTGAGGAGGGAGTGCTTTCTTTCTACTCTCATTCA CATTCTCTCTTCTGTTCCCTA
697	PLP1_ex7_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NCAGCATTGTAGGCTGTGTGGTTAGAGCCTCGCTATTAGA GAAAGGGGGATTCTACGGGG
698	KRAS_ex1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NTGTTACCTTTAAAAGACATCTGCTTTCTGCCAAAATTAA TGTGCTGAACCTAAACTTACC
699	KRAS_ex1_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NTTCCCAGTAAATTACTCTTACCAATGCAACAGACTTTAA AGAAGTTGTGTTTACAATGC
700	KRAS_ex2_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NTAAATGACATAACAGTTATGATTTTGCAGAAAACAGAT CTGTATTTATTTCAAGTGTACT
701	KRAS_ex2_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NGACAGGTTTTGAAAGATATTTGTGTTACTAATGACTGTG CTATAACTTTTTTTCTTTCC
702	KRAS_ex3_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NACTCAAAAAATAAAAACTATAATTACTCCTTAATGTCAG CTTATTATATTCAATTTAAAC
703	KRAS_ex3_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NAACACCTTTTTTTGAAGTAAAAGGTGCACTGTAATAATCC AGACTGTGTTTCTCCCTTCTC
704	KRAS_ex4_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NGAAACCTTTATCTGTATCAAAGAATGGTCCTGCACCAGT AATATGCATATTAAACAAGA
705	KRAS_ex4_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NGTGTATTAACCTTATGTGTGACATGTTCTAATATAGTCA CATTTTCATTATTTTATTAT
706	MYC_r1_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NCCCCAGCCAGCGGTCCGCAACCCTTGCCGCATCCACGA AACTTTGCCCATAGCAGCGGGC
707	MYC_r1_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN Ncgactcatctcagcattaaagtataaaaaataaataaaaggcaagtggactcgggtg
708	MYC_r2_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN Nctgtggcgcgcactgcgcgtgccaggttccgcaccaagacccttaactcaagac
709	MYC_r2_F2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NTTCTACTGCGACGAGGAGGAGAACTTCTACCAGCAGCA GCAGCAGAGCGAGCTGCAGCCC
710	MYC_r2_F3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNN NACCGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTT CATCTGCGACCCGGACGACGAG

10

20

30

40

【表 3 9 - 4】

711	MYC_r2_F4	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGCCGCCGCTCAGAGTGCATCGACCCCTCGGTGGTCTTC CCCTACCCTCTCAACGACAGC
712	MYC_r2_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGGCGGCTAGGGGACAGGGGCGGGGTGGGCAGCAGCTC GAATTTCTTCCAGATATCCTCGC
713	MYC_r2_R2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NAGACGAGCTTGCGGCGGCCGAGAAGCCGCTCCACATA CAGTCCTGGATGATGATGTTTT
714	MYC_r2_R3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NAGGAGAGCAGAGAATCCGAGGACGGAGAGAAGGCGCT GGAGTCTTGCGAGGCGCAGGACT
715	MYC_r2_R4	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN Ntaagagtggcccgttaataagctgccaatgaaaatgggaaaggtatccagccgccact
716	MYC_r3_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN Nttgtattgtacagcattaatctggttaattgattattttaatgtaaccttgctaaaggag
717	MYC_r3_F2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTC AAGAGGTGCCACGTCTCCACAC
718	MYC_r3_F3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTG CGTGACCAGATCCCGGAGTTGG
719	MYC_r3_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NTCCAACCTTGACCCTCTTGGCAGCAGGATAGTCCTTCCGA GTGGAGGGAGGCGCTGCGTAG
720	MYC_r3_R2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGCTTGGACGGACAGGATGTATGCTGTGGCTTTTTTAAGG ATAACTACCTTGGGGGCGCTTT
721	MYC_r3_R3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNN NGCATTTGATCATGCATTTGAAACAAGTTCATAGGTGATT GCTCAGGACATTTCTGTTAGA

【0527】

方法、結果および考察

【0528】

I 部：4 種の gDNA (F、S、C、および L) を、150 μ l の最終容量中に 20 ng/ μ l に希釈した。試料を、500 bp に超音波処理し、125 μ l を、125 μ l のビーズで精製した。出発材料および精製され、断片化された gDNA を、図 37 に示す。gDNA の濃度は、(1) F：137 ng/ μ l；(2) S：129 ng/ μ l；(3) C：153 ng/ μ l；および (4) L：124 ng/ μ l であった。

【0529】

捕捉のために、10 μ l の gDNA 試料を、2 分間にわたり 98 $^{\circ}$ C まで加熱して（鎖の解離を達成し）、氷上で冷却した。4 倍濃度の結合緩衝液 5 μ l および 49 種のプローブプール（配列番号 150～198）5 μ l（50 nM のユニバーサルオリゴ 61 と組み合わせた、1 nM の各プローブ）を添加し、ミックスをアニーリングさせた（98 $^{\circ}$ C で 2 分間に続いて、1 $^{\circ}$ C ずつ逐次的に 69 $^{\circ}$ C まで温度を低下させる、4 分間にわたるインキュベーション）。複合体を、2 μ l の MyOne ストレプトアビジンビーズに結合させ、これ

10

20

30

40

50

を、180 μ l の TE zero 中に、30 分間にわたり懸濁させ（総容量を 200 μ l とする）、25% のホルムアミド洗浄液で5分間ずつ4回にわたり洗浄し、TE zero で1回洗浄し、上清をビーズ複合体から抜き取った。

【0530】

プロセッシングおよびアダプターライゲーションのために、60 μ l の水；10 μ l の NEB「Cut Smart」緩衝液；50% の PEG 8000 15 μ l；10mM の ATP 10 μ l；1mM の dNTP ブレンド 1 μ l；1 μ l の T4 遺伝子 32 タンパク質（NEB）；および 0.5 μ l の T4 DNA ポリメラーゼ（NEB）を含有する、100 μ l の T4 ミックスを作製した。25 μ l のミックスを、4 例の試料の各々に添加し、20 $^{\circ}$ C で15分間にわたりインキュベートするの続き、70 $^{\circ}$ C で10分間にわたるインキュベーションを施して、T4 ポリメラーゼを熱不活化した。不活化ステップの後、1.25 μ l のアダプター（10 μ M）および1.25 μ l の HC T4 DNA リガーゼを添加した。この混合物を、22 $^{\circ}$ C で30分間および65 $^{\circ}$ C で10分間にわたり、さらにインキュベートした。

10

【0531】

ライブラリーフリー法の1つの魅力的な特色は、プロセッシングされた複合体を、ビーズにさらに接合させることである。ビーズをライゲーション緩衝液から引き寄せ、200 μ l の TE zero で1回洗浄した。次いで、複合体を2 μ l 中に再懸濁させた。増幅のために、20 μ l の容量中の単一プライマーによる増幅を使用して、標的断片を増幅し、かつ、プローブ「スタブ」を超える長いゲノム断片を濃縮した。増幅の後、全長プライマーを伴う大容量の PCR 反応物を使用して、「配列準備された」ライブラリーを創出した。

20

【0532】

57 μ l の水と、5 倍濃度の Q5 反応緩衝液 20 μ l と、10 μ l の単一プライマー 117（表 38 を参照されたい）と、10mM の dNTP 2 μ l と、1 μ l の Q5 ホットスタートポリメラーゼを組み合わせることにより、Q5 ベースの単一プライマーによる PCR 増幅緩衝液を作製した。18 μ l ずつを、各チューブに添加するの続き、20 サイクルにわたり増幅した（98 $^{\circ}$ C で30秒間；98 $^{\circ}$ C で10秒間、69 $^{\circ}$ C で10秒間、72 $^{\circ}$ C で10秒間の20 サイクルにわたり増幅し、10 $^{\circ}$ C で保持した）。PCR の後、ビーズをプルアウトし、20 μ l の増幅前上清を、163.5 μ l の水、5 倍濃度の Q5 緩衝液 60 μ l、15 μ l のフォワードプライマー 118（10 μ M）、15 μ M のリバースプライマー 119（10 μ M）、10mM の dNTP 6 μ l、13.5 μ l の EvaGreen+ROX 色素ブレンド（ROX を1に対して EG を1.25 とする）、および 3 μ l の Q5 ホットスタートポリメラーゼを含有する（全ての反応物への色素の添加は意図しなかった）、280 μ l の PCR ミックスに移した。100 μ l のアリコート 2 つは、従来の PCR（98 $^{\circ}$ C で10秒間、69 $^{\circ}$ C で10秒間、72 $^{\circ}$ C で10秒間）で増幅し、4 連の 10 μ l アリコートは、qPCR 条件下で増幅した。図 38 は、4 例の試料全てについて観察された増幅プロットを示す。反応は、PCR を想起させる変曲点／プラトーを経ていると考えられ、従来の反応は、20 サイクルで停止させた（これで、合計 40 のサイクルの PCR）。図 39 は、これらの増幅反応の産物を含有する、2% のアガロースゲルを示す。図 40 は、ビーズによる精製の後における増幅産物を含有する、2% のアガロースゲルを示す。

30

40

【0533】

本明細書の他の箇所で記載されている、十分に検証された qPCR 捕捉アッセイを使用して、ライブラリーフリー試料をアッセイして、遺伝子特異的標的が捕捉され、選択的に増幅されるのかどうかを決定した。アッセイ 1～16 の標的領域を表 40 に示す。

【0534】

【表 4 0】

表40: アッセイ1～16の標的領域

アッセイ	配列
アッセイ 1	PLP1 エキソン2
アッセイ 2	PLP1 エキソン2
アッセイ 3	PLP1 エキソン2
アッセイ 4	PLP1 エキソン2の上流
アッセイ 5	PLP1 エキソン2の下流
アッセイ 6	PLP1 エキソン2の200bp下流
アッセイ 7	PLP1 エキソン3
アッセイ 8	Chr 9, 標的外
アッセイ 9	CYP2D6
アッセイ 10	Chr X: 154376051
アッセイ 11	Chr X: 154376051
アッセイ 12	Chr X: 6929264
アッセイ 13	KRAS 領域1
アッセイ 14	KRAS 領域2
アッセイ 15	Myc 領域2
アッセイ 16	Myc 領域2

10

20

【0 5 3 5】

qPCR解析のために、試料Fから10ng/μl (2μlを8μlのPCRミックスに添加して、それぞれ、10μlおよび2ng/μlの最終容量および最終濃度をもたらす) で得られるゲノムDNAを、対照として使用した。F試料およびS試料から精製されたプロセッシング材料を、0.01ng/μl=10pg/μlに希釈し、2μlを、8μlずつのPCR反応に添加して、2pg/μlの最終濃度をもたらした。結果を表41に示す。

30

【表 4 1】

表41: qPCRの結果

Cqs	アッセイ1	アッセイ2	アッセイ3	アッセイ4	アッセイ5	アッセイ6	アッセイ7	アッセイ8
gDNA	23	22	25	22	22	23	23	19
F-lib	13	13	13	13	13	20	12	35
S-lib	14	14	13	13	14	20	13	N/A
	アッセイ9	アッセイ10	アッセイ11	アッセイ12	アッセイ13	アッセイ14	アッセイ15	アッセイ16
gDNA	21	21	22	22	22	21	21	21
F-lib	11	13	13	16	18	17	14	13
S-lib	11	13	14	17	17	16	13	12
Abs	アッセイ1	アッセイ2	アッセイ3	アッセイ4	アッセイ5	アッセイ6	アッセイ7	アッセイ8
gDNA	16	24	4	17	22	15	8	146
F-lib	12150	13296	8874	16739	13738	79	22495	0
S-lib	6142	7704	10410	10093	6425	132	14648	
	アッセイ9	アッセイ10	アッセイ11	アッセイ12	アッセイ13	アッセイ14	アッセイ15	アッセイ16
gDNA	60	39	18	30	17	43	49	37
F-lib	45447	11671	11034	1288	440	853	6373	16448
S-lib	68935	12477	7634	678	1077	1203	10171	19305
Adj	アッセイ1	アッセイ2	アッセイ3	アッセイ4	アッセイ5	アッセイ6	アッセイ7	アッセイ8
gDNA	16	24	4	17	22	15	8	146
F-lib	12150333	13296288	8874203	16739230	13737864	79482	22495321	2
S-lib	6141637	7703565	10410406	10093396	6424647	131876	14648060	0
	アッセイ9	アッセイ10	アッセイ11	アッセイ12	アッセイ13	アッセイ14	アッセイ15	アッセイ16
gDNA	60	39	18	30	17	43	49	37
F-lib	45447322	11670722	11034289	1288163	440236	852919	6373326	16447972
S-lib	68934509	12476684	7634492	677562	1077493	1203428	10170993	19304903
Adj	アッセイ1	アッセイ2	アッセイ3	アッセイ4	アッセイ5	アッセイ6	アッセイ7	アッセイ8
F-lib	778820	562525	2393851	987542	627702	5247	2657954	0
S-lib	393671	325914	2808248	595466	293551	8706	1730754	0
	アッセイ9	アッセイ10	アッセイ11	アッセイ12	アッセイ13	アッセイ14	アッセイ15	アッセイ16
F-lib	757399	300823	627284	42799	25665	19716	130461	439562
S-lib	1148823	321597	434010	22512	62817	27819	208198	515912

【0 5 3 6】

qPCRデータにより、ライブラリーフリー技術が、標的化ゲノム領域を回収し、標的外領域を棄却するのに極めて効果的である（例えば、アッセイ6、8）ことが指し示された。＞500、000倍であることが多い精製倍数は、本明細書の他の箇所で開示されるライブラリーにより作成された、早期の実験から得られるデータと直接比較可能であった。

【0 5 3 7】

II部：LOO（leave-one-out）解析：複合体プロセシングの酵素要件を査定した。実験のデザインを、下記の表42に示す。

【表 4 2】

表42: 実験デザイン

実験:	1	2	3	4	5
T4 DNAポリメラーゼ	なし	なし	あり	あり	あり
T4遺伝子の32のタンパク質	なし	あり	なし	あり	あり
T4 DNA リガーゼ	なし	あり	あり	なし	あり

【0 5 3 8】

解析のための捕捉複合体を作製するため、12種の同一な反応物を創出した。135 ng/μlの超音波処理したgDNA 10 μlを、上記で記載した通りに、融解させ、タグ付けされた捕捉プローブとアニーリングさせ、ストレプトアビジンでコーティングされたビーズに結合させ、洗浄し、TE zero中に再懸濁させた。500 μlのプロセシ

グマスターミックスを、 $270\mu\text{l}$ の水と、 $50\mu\text{l}$ の10倍濃度のCutSmart緩衝液と、 10mM のATP $50\mu\text{l}$ と、 50% のPEG8000 $75\mu\text{l}$ と、 10mM のdNTP $5\mu\text{l}$ とを組み合わせることにより調製した。この緩衝液を、 10 の $90\mu\text{l}$ アリコートに分け（2連の試験を実行した）、酵素を、上記で記載した量で添加した（マスターミックス $90\mu\text{l}$ 当たり、 $1\mu\text{l}$ のT4遺伝子32タンパク質、 $0.5\mu\text{l}$ のT4ポリメラーゼ、 $5\mu\text{l}$ のアダプター、および／または $5\mu\text{l}$ のHC T4リガーゼを添加した）。上記で記載したT4による埋め込みおよびライゲーションの後、複合体を、プロセッシングミックスを含まないTEZero中で洗浄し、 $2\mu\text{l}$ のTEZero中に再懸濁させた。複合体を、最終容量 $20\mu\text{l}$ ずつの、単一プライマーによる増幅ミックス中に再懸濁させ、上記で記載した通りに、 20 サイクルにわたり増幅した。次いで、磁石を使用してビーズを引き寄せ、 $20\mu\text{l}$ の洗浄された増幅物を、 $180\mu\text{l}$ の全長F+R（ $118+119$ ）PCR増幅ミックスに希釈した。 $50\mu\text{l}$ をqPCR解析のために取り置き、残りの $150\mu\text{l}$ を2つに分割し、従来のPCRにより増幅した。 $50\mu\text{l}$ のqPCR試料を、 $2.5\mu\text{l}$ の色素ブレンドと混合し、 $10\mu\text{l}$ のアリコートを、蛍光の変化によりモニタリングした。この実験のトレースを、図41に示す。

10

【0539】

2つの従来のPCRのためのアリコートのうちの一方を、 10 サイクル目に取り置き、他方をPCRの 16 サイクル目に取り置いた。これらの生のPCR反応物（ $5\mu\text{l}$ ずつの反応物）のアリコートを、 2% のアガロースゲル上で解析した。結果を、図42に示す。驚くべき結果は、3種の酵素全てが、増幅用のライブラリー材料の効率的な作製に必要とされることである。より微細な観察は、 10 サイクル目における3種全ての酵素による材料のサイズ分布は、 16 サイクル目で現れる、P+L単独によるサイズ分布より著明に大きいことである。

20

【0540】

初期の探索から得られたqPCRデータと併せた、これらのデータにより、分子密集剤PEG8000の存在下における、T4遺伝子32タンパク質を伴うT4 DNAポリメラーゼ（前者の寄与は査定しなかった）は、捕捉プローブ上で捕捉されたゲノム材料を効率的にコピーすることが可能であるという解釈が実証される。

【0541】

III部：ライブラリーフリーシーケンシングライブラリーの作製：上記で記載した方法を使用して、本報告の「材料」節に示される4例のCoriel試料を伴うDNAシーケンシングライブラリーを作製した。4例の試料のうちの各1例を、最終的なPCRステップにおいて、個別のインデックスコードでコード処理した。最終的なライブラリー構成要素（プーリングの前に個別に示される）を、図43のゲル画像に示す。「通常の」ライブラリースメアは通例、 175bp から上方に広がる。この図では、最も小さな断片でも、 $>300\text{bp}$ である。同様に、最も大きな断片は、 750bp 以上であると考えられる。大きな断片は、最適のライブラリーをもたらさない。これらの試料は全て、 80% のビーズ：試料比で2回にわたり精製した。これらの試料を、 $16.9\text{ng}/\mu\text{l}$ （推定平均インサートサイズを 400bp として、約 65nM である）のプールにプールした。試料をシーケンシングした。

30

40

【0542】

ライブラリーフリー法は、CNV解析に良好に働いた。X連鎖遺伝子であるPLP1に独特のリードカウントを、常染色体の遺伝子座であるKRASおよびMYCに照らして標準化した。これらのデータのプロットを、図44に示す。データは、ライブラリーフリー手順では、絶対コピー数が失われることを例示する（KRASの、MYCに照らした「コピー」は、もはや同等ではない）。しかし、相対コピー数（PLP1の、常染色体の標準化子と比べた変化）は、頑健に検出される。シーケンシング結果はまた、プローブに照らしたリード開始部位と関連する顕著な特色も示した。図45は、リードは、プローブから 900bp の距離にある限りにおいて、検出され、座標 $1100\sim 1300$ の間では、あらゆる単一の開始点が、複数回にわたり使用されることを示す。これらのデータによ

50

り、リードは、あらゆる単一の可能な塩基位置において開始され、ライゲーション／プロセシングバイアスは、ほとんど見られないことが指し示された。加えて、プローブから100bp以内で開始されるリードが極めて少数であることは、ゲル上で観察される、ライブラリーの極めて大きなサイズ分布と符合する。

【0543】

(実施例20)

標的化遺伝子発現解析

総括

本実施例は、標的化遺伝子発現ライブラリーの開発について実証する。投入するのは、RNAであり、DNAではなく、したがって、二本鎖cDNA合成ステップが必要とされる。好ましい方法は、RNAアーゼH⁻リバーストランスクリプターゼまたはRNAアーゼH⁻様活性を呈示するキット（例えば、Promega製のGoScript）を使用する第1鎖合成およびランダムヘキサマーによるプライミングである。第2鎖合成に好ましい方法は、E. coli DNAポリメラーゼホロ酵素である、NAD⁺依存性リガーゼおよびRNAアーゼHを含むキット（例えば、New England Biolabs第2鎖cDNA合成モジュール）を使用することである。

【0544】

極めて広範にわたる転写物コピーが存在するため、これに応じて、シアリングされ、末端修復されたcDNAに対するアダプター上に導入される、広範にわたるランダムタグが存在しなければならない。したがって、ランダム8マー（65,536種の可能な配列）を使用した。アダプターを、ランダム8マー配列に続く10～12の固定塩基であって、ライゲーションを容易とする、相補的な10～12塩基のためのアニーリング部位として用いられる可能性があり、かつ、マルチプレックス化された試料の場合に試料識別子として使用される固定塩基で操作した。

【0545】

重複リードと対比した独特のリードの実際の数（言い換えるとリードの統計学的分布）は、発現レベルの決定において1つの重要な因子である。1つの潜在的なエラーの発生源は、捕捉イベントの後において重複するリードである。これらのエラーを同定するため、各捕捉イベントを標識するように、ランダムタグを、捕捉プローブに添加した。

【0546】

標的化RNA-seqライブラリーの処理およびシーケンシングも、ゲノムライブラリーの処理と同じ手順に従う。

【0547】

インフォマティクス解析は、捕捉後の重複リードの除去および標的のトランスクリプトームに照らしたアラインメントと共に始まる。次いで、整列しているリードの間の独特のリードカウントを決定する。次いで、データが統計学的分布に適合しうる一方で、生の独特のリードカウントが、実際の発現レベルに対する極めて緊密な近似であることが分かった。

【0548】

目的

これらの実験の目的は、標的化発現シーケンシングライブラリーを、心臓および肝臓の全RNAから作製し、全RNAライブラリーおよび標的化RNAライブラリーの両方を、直接的な比較がなされうる、同じ出発材料から作製することであった。

【0549】

概要

全RNAライブラリーから得られるRNAカウントと、標的化RNAライブラリーから得られるRNAカウントとは、2つのパラメータに沿って良好な一致を示した。第1に、心臓試料の発現比と、肝臓試料の発現比とは、良好に相関した。第2に、全RNAカウントを、標的化RNAカウントと比較したところ、所与の試料中の異なる転写物の絶対存在量の測定は、良好に一致した。これらの初回通過データにより、定量的な標的化核酸法が

、ゲノムDNAを越えて、cDNAライブラリーについての解析に広がりうることが指示された。

【0550】

戦略

rRNAを枯渇させた、妥当な全RNAライブラリーを創出するために、dTプライミングを使用した。標的化RNAライブラリーを創出するために、全RNA試料を、IDTから供給されているランダムヘキサマーでまずプライミングした。ランダムヘキサマープライミングは、トランスクリプトームの最も包括的なカバレッジをもたらす可能性が高い。P1フローセル配列およびP2フローセル配列を導入するPCRプライマーによる増幅の後で、全RNAライブラリーをシーケンシングした。標的化解析のために、本明細書の他の箇所想定される、捕捉ステップ、洗浄ステップ、プロセッシングステップ、および増幅ステップを実行した。次いで、標的化クローンをシーケンシングした。

10

【0551】

方法

オリゴヌクレオチド：全RNAライブラリーには、ポリdTプライマー：TTTTTTT TTTTTTTTTTTTTTVN（配列番号722）を使用した。標的化RNA-seqには、最初の8塩基がランダムであり、次の12塩基が「コード」配列として用いられ、したがって、ライゲーション可能な二重鎖を形成しうる12マーのパートナー鎖オリゴのアンカー配列としても用いられる、アダプターデザインを創出した。

20

【0552】

これらのアダプターの配列は、以下の通りであった。

【化6】

配列番号	名称	配列
723	RNA_L_1	TGCAGGACCAGAGAATTCGAATACANNNNNNNNNGACTCCGATCCC
724	RNA_P_1	GGGATCGGAGTC
725	RNA_L_2	TGCAGGACCAGAGAATTCGAATACANNNNNNNNNCGGAACTCGGAG
726	RNA_P_2	CTCCGAGTTCCG

【0553】

30

cDNAライブラリー構築：以下の方法を使用して、以下の4種のcDNAライブラリー：（1）心臓の全RNA（dTプライミングされた）；（2）心臓の標的化RNA（N₆プライミングされた）；（3）肝臓の全RNA；および（4）肝臓の標的化RNA（N₆プライミングされた）を合成した。1μg/μlの全RNAを、TEz中100ng/μlに10倍に希釈した。以下の成分：2μlの希釈された全RNA（100ng）；2μlの5uMのポリdTVNプライマー、または50uMのN6（IDT）2μl；および6μlの水を、10μlの総容量で組み合わせ、65℃まで加熱し、氷に移した。

【0554】

ミックスを、10μlの第1鎖カクテル（5倍濃度のGoScript緩衝液4μl；25mMのMg⁺⁺1.6μl（2mMの最終濃度）；10mMのdNTP 1.0μl（500uMの最終濃度）；1.0μlのGoScript酵素；および2.4μlの水）と組み合わせ、42℃で30分間にわたり、次いで、70℃で10分間にわたりインキュベートした。60マイクロリットルの第2鎖合成試薬（48μlの水、ミックス；10倍濃度の第2鎖合成緩衝液8μl；4μlの第2鎖酵素ミックス）を、各反応物に添加し、16℃で2時間にわたりインキュベートした。

40

【0555】

第2鎖を合成した後、55μlのTEzを、各反応物に添加し、反応物を、ガラス製のCovaris超音波処理チューブに移し、約500bpに超音波処理した。125μlの超音波処理試料を、PCRストリップチューブに移し、125μlのビーズを添加した。精製の後、試料を、20μlの最終容量に再懸濁させた。

50

【0556】

本明細書で想定される方法を使用して、19 μ lの反応物について、末端修復を実行した。次いで、末端修復された断片を、22℃で30分間にわたり、アダプターにライゲーションし、65℃で10分間にわたり熱不活化した。40 μ lの最終容量：25 μ lの修復断片；10 μ Mのアダプター2 μ l（10 μ MのL鎖、20 μ MのP鎖）；10倍濃度の緩衝液4 μ l；50%のPEG8000 6 μ l；1 μ lの水；および2 μ lのHCT4リガーゼ中でライゲーションを実行した。60 μ lのTEZおよび100 μ lのビーズを、各反応物に添加し、試料を、最終容量の20 μ lに精製した。

【0557】

qPCRを使用して、ライブラリーによる増幅をモニタリングし、20 μ lの精製されたライゲーションミックスを、130 μ lのPCRミックス（75 μ lの2倍濃度のNEBNextマスターミックス、15 μ lのACA2-20、40 μ lの水）と組み合わせることにより、PCRを介して、各ライブラリーを増幅した。50 μ lを、2.5 μ lのEvaGreen+ROXを含有するウェルにアリコート分割し、qPCRプレート内で10 μ lにさらにアリコート分割した。残りの100 μ lは、PCRストリップチューブ内で保存した。72℃で30秒間、98℃で30秒間にわたり、かつ、98℃で10秒間、60℃で10秒間、72℃で10秒間の可変サイクルでPCR増幅を実行した。

【0558】

dTライブラリーのために、100 μ lのPCR反応物を、120 μ lのビーズで精製した。ACA2-20（20ヌクレオチドのPCRプライマー）で増幅された材料を20倍に（5 μ lを、50 μ lの2倍濃度のNEBNextマスターミックス、5 μ lのFプライマー、5 μ lのRプライマー、および35 μ lの水を含有する95 μ lのPCRミックスに）希釈した。Fプライマーは、ACA2__FLFP AATGATACGGCGACCCACCGAGATCTACACGTCATGCAAGGACCCAGAGAAATTCGAATACA（配列番号69）である、オリゴ8であり、リバースプライマーは、エキソームCAC3__FLRP CAAGCAGAAAGACGGCATACGAGATGTGACTGGCACCGGGACCCAGAGAAATTCGAATACA（配列番号74）である、オリゴ63である。増幅は、8サイクルにわたり実行した。このステップは、短い、20bpのACA2末端配列を、フローセル適合的で、シーケンシング可能な、長いP1配列およびP2配列に成長させるように組み入れた。2種の異なるプライマーを取る構築物が増幅されるのに対し、1種の配列だけを有する構築物は、抑制されるであろう。100 μ lのビーズを、100 μ lのPCR反応物に添加し、最終容量の50 μ l中に再懸濁させることにより、結果として得られるDNAを精製した。

【0559】

DNAを、Qubitにより定量化し、ゲル電気泳動により検討した。ACA2__FSP ACACGTCATGCAAGGACCCAGAGAAATTCGAATACA（配列番号68）であるフォワードプライマーオリゴ7、およびエキソームCAC3__RSP GTGACTGGCACCGGGACCCAGAGAAATTCGAATACA（配列番号73）であるリバースプライマーオリゴ62を使用して、DNAをシーケンシングした。dTプライミングされたRNAを、ラン__48および49においてシーケンシングした。

【0560】

DNAゲル試料を、図46に示す。dTプライミングされた全RNAライブラリーの大型断片のサイズ分布は、やや驚くべきものであった。

【0561】

標的化RNAシーケンシングのために、N₈プライミングしたライブラリーを、40 μ lのTEZ中に再懸濁させた。心臓ライブラリー中および肝臓ライブラリー中の断片含量を定量化した：153 fg/ μ lの心臓試料cDNAおよび760 fg/ μ lの肝臓試料cDNA。これらのデータに基づき、40 μ lの心臓ライゲーション試料および8 μ lの肝臓ライゲーション試料を、後段のPCR増幅に送り込んだ。

【0562】

10

20

30

40

50

qPCRを使用して、ライブラリーによる増幅の進行をモニタリングし、40 μ lの精製されたライゲーションミックス（心臓）または8 μ lのライゲーションミックス+32 μ lのTE2（肝臓）を、210 μ lのPCRミックス（125 μ lの2倍濃度のNEB Nextマスターミックス、25 μ lのACA2、60 μ lの水）と組み合わせるPCRにより、ライブラリーを増幅した。50 μ lを、2.5 μ lのEvaGreen+ROXを含有するウェルにアリコート分割し、qPCRプレート内で10 μ lのアリコートにさらにアリコート分割した。残りの100 μ lを、PCRストリップチューブに入れた。72℃で30秒間、98℃で30秒間にわたり、かつ、98℃で10秒間、60℃で10秒間、72℃で10秒間の可変サイクルでPCR増幅を実行した。200 μ lのPCR産物を、200 μ lのビーズで精製し、最終容量の25 μ l中に再懸濁させた。PCR産物の濃度は、心臓ライブラリーが41 ng/ μ lであり、肝臓ライブラリーが42 ng/ μ lであった。

10

【0563】

捕捉のために、心臓試料と肝臓試料とを組み合わせ、タグ付けされたRNA-seq特異的プローブ（配列については、以下の付録を参照されたい）を伴う、2種の「2重」捕捉反応を実行し、洗浄し、プロセッシングし（C+P）（最終的な収率=40 μ lの23 ng/ μ l）、240~600 bpの断片を、Pippin自動式DNAサイズセクター上でサイズ選択した。5.4 ng/ μ lの断片を、Pippin=20.8 nMから回収した。フローセルに、断片をロードし、51ヌクレオチドの第1のリードおよび24ヌクレオチドの第2のリードを回収した。

20

【0564】

結果および考察

有用なRNA-seqデータを決定するため、肝臓試料と対比した心臓試料を、比較のために選択した。一方の組織内または他方の組織内のそれらの絶対存在量（心臓または肝臓におけるRPKM値を約100、10、1などとする）、およびそれらの組織間の比（ここでもまた、肝臓と対比した心臓の比を、約100、10、1、0.1、0.01とする）に基づき、RNA-seqアトラス（medicalgenomics.org/rna_seq_atlas）において報告された転写物のうちの21種の転写物を解析した。候補転写物のリスト、およびそれらの報告されたRPKM値を、表43に示す。

【表 4 3】

表43:候補RNA-seqの転写物

hgnc記号	転写物	転写物の長さ	心臓	肝臓	h/l	l/h
MYH7	NM_000257	6030	2137.51	0.007	305358.6	0.0
NESL	NM_006393	9213	115.617	0.097	1191.9	0.0
MGP	NM_000900	661	1059.49	11.155	95.0	0.0
DKK3	NM_013253	2746	118.464	1.136	104.3	0.0
SVES	NM_147147	5505	11.27	0.11	102.5	0.0
POE4DIP	NM_001002812	4824	108.053	10.053	10.7	0.1
HAND2	NM_021973	2368	10.069	0.992	10.2	0.1
RP9P	NR_003500	1297	1.279	0.125	10.2	0.1
SOD2	NM_000636	1593	492.028	500.782	1.0	1.0
ATP5E	NM_006886	417	101.927	93.268	1.1	0.9
TRAPPC1	NM_021210	819	10.197	9.646	1.1	0.9
TAF1A	NM_005681	1879	1.101	1.035	1.1	0.9
MGST1	NM_145764	910	43.368	429.307	0.1	9.9
PDIA4	NM_004911	2952	14.701	154.615	0.1	10.5
STARD10	NM_006645	1988	1.163	11.147	0.1	9.6
TMEM14A	NM_014051	1014	0.11	1.119	0.1	10.2
APOB	NM_000384	14121	5.822	765.522	0.0	131.5
SERPINF2	NM_001165921	2092	1.034	102.91	0.0	99.5
TFR2	NM_003227	2888	0.142	13.762	0.0	96.9
FGL1	NM_147203	1337	1.102	1305.36	0.0	1184.5
FGA	NM_000508	3655	24.14	9012.08	0.0	373.3

10

20

【0 5 6 5】

標的化RNA-seqライブラリーを、同じ全RNA試料から作製された、標的化されていない全RNAライブラリーと比較した。ポリdTプライミングを使用して、全RNA、主に、rRNAを、非rRNA転写物ライブラリーに転換した。標的化RNA-seqには、ランダムヘキサマーを使用した。dTプライミング全RNAライブラリーでは、リードを、それらの一部は極めて長い、転写物の全長に沿って導出することができる。例として述べると、心臓内のMYH7に沿ったリードの分布を検討し、この長い転写物の5'端の近傍から得られるリードを見い出した。1つの（長い）転写物を、別の（短い）転写物と比較するために、カウントを、転写物の長さで標準化（reads per kb per million法またはRPKM法と称することが多い）した。この第1度の標準化の後、カウントを、全試料と標的化試料との間でもまた標準化した。最終的なリードカウントデータセットを、表44に示す。

30

【0 5 6 6】

【表 4 4】

表44: 図49および50のリードカウントデータ

心臓	アトラス	合計	標的化	肝臓	アトラス	合計	標的化
MYH7	50188	71792	40164	MYH7	0	374	721
NEEL	2715	936	3460	NEEL	1	27	24
MGP	24876	1315	13162	MGP	90	42	393
DKK3	2781	18092	20763	DKK3	9	36	211
BVES	265	179	650	BVES	1	5	62
PDE4DIP	2537	850	5029	PDE4DIP	81	88	294
HAND2	236	479	1052	HAND2	8	42	279
RP9P	30	121	63	RP9P	1	20	12
SOD2	11553	780	5437	SOD2	4031	1503	8802
ATP5E	2393	1607	2363	ATP5E	751	451	2173
TRAPPC1	239	536	1970	TRAPPC1	78	80	397
TAF1A	26	22	34	TAF1A	8	7	48
MGST1	1018	1312	2862	MGST1	3455	5514	9281
PDIA4	345	603	1310	PDIA4	1244	1197	2144
STARD10	27	908	1214	STARD10	90	1857	3338
TMEM14A	3	74	57	TMEM14A	9	110	74
APOB	137	56	107	APOB	6162	5243	17030
SERPINF2	24	119	89	SERPINF2	828	7474	7795
TFR2	3	20	11	TFR2	111	3395	3489
FGL1	26	50	44	FGL1	10507	21763	11420
FGA	567	131	160	FGA	72538	50772	32014

【0 5 6 7】

目視により、3種のデータ全て（アトラスデータ、全試料データ、および標的化試料データ）の間の良好な相関が明らかとなった。いずれのデータセットも、同じ全RNA試料から導出されたため、1つの重要な比較は、本明細書で調製される全RNA-seq試料と、本明細書で調製される標的化RNA-seq試料との比較であった。比較の2つの重要点は、（1）心臓における実際の発現比と、肝臓における実際の発現比との相関；および（2）特異的試料内の転写物の絶対存在量の、全試料カウントと標的化試料カウントとの間の相関を含む。

【0 5 6 8】

第1の点では、発現プロファイルの保存に取り組むが、比較されるカウントの実際の大きさは無視する。心臓における発現比を、肝臓における発現比と対比した比較プロットを、標的化試料における発現比と対比した、全試料における発現比について、図47に示す。このプロットは、2つの方法により作成された「発現プロファイル」の間の例外的な相関（ $r^2 = 0.95$ ）を示す。

【0 5 6 9】

第2の点は、2つの方法の間の絶対比較について問うため、より厳密でありうる。全RNA-seqまたは標的化RNA-seqにおいて測定される絶対発現レベルの比較を、図48に示すが、ここで、 $\log_{10}$ （カウント）値は、互いに対してプロットされる。この比較は、標的化に対して高感度であるだけでなく、また、RNA-seqライブラリーを根本的に異なる方法（全RNA-seqライブラリーには、dTプライミングであり、標的化RNA-seqライブラリーには、ランダムヘキサマープライミングであった）で調製したという事実に対しても高感度であった。異なる調製法にもかかわらず、2つの方法の間には、優れた相関が認められた。

【0 5 7 0】

本研究は、配列特異的捕捉と組み合わせられたランダムタグで標識づけするコア法により、絶対発現存在量を保存し、転写物特異的配列情報を明らかにし、トランスクリプトームデータの複雑性を劇的に軽減する、標的特異的RNA転写物データを作成しうることを実

証した。

【0571】

【化7-1】

付録:RNA-seq解析のために使用されたプローブ配列

配列 番号	名称	配列
727	APOB_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNGCCGAAG AGGAAATGCTGGAAAATGTCAGCCTGGTCTGTCCAAAAGATGCGACCC GATTC
728	APOB_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNCAGCTGAC CTCATCGAGATTGGCTTGGAAGGAAAAGGCTTTGAGCCAACATTGGAA GCT
729	ATP5E_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNTACTGGAG ACAGGCTGGACTCAGCTACATCCGATACTCCAGATCTGTGCAAAAGC AGTG
730	ATP5E_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNCTGAAGA CAGAATTCAAAGCAAATGCTGAGAAGACTTCTGGCAGCAACGTAAAAA TTGTG
731	BVES_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNACAGAGT CCAGCCCATTGAGAGAATCAACTGCCATAGGTTTTACACCTGAGTTAGA AAGT
732	BVES_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNAGCCAGCT CCAGTGACAGTGACGACGGCTTGACACCAGTTTCTTCGGGGTACCTCCAG

10

20

30

【化 7 - 2】

		CA	
733	DKK3_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNGACACGC AGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCT GCTGC	
734	DKK3_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNTGGTGTAT GTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATC CTG	10
735	FGA_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNATGAGGA TCGTCTGCCTGGTCCTAAGTGTGGTGGGCACAGCATGGACTGCAGATAG TGGT	
736	FGA_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNATTTTTTC TGTTTATTGCGATCAAGAGACCAGTTTGGGAGGATGGCTTTTGATCCAG CA	20
737	FGL1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNGTGTTTCAG TTTCATCCTTGTTACCACCGCTCTGACAATGGGCAGGGAAATTCGGCG CTC	
738	FGL1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNTGCAAACC TGAATGGTGTATACTACAGCGGCCCTACACGGCTAAAACAGACAATG GGA	30
739	HAND2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNCATCCCCA ACGTACCCGCCGACACCAAACCTCTCCAAAATCAAGACCCTGCGCCTGG CCA	
740	HAND2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNCGACCAG AATGGCGAGGCGGAGGCCTTCAAGGCAGAGATCAAGAAGACCGACGT GAAAG	
741	MGP_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNCTGATCCT TCTTGCCATCCTGGCCGCCTTAGCGGTAGTAACTTTGTGTTATGAATCAC AT	40
742	MGP_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNCGAGCTCA	

【化 7 - 3】

		ATAGGGAAGCCTGTGATGACTACAGACTTTGCGAACGCTACGCCATGG TTT
743	MGST1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNAAGAAGT ATCTTCGAACAGATGACAGAGTAGAACGTGTACGCAGAGCCCACCTGA ATGA
744	MGST1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNTATTCCTT GAGTGGTCCCGACCCCTCTACAGCCATCCTGCACTTCAGACTATTTGTC GG
745	MYH7_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNCAGACCA GGCCTTTTGACCTCAAGAAGGATGTCTTCGTGCCTGATGACAAACAGGA GTTT
746	MYH7_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNAGAGCTCT TCAAACCTCAAGAACGCCTATGAGGAGTCCCTGGAACATCTGGAGACCT TCA
747	NEBL_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNCAGGTCTT CTATAAGCCTGTTATTGAAGACTTAAGCATGGAATTGGCCAGAAAATGC ACG
748	NEBL_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNAAAGAGC AAAACCTACAAGGCCACTCCGGTAAGCATGACCCCGGAGATAGAGAGAG TGAG
749	PDE4DIP_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNGGATATCG CACTCTGTCCCAGCACCTCAATGACCTGAAGAAGGAGAACTTCAGCCTC AAG
750	PDE4DIP_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNGAAAGCC AAGCTGCTGCAGAGAAGTTGGTGCAAGCCTTAATGGAAAGAAATTCAG AATT
751	PDIA4_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNCCGGACG AGGATTCTTCTAACAGAGAAAAATGCCATTGAGGATGAAGAGGAGGAGG AGGAG

10

20

30

40

【化 7 - 4】

752	PDIA4_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNCGTCAAG GTCGTGGTGGGAAAGACCTTTGACTCCATTGTGATGGACCCCAAGAAG GACG
753	RP9P_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNAGATCAC GAGTTATGAAAAACCTCCTCCTGGGCTTATCAAGGTTGGCATTGCAAAT GGT
754	RP9P_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNAAAGTTTC TTCATCGTTGTCCTCCCTGCTGGTCACATGAGTTTACGATTCCTTAGAAG TG
755	SERPINF2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNCCCGGCGC GTTCCGACTGGCTGCCAGGATGTACCTGCAGAAAGGATTTCCCATCAAA GA
756	SERPINF2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNGCCTGAA AGGCTTCCCCCGCGGAGACAAGCTTTTCGGCCCTGACTTAAAACTTGTG CCC
757	SOD2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNAACAACCT GAACGTCACCGAGGAGAAGTACCAGGAGGCGTTGGCCAAGGGAGATGT TAC
758	SOD2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNAGGAGAA GCTGACGGCTGCATCTGTTGGTGTCCAAGGCTCAGGTTGGGGTTGGCTT GGT
759	STARD10_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNGACCAAG ACTTTCGCAGCTTCCGGTCAGAGTGTGAGGCTGAGGTGGGCTGGAACCT GACC
760	STARD10_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNAGATGTAC AAGGCGTGCCTCAAGTACCCCGAGTGGAAACAGAAGCACCTGCCTCAC TTCA
761	TAF1A_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNTTCAGTGA AGAATTAAAAGGGCCTGTGACAGATGATGAAGAAGTGGAACATCTGT

10

20

30

40

【化 7 - 5】

		GCTC	
762	TAF1A_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNGAAACCA CCTTGCGTGGGTTCAAGAAGAGTGGAACCTCCAGGAAAAACTGGTGGCC AGGC	
763	TFR2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNCTTTGGGG TCTATTCCAGAGAGCGCAACAACTGTCCCCAAGATCCTCTCAGACCGTC TAC	10
764	TFR2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNAGGCCTAC CCATTCCCTGCACACAAAGGAGGACACTTATGAGAACCTGCATAAGGTG CTG	
765	TMEM14A_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNATCGGTTT TGTTTATGCAGCCCTCGTGACATTTGGAAGCATTTTGGATATAAGCGG AGA	20
766	TMEM14A_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNGCTTCTT CCTGGCTACCATAATGGGTGTGAGATTTAAGAGGTCCAAGAAAATAAT GCC	
767	TRAPPC1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNCACAACCT GTACCTGTTTGACCGGAATGGAGTGTGTCTGCACTACAGCGAATGGCAC CGC	30
768	TRAPPC1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNATCCGAG ATGTGCTGCACCACATCTACAGTGCCTGTATGTGGAGCTGGTGGTGAA GAA	
769	APOB_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNGCCGAAG AGGAAATGCTGGAAAATGTCAGCCTGGTCTGTCCAAAAGATGCGACCC GATTC	
770	APOB_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNCAGCTGAC CTCATCGAGATTGGCTTGGAAGGAAAAGGCTTTGAGCCAACATTGGAA GCT	40
771	ATP5E_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNTACTGGAG	

【化 7 - 6】

		ACAGGCTGGACTCAGCTACATCCGATACTCCCAGATCTGTGCAAAAGC AGTG	
772	ATP5E_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNCTGAAGA CAGAATTCAAAGCAAATGCTGAGAAGACTTCTGGCAGCAACGTAAAAA TTGTG	
773	BVES_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNACAGAGT CCAGCCCATTGAGAGAATCAACTGCCATAGGTTTTACACCTGAGTTAGA AAGT	10
774	BVES_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNAGCCAGCT CCAGTGACAGTGACGACGGCTTGACACCAGTTTCTTCGGGGTACCTCCAG CA	
775	DKK3_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNGACACGC AGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCT GCTGC	20
776	DKK3_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNTGGTGTAT GTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATC CTG	
777	FGA_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNATGAGGA TCGTCTGCCTGGTCCTAAGTGTGGTGGGCACAGCATGGACTGCAGATAG TGGT	30
778	FGA_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNATTTTTTC TGTTTATTGCGATCAAGAGACCAGTTTGGGAGGATGGCTTTTGATCCAG CA	
779	FGL1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNGTGTTTCAG TTTCATCCTTGTTACCACCGCTCTGACAATGGGCAGGGAAATTTGGGCG CTC	40
780	FGL1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNTGCAAACC TGAATGGTGTATACTACAGCGGCCCTACACGGCTAAAACAGACAATG GGA	

【化 7 - 7】

781	HAND2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNCATCCCCA ACGTACCCGCCGACACCAAACCTCTCCAAAATCAAGACCCTGCGCCTGG CCA
782	HAND2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNCGACCAG AATGGCGAGGCGGAGGCCTTCAAGGCAGAGATCAAGAAGACCGACGT GAAAG
783	MGP_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNCTGATCCT TCTTGCCATCCTGGCCGCCTTAGCGGTAGTAACCTTTGTGTTATGAATCAC AT
784	MGP_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNCGAGCTCA ATAGGGAAGCCTGTGATGACTACAGACTTTGCGAACGCTACGCCATGG TTT
785	MGST1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNAAGAAGT ATCTTCGAACAGATGACAGAGTAGAACGTGTACGCAGAGCCACCTGA ATGA
786	MGST1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNTATTCCTT GAGTGGTCCCGACCCCTCTACAGCCATCCTGCACTTCAGACTATTTGTC GG
787	MYH7_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNCAGACCA GGCCTTTTGACCTCAAGAAGGATGTCTTCGTGCCTGATGACAAACAGGA GTTT
788	MYH7_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNAGAGCTCT TCAAACCTCAAGAACGCCTATGAGGAGTCCCTGGAACATCTGGAGACCT TCA
789	NEBL_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNCAGGTCTT CTATAAGCCTGTTATTGAAGACTTAAGCATGGAATTGGCCAGAAAATGC ACG
790	NEBL_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNAAAGAGC AAAACCTACAAGGCCACTCCGGTAAGCATGACCCCGGAGATAGAGAGAG

10

20

30

40

【化 7 - 8】

		TGAG	
791	PDE4DIP_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNGGATATCG CACTCTGTCCCAGCACCTCAATGACCTGAAGAAGGAGAACTTCAGCCTC AAG	
792	PDE4DIP_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNGAAAGCC AAGCTGCTGCAGAGAAGTTGGTGCAAGCCTTAATGGAAAGAAATTCAG AATT	10
793	PDIA4_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNCCGGACG AGGATTCTTCTAACAGAGAAAATGCCATTGAGGATGAAGAGGAGGAGG AGGAG	
794	PDIA4_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNCGTCAAG GTCGTGGTGGGAAAGACCTTTGACTCCATTGTGATGGACCCCAAGAAG GACG	20
795	RP9P_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNAGATCAC GAGTTATGAAAAACCTCCTCCTGGGCTTATCAAGGTTGGCATTGCAAAT GGT	
796	RP9P_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNAAAGTTTC TTCATCGTTGTCCTCCCTGCTGGTCACATGAGTTTACGATTCCTTAGAAG TG	30
797	SERPINF2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNCCCGGCGC GTTCCGACTGGCTGCCAGGATGTACCTGCAGAAAGGATTTCCCATCAAA GA	
798	SERPINF2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNGCCTGAA AGGCTTCCCCCGCGGAGACAAGCTTTTCGGCCCTGACTTAAACTTGTG CCC	
799	SOD2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNAACAACCT GAACGTCACCGAGGAGAAGTACCAGGAGGCGTTGGCCAAGGGAGATGT TAC	40
800	SOD2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACNNNNNNNAGGAGAA	

【化 7 - 9】

		GCTGACGGCTGCATCTGTTGGTGTCCAAGGCTCAGGTTGGGGTTGGCTT GGT	
801	STARD10_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNGACCAAG ACTTTCGCAGCTTCCGGTCAGAGTGTGAGGCTGAGGTGGGCTGGAACCT GACC	
802	STARD10_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNAGATGTAC AAGGCGTGCCTCAAGTACCCCGAGTGGAACAGAAGCACCTGCCTCAC TTCA	10
803	TAF1A_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNNTTCAGTGA AGAATTAAAAGGGCCTGTGACAGATGATGAAGAAGTGGAACATCTGT GCTC	
804	TAF1A_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNGAAACCA CCTTGCGTGGGTTCAAGAAGAGTGGAAGTCCAGGAAAACTGGTGGCC AGGC	20
805	TFR2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNCTTTGGGG TCTATTCCAGAGAGCGCAACAACTGTCCCCAAGATCCTCTCAGACCGTC TAC	
806	TFR2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNAGGCCTAC CCATTCTGCACACAAAGGAGGACACTTATGAGAACCTGCATAAGGTG CTG	30
807	TMEM14A_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNATCGGTTT TGGTTATGCAGCCCTCGTGACATTTGGAAGCATTTTTGGATATAAGCGG AGA	
808	TMEM14A_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNGCTTTCTT CCTGGCTACCATAATGGGTGTGAGATTTAAGAGGTCCAAGAAAATAAT GCC	40
809	TRAPPC1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNCACAACCT GTACCTGTTTGACCGGAATGGAGTGTGTCTGCACTACAGCGAATGGCAC CGC	

【化 7 - 1 0】

810	TRAPPC1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACNNNNNNATCCGAG ATGTGCTGCACCACATCTACAGTGCGCTGTATGTGGAGCTGGTGGTGAA GAA
-----	-----------	---

【 0 5 7 2】

以下の特許請求の範囲では一般に、使用される用語は、特許請求の範囲を、本明細書および特許請求の範囲で開示される特殊な実施形態に限定するものとみなされるべきではなく、このような特許請求の範囲が権利を付与される均等物の全範囲に沿って、全ての可能な実施形態を含むものとみなされるべきである。したがって、特許請求の範囲は、本開示により限定されない。

10

【図 1】

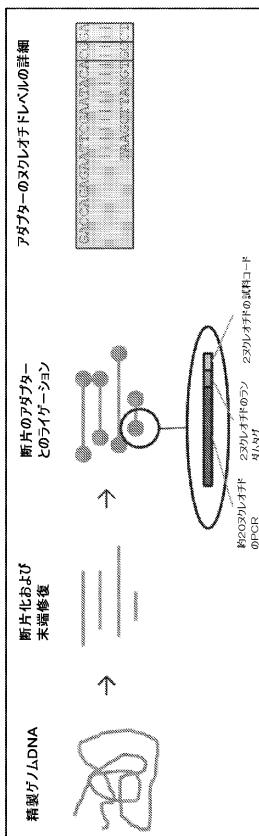


FIG. 1

【図 2】

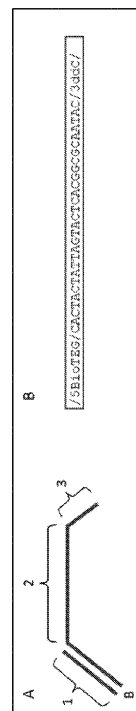


FIG. 2

【図 3】

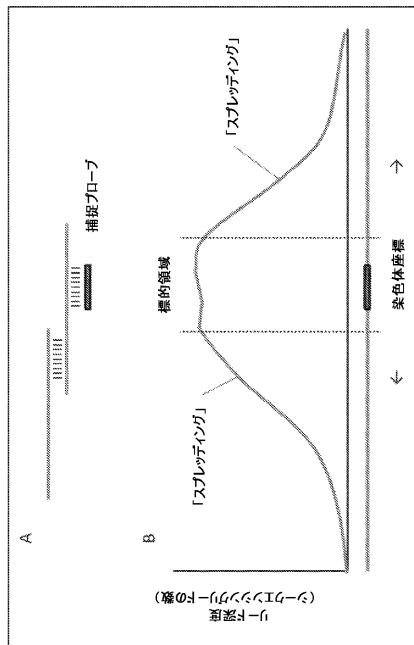


FIG. 3

【図 4】

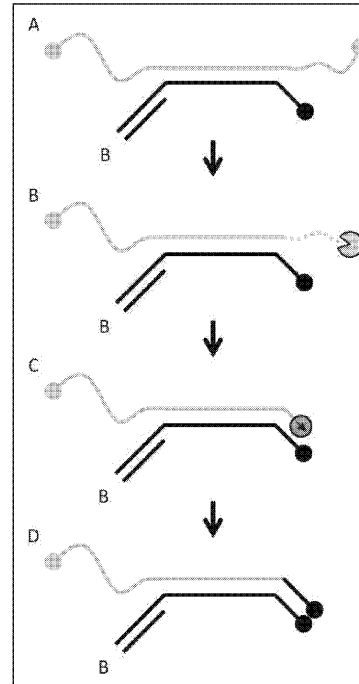


FIG. 4

【図 5】

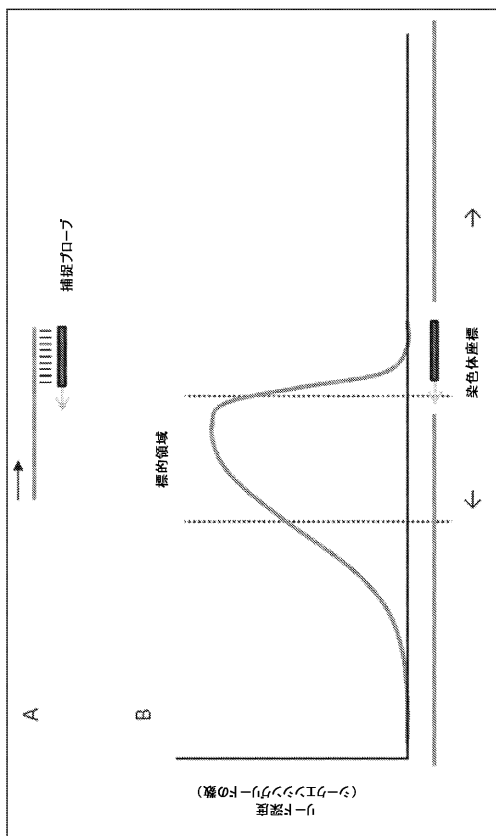


FIG. 5

【図 6】

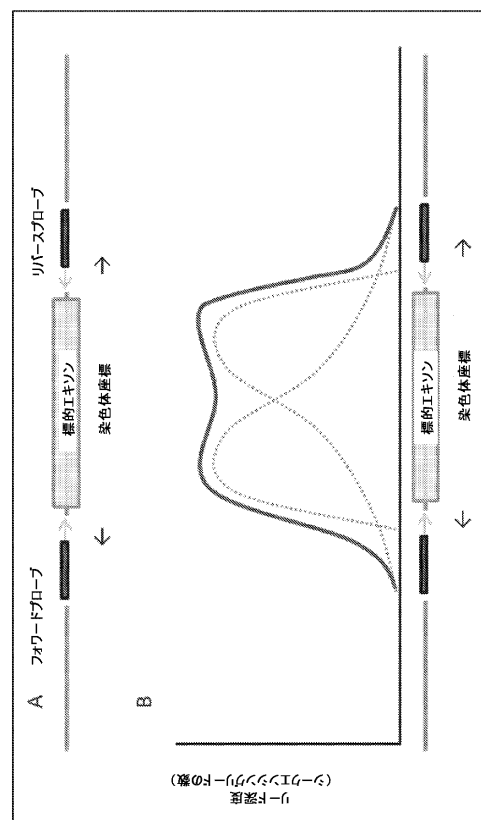


FIG. 6

【図 7】

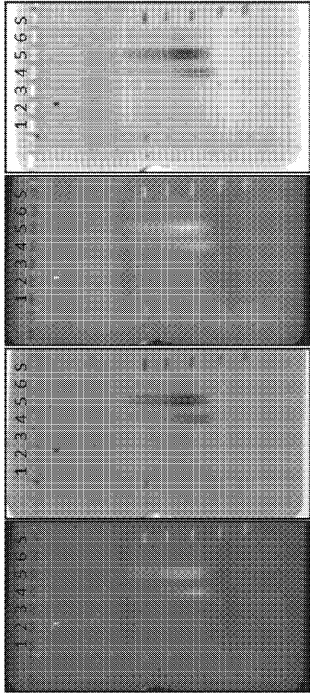


FIG. 7

【図 8】

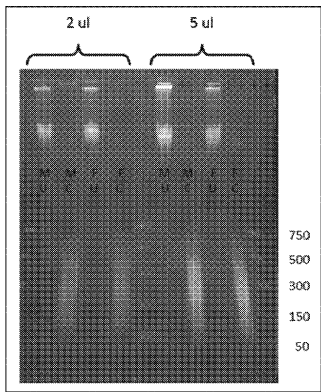


FIG. 8

【図 9】

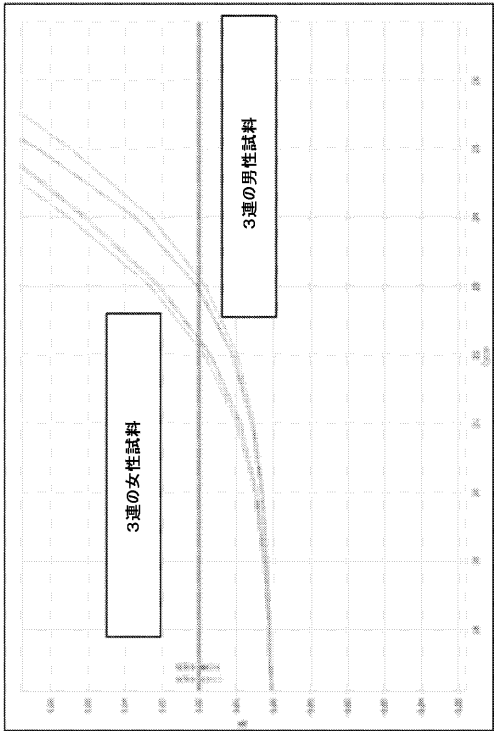


FIG. 9

【図 10】

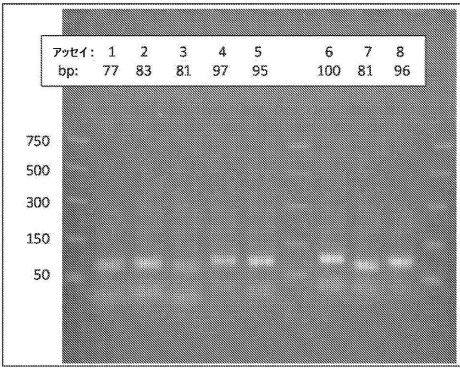
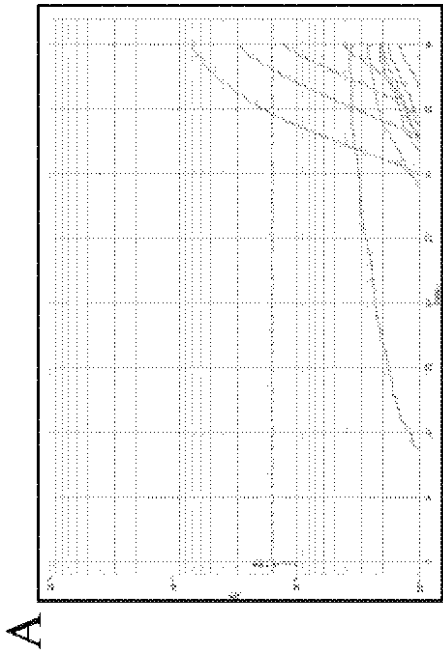
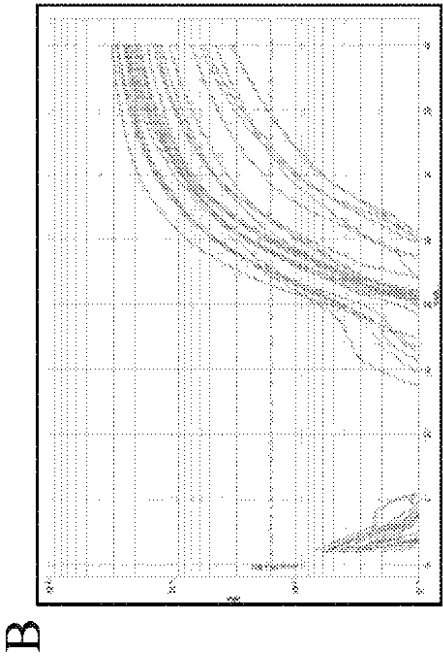


FIG. 10

【図 1 1 A】



【図 1 1 B】



【図 1 2】

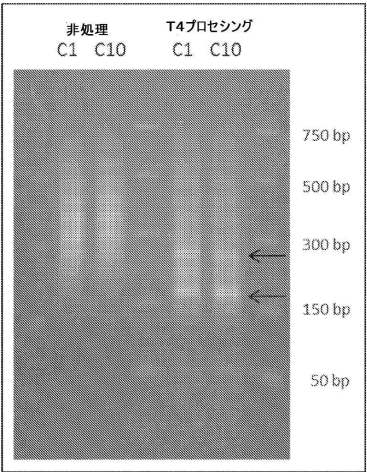


FIG. 12

【図 1 3】

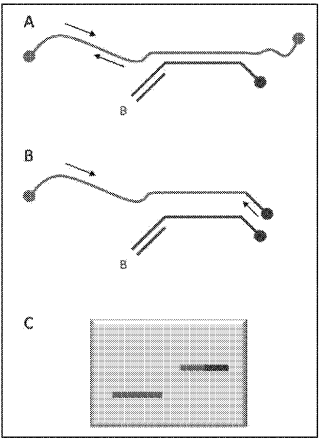


FIG. 13

【図 14】

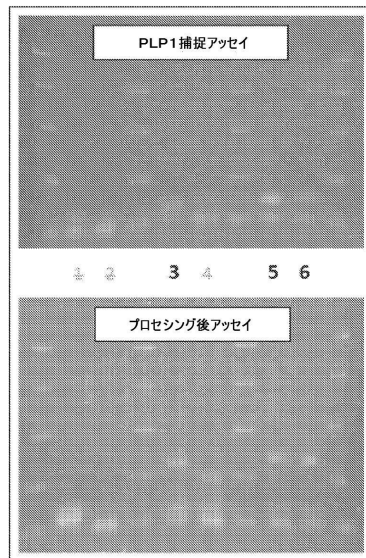


FIG. 14

【図 15】

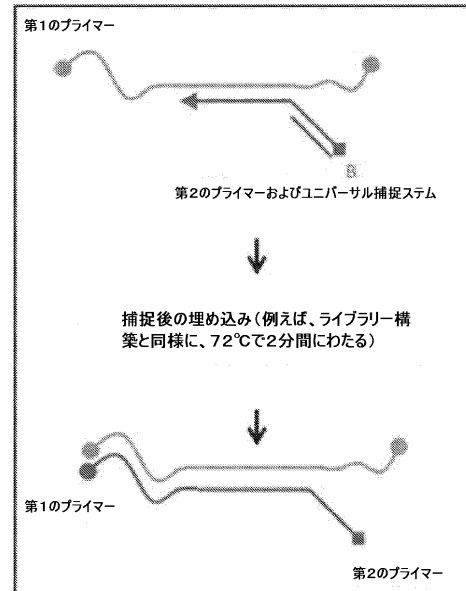


FIG. 15

【図 16】

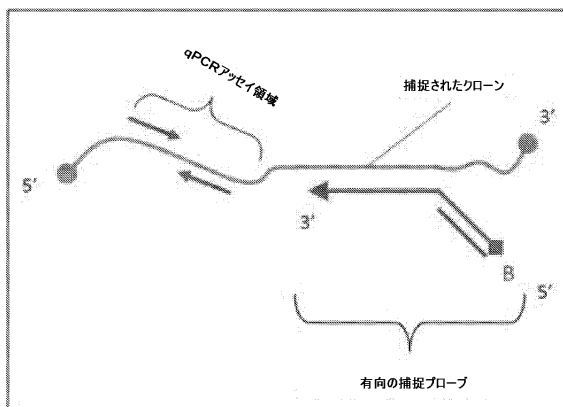


FIG. 16

【図 17】

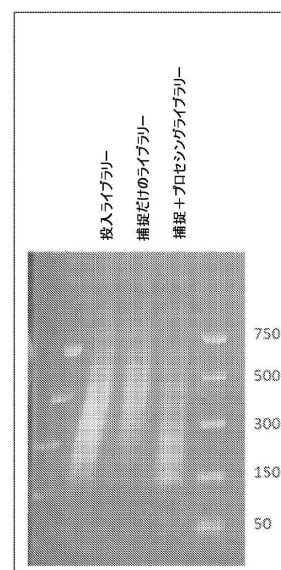


FIG. 17

【図 18】

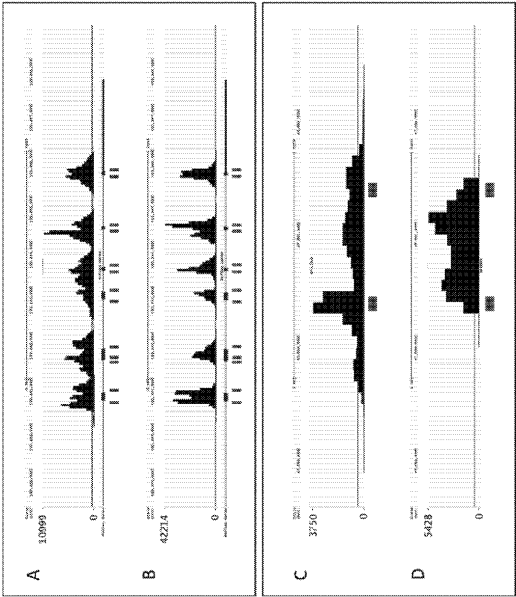


FIG. 18

【図 19】

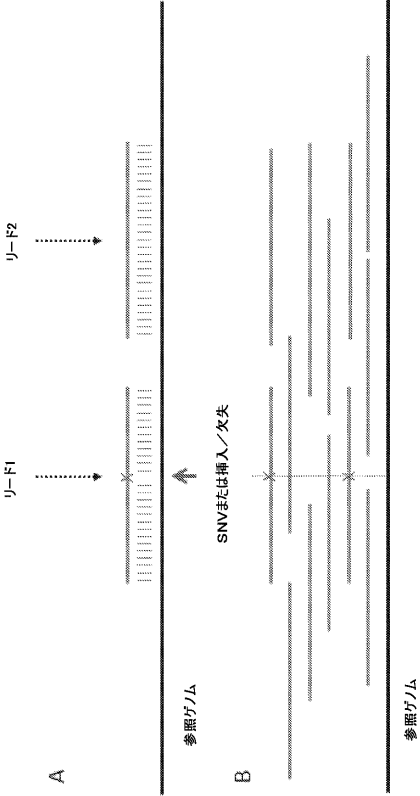


FIG. 19

【図 20】



FIG. 20

【図 21】

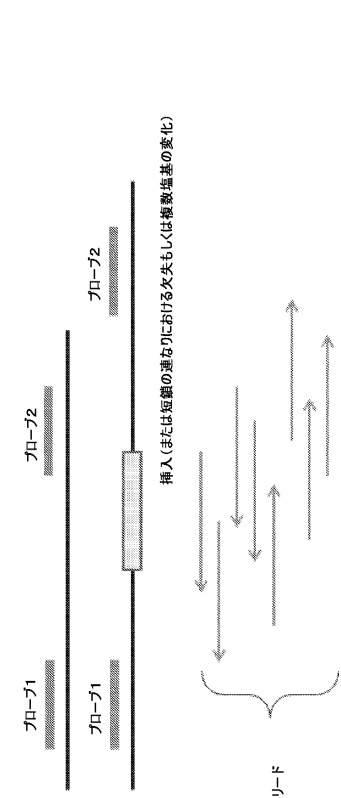


FIG. 21

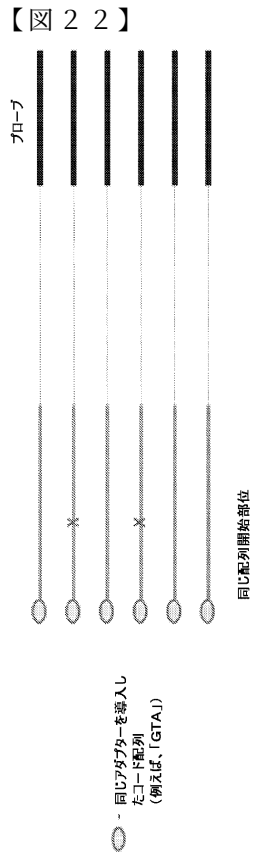
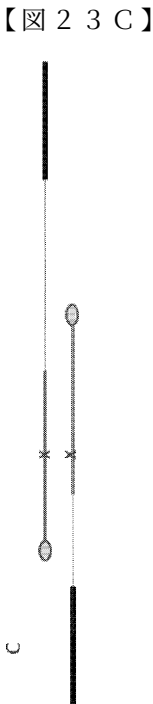
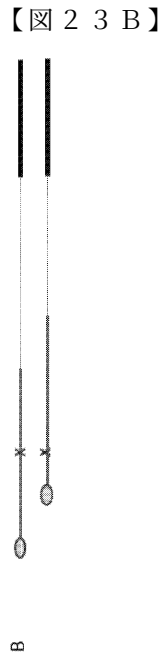


FIG. 22



【図 24】

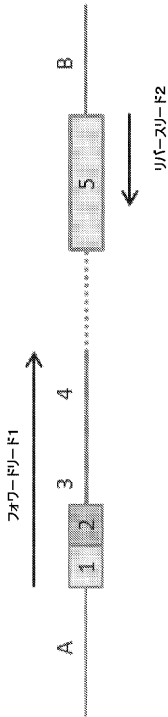


FIG. 24

【図 25】

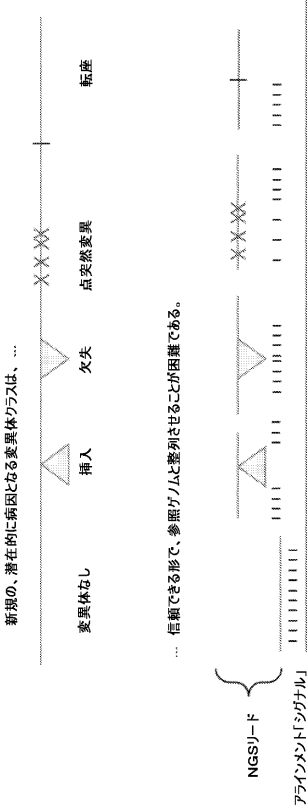


FIG. 25

【図 26】

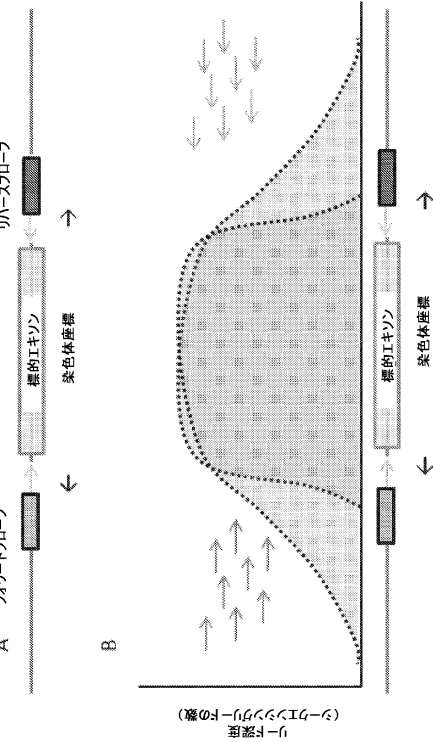
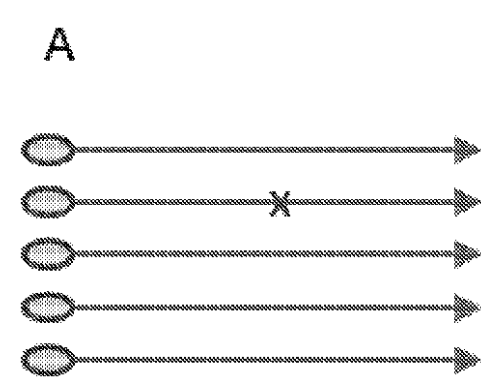
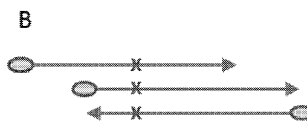


FIG. 26

【図 27 A】



【図 27 B】



【図 28】

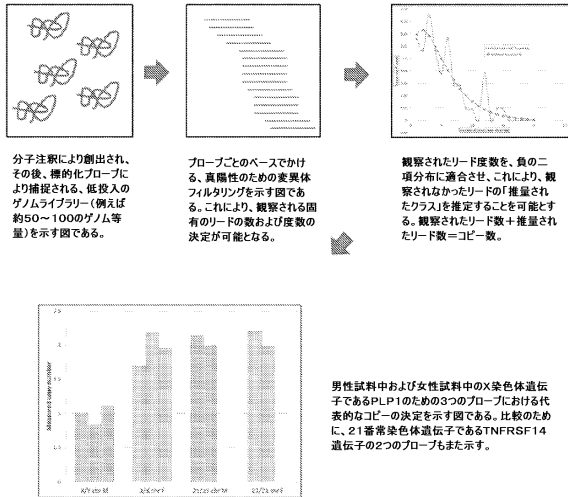


FIG. 28

【図 29】

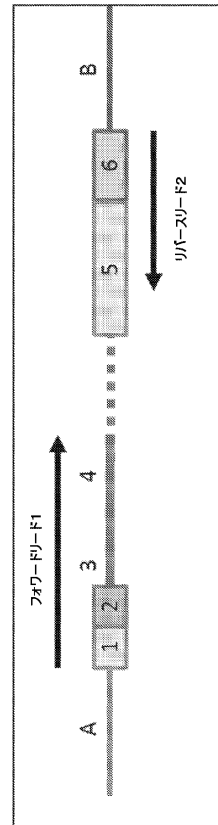


FIG. 29

【図 30】

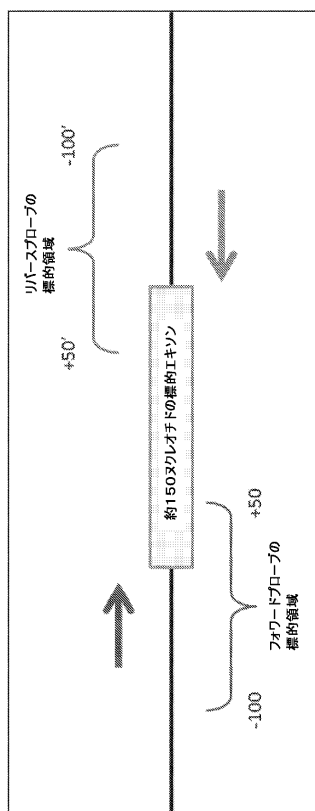


FIG. 30

【図 3 1】

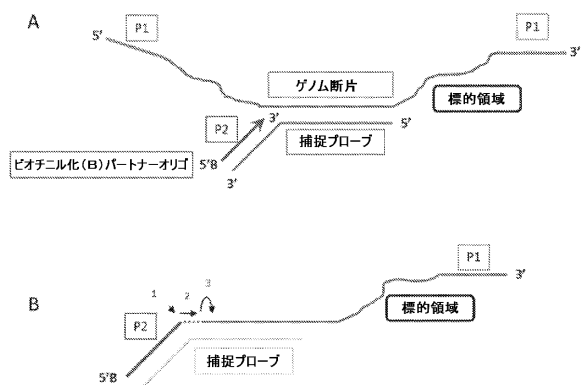


FIG. 31

【図 3 2】

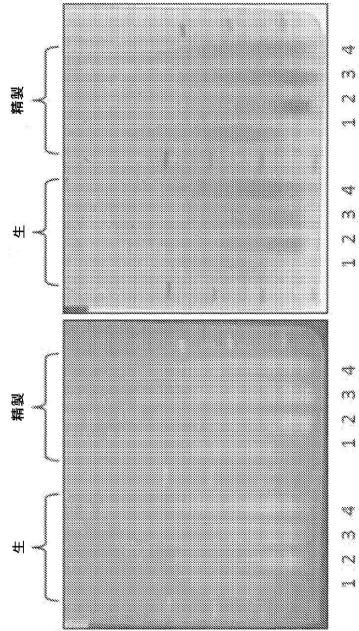


FIG. 32

【図 3 3】

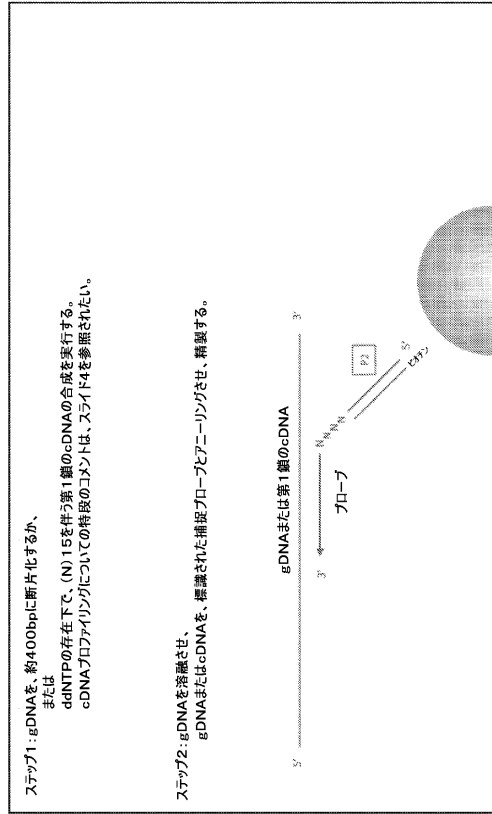


FIG. 33

【図 3 4】

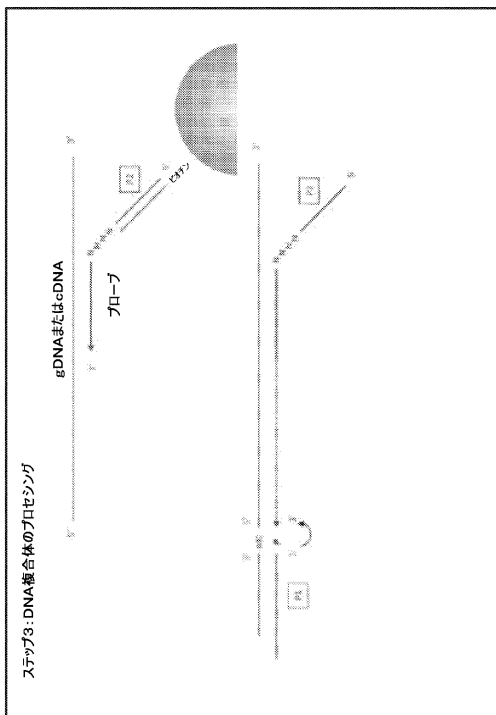


FIG. 34

【図 3 5】

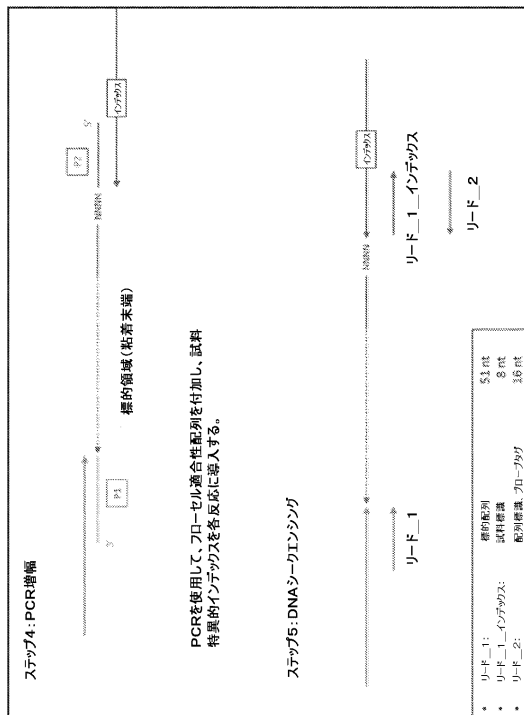


FIG. 35

【図 36】

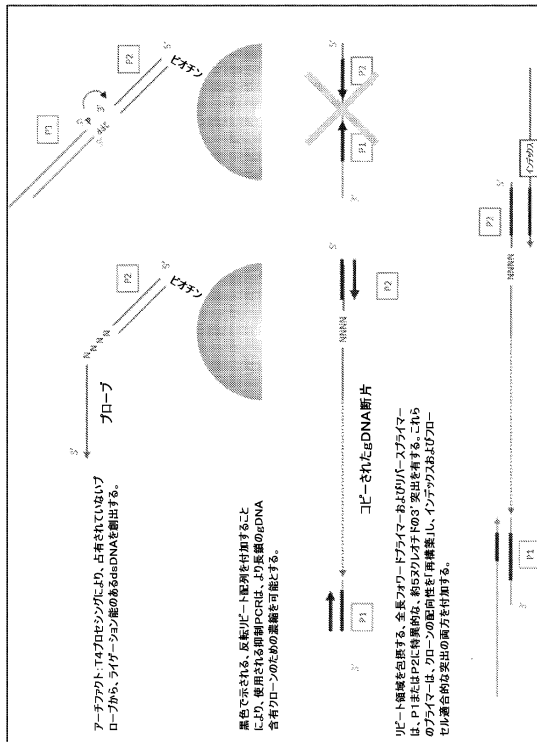


FIG. 36

【図 37】

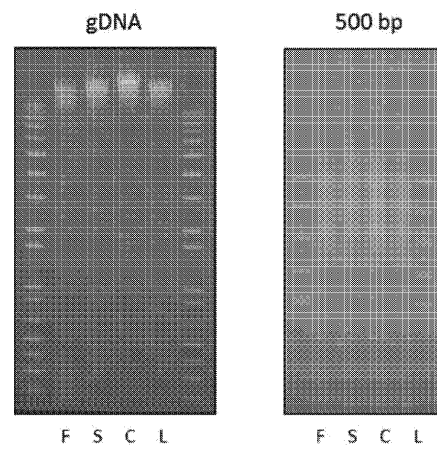


FIG. 37

【図 38】

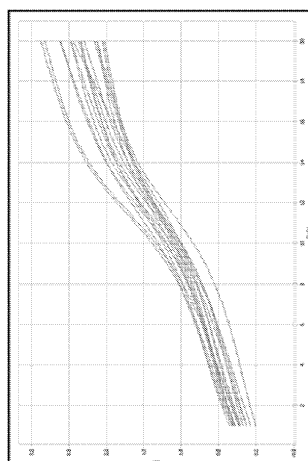


FIG. 38

【図 39】

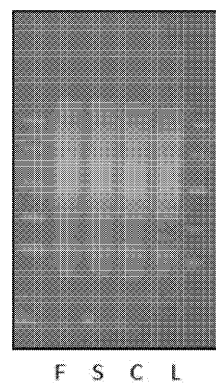


FIG. 39

【図 4 0】

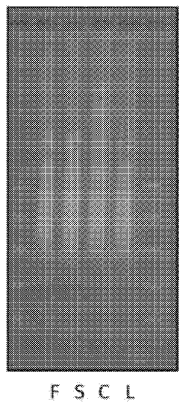


FIG. 40

【図 4 1】

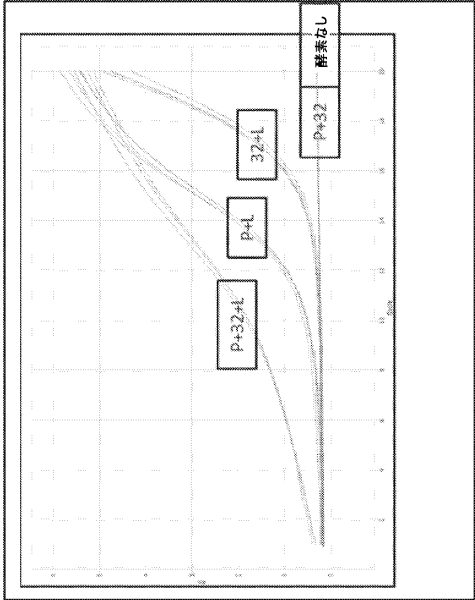


FIG. 41

【図 4 2】

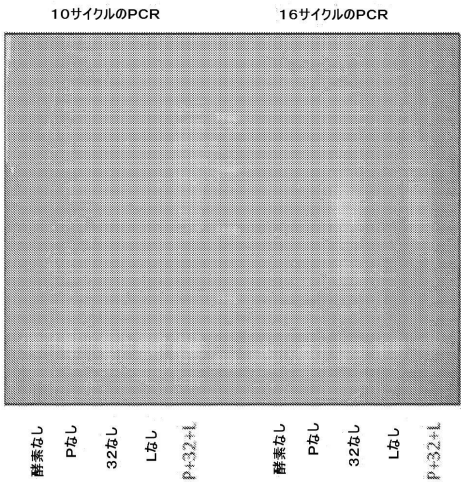


FIG. 42

【図 4 3】

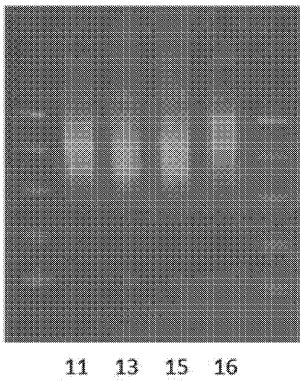


FIG. 43

【図 4 4】

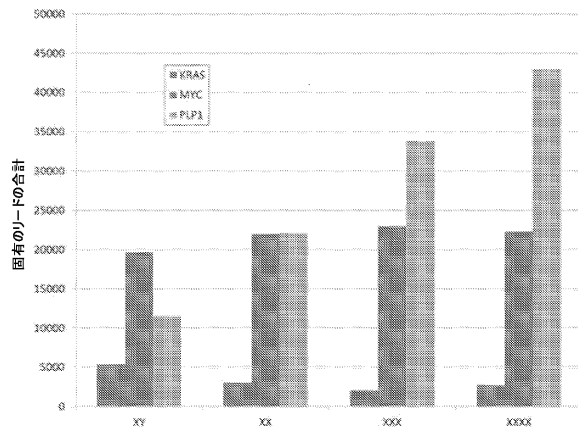


FIG. 44

【図 4 5】

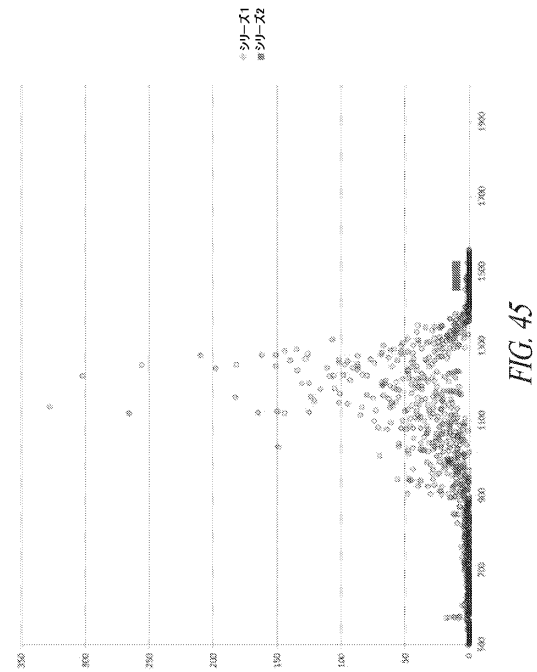


FIG. 45

【図 4 6】

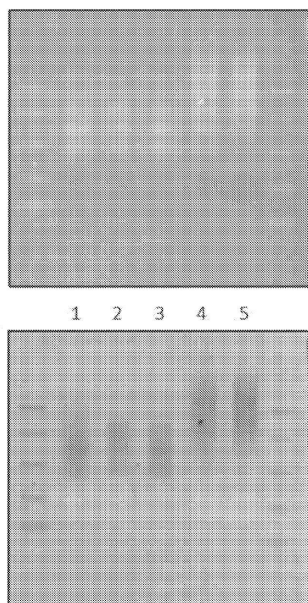


図1: RNA-seq試料についてのDNAゲル

1. 捕捉後RNA-seq試料の捕捉+プロセッシング
2. Pippin調製レーン4
3. Pippin調製レーン5
4. ACA2~20のdTライブラリー
5. ACA2_FLFP+エクソームCAC3_FLRP dTライブラリー

FIG. 46

【図 4 7】

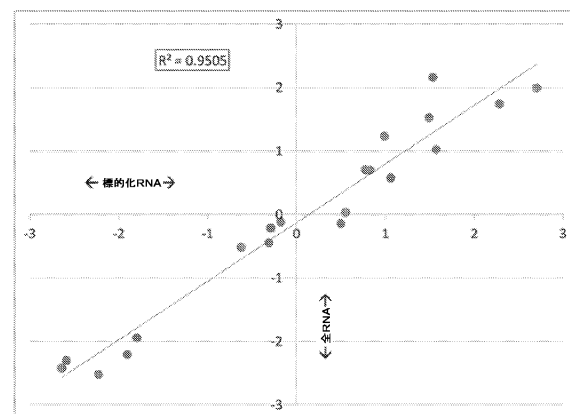


FIG. 47

【図 4 8】

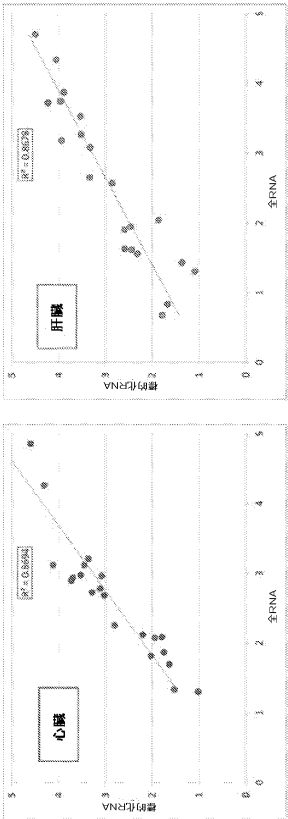


FIG. 48

【配列表】

0006410726000001.app

フロントページの続き

(74)代理人 230113332

弁護士 山本 健策

(72)発明者 レイモンド, クリストファー ケー.

アメリカ合衆国 ワシントン 98115, シアトル, エヌイー 86ティーエイチ ストリート 2626

(72)発明者 アーマー, クリストファー ディー.

アメリカ合衆国 ワシントン 98005, ベルビュー, 120ティーエイチ アベニュー エヌイー 2023

(72)発明者 リム, リー ピー.

アメリカ合衆国 ワシントン 98034, カークランド, エヌイー 120ティーエイチ ストリート 9520, アpartment エー304

審査官 池上 文緒

(56)参考文献 国際公開第2010/129937 (WO, A1)

Nucleic Acids Res. (2007) vol.35, no.15, e97

Proc. Nat. Acad. Sci. U.S.A. (Jan 2012) vol.109, no.4, p.1347-1352, Supporting Information

(58)調査した分野(Int.Cl., DB名)

C12N 15/00-15/90

C12Q 1/6806

C12Q 1/686

PubMed

CAPLUS/MEDLINE/EMBASE/BIOSIS (STN)

WPIDS/WPIX (STN)

JSTPLUS/JMEDPLUS/JST7580 (JDreamIII)