

NextSeq™ 550Dx Instrument

PARA UTILIZAÇÃO EM DIAGNÓSTICO IN VITRO
APENAS PARA EXPORTAÇÃO

Catálogo n.º 20005715

Utilização prevista

O NextSeq 550Dx Instrument destina-se à sequenciação de bibliotecas de ADN quando utilizado com ensaios de diagnóstico *in vitro*. O NextSeq 550Dx Instrument destina-se a ser utilizado com reagentes de diagnóstico *in vitro* específicos registados, certificados ou aprovados e software de análise.

Princípios do procedimento

O NextSeq 550Dx Instrument da Illumina destina-se à sequenciação de bibliotecas de ADN com ensaios de diagnóstico *in vitro* e deve ser utilizado por pessoal de laboratório clínico qualificado e com a devida formação para trabalhar com procedimentos de diagnóstico *in vitro* realizados num laboratório clínico. Para a entrada, o NextSeq 550Dx utiliza as bibliotecas geradas do ADN em que os índices de amostras e as sequências de captura serão adicionados aos alvos amplificados. As bibliotecas de amostra são capturadas numa célula de fluxo e sequenciadas no instrumento, utilizando a química de sequenciação por síntese (sequencing by synthesis, SBS). A química SBS utiliza um método de terminador reversível para detetar bases únicas de nucleótidos com marcação fluorescente à medida que são incorporadas em cadeias de ADN crescentes. O Real-Time Analysis Software (RTA) executa uma análise da imagem e uma identificação de bases e atribui uma pontuação de qualidade a cada base de cada ciclo de sequenciação. Quando a análise primária terminar, a análise secundária pode ser executada no instrumento para processar identificações de base. O NextSeq 550Dx utiliza diferentes módulos de análise secundária consoante o fluxo de trabalho. Para o Germline Variant Module ou o Somatic Variant Module, o processamento inclui desmultiplexagem, geração de ficheiros FASTQ, alinhamento, identificação de variantes e geração de ficheiros em formato de identificação de variantes (VCF e gVCF). Os ficheiros VCF e gVCF contêm informações sobre as variantes encontradas em posições específicas num genoma de referência.

Configuração Dual Boot

O NextSeq 550Dx inclui uma configuração Dual Boot para permitir a utilização do instrumento no modo de diagnóstico (Dx) ou no modo RUO (research use only, apenas para investigação). Os ensaios de sequenciação de diagnóstico *in vitro*, incluindo o Germline Variant Module e o Somatic Variant Module, são executados no modo de diagnóstico. Apenas os reagentes de sequenciação de diagnóstico *in vitro* (IVD) podem ser utilizados no modo de diagnóstico. As características de desempenho e as limitações de procedimento do NextSeq 550Dx Instrument foram estabelecidas utilizando o Germline Variant Module e o Somatic Variant Module no modo de diagnóstico.

Limitações do procedimento

- 1 Para utilização em diagnóstico *in vitro*.
- 2 O Germline Variant Module e o Somatic Variant Module, quando utilizados com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) ou com o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), têm capacidade para:
 - ▶ Resultado de sequenciação ≥ 90 gigabases (Gb)

- ▶ Tamanho de leitura (em ensaios de extremidade emparelhada) 2 x 150 pares de bases (bp)
 - ▶ Bases iguais ou superiores a Q30 $\geq 75\%$ num tamanho de leitura de 2 x 150 bp
75% ou mais bases têm pontuações de qualidade na escala Phred ≥ 30 , indicando uma precisão da identificação de bases superior a 99,9%
- 3 Leituras com indels (inserções, eliminações ou combinações) cujo tamanho do conteúdo é > 25 bp não são alinhadas pelo Assay Software. Consequentemente, indels com tamanho > 25 bp não são detetáveis pelo Assay Software.
 - 4 O Assay Software poderá não alinhar leituras de fragmento amplificado com conteúdo extremo de variantes, resultando na definição da região a reportar como sendo do tipo selvagem. O conteúdo extremo inclui:
 - ▶ Leituras que contenham mais de três indels
 - ▶ Leituras com um tamanho de pelo menos 30 bp com conteúdo de variante de nucleótido único (SNV) $> 4\%$ do tamanho total do alvo de fragmento amplificado (excluindo regiões de investigação)
 - ▶ Leituras com tamanho < 30 bp com conteúdo SNV $> 10\%$ do tamanho total do fragmento amplificado (incluindo regiões de investigação)
 - 5 Variantes maiores, incluindo variantes de multinucleótidos (MNVs) e indels grandes, podem ser reportados como variantes separadas mais pequenas no ficheiro de saída VCF.
 - 6 É possível ignorar ou filtrar variantes de eliminação ao cruzar dois fragmentos amplificados sobrepostos se o tamanho da eliminação for superior ou igual à sobreposição entre os fragmentos amplificados sobrepostos.
 - 7 O sistema não consegue detetar indels se ocorrerem diretamente adjacentes a um primer e se não houver um fragmento amplificado de sobreposição. Para as regiões com fragmentos amplificados de sobreposição, o ensaio não consegue detetar eliminações quando a região de sobreposição é mais pequena do que o tamanho da eliminação a ser detetada. Por exemplo, se a região da sobreposição entre dois fragmentos amplificados adjacentes for duas bases, o ensaio não consegue detetar nenhuma eliminação incluindo ambas as bases.
A eliminação de uma base única em cada uma dessas bases pode ser detetada.
 - 8 Tal como com qualquer fluxo de trabalho de preparação de bibliotecas à base de hibridização, os polimorfismos, as mutações, as inserções ou as eliminações subjacentes em regiões de ligação oligonucleótida podem afetar os alelos investigados e as identificações obtidas durante a sequenciação. Por exemplo:
 - ▶ Uma variante em fase com uma variante na região do primer pode não ser amplificada, resultando num falso negativo.
 - ▶ As variantes na região do primer podem impedir a amplificação do alelo de referência, resultando numa identificação incorreta de variante homozigótica.
 - ▶ As variantes de indel na região do primer podem causar um falso positivo no fim da leitura adjacente ao primer.
 - 9 Os indels podem ser filtrados devido à tendência da cadeia se ocorrerem perto do fim de uma leitura e se forem levemente recortados durante o alinhamento.
 - 10 MNV pequenas podem não ter sido validadas e só são reportadas no Somatic Variant Module.
 - 11 As eliminações são reportadas no VCF na coordenada que antecede a base de acordo com o formato VCF. Assim, devem ser consideradas as variantes adjacentes antes de reportar que uma identificação individual de base é uma referência homozigótica.
 - 12 Limitações específicas do fluxo de trabalho da linha germinal:
 - ▶ O NextSeq 550Dx Instrument, utilizando o Germline Variant Module do Local Run Manager para o NextSeq 550Dx, foi concebido para fornecer resultados qualitativos para a identificação de variantes da linha germinal (p. ex., homozigóticas, heterozigóticas, de tipo selvagem).
 - ▶ Quando utilizado com o Germline Variant Module, a cobertura mínima necessária por fragmento amplificado para uma identificação de variantes precisa é de 150x. Como resultado, são necessários 150 fragmentos de ADN, que são equivalentes a 300 leituras de sobreposição de extremidade emparelhada. O número de amostras e o número total de bases alvo afetam a cobertura. O conteúdo GC (guanina-citosina) e outro conteúdo genómico podem afetar a cobertura.

- ▶ O número de cópia da variação pode afetar a identificação de uma variante como homozigótica ou heterozigótica.
 - ▶ As variantes de alguns contextos repetitivos são filtradas nos ficheiros VCF. O filtro de repetição RMxN é utilizado para filtrar variantes se parte da sequência ou toda a sequência da variante estiver presente repetidamente no genoma de referência adjacente à posição da variante. No caso da identificação de variantes de linha germinal, são necessárias pelo menos nove repetições na referência para que uma variante seja filtrada. Apenas as repetições com um tamanho até 5 bp são consideradas (R5x9).
 - ▶ Um indel e uma SNV num único local podem fazer com que apenas uma variante seja reportada.
- 13 Limitações específicas do fluxo de trabalho somático.
- ▶ O NextSeq 550Dx Instrument, utilizando o Local Run Manager Somatic Variant Module para o NextSeq 550Dx, foi concebido para fornecer resultados qualitativos para a identificação da variante somática (p. ex., presença de uma variante somática com uma frequência de variante igual ou superior a 0,026 com um limite de deteção de 0,05).
 - ▶ Quando utilizado com o Somatic Variant Module, a cobertura mínima necessária por fragmento amplificado para uma identificação de variantes precisa é de 450x por pool de oligonucleótidos. Como resultado, são necessários 450 fragmentos de ADN por pool de oligonucleótidos, que são equivalentes a 900 leituras de sobreposição de extremidade emparelhada. O número de amostras e o número total de bases alvo afetam a cobertura. O conteúdo GC (guanina-citosina) e outro conteúdo genómico podem afetar a cobertura.
 - ▶ Para identificação de variantes somáticas, são necessárias pelo menos seis repetições em referência para que uma variante seja filtrada e apenas as repetições com 3 bp são consideradas (R3x6).
 - ▶ O Somatic Variant Module não consegue distinguir entre variantes de linha germinal e somáticas. O módulo foi concebido para detetar variantes ao longo de um intervalo de frequências de variantes, mas a frequência de variantes não pode ser utilizada para distinguir entre variantes somáticas e variantes de linha germinal.
 - ▶ O tecido normal na amostra afeta a deteção de variantes. O limite de deteção reportado é baseado numa frequência de variantes relativa ao ADN total extraído de tecido tumoral e de tecido normal.

Componentes do produto

- 1 NextSeq 550Dx Instrument (Catálogo n.º 20005715)
- 2 Componentes do software do NextSeq 550Dx Instrument, incluindo o seguinte:

Aplicação de software	Função	Descrição
NextSeq 550Dx Operating Software (NOS)	Controla o funcionamento do instrumento	A aplicação do NOS Software gere o funcionamento do instrumento durante uma sequenciação e gera imagens para serem utilizadas pelo Real-Time Analysis Software (RTA).
Real-time Analysis Software (RTA)	Executa análise primária	A aplicação do RTA Software converte as imagens geradas pelo NOS de cada secção por ciclo de ensaio de sequenciação em ficheiros de identificação de bases, que são entradas para os módulos de análise do Local Run Manager. A aplicação do RTA Software não contém uma interface de utilizador.
Local Run Manager	Interface para seleção de módulos	O Local Run Manager Software é uma solução integrada no instrumento para gestão de utilizadores, que seleciona o módulo de análise apropriado e monitoriza o estado.
Somatic Variant Module	Executa análise secundária	O software do módulo de análise do Local Run Manager processa a identificação de bases através de uma análise secundária. O processamento inclui desmultiplexagem, geração de ficheiros FASTQ, alinhamento, identificação de variantes e a criação de relatórios. O identificador de variantes (Pisces) gera ficheiros VCF que contêm informações sobre variantes encontradas em posições específicas num genoma de referência e inclui a frequência de variantes medida.

Aplicação de software	Função	Descrição
Germline Variant Module	Executa análise secundária	O software do módulo de análise do Local Run Manager processa a identificação de bases através de uma análise secundária. O processamento inclui desmultiplexagem, geração de ficheiros FASTQ, alinhamento, identificação de variantes e a criação de relatórios. O identificador de variantes (Pisces) gera ficheiros VCF que contêm informações sobre variantes encontradas em posições específicas num genoma de referência e identifica cada variante como heterozigótica ou homozigótica.

Condições de funcionamento

Elemento	Especificação
Temperatura	Mantenha a temperatura do laboratório entre os 19 °C e os 25 °C (22 °C \pm 3 °C). Esta é a temperatura de funcionamento do instrumento. Durante uma execução, não permita que a temperatura ambiente varie mais do que \pm 2 °C.
Humidade	Mantenha uma humidade relativa sem condensação entre os 20 e os 80%.

Equipamento e materiais

Equipamento e materiais necessários, vendidos em separado

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

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

Equipamento e materiais necessários, não fornecidos

Consumíveis fornecidos pelo utilizador para ensaios de sequenciação

Consumível	Fornecedor	Finalidade
Toalhetas com álcool isopropílico a 70% ou Etanol a 70%	VWR, catálogo n.º 95041-714 (ou equivalente) Fornecedor geral do laboratório	Finalidade geral e limpeza da célula de fluxo
Pano de laboratório, libertação reduzida de pelo	VWR, catálogo n.º 21905-026 (ou equivalente)	Finalidade geral e limpeza da célula de fluxo

Consumíveis fornecidos pelo utilizador para a manutenção do instrumento

Consumível	Fornecedor	Finalidade
NaOCl, 5% (hipoclorito de sódio)	Sigma-Aldrich, catálogo n.º 239305 (ou equivalente de grau laboratorial)	Lavar o instrumento utilizando a lavagem manual pós-ensaio; diluído para 0,12%
Tween 20	Sigma-Aldrich, catálogo n.º P7949	Lavar o instrumento utilizando as opções de lavagem manual; diluído para 0,05%
Água, grau laboratorial	Fornecedor geral do laboratório	Lavar o instrumento (lavagem manual)
Filtro de ar	Illumina, catálogo n.º 20022240	Limpar o ar que o instrumento aspira para efeitos de arrefecimento

Diretrizes para água laboratorial

Utilize água laboratorial ou desionizada para realizar procedimentos no instrumento. Nunca utilize água da torneira. Utilize apenas água dos seguintes graus ou equivalente:

- ▶ Água desionizada
- ▶ Illumina PW1
- ▶ Água de 18 Megaohms (M Ω)
- ▶ Água Milli-Q
- ▶ Água Super-Q
- ▶ Água para biologia molecular

Avisos e precauções

ATENÇÃO A Lei Federal Americana só permite a venda deste dispositivo mediante receita médica ou por um médico ou outro profissional de saúde autorizado pela legislação do Estado onde este exerce, para utilizar ou prescrever este dispositivo.

- 1 **Alguns componentes dos reagentes fornecidos pela Illumina para serem utilizados com o NextSeq 550Dx Instrument contêm químicos potencialmente perigosos. Podem ocorrer lesões pessoais por inalação, ingestão, contacto da pele e contacto ocular. Use equipamento de proteção, incluindo proteção ocular, luvas e bata de laboratório adequados para o risco de exposição. Manuseie os reagentes usados como resíduos químicos e elimine-os de acordo com a legislação e os regulamentos locais, regionais e nacionais aplicáveis.** Para obter informações adicionais ambientais, de segurança e de saúde, consulte as Fichas de dados de segurança (FDS) em support.illumina.com/sds.html.
- 2 Comunique imediatamente quaisquer incidentes graves relacionados com este produto à Illumina e às autoridades competentes dos Estados-Membros nos quais o utilizador e o paciente estão estabelecidos.
- 3 Manuseie todas as amostras sanguíneas como se fossem conhecidas por estarem infetadas com o VIH (vírus da imunodeficiência humana), o vírus humano da hepatite B (VHB) e com outros agentes patogénicos transmitidos pelo sangue (precauções universais).
- 4 O não seguimento dos procedimentos da forma descrita pode resultar em resultados erróneos ou na redução significativa da qualidade das amostras.
- 5 Aplique as precauções de rotina do laboratório. Não coloque a pipeta na boca. Não coma, beba ou fume nas áreas designadas para trabalho. Use luvas descartáveis e batas de laboratório quando manusear amostras e kits de reagentes. Lave bem as mãos depois de manusear amostras e kits de reagentes.
- 6 É necessário que sejam implementadas as devidas práticas laboratoriais e as boas práticas de higiene laboratorial para evitar que os produtos PCR contaminem reagentes, instrumentos e amostras genómicas de ADN. A contaminação PCR pode causar resultados imprecisos e não fiáveis.
- 7 Para prevenir a contaminação, certifique-se de que as áreas de pré e pós-amplificação têm equipamento e consumíveis dedicados (p. ex., pipetas, pontas de pipeta, blocos de aquecimento, agitadores por vórtice e centrífugas).
- 8 O emparelhamento entre os índices e amostras tem de corresponder exatamente à impressão da disposição das placas. O Local Run Manager preenche automaticamente os primers de indexação associados aos nomes das amostras, quando introduzidas no módulo. Recomenda-se que o utilizador confirme se os primers de indexação são associados às amostras antes de iniciar o ensaio de sequenciação. As divergências entre a amostra e a disposição das placas resulta na perda da identificação positiva da amostra e em relatórios com resultados incorretos.
- 9 Recomenda-se vivamente ao utilizador a instalação de software antivírus para proteger o computador contra vírus. Consulte o manual do utilizador para obter instruções de instalação.
- 10 Não utilize o NextSeq 550Dx se algum dos painéis for removido. A utilização do instrumento com qualquer um dos painéis removidos cria uma potencial exposição à tensão de linha e a tensões de CC.

- 11 Não toque na plataforma da célula de fluxo no compartimento da mesma. O aquecedor neste compartimento funciona entre os 22 °C e os 95 °C e pode causar queimaduras.
- 12 O instrumento pesa cerca de 83 kg (185 lb) e pode causar ferimentos graves caso se deixe cair ou caso seja utilizado incorretamente.

Instruções de utilização

As seguintes instruções de utilização destinam-se à utilização do Germline Variant Module e do Somatic Variant Module no modo de diagnóstico do NextSeq 550Dx Instrument, utilizando o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) ou o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Introduzir informações do ensaio

Para obter informações detalhadas, consulte o Manual de referência do NextSeq 550Dx Instrument (documento n.º 100000009513) e o manual do módulo do Local Run Manager aplicável.

Definir parâmetros

- 1 Inicie sessão no Local Run Manager.
- 2 Selecione **Create Run** (Criar ensaio) e selecione **Somatic Variant** (Variante somática) ou **Germline Variant** (Variante da linha germinal).
- 3 Introduza um nome para o ensaio que o identifique a partir da sequenciação através da análise. Utilize caracteres alfanuméricos, espaços, sublinhados ou travessões.
- 4 **[Opcional]** Introduza a descrição do ensaio para ajudar a identificar o ensaio. Utilize caracteres alfanuméricos, espaços, sublinhados ou travessões.
- 5 Selecione o número de amostras e o conjunto de índices na lista pendente. Considere as seguintes informações quando selecionar uma opção.
 - ▶ A lista pendente contém números de amostras com um conjunto de índice. Por exemplo, 24-Set 1 indica que existem 24 amostras para testar com índices a partir do conjunto de índices 1.
 - ▶ Os números do conjunto de índices referem-se a conjuntos diferentes de pares de índices i5 e i7. O Set 1 (Conjunto 1) e o Set 2 (Conjunto 2) fornecem ambos uma diversidade de índices. São fornecidos dois conjuntos de índices para ajudar a evitar a diminuição de um único conjunto.
 - ▶ Escolha o número de amostras mais aproximado do número de amostras que está a testar. Se o número exato de amostras não estiver na lista, selecione o número mais próximo, mas inferior ao número que está a testar. Por exemplo, se pretende testar 18 amostras, selecione 16 amostras.
 - ▶ Os poços de amostras sugeridos e as combinações de índices que cumprem os requisitos de diversidade de índices estão realçados a verde.

Importar ficheiros de manifesto para o ensaio

- 1 Certifique-se de que os manifestos que pretende importar se encontram disponíveis numa localização de rede acessível ou numa unidade USB.
- 2 Selecione **Import Manifests** (Importar manifestos).
- 3 Aceda ao ficheiro de manifesto e selecione os manifestos que pretende adicionar.

NOTA Para tornar os ficheiros de manifesto disponíveis para todos os ensaios utilizando o módulo de análise Germline Variant ou Somatic Variant, adicione os manifestos com a funcionalidade Module Settings (Definições do módulo). Esta funcionalidade requer uma permissão de nível de utilizador administrador. Para obter mais informações, consulte o *Manual de referência do instrumento NextSeq 550Dx (documento n.º 100000009513)*.

Especificar amostras para o ensaio

Especifique amostras para o ensaio utilizando uma das opções e instruções que se seguem.

- ▶ **Introduza as amostras manualmente**— Utilize a tabela em branco no ecrã Create Run (Criar ensaio).
- ▶ **Importar amostras**— Aceda a um ficheiro externo em formato de valores separados por vírgulas (*.csv). Está disponível um modelo para transferência no ecrã Create Run (Criar ensaio).

Introduzir amostras manualmente

- 1 Introduza um nome único para a amostra (**módulo de análise da Somatic Variant**) ou ID de amostra (**módulo de análise da Germline Variant**).
Utilize caracteres alfanuméricos, travessões ou sublinhados.
- 2 **[Opcional]** No caso de amostras de controlo positivo ou negativo, clique com o botão direito e selecione o tipo de controlo.
O controlo de um poço de amostra preenche automaticamente o poço no outro pool com o mesmo controlo.
- 3 **[Opcional]** Introduza uma descrição da amostra no campo Sample Description (Descrição da amostra).
Utilize caracteres alfanuméricos, travessões ou sublinhados.
- 4 Selecione um adaptador de Índice 1 na lista pendente do Índice 1 (i7).
Quando utilizar os poços de amostra sugeridos, o software preenche automaticamente os adaptadores de índice i7 e i5 que cumprem os requisitos de diversidade de índices. Se o número de amostras exato que está a testar não estiver na lista, certifique-se de que seleciona os adaptadores de índice para poços extra.
- 5 Selecione um adaptador de Índice 2 na lista pendente do Índice 2 (i5).
- 6 Selecione um ficheiro de manifesto na lista pendente do manifesto.
As amostras do Pool A requerem um manifesto diferente do das amostras do Pool B.
- 7 Selecione uma opção para ver, imprimir ou guardar a disposição das placas, como referência para preparar bibliotecas:
 - Selecione o ícone  **Print** (Imprimir) para ver a disposição das placas. Selecione **Print** (Imprimir) para imprimir a disposição das placas.
 - Selecione **Export** (Exportar) para exportar as informações da amostra para um ficheiro externo.
- 8 Selecione **Save Run** (Guardar ensaio).

Importar amostras

- 1 Selecione **Import Samples** (Importar amostras) e aceda à localização do ficheiro de informações da amostra. Existem dois tipos de ficheiros que pode importar.
 - Selecione **Template** (Modelo) no ecrã Create Run (Criar ensaio) para criar uma nova disposição de placas. O ficheiro modelo contém os cabeçalhos de coluna corretos para importação. Introduza informações em cada coluna das amostras do ensaio. Elimine as informações de exemplo das células não utilizadas e, em seguida, guarde o ficheiro.
 - Utilize um ficheiro com informações da amostra que tenha sido exportado do módulo Germline Variant ou Somatic Variant através da função Export (Exportar).
- 2 Selecione o ícone  **Print** (Imprimir) para ver a disposição das placas.
- 3 Selecione **Print** (Imprimir) para imprimir a disposição das placas como referência para preparar bibliotecas.
- 4 Selecione **Save Run** (Guardar ensaio).

Preparar o cartucho de reagentes

Certifique-se de que segue as instruções do cartucho de reagentes cuidadosamente para uma sequenciação bem-sucedida.

- 1 Retire o cartucho de reagentes do armazenamento de -25 °C a -15 °C.
- 2 Selecione um dos seguintes métodos para descongelar os reagentes. Não mergulhe o cartucho. Depois de o cartucho estar descongelado, seque-o antes de avançar para o passo seguinte.

Temperatura	Tempo de descongelamento	Limite de estabilidade
Banho com água de 15 °C a 30 °C	60 minutos	Não deve ultrapassar 6 horas
2 °C a 8 °C	7 horas	Não deve ultrapassar 5 dias

NOTA Se estiver a descongelar mais de um cartucho no mesmo banho de água, o tempo de descongelamento será maior.

- 3 Inverta o cartucho cinco vezes para misturar os reagentes.
- 4 Inspeccione a parte de baixo do cartucho para verificar que os reagentes estão descongelados e que não contêm precipitados. Confirme que as posições 29, 30, 31 e 32 estão descongeladas, pois são as maiores e demoram mais tempo para descongelar.
- 5 Bata levemente na bancada para reduzir as bolhas de ar.
Para obter os melhores resultados, avance diretamente para o carregamento de amostras e configuração do ensaio.

Preparar a célula de fluxo

- 1 Retire uma nova caixa de célula de fluxo do armazenamento de 2 °C a 8 °C.
- 2 Remova a embalagem de alumínio da caixa e mantenha à temperatura ambiente durante 30 minutos.

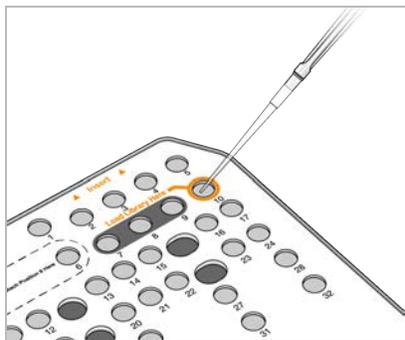
Preparar bibliotecas para sequenciação

Proceda à desnaturação e diluição das suas bibliotecas até um volume de carregamento de 1,3 ml. Na prática, a concentração de carregamento pode variar dependendo da preparação da biblioteca e dos métodos de quantificação. A diluição das bibliotecas de amostras depende da complexidade dos pools de oligonucleótidos. Para obter instruções para preparar bibliotecas de amostras para sequenciação, incluindo pooling e diluição de bibliotecas, consulte a secção das Instruções de utilização do kit de preparação de bibliotecas aplicável. É necessária a otimização da densidade do cluster no NextSeq 550Dx.

Carregar bibliotecas para o cartucho de reagentes

- 1 Limpe o selo de alumínio que cobre o reservatório n.º 10 com a etiqueta **Load Library Here** (Carregar a biblioteca aqui) utilizando um pano com libertação reduzida de pelo.
- 2 Perfure o selo com a ponta de uma pipeta limpa de 1 ml.
- 3 Carregue 1,3 ml de bibliotecas preparadas no reservatório n.º 10 com a etiqueta **Load Library Here** (Carregar biblioteca aqui). Evite tocar no selo de alumínio enquanto distribui as bibliotecas.

Figura 1 Carregar bibliotecas



Configurar um ensaio de sequenciação

- 1 Inicie sessão no NextSeq 550Dx com a palavra-passe do Local Run Manager Software.
- 2 No ecrã Home (Início) do NOS Software, selecione **Sequence** (Sequência).
- 3 Selecione um ensaio na lista e, em seguida, selecione **Next** (Seguinte).
São abertos vários ecrãs de configuração de ensaios pela seguinte ordem: Load Flow Cell (Carregar célula de fluxo), Load Buffer Cartridge (Carregar cartucho de tampão), Load Reagent Cartridge (Carregar cartucho de reagentes) e Pre-run Check (Verificação pré-ensaio).

- 4 Quando o ecrã Load Flow Cell (Carregar célula de fluxo) for apresentado, limpe e carregue a célula de fluxo.
 - ▶ Remova a célula de fluxo da embalagem de alumínio.
 - ▶ Abra a embalagem de plástico transparente e remova a célula de fluxo
 - ▶ Limpe a superfície de vidro da célula de fluxo com um pano sem pelo e com álcool. Seque o vidro com um pano com libertação reduzida de pelo
 - ▶ Certifique-se de que a superfície de vidro da célula de fluxo está limpa. Se necessário, repita o passo de limpeza.
 - ▶ Remova a célula de fluxo usada num ensaio anterior.
 - ▶ Alinhe a célula de fluxo com os pinos de alinhamento e coloque a célula de fluxo na plataforma.
- 5 Selecione **Load** (Carregar).

A porta fecha automaticamente, o ID da célula de fluxo é apresentado no ecrã e os sensores são verificados.
- 6 Siga as instruções do software para esvaziar o recipiente de reagentes gastos, carregar o cartucho de tampão NextSeq 550Dx e carregar o cartucho de reagentes NextSeq 550Dx.

Quando os cartuchos de tampão e de reagentes do NextSeq 550Dx são carregados, o software lê e regista o RFID. Os IDs dos cartuchos de tampão e de reagentes são apresentados no ecrã e os sensores são verificados.
- 7 Quando a verificação de pré-ensaio automática estiver concluída, selecione **Start** (Iniciar). (Não requerido se estiver configurado para iniciar automaticamente.)
- 8 O ecrã Sequencing (Sequenciação) é apresentado quando o ensaio começar. Este ecrã fornece uma representação visual do ensaio em curso, incluindo as intensidades e as pontuações de qualidade (Pontuações Q).

Resultados

O RTA (Real-Time Analysis) é um software integrado que executa uma análise da imagem e uma identificação de bases e atribui uma pontuação de qualidade a cada base de cada ciclo de sequenciação. Quando a análise primária terminar, o módulo Local Run Manager selecionado no NextSeq 550Dx Instrument inicia a análise secundária automaticamente. Os processos da análise secundária aqui descritos destinam-se ao Germline Variant Module e ao Somatic Variant Module.

Desmultiplexagem

A desmultiplexagem compara cada sequência de leitura de índice com as sequências de índice especificadas para o ensaio. Não são considerados quaisquer valores de qualidade neste passo.

As leituras de índice são identificadas através dos seguintes passos:

- ▶ As amostras são numeradas a partir do número 1, na ordem pela qual são listados para o ensaio.
- ▶ O número de amostra 0 está reservado aos clusters que não foram atribuídos a uma amostra.
- ▶ Os clusters são atribuídos a uma amostra quando a sequência de índice corresponde exatamente ou quando existe pelo menos uma única divergência por leitura de índice.

Geração de ficheiros FASTQ

Após a desmultiplexagem, o software gera ficheiros de análise intermédia em formato FASTQ, que é um formato de texto usado para representar sequências. Os ficheiros FASTQ contêm leituras de cada amostra e as pontuações de qualidade associadas. Os clusters que não passarem no filtro serão excluídos.

Cada ficheiro FASTQ contêm leituras de apenas uma amostra e o nome dessa amostra está incluído no nome de ficheiro FASTQ. No Germline Variant Module (módulo de variante de linha germinal) e no Somatic Variant Module (módulo de variante somática), são gerados oito ficheiros FASTQ por amostra e por pool de

oligonucleótidos, quatro da Leitura 1 e quatro da Leitura 2, o que resulta num total de 8 e 16 ficheiros FASTQ por amostra para Germline e Somatic, respetivamente. Os ficheiros FASTQ são a principal entrada para alinhamento.

Alinhamento

Durante o passo de alinhamento, o algoritmo de Smith-Waterman com faixas alinha os clusters de cada amostra com as sequências de fragmentos amplificadas especificadas no ficheiro de manifesto.

O algoritmo de Smith-Waterman com faixas executa alinhamentos de sequências semiglobais para determinar regiões semelhantes entre duas sequências. Em vez de comparar a totalidade da sequência, o algoritmo de Smith-Waterman compara segmentos de todos os tamanhos possíveis.

Cada leitura de extremidade emparelhada é avaliada em termos de alinhamento em relação às sequências de investigação relevantes dessa leitura.

- ▶ A Leitura 1 é avaliada em relação ao complemento inverso dos DLSO (Downstream Locus-Specific Oligos, oligonucleótidos a jusante específicos a um locus).
- ▶ A Leitura 2 é avaliada em relação aos ULSO (Upstream Locus-Specific Oligos, oligonucleótidos a montante específicos a um locus).
- ▶ Se o início de uma leitura corresponder a uma sequência de investigação com até uma divergência, o tamanho total da leitura é alinhado em relação ao alvo amplicon dessa sequência.
- ▶ Se o início de uma leitura corresponder a uma sequência de investigação com até três diferenças (divergências ou mudanças devido aos primeiros indels), o tamanho total da leitura é alinhado em relação ao fragmento amplificado alvo dessa sequência.
- ▶ Os indels dos DLSO e dos ULSO não são observados devido à química do ensaio.

Os alinhamentos são filtrados a partir dos resultados de alinhamento com base nas taxas de divergência na região de interesse ou de todo o fragmento amplificado, consoante o tamanho do mesmo. Os alinhamentos filtrados são gravados em ficheiros de alinhamento como não alinhados e não são utilizados na identificação de variantes.

Identificação de variantes

O identificador de variantes Pisces foi concebido para fazer identificações de variantes SNV (single nucleotide variant, variante de nucleótido único) e de indel de bibliotecas preparadas para o instrumento.

Relatórios e ficheiros de saída adicionais

Os módulos de análise de variantes produzem relatórios em PDF e relatórios delimitados por tabulações (*.txt) que apresentam métricas, tais como a profundidade de sequenciação e as contagens de variantes. Os módulos também produzem ficheiros de saída VCF e gVCF (genome Variant Call Format, formato de identificação de variantes genómicas) para aplicações de identificação de variantes.

Procedimentos de controlo de qualidade

O NextSeq 550Dx Software avalia cada ensaio, amostra e identificação de bases em relação às métricas de controlo de qualidade. Os controlos positivo e negativo também são recomendados na preparação de bibliotecas e devem ser avaliados. Avalie os controlos da seguinte forma:

- **Controlo negativo (controlo sem modelo) ou outro controlo negativo**— Tem de gerar o resultado esperado. Se o controlo negativo gerar um resultado diferente do esperado, poderá ter ocorrido um erro no controlo de amostras, um registo incorreto dos primers de indexação ou contaminação.
- **Amostra de controlo positivo**— Tem de gerar o resultado esperado. Se o controlo positivo gerar um resultado diferente do esperado, poderá ter ocorrido um erro no controlo de amostras ou um registo incorreto dos primers de indexação.

Caraterísticas de desempenho

As caraterísticas de desempenho do NextSeq 550Dx Instrument foram estabelecidas com o Germline Variant Module e o Somatic Variant Module com o TruSeq Custom Amplicon Kit Dx e o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) e confirmadas com o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Os estudos incluíram indexação de amostras, contaminação cruzada de amostras, entrada de ADN, sensibilidade analítica (limite de vazio/limite de deteção), precisão, exatidão, comparação de métodos e reprodutibilidade.

Os estudos analíticos com o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) foram concebidos para avaliar as alegações de desempenho anteriormente estabelecidas com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Os resultados demonstram que os kits de reagentes (v2 e v2.5) têm um desempenho comparável utilizando o TruSeq Custom Amplicon Kit Dx. Consulte o *folheto informativo do TruSeq Custom Amplicon Kit Dx* para ver as caraterísticas de desempenho relacionadas com fatores pré-clínicos, tais como métodos de extração ou substâncias interferentes.

Definições dos cálculos utilizados nas caraterísticas de desempenho

- 1 A PPA (Positive Percent Agreement, percentagem de concordância positiva) é calculada como a proporção, reportada corretamente pelo ensaio, de loci classificados como variantes por um método de referência.
 - ▶ $(\text{N.º de loci variantes corretamente reportados pelo ensaio}) / (\text{n.º total de loci variantes})$
Os loci variantes reportados pelo ensaio que são concordantes com o método de referência são verdadeiros positivos (true positives, TPs). Os loci variantes reportados como identificações de referência ou como identificações de variantes diferentes pelo ensaio são falsos negativos (false negatives, FNs).
- 2 A NPA (Negative Percent Agreement, percentagem de concordância negativa) é calculada como a proporção, reportada corretamente pelo ensaio, de loci classificados como sendo de tipo selvagem por um método de referência.
 - ▶ $(\text{N.º de loci de tipo selvagem corretamente reportados pelo ensaio}) / (\text{n.º total de loci de tipo selvagem})$
Os loci de tipo selvagem reportados pelo ensaio que são concordantes com o método de referência são verdadeiros negativos (TNs). Os loci de tipo selvagem reportados como variantes pelo ensaio são falsos positivos (FPs).
- 3 A OPA (Overall Percent Agreement, percentagem de concordância geral) é calculada como a proporção, corretamente reportada pelo ensaio, de loci relativamente a um método de referência.
 - ▶ $([\text{N.º de loci variantes corretamente reportados pelo ensaio}] + [\text{n.º de loci de tipo selvagem corretamente reportados pelo ensaio}]) / ([\text{n.º total de loci variantes}] + [\text{n.º total de loci de tipo selvagem}])$
- 4 Os cálculos de PPA, NPA e OPA não incluem quaisquer identificações (loci variante ou de referência que não cumpram um ou mais filtros de qualidade).
- 5 A taxa de identificação autossômica é calculada como o número total de loci que cumpre os filtros dividido pelo número total de posições sequenciadas para os cromossomas 1–22. Os cromossomas X e Y estão excluídos. Esta métrica não considera a concordância de identificações com o método de referência.

Desempenho do NextSeq 550Dx High Output Reagent Kit v2 (300 Cycles)

Indexação de amostras

Os primers de indexação de amostras, adicionados durante a preparação de bibliotecas, atribuem uma sequência única a cada amostra de ADN. Estas sequências únicas permitem analisar múltiplas amostras em pool num único ensaio de sequenciação. A indexação de amostras é utilizada em fluxos de trabalho somáticos e de linha germinal. A finalidade deste estudo é estabelecer os números mínimo (8) e máximo (96) de amostras que podem ser processadas num único ensaio de sequenciação pelo NextSeq 550Dx Instrument. Foram testadas oito amostras únicas de Platinum Genome com 12 combinações diferentes do primer de indexação por amostra. Os resultados das amostras de quatro ensaios de sequenciação utilizando o Germline Variant Module foram comparados com o Platinum Genome versão 2016-1.0.

No primeiro conjunto de ensaios, foram testadas 96 bibliotecas de amostras exclusivamente indexadas com um ensaio representativo concebido para analisar uma variedade de genes abrangendo 12 588 bases por cadeia ao longo dos 23 cromossomas humanos, para verificar a capacidade do ensaio de fazer uma identificação de genotipagem consistentemente numa determinada amostra, em combinações diferentes do primer de indexação. No segundo conjunto de ensaios, foram sequenciadas oito bibliotecas de amostras exclusivamente indexadas em dois ensaios de sequenciação, para verificar o número mínimo de índices suportado.

Nos ensaios de 96 índices, a PPA para SNVs variou entre os 98,7% e os 100%, a PPA para inserções e eliminações foi de 100% e a NPA foi de 100% para cada uma das combinações de 96 índices. Os ensaios de 8 índices tinham valores de PPA de 100% (SNVs, inserções e eliminações) e NPA de 100% para cada uma das combinações de oito índices.

Contaminação cruzada de amostras

O NextSeq 550Dx Instrument permite a sequenciação de amostras múltiplas e controlos num único ensaio de sequenciação. Foi realizado um estudo para avaliar a dimensão da contaminação cruzada de amostras num ensaio de sequenciação (no mesmo ensaio) e entre ensaios de sequenciação (de ensaio para ensaio). Foram analisadas duas amostras Platinum Genome, uma masculina e outra feminina, com um ensaio representativo concebido para examinar vários genes abrangendo 12 588 bases (150 fragmentos amplificados) ao longo de 23 cromossomas diferentes, incluindo ambos os cromossomas sexuais. As bibliotecas foram sequenciadas no NextSeq 550Dx Instrument utilizando o Germline Variant Module. Foi observada a contaminação cruzada de amostras masculinas para amostras femininas através da presença de leituras de fragmento amplificado do cromossoma Y em amostras femininas.

A contaminação cruzada no mesmo ensaio pode ocorrer durante a geração de clusters, identificação de bases do ciclo de índices e durante a desmultiplexagem de amostras. Para testar a contaminação cruzada de amostras num ensaio de sequenciação, foi sequenciado um pool de bibliotecas compostas por 46 réplicas cada de amostras masculinas e femininas, mais quatro controlos sem modelo, uma vez no NextSeq 550Dx Instrument. Foi avaliada a contaminação cruzada de amostras no mesmo ensaio comparando a cobertura de fragmentos amplificados do cromossoma Y de cada réplica feminina com a média da cobertura de fragmentos amplificados do cromossoma Y de todas as réplicas masculinas no pool. A mediana observada de contaminação cruzada no mesmo ensaio foi de 0,084%.

Para testar a contaminação cruzada de amostras de ensaio para ensaio, foram preparados e sequenciados consecutivamente dois pools de bibliotecas no NextSeq 550Dx Instrument. O primeiro pool continha 46 réplicas de amostras femininas e dois controlos sem modelo. O segundo pool continha 46 réplicas de amostras masculinas e dois controlos sem modelo. Ambos os pools utilizaram o mesmo conjunto de adaptadores de índice. O pool feminino foi sequenciado primeiro, seguido por um ensaio de sequenciação subsequente com o pool masculino, seguido por outra repetição do ensaio de sequenciação do pool feminino. Foi avaliada a contaminação cruzada de amostras de ensaio para ensaio através da comparação da cobertura de fragmentos amplificados do cromossoma Y entre as réplicas correspondentes do ensaio de repetição do pool feminino e do ensaio do pool masculino. A mediana observada de contaminação cruzada de ensaio para ensaio foi de 0,0076%.

Entrada de ADN

Sangue (linha germinal)

O intervalo de entrada de ADN sanguíneo para a preparação de bibliotecas do TruSeq Custom Amplicon Kit Dx utilizando o fluxo de trabalho do Germline Variant Module foi estabelecido para o NextSeq 550Dx Instrument. Este intervalo foi avaliado executando um estudo de diluição de série com 13 amostras Platinum Genome com um ensaio representativo concebido para examinar diversos genes, abrangendo 12 588 bases nos 23 cromossomas diferentes. A biblioteca foi sequenciada em dois instrumentos NextSeq 550Dx utilizando um lote do NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Foram testadas cinco amostras em duplicado em cinco níveis de entrada de ADN, entre 250 ng e 12 ng (250 ng, 100 ng, 50 ng, 25 ng e 12 ng). Foram testadas oito amostras como única réplica em cada um dos cinco níveis de entrada de ADN. Para determinação da precisão, os genótipos das amostras foram comparados ao Platinum

Genome, versão 2016-1.0. Os resultados foram determinados para cada nível de entrada. A PPA de cada tipo de variante (SNVs, inserções e eliminações) é apresentada na [Tabela 1](#); a NPA é apresentada na [Tabela 2](#). Todos os níveis de entrada tinham uma precisão semelhante. A entrada de ADN recomendada para o TruSeq Custom Amplicon Kit Dx é de 50 ng com 25 ng e 100 ng desde que exista um limite inferior e superior para cumprir as características de desempenho.

Tabela 1 Resultados da PPA para cada entrada de ADN por tipo de variante

Entrada de ADN (ng)	Tipo de variante	Variantes esperadas	TP	FN	Não identificação de variantes	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	Inserção	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	Eliminação	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

Tabela 2 NPA para cada entrada de ADN

Entrada de ADN (ng)	TN	FP	Não identificação de referências	NPA (%)
12	430940	4	26	> 99,9
25	430936	0	34	100
50	430936	2	32	> 99,9
100	430942	0	28	100
250	430942	0	28	100

FFPE (somático)

O intervalo de entrada de ADN fixado em formalina e incorporado em parafina (FFPE, formalin-fixed paraffin-embedded) para a preparação de bibliotecas do TruSeq Custom Amplicon Kit Dx utilizando o fluxo de trabalho do Somatic Variant Module foi estabelecido para o NextSeq 550Dx Instrument. O intervalo de entrada de ADN foi avaliado executando um estudo de diluição de série com três amostras Platinum Genome com um ensaio representativo concebido para examinar diversos genes, abrangendo 12 588 bases nos 23 cromossomas diferentes. As linhas de células GM12878 e GM12877 do Platinum Genome foram fixadas em formalina e incorporadas em parafina, seguindo-se a extração de ADN. A GM12878 foi diluída com a GM12877 de forma a que as frequências de alelos de variantes (VAFs) de 81 variantes (55 SNVs, 10 inserções e 16 eliminações) ficassem perto de 0,025, 0,05 ou 0,10. Além disso, cada amostra tinha 91 variantes com frequências de alelos de variantes mais elevadas de até 1,0 VAF. As amostras foram processadas em duplicado em cinco níveis de entrada de ADN com um ciclo quantitativo delta médio (dCq) de 2,1, 3,6, 4,6, 6,0 e 7,8 conforme medido pelo TruSeq Custom Amplicon Dx - FFPE QC Kit. Cada biblioteca foi sequenciada em dois instrumentos NextSeq 550Dx utilizando dois lotes de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Para determinação da precisão, foram comparadas as identificações de variantes de amostras com o Platinum Genome, versão 2016-1.0. A PPA de cada tipo de variante (SNVs, inserções e eliminações) é apresentada na [Tabela 3](#); a NPA é apresentada na [Tabela 4](#). A entrada de ADN recomendada para variantes de 0,05 VAF ou mais é de dCq \leq 4 com 4,6, desde que exista um limite inferior para cumprir as características de desempenho.

Tabela 3 Resultados da PPA para cada entrada de ADN por tipo de variante

dCq médio	Tipo de variante	Variantes esperadas	Não identificações esperadas	VAF de diluição alvo					
				0,025		0,05		0,10	
				Não identificação de variantes	PPA (%)	Não identificação de variantes	PPA (%)	Não identificação de variantes	PPA (%)
2,1	SNV	808	Não aplicável.	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	Inserção	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	Eliminação	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

Tabela 4 NPA para cada entrada de ADN

dCq médio	Tipo selvagem esperado	VAF de diluição alvo					
		0,025		0,05		0,10	
		Não identificação de referências	NPA (%)	Não identificação de referências	NPA (%)	Não identificação de referências	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

Sensibilidade analítica (limite de vazão [LoB] e limite de detecção [LoD])

Este estudo foi realizado para avaliar o limite de vazão (LoB) e o limite de detecção (LoD) do Somatic Variant Module no NextSeq 550Dx Instrument. Foi realizado utilizando um ensaio representativo concebido para examinar diversos genes abrangendo 12 588 bases em 23 cromossomas diferentes. As linhas de células GM12878 e GM12877 do Platinum Genome foram fixadas em formalina e incorporadas em parafina, seguindo-se a extração de ADN. GM12878 foi diluída com a GM12877 de forma a que as frequências de 74 variantes (53 SNVs, 7 inserções e 14 eliminações) fossem de $0,05 \pm 0,02$. A GM12878 (GM12878-D) diluída e a GM12877 foram analisadas em seis dias de início consecutivos com um único instrumento, alternando entre dois lotes de NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), num total de seis ensaios de sequenciação. Este teste resultou em 60 réplicas de cada variante na GM12878-D e em 72 réplicas de cada coordenada de tipo selvagem correspondente na GM12877 para cada lote de reagentes. O LoB e o LoD foram calculados através da abordagem clássica indicada em CLSI EP17-A2 utilizando a opção não paramétrica. O LoB e o LoD foram calculados para SNVs, inserções e eliminações em separado através do pool das frequências de um determinado tipo de variante. O erro de Tipo I foi definido como 0,01 e o erro de Tipo II foi definido como 0,05.

No caso do LoB, o pool de frequências de variantes foi organizado do mais baixo para o mais alto e foi calculada a posição 99 da classificação de cada lote de reagentes para cada tipo de variante (Tabela 5). O Somatic Variant Module utiliza um limite (o LoB efetivo) de 0,026 VAF para determinar a detecção qualitativa de variantes. O LoB calculado verificou que este limite resulta num erro de Tipo I não superior a 0,01.

Tabela 5 Limite de vazão

Tipo de variante	Observações totais	Lote de reagentes LoB 1 (%)	Lote de reagentes LoB 2 (%)
SNV	3816	0,77	0,77
Inserção	504	0,56	0,56
Eliminação	1008	1,20	1,20

No caso do LoD, foi calculada a percentagem da frequência de mutação individual de cada lote de reagentes por cada tipo de variante abaixo do limite de 0,026 (Tabela 6). Como as percentagens eram inferiores ao erro de Tipo II de 5% (0,05), a mediana das frequências de variantes combinadas foi calculada como LoD (Tabela 6). O LoD de cada tipo de variante foi considerado como o maior de dois valores calculados para os dois lotes de reagentes – 4,97% para SNVs, 5,12% para inserções e 5,26% para eliminações.

Tabela 6 Limite de detecção

Lote de reagentes	Tipo de variante	Observações totais	N.º de medições VAF < 2,6%	% de medições VAF < 2,6%	Limite de detecção (%)
1	SNV	3180	53	1,7	4,94
	Inserção	420	6	1,4	5,08
	Eliminação	840	7	0,8	5,22
2	SNV	3180	51	1,6	4,97
	Inserção	420	5	1,2	5,12
	Eliminação	840	7	0,80	5,26

Precisão

Linha germinal

O seguinte estudo foi realizado para avaliar a precisão da identificação de variantes do Germline Variant Module no NextSeq 550Dx Instrument utilizando o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Foram testadas treze amostras únicas do Platinum Genome utilizando um ensaio representativo concebido para examinar vários genes abrangendo 12 588 bases (150 fragmentos amplificados) em 23 cromossomas diferentes. Foram realizados nove ensaios no total utilizando três instrumentos de sequenciação, três lotes de reagentes e três operadores em cinco dias de início. Foi determinada a precisão para SNVs, inserções e eliminações comparando os resultados a um método de referência composto e bem caracterizado, Platinum Genome versão 2016-1.0. As regiões de confiança genómica foram definidas com base neste método de referência, exceto se especificado de outro modo.

Tabela 7 Resumo de concordância da linha germinal

Crítérios	Observações totais ¹	Resultado por observação ²	Resultado por ensaio ³
PPA para SNV	819	98,7	> 99,9
PPA para inserções	819	95,0	98,9
PPA para eliminações	819	100	100
NPA	819	100	100
OPA	819	> 99,9	> 99,9

¹Calculado como número de amostras por ensaio (91) x número de ensaios (9) = 819.

²Valor mais baixo observado por réplica de amostra nos 9 ensaios.

³Valor mais baixo quando os dados de cada ensaio são analisados em agregado.

A **Tabela 8** contém os dados do estudo apresentados com uma percentagem de concordância positiva e negativa por amostra, em que os resultados da variante são comparados ao Platinum Genome versão 2016-1.0 em cálculos de PPA. Os três tipos de variantes (SNVs, inserções e eliminações) são combinados. Uma vez que o método de referência só fornece resultados para as variantes de nucleótido único e inserções/eliminações, os resultados de base não variante são comparados ao conjunto hg19 de sequenciação de referência do genoma humano para cálculos de NPA.

Tabela 8 Concordância da linha germinal por amostra

Amostra	Taxa média de identificações	Variante esperadas ¹	TP	FN	Não identificação de variantes	TN	FP	PPA	NPA	OPA
NA12877	> 99,9	4788	4788	0	0	756762	0	100	100	100
NA12878	> 99,9	8505	8379	1	125	751464	0	> 99,9	100	> 99,9

Amostra	Taxa média de identificações	Variantes esperadas ¹	TP	FN	Não identificação de variantes	TN	FP	PPA	NPA	OPA
NA12879	> 99,9	6048	5985	5	58	757701	0	99,9	100	> 99,9
NA12880	> 99,9	6993	6930	0	63	757638	0	100	100	100
NA12881	> 99,9	7875	7811	3	61	751653	0	> 99,9	100	> 99,9
NA12882	> 99,9	6300	6174	3	123	754803	0	> 99,9	100	> 99,9
NA12883	> 99,9	7119	7056	0	63	751905	0	100	100	100
NA12884	> 99,9	7182	7119	6	57	754146	0	99,9	100	> 99,9
NA12885	> 99,9	7686	7560	2	124	754173	0	> 99,9	100	> 99,9
NA12886	> 99,9	7245	7182	7	56	752469	0	99,9	100	> 99,9
NA12887	> 99,9	7119	7119	0	0	750645	0	100	100	100
NA12888	> 99,9	6804	6804	0	0	756065	0	100	100	100
NA12893	> 99,9	7434	7371	1	62	750015	0	> 99,9	100	> 99,9

¹ O número total de variantes em todas as réplicas de amostra nos 9 ensaios.

A **Tabela 9** contém os dados do estudo apresentados por amostra, em que os resultados das variantes são comparados com o método de referência composto e bem caracterizado. A deteção é avaliada para cada tipo de variante – SNVs, inserções e eliminações – separadamente. As posições de referência são excluídas.

Tabela 9 Concorrência da linha germinal por amostra e por tipo de variante

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

As amostras foram analisadas novamente para identificar pequenas inserções e eliminações (indels). É apresentado um resumo geral na **Tabela 10**. Foram identificados 71 indels no total com tamanho entre 1–24 bp no caso das inserções, e 1–25 bp no caso das eliminações.

Tabela 10 Resumo da deteção de indels da linha germinal

Tipo de variante	Variantes esperadas	TP	FN	Não identificação de variantes	PPA
Inserção	18522	18018	27	477	99,9
Eliminação	17388	17073	0	315	100

O ensaio representativo era composto por 150 fragmentos amplificados concebidos para abranger o diverso conteúdo genómico. O conteúdo GC dos fragmentos amplificados variou entre 0,19–0,87. Os fragmentos amplificados também tinham uma gama de nucleótido único (p. ex. PolyA, PolyT), repetições de dinucleótidos e trinucleótidos. Os dados foram compilados por fragmento amplificado (Tabela 11) para determinar o efeito do conteúdo genómico na percentagem de identificações corretas. A percentagem de identificações corretas consiste em identificações de referência e variantes e é inferior a 100% se ocorrerem identificações incorretas ou não identificações.

Tabela 11 Precisão ao nível do fragmento amplificado da linha germinal

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
1	1	36450499	36450591	93	93	Indel	0,22	76167	0	0	100
2	1	109465122	109465200	79	79	Poly A (5), Poly C (5), indel	0,38	64701	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	74529	0	0	100
4	1	223906657	223906748	92	92	Indel	0,49	75348	0	0	100
5	1	228526602	228526682	81	81	Poly G (5)	0,69	66339	0	0	100
6	1	236372039	236372108	70	70	Poly T (10), indel	0,39	57330	0	0	100
7	1	247812041	247812128	88	88	Poly A (5), CT(3), TAA(3), indel	0,27	72072	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	73710	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	65520	0	0	100
10	2	177016721	177016805	85	81	N/D	0,65	66339	0	0	100
11	2	186625727	186625801	75	75	Poly A (8)	0,35	61425	0	0	100
12	2	190323504	190323591	88	88	Poly T (5)	0,42	72072	0	0	100
13	2	200796740	200796826	87	87	Poly T (5), indel	0,31	71253	0	0	100
14	2	212245049	212245139	91	91	Poly T (5), Poly A (6), indel	0,3	74529	0	0	100
15	2	228147052	228147144	93	93	Indel	0,43	76167	0	0	100
16	2	235016350	235016422	73	73	Poly T (5), indel	0,42	59787	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	74823	0	1344	98,2
18	3	46620561	46620643	83	83	N/D	0,43	67977	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	57330	0	0	100

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
20	3	189713161	189713248	88	88	Poly A (5), Poly T (5), Poly A (9), TG (3)	0,41	72072	0	0	100
21	3	190106030	190106104	75	74	Indel	0,57	60543	0	63	99,9
22	4	2233667	2233744	78	78	Poly A (6)	0,26	63882	0	0	100
23	4	7780541	7780637	97	97	Poly G (6), Poly T (5), Poly A (5)	0,42	79443	0	0	100
24	4	15688604	15688681	78	78	N/D	0,29	63882	0	0	100
25	4	56236521	56236586	66	62	Poly A (5), indel	0,36	50778	0	0	100
26	4	102839244	102839314	71	69	Poly A (5)	0,46	56511	0	0	100
27	4	164446743	164446804	62	62	Poly A (7), indel	0,27	50778	0	0	100
28	5	1882081	1882158	78	75	N/D	0,78	61425	0	0	100
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	68796	0	0	100
30	5	41069808	41069871	64	64	N/D	0,39	52416	0	0	100
31	5	74077114	74077196	83	83	Poly A (6), indel	0,3	67977	0	0	100
32	5	147475343	147475409	67	67	Poly T (5)	0,37	54873	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	74529	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	61425	0	0	100
35	6	6318713	6318814	102	102	Poly G (6)	0,68	83538	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	75348	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	76608	0	378	99,5
38	6	32147987	32148084	98	98	Poly T (5), TCT (3), CTT(3)	0,55	80262	0	0	100
39	6	32986864	32986958	95	95	Indel	0,53	77805	0	0	100
40	6	33408498	33408583	86	86	Poly C (6)	0,7	70434	0	0	100
41	6	41647401	41647495	95	94	Poly G (5), indel	0,61	76986	0	0	100
42	6	112435865	112435955	91	91	Poly A (5)	0,44	74529	0	0	100

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
43	7	22202076	22202148	73	73	N/D	0,44	59787	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	72072	0	0	100
45	7	77365735	77365821	87	87	Poly A (7), AG(4)	0,26	71253	0	0	100
46	7	110939946	110940030	85	85	Indel	0,38	69615	0	0	100
47	7	128533468	128533557	90	90	Poly G (5), indel	0,62	73710	0	0	100
48	7	149503875	149503965	91	91	Poly G (6), Poly C (6), indel	0,71	74529	0	0	100
49	7	154404519	154404599	81	66	N/D	0,31	54054	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	76167	0	0	100
51	8	1817312	1817394	83	83	N/D	0,42	67977	0	0	100
52	8	24811020	24811109	90	89	Poly G (7), CTC (4), indel	0,61	72171	0	720	99,0
53	8	76518625	76518691	67	67	Indel	0,3	54873	0	0	100
54	9	103054909	103055006	98	98	Poly G (6)	0,67	80262	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	53235	0	0	100
56	9	107620823	107620918	96	96	N/D	0,49	78624	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	67977	0	0	100
58	9	138995345	138995441	97	97	Poly C (6), indel	0,68	79443	0	0	100
59	10	5987120	5987198	79	78	Poly G (5), indel	0,47	63882	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	74529	0	0	100
61	10	27317777	27317855	79	79	Poly T (5)	0,3	64701	0	0	100
62	10	33018351	33018440	90	90	Poly A (5), Poly T (5)	0,2	73710	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	77805	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	71747	0	325	99,5
65	10	101611250	101611329	80	80	N/D	0,49	65520	0	0	100
66	10	118351373	118351453	81	81	N/D	0,51	66339	0	0	100

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
67	11	8159816	8159912	97	96	N/D	0,45	78624	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	57330	0	0	100
69	11	47470345	47470444	100	100	N/D	0,65	81900	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	50778	0	0	100
71	11	64418856	64418957	102	102	N/D	0,59	83538	0	0	100
72	11	93529612	93529684	73	73	Poly A (5)	0,4	59787	0	0	100
73	11	101347052	101347136	85	85	N/D	0,42	69615	0	0	100
74	11	102477336	102477426	91	91	Poly G (6)	0,55	74529	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	69615	0	0	100
76	11	120357801	120357885	85	85	Poly A (5), CA(3), indel	0,34	69615	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	69615	0	0	100
78	12	2834770	2834853	84	84	Poly C (5), indel	0,52	68796	0	0	100
79	12	26811004	26811096	93	93	Poly A (7), AC(4)	0,33	76167	0	0	100
80	12	30881766	30881846	81	81	N/D	0,49	66339	0	0	100
81	12	88474105	88474175	71	71	Poly A (6)	0,35	58149	0	0	100
82	12	120966872	120966966	95	95	Poly G (5)	0,68	77805	0	0	100
83	13	24167504	24167576	73	73	N/D	0,52	59787	0	0	100
84	13	25816961	25817049	89	88	Poly A (5), Poly T (7), Poly A (7), indel	0,22	72072	0	0	100
85	13	44880112	44880200	89	89	Indel	0,49	72891	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	63063	0	0	100
87	14	31619327	31619393	67	67	GA(3),TA(3)	0,39	54873	0	0	100
88	14	39517884	39517966	83	83	N/D	0,25	67977	0	0	100
89	14	46958962	46959034	73	72	Poly T (5), indel	0,19	58642	0	326	99,4
90	14	58050030	58050110	81	81	Indel	0,38	66339	0	0	100

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
91	14	82390559	82390649	91	91	Indel	0,35	74529	0	0	100
92	14	92549544	92549609	66	66	Poly A (5)	0,41	54054	0	0	100
93	14	102808496	102808589	94	94	Indel	0,62	76986	0	0	100
94	15	43170751	43170848	98	96	Poly C (5)	0,45	78624	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	55692	0	0	100
96	15	77879807	77879901	95	93	Poly G (5), indel	0,68	76167	0	0	100
97	15	81625334	81625428	95	95	Poly T (6)	0,43	77805	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	58149	0	0	100
99	15	89817413	89817503	91	91	N/D	0,36	74529	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	57330	0	0	100
101	16	1894910	1894972	63	63	N/D	0,27	51597	0	0	100
102	16	28997904	28997998	95	95	Poly C (5)	0,67	77805	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0,41	71253	0	0	100
104	16	57954406	57954509	104	104	Poly C (5)	0,67	85176	0	0	100
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	74529	0	0	100
106	17	3563920	3564008	89	89	GC(3)	0,64	72891	0	0	100
107	17	3594191	3594277	87	87	Poly C (5), indel	0,67	71247	0	6	100
108	17	3970090	3970180	91	91	Indel	0,46	74529	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	76167	0	0	100
110	17	33998759	33998849	91	89	Poly T (5)	0,54	72891	0	0	100
111	17	39589691	39589774	84	82	Poly A (13), indel (x2)	0,29	66343	27	788	98,8
112	17	41244394	41244484	91	91	Poly A (5)	0,34	74529	0	0	100
113	17	45438866	45438957	92	92	Poly A (7), AT(3), AT(4), AT(4), indel	0,26	75348	0	0	100
114	17	61502432	61502510	79	79	Indel	0,41	64413	0	288	99,6
115	17	64023582	64023667	86	86	Poly T (7)	0,22	70434	0	0	100

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
116	17	72308237	72308320	84	84	GAG(3)	0,62	68796	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	54873	0	0	100
118	18	6980478	6980568	91	91	N/D	0,37	74529	0	0	100
119	18	9888026	9888094	69	69	Poly A (6), TG(3)	0,43	56511	0	0	100
120	18	38836999	38837073	75	75	Poly A (5), indel	0,37	61425	0	0	100
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	66339	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0,45	69615	0	0	100
123	18	59773996	59774060	65	65	N/D	0,48	53235	0	0	100
124	19	625143	625241	99	99	N/D	0,59	81081	0	0	100
125	19	18121418	18121491	74	74	N/D	0,68	60605	1	0	100
126	19	18186574	18186643	70	70	N/D	0,64	57330	0	0	100
127	20	746056	746149	94	94	N/D	0,61	76986	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	67158	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	62244	0	0	100
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0,46	57330	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	82719	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	54873	0	0	100
133	20	62331904	62331994	91	88	Poly G (6)	0,73	72072	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	71253	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	54054	0	0	100
136	21	33694176	33694273	98	98	Poly T (6), CA(3)	0,54	80262	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	71253	0	0	100
138	21	46644924	46644992	69	69	Poly A (6), AG(3), indel	0,32	56439	0	72	99,9
139	21	46705575	46705664	90	90	Poly T (5), Poly A (6)	0,5	73710	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	81900	0	0	100

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
141	22	32439233	32439329	97	97	N/D	0,68	79443	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	79443	0	0	100
143	22	37637596	37637694	99	99	N/D	0,6	81081	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	75348	0	0	100
145	X	15870424	15870492	69	69	Poly T (5)	0,26	56511	0	0	100
146	X	135288543	135288611	69	69	Poly C (5)	0,62	56511	0	0	100
147	X	135290777	135290847	71	71	N/D	0,52	58149	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	Poly A (5)	0,37	0	0	0	N/D

Os resultados de sequenciação da amostra NA12878 foram comparados a um genótipo com um nível de confiança muito elevado para NA12878 e estabelecido pelo NIST (National Institute of Standards and Technology – Instituto Nacional de Normas e Tecnologia) (v.2.19). Dos 150 fragmentos amplificados, 92 fragmentos amplificados foram totalmente incluídos em regiões genómicas com um nível de confiança muito elevado, 41 fragmentos amplificados tinham sobreposição parcial e 17 fragmentos amplificados não tinham sobreposição na sequência do NIST. Isto resultou em 10 000 coordenadas por réplica para comparação. Foram comparadas identificações de base não variantes com o genoma humano de referência do conjunto hg19. Os resultados de precisão são apresentados na **Tabela 12**.

Tabela 12 Concordância da linha germinal da amostra NA12878 com a base de dados do NIST

Amostra	N.º de fragmentos amplificados	Taxa média de identificações	TP	FN	TN	FP	PPA	NPA	OPA
NA12878	133	> 99,9	6552	1	610470	0	> 99,9	100	> 99,9

Com base nos dados fornecidos por este estudo de linha germinal de nove ensaios, o NextSeq 550Dx Instrument pode sequenciar de forma consistente:

- ▶ Conteúdo GC \geq 19% (todas as bases identificadas em 819 fragmentos amplificados sequenciados com 19% de conteúdo GC identificado corretamente com uma taxa de não identificação de 0,6%)
- ▶ Conteúdo GC \leq 87% (todas as bases identificadas em 819 fragmentos amplificados sequenciados com 87% de conteúdo GC identificado corretamente com zero não identificações)
- ▶ Tamanhos PolyA \leq 9 (todas as bases identificadas em 819 fragmentos amplificados sequenciados contendo uma repetição PolyA de nove nucleótidos identificados corretamente com zero não identificações)
- ▶ Tamanhos PolyT \leq 10 (todas as bases identificadas em 819 fragmentos amplificados sequenciados contendo uma repetição PolyT de dez nucleótidos identificados corretamente com zero não identificações)
- ▶ Tamanhos PolyG \leq 7 (todas as bases identificadas em 819 fragmentos amplificados sequenciados contendo uma repetição PolyG de sete nucleótidos identificados corretamente com uma taxa de não identificação de 1,0%)
- ▶ Tamanhos PolyC \leq 6 (todas as bases identificadas em 2457 fragmentos amplificados sequenciados contendo uma repetição PolyC de seis nucleótidos identificados corretamente com zero não identificações)
- ▶ Tamanhos de repetição de dinucleótidos \leq 11x (todas as bases identificadas em 819 fragmentos amplificados sequenciados contendo uma repetição de dinucleótidos de 11x foram identificadas corretamente com uma taxa de não identificação de 0,5%)
- ▶ Tamanhos de repetição de trinucleótidos \leq 5x (todas as bases identificadas em 819 fragmentos amplificados sequenciados contendo uma repetição de trinucleótidos de 5x foram identificadas corretamente com uma taxa de não identificação de 0,5%)
- ▶ Tamanhos de inserção \leq 24 (66 343 de 66 370 bases identificadas em 819 fragmentos amplificados sequenciados contendo uma inserção de 24 nucleótidos foram identificadas corretamente com uma taxa de não identificação de 1,2%; não ocorreram identificações incorretas na região que continha a inserção de 24 nucleótidos)
- ▶ Tamanhos de eliminação \leq 25 (todas as bases identificadas em 2457 fragmentos amplificados sequenciados contendo uma eliminação de 25 nucleótidos foram identificadas corretamente com zero não identificações)

Somático

O estudo aqui descrito foi utilizado para avaliar a precisão da identificação de variantes do Somatic Variant Module no NextSeq 550Dx Instrument utilizando o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Este estudo utilizou um ensaio representativo concebido para analisar uma variedade de genes abrangendo 12 588 bases (150 fragmentos amplificados) em 23 cromossomas diferentes. O ADN do Platinum Genome foi extraído de blocos tratados com FFPE para gerar seis amostras únicas para avaliação no estudo.

A amostra de ADN GM12877 foi diluída na amostra de ADN GM12878 para criar GM12877-D5 e GM12877-D7 como um conjunto único de variantes heterozigóticas com frequência de variantes perto de 5% e 7%. A amostra de ADN GM12878 foi diluída de forma semelhante na amostra de ADN GM12877 para criar GM12878-D5 e GM12878-D7. Cada uma das amostras foi testada em triplicado, à exceção das amostras diluídas, que foram testadas em réplicas de seis. Foram realizados nove ensaios no total utilizando três instrumentos de sequenciação, três lotes de reagentes e três operadores em cinco dias de início. Foi determinada a precisão para SNVs, inserções e eliminações comparando os resultados ao método de referência composto e bem caracterizado, Platinum Genome versão 2016- 1.0. Foram definidas regiões genómicas de confiança com base neste método de referência, exceto quando especificado em contrário.

Tabela 13 Resumo da concordância somática

Critérios	Observações totais ¹	Resultado por observação ²	Resultado por ensaio ³
PPA para SNV	378	98,9	99,9
PPA para inserções	378	96,9	99,9
PPA para eliminações	378	97,1	99,9
NPA	378	> 99,9	> 99,9
OPA	378	> 99,9	> 99,9

¹Calculado como número de amostras por ensaio (42) x número de ensaios (9) = 378.

²Valor mais baixo observado por réplica de amostra nos 9 ensaios.

³Valor mais baixo quando os dados de cada ensaio são analisados em agregado.

A **Tabela 14** contém os dados do estudo apresentados com uma percentagem de concordância positiva e negativa por amostra, em que os resultados da variante são comparados ao método de referência composto e bem caracterizado para cálculos de PPA. Os três tipos de variantes (SNVs, inserções e eliminações) são combinados. Uma vez que o método de referência só fornece resultados para as variantes de nucleótido único e inserções/eliminações, os resultados de base não variante são comparados ao conjunto hg19 de sequenciação de referência do genoma humano para cálculos de NPA.

Tabela 14 Concordância somática por amostra

Amostra	Taxa média de identificações	Esperado	TP	FN	Não identificação de variantes	TN	FP	PPA	NPA	OPA
GM12877	98,7	2052	2025	0	27	318682	15	100	> 99,9	> 99,9
GM12878	98,8	3645	3564	0	81	317645	0	100	100	100
GM12879	99,8	2592	2538	0	54	323614	2	100	> 99,9	> 99,9
GM12884	99,8	3078	3024	0	54	322038	5	100	> 99,9	> 99,9
GM12885	99,8	3294	3213	0	81	322121	0	100	100	100
GM12888	99,8	2916	2889	0	27	323048	2	100	> 99,9	> 99,9
GM12877-D5	99,8	9288	8930	0	358	630621	0	100	100	100
GM12877-D7	99,7	9288	9032	0	256	629719	0	100	100	100
GM12878-D5	99,5	9288	8699	42	547	628582	0	99,5	100	> 99,9
GM12878-D7	99,7	9288	9108	0	180	629803	0	100	100	100

A **Tabela 15** contém os dados do estudo apresentados por amostra, em que os resultados das variantes são comparados com o método de referência composto e bem caracterizado. A deteção é avaliada para cada tipo de variante – SNVs, inserções e eliminações – separadamente. As posições de referência são excluídas.

Tabela 15 Concordância somática por amostra e por tipo de variante

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

As dez amostras foram analisadas novamente para identificar pequenas inserções e eliminações (indels) (Tabela 16). Foram identificados 71 indels no total com tamanho entre 1–24 bp no caso das inserções, e 1–25 bp no caso das eliminações.

Tabela 16 Resumo da detecção de indels somáticos

Tipo de variante	Variantes esperadas	TP	FN	Não identificação de variantes	PPA
Inserção	10773	10282	9	482	99,2
Eliminação	11502	10667	5	830	> 99,9

Os 150 fragmentos amplificados foram concebidos para abranger conteúdo genómico diverso. O conteúdo GC dos fragmentos amplificados variou entre 0,19–0,87%. Os fragmentos amplificados também tinham uma gama de repetições de nucleótido único (p. ex. PolyA, PolyT), dinucleótidos e trinucleótidos. Os dados foram compilados por fragmento amplificado (Tabela 17) para determinar o efeito do conteúdo genómico na percentagem de identificações corretas. A percentagem de identificações corretas consiste em identificações de referência e variantes e é inferior a 100% se ocorrerem identificações incorretas ou não identificações.

Tabela 17 Precisão ao nível do fragmento amplificado somático

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
1	1	36450499	36450591	93	93	Indel	0,22	35066	0	88	99,7
2	1	109465122	109465200	79	79	Poly A (5), Poly C (5), indel	0,38	29827	0	35	99,9
3	1	218353867	218353957	91	91	Indel	0,4	34202	0	283	99,2
4	1	223906657	223906748	92	92	Indel	0,49	34613	0	163	99,5
5	1	228526602	228526682	81	81	Poly G (5)	0,69	30571	0	47	99,8
6	1	236372039	236372108	70	70	Poly T (10), indel	0,39	26452	0	8	100,0
7	1	247812041	247812128	88	88	Poly A (5), CT (3), TAA(3), indel	0,27	33148	0	116	99,7
8	2	55862774	55862863	90	90	Indel	0,28	33928	0	92	99,7
9	2	87003930	87004009	80	80	Indel	0,38	30218	0	22	99,9
10	2	177016721	177016805	85	81	N/D	0,65	30616	0	2	> 99,9
11	2	186625727	186625801	75	75	Poly A (8)	0,35	28017	0	499	98,3
12	2	190323504	190323591	88	88	Poly T (5)	0,42	33207	0	57	99,8
13	2	200796740	200796826	87	87	Poly T (5), indel	0,31	32524	9	718	97,8
14	2	212245049	212245139	91	91	Poly T (5), Poly A (6), indel	0,3	33972	0	456	98,7
15	2	228147052	228147144	93	93	N/D	0,43	35051	0	103	99,7
16	2	235016350	235016422	73	73	Poly T (5), indel	0,42	27459	0	136	99,5

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
17	3	4466229	4466321	93	93	AT(3), indel	0,27	34534	0	620	98,2
18	3	46620561	46620643	83	83	N/D	0,43	31339	0	44	99,9
19	3	49851331	49851400	70	70	CT(3), indel	0,49	26373	0	87	99,7
20	3	189713161	189713248	88	88	Poly A (5), Poly T (5), Poly A (9), TG(3)	0,41	32829	0	857	97,5
21	3	190106030	190106104	75	74	Indel	0,57	27925	0	47	99,8
22	4	2233667	2233744	78	78	Poly A (6)	0,26	29327	4	162	99,4
23	4	7780541	7780637	97	97	Poly G (6), Poly T (5), Poly A (5)	0,42	36585	0	117	99,7
24	4	15688604	15688681	78	78	N/D	0,29	29427	0	57	99,8
25	4	56236521	56236586	66	62	Poly A (5), indel	0,36	23356	5	75	99,7
26	4	102839244	102839314	71	69	Poly A (5)	0,46	25942	0	140	99,5
27	4	164446743	164446804	62	62	Poly A (7), indel	0,27	22944	0	560	97,6
28	5	1882081	1882158	78	75	N/D	0,78	28299	0	53	99,8
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	31658	0	94	99,7
30	5	41069808	41069871	64	64	N/D	0,39	24120	0	72	99,7
31	5	74077114	74077196	83	83	Poly A (6), indel	0,3	31297	0	77	99,8
32	5	147475343	147475409	67	67	Poly T (5)	0,37	25277	0	55	99,8
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	34308	0	90	99,7
34	5	155662213	155662287	75	75	Indel	0,43	28266	0	163	99,4
35	6	6318713	6318814	102	102	Poly G (6)	0,68	38489	0	67	99,8
36	6	24949983	24950074	92	92	Indel	0,63	34730	0	46	99,9
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	35057	0	483	98,6

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
38	6	32147987	32148084	98	98	Poly T (5), TCT (3), CTT(3)	0,55	36647	0	406	98,9
39	6	32986864	32986958	95	95	Indel	0,53	35681	0	238	99,3
40	6	33408498	33408583	86	86	Poly C (6)	0,7	32438	0	70	99,8
41	6	41647401	41647495	95	94	Poly G (5), indel	0,61	35441	0	91	99,7
42	6	112435865	112435955	91	91	Poly A (5)	0,44	34354	0	44	99,9
43	7	22202076	22202148	73	73	N/D	0,44	27575	0	28	99,9
44	7	66276100	66276187	88	88	Indel	0,35	33060	0	213	99,4
45	7	77365735	77365821	87	87	Poly A (7), AG (4)	0,26	32423	0	489	98,5
46	7	110939946	110940030	85	85	Indel	0,38	32074	0	56	99,8
47	7	128533468	128533557	90	90	Poly G (5), indel	0,62	33791	0	281	99,2
48	7	149503875	149503965	91	91	Poly G (6), Poly C (6), indel	0,71	34316	0	82	99,8
49	7	154404519	154404599	81	66	N/D	0,31	24901	0	47	99,8
50	7	156476507	156476599	93	93	Indel	0,35	35067	0	87	99,8
51	8	1817312	1817394	83	83	N/D	0,42	31365	0	9	> 99,9
52	8	24811020	24811109	90	89	Poly G (7), CTC (4), indel	0,61	32781	0	890	97,4
53	8	76518625	76518691	67	67	Indel	0,3	25228	0	146	99,4
54	9	103054909	103055006	98	98	Poly G (6)	0,67	36968	0	76	99,8
55	9	105586150	105586214	65	65	Indel	0,32	24472	0	100	99,6
56	9	107620823	107620918	96	96	N/D	0,49	36203	0	85	99,8
57	9	123769149	123769231	83	83	AT(3)	0,37	31329	0	45	99,9
58	9	138995345	138995441	97	97	Poly C (6), indel	0,68	36472	0	201	99,5

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
59	10	5987120	5987198	79	78	Poly G (5), indel	0,47	29473	0	11	> 99,9
60	10	11784629	11784726	98	91	GC(3)	0,87	34188	0	213	99,4
61	10	27317777	27317855	79	79	Poly T (5)	0,3	29843	0	19	99,9
62	10	33018351	33018440	90	90	Poly A (5), Poly T (5)	0,2	33968	0	68	99,8
63	10	45084159	45084253	95	95	Indel	0,35	35829	0	81	99,8
64	10	55892599	55892687	89	88	AC(11), indel	0,42	32098	88	2048	93,8
65	10	101611250	101611329	80	80	N/D	0,49	30217	0	28	99,9
66	10	118351373	118351453	81	81	N/D	0,51	30531	0	96	99,7
67	11	8159816	8159912	97	96	N/D	0,45	36105	0	192	99,5
68	11	30177648	30177717	70	70	Indel	0,46	26318	0	153	99,4
69	11	47470345	47470444	100	100	N/D	0,65	37785	0	24	99,9
70	11	59837679	59837740	62	62	Indel	0,37	23368	0	68	99,7
71	11	64418856	64418957	102	102	N/D	0,59	38546	0	10	> 99,9
72	11	93529612	93529684	73	73	Poly A (5)	0,4	27516	0	78	99,7
73	11	101347052	101347136	85	85	N/D	0,42	32083	0	48	99,9
74	11	102477336	102477426	91	91	Poly G (6)	0,55	34047	0	369	98,9
75	11	118406285	118406369	85	85	Indel	0,53	32065	0	74	99,8
76	11	120357801	120357885	85	85	Poly A (5), CA (3), indel	0,34	32083	0	47	99,9
77	11	125769313	125769397	85	85	GA(3)	0,52	32103	0	27	99,9
78	12	2834770	2834853	84	84	Poly C (5), indel	0,52	31645	16	525	98,3
79	12	26811004	26811096	93	93	Poly A (7), AC (4)	0,33	34824	0	330	99,1
80	12	30881766	30881846	81	81	N/D	0,49	30497	0	121	99,6
81	12	88474105	88474175	71	71	Poly A (6)	0,35	26773	0	65	99,8

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
82	12	120966872	120966966	95	95	Poly G (5)	0,68	35830	9	72	99,8
83	13	24167504	24167576	73	73	N/D	0,52	27498	0	114	99,6
84	13	25816961	25817049	89	88	Poly A (5), Poly T (7), Poly A (7), indel	0,22	32824	0	566	98,3
85	13	44880112	44880200	89	89	Indel	0,49	33574	0	77	99,8
86	13	77665218	77665294	77	77	Indel	0,39	29075	0	31	99,9
87	14	31619327	31619393	67	67	GA(3),TA(3)	0,39	25313	0	13	99,9
88	14	39517884	39517966	83	83	N/D	0,25	31360	0	22	99,9
89	14	46958962	46959034	73	72	Poly T (5), indel	0,19	26499	0	717	97,4
90	14	58050030	58050110	81	81	Indel	0,38	30494	0	133	99,6
91	14	82390559	82390649	91	91	Indel	0,35	34313	0	86	99,7
92	14	92549544	92549609	66	66	Poly A (5)	0,41	24555	0	1527	94,1
93	14	102808496	102808589	94	94	Indel	0,62	35472	0	69	99,8
94	15	43170751	43170848	98	96	Poly C (5)	0,45	36264	0	24	99,9
95	15	63446149	63446216	68	68	Indel	0,25	25667	0	37	99,9
96	15	77879807	77879901	95	93	Poly G (5), indel	0,68	34745	0	432	98,8
97	15	81625334	81625428	95	95	Poly T (6)	0,43	35870	0	40	99,9
98	15	85438263	85438334	72	71	Indel	0,65	26762	0	76	99,7
99	15	89817413	89817503	91	91	N/D	0,36	34286	0	112	99,7
100	15	89864274	89864343	70	70	Indel	0,56	26449	0	11	> 99,9
101	16	1894910	1894972	63	63	N/D	0,27	23809	0	5	> 99,9
102	16	28997904	28997998	95	95	Poly C (5)	0,67	35860	0	50	99,9
103	16	53682908	53682994	87	87	TA(3)	0,41	32835	0	60	99,8
104	16	57954406	57954509	104	104	Poly C (5)	0,67	39177	0	144	99,6

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	34075	0	323	99,1
106	17	3563920	3564008	89	89	GC(3)	0,64	33632	0	11	> 99,9
107	17	3594191	3594277	87	87	Poly C (5), indel	0,67	32752	0	134	99,6
108	17	3970090	3970180	91	91	Indel	0,46	34343	0	82	99,8
109	17	16084945	16085037	93	93	Indel	0,26	35077	0	78	99,8
110	17	33998759	33998849	91	89	Poly T (5)	0,54	33553	0	89	99,7
111	17	39589691	39589774	84	82	Poly A (13), indel (x2)	0,29	30554	53	2296	92,9
112	17	41244394	41244484	91	91	Poly A (5)	0,34	34360	0	38	99,9
113	17	45438866	45438957	92	92	Poly A (7), AT (3), AT(4), AT(4), indel	0,26	34367	0	418	98,8
114	17	61502432	61502510	79	79	Indel	0,41	29751	0	119	99,6
115	17	64023582	64023667	86	86	Poly T (7)	0,22	32176	0	340	99,0
116	17	72308237	72308320	84	84	GAG(3)	0,62	31604	7	141	99,5
117	18	2616456	2616522	67	67	GA(3)	0,31	25273	8	45	99,8
118	18	6980478	6980568	91	91	N/D	0,37	34386	0	12	> 99,9
119	18	9888026	9888094	69	69	Poly A (6), TG (3)	0,43	25692	0	399	98,5
120	18	38836999	38837073	75	75	Poly A (5), indel	0,37	27923	0	893	96,9
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	30598	0	20	99,9
122	18	54815665	54815749	85	85	CT(3), indel	0,45	31969	0	161	99,5
123	18	59773996	59774060	65	65	N/D	0,48	24531	0	48	99,8
124	19	625143	625241	99	99	N/D	0,59	37298	0	124	99,7
125	19	18121418	18121491	74	74	N/D	0,68	27881	0	109	99,6

Fragmento amplificado	Cromossoma	Fragmento amplificado inicial	Fragmento amplificado final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas
126	19	18186574	18186643	70	70	N/D	0,64	26442	0	26	99,9
127	20	746056	746149	94	94	N/D	0,61	35501	0	31	99,9
128	20	10633195	10633276	82	82	AC(3)	0,59	30951	0	72	99,8
129	20	17705633	17705708	76	76	CT(3)	0,58	28686	0	42	99,9
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0,46	26372	0	88	99,7
131	20	25278421	25278521	101	101	Indel	0,63	38159	0	20	99,9
132	20	50897302	50897368	67	67	Indel	0,36	25188	0	544	97,9
133	20	62331904	62331994	91	88	Poly G (6)	0,73	32969	0	309	99,1
134	20	62690860	62690946	87	87	Indel	0,57	32818	0	77	99,8
135	21	30300823	30300888	66	66	Indel	0,35	24758	9	181	99,2
136	21	33694176	33694273	98	98	Poly T (6), CA (3)	0,54	36902	0	160	99,6
137	21	36710706	36710792	87	87	GT(3), indel	0,39	32841	0	48	99,9
138	21	46644924	46644992	69	69	Poly A (6), AG (3), indel	0,32	25939	0	280	98,9
139	21	46705575	46705664	90	90	Poly T (5), Poly A (6)	0,5	33942	0	78	99,8
140	22	25750774	25750873	100	100	Indel	0,63	37733	0	86	99,8
141	22	32439233	32439329	97	97	N/D	0,68	36617	0	49	99,9
142	22	37409844	37409940	97	97	Indel	0,46	36525	0	162	99,6
143	22	37637596	37637694	99	99	N/D	0,6	37398	0	24	99,9
144	22	47081347	47081438	92	92	Indel	0,66	34754	0	22	99,9
145	X	15870424	15870492	69	69	Poly T (5)	0,26	26046	0	36	99,9
146	X	135288543	135288611	69	69	Poly C (5)	0,62	26019	0	63	99,8
147	X	135290777	135290847	71	71	N/D	0,52	26780	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	Poly A (5)	0,37	0	0	0	N/D

Os resultados de sequenciação da amostra GM12878 foram comparados a um genótipo com um nível de confiança muito elevado para NA12878 e estabelecido pelo NIST (National Institutes of Standards and Technology, Instituto nacional de normas e tecnologia) (v.2.19). Dos 150 fragmentos amplificados, 92 fragmentos amplificados foram totalmente incluídos em regiões genômicas com um nível de confiança muito elevado, 41 fragmentos amplificados tinham sobreposição parcial e 17 fragmentos amplificados não tinham sobreposição na sequência do NIST. Isto resultou em 10 000 coordenadas por réplica para comparação. Foram comparadas identificações de base não variantes com o genoma humano de referência do conjunto hg19. Os resultados de precisão são apresentados na **Tabela 18**.

Tabela 18 Concordância somática da amostra GM12878 com a base de dados do NIST

Amostra	N.º de fragmentos amplificados	Taxa média de identificações	TP	FN	TN	FP	PPA	NPA	OPA
GM12878	133	98,8	2808	0	258488	0	100	100	100

Com base nos dados fornecidos por este estudo somático de nove ensaios, o NextSeq 550Dx Instrument pode sequenciar de forma consistente:

- ▶ Conteúdo GC $\geq 19\%$ (todas as bases identificadas em 378 fragmentos amplificados sequenciados com 19% de conteúdo GC identificado corretamente com uma taxa de não identificação de 2,6%)
- ▶ Conteúdo GC $\leq 87\%$ (todas as bases identificadas em 378 fragmentos amplificados sequenciados com 87% de conteúdo GC identificado corretamente com uma taxa de não identificação de 0,6%)
- ▶ Tamanhos PolyA ≤ 9 (todas as bases identificadas em 378 fragmentos amplificados sequenciados contendo uma repetição PolyA de nove nucleótidos identificados corretamente com uma taxa de não identificação de 2,5%)
- ▶ Tamanhos PolyT ≤ 10 (todas as bases identificadas em 378 fragmentos amplificados sequenciados contendo uma repetição PolyT de dez nucleótidos identificados corretamente com uma taxa de não identificação inferior a 0,1%)
- ▶ Tamanhos PolyG ≤ 6 (todas as bases identificadas em 2268 fragmentos amplificados sequenciados contendo uma repetição PolyG de seis nucleótidos identificados corretamente com uma taxa de não identificação de 0,5%)
- ▶ Tamanhos PolyC ≤ 6 (todas as bases identificadas em 756 fragmentos amplificados sequenciados contendo uma repetição PolyC de seis nucleótidos identificados corretamente com uma taxa de não identificação de 0,4%)
- ▶ Tamanhos de repetição de dinucleótidos $\leq 4x$ (todas as bases identificadas em 1890 fragmentos amplificados sequenciados contendo uma repetição de dinucleótidos de 4x foram identificadas corretamente com uma taxa de não identificação de 0,9%)
- ▶ Tamanhos de repetição de trinucleótidos $\leq 5x$ (todas as bases identificadas em 378 fragmentos amplificados sequenciados contendo uma repetição de trinucleótidos de 5x foram identificadas corretamente com uma taxa de não identificação de 1,4%)
- ▶ Tamanhos de inserção ≤ 23 (todas as bases identificadas em 378 fragmentos amplificados sequenciados contendo uma inserção de 23 nucleótidos identificada corretamente com uma taxa de não identificação de 0,8%)
- ▶ Tamanhos de eliminação ≤ 25 (todas as bases identificadas em 1134 fragmentos amplificados sequenciados contendo uma eliminação de 25 nucleótidos foram identificadas corretamente com uma taxa de não identificação de 0,7%)

Precisão

A precisão do NextSeq 550Dx Instrument foi determinada através da análise de 13 amostras únicas de Platinum Genome utilizando três instrumentos, três lotes de reagentes e três operadores para gerar nove ensaios de sequenciação ao longo de cinco dias de início. O ensaio representativo, as amostras e o método de referência são os mesmos descritos para o estudo de precisão da linha germinal. Os contributos de precisão foram determinados através da análise do componente de variação utilizando VAF como variável de resposta e do cálculo dos desvios padrão ao nível do componente do instrumento, do lote de reagentes, do operador e do

dia de início (Tabela 19). O número total de observações utilizadas na análise de cada componente do instrumento, operador ou variabilidade do lote de reagentes foi de 699, 176 e 235 para SNVs, inserções e eliminações, respetivamente.

Tabela 19 Resultados de precisão para o NextSeq 550Dx Instrument (Desvio padrão)

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

Comparação de métodos (Plataforma de sequenciação)

Foram avaliadas amostras de sangue total e FFPE no NextSeq 550Dx Instrument e no instrumento MiSeqDx utilizando os fluxos de trabalho de linha germinal e somático do TruSeq Custom Amplicon Kit Dx. Foi avaliada a concordância da frequência de variantes em amostras sanguíneas e FFPE utilizando ensaios múltiplos representativos. A Figura 2 traça a correlação VAF entre os dois instrumentos para um ensaio representativo e a Tabela 20 resume esta correlação por painel de ensaio. Com base na forte correlação entre o instrumento MiSeqDx e o NextSeq 550Dx Instrument, as características de desempenho relacionadas com fatores pré-analíticos (p. ex., métodos de extração ou substâncias interferentes) foram determinadas para serem aplicáveis a ambos os instrumentos. Consulte o folheto informativo do TruSeq Custom Amplicon Kit Dx para obter detalhes adicionais.

Figura 2 Correlação VAF de instrumentos MiSeqDx e NextSeq 550Dx para amostras FFPE (esquerda) e sanguíneas (direita) utilizando o ensaio 1

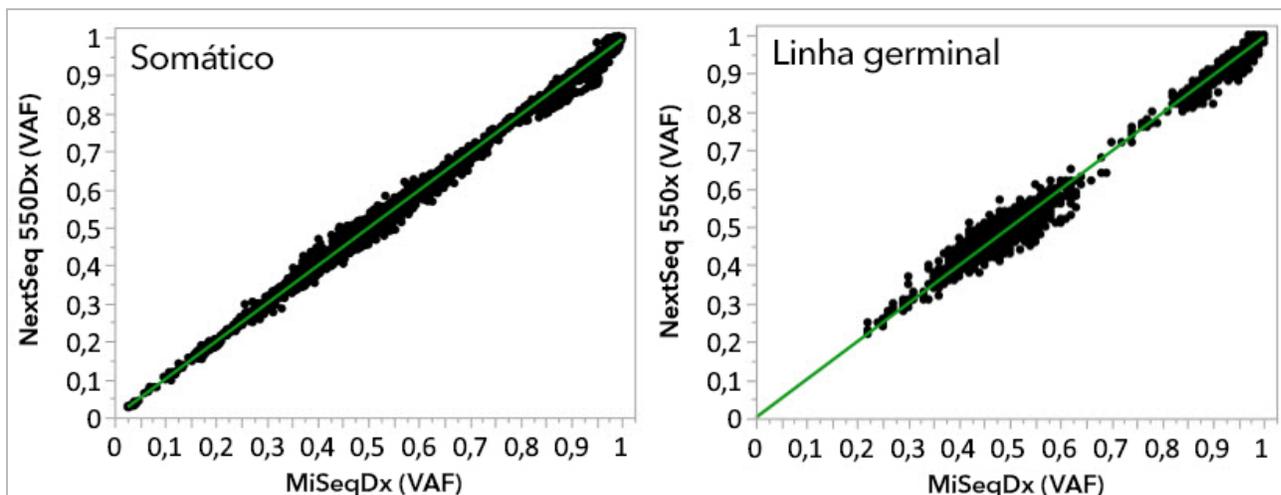


Tabela 20 Resultados da comparação de métodos com amostras únicas de sangue e FFPE

Fonte de gDNA	Ensaio (painel de oligonucleótidos)	Réplicas biológicas (amostras)	Réplicas técnicas (por amostra)	Observações (n.º de variantes)	Tangente	Interceção	Correlação (R ²)
Sangue	Ensaio 1	45	2	8369 ¹	0,992	0,002	0,995 ²
Sangue	Ensaio 2	45	2	5457	0,995	0,005	0,981
FFPE	Ensaio 1	46	2	8319	0,993	0,000	0,997 ²
FFPE	Ensaio 3	40	1	280	0,969	0,015	0,978

¹Foram removidos dois pontos de dados com base na limitação indicada para o Germline Variant Module.

²Coefficiente de determinação dos traçados VAF conforme ilustrado na Figura 2.

Reprodutibilidade

A reprodutibilidade do NextSeq 550Dx Instrument foi avaliada utilizando amostras de Platinum Genome com um ensaio representativo concebido para analisar, utilizando 150 fragmentos amplificados, uma variedade de genes abrangendo 12 588 em 23 cromossomas diferentes. O teste da linha germinal incluiu sete réplicas de 13 amostras; o teste somático incluiu seis réplicas de sete amostras em diferentes níveis VAF. As amostras foram preparadas utilizando o TruSeq Custom Amplicon Kit Dx.

Os testes foram realizados em três centros clínicos externos utilizando um lote do NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Foi utilizado um único NextSeq 550Dx Instrument em cada centro clínico. Dois operadores realizaram os testes em cada centro clínico. Cada operador realizou os testes em três dias de início não consecutivos em cada tipo de amostra para um total de 36 ensaios nos três centros clínicos. Estes testes resultaram em 18 ensaios para cada um dos fluxos de trabalho somático e da linha germinal.

Linha germinal

As variantes da linha germinal com nível VAF $\geq 0,2$ são reportadas como positivas (variante). No caso das variantes positivas da linha germinal, os dados foram avaliados em relação à taxa de não identificação e à taxa de identificação correta e positiva em cada tipo de variante (SNV, inserção, eliminação). A [Tabela 21](#) resume as taxas observadas, juntamente com os níveis de confiança inferiores e superiores de 95% (LCL/UCL) calculados utilizando o método de pontuação de Wilson, para cada tipo de variante.

Tabela 21 Observações de identificação de linha germinal para resultados positivos esperados por tipo de variante

Tipo de variante	Não identificação			Identificação correta e positiva				
	Observado	Total	Porcentagem	Observado	Total	Porcentagem	95% LCL	95% UCL
SNV	16	110 376	0,014	110 349	110 360	99,99	99,98	99,99
Inserções	1026	37 044	2,77	36 018	36 018	100	99,99	100,00
Eliminações	648	34 776	1,86	34 128	34 128	100	99,99	100,00

As variantes da linha germinal com nível VAF $< 0,2$ são reportadas como negativas (tipo selvagem). No caso das localizações de linha germinal negativas esperadas, os dados foram avaliados em relação a taxas de não identificação e identificação correta do tipo selvagem. A [Tabela 22](#) resume as taxas observadas, juntamente com os níveis de confiança inferiores e superiores de 95% (LCL/UCL) calculados utilizando o método de pontuação de Wilson.

Tabela 22 Observações de identificação de linha germinal para resultados negativos esperados

Tipo de variante	Não identificação			Identificação correta e negativa				
	Observado	Total	Porcentagem	Observado	Total	Porcentagem	95% LCL	95% UCL
Tipo selvagem	4883	19 600 182	0,025	19 595 299	19 595 299	100	100,00	100,00

As variantes de linha germinal com um nível VAF $\geq 0,2$ e $< 0,7$ são identificadas como heterozigóticas positivas no que respeita a variante, e as variantes com nível VAF $\geq 0,7$ são identificadas como homozigóticas positivas no que respeita a variante. As amostras de linha germinal com variantes heterozigóticas foram utilizadas para determinar se a variabilidade inerente do ensaio afetaria a identificação do genótipo. O Cx foi determinado para ambos os limites (0,2 para genótipos heterozigóticos e 0,7 para homozigóticos), em que x corresponde à proporção dos testes repetidos que ultrapassaram o limite. Em relação ao limite inferior de 0,2 VAF, o Cx era $\geq 99,999\%$, indicando que $\geq 99,999\%$ de variantes heterozigóticas seriam identificadas como heterozigóticas. A respeito do limite superior de 0,7 VAF, o Cx era $\leq 0,001\%$, indicando assim que $\leq 0,001\%$ das variantes heterozigóticas seriam identificadas como homozigóticas. A [Tabela 23](#) resume os resultados por tipo de variante.

As variantes de linha germinal com um nível VAF $\geq 0,2$ e $< 0,7$ são identificadas como heterozigóticas positivas no que respeita a variante, e as variantes com nível VAF $\geq 0,7$ são identificadas como homozigóticas positivas no que respeita a variante. As amostras de linha germinal com variantes heterozigóticas foram utilizadas para determinar se a variabilidade inerente do ensaio afetaria a identificação do genótipo. O Cx foi determinado para ambos os limites (0,2 para genótipos heterozigóticos e 0,7 para homozigóticos), em que x corresponde à proporção dos testes repetidos que ultrapassaram o limite. A respeito do limite inferior de 0,2 VAF, o Cx era $\geq 99,999\%$, indicando assim que $\geq 99,999\%$ das variantes heterozigóticas seriam identificadas como heterozigóticas. Em relação ao limite superior de 0,7 VAF, o Cx era $\leq 0,001\%$, indicando que $\leq 0,001\%$ das variantes heterozigóticas seriam identificadas como homozigóticas. A [Tabela 23](#) resume os resultados por tipo de variante.

Tabela 23 Valores Cx da linha germinal para variantes heterozigóticas

Tipo de variante	Limite em 0,2 VAF	Limite em 0,7 VAF
	$\geq C99.999\%$	$\leq C0.001\%$
SNV	94/94	94/94
Inserções	24/24	24/24
Eliminações	35/35	35/35
Total	153	153

Somático

As variantes somáticas com níveis VAF $\geq 0,026$ são reportadas como positivas (variante). As observações com níveis VAF $\geq 0,01$ e $< 0,026$ foram consideradas como equívocas para o objetivo desta análise (nem positivas nem negativas, identificadas como frequência de baixa variante). Para avaliar o desempenho, os resultados foram calculados de três formas:

- ▶ Melhor cenário: qualquer resultado equívoco foi considerado como uma identificação correta e positiva (concordância com os resultados esperados)
- ▶ Pior cenário: qualquer resultado equívoco foi considerado como uma identificação incorreta (divergência com os resultados esperados)
- ▶ Caso de exclusão: qualquer resultado equívoco foi excluído da análise

As três tabelas, [Tabela 24](#), [Tabela 25](#) e [Tabela 26](#), resumem os resultados da identificação para o melhor cenário, pior cenário e caso de exclusão, respetivamente, juntamente com os níveis de confiança inferiores e superiores de 95% (LCL/UCL) calculados utilizando o método de pontuação de Wilson.

Tabela 24 Observações de identificação somática para resultados positivos esperados por tipo de variante (melhor cenário)

Tipo de variante	Identificação correta e positiva				
	Observado	Total	Percentagem	95% LCL	95% UCL
SNV	54 346	54 346	100	99,99	100,00
Inserções	18 036	18 036	100	99,98	100,00
Eliminações	18 381	18 381	100	99,98	100,00

Tabela 25 Observações de identificação somática para resultados positivos esperados por tipo de variante (pior cenário)

Tipo de variante	Identificação correta e positiva				
	Observado	Total	Percentagem	95% LCL	95% UCL
SNV	54 346	54 346	100	99,99	100,00
Inserções	18 000	18 036	99,8	99,72	99,86
Eliminações	18 381	18 381	100	99,98	100,00

Tabela 26 Observações de identificação somática para resultados positivos esperados por tipo de variante (identificações equívocas removidas)

Tipo de variante	Identificação correta e positiva				
	Observado	Total	Percentagem	95% LCL	95% UCL
SNV	54 346	54 346	100	99,99	100,00
Inserções	18 000	18 000	100	99,98	100,00
Eliminações	18 381	18 381	100	99,98	100,00

As variantes somáticas com nível VAF < 0,01 são reportadas como identificações negativas (tipo selvagem). No caso das localizações somáticas negativas esperadas, os dados foram avaliados em relação à taxa de não identificação e à taxa de identificação correta do tipo selvagem. As identificações corretas do tipo selvagem foram determinadas excluindo as não identificações e subtraindo as identificações observadas que caíram na zona equívoca (níveis VAF $\geq 0,01$ e < 0,026), bem como as identificações incorretas que estavam acima do limite (níveis VAF $\geq 0,026$) do total. A [Tabela 27](#) resume os resultados observados, totais e de percentagem em relação às localizações somáticas negativas para uma taxa de não identificação e uma taxa de identificação correta do tipo selvagem, juntamente com níveis de confiança superiores e inferiores de 95% (LCL/UCL) calculados utilizando o método de pontuação de Wilson.

Tabela 27 Observações da identificação somática para resultados negativos esperados

Tipo de variante	Não identificação			Identificação correta						
	Observado	Total	Percentagem	Equívoca	Incorreta	Correta	Total	Percentagem	95% LCL	95% UCL
Tipo selvagem	36 326	8 909 676	0,408	2254	121	8 870 975	8 873 350	99,97	99,972	99,974

As amostras somáticas em diferentes níveis VAF da mesma variante foram avaliadas para determinar o C95 do ensaio (no mesmo tipo de variante). Para avaliar a variabilidade perto do limite do ensaio, foram utilizadas as amostras que tinham níveis VAF esperados entre 0,02 e 0,07. O C95 foi determinado para cada variante, com o C95 mais elevado para cada tipo de variante reportado na [Tabela 28](#).

Tabela 28 Resumo do C95 somático

Tipo de variante	N	C95
SNV	74	0,0613
Inserção	24	0,0573
Eliminação	33	0,0575

Desempenho do NextSeq 550Dx High Output Reagent Kit v2.5 (300 Cycles)

Descrição geral

O NextSeq 550Dx é suportado por dois kits de reagentes: o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) e o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Para demonstrar que o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) pode cumprir os requisitos de desempenho analítico verificados e validados com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles), foram realizados estudos com o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Foram realizadas preparações de duas bibliotecas com o TruSeq Custom Amplicon Kit Dx, uma com o fluxo de trabalho da linha germinal e outra com o fluxo de trabalho somático. Foram testadas bibliotecas de cada fluxo de trabalho com três lotes do NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) utilizando três instrumentos NextSeq 550Dx. Além disso, os testes de cada fluxo de trabalho incluíram um ensaio único com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Sensibilidade analítica (limite de vazío [LoB] e limite de detecção [LoD])

A verificação com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) demonstrou que o NextSeq 550Dx Instrument deteta variantes a 0,05 VAF (Variant Allele Frequency, frequência de alelos de variantes) com um erro de Tipo II $\leq 0,05$ e que o limite de 0,026 VAF utilizado pelo Somatic Variant Module (LoB efetivo) suporta um erro de Tipo I $\leq 0,01$. Com base nestas alegações, prevê-se que uma variante a 0,05 VAF seja igual ou superior a 0,026 VAF em 95% do tempo e que uma posição de tipo selvagem seja inferior a 0,026 VAF em 99% do tempo. Para garantir que estas alegações foram cumpridas com o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), foram realizadas medições repetidas no NextSeq 550Dx Instrument com amostras de tipo selvagem (amostras LoB) e com amostras com variantes a 0,05 VAF (amostras LoD) utilizando o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). A proporção de identificações acima e abaixo do limite de 0,026 foi depois comparada com as alegações estabelecidas com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Os testes incluíram duas amostras LoD, cada uma com um conjunto único de variantes específicas de 0,05 VAF e as amostras LoB correspondentes que eram do tipo selvagem para as variantes especificadas. Para a preparação de bibliotecas, foram processadas amostras LoD e LoB em réplicas de oito e sete, respetivamente, utilizando o TruSeq Custom Amplicon Kit Dx. As bibliotecas foram inicialmente sequenciadas utilizando o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) para identificar variantes/coordenadas genómicas para avaliação LoB/LoD com o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Todas as variantes com um VAF médio entre 0,045 – 0,055 (variantes LoD) com base nos resultados do NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) foram utilizadas para análise LoD (N = 51 variantes). Para análise LoB, foram avaliadas as 51 coordenadas genómicas correspondentes.

Para avaliação do NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles), foram sequenciadas bibliotecas em três ensaios ao longo de três dias consecutivos utilizando o mesmo instrumento e lote de kit de reagentes. Estes testes totalizaram 24 réplicas para cada uma das 51 variantes LoD e 21 réplicas para cada uma das posições de tipo selvagem correspondentes. A proporção das identificações de tipo selvagem com VAF $< 0,026$ é fornecida na [Tabela 29](#). A proporção das identificações de variantes LoD com VAF igual ou superior a 0,026 é apresentada na [Tabela 30](#).

Tabela 29 Proporção de identificações < 0,026 para posições de tipo selvagem (avaliação de alegação LoB)

Tipo de variante	Posições avaliadas	Observações totais	N.º de medições VAF ≥ 2,6%	Proporção < 2,6%	Proporção 95% Intervalo de confiança
SNV	32	672	0	1	0,994 – 1
Inserção	11	231	0	1	0,984 – 1
Eliminação	8	168	0	1	0,978 – 1

Tabela 30 Proporção de identificações ≥ 0.026 VAF para variantes LoD (avaliação de alegações LoD)

Tipo de variante	Posições avaliadas	Observações totais	N.º de medições VAF < 2,6%	N.º de medições VAF ≥ 2,6%	Proporção ≥ 2,6%	Proporção 95% Intervalo de confiança
SNV	32	768	1	767	0,999	0,993 – 1
Inserção	11	264	3	261	0,989	0,967 – 0,996
Eliminação	8	192	2	190	0,99	0,963 – 0,997

Precisão

Linha germinal

O seguinte estudo foi realizado para avaliar a precisão da identificação de variantes com o Germline Variant Module utilizando o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Foram testadas doze amostras únicas do Platinum Genome utilizando um ensaio representativo. Foram realizados 11 ensaios no total utilizando três instrumentos NextSeq 550Dx e três NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

A precisão foi determinada para SNVs, inserções e eliminações comparando os resultados com um método de referência composto bem caracterizado, Platinum Genome versão 2016-1.0. São fornecidos, para referência, os resultados de precisão de um único ensaio de sequenciação com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Um resumo dos resultados é fornecido na [Tabela 31](#).

Tabela 31 Resumo de concordância da linha germinal

Critérios	Observações totais (v2.5) ¹	Resultado por observação (v2.5) ²	Resultado por observação (v2) ³	Resultado por ensaio (v2.5) ⁴	Resultado por ensaio (v2) ⁴
PPA para SNV	1056	98,7	98,7	> 99,9	> 99,9
PPA para inserções	1056	100	100	100	98,9
PPA para eliminações	1056	95,2	95,2	> 99,9	100
NPA	1056	100	100	100	100
OPA	1056	> 99,9	> 99,9	> 99,9	> 99,9

¹ Calculado como o número de amostras por ensaio x número de ensaios (96 amostras por ensaio x 11 ensaios = 1056 observações).

² Valor mais baixo observado por réplica de amostra em todos os ensaios (com base em 11 ensaios do NextSeq 550Dx High Output Reagent Kit v2.5).

³ Valor mais baixo observado por réplica de amostra num ensaio (96 observações totais).

⁴ Valor mais baixo quando os dados de cada ensaio são analisados em agregado.

Somático

O seguinte estudo foi realizado para avaliar a precisão da identificação de variantes do Somatic Variant Module no NextSeq 550Dx Instrument utilizando o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles). Foram testadas dez amostras FFPE Platinum Genome (duas com variantes diluídas para 0,05 VAF) utilizando um ensaio representativo. Foram realizados 11 ensaios no total utilizando três instrumentos NextSeq 550Dx e três lotes do NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

A precisão foi determinada para SNVs, inserções e eliminações comparando os resultados com um método de referência composto bem caracterizado, Platinum Genome versão 2016-1.0. São fornecidos, para referência, os resultados de precisão de um único ensaio de sequenciação com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles). Um resumo dos resultados é fornecido na [Tabela 32](#).

Tabela 32 Resumo da concordância somática

Critérios	Observações totais (v2.5) ¹	Resultado por observação (v2.5) ²	Resultado por observação (v2) ³	Resultado por ensaio (v2.5) ⁴	Resultado por ensaio (v2) ⁴
PPA para SNV	528	100	100	100	100
PPA para inserções	528	96,9	96,9	> 99,9	> 99,9
PPA para eliminações	528	100	100	100	100
NPA	528	> 99,9	> 99,9	> 99,9	> 99,9
OPA	528	> 99,9	> 99,9	> 99,9	> 99,9

¹Calculado como o número de amostras por ensaio x número de ensaios (48 amostras por ensaio x 11 ensaios = 528 observações).

²Valor mais baixo observado por réplica de amostra em todos os ensaios (com base em 11 ensaios do NextSeq 550Dx High Output Reagent Kit v2.5).

³Valor mais baixo observado por réplica de amostra num ensaio (96 observações totais).

⁴Valor mais baixo quando os dados de cada ensaio são analisados em agregado.

Precisão

Linha germinal

A precisão do NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) com o Germline Variant Module foi avaliada com amostras do Platinum Genome e um ensaio representativo. Os testes incluíram uma única preparação de bibliotecas utilizando o TruSeq Custom Amplicon Kit Dx e incluiu 12 amostras processadas com oito réplicas cada. As bibliotecas foram sequenciadas com três lotes do NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) e três instrumentos NextSeq 550Dx para um total de nove ensaios de sequenciação.

As amostras com variantes heterozigóticas foram utilizadas para determinar se a variabilidade inerente do ensaio afetaria a identificação do genótipo (N = 153 variantes heterozigóticas únicas). O Cx foi determinado para ambos os limites do Germline Variant Module (0,2 para genótipos heterozigóticos e 0,7 para homozigóticos), em que x corresponde à proporção dos testes repetidos que ultrapassaram o limite. No caso do limite inferior de 0,2 VAF, a variante com o Cx mínimo para o NextSeq 550Dx Reagent Kit v2.5 (300 cycles) foi > 99,9%, indicando que > 99,9% de variantes heterozigóticas seriam identificadas como heterozigóticas. No caso do limite superior de 0,7 VAF, a variante com o Cx máximo para o NextSeq 550Dx Reagent Kit v2.5 (300 cycles) foi < 1,5%, indicando que ≤ 1,5% de variantes heterozigóticas seriam identificadas como homozigóticas. A [Tabela 33](#) resume os resultados por tipo de variante. São fornecidos como referência os valores Cx de um ensaio único de sequenciação utilizando o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Tabela 33 Valores Cx da linha germinal para variantes heterozigóticas

Tipo de variante	N	Limite em 0,2 VAF		Limite em 0,7 VAF	
		Cx mín. (v2.5) ¹	Cx mín. (v2) ²	Cx máx. (v2.5) ¹	Cx máx. (v2) ²
SNV	94	> 99,9%	> 99,9%	1,5%	1,0%
Inserções	24	100%	100%	0%	< 0,1%
Eliminações	35	100%	> 99,9%	< 0,1%	< 0,1%

¹Valores Cx baseados em estimativas de desvio padrão total da análise de variações de componentes.

²Valores Cx baseados nos desvios padrão das amostras.

Somático

A precisão do NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) com o Somatic Variant Module foi avaliada com amostras do Platinum Genome FFPE e um ensaio representativo. Os testes incluíram uma única preparação de bibliotecas utilizando o TruSeq Custom Amplicon Kit Dx e duas amostras com oito réplicas cada. As bibliotecas foram sequenciadas utilizando três lotes do NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) e três instrumentos NextSeq 550Dx para um total de nove ensaios de sequenciação.

Foram utilizadas as variantes somáticas com níveis VAF esperados $\leq 0,10$ VAF (N = 131 variantes únicas) para avaliar a variabilidade do instrumento perto do limite de VAF do Somatic Variant Module (as variantes somáticas com nível VAF $\geq 0,026$ são identificadas como positivas para a variante). Os valores da C95 foram determinados para cada uma das variantes somáticas. Os valores da C95 representam a VAF em que a probabilidade de ser superior ao limite VAF do Somatic Variant Module é de 95%. Os valores mais elevados da C95 por tipo de variante são reportados na Tabela 34. São fornecidos como referência os resultados de C95 de um único ensaio de sequenciação utilizando o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles).

Tabela 34 Resumo do C95 somático

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

¹Valores de C95 baseados em estimativas de desvio padrão total da análise de variações de componentes.

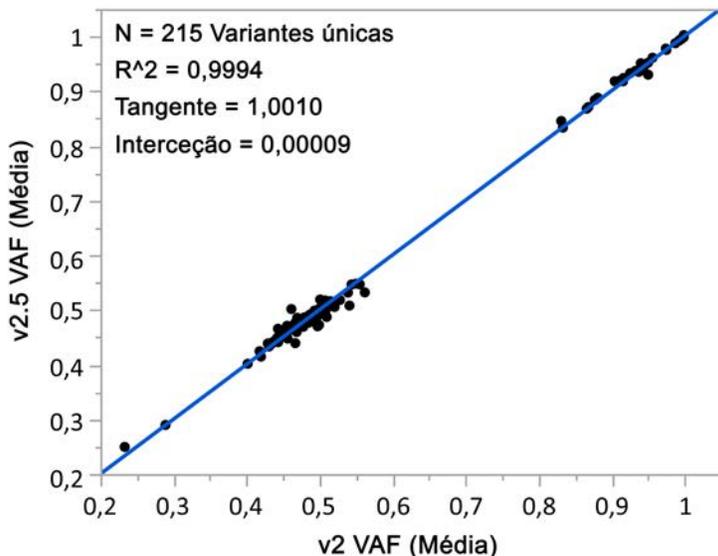
²Valores C95 baseados nos desvios padrão das amostras.

Comparação de métodos (kit de reagentes)

Linha germinal

Foram avaliadas as médias de VAFs de 215 variantes únicas no NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) e no NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) utilizando os resultados gerados a partir do Germline Variant Module. As médias de VAF foram calculadas a partir de 11 ensaios de sequenciação (v2.5) e um ensaio de sequenciação (v2). Foram utilizadas pelo menos oito réplicas para calcular a média de cada variante. Figura 3 traça a correlação VAF entre os dois kits de reagentes. Com base na forte correlação linear VAF e a semelhança dos resultados entre os kits de reagentes, determinou-se que as características de desempenho inicialmente verificadas e validadas com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) com o Germline Variant Module são aplicáveis ao NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

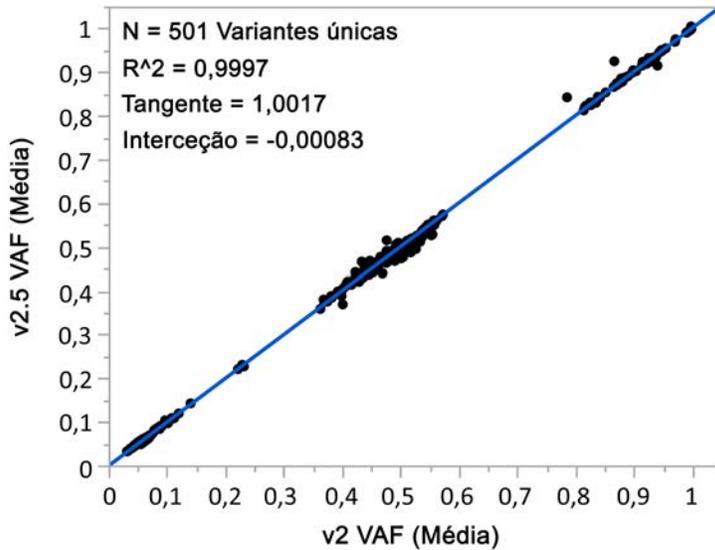
Figura 3 A correlação da VAF (Variant Allele Frequency, frequência de alelos de variantes) do Germline Variant Module entre o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) e o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).



Somático

Foram avaliadas as médias de VAFs para 501 variantes únicas no NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) e no NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles) utilizando os resultados gerados a partir do Somatic Variant Module. As médias de VAF foram calculadas a partir de 11 ensaios de sequenciação (v2.5) e um ensaio de sequenciação (v2). Foram utilizadas pelo menos três réplicas para calcular a média de cada variante única. [Figura 4](#) traça a correlação VAF entre os dois kits de reagentes. Com base na correlação VAF e a semelhança dos resultados entre os kits de reagentes, determinou-se que as características de desempenho verificadas e validadas com o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) com o Somatic Variant Module são aplicáveis ao NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).

Figura 4 A correlação da VAF (Variant Allele Frequency, frequência de alelos de variantes) do Somatic Variant Module entre o NextSeq 550Dx High Output Reagent Kit v2 (300 cycles) e o NextSeq 550Dx High Output Reagent Kit v2.5 (300 cycles).



Histórico de revisões

Documento	Data	Descrição da alteração
Document n.º 1000000030326 v06	Maio de 2022	Foram realizadas atualizações para retificar conteúdo adicionado inadvertidamente a partir de software de origem.
Documento n.º 1000000030326 v05	Novembro de 2021	Foi adicionada a declaração de Avisos e precauções sobre a comunicação de incidentes graves. Foi adicionada uma declaração aos Princípios do procedimento que indica o utilizador previsto. Foi removida uma referência ao High Output Reagent Kit v2 (300 cycles). Foi adicionada uma referência ao High Output Reagent Kit v2.5 (75 cycles).
Document n.º 1000000030326 v04	Agosto de 2021	Foi adicionada a tabela do Histórico de revisões. Foi atualizada a morada do Representante autorizado na UE.

Patentes e marcas comerciais

Este documento e respetivo conteúdo são propriedade da Illumina, Inc. e das suas afiliadas ("Illumina") e destinam-se unicamente à utilização contratual por parte dos clientes relativamente à utilização dos produtos descritos no presente documento e para nenhum outro fim. Este documento e respetivo conteúdo não podem ser utilizados ou distribuídos para qualquer outro fim e/ou de outra forma transmitidos, divulgados ou reproduzidos por qualquer via, seja de que natureza for, sem a autorização prévia por escrito da Illumina. A Illumina não concede qualquer licença ao abrigo da sua patente, marca comercial, direito de autor ou direitos de jurisprudência nem direitos semelhantes de quaisquer terceiros por via deste documento.

As instruções contidas neste documento têm de ser estrita e explicitamente seguidas por pessoal qualificado e com a devida formação para garantir a utilização adequada e segura dos produtos aqui descritos. Todo o conteúdo deste documento tem de ser integralmente lido e compreendido antes da utilização dos referidos produtos.

A NÃO OBSERVÂNCIA DA RECOMENDAÇÃO PARA LER INTEGRALMENTE E SEGUIR EXPLICITAMENTE TODAS AS INSTRUÇÕES AQUI CONTIDAS PODE RESULTAR EM DANOS NOS PRODUTOS, LESÕES EM PESSOAS, INCLUINDO NOS UTILIZADORES OU OUTROS, E EM DANOS MATERIAIS, E IRÁ ANULAR QUALQUER GARANTIA APLICÁVEL AOS PRODUTOS.

A ILLUMINA NÃO ASSUME QUALQUER RESPONSABILIDADE RESULTANTE DA UTILIZAÇÃO INADEQUADA DOS PRODUTOS AQUI DESCRITOS (INCLUINDO PARTES DOS MESMOS OU DO SOFTWARE).

© 2022 Illumina, Inc. Todos os direitos reservados.

Todas as marcas comerciais são propriedade da Illumina, Inc. ou dos respetivos proprietários. Para obter informações específicas sobre marcas comerciais, consulte www.illumina.com/company/legal.html.

Informações de contacto



Illumina

5200 Illumina Way

San Diego, Califórnia 92122 EUA

+1 800 809 ILMN (4566)

+1 858 202 4566 (fora da América do Norte)

techsupport@illumina.com

www.illumina.com



Illumina Netherlands B.V.

Steenoven 19

5626 DK Eindhoven

Países Baixos

Patrocinador australiano

Illumina Australia Pty Ltd

Nursing Association Building

Level 3, 535 Elizabeth Street

Melbourne, VIC 3000

Austrália

Etiquetas do produto

Para uma referência completa dos símbolos que podem ser apresentados nas embalagens e etiquetas do produto, consulte a chave de símbolos para o seu kit em support.illumina.com.