

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2018-509178
(P2018-509178A)

(43) 公表日 平成30年4月5日(2018.4.5)

(51) Int. Cl.	F I	テーマコード (参考)
C 1 2 Q 1/68 (2018.01)	C 1 2 Q 1/68 Z N A A	4 B 0 6 3
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/00 A	

審査請求 未請求 予備審査請求 未請求 (全 75 頁)

(21) 出願番号 特願2017-560901 (P2017-560901)
 (86) (22) 出願日 平成28年2月14日 (2016. 2. 14)
 (85) 翻訳文提出日 平成29年10月6日 (2017. 10. 6)
 (86) 国際出願番号 PCT/US2016/017920
 (87) 国際公開番号 W02016/131030
 (87) 国際公開日 平成28年8月18日 (2016. 8. 18)
 (31) 優先権主張番号 62/116, 302
 (32) 優先日 平成27年2月13日 (2015. 2. 13)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 62/135, 923
 (32) 優先日 平成27年3月20日 (2015. 3. 20)
 (33) 優先権主張国 米国 (US)

(71) 出願人 517283444
 パテール, アビジット アジット
 アメリカ合衆国 06443 コネチカ
 ト州 マディソン アリソン ドライブ
 27
 (74) 代理人 110000176
 一色国際特許業務法人
 (72) 発明者 パテール, アビジット アジット
 アメリカ合衆国 06443 コネチカ
 ト州 マディソン アリソン ドライブ
 27
 Fターム(参考) 4B063 QA01 QQ42 QQ52 QR08 QR32
 QR62 QS25 QX02

最終頁に続く

(54) 【発明の名称】 核酸の高度並列型および正確な測定方法

(57) 【要約】

本文書は、核酸の単純化され、高感度な、且つ正確な定量を可能にする方法および組成物に関する。一部の方法は、複数のサンプルから標的化された複数のリボ核酸の高度な並列測定を可能にする。さらなる方法は、核酸分子の複雑な混合物から低存在量の核酸変異体の高感度測定を可能にする。

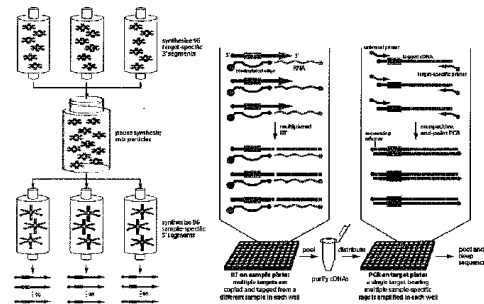


FIG. 1

【特許請求の範囲】

【請求項 1】

逆転写の間に各サンプル内で標的化されるRNA分子のパネルにサンプル特異的のカウンティングタグを配分する段階と；

異なるサンプルからのRNAをコピーするために、一貫した比率で相補的DNAに対してモジュール式オリゴヌクレオチド合成を使用する段階と；

すべてのサンプルから標識cDNAをプールし、精製する段階と；

競合的なエンドポイントPCRにより各cDNA標的を別々に増幅させる段階と；
を含むRNA定量方法。

【発明の詳細な説明】

10

【技術分野】

【0001】

財政支援研究と関連した陳述

本発明は、アメリカ国立衛生研究所により授与された、TR000140およびTR000142の下での政府支援によってなされた。政府は、発明に特定の権利を有する。

【0002】

関連出願に対する相互参照

本出願は、2015年2月13日に出願された仮出願番号62/116,302および2015年3月20日に出願された仮出願番号62/135,923の利益を主張する。

【0003】

20

技術分野

本発明は、溶液中の核酸の同定および定量に関する。

【背景技術】

【0004】

生体医学研究および臨床医学において多くの適用が核酸の正確な検出および定量に依存する。一部の適用は、遺伝子活性または遺伝子発現に関する情報を提供するために、リボ核酸（「RNA」）水準の測定に依存する。他の適用は、点突然変異、挿入、欠失、転位（translocation）、多型またはコピー数多型のようなゲノム異常の存在を示す変異体デオキシリボ核酸（「DNA」）またはRNA配列の測定に依存する。技術的および実用的観点の両方から、色々な課題が核酸の測定に存在する。多くの場合、測定は、大多数のサンプルから行わなければならない。さらに、もし核酸分子の複雑な混合物を含有する制限されたサンプルに、目的とする特定の核酸配列のコピー体が非常に少数で存在すると、低存在量の（low abundance）変異体を確実に確認し、定量することが困難になり得る。

30

【0005】

多様な臨床的および研究標本内の遺伝子発現の分析は、細胞生理学に対する本発明者らの理解を裏付け、疾患に対する本発明者らの取り組みに影響を及ぼす。複雑な生物学的システム内の見識ある有意義な遺伝子発現パターンは、通常、二次元：交差多重リボ核酸および交差多重サンプルの統計学的比較を含む。第1RNA次元の高度並列型分析に対して発達した技術が存在する一方で、処理効率は、第2サンプル次元で制限されたままである。RNA発現のゲノムワイド図が転写体シーケンシング（「RNA-Seq」）または微細配列のような技法を使用して得られる。しかし、これらの接近法が各サンプルの別途の多段階プロセッシングを含むため、大規模サンプルの複合化を容易にするためにデザインされていない。ひいては、RNA-Seqの費用削減がその広範囲な使用を助長する一方で、配列深さ（sequence depth）とサンプル当たりの費用とのトレードオフ（trade-off）が残っていて、それは、希少な転写物の測定感度を制限することがある。

40

【0006】

さらに大きいサンプルセットにわたる標的化されたRNAの評価が、示差的に発現したRNAの部分集合がグローバルプロファイリング方法により確認された後、定量的逆転写

50

ポリメラーゼ連鎖反応(「qRT-PCR」)を使用して行われる。qRT-PCRの精度、感度および広域力動範囲が、そのような転写物の確認および追加の試験のための選択方法にする。しかし、蛍光のリアルタイムモニタリングが別途の反応体積に対して行われる必要があるため、大多数のサンプルに多重遺伝子qRT-PCRアッセイを適用することは、費用がかかり、手間がかかる。たとえ処理量が自動化または微小流体を通じて改善され得るとしても、別途の指数関数的増幅がサンプル間変動性の傾向が依然として残っている。

【0007】

また、核酸分子の複雑な混合物から低存在量の変異体核酸配列を検出し、定量することが困難になり得る。希少な変異配列の検出のための高分析感度を達成することは、提供されたサンプル中のDNAまたはRNAの量が制限される状況では、特に困難になり得る。そのような方法の適用は、癌にかかった個体の血液中の腫瘍由来のDNAまたはRNA分子の少量を検出することである。DNAおよびRNAの断片化された分子は、さまざまな種類の悪性腫瘍を有する患者の癌細胞が死ぬことによって血流中に放出されると知られている。そのような循環腫瘍由来の核酸は、非侵襲性癌のバイオマーカーとして優れた展望を示す。血流において、腫瘍由来の核酸は、腫瘍特異的突然変異の存在を基準として正常な背景DNAまたはRNAと区別され得る。しかし、そのような突然変異核酸のコピー体は、通常、比較的豊富な正常(野生型分子)分子の背後で少量で存在する。しばしば、突然変異腫瘍由来のコピー体は、血漿のトータルDNAまたはRNAを1%未満含み、時には、存在量が0.01%以下と低くてもよい。したがって、非常に高い分析感度を有するアッセイが、そのような低存在量のDNAまたはRNAの検出に關与する。

10

20

【発明の概要】

【0008】

本発明は、広域パネルのマイクロRNA(「miRNA」)、メッセンジャーRNA(「mRNA」)および他の部類のRNAを同時におよび高度の並列方式で多くのサンプルから定量することを可能にする方法および組成物に関する。これらの方法は、従来のデジタルプロファイリング接近法よりずっと少ない配列深さを使用する。一実施態様では、定量タグが逆転写中に配分され、競合増幅およびディープシーケンシング前に、先行サンプルプーリングが許容される。このような接近法は、大規模の遺伝子発現研究をさらに実用的な範囲内でもたらすためにデザインされる。

30

【0009】

また、本発明は、核酸分子の複雑な混合物から低存在量の変異体核酸配列の定量を可能にする方法および組成物に関する。

【図面の簡単な説明】

【0010】

【図1】図1は、開示されたRNAプロファイリング方法を概略的に示す。

【図2】図2は、開示されたRNAプロファイリング方法を使用して複合化されたRNA定量の正確度を支持するデータを提供する。

【図3】図3は、ヒト組織および基準サンプルからのRNAを使用する方法の定量性を検証するためのデータを示す。

40

【図4】図4は、ヒト血液における放射線誘発遺伝子発現変化の高処理測定の結果を示す。

【図5】図5は、色々なmiRNAプロファイリングプラットフォームを比較するためのデータを提供する。

【図6】図6は、ヒト組織のmiRNAの絶対定量の結果を示す。

【図7】図7は、校正用ポリメラーゼの3'から5'エキソヌクレアーゼ活性による末端プロッキング基の消化に抵抗するためにデザインされたRNAseH2活性化可能なプライマーを概略的に示す。

【図8】図8は、系統追跡されたPCRの概略的な説明を提供する。

【図9】図9は、系統追跡されたPCR実験の結果を示す。

50

【図10】図10は、ビーズ特有のバーコードを含有する熱剥離性プライマーがマイクロビーズ上でどのように生成され得るかを示す一例を示す。

【図11】図11は、熱変性により放出され得る一時的に固定されたオリゴヌクレオチドの製造方法を示す。

【図12】図12は、同じ反応体積で共増幅されるPCR生成物に区画固有タグを添加するためのプライマーとして機能できる、クローンタグ付きオリゴヌクレオチドをマイクロコンパートメント内に伝達するためのイン-ソリューション方法を示す。

【図13】図13は、異なる標的がPCR増幅のためにどのように液滴またはマイクロウェル内に無作為に区画化され得るかを示す一例を示す。

【図14】図14は、単一反応区画（例えば、マイクロウェルまたは液滴）の内容物の実例を示す。

【図15】図15AおよびBは、バーコード化プライマーを担持する単一マイクロビーズを含有するマイクロコンパートメント内で実行される系統追跡PCRの二つの例示的シナリオを示す。

【図16】図16AおよびBは、バーコード化プライマーを担持する単一マイクロビーズを含有するマイクロコンパートメント内で実行される系統追跡PCRの二つの追加の実例シナリオを示す。

【図17】図17は、二つ（またはそれ以上）の異なるバーコード化プライマーが所定の区画にある場合、マイクロコンパートメント内で系統追跡PCRの分析がどのように行われるかを示す。

【発明を実施するための形態】

【0011】

本発明は、広域パネルのマイクロRNA（「miRNA」）、メッセンジャーRNA（「mRNA」）および他の部類のRNAを同時におよび高度の並列方式で多くのサンプルから定量することを可能にする方法および組成物に関する。これらの方法は、従来のデジタルプロファイリング接近法よりずっと少ない配列深さを使用する。一実施態様では、定量タグを逆転写反応中に配分され、競合競合増幅およびディープシーケンシング前に、先行サンプルプーリングを可能にする。このような接近法は、大規模の遺伝子発現研究をさらに実用的な範囲内でもたすためにデザインされる。

【0012】

また、本発明は、次世代シーケンシングおよび医学診断に関連した組成物および方法に関する。その方法は、核酸変異体、特に低存在量で利用可能なものまたは豊富な野生型配列により不明瞭にされたものを確認し、定量する段階を含む。また、本発明は、複数のサンプルのうち複数の配列からの特定の配列を確認し定量することに関連した方法に関する。また、本発明は、ポリメラーゼ誤取込（*misincorporation*）エラー、シーケンサーシーケンサーエラーおよびサンプル誤分類エラーから純粋な核酸変異体を検出し区別することに関する。一実施態様では、方法は、サンプル内の標的化された核酸へのバーコードおよび分子系統タグ（MLT）の早期付着を含む。また、方法は、DNA配列を標的化するために高度に特異的な混成化時に遮断されないので、MLTの配分を可能にする一方で、ポリメラーゼ連鎖反応（PCR）中に疑似（*spurious*）増幅生成物を最小化する3'-遮断されたプライマーのネストされたペア（*nested pair*）の使用を含む。方法は、反応の後期サイクルでMLT含有プライマーの関与を回避するために、PCRの最初の数サイクル後にアニーリング温度を上昇させることを含む。また、方法は、配列冗長性を達成するために、クローン性重複ペアエンドシーケンシングを含む。また、方法は、分子系統のトラッキングを可能にするために、多数の小さい反応区画（例えば、オイル中の水性液滴または微小流体装置内の微細な反応体積）にPCR増幅を分けることを含む。追加の方法は、DNA二重螺旋の両鎖で突然変異確認することによって、分析感度を向上させるために、微細な反応体積内に二本鎖DNA断片の両鎖を増幅およびタグ付けを含む。また、方法は、反応体積内に核酸内容物の区画固有のタグ付けを容易にするために、多数の小さい反応体積（例えば、マイクロコンパートメント）に

10

20

30

40

50

クローntag付きオリゴヌクレオチドの多重コピー体の導入を含む。一実施態様では、そのようなクローntag付きオリゴヌクレオチドは、マイクロビーズまたは区画壁のような表面に付着させることなく区画に導入され得る。

【0013】

一実施態様では、方法は、複合化されたPCRにおいて低存在量の鋳型核酸をタグ付けし、増幅させることによって、核酸変異体を測定することを含む。低存在量の鋳型核酸は、母胎循環中の胎児DNA、循環腫瘍DNA (ctDNA)、循環腫瘍RNA、エキソソーム由来RNA、ウイルスRNA、ウイルスDNA、移植された器官からのDNAまたは細菌DNAであり得る。多重PCRは、突然変異を起こしやすいゲノム領域に対する遺伝子特異的プライマーを含むことができる。一実施態様では、突然変異を起こしやすい領域は、癌と関連して変更する遺伝子内にあり得る。

10

【0014】

一実施態様では、プライマーは、バーコードおよび/または分子系統タグ (MLT) を含む。一実施態様では、MLTは、2~10個のヌクレオチドであり得る。他の実施では、MLTは、6、7または8個のヌクレオチドであり得る。一実施態様では、バーコードは、鋳型核酸の起源のサンプルを確認できる。一実施態様では、プライマー伸長反応は、標的化された初期バーコーディングを使用する。標的化された初期バーコーディングでは、異なる核酸領域に特異的な複数の異なるプライマーは、全部同じバーコードを有する。同じバーコードは、特定のサンプルから核酸を確認する。一実施態様では、標的化された初期バーコーディングに対して使用されたプライマーは、特有なバーコード含有オリゴヌクレオチドセグメントを遺伝子特異的プライマーセグメントの均一な混合物とモジュール方式で組み合わせることによって生成される。

20

【0015】

一実施態様では、開示された分析法は、臨床目的に使用され得る。一実施態様では、血液内の核酸変異体は、処理前と後に確認および測定され得る。癌の一実例で、核酸変異体 (例えば、癌関連突然変異) は、治療 (例えば、化学療法、放射線療法、手術、生物学的療法、それらの組合せ) 前に確認および/または測定され得る。次に、治療後に、同じ核酸変異体が確認されるか、測定され得る。治療後に、核酸変異体の量的変化は、治療法が成功であることを示すことができる。

【0016】

用語「分子系統タグ」(「MLT」)の説明

用語「分子系統タグ」(「MLT」)は、合成オリゴヌクレオチド (例えば、プライマー) 内に含有された一連の配列を示すために使用され、鋳型核酸分子のコピー体に対して多様な配列タグを配分するために使用される。MLTの配分は、コピーされた (または増幅された) DNA配列の系統がPCRの最初の数サイクル中に鋳型核酸分子から形成された初期のコピー体を追跡することを可能にする。分子系統タグは、縮重および/またはあらかじめ定められたDNA配列を含有できるが、多様なタグ集団が色々な縮重位置を組み込むことによって、最も簡単に行われる。分子系統タグは、2~14個の縮重塩基位置を有するようにデザインされるが、好ましくは、6~8個の塩基位置を有する。塩基は、連続的である必要はなく、一定の配列によって分離され得る。オリゴヌクレオチド分子集団で生成され得る可能なMLTの数は、一般的に、MLT配列の長さおよび各縮重位置においての可能な塩基の数によって決定される。例えば、もしMLTが8塩基の長さであり、各位置でA、C、GまたはTを有する可能性がほぼ同一である場合、可能な配列の数は、 $4^8 = 65,536$ 個である。MLTは、それぞれのコピーされた鋳型分子に対する完全に特有な配列タグの配分を保証するために、十分な多様性を有する必要はないが、任意の所定のMLT配列を特定の分子に配分する可能性は低くなければならない。可能なMLT配列の数が多きほど、所定の鋳型分子に配分される任意の特定の配列の可能性が低くなる。多くの鋳型分子がコピーされ、タグ付けされるとき、同じMLT配列が一つ以上の鋳型分子に配分され得る。MLT配列は、初期コピーから増幅、プロセッシングおよびシーケンシングまでの分子系統を追跡するために使用される。それらは、純粋な突然変異鋳

30

40

50

型分子に由来する配列からのポリメラーゼ誤取込またはシーケンサーシーケンサーエラーから発生する配列を区別するために使用され得る。また、MLTは、増幅されたPCR生成物が一本のDNA鎖または二つ以上のDNA鎖からコピーされるとき（例えば、鋳型核酸断片の単一コピー体が小さい反応区画内で増幅されるとき）、確認のために使用され得る。また、MLTは、増幅プール中にバーコードのクロスオーバーの結果として誤ったバーコード配分を有する配列を区別するために使用され得る。

【0017】

用語「分子系統タグ付け」は、核酸鋳型分子に対して分子系統タグを配分する過程を示す。MLTは、プライマー内に組み込むことができ、鋳型に対するプライマーの特異的な伸長により標的化された鋳型核酸断片から形成されたコピー体に付着され得る。

10

【0018】

方法

高処理量RNA定量：

qRT-PCRの定量の長所を保有する一方で、シーケンシングベースの読み出しを使用することによって得られるプールサンプルに関連した単一反応の単純性、スケラビリティおよび均一性に影響を及ぼすRNA定量戦略が記述される。図1は、開示されたRNAプロファイリング方法を概略的に示す。实例は、96個のサンプルからの96個のmiRNAの測定を描写する。図1(A)は、モジュールRTプライマーミックスが二段階で合成されることを示す：標的特異的配列を含有する96個の部分的に合成された3'プライマーセグメントがプールされた後、サンプルマーカとして使用される96個の5'タグセグメントの添加のために再分布される。結果的に、生成された96個のプライマーミックスは、それぞれ異なるタグを有する。合成の第2段階が各カラムで3'セグメントの同じ均一混合物で始めるため、最終プライマーミックスは全部、類似な比率の標的特異的配列を共有する。図1(B)は、各サンプルが標的RNA存在量に比例してcDNAに対してサンプル特異的計量タグを配分するために、まず、サンプル特異的モジュールプライマーミックスを使用して多重逆転写(RT)されることを示す。すべてのサンプルからのタグ付きcDNAは、単一体積に組み合わせられ、プライマー伸張配列に相補的であるピオチン標識オリゴヌクレオチドを使用して溶液中のハイブリッドキャプチャにより精製される。多数のサンプルからのタグを有するプールcDNAは、次に、プラトー相(plateau phase)とされる各標的の競合的な単一構成のPCRで共同増幅される。ディープシーケンシングされたアンプリコンからのタグ/標的組合せのカウンティングは、すべてのサンプルにわたるRNAの相対的存在量(relative abundance)を示す。

20

30

【0019】

特定の実施態様で、方法は、マイクロRNA(miRNA)、メッセンジャーRNA(mRNA)、長鎖ノンコーディングRNA(lncRNA)または他のRNA部類を定量できる。方法は、別途のエンドポイントPCRが低および高存在量(low-and high-abundance)のRNA種の総コピー体を略平等にするため、他の標的化されたまたは全体転写体のシーケンシング方法より塩基当たり平均配列深さが非常に少ないことを必要とする。したがって、希少な転写物は、豊富なサンプルをオーバーサンプリングする必要なしに、適切にサンプリングされ、広い動力範囲を得る一方で、配列の経済性を最大化できる。下記の表1から分かるように、Ion Torrentパーソナルベンチ-トップシーケンサーの最低出力モード(1百万より少ない読み込み)が96個のサンプルからの96個のRNAを迅速で且つ低価で定量するために使用され得、9,216個の個別的なqRT-PCR分析と同等なデータを提供する。さらに多数のサンプルセットの分析は、ひいては、反応チューブの数がRNAの数および評価されるサンプルの数の合計-生成物ではない-として増加されるので、qRT-PECと比較してこの接近法の単純性を強調する。

40

【0020】

[表1]

50

試験されたサンプルに対する配列深さ

RNA供給源	標的類型	使用されたシーケンシングチップ	濾過された読み込み	試験されたサンプルx (複製物)	標的化されたRNA	複製物当たり標的-サンプルビン当たり平均読み込み
合成 (図2、aおよびc)	miRNA	318チップ x 2	4.53M+2.42M	96 x (2)	96	377
合成 (図2、bおよびd)	miRNA	314チップ x 2	411K+419K	96 x (2)	96	45
正常ヒト組織 (図3、aおよびb)	miRNA	314チップ x 1	446K	20 x (3)	96	77
正常ヒト組織+合成標準サンプル (図6)	miRNA	314チップ x 1	387K	21 x (3)	96	64
MAQC参照サンプル (図3、cおよびd)	mRNA	314チップ x 0.5 (他のサンプルと共有される)	~217K	4 x (4)	30 標的+4 リファレンス領域 (2個のリファレンス遺伝子の)	399
放射線照射されたヒト血液 (図4)	mRNA	314チップ x 2	318K+326K	108 x (2)	23 標的+2 リファレンス遺伝子	119

10

20

30

40

50

【0021】

一実施態様では、方法は、先行サンプル並列化を可能にし、シーケンシングの直前に、サンプルを組み合わせる接近法より優れた色々な利点をもたらす。作業の流れは、非常に単純化され、微小流体装置または自動化を必要としない。すべてのRT後の段階においてのプールプロセッシングは、サンプルにわたって定量的可変性を減少させると予想される。各標的のPCRを完了することによって、配列深さは、大部分豊富な転写物により消費されるのではなく、すべての標的にわたって均一に分布する。したがって、配列深さに関連したサンプル当たりの費用は、最小化される一方で、サンプル深さは、低存在量の転写物のうちサンプル内での差異を正確に定量するために保存される。

【0022】

方法は、初期サンプルプーリングを利用し、非常に少ない配列深さを使用し、短いmi

R N Aを標的化するため、従来の標的シーケンシング接近法と異なっている。また、それは、q R T - P C Rまたは大部分の混成化に基づく接近法より配列変異体（特にさらに長いm R N A標的）の識別に一層適している。方法は、次世代シーケンサが他の微小流体または直接分子カウティング技術により使用された特殊化された多くの機器より制度的な核施設でさらに通常的に入手可能であるため、大部分の研究所にて広範囲にアクセス可能でなければならない。また、方法は、異なるシーケンシングプラットフォームにも容易に適用可能であり、多様な機能的R N A部類を分析するために拡大され得、最小限の計算インフラおよび専門知識を使用する。

【 0 0 2 3 】

低存在量の核酸変異体の定量：

核酸配列変異体を確認し定量する方法および組成物が開示される。D N AまたはR N Aの複合混合物から低存在量の配列変異体を確認し定量する方法が開示される。方法は、多様な種類の癌にかかった患者の循環系で発見され得る少量の腫瘍由来D N Aを測定できる。

10

【 0 0 2 4 】

希少な変異型D N A配列の評価は、多くの生物学および医学分野において重要である。少量の胎児D N Aが妊婦の循環系で発見され得る。一実施は、胎児の疾患関連遺伝的特徴または性別を評価するために使用され得る希少な胎児D N Aを分析することを含む。受容体により拒否反応を示す器官は、少量のD N Aを血液に放出させることができ、このドナー由来のD N Aは、ドナーと受容体間の遺伝的差異を基に区別され得る。一実施は、臓器拒否反応および治療の有効性に関する情報を提供するために、ドナー由来のD N Aを測定することを含む。他の実施で、核酸は、患者サンプルの感染病原体（例えば、バクテリア、ウイルス、真菌、寄生虫など）から検出され得る。病原体由来核酸の変異に関する遺伝情報は、感染の特徴をよりよく特定し、治療決定に役立つ。例えば、患者に感染するバクテリアゲノム中の抗生物質耐性遺伝子の検出は、抗生物質治療を指示することができる。

20

【 0 0 2 5 】

低存在量の突然変異の検出および測定は、腫瘍学分野において多くの場合、重要な用途に適用される。腫瘍は、体細胞性突然変異を取得すると知られているが、それらの一部は、癌細胞の制御されない増殖を促進する。そのような突然変異の確認および定量は、腫瘍学分野において核心的な診断目標になった。同伴診断は、癌の突然変異の原因を特定し、次に、その特定の突然変異に対する効果的な治療法を施すに際して、重要手段になった。ひいては、一部の腫瘍は、標的化された治療法に対する耐性を付与する新しい突然変異を取得する。したがって、腫瘍の突然変異状態の正確な測定は、所定の患者に対する特定治療法の適切性を決定する上で重要な要因であり得る。しかし、腫瘍特異的体細胞性突然変異を検出することは、特に生検または切除標本から得られた腫瘍組織が大きい背景においての間質細胞で少数の腫瘍細胞を有する場合、難しい。腫瘍由来の突然変異型D N Aは、血液、痰、尿、便、胸水または他の生物学的サンプルにおいてごく微量に存在するときに測定することがさらに困難であり得る。

30

【 0 0 2 6 】

腫瘍由来のD N Aは、多様な種類の悪性腫瘍がある患者の死んでいく癌細胞から血流中に放出される。循環腫瘍D N A（c t D N A）の検出は、限定するものではないが、悪性腫瘍の存在の検出、予後の通知、治療効果の評価、腫瘍突然変異等級の変化追跡および疾患再発または進行に対するモニタリングなどを含む色々な用途を有する。特有の体細胞性突然変異が血漿で正常背景D N Aと腫瘍由来のD N Aとを区別するために使用されるため、そのような循環腫瘍由来のD N Aは、従来の血清蛋白質マーカーの臨床用途を補完できる臨床用途を有する高度に特異的な癌生体マーカーの新しい部類を示す。一実施態様では、方法は、腫瘍特異的、体細胞性突然変異の存在に対してc t D N Aをスクリーニングすることを含む。そのような実施で、偽陽性結果は、健康な個体の血漿D N Aの癌関連突然変異を発見する可能性が非常に少ないため、非常に珍しい。開示された方法は、高い分析感度および特異性で癌細胞から血流中に入る希少な突然変異型D N A分子を測定する方法

40

50

を含む。非常に高い検出感度を達成することは、初期（およびさらに硬化しやすい）段階で小さい腫瘍の検出するためには特に重要である。

【0027】

体細胞性突然変異が多様な癌関連遺伝子内で多数の可能な位置で発生し得るため、c t DNAを分析するための臨床的に有用な試験は、多くの遺伝子で同時に、好ましくは、多くのサンプルから同時に突然変異を評価できる必要がある。複数のサンプルから複数の突然変異を起こしやすい領域の分析は、大量並列シーケンシング技術を使用して得られる大量の配列データのさらに効果的な使用を可能にする。一実施態様では、所定のサンプルから発生する分子をバーコードまたはインデックスとしても知られているサンプル特異的DNA配列タグで標識することは、一つ以上のサンプルの同時分析を容易にする。異なるサンプル由来の分子を標識するために、区別されるバーコード配列を使用することによって、分子を組み合わせ、混合物に対して大量並列シーケンシングを行うことが可能である。結果的な配列は、次にどの配列がどんなサンプルに由来したかを測定するために、バーコード同一性を基準として分類され得る。誤分類の機会を最小化するために、バーコードは、任意の所定のバーコードが最低二つの位置で異なる塩基を有することによって、セットのすべての他のバーコードと確実に区別され得るようにデザインされる。

10

【0028】

大量並列シーケンシングのためにサンプルを製造するために現在使用される大部分のプロトコルで、バーコードは、色々な段階のサンプルプロセッシング（例えば、精製、増幅、エンドリペアなど）後に付着される。バーコードは、バーコーディングされたシーケンシングアダプタの結紮により、または、関心ある核酸のコピー体を形成するために使用されるプライマー内へのバーコードの統合により付着され得る。二つの接近法は、いずれも、通常、バーコードが付着され得る前に、各サンプル由来の核酸に対して別々に行われる色々なプロセッシング工程を使用する。バーコードが付着された後にのみ、サンプルが混合され得る。

20

【0029】

一実施態様では、バーコードは、サンプルプロセッシングの非常に初期の段階で標的化された分子に配分される。標的化された初期バーコード付着は、多数のサンプルのシーケンシングがバッチで行うこと可能にするだけでなく、多くのプロセッシング段階が組合せられた反応体積で行われることを可能にする。ひとまずバーコードがサンプル特異的な方式で核酸分子に付着されると、分子を混合することができ、それ以降の全ての段階が単一チューブで行うことができる。もし大多数のサンプルが分析されると、標的化された初期バーコーディングは、作業の流れを大きく単純化することができる。すべての分子が単一チューブで同じ条件の下に処理され得るため、分子は、均一な実験条件を経験し、サンプル間の変化も最小化される。一実施態様では、異なるサンプルからの核酸のタグ付けは、一定の割合で達成され得、次に、サンプルにわたって核酸濃度の定量的比較を可能にするために使用される。したがって、初期バーコーディングは、多くのサンプルにわたって、変異体だけでなく、多様な標的化された核酸の総量を定量するために使用され得る。

30

【0030】

一実施態様では、サンプル特異的バーコードおよび一定比率の遺伝子特異的セグメントの組合せを含有するプライマーの明確に定義された混合物が製造される。そのようなプライマーは、標的化された初期バーコーディングおよびその後のバッチ式サンプルプロセッシングに使用され得る。これらのプライマーは、また、異なるサンプルでDNAまたはRNAの定量に使用され得る。一実施態様では、そのようなプライマーは、単純化されたおよび均一な方式で多数のサンプルから多数の突然変異を起こしやすいゲノム標的領域の並列プロセッシングおよび分析を許容する。

40

【0031】

現在の開示された方法は、突然変異型DNAの有無を単純に測定するのではなく、それを正確に定量する方法を含む。一実施態様では、突然変異型DNAの量は、腫瘍量および予後に関する情報を提供する。現在の開示された方法は、アポトーシス性死滅を進行す

50

る細胞からの放出時の分解によるものだけでなく、血液由来のヌクレアーゼによる分解により高度に断片化されるDNAを分析できる。体細胞性突然変異が多様な癌関連遺伝子内で多数の可能な位置で発生し得るため、一実施は、所定のサンプルから同時に多くの遺伝子の突然変異を評価できる。現在の開示された方法は、どんな突然変異が患者の腫瘍に存在するかを事前に知らなくてもc t DNAの突然変異を発見できる。一実施は、腫瘍特異的体細胞性突然変異を発達させやすいゲノムDNAの多重領域を評価することによって、多くの異なるタイプの癌に対してスクリーニングすることができる。一実施は、サンプル間の変化を最小化するために、同じ反応チューブ内で一緒に組合せられた複数のサンプルを含む。

【0032】

たとえ現在開示された方法が癌を有する患者の血漿または血清で正常（野生型）無細胞DNAをベースに少量の突然変異循環腫瘍DNA（c t DNA）の測定に最適化されたが、方法は、多様な供給源からの核酸変異体の分析のために、より広範囲に適用され得ることが認知される。そのような供給源の実例としては、限定するものではないが、リンパ節、腫瘍辺縁、胸水、尿、便、血清、骨髄、末梢白血球、口腔粘膜検体採取（c h e e k s w a b）、循環腫瘍細胞、脳脊髄液、腹水、羊水、嚢胞液、冷凍腫瘍標本、およびホルマリン固定されているおよびパラフィン包埋されている腫瘍標本を含む。

【0033】

方法

高処理量RNA定量：

単純化されたおよび正確なRNA測定のための先行サンプル並列化：

一実施態様では、高処理量RNA定量方法は、次の基本的な段階を通じて行われることができる。

【0034】

一実施態様では、作業の流れの早期並列化を可能にするために、サンプル特異的カウンティングタグ（バーコード）が逆転写（RT）中に各サンプル内で標的化されるRNA分子のパネルに配分される。一実施態様では、遺伝子特異的プライマーは、逆転写のために関心あるRNAを標的化するのに使用される。一実施態様では、関心あるRNAは、マイクロRNA、メッセンジャーRNA、長鎖ノンコーディングRNA（l n c RNA）または任意の他のRNA類型であり得る。一実施態様では、遺伝子特異的プライマーは、サンプル特異的バーコードで標識される。一実施態様では、サンプル特異的バーコードは、逆転写中に相補的DNA（c DNA）に配分される。一実施態様では、c DNAに配分された所定のサンプル特異的タグの量は、サンプル中の該当するRNAの存在量に比例する。一実施態様では、逆転写に対して使用されたプライマーの遺伝子特異的ハイブリダイゼーション領域は、6ヌクレオチドほど短くてもよく、40ヌクレオチドほど長くてもよい。特定の実施で、遺伝子特異的混成化配列は、短いマイクロRNA標的を逆転写するために使用されるとき、6ヌクレオチド長である。一実施態様では、6-塩基対RNA/DNA相互作用の特異性および安定性を増大させるために、マイクロRNAに結合しないプライマー塩基は、共通プライマー配列に相補的であるビオチン化オリゴヌクレオチドをアニーリングすることによって、マスキングされ得る。特定の実施態様では、遺伝子特異的混成化配列は、さらに長いメッセンジャーRNAまたはl n c RNA標的を逆転写するために使用されるとき、15~25ヌクレオチド長である。一実施態様では、多重遺伝子特異的プライマーが多重RNA配列の標的化された逆転写（RT）を行うために同じ反応体積で使用され得る。一実施態様では、所定のサンプルからのRNAを逆転写させるために使用されたすべてのプライマーは、同じサンプル特異的バーコード（タグ）を含有する。一実施態様では、複数のサンプルは、別々の反応体積で同時に逆転写され得る。一実施態様では、逆転写が完了するとき、すべてのサンプルからのすべてのタグ付きc DNAコピー体は、単一体積で組み合わせられ、精製され得る。一実施態様では、プールc DNAは、表面に固定されたストレプトアビジンまたはその類似体を使用してビオチン捕捉により精製され得る。

10

20

30

40

50

【 0 0 3 5 】

一実施態様では、異なるサンプル由来のRNAが一定の割合で相補的DNA (cDNA) に対してコピーされることを保証するために、モジュールオリゴヌクレオチド合成スキームが使用される。一実施態様では、j サンプルから逆転写中に、i RNAの複合化標的の標識を可能にするために、サンプル特異的タグに結合された標的的特異的配列のi x j 組合せを有するRTプライマーを生成する必要がある。一実施態様では、定量の一貫性を保証するために、同じ比率のすべての標的的特異的配列を有する特有にタグ付きプライマーミックスを使用して異なるサンプルを逆転写させることが重要であった。個別的に形成されたプライマーを単に数千個混合することは、非実用的であり、不正確な比率を生成するため、2段階モジュール式オリゴヌクレオチド合成戦略が使用された。一実施態様では、オリゴヌクレオチド合成は、色々な異なる標的的特異的プライマー配列を形成した後、停止することができる。一実施態様では、合成器が停止することができ、部分的に合成されたオリゴヌクレオチドを含有した粒子が混合され、色々な新しい合成カラムに分配され得る。一実施態様では、特有のサンプル特異的タグおよびユニバーサルPCRプライマー結合部位を含む各カラムに配列を加えることにより合成を再開することができる。一実施態様では、色々なプライマーミックスが生成されるが、それぞれ5'セグメントに特有なサンプル特異的タグおよび3'セグメントに色々な標的的特異的プライマー配列の均一な組成を有する。

10

【 0 0 3 6 】

一実施態様では、すべてのサンプル由来のタグ付きcDNAがプールされ、精製される。一実施態様では、cDNAプールが、別途の、シングルプレックス、エンドポイントPCR (安定期に対して取られる) によって各標的の増幅のために別々の反応体積に分配される。所定のcDNA種に関連したすべてのサンプル特異的タグが単一体積で競合競合的に増幅されるため、RNA存在量をコード化するタグ比が保存される。一実施態様では、PCRプライマーの5'末端でシーケンシングアダプタの統合は、結果的に生成されるアンプリコンがプールされ、ゲル精製され、追加のライブラリー製造段階なしに大量並列シーケンシングに対する鋳型として直接使用されることを可能にする。

20

【 0 0 3 7 】

一実施態様では、多様なサンプル中のRNAの相対量は、PCR生成物の大量並列シーケンシングにより得られた各cDNA配列に関連したサンプル特異的タグを計数することによって推論され得る。

30

【 0 0 3 8 】

モジュール式プライマーミックスの効用および組成：

RNAプロファイリング方法に対して、標的化された核酸分子 (特に、RNA鋳型からコピーされたcDNA) にサンプル特異的タグを配分するために、モジュール式プライマーミックスが使用された。しかし、そのようなモジュール式プライマーミックスは、広範囲な用途を有することができる。それらは、より一般的には、標的化された核酸分子 (RNAまたはDNA) の分布または頻度を確認、範疇化、分類、区分、カウンティングまたは測定することを補助できるタグを配分するために使用され得る。モジュール式プライマーミックスは、3'セグメントに多数の異なる標的的特異的配列を有し、5'セグメント特有のタグ配列を有するプライマーの混合物である。しばしば、いくつかのモジュール式プライマーミックスは、それぞれのプライマーミックスが別個のタグを有し、すべてのミックスが標的的特異的配列の同一組成を有するようにセットとして形成される。標的およびタグの数が大きくなると、プライマーを個別的に合成し、次に、それらを混合することが実用的でないことがある。

40

【 0 0 3 9 】

モジュール式プライマーミックス中に統合されるタグ (また、バーコードまたは標識とも言及される) は、任意的な配列で構成され得るが、通常、確実に区別され得る所定の配列を含む。例えば、RNAプロファイリング方法において、各タグは、少なくとも二つのヌクレオチド位置がセットのすべての他のタグと異なるようにデザインされ、シーケン

50

シングエラーはタグの誤分類をめったに誘導しない。タグは、塩基の単一の、連続的ストレッチ内に含まれる必要はない。特定の実施では、タグ配列を含むヌクレオチド位置は、モジュール式プライマーミックスの5'セグメントの非連続領域にわたって分布することができる。タグは、また、無作為または縮重位置を含有する（縮重位置は、例えば四個のヌクレオチドA、T、CおよびGがオリゴヌクレオチド合成中に同等な可能性で統合される所である）。しかし、モジュール式プライマーミックス内のタグは、事前規定された（縮重でない）配列を有する少なくとも一部の位置を含有しなければならない。

【0040】

モジュール式プライマーミックス内で、タグは、サンプル特異的である必要はない。例えば、タグは、サンプル、分子、位置または区画に配分され得る。また、タグは、サンプルセット、分子セット、位置セットまたは区画セットに配分され得る。用途によって、タグの配分は、無作為であるか（例えば、任意のタグが任意のサンプル、分子、位置または区画に無作為に配分される）、または事前規定されたものであってもよい（例えば、当業者は、特定のサンプル、分子、位置または区画に特定のタグを配分するように決定できる）。タグの特有な配分必ずしも必要なものではない。一部の用途に対して、各サンプル、分子、位置または区画には、特有のタグを配分しなければならない。一部の他の用途に対して所定のタグが一つ以上のサンプル、分子、位置または区画に配分されることを可能にする。

10

【0041】

一部の用途で、一つ以上のモジュール式プライマーミックスが標的または標的セットを標識するために使用され得る。例えば、モジュール式プライマーミックスは、PCR増幅反応の順方向および逆方向プライマーセットの両方として使用され、標的に対する二つの別個のタグの配分を許容できる。標識の大きい多様性は、タグ付き順方向および逆方向プライマーミックスの多様な組合せを使用することによって達成され得る。

20

【0042】

複合混合物から低存在量の突然変異型DNAの定量
鑄型DNAの分離：

多様な臨床または実験標本からDNAまたはRNAの精製または分離方法が開示される。多くのキットおよび試薬が核酸精製を容易にするために市販されている。分析しようとするサンプルのタイプによって、適切な核酸分離技法が選択され得る。後続の酵素反応段階（例えば、重合反応）を抑制する物質は、除去されるか、精製されたDNAまたはRNAサンプルで非抑制性濃度に減少しなければならない。核酸の収率は、可能な限り、最大化されなければならない。精製中にDNAが消失されることは、消失されたDNAが希少な変異型DNAを含むこともできるため、不利である。血漿からDNAを分離するとき、約1 ng ~ 100 ngの無細胞DNAは、1 mLの血漿から精製され得、それは、約350 ~ 35,000ゲノムコピー体に該当する。DNA収率は、特に癌のような進行中の疾患過程を有する患者にたいして劇的に変化できる。

30

【0043】

一実施態様では、DNAは、また、限定するものではないが、次を含む他のサンプル類型から分析され得る：胸水、尿、便、血清、骨髄、末梢白血球、循環腫瘍細胞、脳脊髄液、腹水、羊水、嚢胞液、リンパ節、冷凍腫瘍標本、およびホルマリン固定されているおよびパラフィン包埋されている腫瘍標本。

40

【0044】

系統追跡PCR

一実施態様では、標的化された鑄型DNA分子が遺伝子特異的プライマーを使用して「分子系統タグ」（MLT）で標識されることを可能にし、これらのタグ付きコピー体が次にユニバーサルプライマーを使用して追加にコピーされる（増幅される）ことを可能にする方法が提供される。一実施態様では、このような反応は、試薬の伝達なしに単一反応体積で行われ、それは、過程の単純性という顕著な長所を提供する。図8に示されているように、MLT配列を含有する色々な遺伝子特異的プライマーが、関心ある複数の標的化さ

50

れたゲノム領域（例えば、癌の体細胞性突然変異を起こしやすい領域）を同時にコピーし標識するために使用される。遺伝子特異的プライマーは、ユニバーサルプライマーの溶融温度より低い溶融温度（標的遺伝子配列に対する混成化のために）を有する。標的化された鋳型DNA断片のコピーおよびMLT配列の配分は、PCRの最初の数（2～4）サイクルの間にさらに低いアニーリング温度を使用することによって促進される。後続のPCRサイクルで、アニーリング温度は、MLT含有遺伝子特異的プライマーの反応への追加の関与を阻止するために上昇される。順方向の遺伝子特異的プライマーの5'部分は、順方向のユニバーサルプライマー配列の3'部分と同一の共通配列を含有する。逆方向の遺伝子特異的プライマーの5'部分は、ユニバーサル逆方向プライマー配列の3'部分と同一の第2（異なる）共通配列を含有する。

10

【0045】

ユニバーサルプライマー配列は、遺伝子特異的プライマーより高い溶融温度を有するようにデザインされる。一実施態様では、ユニバーサルプライマーは、混成化安定性を増加させるために、一部の位置でヌクレオチド類似体、例えば架橋型人工核酸（locked nucleic acid, LNA）残基で変形され得る。あるいは、ユニバーサルプライマーは、溶融温度を増加させるために、単に、さらに長い配列および/またはさらに大きいG/C含量を有することができる。PCRの後期サイクル（最初の2～4サイクル後）の間に、熱サイクリングのアニーリング温度は、ユニバーサルプライマーが効果的に混成化できる水準に上昇し得るが、遺伝子特異的プライマーは、そうできない。したがって、最初の数PCRサイクルに生成されたMLT標識されたコピー体は、増幅され始め、

20

【0046】

一実施態様では、遺伝子特異的プライマーは、相対的に低い濃度（それぞれ～10から～50 nM）でPCRカクテルに存在することができるが、バーコーディングされたユニバーサルプライマーは、さらに高い濃度（それぞれ～200から～500 nM）で存在する。一実施態様では、バーコードおよびアダプタ配列を欠く短いユニバーサルプライマーは、また、相対的に高い濃度（それぞれ～100 nMから～500 nM）でカクテルに添加され得る。低濃度の遺伝子特異的プライマーの混成化および延長のための十分な時間を許容するために、さらに長いアニーリング時間が最初の数PCRサイクルに使用され得、この際、選択的にアニーリング温度にゆっくり冷却される。後続のPCRサイクルの間に、ユニバーサルプライマーのさらに高い濃度のために、さらに速いアニーリング時間が使用され得る。

30

【0047】

標的外混成化および遺伝子特異的プライマーの延長の最小化は、この方法の成功に重要である。同じ反応カクテル内においてのユニバーサルプライマーの存在のため、遺伝子特異的プライマーの相互間の混成化および延長（すなわちプライマー二量体の形成）を最小化することが特に重要である。遺伝子特異的プライマーのうち極めて少量の二量体の形成でも、反応には致命的であることがあるが、そのような二量体がユニバーサルプライマーにより指数関数的にコピーされ、増幅され得るためである。もし二量体の増幅が反応より優位に立つと、標的化された遺伝子領域は、十分に増幅されなくてもよい。遺伝子特異的プライマーの標的外混成化および延長を最小化するために、一実施態様では、遮断された遺伝子特異的プライマーが使用される。そのようなプライマーの3'末端は、PCRポリメラーゼにより延長され得ない一つ以上の残基で遮断される。遮断基は、ポリメラーゼの3'-5'エキソヌクレアーゼ活性により消化されてはならないことがまた重要である。この目的のために、一実施態様では、二つのヌクレオチドがプライマーの末端で逆方向に付着される（これにより、最後から2番目の結合は、3'-3'である）。図7に示されているように、単一RNA残基がDNAオリゴヌクレオチドに導入され、遮断基がプライマーの標的的特異的混成化時に熱安定性RNAse H2酵素により切断され得る。遮断基が切断されるとき、プライマーは、その目的の標的に対して延長することができる。一部の疑似混成化および延長が相変らず発生し得る一方で、そのような手段は、反応に及ぼす

40

50

影響を最小化できる。

【0048】

図7は、校正用ポリメラーゼの3'から5'エキソヌクレアーゼ活性による末端遮断基の消化に抵抗するようにデザインされたRNAseH2活性化可能プライマーを概略的に示す。遮断基は、プライマーの非特異的延長を防止するために、特にプライマー二量体の形成を回避するために、プライマーの3'末端に添加される。プライマーがその標的DNA配列に特異的に混成化されるとき、熱安定性RNAseH2酵素は、単一RNAヌクレオチドでプライマーを切断し、次にポリメラーゼにより延長することができる3'ヒドロキシル末端末端を生成する。「D」で表示された位置は、標的配列に相補的であるDNAヌクレオチドを示す。「r」で表示された位置は、標的配列に相補的であるRNAヌクレオチドを示す。「XX」により表示された遮断基は、逆方向に付着された二つのヌクレオチドを示す(最後から2番目の結合は、3'-3'結合であり、末端「X」は、フリー5'ヒドロキシルを有する)。XX位置は、5'-CE(ベータ-シアノエチル)ホスホラミダイトを使用して合成される。dA-5'ホスホラミダイトが使用されたが、当業者は、また、dC-5'、dT-5'またはdG-5'を使用できる。ポリメラーゼは、5'末端から延長せず、または校正用(proofreading)3'-5'エキソヌクレアーゼ活性は、そのような末端を消化しない。この実例で、プライマーの5'領域は、縮重分子系統タグおよびユニバーサルプライマー配列を有するものと描写されるが、これらの特徴は、選択的であり、且つサンプル特異的バーコードのような他の特徴も含まれることができる。

10

20

【0049】

図8は、系統追跡PCRの概略的な説明を提供する。系統追跡されたPCRの目標は、分子系統タグ(MLT)をPCRの最初の数サイクルの間に鋳型分子に配分し、次にこれらタグ付きコピー体を後続のPCRサイクルの間に(追加のMLTの統合を最小化しつつ)、ユニバーサルプライマーを使用して増幅させることである。このような戦略は、ポリメラーゼ誤取込エラーおよびシーケンサーエラーから純粋な鋳型由来の突然変異を区別するために使用され得る。また、戦略は、二本鎖DNA鋳型の両鎖が液滴またはマイクロウエルのような小さい反応体積内でタグ付けされ、増幅されたことを確認するために使用され得る。系統追跡されたPCRは、試薬を伝達するか、添加することなく、連続的な熱サイクリングプログラムを使用して単一反応体積でまたは多重微細反応体積で行われることができる。方法は、低溶融温度(例えば60)を有する遺伝子特異的プライマー、およびさらに高い溶融温度(例えば72)を有するユニバーサルプライマーを使用する。遺伝子特異的プライマーは、5'領域にユニバーサルプライマー配列だけでなく、MLT配列を含有する。PCRの少なくとも最初の2(しかし、最初の4ほどまで)サイクルは、低いTm(例えば、60)で行われ、MLT含有遺伝子特異的プライマーの混成化および延長を可能にする。後続するPCRの~30サイクルに対して、ユニバーサルプライマーの優先的な使用を促進し、追加のMLTの統合を最小化するために、さらに高いTm(例えば、72)が使用される。ユニバーサルプライマーによる疑似生成物の増幅を回避するために、遺伝子特異的プライマーからプライマー二量体の形成を最小化することが重要である。したがって、プライマー特異性を増大させるためのスキーム、例えばRNAseH2活性化可能な遺伝子特異的プライマーが使用されなければならない。また、ユニバーサルプライマーがRNAseH2活性化可能であってもよいが、選択的である。ここで、ユニバーサルプライマーは、サンプル特異的バーコードを含有することが示されるが、プライマーのこの部分は、省略されるか、または他の特徴が意図的用途に応じて統合され得る。Tm = 溶融温度。MLT = 分子系統タグ。

30

40

【0050】

図9は、系統追跡されたPCR実験の結果を示す。図9(A)は、単一チューブの系統追跡されたPCR実験からの増幅生成物が2%アガロースゲル上で予想されたサイズで移動するバンドを生成することを示す。図9(B)は、系統追跡されたPCR増幅生成物から生成された次世代シーケンシングデータの分析が棒グラフでMLTコピー体の予想さ

50

れた分布パターンを示すことを示す。分析されたサンプルは、健康な支援者の血漿由来の ~ 6000 ゲノム当量の二本鎖野生型 DNA のにスパイクされた既知の K R A S G 1 2 C 突然変異を含有する ~ 20 ゲノム当量の二本鎖 DNA で構成された。X 軸は、所定の M L T 配列ペアが発見された K R A S G 1 2 C 突然変異の読み込みの数を示す。Y 軸は、所定の数の読み込みコピー体を有する特有の M L T 配列ペア (異なるタグ) の数を示す。略 20 個の二本鎖突然変異型 DNA コピー体が反応に添加されたので、~ 40 の異なる M L T 配列ペアが、観察されたように、複数の読み込みカウントを有するものと予想された。

【0051】

また、一実施態様では、ユニバーサルプライマーの特異性は、RNAse H2 - 切断可能な遮断基をプライマー内に組み込むことによって増大することができる。また、一実施態様では、ユニバーサルプライマーは、サンプル特異的バーコードで標識され得、異なるサンプルに対する異なるバーコード化プライマーの使用が、PCR 生成物がプールされ、バッチ式で次世代シーケンシングが行われることを許容する。配列データは、次にバーコード同一性をベースにサンプル特異的ビン (bins) に分類され得る。また、一実施態様では、ユニバーサルプライマーは、アダプタ配列を含有でき、それは、選択された次世代シーケンシング (NGS) プラットホーム上でのシーケンシングを容易にする。一実施態様では、長い (サンプル特異的バーコードおよびアダプタ配列を含有する) および短い (バーコードおよびアダプタを欠く) ユニバーサルプライマーの混合物が使用され得る。短いプライマーがさらに速い混成化速度 (hybridization kinetics) を有するため、それらは、PCR の初期サイクル中の増幅効率を増大させることができる。

10

20

【0052】

特定の実施で、DNA 生成物は、大量並列シーケンシングが行われる前に、所望サイズの生成物を選択し、使用しないプライマーを除去するためにゲル精製される。特定の実施で、限定するものではないが、ビオチンタグ付き相補性オリゴヌクレオチドを使用するハイブリッド捕捉、高性能液体クロマトグラフィー、毛細管電気泳動、シリカ膜分割、または磁性固相可逆的固定化 (SPRI) ビーズへの結合を含んで他の精製接近法が使用され得る。

【0053】

一実施態様では、次世代シーケンサーが、タグされ、増幅され、精製された PCR 生成物から大多数の配列を得るために使用される。そのようなシーケンサーにより生成されたクローン配列 (各配列は、単一核酸分子から発生する) は、ウルトラディープシーケンシングとして知られている接近法を使用して変異体分子を確認し定量するために使用され得る。原則的に、多数の配列が各標的部位および各サンプル毎に得られるため、希少な変異体が検出され、測定され得る。しかし、シーケンサーのエラーが純粋な変異体として誤判断され得るため、シーケンサーのエラー速度は、検出感度を制限できる。シーケンサーエラーの寄与を最小化するために、一実施は、クローン重複ペアエンド配列を使用する。各クローン集団からの DNA の反対鎖を別々にシーケンシングすることによって、且つ、配列の重複領域を比較することによって、シーケンサーエラーから発生する変異体の大多数が除去され得る。一実施態様では、配列重複領域は、突然変異を起こしやすい区域にあるようにデザインされる。一実施態様では、重複領域で完全に一致する読み込みペアだけが追加の分析のために保有される。そのような分析のために、クローンペアエンド読み込みを生成するシーケンサーが有用である。特定の実施で、他の大量並列シーケンシングプラットホームが、また、活用され得る。

30

40

【0054】

一実施態様では、PCR 増幅、プロセッシングまたはシーケンシング中に導入されたエラーは、変異体配列と関連した分子系統タグ (MLT) の分布を分析することによって、純粋な鋳型由来の突然変異配列と区別され得る。もし所定の標的サンプルビン (bin) に対する獲得された NGS 読み込みの数が、そのサンプル内の標的化された鋳型 DNA コピー体の数倍さらに大きい場合、本来配分された M L T が多重コピー体で存在すると予

50

想され得る。したがって、もし突然変異鋳型DNA断片がPCRの早期サイクルの間にMLT配列で標識された場合、配列データは、そのMLT配列および突然変異を有する多重読み込みを含有すると予想され得る。逆に、PCRエラーまたはシーケンサーエラーから発生する変異体は、同じMLT配列を有するさらに少ない数の読み込みを含有すると予想され得る（通常、各MLT配列は、1回のみ発生し得る）。また、一実施態様では、MLTは、プール増幅の間にクロスオーバー事象により不正確なサンプル特異的バーコードを有する配列を区別するに使用され得る。

【0055】

DNA二重螺旋の両鎖上で一致する突然変異を確認するための区画化されたPCRおよび引き続くNGS

たとえ前記記述された系統追跡PCR方法が大部分のPCRエラーとシーケンサーエラーから純粋な鋳型由来の突然変異を区別できるが、PCRの最初の数サイクルの間に発生する誤取込を確認することは難しい。そのような初期誤取込エラーから発生する変異体配列は、純粋な鋳型由来の突然変異に対して予想された多重MLTコピー体と同様に、相対的に高い数のMLTコピー体と関連付けることができる。このような制限を改善するために、鋳型由来の突然変異を確認するための代替戦略は、所定の二本鎖の鋳型DNA断片の両鎖で同じ突然変異が存在することを確認することである。PCRからまたは鋳型DNAの塩基損傷から発生するエラーが同じ鋳型断片の両鎖のコピー体に対する相補性変更を生成する可能性は、殆どないものと見られる。

【0056】

一実施態様では、区画化、タグ付け、増幅およびシークエンシング戦略が、突然変異が二本鎖鋳型DNA断片の両鎖に存在することを証明するために使用される。一実施態様では、PCR反応カクテルは、前記系統追跡されたPCRに対して使用されたもの（ユニバーサルプライマーおよびMLT配列を含有するRNAseH2-活性化可能な遺伝子特異的プライマーの混合物を含有する）と類似している。しかし、決定的な違いは、長いユニバーサルバーコードプライマー（順方向または逆方向）のうち一つがカクテルから省かれ、区画固有バーコードを含有するプライマーがその代わりに使用され得ることである。一実施態様では、PCR反応カクテル（鋳型DNA断片を含む）は、多くの微小流体区画に分けられ、任意の所定の区画が特定の標的化された鋳型DNA断片の一つ以上のコピー体を含有する可能性が、非常に低い。図13に示されたように、区画は、多数の増幅可能な標的化された断片（異なる標的）を有することができるが、同じ標的の一つ以上のコピー体を有することはめったにない。例えば、もし所定の標的のコピー体が10個の区画のうち略一つにのみ発見されると、同じ区画でその標的の二つのコピー体を発見する可能性は、 $\sim 1/100$ になり得る。すべての区画は、ユニバーサルプライマーおよび遺伝子特異的プライマーの全パネルを含有しているため、区画内のすべての増幅可能な標的がタグ付けされ、コピーされ、増幅される。一実施態様では、すべての区画は、同時に同じ熱サイクリングプロトコル（系統追跡されたPCRに対して使用されたものと類似している）が行われる。

【0057】

図13は、異なる標的がPCR増幅のためにどのように液滴またはマイクロウェル内に無作為に区画化されるかの実例を示す。各文字は、標的化された鋳型DNA断片を示し、それぞれの文字発生は、その標的の単一コピー体を示す。増幅反応の区画化は、所定の増幅可能な、標的化された鋳型DNA断片の通常0または一つ（および時には二つ以上）のコピー体が区画内に存在するように行われる。しかし、多重ゲノム領域が同時に標的化されるため、色々な異なる標的化されたDNA断片（通常、それぞれの単一コピー体の、時には一つ以上のコピー体）が区画内に存在することができる。

【0058】

図14は、単一反応区画（例えば、マイクロウェルまたは液滴）の内容の実例を示す。図示されたものは、MLT含有の遺伝子特異的プライマー、ユニバーサルプライマー、標的化された鋳型DNA断片（および他の非標的化されたDNA断片）、およびビーズ特有

10

20

30

40

50

のバーコードを有するビーズ担持熱放出性プライマーである。これに加えて、反応区画は、反応緩衝液、dNTP、RNAseH2酵素およびポリメラーゼ（例えば、Phusion Hot Start）を含有する。すべての区画は、遺伝子特異的プライマーの全体パネルを含有する。各遺伝子特異的プライマーは、MLT配列を含有し、また、ユニバーサルプライマー配列の一部を有する。各遺伝子特異的プライマーは、5～50nMのような相対的に低い濃度で存在する。ユニバーサルプライマーは、高濃度（例えば、200～500nM）で存在する。ビーズから放出されたバーコード化プライマーは、区画で相対的に低い濃度（～5～50nM）を有するものと予想される。二本鎖DNA鋳型断片は、最も強固なエラー抑制を許容するが、一本鎖鋳型がまた使用され得る。任意の所定のマイクロビーズは、同じビーズ特有のバーコードを有するプライマーの多重コピー体を担持する。区画内のビーズ分布がほぼ無作為であるため、多くの区画が一つ以上のマイクロビーズを含有し、少数の区画が何も含有しない（ポアソン統計によって測定される）。この実例で、ピオチン標識増幅生成物は、次にストレプトアビジンコーティングされたビーズを使用して捕捉され、分離される。

10

20

30

40

50

【0059】

図15AおよびBは、バーコード化プライマーを運搬する単一マイクロビーズを含有するマイクロコンパートメント内で行われる系統追跡PCRの二つの実例のシナリオを示す。パネルAは、二重螺旋の両鎖に純粋な突然変異を含有する二本鎖標的化されたDNA断片のタグ付けおよび増幅を示す（二重螺旋の二本鎖は、完全に相補的である）。この場合、同じビーズ特有のバーコードがすべての増幅生成物に配分される。二つの別個のMLTペア（すなわちA-B、およびC-D）を含有する多重読み込みの突然変異の存在は、突然変異が鋳型DNAの両鎖に存在したことを示す。パネルBは、野生型二本鎖DNA断片の類似なタグ付けAC増幅を示す。この場合、増幅生成物は、少数のポリメラーゼエラーを含有するが、配列がビーズ特有のバーコードによりグループ化されるとき、一貫した突然変異は見られない。異なる文字で標識されたMLTおよびバーコード（例えば、MLTGまたはバーコードW）は、異なるヌクレオチド配列タグを示す。単純化のために、各タグまたはバーコードは、単一文字のアルファベットにより確認されるが、実際には、各タグは、通常、6～10個の塩基のストレッチで構成される。

【0060】

図16AおよびBは、バーコード化プライマーを担持する単一マイクロビーズを含有するマイクロコンパートメント内で系統追跡されたPCRが行われる二つの追加実例のシナリオを示す。パネルAは、二つのDNA鋳型鎖のうち一つをコピーするとき、ポリメラーゼ誤取込エラーがPCRの第1サイクルの間に発生した野生型二本鎖DNA断片のタグ付けおよび増幅を示す。これは、PCRの第1サイクルの間に発生したとしても、エラーがどのように区別され得るかに対する極端な実例として示される。この場合、増幅生成物は、純粋な突然変異が鋳型DNA二重螺旋の両鎖からコピーされたとすれば、予想し得るように、二つのMLTペア全部（すなわちI-JおよびK-L）でなく、二MLTペアのうち一つ（すなわちI-J）にのみ関連したエラーを示す。パネルBは、ポリメラーゼ誤取込エラーがPCRの第1サイクルの間に発生した野生型一本鎖DNA断片のタグ付けおよび増幅を示す。この場合、たとえエラーがその区画内で増幅されたコピー体の集団全体で発見され得るが（バーコードZでタグ付けされる）、コピー体はすべて、全部鋳型DNA二重螺旋の両鎖からコピーされた純粋な突然変異に対して予想し得るように、二つ（またはそれ以上）のMLTペアでない、単一MLTペア（すなわちM-N）を有する。

【0061】

図17は、所定の区画で二つ（またはそれ以上）のビーズ上に二つ（またはそれ以上）の異なるバーコード化プライマーがあった場合、分析がどのように行われるかを示す。ビーズは、Poisson分布によって異なる区画内に分布すると予想され、この際、一部の区画は、ビーズを含有せず、一部の区画は、単一ビーズを含有し、一部の区画は、二つ以上を含有する。ビーズを含有しない区画の数を減少させるために、当業者は、区画当たり二つまたは三つのビーズの中間値を達成することを目標とすることができる。あるいは

、単一区画内に単一ピーズを分配するために、P o i s s o n 統計を克服するための方法が存在するが、これらの接近法は、複雑な微小流体操作またはプライマーの規定された反応チャンパー内への事前提供を含む。一つ以上のバーコード化プライマーセットが存在する区画は、配列データの後続的なコンピュータ分析中に確認され得る。所定の M L T ペアが一つ以上の区画に由来する配列で発見された可能性が非常に低いため、そのようなペアに関連したすべての区画固有バーコードは、単一区画から誘導されたものと推論され得る。

【 0 0 6 2 】

一実施態様では、分子系統タグ (M L T) は、遺伝子特異的プライマーを通じて鋳型分子に配分され、次にこれらのタグ付きコピー体は、系統追跡された P C R に対して記述されたように、ユニバーサルプライマーにより増幅される。区画内で、もし一般的に所定の標的化された二本鎖鋳型 D N A 断片のコピー体の一つ以上存在しなければ、M L T は、二つの異なる鎖のコピー体から発生する増幅された配列を確認するために使用され得る (図 1 5 に示される) 。一実施態様では、一つまたは少数の区画固有タグを含有するプライマーが所定の反応区画内で生成されたアンプリコンを確認するために使用され得る。したがって、そのようなタグ付け計画を使用して、同じ変異体配列が同じ区画内で D N A の二つの異なる鎖からコピーされたことを確認することが可能である。

10

【 0 0 6 3 】

P C R カクテルは、多様な方法で微小流体区画内に分けられることができる。一実施態様では、区画は、10ピコリットルほど小さく、10ナノリットルほど大きくてもよい。特定の実施で、区画は、~ 0 . 1 から 1 ナノリットルの体積である。理想的には、所定の実験に対して区画の体積は、均一でなければならない。区画の数は、鋳型 D N A 分子の用途および予想された濃度によって数千~数百万の範囲であり得る。一実施態様では、P C R 区画は、微小流液体滴発生装置を使用してオイル中の P C R カクテルの液滴として製造され得る。ミネラルオイルがこの目的のために使用され、またはフッ素化油が使用され得る。界面活性剤が P C R の前または途中で、液滴を安定化させ、液滴の結合を防止するために使用され得る。一実施態様では、オイル中の P C R カクテルのエマルジョンが、また、混合物を激しく攪拌することによって簡単に生成できる (しかし、この接近法は、不均一なサイズの液滴を生成する短所を有する) 。他の実施で、P C R カクテルは、微小流体装置でマイクロウェルに区画化され得る。一実施態様では、数千のナノリットルサイズのウェルを有するパターン化ポリジメチルシロキサン (P D M S) を含有するスライドが使用され得る。一実施態様では、反応体積がオイルまたは空気により分離される、狭い蛇行チャンネルを含有する微小流体装置が使用され得る。一実施態様では、P C R カクテルがチャンネル内に導入された後、チャンネルが同時に閉じられる数千個の微細バルブにより別途の反応チャンパーに分けられることができる同様の微小流体装置が使用され得る。P C R は、マイクロコンパートメントを同時に熱サイクリングすることによって行われることができる。

20

30

【 0 0 6 4 】

一実施態様では、区画固有タグ (またはバーコード) を含有するクローンプライマーは、マイクロピーズを通じて区画に導入され得る。それぞれ均一にタグ付きプライマーの多くのコピー体を担持する多数のマイクロピーズを製造することが可能であるが、非常に多様なタグが、異なるピーズに存在する。所定のピーズが表面でタグ付きプライマーのクローン集団を担持するが (全部同じタグを有する) 、異なるピーズが異なるタグを有するプライマーを担持する。一実施態様では、マイクロピーズは、P C R カクテルと混合され得、カクテルで区画化され得る。一実施態様では、ピーズの濃度は、平均 2 または 3 個のピーズが各区画に伝達されるように調整され得る (少数の区画はピーズを有しない) 。ピーズの区画内への分布は、P o i s s o n 統計に従うと予想される。一実施態様では、プライマーは、ピーズ表面から加熱により (ピーズに付着された相補性 D N A 鎖からプライマーを溶融させることによって) 区画化された溶液中に放出され得る。他の実施で、プライマーは、光切断によりピーズ表面から区画化された溶液中に放出され得る (光切断可能な

40

50

ホスホラミダイトがオリゴヌクレオチドをビーズ表面に連結させるために使用され得る)。他の実施で、プライマーをビーズに付着されたままにし、混成化および重合反応は、ビーズ表面で行われることができる。一実施態様では、超常磁性ビーズ(交差結合されたポリスチレンでコーティングされ、表面は、アミンまたはヒドロキシル基で活性化される)が使用され得る。他の実施で、限定するものではないが、アガロース、ポリアクリルアミド、ポリスチレンまたはポリメチルメタクリレートを含む物質で構成されたビーズが使用され得る。一実施態様では、ビーズは、ビオチン標識オリゴヌクレオチドに結合するために、ストレプトアビジンでコーティングされ得る。特定の実施で、ビーズは、0.5マイクロメートルから100マイクロメートルのサイズであり得る。特定の実施で、ビーズは、1マイクロメートルから5マイクロメートルのサイズである。特定の実施で、所定の実験で使用されたビーズは、相対的に均一なサイズで各ビーズ上に相対的に均一な数のプライマーコピー体を担持する。

10

【0065】

図10は、ビーズ特有のバーコードを含有する熱放出性プライマーがどのようにマイクロビーズ上で製造され得るかの実例を示す。まず、オリゴヌクレオチドが自動オリゴヌクレオチド合成器でマイクロビーズの表面で標準ホスホラミダイト化学を使用して合成され得る。マイクロビーズ表面は、例えばアミンまたはヒドロキシル基で官能化され得、それは、ホスホラミダイト単量体と共有結合を形成する。追加のホスホラミダイト単量体が以後に順次に標準合成プロトコルを使用して添加され得る。ビーズ結合されたオリゴヌクレオチドの所望の方向によって、標準または5'-ベータ-シアノエチルホスホラミダイト単量体のいずれかが使用され得る。オリゴヌクレオチドとビーズ表面の間に若干の距離を導入させるために、一つまたは多数のスペーサーホスホラミダイトがビーズ表面に添加された後、ヌクレオチド単量体が添加され得る。方法項で記述されたように、スプリットおよびプール合成がビーズ特有のバーコードをオリゴヌクレオチドに組み込むために使用され得る。もしマイクロビーズが小さすぎて自動オリゴヌクレオチド合成器のカラムに使用されたフリットにより保有されない場合、当業者は、磁石により所定の位置に保有された超常磁性マイクロビーズを使用できる。共通プライミング配列(および選択的なビオチン基)を含有する第2オリゴヌクレオチドがDNAポリメラーゼを使用してビーズ結合されたオリゴヌクレオチドをコピーすることができる。このようにして、延長したプライマーは、ユニバーサルプライマー配列だけでなく、ビーズ特有のバーコード配列を含有する。ビーズが液滴またはマイクロウェルのようなさらに小さい反応体積に区画化された後、ビーズ特有のバーコードを含有する延長したプライマーが熱変性によりビーズから放出され得る(例えば、PCRの途中に)。他のプライマー放出方式、例えば光切断および化学的脱カップル化がまた使用され得る。

20

30

【0066】

図11は、熱変性により放出され得る一時的に固定化されたオリゴヌクレオチドを製造するための代替方法を示す。切断可能な基(例えば光切断可能なリンカー)を含有するオリゴヌクレオチドは、表面(例えば、マイクロビーズ)で直接合成されるか、または表面、粒子または分子に共有結合またはビオチン親和性捕捉を通じて合成後に結合され得る。規定されたバーコード配列または縮重タグ配列(例えば、MLT)のセットがオリゴヌクレオチドに統合され得る。タグは、また、スプリットおよびプール合成を通じて合成され、所定のビーズ(または粒子)上に同じタグの多数のコピー体を有する非常に多様なタグが製造され得る。オリゴヌクレオチドは、自己相補性の領域を有するようにデザインされ、切断されたオリゴヌクレオチドが塩基対相互作用(混成化)を通じて付着されたまま維持される。オリゴヌクレオチドは、熱変性により後に溶液中に放出され得る。オリゴヌクレオチドは、下流の用途によって5'から3'に、または3'から5'に合成され得る。

40

【0067】

一実施態様では、多様なセットのクローンタグ付きプライマー(1ビーズ、1タグ)を担持するビーズ集団がスプリットおよびプールオリゴヌクレオチド合成接近法を使用して合成され得る。共通プライマー配列が自動オリゴヌクレオチド合成器で標準ホスホラミダ

50

イト化学を使用して合成され得る。プライマーは、適切なホスホラミダイトを使用して5'から3'、または3'から5'方向に合成され得る。一実施態様では、ホスホラミダイトは、その表面がアミンまたはヒドロキシル基で変形されたビーズを使用することによって、ビーズに共有結合され得る。一実施態様では、永久磁石または電気磁石が自動オリゴヌクレオチド合成器で合成カラム内に磁性マイクロビーズを保有するために使用され得る（ビーズがフリットにより保有されるには、あまりに小さい可能性があるからである）。一実施態様では、スプリットおよびプール合成接近法がビーズ上にクロンの多様なタグを生成するために使用される。プライマーの共通領域が形成された後、合成器がタグ配列が始まる所で停止する。一実施態様では、ビーズがプールされた後、4個の異なる新しいカラムに分けられ、異なるホスホラミダイト（dA、dT、dCまたはdG）が4個のカラムに添加される（カラム当たり一つのホスホラミダイト）。他の実施で、4以上または未満のカラムおよび4つのホスホラミダイトが使用され得る（所定の位置で可能な残基数を増加または減少させるためである）。タグ領域内で各カップリングサイクル後に、ビーズは、プールされ、次のサイクルのために新しいカラムに再分布される。この方法で、所定のビーズにカップリングされたオリゴヌクレオチドは、所定のサイクルで同じ塩基を受けが、所定の位置に添加されたその塩基は、無作為に選択される。一実施態様では、ビーズ特有のタグ配列は、1～15塩基の長さであり得る。特定の実施で、ビーズ特有のタグ配列は、8～12塩基長さであり得る。一実施態様では、相補性プライマーがビーズ結合されたオリゴヌクレオチドに混成化され得、ポリメラーゼを使用して延長し、図10に概略的に示されたように、タグ配列および追加のプライマー配列がコピーされる。延長したプライマーは、ビーズ特有のバーコードを有する熱放出性プライマーとして作用する。一実施態様では、この熱放出性バーコード化プライマーは、区画内でPCR増幅された標的上に混成化し、延長させるために使用され得る（熱放出性プライマーの3'末端は、標的化されたアンプリコンとの混成化を容易にするために、ユニバーサルプライマー配列の一部を含有する）。

10

20

30

40

50

【0068】

他の実施で、区画固有タグを含有するプライマーは、区画内に事前分布され得る。例えば、もしPCRカクテルが微小流体装置のマイクロウェルに分けられると、区画固有タグを含有するプライマーが各マイクロウェルに添加された後、PCRカクテルが添加され得る。一実施態様では、プライマーは、マイクロウェルの表面または壁に化学的にカップリングされるか、またはビオチン-ストレプトアビジン相互作用を通じてカップリングされ得る。一実施態様では、プライマーは、加熱により（前記記述されたように固定された相補性オリゴヌクレオチドの溶融により）、光切断により、または他の手段によりマイクロウェルから放出され得る。一実施態様では、プライマーは、壁の表面に付着されたまま、維持され得、重合化は、表面で行われることができる。

【0069】

一実施態様では、タグ付き増幅生成物は、多数の小さい反応体積の内容物を組み合わせることによって、PCR後にプールされる。一実施態様では、これは、オイル中の水性液滴を結合させる試薬（例えば、クロロホルム）を添加することによって行われることができる。一実施態様では、反応体積は、微小流体装置上のマイクロウェルからの反応生成物を採取することによって組み合わせられることができる。一実施態様では、プールされ、増幅されたDNA生成物は、ゲル精製され、所望のサイズの生成物が選択され、大量並列シーケンシングが行われる前に、使用されないプライマーが除去される。特定の実施で、精製するための他の接近法として、例えば限定するものではないが、ビオチンタグ付き相補性オリゴヌクレオチドを使用するハイブリッド捕獲、高性能液体クロマトグラフィ、毛細管電気泳動、シリカ膜分割、または磁性の可逆的固定化（SPRI）ビーズへの結合が使用され得る。

【0070】

一実施態様では、次世代シーケンシング（NGS）がタグ付けされ、増幅され、精製されたPCR生成物から大多数の配列を得るために使用される。一実施態様では、（前記

記述されたような) クローン重複ペアエンドシーケンシング接近法が使用され、シーケンサー誘導されたエラーを含有する読み込みが除外することができる。一実施態様では、配列データは、標的化された二本鎖鋳型DNA断片の両鎖をコピーすることから得られる純粋な突然変異を確認するために分析される。そのような純粋な突然変異を確認するために使用された戦略は、図15~図17を参照して理解され得る。次の論理が使用される：

【0071】

1. 一実施態様では、MLTパターンは、マイクロコンパートメント内の増幅されたPCR生成物が一つの鋳型鎖または二つの鋳型鎖をコピーすることから得られたか否かを判定するために使用され得る。一実施態様では、もし単一MLT配列ペアが所定の区画から増幅された配列で見られると、増幅された配列は、その区画内で増幅されたDNAの一本鎖に由来したと推論され得る。一実施態様では、もし二つ(またはそれ以上)のMLT配列ペアが所定の区画から増幅された配列に見られると、増幅された配列は、その区画内で増幅された二つ(またはそれ以上)のDNAの鎖から得られたと推論され得る。

10

【0072】

2. 一実施態様では、PCR増幅された配列は、区画固有バーコードの分析をベースに所定の区画から得られたものと確認され得る。一実施態様では、区画に配分された単一バーコードがあり得る。他の実施で、区画に配分された一つ以上のバーコードがあり得る。もし一つ以上のバーコードがあると、バーコードの組合せは、同じ区画に由来したものであって、PCR生成物を確認するために使用され得る。

20

【0073】

3. 一実施態様では、突然変異は、もし(a)所定の区画から得られた増幅された配列の大部分が突然変異を含有し、(b)観察されたMLTパターンが増幅された配列が一つ以上の鋳型鎖から得られたことを確認すると、真正な鋳型由来の突然変異であると見なされる。区画が一つ以上のDNA断片を含有する可能性が殆どないため、一つ以上の鋳型鎖から誘導された配列は、二重螺旋DNA断片の相補鎖から誘導されたものとほぼ確実に推論され得る。

【0074】

異なる区画にクローntag付きオリゴヌクレオチドを伝達する方法：

異なる区画にクローntag付きプライマーを伝達するためにビーズを使用することは、色々な短所を有する。このようなビーズ集団の合成は、特にスプリットおよびプール工程が使用されるため、複雑であり得る。また、ビーズの区画内への無作為分布を保証することが難しいことがあるが、それはビーズが沈降するか、凝集し、Poisson統計に従わない分布に至るためである。より無作為なビーズ分布を達成するために、ビーズスラリーが連続的に攪拌される必要があり得、または区画化が迅速に行われ、ビーズの沈降を最小化できる。

30

【0075】

クローntag付きプライマーをマイクロコンパートメントに事前分配することは、過程の複雑性という短所を有する。プライマーは、異なるタグで別々に合成されなければならない。これは、特殊なロボット装置の使用を含む。タグ付きプライマーを数百または数千個のマイクロウェルに分布させることは、実現可能になり得るが、さらに多くの(例えば、数百万)の区画に対してこれを達成することは難しい。

40

【0076】

クローntag付きオリゴヌクレオチドを、オリゴヌクレオチドを表面(例えば、ビーズまたはマイクロウェル壁)に付着させることを必要とせずにマイクロコン区画に伝達する方法および組成物が開示される。溶液中のオリゴヌクレオチドの使用は、タグの区画内へのさらに均一な分布を保証し、実施するに非常に簡単であるため、有益である。計画は、図12に概略的に示される。

【0077】

図12は、クローntag付きオリゴヌクレオチドをマイクロコンパートメント内に伝達

50

するイン・ソリューション方法を示すが、オリゴヌクレオチドは、同じ反応体積で共同増幅されるPCR生成物に区画固有タグを添加するためのプライマーとして機能することができる。縮重タグ配列を含有する鑄型オリゴヌクレオチドは、PCRカクテルに添加され、PCRカクテルが区画化されるとき、少数の個別的な鑄型オリゴヌクレオチド分子（例えば平均2～3分子）が各区画に分配され得る。鑄型オリゴヌクレオチドを増幅させることができるプライマーが、また、反応カクテルに含まれる。したがって、PCRが行われるとき、各区画内の少数の鑄型オリゴヌクレオチドが少数のクローン区画固有タグを含有する多数のコピー体を生成するために増幅される。これらのクローントグ付きオリゴヌクレオチドは、同一反応体積内で共同増幅された他のPCR生成物に区画固有タグを配分するためのプライマーとして使用され得る（例えば、多重ゲノム領域の系統追跡PCRを通じて）。

10

【0078】

一実施態様では、均一なタグ付きオリゴヌクレオチド配列の多くのコピー体が、該タグ付きDNA配列の単一分子を区画内に導入させた後、それを区画内で短いプライマーを使用して（PCRを通じて）コピーし、増幅させることによって、区画で製造され得る。鑄型として単一タグ付きDNA分子を開始することによって、区画内で増幅されたコピー体は、クローン的であり、且つ、鑄型分子として同じタグを内部に有する。一実施態様では、タグ付き鑄型DNAは、二本鎖であり得る。他の実施で、鑄型DNAは、上部または下部相補鎖のいずれか一つで構成される、一本鎖であり得る。一実施態様では、鑄型分子の集団内のタグ（またはバーコード）配列は、オリゴヌクレオチド合成中に縮重位置を組み込むことによって（例えば、多重「N」位置を組み込むことによって、Nは、T、C、GまたはA塩基を略同一にカップリングさせる可能性を示す）、生成され得る。一実施態様では、事前規定されたバーコードが、また、鑄型分子に統合され得る。一実施態様では、一つ以上の異なるタグ付き分子が区画内で鑄型として使用され得、この場合、区画内で増幅されたオリゴヌクレオチドが一つ以上のタグ配列を含有する。特定の実施で、タグ付き鑄型分子を含有しない区画の数を最小化するために、平均二つまたは三つの異なるタグ付き鑄型分子が区画に導入され得る（Poisson統計に従って分布する）。一実施態様では、その結果として区画内で得られた増幅されたクローントグ付きオリゴヌクレオチドコピー体が区画内で他のDNA配列に混成化し、それをコピーすることによって、プライマーとして機能することができる。一実施態様では、そのようなプライマーは、区画内で増幅生成物に区画固有タグを配分させるために使用され得る。もし一つ以上の区画固有タグ（バーコード）を含有するプライマーが区画内に存在すると、タグの組合せは、所定の区画から得られるものであって、増幅生成物を確認するのに使用され得る。一実施態様では、区画内でタグ付き鑄型分子を増幅させるために、順方向および逆方向の短いプライマーの不均一な濃度が使用され得る。一実施態様では、順方向プライマーは、逆方向プライマーより2倍～20倍さらに濃縮され得る（またはその反対）。不均一な濃度のプライマーの使用は、「非対称PCR」を生じさせ、一つの増幅された鎖のその補体よりさらに多いコピー体が生成される。一実施態様では、そのような非対称増幅は、増幅されたクローントグ付きオリゴヌクレオチドの区画の他のDNA配列との混成化を促進できる（これにより、増幅されたオリゴヌクレオチドがタグ付きプライマーとして作用することが可能になる）。図12は、このような接近法を示す。

20

30

40

【0079】

クローントグ付きオリゴヌクレオチド配列の多くのコピー体を反応区画に導入させるためのこの方法は、多くの潜在的用途を有する。一実施態様では、この方法は、前記記述されたように、低存在量の突然変異型DNA分子の測定を補助するために使用され得る。他の実施で、この方法は、単一細胞ゲノムデータを生成するために、異なる区画で単一細胞から増幅されたDNA生成物をタグするために使用され得る。他の実施で、方法は、単一細胞の高処理量RNAプロファイリングを容易にするために、異なる区画においての単一細胞からの相補的DNA（cDNA）のコピー体を標識するために使用され得る。他の実施で、方法は、ゲノム配列アセンブリを容易にするために、区画内でさらに大きい染色体

50

断片に由来する多重アンプリコンに同じタグを配分するために使用され得る。

【0080】

他の実施で、区画固有DNAタグ付け方法は、高度に複合化された単一細胞プロテオミクスを容易にするために使用され得る。この接近法で、異なる蛋白質を標的とする抗体は、両側に共通プライマー結合配列が隣接する抗体特異的バーコード配列を含有するオリゴヌクレオチドで標識され得る。抗体の複合化されたパネルは、無傷細胞の表面上の、または固定型且つ透過性細胞内部の蛋白質に結合され得る。パネルの各抗体は、異なる抗体特異的タグを含有するオリゴヌクレオチドで標識される。余分の抗体が洗浄された後、細胞は、各区画が一つ以上の細胞を含有する可能性がないように区画化される（例えばオイル中の水性液滴に、または微小流体装置のマイクロウェルに）。区画内の共通PCRプライマーは、共通プライマー結合配列を通じてすべての抗体-結合バーコーディングされたオリゴヌクレオチドを同時に増幅させるに使用され得る。区画内の増幅されたタグの相対存在量は、細胞内でその蛋白質標的に結合された対応抗体の相対存在量を反映する。次いで区画固有バーコードは、異なる単一細胞で蛋白質の定量を可能にするために導入され得る。非常に多様な抗体特異的タグが生成され得るため、異なる抗体に対する複合化の容量は、事実上無限である。

10

【0081】

より一般的に、記述された方法は、区画内で核酸分子が区画固有タグで標識される必要がある任意の用途に対して使用され得る。

【0082】

実施例

本技術は、次の実施例を参照することによりさらによく理解され得る。

これらの実施例は、具体的な実施を示すことが意図される。

20

【0083】

実施例 1

この実施例は、高処理量RNA定量方法の適用を記述する。方法は、下流サンプルプロセッシングおよび分析を簡素化し、費用を減少させるために、多重RNA-含有サンプルの先行並列化を可能にする。

【0084】

材料および方法

RTプライマーミックスのモジュール合成：

プライマーの混合物を生成するために、2段階モジュール式オリゴヌクレオチド合成戦略を使用し、この際、各混合物は、5'セグメントに別個のサンプル特異的バーコードおよび3'セグメントに均一な割合の多重標的的特異的配列を有する（図1a）。まず、色々な標的的特異的3'セグメントを別途のオリゴヌクレオチド合成カラムで形成した。合成を標準ホスホラミダイト化学を使用して3'から5'方向に40ナノモルポリスチレン支持カラム（Prime Synthesis, Aston, PA）でDr. Oligo 192自動合成器を使用して行った。合成を3'セグメントのオリゴマー化が完了した後に中断し、部分的に合成されたオリゴヌクレオチドが依然としてジメトキシトリチル（DMT）基で保護された状態でポリスチレン支持体に残した。

30

40

【0085】

アルゴンガスをカラムを通じて吹き込んで、ポリスチレン支持体を乾燥させた後、カラムを切断し、ポリスチレン粉末を通常のガラスバイアルに注入した。粒子をポリスチレンを中性に浮揚させるために、2：1～3：1のジクロロメタン：アセトニトリル混合物に懸濁させ滴定しておいた。スラリーを一定に攪拌し、均一な混合を保証する一方で、ピペットを使用して等量の体積のスラリーを新しい合成カラム（下部フリットが所定の位置にある）に分配した。次に、カラムをアセトニトリルでフラッシングし、すべてのポリスチレン粒子が底に沈むようにした。アセトニトリルを重力により完全に流した後、上部フリットを所定の位置に配置し、粉末をカラムに固定した。各サンプル特異的バーコードに対して一つのカラムを形成した。

50

【 0 0 8 6 】

新しいカラムを、合成を継続するために自動合成器に再び載置した。下記の表 6 に示されたように、異なるバーコード配列を、プライマーミックスの 5' セグメントに組み込むための各カラムに配分した。バーコードは、8 個のヌクレオチド長さになるようにデザインし、各バーコードは、セット内の他のすべてのバーコードと最低 2 つの位置で異なる（シーケンサーエラーによって引き起こされる誤分類の可能性を最小化するためである）。また、ユニバーサル PCR プライマー結合配列を各オリゴヌクレオチド混合物の 5' セグメントに添加した。合成器をポリスチレン支持体上に既に存在する部分的に合成されたオリゴヌクレオチドを解明するために、3' 末端に追加の「仮想塩基」が存在するようにプログラミングした。

【 0 0 8 7 】

[表 6]

バーコード目録

BC #	配列	BC #	配列	BC #	配列
1	AATA TCGG	49	CGCA TACA	97	AATCCTGT
2	AATA TGGC	50	CGGA TACT	98	CACTATGT
3	AATC CAGA	51	CGGT AACT	99	GACAATGT
4	AATC GAGT	52	CGGT ATCA	100	GAGTATGT
5	AATC GTGA	53	GACA TAGT	101	GCCAATAT
6	AATG CAGT	54	GACT AAGT	102	GCGTATAT
7	AATG CTGA	55	GACT ATGA	103	TATCCAGT
8	AATG GAGA	56	GAGA ATGA	104	TCAGCAAT
9	ACAC CAAT	57	GAGA TAGA	105	TCTCGTAT
10	ACAC CTAA	58	GCCA TAAT	106	TCTGCTAT
11	ACAC GTAT	59	GCCT ATAA	107	TGACGACT
12	ACAG CTAT	60	GCGA ATAA	108	TGTCCACT
13	ACAG GAAT	61	GGCA TACT	注：BC# 97-108は、放射線照射された血液サンプルだけに対して使用した。	
14	ACAG GTAA	62	GGCT AACT		
15	ACAT AGAG	63	GGCT ATCA		
16	ACTA AGAG	64	TAAC CAGA		
17	ACTA TCAG	65	TAAC GTGA		
18	ACTA TGAC	66	TAAG CTGA		
19	ACTC CTAT	67	TAAG GAGA		
20	ACTC GAAT	68	TAAT ACGG		
21	ACTC GTAA	69	TAAT AGGC		
22	ACTG CAAT	70	TATA ACGG		

10

20

30

40

23	ACTG CTAA	71	TATA AGGC		
24	ACTG GTAT	72	TATC CTGA		
25	AGAC CACT	73	TATC GAGA		
26	AGAC CTCA	74	TATG CAGA		
27	AGAC GACA	75	TATG GTGA		
28	AGAG CACA	76	TCAA TCAG		
29	AGAG GACT	77	TCAA TGAC		
30	AGAG GTCA	78	TCAC GTAA		
31	AGAT AGCG	79	TCAG CTAA		
32	AGTA AGCG	80	TCAT ACAG		
33	AGTA TCCG	81	TCAT AGAC		
34	AGTA TGCC	82	TGAA TCCG		
35	AGTC CACA	83	TGAA TGCC		
36	AGTC GACT	84	TGAC CACA		
37	AGTC GTCA	85	TGAC GTCA		
38	AGTG CACT	86	TGAG CTCA		
39	AGTG CTCA	87	TGAG GACA		
40	AGTG GACA	88	TGAT ACCG		
41	CACA ATGA	89	TGAT AGCC		
42	CACA TAGA	90	TGTA ACCG		
43	CAGA TAGT	91	TGTA AGCC		
44	CAGT AAGT	92	TGTA TGCG		
45	CAGT ATGA	93	TGTC CTCA		
46	CCGA TAAT	94	TGTC GACA		

10

20

30

40

47	CCGT ATAA	95	TGTG CACA		
48	CGCA ATCA	96	TGTG GTCA		

【0088】

モジュール式合成の第2段階が完了した後、オリゴヌクレオチド混合物を、DMT基をのこしたままポリスチレン支持体から切断した。各混合物に迅速な脱保護を行い、引き続いて、別のGlen-Pak DNA逆相カートリッジ(Glen Research, Sterling, VA)で精製した。カートリッジは、完了したオリゴヌクレオチドの5'末末端に選択的に疎水性DMT基を保有し、全長生成物を豊富にした。DMT基を精製完了時に除去した。次に、精製されたオリゴヌクレオチド混合物を乾燥させ、10 mM トリス(pH 6.0)に再懸濁し、10x作業ストックを形成した。miRNAおよびmRNAモジュール式プライマーセグメントの配列を下記の表3、表5および表8に列挙する。

10

【0089】

[表3]

miRNA標的の逆転写のためのモジュール式プライマー

第1段階：部分的に合成されたオリゴヌクレオチド3'セグメント

RNA標的	配列	RNA標的	配列
対照標準A	ATACGACAGC TAG	miR-199a- 5p	ATACGACTAA CCA
対照標準B	ATACGACATG AGT	miR-19a	ATACGACTCA GTT
cel-mir-2	ATACGACGCA CAT	miR-200b	ATACGACTCA TCA
RNU44	ATACGACAGT CAG	miR-203	ATACGACCTA GTG
RNU48	ATACGACGGT CAG	miR-204	ATACGACAGG CAT
U6	ATACGACCCC ACC	miR-205	ATACGACCAG ACT
let-7a	ATACGACAAC TAT	miR-21	ATACGACTCA ACA
let-7b	ATACGACAAC CAC	miR-210	ATACGACTCA GCC
miR-100	ATACGACCAC AAG	miR-218	ATACGACACA TGG
miR-103	ATACGACTCA TAG	miR-22	ATACGACACA GTT
miR-105	ATACGACACC ACA	miR-221	ATACGACGAA ACC
miR-10b	ATACGACCAC AAA	miR-222	ATACGACACC CAG
miR-125a- 5p	ATACGACTCA CAG	miR-224	ATACGACAAC GGA
miR-125b	ATACGACTCA CAA	miR-23a	ATACGACGGA AAT
miR-126	ATACGACCGC ATT	miR-24	ATACGACCTG TTC
miR-128	ATACGACAAA GAG	miR-25	ATACGACTCA GAC
miR-129-5 p	ATACGACGCA AGC	miR-26a	ATACGACAGC CTA
miR-130a	ATACGACATG CCC	miR-27b	ATACGACGCA GAA
miR-132	ATACGACCGA CCA	miR-28-5p	ATACGACCTC AAT
miR-133a、 mir-191	ATACGACCAG CTG	miR-299-5 p	ATACGACATG TAT
miR-134	ATACGACCCC CTC	miR-29c	ATACGACTAA CCG
miR-135b、 mir-216b	ATACGACTCA CAT	miR-301a	ATACGACGCT TTG

10

20

30

40

miR-136	ATACGACTCC ATC	miR-302a, mir-367	ATACGACTCA CCA	
miR-137	ATACGACCTA CGC	miR-30b	ATACGACAGC TGA	
miR-140-5 p	ATACGACCTA CCA	miR-31	ATACGACAGC TAT	
miR-141	ATACGACCCA TCT	miR-32	ATACGACTGC AAC	
miR-142-3 p	ATACGACTCC ATA	miR-342-3 p	ATACGACACG GGT	10
miR-143	ATACGACGAG CTA	miR-34a	ATACGACACA ACC	
miR-145	ATACGACAGG GAT	miR-34c-5 p	ATACGACGCA ATC	
miR-146a	ATACGACAAC CCA	miR-363	ATACGACTAC AGA	
miR-148a	ATACGACACA AAG	miR-365	ATACGACATA AGG	
miR-149	ATACGACGGG AGT	miR-372	ATACGACACG CTC	20
miR-150	ATACGACCCAC TGG	miR-373	ATACGACACA CCC	
miR-154	ATACGACCGA AGG	miR-375	ATACGACTCA CGC	
miR-16	ATACGACCGC CAA	miR-381	ATACGACACA GAG	
miR-181a	ATACGACACT CAC	miR-382	ATACGACCGA ATC	
miR-182	ATACGACAGT GTG	miR-383	ATACGACAGC CAC	30
miR-183	ATACGACAGT GAA	miR-422a	ATACGACGCC TTC	
miR-184	ATACGACACC CTT	miR-424	ATACGACTTC AAA	
miR-185	ATACGACTCA GGA	miR-488	ATACGACGAC CAA	
miR-186	ATACGACAGC CCA	miR-7	ATACGACACA ACA	
miR-18a	ATACGACCTA TCT	miR-9	ATACGACTCA TAC	40
miR-192	ATACGACGGC TGT	miR-92a	ATACGACACA GGC	
miR-194	ATACGACTCC ACA	miR-93, mir-106a	ATACGACCTA CCT	
miR-196b	ATACGACCCC AAC	miR-95	ATACGACTGC TCA	
miR-199a- 3p	ATACGACGAA CAG	miR-96	ATACGACAGC AAA	

第2段階：次の配列を96新しいカラムで部分的に合成されたオリゴの混合物に添加した：

CCATCTCATCCCTGCGTGTCTCCGACTCAG [BC] CGTGAGC

独特な8-塩基バーコード（[BC]により表示される）を各カラムに対して使用した。バーコードは、表6にある。

相補的マスク用オリゴ： GTCGTATGCTCACGAAAAAAAAACTGAGTCG
GAGACACGCAGGGATGAGATGG-3'-ビオチン

【0090】

[表5]

MAQC mRNA標的の逆転写のためのモジュール式プライマー

第1段階：部分的に合成されたオリゴヌクレオチド3'セグメント

標的遺伝子	配列
ANXA5	TATCTGCAAGGTAGGCAGGTATAC
B2M	TATCCATGATGCTGCTTACATGTC
CDK9	TATCGTACAGCTCATAGTTGTCCAC
CDKN1A	TATGTTTGGAGTGGTAGAAATCTGTC
CTPS	TATCTCTTGAAGAATCATCCGCTAC
CYP1B1	TATGTCTGCACATCAGGATACCTG
CYP2D6	TATCTCATCCTTCAGCACCGATG
DAD1	TATAGGAAATTCAAAGAGTGAACATTC
ELAVL1	TATGCCTCCGACCGTTTGTCAAAC
FANCG	TATGAAGCAGGTGAAAGTAAGTGTC
FGF9	TATCTGTTCTCTGAATACACACTCTTG
ICAM1	TATCAGTCGGGGGCCATACAG
IGF2R	TATACGTTGGAACCTTCTCCTACAG
IGFBP2	TATGCCCGTTCAGAGACATCTTG
IGFBP5	TATTCACGCACCAGCAGATG
IL1B	TATGCCACAGGTATTTTGTCAATTAC
IL8	TATTTTATGAATTCTCAGCCCTCTTC
KIT	TATTCATTATGTCATACATTTTCAGCAG
MAP2K6	TATGTAGGCCGTTCTTTGGAATTC
MAP3K14	TATGGCTGCAGCTGGGATCTG
MX2	TATTCACTTTTATGTCTTCAATCGTG
MYB	TATGTACTGCTACAAGGCTGCAAG
RARA	TATCGGCTGCTCCAGGTCCTG
RB1	TATTGAGGTATCCATGCTATCATTC
SLC15A2	TATGGAGCACAGATTTTCATGCTAG
SLC2A1	TATCCACGATGCTCAGATAGGAC
SOD1	TATCAGCGTTTCTGTCTTTGTAC
TGFBR2	TATCACTCAGTCAACGTCTCAC
TGM1	TATGCGCAGTGTCACTGTTTC
TYMS	TATGTATAAAGTCACCTGGCTTCAG
POLR2領域1	TATTCATTGTCTTCACCAGGAAGCCCAC

10

20

30

40

POLR2領域2	TATGCACCCACCTGGTGAAGGGATGTAGG
ACTB領域1	TATCTTGCACATGCCGGAGCCGTTGTC
ACTB領域2	TATCGCAACTAAGTCATAGTCCGCCTAGAA
<p>第2段階：次の配列を96個の新しいカラムで部分的に合成されたオリゴの混合物に添加した： CCATCTCATCCCTGCGTGTCTCCGACTCAG [BC] CAA 独特な8-塩基バーコード ([BC] により表示される) を各カラムに対して使用した。バーコード配列は、表6に列挙する。</p>	

【0091】

[表8]

放射線反応性mRNA標的の逆転写のためのモジュール式プライマー

第1段階：部分的に合成されたオリゴヌクレオチド3'セグメント

標的遺伝子	配列
DHCR24	TATCATACACTGCCACACCCATCTCCAC
GINS2	TATCTGAATCCATTCCTTGCACCATCACA C
DB160230	TATATTCCGTCTCTATCGTTGTGTGAACG G
PADI4	TATAGTAGAGCTTTGACTGGTGGAGTCTT G
PLXNA2	TATCCCATCACTTGTCCCTTCCATCTGAGA C
GPR84	TATGATGGAGAGACTTGATTTGGGAGGCT G
AK024898	TATAAAGGCCTTAAGTCACTCCCAAAGCA
RPS27L	TATAAATGAACACCCTTCTGTGAGTCTGG C
CDKN1A (V a e 2)	TATTGGAGTGGTAGAAATCTGTTCATGCTG G
GADD45A	TATACCCATTGATCCATGTAGCGACTTTC C
TMEM30A	TATACAGCCTGAGTATTTCCAAAGCTGAA GT
AK074467	TATGAAACTAAGGTAAATAAGTGCCTGGG TTGG
TRIAP1	TATACACTGAGAGCTCTGAAATGGTGTTC A
ISG20L1	TATGATATCCGTGAACCTTGCTGCTGTGC
CDKN1A (V a r 1)	TATTCCATAGCCTCTACTGCCACCATCTT A
PLK2	TATATGGCCATAGTTCACAGTTACGCAGC
MGC5370	TATTTAGGAAACCTCTGCCATGTCTGCAT C
LOC392454	TATCCTTCTTCATCCTCCATCTTGGGAGC
BE646426	TATGGATATCCGTGTGTCTTGTCTGTGGC
PCNA	TATTATGCTGGCATCTTAGAAGCAGTTCT C
PHLDA3	TATAAAGGACCTAAGCAGCAGGAGACC
DDB2	TATAGATAACCTTGGTTCCAGGCTAGATA CAG
FDXR	TATTTCCAGCATGTTCCCAACCTGGTGAC
ACTB	TATCGCAACTAAGTCATAGTCCGCCTAGA A
GAPDH	TATGCCTGCTTCACCACCTTCTTGATG

10

20

30

40

第2段階：次の配列を108新しいカラムで部分的に合成されたオリゴの混合物に添加した：CATCTCATCCCTGCGTGTCTCCGACTCAG [BC] CAA
独特な8-塩基バーコード ([BC] により表示される) を各カラムに対して使用した。バーコード配列は、表6に列挙する。

【0092】

合成RNAサンプルの製造：

下記の表2に示されたように、90個のマイクロRNAおよび6個の対照RNA配列で構成されたRNAオリゴヌクレオチドをYale Keckオリゴヌクレオチド合成コア設備で40nMol規模で合成して精製し、この際、2'-脱保護された。Tecan Freedom Evo 200ロボットリキッドハンドラーを、事前規定された量の各RNAを96ウェルプレートのウェルに分配するようにプログラミングし、ヒートマップ(heat map)上のバラの画像が製造されるようにデザインされたパターンで4~0.08nM範囲の最終濃度を達成した。RNAをRNAseフリーの水に10mMのTris(pH 7.6)、0.1mMのEDTA、および300ng/mLのキャリアRNA(Qiagen)を含有している緩衝液に溶かした。合成RNA溶液をRTに必要な時まで-80で保管した。

10

【0093】

[表2]

合成RNAオリゴヌクレオチド

合成RNA	配列	合成RNA	配列
対照標準A	GUUACUUAU GAGAGUGGC UAGCU	miR-199a-5p	ACAGUAGU CUGCACAU UGGUUA
対照標準B	UGAUCAUUAU CCUGUGCAC UCAU	miR-19a	UGUGCAAA UCUAUGCA AAACUGA
cel-mir-2	UAUCACAGC CAGCUUUGA UGUGC	miR-200b	UAAUACUG CCUGGUAA UGAUGA
RNU44	GAAGGUCUU AAUUAGCUC UAACUGACU	miR-203	GUGAAAUG UUUAGGAC CACUAG
RNU48	CCAUCACCG CAGCGCUCU GACC	miR-204	UUCCCUUU GUCAUCCU AUGCCU
U6	GGCAUCUCG AGCUAAUCU GGUGGG	miR-205	UCCUUCAU UCCACCGG AGUCUG
let-7a	UGAGGUAGU AGGUUGUAU AGUU	miR-21	UAGCUUAU CAGACUGA UGUUGA
let-7b	UGAGGUAGU AGGUUGUGU GGUU	miR-210	CUGUGCGU GUGACAGC GGCUGA
miR-100	AACCCGUAG AUCCGAACU UGUG	miR-216b	AAAUCUCU GCAGGCAA AUGUGA
miR-103	AGCAGCAUU GUACAGGGC UAUGA	miR-218	UUGUGCUU GAUCUAAC CAUGU
miR-105	UCAA AUGCU CAGACUCCU GUGGU	miR-22	AAGCUGCC AGUUGAAG AACUGU
miR-106a	AAAAGUGCU UACAGUGCA GGUAG	miR-221	AGCUACAU UGUCUGCU GGGUUUC
miR-10b	UACCCUGUA GAACCGAAU UUGUG	miR-222	AGCUACAU CUGGCUAC UGGGU
miR-125a-5p	UCCUGAGA CCCUUUAAC CUGUGA	miR-224	CAAGUCAC UAGUGGUU CCGUU
miR-125b	UCCUGAGA CCCUAACUU GUGA	miR-23a	AUCACAUU GCCAGGGA UUUCC

10

20

30

40

m i R-1 2 6	UCGUACCGU GAGUAAUAA UGCG	m i R-2 4	UGGCUCAG UUCAGCAG GAACAG	
m i R-1 2 8	UCACAGUGA ACCGGUCUC UUU	m i R-2 5	CAUUGCAC UUGUCUCG GUCUGA	
m i R-1 2 9-5 p	CUUUUUGCG GUCUGGGCU UGC	m i R-2 6 a	UUCAAGUA AUCCAGGA UAGGCU	
m i R-1 3 0 a	CAGUGCAAU GUUAAAAGG GCAU	m i R-2 7 b	UUCACAGU GGCUAAGU UCUGC	10
m i R-1 3 2	UAACAGUCU ACAGCCAUG GUCG	m i R-2 8-5 p	AAGGAGCU CACAGUCU AUUGAG	
m i R-1 3 3 a	UUUGGUCCC CUUCAACCA GCUG	m i R-2 9 9-5 p	UGGUUUAC CGUCCCAC AUACAU	
m i R-1 3 4	UGUGACUGG UUGACCAGA GGGG	m i R-2 9 c	UAGCACCA UUUGAAAU CGGUUA	20
m i R-1 3 5 b	UAUGGCUUU UCAUCCUA UGUGA	m i R-3 0 1 a	CAGUGCAA UAGUAUUG UCAAAGC	
m i R-1 3 6	ACUCCAUUU GUUUUGAUG AUGGA	m i R-3 0 2 a	UAAGUGCU UCCAUGUU UUGGUGA	
m i R-1 3 7	UUAUUGCUU AAGAAUACG CGUAG	m i R-3 0 b	UGUAAACA UCCUACAC UCAGCU	30
m i R-1 4 0-5 p	CAGUGGUUU UACCCUAUG GUAG	m i R-3 1	AGGCAAGA UGCUGGCA UAGCU	
m i R-1 4 1	UAACACUGU CUGGUAAAG AUGG	m i R-3 2	UAUUGCAC AUUACUAA GUUGCA	
m i R-1 4 2-3 p	UGUAGUGUU UCCUACUUU AUGGA	m i R-3 4 2-3 p	UCUCACAC AGAAAUCG CACCCGU	40
m i R-1 4 3	UGAGAUGAA GCACUGUAG CUC	m i R-3 4 a	UGGCAGUG UCUUAGCU GGUUGU	
m i R-1 4 5	GUCCAGUUU UCCCAGGAA UCCCU	m i R-3 4 c-5 p	AGGCAGUG UAGUUAGC UGAUUGC	
m i R-1 4 6 a	UGAGAACUG AAUCCAUG GGUU	m i R-3 6 3	AAUUGCAC GGUAUCCA UCUGUA	

m i R-148 a	UCAGUGCAC UACAGAACU UUGU	m i R-365	UAAUGCCC CUAAAAAU CCUUAU	
m i R-149	UCUGGCUCC GUGUCUUCA CUCCC	m i R-367	AAUUGCAC UUUAGCAA UGGUGA	
m i R-150	UCUCCCAAC CCUUGUACC AGUG	m i R-372	AAAGUGCU GCGACAUU UGAGCGU	
m i R-154	UAGGUUAUC CGUGUUGCC UUCG	m i R-373	GAAGUGCU UCGAUUUU GGGGUGU	10
m i R-16	UAGCAGCAC GUAAAUUU GGCG	m i R-375	UUUGUUCG UUCGGCUC GCGUGA	
m i R-181 a	AACAUUCA CGCUGUCGG UGAGU	m i R-381	UAUACAAG GGCAAGCU CUCUGU	
m i R-182	UUUGGCAAU GGUAGAACU CACACU	m i R-382	GAAGUUGU UCGUGGUG GAUUCG	20
m i R-183	UAUGGCACU GGUAGAAUU CACU	m i R-383	AGAUCAGA AGGUGAUU GUGGCU	
m i R-184	UGGACGGAG AACUGAUAA GGGU	m i R-422 a	ACUGGACU UAGGGUCA GAAGGC	
m i R-185	UGGAGAGAA AGGCAGUUC CUGA	m i R-424	CAGCAGCA AUUCAUGU UUUGAA	30
m i R-186	CAAAGAAUU CUCCUUUUG GGCU	m i R-488	UUGAAAGG CUAUUUCU UGGUC	
m i R-18 a	UAAGGUGCA UCUAGUGCA GAUAG	m i R-7	UGGAAGAC UAGUGAUU UUGUUGU	
m i R-191	CAACGGAAU CCCAAAGC AGCUG	m i R-9	UCUUUGGU UAUCUAGC UGUAUGA	40
m i R-192	CUGACCUAU GAAUUGACA GCC	m i R-92 a	UAUUGCAC UUGUCCCG GCCUGU	
m i R-194	UGUAACAGC AACUCCAUG UGGA	m i R-93	CAAAGUGC UGUUCGUG CAGGUAG	
m i R-196 b	UAGGUAGUU UCCUGUUGU UGGG	m i R-95	UUCAACGG GUAUUUAU UGAGCA	

miR-199a-3p	CCCAGUGUU CAGACUACC UGUUC	miR-96	UUUGGCAC UAGCACAU UUUUGCU
-------------	---------------------------------	--------	---------------------------------

【0094】

組織および細胞株RNAサンプル：

第1選択型ヒト全RNA調査パネル(Ambion)を20個の正常ヒト組織由来のトータルRNAの供給源として使用した。MAQCリファレンスサンプルは、StratageneユニバーサルヒトリファレンスRNA(10個のヒト細胞株からのトータルRNAで構成される)、およびAmbion第1選択型ヒト脳リファレンスRNAで構成された。

10

【0095】

放射線照射された血液サンプルからのRNA：

エール大学の人体調査委員会(Human Investigation Committee)の承認下に18人の健康な支援者からの書面同意を得た後、末梢血をクエン酸ナトリウムを含有するチューブに収集した。血液を2mLの部分標本に分け、採血した後、1時間以内に分当たり1.79Gyの容量速度で0、0.1、0.5、2、4または8GyのX放射線を照射した。次に、血液を10%ウシ胎仔血清を含有する等容量のRPMI1640培地を添加した後、24時間37℃でインキュベーションした。末梢血単核細胞をフィコール勾配遠心分離を使用して分離し、全RNAをRNeasyミニキット(Qiagen)を使用してこれら細胞から製造した。

20

【0096】

miRNAサンプルのプロセッシング：

方法の第1段階で、多重RNA標的を各サンプルに対して単一チューブで逆転写させた。所定のサンプルに対して使用されたRTプライマーミックスは、表3に示されたように、5'セグメントにサンプル特異的タグ、および3'セグメントに一貫した割合の多重標的的特異的プライマー配列を有する。プライマーを短いmiRNA(および対照標準RNA)標的の3'末末端で6個のヌクレオチドと混成化するようにデザインした。5'-ピオチン標識オリゴヌクレオチドを隣接する相補的共通プライマー配列にアニーリングし、積層(stackings)塩基を延長させることによって、短いRNA/プライマー異種二重螺旋を安定化させた。

30

【0097】

各逆転写カクテルは、RNAseフリーの水に5μMのタグ付きプライマーミックス(～50nMの各標的的特異的プライマー)、7.5μMのピオチン標識オリゴヌクレオチド、1xRT緩衝液、3mMのMgCl₂、250μMの各dNTP、5mMのジチオトレイトール(DTT)、30ng/μLの担体RNA(Qiagen)、鋳型RNA、および5ユニット/μLのMultiscribe逆転写酵素(Life Technologies)で構成された。各RTを10μLの最終体積で行った。鋳型RNA、DTTおよび逆転写酵素を添加する前に、カクテルを95℃で2分間加熱した後、室温に冷却させることによって、ピオチン標識オリゴヌクレオチドをプライマーミックスにアニーリングした。最後の組み立てられたRTカクテルに対して16℃で2分、42℃で1分、および50℃で1秒の40サイクルを行った。反応を、65℃で20分間加熱し、EDTAを10mMの最終濃度で添加することによって停止させた。次に、すべての別途のRT反応の生成物を単一体積で組み合わせた。

40

【0098】

プールcDNAを、高容量のストレプトアビジンがコーティングされたアガロース樹脂(Thermo Scientific)(10μLのRT反応当たり添加された5μLk樹脂スラリー)を使用して相補的ピオチン標識されたオリゴヌクレオチドの捕捉により精製した。室温で少なくとも2時間チューブを徐々にくるくる回転させることによって、樹脂粒子を溶液中に懸濁された状態を維持させて、ピオチン結合を促進した。次に、粒子

50

を 10 mM の Tris pH 7.6 および 50 mM の NaCl を含有する緩衝液で洗浄した。cDNA を 95 で 2 分間の熱変性させることにより樹脂結合オリゴから新しい体積の同一緩衝液（樹脂スラリーの体積の 2 倍）に放出された。延長されていない RT プライマーを除去するために、下記の表 4 に示されたように、プライマー延長配列（各 100 nM）に相補的なビオチン標識オリゴヌクレオチドの混合物を使用して第 2 ラウンドの選択的アニーリング、捕捉、洗浄および溶出を行った。

【 0 0 9 9 】

[表 4]

miRNA標的に対するハイブリッド捕捉オリゴヌクレオチドおよびPCRプライマー

ハイブリッド捕捉オリゴマーに対して：X=5' -ビオチン

標的的特異的PCRプライマーに対して：X=CCTCTCTATGGGCAGTCGGTGAT

ユニバーサルPCRプライマー：CCATCTCATCCCTGCGTGTCTCCGACT

標的	オリゴヌクレオチド配列	標的	オリゴヌクレオチド配列
対照標準A	X-GTTACT TATGAGAG TGGCTAG	miR-199a-5p	X-ACAGTA GTCTGCAC ATTGG
対照標準B	X-TGATCA TATCCTGT GCACT	miR-19a	X-TGTGCA AATCTATG CAAACCT
cel-mir-2	X-TATCAC AGCCAGCT TTGATG	miR-200b	X-TAATAC TGCCTGGT AATGATG
RNU44	X-GAAGGT CTTAATTA GCTCTAAC TG	miR-203	X-GTGAAA TGTTTAGG ACCACTA
RNU48	X-TCACCG CAGCGCTC TGA	miR-204	X-TTCCCT TTGTTCATC CTATGC
U6	X-GGCATC TCGAGCTA ATCT	miR-205	X-TCCTTC ATTCCACC GGAG
let-7a	X-TGAGGT AGTAGGTT GTATAG	miR-21	X-TAGCTT ATCAGACT GATGTTG
let-7b	X-TGAGGT AGTAGGTT GTGT	miR-210	X-CTGTGC GTGTGACA GCG
miR-100	X-AACCCG TAGATCCG AACTT	miR-216b	X-AAATCT CTGCAGGC AAATG
miR-103	X-AGCAGC ATTGTACA GGGCT	miR-218	X-TTGTGC TTGATCTA ACCAT
miR-105	X-TCAAAT GCTCAGAC TCCTG	miR-22	X-AAGCTG CCAGTTGA AGAAC
miR-106a	X-AAAAGT GCTTACAG TGCAGG	miR-221	X-AGCTAC ATTGTCTG CTGG
miR-10b	X-TACCCT GTAGAACC GAATTTG	miR-222	X-AGCTAC ATCTGGCT ACTG

10

20

30

40

m i R-1 2 5 a-5 p	X-TCCCTG AGACCCTT TAACC	m i R-2 2 4	X-CAAGTC ACTAGTGG TTCC	
m i R-1 2 5 b	X-TCCCTG AGACCCTA ACTTG	m i R-2 3 a	X-ATCACA TTGCCAGG GATT	
m i R-1 2 6	X-TCGTAC CGTGAGTA ATAATGC	m i R-2 4	X-TGGCTC AGTTCAGC AGGA	
m i R-1 2 8	X-TCACAG TGAACCGG TCTC	m i R-2 5	X-CATTGC ACTTGTCT CGGTC	10
m i R-1 2 9-5 p	X-CTTTTT GCGGTCTG GGC	m i R-2 6 a	X-TTCAAG TAATCCAG GATAGG	
m i R-1 3 0 a	X-CAGTGC AATGT TAA AAGGG	m i R-2 7 b	X-TTCACA GTGGCTAA GTTCT	
m i R-1 3 2	X-TAACAG TCTACAGC CATGG	m i R-2 8-5 p	X-AAGGAG CTCACAGT CTATTG	20
m i R-1 3 3 a	X-TTTGGT CCCCTTCA ACCA	m i R-2 9 9-5 p	X-TGGTTT ACCGTCCC ACATA	
m i R-1 3 4	X-TGTGAC TGGTTGAC CAGAG	m i R-2 9 c	X-TAGCAC CATTTGAA ATCGGT	
m i R-1 3 5 b	X-TATGGC TTTTCAATT CCTATGT	m i R-3 0 1 a	X-CAGTGC AATAGTAT TGTC AAAG	30
m i R-1 3 6	X-ACTCCA TTTGTTTT GATGATG	m i R-3 0 2 a	X-TAAGTG CTTCCATG TTTTGG	
m i R-1 3 7	X-TTATTG CTTAAGAA TACGCGT	m i R-3 0 b	X-TGTAAA CATCCTAC ACTCAGC	
m i R-1 4 0-5 p	X-CAGTGG TTTTACCC TATGGT	m i R-3 1	X-AGGCAA GATGCTGG CATA	
m i R-1 4 1	X-TAACAC TGTCTGGT AAAGATG	m i R-3 2	X-TATTGC ACATTACT AAGTTGC	40
m i R-1 4 2-3 p	X-TGTAGT GTTTCCTA CTTTATGG	m i R-3 4 2-3 p	X-TCTCAC ACAGAAAT CGCAC	
m i R-1 4 3	X-TGAGAT GAAGCACT GTAGC	m i R-3 4 a	X-TGGCAG TGTCTTAG CTGG	

m i R-145	X-GTCCAG TTTTCCCA GGAATC	m i R-34c-5p	X-AGGCAG TG TAGTTA GCTGAT	
m i R-146a	X-TGAGAA CTGAATTC CATGG	m i R-363	X-AATTGC ACGGTATC CATCT	
m i R-148a	X-TCAGTG CACTACAG AACTT	m i R-365	X-TAATGC CCCTAAAA ATCCTTA	
m i R-149	X-TCTGGC TCCGTGTC TTCA	m i R-367	X-AATTGC ACTTTAGC AATGGT	10
m i R-150	X-TCTCCC AACCCCTG TACCA	m i R-372	X-AAAGTG CTGCGACA TTTGAG	
m i R-154	X-TAGGTT ATCCGTGT TGCCT	m i R-373	X-GAAGTG CTTCGATT TTGGG	
m i R-16	X-TAGCAG CACGTAAA TATTGG	m i R-375	X-TTTGTT CGTTCGGC TCG	20
m i R-181a	X-AACATT CAACGCTG TCGG	m i R-381	X-TATACA AGGGCAAG CTCTC	
m i R-182	X-TTTGGC AATGGTAG AACTCAC	m i R-382	X-GAAGTT GTTCGTGG TGGAT	
m i R-183	X-TATGGC ACTGGTAG AATTCA	m i R-383	X-AGATCA GAAGGTGA TTGTGG	30
m i R-184	X-TGGACG GAGAACTG ATAAG	m i R-422a	X-ACTGGA CTTAGGGT CAGAA	
m i R-185	X-TGGAGA GAAAGGCA GTTCC	m i R-424	X-CAGCAG CAATTCAT GTTTTG	
m i R-186	X-CAAAGA ATTCTCCT TTTGGG	m i R-488	X-TTGAAA GGCTATTT CTTGGT	
m i R-18a	X-TAAGGT GCATCTAG TGCAGA	m i R-7	X-TGGAAG ACTAGTGA TTTTGTT	40
m i R-191	X-CAACGG AATCCCAA AAGCA	m i R-9	X-TCTTTG GTTATCTA GCTGTAT	
m i R-192	X-CTGACC TATGAATT GACAGC	m i R-92a	X-TATTGC ACTTGTCC CGGC	

miR-194	X-TGTAAC AGCAACTC CATGT	miR-93	X-CAAAGT GCTGTTCTG TGCA
miR-196b	X-TAGGTA GTTTCCTG TTGTTG	miR-95	X-TTCAAC GGGTATTT ATTGAG
miR-199a-3p	X-CCCAGT GTTTCAGAC TACCT	miR-96	X-TTTGGC ACTAGCAC ATTTTTG

10

【0100】

精製されたcDNAプールを各cDNA標的のシングルプレックスエンドポイントPCRのために96個の別途のチューブに分配した。所定の標的に関連したすべてのサンプル特異的タグが単一反応体積で競合競合増幅を進行したため、タグ比率が維持された。各PCRに使用されたプライマーペアは、図1b(表4)に示されたように、ユニバーサル順方向プライマーと区別される標的特異的逆方向プライマーで構成された。シーケンシングアダプタをプライマーの5'末端に統合させて、OCR生成物の直接的なシーケンシングを可能にした。各PCRカクテルは、10 μ L体積の1x AccuPrime PCR緩衝液I(dNTPsおよびMgCl₂を含む)、100nMのユニバーサル順方向プライマー、100nMの標的特異的逆方向プライマー、2 μ LのプールcDNA鑄型および0.2 μ LのAccuPrime Taq DNAポリメラーゼ(Invitrogen)で構成された。ミネラルオイルを添加し、蒸発を最小限に抑えた。熱サイクリング媒介変数は、2分間94 $^{\circ}$ C、30秒間60 $^{\circ}$ C、20秒間72 $^{\circ}$ C、次いで20秒間94 $^{\circ}$ C、30秒間65 $^{\circ}$ Cおよび20秒間72 $^{\circ}$ Cの40サイクルであった。最終延長段階を72 $^{\circ}$ Cで2分間行い、引き続き、4 $^{\circ}$ Cに冷却させた後、EDTA(最終10mM)を添加し、ポリメラーゼ活性を停止させた。

20

【0101】

すべてのPCR体積を組み合わせて、20 μ Lの部分標本のプール反応生成物を2%低融点アガロースゲルで精製した。DNAを切除したゲルスライスからQIAquickゲル抽出キット(Qiagen)を使用して抽出した。濃度をBioanalyzer 2100(Agilent)を使用して推定し、Ion TorrentエマルジョンPCRに対して推奨された水準に調整した。

30

【0102】

mRNAサンプルのプロセッシング:

mRNAサンプルのプロセッシングのための全体計画は、miRNAサンプルに対して前記記述されたものと同一であり、ただいくつかの顕著な変形があった。mRNAがmiRNAよりはるかに大きかったため、プライマーを、約100ヌクレオチド標的領域を増幅するようにデザインすることができた。したがって、さらに長い遺伝子特異的RTプライマーを使用した(表5および表8)。これは、延長した塩基積層を通じて安定性を増大させるために、相補的ピオチン化オリゴヌクレオチドを必要とすることなく、熱安定性ポリメラーゼを使用して高温でRTが行われることを可能にした。各RT反応は、RNaseフリーの水にタグ付きプライマーミックス(各~50nMの標的特異的プライマー)、1x第1鎖の緩衝液、500 μ Mの各dNTP、5mMのDTT、鑄型RNA、および10ユニット/ μ LのSuperScriptIII逆転写酵素(Invitrogen)で構成された10 μ L体積で行った。プライマーを室温で組み合わせた後、55 $^{\circ}$ Cで1時間インキュベーションするときに添加された緩衝液、DTTおよびポリメラーゼなしに、5分間65 $^{\circ}$ Cに加熱することによって、RNA標的にアニーリングした。反応物を、75 $^{\circ}$ Cで20分間、95 $^{\circ}$ Cで1分間加熱し、EDTA(最終10mM)を添加することによって、ポリメラーゼを不活性化させた後にプールした。

40

【0103】

50

R T中のビオチン標識オリゴヌクレオチドの不在は、プライマー延長された配列に相補的であるビオチン化オリゴヌクレオチドを使用する単一段階でc D N Aの捕捉を可能にした(表7および表9)。プールされ、精製されたc D N A鑄型を表7および表9で列挙したプライマーを使用する各標的のシングルプレックスエンドポイントP C Rのために別途のチューブに分配した。熱サイクリング媒介変数は、前記m i R N Aに対して記述したものと同一であるが、第1サイクルに対して60 の代わりに63 のアニーリング温度を使用した。

【 0 1 0 4 】

[表7]

MAQC mRNA標的に対するハイブリッド捕捉オリゴヌクレオチドおよびPCRプライマー

ハイブリッド捕捉オリゴに対して: X=5'-biotin

標的的特異的PCRプライマーに対して: X=CCTCTCTATGGGCAGTCGGTGAT

ユニバーサルPCRプライマー配列: CCATCTCATCCCTGCGTGTCTCCGACT

標的	オリゴヌクレオチド配列
ANXA5	X-ACCATTGACCGCGAGACTTCTGGCAA
B2M	X-TGCCTGCCGTGTGAACCATGTGACT
CDK9	X-GCAGCACCAACTCGCCCTCATCAGT
CDKN1A	X-TCTTGTACCCTTGTGCCTCGCTCAGG
CTPS	X-CCTATAGTGACAGGAGTGGAAGCAGCTC
CYP1B1	X-ATCACTGACATCTTCGGCGCCAGCC
CYP2D6	X-TGACATCGAAGTACAGGGCTTCCGCA
DAD1	X-TGCCAGCACCATCCTGCACCTTGTT
ELAVL1	X-GGATCATCAACTCGCGGGTCTCTCGT
FANCG	X-TAGCCAGCGGCCAGGATACCAAAGC
FGF9	X-AGCATTCGAGGCGTGGACAGTGGAC
ICAM1	X-TCTCCTGCTCTGCAACCCTGGAGGT
IGF2R	X-GCCTGCTGGCCCTGTTGCTCTACAA
IGFBP2	X-AGCACCTCTACTCCCTGCACATCCC
IGFBP5	X-ACCTGCCCAATTGTGACCGCAAAGGA
IL1B	X-AAGCTCTCCACCTCCAGGGACAGGA
IL8	X-TGGACCCCAAGGAAAAGTGGGTGCAG
KIT	X-AAGCAGCCCCTATCCTGGAATGCCG
MAP2K6	X-CCATCGCCACAACCTCCCAGCAGACA
MAP3K14	X-ACATCCGGGAGTTCCACCGGGTCAA
MX2	X-CCAGCAAGCTTTTCATTAACGTGGCCAAA
MYB	X-AAGCTCCGTTTTTAATGGCACCCAGCA
RARA	X-TGCTGCCCTGGAGATGGATGATGC
RB1	X-TCTCCCAGGAGAGTCCAAATTTTCAGCA
SLC15A2	X-ATGCCCTGGTTACAGCTGGGGAGGT
SLC2A1	X-GGCATGGCGGGTTGTGCCATACTCA
SOD1	X-GGAGACCATTGCATCATTGGCCGCACA
TGFBR2	X-TCGAGGGCGACCAGAAATTTCCAGC
TGM1	X-TGTGCTCTTCCGGCTCGAAGGCTCT

10

20

30

40

TYMS	X-GACATGGGCCTCGGTGTGCCTTTCA
POLR2F	X-GGACTGGCCGCTGCCAAACATGTGC
POLR2B	X-CAGCGGCTTCAGCCCAGGTTACTCCC
ACTBF	X-GAGCACAGAGCCTCGCCTTTGCCGAT
ACTBB	X-CCTCGCTGTCCACCTTCCAGCAGATGT

【 0 1 0 5 】

[表 9]

ハイブリッド捕捉オリゴに対して：X = 5' - ビオチン 標的的特異的PCRプライマーに対して：X = CCTCTCTATGGGCAGTCGGTGAT ユニバーサルPCRプライマー配列：CCATCTCATCCCTGCGTGTCTCCGACT	
標的	オリゴヌクレオチド配列
DHCR24	X-AGCCAGTTTCTTTGGCCAGAAGGATGA
GJNS2	X-CGCTCAGGACGTGATGAGGTA CT CGTG G
DB160230	X-ACAGCAAAGCAGCCACCAAGATGGACC
PADI4	X-ACCCTGACGATGAAAGTGGCCAGTGGT
PLXNA2	X-GAGCTGAGAGGAGGAGCCTCGCATTC
GPR84	X-GGACTGTCTCCTCCAGGACCAAAGTGG C
AK024898	X-TACATAAGGGTGGCATGCCCACTGGCT
RPS27I	X-TGTCCAGGTTGCTACAAGATCACACG GTT
CDKN1A (Var2)	X-TGTCTTGTACCCTTGTGCCTCGCTCAG G
GADD45A	X-CTGCACTGCGTGCTGGTGACGAATCCA
TMEM30A	X-TGCTTGGCAGACCTTCATCTTCTGCCT CA
AK074467	X-AATGGAGTGAGCCCTGGATTGGGAGC
TRIAP1	X-ACACAGTTCCCTGCCTTCACAAGAGGT GT
ISG20L1	X-CTGAGCAGCTGTGGCCCAGACAGA ACT
CDKN1A (Var1)	X-ATCGTCCAGCGACCTTCCTCATCCACC
PLK2	X-GGGAGAAGGGAGGAAGCTCCCATGTTG T
MGC5370	X-GTTGCAGGCAAAGGAACGCAGCTGGAA
LOC392454	X-AGAGGAGGAAGCTGTTACCATGGAGAT GA
BE646426	X-TTGCTCTGTTGTCACCTCCCGCACAGT
PCNA	X-AAGCCACTCCACTCTCTTCAACGGTGA
PHLDA3	X-CCACTCCAGAATGGCCTCTGGACTCAC C
DDB2	X-TGCTCTGGACTTGCCTCCAGAGACTGC
FDXR	X-CAGGTGGAGGTGTGGGCCGATCTAACC
ACTB	X-CCTCGCTGTCCACCTTCCAGCAGATGT
GAPDH	X-AACGTGTCAGTGGTGGACCTGACCTGC

10

20

30

40

50

【0106】

次世代シーケンシング：

鋳型を Ion Torrent シーケンシングのために自動 Ion OneTouch システム (Life Technologies) を使用して製造した。ゲル精製されたアンプリコンを機器にローディングする前に、製造業者により推奨された濃度に希釈した。自動エマルジョンPCRは、イオンスフィア粒子 (ISP) に対して大量並列クローン増幅を可能にした。ポリクローナルISPを最小化するために、鋳型希釈を10%~3

0% 鋳型 - ポジティブ I S P をなすように調整した。ワンタッチエンリッチメントシステムを使用して鋳型 - ポジティブ I S P を分離し、それをシーケンシングのために半導体チップ上にローディングした。所望の配列深さに応じて、314 低容量チップまたは 318 高容量チップのうち一つを使用した。シーケンシングを Ion Torrent PGM (Life Technologies) で 200bp 試薬キットを使用して行った。

【0107】

配列のビニング (binning) およびカウンティング :

各標的 / バーコードピンに属する読み込み数を測定するために、Torrent Suite ソフトウェア (バージョン 4.0) の一部として提供された Torrent Mapping Alignment Program (TMAP) を使用した。所定のデータセットの分析のために 3 ファイルのアップロードが必要であった : ユーザ規定されたバーコードおよびアダプタ配列を含有するテキストファイル、miRNA または mRNA 基準配列を列挙する FASTA フォーマットファイル、および標的領域を規定する BED ファイル。読み取りを標的基準配列と整列させた後、包含範囲分析プラグ - インモジュールを作動させ、その結果のバーコード / アンプリコン包含範囲マトリックスをダウンロードした。このマトリックスは、各ピンに対する読み込みカウントを含有し、マイクロソフトエクセルで開かれ、さらに操作され得た。

10

【0108】

配列データのダウンサンプリングが Torrent Suite ソフトウェアでは可能ではなかったため、図 2e に対する読み込みの規定された下位セットからビニングされたカウントを得るために、代替案を使用した。この目的のために、マイクロソフトエクセルで「countifs」機能を探索した。TMAP 分析と比較したこの接近法との重要な違いは、完全な配列マッチのみがカウントされたということであった。したがって、シーケンサーのエラーによる不完全一致の可能性を最小限に抑えるために、約 10 ~ 12ヌクレオチドの短い基準配列を使用した。基準配列は、疑似 PCR 生成物 (例えば、プライマー二量体) の計数を回避するために、任意の単一プライマーに含まれた配列を超えて延長させるために選択した。各基準配列が単に単一標的にのみ一致することを保証するために、さらに注意した。

20

【0109】

ビニングされた配列カウントの正常化および標準化 :

図 2a および図 2b のローズの画像を示すヒートマップを生成するために、二つの複製実験からの計数を 9, 216 個のデータピンのそれぞれに対して平均化した。次いで、カウントを分配された合成 RNA の公知の総量に対して列と行にわたって正常化した。まず、所定の列のカウントをその列に分配された RNA の総量に対するカウントの和の比で乗算した。二番目に、所定の行のその結果の値をその行に分配された RNA の総量に対する値の和の比率で乗算した。最後に、これら正常化された値の二進対数を計算し、ヒートマップに図表化した。

30

【0110】

ヒト組織および血液サンプルからの miRNA および mRNA 測定の正常化および標準化 (図 3a、図 3b、図 5) を下記の説明の通り行った。まず、複製値を各データピンに対して平均化した。二番目に、各標的に対する異なるシングルプレックス PCR によって生成された総カウントを同等にするために、所定の列にわたる値は、その列の値の合計を 1000 にするために共通因子で乗算した。三番目に、データのフローリング (flooring) をすべてのピンに対して 0.01 づつ足すことによって (それによって 0 値を除去する) 達成した。これは、35 より大きい Cq 値を 35 に変換させる qRT - PCR 実験の通常の方法と類似していた。四番目に、miRNA 水準を正常化するために、所定のサンプルのすべての miRNA DP に対する平均発現値を正常化因子として使用した。放射線照射された血液サンプルからの mRNA を二つのハウスキーピング遺伝子、ACTB および GAPDH の平均発現値に対して正常化した。五番目に、 \log_{10} (倍数 - 変化) 値

40

50

をすべてのデータピンに対して計算した。六番目に、平均センタリングを各値から列平均を差し引くことによって行った。最後に、各値を列にわたり標準偏差で割ることによって各値を自動縮尺した。

【0111】

正常なヒト組織のmiRNAの絶対量を測定するために(図6)、各合成miRNAの約15,000の複製物を含む定量参照標準サンプルを逆転写させ、50ngの組織由来の全RNAサンプルと競合競合的に増幅させた。すべてのサンプルを3回の技術的複製で分析した。読み込みカウントを複製物に対して平均化した。所定の組織サンプルの標的に対する平均カウントを対照サンプルの同一標的に対する平均カウントで割った。次いで、その結果値に15,000を乗算してその組織サンプルの50ngのトータルRNA

10

【0112】

ヒートマップの図表化:

すべてのヒートマップをクラスタリングすることなく、TreeViewソフトウェア(ウェブサイト: <http://rana.lbl.gov/EisenSoftware.htm>からダウンロード)を使用して作成した。公開されたqRT-PCR研究からの生(Raw)Cq値をmiRNAボディマップウェブサイト(www.mirnabodymap.org)から得た。その値を35でフローリングし、前記において概略記述したのと同じ正常化および標準化段階を第4段階から開始して行った。公開されたお

20

【0113】

MAQCサンプルのmRNAの分析:

mRNA分析に対する標的遺伝子を、通常的にMAQCデータセットで報告された三個のすべての定量(非微細配列)プラットフォームにわたって試験された48個の遺伝子の中から選択した。これら48個の遺伝子の中から、発現が3個のプラットフォームにわたって一貫した水準(having a low coefficient of variance)で測定された30個を選択した。標的化された遺伝子は、表5に列挙されている。

30

【0114】

4回繰り返した実験からのピニングされた配列カウントを4個のMAQCサンプル(A、B、CおよびD)のそれぞれに対して平均化した。所定の遺伝子に対する平均カウントをその遺伝子に対する値の和を作るために、共通因子で乗算し、1000にした。フローリングは適用しなかった。ただし、30個のみの標的を分析したため、サンプルにわたる全般的な平均発現水準に対する標準化は推奨されなかった。したがって、所定のサンプルに対する発現値をそのサンプルに対するPOLR2およびACTB参照遺伝子の平均測定に対して正常化した。

40

【0115】

正常化した発現値を使用してヒトのユニバーサルリファレンスRNA(サンプルA)とヒト以外のリファレンスRNA(サンプルB)間の30遺伝子すべてに対する倍数-変化を計算した。相対的な正確度をサンプルCおよびDの測定に基づいて、メインテキストで記述された通りに計算した。

【0116】

結果

合成RNA混合物を使用した正確度評価:

開示されたRNAプロファイリング方法の性能を、まず公知量の合成miRNAの混合物に対して試した。90種類のヒトmiRNAの代表的なパネルをmiRBase登録所

50

から選択したが、優先的に、先に発見され、より良く規定された生物学的機能を有するものを選択した。さらに6つのRNAを対照標準として含ませた：3個のヒトの小核/核であるRNA断片、C.エレガンス(C. elegans) miRNAおよび自然で発見されていない2個の任意の配列(表2)。これら合成RNAオリゴヌクレオチドをそれぞれロボットリキッドハンドラーを使用して96個の別個のチューブに、300 ng/mLポリ-A担体RNAをベースに4~0.08 nMの範囲の最終濃度になるように異なる量で分配した。RNAを方法の容量および正確度の複合化を簡単に目視で評価するために、デザインされたパターンで分布させた；定量し、ヒートマップに図表化した時、RNA混合物は、ローズ(rose)の映像を再現する。

【0117】

jサンプルからの逆転写中にiRNAの複合化された標的化された標識化を可能にするために、サンプル特異的タグに付着した標的特異的配列のi×j組み合わせを有するRTプライマーを生成する必要があった。さらに、定量一貫性を保証するために、同じ比率のすべての標的特異的配列を有する特有の標的化されたプライマーミックスを使用して異なるサンプルを逆転写させることが重要であった。個別的に形成された数千個のプライマーを単純に混合することが非実用的で且つ不正確な比率を算出するため、2段階モジュールオリゴヌクレオチド合成戦略を考案した(図1a)。合成を、それぞれが3'末末端で異なる標的特異的プライマー配列を含有する、96個の部分的なオリゴヌクレオチドを形成した後に停止した。部分的に合成されたオリゴヌクレオチドを含有するすべてのポリスチレン粒子を完全に混合し、96個の新しいカラムに分配した。その後、合成を再開し、特有なサンプル特異的タグおよびユニバーサルPCRプライマー結合部位を含んだ配列を各カラムに添加した。最後に、オリゴヌクレオチドを、全長生成物を豊富にするためにソリッド支持体から切断し、脱保護させ、カートリッジ精製した。この接近法で96個のプライマーミックスが生成されたが(表3)、それぞれ5'セグメントに特有のサンプル特異的タグおよび3'セグメントに96個の標的特異的プライマー配列の均一な組成を有した。一度形成された後、プライマーセットを数百もの反応に使用することができた。

【0118】

開示されたRNAプロファイリング方法の第1段階において、96個の標的化されたRNAがすべて各サンプルに対して単一ウェルで同時に逆転写された(図1b)。RTプライマーをそれぞれの短いmiRNA標的の3'末末端で6個のヌクレオチドに混成化するようにデザインした。標的特異的プライマー配列の比率がすべての反応において類似していると見なされるので、タグ付きcDNA複製物の比率は、それぞれのサンプルでRNAの存在量を忠実に反映しなければならない。RNA/DNA相互作用の特異性および安定性を増大させるために、RNAに結合しないプライマーの塩基を共通プライマー配列に相補的であるピオチン化オリゴヌクレオチドをアニーリングすることによってマスキングした；これは、また、塩基スタッキングの領域を延長させると予想される。RTが完了した時、96個のすべてのサンプルからのタグ付きcDNAを単一チューブにプールし、ストレプトアビジンアガロース樹脂を使用して混成化されたピオチン化オリゴヌクレオチドのプルダウン(pull-down)によって精製した。cDNAの熱溶出後に、第2ラウンドの選択的混成化および捕捉をプライマー延長した配列に相補的であるピオチン化オリゴヌクレオチドを使用して行った(表4)。

【0119】

次に、cDNAプールを別途のエンドポイントPCRによる各標的の増幅のために(安定期として採択する)96ウェルプレートのウェルに分配した。重要なことに、所定のcDNA種と関連したすべてのタグが単一体積で競合的に増幅したので、RNA豊富をコーディングするタグ比率が保存された。PCRプライマー(表4)の5'末端でシーケンシングアダプタの統合は、結果として生成されたアンプリコンがプールされ、ゲル精製され、追加のライブラリーの製造段階なしに大量並列シーケンシングに対する鋳型として直接使用されることを可能にした。

【0120】

10

20

30

40

50

96個すべての反応からプールされたアンプリコンを、Ion Torrent PGMで低容量314または高容量318チップのうちいずれか一つを使用してシーケンシングし、それぞれ作動当たり平均0.42Mまたは3.48Mの濾過された読み込みを生成した(表1)。読み込みをそれらの標的およびタグ配列を基準としてピニングした。Ion Torrent TMAP包含範囲分析モジュールを使用して、9,216個のすべてのピンに対する読み込みカウントの表を作成した。各チップの大きさについて、二つの複製実験の平均カウントを使用して各値の正常化およびログ変換後にヒートマップを作成した(方法にて詳細に説明した)。

【0121】

その結果の図表は、意図されたローズの画像(図2、aおよびb)を再現し、大多数のサンプルにわたって複雑な合成RNA混合物の正確かつ高度の並列定量を確認した。ピクセルシェードの微妙な差の一貫したレンダリングは、RNA量において相対的に小さい変化を区別する分析能力を証明した。低容量チップで生成された画像は、低光量で撮った写純粋なように、より不鮮明に見えるが、依然としてノイズ以上の強力な定量信号を示した。サンプルに添加された合成RNAの量とその測定水準の間の一貫性を評価するために、各RNAに対する平均に対する公知および測定された値の倍数変化に対する比較を行った(図2、cおよびd)。回帰分析の結果、318チップデータに対して0.82および0.88、314チップデータに対して0.89および0.84の傾きおよび R^2 をそれぞれ得た。次いで、測定の正確度に及ぼす配列深さの効果を知らるために、Pearson相関係数を、使用された読み込みの総数を異ならせながら公知および測定値の間で計算した(図2e)。この分析は、単に概略500,000の総読み込み(標的/サンプルピン当たり、単に約54読み込みの平均に該当する)以上で正確度のわずかな改善しか示さなかった。

【0122】

図2は、開示されたRNAプロファイリング方法を使用した複合化されたRNA定量の正確度を支持するデータを提供する。パネルAは、96サンプル内で特殊な比率で混合された96個の合成miRNAと対照RNAの測定に基づくローズの9,216ピクセル映像を示したヒートマップを示す。二つの複製実験の平均値が表わされるが、それぞれ高容量318チップを使用してシーケンシングされた。正常化は、方法で記述された。RNAは、表2に列挙されたものと同じ順序であった。パネルBは、低容量314チップを使用した二つの複製物から作成された類似のヒートマップを示す。パネルCおよびDは、318(パネルC)または314チップ(パネルD)を使用してサンプルに添加された合成RNAの量とその測定水準の間の一貫性を示す。倍数変化は、各RNAに対する平均と相対的である。パネルEは、公知および測定されたRNA水準の間のPearson相関係数によって定義された、定量正確度に及ぼす配列深さの効果を示す。

【0123】

ヒト組織のmiRNAの多重分析:

人工RNAサンプルを越えて移動し、分析の性能を次に20個の正常ヒト組織から得られたmiRNAに対して試した。これらのサンプルを、開示されたRNAプロファイリング方法を使用して行った測定が妥当であり得る、独立的に公開されたqRT-PCRデータの活用度を基準に選択した。入力は、各サンプルからの50ngのトータルRNAとなり、結果としての読み込みカウントに対して前述されたように、全般的な平均正常化、平均-センタリングおよび自動縮尺を行った。結果を、開示されたRNAプロファイリング方法を使用して行われる測定が対角線上のスプリットピクセルの二つの半分で公開された値と比較された変形したヒートマップを使用して示す(図3a)。データセット間の一貫性は、赤色および緑色半分の組み合わせを有するピクセルの欠乏で証明される。Pearson相関係数の分析は、所定の組織に対して開示されたRNAプロファイリング方法対qRT-PCRによって測定されたRNA水準の間に良好な一致があることを示す(図3b)。また、相関は、関連した組織(例えば、結腸および小腸または卵巣および精巣)の間でも観察された。他のプラットフォームからのデータに対する比較は、図5に示す。測

10

20

30

40

50

定は、試験された多様なプラットフォームにわたって良好な一貫性を示した。ペア方式の相関係数の類似の範囲が開示されたRNAプロファイリング方法を4個の直角プラットフォームと比較した時に発見され、そのような直角プラットフォームを互いに比較した時に発見されたのと同じであった(図5)。また、公知の、等モル量のすべての合成miRNAを定量参照標準として含有するサンプルを共増幅させることによって、相対濃度よりは絶対濃度を測定することが可能であった(図6)。このような分析に基づいて、分析が少なくとも4~5桁の濃度範囲に亘ってmiRNAを測定することが可能であると示された。

【0124】

図3は、ヒト組織および参照サンプルからのRNAで方法の定量性を証明するためのデータを図示する。パネルAは、公開されたqRT-PCR測定に対して20個の正常ヒト組織からの3個の技術的複製物として測定された90miRNAの水準を比較するピクセルが分けられたヒートマップを示す。二つのデータセットのすべて標準化された。パネルBは、同じ組織からの(対角線の)または異なる組織間の(対角線の外)開示されたRNAプロファイリング方法対qRT-PCRによって測定されたmiRNA水準の相関係数のヒートマップを示す。色相図式および組織の順序は、aと同じである。パネルCは、開示されたRNAプロファイリング方法(4回繰返して)対3個の他のプラットフォームによって測定されるところ、MAQC参照サンプルでのmRNA水準の倍数-差のペア方式相関を示す。すべてのプラットフォームに対して30mRNAを試した。線形回帰フィットが図示される。UHR = ユニバーサルヒトリファレンスRNA; HBR = ヒト脳リファレンスRNA。パネルDは、サンプルAおよびB¹²の測定に基づいて予測された水準と比較したMAQCサンプルCおよびDのmRNAの測定された水準の間の%差として規定された、相対的な正確度のボックス図表を示す(同じ30遺伝子に対するものである)。予測された水準は、 $C' = 0.75A + 0.25B$ および $D' = 0.25A + 0.75B$ と計算された。水平線 = 中間; ボックス = 四分位数範囲; ウィスカー = 10番目 ~ 90番目の百分位数; 点 = 特異値。

10

20

【0125】

図5は、複数のmiRNAプロファイリングプラットフォームを比較するためのデータを提供する。パネルAは、開示されたRNAプロファイリング方法を含み、5つの直角プラットフォームを使用して正常な脳および肝から得られるトータルRNAで測定されたmiRNA水準を示す。他の4個のプラットフォームからのデータは、独立した研究所によって報告された。異なるプラットフォームによって測定された脳対肝でのmiRNA水準間の \log_2 (倍数差)に対する値をヒートマップに示す。分析は、すべての分析パネルに共通するmiRNAに制限されていた。miRNAは、その水準が所定のプラットフォームに対して二つのサンプルのすべてにおいて検出限界を下回っている場合、排除させた。報告された値は、3個の技術的複製物の平均である。パネルBは、開示されたRNAプロファイリング方法対4個の直角プラットフォームによって測定されたところ、脳と肝間のmiRNA水準の倍数差の双方式相関を示す。パネルCは、すべてのプラットフォームの組み合わせに対する双方式 R^2 値を示す。外部データセットをTaqaMan qRT-PCR⁸に対してmirnabodymap.orgから、およびIllumina RNA-Seq (GSE49816), Affymetrix配列(GSE49661)およびNanoString (GSE49600)に対するGene Expression Omnibusからダウンロードした。D.M. = 開示されたRNAプロファイリング方法。

30

40

【0126】

図6は、ヒト組織においてmiRNAの絶対定量を図示する。それぞれのmiRNA種の約15,000個の合成複製物を含有する共増幅された定量参照標準サンプルに対して正常化することによって、絶対miRNA濃度を推定することができた。トータルRNA入力は、組織サンプル当たり50ngであった。各値を平均3回の複製RT反応から得て、シングルプレックスPCRに対してプールした。Hsa-mir-381は、増幅が不十分であったため、分析から除外させた。陰影スケールは、mRNAの存在量を示し、挿入された棒グラフは、これらの値の頻度分布を同じ規模で示す。

50

【0127】

参照標準でmRNAの測定：

mRNAを定量するための方法を調整することは難しくなかった。標的長さ制限がないため、RTを高温でさらに長い遺伝子特異的プライマーを使用して行われることができた(表5)。他のマイナーな変形は、方法のセクションで詳細に説明した。検証ベンチマークを提供するために、その発現を一貫した水準で測定した30個の遺伝子をMicroArray品質管理(MAQC)コンソーシアムプロジェクトの一環として、3個の異なる定量プラットフォームを使用して標的化した。分析を、(A)StratageneユニバーサルヒトRNA、(B)Ambionヒト脳RNA、および(C)3:1および(D)1:3の比率のこれら両サンプルの混合物で構成された4個のMAQC参照サンプルから100ngのトータルRNAを使用して4回繰り返して行った。発現水準をACTBおよびPOLR2Aの平均水準に対して正常化した。開示されたRNAプロファイリング方法と三つの定量MAQCプラットフォームのそれぞれの間の倍数変化測定の相関関係の評価するために、サンプルAとBの間の倍数差の双方式回帰分析を行った(図3c)。30遺伝子の共通セットに対し、開示されたRNAプロファイリング方法対Ta q M a nに対するそれぞれの傾きおよび R^2 は、1.02および0.89であり；対比St a R T - P C Rに対しては0.97および0.91であり、対比Q u a n t i G e n eに対しては0.92および0.88であった。サンプルCおよびDは、規定された比率のサンプルAとBで構成されるので、分析の相対的な正確度(RA)は、CおよびDに対する観察された発現水準をAおよびBの測定から計算された予測水準と比較することによって評価することができた。遺伝子に対するRA点数は、 $C = (C - C') / C'$ および $D = (D - D') / D'$ で定義し、ここで、CおよびDは、遺伝子の測定された水準であり、C'およびD'は、予測水準である。30個のmRNAのパネルに対するRA点数のボックス図表は、値が略0周辺に分布していることを示す(図3d)。

10

20

【0128】

放射線被ばくの高処理量評価：

最後に、臨床サンプルに対する開示されたRNAプロファイリング方法の活用度を評価するために、放射線誘発遺伝子発現変化をヒトの血液で測定した。これは、大規模の核の災難後に、全身被爆線量の推定方法として提案されたことがある；しかし、サンプル処理量の最適化は、潜在的に被ばくした数千名の個体の分類を可能にする必要がある。この目的に対して開示されたRNAプロファイリング方法を使用する実現可能性を調べるために、以前に確認された23個の放射線反応性転写物のパネルの発現変化を定量するための分析法が開発された。この分析を、18名の個体から生体外放射線照射された血液サンプル(それぞれ6容量水準)の並列分析を行うために使用した。入力は、全血の放射線照射24時間後に分離された末梢血単核細胞由来の400ngのトータルRNAでなっていた。予想通り、信号が18名のすべての個体にわたって平均化された時に発現の容量依存度の増加がパネルのすべての遺伝子に対して観察された(図4、aおよびb)。また、各個体に対する発現パターンは、全体傾向と良好な一貫性を示した(図4c)。

30

【0129】

図4は、ヒトの血液で放射線誘発遺伝子発現変化の高処理量測定の結果を図示する。以前に確認された放射線反応性遺伝子のパネルで発現水準変化を、18名の個体からの108個の血液サンプルの生体外放射線照射24時間後に測定した。すべてのサンプルを二つの複製RNAプロファイリング実験で並列に処理および測定した。パネルAは、mock照射されたサンプルに対して、多様な放射線容量で遺伝子発現の平均倍数誘導を示す。エラーバーは、SEMを示す。パネルBは18対象体に対して平均化された異なる容量での標準化された遺伝子発現値のヒートマップを示すが、それらのそれぞれの値は、パネルCに別途図示される。平均センタリングおよび自動縮尺を各対象体からのサンプルに対して別途に行った。

40

【0130】

実施例2

50

この実施例は、複雑な核酸混合物内の低存在量の変異体配列の高感度かつ効果的な測定に対して指示された方法およびシステムを記述する。本発明者などは、この実施例で記述された方法を「系統追跡PCR」(LT-PCR)と言及する。LT-PCRの目標は、分子特異的タグ(分子系統タグまたはMLTと言及される)をPCRの最初の数サイクルの間に鋳型DNA分子に配分し、純粋な鋳型由来の突然変異をシーケンサーまたはPCRエラーから区別することを可能にさせることである。この実施例は、癌にかかった患者から得られた血液サンプルからのDNAの分析を記述するが、この方法は、腫瘍組織、細胞、尿などのような他の供給源からのサンプルにも、より一般的に適用され得る。この方法は、一本鎖または二本鎖DNA鋳型に適用され得、さらにRNAの逆転写によって生成された相補的DNA(cDNA)にも適用され得る。

10

【0131】

材料および方法：

患者血漿サンプルの収集およびプロセッシング

血液をカルシウム-EDTAを含有する真空管内への静脈穿刺によって収集した。多様なチューブサイズ、典型的には、3 mL ~ 10 mLの間のもを使用した。血液を収集時にチューブで数回反転させて K_2 -EDTAを均一に混合させた。サンプルを一時的に保管してから、血漿を分離する前に室温(20 ~ 25)で移した。血漿を分離し血液を収集した後、できるだけ迅速に、好ましくは3または4時間内に冷凍させた。収集チューブを1000 x gで10分間、ゆっくりと加速し減速させる揺動バケットローター付き臨床遠心分離機から遠心分離した(ブレーキオフ)。細胞がチューブの底で邪魔にならないように(増加した背景野生型DNA水準を誘導する白血球の吸引を回避するために)注意しながら、血漿を赤血球およびパフィーコートから1 mLピペットを使用して除去した。血漿を0.5 ~ 1 mL部分標本で1.5 mLのクライオバイアルに分配した。次に、血漿を追加のプロセッシングに必要となるまで-80 で冷凍させた。

20

【0132】

血漿からDNAの抽出および精製

血漿を-80 冷凍装置から取り出し、DNA抽出を進行する前に室温で15 ~ 30分間解凍した。

【0133】

血漿からDNAの抽出および精製された血漿を次に6800 x gで3分間遠心分離して任意の低温沈殿物を除去した。上清を追加のプロセッシングのために新しいチューブに移した。

30

【0134】

1 mLまでの血漿体積からDNAの抽出のために、QiaAmpMinEluteウイルス真空キット(Qiaagen)を使用した(溶出体積は20 μL程度に低い)。5 mLまでの血漿の大量の体積に対しては、DNA精製のためにQiaAmp循環核酸キットを使用した(溶出体積は20 μL程度と低い)。すべてのキットを製造業者の説明書に従って使用し、一般的に最低推奨体積(好ましくは20 μL)でDNAを溶出した。1 mLの血漿をQiaAmp MinEluteウイルス真空キットを使用して処理するために、5マイクログラムの担体RNA(cRNA; Qiaagen)を1 mL当たり添加し、Qiaagenウェブサイトで探した使用者開発プロトコルに従った。

40

【0135】

遮断された3'末端を有するユニバーサルプライマーおよびMLT-含有遺伝子特異的プライマーの合成：

オリゴヌクレオチドをPCRによる増幅のために、ゲノムDNAの特異的突然変異を起こしやすい領域を標的化するようにデザインした。プライマーを自動DNAオリゴヌクレオチド合成器(Dr. Oligo 192)で標準ホスホラミダイト化学を使用して3'から5'方向にユニバーサルポリスチレン支持体III(Glen Research)で200ナノモル規模で合成した。プライマーのデザインは、図7および図8に概略的に図示される。遺伝子特異的プライマーは、3'-末端に遺伝子特異的配列を有し、それら

50

は、M L Tを含む7つの縮重位置を含有し、ユニバーサルプライマー配列の一部のを含有する。ユニバーサルプライマーは、溶融温度を高めるためにL N A変形を含有する。プライマー配列は、下記の表10で列挙する。プライマーをゲル精製するか、カートリッジ精製した。この方法が複数の標的を同時に分析できるかを証明するために、プライマーを癌で頻繁に突然変異した8つのゲノム領域を標的化するようにデザインした：K R A Sの1領域、B R A Fの1領域、P P P 2 R 1 Aの1領域、P I K 3 C Aの2領域、およびE G F Rの3領域。この実施例において、8個のゲノム領域が標的化されたが、この方法は数十または数百または数千個の標的アンプリコンに容易に拡大することができる。

【 0 1 3 6 】

[表 1 0]

系統追跡されたPCRでおよび区画化されたPCR実験で使用されたプライマーの目録

標的化された遺伝子	プライマー配列	
遺伝子特異的順方向プライマー:		
KRAS	CTACACGACGCTCTTCCGATCTNNNNNN NAGGCCTGCTGAAAATGACTGAATATA a ACTTXX	10
BRAF	CTACACGACGCTCTTCCGATCTNNNNNN NCCTCACAGTAAAAATAGGTGATTTTG g TCTAXX	
PPP2R1A	CTACACGACGCTCTTCCGATCTNNNNNN NGACTCCAGGTACTTCCGG a ACCTXX	
PIK3CA領域1	CTACACGACGCTCTTCCGATCTNNNNNN NCAGCTCAAAGCAATTTCTACACG a GAT CXX	20
PIK3CA領域2	CTACACGACGCTCTTCCGATCTNNNNNN NGCAAGAGGCTTTGGAGTATTTTCAT g AA ACXX	
EGFR領域1	CTACACGACGCTCTTCCGATCTNNNNNN NGGATCCCAGAAGGTGAGAAAGTTAA a A TTCXX	30
EGFR領域2	CTACACGACGCTCTTCCGATCTNNNNNN NAAAACACCGCAGCATGTCAA g ATCAXX	
EGFR領域3	CTACACGACGCTCTTCCGATCTNNNNNN NCATCTGCCTCACCTCCA c CGTGXX	
遺伝子特異的逆方向プライマー:		40
KRAS	CAGACGTGTGCTCTTCCGATCTNNNNNN NCTGAATTAGCTGTATCGTCAAG g CACT XX	
BRAF	CAGACGTGTGCTCTTCCGATCTNNNNNN NACTGTTCAAACCTGATGGGAC c CACTXX	

PPP2R1A	CAGACGTGTGCTCTTCCGATCTNNNNNN NCTTGGCAAACCTCCCCCAgCTTGXX	
PIK3CA領域1	CAGACGTGTGCTCTTCCGATCTNNNNNN NATCTCCATTTTAGCACTTACCTgTGAC XX	
PIK3CA領域2	CAGACGTGTGCTCTTCCGATCTNNNNNN NTCAATGCATGCTGTTTAATTGTgTGGA XX	10
EGFR領域1	CAGACGTGTGCTCTTCCGATCTNNNNNN NAGCAGAAACTCACATCGAGGaTTTCXX	
EGFR領域2	CAGACGTGTGCTCTTCCGATCTNNNNNN NTGCCTCCTTCTGCATGGTATTcTTTCX X	
EGFR領域3	CAGACGTGTGCTCTTCCGATCTNNNNNN NAGCCAATATTGTCTTTGTGTTcCCGGX X	20
ユニバーサルプライマー:		
短いユニバーサル順方向	ACAC+TCT+TTCCC+TACACGACGCTGT TCCgATCTXX	
短いユニバーサル逆方向	G+TGAC+TGGAGT+TCAGACGTGTGCTG TTCCgATCTXX	30
バーコード&シークエンシングアダプタがある長いユニバーサル順方向	AATGATACGGCGACCACCGAGATCTACA C [FWDBC] ACAC+TCT+TTCCC+TACAC GACG-CTCTTCCgATCTXX	
バーコード&シークエンシングアダプタがある長いユニバーサル逆方向	CAAGCAGAAGACGGCATACGAGAT [REV BC] G+TGAC+TGGAGT+TCAGACGTGTG CTC-TTCCgATCTXX	40
注:	X = dA-5' -CEホスホラミダイト (Glen Researc h) を使用した反対方向の dA。 下部の場合の残基はRNA ; 上部の場合の残基はDNAである。 FWDBC = 順方向バーコード ; REVBC = 逆方向バーコード。 順方向および逆方向バーコードを表6から選択した。	

	<p>N=A、T、CまたはGを統合させる可能性が同等な縮重位置。 A残基の前方で「+」はその位置でLNAヌクレオチドを指す。 すべてのプライマーをユニバーサルポリスチレン支持体III (Glen Research) で合成した。</p>
--	--

【0137】

系統追跡されたPCRタグ付けおよび増幅：

変形されたポリメラーゼ連鎖反応(PCR)を各DNA鋳型サンプルに対して、下記で概略的に説明する条件を使用して単一反応チューブで行った： 10

【0138】

系統追跡されたPCR設定(20μL反応)：

精製された鋳型DNA(共溶出された坦体RNA[crRNA]を含有することができる)10μL(またはそれ未満)

5x濃縮されたPhusion HF緩衝液(Thermo)4μL

16個の遺伝子特異的プライマーのミックス(ストックはそれぞれ200nMを有する)2μL

サンプル特異的バーコードおよびシークエンシングアダプタを有するユニバーサル順方向および逆方向プライマーの混合物(ストックはそれぞれ5μMを有する)2μL 20

4dNTPの混合物(ストックはそれぞれ10mM)0.4μL

Phusion Hot Start II DNAポリメラーゼ(Thermo)(2U/μLストック)0.2μL

RNAse H2(Integrated DNA Technologies)(20mU/μLストック)1μL

水(20μLの最終体積を形成するためである)。

【0139】

一部の反応のために、さらに長いユニバーサルプライマーに加えてさらに短いユニバーサルプライマー(バーコードおよびシークエンシングアダプタがない[表10])をそれぞれ200nMの最終濃度で、一緒に添加した。さらに迅速な混成化速度を有するさらに短いユニバーサルプライマーを含むことは、MLT標識された複製物のより有効である早期増幅を促進するために意図された。 30

【0140】

温度サイクリング条件：

a. 30秒間98

b. 10秒間98

c. 70 から、10秒当たり1 の速度で徐々に60 に低下する

d. 1分間60

e. 30秒間72

f. 工程b~eを2サイクル以上間繰り返す(合計3サイクル) 40

g. 10秒間98

h. 60秒間72

i. 34回以上のサイクルの間に工程g~hをさらに繰り返す(合計35サイクル)

g. 4 に維持

【0141】

熱サイクリングが終わった時、各反応体積に2μLの100mM EDTA含有緩衝液を添加し、ポリメラーゼ活性を不活性化させた。次いで、各サンプルからのおおよそ10μLの増幅生成物を増幅されたDNAの後続する精製のために単一チューブにプールした。

【0142】

次世代シーケンシングのためのDNAの製造：

プールされたPCR反応生成物をエチジウムブロマイドおよび1×TBE緩衝液を含む2%アガロースゲルで精製した。すべてのPCR生成物が同様の最終長さのものであったため、プールされた生成物は、ゲルで多少拡散帯として現れた。この拡散帯を新しいメスの刃を使用してゲルから切断し、ゲルが可視帯域の数ミリメートル上下を切断してさらに早くまたはさらに遅く作動され得、十分に可視化されていない任意の低密度帯までも含むことを保証した。製造業者の説明書に従って、QIAquickゲル抽出キット(Qiagen)を使用してDNAをゲルスライスから分離した。DNAを50μLの溶出緩衝液、EBに溶出させた。

【0143】

次世代シーケンシング

Illumina HiSeq流れセル上にローディングするためのサンプルを製造するために、DNAの濃度をAgilent Bioanalyzerを使用して測定し、DNAをIlluminaによって推奨された濃度に希釈した。クラスターの形成をIlluminaのプロトコルにしたがってフローセルで行った。サンプルをフローセルが単一のレーン上にローディングした。シーケンシングをHiSeq 2000機器で多重化されたペアエンド方式で行い、この時、各方向の読み込み長さは75塩基対であった。追加の実験で、シーケンシングをIllumina MiSeq機器で行い、各方向に100、150、200または250塩基対のペアエンド読み込み長さが活用された。また、2回のインデックス読み込みを行い、インデックス読み込み長さは標準7サイクルから9サイクルまで増加して、本発明のさらに長いバーコード(インデックス)配列を適切に読み込むことができた。

【0144】

実施例3

実施例2と同様に、実施例3は、複雑な核酸混合物内の低存在量の変異体配列の高感度で効果的な測定に対して指示された方法およびシステムを記述する。この実施例は実施例2で記述したように、「系統追跡されたPCR」(LT-PCR)を統合するが、追加で分析感度を改善するために区画化戦略を使用する。PCRを多くの小さい反応体積に分けたため、所定の反応体積で特定の標的化されたDNA断片の一つ以上の複製物を有する可能性が非常に低くなった。変異体配列の増幅された複製物が所定の反応区画内で二本鎖鋳型DNA断片の両鎖から発生することを確認することを可能にするように作ったタグ付け戦略を使用した。この実施例は、癌にかかった患者から得られた血液サンプルからのDNAの分析を記述するが、この方法は、腫瘍組織、細胞、尿などのような他の供給源からのサンプルにもより一般的に適用され得る。この方法はまた、一本鎖DNA鋳型におよびRNAの逆転写によって生成された相補的DNA(cDNA)にも適用され得るが、エラー抑制の強固さに妥協がなければならない。

【0145】

材料および方法

患者血漿サンプルの収集およびプロセッシング

血液を実施例2に記述の方法と同じ方法を使用して収集した。

【0146】

血漿からのDNAの抽出および精製

DNAを実施例2記述の方法と同じ方法を使用して患者の血漿サンプルから抽出した。

【0147】

遮断された3'末端を有するユニバーサルプライマーおよびMLT含有遺伝子特異的プライマーの合成

実施例2で合成された同じプライマー(表10)をこの実施例で使用するものの、長い順方向ユニバーサルプライマー(バーコードおよびシーケンシングアダプタを含有する)は例外とした。プライマーの合成を実施例2記述の方法と同じ方法を使用して行った。

【0148】

10

20

30

40

50

磁性ビーズ上のビーズ特有のバーコードを含有するオリゴヌクレオチドのスプリットおよびプール合成

磁性のマイクロビーズを使用してバーコーディングされた順方向ユニバーサルプライマーを異なるPCR微細区画(例えば、液滴またはマイクロウェル)に伝達した。各ビーズを多くのプライマー複製物がすべて同じビーズ特有のバーコード(BSBC)を有するようにデザインした。所望の順方向ユニバーサルプライマー配列の配列は、次のとおりである:

【0149】

5' - ビオチン - A A T G A T A C G G C G A C C A C C G A G A T C T A C A C [B S B C] A C A C T C T T T C C C T A C A C G A C G C T C T T C C - 3'

10

【0150】

約百万ビーズ特有のバーコードを有する数百万個の磁性のマイクロビーズを生成するために、オリゴヌクレオチド合成をバーコード配列を生成するためのスプリットおよびプール接近法を使用してビーズの表面で直接行った。アミン変形を有する表面活性化された超常磁性2.8 μmビーズ(Dynabeads M-270 Amine [Thermo Scientific])をオリゴヌクレオチド合成のためのソリッド支持体として使用した。合成の各バッチに対して、50 μLのビーズスラリーを製造業者によって提供された通りに使用した(約100 × 10⁶ビーズ)。ビーズは、フリットによって合成カラムに保有されるには小さすぎるため、ドーナツ状のネオジウム磁石をカラム周辺に置いて磁性ビーズをカラムの側面の位置に保持させた。スペーサー9ホスホラミダイト(Glen Research)は直接アミン変形されたビーズと反応してアンモニウムヒドロキシド/メチルアミン(AmA)で標準脱保護される間に切断されないホスホラミデート結合を生成した。追加のホスホラミダイトは、このスペーサーに結合されて所望のオリゴヌクレオチド鎖が成長された。合成されたオリゴヌクレオチドは、合成が完了した時、ビーズに付着したままであった。次の配列をビーズの表面で合成した:

20

【0151】

5' - スペーサー9 - T T T T T T T T T T - スペーサーC3 - G G A A G A G C G T C G T G T A G G G A A A G A G T G T [B S B C] G T G T A G A T C T C G G T G G T C G C C G T A T C A T T - 3'

【0152】

5'から3'方向にオリゴヌクレオチドを合成するために、5'-CEホスホラミダイト(Glen Research)を使用した。オリゴヌクレオチド配列は、ビーズの表面から追加の空間を導入するために、10dTの残基を含有する。ビーズ特有のバーコード(BSBC)は、スプリットおよびプール合成を使用して合成された10個の残基で構成された。これら10個の位置のそれぞれにおいてホスホラミダイトカップリングのために、合成を停止させ、磁性ビーズをプールした後、4個のカラムに分けた。4個の異なるカラムは、4個の異なるホスホラミダイト(5'-dA、5'-dT、5'-dCおよび5'-dG)を受けた。合成を10回のカップリングサイクルのそれぞれの間に停止してビーズがプールされ、4個のカラムに同等に再分布されることを許容した。合成が完了した後、ビーズ結合されたオリゴヌクレオチドをAmAで65で10分間脱保護させた。次いで、ビーズを脱イオン水で洗浄し、次いで10mMトリスpH7.6緩衝液に懸濁した。

30

40

【0153】

マイクロビーズの表面で熱放出性相補的バーコード化プライマーを合成するために、次のプライマーをビーズ結合されたオリゴヌクレオチドにアニーリングし、クレノウ断片(Exo-)(New England Biolabs)を使用して延長させた。

【0154】

5' - ビオチン - A A T G A T A C G G C G A C C A C C G A G A T C - 3'

【0155】

ビーズを0.2mM dNTPが補充された50μLのNEB緩衝液2(1×濃度)に

50

再懸濁した。プライマー延長反応を製造業者の指示にしたがって行い、クレノウポリメラーゼを添加した後に37℃で30分間反応をインキュベーションした。次いで、ビーズを50 mM NaClおよび10 mM トリス pH 7.6を含有する緩衝液で洗浄し、そこで再懸濁した。

【0156】

クローンタグ付きプライマーを区画に伝達するためのビーズフリー方法

一部の実験で、ビーズを使用する代わりに、代替接近法を使用して区画内のPCR生成物に区画固有タグを導入した。ビーズベースの伝達と同様に、目標は、異なる区画に次のプライマー配列を伝達することであった：

【0157】

5' - ビオチン - A A T G A T A C G G C G A C C A C C G A G A T C T A C A C [C S B C] A C A C T C T T T C C C T A C A C G A C G C T C T T C C - 3'

【0158】

所定の区画で、このプライマーの多数の複製物を導入し、この時、クローン複製物は一つまたは少数の区画固有バーコード(CSBC)を含有する。そのようなプライマーを製造するために、非常に希釈された鋳型DNAをPCRカクテルに添加した後、平均2乃至~3増幅可能な複製物(分子)が各溝画に(Poisson分布に沿って)分配されることを許容する濃度で区画化した。鋳型DNAは、次の配列で構成された：

【0159】

Degen 鋳型：

5' - A A T G A T A C G G C G A C C A C C G A G A T C T A C A C N N N N N N N N N N A C A C T C T T T C C C T A C A C G A C G C T C T T C C - 3'

次のプライマーをさらにカクテルに添加した：

Bio-ShortFWD：

5' - ビオチン - A A + T G + A T + A C G G C G A C C A C C G A G a T C T A X X - 3' (100 nM 最終濃度で添加する)

ShortREV：

5' - G G A + A G A G C G + T C G + T G T A G G G A A a G A G T X X - 3' (20 nM 最終濃度で添加する)

X = dA - 5' - CEホスホラミダイト(Glen Research)を使用した反対方向のdA。

下部の場合の残基はRNA；上部の場合の残基はDNAである。

N = A、T、CまたはGを組み込む可能性が同等な縮重位置。

A残基の前方で「+」は、その位置でLNAヌクレオチドを指す。

【0160】

マイクロコンパートメントに対して熱サイクリングを遂行することによって、少数のタグ付き鋳型分子がクローン増幅されて、区画固有タグを含有する所望のプライマーの多くの複製物が生成された。ビオチニル化された短い前方プライマーが短い逆方向プライマーと比較して5' - 倍過剰に添加されたため、逆方向鎖の複製物よりもさらに多い前方鎖複製物が(非対称PCRを通じて)形成された。したがって、前方鎖の過剰複製物は、次に同一区画で共増幅された遺伝子特異的PCR生成物に混成化されることによってさらに延長され得る。このような方法で、区画の遺伝子特異的PCR生成物は区画固有タグで標識された。この接近法は、図12で図式化される。

【0161】

PCRカクテル

この実施例で使用されたPCRカクテルは、マイクロビーズが区画固有プライマーを伝達するために使用されたか、またはビーズフリー接近法が使用されたかによって左右される。

【0162】

ビーズベース接近法に対して、次のPCRカクテルを使用した：

10

20

30

40

50

精製された鋳型 DNA (共溶出された担体 RNA [cRNA] を含有することができる) 10 μ L (またはそれ未満)

5x 濃縮された Phusion HF 緩衝液 (Thermo) 4 μ L

16 個の遺伝子特異的プライマーの混合物 (ストックはそれぞれ 200 nM を有する) 2 μ L

短いユニバーサル順方向および逆方向プライマー (ストックはそれぞれ 10 μ M) 1 μ L

サンプル特異的バーコードおよびシーケンシングアダプタを有する長いユニバーサル逆方向プライマー (10 μ M ストック) 1 μ L

4 dNTP の混合物 (ストックはそれぞれ 10 mM) 0.4 μ L

Phusion Hot Start II DNA ポリメラーゼ (Thermo) (2 U/ μ L ストック) 0.2 μ L

RNAse H2 (Integrated DNA Technologies) (20 mU/ μ L ストック) 1 μ L

水 (20 μ L の最終体積を作るためである)

【0163】

(プライマー配列は表 10 に列挙される)

【0164】

タグ付きプライマーを担持するビーズを区画化直前にカクテルに添加し、ビーズが区画に均一に分布することを促進するためによく混合した。ビーズの数を調整し、平均 ~ 2 から ~ 3 個のビーズがマイクロコンパートメントに分布され得るようにした。

【0165】

ビーズフリー接近法を使用して区画固有タグを含有するクローンプライマーを導入したとき、次の PCR カクテルを使用した:

精製された鋳型 DNA (共溶出された担体 RNA [cRNA] を含有できる) 8 μ L (またはそれ未満)

5x 濃縮された Phusion HF 緩衝液 (Thermo) 4 μ L

16 個の遺伝子特異的プライマーの混合物 (ストックはそれぞれ 200 nM を有する) 2 μ L

短いユニバーサル順方向 (ストック 5 μ M) および短いユニバーサル逆方向プライマー (ストックは 10 μ M) の混合物 1 μ L

サンプル特異的バーコードおよびシーケンシングアダプタを有する長いユニバーサル逆方向プライマー (10 μ M ストック) 1 μ L

Degen 鋳型 (ストック濃度は下記記述されるもののように調整する) 1 μ L

Bio-Short FWD (1 μ M ストック) および短い REV (0.2 μ M ストック) の混合物 1 μ L

4 dNTP の混合物 (ストックはそれぞれ 10 mM) 0.4 μ L

Phusion Hot Start II DNA ポリメラーゼ (Thermo) (2 U/ μ L ストック) 0.2 μ L

RNAse H2 (Integrated DNA Technologies) (20 mU/ μ L ストック) 1 μ L

水 (20 μ L の最終体積を作るためである)

【0166】

「Degen 鋳型」プライマーのストック溶液の濃度を平均 2 ~ 3 個の増幅可能な分子がそれぞれの区画に分配されるように調整した。デジタル PCR 実験をこの鋳型の連続希釈を使用して行い、増幅可能な分子の濃度を正確に測定した。

【0167】

PCR の微小流体区画化

二つの異なる接近法を PCR カクテルを熱サイクリングする前に、微細反応体積で区画化するために使用した。一つ目の接近法は、オイル中に水性 PCR カクテル (選択的にマ

10

20

30

40

50

イクロビーズを含有する)の微小流体性液滴を製造するものであった。二つ目の接近法は、微小流体装置のマイクロウェルにPCRカクテル(選択的にマイクロビーズを含有する)を分割するものであった。二つの接近法両方において、それぞれ概略1ナノリットルの概略20,000個の異なるの微細な反応体積が20マイクロリットルのPCRカクテルから生成された。区画の総数および大きさは、分析されるゲノム同等物の数によって将来の実験で調整され得る。この実施例で使用した区画化図式は、ゲノム鋳型DNA(~3000ゲノム同等物)の約8~10ngの推定を基盤とした。

【0168】

PCRカクテルをオイル中の水性液滴に区画化するために、BioRad QX100液滴発生器を製造業者の説明書から一部変形して使用した。一部変形とは、前記PCRカクテル(マイクロビーズがある、またはない)を製造業者の推奨されたPCRスーパーミックスの代わりに使用したものであった。EvaGreenに対する液滴生成オイルを使用した。熱サイクリングを0.2mL薄膜のPCRチューブで行った。

10

【0169】

PCRカクテルをマイクロウェルに区画化するために、本発明者などは、それぞれ約1nL体積を保有する20,000個のマイクロウェルを生成するために、その上にポリジメチルシロキサン(PDMS)がパターン化されたカスタム微細加工透明スライドを使用した。PDMS表面をPCRカクテルのマイクロウェルへの均一な分布を促すために表面を親水性にする処理をした。カバースリップをPDMSパターンを挟むために添加し、それによってマイクロウェルを熱サイクリングのために密封した。

20

【0170】

熱サイクリング

実施例2で使用したプロトコルに類似する熱サイクリングプロトコルを使用し、ただし、最後の2回のサイクルは、区画固有タグを含有するビオチン標識されたプライマーの混成化および延長を促進するためにさらに低いアニーリング温度を有した。

【0171】

温度サイクリング条件：

- a . 30 秒間 9 8
- b . 10 秒間 9 8
- c . 70 から、10 秒当たり 1 の速度で徐々に 60 に低下する
- d . 1 分間 6 0
- e . 30 秒間 7 2
- f . 工程 b ~ e をさらに 2 サイクルの間に繰り返す (合計 3 サイクル)
- g . 10 秒間 9 8
- h . 60 秒間 7 2
- i . 3 4 回以上のサイクルの間に工程 g ~ h を繰り返す (合計 3 5 サイクル)
- j . 10 秒間 9 8
- k . 60 秒間 6 0
- l . 1 回以上のサイクルの間に i ~ k を繰り返す (合計 2 サイクル)
- m . 4 維持

30

40

【0172】

すべての区画からタグ付き生成物の組み合わせ

熱サイクリングを完了した時、区画化された反応体積を組み合わせ、EDTA含有緩衝液を組み合わせられた体積に添加して(~10mMの最終濃度)ポリメラーゼ活性を不活性化した。オイル中の液滴を合体するために、クロロホルムを添加し、そのエマルジョンをボルテクサーで攪拌した後、Bio-Rad推奨プロトコルに従って高速で遠心分離した。マイクロウェルからのPCR生成物を組み合わせるために、カバースリップを除去し、マイクロウェルを約200μLのEDTA含有緩衝液で洗浄した。万一、磁性ビーズがカクテルに添加されていれば、これを磁石を使用して溶液から除去した。

【0173】

50

次世代配列化のためのDNAの製造：

プールされたPCR反応生成物をエチジウムブロマイドおよび1×TBE緩衝液を含んだ2%アガロースゲルで精製した。予想されたサイズ(隣接したレーンで走ったサイズマーカーを基準に)のバンドを新しいメス刃を使用してゲルから切断した。製造業者の説明書に従ってQIAquick(登録商標)ゲル抽出キット(Qiagen)を使用してDNAをゲルスライスから分離した。DNAを50μLの溶出緩衝液、EB(Qiagen)に溶出させた。

【0174】

一部の実験で、高容量ストレプトアビジンアガロース樹脂スラリー(5μL)(Thermo Scientific)を各反応体積に添加してビオチン標識反応生成物を捕捉した。次いで、ビーズを10mMトリスpH7.6で洗浄した後、ビオチン化鎖に相補的であるDNA鎖をビーズ表面から50μLの溶出緩衝液EB(Qiagen)での熱変性により溶出した。

10

【0175】

次世代シーケンシング

Illumina HiSeqフローセル上にローディングするためのサンプルを製造するために、DNAの濃度をAgilent Bioanalyzer(登録商標)を使用して測定し、DNAをIlluminaによって推奨された濃度に希釈した。シーケンシングを実施例2で記述した通りに行った。

【0176】

20

配列分析のためのアルゴリズムの概要

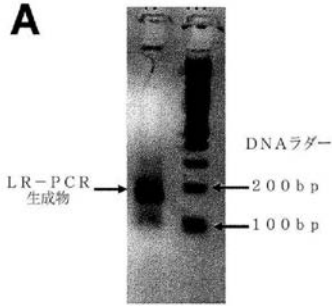
コンピュータを利用した分析を両鎖からマッチング突然変異配列を生成した突然変異二本鎖DNA断片を確認し定量するために、結果的な配列データに対して行った。この分析に使用された基礎となる論理は、「方法」のセクションで記述される。

【0177】

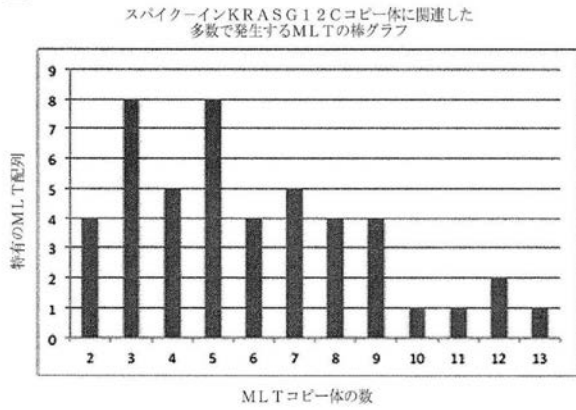
本発明が特定具体例の観点で記述されたが、発明はそのような具体例に制限されるものではない。発明の要旨を逸脱しない変形、変更は技術分野の当業者に自明である。開示された具体例の先行技術は、技術分野の任意の当業者が本開示を製造するか使用することを可能にするために提供されるのは自明である。これら具体例に対する多様な変形は技術分野の当業者に明白であり、本願で定義された一般的な原理は開示の思想または範疇を逸脱することなく他の具体例適用され得る。したがって、本開示は本願に提示された具体例に制限されるものと意図されないが、本願に開示された原理および新規の特徴と一貫した最も広範囲な範疇に従うものであるべきである。

30

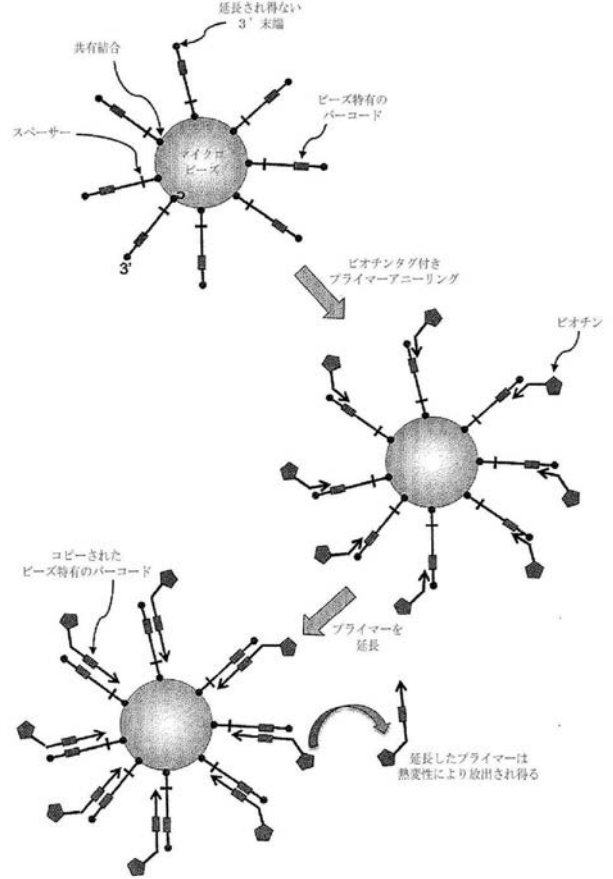
【 図 9 】



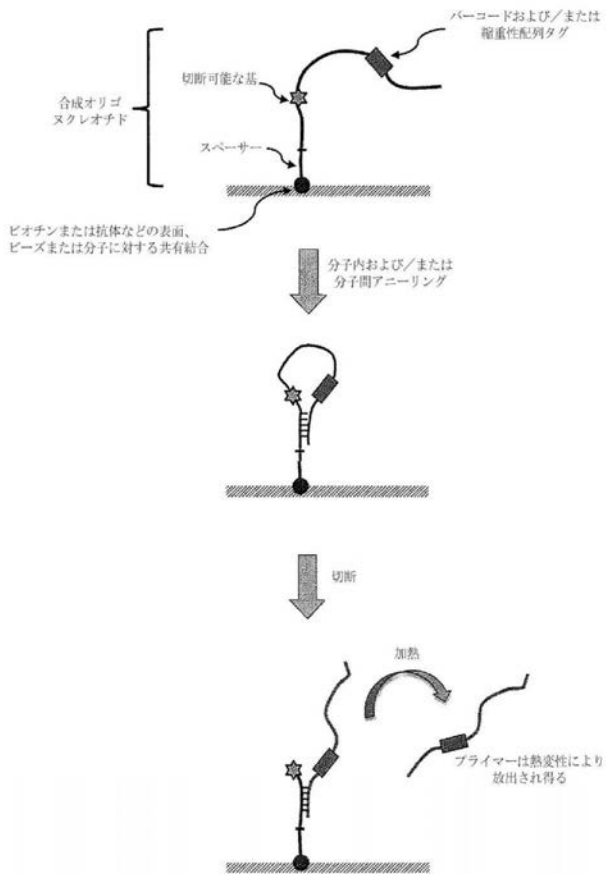
B



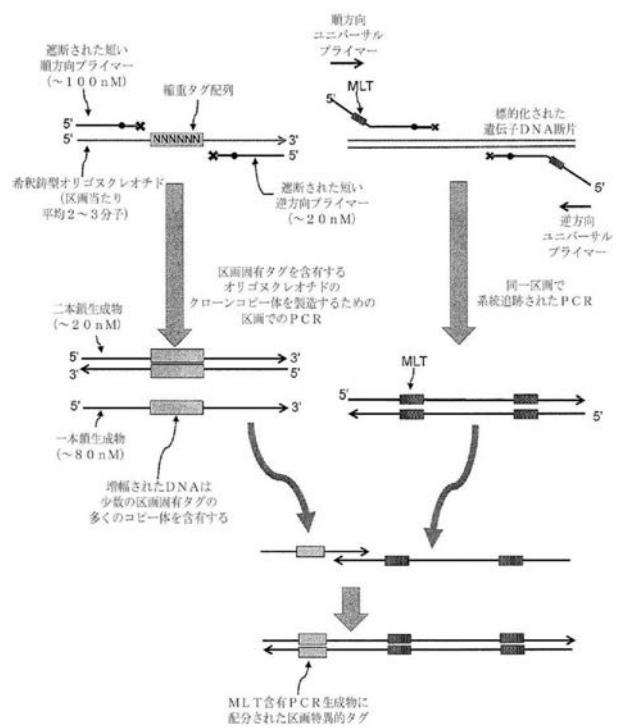
【 図 10 】



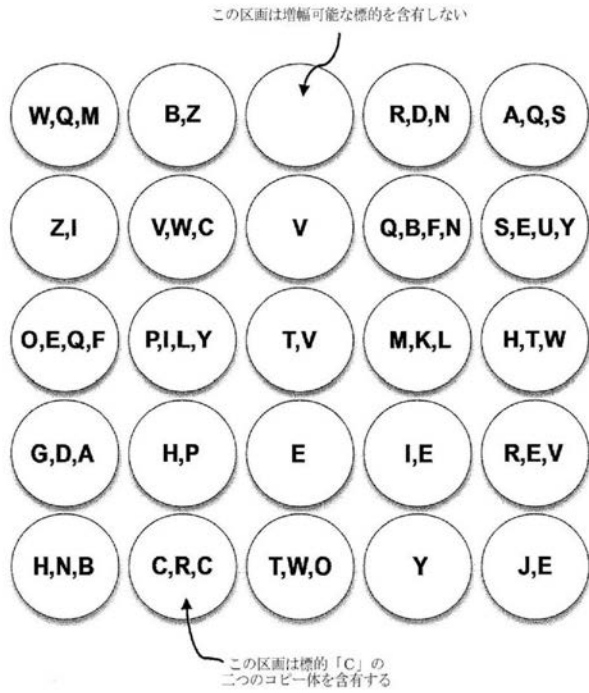
【 図 11 】



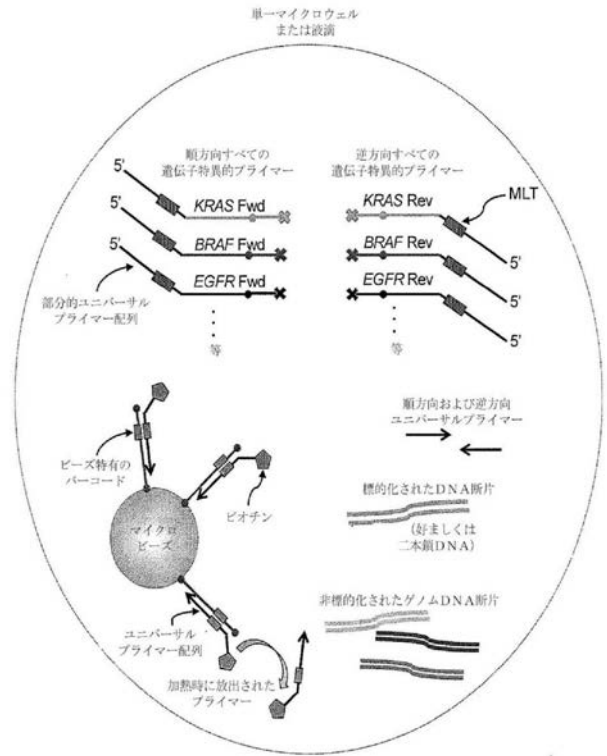
【 図 12 】



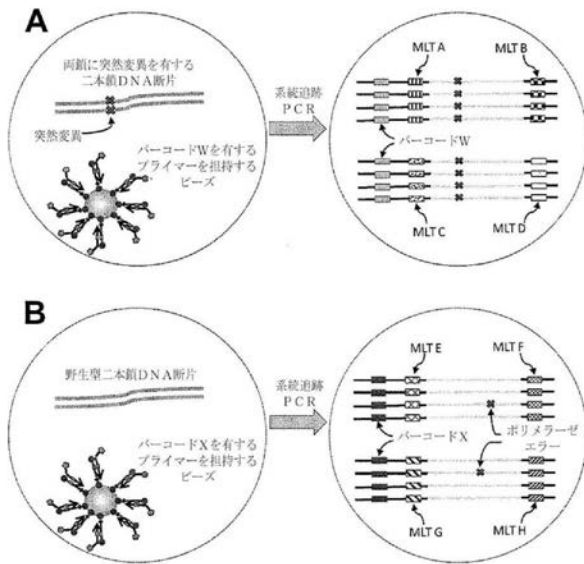
【 図 1 3 】



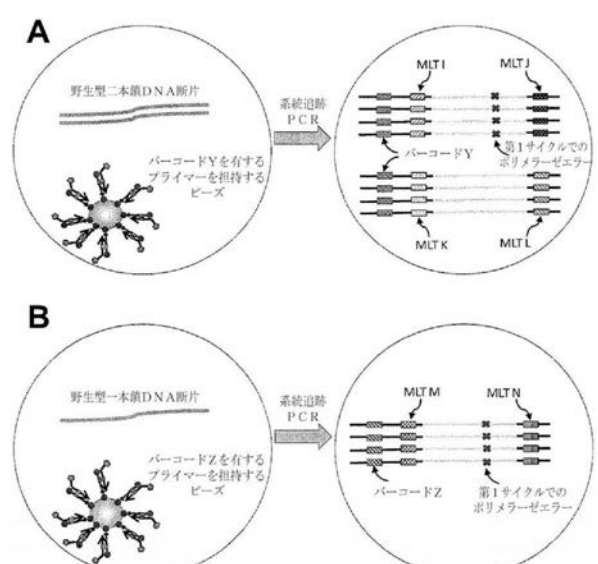
【 図 1 4 】



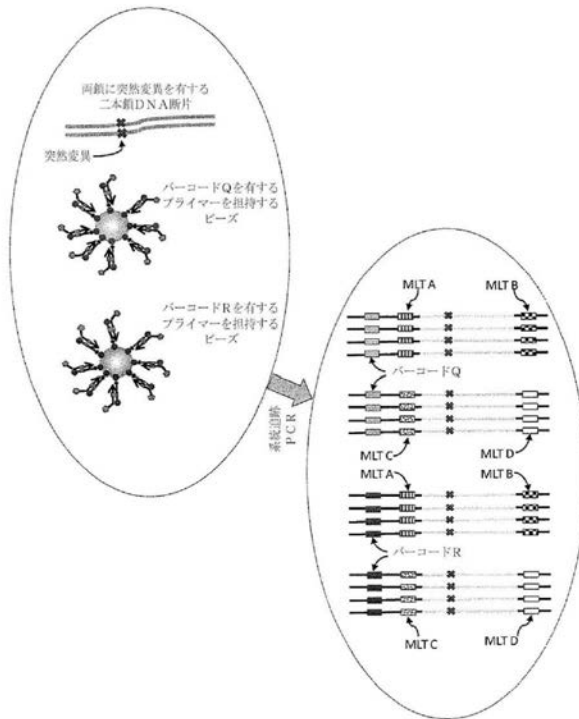
【 図 1 5 】



【 図 1 6 】



【 図 1 7 】



【 手続補正書 】

【 提出日 】平成29年10月11日(2017.10.11)

【 手続補正 1 】

【 補正対象書類名 】特許請求の範囲

【 補正対象項目名 】全文

【 補正方法 】変更

【 補正の内容 】

【 特許請求の範囲 】

【 請求項 1 】

二本鎖核酸断片のペア鎖から得られた配列の同定方法であって、

複数の二本鎖核酸断片を水溶液に溶解させる段階と；

前記溶液を複数の区画に分配する段階であって、区画は、二つ以上の二本鎖核酸断片を含有する可能性が低く、二本鎖核酸断片の増幅生成物が同じゲノム基準配列に対して配列される段階と；

区画化された二本鎖核酸断片の両鎖をPCRを行うことによってコピーし、増幅させる段階と；

一つ以上の区画固有DNA配列タグを増幅されたDNAコピー体に付着させて、二本鎖核酸断片の両鎖のコピー体に同じタグまたはタグのセットが付着される段階と；

増幅され、タグ付着されたDNAコピー体を含有する区画を組み合わせる段階と；

全部または下位セットの増幅され、タグ付きDNAコピー体をシーケンシングする段階と；

二本鎖核酸断片のペア鎖から得られた配列を共通区画固有DNA配列タグまたはタグのセットの共有を基準として同定する段階と；

を含む方法。

【 請求項 2 】

前記二本鎖核酸断片のペア鎖から得られた配列の比較は、二本鎖核酸断片の配列を測定するとき、エラーの減少を可能にする、請求項 1 に記載の方法。

【請求項 3】

前記核酸配列の測定においてエラー減少は、核酸配列の混合物内で低存在量 (low abundance) の配列変異体を同定するために使用される、請求項 2 に記載の方法。

【請求項 4】

前記二本鎖核酸断片は、RNA である、請求項 1 に記載の方法。

【請求項 5】

前記二本鎖核酸断片は、DNA である、請求項 1 に記載の方法。

【請求項 6】

前記二本鎖核酸断片は、ゲノム DNA である、請求項 1 に記載の方法。

【請求項 7】

前記二本鎖核酸断片は、血液由来のゲノムセルフリー DNA である、請求項 1 に記載の方法。

【請求項 8】

前記二本鎖核酸断片は、腫瘍組織由来のゲノム DNA である、請求項 1 に記載の方法。

【請求項 9】

前記二本鎖核酸断片は、ホルマリン固定された、パラフィン挿入された腫瘍組織由来のゲノム DNA である、請求項 1 に記載の方法。

【請求項 10】

前記二本鎖核酸断片は、合成アダプタ分子が結紮されているゲノム DNA を含む、請求項 1 に記載の方法。

【請求項 11】

前記合成アダプタ分子は、次の成分：DNA、RNA、変形塩基、および自然的に発生する核酸で発見されないオリゴヌクレオチド変形のうち少なくとも一つを含む、請求項 10 に記載の方法。

【請求項 12】

前記二本鎖核酸断片の結紮されたアダプタの部分は、増幅前に一つ以上の公知されたミスマッチ塩基対を含有し、増幅されたコピー体の系統が上部または下部鎖に対して追跡されることを可能にする、請求項 10 に記載の方法。

【請求項 13】

前記二本鎖核酸断片は、両鎖の全体長さに沿って、または両鎖の長さの一部に沿って相補的塩基対を有する、請求項 1 に記載の方法。

【請求項 14】

前記水溶液は、オイル中の水性液滴に区画化される、請求項 1 に記載の方法。

【請求項 15】

前記水溶液は、固体分離器、半固体分離器、または固体および半固体分離器を使用してチャンバーに区画化される、請求項 1 に記載の方法。

【請求項 16】

増幅生成物が同じゲノム基準配列に対して配列される二つ以上の二本鎖核酸断片を区画が含有する可能性が 10% 未満で存在する、請求項 1 に記載の方法。

【請求項 17】

増幅生成物が同じゲノム基準配列に対して配列される二つ以上の二本鎖核酸断片を区画が含有する可能性が 1% 未満で存在する、請求項 1 に記載の方法。

【請求項 18】

区画化された二本鎖核酸断片は、特異的なゲノム配列を標的化する一つ以上のプライマーペアを使用する PCR により増幅される、請求項 1 に記載の方法。

【請求項 19】

区画化された二本鎖核酸断片は、結紮されたアダプタ配列を標的化する一つ以上のプラ

イマーペアを使用するPCRにより増幅される、請求項1に記載の方法。

【請求項20】

前記区画固有タグまたは区画固有タグのセットは、異なる区画内で増幅されるDNAコピー体を区別するために使用され得る、請求項1に記載の方法。

【請求項21】

前記区画固有DNA配列タグは、プライマー配列内に統合され、前記プライマーを使用するPCRにより増幅されたDNAコピー体に付着される、請求項1に記載の方法。

【請求項22】

前記区画は、区画当たり平均10個未満の区画固有タグを含有する、請求項1に記載の方法。

【請求項23】

前記区画は、区画当たり平均2～3個の区画固有タグを含有する、請求項1に記載の方法。

【請求項24】

前記増幅され、タグ付着されたDNAコピー体の下位セットは、ハイブリッド捕捉のような標的豊富化方法により、またはイン-ソリューション捕捉でシーケンシングされるために選択される、請求項1に記載の方法。

【請求項25】

複数の区画内に分配される標的化されたDNA分子のコピー体に一つ以上の区画固有DNA配列タグを付着させる方法であって、

複数の希釈された鋳型オリゴヌクレオチド(DTO)分子を含有する水溶液を製造する段階であって、DTO分子は、共通配列が両側にある縮重または部分的に縮重タグ配列を含む段階と；

前記溶液を複数の区画に分配する段階と；

DTO分子をDTO分子の共通配列を標的化するプライマーを使用するPCRによりコピーし、増幅させて、区画内に同じタグ配列を有する複数のクローンDTOコピー体をもたらす段階と；

クローン的に増幅されたDTOコピー体を区画内で標的化されたDNA分子のコピー体に一つ以上の区画固有配列タグを付着させるためのプライマーとして使用する段階と；を含む方法。

【請求項26】

前記DTO分子の濃度は、区画化の前に調整され、区画が増幅前に区画当たり平均10個未満のDTO分子を含有する、請求項25に記載の方法。

【請求項27】

前記DTO分子の濃度は、区画化の前に調整され、区画が増幅前に区画当たり平均2～3個のDTO分子を含有する、請求項25に記載の方法。

【請求項28】

前記DTO増幅の特異性は、遮断要素が全体的にまたは部分的に相補的DNA鎖にプライマーが混成化されながら、除去される時まで延長され得ないプライマーを使用することによって増加する、請求項25に記載の方法。

【請求項29】

標的化されたDNA分子は、区画化の前にDTO分子と同じ溶液に溶解される、請求項25に記載の方法。

【請求項30】

DTO分子を増幅させるために使用されたプライマーは、区画化の前にDTO分子と同じ溶液に溶解される、請求項25に記載の方法。

【請求項31】

DTO分子を増幅させるために使用されたプライマーは、同等でない濃度であり、標的化されたDNA分子に混成化するために活用され得る一本鎖DTOコピー体の生成をもたらす、請求項25に記載の方法。

【請求項 3 2】

複数のサンプルからの標的化されたRNAを並列で定量する方法であって、
複数のRNAサンプルを得る段階と；

モジュール式オリゴヌクレオチド合成戦略を使用してプライマーを合成する段階であって、合成は、複数の異なる標的的特異的プライマーが形成された後、停止され、部分的に合成されたプライマーが混合され、複数の別途の体積に分配された後、少なくとも特有のサンプル特異的タグ配列をそれぞれの別途の体積のプライマー混合物に添加するために合成が再開される段階と；

逆転写時に、各サンプルにおいて、モジュール式に合成されたプライマーを使用して、標的化されたRNAからコピーされたサンプル特異的タグを一定の割合で相補的DNAに付与する段階と；

すべてのサンプルからタグ付き相補的DNAをプールし、精製する段階と；

エンドポイントPCRによりそれぞれの相補的DNA標的を別に増幅させる段階と；

増幅生成物をプールし、シーケンシングする段階と；

異なる標的配列に関連したサンプル特異的タグの数をカウンティングし、すべてのサンプルにわたる標的化されたRNAの相対存在量を測定する段階と；

を含む方法。

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2016/017920
A. CLASSIFICATION OF SUBJECT MATTER C12Q 1/68(2006.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) C12Q 1/68; C07H 21/02; C12P 19/34		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models Japanese utility models and applications for utility models		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS(KIPO internal) & keywords: RNA quantitation, counting tag, reverse transcription, complementary DNA, competitive end-point PCR		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2014-152155 A1 (THE BROAD INSTITUTE, INC. et al.) 25 September 2014 See claims 1, 8, 24.	1
A	US 7270983 B1 (BALLANTYNE et al.) 18 September 2007 See claims 1, 7.	1
A	US 2006-0019258 A1 (YEAKLEY) 26 January 2006 See claims 1, 17-19 and figure 1.	1
A	US 2004-0067492 A1 (PENG et al.) 08 April 2004 See claims 1, 4, 27-28, 47.	1
A	US 8088580 B2 (KINOSHITA et al.) 03 January 2012 See claims 1-6, 8, 10-11, 13 and figures 4, 18-19, 21-22.	1
A	US 8435742 B2 (SHIRAI et al.) 07 May 2013 See claims 1, 10 and figure 6.	1
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 08 July 2016 (08.07.2016)		Date of mailing of the international search report 08 July 2016 (08.07.2016)
Name and mailing address of the ISA/KR International Application Division Korean Intellectual Property Office 189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea Facsimile No. +82-42-481-8578		Authorized officer KIM, Seung Beom Telephone No. +82-42-481-3371

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2016/017920

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2014-152155 A1	25/09/2014	US 2016-0024572 A1	28/01/2016
US 7270983 B1	18/09/2007	US 2008-0050741 A1 US 7582435 B2	28/02/2008 01/09/2009
US 2006-0019258 A1	26/01/2006	WO 2006-014625 A1	09/02/2006
US 2004-0067492 A1	08/04/2004	EP 1556511 A2 EP 1556511 A4 JP 2006-501844 A JP 4718181 B2 WO 2004-033629 A2 WO 2004-033629 A3	27/07/2005 30/04/2008 19/01/2006 06/07/2011 22/04/2004 22/07/2004
US 8088580 B2	03/01/2012	US 2009-0137406 A1 WO 2007-141912 A1	28/05/2009 13/12/2007
US 8435742 B2	07/05/2013	CN 101835905 A CN 101835905 B EP 2206793 A1 EP 2206793 A4 JP 2009-112195 A JP 5299986 B2 KR 10-2010-0075524 A US 2011-0053150 A1 WO 2009-057330 A1	15/09/2010 20/03/2013 14/07/2010 02/02/2011 28/05/2009 25/09/2013 02/07/2010 03/03/2011 07/05/2009

フロントページの続き

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