

Geneious Prime でシーケンス解析



第 6 回 クロマトグラムのアSEMBL(その 1)

サンガーシーケンサーの名機 310/3100xl のメーカーサポートもついに終了し、修理や消耗品についても在庫を残すのみということになっているようですが、後継機への切り替えのご準備はお済みでしょうか。今回からは数回にわたり、クロマトグラムを編集し、アライン/アSEMBLする手法をご紹介します予定ですが、これらは系統樹の再構築やヌクレオチドの多様性の計算など、下流の解析のために必要な作業となりますが、まず今回は、典型的なサンガーの生リードからクオリティの低い配列のトリミング、アラインメントやアSEMBリからシーケンスを編集する方法についてご説明します。

例として使用するアオガラ(*C. caeruleus* と *C. teneriffae*)のミトコンドリア DNA 制御領域の 34 のシーケンスと、アウトグループとしてシジュウカラ(*Parus major*)のシーケンスは、Geneious の Local フォルダ → Tutorials フォルダ → Assembling_Chromatograms フォルダの Cyanistes CR sequences ドキュメントに含まれています。

Code	Species	Origin
CEH	<i>C. teneriffae</i>	Canary Islands - El Hierro
CFU	<i>C. teneriffae</i>	Canary Islands - Fuerteventura
CGC	<i>C. teneriffae</i>	Canary Islands - Gran Canaria
CLG	<i>C. teneriffae</i>	Canary Islands - La Gomera
CLP	<i>C. teneriffae</i>	Canary Islands - La Palma
CLA	<i>C. teneriffae</i>	Canary Islands - Lanzarote
CTE	<i>C. teneriffae</i>	Canary Islands - Tenerife
MCE	<i>C. teneriffae</i>	Morocco - Ceuta
ECA	<i>C. caeruleus</i>	Spain - Cadiz
SRE	<i>C. caeruleus</i>	Sweden - Revinge
GB	<i>C. caeruleus</i>	Great Britain - Oxford
Pmaj	<i>P. major</i>	Sweden - Kvismaren

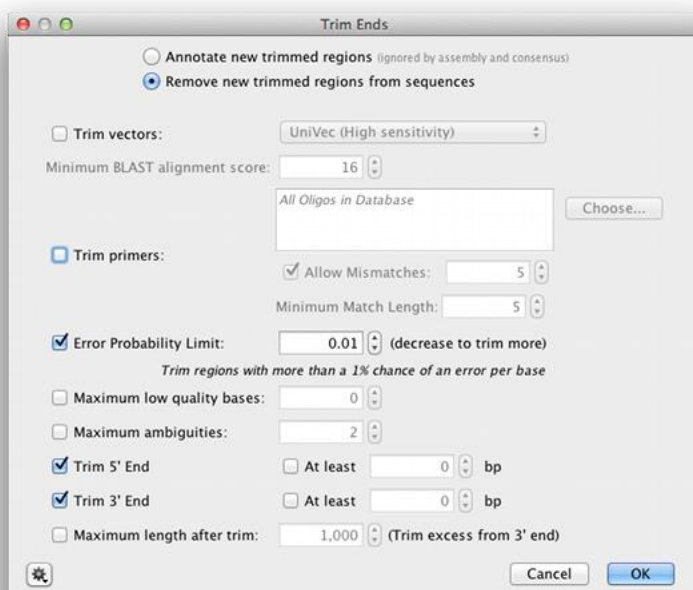
アオガラの種群には、ヨーロッパ全域に生息する *C. caeruleus*、北アフリカとカナリア諸島に生息する *C. teneriffae*、アジアと東ヨーロッパに生息する *C. cyanus* があり、ミトコンドリア DNA データは、これらの種の系統地理や集団構造を調べるために利用することができます。



まず Cyanistes CR sequences ドキュメントをクリックして選択します(ダブルクリックして別ウィンドウで表示することもできます)。表示されたシーケンスビューの右側にあるオプションパネルの **General** タブで、**Base Call Quality** を選択することで、各塩基のクオリティに応じてベースコールが強調表示されます(青色が濃いほどクオリティが低い)。

ズームアウトされた状態では、個々の塩基やクロマトグラム的高峰は見えなくなりますが、シーケンスのクオリティを示すグラフが表示されます。シーケンスを下方にスクロールしてみると、どのシーケンスも最後の方でクオリティが非常に低下していることがわかります。50%以上にズームインしてみると、クオリティの高い領域と低い領域のクロマトグラムがどのように見えるかがわかります。また CLG3 にはシーケンスがなく、シーケンシングが失敗したことを示しているため、このようなシーケンスはリストから削除します。また SRE1 も序盤以降、ピークが重複しており、クオリティの高いシーケンスが非常に短いため、このようなシーケンスも削除します。

Annotate and Predict メニュー → **Trim Ends** をクリックして、クオリティの低い塩基をシーケンスの両末端からトリミングします。**Remove new trimmed regions from sequences** を選択し、**Error probability limit** を 0.01 に設定して、**OK** をクリックします。



ここからは、一旦アラインメントしたシーケンスを編集してクリーニングする方が効率的ですので、Cyanistes CR sequences を選択し、**Align/Assemble** ボタン → **Multiple Align** をクリック、アライメントアルゴリズムとして **MUSCLE** を選択し、デフォルト設定で実行します。

作成されたアラインメントをダブルクリックして開き、ベースコールとクロマトグラムが見えるように50%程度にズームインします。3' 末端の塩基までスクロールすると、GGGGGGGAAGGGGGGGG モチーフの後でベースコールが弱くなっているのがわかります。シーケンスビュー上部にある

