

令和 5 年 6 月 15 日現在

機関番号：82101

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K06102

研究課題名（和文）環境DNA分析による検出誤差を踏まえた種多様性評価手法の開発と検証

研究課題名（英文）Development and validation of the method for species diversity assessment based on environmental DNA analysis accounting for species detection errors

研究代表者

深谷 肇一（Fukaya, Keiichi）

国立研究開発法人国立環境研究所・生物多様性領域・主任研究員

研究者番号：30708798

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）：環境DNAメタバーコーディングは、多数の種の分布データを効率的に収集できる環境DNA分析の技術である。しかし、本手法の応用において生じる種検出の誤差（偽陰性と偽陽性）は得られるデータの解釈を困難にしており、信頼性の高い調査を実現するためには種検出の誤差に対処するための手法開発が求められる。本研究では環境DNAメタバーコーディングにおける偽陰性の種検出誤差に着目し、これに適切に対処し、正確かつ効率的に種多様性の評価・予測を行うための統計的手法を構築した。また、それを簡便に利用するためのソフトウェアを開発した。

研究成果の学術的意義や社会的意義

生物多様性の保全において、野外の種多様性を定量的に把握・監視することは重要である。しかし、そのための生態調査は高価かつ侵襲的であることが多く、効率的な保全計画の実現を困難なものにしている。環境DNAメタバーコーディングはこの問題に対する解決策として大きな可能性を秘めているが、解決すべき技術的課題も残されており、種の検出誤差の問題はその1つである。偽陰性を踏まえた種多様性評価のためのデータ解析手法を提供する本研究は、環境DNAメタバーコーディングを検出誤差に対して頑健な手法に変えることで、その費用対効果と信頼性の向上に貢献する。

研究成果の概要（英文）：Environmental DNA metabarcoding is an environmental DNA analysis technique that efficiently collects occurrence data for many species. However, the species detection errors (i.e., false negatives and false positives) that can occur in the application of this approach make interpreting the resulting data challenging, and the development of methods to deal with species detection errors is required to achieve a reliable survey. This study focused on false-negative errors in environmental DNA metabarcoding and developed a statistical method to appropriately address these errors for accurate and efficient assessment and prediction of species diversity. The proposed method was implemented in a software package to make it easier to use.

研究分野：統計生態学

キーワード：環境DNAメタバーコーディング サイト占有モデル 調査設計 魚類群集

## 様式 C-19、F-19-1、Z-19（共通）

### 1. 研究開始当初の背景

生物多様性の保全においては、野外の種多様性を定量的に把握・監視することがまずもって重要である。しかし、そのための生態調査が一般的に高価かつ侵襲的であることは、効率的な保全計画の実現を困難なものにしている。ところが近年、水や土壌、空気などの環境中に含まれる DNA（環境 DNA）を分析することにより、魚や鳥、哺乳類などのマクロな野生生物の分布を把握できることが分かってきた。目視や捕獲による従来手法に比べると、野外での環境試料の収集にかかる労力は極めて小さく、また非侵襲的であることから、環境 DNA 分析は野外生態調査の在り方を変革する新技術として大きな注目を集めている。特に、ユニバーサルプライマーと次世代シーケンサーを用いて生息種を網羅的に検出するアプローチは「環境 DNA メタバーコーディング」と呼ばれており、多数の種の分布データを一度に収集・蓄積できるという利点がある。これにより、従来は見込めなかった規模での種多様性調査も可能となりつつある。実際、環境 DNA メタバーコーディングは多くの学術研究で用いられているだけでなく、企業や省庁、自治体等の事業においても、生物多様性評価や生態系モニタリングへの活用が検討されている。

一方で、環境 DNA メタバーコーディングによるマクロ生物の生態調査法は新しい技術であり、未だ多くの課題が残されている。特に、野外における DNA 濃度の低さや PCR の増幅誤差などに由来する種検出誤差（偽陰性と偽陽性）は、本手法の適用における懸念事項の 1 つである。すなわち、そこに生息している種の DNA が不検出となる場合や、逆に、本来は存在しない種の DNA が検出されてしまうことが生じうるため、生物の捕獲や視認を伴わない環境 DNA メタバーコーディングデータの解釈は容易でない。環境 DNA 分析による生態調査の信頼性をより高いものとするために、環境 DNA メタバーコーディングにおける種検出の誤差に対処するための手法開発が求められる。

### 2. 研究の目的

本研究では環境 DNA メタバーコーディングにおける偽陰性の種検出誤差に着目し、これに適切に対処し、正確かつ効率的に種多様性の評価・予測を行うための統計的手法を構築すること、およびそれを簡便に利用するためのソフトウェアを開発することを目的とした。

### 3. 研究の方法

(1) 環境 DNA メタバーコーディングに固有のデータ形式である配列リード計数を説明する統計モデルの新しいクラスを定式化する。生態学では、逐次的なデータ生成過程において生じる偽陰性を説明する統計モデルとしてサイト占有モデルが知られており、検出誤差を踏まえて野生生物の生息確率を評価するためのツールとして広く利用されている。そこで、本研究ではまず、配列リード計数データに適合するようなサイト占有モデルの拡張を提案する。さらに、サイト占有モデルを演繹的な確率モデルとみなせば任意の研究デザインの下での期待検出種数を評価できることから、提案モデルに基づき最適な研究デザインを導く手法を構築する。

(2) 淡水魚類を対象とした環境 DNA メタバーコーディングをモデルケースとして、本研究で構築する手法の有効性を確認する。具体的には、環境条件や種構成が異なる霞ヶ浦水系の 50 定点で採水を行い、魚類に対して有効なユニバーサルプライマー MiFish を用いて環境 DNA メタバーコーディングによる種検出を行う。得られたデータにサイト占有モデルを当てはめることで、各魚種の地理分布と局所・地域種多様性を評価する。また、種の検出に関連するパラメータの変動を評価することで環境 DNA メタバーコーディングにおける種検出の潜在的なボトルネックを特定し、研究デザインにおいて効果的な種検出につながる要因を特定する。

(3) 本研究で構築する手法を利用しやすい形で一般に提供することで、検出誤差に頑健な環境 DNA メタバーコーディングの普及を目指す。具体的には、(a) サイト占有モデルの構築と当てはめ、(b) その結果に基づく種多様性評価、(c) 種検出効率の観点からの研究デザインの評価と最適化を容易に行うためのプログラム群を、統計解析に特化したフリーソフトウェアである R の拡張パッケージとして実装して公開する。

### 4. 研究成果

(1) 多種サイト占有モデルと呼ばれる不完全な検出を群集レベルで説明する統計モデルを拡張することで、環境 DNA メタバーコーディングに固有の種検出過程を表す統計モデルを定式化した (図 1)。多地点で収集された環境 DNA メタバーコーディングデータにこのモデルを当てはめることにより、環境 DNA メタバーコーディングによる個々の種の検出可能性と、偽陰性を考慮した種の生息確率を定量的に評価できるようになった。また、このモデルの下で予測される検出

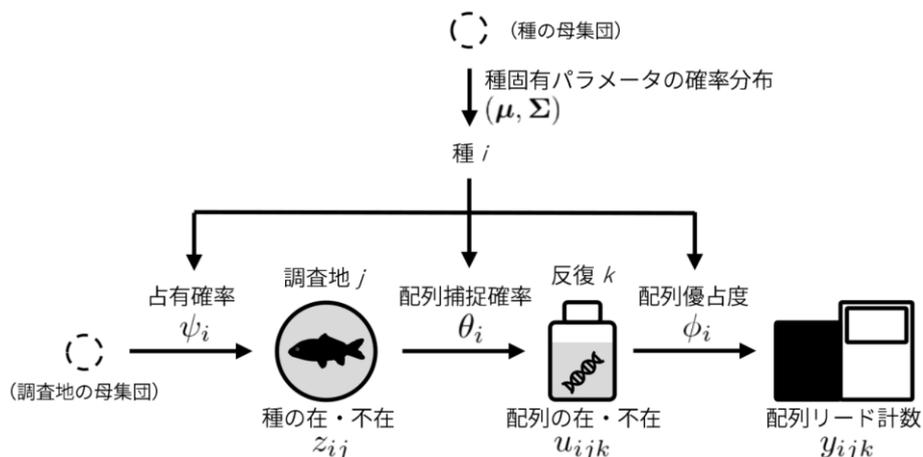


図 1. 環境 DNA メタバーコーディングのデータ生成過程を表す統計モデルの概要 (Fukaya et al. 2022)。種固有パラメータ ( $\psi_i, \theta_i, \phi_i$ ) の推定値に基づき異なる段階での偽陰性の生じやすさが、潜在変数  $z_{ij}$  の推定値に基づき偽陰性を考慮した種の在・不在が、それぞれ評価される。

種数の値に基づき、対象の系について最適な反復数や配列リード総数を特定するための一般的な手続きを構築した。

(2) 上記の提案手法を、霞ヶ浦水系に生息する淡水魚類を対象とした環境 DNA メタバーコーディングのデータに適用した。DNA 配列が検出された 50 分類群のそれぞれについて環境 DNA メタバーコーディングによる検出可能性が推定された結果、環境 DNA による生息種の検出しやすさには大きな不均一性があり、生息種の不検出 (偽陰性) を防ぐ上で同一地点から反復して採水を行うことが有効であることなどが明らかとなった。

(3) 提案手法を簡便に扱うことのできるフリーの解析ソフトウェアを統計解析向けのプログラム言語 R の拡張パッケージとして開発し、そのアルファ版を公開した (<https://github.com/fukayak/occumb>)。計画している主要な機能の大部分は実装済みであるが、今後、ドキュメントの整備や既知のバグ修正などをさらに進めて正式版をリリースする予定である。

(4) 本研究で構築した手法とソフトウェアは、令和 2~4 年度に実施された環境省の環境 DNA 技術標準化・一般化業務の重点調査データの解析に利用され、二次的自然環境に生息する淡水魚類を対象とした環境 DNA 調査のためのマニュアル「環境 DNA 分析技術を用いた淡水魚類調査手法の手引き」の改定に貢献した。

(5) 提案手法の有効性を示す追加的な事例研究として、魚類に比べて環境 DNA メタバーコーディング手法の最適化が進んでいない水生昆虫の環境 DNA メタバーコーディングデータに提案手法を適用することで、水生昆虫種の効果的な検出方法に関する検討を行った。環境 DNA メタバーコーディングによる種の検出しやすさは、水生昆虫においても大きなばらつきがあることが明らかとなり、偽陰性のリスクを減少させる上で効果的な研究デザインについての示唆が得られた。

#### <引用文献>

K. Fukaya, N. I. Kondo, S. S. Matsuzaki and T. Kadoya (2022) Multispecies site occupancy modeling and study design for spatially replicated environmental DNA metabarcoding. *Methods in Ecology and Evolution* 13:183–193.

5. 主な発表論文等

〔雑誌論文〕 計2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 2件）

1. 著者名 K. Fukaya, N. I. Kondo, S. S. Matsuzaki, T. Kadoya	4. 巻 13
2. 論文標題 Multispecies site occupancy modelling and study design for spatially replicated environmental DNA metabarcoding	5. 発行年 2022年
3. 雑誌名 Methods in Ecology and Evolution	6. 最初と最後の頁 183-193
掲載論文のDOI（デジタルオブジェクト識別子） 10.1111/2041-210X.13732	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 -

1. 著者名 深谷肇一	4. 巻 47
2. 論文標題 マクロ生物調査のための環境DNA分析 種の検出と定量およびその他の応用における可能性と課題	5. 発行年 2022年
3. 雑誌名 全国環境研究会誌	6. 最初と最後の頁 159-165
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
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11件（うち招待講演 3件/うち国際学会 1件）

1. 発表者名 深谷肇一, 今藤夏子, 松崎慎一郎, 角谷拓
2. 発表標題 環境DNAメタバーコーディングにおける種検出誤差を考慮した階層モデリングとその実践
3. 学会等名 環境DNA学会第4回大会
4. 発表年 2021年

1. 発表者名 松本岳大, 深谷肇一, 坂田雅之, 沖津二郎, 稲川崇史, 平岡康介, 一柳英隆, 源利文
2. 発表標題 ダム湖の魚類調査における環境DNAメタバーコーディング分析の実装に向けて
3. 学会等名 環境DNA学会第4回大会
4. 発表年 2021年

1. 発表者名 深谷肇一
2. 発表標題 生態学的ビッグデータと統計モデリングによる生物多様性評価の高度化
3. 学会等名 第69回日本生態学会大会
4. 発表年 2022年

1. 発表者名 松本岳大, 深谷肇一, 坂田雅之, 沖津二郎, 稲川崇史, 平岡康介, 一柳英隆, 源利文
2. 発表標題 多種サイト占有モデルを用いた環境DNA分析における調査デザインの最適化
3. 学会等名 第69回日本生態学会大会
4. 発表年 2022年

1. 発表者名 深谷肇一, 今藤夏子, 松崎慎一郎, 角谷拓
2. 発表標題 環境DNAによる生物多様性評価のための階層モデルとベイズ決定分析
3. 学会等名 予測モデリングの理論と応用. 科学研究費シンポジウム(基盤研究(A)20H00576)「大規模複雑データの理論と方法論の革新的展開」
4. 発表年 2021年

1. 発表者名 深谷肇一
2. 発表標題 環境DNA分析による生態系評価の高度化: 統計モデリングによるアプローチ
3. 学会等名 統計数理研究所共同利用研究集会「生物統計と数理生物: 似ているようで異なる生物へのアプローチ」(招待講演)
4. 発表年 2022年

1. 発表者名 深谷肇一
2. 発表標題 生物多様性の理解におけるプレストン不備の解決に向けて
3. 学会等名 環境DNA学会第3回大会第36回個体群生態学会大会合同大会（招待講演）
4. 発表年 2020年

1. 発表者名 深谷肇一
2. 発表標題 階層モデリングをデータ解析と調査設計に活用する
3. 学会等名 第68回日本生態学会大会
4. 発表年 2021年

1. 発表者名 松本岳大, 深谷肇一, 坂田雅之, 稲川崇史, 沖津二郎, 平岡康介, 一柳英隆, 源利文
2. 発表標題 ダム湖において環境DNAメタバーコーディング手法を用いる際の最適採水地点数の検討
3. 学会等名 第68回日本生態学会大会
4. 発表年 2021年

1. 発表者名 深谷肇一
2. 発表標題 生態学における観測・推測・予測と統計モデリング
3. 学会等名 第68回日本生態学会大会（招待講演）
4. 発表年 2021年

1. 発表者名 K. Fukaya, N. I. Kondo, S. S. Matsuzaki, T. Kadoya
2. 発表標題 A hierarchical modeling approach for environmental DNA metabarcoding: inference of species detection process, site occupancy, and study design
3. 学会等名 International Statistical Ecology Conference (ISEC) 2022 (国際学会)
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

空間的に反復された環境DNAメタバーコーディングのための多種サイト占有モデリングと研究デザイン <a href="https://www.nies.go.jp/biology/research/publish/202113.html">https://www.nies.go.jp/biology/research/publish/202113.html</a>
--

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
研究分担者	今藤 夏子  (Kondo Natsuko)  (10414369)	国立研究開発法人国立環境研究所・生物多様性領域・主任研究員    (82101)	
研究分担者	角谷 拓  (Kadoya Taku)  (40451843)	国立研究開発法人国立環境研究所・生物多様性領域・室長    (82101)	

	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
研究協力者	長谷部 勇太  (Hasebe Yuta)		

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
---------	---------