

令和 4 年 6 月 6 日現在

機関番号：17102

研究種目：基盤研究(C) (一般)

研究期間：2019～2021

課題番号：19K06782

研究課題名(和文)連続的に分布する生物の集団遺伝学

研究課題名(英文)Population genetics of continuously distributed populations

研究代表者

手島 康介(Teshima, Kosuke)

九州大学・理学研究院・教授

研究者番号：20447593

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：生物は広がりを持った空間に生息し、その距離に応じて交流を保ちつつ、遺伝的多様性を保持している。現在の集団遺伝学的解析では、本来は連続的に分布している生物集団を離散的な集団の組み合わせに近似して解析しているが、この方法は現実を適切に把握できているのだろうか。本研究ではシミュレーションを用いた解析を実行し、問題点とその対策について研究を進めた。その結果、集団構造の測定及び統一的な解釈のためには特定の統計量の組を使う必要があることを示した。さらにデータの種類にも留意する必要がある。タイピングデータを用いた場合は不可避免的に情報が歪むが、マーカーの作成方法を適切に行うことでこの影響を抑えることができる。

研究成果の学術的意義や社会的意義

生物の多様性は、そのゲノム多様性に起因している。ゲノム多様性は集団の構造や歴史に影響されるが、現在の解析方法では近似的解析であるため、本来の姿を適切に反映しているかどうか分からない部分がある。本研究ではシミュレーションを行って問題点の把握と解決策の提案を行った。集団構造の測定に用いられる統計量や解析に用いるデータが解析を歪ませることを量的に明確に示したことは有意義である。さらにこれらの影響を避けるための方法を提案することで今後の研究に指針を提示したという意味を持つ

研究成果の概要(英文)：Living organisms inhabit a continuous space and maintain genetic diversity. In the current population genetic analysis, the originally continuously distributed population of organisms is approximated as a combination of discrete sub-populations. Does this method adequately represent the reality? In this study, I performed a simulation-based analysis to conduct these issues. The results showed that it is necessary to use a specific set of  $F_{st}$  statistics for measuring and interpreting population structure. Furthermore, the type of data must be taken into account. The use of genotyped data inevitably distorts the information, but this effect can be minimized by using appropriate methods of creating markers.

研究分野：集団遺伝学

キーワード：集団遺伝 集団構造 シミュレーション

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

多くの生物は広がりを持った空間に生息し、その距離に応じて交流を保ちつつ、遺伝的多様性を保持している。集団は一般に複雑な構造を持ち、かつその構造は時間とともに移り変わるのが普通である。現在の集団遺伝学的解析では、この本来は連続的に分布している生物集団を離散的な集団の組み合わせに近似して解析している。実際の生物は時空間的にはほぼ連続的に存在しているのに対し、その解析理論は少数集団・少数遺伝子座・比較的限られたサンプルを元にして構築された理論を用いており、ここに現実の生物の状況と解析方法の乖離が存在していた。

### 2. 研究の目的

我々がこれまでに行ってきた遺伝的多様性解析はそもそも現実を適切に把握できているのだろうか。逆の言い方をすると、現実を的確に認識するためにはどのような解析を行えばいいのだろうか。この点について明らかにすることが本研究の目的である。従来の単純化した解析では把握できていない問題に対し、どうすれば検出することができるのであろうか。その解析方法を提案することがさらなる目的である。

### 3. 研究の方法

本研究ではシミュレーションを作成・実行して、様々な条件下の多様性データを作成した。必要に応じてスクリプトを作成してバイアスの導入及び除去、データの変換を行った。それらのデータから要約統計量を計算し、解析に用いた。

### 4. 研究成果

#### (1)

多くの生物は広がりを持った空間に生息し、その距離に応じて交流を保ちつつも遺伝的多様性を保持している。集団遺伝学的解析では、この本来は連続的に分布している生物集団を離散的な集団の組み合わせに近似して解析しているが、そのためには本来の集団の状況を適切に把握しなければならない。本研究では理論やシミュレーションを用いて、集団の状況を適切に把握するための研究を遂行した。特に近年は全ゲノム SNP データを用いた解析が多いが、これら SNP データはその取得方法に気をつけなければバイアスを含む可能性がある。そこで、タイピング法を用いて作成された SNP データはどの程度正確に元の集団の情報を反映しているのか、バイアスを含んだデータを使って解析した場合どの程度歪んだ結果を導くのか、バイアスのない解析を行うためにはどのような方法が望ましいのか、という点について研究を進めた。

一連のシミュレーションの結果、タイピングデータは不可避免的に情報が歪むことを明確にした。さらにタイピングデータを用いて集団構造の解析を行うと、そのデータの作成プロセスに応じて歪んだ構造推定が行われることが明らかとなった。これらのひずみはデータ作成プロセスに注意を払うことで、影響を小さく抑えることができる。具体的には、サンプルは特定の集団に集中させず、かつマーカーは全集団の多様性情報を総合して作成することが重要である。これらの結果は論文に取りまとめて発表した(Dokan et al. G3, 2021)。

#### (2)

連続空間モデルの下での集団遺伝学は非常に多くのパラメータを含みうる柔軟な理論枠組みであるため、伝統的な理論集団遺伝学のように解析的に式で解くことは現実的ではない。むしろ大規模なシミュレーション、すなわち仮想的進化実験を行うことで、これまでは複雑すぎて定式化できずに研究されてこなかった現象を理解し、特徴を検出することができる。申請者がかつて開発したシミュレーター-continuous-pop(Teshima 未発表)を拡張し、局所的な適応進化に関しては申請者が既に公開している mbs(Teshima and Innan 2009)や準備中の multipop-mbs(Yamamoto and Teshima, in prep)を実装した。これにより、複数集団において集団分化や集団サイズの変動の影響下で、自然選択の影響を考慮した配列多様性データを生成することが可能となった。本プログラムは発表準備中である。

#### (3)

集団構造の把握には、その時に用いる統計量にも影響する。これまで集団構造の把握には  $F_{st}$  等

の統計量が用いられてきたが、その具体的計算方法は複数存在する。これらの量はそれぞれ特性が異なり、得られる分化程度が異なるため、解析時には個々の統計量の実装に留意する必要がある。集団分岐過程や集団構造の影響と、これらの統計量との関係を明らかにした。これまでに使われてきたもののうち普遍的で比較可能な統計量の組を明らかにした。今後、集団構造の解析にはこれらの統計量に限って用いるべきであることを明らかにした。この結果は論文に取りまとめ投稿予定である

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Dokan Kotaro, Kawamura Sayu, Teshima Kosuke M	4. 巻 11
2. 論文標題 Effects of single nucleotide polymorphism ascertainment on population structure inferences	5. 発行年 2021年
3. 雑誌名 G3 Genes Genomes Genetics	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/g3journal/jkab128	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計5件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 中村遥、早川敏之、手島康介
2. 発表標題 適応進化とハプロタイプ長の関係についての考察
3. 学会等名 日本遺伝学会
4. 発表年 2020年

1. 発表者名 道管公太郎、河村瑛友、手島康介
2. 発表標題 バイアスを含んだデータのFSTへの影響
3. 学会等名 日本遺伝学会
4. 発表年 2020年

1. 発表者名 道管公太郎、河村瑛友、手島康介
2. 発表標題 集団構造推論におけるアサートメントバイアスの影響
3. 学会等名 日本遺伝学会
4. 発表年 2021年

1. 発表者名 田中智崇, 手島康介
2. 発表標題 自然選択の検出力の評価
3. 学会等名 日本遺伝学会
4. 発表年 2021年

1. 発表者名 中村遥、手島康介、早川敏之
2. 発表標題 適応進化とハプロタイプ長の関係
3. 学会等名 日本遺伝学会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------