

illumina®

TruSight Whole Genome Analysis Application

Documentação do produto

PROPRIEDADE DA ILLUMINA
Documento n.º 200049931 v00
Abril de 2024

PARA UTILIZAÇÃO EM DIAGNÓSTICO IN VITRO.

Histórico de revisões

Documento	Data	Descrição da alteração
Documento n.º 200049931 v00	Abril de 2024	Edição inicial.

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

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

A NÃO OBSERVÂNCIA DA RECOMENDAÇÃO PARA LEITURA INTEGRAL E SEGUIMENTO EXPLÍCITO DE TODAS AS INSTRUÇÕES AQUI CONTIDAS PODE RESULTAR EM DANOS NO(S) PRODUTO(S), LESÕES EM PESSOAS, INCLUINDO NOS UTILIZADORES OU OUTROS, E EM DANOS MATERIAIS, E IRÁ ANULAR QUALQUER GARANTIA APLICÁVEL AO(S) PRODUTO(S).

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

© 2024 Illumina, Inc. Todos os direitos reservados.

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

Índice

Histórico de revisões	ii
Descrição geral	1
Introdução	1
Definições	2
Criação de ensaio	4
Revisão do ensaio	5
Recolocar a análise em fila de espera	6
Ver ensaio e resultados	7
Resumo do ficheiro de saída	9
Informações do relatório de CQ	9
Informações do ficheiro de identificação de variantes	13
Ficheiros FASTQ	22
Ficheiros CRAM	23
Assistência Técnica	25

Descrição geral

O TruSight Whole Genome Analysis Application é utilizado para planear ensaios de sequenciação TruSight Whole Genome e iniciar automaticamente a análise após a conclusão do ensaio. A análise inclui desmultiplexação, geração FASTQ, mapeamento de leitura, alinhamento com o genoma de referência humana GrCh38/hg38 ativado por gráfico e identificação de variantes utilizando o Illumina DRAGEN Server for NovaSeq 6000Dx.

Em diferentes fases do fluxo de trabalho da análise, a aplicação realiza o controlo de qualidade (CQ) de acordo com as métricas definidas de sequenciação, FASTQ e banco de amostras e gera relatórios com os resultados. Para amostras que passam em todos os passos de CQ, a aplicação gera ficheiros de saída de suporte para utilização em aplicações de linha germinativa a jusante.

O TruSight Whole Genome Analysis Application executa identificadores de variantes DRAGEN, incluindo o Identificador de Variantes Pequenas, o Identificador de Variantes de Números de Cópias (CNV) e a Detecção de Expansão Repetida com o ExpansionHunter.

A aplicação também efetua anotações de níveis de confiança baixos, intermédios ou elevados para variantes pequenas e inclui esta anotação no ficheiro de saída.

Introdução

Certifique-se de que o TruSight Whole Genome Analysis Application está instalado no NovaSeq 6000Dx Instrument que será utilizado para sequenciação como parte do TruSight Whole Genome. As aplicações instaladas podem ser encontradas no ecrã Applications (Aplicações) no NovaSeq 6000Dx Instrument ou no Illumina Run Manager utilizando um navegador num computador ligado à rede. Para obter assistência para agendar a instalação, contacte o seu Representante Illumina local.

Requisitos de armazenamento de dados

Consulte NovaSeq 6000Dx, Documentação do produto (documento n.º 200010105) e Documentação do produto DRAGEN Server for NovaSeq 6000Dx (documento n.º 200014171) para obter informações gerais sobre o resultado e armazenamento dos dados.

O TruSight Whole Genome Analysis Application envia os dados para uma Run Folder (Pasta de Ensaio) e uma Analysis Folder (Pasta de Análise) no armazenamento externo. Os requisitos mínimos de armazenamento podem ser aproximados a partir do tamanho dos dados produzidos em cada pasta para um único ensaio de sequenciação, como se mostra abaixo.

Configuração	Pasta de Ensaio (GB)	Pasta de Análise (GB)
Célula de fluxo S2 (6 amostras)	~430	~350
Célula de fluxo S4 (16 amostras)	~1110	~890

Tempo aproximado de análise

A análise começa automaticamente após a conclusão de um ensaio de sequenciação e ocorre sequencialmente em amostras dentro de um ensaio. Os ficheiros de saída de dados estarão disponíveis no armazenamento externo assim que a análise estiver concluída para todas as amostras num ensaio e a transferência de cópia para o armazenamento externo estiver concluída. Ao iniciar um ensaio de sequenciação no lado A e no lado B ao mesmo tempo, a sequenciação será realizada simultaneamente. A análise destes ensaios de sequenciação será realizada sequencialmente pelo TruSight Whole Genome Analysis Application após a conclusão da sequenciação. O ensaio que concluir primeiro a sequenciação e a transferência será analisado em primeiro lugar. O segundo ensaio de sequenciação será transferido e colocada em fila de espera para análise após a conclusão da primeira análise. Consulte [Ver ensaio e resultados na página 7](#) para saber como determinar o estado dos ensaios ativos ou falhados.

O tempo aproximado até que os resultados da análise estejam disponíveis após a conclusão da sequenciação é mostrado abaixo para a situação em que o lado A e o lado B são carregados simultaneamente com a mesma configuração.

Configuração	Ensaio de análise 1 (horas)	Ensaio de análise 2 (horas)
Célula de fluxo S2 (6 amostras)	~12	~24
Célula de fluxo S4 (16 amostras)	~24	~48

Definições

Selecione o TruSight Whole Genome Analysis Application no ecrã Applications (Aplicações) para ver a configuração atual e alterar permissões.

Configuração

O ecrã de configuração mostra as seguintes definições da aplicação:

- **Application name** (Nome da aplicação)
- **Application version** (Versão da aplicação)
- **DRAGEN Version** (Versão DRAGEN)
- **RTA Version** (Versão RTA)
- **Release Date** (Data de edição)
- **Organization** (Organização)
- **Device Identifier** (Identificador do dispositivo)
- **Production Identifier** (Identificador de produção)

- **Library Prep Kits** (Kits de preparação de bancos) — Mostra o kit de preparação do banco predefinido. Esta predefinição não pode ser alterada.
- **Index Adapter Kits** (Kits de adaptador de índice) — Exibe os conjuntos de adaptadores de índice disponíveis para uso.
- **Index Reads** (Leituras de índice)
- **Read Type** (Tipo de leitura)
- **Index Lengths** (Comprimentos do índice)
- **Read Lengths** (Comprimentos de leitura) — Os comprimentos de leitura são definidos por padrão quando o conjunto de índices é selecionado. Esta predefinição não pode ser alterada.

Autorizações

O administrador designado tem acesso a Permissões e pode utilizar as caixas de verificação no ecrã Permissions (Permissões) para gerir o acesso do utilizador para o TruSight Whole Genome Analysis Application.

Para obter mais informações sobre permissões e gestão de utilizadores, consulte a secção System Configuration (Configuração do Sistema) do NovaSeq 6000Dx, Documentação do produto (documento n.º 200010105).

Criação de ensaio

Crie novos ensaios em modo IVD no instrumento ou acedendo a Illumina Run Manager (IRM) utilizando um navegador num computador ligado à rede. Para aceder remotamente ao instrumento, utilize a informação de endereço e conta de utilizador facultada pelo seu representante Illumina. Consulte o NovaSeq 6000Dx, Documentação do produto (documento n.º 200010105) para obter mais informações.

Criar ensaio é o método recomendado para planeamento do ensaio. Não se recomenda a importação de folhas de amostras. Os ficheiros de folha de amostras produzidos nas pastas de ensaio e análise não são adequados para importação durante o planeamento do ensaio.

Criar ensaios

1. No ecrã Runs (Ensaaios), selecione **Create Runs** (Criar Ensaaios).
2. Selecione TruSight Whole Genome Analysis Application e, em seguida, selecione **Next** (Seguinte).
3. No ecrã Run Settings (Predefinições de Ensaaios), introduzir um nome do ensaio. O nome do ensaio é o nome que o identifica desde a sequenciação à análise.
4. [Opcional] Introduza uma descrição do ensaio para ajudar a identificá-lo. O kit de preparação da biblioteca está predefinido como TruSight Whole Genome e não pode ser alterado.
5. Selecione o conjunto de índice TruSight Whole Genome desejado no menu suspenso **Index Adapter Kit** (Kit de adaptador de índices).
A duração da leitura será definida por predefinição e não pode ser alterada. (Leituras 1 e 2 utilizam 151 ciclos; Índices 1 e 2 utilizam 10 ciclos).
6. Introduza uma Library Tube ID (ID do tubo da biblioteca) (formato recomendado como DX1234567-LIB) e, em seguida, selecione **Next** (Seguinte).
Se não for especificada nenhuma Library Tube ID (ID do tubo da biblioteca) neste passo, o ensaio planeado terá de ser selecionado antes do carregamento dos consumíveis de sequenciação. Se for introduzida a Library Tube ID (ID do tubo da biblioteca) incorreta neste passo, o ensaio planeado tem de ser corrigido antes de carregar os consumíveis. Consulte [Revisão do ensaio na página 5](#) para obter o protocolo para corrigir o ensaio quando estiver pronto para carregar consumíveis.
7. Nos ecrãs Sample Data (Dados da Amostra) e Sample Settings (Definições da Amostra), serão introduzidas informações da amostra. Os dados de amostra podem ser introduzidos manualmente ou através da importação de um ficheiro de dados da amostra. A ID da amostra tem de ser única para cada amostra e só pode conter caracteres alfanuméricos, sublinhados e traços. Não inclua espaços. A Well Position (Posição do Poço) refere-se ao poço no formato A01 a H04 da placa de índice. As informações da sequência de índice serão preenchidas automaticamente quando a Well Position (Posição do Poço) da placa de índice for introduzida. O sexo tem de ser introduzido como

Male (Masculino), Female (Feminino) ou Unknown (Desconhecido). A Library Plate ID (ID da placa da biblioteca) e a Library Well ID (ID do poço da biblioteca) (por exemplo, formato A01) são campos obrigatórios.

- Para introduzir manualmente os dados da amostra, adicione linhas (num total de 6 para S2 ou 16 para célula de fluxo S4) e introduza as informações necessárias nos Campos Sample ID (ID da Amostra) e Well Position (Posição do Poço). As informações também podem ser copiadas e coladas do Excel. Selecione **Next** (Seguinte). No ecrã Sample Settings (Definições da Amostra), introduza Library Plate ID (ID da placa da biblioteca), Library Well ID (ID do poço da biblioteca) e Sex (Sexo.) Selecione **Next** (Seguinte).
 - Para importar um ficheiro de dados da amostra, selecione **Import Samples** (Importar Amostras) e carregue o ficheiro de dados da amostra. As informações serão preenchidas automaticamente em linhas. Está disponível um modelo (*.csv) para transferência neste ecrã. Selecione **Next** (Seguinte). No ecrã Sample Settings (Definições da Amostra), as informações serão preenchidas automaticamente nas linhas a partir do ficheiro de dados da amostra importado. Selecione **Next** (Seguinte).
8. No ecrã Analysis Settings (Definições de Análise), introduza o Batch Name (Nome do lote) registado durante o planeamento do lote e do ensaio.
 9. [Opcional] Selecione o Tipo de Célula de Fluxo, S2 ou S4.
 10. Confirme ou desmarque a caixa de verificação para Gerar FASTQ comprimidos ORA e, em seguida, selecione **Next** (Seguinte).

NOTA O TruSight Whole Genome Analysis Application gera FASTQ comprimidos ORA por predefinição. Alterar esta definição irá aumentar o tamanho da saída de dados final.

11. No ecrã Run Review (Revisão do Ensaio) reveja as informações introduzidas. Se não forem necessárias alterações, selecione **Save** (Guardar). Se forem necessárias alterações, selecione **Back** (Voltar) conforme necessário para regressar ao ecrã adequado.



ATENÇÃO

TruSight Whole Genome foi validado para 6 amostras quando se utiliza a célula de fluxo S2 NovaSeq 6000Dx e 16 amostras quando se utiliza a célula de fluxo S4 NovaSeq 6000Dx. Certifique-se de que é introduzido o número correto de amostras para a configuração da célula de fluxo selecionada.

Revisão do ensaio

Se forem necessárias alterações após a criação do ensaio e antes de carregar consumíveis para sequenciação, reveja os ensaios no modo IVD no instrumento ou acedendo a Illumina Run Manager (IRM) utilizando um browser num computador ligado à rede.

1. Selecione **Runs** (Ensaio).

2. Selecione o Run name (Nome do ensaio) no separador Planned Runs (Ensaio Planeados).
3. Selecione **Edit** (Editar).
4. Atualize as informações do ensaio ou da amostra conforme necessário. Por exemplo, introduza ou corrija o Library Tube ID (ID do tubo da biblioteca) para corresponder ao utilizado ao concluir o fluxo de trabalho.
5. Selecione **Next** (Seguinte) para Revisão do Ensaio.
6. Selecione **Save** (Guardar).
7. Selecione **Exit** (Sair).

Regresse a Sequenciação no modo IVD para repetir o carregamento dos consumíveis. O ensaio deve agora ser automaticamente realçado.

Se atualizar a Library Tube ID (ID do tubo da biblioteca) durante o carregamento de consumíveis, volte à Run Selection (Seleção do Ensaio) no software de controlo e selecione **Refresh** (Atualizar) para a coluna associada, A ou B. O ensaio deve agora ser automaticamente realçado. Caso contrário, selecione **Back** (Voltar) para repetir Carregar consumíveis.

Recolocar a análise em fila de espera

Consulte a secção Resolução de problemas no Folheto informativo TruSight Whole Genome (documento n.º 200050132) para determinar o tipo de Recolocar a análise em fila de espera mais adequado.

Recoloque as análises em fila de espera, sem qualquer alteração

1. Selecione o nome do Completed Run (Ensaio Concluído) para ver os Run Details (Detalhes do Ensaio).
2. Selecione **Requeue Analysis** (Recolocar análise em fila de espera).
3. Selecione **Requeue analysis with no changes** (Recolocar as análises em fila de espera, sem qualquer alteração).
4. Forneça detalhes no campo Reanalysis Reason (Motivo da Reanálise).
5. Selecione **Requeue Analysis** (Recolocar análise em fila de espera).
6. Saia da página e navegue até à página Active Runs (Ensaio Ativos) para confirmar que a recolocação em fila de espera está em curso.

Recolocar as análises em fila de espera com alterações

1. Selecione o nome do Completed Run (Ensaio Concluído) para ver os Run Details (Detalhes do Ensaio).
2. Selecione **Requeue Analysis** (Recolocar análise em fila de espera).

3. Selecione **Edit run settings** e **requeue analysis** (Editar definições do ensaio e de recolocação de análise em fila de espera)
4. Forneça detalhes no campo Reanalysis Reason (Motivo da Reanálise).
5. Selecione **Requeue Analysis** (Recolocar análise em fila de espera).
6. Confirme ou atualize as Run Settings (Definições de Ensaio) e, em seguida, selecione **Next** (Seguinte).
7. Corrija as informações da amostra conforme necessário atualizando manualmente os campos ou selecione **Download Template** (Transferir Modelo) para criar um ficheiro `sampledata.csv` com as informações atuais. Corrija as informações e elimine as linhas existentes no separador Sample Data (Dados da Amostra) antes de utilizar Import Samples (Importar Amostras) para preencher os dados de amostra corrigidos.
8. Reveja as informações no ecrã Run Review (Revisão do Ensaio) e selecione **Save** (Guardar) para iniciar a reanálise.
9. Selecione **Exit** (Sair) e navegue até à página Active Runs (Ensaio Ativos) para confirmar que a recolocação em fila de espera está em curso.

A pasta de dados do ensaio original tem de estar presente no local de armazenamento externo especificado em Run Details (Detalhes do Ensaio) para que a reanálise seja concluída com sucesso. Se a reanálise falhar, certifique-se de que o ensaio não foi movido ou eliminado.

Ver ensaio e resultados

1. No ecrã Illumina Run Manager principal no modo IVD, selecione **Runs** (Ensaio).
2. No separador Completed Runs (Ensaio Concluídos), selecione o nome do ensaio. Este separador também apresenta ensaios concluídos devido a falha de sequenciação, transferência de dados ou análise. Os ensaios ativos e o seu estado são apresentados no separador Active Runs (Ensaio Ativos). Consulte o NovaSeq 6000Dx, Documentação do produto (documento n.º 200010105) para obter mais informações.
3. Selecione nome do ensaio no separador Completed Runs (Ensaio Concluídos) para visualizar os Detalhes do Ensaio e os Resultados para o caminho para a Pasta de Resultados da Análise. Para ensaios falhados, reveja o Estado para cada passo e, em seguida, consulte a secção Resolução de problemas do Folheto informativo TruSight Whole Genome (documento n.º 200050132).
4. Navegue até à pasta de análise na sua unidade local e abra o Relatório Consolidado para rever o resultado PASS/FAIL (PASSOU/FALHO) para cada passo de CQ da seguinte forma:
 - Para o CQ do ensaio de sequenciação, consulte Resumo dos resultados do CQ da sequenciação
 - Para o CQ FASTQ para cada amostra no ensaio, consulte o Resumo dos resultados do CQ FASTQ
 - Para o CQ do banco para cada amostra no ensaio, consulte o Resumo dos resultados do CQ do banco de amostras

Se for observado um resultado de FAIL (FALHOU), anote o passo de CQ e consulte a secção Resolução de problemas do Folheto informativo TruSight Whole Genome (documento n.º 200050132).

Resumo do ficheiro de saída

O TruSight Whole Genome Analysis Application guarda os seguintes ficheiros de saída principais. Consulte as secções de informação de ficheiros abaixo para a localização dos ficheiros de saída principais.

Os ensaios e amostras que não cumprem os critérios de validade não produzem ficheiros CRAM, ROH bed ou *genome.vcf.

Ficheiro de saída	Descrição
Relatório consolidado (*.csv)	Contém métricas de qualidade utilizadas para a validade do ensaio (incluindo rendimento total e Q30), métricas de validade da amostra (incluindo rendimento FASTQ), métricas de CQ do banco e métricas Apenas para informação (FIO) para todas as amostras no ensaio.
Relatório de amostras (*.csv)	Contém resultados de CQ de sequenciação, FASTQ e CQ do banco de amostras. O relatório também contém métricas de concordância de ploidia e FIO para a amostra individual, bem como o ensaio de sequenciação associado.
Pequena variante e mSNV VCF (*.annotated.hard-filtered-gvcf.gz)	Contém informações de identificação de variantes para pequenas variantes (SNV, indels) e SNV mitocondriais.
CNV VCF (*.cnv.vcf.gz)	Contém informações de identificação de variantes para ganhos e perdas de números de cópias.
Repete VCF (*.repeats.vcf.gz)	Contém informações de identificação de variantes em expansões STR e SMN1.
ROH BED (*.roh.bed)	Contém informações para regiões de homozigotia.
FASTQ (*.fastq.gz ou *.fastq.ora)	Ficheiros intermédios que contêm identificações de base com pontuação de qualidade. Os ficheiros FASTQ são a principal entrada para o passo de alinhamento. Se for selecionada a compressão ORA, isto será refletido no nome do ficheiro.
CRAM de alinhamento (*.cram)	Contém leituras alinhadas de uma determinada amostra.

Informações do relatório de CQ

O Relatório Consolidado <<RunID>>_Consolidated_Report.csv está localizado no diretório TruSightWholeGenomeAnalysis_x.x.x_run-complete e contém informações sobre métricas de qualidade utilizadas para aprovar ou falhar amostras em diferentes fases da análise. Relatórios de

amostras individuais <<Sample_ID>>_Sample_Report.csv podem ser encontrados nas pastas <Sample_ID> no diretório TruSightWholeGenomeAnalysis_x.x.x_run-complete.

Os cabeçalhos do relatório incluem as seguintes informações sobre o ensaio: a versão da aplicação, o nome do lote, a ID do tubo do conjunto de banco, o nome do ensaio de sequenciação, a ID do ensaio de sequenciação e o tipo de célula de fluxo. As tabelas seguintes descrevem as informações incluídas no Relatório Consolidado. O Relatório de Amostras individual inclui as mesmas informações, exceto as Métricas Demultiplex.

Tabela 1 Métricas de CQ de sequenciação

Métricas	Especificações	Descrição
Non-Indexed Total Yield (GB) (Rendimento total não indexado ([GB]))	ND	Nenhuma especificação, uma vez que ensaios de menor rendimento podem resultar em bancos de amostras aprovados. Previsão ≥ 3000 GBp para S4 e ≥ 1000 GB para célula de fluxo S2.
% total \geq Q30	≥ 85	Medição da qualidade da base ao nível do ensaio. A especificação mínima é definida uma vez que ensaios com %Q30 demasiado baixos não passarão as bases Q30 no Controlo de Qualidade do Banco de Amostras.
Resultado de CQ de sequenciação de resumo	PASS (PASSOU) ou FAIL (FALHOU)	Para falha do CQ de sequenciação, consulte a secção Resolução de problemas no Folheto informativo TruSight Whole Genome (documento n.º 200050132).

Tabela 2 Métricas de demultiplexagem

Métricas	Especificações	Descrição
Leituras percentuais identificadas	ND	Fração total de leituras de filtro de passagem no ensaio que foram atribuídas a amostras durante a desmultiplexagem.
Percentagem CV	ND	Fornece uma medida da uniformidade das leituras desmultiplexadas para cada par de índices no ensaio. Previsão $< 25\%$ para ensaios sem falhas de resultados de CQ FASTQ.

Tabela 3 Métricas de CQ FASTQ

Métricas	Especificações	Descrição
Rendimento por amostra (bps)	$\geq 90.000.000.000$	O mínimo é definido para ser equivalente a uma cobertura autossômica média de ~26x para fazer a triagem dos bancos de amostras que não passarão o CQ para reduzir o tempo de análise.
Resumo do resultado de CQ FASTQ	PASS (PASSOU) ou FAIL (FALHOU)	Para falha do CQ FASTQ, consulte a secção Resolução de problemas no Folheto informativo TruSight Whole Genome (documento n.º 200050132).

Tabela 4 Métricas de CQ do banco de amostras

Métricas	Especificações	Descrição
Cobertura autossômica média	≥ 35	Cobertura média nos autossomas. A especificação mínima é definida para garantir o desempenho analítico.
Percentagem de autossomas com cobertura superior a 20X	$\geq 93,94$	Medida da uniformidade da cobertura que deteta problemas não necessariamente relacionados com a tendência GC. A especificação mínima é definida para garantir o desempenho analítico.
Cobertura normalizada em recipientes GC de 60% a 79%	$0,82 \leq x \leq 1,13$	Medição da uniformidade da cobertura que deteta a tendência GC, especificamente uma perda de cobertura em áreas do genoma com maior % GC e menor % de composição base AT. São definidas especificações mínimas e máximas para garantir o desempenho analítico.
Cobertura normalizada em recipientes GC de 20% a 39%	$0,97 \leq x \leq 1,06$	Medida da uniformidade de cobertura que deteta a tendência GC, especificamente uma perda de cobertura em áreas do genoma com menor % GC e maior % de composição de base AT. São definidas especificações mínimas e máximas para garantir o desempenho analítico.
Cobertura mitocondrial média	≥ 500	Cobertura do cromossomo mitocondrial. A especificação mínima é definida para garantir o limite de deteção de SNV mitocondrial.
Percentagem de bases Q30	≥ 85	Medição da qualidade da base. A especificação mínima é definida para garantir o desempenho analítico.
Contaminação estimada da amostra	$\leq 0,005$	Deteta leituras contaminantes de outras amostras. A especificação máxima é definida para garantir o limite de deteção de SNV mitocondrial (o tipo de variante com a maior sensibilidade à contaminação).

Métricas	Especificações	Descrição
Resumo do resultado de CQ do banco de amostras	PASS (PASSOU) ou FAIL (FALHOU)	Para uma falha do CQ do banco de amostras, consulte a secção Resolução de problemas no Folheto informativo TruSight Whole Genome (documento n.º 200050132).

Tabela 5 Métricas de CQ de Ploidia

Métricas	Especificações	Descrição
Forneceu ploidia de cromossomas sexuais	ND	Sexo fornecido pelo operador durante a Criação do Ensaio (Feminino, Masculino, Desconhecido).
Estimativa da ploidia	ND	Ploidia sexual estimada por DRAGEN.
Resumo do resultado da ploidia	CONCORDANTE, DISCORDANTE ou ND	CONCORDANTE indica concordância entre a ploidia sexual fornecida e estimada. ND indica o sexo fornecido como Desconhecido ou estimativa diferente de XX ou XY. Para um resultado DISCORDANTE, ou foi introduzido o sexo errado durante a criação do ensaio ou pode ter ocorrido uma troca de amostras. Consulte a secção Resolução de problemas no Folheto informativo TruSight Whole Genome (documento n.º 200050132).

Tabela 6 Métricas Apenas para Informação

Métricas	Descrição
Inserir mediana de comprimento	O alvo é 450 bp, mas espera-se uma variação consoante o ensaio da sequenciação e o operador. É aceitável um intervalo de aproximadamente 360 a 550 bp. Uma operação consistente fora deste intervalo pode levar a uma maior incidência de falhas na amostra.
Percentagem de leituras mapeadas	Percentagem de leituras que mapeiam para o genoma de referência. Pode ser reduzido em resposta a contaminação de gADN não humano, fraca qualidade da amostra ou um comprimento de inserção demasiado pequeno, levando a problemas de mapeamento.
Percentagem de leituras com alinhamentos suplementares	Percentagem de leituras com mapeamento que se divide em diferentes locais no genoma de referência.

Métricas	Descrição
Percentagem de leituras marcadas duplicadas	Previsão < 20%. Pode ser elevado se o rendimento de preparação do banco for baixo ou se agrupar menos do que o volume necessário, ou em resposta a problemas relacionados com a sequenciação.
Percentagem de bases com corte suave Leitura 1	Útil no diagnóstico da causa raiz da falha da cobertura autossômica média.
Percentagem de bases com corte suave Leitura 2	Útil no diagnóstico da causa raiz da falha da cobertura autossômica média.
Percentagem de bases cortadas Leitura 1	Útil no diagnóstico da causa raiz da falha da cobertura autossômica média.
Percentagem de bases cortadas Leitura 2	Útil no diagnóstico da causa raiz da falha da cobertura autossômica média.

Informações do ficheiro de identificação de variantes

Ficheiros VCF

Os ficheiros em formatos de identificação de variantes (*.vcf) contêm informações sobre variantes encontrados em posições específicas num genoma de referência e podem ser encontrados no diretório <Sample_ID>/Analysis.

O cabeçalho do ficheiro VCF inclui a versão do formato do ficheiro, a versão identificadora da variante e lista as anotações no restante do ficheiro. A última linha no cabeçalho contém os títulos das colunas, para as linhas de dados. Cada linha do ficheiro VCF contém informação sobre uma posição de referência única.

Todos os ficheiros VCF contêm um cabeçalho com descrições das colunas de saída e dados de identificação de variantes em colunas rotuladas como CHROM, POS, ID, REF, ALT, QUAL, FILTER (FILTRO), INFO, FORMAT (FORMATO), SAMPLE (AMOSTRA). As definições dos valores de coluna podem variar consoante os identificadores de variantes.

Pequena variante e mSNV VCF

A saída é guardada no ficheiro <Sample_ID>.annotated.hard-filtered.gvcf.gz no diretório <Sample_ID>/Analysis.

Um ficheiro genómico VCF (gVCF) contém informações sobre variantes e posições determinadas como homocigóticas para o genoma de referência. Para regiões homocigóticas, o ficheiro gVCF inclui estatísticas que indicam quão bem as leituras suportam a ausência de variantes ou alelos alternativos. O ficheiro gVCF inclui um alelo <NON_REF> artificial. As leituras que não suportam a referência ou quaisquer variantes são atribuídas ao alelo <NON_REF>. O DRAGEN utiliza estas leituras para determinar se a posição pode ser identificada como uma referência homocigótica, em vez de permanecer não identificada. A pontuação resultante representa o nível de confiança com escala Phred numa identificação de referência homocigótica. No modo de linha germinativa, a pontuação é FORMATO/GQ.

DRAGEN fornece filtragem de variantes pós-VCF com base nas anotações presentes nos registos VCF. A filtragem rígida de variantes é descrita abaixo. No entanto, devido à natureza dos algoritmos de DRAGEN, que incorporam a hipótese de erros correlacionados a partir do núcleo do identificador de variantes, o pipeline tem capacidades melhoradas para distinguir as variantes verdadeiras do ruído e, por conseguinte, a dependência da filtragem pós-VCF é substancialmente reduzida.

O TruSight Whole Genome Analysis Application fornece anotação da pontuação de confiança e do nível de confiança para pequenas variantes que podem ser utilizadas para melhorar ainda mais o desempenho. A anotação do nível de confiança não é um filtro de qualidade e, como tal, não se reflete diretamente no estado de qualidade das identificações de variantes. Por conseguinte, é possível ver a aprovação de identificações de variantes que, no entanto, são anotadas como de baixa confiança.

Tabela 7 Cabeçalhos do ficheiro VCF

Cabeçalho	Descrição
CHROM	O cromossoma do genoma de referência. Os cromossomas aparecem pela mesma ordem que os do ficheiro de referência FASTA.
POS	A posição de base única da variante no cromossoma de referência. Para variantes de nucleótido único (SNV), esta posição é a base de referência com a variante. Para os indels, esta posição é a base de referência imediatamente precedente da variante.
ID	Sempre .
REF	O genótipo de referência. Por exemplo, uma eliminação de um único T é representada como referência TT e T alternativo. Um A para variante de nucleótido de T único é representado como referência A e T alternativo.
ALT	Os alelos que diferem da leitura de referência. Por exemplo, uma inserção de um único T é representada como referência A e AT alternativa. Um A para variante de nucleótido de T único é representado como referência A e T alternativa.

Cabeçalho	Descrição
QUAL	Uma pontuação de qualidade com escala Phred atribuída pelo identificador de variante. Pontuações mais altas indicam um grau de confiança maior na variante e menor probabilidade de erros. Para uma pontuação de qualidade de Q, a probabilidade estimada de ocorrerem erros é de $10^{-(Q/10)}$. Por exemplo, o conjunto de identificações Q30 tem uma taxa de erro de 0,1%. Muitos identificadores de variantes atribuem pontuações de qualidade com base nos respetivos modelos estatísticos, que são elevados em relação à taxa de erros observada.

Tabela 8 Anotações de ficheiros VCF

Cabeçalho	Descrição
FILTER (FILTRO)	<ul style="list-style-type: none"> • PASS (PASSOU): Todos os filtros foram aprovados. • DRAGENSnpHardQUAL — Definir se verdadeiro:QUAL < 10,41 • DRAGENIndelHardQUAL — Definir se verdadeiro:QUAL < 7,83 • LowDepth — Definir se verdadeiro:DP <= 1 • LowGQ — Defina se verdadeiro:GQ = 0 • PloidyConflict — Identificação do genótipo pelo identificador de variantes não consistente com a ploidia de cromossomas • base_quality — Filtrado no local porque a qualidade mediana base das leituras ALT neste locus não atinge o limiar • filtered_reads — Filtrado no local porque uma fração de leituras demasiado grande foi excluída pela filtragem • fragment_length — Filtrado no local porque a diferença absoluta entre o comprimento mediano do fragmento das leituras ALT e REF, neste locus, excede o limiar • low_af — A frequência de alelos não cumpre o limiar • low_depth — Filtrado no local porque a profundidade da leitura é demasiado baixa • low_frac_info_reads — Filtrado no local porque a fração de leituras informativas se encontra abaixo do limiar • low_normal_depth — Filtrado no local porque a profundidade da leitura normal da amostra é demasiado baixa • long_indel — Filtrado no local porque o comprimento do Indel é demasiado longo • mapping_quality — Filtrado no local porque a qualidade mediana do mapeamento das leituras ALT neste locus não atinge o limiar • multiallelic — Filtrado no local porque mais de dois alelos ALT passam o LoD do tumor • non_homref_normal — Filtrado no local porque o genótipo normal da amostra não é referência homocigótica • no_reliable_supporting_read — Filtrado no local porque não existe uma leitura somática de apoio fiável • panel_of_normals — Visto, pelo menos, numa amostra no painel dos VCF normais • read_position — Filtrado no local porque a mediana de distâncias entre o início/fim da leitura e este locus encontra-se abaixo do limiar • RMxNRepeatRegion — Filtrado no local porque todos ou parte dos alelos da variante são uma repetição da referência • strand_artifact — Filtrado no local pela tendência grave da cadeia • str_contraction — Filtrado no local devido a suspeita de erro de PCR, onde o alelo ALT é uma unidade de repetição inferior à referência • too_few_supporting_reads — Filtrado no local porque existem poucas leituras de apoio na amostra do tumor • weak_evidence — Pontuação somática da variante não atinge o limite

Cabeçalho	Descrição
INFO	<ul style="list-style-type: none"> • DB — Adesão dbSNP. • FS — Valor de p com escala Phred, utilizando o teste exato de Fisher para detetar tendências da cadeia. • QD — Fiabilidade da variante/qualidade por profundidade. • R2_5P_bias — Pontuação com base em tendência mate e distância da extremidade principal 5. • SOR — Taxa de probabilidade simétrica de tabela de contingência 2x2, para detetar tendências da cadeia. • DP — Profundidade aproximada de leitura (informativa e não-informativa); algumas leituras podem ter sido filtradas com base no MAPQ, etc. • END — Posição stop do intervalo. • FractionInformativeReads — A fração de leituras informativas do total de leituras. • MQ — Qualidade RMS do mapeamento. • MQRankSum — Pontuação Z do teste de soma de ranks de Wilcoxon de ALT vs. qualidades de mapeamento de leitura REF. • ReadPosRankSum — Pontuação Z do teste de soma de ranks de Wilcoxon de ALT vs. tendências de posição de leitura REF. • SOMATIC — Pelo menos uma variante nesta posição é somática. • ILENS — Comprimentos indel para cada variante ALT. • SCORE — Pontuação de confiança para cada tipo de variante presente no centro como (tipo de variante):(pontuação de confiança). • TIER — Nível de confiança para cada tipo de variante presente no centro como (tipo de variante):(Nível de confiança).

Cabeçalho	Descrição
FORMAT (FORMATO)	<p>O FORMATO da coluna indica os campos separados por dois pontos, por exemplo GT:GQ.</p> <ul style="list-style-type: none"> • AD — Profundidades dos alelos (contando apenas leituras informativas do total de leituras) para os alelos REF e ALT, na ordem listada. • AF — Frações de alelos para os alelos ALT, na ordem listada. • DP — Profundidade aproximada de leitura (leituras com MQ=255 ou com maus mates são filtradas). • F1R2 — Contagem de leituras em orientação de par F1R2 de apoio a cada alelo. • F2R1 — Contagem de leituras em orientação de par F2R1 de apoio a cada alelo. • GP — Probabilidades com escala Phred posteriores para genótipos, como definido na especificação do VCF. • GQ — Qualidade de genótipo. • GT—Genótipo. • ICNT — Contagens de leituras informativas INDEL com base no modelo de confiança de referência. • MB — Estatísticas de componente por amostra, para detetar a tendência mate. • MIN_DP — DP Mínima observada dentro do bloco GVCF. • PL — Probabilidades normalizadas com escala Phred para genótipos, como definido na especificação do VCF. • PRI — Probabilidades anteriores com escala Phred para genótipos. • PS — ID de informação física faseada, onde cada ID único, numa determinada amostra (mas não através das amostras), se liga aos registos dentro de um grupo faseado. • SB — Estatísticas de componente por amostra que abrange o Teste Exato de Fisher, para detetar as tendências da cadeia. • SPL —Probabilidades normalizadas com escala Phred para SNP com base no modelo de confiança de referência. • SQ — Qualidade somática.
SAMPLE (AMOSTRA)	A coluna amostra dá-lhe os valores especificados na coluna FORMATO.

Variantes no número de cópias VCF

A fase de contagens alvo é a primeira fase de processamento para o pipeline CNV DRAGEN, produzindo `<Sample_ID>.target.counts.gz`, depois é realizada a Correção de Tendência GC, gerando um ficheiro `*.target.counts.gc-corrected.gz`. A fase de normalização produz um ficheiro `*.tn.tsv.gz`. O Software Anfitrião DRAGEN gera muitos ficheiros intermédios.

*.seg.called.merged é o ficheiro de identificação final que contém os eventos de amplificação e eliminação.

Além do arquivo de segmento, DRAGEN emite as identificações no formato VCF padrão. O resultado é guardado em <Sample_ID>.cnv.vcf.gz no diretório <Sample_ID>/Analysis.

Definições de colunas específicas do identificador CNV:

A coluna POS é a posição inicial da variante. De acordo com a especificação VCF, se qualquer um dos alelos ALT for um alelo simbólico, tal como , então a base de preenchimento é necessária e POS indica a coordenada da base que precede o polimorfismo. Todas as coordenadas no VCF são baseadas em 1.

A coluna ID é utilizada para representar o evento. O campo ID codifica o tipo de evento e coordenadas do evento.

A coluna REF contém um N para todos os eventos CNV.

A coluna ALT especifica o tipo de evento CNV. Como o VCF contém apenas eventos CNV, apenas a entrada DEL ou DUP é utilizada.

A coluna QUAL contém uma pontuação de qualidade estimada para o evento CNV, que é utilizada na filtragem rígida.

A coluna FILTER (FILTRO) contém PASS (PASSOU) se o evento CNV passar em todos os filtros, caso contrário, a coluna contém o nome do filtro com falha.

A coluna INFO contém informações que representam o evento. A entrada REFLEN indica a duração do evento. A entrada SVTYPE é sempre CNV. A entrada FIM indica a posição final do evento.

Os campos FORMATO são descritos no cabeçalho.

- GT — Genótipo
- SM — Rácio de cópia linear da média do segmento
- CN — Número de cópia estimado
- BC — Número de recipientes na região
- PE — Número de leituras de extremidades emparelhadas incorretamente nos pontos de interrupção de início e paragem

Repete VCF

O ExpansionHunter realiza um realinhamento baseado em gráficos de sequência de leituras que têm origem no interior e à volta de cada repetição do alvo. O ExpansionHunter faz então o genótipo do comprimento da repetição em cada alelo com base nestes alinhamentos gráficos.

Estão disponíveis mais informações e análises nos seguintes documentos do ExpansionHunter:

- Dolzhenko et al., *Deteção de expansões de repetição prolongada a partir de dados de sequência de genoma completo sem PCR* de 2017

- Dolzhenko et al., *ExpansionHunter: Uma ferramenta baseada em gráfico de sequência para analisar a variação em pequenas regiões de repetição em tandem* 2019

O catálogo de variantes STR de TruSight Whole Genome Analysis Application contém especificações sobre repetições causadoras de doença localizadas nos genes AFF2, AR, ATN1, ATXN1, ATXN10, ATXN2, ATXN3, ATXN7, ATXN8OS, C9ORF72, CACNA1A, CBL, CNBP, CSTB, DIP2B, DMPK, FMR1, FXN, GLS, HTT, JPH3, NIPA1, NOP56, NOTCH2NL, PABPN1, PHOX2B, PPP2R2B e TBP.

Os resultados da genotipagem de repetição são apresentados como um ficheiro VCF separado, que fornece a duração de cada alelo em cada repetição identificável definida no ficheiro de catálogo de especificações de repetição. O nome do ficheiro é <Sample_ID>.repeats.gz e pode ser encontrado no diretório <Sample_ID>/Analysis.

Algumas colunas são específicas para repetir o identificador de expansão:

Tabela 9 Campos VCF principais

Campo	Descrição
CHROM	Identificador do cromossoma
POS	Posição da primeira base antes da região de repetição na referência
ID	Sempre .
REF	A base de referência na posição POS
ALT	Lista de alelos repetidos no formato <STR _n >. N é o número de unidades de repetição.
QUAL	Sempre .
FILTER (FILTRO)	O filtro LowDepth é aplicado quando a profundidade global do locus é inferior a 10x ou o número de leituras que abrangem uma ou ambas as extremidades de corte é inferior a 5.

Tabela 10 Campos INFO adicionais

Campo	Descrição
FIM	Posição da última base da região de repetição na referência
REF	Número de cópia de referência
REPID	ID da variante do catálogo de variantes
RL	Comprimento de referência em bp
RU	Unidade de repetição na orientação de referência
VARID	ID da variante do catálogo de variantes

Tabela 11 Campos GENÓTIPO (por amostra)

Campo	Descrição
AD	Profundidades alélicas para os alelos ref e alt pela ordem listada
ADFL	Número de leituras de flanco consistentes com o alelo
ADIR	Número de leituras em repetição consistentes com o alelo
ADSP	Número de leituras de amplitude consistentes com o alelo
DST	Resultados (+ detetados, - não detetados, ? indeterminados) do teste representado pela variante
GT	Genótipo
LC	Cobertura do locus
REPCI	Intervalo de confiança para REPCN
REPCN	Número de unidades de repetição abrangidas pelo alelo
RPL	Razão de probabilidade logarítmica para a presença do alelo de referência
SO	Tipo de leituras que suportam o alelo. Os valores podem ser <code>SPANNING</code> (ABRANGER), <code>FLANKING</code> (FLANQUEAR), ou <code>INREPEAT</code> (NA REPETIÇÃO). Estes valores indicam se as leituras abrangem, flanqueiam ou estão totalmente contidas na repetição.

O arquivo `<Sample_ID>.repeats.vcf.gz` inclui a saída SMN juntamente com quaisquer repetições direcionadas. A saída SMN é representada como uma identificação SNV única na posição de afetação da junção em SMN1 (NM_000344.3:c.840C/T) com estado de Atrofia Muscular Espinal (AME) nos seguintes campos personalizados.

Tabela 12 Resultados SMA no ficheiro de saída repeats.vcf

Campo	Descrição
VARID	SMN marca a identificação SMN.
GT	Identificação de genótipo nesta posição utilizando um modelo de genótipo normal (diploide).
DST	Identificação de estado SMA: + indica detetado - indica não detetado ? indica indeterminado
AD	Total de contagens de leitura que suportam os alelos C e T.
RPL	Razão de probabilidade Log10 entre os modelos não afetados e afetados. As pontuações positivas indicam que o modelo não afetado é mais provável.

ROH BED

As regiões de homoziguidade (ROH) são detetadas como parte do identificador de variantes pequenas. O identificador deteta e produz os ensaios de homoziguidade a partir de identificações de todo o genoma em cromossomas humanos autossómicos. Os cromossomas sexuais são ignorados a menos que o cariótipo sexual da amostra seja XX, conforme determinado pelo Estimador de Ploidia. O resultado do ROH permite que as ferramentas a jusante analisem e prevejam a consanguinidade entre os progenitores do participante probando.

Uma região é definida como identificações de variantes consecutivas no cromossoma sem grande intervalo entre estas variantes. Por outras palavras, as regiões são discriminadas por cromossoma ou por grandes intervalos sem identificações SNV. O tamanho do intervalo está definido para 3 Mbases.

O identificador ROH produz um ficheiro de saída ROH denominado `<Sample_ID>.roh.bed` no diretório `<Sample_ID>/Analysis`. Cada linha representa uma região de homoziguidade. O ficheiro bed contém as seguintes colunas:

```
Cromossoma Início Fim Pontuação #Homozigótico #Heterozigótico
```

em que

- A pontuação é uma função do número de variantes homozigóticas e heterozigóticas, em que cada variante homozigótica aumenta a pontuação em 0,025 e cada variante heterozigótica reduz a pontuação em 0,975.
- As posições início e fim são um intervalo semi-aberto baseado em 0.
- #Homozigótico é o número de variantes homozigóticas na região.
- #Heterozigótico é o número de variantes heterozigóticas na região.

O identificador também produz um ficheiro de métrica chamado `<Sample_ID>.roh_metrics.csv` que lista o número de ROH grandes e a percentagem de SNP em ROH grandes (> 3 MB).

Métricas de Estimativa de Ploidia

O Estimador de Ploidia é executado por predefinição. O Estimador de Ploidia utiliza leituras do mapeador/alinhador para calcular a profundidade de cobertura da sequenciação para cada autossoma e alosoma no genoma humano. O cariótipo sexual da amostra é então estimado utilizando os rácios das coberturas dos cromossomas sexuais medianos para a cobertura autossómica mediana. XX ou XY, e CONCORDANTE, DISCORDANTE ou ND (Não Determinado) em comparação com os dados de amostra fornecidos são comunicados no relatório consolidado. Os resultados detalhados, incluindo cada cobertura mediana normalizada por contig, são apresentados no ficheiro `<Sample_ID>.ploidy_estimation_metrics.csv`.

Ficheiros FASTQ

FASTQ (*.fastq.gz, *.fastq.ora) é um formato de ficheiro baseado em texto, que contém identificações de base e valores de qualidade por leitura. Cada ficheiro contém a seguinte informação:

- O identificador da amostra

- A sequência
- Um sinal de adição (+)
- As pontuações de qualidade Phred em formato com codificação ASCII + 33

O software gera um ficheiro FASTQ para cada amostra, leitura e faixa. Por exemplo, para cada amostra num ensaio de extremidade emparelhada, o software gera dois ficheiros FASTQ: um para a Leitura 1 e outro para a Leitura 2. Além destes ficheiros FASTQ de amostra, o software gera dois ficheiros FASTQ por faixa contendo todas as amostras desconhecidas. Os ficheiros FASTQ para Leitura de índice 1 e Leitura de índice 2 não são gerados porque a sequência é incluída no cabeçalho de cada entrada FASTQ. O formato do nome do ficheiro é construído a partir dos campos especificados na folha de amostras e utiliza o formato de nomeação do ficheiro de `<Sample_ID>_S#_L00#_R#_001.fastq.gz`. Os ficheiros FASTQ são guardados no diretório `<Sample_ID>/Conversion`. No diretório FASTQ da pasta de análise, pode encontrar o diretório Registos com registos de conversão BCL-para-FASTQ e o diretório Relatórios que contém vários ficheiros de métricas de leitura e `SampleSheet.csv` utilizado para conversão FASTQ. Os ficheiros FASTQ de leituras indeterminadas encontram-se no diretório `Indeterminado/Conversão da pasta Análise`.

O identificador da amostra está formatado como segue:

```
@Instrument:RunID:FlowCellID:Lane:Tile:X:Y ReadNum:FilterFlag:0:SampleNumber
```

Exemplo:

```
@SIM:1:FCX:1:15:6329:1045 1:N:0:2
```

```
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
```

```
+
```

```
<>;##=><9=AAAAAAAAAAA9#:<#<;<<<????#=#
```

Ficheiros CRAM

Os ficheiros CRAM ou Mapa de alinhamento orientado para a referência comprimido (*.cram) são armazenados no diretório `<Sample_ID>/Analysis` e contêm cabeçalhos e registos de alinhamento relativos ao ficheiro de referência genómica utilizado durante o alinhamento. O caminho para o ficheiro de referência é listado no ficheiro `<Sample_ID>/Analysis/<Sample_ID>-replay.json`, como um parâmetro `--ht-reference`, por predefinição definido para `hg38.fa`.

Os ficheiros CRAM contêm uma secção de cabeçalho e uma secção de alinhamento:

- **Header** (Cabeçalho) — Contém informações sobre todo o ficheiro, como, por exemplo, o nome da amostra, o tamanho da amostra e o método de alinhamento. Os alinhamentos da secção de alinhamentos estão associados a informações específicas da secção de cabeçalho.
- **Alignments** (Alinhamentos) — Contém o nome da leitura, a sequência da leitura, a qualidade da leitura, informações de alinhamento e etiquetas personalizadas. O nome da leitura inclui o cromossoma, a coordenada de início, qualidade do alinhamento e o descritor de correspondências.

A secção de alinhamentos inclui a seguinte informação para cada leitura ou par de leituras:

- AS: Extremidades emparelhadas com qualidade de alinhamento.
- RG: Grupo de leitura, que indica o número de leituras numa determinada amostra.
- BC: Etiqueta de codificação de barras, que indica o ID da amostra da demultiplexagem, associado à leitura.
- SM: Extremidade única com qualidade de alinhamento.
- XC: Descritor de correspondências.
- XN: O nome da etiqueta do amplicão, que regista o ID do amplicão associado à leitura

Para visualizar os registos de alinhamento, `samtools` pode ser utilizado como `samtools view --reference <path_to_reference_folder>/hg38.fa <Sample_ID>.cram`.

Também são gerados um ficheiro de índice e um ficheiro de soma de verificação.

Assistência Técnica

Para obter assistência técnica, contacte o Suporte Técnico da Illumina.

Website: www.illumina.com

E-mail: techsupport@illumina.com

Fichas de dados de segurança (SDS) — Disponíveis no website da Illumina em support.illumina.com/sds.html.

Documentação do produto — Disponível para descarregamento em support.illumina.com.



Illumina, Inc.
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
The Netherlands

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

PARA UTILIZAÇÃO EM DIAGNÓSTICO IN VITRO.

© 2024 Illumina, Inc. Todos os direitos reservados.

illumina[®]