

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 26 日現在

機関番号：17102

研究種目：若手研究(B)

研究期間：2013～2016

課題番号：25840129

研究課題名(和文)非平衡状態にある非モデル生物の集団史の推定

研究課題名(英文) Inference of population history of non-model organisms under non-equilibrium condition

研究代表者

手島 康介 (Teshima, Kosuke)

九州大学・理学研究院・助教

研究者番号：20447593

交付決定額(研究期間全体)：(直接経費) 2,500,000円

研究成果の概要(和文)：生物進化を理解する上で、その生物種がどのような集団の構造を経てきたのかを理解することは欠かせない。本研究ではゲノム多様性データを利用して、先行する知見の限られた非モデル生物が経てきた集団の歴史を推定する方法について研究し、実際のデータ解析を行った。特に集団の分化を測るためにも用いられている統計量の性質について整理した。現在広く使われている指標は複数のグループに分けることができることを明らかにした。ゲノムデータを用いた集団モデル構築についての研究を行い、ゲノム多様性データから集団モデルを選択する方法について検討した。さらには温帯樹木種のゲノムデータ解析を行って集団史の推定を行った。

研究成果の概要(英文)：Population history is one of the main forces which contribute to the evolution of organisms. In this research, I investigated the pattern of summary statistics which measure population divergence and classified related measurements into several categories. It was found that some of them are compatible with each other but others are not. I also considered a method to select a population model. With this method, a reasonable population model can be selected by analyzing genome-wide polymorphism data. And finally, empirical NGS data of temperate forests are analyzed and population structure and population history are estimated.

研究分野：集団遺伝学

キーワード：集団遺伝 集団構造 ゲノム多様性 シミュレーション 集団分化

1. 研究開始当初の背景

2000年代から、塩基配列解析装置の発展と一般への普及が著しい。それまではヒトゲノム計画のような官民アカデミア一体となった巨大プロジェクトでなければ実現不可能であったゲノム配列の解析が、ここの研究施設のレベルで可能となりつつある状態であった。種を代表する参照配列の作成のみならず、種内の他サンプルからゲノム多様性情報を取得することも可能となりつつある状態であった。技術そのものの発展と価格の低下が大きな要因である。さらにはゲノムレベルのデータ解析環境の充実も進んだ。計算機の性能の向上に伴い、ギガレベルのデータを個々の研究室単位で解析することが可能となった。これらの発展により、ゲノム多様性情報の取得と解析が研究室単位で行うことができるようになった。このことはそれまで先行する知見の豊富なモデル生物に限られていたゲノムレベルの解析が、非モデル生物にも応用可能となったことも意味する。このように、データの量や解析環境に大きな変革が生じていた。

生物が進化的に変化するメカニズムを理解し、その歴史を再構築することは、進化学の研究における目的の大きな目的の一つである。特にここ10年余りの間に発展したゲノム多様性データを用いた解析は、進化研究にも大きな変化をもたらした。個々の遺伝領域の研究から、ゲノム全体にわたる包括的な統計的解析を行うことが可能となった。これにより仮説検定を中心とした伝統的な集団遺伝学的解析から、個々の進化パラメータの推定が可能となった。従来と比較的簡単な進化モデルを過程した解析から、より詳細な推定が可能になりつつある状態であった。

2. 研究の目的

従来用いられていた塩基解析の解析手法は単一あるいは少数の遺伝子領域を詳細に解析することを目指すものであった。これは広くゲノムワイドな情報を得ることが難しかったという理由から来ている。例えば最尤法を用いた推定方法はデータそのものを得られる確率を最大化するという意味で、データの持つ情報を最大限活用するが、その反面実行時間が非常に長くなる。このような伝統的な方法を昨今のゲノムデータにそのまま応用すると、解析が煩雑になり時間がかかりすぎる。そこで既存の方法よりも効率よくかつ詳細に進化過程を再構築する方法を提案し、ゲノムレベルのNGSデータ解析を行うことを目的とする。特に生物集団中の個体数の増減や分集団構造、分集団間の移住といった集団レベルの進化過程の検出と集団史の再構築を行うことに焦点を絞り、そのために必要な方法論に関する基礎的な研究と、実データ

を用いたデータ解析を行う。

3. 研究の方法

本研究ではゲノム多様性データを適宜要約し、そのパターンから特徴を抽出して集団史の推定に用いる。そのために、ゲノムデータ解析の方法とシミュレーションを用いた理論研究の両面から研究を進めた。

- (1) まず集団分岐モデルや集団構造モデルを仮定し、それらの集団モデルの下でシミュレーションを行って仮想データを作成した。マイクロサテライトやSNP、塩基配列などを想定してシミュレーションデータを変換し、集団分岐の程度を測る統計量を計算した。これら必要なスクリプトは自作のものを用いた。
- (2) 集団モデルの選択に関する研究では、集団史をモデル化し、その集団史モデルの下でシミュレーションを行った。基本的にシミュレーションには汎用のシミュレーションソフトを用いたが、統計量の計算や推定については適宜スクリプトを作成した。
- (3) ゲノムデータ解析においてはサンプリング及びNGS利用とデータ取得に関しては共同研究者からの提供を受けた。NGSデータを加工、変換し、そこから変異データを抽出する過程、及び集団遺伝学的解析に用いることのできる多様性データを生成するために必要なスクリプトは自作した。得られたゲノムワイドなゲノム多様性データを解析し、集団構造の推定及び進化パラメータの推定を行った。

4. 研究成果

- (1) 集団遺伝学において用いられる多くのようやく統計量のうち、特に集団分化の程度を測る時に用いられている指標について、その性質の違いを詳細に調べた。伝統的には F_{ST} という量が広く用いられているが、その概念が提唱された時代からは得られるデータの質量ともに大きく変化している。その結果、新しいタイプの分子多様性データに対する応用が、多数の研究者によってもたらされた。その結果、微妙な目的の違い・定義の違い・前提とするデータの違いによって新たな量が乱立する事態となっていた。本研究では同一の集団進化モデルの下で複数の分子マーカーを想定したシミュレーションデータを発生させ、それぞれに対して F_{ST} から派生した統計量を計算し比較することで、性質の違いを整理した。その結果、これらの指標は類似のパターンを示す2つのグループとそれ以外の合

計3カテゴリーに分類されることが確認された。分化の指標を集団分化の程度や時期を反映するパラメータと考えるならば、そのうちのひとつのカテゴリーに属する量を用いるのが望ましいことがわかった。また逆にプロタイプ情報から計算する量は領域依存性が高く、集団分岐の状態を測る量としては不適であることも確認した。

- (2) 集団史を推定する方法はすでいくつか開発されている。それらの方法は研究者がモデルを作成し、得られた多様性データに対しパラメータをフィットさせる。どのようなモデルを用いるかという点に関しては研究者に任されている。そこで本研究では、得られたゲノム多様性データそのものから集団モデルについての情報を抽出する方法について開発を行った。特に単一の集団からサンプルが取られた時に、そのサンプルが得られた集団以外の、隠れた集団の存在をデータから推定する可能性について検討した。その結果、集団史の細部まで再現することは難しいものの、隠れた集団の存在を検出すること・集団構造の有無・集団構造の変化については可能であることを明らかにした。
- (3) スギ、ヌマスギ、イスノキといった主に温帯樹木種を対象に、PCR産物を利用してゲノムを濃縮するアンプリコンシーケンス法を用いて取得したゲノムデータの解析を行った。これらは非モデル生物であり、参照ゲノム配列も存在せず、また、先行する知見も限られている。そこでまずは正確なゲノム多様性データを取得するために必要なパイオインフォマティックスパイプラインの構築を行なった。NGSデータからエラーを排除して、バイアスを排除した多様性データの取得を行なった。次に得られた多様性データを用いて集団構造の把握と集団史の推定、進化パラメータの推定を行った。これらの結果は論文にまとめて投稿した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 3 件)

Ikezaki Y, Suyama Y, Middleton BA, Tsumura Y, Teshima K, Tachida H, and Kusumi J
Inferences of population structure and demographic history for *Taxodium distichum*, a coniferous tree in North America, based on amplicon sequencing analysis.

American Journal of Botany, 査読あり
103(11), 2016, 1937-1949

Ng WL, Onishi Y, Inomata N, Teshima KM, Chan HT, Baba S., Changtragoon S, Siregar IZ and Szmidt AE
Closely related and sympatric but not all the same: genetic variation of Indo-West Pacific *Rhizophora* mangroves across the Malay Peninsula
Conservation Genetics, 査読あり
16(1), 2014, 137-150

Urashi C, Teshima KM, Minobe S, Koizumi O, and Inomata N
Inferences of evolutionary history of a widely distributed mangrove species, *Bruguiera gymnorhiza*, in the Indo-West Pacific region
Ecology and Evolution, 査読あり
3(7), 2013, 2251-2261

〔学会発表〕(計 8 件)

手島康介
集団の遺伝的分化の指標 F_{st} に関する研究
遺伝研研究集会「分子進化学の現状と今後の展望」、2016年8月20-21日、三島市

中村遥奈, 手島康介, 舘田英典
集団サイズの周期的変動による塩基配列多型への影響
日本遺伝学会、2015/09/24-2015/09/26、仙台

河村瑳友, 手島康介
集団文化の程度を測る指標の挙動パターンに関する研究
日本遺伝学会、2015/09/24-2015/09/26、仙台

家入雄樹, 手島康介, 舘田英典
クラスタリングと判別分析による集団モデルの選択
日本遺伝学会、2014/09/17-2014/09/19、長浜市

池崎由佳, 陶山佳久, Beth A. Middleton, 津村義彦, 手島康介, 舘田英典, 楠見淳子
アンプリコンシーケンス解析による針葉樹ヌマスギの集団構造、及び分化の歴史の推定
日本遺伝学会、2014/09/17-2014/09/19、長浜市

Kosuke M. Teshima, Eri Sato, Atsuya

Kamei, Setsuko Suzuki, Kyoko Sugai,
Atsushi Sakai, Nobuyuki Inomata,
Shin-ichi Yamamoto, Hidenori Tachida
Population genetics analysis of
Distylium racemosum, a climax tree
species in Japanese evergreen
broadleaf forests
SMBE satellite meeting/NIG
international symposium: The causes of
Genome Evolution,
2014/03/14-2014/03/17、三島市

池崎由佳, 手島康介, 舘田英典, 楠見淳
子
アンプリコンシーケンス解析による塩
基配列多型データに基づく針葉樹ヌマス
ギの集団構造の推定
日本遺伝学会、2013/09/19-2013/09/21、
横浜市

佐藤衣里, 亀井敦哉, 手島康介, 舘田英
典
常緑広葉樹イスノキ(Distylium
racemosum)及びシマイスノキ
(D. lepidotum)の集団遺伝学的解析
日本遺伝学会、2013/09/19-2013/09/21、
横浜市

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等

6. 研究組織
(1)研究代表者

手島 康介 (TESHIMA, Kosuke)

九州大学・大学院理学研究院・助教
研究者番号：20447593

(2)研究分担者
()

研究者番号：

(3)連携研究者
()

研究者番号：

(4)研究協力者
()