猫も杓子も

Geneious Prime でシーケンス解析

第 29 回 リファレンスマッピングと SNP 検出 (その 4)



NGS リードをリファレンス配列にマッピングし、SNP を検出するための一般的なワークフローについてご紹介しています。

[第 26 回 リファレンスマッピングと SNP 検出 \(その 1\)](#)

[第 27 回 リファレンスマッピングと SNP 検出 \(その 2\)](#)

[第 28 回 リファレンスマッピングと SNP 検出 \(その 3\)](#)

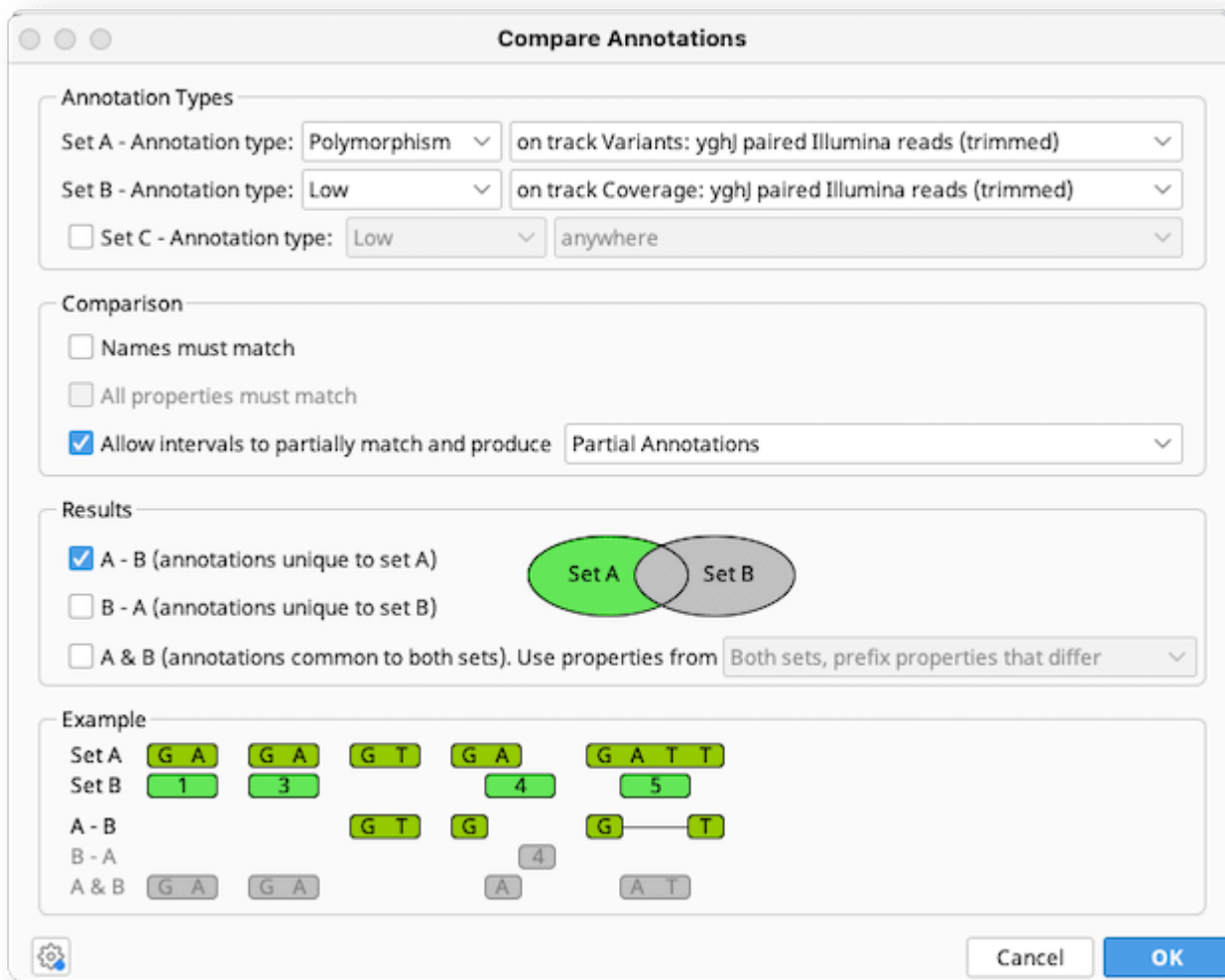
4 回目となる今回は、アノテーションの比較機能を使用して、より確度の高い SNP を検出する手法についてです。

Geneious Prime では **Compare Annotations** 機能を使用することで、別のアノテーショントラックやアノテーションタイプとのオーバーラップに基づいてアノテーションを比較し、フィルタリングすることができます。

前々回の最後で、平均カバレッジから 2 標準偏差以下の領域を特定するために、Coverage というアノテーショントラックを作成しています。これと Variants のアノテーショントラックを Compare Annotations 機能で比較することで、カバレッジの低い領域で検出されている SNP のアノテーションを除外します。

リファレンス配列を選択すると、前々回に作成した Coverage のトラックと、前回作成した Variants のトラックが表示されますので、**Annotate & Predict** → **Compare Annotations** に進みます。**Annotation Types** パネルで比較したいアノテーションを指定できます。チュートリアルデータの場合、**Set A** に”Polymorphism” ”on track Variants: yghJ paired Illumina reads”を、**Set B** に”Low” ”on track Coverage: yghJ paired Illumina reads”を選択します。Comparison パネルでは、**Names must match** のチェックを外してください。**Allow intervals to partially match...**にチェックを入れると、低カバレッジアノテーションの一部にしか含まれないインデルのような多型が検出されます。

Results パネルで **A-B** にチェックすることで、低カバレッジアノテーション(B)と重ならない多型アノテーション(A)が検出されます。



OK をクリックすると、リファレンス配列上に "Variants: yghJ paired Illumina reads - Coverage: yghJ paired Illumina reads" という 3 つ目のアノテーショントラックが表示されます。このトラックは、元の Variants トラックよりも多型のアノテーションが少なくなっており、低カバレッジとマーキングされた配列の領域までスクロールすると、この領域では元の Variants トラックの多型アノテーションが新しいトラックから除外されていることがわかります。



Save をクリックして、リファレンス配列に新しいトラックを保存します。

実は……本連載記事では機能のご紹介のため、Compare Annotations を使用してカバレッジの低い領域の SNP を除外していますが、このステップを実行しなくても、Variant Finder 機能での SNP 検出時に、カバレッジの低い SNP を自動的にフィルタリングするように設定することが可能です。そのためには、**Find Variations/SNPs** の **Minimum Coverage** ボックスにチェックして、必要な最小カバレッジを入力してください。

次回は番外編として、サンプルの SNP を反映した仮想ゲノム配列を作成し、SNP に基づく系統樹を再構築する方法についてご紹介する予定です。

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』過去の記事は[こちらでチェック!](#)

TDB News 3. 2024
トミーデジタルバイオロジー株式会社
Phone 03-6240-0843 Fax 03-6240-0461