

Illumina Complete Long Read Prep, Humanを用いた 包括的な全ゲノム シーケンス

難易度が高く、マッピングが困難な
領域での優れた性能

illumina[®]

はじめに

次世代シーケンサー (NGS) により、サイエンティストがゲノムを解読し、より深く生物学を理解することが可能となっています。実績のあるイルミナ Sequence by Synthesis (SBS) ケミストリーと受賞歴のある DRAGEN™ 二次解析との組み合わせから、非常に高い精度で全ゲノムシーケンス (WGS) データが生み出されます。¹ それでもなお、非常にわずかなゲノム領域では、より長いリード長を活用することで、これら困難な領域のより正確な解析とマッピングを行うことができます。しかし、多くのロングリード対応シーケンスソリューションは複雑なワークフローであり、多量のDNAインプット量を必要とし、結果がばらつくことがあります。²⁻⁵

ILLUMINA Complete Long Readsなら、ゲノム研究においてロングリードシーケンスをより利用しやすく、効率化できるようになります。ILLUMINA Complete Long Read Prep, Humanは、この新しいロングリードケミストリーに基づいた初めての製品です。標準NGSワークフローを用いるWGS用の高性能アッセイが、NovaSeq™ 6000システムおよびNovaSeq Xシリーズ上で連続的ロングリードシーケンスを生成します (図1)。効率のかつ所要時間1日のライブラリー調製プロトコルは、ハイスループット研究に対応して簡単にスケールアップでき、わずか10 ngのDNAインプットしか必要とせず、特別な抽出、断片化、サイズ選択は不要です。

本テクニカルノートでは、幅広い範囲のインプットDNA量、DNA品質およびサンプルタイプを用いた場合も、ILLUMINA Complete Long Read Prep, Humanが生殖細胞バリエーション解析のための高精度かつ包括的なWGSにおいて、優れた性能を持つことを示します。

メソッド

標準サンプルのHG002/NA24385は、精製済みゲノムDNA (gDNA) としてGenome in a Bottleコンソーシアム (GIAB) から入手しました。これとは別に、血液と唾液サンプルを入手し、市販の標準的分子量用または高分子量用抽出キットのいずれかを用いて、メーカーの指示に従ってgDNAを抽出しました。

ライブラリー調製

ILLUMINA Complete Long Read Prep, Human (ILLUMINA、カタログ番号: 20089108) を用いてさまざまな量のインプットgDNAから目印付きのライブラリーを調製しました。概要としては、長い単一分子のDNA断片に、酵素的に固有のパターン (目印、ランドマーク) を付けます。目印のないライブラリーは、元の単一分子断片を表す長い連続リードを生成するデータ解析ステップで使用するために、ILLUMINA DNA PCR-Free Prep, Tagmentation (ILLUMINA、カタログ番号: 20041794) を用いて調製しました。

シーケンス

目印付きと目印なしのライブラリーは、150 bp × 2のリード長を用いてNovaSeq 6000システムで別々にシーケンスしました。

データ解析

目印付きのリードのアセンブリとレンダリングはBaseSpace™ Sequence HubのILLUMINA Complete Long Read WGSアプリを使って実施しました。データ解析と視覚化にはInteractive Genomics Viewer (IGV) を使用しました。



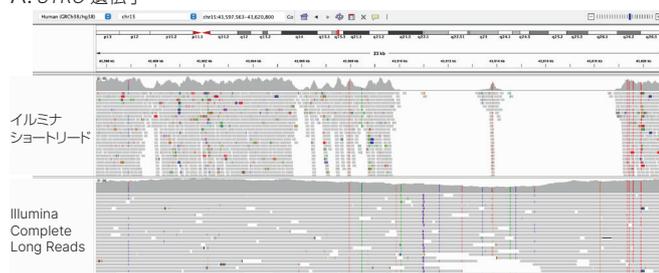
図1: ILLUMINA Complete Long Read Prep, Humanワークフロー: ILLUMINA Complete Long Read Prep, Humanは、DNA抽出、ライブラリー調製、シーケンスおよびデータ解析の4つの手順からなる効率的なワークフローを提供します。

a. 最少インプットは10 ngですが、50 ngのDNAを推奨します

困難な領域での高品質データ

Illumina Complete Long Read Prep, Humanはショートリードでは明確にマッピングできない領域のカバレッジを可能にします。これにより、*STRC* (図2A)、*SEMG1* (図2B) および *SULTIA1* (図2C) などのショートリードのみでのマッピングに問題を生じることが知られている疾患関連遺伝子と偽遺伝子の連続的なシーケンスが実現します。

A. *STRC* 遺伝子



B. *SEMG1* 遺伝子



C. *SULTIA1* 遺伝子

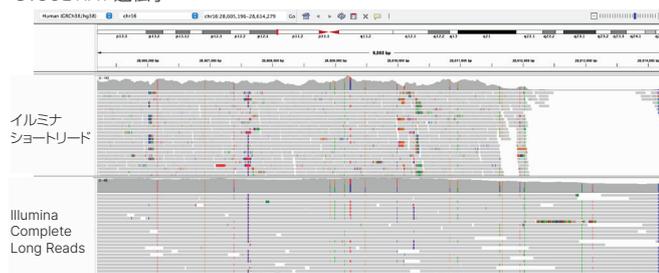


図2：包括的なカバレッジによって困難な領域で複数のバリエーションタイプを正確に検出することが可能：Illumina Complete Long Read Prep, Humanでは (A) *STRC*、(B) *SEMG1*、(C) *SULTIA1*などの困難なゲノム領域にある疾患関連遺伝子の正確なシーケンスが可能で。

さらに、Illumina Complete Long Read Prep, Humanは一般的な遺伝子サイズよりも非常に大きなフェーズブロックを提供します。これにより、高い配列相同性、数多くの偽遺伝子、高い変異性、および数多くのアリルを含む高度な多型性を示す領域である、ヒト白血球抗原 (HLA) 座位全体を均一にカバーできます。正確にシーケンスすることが困難であっても、Illumina Complete Long Read Prep, HumanはHLA領域全体の均一なカバレッジを生み出し、ゲノムフェージングを行い、母系と父系の染色体間のアリルを明確に視覚化し、識別できます (図3)。

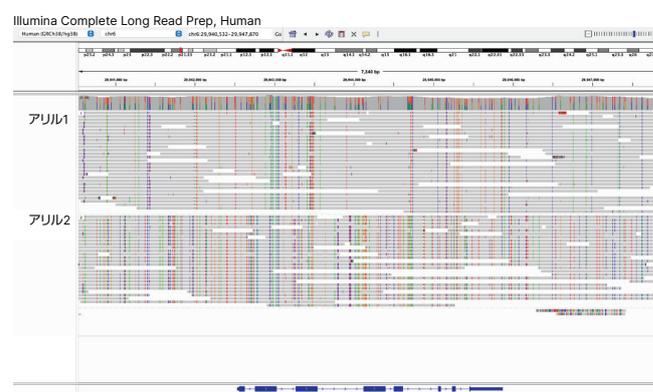


図3：均一なカバレッジにより明確なハプロタイプフェージングが実現：Illumina Complete Long Read Prep, Humanは、HLA領域の包括的なカバレッジを提供し、HLA領域内のアリルの明確かつ正確なフェージングが可能になります。

バリエーションコール精度をより包括的に評価するために、PrecisionFDA Truth Challenge V2でIllumina Long Read Prep, Humanについて他社のロングリードプラットフォームとイルミナショートリードとの比較を行いました。このチャレンジは、PrecisionFDA、GIABおよび米国国立標準技術研究所 (NIST) が主催しており、マッピング困難領域に重点を置き、共通のフレームで小さなバリエーションコールパイプラインの性能を評価するために始まりました。このデータセットを用いた場合、Illumina Complete Long Read Prep, Humanは非常に高い精度を示し、F1スコアが99.87%という他のロングリードプラットフォームをしのぐ結果を示します (図4)。

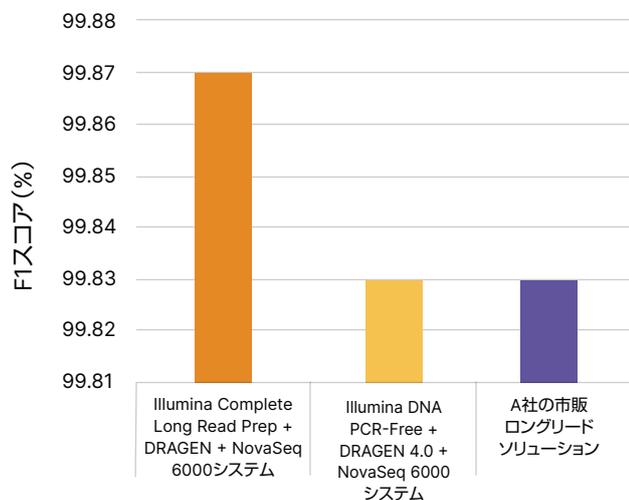


図4：精度における新基準：Illumina Complete Long Read Prep, Humanは、F1スコア (%) (WGSの精度とコール率を反映する値) によって判断されるバリエーションコールに対して非常に高い精度を生み出します。¹

さまざまなDNAインプット量での頑健な性能

Illumina Complete Long Read Prep, HumanをさまざまなDNAインプット量で評価しました。ライブラリーは10、25、100、600および1,200 ngのDNAからトリプレートで調製しました。シーケンス結果より、わずか5 ngのインプット量から調製したライブラリーでも、N50とフェージングブロックN50を含む高品質なメトリクスを示したことが明らかになりました (図5)。イルミナでは50 ngのDNAインプット量を推奨していますが、さらに少量のインプットも使用することができます。

DNA品質のばらつきにも対応

損傷したDNAをシミュレーションするために、Covaris超音波破碎装置を使用して1秒または3秒間、高品質gDNAを断片化しました。DNA品質はFragment Analyzerを用いて評価しました (図6)。ライブラリーは断片化したgDNAインプットとIllumina Complete Long Read Prep, Humanを用いて調製しました。シーケンス結果より、1秒間断片化されたDNAは高品質なデータを生成し、精度の高いバリエーションコールが可能でした (図7)。さらに、凍結融解サイクルを増やしたインプットgDNAを用いた場合も、高品質なデータを生成しました (図8)。

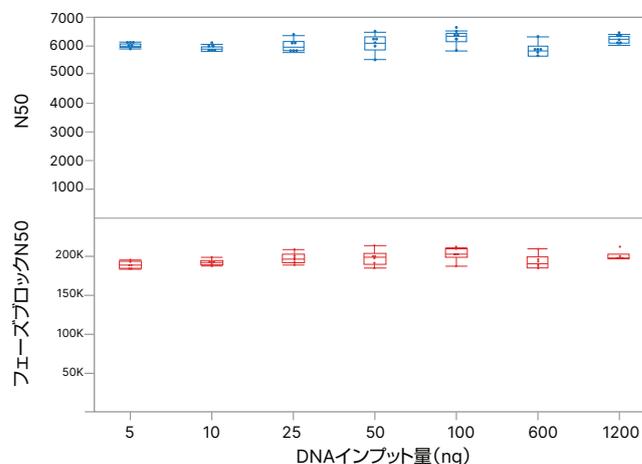


図5：幅広いDNAインプット量における高品質な性能：Illumina Complete Long Read Prep, Humanは5 ng~1,200 ngのDNAインプットから (トリプレートで) 調製した高品質ライブラリーを生成し、N50およびフェーズブロックN50について同様のデータ品質を生成します。N50はアセンブリの合計長の半分となる時の最も短いコンティグの配列長として定義され、フェーズブロックの基準としても使用することができます。

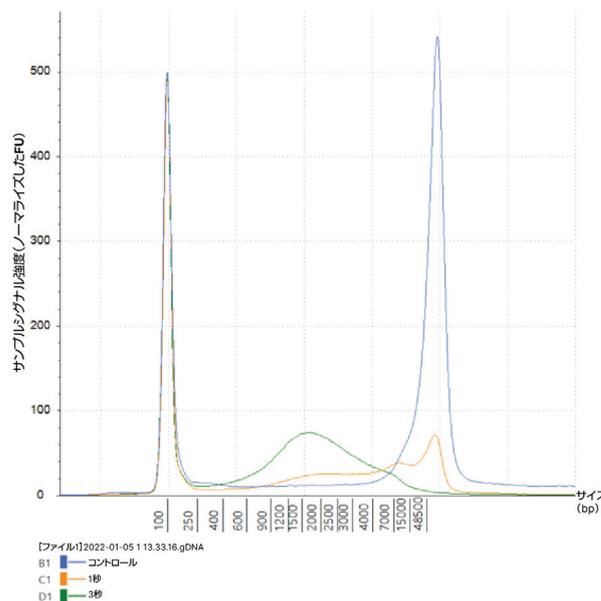


図6：断片化DNA品質：Fragment AnalyzerでのDNA解析より、断片化時間を増やすと品質が低下することが示されます。DNAサンプルを1秒間または3秒間、Covarisによって断片化し、品質はFragment Analyzerを用いて測定しました。

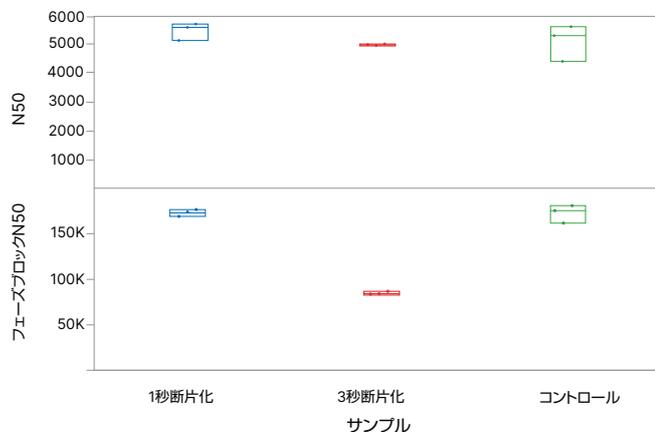


図7: 断片化DNAシーケンスメトリクス: 1秒間または3秒間断片化したDNAをIllumina Complete Long Read Prep, Humanのインプットに使用しました。得られたライブラリーは、N50とフェーズブロックN50の判定で分かる通り、高品質なロングリードシーケンスデータを生成しました。

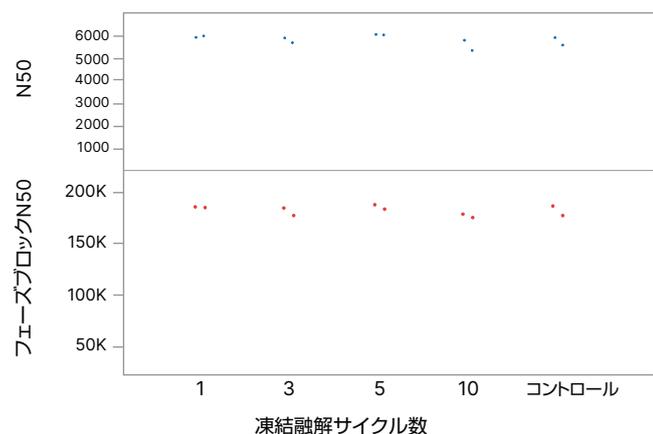


図8: 凍結融解されたDNAのシーケンスメトリクス: 凍結融解サイクルを1、3、5、10回行ったDNAをIllumina Complete Long Read Prep, Humanのインプットに使用しました。得られたライブラリーは、N50とフェーズブロックN50の判定の通り、高品質なロングリードシーケンスデータを生成しました。

さまざまなサンプルタイプに柔軟に対応

Illumina Complete Long Read Prep, Humanは血液および唾液を含むさまざまな種類のサンプルに対応します。標準的な分子量または高分子量 (HMW) 用キットのいずれかを用いてDNAを抽出した後、DNA品質はFragment Analyzerを用いて評価しました (図9)。シーケンス結果より、血液および唾液からのDNAは高品質データを生成したことが示されました (図10)。注目すべきは、標準的な分子量用キットとHMW用キットで抽出したDNA間に性能の違いは認められなかったことです (図10)。

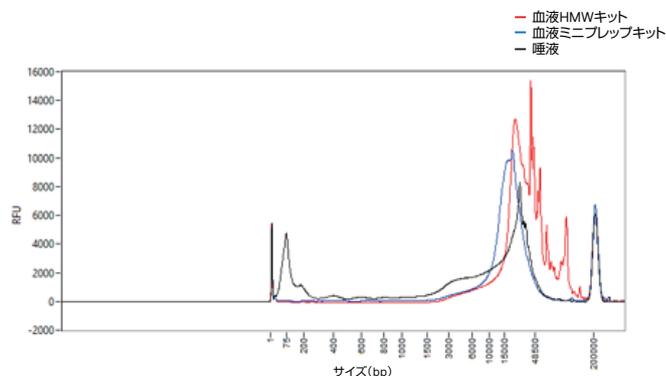


図9: 血液と唾液のDNA品質: 血液 (HMW、ミニ) または唾液から抽出したDNAはFragment Analyzerで解析しました。

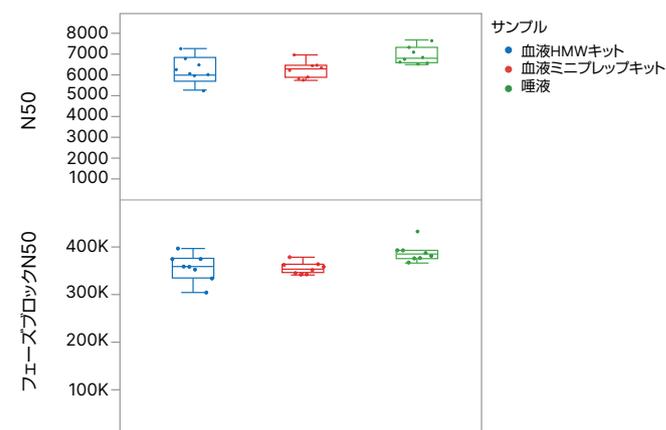


図10: 異なる種類のサンプルを用いた高品質な性能: 血液 (HMW、ミニ) および唾液から生成したIllumina Complete Long Read Prep, Humanライブラリーは、N50とフェーズブロックN50の判定に見られる通り、高品質なデータを生成しました。

まとめ

ロングリードシーケンスケミストリーは、ショートリードNGSで解明することが困難なゲノム領域に対してさらなる情報を提供します。Illumina Complete Long Readsを用いることで、サイエンティストはゲノム研究においてロングリードシーケンスにもアクセスしやすくなり、効率化できるようになります。新しいケミストリーに基づいて構築された初の製品である、Illumina Complete Long Read Prep, Humanは、効率的で馴染みのあるワークフローと、実績あるイルミナSBSケミストリーとDRAGEN解析による相乗効果を提供します。この非常に革新的なケミストリーによって、1台の装置上でショートリードとロングリードシーケンスを実施することが可能となり、ゲノム研究においてロングリードNGSにアクセスしやすくなります。Illumina Complete Long Read Prep, Humanは異なるサンプルソースからのさまざまな量と品質のDNAに頑健な性能を示します。その結果は高い拡張性があり、精度の高いヒトWGSソリューションです。

詳細はこちら

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

参考文献

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

イルミナ株式会社

〒108-0014 東京都港区芝 5-36-7 三田ベルジュビル 22階
Tel (03) 4578-2800 Fax (03) 4578-2810
jp.illumina.com

 www.facebook.com/illuminakk

販売店

本製品の使用目的は研究に限定されます。診断での使用はできません。 販売条件： jp.illumina.com/tc

© 2023 Illumina, Inc. All rights reserved.

すべての商標および登録商標は、illumina, inc. または各所有者に帰属します。
商標および登録商標の詳細は jp.illumina.com/company/legal.html をご覧ください。
予告なしに仕様および希望販売価格を変更する場合があります。

