

【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19年6月14日(2007.6.14)

【公表番号】特表2006-525814(P2006-525814A)

【公表日】平成18年11月16日(2006.11.16)

【年通号数】公開・登録公報2006-045

【出願番号】特願2006-514923(P2006-514923)

【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 19/00 (2006.01)

【F I】

C 1 2 N 15/00 Z N A A

C 1 2 Q 1/68 A

G 0 6 F 19/00 6 0 0

【手続補正書】

【提出日】平成19年4月11日(2007.4.11)

【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】

複数の核酸分子であって、

(a) 該複数はNの核酸分子よりなり；

(b) 該複数の核酸分子の各々は、Z塩基対のゲノム中の配列に特異的にハイブリダイズするヌクレオチド配列を有し；および

(c) 該複数の核酸分子の少なくともP%は、

(i) Kヌクレオチドの長さを有し；

(i i) 該ゲノムに由来する表示に存在する、または存在すると予測される少なくとも1つの核酸分子に特異的にハイブリダイズし、該表示は該ゲノムの複雑性のR%超を有し；および

(i i i) 該ゲノムに対して L_1 ヌクレオチドのX以下の正確なマッチおよび該ゲノムに対して L_1 ヌクレオチドのY以上の正確なマッチを有し；かつ

(A) $N \geq 500$ ；

(B) $Z \geq 1 \times 10^8$ ；

(C) $300 \geq K \geq 30$ ；

(D) $70 \geq R \geq 0.001$ ；

(E) $P = (N \times R + (3 \times \text{シグマ})) / N$ ；

(F) シグマは $(N \times R \times (1 - R))$ の平方根であり；

(G) $(\log_4(Z) + 2)$ に最も近い整数 $\geq L_1 \geq \log_4(Z)$ に最も近い整数；

(H) Xは $D_1 \times (K - L_1 + 1)$ に最も近い整数であり；

(I) Yは $D_2 \times (K - L_1 + 1)$ に最も近い整数であり；

(J) $1.5 \geq D_1 \geq 1$ ；および

(K) $1 \geq D_2 \geq 0.5$

である該核酸分子。

【請求項 2】

N が少なくとも 500；少なくとも 1,000；少なくとも 2,500；少なくとも 5,000；少なくとも 10,000；少なくとも 25,000；少なくとも 50,000；少なくとも 85,000；少なくとも 190,000；少なくとも 350,000；および少なくとも 550,000 の核酸分子よりなる群から選択される請求項 1 記載の複数の核酸分子。

【請求項 3】

Z が少なくとも 3×10^8 、少なくとも 1×10^9 、少なくとも 1×10^{10} および少なくとも 1×10^{11} よりなる群から選択される請求項 1 記載の複数の核酸分子。

【請求項 4】

前記ゲノムが哺乳動物ゲノムである請求項 1 記載の複数の核酸分子。

【請求項 5】

前記ゲノムがヒトゲノムである請求項 4 記載の複数の核酸分子。

【請求項 6】

R が 0.001、1、2、4、10、15、20、30、40、50 および 70 よりなる群から選択される請求項 1 記載の複数の核酸分子。

【請求項 7】

P が少なくとも 70、少なくとも 80、少なくとも 90、少なくとも 95、少なくとも 97、および少なくとも 99 よりなる群から選択される請求項 1 記載の複数の核酸分子。

【請求項 8】

D_1 が 1 である請求項 1 記載の複数の核酸分子。

【請求項 9】

D_2 が 1 である請求項 1 記載の複数の核酸分子。

【請求項 10】

L_1 が 15、16、17、18、19、20、21、22、23 および 24 よりなる群から選択される請求項 1 記載の複数の核酸分子。

【請求項 11】

前記複数の核酸分子の前記 P % の各々が、さらに、前記ゲノムに対して L_2 ヌクレオチドの A 以下の正確なマッチ、および該ゲノムに対して L_2 ヌクレオチドの B 以上の正確なマッチを有し；および

ここで、

(a) $L_1 > L_2 \geq \log_4(Z) - 3$ に最も近い整数；

(b) A は $D_3 \times ((K - L_2 + 1) \times (Z / 4^{L_2}))$ に最も近い整数であり；

(c) B は $D_4 \times ((K - L_2 + 1) \times (Z / 4^{L_2}))$ に最も近い整数であり；

(d) $4 \geq D_3 \geq 1$ ；および

(e) $1 > D_4 \geq 0.5$

である請求項 1 記載の複数の核酸分子。

【請求項 12】

$D_3 \leq 3$ 、2 または 1.5 である請求項 11 記載の複数の核酸分子。

【請求項 13】

前記複数の核酸分子の前記 P % が、前記表示に存在する、または存在すると予測される少なくとも 1 つの核酸分子に対して、少なくとも 91、92、93、94、95、96、97、98、99 または 100 % の配列同一性を有する請求項 1 記載の複数の核酸分子。

【請求項 14】

K が 40、50、60、70、80、90、100、110、120、140、160、180、200 および 250 よりなる群から選択される請求項 1 記載の複数の核酸分子。

【請求項 15】

複数の核酸分子であって、

(a) 該複数の核酸分子は少なくとも 100 の核酸分子よりなり；

(b) 該複数の核酸分子の各々は少なくとも Z 塩基対のゲノム中の配列に少なくとも 9

0 % 同一であるヌクレオチド配列を有し；および

(c) 該複数の核酸分子の少なくとも P % は、

(i) Kヌクレオチドの長さ；

(i i) 該ゲノムに由来する表示に存在する、または存在すると予測される少なくとも 1 つの核酸分子に対して少なくとも 90 % 配列同一性であって、該表示は該ゲノムの複雑性の R % 以下を有する、配列同一性；および

(i i i) 該表示に対して L_1 ヌクレオチドの X 以下の正確なマッチ、および該表示に対して L_1 ヌクレオチドの Y 以上の正確なマッチ；を有し；および

(A) $Z \geq 1 \times 10^8$ ；

(B) $300 \geq K \geq 30$ ；

(C) $70 \geq R \geq 0.001$ ；

(D) $P \geq 90 - R$ ；

(E) $(\log_4 ((Z \times R) / 100) + 2)$ に最も近い整数 $\geq L_1 \geq \log_4 ((Z \times R) / 100)$ に最も近い整数；

(F) X は $D_1 \times (K - L_1 + 1)$ に最も近い整数であり；

(G) Y は $D_2 \times (K - L_1 + 1)$ に最も近い整数であり；

(H) $1.5 \geq D_1 \geq 1$ ；および

(I) $1 > D_2 \geq 0.5$ ；

である該複数の核酸分子。

【請求項 16】

少なくとも 500；少なくとも 1,000；少なくとも 2,500；少なくとも 5,000；少なくとも 10,000；少なくとも 25,000；少なくとも 50,000；少なくとも 85,000；少なくとも 190,000；少なくとも 350,000；または少なくとも 550,000 の核酸分子を含む請求項 15 記載の複数の核酸分子。

【請求項 17】

Z が少なくとも 3×10^8 、少なくとも 1×10^9 、少なくとも 1×10^{10} および少なくとも 1×10^{11} よりなる群から選択される請求項 15 記載の複数の核酸分子。

【請求項 18】

前記ゲノムが哺乳動物ゲノムである請求項 15 記載の複数の核酸分子。

【請求項 19】

前記ゲノムがヒトゲノムである請求項 18 記載の複数の核酸分子。

【請求項 20】

R が 0.001、1、2、4、10、15、20、30、40、50 および 70 よりなる群から選択される請求項 15 記載の複数の核酸分子。

【請求項 21】

P が少なくとも 70、少なくとも 80、少なくとも 90、少なくとも 95、少なくとも 97 および少なくとも 99 よりなる群から選択される請求項 15 記載の複数の核酸分子。

【請求項 22】

D_1 が 1 である請求項 15 記載の複数の核酸分子。

【請求項 23】

D_2 が 1 である請求項 15 記載の複数の核酸分子。

【請求項 24】

L_1 が 15、16、17、18、19、20、21、22、23 および 24 よりなる群から選択される請求項 15 記載の複数の核酸分子。

【請求項 25】

前記複数の核酸分子の前記 P % の各々が、さらに、前記ゲノムに対して L_2 ヌクレオチドの A 以下の正確なマッチ、および該ゲノムに対して L_2 ヌクレオチドの B 以上の正確なマッチを有し；および

(a) $L_1 > L_2 \geq \log_4 (Z) - 3$ に最も近い整数；

- (b) A は $D_3 \times ((K - L_2 + 1) \times (Z / 4^{L_2}))$ に最も近い整数であり；
- (c) B は $D_4 \times ((K - L_2 + 1) \times (Z / 4^{L_2}))$ に最も近い整数であり；
- (d) $4 \geq D_3 \geq 1$ ；および
- (e) $1 > D_4 \geq 0.5$ ；

である請求項 15 記載の複数の核酸分子。

【請求項 26】

前記複数の核酸分子の前記 P % が前記表示に存在する、または存在すると予測される少なくとも 1 つの核酸分子に対して少なくとも 91、92、93、94、95、96、97、98、99 または 100 % の配列同一性を有する請求項 15 記載の複数の核酸分子。

【請求項 27】

K が 40、50、60、70、80、90、100、110、120、140、160、180、200 および 250 よりなる群から選択される請求項 15 記載の複数の核酸分子。

【請求項 28】

前記表示が前記ゲノムの配列特異的切断によって生じる請求項 1 記載の複数の核酸分子。

【請求項 29】

配列特異的切断が制限エンドヌクレアーゼによって達成される請求項 28 記載の複数の核酸分子。

【請求項 30】

前記表示が化合物表示である請求項 1 記載の複数の核酸分子。

【請求項 31】

前記複数の核酸分子が固相の表面に固定化された請求項 1 記載の複数の核酸分子。

【請求項 32】

前記固相がナイロン膜、ニトロセルロース膜、ガラススライド、およびマイクロスフィアよりなる群から選択される請求項 31 記載の複数の核酸分子。

【請求項 33】

前記固相上の前記複数の核酸分子の位置が公知である請求項 31 記載の複数の核酸分子。

【請求項 34】

前記複数の核酸分子がマイクロアレイ上にある請求項 33 記載の複数の核酸分子。

【請求項 35】

前記複数の核酸分子がマイクロスフィアに固定化されている請求項 33 記載の複数の核酸分子。

【請求項 36】

(a) 請求項 1 記載の複数の核酸分子に試料をハイブリダイズさせ；次いで、
(b) 該複数の核酸分子のいずれに該試料がハイブリダイズするかを決定する；
ことを包含する、核酸試料を分析する方法。

【請求項 37】

前記試料が表示である請求項 36 記載の方法。

【請求項 38】

前記複数の核酸分子が固相の表面に固定化された請求項 36 記載の方法。

【請求項 39】

前記固相がナイロン膜、ニトロセルロース膜、ガラススライド、およびマイクロスフィアよりなる群から選択される請求項 38 記載の方法。

【請求項 40】

前記固相上の前記複数の核酸分子の位置が公知である請求項 38 記載の方法。

【請求項 41】

前記複数の核酸分子がマイクロアレイ上にある請求項 40 記載の方法。

【請求項 42】

前記複数の核酸分子がマイクロスフィアに固定化された請求項 38 記載の方法。

【請求項 43】

2つのゲノムの間のゲノム配列のコピー数変動を分析する方法であって、

(a) 第一のゲノムおよび第二のゲノムを供し；

(b) 少なくとも1つの同一制限酵素を用いて各ゲノムの検出可能に標識された表示を調製し；

(c) 該表示を請求項1または31記載の複数の核酸分子と接触させて、該表示および該複数の核酸分子の間のハイブリダイゼーションを可能とし；次いで、

(d) 該表示のハイブリダイゼーションのレベルを比較することを包含し、

ここに、該レベルの差が、該メンバーによって標的とされるゲノム配列に関して2つのゲノムの間のコピー数変動を示す、該方法。

【請求項44】

前記2つの表示が区別可能に標識された請求項43記載の方法。

【請求項45】

前記表示を前記複数の核酸分子と同時に接触させる請求項44記載の方法。

【請求項46】

2つのゲノムの間のゲノム配列のメチル化状態を比較する方法であって、

(a) 第一および第二のゲノムを提供する工程；

(b) 少なくとも1つの同一酵素を用いて各ゲノムの検出可能に標識された表示を調製し、ここで、該表示はメチル化感受性方法によって調製する工程；

(c) 該表示を請求項1または31記載の複数の核酸分子と接触させて、該表示および該複数の核酸分子の間のハイブリダイゼーションを可能とする工程；および

(d) 該表示のハイブリダイゼーションのレベルを比較する工程を包含する方法であって、

ここで、該レベルの差は、該メンバーによって標的化されたゲノム配列に関して2つのゲノムの間のメチル化状態の差を示す、方法。

【請求項47】

前記メチル化感受性方法が、第一の制限酵素を用いて第一の表示を、および第二の制限酵素を用いて第二の表示を調製する工程を包含し、ここで、該第一および第二の制限酵素は同一制限部位を認識するが、一方はメチル化感受性であって、他方はそうではない請求項46記載の方法。

【請求項48】

前記メチル化感受性方法が、メチル化されたゲノムに由来する表示が非メチル化ゲノムに由来する表示と区別できるように、非メチル化感受性制限酵素で表示を作成した後に、メチル-C配列の化学的切断を含む請求項46記載の方法。

【請求項49】

(a) Kヌクレオチドの長さ；

(b) 少なくともZ塩基対のゲノムに由来する表示に存在する、または存在すると予測される少なくとも1つの核酸分子に対して少なくとも90%の配列同一性；および

(c) 該ゲノムに対して L_1 ヌクレオチドのX以下の正確なマッチ、および該ゲノムに対して L_1 ヌクレオチドのY以上の正確なマッチ；

を有し、

ここで、

(i) $Z \geq 1 \times 10^8$ ；

(ii) $300 \leq K \leq 30$ ；

(iii) $(\log_4(Z) + 2)$ に最も近い整数 $\geq L_1 \geq \log_4(Z)$ に最も近い整数；

(iv) Xは $D_1 \times (K - L_1 + 1)$ に最も近い整数；

(v) Yは $D_2 \times (K - L_1 + 1)$ に最も近い整数；

(vi) $1.5 \geq D_1 \geq 1$ ；および

(vii) $1 > D_2 \geq 0.5$ ；

であるオリゴヌクレオチドを同定する方法であって、

(A) 該ゲノムを制限酵素で *in silico* 切断して、複数の予測された核酸分子を生じさせる工程、

(B) 各々が包括的に 200 ないし 1,200 塩基対の長さを有する予測された核酸分子を同定することによって該ゲノムの実質的表示を生じさせる工程；

(C) 包括的に 30 ないし 300 ヌクレオチドの長さ、および (B) における予測された核酸分子に対して少なくとも 90% の配列同一性を有するオリゴヌクレオチドを選択する工程；

(D) 該オリゴヌクレオチドで起こる L_1 ヌクレオチドの全てのストレッチを同定する工程；および

(E) 該ゲノムで起こる該ストレッチの各々の回数の数が (c) の要件を満たすことを確認する工程、を包含する、方法。

【請求項 50】

工程 (E) が：

前記ゲノムの圧縮されたトランスフォームを提供する工程；

該ゲノムに関連する情報を含む補助的データ構造を提供する工程；および

該圧縮されたトランスフォームおよび補助的データ構造を用いて L_1 ヌクレオチドについての単語カウントを決定する工程、
を包含する請求項 49 記載の方法。

【請求項 51】

工程 (E) が：

前記ゲノムの圧縮されたトランスフォームを提供する工程；

最後のヌクレオチドで開始し、第一のヌクレオチドまで反復当たり 1 文字で進行して、 L_1 ヌクレオチドの前記ストレッチの各ヌクレオチドを通じて反復する工程であって、ここに、特定の反復に対応するヌクレオチドは指標ヌクレオチドとして貯蔵される、工程であって、該反復は、さらに：

該トランスフォーム内の連続した範囲のヌクレオチドを示すサーチ領域を規定する工程；

該サーチ範囲で起こる該指標ヌクレオチドに先行するヌクレオチドの回数の数をカウントする工程、を包含し；

ここで、該指標ヌクレオチドに先行するヌクレオチドの出現が該サーチ範囲で起こらない場合、該反復を終了する；および

L_1 ヌクレオチドの該ストレッチの第一のヌクレオチドがカウントされる回数の数を出力する工程であって、この数は該ゲノムで出現する L_1 ヌクレオチドの該ストレッチの回数の数と同等である、工程、
を包含する請求項 49 記載の方法。

【請求項 52】

補助的データ構造を提供する工程をさらに包含し、該補助的データ構造は：

前記圧縮されたトランスフォーム中の特定の所定の位置まで、およびそれを含めて、該トランスフォーム中で出現した各ヌクレオチドのランニング合計を維持する K 間隔データ構造；および

速いルックアップアクセスを圧縮されたトランスフォームに提供する辞書カウントデータ構造を含み；

ここに、前記カウントする工程および前記規定する工程は、該補助的データ構造および該圧縮されたトランスフォームを用いて行われる請求項 51 記載の方法。

【請求項 53】

該トランスフォームは該カウンティングが実行される間は圧縮されたままである請求項 52 記載の方法。

【請求項 54】

前記圧縮解除されたトランスフォーム中の 3 つの文字ごとに圧縮されてバイトを形成するように該圧縮されたトランスフォームが圧縮され、該カウンティングは該反復の 1 つの間

にせいぜい1つのそのようなバイトを圧縮解除する請求項52記載の方法。

【請求項55】

前記ゲノムが少なくとも30億文字を含む請求項51記載の方法。

【請求項56】

前記圧縮されたトランスフォームが前記ゲノムのBurrows-Wheelerトランスフォームである請求項51記載の方法。

【請求項57】

前記トランスフォームに基づくデータを提供する工程をさらに包含し、ここに、前記規定する工程は該データおよび前記標識ヌクレオチドを用いて前サーチ領域を規定する工程を包含する、請求項51記載の方法。

【請求項58】

さらに、

前記トランスフォームに基づくデータを提供する工程；および

先行ヌクレオチドカウントを決定する工程であって、該先行ヌクレオチドカウントは、前記サーチ領域の開始前に該トランスフォームで起こる指標ヌクレオチドに先行するヌクレオチドの回数の数である、工程、
を包含する方法であって、

ここに、前記規定する工程は、該データ、該指標ヌクレオチド、および該先行ヌクレオチドカウントを用いて、該サーチ領域を規定する工程を包含する、請求項51記載の方法。

【請求項59】

前記先行ヌクレオチドカウントが、K間隔を用いて得られ、該K間隔は前記トランスフォームに沿った所定の位置に貯蔵され、特定の所定の位置まで、およびそれを含めて、該トランスフォームで出現した各ヌクレオチドのランニング合計を維持する、請求項58記載の方法。

【請求項60】

各々が請求項49記載の方法によって生産される複数のオリゴヌクレオチドであって、該複数は、少なくとも500オリゴヌクレオチドを含む、オリゴヌクレオチド。

【請求項61】

各々が請求項49記載の方法によって生産される複数のオリゴヌクレオチドであって、該複数は、少なくとも1,000オリゴヌクレオチド；少なくとも2,500オリゴヌクレオチド；少なくとも5,000オリゴヌクレオチド；少なくとも10,000オリゴヌクレオチド；少なくとも25,000オリゴヌクレオチド；少なくとも50,000オリゴヌクレオチド；少なくとも85,000オリゴヌクレオチド；少なくとも190,000オリゴヌクレオチド；少なくとも350,000オリゴヌクレオチド；または少なくとも550,000オリゴヌクレオチドを含む、オリゴヌクレオチド。