

令和 5 年 5 月 30 日現在

機関番号：82708

研究種目：基盤研究(C) (一般)

研究期間：2019～2022

課題番号：19K06862

研究課題名(和文) 家系生態学：血縁情報から集団構造を解明する新しい理論的基盤の構築とその応用

研究課題名(英文) Ecological inference based on kinship assignment

研究代表者

秋田 鉄也 (Akita, Tetsuya)

国立研究開発法人水産研究・教育機構・水産資源研究所(横浜)・主任研究員

研究者番号：60625507

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：対象とする生物種がどのような単位でまとまっているのを知ることが、その生物の生き様に関する理解のみならず、個体数の維持や持続的な利用など、応用の面でも重要である。しかしながら、1単位のまとまり(=集団)の大きさを調べるのは非常に難しい。近年、個体数が小さな集団ほど近親関係が見つかりやすいことを利用して、集団の個体数を調べる方法が出現した。本研究ではこの方法を拡張して、半兄弟関係や親子関係から、個体数(N)・進化的な単位として機能する個体数( $N_e$ )・ $N_e/N$ ・集団間の移動個体数、のそれぞれを高精度で推定する理論を開発した。さらに、機械学習を用いて効率的に近親関係を見つけるアルゴリズムを開発した。

研究成果の学術的意義や社会的意義

近親関係から個体数や集団構造を推定する手法について、水産資源種など個体数が大きな集団を対象とする場合は近親関係が見つかりにくく、現実的な手法ではなかった。本研究では、この問題点を克服するため、複数の近親関係の同時利用や少ないサンプルサイズでも機能する推定理論を開発した。これらの成果により、様々な状況に応じたサンプル数や予算規模のもとで個体数や集団構造を明らかにする理論的基盤を整備した。開発された移動個体数推定理論は、遺伝的に分化していない2集団を対象にできることから、頻繁に移動がある生物集団の管理単位を定量的に評価するのに利用可能である。

研究成果の概要(英文)：It is important to know the characteristics of populations not only for understanding how the organisms live, but also for applications such as wildlife management and sustainable use. However, it is usually very difficult to determine the population size. Recently, a method has emerged to estimate the number of mature individuals in a population by utilizing the fact that kinship pair is more likely to be found in a smaller population. In this study, we extended this method and developed a theory to estimate the census population size (N), the effective population size functioning as an evolutionary unit ( $N_e$  or  $N_b$ ),  $N_e/N$  (or  $N_b/N$ ), and the number of migrants moving between populations, from the number of half-sibling and/or parent-offspring pairs in a sample. In addition, we developed an algorithm to efficiently find kinship relationships using machine learning.

研究分野：集団ゲノミクス

キーワード：近親関係 個体数推定 集団構造 ランダムフォレスト 有効集団サイズ 有効繁殖サイズ 移動率推定 不偏推定量

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

## 1. 研究開始当初の背景

生態学とは、「生物の分布と密度を規定している相互作用を解明する科学」である<sup>1)</sup>。現存する生物のゲノム情報には、その生物の過去の個体数・空間的な分布・移動・進化過程の情報など(以下、集団構造と呼ぶ)に関する情報が含まれている。ゲノム解析技術の発展に伴い、ゲノム情報から生態学的な問いに挑戦することが可能になってきた。

近年、近親標識法<sup>2)</sup>と呼ばれる新たな方法論が出現したことで、ゲノム情報を生態学へ応用できる範囲が格段に広がりつつある。この方法論は、「放流した“標識”がサンプル中に見出されやすいほど、集団サイズは小さい」ことを利用した標識放流再捕法の応用である。すなわち、ゲノム情報をもとに見出される遺伝的な近親関係を“再捕”とみなし、集団構造や繁殖生態に関する情報を推定する方法であり、主に水産資源解析の文脈において発達してきた。

近親標識法は非常に新しい方法論であり、繁殖成功に関する過分散の検定<sup>3)</sup>や集団サイズの不偏推定量の開発<sup>4)</sup>など理論的側面については研究代表者の貢献があったものの、申請時点における実際の応用例は、豪州研究者が主導して進めた2例に限られていた。1例目は親子関係からミナミマグロの親魚尾数を推定した例<sup>5)</sup>であり、2例目は半兄弟関係からホオジロザメの親魚尾数と生存率を推定した例<sup>6)</sup>である。現状の近親標識法を応用する場合、ある程度の信頼性のもとで集団構造を推定するには、非常に多くのサンプル数と、近親関係を検出するための多額の予算が必要である。例えば、真の集団サイズを1万、10万、100万とした場合、変動係数15%の精度で親子関係から集団サイズを推定するには、母集団からそれぞれ1千個体、3千個体、1万個体をサンプルして、近親判別をする必要がある。上記の応用例については、社会的関心が高く潤沢な予算が配分され、かつ、長期間のサンプリング体制が維持されていたために近親標識法が適用されたが、希少種など十分なサンプリングが困難な種を扱う場合や、予算が限られている場合については、現状の近親標識法を適用するのは現実的ではない。

以上より、近親標識法を広範に应用するためには、ある程度の推定精度を担保しつつも、より少ないサンプル数で実施可能な方法論を開発することが急務である。この目的のためには、複数の近親関係を抽出し、これらを同時に推定に利用することが有効である。しかしながら、そのためには下記のような2点の克服すべき問題点が存在する。1つ目は、親子もしくは半兄弟以外の近親関係から集団サイズ等を推定する理論が存在しない点である。2つ目は、親子以外の近親判別を対象とする場合、必要なDNAマーカー数の増加に伴い、コストが格段に増加する点<sup>7)</sup>である。

## 2. 研究の目的

本申請では、上記で挙げた問題点を克服する方法論の開発に取り組むことで、近親標識法が対象とする生物集団の範囲を広げることを目指した。具体的には、

(1): 複数の近親関係を同時に用いた集団構造推定法の開発

(2): より少ないDNAマーカー数から近親判別を実施するアルゴリズムの開発  
という2つの課題に取り組んだ。

## 3. 研究の方法

(1) まず初めに、集団中の同一コホートからサンプルした2個体が半兄弟関係にある確率( )を導出した。

$$\pi \approx \frac{c}{N + c - 1} \quad (1)$$

$c$ は $(1 + \phi^{-1})E[\lambda^2]/E[\lambda]^2$ と表される係数であり、親の繁殖成功度の期待値( )の不均一性と、家系間の運不運を考慮した親の個体差( は負の二項分布の過分散パラメータ)を要約したものである(図1A)。次に、進化的な単位である有効集団サイズを年あたりに換算した有効繁殖サイズ( $N_b$ )について、サンプルから見出された半兄弟ペア数から不偏推定する理論を開発した<sup>8)</sup>。加えて、半兄弟ペア数に加えて見出された親子ペア数も同時に用いて、有効繁殖サイズと成熟個体数の比( $N_b/N$ )を不偏推定する理論を開発した<sup>9)</sup>。

$$\widehat{N}_b = \frac{2n_0(n_0 - 1) + 1}{H_{HS} + 1}, \quad \widehat{N}_b = \frac{\widehat{N}_b H_{PO}}{2n_p n_0} \quad (2)$$

$n_p$ 、 $n_0$ 、 $H_{HS}$ 、 $H_{PO}$ はそれぞれ、親のサンプルサイズ、子のサンプルサイズ、見出された半兄弟ペア数、見出された親子ペア数である。また、年を跨いで親が移動する集団を対象に、異なる集団間で見出された半兄弟ペアおよび親子ペアから、一方向に移動した成熟個体数( $M$ )を推定する理論を開発した<sup>10)</sup>(図1B)。

$$\widehat{M} = \frac{2n_{p,1}n_{0,2}n_{p,2}(H_{HS,bet} + H_{PO,bet})}{(2n_{0,2} + n_{p,2})(H_{PO,1} + 1)(H_{PO,2} + 1)} \quad (3)$$

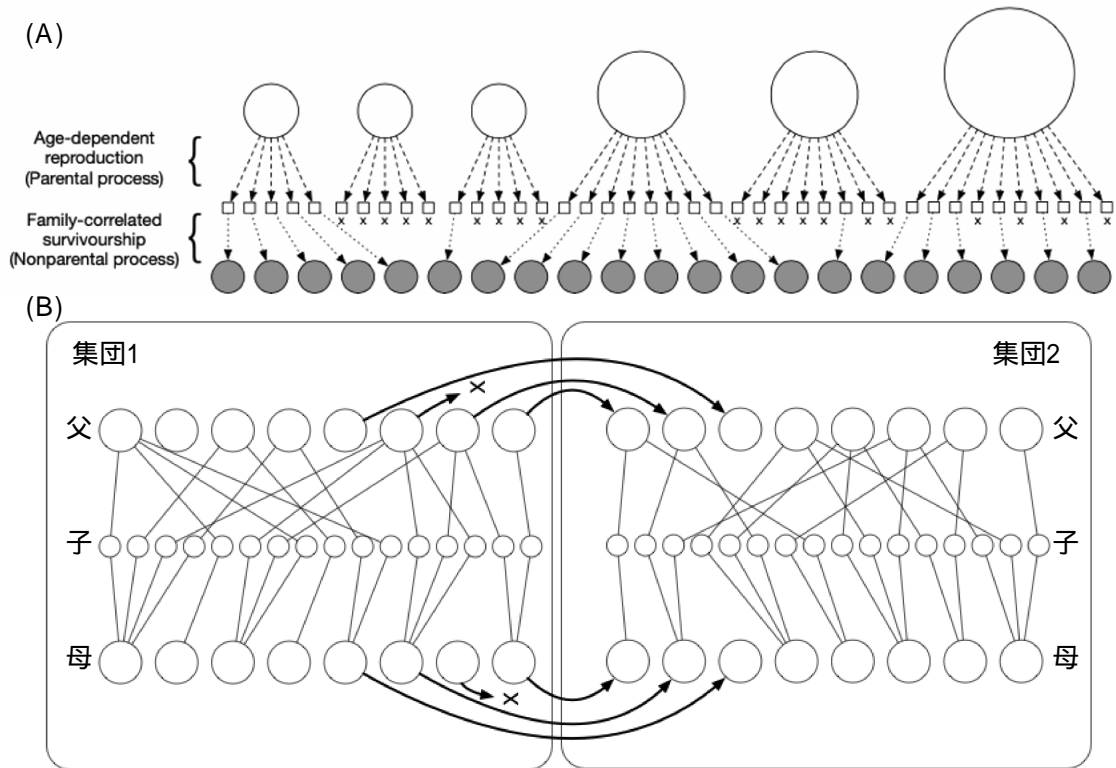


図 1 : (A) 親の繁殖成功度の不均一性と親間の運不運を考慮した繁殖過程の模式図。白丸が親個体で、大きさが繁殖成功度を表す。四角は卵の数を、色付きの丸は初期生残過程を超えた子をそれぞれ意味する。モデルでは、半兄弟ペアは子の中から、親子ペアは子と親から見出される。(B) 親が年を跨いで移動する場合の模式図。親も子も、初年度は集団 1 から、次年度は集団 2 からサンプルされる。

式 3 で使われている文字について、添字の 1 と 2 は移動前の集団と移動後の集団でサンプルされた/見出された近親関係であることを意味し、添字の bet は移動前の集団と移動後の集団との間で見出された近親関係であることを意味する。添字の P0 と HS は、それぞれ親子関係と半兄弟関係を意味する。この推定量は、親の繁殖成功度が不均一であっても、繁殖成功度が移動率と独立である限り成立する。

( 2 ) 近縁判別に必要なマーカー数を抑えつつ精度および頑強性の向上を目指して、サンプルの個体ペアの IBD ( identity by descent ) の情報を用いて、ランダムフォレスト法 <sup>11)</sup> による機械学習によって、親子・全兄弟・半兄弟およびそれ以外の近縁関係に分別するアルゴリズムを開発した ( 図 2 A )。本手法は半兄弟検出の精度向上を目的として、「遠縁ペア」の存在を考慮する。遠縁ペアは祖父母-孫、異母兄弟叔父-甥など、比較的弱い近縁関係を持つペアであり、IBD 等の各種近縁指標において半兄弟と近い値を示し、サンプル中に無視できない頻度で出現するため、しばしば半兄弟検出の精度を低下させる。機械学習による分別に、本来の目的である親子・全兄弟・半兄弟に加えて遠縁ペアを含めることで、偽陽性を抑制できる。本手法では、遠縁ペアの典型例として半従兄弟を使用する。機械学習に必要な学習データとして、サンプルから得られた DNA マーカーの情報にもとづくシミュレーションによって、親子・全兄弟・半兄弟・半従兄弟・非血縁の 5 種類の近縁関係の遺伝子型データを十分な個体数生成し、最尤法 <sup>12)</sup> によりそれらの IBD を求める。これにランダムフォレスト法を適用し、5 種類の近縁関係を分別する予測モデルを構築する ( 図 2 B 右 )。最後に、観察されたサンプルペアについて IBD を計算し、予測モデルを用いて各ペアを 5 種類の近縁関係のいずれかに割り当てる ( 図 2 B 左 )。

#### 4 . 研究成果

( 1 ) 個体ベースシミュレーションによって、推定量の不偏性が確かめられ、サンプル数と不確実性との関係性が調べられた。その結果、特に  $N_b$  の推定について、既存のモーメント推定量よりも遥かに高精度であることが判明した。すなわち、十分なサンプル数を確保できない状況下においても推定精度が担保されるようになり、 $N_b$  の推定の適用範囲が広がったとみなせる。また、1 回のサンプリングに基づく推定量としては、世界初の不偏推定量の導出であると考えられる。近親関係だけを用いて  $N_b$  や  $N_b/N$  を推定するこれらの試みは、有効繁殖サイズの開発者であり当該分野の第一人者である Robin Waples 博士によって、レビュー論文で紹介された <sup>13)</sup>。

移動個体数の推定について、遺伝的に分化のない集団間における親の移動を推定する理論は皆無であった。唯一の例外は標識を用いた方法であるが、様々な不確実性が移動率推定を困難にすることが知られている。今回開発した推定量によって、移動が頻繁にある集団間を対象

に推定可能であることを示した。この成果は、近親関係が移動の情報有することを示したのみならず、水産資源の管理単位の妥当性検証に資する新たな方法論の提示と考えられる。

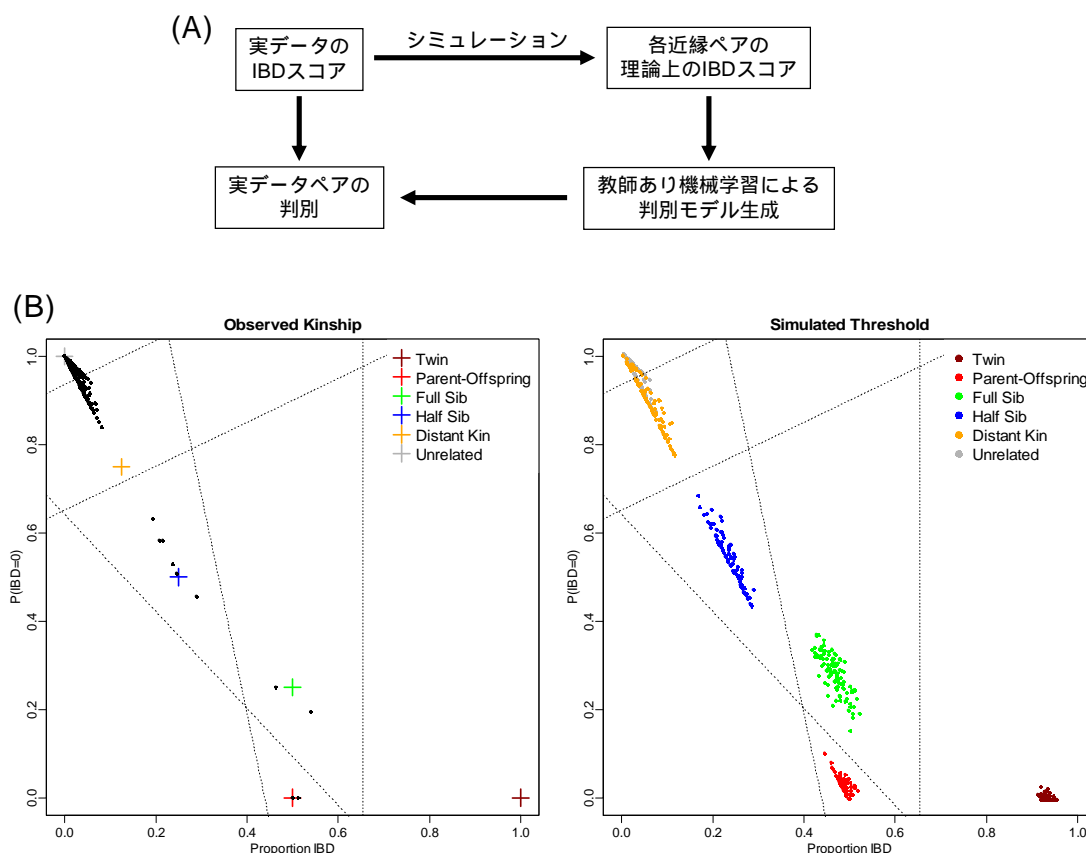


図 2: (A) 機械学習による近縁判別のフローチャート。(1)シミュレーションによる各近縁関係の IBD の理論値算出、(2)理論値を学習データとした判別モデル生成、(3)実データの IBD に判別モデルを適用、の 3 工程からなる。(B) IBD による判別モデルの概念図。右図がシミュレーションによる各近縁関係の IBD の理論値の分布および、そこから学習した判別境界(点線)である。左図が観測された IBD の分布であり、右図で得られた判別境界を適用する。この図は IBD を平面図に縮約したものであり、実際の IBD は 3 次元空間に分布する。また判別境界も単純な直線ではなく、3 次元空間内の複雑な曲面となる。

(2) 機械学習による近縁判別アルゴリズムを R ソフトウェア fraRF (flexible relationship analyzer by random forest) として実装した。シミュレーションにより、fraRF の性能評価を行い、近縁検出の精度および頑強性と必要なマーカー数・サンプル数を調査し、現実的なサンプリングによって十分な検出力を安定して得られることを確認した。また、既存の近縁判別手法として CKMRsim (尤度比法)<sup>2,14-15</sup> および COLONY (家系ベース尤度法)<sup>16-17</sup> との比較を行い、判別の特性的の違いを確認した<sup>18</sup>。CKMRsim は偽陽性の抑制を重視する閾値により判別するため近縁ペアを過小評価する傾向があるのに対し、fraRF と COLONY はデータとのフィッティングにより判別するため偽陽性・偽陰性を同等に含み、より多くのペアを近縁と判別する。特に COLONY は遠縁ペアを考慮しないため半兄弟を過大評価しやすい。これら手法の特性的の違いについてはなお検討が必要である。

fraRF は柔軟性の高い判別アルゴリズムであり、今回の親子・兄弟に加えて、祖父母-孫など様々な近縁関係の判別に拡張可能である。また計算速度も実用的であり、数千個体を用いた解析でも小型ワークステーションを用いて 1 日未満で完了する。これにより、資源量評価や種苗放流効果推定など幅広い適用が期待できる。

#### <引用文献>

- 1) ベゴンら、生態学(原著第 3 版)、京都大学学術出版 (2003).
- 2) Bravington MV. et al., Close-kin mark-recapture. Stat. Sci. 31:259-275 (2016).
- 3) Akita T., Statistical test for detecting overdispersion in offspring number based on kinship information. Popl. Ecol. 60:1-12 (2018).
- 4) Akita T., Nearly unbiased estimator of adult population size based on within-cohort half-sibling pairs incorporating flexible reproductive variation. Preprint (doi:

<https://doi.org/10.1101/422659>).

- 5) Bravington MV. et al., Close-kin mark-recapture: estimating the abundance of Bluefin tuna from parent-offspring pairs. *Nat. Commun.* 7:13162 (2016).
- 6) Hillary RM. et al., Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. *Sci. Rep.* 8:2661 (2018).
- 7) 關野&中道ら、DNA 親子判別における一塩基多型(SNPs)の有効性の検証、水産育種 46:79-85 (2017).
- 8) Akita T., Nearly unbiased estimator of contemporary effective mother size using within-cohort maternal sibling pairs incorporating parental and nonparental reproductive variations. *Heredity* 124:299-312 (2020).
- 9) Akita T., Nearly unbiased estimator of contemporary  $N_e/N$  based on kinship relationships, *Ecology and Evolution* 10:10342-10352 (2020).
- 10) Akita T., Estimating contemporary migration numbers of adults based on kinship relationships in iteroparous species, *Molecular Ecology Resources*, 22:8, 3006-3017 (2022).
- 11) Breiman L., Random Forests. *Machine Learning* 45:5-32 (2001).
- 12) Milligan BG., Maximum-likelihood estimation of relatedness. *Genetics* 163:1153-1167 (2003).
- 13) Waples RS. & Feutry P., Close-kin methods to estimate census size and effective population size. *Fish and Fisheries* 23:273–293 (2021).
- 14) Marshall TC. et al., Statistical confidence for likelihood based paternity inference in natural populations. *Molecular Ecology* 7: 639 - 65 (1998).
- 15) Anderson EC., CKMRsim. <https://github.com/eriqande/CKMRsim/> (2022).
- 16) Wang J. & Santure AW., Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics*, 181, 1579–1594 (2009).
- 17) Jones O. & Wang J., COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* 10, 551–555 (2010).
- 18) Tsukahara Y. et al. Comparison of kinship-identification methods for robust stock assessment using close-kin mark-recapture data for Pacific bluefin tuna. Preprint (doi: 10.22541/au.168267602.25993369/v1).

## 5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 4件/うち国際共著 0件/うちオープンアクセス 4件）

1. 著者名 Akita Tetsuya	4. 巻 22
2. 論文標題 Estimating contemporary migration numbers of adults based on kinship relationships in iteroparous species	5. 発行年 2022年
3. 雑誌名 Molecular Ecology Resources	6. 最初と最後の頁 3006 ~ 3017
掲載論文のDOI (デジタルオブジェクト識別子) 10.1111/1755-0998.13682	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 秋田 鉄也	4. 巻 -
2. 論文標題 家系生態学：ゲノム情報から近年の生態情報を抽出して希少種保全に貢献する	5. 発行年 2023年
3. 雑誌名 Jxiv	6. 最初と最後の頁 1 ~ 18
掲載論文のDOI (デジタルオブジェクト識別子) 10.51094/jxiv.279	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 -
1. 著者名 Naito Takuya, Nakayama Kouji, Takeshima Hirohiko, Hashiguchi Yasuyuki, Akita Tetsuya, Yamasaki Yo Y., Mishina Tappei, Takeshita Naohiko, Nagano Atsushi J., Takahashi Hiroshi	4. 巻 -
2. 論文標題 The detailed population genetic structure of the rare endangered latid fish akame Lates japonicus with extremely low genetic diversity revealed from single-nucleotide polymorphisms	5. 発行年 2023年
3. 雑誌名 Conservation Genetics	6. 最初と最後の頁 1 ~ 13
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s10592-023-01517-2	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Akita Tetsuya	4. 巻 10
2. 論文標題 Nearly unbiased estimator of contemporary Ne/N based on kinship relationships	5. 発行年 2020年
3. 雑誌名 Ecology and Evolution	6. 最初と最後の頁 10343-10352
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/ece3.6421	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 Akita Tetsuya	4. 巻 124
2. 論文標題 Nearly unbiased estimator of contemporary effective mother size using within-cohort maternal sibling pairs incorporating parental and nonparental reproductive variations	5. 発行年 2020年
3. 雑誌名 Heredity	6. 最初と最後の頁 299 ~ 312
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41437-019-0271-6	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 Yohei Tsukahara, Reiichiro Nakamichi, Aiko Matsuura, Tetsuya Akita, Atushi Fujiwara, Nobuaki Suzuki	4. 巻 -
2. 論文標題 Comparison of kinship-identification methods for robust stock assessment using close-kin mark-recapture data for Pacific bluefin tuna	5. 発行年 2023年
3. 雑誌名 AUTHOREA (preprint)	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) 10.22541/au.168267602.25993369/v1	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8件 (うち招待講演 5件 / うち国際学会 3件)

1. 発表者名 秋田 鉄也
2. 発表標題 遺伝情報から近年の有効集団サイズを推定する方法およびその代替指標
3. 学会等名 第70回日本生態学会大会 (招待講演)
4. 発表年 2023年

1. 発表者名 秋田 鉄也
2. 発表標題 家系生態学:ゲノム情報から近年の生態情報を抽出して希少種保全に貢献する
3. 学会等名 令和4年度 水産育種研究会 定例シンポジウム (招待講演)
4. 発表年 2022年

1. 発表者名 塚原 洋平、松浦 愛子、秋田 鉄也
2. 発表標題 Close-kin mark-recapture method for estimating spawning stock biomass: theory and its preliminary application to Pacific bluefin tuna
3. 学会等名 第38回個体群生態学会大会（招待講演）（国際学会）
4. 発表年 2022年

1. 発表者名 水野 ひなの、中山 耕至、秋田 鉄也、橋口 康之、大杉 奉功、南野 洋孝、鈴木 厚裕、武島 弘彦
2. 発表標題 Detailed kinship estimation for detecting bias among breeding families in a reintroduced population of endangered bagrid catfish, <i>Tachysurus Ichikawai</i>
3. 学会等名 第38回個体群生態学会大会（招待講演）（国際学会）
4. 発表年 2022年

1. 発表者名 野田 祥平、秋田 鉄也、武島 弘彦、佐藤 拓哉
2. 発表標題 Causes and consequences of life-history variation in wild populations: relating individual fitness to population dynamics and local adaptation using a pedigree analysis
3. 学会等名 第38回個体群生態学会大会（招待講演）（国際学会）
4. 発表年 2022年

1. 発表者名 奥村 健太、野原 健司、井戸 啓太、坂本 正吾、稲川 崇史、沖津 二郎、松崎 厚史、佐々木良浩、大杉 奉功、秋田 鉄也、田原 大輔、武島 弘彦
2. 発表標題 ダム湖の外來魚オオクチバスにおける新たな駆除効果の評価方法の開発；大規模DNA分析を用いた高精度近親関係解析による個体数推定法の適用
3. 学会等名 日本緑化工学会・日本景観生態学会・応用生態工学会3学会合同大会
4. 発表年 2022年



1. 発表者名 野田 祥平、秋田 鉄也、武島 弘彦、佐藤 拓哉
2. 発表標題 ゲノムワイドSNPを用いた家系解析手法の確立とアマゴ野生集団への適用
3. 学会等名 生物多様性のDNA情報学 2022
4. 発表年 2022年

1. 発表者名 秋田 鉄也
2. 発表標題 近親標識法による集団間の移動個体数推定
3. 学会等名 第37回個体生態学会大会
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
研究 分担 者	中道 礼一郎  (Nakamichi Reichiro)  (70401255)	国立研究開発法人水産研究・教育機構・水産資源研究所(横浜)・主任研究員   (82708)	

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1件

国際研究集会 Ecological inference based on kinship assignment: a new approach to obtain contemporary information from genetic data	開催年 2022年～2022年
---	--------------------

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