

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

Ecological inference based on kinship assignment: a new approach to obtain contemporary information for conservation purposes

秋田鉄也（水産研究・教育機構）*

プレプリント (February 10, 2023)

Abstract

In this report, we review the close-kin mark-recapture (CKMR) method, which is recently developed for estimating adult population size and life-history parameters in the context of fisheries stock assessment. The rationale is that kin pairs in the sample are less likely to be observed in larger populations; thus, the number of kin pairs may reflect the number of adults in the population. Here, we provide an overview of the method based on parent–offspring pairs, which is used in most of the literature on practical examples of the CKMR method. First, we explain the simple model using analogies to the classical mark-recapture method. Next, we show the conditions under which the complex model is needed, instead of the simple model. Then, we review literatures that apply the CKMR method using parent–offspring relationships. Finally, we discuss the utility and uncertainty of the method. Further rapid progress of the CKMR method is expected from the broad application of other types of kinship relationships, including half-sibling pairs.

*連絡先：236-8648 神奈川県横浜市金沢区福浦 2-12-4 国立研究開発法人 水産研究・教育機構 水産資源研究所 漁業情報解析部 秋田鉄也 Tel 045-788-7919 akita.tetsuya81@fra.go.jp

Key words: close-kin mark-recapture; population size estimation; parentage

はじめに

近親情報の利用は、育種における遺伝率や育種価の推定のみならず、社会性、移動や拡散、繁殖成功率、有効繁殖サイズ、といった生態学的特性の理解に有用であることから、希少種保全の文脈においても重要である^{1,2}。近親情報を野外のサンプルに適用して解析する場合、マイクロサテライトマーカーを用いて推定される親子関係の利用が主流であったが³、近年の技術革新により一塩基多型 (Single Nucleotide Polymorphism、SNP) マーカーが大量かつ安価で取得できるようになり、解析に使用可能な情報量が飛躍的に増加した。その結果、(半)兄弟関係やより複雑な近親関係を野生集団に適用することが現実的になってきた。したがって、近親情報を用いて生態学的な問いにチャレンジする試み(著者はこの方法論を「家系生態学」と呼んでいる)は、今後飛躍的に進むことが期待される。

このような背景のもと、近親標識再捕 (Close-Kin Mark-Recapture, CKMR) 法と呼ばれる近親関係を用いた個体数推定法が水産資源分野において出現し⁴⁻¹⁰、資源評価における実践や他分野への応用が期待されている。CKMR法は、標識再捕法とのアナロジーとして捉えることが可能である。すなわち、近親関係の発見を「標識個体の再捕獲」とみなすことで、未知の集団サイズや関連するパラメータを推定する手法である。一般に、近親関係は集団内の成熟個体の数が少ない程見つけやすい。そのため、希少種集団を対象とした場合、比較的少ないサンプル数から多数の近親関係が見出されることが期待され、推定精度も高くなると考えられる。CKMR法は、標識再捕法とは異なり物理的なマーキングが不要であり、一度きりのサンプリングで十分であるという点からも、希少種集団を解析対象にしやすい手法と言える。しかしながら、水産資源評価会議で用いられる会議文書を介して知見が蓄積されたこと、2016年に刊行された最初の原著論文⁷の読み解きには数理統計の素養が必須であることなどから、CKMR法の理解が国内で広がっているとは言えない(著者の知る限り、日本語での解説は入江(2016)¹¹のみである)。

本報告では、親子関係から成熟個体の数を推定するCKMR法に注目して、その方法論や応用例について概説することを目的とする。まずはじめに、標識再捕法とのアナロジーを用いて、母子関係に基づくCKMR法の理論を平易に説明する。次に、親子関係に基づくCKMR法について単純なモデルと複雑なモデルを紹介する。続いて、水圏生物を対象にしたCKMR法の応用例を紹介する。最後に、半兄弟関係を用いたCKMR法や関連す

る不確実性について整理する。

母子関係に基づく母親個体数の推定：標識再捕法とのアナロジー

標識再捕法による個体数推定

親子関係から成熟個体数を推定する CKMR 法は、標識再捕法による個体数推定のアナロジーとして語られることが多い⁷。それを実感するために、標識を付けた個体を1度だけサンプルする標識再捕法 (single-tag recovery) による個体数の推定式を導出し、その式と全く同じ形の式を用いることで、母子関係から成熟雌個体数を推定できることを示す。

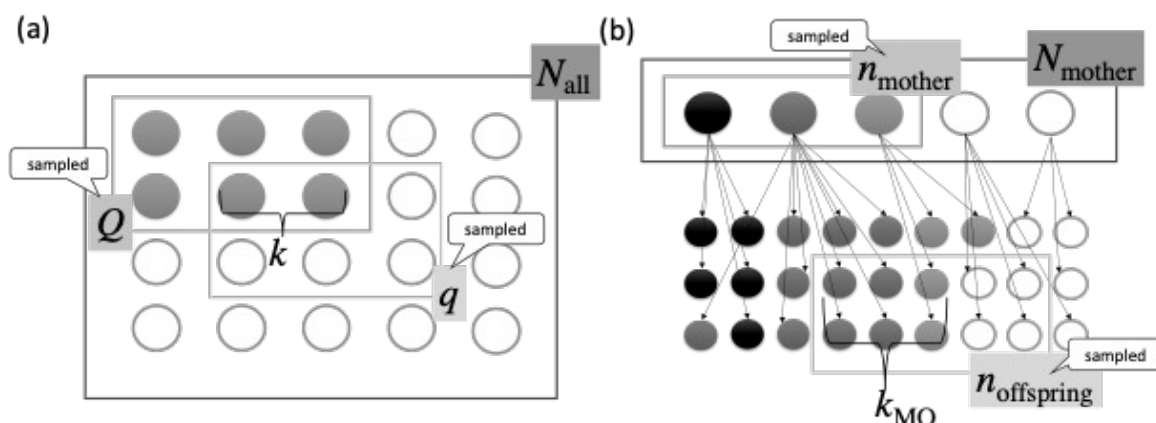


図 1: (a) 標識再捕法による個体数推定。 N_{all} : 個体数、 Q : 標識個体数、 q : 再捕獲数、 k : 再捕獲された標識個体数。(b) 母子関係に基づく母親個体数の推定。矢印は母子関係を表す。 N_{mother} : 母親個体数、 n_{mother} : サンプルされた母親個体数、 $n_{\text{offspring}}$: サンプルされた子供個体数、 k_{MO} : 見出された親子ペアの数。

調査域に生息する対象種の個体数 N_{all} を推定することを考えよう。 N_{all} 個体のうち Q 個体を捕獲し、標識を付けた後に調査域に戻す。標識個体とそうでない個体が十分混じったのちに q 個体を再捕獲する。再捕獲した n 個体のうち、 k 個体が標識付きの個体であるとしよう (図 1a)。このとき、下記のような関係性が期待されることは直感的にも明らかであろう。

$$\frac{Q}{N_{\text{all}}} = \frac{E[k]}{q} \quad (1)$$

個体数 N_{all} は未知のパラメータであり推定したい値である。標識個体数 Q と再捕獲数 q

は（労力を気にしなければ）コントロール可能な既知のパラメータである。再捕獲に含まれる標識個体数 k は確率変数であるためその期待値 $E[k]$ が式に現れている。すなわち、試行を繰り返すごとに k の値は変化するだろうが、その期待値を再捕獲数 q で割った値は、標識個体の割合と等しくなることを式 (1) は意味している。さて、実際に標識再捕実験を実施したとして、再捕獲に含まれていた標識個体数の観測値を k_{obs} としよう。式 (1) の $E[k]$ を k_{obs} で代用して式変形することで、Petersen–Lincoln 推定量と呼ばれる N_{all} の推定量が得られる¹²（ハットは推定量を意味する）。

$$\hat{N}_{\text{all}} = \frac{qQ}{k_{\text{obs}}} \quad (2)$$

母子関係に基づく母親個体数の推定

次に、調査域に生息する対象種の成熟雌（母親と呼ぶ）の個体数 N_{mother} を推定することを考える。調査域には未成熟個体（子供と呼ぶ）が複数おり、子供の母親は N_{mother} に必ず含まれているとしよう。このような状況下で、母親を n_{mother} 個体、子供を $n_{\text{offspring}}$ 個体サンプルする。サンプルされた母親個体と子供個体の全ペア ($n_{\text{mother}} \times n_{\text{offspring}}$) について母子関係を調べた結果、 k_{MO} ペアの母子関係が発見されるとしよう。このことを模式的に示した図 1b から明らかのように、下記の関係が成立する。

$$\frac{n_{\text{mother}}}{N_{\text{mother}}} = \frac{E[k_{\text{MO}}]}{n_{\text{offspring}}} \quad (3)$$

この式は、使用している文字の意味は異なるが、標識再補法で示した式 (1) と全く同一である。すなわち、標識再補法と比較すると、サンプルされた母親の数 n_{mother} が標識された個体数 Q に、サンプルされた子供の数 $n_{\text{offspring}}$ が再捕獲された個体数 q に、発見された母子ペア数 k_{MO} が再捕獲された個体に含まれる標識個体の数 k にそれぞれ相当する。上記の式でピンと来なければ、下記のように考えることも可能である。それぞれランダムにサンプルされた母親と子供が真の母子関係にある確率は、仮定より $1/N_{\text{mother}}$ である。真の母子かどうか比較すべきペア数の総数は $n_{\text{mother}} \times n_{\text{offspring}}$ である。したがって、発見される真の母子ペア数の期待値 $E[k_{\text{MO}}]$ は $(1/N_{\text{mother}}) \times (n_{\text{mother}} \times n_{\text{offspring}})$ と表せ、これを変形することで式 (3) を得る。標識再補法と同様に、式 (3) から下記のような推定量が得られる。

$$\hat{N}_{\text{mother}} = \frac{n_{\text{mother}} n_{\text{offspring}}}{k_{\text{MO,obs}}} \quad (4)$$

親子関係に基づく親個体数の推定

単純なモデル

母親個体数を推定した場合と同様の状況を仮定して、雌雄合わせた成熟個体（親と呼ぶ）の数 (N_{adult}) を推定する方法を考えよう。子供の父親と母親は N_{adult} に必ず含まれているとするとし、親を n_{parent} 個体、子供を $n_{\text{offspring}}$ 個体サンプルする。サンプルされた親個体と子供個体の全ペア ($n_{\text{parent}} \times n_{\text{offspring}}$) について親子関係を調べた結果、 k_{PO} ペアの親子関係が発見されるとする。性比が集団中の全親個体とサンプル集団で同じだと仮定すると、下記の関係が成立する。

$$\frac{n_{\text{parent}}}{N_{\text{adult}}} = \frac{E[k_{\text{PO}}]}{2n_{\text{offspring}}} \quad (5)$$

サンプルされた子供の数を2倍するのは、子供から見て親は父と母の2個体が集団中に存在するからである。ランダムにサンプルされた親個体と子供が真の親子関係にある確率は $2/N_{\text{adult}}$ となる。より詳細な導出は付録に記した。このような関係式が成立する状態について、以降「単純なモデル」と呼ぶ。この関係式から、よく知られた推定量⁷が導出される。

$$\hat{N}_{\text{adult}} = \frac{2n_{\text{parent}}n_{\text{offspring}}}{k_{\text{PO,obs}}} \quad (6)$$

変動係数による精度評価と不偏性

各親が残す子供の数はポアソン分布に従うという仮定を与えることで、 \hat{N}_{adult} の変動係数 (Coefficient of Variation、CV) が近似的に得られる (導出は付録を参照)。

$$CV[\hat{N}_{\text{adult}}] \approx \sqrt{\frac{N_{\text{adult}}}{2n_{\text{parent}}n_{\text{offspring}}}} \quad (7)$$

推定したい親個体数の大体のオーダーがわかっている場合、この変動係数を用いることで、どのくらいのサンプル数が必要であるのかを事前に見積もることが可能である。例えば、CV=20%の精度で N_{adult} を推定する場合を考えよう。簡単のため、親個体と子供個体は同数サンプルするとする ($n_{\text{parent}} = n_{\text{offspring}}$)。親個体数が1000–10000の範囲であると事前に見積もられている場合、CV=20%の精度を保つのに必要なサンプルサイズ (親と子を合算) の範囲は、式(7)から約224–708とわかる。一方、親個体数に関する事前の情報がない場合もあるだろう。そのような場合、下記のような変動係数の推定量が有用である。

$$CV[\hat{N}_{\text{adult}}] \approx \sqrt{\frac{1}{k_{\text{PO,obs}}}} \quad (8)$$

この推定量を用いると、見出された親子ペア数が 11 ペア、25 ペアの時、親個体数の推定値の変動係数はそれぞれ 30%、20%と見積もることができる。すなわち、見出された親子ペア数が二桁に満たない場合、推定された親個体数の値の信頼性は高くないと考えて良いだろう。

紹介した変動係数やその推定値は、各親子ペアが独立に見つかる（すなわち、各親が残す子供の数はポアソン分布に従う）という仮定がある点に注意しよう。言い換えると、真の変動係数の値はこれらよりも大きいと考えるべきである⁷。一方、式(4)で表される N_{adult} の推定量は、子供の残し方に依らず、不偏性の観点からは安定していることが知られている。不偏性があるとは、推定量の期待値が真の値と一致することを意味する。これは一体どういうことだろうか？繁殖力が母親間や父親間でばらついている状況を考えよう。親の年齢と繁殖力に相関がある場合や、単に運不運で多くの親個体が子供を残せない場合などが該当するだろう。極端なケースとして、全ての子供が 1 個体の父親（もしくは母親）に由来する場合は考えられる。そのような場合、次世代に強く貢献した親をサンプルできたかどうかで、見つかる親子ペア数は確率的に大きく揺らぐ。言い換えると、各親子ペアは独立とは言えなくなる。そのような極端な状況下でも、任意の親子ペアが真の親子である確率は $2/N_{\text{adult}}$ であることに変わりなく、サンプル数の増加に伴って推定値の期待値は真の値に漸近するのである。

一方、サンプル数が極端に少ない場合は、 \hat{N}_{adult} の不偏性が保たれず過大推定されることが知られている。これは式(2)で紹介した標識再補法における Petersen–Lincoln 推定量においても同様であり、Chapman 型の補正をすることで不偏性が担保される¹²。親子関係に基づく個体数推定においても同様に、

$$\hat{N}_{\text{adult}} = \frac{2n_{\text{parent}}n_{\text{offspring}}}{k_{\text{PO,obs}} + 1} \quad (9)$$

と補正することで、サンプル数が少ない場合に生じるバイアスが劇的に緩和される^{13,14}（類似の導出は Akita (2020)¹⁵ を参照）。

単純なモデルが成立するための条件

ここまで、単純なモデルが成立すると仮定して、親個体数の推定量とその推定精度の評価法を紹介してきた。ここでは Bravington ら (2016)⁷ に基づき、単純なモデルが用いた仮定について整理し、どのような条件下では仮定が満たされないのかを具体的に紹介する。なお、ここでは説明を簡単にするために、再び母子関係から母親個体数を推定するモデルを考える。また、サンプリングは致命的とする。繁殖に貢献する親個体やその年に生ま

表 1: 単純なモデルに基づく式 (4) が利用可能となる主な条件

繁殖成功度が年齢 に依らず一定	母親のサンプルされやすさ が繁殖成功度と無関係	死亡率が年齢に 依らず一定	複数年サンプ リング
Yes	—	—	—
No	Yes	Yes	—
No	Yes	No	No*

「—」は任意であることを意味する。

*: 繁殖期間終了後すぐに母親をサンプルすることが求められる。

れた子供を翌年以降にサンプリングすることを「複数年サンプリング」と呼び、そのようなサンプルデザインもここでは想定する。一般に、推定対象集団の個体数が大きい場合、必要となるサンプル数が膨大になることから複数年サンプルリングが求められる場合が多い。

最も制約の強い仮定は、繁殖成功度が母親間で一定であるという条件である。言い換えると、母親あたりの子の数の期待値が、母親間で同じであるという条件である。この仮定の下では、ほとんどの場合において単純なモデルが利用可能である。複数年サンプルリングを実施する場合であっても、任意の親子候補ペアが真の親子である確率は $2/N_{\text{adult}}$ である。例外の1つとして、「サンプルされやすさが遺伝する場合」などが考えられるが、詳細には立ち入らない。

すでに述べた通り、繁殖成功度が一定でなくても単純なモデルは利用可能である。その場合、母親のサンプルされやすさが繁殖成功度と無関係であることが要請される。したがって、年齢（体長）が繁殖成功度と相関する種については注意が必要である。例えば、サンプリングが特定の年齢（体長）の母個体に偏る場合や、繁殖期間終了後にサンプリングされるまでの期間（複数年サンプリングも含む）の死亡率が年齢（体長）に依存する場合などである。後者の制約は、母親の繁殖成功度の分布が変化しないという仮定に由来する。漁業対象種であれば、漁獲における体長の選択性 (selectivity) を考慮する必要があり、親魚の死亡率を一定とする仮定には注意を要する。単純なモデルが利用可能となる場合の諸条件を表1にまとめた。

複雑なモデル：サンプルの年齢情報が求められる

ここでは、表1に示したような諸条件が成立しない場合を考える。以下、繁殖成功度が年齢で決まる ($\beta(a)$, a : 年齢) と仮定し、サンプルされた個体の年齢が既知であるという前提のもと、単純なモデルに替わる新しいモデルを構築する。引き続きサンプリングは致死的とし、母子関係から母親個体数を推定することを考える。具体的な例として、2歳で100%性成熟し、2歳と3歳の時に繁殖し、3歳の間に100%死亡するような仮想的な生物を想定する。そのような生物集団を対象に、2020年と2021年の2年間にサンプリングを実施し、2020年における繁殖直後の母親個体数を推定することを考えよう。図2にサンプリングスキームを示す。2020年の繁殖期間後に生息する2歳および3歳の母親個体数を $N_{M,a=2}^{(2020)}$ および $N_{M,a=3}^{(2020)}$ などと記載する。単純なモデルでは上記を合算したものを直接推定していたが、ここでは年齢ごとに区別して立式をする。母親サンプルは、2020年の2歳および3歳個体 ($n_{M,a=2}^{(2020)}$ および $n_{M,a=3}^{(2020)}$) と2021年の3歳個体 ($n_{M,a=3}^{(2021)}$) を用いる。子供サンプルは、2020年の0歳個体 ($n_{O,a=0}^{(2020)}$) と2021年にサンプルされた1歳の個体 ($n_{O,a=1}^{(2021)}$) を用いる (Mは母を、Oは子供を意味している)。

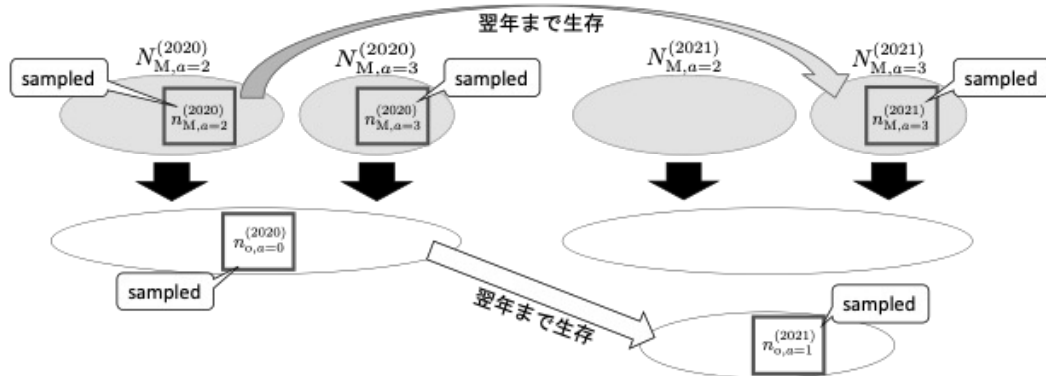


図2: 2020年の母親個体数を推定するためのサンプリングスキームの一例。影付きの楕円が母親を、白抜き楕円が子供を意味している。2020年に2歳と3歳の母と0歳の子供を、2021年に3歳の母親と1歳の子供をサンプルしている。記号の意味は本文中を参照。

2020年に繁殖した2歳の母親集団からサンプルされたある個体と、2020年に生まれ2020年に子供集団からサンプルされたある個体が真の母子関係である確率を考えよう。2歳と3歳では、母親の繁殖成功度が異なることから、その確率は以下のように与えられる。

$$\frac{\beta(a=2)}{\beta(a=2)N_{M,a=2}^{(2020)} + \beta(a=3)N_{M,a=3}^{(2020)}} \quad (10)$$

すなわち、3歳の母親と比べて、2歳の母親の繁殖成功度が相対的に小さい場合、この確率は小さい値となる。2020年の母親個体数で重み付けした平均繁殖成功度を $\bar{\beta}$ とすると、

この確率は下記のように書くことができる。

$$\frac{\beta(a=2)}{\bar{\beta}N_M^{(2020)}} \quad (11)$$

分母の $N_M^{(2020)}$ は、2020年における繁殖直後の母親個体数を全ての年齢で合算した値であり、推定しようとしている値である。この確率は、2021年に繁殖した3歳の母親集団からサンプルされた個体や、2020年に生まれ2021年に子供集団からサンプルされた個体を用いた場合も成立する。これは、年齢が同じであれば繁殖成功度も同じであるという前提があるからであり、サンプル数が足りなければ翌年以降に追加しても良いわけである。以上より、比較可能なペア数は $(n_{O,a=0}^{(2020)} + n_{O,a=1}^{(2021)}) \times (n_{M,a=2}^{(2020)} + n_{M,a=3}^{(2021)})$ となり、2020年の2歳の母親による繁殖と関連する母子ペア数の期待値は下記のようになる。

$$E[k_{MO,a=2}] = \frac{\beta(a=2)}{\bar{\beta}N_M^{(2020)}} (n_{O,a=0}^{(2020)} + n_{O,a=1}^{(2021)}) (n_{M,a=2}^{(2020)} + n_{M,a=3}^{(2021)}) \quad (12)$$

同様に、2020年の3歳の母親による繁殖と関連する母子ペア数の期待値は下記のようになる。

$$E[k_{MO,a=3}] = \frac{\beta(a=3)}{\bar{\beta}N_M^{(2020)}} (n_{O,a=0}^{(2020)} + n_{O,a=1}^{(2021)}) n_{M,a=3}^{(2020)} \quad (13)$$

この場合、3歳の間には全ての母親が死亡することから、2021年の母親サンプルが使えないことに注意しよう。以上をまとめると、以下のような推定量を得ることができる¹⁴。

$$\hat{N}_M^{(2020)} = \frac{n_{O,a=0}^{(2020)} + n_{O,a=1}^{(2021)}}{\bar{\beta}k_{MO,obs}} \left(\beta(a=2) (n_{M,a=2}^{(2020)} + n_{M,a=3}^{(2021)}) + \beta(a=3)n_{M,a=3}^{(2020)} \right) \quad (14)$$

$k_{MO,obs}$ は、見出された母子ペアを全て合算した値である。式(14)を利用する場合、各年齢ごとの母親の繁殖成功度 ($\beta(a)$) や、2020年における年齢構成で重み付けされた平均繁殖成功度 ($\bar{\beta}$) の情報が必要である。前者については、情報がなければ体長式などから推定する必要がある。後者については、集団と同様の年齢構成の比率でサンプルできると仮定すると、下記のように推定することが可能である¹⁴。

$$\bar{\beta} = \frac{\beta(a=2)n_{M,a=2}^{(2020)} + \beta(a=3)n_{M,a=3}^{(2020)}}{n_{M,a=2}^{(2020)} + n_{M,a=3}^{(2020)}} \quad (15)$$

単純なモデルが仮定した状況(表1)を満たさない場合でも、親の年齢に応じて個別にモデル化することでCKMR法は利用可能であることを示した。ただし、サンプルの年齢情報が必要となる。以降、このようなモデルを「複雑なモデル」と呼ぶ。しかしながら、偏りの少ない推定の実施には多くの仮定が必要であり、これらについては本報告の後半でまとめた。また、親の年齢情報がない場合の取り扱い方法については、より複雑なモデル

ングが必要となるため、詳細を示した原著^{7,16}にあたって欲しい。紹介した方法では推定量を直接導出したが、真の母子となる確率を元に尤度関数を定義して、最尤推定することも可能である⁷。母親集団だけでなく雌雄合わせた親集団を考える場合や、非致死的なサンプリングの場合についても、これまで紹介した方法を応用することで容易に理解できるだろう。

親子関係に基づく CKMR 法の実例

ここでは、親子関係に基づいた CKMR 法の実例として、6 報の文献を簡単に紹介する。前半の 4 報が河川や湖を対象にしたもの、後半の 2 報が海産資源を対象にしたものであり、基本的には後半に進むにつれてモデルが複雑になっていく。

Prystupa et al. (2021)¹³ は、カナダの河川に生息するキタカワヒメマス *Thymallus arcticus* を対象に、単純なモデルを用いて親魚尾数の推定した。親子判別ソフトウェアを複数用いても見出される親子ペア数はほぼ同じであり、したがって推定親魚尾数の結果も同様であった（約 1800 尾、CV=0.16）。加えて、同河川における有効集団サイズ (N_e) を LD 法^{17,18} を用いて推定し、有効集団サイズと親魚尾数の比 (N_e/N_{adult}) について約 0.41 と比較的高めの推定値を得た。親魚尾数推定に用いたサンプル数は 967 個体、親子判定に用いた遺伝子マーカー数はマイクロサテライト 38 座、見出された親子ペア数は約 38 ペアであった。

Wacker et al. (2021)¹⁹ は、ノルウェーの河川を遡上するタイセイヨウサケ *Salmo salar* の親魚尾数の推定について、産卵調査によるカウント（従来法と呼ぶ）・産卵調査時に親魚から組織をサンプルする CKMR 法（CKMR 法と呼ぶ）・PIT タグ装着と同時に組織をサンプルする CKMR 法（PIT-CKMR 法と呼ぶ）の 3 つの手法を比較検討した。その結果、推定された親魚尾数の大小は、従来法（439 尾）< CKMR 法（621 尾、CV=0.12）< PIT-CKMR 法（1501 尾、CV=0.22）となった。サンプリングの偏りを考慮して、体長から親の繁殖成功度を 2 種類に区別した複雑なモデルを用いても、手法による差が大きく縮まることはなかった。手法間の推定結果の違いについて、各手法が対象とする親魚集団が時空間的に一致しなかったことに起因すると著者らは考察している。親魚尾数推定に用いたサンプル数は 391 個体、親子判定に用いた遺伝子マーカー数は 164 SNP、見出された親子ペア数は 60 ペア（CKMR 法）および 20 ペア（PIT-CKMR 法）であった。

Marcy-Quay et al. (2020)²⁰ は、ニューヨーク州の湖に生息するカワマス *Salvelinus fontinalis* の親魚尾数の推定について、単純なモデルを用いた CKMR 法の推定結果を標識再捕法による推定結果と比較した。サンプリングは親子区別することなく、PIT タグを

用いた非致死的な方法で3年間実施された。タグを装着した翌年や翌々年に繁殖する場合もあるため、モデル式には生存率に関するパラメータも含まれ、親魚尾数と一緒に推定された。CKMR法で推定された生存率および5年分の親魚尾数(約1500尾)は、標識再捕法による推定結果とほぼ同じであった。この結果から、CKMR法の頑健性が裏付けられたと著者らは結論づけている。親魚尾数推定に用いたサンプル数は384個体、親子判定に用いた遺伝子マーカー数はマイクロサテライト44座、見出された親子ペア数は72ペアであった。

Ruzzante et al. (2019)¹⁴は、カナダの7つの河川に生息するカワマスを対象に、複雑なモデルを用いたCKMR法による親魚尾数の推定結果を標識再捕法による推定結果と比較した。CKMR法を用いて2015年の親魚尾数を推定するため、親を2014年から2017年まで毎年サンプルし、子供を2017年にサンプルした。サンプリングは非致死的に実施された。推定に必要な年齢ごとの繁殖成功率や生存率は、既知のデータの値を用いた。7河川のうち6河川において、CKMR法と標識再捕法で推定結果の差異がほとんど見られなかったことから、CKMR法の頑健性が裏付けられたと著者らは結論づけている。親魚尾数推定に用いたサンプル数は2,041個体、親子判定に用いた遺伝子マーカー数はマイクロサテライト33座、見出された親子ペア数は159ペアであった。

Trenkel et al. (2022)¹⁶は、イベリア半島の北岸ビスケー湾に生息するイボガンギエイ *Raja clavata* を対象に、やや複雑なモデルを用いたCKMR法による親魚尾数推定を実施した。主に、商業漁獲された個体から得られた10年分のサンプルから親子判別をし、2013年の親魚尾数を推定した。漁獲制限のため若い個体の利用が困難であり、漁獲対象の高齢個体を利用せざるを得なかったことから、サンプルの年齢情報には大きな不確実性が見込まれた。そのため、単純なモデルに体長と年齢の関係を確率的に組み込んだものが使用された。一方、繁殖成功率については各年齢で一定と仮定された。推定された2013年の親魚尾数は約135,000尾(CV=0.19)であり、2017年までの年トレンドについても同時に推定された。また、見出された複数の近親ペアの空間的分布から集団構造の存在が示唆され、これらを一緒にして推定すると親魚尾数の年トレンドの方向を見誤る可能性が指摘された。親魚尾数推定に用いたサンプル数は6,555個体、親子判定に用いた遺伝子マーカー数は3,668 SNP、見出された親子ペア数は73ペアであった。

最後に、大規模な集団を対象にCKMR法を応用した最初の例として、Branvington et al. (2016)⁶によるミナミマグロを対象にした成果を紹介する。ミナミマグロは集団構造がなく産卵場所が単一と考えられており、漁獲を通じて親魚と幼魚のサンプルが多数利用可能であるという利点がある。2006年から2010年にかけて漁獲された親魚5,755尾と3歳

の幼魚 7,448 尾のサンプルから、2002 年から 2010 年までの親魚尾数を推定された。年齢情報については、Trenkel et al. (2022)¹⁶ と同様に、体長と年齢の関係が確率的に組み込まれたモデルが使用された。産卵場滞在時間に 1 日あたりの卵放出量を乗じたものを雌の繁殖成功度とし、雄の繁殖成功度については年生存率（一定と仮定）や親魚尾数と一緒に推定された。推定された親魚尾数は、CKMR 法を適用しないで推定された資源評価結果よりも大きい値になった。加えて、雌は高齢魚ほど体重当たりの繁殖成功度が大きいという、繁殖生態に関する新たな知見も得られた。親魚尾数推定に用いたサンプル数は 13,203 個体、親子判定に用いた遺伝子マーカー数はマイクロサテライト 25 座、見出された親子ペア数は 45 ペアであった。

親子関係に基づく CKMR 法の留意点

ここでは、CKMR 法の利用に関する留意点について、これまでの説明と重複する部分もあるが説明する。

CKMR 法は親個体数の推定にフォーカスした手法であるが、推定されるのは繁殖直後の親個体数である。一方、繁殖生態に基づいた適切なモデル選択がなされていれば、親や子供を繁殖期間直後にサンプルする必要はない。しかしながら、推定したい年の繁殖期間の終了を待たずに親を（致命的に）サンプルすると、親子関係は見つかりにくくなり親個体数を過大推定してしまうだろう。

サンプリングについては、タイミングだけでなく空間的な配慮も重要である。CKMR 法は標識再補法と同様に、十分に個体が混合していることを前提としている。混合が不十分な状態でサンプリングを実施すると、親個体数の推定結果は偏るだろう。一方、魚類のように産卵場が複数あり、その産卵場由来の子供やそこで産卵した親がサンプルできれば、見出された親子ペアの空間的なパターンから集団構造の理解が進むかもしれない²¹。

ここまで、親子判別における不確実性については言及してこなかった。野生集団を対象にする場合、近親関係の判別は遺伝マーカーを用いて実施される。親子関係に限れば、前述した実例と同程度のマーカー数が確保できれば、親子判別に伴う不確実性は少ないと言える³。CKMR 法の応用において、親子かどうか比較すべきペア数は、実際に見つかる親子ペア数に比べてずっと多い。そのため、親子判別ソフトウェアの利用においては、偽陽性となる確率を下げる設定が必須となる^{6,7}。また、マーカー数が少ないと兄弟関係と親子関係の区別が難しいが、サンプルの年齢やサイズ情報を併用することで、近親判別の間違いを減らせる場合がある。

CKMR 法の適用可能性や有効性については、対象種の生活史によって決定される⁷。

生活史の中で単為生殖や一卵性の繁殖をする種は、個体を遺伝的に識別できないために対象外となる。一回繁殖 (iteroparity) の種を対象にする場合、特に魚類では親のサンプルしやすさが繁殖経験の有無で異なる可能性が高いため、注意が必要である。子供が若齢時にサンプルできない種では、ある程度大きくなってからサンプルせざるを得ない。その場合、推定される親個体数は、サンプルされた子供が産まれた年の値になり、直近年の値とは離れてしまう。親の個体数が非常に大きい種については、必要なサンプルサイズも非常に大きくなるため現実的ではないだろう。反対に、親の個体数が極めて小さい種については、サンプルサイズは小さくても集団全体から大きな割合をサンプルする必要がある。

おわりに

本報告では、CKMR 法による親個体数推定について、その推定ロジックと実例を親子関係の利用に限定して紹介した。特に、単純なモデルが利用可能な条件や、複雑なモデルが必要となる条件の説明に力点を置いた。これを一言でまとめると、複数年のサンプルを利用してある年の親個体数を推定する場合は複雑なモデルが望ましく、その際はサンプルの年齢情報や年齢ごとの繁殖成功度の情報が必要となる、ということである。十分な情報がないために、単純なモデルを利用する場合もあり得るだろう。その場合は、どのような仮定の上に推定結果が得られたのかに留意する必要がある。

本報告では、半兄弟関係に基づく CKMR 法については述べなかった。この方法は、基本的にはコホート間で見つかる半兄弟ペア数から親個体数を推定するものであり、いくつか実例も存在する^{9,22}。親子関係を利用する場合と比較すると、親個体のサンプルが不要なこと、親個体数に加えて親個体の生存率をより直接的に推定できることが主な利点である⁷。加えて、コホート内の半兄弟関係から有効繁殖サイズ (N_b) が推定でき²³、この値から N_e を推定したり²⁴、コホート間の半兄弟ペアもしくは親子ペアと合わせて N_e/N_{adult} を推定することも可能である^{10,25}。一方、半兄弟ペアの判定は親子ペアよりも困難であるという大きな障壁が存在するが、詳細の説明については別の機会に譲りたい。

ポスト 2020 生物多様性枠組では、野生生物の有効集団サイズのモニタリングについて言及されている²⁶。我が国においても、有効集団サイズ推定のために希少種のゲノム情報蓄積が加速度的に進むことが見込まれる。CKMR 法も含め、近親関係を希少種保全に有効利用する技術について、今後の発展を期待する。

謝辞

本報告は、令和4年12月10日に開催された水産育種研究会シンポジウムにおける講演内容に基づいている。CKMR法の理論的側面について、多くのご指導をいただいた東京海洋大学の北門利英教授、東京大学の入江貴博助教、丹羽智洋博士に感謝する。本研究は、JSPS 科研費 19K06862 および水産研究・教育機構水産資源研究センター運営費交付金プロジェクト「ゲノム情報を用いた資源学的研究の統一的基盤の構築およびガイドラインの作成」の助成を受けた。

References

- [1] Fred W. Allendorf, Paul A. Hohenlohe, and Gordon Luikart. Genomics and the future of conservation genetics. *Nature Reviews Genetics*, 11(10):697–709, 2010.
- [2] Stephanie J. Galla, Liz Brown, Yvette Couch-Lewis (Ngāi Tahu: Te Hapū o Ngāti Wheke, Ngāti Waewae), Ilina Cubrinovska, Daryl Eason, Rebecca M. Gooley, Jill A. Hamilton, Julie A. Heath, Samantha S. Hauser, Emily K. Latch, Marjorie D. Matocq, Anne Richardson, Jana R. Wold, Carolyn J. Hogg, Anna W. Santure, and Tammy E. Steeves. The relevance of pedigrees in the conservation genomics era. *Molecular Ecology*, 31(1):41–54, 2022.
- [3] Sarah P. Flanagan and Adam G. Jones. The future of parentage analysis: From microsatellites to SNPs and beyond. *Molecular Ecology*, 28(3):544–567, 2019.
- [4] Hans J Skaug. Allele-sharing methods for estimation of population size. *Biometrics*, 57(3):750–756, 2001.
- [5] Rasmus Nielsen, David K Mattila, Philip J Clapham, and Per J Palsbøll. Statistical approaches to paternity analysis in natural populations and applications to the north atlantic humpback whale. *Genetics*, 157(4):1673–1682, 2001.
- [6] Mark V. Bravington, Peter M. Grewe, and Campbell R. Davies. Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. *Nat Commun*, 7:13162, 2016.
- [7] Mark V Bravington, Hans J Skaug, and Eric C Anderson. Close-kin mark-recapture. *Stat Sci*, 31(2):259–274, 2016.

- [8] Hans J Skaug. The parent–offspring probability when sampling age-structured populations. *Theor popul biol*, 118:20–26, 2017.
- [9] R. M. Hillary, M. V. Bravington, T. A. Patterson, P. Grewe, R. Bradford, P. Feutry, R. Gunasekera, V. Peddemors, J. Werry, M. P. Francis, and et al. Genetic relatedness reveals total population size of white sharks in eastern australia and new zealand. *Scientific Reports*, 8(1):2661, 2018.
- [10] Robin S. Waples and Pierre Feutry. Close-kin methods to estimate census size and effective population size. *Fish and Fisheries*, 23(2):273–293, 2022.
- [11] 入江貴博. 中立遺伝マーカーを用いた近親判別に基づく個体数推定の可能性. **月刊海洋**, 549:340–348, 2016.
- [12] George Arthur Frederick Seber. *The Estimation of Animal Abundance and Related Parameters*. Griffin London, second edition, 2002.
- [13] Samuel Prystupa, Gregory McCracken, Robert Perry, and Daniel Ruzzante. Population abundance in arctic grayling using genetics and close-kin mark-recapture. *Ecology and Evolution*, 11, 2021.
- [14] Daniel E. Ruzzante, Gregory R. McCracken, Brage Følrand, John MacMillan, Daniela Notte, Colin Buhariwalla, Joanna Mills Flemming, and Hans Skaug. Validation of close-kin mark–recapture (CKMR) methods for estimating population abundance. *Methods in Ecology and Evolution*, 10(9):1445–1453, 2019.
- [15] Tetsuya Akita. 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.
- [16] Verena M Trenkel, Grégory Charrier, Pascal Lorance, and Mark V Bravington. Close-kin mark–recapture abundance estimation: practical insights and lessons learned. *ICES Journal of Marine Science*, 2022. fsac002.
- [17] Robin S. Waples and Chi Do. Linkage disequilibrium estimates of contemporary ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3(3):244–262, 2010.

- [18] Robin S Waples, Tiago Antao, and Gordon Luikart. Effects of Overlapping Generations on Linkage Disequilibrium Estimates of Effective Population Size. *Genetics*, 197(2):769–780, 06 2014.
- [19] Sebastian Wacker, Hans J. Skaug, Torbjørn Forseth, Øyvind Solem, Eva M. Ulvan, Peder Fiske, and Sten Karlsson. Considering sampling bias in close-kin mark-recapture abundance estimates of atlantic salmon. *Ecology and Evolution*, 11(9):3917–3932, 2021.
- [20] Benjamin Marcy-Quay, Suresh A. Sethi, Nina O. Therkildsen, and Clifford E. Kraft. Expanding the feasibility of fish and wildlife assessments with close-kin mark-recapture. *Ecosphere*, 11(10):e03259, 2020.
- [21] Tetsuya Akita. Estimating contemporary migration numbers of adults based on kinship relationships in iteroparous species. *Molecular Ecology Resources*, 22(8):3006–3017, 2022.
- [22] Aurélien Delaval, Victoria Bendall, Stuart J. Hetherington, Hans J. Skaug, Michelle Frost, Catherine S. Jones, and Leslie R. Noble. Evaluating the suitability of close-kin mark-recapture as a demographic modelling tool for a critically endangered elasmobranch population. *Evolutionary Applications*, in press.
- [23] Jinliang Wang. A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular Ecology*, 18(10):2148–2164, 2009.
- [24] Robin S. Waples, Gordon Luikart, James R. Faulkner, and David A. Tallmon. Simple life-history traits explain key effective population size ratios across diverse taxa. *Proceedings of the Royal Society B: Biological Sciences*, 280(1768):20131339, 2013.
- [25] Tetsuya Akita. Nearly unbiased estimator of contemporary N_e/N based on kinship relationships. *Ecology and Evolution*, 10(19):10343–10352, 2020.
- [26] Sean Hoban, Michael Bruford, Josephine D’Urban Jackson, Margarida Lopes-Fernandes, Myriam Heuertz, Paul A. Hohenlohe, Ivan Paz-Vinas, Per Sjögren-Gulve, Gernot Segelbacher, Cristiano Vernesi, Sally Aitken, Laura D. Bertola, Paulette Bloomer, Martin Breed, Hernando Rodríguez-Correa, W. Chris Funk, Catherine E. Grueber, Margaret E. Hunter, Rodolfo Jaffe, Libby Liggins, Joachim Mergeay,

Farideh Moharrek, David O'Brien, Rob Ogden, Clarisse Palma-Silva, Jennifer Pier-son, Uma Ramakrishnan, Murielle Simo-Droissart, Naoki Tani, Lisette Waits, and Linda Laikre. Genetic diversity targets and indicators in the cbd post-2020 global biodiversity framework must be improved. *Biological Conservation*, 248:108654, 2020.

付録

単純なモデルから得られる親個体数推定値の変動係数

N_{adult} 個体からなる親集団を考える。繁殖期間終了後にサンプルされた親 n_{parent} 個体と子供 $n_{\text{offspring}}$ 個体から k_{PO} ペアの親子関係が発見される確率は、下記のように超幾何分布で表される。

$$\Pr[k_{\text{PO}}] = \frac{\binom{n_{\text{parent}}}{k_{\text{PO}}} \binom{N_{\text{adult}} - n_{\text{parent}}}{2n_{\text{offspring}} - k_{\text{PO}}}}{\binom{N_{\text{adult}}}{2n_{\text{offspring}}}} \quad (\text{A1})$$

数式中の2は、子のゲノムが父と母由来の2セットの染色体から構成されていることに由来する。その平均は、

$$E[k_{\text{PO}}] = \frac{2n_{\text{parent}}n_{\text{offspring}}}{N_{\text{adult}}} \quad (\text{A2})$$

となり、これを変形したのが式(5)である。一般にCKMR法の適用対象は、 N_{adult} が n_{parent} や $n_{\text{offspring}}$ に比べてずっと大きい場合が通常である。この時、 k_{PO} が従う分布はポアソン分布で近似できることが知られており、

$$V[k_{\text{PO}}] \approx E[k_{\text{PO}}] \quad (\text{A3})$$

としても実用上は問題にならない。さて、ここで考えたいのは、本文中に式(6)で示した推定量 \hat{N}_{adult} の推定精度についてである。この推定量の分散は、下記のように求められる。

$$\begin{aligned} V[\hat{N}_{\text{adult}}] &= V\left[\frac{2n_{\text{parent}}n_{\text{offspring}}}{k_{\text{PO}}}\right] \\ &= (2n_{\text{parent}}n_{\text{offspring}})^2 V\left[\frac{1}{k_{\text{PO}}}\right] \\ &\approx (2n_{\text{parent}}n_{\text{offspring}})^2 \frac{V[k_{\text{PO}}]}{(E[k_{\text{PO}}])^4} \\ &\approx \frac{(2n_{\text{parent}}n_{\text{offspring}})^2}{(E[k_{\text{PO}}])^3} \\ &= \frac{N_{\text{adult}}^3}{2n_{\text{parent}}n_{\text{offspring}}} \end{aligned} \quad (\text{A4})$$

式変形について補足しよう。1行目から2行目：分散の中から係数を外に出す場合は2乗がつく。2行目から3行目：デルタ法という近似を用いた $(1/k_{PO} \approx 1/E[k_{PO}] - (k_{PO} - E[k_{PO}])/E[k_{PO}]^2)$ と、期待値まわりで1次の項までテーラー展開し、両辺の分散をとることで得られる)。3行目から4行目： k_{PO} は近似的にポアソン分布に従うとして、分散は平均と等しいとした。4行目から5行目：式 (A2) を代入した。この分散の平方根を N_{adult} で割ると、本文中の式 (7) で出てきた変動係数が得られる。なお、得られた変動係数の右辺の N_{adult} を \hat{N}_{adult} で置き換えることで、本文中の式 (8) で示した変動係数の推定量（プラグイン推定量と呼ばれる）が得られる。