

研究課題名：希少野生動物のゲノム解析

課題代表者：国立環境研究所生物・生態系環境研究センター 中嶋信美
 共同研究者：国立環境研究所生物・生態系環境研究センター 大沼 学
 酪農学園大学・獣医学群 遠藤大二
 京都大学野生動物研究センター 村山美穂
 東北大学大学院農学研究科 福田智一
 筑波大学生命環境系 田島淳史・浅野敦之

実施年度：平成 25 年度～平成 25 年度

1. 研究目的

本年度、生物・生態系環境研究センターでは次世代シーケンサーとして Ion Torrent 一台と MySeq 一台を導入した。さらに、動物園関係者等に情報を公開することにより、希少野生動物のゲノムデータを集積および分析する基盤を整備した。ゲノム全長のデータは ENSEMBL サイトに 40 種の哺乳動物と 5 種の鳥類がすでに掲載されているが、野生動物の進化および種の保全に関する知識としての実用性を求めた場合、10,000 種が取得目標とされている（10k プロジェクト）。その実現のためには、次世代シーケンサーによる得られるランダムな塩基配列データを連結して基本的には重複の無い遺伝子情報とし（アセンブル）、それらの配列を比較的近縁の動物種のゲノムとの相同性を検索することにより染色体の塩基配列を予測していく必要がある。本研究では、アセンブルプログラムおよび相同性検索ソフトをスーパーコンピュータ上に展開し、次世代シーケンサーデータを元に野生動物のゲノム配列を構築することを目的とする。

2. 研究計画

前年度においては、ヤンバルクイナ (*Gallirallus okinawae*) の次世代シーケンサーの生データをアセンブルプログラム (SparseAssembler) により連結した。今年度は、アセンブルプログラムにより作出された Contig 配列を、相同性検索プログラム (BLAT) により、既知のゲノムデータ (ニワトリゲノム) 上に位置づける。位置づけの信頼性について、隣接する Contig 間で PCR を実施することにより、検証する。Contig の予測染色体上の位置について十分な信頼性が得られれば、ニワトリゲノムを基本としてヤンバルクイナのドラフトゲノム配列を、ニワトリゲノムを元に構築する。

3. 進捗状況

前年度において、ヤンバルクイナゲノムから次世代シーケンサーより得られた 142GB 約 2.75 億本の塩基

配列データをアセンブルして合計 10^9 塩基、1,770,973 本 (平均 1,474bp) の Contig を得た。これらの Contig 配列を、鳥類の典型例としてのニワトリのゲノムとの相同性を確認したところ、ニワトリゲノムの 60% について Contig との相同性が認められた。隣接することが予測された Contig のうち、ニワトリゲノム上で近接している Contig について、連結する形での PCR を実施した。67 組の隣接 Contig 中 53 組について増幅が見られたため、ニワトリゲノム上への Contig のマッピングは、80% 程度の確率で信頼できることが示された。また、増幅された PCR 断片の塩基配列から、Contig の末端が同じ配列を持ち、重なっていることが示された。



図 1. ニワトリゲノムとの相同性により隣接することが予測されたヤンバルクイナ Contig 間領域の増幅。SparseAssembler により構築された Contig をニワトリゲノム上に相同性検索によってマッピングし、隣接 Contig 間で PCR を実施した。

4. 今後の計画

アセンブルされた配列をゲノムが既知の配列へのマッピングを行うことによりゲノムドラフトが整備されていくことが示された。作業の一部はスーパーコンピュータが用いられたが、バックグラウンドでのバッチジョブの設定が難しいため、多くは研究者の手元のデスクトップパソコンで実施された。今後は、スーパーコンピュータ上での相同性検索をバックグラウンドでのバッチジョブ化する作業を進める。その上で、研究所内の次世代シーケンサーを用いて追加で得られた塩

基配列を追加することによりヤンバルクイナゲノムの
ドラフトを構築する。

**5. 昨年度計算機資源の利用状況（2013年6
月1日～2014年3月31日）**

実行ユーザ数: 3

CPU時間 v_deb: 0 hours, v_cpu: 0 hours, v_8cpu: 0 hours,
v_16cpu: 0 hours, 計: 0 hours