猫も杓子も

## Geneious Prime でシーケンス解析

### 第 30 回 リファレンスマッピングと SNP 検出 (おまけ)



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 \(その 3\)](#)

[第 29 回 リファレンスマッピングと SNP 検出 \(その 4\)](#)

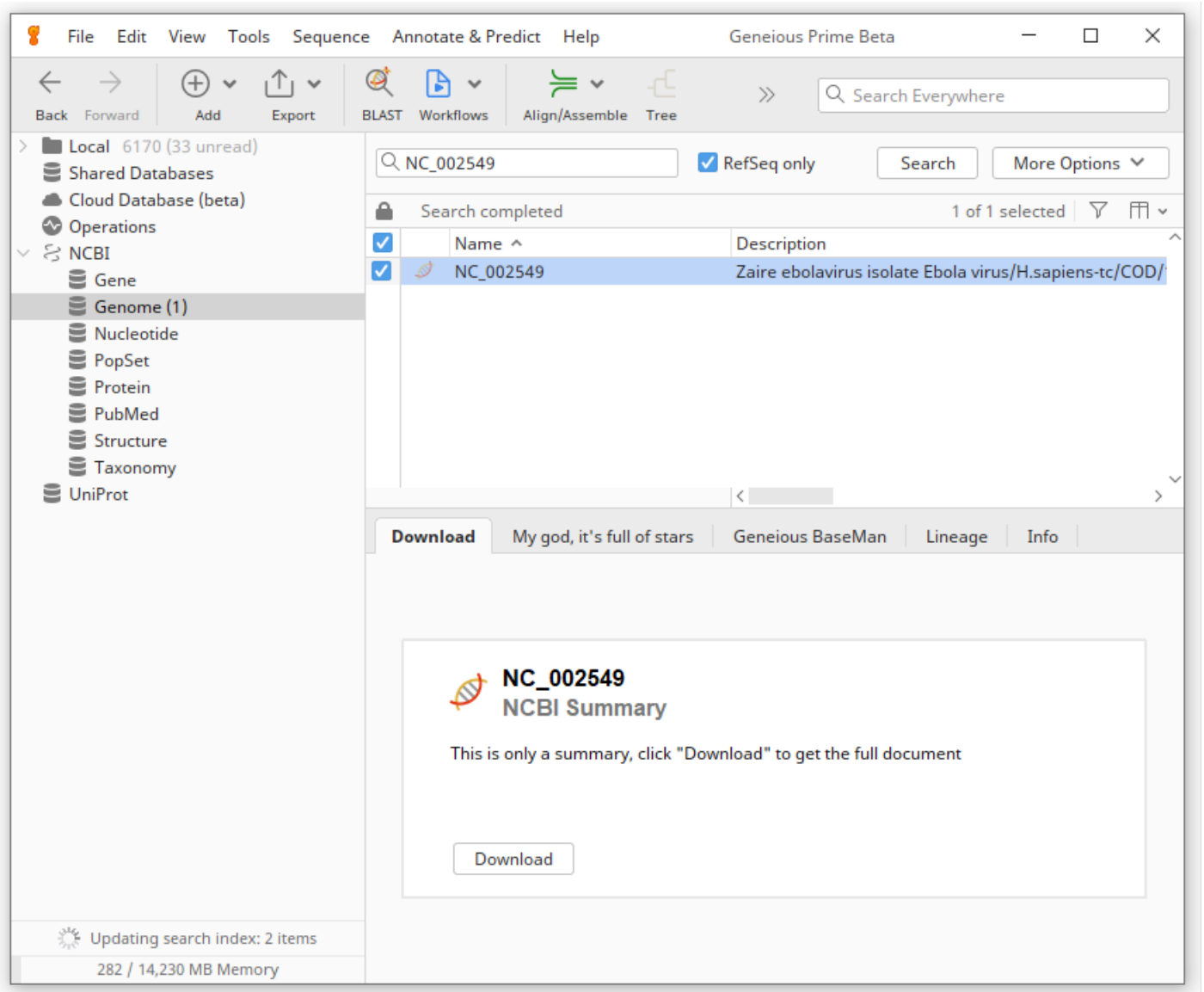
今回はその番外編として、サンプルの SNP を反映した仮想ゲノム配列を作成し、SNP に基づく系統樹を再構築する方法についてご説明いたします。

この説明では、2014 年西アフリカでのエボラ発生時に分離されたウイルスから、米国陸軍医療感染症研究所によって取得された NGS リードデータセットと、公開済みエボラゲノム(NC\_002549)をリファレンスとして使用します。

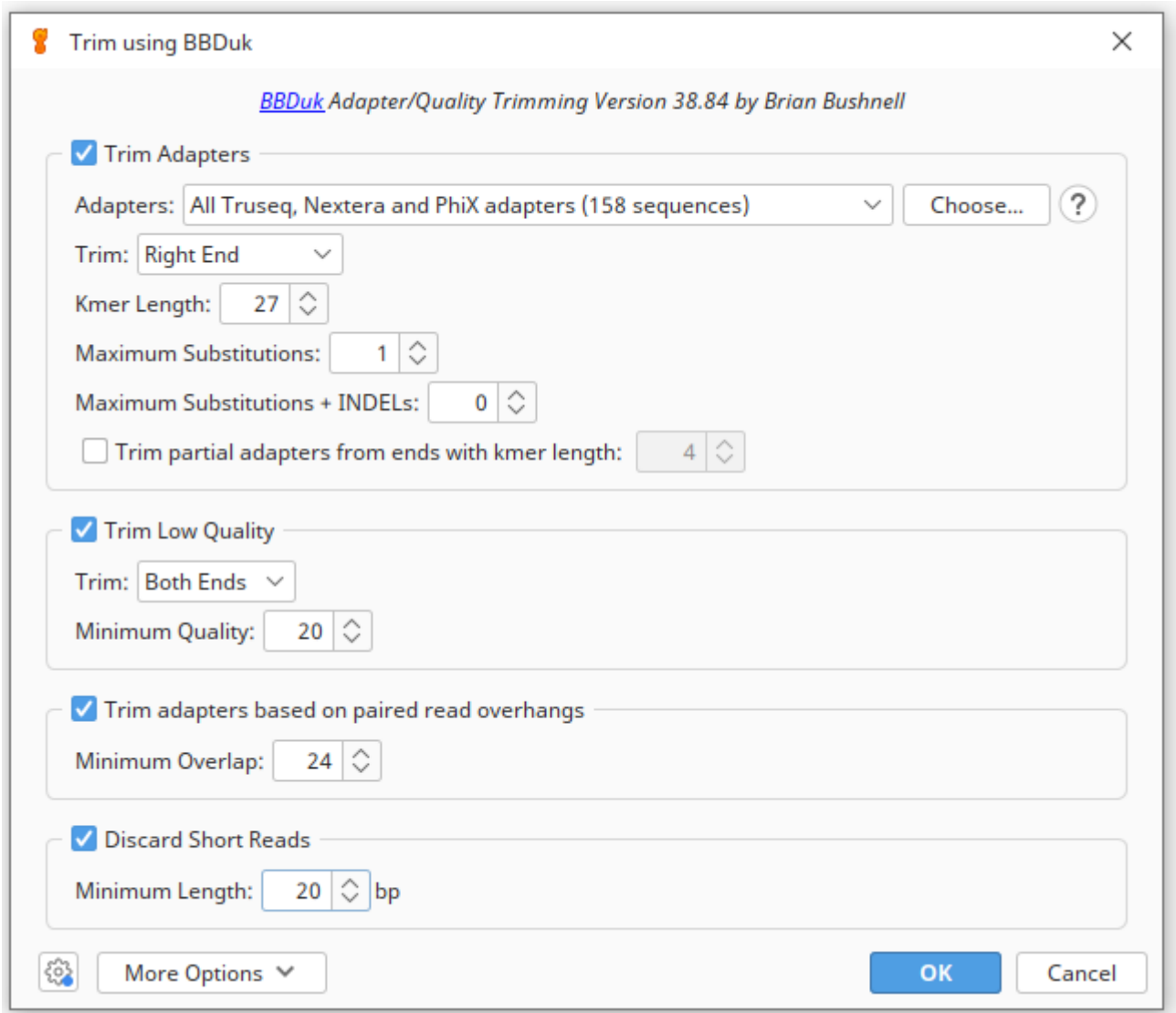
以下のリンクから、各分離株の NGS リードデータセットをそれぞれダウンロードし、任意のフォルダにインポートします。**Expected Distance** は 300 bp のペアリードとして設定してください。

LIBR0284	<a href="http://www.ebi.ac.uk/ena/data/view/SRR3152967">http://www.ebi.ac.uk/ena/data/view/SRR3152967</a>
LIBR0380	<a href="http://www.ebi.ac.uk/ena/data/view/SRR3152978">http://www.ebi.ac.uk/ena/data/view/SRR3152978</a>
LIBR10279	<a href="http://www.ebi.ac.uk/ena/data/view/SRR3152980">http://www.ebi.ac.uk/ena/data/view/SRR3152980</a>
LIBR11079	<a href="http://www.ebi.ac.uk/ena/data/view/SRR3152981">http://www.ebi.ac.uk/ena/data/view/SRR3152981</a>
LIBR11177	<a href="http://www.ebi.ac.uk/ena/data/view/SRR3152982">http://www.ebi.ac.uk/ena/data/view/SRR3152982</a>
LIBR0622	<a href="http://www.ebi.ac.uk/ena/data/view/SRR3152984">http://www.ebi.ac.uk/ena/data/view/SRR3152984</a>
LIBR0707	<a href="http://www.ebi.ac.uk/ena/data/view/SRR3152985">http://www.ebi.ac.uk/ena/data/view/SRR3152985</a>
LIBR010180	<a href="http://www.ebi.ac.uk/ena/data/view/SRR3153068">http://www.ebi.ac.uk/ena/data/view/SRR3153068</a>

リファレンスとなるエボラゲノム(NC\_002549)を、Geneious Sources パネルの NCBI 検索ツールを使用して検索、ダウンロードして、先程のフォルダに移動します。ダウンロードの手順は、NCBI フォルダ → **Genome** フォルダ → **Search** ボックスに NC\_002549 を入力 → **Search** ボタンをクリック → 検索結果の NC\_002549 を選択して **Download** ボタンをクリック、となります。




インポートしたペアリードリストを全て選択し、Annotate & Predict メニュー → Trim using BBDuk より、下図の設定でトリムします。



トリムしたNGSデータセットリストとリファレンスを全て選択し、**Align/Assemble**ボタン → **Map to Reference** でマッピングします。**Assemble each sequence list separately** オプションにチェックします。その他の設定はデフォルトのままです。

マッピングされたコンティグを全て選択し、**Annotate & Predict** メニュー → **Find Variants/SNPs** より、下図の設定で SNP を検出します。**Analyze effect of polymorphisms on Translation** オプションをオフにし、**Use separate annotations for each variant at a position** オプションをオンにしてください。

 Find Variations/SNPs
✕

**Find Variants**

Minimum Coverage:

Minimum Variant Frequency:

Maximum Variant P-value:  $10^{-6}$  (0.0001% to see variant by chance) ?

Minimum Strand-Bias P-value:  $10^{-5}$  when exceeding  % bias ?

Find Variants

In selected region only

**Analyze Effects on Translations**

Analyze effect of variants on translations

Default Genetic Code:

Note: Variations found within overlapping CDS annotations will create separate annotations for each CDS

Calculate Variant P-values

Assumed quality of bases without quality:  (99.0% correct)

P-value calculation method:

Homopolymer quality reduction for 454 / Ion Torrent:  % ?

**Advanced**

Only Find SNVs

Merge adjacent variations

Ignore reference sequence (only find variations within the sample)

Exclude paired reads over  % from their expected distance

Use separate annotations for each variant at a position

Record names of all contributing sequences for each variant


Ignore reads mapped to multiple locations

Don't find variations in annotation types:

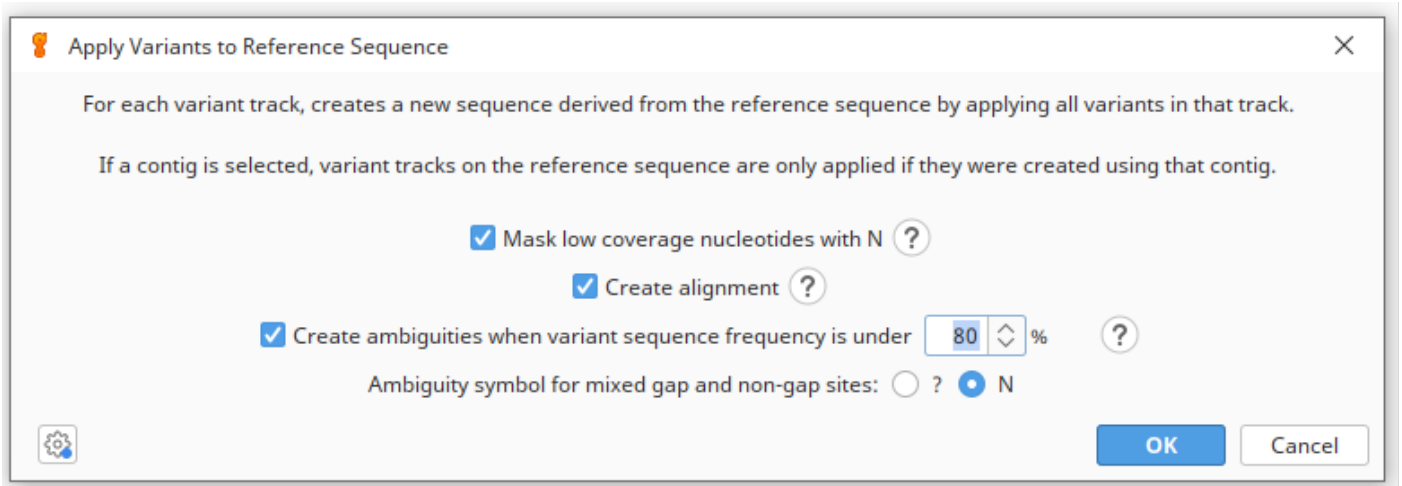
Only find variations in annotation types:

Also find variations within  bases of those types

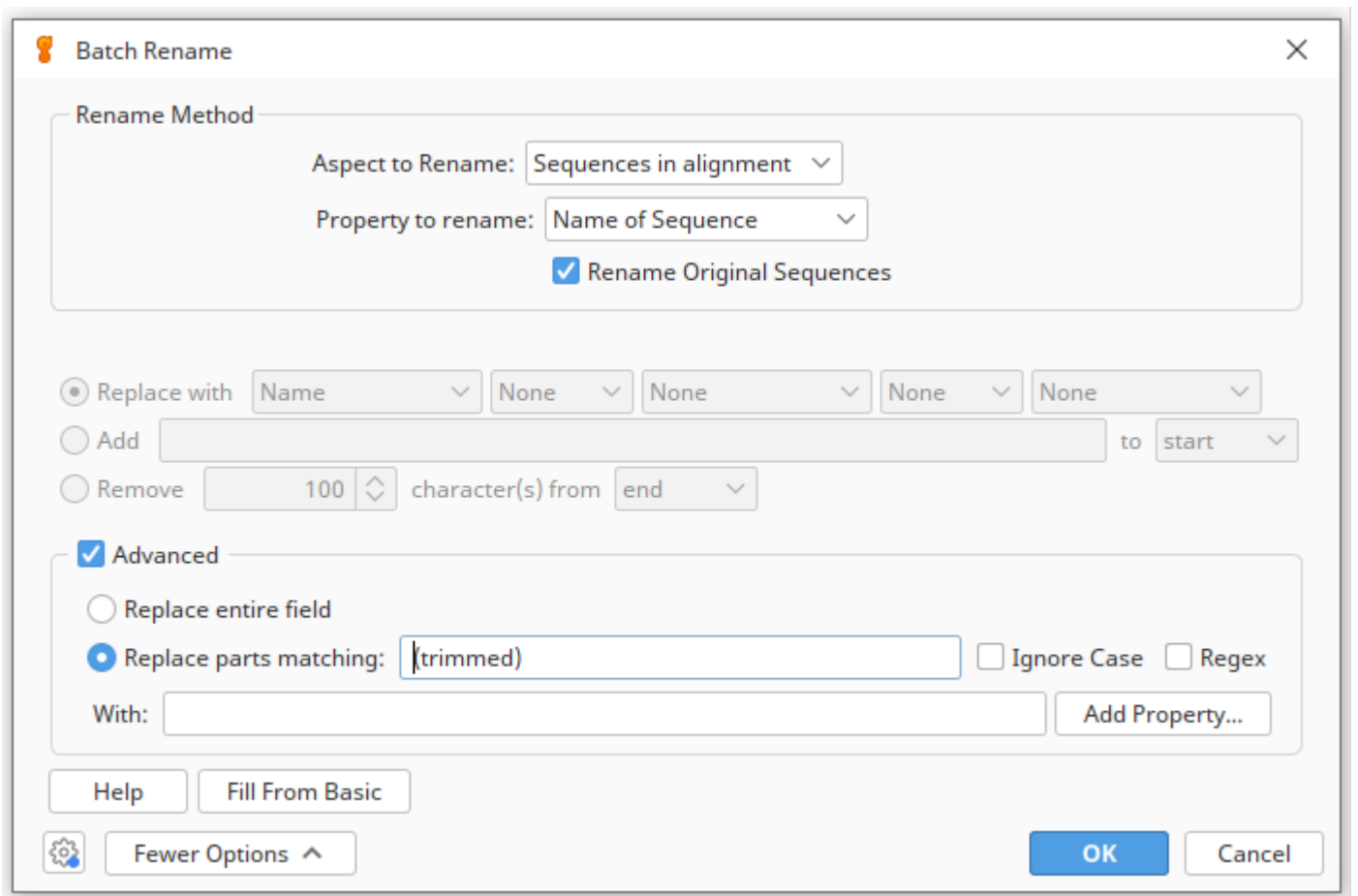
CDS Properties to Copy:

 Fewer Options ^

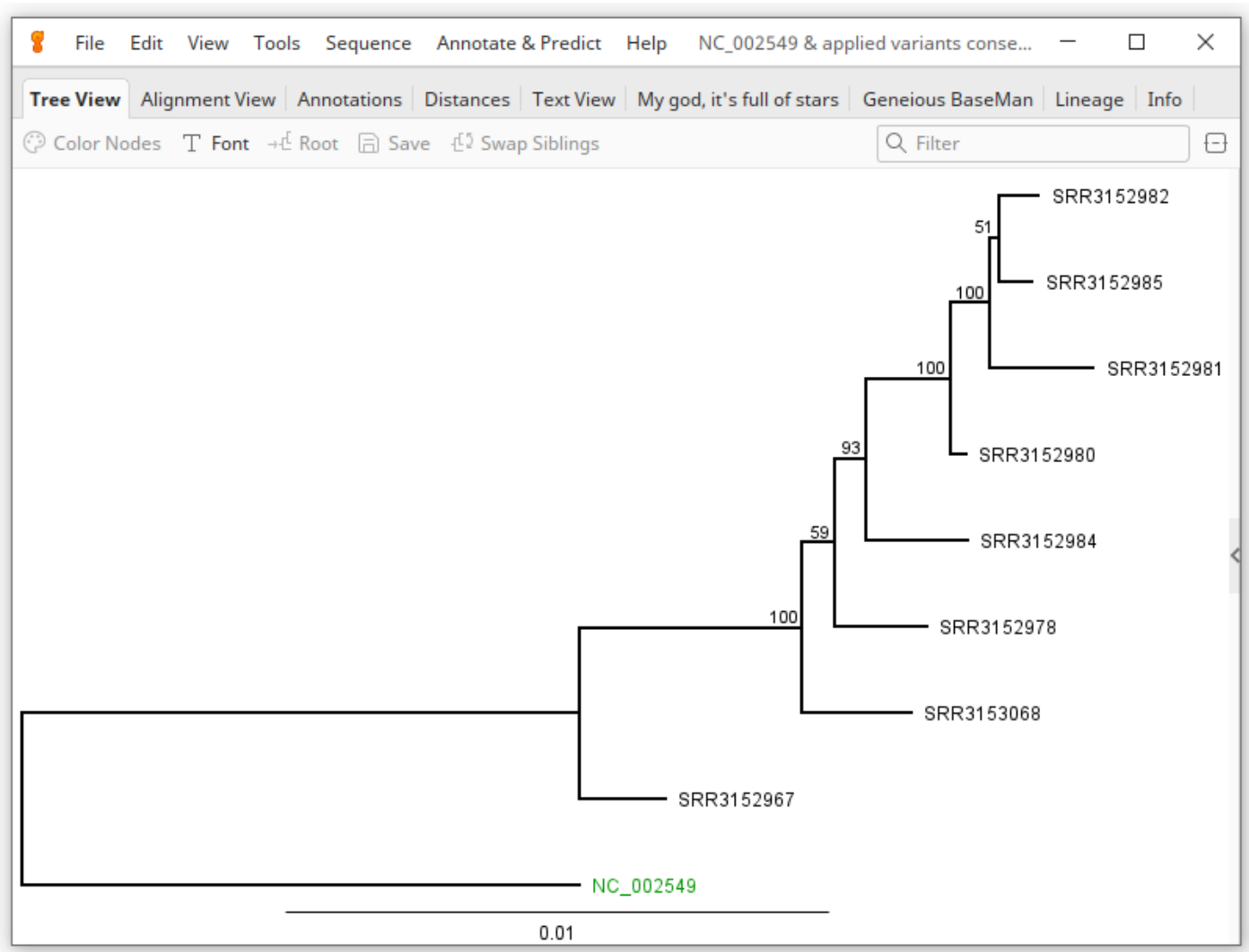
検出した SNP をトラックとして保存したリファレンス配列を選択し、**Workflow** ボタン → **Apply Variants to Reference Sequence** で、リファレンス配列に基づき、検出した SNP を反映させた仮想ゲノム配列を作成します。



必須の操作ではなく、任意のオプションになりますが、シーケンス名のリネームを行います。BBDukでのトリミングステップで、シーケンス名に「trimmed~」という文言が追加されており、それが仮想ゲノム配列まで持ち越されていますので、作成された仮想ゲノム配列アラインメントを選択し、**Edit**メニュー→ **Batch Rename** で元のシーケンス名がツリーに表示されるように変更します。**Advanced** 設定で、**Replace parts matching** に、半角スペース+(trimmed)を入力します。



作成された仮想ゲノム配列アラインメントから、任意のアルゴリズムで SNP に基づく系統樹を再構築することができます。



次回からはシークエンス間のアノテーションの転送についてシリーズで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 で猫も杓子もシークエンス解析』過去の記事は[こちらでチェック!](#)