

Umfassende Genom- sequenzierung mit Illumina Complete Long Read Prep, Human

Herausragende Leistung
in anspruchsvollen, schwer
zu mappenden Regionen



Einleitung

Die Sequenzierung der nächsten Generation (NGS, Next-Generation Sequencing) ermöglicht Wissenschaftlern tiefgreifendere biologische Erkenntnisse durch die Entschlüsselung des Genoms. Die bewährte SBS-Chemie (Sequencing by Synthesis, Sequenzierung durch Synthese) von Illumina in Kombination mit der preisgekrönten DRAGEN™-Sekundäranalyse liefert WGS-Daten (Whole-Genome Sequencing, Genomsequenzierung) mit herausragender Genauigkeit.¹ Dennoch sind bei einem kleinen Teil des Genoms u. U. größere Read-Längen von Vorteil, da diese eine höhere Auflösung sowie ein genaueres Mapping dieser anspruchsvollen Regionen ermöglichen. Jedoch sind bei zahlreichen Long-Read-Sequenzierungslösungen die Workflows komplex, die Anforderungen an die DNA-Zugabe hoch und Ergebnisse extrem variabel.⁴⁻⁷

Illumina Complete Long Reads rationalisiert die Long-Read-Sequenzierung und macht sie für die Genomforschung verfügbar. Illumina Complete Long Read Prep, Human ist das erste Produkt dieser neuartigen Long-Read-Chemie. Dieser WGS-Hochleistungsassay generiert anhand eines standardmäßigen NGS-Workflows auf dem NovaSeq™ 6000 System und der NovaSeq X Series zusammenhängende Long-Read-Sequenzen (Abbildung 1). Das effiziente, an einem Tag durchführbare Protokoll zur Bibliotheksvorbereitung lässt sich leicht auf Hochdurchsatzstudien abstimmen und erfordert eine DNA-Zugabe von nur 10 ng ohne spezielle Extraktionen, Scherung oder Größenauswahl.

Dieser technische Hinweis verdeutlicht die herausragende Leistung von Illumina Complete Long Read Prep, Human bei einem breiten Spektrum an DNA-Zugaben, DNA-Qualitäten und Probenotypen für die hochgenaue, umfassende WGS zur Analyse von Keimbahnvarianten.

Methoden

Die Referenzprobe HG002/NA24385 wurde vom Genome in a Bottle Consortium (GIAB) als gereinigte genomische DNA (gDNA) bereitgestellt. Unabhängig davon wurden Blut- und Speichelproben entnommen und gDNA wurde mithilfe handelsüblicher Kits mit normaler oder hoher Molekulargewichtsextraktion gemäß den Anweisungen des Herstellers extrahiert.

Bibliotheksvorbereitung

Bibliotheken mit „Land-Marks“ wurden aus variierenden Mengen an zugegebener gDNA mit Illumina Complete Long Read Prep, Human (Illumina, Katalog-Nr. 20089108) vorbereitet. Kurz gesagt: lange Einzelmolekül-DNA-Fragmente werden enzymatisch mit eindeutigen Mustern (sogenannten „Land-Marks“) gekennzeichnet. Ungezeichnete Bibliotheken, die während der Analyse zur Erstellung langer zusammenhängender Reads verwendet wurden, die das ursprüngliche Einzelmolekülfragment darstellen, wurden mit Illumina DNA PCR-Free Prep, Tagmentation (Illumina, Katalog-Nr. 20041794) vorbereitet.

Sequenzierung

Bibliotheken mit und ohne „Land-Marks“ wurden während separater Läufe auf dem NovaSeq 6000 System mit einer Read-Länge von 2×150 bp sequenziert.

Datenanalyse

Die Assemblierung und das Rendering der Reads mit „Land-Marks“ wurden mit der Illumina Complete Long Read WGS-App auf BaseSpace™ Sequence Hub durchgeführt. Für die Datenanalyse und -visualisierung wurde Interactive Genomics Viewer (IGV) verwendet.

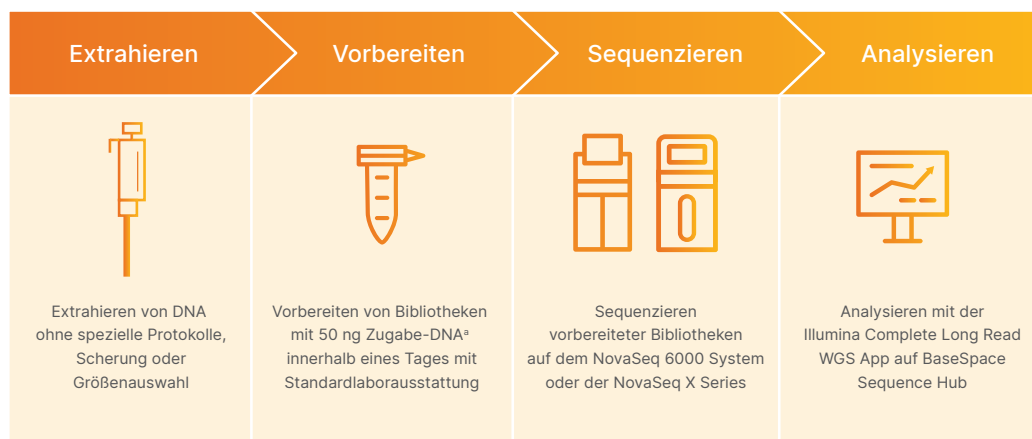


Abbildung 1: Workflow von Illumina Complete Long Read Prep, Human: Illumina Complete Long Read Prep, Human bietet einen optimierten Workflow mit vier Schritten: DNA-Extraktion, Bibliotheksvorbereitung, Sequenzierung und Datenanalyse.

a. Zugabemengen von nur 10 ng sind zwar möglich, Illumina empfiehlt jedoch die Verwendung von 50 ng DNA.

Hochwertige Daten in anspruchsvollen Regionen

Illumina Complete Long Read Prep, Human sorgt für Coverage in Regionen, in denen kurze Reads nicht eindeutig zugeordnet werden können. Dies ermöglicht eine zusammenhängende Sequenzierung von krankheitsassoziierten Genen und Pseudogenen, bei denen das Mapping ausschließlich mit kurzen Reads bekanntermaßen problematisch ist, darunter *STRC*, (Abbildung 2A), *SEMG1* (Abbildung 2B) und *SULTIA1* (Abbildung 2C).

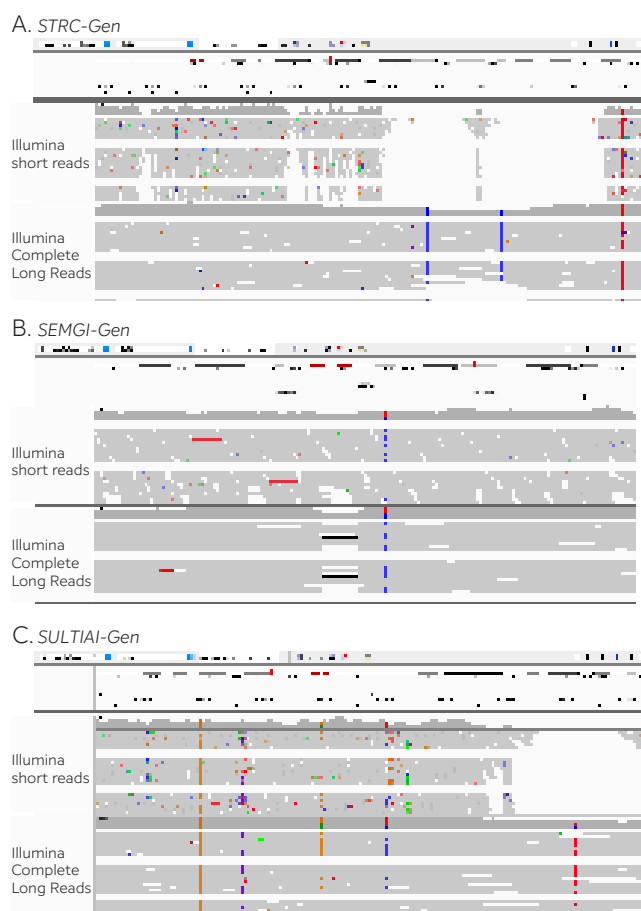


Abbildung 2: Die umfassende Coverage ermöglicht den genauen Nachweis mehrerer Variantentypen in anspruchsvollen Regionen: Illumina Complete Long Read Prep, Human ermöglicht die genaue Sequenzierung krankheitsassoziierten Gene in anspruchsvollen genomischen Regionen, darunter (A) *STRC*, (B) *SEMG1* und (C) *SULTIA1*.

Darüber hinaus liefert Illumina Complete Long Read Prep, Human Phasenblöcke, die die typische Gengröße deutlich übertreffen. Dies ermöglicht eine einheitliche Coverage des gesamten HLA-Locus (humanes Leukozytenantigen), einer dicht polymorphen Region mit hoher Sequenzhomologie, Pseudogenen, Variabilität und zahlreichen Allelen. Eine genaue Sequenzierung ist schwierig. Illumina Complete Long Read Prep, Human bietet jedoch eine einheitliche Coverage der gesamten HLA-Region und ermöglicht eine Genom-Phasierung, mit der sich sämtliche Allele klar visualisieren sowie nach mütterlichen und väterlichen Chromosomen unterscheiden lassen (Abbildung 3).

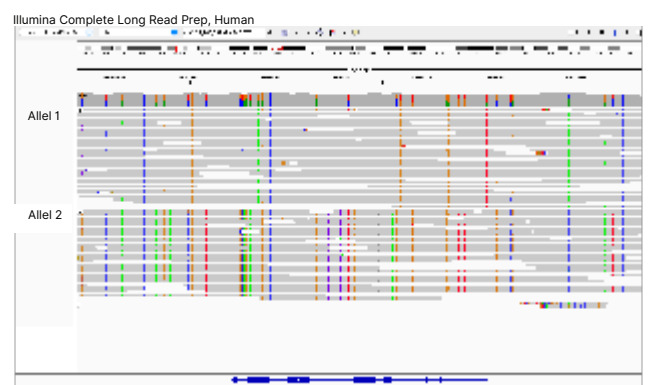


Abbildung 3: Die einheitliche Coverage ermöglicht eine klare Haplotyp-Phasierung: Illumina Complete Long Read Prep, Human bietet eine umfassende Coverage der HLA-Region und ermöglicht eine klare und genaue Phasierung der dort enthaltenen Allele.

Eine umfassendere Bewertung der Varianten-Calling-Genauigkeit erfolgte im Rahmen der PrecisionFDA Truth Challenge V2, bei der Illumina Long Read Prep, Human mit einer Long-Read-Plattform von Drittanbietern und Illumina-Short-Reads verglichen wurde. Diese von PrecisionFDA, GIAB und dem National Institute of Standards and Technology (NIST) veranstaltete Challenge diente dem Vergleich der Leistung von Pipelines für das Calling kleiner Varianten in einem gemeinsamen Bezugssystem, wobei der Schwerpunkt auf schwer zu mappenden Regionen lag. Bei diesem Datensatz weist Illumina Complete Long Read Prep, Human eine herausragende Genauigkeit auf und übertrifft mit einem F1-Score von 99,87 % die alternative Long-Read-Plattform (Abbildung 4).

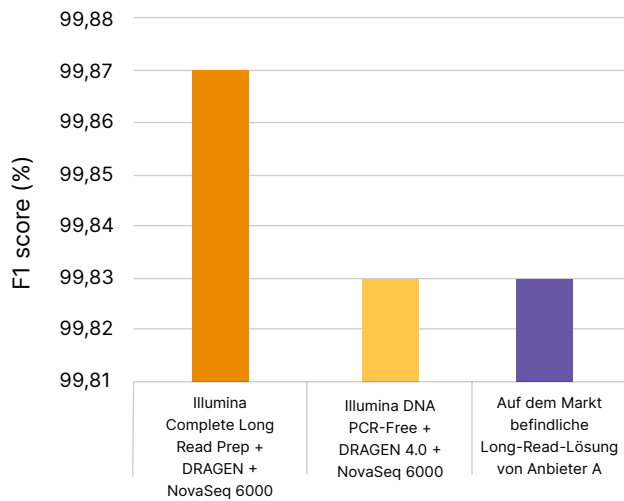


Abbildung 4: Ein neuer Maßstab für Genauigkeit: Illumina Complete Long Read Prep, Human zeichnet sich durch eine bislang unerreichte Genauigkeit beim Varianten-Calling aus. Diese wird als F1-Score (%) gemessen, der aus Präzisions- und Recall-Werten in Bezug auf die WGS gebildet wird.¹

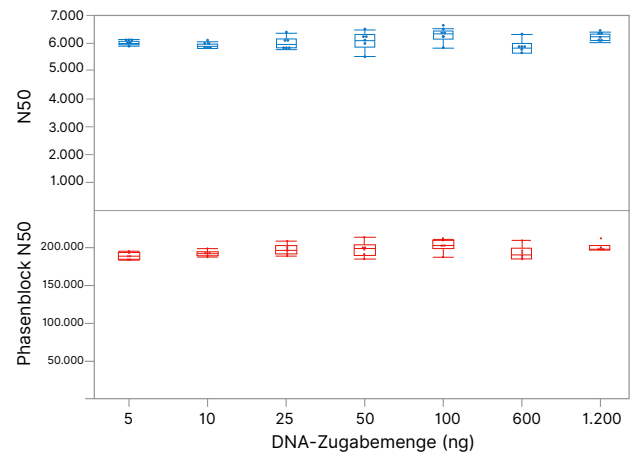


Abbildung 5: Hochwertige Leistung bei einem breiten Spektrum an DNA-Zugaben: Illumina Complete Long Read Prep, Human erzeugt hochwertige Bibliotheken, die aus DNA-Zugaben von 5 ng bis 1.200 ng (in dreifacher Ausführung) vorbereitet wurden, und liefert für N50 und Phasenblock N50 eine vergleichbare Datenqualität. N50 ist definiert als die Sequenzlänge des kürzesten Contigs bei 50 % der Gesamt-Assembly-Länge und kann auch als Maß für Phasenblocks verwendet werden.

Zuverlässige Leistung bei allen DNA-Zugabemengen

Die Leistung von Illumina Complete Long Read Prep, Human wurde anhand einer Reihe von DNA-Zugabemengen bewertet. Bibliotheken wurden aus 10, 25, 100, 600 und 1200 ng DNA in dreifacher Ausfertigung vorbereitet. Die Sequenzierungsergebnisse zeigten, dass Bibliotheken, die aus nur 5 ng Zugabe vorbereitet wurden, qualitativ hochwertige Metriken lieferten, einschließlich N50 und Phasierungsblock N50 (Abbildung 5). Illumina empfiehlt zwar eine Zugabe von 50 ng DNA, jedoch können auch niedrigere Zugabemengen verwendet werden.

Eignung für unterschiedlicher DNA-Qualitäten

Um beschädigte DNA zu simulieren, wurde hochwertige gDNA eine bzw. drei Sekunden lang einer Scherung mit einem Covaris-Ultraschallgerät unterzogen. Die DNA-Qualität wurde mit einem Fragment Analyzer beurteilt (Abbildung 6). Die Bibliotheken wurden mit gescherter gDNA-Zugabe und Illumina Complete Long Read Prep, Human vorbereitet. Die Sequenzierungsergebnisse zeigten, dass für eine Sekunde gescherter DNA qualitativ hochwertige Daten lieferte und ein genaues Varianten-Calling ermöglichte (Abbildung 7). Zugabe-gDNA, die wiederholt gefroren und aufgetaut wurde, lieferte ebenfalls qualitativ hochwertige Daten (Abbildung 8).

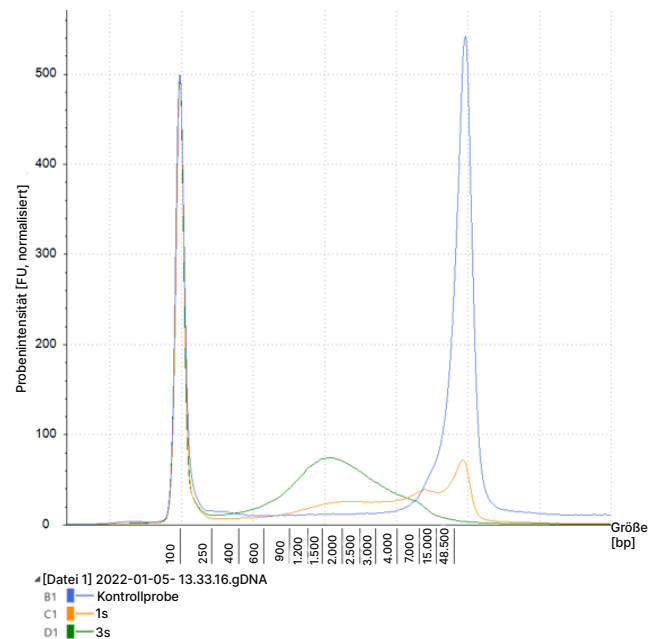


Abbildung 6: Qualität von gescherter DNA: Bei der Analyse von DNA mit einem Fragment Analyzer zeigt sich, dass die Qualität bei längerer Scherdauer abnimmt. DNA-Proben wurden eine oder drei Sekunden lang einer Covaris-Scherung unterzogen und die Qualität wurde mit einem Fragment Analyzer gemessen.

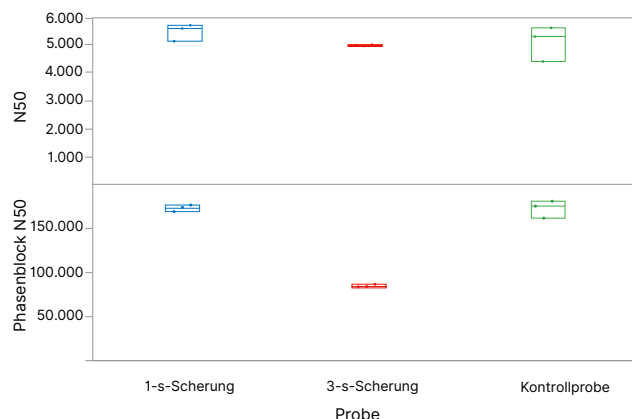


Abbildung 7: Sequenzierungsmetriken für gescherte DNA: Als Zugabe für Illumina Complete Long Read Prep, Human wurde für eine und drei Sekunden gescherte DNA verwendet. Die resultierenden Bibliotheken lieferten qualitativ hochwertige Long-Read-Sequenzierungsdaten (gemessen anhand von N50 und Phasenblock N50).

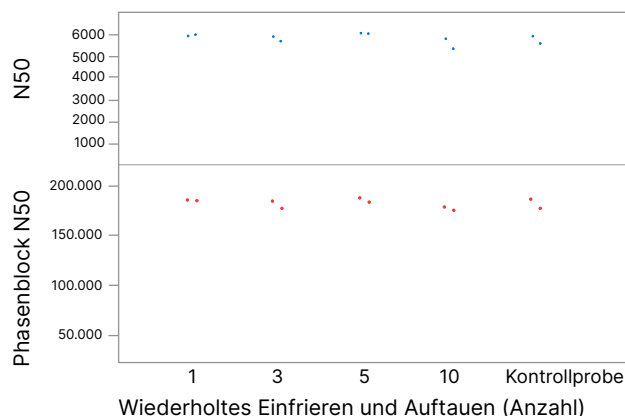


Abbildung 8: Sequenzierungsmetriken für gefrorene und aufgetaute DNA: Als Zugabe für Illumina Complete Long Read Prep, Human wurde DNA verwendet, die ein-, drei-, fünf- bzw. zehnmal gefroren und wieder aufgetaut wurde. Die resultierenden Bibliotheken lieferten qualitativ hochwertige Long-Read-Sequenzierungsdaten (gemessen anhand von N50 und Phasenblock N50).

Flexible Eignung für unterschiedliche Probentypen

Illumina Complete Long Read Prep, Human eignet sich für unterschiedliche Probentypen, einschließlich Blut und Speichel. Nach der Extraktion mit einem Standard-Kit bzw. einem Kit für hohes Molekulargewicht (HMW, High Molecular Weight) wurde die DNA-Qualität mit einem Fragment Analyzer bewertet (Abbildung 9). Die Sequenzierungsergebnisse zeigten, dass DNA aus Blut und Speichel qualitativ hochwertige Daten lieferte (Abbildung 10). Es gab keinen Unterschied in der Leistung zwischen mit einem Standard-Kit und mit einem Kit für hohes Molekulargewicht extrahierter DNA (Abbildung 10).

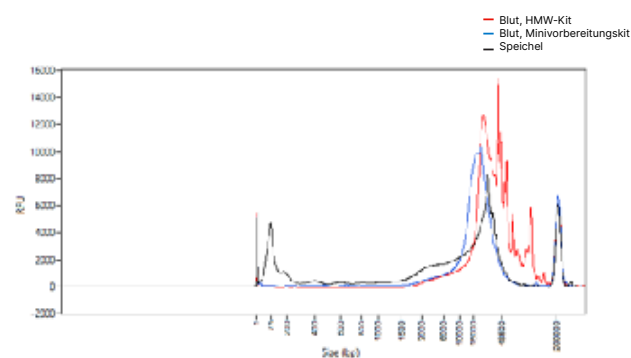


Abbildung 9: Qualität der DNA aus Blut und Speichel: Aus Blut (HMW, mini) oder Speichel extrahierte DNA wurde auf einem Fragment Analyzer analysiert.



Abbildung 10: Hochwertige Leistung bei verschiedenen Probentypen: Bibliotheken von Illumina Complete Long Read Prep, Human, die aus Blut (HMW, mini) und Speichel generiert wurden, liefern hochwertige Daten, gemessen anhand von N50 und Phasenblock N50.

Zusammenfassung

Die Long-Read-Sequenzierungsschemie liefert zusätzliche Informationen zu Regionen des Genoms, die mit Short-Read-NGS schwer zu analysieren sind. Illumina Complete Long Reads macht die Long-Read-Sequenzierung für die Genomforschung zugänglicher und einfacher. Illumina Complete Long Read Prep, Human, das erste auf dieser neuartigen Chemie basierende Produkt, bietet einen rationalisierten, vertrauten Workflow sowie Synergieeffekte mit der bewährten Illumina SBS-Chemie und der entsprechenden DRAGEN-Analyse. Diese hochinnovative Chemie ermöglicht die Short-Read- und Long-Read-Sequenzierung auf einem einzigen Gerät, wodurch Long-Read-NGS für Genomiklabore zugänglich wird. Illumina Complete Long Read Prep, Human zeigt eine zuverlässige Leistung bei verschiedenen DNA-Mengen und -Qualitäten aus unterschiedlichen Probenquellen. Damit steht eine hochskalierbare und genaue Lösung für die Sequenzierung des Humangenoms zur Verfügung.

Weitere Informationen

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

Quellen

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. Aufgerufen am 12. Januar 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. Veröffentlicht 2022. Aufgerufen am 12. Januar 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. Veröffentlicht 2022. Aufgerufen am 12. Januar 2023.
4. Oxford Nanopore Technologies. Ligation Sequencing Kit. store.nanoporetech.com/us/ligation-sequencing-kit110.html. Aufgerufen am 12. Januar 2023.
5. Pacific Biosciences. Low Yield Troubleshooting Guide. pacb.com/wp-content/uploads/Guide-Low-Yield-Troubleshooting.pdf. Veröffentlicht 2018. Aufgerufen am 12. Januar 2023.



1 800 8094566 (USA, gebührenfrei) | +1 858 2024566 (Tel. außerhalb der USA)
techsupport@illumina.com | www.illumina.com

© 2023 Illumina, Inc. Alle Rechte vorbehalten. Alle Marken sind Eigentum von Illumina, Inc. bzw. der jeweiligen Inhaber. Spezifische Informationen zu Marken finden Sie unter www.illumina.com/company/legal.html.
M-GL-01421 DEU v1.0