

# 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 bibliotecas de ADN cuando se utiliza con ensayos de diagnóstico *in vitro*. Debe usarse con reactivos de diagnóstico *in vitro* registrados, certificados o aprobados y con software analítico específicos.

## Principios de procedimiento

El instrumento Illumina NextSeq 550Dx está concebido para la secuenciación de bibliotecas de ADN con ensayos de diagnóstico *in vitro*. Para la entrada, el NextSeq 550Dx utiliza bibliotecas generadas a partir de ADN en las que los índices de muestra y las secuencias de captura se añaden a los objetivos amplificados. Las bibliotecas 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). El proceso químico de SBS utiliza un método basado en terminadores reversibles para detectar bases de nucleótidos únicos marcados con tinta 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 llamadas 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. El NextSeq 550Dx utiliza 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, las llamadas 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). 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) 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), al utilizarse con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) o el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos), pueden 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  $\geq 75$  % a una longitud de lectura de  $2 \times 150$  pb  
Un 75 % o más de las bases tiene una puntuación de calidad según la escala de Phred  $\geq 30$ , lo que implica una precisión de la llamada de bases superior al 99,9 %
- 3 El software del ensayo no alineará las lecturas con inserciones, deleciones o modificaciones cuya longitud de contenido sea superior a los 25 pb. Por lo tanto, el software del ensayo no detectará las inserciones y deleciones con una longitud superior a los 25 pares de bases.
- 4 Es posible que el software del ensayo no alinee las lecturas de amplicones que contienen una cantidad excesiva de variantes y, en consecuencia, se catalogará la región como de estado natural. El contenido excesivo incluye:
  - ▶ Lecturas que contienen más de tres inserciones y deleciones
  - ▶ Lecturas cuya longitud es de, al menos, 30 pares de bases y cuyo contenido de una 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 los 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) y las inserciones y deleciones 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 inserciones y deleciones si son directamente adyacentes a un cebador y no existe solapamiento de amplicones. Con respecto a las regiones con solapamiento de amplicones, el ensayo no puede detectar deleciones cuando la región de solapamiento es más pequeña que 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. Se puede detectar una deleción de una sola base en cualquiera de esas dos bases.
- 8 Como con cualquier flujo de trabajo de preparación de bibliotecas 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 están sondando 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 errónea.
  - ▶ Las variantes de deleciones e inserciones 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 a un cebador.
- 9 Las inserciones y deleciones 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 variantes 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 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:
  - ▶ El instrumento NextSeq 550Dx, que utiliza el módulo Germline Variant (Variante germinal) de la aplicación Local Run Manager para NextSeq 550Dx, se ha concebido para ofrecer resultados cualitativos en lo que se refiere a las llamadas de variantes germinales (por ejemplo, homocigóticas, heterocigóticas y en estado natural).

- ▶ Para utilizar el módulo Germline Variant (Variante germinal) es necesaria una cobertura mínima de 150× por amplicón para que la llamada de variantes sea precisa. 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 que se van a estudiar influirán en la cobertura. El contenido de GC, así como otros contenidos genómicos, también influyen en la cobertura.
- ▶ La variación en el número de copias puede influir 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 utiliza 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 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 (R5×9).
- ▶ Una inserción y deleción y una variante de nucleótido único en un solo locus pueden tener como resultado que solo se informe de una sola variable.

### 13 Limitaciones específicas de Somatic.

- ▶ El instrumento NextSeq 550Dx, que utiliza el módulo Somatic Variant (Variante somática) de Local Run Manager para NextSeq 550Dx, se ha concebido para ofrecer resultados cualitativos en lo que se refiere a las llamadas de variantes somáticas (por ejemplo, la presencia de una variante somática con una frecuencia de variante mayor o igual a 0,026 con un límite de detección de 0,05).
- ▶ Para utilizar el módulo Somatic Variant (Variante somática) es necesaria una cobertura mínima por amplicón de 450× por grupo de oligonucleótidos para que la llamada de variantes sea precisa. 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 que se van a estudiar influirán en la cobertura. El contenido GC, así como otros contenidos genómicos, también influyen en la cobertura.
- ▶ En el caso de las llamadas de 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 tres pares de bases (R3×6).
- ▶ El módulo Somatic Variant (Variante somática) no es capaz de diferenciar entre las variantes germinales y las variantes somáticas. Aunque el módulo se ha diseñado para detectar las variantes dentro de un espectro de frecuencias de variantes, no es posible utilizar las frecuencias para diferenciar las variantes germinales de las somáticas.
- ▶ El tejido común que se encuentre en la muestra puede influir en la detección de las variantes. El límite de detección 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 software	Función	Descripción
Software operativo de NextSeq 550Dx (NOS)	Controla el funcionamiento del instrumento	La aplicación de software NOS controla el funcionamiento del instrumento durante la secuenciación y genera imágenes para que las utilice el software de análisis en tiempo real (RTA).
Software de análisis en tiempo real (RTA)	Realiza el análisis principal	La aplicación de software RTA convierte las imágenes que genera NOS de cada placa por ciclo de experimento de secuenciación en archivos de llamada de bases, que son entradas de los módulos de análisis de Local Run Manager. La aplicación de software RTA no tiene interfaz de usuario.
Local Run Manager	Interfaz para la selección del módulo	El software 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 software 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, las llamadas 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 software 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, las llamadas 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 la 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 $\pm$ 3 °C). Se trata de la temperatura de funcionamiento del instrumento. Durante un experimento, no deje que la temperatura ambiente varíe más de $\pm$ 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

Kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos), n.º de catálogo 20019554

Kit de reactivos de rendimiento elevado NextSeq 550Dxv2.5 (300 ciclos), n.º de catálogo 20028871

### Materiales y equipo necesarios no suministrados

#### Consumibles proporcionados 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 usos múltiples

#### Consumibles proporcionados por el usuario para el mantenimiento del instrumento

Consumible	Proveedor	Finalidad
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	Illumina, n.º de catálogo 20022240	Limpieza del aire que el instrumento recibe para refrigerarse

### Directrices para el agua de laboratorio

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

- ▶ Agua desionizada
- ▶ Illumina PW1
- ▶ Agua de 18 Megaohmios (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 utilizar u ordenar la utilización de este dispositivo.

- 1 **Algunos componentes de los reactivos proporcionados por Illumina que se utilizan 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.**

**Utilice un equipo de protección, incluidos gafas, guantes y batas de laboratorio adecuados para el riesgo de exposición. Manipule los reactivos utilizados 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](http://support.illumina.com/sds.html).

- 2 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).
- 3 El incumplimiento de los procedimientos descritos puede provocar resultados erróneos o una reducción considerable de la calidad de las muestras.
- 4 Utilice 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.
- 5 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 por PCR puede causar resultados poco precisos y fiables.
- 6 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 calientes, mezcladores vorticiales y centrifugadoras).
- 7 El emparejamiento de índice a muestra debe coincidir exactamente con la disposición de placas impresa. Local Run Manager se encarga de rellenar de forma automática los cebadores de índices que se han asociado con los nombres de las muestras cuando estos se introducen en el módulo. Es 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.
- 8 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.
- 9 No utilice el NextSeq 550Dx si se ha retirado alguno de los paneles. Poner en funcionamiento el instrumento estando alguno de estos paneles retirado constituye una posible exposición a la tensión de línea, así como a tensiones de CC.
- 10 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.
- 11 El instrumento pesa aproximadamente 83 kg (185 libras) y puede provocar lesiones graves si se cae o se maneja de forma indebida.

## Instrucciones de uso

Las siguientes instrucciones de uso describen cómo utilizar los módulos Germline Variant (Variante germinal) y Somatic Variant (Variante somática) en el modo de diagnóstico en el instrumento NextSeq 550Dx con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) o el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos).

## Introducir la información sobre el 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, a continuación, **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 Haga clic en **Import Manifest** (Importar manifiesto).
- 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)*.

#### Definir las muestras para el experimento


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

- ▶ **Introducción manual de las muestras:** utilice la tabla en blanco que aparece en la pantalla Create Run (Crear experimento).
- ▶ **Importar las 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.


#### Introducir 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 del campo 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 bibliotecas:
  - 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).

#### Importar las muestras

- 1 Seleccione **Import Samples** (Importar muestras) y busque la ubicación del archivo que contiene la información sobre 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 Seleccione 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 bibliotecas.
- 4 Seleccione **Save Run** (Guardar Experimento).

## Preparación del cartucho de reactivo

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

- 1 Extraiga el cartucho de reactivo 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 Golpetee ligeramente en el banco 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 Extraiga 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.

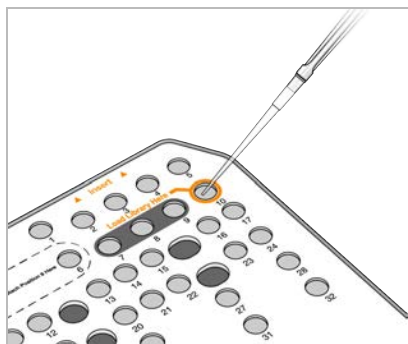
## Preparación de bibliotecas para la secuenciación

Desnaturalice y diluya las bibliotecas 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 bibliotecas. La dilución de las bibliotecas de muestras depende de la complejidad de los grupos de oligonucleótidos. Para obtener información sobre la preparación de las bibliotecas de muestras para la secuenciación, incluidas la agrupación y la dilución de bibliotecas, consulte la sección Instrucciones de uso del kit de preparación de bibliotecas aplicable. El NextSeq 550Dx precisa la optimización de la densidad de grupos.

## Carga de bibliotecas en el cartucho de reactivo

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

Figura 1 Carga de bibliotecas



## 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 Elija uno de los experimentos del listado y, a continuación, seleccione **Next** (Siguiendo).  
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 reactivo) 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.
  - ▶ Extraiga la celda de flujo del embalaje de aluminio.
  - ▶ Abra el estuche de plástico transparente y extraiga 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 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 reactivo NextSeq 550Dx.  
Tras cargar los cartuchos de tampones y de reactivo del NextSeq 550Dx, el software lee y registra la RFID. Aparecen en la pantalla los ID de los cartuchos de tampones y de reactivo 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 analiza imágenes y realiza llamadas de bases, además de asignar una puntuación de calidad a cada base para 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 la secuencia de índice que se ha 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 al experimento.
- ▶ 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

Tras el 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 incluyen 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 y por grupo de oligonucleótidos (es decir, cuatro de la lectura 1 y cuatro de la lectura 2). Esto se traduce en un total de ocho archivos FASTQ por muestra para el módulo de variante germinal y de 16 archivos para el módulo de variante somática. Los archivos FASTQ son la entrada principal para la alineación.

## Alineación

A lo largo del proceso de alineación, el algoritmo Smith-Waterman alinea las agrupaciones de cada muestra con las secuencias del amplicón que determina el archivo de manifiesto.

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

En el caso de la 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 de acuerdo con el complemento inverso de los oligonucleótidos descendentes de locus específico (DLSO).
- ▶ La Lectura 2 se evalúa teniendo en cuenta oligonucleótidos ascendentes de locus específico (ULSO).
- ▶ 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 inserciones y deleciones), la lectura completa se alineará con el amplicón objetivo de esa secuencia.
- ▶ Las inserciones y deleciones dentro del DLSO y del 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 utilizan para la llamada de variantes.

## Llamadas de variantes

El llamador de variantes PISCES se ha diseñado para hacer llamadas de variantes de SNV e inserciones y deleciones a partir de las bibliotecas 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 cálculo de variantes. Los módulos generan también archivos de resultados, como los VCF y los gVCF (formato de llamada de variantes de genoma), 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 bibliotecas, se recomiendan controles positivos y negativos y su evaluación. Evalúe los controles como se indica a continuación:

- **Control 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, el 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 utilizando los módulos Germline Variant (Variante germinal) y Somatic Variant (Variante somática) con el kit de amplicones personalizados TruSeq Dx y el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) y se confirmaron usando el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos). 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 utilizan el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos) se diseñaron para evaluar las afirmaciones de rendimiento previamente establecidas con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos). Los resultados demuestran que los kits de reactivos (v2 y v2.5) tienen un rendimiento comparable con el kit de amplicones personalizados TruSeq Dx. Consulte el *prospecto del kit de amplicones personalizados TruSeq Dx* para conocer las características de rendimiento relacionadas con los factores preanalíticos, por ejemplo, los métodos de extracción o las sustancias interferentes.

## Definiciones de los cálculos que se utilizan en las características de rendimiento

- 1 La coincidencia de porcentaje positivo (PPA) 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.
  - ▶  $(N.^{\circ} \text{ de locus en los que se han detectado variantes de forma correcta en el ensayo}) / (\text{cantidad total de locus en los que existen variantes})$   
 Los locus en los que se han detectado variantes en el ensayo y que concuerdan con el método de referencia son positivos verdaderos (TP). Los locus en los que existía una variante, pero que en el ensayo se han detectado como llamada de referencia o como distintas llamadas de variantes son falsos negativos (FN).
- 2 La coincidencia de porcentaje negativo (NPA) es la proporción de locus que se clasifican como sin mutaciones por medio de un método de referencia y que se detectan de forma correcta en el ensayo.
  - ▶  $(N.^{\circ} \text{ de locus en estado natural detectados de forma correcta en el ensayo}) / (\text{cantidad total de locus en estado natural})$   
 Los locus en estado natural que se han detectado en el ensayo y que concuerdan con el método de referencia son negativos verdaderos (TN). Los locus en estado natural que en el ensayo se detectan como variantes son falsos positivos (FP).
- 3 La coincidencia total de porcentajes (OPA) es la proporción de locus que se detectan correctamente en el ensayo con respecto al método de referencia.
  - ▶  $([N.^{\circ} \text{ de locus en los que se han detectado variantes de forma correcta en el ensayo}] + [n.^{\circ} \text{ de locus en estado natural detectados de forma correcta en el ensayo}]) / ([n.^{\circ} \text{ total de locus en los que existen variantes}] + [n.^{\circ} \text{ total de locus en estado natural}])$
- 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 superan 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 superado 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. Esta métrica no tiene en cuenta la coincidencia de las llamadas con el método de referencia.

## Rendimiento del kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos)

### Indexación de muestras

Los cebadores de índice de muestras, añadidos durante la preparación de bibliotecas, 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 utiliza en los flujos de trabajo tanto germinal como somático. El objetivo de este estudio fue establecer el número mínimo (8) y máximo (96) de muestras que se pueden procesar en un único experimento de secuenciación en el instrumento NextSeq 550Dx. Se probaron 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 bibliotecas de muestras indexadas de forma única utilizando un ensayo típico 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 diferentes de cebadores de índice. Para el segundo conjunto de experimentos, se secuenciaron ocho bibliotecas 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 % en cada una de las combinaciones de ocho índices.

## Arrastre de muestras

El instrumento NextSeq 550Dx permite la secuenciación de varias muestras, así como el control 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 típico 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 bibliotecas 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, las llamadas 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 secuencian una sola vez un grupo de bibliotecas 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 es de 0,084 %.

Para analizar el arrastre de muestras entre experimentos, se preparan dos grupos de bibliotecas y se secuencian 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 emplean el mismo conjunto de adaptadores de índices. El grupo de muestras femeninas se secuencian primero seguido de un experimento de secuenciación posterior con el grupo de muestras masculinas. Por último, se repite 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 repetido del grupo de muestras femeninas y el experimento del grupo de muestras masculinas. La mediana observada de arrastre entre experimentos es de 0,0076 %.

## Entrada de ADN

### Sangre (Germline)

Se ha establecido el rango de cantidad de entrada de ADN de sangre para la preparación de bibliotecas del kit de amplicones personalizados TruSeq 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 típico diseñado para buscar una variedad de genes en 12 588 bases de 23 cromosomas diferentes. La biblioteca se secuenció en dos instrumentos NextSeq 550Dx con un lote del kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos).

Cinco muestras se evaluaron por duplicado en cinco niveles de cantidad de entrada de ADN que oscilaban entre 250 ng y 12 ng (250 ng, 100 ng, 50 ng, 25 ng y 12 ng). Ocho muestras se evaluaron como un único duplicado en cada uno de los cinco niveles de cantidad 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 cantidad de entrada. La PPA en 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 cantidad de entrada se observó una exactitud parecida. La cantidad de entrada de ADN recomendada para el kit de amplicones personalizados TruSeq 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 cada 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)

Se ha establecido el rango de cantidad de entrada de ADN fijada en formol y embebida en parafina (FFPE) para la preparación de bibliotecas con el kit de amplicones personalizados TruSeq Dx mediante el flujo de trabajo del módulo Somatic Variant (Variante somática) en el instrumento NextSeq 550Dx. Dicho rango de cantidad de entrada de ADN se evaluó llevando a cabo un estudio de dilución en serie con tres muestras de Platinum Genomes con un ensayo típico diseñado para buscar varios genes en 12 588 bases de 23 cromosomas diferentes. Las estirpes celulares de 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 variantes de alelos (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 variantes de hasta 1,0 VAF. Las muestras se procesaron por duplicado en cinco niveles de cantidad 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 el kit de amplicones personalizados TruSeq Dx: control de calidad FFPE. Cada biblioteca se secuenció en dos instrumentos NextSeq 550Dx con dos lotes de kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos). Para determinar la exactitud, se compararon llamadas de variantes con la versión 1.0 de Platinum Genomes de 2016. La PPA en cada tipo de variante (SNV, inserciones y delecciones) se representa en la [Tabla 3](#) y la NPA se representa en la [Tabla 4](#). La cantidad de entrada de ADN recomendada para las variantes con 0,05 VAF o un valor superior es de dCq ≤ 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 cada tipo de variante

dCQ medio	Tipo de variante	Variantes previstas	Ausencia de llamadas previstas	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

dCQ medio	Estado natural esperado	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	93688	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 típico diseñado para buscar una variedad de genes en 12 588 bases de 23 cromosomas diferentes. Las estirpes celulares de 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 variantes de 74 variantes (53 SNV, siete inserciones y 14 delecciones) fueran de  $0,05 \pm 0,02$ . GM12877 y la estirpe celular diluida GM12878 (GM12878-D) se analizaron durante seis días de inicio consecutivos con un solo instrumento, alternando entre dos lotes del kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos), para un total de seis experimentos de secuenciación. Los resultados obtenidos fueron de 60 duplicados por cada variante en la muestra GM12878-D y 72 duplicados de cada coordenada en estado natural correspondiente en la muestra 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 variantes por cada tipo de variante determinado. El error del Tipo I se definió como 0,01 y el error del 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 establecer 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 Reactivo 1 (%)	LoB del Reactivo 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 por cada lote de reactivos de 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 de 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 las SNV, 5,12 % en las inserciones y 5,26 % en las deleciones.

Tabla 6 Límite de detección

Lote de reactivo	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

### 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 con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos). Se analizaron 13 muestras únicas de Platinum Genomes con un ensayo típico diseñado para buscar varios genes en 12 588 bases (150 amplicones) de 23 cromosomas diferentes. Se llevaron 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 ha determinado 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 han definido regiones genómicas fiables según este método de referencia, a menos que se haya indicado lo contrario.

Tabla 7 Resumen de la coincidencia germinal

Criterios	Total de observaciones <sup>1</sup>	Resultado por observación <sup>2</sup>	Resultado por experimento <sup>3</sup>
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

<sup>1</sup> Se calcula de la siguiente manera: número de muestras por experimento (91) × número de experimentos (9) = 819.

<sup>2</sup> Valor observado más bajo por duplicado de muestra en los nueve experimentos.

<sup>3</sup> Valor 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 las 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 y las deleciones, los resultados de la base sin variantes se comparan con el conjunto hg19 de la secuencia de referencia del genoma humano para los cálculos de la coincidencia de porcentaje negativo.

Tabla 8 Coincidencia de variantes germinales por muestra

Muestra	Promedio del índice de llamada	Variantes previstas <sup>1</sup>	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

<sup>1</sup> Número 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 las 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 variantes germinales 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

>Muestra	SNV			Inserciones			Deleciones		
	>Previsto	>TP	>FN	>Previsto	>TP	>FN	Previsto	TP	FN
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 (inserciones y deleciones). En la [Tabla 10](#) se incluye un resumen general. Se detectó un total de 71 inserciones y deleciones 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 inserciones y deleciones germinales

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 diverso contenido genómico. El contenido de GC de los amplicones oscilaba entre 0,19 y 0,87. Los amplicones también contenían una serie de repeticiones de nucleótido único (por ejemplo, poli-A y poli-T), dinucleótidas y trinucleótidas. 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 no hay llamadas.

Tabla 11 Exactitud en el nivel del amplicón de la variante 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	36450499	36450591	93	93	Inserción y delección	0,22	76 167	0	0	100
2	1	109465122	109465200	79	79	Poli-A (5), Poli-C (5), inserción y delección	0,38	64 701	0	0	100
3	1	218353867	218353957	91	91	Inserción y delección	0,4	74 529	0	0	100
4	1	223906657	223906748	92	92	Inserción y delección	0,49	75 348	0	0	100
5	1	228526602	228526682	81	81	Poli-G (5)	0,69	66 339	0	0	100
6	1	236372039	236372108	70	70	Poli-T (10), inserción y delección	0,39	57 330	0	0	100
7	1	247812041	247812128	88	88	Poli-A (5), CT (3), TAA (3), inserción y delección	0,27	72 072	0	0	100
8	2	55862774	55862863	90	90	Inserción y delección	0,28	73 710	0	0	100
9	2	87003930	87004009	80	80	Inserción y delección	0,38	65 520	0	0	100
10	2	177016721	177016805	85	81	N/D	0,65	66 339	0	0	100
11	2	186625727	186625801	75	75	Poli-A (8)	0,35	61 425	0	0	100
12	2	190323504	190323591	88	88	Poli-T (5)	0,42	72 072	0	0	100
13	2	200796740	200796826	87	87	Poli-T (5), inserción y delección	0,31	71 253	0	0	100
14	2	212245049	212245139	91	91	Poli-T (5), Poli-A (6), inserción y delección	0,3	74 529	0	0	100
15	2	228147052	228147144	93	93	Inserción y delección	0,43	76 167	0	0	100
16	2	235016350	235016422	73	73	Poli-T (5), inserción y delección	0,42	59 787	0	0	100
17	3	4466229	4466321	93	93	AT(3), inserción y delección	0,27	74 823	0	1344	98,2
18	3	46620561	46620643	83	83	N/D	0,43	67 977	0	0	100

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
19	3	49851331	49851400	70	70	CT(3), inserción y deleción	0,49	57 330	0	0	100
20	3	189713161	189713248	88	88	Poli-A (5), Poli-T (5), Poli-A (9), TG (3)	0,41	72 072	0	0	100
21	3	190106030	190106104	75	74	Inserción y deleción	0,57	60 543	0	63	99,9
22	4	2233667	2233744	78	78	Poli-A (6)	0,26	63 882	0	0	100
23	4	7780541	7780637	97	97	Poli-G (6), Poli-T (5), Poli-A (5)	0,42	79 443	0	0	100
24	4	15688604	15688681	78	78	N/D	0,29	63 882	0	0	100
25	4	56236521	56236586	66	62	Poli-A (5), inserción y deleción	0,36	50 778	0	0	100
26	4	102839244	102839314	71	69	Poli-A (5)	0,46	56 511	0	0	100
27	4	164446743	164446804	62	62	Poli-A (7), inserción y deleción	0,27	50 778	0	0	100
28	5	1882081	1882158	78	75	N/D	0,78	61 425	0	0	100
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	68 796	0	0	100
30	5	41069808	41069871	64	64	N/D	0,39	52 416	0	0	100
31	5	74077114	74077196	83	83	Poli-A (6), inserción y deleción	0,3	67 977	0	0	100
32	5	147475343	147475409	67	67	Poli-T (5)	0,37	54 873	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	74 529	0	0	100
34	5	155662213	155662287	75	75	Inserción y deleción	0,43	61 425	0	0	100
35	6	6318713	6318814	102	102	Poli-G (6)	0,68	83 538	0	0	100
36	6	24949983	24950074	92	92	Inserción y deleción	0,63	75 348	0	0	100
37	6	31084900	31084999	100	94	GCT(5), inserción y deleción	0,61	76 608	0	378	99,5
38	6	32147987	32148084	98	98	Poli-T (5), TCT(3), CTT(3)	0,55	80 262	0	0	100
39	6	32986864	32986958	95	95	Inserción y deleción	0,53	77 805	0	0	100
40	6	33408498	33408583	86	86	Poli-C (6)	0,7	70 434	0	0	100

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
41	6	41647401	41647495	95	94	Poli-G (5), inserción y delección	0,61	76 986	0	0	100
42	6	112435865	112435955	91	91	Poli-A (5)	0,44	74 529	0	0	100
43	7	22202176	22202148	73	73	N/D	0,44	59 787	0	0	100
44	7	66276100	66276187	88	88	Inserción y delección	0,35	72 072	0	0	100
45	7	77365735	77365821	87	87	Poli-A (7), AG (4)	0,26	71 253	0	0	100
46	7	110939946	110940030	85	85	Inserción y delección	0,38	69 615	0	0	100
47	7	128533468	128533557	90	90	Poli-G (5), inserción y delección	0,62	73 710	0	0	100
48	7	149503875	149503965	91	91	Poli-G (6), Poli-C (6), inserción y delección	0,71	74 529	0	0	100
49	7	154404519	154404599	81	66	N/D	0,31	54 054	0	0	100
50	7	156476507	156476599	93	93	Inserción y delección	0,35	76 167	0	0	100
51	8	1817312	1817394	83	83	N/D	0,42	67 977	0	0	100
52	8	24811020	24811109	90	89	Poli-G (7), CTC (4), inserción y delección	0,61	72 171	0	720	99,0
53	8	76518625	76518691	67	67	Inserción y delección	0,3	54 873	0	0	100
54	9	103054909	103055006	98	98	Poli-G (6)	0,67	80 262	0	0	100
55	9	105586150	105586214	65	65	Inserción y delección	0,32	53 235	0	0	100
56	9	107620823	107620918	96	96	N/D	0,49	78 624	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	67 977	0	0	100
58	9	138995345	138995441	97	97	Poli-C (6), inserción y delección	0,68	79 443	0	0	100
59	10	5987120	5987198	79	78	Poli-G (5), inserción y delección	0,47	63 882	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	74 529	0	0	100
61	10	27317777	27317855	79	79	Poli-T (5)	0,3	64 701	0	0	100
62	10	33018351	33018440	90	90	Poli-A (5), Poli-T (5)	0,2	73 710	0	0	100
63	10	45084159	45084253	95	95	Inserción y delección	0,35	77 805	0	0	100

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
64	10	55892599	55892687	89	88	AC(11), inserción y delección	0,42	71 747	0	325	99,5
65	10	101611250	101611329	80	80	N/D	0,49	65 520	0	0	100
66	10	118351373	118351453	81	81	N/D	0,51	66 339	0	0	100
67	11	8159816	8159912	97	96	N/D	0,45	78 624	0	0	100
68	11	30177648	30177717	70	70	Inserción y delección	0,46	57 330	0	0	100
69	11	47470345	47470444	100	100	N/D	0,65	81 900	0	0	100
70	11	59837679	59837740	62	62	Inserción y delección	0,37	50 778	0	0	100
71	11	64418856	64418957	102	102	N/D	0,59	83 538	0	0	100
72	11	93529612	93529684	73	73	Poli-A (5)	0,4	59 787	0	0	100
73	11	101347052	101347136	85	85	N/D	0,42	69 615	0	0	100
74	11	102477336	102477426	91	91	Poli-G (6)	0,55	74 529	0	0	100
75	11	118406285	118406369	85	85	Inserción y delección	0,53	69 615	0	0	100
76	11	120357801	120357885	85	85	Poli-A (5), CA(3), inserción y delección	0,34	69 615	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	69 615	0	0	100
78	12	2834770	2834853	84	84	Poli-C (5), inserción y delección	0,52	68 796	0	0	100
79	12	26811004	26811096	93	93	Poli-A (7), AC (4)	0,33	76 167	0	0	100
80	12	30881766	30881846	81	81	N/D	0,49	66 339	0	0	100
81	12	88474105	88474175	71	71	Poli-A (6)	0,35	58 149	0	0	100
82	12	120966872	120966966	95	95	Poli-G (5)	0,68	77 805	0	0	100
83	13	24167504	24167576	73	73	N/D	0,52	59 787	0	0	100
84	13	25816961	25817049	89	88	Poli-A (5), Poli-T (7), Poli-A (7), inserción y delección	0,22	72 072	0	0	100
85	13	44880112	44880200	89	89	Inserción y delección	0,49	72 891	0	0	100
86	13	77665218	77665294	77	77	Inserción y delección	0,39	63 063	0	0	100
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	54 873	0	0	100
88	14	39517884	39517966	83	83	N/D	0,25	67 977	0	0	100

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
89	14	46958962	46959034	73	72	Poli-T (5), inserción y delección	0,19	58 642	0	326	99,4
90	14	58050030	58050110	81	81	Inserción y delección	0,38	66 339	0	0	100
91	14	82390559	82390649	91	91	Inserción y delección	0,35	74 529	0	0	100
92	14	92549544	92549609	66	66	Poli-A (5)	0,41	54 054	0	0	100
93	14	102808496	102808589	94	94	Inserción y delección	0,62	76 986	0	0	100
94	15	43170751	43170848	98	96	Poli-C (5)	0,45	78 624	0	0	100
95	15	63446149	63446216	68	68	Inserción y delección	0,25	55 692	0	0	100
96	15	77879807	77879901	95	93	Poli-G (5), inserción y delección	0,68	76 167	0	0	100
97	15	81625334	81625428	95	95	Poli-T (6)	0,43	77 805	0	0	100
98	15	85438263	85438334	72	71	Inserción y delección	0,65	58 149	0	0	100
99	15	89817413	89817503	91	91	N/D	0,36	74 529	0	0	100
100	15	89864274	89864343	70	70	Inserción y delección	0,56	57 330	0	0	100
101	16	1894910	1894972	63	63	N/D	0,27	51 597	0	0	100
102	16	28997904	28997998	95	95	Poli-C (5)	0,67	77 805	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0,41	71 253	0	0	100
104	16	57954406	57954509	104	104	Poli-C (5)	0,67	85 176	0	0	100
105	16	85706375	85706465	91	91	Poli-T (5), inserción y delección	0,37	74 529	0	0	100
106	17	3563920	3564008	89	89	GC(3)	0,64	72 891	0	0	100
107	17	3594191	3594277	87	87	Poli-C (5), inserción y delección	0,67	71 247	0	6	100
108	17	3970090	3970180	91	91	Inserción y delección	0,46	74 529	0	0	100
109	17	16084945	16085037	93	93	Inserción y delección	0,26	76 167	0	0	100
110	17	33998759	33998849	91	89	Poli-T (5)	0,54	72 891	0	0	100
111	17	39589691	39589774	84	82	Poli-A (13), inserción y delección (x2)	0,29	66 343	27	788	98,8

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
112	17	41244394	41244484	91	91	Poli-A (5)	0,34	74 529	0	0	100
113	17	45438866	45438957	92	92	Poli-A (7), AT (3), AT (4), AT (4), inserción y deleción	0,26	75 348	0	0	100
114	17	61502432	61502510	79	79	Inserción y deleción	0,41	64 413	0	288	99,6
115	17	64023582	64023667	86	86	Poli-T (7)	0,22	70 434	0	0	100
116	17	72308237	72308320	84	84	GAG(3)	0,62	68 796	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	54 873	0	0	100
118	18	6980478	6980568	91	91	N/D	0,37	74 529	0	0	100
119	18	9888026	9888094	69	69	Poli-A (6), TG (3)	0,43	56 511	0	0	100
120	18	38836999	38837073	75	75	Poli-A (5), inserción y deleción	0,37	61 425	0	0	100
121	18	47405382	47405462	81	81	CTC(3), inserción y deleción	0,47	66 339	0	0	100
122	18	54815665	54815749	85	85	CT(3), inserción y deleción	0,45	69 615	0	0	100
123	18	59773996	59774060	65	65	N/D	0,48	53 235	0	0	100
124	19	625143	625241	99	99	N/D	0,59	81 081	0	0	100
125	19	18121418	18121491	74	74	N/D	0,68	60 605	1	0	100
126	19	18186574	18186643	70	70	N/D	0,64	57 330	0	0	100
127	20	746056	746149	94	94	N/D	0,61	76 986	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	67 158	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	62 244	0	0	100
130	20	21766821	21766890	70	70	GT(3), TG(4), inserción y deleción	0,46	57 330	0	0	100
131	20	25278421	25278521	101	101	Inserción y deleción	0,63	82 719	0	0	100
132	20	50897302	50897368	67	67	Inserción y deleción	0,36	54 873	0	0	100
133	20	62331904	62331994	91	88	Poli-G (6)	0,73	72 072	0	0	100
134	20	62690860	62690946	87	87	Inserción y deleción	0,57	71 253	0	0	100
135	21	30300823	30300888	66	66	Inserción y deleción	0,35	54 054	0	0	100



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
136	21	33694176	33694273	98	98	Poli-T (6), CA (3)	0,54	80 262	0	0	100
137	21	36710706	36710792	87	87	GT(3), inserción y delección	0,39	71 253	0	0	100
138	21	46644924	46644992	69	69	Poli-A (6), AG (3), inserción y delección	0,32	56 439	0	72	99,9
139	21	46705575	46705664	90	90	Poli-T (5), Poli-A (6)	0,5	73 710	0	0	100
140	22	25750774	25750873	100	100	Inserción y delección	0,63	81 900	0	0	100
141	22	32439233	32439329	97	97	N/D	0,68	79 443	0	0	100
142	22	37409844	37409940	97	97	Inserción y delección	0,46	79 443	0	0	100
143	22	37637596	37637694	99	99	N/D	0,6	81 081	0	0	100
144	22	47081347	47081438	92	92	Inserción y delección	0,66	75 348	0	0	100
145	X	15870424	15870492	69	69	Poli-T (5)	0,26	56 511	0	0	100
146	X	135288543	135288611	69	69	Poli-C (5)	0,62	56 511	0	0	100
147	X	135290777	135290847	71	71	N/D	0,52	58 149	0	0	100
148	Y	2655397	2655461	65	0	N/D	0,55	0	0	0	N/D
149	Y	2655519	2655609	91	0	N/D	0,48	0	0	0	N/D
150	Y	2655609	2655679	71	0	Poli-A (5)	0,37	0	0	0	N/D

Los resultados de la secuenciación de la muestra NA12878 se compararon con un genotipo muy seguro definido para la NA12878 por el Instituto de estándares y tecnología de EE. UU. (NIST, del inglés "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 del NIST. El resultado fueron 10 000 coordenadas por duplicado para llevar a cabo la comparación. Las llamadas de bases no variantes se compararon con el conjunto hg19 de la secuencia de referencia del genoma humano. La exactitud de los resultados se muestra en la [Tabla 12](#).

Tabla 12 Coincidencia germinal de la muestra NA12878 con la base de datos del NIST

Muestra	N.º de amplicones	Promedio del 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 variante germinal de nueve experimentos, el instrumento NextSeq 550Dx puede secuenciar de forma sistemática:

- ▶ Contenido de GC  $\geq 19$  % (se llamaron correctamente todas las bases de 819 de los amplicones secuenciados con un 19 % de contenido de GC y se obtuvo un índice de ausencia de llamadas del 0,6 %)
- ▶ Contenido de GC  $\leq 87$  % (se llamaron correctamente todas las bases de 819 de los amplicones secuenciados con un 87 % de contenido de GC y no hubo ninguna ausencia de llamadas)
- ▶ Longitudes de poli-A  $\leq 9$  (se llamaron correctamente todas las bases de 819 de los amplicones secuenciados con una repetición de poli-A de nueve nucleótidos y no hubo ninguna ausencia de llamadas)
- ▶ Longitudes de poli-T  $\leq 10$  (se llamaron correctamente todas las bases de 819 de los amplicones secuenciados con una repetición de poli-T de diez nucleótidos y no hubo ninguna ausencia de llamadas)
- ▶ Longitudes de poli-G  $\leq 7$  (se llamaron correctamente todas las bases de 819 de los amplicones secuenciados con una repetición de poli-G de siete nucleótidos y se obtuvo un índice de ausencia de llamadas del 1,0 %)
- ▶ Longitudes de poli-C  $\leq 6$  (se llamaron correctamente todas las bases de 2457 de los amplicones secuenciados con una repetición de poli-C de seis nucleótidos y no hubo ninguna ausencia de llamadas)
- ▶ Longitudes de repetición dinucleótida  $\leq 11\times$  (se llamaron correctamente todas las bases de 819 de los amplicones secuenciados con una repetición dinucleótida de  $11\times$  y se obtuvo un índice de ausencia de llamadas del 0,5 %)
- ▶ Longitudes de repetición trinucleótida  $\leq 5\times$  (se llamaron correctamente todas las bases de 819 de los amplicones secuenciados con repetición trinucleótida de  $5\times$  y se obtuvo un índice de ausencia de llamadas del 0,5 %)
- ▶ Longitudes de inserción  $\leq 24$  (se llamaron correctamente 66 343 de 66 370 bases de 819 de los amplicones secuenciados con 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 incluía la inserción de 24 nucleótidos)
- ▶ Longitudes de delección  $\leq 25$  (se llamaron correctamente todas las bases de 2457 de los amplicones secuenciados con una delección de 25 nucleótidos y no hubo ninguna ausencia de llamadas)

### Somático

El estudio descrito a continuación se utilizó para evaluar la exactitud de la llamada de variantes del módulo Germline Somatic (Variante somática) en el instrumento NextSeq 550Dx con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos).

En este estudio se empleó un ensayo típico 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 muestra GM12877 se diluyó con el ADN de muestra GM12878 para crear GM12877-D5 y GM12877-D7 como un conjunto de variantes heterocigóticas únicas con frecuencias de variantes próximas a entre el 5 % y el 7 %. El ADN de muestra GM12878 se diluyó del mismo modo con el ADN de muestra GM12877 para crear GM12878-D5 y GM12878-D7. Cada una de las muestras se evaluó por triplicado excepto las muestras diluidas, que se evaluaron por sextuplicado. Se llevaron 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 ha determinado 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 han definido regiones genómicas fiables según este método de referencia, a menos que se haya indicado lo contrario.

Tabla 13 Resumen de la coincidencia somática

Criterios	Total de observaciones <sup>1</sup>	Resultado por observación <sup>2</sup>	Resultado por experimento <sup>3</sup>
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

<sup>1</sup> Se calcula de la siguiente manera: número de muestras por experimento (42) × número de experimentos (9) = 378.

<sup>2</sup> Valor observado más bajo por duplicado de muestra en los nueve experimentos.

<sup>3</sup> Valor 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, en las que los resultados de las variantes se comparan con el método de referencia compuesto y con características bien definidas que se utiliza para llevar a cabo los cálculos del 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 y las deleciones, los resultados de la base sin variantes se comparan con el conjunto hg19 de la secuencia de referencia del genoma humano para los cálculos de la coincidencia de porcentaje negativo.

Tabla 14 Coincidencia de variantes somáticas por muestra

Muestra	Promedio del 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 las 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 variantes somáticas 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

Muestra	SNV			Inserciones			Deleciones		
	Previsto	TP	FN	Previsto	TP	FN	Previsto	TP	FN
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 (Tabla 16). Se detectó un total de 71 inserciones y deleciones 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 inserciones y deleciones somáticas

Tipo de variante	Variantes 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), dinucleótidas y trinucleótidas. 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 no hay llamadas.

Tabla 17 Exactitud en el nivel del amplicón de la variante 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	36450499	36450591	93	93	Inserción y delección	0,22	35 066	0	88	99,7
2	1	109465122	109465200	79	79	Poli-A (5), Poli-C (5), inserción y delección	0,38	29 827	0	35	99,9
3	1	218353867	218353957	91	91	Inserción y delección	0,4	34 202	0	283	99,2
4	1	223906657	223906748	92	92	Inserción y delección	0,49	34 613	0	163	99,5
5	1	228526602	228526682	81	81	Poli-G (5)	0,69	30 571	0	47	99,8
6	1	236372039	236372108	70	70	Poli-T (10), inserción y delección	0,39	26 452	0	8	100,0
7	1	247812041	247812128	88	88	Poli-A (5), CT (3), TAA (3), inserción y delección	0,27	33 148	0	116	99,7
8	2	55862774	55862863	90	90	Inserción y delección	0,28	33 928	0	92	99,7
9	2	87003930	87004009	80	80	Inserción y delección	0,38	30 218	0	22	99,9
10	2	177016721	177016805	85	81	N/D	0,65	30 616	0	2	>99,9
11	2	186625727	186625801	75	75	Poli-A (8)	0,35	28 017	0	499	98,3
12	2	190323504	190323591	88	88	Poli-T (5)	0,42	33 207	0	57	99,8
13	2	200796740	200796826	87	87	Poli-T (5), inserción y delección	0,31	32 524	9	718	97,8
14	2	212245049	212245139	91	91	Poli-T (5), Poli-A (6), inserción y delección	0,3	33 972	0	456	98,7
15	2	228147052	228147144	93	93	N/D	0,43	35 051	0	103	99,7
16	2	235016350	235016422	73	73	Poli-T (5), inserción y delección	0,42	27 459	0	136	99,5
17	3	4466229	4466321	93	93	AT(3), inserción y delección	0,27	34 534	0	620	98,2

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
18	3	46620561	46620643	83	83	N/D	0,43	31 339	0	44	99,9
19	3	49851331	49851400	70	70	CT(3), inserción y deleción	0,49	26 373	0	87	99,7
20	3	189713161	189713248	88	88	Poli-A (5), Poli-T (5), Poli-A (9), TG (3)	0,41	32 829	0	857	97,5
21	3	190106030	190106104	75	74	Inserción y deleción	0,57	27 925	0	47	99,8
22	4	2233667	2233744	78	78	Poli-A (6)	0,26	29 327	4	162	99,4
23	4	7780541	7780637	97	97	Poli-G (6), Poli-T (5), Poli-A (5)	0,42	36 585	0	117	99,7
24	4	15688604	15688681	78	78	N/D	0,29	29 427	0	57	99,8
25	4	56236521	56236586	66	62	Poli-A (5), inserción y deleción	0,36	23 356	5	75	99,7
26	4	102839244	102839314	71	69	Poli-A (5)	0,46	25 942	0	140	99,5
27	4	164446743	164446804	62	62	Poli-A (7), inserción y deleción	0,27	22 944	0	560	97,6
28	5	1882081	1882158	78	75	N/D	0,78	28 299	0	53	99,8
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	31 658	0	94	99,7
30	5	41069808	41069871	64	64	N/D	0,39	24 120	0	72	99,7
31	5	74077114	74077196	83	83	Poli-A (6), inserción y deleción	0,3	31 297	0	77	99,8
32	5	147475343	147475409	67	67	Poli-T (5)	0,37	25 277	0	55	99,8
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	34 308	0	90	99,7
34	5	155662213	155662287	75	75	Inserción y deleción	0,43	28 266	0	163	99,4
35	6	6318713	6318814	102	102	Poli-G (6)	0,68	38 489	0	67	99,8
36	6	24949983	24950074	92	92	Inserción y deleción	0,63	34 730	0	46	99,9
37	6	31084900	31084999	100	94	GCT(5), inserción y deleción	0,61	35 057	0	483	98,6
38	6	32147987	32148084	98	98	Poli-T (5), TCT(3), CTT (3)	0,55	36 647	0	406	98,9

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
39	6	32986864	32986958	95	95	Inserción y delección	0,53	35 681	0	238	99,3
40	6	33408498	33408583	86	86	Poli-C (6)	0,7	32 438	0	70	99,8
41	6	41647401	41647495	95	94	Poli-G (5), inserción y delección	0,61	35 441	0	91	99,7
42	6	112435865	112435955	91	91	Poli-A (5)	0,44	34 354	0	44	99,9
43	7	22202176	22202148	73	73	N/D	0,44	27 575	0	28	99,9
44	7	66276100	66276187	88	88	Inserción y delección	0,35	33 060	0	213	99,4
45	7	77365735	77365821	87	87	Poli-A (7), AG (4)	0,26	32 423	0	489	98,5
46	7	110939946	110940030	85	85	Inserción y delección	0,38	32 074	0	56	99,8
47	7	128533468	128533557	90	90	Poli-G (5), inserción y delección	0,62	33 791	0	281	99,2
48	7	149503875	149503965	91	91	Poli-G (6), Poli-C (6), inserción y delección	0,71	34 316	0	82	99,8
49	7	154404519	154404599	81	66	N/D	0,31	24 901	0	47	99,8
50	7	156476507	156476599	93	93	Inserción y delección	0,35	35 067	0	87	99,8
51	8	1817312	1817394	83	83	N/D	0,42	31 365	0	9	>99,9
52	8	24811020	24811109	90	89	Poli-G (7), CTC (4), inserción y delección	0,61	32 781	0	890	97,4
53	8	76518625	76518691	67	67	Inserción y delección	0,3	25 228	0	146	99,4
54	9	103054909	103055006	98	98	Poli-G (6)	0,67	36 968	0	76	99,8
55	9	105586150	105586214	65	65	Inserción y delección	0,32	24 472	0	100	99,6
56	9	107620823	107620918	96	96	N/D	0,49	36 203	0	85	99,8
57	9	123769149	123769231	83	83	AT(3)	0,37	31 329	0	45	99,9
58	9	138995345	138995441	97	97	Poli-C (6), inserción y delección	0,68	36 472	0	201	99,5
59	10	5987120	5987198	79	78	Poli-G (5), inserción y delección	0,47	29 473	0	11	>99,9
60	10	11784629	11784726	98	91	GC(3)	0,87	34 188	0	213	99,4
61	10	27317777	27317855	79	79	Poli-T (5)	0,3	29 843	0	19	99,9

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
62	10	33018351	33018440	90	90	Poli-A (5), Poli-T (5)	0,2	33 968	0	68	99,8
63	10	45084159	45084253	95	95	Inserción y delección	0,35	35 829	0	81	99,8
64	10	55892599	55892687	89	88	AC(11), inserción y delección	0,42	32 098	88	2048	93,8
65	10	101611250	101611329	80	80	N/D	0,49	30 217	0	28	99,9
66	10	118351373	118351453	81	81	N/D	0,51	30 531	0	96	99,7
67	11	8159816	8159912	97	96	N/D	0,45	36 105	0	192	99,5
68	11	30177648	30177717	70	70	Inserción y delección	0,46	26 318	0	153	99,4
69	11	47470345	47470444	100	100	N/D	0,65	37 785	0	24	99,9
70	11	59837679	59837740	62	62	Inserción y delección	0,37	23 368	0	68	99,7
71	11	64418856	64418957	102	102	N/D	0,59	38 546	0	10	>99,9
72	11	93529612	93529684	73	73	Poli-A (5)	0,4	27 516	0	78	99,7
73	11	101347052	101347136	85	85	N/D	0,42	32 083	0	48	99,9
74	11	102477336	102477426	91	91	Poli-G (6)	0,55	34 047	0	369	98,9
75	11	118406285	118406369	85	85	Inserción y delección	0,53	32 065	0	74	99,8
76	11	120357801	120357885	85	85	Poli-A (5), CA(3), inserción y delección	0,34	32 083	0	47	99,9
77	11	125769313	125769397	85	85	GA(3)	0,52	32 103	0	27	99,9
78	12	2834770	2834853	84	84	Poli-C (5), inserción y delección	0,52	31 645	16	525	98,3
79	12	26811004	26811096	93	93	Poli-A (7), AC (4)	0,33	34 824	0	330	99,1
80	12	30881766	30881846	81	81	N/D	0,49	30 497	0	121	99,6
81	12	88474105	88474175	71	71	Poli-A (6)	0,35	26 773	0	65	99,8
82	12	120966872	120966966	95	95	Poli-G (5)	0,68	35 830	9	72	99,8
83	13	24167504	24167576	73	73	N/D	0,52	27 498	0	114	99,6
84	13	25816961	25817049	89	88	Poli-A (5), Poli-T (7), Poli-A (7), inserción y delección	0,22	32 824	0	566	98,3
85	13	44880112	44880200	89	89	Inserción y delección	0,49	33 574	0	77	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
86	13	77665218	77665294	77	77	Inserción y delección	0,39	29 075	0	31	99,9
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	25 313	0	13	99,9
88	14	39517884	39517966	83	83	N/D	0,25	31 360	0	22	99,9
89	14	46958962	46959034	73	72	Poli-T (5), inserción y delección	0,19	26 499	0	717	97,4
90	14	58050030	58050110	81	81	Inserción y delección	0,38	30 494	0	133	99,6
91	14	82390559	82390649	91	91	Inserción y delección	0,35	34 313	0	86	99,7
92	14	92549544	92549609	66	66	Poli-A (5)	0,41	24 555	0	1527	94,1
93	14	102808496	102808589	94	94	Inserción y delección	0,62	35 472	0	69	99,8
94	15	43170751	43170848	98	96	Poli-C (5)	0,45	36 264	0	24	99,9
95	15	63446149	63446216	68	68	Inserción y delección	0,25	25 667	0	37	99,9
96	15	77879807	77879901	95	93	Poli-G (5), inserción y delección	0,68	34 745	0	432	98,8
97	15	81625334	81625428	95	95	Poli-T (6)	0,43	35 870	0	40	99,9
98	15	85438263	85438334	72	71	Inserción y delección	0,65	26 762	0	76	99,7
99	15	89817413	89817503	91	91	N/D	0,36	34 286	0	112	99,7
100	15	89864274	89864343	70	70	Inserción y delección	0,56	26 449	0	11	>99,9
101	16	1894910	1894972	63	63	N/D	0,27	23 809	0	5	>99,9
102	16	28997904	28997998	95	95	Poli-C (5)	0,67	35 860	0	50	99,9
103	16	53682908	53682994	87	87	TA(3)	0,41	32 835	0	60	99,8
104	16	57954406	57954509	104	104	Poli-C (5)	0,67	39 177	0	144	99,6
105	16	85706375	85706465	91	91	Poli-T (5), inserción y delección	0,37	34 075	0	323	99,1
106	17	3563920	3564008	89	89	GC(3)	0,64	33 632	0	11	>99,9
107	17	3594191	3594277	87	87	Poli-C (5), inserción y delección	0,67	32 752	0	134	99,6
108	17	3970090	3970180	91	91	Inserción y delección	0,46	34 343	0	82	99,8
109	17	16084945	16085037	93	93	Inserción y delección	0,26	35 077	0	78	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
110	17	33998759	33998849	91	89	Poli-T (5)	0,54	33 553	0	89	99,7
111	17	39589691	39589774	84	82	Poli-A (13), inserción y deleción (×2)	0,29	30 554	53	2296	92,9
112	17	41244394	41244484	91	91	Poli-A (5)	0,34	34 360	0	38	99,9
113	17	45438866	45438957	92	92	Poli-A (7), AT (3), AT (4), AT (4), inserción y deleción	0,26	34 367	0	418	98,8
114	17	61502432	61502510	79	79	Inserción y deleción	0,41	29 751	0	119	99,6
115	17	64023582	64023667	86	86	Poli-T (7)	0,22	32 176	0	340	99,0
116	17	72308237	72308320	84	84	GAG(3)	0,62	31 604	7	141	99,5
117	18	2616456	2616522	67	67	GA(3)	0,31	25 273	8	45	99,8
118	18	6980478	6980568	91	91	N/D	0,37	34 386	0	12	>99,9
119	18	9888026	9888094	69	69	Poli-A (6), TG (3)	0,43	25 692	0	399	98,5
120	18	38836999	38837073	75	75	Poli-A (5), inserción y deleción	0,37	27 923	0	893	96,9
121	18	47405382	47405462	81	81	CTC(3), inserción y deleción	0,47	30 598	0	20	99,9
122	18	54815665	54815749	85	85	CT(3), inserción y deleción	0,45	31 969	0	161	99,5
123	18	59773996	59774060	65	65	N/D	0,48	24 531	0	48	99,8
124	19	625143	625241	99	99	N/D	0,59	37 298	0	124	99,7
125	19	18121418	18121491	74	74	N/D	0,68	27 881	0	109	99,6
126	19	18186574	18186643	70	70	N/D	0,64	26 442	0	26	99,9
127	20	746056	746149	94	94	N/D	0,61	35 501	0	31	99,9
128	20	10633195	10633276	82	82	AC(3)	0,59	30 951	0	72	99,8
129	20	17705633	17705708	76	76	CT(3)	0,58	28 686	0	42	99,9
130	20	21766821	21766890	70	70	GT(3), TG(4), inserción y deleción	0,46	26 372	0	88	99,7
131	20	25278421	25278521	101	101	Inserción y deleción	0,63	38 159	0	20	99,9

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
132	20	50897302	50897368	67	67	Inserción y delección	0,36	25 188	0	544	97,9
133	20	62331904	62331994	91	88	Poli-G (6)	0,73	32 969	0	309	99,1
134	20	62690860	62690946	87	87	Inserción y delección	0,57	32 818	0	77	99,8
135	21	30300823	30300888	66	66	Inserción y delección	0,35	24 758	9	181	99,2
136	21	33694176	33694273	98	98	Poli-T (6), CA (3)	0,54	36 902	0	160	99,6
137	21	36710706	36710792	87	87	GT(3), inserción y delección	0,39	32 841	0	48	99,9
138	21	46644924	46644992	69	69	Poli-A (6), AG (3), inserción y delección	0,32	25 939	0	280	98,9
139	21	46705575	46705664	90	90	Poli-T (5), Poli-A (6)	0,5	33 942	0	78	99,8
140	22	25750774	25750873	100	100	Inserción y delección	0,63	37 733	0	86	99,8
141	22	32439233	32439329	97	97	N/D	0,68	36 617	0	49	99,9
142	22	37409844	37409940	97	97	Inserción y delección	0,46	36 525	0	162	99,6
143	22	37637596	37637694	99	99	N/D	0,6	37 398	0	24	99,9
144	22	47081347	47081438	92	92	Inserción y delección	0,66	34 754	0	22	99,9
145	X	15870424	15870492	69	69	Poli-T (5)	0,26	26 046	0	36	99,9
146	X	135288543	135288611	69	69	Poli-C (5)	0,62	26 019	0	63	99,8
147	X	135290777	135290847	71	71	N/D	0,52	26 780	0	58	99,8
148	Y	2655397	2655461	65	0	N/D	0,55	0	0	0	N/D
149	Y	2655519	2655609	91	0	N/D	0,48	0	0	0	N/D
150	Y	2655609	2655679	71	0	Poli-A (5)	0,37	0	0	0	N/D

Los resultados de la secuenciación de la muestra GM12878 se compararon con un genotipo muy fiable definido para la NA12878 por el Instituto de estándares y tecnología de EE. UU. (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 del NIST. El resultado fueron 10 000 coordenadas por duplicado para llevar a cabo la comparación. Las llamadas de bases no variantes se compararon con el conjunto hg19 de la secuencia de referencia del genoma humano. La exactitud de los resultados se muestra en la [Tabla 18](#).

Tabla 18 Coincidencia somática de la muestra GM12878 con la base de datos del NIST

Muestra	N.º de amplicones	Promedio del 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 variante somática de nueve experimentos, el instrumento NextSeq 550Dx puede secuenciar de forma sistemática:

- ▶ Contenido de GC  $\geq 19\%$  (se llamaron correctamente todas las bases de 378 de los amplicones secuenciados con un 19 % de contenido de GC y se obtuvo un índice de ausencia de llamadas del 2,6 %)
- ▶ Contenido de GC  $\leq 87\%$  (se llamaron correctamente todas las bases de 378 de los amplicones secuenciados con un 87 % de contenido de GC y se obtuvo un índice de ausencia de llamadas del 0,6 %)
- ▶ Longitudes de poli-A  $\leq 9$  (se llamaron correctamente todas las bases de 378 de los amplicones secuenciados con una repetición de poli-A de nueve nucleótidos y se obtuvo un índice de ausencia de llamadas del 2,5 %)
- ▶ Longitudes de poli-T  $\leq 10$  (se llamaron correctamente todas las bases de 378 de los amplicones secuenciados con 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  $\leq 6$  (se llamaron correctamente todas las bases de 2268 de los amplicones secuenciados con una repetición de poli-G de seis nucleótidos y se obtuvo un índice de ausencia de llamadas del 0,5 %)
- ▶ Longitudes de poli-C  $\leq 6$  (se llamaron correctamente todas las bases de 756 de los amplicones secuenciados con una repetición de poli-C de seis nucleótidos y se obtuvo un índice de ausencia de llamadas del 0,4 %)
- ▶ Longitudes de repetición dinucleótida  $\leq 4x$  (se llamaron correctamente todas las bases de 1890 de los amplicones secuenciados con una repetición dinucleótida de  $4x$  y se obtuvo un índice de ausencia de llamadas del 0,9 %)
- ▶ Longitudes de repetición trinucleótida  $\leq 5x$  (se llamaron correctamente todas las bases de 378 de los amplicones secuenciados con repetición trinucleótida de  $5x$  y se obtuvo un índice de ausencia de llamadas del 1,4 %)
- ▶ Longitudes de inserción  $\leq 23$  (se llamaron correctamente todas las bases de 378 de los amplicones secuenciados con una inserción de 23 nucleótidos y se obtuvo un índice de ausencia de llamadas del 0,8 %)
- ▶ Longitudes de delección  $\leq 25$  (se llamaron correctamente todas las bases de 1134 de los amplicones secuenciados con 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 comprobando 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 germinal. Se determinaron las contribuciones de precisión analizando los componentes de la varianza con el valor VAF como variable de respuesta y calculando las desviaciones estándar 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 delecciones, 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 total y de FFPE se analizaron en los instrumentos NextSeq 550Dx y MiSeqDx con los flujos de trabajo germinal y somático del kit de amplicones personalizado TruSeq Dx. Se evaluó la coincidencia de frecuencia de variantes de muestras de sangre y FFPE utilizando varios ensayos típicos. La Figura 2 representa la correlación de VAF entre dos instrumentos para un ensayo típico y la Tabla 20 resume dicha correlación mediante un panel de ensayo. Basándose en la estrecha correlación entre los instrumentos MiSeqDx y NextSeq 550Dx, se ha determinado que las características de rendimiento relacionadas con los factores preanalíticos (por ejemplo, los métodos de extracción o las sustancias interferentes) son aplicables a ambos. Consulte el prospecto del kit de amplicones personalizados TruSeq Dx para obtener más información.

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

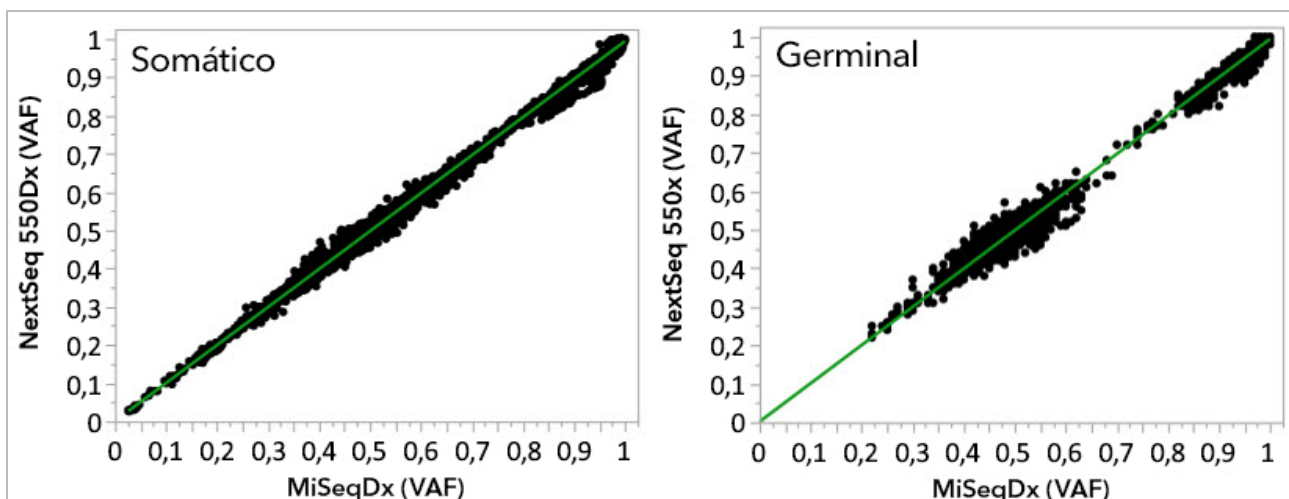


Tabla 20 Resultados de la comparación de métodos con muestras de sangre y FFPE únicas

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

<sup>1</sup> Se eliminaron dos puntos de datos según la limitación indicada para el módulo Germline Variant (Variante germinal).

<sup>2</sup> Coeficiente de determinación de los diagramas de VAF como se muestra en la figura 2.

## Reproducibilidad

La reproducibilidad del instrumento NextSeq 550Dx se evaluó utilizando muestras de Platinum Genomes con un ensayo típico diseñado para buscar una variedad de genes en 12 588 bases (150 amplicones) de 23 cromosomas diferentes. Las pruebas de variantes germinales 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 el kit de amplicones personalizado TruSeq Dx.

Las pruebas se llevaron a cabo en tres sitios externos con un lote del kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos). En cada sitio, se utilizó un único instrumento NextSeq 550Dx. Dos operadores realizaron las pruebas en cada sitio. Cada operador llevó a cabo las pruebas en tres días de inicio no consecutivos para cada tipo de muestra durante un total de 36 experimentos en cada uno de los tres sitios. El resultado fueron 18 experimentos para cada flujo de trabajo germinal y somático.

### Germinal

Variantes germinales con nivel de VAF  $\geq 0,2$  indicadas como llamadas positivas (variante). Para las variantes germinales positivas previstas, se evaluaron los datos de 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) al 95 % calculados con el método Wilson para cada tipo de variante.

Tabla 21 Observaciones de llamadas de variantes germinales 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

Variantes germinales con nivel de VAF  $< 0,2$  indicadas como llamadas negativas (estado natural). Para las variantes germinales negativas previstas, se evaluaron los datos de índice de ausencia de llamada e 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 Wilson.

Tabla 22 Observaciones de llamadas de variantes germinales 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 variantes germinales con nivel de VAF  $\geq 0,2$  y  $< 0,7$  se conocen como heterocigóticas positivas y las variantes con nivel de VAF  $\geq 0,7$  se conocen como homocigóticas positivas. Se utilizaron muestras de variantes germinales 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 es  $\geq 99,999\%$ , lo que indica que  $\geq 99,999\%$  de las variantes heterocigóticas serán heterocigóticas. En relación con el corte superior de VAF de 0,7, el valor Cx es  $\leq 0,001\%$ , lo que indica que  $\leq 0,001\%$  de las variantes heterocigóticas serán homocigóticas. La [Tabla 23](#) resume los resultados por tipo de variante.

Las variantes germinales con nivel de VAF  $\geq 0,2$  y  $< 0,7$  se conocen como heterocigóticas positivas y las variantes con nivel de VAF  $\geq 0,7$  se conocen como homocigóticas positivas. Se utilizaron muestras de variantes germinales 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 es  $\geq 99,999\%$ , lo que indica que  $\geq 99,999\%$  de las variantes heterocigóticas serán heterocigóticas. En relación con el corte superior de VAF de 0,7, el valor Cx es  $\leq 0,001\%$ , lo que indica que  $\leq 0,001\%$  de las variantes heterocigóticas serán homocigóticas. La [Tabla 23](#) resume los resultados por tipo de variante.

Tabla 23 Valores Cx de variantes germinales para variantes heterocigóticas

Tipo de variante	Corte con VAF de 0,2	Corte con 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

### Somático

Variantes somáticas con nivel de VAF  $\geq 0,026$  indicadas 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 (falta de coincidencia 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 Wilson.

Tabla 24 Observaciones de llamadas de variantes somáticas 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 variantes somáticas 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 variantes somáticas 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

Variantes somáticas con nivel de VAF <0,01 indicadas como llamadas negativas (estado natural). Para las variantes somáticas negativas previstas, se evaluaron los datos de índice de ausencia de llamada e índice de llamadas en estado natural correctas. Se determinaron las llamadas en estado natural correctas excluyendo el valor de ausencia de llamadas y restando las llamadas observadas que cayeron en la zona dudosa (niveles de VAF  $\geq 0,01$  y <0,026), así como 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 el porcentaje de las variantes somáticas negativas, 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 Wilson.

Tabla 27 Observaciones de llamadas de variantes somáticas para los resultados negativos previstos

Tipo de variante	Ausencia de llamadas			Llamada correcta						
	Observado	Total	Porcentaje	Dudosa	Incorrecto	Correcto	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 variantes somáticas 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 utilizaron 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 indicada en la [Tabla 28](#).

Tabla 28 Resumen de C95 de variantes somáticas

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



## Rendimiento del kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos)

### Descripción general

El sistema NextSeq 550Dx cuenta con dos kits de reactivos: el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) y el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos). Para demostrar que el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos) puede cumplir los requisitos de rendimiento analítico verificados y validados con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos), se han realizado estudios con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos). Se llevaron a cabo dos preparaciones de bibliotecas con el kit de amplicones personalizados TruSeq Dx, una con el flujo de trabajo germinal y la otra con el flujo de trabajo somático. Las bibliotecas de cada flujo de trabajo se analizaron con tres lotes del kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos) utilizando tres instrumentos NextSeq 550Dx. Además, las pruebas de cada flujo de trabajo incluyeron un único experimento con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos).

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

La verificación con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) demostró que el instrumento NextSeq 550Dx podría detectar variantes a una VAF de 0,05 con un error de tipo II  $\leq 0,05$  y que el corte con 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 a una VAF de 0,05 sea superior o igual a una VAF de 0,026 el 95 % del tiempo y que una posición en estado natural sea inferior a una VAF de 0,026 el 99 % del tiempo. Para cerciorarse de que estas afirmaciones se cumplieran con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos), se repitieron las mediciones en el instrumento NextSeq 550Dx con muestras en estado natural (muestras de LoB) y con muestras que contenían variantes a una VAF de 0,05 (muestras de LoD) utilizando el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos). Después, la proporción de llamadas por encima y por debajo del corte de 0,026 se comparó con las afirmaciones determinadas con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos).

La prueba incluyó dos muestras de LoD, cada una con un conjunto único de variantes objetivo a una VAF de 0,05, y las muestras de LoB correspondientes en estado natural para las variantes objetivo. Para la preparación de bibliotecas, las muestras de LoD y LoB se procesaron por octuplicado y septuplicado, respectivamente, con el kit de amplicones personalizados TruSeq Dx. Las bibliotecas se secuenciaron inicialmente con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) para identificar variantes o coordenadas genómicas para evaluar el LoB/LoD con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos). Todas las variantes con una VAF promedio de entre 0,045 y 0,055 (variantes de LoD) basadas en los resultados del kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) se usaron para el análisis del LoD (N = 51 variantes). Para el análisis del LoB, se evaluaron las 51 coordenadas genómicas correspondientes.

Para la evaluación del kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos), las bibliotecas se secuenciaron en tres experimentos durante tres días consecutivos usando el mismo instrumento y el mismo lote de kit de reactivos. Esta prueba alcanzó un total de 24 duplicados por cada una de las 51 variantes de LoD y 21 duplicados por cada una de las posiciones en estado natural correspondientes. En la [Tabla 29](#), se indica la proporción de llamadas en estado natural con una VAF  $< 0,026$ . En la [Tabla 30](#), se indica la proporción de llamadas de variantes de LoD con una VAF superior o igual a 0,026.

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

Tipo de variante	Posiciones evaluadas	Total de observaciones	N.º de mediciones de VAF $\geq 2,6$ %	Proporción $< 2,6$ %	Intervalo de confianza del 95 % (proporción)
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 LoD con una VAF  $\geq 0,026$  (evaluación de afirmaciones sobre el 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 $\geq 2,6$ %	Intervalo de confianza del 95 % (proporción)
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

### Germinal

El siguiente estudio se llevó a cabo para evaluar la exactitud de la llamada de variantes del módulo Germline Variant (Variante germinal) con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos). Se analizaron doce muestras únicas de Platinum Genome con un ensayo típico. Se realizaron un total de 11 experimentos con tres instrumentos NextSeq 550Dx y tres de los kits de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos). Se determinó la exactitud de las SNV, inserciones y deleciones comparando los resultados con un método de referencia compuesto y con características bien definidas, Platinum Genomes versión 2016-1.0. A modo de referencia, se proporcionan los resultados de precisión de un único experimento de secuenciación con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos). En la [Tabla 31](#) se ofrece un resumen de los resultados.

Tabla 31 Resumen de la coincidencia germinal

Criterios	Total de observaciones (v2.5) <sup>1</sup>	Resultado por observación (v2.5) <sup>2</sup>	Resultado por observación (v2) <sup>3</sup>	Resultado por experimento (v2.5) <sup>4</sup>	Resultado por experimento (v2) <sup>4</sup>
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

<sup>1</sup> Calculado como número de muestras por experimento  $\times$  número de experimentos (96 muestras por experimento  $\times$  11 experimentos = 1056 observaciones).

<sup>2</sup> Valor observado más bajo por duplicado de muestra en todos los experimentos (basado en 11 experimentos con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5).

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

<sup>4</sup> Valor más bajo cuando los datos de cada experimento se analizan en conjunto.

### Somático

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 con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos). Se analizaron diez muestras de Platinum Genome FFPE (dos con variantes diluidas hasta una VAF de 0,05) utilizando un ensayo típico. Se realizaron un total de 11 experimentos con tres instrumentos NextSeq 550Dx y tres lotes del kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos).

Se determinó la exactitud de las SNV, inserciones y deleciones comparando los resultados con un método de referencia compuesto y con características bien definidas, Platinum Genomes versión 2016-1.0. A modo de

referencia, se proporcionan los resultados de precisión de un único experimento de secuenciación con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos). En la [Tabla 32](#) se ofrece un resumen de los resultados.

Tabla 32 Resumen de la coincidencia somática

Criterios	Total de observaciones (v2.5) <sup>1</sup>	Resultado por observación (v2.5) <sup>2</sup>	Resultado por observación (v2) <sup>3</sup>	Resultado por experimento (v2.5) <sup>4</sup>	Resultado por experimento (v2) <sup>4</sup>
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

<sup>1</sup> Calculado como número de muestras por experimento × número de experimentos (48 muestras por experimento × 11 experimentos = 528 observaciones).

<sup>2</sup> Valor observado más bajo por duplicado de muestra en todos los experimentos (basado en 11 experimentos con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5).

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

<sup>4</sup> Valor más bajo cuando los datos de cada experimento se analizan en conjunto.

## Precisión

### Germinal

La precisión del kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos) con el módulo Germline Variant (Variante germinal) se evaluó usando muestras de Platinum Genome y un ensayo típico. La prueba consistió en una única preparación de bibliotecas con el kit de amplicones personalizados TruSeq Dx e incluyó 12 muestras procesadas con ocho duplicados cada una. Las bibliotecas se secuenciaron con tres lotes del kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos) y tres instrumentos NextSeq 550Dx para llegar a un total de nueve experimentos de secuenciación.

Se utilizaron 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 del módulo Germline Variant (Variante germinal) (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, la variante con el valor Cx mínimo para el kit de reactivos NextSeq 550Dx v2.5 (300 ciclos) fue >99,9 %, lo que indica que >99,9 % de las variantes heterocigóticas se considerarían heterocigóticas. En relación con el corte superior de VAF de 0,7, la variante con el valor Cx máximo para el kit de reactivos NextSeq 550Dx v2.5 (300 ciclos) fue <1,5 %, lo que indica que ≤1,5 % de las variantes heterocigóticas se considerarían homocigóticas. La [Tabla 33](#) resume los resultados por tipo de variante. A modo de referencia, se proporcionan los valores Cx del único experimento de secuenciación con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos).

Tabla 33 Valores Cx de variantes germinales para variantes heterocigóticas

Tipo de variante	N	Corte con VAF de 0,2		Corte con VAF de 0,7	
		Valor Cx mínimo (v2.5) <sup>1</sup>	Valor Cx mínimo (v2) <sup>2</sup>	Valor Cx máximo (v2.5) <sup>1</sup>	Valor Cx máximo (v2) <sup>2</sup>
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 %

<sup>1</sup> Valores Cx basados en estimaciones de la desviación estándar total del análisis de componentes de la varianza.

<sup>2</sup> Valores Cx basados en las desviaciones estándar de las muestras.

## Somático

La precisión del kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos) con el módulo Somatic Variant (Variante somática) se evaluó usando muestras de Platinum Genome FFPE y un ensayo típico. La prueba consistió en una única preparación de bibliotecas con el kit de amplicones personalizados TruSeq Dx e incluyó dos muestras con ocho duplicados cada una. Las bibliotecas se secuenciaron con tres lotes del kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos) y tres instrumentos NextSeq 550Dx para llegar a un total de nueve experimentos de secuenciación.

Se utilizaron variantes somáticas con niveles previstos de VAF  $\leq 0,10$  VAF (N = 131 variantes únicas) 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 consideran positivas para la variante). Se determinaron los valores C95 para cada una de las variantes somáticas. Los valores C95 representan la VAF en la que la probabilidad de superar el corte de VAF del módulo Somatic Variant (Variante somática) es del 95 %. En la [Tabla 34](#), se indican los valores C95 más altos por tipo de variante. A modo de referencia, se proporcionan los resultados de C95 de un único experimento de secuenciación con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos).

Tabla 34 Resumen de C95 de variantes somáticas

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

<sup>1</sup> Valores C95 basados en estimaciones de la desviación estándar total del análisis de componentes de la varianza.

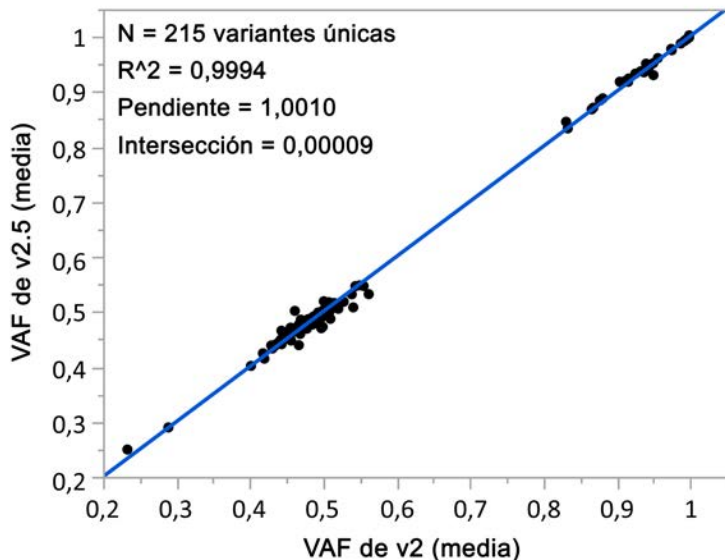
<sup>2</sup> Valores C95 basados en las desviaciones estándar de las muestras.

## Comparación de métodos (kit de reactivos)

### Germinal

Se evaluaron las VAF medias de 215 variantes únicas con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) y el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos) utilizando 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 de cada variante. La [Figura 3](#) muestra la correlación de VAF entre los dos kits de reactivos. Según la sólida correlación de VAF lineal y la similitud de los resultados entre los kits de reactivos, las características de rendimiento verificadas y validadas en un primer momento con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) con el módulo Germline Variant (Variante germinal) se consideraron aplicables al kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos).

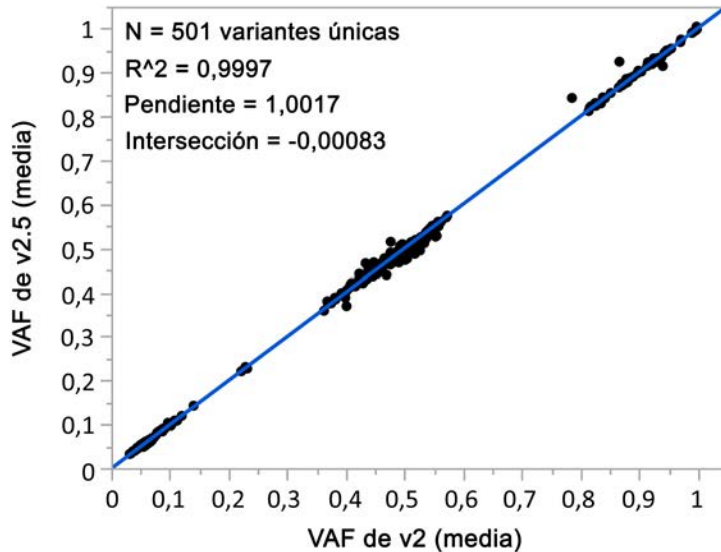
Figura 3 Correlación de la frecuencia de variantes de alelos (VAF) del módulo Germline Variant (Variante germinal) entre el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) y el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos).



### Somático

Se evaluaron las VAF medias de 501 variantes únicas con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) y el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos) utilizando 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 de cada variante única. La [Figura 4](#) muestra la correlación de VAF entre los dos kits de reactivos. Según la correlación de VAF y la similitud de los resultados entre los kits de reactivos, las características de rendimiento verificadas y validadas con el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) con el módulo Somatic Variant (Variante somática) se consideraron aplicables al kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos).

Figura 4 Correlación de la frecuencia de variantes de alelos (VAF) del módulo Somatic Variant (Variante somática) entre el kit de reactivos de rendimiento elevado NextSeq 550Dx v2 (300 ciclos) y el kit de reactivos de rendimiento elevado NextSeq 550Dx v2.5 (300 ciclos).



## Patentes y marcas comerciales

Este documento y su contenido son propiedad 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 utilizarán 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).

© 2021 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](http://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

Para obtener una información detallada sobre los símbolos que aparecen en las etiquetas o en el embalaje del producto, consulte la leyenda que se ofrece en [support.illumina.com](http://support.illumina.com) en la ficha *Documentation and Literature* (Documentación y publicaciones) del kit.