

Sequenciamento do genoma completo abrangente com Illumina Complete Long Read Prep, Human

Desempenho excepcional em
regiões difíceis de mapear



Introdução

O sequenciamento de última geração (NGS) permite que os cientistas decifrem o genoma para uma compreensão mais profunda da biologia. A comprovada química de sequenciamento por síntese (SBS) da Illumina combinada com a premiada análise secundária DRAGEN™ fornece dados de sequenciamento do genoma completo (WGS) com precisão excepcional.¹ Ainda assim, uma pequena fração do genoma pode se beneficiar de duração da leitura mais longa para permitir uma resolução e mapeamento mais precisos dessas regiões desafiadoras. No entanto, muitas soluções de sequenciamento de leitura longa têm fluxos de trabalho complexos, altos requisitos de entrada de DNA e resultados altamente variáveis.²⁻⁵

O Illumina Complete Long Reads torna o sequenciamento de leitura longa acessível e simplificado para cientistas genômicos. O Illumina Complete Long Read Prep, Human é o primeiro produto baseado nesta nova química de leitura longa. O ensaio de WGS de alto desempenho usa um fluxo de trabalho NGS padrão para gerar sequências de leitura longa contíguas no NovaSeq™ 6000 System e NovaSeq X Series (Figura 1). O protocolo eficiente de preparação de biblioteca de um dia é fácil de dimensionar para estudos de alto rendimento e requer apenas 10 ng de entrada de DNA sem extrações especializadas, cisalhamento ou seleção de tamanho.

Esta nota técnica demonstra o desempenho excepcional do Illumina Complete Long Read Prep, Human com uma ampla variedade de entradas de DNA, qualidade de DNA e tipos de amostras para WGS altamente preciso e abrangente para análise de variantes de linha genética.

Métodos

A amostra de referência HG002/NA24385 foi obtida do Genome in a Bottle consortium (GIAB) como DNA genômico purificado (gDNA). Separadamente, foram obtidas amostras de sangue e saliva e o gDNA foi extraído usando kits comercialmente disponíveis com extração regular ou de alto peso molecular, seguindo as instruções do fabricante.

Preparação da biblioteca

As bibliotecas “marcadas” foram preparadas a partir de quantidades variáveis de gDNA de entrada usando Illumina Complete Long Read Prep, Human (Illumina, Catálogo nº 20089108). Resumidamente, fragmentos de DNA longos de molécula única são marcados enzimaticamente com padrões exclusivos (marcados). Bibliotecas não marcadas, usadas durante a análise para produzir leituras longas e contíguas representando o fragmento original de molécula única, foram preparadas usando Illumina DNA PCR-Free Prep, Tagmentation (Illumina, Catálogo nº 20041794).

Sequenciamento

As bibliotecas marcadas e não marcadas foram sequenciadas durante corridas separadas no NovaSeq 6000 System com duração da leitura de 2 × 150 bp.

Análise de dados

A montagem e a renderização de leituras com marcação de referência foram realizadas usando o aplicativo Illumina Complete Long Read WGS no BaseSpace™ Sequence Hub. O Interactive Genomics Viewer (IGV) foi usado para análise e visualização de dados.



Figura 1: Fluxo de trabalho do Illumina Complete Long Read Prep, Human: o Illumina Complete Long Read Prep oferece um fluxo de trabalho simplificado que consiste em quatro etapas: Extração de DNA, preparação de bibliotecas, sequenciamento e análise de dados.

a. Embora entradas tão baixas quanto 10 ng sejam possíveis, a Illumina recomenda 50 ng de DNA

Dados de alta qualidade em regiões desafiadoras

O Illumina Complete Long Read Prep, Human fornece cobertura em regiões em que leituras curtas não podem ser mapeadas de forma inequívoca. Isso permite o sequenciamento contíguo de genes e pseudogenes associados à doença com problemas de mapeamento conhecidos apenas com leituras curtas, incluindo *STRC*, (Figura 2A), *SEM1* (Figura 2B) e *SULTIA1* (Figura 2C).

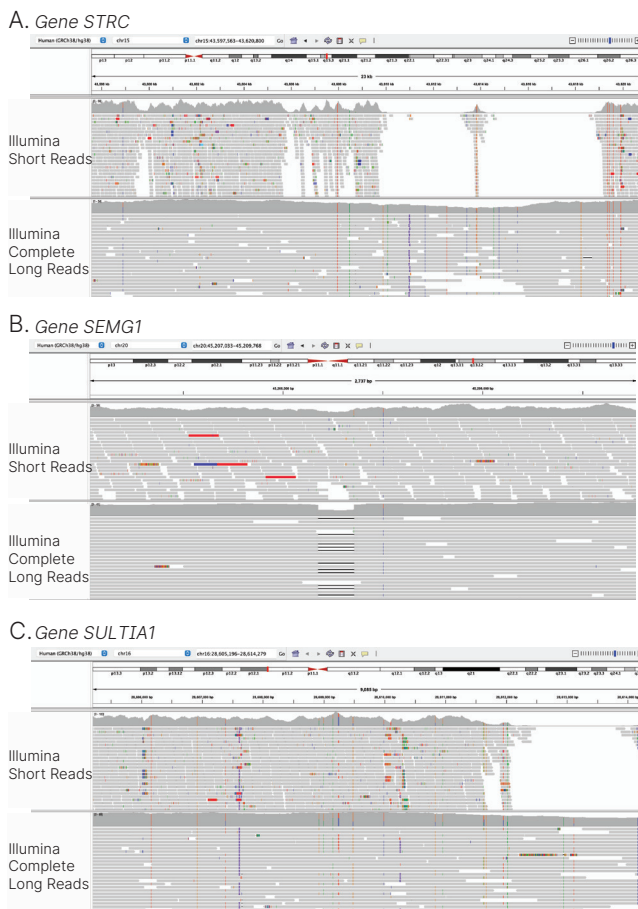


Figura 2: A cobertura abrangente permite a detecção precisa de vários tipos de variantes em regiões desafiadoras: o Illumina Complete Long Read Prep, Human permite o sequenciamento preciso de genes associados a doenças em regiões genômicas desafiadoras, incluindo (A) *STRC*, (B) *SEM1* e (C) *SULTIA1*.

Além disso, o Illumina Complete Long Read Prep, Human fornece blocos de fase significativamente maiores do que o tamanho típico do gene. Isso permite uma cobertura uniforme em todo o loco do antígeno leucocitário humano (HLA), uma região densamente polimórfica com altos níveis de homologia de sequência, pseudogenes, variabilidade e vários alelos. Embora seja difícil sequenciar com precisão, o Illumina Complete Long Read Prep, Human oferece cobertura uniforme em toda a região do HLA e permite que o faseamento do genoma visualize e distinga claramente os alelos entre cromossomos maternos e paternos (Figura 3).

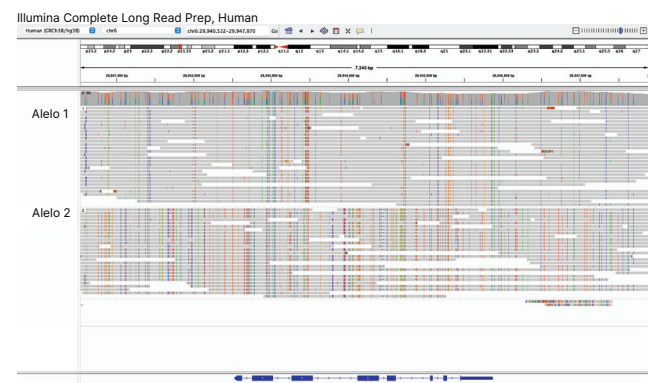


Figura 3: A cobertura uniforme permite o faseamento claro do haplótipo: o Illumina Complete Long Read Prep, Human fornece cobertura abrangente na região do HLA e permite o faseamento claro e preciso dos alelos dentro da região do HLA.

Para uma avaliação mais abrangente da precisão de identificação de variantes, o Illumina Long Read Prep, Human foi comparado com uma plataforma de leitura longa de terceiros e leituras curtas da Illumina no PrecisionFDA Truth Challenge V2. Patrocinado pela PrecisionFDA, pela GIAB e pelo National Institute of Standards and Technology (NIST), esse desafio foi lançado para avaliar o desempenho do pipeline de identificação de pequenas variantes em um quadro comum de referência, com foco em regiões de “difícil mapeamento”. Usando esse conjunto de dados, Illumina Complete Long Read Prep, Human demonstra precisão excepcional e supera a plataforma alternativa de leitura longa com uma pontuação F1 de 99,90% (Figura 4).

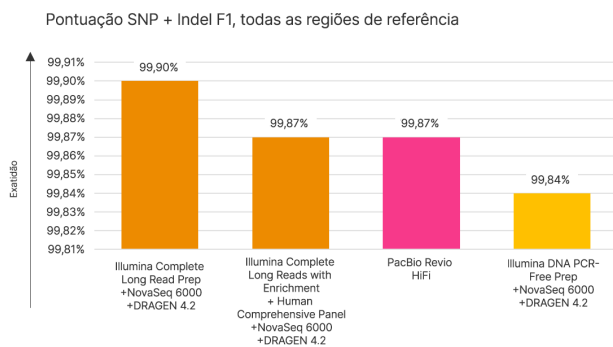


Figura 4: Um novo padrão de precisão: o Illumina Complete Long Read Prep, Human oferece precisão sem precedentes para identificação de variantes, conforme medido pela pontuação F1 (%), refletindo precisão e recall para WGS.¹

Desempenho robusto em quantidades de entrada de DNA

O desempenho do Illumina Complete Long Read Prep, Human foi avaliado em uma variedade de quantidades de entrada de DNA. As bibliotecas foram preparadas a partir de 5, 10, 25, 50, 100, 600 e 1.200 ng de DNA em triplicata. Os resultados do sequenciamento demonstraram que as bibliotecas preparadas a partir de apenas 5 ng de entrada exibiram métricas de alta qualidade, incluindo N50 e bloco de faseamento N50 (Figura 5). Embora a Illumina recomende uma entrada de 50 ng de DNA, entradas mais baixas podem ser usadas.

Suporte para qualidade de DNA variável

Para simular o DNA danificado, o gDNA de alta qualidade foi submetido a cisalhamento usando um ultrassônico Covaris por um ou três segundos. A qualidade do DNA foi avaliada usando um analisador de fragmentos (Figura 6). As bibliotecas foram preparadas usando entrada de gDNA cisalhado e Illumina Complete Long Read Prep, Human. Os resultados do sequenciamento mostraram que o DNA cisalhado por um segundo gerou dados de alta qualidade e permitiu a identificação precisa de variantes (Figura 7). Além disso, o gDNA de entrada que foi submetido a ciclos crescentes de congelamento e descongelamento produziu dados de alta qualidade (Figura 8).

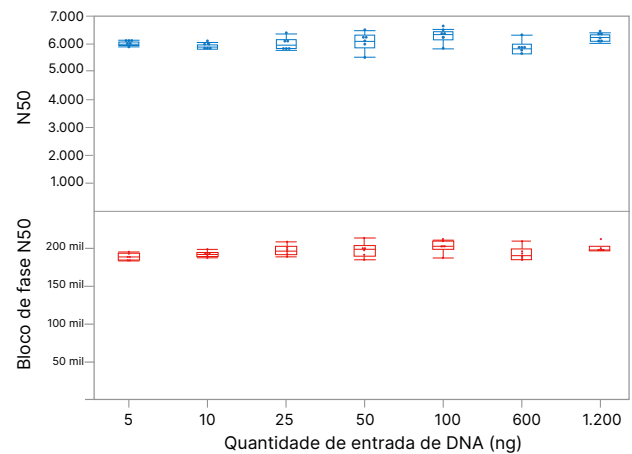


Figura 5: Desempenho de alta qualidade em uma ampla variedade de entradas de DNA: o Illumina Complete Long Read Prep, Human produz bibliotecas de alta qualidade preparadas a partir de entradas de DNA de 5 ng a 1.200 ng (em triplicata) que geram qualidade dos dados semelhante para N50 e bloco de fase N50. N50 é definido como o comprimento da sequência do contig mais curto a 50% do comprimento total da montagem; também pode ser usado como uma medida dos blocos de fase.

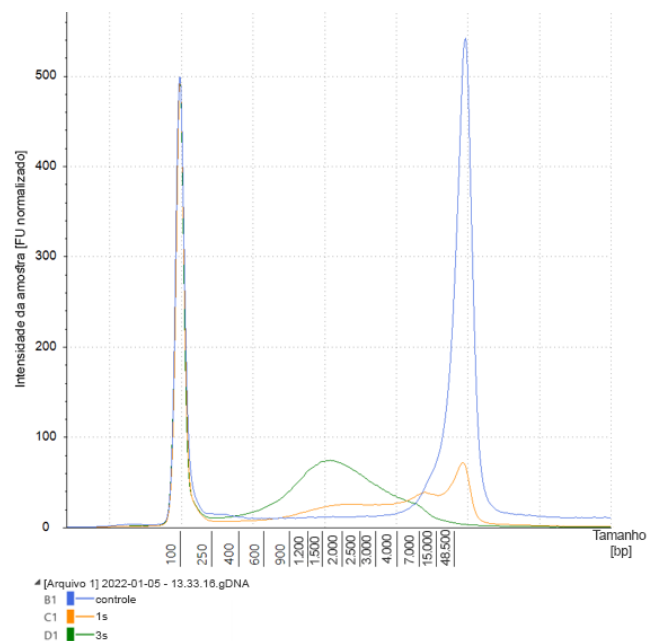


Figura 6: Qualidade do DNA cisalhado: a análise do DNA em um analisador de fragmentos mostra diminuição da qualidade com o aumento do tempo de cisalhamento. As amostras de DNA foram submetidas a cisalhamento de Covaris por 1 ou 3 segundos e a qualidade foi medida usando um analisador de fragmentos.



Figura 7: Métricas de sequenciamento de DNA cisalhado: o DNA cisalhado por 1 e 3 segundos foi usado como entrada para o Illumina Complete Long Read Prep, Human. As bibliotecas resultantes geraram dados de sequenciamento de leitura longa de alta qualidade, conforme medidos por N50 e bloco de fase N50.

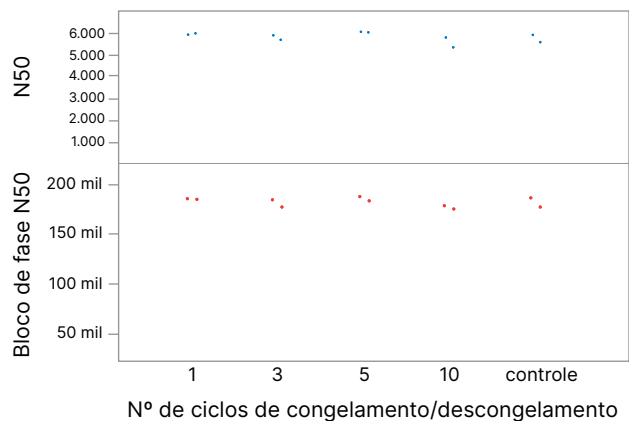


Figura 8: Métricas de sequenciamento de DNA congelado-descongelado: o DNA submetido a 1, 3, 5 e 10 ciclos de congelamento-descongelamento foi usado como entrada para o Illumina Complete Long Read Prep, Human. As bibliotecas resultantes geraram dados de sequenciamento de leitura longa de alta qualidade, conforme medidos por N50 e bloco de fase N50.

Suporte flexível para vários tipos de amostras

Illumina Complete Long Read Prep, Human é compatível com vários tipos de amostras, incluindo sangue e saliva. Após a extração usando kits padrão ou de alto peso molecular (HMW), a qualidade do DNA foi avaliada usando um analisador de fragmentos (Figura 9). Os resultados do sequenciamento mostraram que o DNA do sangue e da saliva gerou dados de alta qualidade (Figura 10). É importante observar que não houve diferença no desempenho entre o DNA extraído com um kit padrão e um kit de HMW (Figura 10).

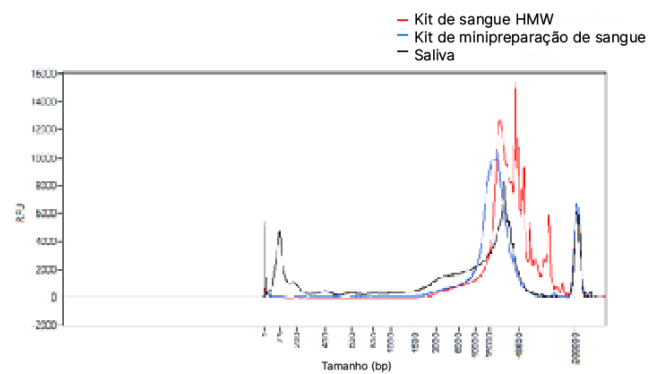


Figura 9: Qualidade do DNA de sangue e saliva: o DNA extraído do sangue (HMW, mini) ou saliva foi analisado em um analisador de fragmentos.

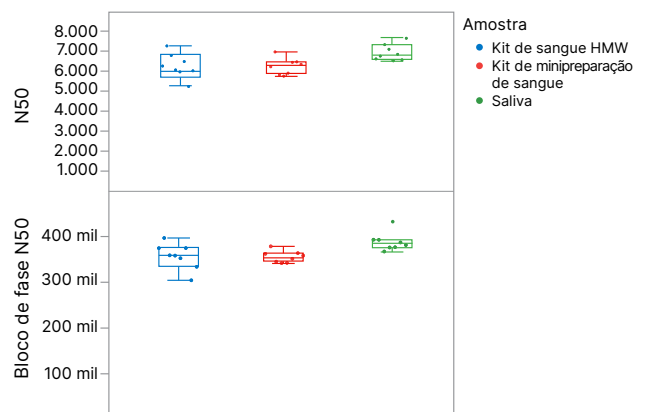


Figura 10: Desempenho de alta qualidade com diferentes tipos de amostras: Illumina Complete Long Read Prep, bibliotecas humanas geradas a partir de sangue (HMW, mini) e saliva fornecem dados de alta qualidade, conforme medidos por N50 e bloco de fase N50.

Resumo

A química de sequenciamento de leitura longa fornece informações adicionais para regiões do genoma que são difíceis de resolver com NGS de leitura curta. O Illumina Complete Long Reads torna o sequenciamento de leitura longa acessível e simplificado para cientistas genômicos. O Illumina Complete Long Read Prep, Human, o primeiro produto construído com base nessa nova química, oferece um fluxo de trabalho e sinergia simplificados e familiares com química comprovada Illumina SBS e análise DRAGEN. Essa química altamente inovadora permite que o sequenciamento de leitura curta e longa seja realizado em um único instrumento, tornando o NGS de leitura longa acessível para laboratórios genômicos. Illumina Complete Long Read Prep, Human demonstra desempenho robusto com DNA de quantidade e qualidade variáveis de diferentes fontes de amostra. O resultado é uma solução WGS humana altamente escalável e precisa.

Saiba mais

[Illumina Complete Long Read Prep, Human](#)

Referências

1. Mehio R, Ruehle M, Catreux S, et al. DRAGEN Wins at Precision- FDA Truth Challenge V2 Showcase Accuracy Gains from Alt-aware Mapping and Graph Reference Genomes. illumina.com/science/genomics-research/dragen-wins-precisionfda-challenge-showcase-accuracy-gains.html. Acessado em 12 de janeiro de 2023.
2. Pacific Biosciences. Preparing DNA for PacBio HiFi sequencing—Extraction and quality control. pacb.com/wp-content/uploads/Technical-Note-Preparing-DNA-for-PacBio-HiFi-Sequencing-Extraction-and-Quality-Control.pdf. Publicado em 2022. Acessado em 12 de janeiro de 2023.
3. Pacific Biosciences. Preparing whole genome and metagenome libraries using SMRTbell prep kit 3.0. pacb.com/wp-content/uploads/Procedure-checklist-Preparing-whole-genome-and-metagenome-libraries-using-SMRTbell-prep-kit-3.0.pdf. Publicado em 2022. Acessado em 12 de janeiro de 2023.
4. Oxford Nanopore Technologies. Ligation Sequencing Kit. store.nanoporetech.com/us/ligation-sequencing-kit110.html. Acessado em 12 de janeiro de 2023.
5. Pacific Biosciences. Low Yield Troubleshooting Guide. pacb.com/wp-content/uploads/Guide-Low-Yield-Troubleshooting.pdf. Publicado em 2018. Acessado em 12 de janeiro de 2023.



1 800-809-4566, ligação gratuita (EUA) | tel. +1 (858) 202-4566
techsupport@illumina.com | www.illumina.com

© 2024 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.
M-GL-01421 PTB v2.0