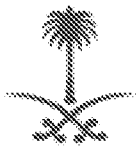
الهيئة السعودية للملكية الفكرية
Saudi Authority for Intellectual Property



المملكة العربية السعودية
Kingdom of Saudi Arabia

براءة اختراع

إن الرئيس التنفيذي للهيئة السعودية للملكية الفكرية و بموجب أحكام نظام براءات الإختراع و التصميمات التخطيطية للدارات المتكاملة و الأصناف النباتية و النماذج الصناعية الصادر بالمرسوم الملكي الكريم رقم م/27 وتاريخ 1425/05/29هـ و المعدل بقرار مجلس الوزراء رقم 536 و تاريخ 1439/10/19هـ ، و لائحته التنفيذية.
يقرر منح :

الومينا، إنك.
Illumina, Inc.

براءة اختراع رقم : SA 7780

بتاريخ : 1442/08/18 هـ
الموافق : 2021/03/31 م

عن الإختراع المسمى :

طرق وأنظمة لتحليل بيانات توالي الحمض النووي

METHODS AND SYSTEMS FOR ANALYZING NUCLEIC ACID SEQUENCING DATA

وفق ما هو موضح في وصف الاختراع المرفق، ولمالك البراءة الحق في الانتفاع بكامل الحقوق النظامية في المملكة العربية السعودية خلال فترة سريان الحماية.

الرئيس التنفيذي

د. عبدالعزيز بن محمد السويلم

الهيئة السعودية للملكية الفكرية
Saudi Authority for Intellectual Property

[19] **الهيئة السعودية للملكية الفكرية**
[11] **رقم البراءة:** SA 7780 B1

[12] **براءة اختراع**

[45] **تاريخ المنح:** 1442/08/18 هـ
الموافق: 2021/03/31 م

[21] رقم الطلب: 517381091
[22] تاريخ دخول المرحلة الوطنية: 1438/06/15 هـ
الموافق: 2017/03/14 م
[30] بيانات الأسبقية:
US 62/052.189 2014/09/18 م
[72] اسم المخترع: بروند جوسيلين، شليسينجير جوهان فيليكس
[73] مالك البراءة: الومينا، إنك.
عنوانه: 5200 لومينا واى، صن ديجو، كاليفورنيا 92122، كاليفورني، الولايات المتحدة الامريكية
جنسيته: امريكية
[74] الوكيل: شركة الهدف لخدمات العلامات المحدودة

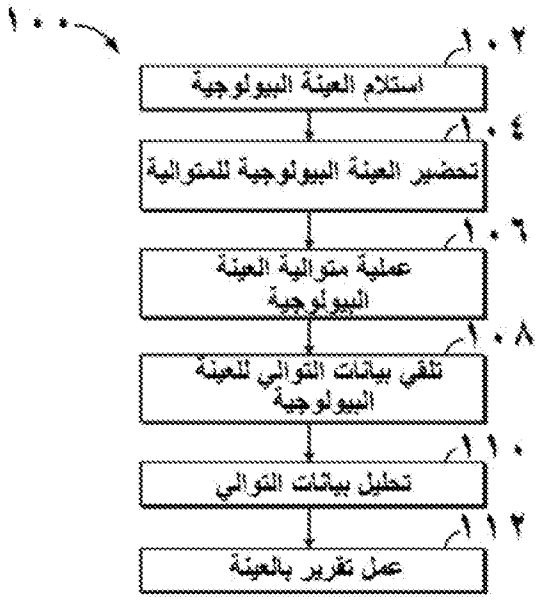
[86] رقم الطلب الدولي: PCT/US2015/050129
تاريخ إيداع الطلب الدولي: 2015/09/15 م
[87] رقم النشر الدولي: WO 2016/044233 A1
تاريخ النشر الدولي: 2016/03/24 م
[51] التصنيف الدولي(IPC®):
C12Q 001/068, G16B 045/000
[56] المراجع:
US 2014206547, EP 2287307
US 2014163900, WO 2013052557
WO 2013148970, US 2012053845

الفاحص: آمنه بنت ناصر الدوسري

[54] اسم الاختراع: طرق وأنظمة لتحليل بيانات توالي الحمض النووي
METHODS AND SYSTEMS FOR ANALYZING NUCLEIC ACID SEQUENCING DATA

[57] الملخص: طريقة تتضمن تلقي بيانات توالي شاملة مجموعة من قراءات عينة التي لها متواليات مناظرة لنيكلوتيدات nucleotides وتعيين قراءات العينة إلى المواقع المعينة. تتضمن الطريقة أيضا تحليل القراءات المعينة لكل موقع معين لتحديد المناطق محل الاهتمام regions-of-interest (ROIs) المناظرة في القراءات المعينة. يتضمن كل من المناطق محل الاهتمام واحد أو أكثر من سلسلة من بواعث تكرار repeat motifs. تتضمن الطريقة أيضا تصنيف القراءات المعينة بناء على متواليات المناطق محل الاهتمام بحيث تكون المناطق محل الاهتمام ذات المتواليات المختلفة معينة كألائل alleles محتملة مختلفة. تتضمن الطريقة أيضا تحليل، لمواقع معينة لها ألائل محتملة متعددة، متواليات الألائل المحتملة لتحديد ما إذا كانت الأليلة allele الأولى من الألائل المحتملة عبارة عن منتج مضطرب مشكوك فيه من أليلة ثانية من الألائل المحتملة. الشكل (1)

عدد عناصر الحماية (26)، عدد الأشكال (17)



طرق وأنظمة لتحليل بيانات توالي الحمض النووي

METHODS AND SYSTEMS FOR ANALYZING NUCLEIC ACID SEQUENCING DATA

الوصف الكامل

خلفية الاختراع

تم تحديد العديد من المواقع الجينية genetic loci بأنها مفيدة في تمييز الأفراد في مجموعة من الأنواع (مثلا، البشر) أو توفير معلومات مفيدة أخرى بشأن المجموعة أو الأفراد في المجموعة. مثلا، يمكن أن يكون لموقع جيني genetic locus عدد من الصور المتغيرة، المسماة ألائل، وكل فرد في مجموعة يمكن أن يتضمن واحد أو أكثر من الألائل alleles لموقع معين. يمكن أن تختلف أليلة allele لموقع عن ألائل أخرى لنفس الموقع في الطول (أي، العدد الكلي لنيكلوتيدات nucleotides) و/أو في متوالية النيكلوتيدات. هناك العديد من التطبيقات الجينية genetic applications التي تحلل ألائل المواقع الجينية. تتضمن هذه التطبيقات الجينية اختبار الأبوة، التعريف البشري (مثلا، تحليل شرعي forensic analysis)، مراقبة الخيمر chimera (مثلا، مراقبة زراعة النسيج tissue transplantation)، وغيرها من التطبيقات الجينية في بحوث النبات والحيوان. هناك العديد من التطبيقات الجينية التي تقوم بتحليل مواقع تتضمن تكرارات ترادفية قصيرة short tandem repeats (STRs) و/أو تعددات شكلية لنيكلوتيد فردي single nucleotide polymorphisms (SNPs). تكرارات ترادفية قصيرة عبارة عن مناطق تكرارية للحمض النووي الريبي منقوص الأكسجين Deoxyribonucleic acid (DNA) تتضمن بواعث تكرار repeat motifs. يمكن أن تكون بواعث التكرار، مثلا، عبارة عن اثنين إلى ست نيكلوتيدات في الطول، على الرغم من وجود بواعث تكرار من أحجام أخرى.

على الرغم من أن تحليل تكرار ترادفي قصير short tandem repeat (STR) و/أو تعدد شكلي لنيكلوتيد فردي nucleotide polymorphism (SNP) قد تحسن في السنوات الأخيرة، إلا أن هناك تحديات موجودة. مثلا، لم يتضمن تحاليل تحليل تكرار ترادفي قصير المتوالية الفعلية للنيكلوتيدات. يتم تحليل تكرارات ترادفية قصيرة بشكل مثالي باستخدام أنظمة الاستشراد الكهربي electrophoresis (CE) المعياري. تقوم أنظمة الاستشراد الكهربي فقط بتحديد طول الأليلة، مع ذلك، ولا تحدد متوالية الأليلة. بالتالي، من الممكن أن تشير بيانات الاستشراد الكهربي إلى أن الفرد متماثل

الجينات لأليلة معينة، في الواقع، عندما يتمتع الفرد باثنين من الألائل المختلفة التي لها نفس الطول ولكن المتواليات مختلفة.

يمكن أن تكون التحديات الخاصة بالتحكم في الجودة موجودة أيضا لأنظمة تقوم بتحليل متواليات الحمض النووي nucleic acid sequences. مثلا، تتضمن بعض الفحوص تحضير عينة بيولوجية biological sample، تكبير ألائل تكرار ترادفي قصير للعينة البيولوجية، وبعد ذلك توالي نواتج التكبير amplicons الناتجة. بعد تحضير وتكبير العينة، قد يكون من الممكن أن يتم تطوير واحد أو أكثر من نواتج التكبير خلال دايمر أولي primer dimer و/أو يتضمن الأحماض النووية nucleic acids من أكثر من مصدر واحد (مثلا، خيمرات chimeras) للحصول على البيانات غير الاعتمادية المناظرة. إذا كانت البيانات غير المطلوبة وتم ترشيحها، سوف يكون من الصعب، مثلا، توفير هيئة جينية genetic profile دقيقة للمصدر أو تحديد أن هناك العديد من المصادر. إذا كانت البيانات غير المطلوبة محددة، سوف يتم ترشيح البيانات والتخلص منها، ولكن دون المزيد من التحليل. بالمثل، يمكن أيضا أن تقوم الأخطاء التي تتم أثناء التوالي بجعل التحليل أصعب ويتم التخلص من هذه البيانات بشكل مثالي. أخيرا، يمكن أن يكون من الصعب تحديد نوع فرد من مصدر غير معروف.

بالتالي، هناك حاجة لطرق وأنظمة محسنة لتحليل بيانات توالي.

توفر البراءة الأمريكية 20140163900أ1 طرق وتركيبات ذات الصلة باستخدام مناطق التكرار الترادفي القصير للتطبيقات الجينية. توفر البراءة الأوروبية 2287307 مواد وطرق لتحديد وتحليل علامات الحمض النووي الريبي منقوص الأكسجين المتكررة الترادفية الوسيطة. توفر البراءة الأمريكية 2012053845 طرق وأنظمة لتحليل وتصحيح الأخطاء للمتواليات البيولوجية واستنتاجات العلاقات لعينات متعددة.

توفر البراءة الدولية 2013148970 طرق وأنظمة لتحديد وجود أو عدم وجود اختلال الصيغة الصبغية في الجنين. توفر البراءة الدولية 2013052557 طرق لتحديد حالة الصيغة الصبغية في الجنين عند الكروموسوم chromosome من عينة الحمض النووي الريبي منقوص الأكسجين من جنين. توفر البراءة الدولية 2014206547 طرق لتحديد المنطقة الجينومية genomic region بأكملها لموضع مستضد الكريات البيض البشرية HLA (HLA) معين بما في ذلك كل من

الإنترونات introns والإكسونات exons. يوفر الاختراع الحالي طرق وأنظمة محسنة لتحليل بيانات المتواليات.

الوصف العام للاختراع

في أحد النماذج، يتم توفير طريقة تتضمن تلقي بيانات توالي تتضمن مجموعة من قراءات عينة sample reads التي لها متواليات مناظرة لنيكلوتيدات. تتضمن الطريقة أيضا تعيين قراءات العينة إلى المواقع المعينة بناء على متوالية النيكلوتيدات، حيث قراءات العينة المعينة إلى موقع معين مناظر عبارة عن قراءات معينة للموقع المعين المناظر. تتضمن الطريقة أيضا تحليل القراءات المعينة لكل موقع معين لتحديد المناطق محل الاهتمام regions-of-interest (ROIs) المناظرة في القراءات المعينة. تتضمن كل من المناطق محل الاهتمام واحد أو أكثر من سلسلة من بواعث تكرار حيث كل باعث تكرار repeat motif لسلسلة مناظرة تتضمن مجموعة متطابقة من النيكلوتيدات. تتضمن الطريقة أيضا تصنيف، لمواقع معينة لها قراءات معينة متطابقة، القراءات المعينة بناء على متواليات المناطق محل الاهتمام بحيث تكون المناطق محل الاهتمام ذات المتواليات المختلفة معينة كألائل محتملة مختلفة. كل أليلة محتملة لها متوالية مختلفة عن متواليات ألائل محتملة أخرى في الموقع المعين. تتضمن الطريقة أيضا تحليل، لمواقع معينة لها ألائل محتملة متعددة، متواليات الألائل المحتملة لتحديد ما إذا كانت الأليلة الأولى من الألائل المحتملة عبارة عن منتج مضطرب مشكوك فيه من أليلة ثانية من الألائل المحتملة. الأليلة الأولى عبارة عن المنتج المضطرب المشكوك فيه للأليلة الثانية إذا كانت بواعث تكرار k في المتواليات المناظرة والتي تمت إضافتها أو إسقاطها بين الألائل الأولى والثانية، حيث k عبارة عن عدد كامل. اختياريا، k يساوي 1 أو 2.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي بيانات توالي لها مجموعة من قراءات عينة لنواتج تكبير التي تناظر مجموعة من مواقع جينية. قراءات العينة يتضمن أزواج قراءة حيث كل زوج قراءة لناتج تكبير amplicon مناظر يتضمن قراءات أولى وثانية لناتج التكبير المناظر. كل من القراءات الأولى والثانية لها متوالية قراءة مناظرة. تتضمن الطريقة أيضا تحديد مواقع جينية محتملة للقراءات الأولية بناء على تحليل متواليات القراءة للقراءات الأولية. المواقع الجينية المحتملة من مجموعة المواقع الجينية. تتضمن الطريقة أيضا تحديد، لكل للقراءات الأولية لها موقع محتمل واحد على الأقل، ما إذا كانت القراءة الأولى تتحاذى مع متوالية مرجعية لكل من المواقع الجينية

المحتملة. إذا كانت القراءة الأولى تتحاذى مع متوالية مرجعية لموقع جيني واحد فقط، تتضمن الطريقة تحديد أن القراءة الأولى تتضمن أليلة محتملة للموقع الجيني الواحد. إذا كانت القراءة الأولى تتحاذى مع أكثر من متوالية مرجعية reference sequence واحدة، تتضمن الطريقة تحديد أن القراءة الأولى تتضمن أليلة محتملة للموقع الجيني لها المتوالية المرجعية تتحاذى بشكل أفضل مع القراءة الأولى. إذا كانت القراءة الأولى لا تتحاذى مع متوالية مرجعية، تتضمن الطريقة تعيين القراءة الأولى كقراءة غير متحاذية وتحليل القراءة غير المتحاذية لتحديد موقع جيني من المواقع الجينية المحتملة التي تتلاءم بشكل أفضل مع القراءة غير المتحاذية. تتضمن الطريقة أيضا توليد هيئة جينية تتضمن أنماط جينية genotypes مسماة لمجموعة على الأقل من المواقع الجينية، حيث الأنماط الجينية المسماة قائمة على أساس الألائل المحتملة للمواقع الجينية المناظرة. الهيئة الجينية تتضمن أيضا واحد أو أكثر من إخطارات للمواقع الجينية لها قراءات غير متحاذية.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي بيانات توالي لها مجموعة من قراءات عينة لنواتج تكبير التي تناظر مجموعة من مواقع جينية. قراءات العينة يتضمن أزواج قراءة حيث كل زوج قراءة لناتج تكبير مناظر يتضمن قراءات أولى وثانية لناتج التكبير المناظر. كل من القراءات الأولى والثانية لها متوالية قراءة مناظرة. تتضمن الطريقة أيضا تحديد مواقع جينية محتملة للقراءات الأولية بناء على تحليل متواليات القراءة للقراءات الأولية. المواقع الجينية المحتملة من مجموعة المواقع الجينية. تتضمن الطريقة أيضا تحديد، لكل للقراءات الأولية لها موقع محتمل واحد على الأقل، ما إذا كانت القراءة الأولى تتحاذى مع متوالية مرجعية لكل من المواقع الجينية المحتملة. تتضمن الطريقة أيضا تعيين القراءات الأولية التي لا تتحاذى مع متوالية مرجعية كقراءات غير متحاذية. تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد موقع جيني من المواقع الجينية المحتملة التي تتلاءم بشكل أفضل مع القراءة غير المتحاذية. تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد ما إذا كان إسقاط الأليلة المحتملة موجود للموقع الجيني ذو الملائمة الأفضل.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي توزيع قراءة لكل موقع جيني لمجموعة من مواقع جينية. توزيع القراءة تتضمن مجموعة من ألائل محتملة، حيث كل أليلة محتملة لها متوالية أليلة وعدد قراءة. يمثل عدد القراءة عدد من قراءات عينة من بيانات توالي تم تحديدها على أنها تتضمن الأليلة المحتملة. يمكن أن تتضمن الطريقة أيضا تحديد، لكل موقع جيني لمجموعة مواقع جينية،

واحد من الألائل المحتملة لتوزيع القراءة التي لها عدد قراءة أقصى. يمكن أن تتضمن الطريقة أيضا تحديد، لكل موقع جيني لمجموعة مواقع جينية، ما إذا كان عدد القراءة الأقصى يتخطى قيمة تفسير حدية interpretation threshold. إذا كانت أقصى قراءة تتخطى قيمة التفسير الحدية، تتضمن الطريقة تحليل الأليلة المحتملة للموقع الجيني المناظر لتسمية نمط جيني للموقع الجيني. إذا كانت أقصى قراءة أقل من قيمة التفسير الحدية، تتضمن الطريقة توليد تنبيه أن الموقع الجيني له تغطية منخفضة. تتضمن الطريقة أيضا توليد هيئة جينية التي لها الأنماط الجينية لكل للمواقع الجينية التي تم من أجلها تسمية النمط الجيني والتنبيه للمواقع الجينية التي لها تغطية منخفضة.

في أحد النماذج، يتم توفير طريقة تتضمن، (أ) تلقي توزيع قراءة لموقع جيني. توزيع القراءة تتضمن مجموعة من ألائل محتملة، حيث كل أليلة محتملة لها متوالية أليلة ودرجة عد. درجة العد قائم على عدد من قراءات عينة من بيانات توالي تم تحديدها على أنها تتضمن الأليلة المحتملة. تتضمن الطريقة أيضا، (ب) تحديد ما إذا كان الموقع الجيني له تغطية منخفضة بناء على درجة العد لأليلة أخرى من الألائل المحتملة. إذا كان الموقع الجيني له تغطية منخفضة، تتضمن الطريقة توليد إخطار أن الموقع الجيني له تغطية منخفضة. إذا كان الموقع الجيني لا يتضمن تغطية منخفضة، تتضمن الطريقة تحليل درجات العد للألائل المحتملة لتحديد نمط جيني genotype للموقع الجيني. تتضمن الطريقة أيضا، (د) توليد هيئة جينية تتضمن النمط الجيني للموقع الجيني أو التنبيه أن الموقع الجيني له تغطية منخفضة.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي توزيع قراءة لموقع جيني. توزيع القراءة تتضمن مجموعة من ألائل محتملة، حيث كل أليلة محتملة لها متوالية أليلة allele sequence وعدد قراءة. يمثل عدد القراءة عدد من قراءات عينة من بيانات توالي والتي تم تعيينها إلى الموقع الجيني. يمكن أن تتضمن الطريقة أيضا تحديد درجة عد لكل للألائل المحتملة. يمكن أن تكون درجة العد قائمة على عدد القراءة للأليلة المحتملة. يمكن أن تتضمن الطريقة أيضا تحديد ما إذا كان درجات العد للألائل المحتملة تتخطى قيمة حدية تحليلية analytical threshold. إذا كانت درجة العد لأليلة محتملة مناظرة لا تتخطى القيمة الحدية التحليلية، تتضمن الطريقة التخلص من الأليلة المحتملة المناظرة. إذا كانت درجة العد لأليلة محتملة مناظرة يتخطى القيمة الحدية التحليلية، تتضمن الطريقة تعيين الأليلة المحتملة كأليلة معينة للموقع الجيني.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي توزيع قراءة لموقع جيني. توزيع القراءة تتضمن مجموعة من ألائل محتملة، حيث كل أليلة محتملة لها متوالية أليلة وعدد قراءة. يمثل عدد القراءة عدد من قراءات عينة من بيانات توالي والتي تم تعيينها إلى الموقع الجيني. تتضمن الطريقة أيضا تحديد ما إذا كان أعداد القراءة تتخطى قيمة حدية تحليلية. إذا كان عدد القراءة لأليلة محتملة مناظرة أقل من القيمة الحدية التحليلية، تتضمن الطريقة تعيين الأليلة المحتملة المناظرة كأليلة ضوضائية noise allele. إذا كان عدد القراءة لأليلة محتملة مناظرة يتخطى القيمة الحدية التحليلية، تتضمن الطريقة تعيين الأليلة المحتملة كأليلة للموقع الجيني. تتضمن الطريقة أيضا تحديد ما إذا كان إجمالي أعداد القراءة للألائل الضوضائية noise alleles تتخطى قيمة حدية ضوضائية. إذا كان الإجمالي تتخطى القيمة الحدية الضوضائية، تتضمن الطريقة توليد تنبيه أن الموقع الجيني يتضمن ضوضاء مفرطة.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي بيانات موقع لكل موقع جيني لمجموعة من مواقع جينية. بيانات الموقع تتضمن واحد أو أكثر من ألائل معينة للموقع الجيني المناظر. كل أليلة معينة قائم على أعداد القراءة التي تم الحصول عليها من بيانات توالي. تتضمن الطريقة أيضا تحديد، لكل موقع جيني لمجموعة مواقع جينية، ما إذا كان عدد ألائل معينة للموقع الجيني المناظر أكبر من عدد أقصى محدد سلفا للألائل المقبولة للموقع الجيني المناظر. يمكن أن تتضمن الطريقة توليد تنبيه عدد الألائل allele-number alert إذا كان عدد ألائل معينة تتخطى العدد الأقصى المحدد سلفا للألائل المقبولة. تتضمن الطريقة أيضا تحديد، لكل موقع جيني لمجموعة مواقع جينية، ما إذا كان تناسب أليلة من الألائل المعينة غير كافي. يمكن أن يكون تناسب الأليلة قائم على أعداد قراءة الألائل المعينة. يمكن أن تتضمن الطريقة أيضا توليد تنبيه تناسب أليلة allele-proportion alert إذا كان تناسب الأليلة غير متوازن. يمكن أن تتضمن الطريقة أيضا تحديد أن العينة تتضمن خليط من مجموعة من مصادر بناء على عدد تنبيه عدد الألائل وتنبيهات تناسب أليلة لمجموعة المواقع الجينية.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي بيانات موقع لمجموعة من مواقع Y-. بيانات الموقع يتضمن ألائل معينة للمواقع Y-. كل أليلة معينة قائم على أعداد القراءة التي تم الحصول عليها من بيانات توالي. تتضمن الطريقة أيضا مقارنة عدد ألائل معينة لكل موقع Y- بعدد غير متوقع من ألائل المواقع Y-. تتضمن الطريقة أيضا توليد توقع أن العينة عبارة عن ذكر

أو أنثى بناء على نتائج من عملية المقارنة. اختياريا، تتضمن المواقع الجينية مواقع تكرار ترادفي قصير ومواقع تعدد شكلي لنيكلوتيد فردي.

<u>شرح مختصر للرسومات</u>

الشكل 1 عبارة عن مخطط تدفق يشرح طريقة وفقا لأحد النماذج.

الشكل 2 عبارة عن مخطط تدفق يشرح طريقة تعيين أنواع مختلفة من قراءات عينة لعمليات التحليل المختلفة.

الشكل 3 عبارة عن مخطط بياني يشرح جزء من الطريقة وفقا للشكل 2.

الشكل 4 عبارة عن مخطط بياني يشرح كيف يمكن تحديد منطقة محل الاهتمام region of interest (ROI) وفقا لأحد النماذج.

الشكل 5 عبارة عن مخطط يبين العديد من أخطاء سوء المحاذاة التي يمكن أن تتم إذا كانت المنطقة المجنحة flanking region مجاورة مباشرة لتكرار ترادفي قصير يستخدم لبذر المحاذاة.

الشكل 6أ عبارة عن مجموعة من رسومات بيانية يبين تسميات تكرار ترادفي قصير فعلية مقارنة بنتائج نظرية بناء على مدخل عينة من خليط العينات.

الشكل 6ب عبارة عن مجموعة أخرى من رسومات بيانية يبين تسميات تكرار ترادفي قصير فعلية مقارنة بنتائج نظرية بناء على مدخل عينة من خليط العينات.

الشكل 6ج عبارة عن مجموعة أخرى من رسومات بيانية يبين تسميات تكرار ترادفي قصير فعلية مقارنة بنتائج نظرية بناء على مدخل عينة من خليط العينات.

الشكل 6د عبارة عن مجموعة أخرى من رسومات بيانية يبين تسميات تكرار ترادفي قصير فعلية مقارنة بنتائج نظرية بناء على مدخل عينة من خليط العينات.

الشكل 7 عبارة عن جدول يبين توافق لتسميات أليلة لمواقع معروفة لخمسة من عينات الحمض النووي الريبي منقوص الأكسجين ضابطة.

الشكل 8 عبارة عن مخطط تدفق يشرح طريقة of تحديد منتج مضطرب في قراءات عينة وفقا لأحد النماذج.

الشكل 9 تتضمن جدول يشرح أعداد قراءة لألائل محتملة للموقع D1S1656.

الشكل 10 تتضمن رسم بياني يقوم على البيانات الموجودة في الجدول وفقا للشكل 9.

الشكل 11 عبارة عن مخطط تدفق يشرح طريقة تحليل قراءات عينة لتحديد نمط جيني لواحد أو أكثر من مواقع جينية.

الشكل 12 عبارة عن مخطط تدفق يشرح طريقة توليد تقرير عينة تتضمن مجموعة من تسميات النمط الجيني.

الشكل 13 عبارة عن مخطط تدفق يشرح طريقة الكشف عن ما إذا كانت عينة تتضمن خليط المصادر.

الشكل 14 عبارة عن مخطط تدفق يشرح طريقة تحديد نوع عينة.

الشكل 15 يشرح نظام مكون وفقا لأحد النماذج الذي يمكن استخدامه لتنفيذ العديد من الطرق المذكورة أعلاه.

الشكل 16أ يشرح جزء من تقرير عينة sample report وفقا لواحد أو أكثر من النماذج.

الشكل 16ب يشرح جزء آخر من تقرير العينة وفقا لواحد أو أكثر من النماذج.

الشكل 17أ يشرح جزء من تقرير عينة وفقا لواحد أو أكثر من النماذج.

الشكل 17ب يشرح جزء آخر من تقرير العينة.

الشكل 17ج يشرح جزء آخر من تقرير العينة.

الشكل 17د يشرح جزء آخر من تقرير العينة.

الشكل 17هـ يشرح جزء آخر من تقرير العينة.

الشكل 17و يشرح جزء آخر من تقرير العينة.

الوصف التفصيلي:

يتضمن الطلب الحالي الموضوع المذكور المشابه لذلك الموضوع الموصوف في الطلب الأمريكي رقم 030867/2013 (الطلب الدولي رقم 142831/2014)، المودع في 15 مارس 2013 وبعنوان "طرق وأنظمة لمحاذاة عناصر الحمض النووي الريبي منقوص الأكسجين تكرارية،" والتي تم دمجها هنا بالإشارة إليها كمرجع.

يمكن أن تكون النماذج المذكورة أعلاه قابلة للتطبيق على تحليل متواليات الحمض النووي nucleic acid sequences لمتوالية تحديد الاختلافات. يمكن استخدام النماذج لتحليل ألائل محتملة لموقع جيني وتحديد نمط جيني للموقع الجيني أو، بمعنى آخر، توفير تسمية نمط جيني للموقع. في بعض الحالات، يمكن أن تقوم الطريقة وأنظمة المذكورة أعلاه بتوليد تقارير عينة أو هيئات عينة

تتضمن مجموعة من تسميات النمط الجيني هذه. يمكن أيضا أن تنطبق النماذج على مراقبة جودة فحص يتضمن توالي و/أو تحليل متواليات الحمض النووي، مثل تلك التي تتضمن اختلافات متوالية. يمكن أن تتضمن اختلافات المتوالية تعددات شكلية لنيكلوتيد فردي أو عناصر تكرارية متعددة الشكل، مثل تكرارات ترادفية قصيرة. يمكن وضع اختلافات المتوالية في مواقع جينية معينة، مثل تلك الموجودة في قاعدة بيانات نظام مؤشر الحمض النووي الريبي منقوص الأكسجين المدمج Combined DNA Index System (CODIS) أو غيرها من المستخدم في التحليل. مثلا، يمكن أن تتضمن اختلافات المتوالية تكرارات ترادفية قصيرة المختارة من مواقع تكرار ترادفي قصير صبغية جسدية لنظام مؤشر الحمض النووي الريبي منقوص الأكسجين المدمج، مواقع نظام مؤشر الحمض النووي الريبي منقوص الأكسجين المدمج واي-Y تكرار ترادفي قصير، مواقع إي يو EU تكرار ترادفي قصير صبغية جسدية، مواقع EU Y- تكرار ترادفي قصير وما شابه. النظام المؤشر الحمض النووي الريبي منقوص الأكسجين المدمج عبارة عن مجموعة من مواقع تكرار ترادفي قصير لبية محددة بواسطة مجموعة إف بي آى FBI وتتضمن 13 مواقع، سي بي إف 1 بي أوه CSF1PO، إف جي ايه FGA، تي اتش 01 TH01، تي بي أوه أكس TPOX، في دبليو ايه VWA، دي 3إس1358 D3S1358، دي5إس818 D5S818، دي7إس820 D7S820، دي8إس1179 D8S1179، دي13إس317 D13S317، دي16إس539 D16S539، دي18إس51 D18S51 و دي21إس11 D21S11 . يمكن أن تتضمن تكرارات ترادفية قصيرة الإضافية محل الاهتمام بي إي إن تي ايه دي PENTA D و بي إي إن تي ايه إي PENTA E، مع ذلك، يمكن تحليل تكرارات ترادفية قصيرة أخرى بواسطة النماذج المذكورة أعلاه. يمكن أن تكون تعددات شكلية لنيكلوتيد فردي في قواعد بيانات معروفة، مثل قاعدة بيانات المركز الوطني لمعلومات التقنية الحيوية National Center for Biotechnology Information (NCBI) دي بي إس إن بي dbSNP. يمكن تحديد تكرارات ترادفية قصيرة وتعددات شكلية لنيكلوتيد فردي في البحوث المستقبلية أيضا.

حسب الاستخدام هنا، يتضمن أو يمثل المصطلح "متوالية" جديلة لنيكلوتيدات مقترنة ببعضها البعض. يمكن أن تكون النيكلوتيدات بناء على الحمض النووي الريبي منقوص الأكسجين أو رنا RNA. يجب فهم أن متوالية واحدة يمكن أن تتضمن العديد من متواليات ثانوية. مثلا، عينة قراءة فردية (مثلا، لناتج تكبير تفاعل سلسلة بوليمرات polymerase chain reaction (PCR)) يمكن أن

تتضمن متوالية لها 350 نيكلوتيد. يمكن أن تتضمن عينة القراءة العديد من المتواليات الثانوية في هذه الـ 350 نيكلوتيد. مثلا، يمكن أن تتضمن عينة القراءة متواليات مجنحة أولى وثانية ثانوية لها، مثلا، 20-50 نيكلوتيد. يمكن وضع المتواليات المجنحة الأولى والثانية الثانوية على أي من جوانب القطاع التكراري الذي له متوالية ثانوية مناظرة (مثلا، 40-100 نيكلوتيد). يمكن أن تتضمن كل من المتواليات الثانوية المجنحة (أو تتضمن أجزاء من) متوالية بادئة ثانوية (مثلا 10-30 نيكلوتيد). لسهولة القراءة، سوف تتم الإشارة إلى المصطلح "متوالية ثانوية" بـ "متوالية،" ولكن سيتم فهم أن هناك اثنين من المتواليات بالضرورة غير منفصلة عن بعضها البعض على جديلة مشتركة. لتمييز المتواليات المختلفة الموضحة هنا، يمكن إعطاء المتواليات علامات مختلفة (مثلا، متوالية مستهدفة target sequence، متوالية بادئة primer sequence، متوالية مجنحة flanking sequence، متوالية مرجعية، وما شابه). يمكن استخدام مصطلحات أخرى، مثل "أليلة،" للتمييز بين الأشياء المتشابهة.

حسب الاستخدام هنا، يتضمن المصطلح "المنطقة محل الاهتمام" أو "ROI" قطاع تكراري لعينة القراءة تتضمن واحد أو أكثر من سلسلة من بواعث تكرار. يمكن أن تكون السلسلة من بواعث التكرار عبارة عن تكرار ترادفي قصير. في بعض النماذج، المنطقة محل الاهتمام عبارة عن القطاع التكراري فقط (مثلا، التكرار الترادفي القصير). في النماذج الأخرى، مع ذلك، يمكن أن يتضمن المنطقة محل الاهتمام متواليات ثانوية للمناطق المجنحة. مثلا، يمكن أن تتضمن المنطقة محل الاهتمام القطاع التكراري، حوالي 1-5 نيكلوتيدات للمنطقة المجنحة الأولى التي تمتد من أحد أطراف القطاع التكراري، وحوالي 1-5 نيكلوتيدات للمنطقة المجنحة الثانية التي تمتد من الطرف المقابل للقطاع التكراري.

يجب فهم أن القطاع التكراري غير مطلوب ليتضمن نفس البواعث فيه. يمكن أن يتضمن القطاع التكراري سلسلة من بواعث X-، بعد ذلك سلسلة من بواعث Y-، بعد ذلك سلسلة من بواعث Z- (أو سلسلة أخرى من بواعث X-)، إلخ. القطاع التكراري لـ [TAGA]11[TAGG]1[TG]5 عبارة عن مصال واحد محدد لما هو مذكور أعلاه. يجب فهم أن القطاع التكراري غير مطلوب أن يتضمن بواعث تكرار فيه. كما هو مبين في المثال أعلاه، يمكن أن يتضمن القطاع التكراري بواعث تكرار والذي تتم مقاطعته بواسطة تكرار غير باعث. [تي ايه جي جي TAGG] في المثال أعلاه عبارة عن هذا التكرار غير الباعث.

حسب الاستخدام هنا، يشير المصطلح "قيمة حدية" إلى نقطة يمكن عندها تغيير سياق التحليل و/أو نقطة يتم عندها بدء الأثر. القيمة الحدية غير مطلوب أن تكون محددة سلفا. بدلا من ذلك، يمكن أن تكون القيمة الحدية، مثلا، عبارة عن وظيفة قائمة على مجموعة من عوامل. بمعنى آخر، يمكن أن تكون القيمة الحدية قابلة للتكيف مع الظروف. كمثال، عند تحديد ما إذا كانت مجموعة من قراءات العينة تشكل ضوضاء يجب تجاهلها أو بيانات يجب تحليلها، يمكن أن تكون القيمة الحدية إما عدد محدد (مثلا، 10 قراءات عينة) أو وظيفة قائمة على عوامل مختلفة، مثل إجمالي عدد قراءات الموقع الجيني المناظر والمعرفة التاريخية للموقع الجيني. علاوة على ذلك، يمكن أن تشير القيمة الحدية إلى حد أعلى، حد أدنى، أو نطاق بين الحدود. يمكن أن يتضمن الأثر الذي يتم بدؤه، مثلا، إخطار مستخدم نهائي بأن العينة مشكوك بأنها تتضمن منتج مضطرب، حيث تحتوي العينة على خليط المصادر، حيث يتضمن الفحص مناطق مشكلة محددة، حيث أن العينة ذات جودة ضعيفة، إلخ.

في بعض النماذج، يقوم القياس أو الدرجات على بيانات توالي يمكن مقارنتها بالقيمة الحدية. حسب الاستخدام هنا، يمكن أن تتضمن المصطلحات "قياس" أو "درجة" قيم أو نتائج تم تحديدها من بيانات التوالي أو يمكن أن تتضمن وظائف تقوم على أساس القيم أو النتائج التي تم تحديدها من بيانات التوالي. مثل القيمة الحدية، يمكن أن يكون القياس أو الدرجة متكيفين مع الظروف. مثلا، يمكن أن يكون القياس أو الدرجة عبارة عن قيمة معايرة.

كمثال على الدرجة أو القياس، يمكن أن يقوم واحد أو أكثر من النماذج باستخدام درجات العد عند تحليل البيانات. يمكن أن تكون درجة العد قائمة على عدد قراءات العينة. يمكن أن تخضع قراءات العينة إلى واحد أو أكثر من مراحل الترشيح filtering stages بحيث تكون قراءات العينة لها على الأقل خاصية أو جودة واحدة مشتركة. مثلا، يمكن أن تتحاذى كل من قراءات العينة المستخدمة لتحديد درجة العد مع متوالية مرجعية أو يمكن تعيينها كأليلة محتملة. يتضمن عدد قراءات العينة خاصية مشتركة يمكن عدها لتحديد عدد قراءة. يمكن أن تكون درجات العد بناء على عدد القراءة. في بعض النماذج، يمكن أن تكون درجة العد عبارة عن قيمة تساوي عدد القراءة. في النماذج الأخرى، يمكن أن تكون درجة العد قائمة على عدد القراءة والمعلومات الأخرى. مثلا، يمكن أن تكون درجة العد بناء على عدد القراءة لأليلة معينة لموقع جيني والعدد الكلي لقراءات للموقع الجيني. في بعض النماذج، يمكن أن تكون درجة العد قائمة على عدد القراءة وبيانات تم الحصول

عليها سلفا للموقع الجيني. في بعض النماذج، يمكن أن تكون درجات العد عبارة عن درجات معايرة بين قيم محددة سلفا. يمكن أيضا أن تكون درجة العد عبارة عن وظيفة أعداد قراءة من مواقع أخرى للعينة أو وظيفة أعداد قراءة من عينات أخرى والتي تم تشغيلها بشكل متزامن حاليا مع العينة محل الاهتمام. مثلا، يمكن أن تكون درجة العد عبارة عن وظيفة لعدد قراءة أليلة معينة وأعداد القراءة غير المواقع في العينة و/أو أعداد القراءة من عينات أخرى. كمثال، يمكن استخدام أعداد القراءة من مواقع أخرى و/أو أعداد القراءة من عينات أخرى لمعايرة درجة العد لأليلة معينة.

يتم تحديد أعداد القراءة بشكل مثالي من بيانات التوالي. يمكن أن يكون عدد القراءة، مثلا، عبارة عن عدد من قراءات عينة تم تحديدها لتكون هي نفس منطقة محل الاهتمام لتتضمن المنطقة محل الاهتمام. يمكن استخدام عدد القراءة (مثلا، 350 قراءات عينة) لحساب قياس مضطرب والذي يقارن بعد ذلك بقيمة حدية معينة. مثلا، يمكن تحديد القياس المضطرب بواسطة ضرب عدد القراءة في عامل معين بناء على المعرفة التاريخية، معرفة العينة، معرفة الموقع، إلخ. يمكن أن يكون القياس المضطرب عبارة عن قيمة معايرة لعدد القراءة.

سوف يتم فهم الوصف التفصيلي أعلاه والتالي للنماذج المتعددة بشكل أفضل عند القراءة فيما يتعلق بالأشكال الملحقة. بالقدر الذي تقوم به الأشكال بتوضيح المخططات الخاصة بالإطارات الوظيفية للعديد من النماذج، ليس من الضروري أن تشير الإطارات الوظيفية إلى قسمة بين دارات الأجهزة. بالتالي، مثلا، يمكن تنفيذ واحد أو أكثر من الإطارات الوظيفية (مثلا، الوحدات modules والمعالجات processors أو وحدات الذاكرة memories) في قطعة واحدة من الجهاز (مثلا، معالج إشارة signal processor ذات غرض عام أو إطار ذاكرة وصول عشوائي random access memory (RAM) أو قرص صلب hard disk، أو ما شابه) أو العديد من قطع الأجهزة. بالمثل، يمكن أن تكون البرامج منفصلة بذاتها، ويمكن دمجها كروتينات ثانوية subroutines في نظام تشغيل operating system، ويمكن أن تكون عبارة عن وظائف في مجموعة برمجيات تم تنصيبها، وما شابه. يجب فهم أن العديد من النماذج غير مقيدة بالتجهيزات والأدوات المبينة في الأشكال.

يصف الطلب الحالي العديد من الطرق وأنظمة لتنفيذ الطرق. تم شرح على الأقل بعض الطرق في الأشكال كمجموعة من الخطوات. مع ذلك، يجب فهم أن النماذج غير مقيدة بالخطوات المبينة في الأشكال. يمكن استبعاد الخطوات، يمكن تعديل الخطوات، و/أو إضافة خطوات أخرى. مثلا، على الرغم من أن بعض النماذج الموضحة هنا قد يتضمن تحضير وتوالي عينة للحصول على بيانات

توالي، يمكن أن تتضمن النماذج الأخرى تلقي بيانات التوالي مباشرة، بدون تحضير العينة و/أو توالي العينة. علاوة على ذلك، يمكن دمج الخطوات الموضحة هنا، يمكن تنفيذ الخطوات آنيا، يمكن تنفيذ الخطوات بشكل متزامن، يمكن فصل الخطوات إلى خطوات ثانوية متعددة، يمكن تنفيذ الخطوات في ترتيب مختلف، أو إعادة تنفيذ الخطوات (أو سلسلة من الخطوات) بطريقة متتابعة. علاوة على ذلك، على الرغم من ذكر الطرق المختلفة أعلاه، يجب فهم أنه يمكن دمج الطرق المختلفة (أو خطوات الطرق المختلفة) في النماذج الأخرى.

الشكل 1 يشرح طريقة 100 وفقا لأحد النماذج. الطريقة 100 تتضمن تلقي، عند 102، عينة بيولوجية تتضمن أو مشكوك بأنها تتضمن الأحماض النووية، مثل الحمض النووي الريبي منقوص الأكسجين. يمكن أن تكون العينة البيولوجية من مصدر معروف أو غير معروف، مثل حيوان (مثلا، بشري)، نبات، بكتريا bacteria، أو فطر fungus. يمكن أخذ العينة البيولوجية مباشرة من المصدر. مثلا، يمكن أخذ الدم blood أو اللعاب saliva مباشرة من فرد. بشكل بديل، يمكن ألا يتم الحصول على العينة مباشرة من المصدر. مثلا، يمكن الحصول على العينة البيولوجية من مسرح جريمة أو بقايا حفر أو مناطق أخرى يتم فحصها (مثلا، موقع تاريخي). حسب الاستخدام هنا، يتضمن المصطلح "عينة بيولوجية" احتمالية أن تتضمن العينة البيولوجية العديد من العينات البيولوجية من مصادر مختلفة. مثلا، يمكن أن تتضمن العينة البيولوجية التي تم الحصول عليها من مسرح جريمة خليط من الحمض النووي الريبي منقوص الأكسجين من أفراد مختلفين.

يمكن أيضا أن تتضمن الطريقة 100 تحضير، عند 104، العينة للتوالي. يمكن أن تتضمن عملية التحضير 104 إزالة المواد الخارجية و/أو عزل مادة معينة (مثلا، الحمض النووي الريبي منقوص الأكسجين). يمكن تحضير العينة البيولوجية لتتضمن خواص مطلوبة لفحص معين. مثلا، يمكن تحضير العينة البيولوجية للتوالي بواسطة التخليق sequencing-by-synthesis (SBS). في نماذج معينة، يمكن أن يتضمن التحضير تكبير مناطق معينة من الجينوم genome. مثلا، يمكن أن يتضمن التحضير، عند 104، تكبير مواقع جينية محددة سلفا معروفة بأنها تتضمن تكرارات ترادفية قصيرة و/أو تعددات شكلية لنيكلوتيد فردي. يمكن تكبير المواقع الجينية باستخدام متواليات أولية محددة سلفا.

عند 106، يمكن توالي العينة. يمكن تنفيذ التوالي خلال مجموعة من بروتوكولات توالي sequencing protocols معروفة. في نماذج معينة، يتضمن التوالي بواسطة التخليق. في التوالي

بواسطة التخليق، يتم استخدام مجموعة من نيكلوتيدات معلمة بشكل متألق إلى متوالية مجموعة من تكتلات من الحمض النووي الريبي منقوص الأكسجين مكبرة (بشكل محتمل ملايين من التكتلات) الموجودة على سطح ركيزة بصرية optical substrate (مثلا، سطح يحدد جزئيا على الأقل قناة في خلية تدفق flow cell). يمكن أن تحتوي تدفقات الخلية على عينات الحمض النووي للتوالي حيث يتم وضع تدفقات الخلية في حوامل تدفق الخلية المناسبة. يمكن أن تكون العينات الخاصة بالتوالي في صورة جزيئات حمض نووي فردية والتي يتم فصلها عن بعضها البعض لتكون قابل للحر بشكل فردي، المجموعات المكبرة لجزيئات الحمض النووي في صورة تكتلات أو خواص أخرى، أو كريات ملحقة بواحد أو أكثر من جزيئات الحمض النووي.

يمكن تحضير الأحماض النووية بحيث تتضمن بادئ أوليجو نيكلوتيد oligonucleotide primer، والذي يمكن والذي يمكن الإشارة إليه بمتوالية بادئة primer sequence، والتي تجأور متوالية مستهدفة غير معروفة. لبدء دورة توالي التوالي بواسطة التخليق الأولية، يمكن تدفق واحد أو أكثر من النيكلوتيدات المعلمة بشكل مختلف، وبوليمراز polymerase الحمض النووي الريبي منقوص الأكسجين، إلخ.، في/خلال خلية التدفق بواسطة نظام ثانوي لتدفق مائع fluid flow subsystem (غير مبين). يمكن إضافة أي نوع واحد من انيكلوتيد في المرة، أو النيكلوتيدات المستخدمة في إجراء التوالي المعين خصيصا ليتضمن خاصية إنهاء قابلة للعكس ، بالتالي يتم السماح لكل دورة من تفاعل التوالي بأن تتم آنيا في وجود أنواع متعددة من نيكلوتيدات معلمة (مثلا A، C، T، G). يمكن أن تتضمن النيكلوتيدات شقوق علامة قابلة للكشف عنها مثل الفلوروفورات fluorophores. حيث يتم خلط النيكلوتيدات الأربعة معا، يكون البوليمراز قادرا على اختيار القاعدة الصحيحة للدمج ويتم تمديد كل متوالية بواسطة قاعدة فردية. يمكن كسح النيكلوتيدات غير المدمجة بواسطة دفق محلول غسل خلال خلية التدفق. يمكن أن يقوم واحد أو أكثر من أشعات الليزر بإثارة الأحماض النووية والتسبب في تألق. يقوم التألق المنبعث من الأحماض النووية على الفلوروفورات الخاصة بالقاعدة المدمجة، ويمكن أن تقوم الفلوروفورات بإرسال أضواء انبعاث ذات أطوال موجية مختلفة. يمكن إضافة عامل كاشف مانع للإعاقة إلى خلية التدفق لإزالة المجموعات الناهية القابلة للانعكاس من جدائل الحمض النووي الريبي منقوص الأكسجين التي تم مدها والكشف عنها. يمكن يعد ذلك كسح العامل الكاشف المانع للإعاقة deblocking reagent بواسطة دفق محلول الغسل خلال خلية التدفق. بعد ذلك تكون خلية التدفق جاهزة للمزيدلدورة أخرى لبدء التوالي مع إدخال

نيكلوتيد معلم كما هو مذكور أعلاه. يمكن تكرار الخطوة المائعة وخطوة الكشف عدة مرات لإكمال دورة التوالي. تم وصف طرق التوالي المثالية، مثلا، في Bentley et al., Nature 456:53-59 (2008)، المنشور الدولي رقم 018497/04؛ براءة الاختراع الأمريكية رقم 7،057،026؛ المنشور الدولي رقم 06678/91؛ المنشور الدولي رقم 123744/07؛ براءة الاختراع الأمريكية رقم 7،329،492؛ براءة الاختراع الأمريكية رقم 7،211،414؛ براءة الاختراع الأمريكية رقم 7،315،019؛ براءة الاختراع الأمريكية رقم 7،405،281، والمنشور الأمريكي رقم 0108082/2008، وقد تم دمج كل منها بالإشارة إليها كمراجع.

في بعض النماذج، يمكن إلحاق الأحماض النووية بسطح وتكبيرها قبل أو أثناء توالي. مثلا، يمكن تنفيذ التكبير باستخدام التكبير المجسر لتكوين تكتلات الحمض النووي على سطح. تم وصف طرق التكبير المجسر المفيدة، مثلا، في براءة الاختراع الأمريكية رقم 5،641،658؛ منشور براءة الاختراع الأمريكية رقم 0055100/2002؛ براءة الاختراع الأمريكية رقم 7،115،400؛ منشور براءة الاختراع الأمريكية رقم 0096853/2004؛ منشور براءة الاختراع الأمريكية رقم 0002090/2004؛ منشور براءة الاختراع الأمريكية رقم 0128624/2007؛ ومنشور براءة الاختراع الأمريكية رقم 0009420/2008، وقد تم دمج كل منها هنا بالإشارة إليها كمرجع. ثمة طريقة أخرى مفيدة لتكبير الأحماض النووية على السطح هي تكبير الدائرة الدوارة rolling circle amplification (RCA)، مثلا، كما هو موضح في Lizardi et al., Nat. Genet. 19:225-232 (1998) ومنشور براءة الاختراع الأمريكية رقم 0099208/2007 A1، وقد تم دمج كل منها بالإشارة إليها كمراجع.

يقوم بروتوكول التوالي بواسطة التخليق المحدد تحديدا باستغلال نيكلوتيدات معدلة لها كتل 3' قابلة للإزالة، مثلا، كما هو موضح في المنشور الدولي رقم 018497/04، منشور براءة الاختراع الأمريكية رقم 0166705/2007أ1، وبراءة الاختراع الأمريكية رقم 7،057،026، وقد تم دمج كل منها بالإشارة إليها كمراجع. مثلا، يمكن توصيل دورات متكررة من العوامل الكاشفة reagents التوالي بواسطة التخليق إلى خلية تدفق لها أحماض نووية مستهدفة ملحقة بها، مثلا، نتيجة لبروتوكول التكبير المجسر bridge amplification protocol. يمكن تحويل تكتلات الحمض النووي nucleic acid clusters إلى صورة مجدولة واحدة باستخدام محلول استقامة. يمكن أن يحتوي محلول الاستقامة على، مثلا، إندو نيوكلياز endonuclease تقييد قادر على شق جديلة واحدة لكل

تكتل. يمكن استخدام طرق أخرى للشق التي يمكن استخدامها كبديل لإنزيمات التقييد أو إنزيمات تحزيز، شاملة من بين غيرها شق كيمأوي (مثلا، شق رابط داي ول diol linkage مع بيريودات periodate)، مواقع الشق الأليفاتية القاعدية بواسطة الشق مع إندونوكلياز (مثلا 'USER'، المقدم بواسطة NEB، Ipswich، MA، USA، عدد الجزء M5505S)، بواسطة التعرض للحرارة أو لمادة قلوية، شق ريبو نيكلوتيدات ribonucleotides مدمجة في منتجات تكبير مكون من دي peptide linker وكسي ريبو نيكلوتيدات deoxyribonucleotides، الشق الكيمأوي الضوئي photochemical cleavage أو شق رابط ببتيدي. بعد خطوة التقويم linearization step يمكن توصيل بادئ توالي إلى خلية التدفق في ظروف لتهجين التوالي البادئ إلى أحماض نووية مستهدفة سيتم تواليها.

يمكن ملامسة خلية التدفق مع عامل كاشف لامتداد التوالي بواسطة التخليق له نيكلوتيدات معدلة ذات كتل 3' قابلة للإزالة وعلامات متألقة تحت ظروف لمد بادئ مهجن إلى كل حمض نووي مستهدف بواسطة إضافة نيكلوتيد واحد. تتم إضافة فقط نيكلوتيد واحد إلى كل بادئ لأنه بمجرد دمج النيكلوتيد المعدل في سلسلة البولي نيكلوتيد polynucleotide chain النامي المكمل للمنطقة الخاصة بالقالب الذي يتم تواليه لا يكون هناك مجموعة خالية من 3'-OH متاحة لتوجيه عملية مد المتوالية مباشرة وبالتالي لا يمكن إضافة البوليمراز أكثر إلى النيكلوتيدات. يمكن إزالة العامل الكاشف لامتداد التوالي بواسطة التخليق واستبداله بعامل كاشف للمسح scan reagent حتوي على مكونات تقوم بحماية العينة تحت إثارة بالإشعاع. تم وصف المكونات المثالية لالعامل الكاشف للمسح في المنشور الأمريكي رقم 0280773/2008 أ1 والرقم المسلسلة الأمريكي 13018255، وقد تم دمج كل منها بالإشارة إليها كمراجع. يمكن بعد ذلك الكشف عن الأحماض النووية الممتدة بطريقة متألقة في وجود العامل الكاشف للمسح. بمجرد الكشف عنها، يمكن إزالة العائق 3' باستخدام عامل كاشف نازع للعائق مناسب للمجموعة المعيقة المستخدمة. تم وصف عوامل كاشفة نازعة للعائق مفيدة للمجموعات المعيقة المناظرة في براءة الاختراع الدولية رقم 04018497، وبراءات الاختراع الأمريكية رقم 0166705/2007أ1 و رقم 7057026، وقد تم دمج كل منها بالإشارة إليها كمراجع. يمكن كسح عامل كاشف نازع للعائق لترك الأحماض النووية المستهدفة مهجنة إلى مواد بادئة ممتدة لها مجموعات 3' OH تكون الآن مستحقة لإضافة نيكلوتيد آخر. بالتالي يمكن تكرار دورات إضافة العامل الكاشف للامتداد، العامل الكاشف للمسح، وعامل كاشف نازع للعائق، مع عمليات غسل اختيارية بين واحد أو أكثر من الخطوات، لحين الحصول على

المتوالية المطلوبة. يمكن تنفيذ الدورات المذكورة أعلاه باستخدام خطوة توصيل عامل كاشف للامتداد واحد لكل دورة حتى عندما تتضمن كل من النيكلوتيدات المعدلة علامة مختلفة ملحقة بها، والمعروفة بأنها تناظر قاعدة معينة. تقوم العلامات المختلفة بتسهيل التمييز بين النيكلوتيدات المضافة أثناء كل خطوة دمج. بشكل بديل، يمكن أن تتضمن كل دورة خطوات منفصلة خاصة بتوصيل عامل كاشف للامتداد متبوع بواسطة خطوات منفصلة من توصيل العامل الكاشف للمسح والكشف، في هذه الحالة يمكن أن يكون لاثنين أو أكثر من النيكلوتيدات نفس العلامة ويمكن أن يتم تمييزها بناء على الترتيب المعروف للتوصيل.

بالاستمرار في المثال على تكتلات الحمض النووي في خلية تدفق، يمكن معالجة الأحماض النووية أيضا للحصول على قراءة ثانية من الطرف المقابل بطريقة معروفة بـ "توالي مزدوج الطرف." يقوم التوالي مزدوج الطرف للمستخدم بإتاحة المتوالية بكلا طرفي الشظية المستهدفة. يمكن أن يقوم التوالي مزدوج الطرف بتسهيل الكشف عن عمليات إعادة ترتيب جينومية وقطاعات تكرارية. بالإضافة إلى عمليات دمج جين ونواتج تحويل جديدة. تم وصف طريقة لتوالي مزدوج الطرف في منشور الدولي رقم 07010252، طلب نشر براءة الاختراع البريطانية برقم مسلسل 003798/2007 ومنشور طلب براءة الاختراع الأمريكية رقم 0088327/2009، وقد تم دمج كل منها هنا بالإشارة إليها كمرجع. في أحد الأمثلة، يمكن تنفيذ سلسلة من الخطوات كما يلي؛ (أ) توليد تكتلات الأحماض النووية؛ (ب) تقويم الأحماض النووية؛ (ج) تهجين بادئ توالي أول وتنفيذ دورات متكررة للامتداد والمسح ونزع الإعاقة، كما هو مذكور أعلاه؛ (د) "تحويل الأحماض النووية المستهدفة على سطح خلية التدفق بواسطة تخليق نسخة مكملة؛ (هـ) تقويم الجديلة المعاد تخليقها؛ و(و) تهجين بادئ توالي ثاني وتنفيذ دورات متكررة من مد، مسح ونزع الإعاقة، كما هو مذكور أعلاه. يمكن تنفيذ خطوة التحليل بواسطة عوامل كاشفة reagents موصلة كما هو مذكور أعلاه لدورة واحدة من التكبير المجسر.

على الرغم من تقديم مثال على عملية التوالي، عند 106، أعلاه نسبة إلى بروتوكول التوالي بواسطة التخليق معين، سوف يتم فهم أن هناك بروتوكولات أخرى لتوالي أي مجموعة من عمليات تحليل جزيئية أخرى يمكن تنفيذها حسب الحاجة. مثلا، يمكن استخدام عملية استحلاب تفاعل سلسلة بوليمرات على كريات، مثلا كما هو موضح في Dressman et al., Proc. Natl. Acad. Sci. USA 100:8817-8822 (2003)، براءة الاختراع الدولية رقم 01045/05، أو منشور براءة الاختراع

الأمريكية أرقام 0130173/2005 أو 0064460/2005، وقد تم دمج كل منها هنا بالإشارة إليها كمرجع. هناك تقنيات توالي أخرى يمكن استخدامها مع الطرق والأنظمة المذكورة أعلاه وهي التوالي الحراري، توالي المسام بحجم النانو، وتوالي بواسطة الارتباط. تم وصف التقنيات المثالية للتوالي الحراري والعينات المفيدة تحديدا في منشور براءة الاختراع الأمريكية أرقام 6210891؛ و 6258568؛ و 6274320 و Ronaghi, جينوم(2001) 11:3-11 Research ، وقد تم دمج كل منها بالإشارة إليها كمراجع. تتضمن التقنيات المثالية على تقنيات المسام بحجم النانو والعينات تلك المفيدة التي تم وصفها في Deamer et al., Acc. Chem. Res. 35:817-825 (2002); Li et al., Nat. Mater. 2:611-615 (2003); Soni et al., Clin Chem. 53:1996-2001 (2007) Healy et al., Nanomed. 2:459-481 (2007) and Cockroft et al., J. am. Chem. Soc. 130:818-820؛ وبراءة الاختراع الأمريكية رقم 7001792، وقد تم دمج كل منها بالإشارة إليها كمراجع. تحديدا، تقوم هذه طرق باستخدام الخطوات المكررة لتوصيل العامل الكاشف. يمكن تهيئة أداة أو طريقة المذكورة أعلاه بخزانات، صمامات، خطوط مائعة ومكونات أخرى على طول أنظمة التحكم لتلك المكونات لإدخال عوامل كاشف والكشف عن الإشارات البصرية وفقا لبروتوكول مطلوب مثل تلك المذكورة في المراجع المذكورة أعلاه. يمكن استخدام أي من مجموعة العينات في هذه الأنظمة مثل الركائز التي لها كريات تم توليدها بواسطة استحلاب تفاعل سلسلة بوليمرات، الركائز لها موجهات موجية في الوضع الصفري، الركائز لها وسائل كشف سي إم أوه إس CMOS مدمجة، الركائز لها مسام بيولوجية بحجم النانو في طبقات دهنية مزدوجة lipid bilayers، الركائز في الحالة الصلبة لها مسام تخليقية بحجم النانو، وغيرها من المعروف في المجال. تم وصف عينات أخرى في سياق تقنيات التوالي المختلفة في المراجع المذكورة أعلاه وأيضا في براءات الاختراع الأمريكية رقم 0042648/2005؛ و 0079510/2005؛ و 0130173/2005؛ و براءة الاختراع الدولية رقم 010145/05، وقد تم دمج كل منها بالإشارة إليها كمراجع.

أنظمة يمكن أن تكون قادرة على تنفيذ واحد أو أكثر من بروتوكولات التوالي بواسطة التخليق الموضحة أعلاه شاملة الأنظمة المطورة بواسطة Illumina, Inc.، مثل أنظمة MiSeq، HiSeq 2500، HiSeq X Ten، NeoPrep، HiScan، وiScan. تم وصف الأنظمة القادرة على تنفيذ واحد أو أكثر من بروتوكولات التوالي بواسطة التخليق الموضحة أعلاه في طلب براءة الاختراع الأمريكية رقم 273666/13 و905633/13، وقد تم دمج كل منها هنا بالإشارة إليها كمرجع.

عند 108، يمكن تلقي بيانات التوالي للتحليل التالي عند 110. يمكن أن تتضمن بيانات التوالي، مثلا، عدد من قراءات عينة. يمكن أن تتضمن كل عينة قراءة متوالية لنيكلوتيدات، والذي يمكن الإشارة إليه بعينة متوالية أو متوالية مستهدفة. يمكن أن تتضمن عينة المتوالية، مثلا، متواليات أولية، متواليات مجنحة، ومتوالية مستهدفة. يمكن أن يتضمن عدد النيكلوتيدات في عينة المتوالية 30، 40، 50، 60، 70، 80، 90، 100 أو أكثر. في بعض النماذج، واحد أو أكثر من قراءات العينة (أو متواليات عينة) تتضمن على الأقل 150 نيكلوتيدات، 200 نيكلوتيدات، 300 نيكلوتيدات، 400 نيكلوتيدات، 500 نيكلوتيدات، أو أكثر. في بعض النماذج، يمكن أن تتضمن قراءات العينة أكثر من 1000 نيكلوتيدات، 2000 نيكلوتيدات، أو أكثر. يمكن أن تتضمن قراءات العينة (أو عينة المتواليات) متواليات أولية عند واحد من أو كلا الطرفين. في نماذج معينة، يمكن أن تكون كل عينة قراءة مرتبطة بقراءة أخرى في الاتجاه المقابل على طول القالب. مثلا، يمكن أن يتضمن التوالي، عند 106، توالي مزدوج الطرف حيث يتم تنفيذ قراءة أولى (قراءة 1) متبوعة قراءة ثانية (قراءة 2) في الاتجاه المقابل. يمكن أن تتضمن كل من القراءات الأولى والثانية متوالية مستهدفة كاملة أو متوالية مستهدفة كاملة تقريبا. مع ذلك، في النماذج الأخرى، يمكن استخدام التوالي مزدوج الطرف "غير المتناظر" حيث تتضمن القراءة الثانية فقط جزء مما يمكن الحصول عليه. مثلا، يمكن أن تتضمن القراءة الثانية عدد محدد من النيكلوتيدات لتأكيد هوية المتوالية البادئة الموضوعة بالقرب من بداية المتوالية للقراءة الثانية. مثلا، يمكن أن تتضمن القراءة الأولى 300-500 نيكلوتيدات، ولكن يمكن أن تتضمن القراءة الثانية فقط 20-50 نيكلوتيدات.

تم وصف التحليل، عند 110، بالمزيد من التفصيل أدناه. يمكن أن يتضمن التحليل، عند 110، بروتوكول واحد أو توليفة من بروتوكولات تقوم بتحليل قراءات العينة بطريقة معينة للحصول على المعلومات المطلوبة. يمكن أن تتضمن الأمثلة غير المقيدة على التحليل، عند 110، تحليل قراءات العينة لتعيين قراءات العينة إلى (أو تعيين قراءات العينة لـ) مواقع جينية معينة؛ تحليل قراءات العينة لتحديد طول و/أو متوالية لقراءات العينة؛ تحليل قراءات العينة لتصنيف المناطق محل الاهتمام المرتبطة بألائل مستهدفة من موقع معين؛ تحليل قراءات العينة (أو المناطق محل الاهتمام) الخاصة بألائل مستهدفة مختلفة لتحديد إذا كانت المنطقة محل الاهتمام لأليلة مستهدفة واحدة عبارة عن منتج مضطرب مشكوك فيه لمنطقة محل الاهتمام لأليلة مستهدفة أخرى؛ تحديد نمط جيني لموقع جيني؛ و/أو مراقبة صحة أو التحكم في جودة الفحص.

يمكن أن تتضمن الطريقة 100 أيضا توليد أو توفير، عند 112، تقرير عينة. يمكن أن يتضمن تقرير العينة، مثلا، معلومات تتعلق بمجموعة من مواقع جينية نسبة إلى العينة. مثلا، لكل موقع جيني من مجموعة محددة سلفا من مواقع جينية، يمكن أن يتضمن تقرير العينة واحد على الأقل من عملية توفير تسمية نمط جيني؛ الإشارة إلى أن تسمية نمط جيني لا يمكن تنفيذها؛ توفير درجة ثقة بناء على التأكد من تسمية النمط الجيني؛ أو الإشارة إلى المشاكل المحتملة مع فحص يتعلق بواحد أو أكثر من مواقع جينية. يمكن أيضا أن يشير تقرير العينة إلى نوع فرد قام بتقديم عينة و/أو الإشارة إلى أن العينة تتضمن العديد من المصادر. حسب الاستخدام هنا، يمكن أن يتضمن "تقرير العينة" بيانات رقمية (مثلا، ملف بيانات) لموقع جيني أو مجموعة محددة سلفا من موقع جيني و/أو تقرير مطبوع للموقع الجيني أو مجموعة المواقع الجينية. بالتالي، يمكن أن تتضمن عملية التوليد أو التوفير، عند 112، تكوين ملف بيانات و/أو طباعة تقرير العينة، أو عرض تقرير العينة.

الشكل 2 عبارة عن مخطط تدفق يشرح الطريقة 150 لتحليل بيانات توالي لقراءات عينة لها اختلافات متوالية. تم شرح الشكل 2 أدناه بالإضافة إلى الشكل 3، والذي يشرح أيضا الخطوات المختلفة وفقا للشكل 1. تتضمن الطريقة 150 تلقي، عند 152، بيانات توالي من واحد أو أكثر من مصادر. يمكن أن تتضمن بيانات التوالي مجموعة من قراءات عينة تتضمن متواليات عينة مناظرة للنيكلوتيدات. الشكل 3 يبين مثال على عينة قراءة 180. تمثل المصطلحات "متوالية تحديد" و"اختلاف متوالية" أجزاء من عينة المتوالية. على الرغم من توضيح عينة قراءة 180 واحدة، يجب فهم أنه يمكن أن تتضمن بيانات التوالي، مثلا، المئات أو الآلاف أو مئات الآلاف أو الملايين من قراءات العينة. يمكن أن تتضمن قراءات العينة المختلفة أعداد مختلفة من النيكلوتيدات. مثلا، يمكن أن تتفاوت عينة القراءة بين 10 نيكلوتيدات إلى حوالي 500 نيكلوتيدات أو أكثر. مع ذلك، يمكن أن تتضمن قراءات العينة نيكلوتيدات أكثر في النماذج الأخرى. يمكن أن تقوم قراءات العينة بلف الجينوم الكامل للمصدر. في نماذج معينة، تتعلق قراءات العينة بمواقع جينية محددة سلفا، مثل تلك المواقع الجينية التي لها تكرارات ترادفية قصيرة مشكوك فيها أو تعددات شكلية لنيكلوتيد فردي مشكوك فيها. يمكن اختيار قراءات العينة بناء على متواليات أولية معروفة مرتبطة بالمواقع الجينية محل الاهتمام. مثلا، يمكن أن تتضمن قراءات العينة نواتج تكبير تفاعل سلسلة بوليمرات التي تم الحصول عليها باستخدام المتواليات الأولية المرتبطة بالمواقع الجينية محل الاهتمام.

عند 154، يمكن تعيين كل من قراءات العينة إلى مواقع جينية مناظرة. يمكن تعيين قراءات العينة إلى مواقع جينية مناظرة بناء على متوالية النيكلوتيدات لعينة القراءة أو، بمعنى آخر، ترتيب النيكلوتيدات في عينة القراءة (مثلا، A، C، G، T). بناء على هذا التحليل، يمكن تعيين عينة القراءة على أنها تتضمن أليلة ممكنة لموقع جيني معين. يمكن جمع عينة القراءة (أو تراكمها أو تثبيتها) مع قراءات عينة أخرى تم تصميمها على أنها تتضمن ألائل ممكنة للموقع الجيني. تم تمثيل المواقع الجينية المختلفة كصناديق 182 في الشكل 3. يمكن أن تكون المواقع الجينية عبارة عن مجموعة محددة سلفا من المواقع الجينية المستخدمة لفحص معين. مثلا، يمتلك مكتب التحقيقات الفيدرالي (13) من مواقع تكرار ترادفي قصير المعرفة التي يمكن استخدامها لتوليد هيئات عينة لمشتبهين محتملين في الجرائم. باستخدام معيار FBI كمثال، يمكن أن تقوم الطريقة 150 بتعيين كل من قراءات العينة، إن أمكن، إلى واحد من الصناديق الثلاثة عشر.

يمكن أن تخضع قراءات العينة للصناديق المختلفة إلى عملية تحليل مختلفة. مثلا، يمكن تعيين قراءات العينة إلى مواقع جينية تتضمن تكرارات ترادفية قصيرة. يمكن الإشارة إلى هذه المواقع بمواقع تكرار ترادفي قصير. مع ذلك، يمكن تعيين قراءات العينة إلى مواقع جينية تتضمن تعددات شكلية لنيكلوتيد فردي. بحيث تتم الإشارة إلى هذه المواقع بمواقع تعدد شكلي لنيكلوتيد فردي. بالنسبة لعينة قراءة مثالية، سوف يتم تعيين عينة القراءة إلى موقع جيني واحد فقط (أو صندوق واحد). في هذه الظروف، سوف تخضع قراءات العينة إلى تحليل مهيأ لنوع المواقع الجينية. بشكل أكثر تحديدا، سوف تخضع قراءات العينة المعينة لموقع تكرار ترادفي قصير إلى تحليل تكرار ترادفي قصير، حيث تخضع قراءات العينة المعينة إلى موقع تعدد شكلي لنيكلوتيد فردي إلى تحليل تعدد شكلي لنيكلوتيد فردي. في بعض الظروف، مع ذلك، قد يكون من الممكن لعينة قراءة أن تكون معينة إلى أكثر من موقع جيني واحد و، بالتالي، يمكن أن تخضع عينة القراءة إلى أكثر من نوع واحد من التحليل.

يمكن أيضا الإشارة إلى عملية التعيين، عند 154، بتسمية موقع حيث يتم تحديد عينة القراءة على أنها مرتبطة بشكل ممكن بموقع جيني معين. يمكن تحليل قراءات العينة لوضع واحد أو أكثر من متواليات التحديد (مثلا، متواليات أولية primer sequences) لنيكلوتيدات تقوم بتمييز عينة القراءة من قراءات عينة أخرى. بشكل أكثر تحديدا، يمكن أن تقوم متوالية التحديد بتحديد عينة القراءة من قراءات عينة أخرى على أنه مرتبط بموقع جيني معين. يمكن أن تتضمن متوالية التحديد أو

وضعها بالقرب (مثلا، في 10-30 نيكلوتيدات) من طرف آخر من عينة القراءة. في نماذج معينة، متواليات التحديد لعينة القراءة قائمة على أساس متواليات أولية التي تم استخدامها لتكبير متواليات بشكل انتقائي من المصدر أو مصادر. مع ذلك، في النماذج الأخرى، يمكن ألا يتم وضع متوالية التحديد بالقرب من أحد أطراف عينة القراءة sample read.

في بعض النماذج، تتم مقارنة متواليات التحديد بمجموعة من متواليات محددة سلفا لتحديد ما إذا كان أي من متواليات التحديد متطابقة مع أو قريبة من واحد من المتواليات المحددة سلفا. مثلا، يمكن مقارنة كل متوالية تحديد بقائمة من متواليات محددة سلفا في قاعدة بيانات 184 (مثلا، انظر الجدول). يمكن أن تتعلق المتواليات المحددة سلفا إلى مواقع جينية معينة. تتم الإشارة إلى المتواليات المحددة سلفا من قاعدة البيانات لاختيار المتواليات. تمثل كل متوالية اختيار متوالية لنيكلوتيدات. إذا كانت متوالية التحديد متوافقة بفاعلية مع أي من متواليات الاختيار، يمكن تعيين عينة القراءة لها متوالية التحديد إلى الموقع الجيني الذي يتعلق بمتوالية الاختيار. قد يكون من الممكن أن تتوافق عينة قراءة بفاعلية أكثر من متوالية اختيار واحدة. في هذه الحالات، يمكن تعيين عينة إلى كل من المواقع الجينية إلى متواليات اختيار وتخضع أكثر للتحليل لتحديد ما إذا كان واحد من المواقع الجينية لعينة القراءة يمكن تسميته.

يمكن أن يكون هناك عدد محدد سلفا من متواليات اختيار مستخدمة أثناء التحليل. مثلا، يمكن أن تتضمن الهيئة الجينية المولدة بواسطة النماذج المذكورة أعلاه تحليل من حوالي 5 إلى حوالي 300 مواقع جينية. في نماذج معينة، يمكن أن يكون عدد المواقع الجينية من حوالي 5 إلى حوالي 100 موقع جيني أو، بشكل أكثر تحديدا، من حوالي 10 إلى حوالي 30 مواقع جينية. مع ذلك، يمكن استخدام أعداد أخرى من المواقع الجينية. يمكن أن يتضمن كل موقع جيني عدد محدود من متواليات اختيار مرتبطة بالموقع الجيني. مع عدد محدود من المواقع الجينية وعدد محدود من متواليات الاختيار المرتبطة بكل موقع جيني، يمكن أن تكون قراءات العينة مسماة للمواقع الجينية بدون استخدام مفرط للمصادر الحسابية. في بعض النماذج، تقوم متواليات الاختيار على أساس متواليات أولية تم استخدام لتكبير متواليات محددة سلفا من الحمض النووي الريبي منقوص الأكسجين بشكل انتقائي.

على الرغم من أن كل متوالية اختيار يمكن أن تكون قائمة على متوالية تحديد للموقع الجيني (مثلا، متوالية بادئة)، يمكن ألا تتضمن متوالية الاختيار كل نيكلوتيد من متوالية التحديد. كمثال،

يمكن أن تتضمن متوالية الاختيار سلسلة من نيكلوتيدات n من متوالية التحديد لواحد من قراءات العينة. في نماذج معينة، يمكن أن تتضمن متواليات الاختيار نيكلوتيدات n الأولى لمتوالية التحديد. يمكن أن يكون عدد n كافي لتمييز الألائل الخاصة بموقع جيني واحد من ألائل من موقع مستهدف آخر. في بعض النماذج، يكون عدد n ما بين 10 و30.

يمكن أن تتضمن عملية التعيين، عند 154، تحليل سلسلة من n نيكلوتيدات لمتوالية التحديد لتحديد إذا كانت السلسلة من نيكلوتيدات n من متوالية التحديد تتوافق بفاعلية مع واحد أو أكثر من متواليات الاختيار. في نماذج معينة، يمكن أن تتضمن عملية التعيين، عند 154، تحليل نيكلوتيدات n الأولى لعينة المتوالية لتحديد إذا كانت نيكلوتيدات n الأولى لعينة المتوالية بفاعلية تتوافق مع واحد أو أكثر من متواليات الاختيار. يمكن أن يتضمن العدد n مجموعة من القيم، والتي يمكن برمجتها في البروتوكول أو إدخالها بواسطة مستخدم. مثلا، يمكن تحديد العدد n كعدد نيكلوتيدات أقصر متوالية اختيار في قاعدة البيانات. يمكن أن يكون العدد n عبارة عن عدد محدد سلفا. يمكن أن يكون عدد محدد سلفا، مثلا، 10، 11، 12، 13، 14، 15، 16، 17، 18، 19، 20، 21، 22، 23، 24، 25، 26، 27، 28، 29، أو 30 نيكلوتيدات. مع ذلك، يمكن استخدام نيكلوتيدات أقل أو أكثر في النماذج الأخرى. يمكن أيضا اختيار العدد n بواسطة فرد، مثل مستخدم النظام. يمكن أن يكون العدد n قائم على واحد أو أكثر من الظروف. مثلا، يمكن تحديد العدد n كعدد النيكلوتيدات لأقصر متوالية بادئة في قاعدة البيانات أو عدد معين، والذي لم يكن هو العدد الأصغر. في بعض النماذج، يمكن استخدام القيمة الأقل لـ n، مثل 15، بحيث يمكن تعيين أي متوالية بادئة من 15 نيكلوتيدات كاستثناء.

في بعض الحالات، يمكن ألا تتوافق سلسلة النيكلوتيدات n من متوالية محددة بدقة مع نيكلوتيدات متوالية الاختيار. مع ذلك، يمكن أن تتوافق متوالية التحديد بفاعلية مع متوالية الاختيار إذا كانت متوالية التحديد شبه متطابقة مع متوالية الاختيار. مثلا، يمكن تسمية عينة القراءة لموقع جيني إذا كانت سلسلة النيكلوتيدات n (مثلا، نيكلوتيدات n الأولى) من متوالية التحديد تتوافق مع متوالية اختيار بما لا يزيد عن العدد المعين من حالات عدم التوافق (مثلا، 3) و/أو عدد معين من الإزاحات (مثلا، 2). يمكن أن تطبيق القوانين بحيث تكون كل حالة عدم توافق أو إزاحة يمكن أن تكون محسوب كفرق بين عينة القراءة والمتوالية البادئة. إذا كان عدد الفروق أقل من عدد معين، بعد ذلك يمكن تسمية عينة القراءة للموقع الجيني المناظر (أي، المعين إلى الموقع الجيني

المناظر). في بعض النماذج، يمكن أن تكون درجة التوافق محددة بناء على عدد الفروق بين متوالية التحديد لعينة القراءة ومتوالية الاختيار مرتبطة بموقع جيني. إذا كانت درجة التوافق تتخطى قيمة حدية متوافقة معينة، فإن الموقع الجيني الذي يناظر متوالية الاختيار يمكن تعيينه كموقع محتمل لعينة القراءة. في بعض النماذج، يمكن تنفيذ التحليل التالي لتحديد ما إذا كان عينة القراءة مسماة للموقع الجيني.

يمكن أن يكون العدد المعين من الفروق بين متوالية التحديد ومتوالية الاختيار، مثلا، عدد أقل من 20٪ من العدد الكلي للنيكلوتيدات في متوالية الاختيار المناظرة، أو، بشكل أكثر تحديدا، عدد أقل من 15٪ من العدد الكلي لنيكلوتيدات في متوالية الاختيار المناظرة. يمكن أن يكون العدد المعين من الفروق عبارة عن قيمة محددة سلفا، مثل 6، 5، 4، 3، أو 2. بالتالي، العبارة "تتوافق بفاعلية" تتضمن عينة متوالية لها سلسلة من نيكلوتيدات n التي تتوافق بالضبط مع متوالية اختيار أو تتوافق تقريبا مع متوالية اختيار مع عدد محدد من الفروق بين متوالية الاختيار وسلسلة النيكلوتيدات n.

إذا كانت عينة القراءة تتوافق بفاعلية مع واحد من متواليات الاختيار في قاعدة البيانات (أي، بالضبط تتوافق مع أو تتوافق تقريبا مع كما هو موضح أعلاه)، يتم بعد ذلك تعيين عينة القراءة أو تعيينها إلى الموقع الجيني التي تتعلق بمتوالية الاختيار. يمكن الإشارة إلى هذا بتسمية موقع أو تسمية مؤقتة للموقع، حيث تتم تسمية عينة القراءة للموقع الجيني التي تتعلق بمتوالية الاختيار. مع ذلك، كما هو مذكور أعلاه، يمكن تسمية عينة قراءة لأكثر من موقع جيني واحد. في هذه النماذج، يمكن تنفيذ المزيد من التحليل لتسمية أو تعيين عينة القراءة لواحد فقط من المواقع الجينية المحتملة.

في بعض النماذج، عينة القراءة التي تتم مقارنتها بقاعدة البيانات هي القراءة الأولى من توالي مزدوج الطرف. بالنسبة لنماذج محددة أكثر، يمكن تحليل القراءة الثانية التي تتعلق بعينة القراءة لتأكيد أن متوالية التحديد في القراءة الثانية تتوافق بفاعلية مع متوالية الاختيار من قاعدة البيانات. يمكن أن تكون متواليات الاختيار في قاعدة البيانات للقراءات الثانية مختلفة عن متواليات الاختيار المستخدمة للقراءات الأولية. في بعض النماذج، تتم تسمية عينة القراءة للموقع الجيني فقط بعد تأكيد أن القراءة الثانية تتوافق أيضا بفاعلية مع متوالية اختيار في قاعدة البيانات. يمكن تنفيذ عملية تحديد ما إذا كان القراءة الثانية تتوافق بفاعلية مع متوالية اختيار بطريقة مشابهة كما هو موضح أعلاه. بواسطة تأكيد أن القراءة الثانية تتوافق بفاعلية مع متوالية اختيار، يمكن ترشيح

قراءات العينة المنفصلة عن الهدف (مثلا، نواتج التكبير المنفصلة عن الهدف off-target amplicons) من تحليل آخر.

يمكن الإشارة إلى قراءات العينة التي تتضمن تلك المسماة لموقع جيني معين بـ "قراءات معينة" للموقع الجيني المحدد. عند هذه المرحلة، على الرغم من أن القراءات المعينة قد تم تحديدها على أنها تتعلق بشكل محتمل بموقع جيني معين، من الممكن ألا تكون القراءة المعينة مناسبة للمزيد من التحليل. بشكل أكثر تحديدا، يمكن إزالة قراءة معينة (أو قراءات) بعد ذلك من تحليل آخر بناء على عوامل أخرى.

بعد تعيين، عند 154، القراءات المعينة إلى مواقع جينية مناظرة، يمكن بعد ذلك تحليل قراءات العينة بشكل أكبر. يمكن أن يكون التحليل التالي الذي يتم تنفيذه مع القراءات المعينة بناء على نوع الموقع الجيني المسمى للقراءة المعينة. مثلا، إذا كان موقع جيني معروف بأنه يتضمن تعددات شكلية لنيكلوتيد فردي، فإن القراءات المعينة التي تمت تسميتها للموقع الجيني يمكن أن تخضع للتحليل، عند 156، لتحديد التعددات الشكلية لنيكلوتيد فردي للقراءات المعينة. إذا كان الموقع الجيني معروف بأنه يتضمن عناصر الحمض النووي الريبي منقوص الأكسجين تكرارية متعددة الشكل، فإنه يمكن تحليل القراءات المعينة، عند 158، لتحديد أو وصف عناصر الحمض النووي الريبي منقوص الأكسجين التكرارية متعددة الشكل في قراءات العينة. في بعض النماذج، إذا كانت قراءة معينة تتوافق بفاعلية مع موقع تكرار ترادفي قصير وموقع تعدد شكلي لنيكلوتيد فردي، يمكن تعيين إنذار أو علم إلى عينة القراءة. يمكن تعيين عينة القراءة بأنه كل من موقع تكرار ترادفي قصير وموقع تعدد شكلي لنيكلوتيد فردي التي تخضع، مثلا، لتحليل عند 156 وتحليل عند 158.

في بعض النماذج، يمكن تنفيذ تحليل تكرار ترادفي قصير باستخدام البروتوكول الموضح أدناه نسبة إلى الأشكال 4-7. يمكن أن يتضمن التحليل عند 158 تحليل قراءات العينة لتحديد المناطق محل الاهتمام، والتي يمكن أن تتضمن متواليات التحديد لمناطق محل الاهتمام و/أو أطوال المناطق محل الاهتمام. يمكن أن تكون المناطق محل الاهتمام عبارة عن متواليات لقراءات العينة (مثلا، متواليات ثانوية من متواليات عينة). يمكن أن يتضمن المناطق محل الاهتمام قطاعات تكرارية repetitive segments. يمكن أن تكون المناطق محل الاهتمام عبارة عن متواليات من نيكلوتيدات التي تتضمن فقط واحد أو أكثر من سلسلة من بواعث تكرار (أي، القطاع التكراري repetitive segment) أو تتضمن واحد أو أكثر من سلسلة من بواعث تكرار بالإضافة إلى عدد معين من

نيكلوتيدات ممتدة من واحد من أو كلا طرفي القطاع التكراري. بشكل أكثر تحديدا، يمكن أن يتضمن كل من المناطق محل الاهتمام واحد أو أكثر من سلسلة من بواعث تكرار حيث يتضمن كل باعث تكرار مجموعة متطابقة من نيكلوتيدات (مثلا، اثنين، ثلاثة، أربعة، خمسة، ست نيكلوتيدات، أو أكثر) من نيكلوتيدات. تتضمن بواعث التكرار المستخدمة بشكل مشترك تترا نيكلوتيدات، ولكن يمكن استخدام بواعث أخرى، مثل نيكلوتيدات أحادية-، ثنائية-، ثلاثية، خماسية، أو سداسية. في نماذج معينة، بواعث التكرار يتضمن تترا نيكلوتيد tetranucleotide.

يمكن أن يتضمن التحليل، عند 158، تحليل القراءات المعينة لكل موقع معين لتحديد المناطق محل الاهتمام مناظرة في القراءات المعينة. بشكل أكثر تحديدا، يمكن تحديد طول و/أو متوالية المناطق محل الاهتمام. يمكن أن يتضمن التحليل، عند 158، محاذاة القراءات المعينة وفقا لبروتوكول محاذاة alignment protocol لمتواليات التحديد و/أو أطوال القراءات المعينة. يمكن أن يتضمن بروتوكول المحاذاة الطريقة الموضحة في طلب النشر الدولي لبراءة الاختراع الأمريكية رقم 030867/2013 (منشور لبراءة الاختراع الدولية رقم 142831/2014)، المودع في 15 مارس 2013، والذي تم دمجه هنا بالإشارة إليه كمرجع.

مع ذلك، يمكن استخدام بروتوكولات محاذاة أخرى. مثلا، هناك بروتوكول محاذاة معروف يقوم بمحاذاة عينة قراءة مع متوالية مرجعية. ثمة منهج آخر حالي يقوم بمحاذاة عينة القراءة مع بنية متسلسلة مرجعية. في هذا المثال، يتم تكوين "جينوم مرجعي reference genome" بواسطة بناء بنية متسلسلة من كل ألائل تكرار ترادفي قصير المعروفة ومحاذاة القراءات مع الجينوم المرجعي، كما هو منفذ بشكل مثالي مع بيانات متوالية جينوم NGS كامل أو توالي مستهدف لمناطق الحمض النووي الريبي منقوص الأكسجين غير متكررة. ثمة طريقة أخرى التي يمكن استخدامها مع النماذج المذكورة أعلاه معروفة كـ lobSTR. تقوم طريقة lobSTR بعد ذلك بتسمية كل تكرارات ترادفية قصيرة الموجودة من بيانات توالي لعينة جديدة فردية، بدون معرفة سابقة لـ تكرارات ترادفية قصيرة (انظر Gymrek et al. 2012 Genome Research 22:1154-62)، والذي تم دمجه هنا بالإشارة إليه كمرجع.

تقوم طريقة المحاذاة المذكورة أعلاه في الطلب الدولي لبراءة الاختراع الأمريكية رقم 030867/2013 (منشور براءة الاختراع الدولية رقم 142831/2014) الآن للمواقع الجينية التي تتضمن المناطق محل الاهتمام. لسهولة القراءة، يمكن الإشارة إلى مواقع جينية هذه بمواقع تكرار

ترادفي قصير. في بعض النماذج، يتم استخدام الجنيحات المحفوظة لمواقع تكرار ترادفي قصير لتحديد متوالية القطاع التكراري بفاعلية. بعد التعيين، عند 154، يمكن أن تقوم قراءات العينة إلى مواقع تكرار ترادفي قصير المناظرة، يمكن أن تقوم النماذج بمحاذاة قطاعات متواليات مجنحة على كل من جوانب القطاع التكراري المناظر لتحديد طول ومتوالية القطاع التكراري. يمكن بذر عمليات المحاذاة باستخدام استراتيجية وحدة k. يمكن أن تكون مناطق البذر seed regions، مثلا، في منطقة مختارة عالية التعقيد من المتوالية المجنحة flanking sequence، بالقرب من القطاع التكراري، ولكن تجنب المتوالية منخفضة التعقيد low-complexity sequence مع التجانس مع القطاع التكراري. هذه المناهج يمكن أن تتجنب عدم توافق متواليات مجنحة منخفضة التعقيد قريبة من القطاع التكراري.

يمكن أن تقوم النماذج باستخدام متواليات معروفة في جنيحات التكرار الترادفي القصير نفسها، والتي تم تحديدها سلفا بناء على الاختلافات المعروفة الموجودة بين المجموعة البشرية. بشكل مفيد، تنفيذ محاذاة امتداد قصير للمنطقة المجنحة أسرع حسابيا من طرق أخرى. مثلا، محاذاة برمجة ديناميكية (من نوع Smith-Waterman) من القراءة الكاملة يمكن أن تكون حساسة لسي بي يو CPU، مستهلكة للزمن، خصوصا عند محاذاة العديد من متواليات عينة. علاوة على ذلك، محاذاة الوقت المنصرم مع متوالية كاملة (التي يمكن ألا تكون هناك إشارة لها) التي تتضمن مصادر حوسبة قيمة.

يمكن أن تقوم النماذج باستخدام المعرفة السابقة لمتوالية مجنحة لضمان التسمية الملائمة لأليلة تكرار ترادفي قصير. على العكس، الطرق الحالي، التي تعتمد على متوالية مرجعية كليا لكل أليلة، تواجه معدلات فشل ملحوظة في مواقف حيث يكون هناك إرة غير كاملة. هناك العديد من الألائل غير المعروفة للمتوالية، وبشكل محتمل بعض من ألائل غير معروفة. للتوضيح، افتراض قطاع تكراري مع باعث تكرار[تي سي تي ايه TCTA] بسيط لها جنيح 3' بادئ مع المتوالية TCAGCTA. بالتالي، يمكن أن تتضمن الإشارة يتضمن هذه المتواليات ك [flank1][TCTA]$_n$TCAGCTA[rest_of_flank2]، حيث "n" عبارة عن أعداد تكرار في الأليلة. قد تختلف الأليلة 9.3 من أليلة 10 بواسطة أن يكون لها حذف في مكان آخر على طول المتوالية. يمكن أن يتم تضمين هذا في الإشارة، على الرغم من أنه يمكن ألا يكون موجودا على الإطلاق. [TCTA]$_7$TCA[TCTA]$_2$ عبارة عن مثال على هذه الأليلة. في بروتوكولات المحاذاة الحالية، أي

قراءة تنتهي بعد $[TCTA]_7$ وقبل [TCTA] النهائية، وسوف تقوم بمحاذاة [flank1][TCTA]$_7$TCAGCTA، بما يجعل هذا الأمر تسمية غير ملائمة.

تسمح النماذج المقدمة هنا لتحديد طول عنصر الحمض النووي الريبي منقوص الأكسجين تكراري متعدد الشكل أو قطاع تكراري موضوع بين المنطقة المجنحة المحفوظة الأولى والمنطقة المجنحة المحفوظة الثانية. في أحد النماذج، تتضمن الطرق توفير مجموعة بيانات تتضمن عينة قراءة واحدة على الأقل عنصر الحمض النووي الريبي منقوص الأكسجين تكراري متعدد الشكل؛ توفير متوالية مرجعية تتضمن المنطقة المجنحة المحفوظة الأولى والمنطقة المجنحة المحفوظة الثانية؛ محاذاة جزء من المنطقة المجنحة الأولى للمتوالية المرجعية إلى عينة القراءة؛ محاذاة جزء من المنطقة المجنحة الثانية للمتوالية المرجعية إلى عينة القراءة؛ وتحديد طول و/أو متوالية القطاع التكراري. في النماذج المثالية، يتم تنفيذ واحد أو أكثر من الخطوات في الطريقة باستخدام كومبيوتر مبرمج بشكل مناسب.

حسب الاستخدام هنا، يشير المصطلح "عينة قراءة" إلى بيانات متوالية التي يتم من أجلها تحديد طول و/أو هوية العنصر التكراري. يمكن أن تكون عينة القراءة قائمة على الحمض النووي الريبي منقوص الأكسجين أو RNA. يمكن أن يتضمن عينة القراءة كل من العنصر التكراري repetitive element، أو جزء منه. يمكن أن تتضمن عينة القراءة أيضا المنطقة المجنحة المحفوظة على أحد أطراف العنصر التكراري (مثلا، المنطقة المجنحة 5')، يمكن أن تتضمن عينة القراءة أيضا المنطقة المجنحة المحفوظة الإضافية على طرف آخر للعنصر التكراري (مثلا، المنطقة المجنحة 3'). في النماذج المثالية، تتضمن عينة القراءة تتضمن بيانات متوالية من ناتج تكبير تفاعل سلسلة بوليمرات لها متوالية بادئة أمامية وخلفية. يمكن الحصول على بيانات المتوالية من أي طريقة اختيار متوالية. يمكن أن تكون عينة القراءة، مثلا، من توالي بواسطة تفاعل التخليق (التوالي بواسطة التخليق)، تفاعل توالي-بواسطة-الارتباط، أو أي طريقة توالي مناسبة التي يكون من أجلها مطلوبا لتحديد طول و/أو هوية العنصر التكراري. يمكن أن تكون عينة القراءة عبارة عن متوالية توافقية (مثلا، متوسطة أو موزونة) مشتقة من قراءات عينة متعددة. في نماذج معينة، توفير متوالية مرجعية تتضمن تحديد موقع محل الاهتمام بناء على المتوالية البادئة لناتج تكبير تفاعل سلسلة بوليمرات.

حسب الاستخدام هنا، يشير المصطلح "عنصر الحمض النووي الريبي منقوص الأكسجين تكراري متعدد الشكل" إلى أي متوالية الحمض النووي الريبي منقوص الأكسجين تكرارية، والذي يمكن الإشارة إليه بقطاع تكراري. يمكن استخدام الطرق المتوفرة هنا لمحاذاة المناطق المجنحة المناظرة لأي متوالية الحمض النووي الريبي منقوص الأكسجين التكرارية. يمكن استخدام الطرق المقدمة هنا لأي منطقة صعبة في المحاذاة، بغض النظر عن فئة التكرار. الطريقة المقدمة هنا مفيدة خصوصا لمنطقة لها مناطق مجنحة محفوظة. بشكل إضافي أو بشكل بديل، تلك الطرق المقدمة هنا مفيدة خصوصا لقراءات عينة والتي تمتد على طول القطاع التكراري شاملا جزء على الأقل من كل المنطقة المجنحة. في النماذج المثالية، عنصر الحمض النووي الريبي منقوص الأكسجين التكراري عبارة عن تكرار ترادفي عددي متغير variable number tandem repeat (VNTR). تكرارات ترادفية عددية متغيرة variable number tandem repeats (VNTRs) عبارة عن تعددات شكلية حيث يتم متوالية تكرار معينة عند تلك الأوقات المتعددة. تتضمن بعض تكرارات ترادفية عددية متغيرة توابع صغيرة و توابع دقيقة، المعروفة أيضا كمتوالية تكرار بسيطة simple sequence repeats (SSRs) أو تكرارات ترادفية قصيرة. في بعض النماذج، القطاع التكراري أقل من 100 نيكلوتيدات، على الرغم من إمكانية محاذاة قطاعات تكرارية أكبر. يمكن أن تكون وحدة التكرار (مثلا، باعث تكرار) للقطاع التكراري 2، 3، 4، 5، 6، 7، 8، 9، 10، 11، 12، 13، 14، 15، 16، 17، 18، 19، 20 أو أكثر من النيكلوتيدات، ويمكن أن يتم التكرار بما يصل إلى 2، 3، 4، 5، 6، 7، 8، 9، 10، 11، 12، 13، 14، 15، 16، 17، 18، 19، 20، 21، 22، 23، 24، 25، 26، 27، 28، 29، 30، 40، 45، 50، 55، 60، 65، 70، 75، 80، 85، 90، 95 أو ما يصل إلى على الأقل 100 مرة أو أكثر.

في نماذج معينة، عنصر الحمض النووي الريبي منقوص الأكسجين تكراري متعدد الشكل عبارة عن تكرار ترادفي قصير. في بعض النماذج، يتم استخدام التكرار الترادفي القصير للأغراض الشرعية. في النماذج المثالية لتطبيقات شرعية، مثلا، يتضمن عنصر الحمض النووي الريبي منقوص الأكسجين تكراري متعدد الشكل بواعث تكرار تترا - أو بنتا-نيكلوتيد، مع ذلك، هذه الطرق المقدمة هنا مناسبة لأي طول لباعث التكرار. في نماذج معينة، القطاع التكراري عبارة عن تكرار ترادفي قصير مثل، مثلا، تكرار ترادفي قصير مختار من مواقع تكرار ترادفي قصير صبغية جسدية لنظام مؤشر الحمض النووي الريبي منقوص الأكسجين المدمج، مواقع نظام مؤشر الحمض النووي

الريبي منقوص الأكسجين المدمج Y- تكرار ترادفي قصير، مواقع EU تكرار ترادفي قصير صبغية جسدية، مواقع EU Y- تكرار ترادفي قصير وما شابه. كمثال، (نظام مؤشر الحمض النووي الريبي منقوص الأكسجين مدمج) قاعدة بيانات عبارة عن مجموعة من مواقع تكرار ترادفي قصير لبية للتحديد بواسطة مجموعة FBI وتتضمن 13 مواقع، CSF1PO، FGA، TH01، TPOX، VWA، D3S1358، D5S818، D7S820، D8S1179، D13S317، D16S539، D18S51 و D21S11. تكرارات ترادفية قصيرة محل الاهتمام الإضافية إلى المجتمع الشرعي والتي يمكن محاذاتها باستخدام الطرق وأنظمة المقدمة هنا تتضمن بنتا D وبنتا E. يمكن تطبيق الطرق وأنظمة المقدمة هنا على أي عنصر الحمض النووي الريبي منقوص الأكسجين تكراري ويتم تقييدها بالـ تكرارات ترادفية قصيرة الموضحة أعلاه.

حسب الاستخدام هنا، يشير المصطلح "متوالية مرجعية" إلى متوالية معروفة والتي تعمل كبنية متراصة في مقابل عينة متوالية يمكن محاذاتها. في النماذج المثالية للطرق وأنظمة المقدمة هنا، تتضمن المتوالية المرجعية على الأقل المنطقة المجنحة المحفوظة الأولى والمنطقة المجنحة المحفوظة الثانية. يشير المصطلح "المنطقة المجنحة المحفوظة" إلى منطقة من متوالية خارج القطاع التكراري (مثلا، تكرار ترادفي قصير). يتم حفظ المنطقة بشكل مثالي عبر العديد من الألائل، حتى من خلال جعل القطاع التكراري متعدد الشكل. سوف تقوم المنطقة المجنحة المحفوظة حسب الاستخدام هنا بشكل مثالي تعقيد أعلى من القطاع التكراري. في النماذج المثالية، يمكن استخدام متوالية مرجعية فردية لمحاذاة كل الألائل في الموقع الجيني. في بعض النماذج، يتم استخدام أكثر من متوالية مرجعية واحدة لمحاذاة متواليات عينة لموقع جيني بسبب الاختلاف في المنطقة المجنحة. مثلا، القطاع التكراري لأميلوجينين Amelogenin له فرق في الجنيحات بين X و Y، خلال مرجع فردي يمكن أن يمثل القطاع التكراري إذا تم تضمين منطقة أطول في المرجع.

في النماذج المقدمة هنا، تتم محاذاة جزء من المنطقة المجنحة لمتوالية مرجعية مع عينة المتوالية. يتم تنفيذ المحاذاة بواسطة تحديد موضع المنطقة المجنحة المحفوظة وبعد ذلك تنفيذ متوالية محاذاة لذلك الجزء من المنطقة المجنحة مع الجزء المناظر من عينة القراءة. يتم تنفيذ محاذاة لجزء من المنطقة المجنحة وفقا لطرق محاذاة معروفة. في نماذج معينة، تتضمن عملية محاذاة جزء من المنطقة المجنحة (مثلا، المناطق المجنحة الأولى والثانية)، (i) تحديد موضع المنطقة المجنحة المحفوظة على عينة القراءة باستخدام توافق وحدة k محددة لمنطقة بذرية تتداخل مع أو تجأور

القطاع التكراري؛ و(ii) محاذاة المنطقة المجنحة مع عينة القراءة. في بعض النماذج، يمكن أن تتضمن المحاذاة محاذاة كل من المتوالية المجنحة ومنطقة مجاورة قصيرة تتضمن جزء من القطاع التكراري.

تم شرح ثمة مثال على هذا المنهج في الشكل 4. تم توضيح ناتج تكبير ("قالب")، والتي يمكن الإشارة إليها كعينة القراءة، في الشكل 4 الذي له تكرار ترادفي قصير لطول و/أو هوية غير معروفة. كما هو موضح أعلاه نسبة إلى الشكل 2، يمكن تحليل عينة القراءة لتعيين عينة القراءة إلى موقع جيني، والتي تكون معروفة بأنها تتضمن تكرار ترادفي قصير في هذه الحالة. بعد تحديد الموقع الجيني لعينة القراءة، يمكن أن يتضمن بروتوكول المحاذاة محاذاة متوالية محددة سلفا لعينة القراءة مع متوالية محددة سلفا للمتوالية المرجعية. مثلا، يتم شرح المواد البادئة كـ p1 وp2، والتي تقوم على أساس المتواليات الأولية التي تم استخدامها لتوليد ناتج التكبير. في النموذج المبين في الشكل 4، يتم استخدام p1 منفردا أثناء خطوة المحاذاة. في بعض النماذج، يتم استخدام p2 فقط لبادئ محاذاة primer alignment. في النماذج الأخرى، يتم استخدام كل من p1 وp2 لبادئ محاذاة. أيضا في النماذج الأخرى، يمكن استخدام متواليات أخرى لخطوة المحاذاة الأولية initial alignment step.

بعد المحاذاة الأولية، تتم محاذاة الجزء المجنح 1، المعين في الشكل 4 مثل "$f1_{al}$". يمكن استباق محاذاة الجزء المجنح 1 بواسطة بذر الجزء المجنح 1، المعين في الشكل 4 مثل "$f1_{seed}$". بذر الجزء المجنح 1 هو تصحيح لعدد صغير (المعين كـ "e") لإندلات بين بداية عينة المتوالية والـ تكرار ترادفي قصير. يمكن أن تكون المنطقة البذرية seeding region موجودة مباشرة بجوار بداية التكرار الترادفي القصير، أو يمكن أن تنطلق (كما في الشكل) لتجنب مناطق منخفضة التعقد. يمكن أن يتم البذر بواسطة توافق وحدة k محددة. تستمر محاذاة الجزء المجنح 1 لتحديد موضع البدء لمتوالية تكرار ترادفي قصير. إذا تم حفظ نمط التكرار الترادفي القصير لتوقع أول نيكلوتيدات nucleotides (s1)، تمت إضافتها إلى المحاذاة بدقة محسنة.

بحيث يكون طول التكرار الترادفي القصير غير معروف، يتم تنفيذ المحاذاة للجزء المجنح 2 كما يلي. يتم تنفيذ بذر الجزء المجنح 2 للعثور على مواضع نهائية محتملة لتكرار ترادفي قصير. حيث البذر للجزء المجنح 1، يمكن أن يتم بدء البذر لتجنب مناطق منخفضة التعقد وسوء المحاذاة. يتم التخلص من بذور الجزء المجنح 2 التي تفشل في المحاذاة. بمجرد تحاذي الجزء المجنح 2 بشكل

مناسب، يمكن تحديد الموضع الطرفي (s2) لتكرار ترادفي قصير. مع بداية متوالية تكرار ترادفي قصير المعروفة عند s1 ونهاية متوالية تكرار ترادفي قصير المعروفة عند s2، يمكن حساب طول التكرار الترادفي القصير.

يمكن تجأور المنطقة البذرية مباشرة بشكل مجاور للقطاع التكراري (مثلا، التكرار الترادفي القصير) و/أو تتضمن جزء من القطاع التكراري. في بعض النماذج، سوف يعتمد موضع المنطقة البذرية على تعقيد المنطقة المباشرة بجوار القطاع التكراري. يمكن ارتباط بداية أو طرف تكرار ترادفي قصير بواسطة متواليات تتضمن تكرارات إضافية أو التي لها تعقيد منخفض. بالتالي، قد يكون مفيدا بدء بذر المنطقة المجنحة لتجنب مناطق منخفضة التعقيد. حسب الاستخدام هنا، يشير المصطلح "منخفضة التعقيد" إلى منطقة ذات متوالية تشابه تلك الخاصة ببواعث التكرار و/أو القطاع التكراري. بشكل إضافي أو بشكل بديل، تقوم منطقة منخفضة التعقيد بدمج تنوعية منخفضة للنيكلوتيدات. مثلا، في بعض النماذج، منطقة منخفضة التعقيد تتضمن متوالية لها أكثر من 30٪، 40٪، 50٪، 60٪، 70٪ أو أكثر من 80٪ هوية متوالية إلى متوالية التكرار. في النماذج المثالية، تقوم المنطقة منخفضة التعقيد بدمج كل من النيكلوتيدات الأربعة عند تردد أقل من 20٪، 15٪، 10٪ أو أقل من 5٪ من كل النيكلوتيدات في المنطقة. يمكن استخدام أي طريقة مناسبة لتحديد منطقة منخفضة التعقيد. هناك طرق لتحديد منطقة منخفضة التعقيد معروفة في المجال، كما هو مبين بواسطة الطرق الموضحة في Morgulis et al., (2006) Bioinformatics. 22(2):134-41، والتي تم دمجها هنا بالإشارة إليها كمرجع. مثلا، كما هو موضح في المواد المدمجة لـ Morgulis et al.، يمكن استخدام لوغاريتم مثل دي يو إس تي DUST لتحديد مناطق في متوالية نيكلوتيد معينة تتمتع بانخفاض التعقيد.

في بعض النماذج، يتم بدء البذر بداية التكرار الترادفي القصير بواسطة على الأقل 1، 2، 3، 4، 5، 6، 7، 8، 9، 10، 15، 20، 25، 30، 35، 40 أو أكثر نيكلوتيدات. في بعض النماذج، يتم تقييم المنطقة المجنحة لتحديد منطقة عالية التعقيد. حسب الاستخدام هنا، يشير المصطلح "منطقة عالية التعقيد high-complexity region" إلى منطقة مع متوالية مختلفة بما يكفي من تلك الخاصة بباعث تكرار و/أو قطاع تكراري يقلل احتمالية حالات سوء المحاذاة. بشكل إضافي أو بشكل بديل، تقوم المنطقة عالية التعقيد بدمج تشكيلة من النيكلوتيدات. مثلا، في بعض النماذج، منطقة عالية التعقيد تتضمن متوالية لها أقل من 80٪، 70٪، 60٪، 50٪، 40٪، 30٪، 20٪ أو أقل من

10٪ هوية إلى متوالية تكرار. في النماذج المثالية، تتضمن المنطقة عالية التعقيد كل من النيكلوتيدات الأربعة عند تردد على الأقل 10٪، 15٪، 20٪، أو على الأقل 25٪ من كل النيكلوتيدات في المنطقة.

حسب الاستخدام هنا، يشير المصطلح "توافق الوحدة k المحددة" إلى طريقة للعثور على محاذاة مطلقة باستخدام طريقة كلامية حيث طول الكلمة محدد بأن له قيمة k. في بعض النماذج، تبلغ قيمة k 3، 4، 5، 6، 7، 8، 9، 10، 11، 12، 13، 14، 15، 16، 17، 18، 19، 20، 21، 22، 23، 24، 25، 26، 27، 28، 29، 30، 31، 32، 33، 34، 35، 36، 37، 38، 39، 40 أو أكثر نيكلوتيدات في طول. في النماذج المثالية، تتمتع k بقيمة ما بين 5 و30 نيكلوتيدات في طول. في بعض النماذج، تتضمن k بقيمة بين 5 و16 نيكلوتيدات في طول. في نماذج معينة، يتم اختيار k بواسطة النظام أو المستخدم بناء على واحد أو أكثر من العوامل. مثلا، إذا كانت المنطقة المجنحة قصيرة، مثل عند وضع المتوالية البادئة قريبة نسبيا إلى متوالية تكرار ترادفي قصير، يمكن تقليل k بشكل مناسب. في النماذج المثالية، يتم اختيار k للعثور ع7لى كل حالات توافق في مسافة -/+ e.

تقوم الطرق الخاصة بالكلمة بتحديد سلسلة من متواليات ثانوية قصيرة غير متداخلة ("كلمات") في المتوالية الاستفهامية التي يتم توافقها بعد ذلك مع قاعدة بيانات متواليات مرشحة. يتم استخلاص المواضع النسبية للكلمة في اثنين من المتواليات التي تتم مقارنتها للحصول على بداية؛ وسوف يشير هذا إلى منطقة محاذاة إذا كانت الكلمات المميزة المتعددة لنفس البداية. فقط إذا تم الكشف عن هذه المنطقة يمكن أن تكون هذه الطرق بتطبيق معايير محاذاة حساسة؛ بالتالي، هناك العديد من المقارنات غير الضرورية مع متواليات ذات تشابه مفهوم مستبعدة. هناك طرق لتنفيذ توافق وحدة k، شاملة توافق وحدة k محددة، معروفة في المجال، كما هو مبين بواسطة الكشف الخاص بـ Lipman, et al., (1985) *Science* 227:1435–41, and of Altschul, et al., (1990) *Journal of Molecular Biology* 215:403–410، وقد تم دمج محتوى كل منها هنا بالإشارة إليها كمرجع.

حسب الاستخدام هنا، يشير المصطلح "ناتج تكبير amplicon" إلى أي منتج تكبير مناسب يتم الحصول على المتوالية الخاصة بها. بشكل مثالي، منتج التكبير عبارة عن منتج لطريقة تكبير انتقائية، باستخدام مواد بادئة مستهدفة معينة، مثل مواد بادئة تفاعل سلسلة بوليمرات. في نماذج معينة، تكون بيانات المتوالية من ناتج تكبير تفاعل سلسلة بوليمرات لها متواليات أولية متجهة إلى

الأمام والخلف. في بعض النماذج، يمكن أن تتضمن طريقة التكبير الانتقائية واحد أو أكثر من خطوات التكبير غير الانتقائية. مثلا، يمكن إتباع عملية تكبير باستخدام مواد بادئة عشوائية أو متحللة بواسطة واحد أو أكثر من دورات تكبير باستخدام مواد بادئة مستهدفة معينة. تتضمن الطرق المناسبة لعملية تكبير انتقائية، على سبيل المثال لا الحصر، تفاعل سلسلة بوليمرات، تكبير إزاحة جديلةstrand displacement amplification (SDA)، تكبير ناتج الانتساخ transcription mediated amplification (TMA) والتكبير القائمةbased amplification على أساس متوالية الحمض النوويnucleic acid sequence (NASBA)، كما هو موضح في براءة الاختراع الأمريكية رقم 8،003،354، والتي تم دمجها هنا بالإشارة إليها كمرجع. يمكن استخدام طرق التكبير المذكورة أعلاه لتكبير واحد أو أكثر من الأحماض النووية محل الاهتمام بشكل انتقائي. مثلا، تفاعل سلسلة بوليمرات، شاملا مضاعفة تفاعل سلسلة بوليمرات، تكبير إزاحة جديلة، تكبير ناتج الانتساخ، التكبير القائمة على أساس متوالية الحمض النووي وما شابه يمكن استخدامها لتكبير واحد أو أكثر من الأحماض النووية محل الاهتمام بشكل انتقائي. في هذه النماذج، يتم تضمين المواد البادئة المتعلقة خصيصا بالحمض النووي محل الاهتمام في تفاعل التكبير. يمكن أن تتضمن الطرق المناسبة الأخرى لتكبير الأحماض النووية مد وارتباط أوليجو نيكلوتيد، تكبير الدائرة الدوارة (Lizardi et al., Nat. Genet. 19:225-232 (1998)، والتي تم دمجها هنا بالإشارة إليها كمرجع) وتقنيات فحص ارتباط أوليجو نيكلوتيد oligonucleotide ligation assay (OLA) (انظر بشكل عام براءة الاختراع الأمريكية أرقام 7582420، 5185243، 5679524 و5573907؛ براءات الاختراع الأوربية أرقام 308 320 0 ب1؛ و 731 336 0 ب1؛ و182 439 0 ب1 ؛ براءات الاختراع الدولية أرقام 01069/90؛ و12696/89؛ و 09835/89، والتي تم دمجها هنا بالإشارة إليها كمرجع).

سوف يتم فهم أن طرق التكبير يمكن تعيينها لتكبير حمض نووي مستهدف محل الاهتمام انتقائيا. مثلا، في بعض النماذج، يمكن أن يتضمن تكبير الطريقة الانتقائية مسبار ارتباط تكبير أو فحص تفاعل ارتباط أوليجو نيكلوتيد التي تحتوي على مواد بادئة موجهة خصوصا إلى الحمض النووي محل الاهتمام. في بعض النماذج، يمكن أن تتضمن طريقة التكبير الانتقائية تفاعل ارتباط بادئ مد الذي يحتوي على مواد بادئة موجهة خصيصا بالحمض النووي محل الاهتمام. كمثال غير محدد على بادئ مد وارتباط المواد البادئة التي يمكن تعيينها خصيصا لتكبير الحمض النووي محل

الاهتمام، يمكن أن يتضمن التكبير مواد بادئة مستخدمة لفحص GoldenGate™ (Illumina, Inc.، سان دييجو، سي ايه CA)، كما هو موضح في براءة الاختراع الأمريكية رقم 7582420، والتي تم دمجها هنا بالإشارة إليها كمرجع. الطرق الحالية غير مقيدة بأي تقنية تكبير وتقنيات التكبير الموضحة هنا وهي مثالية فقط نسبة إلى طرق ونماذج الكشف الحالي.

تتهجن مواد بادئة Primers لتكبير عنصر الحمض النووي الريبي منقوص الأكسجين تكراري بشكل مثالي إلى المتواليات الفريدة للمناطق المجنحة. يمكن تعيين المواد البادئة وتوليدها وفقا لأي طريقة مناسبة. سوف يكون معروفا تصميم مواد بادئة للمناطق المجنحة لقطاعات تكرارية في المجال، كما هو مبين بواسطة Zhi, et al. (2006) *Genome Biol*, 7(1):R7، والتي تم دمجها هنا بالإشارة إليها كمرجع. مثلا، يمكن تصميم المواد البادئة يدويا. وهذا يتضمن بحث متوالية الحمض النووي الريبي منقوص الأكسجين الجينومية لتكرارات تابع، والتي يمكن تنفيذها بواسطة العين أو باستخدام الأدوات الآلية مثل برنامج RepeatMasker. بمجرد تحديد القطاعات التكرارية والمناطق المجنحة المناظرة، يمكن استخدام المتواليات المجنحة flanking sequences لتصميم أوليجو نيكلوتيدات بادئة oligonucleotide primers والتي سوف تقوم بتكبير التكرار المعين في تفاعل سلسلة بوليمرات.

يقوم ما يلي بوصف الأمثلة التي تم تنفيذها وفقا للوصف أعلاه.

المثال 1 - محاذاة الموقع D18S51

يصف هذا المثال محاذاة الموقع D18S51 وفقا لأحد النماذج. تتضمن بعض المواقع متواليات مجنحة والتي كانت منخفضة التعقيد وتشابه متوالية تكرار ترادفي قصير. يمكن أن يؤدي هذا إلى عدم توالي المتوالية المجنحة (أحيانا مع متوالية تكرار ترادفي قصير نفسها) وبالتالي يمكن إساءة تسمية الأليلة. ثمة مثال على موقع مسبب للمشاكل وهو D18S51. باعث التكرار هو [AGAA]n AAAG AGAGAG. تم توضيح المتوالية المجنحة أدناه مع "مشكلة" المتوالية منخفضة التعقيد الموضوع تحتها خط،

GAGACCTTGTCTC (STR)

GAAAGAAAGAGAAAAAGAAAAGAAATAGTAGCAACTGTTAT

إذا تم استخدام المنطقة المجنحة المجاورة مباشرة لـ تكرار ترادفي قصير لبذر المحاذاة، سوف يتم توليد وحدات k مثل GAAAG، AAAGAA، AGAGAAA، والتي تخطط متوالية تكرار ترادفي قصير. وهذا يعيق الأداء حيث يتم الحصول على العديد من الاحتماليات من البذر، ولكن الأهم،

يقوم المنهج بتكوين حالات سوء المحاذاة، مثل تلك المبينة في الشكل 5. في المتواليات المبينة في الشكل 5، يتم تحديد متوالية تكرار ترادفي قصير الحقيقة، تم وضع خط تحت متوالية تكرار ترادفي قصير الناتجة من سوء المحاذاة وتم توضيح أخطاء القراءة بخط سميك.

بالنسبة لهذه الأجزاء المجنحة منخفضة التعقيد، تم التأكيد على أن المناطق البذرية seeding regions في المنطقة منخفضة التعقيد بواسطة ضغطها بعيدا عن متوالية تكرار ترادفي قصير. بينما يتطلب هذا قراءات أطول لتسمية الـ تكرار ترادفي قصير، سوف تضمن الدقة العالية وتمنع سوء المحاذاة للمنطقة المجنحة مع متوالية تكرار ترادفي قصير (أو أجزاء أخرى من الجزء المجنح). لا تزال تتم محاذاة الجزء المجنح منخفض التعقيد مع القراءة للعثور على الموضع الطرفي لتكرار ترادفي قصير ولكن بسبب بذر المحاذاة مع متوالية عالية التعقيد، يجب أن يكون في الموضع الصحيح.

المثال 2 - محاذاة الموقع بنتا-D بواسطة إضافة متوالية تكرار ترادفي قصير قصيرة

تميل مجموعة من متواليات بنتا-D إلى أن تتضمن تكرارات ترادفية قصيرة التي كانت أقصر بمقدار 1 nt مما هو متوقع. عند المزيد من الفحص، سوف يتم اكتشاف أن كل من الأجزاء المجنحة احتوت على امتداد متعددة A- وخطوات توالي / تكبير غالبا ما تتم إزالة واحد من A's في هذه الامتدادات. كما هو مبين في المتوالية أدناه، يتم العثور على امتدادات A متجانسة متعددة الشكل على كل الأجزاء المجنحة.

[AAAGA]n CAAGAAAGAAAAAAAAG ...

...AAAAACGAAGGGGAAAAAAAGAGAAT

سوف يؤدي خطأ القراءة باستخدام حذف في الجزء المجنح الأول إلى اثنين من عمليات المحاذاة الحيوية بنفس الطريقة،

قراءة، ...CAAGAAAGAAAAAAA-GA...

الجزء المجنح، ...CAAGAAAGAAAAAAAAG- (2 إندلات)

قراءة، ...CAAGAAAGAAAAAAAGA... (2 حالة عدم توافق)

الجزء المجنح، ...CAAGAAAGAAAAAAAAG

إجبار القاعدة الأقرب إلى التكرار الترادفي القصير لتكون عبارة عن توافق لا يعمل بسبب واحد من الأجزاء المجنحة في واحد من التكرارات الترادفية القصيرة التي انتهت لها تعدد شكلي لنيكلوتيد فردي فيها. سوف يتم اكتشاف أن إضافة اثنين من النيكلوتيدات فقط من متوالية تكرار ترادفي قصير قد حلت المشكلة،

قراءة، CAAGAAAGAAAAAAA-GAA...

الجزء المجنح، CAAGAAAGAAAAAAAAGAA... (1 إندل) ✔

قراءة، CAAGAAAGAAAAAAAG-AA... (1 إندل + 1 حالة عدم توافق)

الجزء المجنح، CAAGAAAGAAAAAAAAGAA...

المثال 3 – تحليل خليط من عينات الحمض النووي الريبي منقوص الأكسجين

تم تحليل خليط العينات باستخدام الطرق المقدمة هنا لعمل تسميات لكل موقع في لوحة تكرارات ترادفية قصيرة شرعية. لكل موقع، تم عد عدد القراءات المناظرة لكل أليلة ولكل متوالية مختلفة لتلك الأليلة.

تم توضيح النتائج المثالية في الأشكال 6أ–6د. كما هو مبين، يمثل العمود على اليمين من كل زوج البيانات الفعلية التي تم الحصول عليها، التي تشير إلى تناسب قراءات لكل أليلة. تمثل الظلال المختلفة متواليات مختلفة. يتم استبعاد الألائل التي لها أقل من 0.1٪ من عدد قراءة الموقع ومتواليات التي لها أقل من 1٪ من عدد الألائل. العمود على يسار كل زوج يمثل النسب النظرية (غير المضطربة). تمثل الظلال المختلفة الحمض النووي الريبي منقوص الأكسجين ضابطة مختلفة في المدخل كما هو مشار إليه في مفتاح الشكل. في الشكل 6أ–6د، المحور x خاص بالأليلة، والمحور Y يشير إلى تناسب قراءات مع الأليلة المشار إليها.

كما هو مبين في الشكل، قام منهج تسمية التكرار الترادفي القصير باستخدام الطرق المقدمة هنا بتحقيق تسميات يتم تحقيقها بشكل مدهش لكل أليلة في اللوح.

المثال 4 – تحليل لوح تكرار ترادفي قصير شرعي

تم تحليل لوح ذو 15 موقع مختلف في 5 من العينات المختلفة. تم الحصول على العينات من Promega Corp، وعينات متضمنة 9947A، K562، 2800M، NIST، A و B (SRM 2391c).

تم اختيار المواقع من مرقمات الـ نظام المؤشر للحمض النووي الريبي منقوص الأكسجين المدمج STR الشرعية والتي تتضمن CSF1PO، D3S1358، D7S820، D16S539، D18S51، FGA، PENTAE، TH01، vWA، D5S818، D8S1179، D13S317، D21S11، PENTAD وTPOX باستخدام طريقة المحاذاة المقدمة هنا. بشكل مختصر، تم تكبير المرقمات باستخدام مواد بادئة معيارية standard primers، كما هو مذكور في Krenke, et al. (2002) *J. Forensic Sci.* 47(4): 773-785، والتي تم دمجها هنا بالإشارة إليها كمرجع. تم نقع نواتج التكبير وتم الحصول على بيانات التوالي باستخدام 460x1 دورات على توالي أداة MiSeq (Illumina، سان دييجو، CA).

تم تنفيذ المحاذاة وفقا للطرق المقدمة هنا. كما هو مذكور في الشكل 7، تم توضيح نسبة 100٪ توافق للعينات الضابطة هذه مقارنة بالعينات الضابطة. علاوة على ذلك، قامت هذه طريقة بتحديد تعدد شكلي لنيكلوتيد فردي غير معروف سلفا في أحد العينات للمرقم D8S1179، والتي تبين أيضا أداة قوية لتحليل تكرار ترادفي قصير قائم على أساس متوالية عند الدمج بطرق المحاذاة المقدمة هنا.

الشكل 8 يشرح طريقة 160 لتحديد منتج مضطرب. بعد تحديد المناطق محل الاهتمام في القراءات المعينة، يمكن أن تقوم النماذج المذكورة أعلاه بتصنيف، عند 162، المناطق محل الاهتمام (أو القراءات المعينة) بناء على متواليات المناطق محل الاهتمام. كما هو موضح أعلاه، في ظروف معينة، يمكن أن يقوم بروتوكول المحاذاة بتحليل جزء من واحد أو كل من المناطق المجنحة flanking regions بالإضافة إلى متوالية القطاع التكراري. بالتالي، في نماذج معينة، يمكن أن يتضمن التصنيف، عند 162، تصنيف بناء على متوالية القطاع التكراري ومتوالية ثانوية لواحد أو كل من المناطق المجنحة. كمثال، يمكن أن يتضمن التصنيف تحليل القطاع التكراري والقليل من النيكلوتيدات لكل من المناطق المجنحة التي تمتد من القطاع التكراري. في النماذج الأخرى، يمكن أن يتضمن التصنيف، عند 162، تصنيف يقوم على منطقة محل الاهتمام يتضمن فقط متوالية من القطاع التكراري.

يمكن تصنيف المناطق محل الاهتمام (أو قطاعات تكرارية) بحيث يتم تعيين الـ المناطق محل الاهتمام (أو قطاعات تكرارية) ذات المتواليات المختلفة على أنها ألائل محتملة أو (مشكوك فيها). مثلا، يمكن أن تتضمن كل أليلة محتملة عينة متوالية فريدة و/أو طول فريد. بشكل أكثر تحديدا، يمكن أن تتضمن كل أليلة محتملة متوالية فريدة من المنطقة محل الاهتمام أو قطاع تكراري و/أو

طول فريد من المنطقة محل الاهتمام أو قطاع تكراري. كما هو موضح أدناه، في بعض النماذج، يمكن ترتيب القطاعات التكرارية بناء على اسم أليلة الاستشراد الكهربي.

يمكن تنفيذ التصنيف، عند 162، لكل موقع معين. بعد تعيين قراءات العينة إلى المواقع الجينية المناظرة، يمكن أن يتضمن كل موقع جيني مجموعة من قراءات معينة مرتبطة بالموقع الجيني. مثلا، في بعض النماذج، يمكن أن يتضمن واحد أو أكثر من المواقع الجينية مئات من القراءات المعينة التي يتم تجميعها أو دمجها مع بعضها البعض. كما هو معروف، يمكن أن يتضمن موقع جيني مناظر، مثل موقع تكرار ترادفي قصير معروف، مجموعة من ألائل حيث تتضمن كل أليلة متوالية مختلفة. بواسطة التحليل الجماعي لمجموعة القراءات المعينة التي تم تحديدها على أنها جزء من موقع جيني مشترك، يمكن تحليل مجموعة القراءات المعينة لتوفير تسمية نمط جيني لفرد أو مجموعة من الأفراد.

يمكن أن تتضمن الطريقة 160 عد (أو اختصار)، عند 164، القراءات المعينة لموقع جيني مشترك يتضمن متوالية مشتركة. يمكن أن يتضمن العد، عند 164، تحديد درجات عد كما هو موضح هنا. مثلا، يتضمن الشكل 9 جدول 190 يتضمن الألائل المحتملة لموقع D1S1656، ويتضمن الشكل 10 رسم بياني 192 يشرح توزيع ألائل الاستشراد الكهربي. تتم تسمية ألائل الاستشراد الكهربي وفقا للاتفاقية و، كما هو مبين في الشكل 10، يتضمن بشكل محتمل منتج مضطرب. في هذا المثال، بعد توالي الأحماض النووية من مصدر منفرد، تم تحليل قراءات العينة لتحديد المناطق محل الاهتمام (مثلا، قطاعات تكرارية) للموقع D1S1656. تم تصنيف الـ المناطق محل الاهتمام وتم العد لتحديد عدد الألائل المحتملة في الموقع D1S1656. في هذا المثال، لم يتم اعتبار الألائل التي لها أعداد أقل من 1٪ من العدد الكلي لقراءات معينة للموقع D1S1656. كما هو مبين في الشكل 9، تضمنت القراءات المعينة المرشحة إجمالي أربعة متواليات فريدة، والتي يمكن اعتبارها ألائل محتملة للموقع D1S1656. بعد التحليل، كما هو موضح أدناه، تسمية النمط الجيني للموقع غير متجانسة 13/12.

في بعض النماذج، بناء على درجات العد للألائل المحتملة لموقع جيني، يمكن إجراء تسمية نمط جيني للموقع الجيني. في بعض النماذج، مع ذلك، يمكن تنفذي المزيد من التحليل للمتواليات. مثلا، يمكن أن تتضمن الطريقة 160 تحليل، عند 166، متواليات الألائل المحتملة لتحديد ما إذا كانت الأليلة الأولى عبارة عن منتج مضطرب مشكوك فيه من أليلة ثانية. الاضطراب عبارة عن

ظاهرة تتم أثناء تكبير الحمض النووي، خصوصا الأحماض النووية التي تتضمن واحد أو أكثر من سلسلة من بواعث تكرار، مثل تلك الموجودة في ألائل تكرار ترادفي قصير. المنتجات المضطربة لها متواليات بشكل مثالي عبارة عن واحد أو أكثر من بواعث تكرار أقل في الحجم (أو أكبر في الحجم) من الأليلة الحقيقة. أثناء تكاثر متوالية الحمض النووي، يمكن أن تنفصل اثنين من الجدائل على طول الـ تكرار ترادفي قصير. حيث كل باعث تكرار هو نفسه، يمكن أن تتم إعادة تصلب الاثنين من الجدائل بشكل غير مناسب بحيث تكون الجديلتين مزاحتين بواسطة واحد أو أكثر من بواعث تكرار. بالتالي، يختلف المنتج الناتج، والذي يمكن تكبيره بشكل أكبر، عن المتوالية الحقيقية بواسطة واحد أو أكثر من بواعث تكرار.

لأن المنتجات المضطربة هي تقريبا هي ذات نفس الحجم مثل الأليلة الحقيقة، قد يكون من الصعب تحديد ما إذا كان منتج مضطرب عبارة عن أليلة حقيقية للموقع الجيني أو منتج مضطرب لأليلة مجاورة. بالتالي، يمكن أن يقوم المنتج المضطرب بتقليل ثقة تسمية نمط جيني. في ظروف معينة، يمكن أن يقوم المنتج المضطرب بمنع توفير تسمية نمط جيني أو التسبب في تسمية نمط جيني غير صحيحة. يمكن أن يقوم منتج مضطرب بالحصول على تسميات النمط الجيني للعينات التي تتضمن مجموعة من مصادر صعبة تحديدا.

يمكن أن يقوم التحليل، عند 166، بتحديد ما إذا كانت الأليلة الأولى عبارة عن منتج مضطرب مشكوك فيه من أليلة ثانية. في بعض النماذج، يتضمن التحليل استخدام واحد أو أكثر من القواعد أو الظروف على متواليات الألائل الأولى والثانية. مثلا، يمكن أن تكون الأليلة الأولى عبارة عن منتج مضطرب مشكوك فيه للأليلة الثانية إذا تم تحديدها، عند 171، حيث بواعث تكرار k والتي تمت إضافتها أو إسقاطها بين الألائل الأولى والثانية. العدد k عبارة عن عدد كامل. في نماذج معينة، العدد k عبارة عن 1 أو 2. على الرغم من أن المنتج المضطرب بشكل مثالي يتضمن باعث تكرار واحد أقل، يمكن أيضا أن يتضمن المنتج المضطرب اثنين من بواعث التكرار الأقل أو باعث تكرار واحد مضاف. من الممكن أن يتضمن المنتج المضطرب أيضا فروق أخرى في بواعث تكرار. يمكن أن يتضمن التحليل، عند 166، مقارنة كل أليلة محتملة مرتبطة بموقع جيني بكل أليلة أخرى محتملة من نفس الموقع الجيني.

في بعض النماذج، يمكن أن يتضمن التحليل، عند 166، تحديد، عند 172، باعث التكرار الذي تمت إضافته أو إسقاطها. تحديد، عند 172، يمكن أن تتضمن بواعث التكرار التي تمت إضافتها

أو إسقاطها محاذاة المتواليتين لاثنين من الألائل على طول المناطق محل الاهتمام أو قطاعات تكرارية لتحديد عند إسقاط أو إضافة باعث تكرار. مثلا، يمكن محاذاة المتواليات مع بعضها البعض عند أحد الطراف لتحديد عند إضافة أو إسقاط باعث تكرار.

بشكل بديل أو بالإضافة إلى ما سبق، يمكن أن يتضمن التحليل، عند 173، مقارنة أطوال القطاعات التكرارية للألائل الأولى والثانية لتحديد إذا كانت أطوال القطاعات التكرارية بين الألائل الأولى والثانية تختلف بمقدار طول باعث تكرار واحد أو بواعث تكرار متعددة. مثلا، في المثال المبين في الشكل 9، باعث التكرار عبارة عن تي ايه جي ايه TAGA، والذي هو عبارة عن تترا نيكلوتيد له أربعة نيكلوتيدات. تم توضيح أطوال متوالية الألائل المستهدفة في الشكل 9. تتضمن كل من الأليلة 1 والأليلة 2 62 نيكلوتيدات، وكل من الأليلة 3 والأليلة 4 تتضمن 58 نيكلوتيدات. بالتالي، يختلف طول المتوالية للأليلة 1 من متوالية أليلة 3 ومتوالية أليلة 4 بواسطة أربعة نيكلوتيدات أو، بمعنى آخر، طول باعث التكرار. بالمثل، يختلف طول المتوالية للأليلة 2 عن متوالية أليلة 3 ومتوالية أليلة 4 بواسطة طول باعث التكرار.

في بعض النماذج، يمكن أن يتضمن التحليل، عند 166، تحديد، عند 174، ما إذا كانت باعث التكرار الذي تمت إضافته أو إسقاطه مطابق لباعث تكرار مجاور في نفس المتوالية. كما هو موضح أعلاه، يمكن تحديد باعث التكرار الذي تمت إضافته أو إسقاطه بواسطة محاذاة متواليات أليلة لتحديد باعث التكرار التي تمت إضافتها أو إسقاطها. بعد محاذاة المتواليات، قد يتم تحديد أن باعث التكرار الذي تمت إضافته/إسقاطه متطابق مع باعث التكرار المجاور له. في بعض النماذج، يمكن تحقيق المحاذاة باستخدام لوغاريتم شره.

تتضمن الأليلة الأولى (أو الأليلة المشكوك فيها بأنها عبارة عن منتج مضطرب) بشكل مثالي تتضمن عدد قراءة (أو درجة عد) أقل من عدد القراءة (أو درجة العد) للأليلة الثانية. في ظروف معينة، مثل عندما العينة مساهم أقل، يمكن ألا يكون هذا حقيقا. في بعض الحالات، يمكن أن يكون المنتج المضطرب للأليلة أقل من قيمة حدية مضطربة معينة أو تقع في نطاق محدد سلفا للموقع و/أو الأليلة. يمكن أن تقوم القيمة الحدية المضطربة على، مثلا، عدد أعداد قراءة للأليلة الثانية، البيانات التاريخية للموقع و/أو الأليلة المناظرة، و/أو ملاحظات على الموقع المناظر و/أو الأليلة أثناء الفحص. لتوفير مثال يتعلق بالبيانات التاريخية أو ملاحظات الأليلة، يمكن تحديها خلال الخبرة التي تتعلق بفحص معين لأليلة يقدم كمية محددة سلفا من حالة اضطرابية أكبر من

أو أقل مما هو متوقع بشكل عام. يمكن استخدام هذه البيانات و/أو الملاحظات لتعديل القيمة الحدية. كمثال آخر حيث يمكن أن تقوم معرفة الأليلة بالتأثير على القيمة الحدية المضطربة، يمكن أن تقدم الألائل ذات الأطول المتوسطة نسبة أكبر من منتج مضطرب. بالتالي، يمكن تغيير القيمة الحدية المضطربة بناء على طول الأليلة.

في بعض النماذج، يمكن أن يتضمن التحليل، عند 166، تحديد، عند 175، ما إذا كانت درجات العد للأليلة الأولى تقع في نطاق محدد سلفا من درجات العد للأليلة الثانية. مثلا، إذا كانت درجات العد (مثلا، أعداد قراءة) للأليلة الأولى في نطاق نسبة محددة سلفا من درجات العد (مثلا، أعداد قراءة) للأليلة الثانية، فإن الأليلة الأولى يمكن أن تكون عبارة عن منتج مضطرب مشكوك فيه. يمكن أن تتفاوت نسبة محددة سلفا بين حوالي 5٪ وحوالي 40٪. في نماذج معينة، يمكن أن يكون نطاق النسبة المحدد سلفا ما بين حوالي 10٪ وحوالي 30٪ أو بين حوالي 10٪ وحوالي 25٪. يمكن حساب نطاق النسبة المحدد سلفا أو يتم الحصول عليه باستخدام البيانات التاريخية أو ملاحظات موقع تكرار ترادفي قصير المناظر أثناء الفحص. بالمثل، إذا كانت درجات العد للأليلة الأولى أقل من القيمة الحدية المضطربة المعينة التي تقوم على درجات العد للأليلة الثانية، يمكن بالتالي أن تكون الأليلة الأولى عبارة عن منتج مضطرب مشكوك فيه. مثلا، يمكن أن تكون القيمة الحدية المضطربة المعينة بناء على تناسب محدد سلفا لدرجة العد للأليلة الثانية. مثلا، يمكن أن يكون التناسب المحدد سلفا حوالي 20٪، 25٪، 30٪، 35٪، أو 40٪. يمكن تحديد التناسب المحدد سلفا أو الحصول عليه باستخدام البيانات التاريخية لتكرار ترادفي قصير مناظر أو ملاحظات لموقع تكرار ترادفي قصير المناظر أثناء الفحص.

في بعض النماذج، يمكن استخدام درجات العد للأليلة المحتملة لتحديد قياس مضطرب (أو درجة الاضطراب). يمكن أن يكون القياس المضطرب عبارة عن قيمة أو وظيفة قائمة على درجات عد للأليلة الأولى. يمكن أن يكون القياس المضطرب بناء على درجات العد للأليلة الثانية. يمكن مقارنة القياس المضطرب بالقيمة الحدية المضطربة المعينة لتحديد ما إذا كانت الأليلة المحتملة المناظرة عبارة عن منتج مضطرب مشكوك فيه. إذا كان القياس المضطرب أقل من القيمة الحدية المضطربة المعينة، فإن الأليلة الأولى يمكن أن تكون عبارة عن منتج مضطرب مشكوك فيه للأليلة الثانية. إذا لم يكن القياس المضطرب أقل من القيمة الحدية المضطربة المعينة، فإنه يمكن

اعتبار الأليلة الأولى على أنها أليلة محتملة. في هذه الحالة، يمكن أن تكون كل من الأليلة الأولى والأليلة الثانية عبارة عن ألائل حقيقية للموقع.

يمكن تطبيق ظروف إضافية لتحديد ما إذا كان أليلة واحدة عبارة عن المنتج المضطرب لأليلة أخرى. مثلا، يمكن أن يتضمن التحليل، عند 166، تحديد، عند 176، أنه لا توجد حالات عدم توافق موجودة بين متواليات الألائل الأولى والثانية. يمكن أن يتم تحليل المناطق محل الاهتمام أو، بشكل أكثر تحديدا، القطاعات التكرارية لتحديد أي حالات عدم توافق بين المتواليات المناظرة. مثلا، إذا كان نيكلوتيد لمتوالية واحدة غير متوافق بواسطة النيكلوتيد الخاص بالمتوالية الأخرى (غير باعث تكرار الذي تمت إضافته/إسقاطه)، فإن المتواليات يمكن ألا تكون عبارة عن منتجات مضطربة.

في النماذج الأخرى، قد يتم تحديد أن المنتج المضطرب المشكوك فيه ليس منتج مضطرب من أليلة ثانية. بدلا من ذلك، يمكن أن يكون المنتج المضطرب المشكوك فيه من مساهم آخر أو يمكن التسبب فيه بواسطة خطأ توالي. مثلا، يمكن أن يقوم واحد أو أكثر من النماذج بتحديد أن المنتج المضطرب المشكوك فيه من مساهم آخر إذا كانت القياس المضطرب (مثلا، درجة العد أو وظيفة أخرى بناء على درجة العد) للأليلة الأولى أكبر من القيمة الحدية المضطربة المعينة. يمكن أن تكون القيمة الحدية المعينة قائمة على درجة العد للأليلة الثانية ووظيفة مضطربة محددة سلفا، والتي يمكن أن تكون قائمة على البيانات التاريخية و/أو البيانات في الفحص محل الاهتمام. يمكن أن يقوم واحد أو أكثر من النماذج بتحديد أن المنتج المضطرب المشكوك فيه عبارة عن خطأ توالي إذا كان القياس المضطرب للأليلة الأولى أقل من قيمة أساسية. يمكن أن تكون القيمة الأساسية قائمة على درجة العد للأليلة الثانية ووظيفة مضطربة محددة سلفا، والتي يمكن أن تكون قائمة على البيانات التاريخية و/أو البيانات في الفحص محل الاهتمام. مثلا، يمكن أن يتضمن موقع معين نطاق مضطرب بمقدار 10-30٪. إذا كان عدد القراءة للأليلة الثانية للموقع المعين يبلغ 100، بعد ذلك يمكن أن تكون الأليلة الأولى عبارة عن خطأ توالي إذا كان عدد القراءة أقل من 10. يمكن أن تكون الأليلة الأولى بشكل محتمل من مساهم آخر إذا كان عدد القراءة أكبر من 30.

في نماذج معينة، يتم اعتبار الأليلة الأولى عبارة عن منتج مضطرب من أليلة ثانية، إذا كان، (أ) متواليات الأليلة للألائل الأولى والثانية تختلف في الطول بواسطة بواعث تكرار k؛ (ب) يتطابق

باعث تكرار الذي تم إسقاطه أو إرساله مع باعث التكرار المجاور؛ (ج) لا يوجد هناك حالات عدم توافق أخرى بين اثنين من ألائل (مثلا، المناطق محل الاهتمام أو قطاعات تكرارية)؛ و، اختياريا، (د) القياس المضطرب للأليلة الأولى موجود في نطاق مضطرب محدد سلفا (أو أقل من القيمة الحدية المضطربة المعينة) للقياس المضطرب للأليلة الثانية.

بالعودة إلى المثال المبين في الشكل 9، سوف تكون المتواليات للاثنين من الألائل الحقيقية للموقع D1S1656 هي [TAGA]11[TAGG]1[TG]5 للأليلة 12 و[TAGA]13[TG]5 للأليلة 13. تتضمن الأليلة 12 تعدد شكلي لنيكلوتيد فردي في آخر وحدة تكرار "TAGA". بالتالي، يمكن أن نحدد أن متوالية الأليلة 12 [TAGA]12[TG]5 في الحقيقة هو −1 مضطرب من الأليلة 13، ومتوالية الأليلة 13 [TAGA]12[TAGG]1[TG]5 هو +1 مضطرب من الأليلة 12. كما هو موضح، يمكن أن تكون النماذج المذكورة أعلاه مفيدة عبر أنظمة الاستشراد الكهربي. بشكل أكثر تحديدا، لن تكون أنظمة الاستشراد الكهربي قادرة على أن متوالية الأليلة 12 [TAGA]12[TG]5 هو −1 مضطرب من أليلة 13، ومتوالية الأليلة 13 [TAGA]12[TAGG]1[TG]5 هو +1 مضطرب من الأليلة 12.

الشكل 11 يشرح طريقة 200 لتحليل بيانات توالي وفقا لأحد النماذج. يمكن دمج الطريقة 200 مع النماذج الأخرى المذكورة أعلاه. الطريقة 200 تتضمن تلقي، عند 202، بيانات توالي شاملة مجموعة من قراءات عينة تمت تهيئتها لتناظر مجموعة من المواقع الجينية. يمكن تهيئة مجموعة المواقع الجينية لتطبيق جيني محدد سلفا، مثل الاختبارات الشرعية أو إثبات النسب. يمكن أن تقوم قراءات العينة بتكوين أزواج قراءة تناظر نواتج تكبير حيث يتضمن كل زوج قراءة أولى وقراءة ثانية لناتج التكبير المناظر. مثلا، يمكن الحصول على أزواج القراءة الأول والثاني من توالي مزدوج الطرف و، في نماذج معينة، توالي غير متناظر طرفي مزدوج. يمكن أن تتضمن كل من القراءات الأولى والثانية متوالية مناظرة، المشار إليها فيما يلي بمتوالية قراءة. يمكن أن تتضمن كل متوالية قراءة، مثلا، متواليات التحديد (مثلا، متواليات أولية) ومتوالية تتضمن اختلاف متوالية، مثل تعدد شكلي لنيكلوتيد فردي أو تكرار ترادفي قصير.

يمكن أن تتضمن الطريقة 200 تحديد، عند 204، واحد أو أكثر من مواقع جينية محتملة لقراءات العينة. يمكن أن تكون عملية التحديد مشابهة للتعيين، عند 154، الموضحة أعلاه نسبة إلى الشكل 2. مثلا، عند 204، يمكن تعريف واحد أو أكثر من المواقع الجينية لقراءة أولى أو زوج قراءة

بشكل مؤقت. يمكن مقارنة القراءة الأولى لكل زوج قراءة بمتواليات اختيار لقاعدة بيانات (مثلا، انظر الجدول). يمكن أن تناظر كل من متواليات الاختيار لقاعدة البيانات موقع جيني معين لمجموعة المواقع الجينية. إذا كانت متوالية قراءة القراءة الأولى تتوافق بفاعلية مع واحد أو أكثر من متواليات الاختيار، بعد ذلك يمكن تسمية القراءة الأولى بشكل مؤقت للمواقع الجينية الذي يناظر متواليات الاختيار. مثلا، إذا كانت سلسلة النيكلوتيدات n (مثلا، نيكلوتيدات n الأولى) من متوالية تحديد للقراءة الأولى تتوافق بفاعلية مع واحد أو أكثر من متواليات الاختيار، يمكن بالتالي تسمية القراءة الأولى بشكل مؤقت لتلك المواقع الجينية المناظرة. يمكن الإشارة إلى الموقع الجيني المناظر (أو المواقع) بموقع (أو مواقع) معينة مؤقتة.

إذا كانت القراءة الأولى لا تتوافق بفاعلية مع أي من متواليات الاختيار، يمكن تجاهل القراءة غير المعينة. اختياريا، يمكن جمع أو تراكم القراءة غير المعينة، والتي يمكن أن تكون هي القراءة الأولى و/أو القراءة الثانية المناظرة، مع قراءات أخرى غير معينة. عند 206، يمكن تحليل القراءات غير المعينة للتحكم في الجودة. مثلا، يمكن تحليل متواليات القراءة للقراءة الأولى لتحديد سبب عدم تعيين القراءة الأولى.

يمكن أن تتضمن الطريقة 200 أيضا تحديد، عند 208، لكل للقراءات الأولية التي لها موقع جيني محتمل، ما إذا كانت القراءة الأولى تتحاذى مع واحدة أو أكثر من المتواليات المرجعية للمواقع الجينية المحتملة. يمكن أن يكون التحديد، 208، باستخدام واحد أو أكثر من بروتوكولات محاذاة. مثلا، يمكن أن يتضمن التحديد، عند 208، محاذاة القراءات الأولية مع متواليات مرجعية مناظرة للمواقع الجينية المحتملة كما هو موضح أعلاه نسبة إلى الأشكال 3-7. إذا كانت القراءة الأولى تتحاذى مع المتوالية المرجعية فقط لواحد من المواقع المحتملة، بعد ذلك يمكن توفير القراءة الأولى المؤقتة المعينة كقراءة صالحة للموقع الجيني الواحد ويمكن أن تستمر الطريقة إلى الخطوة 210. في النماذج الأخرى، يمكن تعيين القراءة الأولى المعين كقراءة صالحة للموقع الجيني الواحد ويمكن أن تنتقل الطريقة إلى الخطوة 212.

مع ذلك، إذا كانت القراءة الأولى بفاعلية تتحاذى مع أكثر من متوالية مرجعية واحدة، يمكن أن يتضمن التحديد، عند 208، تحديد المتوالية المرجعية حيث تتحاذى القراءة الأولى بشكل أفضل. بشكل أكثر تحديدا، على الرغم من أن القراءة الأولى يمكن أن تتحاذى بفاعلية مع العديد من المتواليات المرجعية، يمكن أن تكون إحدى عمليات المحاذاة أفضل من غيرها. كمثال بسيط، يمكن

أن يقوم تحليل المحاذاة بتحليل القراءة الأولى ومحاذاة القراءة الأولى إلى ثلاثة متواليات مرجعية، Ref Seq A، Ref Seq B، و Ref Seq C، والتي هي عبارة عن المتواليات المرجعية المرتبطة بثلاثة مواقع جينية محتملة محددة عند 204. يمكن أن تقوم محاذاة التحليل بتحديد أن القراءة الأولى بفاعلية تتحاذى مع Ref Seq A مع إجمالي ثلاثة فروق بين Ref Seq A والقراءة الأولى. يمكن أن تقوم محاذاة التحليل بتحديد أن القراءة الأولى بفاعلية تتحاذى مع Ref Seq B مع إجمالي أربعة فروق بين Ref Seq B والقراءة الأولى. يمكن أن تقوم محاذاة التحليل بتحديد أن القراءة الأولى و Ref Seq C لا تتحاذى مع بعضها البعض. مثلا، يمكن أن يكون هناك عدد متتابع للفروق (مثلا، أكبر من 10) بين القراءة الأولى و Ref Seq C. كمثال آخر، يمكن أن يكون هناك تناسب مفرط أو نسبة فروق (مثلا، يتناسب عدد من هذه الفروق مع العدد الكلي لنيكلوتيدات في القراءة أو في المتوالية المرجعية) موجودة بين القراءة الأولى و Ref Seq C. بناء على هذه البيانات، يمكن أن تحدد الطريقة أن القراءة الأولى تتحاذى بشكل أفضل مع Ref Seq A مما مع Ref Seq B. بالتالي، يمكن أن تكون القراءة الأولى معينة مؤقتا كقراءة صالحة للموقع الجيني التي تناظر Ref Seq A.

في بعض النماذج، تحديد أي متوالية مرجعية تتحاذى بشكل أفضل مع القراءة الأولى والتي يمكن أن تتضمن حساب درجات المحاذاة لكل من المتواليات المرجعية، حيث درجات المحاذاة قائمة على أساس عدد الفروق. كما هو موضح أعلاه، يمكن أن تكون درجة المحاذاة عبارة عن عدد خام (مثلا، عدد الفروق). في النماذج الأخرى، يمكن أن تكون درجة المحاذاة عبارة عن وظيفة لعدد و/أو أنواع الفروق. مثلا، يمكن أن تتفاوت الإندلات وحالات عدم التوافق بشكل مختلف.

اختياريا، يمكن أن تتضمن الطريقة 200 تحليل، عند 210، القراءة الثانية لتأكيد أن القراءة الأولى يجب تسميتها للموقع الجيني المعين بشكل مؤقت. يمكن تحليل القراءة الثانية بطريقة مشابهة مثل القراءة الأولى لزوج القراءة المناظر. يمكن تحليل القراءة الثانية لتحديد إذا كانت متوالية تحديد القراءة الثانية تتوافق بفاعلية مع واحد أو أكثر من متواليات اختيار لقاعدة بيانات. إذا كانت متوالية التحديد للقراءة الثانية تتوافق بفاعلية مع متوالية اختيار واحدة فقط، يمكن أن تتضمن الطريقة تحديد الموقع الجيني التي تناظر متوالية الاختيار الواحدة. إذا كان الموقع الجيني هو نفس الموقع الجيني حيث تم تعيين القراءة الأولى بشكل مؤقت، بعد ذلك يمكن الإشارة إلى الموقع الجيني بالموقع الجيني للقراءة الأولى ويمكن تعيين القراءة الأولى، عند 212، على أنها قراءة صالحة للموقع الجيني.

مع ذلك، إذا كانت متوالية التحديد للقراءة الثانية تتوافق بفاعلية مع العديد من متواليات الاختيار، يمكن أن تتضمن الطريقة تحديد المواقع الجينية التي تناظر متواليات الاختيار المتعددة. إذا كان واحد من هذه المواقع الجينية هو نفس الموقع الجيني الذي تم تعيين القراءة الأولى بشكل مؤقت له، فإنه يمكن الإشارة إلى الموقع الجيني بالموقع الجيني للقراءة الأولى ويمكن تعيين القراءة الأولى، عند 212، على أنها قراءة صالحة للموقع الجيني.

إذا كان التحليل، عند 210، لا يؤكد على أن القراءة الثانية تناظر الموقع المعين بشكل مؤقت للقراءة الأولى، فإن الطريقة 200 يمكن أن تتضمن تعيين القراءة الأولى المناظرة كقراءة غير مؤكدة. يمكن جمع القراءات غير المؤكدة و، اختياريا، تحليل، عند 214، للتحكم في الجودة. مثلا، أزواج قراءة متوافقة بفاعلية مع متوالية اختيار أولى للموقع المعين بشكل مؤقت، ولكنها لا تتوافق بفاعلية مع متوالية اختيار ثانية للموقع المعين بشكل مؤقت ويمكن أن تشير إلى اعتبارات في الفحص. يمكن أن تشير القراءات غير المؤكدة إلى واحد أو أكثر من نواتج التكبير البعيدة عن الهدف. يمكن تحليل أزواج القراءة، عند 214، لتحديد، مثلا، ما إذا كانت مشكلة التحكم في الجودة موجودة في الفحص أو تشير إلى إسقاط أليلة.

مع ذلك، إذا كانت القراءة الأولى لا تتحاذى مع متوالية مرجعية لموقع جيني محتمل عند 208، بعد ذلك يمكن أن تتضمن الطريقة تعيين، عند 216، القراءة الأولى كموجه غير متحاذي. يمكن أن تمثل القراءات غير المتحاذية قراءات أولى مرت بمرحلة ترشيح واحدة، ولكنها لم تكن قادرة على المحاذاة مع متوالية مرجعية. تحديدا، يمكن أن تكون القراءات غير المتحاذية عبارة عن قراءات أولى تم تأكيدها على أن لها متوالية تحديد تتوافق بفاعلية مع واحد أو أكثر من متواليات الاختيار، ولكنها لم تكن قادرة على المحاذاة مع متوالية مرجعية.

اختياريا، يمكن أن تتضمن الطريقة 200 تحليل، عند 218، كل من القراءات غير المتحاذية لتحديد موقع جيني ذو ملائمة أفضل للقراءة غير المتحاذية المناظرة. كما هو موضح أعلاه، يمكن أن تتوافق متوالية التحديد بفاعلية مع أكثر من متوالية اختيار واحدة. يمكن أن يتضمن التحليل، عند 218، مقارنة متوالية التحديد للقراءة غير المتحاذية مع متواليات الاختيار التي تم تحديدها سلفا عند 204. يمكن أن يكون الموقع الجيني ذو الملائمة الأفضل عبارة عن الموقع الجيني الذي يناظر متوالية الاختيار التي تتوافق بشكل أفضل مع أو تتوافق بشكل أفضل مع متوالية التحديد للقراءة غير المتحاذية. بالتالي، عند 218، يمكن أن تقوم الطريقة بتحديد أي متوالية اختيار، من

بين العديد من متواليات الاختيار، تتوافق بشكل أفضل مع متوالية التحديد. مثلا، يمكن أن يكون الموقع الجيني ذو الملائمة الأفضل عبارة عن الموقع الجيني الذي يناظر متوالية الاختيار التي لها أقل فروق مع متوالية التحديد. في بعض النماذج، يمكن أن يتضمن التحليل، عند 218، تحديد درجة التوافق لكل من متواليات الاختيار نسبة إلى متوالية التحديد. يمكن تعيين الموقع الجيني الذي يناظر متوالية الاختيار مع أكبر درجة توافق معينة مثل الموقع الجيني ذو الملائمة الأفضل.

عند 220، يمكن تحليل القراءة الثانية المرتبطة بالقراءة غير المتحاذية (أي، القراءة الأولى) لتحديد إذا كانت القراءة الثانية تؤكد أفضل ملائمة للموقع المحدد عند 218. يمكن تحليل القراءة الثانية لتحديد إذا كانت متوالية التحديد للقراءة الثانية تتوافق بفاعلية مع واحد أو أكثر من متواليات الاختيار. إذا كانت متوالية التحديد للقراءة الثانية تتوافق بفاعلية مع متوالية اختيار وأن متوالية الاختيار تناظر الموقع الجيني ذو الملائمة الأفضل، فإنه يمكن تعيين القراءة غير المتحاذية، عند 222، كقراءة اثنين على الهدف غير متحاذية (المشار إليها أيضا بزوج على هدف قراءة غير متحاذية). يمكن أن تمثل قراءة اثنين على الهدف غير متحاذية قراءة غير متحاذية لها متواليات قريبة من كلا طرفي القراءة غير المتحاذية التي تتوافق بفاعلية مع متواليات اختيار من قاعدة البيانات. على الرغم من التوافق بفاعلية مع اثنين من متواليات الاختيار، لم تكن المنطقة محل الاهتمام للقراءة غير المتحاذية قادرة على المحاذاة مع متوالية مرجعية.

مع ذلك، إذا كانت متوالية التحديد للقراءة الثانية لا تتوافق بفاعلية مع متوالية اختيار تناظر الموقع الجيني ذو الملائمة الأفضل، فإنه يمكن تعيين القراءة غير المتحاذية، عند 224، كقراءة واحد على هدف غير متحاذية. يمكن أن تمثل قراءة واحد على هدف غير متحاذية قراءة غير متحاذية لها متوالية تحديد واحدة فقط تتوافق بفاعلية مع متوالية اختيار من قاعدة البيانات.

يمكن تحليل كل من قراءات اثنين على هدف غير متحاذية وقراءات الواحد على هدف غير متحاذية، عند 226 و228، على التوالي، لأغراض التحكم في الجودة. يمكن أن يتضمن التحليل، عند 226 أو 228، تحليل العدد الكلي لقراءات غير متحاذية (أو درجة مقارنة) و/أو تحليل متواليات الـ منطقة محل الاهتمام للقراءات غير المتحاذية. مثلا، يمكن تحليل قراءات الواحد على هدف غير متحاذية، عند 228، لتحديد صحة الفحص. بشكل أكثر تحديدا، يمكن تحليل قراءات الواحد على هدف غير متحاذية لتحديد إذا كانت الخيمرات موجودة و/أو إذا كانت الدايمرات لأولية موجودة. يمكن أن يقوم عدد متتابع لخيمر و/أو دايمرات أولية بإرسال الإشارة إلى أن الفحص

ضعيف (مثلا، مشكلة التكبير) أو أن عينة الحمض النووي الريبي منقوص الأكسجين منخفض الجودة. اختياريا، يمكن أن يتضمن التحليل، عند 228، تحليل القراءات غير المؤكدة لـ 214 لتحديد صحة الفحص. يمكن أن يتضمن التحليل، عند 228، جمعا تحليل القراءات غير المؤكدة وقراءات الواحد على هدف غير متحاذية. بشكل بديل، يمكن أن يتضمن التحليل، عند 228، بشكل منفصل تحليل القراءات غير المؤكدة وقراءات الواحد على هدف غير متحاذية.

نسبة إلى قراءات اثنين على هدف غير متحاذية، يمكن أن يشير عدد متتابع لهذه القراءات إلى إسقاط أليلة محتمل. في بعض النماذج، يمكن أن يتضمن التحليل، عند 226، تحديد إذا كان عدد قراءات اثنين على هدف غير متحاذية يتخطى نسبة قراءات كلية للموقع المعين، بعد ذلك قد يتم تحديد أن هناك مشكلة موجودة مع الموقع المعين. يمكن أن تكون "القراءات الكلية" للموقع المعين عبارة عن وظيفة للقراءات الصالحة، المعينة عند 212، والقراءات غير المتحاذية، المعينة عند 216. مثلا، يمكن أن تكون القراءات الكلية مساوية لإجمالي القراءات الصالحة والقراءات غير المتحاذية. في النماذج الأخرى، يمكن أن تكون القراءات الكلية أيضا وظيفة للقراءات غير المؤكدة. عند 226، يمكن مقارنة عدد قراءات اثنين على هدف غير متحاذية (أو درجة مقارنة) بقيمة حدية لتحديد ما إذا كانت هناك مشكلة (مثلا، إسقاط أليلة) موجودة مع الموقع المعين.

عند 230، يمكن توفير إخطار يتعلق بجودة الفحص و/أو الثقة في الهيئة الجينية. مثلا، يمكن أن يقوم الإخطار بإخطار مستخدم بعدد القراءات غير المتحاذية. في نماذج معينة، يمكن أن يقوم الإخطار بإعلام المستخدم عدد قراءات واحد على هدف غير متحاذية و/أو عدد قراءات اثنين على هدف غير متحاذية. في بعض الحالات، يمكن أن تقوم الطريقة بمقارنة عدد قراءات غير متحاذية (أو درجة مقارنة)، عدد قراءات واحد على هدف غير متحاذية (أو درجة مقارنة)، و/أو عدد قراءات اثنين على هدف غير متحاذية (أو درجة مقارنة) بقيم حدية معينة. إذا كانت الأعداد أو الدرجات تتخطى القيم الحدية، يمكن أن يتضمن الإخطار تحذير معين أو توجيه معين للمستخدم. مثلا، يمكن أن يقوم الإخطار بإعلام المستخدم يشير هذا الدليل إلى أن العينة ذات جودة منخفضة و/أو كانت منخفضة في الكمية. يمكن توجيه الإخطار إلى الفحص ككل أو يمكن أن يكون انتقائي لمواقع معينة. بشكل أكثر تحديدا، يمكن أن يشير عدد متتابع لقراءات واحد على هدف غير متحاذية إلى مشكلة مع الفحص، حيث يمكن أن يشير العدد المتتابع من القراءات اثنين على هدف غير متحاذية إلى إسقاط أليلة.

عند 232، يمكن تصنيف القراءات الصالحة لتكوين توزيع قراءة للموقع المعين. يتضمن توزيع القراءة بشكل مثالي العديد من قراءات عينة التي تم تمريرها خلال العديد من مراحل الترشيح وتم تعيينها إلى الموقع المعين. مثلا، يمكن أن يتضمن توزيع القراءة العشرات أو المئات أو الآلاف من القراءات الأولى التي تم تعيينها إلى الموقع المعين. يمكن جمع توزيع القراءة في ملف (مثلا، "ملف توزيع distribution file") ويتضمن معلومات تتعلق بتوزيع قراءات العينة، مثل ألائل محتملة مختلفة، متواليات الألائل، ودرجة عد (مثلا، عدد قراءة أو قيمة/وظيفة أخرى بناء على عدد القراءة) لكل أليلة محتملة. مثلا، عند تصنيف القراءات الصالحة لتوزيع القراءة، يمكن تصنيف القراءات الصالحة بناء على متوالية. يمكن أن تتضمن القراءات الصالحة عدد متواليات مختلفة والتي، على الرغم من اختلافها، قد تم تعيينها إلى الموقع المعين. تمثل كل متوالية أليلة مختلفة محتملة للموقع المعين. يمكن أن تكون واحدة أو أكثر من المتواليات ضوضائية (مثلا، خطأ توالي)، يمكن أن تكون واحدة أو أكثر من المتواليات عبارة عن منتج مضطرب، ويمكن أن تكون واحدة أو أكثر من المتواليات عبارة عن ألائل حقيقية.

يمكن مراكمة القراءات الصالحة مع قراءات صالحة أخرى تتضمن نفس المتوالية. يمكن عد عدد القراءات الصالحة لها نفس المتوالية للمتوالية المعينة. مثلا، بافتراض موقع جيني له 1000 قراءة صالحة معينة له، يمكن أن يشير توزيع القراءة إلى أن ثمانية متواليات مختلفة موجودة. يمكن توزيع القراءات الصالحة بين المتواليات الثمانية المختلفة. مثلا، أليلة 1 يمكن أن تتضمن 10 قراءة صالحة؛ أليلة 2 يمكن أن تتضمن 20 قراءة صالحة؛ أليلة 3 يمكن أن تتضمن 10 قراءة صالحة؛ أليلة 4 يمكن أن تتضمن 400 قراءة صالحة؛ أليلة 5 يمكن أن تتضمن 15 القراءات الصالحة؛ أليلة 6 يمكن أن تتضمن 500 قراءة صالحة؛ أليلة 7 يمكن أن تتضمن 25 القراءات الصالحة؛ وأليلة 8 يمكن أن تتضمن 20 قراءة صالحة. يمكن أن يحدد المزيد من التحليل، كما هو موضح أدناه، أن بعض الألائل عبارة عن ضوضاء و/أو منتج مضطرب.

في بعض النماذج، يمكن أن تقوم الألائل المحتملة بتوفير اسم أليلة الاستشراد الكهربي بناء على ممارسات التسمية التقليدية في الاستشراد الكهربي. يمكن أن يقوم اسم أليلة الاستشراد الكهربي لأليلة محتملة، جزئيا، على عدد بواعث تكرار في المتوالية. يمكن أيضا أن تكون تسمية أليلة الاستشراد الكهربي عبارة عن استخدام قائم على التاريخ. في بعض النماذج، يتم ترتيب الألائل المحتملة في توزيع القراءة بناء على اسم أليلة الاستشراد الكهربي. مثلا، تتضمن أسماء أليلة

الاستشراد الكهربي بشكل مثالي قيمة عددية. يمكن ترتيب الألائل المحتملة بناء على القيم الرقمية. على سبيل المثال، يشرح الرسم البياني 192 المبين في الشكل 10 يشرح توزيع قراءة واحدة. كما هو مبين، تتضمن الألائل المحتملة 11، 11.2، 12، 13، و14. يمكن ترتيب توزيع القراءة للموقع الجيني الموضح في رسم بياني 192 11، 11.2، 12، و13.

في بعض الظروف، يمكن أن تتضمن الألائل المحتملة المختلفة نفس اسم أليلة الاستشراد الكهربي. مثلا، بناء على ممارسات التسمية التقليدية، يمكن إعطاء الألائل المحتملة نفس اسم أليلة الاستشراد الكهربي. في بعض النماذج، يمكن أن يشير توزيع القراءة إلى أن المتواليتين مختلفة لهما نفس اسم أليلة الاستشراد الكهربي. مثلا، يمكن أن يشير توزيع القراءة إلى اسم أليلة الاستشراد الكهربي (مثلا، 13) وبعد ذلك يتم ذكر قائمة المتواليات المختلفة التي تناظر نفس اسم أليلة الاستشراد الكهربي.

بعد تصنيف القراءة لتكوين توزيعات قراءة، يمكن بعد ذلك توصيل توزيعات القراءة إلى عمليات التحليل المختلفة. مثلا، يمكن توجيه المواقع الجينية المعروفة بأنها تتضمن تعددات شكلية لنيكلوتيد فردي يمكن توجيهها خلال تحليل تعدد شكلي لنيكلوتيد فردي. يمكن توجيه المواقع الجينية المعروفة لـ تكرارات ترادفية قصيرة خلال تحليل تكرار ترادفي قصير. على الرغم من أن تحليل تعدد شكلي لنيكلوتيد فردي وتكرار ترادفي قصير يمكن أن يتضمن خطوات مختلفة، يمكن أن يتضمن التحليل أيضا خطوات مشابهة.

يشرح الشكل 12 طريقة 240 لتحليل بيانات توالي وفقا لأحد النماذج. تحديدا، الطريقة 240 تتضمن تحليل توزيعات قراءة لمواقع معينة. يمكن أن تكون توزيعات القراءة عبارة عن مواقع تكرار ترادفي قصير، تعدد شكلي لنيكلوتيد فردي مواقع، أو مواقع أخرى مرتبطة باختلافات متوالية. تتضمن الطريقة 240 تلقي، عند 242، توزيعات قراءة للمواقع المعينة. نسبة إلى الخطوات التالية، يمكن أن تكون كل الخطوة قائمة جزئيا على الأقل على الموقع المعين. مثلا، هناك العديد من الوظائف (مثلا، القيم الحدية thresholds) التي تنطبق حيث تقوم هذه الوظائف على أساس الموقع المعين. بشكل أكثر تحديدا، يمكن ألا تكون وظيفة أحد المواقع الجينية هي نفس وظيفة الموقع الجيني.

اختياريا، تتضمن الطريقة 240 تحديد، عند 244، درجة عد لكل للألائل المحتملة لموقع جيني معين. يمكن أن تكون درجة العد قائمة على عدد القراءة للأليلة المحتملة. يمثل عدد القراءة عدد

القراءات الصالحة التي تتضمن متوالية مشتركة. في بعض النماذج، درجة العد عبارة عن قيمة تساوي عدد القراءة للأليلة المحتملة. مثلا، إذا كان عدد القراءة 300، فإنه درجة العد يمكن أن تكون 300. في النماذج الأخرى، يمكن أن تكون درجة العد لأليلة محتملة قائمة على عدد القراءة والعدد الكلي لقراءات للموقع الجيني. يمكن أن يكون العدد الكلي للقراءات، مثلا، عبارة عن العدد الكلي لقراءات في توزيع القراءة لكل الألائل المحتملة. في بعض النماذج، يمكن أن تكون درجة العد للأليلة المحتملة قائما على عدد القراءة وبيانات تم الحصول عليها سلفا للموقع الجيني. في نماذج معينة، يمكن أن تكون درجات العد عبارة عن درجات معايرة بين أعداد محددة سلفا (مثلا، 0 و1). يمكن أن تكون درجات المعايرة قائمة على العدد الكلي لقراءات للموقع الجيني. اختياريا، درجات المعايرة عبارة عن وظيفة لأعداد القراءة من مواقع أخرى و/أو أعداد القراءة من عينات أخرى. يمكن أن تكون درجة العد عبارة عن وظيفة لأعداد القراءة من مواقع أخرى للعينة أو وظيفة لأعداد القراءة من عينات أخرى التي تمت تشغيلها بشكل متزامن مع العينة محل الاهتمام. يمكن أن تكون درجة العد أيضا عبارة عن وظيفة للبيانات التاريخية. مثلا، يمكن تشغيل أنواع مختلفة من الفحوص للحصول على أعداد قراءة. في بعض النماذج، تقوم درجة العد قائم على البيانات التاريخية التي تتعلق بفحص معين.

يمكن أن تتضمن الطريقة 240 أيضا تحديد، عند 245، ما إذا كانت واحد أو أكثر من درجات العد للألائل المحتملة يتخطى قيمة تفسير حدية. يمكن أن تكون قيمة التفسير الحدية عبارة عن قيمة محددة سلفا أو يمكن أن تكون عبارة عن وظيفة قائمة على مجموعة من العوامل. مثلا، يمكن أن تكون قيمة التفسير الحدية قائمة على عدد القراءات الكلية التي تناظر الموقع المعين. يمكن أن يتضمن عدد القراءات الكلية القراءات الصالحة لكل الألائل المحتملة في موقع. في بعض النماذج، يمكن أنيتضمن عدد القراءات الكلية القراءات الصالحة للموقع والقراءات غير المتحاذية للموقع. في نماذج معينة، يمكن أن يتضمن عدد القراءات الكلية القراءات الصالحة، القراءات غير المتحاذية، والقراءات غير المؤكدة للموقع. إذا كان واحد من درجات العد يتخطى قيمة التفسير الحدية عند 245، بعد ذلك يمكن أن تنتقل الطريقة 240 إلى الخطوة 246 أو خطوة أخرى تالية. في بعض النماذج، يمكن أن تكون قيمة التفسير الحدية قائمة على العدد الكلي لقراءات في عينة. في بعض النماذج، يمكن أن تكون قيمة التفسير الحدية قائمة على العدد الكلي لقراءات في مجموعة من عينات.

إذا لم يكن أي من درجات العد يتخطى قيمة التفسير الحدية، عند 245، حينئذ يمكن أن تقدم الطريقة 240، عند 248، تنبيه أو إخطار آخر يتعلق بالموقع المعين. مثلا، يمكن أن يقوم التنبيه بإخطار مستخدم أن الموقع المعين له تغطية منخفضة. بشكل أكثر تحديدا، يمكن أن يقوم التنبيه بإخطار المستخدم بأن كمية البيانات التي تتعلق بالموقع المعين يمكن أن تكون غير كافية لتوفير تسمية نمط جيني.

في نموذج معين، تتضمن الطريقة 240 تحديد الأليلة المحتملة التي لها عدد قراءة أقصى (أو عد الأليلة) في توزيع القراءة. يمثل عدد القراءة عدد القراءات الصالحة التي تتضمن متوالية مشتركة. نسبة إلى تكرارات ترادفية قصيرة، يمكن أن يمثل عدد القراءة عدد القراءات الصالحة التي تتضمن متوالية مشتركة من المنطقة محل الاهتمام أو القطاع التكراري. يمكن أن تتضمن الطريقة 240 أيضا مقارنة عدد القراءة الأقصى بقيمة التفسير الحدية. إذا كان عدد القراءة الأقصى يتخطى قيمة التفسير الحدية، عند 245، بعد ذلك يمكن أن تنتقل الطريقة 240 إلى الخطوة 246 أو خطوة أخرى تالية. إذا كان أقصى عدد أليلة يتخطى قيمة التفسير الحدية، يمكن بعد ذلك أن تقدم الطريقة 240، عند 248، تنبيه أو إخطار آخر يتعلق بالموقع المعين كما هو موضح أعلاه.

في النماذج الأخرى، يمكن مقارنة درجة العد بقيمة حدية أخرى، مثل القيمة الحدية التحليلية الموضحة أدناه. تكون القيمة الحدية التحليلية أسهل بشكل مثالي في التخطي من قيمة التفسير الحدية. إذا لم يكن أحد الألائل المحتملة له درجة عد تتخطى القيمة الحدية التحليلية، فإنه قد يتم تحديد أن الموقع الجيني له تغطية منخفضة. كمثال آخر لتحديد ما إذا كان الموقع الجيني تتضمن تغطية كافية، يمكن مقارنة العدد الكلي لقراءات الموقع الجيني (مثلا، القراءات الصالحة) بقراءة قيمة حدية. يمكن أن تكون قراءة القيمة الحدية قائمة على العدد الكلي لقراءات في العينة و/أو البيانات التاريخية. إذا كانت العدد الكلي لقراءات للموقع الجيني لا تتخطى قراءة القيمة الحدية، بعد ذلك قد يتم تحديد أن الموقع الجيني له تغطية منخفضة. في النماذج الأخرى، يمكن استخدام توليفة من واحد أو أكثر من الخطوات، مثل تلك الموضحة أعلاه، لتحديد ما إذا كان الموقع الجيني له تغطية منخفضة.

اختياريا، عند 246، يمكن مقارنة كل من درجات العد أو أعداد القراءة المناظرة في توزيع القراءة بقيمة حدية تحليلية analytical threshold. يمكن أن تكون قيمة التفسير الحدية، القيمة الحدية التحليلية هي نفس القيمة المحددة سلفا أو وظيفة قائمة على مجموعة من عوامل، مثل العدد الكلي

للقراءات (مثلا، العدد الكلي للقراءات الصالحة) للموقع و/أو المعرفة التاريخية للموقع المعين. يمكن أن تكون القيمة الحدية التحليلية أقل قوة (مثلا، أسهل في تخطي) من قيمة التفسير الحدية. بشكل أكثر تحديدا، يمكن أن تتطلب قيمة التفسير الحدية عدد قراءة أكبر لتخطي القيمة الحدية التحليلية.

بعدد تخطي القيمة الحدية التحليلية عند 246، يمكن أن تتضمن الطريقة 240 تحديد، عند 247، ما إذا كانت الأليلة المحتملة عبارة عن منتج مضطرب مشكوك فيه. يمكن استخدام العديد من القواعد أو الظروف لتحديد ما إذا كان الأليلة المحتملة عبارة عن منتج مضطرب مشكوك فيه. مثلا، يمكن تطبيق واحد أو أكثر من العوامل 171-175 الموضحة أعلاه نسبة إلى الشكل 8. في نماذج معينة، يتضمن التحديد، عند 247، تحديد ما إذا كانت الأليلة الأولى لها باعث تكرار تمت إضافتها أو إسقاطها نسبة إلى أليلة ثانية.

إذا كانت الأليلة المحتملة غير المشكوك فيها عبارة عن جزء منتج مضطرب، يتم تعيين الأليلة المحتملة، عند 250، على أنها أليلة معينة أو مسماة للموقع. إذا كان مشكوكا في أن الأليلة المحتملة عبارة عن جزء منتج مضطرب، تتضمن الطريقة 240 تحديد، عند 249، ما إذا كانت درجة العد للأليلة الأولى أقل من قيمة حدية معينة. يمكن أن تكون درجة العد عبارة عن عدد القراءة أو وظيفة بناء على عدد القراءة. يمكن أن تقوم القيمة الحدية المعينة على درجة العد للأليلة الثانية. في نماذج معينة، يمكن أن يتضمن التحديد، عند 249، تحديد ما إذا كان درجة العد للأليلة الأولى في نطاق محدد سلفا (مثلا، 10٪-30٪) لدرجة العد للأليلة الثانية.

على الرغم من عدم الإشارة إليه في الشكل 12، إذا كانت الأليلة المحتملة أقل من القيمة الحدية المعينة أو في النطاق المحدد سلفا، يمكن أن تكون الأليلة المحتملة معينة كمنتج مضطرب للأليلة الثانية. يمكن أن تتم ملاحظة المنتج المضطرب مع تسمية النمط الجيني للموقع. مثلا، يمكن أن يتضمن تقرير العينة النمط الجيني للموقع، مع إشارة إلى أن المنتج المضطرب موجود. يمكن تقديم معلومات تتعلق بالمنتج المضطرب (مثلا، متوالية ونسبة الأليلة الثانية) في تقرير العينة. مع ذلك، إذا كانت درجة العد أو عدد القراءة يتخطى قيمة حدية معينة (أو في نطاق محدد سلفا)، بعد ذلك يمكن تعيين الأليلة المحتملة، عند 250، كأليلة معينة للموقع الجيني.

في بعض النماذج، يتم جمع درجات العد للألائل الضوضائية noise alleles، عند 252. يمكن أن تتضمن الألائل الضوضائية الألائل المحتملة التي لم تتخطى القيمة الحدية التحليلية، عند 246،.

في بعض النماذج، يمكن أن تتضمن الألائل الضوضائية أيضا درجات عد من القراءات غير المتحاذية، اختياريا، والقراءات غير المؤكدة الموضحة أعلاه. يمكن جمع درجات العد للألائل الضوضائية، عند 252، ويتم تحليلها، عند 254، لتحديد إذا كان العدد المتتابع للقراءات يشير إلى مشكلة محتملة مع الموقع المناظر. مثلا، يمكن تلخيص درجات العد لكل الألائل الضوضائية ومقارنتها بقيمة محددة سلفا حدية ضوضائية. يمكن أن تقوم القيمة الحدية الضوضائية على العدد الكلي لقراءات و/أو البيانات التاريخية. إذا تم تخطي القيمة الحدية الضوضائية، عند 254، يمكن توفير التنبيه، عند 256، التي لها موقع له كمية مفرطة من الضوضاء.

في بعض النماذج، يمكن تحليل الألائل الضوضائية، عند 258، للتحكم في الجودة. في نماذج معينة، يمكن تقسيم الألائل الضوضائية ثانويا لموقع تكرار ترادفي قصير في ألائل ضوضائية لها متواليات لها نفس طول مثل الأليلة وألائل ضوضائية لها متواليات ليست هي نفس طول الألائل المسماة. يمكن أن يقوم فصل الألائل الضوضائية بتوفير المعلومات الإضافية فيما يتعلق بسبب وجود ضوضاء مفرطة مع الموقع المناظر.

بعد تحديد الألائل المعينة، عند 250، يمكن أن تتضمن الطريقة 240 تحليل الألائل المعينة أيضا قبل تسمية نمط جيني للموقع المعين. سوف تتضمن تسمية النمط الجيني بشكل مثالي تسمية غير متجانسة (أي، اثنين من الألائل المتخلفة) أو تسميات الجينات المتماثلة (أي، أليلة واحدة ملحوظة). للتسميات غير المتجانسة، سوف تقوم البيانات بشكل مثالي بدعم أن القراءات متناسبة بشكل متساوي. إذا كان الاثنين من الألائل غير موجودين بشكل متساوي أساسا في البيانات، يمكن وجود مشكلة مع الموقع. بالتالي، في بعض النماذج، يمكن أن تتضمن الطريقة 240 تحليل، عند 260، الألائل المسماة لتحديد إذا كانت الألائل المسماة متوازنة أو متناسبة. مثلا، يمكن حساب نسبة الألائل المسماة لتحديد إذا كانت النسبة تلبي قيمة حدية متوازنة. مثلا فقط، إذا كانت درجة العد (مثلا، عدد قراءة) لأليلة واحدة أقل من 50٪ أو أقل من 75٪ من درجة العد (مثلا، عدد قراءة) للأليلة الأخرى، يمكن أن يتم تعيين الألائل على أنها غير متناسبة. بالتالي، يمكن تقديم تنبيه تناسب الأليلة، عند 262، بما يشير إلى أن الألائل غير متوازنة. كما هو ملحوظ أدناه، يمكن تحليل تنبيه تناسب الأليلة بواسطة دليل آخر (مثلا، تنبيهات أخرى) لتحديد ما إذا كانت العينة تتضمن مجموعة من مصادر.

في بعض النماذج، يمكن أن تتضمن الطريقة 240 تحديد، عند 264، ما إذا كان عدد النسخ للموقع يتخطى القيمة الحدية للنسخ copy threshold. بالنسبة للمواقع الكروموسومية الذاتية autosomal loci، سوف يكون عدد النسخ بشكل مثالي اثنين على الأكثر. بالنسبة للمواقع غير الكروموسومية الذاتية non-autosomal loci، مثل مواقع X- أو مواقع Y-، يمكن أن تكون عدد النسخ مختلفا. مثلا، يمكن أن يكون عدد النسخ لموقع Y- واحد على الأكثر. يجب أن يكون عدد النسخ لموقع X- اثنين على الأكثر. كما هو موضح أدناه، في بعض الحالات، يمكن توقع نوع العينة وبعد ذلك يتم الاستخدام عند السؤال عما إذا كانت مجموعة من مصادر موجودة في العينة. بالتالي، يمكن أن يتضمن التحديد، عند 264، الحصول على عدد نسخ للموقع المعين (مثلا، 0، 1، أو 2) ومقارنة عدد ألائل الموقع المعين المسماة بعدد النسخ. إذا كان عدد الألائل المسماة يتخطى عدد النسخ، يمكن توفير تنبيه عدد الألائل، عند 266، حيث يتضمن الموقع عدد متتابع من الألائل. كما هو ملحوظ أدناه، يمكن تحليل تنبيه عدد الألائل مع دليل آخر (مثلا، تنبيهات أخرى) لتحديد ما إذا كانت العينة تتضمن مجموعة من مصادر.

عند 268، يمكن تسمية النمط الجيني للموقع. تقوم تسمية النمط الجيني على الألائل المعينة، عند 250، وسوف يكون بشكل مثالي عبارة عن واحد أو اثنين من الألائل. مع ذلك، في بعض النماذج، سوف تتضمن تسمية النمط الجيني أكثر من اثنين من الألائل. يمكن أن تتضمن تسميات النمط الجيني مع أكثر من اثنين من ألائل إخطارات التي يمكن أن تشير إلى مشكلة يمكن أن تكون موجودة عند الموقع أو مع العينة بشكل عام. عند 270، يمكن توليد تقرير عينة تتضمن تسمية نمط جيني، إن أمكن، بالنسبة للمواقع الجينية للمجموعة المحددة سلفا. يمكن أيضا أن يتضمن تقرير العينة عدد إخطارات (مثلا، تنبيهات) تتضمن تلك المحددة بواسطة الطريقة 240 أو الطريقة 200 (الشكل 11). في بعض النماذج، يمكن توفير تسمية نمط جيني لموقع بالإضافة إلى مؤشر يقوم بإخطار القارئ بمشكلة محتملة (مثلا، التغطية، ضوضاء، إسقاط الأليلة، الاضطراب، إلخ.) تتعلق بالموقع. في النماذج الأخرى، لم يتم توفير تسمية نمط جيني لموقع جيني إذا تم تحديد تنبيهات معينة للموقع الجيني (مثلا، التغطية أو الضوضاء). في بعض النماذج، يمكن أن يتضمن تقرير المتواليات عينة الألائل المسماة، اختياريا، ومتواليات المنتجات المضطربة و/أو ألائل أخرى محتملة محددة. في بعض النماذج، يمكن أن يتضمن تقرير العينة درجة ثقة نسبة إلى العينة ككل.

مثلا، إذا كان عدد كبير من قراءات واحد على هدف غير متحاذية موجودة، يمكن أن يشير تقرير العينة إلى أن العينة يمكن أن تكون ذات جودة منخفضة.

الشكل 13 عبارة عن مخطط تدفق يشرح طريقة 300 لنوع المصدر المتوقع للعينة. تفترض الطريقة 300 أن العينة من مصدر فردي. إذا تم تحديد أن العينة من مصادر متعددة، كما هو موضح أدناه، حينها يمكن إزالة توقع النوع. في بعض النماذج، بعد تحديد أن العينة تتضمن العديد من المصادر، يمكن أن تتوقع الطريقة أن كل مصادر العينة من نوع واحد، مثل ذكر.

يمكن أن يتم دمج الطريقة 300 مع الطريقة 240 (الشكل 12). يمكن تنفيذ الطريقة 300 بعد تحديد الألائل المعينة لكل موقع جيني من مجموعة من مواقع جينية. مثلا، يمكن تنفيذ الطريقة 300 بعد حدوث الخطوة 250 في الشكل 12 لكل الألائل المحتملة لمجموعة من المواقع الجينية في مجموعة من مواقع جينية (أو لكل المواقع الجينية في المجموعة). تتضمن الطريقة 300 تلقي، عند 302، بيانات موقع لمجموعة من مواقع جينية. يمكن أن تتضمن بيانات الموقع واحد أو أكثر من ألائل معينة (أو مسماة) للمواقع الجينية المناظرة. يمكن أن تكون مجموعة المواقع الجينية عبارة عن مواقع متوقع أن تتضمن أعداد مختلفة من الألائل بناء على نوع العينة. بمعنى آخر، يمكن أن تناظر بيانات الموقع المواقع X- وY-. يمكن أن تتضمن المواقع X- تعدد شكلي لنيكلوتيد فردي معروفة أو مواقع تكرار ترادفي قصير على الكروموسوم -X. يمكن أن تتضمن المواقع Y- تعدد شكلي لنيكلوتيد فردي معروفة أو مواقع تكرار ترادفي قصير على الكروموسوم -Y.

يمكن أن تتضمن الطريقة 300 مقارنة، عند 304، عدد ألائل معينة لكل موقع Y- بعدد غير متوقع إذا كانت العينة عبارة عن ذكر و/أو بعدد غير متوقع إذا كانت العينة عبارة عن أنثى. يمكن أن يكون العدد المتوقع عبارة عن عدد محدد سلفا بناء على البيانات التاريخية. يمكن أن يكون العدد المتوقع للألائل المعينة لعينة ذكر بناء على مرات عد الموقع أو الألائل الظاهرة على الكروموسوم -Y. على الرغم من أن هذا واحد بشكل مثالي، يمكن أن يكون أكثر من واحد (مثلا، اثنين). العدد المتوقع للألائل المعينة لعينة أنثى في موقع Y- هو صفر.

اختياريا، يمكن أن تتضمن الطريقة 300 مقارنة، عند 306، عدد ألائل معينة لكل موقع X- بعدد غير متوقع إذا كانت العينة عبارة عن ذكر و/أو بعدد غير متوقع إذا كانت العينة عبارة عن أنثى. العدد المتوقع للألائل المعينة لعينة ذكر في موقع X- بشكل مثالي عبارة عن واحد، ولكن يمكن أن يكون أكثر من واحد إذا كان الموقع أو الأليلة يظهر أكثر من مرة على الكروموسوم X-. يبلغ

العدد المتوقع للألائل المعينة لعينة أنثى في موقع X- بشكل مثالي اثنين، ولكن يمكن أن تكون أكثر إذا كان الموقع/الأليلة يظهر أكثر من مرة على الكروموسوم X-.

تتضمن الطريقة 300 أيضا توقع، عند 308، نوع العينة بناء على النتائج من المقارنة، عند 304، و/أو النتائج من المقارنة، عند 306. بشكل مثالي، سوف تتضمن كل من المواقع Y- أليلة معينة واحدة إذا كانت العينة ذكر وسوف تتضمن ألائل معينة صفرية إذا كانت العينة أنثى. بالمثل، سوف تتضمن كل من المواقع X- بشكل مثالي أليلة واحدة معينة إذا كانت العينة ذكر وسوف تتضمن واحد أو اثنين من الألائل المعينة إذا كانت العينة أنثى. مع ذلك، بسبب خطأ التوالي، التلوث، التحليل غير المناسب، إلخ.، قد يكون من الممكن أن المواقع X- والمواقع Y- لن تكون متوافقة في توقع نوع العينة. في بعض الحالات، يمكن أن يقوم التحليل باعتبار العديد من المواقع الجينية. مثلا، يمكن أن يكون هناك حوالي خمسة (5) إلى عشرة (10) مواقع Y- وحوالي عشرين (20) إلى ثلاثين (30) من مواقع X-. بالتالي، على الرغم من أن العينة يمكن أن تكون ذكر، من الممكن أن يتضمن واحد أو أكثر من المواقع Y- ألائل معينة صفرية. بالمثل، على الرغم من أن العينة يمكن أن تكون أنثى، من الممكن أن يتضمن واحد أو أكثر من المواقع Y- أليلة معينة.

بالتالي، يمكن أن يتضمن تحليل توقع نوع العينة تحليل الدليل كاملا لتوقع نوع العينة. مثلا، يمكن أن يتضمن التحليل عد واحد على الأقل من (i) عدد مواقع Y- متوافقة مع العينة الذكر؛ (ii) عدد مواقع Y- متوافقة مع العينة الأنثى؛ (iii) عدد مواقع X- متوافقة مع العينة الذكر؛ (iv) أو عدد مواقع X- متوافقة مع العينة الذكر. في بعض النماذج، فقط يمكن اعتبار أعداد المواقع Y- في التحليل، عند 308، أو، بشكل بديل، فقط يمكن اعتبار أعداد المواقع X-. في بعض النماذج، يمكن اعتبار أعداد كل من مواقع X- وY- في التحليل، عند 308. في بعض النماذج، يمكن أن يتم إعطاء واحد أو أكثر من المواقع X- و/أو واحد أو أكثر من المواقع Y- وزن أكبر من مواقع أخرى.

على سبيل المثال، يمكن أن يقوم التحليل بمراجعة عشرة مواقع Y-. إذا كانت تسعة من أصل عشرة مواقع Y- تتضمن أليلة معينة، متوافقة مع العينة الذكر، يمكن توقع أن يكون نوع العينة ذكر. إذا كان واحد من أصل عشرة مواقع Y- تتضمن أليلة معينة، يمكن توقع أن يكون نوع العينة أنثى. في بعض النماذج، يمكن أن يحدد التحليل أن العينة تتضمن خليط. مثلا، إذا كان التحليل،

عند 308، يحدد عدد مواقع Y- وعدد مواقع X- تدعم كل من عينات ذكر وأنثى، يمكن توقع خليط المصادر.

الشكل 14 عبارة عن مخطط تدفق يشرح طريقة 320 للكشف عما إذا كانت عينة تتضمن خليط المصادر. يمكن دمج الطريقة 320 مع الطريقة 240 (الشكل 12) و، اختياريا، يمكن أن يتم تنفيذها بعد توقع نوع العينة. تتضمن الطريقة 300 تلقي، عند 322، بيانات موقع لكل موقع جيني لمجموعة من مواقع جينية. يمكن أن تتضمن بيانات الموقع واحد أو أكثر من الألائل المعينة أو المسماة للموقع الجيني المناظر. يمكن أن تتضمن بيانات الموقع درجات عد (مثلا، أعداد قراءة) للألائل المعينة، يمكن الحصول على درجات العد للألائل ضوضائية، ودرجات العد للمنتج المضطرب. يمكن الحصول على درجات العد كما هو موضح هنا.

بالنسبة لكل موقع جيني، يمكن أن تتضمن الطريقة 320 تحديد، عند 324، ما إذا كان عدد النسخ للموقع الجيني يتخطى أقصى عدد مسموح به للألائل (المشار إليها هنا فيما يلي بـ "أقصى عدد ألائل"). كما هو موضح أعلاه، يبلغ أقصى عدد ألائل للمواقع الكروموسومية الذاتية بشكل مثالي اثنين. أقصى عدد ألائل للمواقع X- أو مواقع Y- قائم على ما إذا كانت العينة (بافتراض عينة أحادية المصدر) عبارة عن ذكر أو أنثى. إذا كانت العينة عبارة عن ذكر، يبلغ أقصى عدد ألائل للموقع Y- واحد ويبلغ أقصى عدد ألائل للموقع X- هو وواحد. إذا كانت العينة عبارة عن أنثى، يبلغ أقصى عدد ألائل للموقع Y- صفر ويبلغ أقصى عدد ألائل للموقع X- اثنين. يمكن توقع أن تكون العينة ذكر أو يمكن توقع أن تكون أنثى بناء على الطريقة 300 الموضحة أعلاه.

بالتالي، يمكن أن يتضمن التحديد، عند 324، الحصول على أقصى أعداد ألائل للمواقع الجينية (مثلا، 0، 1، 2) ومقارنة عدد النسخ (أي، عدد ألائل معينة) لكل من المواقع الجينية إلى أقصى عدد ألائل مناظر. إذا كان عدد النسخ يتخطى أقصى عدد ألائل، يمكن توفير تنبيه عدد الألائل أو علم للموقع الجيني.

لكل موقع جيني، يمكن أن تتضمن الطريقة 300 تحديد، عند 326، ما إذا كان تناسب أليلة من الألائل المعينة غير متوازن. كما هو موضح أعلاه، يمكن أن يقوم تناسب الأليلة من موقع جيني على درجة عد (مثلا، عدد قراءة) لأليلة أولى معينة ودرجة عد (مثلا، عدد قراءة) لأليلة ثانية معينة. قد يتم توقع أن عينة أحادية المصدر يمكن أن تكون متماثلة الجينات عند موقع جيني أو غير متجانسة عند موقع جيني. إذا كانت غير متجانسة، قد يتم توقع أن تناسب الأليلة سوف يبلغ

نسبة حوالي 1،1. يمكن أن يشير عدم التناسب أساسا إلى أن الموقع الجيني غير متجانس أو العينة شاملة أكثر من مصدر واحد. بشكل أكثر تحديدا، كلما زادت النسبة المحسوبة التي تنحرف عن 1،1، كلما زادت احتمالية أن يكون الموقع الجيني إما متماثل الجينات أو العينة، ككل، يتضمن خليط المصادر. كما هو موضح أدناه، يقوم تحديد أن العينة تتضمن خليط المصادر على تحليل العديد من المواقع الجينية (مثلا كل المواقع الجينية التي تمت تسميتها).

في بعض النماذج، يمكن أن يتضمن التحديد، عند 326، حساب درجة توازن يقوم على نسبة درجات العد بين الاثنين من ألائل معينة للموقع الجيني. إذا كانت درجة التوازن ليست في نطاق معين، مثل 0.8،1.0 إلى حوالي 1.2،1.0، يمكن أن تشير درجة التوازن إلى أن تناسب الأليلة غير متوازن. إذا تم تحديد أن الموقع الجيني ليتضمن تناسب أليلة غير متوازن، يمكن توليد تنبيه تناسب أليلة للموقع الجيني. في بعض النماذج، يمكن مقارنة درجة التوازن بقيمة حدية معينة لتحديد ما إذا كان تناسب الأليلة غير متوازن.

يمكن أن تتضمن الطريقة 320 أيضا تحليل، عند 328، نتائج التحديد، عند 324، والتحديد، عند 326، لتحديد ما إذا كانت مجموعة من المصادر موجودة في العينة. يمكن أن يقوم التحليل، عند 328، على عدد تنبيه عدد الألائل وعدد تنبيهات تناسب أليلة لمجموعة المواقع الجينية. في أحد النماذج، يمكن حساب العدد الكلي للتنبيهات. إذا كان العدد الكلي للتنبيهات يتخطى القيمة الحدية للخليط، حينها يمكن أن يتم تعليم العينة بأن لها مجموعة من المصادر. يمكن أن تقوم القيمة الحدية للخليط على عدد المواقع الجينية التي تم تحليلها (أي، عدد المواقع الجينية في مجموعة المواقع الجينية). في نماذج معينة، يمكن أن تقوم القيمة الحدية للخليط على عدد المواقع الجينية التي تمت تسميتها. في بعض النماذج، القيمة الحدية للخليط قائمة على البيانات التاريخية أو المعرفة نسبة إلى فحص معين.

في بعض النماذج، يمكن أن تتضمن مجموعة المواقع الجينية، مثلا، 10، 20، 30، 40، 50، 60، 70، 80، 90، 100 مواقع جينية أو أكثر. في نماذج معينة، يمكن أن تتضمن مجموعة المواقع الجينية 120، 140، 160، 180، 200 مواقع جينية أو أكثر. في نماذج أكثر تحديدا، يمكن أن تتضمن مجموعة المواقع الجينية 250، 300، 350 مواقع جينية أو أكثر.

في بعض النماذج، القيمة الحدية للخليط عبارة عن قيمة محددة سلفا تساوي تناسب محدد سلفا للمواقع الجينية في المجموعة. يمكن أن يكون التناسب المحدد سلفا على الأقل، مثلا، 5٪، 10٪، 15٪، 20٪، 25٪، 30٪، 35٪، 40٪، 50٪، 60٪، 70٪، أو أكثر.

في بعض النماذج، يمكن أن يتضمن تنبيه عدد الألائل درجة عدد أليلة قائمة على عدد ألائل معينة. بشكل أكثر تحديدا، يمكن أن تزيد احتمالية العينة شاملة خليط مع زيادة عدد الألائل المعين وراء العدد الأقصى للألائل المقبولة للموقع الجيني. للتوضيح، إذا كان عدد الألائل المعينة لموقع جيني أول يبلغ ثلاثة (3) وعدد ألائل معينة لموقع جيني ثاني يبلغ (4)، يمكن تعيين درجة عدد أليلة للموقع الجيني الثاني بقيمة أكبر (أو وزن أكبر) من درجة عدد أليلة للموقع الجيني الأول عند تحديد ما إذا كان الخليط موجود.

في بعض النماذج، يمكن أن يتضمن تنبيه تناسب الأليلة درجة تناسب أليلة يقوم على التناسب من الألائل المعينة لموقع جيني. بشكل أكثر تحديدا، يمكن أن تزيد احتمالية أن تكون العينة التي تتضمن الخليط مع التناسب من ألائل معينة عندما تكون غير متناسبة. مثلا، إذا كان تناسب الأليلة لموقع جيني أول 1.0،1.3 وكان تناسب الأليلة لموقع جيني ثاني 1.0،2.0، يمكن تعيين قيمة أكبر لدرجة عدد أليلة للموقع الجيني الثاني (أو وزن أكبر) من درجة تناسب الأليلة للموقع الجيني الأول عند تحديد ما إذا كان خليط موجود.

في بعض النماذج، يمكن أن يتضمن تقرير العينة خليط تنبيه يقوم بإخطار المستخدم بأن العينة مشكوك في أنها تحتوي على مجموعة من المصادر. في بعض النماذج، يمكن أن يكون خليط التنبيه مصحوبا بواسطة درجة ثقة تقوم بإخطار المستخدم بمستوى الثقة في خليط التنبيه. يمكن أن تكون درجة الثقة قائمة على واحد على الأقل من عدد تنبيه أعداد الألائل، درجة أعداد الأليلة مرتبطة بتنبيه أعداد الألائل، عدد تنبيهات تناسب الأليلة، ودرجة تناسب الأليلة المرتبطة بتنبيهات تناسب الأليلة.

الشكل 15 يشرح نظام 400 مكون وفقا لأحد النماذج الذي يمكن استخدامه لتنفيذ العديد من الطرق المذكورة أعلاه. مثلا، يمكن استخدام النظام 400 لتنفيذ واحد أو أكثر من الطرق 100 (الشكل 1)، 150 (الشكل 1)، 160 (الشكل 8)، 200 (الشكل 11)، 240 (الشكل 12)، 300 (الشكل 13)، و340 (الشكل 14). يمكن أن تكون العديد من الخطوات آلية بواسطة النظام 400، مثل توالي، حيث يتطلب تنفيذ واحد أو أكثر من الخطوات يدويا أو تتطلب تفاعل المستخدم. في نماذج معينة،

يمكن أن يقوم المستخدم بتوفير عينة (مثلا، دم، لعاب، الشعر أو المني، إلخ.) ويمكن أن يقوم النظام 400 بتحضير وتوالي وتحليل العينة وتوفير هيئة جينية لمصدر العينة. في بعض النماذج، النظام 400 عبارة عن نظام منفرد متكامل موضوع عند موقع واحد. في النماذج الأخرى، يتم وضع واحد أو أكثر من مكونات النظام بعيدة عن بعضها البعض.

كما هو مبين، يتضمن النظام 400 مولد عينة sample generator 402، جهاز توالي sequencer 404، ووسيلة تحليل عينة sample analyzer 406. يمكن أن يقوم مولد العينة 402 بتحضير العينة لتوالي بروتوكول معين. مثلا، يمكن أن يقوم مولد العينة بتحضير العينة لتوالي بواسطة التخليق. يمكن أن يقوم جهاز التوالي 404 بتنفيذ التوالي لتوليد بيانات التوالي. كما هو موضح أعلاه، يمكن أن تتضمن بيانات التوالي مجموعة من قراءات عينة. يمكن أن تتضمن كل عينة قراءة عينة متوالية. في نماذج معينة، تشكل قراءات العينة أزواج قراءة التي يتم توليدها من توالي مزدوج الطرف أو، بشكل أكثر تحديدا، توالي غير متناظر طرفي مزدوج.

يمكن أن يتلقى وسيلة تحليل العينة 406 بيانات التوالي من جهاز التوالي 404. يتضمن الشكل 15 مخطط إطاري لوسيلة تحليل عينة 406 مكون وفقا لأحد النماذج. يمكن استخدام وسيلة تحليل العينة 406، مثلا، لتحليل بيانات توالي إلى توفير تسمية نمط جيني لموقع معين أو توليد هيئة جينية للعينة. تتضمن وسيلة تحليل العينة 406 وسيلة تحكم في النظام system controller 412 وسطح بيني للمستخدم user interface 414. وسيلة التحكم في النظام 412 مقترنة توصيليا بالسطح البيني للمستخدم 414 ويمكن أيضا أن يتم إقرانها توصيليا بجهاز التوالي 404 و/أو مولد العينة 402.

في نموذج مثالي، تتضمن وسيلة التحكم في النظام 412 واحد أو أكثر من معالجات/وحدات مهيأة لمعالجة و، اختياريا، تحليل بيانات توالي وفقا لواحد أو أكثر من طرق المذكورة أعلاه. مثلا، يمكن أن تتضمن وسيلة التحكم في النظام 412 واحد أو أكثر من وحدات مهيأة لتنفيذ مجموعة من تعليمات يتم تخزينها في واحد أو أكثر من عناصر التخزين storage elements (مثلا، التعليمات المخزنة على وسط تخزين قابل للقراءة readable storage medium بواسطة كومبيوتر ملموس و/أو غير انتقالي باستثناء الإشارات) لمعالجة بيانات التوالي. يمكن أن تتضمن مجموعة التعليمات العديد من الأوامر التي تقوم بإرشاد وسيلة التحكم في النظام 412 كماكينة معالجة لتنفيذ عمليات معينة مثل تدفقات العمل والعمليات والطرق الموضحة هنا. مثلا، يمكن أن تتضمن وسيلة تحليل

العينة 406 كومبيوتر مكتبي أو لابتوب أو مفكرة إلكترونية أو كومبيوتر لوحي أو هاتف ذكي. يمكن أن يتضمن السطح البيني للمستخدم 414 أجهزة وبرمجيات أو توليفات مما سبق والتي تسمح للفرد (مثلا، مستخدم) للتحكم مباشرة أو بشكل غير مباشرة في تشغيل وسيلة التحكم في النظام 412 والمكونات المختلفة منها. كما هو مبين، يتضمن السطح البيني للمستخدم 414 شاشة مشغل operator display 410.

في النموذج الموضح، تتضمن وسيلة التحكم في النظام 412 مجموعة من وحدات أو وحدات ثانوية تتحكم في تشغيل وسيلة التحكم في النظام 412. مثلا، يمكن أن تتضمن وسيلة التحكم في النظام 412 وحدات 421-426 ونظام تخزين storage system 426 يتصل بالبعض على الأقل من الوحدات 421-426. تتضمن الوحدات وحدة مرشح filter module أول 421، وحدة محاذي aligner module 422، وحدة مرشح ثاني 423، وحدة اضطراب stutter module 424، وحدة كاشف detector module 425، ووحدة تحليل analysis module 426. يمكن أن يتضمن النظام 400 وحدات أخرى أو وحدات ثانوية مهيأة لتنفيذ العمليات الموضحة هنا. وحدة المرشح الأولى 421 مهيأة لتحليل قراءات عينة لتحديد ما إذا كان قراءات العينة عبارة عن قراءات مؤكدة لموقع معين كما هو مذكور أعلاه. تتم تهيئة وحدة المحاذي 422 لتحليل القراءات المؤكدة وتحديد ما إذا كانت القراءات المؤكدة عبارة عن قراءات محاذية للموقع المعين كما هو مذكور أعلاه. وحدة المرشح الثاني 423 مهيأة لتلقي القراءات الصالحة وتحديد ما إذا كانت القراءات الصالحة تمثل ألائل محتملة للموقع المناظر كما هو مذكور أعلاه. تمت تهيئة وحدة الاضطراب 424 لتحديد ما إذا كانت القراءة الصالحة عبارة عن منتج مضطرب لأليلة أخرى كما هو مذكور أعلاه. وحدة الكاشف 425 مهيأة لتحديد ما إذا كان يجب الإشارة إلى أي أخطاء أو تنبيهات لمواقع مناظرة كما هو مذكور أعلاه. مثلا، يمكن أن تحدد وحدة الكاشف 425 أن موقع له عدد متتابع من قراءات غير متحاذية، تغطية منخفضة، عدد متتابع لألائل ضوضائية، ألائل غير متوازنة، و/أو خليط من ألائل من مصادر مختلفة. تتم تهيئة الوحدة تحليل 426 لتحديد نمط جيني للمواقع الجينية كما هو موضح هنا.

حسب الاستخدام هنا، يمكن أن تتضمن المصطلحات "وحدة"، "نظام،" أو "وسيلة تحكم في النظام" نظام جهاز و/أو برمجيات ودارة تعمل على تشغيل واحد أو أكثر من وظائف. مثلا، يمكن أن تتضمن وحدة، نظام، أو وسيلة تحكم في النظام معالج كومبيوتر computer processor، وسيلة

تحكم controller، أو جهاز آخر ذو منطق والذي يقوم بتنفيذ عمليات بناء على تعليمات مخزنة على وسط تخزين قراءة كومبيوتري ملموس وغير انتقالي، مثل ذاكرة كومبيوتر. بشكل بديل، يمكن أن تتضمن وحدة، نظام، أو جهاز سلكي متحكم في النظام والذي يقوم بتنفيذ عمليات بناء على المنطق الموصل سلكيا ودارة. يمكن أن تمثل الوحدة، النظام، أو وسيلة التحكم في النظام المبين في الأشكال الملحقة الجهاز ودارة تعمل بناء على برمجيات أو تعليمات سلكية، البرمجيات التي تقوم بتوجيه الجهاز لتنفيذ العمليات، أو توليفة مما سبق. يمكن أن تتضمن الوحدة، النظام، أو وسيلة التحكم في النظام يتضمن أو تمثل دارات جهاز hardware circuits أو دارات تتضمن و/أو تتصل بواحد أو أكثر من معالجات processors، مثل معالجة كومبيوترية دقيقة computer microprocessors.

[0100] حسب الاستخدام هنا، المصطلحات "برمجيات" و"برنامج" قابلة للتبادل، وتتضمن أي برنامج كومبيوتر مخزن في الذاكرة للتنفيذ بواسطة كومبيوتر، شاملة ذاكرة وصول عشوائي، ذاكرة القراءة فقط Read Only Memory (ROM)، ذاكرة القراءة فقط القابلة للبرمجة والمسح Erasable Programmable Read-Only Memory (EPROM)، ذاكرة قابلة للبرمجة والمسح كهربائيا Electrically Erasable Programmable Read-Only Memory (EEPROM)، و ذاكرة ذاكرة وصول عشوائي غير متطاير non-volatile random access memory (NVRAM). ما سبق عبارة عن أمثلة على أنواع ذاكرة، وهي غير مقيدة بالتالي لأنواع الذاكرة القابلة للاستخدام لتخزين برنامج كومبيوتر.

في بعض النماذج، يمكن فهم وحدة المعالجة processing unit، معالج processor، وحدة، أو نظام حوسبة "مهيأ لـ" تنفيذ هدف أو عملية على أنه مبني خصيصا لتنفيذ هدف أو عملية (مثلا، لها واحد أو أكثر من البرامج أو التعليمات المخزنة عليه أو المستخدمة فيما يتعلق بـ أو المصممة لتنفيذ الهدف أو العملية، و/أو لها تجهيزة ذات دارة معالجة processing circuitry مخصصة أو مطلوبة منها تنفيذ الهدف أو العملية). للتوضيح وتجنب الشك، لن تتم تهيئة كومبيوتر ذو غرض عام (والذي يمكن أن يكون "مهيأ لـ" تنفيذ الهدف والعملية إذا كانت مبرمجة بشكل مناسب) "مهيأ لـ" تنفيذ مهمة أو عملية ما لم أو لحين برمجة كومبيوتر خصيصا أو تعديله بنيويا لتنفيذ المهمة أو العملية.

تشرح الأشكال 16أ، 16ب و17أ-17و أمثلة على تقارير عينة sample reports 500، 520 والتي يمكن توليدها بواسطة النماذج الموضحة هنا. يمكن تصنيف التقارير عينة 500، 520 في واحد أو أكثر من الملفات ويتم بثها شبكة اتصال. يمكن عرض تقارير العينة 500، 520، مثلا، على شاشة أو تتم طباعتها على ورق. تشرح الأشكال 16أ و16ب فقط جزء من تقرير العينة 500 الكامل. كما هو مبين، يمكن أن يتضمن تقرير العينة 500 لمحة أو تحليل ملخص والذي يعتقد أنه عبارة عن عينة أحادية المصدر. يتضمن تقرير العينة 500 قطاع أول 511 لتحليل تكرار ترادفي قصير وقطاع ثاني 512 لتحليل تعدد شكلي لنيكلوتيد فردي. يمكن أن يقوم تقرير العينة 500 بالتأكيد على أن العينة عبارة عن مصدر فردي ذو علم أو مؤشر 510.

يتضمن تقرير العينة 500 تتضمن مصفوفة 502 من صناديق التسمية 504. يمكن أن يتعلق كل صندوق تسمية call box 504 بموقع جيني معين. مثلا، صندوق التسمية 504أ تناظر الموقع الجيني أميلوجينين genetic locus Amelogenin، وصندوق التسمية 504ب يناظر الموقع الجيني TPOX. كل من تتضمن صناديق التسمية call boxes 504 تسمية نمط جيني 506 للموقع الجيني. تسمية النمط الجيني 506 لأميلوجينين عبارة عن X، Y، وتسمية النمط الجيني لـ TPOX عبارة عن ألائل 11، 11. يمكن أن تكون أسماء الألائل قائمة على التسمية التقليدية خلال بروتوكولات التسمية الأخرى (مثلا، بروتوكول الملكية proprietary protocol).

يمكن أن تشير كل من صناديق التسمية 504 إلى ما إذا تم ربط علم أو إخطار بالموقع الجيني. مثلا، في الشكل 16، يتم تكويد صناديق التسمية 504 باللون للإشارة إلى ما إذا كانت هناك علم أو إخطار موجود. صندوق التسمية 504أ رمادي، وصندوق التسمية 504ج برتقالي أو أحمر. يمكن استخدام طرق أخرى للإشارة يمكن استخدامها في النماذج البديلة. في الشكل 16، كل من صناديق التسمية 504 المكودة بألوان تتضمن علم 508. تتم الإشارة إلى الأعلام 508 أعلاه في مفتاح الشكل 516 الذي يحدد الأعلام 508. مثلا، يتضمن تقرير العينة 500 أعلام 508 "اضطراب،" "عدد أليلة،" "غير متوازن،" "تغطية منخفضة،" "قيمة تفسير حدية،" و"مستخدم معدل." يمكن تعيين هذه الأعلام 508 إلى صناديق التسمية 504، مثلا، بعد التحليل الموضح هنا.

تقدم الأشكال 17أ-17و تحليل تفصيلي أكثر للمواقع الجينية. في بعض النماذج، يمكن أن يكون تقرير العينة 520 عبارة عن جزء من تقرير العينة 500 (الشكل 16). كما هو مبين، يتم تعيين رسم بياني 522 إلى كل من المواقع الجينية التي تمثل البيانات للموقع الجيني المناظر. في

النموذج الموضح، الرسم البياني 522 عبارة عن مخطط عمودي، ولكن يمكن استخدام رسومات بيانية أخرى لتمثيل البيانات بصريا. يشرح الرسم البياني 522 خصوصا شدة قراءة نسبة إلى الألائل المختلفة. يمكن أن تكون شدة القراءة هي درجة العد أو بناء على درجة العد كما هو موضح أعلاه. في بعض النماذج، شدة القراءة /درجة عد عبارة عن عدد القراءة.

يمكن أن تشير الرسومات البيانية 522 إلى قيمة تفسير حدية وقيمة حدية تحليلية نسبة إلى شدة القراءة (أو درجة العد). مثلا، يتضمن موقع D2S441 قيمة تفسير حدية 530 وقيمة حدية تحليلية 532. يمكن أن يكون التفسير والقيم الحدية التحليلية 530، 532 متشابهين مع التفسير والقيم الحدية التحليلية الموضحة أعلاه. كما هو مبين في الشكل 17، يمكن أن يكون التفسير والقيم الحدية التحليلية مختلفين بالنسبة للمواقع. مثلا، يتضمن موقع D21S11 قيمة تفسير حدية 550 أكبر من قيمة تفسير حدية 551 لموقع PENTAE. كما هو موضح أعلاه، يمكن أن تقوم قيمة التفسير الحدية و/أو القيمة الحدية التحليلية على (أي، وظيفة) العدد الكلي لقراءات التي تناظر الموقع المعين. اختياريا، يمكن أن تكون قيم التفسير الحدية و/أو القيم الحدية التحليلية عبارة عن وظيفة لأعداد القراءة لموقع معين وأيضا وظيفة لأعداد قراءة من مواقع أخرى و/أو أعداد قراءة من عينات أخرى.

في بعض النماذج، يمكن أن تشير الرسومات البيانية 522 أيضا إلى منتج مضطرب. يمكن أن تقوم الرسومات البيانية 522 بتمييز المنتج المضطرب من الألائل المختلفة. مثلا، يتضمن موقع D1S1656 أعمدة 541–543 تتعلق بألائل الاستشراد الكهربي 11، 12، و13، على التوالي، لـ D1S1656. يمكن أن تشير الأعمدة 541–543 إلى شدة قراءة (أو درجة عد) للألائل المناظرة. وقد قامت ألائل الموقع D1S1656 المبينة في الشكل 17 تاريخيا على البيانات الاستشراد الكهربي وتم تعليمها، بشكل تقليدي، 11، 12، و13. كما هو مشار إليه بواسطة ألوان مختلفة (مثلا، زرقاء وبنية) في الشكل 17، يمكن أن تتضمن الألائل للموقع D1S1656 منتج مضطرب. بشكل أكثر تحديدا، العمود 541 عبارة عن منتج مضطرب ولا يتخطى قيمة التفسير الحدية 555 للموقع D1S1656. يتضمن العمود 542 أجزاء عمودية 546، 547. كل من أجزاء العمود 546، 547 بصريا تمثل شدة القراءة. على الرغم من أن القراءات التي تناظر أجزاء العمود 546، 547 لها نفس طول المتوالية، القراءات التي تناظر أجزاء العمود 546، 547 لها متواليات مختلفة. يمثل جزء العمود 546 منتج مضطرب. مع ذلك، كما هو موضح أعلاه، يمكن أن يكون المنتج

المضطرب الممثل بواسطة جزء العمود 546 عبارة عن أليلة، مثل أليلة الاستشراد الكهربي 13. بالتالي، يمكن أن يقوم كود التلوين color coding (أو مؤشر آخر يقوم بتمييز المنتج المضطرب والألائل الحقيقية) بإخطار أو تنبيه المستخدم لتحليل المتواليات المختلفة للألائل الاستشراد الكهربي 11، 12، 13 لتوفير عملية تحديد أكثر ثقة للتسمية الجينية genetic call. في الشكل 17، التسمية الجينية للموقع D1S1656 هي 13/12. في حالات أخرى، مع ذلك، يمكن أن يقوم تحليل متواليات المنتج المضطرب بتغيير التسمية الجينية. بشكل أكثر تحديدا، في بعض الحالات، سوف تكون التسمية الجينية باستخدام عمليات الاستشراد الكهربي المعروفة غير صحيحة. يمكن أن تكون النماذج المذكورة أعلاه قادرة على توفير التسمية الجينية الصحيحة.

يقوم تقرير العينة 520 أيضا بتوفير أعلام أو إخطارات للمواقع الجينية المختلفة. يحدد مفتاح الشكل 524 الإخطارات. على سبيل المثال، يتضمن موقع D21S11 أعلام لـ "عد التوازن" و"عدد أليلة". بمعنى آخر، يشير تقرير العينة 520 إلى أنه من غير المتوقع أن تكون عدد الألائل وأن من غير المتوقع توازن الألائل. قد يرغب أحدهم في فحص التي تتعلق بموقع D21S11 بشكل أكبر.

في أحد النماذج، يتم توفير طريقة. تتضمن الطريقة تلقي بيانات توالي تتضمن مجموعة من قراءات عينة التي لها متواليات مناظرة لنيكلوتيدات. تتضمن الطريقة أيضا تعيين قراءات العينة إلى المواقع المعينة بناء على متوالية النيكلوتيدات، حيث قراءات العينة المعينة إلى موقع معين مناظر عبارة عن قراءات معينة للموقع المعين المناظر. تتضمن الطريقة أيضا تحليل القراءات المعينة لكل موقع معين لتحديد المناطق محل الاهتمام المناظرة في القراءات المعينة. تتضمن كل من الـ المناطق محل الاهتمام واحد أو أكثر من سلسلة من بواعث تكرار حيث كل باعث تكرار لسلسلة مناظرة تتضمن مجموعة متطابقة من النيكلوتيدات. تتضمن الطريقة أيضا تصنيف، لمواقع معينة لها قراءات معينة متطابقة، القراءات المعينة بناء على متواليات الـ المناطق محل الاهتمام بحيث تكون الـ المناطق محل الاهتمام ذات المتواليات المختلفة معينة كألائل محتملة مختلفة. كل أليلة محتملة لها متوالية مختلفة عن متواليات ألائل محتملة أخرى في الموقع المعين. تتضمن الطريقة أيضا تحليل، لمواقع معينة لها ألائل محتملة متعددة، متواليات الألائل المحتملة لتحديد ما إذا كانت الأليلة الأولى من الألائل المحتملة عبارة عن منتج مضطرب مشكوك فيه من أليلة ثانية من الألائل المحتملة. الأليلة الأولى عبارة عن المنتج المضطرب المشكوك فيه للأليلة الثانية إذا كانت

بواعث تكرار k في المتواليات المناظرة والتي تمت إضافتها أو إسقاطها بين الألائل الأولى والثانية، حيث k عبارة عن عدد كامل. اختياريا، k يساوي 1 أو 2.

في أحد الجوانب، تحليل، للمواقع المعينة التي لها ألائل محتملة متعددة، يمكن أن تتضمن متواليات الألائل المحتملة لتحديد سواء كانت الأليلة الأولى عبارة عن المنتج المضطرب المشكوك فيه للأليلة الثانية مقارنة أطوال الـ المناطق محل الاهتمام للألائل الأولى والثانية لتحديد إذا كانت أطوال الـ المناطق محل الاهتمام للألائل الأولى والثانية تختلف بمقدار باعث تكرار واحد أو بواعث تكرار متعددة.

في جانب آخر، تحليل، للمواقع المعينة التي لها ألائل محتملة متعددة، يمكن أن تتضمن متواليات الألائل المحتملة لتحديد سواء كانت الأليلة الأولى عبارة عن المنتج المضطرب المشكوك فيه للأليلة الثانية تحديد باعث التكرار التي تمت إضافتها أو إسقاطها وتحديد ما إذا كان باعث التكرار الذي تمت إضافته أو إسقاطه مطابق لباعث تكرار مجاور في المتواليات المناظرة.

في جانب آخر، يمكن أن تكون الأليلة الأولى عبارة عن المنتج المضطرب للأليلة الثانية إذا لم تكن هناك أي حالات عدم توافق موجودة بين متواليات المناطق محل الاهتمام للألائل الأولى والثانية.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا توليد هيئة نمط جيني، هيئة النمط الجيني تسمي نمط جيني لمجموعة على الأقل من المواقع المعينة، حيث المواقع المعينة التي لها منتج مضطرب مشكوك فيه مشار إليها بأن لها المنتج المضطرب المشكوك فيه.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا توفير تسميات النمط الجيني لمجموعة على الأقل من المواقع المعينة، حيث يشير واحد على الأقل من تسميات النمط الجيني إلى أن هناك منتجات مضطربة مشكوك فيها موجودة للموقع المعين لتسمية نمط جيني واحد على الأقل.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا عد، لكل موقع معين لها ألائل محتملة متعددة، العدد الكلي لقراءات العينة المسماة للأليلة المحتملة. يمكن أن تكون الأليلة الأولى عبارة عن المنتج المضطرب للأليلة الثانية إذا كانت قراءات العينة للأليلة الأولى أقل من قيمة حدية معينة لقراءات العينة للأليلة الثانية. اختياريا، القيمة الحدية المعينة تبلغ حوالي 40٪ من قراءات العينة للأليلة الثانية. اختياريا، يتم تعيين المنتج المضطرب المشكوك فيه على أنه من مساهم آخر إذا كانت قراءات العينة للأليلة الأولى يتخطى تناسب محدد سلفا من قراءات العينة للأليلة الثانية. اختياريا،

يتم تعيين المنتج المضطرب المشكوك فيه على أنه ضوضائي إذا كانت قراءات العينة للأليلة الأولى أقل من تناسب محدد سلفا من قراءات العينة للأليلة الثانية.

في جانب آخر، تتضمن القراءات المعينة مناطق مجنحة أولى وثانية محفوظة لها قطاع تكراري مناظر موضوع بينهما. لكل قراءة معينة، يمكن أن تتضمن الطريقة (أ) توفير متوالية مرجعية تتضمن المنطقة المجنحة المحفوظة الأولى والمنطقة المجنحة المحفوظة الثانية؛ (ب) محاذاة جزء من المنطقة المجنحة الأولى للمتوالية المرجعية مع القراءة المعينة المناظرة؛ (ج) محاذاة جزء من المنطقة المجنحة الثانية للمتوالية المرجعية مع القراءة المعينة المناظرة؛ و(د) تحديد طول و/أو متوالية القطاع التكراري.

اختياريا، تتضمن محاذاة الجزء من المنطقة المجنحة في واحد أو كل من الخطوات (ب) و(ج)، (i) تحديد موضع المنطقة المجنحة المحفوظة المناظرة على القراءة المعينة باستخدام توافق وحدة k محددة لمنطقة بذرية تتداخل مع أو تجأور القطاع التكراري و(ii) محاذاة المنطقة المجنحة مع القراءة المعينة.

اختياريا، يمكن أن تتضمن المنطقة البذرية تتضمن منطقة عالية التعقيد للمنطقة المجنحة المحفوظة. مثلا، المنطقة عالية التعقيد متوالية مميزة بشكل مناسب عن القطاع التكراري لتجنب سوء المحاذاة. كمثال آخر، المنطقة عالية التعقيد يتضمن متوالية لها خليط متعدد الأشكال لنيكلوتيدات.

اختياريا، تتجنب المنطقة البذرية مناطق منخفضة التعقد من المنطقة المجنحة المحفوظة المناظرة. مثلا، يمكن أن تتضمن المناطق منخفضة التعقد متواليات تقوم أساسا بتجميع مجموعة من بواعث التكرار.

اختياريا، يمكن أن تتضمن المنطقة البذرية مجاورة مباشرة للقطاع التكراري؛ المنطقة البذرية جزء من القطاع التكراري؛ أو المنطقة البذرية مزاحة من القطاع التكراري.

في جانب آخر، يمكن أن تكون قراءات العينة عبارة عن نواتج تكبير تفاعل سلسلة بوليمرات لها متواليات أولية متجهة إلى الأمام والخلف.

في جانب آخر، يمكن أن يتضمن تعيين قراءات العينة إلى المواقع المعينة متواليات التحديد لقراءات العينة التي تناظر متواليات أولية لنواتج تكبير تفاعل سلسلة بوليمرات.

في جانب آخر، يمكن أن تكون بيانات التوالي من فحص توالي بواسطة التخليق.

في جانب آخر، الـ منطقة محل الاهتمام تتضمن تكرار ترادفي قصير. اختياريا، يتم اختيار التكرار الترادفي القصير من واحد على الأقل من مواقع تكرار ترادفي قصير صبغية جسدية لنظام مؤشر الحمض النووي الريبي منقوص الأكسجين المدمج، مواقع نظام مؤشر الحمض النووي الريبي منقوص الأكسجين المدمج Y- تكرار ترادفي قصير، المواقع EU تكرار ترادفي قصير صبغية جسدية، أو المواقع EU Y- تكرار ترادفي قصير.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي بيانات توالي لها مجموعة من قراءات عينة لنواتج تكبير التي تناظر مجموعة من مواقع جينية. قراءات العينة يتضمن أزواج قراءة حيث كل زوج قراءة لناتج تكبير مناظر يتضمن قراءات أولى وثانية لناتج التكبير المناظر. كل من القراءات الأولى والثانية لها متوالية قراءة مناظرة. تتضمن الطريقة أيضا تحديد مواقع جينية محتملة للقراءات الأولية بناء على تحليل متواليات القراءة للقراءات الأولية. المواقع الجينية المحتملة من مجموعة المواقع الجينية. تتضمن الطريقة أيضا تحديد، لكل للقراءات الأولية لها موقع محتمل واحد على الأقل، ما إذا كانت القراءة الأولى تتحاذى مع متوالية مرجعية لكل من المواقع الجينية المحتملة. إذا كانت القراءة الأولى تتحاذى مع متوالية مرجعية لموقع جيني واحد فقط، تتضمن الطريقة تحديد أن القراءة الأولى تتضمن أليلة محتملة للموقع الجيني الواحد. إذا كانت القراءة الأولى تتحاذى مع أكثر من متوالية مرجعية واحدة، تتضمن الطريقة تحديد أن القراءة الأولى تتضمن أليلة محتملة للموقع الجيني لها المتوالية المرجعية تتحاذى بشكل أفضل مع القراءة الأولى. إذا كانت القراءة الأولى لا تتحاذى مع متوالية مرجعية، تتضمن الطريقة تعيين القراءة الأولى كقراءة غير متحاذية وتحليل القراءة غير المتحاذية لتحديد موقع جيني من المواقع الجينية المحتملة التي تتلاءم بشكل أفضل مع القراءة غير المتحاذية. تتضمن الطريقة أيضا توليد هيئة جينية تتضمن أنماط جينية مسماة لمجموعة على الأقل من المواقع الجينية، حيث الأنماط الجينية المسماة قائمة على أساس الألائل المحتملة للمواقع الجينية المناظرة. الهيئة الجينية تتضمن أيضا واحد أو أكثر من إخطارات للمواقع الجينية لها قراءات غير متحاذية.

في أحد الجوانب، يتضمن واحد على الأقل من الإخطارات درجة ثقة مرتبطة بالموقع الجيني المناظر. يمكن أن تقوم درجة الثقة على عدد قراءات غير متحاذية التي تتلاءم بشكل أفضل مع الموقع الجيني المناظر، حيث يشير عدد أكبر من القراءات غير المتحاذية إلى أن النمط الجيني المسمى أقل أهمية.

في جانب آخر، يمكن أن يتضمن تحليل القراءة غير المتحاذية لتحديد موقع جيني من المواقع الجينية المحتملة التي تتلاءم بشكل أفضل مع القراءة غير المتحاذية تحليل متوالية تحديد ثانوية للقراءة غير المتحاذية لتحديد الموقع الجيني التي تتلاءم بشكل أفضل مع تحديد المتوالية الثانوية.

في جانب آخر، يتضمن تحديد المتوالية الثانوية جزء على الأقل من متوالية بادئة primer sequence.

في جانب آخر، تحديد مواقع جينية محتملة للقراءات الأولية تتضمن تحديد أن متواليات أولية للقراءات الأولية تتوافق بفاعلية مع متواليات مرتبطة بالمواقع الجينية المحتملة.

في جانب آخر، يتم توليد بيانات التوالي خلال توالي غير متناظر طرفي مزدوج.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد ما إذا كان إسقاط الأليلة المحتملة موجود.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد صحة الفحص.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد ما إذا كان القراءات غير المتحاذية تشير إلى خيمر.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد عدد دايمرات أولية primer dimers.

في جانب آخر، يمكن أن يتضمن تحديد أن القراءة الأولى تتضمن أليلة محتملة للموقع الجيني تأكيد أن القراءة الثانية المناظرة للقراءة الأولى تتعلق أيضا بالموقع الجيني.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد إذا كانت القراءات غير المتحاذية عبارة عن قراءات واحد على هدف أو قراءات زوج على هدف. قراءات الزوج على الهدف يمكن أن تتضمن متواليات تحديد ثانوية أولى وثانية تتوافق بفاعلية مع متواليات اختيار أولى وثانية لقاعدة بيانات. قراءات الواحد على هدف يمكن أن تتضمن فقط متوالية التحديد الثانوية الأولى تتوافق بفاعلية مع متوالية الاختيار الأولى لقاعدة بيانات.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي بيانات توالي لها مجموعة من قراءات عينة لنواتج تكبير التي تناظر مجموعة من مواقع جينية. قراءات العينة يتضمن أزواج قراءة حيث كل زوج قراءة لناتج تكبير مناظر يتضمن قراءات أولى وثانية لناتج التكبير المناظر. كل من القراءات

الأولى والثانية لها متوالية قراءة مناظرة. تتضمن الطريقة أيضا تحديد مواقع جينية محتملة للقراءات الأولية بناء على تحليل متواليات القراءة للقراءات الأولية. المواقع الجينية المحتملة من مجموعة المواقع الجينية. تتضمن الطريقة أيضا تحديد، لكل للقراءات الأولية لها موقع محتمل واحد على الأقل، ما إذا كانت القراءة الأولى تتحاذى مع متوالية مرجعية لكل من المواقع الجينية المحتملة. تتضمن الطريقة أيضا تعيين القراءات الأولية التي لا تتحاذى مع متوالية مرجعية كقراءات غير متحاذية. تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد موقع جيني من المواقع الجينية المحتملة التي تتلاءم بشكل أفضل مع القراءة غير المتحاذية. تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد ما إذا كان إسقاط الأليلة المحتملة موجود للموقع الجيني ذو الملائمة الأفضل.

في أحد الجوانب، يمكن أن تتضمن الطريقة أيضا تحليل القراءات غير المتحاذية لتحديد إذا كانت القراءات غير المتحاذية عبارة عن قراءات واحد على هدف أو قراءات زوج على هدف. قراءات الزوج على الهدف يمكن أن تتضمن متواليات تحديد ثانوية أولى وثانية تتوافق بفاعلية مع متواليات اختيار أولى وثانية لقاعدة بيانات. قراءات الواحد على هدف يمكن أن تتضمن فقط متوالية التحديد الثانوية الأولى تتوافق بفاعلية مع متوالية الاختيار الأولى لقاعدة بيانات. يمكن أن يقوم تحليل القراءات غير المتحاذية لتحديد ما إذا كان إسقاط الأليلة المحتملة موجود للموقع الجيني ذو الملائمة الأفضل على عدد قراءات زوج على هدف.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي توزيع قراءة لكل موقع جيني لمجموعة من مواقع جينية. توزيع القراءة تتضمن مجموعة من ألائل محتملة، حيث كل أليلة محتملة لها متوالية أليلة وعدد قراءة. يمثل عدد القراءة عدد من قراءات عينة من بيانات توالي تم تحديدها على أنها تتضمن الأليلة المحتملة. يمكن أن تتضمن الطريقة أيضا تحديد، لكل موقع جيني لمجموعة مواقع جينية، واحد من الألائل المحتملة لتوزيع القراءة التي لها عدد قراءة أقصى. يمكن أن تتضمن الطريقة أيضا تحديد، لكل موقع جيني لمجموعة مواقع جينية، ما إذا كان عدد القراءة الأقصى يتخطى قيمة تفسير حدية. إذا كانت أقصى قراءة تتخطى قيمة التفسير الحدية، تتضمن الطريقة تحليل الأليلة المحتملة للموقع الجيني المناظر لتسمية نمط جيني للموقع الجيني. إذا كانت أقصى قراءة أقل من قيمة التفسير الحدية، تتضمن الطريقة توليد تنبيه أن الموقع الجيني له تغطية منخفضة. تتضمن

الطريقة أيضا توليد هيئة جينية التي لها الأنماط الجينية لكل للمواقع الجينية التي تم من أجلها تسمية النمط الجيني والتنبيه للمواقع الجينية التي لها تغطية منخفضة.

في أحد الجوانب، يمكن أن يتضمن تحليل الأليلة المحتملة للموقع الجيني المناظر لتسمية النمط الجيني للموقع الجيني مقارنة عدد الألائل المحتملة لكل موقع جيني إلى عدد أقصى محتمل من ألائل مقبولة للموقع الجيني وتوليد تنبيه أن الموقع الجيني له عدد متتابع من ألائل إذا كان عدد الألائل المحتملة يتخطى العدد الأقصى المحدد سلفا للألائل المقبولة.

في جانب آخر، يمكن أن يتضمن تحليل الأليلة المحتملة للموقع الجيني المناظر لتسمية النمط الجيني للموقع الجيني توليد تنبيه أن الموقع الجيني غير متوازن إذا كان الموقع الجيني يتضمن مجموعة من ألائل محتملة تتضمن نسب غير كافية نسبة إلى بعضها البعض.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا تحديد، لكل موقع جيني لمجموعة مواقع جينية، ما إذا كان أعداد القراءة للألائل المحتملة تتخطى قيمة حدية تحليلية. يمكن أن تكون القيمة الحدية التحليلية أسهل في التخطي من قيمة التفسير الحدية.

في جانب آخر، يتم تعيين الألائل المحتملة لها أعداد قراءة لا تتخطى قيمة التفسير الحدية على أنها ألائل ضوضائية، تتضمن الطريقة أيضا مقارنة إجمالي أعداد القراءة للألائل الضوضائية إلى قيمة حدية ضوضائية وتوليد تنبيه أن الموقع الجيني يتضمن ضوضاء مفرطة إذا كان الإجمالي تتخطى القيمة الحدية الضوضائية noise threshold.

اختياريا، تتضمن المواقع الجينية مواقع تكرار ترادفي قصير ومواقع تعدد شكلي لنيكلوتيد فردي.

في أحد النماذج، يتم توفير طريقة تتضمن، (أ) تلقي توزيع قراءة لموقع جيني. توزيع القراءة تتضمن مجموعة من ألائل محتملة، حيث كل أليلة محتملة لها متوالية أليلة ودرجة عد. درجة العد قائم على عدد من قراءات عينة من بيانات توالي تم تحديدها على أنها تتضمن الأليلة المحتملة. تتضمن الطريقة أيضا، (ب) تحديد ما إذا كان الموقع الجيني له تغطية منخفضة بناء على درجة العد لأليلة أخرى من الألائل المحتملة. إذا كان الموقع الجيني له تغطية منخفضة، تتضمن الطريقة توليد إخطار أن الموقع الجيني له تغطية منخفضة. إذا كان الموقع الجيني لا يتضمن تغطية منخفضة، تتضمن الطريقة تحليل درجات العد للألائل المحتملة لتحديد نمط جيني للموقع الجيني. تتضمن الطريقة أيضا، (د) توليد هيئة جينية تتضمن النمط الجيني للموقع الجيني أو التنبيه أن الموقع الجيني له تغطية منخفضة.

في أحد الجوانب، يمكن أن يتضمن تحديد ما إذا كان الموقع الجيني له تغطية منخفضة تحديد ما إذا كان واحد أو أكثر من درجات العد للألائل المحتملة يتخطى قيمة تفسير حدية. إذا كانت واحد على الأقل من درجات العد يتخطى قيمة التفسير الحدية، يمكن أن تتضمن الطريقة أيضا تحليل الألائل المحتملة للموقع الجيني المناظر لتسمية نمط جيني للموقع الجيني. إذا لم تتخطى أي من درجات العد يتخطى قيمة التفسير الحدية، يمكن أن تتضمن الطريقة توليد الإخطار أن الموقع الجيني له تغطية منخفضة.

في جانب آخر، تحديد ما إذا كان الموقع الجيني له تغطية منخفضة تتضمن تحديد ما إذا كان واحد أو أكثر من درجات العد للألائل المحتملة يتخطى قيمة حدية تحليلية. إذا كانت واحد على الأقل من درجات العد يتخطى القيمة الحدية التحليلية، يمكن أن تتضمن الطريقة أيضا تحليل الألائل المحتملة للموقع الجيني المناظر لتسمية نمط جيني للموقع الجيني. إذا لم تتخطى أي من درجات العد القيمة الحدية التحليلية، يمكن أن تتضمن الطريقة أيضا توليد الإخطار أن الموقع الجيني له تغطية منخفضة.

في جانب آخر، تحديد ما إذا كان الموقع الجيني له تغطية منخفضة تتضمن مقارنة العدد الكلي لقراءات محاذية للموقع الجيني بقراءة قيمة حدية. إذا كانت العدد الكلي لقراءات محاذية يتخطى قراءة القيمة الحدية، يمكن أن تتضمن الطريقة تحليل الألائل المحتملة للموقع الجيني المناظر لتسمية نمط جيني للموقع الجيني. إذا كانت العدد الكلي لقراءات محاذية لا تتخطى قراءة القيمة الحدية، يمكن أن تتضمن الطريقة توليد الإخطار أن الموقع الجيني له تغطية منخفضة.

في جانب آخر، كل من درجات العد عبارة عن قيمة تساوي عدد قراءة للألية المحتملة المناظرة.

في جانب آخر، كل من درجات العد عبارة عن وظيفة قائمة على عدد قراءة والعدد الكلي لقراءات للموقع الجيني.

في جانب آخر، كل من درجات العد عبارة عن وظيفة قائمة على عدد قراءة وبيانات تم الحصول عليها سلفا للموقع الجيني.

في جانب آخر، كل من درجات العد عبارة عن وظيفة قائمة على أعداد قراءة من مواقع جينية أخرى للعينة.

في جانب آخر، كل من درجات العد عبارة عن وظيفة قائمة على أعداد قراءة للموقع الجيني من عينات أخرى.

في جانب آخر، تحليل الألائل المحتملة للموقع الجيني لتسمية النمط الجيني للموقع الجيني تتضمن أيضا مقارنة عدد الألائل المحتملة للموقع الجيني إلى عدد أقصى محتمل من ألائل مقبولة للموقع الجيني وتوليد تنبيه أن الموقع الجيني له عدد متتابع من ألائل إذا كان عدد الألائل المحتملة يتخطى العدد الأقصى المحدد سلفا للألائل المقبولة.

في جانب آخر، يمكن أن يتضمن تحليل الألائل المحتملة للموقع الجيني لتسمية النمط الجيني للموقع الجيني توليد إخطار أن الموقع الجيني غير متوازن إذا كان الموقع الجيني يتضمن مجموعة من ألائل محتملة تتضمن نسب غير كافية نسبة إلى بعضها البعض.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا تحديد ما إذا كان درجات العد للألائل المحتملة تتخطى قيمة حدية تحليلية. يمكن أن تكون القيمة الحدية التحليلية أسهل في التخطي من قيمة التفسير الحدية. اختياريا، يتم تعيين الألائل المحتملة لها درجات عد لا تتخطى القيمة الحدية التحليلية على أنها ألائل ضوضائية. يمكن أن تتضمن الطريقة أيضا مقارنة درجة ضوضائية إلى قيمة حدية ضوضائية وتوليد تنبيه أن الموقع الجيني تتضمن ضوضاء مفرطة إذا كانت الدرجة الضوضائية تتخطى القيمة الحدية الضوضائية. يمكن أن تكون الدرجة الضوضائية قائمة على درجات العد للألائل الضوضائية.

اختياريا، الموقع الجيني عبارة عن موقع تكرار ترادفي قصير أو موقع تعدد شكلي لنيكلوتيد فردي.

في جانب آخر، تتضمن الطريقة تكرار (أ)-(ج) لمجموعة من مواقع جينية، حيث توليد الهيئة الجينية تتضمن تسمي نمط جيني لكل للمواقع الجينية أو توفير إخطار لكل للمواقع الجينية لها تغطية منخفضة.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي توزيع قراءة لموقع جيني. توزيع القراءة تتضمن مجموعة من ألائل محتملة، حيث كل أليلة محتملة لها متوالية أليلة وعدد قراءة. يمثل عدد القراءة عدد من قراءات عينة من بيانات توالي والتي تم تعيينها إلى الموقع الجيني. يمكن أن تتضمن الطريقة أيضا تحديد درجة عد لكل للألائل المحتملة. يمكن أن تكون درجة العد قائمة على عدد القراءة للأليلة المحتملة. يمكن أن تتضمن الطريقة أيضا تحديد ما إذا كان درجات العد للألائل المحتملة تتخطى قيمة حدية تحليلية. إذا كانت درجة العد لأليلة محتملة مناظرة لا تتخطى القيمة الحدية التحليلية، تتضمن الطريقة التخلص من الأليلة المحتملة المناظرة. إذا كانت درجة العد لأليلة

محتملة مناظرة يتخطى القيمة الحدية التحليلية، تتضمن الطريقة تعيين الأليلة المحتملة كأليلة معينة للموقع الجيني.

في أحد الجوانب، التخلص من الأليلة المحتملة المناظرة تتضمن تعيين الأليلة المحتملة كأليلة ضوضائية. يمكن أن تتضمن الطريقة أيضا تحديد ما إذا كان درجات العد للألائل الضوضائية تتخطى جمعا قيمة حدية ضوضائية. إذا كانت درجات العد تتخطى جمعا القيمة الحدية الضوضائية، يمكن أن تتضمن الطريقة توليد تنبيه أن الموقع الجيني يتضمن ضوضاء مفرطة.

في جانب آخر، كل من درجات العد عبارة عن قيمة تساوي عدد القراءة للأليلة المحتملة المناظرة.

في جانب آخر، كل من درجات العد عبارة عن وظيفة قائمة على عدد القراءة والعدد الكلي لقراءات للموقع الجيني.

في جانب آخر، كل من درجات العد عبارة عن وظيفة قائمة على عدد القراءة وبيانات تم الحصول عليها سلفا للموقع الجيني.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا مقارنة عدد ألائل معينة إلى عدد أقصى محتمل من ألائل مقبولة للموقع الجيني وتوليد تنبيه أن الموقع الجيني له عدد متتابع من ألائل إذا كان عدد ألائل معينة تتخطى العدد الأقصى المحدد سلفا للألائل المقبولة.

في جانب آخر، تتضمن الطريقة أيضا توليد تنبيه أن الموقع الجيني غير متوازن إذا كان الموقع الجيني يتضمن مجموعة من ألائل معينة تتضمن نسب غير كافية نسبة إلى بعضها البعض.

اختياريا، تتضمن المواقع الجينية مواقع تكرار ترادفي قصير ومواقع تعدد شكلي لنيكلوتيد فردي.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي توزيع قراءة لموقع جيني. توزيع القراءة تتضمن مجموعة من ألائل محتملة، حيث كل أليلة محتملة لها متوالية أليلة وعدد قراءة. يمثل عدد القراءة عدد من قراءات عينة من بيانات توالي والتي تم تعيينها إلى الموقع الجيني. تتضمن الطريقة أيضا تحديد ما إذا كان أعداد القراءة تتخطى قيمة حدية تحليلية. إذا كان عدد القراءة لأليلة محتملة مناظرة أقل من القيمة الحدية التحليلية، تتضمن الطريقة تعيين الأليلة المحتملة المناظرة كأليلة ضوضائية. إذا كان عدد القراءة لأليلة محتملة مناظرة يتخطى القيمة الحدية التحليلية، تتضمن الطريقة تعيين الأليلة المحتملة كأليلة للموقع الجيني. تتضمن الطريقة أيضا تحديد ما إذا كان إجمالي أعداد القراءة للألائل الضوضائية تتخطى قيمة حدية ضوضائية. إذا كان الإجمالي تتخطى القيمة الحدية الضوضائية، تتضمن الطريقة توليد تنبيه أن الموقع الجيني يتضمن ضوضاء مفرطة.

في أحد الجوانب، يمكن أن تتضمن الطريقة أيضا مقارنة عدد ألائل معينة إلى عدد أقصى محتمل من ألائل مقبولة للموقع الجيني وتوليد تنبيه أن الموقع الجيني له عدد متتابع من ألائل إذا كان عدد ألائل معينة تتخطى العدد الأقصى المحدد سلفا للألائل المقبولة.

في جانب آخر، يمكن أن تتضمن الطريقة أيضا توليد تنبيه أن الموقع الجيني غير متوازن إذا كان الموقع الجيني يتضمن مجموعة من ألائل معينة تتضمن نسب غير كافية نسبة إلى بعضها البعض.

اختياريا، تتضمن المواقع الجينية مواقع تكرار ترادفي قصير ومواقع تعدد شكلي لنيكلوتيد فردي.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي بيانات موقع لكل موقع جيني لمجموعة من مواقع جينية. بيانات الموقع تتضمن واحد أو أكثر من ألائل معينة للموقع الجيني المناظر. كل أليلة معينة قائم على أعداد القراءة التي تم الحصول عليها من بيانات توالي. تتضمن الطريقة أيضا تحديد، لكل موقع جيني لمجموعة مواقع جينية، ما إذا كان عدد ألائل معينة للموقع الجيني المناظر أكبر من عدد أقصى محدد سلفا للألائل المقبولة للموقع الجيني المناظر. يمكن أن تتضمن الطريقة توليد تنبيه عدد الألائل إذا كان عدد ألائل معينة تتخطى العدد الأقصى المحدد سلفا للألائل المقبولة. تتضمن الطريقة أيضا تحديد، لكل موقع جيني لمجموعة مواقع جينية، ما إذا كان تناسب أليلة من الألائل المعينة غير كافي. يمكن أن يكون تناسب الأليلة قائم على أعداد قراءة الألائل المعينة. يمكن أن تتضمن الطريقة أيضا توليد تنبيه تناسب أليلة إذا كان تناسب الأليلة غير متوازن. يمكن أن تتضمن الطريقة أيضا تحديد أن العينة تتضمن خليط من مجموعة من مصادر بناء على عدد تنبيه عدد الألائل وتنبيهات تناسب أليلة لمجموعة المواقع الجينية.

في أحد الجوانب، تحديد أن العينة تتضمن خليط من مجموعة من مصادر تتضمن تحديد أن العدد الكلي للتنبيهات يتخطى قيمة حدية للخليط. اختياريا، القيمة الحدية للخليط قائمة على عدد مواقع جينية في مجموعة المواقع الجينية. اختياريا، القيمة الحدية للخليط قيمة محددة سلفا تساوي تناسب محدد سلفا للمواقع الجينية في المجموعة.

في جانب آخر، يتضمن توليد تنبيه عدد الألائل توفير درجة عدد أليلة قائمة على عدد ألائل معينة. يمكن أن يتضمن تحديد أن العينة تتضمن خليط من مجموعة من مصادر تحليل درجة عدد أليلة. اختياريا، تزيد احتمالية للعينة شاملة خليط حيث يزيد عدد ألائل معينة وراء العدد الأقصى للألائل المقبولة.

في جانب آخر، توليد تنبيه تناسب أليلة تتضمن توفير درجة تناسب أليلة تقوم على تناسب الأليلة. تحديد أن العينة تتضمن خليط من مجموعة من مصادر تتضمن تحليل درجة تناسب الأليلة. اختياريا، تزيد احتمالية للعينة شاملة خليط بينما يزيد عدم التناسب بين الألائل.

اختياريا، تتضمن المواقع الجينية مواقع تكرار ترادفي قصير ومواقع تعدد شكلي لنيكلوتيد فردي.

في أحد النماذج، يتم توفير طريقة تتضمن تلقي بيانات موقع لمجموعة من مواقع Y-. بيانات الموقع يتضمن ألائل معينة للمواقع Y-. كل أليلة معينة قائم على أعداد القراءة التي تم الحصول عليها من بيانات توالي. تتضمن الطريقة أيضا مقارنة عدد ألائل معينة لكل موقع Y- بعدد غير متوقع من ألائل المواقع Y-. تتضمن الطريقة أيضا توليد توقع أن العينة عبارة عن ذكر أو أنثى بناء على نتائج من عملية المقارنة. اختياريا، المواقع الجينية يتضمن مواقع تكرار ترادفي قصير ومواقع تعدد شكلي لنيكلوتيد فردي.

في واحد أو أكثر من النماذج، يتم توفير نظام يتضمن وسيلة تحليل عينة مهيأة لتنفيذ واحد أو أكثر من عناصر الحماية المذكورة أعلاه.

في هذا الطلب تمت الإشارة إلى العديد من المنشورات الأمريكية وبراءات الاختراع و/أو طلبات براءات الاختراع. وقد تم دمج الكشف الخاص بهذه المنشورات هنا في هذا الطلب كليا بالإشارة إليها كمراجع.

حسب الاستخدام هنا، يقصد بالمصطلحات "تتضمن،" "شاملة،" و"لها،" وما شابه أن تكون ذات نهاية مفتوحة، ولا تشمل فقط العناصر المذكورة، ولكن يمكن أن تتضمن عناصر إضافية .

سوف يتم فهم أن الوصف السابق للتوضيح وليس التقييد. مثلا، يمكن استخدام النماذج السابقة (و/أو جوانبه) في توليفة مع بعضها البعض. علاوة على ذلك، يمكن تنفيذ العديد من التعديلات لتهيئة موقف أو مادة معينة مع كل التعليمات الخاصة بالاختراع دون الخروج من منظوره. يقصد بالأبعاد وأنواع المواد وتوجيهات المكونات المختلفة وعدد ومواضع المكونات المختلفة الموضحة هنا أن تقوم بتحديد متغيرات النماذج المعينة، ولا يقصد بها أن تكون مقيدة وهي عبارة عن نماذج مثالية. سوف تكون هناك العديد من النماذج الأخرى والتعديلات في منظور وروح عناصر الحماية واضحة للخبير في المجال عند مراجعة الوصف السابق. يجب بالتالي تحديد منظور الاختراع، بالتالي، بالإضافة إلى عناصر الحماية الملحقة، بالإضافة إلى المنظور الكامل للنماذج المكافئة لعناصر الحماية هذه.

حسب الاستخدام في الوصف، العبارات "في نموذج مثالي،" "في بعض النماذج،" "في نماذج معينة،" وما شابه تعني أن النموذج الموضح عبارة عن أمثلة على النماذج التي يمكن تكوينها أو تنفيذها وفقا للطلب الحالي. لا يقصد بالعبارة أن تقيد الموضوع الابتكاري للنموذج. بشكل أكثر تحديدا، يمكن ألا تتضمن النماذج الأخرى للموضوع الابتكاري الخاصية المذكورة أو البنية الموضحة بالإشارة إلى نموذج معين.

في عناصر الحماية الملحقة، يتم استخدام المصطلحات "شاملة" و"حيث" كمكافئات إنجليزية للمصطلح المناظر "تتضمن" و"حيث." علاوة على ذلك، في عناصر الحماية التالي، يتم استخدام المصطلحات "أول،" "ثاني،" و"ثالث،" إلخ. بالكاد كعلامات، ولا يقصد بها أن تضع قيد رقمي على أهدافها. أيضا، لم تتم كتابة القيود في صيغ وظيفة + وسيلة ولن يتم تفسيرها بناء على المادة 35 من القانون الأمريكي § 112 (و) ما لم تتم الإشارة صراحة إلى هذا القيد بواسطة "وسيلة لـ" متبوعة ببيان وظيفة خاصة ببنية أخرى.

تشير عناصر الحماية التالية إلى واحد أو أكثر من نماذج الطلب الحالي والتي تم دمجها هنا في وصف الطلب الحالي.

عناصر الحماية

1. مُحلل عينات sample analyzer يشتمل على:

وحدة تحكم في نظام system controller؛ و

واجهة مستخدم user interface؛

حيث تشتمل وحدة التحكم في النظام system controller على مجموعة من الوحدات، حيث تشتمل مجموعة الوحدات على:

وحدة لتلقي بيانات توالي شاملة مجموعة من قراءات عينة التي لها متواليات مناظرة لنيكلوتيدات nucleotides؛

وحدة مرشح filter module أولى لتعيين قراءات العينة إلى المواقع المعينة بناء على متوالية النيكلوتيدات nucleotides، حيث قراءات العينة المعينة إلى موقع معين مناظر عبارة عن قراءات معينة للموقع المعين المناظر؛

وحدة اصطفاف aligner module لتحليل القراءات المعينة لكل موقع معين لتحديد المناطق محل الاهتمام regions-of-interest (ROIs) المناظرة في القراءات المعينة، كل من المناطق محل الاهتمام regions-of-interest لها واحد أو أكثر من سلسلة من بواعث تكرار repeat motifs حيث كل باعث تكرار repeat motif لسلسلة مناظرة تتضمن مجموعة متطابقة من النيكلوتيدات nucleotides؛

وحدة مرشح filter module ثانية لتصنيف، لمواقع معينة لها قراءات معينة متطابقة، القراءات المعينة بناء على متواليات المناطق محل الاهتمام regions-of-interest بحيث تكون المناطق محل الاهتمام regions-of-interest ذات المتواليات المختلفة معينة كألائل alleles محتملة مختلفة، كل أليلة allele محتملة لها متوالية مختلفة عن متواليات ألائل alleles محتملة أخرى في الموقع المعين؛ و

وحدة تقطع stutter module لتحليل، لمواقع معينة لها ألائل alleles محتملة متعددة، متواليات الألائل alleles المحتملة لتحديد ما إذا كانت الأليلة allele الأولى من الألائل alleles المحتملة عبارة عن منتج مضطرب مشكوك فيه suspected stutter product من أليلة allele ثانية من الألائل alleles المحتملة، الأليلة allele الأولى عبارة عن المنتج المضطرب المشكوك فيه suspected stutter produc للأليلة allele الثانية إذا كانت بواعث تكرار k repeat motifs في

المتواليات المناظرة والتي تمت إضافتها أو إسقاطها بين الألائل alleles الأول والثاني، حيث k عبارة عن عدد كامل.

2. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1، حيث تحليل، للمواقع المعينة التي لها ألائل alleles محتملة متعددة، متواليات الألائل alleles المحتملة لتحديد سواء كانت الأليلة allele الأولى عبارة عن المنتج المضطرب المشكوك فيه suspected stutter product للأليلة allele الثانية تتضمن مقارنة أطوال المناطق محل الاهتمام regions-of-interest للألائل alleles الأول والثاني لتحديد إذا كانت أطوال المناطق محل الاهتمام regions-of-interest للألائل alleles الأول والثاني تختلف بمقدار باعث تكرار repeat motif واحد أو بواعث تكرار repeat motifs متعددة.

3. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث تحليل، للمواقع المعينة التي لها ألائل alleles محتملة متعددة، متواليات الألائل alleles المحتملة لتحديد سواء كانت الأليلة allele الأولى عبارة عن المنتج المضطرب المشكوك فيه suspected stutter product للأليلة allele الثانية تتضمن تحديد باعث التكرار repeat motif التي تمت إضافتها أو إسقاطها وتحديد ما إذا كان باعث التكرار repeat motif الذي تمت إضافته أو إسقاطه مطابق لباعث تكرار repeat motif مجاور في المتواليات المناظرة.

4. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث k يساوي 1 أو 2.

5. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث الأليلة allele الأولى عبارة عن المنتج المضطرب stutter product للأليلة allele الثانية إذا لم تكن هناك أي حالات عدم توافق موجودة بين متواليات المناطق محل الاهتمام regions-of-interest للألائل alleles الأول والثاني.

6. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث تتضمن وحدة التحكم في النظام system controller أيضا وحدة لتوليد هيئة نمط جيني genotype profile، هيئة النمط الجيني genotype profile تسمي نمط جيني genotyp لمجموعة على الأقل من المواقع المعينة، حيث المواقع المعينة التي لها منتج مضطرب مشكوك فيه suspected stutter product مشار إليها بأن لها المنتج المضطرب المشكوك فيه suspected stutter product.

7. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث تتضمن وحدة التحكم في النظام system controller أيضا وحدة لتوفير تسميات النمط الجيني genotype لمجموعة على الأقل من المواقع المعينة، حيث يشير واحد على الأقل من تسميات النمط الجيني genotype إلى أن هناك منتجات مضطربة مشكوك suspected stutter product فيها موجودة للموقع المعين لتسمية نمط جيني genotype واحد على الأقل.

8. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، يتضمن أيضا وحدة لعد، لكل موقع معين لها ألائل alleles محتملة متعددة، العدد الكلي لقراءات العينة المسماة للأليلة allele المحتملة، حيث الأليلة allele الأولى عبارة عن المنتج المضطرب stutter product للأليلة allele الثانية إذا كانت قراءات العينة للأليلة allele الأولى أقل من قيمة حدية معينة لقراءات العينة للأليلة allele الثانية.

9. مُحلل العينات sample analyzer وفقا لعنصر الحماية 8، حيث القيمة الحدية المعينة تبلغ حوالي 40٪ من قراءات العينة للأليلة allele الثانية

10. مُحلل العينات sample analyzer وفقا لعنصر الحماية 8، حيث يتم تعيين المنتج المضطرب المشكوك فيه suspected stutter product على أنه من مساهم آخر إذا كانت قراءات العينة للأليلة allele الأولى يتخطى تناسب محدد سلفا من قراءات العينة للأليلة allele الثانية.

11. مُحلل العينات sample analyzer وفقا لعنصر الحماية 8، حيث يتم تعيين المنتج المضطرب المشكوك suspected stutter product فيه على أنه ضوضائي إذا كانت قراءات العينة للأليلة allele الأولى أقل من تناسب محدد سلفا من قراءات العينة للأليلة allele الثانية.

12. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث تتضمن القراءات المعينة مناطق مجنحة flanking regions أولى وثانية محفوظة لها قطاع تكراري repetitive segment مناظر موضوع بينهما، حيث، لكل قراءة معينة، تتضمن وحدة التحكم في النظام system controller أيضا واحدة أو أكثر من الوحدات من أجل،

أ) توفير متوالية مرجعية reference sequence تتضمن المنطقة المجنحة flanking region المحفوظة الأولى والمنطقة المجنحة flanking region المحفوظة الثانية؛

ب) محاذاة جزء من المنطقة المجنحة flanking region الأولى للمتوالية المرجعية reference sequence مع القراءة المعينة المناظرة؛

ج) محاذاة جزء من المنطقة المجنحة flanking region الثانية للمتوالية المرجعية reference sequence مع القراءة المعينة المناظرة؛ و

د) تحديد طول و/أو متوالية القطاع التكراري repetitive segment.

13. مُحلل العينات sample analyzer وفقا لعنصر الحماية 12، حيث تتضمن محاذاة الجزء من المنطقة المجنحة flanking region في واحد أو كل من الخطوات (ب) و(ج) على ما يلي:

(i) تحديد موضع المنطقة المجنحة flanking region المحفوظة المناظرة على القراءة المعينة باستخدام توافق وحدة k محددة لمنطقة بذرية seeding region تتداخل مع أو تجاور القطاع التكراري repetitive segment؛ و

(ii) محاذاة المنطقة المجنحة flanking region مع القراءة المعينة.

14. مُحلل العينات وفقا لعنصر الحماية 13، حيث المنطقة البذرية seeding region تتضمن منطقة عالية التعقيد high-complexity region للمنطقة المجنحة flanking region المحفوظة، حيث تشتمل المنطقة عالية التعبيق high-complexity region على كل من أربعة نيوكليوتيدات

nucleotides عند تكرار يكون على الأقل 10٪، 15٪، 20 ٪، أو 25٪ من كل النيوكليوتيدات nucleotides في المنطقة المجنحة المحافظة conserved flanking region.

15. مُحلل العينات sample analyzer وفقا لعنصر الحماية 14، حيث المنطقة عالية التعقيد high-complexity region تتضمن متوالية مميزة بشكل مناسب عن القطاع التكراري repetitive segment لتجنب سوء المحاذاة.

16. مُحلل العينات sample analyzer وفقا لعنصر الحماية 14، حيث المنطقة عالية التعقيد high-complexity region تتضمن متوالية لها خليط متعدد الأشكال diverse mixture لنيكلوتيدات nucleotides.

17. مُحلل العينات sample analyzer وفقا لعنصر الحماية 13، حيث لا تشتمل المنطقة البذرية seeding region مناطق منخفضة التعقد low-complexity regions من المنطقة المجنحة flanking region المحفوظة المناظرة، حيث أن المناطق منخفضة التعقد low-complexity regions المذكورة تشتمل على متوالية بها أكثر من 30٪، 40٪، 50٪، 60٪، 70٪ أو أكثر من 80٪ من الهوية الخاصة بالمتوالية بالنسبة للجزء المتكرر repetitive segment.

18. مُحلل العينات sample analyzer وفقا لعنصر الحماية 17، حيث المناطق منخفضة التعقد low-complexity regions لها متواليات تقوم أساسا بتجميع مجموعة من بواعث التكرار repeat motifs.

19. مُحلل العينات sample analyzer وفقا لعنصر الحماية 13، حيث المنطقة البذرية seeding region مجاورة مباشرة للقطاع التكراري repetitive segment.

20. مُحلل العينات sample analyzer وفقا لعنصر الحماية 13، حيث المنطقة البذرية seeding region تتضمن جزء من القطاع التكراري repetitive segment.

21. مُحلل العينات sample analyzer وفقا لعنصر الحماية 13، حيث المنطقة البذرية seeding region مزاحة من القطاع التكراري repetitive segment.

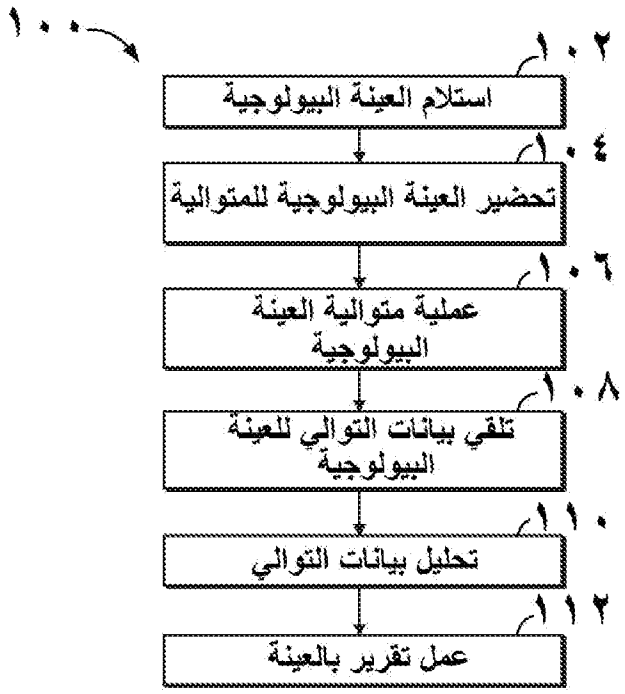
22. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث قراءات العينة عبارة عن نواتج تكبير amplicons لتفاعل البلمرة المتسلسل Polymerase chain reaction (PCR) لها متواليات أولية primer sequences متجهة إلى الأمام والخلف.

23. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث تعيين قراءات العينة إلى المواقع المعينة تتضمن تحديد متواليات لقراءات العينة التي تناظر متواليات أولية primer sequences لنواتج تكبير amplicons لتفاعل البلمرة المتسلسل Polymerase chain reaction.

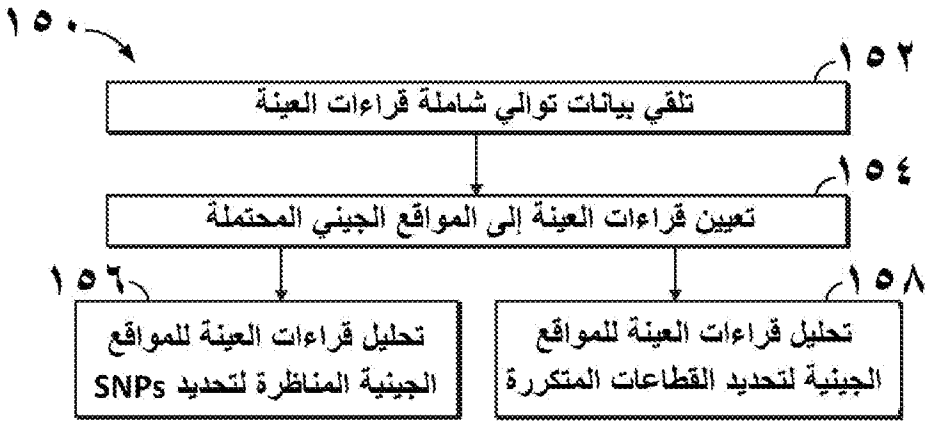
24. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث بيانات التوالي من فحص توالي بواسطة التخليق sequencing-by-synthesis (SBS).

25. مُحلل العينات sample analyzer وفقا لعنصر الحماية 1 أو عنصر الحماية 2، حيث المنطقة محل الاهتمام region of interest عبارة عن تكرار ترادفي قصير short tandem repeat (STR).

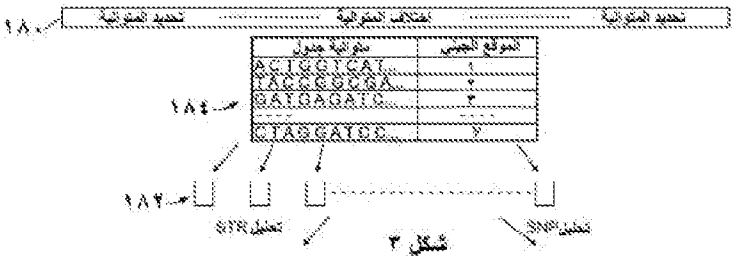
26. مُحلل العينات sample analyzer وفقا لعنصر الحماية 25، حيث يتم اختيار تكرار ترادفي قصير short tandem repeat من واحد على الأقل من مواضع تكرار ترادفي قصير short tandem repeat لصبغة جسدية ذاتية من CODIS، ومواضع CODIS Y-STR و EU و مواضع تكرار ترادفي قصير short tandem repeat للصبغة الجسدية أو مواضع EU Y-STR.


١٠٠
١٠٢
استلام العينة البيولوجية
١٠٤
تحضير العينة البيولوجية للمتوالية
١٠٦
عملية متوالية العينة البيولوجية
١٠٨
تلقي بيانات التوالي للعينة البيولوجية
١١٠
تحليل بيانات التوالي
١١٢
عمل تقرير بالعينة

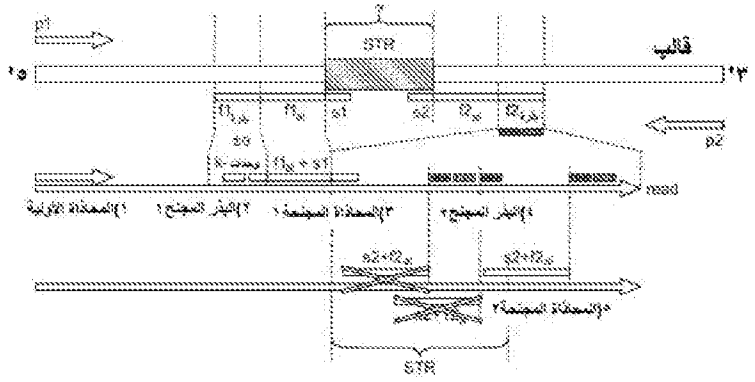
شكل ١


١٥٠
١٥٢
تلقي بيانات توالي شاملة قراءات العينة
١٥٤
تعيين قراءات العينة إلى المواقع الجيني المحتملة
١٥٦
تحليل قراءات العينة للمواقع الجينية المناظرة لتحديد SNPs
١٥٨
تحليل قراءات العينة للمواقع الجينية لتحديد القطاعات المتكررة

شكل ٢



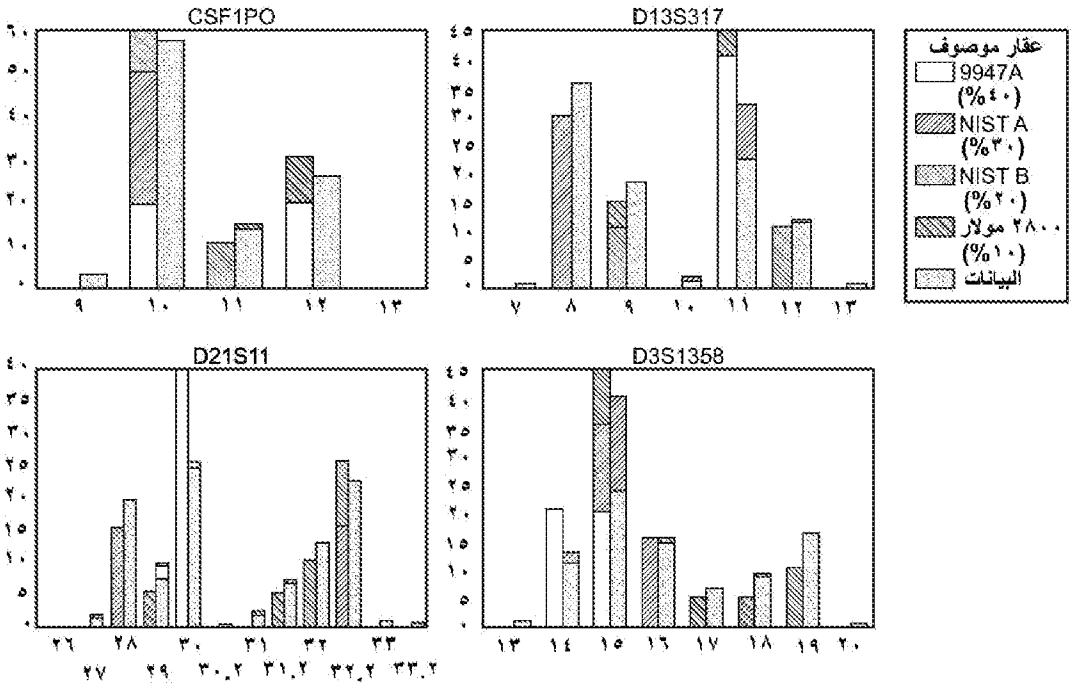
شكل ٣



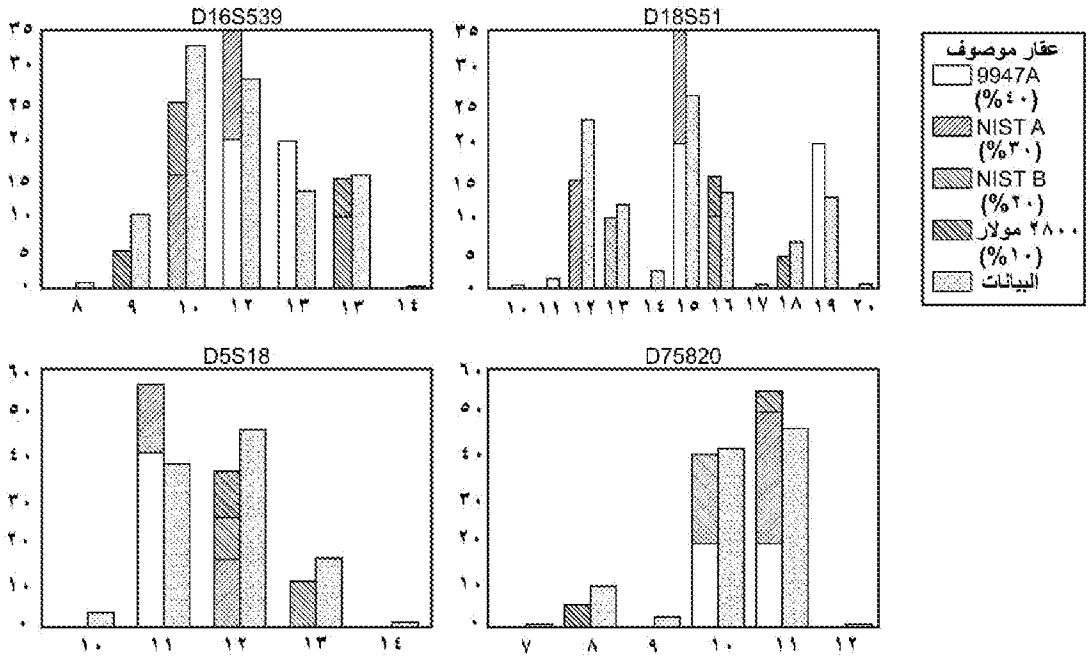
شكل ٤

يجب أن يكون قصيرا جدا لتتم تسميته AGAAAGAAAGAAAGAAAGAAA... :قراءة
طرف مجاميع: GAAAGAAAGAGAAAAAGAAAAGAAATAGTAGCAACTGTTAT...
قراءة: ...AGAAAGAAAGAAAGAAAGAGAAAAAGAGAGGAAAGAAAGAGAAAAAGAAAAGAAATAGTAG...
طرف مجاميع٢: GAAAGAAAGAGAAAAAGAAAAGAAATAGTAGCAAC...
قراءة: ...AGAAAGAAAAAGAGAGAGGAAAGAAAGAAAAAAAGAAAAAAAATAGTAG...
طرف مجاميع٢: GAAAGAAAGAGAAAAAGAAAAGAAATA...

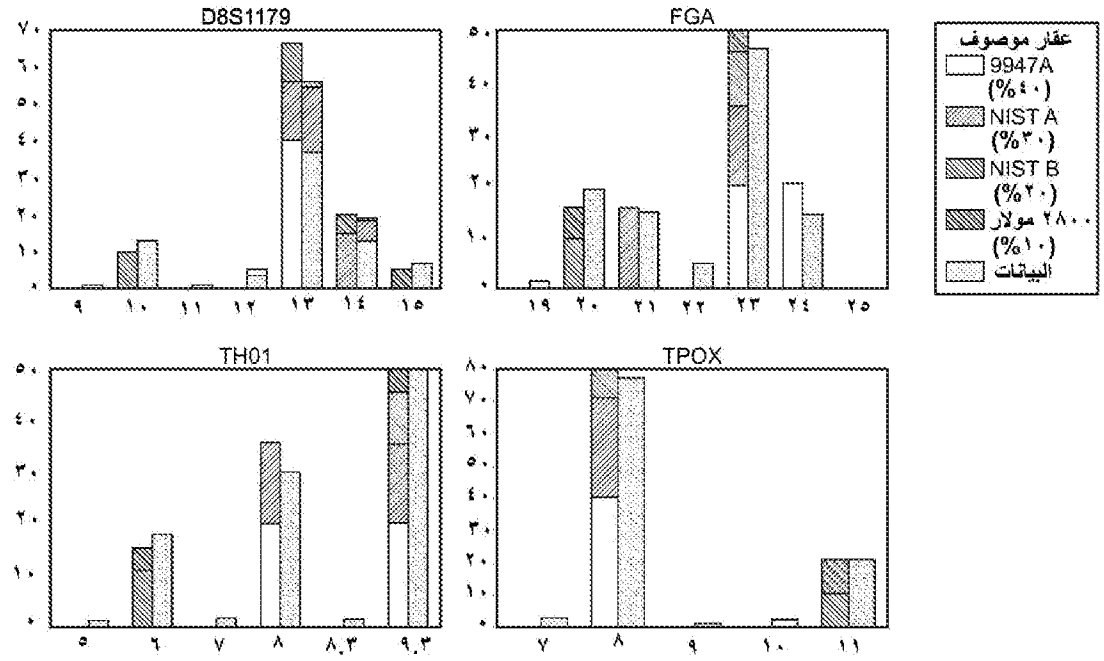
شكل ٥


CSF1PO
D13S317
D21S11
D3S1358
عقار موصوف
9947A
(%٤٠)
NIST A
(%٣٠)
NIST B
(%٢٠)
٢٨٠٠ مولار
(%١٠)
البيانات

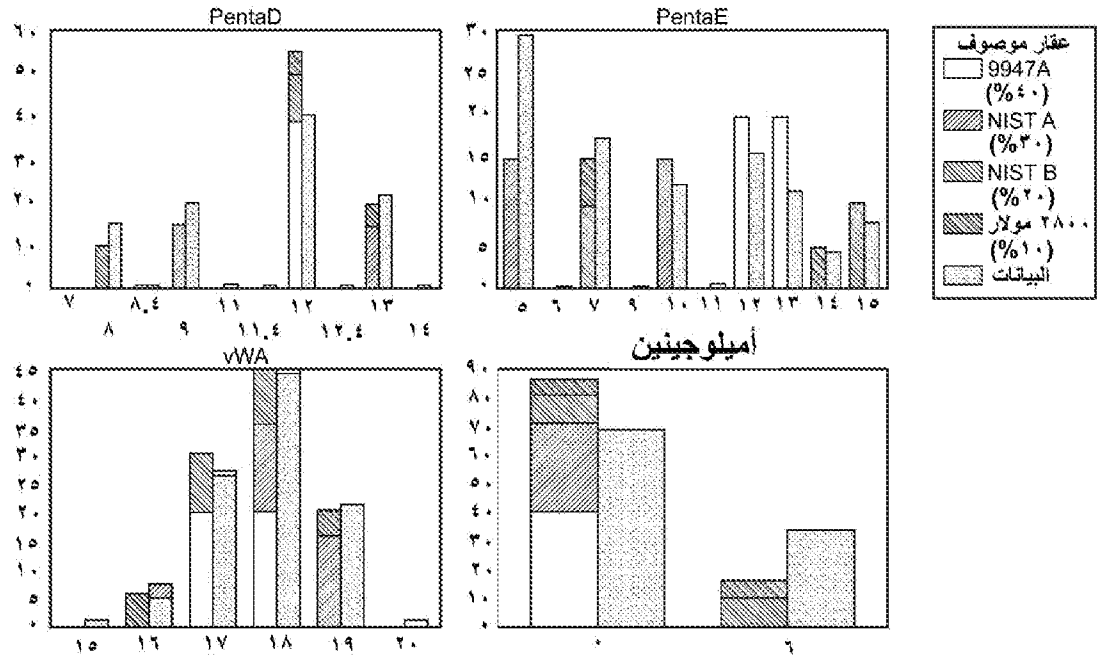
شكل ٦أ


D16S539
D18S51
D5S18
D75820
عقار موصوف
9947A (%٤٠)
NIST A (%٣٠)
NIST B (%٢٠)
٢٨٠٠ مولار (%١٠)
البيانات

شكل ٦ب


D8S1179
FGA
TH01
TPOX
عقار موصوف
9947A
(%٤٠)
NIST A
(%٣٠)
NIST B
(%٢٠)
٢٨٠٠ مولار
(%١٠)
البيانات

شكل ٦جـ


PentaD
PentaE
vWA
أميلوجينين
عقار موصوف
9947A (%٤٠)
NIST A (%٣٠)
NIST B (%٢٠)
٢٨٠٠ مولار (%١٠)
البيانات

شكل ٦د

عينة DNA	CSF1PO	D351358	D7S820	D16S539	D18S51	FGA	PentaE	TH01
9947A	١٠،١٢	١٤،١٥	١٠،١١	١١،١٢	١٥،١٩	٢٣،٢٤	١٢،١٣	٨،٩،٣
2800M	١٢	١٧،١٨	٨،١١	٩،١٣	١٦،١٨	٢٠،٢٣	٧،١٤	٦،٩،٣
NIST A	١٠	١٥،١٦	١١	١٠،١١	١٢،١٥	٢١،٢٣	٥،١٠	٨،٩،٣
NIST B	١٠،١١	١٥،١٩	١٠	١٠،١٣	١٣،١٦	٢٠،٢٣	٧،١٥	٦،٩،٣
NIST C	١٠،١٢	١٦،١٨	١٠،١٢	١٠	١٦،١٩	٢٤،٢٦	١٢،١٣	٦،٨

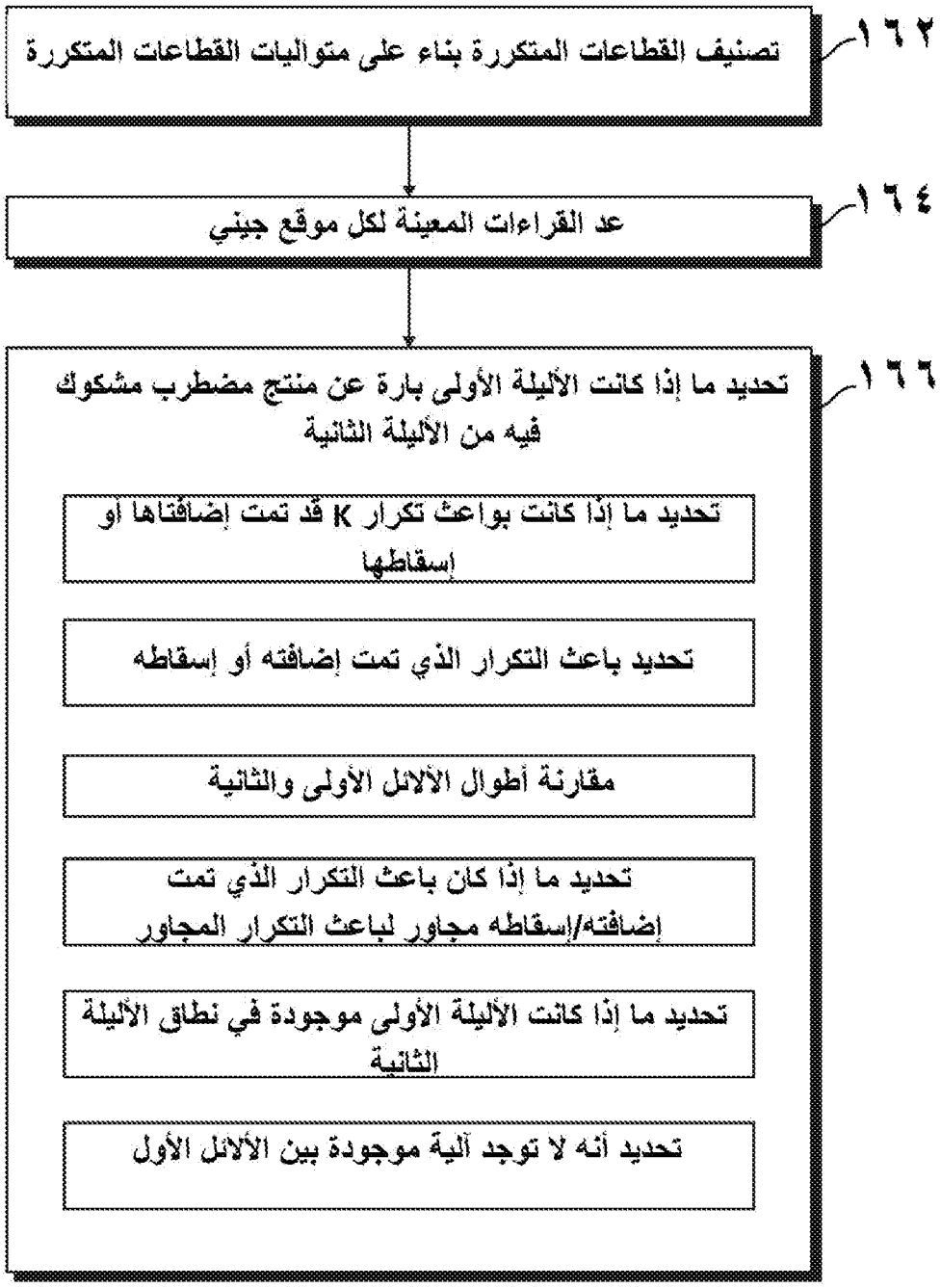
vWA	D5S818	D8S1179	D13S317	D21S11	PentaD	TPOX
١٩،٢٠	١١	١٣،١٣	١١	٣٠	١٢	٨
١٦،١٩	١٢	١٤،١٥	٩،١١	٢٩،٣١،٢	١٢،١٣	١١
١٨،١٩	١١،١٢	١٣،١٤	٨	٢٨،٣٢،٢	٩،١٣	٨
١٧،١٨	١٢،١٣	١٠،١٣	٩،١٢	٣٢،٣٢،٢	٨،١٢	٨،١١
١٦،١٨	١٠،١١	١٠،١٧	١١	٢٩،٣٠	١٠،١١	١١

A

A

[TCTA]11[TCTG]1[TCTA]1[TCTA]٣،٥٦%[TCTA] ٤٦% :SNP المكتشف في نصف التكرارات

شكل ٧

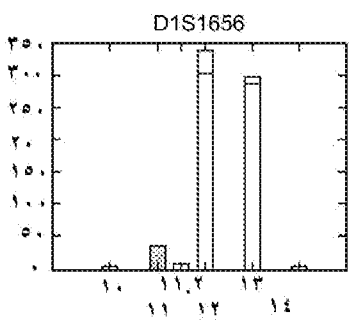

١٦٢
تصنيف القطاعات المتكررة بناء على متواليات القطاعات المتكررة
١٦٤
عد القراءات المعينة لكل موقع جيني
١٦٦
تحديد ما إذا كانت الأليلة الأولى بارة عن منتج مضطرب مشكوك فيه من الأليلة الثانية
تحديد ما إذا كانت بواعث تكرار K قد تمت إضافتها أو إسقاطها
تحديد باعث التكرار الذي تمت إضافته أو إسقاطه
مقارنة أطوال الألائل الأولى والثانية
تحديد ما إذا كان باعث التكرار الذي تمت إضافته/إسقاطه مجاور لباعث التكرار المجاور
تحديد ما إذا كانت الأليلة الأولى موجودة في نطاق الأليلة الثانية
تحديد أنه لا توجد آلية موجودة بين الألائل الأول

شكل ٨

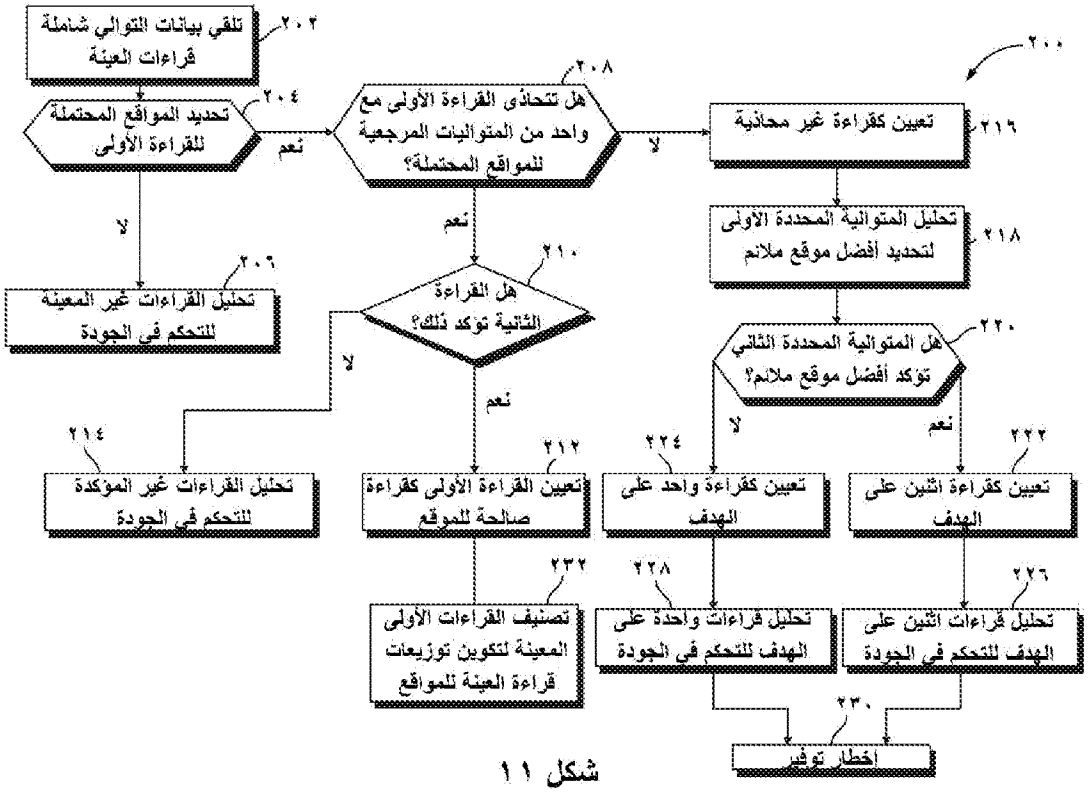
١٩٠

قراءة الأعداد	طول المتوالية	المتوالية	اسم آلية CE	الآلية المحتملة
٢٨٧(%٩٦.٦٣)	٦٢	[TAGA]13[TG]5	١٣	١
٤(%١.٣٥)	٦٢	[TAGA]12[TAGG]1[TG]5	١٣	٢
٣٠٣(%٨٩.٣٨)	٥٨	[TAGA]11[TAGG]1[TG]5	١٢	٣
٢٨(%٨.٢٦)	٥٨	[TAGA]12[TG]5	١٢	٤

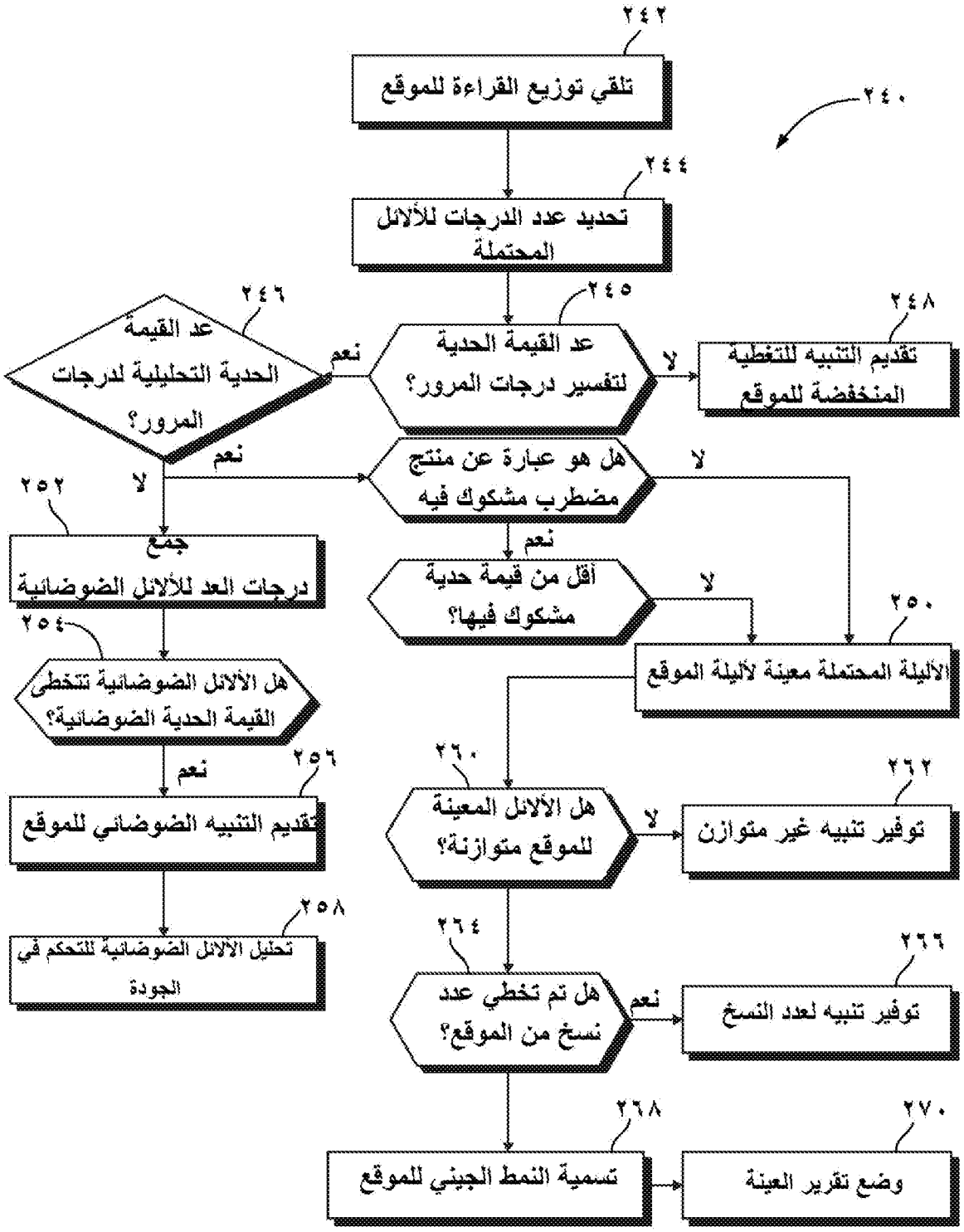
شكل ٩



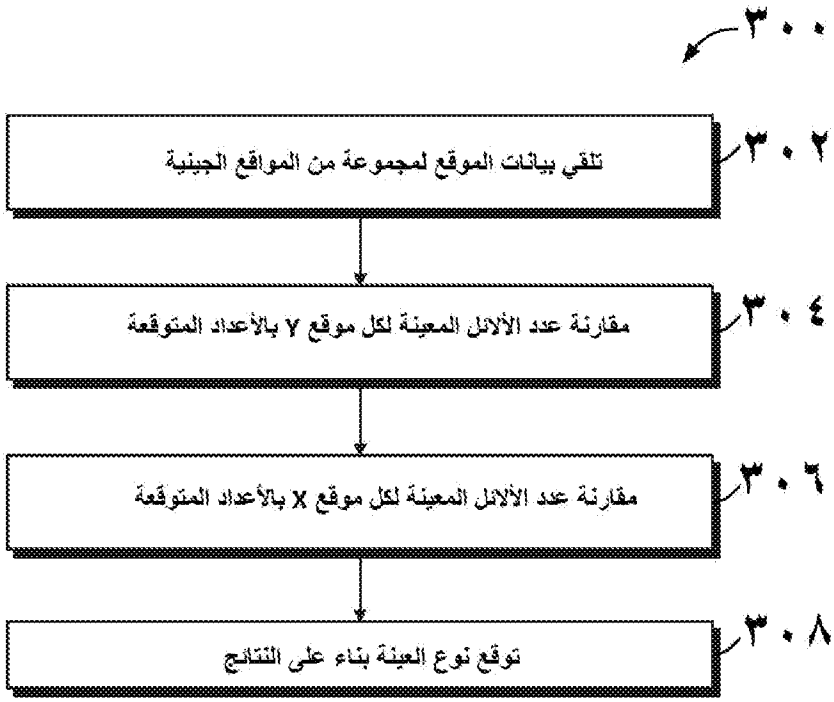
شكل ١٠


٢٠٠
٢٠٢
تلقي بيانات التوالي شاملة قراءات العينة
٢٠٤
تحديد المواقع المحتملة للقراءة الأولى
٢٠٨
هل تتحاذى القراءة الأولى مع واحد من المتواليات المرجعية للمواقع المحتملة؟
٢١٦
تعيين كقراءة غير محاذية
٢١٨
تحليل المتوالية المحددة الأولى لتحديد أفضل موقع ملائم
٢٢٠
هل المتوالية المحددة الثاني تؤكد أفضل موقع ملائم؟
٢٠٦
تحليل القراءات غير المعينة للتحكم في الجودة
٢١٠
هل القراءة الثانية تؤكد ذلك؟
٢١٤
تحليل القراءات غير المؤكدة للتحكم في الجودة
٢١٢
تعيين القراءة الأولى كقراءة صالحة للموقع
٢٢٤
تعيين كقراءة واحد على الهدف
٢٢٢
تعيين كقراءة اثنين على الهدف
٢٣٢
تصنيف القراءات الأولى المعينة لتكوين توزيعات قراءة العينة للمواقع
٢٢٨
تحليل قراءات واحدة على الهدف للتحكم في الجودة
٢٢٦
تحليل قراءات اثنين على الهدف للتحكم في الجودة
٢٣٠
اخطار توفير
نعم
لا
نعم
لا
نعم
لا
نعم
لا

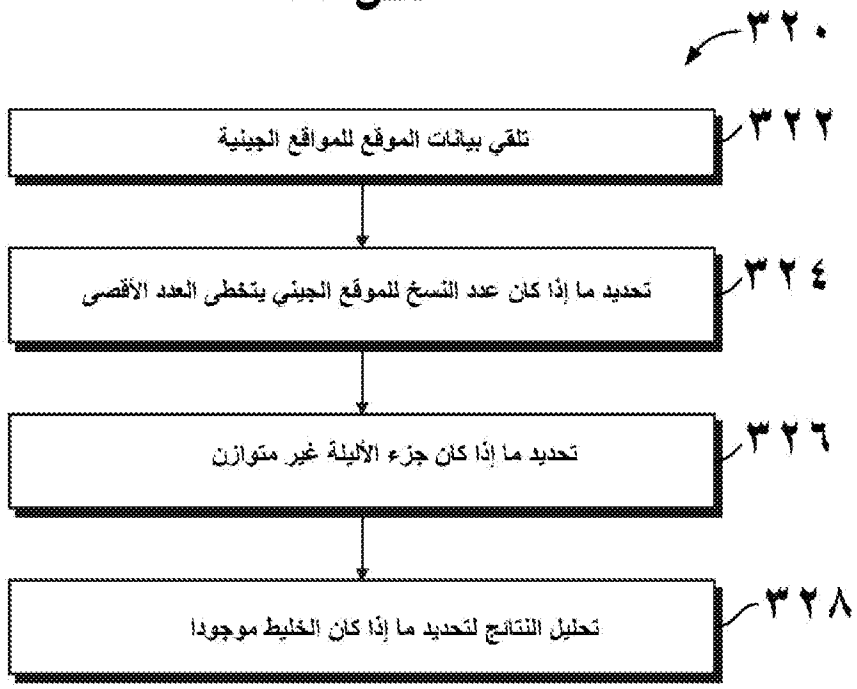
شكل ١١


٢٤٠
٢٤٢
تلقي توزيع القراءة للموقع
٢٤٤
تحديد عدد الدرجات للألائل المحتملة
٢٤٥
عد القيمة الحدية لتفسير درجات المرور؟
٢٤٦
عد القيمة الحدية التحليلية لدرجات المرور؟
٢٤٨
تقديم التنبيه للتغطية المنخفضة للموقع
نعم
لا
هل هو عبارة عن منتج مضطرب مشكوك فيه
نعم
لا
أقل من قيمة حدية مشكوك فيها؟
نعم
لا
٢٥٠
الأليلة المحتملة معينة لأليلة الموقع
٢٥٢
جمع درجات العد للألائل الضوضائية
٢٥٤
هل الألائل الضوضائية تتخطى القيمة الحدية الضوضائية؟
نعم
٢٥٦
تقديم التنبيه الضوضائي للموقع
٢٥٨
تحليل الألائل الضوضائية للتحكم في الجودة
٢٦٠
هل الألائل المعينة للموقع متوازنة؟
لا
٢٦٢
توفير تنبيه غير متوازن
٢٦٤
هل تم تخطي عدد نسخ من الموقع؟
نعم
٢٦٦
توفير تنبيه لعدد النسخ
٢٦٨
تسمية النمط الجيني للموقع
٢٧٠
وضع تقرير العينة

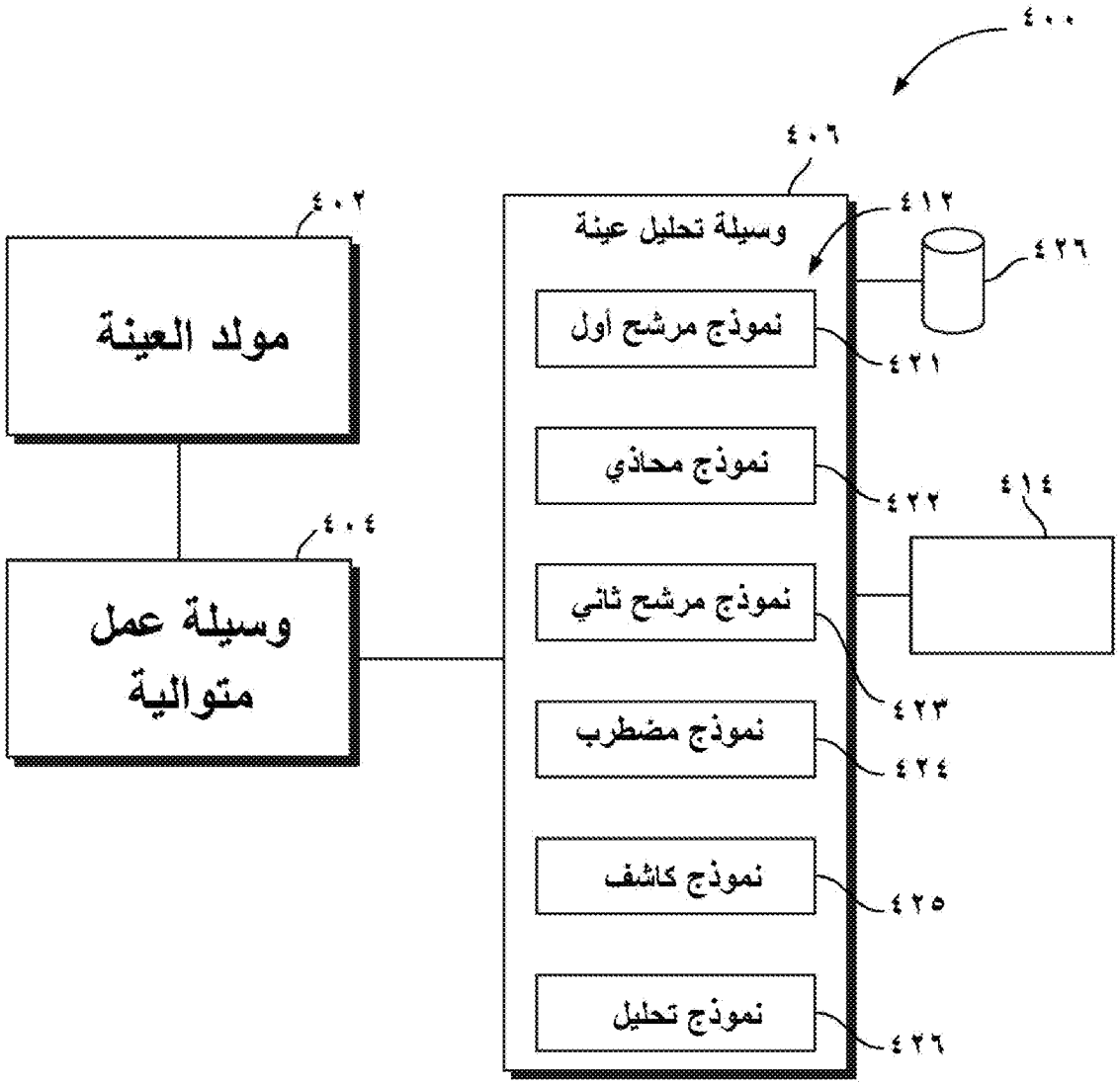
شكل ١٢


٣٠٠
٣٠٢
تلقي بيانات الموقع لمجموعة من المواقع الجينية
٣٠٤
مقارنة عدد الألائل المعينة لكل موقع Y بالأعداد المتوقعة
٣٠٦
مقارنة عدد الألائل المعينة لكل موقع X بالأعداد المتوقعة
٣٠٨
توقع نوع العينة بناء على النتائج

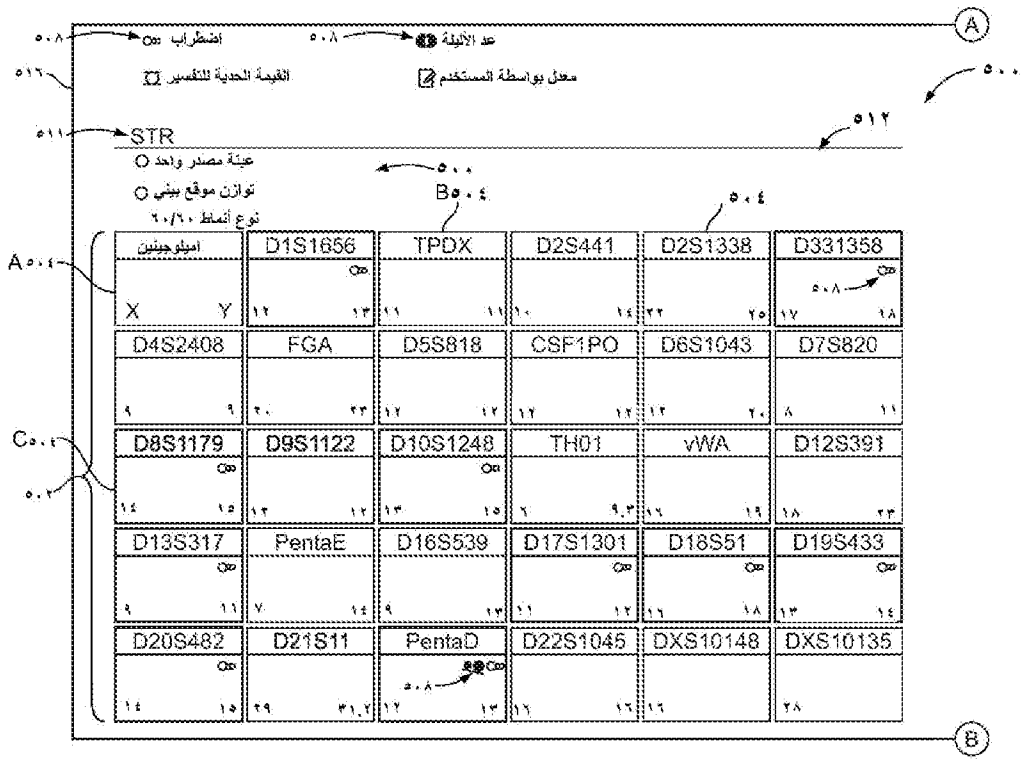
شكل ١٣


٣٢٠
٣٢٢
تلقي بيانات الموقع للمواقع الجينية
٣٢٤
تحديد ما إذا كان عدد النسخ للموقع الجيني يتخطى العدد الأقصى
٣٢٦
تحديد ما إذا كان جزء الأليلة غير متوازن
٣٢٨
تحليل النتائج لتحديد ما إذا كان الخليط موجودا

شكل ١٤

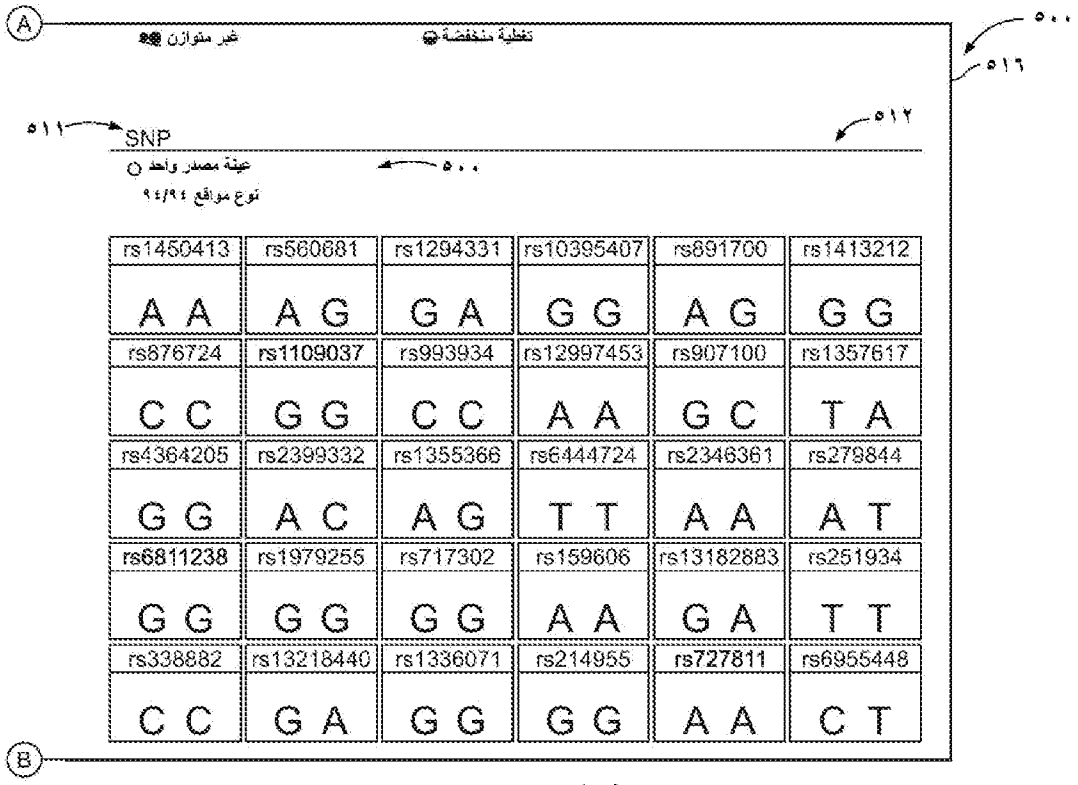

٤٠٠
٤٠٦
٤٠٢
مولد العينة
وسيلة تحليل عينة
٤١٢
٤٢٦
نموذج مرشح أول
٤٢١
نموذج محاذي
٤٢٢
٤١٤
٤٠٤
وسيلة عمل متوالية
نموذج مرشح ثاني
٤٢٣
نموذج مضطرب
٤٢٤
نموذج كاشف
٤٢٥
نموذج تحليل
٤٢٦

شكل ١٥



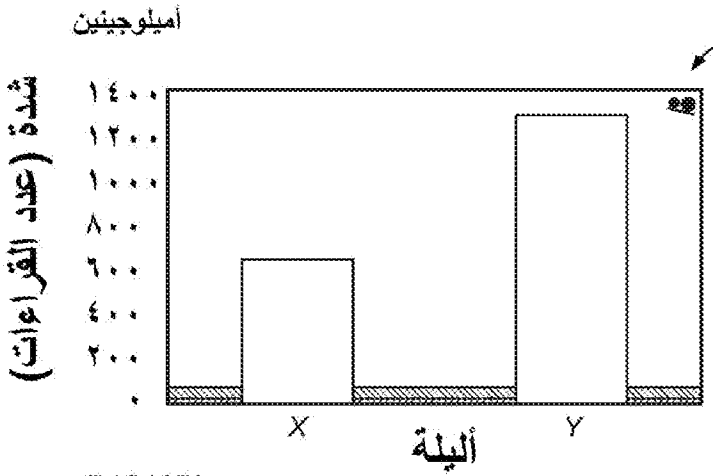
اميلوجينين	D1S1656	TPOX	D2S441	D2S1338	D331358
X Y	١٢ ١٣	١١ ١١	١٠ ١٤	٢٢ ٢٥	١٧ ١٨
D4S2408	FGA	D5S818	CSF1PO	D6S1043	D7S820
٩ ٩	٢٠ ٢٣	١٢ ١٢	١٢ ١٢	١٢ ٢٠	٨ ١١
D8S1179	D9S1122	D10S1248	TH01	vWA	D12S391
١٤ ١٥	١٢ ١٢	١٣ ١٥	٦ ٩,٣	١٦ ١٩	١٨ ٢٣
D13S317	PentaE	D16S539	D17S1301	D18S51	D19S433
٩ ١١	٧ ١٤	٩ ١٣	١١ ١٢	١٦ ١٨	١٣ ١٤
D20S482	D21S11	PentaD	D22S1045	DXS10148	DXS10135
١٤ ١٥	٢٩ ٣١,٢	١٢ ١٣	١١ ١٦	١٦	٢٨

شكل ١٦أ

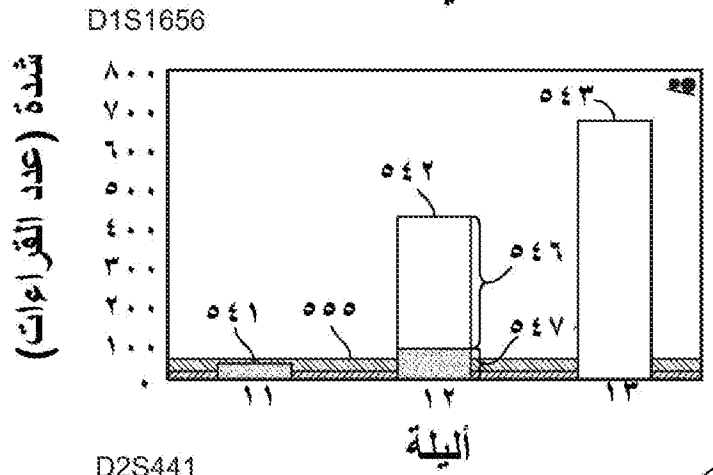


rs1450413	rs560681	rs1294331	rs10395407	rs691700	rs1413212
A A	A G	G A	G G	A G	G G
rs876724	rs1109037	rs993934	rs12997453	rs907100	rs1357617
C C	G G	C C	A A	G C	T A
rs4364205	rs2399332	rs1355366	rs6444724	rs2346361	rs279844
G G	A C	A G	T T	A A	A T
rs6811238	rs1979255	rs717302	rs159606	rs13182883	rs251934
G G	G G	G G	A A	G A	T T
rs338882	rs13218440	rs1336071	rs214955	rs727811	rs6955448
C C	G A	G G	G G	A A	C T

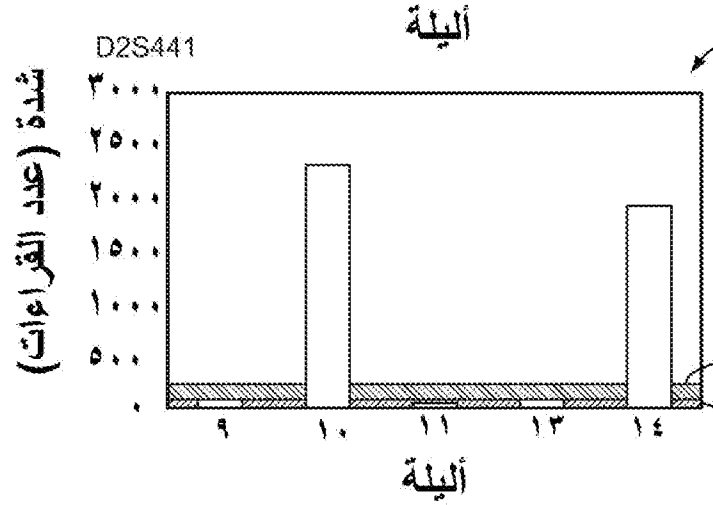
شكل ١٦ب



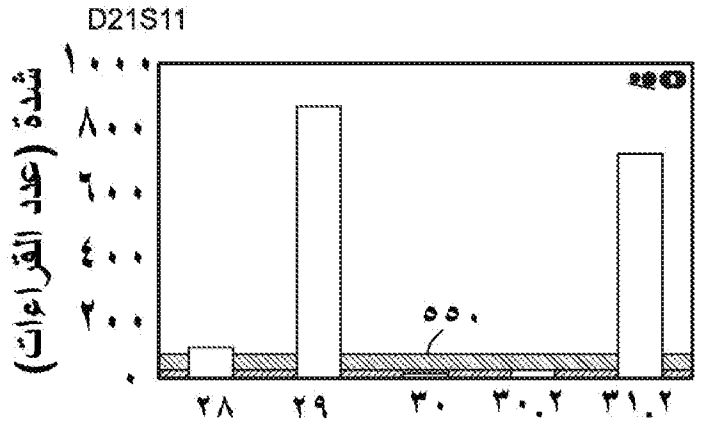
شكل ١٧أ



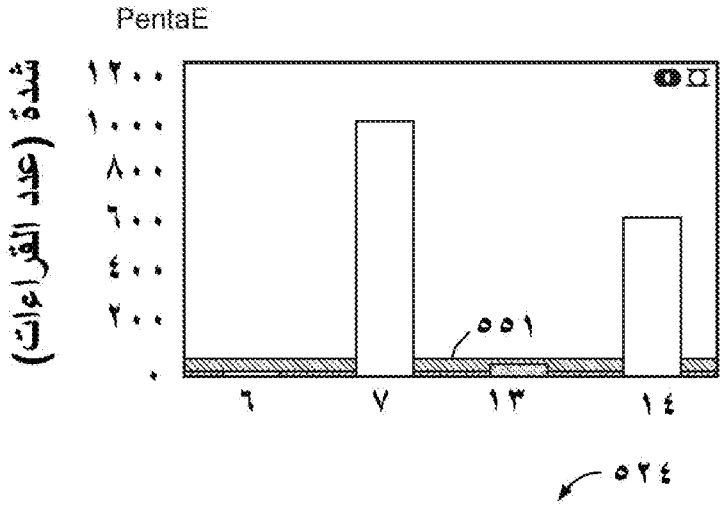
شكل ١٧ب



شكل ١٧جـ



شكل ١٧د



شكل ١٧هـ

شكل ١٧و

مدة سريان هذه البراءة عشرون سنة من تاريخ إيداع الطلب

وذلك بشـرط تسـديد المقابل المالي السـنوي للبراءة وعدم بطلانها أو سـقوطها لمخالفتها لأي من أحكام نظام براءات الاختراع والتصميمات التخطيطية للدارات المتكاملة والأصناف النباتية والنماذج الصناعية أو لائحته التنفيذية.

صادرة عن

الهيئة السعودية للملكية الفكرية

ص ب ٦٥٣١ ، الرياض ١٣٣٢١ ، المملكة العربية السعودية

SAIP@SAIP.GOV.SA