ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Dispersal in a house sparrow metapopulation: An integrative case study of genetic assignment calibrated with ecological data and pedigree information

Dilan Saatoglu[1] | Alina K. Niskanen[1,2] | Markku Kuismin[3,4] | Peter S. Ranke[1] | Ingerid J. Hagen[1,5] | Yimen G. Araya-Ajoy[1] | Thomas Kvalnes[1] | Henrik Pärn[1] | Bernt Rønning[1] | Thor Harald Ringsby[1] | Bernt-Erik Sæther[1] | Arild Husby[1,6] | Mikko J. Sillanpää[3,4,7] | Henrik Jensen[1]

[1]Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway

[2]Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland

[3]Research Unit of Mathematical Sciences, University of Oulu, Oulu, Finland

[4]Biocenter Oulu, University of Oulu, Finland

[5]Norwegian Institute for Nature Research, Trondheim, Norway

[6]Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden

[7]Infotech Oulu, University of Oulu, Oulu, Finland

Correspondence
Henrik Jensen, Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway.
Email: henrik.jensen@ntnu.no

Abstract

Dispersal has a crucial role determining ecoevolutionary dynamics through both gene flow and population size regulation. However, to study dispersal and its consequences, one must distinguish immigrants from residents. Dispersers can be identified using telemetry, capture-mark-recapture (CMR) methods, or genetic assignment methods. All of these methods have disadvantages, such as high costs and substantial field efforts needed for telemetry and CMR surveys, and adequate genetic distance required in genetic assignment. In this study, we used genome-wide 200K Single Nucleotide Polymorphism data and two different genetic assignment approaches (GSI_SIM, Bayesian framework; BONE, network-based estimation) to identify the dispersers in a house sparrow (*Passer domesticus*) metapopulation sampled over 16 years. Our results showed higher assignment accuracy with BONE. Hence, we proceeded to diagnose potential sources of errors in the assignment results from the BONE method due to variation in levels of interpopulation genetic differentiation, intrapopulation genetic variation and sample size. We show that assignment accuracy is high even at low levels of genetic differentiation and that it increases with the proportion of a population that has been sampled. Finally, we highlight that dispersal studies integrating both ecological and genetic data provide robust assessments of the dispersal patterns in natural populations.

KEYWORDS
dispersal, genetic assignment, house sparrow, metapopulation, SNP

---

1 | INTRODUCTION

Dispersal is an important life-history trait and a key process determining both ecological and evolutionary dynamics of populations as well as their conservation status (Bonte & Dahirel, 2017; Van Dyck & Baguette, 2005). Dispersal is defined as permanent movement of individuals from birthplace to reproduction site (natal dispersal) or between reproduction sites (breeding dispersal; Greenwood, 1980; Ronce, 2007). When dispersers reproduce, dispersal leads to gene flow between populations which is expected to result in increased genetic diversity within a population and decreased genetic differentiation between populations (Holsinger & Weir, 2009). Gene flow may prevent the fixation of deleterious mutations by counteracting genetic drift and introducing new genetic variation. Immigration may, thereby, increase mean fitness and adaptive potential of populations, which in turn is expected to improve their long-term viability (Whiteley et al., 2015). However, high rates of gene flow may also impede local adaptation and have a negative effect on population persistence (Alleaume-Benharira et al., 2006; Berdahl et al., 2015; Berg et al., 2010). In addition, dispersal is expected to have direct effects on population demography; while immigrants may increase the local population size, emigrants may reduce it (Millon et al., 2019; Vance, 1984), depending on the relationships between population density and emigration/immigration rates. Consequently, dispersal plays a vital role in the ecological dynamics of structured populations such as metapopulations (Hanski & Gaggiotti, 2004). Thus, understanding the causes and consequences of dispersal is fundamental to be able to predict the short- and long-term viability of populations in rapidly changing environments (Akçakaya, 2000; Lowe et al., 2017; Travis et al., 2013).

To study dispersal, it is essential to collect high-quality dispersal data where dispersers are correctly distinguished from residents. This poses considerable challenges in many organisms because the study area needs to be large enough to cover normal dispersal distances (Armansin et al., 2020; Matthysen, 2012). The movement of a high proportion of individuals within and between a set of study populations should be monitored, and monitoring must start before the dispersal event and continue for a sufficiently long time period to include the settlement and establishment phase (Holyoak et al., 2008). Telemetry and capture-mark-recapture (CMR) methods are commonly used ecological monitoring tools to identify dispersing individuals. Telemetry methods may provide accurate information on movement, dispersal distance and direction (Cagua et al., 2015; Hayden et al., 2014; Iwajomo et al., 2018; Mauritzen et al., 2002). However, high cost of electronic telemetry devices and the need for extensive field efforts often result in data with relatively small sample sizes that cover a restricted geographic area and span short time periods, relative to dispersal distances and the length of dispersal and establishment processes (Cayuela et al., 2018). Similarly, implementing CMR surveys often have limitations because of logistic challenges and high costs due to the substantial field effort needed (Muriel et al., 2015; Truve & Lemel, 2003). Detecting dispersers using telemetry and CMR methods may therefore not be sufficient to obtain unbiased information on individual dispersal events and dispersal dynamics in many systems.

The use of genetic clustering methods that delineate the genetic structure of populations and genetic assignment methods that determine the origin of individuals based on genetic clusters has increased during the last two decades as a result of decreased costs of genotyping and sequencing in non-model organisms (Chen et al., 2018; Corander et al., 2003; François & Waits, 2015; Pritchard et al., 2000). Compared to ecological monitoring by use of telemetry and CMR, genetic assignment methods may offer a less labor-intensive and more cost-efficient way to identify dispersers. In principle, conventional genetic assignment methods assign the individuals of interest to the most likely populations/groups based on the expected probabilities of belonging (genetically) to each of the predefined source populations (Anderson et al., 2008; Manel et al., 2005; Paetkau et al., 2004). Although genetic assignment has been successfully used in a number of studies of dispersal in natural population of vertebrates (Riley et al., 2006; Roffler et al., 2014; Schwartz et al., 2002), plants (Hanaoka et al., 2014; Orantes et al., 2012; Sinclair et al., 2018), and insects (Marchi et al., 2013; Vanden Broeck et al., 2017), there are some challenges and limitations. For instance, reduced accuracy due to low genetic differentiation (low F_{ST} values) and unbalanced sample sizes from sampled clusters, which both are common characteristics of empirical data from natural populations, are challenges for most genetic approaches (Araujo et al., 2014; Broquet & Petit, 2009; Paetkau et al., 2004; Putman & Carbone, 2014). Additionally, many of the existing genetic assignment methods have computational limitations and are not suitable for high numbers of molecular markers such as single nucleotide polymorphisms (SNPs; Piry et al., 2004; Pritchard et al., 2000). Some recently developed genetic assignment tools (Chen et al., 2018; Kuismin et al., 2020; Moran & Anderson, 2019) seem to deal with some of the potential biases and challenges caused by, for instance, high-density marker genotype data and unbalanced sample sizes. However, to better understand whether and when genetic assignment methods provide accurate assignments, we do not only need simulation studies but also high-quality empirical data sets from structured populations with information on many resident and dispersing individuals with known natal and adult populations, which can be used to explore the effects of sampling and population characteristics on the assignment performance (Cayuela et al., 2018).

In this study, we combine genetic assignment results and ecological CMR data to reveal the true dispersers and explore dispersal patterns in a long-term study of an insular house sparrow metapopulation off the coast of northern Norway (Baalsrud et al., 2014; Pärn et al., 2012; Ringsby et al., 2002). To do so, we use two genetic assignment methods implementing different model-based frameworks: (i) a commonly used genetic stock identification method based on Bayesian inference "GSI_SIM" (Anderson et al., 2008; see Moran & Anderson, 2019 for implementation of GSI_SIM in a recently released R package "RUBIAS"), and (ii) a recently developed network-based estimation method, BONE (Kuismin et al., 2020). We selected these methods because they can process high-density

genome-wide SNP genotype data for a large number of individuals and they performed well for a subset of our data in Kuismin et al., (2020). First, we tested the accuracy of these genetic assignment methods and investigated the possible sources of error using individuals with known natal population and dispersal status. These analyses are based on individual CMR data on recruiting offspring, as natal dispersal occurs during the first autumn in an individual's life in the study metapopulation. Second, we also investigated the effects of population genetic differentiation, heterozygosity, population size, and proportion of population sampled on the accuracy of the genetic assignment to portray when and how genetic assignment methods could and should be used. For example, the reliability of genetic assignment is expected to be poor with small sample sizes from potential source populations or low genetic differentiation between source populations (Araujo et al., 2014; Kalinowski 2004). Our results may provide general guidelines about when genetic assignment methods are expected to lead to erroneous assignments. Third, we constructed an extensive dispersal data set by combining results from the genetic assignment method that performed best with ecological CMR data and parentage information from a SNP-based metapopulation-level pedigree (Niskanen et al., 2020). We used this combined dispersal data set to examine some patterns of variation in dispersal in the study metapopulation: (i) Does the proportion of

dispersing recruits vary across years? If so, such temporal variation could indicate annual variation in environmental conditions that affect emigration probability and/or establishment success (Bowler & Benton, 2005). (ii) Are there differences between populations and/or habitat types in proportions of dispersing recruits, which may suggest the existence of source-sink dynamics in the study metapopulation (Dias, 1996)? (iii) Is the proportion of dispersers female-biased, as found in most bird species (Clarke et al., 1997; Pusey, 1987)?

2 | MATERIALS AND METHODS

2.1 | Study metapopulation and data collection

The study was carried out in an insular house sparrow (*Passer domesticus*) metapopulation in the Helgeland archipelago in northern Norway, which has been monitored since 1993 (Figure 1). The study metapopulation consists of 18 island populations, interconnected almost exclusively by natal dispersal. Moreover, the study area spans over more than 1600 km^2 and covers a large area relative to the average house sparrow dispersal distances (10–15 km in our study system; Baalsrud et al., 2014; Pärn et al., 2012; Tufto et al., 2005; Ranke et al., 2021). Data were collected every year (from 1993 to present)
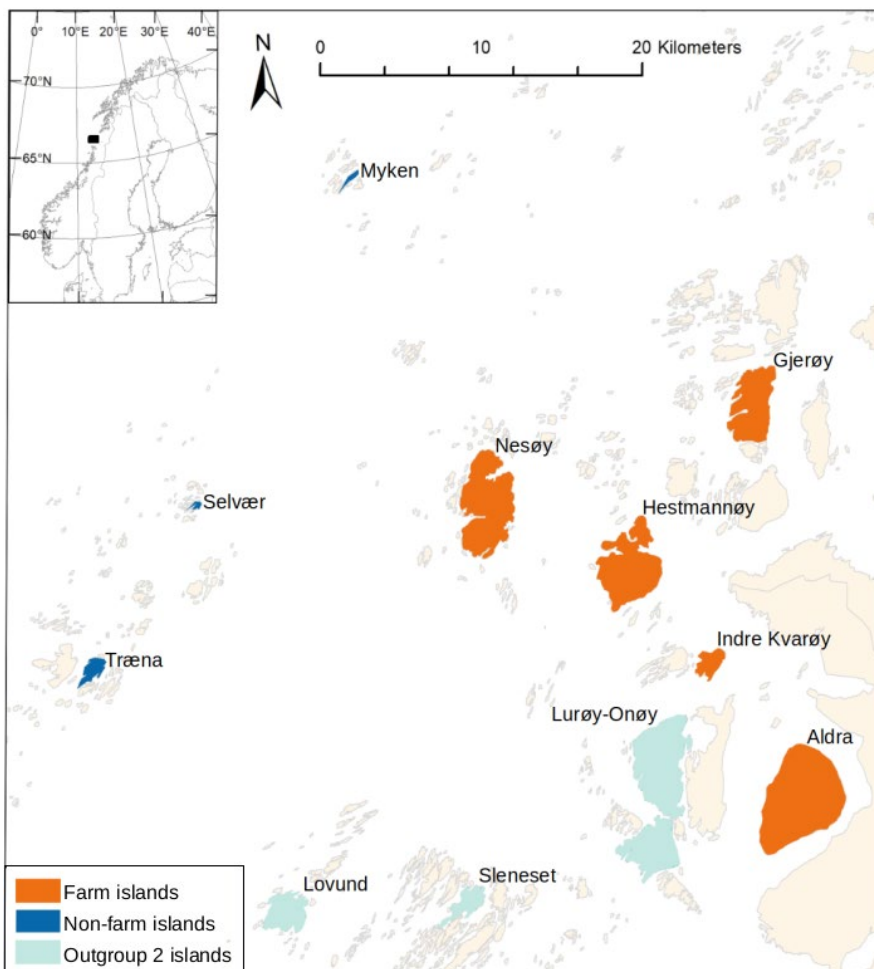

FIGURE 1 The study area in the Helgeland archipelago in northern Norway. Nonfarm islands (dark blue), farm islands (orange) and out-group 2 islands (teal) used in this study are shown

in the breeding season from May to August and in the autumn between mid-September and mid-November. Offspring were ringed as nestlings (in nest boxes or cavities under the roofs) in the summer season, whereas pre-/post-moult juveniles and any unringed recruits were captured using mist nets during summer and autumn seasons. Individuals were marked with a metal ring with a unique number and three coloured plastic rings to provide a unique colour combination. Most (~90%) house sparrows in the study metapopulation were individually marked, and the average resighting/recapture rate in the study metapopulation was 74% (Niskanen et al., 2020). Blood samples (~25 μl) were collected from the brachial vein located under the wing for genotyping. These extensive fieldwork procedures ensured high recapture rates, as well as high-quality data on recruitment, annual survival of adult birds, and dispersal within the study system (Holand et al., 2016; Jensen et al., 2004, 2008; Kvalnes et al., 2018; Pärn et al., 2009, 2012; Ringsby et al., 2002; Stubberud et al., 2017). House sparrows are sedentary species and they are only found around human settlements or agricultural areas which makes them easy to monitor (Summers-Smith, 1988).

A custom Affymetrix Axiom 200,000 SNP array was developed by using the house sparrow reference genome (Elgvin et al., 2017) and whole genome resequencing data from 33 house sparrows sampled at different localities in Norway ($n = 29$) and Finland ($n = 4$; Lundregan et al., 2018). A total of 3253 individuals recorded as adults on a subset of islands in our study metapopulation between the years 1998–2013 were genotyped with this custom Affymetrix Axiom SNP array (Figure 1; Table 1). Based on Affymetrix' MonoHigh and PolyHigh quality criteria, 185 587 SNPs were passed on to further quality control (Lundregan et al., 2018). After removing the birds with low sample quality (genotyping rate < 0.90, $n = 68$), potential duplicate samples based on identity by state (IBS) above 0.98 ($n = 20$), and excluding the loci likely to have a relatively high level of genotyping errors (SNP call rate < 95%; Mendelian error rate based on parental relationships > 5%) or low minor allele frequency (MAF < 0.01), 3116 individuals and 183,145 SNPs were found suitable for further analyses (Niskanen et al., 2020).

On five islands (Aldra, Gjerøy, Hestmannøy, Indre Kvarøy and Nesøy), we SNP-genotyped adults sampled from 1998 to 2013 (Table 1). These islands are denoted with the habitat type "farm island" because sparrows mainly live and breed at or near dairy farms. On three other islands (Myken, Selvær, and Træna) we SNP-genotyped adults sampled from 2003 (Selvær, Træna) and 2004 (Myken) to 2013 (Table 1). On these three islands, denoted with the habitat type "nonfarm island", there are no dairy farms, and the house sparrows breed in nest boxes or in cavities under the roofs of houses and forage in gardens. Because birds on farm islands live on farms, they experience higher local densities (in "colonies") during the breeding season, but also have better access to shelter during bad weather conditions (especially in winter) compared to birds on nonfarm islands (see Araya-Ajoy et al., 2019; Pärn et al., 2012). Because house sparrows may be expected to move more around on nonfarm islands, and any effects of population size (e.g., density) may be stronger on farm-islands, we expect that dispersal patterns may differ between habitat types. Population sizes were estimated as described in Baalsrud et al. (2014) and Stubberud et al. (2017).

2.2 | Dispersal data

We defined a disperser as a bird that emigrated from its natal island and recruited on another island based on CMR (i.e., initially ringed on natal island and subsequently recaptured after recruitment). Natal dispersal comprises nearly all dispersal in the metapopulation (only c. 0.2% of recruited birds perform breeding dispersal as adults; Altwegg et al., 2000; Tufto et al., 2005). In the study metapopulation, natal dispersal usually occurs between early autumn and early winter (i.e., August–December). Hence, many dispersal events have therefore already occurred during the period between the summer and autumn fieldwork (i.e., between approximately mid-August and October). Consequently, ecological CMR data can only be used to determine individuals' natal island with certainty for nestlings and fledged juveniles captured during summer fieldwork (May–August). For individuals captured for the first time in the autumn (any unringed post-moult juveniles or adults), natal island cannot be determined based on the ecological data (i.e., the island of first capture may in some cases not be the bird's natal island). Our data set includes 1645 recruits with known natal island; 1354 of these recruited on their natal island (i.e., residents) and 291 dispersed to another island before recruitment (i.e., dispersers). Furthermore, there were 1096 recruits for which the recruitment island but not the natal island was known; some of these birds could be dispersers that we were not able to detect using our ecological CMR data.

2.3 | Genetic assignment methods

Our aim was to construct a complete and accurate dispersal data set for the house sparrow study system. Therefore, we generated genetic assignments of individuals to their natal island with two different approaches and compared the assignment accuracy with known natal island and dispersal information generated from our extensive CMR data. We utilized: (i) the GSI_SIM software, which is based on the conditional maximum likelihood algorithm of the ONCOR software but further developed to handle a high number of markers (Anderson et al., 2008; see Moran & Anderson, 2019 for an R package "RUBIAS" which is equivalent to GSI_SIM), and (ii) the recently developed BONE method (Kuismin et al., 2020). BONE uses sparse multinomial least absolute shrinkage and selection operator (LASSO) regression, and is based on estimating network communities from genetic data (see also CONE "Community Oriented Network Estimation" method in Kuismin et al., 2017). GSI_SIM uses an "individual assignment test" to assign every individual in a "mixture population" (i.e., individuals with unknown population origin) to the population among a set of "baseline populations" (i.e., samples of individuals from pre-defined potential populations of origin, that should not include the individuals in the "mixture population") that

TABLE 1 House sparrow populations included in the current study with information on whether or not it was selected for high-density SNP-genotyping, its habitat type (Type) and its classification in the genetic assignment analyses; a population was either included as a separate baseline population per year or in one of the two outgroup baseline populations for the time periods 1998–2003 and 2004–2012

| Island | Selected for SNP-genotyping | Type | Separate baseline population per year | | Total number of adults SNP-genotyped |
			Years 1998–2003	Years 2004–2012	
Aldra	Yes	Farm	Yes	Yes	181
Indre Kvarøy	Yes	Farm	Yes	Yes	344
Hestmannøy	Yes	Farm	Yes	Yes	1014
Gjerøy	Yes	Farm	Yes	Yes	528
Nesøy	Yes	Farm	Yes	Yes	122
Selvær	Yes	Nonfarm	No (included in outgroup 1)	Yes	247
Træna	Yes	Nonfarm	No (included in outgroup 1)	Yes	223
Myken	Yes	Nonfarm	No (included in outgroup 1)	Yes	79
Lovund	No	Nonfarm	-	No (included in outgroup 2)	24
Sleneset	No	Nonfarm	-	No (included in outgroup 2)	22
Lurøy-Onøy	No	Farm	-	No (included in outgroup 2)	25

Total number of SNP-genotyped adult individuals indicates the number of genotyped adults for the islands selected for SNP-genotyping (i.e., total number of unique adult individuals in all the baseline populations for that island), whereas for the islands not selected for SNP-genotyping (outgroup 2) it indicates the number of individuals that hatched on the island (as determined by ecological CMR data) and emigrated to one of the islands selected for SNP-genotyping.

the individual had the highest posterior probability of belonging to given its multilocus genotype and allele frequencies in the baseline populations (Anderson et al., 2008). BONE uses similar "mixture" and "baseline" populations, but is a multiphasic approach that uses model and tuning parameter selection for estimating undirected genetic relationship networks based on high-density genotype data (Kuismin et al., 2020). The BONE approach then uses the nodes and edges from the estimated networks to infer the probability of an individual belonging to each of the different baseline populations, calculated from the individual's node degrees and number of neighbours. The individual is then assigned to the baseline population which it has the highest probability of belonging to. Importantly, the BONE approach can also identify individuals originating from a population not among any of the baseline populations.

Because sampling of the different island groups for high-density SNP-genotyping started in different years, data for the baseline and mixture populations were assembled in four different stages when utilizing the two genetic assignment methods (see Table 1). First, adults from farm islands were included in the baseline populations for each of the years 1998–2012, since data were available for these islands throughout all these study years. Second, to be able to identify immigrants from nonfarm islands before 2004, that is, before the systematic sampling for these islands started (Træna and Selvær in 2003; Myken in 2004), we created an outgroup (denoted as "outgroup 1"; Table 1). "Outgroup 1" consisted of a small set of individuals (27 < n < 74) which were present on these three nonfarm islands in the years 1998–2003 (Table S1). This outgroup was included as one of the baseline populations for annual assignment analyses before

2004. For the assignment analyses for 2004 and later, each of the three nonfarm islands (Træna, Selvær and Myken) were included as a separate baseline population. Third, a new outgroup was created for the years 2004–2012 (denoted as "outgroup 2"; Table 1), including SNP-genotyped individuals that had hatched on an island within the metapopulation that was not actively selected for genotyping (outgroup 2 islands: Lovund, Lurøy-Onøy and Sleneset). The outgroup 2 baseline populations were constructed separately for each year by pooling the SNP-genotyped adults (23 < n < 66) that were known to have hatched on any of the three islands (based on ecological data) from 1997 up to the year of the analysis (e.g., the outgroup 2 baseline population for 2004 was constructed by including the adult individuals that were ringed as nestlings or juveniles May–August on Lovund, Lurøy-Onøy or Sleneset from 1997 to 2003). Although individuals present on the islands and years included in the two outgroups had not been selected for high-density SNP-genotyping to represent samples from the outgroups, some genotyped adult individuals from these islands and years were present in our data set because they had dispersed from one of the outgroup islands to an island where adult individuals were genotyped (see Table 1). Finally, as a last step, any duplicates that occurred both in the baseline and pooled set of one year old recruits (i.e., mixture populations; see below for details) due to the way we constructed the outgroups, were removed from baseline populations in order to avoid self-assigning of these individuals.

After preparing the data, we ran the assignment analyses according to the following procedure: a mixture population was assembled separately for each year ($t + 1$) by pooling all the recruited individuals

in our data that were born in the previous year (t). Every individual in the mixture population was then assigned to the baseline populations for year t using the two genetic assignment methods described above. The assumption was that one of these baseline populations should include parents and/or other ancestors and relatives of the recruit (see Figure S1). Note that for the individuals in the mixture population, no population-specific information was provided in the analyses. This procedure was carried out for each year (t) 1998–2012.

2.4 | Accuracy of the genetic assignment methods

In order to evaluate the accuracy of the genetic assignment analyses carried out using GSI_SIM and BONE, the natal island and dispersal status obtained from each genetic assignment method were compared with the individuals' known natal island and dispersal status obtained from the CMR data. In these analyses we used the 1645 (60%) recruits for which ecological CMR data can reliably be used to determine natal island (i.e., nestlings and fledged juveniles captured during the summer season). If the natal and "adult" islands (i.e., the island on which they recruited; see above) were the same, the individual was defined as a "resident". Whereas in cases where natal and adult islands of an individual differed, the individual was regarded as a "disperser". These criteria were used both for the ecological CMR data and the genetic assignment data. We distinguished two types of misassignments. We refer to false positives (type I error), when the ecological CMR data showed that the individual was a resident and the genetic assignment wrongly indicated it was a disperser, and false negatives (type II error) when the ecological CMR data showed the individual was a disperser and the genetic assignment wrongly indicated it was a resident.

2.5 | Causes of errors in genetic assignment

After determining which method yields the highest assignment accuracy, causes of any errors in genetic assignments were investigated using assignment results from the approach performing best. We tested whether the annual number of misassignments per population or population pair were related to genetic and ecologic population characteristics such as genetic differentiation between pairs of populations, within-population genetic diversity and sample size. Furthermore, the effect of the proportion of the population sampled (i.e., the proportion of the adult population in year t that was successfully SNP-genotyped and included in the baseline population that year) on the number of errors was also tested to examine the effect of sampling regime on the accuracy of genetic assignment. To investigate the relationships between these variables and genetic assignment accuracy, statistical models were fitted for each variable separately and together (all variables except genetic differentiation, to account for any effects of the other variables) using the "glmmTMB" package in R (Brooks et al., 2017). Correlations among the explanatory variables were examined using the "psych" package in R

(Revelle, 2018; Figure S9). Additionally, we carried out similar analyses using proportion of misassignments instead of number of misassignments (see Supporting Information Materials for description of methods and results from these analyses). Parameter estimates are given ±1 SE.

In order to estimate pairwise genetic differentiation (F_{ST}) between baseline populations, and the genetic diversity (i.e., observed heterozygosity) within baseline populations, we used a subset of the 183,145 SNPs that were highly variable and unlinked to decrease computational time. The subset was generated using PLINK 1.09 (Purcell et al., 2007) and consisted of 5807 SNPs with high MAF (>0.2) that were in low linkage disequilibrium (LD; variance inflation factor < 1.04 using a sliding window approach; 50 SNPs of window size and window overlap of five SNPs) across all islands and years combined.

2.5.1 | Pairwise number of misassignments vs. pairwise genetic differentiation

Pairwise genetic differentiation between baseline populations was estimated for each year t based on the F_{ST} estimator of Weir and Cockerham (1984) implemented in the R-package "hierfstat" (Goudet, 2005). Misassignments between pairs of populations were calculated for each year t by taking the sum of misassignments for a given pair of baseline populations that year (i.e., the number of birds known to have hatched in each of the two population with the other population wrongly assigned as their natal population). To examine the effect of genetic differentiation between population pairs on number of pairwise misassignments, we fitted a zero-inflated Poisson generalized linear mixed model (ZIP GLMM), where pairwise number of misassignments was fitted as the response and genetic differentiation (F_{ST}) as a continuous covariate. We included island pair identity as a random factor to account for nonindependence of island pairs across years.

2.5.2 | Number of misassignments vs. observed heterozygosity

To assess the effect of genetic diversity within a baseline population on the number of misassignments, we calculated observed mean heterozygosity across SNPs using the R-package hierfstat (Goudet, 2005). In order to separate and identify the reasons behind the reduced accuracy in the genetic assignment method correctly, the number of misassignments were calculated separately for each year t in each baseline population in two ways; (i) the sum of misassignments "given away" (i.e., individuals known to have hatched in the focal population in year t based on ecological CMR data, but assigned to another baseline population), and (ii) the sum of misassignments "received" (i.e., individuals assigned to the focal baseline population in year t, but known to have hatched in another population based on ecological CMR data). The number of misassignments

was used as the response variable in two sets of models (misassignments given away in one model, misassignments received in another) where the observed heterozygosity was fitted as a continuous covariate in a ZIP GLMM.

2.5.3 | Number of misassignments vs. sample size of baseline populations

The effect of baseline populations' sample size (i.e., number of adults in the baseline population in year t that were SNP-genotyped) on the number of misassignments was examined using annual data for each baseline population. Number of misassignments (given away and received in separate models, as described above) was fitted as the response variable whereas sample size was a continuous covariate in a ZIP GLMM.

2.5.4 | Number of misassignments vs. proportion of the population included in the baseline

Finally, we examined how the proportion of the sampled adult population included in the baseline population ("baseline proportion") affected the number of misassignments. Baseline proportion was estimated for each island and year t separately by dividing the number of individuals included in a baseline population with the population size estimate for that island and year. ZIP GLMMs were fitted with separate models for misassignments given away and misassignments received. In both models, baseline proportion was included as a continuous covariate.

2.6 | Constructing an extended dispersal data set

To determine the natal island and dispersal status of the 1096 (40%) recruits with unknown natal island based on the ecological CMR data set (i.e., unringed new recruits captured in year $t + 1$ and recruits first captured as fledged juveniles in the autumn season of year t), we used results from the genetic assignment method with higher assignment accuracy.

Furthermore, to verify and correct the genetically assigned natal island and dispersal status of the individuals with unknown natal island based on the ecological CMR data set, we utilized a metapopulation-level pedigree that was recently constructed based on parentage analyses using high-density SNP-genotype data (Niskanen et al., 2020).

Verifications and corrections based on the pedigree could be made for the 743 recruits with at least one parent in the pedigree (the other 353 recruits with unknown natal island based on the ecological CMR data also had unknown parents). For these 743 recruits, each recruit's genetically assigned island was compared with the island on which its mother and/or father was present in the year that the recruit was born (based on the ecological CMR data). In cases

where the genetically assigned natal island of a recruit did not match with the adult island of its mother and/or father, the genetically assigned natal island was assumed to be wrong, and the true natal island was assumed to be the island on which its parent(s) were present in its year of birth. Finally, when the verifications of the recruits' natal island had been carried out, each recruit's dispersal status was determined by comparing its (corrected) natal island with its adult island. When a similar check against the pedigree was done for the 1460 recruits (out of 1645) with known hatch island based on the ecological CMR data and at least one known parent, we found that only two individuals (0.14%) had hatched on an island that did not match with the adult island where their mother and/or father were recorded.

2.7 | General dispersal patterns in the house sparrow metapopulation

To examine (i) the temporal and spatial variation in interisland dispersal, (ii) proportion of recruits that dispersed between islands within the same and different habitat types (farm islands and nonfarm islands), and (iii) differences in the sex-ratio between dispersers and residents, generalized linear mixed models with a binomial error distribution were fitted using the "glmmTMB" package in R (Brooks et al., 2017). We expect spatial variation in dispersal because previous results from the study metapopulation suggest that there are differences between islands and/or habitat types (farm and nonfarm islands) in morphological traits and their heritable genetic basis (Araya-Ajoy et al., 2019; Muff et al., 2019) as well as in demographic parameters (Holand et al., 2016; Pärn et al., 2012; Ringsby et al., 1999, 2002; Sæther et al., 1999). All sampling years were used in investigation of spatiotemporal variation and differences in the sex-ratio. However, because the sampling period for nonfarm islands did not include years before 2004, we chose to use only data from years 2004–2012 in the examination of differences between the habitat types. Finally, random intercepts were included for natal island and year when investigating whether the sex-ratio differed between dispersers and residents.

3 | RESULTS

3.1 | Accuracy of the genetic assignment methods

We used GSI_SIM and BONE genetic assignment methods to assign a total of 2741 recruits to their most likely natal island population within our metapopulation. When testing the accuracy of the assignment methods, we only used individuals with known natal island based on the ecological CMR data ($n = 1645$; Figure S2).

Based on the recruits with known natal island and dispersal status, the error rates for GSI_SIM were 9.5% (total), 7.3% (type I), and 2.2% (type II), whereas corresponding error rates for BONE were 4.8%, 3.6%, and 1.2% (see Supporting Information Material

and Figures S3 and S12 for details on GSI_SIM assignment results). Because the BONE method gave considerably lower error rates, and thus more accurate results than GSI_SIM, we constructed our new dispersal data set using the BONE genetic assignment method.

In the BONE analysis we used the so called "Winner Takes it All" (WTA) approach that assigns a given individual to the most probable genetic network, and can also detect any immigrants from any unsampled populations not included among the baseline populations (see Kuismin et al., 2020 for details). However, BONE did not identify any such individuals in our data set.

3.2 | Causes of errors in genetic assignment

3.2.1 | Pairwise number of misassignments vs. pairwise genetic differentiation

Among 414 population pair-years, 85.5% had no misassignments and 9.2% had only one misassigned individual (Figure 2). The maximum number of misassigned individuals between two populations was 7 (between Træna and Selvær in 2011, with pairwise $F_{ST} = 0.005$), and 53% of all misassignments were between populations with pairwise $F_{ST} < 0.013$. Accordingly, there was very strong evidence that genetic differentiation, estimated as annual pairwise F_{ST}, explained variation in number of misassignments between population pairs ($\chi^2_{(1)} = 18.534$; $p < .001$). The number of misassignments for a given population pair increased when populations were genetically similar, in other words, when pairwise F_{ST} values were lower ($\beta = -54.792 \pm 12.125$; $p < .001$; Figure 2; see Table S6 and Figure S10 for the impact of pairwise F_{ST} on pairwise proportion of misassignments). For example, the model predicts on average 0.16 misassigned individuals per year between a pair of islands, if the pairwise F_{ST} among the islands is equal to the mean in our study metapopulation ($F_{ST} = 0.036$), and 0.51 misassigned individuals between islands when pairwise $F_{ST} = 0.010$).

3.2.2 | Number of misassignments vs. observed heterozygosity, sample size of baseline populations, and proportion of the population included in the baseline

To examine how within-population characteristics may cause errors in genetic assignments, we tested whether the number of misassignments in a given year (misassignments given away and received) was related to the population's level of heterozygosity (mean: 0.415; range: 0.391–0.433), sample size (mean: 51.431; range: 2–231), and proportion of the population that was included in the baseline population (mean: 0.884; range: 0.429–1).

There was evidence for positive relationship between the number of misassignments (both given away and received) and the observed heterozygosity of baseline populations (p-values <.05; Figure 3a, b; Table S3; Table S4). Furthermore, there was a positive relationship between the number of misassignments received and sample size of the baseline population ($p < .05$; Figure 3d; Table S4). Finally, we found that as the proportion of the sampled adult population increased, there were fewer misassignments given away (Figure 3e, Table S3).

Results were generally similar when models were fitted with all three population characteristics as covariates simultaneously, to test whether each of them explained variation in the number of misassignments (given away and received) when any effects of the other two characteristics were controlled for. The positive relationship between observed heterozygosity and misassignments given away was still present ($\beta = 69.540$, $p < .05$, Table S5), whereas there now was little evidence that heterozygosity had an effect on misassignments received ($\beta = 41.758$, $p = .126$, Table S5). Additionally, the relationship between sample size and the assignment accuracy did not change ($\beta = -0.192$, $p = .363$ for misassignments given away; $\beta = 0.553$, $p < .01$ for misassignments received). However, these analyses suggested that the proportion of the adult population sampled had a negative effect on the assignment accuracy for both misassignments given and received ($\beta = -1.939$, $p = .015$ for
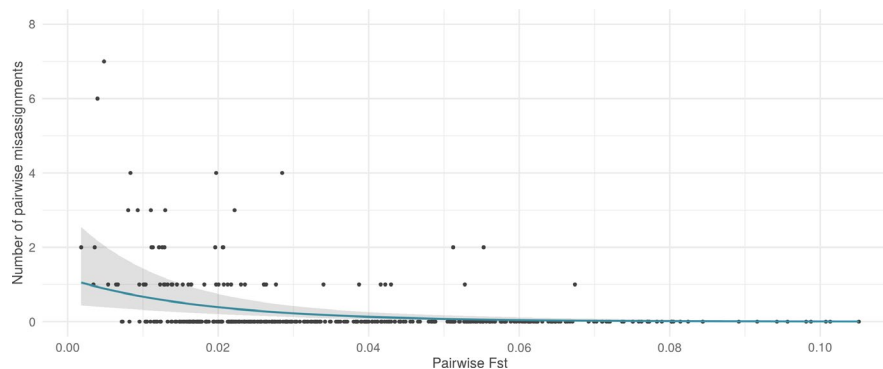

FIGURE 2 Relationship between the annual number of genetically misassigned recruits and F_{ST} for each population pair in an insular house sparrow metapopulation. The line shows the predicted values for the number of pairwise misassignments, and the grey area shows its 95% confidence interval from a ZIP GLMM where island pair ID was included as a random factor
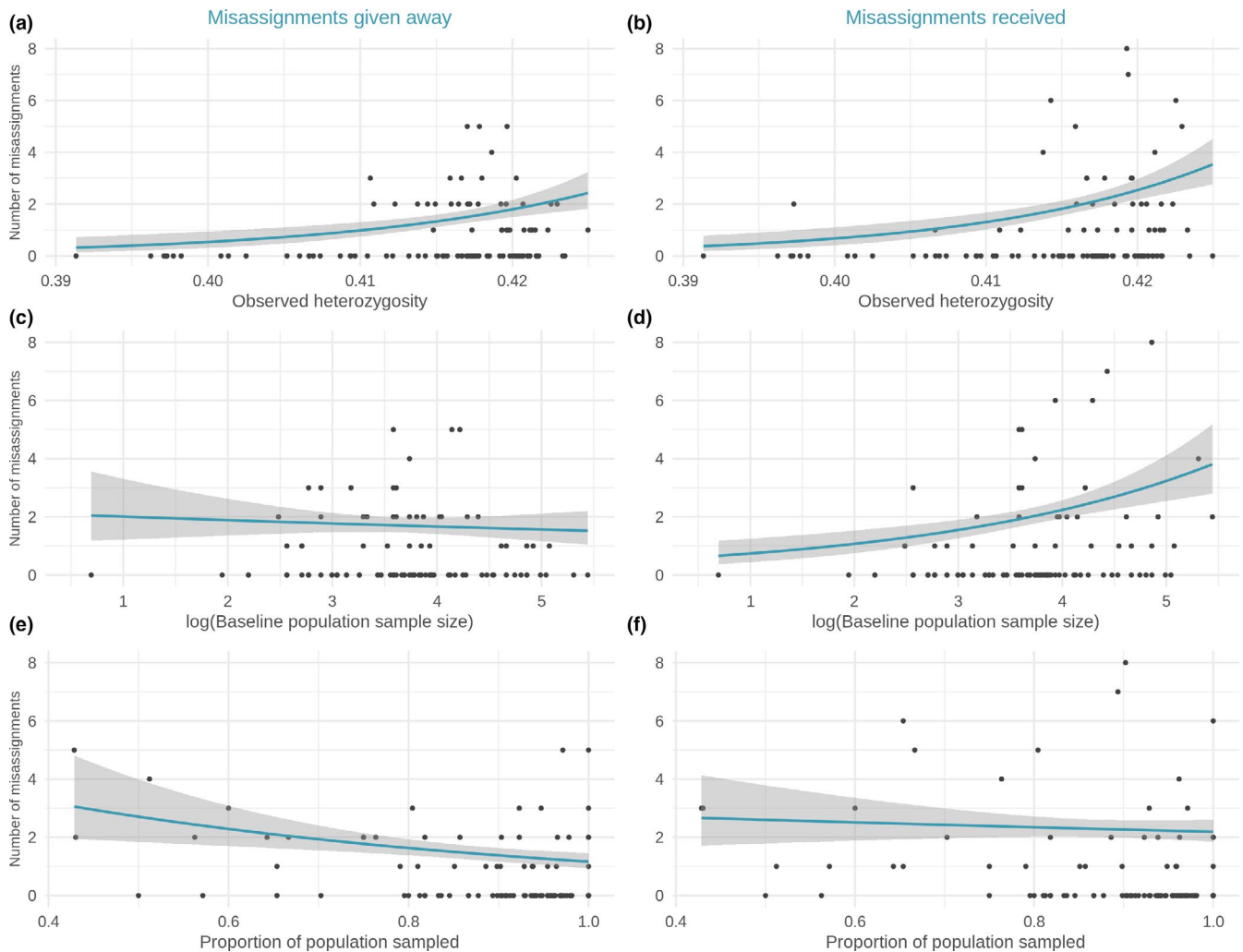
FIGURE 3 The number of recruits in an insular house sparrow metapopulation, among the ones that were known to have hatched on a focal island based on the ecological CMR data ($n = 1645$), that were either wrongly assigned to another island population (misassignments given away; a, c, e), or wrongly assigned to a focal island population (misassignments received; b, d, f). Relationships are shown for the number of misassignments per population and year against observed heterozygosity (a, b), log_e of the baseline population sample size (c, d), and the proportion of the population included in the baseline population (e, f). Blue lines show the predicted relationships, and grey areas show the 95% confidence intervals from ZIP GLMMs

misassignments received, Table S5; see Tables S7, S8 and Figure S11 for the impacts of observed heterozygosity, log_e of the annual sample sizes of the baseline populations, and the proportion of the population that was sampled on proportion of misassignments given away and received).

3.3 | The extended dispersal data set

The number of dispersers identified by ecological CMR monitoring was 291. This constituted 17.7% of the subset of recruits with eco-logically known natal island ($n = 1645$) and 10.6% of the total num-ber of recruits ($n = 2741$) included in the study. Genetic assignment of recruits' natal island using the BONE method suggested however that 658 of the recruits (24.0%) had hatched on a different island in the study metapopulation than the one on which they recruited

(Table S2). In order to reduce the number of potential errors in the final extended dispersal data set, we first checked the genetic as-signment results against the ecological CMR data for individuals where this was available ($n = 1645$; see detailed information in the section on accuracy of genetic assignment methods above). This en-abled us to correct some assignment errors, and reduced the number of dispersers to 621 (22.7% of the recruits). Second, we used the SNP-pedigree either to confirm or potentially correct the recruits' natal islands. This was done for the individuals without ecological CMR data with at least one parent in the SNP-pedigree ($n = 743$). For these recruits, the majority of genetically assigned natal islands was consistent with the island on which their parent(s) was breeding (97.8%; $n = 727$); for the other 16 recruits the natal island was cor-rected to the island where their parent(s) was breeding. Note that seven out of these 16 recruits had parents that were both dispersers themselves, and of these (7) recruits three were actually genetically

assigned to the natal island of their parent(s). Finally, for the remaining recruits (n = 353), which had no available ecological CMR or pedigree information, we only used the results from the genetic assignment analyses to determine their natal island. In total, this resulted in an extended dispersal data set with 607 dispersers (22.2% of recruits), which is about twice the number identified with ecological CMR information (Table S2). The proportions of recruits with different levels of natal island reliability are shown in Figure S4. Finally, in the extended dispersal data set, the annual proportion of dispersing recruits that were only identified using genetic assignment with pedigree correction was on average 54.1% and ranged from 19.1% (recruits born in 2012) to 91.7% (recruits born in 1998; Figure S5).

3.4 | Dispersal patterns in the house sparrow metapopulation

There was considerable variation between years in both the number (Figure S5) and the proportion of dispersers among recruits (Figure S6). There was strong evidence that the proportion of recruits that dispersed varied among years ($\chi^2_{(14)} = 34.666$; $p < .01$). Dispersal was most frequent in 2004 and 2006, when nearly 30% of the recruits dispersed, whereas the lowest dispersal rates were observed in 2000, 2002 and 2008, when 15%–16% of recruits dispersed (Figure S6). There was also very strong evidence for considerable variation among islands in the number and proportions of recruits that emigrated (test for differences in proportions among islands: $\chi^2_{(9)} = 510.25$; $p < .001$), and in the proportion of individuals that recruited to a population that were immigrants from each of the other island populations in the metapopulation (Figure 4). The proportion of recruits born on an island that dispersed to another island in the study system ranged from less than 6% (Aldra) to ~40% (Nesøy and Selvær), and on the two populations with the largest average population sizes across the study period (Gjerøy and Hestmannøy) about 11% of the recruits dispersed to another island. Furthermore, the proportion of recruits on an island that were immigrants ranged from c. 13% (Aldra) to 42% (Træna). For the years 2004–2012, emigration was slightly lower than immigration on six of the study islands, with net dispersal rates (difference between number of immigrants and number of emigrants, divided by the number of birds that recruited on an island) ranging from 2% to 8%. However, on two islands the net dispersal rate was highly negative; −15% on Selvær and −39% on Nesøy.

The majority of dispersal events occurred within each of the two habitat types (farm to farm and nonfarm to nonfarm; Figure 5). There was strong evidence that a higher proportion of recruits dispersed between nonfarm islands (24.3%; 131 out of 540 recruits) than between farm islands (9.6%; 134 out of 1394 recruits; 2004–2012 data: $\beta = 1.199 \pm 0.137$; $p < .001$). Furthermore, the proportion of recruits dispersing from nonfarm islands to farm islands was higher (7.2%; 39 out of 540 recruits) than from farm islands to nonfarm islands (2.1%; 29 out of 1394 recruits; 2004–2012 data: $\beta = 1.357 \pm 0.251$; $p < .001$).

Finally, we found strong evidence that dispersal was female biased across all years and islands in the study metapopulation ($\chi^2_{(1)} = 14.97$; $p < .001$; Figure S7). Overall, 58.6% of the dispersing recruits were females whereas only 49.1% of the resident recruits were females. The female bias in dispersal was slightly higher in this data set, which included genetic assignment information, than the data set that included only ecological CMR data (55% female dispersing recruits; G = 0.114, df = 1, $p > .05$; Figure S7).

4 | DISCUSSION

We have shown that genetic assignment can be a successful approach to identify the natal population and dispersal status of individuals in geographically structured populations even with considerable amounts of gene flow. Our results corroborate previous studies showing that genetic assignment is a powerful tool to explore key ecological and genetic processes in fragmented populations in plants, insects and mammals (Roffler et al., 2014; Sinclair et al., 2018; Vanden Broeck et al., 2017). Using genetic assignment tools can be crucial to accurately track animal and plant movements as well as examine how dispersal affects adaptive and nonadaptive divergence of populations. This tool will thus improve our ability to develop optimal conservation and management strategies for spatially structured populations (Shafer et al., 2016).

4.1 | Accuracy of genetic assignment

In the present study, we applied two different genetic assignment approaches on a unique and extensive long-term individual-based ecological data set to study the accuracy of these methods in a wild vertebrate metapopulation. Using high-quality ecological CMR data on 1645 individuals, we showed that the error rate of BONE was lower than for GSI_SIM (4.8% and 9.5%, respectively; Figure S3). Our results agree with a recent study using both simulations and empirical data on Chinook salmon and house sparrows (data from year 2012, also included in the current study), where the assignment accuracy of the BONE approach was found to be more robust than three other genetic assignment methods (including GSI_SIM; Kuismin et al., 2020). Therefore, the results of the BONE assignment method were chosen to complement our ecological CMR data when identifying the natal island and dispersal status of the recruiting individuals without known origin. The lower error rate of the BONE genetic assignment approach could have two main causes. First, the adapted network estimation method allows the user to use more flexible model selection properties compared to more traditional genetic assignment methods that are based on the Bayesian framework. Second, analyses of different empirical and simulated data sets indicate that the genetic assignment algorithm of BONE is rather robust to uneven sample sizes, which is often the case when sampling wild populations (Kuismin et al., 2017; Lawson et al., 2018).
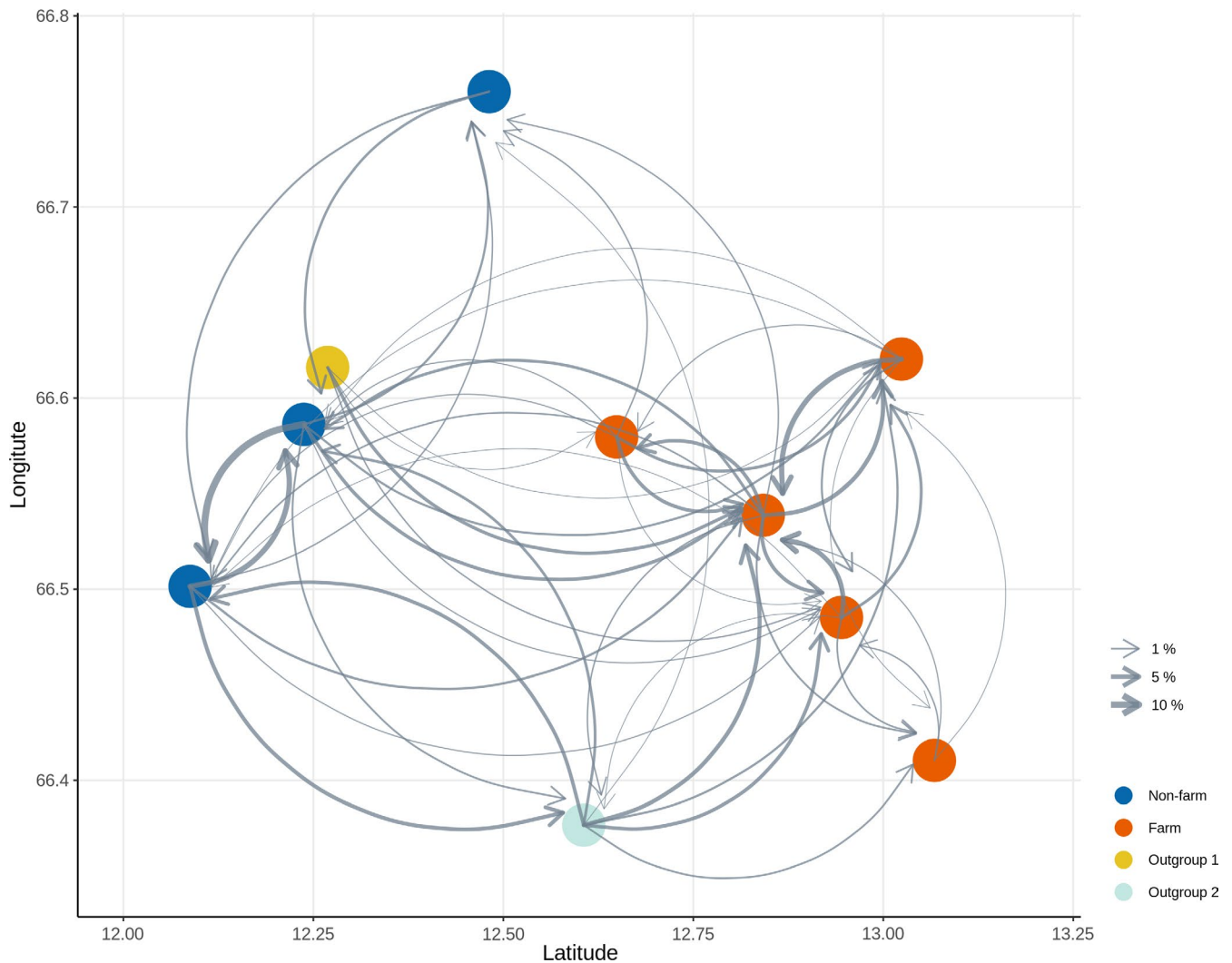
FIGURE 4 Spatial variation in dispersal in an insular house sparrow metapopulation in northern Norway. The direction of arrows show the direction of dispersal from natal island to adult island. Different habitat types and outgroups are shown in different colours. The thicknesses of arrows indicate the proportion of recruits born on an island that dispersed to the other island connected by the arrow across 15 years of study (cohorts 1998–2012 for the five farm islands, and cohorts 2003–2012 for the three nonfarm islands; arrow thicknesses shown as examples in the legends correspond to 1%, 5% and 10% of the recruits, respectively). See Figure 1 for the study islands' true geographical structure

We investigated plausible reasons for wrong genetic assignments obtained using the BONE genetic assignment method. First, a strong negative relationship was observed between pairwise F_{ST} values and the number of misassignments shared by each pair of islands (i.e., baseline populations; Figure 2). Using the BONE genetic assignment method, very few population pairs had any wrong assignments: 94.3% of populations with a pairwise F_{ST} value higher than 0.025, and 77.5% of populations with a pairwise F_{ST} value between 0.010–0.025 had no wrong assignments. However, this percentage dropped to 25% for populations with a pairwise F_{ST} value below 0.01. We also noted that a large proportion of wrong assignments (27%) were observed between the islands Træna and Selvær, that are known to be highly connected through natal dispersal (Pärn et al., 2012) and have the lowest mean pairwise F_{ST} among population pairs in the study system (0.006; Figure S8). This means that the genetic assignment

method identifies an individual's source population more reliably when populations are genetically different from each other. Still, the BONE method appears remarkably reliable even when the level of differentiation (F_{ST}) is as low as 0.01 in our study system. Similar positive effects of genetic differentiation on assignment reliability are expected from theory and simulations also in other geographically structured populations (Anderson, 2010; Araujo et al., 2014; Latch et al., 2006; Paetkau et al., 2004; Waples & Gaggiotti, 2006), and have been documented in other wild populations, such as a grand skink (*Oligosoma grande*) metapopulation (Berry et al., 2004), arctic char (*Salvelinus alpinus*) populations (Moore et al., 2017), and in seven other species of vertebrates and three species of insects (Manel et al., 2002). In general, results from our and other studies suggest that genetic assignment is reliable when the genetic differentiation (F_{ST}) between individuals' potential source populations is

FIGURE 5 The annual percentages of dispersers between and within habitat types among all birds that recruited on the eight islands in the study metapopulation. Abbreviations used for the habitat types are: "F" for "Farm", "NF" for "nonfarm", "OG1" for "Outgroup 1" and "OG2" for "Outgroup 2". See Table S9 for details and actual numbers
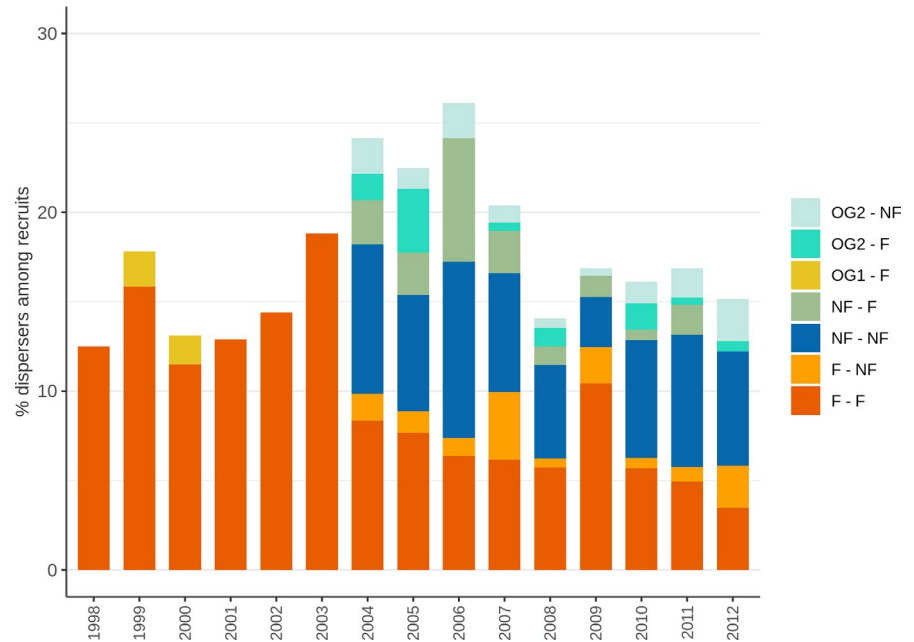
in the order of 0.02–0.07, with the lower limit of differentiation depending on, for example, assignment method and number and type of genetic markers.

Second, the number of misassignments given away and received for a population were positively related to the average observed genome-wide level of heterozygosity of the population (H_o; Figure 3a, b; Table S3; Table S4). However, after controlling for sample size and proportion of the baseline population sampled, we observed a positive relationship only between H_o and misassignments given away (Table S5). In contrast to what we found, it has been previously shown that genetic diversity may positively affect the accuracy of genetic assignments especially if the genetic differentiation between the populations is moderate ($F_{ST} > 0.08$; Berry et al., 2004; Manel et al., 2002). In our metapopulation, annual pairwise F_{ST} estimates were lower than 0.08 for most of the population pairs (Figure 2). Thus, the observed increase in misassignments with increasing H_o may be a consequence of relatively low genetic differentiation in our study metapopulation. Accordingly, populations with high H_o actually appear to be more closely genetically connected and genetically similar to other populations in the system (correlation between annual pairwise F_{ST} and average H_o of population pairs: $r = -0.68, p < .001$). As a consequence, a high population-specific H_o in our system indicates relatively high level of gene flow and low F_{ST} with at least one population. This results in that for some individuals there is another baseline population that provides a sufficiently good but incorrect genetic network match in BONE; the misassignments occur most likely with populations in which relatives of an individual are present (Kuismin et al., 2020).

Third, we found a positive relationship between the number of wrong assignments received and the sample size of the baseline population (Figure 3d; Table S4); this relationship was robust and remained the same also after controlling for other covariates (Table S5). In our study metapopulation, there is a strong positive

correlation between the sample size and the adult population size (i.e., the annual estimated number of adults within each island population; $n = 102$ island years; $r = 0.98$). A positive relationship between genetic diversity and population size is expected from theory (Soulé, 1976) and has been reported before (Frankham, 1996). One possibility could therefore be that the positive relationship between sample (~population) size and wrong assignments received simply reflects the increase in misassignments with higher genetic diversity discussed above. On the other hand, this may not be the only explanation in our case as the relationship with baseline sample sizes remained even when genetic diversity (H_o) was accounted for (Table S5). One plausible complementary reason could be that subpopulations with large population size harbour individuals with more diverse genetic backgrounds that consequently may result in sufficient genetic resemblance between an individual and a large population even if this population is not the individual's true natal population. Another reason might be that this is due to the negative correlation between sample size and level of genetic differentiation in the system (correlation between annual pairwise F_{ST} and average sample size of population pairs: $r = -0.44, p < .001$).

Fourth, a negative relationship was found between the number of misassignments given away and the proportion of a population included in the baseline population (Figure 3e; Table S3). Moreover, when the effects of H_o and sample size were controlled for, a strong negative relationship was observed between the proportion of the population that is sampled and both misassignments given away and received (Table S5). This clearly indicates that more realistic representations of the genetic variation and relatedness structure of a population will cause fewer assignment errors in network based genetic assignment analysis such as BONE, and hence increase the accuracy of the genetic assignments in such situations. However, one should note that this may not be the case for the assignment methods utilizing a Bayesian framework: it has in fact been shown

that inclusion of close relatives in baseline populations can decrease the accuracy of genetic assignments using methods such as ONCOR (Östergren et al., 2019).

4.2 | Dispersal in the house sparrow metapopulation

After merging the ecological CMR data with the genetic assignments that had been checked and corrected with information from a metapopulation-level SNP-genotype-based pedigree, the number of dispersers in our data set increased from 291 recruits (17.7% of the ecological CMR data only) to 607 recruits (22.2% of all recruits included in the study; Table S2). This is a considerable increase, and mostly due to the fact that a high proportion of the dispersers did not have natal island information from ecological CMR data. Natal island may be unknown because an individual was captured for the first time in autumn and could have dispersed before/during this period of the year, or that it was first captured and ringed after the natal dispersal season as a one year old adult (i.e., a new recruit; Figure S4). Even though the assignment accuracy was very high (95.2% of the assignments were correct based on the ecological CMR data), some errors in the assignment of individuals as dispersers or residents are still expected to be present in our extended dispersal data set (Figure S3), and the final number of dispersers may therefore be slightly overestimated. Most of these incorrectly defined dispersers are expected to be from population pairs with low levels of genetic differentiation (see Figure 2), such as the islands of Træna and Selvær (Figure S8). Although such populations may be considered as a single population in a population genetic context, it can be important to treat them as separate populations, with a high level of interpopulation dispersal, in an ecological context, at least when dispersal is not occurring throughout the whole year. This is because most ecological and social processes, such as any competition over mates and/or breeding sites or other resources, will happen within the local populations. Nevertheless, even if one assumes that 3.6% of the recruits were wrongly assigned as dispersers, and 1.2% of the recruits were wrongly assigned as residents (see Figure S3), the overall proportion of recruits that dispersed would still be c. 19.8%. Furthermore, note that the offspring of dispersers in our house sparrow metapopulation tend to have a higher probability of dispersing than the offspring of resident birds (Saatoglu et al., 2021). When this is the case, some individuals may have been assigned to a wrong island simply because their parents were immigrants from that island.

We found variation in the proportion of dispersers in both time and space in our house sparrow metapopulation (Figures 4, 5). The percentage of dispersers among cohorts of recruits ranged between 15% and 30% across 16 study years (Figure S6). Interestingly, we found that six out of our eight island populations had approximately equal emigration and immigration rates, whereas the emigration rate was considerably higher than the immigration rate on two islands (Nesøy and Selvær). Furthermore, dispersal rates were higher among islands with nonfarm habitat than among islands with farm habitat,

and higher from nonfarm habitat islands to farm-habitat islands than vice versa (Figure 5; Tables S9, S10). These results may help explain interisland differences in inbreeding levels (Niskanen et al., 2020) and the differences in mean morphology that appear to exist between birds in the two habitat types (Araya-Ajoy et al., 2019; Muff et al., 2019). Furthermore, the results may suggest the existence of habitat differences in productivity and dispersal probability, and hence that there could perhaps be source and sink populations in our study metapopulation. However, to properly identify populations as sources and sinks, other demographic parameters like reproduction and mortality rates also need to be quantified (Dias, 1996; Furrer & Pasinelli, 2016). In any case, the considerable spatiotemporal variation in dispersal rates we observed suggests that population properties, such as population size/density (Pärn et al., 2012) and relatedness (through inbreeding avoidance; Niskanen et al., 2020), environmental conditions, such as weather (Pärn et al., 2012), and/ or individual phenotypic characteristics, such as wing length (Araya-Ajoy et al., 2019), may be important determinants of dispersal in our study system (Benton & Bowler, 2012; Cote et al., 2017). Examining the causes of variation in dispersal rates and individual dispersal decisions will help us to understand both the environmental (e.g., population size and/or density, food abundance, weather/climate conditions, social structure) and intrinsic drivers (e.g., genetic basis for dispersal, and inbreeding avoidance) of dispersal. Our new dispersal data set allows such investigations of both causes and consequences of dispersal but this is outside the scope of the current study and will be the focus of future studies.

We found that a higher proportion of dispersers were females (0.59). This met with our expectation since it has been shown that female-biased dispersal is common in natural bird populations (species level, 0.70; family level 0.65; Pusey, 1987).

The increased number of dispersers identified with high accuracy using a genetic assignment method, shows that even with extensive ecological CMR monitoring, the natal population (and dispersal status) of a relatively large number of individuals may remain unidentified (e.g., 40% of recruits in our house sparrow metapopulation). The importance of duration and timing in CMR studies has been widely emphasized, and it has been shown that for instance in bird metapopulations, capturing the nestlings in the nest or fledged juveniles with mist-nets during the breeding season is an efficient way to maximize the accuracy of the dispersal status information gathered (Chambert et al., 2012; Dupont et al., 2019). However, in long-term studies, it is challenging to carry out CMR monitoring during the whole year and in a sufficiently large study area to document dispersal events due to cost and effort limitations. Accordingly, a review study by Driscoll et al., (2014) that examined how dispersal was treated in conservation biology, found that half of the ~600 studies had a dispersal knowledge gap (including basic information on individual status as disperser or resident) due to deficiencies in the information gathering stage. Moreover, it has been shown that the consequences of inadequate dispersal knowledge may be incomplete or biased conclusions of the results from empirical studies, including effects on eco-evolutionary dynamics (Driscoll et al., 2014).

Genetic approaches enable us to estimate dispersal rates directly over the sampled individuals when the sampling is adequate and timely, whereas dispersal rates over CMR data are usually estimated indirectly (i.e., with CMR models; Broquet & Petit, 2009). CMR models have been widely criticized in the literature because many are likely to give biased estimates due to inadequate dispersal data, limiting or unrealistic model assumptions, and unsuitable sampling times (Cayuela et al., 2018; Dupont et al., 2019; Lebreton et al., 2009). Moreover, many studies have a geographical scale that is suboptimal for obtaining unbiased estimates of dispersal, regardless of methods used to identify dispersers. However, the scale of our study system is large relative to normal dispersal distances of the house sparrow (Tufto et al., 2005), which allowed accurate identification of nearly all dispersers using the BONE genetic assignment methods. Therefore, as shown in this study, it may be very useful to complement even extensive ecological CMR monitoring with genetic assignment approaches to reduce any biases in estimates of dispersal.

5 | CONCLUSIONS

To conclude, determining the true identity and number of dispersers with a high accuracy is crucially important to increase the statistical power of both empirical studies as well as trainee data sets for theoretical studies, and thus help improve our understanding of ecological and evolutionary dynamics. We have shown that the recently designed genetic assignment software BONE can identify the source population of individuals with a high accuracy. Nonetheless, one cannot completely eliminate genetic assignment errors, unless the study system consists of (at least) moderately differentiated populations with relatively few dispersal events. This is simply because the offspring of the dispersers will, with some probability, be assigned to the natal population(s) of their immigrant parent(s). Using a high-quality metapopulation-level pedigree is another way to test and increase the assignment accuracy (see also Berry et al., 2004). However, none of these findings would have been testable or open to discussion with ecological or genetic data alone. We believe that it will be increasingly important to integrate ecological and genetic approaches to improve our understanding, not only of dispersal dynamics of wild populations, but also of ecological and evolutionary dynamics (Cayuela et al., 2018; Moore et al., 2017; Shafer et al., 2016).

DATA AVAILABILITY STATEMENT
SNP genotype data for the house sparrow study metapopulation have been made available in the Dryad repository (https://doi.org/10.5061/dryad.hp758snand https://doi.org/10.5061/dryad.m0cfxpp10. Relevant R scripts and analysis files are deposited in Dryad (https://doi.org/10.5061/dryad.xgxd254h1.

ORCID
Dilan Saatoglu https://orcid.org/0000-0003-3936-287X
Alina K. Niskanen https://orcid.org/0000-0003-2017-2718
Markku Kuismin https://orcid.org/0000-0001-9074-7420
Peter S. Ranke https://orcid.org/0000-0003-3757-8626
Ingerid J. Hagen https://orcid.org/0000-0003-1028-3940
Thor Harald Ringsby https://orcid.org/0000-0001-9089-7592
Bernt-Erik Sæther https://orcid.org/0000-0002-0049-9767
Arild Husby https://orcid.org/0000-0003-1911-8351
Mikko J. Sillanpää https://orcid.org/0000-0003-2808-2768
Henrik Jensen https://orcid.org/0000-0001-7804-1564

REFERENCES
Akçakaya, H. R. (2000). Population viability analyses with demographically and spatially structured models. *Ecological Bulletins*, *48*, 23–38.
Alleaume-Benharira, M., Pen, I. R., & Ronce, O. (2006). Geographical patterns of adaptation within a species' range: Interactions between drift and gene flow. *Journal of Evolutionary Biology*, *19*(1), 203–215. https://doi.org/10.1111/j.1420-9101.2005.00976.x
Altwegg, R., Ringsby, T. H., & Sæther, B.-E. (2000). Phenotypic correlates and consequences of dispersal in a metapopulation of house sparrows *Passer domesticus*. *Journal of Animal Ecology*, *69*(5), 762–770. https://doi.org/10.1046/j.1365-2656.2000.00431.x
Anderson, E. C. (2010). Assessing the power of informative subsets of loci for population assignment: Standard methods are upwardly biased. *Molecular Ecology Resources*, *10*(4), 701–710. https://doi.org/10.1111/j.1755-0998.2010.02846.x

Anderson, E. C., Waples, R. S., & Kalinowski, S. T. (2008). An improved method for predicting the accuracy of genetic stock identification. *Canadian Journal of Fisheries and Aquatic Sciences*, 65(7), 1475–1486. https://doi.org/10.1139/F08-049

Araujo, H. A., Candy, J. R., Beacham, T. D., White, B., & Wallace, C. (2014). Advantages and challenges of genetic stock identification in fish stocks with low genetic resolution. *Transactions of the American Fisheries Society*, 143(2), 479–488. https://doi.org/10.1080/00028 487.2013.855258

Araya-Ajoy, Y. G., Ranke, P. S., Kvalnes, T., Rønning, B., Holand, H., Myhre, A. M., Pärn, H., Jensen, H., Ringsby, T. H., Sæther, B.-E., & Wright, J. (2019). Characterizing morphological (co)variation using structural equation models: Body size, allometric relationships and evolvability in a house sparrow metapopulation. *Evolution*, 73(3), 452–466. https://doi.org/10.1111/evo.13668

Armansin, N. C., Stow, A. J., Cantor, M., Leu, S. T., Klarevas-Irby, J. A., Chariton, A. A., & Farine, D. R. (2020). Social barriers in ecological landscapes: The social resistance hypothesis. *Trends in Ecology and Evolution*, 35(2), 137–148. https://doi.org/10.1016/j.tree.2019.10.001

Baalsrud, H. T., Sæther, B.-E., Hagen, I. J., Myhre, A. M., Ringsby, T.-H., Pärn, H., & Jensen, H. (2014). Effects of population characteristics and structure on estimates of effective population size in a house sparrow metapopulation. *Molecular Ecology*, 23, 2653–2668. https://doi.org/10.1111/mec.12770

Benton, T. G., & Bowler, D. E. (2012). Dispersal in invertebrates: influences on individual decisions. In M. Baguette T. G. Benton & J. M. Bullock (Eds.), *Dispersal ecology and evolution*. (pp. 41–49). Oxford University Press.

Berdahl, A., Torney, C. J., Schertzer, E., & Levin, S. A. (2015). On the evolutionary interplay between dispersal and local adaptation in heterogeneous environments. *Evolution*, 69(6), 1390–1405. https://doi.org/10.1111/evo.12664

Berg, M. P., Kiers, E. T., Driessen, G., van der HEIJDEN, M., Kooi, B. W., Kuenen, F., Liefting, M., Verhoef, H. A., & Ellers, J. (2010). Adapt or disperse: Understanding species persistence in a changing world. *Global Change Biology*, 16(2), 587–598. https://doi.org/10.1111/j.1365-2486.2009.02014.x

Berry, O., Tocher, M. D., & Sarre, S. D. (2004). Can assignment tests measure dispersal? *Molecular Ecology*, 13(3), 551–561. https://doi.org/10.1046/j.1365-294X.2004.02081.x

Bonte, D., & Dahirel, M. (2017). Dispersal: a central and independent trait in life history. *Oikos*, 126(4), 472–479. https://doi.org/10.1111/oik.03801

Bowler, D. E., & Benton, T. G. (2005). Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, 80, 205–225. https://doi.org/10.1017/S1464 793104006645

Brooks, M. E., Kristensen, K., Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2), 378–400. https://doi.org/10.32614/RJ-2017-066

Broquet, T., & Petit, E. J. (2009). Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 193–216. https://doi.org/10.1146/annurev.ecolsys.110308.120324

Cagua, E. F., Cochran, J. E. M., Rohner, C. A., Prebble, C. E. M., Sinclair-Taylor, T. H., Pierce, S. J., & Berumen, M. L. (2015). Acoustic telemetry reveals cryptic residency of whale sharks. *Biology Letters*, 11(4), 20150092. https://doi.org/10.1098/rsbl.2015.0092

Cayuela, H., Rougemont, Q., Prunier, J. G., Moore, J. S., Clobert, J., Besnard, A., & Bernatchez, L. (2018). Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Molecular Ecology*, 27(20), 3976–4010. https://doi.org/10.1111/mec.14848

Chambert, T., Pardo, D., Choquet, R., Staszewski, V., McCoy, K. D., Tveraa, T., & Boulinier, T. (2012). Heterogeneity in detection probability along the breeding season in Black-legged Kittiwakes: Implications for sampling design. *Journal of Ornithology*, 152(SUPPL. 2), 371–380. https://doi.org/10.1007/s10336-010-0542-8

Chen, K. Y., Marschall, E. A., Sovic, M. G., Fries, A. C., Gibbs, H. L., & Ludsin, S. A. (2018). assignPOP: An r package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. *Methods in Ecology and Evolution*, 9(2), 439–446. https://doi.org/10.1111/2041-210X.12897

Clarke, A. L., Sæther, B.-E., Røskaft, E., Saether, B.-E., & Roskaft, E. (1997). Sex biases in avian dispersal: A reappraisal. *Oikos*, 79(3), 429–438. https://doi.org/10.2307/3546885

Corander, J., Waldmann, P., & Sillanpää, M. J. (2003). Bayesian analysis of genetic differentiation between populations. *Genetics*, 163(1), 367–374. https://doi.org/10.1093/bioinformatics/bth250

Cote, J., Bestion, E., Jacob, S., Travis, J., Legrand, D., & Baguette, M. (2017). Evolution of dispersal strategies and dispersal syndromes in fragmented landscapes. *Ecography*, 40(1), 56–73. https://doi.org/10.1111/ecog.02538

Dias, P. C. (1996). Sources and sinks in population biology. *Trends in Ecology & Evolution*, 11(8), 326–330. https://doi.org/10.1016/0169-5347(96)10037-9

Driscoll, D. A., Banks, S. C., Barton, P. S., Ikin, K., Lentini, P., Lindenmayer, D. B., Smith, A. L., Berry, L. E., Burns, E. L., Edworthy, A., Evans, M. J., Gibson, R., Heinsohn, R., Howland, B., Kay, G., Munro, N., Scheele, B. C., Stirnemann, I., Stojanovic, D., ... Westgate, M. J. (2014). The trajectory of dispersal research in conservation biology. Systematic review. *PLoS One*, 9(4), e95053. https://doi.org/10.1371/journal.pone.0095053

Dupont, P., Milleret, C., Gimenez, O., & Bischof, R. (2019). Population closure and the bias-precision trade-off in spatial capture–recapture. *Methods in Ecology and Evolution*, 10(5), 661–672. https://doi.org/10.1111/2041-210X.13158

Elgvin, T. O., Trier, C. N., Tørresen, O. K., Hagen, I. J., Lien, S., Nederbragt, A. J., Ravinet, M., Jensen, H., & Sætre, G.-P. (2017). The genomic mosaicism of hybrid speciation. *Science Advances*, 3(6), e1602996. https://doi.org/10.1126/sciadv.1602996

François, O., & Waits, L. P. (2015). Clustering and assignment methods in landscape genetics. *Landscape Genetics*, 114–128. https://doi.org/10.1002/9781118525258.ch07

Frankham, R. (1996). Relationship of genetic variation to population size in wildlife. *Conservation Biology*, 10(6), 1500–1508. https://doi.org/10.1046/j.1523-1739.1996.10061500.x

Furrer, R. D., & Pasinelli, G. (2016). Empirical evidence for source–sink populations: A review on occurrence, assessments and implications. *Biological Reviews*, 91, 782–795. https://doi.org/10.1111/brv.12195

Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F statistics. *Molecular Ecology Notes*, 5(1), 184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x

Greenwood, P. J. (1980). Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, 28(4), 1140–1162. https://doi.org/10.1016/S0003-3472(80)80103-5

Hanaoka, S., Chien, C. T., Chen, S. Y., Watanabe, A., Setsuko, S., & Kato, K. (2014). Genetic structures of *Calophyllum inophyllum* L., a tree employing sea-drift seed dispersal in the northern extreme of its distribution. *Annals of Forest Science*, 71(5), 575–584. https://doi.org/10.1007/s13595-014-0365-5

Hanski, I. A., & O. E. Gaggiotti (Eds.) (2004). *Ecology, genetics and evolution of metapopulations*. Academic Press.

Hayden, T. A., Holbrook, C. M., Fielder, D. G., Vandergoot, C. S., Bergstedt, R. A., Dettmers, J. M., Krueger, C. C., & Cooke, S. J. (2014). Acoustic telemetry reveals large-scale migration patterns of walleye in Lake Huron. *PLoS One*, 9(12), 1–19. https://doi.org/10.1371/journal.pone.0114833

Holand, H., Kvalnes, T., Gamelon, M., Tufto, J., Jensen, H., Pärn, H., Ringsby, T. H., & Sæther, B.-E. (2016). Spatial variation in senescence rates in a bird metapopulation. *Oecologia*, 181(3), 865–871. https://doi.org/10.1007/s00442-016-3615-4

Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: Defining, estimating and interpreting F_{st}. *Nature Reviews Genetics*, 10(9), 639–650. https://doi.org/10.1038/nrg2611

Holyoak, M., Casagrandi, R., Nathan, R., Revilla, E., & Spiegel, O. (2008). Trends and missing parts in the study of movement ecology. *Proceedings of the National Academy of Sciences*, 105(49), 19060–19065.

Iwajomo, S. B., Willemoes, M., Ottosson, U., Strandberg, R., & Thorup, K. (2018). Intra-African movements of the African cuckoo *Cuculus gularis* as revealed by satellite telemetry. *Journal of Avian Biology*, 49(1), 1–5. https://doi.org/10.1111/jav.01616

Jensen, H., Sæther, B.-E., Ringsby, T. H., Tufto, J., Griffith, S. C., & Ellegren, H. (2004). Lifetime reproductive success in relation to morphology in the house sparrow *Passer domesticus*. *Journal of Animal Ecology*, 73(4), 599–611. https://doi.org/10.1111/j.0021-8790.2004.00837.x

Jensen, H., Steinsland, I., Ringsby, T. H., & Sæther, B.-E. (2008). Evolutionary dynamics of a sexual ornament in the house sparrow (*Passer domesticus*): The role of indirect selection within and between sexes. *Evolution*, 62(6), 1275–1293. https://doi.org/10.1111/j.1558-5646.2008.00395.x

Kalinowski, S. T. (2004). Counting Alleles with Rarefaction: Private Alleles and Hierarchical Sampling Designs. *Conservation Genetics*, 5, 539–543. https://doi.org/10.1023/B:COGE.0000041021.91777.1a

Kuismin, M., Ahlinder, J., & Sillanpää, M. J. (2017). CONE: Community oriented network estimation is a versatile framework for inferring population structure in large-scale sequencing data. *G3 Genes|Genomes|Genetics*, 7(10), 3359–3377. https://doi.org/10.1534/g3.117.300131

Kuismin, M., Saatoglu, D., Niskanen, A. K., Jensen, H., & Sillanpää, M. J. (2020). Genetic assignment of individuals to source populations using network estimation tools. *Methods in Ecology and Evolution*, 11(2), 333–344. https://doi.org/10.1111/2041-210X.13323

Kvalnes, T., Røberg, A. Å., Jensen, H., Holand, H., Pärn, H., Sæther, B.-E., & Ringsby, T. H. (2018). Offspring fitness and the optimal propagule size in a fluctuating environment. *Journal of Avian Biology*, 49(7), 1–6. https://doi.org/10.1111/jav.01786

Latch, E. K., Dharmarajan, G., Glaubitz, J. C., & Rhodes, O. E. (2006). Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics*, 7(2), 295–302. https://doi.org/10.1007/s10592-005-9098-1

Lawson, D. J., van Dorp, L., & Falush, D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications*, 9(1), 1–11. https://doi.org/10.1038/s41467-018-05257-7

Lebreton, J., Nichols, J. D., Barker, R. J., Pradel, R., & Spendelow, J. A. (2009). Modeling individual animal histories with multistate capture–recapture models. *Advances in Ecological Research*, 41, 87–173.

Lowe, W. H., Kovach, R. P., & Allendorf, F. W. (2017). Population genetics and demography unite ecology and evolution. *Trends in Ecology and Evolution*, 32(2), 141–152. https://doi.org/10.1016/j.tree.2016.12.002

Lundregan, S. L., Hagen, I. J., Gohli, J., Niskanen, A. K., Kemppainen, P., Ringsby, T. H., Kvalnes, T., Pärn, H., Rønning, B., Holand, H., Ranke, P. S., Båtnes, A. S., Selvik, L.-K., Lien, S., Saether, B.-E., Husby, A., & Jensen, H. (2018). Inferences of genetic architecture of bill morphology in house sparrow using a high density SNP array point to a polygenic basis. *Molecular Ecology*, 27, 3498–3514. https://doi.org/10.1111/mec.14811

Manel, S., Berthier, P., & Luikart, G. (2002). Detecting wildlife poaching: Identifying the origin of multilocus genotypes. *Conservation Biology*, 16(3), 650–659.

Manel, S., Gaggiotti, O. E., & Waples, R. S. (2005). Assignment methods: Matching biological questions with appropriate techniques. *Trends in Ecology and Evolution*, 20, 136–142. https://doi.org/10.1016/j.tree.2004.12.004

Marchi, C., Andersen, L. W., & Loeschcke, V. (2013). Effects of land management strategies on the dispersal pattern of a beneficial arthropod. *PLoS One*, 8(6), 2–8. https://doi.org/10.1371/journal.pone.0066208

Matthysen, E. (2012). Multicausality of dispersal: A review. In M. Baguette T. G. Benton & J. M. Bullock (Eds.), *Dispersal ecology and Evolution*, Vol. 27 (pp. 3–18). Oxford University Press.

Mauritzen, M., Derocher, A. E., Wiig, Ø., Belikov, S. E., Boltunov, A. N., Hansen, E., & Garner, G. W. (2002). Using satellite telemetry to define spatial population structure in polar bears in the Norwegian and western Russian Arctic. *Journal of Applied Ecology*, 39, 79–90. https://doi.org/10.1046/j.1365-2664.2002.00690.x

Millon, A., Lambin, X., Devillard, S., & Schaub, M. (2019). Quantifying the contribution of immigration to population dynamics: a review of methods, evidence and perspectives in birds and mammals. *Biological Reviews*, 94(6), 2049–2067. https://doi.org/10.1111/brv.12549

Moore, J.-S., Harris, L. N., Le Luyer, J., Sutherland, B. J. G., Rougemont, Q., Tallman, R. F., Fisk, A. T., & Bernatchez, L. (2017). Genomics and telemetry suggest a role for migration harshness in determining overwintering habitat choice, but not gene flow, in anadromous Arctic Char. *Molecular Ecology*, 26(24), 6784–6800. https://doi.org/10.1111/mec.14393

Moran, B. M., & Anderson, E. C. (2019). Bayesian inference from the conditional genetic stock identification model. *Canadian Journal of Fisheries and Aquatic Sciences*, 76, 551–560. https://doi.org/10.1139/cjfas-2018-0016

Muff, S., Niskanen, A. K., Saatoglu, D., Keller, L. F., & Jensen, H. (2019). Animal models with group-specific additive genetic variances: Extending genetic group models. *Genetics Selection Evolution*, 51(1), 1–16. https://doi.org/10.1186/s12711-019-0449-7

Muriel, R., Morandini, V., Ferrer, M., & Balbontín, J. (2015). Independence and juvenile dispersal distances in wild and reintroduced Spanish imperial eagles. *Biological Conservation*, 191, 300–305. https://doi.org/10.1016/j.biocon.2015.07.020

Niskanen, A. K., Billing, A. M., Holand, H., Hagen, I. J., Araya-Ajoy, Y. G., Husby, A., Rønning, B., Myhre, A. M., Ranke, P. S., Kvalnes, T., Pärn, H., Ringsby, T. H., Lien, S., Sæther, B.-E., Muff, S., & Jensen, H. (2020). Consistent scaling of inbreeding depression in space and time in a house sparrow metapopulation. *Proceedings of the National Academy of Sciences*, 117(25), 14584–14592. https://doi.org/10.1073/pnas.1909599117

Orantes, L. C., Zhang, W., Mian, M. A. R., & Michel, A. P. (2012). Maintaining genetic diversity and population panmixia through dispersal and not gene flow in a holocyclic heteroecious aphid species. *Heredity*, 109(2), 127–134. https://doi.org/10.1038/hdy.2012.21

Östergren, J., Palm, S., Gilbey, J., & Dannewitz, J. (2019). Close relatives in population samples: Evaluation of the consequences for genetic stock identification. *Molecular Ecology Resources*, 20(2), 498–510. https://doi.org/10.1111/1755-0998.13131

Paetkau, D., Slade, R., Burden, M., & Estoup, A. (2004). Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, 13, 55–65. https://doi.org/10.1046/j.1365-294X.2003.02008.x

Pärn, H., Jensen, H., Ringsby, T.-H., & Sæther, B.-E. (2009). Sex-specific fitness correlates of dispersal in a house sparrow metapopulation. *Journal of Animal Ecology*, 78(6), 1216–1225. https://doi.org/10.1111/j.1365-2656.2009.01597.x

Pärn, H., Ringsby, T.-H., Jensen, H., & Sæther, B.-E. (2012). Spatial heterogeneity in the effects of climate and density-dependence on dispersal in a house sparrow metapopulation. *Proceedings of the*

Royal Society B: Biological Sciences, *279*(1726), 144–152. https://doi.org/10.1098/rspb.2011.0673

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

Pritchard, J. K., Stephens, M., Rosenberg, N. A., & Donnelly, P. (2000). Association mapping in structured populations. *American Journal of Human Genetics*, *67*(1), 170–181. https://doi.org/10.1086/302959

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Pusey, A. E. (1987). Sex-biased dispersal and inbreeding avoidance in birds and mammals. *Trends in Ecology & Evolution*, *2*(10), 295–299. https://doi.org/10.1016/0169-5347(87)90081-4

Putman, A. I., & Carbone, I. (2014). Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution*, *4*(22), 4399–4428. https://doi.org/10.1002/ece3.1305

Ranke, P. S., Araya-Ajoy, Y. G., Ringsby, T. H., Pärn, H., Rønning, B., Jensen, H. & Sæther, B.-E. (2021). Spatial structure and dispersal dynamics in a house sparrow metapopulation. Manuscript submitted for publication.

Revelle, W. (2018). *R-package 'Psych': Procedures for personality and psychological research*. Northwestern University.

Riley, S. P. D., Pollinger, J. P., Sauvajot, R. M., York, E. C., Bromley, C., Fuller, T. K., & Wayne, R. K. (2006). A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology*, *15*(7), 1733–1741. https://doi.org/10.1111/j.1365-294X.2006.02907.x

Ringsby, T. H., Sæther, B.-E., Tufto, J., Jensen, H., & Solberg, E. J. (2002). Asynchronous spatiotemporal demography of a House Sparrow metapopulation in a correlated environment. *Ecology*, *83*(2), 561–569. https://doi.org/10.1890/0012-9658(2002)083(0561:ASDOAH)2.0.CO;2

Roffler, G. H., Talbot, S. L., Luikart, G., Sage, G. K., Pilgrim, K. L., Adams, L. G., & Schwartz, M. K. (2014). Lack of sex-biased dispersal promotes fine-scale genetic structure in alpine ungulates. *Conservation Genetics*, *15*(4), 837–851. https://doi.org/10.1007/s10592-014-0583-2

Ronce, O. (2007). How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics*, *38*, 231–253. https://doi.org/10.1146/annurev.ecolsys.38.091206.095611

Saatoglu, D., Lundregan, S. L., Fetterplace, E., Niskanen, A. K., Ranke, P. S., Ringsby, T.-H. & Jensen, H. (2021). The genetic basis for dispersal in a house sparrow metapopulation. Manuscript in preparation.

Schwartz, M. K., Mills, L. S., McKelvey, K. S., Ruggiero, L. F., & Allendorf, F. W. (2002). DNA reveals high dispersal synchronizing the population dynamics of Canada lynx. *Nature*, *415*(6871), 520–522.

Shafer, A. B. A., Northrup, J. M., Wikelski, M., Wittemyer, G., & Wolf, J. B. W. (2016). Forecasting ecological genomics: High-tech animal instrumentation meets high-throughput sequencing. *PLoS Biology*, *14*(1), 1–11. https://doi.org/10.1371/journal.pbio.1002350

Sinclair, E. A., Ruiz-Montoya, L., Krauss, S. L., Anthony, J. M., Hovey, R. K., Lowe, R. J., & Kendrick, G. A. (2018). Seeds in motion: Genetic assignment and hydrodynamic models demonstrate concordant patterns of seagrass dispersal. *Molecular Ecology*, *27*(24), 5019–5034. https://doi.org/10.1111/mec.14939

Soulé, M. (1976). Allozyme variation, its determinants in space and time. In F. J. Ayala (Ed.), *Molecular evolution*. Sinauer Associates.

Stubberud, M. W., Myhre, A. M., Holand, H., Kvalnes, T., Ringsby, T. H., Sæther, B.-E., & Jensen, H. (2017). Sensitivity analysis of effective population size to demographic parameters in house sparrow populations. *Molecular Ecology*, *26*(9), 2449–2465. https://doi.org/10.1111/mec.14057

Summers-Smith, J. D. (1988). *The Sparrows: A study of the genus*. Passer. T & AD Poyser.

Travis, J. M. J., Delgado, M., Bocedi, G., Baguette, M., Bartoń, K., Bonte, D., Boulangeat, I., Hodgson, J. A., Kubisch, A., Penteriani, V., Saastamoinen, M., Stevens, V. M., & Bullock, J. M. (2013). Dispersal and species' responses to climate change. *Oikos*, *122*(11), 1532–1540. https://doi.org/10.1111/j.1600-0706.2013.00399.x

Truve, J., & Lemel, J. (2003). Timing and distance of natal dispersal for wild boar. *Wildlife Biology*, *9*, 51–57.

Tufto, J., Ringsby, T., Dhondt, A. A., Adriaensen, F., & Matthysen, E. (2005). A parametric model for estimation of dispersal patterns applied to five passerine spatially structured populations. *The American Naturalist*, *165*(1), E13–E26. https://doi.org/10.1086/426698

Van Dyck, H., & Baguette, M. (2005). Dispersal behaviour in fragmented landscapes: Routine or special movements? *Basic and Applied Ecology*, *6*(6), 535–545. https://doi.org/10.1016/j.baae.2005.03.005

Vance, R. R. (1984). The effect of dispersal on population stability in one-species, discrete-space population growth models. *American Naturalist*, *123*(2), 230–254. https://doi.org/10.1086/284199

Vanden Broeck, A. N., Maes, D., Kelager, A., Wynhoff, I., WallisDeVries, M. F., Nash, D. R., Oostermeijer, J. G. B., Van Dyck, H., & Mergeay, J. (2017). Gene flow and effective population sizes of the butterfly *Maculinea alcon* in a highly fragmented, anthropogenic landscape. *Biological Conservation*, *209*, 89–97. https://doi.org/10.1016/j.biocon.2017.02.001

Waples, R. S., & Gaggiotti, O. (2006). 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

Weir, B., & Cockerham, C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358–1370. https://doi.org/10.2307/2408641

Whiteley, A. R., Fitzpatrick, S. W., Funk, W. C., & Tallmon, D. A. (2015). Genetic rescue to the rescue. *Trends in Ecology and Evolution*, *30*(1), 42–49. https://doi.org/10.1016/j.tree.2014.10.009

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.