

【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5月28日 (2015.5.28)

【公表番号】特表2014-516514(P2014-516514A)

【公表日】平成26年7月17日 (2014.7.17)

【年通号数】公開・登録公報2014-038

【出願番号】特願2014-505385(P2014-505385)

【国際特許分類】

C 1 2 N 15/09 (2006.01)

C 1 2 Q 1/68 (2006.01)

G 0 6 F 19/22 (2011.01)

【 F I 】

C 1 2 N 15/00 Z N A A

C 1 2 Q 1/68 Z

G 0 6 F 19/22

【手続補正書】

【提出日】平成27年4月6日 (2015.4.6)

【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、

1 種または複数のコンピューティングデバイスが、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 の断片に由来する配列とアリコート特異的タグ配列を含み、各アリコートはゲノム DNA の 1 つ未満のハプロイドゲノム相当物を含有する；

アリコート特異的タグ配列を同定することによって、リードが出自するアリコートを決定すること；

1 種または複数のコンピューティングデバイスが、

生物のゲノムの少なくとも一部分に対応する複数のヘテロ接合性遺伝子座を同定すること、および

前記複数のヘテロ接合性遺伝子座をフェージングして、第 1 のハプロタイプと第 2 のハプロタイプを生成すること、ここでフェージングは、前記複数のヘテロ接合性遺伝子座に対応するリードの出自のアリコートをを用いて、前記ヘテロ接合性遺伝子座のどのアレルが同一ハプロタイプ上にあるかを決定し、フェージングされた配列は生物のゲノムの少なくとも一部分に対応するものである、

によって、リードからフェージングされた配列を生成すること、を含む、方法。

【請求項 2】

前記複数のヘテロ接合性遺伝子座をフェージングすることが、

少なくとも 3 種のヘテロ接合性遺伝子座をフェージングすること；および

前記少なくとも 2 種のヘテロ接合性遺伝子座のフェージングと一致しない第 1 の遺伝子座の第 1 の配列変異体をエラーと同定すること、を含む、請求項 1 に記載の方法。

【請求項 3】

前記ヘテロ接合性遺伝子座の少なくとも70パーセントがフェージングされる、請求項1に記載の方法。

【請求項 4】

領域がショートタンデムリピートを含む、請求項1に記載の方法であって、
前記領域の第1のハプロタイプのリードの第1の数を決定すること；
前記領域の第2のハプロタイプのリードの第2の数を決定すること；
前記第1の数を、前記第2の数と比較すること；および
比較に基づいて、第1のハプロタイプまたは第2のハプロタイプにおけるショートタンデムリピートの拡大を同定すること、
をさらに含む、請求項1に記載の方法。

【請求項 5】

1種または複数のコンピューティングデバイスが、ゲノムの重複領域にアラインする複数のアセンブルされた配列を生成すること、ここで前記重複領域における各アセンブルされた配列は、異なるアリコートに対応し、前記複数のヘテロ接合性遺伝子座はN個のヘテロ接合性遺伝子座を含み、ここでNは1より大きい整数である；
をさらに含む、請求項1に記載の方法であって、
前記複数の遺伝子座をフェージングすることが、
それぞれのアセンブルされた配列についてN個のヘテロ接合性遺伝子座のアレルに基づき2Nから4Nの可能性の空間においてアセンブルされた配列をクラスタリングし、それによって複数のクラスターを作製すること、および
最高密度を有する2つのクラスターを同定すること
を含む、方法。

【請求項 6】

前記フェージングすることが、
N次元のマトリックスを計算すること、ここで各次元はヘテロ接合性遺伝子座に対応し、各マトリックス要素は前記マトリックス要素に対応するアレルの組合せを有する多数のアセンブルされた配列に対応する；
それぞれ前記2つのクラスターのうちの1つの中心である、第1のマトリックス要素および第2のマトリックス要素を同定すること；
前記第1のマトリックス要素からN個のヘテロ接合性遺伝子座の第1のハプロタイプを決定すること；および
前記第2のマトリックス要素からN個のヘテロ接合性遺伝子座の第2のハプロタイプを決定すること、
をさらに含む、請求項5に記載の方法。

【請求項 7】

生物が2倍体哺乳類である、請求項1に記載の方法であって、
フェージングされた配列を用いて第1および第2のハプロタイプについてアセンブルされた配列を生成すること、ここで前記アセンブルされた配列は、70%以上のゲノムコールレートおよび70%以上のエキソームコールレートを含む、
をさらに含む、請求項1に記載の方法。

【請求項 8】

アセンブルされた配列が、メガベースあたり1未満の偽の単一ヌクレオチド変異体を含む、請求項7に記載の方法。

【請求項 9】

アセンブルされた配列が、ギガベースあたり600未満の偽の単一ヌクレオチド変異体を含む、請求項7に記載の方法。

【請求項 10】

前記アセンブルされた配列の一定位置に、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；および

ベースコールが前記 2 種以上のアリコートからのリード中に 3 回以上存在する場合に、
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 7 に記載の方法。

【請求項 1 1】

前記複数のヘテロ接合性遺伝子座をフェージングすることが、
ヘテロ接合性遺伝子座の複数の対のそれぞれについて、
前記対のヘテロ接合性遺伝子座のアレル間で共有される多数のアリコートのマトリックスを決定すること、ここで前記対のヘテロ接合性遺伝子座は互いに特定の距離内に位置している、
を含む、請求項 1 に記載の方法。

【請求項 1 2】

前記複数のヘテロ接合性遺伝子座をフェージングすることが、
前記各マトリックスを用いて、ヘテロ接合性遺伝子座のそれぞれの対についてスコアおよび配向を計算すること；および
前記スコアおよび前記配向を用いて、第 1 のハプロタイプおよび第 2 のハプロタイプを決定すること、
をさらに含む、請求項 1 1 に記載の方法。

【請求項 1 3】

前記スコアおよび前記配向を用いて、前記ハプロタイプを決定することが、
前記スコアおよび前記配向に基づいて、ヘテロ接合性遺伝子座の対間の結合のグラフを最適化することを含む、請求項 1 2 に記載の方法。

【請求項 1 4】

前記複数のヘテロ接合性遺伝子座からフェージングされた SNP を同定すること、ここで前記フェージングされた SNP は第 1 のアレルおよび第 2 のアレルを有する；
前記フェージングされた SNP に隣接する遺伝子座であって、コールなしとしての遺伝子座を同定すること、ここで前記遺伝子座は第 3 のアレルおよび第 4 のアレルを有するリードを有する；
前記フェージングされた SNP に前記第 1 のアレルおよび前記遺伝子座に前記第 3 のアレルを含む、共有されるアリコートの第 1 の数を計算すること；および
前記共有されるアリコートの第 1 の数に基づいて前記第 3 の塩基が前記遺伝子座にあると決定すること、
をさらに含む、請求項 1 に記載の方法。

【請求項 1 5】

共有されるアリコートの前記第 1 の数が閾値を超え、前記閾値が 2 以上である場合に、
前記第 3 の塩基が前記遺伝子座にあると決定すること、
をさらに含む、請求項 1 4 に記載の方法。

【請求項 1 6】

前記フェージングされた SNP に前記第 2 のアレルおよび前記遺伝子座に前記第 3 のアレルを含む、共有されるアリコートの第 2 の数を計算すること；
前記フェージングされた SNP に前記第 1 のアレルおよび前記遺伝子座に前記第 4 のアレルを含む共有されるアリコートの第 3 の数を計算すること；
前記第 1 の数および前記第 2 の数が閾値よりも大きく、前記第 3 の数が前記閾値未満である場合に、前記遺伝子座が前記第 3 のアレルについてホモ接合性であると決定すること、
をさらに含む、請求項 1 4 に記載の方法。

【請求項 1 7】

前記フェージングされた SNP に前記第 2 のアレルおよび前記遺伝子座に前記第 4 のアレルを含む、共有されるアリコートの第 2 の数を計算すること；および
前記第 3 のアレルを有するリードの全てが前記第 1 のアレルを有するアリコートを共有し、かつ前記第 4 のアレルを有するリードの全てが前記第 2 のアレルを有するアリコート

を共有する場合に、前記遺伝子座が前記第 3 のアレルおよび前記第 4 のアレルにつ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 1 4 に記載の方法。

【請求項 1 8】

前記複数のヘテロ接合性の遺伝子座の少なくとも 7 0 % をフェージングすること、
をさらに含む、請求項 1 に記載の方法。

【請求項 1 9】

前記各アリコート特異的タグがエラー修正コードを含み、各リードが、正しいタグ配列データまたは 1 種もしくは複数のエラーを含む不正確なタグ配列データを含む、請求項 1 に記載の方法であって、

エラー修正コードを使用して、不正確なタグ配列データを修正し、それによって、修正されたタグ配列データおよび修正され得ないタグ配列データを生成すること、

タグ配列データを必要とし、かつ第 1 のアウトプットを生成する第 1 のコンピュータプロセスにおいて、正しいタグ配列データおよび修正されたタグ配列データを含むリードを使用すること、および

タグ配列データを必要とせず、かつ第 2 のアウトプットを生成する第 2 のコンピュータプロセスにおいて、修正され得ないタグ配列データを含むリードを使用すること、
をさらに含む、方法。

【請求項 2 0】

前記第 1 のコンピュータプロセスが、サンプルマルチプレックス化、ライブラリーマルチプレックス化、フェージングおよびタグ配列データを使用するエラー修正プロセスを含むリストから選択され、第 2 のコンピュータプロセスが、マッピング、アセンブリおよびプールに基づく統計学を含む、請求項 1 9 に記載の方法

【請求項 2 1】

前記フェージングされた配列が生物のゲノムの第 1 の領域のものであり、前記第 1 の領域がショートタンデムリピートを含む、請求項 1 に記載の方法であって、

前記領域の第 1 のハプロタイプのリードを、第 2 のハプロタイプのリードと比較すること、および

比較に基づいて、第 1 のハプロタイプまたは第 2 のハプロタイプの一方におけるショートタンデムリピートの拡大を同定すること、
をさらに含む、方法。

【請求項 2 2】

前記ゲノム DNA が、前記生物のゲノム、前記生物のエキソーム、前記生物を含む種々の生物のゲノムの混合物、前記生物の種々の細胞種のゲノムの混合物、およびそのサブセットからなる群に由来する、請求項 1 に記載の方法。

【請求項 2 3】

前記生物が哺乳類である、請求項 1 に記載の方法。

【請求項 2 4】

前記生物がヒトである、請求項 1 に記載の方法。

【請求項 2 5】

前記ゲノム DNA を前記複数のアリコートに分注すること、

前記各アリコートにおけるゲノム DNA の断片を対応するアリコート特異的タグ配列でタグ付けすること、および

前記各アリコートにおいてゲノム DNA の断片を増幅すること、
をさらに含む、請求項 1 に記載の方法。

【請求項 2 6】

プロセッサが操作を実行するのを制御するための複数の命令であって、請求項 1 ~ 2 5 のいずれか一項に記載の工程を含む命令を保存する、コンピュータによって読み取り可能なメディア。