

مدير التشغيل المحلي وحدة تحليل متغير السلالة الجرثومية دليل سير عمل جهاز NextSeq 550Dx

لأغراض الاستخدام في التشخيص المختبري فقط

3	نظرة عامة
3	إدخال معلومات التشغيل
5	طرق التحليل
6	عرض عملية التشغيل وبيانات العينة
7	تقرير التحليل
8	ملفات مخرجات التحليل
14	سجل المراجعة
15	المساعدة الفنية



هذا المستند ومحتوياته مملوك لشركة Illumina, Inc، والشركات التابعة لها ("Illumina"). ويهدف إلى الاستخدام التعاقدى لعملائها فقط فيما يتعلق باستخدام المنتج (المنتجات) الموضح هنا وليس لأي غرض آخر. يجب ألا يتم استخدام هذا المستند ومحتوياته أو توزيعه لأي غرض آخر و/أو إرساله، أو الكشف عنه، أو نسخه بأي شكل آخر دون موافقة خطية مسبقة من شركة Illumina. لا تُقدم شركة Illumina أي تراخيص تتعلق ببراءات الاختراع، أو العلامات التجارية أو حقوق التأليف والنشر، أو حقوق القانون العام ولا الحقوق المماثلة لأي أطراف أخرى بموجب هذا المستند.

يجب على الموظفين المؤهلين والمدربين بشكل جيد اتباع التعليمات الواردة في هذا المستند بشكل صارم وصريح من أجل ضمان الاستخدام السليم والأمن للمنتج (المنتجات) الموضحة به. يجب قراءة جميع محتويات هذا المستند وفهمها بشكل كامل قبل استخدام هذا المنتج (هذه المنتجات).

وقد يؤدي عدم قراءة التعليمات الواردة هنا بشكل كامل واتباعها بوضوح إلى حدوث تلف في المنتج (المنتجات)، أو إصابة للأشخاص، بما في ذلك المستخدم أو أشخاص آخرين، والحاق الضرر بممتلكات أخرى، وستفقد أي ضمان ينطبق على المنتج (المنتجات).

لا تتحمل شركة ILLUMINA أي مسؤولية ناجمة عن سوء استخدام المنتج (المنتجات) الموضح هنا (بما في ذلك البرامج أو أجزاء منها).

حقوق الطبع والنشر © لشركة Illumina, Inc 2021، جميع الحقوق محفوظة.

جميع العلامات التجارية مملوكة لشركة أو Illumina, Inc، أصحابها المعنيين. للحصول على معلومات محددة حول العلامات التجارية، راجع www.illumina.com/company/legal.html

نظرة عامة

إن وحدة تحليل مدير التشغيل المحلي Germline Variant مخصصة للاستخدام مع اختبار Illumina TruSeq Custom Amplicon Kit Dx وجهاز NextSeq 550Dx. وعند استخدام الاختبار مع وحدة Germline Variant Module، فإنه يكون مخصصًا لتحضير المكتبات المستخدمة لتسلسل الحمض النووي (DNA) من عينات الدم المحيطي الكامل.

وتقوم وحدة التحليل بتقييم مناطق قصيرة من الحمض النووي (DNA) المضخم، أو الأمبليكونات، بحثًا عن المتغيرات. كما يتيح التسلسل المرکز للأمبليكون تغطية عالية لمناطق محددة. انظر نشرة العبوة الخاصة بمجموعة *TruSeq Custom Amplicon Kit Dx* (المستند رقم 1000000029772).

تتطلب وحدة تحليل Germline Variant مستلزمات تسلسل تبلغ 300 دورة. لمزيد من المعلومات، انظر نشرة العبوة الخاصة بـ *NextSeq 550Dx High-Output Reagent Kit v2.5* أو *NextSeq 550Dx High Output Reagent Kit v2*.

حول هذا الدليل

يوفر هذا الدليل تعليمات بشأن إعدادات التشغيل للتسلسل والتحليل لوحدة تحليل متغير السلالة الجرثومية. وللحصول على معلومات حول لوحة معلومات مدير التشغيل المحلي وإعدادات النظام، انظر *الدليل المرجعي لجهاز NextSeq 550Dx* (مستند رقم 1000000009513).

عرض مدير التشغيل المحلي

تُعرض واجهة مدير التشغيل المحلي داخل تطبيق برنامج تشغيل NextSeq 550Dx (NOS) أو عبر أي متصفح ويب. متصفح الويب المدعوم هو Chromium.

ملاحظة



إذا كنت تستخدم متصفحًا غير مدعوم، فعليك تنزيل المتصفح المدعوم عندما يُطلب منك ذلك في رسالة "Confirm Unsupported Browser" ("تأكيد متصفح غير مدعوم"). اختر "هنا" لتنزيل إصدار Chromium المدعوم.

عرض على شاشة مراقبة الجهاز

- 1 لعرض واجهة مدير التشغيل المحلي على شاشة مراقبة الجهاز، حدد أحد الخيارات التالية:
 - ◀ من الشاشة الرئيسية لتطبيق NOS، حدد **Local Run Manager** (مدير التشغيل المحلي).
 - انقر فوق علامة "X" الموجودة أعلى الزاوية اليمنى للعودة إلى تطبيق NOS عند الانتهاء.
 - ◀ حدد أيقونة تصغير NOS، ثم افتح متصفح ويب Chromium على الجهاز، وأدخل <http://localhost> في شريط العنوان. يمكن للمستخدمين المسؤولين فقط تصغير NOS.

العرض من جهاز كمبيوتر متصل بالشبكة

- 1 افتح متصفح ويب Chromium على جهاز كمبيوتر متصل بالشبكة نفسها المتصل بها الجهاز، واتصل باستخدام عنوان IP الخاص بالجهاز أو اسم الجهاز. على سبيل المثال، <http://myinstrument>.

إدخال معلومات التشغيل

تعيين المعلومات

- 1 سجّل الدخول إلى مدير التشغيل المحلي.
- 2 حدد **Create Run** (إنشاء عملية تشغيل)، وحدد **Germline Variant**.
- 3 أدخل اسم عملية التشغيل التي تُحدد عملية التشغيل بدايةً من التسلسل حتى التحليل. استخدم الأحرف الأبجدية الرقمية، أو المسافات، أو الشرطات السفلية، أو الشرطات.
- 4 [اختياري] أدخل وصفاً للتشغيل للمساعدة في تحديد هوية التشغيل. استخدم الأحرف الأبجدية الرقمية، أو المسافات، أو الشرطات السفلية، أو الشرطات.

- 5 حدد عدد العينات ومجموعة المؤشرات من القائمة المنسدلة.
راع المعلومات التالية عند إجراء التحديد.
- ◀ تتضمن القائمة المنسدلة أعداد العينات مع مجموعة المؤشرات. على سبيل المثال، يشير الخيار Set 1-24 إلى 24 عينة سيتم اختبارها، مع استخدام مؤشرات من مجموعة المؤشرات 1.
- ◀ تشير أرقام مجموعات المؤشرات إلى مجموعات مختلفة من مؤشرات i5. توفر كل من Set 1 (المجموعة 1) و Set 2 (المجموعة 2) توعاً للمؤشرات. يتم توفير مجموعتين من المؤشرات للمساعدة في منع استنفاد مجموعة واحدة.
- ◀ اختر عدد العينات الأقرب لعدد العينات التي تختبرها. وإذا لم يكن العدد الدقيق للعينات موجوداً في القائمة، فحدد العدد الأقرب ولكن الأقل من عدد العينات التي تختبرها. على سبيل المثال، إذا كنت تريد اختبار 18 عينة، فحدد 16 عينة.
- ◀ يتم تمييز آبار العينات ومجموعات المؤشرات التي تستوفي متطلبات تنوع المؤشرات باللون الأخضر. وإذا حددت آباراً وتوليفات مؤشرات أخرى، فستلقى إشعاراً عند حفظ التشغيل إذا لم يتم استيفاء متطلبات تنوع المؤشرات.

استيراد ملفات البيان لعملية التشغيل

- 1 تأكد من أن ملفات البيان التي ترغب في استيرادها متاحة في أحد مواقع الشبكة التي يُمكن الوصول إليها أو على محرك USB.
- 2 حدد **Import Manifests** (استيراد ملفات البيان).
- 3 انتقل إلى ملف البيان، وحدد ملفات البيان التي تريد إضافتها.

ملاحظة



لجعل ملفات البيان متاحة لكافة عمليات التشغيل التي تستخدم وحدة تحليل متغير السلالة الجراثومية، أضف ملفات البيان باستخدام ميزة Module Settings (إعدادات الوحدة). تتطلب هذه الميزة أذونات المسؤول على مستوى المستخدم. لمزيد من المعلومات، راجع *الدليل المرجعي لجهاز NextSeq 550Dx (مستند رقم 1000000009513)*.

تحديد العينات لعملية التشغيل

- حدد العينات لعملية التشغيل باستخدام أحد الخيارات والتوجيهات التالية.
 - ◀ **إدخال العينات يدوياً** - استخدم الجدول الفارغ الموجود على شاشة Create Run (إنشاء عملية التشغيل).
 - ◀ **استيراد العينات** - انتقل إلى ملف خارجي بتنسيق (CSV.*) مع قيم مفصولة بفاصلة. يتوفر قالب للتنزيل على شاشة Create Run (إنشاء عملية تشغيل).
- وبعد تعبئة جدول العينات، يمكنك تصدير معلومات العينة إلى ملف خارجي. استخدم الملف كمرجع عند تحضير المكتبات أو استورد الملف لإجراء تشغيل آخر.

إدخال العينات يدوياً

- 1 أدخل مُعرّف العينة الفريد في حقل Sample ID (مُعرّف العينة).
- 2 **[اختياري]** بالنسبة للعينات ذات عنصر التحكم الإيجابي أو السلبي، انقر بزر الماوس الأيمن على نوع التحكم وحدده.
- 3 **[اختياري]** أدخل وصف العينة في حقل Sample Description (وصف العينة).
- استخدم الأحرف الأبجدية الرقمية، أو الشرطات أو الشرط السفلية.
- تُعدّ أوصاف العينة مرتبطة باسم العينة. وتُستبدل أوصاف العينة في حال أُعيد استخدام اسم العينة نفسه في عملية تشغيل لاحقة.
- 4 حدد محوّل المؤشر 1 من القائمة المنسدلة للمؤشر 1 (i7).
- وعند استخدامك مجمعات العينة المقترحة، سيعمل البرنامج على التعبئة التلقائية لمحولات المؤشر i7 و i5 التي تُلبي متطلبات التنوع للمؤشر. وإذا لم يكن العدد الدقيق للعينات التي تختبرها مُدرجاً في القائمة، فتأكد من تحديد محولات المؤشر للمجموعات الإضافية. و إذا كنت بحاجة إلى تحديد المؤشرات للمجمعات الإضافية، أو لا تستخدم المجموعات المدمجة الموصى بها لمحوّل المؤشر، فتأكد قبل اختيارك للمؤشرات، من قراءة *الاستدعاء الأساسي وتنوع المؤشر في صفحة 13*.
- 5 حدد محوّل المؤشر 2 من القائمة المنسدلة للمؤشر 2 (i5).

6 حدد أحد ملفات البيان من القائمة المنسدلة لملفات البيان.

7 اختر أحد الخيارات لعرض مخطط اللوحة، أو طباعته أو حفظه كمرجع لإعداد المكتبات:

◀ حدد رمز **Print** (طباعة) لعرض مخطط اللوحة. حدد **Print** (طباعة) لطباعة مخطط اللوحة.

◀ حدد **Export** (تصدير) لتصدير معلومات العينة إلى ملف خارجي.

تأكد من أن معلومات البيان والعينة صحيحة. قد تؤثر المعلومات غير الصحيحة على النتائج.

8 حدد **Save Run** (حفظ عملية التشغيل).

استيراد العينات

1 حدد **Import Samples** (استيراد العينات) واستعرض للوصول إلى موقع ملف معلومات العينة. هناك نوعان من الملفات التي يمكنك استيرادها.

◀ حدد **Template** (قالب) من شاشة **Create Run** (إنشاء تشغيل) لإنشاء تخطيط لوحة جديد. ويحتوي ملف القالب على رؤوس الأعمدة الصحيحة للاستيراد. أدخل معلومات العينة في كل عمود للعينات الموجودة في التشغيل. احذف المعلومات التجريبية في الخلايا غير المستخدمة، ثم احفظ الملف.

◀ استخدم ملفًا لمعلومات العينة تم تصديره من وحدة متغير السلالة الجراثومية باستخدام ميزة **Export** (تصدير).

2 حدد رمز **Print** (طباعة) لعرض مخطط اللوحة.

3 حدد **Print** (طباعة) لطباعة مخطط اللوحة كمرجع لإعداد المكتبات.

4 [اختياري] حدد **Export** (تصدير) لتصدير معلومات العينة إلى ملف خارجي.

تأكد من أن معلومات البيان والعينة صحيحة. قد تؤثر المعلومات غير الصحيحة على النتائج.

5 حدد **Save Run** (حفظ عملية التشغيل).

تحرير عملية تشغيل

للحصول على تعليمات بشأن تحرير المعلومات في التشغيل قبل إجراء التسلسل، انظر *الدليل المرجعي لجهاز NextSeq 550Dx* (مستند رقم 1000000009513).

طرق التحليل

تقوم وحدة تحليل متغير السلالة الجراثومية بتنفيذ خطوات التحليل التالية، ثم تكثف ملفات مخرجات في مجلد **Alignment**.

◀ فك تعدد إرسال قراءات المؤشر

◀ إنشاء ملفات FASTQ

◀ المحاذاة إلى مرجع

◀ تحديد المتغيرات

نظام موزع البيانات

يُعارن موزع البيانات كل تسلسل لقراءة المؤشر بتسلسلات المؤشر المُحددة لعملية التشغيل. ولا يتم النظر في قيم الجودة في هذه الخطوة.

تُحدّد قراءات المؤشر باستخدام الخطوات التالية:

◀ تُرَقم العينات بدءًا من 1 حسب الترتيب المُدرّج بها لعملية التشغيل.

◀ يُعدّ رقم العينة 0 مخصصًا للعناقيد التي لم تُعيّن إلى إحدى العينات.

◀ تُعيّن العناقيد إلى إحدى العينات عندما يتطابق تسلسل المؤشر تمامًا أو عندما يكون هناك ما يصل إلى حالة عدم تطابق واحدة لكل قراءة مؤشر.

إنشاء ملف FASTQ

بعد توزيع البيانات، يُنشئ البرنامج ملفات التحليل الوسيطة بتنسيق FASTQ، وهو تنسيق نصي يُستخدم لتمثيل عمليات التسلسل. تتضمن ملفات FASTQ قراءات لكل عينة ودرجات الجودة المرتبطة بها. وتُستثنى العناقيد التي لم تمر من الفلتر. يحتوي كل ملف من ملفات FASTQ على قراءات لعينة واحدة فقط، ويضمن اسم تلك العينة في اسم ملف FASTQ. تُعدّ ملفات FASTQ بمثابة الإدخال الأولي للمحاذاة. يتم إنشاء ثمانية ملفات FASTQ لكل عينة، بواقع أربعة ملفات من القراءة 1 (Read 1) وأربعة ملفات من القراءة 2 (Read 2).

المحاذاة

خلال خطوة المحاذاة، تعمل خوارزمية سميث ووترمان ذات التنسيق الخاص على محاذاة العناقيد من كل عينة مقابل تسلسلات الأمبليكون المُحددة في ملف البيان.

وتجري خوارزمية Smith-Waterman الشريطية عمليات محاذاة تسلسل شبه عالمية لتحديد المناطق المتشابهة بين تسلسلين. وبدلاً من مقارنة التسلسل الإجمالي، تُقارن خوارزمية سميث ووترمان مقاطع من جميع الأطوال الممكنة.

يجري تقييم كل قراءة ذات نهاية مُقترنة فيما يتعلّق بالمحاذاة الخاصة بها مع تسلسلات المسبار ذات الصلة بهذه القراءة.

◀ يجري تقييم القراءة رقم 1 مقابل المتمم المعكوس الخاص بقليل النيوكليوتيد ذي الموقع الكروموسومي المحدد المتماشي مع الاتجاه (DLSO).

◀ يجري تقييم القراءة رقم 2 مقابل قليل النيوكليوتيد ذي الموقع الكروموسومي المحدد عكس الاتجاه (ULSO).

◀ إذا كانت بداية القراءة تُطابق تسلسل المسبار بما لا يتجاوز ثلاثة اختلافات (حالات عدم التطابق أو التحولات نتيجة الإنديل الأساسي)، تتم محاذاة الطول الكامل للقراءة مقابل الأمبليكون المستهدف لهذا التسلسل.

◀ لا تتم ملاحظة الإنديل الموجود في نطاق DLSO وULSO بسبب كيمياء الفحص.

تتم تصفية عمليات المحاذاة من نتائج المحاذاة حسب معدلات عدم التطابق على منطقة الاهتمام أو الأمبليكون بالكامل، وفقاً لطول الأمبليكون. وتُكتب عمليات المحاذاة التي تمت تصفيتها في ملفات المحاذاة بوصفها "غير محاذاة" كما أنها ليست مستخدمة في استدعاء المتغير.

استدعاء المتغير

تقوم أداة Pisces Variant Caller، التي طورتها شركة Illumina، بتحديد المتغيرات الموجودة في عينة الحمض النووي (DNA).

تتعرف أداة Pisces Variant Caller على تبديلات النيوكليوتيدات المفردة (SNVs)، وتبديلات النيوكليوتيدات المتعددة (MNVs)، وطفرة الإدخال والحذف (indels) الصغيرة في 3 خطوات:

◀ تراعي كل موضع في الجينوم المرجعي بشكل منفصل

◀ تحسب القواعد عند الموضع المحدد للقراءات التي تمت محاذاتها والتي تتداخل مع ذلك الموضع

◀ تحسب درجة سجل جودة المتغير التي تقيس جودة الاستدعاء باستخدام نموذج بواسون (Poisson model). وتُستثنى المتغيرات التي تقل درجة جودة سجلها عن Q20.

وإذا اجتاز المتغير كافة المرشحات، يتم وضع علامة PASS (اجتياز) عليه في ملف بتنسيق (VCF).

لمزيد من المعلومات، راجع github.com/Illumina/Pisces/wiki.

عرض عملية التشغيل وبيانات العينة

1 من لوحة معلومات مدير التشغيل المحلي، انقر فوق اسم التشغيل.

2 من علامة التبويب Run Overview (نظرة عامة على التشغيل)، راجع مقاييس تشغيل التسلسل.

3 [اختياري] انقر فوق أيقونة **Copy to Clipboard** (نسخ إلى الحافظة) لنسخ مسار مجلد مخرجات التشغيل.

4 انقر فوق علامة التبويب Sequencing Information (معلومات التسلسل) لمراجعة معلمات التشغيل ومعلومات المستلزمات.

5 انقر فوق علامة التبويب Samples and Results (العينات والنتائج) لعرض موقع تقرير التحليل.

◀ إذا تكرر التحليل، فقم بتوسيع القائمة المنسدلة Select Analysis (تحديد التحليل) وحدد التحليل المناسب.

6 انقر فوق أيقونة **Copy to Clipboard** (نسخ إلى الحافظة) لنسخ مسار مجلد Analysis (التحليل).

لمزيد من المعلومات حول علامتي التبويب Run Overview (نظرة عامة على التشغيل) و Sequencing Information (معلومات التسلسل)، وكيفية إعادة طلب التحليل، راجع الدليل المرجعي لجهاز NextSeq 550Dx (مستند رقم 1000000009513).

تقرير التحليل

يتم تلخيص نتائج التحليل في علامة التبويب Samples and Results (العينات والنتائج) وكتقرير مجمع في مجلد المحاذاة. يتوفر تقرير لكل عينة أيضاً بتنسيق ملف PDF لكل عينة.

العينات علامة تبويب النتائج

1 انقر فوق عينة في القائمة لعرض تقرير العينة.

الجدول 1 معلومات التشغيل والعينة

عنوان العمود	الوصف
حالة التشغيل	يذكر ما إذا كان تشغيل التسلسل قد نجح أم فشل.
النتائج الإجمالي (جيبايت)	عدد القواعد التي تم استدعاؤها في تشغيل التسلسل. يوضح حد المرور، وحالة النجاح أو الفشل.
Q30 ≤ %	النسبة المئوية للقراءات في تشغيل التسلسل التي سجلت درجة جودة تبلغ 30 (Q30) أو أكثر. يوضح حد المرور، وحالة النجاح أو الفشل.
مُعرّف العينة	مُعرّف العينة الذي تم تقديمه عند إنشاء التشغيل.
إجمالي قراءات مرشح تعيير القراءات (PF)	إجمالي عدد القراءات التي مرت عبر المرشح.
نسبة Read 1 (القراءة 1) Q30 ≤	النسبة المئوية للقراءات في Read 1 (القراءة 1) التي سجلت درجة جودة تبلغ 30 (Q30) أو أكثر للعينة.
نسبة Read 2 (القراءة 2) Q30 ≤	النسبة المئوية للقراءات في Read 2 (القراءة 2) التي سجلت درجة جودة تبلغ 30 (Q30) أو أكثر للعينة.
معدل استدعاء الكروموسومات الجسدية	عدد المواقع الجينومية عبر الكروموسومات الجسدية (الكروموسومات من 1 إلى 22) التي تستوفي حد قيمة الثقة المحدد مسبقاً، مقسوماً على إجمالي عدد المواقع الجينومية الجسدية التي تم فحصها. ويتم وصف معدل الاستدعاء على أساس كل عينة ويُذكر في شكل نسبة مئوية يتم حسابها كـ 1 ناقص (عدد المواقع الجسدية ذات الاستدعاءات غير المكتملة مقسوماً على إجمالي عدد المواقع الجسدية التي تم تسلسلها).

الجدول 2 نموذج تقرير معلومات

عنوان العمود	الوصف
العينة	مُعرّف العينة الذي تم تقديمه عند إنشاء التشغيل.
تاريخ التقرير	تاريخ إنشاء التقرير.
معلومات العينة	مُعرّف العينة الذي تم تقديمه عند إنشاء التشغيل، وإجمالي القراءات التي مرت من الفلتر في العينة، والنسبة المئوية للقراءات الخاصة بالعينة التي سجلت درجة جودة تبلغ 30 (Q30) أو أكثر، ومعدل استدعاء الكروموسومات الجسدية.
ملخص الأمبليكون	إجمالي عدد مناطق الأمبليكون التي تم تسلسلها، وإجمالي الطول بالأزواج القاعدية للأمبليكونات التي تم تسلسلها في المناطق المستهدفة، للعينة وملف البيان. ويحدد ملف البيان الجينوم المرجعي والمناطق المرجعية المستهدفة المستخدمة في خطوة المحاذاة.
إحصائيات مستوى القراءة	عدد ونسبة القراءات للعينة التي تغطي كل موضع في المرجع، لكل من Read 1 (القراءة 1) و Read 2 (القراءة 2).
ملخص المتغيرات	عدد تبادلات النيوكليوتيدات المفردة (SNVs)، وعمليات الإدخال، وعمليات الحذف المكتشفة للعينة التي اجتازت القيم المقترحة لتحديد ما إذا كانت نتائج الجودة ضمن نطاق مقبول.
ملخص التغطية	إجمالي عدد القواعد التي تمت محادتها مقسوماً على حجم المنطقة المستهدفة، والنسبة المئوية لمناطق الأمبليكون ذات قيم تغطية أكبر من حد التغطية المنخفض البالغ 0.2 * متوسط تغطية الأمبليكون، للعينة.
مخططات التغطية	يوضح مخطط التغطية حسب منطقة الأمبليكون (Coverage by Amplicon Region) التغطية عبر مناطق الأمبليكون للعينة. ويتم تمييز المناطق التي تقل قيم تغطيتها عن حد التغطية باللون الأحمر. ويُشار إلى متوسط كافة القيم بخط برتقالي.
إصدارات البرامج	إصدارات البرامج عند تسلسل العينة. وتتضمن برنامج تشغيل NextSeq 550Dx (NOS)، وبرنامج Local Run Manager، وبرنامج RTA، وإصدار Germline Variant Module.

ملفات مخرجات التحليل

يتم إنشاء ملفات مخرجات التحليل التالية لوحدت تحليل متغير السلالة الجرثومية، وهي توفر نتائج التحليل للمحاذاة واستدعاء المتغيرات. وتوجد ملفات مخرجات التحليل في مجلد Alignment.

اسم الملف	الوصف
فك تعدد الإرسال (*.txt)	ملفات بسيطة تحتوي على نتائج ملخص فك تعدد الإرسال.
FASTQ (*.fastq.gz)	ملفات بسيطة تحتوي على استدعاءات قاعدية ذات سجل درجات جودة. تُعدّ ملفات FASTQ بمثابة الإدخال الأولي لخطوة المحاذاة.
ملفات المحاذاة بتنسيق BAM (*.bam)	تحتوي على القراءات التي تمت محاذاتها لعينة محددة.
ملفات استدعاء المتغيرات بتنسيق VCF الخاص بالجينوم (*.genome.vcf.gz)	تحتوي على النمط الجيني لكل موضع، سواء تم استدعاؤه كمتغير أو استدعاؤه كمرجع.
ملفات استدعاء المتغيرات بتنسيق VCF (*.vcf.gz)	تحتوي على كافة المتغيرات التي تم استدعاؤها عبر المنطقة المستهدفة.
AmpliconCoverage_M1.tsv	يحتوي على معلومات حول التغطية لكل أمبليكون لكل عينة لكل ملف بيان مُقدم. ويمثل الرمز #M رقم ملف البيان.

تنسيق ملف فك تعدد الإرسال

تقرأ عملية فك تعدد الإرسال تسلسل المؤشر المرفق بكل عنقود لتحديد العينة التي نشأ منها العنقود. وتُكتب عملية المطابقة بين العناقيد ورقم العينة في ملف فك تعدد إرسال (demux.*) لكل شريحة من شرائح خلية التدفق.

تنسيق تسمية ملف فك تعدد الإرسال هو s_1_X.demux، حيث يمثل الرمز X رقم الشريحة.

وتبدأ ملفات فك تعدد الإرسال برأس يتكون مما يلي:

◀ الإصدار (عدد صحيح بطول 4 بايت)، حاليًا 1

◀ عدد العناقيد (عدد صحيح بطول 4 بايت)

ويتكون باقي الملف من أرقام العينات لكل عنقود من الشريحة.

وعند اكتمال خطوة فك تعدد الإرسال، يُنشئ البرنامج ملف فك تعدد إرسال باسم DemultiplexSummaryF1L1.txt.

◀ في اسم الملف، يمثل الرمز F1 رقم خلية التدفق.

◀ في اسم الملف، يمثل الرمز L1 رقم المر.

◀ تؤدي عملية فك تعدد الإرسال إلى جدول به صف واحد لكل شريحة وعمود واحد لكل عينة، بما في ذلك العينة 0.

◀ التسلسلات الأكثر شيوعًا في قراءات المؤشر.

تنسيق ملف FASTQ

FASTQ هو تنسيق ملف نصي يحتوي على استدعاءات قاعدية وقيم الجودة لكل قراءة. ويحتوي كل سجل على 4 أسطر:

◀ المعرف

◀ التسلسل

◀ علامة زائد (+)

◀ سجلات جودة Phred بتنسيق مشفر بترميز ASCII + 33

يتم تنسيق المعرف على النحو التالي:

@Instrument:RunID:FlowCellID:Lane:Tile:X:Y ReadNum:FilterFlag:0:SampleNumber

مثال:

```
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAAAAA9#:<#<;<<<???)#=#
```

تنسيق ملف BAM

يُعد ملف BAM (الذي يحمل الامتداد *.bam) النسخة الثنائية المضغوطة من ملف SAM التي تُستخدم لتمثيل التسلسلات التي تمت محادتها والتي تصل سعتها إلى 128 ميجابايت. وقد تم وصف تنسيقات SAM وBAM بالتفصيل في الموقع الإلكتروني [.samtools.github.io/hts-specs/SAMv1.pdf](https://samtools.github.io/hts-specs/SAMv1.pdf).

تستخدم ملفات BAM تنسيق تسمية الملفات **SampleName_S#.bam**؛ حيث يمثل الرمز # رقم العينة المحدد بناءً على الترتيب الذي تم إدراج العينات به للتشغيل.

وتحتوي ملفات BAM على قسم العنوان (header) وقسم المحاذاة (alignment):

◀ **العنوان** - يحتوي على معلومات حول الملف بأكمله، مثل اسم العينة، وطول العينة، وطريقة المحاذاة. وترتبط عمليات المحاذاة الموجودة في قسم المحاذاة بمعلومات محددة في قسم العنوان.

◀ **المحاذاة** - تحتوي على اسم القراءة، وتسلسل القراءة، وجودة القراءة، ومعلومات المحاذاة، والعلامات المخصصة. ويتضمن اسم القراءة الكروموسوم، وإحداثي البدء، وجودة المحاذاة، وسلسلة واصف المطابقة.

يتضمن قسم المحاذاة المعلومات التالية لكل قراءة أو زوج قراءات:

◀ **AS**: جودة المحاذاة ذات النهاية المقترنة

◀ **RG**: مجموعة القراءة، والتي تشير إلى عدد القراءات لعينة محددة.

◀ **BC**: علامة الرمز الشريطي، والتي تشير إلى معرف العينة الذي تم فك تعدد إرساله والمقترن بالقراءة.

◀ **SM**: جودة المحاذاة ذات النهاية المفردة.

◀ **XC**: سلسلة واصف المطابقة

◀ **XN**: علامة اسم الأمبليكون، التي تسجل معرف الأمبليكون المقترن بالقراءة

توفر ملفات فهرس BAM (التي تحمل الامتداد *.bam.bai) فهرسًا لملف BAM المقابل.

تنسيق ملف VCF

يُعد تنسيق استدعاء المتغيرات (VCF) تنسيق ملف شائعًا تم تطويره بواسطة المجتمع العلمي لعلم الجينوم. ويحتوي على معلومات حول المتغيرات الموجودة في مواقع محددة في جينوم مرجعي. وتنتهي ملفات VCF باللاحقة .vcf. ويحتوي على معلومات حول المتغيرات الموجودة في مواقع محددة في جينوم مرجعي. وتنتهي ملفات VCF باللاحقة .vcf.

يتضمن عنوان ملف VCF إصدار تنسيق ملف VCF وإصدار أداة استدعاء المتغيرات، ويسرد الشرح التوضيحية المستخدمة في الجزء المتبقي من الملف. ويتضمن عنوان تنسيق استدعاء المتغيرات (VCF) أيضًا ملف الجينوم المرجعي وملف BAM. يحتوي السطر الأخير في العنوان على رؤوس الأعمدة لخطوط البيانات. ويحتوي كل خط من خطوط بيانات ملف VCF على معلومات حول متغير واحد.

عناوين ملف VCF

العنوان	الوصف
CHROM	الكروموسوم الخاص بالجينوم المرجعي. تظهر الكروموسومات بنفس ترتيب ملف FASTQ المرجعي.
POS	موقع القاعدة الواحدة للمتغير في الكروموسوم المرجعي. بالنسبة لتعدد أشكال النيوكليوتيدات المفردة (SNPs)، يكون هذا الموقع هو القاعدة المرجعية مع المتغير؛ وبالنسبة لطفرات الإدخال والحذف أو عمليات الحذف، يكون هذا الموقع هو القاعدة المرجعية التي تسبق المتغير مباشرةً.
ID	رقم rs للمتغير الذي تم الحصول عليه من dbSNP.txt، إذا كان ذلك منطبقاً. في حالة وجود أرقام rs متعددة في هذا الموقع، يتم الفصل بين القائمة بفاصلة منقوطة. في حالة عدم وجود إدخال dbSNP في هذا الموقع، يتم استخدام علامة قيمة مفقودة ('.').
REF	النمط الجيني المرجعي. على سبيل المثال، يتم تمثيل حذف حرف T واحد كمرجع TT وحرف T بديل. ويتم تمثيل متغير النيوكليوتيد المفرد من A إلى T كمرجع A وحرف T بديل.
ALT	الآليات التي تختلف عن القراءة المرجعية. على سبيل المثال، يتم تمثيل إدخال حرف T واحد كمرجع A وAT بديل. يتم تمثيل متغير النيوكليوتيد المفرد من A إلى T على أنه A المرجعي وT البديل.
QUAL	درجة جودة مُقاسة بمقياس Phred يتم تحديدها بواسطة مُستدعي المتغيرات. تشير السجلات الأعلى إلى أعلى ثقة أعلى في المتغير واحتمالية أقل لوقوع أخطاء. بالنسبة لسجل جودة بقيمة Q، تكون الاحتمالية المقدرة للخطأ هي $10^{-Q/10}$. على سبيل المثال، تحتوي مجموعة استدعاءات Q30 على معدل خطأ بنسبة 0.1%. تقوم العديد من أدوات استدعاء المتغيرات بتعيين سجلات الجودة بناءً على نماذجها الإحصائية، والتي تكون عالية بالنسبة لمعدل الخطأ الملحوظ.

تعليقات ملفات VCF

العنوان	الوصف
الترشيح	<p>إذا تم اجتياز جميع عمليات الترشيح، يتم كتابة PASS (اجتياز) في عمود المرشح.</p> <ul style="list-style-type: none"> LowDP - يُطبق على المواقع التي يقل عمق تغطيتها عن 150 ضعفًا. وبالنسبة لمواقع الأمبليكون التي تغطيها كل من القراءة الأمامية والعكسية، فإن هذا يعادل 300 قراءة ذات نهاية مقترنة متداخلة. q20 - سجل درجة الجودة > 20. MultiAllelicSite - المتغير لا يتوافق مع النموذج ثنائي الصيغة الصغرية. R5x9 - عدد التكرارات المتجاورة (بطول يتراوح من 1 إلى 5 أزواج قاعدية) لاستدعاءات المتغير ≤ 9. SB - انحياز الشريط أكبر من الحد المحدد.
المعلومات	<p>تتضمن الإدخالات المحتملة في عمود المعلومات ما يلي:</p> <ul style="list-style-type: none"> AC - عدد الأليلات في الأنماط الجينية لكل الأليل بديل (ALT)، بنفس الترتيب المدرج. AF - تردد الأليل لكل الأليل بديل (ALT)، بنفس الترتيب المدرج. AN - العدد الإجمالي للأليلات في الأنماط الجينية التي تم استدعاؤها. CD - علامة تشير إلى أن تعدد أشكال النيوكليوتيدات المفردة (SNP) يحدث داخل منطقة التشفير لإدخال واحد على الأقل في RefGene. DP - العمق (عدد استدعاءات القواعد المحاذية لموقع ما والمستخدمة في الاستدعاء المتغير). Exon - قائمة مفصلة بفاصلة لمناطق الإكسون (Exon) المقروءة من RefGene. FC - العاقبة الوظيفية. GI - قائمة مفصلة بفاصلة لمعرفات الجينات المقروءة من RefGene. QD - ثقة المتغير/الجودة حسب العمق. TI - قائمة مفصلة بفاصلة لمعرفات الجينات المقروءة من RefGene.
التنسيق	<p>يسرد عمود التنسيق الحقول مفصلة بنقطتين رأسيين. على سبيل المثال، GT:GQ. وتعتمد قائمة الحقول المقدمة على أداة استدعاء المتغير المستخدمة. تشمل الحقول المتاحة ما يلي:</p> <ul style="list-style-type: none"> AD - إدخال في شكل X,Y، حيث يمثل X عدد استدعاءات المرجع، ويمثل Y عدد استدعاءات البديل. DP - يعقب القراءة التقريبي؛ حيث يتم ترشيح القراءات ذات MQ=255 أو ذات القرناء السينيين. GQ - جودة النمط الجيني. GT - النمط الجيني. يمثل الرقم 0 القاعدة المرجعية، بينما يمثل الرقم 1 أول إدخال في عمود البديل (ALT)، وهكذا. وتشير الشرطة المائلة للأمام (/) إلى عدم توفر معلومات الطور (Phasing). NC - جزء القواعد التي لم يتم استدعاؤها، أو التي تقل جودة استدعاء قاعدتها عن الحد الأدنى. NL - مستوى الضوضاء؛ وهو تقدير لضوضاء استدعاء القواعد في هذا الموقع. SB - انحياز الشريط في هذا الموقع. تشير القيم السالبة الأكبر إلى انحياز أقل، بينما تشير القيم القريبة من 0 إلى انحياز أكبر. VF - تردد المتغير؛ وهو النسبة المئوية للقراءات التي تدعم الأليل البديل.
العينة	يعطي عمود العينة القيم المحددة في عمود التنسيق (FORMAT).

ملفات VCF الجينومية

ملفات VCF الجينومية (gVCF) هي ملفات بتنسيق VCF الإصدار 4.1 تتبع مجموعة من الاصطلاحات لتمثيل جميع المواقع داخل الجينوم بتنسيق مضغوط بدرجة معقولة. وتتضمن ملفات (*.genome.vcf.gz) (gVCF) جميع المواقع داخل المنطقة ذات الأهمية في ملف واحد لكل عينة.

يُظهر ملف gVCF حالات عدم الاستدعاء (No-calls) في المواقع التي لا تجتاز جميع عوامل التصفية. وتشير علامة النمط الجيني (GT) /. إلى حالة عدم استدعاء (No-call).

لمزيد من المعلومات، راجع sites.google.com/site/gvcftools/home/about-gvcf.

ملف تغطية الأمبليكون

يتم إنشاء ملف تغطية أمبليكون لكل ملف بيان. يمثل الرمز M# في اسم الملف رقم ملف البيان.

يتضمن كل ملف صف عنوان يحتوي على معرفات العينات المرتبطة بملف البيان. ويحتوي الملف على المعلومات التالية.

◀ معرف الهدف كما هو مدرج في ملف البيان.

◀ عمق تغطية القراءات التي تجتاز المرشح.

ملفات المخرجات التكميلية

توفر ملفات المخرجات التكميلية التالية معلومات إضافية، أو تلخص نتائج عملية التشغيل وأخطاء التحليل. وعلى الرغم من أن هذه الملفات ليست مطلوبة لتقييم نتائج التحليل، إلا أنه يمكن استخدامها لأغراض استكشاف الأخطاء وإصلاحها. وتوجد جميع الملفات في مجلد المحاذاة ما لم يُنص على خلاف ذلك.

اسم الملف	الوصف
AnalysisLog.txt	سجل المعالجة الذي يصف كل خطوة حدثت أثناء تحليل مجلد عملية التشغيل الحالي. لا يحتوي هذا الملف على رسائل خطأ. ويقع في مجلد المحاذاة.
AnalysisError.txt	سجل المعالجة الذي يسرد أي أخطاء حدثت أثناء التحليل. وسيكون هذا الملف فارغاً في حال عدم حدوث أي أخطاء. ويقع في مجلد المحاذاة.
DemultiplexSummaryF1L1#.txt	يقدم تقريراً عن نتائج فك التجميع (Demultiplexing) في جدول يحتوي على صف واحد لكل شريحة وعمود واحد لكل عينة. يمثل الرمز # الممر رقم 1 أو 2 أو 3 أو 4 في خلية التدفق. ويقع في مجلد المحاذاة.
AmpliconRunStatistics.xml	يحتوي على إحصائيات ملخصة خاصة بعملية التشغيل. ويقع في مجلد المحاذاة.

مجلد التحليل

يحتوي مجلد التحليل على الملفات التي يتم إنشاؤها بواسطة برنامج مدير التشغيل المحلي. وتتخصص العلاقة بين مجلد المخرجات ومجلد التحليل فيما يلي:















- ◀ أثناء عملية التسلسل، يقوم برنامج (RTA) Real-Time Analysis بتزويد مجلد المخرجات بالملفات التي يتم إنشاؤها أثناء تحليل الصور، والاستدعاء القاعدي، وتسجيل درجات الجودة.
- ◀ يقوم برنامج RTA بنسخ الملفات إلى مجلد التحليل في الوقت الفعلي. وبعد أن يقوم برنامج RTA بتعيين سجل درجات الجودة لكل قاعدة في كل دورة، يكتب البرنامج ملف RTAComplete.txt في كلا المجلدين.
- ◀ عندما يتوفر ملف RTAComplete.txt، تبدأ عملية التحليل.
- ◀ ومع استمرار عملية التحليل، يكتب مدير التشغيل المحلي ملفات المخرجات في مجلد التحليل، ثم ينسخها مرة أخرى إلى مجلد المخرجات.

مجلدات المحاذاة

في كل مرة يتم فيها إعادة إدراج التحليل في قائمة الانتظار، ينشئ مدير التشغيل المحلي مجلد محاذاة يسمى **Alignment_N**، حيث يمثل الحرف N رقماً تسلسلياً.

هيكل المجلد

- Alignment - يحتوي على ملفات .bam و .vcf و .FASTQ، والملفات الخاصة بوحدة التحليل.
- Date and Time Stamp - ختم تاريخ_وقت التحليل بصيغة YYYYMMDD_HHMMSS
- AnalysisError.txt
- AnalysisLog.txt
- aggregate.report.html
- aggregate.report.pdf
- aggregate.summary.csv
- AmpliconCoverage_M#.tsv
- AmpliconRunStatistics.xml
- Sample1.genome.vcf.gz

- Sample1.coverage.csv 
- Sample1.report.pdf 
- Sample1.summary.csv 
- Sample1.vcf.gz 
- Sample1.bam 
- FASTQ** 
- Sample1** 
- Sample1_L001_R1_001_fastq.gz 
- Stats** 
- DemuxSummaryF1L1.txt 
- FastqSummaryF1L1.txt 
- Data** 
- Intensities** 
- BaseCalls** 
- L001** - يحتوي على ملفات *.bcl. 
- L001** - يحتوي على ملفات *.locs. 
- RTA Logs** - يحتوي على ملفات السجل الخاصة بتحليل برنامج التحليل في الوقت الفعلي (RTA). 
- InterOp** - يحتوي على ملفات ثنائية تُستخدم للإبلاغ عن مقاييس عملية تشغيل التسلسل. 
- Logs** - يحتوي على ملفات سجل تصف الخطوات التي تم تنفيذها أثناء عملية التسلسل. 
- RTAComplete.txt 
- RunInfo.xml 
- RunParameters.xml 

الاستدعاء الأساسي وتنوع المؤشر

عند إجراء تسلسل للعينات على جهاز NextSeq 550Dx، يحدد الاستدعاء القاعدي قاعدة (A أو C أو G أو T) لكل عنقود في شريحة معينة، أو منطقة تصوير في خلية التدفق، في دورة محددة. ويستخدم جهاز NextSeq 550Dx التسلسل ذو القناتين، والذي يتطلب صورتين فقط لتشفير البيانات للقواعد الأربعة للحمض النووي (DNA)، صورة من القناة الحمراء وصورة من القناة الخضراء. تختلف قراءات مؤشر عملية الاستدعاء الأساسي من الاستدعاء الأساسي خلال القراءات الأخرى. يجب أن تبدأ قراءات الفهرسة بأساس واحد على الأقل بخلاف G في أيّ من الدوريتين الأوليين. إذا بدأت قراءة الفهرسة باستدعائين أساسيين لـ G، فلن يتم بث إشارة قوية. يجب أن تكون الإشارة موجودة في كلٍ من الدوريتين الأوليين لضمان إزالة تعدد الإرسال. عند اختيار المؤشرات أثناء إنشاء عملية التشغيل، سيظهر تحذير قلة التنوع إذا كانت المؤشرات لا تُلبي متطلبات التنوع. ولمنع ظهور تحذير قلة التنوع، حدد تسلسلات المؤشر التي توفر إشارة في كلتا القناتين لكل دورة.

◀ القناة الحمراء - A أو C

◀ القناة الخضراء - A أو T

تضمن عملية الاستدعاء الأساسي الدقة عند تحليل نماذج إرسال الإشارات الضعيفة. ولمزيد من المعلومات حول تسلسلات المؤشرات الخاصة بك، راجع النشرة الداخلية لفحص مجموعة *TruSeq Custom Amplicon Kit Dx* (المستند رقم 1000000029772).

أثناء إنشاء عملية التشغيل في مدير التشغيل المحلي، ستختار عدد العينات التي سيتم اختبارها. يعمل البرنامج على تعينة مجموعات المؤشر المقترحة التي تُلبي متطلبات تنوع المؤشر تلقائيًا. وعلى الرغم من أنك لست ملزمًا باستخدام تركيبات المؤشرات المقترحة، إلا أنه يُوصى بذلك.

سجل المراجعة

المستند	التاريخ	وصف التغيير
مستند رقم 1000000030329 إصدار 04	أغسطس 2021	تم تحديث عنوان الممثل المعتمد للاتحاد الأوروبي.
مستند رقم 1000000030329 إصدار 03	ديسمبر 2019	تم تحديث عنوان الممثل المعتمد للاتحاد الأوروبي. تم تحديث عنوان الراعي الأسترالي.
مستند رقم 1000000030329 إصدار 02	يناير 2019	أضيفت معلومات حول مجموعات كاشف إصدار 2.5.
المستند رقم 1000000030329 الإصدار 01	أغسطس 2018	تم تحديث العلامات التنظيمية.
المستند رقم 1000000030329 الإصدار 00	نوفمبر 2017	الإصدار المبدئي.

المساعدة الفنية

للحصول على المساعدة الفنية، اتصل Illumina بالدعم الفني.

الموقع الإلكتروني: www.illumina.com
عنوان البريد الإلكتروني: techsupport@illumina.com

أرقام الهواتف الخاصة بدعم عملاء شركة Illumina

المنطقة	الرقم المجاني	الإقليمي
أمريكا الشمالية	+1.800.809.4566	
أستراليا	+1.800.775.6888	
النمسا	+43 800006249	+43 19286540
بلجيكا	+32 80077160	+32 34002973
الصين	400.066.5835	
الدانمارك	+45 80820183	+45 89871156
فنلندا	+358 800918363	+358 974790110
فرنسا	+33 805102193	+33 170770446
ألمانيا	+49 8001014940	+49 8938035677
هونغ كونج	800960230	
أيرلندا	+353 1800936608	+016950506 353
إيطاليا	+39 800985513	+39 236003759
اليابان	0.800.111.5011	
هولندا	+31 8000222493	+31 207132960
نيوزيلندا	0.800.451.650	
النرويج	+47 800 16836	+47 21939693
سنغافورة	+1.800.579.2745	
إسبانيا	+34 911899417	+34 800300143
السويد	+46 850619671	+46 200883979
سويسرا	+41 565800000	+41 800200442
تايوان	00806651752	
المملكة المتحدة	+44 8000126019	+44 2073057197
دول أخرى	+44.1799.534000	

تتوفر صحائف بيانات السلامة (SDS) - على موقع Illumina الإلكتروني على support.illumina.com/sds.html.

مستندات المنتج - متوفرة للتنزيل بصيغة PDF من الموقع الإلكتروني لشركة Illumina. انتقل إلى support.illumina.com، وحدد منتجًا، ثم حدد Documentation & Literature (الوثائق والمؤلفات العلمية).

الجهة الراعية الأسترالية
Illumina Australia Pty Ltd
Nursing Association Building
Level 3, 535 Elizabeth Street
Melbourne, VIC 3000
أستراليا

CE

IVD

EC REP

Illumina Netherlands B. V
Steenoven 19
DK Eindhoven 5626
هولندا



Illumina
5200 Illumina Way
San Diego, California 92122
U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (خارج أمريكا الشمالية)
techsupport@illumina.com
www.illumina.com

illumina®

لأغراض الاستخدام في التشخيص المختبري فقط
حقوق الطبع والنشر © لشركة Illumina, Inc 2021. جميع الحقوق محفوظة.