
 猫も杓子も

Geneious Prime でシーケンス解析

第 22 回 De novo アセンブリ (その 2: 前処理の実例)

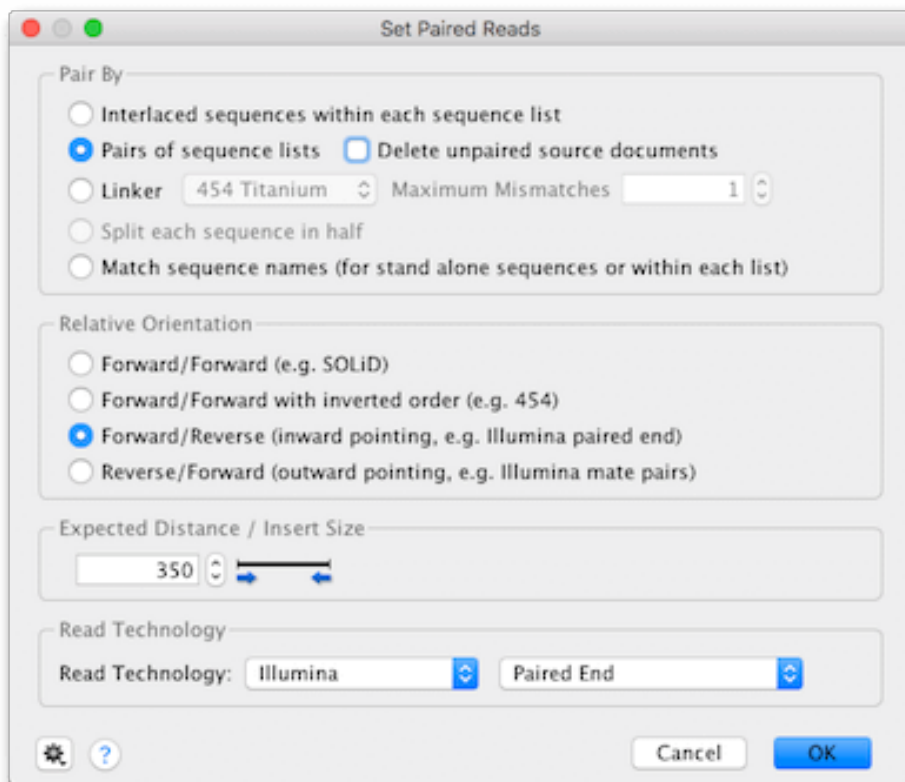


前回:[第 21 回 De novo アセンブリ \(その 1: 前処理の概要\)](#)より、NGS リードを処理し、de novo アセンブルする一般的な流れについてご紹介しています。2 回目となる今回は、NGS リードのペアリングとトリミングの前処理ステップの実例についてです。

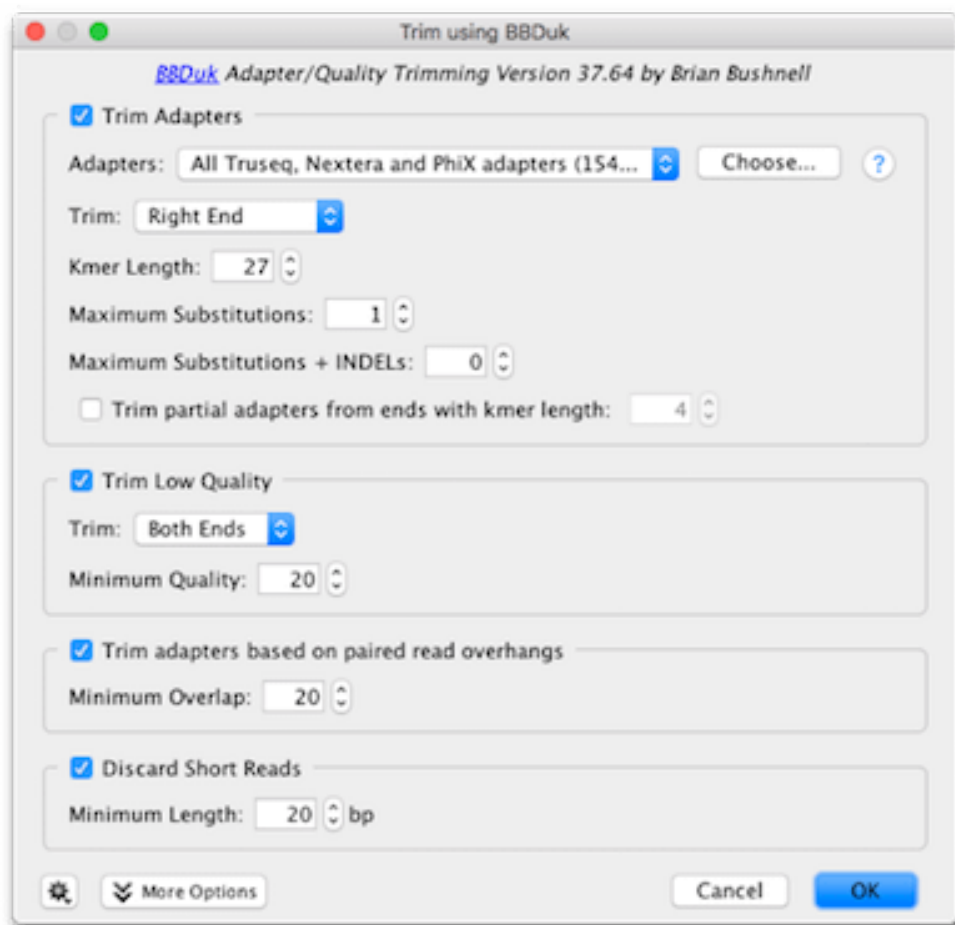
ご紹介している内容はチュートリアルとしてまとめられていますので、ご自身で試してみたい方は、[こちらからダウンロード](#)し、ZIP ファイルを解凍せずにそのまま Geneious Prime にドラッグ&ドロップしてインストールしてください。

NGS リードのペアリングは 2 つのリストを選択し、**Sequence メニュー → Set Paired Reads** に進みます。チュートリアルデータの場合は、ペアリードとなる SRR513053 subset F と SRR513053 subset R の 2 つのリストが用意されています。

Pairs of sequence lists を選択して、**Forward/Reverse (inward pointing)**(イルミナペアエンドデータの場合)で **Expected Distance** を 350 と設定し、**Read Technology** が **Illumina Paired End** に設定されていることを確認して、**OK** をクリックすると、**SRR513053 subset** という新しいペアリストが作成されます。

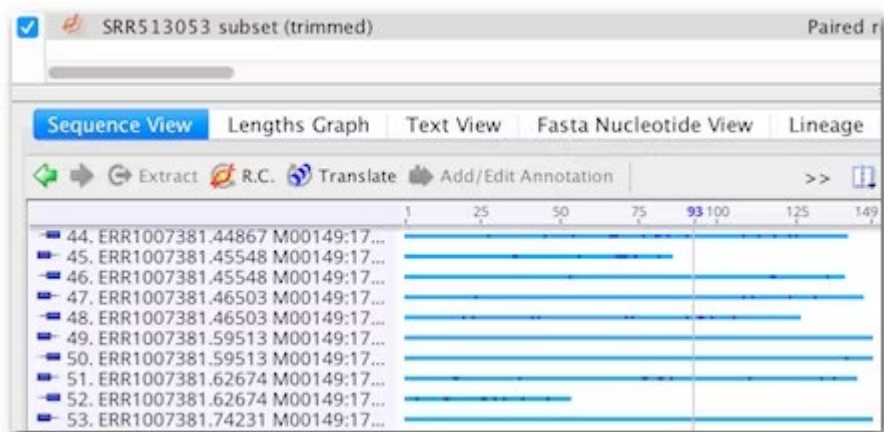


次に、作成された新しいペアリストを選択し、**Annotate & Predict** メニュー → **Trim using BBDuk** に進みます。Trim Adapters にチェックを入れて下図のように設定します。次に Trim Low Quality にチェックを入れ、Minimum Quality を 20 に設定します(Nanopore リードの場合は 6 が推奨値)。Trim adapters based on paired read overhangs では、Minimum Overlap を 20 に設定します。最後に Discard Short reads にチェックを入れ、Minimum Length を 20 に設定します。



この設定により、アダプター配列とクオリティの低い配列、極端に短いリードが除去された、**SRR513053 subset (trimmed)**という新しいペアリードリストが作成されます。

トリムされたリストを選択し、リストをスクロールすると、多くのリードが元の長さの 149 ヌクレオチドより短くトリムされていることがわかります。BBDuk のコマンドライン出力(トリミング中に削除された塩基数やリードのリストなど)は、ビューアの上の **Info** タブをクリックすることで確認できます。



今回はトリミングされたペアリードを Geneious de novo アセンブラを使ってアセンブルする実例をご紹介します予定です。

Geneious 製品概要・フリートライアルリクエストについては[こちら](#)

『Geneious Prime で猫も杓子もシーケンス解析』連載 第 22 回！！
皆様の解析のサポートに、ネコの手を。過去の記事は[こちらでチェック！](#)



いつもご愛読ありがとうございます