

Estimating contemporary migration numbers of adults based on kinship relationships in iteroparous species

メタデータ	言語: English 出版者: 公開日: 2023-01-24 キーワード (Ja): キーワード (En): close-kin mark-recapture; dispersion; kinship assignment; reproductive potential 作成者: 秋田, 鉄也 メールアドレス: 所属: 水産研究・教育機構
URL	https://fra.repo.nii.ac.jp/records/124

Estimating contemporary migration numbers of adults based on kinship relationships in iteroparous species

Tetsuya Akita 

Fisheries Resources Institute, Japan
Fisheries Research and Education Agency,
Kanagawa, Japan

Correspondence

Tetsuya Akita, Fisheries Resources
Institute, Japan Fisheries Research and
Education Agency, Kanagawa 236-8648,
Japan.

Email: akita_tetsuya81@fra.go.jp

Funding information

Japan Society for the Promotion of
Science, Grant/Award Number: 19K06862
and 20H03012

Handling Editor: Frederic Austerlitz

Abstract

This paper describes the development of estimators for the contemporary migration number and rate of adults between two populations in iteroparous species. The proposed estimators are based on known half-sibling (HS) and/or parent-offspring (PO) relationships observed between populations across breeding seasons. The rationale is that HS and PO pairs exhibit information about the occurrence frequency of parental movements during the breeding interval. The proposed method allows for variance in the average number of offspring per parent within and between populations. In addition, coupled with the PO pairs found within the population, the estimators can be obtained using only genetic data. Generally, a sample size representing the square root of the population size is required to obtain meaningful migration information. We describe a detailed evaluation of the performance of the proposed estimators by running an individual-based model, and the results provide guidance regarding sample sizes to ensure the required accuracy and precision. In addition, given that there are few effective methods to estimate adult movement (especially when populations cannot be genetically distinct), we discuss the usefulness of the proposed kinship assignment method in terms of conservation biology and wildlife management.

KEYWORDS

close-kin mark-recapture, dispersion, kinship assignment, reproductive potential

1 | INTRODUCTION

In conservation and wildlife management, estimating current ecological information is required to effectively monitor population levels and propose effective management strategies. The contemporary migration rate, that is the recent rate of movement of individuals or gametes between predefined populations, is a critical component when determining the current degree of gene flow and synchrony among the population (Lowe & Allendorf, 2010; Waples & Gaggiotti, 2006). Thus, contemporary migration rate estimations provide information about the degree of genetic differentiation and demographic dependency, which can delineate evolutionarily significant units and identify appropriate management units (Moritz, 1994; Palsbøll et al., 2007). In addition, the availability of

a migration rate estimator can link to underlying population dynamics, thereby allowing for precise and flexible evaluations of management practices, for example stock assessment in fisheries (Hampton & Fournier, 2001) and invasive species control (Sakai et al., 2001).

Essentially, there are two primary approaches to estimating the contemporary migration rate. The first approach involves using the mark-recapture (MR) method, which estimates the migration rate and other population parameters (Kéry & Schaub, 2011; Thorson et al., 2021). Here, the rationale is that, for example, conventional tags provide release and recovery location information for known release and recovery dates, which generates a movement fraction matrix among strata per the given time interval. This direct method is relatively easy to interpret; however, it is generally hampered by

several uncertainties, for example tag loss, tagging-related mortality and time-varying reporting rates (Hilborn & Walters, 1992). In addition, the practical application of the MR method to infer information about parental movements is generally limited to situations where the collection of a sufficient number of adult samples at multiple sites is possible over a short period.

The second approach involves using numerous genetic markers to assign individuals to source populations, thereby allowing the inference of recent migration (since Paetkau et al., 1995). The previously developed model of Wilson and Rannala (2003), which a population assignment method implemented in the "BayesAss" software, can output the point estimate as an element of the current migration matrix. However, according to several papers (Faubet et al., 2007; Wang, 2014), the accuracy of the migration rates estimated by this model is valid only when populations are highly differentiated (e.g. $F_{ST} \geq 0.05$). This suggests that it may be difficult to infer the degree of demographic dependency among populations because demographic independency is realized even if there are many migrants.

Parental assignments, which are based on genetic markers and are frequently used to complement the population assignment method in ecological studies, also provide information about the current level of migration. In contrast to the population assignment method, the parental assignment method does not require population differentiation (Wang, 2014); however, the current methodology focuses on estimating dispersal kernels and, to estimate the migration rate, requires the assumption that adults are not migrant, for example pollen and seed dispersal in plants (Ashley, 2010) or larval drift in marine animals (Gagnaire et al., 2015). Thus, this method may not consider adult movement of iteroparous species, that is multiple reproductive cycles during the lifetime, which may change the spawning ground in each breeding season.

Close-kin mark-recapture (CKMR) is a recently developed method to estimate adult population size with life-history parameters. The CKMR method utilizes the known kinship information in a sample (Bravington, Grewe, & Davies, 2016; Bravington, Skaug, & Anderson, 2016; Hillary et al., 2018; Skaug, 2017; Waples & Feutry, 2022) although similar methods have been proposed in the beginning of the 21st century (Nielsen et al., 2001; Pearse et al., 2001; Skaug, 2001). In the CKMR method, the presence of a kinship pair in the sample is analogous to recapture of a marked individual in the MR method. Note that kinship pairs in the sample are less likely to be observed in larger populations; thus, the number of kinship pairs may reflect the number of adults in the population. In addition, the CKMR method has the potential to infer information about parental movements even when the populations are not highly differentiated because the number of kinship pairs found between populations may contain information about the frequency of parental movements during the breeding interval. Recent CKMR studies have argued that there is availability of kinship information to identify metapopulation structure (Conn et al., 2020; Trenkel et al., 2022); however, CKMR-based migration estimation

has not been well developed or investigated extensively in simulation studies.

Thus, in this paper, we propose a method to estimate the contemporary migration number or the rate of adults between two pre-defined populations in iteroparous species as an extension of current CKMR techniques. Under the assumption that kinships are genetically detected without error, the proposed method is based on the numbers of half-sibling (HS) and parent-offspring (PO) pairs in a sample. Here, sampling can be either invasive or non-invasive (hereafter, non-invasive is considered to mean returning the sampled individual to the population) and is completed in two breeding seasons. Sampling offspring (young-of-year individuals) and parents likely share a PO relationship with the offspring sampled in one population in the first breeding season and in the other population in the second breeding season. Our model explicitly incorporates reproductive variation both within and between populations, which makes it possible to target a species whose fertility is affected by environmental differences between populations. In the following, we first explain the modelling assumption and sampling scheme, where migration direction is specified. We then analytically determine the estimators of the contemporary migration number or rate of adults, which are based on the numbers of HS and/or PO pairs. Finally, by running an individual-based model, we investigate the performance of the estimator and provide a guide for sample size. Note that the proposed modelling framework can be applied to diverse animal species; however, in this paper, we describe the model relative to fish species, which are currently the best target candidate for the proposed method.

2 | THEORY

Here, we present the theoretical foundation of estimating a migrant number and migration rate of iteroparous species using HS or PO pairs found between two populations with different sample timing. Note that the estimators can hold under flexible assumptions for reproduction (Appendix B); however, here, we assume a relatively simple case. The main symbols used in this paper are summarized in Table 1.

2.1 | Hypothetical population

Assume we have a set of populations 1 and 2, where random mating occurs within each population and parents can move to the other population after the reproductive season ends. Without loss of generality, we consider the movement of parents from population 1 to population 2, and we focus on estimating the migration number or rate. In this framework, we demonstrate that considering two reproductive seasons (referred to as "the first year" and "the second year," respectively) is sufficient to estimate the migration number or rate. Figure 1a shows a schematic representation of the kinship relationships and parent movements.

TABLE 1 List of mathematical symbols employed in the main text

$n_{p,1}, n_{p,2}$	Sampled number of parents from populations 1 and 2
$n_{o,1}, n_{o,2}$	Sampled number of offspring from populations 1 and 2
N_1, N_2	Number of parents in populations 1 and 2 when sampled offspring are born
M	Number of survived migrants of parents from population 1 to population 2
m	Migration rate of parents from population 1 to population 2, defined by M/N_1
r	Sex ratio
$\pi_{PO,1}, \pi_{PO,2}$	Probability that a randomly selected pair (parent and offspring) Shares a parent-offspring relationship within populations 1 and 2
$\pi_{PO,bet}$	Probability that a randomly selected pair (parent and offspring) Shares a parent-offspring relationship between populations 1 and 2
$\pi_{HS,bet}$	Probability that a randomly selected pair (two offspring) Shares a half-sibling relationship between populations 1 and 2
$\lambda_{i,1}, \lambda_{j,2}$	Expected number of surviving offspring of parent i and j at sampling in populations 1 and 2
$\lambda_{M,l,1}, \lambda_{M,l,2}$	Expected number of surviving offspring of migrant l at sampling in populations 1 and 2
$k_{i,1}, k_{j,2}$	Number of surviving offspring born to parent i and j in populations 1 and 2
$H_{PO,1}, H_{PO,2}$	Number of parent-offspring pairs observed in samples within populations 1 and 2
$H_{PO,bet}$	Number of parent-offspring pairs observed in samples between populations 1 and 2
$H_{HS,bet}$	Number of half-sibling pairs observed in samples between populations 1 and 2

Note: Subscripts "1" and "2" indicate the quantity in population 1 during the first year and in population 2 during the second year, respectively.

Assume there are N_1 parents in population 1 at the beginning of the first year, and each parent produces a number of offspring that is governed by the parent's reproductive potential, which is denoted

$\lambda_{i,1}$ ($i = 1, 2, \dots, N_1$). After the reproductive season, some parents begin to move toward population 2. Here, M survived migrants arrive in population 2; thus, there are N_2 parents in population 2 at the beginning of the second year ($M \leq N_1$ and $M \leq N_2$). Similar to population 1, N_2 parents produce a number of offspring governed by $\lambda_{j,2}$ ($j = 1, 2, \dots, N_2$). Note that reproductive potential is determined by several factors. Additional details about reproductive potential can be found in Appendix A. Theoretically, the number of offspring for each parent is set as a random variable with mean $\lambda_{i,1}$ or $\lambda_{j,2}$.

2.2 | Sampling

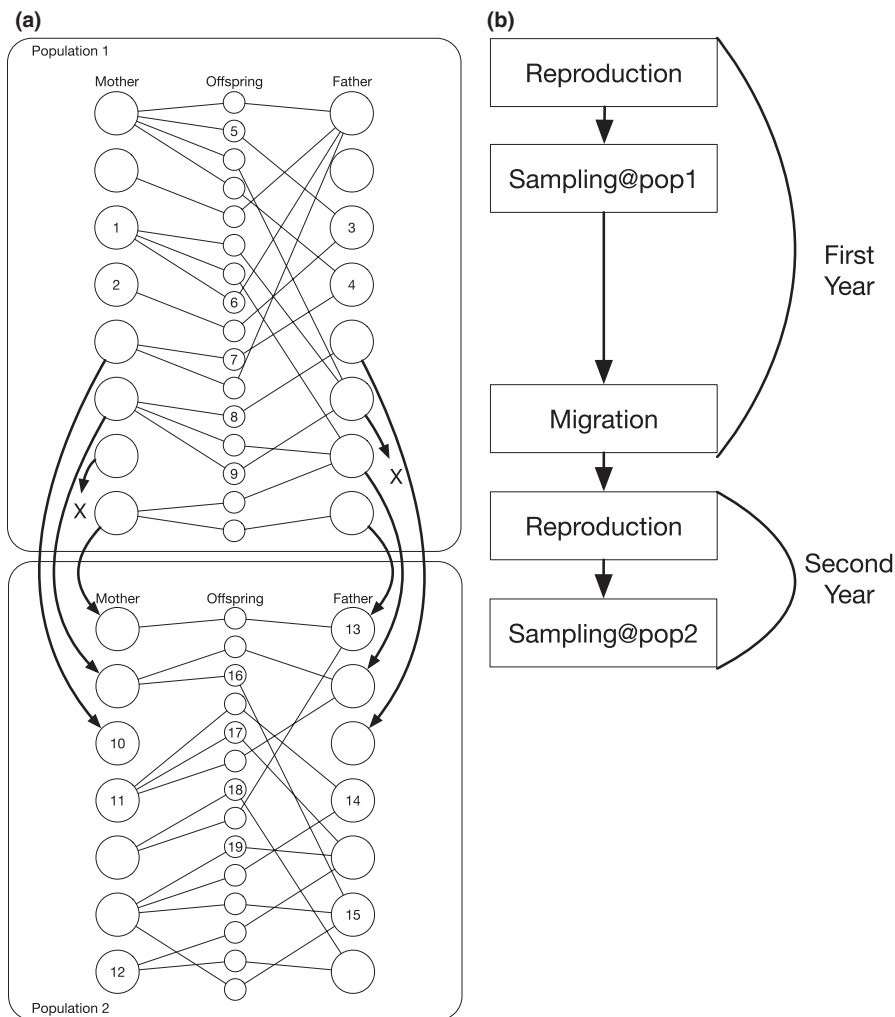
To estimate the migration number or rate, we utilize the number of HS and/or PO pairs observed in a given sample. In both populations, $n_{o,1}$ and $n_{o,2}$ offspring are sampled randomly in the first and second year, respectively, which are identified as young-of-year individuals without error. In addition, $n_{p,1}$ and $n_{p,2}$ parents are sampled randomly immediately after the end of the reproductive season in the first and second years, respectively. For mathematical tractability, the parents must survive the reproductive season; thus, both the mother and father of a given offspring have the potential to be sampled in the same year. Note that all four sample types, that is $n_{p,1}$, $n_{p,2}$, $n_{o,1}$ and $n_{o,2}$, are not always required to estimate the migration number or rate. The types of required samples depend on the given situation, which is explained later. Up to Section 2.6, we focus on only the HS and PO pairs found between populations 1 and 2. To apply the proposed method to fishery assessment, we assume an invasive sampling procedure, which prevents us from finding a PO pair such that a parent and offspring are sampled in populations 1 and 2, respectively. Figure 1b shows the timeline of the sampling scheme for these hypothetical populations.

In the example shown in Figure 1a, five offspring and four parents are sampled in population 1, four offspring and six parents are sampled in population 2, and two HS pairs and a single PO pair are observed between populations 1 and 2. In addition, several PO pairs are observed in the same population, which are available to estimate parent numbers (Section 2.7). In our modelling framework, a full-sibling relationship is counted as two HS pairs. The numbers of HS and PO pairs found between the two populations are determined via pairwise comparison of all sample individuals, that is comparisons of $n_{o,1} \times n_{o,2}$ and $n_{o,1} \times n_{p,2}$, respectively.

2.3 | HS pair-based model

Here, we consider the probability that two offspring sampled in populations 1 and 2 will share a HS relationship with an arbitrary mother or father, denoted $\pi_{HS,bet}$. Note that we assume equal reproductive potential among parents (up to Section 2.7) although this assumption can be relaxed for most of the case (refer to Appendix B for additional information). Here, $\pi_{HS,bet}$ can be partitioned into three probabilities, that is (i) the probability that a sampled offspring in

FIGURE 1 (a) Hypothetical populations with $N_1 = 16$, $N_2 = 14$ and $M = 6$. Upper and lower areas indicate individuals in population 1 before migration (at the sample timing in the first year) and individuals in population 2 after reproduction (at the sample timing in the second year), respectively. Open circles on the left, right and centre represent mothers, fathers and their offspring, respectively. The thin line denotes PO relationship. Bold arrows denote migration, and x denotes failure to survive at the sampling in the second year. Sampled individuals are labelled with an index number. Number of sampled individuals in this example: $n_{p,1} = 4$, $n_{o,1} = 5$, $n_{p,2} = 6$, $n_{o,2} = 4$; numbers of kinship pairs: $H_{PO,bet} = 1$ (i.e. "7-10" pair), $H_{HS,bet} = 2$ (i.e. "8-16" and "9-16" pairs), $H_{PO,1} = 3$ (i.e. "1-6", "3-5" and "4-7" pairs) and $H_{PO,2} = 2$ (i.e. "11-17" and "15-16" pairs). (b) Phases of events relevant to this study in the given timeline



population 1 is born to a parent that safely arrives in population 2 (hereafter, referred to as a "migrant"); (ii) the probability that a sampled offspring in population 2 is born to a migrant parent; and (iii) the probability that the migrant parent of a sampled offspring in population 1 and the migrant parent of a sampled offspring in population 2 are identical.

To assess these probabilities, recall that there must be two parents of arbitrary offspring. The first probability is the sum of the father-offspring and mother-offspring relationship, which can be expressed as $r_M M / (r_{S,1} N_1) + (1 - r_M) M / ((1 - r_{S,1}) N_1)$, where r_M and $r_{S,1}$ indicate the sex ratio of the parents in the migrants and in all of population 1, respectively. Under the assumption that $r_M = r_{S,1}$, the probability is simplified to $2M / N_1$. Similarly, the second probability is simplified to $2M / N_2$, which implies the equal sex ratio assumption in migrants and others. The third probability corresponds to the probability that two randomly selected migrants are identical, that is $1 / M$. Taken together, we obtain the following:

$$\pi_{HS,bet} = \frac{2M}{N_1} \frac{2M}{N_2} \frac{1}{M} = \frac{4M}{N_1 N_2} \quad (1)$$

This form holds under a flexible setting for reproductive potential, for example a situation with random variable of $\lambda_{i,1}$ and $\lambda_{j,2}$ (Appendix B). If all N_1 parents safely move to population 2 (i.e. $M = N_1$), $\pi_{HS,bet}$ equals $4 / N_2$. In addition, if population 2 comprises only migrants (i.e. $M = N_2$), $\pi_{HS,bet}$ equals $4 / N_1$. In extreme cases, these probabilities take similar forms as the HS probability, which is sampled randomly from two different cohorts within a population (Bravington, Skaug, & Anderson, 2016).

Here, let $H_{HS,bet}$ be the number of HS pairs found in the offspring samples of size $n_{O,1}$ and $n_{O,2}$. Assuming the total number of HS pairs between the two populations is much greater than $H_{HS,bet}$, the distribution is approximated by a binomial form (i.e. $H_{HS,bet} \sim \text{Binom}[\pi_{HS,bet}, n_{O,1} n_{O,2}]$). Thus, the theoretical expectation of $H_{HS,bet}$ is given as follows:

$$\mathbb{E}[H_{HS,bet}] = \pi_{HS,bet} n_{O,1} n_{O,2} = \frac{4n_{O,1} n_{O,2} M}{N_1 N_2} \quad (2)$$

$$= \frac{4n_{O,1} n_{O,2} m}{N_2}, \quad (3)$$

where m is the migration rate satisfying $M = mN_1$. The observed number of HS pairs in a sample found between populations 1 and 2 is defined by $\tilde{H}_{HS,bet}$, and $E[H_{HS,bet}]$ in Equation 2 is replaced by $\tilde{H}_{HS,bet}$, thereby generating the linear estimator of M :

$$\hat{M}_1 = \frac{N_1 N_2 \tilde{H}_{HS,bet}}{4n_{O,1}n_{O,2}}. \quad (4)$$

In this paper, the “tilde” and “hat” symbols indicate the observation and estimator of a variable, respectively. Similarly, $E[H_{HS,bet}]$ in Equation 3 is replaced by $\tilde{H}_{HS,bet}$, which generates the linear estimator of m :

$$\hat{m}_1 = \frac{N_2 \tilde{H}_{HS,bet}}{4n_{O,1}n_{O,2}}. \quad (5)$$

Note that bidirectional migration does not affect the value of these estimators because any reduction to N_2 (via parental movements from population 2 to population 1) is cancelled by increasing $E[H_{HS,bet}]$. The subscripts associated with these estimators indicate the numbering of the proposed estimators, which is summarized in Table 2.

2.4 | PO pair-based model

In the following, we consider the probability that offspring sampled in population 1 and a parent sampled in population 2 share a PO relationship, denoted $\pi_{PO,bet}$. Here, $\pi_{PO,bet}$ can be partitioned into the above-mentioned probabilities. Note that probabilities (i) and (iii) are the same as that introduced in the Section 2.3, that is $2M/N_1$ and $1/M$. In addition, probability (ii) is M/N_2 by definition. Taken together, we obtain the following:

$$\begin{aligned} \pi_{PO,bet} &= \frac{2M}{N_1} \frac{M}{N_2} \frac{1}{M} \\ &= \frac{2M}{N_1 N_2}. \end{aligned} \quad (6)$$

This form holds under a flexible setting for the reproductive potential, as noted in Appendix B. Here, if all N_1 parents safely move to

population 2 (i.e. $M = N_1$), $\pi_{PO,bet}$ equals $2/N_2$. In addition, if population 2 comprises only migrants (i.e. $M = N_2$), $\pi_{PO,bet}$ equals $2/N_1$. In extreme cases, these probabilities take similar forms to the PO probability that is sampled randomly within a population (Bravington, Skaug, & Anderson, 2016).

Here, let $H_{PO,bet}$ be the number of PO pairs found in offspring samples of size $n_{O,1}$ and parent samples of size $n_{P,2}$. Under the assumption that the total number of PO pairs between the two populations is much greater than $H_{PO,bet}$, the distribution is approximated by a binomial form (i.e. $H_{PO,bet} \sim \text{Binom}[\pi_{PO,bet}, n_{O,1}n_{P,2}]$). Thus, the theoretical expectation of $H_{PO,bet}$ is expressed as follows:

$$\begin{aligned} E[H_{PO,bet}] &= \pi_{PO,bet} n_{O,1} n_{P,2} \\ &= \frac{2n_{O,1}n_{P,2}M}{N_1 N_2} \end{aligned} \quad (7)$$

$$= \frac{2n_{O,1}n_{P,2}m}{N_2}. \quad (8)$$

The observed number of PO pairs in a sample is defined by $\tilde{H}_{PO,bet}$, and $E[H_{PO,bet}]$ in Equations 7 and 8 is replaced by $\tilde{H}_{PO,bet}$, which generates the linear estimators of M and m as follows:

$$\hat{M}_2 = \frac{N_1 N_2 \tilde{H}_{PO,bet}}{2n_{O,1}n_{P,2}}, \quad (9)$$

and

$$\hat{m}_2 = \frac{N_2 \tilde{H}_{PO,bet}}{2n_{O,1}n_{P,2}}. \quad (10)$$

2.5 | Required sample size

The proposed estimators are based on the observed number of kinship pairs. Their expected number is determined linearly by the number of pairwise comparison (Equations 2–3 and 7–8), which provides sample size guidance to ensure the condition that at least one or more kinship pairs can be found. The conditions are given as follows:

$$n_{O,1}n_{O,2} > \frac{N_2}{4m}, \quad (11)$$

TABLE 2 Summary of proposed estimators for required parameters and kinship types

Estimator	N_1	N_2	$n_{O,1}$	$n_{O,2}$	$n_{P,1}$	$n_{P,2}$	Required kinship type
\hat{M}_1	Given	Given	✓	✓			HSP
\hat{M}_2	Given	Given	✓			✓	POP
\hat{M}_3	Given	Given	✓	✓		✓	HSP & POP
\hat{M}_4	Given	Estimated	✓	✓	✓	✓	HSP & POP
\hat{M}_5	Estimated	Estimated	✓	✓		✓	HSP & POP
\hat{m}_1	—	Given	✓	✓			HSP
\hat{m}_2	—	Given	✓			✓	POP
\hat{m}_3	—	Given	✓	✓		✓	HSP & POP
\hat{m}_4	—	Estimated	✓	✓		✓	HSP & POP

Abbreviations: HSP, half-sibling pair; POP, parent-offspring pair.

and

$$n_{O,1}n_{P,2} > \frac{N_2}{2m}. \quad (12)$$

Generally, the sample size required to ensure the above conditions is $n > \sqrt{N_2}$, where $n = n_{O,1} = n_{O,2} = n_{P,2}$. Note that a very small m value, which dramatically increases the required sample size, produces significant genetic differences between the two populations. In such cases, there are several methods to estimate contemporary migration rate using population assignment technique such as "BayesAss" (Wilson & Rannala, 2003), "GENECLASS2" (Piry et al., 2004), or "BIMr" (Faubet & Gaggiotti, 2008). Alternatively, in this paper, we focus on a moderate-to-large migration rate, which yields little genetic differences between the two populations.

Other guidance for the required sample size is obtained by an approximate lower bound on the coefficient of variation (CV) of M_1 or m_1 , which is defined as $1/\sqrt{H_{HS,bet}}$. This is applied in the context of classic MR (Seber, 2002) or CKMR (Bravington, Skaug, & Anderson, 2016). For example, to achieve a 30% CV, the target of $H_{HS,bet}$ is >10 , thereby providing the required sample size when the population parameters are given (e.g. N_1 , N_2 and M). Similarly, the CV of M_2 or m_2 , defined by $1/\sqrt{H_{PO,bet}}$, provides the sample size required to estimate M_2 or m_2 .

2.6 | Efficient use of kinship pairs found between populations

When both HS and PO pairs, which are found between populations, are available simultaneously, we can obtain efficient estimators by combining \hat{M}_1 and \hat{M}_2 for the migration number:

$$\hat{M}_3 = \frac{N_1 N_2 (\tilde{H}_{HS,bet} + \tilde{H}_{PO,bet})}{2n_{O,1}(2n_{O,2} + n_{P,2})}, \quad (13)$$

and by combining \hat{m}_1 and \hat{m}_2 for the migration rate:

$$\hat{m}_3 = \frac{N_2 (\tilde{H}_{HS,bet} + \tilde{H}_{PO,bet})}{2n_{O,1}(2n_{O,2} + n_{P,2})}. \quad (14)$$

For these estimators, $\tilde{H}_{HS,bet}$ and $\tilde{H}_{PO,bet}$ are weighted by $2n_{O,2}$ and $n_{P,2}$, respectively.

2.7 | Estimation of parent number by PO pairs in a population

To this point, we have stated that the formulation of the estimators presented in Equations 4–5 and 9–10 is a function with parent numbers for each population (N_1 and/or N_2). In other words, such estimators are only available when the parent number or numbers are known. Here, we describe how we can estimate unknown parent numbers by also using PO pairs to produce estimators for both migration number and rate that can be obtained from only genetic data.

When PO pairs found in offspring and parent samples from population 2 in the second year are available, the standard parental number estimator can be obtained (Bravington, Skaug, & Anderson, 2016) as follows:

$$\hat{N}_2 = \frac{2n_{O,2}n_{P,2}}{\tilde{H}_{PO,2} + 1}, \quad (15)$$

where $\tilde{H}_{PO,2}$ is the observed number of PO pairs found in the offspring samples of size $n_{O,2}$ and parent samples of size $n_{P,2}$. Here, the "+1" term reduces bias, especially when $\tilde{H}_{PO,2}$ is small (e.g. Prystupa et al., 2021), where a similar derivation of this bias correction is provided in the literature Akita, 2020a. By replacing N_2 with \hat{N}_2 in Equations 13 and 14, we obtain the following estimators:

$$\hat{M}_4 = \frac{n_{O,2}n_{P,2}N_1(\tilde{H}_{HS,bet} + \tilde{H}_{PO,bet})}{n_{O,1}(2n_{O,2} + n_{P,2})(\tilde{H}_{PO,2} + 1)}, \quad (16)$$

and

$$\hat{m}_4 = \frac{n_{O,2}n_{P,2}(\tilde{H}_{HS,bet} + \tilde{H}_{PO,bet})}{n_{O,1}(2n_{O,2} + n_{P,2})(\tilde{H}_{PO,2} + 1)}. \quad (17)$$

The estimator of migration number (Equation 16) requires the (unknown) parent number in population 1 (N_1); however, the estimator of migration rate (Equation 17) can be obtained using only the observed number of HS and PO pairs.

Similar to estimating N_2 , when PO pairs found in offspring and parent samples from population 1 in the first year are available, we can obtain the estimator of N_1 as follows:

$$\hat{N}_1 = \frac{2n_{O,1}n_{P,1}}{\tilde{H}_{PO,1} + 1}, \quad (18)$$

where $\tilde{H}_{PO,1}$ is the observed number of PO pairs in the offspring samples of size $n_{O,1}$ and parent samples of size $n_{P,1}$. By replacing N_1 with \hat{N}_1 in Equation 16, we obtain a migration number estimator that can be calculated from only genetic data, which is expressed as follows:

$$\hat{M}_5 = \frac{2n_{O,1}n_{P,1}n_{O,2}n_{P,2}(\tilde{H}_{HS,bet} + \tilde{H}_{PO,bet})}{n_{O,1}(2n_{O,2} + n_{P,2})(\tilde{H}_{PO,1} + 1)(\tilde{H}_{PO,2} + 1)}. \quad (19)$$

Table 2 summarizes the conditions for estimators in terms of whether N_1 and/or N_2 are known and which type of samples is required for estimation, and Table 2 identifies the kinship type required to calculate the estimator.

2.8 | Variation in reproductive potential among individuals and populations

The proposed estimators are derived under the assumption that reproductive potential is equal among individuals. As described in Appendix B, this assumption can be relaxed, and the estimators still hold, which is exemplified in iteroparous species that may

show significant variation in reproductive potential among individuals. In addition, the estimators hold for most cases where the mean reproductive potential differs between populations 1 and 2, for example when the environmental conditions for reproductive success vary between populations. Here, the required condition is that the distribution of reproductive potential does not differ between sexes and that the migration event does not depend on the degree of reproductive potential (refer to Appendix C for additional details).

2.9 | Individual-based model

We developed an individual-based model that tracks kinship relationships to evaluate the performance of the estimators. Here, the population structure was assumed to be identical to that in the development of the estimators. Populations 1 and 2 comprised N_1 and N_2 parents with an equal-sex ratio, and their offspring number was assumed to follow the geometric distribution with mean $\bar{\lambda}_1$ and $\bar{\lambda}_2$ (i.e. Poisson reproduction with mean $\lambda_{i,1}$ and $\lambda_{i,2}$, which follows the exponential distribution with mean $\bar{\lambda}_1$ and $\bar{\lambda}_2$, respectively). Migrant parents were selected randomly from population 1 at the end of the first year, and each offspring retained the IDs of the parents, thereby making it possible to trace the HS and PO relationship.

Here, the parameter set ($N_1, N_2, M, \bar{\lambda}_1, \bar{\lambda}_2, n_{O,1}, n_{O,2}, n_{P,1}, n_{P,2}$) was given. We simulated a population history and sampling process,

which generates the proposed estimators. Note that this process was repeated 1000 times, which allowed us to construct the distribution of the estimators for each parameter set. All scripts (written in the C++ and R languages) and documentation for these analyses are available at <https://doi.org/10.5281/zenodo.6668905>.

3 | RESULTS

We evaluated the performance of \hat{M}_s or \hat{m}_s numerically for a case with variable reproductive potential among parents. Here, the scaled statistical properties of $\hat{m}_1, \hat{m}_2, \hat{m}_3$ and \hat{m}_4 were the same as $\hat{M}_1, \hat{M}_2, \hat{M}_3$ and \hat{M}_4 . Thus, hereafter, we only discuss the results for $\hat{M}_1 - \hat{M}_5$. Figure 2 shows the distribution of the relative bias of \hat{M} for limiting cases where the parent and offspring sample numbers were identical (i.e. $n = n_{P,1} = n_{P,2} = n_{O,1} = n_{O,2}$) and parent numbers in the two populations were also identical (i.e. $N = N_1 = N_2$). The relative bias was calculated by applying the outputs of the individual-based model, which is defined as “(averaged estimator – true value)/true value.” Refer to Table S1 for a full list of the parameter sets used to evaluate performance (relative bias and CV of \hat{M}_s).

First, we evaluated the accuracy of \hat{M}_s based on the relative bias. As expected, for most of the investigated parameter sets, we observed that their relative bias was <5%, as shown in Table S1. Thus, when the assumptions are satisfied, that is (i) offspring and parents are sampled randomly, (ii) kinships are detected without error, and

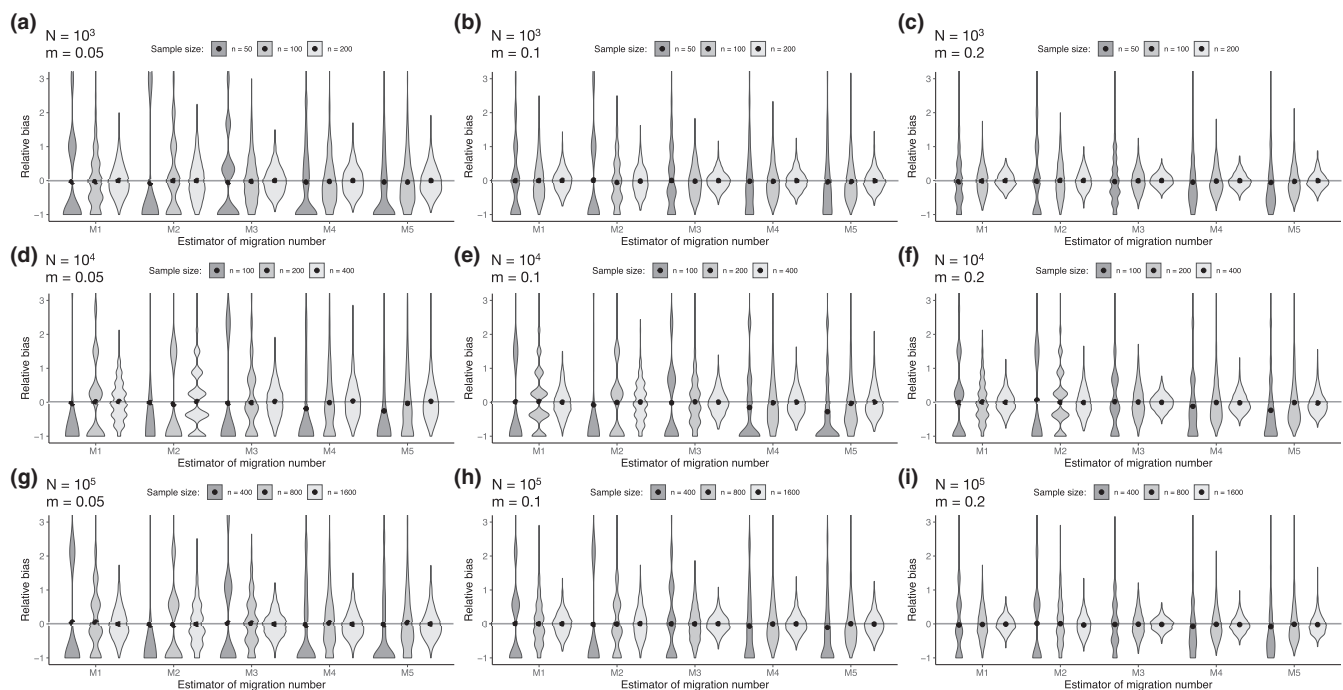


FIGURE 2 Violin plots showing the distribution of relative bias in the estimator of M for various sample sizes, parent numbers and migration rates. Filled circles represent mean values. The sample sizes for parents and offspring are identical (i.e. $n = n_{P,1} = n_{P,2} = n_{O,1} = n_{O,2}$), and the parent sizes in the two populations are identical (i.e. $N = N_1 = N_2$), as indicated in the legend. Migration rate m is specified by N_1/M . For demonstration purposes, the upper side of the distribution is truncated, although the mean values are calculated (including the truncated values)

(iii) the migration event does not depend on the degree of reproductive potential, it is reasonable to consider \hat{M} s nearly unbiased estimators.

Next, we evaluated the precision of \hat{M} s based on the CV value. Table S1 shows the CV value, and the violin plots in Figure 2 visualize the degree of precision. For each estimator \hat{M} , we found that precision increased with an increasing sample size. Note that the total sample size depends on the estimator. For example, $n = 50$ in Figure 2, the total sample sizes of \hat{M}_1 , \hat{M}_2 , \hat{M}_3 , \hat{M}_4 and \hat{M}_5 were 100 ($= n_{O,1} + n_{O,2}$), 100 ($= n_{O,1} + n_{P,2}$), 150 ($= n_{O,1} + n_{O,2} + n_{P,2}$), 150 ($= n_{O,1} + n_{O,2} + n_{P,2}$) and 200 ($= n_{O,1} + n_{O,2} + n_{P,1} + n_{P,2}$), respectively.

We found that precision increases as the sample size or number of migrants increased, and the shape of the distribution asymptotically became symmetric (Figure 2) because an increasing sample size or increasing number of migrants is likely to increase the observed number of kinship pairs found between the two populations ($\hat{H}_{HS,bet}$ or $\hat{H}_{PO,bet}$), which reduces the variance of those kinship pair numbers. In addition, the number of PO pairs in a given population ($\hat{H}_{PO,1}$ or $\hat{H}_{PO,2}$) contributes to the precision of estimating N_1 or N_2 , respectively, thereby providing \hat{M}_4 and \hat{M}_5 with relatively high precision.

Note that our simulation can handle cases with and without invasive sampling. Invasive sampling potentially affects the level of m because the sampled parents from population 1 have no chance to move to population 2; thus, such parents cannot be sampled. In this case, m might be defined by $M / (N_1 - n_{P,1})$ rather than M / N_1 , although this does not affect the estimator \hat{M} s, as shown in Table S1.

Finally, we investigated the case where the parental number in population 1 is much greater than that in population 2 (i.e. $N_1 \gg N_2$); thus, migrants represent a large proportion of N_2 . For such cases, for example $(N_1, N_2, M) = (10^4, 10^3, 500)$, we confirmed a similar property of \hat{M} s as observed in the $N_1 = N_2$ case, as shown in Table S1, which suggests the robustness of \hat{M} s in the presence of heterogeneous population sizes.

4 | DISCUSSION

In this study, we theoretically developed estimators for the contemporary migration number ($\hat{M}_1 - \hat{M}_5$) and migration rate ($\hat{m}_1 - \hat{m}_4$) of parents between two predefined populations in iteroparous species. The proposed estimators are based on the known PO relationship and HS relationships observed between and within the two populations without error in terms of kinship assignment, which is presented in the context of the CKMR method. Users can select the appropriate estimator for cases where the parental number of population 1 (N_1) and/or that of population 2 (N_2) is known (Table 2). The performance of the estimator (both accuracy and precision) was evaluated quantitatively by performing an individual-based simulation (Figure 2 and Table S1). The proposed modelling framework utilizes several types of reproductive variations (i.e. the number of survived offspring per parent), including the variance of reproductive potential within and between populations, in consideration of

several situations, including body-size structure or environmental heterogeneity for reproductive success.

Our primary contributions are summarized as follows. First, we have formulated the probabilities of kinship pairs selected randomly between two populations ($\pi_{HS,bet}$ and $\pi_{PO,bet}$), which provides migration number or rate estimators. While similar derivations may be found in the (nongenetic) MR method, these are limited to cases where sampling is non-invasive and adult individuals must be sampled. The proposed method can avoid these limitations, where such an advantage is characterized by the CKMR method to estimate population sizes. Second, we have demonstrated that the probabilities of kinship pairs are approximately independent of the reproductive potential; thus, information about it is not required to realize the estimation provided that migration is assumed to occur independently of the reproductive potential. This is a useful property of the estimators because it is natural that heterogeneity of reproductive potential within and between populations exists in iteroparous species; however, its observation is generally difficult. Third, we have demonstrated that estimators can be obtained using only genetic data (i.e. \hat{M}_5 and \hat{m}_4). There are several advantages to using the proposed estimator rather than estimating $\pi_{HS,bet}$ (and/or $\pi_{PO,bet}$) separately (via a genetic method) and N_1 (and/or N_2) (via a nongenetic method), including simplified sampling processes and analysing designs, and the availability of a unified framework for genetic analyses to detect HS and PO pairs (a similar discussion was given by Akita, 2020b relative to estimating the ratio of effective breeding size to the census size, N_b / N).

To estimate the contemporary migration number or migration rate, our simulation-based results provide sample size guidance to ensure that the required accuracy and precision are realized, especially if the order of the number of migrant parents and parental sizes are approximately known (Table S1). For example, when $m = 0.1$ and $N_1 = N_2 = 10^3$, sampling 10% and 20% of the parents and an equal number of offspring in both populations leads to 70% and 39% CV of \hat{M}_5 , respectively (in the invasive sampling case). Even if no information is available about these numbers, $1 / \sqrt{\hat{H}_{HS,bet}}$ (or $1 / \sqrt{\hat{H}_{PO,bet}}$) provides an approximate lower bound on the CV, which can be used as an indicator of the precision of \hat{M} or \hat{m} . In addition, the condition that $n > \sqrt{N_2}$ is also used as rule of thumb, especially when planning a research project ($n = n_{O,1} = n_{O,2} = n_{P,2}$; refer to Equations 11 and 12). Note that the guidance is to estimate the information about one-way migration between two populations; thus, estimating two-way migrations among two or three populations would require twice and three times the number of samples, respectively, compared with our sample size guidance.

We believe there are several scenarios where the proposed estimator may be beneficial. The first case is where there is a large number of migrants between populations, for example $M > 100$, which eliminates population differentiation and thus hampers detection of the M via a population genetics method if M is interpreted as an effective migration number. Information about movement between populations is essential to assess population dynamics in the conservation and management contexts, even if the migrant number

is sufficiently large that it cannot be genetically assigned to two populations. Note that demographic independency is realized even if there are many migrants; for example, Hastings (1993) theoretically examined the dynamics of coupled patches and demonstrated that demographic independency is realized as long as the relationship $m < 0.1$ is satisfied. Candidate examples that fall into this category are the many highly migratory stocks that spawn in multiple seasons, which often show no clear genetic differentiation, even when the sample sites or spawning grounds are spatially distant, for example yellowfin tuna (Anderson et al., 2019; Barth et al., 2017; Mullins et al., 2018). The second case involves genetic monitoring performed for each reproductive period (e.g. annually) because the proposed estimators consider information about contemporary migration that explicitly specified the timing (e.g. year); thus, time-series data for migration may reflect environmental changes, which would provide insights into the underlying ecological processes. In addition, HS pairs found within the same cohort in the genetic population provide N_b (Wang, 2009; Waples & Waples, 2011), which is also used to assess genetic health. The third case involves the development of integrated models that combine several data sources into a single analysis (Maunder & Punt, 2013). An example of this direction is to provide additional information about migration to ongoing CKMR projects for stock assessment in fisheries (e.g. Bravington, Grewe, & Davies, 2016; Hillary et al., 2018; Prystupa et al., 2021; Trenkel et al., 2022), which would utilize kinship pairs for simultaneous estimation of population parameters. If the population structure is defined hypothetically, the data accumulated in such projects could be readily connected to our theory via likelihood methods. The fourth case is where sampling adults is difficult due to conservation practices or other reasons, but sampling offspring is relatively easy, although N_2 (and N_1) must be given externally to employ \hat{m}_1 (and \hat{M}_1). In such cases, very few methods are available to estimate adult movement, especially when two populations cannot be genetically distinct. One exception is a method that estimates parent-offspring dispersal from the spatial distribution of close-kin dyads without pre-specification of the population structure (Jasper et al., 2019; Jasper et al., 2022), suggesting that CKMR-based methods have the potential to expand the scope of current population monitoring techniques.

The proposed estimators are limited to detecting parental movements in a period between breeding seasons in the given populations. Thus, the estimation of fine-scale spatiotemporal movements, which is available to integrate data sources (Thorson et al., 2021), is beyond the scope of our modelling framework. Although population dynamics models with coarser spatial resolution than the spatial scale of environmental layers are used frequently in assessment models, that is the target application of the proposed method, the proposed estimators require pre-specification of the population structure. Kinship relationships with sample location information potentially realize the ability to explore a plausible population structure by estimating the migration number or rate between hypothetical populations, which is expected to contribute to the determination of an effective management unit.

Finally, in the following, we discuss some caveats in terms of applying the proposed method. Our theory for developing the estimators depends on the assumption that kinships are detected without error. There are two primary situations in which inaccurate kinship detections are made. The first is that there is a large number of pairs in the avuncular genetic category (e.g. uncle-nephew) in the population because these pairs are genetically indistinguishable from the HS pair. These pairs are generated by particular life-history traits (e.g. females storing the sperm of a given male) causing a large number of full-sibling pairs (Trenkel et al., 2022). To apply the proposed method to such cases, it would be preferable to use only PO pairs (rather than HS pairs). The second situation is due to a lack of quality and quantity of genetic markers. Even if full-sibling pairs do not occur frequently, iteroparous species may have several kinship types, for example half-uncle-nephew or half-cousins, which are expected to appear frequently and should be accurately differentiated from HS pairs. It is also desirable to pre-estimate how many markers are required to realize kinship detection in the target populations associated with the simulation of pedigree reconstruction (e.g. Anderson, 2022). Even if the quality and quantity of genetic markers are increased, uncertainty about kinship detection remains. There are many algorithms that can be used to detect kinship pairs from single nucleotide polymorphisms or short tandem repeats (e.g. Huisman, 2017; Wang & Santure, 2009). It may be beneficial to obtain the upper and lower bounds of the proposed estimator using a combination of multiple algorithms to handle the uncertainty. In addition, if the kinship detection algorithm outputs the frequency of a kinship type for a given pair, the probability distribution of the proposed estimator is also available. These sensitivity analyses will be a future task. The current theory behind the estimators does not assume desynchronized reproduction within a population (e.g. skip spawning), the different distribution of reproductive potential between sexes, correlation between individual mobility and fertility, and nonrandom sampling of parents (e.g. mothers with a large number of offspring are likely to be sampled). These issues will also be the focus of future work.

AUTHOR CONTRIBUTIONS

Tetsuya Akita: Conceptualization (lead); formal analysis (lead); funding acquisition (lead); methodology (lead); writing-original draft (lead); writing-review and editing (lead).

ACKNOWLEDGEMENTS

The author thanks N. Shinohara, N. Suzuki and Y. Tsukahara for their fruitful discussions. The author also thanks W.M. Iwasaki for assisting with the C++ coding. Finally, the author thanks three anonymous reviewers for their constructive feedback, which substantially improved the quality of the manuscript. This work was supported by JSPS KAKENHI Grant Numbers 19K06862 and 20H03012.

CONFLICT OF INTEREST

The author declares no conflict of interest.

DATA AVAILABILITY STATEMENT

Code to conduct the simulations and reproduce the reported results is available at <https://doi.org/10.5281/zenodo.6668905>

ORCID

Tetsuya Akita  <https://orcid.org/0000-0003-4940-886X>

REFERENCES

- Akita, T. (2020a). Nearly unbiased estimator of contemporary effective mother size using within-cohort maternal sibling pairs incorporating parental and nonparental reproductive variations. *Heredity*, 124, 299–312. <https://doi.org/10.1038/s41437-019-0271-6>
- Akita, T. (2020b). Nearly unbiased estimator of contemporary n_e/N based on kinship relationships. *Ecology and Evolution*, 10(19), 10343–10352.
- Anderson, E. C. (2022). CKMRpop: Forward-in-time simulation and tabulation of pairwise kin relationships in age-structured populations. *Molecular Ecology Resources*, 22(3), 1190–1199. <https://doi.org/10.1111/1755-0998.13513>
- Anderson, G., Lal, M., Hampton, J., Smith, N., & Rico, C. (2019). Close kin proximity in yellowfin tuna (*Thunnus albacares*) as a driver of population genetic structure in the tropical western and central Pacific Ocean. *Frontiers in Marine Science*, 6, 341. <https://doi.org/10.3389/fmars.2019.00341>
- Ashley, M. V. (2010). Plant parentage, pollination, and dispersal: How DNA microsatellites have altered the landscape. *Critical Reviews in Plant Sciences*, 29(3), 148–161. Retrieved from. <https://doi.org/10.1080/07352689.2010.481167>
- Barth, J. M. I., Damerou, M., Matschiner, M., Jentoft, S., & Hanel, R. (2017). Genomic differentiation and demographic histories of Atlantic and Indo-Pacific yellowfin tuna (*Thunnus albacares*) populations. *Genome Biology and Evolution*, 9(4), 1084–1098. <https://doi.org/10.1093/gbe/evx067>
- Bravington, M. V., Grewe, P. M., & Davies, C. R. (2016). Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. *Nature Communications*, 7, 13162. <https://doi.org/10.1038/ncomms13162>
- Bravington, M. V., Skaug, H. J., & Anderson, E. C. (2016). Close-kin mark-recapture. *Statistical Science*, 31(2), 259–274.
- Conn, P. B., Bravington, M. V., Baylis, S., & Ver Hoef, J. M. (2020). Robustness of close-kin mark-recapture estimators to dispersal limitation and spatially varying sampling probabilities. *Ecology and Evolution*, 10(12), 5558–5569. <https://doi.org/10.1002/ece3.6296>
- Faubet, P., & Gaggiotti, O. E. (2008). A new Bayesian method to identify the environmental factors that influence recent migration. *Genetics*, 178(3), 1491–1504. <https://doi.org/10.1534/genetics.107.082560>
- Faubet, P., Waples, R. S., & Gaggiotti, O. E. (2007). Evaluating the performance of a multi-locus Bayesian method for the estimation of migration rates. *Molecular Ecology*, 16(6), 1149–1166. <https://doi.org/10.1111/j.1365-294X.2007.03218.x>
- Gagnaire, P.-A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S., & Bierne, N. (2015). Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications*, 8(8), 769–786. <https://doi.org/10.1111/eva.12288>
- Hampton, J., & Fournier, D. A. (2001). A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Marine and Freshwater Research*, 52(7), 937–963.
- Hastings, A. (1993). Complex interactions between dispersal and dynamics: Lessons from coupled logistic equations. *Ecology*, 74(5), 1362–1372. <https://doi.org/10.2307/1940066>
- Hilborn, R., & Walters, C. J. (1992). *Quantitative fisheries stock assessment: Choice, dynamics and uncertainty*. Chapman and Hall.
- Hillary, R. M., Bravington, M. V., Patterson, T. A., Grewe, P., Bradford, R., Feutry, P., Gunasekera, R., Peddemors, V., Werry, J., Francis, M. P., Duffy, C. A. J., & Bruce, B. D. (2018). Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. *Scientific Reports*, 8(1), 2661. <https://doi.org/10.1038/s41598-018-20593-w>
- Huisman, J. (2017). Pedigree reconstruction from SNP data: Parentage assignment, sibship clustering and beyond. *Molecular Ecology Resources*, 17(5), 1009–1024. <https://doi.org/10.1111/1755-0998.12665>
- Jasper, M., Schmidt, T. L., Ahmad, N. W., Sinkins, S. P., & Hoffmann, A. A. (2019). A genomic approach to inferring kinship reveals limited intergenerational dispersal in the yellow fever mosquito. *Molecular Ecology Resources*, 19(5), 1254–1264. <https://doi.org/10.1111/1755-0998.13043>
- Jasper, M. E., Hoffmann, A. A., & Schmidt, T. L. (2022). Estimating dispersal using close kin dyads: The kindisperse R package. *Molecular Ecology Resources*, 22(3), 1200–1212. <https://doi.org/10.1111/1755-0998.13520>
- Kéry, M., & Schaub, M. (2011). *Bayesian population analysis using winbugs: A hierarchical perspective*. Academic Press.
- Lowe, W. H., & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology*, 19(15), 3038–3051. <https://doi.org/10.1111/j.1365-294X.2010.04688.x>
- Maunder, M. N., & Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, 142, 61–74. <https://doi.org/10.1016/j.fishres.2012.07.025>
- Moritz, C. (1994). Defining 'evolutionarily significant units' for conservation. *Trends in Ecology & Evolution*, 9(10), 373–375.
- Mullins, R. B., McKeown, N. J., Sauer, W. H. H., & Shaw, P. W. (2018). Genomic analysis reveals multiple mismatches between biological and management units in yellowfin tuna (*Thunnus albacares*). *ICES Journal of Marine Science*, 75(6), 2145–2152. <https://doi.org/10.1093/icesjms/fsy102>
- Nielsen, R., Mattila, D. K., Clapham, P. J., & Palsbøll, P. J. (2001). Statistical approaches to paternity analysis in natural populations and applications to the North Atlantic humpback whale. *Genetics*, 157(4), 1673–1682.
- Paetkau, D., Calvert, W., Stirling, I., & Strobeck, C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, 4(3), 347–354. <https://doi.org/10.1111/j.1365-294X.1995.tb00227.x>
- Palsbøll, P. J., Bérubé, M., & Allendorf, F. W. (2007). Identification of management units using population genetic data. *Trends in Ecology and Evolution*, 22(1), 11–16. <https://doi.org/10.1016/j.tree.2006.09.003>
- Pearse, D., Eckerman, C., Janzen, F., & Avise, J. (2001). A genetic analogue of 'mark-recapture' methods for estimating population size: An approach based on molecular parentage assessments. *Molecular Ecology*, 10(11), 2711–2718.
- Piry, S., Alapetite, A., Cornuet, J.-M., Paetkau, D., Baudouin, L., & Estoup, A. (2004). GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, 95(6), 536–539. <https://doi.org/10.1093/jhered/esh074>
- Prystupa, S., McCracken, G., Perry, R., & Ruzzante, D. (2021). Population abundance in arctic grayling using genetics and close-kin mark-recapture. *Ecology and Evolution*, 11, 4763–4773. <https://doi.org/10.1002/ece3.7378>
- Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., Baughman, S., Cabin, R. J., Cohen, J. E., Ellstrand, N. C., McCauley, D. E., O'Neil, P., Parker, I. M., Thompson, J. N., & Weller, S. G. (2001). The population biology of invasive species. *Annual Review of Ecology and Systematics*, 32(1), 305–332. <https://doi.org/10.1146/annurev.ecolsys.32.081501.114037>
- Seber, G. A. F. (2002). *The estimation of animal abundance and related parameters* (2nd ed.). Griffin London.
- Skaug, H. J. (2001). Allele-sharing methods for estimation of population size. *Biometrics*, 57(3), 750–756.

- Skaug, H. J. (2017). The parent-offspring probability when sampling age-structured populations. *Theoretical Population Biology*, 118, 20–26.
- Thorson, J. T., Barbeaux, S. J., Goethel, D. R., Kearney, K. A., Laman, E. A., Nielsen, J. K., & Thompson, G. G. (2021). Estimating fine-scale movement rates and habitat preferences using multiple data sources. *Fish and Fisheries*, 22(6), 1359–1376. <https://doi.org/10.1111/faf.12592>
- Trenkel, V. M., Charrier, G., Lorance, P., & Bravington, M. V. (2022). Close-kin mark-recapture abundance estimation: Practical insights and lessons learned. *ICES Journal of Marine Science*, 79(2), 413–422. <https://doi.org/10.1093/icesjms/fsac002>
- Wang, J. (2009). A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular Ecology*, 18(10), 2148–2164.
- Wang, J. (2014). Estimation of migration rates from marker-based parentage analysis. *Molecular Ecology*, 23(13), 3191–3213. <https://doi.org/10.1111/mec.12806>
- Wang, J., & Santure, A. W. (2009). Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics*, 181(4), 1579–1594. <https://doi.org/10.1534/genetics.108.100214>
- Waples, R. S., & Feutry, P. (2022). Close-kin methods to estimate census size and effective population size. *Fish and Fisheries*, 23(2), 273–293. <https://doi.org/10.1111/faf.12615>
- Waples, R. S., & Gaggiotti, O. (2006). Invited review: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, 15(6), 1419–1439. <https://doi.org/10.1111/j.1365-294X.2006.02890.x>
- Waples, R. S., & Waples, R. K. (2011). Inbreeding effective population size and parentage analysis without parents. *Molecular Ecology Resources*, 11, 162–171. <https://doi.org/10.1111/j.1755-0998.2010.02942.x>
- Wilson, G. A., & Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, 163, 1177–1191. <https://doi.org/10.1093/genetics/163.3.1177>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Akita, T. (2022). Estimating contemporary migration numbers of adults based on kinship relationships in iteroparous species. *Molecular Ecology Resources*, 22, 3006–3017. <https://doi.org/10.1111/1755-0998.13682>

APPENDIX A

REPRODUCTIVE POTENTIAL

Here, we introduce the concept of the reproductive potentials of parents i and j in the populations 1 and 2, respectively, which are defined as the expected number of surviving offspring at the given sampling time, denoted by $\lambda_{i,1}$ and $\lambda_{j,2}$ ($i = 1, \dots, N_1$ and $j = 1, \dots, N_2$). Reproductive potential is determined by several factors, including the parent age, weight and residence time in the spawning ground. Note that the magnitude of this parameter includes information about offspring survival rate, the number of days after egg hatching

and the number of eggs. This implies that the parameter reflects the sample timing. Also note that the modelling framework does not depend on whether the reproductive potential is heritable or not.

APPENDIX B

DERIVATION OF EQUATIONS 1 AND 6 WHEN REPRODUCTIVE POTENTIAL IS VARIABLE AMONG PARENTS

In the main text, we ignore the variation of reproductive potential among parents (i.e. both $\lambda_{i,1}$ and $\lambda_{j,2}$ are constant) to derive $\pi_{\text{HS,bet}}$ and $\pi_{\text{PO,bet}}$. Here, let $k_{i,1}$ and $k_{j,2}$ be the number of surviving offspring of parents i and j at sampling time in populations 1 and 2, and assumed to follow a kind of discrete distribution (e.g. Poisson or negative binomial distribution) with mean $\lambda_{i,1}$ and $\lambda_{j,2}$, respectively. Without loss of generality, we set the index such that parents with $i = 1$ to M in population 1 and $j = 1$ to M in population 2 are identical migrants. For example, the parent with $i = 1$ reproduces $k_{1,1}$ offspring (in population 1) and then reproduce $k_{1,2}$ offspring (in population 2) after migration. Given $k_{i,1}$ and $k_{j,2}$, the conditional probability that two offspring sampled in populations 1 and 2 share a HS relationship is expressed as follows:

$$\pi_{\text{HS,bet}}|_{k_1,k_2} = \frac{2 \sum_{i=1}^M k_{i,1}}{\sum_{i=1}^{N_1} k_{i,1}} \frac{2 \sum_{j=1}^M k_{j,2}}{\sum_{j=1}^{N_2} k_{j,2}} \frac{1}{M}, \quad (\text{A1})$$

where $\mathbf{k}_1 = (k_{1,1}, \dots, k_{M,1}, \dots, k_{N_1,1})$ and $\mathbf{k}_2 = (k_{1,2}, \dots, k_{M,2}, \dots, k_{N_2,2})$. Note that $k_{i,1}$ and $k_{j,2}$ are random variables with mean $\lambda_{i,1}$ and $\lambda_{j,2}$, respectively. By taking the expectation over the distribution of the offspring number, the conditional probability is given approximately as follows:

$$\begin{aligned} \pi_{\text{HS,bet}}|_{\lambda_1,\lambda_2} &= \mathbb{E} \left[\pi_{\text{HS,bet}}|_{k_1,k_2} \right] \\ &= \frac{4}{M} \mathbb{E} \left[\frac{\sum_{i=1}^M k_{i,1}}{\sum_{i=1}^{N_1} k_{i,1}} \frac{\sum_{j=1}^M k_{j,2}}{\sum_{j=1}^{N_2} k_{j,2}} \right] \\ &\approx \frac{4}{M} \frac{\mathbb{E} \left[\sum_{i=1}^M k_{i,1} \sum_{j=1}^M k_{j,2} \right]}{\mathbb{E} \left[\sum_{i=1}^{N_1} k_{i,1} \sum_{j=1}^{N_2} k_{j,2} \right]} \\ &= \frac{4}{M} \frac{\sum_{i=1}^M \lambda_{i,1} \sum_{j=1}^M \lambda_{j,2}}{\sum_{i=1}^{N_1} \lambda_{i,1} \sum_{j=1}^{N_2} \lambda_{j,2}}, \quad (\text{A2}) \end{aligned}$$

where $\lambda_1 = (\lambda_{1,1}, \dots, \lambda_{M,1}, \dots, \lambda_{N_1,1})$ and $\lambda_2 = (\lambda_{1,2}, \dots, \lambda_{M,2}, \dots, \lambda_{N_2,2})$. From the second to third lines, we use the approximation that $\mathbb{E}[g_1(k)/g_2(k)] \approx \mathbb{E}[g_1(k)]/\mathbb{E}[g_2(k)]$. From the third to fourth lines, we

use the relationship that $\mathbb{E} \left[k_{i,1} k_{j,2} |_{\lambda_{i,1}, \lambda_{j,2}} \right] = \mathbb{E} \left[k_{i,1} |_{\lambda_{i,1}} \right] \mathbb{E} \left[k_{j,2} |_{\lambda_{j,2}} \right]$, which implies the statistical independence of the offspring number before and after parental movement. In other words, the conditional probability is not affected by $\mathbb{V}[k]_{\lambda}$. Here, we assume that $\lambda_{i,1}$ and $\lambda_{j,2}$ are also random variables that are followed by an arbitrary function with mean $\bar{\lambda}_1$ and $\bar{\lambda}_2$, respectively. By taking the expectation over λ and applying a similar approximation, the unconditional probability is given as follows:

$$\begin{aligned}
\pi_{\text{HS,bet}} &= \mathbb{E} \left[\pi_{\text{HS,bet}} | \lambda_1, \lambda_2 \right] \\
&= \frac{4}{M} \mathbb{E} \left[\frac{\sum_{i=1}^M \lambda_{i,1}}{\sum_{i=1}^{N_1} \lambda_{i,1}} \frac{\sum_{j=1}^M \lambda_{j,2}}{\sum_{j=1}^{N_2} \lambda_{j,2}} \right] \\
&\approx \frac{4}{M} \frac{\mathbb{E} \left[\sum_{i=1}^M \lambda_{i,1} \sum_{j=1}^M \lambda_{j,2} \right]}{\mathbb{E} \left[\sum_{i=1}^{N_1} \lambda_{i,1} \sum_{j=1}^{N_2} \lambda_{j,2} \right]} \\
&= \frac{4}{M} \frac{M^2 \bar{\lambda}_1 \bar{\lambda}_2}{N_1 N_2 \bar{\lambda}_1 \bar{\lambda}_2} \\
&= \frac{4M}{N_1 N_2}, \tag{A3}
\end{aligned}$$

which provides the same formulation described in Equation 1. Note that, from the third to fourth lines, we assume that $\lambda_{i,1}$ and $\lambda_{i,2}$ ($i = 1, \dots, M$) are independent variables (i.e. $\mathbb{E}[\lambda_{i,1} \lambda_{i,2}] = \bar{\lambda}_1 \bar{\lambda}_2$), which implies variable reproductive potential of an identical parent before and after migration.

Next, we derive the probability that offspring sampled in population 1 and a parent sampled in population 2 share a PO relationship ($\pi_{\text{PO,bet}}$) under the flexible settings of λ , which is similar to the derivation of $\pi_{\text{HS,bet}}$. The conditional probability is expressed as follows:

$$\pi_{\text{PO,bet}} | k_1, k_2 = \frac{2 \sum_{i=1}^M k_{i,1}}{\sum_{i=1}^{N_1} k_{i,1}} \frac{\sum_{j=1}^M k_{j,2}}{\sum_{j=1}^{N_2} k_{j,2}} \frac{1}{M}. \tag{A4}$$

Here, by taking the expectation over k and λ , in the same manner as noted above, the unconditional probability is approximately given as follows:

$$\begin{aligned}
\pi_{\text{PO,bet}} &= \mathbb{E} \left[\mathbb{E} \left[\pi_{\text{PO,bet}} | k_1, k_2 \right] \right] \\
&\approx \frac{2M}{N_1 N_2}, \tag{A5}
\end{aligned}$$

which provides the same formulation described in Equation 6.

APPENDIX C

DIFFERENCE IN REPRODUCTIVE POTENTIAL BETWEEN MIGRANTS AND NONMIGRANTS

In the derivation of $\pi_{\text{HS,bet}}$ and $\pi_{\text{PO,bet}}$ in Appendix B, we ignored co-variation between migration and reproductive potential. Here, we consider the case where migrants have a distinguishable distribution of reproductive potential from nonmigrant parents. Let $\lambda_{i,1}$, $\lambda_{j,2}$, $\lambda_{M,l,1}$ and $\lambda_{M,l,2}$ be the reproductive potential of nonmigrant parents in populations 1 and 2 and migrants in the populations 1 and 2 with mean $\bar{\lambda}_1$, $\bar{\lambda}_2$, $\bar{\lambda}_{M,1}$ and $\bar{\lambda}_{M,2}$, respectively. Under this setting, $\pi_{\text{HS,bet}}$ can be expressed as follows:

$$\begin{aligned}
\pi_{\text{HS,bet}} &\approx \frac{4}{M} \frac{\mathbb{E} \left[\sum_{l=1}^M \lambda_{M,l,1} \sum_{l=1}^M \lambda_{M,l,2} \right]}{\mathbb{E} \left[\left(\sum_{l=1}^M \lambda_{M,l,1} + \sum_{i=M+1}^{N_1} \lambda_{i,1} \right) \left(\sum_{l=1}^M \lambda_{M,l,2} + \sum_{j=M+1}^{N_2} \lambda_{j,2} \right) \right]} \\
&= \frac{4}{M} \frac{M^2 \bar{\lambda}_{M,1} \bar{\lambda}_{M,2}}{\left(M \bar{\lambda}_{M,1} + (N_1 - M) \bar{\lambda}_1 \right) \left(M \bar{\lambda}_{M,2} + (N_2 - M) \bar{\lambda}_2 \right)}. \tag{A6}
\end{aligned}$$

If $\bar{\lambda}_{M,1} = \bar{\lambda}_1$ and $\bar{\lambda}_{M,2} = \bar{\lambda}_2$, Equation A6 can be reduced Equation A3. This formulation includes reproductive potential terms such that the distinct reproductive potential between migrants and nonmigrants eliminates the usefulness of the proposed HS-based estimators (this is also applied to the proposed PO-based estimators).