

Instrumento MiSeq™ Dx

PARA UTILIZAÇÃO EM DIAGNÓSTICO IN VITRO
APENAS PARA EXPORTAÇÃO

Utilização prevista

O instrumento MiSeqDx destina-se à sequenciação alvo das bibliotecas de ADN de ADN genómico humano extraído de sangue total periférico ou de tecido fixado em formalina, incorporado em parafina (FFPE), quando utilizado com ensaios de diagnóstico *in vitro* (IVD) efetuados no instrumento.

O instrumento MiSeqDx não se destina à sequenciação do genoma completo ou *de novo*.

O instrumento MiSeqDx destina-se a ser utilizado com reagentes de IVD específicos registados, autorizados ou aprovados e software de análise.

Princípios do procedimento

O MiSeqDx da Illumina destina-se à nova sequenciação específica de ADN humano utilizando consumíveis de sequenciação da Illumina e bibliotecas preparadas com ADN genómico humano extraído de sangue total periférico ou tecido FFPE utilizando reagentes de IVD registados e listados, autorizados ou aprovados. As bibliotecas são preparadas através da amplificação dos alvos e adição de índices de amostras e seqüências de captura. As bibliotecas de amostra são capturadas numa célula de fluxo e sequenciadas no instrumento utilizando a química de sequenciação por síntese (SBS). A química SBS utiliza um método reversível-terminator para detetar bases únicas de nucleótidos à medida que são incorporadas em cadeias de ADN crescentes. O software RTA (Real-Time Analysis) executa uma análise da imagem e uma identificação de bases, e atribui um índice de qualidade a cada base de cada ciclo de sequenciação. Quando a análise primária terminar, a análise secundária no instrumento MiSeqDx processa identificações de bases.

O processamento inclui caracteristicamente a demultiplexação, a geração de ficheiros FASTQ, o alinhamento, a identificação de variantes e a geração de ficheiros em formatos de identificação de variantes (VCF) que contêm informações sobre variantes encontrados em posições específicas num genoma de referência. O MiSeqDx utiliza diferentes módulos de análise secundária consoante o fluxo de trabalho.

Configuração Dual Boot

A configuração Dual Boot inclui o hardware, o software e os procedimentos de instalação para permitir que o instrumento MiSeqDx execute ensaios de sequenciação de diagnóstico *in vitro* (IVD) e apenas para efeitos de investigação (RUO). A configuração Dual Boot permite ao utilizador alternar entre o

modo de diagnóstico do instrumento e o modo de investigação do instrumento. As identificações de radiofrequência (RFID, radio-frequency identification) nos consumíveis de sequenciação impedem a utilização dos reagentes de sequenciação RUO em ensaios de sequenciação de diagnóstico.

Limitações do procedimento

- Para utilização em diagnóstico *in vitro*.
- Os resultados apresentados nas etiquetas foram obtidos com painéis de ensaios representativos com sangue total periférico ou linhas celulares para desempenho de linha germinal e tecido FFPE ou linhas celulares FFPE para desempenho somático com os reagentes e módulos de software descritos. Os módulos Germline Variant e Somatic Variant foram desenvolvidos com o objetivo de avaliar o desempenho com ensaios representativos. As características de desempenho são fornecidas apenas para efeitos informativos. A testagem de validação apresentada serve apenas para exemplificar as capacidades gerais do instrumento e não estabelece as capacidades ou a adequação do instrumento relativamente a quaisquer reivindicações específicas. Todos os testes de diagnóstico desenvolvidos para utilização neste instrumento exigem validação completa para todos os aspetos do desempenho.
- Este produto é limitado ao fornecimento do seguinte:
 - Saída de sequenciação ≥ 5 Gb a um comprimento de leitura de 2 x 150 bp
 - Lê os filtros de passagem ≥ 15 milhões a um comprimento de leitura de 2 x 150 bp
 - Bases iguais ou superiores a Q30 $\geq 80\%$ num tamanho de leitura de 2 x 150 bp
80% ou mais bases têm índices de qualidade na escala Phred superiores a 30, indicando uma precisão da identificação de bases superior a 99,9%.
- O instrumento MiSeqDx só foi validado para sequenciar bibliotecas de ADN humano extraído de sangue total periférico ou tecido FFPE. As bibliotecas geradas a partir de outros tipos de amostras não devem ser utilizadas com este instrumento para utilização em diagnóstico *in vitro*. O desempenho deste instrumento para a sequenciação de ácidos nucleicos microbianos ou virais de amostras clínicas não foi estabelecido.
- O MiSeqDx destina-se à utilização em diagnóstico *in vitro* com ensaios ou reagentes de IVD registados e listados, autorizados ou aprovados. As características de desempenho e as limitações de reagentes descritas neste folheto informativo têm por base ensaios representativos e módulos de software. Para ensaios IVD, consulte no folheto informativo específico do ensaio para a utilização pretendida, as variantes detetadas e o tipo de amostra.
- O conteúdo indel (inserções, eliminações e combinações) com tamanho superior a 25 bp não é alinhado pelo Assay Software. Consequentemente, indels com tamanho superior a 25 bp não são detetáveis pelo Assay Software.

- O sistema foi validado para a detecção de variantes únicas de nucleótido (single nucleotide variants, SNV) e eliminações até 25 bp e inserções até 24 bp quando utilizado com o módulo de software Somatic Variant Module e Germline. Para a identificação com o Somatic, a uma frequência de variantes de 0,05, foram detetadas eliminações de 25 bp e inserções de 18 bp.
- As leituras de fragmentos amplificados com conteúdo extremo de variantes podem não ser alinhadas pelo Assay Software, o que pode fazer com que uma região seja reportada como selvagem. O conteúdo extremo inclui:
 - Leituras que contenham mais de três indels.
 - Leituras com um tamanho de, pelo menos, 30 bp com conteúdo de SNV superior a 4% do tamanho total do alvo de fragmento amplificado (excluindo regiões de investigação).
 - Leituras com tamanho inferior a 30 bp com conteúdo SNV superior a 10% do tamanho total de fragmento amplificado (incluindo regiões de investigação).
- Variantes maiores, incluindo variantes multinucleótidos (MNV) e indels grandes, podem ser reportadas como variantes separadas mais pequenas no ficheiro de saída VCF.
- É possível ignorar ou filtrar variantes de eliminação ao cruzar dois fragmentos amplificados cobertos se o tamanho da eliminação for superior ou igual à sobreposição entre os fragmentos amplificados cobertos.
- 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 (2) 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.
- Tal como com qualquer fluxo de trabalho de preparação de bibliotecas à base de hibridização, os polimorfismos, mutações, inserções ou eliminações subjacentes em regiões de ligação oligonucleótida podem afetar os alelos investigados. Consequentemente, as identificações obtidas durante a sequenciação também são afetadas. Por exemplo:
 - Uma variante na 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 indel na região do primer podem causar um falso positivo no fim da leitura adjacente ao primer.
- 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.
- MNV pequenas não foram validadas.

- Variações do número de cópias ou variantes estruturais, como fusões ou translocações, não foram validadas.
- Limitações específicas do Germline:
 - O sistema MiSeqDx, utilizando o módulo Germline Variant, foi concebido para fornecer resultados qualitativos para a identificação da variante da linha germinal (p. ex., homozigótica, heterozigótica, selvagem).
 - Quando utilizado com o módulo Germline Variant, a cobertura mínima por fragmento amplificado necessária para uma identificação de variantes precisa é de 150x. O número de amostras e o número total de bases alvo afetam a cobertura. O conteúdo GC 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 R_Mx_N é utilizado para filtrar variantes se a sequência da variante estiver presente total ou parcialmente repetidamente no genoma de referência adjacente à posição da variante. Para identificação de variantes de linha germinal, são necessárias pelo menos 9 repetições na referência para que uma variante seja filtrada e apenas as repetições com 5 bp são consideradas (R5x9).
- Limitações específicas do Somatic:
 - O sistema MiSeqDx utilizando o módulo Somatic Variant foi concebido para fornecer resultados qualitativos de identificação de variantes somáticas (ou seja, a presença de uma variante somática com uma frequência de variantes superior ou igual a 0,026 num limite de deteção de 0,05).
 - Quando utilizado com o módulo Somatic Variant, a cobertura mínima por fragmento amplificado necessária para uma identificação de variantes precisa é de 450x por pool de oligonucleótidos. O número de amostras e o número total de bases alvo afetam a cobertura. O conteúdo GC e outro conteúdo genómico podem afetar a cobertura.
 - As variantes de alguns contextos repetitivos são filtradas nos ficheiros VCF. O filtro de repetição R_Mx_N é utilizado para filtrar variantes se a sequência da variante estiver presente total ou parcialmente repetidamente no genoma de referência adjacente à posição da variante. Para identificação de variantes somáticas, são necessárias pelo menos 6 repetições na 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 tecido normal.

Componentes do produto

O MiSeqDx da Illumina é composto pelo seguinte:

Instrumento MiSeqDx (N.º de catálogo DX-410-1001)

São necessários os seguintes componentes de software para o funcionamento e a análise de dados do instrumento MiSeqDx:

Aplicação de software	Função	Descrição
MiSeq Operating Software (MOS)	Controla o funcionamento do instrumento	A aplicação de software MOS gere o funcionamento do instrumento durante uma sequenciação e gera imagens para serem utilizadas pelo software RTA (Real-Time Analysis). Para mais informações, consulte o <i>Guia de referência do instrumento MiSeqDx para MOS v4 (documento n.º 200010452)</i> .
Real-Time Analysis (RTA)	Executa análise primária	A aplicação de software RTA converte as imagens geradas pelo MOS 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 de software RTA não contém uma interface de utilizador.
Local Run Manager	Interface para seleção de módulos	O software Local Run Manager é uma solução integrada no instrumento para gestão de utilizadores, que executa a análise secundária e monitoriza o estado. Para mais informações, consulte o <i>Guia de referência do software Local Run Manager v3 para o MiSeqDx (documento n.º 200003931)</i> .

Armazenamento e manuseamento

Elemento	Especificação
Temperatura	Transporte e armazenamento: -10 °C a 40 °C (14 °F a 104 °F) Condições de funcionamento: 19 °C a 25 °C (66 °F a 77 °F)
Humidade	Transporte e armazenamento: humidade sem condensação Condições de funcionamento: 30–75% de humidade relativa (sem condensação)

Equipamento e materiais necessários, mas não fornecidos

Consumíveis de sequenciação

Kit de reagentes MiSeqDx v3 (N.º de catálogo 20037124)

MiSeqDx Reagent Kit v3 Micro (N.º de catálogo 20063860)

Consumíveis fornecidos pelo utilizador

Certifique-se de que estão disponíveis os seguintes consumíveis fornecidos pelo utilizador antes do início de um ensaio.

Consumível	Finalidade
Toalhitas com álcool isopropílico a 70% ou Etanol a 70%	Limpar a plataforma e o vidro da célula de fluxo
Pano de laboratório, libertação reduzida de pelo	Limpar a plataforma da célula de fluxo
Papel para lentes, 4 x 6 pol.	Limpar a célula de fluxo
Tween 20	Lavar o instrumento
Pinças de ponta quadrada de plástico (opcional)	Remover a célula de fluxo do contentor de transporte da célula de fluxo
Água, grau laboratorial	Lavar o instrumento

Diretrizes para água laboratorial

Utilize água laboratorial ou água 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



WARNING (AVISO)

Este conjunto de reagentes contém 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 a ficha de dados de segurança (FDS) em support.illumina.com/sds.html.

- 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).
- O não seguimento dos procedimentos da forma descrita poderá resultar em resultados erróneos ou na redução significativa da qualidade das amostras.
- 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.
- 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.

- 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 centrífuga).
- Sempre que seja adequado, o emparelhamento entre os índices e as 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.
- 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.
- Não utilize o MiSeqDx 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.
- 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.
- O instrumento pesa cerca de 57 kg (126 lb) e pode causar ferimentos graves caso se deixe cair ou caso seja utilizado incorretamente.
- Comunique imediatamente quaisquer incidentes graves relacionados com este produto à Illumina e à Autoridade Competente do Estado-membro no qual o utilizador e/ou o paciente estão estabelecidos.

Instruções de utilização

As instruções que se seguem para a utilização do instrumento MiSeqDx requerem reagentes fornecidos no kit de reagentes MiSeqDx v3.

Criar ensaio com o Local Run Manager

Para obter instruções detalhadas para criar um ensaio, consulte o *Guia de referência do software do Local Run Manager v3 para o MiSeqDx (documento n.º 200003931)* e o guia do módulo do Local Run Manager para o módulo de análise utilizado.

Preparar o cartucho de reagentes

As instruções que se seguem descrevem como descongelar os reagentes utilizando um banho com água à temperatura ambiente.

1. Retire o cartucho de reagentes do armazenamento de -15 °C a -25 °C.

2. Coloque o cartucho de reagentes num banho com água que contenha água desionizada à temperatura ambiente suficiente para mergulhar a base no cartucho de reagentes até à linha de água impressa no cartucho de reagentes. Não permita que a água exceda a linha de água máxima.

Figura 1 Linha de água máxima



3. Deixe o cartucho de reagentes descongelar no banho com água à temperatura ambiente durante aproximadamente 60–90 minutos ou até estar completamente descongelado.
4. Remova o cartucho do banho com água e bata-o delicadamente na bancada para deslocar a água da base do cartucho. Seque a base do cartucho. Certifique-se de que não foi salpicada nenhuma água para a parte superior do cartucho de reagentes.

Inspecionar o cartucho de reagentes

1. Inverta o cartucho de reagentes dez vezes para misturar os reagentes descongelados e, em seguida, verifique se todas as posições estão descongeladas.
2. Inspeccione os reagentes nas posições 1, 2 e 4 para garantir que estão completamente misturados e sem precipitados.

NOTA É essencial que os reagentes no cartucho estejam completamente descongelados e misturados para garantir a sequenciação correta.

3. Bata levemente o cartucho na bancada para reduzir as bolhas de ar nos reagentes.

NOTA Os tubos da unidade de aspiração do MiSeqDx vão para o fundo de cada reservatório para aspirar os reagentes, por isso, é importante que os reservatórios estejam isentos de bolhas de ar.

4. Coloque o cartucho de reagentes em gelo ou coloque de lado a uma temperatura entre 2 °C e 8 °C (até seis horas) até estar pronto para preparar o ensaio. Para obter os melhores resultados, avance diretamente para o carregamento de amostras e configuração do ensaio.

Preparar amostras para sequenciação

Para obter instruções para preparar bibliotecas de amostras para sequenciação, incluindo pooling e diluição de bibliotecas, consulte a secção das Instruções de utilização do folheto informativo de preparação de bibliotecas.

A diluição das bibliotecas de amostras está dependente da complexidade dos pools de oligonucleótidos. É necessária a otimização da densidade dos clusters do MiSeqDx, e a densidade ideal dos clusters varia dependendo do ensaio de preparação de bibliotecas em particular.

Carregar bibliotecas de amostras para o cartucho

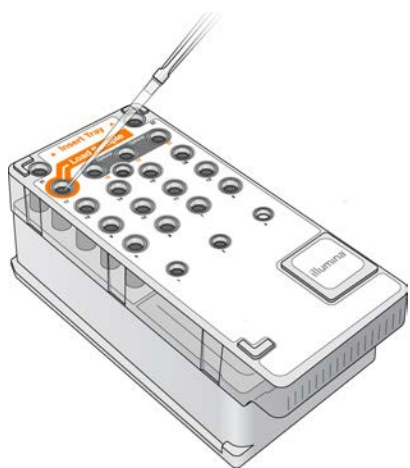
Quando o cartucho de reagentes está totalmente descongelado e pronto para utilizar, o operador pode carregar amostras no cartucho.

1. Utilize uma ponta de pipeta de 1 ml separada, limpa e vazia para perfurar o selo de alumínio sobre o reservatório no cartucho de reagentes identificado com **Load Samples** (Carregar amostras).

NOTA Não perfure qualquer outras posições de reagentes. As outras posições de reagentes são perfuradas automaticamente durante o ensaio.

2. Efetue a pipetagem de 600µl de amostra preparada da biblioteca de amplicões diluídos (DAL, Diluted Amplicon Library) no reservatório **Load Samples** (Carregar amostras). Evite tocar no selo de alumínio.
3. Verifique se estão visíveis bolhas de ar no reservatório depois de carregar a amostra. Se estiverem presentes bolhas de ar, bata delicadamente no cartucho no banco para soltar as bolhas.

Figura 2 Carregar bibliotecas



4. Prossiga diretamente com os passos de configuração do ensaio utilizando a interface do Software operativo MiSeq (MOS).

Configuração do ensaio

Consulte o *Guia de referência do instrumento MiSeqDx para MOS v4 (documento n.º 200010452)* para obter as instruções completas para a configuração do ensaio.

1. Inicie sessão no MiSeqDx com a palavra-passe do software Local Run Manager.
2. No ecrã Home (Início) do software MOS, 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 Reagents (Carregar reagentes), Review (Rever) 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.
5. Feche o trinco da célula de fluxo e a porta do compartimento da célula de fluxo.
Tanto o trinco como a porta do compartimento devem ser fechados antes do início do ensaio. Quando a célula de fluxo estiver carregada, o software lê e regista o RFID. Aparece no canto inferior direito do ecrã uma confirmação de que a RFID foi lida com sucesso.
6. Siga as indicações do software para carregar o frasco de Solução SBS MiSeqDx (PR2), garanta que o frasco de resíduos está vazio e carregue o cartucho de reagentes.
Quando o frasco e o cartucho de reagente da Solução SBS MiSeqDx (PR2) são carregados, o software lê e regista o RFID. Aparece no canto inferior direito do ecrã uma confirmação de que a RFID foi lida com sucesso.
7. 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 um índice de qualidade a cada base de cada ciclo de sequenciação. Quando a análise principal terminar, o módulo no instrumento MiSeqDx selecionado em [Criar ensaio com o Local Run Manager na página 8](#) inicia a análise secundária. Consulte a documentação específica do ensaio relativamente a outros fluxos de trabalho.

Procedimentos de controlo de qualidade

O software MiSeqDx avalia cada ensaio, amostra e identificação de bases em relação às métricas de controlo de qualidade. Sempre que tal seja exigido, os controlos positivo e negativo incluídos na preparação de bibliotecas também devem ser avaliados para os resultados esperados.

Caraterísticas de desempenho

Todos os estudos foram realizados no MiSeqDx.

Os estudos de linha germinal utilizaram os reagentes do ensaio de fibrose cística de 139 variantes do MiSeqDx ou o kit TruSeq Custom Amplicon para a preparação de bibliotecas. Os dois kits utilizam reagentes de preparação de bibliotecas idênticos e têm apenas uma diferença no fluxo de trabalho: o número de ciclos de reação em cadeia da polimerase (PCR) (25 e 28, respetivamente). Os ciclos de PCR permitem uma entrada de ADN mais baixa com o kit TruSeq Custom Amplicon Dx (50 ng) relativamente ao ensaio de fibrose cística de 139 variantes do MiSeqDx (250 ng), conforme demonstrado no estudo de entrada de ADN utilizando o kit TruSeq Custom Amplicon Dx. As bibliotecas preparadas com o ensaio de fibrose cística de 139 variantes do MiSeqDx foram sequenciadas com os reagentes de sequenciação incluídos no kit. As bibliotecas preparadas com o kit TruSeq Custom Amplicon Dx foram sequenciadas com o kit de reagentes MiSeqDx v3. Estes últimos reagentes de sequenciação aumentaram a saída relativamente aos do ensaio de fibrose cística de 139 variantes do MiSeqDx.

O teste engloba os intervalos de produção de amostras suportados pelo MiSeqDx Reagent Kit v3 Micro. O MiSeqDx pode suportar 1 a 96 amostras por ensaio, dependendo do ensaio. O MiSeqDx Reagent Kit v3 Micro foi concebido para suportar produções de amostras mais baixas dentro deste intervalo para os ensaios selecionados.

Os estudos somáticos utilizaram o kit TruSeq Custom Amplicon Dx com o kit de reagentes MiSeqDx v3. Os fluxos de trabalho somático ou de linha germinal, descritos para a preparação de bibliotecas do kit TruSeq Custom Amplicon Dx para sequenciação, foram seguidos de análise com o módulo Germline Variant ou o módulo Somatic Variant, respetivamente, com duas exceções. Os estudos que utilizam um gene (desempenho de linha germinal; ensaio de fibrose cística de 139 variantes do MiSeqDx) ou dois genes (desempenho somático) como painéis de mutação representativos utilizaram fluxos de trabalho e módulos de análise específicos do ensaio.

NOTA O conteúdo genómico dos fragmentos amplificados é resumido relativamente à cadeia genómica sequenciada. Para fragmentos amplificados concebidos contra a cadeia negativa, o conteúdo de genoma de referência é o complemento inverso (por exemplo, regiões PolyA nos fragmentos amplificados de cadeia negativa correspondem a regiões PolyT no genoma de referência).

Definições dos cálculos utilizados nas características de desempenho

- A PPA (percentagem de concordância positiva) é calculada como a proporção de loci classificados como variantes por um método de referência reportados corretamente pelo ensaio.
 - $(N.^{\circ} \text{ de loci variantes corretamente reportadas pelo ensaio}) / (n.^{\circ} \text{ total de loci variantes})$
As loci variantes reportadas pelo ensaio que são concordantes com o método de referência são verdadeiros positivos (TP). Os loci variantes reportados como identificações de referência ou identificações de variantes diferentes pelo ensaio são falsos negativos (FN)
- A NPA (percentagem de concordância negativa) é calculada como a proporção de loci classificados como tipo selvagem por um método de referência reportados corretamente pelo ensaio.
 - $(N.^{\circ} \text{ loci de tipo selvagem corretamente reportadas pelo ensaio}) / (n.^{\circ} \text{ total de loci de tipo selvagem})$
As loci de tipo selvagem reportadas pelo ensaio que são concordantes com o método de referência são verdadeiros negativos (TN). Os loci de tipo selvagem reportados como variantes pelo ensaio são falsos positivos (FP).
- A OPA (percentagem de concordância geral) é calculada como a proporção de loci corretamente pelo ensaio relativo a um método de referência.
 - $([N.^{\circ} \text{ de loci variantes corretamente reportadas pelo ensaio}] + [n.^{\circ} \text{ de loci de tipo selvagem corretamente reportadas pelo ensaio}]) / ([n.^{\circ} \text{ total de loci variantes}] + [n.^{\circ} \text{ total de loci de tipo selvagem}])$
- Para aplicações de identificação de variantes, os cálculos de PPA, NPA e OPA não incluem quaisquer identificações (loci variante ou de referência que não cumpra um ou mais filtros de qualidade). Dois estudos incluíram especificamente as não-identificações na métrica "% de identificações corretas", e esta inclusão de não-identificações é registada para as tabelas aplicáveis.
- A taxa de identificação é calculada como o número total de loci que cumprem os filtros de passagem 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.

Para saber quais as características de desempenho relacionadas com os fatores pré-analíticos (por ex. métodos de extração ou entrada de ADN), consulte o folheto informativo para o método de preparação de bibliotecas aplicável.

Indexação de amostras

Os primers de indexação de amostras, adicionados durante a preparação de bibliotecas, atribuem uma sequência única a cada amostra de ADN, permitindo o pool de várias amostras em conjunto num só ensaio de sequenciação. A indexação de amostras foi testada em fluxos de trabalho somáticos e de

linha germinal.

Foi testado um total de 96 índices de amostras 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. O cromossoma Y não contém regiões de confiança e não foi avaliado. Foram testadas oito amostras únicas com 12 combinações diferentes do primer de indexação por amostra. Os resultados da amostra do módulo de análise Germline foram comparados com os Platinum Genomes versão 2016-01. A PPA (SNV e indels) excedeu 97% (as identificações positivas reais foram pelo menos 70 para SNV, 38 para inserções, 36 para eliminações) e a NPA foi 100% (pelo menos 23.440 posições de referência por combinação de índices) para cada uma das 96 combinações de índices. Independentemente, foi testado um único índice para verificar se a química de sequenciação do kit de reagentes MiSeqDx v3 pode suportar menos de oito amostras (a química precursora no kit universal MiSeqDx 1.0 estava limitada a um mínimo de oito amostras). O índice único apresentava valores PPA de 98,9% (180/182) para SNV, 100% (38/38) para inserções e 100% (46/46) para eliminações. A NPA foi de 100% (23.856/23.856).

Foram testadas doze réplicas (24 bibliotecas) de uma amostra para medir a precisão do índice com variantes somáticas em frequências de 0,05-0,10 utilizando o módulo Somatic Variant (são utilizadas duas combinações de índices por réplica para fazer identificações somáticas). A PPA foi de 100% para SNV (64/64), inserções (11/11) e eliminações (19/19). A NPA foi de 100% (pelo menos 11590 posições de referência por combinação de índices) para todas as combinações de índices.

Contaminação cruzada de amostras

O fluxo de trabalho do instrumento MiSeqDx implica a preparação de bibliotecas e a sequenciação de várias amostras e controlos processados ao mesmo tempo. O estudo de contaminação cruzada de amostras foi realizado para avaliar se os resultados positivos falsos, devido à contaminação cruzada de poço para poço, bem como à contaminação de ensaio para ensaio entre ensaios de sequenciação consecutivos, têm impacto nos resultados dos testes. Foram utilizadas variantes somáticas, pois estas podem ser detetadas em eventos de mais baixa frequência de alelos que as variantes de linha germinal.

As amostras consistiram em quatro amostras de ADN genómico de linhas celulares, contendo cada uma mutações de painel diferentes num ensaio representativo de dois genes. Uma mutação na posição de uma amostra terá uma sequência de referência (tipo selvagem) nas outras.

A contaminação cruzada de poço para poço foi definida como modo de falha potencialmente criado por passos de processamento manuais (pipetagem, troca de amostras, etc.). Para avaliar a contaminação cruzada de um poço de amostra para outro, foram efetuados dois ensaios de teste:

- Uma disposição quadriculada de uma amostra de ADN genómico (gDNA) de entrada elevada com uma mutação no Gene 1 em alternância com uma amostra de gDNA de entrada baixa com uma mutação no Gene 2.

- Uma disposição quadriculada de uma amostra de gDNA de entrada elevada com uma mutação no Gene 2 em alternância com uma amostra de gDNA de entrada baixa com uma mutação no Gene 1.

Em cada ensaio foram avaliadas no total 12 réplicas para falsos positivos (por ex., uma mutação no Gene 1 comunicada num poço designado como amostra de mutação no Gene 2 ou vice-versa).

A contaminação cruzada de ensaio para ensaio é definida como modo de falha potencialmente criado pelo resíduo de um ensaio de sequenciação anterior. Para poder determinar se existe contaminação cruzada entre ensaios de sequenciação, duas placas contendo cada uma 11 réplicas de uma só amostra única de gDNA de entrada elevada mais uma amostra vazia foram preparadas e sequenciadas consecutivamente num instrumento MiSeqDx e avaliadas para falsos positivos.

O primeiro ensaio continha 11 réplicas de uma amostra com mutação no Gene 2 mais uma vazia.

O segundo ensaio continha 11 réplicas de uma amostra com mutação no Gene 1 mais 1 vazia. Primeiro, a biblioteca de amostras com mutação no Gene 2 foi sequenciada, seguindo-se um ensaio de sequenciação com a biblioteca de amostras com mutação no Gene 1, e posteriormente uma repetição do ensaio de sequenciação da biblioteca de amostras com mutação no Gene 2. Se forem observadas mutações no Gene 2 num ensaio único para mutações no Gene 1, e vice-versa, isso indica contaminação cruzada.

Foram comunicados zero falsos positivos (0/24, 0%) devido a contaminação cruzada *de poço para poço*. Foram detetadas todas as mutações esperadas. Foram comunicados zero falsos positivos (0/24, 0%) devido a contaminação cruzada *de ensaio para ensaio*. Foram detetadas todas as mutações esperadas. Foram comunicados zero falsos positivos (0/48, 0%) devido a contaminação cruzada *total* (de poço para poço e de ensaio para ensaio).

Caraterísticas de desempenho da linha germinal

Os estudos aqui descritos utilizaram o módulo Germline Variant para analisar a sequenciação de dados, exceto os que utilizaram um painel de um gene quando foi utilizado um módulo específico do ensaio.

Precisão

O estudo que se segue foi conduzido para avaliar a precisão do instrumento MiSeqDx com o kit de reagentes MiSeqDx v3 e ADN de elevada qualidade. O estudo utilizou um ensaio representativo concebido para analisar uma variedade de genes abrangendo 12.588 bases ao longo de 23 cromossomas diferentes utilizando 150 fragmentos amplificados. O cromossoma Y não contém regiões de confiança e não foi avaliado. As 12 amostras únicas utilizadas neste estudo veem de uma só família – dois pais e dez filhos – frequentemente sequenciada por vários laboratórios e metodologias de sequenciação. Existem cinco amostras femininas e sete masculinas. Cada amostra foi testada em duplicado. A precisão foi determinada para SNV, inserções e eliminações comparando os dados do estudo com uma base de dados de referência bem caracterizada. A sequência da base de dados de referência (Platinum Genomes versão 2016-01) foi derivada da combinação de várias metodologias de sequenciação, dados disponíveis publicamente e informações hereditárias. Foram definidas regiões genómicas de confiança com base neste método de referência, exceto quando especificado em contrário. As amostras foram submetidas a ensaio oito vezes no total; as tabelas apresentadas para demonstrar a precisão têm por base os dados do primeiro ensaio.

A [Tabela 1](#) 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 (SNV, 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 1 Concordância dos resultados da identificação de bases do instrumento MiSeqDx por amostra

Amostra	Taxa média de identificações	N.º total de variantes	N.º total de variantes TP	N.º total de variantes FN	N.º total de não-identificações	N.º total de identificações TN	PPA	NPA	OPA
NA12877	> 99,9	152	152	0	4	24024	100	100	100
NA12878	> 99,9	270	266	0	4	23856	100	100	100

Amostra	Taxa média de identificações	N.º total de variantes	N.º total de variantes TP	N.º total de variantes FN	N.º total de não-identificações	N.º total de identificações TN	PPA	NPA	OPA
NA12879	> 99,9	192	190	1	1	24054	99,5	100	> 99,9
NA12880	> 99,9	222	220	0	6	24052	100	100	100
NA12881	> 99,9	250	247	1	2	23862	99,6	100	> 99,9
NA12882	> 99,9	200	196	2	2	23962	99,0	100	> 99,9
NA12883	> 99,9	226	224	0	6	23870	100	100	100
NA12884	> 99,9	228	226	1	1	23942	99,6	100	> 99,9
NA12885	> 99,9	244	240	2	2	23942	99,2	100	> 99,9
NA12886	> 99,9	230	228	1	1	23888	99,6	100	> 99,9
NA12888	> 99,9	216	216	0	4	24002	100	100	100
NA12893	> 99,9	236	234	0	2	23810	100	100	100

O ensaio representativo era composto por 150 fragmentos amplificados concebidos para abranger uma variedade de conteúdo genómico. O conteúdo GC dos fragmentos amplificados encontrava-se entre 26 e 87%. Os fragmentos amplificados também tinham uma gama de nucleótido único (p. ex. PolyA, PolyT), repetições de dinucleótidos e trinucleótidos. Os dados foram compilados numa base por fragmento amplificado ([Tabela 2](#)) para determinar o efeito do conteúdo genómico na % de identificações corretas. A % de identificações corretas é composta por identificações de referência e variantes e será inferior a 100% se ocorrerem identificações incorretas ou não-identificações. Não ocorrem identificações quando não estão presentes um ou mais filtros para identificação de variantes (por ex. cobertura insuficiente).

Das oito variantes FN da [Tabela 2](#), sete ocorreram com uma inserção 1 bp no fragmento amplificado 111, que também contém homopolímero PolyA e conteúdo GC de 0,29. A FN restante (identificação incorreta) deveu-se a um SNV heterozigótico esperado, no fragmento amplificado 125 com um conteúdo GC de 0,68, identificado como variante homozigótica. A frequência da variante SNV foi de 0,71, acima do limiar de 0,70 para classificação como variante homozigótica. O fragmento amplificado com a % de identificações corretas mais baixa (98,2%) foi o fragmento amplificado 17 com 40 não-identificações e contendo repetições AT e conteúdo GC de 27%.

Tabela 2 Precisão ao nível do fragmento amplificado para o instrumento MiSeqDx

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
1	1	36450499	36450591	93	93	Indel	0,22	2232	0	0	100
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	1896	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	2184	0	0	100
4	1	223906657	223906748	92	92	Indel	0,49	2208	0	0	100
5	1	228526602	228526682	81	81	PolyG (5)	0,69	1944	0	0	100
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	1680	0	0	100
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA (3), indel	0,27	2112	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	2160	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	1920	0	0	100
10	2	177016721	177016805	85	81	N/D	0,65	1944	0	0	100
11	2	186625727	186625801	75	75	PolyA (8)	0,35	1800	0	0	100
12	2	190323504	190323591	88	88	PolyT (5)	0,42	2112	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	2088	0	0	100
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	2184	0	0	100
15	2	228147052	228147144	93	93	N/D	0,43	2232	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	1752	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	2192	0	40	98,2
18	3	46620561	46620643	83	83	N/D	0,43	1992	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	1680	0	0	100
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0,41	2112	0	0	100
21	3	190106030	190106104	75	74	Indel	0,57	1774	0	2	99,9
22	4	2233667	2233744	78	78	PolyA (6)	0,26	1872	0	0	100
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	2328	0	0	100
24	4	15688604	15688681	78	78	N/D	0,29	1872	0	0	100

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	1488	0	0	100
26	4	102839244	102839314	71	69	PolyA (5)	0,46	1656	0	0	100
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	1488	0	0	100
28	5	1882081	1882158	78	75	N/D	0,78	1800	0	0	100
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	2016	0	0	100
30	5	41069808	41069871	64	64	N/D	0,39	1536	0	0	100
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	1992	0	0	100
32	5	147475343	147475409	67	67	PolyT (5)	0,37	1608	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	2184	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	1800	0	0	100
35	6	6318713	6318814	10	10	PolyG (6)	0,68	2448	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	2208	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	2244	0	12	99,5
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT (3)	0,55	2352	0	0	100
39	6	32986864	32986958	95	95	Indel	0,53	2280	0	0	100
40	6	33408498	33408583	86	86	PolyC (6)	0,7	2064	0	0	100
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	2256	0	0	100
42	6	112435865	112435955	91	91	PolyA (5)	0,44	2184	0	0	100
43	7	22202076	22202148	73	73	N/D	0,44	1752	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	2112	0	0	100
45	7	77365735	77365821	87	87	PolyA (7), AG(4)	0,26	2088	0	0	100
46	7	110939946	110940030	85	85	Indel	0,38	2040	0	0	100
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	2160	0	0	100
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	2184	0	0	100
49	7	154404519	154404599	81	66	N/D	0,31	1584	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	2232	0	0	100
51	8	1817312	1817394	83	83	N/D	0,42	1992	0	0	100

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0,61	2113	0	23	98,9
53	8	76518625	76518691	67	67	Indel	0,3	1608	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0,67	2352	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	1560	0	0	100
56	9	107620823	107620918	96	96	N/D	0,49	2304	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	1992	0	0	100
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	2328	0	0	100
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	1872	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	2184	0	0	100
61	10	27317777	27317855	79	79	PolyT (5)	0,3	1896	0	0	100
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	2160	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	2280	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	2102	0	10	99,5
65	10	101611250	101611329	80	80	N/D	0,49	1920	0	0	100
66	10	118351373	118351453	81	81	N/D	0,51	1944	0	0	100
67	11	8159816	8159912	97	96	N/D	0,45	2304	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	1680	0	0	100
69	11	47470345	47470444	100	100	N/D	0,65	2400	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	1488	0	0	100
71	11	64418856	64418957	102	102	N/D	0,59	2448	0	0	100
72	11	93529612	93529684	73	73	PolyA (5)	0,4	1752	0	0	100
73	11	101347052	101347136	85	85	N/D	0,42	2040	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0,55	2184	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	2040	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0,34	2040	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	2040	0	0	100
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	2016	0	0	100
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0,33	2232	0	0	100

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genômico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
80	12	30881766	30881846	81	81	N/D	0,49	1944	0	0	100
81	12	88474105	88474175	71	71	PolyA (6)	0,35	1704	0	0	100
82	12	120966872	120966966	95	95	PolyG (5)	0,68	2280	0	0	100
83	13	24167504	24167576	73	73	N/D	0,52	1752	0	0	100
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	2112	0	0	100
85	13	44880112	44880200	89	89	Indel	0,49	2136	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	1848	0	0	100
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	1608	0	0	100
88	14	39517884	39517966	83	83	N/D	0,25	1992	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	1727	0	1	99,9
90	14	58050030	58050110	81	81	Indel	0,38	1944	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	2184	0	0	100
92	14	92549544	92549609	66	66	PolyA (5)	0,41	1584	0	0	100
93	14	102808496	102808589	94	94	Indel	0,62	2256	0	0	100
94	15	43170751	43170848	98	96	PolyC (5)	0,45	2304	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	1632	0	0	100
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	2232	0	0	100
97	15	81625334	81625428	95	95	PolyT (6)	0,43	2280	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	1704	0	0	100
99	15	89817413	89817503	91	91	N/D	0,36	2184	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	1680	0	0	100
101	16	1894910	1894972	63	63	N/D	0,27	1512	0	0	100
102	16	28997904	28997998	95	95	PolyC (5)	0,67	2280	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0,41	2088	0	0	100
104	16	57954406	57954509	104	104	PolyC (5)	0,67	2496	0	0	100
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	2184	0	0	100
106	17	3563920	3564008	89	89	GC(3)	0,64	2136	0	0	100
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	2088	0	0	100

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genômico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
108	17	3970090	3970180	91	91	Indel	0,46	2184	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	2232	0	0	100
110	17	33998759	33998849	91	89	PolyT (5)	0,54	2136	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	1944	7	17	98,8
112	17	41244394	41244484	91	91	PolyA (5)	0,34	2184	0	0	100
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT(4), AT(4), indel	0,26	2208	0	0	100
114	17	61502432	61502510	79	79	Indel	0,41	1887	0	9	99,5
115	17	64023582	64023667	86	86	PolyT (7)	0,22	2064	0	0	100
116	17	72308237	72308320	84	84	GAG(3)	0,62	2016	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	1608	0	0	100
118	18	6980478	6980568	91	91	N/D	0,37	2184	0	0	100
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0,43	1656	0	0	100
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	1800	0	0	100
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	1944	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0,45	2040	0	0	100
123	18	59773996	59774060	65	65	N/D	0,48	1560	0	0	100
124	19	625143	625241	99	99	N/D	0,59	2376	0	0	100
125	19	18121418	18121491	74	74	N/D	0,68	1775	1	0	99,9
126	19	18186574	18186643	70	70	N/D	0,64	1680	0	0	100
127	20	746056	746149	94	94	N/D	0,61	2256	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	1968	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	1824	0	0	100
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0,46	1680	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	2424	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	1608	0	0	100
133	20	62331904	62331994	91	88	PolyG (6)	0,73	2112	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	2088	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	1584	0	0	100

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genômico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0,54	2352	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	2088	0	0	100
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0,32	1656	0	0	100
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	2160	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	2400	0	0	100
141	22	32439233	32439329	97	97	N/D	0,68	2328	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	2328	0	0	100
143	22	37637596	37637694	99	99	N/D	0,6	2376	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	2208	0	0	100
145	X	15870424	15870492	69	69	PolyT (5)	0,26	1656	0	0	100
146	X	135288543	135288611	69	69	PolyC (5)	0,62	1656	0	0	100
147	X	135290777	135290847	71	71	N/D	0,52	1704	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	PolyA (5)	0,37	0	0	0	N/D

As variantes que constituíram não-identificações estão resumidas na [Tabela 3](#). Os filtros que resultaram nas não-identificações estão enumerados na tabela. A inserção no fragmento amplificado 111 foi filtrada em nove das 16 ocorrências, com as restantes sete ocorrências identificadas como referência, e por conseguinte FN.

Tabela 3 Resumo de não-identificações de variantes

N.º de fragmentos amplificados	Cro:Pos	Variante	Conteúdo de fragmento amplificado correspondente	Filtro	Variantes ignoradas	Variantes esperadas	Identificações FN
64	10:55892600	TAC > T	AC(11), 42% GC	R5x9 ¹	10	10	0
111	17:39589692	C > CA	PolyA (13), 29% GC	R5x9	9	16	7

¹ R5x9: repetir filtro. Uma variante é filtrada se a variante estiver presente total ou parcialmente de forma repetida no genoma de referência adjacente à posição da variante. São necessárias, pelo menos, nove repetições na referência e apenas as repetições com 5 bp são consideradas.

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 bases não-variantes com o genoma humano de referência do conjunto 19. Os resultados de precisão são apresentados na [Tabela 4](#).

Tabela 4 Concordância dos resultados da identificação de bases do instrumento MiSeqDx para a amostra NA12878 com a base de dados do NIST

Amostra	N.º de fragmentos amplificados	Taxa média de identificações	N.º total de variantes TP	N.º total de variantes FN	N.º total de identificações TN	N.º total de identificações FP	PPA	NPA	OPA
NA12878	133	99,98	208	0	19380	0	100	100	100

As amostras foram analisadas novamente para identificar pequenas inserções e eliminações (indels) ([Tabela 5](#)). Em alguns casos, o indel era comum entre duas ou mais amostras, conforme refletido na coluna N.º total de réplicas de amostras com indel. Os resultados de ambas as réplicas das 12 amostras válidas estão incluídas na [Tabela 5](#). Foram identificados 71 indels no total com tamanho entre 1–24 bp no caso das inserções e 1–25 bp no caso das eliminações. Foram detetados 69 indels cada um com uma percentagem de concordância positiva de 100%. Uma eliminação (fragmento amplificado 64; eliminação de 2 bp (chr10 55892600 TAC>T) não apresentou identificações corretas porque cada uma destas variantes era uma não-identificação devido ao filtro R5x9. Assim, não foi possível calcular o PPA, que exclui as não-identificações. Outro indel, inserção 1 bp (chr17 39589692 C>CA no fragmento amplificado 111), também não teve identificações corretas, porque nove variantes foram não-identificações devido ao filtro R5x9 e sete foram identificações FN.

Tabela 5 Resumo da deteção de indel com o instrumento MiSeqDx

Amplicon	Cromossoma	Posição	Tamanho do fragmento analisado	Tipo e tamanho do indel do fragmento amplificado	Indel	N.º total de réplicas de amostras com indel	N.º de não-identificações	N.º total de identificações de indel incorretas	N.º total de identificações de indel corretas	PPA
1	1	36450544	93	Eliminação 25 bp	GAAAATTTAATGAAACACATTGTCCT>G	2	0	0	2	100
2	1	109465165	79	Eliminação 3 bp	ACTT>A	12	0	0	12	100

Amplicon	Cromossoma	Posição	Tamanho do fragmento analisado	Tipo e tamanho do indel do fragmento amplificado	Indel	N.º total de réplicas de amostras com indel	N.º de não-identificações	N.º total de identificações de indel incorretas	N.º total de identificações de indel corretas	PPA
3	1	218353908	91	Inserção 23 bp	T>TTTTAATAGCAAAAAGAGGCTAGA	24	0	0	24	100
4	1	223906701	92	Eliminação 17 bp	GACAGACTGTGAGGAAGA>G	10	0	0	10	100
6	1	236372081	70	Inserção 5 bp	C>CTTAAG	10	0	0	10	100
7	1	247812083	88	Inserção 3 bp	C>CATG	10	0	0	10	100
8	2	55862804	90	Inserção 7 bp	T>TTTGGTAA	14	0	0	14	100
9	2	87003972	80	Eliminação 6 bp	TTATCTC>T	6	0	0	6	100
13	2	200796749	87	Inserção 5 bp	T>TTAAAA	24	0	0	24	100
14	2	212245090	91	Inserção 12 bp	C>CTGAAAATAGGAT	14	0	0	14	100
16	2	235016388	73	Inserção 2 bp	A>ATG	12	0	0	12	100
17	3	4466274	93	Eliminação 23 bp	TAACCTAAAATTACAAAATAACCC>T	2	0	0	2	100
19	3	49851375	70	Inserção 9 bp	C>CCTGGCTCCT	4	0	0	4	100
21	3	190106071	75	Eliminação 1 bp	AG>A	20	0	0	20	100
25	4	56236567	66	Eliminação 8 bp	TAACCGAAA>T	12	0	0	12	100
27	4	164446785	62	Inserção 11 bp	T>TTATGGTATTGA	12	0	0	12	100
31	5	74077155	83	Eliminação 4 bp	TAGTA>T	10	0	0	10	100
34	5	155662255	75	Inserção 8 bp	G>GCCTACTGA	20	0	0	20	100
36	6	24950035	92	Eliminação 21 bp	CCCTGGGTGCTATAGCCACCA>C	10	0	0	10	100
37	6	31084942	100	Eliminação 3 bp	GCTT>G	14	0	0	14	100
39	6	32986905	95	Eliminação 25 bp	CTTTCACCTTCCCGTCTCATGCAAAG>C	12	0	0	12	100
41	6	41647442	95	Eliminação 23 bp	GGCATGAGGCTTGGTGACATGGCA>G	8	0	0	8	100
44	7	66276142	88	Inserção 1 bp	C>CT	16	0	0	16	100
46	7	110939983	85	Eliminação 4 bp	CAAGT>C	12	0	0	12	100
47	7	128533514	90	Inserção 1 bp	T>TC	24	0	0	24	100
48	7	149503916	91	Eliminação 4 bp	GGATA>G	8	0	0	8	100
50	7	156476548	93	Eliminação 11 bp	GAATCTGCACTT>G	12	0	0	12	100
52	8	24811064	90	Eliminação 1 bp	AG>A	24	0	0	24	100
53	8	76518677	67	Inserção 4 bp	T>TACTG	14	0	0	14	100
55	9	105586193	65	Inserção 4 bp	C>CAATT	2	0	0	2	100

Amplicon	Cromossoma	Posição	Tamanho do fragmento analisado	Tipo e tamanho do indel do fragmento amplificado	Indel	N.º total de réplicas de amostras com indel	N.º de não-identificações	N.º total de identificações de indel incorretas	N.º total de identificações de indel corretas	PPA
58	9	138995370	97	Eliminação 21 bp	TCTGGGGGGCAGCCCCTGAGGG>T	14	0	0	14	100
59	10	5987158	79	Eliminação 3 bp	TAAC>T	10	0	0	10	100
63	10	45084202	95	Eliminação 16 bp	AGCGTCTATAACCAAAT>A	12	0	0	12	100
64	10	55892600	89	Eliminação 2 bp	TAC>T	10	10	0	0	N/D
68	11	30177690	70	Inserção 2 bp	C>CTG	10	0	0	10	100
70	11	59837721	62	Inserção 8 bp	T>TTATGAAAA	12	0	0	12	100
75	11	118406328	85	Eliminação 8 bp	CAGTGTGGA>C	10	0	0	10	100
76	11	120357842	85	Eliminação 2 bp	CTT>C	10	0	0	10	100
78	12	2834814	84	Inserção 21 bp	T>TTCTCAGTACGGTGAACCCCAG	24	0	0	24	100
84	13	25817002	89	Inserção 19 bp	C>CAAAATATAAAAAGCTCCCT	24	0	0	24	100
85	13	44880152	89	Inserção 4 bp	C>CCTGT	12	0	0	12	100
86	13	77665265	77	Eliminação 20 bp	ATCTATTTTCTAATAGACGGC>A	14	0	0	14	100
89	14	46958967	73	Eliminação 22 bp	TTTAAATTTGAATGTGATAAAA>T	24	0	0	24	100
90	14	58050081	81	Inserção 4 bp	C>CTGAT	20	0	0	20	100
91	14	82390602	91	Eliminação 16 bp	CTTGCTCTATAAACCGT>C	10	0	0	10	100
93	14	102808554	94	Eliminação 5 bp	CGTGGA>C	10	0	0	10	100
95	15	63446199	68	Eliminação 6 bp	CAAAT>C	12	0	0	12	100
96	15	77879862	95	Eliminação 25 bp	GCCCCTGAGCCAGCCTCCCCTCTTA>G	14	0	0	14	100
98	15	85438311	72	Inserção 3 bp	C>CTTG	8	0	0	8	100
100	15	89864316	70	Inserção 4 bp	G>GCTAC	8	0	0	8	100
105	16	85706416	91	Eliminação 7 bp	ATTATTTTC>A	16	0	0	16	100
107	17	3594276	87	Eliminação 1 bp	TG>T	2	0	0	2	100
108	17	3970133	91	Inserção 18 bp	A>ATCCTATTCTACTCTGAAT	10	0	0	10	100
109	17	16084985	93	Inserção 4 bp	A>AACAC	10	0	0	10	100
111	17	39589692	84	Inserção 1 bp	C>CA	16	9	7	0	0
112	17	39589739	84	Inserção 24 bp	T>TTCTGAAGGTCAAGTCTATCCCTGA	24	0	0	24	100
113	17	45438886	92	Eliminação 4 bp	CAGTG>C	12	0	0	12	100
114	17	61502459	79	Eliminação 12 bp	TTTGATCTGCTG>T	20	0	0	20	100

Amplicon	Cromossoma	Posição	Tamanho do fragmento analisado	Tipo e tamanho do indel do fragmento amplificado	Indel	N.º total de réplicas de amostras com indel	N.º de não-identificações	N.º total de identificações de indel incorretas	N.º total de identificações de indel corretas	PPA
120	18	38837054	75	Inserção 22 bp	T>TGTATCTTAGCAAAGTTTCTCA	24	0	0	24	100
121	18	47405425	81	Inserção 3 bp	T>TGAG	20	0	0	20	100
122	18	54815706	85	Eliminação 2 bp	ACT>A	20	0	0	20	100
130	20	21766863	70	Eliminação 15 bp	TACTTGAGAACTGAGG>T	4	0	0	4	100
131	20	25278464	101	Inserção 5 bp	A>AGTGGG	20	0	0	20	100
132	20	50897361	67	Inserção 11 bp	G>GGAATGTCAGCC	24	0	0	24	100
134	20	62690925	87	Eliminação 16 bp	TCCTGGCTGGCCTGTGG>T	10	0	0	10	100
135	21	30300873	66	Inserção 11 bp	G>GATAAACTTTA	10	0	0	10	100
137	21	36710749	87	Eliminação 21 bp	ACTCAAGATAACTCATGTTATC>A	16	0	0	16	100
138	21	46644985	69	Eliminação 5 bp	GTTGTT>G	8	0	0	8	100
140	22	25750814	100	Inserção 6 bp	C>CAGGGCA	20	0	0	20	100
142	22	37409885	97	Inserção 5 bp	C>CTGTTT	2	0	0	2	100
144	22	47081407	92	Eliminação 10 bp	GGGCACAGGCA>G	12	0	0	12	100

Reprodutibilidade

Foram conduzidos dois estudos para avaliar a reprodutibilidade do instrumento MiSeqDx com linhas celulares (estudo 1 e 2) ou sangue com leucócitos diminuídos contendo linhas celulares (estudo 2). O estudo 1 utilizou vários instrumentos. O estudo 2 incluiu vários centros.

Estudo 1

A reprodutibilidade do instrumento MiSeqDx foi determinada com dois instrumentos, dois operadores e dois lotes de reagentes, num total de oito ensaios. O ensaio representativo, as amostras e o método de referência são os mesmos descritos para o estudo de precisão.

Os resultados são apresentados com base no fragmento amplificado para cada instrumento ([Tabela 6](#)) para demonstrar a reprodutibilidade da identificação em todos os instrumentos. A % de identificações corretas incluiu as identificações incorretas e não-identificações (em que não existe correspondência de um ou mais filtros para a identificação de variantes). Os instrumentos geraram números semelhantes de não-identificações e identificações incorretas dependendo do fragmento amplificado em particular.

Tabela 6 Resultados de reprodutibilidade instrumento a instrumento do estudo para o instrumento MiSeqDx (ao nível do fragmento amplificado)

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
								Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
1	1	36450499	36450591	93	93	Indel	0,22	8928	0	0	8928	0	0
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	7584	0	0	7584	0	0
3	1	218353867	218353957	91	91	Indel	0,4	8736	0	0	8736	0	0
4	1	223906657	223906748	92	92	Indel	0,49	8832	0	0	8832	0	0
5	1	228526602	228526682	81	81	PolyG (5)	0,69	7776	0	0	7776	0	0
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	6720	0	0	6720	0	0
7	1	247812041	247812128	88	88	PolyA (5), CT (3), TAA(3), indel	0,27	8448	0	0	8448	0	0
8	2	55862774	55862863	90	90	Indel	0,28	8640	0	0	8640	0	0
9	2	87003930	87004009	80	80	Indel	0,38	7680	0	0	7680	0	0
10	2	177016721	177016805	85	81	N/D	0,65	7775	1	0	7775	1	0
11	2	186625727	186625801	75	75	PolyA (8)	0,35	7200	0	0	7200	0	0
12	2	190323504	190323591	88	88	PolyT (5)	0,42	8448	0	0	8448	0	0
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	8352	0	0	8352	0	0
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	8736	0	0	8736	0	0
15	2	228147052	228147144	93	93	N/D	0,43	8928	0	0	8928	0	0
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	7008	0	0	7008	0	0
17	3	4466229	4466321	93	93	AT(3), indel	0,27	8761	0	167	8760	0	168
18	3	46620561	46620643	83	83	N/D	0,43	7968	0	0	7968	0	0
19	3	49851331	49851400	70	70	CT(3), indel	0,49	6720	0	0	6720	0	0
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG (3)	0,41	8448	0	0	8448	0	0

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
								Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
21	3	190106030	190106104	75	74	Indel	0,57	7096	0	8	7096	0	8
22	4	2233667	2233744	78	78	PolyA (6)	0,26	7488	0	0	7488	0	0
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	9312	0	0	9312	0	0
24	4	15688604	15688681	78	78	N/D	0,29	7488	0	0	7488	0	0
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	5952	0	0	5952	0	0
26	4	102839244	102839314	71	69	PolyA (5)	0,46	6624	0	0	6624	0	0
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	5952	0	0	5952	0	0
28	5	1882081	1882158	78	75	N/D	0,78	7200	0	0	7200	0	0
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	8064	0	0	8064	0	0
30	5	41069808	41069871	64	64	N/D	0,39	6144	0	0	6144	0	0
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	7968	0	0	7968	0	0
32	5	147475343	147475409	67	67	PolyT (5)	0,37	6432	0	0	6432	0	0
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	8736	0	0	8736	0	0
34	5	155662213	155662287	75	75	Indel	0,43	7200	0	0	7200	0	0
35	6	6318713	6318814	102	102	PolyG (6)	0,68	9792	0	0	9792	0	0
36	6	24949983	24950074	92	92	Indel	0,63	8832	0	0	8832	0	0
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	8979	0	45	8979	0	45
38	6	32147987	32148084	98	98	PolyT (5), TCT (3), CTT(3)	0,55	9408	0	0	9408	0	0
39	6	32986864	32986958	95	95	Indel	0,53	9120	0	0	9120	0	0
40	6	33408498	33408583	86	86	PolyC (6)	0,7	8256	0	0	8256	0	0
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	9024	0	0	9024	0	0
42	6	112435865	112435955	91	91	PolyA (5)	0,44	8736	0	0	8736	0	0
43	7	22202076	22202148	73	73	N/D	0,44	7008	0	0	7008	0	0
44	7	66276100	66276187	88	88	Indel	0,35	8448	0	0	8448	0	0
45	7	77365735	77365821	87	87	PolyA (7), AG (4)	0,26	8352	0	0	8352	0	0

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
								Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
46	7	110939946	110940030	85	85	Indel	0,38	8160	0	0	8160	0	0
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	8550	0	90	8550	0	90
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	8736	0	0	8736	0	0
49	7	154404519	154404599	81	66	N/D	0,31	6336	0	0	6336	0	0
50	7	156476507	156476599	93	93	Indel	0,35	8928	0	0	8928	0	0
51	8	1817312	1817394	83	83	N/D	0,42	7968	0	0	7968	0	0
52	8	24811020	24811109	90	89	PolyG (7), CTC (4), indel	0,61	8452	0	92	8449	0	95
53	8	76518625	76518691	67	67	Indel	0,3	6432	0	0	6432	0	0
54	9	103054909	103055006	98	98	PolyG (6)	0,67	9408	0	0	9408	0	0
55	9	105586150	105586214	65	65	Indel	0,32	6240	0	0	6240	0	0
56	9	107620823	107620918	96	96	N/D	0,49	9216	0	0	9216	0	0
57	9	123769149	123769231	83	83	AT(3)	0,37	7968	0	0	7968	0	0
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	9312	0	0	9312	0	0
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	7488	0	0	7488	0	0
60	10	11784629	11784726	98	91	GC(3)	0,87	8644	1	91	8644	1	91
61	10	27317777	27317855	79	79	PolyT (5)	0,3	7584	0	0	7584	0	0
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	8640	0	0	8640	0	0
63	10	45084159	45084253	95	95	Indel	0,35	9120	0	0	9120	0	0
64	10	55892599	55892687	89	88	AC(11), indel	0,42	8408	0	40	8407	0	41
65	10	101611250	101611329	80	80	N/D	0,49	7680	0	0	7680	0	0
66	10	118351373	118351453	81	81	N/D	0,51	7776	0	0	7776	0	0
67	11	8159816	8159912	97	96	N/D	0,45	9216	0	0	9216	0	0
68	11	30177648	30177717	70	70	Indel	0,46	6720	0	0	6720	0	0
69	11	47470345	47470444	100	100	N/D	0,65	9600	0	0	9600	0	0
70	11	59837679	59837740	62	62	Indel	0,37	5952	0	0	5952	0	0

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
								Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
71	11	64418856	64418957	102	102	N/D	0,59	9792	0	0	9792	0	0
72	11	93529612	93529684	73	73	PolyA (5)	0,4	7008	0	0	7008	0	0
73	11	101347052	101347136	85	85	N/D	0,42	8160	0	0	8160	0	0
74	11	102477336	102477426	91	91	PolyG (6)	0,55	8736	0	0	8736	0	0
75	11	118406285	118406369	85	85	Indel	0,53	8160	0	0	8160	0	0
76	11	120357801	120357885	85	85	PolyA (5), CA (3), indel	0,34	8160	0	0	8160	0	0
77	11	125769313	125769397	85	85	GA(3)	0,52	8160	0	0	8160	0	0
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	8064	0	0	8064	0	0
79	12	26811004	26811096	93	93	PolyA (7), AC (4)	0,33	8928	0	0	8928	0	0
80	12	30881766	30881846	81	81	N/D	0,49	7776	0	0	7776	0	0
81	12	88474105	88474175	71	71	PolyA (6)	0,35	6816	0	0	6816	0	0
82	12	120966872	120966966	95	95	PolyG (5)	0,68	9117	3	0	9119	1	0
83	13	24167504	24167576	73	73	N/D	0,52	7008	0	0	7008	0	0
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	8448	0	0	8448	0	0
85	13	44880112	44880200	89	89	Indel	0,49	8544	0	0	8544	0	0
86	13	77665218	77665294	77	77	Indel	0,39	7392	0	0	7392	0	0
87	14	31619327	31619393	67	67	GA(3),TA(3)	0,39	6432	0	0	6432	0	0
88	14	39517884	39517966	83	83	N/D	0,25	7968	0	0	7968	0	0
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	6830	0	82	6835	0	77
90	14	58050030	58050110	81	81	Indel	0,38	7776	0	0	7776	0	0
91	14	82390559	82390649	91	91	Indel	0,35	8736	0	0	8736	0	0
92	14	92549544	92549609	66	66	PolyA (5)	0,41	6336	0	0	6336	0	0
93	14	102808496	102808589	94	94	Indel	0,62	9024	0	0	9024	0	0
94	15	43170751	43170848	98	96	PolyC (5)	0,45	9216	0	0	9216	0	0
95	15	63446149	63446216	68	68	Indel	0,25	6528	0	0	6528	0	0

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
								Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	8928	0	0	8926	2	0
97	15	81625334	81625428	95	95	PolyT (6)	0,43	9120	0	0	9120	0	0
98	15	85438263	85438334	72	71	Indel	0,65	6816	0	0	6816	0	0
99	15	89817413	89817503	91	91	N/D	0,36	8736	0	0	8736	0	0
100	15	89864274	89864343	70	70	Indel	0,56	6720	0	0	6720	0	0
101	16	1894910	1894972	63	63	N/D	0,27	6048	0	0	6048	0	0
102	16	28997904	28997998	95	95	PolyC (5)	0,67	9120	0	0	9120	0	0
103	16	53682908	53682994	87	87	TA(3)	0,41	8352	0	0	8352	0	0
104	16	57954406	57954509	104	104	PolyC (5)	0,67	9984	0	0	9984	0	0
105	16	85706375	85706465	91	91	PolyT (5), indel	0,37	8736	0	0	8736	0	0
106	17	3563920	3564008	89	89	GC(3)	0,64	8544	0	0	8544	0	0
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	8347	0	5	8347	0	5
108	17	3970090	3970180	91	91	Indel	0,46	8736	0	0	8736	0	0
109	17	16084945	16085037	93	93	Indel	0,26	8928	0	0	8928	0	0
110	17	33998759	33998849	91	89	PolyT (5)	0,54	8544	0	0	8544	0	0
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	7776	7	89	7777	12	83
112	17	41244394	41244484	91	91	PolyA (5)	0,34	8736	0	0	8736	0	0
113	17	45438866	45438957	92	92	PolyA (7), AT (3), AT(4), AT (4), indel	0,26	8832	0	0	8832	0	0
114	17	61502432	61502510	79	79	Indel	0,41	7546	0	38	7547	0	37
115	17	64023582	64023667	86	86	PolyT (7)	0,22	8256	0	0	8256	0	0
116	17	72308237	72308320	84	84	GAG(3)	0,62	8064	0	0	8064	0	0
117	18	2616456	2616522	67	67	GA(3)	0,31	6432	0	0	6432	0	0
118	18	6980478	6980568	91	91	N/D	0,37	8736	0	0	8736	0	0
119	18	9888026	9888094	69	69	PolyA (6), TG (3)	0,43	6624	0	0	6624	0	0
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	7200	0	0	7200	0	0

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
								Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	7776	0	0	7776	0	0
122	18	54815665	54815749	85	85	CT(3), indel	0,45	8160	0	0	8160	0	0
123	18	59773996	59774060	65	65	N/D	0,48	6240	0	0	6240	0	0
124	19	625143	625241	99	99	N/D	0,59	9504	0	0	9504	0	0
125	19	18121418	18121491	74	74	N/D	0,68	7102	2	0	7104	0	0
126	19	18186574	18186643	70	70	N/D	0,64	6718	2	0	6718	2	0
127	20	746056	746149	94	94	N/D	0,61	9024	0	0	9024	0	0
128	20	10633195	10633276	82	82	AC(3)	0,59	7872	0	0	7872	0	0
129	20	17705633	17705708	76	76	CT(3)	0,58	7296	0	0	7296	0	0
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0,46	6720	0	0	6720	0	0
131	20	25278421	25278521	101	101	Indel	0,63	9696	0	0	9696	0	0
132	20	50897302	50897368	67	67	Indel	0,36	6432	0	0	6432	0	0
133	20	62331904	62331994	91	88	PolyG (6)	0,73	8360	0	88	8360	0	88
134	20	62690860	62690946	87	87	Indel	0,57	8352	0	0	8352	0	0
135	21	30300823	30300888	66	66	Indel	0,35	6336	0	0	6336	0	0
136	21	33694176	33694273	98	98	PolyT (6), CA (3)	0,54	9408	0	0	9408	0	0
137	21	36710706	36710792	87	87	GT(3), indel	0,39	8352	0	0	8352	0	0
138	21	46644924	46644992	69	69	PolyA (6), AG (3), indel	0,32	6603	0	21	6601	0	23
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	8640	0	0	8640	0	0
140	22	25750774	25750873	100	100	Indel	0,63	9600	0	0	9600	0	0
141	22	32439233	32439329	97	97	N/D	0,68	9312	0	0	9312	0	0
142	22	37409844	37409940	97	97	Indel	0,46	9312	0	0	9312	0	0
143	22	37637596	37637694	99	99	N/D	0,6	9504	0	0	9504	0	0
144	22	47081347	47081438	92	92	Indel	0,66	8832	0	0	8832	0	0
145	X	15870424	15870492	69	69	PolyT (5)	0,26	6624	0	0	6624	0	0

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
								Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
146	X	135288543	135288611	69	69	PolyC (5)	0,62	6624	0	0	6624	0	0
147	X	135290777	135290847	71	71	N/D	0,52	6816	0	0	6816	0	0
148	Y	2655397	2655461	65	0	N/D	0,55	0	0	0	0	0	0
149	Y	2655519	2655609	91	0	N/D	0,48	0	0	0	0	0	0
150	Y	2655609	2655679	71	0	PolyA (5)	0,37	0	0	0	0	0	0

Os resultados do estudo de reprodutibilidade foram analisados com base no operador utilizando frequência de variantes (Tabela 7). Esta análise demonstrou a consistência das frequências de variantes em todos os operadores. São apresentadas frequências de variantes médias com um desvio padrão de +/- 1.

Tabela 7 Resultados operador a operador para o instrumento MiSeqDx

Intervalo de frequência de variantes	N.º de variantes únicas	N.º total de variantes analisadas operador 1	N.º total de variantes analisadas operador 2	Frequência de variante média (DP) comunicada operador 1	Frequência de variante média (DP) comunicada operador 2
Homozigótico (0,70-1,00)	2424	2424	2422	0,94 +/- 0,07	0,96 +/- 0,05
Heterozigótico (0,20-0,70)	8240	8132	8128	0,48 +/- 0,04	0,49 +/- 0,04

São apresentados resultados do estudo de reprodutibilidade para cada amostra, compostos por todos os oito ensaios (Tabela 8). A deteção é avaliada para cada tipo de variante – SNV, inserções e eliminações – separadamente. As posições de referência são excluídas. Esta análise demonstrou a reprodutibilidade dos resultados das variantes em todas as amostras.

Tabela 8 Concordância dos resultados da identificação de bases do instrumento MiSeqDx por amostra

Amostra	SNV					Inserções			Eliminações			
	N.º total	N.º total de TP	N.º total de F	N.º total de F	N.º total de F	N.º total de TP	N.º total de FP	N.º total de FN	N.º total	N.º total de TP	N.º total de FP	N.º total de FN
NA12877	592	592	0	0	336	336	0	0	288	288	0	0
NA12878	1456	1456	0	0	320	304	0	0	384	368	0	0
NA12879	912	912	0	0	336	320	0	2	288	288	0	0
NA12880	1072	1071	0	0	384	384	0	0	320	304	0	0
NA12881	1248	1247	0	1	384	368	0	0	368	368	0	0
NA12882	944	943	0	1	352	336	0	4	304	288	0	0
NA12883	1088	1087	0	1	368	368	0	0	352	335	0	0
NA12884	1088	1088	0	0	400	384	0	5	336	336	0	0
NA12885	1200	1189	0	7	400	382	0	4	352	336	0	0
NA12886	1104	1102	0	2	368	352	0	3	368	368	0	0
NA12888	1056	1054	0	2	368	368	0	0	304	304	0	0
NA12893	1168	1168	0	0	352	336	0	1	368	368	0	0

Os dados fornecidos pelos oito ensaios neste estudo de reprodutibilidade suportam a alegação de que o instrumento MiSeqDx pode sequenciar de forma consistente:

- Conteúdo GC \geq 19% (todas as bases identificadas em 192 de 192 fragmentos amplificados sequenciados com 19% de conteúdo GC identificado corretamente com uma taxa de não-identificação de 1,1%)
- Conteúdo GC \leq 78% (todas as bases identificadas em 192 de 192 fragmentos amplificados sequenciados com 78% de conteúdo GC identificado corretamente com zero não-identificações)
- Tamanhos PolyA \leq 8 (repetição PolyA de 8 nucleótidos identificada corretamente em 192 de 192 fragmentos amplificados sequenciados contendo PolyA = 8)
- Tamanhos PolyT \leq 10 (repetição PolyT de 10 nucleótidos identificada corretamente em 192 de 192 fragmentos amplificados sequenciados contendo PolyT = 10)
- Tamanhos PolyG \leq 7 (repetição PolyG de 7 nucleótidos identificada corretamente em 192 de 192 fragmentos amplificados sequenciados contendo PolyG = 7)
- Tamanhos PolyC \leq 6 (repetição PolyC de 6 nucleótidos identificada corretamente em 576 de 576 fragmentos amplificados sequenciados contendo PolyC = 6)

- Tamanhos de repetição de dinucleótidos $\leq 11x$ (todas as bases identificadas em 192 de 192 fragmentos amplificados sequenciados com 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 192 de 192 fragmentos amplificados sequenciados com repetição de trinucleótidos de 5x foram identificadas corretamente com uma taxa de não-identificação de 0,5%)
- 24 ou menos inserções de base e 25 ou menos eliminações de base
 - 24 inserções de base identificadas corretamente em 192 de 192 amostras
 - 25 eliminações de base identificadas corretamente em 223 amostras e identificadas incorretamente em 1 amostra de 224 amostras

Estudo 2

Um estudo de reprodutibilidade de centro para centro efetuado com um ensaio representativo, o ensaio de fibrose cística de 139 variantes do MiSeqDx da Illumina, incluiu um subconjunto de variações genéticas clinicamente significativas *CFTR* analisadas com o software MiSeq Reporter com o fluxo de trabalho de sequenciação alvo do ADN MiSeqDx. O estudo em regime de ocultação utilizou 3 centros de ensaio e 2 operadores em cada centro. Foram testados dois painéis bem caracterizados de 46 amostras cada por cada um dos operadores em cada centro, num total de 810 identificações por centro. Os painéis continham uma mistura de ADN genómico de linhas celulares com variantes conhecidas no gene *CFTR*, bem como sangue com leucócitos diminuídos contendo linhas celulares com variantes conhecidas no gene *CFTR*. As amostras de sangue foram fornecidas para permitir a incorporação dos passos de extração utilizados para preparar gDNA que serve como entrada principal para o fluxo de trabalho do ensaio. A taxa de aprovação das amostras, definida como o número de amostras aprovadas nos indicadores de CQ na primeira tentativa, foi de 99,88%. Todos os resultados dos testes são baseados no teste inicial.

Tabela 9 Resumo dos resultados do estudo de reprodutibilidade efetuado com um ensaio representativo de fibrose cística de 139 variantes do MiSeqDx

Painel	N.º da amostra	Genótipo da amostra	Variantes	Identificações totais por centro	Identificações positivas concordantes (variantes)			Identificações negativas concordantes (tipo selvagem)			N.º de identificações incorretas	N.º de não-identificações	Concordância positiva (%)	Concordância negativa (%)	Concordância total (%)
					Centro 1	Centro 2	Centro 3	Centro 1	Centro 2	Centro 3					
					1	2	3	1	2	3					
A	1	S549N (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	2	1812-1G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100

Painel	N.º da amostra	Genótipo da amostra	Variantes	Identificações totais por centro	Identificações positivas concordantes (variantes)			Identificações negativas concordantes (tipo selvagem)			N.º de identificações incorretas	N.º de não-identificações	Concordância positiva (%)	Concordância negativa (%)	Concordância total (%)
					Centro 1	Centro 2	Centro 3	Centro 1	Centro 2	Centro 3					
					1	2	3	1	2	3					
A	3	Q493X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	4 ¹	F508del/2184delA (HET)		810	12	12	12	797	798	798	0	1 ¹	100	100	100
A	5 ²	Y122X/R1158X (HET)		810	12	10	12	798	665	798	0	135 ²	94,44	94,44	94,44
A	6	F508del/2183AA>G (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	7	R75X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	8	I507del/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	9 ³	F508del/W1282X (HET)		810	12	11	12	798	797	798	2 ³	0	97,22	99,96	99,92
A	10 ³	F508del/3272-26A>G (HET)		810	12	11	12	798	797	798	2 ³	0	97,22	99,96	99,92
A	11	F508del/3849+10kbC>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	12	621+1G>T/3120+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	13	E60X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	14	M1101K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	15	M1101K (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	16	F508del (HOM)	I506V, I507V, F508C não presente	828	6	6	6	822	822	822	0	0	100	100	100
A	17	F508del/3659delC (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	18	R117H/F508del (HET)	(TG)10 (T)9/ (TG)12(T)5	816	18	18	18	798	798	798	0	0	100	100	100

Painel	N.º da amostra	Genótipo da amostra	Variantes	Identificações totais por centro	Identificações positivas concordantes (variantes)			Identificações negativas concordantes (tipo selvagem)			N.º de identificações incorretas	N.º de não-identificações	Concordância positiva (%)	Concordância negativa (%)	Concordância total (%)
					Centro 1	Centro 2	Centro 3	Centro 1	Centro 2	Centro 3					
					1	2	3	1	2	3					
A	19	621+1G>T/711+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	20	G85E/621+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	21	A455E/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	22	F508del/R560T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	23	F508del/Y1092X (C>A) (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	24	N1303K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	25	G542X (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	26	G542X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	27	G551D/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	28	3849+10kbC>T (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	29	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
A	30	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	31	1717-1G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	32	R1162X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	33	R347P/G551D (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	34	R334W (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	35	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
A	36	G85E (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	37	I336K (HET)		810	6	6	6	804	804	804	0	0	100	100	100

Painel	N.º da amostra	Genótipo da amostra	Variantes	Identificações totais por centro	Identificações positivas concordantes (variantes)			Identificações negativas concordantes (tipo selvagem)			N.º de identificações incorretas	N.º de não-identificações	Concordância positiva (%)	Concordância negativa (%)	Concordância total (%)
					Centro 1	Centro 2	Centro 3	Centro 1	Centro 2	Centro 3					
					1	2	3	1	2	3					
A	38	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
A	39	F508del/3849+10kbC>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	40	621+1G>T/3120+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	41	F508del/3659delC (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	42	R117H/F508del (HET)	(TG)10 (T)9/ (TG)12(T)5	816	18	18	18	798	798	798	0	0	100	100	100
A	43	G85E/621+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	44	A455E/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	45	N1303K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	46	G551D/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	47	2789+5G>A (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
B	48	CFTR dele2, 3/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	49	F508del/1898+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	50	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	51	F508del/2143delT (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	52	3876delA (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	53	3905insT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	54	394delTT (HET)		810	6	6	6	804	804	804	0	0	100	100	100

Painel	N.º da amostra	Genótipo da amostra	Variantes	Identificações totais por centro	Identificações positivas concordantes (variantes)			Identificações negativas concordantes (tipo selvagem)			N.º de identificações incorretas	N.º de não-identificações	Concordância positiva (%)	Concordância negativa (%)	Concordância total (%)
					Centro 1	Centro 2	Centro 3	Centro 1	Centro 2	Centro 3					
					1	2	3	1	2	3					
B	55	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	56	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	57	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	58	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	59	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	60	L206W (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	61	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	62	G330X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	63	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	64	R347H (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	65	1078delT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	66	G178R/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	67	S549R (c.1647T>G) (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	68	S549N (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	69	W846X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	70	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	71	E92X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	72 ⁴	621+1G>T/1154insTC (HET)		810	12	12	12	798	798	797	0	1 ⁴	100	99,96	99,96
B	73	G542X (HET)		810	6	6	6	804	804	804	0	0	100	100	100

Painel	N.º da amostra	Genótipo da amostra	Variantes	Identificações totais por centro	Identificações positivas concordantes (variantes)			Identificações negativas concordantes (tipo selvagem)			N.º de identificações incorretas	N.º de não-identificações	Concordância positiva (%)	Concordância negativa (%)	Concordância total (%)
					Centro 1	Centro 2	Centro 3	Centro 1	Centro 2	Centro 3					
					1	2	3	1	2	3					
B	74	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	75 ²	F508del (HET)		810	6	5	6	804	670	804	0	135 ²	94,44	94,44	94,44
B	76	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	77	621+1G>T/A455E (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	78	1812-1G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	79	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	80	F508del/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	81	F508del/G551D (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	82	R347P/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	83	R117H/F508del (HET)	(TG)10 (T)9/ (TG)12(T)5	816	18	18	18	798	798	798	0	0	100	100	100
B	84	I507del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	85	2789+5G>A (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
B	86 ⁴	CFTR dele2, 3/F508del (HET)		810	12	12	12	798	797	798	0	1 ⁴	100	99,96	99,96
B	87	F508del/1898+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	88	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	89	F508del/2143delT (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	90	3905insT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	91	394delTT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	92	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
Total				74556	2209			221182			4	273	99,77	99,88	99,88

¹ A localização do tipo selvagem correspondente à variante N1303K para uma réplica resultou numa não-identificação devido a insuficiência da cobertura.

² Uma réplica das amostras 5 e 75 teve uma taxa de identificação de 0%. Uma investigação mais aprofundada indica que as amostras podem não ter sido adicionadas à placa de amostras antes da preparação de bibliotecas, uma vez que os volumes de amostras restantes nos tubos eram consistentes com a não remoção de qualquer volume.

³ A evidência indica que as amostras 9 e 10 foram provavelmente trocadas pelo operador antes da preparação de bibliotecas.

⁴ A localização do tipo selvagem correspondente à variante M1V para uma réplica de cada uma de duas amostras resultou numa não-identificação devido a insuficiência da cobertura.

Caraterísticas do desempenho somático

Os estudos aqui descritos utilizaram o módulo Somatic Variant para analisar a sequenciação de dados, exceto os que utilizaram um painel de dois genes quando foi utilizado um módulo específico do ensaio.

Precisão

Foram conduzidos três estudos para avaliar a precisão do instrumento MiSeqDx com ADN extraído de amostras FFPE.

Estudo 1

O estudo utilizou um ensaio representativo concebido para analisar uma variedade de genes abrangendo 12.588 bases ao longo de 23 cromossomas diferentes utilizando 150 fragmentos amplificados. O cromossoma Y não contém regiões de confiança e não foi avaliado. As cinco amostras únicas utilizadas neste estudo veem de uma só família – dois pais e três filhos – frequentemente sequenciada por vários laboratórios e metodologias de sequenciação. Existem três amostras femininas e duas masculinas. Todas as amostras foram fixadas em formalina e incorporadas em parafina antes da extração do ADN para o estudo. A amostra GM12877 foi diluída, ao nível do ADN, com a amostra GM12878 para criar GM12877-D para conseguir um conjunto de variantes com frequências próximas dos 5% e 10%. Cada uma das amostras foi testada em duplicado, exceto a amostra GM12877-D, que foi testada com cinco réplicas. A precisão foi determinada para SNV, inserções e eliminações comparando os dados do estudo com uma base de dados de referência bem caracterizada. A sequência da base de dados de referência (Platinum Genomes versão 2016-01) foi derivada da combinação de várias metodologias de sequenciação, dados disponíveis publicamente e informações hereditárias. Foram definidas regiões genómicas de confiança com base neste método de referência, exceto quando especificado em contrário. As amostras foram submetidas a ensaio oito vezes no total. As tabelas apresentadas para demonstrar a precisão têm por base os dados do primeiro ensaio.

A **Tabela 10** 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 (SNV, 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 10 Concordância dos resultados da identificação de bases do instrumento MiSeqDx com dados de referência para 6 amostras bem caracterizadas

Amostra	Taxa média de identificações	N.º total de variantes	N.º total de variantes TP	N.º total de variantes FN	N.º total de identificações TN	PPA	NPA	OPA
GM12877	98,7	152	147	0	23719	100	100	100
GM12878	98,4	270	260	0	23482	100	100	100
GM12879	98,7	192	186	0	23744	100	100	100
GM12885	99,1	244	236	0	23713	100	100	100
GM12886	98,7	230	226	0	23652	100	100	100
GM12877-D ¹		675	650	0		100	100	100
GM12877-D ²	98,4	155	155	0	57608	100	100	100

¹ Variantes com frequência superior a 20%.

² Variante com frequência inferior a 20%.

Os 150 fragmentos amplificados foram concebidos para abranger uma variedade de conteúdo genómico. O conteúdo GC dos fragmentos amplificados encontrava-se entre 26 e 87%. Os fragmentos amplificados também tinham uma gama de nucleótido único (p. ex. PolyA, PolyT), repetições de dinucleótidos e trinucleótidos. Foram utilizadas 6 amostras únicas no ensaio. Os dados foram compilados numa base por fragmento amplificado (**Tabela 11**) para determinar o efeito do conteúdo genómico na % de identificações corretas. A % de identificações corretas é composta por identificações de referência e variantes e será inferior a 100% se ocorrerem identificações incorretas ou não-identificações. Não ocorrem identificações quando não estão presentes um ou mais filtros para identificação de variantes (por ex. cobertura insuficiente). Não ocorreram identificações incorretas. O número de não-identificações variou consideravelmente em todos os fragmentos amplificados. O conteúdo GC e várias interações com conteúdo GC foram os preditores mais significativos de não-identificações. 2040/2580 (79%) de não-identificações deveram-se ao não cumprimento da especificação de cobertura. Os fragmentos amplificados com conteúdo GC superior a 78% resultaram no maior número de não-identificações. Um fragmento amplificado representativo com 78% de conteúdo GC teve no total 675 não-identificações.

Um fragmento amplificado representativo com 87% teve no total 1365 não-identificações. A cobertura pode ser aumentada reduzindo o número de amostras carregadas na célula de fluxo, o que pode permitir detecção em fragmentos amplificados com elevado conteúdo GC.

Tabela 11 Dados de precisão ao nível do fragmento amplificado

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
1	1	36450499	36450591	93	93	Indel	0,22	1395	0	0	100
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	1185	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	1364	0	1	99,9
4	1	223906657	223906748	92	92	Indel	0,49	1380	0	0	100
5	1	228526602	228526682	81	81	PolyG (5)	0,69	1215	0	0	100
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	1050	0	0	100
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA (3), indel	0,27	1320	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	1350	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	1200	0	0	100
10	2	177016721	177016805	85	81	N/D	0,65	1215	0	0	100
11	2	186625727	186625801	75	75	PolyA (8)	0,35	1117	0	10	99,1
12	2	190323504	190323591	88	88	PolyT (5)	0,42	1320	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	1302	0	8	99,4
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	1365	0	0	100
15	2	228147052	228147144	93	93	N/D	0,43	1395	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	1095	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	1349	0	46	96,7
18	3	46620561	46620643	83	83	N/D	0,43	1245	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	1050	0	0	100
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0,41	1305	0	30	97,8
21	3	190106030	190106104	75	74	Indel	0,57	1108	0	2	99,8
22	4	2233667	2233744	78	78	PolyA (6)	0,26	1170	0	0	100

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	1455	0	0	100
24	4	15688604	15688681	78	78	N/D	0,29	1169	0	1	99,9
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	930	0	0	100
26	4	102839244	102839314	71	69	PolyA (5)	0,46	1035	0	0	100
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	920	0	10	98,9
28	5	1882081	1882158	78	75	N/D	0,78	450	0	675	40,0
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	1260	0	0	100
30	5	41069808	41069871	64	64	N/D	0,39	960	0	0	100
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	1245	0	0	100
32	5	147475343	147475409	67	67	PolyT (5)	0,37	1005	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	1365	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	1125	0	0	100
35	6	6318713	6318814	102	102	PolyG (6)	0,68	1530	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	1380	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	1383	0	27	98,1
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT (3)	0,55	1455	0	15	99,0
39	6	32986864	32986958	95	95	Indel	0,53	1425	0	0	100
40	6	33408498	33408583	86	86	PolyC (6)	0,7	1290	0	0	100
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	1410	0	0	100
42	6	112435865	112435955	91	91	PolyA (5)	0,44	1365	0	0	100
43	7	22202076	22202148	73	73	N/D	0,44	1095	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	1320	0	0	100
45	7	77365735	77365821	87	87	PolyA (7), AG(4)	0,26	1299	0	6	99,5
46	7	110939946	110940030	85	85	Indel	0,38	1275	0	0	100
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	1350	0	0	100
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	1365	0	0	100

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
49	7	154404519	154404599	81	66	N/D	0,31	990	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	1395	0	0	100
51	8	1817312	1817394	83	83	N/D	0,42	1245	0	0	100
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0,61	1305	0	30	97,8
53	8	76518625	76518691	67	67	Indel	0,3	1005	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0,67	1470	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	973	0	2	99,8
56	9	107620823	107620918	96	96	N/D	0,49	1440	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	1242	0	3	99,8
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	1455	0	0	100
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	1170	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	0	0	1365	0
61	10	27317777	27317855	79	79	PolyT (5)	0,3	1185	0	0	100
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	1350	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	1425	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	1290	0	69	94,9
65	10	101611250	101611329	80	80	N/D	0,49	1200	0	0	100
66	10	118351373	118351453	81	81	N/D	0,51	1215	0	0	100
67	11	8159816	8159912	97	96	N/D	0,45	1440	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	1050	0	0	100
69	11	47470345	47470444	100	100	N/D	0,65	1500	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	930	0	0	100
71	11	64418856	64418957	102	102	N/D	0,59	1530	0	0	100
72	11	93529612	93529684	73	73	PolyA (5)	0,4	1095	0	0	100
73	11	101347052	101347136	85	85	N/D	0,42	1275	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0,55	1365	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	1275	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0,34	1275	0	0	100

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genômico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
77	11	125769313	125769397	85	85	GA(3)	0,52	1275	0	0	100
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	1260	0	14	98,9
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0,33	1395	0	0	100
80	12	30881766	30881846	81	81	N/D	0,49	1215	0	0	100
81	12	88474105	88474175	71	71	PolyA (6)	0,35	1065	0	0	100
82	12	120966872	120966966	95	95	PolyG (5)	0,68	1425	0	0	100
83	12	24167504	24167576	73	73	N/D	0,52	1095	0	0	100
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	1305	0	15	98,9
85	13	44880112	44880200	89	89	Indel	0,49	1335	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	1155	0	0	100
87	14	31619327	31619393	67	67	GA(3),TA(3)	0,39	1005	0	0	100
88	14	39517884	39517966	83	83	N/D	0,25	1245	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	1038	0	42	96,1
90	14	58050030	58050110	81	81	Indel	0,38	1215	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	1365	0	0	100
92	14	92549544	92549609	66	66	PolyA (5)	0,41	975	0	60	94,2
93	14	102808496	102808589	94	94	Indel	0,62	1410	0	0	100
94	15	43170751	43170848	98	96	PolyC (5)	0,45	1440	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	1020	0	0	100
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	1395	0	0	100
97	15	81625334	81625428	95	95	PolyT (6)	0,43	1425	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	1065	0	0	100
99	15	89817413	89817503	91	91	N/D	0,36	1365	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	1050	0	0	100
101	16	1894910	1894972	63	63	N/D	0,27	945	0	0	100
102	16	28997904	28997998	95	95	PolyC (5)	0,67	1425	0	0	100
103	16	3682908	53682994	87	87	TA(3)	0,41	1305	0	0	100
104	16	57954406	57954509	104	104	PolyC (5)	0,67	1560	0	0	100

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	1362	0	3	99,8
106	17	3563920	3564008	89	89	GC(3)	0,64	1335	0	0	100
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	1303	0	2	99,8
108	17	3970090	3970180	91	91	Indel	0,46	1365	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	1395	0	0	100
110	17	33998759	33998849	91	89	PolyT (5)	0,54	1335	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	1215	0	78	94,0
112	17	41244394	41244484	91	91	PolyA (5)	0,34	1365	0	0	100
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT(4), AT(4), indel	0,26	1365	0	15	98,9
114	17	61502432	61502510	79	79	Indel	0,41	1175	0	10	99,2
115	17	64023582	64023667	86	86	PolyT (7)	0,22	1289	0	1	99,9
116	17	72308237	72308320	84	84	GAG(3)	0,62	1260	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	1005	0	0	100
118	18	6980478	6980568	91	91	N/D	0,37	1365	0	0	100
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0,43	1035	0	0	100
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	1121	0	19	98,3
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	1215	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	00,45	1275	0	0	100
123	18	59773996	59774060	65	65	N/D	0,48	975	0	0	100
124	19	625143	625241	99	99	N/D	0,59	1478	0	7	99,5
125	19	18121418	18121491	74	74	N/D	0,68	1110	0	0	100
126	19	18186574	18186643	70	70	N/D	0,64	1050	0	0	100
127	20	746056	746149	94	94	N/D	0,61	1410	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	1230	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	1140	0	0	100
130	20	21766821	21766890	70	70	GT(3),TG(4), indel	0,46	1050	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	1515	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	1005	0	6	99,4

Amplicon	Cromossoma	Amplicon inicial	Amplicon final	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	N.º total de identificações corretas	N.º total de identificações incorretas	N.º total de não-identificações	% de identificações corretas
133	20	62331904	62331994	91	88	Poly G (6)	0,73	1320	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	1305	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	990	0	0	100
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0,54	1470	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	1305	0	0	100
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0,32	1029	0	7	99,3
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	1350	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	1500	0	1	99,9
141	22	32439233	32439329	97	97	N/D	0,68	1455	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	1455	0	0	100
143	22	37637596	37637694	99	99	N/D	0,6	1485	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	1380	0	0	100
145	X	15870424	15870492	69	69	PolyT (5)	0,26	1035	0	0	100
146	X	135288543	135288611	69	69	PolyC (5)	0,62	1035	0	0	100
147	X	135290777	135290847	71	71	N/D	0,52	1065	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	PolyA (5)	0,37	0	0	0	N/D

As variantes que constituíram não-identificações estão resumidas na [Tabela 12](#). Os filtros que resultaram nas não-identificações estão enumerados na tabela.

Tabela 12 Resumo de não-identificações de variantes

N.º de fragmentos amplificados	Cro:Pos	Variante	Conteúdo de fragmento amplificado correspondente	Filtro	Variantes ignoradas	Variantes esperadas
28	5:1882129	T > G	78% GC	LowDP ¹	8	13
52	8:24811064	AG > A	PolyG (7), CTC(4), 61% GC	R3x6 ²	15	15
60	10:11784633	C > T	PolyGC (3), 87% GC	LowDP	13	13
64	10:55892600	TAC > T	AC(11), 42% GC	R3x6	9	9
111	17:39589692	C > CA	PolyA (13), 29% GC	R3x6	13	13

¹ LowDP: pouca cobertura. Uma variante é filtrada se a profundidade de pelo menos um dos pools nesta posição em particular estiver abaixo de 900.

² R3x6: repetir filtro. Uma variante é filtrada se a variante estiver presente total ou parcialmente de forma repetida no genoma de referência adjacente à posição da variante. São necessárias pelo menos seis repetições na referência e apenas as repetições com 3 bp são consideradas.

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 13](#).

Tabela 13 Concordância dos resultados da identificação de bases do instrumento MiSeqDx com a referência NIST para a amostra GM12878

Amostra	N.º de fragmentos amplificados	Taxa média de identificações	N.º total de identificações de variantes TP	N.º total de identificações de variantes FN	N.º total de identificações TN	N.º total de identificações FP	PPA	NPA	OPA
GM12878	150	98,43	206	0	19231	0	100	100	100

As cinco amostras não diluídas foram analisadas novamente para identificar pequenas inserções e eliminações (indels) ([Tabela 14](#)).

Em alguns casos, o indel era comum entre duas ou mais amostras, conforme refletido na coluna N.º total de réplicas de amostras com indel. Os resultados de ambas as réplicas das cinco amostras estão incluídas na [Tabela 14](#). Houve no total 71 indels com tamanhos entre 1-24 bp para inserções e 1-25 bp para eliminações. Foram detetados 68 indels cada uma com uma concordância de percentagem positiva de 1. Três inserções e eliminações não apresentaram identificações corretas porque cada uma destas variantes era uma não-identificação devido ao filtro R3x6. Assim, não foi possível calcular o PPA, que exclui as não-identificações. As três variantes foram eliminação 1 bp (chr8 24811064 AG>A); eliminação 2 bp (chr10 55892600 TAC>T); e inserção 1 bp (chr17 39589692 C>CA).

Tabela 14 Resumo da deteção de indel com o instrumento MiSeqDx

Amplicon	Cromossoma	Posição	Tamanho do fragmento analisado	Tipo e tamanho do indel do fragmento amplificado	Indel	N.º total de réplicas de amostras com indel	N.º de não-identificações	N.º total de identificações de indel incorretas	N.º total de identificações de indel corretas	PPA
1	1	36450544	93	Eliminação 25 bp	GAAAATTTAATGAAACACATTGTCCT>G	7	0	0	7	100
2	1	109465165	79	Eliminação 3 bp	ACTT>A	9	0	0	9	100
3	1	218353908	91	Inserção 23 bp	T>TTTTAATAGCAAAAAGAGGCTAGA	15	0	0	15	100
4	1	223906701	92	Eliminação 17 bp	GACAGACTGTGAGGAAGA>G	11	0	0	11	100

Amplicon	Cromossoma	Posição	Tamanho do fragmento analisado	Tipo e tamanho do indel do fragmento amplificado	Indel	N.º total de réplicas de amostras com indel	N.º de não-identificações	N.º total de identificações de indel incorretas	N.º total de identificações de indel corretas	PPA
6	1	236372081	70	Inserção 5 bp	C>CTTAAG	9	0	0	9	100
7	1	247812083	88	Inserção 3 bp	C>CATG	9	0	0	9	100
8	2	55862804	90	Inserção 7 bp	T>TTTGGTAA	13	0	0	13	100
9	2	87003972	80	Eliminação 6 bp	TTATCTC>T	11	0	0	11	100
13	2	200796749	87	Inserção 5 bp	T>TTAAAA	15	0	0	15	100
14	2	212245090	91	Inserção 12 bp	C>CTGAAAATAGGAT	11	0	0	11	100
16	2	235016388	73	Inserção 2 bp	A>ATG	9	0	0	9	100
17	3	4466274	93	Eliminação 23 bp	TAACCTAAAATTACAAAATAACCC>T	13	0	0	13	100
19	3	49851375	70	Inserção 9 bp	C>CCTGGCTCCT	7	0	0	7	100
21	3	190106071	75	Eliminação 1 bp	AG>A	13	0	0	13	100
25	4	56236567	66	Eliminação 8 bp	TAACCGAAA>T	9	0	0	9	100
27	4	164446785	62	Inserção 11 bp	T>TTATGGTATTGA	9	0	0	9	100
31	5	74077155	83	Eliminação 4 bp	TAGTA>T	7	0	0	7	100
34	5	155662255	75	Inserção 8 bp	G>GCCTACTGA	13	0	0	13	100
36	6	24950035	92	Eliminação 21 bp	CCCTGGGTGCTATAGCCCACCA>C	11	0	0	11	100
37	6	31084942	100	Eliminação 3 bp	GCTT>G	15	0	0	15	100
39	6	32986905	95	Eliminação 25 bp	CTTTCACCTTCCCGTCTCATGCAAAG>C	7	0	0	7	100
41	6	41647442	95	Eliminação 23 bp	GGCATGAGGCTTGGTGACATGGCA>G	11	0	0	11	100
44	7	66276142	88	Inserção 1 bp	C>CT	13	0	0	13	100
46	7	110939983	85	Eliminação 4 bp	CAAGT>C	13	0	0	13	100
47	7	128533514	90	Inserção 1 bp	T>TC	15	0	0	15	100
48	7	149503916	91	Eliminação 4 bp	GGATA>G	7	0	0	7	100
50	7	156476548	93	Eliminação 11 bp	GAATCTGCACTT>G	13	0	0	13	100
52	8	24811064	90	Eliminação 1 bp	AG>A	15	15	0	0	N/D
53	8	76518677	67	Inserção 4 bp	T>TACTG	9	0	0	9	100
55	9	105586193	65	Inserção 4 bp	C>CAATT	13	0	0	13	100
58	9	138995370	97	Eliminação 21 bp	TCTGGGGGCAGCCCCTGAGGG>T	9	0	0	9	100
59	10	5987158	79	Eliminação 3 bp	TAAC>T	11	0	0	11	100

Amplicon	Cromossoma	Posição	Tamanho do fragmento analisado	Tipo e tamanho do indel do fragmento amplificado	Indel	N.º total de réplicas de amostras com indel	N.º de não-identificações	N.º total de identificações de indel incorretas	N.º total de identificações de indel corretas	PPA
63	10	45084202	95	Eliminação 16 bp	AGCGTCTATAACCAAT>A	11	0	0	11	100
64	10	55892600	89	Eliminação 2 bp	TAC>T	9	9	0	0	100
68	11	30177690	70	Inserção 2 bp	C>CTG	7	0	0	7	100
70	11	59837721	62	Inserção 8 bp	T>TTATGAAAA	11	0	0	11	100
75	11	118406328	85	Eliminação 8 bp	CAGTGTGGA>C	9	0	0	9	100
76	11	120357842	85	Eliminação 2 bp	CTT>C	11	0	0	11	100
78	12	2834814	84	Inserção 21 bp	T>TTCTCAGTACGGTGAACCCAG	15	0	0	15	100
84	13	25817002	89	Inserção 19 bp	C>CAAAATATAAAAAGCTCCCT	15	0	0	15	100
85	13	44880152	89	Inserção 4 bp	C>CCTGT	11	0	0	11	100
86	13	77665265	77	Eliminação 20 bp	ATCTATTTTCTAATAGACGGC>A	9	0	0	9	100
89	14	46958967	73	Eliminação 22 bp	TTTAAAATTTGAATGTGATAAAA>T	15	0	0	15	100
90	14	58050081	81	Inserção 4 bp	C>CTGAT	13	0	0	13	100
91	14	82390602	91	Eliminação 16 bp	CTTGCTCTATAAACCGT>C	11	0	0	11	100
93	14	102808554	94	Eliminação 5 bp	CGTGGA>C	9	0	0	9	100
95	15	63446199	68	Eliminação 6 bp	CAAAATT>C	11	0	0	11	100
96	15	77879862	95	Eliminação 25 bp	GCCCCTGAGCCAGCCTCCCGCTCTTA>G	9	0	0	9	100
98	15	85438311	72	Inserção 3 bp	C>CTTG	9	0	0	9	100
100	15	89864316	70	Inserção 4 bp	G>GCTAC	9	0	0	9	100
105	16	85706416	91	Eliminação 7 bp	ATTATTTTC>A	11	0	0	11	100
107	17	3594276	87	Eliminação 1 bp	TG>T	13	0	0	13	100
108	17	3970133	91	Inserção 18 bp	A>ATCCTATTCTACTCTGAAT	11	0	0	11	100
109	17	16084985	93	Inserção 4 bp	A>AACAC	7	0	0	7	100
111	17	39589692	84	Inserção 1 bp	C>CA	13	13	0	0	100
112	17	39589739	84	Inserção 24 bp	T>TTCTGAAGGTCAAGTCTATCCCTGA	15	0	0	15	100
113	17	45438886	92	Eliminação 4 bp	CAGTG>C	7	0	0	7	100
114	17	61502459	79	Eliminação 12 bp	TTTGTATCTGCTG>T	13	0	0	13	100
120	18	38837054	75	Inserção 22 bp	T>TGTATCTTAGCAAAAGTTTCTCA	15	0	0	15	100
121	18	47405425	81	Inserção 3 bp	T>TGAG	11	0	0	11	100

Amplicon	Cromossoma	Posição	Tamanho do fragmento analisado	Tipo e tamanho do indel do fragmento amplificado	Indel	N.º total de réplicas de amostras com indel	N.º de não-identificações	N.º total de identificações de indel incorretas	N.º total de identificações de indel corretas	PPA
122	18	54815706	85	Eliminação 2 bp	ACT>A	13	0	0	13	100
130	20	21766863	70	Eliminação 15 bp	TACTTGAGAACTGAGG>T	9	0	0	9	100
131	20	25278464	101	Inserção 5 bp	A>AGTGGG	13	0	0	13	100
132	20	50897361	67	Inserção 11 bp	G>GGAATGTCAGCC	15	0	0	15	100
134	20	62690925	87	Eliminação 16 bp	TCCTGGCTGGCCTGTGG>T	9	0	0	9	100
135	21	30300873	66	Inserção 11 bp	G>GATAAACTTTA	9	0	0	9	100
137	21	36710749	87	Eliminação 21 bp	ACTCAAGATAACTCATGTTATC>A	9	0	0	9	100
138	21	46644985	69	Eliminação 5 bp	GTTGTT>G	13	0	0	13	100
140	22	25750814	100	Inserção 6 bp	C>CAGGGCA	13	0	0	13	100
142	22	37409885	97	Inserção 5 bp	C>CTGTTT	13	0	0	13	100
144	22	47081407	92	Eliminação 10 bp	GGGCACAGGCA>G	7	0	0	7	100

Estudo 2

Este estudo utilizou amostras de tecido de cancro colorretal FFPE de reserva e um ensaio representativo de dois genes que foi comparado com a sequenciação Sanger bidirecional do método de referência (Sanger). Do total de 1183 indivíduos, 441 indivíduos tinham resultados do ensaio representativo e Sanger válidos. Quando avaliados ao nível do indivíduo ([Tabela 15](#)), 230 dos 441 indivíduos apresentaram resultados positivos na sequenciação Sanger (mutação detetada por Sanger). Destes, 227 apresentaram resultados positivos no ensaio representativo. Os restantes 211 dos 441 indivíduos apresentaram resultados negativos na sequenciação Sanger (sem mutações detetadas por Sanger). Destes, 206 apresentaram resultados negativos no ensaio representativo. O resultado foi uma PPA (percentagem de concordância positiva) de 98,7% e uma NPA (percentagem de concordância negativa) de 97,6% ([Tabela 15](#)).

Tabela 15 Percentagem de concordância positiva e negativa dos resultados ao nível do indivíduo

Ensaio representativo	Sanger		Total
	Positivos	Negativos	
Positivos	227 ¹	5	232
Negativos	3 ²	206	209
Total	230	211	441

Resumo do desempenho		
Estatística de concordância	Estimativa pontual	IC exata de 95%
PPA	227/230 = 98,7%	[96,2%,99,7%]
NPA	206/211 = 97,6%	[94,6%,99,2%]

¹Ocorreram 224 correspondências exatas para resultados ao nível de todas as mutações no mesmo indivíduo. Em dois indivíduos, MiSeqDx detetou a mutação detetada por Sanger e uma mutação adicional. Num indivíduo, MiSeqDx e Sanger detetaram mutações diferentes.

²Num indivíduo, Sanger detetou duas mutações. Em dois indivíduos, Sanger detetou uma mutação.

Estudo 3

Este estudo avaliou bibliotecas de ADN preparadas com amostras FFPE em vários tipos de tecido. Um total de 109 amostras FFPE de oito tecidos diferentes (côlon, ovário, pâncreas, adrenal, bexiga, fígado, tireoide e mama) com pelo menos 11 amostras FFPE a representar cada tipo de tecido. O tecido adrenal incluía metástases de tumores do esófago, pulmão e cólon. Os outros tecidos tinham tumores principais. Este estudo utilizou um ensaio representativo concebido para analisar 26 genes abrangendo 21 577 bases ao longo de 17 cromossomas diferentes. No total, seis genes diferentes (*KRAS*, *NRAS*, *TP53*, *PIK3CA*, *EGFR* e *BRAF*) foram submetidos a sequenciação Sanger; em cada tumor, 1-3 genes foram submetidos a sequenciação Sanger com base na prevalência esperada de mutações somáticas para esse tumor. Os resultados da sequenciação Sanger identificaram 39 mutações somáticas SNV em 33 de 109 amostras FFPE. O MiSeqDx identificou 36 mutações somáticas SNV em 32 de 109 amostras FFPE, com um falso negativo e duas não-identificações de posição da variante. A PPA foi de 97,3%. O MiSeqDx identificou 78.975 bases de referência nas 109 amostras FFPE com 29 falsos positivos relativamente à sequenciação Sanger e 2437 não-identificações. A NPA foi de 99,9%. Entre os dois métodos houve concordância na eliminação de duas bases. A [Tabela 16](#) resume os resultados por tipo de tecido.

Tabela 16 Percentagem de concordância positiva e negativa por tipo de tecido

Tipo de tecido	N.º de amostras	N.º total de variantes	N.º total de variantes TP	N.º total de variantes FN	N.º total de identificações TN	N.º total de identificações FP	N.º total de não-identificações	PPA	NPA
Adrenal	16	6	4	1	11823	2	607	80	>99,9
Bexiga	12	4	4	0	7070	3	273	100	>99,9
Mama	16	3	3	0	13439	7	479	100	99,9
Cólon	11	6	5	0	8720	2	133	100	>99,9
Fígado	13	3	3	0	7984	1	59	100	>99,9
Ovário	13	7	7	0	10581	1	724	100	>99,99
Pâncreas	17	7	7	0	11929	12	489	100	99,9
Tiroide	11	3	3	0	7429	1	652	100	>99,9
Total	109	39	36	1	78974	29	3416	97,3	>99,9

Reprodutibilidade

Foram conduzidos dois estudos para avaliar a reprodutibilidade do instrumento MiSeqDx com ADN extraído de amostras FFPE. O estudo 1 utilizou vários instrumentos. O estudo 2 incluiu vários centros.

Estudo 1

A reprodutibilidade do instrumento MiSeqDx foi determinada com dois instrumentos e dois operadores com formação, num total de oito ensaios. O ensaio representativo, o contexto genómico do fragmento amplificado, as amostras e o método de referência são os mesmos descritos para o estudo de precisão 1 acima. Os resultados são apresentados com base no fragmento amplificado para cada instrumento ([Tabela 17](#)) para demonstra a reprodutibilidade da identificação em todos os instrumentos. A % de identificações corretas incluiu as identificações incorretas e não-identificações (em que não existe correspondência de um ou mais filtros para a identificação de variantes). Os instrumentos geraram números semelhantes de não-identificações dependendo do fragmento amplificado em particular. Foi observada uma única identificação incorreta numa região de confiança conforme definida pelo padrão de referência Platinum Genomes para o MiSeqDx 1. A identificação incorreta foi uma identificação positiva falsa de uma variante de inserção no fragmento amplificado 64 examinando o cromossoma 10 nas posições 55892599 a 55892687. O fragmento amplificado tinha um tamanho de repetição de dinucleótidos de 11.

Tabela 17 Resultados de reprodutibilidade instrumento a instrumento do estudo para o instrumento MiSeqDx (ao nível do fragmento amplificado)

Amplicon	Cromossoma	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
						Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
1	1	93	93	Indel	0,22	5580	0	0	5580	0	0
2	1	79	79	PolyA (5), PolyC (5), indel	0,38	4740	0	0	4740	0	0
3	1	91	91	Indel	0,4	5448	0	12	5453	0	8
4	1	92	92	Indel	0,49	5518	0	2	5518	0	2
5	1	81	81	PolyG (5)	0,69	4858	0	2	4860	0	0
6	1	70	70	PolyT (10), indel	0,39	4200	0	0	4200	0	0
7	1	88	88	PolyA (5), CT(3), TAA (3), indel	0,27	5279	0	1	5279	0	1
8	2	90	90	Indel	0,28	5400	0	0	5400	0	0
9	2	80	80	Indel	0,38	4800	0	0	4800	0	0
10	2	85	81	N/D	0,65	4859	0	1	4859	0	1
11	2	75	75	PolyA (8)	0,35	4468	0	40	4468	0	40
12	2	88	88	PolyT (5)	0,42	5280	0	0	5280	0	0
13	2	87	87	PolyT (5), indel	0,31	5211	0	43	5214	0	40
14	2	91	91	PolyT (5), PolyA (6), indel	0,3	5453	0	7	5449	0	11
15	2	93	93	N/D	0,43	5579	0	1	5579	0	1
16	2	73	73	PolyT (5), indel	0,42	4378	0	2	4379	0	1
17	3	93	93	AT(3), indel	0,27	5396	0	184	5396	0	184
18	3	83	83	N/D	0,43	4980	0	0	4980	0	0
19	3	70	70	CT(3), indel	0,49	4193	0	7	4194	0	6
20	3	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0,41	5220	0	120	5220	0	120

Amplicon	Cromossoma	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
						Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
21	3	75	74	Indel	0,57	4432	0	8	4432	0	8
22	4	78	78	PolyA (6)	0,26	4676	0	4	4676	0	4
23	4	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	5820	0	0	5820	0	0
24	4	78	78	N/D	0,29	4679	0	1	4677	0	3
25	4	66	62	PolyA (5), indel	0,36	3720	0	0	3720	0	0
26	4	71	69	PolyA (5)	0,46	4140	0	0	4140	0	0
27	4	62	62	PolyA (7), indel	0,27	3676	0	45	3671	0	51
28	5	78	75	N/D	0,78	3368	0	1132	3485	0	1015
29	5	84	84	GT(3), CCA(3)	0,62	5040	0	0	5040	0	0
30	5	64	64	N/D	0,39	3840	0	0	3840	0	0
31	5	83	83	PolyA (6), indel	0,3	4979	0	1	4980	0	0
32	5	67	67	PolyT (5)	0,37	4020	0	0	4020	0	0
33	5	91	91	CT(4), AG(3)	0,55	5460	0	0	5460	0	0
34	5	75	75	Indel	0,43	4498	0	6	4500	0	1
35	6	102	102	PolyG (6)	0,68	6120	0	0	6120	0	0
36	6	92	92	Indel	0,63	5520	0	0	5520	0	0
37	6	100	94	GCT(5), indel	0,61	5532	0	108	5532	0	108
38	6	98	98	Poly T (5), TCT(3), CTT(3)	0,55	5820	0	60	5820	0	60
39	6	95	95	Indel	0,53	5697	0	3	5698	0	2
40	6	86	86	PolyC (6)	0,7	5159	0	1	5160	0	0
41	6	95	94	PolyG (5), indel	0,61	5638	0	2	5638	0	2
42	6	91	91	PolyA (5)	0,44	5460	0	0	5460	0	0
43	7	73	73	N/D	0,44	4380	0	0	4380	0	0
44	7	88	88	Indel	0,35	5279	0	1	5276	0	4

Amplicon	Cromossoma	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
						Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
45	7	87	87	PolyA (7), AG(4)	0,26	5184	0	36	5181	0	39
46	7	85	85	Indel	0,38	5100	0	0	5100	0	0
47	7	90	90	PolyG (5), indel	0,62	5398	0	2	5399	0	1
48	7	91	91	PolyG (6), PolyC (6), indel	0,71	5460	0	0	5459	0	1
49	7	81	66	N/D	0,31	3960	0	0	3960	0	0
50	7	93	93	Indel	0,35	5580	0	0	5579	0	1
51	8	83	83	N/D	0,42	4980	0	0	4980	0	0
52	8	90	89	PolyG (7), CTC(4), indel	0,61	5219	0	121	5220	0	120
53	8	67	67	Indel	0,3	4020	0	0	4020	0	0
54	9	98	98	PolyG (6)	0,67	5879	0	1	5880	0	0
55	9	65	65	Indel	0,32	3894	0	6	3895	0	5
56	9	96	96	N/D	0,49	5760	0	0	5760	0	0
57	9	83	83	AT(3)	0,37	4973	0	7	4978	0	2
58	9	97	97	PolyC (6), indel	0,68	5817	0	3	5818	0	2
59	10	79	78	PolyG (5), indel	0,47	4679	0	1	4680	0	0
60	10	98	91	GC(3)	0,87	450	0	5010	632	0	4828
61	10	79	79	PolyT (5)	0,3	4740	0	0	4740	0	0
62	10	90	90	PolyA (5), PolyT (5)	0,2	5400	0	0	5400	0	0
63	10	95	95	Indel	0,35	5699	0	1	5699	0	1
64	10	89	88	AC(11), indel	0,42	5157	0	276	5153	2	273
65	10	80	80	N/D	0,49	4800	0	0	4800	0	0
66	10	81	81	N/D	0,51	4860	0	0	4860	0	0
67	11	97	96	N/D	0,45	5760	0	0	5760	0	0
68	11	70	70	Indel	0,46	4199	0	2	4200	0	1

Amplicon	Cromossoma	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
						Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
69	11	100	100	N/D	0,65	5999	0	1	5998	0	2
70	11	62	62	Indel	0,37	3720	0	0	3720	0	0
71	11	102	102	N/D	0,59	6120	0	0	6118	0	2
72	11	73	73	PolyA (5)	0,4	4380	0	0	4380	0	0
73	11	85	85	N/D	0,42	5100	0	0	5100	0	0
74	11	91	91	PolyG (6)	0,55	5437	0	23	5441	0	19
75	11	85	85	Indel	0,53	5100	0	0	5100	0	0
76	11	85	85	Poly A (5), CA(3), indel	0,34	5100	0	0	5100	0	0
77	11	85	85	GA(3)	0,52	5100	0	0	5100	0	0
78	12	84	84	PolyC (5), indel	0,52	5040	0	60	5038	0	63
79	12	93	93	PolyA (7), AC(4)	0,33	5577	0	3	5573	0	7
80	12	81	81	N/D	0,49	4860	0	0	4860	0	0
81	12	71	71	PolyA (6)	0,35	4260	0	0	4260	0	0
82	2	95	95	PolyG (5)	0,68	5605	0	95	5605	0	95
83	13	73	73	N/D	0,52	4380	0	0	4379	0	1
84	13	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	5220	0	60	5220	0	60
85	13	89	89	Indel	0,49	5340	0	0	5340	0	0
86	13	77	77	Indel	0,39	4620	0	0	4620	0	0
87	14	67	67	GA(3),TA(3)	0,39	4020	0	0	4020	0	0
88	14	83	83	N/D	0,25	4980	0	0	4980	0	0
89	14	73	72	PolyT (5), indel	0,19	4173	0	147	4173	0	147
90	14	81	81	Indel	0,38	4860	0	2	4860	0	0
91	14	91	91	Indel	0,35	5459	0	1	5460	0	0
92	14	66	66	PolyA (5)	0,41	3900	0	240	3900	0	240

Amplicon	Cromossoma	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
						Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
93	14	94	94	Indel	0,62	5637	0	3	5637	0	3
94	15	98	96	PolyC (5)	0,45	5760	0	0	5760	0	0
95	15	68	68	Indel	0,25	4079	0	1	4078	0	2
96	15	95	93	PolyG (5), indel	0,68	5475	0	105	5487	0	93
97	15	95	95	PolyT (6)	0,43	5699	0	1	5700	0	0
98	15	72	71	Indel	0,65	4260	0	0	4260	0	0
99	15	91	91	N/D	0,36	5460	0	0	5460	0	0
100	15	70	70	Indel	0,56	4200	0	0	4200	0	0
101	16	63	63	N/D	0,27	3780	0	0	780	0	0
102	16	95	95	PolyC (5)	0,67	5700	0	0	5700	0	0
103	16	87	87	TA(3)	0,41	5220	0	0	5220	0	0
104	16	104	104	PolyC (5)	0,67	6238	0	3	6238	0	3
105	16	91	91	PolyT (5), indel	0,37	5443	0	17	5444	0	16
106	17	89	89	GC(3)	0,64	5251	0	89	5339	0	1
107	17	87	87	PolyC (5), indel	0,67	5212	0	8	5212	0	8
108	17	91	91	Indel	0,46	5459	0	1	5459	0	1
109	17	93	93	Indel	0,26	5580	0	0	5580	0	0
110	17	91	89	PolyT (5)	0,54	5340	0	0	5340	0	0
111	17	84	82	Poly A (13), indel (x2)	0,29	4860	0	308	4860	0	07
112	17	91	91	PolyA (5)	0,34	5459	0	1	5459	0	1
113	17	92	92	PolyA (7), AT(3), AT (4), AT(4), indel	0,26	5460	0	60	5460	0	60
114	17	79	79	Indel	0,41	4699	0	41	4700	0	40
115	17	86	86	PolyT (7)	0,22	5153	0	7	5156	0	4
116	17	84	84	GAG(3)	0,62	5039	0	1	5039	0	1

Amplicon	Cromossoma	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
						Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
117	18	67	67	GA(3)	0,31	4020	0	0	4020	0	0
118	18	91	91	N/D	0,37	5460	0	0	5460	0	0
119	18	69	69	PolyA (6), TG(3)	0,43	4132	0	8	4131	0	9
120	18	75	75	PolyA (5), indel	0,37	4475	0	85	4480	0	79
121	18	81	81	CTC(3), indel	0,47	4860	0	0	4860	0	0
122	18	85	85	CT(3), indel	0,45	5098	0	2	5098	0	2
123	18	65	65	N/D	0,48	3900	0	0	3900	0	0
124	19	99	99	N/D	0,59	5926	0	14	5924	0	16
125	19	74	74	N/D	0,68	4440	0	0	4438	0	2
126	19	70	70	N/D	0,64	4199	0	1	4200	0	0
127	20	94	94	N/D	0,61	5640	0	1	5638	0	3
128	20	82	82	AC(3)	0,59	4920	0	0	4920	0	0
129	20	76	76	CT(3)	0,58	4559	0	1	4558	0	2
130	20	70	70	GT(3),TG(4), indel	0,46	4200	0	0	4200	0	0
131	20	101	101	Indel	0,63	6060	0	0	6060	0	0
132	20	67	67	Indel	0,36	4020	0	31	4020	0	25
133	20	91	88	PolyG (6)	0,73	5277	0	3	5274	0	6
134	20	87	87	Indel	0,57	5218	0	2	5218	0	2
135	21	66	66	Indel	0,35	3959	0	1	3957	0	3
136	21	98	98	PolyT (6), CA(3)	0,54	5880	0	0	5880	0	0
137	21	87	87	GT(3), indel	0,39	5220	0	0	5220	0	0
138	21	69	69	PolyA (6), AG(3), indel	0,32	4119	0	31	4113	0	37
139	21	90	90	PolyT (5), PolyA (6)	0,5	5399	0	1	5399	0	1
140	22	100	100	Indel	0,63	5998	0	7	5997	0	5

Amplicon	Cromossoma	Tamanho do fragmento analisado	Bases em regiões de confiança	Conteúdo genómico de fragmento amplificado	Conteúdo GC	MiSeqDx 1			MiSeqDx 2		
						Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações	Total de identificações corretas	Total de identificações incorretas	N.º total de não-identificações
141	22	97	97	N/D	0,68	5819	0	1	5819	0	1
142	22	97	97	Indel	0,46	5818	0	2	5816	0	4
143	22	99	99	N/D	0,6	5940	0	0	5940	0	0
144	22	92	92	Indel	0,66	5519	0	1	5519	0	1
145	X	69	69	PolyT (5)	0,26	4139	0	1	4140	0	0
146	X	69	69	PolyC (5)	0,62	4136	0	4	4137	0	3
147	X	71	71	N/D	0,52	4260	0	0	4260	0	0
148	Y	65	0	N/D	0,55	0	0	0	0	0	0
149	Y	91	0	N/D	0,48	0	0	0	0	0	0
150	Y	71	0	PolyA (5)	0,37	0	0	0	0	0	0

Os resultados do estudo de reprodutibilidade foram analisados com base no operador utilizando frequência de variantes (Tabela 18). Esta análise demonstrou a consistência das frequências de variantes em todos os operadores. São apresentadas frequências de variantes médias com um desvio padrão de +/- 1.

Tabela 18 Resultados operador a operador para o instrumento MiSeqDx

Intervalo de frequência de variantes	N.º de variantes únicas	N.º total de variantes analisadas operador 1	N.º total de variantes analisadas operador 2	Frequência de variante média (DP) comunicada operador 1	Frequência de variante média (DP) comunicada operador 2
Alta frequência (~100%)	1112	1072	1072	0,96 +/- 0,05	0,96 +/- 0,05
Frequência média (~50%)	3240	3151	3161	0,49 +/- 0,04	0,49 +/- 0,04

Intervalo de frequência de variantes	N.º de variantes únicas	N.º total de variantes analisadas operador 1	N.º total de variantes analisadas operador 2	Frequência de variante média (DP) comunicada operador 1	Frequência de variante média (DP) comunicada operador 2
Baixa frequência (3-7%)	620	618	612	0,05 +/- 0,01	0,05 +/- 0,01

São apresentados resultados do estudo de reprodutibilidade para cada amostra, compostos por todos os oito ensaios (Tabela 19). A deteção é avaliada separadamente para cada tipo de variante – SNV, inserções e eliminações. As posições de referência são excluídas. Esta análise demonstrou a reprodutibilidade dos resultados das variantes em todas as amostras.

Tabela 19 Concordância dos resultados da identificação de bases do instrumento MiSeqDx por amostra

Amostra	SNV				Inserções				Eliminações			
	N.º total	N.º total de TP	N.º total de FP	N.º total de FN	N.º total	N.º total de TP	N.º total de FP	N.º total de FN	N.º total	N.º total de TP	N.º total de FP	N.º total de FN
GM12877	592	574	2	0	336	336	0	0	228	272	0	0
GM12878	1456	1432	0	0	320	304	0	0	384	352	0	0
GM12879	912	896	0	0	336	320	0	0	288	272	0	0
GM12885	1200	1192	0	0	400	384	0	0	352	320	0	0
GM12886	1104	1104	0	0	368	352	0	0	368	352	0	0
GM12877-D1 ¹	3640	3582	0	0	800	760	0	0	960	880	0	0
GM12877-D2 ²	400	398	0	0	520	516	0	0	560	556	0	0

¹Variante com frequência superior a 20%.

²Variante com frequência inferior a 20%.

Os dados fornecidos pelos 8 ensaios neste estudo de reprodutibilidade suportam a alegação de que o instrumento MiSeqDx pode sequenciar de forma consistente:

- Conteúdo GC \geq 19% (todas as bases identificadas em 120 de 120 fragmentos amplificados sequenciados com 19% de conteúdo GC identificado corretamente com uma taxa de não-identificação de 3,4%)
- Conteúdo GC \leq 73% (todas as bases identificadas em 120 de 120 fragmentos amplificados sequenciados com 73% de conteúdo GC identificado corretamente com uma taxa de não-identificação de 0,1%)
- Tamanhos PolyA \leq 8 (repetição PolyA de 8 nucleótidos identificada corretamente em 120 de 120 fragmentos amplificados sequenciados contendo PolyA = 8)
- Tamanhos PolyT \leq 10 (repetição PolyT de 10 nucleótidos identificada corretamente em 120 de 120 fragmentos amplificados sequenciados contendo PolyT = 10)
- Tamanhos PolyG \leq 6 (repetição PolyG de 6 nucleótidos identificada corretamente em 720 de 720 fragmentos amplificados sequenciados contendo PolyG = 6)
- Tamanhos PolyC \leq 6 (repetição PolyC de 6 nucleótidos identificada corretamente em 359 de 360 fragmentos amplificados sequenciados contendo PolyC = 6, com 1 não-identificação)
- Tamanhos de repetição de dinucleótidos \leq 4x (todas as bases identificadas em 600 de 600 fragmentos amplificados sequenciados com repetição de dinucleótidos de 4x foram identificadas corretamente com uma taxa de não-identificação de 0,4%)
- Tamanhos de repetição de trinucleótidos \leq 5x (todas as bases identificadas em 120 de 120 fragmentos amplificados sequenciados com repetição de trinucleótidos de 5x foram identificadas corretamente com uma taxa de não-identificação de 1,9%)
- 24 ou menos inserções de base e 25 ou menos eliminações de base
 - 24 inserções de base identificadas corretamente em 120 de 120 amostras
 - 25 eliminações de base identificadas corretamente em 182 amostras e comunicadas como não-identificações em 2 amostras de 184 amostras

Estudo 2

Foi efetuado um estudo externo para avaliar a reprodutibilidade do ensaio representativo de dois genes, descrito no estudo de precisão 2, em três centros de teste externos (dois operadores por centro), um lote de reagentes e três dias de teste não consecutivos. Os testes foram executados com seis painéis bem caracterizados de amostras de ADN genómico a partir de amostras clínicas de FFPE ou linhas celulares. Cada painel consistia em 10 membros, num total de 60 membros em todos os painéis.

Os 60 membros dos painéis consistiam em duplicados de quatro amostras de tipo selvagem único (para mutações de painel), 12 amostras com mutação única (com uma única mutação) preparadas em níveis altos e baixos de frequência de mutação e duas amostras com mutação única (com uma única mutação) preparadas apenas a um nível baixo de frequência. Cada amostra de nível de frequência de mutação/amostra única (testada em duplicado em cada ensaio) tinha 36 resultados possíveis (2 réplicas × 2 operadores × 3 dias × 3 centros) se todos os resultados fossem válidos.

A identificação esperada de percentagem (PEC) em todas as variantes positivas e negativas foi avaliada comparando o resultado do ensaio representativo com o resultado esperado da mutação (mutação esperada detetada ou não detetada) em cada amostra. A PEC é calculada como 100% das vezes o número de identificações esperadas divididas pelo número de identificações tentadas. O intervalo de confiança de 95% bilateral é calculado através do método de pontuação de Wilson.

Em centros combinados, as taxas de aprovação das amostras foram $\geq 94,7\%$ para o primeiro ensaio da amostra ou em amostras testadas em ensaios que foram consideradas válidas na primeira aprovação. A PEC ao nível da mutação em todas as amostras com mutação foi de 99,6% (905/909) (IC de 95%; 98,9, 99,8). O número de identificações tentadas em todas as 56 mutações de painel (independentemente de uma mutação detetada ser ou não esperada) para todas as amostras válidas foi 58.856 (56 × 1051). Destas 58 856 observações ao nível da mutação, houve apenas seis incidentes em que os resultados observados e esperados foram discordantes. A PEC ao nível da mutação em todas as variantes positivas e negativas de todos os membros do painel de tipo selvagem e mutantes combinados foi de 99,9% (58.850/58.856).

Sensibilidade analítica (Limite de vazio [LoB] e Limite de deteção [LoD])

Este estudo verificou o limite do ensaio e determinou o Limite de deteção (LoD) para o MiSeqDx com um painel representativo. Resumidamente, as bem caracterizadas linhas celulares GM12878 e GM12877 do Platinum Genome foram fixadas em formalina e incorporadas em parafina, e depois o ADN foi extraído. GM12878 foi diluída com GM12877 a tal ponto que as frequências de variante de setenta variantes (52 SNV, nove inserções e nove eliminações) estiveram perto de 0,05. As duas amostras de ADN foram testadas por dois operadores, utilizando dois instrumentos e dois lotes de reagentes, para um total de 10 ensaios de sequenciação MiSeqDx. Isto resultou em 40 réplicas de cada variante na GM12878 e em 60 réplicas de cada coordenada de tipo selvagem correspondente na GM12877 para cada lote de reagentes. O LoB e o LoD foram calculados utilizando a abordagem clássica indicada em CLSI EP17-A2 utilizando a opção não paramétrica. O LoB e o LoD foram calculados para SNV, inserções e eliminações em separado por 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 20). O software MiSeqDx utiliza um limite (o LoB efetivo) de 0,026 frequência de variantes para determinar a deteção qualitativa de variantes. Os limites de vazio calculados verificaram que este limite resulta num erro de tipo I não superior a 0,01.

Tabela 20 Limite de vazio

Tipo de variante	Número total de frequências de variantes	Lote de reagentes LoB 1 (%)	Lote de reagentes LoB 2 (%)
SNV	3120	0,87	0,75
Inserção	540	0,79	0,60
Eliminação	540	0,96	0,84

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 21). Como as percentagens eram inferiores ao erro de tipo II de 5% (0,05), a mediana das frequências de variantes combinadas foi calculada como o LoD (Tabela 22). O LoD de cada tipo de variante foi considerado como o maior dos dois valores calculados para os dois lotes de reagentes – 5,45% para SNV, 4,88% para inserções e 5,44% para eliminações.

Tabela 21 Limite de deteção

Lote de reagentes	Tipo de variante	Número total de frequências de variantes	N.º de medições VF < 2,6%	% de medições VF < 2,6%	Limite de deteção (%)
1	SNV	2080	5	0,20	5,45
	Inserção	360	0	0,00	4,86
	Eliminação	360	3	0,80	5,44
2	SNV	2080	26	1,30	5,44
	Inserção	360	0	0,00	4,88
	Eliminação	360	0	0,00	5,24

Os estudos que se seguem demonstram características de desempenho do MiSeqDx com outro ensaio representativo que especifica 56 mutações em dois genes de cancro clinicamente relevantes (painel de mutação). O painel de mutação foi concebido para detetar especificamente 56 mutações em dois genes de cancro clinicamente relevantes (Gene 1 e Gene 2). O ensaio determina simultaneamente a presença ou a ausência de cada uma das 56 mutações em todas as amostras sequenciadas. O método de referência para estes estudos foi a sequenciação Sanger bidirecional.

Precisão lote a lote

Foi efetuado um estudo de precisão lote a lote para avaliar o desempenho do instrumento MiSeqDx em vários lotes de kits de reagentes fabricados (que incluem qualificação de amostras, preparação de bibliotecas e sequenciação de reagentes) utilizando o ensaio representativo de dois genes utilizando um painel de cinco amostras FFPE misturadas que cumprem os requisitos de qualificação da amostra. Cada amostra FFPE continha duas mutações únicas: uma com nível baixo de frequência de mutação (aproximadamente 8%) e uma com nível alto de frequência de mutação (aproximadamente 14%). Doze (12) observações de cada uma das cinco misturas de amostras foram recolhidas ao longo de três dias não consecutivos com três lotes de kits de reagentes. O número total de observações para o estudo em todos os lotes de reagentes foi de 180 observações em todas as misturas de amostras e 360 observações em todos os níveis de frequência de mutação. Em todos os lotes e dias, 99,7% (359/360) das observações apresentaram o resultado de mutação esperado. Uma mutação de baixa frequência foi incorretamente identificada como tipo selvagem. Foi executada uma análise de variações de componentes para cada uma das mutações/níveis de frequência de mutação para calcular a variabilidade do sistema. O desvio padrão total foi de 0,011 a 0,029. O componente do lote de reagentes do desvio padrão total foi de 0 a 0,015.

Histórico de revisão

Documento n.º	Data	Descrição da alteração
Documento n.º 200006218 v01	Maio de 2022	Adição do número de catálogo do folheto informativo do MiSeqDx Reagent Kit v3 Micro à secção Equipamento e materiais necessários, mas não fornecidos. Adição de informações de teste do MiSeqDx Reagent Kit v3 Micro à secção Características de desempenho. Remoção da nota de atenção específica dos EUA da secção Avisos e precauções.
Documento n.º 200006218 v00	Novembro de 2021	Edição inicial para suportar o MOS v4.0 e o Local Run Manager v3.0.

Patentes e marcas comerciais

Este documento e respetivo conteúdo são propriedade da Illumina, Inc. e das suas afiliadas ("Illumina") e destinam-se unicamente a 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).

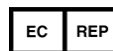
© 2022 Illumina, Inc. Todos os direitos reservados.

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

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.