

Instrumento NextSeq™ 550Dx

PARA USO DIAGNÓSTICO *IN VITRO*
SOLO PARA EXPORTACIÓN

N.º de catálogo 20005715

Uso previsto

El instrumento NextSeq 550Dx está concebido para la secuenciación de librerías de ADN cuando se usa con ensayos de diagnóstico *in vitro*. El instrumento NextSeq 550Dx se ha de usar con reactivos de diagnóstico *in vitro* y *software* de análisis específicos que estén registrados, certificados o aprobados.

Principios del procedimiento

El instrumento NextSeq 550Dx de Illumina está concebido para la secuenciación de librerías de ADN con ensayos de diagnóstico *in vitro* y está concebido para su uso por parte de personal de laboratorio clínico formado que cuenta con formación en el uso de los procedimientos de diagnóstico *in vitro* que se realizan en un laboratorio clínico. Para la entrada, NextSeq 550Dx usa librerías generadas a partir de ADN en las que los índices de muestra y las secuencias de captura se añaden a los objetivos amplificados. Las librerías de muestras se capturan en una celda de flujo y se secuencian en el instrumento por medio del proceso químico de secuenciación por síntesis (SBS, Sequencing By Synthesis). El proceso químico de SBS usa un método basado en terminadores reversibles para detectar bases de nucleótidos únicos marcados con un colorante fluorescente a medida que se incorporan a las cadenas de ADN en crecimiento. El *software* de análisis en tiempo real (RTA) realiza análisis de imágenes y llamada de bases. Asimismo, asigna una puntuación de calidad a cada base de cada ciclo de secuenciación. Cuando finaliza el análisis principal, es posible ejecutar en el instrumento el análisis secundario para procesar las llamadas de bases. NextSeq 550Dx usa módulos de análisis secundarios diferentes según el flujo de trabajo. Para los módulos Germline Variant (variante germinal) o Somatic Variant (variante somática), el procesamiento incluye el demultiplexado, la generación del archivo FASTQ, la alineación, la llamada de variantes y la generación de archivos en formato de llamada de variantes (VCF y gVCF). Los archivos VCF y gVCF contienen información sobre las variantes que se encuentran en posiciones específicas en un genoma de referencia.

Configuración de doble arranque

NextSeq 550Dx incluye una configuración de doble arranque que permite el uso del instrumento en los modos de diagnóstico (Dx) o de uso exclusivo en investigación (RUO, Research Use Only). Los ensayos de secuenciación de diagnóstico *in vitro*, incluidos los módulos Germline Variant (variante germinal) y Somatic Variant (variante somática), se ejecutan en el modo de diagnóstico. Solo los reactivos de secuenciación de diagnóstico *in vitro* (DIV, In Vitro Diagnostic) pueden usarse en dicho modo. Se han establecido características de rendimiento y limitaciones de procedimiento para el instrumento NextSeq 550Dx usando los módulos Germline Variant (variante germinal) y Somatic Variant (variante somática) en el modo de diagnóstico.

Limitaciones del procedimiento

- 1 Para uso diagnóstico *in vitro*.
- 2 Los módulos Germline Variant (variante germinal) y Somatic Variant (variante somática), si se usan con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), pueden llegar a ofrecer lo siguiente:
 - ▶ Rendimiento de secuenciación: ≥90 gigabases (Gb)

- ▶ Longitud de lectura (en experimento "paired-end"): 2 × 150 pares de bases (pb)
 - ▶ Bases iguales o superiores a Q30 ≥ 75 % a una longitud de lectura de 2 × 150 pb
Un 75 % o más de las bases tiene una puntuación de calidad según la escala de Phred ≥ 30 , lo que implica una precisión de la llamada de bases superior al 99,9 %
- 3 El *software* del ensayo no alineará las lecturas con indel (inserciones, deleciones o combinaciones) cuya longitud de contenido sea superior a los 25 pb. Por lo tanto, el *software* del ensayo no podrá detectar las indel con una longitud superior a los 25 pb.
 - 4 Es posible que el *software* del ensayo no alinee las lecturas de amplicones que contienen un contenido excesivo de variantes y, en consecuencia, se informará de que la región se encuentra en estado natural. Tal contenido excesivo incluye:
 - ▶ Lecturas que contienen más de tres indels
 - ▶ Lecturas cuya longitud es de, al menos, 30 pb y cuyo contenido de variante de nucleótido único (SNV) es superior al 4 % de la longitud total del amplicón objetivo (sin tener en cuenta las regiones de sonda)
 - ▶ Lecturas cuya longitud es inferior a 30 pb y cuyo contenido de SNV es superior al 10 % de la longitud total del amplicón (incluidas las regiones de sonda)
 - 5 Es posible que las variantes de mayor tamaño, incluidas las variantes de nucleótidos múltiples (MNV, Multi-Nucleotide Variants) y las indels de gran tamaño, aparezcan en el archivo VCF de resultados por separado como variantes de menor tamaño.
 - 6 Cuando se cubren dos amplicones en la placa, es posible que las variantes de deleción se omitan o queden filtradas si la longitud de la deleción es mayor o igual que la región de solapamiento entre los amplicones de la placa.
 - 7 El sistema no puede detectar indels si son directamente adyacentes a un cebador y no existe solapamiento de amplicones. Con respecto a las regiones con amplicones solapados, el ensayo no puede detectar deleciones cuando la región de solapamiento es más pequeña que el tamaño de la deleción que se desea detectar. Por ejemplo, si la región de solapamiento entre dos amplicones adyacentes tiene dos bases, el ensayo no puede detectar deleciones que incluyan esas dos bases. Una deleción de base individual en cualquiera de esas dos bases se puede detectar.
 - 8 Como con cualquier flujo de trabajo de preparación de librerías basado en hibridación, los polimorfismos, las mutaciones, las inserciones o las deleciones subyacentes en regiones de unión de oligonucleótidos pueden afectar a los alelos que se comprueban y a las llamadas que se han realizado durante la secuenciación. Por ejemplo:
 - ▶ Una variante en fase que contiene una variante en la región del cebador podría no amplificarse, lo que conllevaría un falso negativo.
 - ▶ Es posible que las variantes que se sitúan en la región del cebador impidan la amplificación del alelo de referencia, lo que conllevaría una llamada de variantes homocigóticas incorrecta.
 - ▶ Las variantes de indel que se encuentran en la región del cebador pueden provocar un falso positivo en la llamada en la parte final de la lectura de las posiciones adyacentes al cebador.
 - 9 Las indels pueden quedar filtradas a causa de cortes en la cadena si están situadas en la parte final de una lectura y se han truncado levemente durante la alineación.
 - 10 Las MNV de menor tamaño no se han validado y solo se identifican en el módulo Somatic Variant (variante somática).
 - 11 Las deleciones se identifican en el archivo VCF en la coordenada de la base precedente de cada formato VCF. Por lo tanto, se deben considerar las variantes adyacentes antes de identificar una llamada de bases individual como una referencia homocigótica.
 - 12 Limitaciones específicas de Germline (germinal):
 - ▶ El instrumento NextSeq 550Dx, que usa el módulo Germline Variant (variante germinal) de Local Run Manager para NextSeq 550Dx, está diseñado para ofrecer resultados cualitativos en lo que se refiere a la llamada de Germline Variant (variante germinal) (por ejemplo, homocigóticas, heterocigóticas y en estado natural).

- ▶ Para usar el módulo Germline Variant (variante germinal), es necesaria una cobertura mínima de 150x por amplicón para que la llamada de variantes sea exacta. Por lo tanto, se precisan 150 fragmentos con ADN, lo que equivale a 300 lecturas "paired-end" solapadas. La cantidad de muestras y el número total de bases selectivas afectan a la cobertura. El contenido de GC, así como otros contenidos genómicos, también afectan a la cobertura.
- ▶ La variación en el número de copias puede afectar en la identificación de una variante como homocigótica o heterocigótica.
- ▶ Las variantes que se encuentran en un contexto repetitivo determinado se filtran en los archivos VCF. El filtro RMxN para las repeticiones se usa para filtrar variantes en caso de que la secuencia de la variante aparezca varias veces, en parte o en su totalidad, en el genoma de referencia adyacente a la posición de la variante. En el caso de las llamadas de Germline Variant (variante germinal), es necesario que existan al menos nueve repeticiones en la referencia para que se filtre la variante. Únicamente se tienen en cuenta las repeticiones con una longitud de hasta 5 pb (R5x9).
- ▶ Una indel y una SNV en un único locus pueden tener como resultado que solo se identifique una sola variante.

13 Limitaciones específicas de Somatic (somática).

- ▶ El instrumento NextSeq 550Dx, que usa el módulo Somatic Variant (variante somática) de Local Run Manager para NextSeq 550Dx, está diseñado para ofrecer resultados cualitativos en lo que se refiere a la llamada de Somatic Variant (variante somática) (por ejemplo, la presencia de una Somatic Variant (variante somática) con una frecuencia de variante mayor o igual a 0,026 con un límite de detección de 0,05).
- ▶ Para usar el módulo Somatic Variant (variante somática), es necesaria una cobertura mínima por amplicón de 450x por grupo de oligonucleótidos para que la llamada de variantes sea exacta. Por lo tanto, se precisan 450 fragmentos con ADN por grupo de oligonucleótidos, lo que equivale a 900 lecturas "paired-end" solapadas. La cantidad de muestras y el número total de bases selectivas afectan a la cobertura. El contenido de GC, así como otros contenidos genómicos, también afectan a la cobertura.
- ▶ En el caso de las llamadas de Somatic Variant (variante somática), es necesario que existan al menos seis repeticiones en la referencia para que se filtre la variante y únicamente se tienen en cuenta las repeticiones con una longitud de hasta 3 pb (R3x6).
- ▶ El módulo Somatic Variant (variante somática) no es capaz de diferenciar entre Germline Variants (variantes germinales) y Somatic Variants (variantes somáticas). Aunque el módulo está diseñado para detectar las variantes dentro de un espectro de frecuencias de variante, no es posible usar las frecuencias de variante para diferenciar las Somatic Variants (variantes somáticas) de las Germline Variants (variantes germinales).
- ▶ El tejido común que se encuentre en la muestra puede influir en la detección de las variantes. El límite de detección informado se basa en una frecuencia de variante relativa al total de ADN que se ha extraído del tumor y del tejido común.

Componentes del producto

- 1 Instrumento NextSeq 550Dx (n.º de catálogo 20005715)
- 2 Componentes de *software* para el instrumento NextSeq 550Dx, incluidos los siguientes:

Aplicación de <i>software</i>	Función	Descripción
NextSeq 550Dx Operating Software (NOS)	Controla el funcionamiento del instrumento	La aplicación de <i>software</i> NOS gestiona el funcionamiento del instrumento durante la secuenciación y genera imágenes para que las use el <i>software</i> de análisis en tiempo real (RTA, Real-Time Analysis).
<i>Software</i> de análisis en tiempo real (RTA)	Realiza el análisis principal	La aplicación de <i>software</i> RTA convierte las imágenes que genera NOS de cada placa por ciclo del experimento de secuenciación en archivos de llamada de bases, que son entradas de los módulos de análisis de Local Run Manager. La aplicación de <i>software</i> de RTA no tiene interfaz de usuario.
Local Run Manager	Interfaz para la selección del módulo	El <i>software</i> Local Run Manager es una solución integrada en el instrumento que permite la gestión de usuarios, la selección del módulo de análisis adecuado y la supervisión del estado.
Módulo Somatic Variant (variante somática)	Realiza el análisis secundario	Este <i>software</i> del módulo de análisis de Local Run Manager procesa las llamadas de bases con análisis secundarios. El procesamiento incluye el demultiplexado, la generación de archivos FASTQ, la alineación, la llamada de variantes y la generación de informes. El llamador de variantes (Pisces) genera archivos VCF que contienen información sobre las variantes que se han detectado en posiciones determinadas de un genoma de referencia e incluye también la medición de la frecuencia de variante.
Módulo Germline Variant (variante germinal)	Realiza el análisis secundario	Este <i>software</i> del módulo de análisis de Local Run Manager procesa las llamadas de bases con análisis secundarios. El procesamiento incluye el demultiplexado, la generación de archivos FASTQ, la alineación, la llamada de variantes y la generación de informes. El llamador de variantes (Pisces) genera archivos VCF que contienen información sobre las variantes que se han detectado en posiciones determinadas de un genoma de referencia y determina si cada variante es heterocigótica u homocigótica.

Condiciones de funcionamiento

Elemento	Especificación
Temperatura	Mantenga una temperatura de laboratorio de entre 19 °C y 25 °C (22 °C ± 3 °C). Se trata de la temperatura de funcionamiento del instrumento. Durante un experimento, no deje que la temperatura ambiente varíe más de ±2 °C.
Humedad	Mantenga una humedad relativa sin condensación de entre el 20 % y el 80 %.

Materiales y equipo

Materiales y equipo necesarios que se venden por separado

NextSeq 550Dx High Output Reagent Kit v2.5 (75 cycles), n.º de catálogo 20028870

NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), n.º de catálogo 20028871

Materiales y equipo necesarios, no suministrados

Consumibles suministrados por el usuario para los experimentos de secuenciación

Consumible	Proveedor	Finalidad
Paño humedecido en alcohol isopropilo al 70 % o en etanol al 70 %	VWR, n.º de catálogo 95041-714 (o equivalente) Proveedor de laboratorio general	Limpieza de la celda de flujo y fines generales
Toallita de laboratorio sin pelusa	VWR, n.º de catálogo 21905-026 (o equivalente)	Limpieza de la celda de flujo y fines generales

Consumibles suministrados por el usuario para el mantenimiento del instrumento

Consumible	Proveedor	Uso
NaOCl al 5 % (hipoclorito de sodio)	Sigma-Aldrich, n.º de catálogo 239305 (o equivalente de laboratorio)	Lavado del instrumento con el lavado manual posterior al experimento; diluido al 0,12 %
Tween 20	Sigma-Aldrich, n.º de catálogo P7949	Lavado del instrumento con las opciones de lavado manual; diluido al 0,05 %
Agua de laboratorio	Proveedor de laboratorio general	Lavado del instrumento (lavado manual)
Filtro de aire	Ilumina, n.º de catálogo 20022240	Limpieza del aire que el instrumento recibe para refrigerarse

Guías para el agua de laboratorio

Use siempre agua de laboratorio o agua desionizada para llevar a cabo los procedimientos del instrumento. No use nunca agua corriente. Use solamente los siguientes tipos de agua o equivalentes:

- ▶ Agua desionizada
- ▶ Illumina PW1
- ▶ Agua de 18 Megohmios (MΩ)
- ▶ Agua Milli-Q
- ▶ Agua Super-Q
- ▶ Agua de biología molecular

Advertencias y precauciones

PRECAUCIÓN Las leyes federales limitan la venta de este dispositivo a médicos u otros facultativos, o bajo prescripción de estos, que se encuentren autorizados en virtud de la legislación del estado en el que ejercen su profesión para usar u ordenar el uso de este dispositivo.

- 1 **Algunos componentes de los reactivos proporcionados por Illumina que se usan con el instrumento NextSeq 550Dx contienen sustancias químicas potencialmente peligrosas. Evite su inhalación, ingestión y el contacto con la piel o los ojos, puesto que puede provocar lesiones. Use un equipo de protección, incluidos gafas, guantes y batas de laboratorio adecuados para el riesgo de exposición. Manipule los reactivos usados como residuos químicos y deséchelos de conformidad**

con las normativas y leyes regionales, nacionales y locales aplicables. Para obtener más información medioambiental, sanitaria y de seguridad, consulte las hojas de datos de seguridad (SDS) en support.illumina.com/sds.html.

- 2 Comuníquese de inmediato cualquier incidente grave relacionado con este producto a Illumina y las autoridades competentes de los Estados miembros en los que se encuentren el usuario y el paciente.
- 3 Manipule todas las muestras de sangre como si contuviesen agentes infecciosos del virus de la inmunodeficiencia humana (VIH), del virus de la hepatitis B humana (VHB) o de cualquier otro patógeno de transmisión sanguínea (precauciones universales).
- 4 El incumplimiento de los procedimientos descritos puede provocar resultados erróneos o una reducción considerable de la calidad de las muestras.
- 5 Use las precauciones rutinarias del laboratorio. No pipetee con la boca. No coma, beba ni fume en las zonas de trabajo designadas. Lleve guantes desechables y batas de laboratorio para la manipulación de muestras y los reactivos del kit. Lávese bien las manos tras la manipulación de muestras y reactivos del kit.
- 6 Se precisan prácticas de laboratorio adecuadas y procedimientos óptimos en materia de higiene de laboratorio para evitar que los productos de PCR contaminen los reactivos, los instrumentos y las muestras de ADN genómico. La contaminación mediante PCR puede conllevar resultados imprecisos y poco fiables.
- 7 Para evitar la contaminación, asegúrese de que las zonas de preamplificación y posamplificación dispongan de equipos y consumibles específicos (tales como pipetas, puntas de pipeta, bloques de calentamiento, mezcladores vorticiales y centrifugadoras).
- 8 El emparejamiento de índice a muestra debe coincidir exactamente con la disposición de placas impresa. Local Run Manager completa de forma automática los cebadores de índices que se han asociado a los nombres de las muestras cuando estos se introducen en el módulo. Resulta aconsejable que el usuario haga una comprobación de los cebadores de índices que se han asignado a las muestras antes de comenzar el experimento de secuenciación. Las discrepancias entre los resultados de disposición de muestras y de la placa pueden provocar que no se identifiquen las muestras positivas y la generación de informes con resultados incorrectos.
- 9 Se recomienda encarecidamente instalar un *software* antivirus proporcionado por el usuario para proteger el ordenador frente a virus informáticos. Consulte el manual del usuario para obtener instrucciones acerca de la instalación.
- 10 No ponga en funcionamiento NextSeq 550Dx si se ha retirado alguno de los paneles. La puesta en funcionamiento del instrumento estando alguno de estos paneles retirado constituye una posible exposición a la tensión de línea, así como a tensiones de CC.
- 11 No toque la platina de la celda de flujo en el compartimento de la celda de flujo. El calentador de este compartimento funciona a una temperatura de entre 22 °C y 95 °C, por lo que podría causar quemaduras.
- 12 El instrumento pesa aproximadamente 84 kg (185 libras) y podría provocar lesiones graves si se cae o se maneja de forma indebida.

Instrucciones de uso

Las siguientes instrucciones de uso describen cómo usar los módulos Germline Variant (variante germinal) y Somatic Variant (variante somática) en el modo de diagnóstico en el instrumento NextSeq 550Dx usando NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Introducción de información del experimento

Para obtener instrucciones detalladas, consulte la Guía de referencia del instrumento NextSeq 550Dx (n.º de documento 1000000009513) y la Guía del módulo de Local Run Manager aplicable.

Configuración de parámetros

- 1 Inicie sesión en Local Run Manager.
- 2 Seleccione **Create Run** (Crear experimento) y seleccione **Somatic Variant** (Variante somática) o **Germline Variant** (Variante germinal).

- 3 Introduzca un nombre para el experimento de forma que se pueda identificar desde la secuenciación hasta el análisis.
Utilice caracteres alfanuméricos, espacios o guiones tanto normales como bajos.
- 4 **(Opcional)** Escriba una descripción que ayude a identificar el experimento.
Utilice caracteres alfanuméricos, espacios o guiones tanto normales como bajos.
- 5 Seleccione el número de muestras y la configuración del índice de la lista desplegable.
Tenga en cuenta la siguiente información a la hora de hacer la selección.
 - ▶ La lista desplegable contiene números de muestras con un juego de índices. Por ejemplo, "24-Set 1" indica 24 muestras que analizar, con índices del juego 1 de índices.
 - ▶ Los números de juegos de índices hacen referencia a los distintos juegos de pares de índices i5 e i7. Tanto el juego 1 como el juego 2 proporcionan diversidad de índices. Se ofrecen dos juegos de índices para ayudar a prevenir la eliminación de un único juego.
 - ▶ Escoja el número de muestras que más se acerque a la cantidad de muestras que vaya a analizar. Si la cantidad exacta de muestras no aparece en el desplegable, elija el número más próximo, pero siempre inferior al número de muestras que se van a analizar. Por ejemplo, si va a analizar 18 muestras, seleccione 16.
 - ▶ Las combinaciones de índices y pocillos de muestras sugeridas que cumplen los requisitos de diversidad de índices están resaltadas en color verde.

Importación de archivos de manifiesto para el experimento

- 1 Asegúrese de que los manifiestos que necesita importar están a su disposición a través de una ubicación accesible de la red o de un dispositivo USB.
- 2 Seleccione **Import Manifests** (Importar manifiestos).
- 3 Vaya al archivo de manifiesto y elija los manifiestos que quiere añadir.

NOTA Para que los archivos de manifiesto estén disponibles para cualquier experimento en el que se utilicen los módulos de análisis Germline Variant (Variante germinal) o Somatic Variant (Variante somática), añada los manifiestos por medio de la función Module Settings (Ajustes del módulo). Para acceder a esta función es necesario obtener permisos de administrador. Para obtener más información, consulte la *Guía de referencia del instrumento NextSeq 550Dx (n.º de documento 100000009513)*.

Definición de las muestras para el experimento

Defina las muestras que se van a usar en el experimento por medio de una de las siguientes opciones e instrucciones.

- ▶ **Introducir las muestras de forma manual:** use la tabla en blanco que aparece en la pantalla Create Run (Crear experimento).
- ▶ **Importar muestras:** vaya a un archivo externo que tenga un formato de valores separados por comas (*.csv). En la pantalla Create Run (Crear experimento), se ofrece una plantilla que se puede descargar.

Introducción de las muestras de forma manual

- 1 Introduzca un nombre de muestra (**módulo de análisis Somatic Variant** [Variante somática]) o un ID de muestra (**módulo de análisis Germline Variant** [Variante Germline]) que sean únicos.
Puede utilizar caracteres alfanuméricos y guiones tanto normales como bajos.
- 2 **(Opcional)** Para las muestras de control tanto negativo como positivo, haga clic en el botón derecho del ratón y seleccione el tipo de control.
El control que se asigna al pocillo de la muestra aparece automáticamente en el pocillo correspondiente en el otro conjunto y asigna ese mismo control.
- 3 **(Opcional)** Introduzca la descripción de la muestra en el campo Sample Description (Descripción de la muestra).
Puede utilizar caracteres alfanuméricos y guiones tanto normales como bajos.
- 4 Seleccione un adaptador del Index 1 (Índice 1) del menú desplegable que aparece en el Índice 1 (i7).

Si utiliza los pocillos de muestra sugeridos, el software rellena automáticamente los adaptadores de índices i7 e i5 que cumplen los requisitos de índice de diversidad. Si la cantidad exacta de muestras que está analizando no aparece en el desplegable, asegúrese de seleccionar los adaptadores de índices para los pocillos de más.

- 5 Seleccione un adaptador del Index 2 (Índice 2) del menú desplegable que aparece en el Índice 2 (i5).
- 6 Elija un archivo de manifiesto de la lista que aparece en el desplegable de Manifest (Manifiesto).
Las muestras que se encuentran en el conjunto A precisan un manifiesto diferente al de las muestras del conjunto B.
- 7 Seleccione la opción correspondiente para visualizar, imprimir o guardar la disposición de la placa para utilizarla como referencia a la hora de preparar librerías:
 - Haga clic en el icono  **Print** (Imprimir) para ver la disposición de la placa. Seleccione **Print** (Imprimir) para imprimir la disposición de la placa.
 - Seleccione **Export** (Exportar) para exportar la información sobre las muestras a un archivo externo.
- 8 Seleccione **Save Run** (Guardar experimento).

Importación de las muestras

- 1 Seleccione **Import Samples** (Importar muestras) y busque la ubicación del archivo que contiene la información de la muestra. Se pueden importar dos tipos de archivo.
 - Para crear una nueva disposición de placa, haga clic en **Template** (Plantilla) en la pantalla Create Run (Crear experimento). El archivo que se usa como plantilla incluye los encabezados de columna adecuados para la importación. Introduzca en cada columna la información relativa a las muestras del experimento. Elimine la información que aparece como ejemplo en las celdas que no vaya a utilizar y guarde el archivo.
 - Emplee uno de los archivos de información sobre las muestras que se hayan exportado de los módulos Germline Variant (Variante germinal) o Somatic Variant (Variante somática) por medio de la función Export (Exportación).
- 2 Haga clic en el icono  **Print** (Imprimir) para ver la disposición de la placa.
- 3 Seleccione **Print** (Imprimir) para poder imprimir la disposición de la placa y utilizarla como referencia a la hora de preparar las librerías.
- 4 Seleccione **Save Run** (Guardar experimento).

Preparación del cartucho de reactivos

Asegúrese de seguir con cuidado las indicaciones del cartucho de reactivos para lograr una correcta secuenciación.

- 1 Extraiga el cartucho de reactivos almacenado a una temperatura de entre -25 °C y -15 °C .
- 2 Elija uno de los métodos siguientes para descongelar los reactivos. No sumerja el cartucho. Una vez descongelado el cartucho, séquelo antes de continuar con el siguiente paso.

Temperatura	Tiempo hasta la descongelación	Límite de estabilidad
Baño de agua a una temperatura entre 15 °C y 30 °C	60 minutos	Sin superar las seis horas
Entre 2 °C y 8 °C	7 horas	Sin superar los cinco días

NOTA Si se está descongelando más de un cartucho en el mismo baño de agua, calcule más tiempo hasta la descongelación.

- 3 Invierta el cartucho cinco veces para mezclar los reactivos.
- 4 Realice una inspección de la parte inferior del cartucho para asegurarse de que los reactivos se hayan descongelado y no presenten precipitados. Confirme que las posiciones 29, 30, 31 y 32 están descongeladas, ya que son las más grandes y pueden necesitar más tiempo.
- 5 Golpéelo ligeramente sobre la mesa para reducir las burbujas de aire.
Para obtener unos resultados óptimos, proceda directamente con la carga de la muestra y la configuración del experimento.

Preparación de la celda de flujo

- 1 Retire una nueva caja de celda de flujo de su almacenamiento a una temperatura de entre 2 °C y 8 °C.
- 2 Retire el embalaje metálico de la caja y resérvela a temperatura ambiente durante 30 minutos.

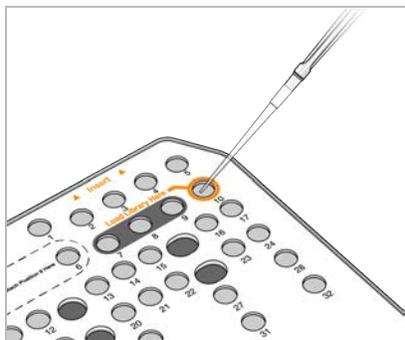
Preparación de librerías para la secuenciación

Desnaturalice y diluya las librerías a un volumen de carga de 1,3 ml. En la práctica, la concentración de carga puede variar en función de los métodos de cuantificación y preparación de librerías. La dilución de las librerías de muestras depende de la complejidad de los grupos de oligonucleótidos. Para obtener información sobre la preparación de las librerías de muestras para la secuenciación, incluidas la agrupación y la dilución de librerías, consulte la sección Instrucciones de uso del kit de preparación de librerías aplicable. NextSeq 550Dx precisa la optimización de la densidad de grupos.

Carga de librerías en el cartucho de reactivos

- 1 Limpie el cierre metálico que cubre el depósito n.º 10 marcado como **Load Library Here** (Cargar librería aquí) con una toallita sin pelusa.
- 2 Perfore el cierre metálico con una punta de pipeta de 1 ml limpia.
- 3 Cargue 1,3 ml de librerías preparadas en el depósito n.º 10 marcado como **Load Library Here** (Cargar librería aquí). Evite tocar el cierre metálico cuando dispense las librerías.

Figura 1 Carga de librerías



Configuración de un experimento de secuenciación

- 1 Inicie sesión en NextSeq 550Dx con su contraseña del *software* Local Run Manager.
- 2 En la pantalla Home (Inicio) del *software* NOS, seleccione **Sequence** (Secuenciar).
- 3 Seleccione uno de los experimentos del listado y, a continuación, seleccione **Next** (Siguiente). Se abrirá un conjunto de pantallas de configuración del experimento en el siguiente orden: Load Flow Cell (Cargar celda de flujo), Load Buffer Cartridge (Cargar cartucho de tampones), Load Reagent Cartridge (Cargar cartucho de reactivos) y Pre-Run Check (Comprobación previa al experimento).
- 4 Cuando aparezca la pantalla Load Flow Cell (Cargar celda de flujo), limpie la celda de flujo y, después, cárguela.
 - ▶ Retire la celda de flujo del embalaje metálico.
 - ▶ Abra el estuche de plástico transparente y retire la celda de flujo.
 - ▶ Limpie la superficie de cristal de la celda de flujo con un paño sin pelusa humedecido en alcohol. Seque el cristal con una toallita de laboratorio sin pelusa.
 - ▶ Asegúrese de que la superficie del cristal de la celda de flujo esté limpia. Si fuera necesario, repita los pasos de limpieza.
 - ▶ Retire la celda de flujo usada de un experimento anterior.
 - ▶ Alinee la celda de flujo sobre los pasadores de alineación y colóquela en la platina.

- 5 Seleccione **Load** (Cargar).
La puerta se cierra de forma automática, el ID de la celda de flujo aparece en la pantalla y se comprueban los sensores.
- 6 Siga las indicaciones del *software* para vaciar el contenedor de reactivos usado y cargar el cartucho de tampones y el cartucho de reactivos de NextSeq 550Dx.
Tras cargar los cartuchos de tampones y de reactivos de NextSeq 550Dx, el *software* lee y registra la RFID. Aparecen en la pantalla los ID de los cartuchos de tampones y de reactivos y se comprueban los sensores.
- 7 Una vez finalizada la comprobación automatizada previa al experimento, seleccione **Start** (Iniciar). (Este paso puede omitirse si se configura el inicio de forma automática).
- 8 La pantalla Sequencing (Secuenciación) se abre cuando empieza el experimento. Esta pantalla proporciona una representación visual del experimento en curso, incluidas las intensidades y las puntuaciones de calidad (puntuaciones Q).

Resultados

El análisis en tiempo real (RTA) es un *software* integrado que realiza análisis de imágenes y llamada de bases. Asimismo, asigna una puntuación de calidad a cada base de cada ciclo de secuenciación. Cuando finaliza el análisis principal, el módulo seleccionado de Local Run Manager en el instrumento NextSeq 550Dx inicia automáticamente el análisis secundario. El análisis secundario descrito aquí sirve para los módulos Germline Variant (variante germinal) y Somatic Variant (variante somática).

Demultiplexado

El demultiplexado compara cada secuencia de lectura de índice con las secuencias de índice que se han fijado para el experimento. En este paso, no se considera ningún valor de calidad.

Las lecturas de índice se pueden identificar por medio de los siguientes pasos:

- ▶ Las muestras están numeradas comenzando por el número uno de acuerdo con el lugar que ocupan en el experimento.
- ▶ La muestra número cero se reserva para los grupos que no se han asignado a una muestra.
- ▶ Los grupos se asignan a una muestra siempre que la secuencia de índice coincida de forma exacta o cuando exista solamente una discrepancia por lectura de índice.

Generación de archivos FASTQ

Después del demultiplexado, el *software* genera archivos de análisis intermedios en el formato FASTQ, que es un formato de texto empleado para representar secuencias. Los archivos FASTQ contienen las lecturas de cada muestra con sus puntuaciones de calidad correspondientes. Los grupos que no han podido pasar los filtros quedan excluidos.

Cada archivo FASTQ contiene las lecturas de una sola muestra, cuyo nombre está incluido en el nombre del archivo FASTQ. En los módulos Germline Variant (variante germinal) y Somatic Variant (variante somática), se generan ocho archivos FASTQ por muestra, por grupo de oligonucleótidos (es decir, cuatro de la Lectura 1 y cuatro de la Lectura 2). Este resultado se traduce en un total de ocho archivos FASTQ por muestra para Germline (germinal) y de 16 archivos FASTQ para Somatic (somática). Los archivos FASTQ incluyen la entrada principal para la alineación.

Alineación

A lo largo del proceso de alineación, el algoritmo de Smith-Waterman de bandas alinea los grupos de cada muestra con las secuencias del amplicón especificadas en el archivo de manifiesto.

El algoritmo de Smith-Waterman de bandas lleva a cabo alineaciones de secuencias semiglobales para determinar las regiones similares que hay en dos secuencias. En lugar de comparar la secuencia total, el algoritmo de Smith-Waterman compara los segmentos de todas las longitudes posibles.

Cada lectura "paired-end" se evalúa teniendo en cuenta su alineación respecto a las secuencias de sonda significativas de dicha lectura.

- ▶ La lectura 1 se evalúa teniendo en cuenta el complemento inverso de los oligonucleótidos descendentes de locus específico (DLSO, Downstream Locus-Specific Oligos).
- ▶ La lectura 2 se evalúa teniendo en cuenta los oligonucleótidos ascendentes de locus específico (ULSO, Upstream Locus-Specific Oligos).
- ▶ Si el comienzo de una lectura coincide con una secuencia de sonda con solo una incoherencia, la longitud completa de la lectura se alineará con el objetivo de amplicones de dicha secuencia.
- ▶ Si el comienzo de una lectura observa una coincidencia con la secuencia de sonda con menos de tres diferencias (discrepancias o cambios producidos por las indels principales), la lectura completa se alineará con el amplicón objetivo de esa secuencia.
- ▶ Las indels dentro de los DLSO y de los ULSO no se observan debido a los procesos químicos del ensayo. Las alineaciones se filtran teniendo en cuenta los resultados de alineación, que se basan en los índices de discrepancia que se encuentran en la región estudiada o en el amplicón completo, dependiendo de la longitud de este último. Las alineaciones filtradas se copian en los archivos de alineación como no alineadas y, por tanto, no se usan para la llamada de variantes.

Llamada de variantes

El llamador de variantes PISCES está diseñado para hacer llamadas de variantes de SNV (Single Nucleotide Variant) e indel a partir de las librerías que se han preparado para el instrumento.

Informes y archivos complementarios de resultados

Los módulos de análisis de variantes generan informes en formato PDF y de texto delimitado por tabulaciones (*.txt) que muestran criterios de medición como la profundidad de secuenciación y el recuento de variantes. Los módulos generan también archivos de resultados como los de VCF y de formato de llamada de variantes de genoma (gVCF, Genome Variant Call Format) para aplicaciones de llamada de variantes.

Procedimientos de control de calidad

El *software* NextSeq 550Dx evalúa cada experimento, cada muestra y cada llamada de bases de acuerdo con los criterios de medición de control de calidad. Asimismo, para la preparación de librerías, se recomiendan controles positivos y negativos y su evaluación. Evalúe los controles como se indica a continuación:

- **Control negativo sin cadena molde u otro control negativo:** debe generar el resultado previsto. Si el control negativo genera un resultado distinto al previsto, significa que se ha producido un posible error en el seguimiento de la muestra, un registro incorrecto de los cebadores de índice o contaminación.
- **Muestra de control positivo:** debe generar el resultado previsto. Si el control positivo genera un resultado distinto al previsto, significa que se ha producido un posible error en el seguimiento de la muestra o el registro incorrecto de los cebadores de índice.

Características de rendimiento

Las características de rendimiento del instrumento NextSeq 550Dx se establecieron usando los módulos Germline Variant (variante germinal) y Somatic Variant (variante somática) con TruSeq Custom Amplicon Kit Dx y NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) y confirmaron usando NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Los estudios incluían indexación de muestras, arrastre de muestras, entrada de ADN, sensibilidad del análisis (límite del blanco/límite de detección), exactitud, precisión, comparación de métodos y reproducibilidad.

Los estudios analíticos que usan NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) se diseñaron para evaluar los valores declarados de rendimiento establecidos previamente con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Los resultados demuestran que los kits de reactivos (v2 y v2.5) tienen un

rendimiento comparable usando TruSeq Custom Amplicon Kit Dx. Consulte las *instrucciones de uso de TruSeq Custom Amplicon Kit Dx* para conocer las características relacionadas con los factores preanalíticos, por ejemplo, los métodos de extracción o las sustancias interferentes.

Definiciones de los cálculos que se usan en las características de rendimiento

- 1 La coincidencia de porcentaje positivo (PPA, Positive Percent Agreement) es la proporción de locus que se clasifican como variantes usando un método de referencia y que se detectan de forma correcta en el ensayo.
 - ▶ $(\text{N.º de locus en los que se han detectado variantes de forma correcta en el ensayo}) / (\text{cantidad total de locus en los que existen variantes})$
Los locus en los que se han detectado variantes en el ensayo y que coinciden con el método de referencia son positivos verdaderos (TP, True Positive). Los locus en los que se han detectado variantes como llamadas de referencia o como distintas llamadas de variantes en el ensayo son falsos negativos (FN, False Negative).
- 2 La coincidencia de porcentaje negativo (NPA, Negative Percent Agreement) es la proporción de locus que se clasifican como en estado natural por medio de un método de referencia y que se detectan de forma correcta en el ensayo.
 - ▶ $(\text{N.º de locus en estado natural detectados de forma correcta en el ensayo}) / (\text{cantidad total de locus en estado natural})$
Los locus en estado natural que se han detectado en el ensayo y que concuerdan con el método de referencia son negativos verdaderos (TN, True Negative). Los locus en estado natural que en el ensayo se detectan como variantes son falsos positivos (FP, False Positive).
- 3 La coincidencia total de porcentajes (OPA, Overall Percent Agreement) es la proporción de locus que se detectan de forma correcta en el ensayo con respecto a un método de referencia.
 - ▶ $([\text{N.º de locus en los que se han detectado variantes de forma correcta en el ensayo}] + [\text{n.º de locus en estado natural detectados de forma correcta en el ensayo}]) / ([\text{n.º total de locus en los que existen variantes}] + [\text{n.º total de locus en estado natural}])$
- 4 Los cálculos que se llevan a cabo para determinar la PPA, la NPA y la OPA no incluyen la ausencia de llamadas (locus de referencia o en los que existen variantes que no pasan uno o varios filtros de calidad).
- 5 El índice de llamada de autosoma se calcula teniendo en cuenta el número total de locus que han pasado los filtros y dividiéndolo por el número total de posiciones secuenciadas en los cromosomas del 1 al 22. Los cromosomas X e Y quedan excluidos. Este criterio de medición no tiene en cuenta la coincidencia de las llamadas con el método de referencia.

Rendimiento de NextSeq 550Dx High Output Reagent Kit v2 (300 Cycles)

Indexación de muestras

Los cebadores de índice de muestras, añadidos durante la preparación de librerías, asignan una secuencia única a cada ADN de muestra. Estas secuencias únicas permiten agrupar varias muestras en un único experimento de secuenciación. La indexación de muestras se usa en los flujos de trabajo tanto de Germline (germinal) como Somatic (somática). El objetivo de este estudio fue establecer el número mínimo (ocho) y máximo (96) de muestras que se pueden procesar en un único experimento de secuenciación en el instrumento NextSeq 550Dx. Se analizaron ocho muestras únicas de Platinum Genomes con 12 combinaciones distintas de cebadores de índice por muestra. Los resultados de las muestras de cuatro experimentos de secuenciación llevados a cabo con el módulo Germline Variant (variante germinal) se compararon con la versión 1.0 de Platinum Genomes de 2016.

Para el primer conjunto de experimentos, se comprobaron 96 librerías de muestras indexadas de forma única usando un ensayo representativo diseñado para buscar una variedad de genes en 12 588 bases por cadena de cada uno de los 23 cromosomas humanos con el fin de verificar la capacidad del ensayo de realizar una llamada de genotipado de manera sistemática con una muestra determinada en combinaciones distintas de cebadores de índice. Para el segundo conjunto de experimentos, se secuenciaron ocho librerías de muestras indexadas de forma única en dos experimentos de secuenciación con el fin de verificar el número mínimo de indexaciones admitido.

Para los experimentos de 96 índices, el valor PPA de SNV osciló entre el 98,7 % y el 100 %, el valor PPA de inserciones y deleciones fue del 100 % y el valor NPA fue también del 100 % para cada una de las combinaciones de 96 índices. Los experimentos de ocho índices obtuvieron valores de PPA del 100 % (SNV, inserciones y deleciones) y valores de NPA del 100 % para cada una de las combinaciones de ocho índices.

Arrastre de muestras

El instrumento NextSeq 550Dx permite la secuenciación de varias muestras, así como los controles en un único experimento de secuenciación. Se realizó un estudio para evaluar el alcance del arrastre de muestras en un experimento de secuenciación (en el experimento) y entre experimentos de secuenciación (entre experimentos). Se analizaron dos muestras de Platinum Genomes, una masculina y una femenina, con un ensayo representativo diseñado para buscar una variedad de genes en 12 588 bases (150 amplicones) de 23 cromosomas diferentes, incluidos los cromosomas de ambos sexos. Se secuenciaron las librerías en el instrumento NextSeq 550Dx con el módulo Germline Variant (variante germinal). El arrastre de muestras masculinas a muestras femeninas se observó por medio de la presencia de lecturas de amplicones de cromosoma Y en muestras femeninas.

El arrastre en el experimento puede introducirse durante la generación de grupos, la llamada de bases de ciclos de índices y el demultiplexado de muestras. Para analizar el arrastre de muestras en un experimento de secuenciación, en el instrumento NextSeq 550Dx, se secuenció una sola vez un grupo de librerías compuesto por 46 duplicados de cada muestra masculina y femenina, además de cuatro controles sin cadena molde. El arrastre de muestras en el experimento se evalúa mediante la comparación de la cobertura de amplicones de cromosoma Y de cada duplicado de muestra femenina con el promedio de cobertura de amplicones de cromosoma Y de todos los duplicados de muestras masculinas del grupo. La mediana observada de arrastre en el experimento fue del 0,084 %.

Para someter a ensayo el arrastre de muestras entre experimentos, se prepararon dos grupos de librerías y se secuenciaron de manera consecutiva en el instrumento NextSeq 550Dx. El primer grupo se compone de 46 duplicados de muestras femeninas, además de dos controles sin cadena molde. El segundo se compone de 46 duplicados de muestras masculinas, además de dos controles sin cadena molde. Ambos grupos emplearon el mismo conjunto de adaptadores de índices. El grupo de muestras femeninas se secuenció primero, seguido de un experimento de secuenciación posterior con el grupo de muestras masculinas. Por último, se repitió otro experimento de secuenciación del grupo de muestras femeninas. El arrastre de muestras entre experimentos se evalúa mediante la comparación de la cobertura de amplicones de cromosoma Y entre los duplicados correspondientes del experimento de repetición del grupo de muestras femeninas y el experimento del grupo de muestras masculinas. La mediana observada de arrastre entre experimentos fue del 0,0076 %.

Entrada de ADN

Sangre (Germline [germinal])

Se estableció el rango de entrada de ADN de sangre para la preparación de librerías de TruSeq Custom Amplicon Kit Dx con el flujo de trabajo del módulo Germline Variant (variante germinal) en el instrumento NextSeq 550Dx. Dicho rango se evaluó llevando a cabo un estudio de dilución en serie con 13 muestras de Platinum Genomes con un ensayo representativo diseñado para buscar varios genes en 12 588 bases de 23 cromosomas diferentes. La librería se secuenció en dos instrumentos NextSeq 550Dx con un lote de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Se analizaron cinco muestras por duplicado en cinco niveles de entrada de ADN que oscilaban entre 250 ng y 12 ng (250 ng, 100 ng, 50 ng, 25 ng y 12 ng). Se analizaron ocho muestras como un único duplicado en cada uno de los cinco niveles de entrada de ADN. Para determinar la exactitud, se compararon genotipos de muestra con la versión 1.0 de Platinum Genomes de 2016. Se determinaron los resultados para cada nivel de entrada. La PPA por cada tipo de variante (SNV, inserciones y deleciones) se representa en la [Tabla 1](#) y la NPA se representa en la [Tabla 2](#). En todos los niveles de entrada se observó una exactitud parecida. La entrada de ADN recomendada para TruSeq Custom Amplicon Kit Dx es de 50 ng, donde los valores 25 ng y 100 ng indican los límites inferior y superior, respectivamente, para cumplir las características de rendimiento.

Tabla 1 Resultados de PPA de cada entrada de ADN por tipo de variante

Entrada de ADN (ng)	Tipo de variante	Variantes previstas	TP	FN	Ausencia de llamadas de variante	PPA (%)
12	SNV	2412	2381	31	0	98,7
25			2404	8	0	99,7
50			2403	9	0	99,6
100			2412	0	0	100
250			2412	0	0	100
12	Inserción	808	784	3	21	99,6
25			781	5	22	99,4
50			786	2	20	99,8
100			786	0	22	100
250			786	0	22	100
12	Delección	758	732	12	14	98,4
25			737	7	14	99,1
50			742	2	14	99,7
100			744	0	14	100
250			744	0	14	100

Tabla 2 NPA de cada entrada de ADN

Entrada de ADN (ng)	TN	FP	Ausencia de llamadas de referencia	NPA (%)
12	430 940	4	26	>99,9
25	430 936	0	34	100
50	430 936	2	32	>99,9
100	430 942	0	28	100
250	430 942	0	28	100

FFPE (Somatic [somática])

Se estableció el rango de entrada de ADN fijado en formol y embebido en parafina (FFPE, Formalin-Fixed Paraffin-Embedded) para la preparación de librerías de TruSeq Custom Amplicon Kit Dx con el flujo de trabajo del módulo Somatic Variant (variante somática) en el instrumento NextSeq 550Dx. Dicho rango de entrada de ADN se evaluó llevando a cabo un estudio de dilución en serie con tres muestras de Platinum Genomes con un ensayo representativo diseñado para buscar varios genes en 12 588 bases de 23 cromosomas diferentes. Las estirpes celulares GM12878 y GM12877 de Platinum Genomes se fijaron en formol y embebieron en parafina y, a continuación, se extrajo ADN. La estirpe celular GM12878 se diluyó con GM12877 de tal manera que las frecuencias alélicas de variantes (VAF) de 81 variantes (55 SNV, 10 inserciones y 16 delecciones) fueran cercanas a 0,025, a 0,05 o a 0,10. Además, cada muestra contaba con 91 variantes con mayores frecuencias de variante de hasta 1,0 de VAF. Las muestras se procesaron por duplicado en cinco niveles de entrada de ADN con una media de ciclos cuantitativos delta (dCq) de 2,1, 3,6, 4,6, 6,0 y 7,8 según lo medido por TruSeq Custom Amplicon Dx - FFPE QC Kit. Cada librería se secuenció en dos instrumentos NextSeq 550Dx con dos lotes de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Para determinar la exactitud, se compararon llamadas de variantes con la versión 1.0 de Platinum Genomes de 2016. La PPA por cada tipo de variante (SNV, inserciones y delecciones) se representa en la [Tabla 3](#) y la NPA se representa en la [Tabla 4](#). La entrada de ADN recomendada para las variantes con una VAF de 0,05 o un valor superior es de $dCq \leq 4$, donde el valor 4,6 indica el límite inferior para cumplir las características de rendimiento.

Tabla 3 Resultados de PPA de cada entrada de ADN por tipo de variante

Media de dCq	Tipo de variante	Variantes previstas	Ausencia de llamadas prevista	VAF de dilución objetivo					
				0,025		0,05		0,10	
				Ausencia de llamadas de variante	PPA (%)	Ausencia de llamadas de variante	PPA (%)	Ausencia de llamadas de variante	PPA (%)
2,1	SNV	808	No aplicable.	196	100	0	100	0	100
3,6				250	99,3	4	100	0	100
4,6				251	94,6	51	99,2	5	100
6,0				257	65,3	213	91,4	100	100
7,8				254	69,3	185	90,7	100	100
2,1	Inserción	264	8	66	96,5	8	100	8	100
3,6				62	97,0	8	100	8	100
4,6				48	96,3	21	100	8	100
6,0				40	80,4	47	98,2	24	95,8
7,8				57	87,0	56	96,2	31	100
2,1	Delección	304	16	58	100	16	100	16	100
3,6				80	100	16	100	16	100
4,6				65	95,4	28	100	16	100
6,0				78	74,8	105	94,0	36	100
7,8				76	75,0	79	95,1	57	98,8

Tabla 4 NPA de cada entrada de ADN

Media de dCq	Estado natural previsto	VAF de dilución objetivo					
		0,025		0,05		0,10	
		Ausencia de llamadas de referencia	NPA (%)	Ausencia de llamadas de referencia	NPA (%)	Ausencia de llamadas de referencia	NPA (%)
2,1	93 688	344	100	260	100	324	100
3,6		400	100	332	100	380	100
4,6		1308	100	1336	100	784	100
6,0		3900	>99,9	3296	>99,9	2996	100
7,8		3020	>99,9	2880	>99,9	2448	>99,9

Sensibilidad analítica (límite del blanco [LoB] y límite de detección [LoD])

Este estudio se llevó a cabo para evaluar el límite del blanco (LoB) y el límite de detección (LoD) del módulo Somatic Variant (variante somática) en el instrumento NextSeq 550Dx. Para ello, se empleó un ensayo representativo diseñado para buscar una variedad de genes en 12 588 bases de 23 cromosomas diferentes. Las estirpes celulares GM12878 y GM12877 de Platinum Genomes se fijaron en formol y embebieron en parafina y, a continuación, se extrajo ADN. La estirpe celular GM12878 se diluyó con GM12877 de tal manera que las frecuencias de variante de 74 variantes (53 SNV, 7 inserciones y 14 delecciones) fueran de $0,05 \pm 0,02$. La estirpe celular GM12877 y la estirpe celular diluida GM12878 (GM12878-D) se analizaron durante seis días de inicio consecutivos con un único instrumento, alternando entre los dos lotes de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), para un total de seis experimentos de secuenciación. Esta prueba dio como resultado 60 duplicados por cada variante en la estirpe celular diluida GM12878-D y 72 duplicados de cada coordenada en estado natural correspondiente en la estirpe celular GM12877 por cada lote de reactivos. El LoB y el LoD se

calcularon usando el enfoque clásico que establece el protocolo CLSI EP17-A2 con la opción no paramétrica. Se calcularon tanto el LoB como el LoD de SNV, inserciones y deleciones de forma independiente por medio de la agrupación de frecuencias de variante por cada tipo de variante determinado. El error de tipo I se definió como 0,01 y el error de tipo II como 0,05.

En el caso del LoB, las frecuencias de variantes agrupadas se clasificaron de la más baja a la más alta y se calculó la posición nonagésima novena de la clasificación para cada lote de reactivos por cada tipo de variante (Tabla 5). El módulo Somatic Variant (variante somática) emplea un corte (LoB eficiente) de VAF de 0,026 para determinar la detección cualitativa de variantes. El LoB que se había calculado confirmó que este corte daba como resultado un error de tipo I de, como máximo, 0,01.

Tabla 5 Límite del blanco

Tipo de variante	Total de observaciones	LoB del Lote de reactivos 1 (%)	LoB del Lote de reactivos 2 (%)
SNV	3816	0,77	0,77
Inserción	504	0,56	0,56
Delección	1008	1,20	1,20

En el caso del LoD, se calculó el porcentaje de frecuencia de mutación individual de cada lote de reactivos por cada tipo de variante que no superaba el corte de 0,026 (Tabla 6). Debido a que los porcentajes eran inferiores al 5 % (0,05) correspondiente al error de tipo II, se calculó la mediana de las frecuencias de variantes combinadas para obtener el LoD (Tabla 6). El LoD por cada tipo de variante se tomó del valor más alto de los dos que se habían calculado para los dos lotes de reactivos: 4,97 % en el caso de las SNV, 5,12 % en el caso de las inserciones y 5,26 % en el caso de las deleciones.

Tabla 6 Límite de detección

Lote de reactivos	Tipo de variante	Total de observaciones	N.º de mediciones de VAF <2,6 %	% de mediciones de VAF <2,6 %	Límite de detección (%)
1	SNV	3180	53	1,7	4,94
	Inserción	420	6	1,4	5,08
	Delección	840	7	0,8	5,22
2	SNV	3180	51	1,6	4,97
	Inserción	420	5	1,2	5,12
	Delección	840	7	0,80	5,26

Exactitud

Germline (germinal)

El siguiente estudio se llevó a cabo para evaluar la exactitud de la llamada de variantes del módulo Germline Variant (variante germinal) en el instrumento NextSeq 550Dx usando NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Se analizaron 13 muestras únicas de Platinum Genomes usando un ensayo representativo diseñado para buscar una variedad de genes en 12 588 bases (150 amplicones) de 23 cromosomas diferentes. Se llevó a cabo un total de nueve experimentos con tres instrumentos de secuenciación, tres lotes de reactivos y tres operadores durante cinco días de inicio. Se determinó la exactitud de las SNV, las inserciones y las deleciones mediante la comparación de los resultados con un método de referencia compuesto y con características bien definidas: la versión 1.0 de Platinum Genomes de 2016. Se definieron las regiones genómicas fiables según este método de referencia, a menos que se indicara lo contrario.

Tabla 7 Resumen de la coincidencia de Germline (germinal)

Criterios	Total de observaciones ¹	Resultado por observación ²	Resultado por experimento ³
PPA de SNV	819	98,7	>99,9
PPA de inserciones	819	95,0	98,9
PPA de deleciones	819	100	100
NPA	819	100	100
OPA	819	>99,9	>99,9

¹Calculado de la siguiente manera: número de muestras por experimento (91) × número de experimentos (9) = 819.

²Valor observado más bajo por duplicado de muestra en los nueve experimentos en total.

³Valor más bajo cuando los datos de cada experimento se analizan en conjunto.

La **Tabla 8** engloba los datos del estudio donde se presentan las coincidencias de porcentaje positivo y negativo por muestra, en los que los resultados de las variantes se comparan con la versión 1.0 de Platinum Genomes de 2016 para llevar a cabo los cálculos de PPA. Se combinan los tres tipos de variantes (SNV, inserciones y deleciones). Dado que el método de referencia solo proporciona resultados de las variantes de nucleótido único y de las inserciones/deleciones, los resultados de la base sin variantes se comparan con el conjunto hg19 de la secuencia de referencia del genoma humano para llevar a cabo los cálculos de NPA.

Tabla 8 Coincidencia de Germline (germinal) por muestra

Muestra	Media de índice de llamada	Variantes previstas ¹	TP	FN	Ausencia de llamadas de variante	TN	FP	PPA	NPA	OPA
NA12877	>99,9	4788	4788	0	0	756 762	0	100	100	100
NA12878	>99,9	8505	8379	1	125	751 464	0	>99,9	100	>99,9
NA12879	>99,9	6048	5985	5	58	757 701	0	99,9	100	>99,9
NA12880	>99,9	6993	6930	0	63	757 638	0	100	100	100
NA12881	>99,9	7875	7811	3	61	751 653	0	>99,9	100	>99,9
NA12882	>99,9	6300	6174	3	123	754 803	0	>99,9	100	>99,9
NA12883	>99,9	7119	7056	0	63	751 905	0	100	100	100
NA12884	>99,9	7182	7119	6	57	754 146	0	99,9	100	>99,9
NA12885	>99,9	7686	7560	2	124	754 173	0	>99,9	100	>99,9
NA12886	>99,9	7245	7182	7	56	752 469	0	99,9	100	>99,9
NA12887	>99,9	7119	7119	0	0	750 645	0	100	100	100
NA12888	>99,9	6804	6804	0	0	756 065	0	100	100	100
NA12893	>99,9	7434	7371	1	62	750 015	0	>99,9	100	>99,9

¹Número total de variantes en todos los duplicados de muestras en los nueve experimentos.

La **Tabla 9** engloba los datos del estudio presentados por muestra, en los que los resultados de las variantes se comparan con el método de referencia compuesto y con características bien definidas. La detección de cada tipo de variante se evalúa de forma independiente: SNV, inserciones y deleciones. Se excluyen las posiciones de referencia.

Tabla 9 Coincidencia de Germline (germinal) por muestra y tipo de variante

>Muestra	SNV			Inserciones			Deleciones		
	>Previsto	>TP	>FN	>Previsto	>TP	>FN	Previsto	TP	FN
NA12877	2331	2331	0	1323	1323	0	1134	1134	0
NA12878	5733	5733	0	1260	1197	1	1512	1449	0
NA12879	3591	3591	0	1323	1260	5	1134	1134	0
NA12880	4221	4221	0	1512	1512	0	1260	1197	0
NA12881	4914	4913	1	1512	1449	2	1449	1449	0
NA12882	3717	3717	0	1386	1323	3	1197	1134	0
NA12883	4284	4284	0	1449	1449	0	1386	1323	0
NA12884	4284	4284	0	1575	1512	6	1323	1323	0
NA12885	4725	4725	0	1575	1512	2	1386	1323	0
NA12886	4347	4347	0	1449	1386	7	1449	1449	0
NA12887	4284	4284	0	1323	1323	0	1512	1512	0
NA12888	4158	4158	0	1449	1449	0	1197	1197	0
NA12893	4599	4599	0	1386	1323	1	1449	1449	0

Las muestras se analizaron con una mayor profundidad para hacer llamadas a inserciones y deleciones de menor tamaño (indels). En la [Tabla 10](#) se incluye un resumen general. Se detectó un total de 71 indels cuyo tamaño oscilaba entre 1 y 24 pb en el caso de las inserciones y entre 1 y 25 pb en el caso de las deleciones.

Tabla 10 Resumen de la detección de indel de Germline (germinal)

Tipo de variante	Variantes previstas	TP	FN	Ausencia de llamadas de variante	PPA
Inserción	18 522	18 018	27	477	99,9
Delección	17 388	17 073	0	315	100

El ensayo representativo incluía 150 amplicones concebidos para cubrir diferentes clases de contenido genómico. El contenido de GC de los amplicones oscilaba entre el 0,19 % y el 0,87 %. Los amplicones también contenían una serie de repeticiones de nucleótido único (por ejemplo, poli-A y poli-T), de dinucleótidos y de trinucleótidos. Los datos se recopilaron por amplicón (Tabla 11) con el objetivo de determinar el efecto del contenido genómico en el porcentaje de llamadas correctas. El porcentaje de llamadas correctas contiene las llamadas de variantes y de referencia y se sitúa por debajo del 100 % si se producen llamadas incorrectas o ausencia de llamadas.

Tabla 11 Exactitud en el nivel del amplicón de Germline (germinal)

Amplicón	Cromosoma	Inicio del amplicón	Final del amplicón	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplicón	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
1	1	36 450 499	36 450 591	93	93	Indel	0,22	76 167	0	0	100
2	1	109 465 122	109 465 200	79	79	Poli-A (5), Poli-C (5), indel	0,38	64 701	0	0	100
3	1	218 353 867	218 353 957	91	91	Indel	0,4	74 529	0	0	100
4	1	223 906 657	223 906 748	92	92	Indel	0,49	75 348	0	0	100
5	1	228 526 602	228 526 682	81	81	Poli-G (5)	0,69	66 339	0	0	100
6	1	236 372 039	236 372 108	70	70	Poli-T (10), indel	0,39	57 330	0	0	100
7	1	247 812 041	247 812 128	88	88	Poli-A (5), CT (3), TAA (3), indel	0,27	72 072	0	0	100
8	2	55 862 774	55 862 863	90	90	Indel	0,28	73 710	0	0	100
9	2	87 003 930	87 004 009	80	80	Indel	0,38	65 520	0	0	100
10	2	177 016 721	177 016 805	85	81	N/A	0,65	66 339	0	0	100
11	2	186 625 727	186 625 801	75	75	Poli-A (8)	0,35	61 425	0	0	100
12	2	190 323 504	190 323 591	88	88	Poli-T (5)	0,42	72 072	0	0	100
13	2	200 796 740	200 796 826	87	87	Poli-T (5), indel	0,31	71 253	0	0	100
14	2	212 245 049	212 245 139	91	91	Poli-T (5), Poli- A (6), indel	0,3	74 529	0	0	100
15	2	228 147 052	228 147 144	93	93	Indel	0,43	76 167	0	0	100
16	2	235 016 350	235 016 422	73	73	Poli-T (5), indel	0,42	59 787	0	0	100

Amplificón	Cromosoma	Inicio del amplificón	Final del amplificón	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificón	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
17	3	4 466 229	4 466 321	93	93	AT (3), indel	0,27	74 823	0	1344	98,2
18	3	46 620 561	46 620 643	83	83	N/A	0,43	67 977	0	0	100
19	3	49 851 331	49 851 400	70	70	CT (3), indel	0,49	57 330	0	0	100
20	3	189 713 161	189 713 248	88	88	Poli-A (5), Poli-T (5), Poli-A (9), TG (3)	0,41	72 072	0	0	100
21	3	190 106 030	190 106 104	75	74	Indel	0,57	60 543	0	63	99,9
22	4	2 233 667	2 233 744	78	78	Poli-A (6)	0,26	63 882	0	0	100
23	4	7 780 541	7 780 637	97	97	Poli-G (6), Poli-T (5), Poli-A (5)	0,42	79 443	0	0	100
24	4	15 688 604	15 688 681	78	78	N/A	0,29	63 882	0	0	100
25	4	56 236 521	56 236 586	66	62	Poli-A (5), indel	0,36	50 778	0	0	100
26	4	102 839 244	102 839 314	71	69	Poli-A (5)	0,46	56 511	0	0	100
27	4	164 446 743	164 446 804	62	62	Poli-A (7), indel	0,27	50 778	0	0	100
28	5	1 882 081	1 882 158	78	75	N/A	0,78	61 425	0	0	100
29	5	14 769 061	14 769 144	84	84	GT (3), CCA (3)	0,62	68 796	0	0	100
30	5	41 069 808	41 069 871	64	64	N/A	0,39	52 416	0	0	100
31	5	74 077 114	74 077 196	83	83	Poli-A (6), indel	0,3	67 977	0	0	100
32	5	147 475 343	147 475 409	67	67	Poli-T (5)	0,37	54 873	0	0	100
33	5	149 323 731	149 323 821	91	91	CT (4), AG (3)	0,55	74 529	0	0	100
34	5	155 662 213	155 662 287	75	75	Indel	0,43	61 425	0	0	100
35	6	6 318 713	6 318 814	102	102	Poli-G (6)	0,68	83 538	0	0	100
36	6	24 949 983	24 950 074	92	92	Indel	0,63	75 348	0	0	100
37	6	31 084 900	31 084 999	100	94	GCT (5), indel	0,61	76 608	0	378	99,5

Amplificón	Cromosoma	Inicio del amplificón	Final del amplificón	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificón	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
38	6	32 147 987	32 148 084	98	98	Poli-T (5), TCT (3), CTT (3)	0,55	80 262	0	0	100
39	6	32 986 864	32 986 958	95	95	Indel	0,53	77 805	0	0	100
40	6	33 408 498	33 408 583	86	86	Poli-C (6)	0,7	70 434	0	0	100
41	6	41 647 401	41 647 495	95	94	Poli-G (5), indel	0,61	76 986	0	0	100
42	6	112 435 865	112 435 955	91	91	Poli-A (5)	0,44	74 529	0	0	100
43	7	22 202 076	22 202 148	73	73	N/A	0,44	59 787	0	0	100
44	7	66 276 100	66 276 187	88	88	Indel	0,35	72 072	0	0	100
45	7	77 365 735	77 365 821	87	87	Poli-A (7), AG (4)	0,26	71 253	0	0	100
46	7	110 939 946	110 940 030	85	85	Indel	0,38	69 615	0	0	100
47	7	128 533 468	128 533 557	90	90	Poli-G (5), indel	0,62	73 710	0	0	100
48	7	149 503 875	149 503 965	91	91	Poli-G (6), Poli-C (6), indel	0,71	74 529	0	0	100
49	7	154 404 519	154 404 599	81	66	N/A	0,31	54 054	0	0	100
50	7	156 476 507	156 476 599	93	93	Indel	0,35	76 167	0	0	100
51	8	1 817 312	1 817 394	83	83	N/A	0,42	67 977	0	0	100
52	8	24 811 020	24 811 109	90	89	Poli-G (7), CTC (4), indel	0,61	72 171	0	720	99,0
53	8	76 518 625	76 518 691	67	67	Indel	0,3	54 873	0	0	100
54	9	103 054 909	103 055 006	98	98	Poli-G (6)	0,67	80 262	0	0	100
55	9	105 586 150	105 586 214	65	65	Indel	0,32	53 235	0	0	100
56	9	107 620 823	107 620 918	96	96	N/A	0,49	78 624	0	0	100
57	9	123 769 149	123 769 231	83	83	AT (3)	0,37	67 977	0	0	100
58	9	138 995 345	138 995 441	97	97	Poli-C (6), indel	0,68	79 443	0	0	100

Amplificación	Cromosoma	Inicio del amplificación	Final del amplificación	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificación	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
59	10	5 987 120	5 987 198	79	78	Poli-G (5), indel	0,47	63 882	0	0	100
60	10	11 784 629	11 784 726	98	91	GC(3)	0,87	74 529	0	0	100
61	10	27 317 777	27 317 855	79	79	Poli-T (5)	0,3	64 701	0	0	100
62	10	33 018 351	33 018 440	90	90	Poli-A (5), Poli-T (5)	0,2	73 710	0	0	100
63	10	45 084 159	45 084 253	95	95	Indel	0,35	77 805	0	0	100
64	10	55 892 599	55 892 687	89	88	AC (11), indel	0,42	71 747	0	325	99,5
65	10	101 611 250	101 611 329	80	80	N/A	0,49	65 520	0	0	100
66	10	118 351 373	118 351 453	81	81	N/A	0,51	66 339	0	0	100
67	11	8 159 816	8 159 912	97	96	N/A	0,45	78 624	0	0	100
68	11	30 177 648	30 177 717	70	70	Indel	0,46	57 330	0	0	100
69	11	47 470 345	47 470 444	100	100	N/A	0,65	81 900	0	0	100
70	11	59 837 679	59 837 740	62	62	Indel	0,37	50 778	0	0	100
71	11	64 418 856	64 418 957	102	102	N/A	0,59	83 538	0	0	100
72	11	93 529 612	93 529 684	73	73	Poli-A (5)	0,4	59 787	0	0	100
73	11	101 347 052	101 347 136	85	85	N/A	0,42	69 615	0	0	100
74	11	102 477 336	102 477 426	91	91	Poli-G (6)	0,55	74 529	0	0	100
75	11	118 406 285	118 406 369	85	85	Indel	0,53	69 615	0	0	100
76	11	120 357 801	120 357 885	85	85	Poli-A (5), CA (3), indel	0,34	69 615	0	0	100
77	11	125 769 313	125 769 397	85	85	GA(3)	0,52	69 615	0	0	100
78	12	2 834 770	2 834 853	84	84	Poli-C (5), indel	0,52	68 796	0	0	100
79	12	26 811 004	26 811 096	93	93	Poli-A (7), AC (4)	0,33	76 167	0	0	100
80	12	30 881 766	30 881 846	81	81	N/A	0,49	66 339	0	0	100
81	12	88 474 105	88 474 175	71	71	Poli-A (6)	0,35	58 149	0	0	100

Amplificón	Cromosoma	Inicio del amplificón	Final del amplificón	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificón	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
82	12	120 966 872	120 966 966	95	95	Poli-G (5)	0,68	77 805	0	0	100
83	13	24 167 504	24 167 576	73	73	N/A	0,52	59 787	0	0	100
84	13	25 816 961	25 817 049	89	88	Poli-A (5), Poli-T (7), Poli-A (7), indel	0,22	72 072	0	0	100
85	13	44 880 112	44 880 200	89	89	Indel	0,49	72 891	0	0	100
86	13	77 665 218	77 665 294	77	77	Indel	0,39	63 063	0	0	100
87	14	31 619 327	31 619 393	67	67	GA(3),TA(3)	0,39	54 873	0	0	100
88	14	39 517 884	39 517 966	83	83	N/A	0,25	67 977	0	0	100
89	14	46 958 962	46 959 034	73	72	Poli-T (5), indel	0,19	58 642	0	326	99,4
90	14	58 050 030	58 050 110	81	81	Indel	0,38	66 339	0	0	100
91	14	82 390 559	82 390 649	91	91	Indel	0,35	74 529	0	0	100
92	14	92 549 544	92 549 609	66	66	Poli-A (5)	0,41	54 054	0	0	100
93	14	102 808 496	102 808 589	94	94	Indel	0,62	76 986	0	0	100
94	15	43 170 751	43 170 848	98	96	Poli-C (5)	0,45	78 624	0	0	100
95	15	63 446 149	63 446 216	68	68	Indel	0,25	55 692	0	0	100
96	15	77 879 807	77 879 901	95	93	Poli-G (5), indel	0,68	76 167	0	0	100
97	15	81 625 334	81 625 428	95	95	Poli-T (6)	0,43	77 805	0	0	100
98	15	85 438 263	85 438 334	72	71	Indel	0,65	58 149	0	0	100
99	15	89 817 413	89 817 503	91	91	N/A	0,36	74 529	0	0	100
100	15	89 864 274	89 864 343	70	70	Indel	0,56	57 330	0	0	100
101	16	1 894 910	1 894 972	63	63	N/A	0,27	51 597	0	0	100
102	16	28 997 904	28 997 998	95	95	Poli-C (5)	0,67	77 805	0	0	100
103	16	53 682 908	53 682 994	87	87	TA(3)	0,41	71 253	0	0	100
104	16	57 954 406	57 954 509	104	104	Poli-C (5)	0,67	85 176	0	0	100

Amplificón	Cromosoma	Inicio del amplificón	Final del amplificón	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificón	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
105	16	85 706 375	85 706 465	91	91	Poli-T (5), indel	0,37	74 529	0	0	100
106	17	3 563 920	3 564 008	89	89	GC(3)	0,64	72 891	0	0	100
107	17	3 594 191	3 594 277	87	87	Poli-C (5), indel	0,67	71 247	0	6	100
108	17	3 970 090	3 970 180	91	91	Indel	0,46	74 529	0	0	100
109	17	16 084 945	16 085 037	93	93	Indel	0,26	76 167	0	0	100
110	17	33 998 759	33 998 849	91	89	Poli-T (5)	0,54	72 891	0	0	100
111	17	39 589 691	39 589 774	84	82	Poli-A (13), indel (x2)	0,29	66 343	27	788	98,8
112	17	41 244 394	41 244 484	91	91	Poli-A (5)	0,34	74 529	0	0	100
113	17	45 438 866	45 438 957	92	92	Poli-A (7), AT (3), AT (4), AT (4), indel	0,26	75 348	0	0	100
114	17	61 502 432	61 502 510	79	79	Indel	0,41	64 413	0	288	99,6
115	17	64 023 582	64 023 667	86	86	Poli-T (7)	0,22	70 434	0	0	100
116	17	72 308 237	72 308 320	84	84	GAG(3)	0,62	68 796	0	0	100
117	18	2 616 456	2 616 522	67	67	GA(3)	0,31	54 873	0	0	100
118	18	6 980 478	6 980 568	91	91	N/A	0,37	74 529	0	0	100
119	18	9 888 026	9 888 094	69	69	Poly A (6), TG (3)	0,43	56 511	0	0	100
120	18	38 836 999	38 837 073	75	75	Poli-A (5), indel	0,37	61 425	0	0	100
121	18	47 405 382	47 405 462	81	81	CTC(3), indel	0,47	66 339	0	0	100
122	18	54 815 665	54 815 749	85	85	CT (3), indel	0,45	69 615	0	0	100
123	18	59 773 996	59 774 060	65	65	N/A	0,48	53 235	0	0	100
124	19	625 143	625 241	99	99	N/A	0,59	81 081	0	0	100
125	19	18 121 418	18 121 491	74	74	N/A	0,68	60 605	1	0	100
126	19	18 186 574	18 186 643	70	70	N/A	0,64	57 330	0	0	100

Amplificón	Cromosoma	Inicio del amplificón	Final del amplificón	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificón	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
127	20	746 056	746 149	94	94	N/A	0,61	76 986	0	0	100
128	20	10 633 195	10 633 276	82	82	AC(3)	0,59	67 158	0	0	100
129	20	17 705 633	17 705 708	76	76	CT(3)	0,58	62 244	0	0	100
130	20	21 766 821	21 766 890	70	70	GT(3),TG(4), indel	0,46	57 330	0	0	100
131	20	25 278 421	25 278 521	101	101	Indel	0,63	82 719	0	0	100
132	20	50 897 302	50 897 368	67	67	Indel	0,36	54 873	0	0	100
133	20	62 331 904	62 331 994	91	88	Poli-G (6)	0,73	72 072	0	0	100
134	20	62 690 860	62 690 946	87	87	Indel	0,57	71 253	0	0	100
135	21	30 300 823	30 300 888	66	66	Indel	0,35	54 054	0	0	100
136	21	33 694 176	33 694 273	98	98	Poli-T (6), CA (3)	0,54	80 262	0	0	100
137	21	36 710 706	36 710 792	87	87	GT(3), indel	0,39	71 253	0	0	100
138	21	46 644 924	46 644 992	69	69	Poli-A (6), AG (3), indel	0,32	56 439	0	72	99,9
139	21	46 705 575	46 705 664	90	90	Poli-T (5), Poli-A (6)	0,5	73 710	0	0	100
140	22	25 750 774	25 750 873	100	100	Indel	0,63	81 900	0	0	100
141	22	32 439 233	32 439 329	97	97	N/A	0,68	79 443	0	0	100
142	22	37 409 844	37 409 940	97	97	Indel	0,46	79 443	0	0	100
143	22	37 637 596	37 637 694	99	99	N/A	0,6	81 081	0	0	100
144	22	47 081 347	47 081 438	92	92	Indel	0,66	75 348	0	0	100
145	X	15 870 424	15 870 492	69	69	Poli-T (5)	0,26	56 511	0	0	100
146	X	135 288 543	135 288 611	69	69	Poli-C (5)	0,62	56 511	0	0	100
147	X	135 290 777	135 290 847	71	71	N/A	0,52	58 149	0	0	100
148	Y	2 655 397	2 655 461	65	0	N/A	0,55	0	0	0	N/A
149	Y	2 655 519	2 655 609	91	0	N/A	0,48	0	0	0	N/A
150	Y	2 655 609	2 655 679	71	0	Poli-A (5)	0,37	0	0	0	N/A

Los resultados de la secuenciación de la muestra NA12878 se compararon con un genotipo muy fiable establecido para la NA12878 por los Institutos Nacionales de Normas y Tecnología (NIST, National Institutes of Standards and Technology) (v.2.19). De los 150 amplicones, 92 estaban incluidos completamente en las regiones genómicas muy fiables, 41 coincidían parcialmente y 17 no coincidían en absoluto con la secuencia de los NIST. El resultado fueron 10 000 coordenadas por duplicado para llevar a cabo la comparación. Las llamadas de bases sin variantes se compararon con el conjunto hg19 de la secuencia de referencia del genoma humano. Los resultados de exactitud se muestran en la [Tabla 12](#).

Tabla 12 Coincidencia de Germline (germinal) de la muestra NA12878 con la base de datos de los NIST

Muestra	N.º de amplicones	Media de índice de llamada	TP	FN	TN	FP	PPA	NPA	OPA
NA12878	133	>99,9	6552	1	610 470	0	>99,9	100	>99,9

Según los datos proporcionados por este estudio de Germline (germinal) de nueve experimentos, el instrumento NextSeq 550Dx puede secuenciar de forma sistemática:

- ▶ Contenido de GC ≥ 19 % (se llamaron correctamente todas las bases llamadas de 819 amplicones secuenciados con el 19 % de contenido de GC y se obtuvo un índice de ausencia de llamadas del 0,6 %)
- ▶ Contenido de GC ≤ 87 % (se llamaron correctamente todas las bases llamadas de 819 amplicones secuenciados con el 87 % de contenido de GC y no hubo ninguna ausencia de llamadas)
- ▶ Longitudes de poli-A ≤ 9 (se llamaron correctamente todas las bases llamadas de 819 amplicones secuenciados que contenían una repetición de poli-A de nueve nucleótidos y no hubo ninguna ausencia de llamadas)
- ▶ Longitudes de poli-T ≤ 10 (se llamaron correctamente todas las bases llamadas de 819 amplicones secuenciados que contenían una repetición de poli-T de diez nucleótidos y no hubo ninguna ausencia de llamadas)
- ▶ Longitudes de poli-G ≤ 7 (se llamaron correctamente todas las bases llamadas de 819 amplicones secuenciados que contenían una repetición de poli-G de siete nucleótidos y se obtuvo un índice de ausencia de llamadas del 1,0 %)
- ▶ Longitudes de poli-C ≤ 6 (se llamaron correctamente todas las bases llamadas de 2457 amplicones secuenciados que contenían una repetición de poli-C de seis nucleótidos y no hubo ninguna ausencia de llamadas)
- ▶ Longitudes de repetición de dinucleótidos $\leq 11\times$ (se llamaron correctamente todas las bases llamadas de 819 amplicones secuenciados que contenían una repetición de dinucleótidos de $11\times$ y se obtuvo un índice de ausencia de llamadas del 0,5 %)
- ▶ Longitudes de repetición de trinucleótidos $\leq 5\times$ (se llamaron correctamente todas las bases llamadas de 819 amplicones secuenciados que contenían una repetición de trinucleótidos de $5\times$ y se obtuvo un índice de ausencia de llamadas del 0,5 %)
- ▶ Longitudes de inserción ≤ 24 (se llamaron correctamente 66 343 de 66 370 bases llamadas de 819 amplicones secuenciados que contenían una inserción de 24 nucleótidos y se obtuvo un índice de ausencia de llamadas del 1,2 %; no se produjeron llamadas incorrectas en la región que contenía la inserción de 24 nucleótidos)
- ▶ Longitudes de delección ≤ 25 (se llamaron correctamente todas las bases llamadas de 2457 amplicones secuenciados que contenían una delección de 25 nucleótidos y no hubo ninguna ausencia de llamadas)

Somatic (somática)

El estudio descrito, a continuación, se usó para evaluar la exactitud de la llamada de variantes del módulo Somatic Variant (variante somática) en el instrumento NextSeq 550Dx usando NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

En este estudio se empleó un ensayo representativo diseñado para buscar una variedad de genes en 12 588 bases (150 amplicones) de 23 cromosomas diferentes. El ADN de Platinum Genomes se extrajo de bloques modificados con FFPE para generar seis muestras únicas para su evaluación en el estudio.

El ADN de la muestra GM12877 se diluyó con el ADN de la muestra GM12878 para crear GM12877-D5 y GM12877-D7 como un conjunto de variantes heterocigóticas únicas con frecuencias de variante próximas a entre el 5 % y el 7 %. El ADN de la muestra GM12878 se diluyó del mismo modo con el ADN de la muestra GM12877 para crear GM12878-D5 y GM12878-D7. Cada una de las muestras se analizó por triplicado, excepto las muestras diluidas, que se analizaron por sextuplicado. Se llevó a cabo un total de nueve experimentos con tres instrumentos de secuenciación, tres lotes de reactivos y tres operadores durante cinco días de inicio. Se determinó la exactitud de las SNV, las inserciones y las deleciones mediante la comparación de los resultados con un método de referencia compuesto y con características bien definidas: la versión 1.0 de Platinum Genomes de 2016. Las regiones genómicas de confianza se definieron teniendo en cuenta este método de referencia, a menos que se especifique lo contrario.

Tabla 13 Resumen de la coincidencia de Somatic (somática)

Crterios	Total de observaciones ¹	Resultado por observación ²	Resultado por experimento ³
PPA de SNV	378	98,9	99,9
PPA de inserciones	378	96,9	99,9
PPA de deleciones	378	97,1	99,9
NPA	378	>99,9	>99,9
OPA	378	>99,9	>99,9

¹Calculado de la siguiente manera: número de muestras por experimento (42) × número de experimentos (9) = 378.

²Valor observado más bajo por duplicado de muestra en los nueve experimentos en total.

³Valor más bajo cuando los datos de cada experimento se analizan en conjunto.

La **Tabla 14** engloba los datos del estudio donde se presentan las coincidencias de porcentaje positivo y negativo por muestra, donde los resultados de las variantes se comparan con el método de referencia compuesto y con características bien definidas que se usa para llevar a cabo los cálculos de PPA. Se combinan los tres tipos de variantes (SNV, inserciones y deleciones). Dado que el método de referencia solo proporciona resultados de las variantes de nucleótido único y de las inserciones/deleciones, los resultados de la base sin variantes se comparan con el conjunto hg19 de la secuencia de referencia del genoma humano para llevar a cabo los cálculos de NPA.

Tabla 14 Coincidencia de Somatic (somática) por muestra

Muestra	Media de índice de llamada	Previsto	TP	FN	Ausencia de llamadas de variante	TN	FP	PPA	NPA	OPA
GM12877	98,7	2052	2025	0	27	318 682	15	100	>99,9	>99,9
GM12878	98,8	3645	3564	0	81	317 645	0	100	100	100
GM12879	99,8	2592	2538	0	54	323 614	2	100	>99,9	>99,9
GM12884	99,8	3078	3024	0	54	322 038	5	100	>99,9	>99,9
GM12885	99,8	3294	3213	0	81	322 121	0	100	100	100
GM12888	99,8	2916	2889	0	27	323 048	2	100	>99,9	>99,9
GM12877-D5	99,8	9288	8930	0	358	630 621	0	100	100	100
GM12877-D7	99,7	9288	9032	0	256	629 719	0	100	100	100
GM12878-D5	99,5	9288	8699	42	547	628 582	0	99,5	100	>99,9
GM12878-D7	99,7	9288	9108	0	180	629 803	0	100	100	100

La **Tabla 15** engloba los datos del estudio presentados por muestra, en los que los resultados de las variantes se comparan con el método de referencia compuesto y con características bien definidas. La detección de cada tipo de variante se evalúa de forma independiente: SNV, inserciones y deleciones. Se excluyen las posiciones de referencia.

Tabla 15 Coincidencia de Somatic (somática) por muestra y tipo de variante

Muestra	SNV			Inserciones			Deleciones		
	Previsto	TP	FN	Previsto	TP	FN	Previsto	TP	FN
GM12877	999	999	0	567	567	0	486	459	0
GM12878	2457	2457	0	540	513	0	648	594	0
GM12879	1539	1539	0	567	540	0	486	459	0
GM12884	1836	1836	0	675	648	0	567	540	0
GM12885	2025	2025	0	675	648	0	594	540	0
GM12888	1782	1782	0	621	621	0	513	486	0
GM12877-D5	5454	5392	0	1782	1647	0	2052	1891	0
GM12877-D7	5454	5406	0	1782	1728	0	2052	1898	0
GM12878-D5	5454	5192	28	1782	1651	9	2052	1856	5
GM12878-D7	5454	5445	0	1782	1719	0	2052	1944	0

Las diez muestras se analizaron con una mayor profundidad para hacer llamadas a inserciones y deleciones de menor tamaño (indels) (Tabla 16). Se detectó un total de 71 indels cuyo tamaño oscilaba entre 1 y 24 pb en el caso de las inserciones y entre 1 y 25 pb en el caso de las deleciones.

Tabla 16 Resumen de la detección de indel de Somatic (somática)

Tipo de variante	Variante previstas	TP	FN	Ausencia de llamadas de variante	PPA
Inserción	10 773	10 282	9	482	99,2
Delección	11 502	10 667	5	830	>99,9

Los 150 amplicones se diseñaron para cubrir diferentes clases de contenido genómico. El contenido de GC de los amplicones oscilaba entre el 0,19 % y el 0,87 %. Los amplicones también contenían una serie de repeticiones de nucleótido único (por ejemplo, poli-A y poli-T), de dinucleótidos y de trinucleótidos. Los datos se recopilaron por amplicón (Tabla 17) con el objetivo de determinar el efecto del contenido genómico en el porcentaje de llamadas correctas. El porcentaje de llamadas correctas contiene las llamadas de variantes y de referencia y se sitúa por debajo del 100 % si se producen llamadas incorrectas o ausencia de llamadas.

Tabla 17 Exactitud en el nivel del amplicón de Somatic (somática)

Amplicón	Cromosoma	Inicio del amplicón	Final del amplicón	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplicón	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
1	1	36 450 499	36 450 591	93	93	Indel	0,22	35 066	0	88	99,7
2	1	109 465 122	109 465 200	79	79	Poli-A (5), Poli-C (5), indel	0,38	29 827	0	35	99,9
3	1	218 353 867	218 353 957	91	91	Indel	0,4	34 202	0	283	99,2
4	1	223 906 657	223 906 748	92	92	Indel	0,49	34 613	0	163	99,5
5	1	228 526 602	228 526 682	81	81	Poli-G (5)	0,69	30 571	0	47	99,8
6	1	236 372 039	236 372 108	70	70	Poli-T (10), indel	0,39	26 452	0	8	100,0
7	1	247 812 041	247 812 128	88	88	Poli-A (5), CT (3), TAA (3), indel	0,27	33 148	0	116	99,7
8	2	55 862 774	55 862 863	90	90	Indel	0,28	33 928	0	92	99,7
9	2	87 003 930	87 004 009	80	80	Indel	0,38	30 218	0	22	99,9
10	2	177 016 721	177 016 805	85	81	N/A	0,65	30 616	0	2	>99,9
11	2	186 625 727	186 625 801	75	75	Poli-A (8)	0,35	28 017	0	499	98,3
12	2	190 323 504	190 323 591	88	88	Poli-T (5)	0,42	33 207	0	57	99,8
13	2	200 796 740	200 796 826	87	87	Poli-T (5), indel	0,31	32 524	9	718	97,8
14	2	212 245 049	212 245 139	91	91	Poli-T (5), Poli-A (6), indel	0,3	33 972	0	456	98,7
15	2	228 147 052	228 147 144	93	93	N/A	0,43	35 051	0	103	99,7

Amplificación	Cromosoma	Inicio del amplificación	Final del amplificación	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificación	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
16	2	235 016 350	235 016 422	73	73	Poli-T (5), indel	0,42	27 459	0	136	99,5
17	3	4 466 229	4 466 321	93	93	AT (3), indel	0,27	34 534	0	620	98,2
18	3	46 620 561	46 620 643	83	83	N/A	0,43	31 339	0	44	99,9
19	3	49 851 331	49 851 400	70	70	CT (3), indel	0,49	26 373	0	87	99,7
20	3	189 713 161	189 713 248	88	88	Poli-A (5), Poli-T (5), Poli-A (9), TG (3)	0,41	32 829	0	857	97,5
21	3	190 106 030	190 106 104	75	74	Indel	0,57	27 925	0	47	99,8
22	4	2 233 667	2 233 744	78	78	Poli-A (6)	0,26	29 327	4	162	99,4
23	4	7 780 541	7 780 637	97	97	Poli-G (6), Poli-T (5), Poli-A (5)	0,42	36 585	0	117	99,7
24	4	15 688 604	15 688 681	78	78	N/A	0,29	29 427	0	57	99,8
25	4	56 236 521	56 236 586	66	62	Poli-A (5), indel	0,36	23 356	5	75	99,7
26	4	102 839 244	102 839 314	71	69	Poli-A (5)	0,46	25 942	0	140	99,5
27	4	164 446 743	164 446 804	62	62	Poli-A (7), indel	0,27	22 944	0	560	97,6
28	5	1 882 081	1 882 158	78	75	N/A	0,78	28 299	0	53	99,8
29	5	14 769 061	14 769 144	84	84	GT (3), CCA (3)	0,62	31 658	0	94	99,7
30	5	41 069 808	41 069 871	64	64	N/A	0,39	24 120	0	72	99,7
31	5	74 077 114	74 077 196	83	83	Poli-A (6), indel	0,3	31 297	0	77	99,8
32	5	147 475 343	147 475 409	67	67	Poli-T (5)	0,37	25 277	0	55	99,8
33	5	149 323 731	149 323 821	91	91	CT (4), AG (3)	0,55	34 308	0	90	99,7

Amplificación	Cromosoma	Inicio del amplificación	Final del amplificación	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificación	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
34	5	155 662 213	155 662 287	75	75	Indel	0,43	28 266	0	163	99,4
35	6	6 318 713	6 318 814	102	102	Poli-G (6)	0,68	38 489	0	67	99,8
36	6	24 949 983	24 950 074	92	92	Indel	0,63	34 730	0	46	99,9
37	6	31 084 900	31 084 999	100	94	GCT (5), indel	0,61	35 057	0	483	98,6
38	6	32 147 987	32 148 084	98	98	Poli-T (5), TCT (3), CTT (3)	0,55	36 647	0	406	98,9
39	6	32 986 864	32 986 958	95	95	Indel	0,53	35 681	0	238	99,3
40	6	33 408 498	33 408 583	86	86	Poli-C (6)	0,7	32 438	0	70	99,8
41	6	41 647 401	41 647 495	95	94	Poli-G (5), indel	0,61	35 441	0	91	99,7
42	6	112 435 865	112 435 955	91	91	Poli-A (5)	0,44	34 354	0	44	99,9
43	7	22 202 076	22 202 148	73	73	N/A	0,44	27 575	0	28	99,9
44	7	66 276 100	66 276 187	88	88	Indel	0,35	33 060	0	213	99,4
45	7	77 365 735	77 365 821	87	87	Poli-A (7), AG (4)	0,26	32 423	0	489	98,5
46	7	110 939 946	110 940 030	85	85	Indel	0,38	32 074	0	56	99,8
47	7	128 533 468	128 533 557	90	90	Poli-G (5), indel	0,62	33 791	0	281	99,2
48	7	149 503 875	149 503 965	91	91	Poli-G (6), Poli-C (6), indel	0,71	34 316	0	82	99,8
49	7	154 404 519	154 404 599	81	66	N/A	0,31	24 901	0	47	99,8
50	7	156 476 507	156 476 599	93	93	Indel	0,35	35 067	0	87	99,8
51	8	1 817 312	1 817 394	83	83	N/A	0,42	31 365	0	9	>99,9
52	8	24 811 020	24 811 109	90	89	Poli-G (7), CTC (4), indel	0,61	32 781	0	890	97,4

Amplión	Cromosoma	Inicio del amplión	Final del amplión	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplión	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
53	8	76 518 625	76 518 691	67	67	Indel	0,3	25 228	0	146	99,4
54	9	103 054 909	103 055 006	98	98	Poli-G (6)	0,67	36 968	0	76	99,8
55	9	105 586 150	105 586 214	65	65	Indel	0,32	24 472	0	100	99,6
56	9	107 620 823	107 620 918	96	96	N/A	0,49	36 203	0	85	99,8
57	9	123 769 149	123 769 231	83	83	AT (3)	0,37	31 329	0	45	99,9
58	9	138 995 345	138 995 441	97	97	Poli-C (6), indel	0,68	36 472	0	201	99,5
59	10	5 987 120	5 987 198	79	78	Poli-G (5), indel	0,47	29 473	0	11	>99,9
60	10	11 784 629	11 784 726	98	91	GC(3)	0,87	34 188	0	213	99,4
61	10	27 317 777	27 317 855	79	79	Poli-T (5)	0,3	29 843	0	19	99,9
62	10	33 018 351	33 018 440	90	90	Poli-A (5), Poli-T (5)	0,2	33 968	0	68	99,8
63	10	45 084 159	45 084 253	95	95	Indel	0,35	35 829	0	81	99,8
64	10	55 892 599	55 892 687	89	88	AC (11), indel	0,42	32 098	88	2048	93,8
65	10	101 611 250	101 611 329	80	80	N/A	0,49	30 217	0	28	99,9
66	10	118 351 373	118 351 453	81	81	N/A	0,51	30 531	0	96	99,7
67	11	8 159 816	8 159 912	97	96	N/A	0,45	36 105	0	192	99,5
68	11	30 177 648	30 177 717	70	70	Indel	0,46	26 318	0	153	99,4
69	11	47 470 345	47 470 444	100	100	N/A	0,65	37 785	0	24	99,9
70	11	59 837 679	59 837 740	62	62	Indel	0,37	23 368	0	68	99,7
71	11	64 418 856	64 418 957	102	102	N/A	0,59	38 546	0	10	>99,9
72	11	93 529 612	93 529 684	73	73	Poli-A (5)	0,4	27 516	0	78	99,7
73	11	101 347 052	101 347 136	85	85	N/A	0,42	32 083	0	48	99,9
74	11	102 477 336	102 477 426	91	91	Poli-G (6)	0,55	34 047	0	369	98,9
75	11	118 406 285	118 406 369	85	85	Indel	0,53	32 065	0	74	99,8

Amplificación	Cromosoma	Inicio del amplificación	Final del amplificación	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificación	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
76	11	120 357 801	120 357 885	85	85	Poli-A (5), CA (3), indel	0,34	32 083	0	47	99,9
77	11	125 769 313	125 769 397	85	85	GA(3)	0,52	32 103	0	27	99,9
78	12	2 834 770	2 834 853	84	84	Poli-C (5), indel	0,52	31 645	16	525	98,3
79	12	26 811 004	26 811 096	93	93	Poli-A (7), AC (4)	0,33	34 824	0	330	99,1
80	12	30 881 766	30 881 846	81	81	N/A	0,49	30 497	0	121	99,6
81	12	88 474 105	88 474 175	71	71	Poli-A (6)	0,35	26 773	0	65	99,8
82	12	120 966 872	120 966 966	95	95	Poli-G (5)	0,68	35 830	9	72	99,8
83	13	24 167 504	24 167 576	73	73	N/A	0,52	27 498	0	114	99,6
84	13	25 816 961	25 817 049	89	88	Poli-A (5), Poli-T (7), Poli-A (7), indel	0,22	32 824	0	566	98,3
85	13	44 880 112	44 880 200	89	89	Indel	0,49	33 574	0	77	99,8
86	13	77 665 218	77 665 294	77	77	Indel	0,39	29 075	0	31	99,9
87	14	31 619 327	31 619 393	67	67	GA(3),TA(3)	0,39	25 313	0	13	99,9
88	14	39 517 884	39 517 966	83	83	N/A	0,25	31 360	0	22	99,9
89	14	46 958 962	46 959 034	73	72	Poli-T (5), indel	0,19	26 499	0	717	97,4
90	14	58 050 030	58 050 110	81	81	Indel	0,38	30 494	0	133	99,6
91	14	82 390 559	82 390 649	91	91	Indel	0,35	34 313	0	86	99,7
92	14	92 549 544	92 549 609	66	66	Poli-A (5)	0,41	24 555	0	1527	94,1
93	14	102 808 496	102 808 589	94	94	Indel	0,62	35 472	0	69	99,8
94	15	43 170 751	43 170 848	98	96	Poli-C (5)	0,45	36 264	0	24	99,9
95	15	63 446 149	63 446 216	68	68	Indel	0,25	25 667	0	37	99,9
96	15	77 879 807	77 879 901	95	93	Poli-G (5), indel	0,68	34 745	0	432	98,8

Amplificación	Cromosoma	Inicio del amplificación	Final del amplificación	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificación	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
97	15	81 625 334	81 625 428	95	95	Poli-T (6)	0,43	35 870	0	40	99,9
98	15	85 438 263	85 438 334	72	71	Indel	0,65	26 762	0	76	99,7
99	15	89 817 413	89 817 503	91	91	N/A	0,36	34 286	0	112	99,7
100	15	89 864 274	89 864 343	70	70	Indel	0,56	26 449	0	11	>99,9
101	16	1 894 910	1 894 972	63	63	N/A	0,27	23 809	0	5	>99,9
102	16	28 997 904	28 997 998	95	95	Poli-C (5)	0,67	35 860	0	50	99,9
103	16	53 682 908	53 682 994	87	87	TA(3)	0,41	32 835	0	60	99,8
104	16	57 954 406	57 954 509	104	104	Poli-C (5)	0,67	39 177	0	144	99,6
105	16	85 706 375	85 706 465	91	91	Poli-T (5), indel	0,37	34 075	0	323	99,1
106	17	3 563 920	3 564 008	89	89	GC(3)	0,64	33 632	0	11	>99,9
107	17	3 594 191	3 594 277	87	87	Poli-C (5), indel	0,67	32 752	0	134	99,6
108	17	3 970 090	3 970 180	91	91	Indel	0,46	34 343	0	82	99,8
109	17	16 084 945	16 085 037	93	93	Indel	0,26	35 077	0	78	99,8
110	17	33 998 759	33 998 849	91	89	Poli-T (5)	0,54	33 553	0	89	99,7
111	17	39 589 691	39 589 774	84	82	Poli-A (13), indel (x2)	0,29	30 554	53	2296	92,9
112	17	41 244 394	41 244 484	91	91	Poli-A (5)	0,34	34 360	0	38	99,9
113	17	45 438 866	45 438 957	92	92	Poli-A (7), AT (3), AT (4), AT (4), indel	0,26	34 367	0	418	98,8
114	17	61 502 432	61 502 510	79	79	Indel	0,41	29 751	0	119	99,6
115	17	64 023 582	64 023 667	86	86	Poli-T (7)	0,22	32 176	0	340	99,0
116	17	72 308 237	72 308 320	84	84	GAG(3)	0,62	31 604	7	141	99,5
117	18	2 616 456	2 616 522	67	67	GA(3)	0,31	25 273	8	45	99,8
118	18	6 980 478	6 980 568	91	91	N/A	0,37	34 386	0	12	>99,9

Amplificación	Cromosoma	Inicio del amplificación	Final del amplificación	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplificación	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
119	18	9 888 026	9 888 094	69	69	Poly A (6), TG(3)	0,43	25 692	0	399	98,5
120	18	38 836 999	38 837 073	75	75	Poli-A (5), indel	0,37	27 923	0	893	96,9
121	18	47 405 382	47 405 462	81	81	CTC(3), indel	0,47	30 598	0	20	99,9
122	18	54 815 665	54 815 749	85	85	CT (3), indel	0,45	31 969	0	161	99,5
123	18	59 773 996	59 774 060	65	65	N/A	0,48	24 531	0	48	99,8
124	19	625 143	625 241	99	99	N/A	0,59	37 298	0	124	99,7
125	19	18 121 418	18 121 491	74	74	N/A	0,68	27 881	0	109	99,6
126	19	18 186 574	18 186 643	70	70	N/A	0,64	26 442	0	26	99,9
127	20	746 056	746 149	94	94	N/A	0,61	35 501	0	31	99,9
128	20	10 633 195	10 633 276	82	82	AC(3)	0,59	30 951	0	72	99,8
129	20	17 705 633	17 705 708	76	76	CT(3)	0,58	28 686	0	42	99,9
130	20	21 766 821	21 766 890	70	70	GT(3),TG(4), indel	0,46	26 372	0	88	99,7
131	20	25 278 421	25 278 521	101	101	Indel	0,63	38 159	0	20	99,9
132	20	50 897 302	50 897 368	67	67	Indel	0,36	25 188	0	544	97,9
133	20	62 331 904	62 331 994	91	88	Poli-G (6)	0,73	32 969	0	309	99,1
134	20	62 690 860	62 690 946	87	87	Indel	0,57	32 818	0	77	99,8
135	21	30 300 823	30 300 888	66	66	Indel	0,35	24 758	9	181	99,2
136	21	33 694 176	33 694 273	98	98	Poli-T (6), CA (3)	0,54	36 902	0	160	99,6
137	21	36 710 706	36 710 792	87	87	GT(3), indel	0,39	32 841	0	48	99,9
138	21	46 644 924	46 644 992	69	69	Poli-A (6), AG (3), indel	0,32	25 939	0	280	98,9
139	21	46 705 575	46 705 664	90	90	Poli-T (5), Poli-A (6)	0,5	33 942	0	78	99,8
140	22	25 750 774	25 750 873	100	100	Indel	0,63	37 733	0	86	99,8

Amplicón	Cromosoma	Inicio del amplicón	Final del amplicón	Tamaño del fragmento analizado	Bases en regiones de confianza	Contenido genómico del amplicón	Contenido de GC	Llamadas correctas	Llamadas incorrectas	Ausencia de llamadas	% de llamadas correctas
141	22	32 439 233	32 439 329	97	97	N/A	0,68	36 617	0	49	99,9
142	22	37 409 844	37 409 940	97	97	Indel	0,46	36 525	0	162	99,6
143	22	37 637 596	37 637 694	99	99	N/A	0,6	37 398	0	24	99,9
144	22	47 081 347	47 081 438	92	92	Indel	0,66	34 754	0	22	99,9
145	X	15 870 424	15 870 492	69	69	Poli-T (5)	0,26	26 046	0	36	99,9
146	X	135 288 543	135 288 611	69	69	Poli-C (5)	0,62	26 019	0	63	99,8
147	X	135 290 777	135 290 847	71	71	N/A	0,52	26 780	0	58	99,8
148	Y	2 655 397	2 655 461	65	0	N/A	0,55	0	0	0	N/A
149	Y	2 655 519	2 655 609	91	0	N/A	0,48	0	0	0	N/A
150	Y	2 655 609	2 655 679	71	0	Poli-A (5)	0,37	0	0	0	N/A

Los resultados de la secuenciación de la muestra NA12878 se compararon con un genotipo muy fiable establecido para la NA12878 por los Institutos Nacionales de Normas y Tecnología (NIST) (v.2.19). De los 150 amplicones, 92 estaban incluidos completamente en las regiones genómicas muy fiables, 41 coincidían parcialmente y 17 no coincidían en absoluto con la secuencia de los NIST. El resultado fueron 10 000 coordenadas por duplicado para llevar a cabo la comparación. Las llamadas de bases sin variantes se compararon con el conjunto hg19 de la secuencia de referencia del genoma humano. Los resultados de exactitud se muestran en la [Tabla 18](#).

Tabla 18 Coincidencia de Somatic (somática) de la muestra GM12878 con la base de datos de los NIST

Muestra	N.º de amplicones	Media de índice de llamada	TP	FN	TN	FP	PPA	NPA	OPA
GM12878	133	98,8	2808	0	258 488	0	100	100	100

Según los datos proporcionados por este estudio de Somatic (somática) de nueve experimentos, el instrumento NextSeq 550Dx puede secuenciar de forma sistemática:

- ▶ Contenido de GC ≥ 19 % (se llamaron correctamente todas las bases llamadas de 378 de los amplicones secuenciados con el 19 % de contenido de GC y se obtuvo un índice de ausencia de llamadas del 2,6 %)
- ▶ Contenido de GC ≤ 87 % (se llamaron correctamente todas las bases llamadas de 378 de los amplicones secuenciados con el 87 % de contenido de GC y se obtuvo un índice de ausencia de llamadas del 0,6 %)
- ▶ Longitudes de poli-A ≤ 9 (se llamaron correctamente todas las bases llamadas de 378 de los amplicones secuenciados que contenían una repetición de poli-A de nueve nucleótidos y se obtuvo un índice de ausencia de llamadas del 2,5 %)
- ▶ Longitudes de poli-T ≤ 10 (se llamaron correctamente todas las bases llamadas de 378 de los amplicones secuenciados que contenían una repetición de poli-T de diez nucleótidos y se obtuvo un índice de ausencia de llamadas de menos del 0,1 %)
- ▶ Longitudes de poli-G ≤ 6 (se llamaron correctamente todas las bases llamadas de 2268 de los amplicones secuenciados que contenían una repetición de poli-G de seis nucleótidos y se obtuvo un índice de ausencia de llamadas del 0,5 %)
- ▶ Longitudes de poli-C ≤ 6 (se llamaron correctamente todas las bases llamadas de 756 de los amplicones secuenciados que contenían una repetición de poli-C de seis nucleótidos y se obtuvo un índice de ausencia de llamadas del 0,4 %)
- ▶ Longitudes de repetición de dinucleótidos $\leq 4x$ (se llamaron correctamente todas las bases llamadas de 1890 de los amplicones secuenciados que contenían una repetición de dinucleótidos de $4x$ y se obtuvo un índice de ausencia de llamadas del 0,9 %)
- ▶ Longitudes de repetición de trinucleótidos $\leq 5x$ (se llamaron correctamente todas las bases llamadas de 378 de los amplicones secuenciados que contenían repetición de trinucleótidos de $5x$ y se obtuvo un índice de ausencia de llamadas del 1,4 %)
- ▶ Longitudes de inserción ≤ 23 (se llamaron correctamente todas las bases llamadas de 378 de los amplicones secuenciados que contenían una inserción de 23 nucleótidos y se obtuvo un índice de ausencia de llamadas del 0,8 %)
- ▶ Longitudes de delección ≤ 25 (se llamaron correctamente todas las bases llamadas de 1134 de los amplicones secuenciados que contenían una delección de 25 nucleótidos y se obtuvo un índice de ausencia de llamadas del 0,7 %)

Precisión

Se determinó la precisión del instrumento NextSeq 550Dx mediante la prueba de 13 muestras únicas de Platinum Genomes con tres instrumentos, tres lotes de reactivos y tres operadores para generar nueve experimentos de secuenciación durante cinco días de inicio. El ensayo representativo, las muestras y el método de referencia son los mismos que se han descrito en el estudio de precisión de Germline (germinal). Se determinaron las contribuciones de precisión analizando los componentes de varianza con el valor de VAF como variable de respuesta y calculando las desviaciones estándares en el nivel del componente del

instrumento, el lote de reactivos, el operador y el día de inicio (Tabla 19). El número total de observaciones empleado en el análisis de la variabilidad del instrumento, el operador o el lote de reactivos fue de 699, 176 y 235 para las SNV, las inserciones y las deleciones, respectivamente.

Tabla 19 Resultados de precisión del instrumento NextSeq 550Dx (desviación estándar)

Componente	Tipo de variante	DE de componente		DE total	
		Máx.	Mediana	Máx.	Mediana
Lote	SNV	0,0076	0,0002	0,0833	0,0154
	Inserción	0,0104	0,0000	0,0410	0,0157
	Delección	0,0046	0,0005	0,0560	0,0187
Instrumento	SNV	0,0114	0,0003	0,0840	0,0153
	Inserción	0,0138	0,0009	0,0407	0,0161
	Delección	0,0079	0,0008	0,0549	0,0187
Operador	SNV	0,0226	0,0008	0,0841	0,0155
	Inserción	0,0344	0,0010	0,0417	0,0164
	Delección	0,0083	0,0013	0,0547	0,0187
Día	SNV	0,0277	0,0012	0,0825	0,0160
	Inserción	0,0235	0,0012	0,0409	0,0169
	Delección	0,0271	0,0014	0,0548	0,0188

Comparación de métodos (plataforma de secuenciación)

Las muestras de sangre completa y FFPE se evaluaron en los instrumentos NextSeq 550Dx y MiSeqDx con los flujos de trabajo de Germline (germinal) y Somatic (somática) de TruSeq Custom Amplicon Kit Dx. Se evaluó la coincidencia de frecuencia de variante en muestras de sangre y FFPE usando varios ensayos representativos. La Figura 2 representa la correlación de VAF entre dos instrumentos para un ensayo representativo y la Tabla 20 resume dicha correlación mediante un panel de ensayo. Debido a la fuerte correlación entre los instrumentos MiSeqDx y NextSeq 550Dx, se determina que las características de rendimiento relacionadas con los factores preanalíticos (por ejemplo, los métodos de extracción y las sustancias interferentes) se pueden aplicar a ambos instrumentos. Consulte las instrucciones de uso de TruSeq Custom Amplicon Kit Dx para obtener más detalles.

Figura 2 Correlación de VAF de los instrumentos MiSeqDx a NextSeq 550Dx para las muestras FFPE (izquierda) y de sangre (derecha) con el Ensayo 1

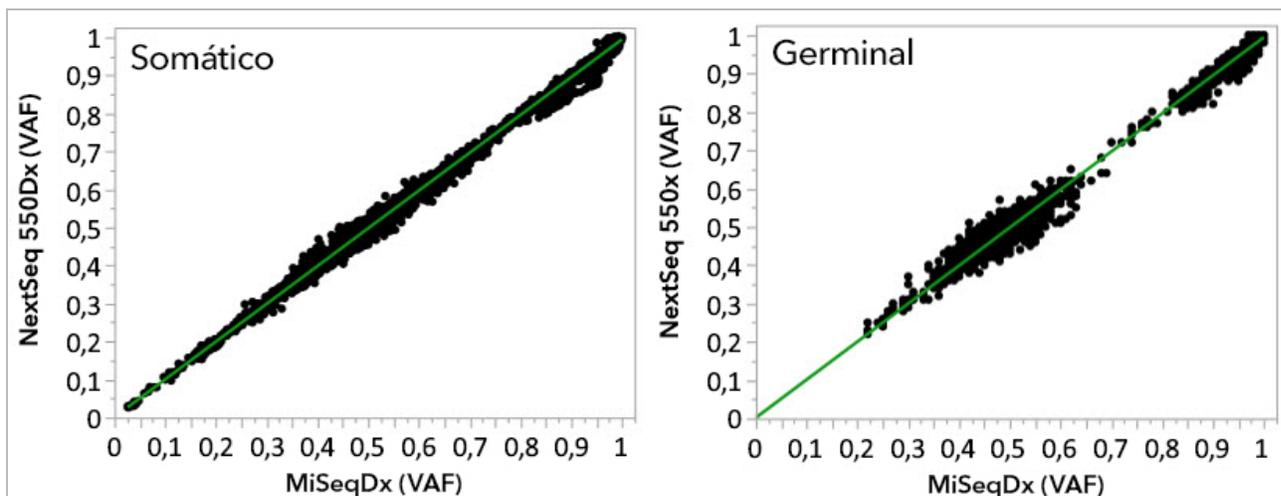


Tabla 20 Resultados de la comparación de métodos usando muestras de sangre y FFPE únicas

Fuente de ADNg	Ensayo (panel de oligonucleótidos)	Duplicados biológicos (muestras)	Duplicados técnicos (por muestra)	Observaciones (n.º de variantes)	Pendiente	Intersección	Correlación (R ²)
Sangre	Ensayo 1	45	2	8369 ¹	0,992	0,002	0,995 ²
Sangre	Ensayo 2	45	2	5457	0,995	0,005	0,981
FFPE	Ensayo 1	46	2	8319	0,993	0,000	0,997 ²
FFPE	Ensayo 3	40	1	280	0,969	0,015	0,978

¹Se eliminaron dos puntos de datos según la limitación indicada para el módulo Germline Variant (variante germinal).

²Coefficiente de determinación de los diagramas de VAF, como se muestra en la figura 2.

Reproducibilidad

La reproducibilidad del instrumento NextSeq 550Dx se evaluó usando muestras de Platinum Genomes con un ensayo representativo diseñado para buscar una variedad de genes en 12 588 bases (150 amplicones) de 23 cromosomas diferentes. Las pruebas de Germline (germinal) se componían de siete duplicados de 13 muestras; las de variantes somáticas de seis duplicados de siete muestras con diferentes niveles de VAF.

Las muestras se prepararon con TruSeq Custom Amplicon Kit Dx.

Las pruebas se llevaron a cabo en tres centros externos con un lote de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). En cada centro, se usó un único instrumento NextSeq 550Dx. Dos operadores llevaron a cabo las pruebas en cada centro. Cada operador llevó a cabo las pruebas en tres días de inicio no consecutivos para cada tipo de muestra en un total de 36 experimentos en cada uno de los tres centros.

El resultado fueron 18 experimentos para cada flujo de trabajo de Germline (germinal) y Somatic (somática).

Germline (germinal)

Las Germline Variants (variantes germinales) con un nivel de VAF $\geq 0,2$ se informan como llamadas positivas (variante). Para las Germline Variants (variantes germinales) positivas previstas, se evaluaron los datos del índice de ausencia de llamada y el índice de llamadas positivas correctas en cada tipo de variante (SNV, inserción y delección). La [Tabla 21](#) resume los índices observados, junto con los niveles de confianza superior e inferior (LCL/UCL, Lower Confidence Level/Upper Confidence Level) al 95 % calculados con el método de puntuación Wilson, por cada tipo de variante.

Tabla 21 Observaciones de llamadas de Germline (germinal) para los resultados positivos previstos por tipo de variante

Tipo de variante	Ausencia de llamadas			Llamada positiva correcta				
	Observado	Total	Porcentaje	Observado	Total	Porcentaje	LCL al 95 %	UCL al 95 %
SNV	16	110 376	0,014	110 349	110 360	99,99	99,98	99,99
Inserciones	1026	37 044	2,77	36 018	36 018	100	99,99	100,00
Delecciones	648	34 776	1,86	34 128	34 128	100	99,99	100,00

Las Germline Variants (variantes germinales) con un nivel de VAF $< 0,2$ se informan como llamadas negativas (estado natural). Para las ubicaciones de Germline (germinal) negativas previstas, se evaluaron los datos del índice de ausencia de llamada y el índice de llamadas en estado natural correctas. La [Tabla 22](#) resume los índices observados, junto con los niveles de confianza superior e inferior (LCL/UCL) al 95 % calculados con el método de puntuación Wilson.

Tabla 22 Observaciones de llamadas de Germline (germinal) para los resultados negativos previstos

Tipo de variante	Ausencia de llamadas			Llamada negativa correcta				
	Observado	Total	Porcentaje	Observado	Total	Porcentaje	LCL al 95 %	UCL al 95 %
Estado natural	4883	19 600 182	0,025	19 595 299	19 595 299	100	100,00	100,00

Las Germline Variants (variantes germinales) con un nivel de VAF $\geq 0,2$ y $< 0,7$ se llaman heterocigóticas positivas para la variante y las variantes con nivel de VAF $\geq 0,7$ se llaman homocigóticas positivas para la variante. Se usaron muestras de Germline (germinal) con variantes heterocigóticas para determinar si la variabilidad intrínseca del ensayo podría afectar a la llamada de genotipo. Se determinó el valor Cx para ambos cortes (0,2 para genotipos heterocigóticos y 0,7 para genotipos homocigóticos), donde la x equivale a la proporción de pruebas repetidas que superan el corte. En relación con el corte inferior de VAF de 0,2, el valor Cx fue $\geq 99,999$ %, lo que indica que $\geq 99,999$ % de las variantes heterocigóticas se llamarían heterocigóticas. Con respecto al corte superior de VAF de 0,7, el valor Cx fue $\leq 0,001$ %, lo que indica que $\leq 0,001$ % de las variantes heterocigóticas se llamarían homocigóticas. La [Tabla 23](#) resume los resultados por tipo de variante.

Las Germline Variants (variantes germinales) con un nivel de VAF $\geq 0,2$ y $< 0,7$ se llaman heterocigóticas positivas para la variante y las variantes con nivel de VAF $\geq 0,7$ se llaman homocigóticas positivas para la variante. Se usaron muestras de Germline (germinal) con variantes heterocigóticas para determinar si la variabilidad intrínseca del ensayo podría afectar a la llamada de genotipo. Se determinó el valor Cx para ambos cortes (0,2 para genotipos heterocigóticos y 0,7 para genotipos homocigóticos), donde la x equivale a la proporción de pruebas repetidas que superan el corte. Con respecto al corte inferior de VAF de 0,2, el valor Cx fue $\geq 99,999$ %, lo que indica que $\geq 99,999$ % de las variantes heterocigóticas se llamarían heterocigóticas. En relación con el corte superior de VAF de 0,7, el valor Cx fue $\leq 0,001$ %, lo que indica que $\leq 0,001$ % de las variantes heterocigóticas se llamarían homocigóticas. La [Tabla 23](#) resume los resultados por tipo de variante.

Tabla 23 Valores Cx de Germline (germinal) para variantes heterocigóticas

Tipo de variante	Corte con una VAF de 0,2	Corte con una VAF de 0,7
	$\geq 99,999$ %	$\leq 0,001$ %
SNV	94/94	94/94
Inserciones	24/24	24/24
Deleciones	35/35	35/35
Total	153	153

Somatic (somática)

Las Somatic Variants (variantes somáticas) con un nivel de VAF $\geq 0,026$ se informan como llamadas positivas (variante). Las observaciones con niveles de VAF $\geq 0,01$ y $< 0,026$ se consideraron como dudosas a efectos de este análisis (ni positivas ni negativas, se marcaron como frecuencia de variante baja). Para evaluar el rendimiento, se calcularon los resultados de tres formas distintas:

- ▶ Caso más favorable: cualquier resultado dudoso se consideró una llamada positiva correcta (coincidencia con los resultados previstos)
- ▶ Caso más desfavorable: cualquier resultado dudoso se consideró una llamada incorrecta (discrepancia con los resultados previstos)
- ▶ Caso de exclusión: cualquier resultado dudoso se excluyó del análisis

Las tres tablas ([Tabla 24](#), [Tabla 25](#) y [Tabla 26](#)) resumen los resultados de llamada del caso más favorable, del más desfavorable y del de exclusión, respectivamente, junto con los niveles de confianza superior e inferior (LCL/UCL) al 95 % calculados con el método de puntuación Wilson.

Tabla 24 Observaciones de llamadas de Somatic (somática) para los resultados positivos previstos por tipo de variante (caso más favorable)

Tipo de variante	Llamada positiva correcta				
	Observado	Total	Porcentaje	LCL al 95 %	UCL al 95 %
SNV	54 346	54 346	100	99,99	100,00
Inserciones	18 036	18 036	100	99,98	100,00
Deleciones	18 381	18 381	100	99,98	100,00

Tabla 25 Observaciones de llamadas de Somatic (somática) para los resultados positivos previstos por tipo de variante (caso más desfavorable)

Tipo de variante	Llamada positiva correcta				
	Observado	Total	Porcentaje	LCL al 95 %	UCL al 95 %
SNV	54 346	54 346	100	99,99	100,00
Inserciones	18 000	18 036	99,8	99,72	99,86
Deleciones	18 381	18 381	100	99,98	100,00

Tabla 26 Observaciones de llamadas de Somatic (somática) para los resultados positivos previstos por tipo de variante (llamadas dudosas eliminadas)

Tipo de variante	Llamada positiva correcta				
	Observado	Total	Porcentaje	LCL al 95 %	UCL al 95 %
SNV	54 346	54 346	100	99,99	100,00
Inserciones	18 000	18 000	100	99,98	100,00
Deleciones	18 381	18 381	100	99,98	100,00

Las Somatic Variants (variantes somáticas) con un nivel de VAF $<0,01$ se informan como llamadas negativas (estado natural). Para las ubicaciones de Somatic (somática) negativas previstas, se evaluaron los datos del índice de ausencia de llamada y el índice de llamadas en estado natural correctas. Se determinaron las llamadas en estado natural correctas mediante la exclusión del valor de ausencia de llamadas y la resta de las llamadas observadas que cayeron en la zona dudosa (niveles de VAF $\geq 0,01$ y $<0,026$), así como de las llamadas incorrectas situadas por encima del corte (niveles de VAF $\geq 0,026$) respecto del total. La [Tabla 27](#) resume los resultados observados, el total y de porcentaje de las ubicaciones de Somatic (somática) negativas para determinar el índice de ausencia de llamadas y el índice de llamadas en estado natural correctas, así como los niveles de confianza inferior y superior (LCL/UCL) al 95 % calculados con el método de puntuación Wilson.

Tabla 27 Observaciones de llamadas de Somatic (somática) para los resultados negativos previstos

Tipo de variante	Ausencia de llamadas			Llamada correcta						
	Observado	Total	Porcentaje	Dudosa	Incorrecta	Correcta	Total	Porcentaje	LCL al 95 %	UCL al 95 %
Estado natural	36 326	8 909 676	0,408	2254	121	8 870 975	8 873 350	99,97	99,972	99,974

Se evaluaron las muestras de Somatic (somática) a niveles de VAF distintos para la misma variante con el fin de determinar el valor C95 del ensayo (en cada tipo de variante). Con el fin de evaluar la variabilidad próxima al corte del ensayo, se usaron muestras con niveles de VAF esperados de entre 0,02 y 0,07. Se determinó el valor C95 para cada variante, con el C95 más alto para cada tipo de variante que se informa en la [Tabla 28](#).

Tabla 28 Resumen de C95 de Somatic (somática)

Tipo de variante	N	C95
SNV	74	0,0613
Inserción	24	0,0573
Delección	33	0,0575

Rendimiento de NextSeq 550Dx High Output Reagent Kit v2.5 (300 Cycles)

Descripción general

NextSeq 550Dx es compatible con dos kits de reactivos: NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) y NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Para demostrar que NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) puede cumplir los requisitos de rendimiento analítico verificados y validados con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), se realizaron estudios con NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Se llevaron a cabo dos preparaciones de librerías usando TruSeq Custom Amplicon Kit Dx, una con el flujo de trabajo de Germline (germinal) y la otra con el flujo de trabajo de Somatic (somática). Las librerías de cada flujo de trabajo se analizaron con tres lotes de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) usando tres instrumentos NextSeq 550Dx. Además, las pruebas de cada flujo de trabajo incluyeron un experimento único con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Sensibilidad analítica (límite del blanco [LoB, Limit Of Blank] y límite de detección [LoD, Limit Of Detection])

La verificación con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) demostró que el instrumento NextSeq 550Dx podría detectar variantes con una VAF (Variant Allele Frequency) de 0,05 con un error de tipo II $\leq 0,05$ y que el corte con una VAF de 0,026 empleado por el módulo Somatic Variant (variante somática) (LoB eficiente) admite un error de tipo I $\leq 0,01$. Sobre la base de estas afirmaciones, se espera que una variante con una VAF de 0,05 sea superior o igual con una VAF de 0,026 el 95 % del tiempo y que una posición en estado natural sea inferior con una VAF de 0,026 el 99 % del tiempo. Para asegurarse de que estas afirmaciones se cumplieran con NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), se repitieron las mediciones en el instrumento NextSeq 550Dx con muestras en estado natural (muestras de LoB) y con muestras que contenían variantes con una VAF de 0,05 (muestras de LoD) usando NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). A continuación, se comparó la proporción de llamadas por encima y por debajo del corte de 0,026 con las afirmaciones establecidas con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Las pruebas incluyeron dos muestras de LoD, cada una con un conjunto único de variantes selectivas con una VAF de 0,05, y las correspondientes muestras de LoB que estaban en estado natural para las variantes selectivas. Para la preparación de librerías, las muestras de LoD y LoB se procesaron en duplicados de ocho y siete, respectivamente, usando TruSeq Custom Amplicon Kit Dx. Las librerías se secuenciaron inicialmente usando NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) para identificar variantes/coordenadas genómicas para la evaluación de LoB/LoD con NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Todas las variantes con una VAF promedio entre 0,045 - 0,055 (variantes de LoD) basadas en los resultados de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) se usaron para el análisis de LoD (N = 51 variantes). En el análisis de LoB, se evaluaron las correspondientes 51 coordenadas genómicas.

En la evaluación de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), las librerías se secuenciaron en tres experimentos durante tres días consecutivos usando el mismo instrumento y lote de kit de reactivos. Estas pruebas supusieron 24 duplicados para cada una de las 51 variantes de LoD y a 21 duplicados para cada una de las posiciones en estado natural correspondientes. La proporción de llamadas en estado natural con una VAF $< 0,026$ se proporciona en la [Tabla 29](#). La proporción de llamadas de variantes de LoD con una VAF superior o igual a 0,026 se proporciona en la [Tabla 30](#).

Tabla 29 Proporción de llamadas con $<0,026$ para las posiciones en estado natural (evaluación de afirmaciones de LoB)

Tipo de variante	Posiciones evaluadas	Total de observaciones	N.º de mediciones de VAF $\geq 2,6$ %	Proporción $< 2,6$ %	Proporción al 95 % Intervalo de confianza
SNV	32	672	0	1	0,994 – 1
Inserción	11	231	0	1	0,984 – 1
Delección	8	168	0	1	0,978 – 1

Tabla 30 Proporción de llamadas de variantes de con una VAF de $\geq 0,026$ para las variantes de LoD (evaluación de afirmaciones de LoD)

Tipo de variante	Posiciones evaluadas	Total de observaciones	N.º de mediciones de VAF $< 2,6$ %	N.º de mediciones de VAF $\geq 2,6$ %	Proporción $\geq 2,6$ %	Proporción al 95 % Intervalo de confianza
SNV	32	768	1	767	0,999	0,993 – 1
Inserción	11	264	3	261	0,989	0,967 – 0,996
Delección	8	192	2	190	0,99	0,963 – 0,997

Exactitud

Germline (germinal)

El siguiente estudio se llevó a cabo para evaluar la exactitud de la llamada de variantes del módulo Germline Variant (variante germinal) usando NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Se analizaron doce muestras únicas de Platinum Genomes con un ensayo representativo. Se llevó a cabo un total de 11 experimentos con tres instrumentos NextSeq 550Dx y tres de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Se determinó la exactitud de las SNV, las inserciones y las deleciones mediante la comparación de los resultados con un método de referencia compuesto y con características bien definidas: la versión 1.0 de Platinum Genomes de 2016. Se proporcionan como referencia los resultados de exactitud de un único experimento de secuenciación con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). En la [Tabla 31](#) se proporciona un resumen de los resultados.

Tabla 31 Resumen de la coincidencia de Germline (germinal)

Criterios	Total de observaciones (v2.5) ¹	Resultado por observación (v2.5) ²	Resultado por observación (v2) ³	Resultado por experimento (v2.5) ⁴	Resultado por experimento (v2) ⁴
PPA de SNV	1056	98,7	98,7	>99,9	>99,9
PPA de inserciones	1056	100	100	100	98,9
PPA de deleciones	1056	95,2	95,2	>99,9	100
NPA	1056	100	100	100	100
OPA	1056	>99,9	>99,9	>99,9	>99,9

¹Calculado de la siguiente manera: número de muestras por experimento x número de experimentos (96 muestras por experimento x 11 experimentos = 1056 observaciones).

²Valor observado más bajo por duplicado de muestra en todos los experimentos (en función de 11 experimentos de NextSeq 550Dx High Output Reagent Kit v2.5).

³Valor observado más bajo por duplicado de muestra en un experimento (96 observaciones en total).

⁴Valor más bajo cuando los datos de cada experimento se analizan en conjunto.

Somatic (somática)

El siguiente estudio se llevó a cabo para evaluar la exactitud de la llamada de variantes del módulo Somatic Variant (variante somática) en el instrumento NextSeq 550Dx usando NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Se analizaron diez muestras FFPE de Platinum Genomes (dos con variantes diluidas hasta con una VAF de 0,05) usando un ensayo representativo. Se llevó a cabo un total de 11 experimentos con tres instrumentos NextSeq 550Dx y tres lotes de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Se determinó la exactitud de las SNV, las inserciones y las deleciones mediante la comparación de los resultados con un método de referencia compuesto y con características bien definidas: la versión 1.0 de Platinum Genomes de 2016. Se proporcionan como referencia los resultados de exactitud de un único experimento de secuenciación con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). En la [Tabla 32](#) se proporciona un resumen de los resultados.

Tabla 32 Resumen de la coincidencia de Somatic (somática)

Crterios	Total de observaciones (v2.5) ¹	Resultado por observación (v2.5) ²	Resultado por observación (v2) ³	Resultado por experimento (v2.5) ⁴	Resultado por experimento (v2) ⁴
PPA de SNV	528	100	100	100	100
PPA de inserciones	528	96,9	96,9	>99,9	>99,9
PPA de deleciones	528	100	100	100	100
NPA	528	>99,9	>99,9	>99,9	>99,9
OPA	528	>99,9	>99,9	>99,9	>99,9

¹Calculado de la siguiente manera: número de muestras por experimento × número de experimentos (48 muestras por experimento × 11 experimentos = 528 observaciones).

²Valor observado más bajo por duplicado de muestra en todos los experimentos (en función de 11 experimentos de NextSeq 550Dx High Output Reagent Kit v2.5).

³Valor observado más bajo por duplicado de muestra en un experimento (96 observaciones en total).

⁴Valor más bajo cuando los datos de cada experimento se analizan en conjunto.

Precisión

Germline (germinal)

La precisión de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) con el módulo Germline Variant (variante germinal) se evaluó usando muestras de Platinum Genomes y un ensayo representativo. Las pruebas consistieron en la preparación de una única librería usando TruSeq Custom Amplicon Kit Dx e incluyeron 12 muestras procesadas con ocho duplicados cada una. Las librerías se secuenciaron con tres lotes de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) y tres instrumentos NextSeq 550Dx para un total de nueve experimentos de secuenciación.

Se usaron muestras con variantes heterocigóticas para determinar si la variabilidad intrínseca del ensayo podría afectar a la llamada de genotipo (N = 153 variantes heterocigóticas únicas). Se determinó el valor Cx para ambos cortes (0,2 para genotipos heterocigóticos y 0,7 para genotipos homocigóticos) del módulo Germline Variant (variante germinal), donde la x equivale a la proporción de pruebas repetidas que superan el corte. Para el corte inferior de VAF de 0,2, la variante con el valor Cx mínimo para NextSeq 550Dx Reagent Kit v2.5 (300 cycles) fue >99,9 %, lo que indica que >99,9 % de las variantes heterocigóticas se denominarían heterocigóticas. Para el corte superior de VAF de 0,7, la variante con el valor Cx máximo para NextSeq 550Dx Reagent Kit v2.5 (300 cycles) fue <1,5 %, lo que indica que ≤1,5 % de las variantes heterocigóticas se denominarían homocigóticas. La [Tabla 33](#) resume los resultados por tipo de variante. Se proporcionan para su uso como referencia los valores Cx del único experimento de secuenciación usando NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Tabla 33 Valores Cx de Germline (germinal) para variantes heterocigóticas

Tipo de variante	N	Corte con una VAF de 0,2		Corte con una VAF de 0,7	
		Valor Cx mínimo (v2.5) ¹	Valor Cx mínimo (v2) ²	Valor Cx máximo (v2.5) ¹	Valor Cx máximo (v2) ²
SNV	94	>99,9 %	>99,9 %	1,5 %	1,0 %
Inserciones	24	100 %	100 %	0 %	<0,1 %
Deleciones	35	100 %	>99,9 %	<0,1 %	<0,1 %

¹Valores Cx basados en las estimaciones de la desviación estándar total a partir del análisis de los componentes de varianza.

²Valores Cx basados en las desviaciones estándares de las muestras.

Somatic (somática)

La precisión de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) con el módulo Somatic Variant (variante somática) se evaluó usando muestras FFPE de Platinum Genomes y un ensayo representativo. Las pruebas consistieron en la preparación de una única librería usando TruSeq Custom Amplicon Kit Dx e incluyeron dos muestras con ocho duplicados cada una. Las librerías se secuenciaron usando tres lotes de NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) y tres instrumentos NextSeq 550Dx para un total de nueve experimentos de secuenciación.

Las variantes somáticas con niveles de VAF previstos $\leq 0,10$ de VAF (N = 131 variantes únicas) se usaron para evaluar la variabilidad del instrumento cerca del corte de VAF del módulo Somatic Variant (variante somática) (las variantes somáticas con un nivel de VAF $\geq 0,026$ se denominan positivas para la variante). Se determinaron los valores C95 para cada variante somática. Los valores C95 representan la VAF en la que la probabilidad de ser mayor que el corte de VAF del módulo Somatic Variant (variante somática) es del 95 %. Los valores C95 más altos por tipo de variante se informan en la [Tabla 34](#). Se proporcionan para su uso como referencia los resultados de C95 de un único experimento de secuenciación usando NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Tabla 34 Resumen de C95 de Somatic (somática)

Tipo de variante	N.º de variantes evaluadas	C95 (v2.5) ¹	C95 (v2) ²
SNV	74	0,064	0,063
Inserciones	24	0,062	0,061
Deleciones	33	0,060	0,060

¹Valores C95 basados en las estimaciones de la desviación estándar total a partir del análisis de los componentes de varianza.

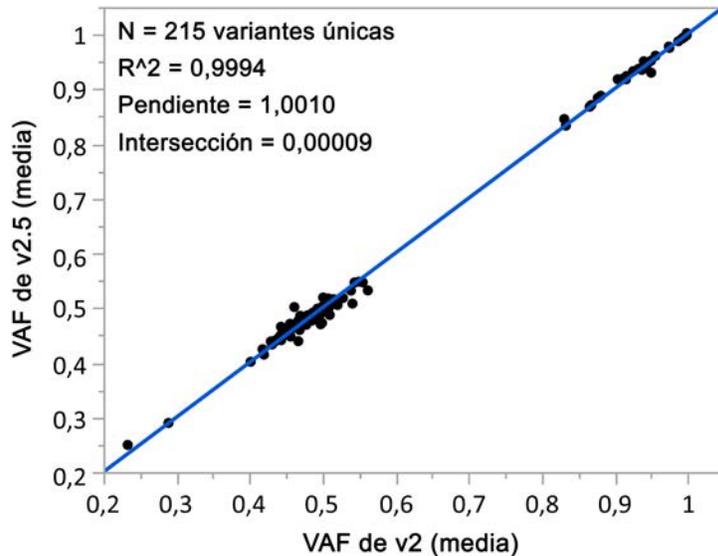
²Valores C95 basados en las desviaciones estándares de las muestras.

Comparación de métodos (kit de reactivos)

Germline (germinal)

Se evaluaron las VAF promedio de 215 variantes únicas en NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) y en NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) usando los resultados generados por el módulo Germline Variant (variante germinal). Los promedios de VAF se calcularon a partir de 11 experimentos de secuenciación (v2.5) y un experimento de secuenciación (v2). Se usaron al menos ocho duplicados para calcular el promedio para cada variante. La [Figura 3](#) representa la correlación de VAF entre los dos kits de reactivos. En función de la fuerte correlación de VAF lineal y de la similitud de los resultados entre los kits de reactivos, se determina que las características de rendimiento verificadas y validadas inicialmente con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) con el módulo Germline Variant (variante germinal) son aplicables a NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

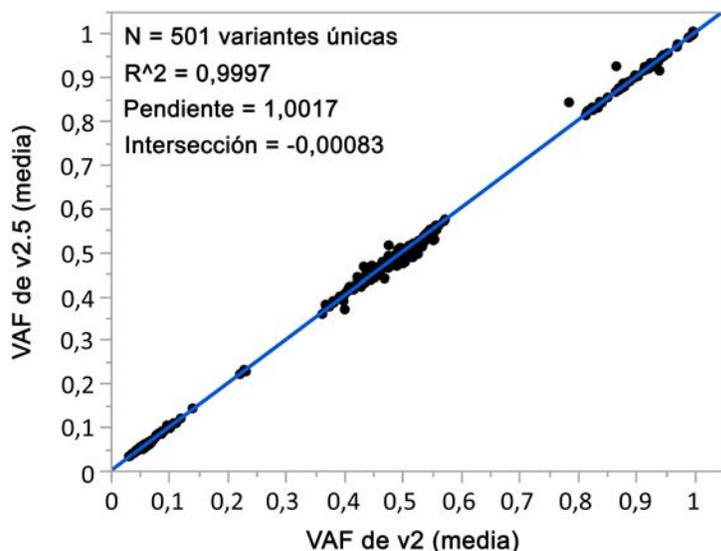
Figura 3 Correlación de frecuencia alélica de variantes (VAF) del módulo Germline Variant (variante germinal) entre NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) y NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).



Somatic (somática)

Se evaluaron las VAF promedio para 501 variantes únicas en NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) y NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) usando los resultados generados por el módulo Somatic Variant (variante somática). Los promedios de VAF se calcularon a partir de 11 experimentos de secuenciación (v2.5) y un experimento de secuenciación (v2). Se usaron al menos tres duplicados para calcular el promedio para cada variante única. La [Figura 4](#) representa la correlación de VAF entre los dos kits de reactivos. En función de la fuerte correlación de VAF y de la similitud de los resultados entre los kits de reactivos, se determina que las características de rendimiento verificadas y validadas con NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) con el módulo Somatic Variant (variante somática) son aplicables a NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Figura 4 Correlación de frecuencia alélica de variantes (VAF) del módulo Somatic Variant (variante somática) entre NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) y NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).



Historial de revisiones

Documento	Fecha	Descripción del cambio
N.º de documento 1000000030326 v06	Mayo de 2022	Se han realizado actualizaciones para corregir contenido añadido por error del software de origen.
N.º de documento 1000000030326 v05	Noviembre de 2021	Se ha añadido una declaración de Advertencias y Precauciones sobre la notificación de incidentes graves. Se ha añadido una declaración a los Principios de Procedimiento en la que se especifica el usuario previsto. Se ha eliminado la referencia a High Output Reagent Kit v2 (300 cycles). Se ha añadido la referencia a High Output Reagent Kit v2.5 (75 cycles).
N.º de documento 1000000030326 v04	Agosto de 2021	Se ha añadido la tabla Historial de revisiones. Se ha actualizado la dirección del representante autorizado en la UE.

Patentes y marcas comerciales

Este documento y su contenido son propiedad exclusiva de Illumina, Inc. y sus afiliados (“Illumina”) y están previstos solamente para el uso contractual de sus clientes en conexión con el uso de los productos descritos en él y no para ningún otro fin. Este documento y su contenido no se usarán ni distribuirán con ningún otro fin ni tampoco se comunicarán, divulgarán ni reproducirán en ninguna otra forma sin el consentimiento previo por escrito de Illumina. Illumina no transfiere mediante este documento ninguna licencia bajo sus derechos de patente, marca comercial, copyright ni derechos de autor o similares derechos de terceros.

Para asegurar el uso correcto y seguro de los productos descritos en este documento, el personal cualificado y adecuadamente capacitado debe seguir las instrucciones incluidas en este de manera rigurosa y expresa. Se debe leer y entender completamente todo el contenido de este documento antes de usar estos productos.

SI NO SE LEE COMPLETAMENTE EL DOCUMENTO Y NO SE SIGUEN EXPRESAMENTE TODAS LAS INSTRUCCIONES DESCRITAS EN ESTE, PODRÍAN PRODUCIRSE DAÑOS EN EL PRODUCTO, LESIONES PERSONALES, INCLUIDOS LOS USUARIOS U OTRAS PERSONAS Y DAÑOS EN OTROS BIENES Y QUEDARÁ ANULADA TODA GARANTÍA APLICABLE AL PRODUCTO.

ILLUMINA NO ASUME RESPONSABILIDAD ALGUNA DERIVADA DEL USO INCORRECTO DE LOS PRODUCTOS AQUÍ DESCRITOS (INCLUIDAS LAS PIEZAS O EL SOFTWARE).

© 2022 Illumina, Inc. Todos los derechos reservados.

Todas las marcas comerciales pertenecen a Illumina, Inc. o a sus respectivos propietarios. Para obtener información específica sobre las marcas comerciales, consulte www.illumina.com/company/legal.html.

Información de contacto



Illumina
5200 Illumina Way
San Diego, California 92122 (EE. UU.)
+ 1 800 809 ILMN (4566)
+ 1 858 202 4566 (fuera de Norteamérica)
techsupport@illumina.com
www.illumina.com



Illumina Netherlands B.V.
Steenoven 19
5626 DK Eindhoven
Países Bajos

Patrocinador australiano

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

Etiquetado de productos

Consulte la clave de símbolos de su kit en support.illumina.com para acceder a una referencia completa de símbolos que podrían aparecer en etiquetas y envases de productos.