

## Folheto informativo

PARA UTILIZAÇÃO EM DIAGNÓSTICO IN VITRO  
SOMENTE PARA EXPORTAÇÃO

## Uso previsto

O NovaSeq 6000Dx Instrument destina-se ao sequenciamento de bibliotecas de DNA quando usado com ensaios de diagnóstico *in vitro* (IVD). O NovaSeq 6000Dx Instrument destina-se ao uso com reagentes IVD específicos registrados, certificados ou aprovados e software analítico.

## Princípios do procedimento

O Illumina® NovaSeq 6000Dx Instrument destina-se ao sequenciamento de bibliotecas de DNA quando usado com ensaios de diagnóstico *in vitro* (IVD). Para a entrada, o NovaSeq 6000Dx usa bibliotecas geradas a partir de DNA onde índices de amostras e seqüências de captura são adicionados a alvos amplificados. As bibliotecas de amostras são capturadas em uma lâmina de fluxo e sequenciadas no instrumento com o uso da química de sequenciamento por síntese (SBS, sequencing by synthesis). A química de SBS usa um método de terminador reversível para detectar bases de nucleotídeo únicas identificadas por fluorescência, à medida que são incorporadas aos filamentos crescentes de DNA. O software Real-Time Analysis (RTA) faz a análise da imagem e a identificação de bases e atribui uma pontuação de qualidade a cada base de cada ciclo de sequenciamento. Quando a análise primária for concluída, a análise secundária poderá ser executada no Illumina DRAGEN Server incluído e necessário para o NovaSeq 6000Dx processar as identificações de bases. O NovaSeq 6000Dx usa módulos diferentes de análise secundária, dependendo do fluxo de trabalho. Para o DRAGEN for Illumina DNA Prep with Enrichment Dx Application, o processamento inclui demultiplexação, geração de arquivos FASTQ, alinhamento, chamada de variante e geração de arquivos de formato de chamada de variante (VCF e gVCF). Os arquivos VCF e gVCF contêm informações sobre variantes de linha genética ou somáticas (dependendo do fluxo de trabalho selecionado) encontradas em posições específicas em um genoma de referência.

## Modo de operação duplo

O NovaSeq 6000Dx inclui um disco rígido de inicialização única com modos separados de diagnóstico *in vitro* (IVD) e somente para pesquisa (RUO). O modo é selecionado usando uma alternância na tela Sequencing (Sequenciamento). O modo selecionado é claramente identificado na interface em todas as telas. Os ensaios de sequenciamento IVD, incluindo o DRAGEN for Illumina DNA Prep with Enrichment Dx Application em fluxos de trabalho de linha genética e/ou somáticos, são executados no modo IVD. Somente reagentes de sequenciamento IVD podem ser usados no modo IVD. As características de desempenho e limitações do procedimento para o NovaSeq 6000Dx foram estabelecidas usando o DRAGEN for Illumina DNA Prep with Enrichment Dx Application no modo IVD.

## Limitações do procedimento

1. Somente para uso em diagnóstico *in vitro*.
2. O DRAGEN for Illumina DNA Prep with Enrichment Dx Application, quando usado com o NovaSeq 6000Dx S2 Reagent v1.5 Kit (300 cycles) e o NovaSeq 6000Dx S4 Reagent v1.5 Kit (300 cycles), é capaz de fornecer:
  - Geração de dados de sequenciamento:
    - $\geq 1,0$  terabases (TB) com o kit S2
    - $\geq 3,0$  TB com o kit S4
  - Comprimento da leitura (em corrida tipo paired-end) 2 x 150 pares de base (bp)
  - Bases superiores a Q30  $\geq 85\%$  no comprimento da leitura de 2 x 150 bp. Igual ou superior a 85% das identificações de bases têm pontuações de qualidade na escala Phred superiores a 30, indicando precisão de identificação de bases superior a 99,9%.
3. Inserções de comprimento > 18 bp e deleções de comprimento > 21 bp não foram validadas.
4. Grandes variantes, incluindo variantes de vários nucleotídeos (MNVs, multi-nucleotide variants) e grandes indels, podem ser reportadas como variantes menores separadas no arquivo VCF de saída.
5. MNVs pequenos são relatados como variantes separadas no arquivo VCF de saída.
6. As deleções são reportadas no arquivo VCF na coordenada da base precedente segundo o formato VCF. Portanto, leve em consideração as variantes adjacentes antes de relatar que uma determinada identificação de bases é uma referência homozigótica.
7. Limitações específicas de linha genética:
  - O NovaSeq 6000Dx que usa o fluxo de trabalho de análise de geração de linha genética FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx Application foi projetado para fornecer resultados qualitativos para identificação de variantes de linha genética (por exemplo, homozigoto, heterozigoto, tipo selvagem).
  - A variação do número de cópias pode afetar o fato de uma variante ser identificada como homozigótica ou heterozigótica.
  - O sistema não relatará mais de duas variantes em um único local, mesmo na presença de variação do número de cópias.
8. Limitações específicas somáticas:
  - O NovaSeq 6000Dx usando o fluxo de trabalho de análise de geração somática FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx Application foi projetado para fornecer resultados qualitativos para identificação de variantes somáticas (ou seja, presença de uma variante somática).
  - O fluxo de trabalho de análise de geração de FASTQ e VCF somático não consegue diferenciar entre variantes de linha genética e somática. O fluxo de trabalho foi projetado para detectar variantes em uma variedade de frequências de variantes, mas a frequência de variantes não pode ser usada para diferenciar variantes somáticas de variantes de linha genética.

- O tecido normal no espécime afeta a detecção de variantes. O limite reportado de detecção é baseado em uma frequência de variante em relação ao DNA total extraído tanto de tecido tumoral quanto de tecido normal.
- Se mais de um alelo variante for chamado no mesmo locus, nenhum dos alelos será relatado como variantes de passagem. Em vez disso, o conjunto completo de alelos será relatado, mas filtrado por meio da tag multialélica.

## Procedimentos de controle de qualidade

O software do NovaSeq 6000Dx avalia cada corrida, amostra e identificação de bases com relação a métricas de controle de qualidade. Também são recomendados controles positivos e negativos na preparação das bibliotecas e que precisam ser avaliados. Avalie os controles como mostrado a seguir.

- Controle negativo (controle sem modelo) ou outro controle negativo: deve gerar o resultado esperado. Se o controle negativo gerar um resultado diferente do esperado, pode ter ocorrido erro no rastreamento das amostras, registro incorreto de primers de índice ou contaminação.
- Amostra de controle positivo: deve gerar o resultado esperado. Se o controle positivo gerar um resultado diferente do esperado, pode ter ocorrido erro no rastreamento das amostras ou registro incorreto de primers de índice.

## Componentes do produto

O Illumina NovaSeq 6000Dx consiste no seguinte:

1. NovaSeq 6000Dx Instrument (n.º do catálogo 20068232)
2. Os componentes do software para o NovaSeq 6000Dx Instrument incluem:

Aplicação do software	Local de instalação	Função	Descrição
NovaSeq Operating Software	NovaSeq 6000Dx	Controla a operação do instrumento	O NovaSeq Operating Software (NVOS) gerencia a operação do instrumento durante o sequenciamento e gera imagens para uso pelo software Real-Time Analysis (RTA).

Aplicação do software	Local de instalação	Função	Descrição
Real-Time Analysis Software (RTA)	NovaSeq 6000Dx	Executa a análise primária	A aplicação de software RTA converte as imagens geradas pelo NVOS para cada bloco por ciclo de corrida de sequenciamento em arquivos de identificação de bases. Os arquivos de identificação de bases são entradas para os módulos de aplicação no Illumina DRAGEN Server para NovaSeq 6000Dx. A aplicação de software RTA não contém uma interface de usuário.
Illumina Run Manager	Illumina DRAGEN Server	Controla configuração e gerenciamento do sequenciamento	O Illumina Run Manager fornece gerenciamento de usuários e instrumentos, hospeda o software de aplicação e permite o uso de módulos de análise genômica secundária acelerada por hardware do DRAGEN.

## Condições de operação

Para obter mais informações sobre as condições operacionais, consulte a seção Considerações ambientais na *documentação do produto do NovaSeq 6000Dx Instrument*.

Elemento	Especificação
Temperatura	Mantenha a temperatura do laboratório entre 19 °C e 25 °C (22 °C ±3 °C). Esta é a faixa de temperatura operacional do instrumento. Durante o sequenciamento, não permita que a temperatura ambiente varie mais do que ±2 °C.
Umidade	Mantenha uma umidade relativa não condensada entre 20% e 80%. O sistema deve ser operado em uma altitude operacional de até 2.000 m.

## Materiais de consumo e equipamentos

Esta seção lista tudo o que é necessário para a corrida de um sequenciamento no NovaSeq 6000Dx. Isso inclui materiais de consumo fornecidos pela Illumina, materiais de consumo auxiliares e equipamentos que você deve comprar de outros fornecedores. Esses itens são necessários para concluir o protocolo e realizar procedimentos de manutenção e solução de problemas.

Para obter informações sobre os símbolos em materiais de consumo ou embalagens de materiais de consumo, consulte a [Legenda de símbolos de IVD da Illumina \(documento n.º 1000000039141\)](#).

## Materiais de consumo de sequenciamento

Uma corrida do NovaSeq 6000Dx requer os seguintes componentes:

- Cartucho de solução tampão
- Cartucho de cluster
- Lâmina de fluxo
- Tubo da biblioteca
- Cartucho SBS

Os materiais de consumo do NovaSeq 6000Dx são embalados nas configurações abaixo. Cada componente usa identificação por radiofrequência (RFID) para obter compatibilidade e um rastreamento de materiais de consumo preciso.

Tabela 1 Materiais de consumo fornecidos pela Illumina

Nome do kit	Conteúdo	Número de catálogo da Illumina
NovaSeq 6000Dx S2 Reagent v1.5 Kit (300 cycles)	Cartucho de cluster S2 Lâmina de fluxo S2 Cartucho SBS S2	20046931
NovaSeq 6000Dx S4 Reagent v1.5 Kit (300 cycles)	Cartucho de cluster S4 Lâmina de fluxo S4 Cartucho SBS S4	20046933
NovaSeq 6000Dx S2 Buffer Cartridge	Cartucho de solução tampão S2	20062292
NovaSeq 6000Dx S4 Buffer Cartridge	Cartucho de solução tampão S4	20062293
NovaSeq 6000Dx Library Tube	Tubo único da biblioteca	20062290
NovaSeq 6000Dx Library Tube, 24 Pack	24 tubos da biblioteca	20062291

Quando receber o material de consumo, armazene prontamente os componentes na temperatura indicada para assegurar o desempenho apropriado.

Tabela 2 Armazenamento do NovaSeq 6000Dx Kit

Material de consumo	Quantidade	Temperatura de armazenamento	Comprimento	Largura	Altura
Lâmina de fluxo	1	2 °C a 8 °C	27,7 cm (10,9 pol.)	17 cm (6,7 pol.)	3,8 cm (1,5 pol.)

Material de consumo	Quantidade	Temperatura de armazenamento	Comprimento	Largura	Altura
Cartucho de cluster	1	-25 °C a -15 °C	29,5 cm (11,6 pol.)	13 cm (5,1 pol.)	9,4 cm (3,7 pol.)
Cartucho SBS	1	-25 °C a -15 °C	30 cm (11,8 pol.)	12,4 cm (4,9 pol.)	11,2 cm (4,4 pol.)
Cartucho de solução tampão	1	15 °C a 30 °C	42,2 cm (16,6 pol.)	20,6 cm (8,1 pol.)	21,1 cm (8,3 pol.)
Tubo da biblioteca	1	15 °C a 30 °C	4,1 cm (1,6 pol.)	2,3 cm (0,9 pol.)	12,4 (4,9 pol.)

## Detalhes dos materiais de consumo

Para identificar componentes compatíveis do kit, as lâminas de fluxo e os cartuchos são rotulados com símbolos que mostram o modo do kit.

Tabela 3 Identificação de compatibilidade

Modo do kit	Marcação na etiqueta	Descrição
Componentes do kit S2		A lâmina de fluxo S2 gera até 4,1 bilhões de leituras únicas que passam pelo filtro com produção de até 1000 Gb a 2 x 150 bp. A lâmina de fluxo S2 fornece sequenciamento rápido para a maioria das aplicações de alto rendimento.
Componentes do kit S4		A lâmina de fluxo S4 gera até 10 bilhões de leituras únicas que passam pelo filtro com produção de até 3.000 Gb a 2 x 150 bp. A lâmina de fluxo S4 é uma versão de quatro canais da lâmina de fluxo, projetada para saída máxima.

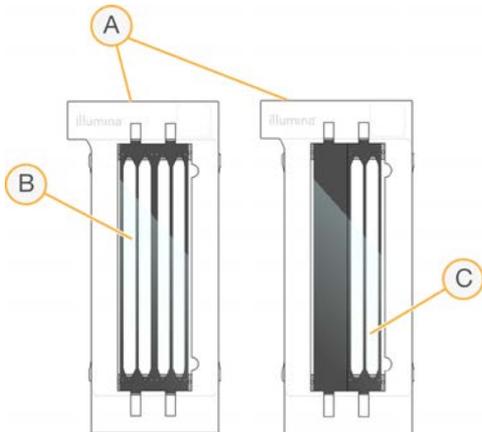
## Lâmina de fluxo

A lâmina de fluxo NovaSeq 6000Dx é uma lâmina de fluxo padronizada revestida por um cartucho. A lâmina de fluxo é um substrato com base em vidro contendo bilhões de nanoporos em um arranjo ordenado. Os clusters são gerados nos nanoporos, nos quais é, então, realizado o sequenciamento.

Cada lâmina de fluxo tem várias canais para fazer o sequenciamento de bibliotecas em pool. A lâmina de fluxo S2 possui duas canais e a lâmina de fluxo S4 possui quatro. Cada canal é captado em várias faixas e, em seguida, o software divide a imagem de cada faixa em porções menores chamadas de blocos.

Alguns arranhões e outros defeitos cosméticos menores na lâmina de fluxo são normais e não devem comprometer a qualidade dos dados e o rendimento. A Illumina recomenda usar essas lâminas de fluxo da forma normal.

Figura 1 Lâminas de fluxo



- A. Cartucho da lâmina de fluxo
- B. Lâmina de fluxo de quatro canais (S4)
- C. Lâmina de fluxo de duas canais (S2)

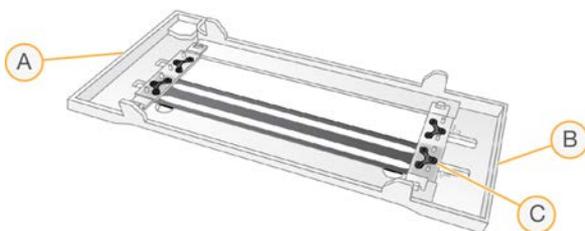
O lado inferior de cada lâmina de fluxo tem várias vedações. As bibliotecas e os reagentes entram nas canais da lâmina de fluxo através das vedações situadas na extremidade de entrada da lâmina de fluxo. Os reagentes usados são expelidos das canais através das vedações situadas na extremidade de saída.



### CUIDADO

Evite tocar nas vedações ao manusear a lâmina de fluxo.

Figura 2 Lâmina de fluxo invertida



- A. Extremidade de saída
- B. Extremidade de entrada
- C. Vedação (uma de quatro)

## Detalhes dos cartuchos de solução tampão, de cluster e de SBS

Os cartuchos de solução tampão, de cluster e de SBS do NovaSeq 6000Dx têm reservatórios com selo de alumínio pré-carregados com reagentes, soluções tampão e solução de limpeza. Os cartuchos de cluster e SBS estão incluídos nos kits de reagentes do NovaSeq 6000Dx. O cartucho de solução tampão é vendido separadamente.

Os cartuchos são carregados diretamente no instrumento, codificados por cores e rotulados para reduzir os erros de carregamento. As guias no refrigerador do reagente e as gavetas de solução tampão garantem a orientação adequada.

Tabela 4 Cartuchos do NovaSeq 6000Dx

Material de consumo	Descrição
Cartucho de solução tampão 	<p>Pré-carregado com soluções tampão de sequenciamento e pesa até 6,8 kg. Uma alça de plástico facilita o transporte, o carregamento e o descarregamento.</p> <p>Os cartuchos de solução tampão contêm reagentes que são sensíveis à luz. Mantenha o receptáculo de solução tampão dentro da embalagem até o uso.</p>
Cartucho de cluster 	<p>Pré-carregado com reagentes do tipo paired-end, de clusterização, de indexação, e de solução de limpeza. Inclui uma posição destinada ao tubo da biblioteca. A rotulagem cor de laranja distingue o cartucho de cluster do cartucho de SBS.</p> <p>Um reagente de desnaturação na posição 30 contém formamida, que é uma amida orgânica e uma toxina reprodutiva. Para facilitar a eliminação segura de qualquer reagente não utilizado após a corrida de sequenciamento, esse reservatório é removível.</p>
Cartucho de SBS 	<p>Pré-carregado com reagentes de sequenciamento em volumes específicos para o número de ciclos que o kit suporta. Cada uma das três posições de reagente tem uma posição adjacente reservada para a limpeza automática após a corrida. A rotulagem cinza diferencia o cartucho de SBS do cartucho de cluster.</p> <p>Os cartuchos de SBS contêm reagentes que são sensíveis à luz. Mantenha o receptáculo de SBS dentro da embalagem até o uso.</p>

## Reservatórios de cartucho de clusters reservados

Três reservatórios são dedicados para primers personalizados e uma posição vazia é dedicada para o tubo da biblioteca. Para a rastreabilidade de amostra, o tubo da biblioteca é carregado no cartucho de cluster durante a configuração de corrida e permanece com o cartucho até o final da corrida.

Figura 3 Reservatórios numerados

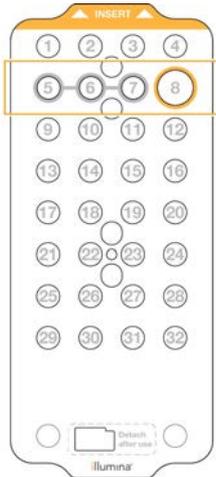


Tabela 5 Reservatórios de cartuchos de cluster

Posição	Reservada para
5, 6 e 7	Primers personalizados opcionais
8	Tubo da biblioteca

## Materiais de consumo e equipamentos fornecidos pelo usuário

Tabela 6 Materiais de consumo

Material de consumo	Fornecedor	Finalidade
Frasco de centrífuga, 500 ml	Fornecedor de itens de uso comum do laboratório	Diluição de Tween 20 para uma limpeza de manutenção.
Tubo de centrífuga, 30 ml	Fornecedor de itens de uso comum do laboratório	Diluição de NaOCl para uma limpeza de manutenção.

<b>Material de consumo</b>	<b>Fornecedor</b>	<b>Finalidade</b>
Luvas descartáveis, sem pó	Fornecedor de itens de uso comum do laboratório	Uso geral.
Lenços umedecidos com álcool isopropílico 70% ou Lenços umedecidos com álcool etanol 70%	VWR, catálogo n.º 95041-714 ou equivalente Fornecedor de itens de uso comum do laboratório	Componentes de limpeza antes de uma corrida e finalidade geral.
Lenço para laboratório com poucos fiapos	VWR, catálogo n.º 21905-026 ou equivalente	Secagem da mesa da lâmina de fluxo e para finalidade geral.
NaOCl com grau de reagente, 5%	Sigma-Aldrich, n.º do catálogo 239305	Realização de uma limpeza de manutenção.
Pontas de pipeta, 2 µl	Fornecedor de itens de uso comum do laboratório	Pipetagem para diluição e carregamento de bibliotecas.
Pontas de pipeta, 20 µl	Fornecedor de itens de uso comum do laboratório	Pipetagem para diluição e carregamento de bibliotecas.
Pontas de pipeta, 200 µl	Fornecedor de itens de uso comum do laboratório	Pipetagem para diluição e carregamento de bibliotecas.
Pontas de pipeta, 1.000 µl	Fornecedor de itens de uso comum do laboratório	Pipetagem para diluição e carregamento de bibliotecas.
Álcool isopropílico reagente ou de grau espectrofotométrico (99%), frasco de 100 ml	Fornecedor de itens de uso comum do laboratório	Limpeza dos componentes ópticos periodicamente e apoio ao cartucho de limpeza da objetiva.
Tween 20	Sigma-Aldrich, n.º de catálogo P7949	Realização de uma limpeza de manutenção.
Água, aprovada para uso em laboratório	Fornecedor de itens de uso comum do laboratório	Diluição de Tween 20 e de hipoclorito de sódio para uma limpeza de manutenção.

Tabela 7 Equipamento

Item	Origem
Congelador, -25 °C a -15 °C	Fornecedor de itens de uso comum do laboratório
Cilindro graduado, 500 ml, estéril	Fornecedor de itens de uso comum do laboratório
Balde de gelo	Fornecedor de itens de uso comum do laboratório
Pipeta, 20 µl	Fornecedor de itens de uso comum do laboratório
Pipeta, 200 µl	Fornecedor de itens de uso comum do laboratório
Pipeta, 1.000 µl	Fornecedor de itens de uso comum do laboratório
Refrigerador, 2 °C a 8 °C	Fornecedor de itens de uso comum do laboratório
Vasilhame, banhos-maria*	Fornecedor de itens de uso comum do laboratório

\*Use um vasilhame que possa acomodar dois cartuchos de reagentes e o nível de água apropriado. Por exemplo, (61 cm x 91,4 cm x 25,4 cm)(24 pol. x 36 pol. x 10 pol.).

## Orientações para água aprovada para uso em laboratório

Utilize sempre água aprovada para uso em laboratório ou água deionizada para realizar procedimentos com instrumentos. Nunca use água da torneira. Utilize apenas os seguintes tipos de água ou equivalentes:

- Água deionizada
- Illumina PW1
- Água de 18 Megaohms (MΩ)
- Água Milli-Q
- Água Super-Q
- Água para biologia molecular

## Instruções de uso

As instruções a seguir são para corrida do NovaSeq 6000Dx Instrument no modo IVD de operação usando as configurações do kit S2 ou S4.

## Criar uma corrida do sequenciamento

Use as etapas a seguir para criar uma corrida usando o Illumina Run Manager no modo IVD ou RUO. Alternativamente, selecione **Import Run** (Importar Corrida) na guia Planned (Planejada) da página Runs (Corridas) e importe uma planilha de amostra. Crie novas corridas no instrumento ou acessando o Illumina Run Manager usando um navegador em um computador em rede.

**OBSERVAÇÃO** As informações exatas exigidas por cada aplicação de análise são diferentes, mas o processo para criar uma corrida inclui as etapas a seguir.

1. Na guia Planned (Planejado) da tela Runs (Corridas), selecione **Create Run** (Criar corrida).
2. Selecione uma aplicação e selecione **Next** (Próximo).
3. Prossiga pelas telas de configurações. Dependendo da aplicação, as telas exibidas podem incluir o seguinte:
  - **Run Settings** (Configurações de corrida): Insira parâmetros de corrida.
  - **Sample Data** (Dados da amostra): Insira os dados da amostra manualmente ou importando um arquivo CSV contendo informações da amostra. Os nomes das amostras devem ser exclusivos.
  - **Analysis settings** (Configurações de análise): Insira configurações para análise.
4. Na tela Review (Revisar), revise as informações da corrida e selecione **Save** (Salvar). A corrida é adicionada à parte superior da lista de corridas na guia Planned (Planejado).

## Preparação de materiais de consumo

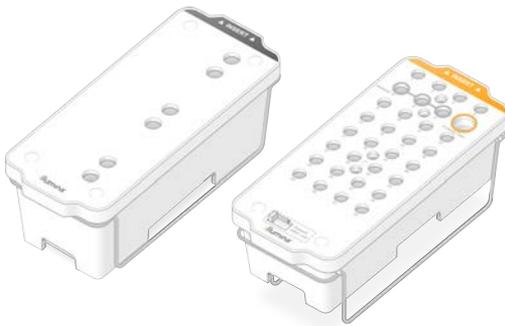
Descongelar os cartuchos de SBS e de cluster

### CUIDADO

Usar água quente para descongelar os reagentes pode causar redução na qualidade dos dados ou falha na corrida.

1. Se uma corrida de sequenciamento estiver em andamento, certifique-se de que ambos os lados do instrumento estejam disponíveis quando o descongelamento estiver concluído.
2. Remova os cartuchos de SBS e de cluster do armazenamento de -25 °C a -15 °C.
3. Coloque cada cartucho em um rack de descongelamento.  
Os racks são fornecidos com o instrumento e evitam que ele vire no banho-maria.

Figura 4 Cartuchos em racks de descongelamento



- Use a tabela a seguir para determinar a duração do descongelamento.

Descongele os cartuchos SBS e de cluster em banho-maria em temperatura ambiente (19 °C a 25 °C) conforme descrito abaixo. Mergulhe os cartuchos aproximadamente até a metade.

Cartucho	Duração do descongelamento
Cartucho de SBS S2	4 horas
Cartucho de cluster S2	Até 2 horas
Cartucho de SBS S4	4 horas
Cartucho de cluster S4	Até 4 horas



### CUIDADO

Não iniciar o sequenciamento em quatro horas após o início do descongelamento dos cartuchos de reagentes pode reduzir a qualidade dos dados.

- Seque completamente as bases do cartucho usando toalhas de papel. Seque entre os poços para que toda a água seja removida.
- Verifique se há presença de água nas vedações de alumínio. Se houver água, seque com um tecido sem fiapos.
- Inspecione o lado de baixo de cada cartucho para certificar-se de que os reservatórios estejam livres de gelo, o que indica que os reagentes estão descongelados.
- Inverta cada cartucho dez vezes para misturar os reagentes.



### CUIDADO

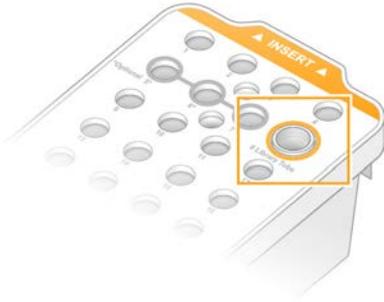
A não inversão completa dos cartuchos pode reduzir a qualidade dos dados.

- Bata levemente a parte inferior de cada cartucho na bancada para reduzir as bolhas de ar.

## Carregar tubo da biblioteca

- Sem mexer com a biblioteca na parte inferior, insira o tubo da biblioteca destampado que contém o pool de bibliotecas desnaturadas e diluídas na posição (n.º 8) do **tubo da biblioteca** do cartucho de cluster.
- Insira o tubo de biblioteca na posição n.º 8 do cartucho de cluster.

Figura 5 Tubo da biblioteca destampado carregado na posição n.º 8



## Esvaziar frascos de reagentes usados

Use as seguintes instruções para esvaziar os frascos de reagentes usados em *cada* corrida de sequenciamento. Se seu sistema estiver configurado para encaminhar externamente os reagentes usados, o frasco pequeno coletará reagentes usados e deverá ser esvaziado em cada corrida de sequenciamento. O frasco grande deve estar no lugar.

1. Remova e esvazie o frasco pequeno de reagente usado conforme as instruções a seguir.
  - a. Levante a alavanca e remova do nicho o frasco pequeno de reagente usado. Segure o frasco pelos lados.
  - b. Remova a tampa rosqueada do suporte da tampa na frente do frasco.
  - c. Vede a abertura do frasco com a tampa para evitar derramamentos.
  - d. Mantendo o conteúdo separado do conteúdo do outro frasco, descarte-o de acordo com as normas aplicáveis para sua região.
  - e. Devolva o frasco destampado para o nicho e, em seguida, abaixe a alavanca. Guarde a tampa no suporte da tampa.
2. Remova e esvazie o frasco grande de reagente usado conforme as instruções a seguir.
  - a. Utilizando a alça superior, remova o frasco grande de reagente usado pelo lado esquerdo da gaveta de solução tampão.
  - b. Remova a tampa rosqueada do suporte da tampa na frente do frasco.
  - c. Vede a abertura do frasco com a tampa para evitar derramamentos.
  - d. Descarte o conteúdo segundo as normas aplicáveis para sua região. Segure ambas as alças ao esvaziar.
  - e. Devolva o frasco destampado para a gaveta de solução tampão. Guarde a tampa no suporte da tampa.

Figura 6 Devolver o frasco vazio



3. Vista um novo par de luvas sem pó.

**CUIDADO**

Sempre coloque um novo par de luvas após manusear o frasco de reagente usado.

4. Feche a gaveta de solução tampão e, em seguida, feche as portas do compartimento de líquidos.

**CUIDADO**

O não esvaziamento dos frascos de reagentes usados pode resultar no encerramento de uma corrida e em transbordamento, o que prejudica o instrumento e representa um risco de segurança.

## Preparar a lâmina de fluxo

1. Retire um novo pacote da lâmina de fluxo do armazenamento de 2 °C a 8 °C.
2. Deixe o pacote da lâmina de fluxo vedado em temperatura ambiente (19 °C a 25 °C) por 10 a 15 minutos. Use a lâmina de fluxo em até 12 horas após retirá-la da embalagem.

## Carregar materiais de consumo

Siga as instruções para iniciar a configuração de corrida e carregar os materiais de consumo.

1. No menu principal, selecione **Sequence** (Sequenciar) e, em seguida, selecione uma corrida de lâmina de fluxo única ou dupla, conforme a seguir.
  - **A+B**: Configura uma corrida de lâmina de fluxo duplo.
  - **A**: Configura uma corrida de lâmina de fluxo simples no lado A.
  - **B**: Configura uma corrida de lâmina de fluxo simples no lado B.

O sistema inicia a configuração de corrida, começando com o carregamento da lâmina de fluxo.

2. Selecione **OK** para reconhecer o aviso e abrir a porta da lâmina de fluxo.



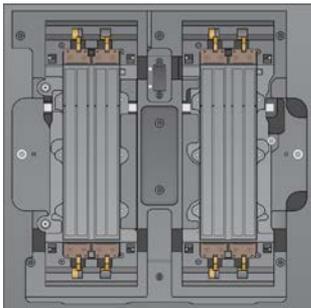
## CUIDADO

Mantenha a superfície limpa durante a corrida de sequenciamento e evite inclinar-se sobre o instrumento. A pressão na porta da lâmina de fluxo pode fazer com que a porta se abra, interrompendo a corrida. Corridas interrompidas não podem ser retomadas.

## Carregar a lâmina de fluxo

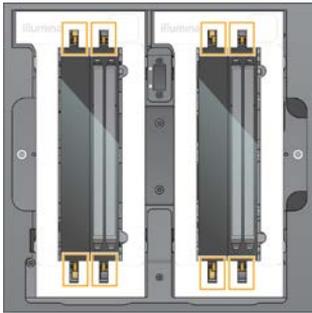
1. Se estiver presente, remova a lâmina de fluxo da corrida anterior.
2. Se houver material particulado visível na mesa da lâmina de fluxo, limpe toda a mesa, inclusive a interface fluídica e a superfície de vidro do alvo do alinhamento óptico com um lenço umedecido com álcool. Seque com um tecido sem fiapos.

Figura 7 Mesa da lâmina de fluxo



3. Remova a lâmina de fluxo da embalagem conforme abaixo.
  - a. Coloque um novo par de luvas sem pó para evitar contaminar a superfície de vidro da lâmina de fluxo.
  - b. Com a embalagem sobre uma superfície plana, abra puxando a película pela aba do canto.
  - c. Remova o retentor plástico transparente que cobre a lâmina de fluxo.
  - d. Remova a lâmina de fluxo da embalagem. Segure a lâmina de fluxo pelas laterais para evitar tocar o vidro ou as vedações inferiores.
  - e. Se houver material particulado visível em cada uma das superfícies do vidro, limpe a respectiva superfície com um lenço umedecido com álcool sem fiapos e seque com um lenço para laboratório que solte poucos fiapos.
  - f. Descarte a embalagem adequadamente.
4. Alinhe a lâmina de fluxo ao longo das quatro fixações levantadas e coloque-a na mesa da lâmina de fluxo.

Figura 8 Lâminas de fluxo carregadas alinhadas sobre fixações



5. Selecione **Close Flow Cell Door** (Fechar porta da lâmina de fluxo).

A porta da lâmina de fluxo se fecha, os sensores e o RFID são verificados e a identificação da lâmina de fluxo é exibida na tela.

## Carregar os cartuchos de SBS e de cluster

1. Abra as portas de compartimentos de líquidos e, em seguida, abra a porta do refrigerador de reagentes.
2. Remova os cartuchos de SBS e de cluster usados, se presentes de uma corrida anterior.  
Os cartuchos usados têm selos de alumínio perfurados.
3. Descarte o conteúdo não utilizado de acordo com as normas aplicáveis.  
Para um descarte seguro da posição n.º 30 do cartucho de cluster, consulte [Remover a posição n.º 30 na página 21](#).
4. Carregue os cartuchos preparados na gaveta do refrigerador de reagentes conforme abaixo, de modo que as etiquetas Insert (Inserir) fiquem voltadas para a parte traseira do instrumento.
  - Coloque o cartucho de SBS (etiqueta cinza) na posição esquerda.
  - Coloque o cartucho de cluster (etiqueta laranja) que contém o tubo da biblioteca destampado na posição direita.

Figura 9 Cartuchos de reagente carregados



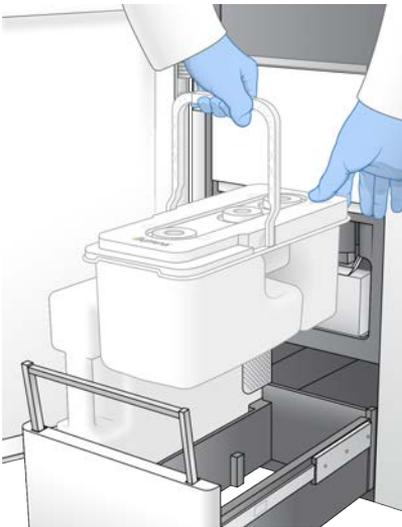
5. Deslize a gaveta para dentro do refrigerador e, em seguida, feche a porta do refrigerador de reagentes.  
Os sensores e os RFIDs são verificados. As identificações do tubo da biblioteca e os dois cartuchos são exibidos na tela.

## Carregar o cartucho de solução tampão

1. Puxe a alça de metal para abrir a gaveta de solução tampão.
2. Remova o cartucho de solução tampão usado pelo lado direito da gaveta de solução tampão.  
O cartucho de solução tampão usado tem selos de alumínio perfurados.
3. Coloque um novo cartucho de solução tampão na gaveta de solução tampão de modo que a etiqueta Illumina fique voltada para a frente da gaveta. Alinhe o cartucho com as guias levantadas no piso e laterais da gaveta.

Quando carregado adequadamente, o cartucho de solução tampão é encaixado uniformemente e a gaveta pode ser fechada.

Figura 10 Carregar o cartucho de solução tampão



4. Se ambos os frascos de reagentes tiverem sido esvaziados, marque a caixa de seleção que informa que ambos os frascos de reagentes usados estão vazios.

**OBSERVAÇÃO** O não esvaziamento dos frascos de reagentes usados pode resultar no encerramento de uma corrida e em transbordamento, o que prejudica o instrumento e representa um risco de segurança.

5. Quando os materiais de consumo tiverem sido adicionados, selecione **Run Selection** (Seleção de corrida) para continuar.

## Selecionar e iniciar a corrida

O instrumento faz a leitura do ID do tubo da biblioteca e pesquisa uma corrida planejada correspondente.

1. Se uma corrida planejada correspondente à ID do tubo da biblioteca for encontrada para cada lado sendo usado, a seleção da corrida será ignorada. Selecione **Review** (Revisar) para continuar.

2. Se não houver corrida correspondente para um ou ambos os lados, selecione **Run Selection** (Seleção de corrida) e selecione uma ou mais corridas planejadas.  
A mesma corrida planejada não pode ser selecionada em ambos os lados.
3. Quando uma ou mais corridas forem selecionadas, selecione **Pre-Run Checks** (Verificações pré-corrída).
4. Aguarde cerca de 5 minutos até que as verificações pré-corrída sejam concluídas.  
A corrida é iniciada automaticamente após uma conclusão bem-sucedida.

**OBSERVAÇÃO** Para evitar encher demais o disco rígido, não copie dados para o C:\ após o início da corrida.

## Erros de verificação pré-corrída

1. Se as verificações pré-corrída falharem devido a um erro do sensor, como lâmina de fluxo não detectada, você precisará sair e reiniciar o fluxo de trabalho.
2. Para outras falhas na verificação pré-corrída, selecione **Retry** (Repetir) para reiniciar a verificação com falha ou **Retry All** (Repetir tudo) para reiniciar todas as verificações.  
A presença de erros exige resolução antes do início da corrida.
3. Selecione o ícone **Error** (Erro) para ver os detalhes do erro.
4. Se a verificação de alinhamento falhar, resolva o erro conforme a seguir.
  - a. Selecione **Reload** (Recarregar) e, em seguida, selecione **OK** para confirmar o retorno à tela Load (Carregar).
  - b. Remova todos os itens da parte de cima do instrumento e, em seguida, selecione **OK**. A porta da lâmina de fluxo é aberta.
  - c. Recarregue a lâmina de fluxo e, em seguida, selecione **Run Setup** (Configuração de corrida).
  - d. Continue em cada tela para reler todos os RFIDs e retorne à tela Pre-Run Checks (Verificações pré-corrída).
  - e. Refaça a verificação.

## Monitorar o progresso da corrida

Os seguintes detalhes são exibidos na tela Sequencing (Sequenciamento) enquanto a corrida está em andamento. A tela Sequencing (Sequenciamento) é acessada por meio do menu principal.

- **Status das etapas da corrida individual**
- **Time to completion** (Hora da conclusão): A data e a hora da conclusão da corrida (aaaa-mm-dd hh:mm).
- **Run progress** (Progresso da corrida): A etapa atual da corrida. O tamanho da barra de progresso não é proporcional à taxa de corrida de cada etapa.
- **Q-Scores**: a distribuição das classificações de qualidade (Q-Scores).

- **Intensity** (Intensidade): o valor das intensidades de cluster do 90º percentil para cada bloco. As cores do gráfico indicam os canais vermelho e verde.
- **Clusters passing filter (%)** (Filtro de passagem de clusters): a porcentagem da passagem de clusters pelo filtro.
- **Projected Total Yield (Gb)** (Rendimento total projetado): o rendimento projetado para a corrida da lâmina de fluxo. Se as métricas por canal forem selecionadas (H), os números exibidos indicarão o rendimento atual por canal e serão atualizados para cada ciclo durante a corrida.
- **Q30**: a porcentagem de identificações de bases para a corrida que tiverem um Q-Score  $\geq 30$ .

## Ícones de status

Um ícone de status na interface do NVOS indica o status da corrida. Um número no ícone indica o número de condições para um status.

Quando um status da corrida muda, o ícone pisca para alertar você. Selecione o ícone para visualizar uma descrição da condição. Selecione **Acknowledge** (Reconhecer) para apagar a mensagem e, em seguida, **Close** (Fechar) para fechar a caixa de diálogo.

Ícone de status	Nome do status	Descrição
	Status OK	O sistema está normal.
	Processing (Processando)	O sistema está em processamento.
	Warning (Advertência)	Ocorreu uma advertência e é necessária atenção. Advertências não interrompem uma corrida nem exigem uma ação antes de continuar.
	Error (Erro)	Ocorreu um erro. Erros exigem uma ação antes de continuar a corrida.
	Information (Informações)	Uma mensagem não crítica está disponível.

## Métricas da corrida

O software exibe métricas geradas durante a corrida. As métricas são exibidas em forma de gráficos, gráficos de pontos e tabelas com base nos dados gerados pela RTA3 e gravados nos arquivos InterOp.

A clusterização leva cerca de duas horas; depois o sequenciamento é iniciado com o ciclo 1. As métricas são atualizadas conforme o sequenciamento continua. A passagem de clusters pelo filtro, o rendimento e as pontuações de qualidade estão disponíveis após o ciclo 26. Antes do ciclo 26, não há valores preenchidos e é exibida a designação N/A.

## Após o sequenciamento

As seções a seguir fornecem instruções sobre as etapas que acontecem após a conclusão do sequenciamento.

### Limpeza automática após a corrida

Depois da conclusão do sequenciamento, o software inicia automaticamente a limpeza após a corrida, que leva cerca de 80 minutos. O sistema bombeia hipoclorito de sódio (NaOCl) 0,24% na posição n.º 17 e o dilui a 0,12%. O NaOCl 0,12% é bombeado para o reagente ExAmp e para as posições da biblioteca, por meio da lâmina de fluxo, e, em seguida, para os frascos de reagentes usados. A limpeza lava o modelo do sistema para evitar contaminação cruzada.

Após a conclusão da limpeza, o sistema é colocado em um estado seguro, e o botão Home (Início) fica ativo. Deixe os materiais de consumo no lugar até a próxima corrida. Após a limpeza, os aspiradores de líquidos permanecem nos cartuchos de SBS e de cluster para impedir a entrada de ar no sistema. Os aspiradores de líquidos no cartucho de solução tampão são levantados para que os frascos de reagentes usados possam ser esvaziados. O tampão de limpeza é então bombeado através de todas as linhas para limpar NaOCl e reagentes do sistema.

**OBSERVAÇÃO** Se ocorrer um erro durante uma limpeza automática após a corrida, e se a limpeza após a corrida for incompleta, será necessária uma limpeza de manutenção.

### Remover a posição n.º 30

O reservatório da posição n.º 30 do cartucho de cluster contém formamida. Ele é removido do cartucho de cluster usado e descartado separadamente.



#### CUIDADO

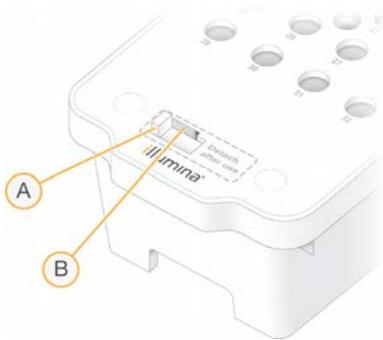
**Esse conjunto de reagentes contém produtos químicos potencialmente perigosos. Podem ocorrer lesões por inalação, ingestão e contato com a pele ou com os olhos. Use equipamento de proteção, incluindo proteção para os olhos, luvas e jaleco adequado para risco de exposição. Manuseie os reagentes usados como resíduo químico e descarte-os de acordo com as leis e regulamentações regionais, nacionais e locais aplicáveis. Para obter mais informações ambientais, de saúde e de segurança, consulte a SDS em [support.illumina.com/sds.html](https://support.illumina.com/sds.html).**

1. Usando luvas, empurre para a direita a aba de plástico branca rotulada **Detach after use** (Remover após o uso).

2. Coloque uma das mãos ou uma superfície sob o reservatório e pressione a aba de plástico transparente em direção à etiqueta da Illumina para soltar o reservatório de sua posição sob o cartucho de cluster.

**OBSERVAÇÃO** Evite empilhar os cartuchos de cluster ao armazenar. O empilhamento pode provocar deslocamento acidental do reservatório.

Figura 11 Posição removível n.º 30



- A. Aba de plástico branca para remover
- B. Aba de plástico transparente para soltar

3. Descarte o reservatório de acordo com as normas aplicáveis.

## Geração de dados de sequenciamento

Durante o sequenciamento, os dados são transferidos automaticamente do NovaSeq 6000Dx Instrument para o Illumina DRAGEN Server. Quando a análise primária termina e a transferência de dados é concluída, a análise secundária no Illumina DRAGEN Server pode começar automaticamente usando as opções de análise definidas pela aplicação selecionada no Illumina Run Manager. Os resultados produzidos dependem das opções escolhidas durante a configuração de corrida. Para visualizar os resultados de uma corrida, selecione o nome da corrida desejada na guia Completed (Concluída) na tela Runs (Corridas). Você também pode encontrar arquivos de dados gerados no local especificado na tela Instrument Settings (Configurações do instrumento).

### Real-Time Analysis

O NovaSeq 6000Dx Instrument executa o RTA3, uma implementação do software Real-Time Analysis, no Compute Engine (CE). O RTA3 extrai as intensidades das imagens recebidas da câmera, realiza identificações de bases, atribui uma pontuação de qualidade às identificações de bases, alinha a PhiX e relata dados em arquivos InterOp.

Para otimizar o tempo de processamento, o RTA3 armazena informações na memória. Se o RTA3 for desligado, o processamento não será restabelecido e qualquer dado de corrida processado na memória será perdido.

## Entradas do RTA3

O RTA3 requer imagens de blocos contidas na memória local do sistema para processamento. O RTA3 recebe informações da corrida e comandos do NVOS.

## Saídas do RTA3

As imagens de cada canal de cor são passadas na memória para o RTA3 como blocos. Nessas imagens, o RTA3 gera um conjunto de arquivos de identificação de bases com pontuação de qualidade e arquivos de filtro. Todas as outras saídas são arquivos de saída de apoio.

Tipo de arquivo	Descrição
Arquivos de identificação de bases	Cada bloco analisado é incluído em um arquivo de identificação de bases concatenado (*.cbcl). Os blocos do mesmo canal e superfície são agregados em um arquivo CBCL para cada canal e superfície.
Arquivos de filtro	Cada bloco produz um arquivo de filtro (*.filter) que especifica se um cluster passa pelos filtros.

O RTA3 fornece métricas em tempo real da qualidade da corrida armazenadas como arquivos InterOp, que são uma saída binária que contém métricas de blocos, ciclos e nível de leitura.

## Tratamento de erros

O RTA3 cria arquivos de registro e os grava na pasta Logs (Registros). Os erros são gravados em um arquivo de texto em formato de arquivo \*.log.

Os arquivos de registro a seguir são transferidos para o destino de saída final ao término do processamento:

- `info_00000.log` resume eventos importantes da corrida.
- `error_00000.log` lista erros que ocorreram durante uma corrida.
- `warning_00000.log` lista avisos que ocorreram durante uma corrida.

## Blocos da lâmina de fluxo

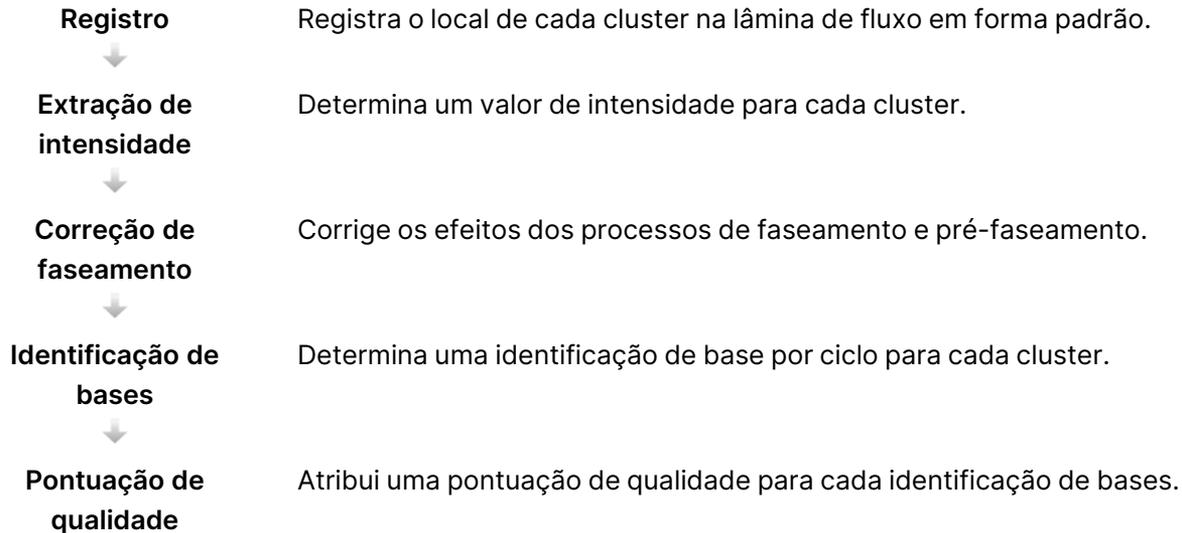
Os blocos são pequenas áreas de imagem na lâmina de fluxo. A câmera captura uma imagem de cada faixa, que é dividida pelo software em blocos para processamento no RTA3. O número total de blocos depende de quantos canais, faixas e superfícies são captadas na lâmina de fluxo.

- As lâminas de fluxo S2 têm um total de 1.408 blocos.
- As lâminas de fluxo S4 têm um total de 3.744 blocos.

Componente da lâmina de fluxo	S2	S4	Descrição
Canais	2	4	É um caminho físico com portas de entrada e de saída.
Superfícies	2	2	As lâminas de fluxo S2 e S4 são captadas em duas superfícies: superior e inferior. A imagem da superfície superior de um bloco é captada primeiro.
Faixas por canal	4	6	Uma faixa é uma coluna em um canal da lâmina de fluxo que a câmera capta como uma imagem.
Blocos por faixa	88	78	Um bloco é uma porção de uma faixa e apresenta uma área com imagens na lâmina de fluxo.
Total de blocos gerados	1.408	3.744	Canais × superfícies × faixas × blocos por faixa é igual ao número total de blocos.

O nome do bloco é um número de cinco dígitos que representa a posição do bloco na lâmina de fluxo. Por exemplo, o nome do bloco 1\_1205 indica canal 1, superfície superior, faixa 2, bloco 5.

- O primeiro dígito é o número da canal:
  - 1 ou 2 para uma lâmina de fluxo S2.
  - 1, 2, 3 ou 4 para uma lâmina de fluxo S4.
- O segundo dígito representa a superfície: 1 para superior e 2 para inferior.
- O terceiro dígito representa o número da faixa:
  - 1, 2, 3 ou 4 para uma lâmina de fluxo S2.
  - 1, 2, 3, 4, 5 ou 6 para uma lâmina de fluxo S4.
- Os últimos dois dígitos representam o número do bloco. A numeração começa com 01 na extremidade de saída da lâmina de fluxo e vai até 88 ou 78 na extremidade de entrada.
  - 01 a 88 para uma lâmina de fluxo S2.
  - 01 a 78 para uma lâmina de fluxo S4.

**Fluxo de trabalho do Real-Time Analysis****Registro**

O registro alinha uma imagem à matriz quadrada rotacionada de nanopozos na lâmina de fluxo, em forma padrão. Devido ao arranjo ordenado dos nanopozos, as coordenadas X e Y para cada cluster em um bloco são predeterminadas. As posições do cluster são gravadas em um arquivo local do cluster (s.locs) para cada corrida.

Se o registro falhar para qualquer imagem em um ciclo, nenhuma identificação de bases será gerada para o bloco no ciclo.

**Extração de intensidade**

Após o registro, a extração de intensidade calcula um valor de intensidade para cada nanopozo em uma determinada imagem. Se o registro falhar, a intensidade para aquele bloco não poderá ser extraída.

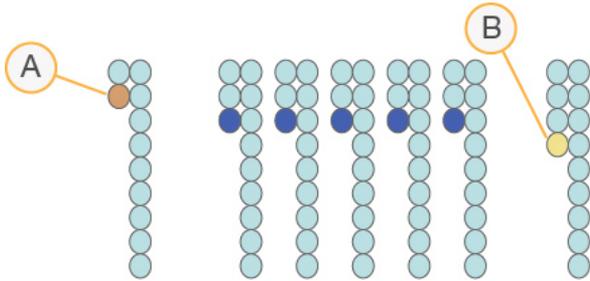
**Correção de faseamento**

Durante a reação de sequenciamento, cada fita de DNA em um cluster se estende em uma base por ciclo. Os processos de faseamento e pré-faseamento ocorrem quando uma fita fica fora de fase com o ciclo de incorporação atual.

O faseamento ocorre quando uma incorporação de base fica para trás.

O pré-faseamento ocorre quando uma incorporação base se adianta.

Figura 12 Faseamento e pré-faseamento



- A. Leitura com uma base em faseamento
- B. Leitura com uma base em pré-faseamento.

O RTA3 corrige os efeitos do faseamento e do pré-faseamento, o que potencializa a qualidade dos dados em cada ciclo ao longo da corrida.

### Identificação de bases

A identificação de bases determina uma base (A, C, G ou T) para cada cluster de um determinado bloco em um ciclo específico. O NovaSeq 6000Dx Instrument utiliza sequenciamento de dois canais, o que requer duas imagens para codificar os dados de quatro bases de DNA, uma imagem do canal vermelho e outra do canal verde.

A ausência de identificação é designada como N. Isso ocorre quando um cluster não passa pelo filtro, o registro falha ou um cluster é deslocado para fora da imagem.

As intensidades para cada cluster são extraídas das imagens vermelhas e verdes e comparadas entre si, o que resulta em quatro populações distintas. Cada população corresponde a uma base. O processo de identificação de bases determina a que população cada cluster pertence.

Figura 13 Visualização de intensidades de clusters

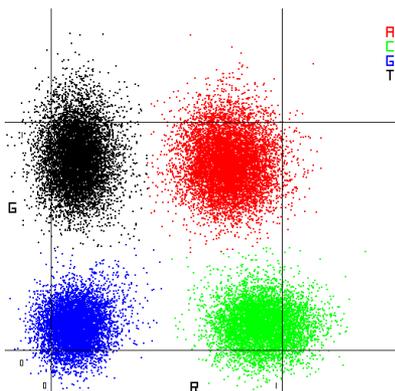


Tabela 8 Identificações de bases em sequenciamento de dois canais

Base	Canal vermelho	Canal verde	Resultado
A	1 (ligado)	1 (ligado)	Clusters que mostram intensidade nos canais vermelho e verde.
C	1 (ligado)	0 (desligado)	Clusters que mostram intensidade apenas no canal vermelho.
G	0 (desligado)	0 (desligado)	Clusters que não mostram nenhuma intensidade em um local de cluster conhecido.
T	0 (desligado)	1 (ligado)	Clusters que mostram intensidade apenas no canal verde.

### Clusters passando por filtros

Durante a corrida, o RTA3 filtra os dados brutos para remover leituras que não estão de acordo com o limite de qualidade dos dados. Clusters de sobreposição e de baixa qualidade são removidos.

Para a análise de dois canais, o RTA3 usa um sistema de base populacional para determinar a pureza (medida de pureza de intensidade) de uma identificação de bases. Os clusters passam pelo filtro (PF) quando não mais que uma identificação de bases nos primeiros 25 ciclos tem uma pureza inferior a um limite fixo. Quando incluído, o alinhamento PhiX é realizado no ciclo 26 em um subconjunto de blocos para clusters que passaram pelo filtro. Os clusters que não passam pelo filtro não passam pelo processo de identificação de bases nem são alinhados.

### Pontuações de qualidade

Uma pontuação de qualidade (Q-Score) é uma previsão da probabilidade de uma identificação de bases errada. Um Q-score mais alto indica que a identificação de bases tem mais qualidade e probabilidade de estar correta. Após a determinação do Q-score, os resultados são registrados em arquivos CBCL.

O Q-Score sucintamente comunica pequenas probabilidades de erro. As pontuações de qualidade são representadas como Q(X), onde X é a pontuação. A tabela a seguir mostra a relação entre uma pontuação de qualidade e a probabilidade de erro.

Q-Score Q(X)	Probabilidade de erro
Q40	0,0001 (1 em 10.000)
Q30	0,001 (1 em 1.000)
Q20	0,01 (1 em 100)
Q10	0,1 (1 em 10)

### Pontuação de qualidade e relatórios

A pontuação de qualidade calcula um conjunto de prognosticadores para cada identificação de bases e usa esses valores para consultar o Q-score em uma tabela de qualidade. Tabelas de qualidade são criadas para fornecer previsões de qualidade com precisão ideal para corridas geradas por uma configuração específica de plataforma de sequenciamento e versão de química.

A pontuação de qualidade se baseia em uma versão modificada do algoritmo Phred.

Para gerar a tabela Q para o NovaSeq 6000Dx Instrument, foram determinados três grupos de identificações de bases, com base na clusterização desses recursos preditivos específicos. Após o agrupamento das identificações de bases, calculou-se empiricamente a taxa média de erros para cada um dos três grupos, e os Q-scores correspondentes foram registrados na Q-table com as funções preditivas correlacionadas àquele grupo. Desse modo, há somente três Q-scores possíveis com RTA3, e esses Q-scores representam a taxa média de erros do grupo. Em geral, isso resulta em uma pontuação de qualidade simplificada, porém altamente precisa. Os três grupos da tabela de qualidade correspondem a identificações de bases marginais (< Q15), médias (~Q20) e de alta qualidade (> Q30), que recebem pontuações específicas de 12, 26 e 34, respectivamente. Além disso, atribui-se uma pontuação nula igual a 2 a todos os não codificantes. Esse modelo de relatório de Q-score reduz os requisitos de espaço de armazenamento e largura de banda sem afetar a precisão e o desempenho.

Figura 14 Pontuação Q-score simplificada com RTA3



## Arquivos de dados gerados no sequenciamento

Tipo de arquivo	Descrição, local e nome do arquivo
Arquivos de identificação de bases	Cada cluster analisado é incluído em um arquivo de identificação de bases, agregado em um arquivo por ciclo, canal e superfície. O arquivo agregado contém a identificação de bases e a pontuação de qualidade codificada. Data\Intensities\BaseCalls\L001\C1.1 L[ <i>lane</i> ][_ <i>surface</i> ].cbcl, por exemplo L001_1.cbcl
Arquivos de localização de cluster	Para cada lâmina de fluxo, um arquivo de localização do cluster binário contém as coordenadas XY para os clusters em um bloco. Um layout hexagonal que corresponde ao layout do nanopoço da lâmina de fluxo predefine as coordenadas. Data\Intensities (Dados\Intensidades) s_[ <i>lane</i> ].locs
Arquivos de filtro	O arquivo de filtro especifica se um cluster passou por filtros. Os arquivos de filtro são gerados no ciclo 26 usando 25 ciclos de dados. Um arquivo de filtro é gerado para cada bloco. Data\Intensities\BaseCalls\L001 s_[ <i>lane</i> ][_ <i>tile</i> ].filter
Arquivo de informações da corrida	Lista o nome da corrida, o número de ciclos em cada leitura, se a leitura é uma leitura de índice e o número de faixas e blocos da lâmina de fluxo. O arquivo de informações da corrida é criado no início da corrida. [Pasta principal],RunInfo.xml
Arquivos de miniaturas	Imagens em miniatura para o primeiro ciclo de cada leitura de sequenciamento. Thumbnail_Images\L001\C[X.1]: Os arquivos são armazenados em uma subpasta para cada ciclo. s_[ <i>lane</i> ][_ <i>tile</i> ][_ <i>channel</i> ].jpg: A imagem em miniatura inclui o número do bloco.

## Estrutura da pasta de gravação de dados de sequenciamento

O NVOS gera o nome da pasta de saída automaticamente.

 **Config:** definições de configuração para a corrida.

 **Logs:** arquivos de log que descrevem etapas operacionais, análise de instrumentos e eventos do RTA3.

 Samplesheet.csv: planilha de amostra ou outro arquivo anexado, se aplicável.

 **Data** (Dados)

 **Intensities** (Intensidades)

 **BaseCalls** (Identificações de bases)

 **L00[X]:** arquivos de identificação de bases (\*.cbcl) agregados em um arquivo por canal, superfície e ciclo.

 s.locs: o arquivo dos locais do cluster para a corrida.

 **InterOp**: arquivos binários.

 **Recipe** (Receita): arquivo de receita específico para a corrida.

 **Thumbnail Images** (Imagens em miniatura): imagens em miniatura para cada 10º bloco.

 **LIMS**: arquivo (\*.json) de configuração da corrida, se aplicável.

 **Audit (Auditoria)**

 AuditInfo.xml

 RTA3.cfg

 RunInfo.xml

 RunParameters.xml

 RTAComplete.txt

 CopyComplete.txt

 SequenceComplete.txt

 IlluminaRunManagerCopyComplete.txt

 Manifest.tsv

## Advertências e precauções



### CUIDADO

A lei federal restringe este dispositivo a ser vendido por ou mediante a ordem de um médico ou outro profissional da área médica licenciado pela lei do estado no qual atua, para usar ou designar o uso do dispositivo.

- **Alguns componentes dos reagentes fornecidos pela Illumina para uso com o NovaSeq 6000Dx Instrument contêm substâncias químicas potencialmente perigosas. Podem ocorrer lesões por inalação, ingestão e contato com a pele ou com os olhos. Use equipamento de proteção, incluindo proteção para os olhos, luvas e jaleco adequado para risco de exposição. Manuseie os reagentes usados como resíduo químico e descarte-os de acordo com as leis e regulamentações regionais, nacionais e locais aplicáveis.** Para obter informações ambientais, de saúde e segurança adicionais, consulte as fichas de dados de segurança (SDS) em [support.illumina.com/sds.html](https://support.illumina.com/sds.html).
- Se os procedimentos não forem seguidos conforme o estabelecido, poderá haver resultados com erro ou uma redução significativa na qualidade da amostra.
- Adote precauções de rotina do laboratório. Não utilize a pipeta com a boca. Não coma, beba ou fume em áreas de trabalho designadas. Use luvas descartáveis e jalecos para laboratório ao manusear espécimes e kits de reagentes. Lave bem as mãos depois de manusear espécimes e kits de reagentes.

- Práticas adequadas e boa higiene no laboratório são necessárias para evitar que os produtos de PCR contaminem reagentes, instrumentos e amostras de DNA genômico. A contaminação por PCR pode causar resultados inexatos e não confiáveis.
- Para evitar a contaminação, certifique-se de que as áreas de pré-amplificação e pós-amplificação possuam equipamentos e materiais de consumo exclusivos (p. ex., pipetas, pontas de pipeta, blocos de calor, agitadores e centrífuga).
- O índice para o emparelhamento de amostras deve corresponder exatamente ao layout da placa impressa. O DNA Prep with Enrichment Application preenche automaticamente os primers de índice associados aos nomes das amostras, quando inseridos durante a configuração de corrida. O usuário é aconselhado a verificar os primers de índice associados às amostras antes de iniciar a corrida de sequenciamento. A incompatibilidade entre a planilha de amostras e o layout da placa provoca a perda da identificação positiva da amostra e a geração de relatórios com resultados incorretos.
- A instalação de software antivírus fornecido pelo usuário é altamente recomendável para proteger o computador contra vírus.
- Não opere o NovaSeq 6000Dx se um dos painéis tiver sido removido. Operar o instrumento sem um dos painéis pode causar exposição à tensão de linha e às tensões CC.
- Não toque na mesa da lâmina de fluxo no compartimento da lâmina de fluxo. O aquecedor desse compartimento opera entre 22 °C e 95 °C e pode causar queimaduras.
- O instrumento pesa aproximadamente 481 kg e pode causar ferimentos graves se cair ou for manuseado de forma imprópria.

## Características de desempenho

As características de desempenho do NovaSeq 6000Dx Instrument foram estabelecidas usando o Illumina DNA Prep with Enrichment Dx para preparação da biblioteca, o NovaSeq 6000Dx S2 Reagent v1.5 Kit (300 cycles) e o NovaSeq 6000Dx S4 Reagent v1.5 Kit (300 cycles) para sequenciamento e o DRAGEN for Illumina DNA Prep with Enrichment Dx Application para análise secundária, incluindo detecção de variantes de linha genética e somáticas. Os estudos incluíram indexação de amostras, transferência de amostras, entrada de DNA, sensibilidade analítica (limite de branco/ limite de detecção), exatidão, precisão, comparação de métodos e reprodutibilidade. Consulte o *folheto informativo do Illumina DNA Prep with Enrichment Dx* para obter características de desempenho relacionadas a fatores pré-analíticos, como métodos de extração ou substâncias interferentes.

## Definições de cálculos usados nas características de desempenho

1. O percentual de concordância positiva (PPA) é calculado como a proporção de locais classificados como variantes por um método de referência que são corretamente relatados pelo ensaio.

- $(\text{N.º de locais de variantes corretamente relatados pelo ensaio}) / (\text{N.º total de locais de variantes})$   
Os locais de variantes relatados pelo ensaio que são concordantes com o método de referência são positivos verdadeiros (TPs, True Positives). Os locais de variantes relatados como identificações de referência ou como identificações de variantes diferentes pelo ensaio são falsos negativos (FNs, False Negatives).
2. O percentual de concordância negativa (NPA) é calculado como a proporção de locais classificados como tipo selvagem por um método de referência que são corretamente relatados pelo ensaio.
    - $(\text{N.º de locais do tipo selvagem corretamente relatados pelo ensaio}) / (\text{n.º total de locais do tipo selvagem})$   
Os locais do tipo selvagem relatados pelo ensaio que são concordantes com o método de referência são negativos verdadeiros (TNs, True Negatives). Os locais do tipo selvagem relatados como variantes pelo ensaio são falsos positivos (FPs).
  3. O percentual geral de concordância (OPA, Overall Percent Agreement) é calculado como a proporção de locais corretamente relatados pelo ensaio, relativos a um método de referência.
    - $([\text{N.º de locais de variantes corretamente relatados pelo ensaio}] + [\text{N.º de locais do tipo selvagem corretamente relatados pelo ensaio}]) / ([\text{N.º total de locais de variantes}] + [\text{N.º total de locais do tipo selvagem}])$
  4. Os cálculos de PPA, NPA e OPA não incluem não identificações (locais de variantes ou de referência que não atendem a um ou mais filtros de qualidade).
  5. A porcentagem de identificações positivas (PPC) é o número de observações com a variante detectada dividido pelo número total de observações testadas, excluindo quaisquer observações inválidas ou aquelas filtradas como de baixa profundidade.
  6. A porcentagem de identificações negativas (PNC) é calculada como o número de observações com referência de passagem como o resultado em uma posição dividido pelo número total de observações testadas, excluindo quaisquer observações inválidas ou aquelas filtradas como de baixa profundidade.
  7. A porcentagem de capacidade de identificação de autossomo é calculada como a porcentagem de posições de referência não N em regiões-alvo em cromossomos autossômicos com uma identificação de genótipo de passagem.

## Indexação de amostras

Os primers de índice de amostras, adicionados durante a preparação das bibliotecas, atribuem uma sequência única a cada amostra de DNA. Essas sequências únicas possibilitam que várias amostras sejam agrupadas em uma única corrida de sequenciamento. A indexação de amostras é usada para os fluxos de trabalho de linha genética e somático. O objetivo deste estudo foi estabelecer o número mínimo (12) e máximo (192) de amostras que podem ser processadas em uma única corrida de sequenciamento pelo NovaSeq 6000Dx Instrument. 12 amostras únicas de DNA Platinum Genome (NA12877–NA12888) foram testadas com pelo menos 12 combinações diferentes de primer de índice por amostra. As bibliotecas de amostras foram preparadas usando um ensaio representativo projetado para consultar uma variedade de genes cobrindo 1.970.505 bases em

todos os 23 cromossomos humanos. Os resultados das amostras de quatro corridas de sequenciamento usando o fluxo de trabalho de análise de geração de linha genética FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx Application foram comparados com Platinum Genomes versão 2016-1.0.

Para o primeiro conjunto de corridas, 192 bibliotecas de amostras indexadas exclusivamente foram sequenciadas em duas corridas do sequenciamento, uma cada com reagentes S2 e S4, para verificar o número máximo de índices suportados e a capacidade do ensaio de fazer uma identificação de genotipagem de forma consistente para uma determinada amostra em diferentes combinações de primers de índice. Para o segundo conjunto de corridas, 12 bibliotecas de amostras indexadas exclusivamente foram sequenciadas em duas corridas do sequenciamento, uma com reagentes S2 e S4, para verificar o número mínimo de índices suportados.

Para as corridas de 192 índices, o PPA para SNVs (single nucleotide variant, variante de nucleotídeo único) variou de 99,7% a 100%, o PPA para inserções foi de 100%, o PPA para deleções variou de 96,7% a 100% e o NPA foi de 100%. Para as corridas de 12 índices, o PPA para SNVs variou de 99,7% a 100%, o PPA para inserções variou de 89,6% a 100%, o PPA para deleções variou de 94,6% a 100% e o NPA foi de 100%.

## Transferência de amostras

O NovaSeq 6000Dx Instrument permite o sequenciamento de várias amostras e de controles em uma única corrida de sequenciamento. Foi realizado um estudo para avaliar a extensão da transferência de amostras em uma corrida de sequenciamento (dentro da corrida) e entre corridas do sequenciamento (de corrida para corrida). 12 amostras de DNA Platinum Genome, seis masculinas e seis femininas, foram testadas com um ensaio representativo projetado para consultar uma variedade de genes cobrindo 1.970.505 bases em todos os 23 cromossomos humanos, incluindo os cromossomos de ambos os sexos. As bibliotecas foram sequenciadas no NovaSeq 6000Dx Instrument usando o fluxo de trabalho de análise de geração de linha genética FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx Application. A transferência de amostras masculinas para amostras femininas foi observada pela presença de leituras alvo do cromossomo Y em amostras femininas.

A transferência dentro da corrida pode ser introduzida durante a clusterização, identificação de base do ciclo de índice e demultiplexação de amostra. Para testes de transferência de amostras em uma corrida de sequenciamento, um pool de bibliotecas consistindo em pelo menos 12 réplicas de cada amostra masculina e feminina exclusiva mais dois controles sem modelo, para um total de 192 bibliotecas indexadas exclusivamente, foi sequenciado no NovaSeq 6000Dx Instrument em duas corridas do sequenciamento, uma cada com reagentes S2 e S4. O transporte de amostras durante a corrida foi avaliado comparando a cobertura alvo do cromossomo Y de cada réplica feminina com a cobertura média do alvo do cromossomo Y de todas as réplicas masculinas no pool. O 95º percentil de transferência observada durante a corrida foi de 0,0090% e 0,041% para reagentes S2 e S4, respectivamente.

Para testar o transferência de amostras entre corridas, dois pools de bibliotecas foram preparados e sequenciados consecutivamente em um NovaSeq 6000Dx Instrument, com o lado A usando reagentes S4 e o lado B usando reagentes S2. O primeiro pool continha pelo menos 12 réplicas de seis amostras femininas exclusivas mais dois controles sem modelo, para um total de 96 bibliotecas indexadas exclusivamente. O segundo pool continha pelo menos 12 réplicas de seis amostras masculinas exclusivas mais dois controles sem

modelo, para um total de 96 bibliotecas indexadas exclusivamente. Ambos os pools usaram o mesmo conjunto de adaptadores de índice. O pool do sexo feminino foi sequenciado primeiro, seguido de uma corrida subsequente do sequenciamento com o pool do sexo masculino, seguido de outra corrida de sequenciamento repetida do pool do sexo feminino. A transferência de amostras entre corridas foi avaliada por tipo de reagente, S2 e S4, comparando a cobertura alvo do cromossomo Y entre as réplicas correspondentes da corrida repetida do pool feminino e da corrida do pool masculino. O 95º percentil de transferência de corrida para corrida observada foi de 0,0089% e 0,012% para reagentes S2 e S4, respectivamente.

## Entrada de DNA

### Sangue (linha genética)

A faixa de entrada de DNA sanguíneo para o kit Illumina DNA Prep with Enrichment Dx usando o DRAGEN for Illumina DNA Prep with Enrichment Dx Application foi estabelecida para o NovaSeq 6000Dx. Isto foi avaliado através da realização de um estudo de diluição em série usando oito amostras de DNA do Platinum Genome (NA12877 - NA12884) com um ensaio representativo projetado para consultar uma variedade de genes cobrindo 1.970.505 bases em todos os 23 cromossomos humanos. As bibliotecas foram sequenciadas em um NovaSeq 6000Dx Instrument usando um lote cada de NovaSeq 6000Dx S2 Reagent v1.5 Kit (300 cycles) e NovaSeq 6000Dx S4 Reagent v1.5 Kit (300 cycles).

Sete amostras foram testadas em duplicata em seis níveis de entrada de DNA variando de 1.000 ng a 10 ng (1.000 ng, 250 ng, 100 ng, 50 ng, 25 ng e 10 ng). Uma oitava amostra (NA12884) foi testada como uma única réplica a 10 ng de entrada e em duplicata para todos os outros níveis de entrada. Para a determinação da precisão, os genótipos da amostra foram comparados com os Platinum Genomes versão 2016-1.0. Os resultados foram determinados para cada nível de entrada. O PPA para cada tipo de variante (SNVs, inserções e deleções) é apresentado nos [Resultados de PPA para cada entrada de DNA sanguíneo por tipo de variante na página 34](#). O NPA é apresentado em [NPA para cada entrada de DNA sanguíneo na página 35](#). Todos os níveis de entrada apresentaram exatidão semelhante. A entrada de DNA sanguíneo recomendada para o Illumina DNA Prep with Enrichment Dx é de 50 a 1.000 ng, com 1.000 ng e 10 ng, fornecendo um limite superior e inferior para atender às características de desempenho quando sequenciado no NovaSeq 6000Dx.

Tabela 9 Resultados de PPA para cada entrada de DNA sanguíneo por tipo de variante

Entrada de DNA (ng)	Tipo de variante	Variantes prováveis	TP	FN	Não identificações de variante	PPA (%)
10	SNV	69612	69538	68	6	99,9
25		74192	74105	75	12	99,9
50		74105	74	13	99,9	
100		74116	72	4	99,9	
250		74113	72	7	99,9	
1.000		74112	73	7	99,9	

Entrada de DNA (ng)	Tipo de variante	Variantes prováveis	TP	FN	Não identificações de variante	PPA (%)
10	Inserção	2732	2732	0	0	100
25		2928	2916	6	6	99,8
50		2914	8	6	99,7	
100		2917	6	5	99,8	
250		2928	0	0	100	
1.000		2.921	5	2	99,8	
10	Deleção	2084	2049	4	31	99,8
25		2240	2200	9	31	99,6
50		2207	3	30	99,9	
100		2199	1	40	>99,9	
250		2201	0	39	100	
1.000		2195	2	43	99,9	

Tabela 10 NPA para cada entrada de DNA sanguíneo

Entrada de DNA (ng)	TN	FP	Não identificações de referência	NPA (%)
10	115449045	384	285751	>99,9
25	123012157	415	438153	>99,9
50	122985299	369	465043	>99,9
100	122976660	321	473730	>99,9
250	122971099	331	479289	>99,9
1.000	122978527	324	471882	>99,9

## FFPE (somático)

A faixa de entrada de DNA fixado em formalina e embebido em parafina (FFPE) para o Illumina DNA Prep with Enrichment Dx Kit usando o DRAGEN for Illumina DNA Prep with Enrichment Dx Application foi estabelecida para o NovaSeq 6000Dx. Esse intervalo foi avaliado com a realização de um estudo de diluição em série com o uso de duas amostras de Platinum Genomes com um ensaio representativo criado para consultar vários genes, cobrindo 1.970.505 bases em 23 cromossomos humanos diferentes. As bibliotecas foram sequenciadas em um NovaSeq 6000Dx Instrument usando um lote cada de NovaSeq 6000Dx S2 Reagent v1.5 Kit (300 cycles) e NovaSeq 6000Dx S4 Reagent v1.5 Kit (300 cycles).

A amostra de DNA GM12877 foi diluída com amostra de DNA GM12878 para criar GM12877-13 com variantes heterozigóticas e homozigóticas exclusivas de GM12877 em frequências próximas de 6,5% e 13%, respectivamente. O GM12877 não diluído também foi testado. GM12877-13 foi testado em duplicata em quatro

níveis de entrada de DNA variando de 1.000 ng a 25 ng (1.000 ng, 250 ng, 50 ng e 25 ng). O GM12877 foi testado como uma única réplica a 250 ng e em duplicata para todos os outros níveis de entrada. Para a determinação da precisão, as identificações das variantes da amostra foram comparadas com os Platinum Genomes versão 2016-1.0. Os resultados foram determinados para cada nível de entrada. O PPA para cada tipo de variante (SNVs, inserções e deleções) é apresentado nos [Resultados de PPA para cada entrada de DNA FFPE por tipo de variante e VAF alvo na página 36](#). O NPA é apresentado em [NPA para cada entrada de DNA FFPE na página 36](#). Todos os níveis de entrada apresentaram exatidão semelhante. Para amostras FFPE com valor  $\Delta Cq \leq 5$ , a entrada de DNA recomendada é de 50 a 1.000 ng para o Illumina DNA Prep with Enrichment Dx Kit com 1.000 ng e 25 ng, fornecendo um limite superior e inferior para atender às características de desempenho quando sequenciado no NovaSeq 6000Dx.

Tabela 11 Resultados de PPA para cada entrada de DNA FFPE por tipo de variante e VAF alvo

Entrada de DNA (ng)	Tipo de variante	Variantes prováveis	VAF de diluição-alvo								
			0,065					0,13			
			TP	FN	Não identificações de variante	PPA (%)	Variantes prováveis	TP	FN	Não identificações de variante	PPA (%)
25	SNV	3.000	2.931	8	61	99,7	624	624	0	0	100
50		3.000	2.930	8	62	99,7	624	622	0	2	100
250		3.000	2.927	8	65	99,7	624	624	0	0	100
1.000		3.000	2.921	8	71	99,7	624	624	0	0	100
25	Inserção	96	96	0	0	100	48	48	0	0	100
50		96	96	0	0	100	48	48	0	0	100
250		96	96	0	0	100	48	48	0	0	100
1.000		96	96	0	0	100	48	48	0	0	100
25	Deleção	88	88	0	0	100	32	32	0	0	100
50		88	88	0	0	100	32	31	0	1	100
250		88	88	0	0	100	32	32	0	0	100
1.000		88	88	0	0	100	32	32	0	0	100

Tabela 12 NPA para cada entrada de DNA FFPE

Entrada de DNA (ng)	Tipo selvagem provável	TN	FP	Não identificações de referência	NPA (%)
25	25354119	25353706	413	5499498	>99,9
50	27538269	27538013	256	3315421	>99,9
250	21562303	21561983	320	1577958	>99,9
1.000	29030903	29030596	307	1822781	>99,9

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

Este estudo foi conduzido para avaliar o limite de branco (LoB) e o limite de detecção (LoD) para o fluxo de trabalho de análise de geração somática FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx Application no NovaSeq 6000Dx Instrument. O estudo foi realizado usando um ensaio representativo projetado para consultar uma variedade de genes cobrindo 1.970.505 bases em todos os 23 cromossomos humanos. As linhagens celulares GM12878 e GM12877 do Platinum Genome foram fixadas em formalina e embebidas em parafina seguidas pela extração de DNA. Diluições de GM12877 em GM12878 foram preparadas para produzir amostras consistindo em 0%, 4%, 6,5% e 13% de GM12877 por volume, de modo que as frequências de variantes de 489 variantes exclusivas de GM12877 (454 SNVs, 17 inserções e 18 deleções) variaram entre 0 e 0,13. As bibliotecas de amostras foram preparadas usando dois lotes de reagentes Illumina DNA Prep with Enrichment Dx Kit e sequenciadas durante seis dias consecutivos de início com dois NovaSeq 6000Dx Instruments e dois lotes cada de NovaSeq 6000Dx S2 Reagent v1.5 Kit (300 cycles) e NovaSeq 6000Dx S4 Reagent v1.5 Kit (300 cycles), para um total de 12 corridas do sequenciamento. Isso resultou em 288 observações para cada variante em cada uma das diluições da amostra. LoB e LoD foram calculados usando a abordagem clássica indicada no CLSI EP17-A2. LoB e LoD foram calculados para reagentes S2 e S4 separadamente, agrupando as frequências de variantes de todas as variantes na corrida de sequenciamento para cada tipo de reagente. O erro do tipo I foi definido como 0,01 e o erro do tipo II foi definido como 0,05.

O LoB foi avaliado para 489 loci de forma independente em dois lotes de sequenciamento para cada tipo de reagente (S2 ou S4) e preparação da biblioteca. Para reagentes S2, o LoB do 95º percentil foi de 2,9%. Para reagentes S4, o LoB do 95º percentil foi de 2,2%.

O LoD foi calculado com sucesso para 478 de 489 variantes para S2 e 485 de 489 variantes para S4. As variantes em que nenhum LoD foi determinado para uma ou ambas as preparações da biblioteca foram excluídas da atribuição final de LoD para o sistema NovaSeq 6000Dx. O LoD do sistema NovaSeq 6000Dx com reagentes S2 e S4 foi determinado pelo 95º percentil dos LoDs de variantes individuais. Para reagentes S2, o 95º percentil em 478 LoDs variantes foi de 4,8%. Para reagentes S4, o 95º percentil em 485 LODs variantes foi de 3,9%.

# Exatidão

## Linha genética

O estudo a seguir foi conduzido para avaliar a precisão da identificação da variante do fluxo de trabalho da análise de geração de linha genética FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx no NovaSeq 6000Dx Instrument usando o NovaSeq 6000Dx S2 Reagent v1.5 Kit (300 cycles). Quatro amostras exclusivas do DNA Platinum Genome foram testadas usando um ensaio representativo projetado para consultar uma variedade de genes que cobrem 1.970.505 bases (9.232 alvos) em todos os 23 cromossomos humanos. Cada uma das amostras foi testada em réplicas de 12, exceto para NA12880, que foram testadas em réplicas de 11. Um total de 18 corridas foram realizadas utilizando três instrumentos de sequenciamento, três lotes de reagentes S2 e dois operadores durante seis dias de início. A precisão foi determinada para SNVs, inserções e deleções comparando os resultados com o Platinum Genomes versão 2016-1.0.

Tabela 13 Resumo de concordância da linha genética

Critérios	Total de observações <sup>1</sup>	Resultado das observações <sup>2</sup>	Resultado por corrida <sup>3</sup>
PPA para SNV	846	99,8	99,9
PPA para inserções	846	97,9	>99,9
PPA para deleções	846	96,9	99,9
NPA	846	>99,9	>99,9
OPA	846	>99,9	>99,9

<sup>1</sup>Calculado como número de amostras por corrida (47) x número de corridas (18) = 846.

<sup>2</sup>Valor mais baixo observado por réplica de amostra em todas as 18 corridas.

<sup>3</sup>Valor mais baixo quando os dados de cada corrida são analisados em conjunto.

[Concordância da linha genética por amostra na página 39](#) contém os dados do estudo apresentados com concordância percentual positiva e negativa por amostra, em que os resultados das variantes são comparados com o Platinum Genomes versão 2016-1.0 para cálculos de PPA. Os três tipos de variante (SNVs, inserções e deleções) são combinados. Como o método de referência só fornece resultados para as variantes de nucleotídeo único e inserções/deleções, os resultados de base não variantes são comparados com sequências de referências do genoma humano versão hg19 para os cálculos do NPA.

Tabela 14 Concordância da linha genética por amostra

Amostra	Capacidade de identificação de autossomo	Variantes prováveis <sup>1</sup>	TP	FN	Não identificações de variante	TN	FP	PPA	NPA	OPA
NA12877	99,4	273672	273452	220	0	414765131	931	99,9	>99,9	>99,9
NA12878	99,4	265680	265208	234	238	414803691	1193	99,9	>99,9	>99,9
NA12879	99,4	261792	261792	0	0	414746986	1429	100	>99,9	>99,9
NA12880	99,4	246114	245551	399	164	380157538	1458	99,8	>99,9	>99,9

<sup>1</sup> Número total de variantes em todas as réplicas de amostras em 18 corridas.

[Concordância da linha genética por amostra por tipo de variante na página 39](#) contém os dados do estudo apresentados por amostra, em que os resultados da variante são comparados com o método de referência composto bem caracterizado. A detecção é avaliada para cada tipo de variante: SNVs, inserções e deleções, separadamente. As posições de referência são excluídas.

Tabela 15 Concordância da linha genética por amostra por tipo de variante

Amostra	SVNs			Inserções			Deleções		
	Provável	TP	FN	Provável	TP	FN	Provável	TP	FN
NA12877	255096	254877	219	10368	10367	1	8208	8208	0
NA12878	250344	250077	221	8424	8424	0	6912	6707	13
NA12879	246024	246024	0	8856	8856	0	6912	6912	0
NA12880	229482	229086	396	9306	9306	0	7326	7159	3

As amostras foram analisadas novamente para a identificação de pequenas inserções e deleções (indels). Um resumo geral é apresentado no [Resumo da detecção de indels na linha genética na página 40](#). Foi encontrado um total de 210 indels variando em tamanho de 1 a 18 pb para inserções e de 1 a 21 pb para deleções.

Tabela 16 Resumo da detecção de indels na linha genética

Tipo de variante	Variantes prováveis	TP	FN	Não identificações de variante	PPA
Inserção	36954	36953	1	0	>99,9
Deleção	29358	28986	16	356	99,9

O ensaio representativo consistiu em 9.232 alvos cobrindo uma variedade de conteúdo genômico. O conteúdo de GC dos alvos variou de 0,20 a 0,86. Os alvos também tinham uma faixa de repetições de nucleotídeo único (por exemplo, PoliA, PoliT), dinucleotídeo e trinucleotídeo. Os dados compilados por cromossomo para determinar o efeito do conteúdo genômico em porcentagem de identificações corretas são apresentados em [Exatidão no nível do cromossomo de linha genética na página 40](#). O percentual de identificações corretas consiste em identificações de variância e de referência e é inferior a 100% no caso de identificações incorretas ou não identificações.

Tabela 17 Exatidão no nível do cromossomo de linha genética

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr1	47	728	138328	Poli A (12), Poli C (7), Poli T (14), Poli G (7), Dinucleotídeo (22), Trinucleotídeo (8), Inserção (18), Deleção (4)	[0,22 - 0,8]; Mediana: 0,51	114888718	34	966860	>99,9	0,83

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr2	39	628	159588	Poli A (46), Poli C (8), Poli T (23), Poli G (7), Dinucleotídeo (22), Trinucleotídeo (8), Inserção (5), Deleção (2)	[0,24 - 0,81]; Mediana: 0,44	132293464	798	460345	>99,9	0,35
chr3	38	650	137627	Poli A (18), Poli C (6), Poli T (18), Poli G (7), Dinucleotídeo (12), Trinucleotídeo (6), Inserção (11), Deleção (1)	[0,25 - 0,86]; Mediana: 0,45	114625053	2	226461	>99,9	0,20

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr4	17	370	73766	Poli A (9), Poli C (7), Poli T (25), Poli G (6), Dinucleotídeo (5), Trinucleotídeo (5), Inserção (2), Deleção (2)	[0,27 - 0,77]; Mediana: 0,45	61872303	0	66741	100	0,11
chr5	25	507	90008	Poli A (10), Poli C (6), Poli T (12), Poli G (7), Dinucleotídeo (10), Trinucleotídeo (8), Inserção (8), Deleção (18)	[0,29 - 0,79]; Mediana: 0,46	75314497	912	153061	>99,9	0,20

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr6	39	453	126721	Poli A (28), Poli C (7), Poli T (33), Poli G (7), Dinucleotídeo (18), Trinucleotídeo (11), Inserção (4), Deleção (2)	[0,24 - 0,79]; Mediana: 0,48	103412695	1	182361	>99,9	0,18
chr7	21	450	161501	Poli A (27), Poli C (8), Poli T (21), Poli G (7), Dinucleotídeo (31), Trinucleotídeo (5), Inserção (1), Deleção (4)	[0,2 - 0,77]; Mediana: 0,46	132534074	19	246884	>99,9	0,19

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr8	18	381	67775	Poli A (19), Poli C (7), Poli T (13), Poli G (7), Dinucleotídeo (5), Trinucleotídeo (9), Inserção (4), Deleção (1)	[0,26 - 0,78]; Mediana: 0,47	56247612	411	170925	>99,9	0,30
chr9	23	347	87100	Poli A (12), Poli C (7), Poli T (27), Poli G (8), Dinucleotídeo (9), Trinucleotídeo (9), Inserção (4), Deleção (1)	[0,27 - 0,83]; Mediana: 0,49	72650800	20	241991	>99,9	0,33
chr10	14	317	66723	Poli A (26), Poli C (7), Poli T (15), Poli G (6), Dinucleotídeo (16), Trinucleotídeo (6), Inserção (1), Deleção (1)	[0,23 - 0,78]; Mediana: 0,44	55539058	1	188216	>99,9	0,34

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr11	29	511	91786	Poli A (28), Poli C (8), Poli T (21), Poli G (7), Dinucleotídeo (26), Trinucleotídeo (7), Inserção (2), Deleção (2)	[0,28 - 0,8]; Mediana: 0,47	75744222	742	259258	>99,9	0,34
chr12	29	577	120365	Poli A (19), Poli C (8), Poli T (40), Poli G (7), Dinucleotídeo (7), Trinucleotídeo (7), Inserção (1), Deleção (5)	[0,26 - 0,77]; Mediana: 0,49	99972530	1	542005	>99,9	0,54
chr13	13	283	58639	Poli A (24), Poli C (6), Poli T (12), Poli G (7), Dinucleotídeo (6), Trinucleotídeo (8), Inserção (14), Deleção (0)	[0,28 - 0,79]; Mediana: 0,42	48503179	1	45666	>99,9	0,09

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr14	11	147	26980	Poli A (21), Poli C (6), Poli T (18), Poli G (11), Dinucleotídeo (6), Trinucleotídeo (6), Inserção (4), Deleção (1)	[0,29 - 0,77]; Mediana: 0,47	22286153	198	147895	>99,9	0,66
chr15	15	266	52091	Poli A (26), Poli C (7), Poli T (13), Poli G (6), Trinucleotídeo (8), Inserção (4), Deleção (6)	[0,29 - 0,76]; Mediana: 0,46	43600279	0	99041	100	0,23
chr16	21	366	80030	Poli A (7), Poli C (7), Poli T (15), Poli G (7), Dinucleotídeo (5), Trinucleotídeo (10), Inserção (15), Deleção (21)	[0,3 - 0,76]; Mediana: 0,54	65490245	16	1438278	>99,9	2,15

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr17	36	645	118062	Poli A (19), Poli C (7), Poli T (18), Poli G (8), Dinucleotídeo (13), Trinucleotídeo (6), Inserção (18), Deleção (16)	[0,28 - 0,82]; Mediana: 0,49	97929929	417	335905	>99,9	0,34
chr18	9	99	19195	Poli A (7), Poli C (7), Poli T (15), Poli G (6), Trinucleotídeo (10), Inserção (4), Deleção (0)	[0,22 - 0,78]; Mediana: 0,44	15967171	312	42077	>99,9	0,26
chr19	30	605	104004	Poli A (19), Poli C (7), Poli T (31), Poli G (7), Dinucleotídeo (5), Trinucleotídeo (7), Inserção (2), Deleção (21)	[0,33 - 0,83]; Mediana: 0,59	85642066	3	678213	>99,9	0,79

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr20	12	179	33795	Poli A (6), Poli C (6), Poli T (7), Poli G (8), Trinucleotídeo (9), Inserção (5), Deleção (0)	[0,31 - 0,84]; Mediana: 0,53	28108712	0	38374	100	0,14
chr21	5	63	30642	Poli A (28), Poli C (6), Poli T (24), Poli G (7), Dinucleotídeo (5), Inserção (2), Deleção (5)	[0,22 - 0,78]; Mediana: 0,52	25319736	50	57434	>99,9	0,23
chr22	10	187	36727	Poli A (26), Poli C (7), Poli T (19), Poli G (7), Dinucleotídeo (5), Trinucleotídeo (6), Inserção (6), Deleção (0)	[0,27 - 0,74]; Mediana: 0,51	30258131	0	42673	100	0,14

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chrX	23	433	83576	Poli A (18), Poli C (8), Poli T (23), Poli G (9), Dinucleotídeo (5), Trinucleotídeo (23), Inserção (3), Deleção (0)	[0,2 - 0,72]; Mediana: 0,48	67318722	0	770544	100	1,13
chrY	0	40	5476	Poli A (11), Poli C (8), Poli T (11), Poli G (5), Inserção (0), Deleção (0)	[0,4 - 0,59]; Mediana: 0,45	0	0	0	N/A	N/A

Os resultados do sequenciamento da amostra NA12878 foram comparados com um genótipo de alta confiança para NA12878, estabelecido pelo National Institutes of Standards and Technology (NIST) (v.2.19). Dos 9.232 alvos, 8.009 alvos estavam totalmente contidos nas regiões genômicas altamente confiáveis, 776 alvos tinham sobreposição parcial e 447 alvos não tinham sobreposição na sequência do NIST. Isso resultou em 1.831.483 coordenadas por réplica para comparação. Identificações de base não variantes foram comparadas com sequências de referências do genoma humano versão hg19. Os resultados de precisão são mostrados na [Concordância da linha genética da amostra NA12878 com o banco de dados do NIST na página 49](#).

Tabela 18 Concordância da linha genética da amostra NA12878 com o banco de dados do NIST

Amostra	Nº de alvos cobertos	Capacidade de identificação de autossomo	TP	FN	TN	FP	PPA	NPA	OPA
NA12878	8785	99,4	247709	218	394262149	4584	>99,9	>99,9	>99,9

Com base nos dados fornecidos por este estudo de linha genética de 18 corridas, o NovaSeq 6000Dx Instrument pode sequenciar de forma consistente:

- Conteúdo de GC  $\geq$  20% (todas as bases identificadas em 1.692 regiões-alvo sequenciadas com 20% de conteúdo de GC identificado corretamente com taxa de não identificação de 0%).
- Conteúdo de GC  $\leq$  86% (todas as bases identificadas em 846 regiões-alvo sequenciadas com 86% de conteúdo de GC identificado corretamente com taxa de não identificação de 0%).
- Comprimentos de PoliA  $\leq$  46 (todas as bases identificadas em 846 regiões-alvo sequenciadas com 46 repetições de PoliA identificadas corretamente com taxa de não identificação de 0,27%).
- Comprimentos de PoliT  $\leq$  40 (13.384.074 de 13.384.321 bases identificadas em 846 regiões-alvo sequenciadas com 40 repetições de PoliT identificadas corretamente com taxa de não identificação de 0,26%).
- Comprimentos de PoliG  $\leq$  11 (todas as bases identificadas em 846 regiões-alvo sequenciadas com 11 repetições de PoliG identificadas corretamente com taxa de não identificação de 0%).
- Comprimentos de PoliC  $\leq$  8 (9.815.030 de 9.815.035 bases identificadas em 5.922 regiões-alvo sequenciadas com 8 repetições de PoliC identificadas corretamente com taxa de não identificação de 0,53%).
- Comprimentos de repetição de dinucleotídeos  $\leq$  31x (32.233.922 de 32.233.926 bases identificadas em 846 regiões-alvo sequenciadas com 31 repetições de dinucleotídeos identificadas corretamente com taxa de não identificação de 0,21%).
- Comprimentos de repetição de trinucleotídeos  $\leq$  23x (todas as bases identificadas em 846 regiões-alvo sequenciadas com 23 repetições de trinucleotídeos identificadas corretamente com taxa de não identificação de 0,21%).
- Comprimentos de inserção  $\leq$  18 (todas as bases identificadas em 1.692 regiões-alvo sequenciadas com 18 inserções identificadas corretamente com taxa de não identificação de 7,71%).
- Comprimentos de deleção  $\leq$  21 (todas as bases identificadas em 1.692 regiões-alvo sequenciadas com 21 deleções identificadas corretamente com taxa de não identificação de 1,14%).

## Somática

O estudo descrito aqui foi usado para avaliar a precisão da identificação de variantes do fluxo de trabalho de análise de geração somática FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx Application no NovaSeq 6000Dx Instrument usando o NovaSeq 6000Dx S4 Reagent v1.5 Kit (300 cycles).

Este estudo utilizou um ensaio representativo projetado para consultar uma variedade de genes cobrindo 1.970.505 bases (9.232 alvos) em todos os 23 cromossomos humanos. Foi extraído o DNA Platinum Genome de blocos tratados com FFPE para gerar quatro amostras exclusivas para avaliação no estudo.

A amostra de DNA GM12877 foi diluída com amostra de DNA GM12878 para criar GM12877-13 com variantes heterozigóticas e homozigóticas exclusivas de GM12877 em frequências próximas de 6,5% e 13%, respectivamente. A amostra de DNA GM12878 foi diluída de forma semelhante com a amostra de DNA GM12877 para criar GM12878-13 com variantes heterozigóticas e homozigóticas únicas de GM12878 em frequências próximas de 6,5% e 13%, respectivamente. GM12877 e GM12878 não diluídos também foram testados. Cada uma das amostras foi testada em réplicas de 12, exceto para GM12878 não diluído, que foi testado em réplicas de 11. Um total de 18 corridas foram realizadas utilizando três instrumentos de sequenciamento, três lotes de reagentes S4 e dois operadores durante seis dias de início. A precisão foi determinada para SNVs, inserções e deleções comparando os resultados com o Platinum Genomes versão 2016-1.0.

Tabela 19 Resumo da concordância somática

<b>Critérios</b>	<b>N.º de observações<sup>1</sup></b>	<b>Resultado das observações<sup>2</sup></b>	<b>Resultado por corrida<sup>3</sup></b>
PPA para SNVs somáticas	846	99,8	98,9
PPA para inserções somáticas	846	100	100
PPA para deleções somáticas	846	100	100
NPA	846	>99,9	>99,9
OPA	846	>99,9	>99,9

<sup>1</sup> Calculado como = Nº de amostras por corrida (47) x Nº de corridas (18) = 846.

<sup>2</sup> Valor mais baixo observado por amostra replicada em todas as 18 corridas.

<sup>3</sup> Valor mais baixo quando os dados de cada corrida são analisados em conjunto.

[Concordância somática por amostra na página 52](#) contém os dados do estudo apresentados com concordância percentual positiva e negativa por amostra, em que os resultados das variantes são comparados com o método de referência composto bem caracterizado para cálculos de PPA. Os três tipos de variante (SNVs, inserções e deleções) são combinados. Como o método de referência só fornece resultados para as variantes de nucleotídeo único e inserções/deleções, os resultados de base não variantes são comparados com sequências de referências do genoma humano versão hg19 para os cálculos do NPA.

Tabela 20 Concordância somática por amostra

Amostra	Capacidade de identificação de autossomo	Variantes prováveis	TP	FN	Não identificações de variante	TN	FP	PPA	NPA	OPA
GM12877	95,4	96228	95022	198	1008	365425810	1203	99,8	>99,9	>99,9
GM12878	94,5	96768	96278	0	490	395002023	1278	100	>99,9	>99,9
GM12877-13	94,7	104976	103029	216	1731	395989324	1286	99,8	>99,9	>99,9
GM12878-13	95,2	96768	96027	0	741	397900884	1218	100	>99,9	>99,9

*Concordância somática por amostra por tipo de variante na página 52* contém os dados do estudo apresentados por amostra, em que os resultados das variantes são comparados com o método de referência composto bem caracterizado. A detecção é avaliada para cada tipo de variante: SNVs, inserções e deleções, separadamente. As posições de referência são excluídas.

Tabela 21 Concordância somática por amostra por tipo de variante

Amostra	SNVs			Inserções			Deleções		
	Provável	TP	FN	Provável	TP	FN	Provável	TP	FN
GM12877	89694	88488	198	3564	3564	0	2970	2970	0
GM12878	92664	92390	0	2160	2160	0	1944	1728	0
GM12877-13	97848	95901	216	3888	3888	0	3240	3240	0
GM12878-13	92664	92139	0	2160	2160	0	1944	1728	0

As quatro amostras foram analisadas novamente para a identificação de pequenas inserções e deleções (indels). Um resumo geral é apresentado no *Resumo da detecção de indels somáticos na página 53*. Foi encontrado um total de 210 indels variando em tamanho de 1 a 18 pb para inserções e de 1 a 21 pb para deleções.

Tabela 22 Resumo da detecção de indels somáticos

Tipo de variante	Variantes prováveis	TP	FN	Não identificações de variante	PPA
Inserção	11772	11772	0	0	100
Deleção	10098	9666	0	432	100

O ensaio representativo consistiu em 9.232 alvos cobrindo uma variedade de conteúdo genômico. O conteúdo de GC dos alvos variou de 0,20 a 0,86. Os alvos também tinham uma faixa de repetições de nucleotídeo único (por exemplo, PoliA, PoliT), dinucleotídeo e trinucleotídeo. Os dados compilados por cromossomo para determinar o efeito do conteúdo genômico na porcentagem de identificações corretas são apresentados em [Exatidão no nível de cromossomo somático na página 53](#). O percentual de identificações corretas consiste em identificações de variância e de referência e é inferior a 100% no caso de identificações incorretas ou não identificações.

Tabela 23 Exatidão no nível de cromossomo somático

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr1	47	728	138328	Poli A (12), Poli C (7), Poli T (14), Poli G (7), Dinucleotídeo (22), Trinucleotídeo (8), Inserção (3), Deleção (0)	[0,22 - 0,8]; Mediana: 0,51	110145939	52	5642613	>99,9	4,9

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr2	39	628	159588	Poli A (46), Poli C (8), Poli T (23), Poli G (7), Dinucleotídeo (22), Trinucleotídeo (8), Inserção (5), Deleção (1)	[0,24 - 0,81]; Mediana: 0,44	126795713	842	5850393	>99,9	4,4
chr3	38	650	137627	Poli A (18), Poli C (6), Poli T (18), Poli G (7), Dinucleotídeo (12), Trinucleotídeo (6), Inserção (1), Deleção (1)	[0,25 - 0,86]; Mediana: 0,45	109902527	593	4889226	>99,9	4,3
chr4	17	370	73766	Poli A (9), Poli C (7), Poli T (25), Poli G (6), Dinucleotídeo (5), Trinucleotídeo (5), Inserção (0), Deleção (1)	[0,27 - 0,77]; Mediana: 0,45	59373461	16	2517412	>99,9	4,1

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr5	25	507	90008	Poli A (10), Poli C (6), Poli T (12), Poli G (7), Dinucleotídeo (10), Trinucleotídeo (8), Inserção (8), Deleção (18)	[0,29 - 0,79]; Mediana: 0,46	72261191	723	3116981	>99,9	4,1
chr6	39	453	126721	Poli A (28), Poli C (7), Poli T (33), Poli G (7), Dinucleotídeo (18), Trinucleotídeo (11), Inserção (0), Deleção (1)	[0,24 - 0,79]; Mediana: 0,48	98593101	687	4890221	>99,9	4,7

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr7	21	450	161501	Poli A (27), Poli C (8), Poli T (21), Poli G (7), Dinucleotídeo (31), Trinucleotídeo (5), Inserção (1), Deleção (4)	[0,2 - 0,77]; Mediana: 0,46	126913574	104	5773856	>99,9	4,4
chr8	18	381	67775	Poli A (19), Poli C (7), Poli T (13), Poli G (7), Dinucleotídeo (5), Trinucleotídeo (9), Inserção (4), Deleção (0)	[0,26 - 0,78]; Mediana: 0,47	53430489	175	2958909	>99,9	5,2

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr9	23	347	87100	Poli A (12), Poli C (7), Poli T (27), Poli G (8), Dinucleotídeo (9), Trinucleotídeo (9), Inserção (0), Deleção (1)	[0,27 - 0,83]; Mediana: 0,49	69594586	74	3260257	>99,9	4,5
chr10	14	317	66723	Poli A (26), Poli C (7), Poli T (15), Poli G (6), Dinucleotídeo (16), Trinucleotídeo (6), Inserção (0), Deleção (0)	[0,23 - 0,78]; Mediana: 0,44	53209592	90	2469444	>99,9	4,4

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr11	29	511	91786	Poli A (28), Poli C (8), Poli T (21), Poli G (7), Dinucleotídeo (26), Trinucleotídeo (7), Inserção (2), Deleção (2)	[0,28 - 0,8]; Mediana: 0,47	72291795	150	3665560	>99,9	4,8
chr12	29	577	120365	Poli A (19), Poli C (8), Poli T (40), Poli G (7), Dinucleotídeo (7), Trinucleotídeo (7), Inserção (0), Deleção (3)	[0,26 - 0,77]; Mediana: 0,49	96109352	101	4331932	>99,9	4,3

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr13	13	283	58639	Poli A (24), Poli C (6), Poli T (12), Poli G (7), Dinucleotídeo (6), Trinucleotídeo (8), Inserção (14), Deleção (0)	[0,28 - 0,79]; Mediana: 0,42	46130028	44	2384839	>99,9	4,9
chr14	11	147	26980	Poli A (21), Poli C (6), Poli T (18), Poli G (11), Dinucleotídeo (6), Trinucleotídeo (6), Inserção (4), Deleção (0)	[0,29 - 0,77]; Mediana: 0,47	21336891	0	1078329	100	4,8
chr15	15	266	52091	Poli A (26), Poli C (7), Poli T (13), Poli G (6), Trinucleotídeo (8), Inserção (4), Deleção (0)	[0,29 - 0,76]; Mediana: 0,46	41918631	184	1753300	>99,9	4,0

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr16	21	366	80030	Poli A (7), Poli C (7), Poli T (15), Poli G (7), Dinucleotídeo (5), Trinucleotídeo (10), Inserção (15), Deleção (21)	[0,3 - 0,76]; Mediana: 0,54	62344351	18	4540539	>99,9	6,8
chr17	36	645	118062	Poli A (19), Poli C (7), Poli T (18), Poli G (8), Dinucleotídeo (13), Trinucleotídeo (6), Inserção (18), Deleção (1)	[0,28 - 0,82]; Mediana: 0,49	93811318	414	4403622	>99,9	4,5
chr18	9	99	19195	Poli A (7), Poli C (7), Poli T (15), Poli G (6), Trinucleotídeo (10), Inserção (0), Deleção (0)	[0,22 - 0,78]; Mediana: 0,44	15007653	6	990633	>99,9	6,2

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr19	30	605	104004	Poli A (19), Poli C (7), Poli T (31), Poli G (7), Dinucleotídeo (5), Trinucleotídeo (7), Inserção (2), Deleção (3)	[0,33 - 0,83]; Mediana: 0,59	81416722	455	4860311	>99,9	5,6
chr20	12	179	33795	Poli A (6), Poli C (6), Poli T (7), Poli G (8), Trinucleotídeo (9), Inserção (5), Deleção (0)	[0,31 - 0,84]; Mediana: 0,53	26833936	7	1301905	>99,9	4,6
chr21	5	63	30642	Poli A (28), Poli C (6), Poli T (24), Poli G (7), Dinucleotídeo (5), Inserção (1), Deleção (0)	[0,22 - 0,78]; Mediana: 0,52	24169250	44	1172087	>99,9	4,6

Cromossomo	Nº de genes	Nº de alvos	Nº de bases	Conteúdo genômico	Faixa de GC	Identificações corretas	Identificações incorretas	Não identificações	% de identificações corretas	% de não identificações
chr22	10	187	36727	Poli A (26), Poli C (7), Poli T (19), Poli G (7), Dinucleotídeo (5), Trinucleotídeo (6), Inserção (6), Deleção (0)	[0,27 - 0,74]; Mediana: 0,51	28887217	86	1392179	>99,9	4,6
chrX	23	433	83576	Poli A (18), Poli C (8), Poli T (23), Poli G (9), Dinucleotídeo (5), Trinucleotídeo (23), Inserção (3), Deleção (0)	[0,2 - 0,72]; Mediana: 0,48	64231080	241	3852253	>99,9	5,7
chrY	0	40	5476	Poli A (11), Poli C (8), Poli T (11), Poli G (5), Inserção (0), Deleção (0)	[0,4 - 0,59]; Mediana: 0,45	0	0	0	N/A	N/A

Os resultados do sequenciamento da amostra GM12878 foram comparados com um genótipo de alta confiança para NA12878, estabelecido pelo National Institutes of Standards and Technology (NIST) (v.2.19). Dos 9.232 alvos, 8.009 alvos estavam totalmente contidos nas regiões genômicas altamente confiáveis, 776 alvos tinham sobreposição parcial e 447 alvos não tinham sobreposição na

sequência do NIST. Isso resultou em 1.831.483 coordenadas por réplica para comparação. Identificações de base não variantes foram comparadas com sequências de referências do genoma humano versão hg19. Os resultados de precisão são mostrados em [Concordância somática da amostra GM12878 com o banco de dados do NIST na página 63](#).

Tabela 24 Concordância somática da amostra GM12878 com o banco de dados do NIST

Amostra	Nº de alvos cobertos	Capacidade de identificação de autossomo	TP	FN	TN	FP	PPA	NPA	OPA
GM12878	8785	94,5	247228	0	375073821	2043	100	>99,9	>99,9

Com base nos dados fornecidos por este estudo somático de 18 corridas, o NovaSeq 6000Dx Instrument pode sequenciar de forma consistente:

- Conteúdo de GC  $\geq$  20% (todas as bases identificadas em 1.692 regiões-alvo sequenciadas com 20% de conteúdo de GC identificado corretamente com taxa de não identificação de 0,34%).
- Conteúdo de GC  $\leq$  86% (todas as bases identificadas em 846 regiões-alvo sequenciadas com 86% de conteúdo de GC identificado corretamente com taxa de não identificação de 4,21%).
- Comprimentos de PoliA  $\leq$  46 (14550082 de 14550083 de bases identificadas em 846 regiões-alvo sequenciadas com 46 repetições de PoliA identificadas corretamente com taxa de não identificação de 4,18%).
- Comprimentos de PoliT  $\leq$  40 (12833489 de 12833491 de bases identificadas em 846 regiões-alvo sequenciadas com 40 repetições de PoliT identificadas corretamente com taxa de não identificação de 4,37%).
- Comprimentos de PoliG  $\leq$  11 (todas as bases identificadas em 846 regiões-alvo sequenciadas com 11 repetições de PoliG identificadas corretamente com taxa de não identificação de 7,59%).
- Comprimentos de PoliC  $\leq$  8 (9405604 de 9405615 de bases identificadas em 5.922 regiões-alvo sequenciadas com 8 repetições de PoliC identificadas corretamente com taxa de não identificação de 4,68%).
- Comprimentos de repetição de dinucleotídeos  $\leq$  31x (30996684 de 30996712 bases identificadas em 846 regiões-alvo sequenciadas com 31 repetições de dinucleotídeos identificadas corretamente com taxa de não identificação de 4,04%).
- Comprimentos de repetição de trinucleotídeos  $\leq$  23x (todas bases identificadas em 846 regiões-alvo sequenciadas com 23 repetições de trinucleotídeos identificadas corretamente com taxa de não identificação de 5,39%).
- Comprimentos de inserção  $\leq$  18 (todas as bases identificadas em 846 regiões-alvo sequenciadas com 18 inserções identificadas corretamente com taxa de não identificação de 1,44%).

- Comprimentos de deleção  $\leq 21$  (todas as bases identificadas em 846 regiões-alvo sequenciadas com 21 deleções identificadas corretamente com taxa de não identificação de 7,86%).

## Precisão

A precisão do NovaSeq 6000Dx Instrument foi avaliada usando amostras do Platinum Genome com um ensaio representativo projetado para consultar uma variedade de genes cobrindo 1.970.505 bases em 23 cromossomos diferentes usando 9.232 oligos alvo. Um total de 1.723 variantes pequenas direcionadas (SNVs, inserções e deleções) foi avaliado. Os testes de linha genética consistiram em 11 ou 12 réplicas de quatro amostras exclusivas do Platinum Genome. Os testes somáticos consistiram em 11 ou 12 réplicas de quatro amostras exclusivas do Platinum Genome tratadas com FFPE em diferentes níveis de VAF. As bibliotecas de amostras foram preparadas usando reagentes Illumina DNA Prep with Enrichment Dx Kit.

O teste foi realizado em um local interno usando três NovaSeq 6000Dx Instrument, três lotes cada de NovaSeq 6000Dx S2 Reagent v1.5 Kit (300 cycles) e NovaSeq 6000Dx S4 Reagent v1.5 Kit (300 cycles) e dois operadores durante seis dias de início. Para cada dia de início, as bibliotecas de amostras de linha genética foram sequenciadas em um lado do instrumento usando reagentes S2 e o fluxo de trabalho de análise de geração de FASTQ e VCF da linha genética do DRAGEN for Illumina DNA Prep with Enrichment Dx Application, e as bibliotecas de amostras somáticas foram sequenciadas no outro lado do instrumento usando reagentes S4 e o fluxo de trabalho de análise de geração de FASTQ e VCF da linha somática do DRAGEN for Illumina DNA Prep with Enrichment Dx Application. Esses testes resultaram em 18 lâminas de fluxo para cada um dos fluxos de trabalho somáticos e da linha genética.

## Linha genética

Para corridas de linha genética, os locais genômicos em que uma variante de linha genética direcionada é detectada são relatados como positivos (variante). Para variantes de linha genética positiva esperada, os dados foram avaliados quanto à taxa de não identificações e porcentagem de identificações positivas (PPC) dentro de cada tipo de variante (SNV, inserção, deleção). [Observações de identificação de linha genética de precisão intralaboratorial para resultados positivos esperados por tipo de variante na página 66](#) resumem as taxas observadas, juntamente com os níveis de confiança inferior e superior de 95% (LCL/UCL) calculados usando o método de pontuação de Wilson, para cada tipo de variante.

Tabela 25 Observações de identificação de linha genética de precisão intralaboratorial para resultados positivos esperados por tipo de variante

Tipo de variante	Não identificações observadas <sup>1</sup>	Total de identificações	Percentual de não identificações	Identificações positivas observadas <sup>2</sup>	Total de identificações avaliáveis	PPC	95% LCL <sup>3</sup>	95% UCL
SNV	6	980316	<0,01	979854	980310	99,95	99,95	99,96
Inserção	0	36738	0	36738	36738	100	>99,99	100
Deleção	18	34434	0,05	32160	34416	93,44	93,18	93,70

<sup>1</sup> Não identificação definida como posição cromossômica direcionada onde uma variante não pode ser determinada (devido à baixa profundidade de cobertura).

<sup>2</sup> Identificação positiva definida como posições cromossômicas direcionadas onde uma variante é detectada.

<sup>3</sup> Intervalos de confiança de 95% bilaterais calculados usando o método de pontuação de Wilson.

Locais genômicos onde uma variante direcionada não é detectada são relatados como negativos (tipo selvagem). Para locais negativos esperados, os dados foram avaliados para taxa de não identificações e porcentagem de identificações negativas (PNC). [Observações de identificação da linha genética de precisão intralaboratorial para resultados negativos esperados na página 66](#) resumem as taxas observadas, juntamente com os níveis de confiança inferior e superior de 95% (LCL/UCL) calculados usando o método de pontuação de Wilson.

Tabela 26 Observações de identificação da linha genética de precisão intralaboratorial para resultados negativos esperados

Tipo de variante	Não identificações observadas <sup>1</sup>	Total de identificações	Percentual de não identificações	Identificações negativas observadas <sup>2</sup>	Total de identificações avaliáveis	PNC	95% LCL <sup>3</sup>	95% UCL
Tipo selvagem	0	406170	0	406170	406170	100	>99,99	100

<sup>1</sup> Não identificação definida como posição cromossômica direcionada onde uma variante não pode ser determinada (devido à baixa profundidade de cobertura).

<sup>2</sup> Identificação negativa definida como posições cromossômicas direcionadas em que uma variante não é detectada.

<sup>3</sup> Intervalos de confiança de 95% bilaterais calculados usando o método de pontuação de Wilson.

A contribuição de cada parâmetro (instrumento, lote de reagentes, dia, réplica da biblioteca) para a variabilidade geral foi determinada pela análise do componente de variância usando a frequência da variante como a variável de resposta. O desvio padrão geral teve uma média de 0,0370. O maior contribuinte para a variabilidade da frequência da variante foi a partir de réplicas de preparação da biblioteca, que contribuíram para 17,1% da variabilidade geral. O dia contribuiu para 1%, enquanto o lote de instrumentos e reagentes contribuiu para menos de 1% da variabilidade total das [Estimativas de componentes de variância de precisão intralaboratorial para frequências de variantes de amostras de linha genética na página 67](#) (DP = desvio padrão).

Tabela 27 Estimativas de componentes de variância de precisão intralaboratorial para frequências de variantes de amostras de linha genética

Componente	DP médio	% média do DP total
Dia	0,0020	1,028
Instrumento	0,0018	0,837
Lote de material de consumo	0,0016	0,712
Réplica da biblioteca	0,0143	17,110
Total	0,0370	100

## Somática

Para corridas somáticas, os locais genômicos em que uma variante somática direcionada é detectada são relatados como positivos (variante). Para amostras diluídas GM12877-13 e GM12878-13 com variantes somáticas positivas esperadas em VAFs entre 6,5% e 13%, os dados foram avaliados quanto à ausência de taxa de identificação e porcentagem de identificação positiva (PPC) em cada tipo de variante (SNV, inserção, deleção). As [Observações de identificações somáticas de precisão intralaboratorial para resultados positivos esperados por tipo de variante \(VAF ≥ 6,5% e ≤ 13%\) na página 67](#) resumem as taxas observadas, juntamente com os níveis de confiança de 95% inferiores e superiores (LCL/UCL) calculados usando o método de pontuação de Wilson para cada tipo de variante.

Tabela 28 Observações de identificações somáticas de precisão intralaboratorial para resultados positivos esperados por tipo de variante (VAF ≥ 6,5% e ≤ 13%)

Tipo de variante	Não identificações observadas <sup>1</sup>	Total de identificações	Percentual de não identificações	Identificações positivas observadas <sup>2</sup>	Total de identificações avaliáveis	PPC	95% LCL <sup>3</sup>	95% UCL
SNV	0	96.939	0	96.069	96.939	99,10	99,04	99,16
Inserção	0	3.004	0	3.004	3.004	100	99,87	100
Deleção	0	2.912	0	2.907	2.912	99,83	99,60	99,93

<sup>1</sup> Não identificação definida como posição cromossômica direcionada onde uma variante não pode ser determinada (devido à baixa profundidade de cobertura).

<sup>2</sup> Identificação positiva definida como posições cromossômicas direcionadas onde uma variante é detectada.

<sup>3</sup> Intervalos de confiança de 95% bilaterais calculados usando o método de pontuação de Wilson.

Locais genômicos onde uma variante somática direcionada não é detectada são relatados como negativos (tipo selvagem). Para locais negativos esperados, os dados foram avaliados para taxa de ausência de identificações e porcentagem de identificações negativas. [Observações de identificações somáticas de precisão intralaboratorial para resultados negativos esperados na página 68](#) resumem as taxas observadas, juntamente com os níveis de confiança inferior e superior de 95% (LCL/UCL) calculados usando o método de pontuação de Wilson, para cada tipo de variante.

Tabela 29 Observações de identificações somáticas de precisão intralaboratorial para resultados negativos esperados

Tipo de variante	Não identificações observadas <sup>1</sup>	Total de identificações	Percentual de não identificações	Identificações negativas observadas <sup>2</sup>	Total de identificações avaliáveis	PNC	95% LCL <sup>3</sup>	95% UCL
Tipo selvagem	0	194.922	0	194.919	194.922	>99,99	>99,99	100

<sup>1</sup> Não identificação definida como posição cromossômica direcionada onde uma variante não pode ser determinada (devido à baixa profundidade de cobertura).

<sup>2</sup> Identificação negativa definida como posições cromossômicas direcionadas em que uma variante não é detectada.

<sup>3</sup> Intervalos de confiança de 95% bilaterais calculados usando o método de pontuação de Wilson.

A contribuição de cada parâmetro (instrumento, lote de reagentes, dia, réplica da biblioteca) para a variabilidade geral foi determinada pela análise do componente de variância usando a frequência da variante como a variável de resposta. O desvio padrão geral teve uma média de 0,0062. As réplicas de preparação de bibliotecas permaneceram a fonte mais significativa de variabilidade, representando 50,7% do total. Dia, instrumento e lote de material de consumo contribuíram para menos de 1% das [Estimativas de componentes de variância de precisão intralaboratorial para frequências de variantes de amostras somáticas na página 68](#) (DP = desvio padrão).

Tabela 30 Estimativas de componentes de variância de precisão intralaboratorial para frequências de variantes de amostras somáticas

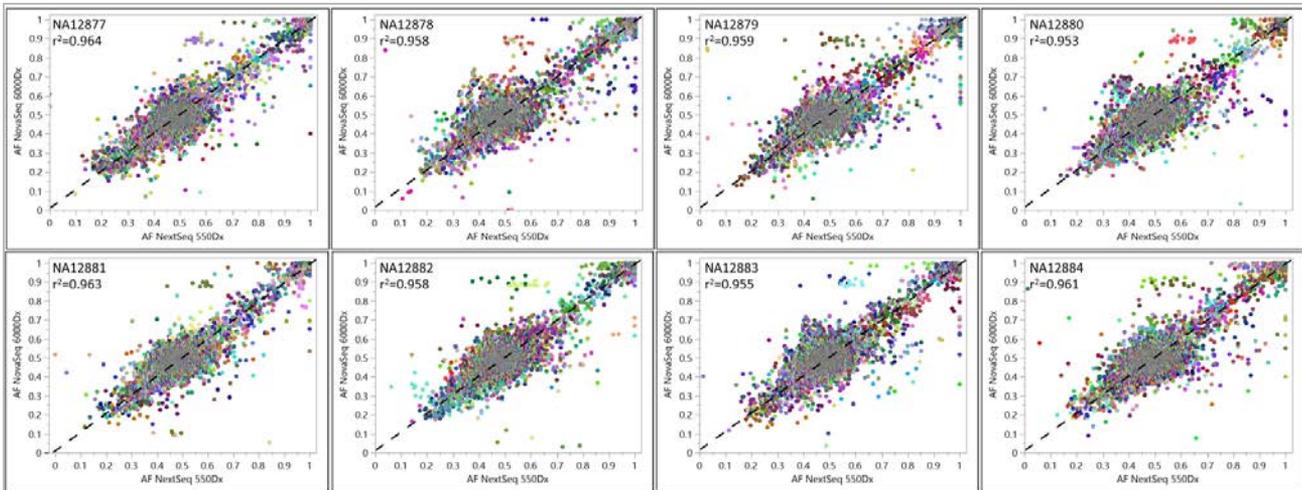
Componente	DP médio	% média do DP total
Dia	0,0002	0,41
Instrumento	0,0002	0,40
Lote de material de consumo	0,0002	0,35
Réplica da biblioteca	0,0044	50,7
Total	0,0062	100

## Comparação de métodos

Foi realizado um estudo para comparar o desempenho entre os NovaSeq 6000Dx e NextSeq 550Dx Instruments. A concordância da frequência de variantes para amostras de sangue foi avaliada usando um ensaio representativo projetado para consultar uma variedade de genes que cobrem 1.970.505 bases em todos os 23 cromossomos humanos. Oito amostras de DNA Platinum Genome foram testadas, sete em réplicas de seis e uma (NA12881) em réplicas de cinco. As bibliotecas foram sequenciadas no NovaSeq 6000Dx Instrument usando o fluxo de trabalho de análise de linha genética FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx Application e no NextSeq 550Dx Instrument usando o módulo DNA Generate FASTQ Dx Local Run Manager. [Gráficos de correlação de frequência de variantes \(os pontos são coloridos por variante](#)

*exclusiva. As variantes podem ser coloridas de forma diferente em cada gráfico individual.)* na página 69 traçam a correlação VAF entre os dois instrumentos para cada amostra. Com base na forte correlação entre o NovaSeq 6000Dx Instrument e o NextSeq 550Dx Instrument, as características de desempenho relacionadas aos fatores pré-analíticos (p. ex., métodos de extração ou substâncias interferentes) são determinadas como sendo aplicáveis em ambos os instrumentos. Consulte o folheto informativo do Illumina DNA Prep with Enrichment Dx para obter detalhes adicionais.

Figura 15 Gráficos de correlação de frequência de variantes (os pontos são coloridos por variante exclusiva. As variantes podem ser coloridas de forma diferente em cada gráfico individual.)



## Reprodutibilidade

A reprodutibilidade do NovaSeq 6000Dx Instrument foi avaliada usando amostras do Platinum Genome com um ensaio representativo projetado para consultar uma variedade de genes cobrindo 1.970.505 bases em 23 cromossomos diferentes usando 9.232 oligos alvo. Um total de 1.723 variantes pequenas direcionadas (SNVs, inserções e deleções) foi avaliado. Os testes de linha genética consistiram em três ou quatro réplicas de 12 amostras exclusivas de platina. Os testes somáticos consistiram em cinco ou seis réplicas de oito amostras exclusivas de Platinum Genome tratadas com FFPE em diferentes níveis de VAF. As bibliotecas de amostras foram preparadas usando reagentes Illumina DNA Prep with Enrichment Dx Kit.

Os testes foram realizados em três locais externos usando um lote cada de NovaSeq 6000Dx S2 Reagent v1.5 Kit (300 cycles) e NovaSeq 6000Dx S4 Reagent v1.5 Kit (300 cycles). Um único NovaSeq 6000Dx Instrument foi usado em cada local. Dois operadores realizaram os testes em cada local. Cada operador realizou testes em três dias de início não consecutivos para cada tipo de amostra, totalizando 36 lâminas de fluxo nos três locais. Para cada dia de início, as bibliotecas de amostras de linha genética foram sequenciadas no lado do instrumento A usando reagentes S2 e o fluxo de trabalho de análise de geração linha genética FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx Application, e as bibliotecas de amostras somáticas foram sequenciadas no lado do instrumento B usando reagentes S4 e o fluxo de trabalho de análise de geração Somatic FASTQ e VCF do DRAGEN for Illumina DNA Prep with Enrichment Dx Application. Esses testes resultaram em 18 lâminas de fluxo para cada um dos fluxos de trabalho somáticos e da linha genética.

## Linha genética

Para corridas de linha genética, os locais genômicos em que uma variante de linha genética direcionada é detectada são relatados como positivos (variante). Para variantes de linha genética positiva esperada, os dados foram avaliados quanto à taxa de não identificações e porcentagem de identificações positivas (PPC) dentro de cada tipo de variante (SNV, inserção, deleção). [Observações de identificação da linha genética para resultados positivos esperados por tipo de variante na página 70](#) resumem as taxas observadas, juntamente com os níveis de confiança inferior e superior de 95% (LCL/UCL) calculados usando o método de pontuação de Wilson, para cada tipo de variante.

Tabela 31 Observações de identificação da linha genética para resultados positivos esperados por tipo de variante

Tipo de variante	Não identificações observadas <sup>1</sup>	Total de identificações	Percentual de não identificações	Identificações positivas observadas <sup>2</sup>	Total de identificações avaliáveis	PPC	95% LCL <sup>3</sup>	95% UCL
SNV	0	991.026	0	990.276	991.026	99,92	99,92	99,93
Inserção	0	38.358	0	38.358	38.358	100	99,99	100
Deleção	0	34.758	0	32.228	34.758	92,72	92,44	92,99

<sup>1</sup> Não identificação definida como posição cromossômica direcionada onde uma variante não pode ser determinada (devido à baixa profundidade de cobertura).

<sup>2</sup> Identificação positiva definida como posições cromossômicas direcionadas onde uma variante é detectada.

<sup>3</sup> Intervalos de confiança de 95% bilaterais calculados usando o método de pontuação de Wilson.

Locais genômicos onde uma variante direcionada não é detectada são relatados como negativos (tipo selvagem). Para locais negativos esperados, os dados foram avaliados para taxa de não identificações e porcentagem de identificações negativas (PNC). [Observações de identificação da linha genética para resultados negativos esperados na página 70](#) resumem as taxas observadas, juntamente com os níveis de confiança inferior e superior de 95% (LCL/UCL) calculados usando o método de pontuação de Wilson.

Tabela 32 Observações de identificação da linha genética para resultados negativos esperados

Tipo de variante	Não identificações observadas <sup>1</sup>	Total de identificações	Percentual de não identificações	Identificações negativas observadas <sup>2</sup>	Total de identificações avaliáveis	PNC	95% LCL <sup>3</sup>	95% UCL
Tipo selvagem	0	393.516	0	393.516	393.516	100	>99,99	100

<sup>1</sup> Não identificação definida como posição cromossômica direcionada onde uma variante não pode ser determinada (devido à baixa profundidade de cobertura).

<sup>2</sup> Identificação negativa definida como posições cromossômicas direcionadas em que uma variante não é detectada.

<sup>3</sup> Intervalos de confiança de 95% bilaterais calculados usando o método de pontuação de Wilson.

## Somática

Para corridas somáticas, os locais genômicos em que uma variante somática direcionada é detectada são relatados como positivos (variante). Para variantes somáticas positivas esperadas, em que a frequência média de alelos variantes (VAF) é maior ou igual a 14% e menor ou igual a 28%, os dados foram avaliados quanto à ausência de taxa de identificações e porcentagem de identificações positivas (PPC) dentro de cada tipo de variante (SNV, inserção, deleção). [Observações da identificação somática para resultados positivos esperados por tipo de variante \(VAF médio  \$\geq 14\%\$  e  \$\leq 28\%\$ \) na página 71](#) resumem as taxas observadas, juntamente com os níveis de confiança inferior e superior de 95% (LCL/UCL) calculados usando o método de pontuação de Wilson, para cada tipo de variante.

Tabela 33 Observações da identificação somática para resultados positivos esperados por tipo de variante (VAF médio  $\geq 14\%$  e  $\leq 28\%$ )

Tipo de variante	Não identificações observadas <sup>1</sup>	Total de identificações	Percentual de não identificações	Identificações positivas observadas <sup>2</sup>	Total de identificações avaliáveis	PPC	95% LCL <sup>3</sup>	95% UCL
SNV	0	71.028	0	70.314	71.028	98,99	98,92	99,07
Inserção	0	1.962	0	1.962	1.962	100	99,80	100
Deleção	0	2.142	0	2.098	2.142	97,95	97,25	98,47

<sup>1</sup> Não identificação definida como posição cromossômica direcionada onde uma variante não pode ser determinada (devido à baixa profundidade de cobertura).

<sup>2</sup> Identificação positiva definida como posições cromossômicas direcionadas onde uma variante é detectada.

<sup>3</sup> Intervalos de confiança de 95% bilaterais calculados usando o método de pontuação de Wilson.

Locais genômicos onde uma variante somática direcionada não é detectada são relatados como negativos (tipo selvagem). Para locais negativos esperados, os dados foram avaliados para taxa de ausência de identificações e porcentagem de identificações negativas. [Observações da identificação somática para resultados negativos esperados na página 71](#) resume as taxas observadas, juntamente com os níveis de confiança inferior e superior de 95% (LCL/UCL) calculados usando o método de pontuação de Wilson, para cada tipo de variante.

Tabela 34 Observações da identificação somática para resultados negativos esperados

Tipo de variante	Não identificações observadas <sup>1</sup>	Total de identificações	Percentual de não identificações	Identificações negativas observadas <sup>2</sup>	Total de identificações avaliáveis	PNC	95% LCL <sup>3</sup>	95% UCL
Tipo selvagem	0	92.718	0	92.714	92.718	>99,99	99,99	100

<sup>1</sup> Não identificação definida como posição cromossômica direcionada onde uma variante não pode ser determinada (devido à baixa profundidade de cobertura).

<sup>2</sup> Identificação negativa definida como posições cromossômicas direcionadas em que uma variante não é detectada.

<sup>3</sup> Intervalos de confiança de 95% bilaterais calculados usando o método de pontuação de Wilson.

## Histórico de revisões

Documento	Data	Descrição da alteração
Documento n.º 200025276 v01	Setembro de 2022	Dados de precisão atualizados para observações de identificações de linha genética.
Documento n.º 200025276 v00	Agosto de 2022	Versão inicial.

## Patentes e marcas comerciais

Este documento e seu conteúdo são de propriedade da Illumina, Inc. e de suas afiliadas ("Illumina") e destinam-se exclusivamente ao uso contratual pelo cliente com relação ao uso dos produtos descritos neste documento e a nenhuma outra finalidade. O documento e seu conteúdo não devem ser usados nem distribuídos para qualquer outra finalidade, tampouco comunicados, divulgados ou reproduzidos de qualquer forma sem o consentimento prévio por escrito da Illumina. A Illumina não concede qualquer licença segundo seus direitos de patente, marca registrada, direitos autorais ou lei comum nem direitos semelhantes de terceiros por meio deste documento.

As instruções no documento devem ser estrita e explicitamente seguidas por pessoal devidamente treinado e qualificado para garantir o uso adequado e seguro dos produtos descritos neste documento. Todo o conteúdo do documento deve ser lido e compreendido por completo antes da utilização de tais produtos.

NÃO LER COMPLETAMENTE E NÃO SEGUIR EXPLICITAMENTE TODAS AS INSTRUÇÕES AQUI CONTIDAS PODE RESULTAR EM DANOS AO(S) PRODUTO(S), LESÕES PESSOAIS, INCLUSIVE USUÁRIOS OU OUTROS, E DANOS EM OUTROS BENS, ANULANDO QUALQUER GARANTIA APLICÁVEL AO(S) PRODUTO(S).

A ILLUMINA NÃO SE RESPONSABILIZA POR QUALQUER PROBLEMA CAUSADO PELO USO INDEVIDO DO(S) PRODUTO(S) MENCIONADO(S) ACIMA (INCLUINDO PARTES SEPARADAS OU O SOFTWARE).

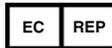
© 2022 Illumina, Inc. Todos os direitos reservados.

Todas as marcas comerciais pertencem à Illumina, Inc. ou aos respectivos proprietários. Para obter informações específicas sobre marcas comerciais, consulte [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).

## Informações de contato



Illumina  
5200 Illumina Way  
San Diego, California 92122, EUA  
+1 (800) 809.ILMN (4566)  
+1 (858) 202.4566 (fora da América do Norte)  
techsupport@illumina.com  
www.illumina.com



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

### Patrocinador australiano

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

## Identificação dos rótulos dos produtos

Para obter explicações completas sobre os símbolos que constam da embalagem e rótulos dos produtos, consulte a legenda de símbolos em [support.illumina.com](http://support.illumina.com) na guia *Documentation* (Documentação) do seu kit.