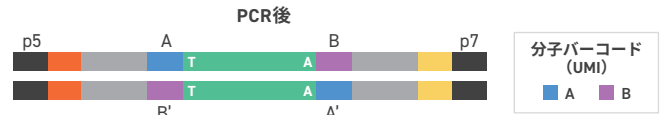


UMIのデータ処理およびコンセンサスに関するガイダンス

Twist UMIアダプターシステムで調製したライブラリをご利用のお客様へ

本ガイドには、Illumina社シーケンシングシステム用にTwist UMIアダプターシステムで調製したサンプルのデータ処理、およびUMIファミリーからのコンセンサスリードの呼び出しに関する推奨事項を記載しています。Twist UMIアダプターシステムで調製したライブラリには、5塩基対の分子バーコード、10塩基対のユニークデュアルインデックス、p5およびp7標準アダプターが含まれています。



以下に示すワークフローは、Fulclum Genomics社のオープンソースのコマンドラインツールキットであるfgbioを活用しています。BWAおよびPicardツールのソフトウェアパッケージも必要です。

まず、FASTQファイルをunaligned BAMファイルに変換してリード末端のUMI配列を抽出し、BAMタグとして保存します。次に、リファレンスゲノムとアライメントするため、unaligned BAMファイルをFASTQファイルに再度変換します。得られたaligned BAMファイルを、UMIタグを含むunaligned BAMファイルと統合し、その後、fgbioパッケージで処理して二本鎖のコンセンサスリードを作成します。

1. **GroupReadsByUmi**を使用して、由来する元の分子が同じであるリードをグループ化します。これにより、マッピング位置（染色体、開始位置と終了位置）およびUMI配列ごとにリードがグループ化されます。
2. **CallDuplexConsensusReads**を使用し、順鎖と逆鎖のコンセンサス読み取りのエビデンスを組み合わせることにより、二本鎖のコンセンサスリードを作成します。

これで、最終コンセンサスBAMファイルを三次データ解析に使用できるようになります。

推奨ワークフロー

- 1 FASTQファイルをunaligned BAMファイルに変換
- 2 UMIをunaligned BAMファイルからBAMタグに抽出
- 3 unaligned BAMファイルをFASTQファイルに変換
- 4 UMIおよびアダプターをトリミングしたリードとリファレンスゲノムとのアライメント
- 5 aligned BAMファイルとunaligned BAMファイルの統合
- 6 MarkDuplicatesでduplicate readを同定
- 7 picardメトリクスを計算
- 8 UMIごとにリードをグループ化
- 9 二本鎖コンセンサスリードの呼び出し
- 10 ヒトゲノムに対する二本鎖コンセンサスリードのリアライメント
- 11 二本鎖コンセンサスリードのpicardメトリクスを計算

Picard: FastqToSam

fgbio: ExtractUmisFromBam (5M2ST+)

Picard: SamToFastq

BWA: bwa mem

Picard: MergeBamAlignment

Picard: MarkDuplicates

Picard: CollectHsMetrics

fgbio: GroupReadsByUmi

fgbio: CallDuplexConsensusReads

Picard: SamToFastq

BWA: bwa mem

Picard: CollectHsMetrics

ExtractUmisFromBamツールにおいて、Twist UMIシステムで調製したライブラリのリード構造は、151ペアエンド (PE) のシーケンシングでは5M2S144Tとなります。ここで、M = UMI配列 (5 bp)、S = シャドー／スキップ塩基 (2 bp)、T = 標的／DNA塩基 (144 bp) を意味します。

ご質問にお答えします

support@twistbioscience.comまでご連絡ください。または、twistbioscience.comでカスタマーサポート担当者がチャットで回答します。

関連製品

- 105094 : Twist UMI Adapters - TruSeq Compatible, 96サンプル
- 105040 : Twist UMI Adapter System - TruSeq Compatible, 16サンプル
- 105041 : Twist UMI Adapter System - TruSeq Compatible, 96サンプル プレートA
- 105042 : Twist UMI Adapter System - TruSeq Compatible, 96サンプル プレートB
- 105043 : Twist UMI Adapter System - TruSeq Compatible, 96サンプル プレートC
- 105044 : Twist UMI Adapter System - TruSeq Compatible, 96サンプル プレートD