

Secuenciación del exoma completo exhaustiva con Illumina Complete Long Read Prep, Human

Rendimiento excepcional en regiones complejas y difíciles de asignar



Introducción

La secuenciación de nueva generación (NGS, Next-Generation Sequencing) permite a los científicos descodificar el genoma para comprender mejor las características biológicas. La química de secuenciación por síntesis (SBS, Sequencing By Synthesis) probada de Illumina, combinada con el galardonado análisis secundario de DRAGEN™, ofrece datos de secuenciación del genoma completo (WGS, Whole-Genome Sequencing) con una precisión excepcional.¹ Aun así, una pequeña fracción del genoma puede beneficiarse de longitudes de lectura más largas para permitir una resolución y asignación más precisas de estas regiones complejas. Sin embargo, muchas soluciones de secuenciación de lectura larga tienen flujos de trabajo complejos, elevados requisitos de aporte de ADN y resultados muy variables.²⁻⁵

Illumina Complete Long Read optimiza la secuenciación de lectura larga y la hace accesible a los científicos genómicos. Illumina Complete Long Prep, Human es el primer producto que se basa en esta novedosa química de lectura larga. Este ensayo de WGS de alto rendimiento emplea un flujo de trabajo de NGS estándar para generar secuencias contiguas de lectura larga en NovaSeq™ 6000 System y NovaSeq X Series (Figura 1). El eficiente protocolo de preparación de librerías en un solo día es fácil de redimensionar para estudios de alta productividad y requiere únicamente 10 ng de aporte de ADN sin extracciones especializadas, fraccionamiento o selección de tamaño.

Esta nota técnica demuestra el rendimiento excepcional de Illumina Complete Long Read Prep, Human con una amplia gama de aportes de ADN, calidad del ADN y tipos de muestras para una WGS completa y de alta precisión para el análisis de variantes germinales.

Métodos

La muestra de control HG002/NA24385 se obtuvo del consorcio Genome in a Bottle (GIAB) como ADN genómico (ADNg) purificado. Por otra parte, se obtuvieron muestras de sangre y saliva y se extrajo ADNg utilizando kits comerciales con extracción de peso molecular normal o elevado, siguiendo las instrucciones del fabricante.

Preparación de librerías

Las librerías con «marcas de referencia» se prepararon a partir de diversas cantidades de aporte de ADNg mediante Illumina Complete Long Read Prep, Human (Illumina, n.º de catálogo 20089108). En pocas palabras, los fragmentos largos de ADN de molécula única se marcan enzimáticamente con patrones únicos (marcas de referencia). Las librerías sin marcar, utilizadas durante el análisis para producir lecturas largas contiguas que representaban el fragmento de molécula única original, se prepararon con Illumina DNA PCR-Free Prep, Tagmentation (Illumina, n.º de catálogo 20041794).

Secuenciación

Las librerías con marcas de referencia y sin marcas se secuenciaron en experimentos separados en NovaSeq 6000 System con una longitud de lectura de 2 × 150 pb.

Análisis de datos

El ensamblaje y la representación de las lecturas marcadas como referencia se llevaron a cabo con la aplicación Illumina Complete Long Read WGS en BaseSpace™ Sequence Hub. Se utilizó Interactive Genomics Viewer (IGV) para el análisis y la visualización de datos.



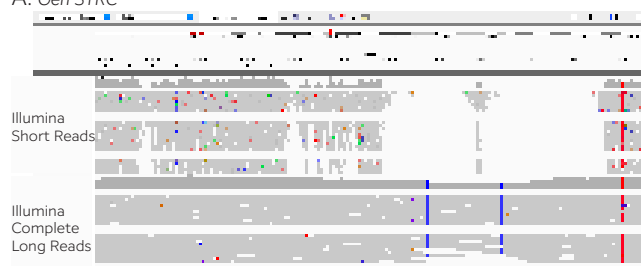
Figura 1: Flujo de trabajo de Illumina Complete Long Read Prep, Human. Illumina Complete Long Read Prep, Human ofrece un flujo de trabajo optimizado que consta de cuatro pasos: extracción de ADN, preparación de librerías, secuenciación y análisis de datos.

a. Aunque son posibles aportes de hasta 10 ng, Illumina recomienda 50 ng de aporte de ADN

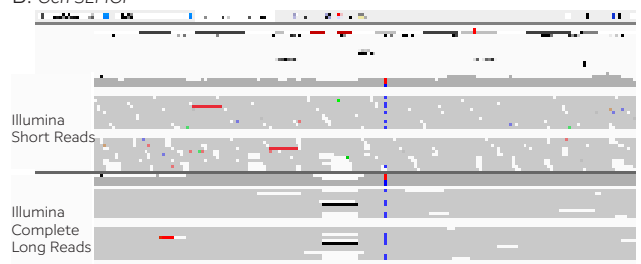
Datos de alta calidad en regiones complejas

Illumina Complete Long Read Prep, Human proporciona cobertura en regiones en las que las lecturas cortas no se pueden asignar sin ambigüedades. Esto permite la secuenciación contigua de genes y pseudogenes asociados a enfermedades con problemas de asignación conocidos cuando solo se emplean lecturas cortas, entre los que se incluyen *STRC*, (Figura 2A), *SEMG1* (Figura 2B) y *SULTIA1* (Figura 2C).

A. Gen *STRC*



B. Gen *SEMG1*



C. Gen *SULTIA1*

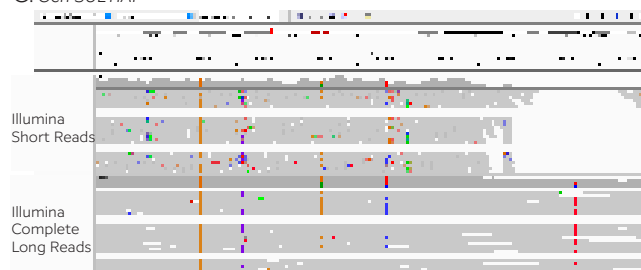


Figura 2: La cobertura completa permite una detección precisa de diversos tipos de variantes en regiones complejas. Illumina Complete Long Read Prep, Human posibilita la secuenciación precisa de genes asociados a enfermedades en regiones genómicas complejas, entre las que se incluyen (A) *STRC*, (B) *SEMG1* y (C) *SULTIA1*.

Además, Illumina Complete Long Read Prep, Human proporciona bloques de fase significativamente más grandes que el tamaño típico del gen. Esto permite una cobertura uniforme en todo el locus del antígeno leucocitario humano (HLA, Human Leukocyte Antigen), una región densamente polimórfica con altos niveles de homología de secuencias, pseudogenes, variabilidad y numerosos alelos. Illumina Complete Long Read Prep, Human proporciona una cobertura uniforme en toda la región del HLA y facilita la fase de hebra retrasada genómica para visualizar claramente y distinguir los alelos entre los cromosomas maternos y paternos, aunque sean difíciles de secuenciar con precisión (Figura 3).

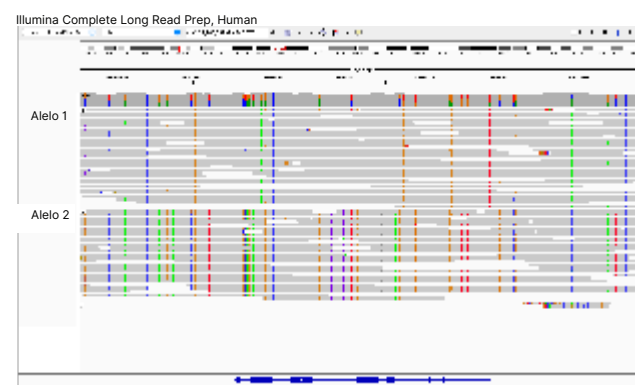


Figura 3: La cobertura uniforme permite una clara fase de hebra retrasada de haplotipo. Illumina Complete Long Read Prep, Human proporciona una cobertura completa en la región del HLA y posibilita una fase de hebra retrasada clara y precisa de los alelos en la región del HLA.

Para una evaluación más completa de la precisión de las llamadas de variantes, se comparó Illumina Long Read Prep, Human con una plataforma de lectura larga de terceros y las lecturas cortas de Illumina en PrecisionFDA Truth Challenge V2. Patrocinado por PrecisionFDA, GIAB y el National Institute of Standards and Technology (NIST), este desafío se lanzó para evaluar el rendimiento del proceso de llamada de variantes pequeños en un marco común de referencia, centrándose en las regiones «difíciles de asignar». Con este conjunto de datos, Illumina Complete Long Read Prep, Human demuestra una precisión excepcional y supera a la plataforma alternativa de lectura larga con una puntuación F1 del 99,87 % (Figura 4).

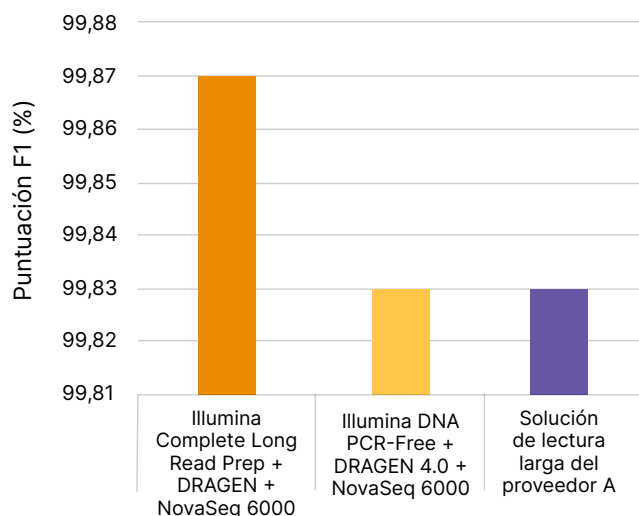


Figura 4: Un nuevo estándar de precisión. Illumina Complete Long Read Prep, Human ofrece una precisión sin precedentes para la llamada de variantes, medida por la puntuación F1 (%), que refleja la precisión y la recuperación para la WGS.¹

Rendimiento sólido en todas las cantidades de aporte de ADN

El rendimiento de Illumina Complete Long Read Prep, Human se evaluó con una serie de cantidades de aporte de ADN. Las librerías se prepararon a partir de 10, 25, 100, 600 y 1200 ng de ADN por triplicado. Los resultados de secuenciación demostraron que las librerías preparadas a partir de tan solo 5 ng de aporte mostraban criterios de medición de alta calidad, incluidos N50 y el bloque de fase de hebra retrasada N50 (Figura 5). Aunque Illumina recomienda un aporte de ADN de 50 ng, se pueden utilizar aportes más bajos.

Compatibilidad con diversas calidades del ADN

A fin de simular el ADN dañado, el ADNg de alta calidad se sometió a fraccionamiento con un dispositivo de ultrasonidos Covaris durante uno o tres segundos. La calidad del ADN se evaluó con un Fragment Analyzer (Figura 6). Las librerías se prepararon con un aporte de ADNg fraccionado e Illumina Complete Long Read Prep, Human. Los resultados de secuenciación mostraron que el ADN fraccionado durante un segundo generaba datos de alta calidad y permitía una llamada de variantes precisa (Figura 7). Además, el aporte de ADNg sometido a ciclos de congelación y descongelación cada vez mayores produjo datos de alta calidad (Figura 8).

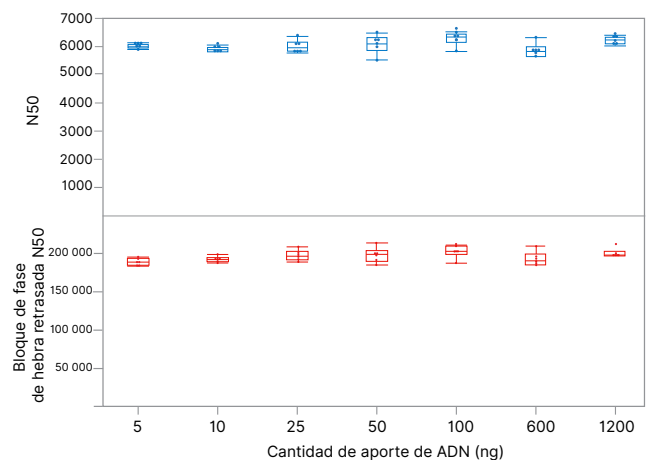


Figura 5: Rendimiento de alta calidad en una amplia gama de aportes de ADN. Illumina Complete Long Read Prep, Human produce librerías de alta calidad preparadas a partir de aportes de ADN de 5 ng a 1200 ng (por triplicado) que generan una calidad de datos similar para N50 y el bloque de fase de hebra retrasada N50. N50 se define como la longitud de secuencia del cóntigo más corto al 50 % de la longitud total del conjunto; también puede usarse como una medida de los bloques de fase de hebra retrasada.

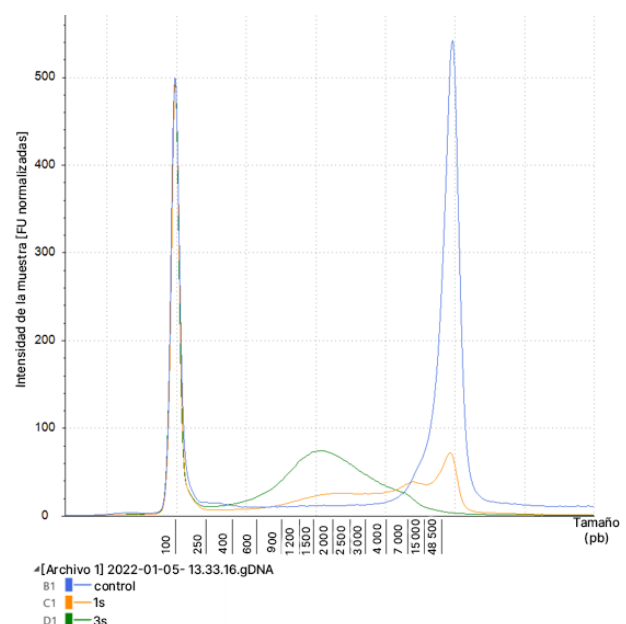


Figura 6: Calidad del ADN fraccionado. El análisis del ADN en un Fragment Analyzer muestra una calidad reducida con un mayor tiempo de fraccionamiento. Las muestras de ADN se sometieron a fraccionamiento en el dispositivo Covaris durante 1 o 3 segundos y la calidad se midió con un Fragment Analyzer.

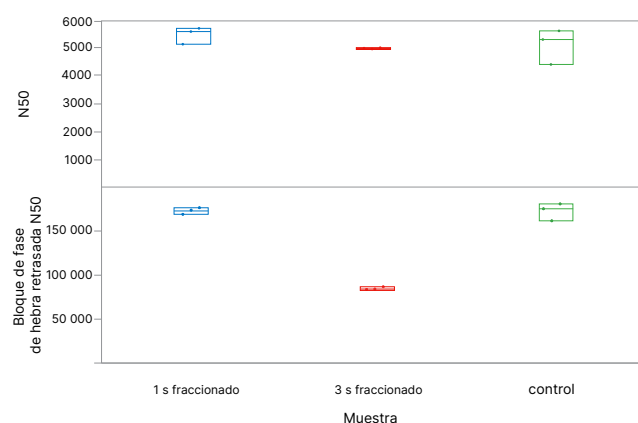


Figura 7: Criterios de medición de secuenciación de ADN fragmentado. Se utilizó ADN fraccionado durante 1 y 3 segundos como aporte para Illumina Complete Long Read Prep, Human. Las librerías resultantes generaron datos de secuenciación de lectura larga de alta calidad, medidos por N50 y el bloque de fase de hebra retrasada N50.

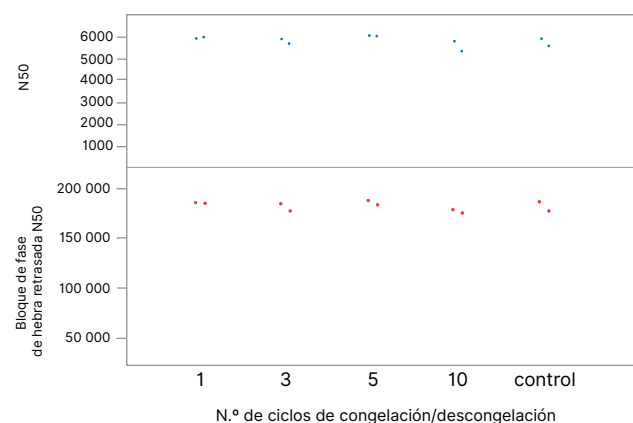


Figura 8: Criterios de medición de secuenciación de ADN congelado-descongelado. Se utilizó ADN sometido a 1, 3, 5 y 10 ciclos de congelación-descongelación como aporte para Illumina Complete Long Read Prep, Human. Las librerías resultantes generaron datos de secuenciación de lectura larga de alta calidad, medidos por N50 y el bloque de fase de hebra retrasada N50.

Soporte flexible para varios tipos de muestras

Illumina Complete Long Read Prep, Human es compatible con varios tipos de muestras, como sangre y saliva. Tras la extracción con kits de peso molecular estándar o elevado (HMW, High Molecular Weight), se evaluó la calidad del ADN con un Fragment Analyzer (Figura 9). Los resultados de secuenciación mostraron que el ADN de la sangre y la saliva generaba datos de alta calidad (Figura 10). Cabe destacar que no hubo diferencias de rendimiento entre el ADN extraído con un kit estándar y un kit para HMW (Figura 10).

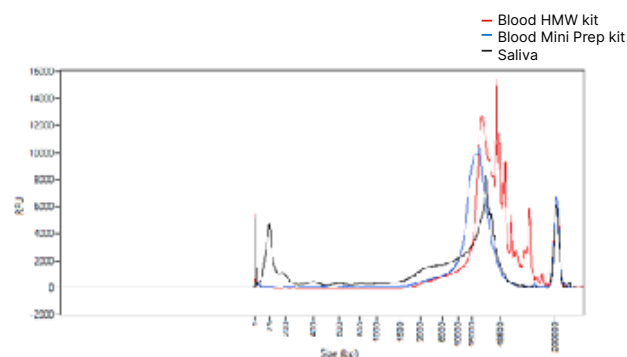


Figura 9: Calidad del ADN de sangre y saliva. El ADN extraído de sangre (HMW, mini) o saliva se analizó en un Fragment Analyzer.



Figura 10: Rendimiento de alta calidad con diferentes tipos de muestras. Las librerías de Illumina Long Read Prep, Human generadas a partir de sangre (HMW, mini) y saliva proporcionan datos de alta calidad, medida mediante N50 y el bloque de fase de hebra retrasada N50.

Resumen

La química de secuenciación de lectura larga proporciona información adicional para las regiones del genoma que son difíciles de resolver con la NGS de lectura corta. Illumina Complete Long Read hace que la secuenciación de lectura larga sea accesible y optimizada para los científicos genómicos. Illumina Complete Long Read Prep, Human, el primer producto basado en esta nueva química, ofrece un flujo de trabajo familiar y optimizado, y sinergia con la química probada de SBS de Illumina y el análisis de DRAGEN. Esta química bastante innovadora permite realizar la secuenciación de lectura corta y larga en un único instrumento, lo que hace que la NGS de lectura larga sea accesible para los laboratorios de genómica. Illumina Complete Long Read Prep, Human demuestra un rendimiento sólido con ADN de diferentes cantidades y calidades de diferentes fuentes de muestras. El resultado es una solución de WGS humana muy flexible y precisa.

Información adicional

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

Bibliografía

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. Fecha de consulta: 12 de enero 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. Año de publicación: 2022. Fecha de consulta: 12 de enero 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. Año de publicación: 2022. Fecha de consulta: 12 de enero de 2023.
4. Oxford Nanopore Technologies. Ligation Sequencing Kit. store.nanoporetech.com/us/ligation-sequencing-kit110.html. Fecha de consulta: 12 de enero de 2023.
5. Pacific Biosciences. Low Yield Troubleshooting Guide. pacb.com/wp-content/uploads/Guide-Low-Yield-Troubleshooting.pdf. Año de publicación: 2018. Fecha de consulta: 12 de enero de 2023.



1 800 809 4566 (llamada gratuita, EE. UU.) | Tel.: +1 858 202 4566
techsupport@illumina.com | www.illumina.com

© 2023 Illumina, Inc. Todos los derechos reservados. Todas las marcas comerciales pertenecen a Illumina, Inc. o a sus respectivos propietarios. Si desea consultar información específica sobre las marcas comerciales, consulte www.illumina.com/company/legal.html.
M-GL-01421 ESP v1.0