

【公報種別】特許法第 17 条の 2 の規定による補正の掲載

【部門区分】第 1 部門第 1 区分

【発行日】平成27年6月11日 (2015.6.11)

【公表番号】特表2014-523233(P2014-523233A)

【公表日】平成26年9月11日 (2014.9.11)

【年通号数】公開・登録公報2014-049

【出願番号】特願2014-505406(P2014-505406)

【国際特許分類】

C 1 2 Q 1/68 (2006.01)

C 1 2 N 15/09 (2006.01)

【F I】

C 1 2 Q 1/68 Z

C 1 2 N 15/00 A

【手続補正書】

【提出日】平成27年4月13日 (2015.4.13)

【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】

生物のゲノム DNA を解析する方法であって、

(a) 生物のゲノム DNA の異なるポリヌクレオチドを含むサンプルを分注して、各アリコートが前記異なるポリヌクレオチドのいくつかを含み、ゲノム DNA のハプロイドゲノム相当物よりも少なく含む、複数のアリコートを生成することと、

各アリコート中の前記ポリヌクレオチドを断片化して、各アリコート中のポリヌクレオチドの断片を生成することと、

前記断片をアリコート特異的タグ配列でタグ付けすることと、

(b) 各アリコート中の断片を配列決定して、各アリコートからゲノム配列およびタグ配列を含む複数のリードを生成することと

(c) 前記リードのアリコート特異的タグ配列を使用して各アリコートからの前記リードをアセンブルして、生物のゲノムの少なくとも一部に対応し、70 パーセント以上のコールレートで、メガベースあたり 1 以下の偽の単一ヌクレオチド変異体を含む、アセンブルされた配列を生成することと

を含む、方法。

【請求項 2】

ゲノムが哺乳類ゲノムであり、アセンブルされた配列が 70 パーセント以上のゲノムコールレートおよび 70 パーセント以上のエキソームコールレートを有する、請求項 1 に記載の方法。

【請求項 3】

ゲノムが少なくとも 1 ギガベースを含む、請求項 1 に記載の方法。

【請求項 4】

アリコート特異的タグ配列がポリヌクレオチドである、請求項 1 に記載の方法。

【請求項 5】

アリコート特異的タグ配列がエラー修正コードまたはエラー検出コードを含む、請求項 4 に記載の方法。

【請求項 6】



配列決定する前に、アリコート中の断片をプールすることをさらに含む、請求項 1 に記載の方法。

【請求項 7】

アセンブルされた配列がアセンブルされた配列の位置にベースコールを含み、ベースコールが 2 種以上のアリコートから出自する場合に、前記ベースコールを真実と同定することをさらに含む、請求項 1 から 6 のいずれかに記載の方法。

【請求項 8】

サンプルが 1 p g から 1 0 n g の間の ゲノム DNA を含む、請求項 1 から 6 のいずれかに記載の方法。

【請求項 9】

各アリコート中の前記 断片を増幅することをさらに含む、請求項 1 から 6 のいずれかに記載の方法。

【請求項 10】

ゲノム DNA が、ゲノム、エキソーム、メチローム、異なる生物のゲノムの混合物、ある生物の異なる細胞種のゲノムの混合物およびそれらのサブセットからなる群から選択されるものに対応する、請求項 1 に記載の方法。

【請求項 11】

アセンブルされた配列が ゲノムの 8 0 × 以下のカバレッジを有する、請求項 1 に記載の方法。

【請求項 12】

前記アセンブルされた配列の位置の塩基が、予備的ベースコールに基づいて、2 種以上のアリコートからの位置にコールされる、請求項 1 に記載の方法。

【請求項 13】

リードをアセンブルすることが、ゲノムの特定の位置の偽のベースコールを同定することを含み、前記偽のベースコールは、前記特定の位置にアリコートの第 2 の数に現れる異なるベースコールも含むアリコートの第 1 の数に現れるとき、アリコートの第 2 の数がアリコートの第 1 の数より大きい場合に同定される、請求項 1 に記載の方法。

【請求項 14】

リードをアセンブルすることが、ゲノムの特定の位置にある第 1 のベースコールの第 1 のスコアを決定することと、ここで、アリコートの第 1 の数に基づく第 1 のスコアが第 1 のベースコールを含む、第 1 のスコアを第 1 の閾値と比較することと、スコアが第 1 の閾値を上回るか、または下回るかどうかに基づいて、第 1 のベースコールが認められる、またはエラーであるかどうかを同定することを含む、請求項 1 に記載の方法。

【請求項 15】

スコアのすべてが第 1 の閾値を下回る場合、特定の位置がコールなしであると決定される、特定の位置でのその他のベースコールの 1 種または複数のその他のスコアを決定することをさらに含む、請求項 14 に記載の方法。

【請求項 16】

2 つのスコアが第 2 の閾値を超える場合、特定の位置がゲノムにおいてヘテロ接合性であると決定することをさらに含む、請求項 14 に記載の方法。

【請求項 17】

7 0 パーセント以上のコールレートで、メガベースあたり 1 より少ない偽の単一ヌクレオチド変異体を含む、哺乳類の ゲノム のアセンブルされた配列。

【請求項 18】

生物の ゲノム DNA を解析する方法であって、

- (a) 1 p g ~ 1 0 n g の ゲノム DNA を含むサンプルを用意することと、
- (b) ゲノム DNA を増幅して、増幅された核酸を生成することと、
- (c) 増幅された核酸を配列決定して、リードを生成することと、



(d) 前記リードをアセンブルして、生物のゲノムの少なくとも70パーセントのコールレートを有するアセンブルされた配列を生成することを含む、方法。

【請求項19】

サンプルが未精製である、請求項18に記載の方法。

【請求項20】

ゲノムDNAの増幅が多重置換増幅によって行われる、請求項18に記載の方法。

【請求項21】

ゲノムDNAを少なくとも1000倍増幅することを含む、請求項18に記載の方法。

【請求項22】

サンプルが、生物の1～20個の細胞または前記細胞から単離されたゲノムDNAを含む、請求項1から21のいずれか一項に記載の方法。

【請求項23】

ゲノムDNAおよび細胞夾雑物を含む細胞を溶解することを含み、ここで、ゲノムDNAの増幅が細胞夾雑物の存在下で行われる、請求項22に記載の方法。

【請求項24】

細胞が、生物の血液からの循環非血液細胞である、請求項22に記載の方法。

【請求項25】

アセンブルされた配列がメガベースあたり2以下の偽の単一ヌクレオチド変異体を含む、請求項18に記載の方法。

【請求項26】

サンプルのゲノムDNAを分注して、各アリコートがゲノムDNAのハプロイドゲノム相当物よりも少なく含む複数のアリコートを生成することを含み、ここで前記ゲノムDNAの増幅が各アリコートにおいて行われる、請求項18に記載の方法。

【請求項27】

各アリコート中の増幅された核酸を断片化して、各アリコート中の増幅された核酸の断片を生成することと、

各アリコート中の増幅された核酸の断片をアリコート特異的タグでタグ付けして、各アリコート中のタグ付けした断片を生成することとを含む、請求項26に記載の方法。

【請求項28】

前記アセンブルされた配列の位置の塩基が、予備的ベースコールに基づいて、2種以上のアリコートからの位置にコールされる、請求項27に記載の方法。

【請求項29】

リードをアセンブルすることが、ゲノムの特定の位置の偽のベースコールを同定することを含み、前記偽のベースコールは、前記特定の位置にアリコートの第2の数に現れる異なるベースコールも含むアリコートの第1の数に現れるとき、アリコートの第2の数がアリコートの第1の数より大きい場合に同定される、請求項27に記載の方法。

【請求項30】

リードをアセンブルすることが、ゲノムの特定の位置にある第1のベースコールの第1のスコアを決定することと、ここで、第1のアリコートに基づく第1のスコアが第1のベースコールを含む、第1のスコアを第1の閾値と比較することと、スコアが第1の閾値を上回るか、または下回るかどうかに基づいて、第1のベースコールが認められる、またはエラーであるかどうかを同定することを含む、請求項27に記載の方法。

【請求項31】

スコアのすべてが第1の閾値を下回る場合、特定の位置がコールなしであると決定される、特定の位置でのその他のベースコールの1種または複数のその他のスコアを決定することをさらに含む、請求項30に記載の方法。



## 【請求項 3 2】

2つのスコアが第2の閾値を超える場合、特定の位置がゲノムにおいてヘテロ接合性であることが決定されることを含む、請求項30に記載の方法。

## 【請求項 3 3】

ベースコールが2種以上のアリコートからのリード中に存在する場合に、アセンブルされた配列の位置でのベースコールが真実として認められる、請求項26に記載の方法。

## 【請求項 3 4】

生物がヒトである、請求項1から33のいずれかに記載の方法。

## 【請求項 3 5】

メガベースあたり1以下の偽の単一ヌクレオチド変異体および少なくとも70パーセントのコールレートを有するアセンブルされた全ヒトゲノム配列であって、1 pg ~ 10 ngの間のヒトゲノムDNAを配列決定することによって生成される、アセンブルされた全ヒトゲノム配列。

## 【請求項 3 6】

複数の染色体を含む生物のゲノムの配列変異体をフェージングする方法であって、  
前記複数の染色体各々のベクターフリー断片の混合物を含むサンプルを用意することと、  
ベクターフリー断片を配列決定して、複数のリードを生成することと、  
複数のリードをアセンブルして、複数の配列変異体を含むゲノム配列を得ることと、  
配列変異体をフェージングすることと、ここで、ゲノム配列は、メガベースあたり1より少ない偽の単一ヌクレオチド変異体を有する  
を含む、方法。

## 【請求項 3 7】

配列変異体の少なくとも70パーセントをフェージングすることを含む、請求項36に記載の方法。

## 【請求項 3 8】

ゲノム配列がゲノムの少なくとも70パーセントのコールレートを有する、請求項36に記載の方法。

## 【請求項 3 9】

サンプルが1 pg ~ 10 ngのゲノムを含む、請求項36に記載の方法。

## 【請求項 4 0】

ヒトゲノムを配列決定する方法であって、  
ヒトゲノムDNAのサンプルを分注して、各アリコートがヒトゲノムDNAを含む複数のアリコートを生成することと、  
各アリコートからの前記ヒトゲノムDNAを配列決定して、各アリコートから複数のリードを生成することと、  
各アリコートからの前記1種または複数のリードをアセンブルして、第1のアセンブルされた配列を生成することと、  
第1のアセンブルされた配列における複数の配列変異体を同定することと、  
配列変異体のうち少なくとも3種をフェージングすることと、  
エラーとして、前記少なくとも2種のその他の配列変異体のフェージングと一致しない配列変異体を同定し、それによって第2のアセンブルされた配列を生成することと  
を含み、第2のアセンブルされた配列が、70パーセント以上のコールレートでメガベースあたり1以下の偽の単一ヌクレオチド変異体を含む、方法。

## 【請求項 4 1】

ヒトゲノムを配列決定する方法であって、  
ヒトゲノムDNAのサンプルを分注して、各アリコートがヒトゲノムDNAを含む複数のアリコートを生成することと、  
各アリコート中の前記ヒトゲノムDNAを断片化して、各アリコート中のDNA断片を生成することと、



各アリコート中のDNA断片をアリコート特異的タグでタグ付けし、それによって、タグ付けされたDNA断片が出自するアリコートが決定可能となることと、

各アリコートからのタグ付けされたDNA断片を配列決定して、各アリコートから複数のリードを生成することと、

各アリコートからの前記1種または複数のリードをアセンブルして、70パーセント以上のコールレートでメガベースあたり1以下の偽の単一ヌクレオチド変異体を含むアセンブルされた配列を生成することと

を含み、前記アセンブルされた配列の位置の塩基が、予備的ベースコールに基づいて、2種以上のアリコートからの位置にコールされる、方法。

【請求項42】

ヒトゲノムを配列決定する方法であって、

ヒトゲノムDNAのサンプルを分注して、各アリコートがヒトゲノムDNAを含む複数のアリコートを生成することと、

各アリコート中の前記ヒトゲノムDNAを断片化して、各アリコート中のDNA断片を生成することと、

各アリコート中のDNA断片をアリコート特異的タグでタグ付けし、アリコート特異的タグによって、タグ付けされたDNA断片が出自するアリコートが決定可能となることと

、  
各アリコートからの前記タグ付けされたDNA断片を配列決定して、各アリコートから複数のリードを生成することと、

各アリコートからの前記1種または複数のリードをアセンブルして、第1のアセンブルされた配列を生成することと、

前記第1のアセンブルされた配列の位置に、2種以上のアリコートからの前記位置についての予備的ベースコールに基づいて、塩基をコールすること、ならびに

前記第1のアセンブルされた配列中の複数の配列変異体を同定すること、配列変異体のうち少なくとも3つをフェージングすること、および少なくとも2種のその他の配列変異体のフェージングと一致しない配列変異体をエラーと同定することによって、

第1のアセンブルされた配列中のエラーを低減して、第2のアセンブルされた配列を生成することと

を含み、第2のアセンブルされた配列が70パーセント以上のコールレートでメガベースあたり1以下の偽の単一ヌクレオチド変異体を含む、方法。

【請求項43】

操作を行うためにプロセッサを制御するための複数の命令を保存するコンピュータによって読み取り可能なメディアであって、前記命令は、

複数のアリコートからゲノムDNAの断片に対応する複数のリードを受け取ること、ここで、ゲノムDNAの各断片はアリコート特異的タグ配列でタグ付けされており、各リードはゲノムDNAの断片およびアリコート特異的タグ配列からの配列を含み、各アリコートはゲノムDNAのハプロイドゲノム相当物よりも少なく含む；

リードがアリコート特異的タグ配列を同定することにより出自するアリコートを決定することと、

前記リードのアリコート特異的タグ配列を使用して各アリコートからの前記リードをアセンブルして、生物のゲノムの少なくとも一部に対応し、70パーセント以上のコールレートで、メガベースあたり1以下の偽の単一ヌクレオチド変異体を含む、アセンブルされた配列を生成することと

を含む、コンピュータによって読み取り可能なメディア。