Dear Contributor,

Here is a low resolution PDF of your chapter and, if applicable, your affiliation details in a 'List of contributors'. If your chapter includes plates that are part of a separate colour section, these are also attached. You are receiving this as the lead author of this chapter. Please forward to any co-authors if required.

Please check these page proofs and submit **essential corrections only:** these include typographical, factual or typesetting errors. Please note that we cannot accept any new, rewritten or revised content or figures at this stage (unless we have specifically requested this from you). Extensive changes are not possible, due to the scheduled publication date for your book, which is now being pre-promoted.

Please also ensure that a cross-reference in the text is included for all figures (if applicable).

Mark your corrections directly in the PDF using the editing function in Adobe Acrobat. Your changes must appear as track changes or comments. (Note: If you wish to use an alternative method to mark your corrections, please consult with your publisher/editor first. Do not supply a replacement Word file with changes already incorporated (i.e. where track changes are not visible), as the book will not be re-typeset.)

Please refer to the email from your publisher or editor for the deadline to submit corrections and any specific queries we may have on your chapter.

Should you have any questions about the proofs, please do not hesitate to contact your publisher/editor.

Kind regards,

CSIRO Publishing Production Team

CSIRO
Australia's National Science Agency

22   Genomic approaches to study dispersal in wild animal populations: implications for wildlife management

Hugo Cayuela, Jérôme G. Prunier and Quentin Rougemont

ABSTRACT

Dispersal is a central eco-evolutionary process determining the demographic and genetic dynamics of wild populations. Therefore, wildlife managers are often interested in quantifying frequency and distances of dispersal within spatially structured populations, by considering dispersal (i.e. movements followed or not by a successful reproduction event) and gene flow (i.e. movements followed by a successful reproduction event). Yet, dispersal is a difficult process to measure through direct observation. Genomic approaches to studying dispersal capitalise on the rich genealogical information obtainable from DNA sequences and its distribution within the landscape. The continued reduction in costs of DNA sequencing, and the development and diversification of analytical techniques means that DNA-based measurement of dispersal is now mainstream activity applied to such diverse applications as conservation, wild harvest, and pest control. Here, we provide a synthetic snapshot of existing genomic methods to assess dispersal and gene flow in wild animal populations. We identify the strengths and weaknesses of each method, provide recommendations about their use, and illustrate these with case studies.

DISPERSAL AND CONSERVATION ISSUES IN THE ERA OF GENOMICS

What is dispersal?

Dispersal designates the movement of an individual from its birth patch to its breeding patch (i.e. natal dispersal), or between successive breeding patches (i.e. breeding dispersal) (Clobert *et al.* 2009; Matthysen 2012). Dispersal is usually modelled as a three-stage process including emigration (i.e. departure), transience (i.e. movement into the landscape), and immigration (i.e. arrival). The evolution of dispersal is determined by the balance between the costs incurred during the three stages and fitness benefits associated with settlement in another patch (Bonte *et al.* 2012). Dispersing individuals often differ from resident individuals phenotypically (i.e. *condition-dependent dispersal*, Clobert *et al.* 2004), leading to complex dispersal syndromes (Ronce and Clobert 2012; Cote *et al.* 2017). Moreover, dispersal is a highly plastic behaviour where individuals adjust their emigration and immigration decisions according to social (e.g. public information; Boulinier *et al.* 2008) and environmental cues (i.e. *context-dependent dispersal*) reflecting fitness prospects at a given site (Clobert *et al.* 2009; Matthysen 2012). In addition, individual movements during the transience step are constrained by landscape composition and configuration and the presence of physical barriers that determine connectivity between patches (Baguette and Van Dyck 2007; Baguette *et al.* 2013).

Individual behavioural processes at the three dispersal steps have tremendous effects on the dynamics of spatially structured populations (SSP, *sensu* Thomas and Kunin 1999), which are populations composed of subpopulations occupying discrete habitat patches linked together by dispersal. Dispersal rate, distance, and direction determine

the level of demographic interdependency of subpopulations, local colonisation–extinction dynamics, as well as the whole SSP persistence (Hansson 1991; Hanski and Gaggiotti 2004). Dispersal is therefore central to the *metapopulation* paradigm used by ecologists and evolutionary biologists for decades to understand the links between landscape structure, demography, and evolution (Hanski 1998).

When it is followed by successful reproduction, dispersal becomes a central evolutionary force (i.e. called *gene flow, migration, or effective dispersal* in the literature; Hanski and Gaggiotti 2004; Ronce 2007; Broquet and Petit 2009; Cayuela *et al.* 2018). In combination with genetic drift, selection, and mutation, gene flow shapes population genetic structure, producing various patterns of genetic clustering and clines at different spatial scales (Bohonak 1999; Slatkin 2017). Gene flow strongly influences adaptive processes, by either eroding local adaptation via the swamping of locally beneficial alleles (Lenormand 2002; Tigano and Friesen 2016), or strengthening adaptation via matching habitat choice ('directed gene flow'; Edelaar *et al.* 2008; Jacob *et al.* 2017). In addition, gene flow mitigates the negative effect of genetic drift on genetic diversity and reduces inbreeding depression in small populations through the introduction of foreign genotypes and the masking of deleterious recessive alleles (Whitlock *et al.* 2000; Keller and Waller 2002). Ultimately, gene flow influences speciation by modulating the level of reproductive isolation among populations and lineages (Mayr 1963; Coyne and Orr 2004).

Methods for measuring dispersal in environmental management

Because of its effects on the dynamics and genetics of SSPs, dispersal is an important parameter to consider when wildlife managers develop conservation strategies for threatened species, when practitioners design harvesting plans for exploited species (e.g. seafood), or when national agencies determine plans for pest and disease control. Demographic studies on dispersal usually aim to quantifying annual or seasonal dispersal rates and distances, as well as the degree of demographic interdependency (i.e. demographic connectivity, *sensu* Lowe and Allendorf 2010) among the subpopulations of an SSP. These studies also evaluate functional connectivity (i.e. landscape influence on movement during transience, *sensu* Taylor *et al.* 1993) and examine the influence of patch quality on emigration and immigration (Baguette *et al.* 2013; Cayuela *et al.* 2018). These demographic approaches may help

environmental managers to quantify the degree of population isolation and identify potential habitat corridors, assess the effect of management operations (e.g. habitat restoration) on connectivity, define the limits of protected areas, and evaluate the risks associated with spread of invasive species and pest.

Studying dispersal is usually achieved through two types of approaches: direct observation methods (i.e. individual monitoring using capture-mark-recapture or telemetry approaches) and indirect observation methods (i.e. genetic approaches). Direct observation methods are based on longitudinal individual data with spatial information that allow quantifying dispersal rates and distances at different time scales, and examining finely the influence of phenotype, environmental variation, and their interactions on the three dispersal steps (reviewed in Cayuela *et al.* 2018). These methods are usually costly in terms of logistic and human resources and are not applicable to organisms of very small body size, having very large population sizes (e.g. most invertebrates) or living in environments that cannot be sampled easily (e.g. most marine species). Furthermore, although genetics may be useful for individual identification in some capture-recapture studies (e.g. Cayuela *et al.* 2021; see Chapters 10 and 11), these genotype-based recognition is often costly, especially when it is used in long-term population monitoring.

In the cases where direct observation methods cannot be used due to technical or financial constraints, genetic methods provide interesting methodological solutions to study dispersal (Broquet and Petit 2009; Cayuela *et al.* 2018). Genetic studies of dispersal aim at estimating ancient or contemporary migration rates, assessing gene flow distance, and studying population genetic structure (neutral or adaptive; Broquet and Petit 2009; Cayuela *et al.* 2018). They also examine genetic connectivity, namely the influence of landscape and especially anthropogenic physical barriers on among-patches genetic differentiation (Schwartz *et al.* 2010; Manel and Holderegger 2013). Like demographic approaches, these genetic studies are particularly useful for managers to assess the level of population fragmentation, define the spatial limits of protected areas and fishing stocks, and reconstruct the history of colonisation of invasive species.

Dispersal studies and high-throughput sequencing technologies

The use of genetic markers to study population differentiation and demographic processes dates to the 1940s, when Theodosius Dobzhansky began to compare changes in the frequency of chromosomal rearrangements in natural
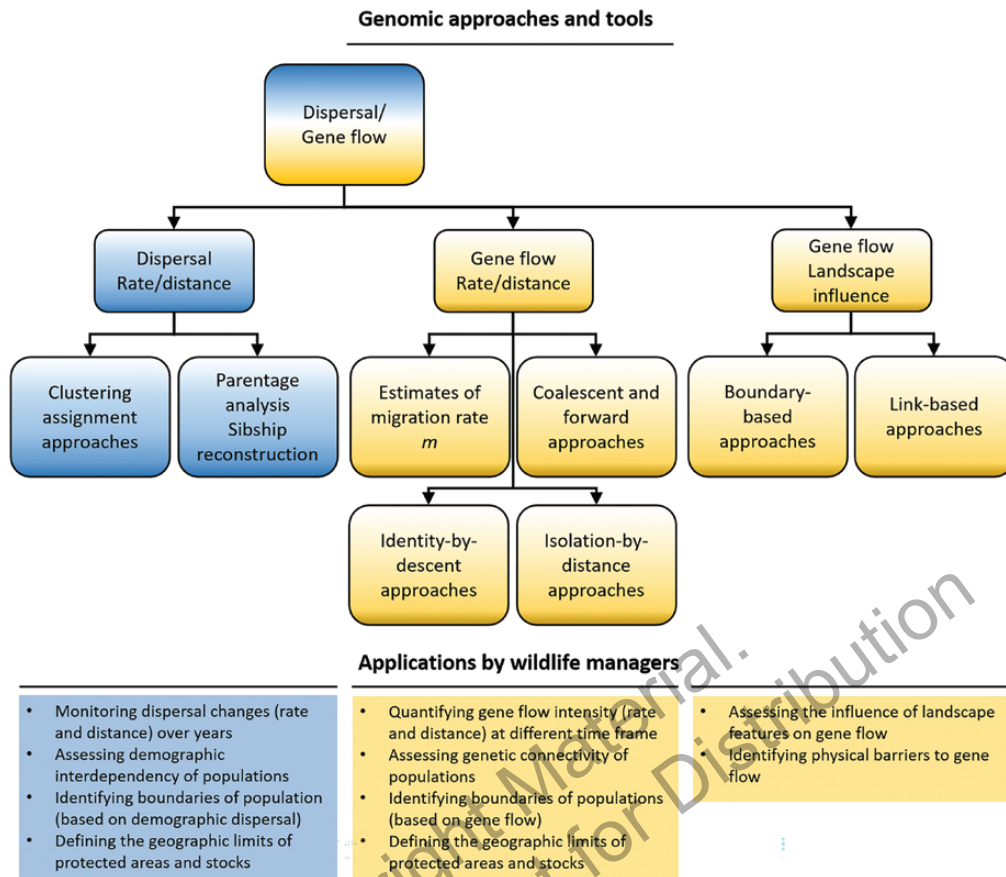
Fig. 22.1. Genomic approaches used to investigate dispersal (in blue) and gene flow (in yellow) and applications by wildlife managers.

populations of *Drosophila pseudoobscura* (Dobzhansky 1943). During the 1990s and 2000s, the use of AFLP (amplified fragment length polymorphism) and microsatellite DNA markers has played a major role in basic and applied studies on dispersal (Sunde *et al.* 2020), allowing the development of completely novel research fields such as *landscape genetics* (Manel *et al.* 2003). During the last two decades, the continuous development of Next Generation Sequencing (NGS) technologies opened up new opportunities for generating more precise estimates of dispersal and gene flow. The use of thousands of genomic markers (i.e. *Single Nucleotide Polymorphisms*, SNP) increases the accuracy of sibship reconstruction, and the robustness of parentage and kinship analyses, used to estimate dispersal (see section 'Parentage and kinship analyses'). In addition, it provides higher statistical power to detect small genetic variation among populations (Waples 1998; Gagnaire *et al.* 2015; Galla *et al.* 2022) and allows increased precision in measures of genetic differentiation used to assess gene flow. Furthermore, NGS approaches facilitate the discovery of markers under selection and hitchhiker markers that usually allow a more efficient delineation of management units

(Flanagan *et al.* 2018; see Chapter 14). Finally, in most methods described below, dense SNP data considerably reduce the needs to sample a large number of individuals in a single sampling locality (e.g. Willing *et al.* 2012). Another benefit of such strong marker density is a better sampling of the whole genome, and a decreased in sampling variance among loci (Morin *et al.* 2004). Here, we review the genomic methods usually considered to investigate dispersal and gene flow (Fig. 22.1). We discuss the strength and limits of each method and provide a general reflection on the use of NGS methods in the framework of dispersal studies.

ESTIMATING DISPERSAL RATES AND DISTANCES USING NEXT GENERATION SEQUENCING
Clustering and assignment approaches
Clustering analyses allow the assignment of sampled individuals to discrete genetic clusters of origin (Pritchard *et al.* 2000; Corander *et al.* 2003), based on their individual genotype and sometimes including spatial information (Guillot *et al.* 2012; Caye *et al.* 2018). These approaches investigate the spatial subdivision of populations and may thus help

identify potential physical barriers to gene flow. The use of these approaches, which are sensitive to a non-exhaustive sample of populations (Puechmaille 2016), is appropriate when populations are effectively fragmented and spatially structured (Lawson *et al.* 2018; Pritchard *et al.* 2000). In such conditions, it is possible to estimate dispersal rates by dividing the number of first-generation dispersers (or F_0, that is, individuals assigned to a cluster different from the ones they were sampled in) by the sample size (Broquet and Petit 2009). Examples of clustering analyses adapted to NGS data are implemented in the program ADMXITURE (Alexander 2009) and the R-package LEA (Frichot and François 2015), among others. Clustering analyses based on SNP data has been used by many studies on broad range of organisms including molluscs (Kess *et al.* 2018), insects (Yadav *et al.* 2021), crustaceans (Vu *et al.* 2021), amphibians (Wei *et al.* 2020), reptiles (van der Zee *et al.* 2021), fishes (Moore *et al.* 2014; Rougemont *et al.* 2019, 2022), birds (Söderquist *et al.* 2017) and mammals (Wultsch *et al.* 2016).

Clustering analyses can nevertheless come with technical limitations, which can be circumvented by some promising new methods. For instance, Isolation-By-Distance (IBD) patterns may lead to an inflation of the signals of population structure since most clustering methods assign individuals to discrete genetic clusters by assuming that allelic frequencies are constant within each cluster and do not consider spatial autocorrelation in allelic frequencies (Frantz *et al.* 2009; Meirmans 2012; but see Galpern *et al.* 2014, Bradburd *et al.* 2018). The R-package *conStruct* (Bradburd *et al.* 2018) provides a statistical framework for the simultaneous inference of continuous and discrete patterns of population structure based on SNP data. This method explicitly addresses the *clines-versus-clusters* problem in modeling population genetic variation. Clustering methods may also show low efficiency when population differentiation is low ($F_{ST} < 0.05$), and/or when the number of genome-wide markers is limited (Paetkau *et al.* 2004; Faubet *et al.* 2007; Meirmans 2014). The R-package *Assigner*, which implements the *gsi_sim* method of Anderson *et al.* (2008), allows assignment analysis to be conducted whose inferences are robust even when population differentiation is low ($F_{ST} < 0.05$).

Parentage and kinship analyses

Parentage analyses (i.e. parentage assignment and sibship reconstruction) use the genotypes of individuals – possibly assessed using SNP data (Snyder-Mackler *et al.* 2016; Grashei *et al.* 2018; Dodds *et al.* 2019; Jasper *et al.* 2019) – to assess parent-offspring relationships and kin relationships

from which dispersal rates and distances can be then quantified (Broquet and Petit 2009; Cayuela *et al.* 2018; see also Chapter 11). Accurate parentage assignments and sibship reconstruction can be obtained from a small number of SNP makers (50–100 SNPs; see Huisman 2017). However, to be efficient, these approaches classically require an extensive individual sampling of both the parents and their progeny (Broquet and Petit 2009) and are therefore generally best suited to small, isolated, and spatially structured populations. Provided that false positives are properly dealt with (Städele and Vigilant 2016), kinship analyses allow assessing familial relationships in the absence of parentage data using various relatedness estimators (Blouin 2003), which can be further used to infer relatedness structure and IBD within spatially structured populations (e.g. amphibians: Unglaub *et al.* 2021; lek-breeding birds: Cayuela *et al.* 2019). For instance, Prunier *et al.* (2021) used kinship reconstruction in the freshwater fish ectoparasite *Tracheliastes polycolpus* to assess the distribution of full-sibs within and among fish hosts and to estimate individual dispersal distances at various spatial scales. At local scales, they showed that *T. polycolpus* sibs tend to be aggregated within fish shoals but not within host individuals, a dispersal strategy that may limit the risk of inbreeding depression. At the river scale, they detected a strong IBD pattern, resulting from a combination of passive upstream-to-downstream dispersal events (passive drift with current) and some active host-driven downstream-to-upstream dispersal events. Parentage and kinship analyses performed to infer dispersal were widely based on microsatellite DNA markers (e.g. Telfer *et al.* 2003; Lepais *et al.* 2010; Woltmann *et al.* 2012; Almany *et al.* 2017; Unglaub *et al.* 2021) whereas the use of SNP markers is still relatively rare (e.g. birds: Aguillon *et al.* 2017; butterflies: Fountain *et al.* 2018; fishes: Feutry *et al.* 2017; Beacham *et al.* 2022).

ESTIMATING GENE FLOW INTENSITY AND DISTANCE USING NEXT GENERATION SEQUENCING

NGS approaches to quantify gene flow

Genetic tools have been historically designed to infer gene flow (Wright 1943; Slatkin 1981). Many of the currently available methods have been originally designed for microsatellite DNA data and were then adapted to be used with SNP data. Some of these approaches allow quantification of gene flow patterns at historical timescales (e.g. coalescent and forward approaches) whereas other enable estimating more contemporary gene flow (e.g. identity-by-descendent

approaches). Landscape genetic/genomic methods provide a framework to evaluate the influence of landscape features on gene flow and to identify physical barriers shaping population genetic structure.

Isolation-By-Distance

Both movement capacity and behaviour of propagules (or of their dispersal vectors) generally make gene flow a spatially limited process, leading to spatial autocorrelation in the distribution of allelic frequencies (Wright 1943). Inter-individual or inter-population genetic measures of relatedness tend to decrease with distance, a process known as Isolation-By-Distance (IBD; Wright 1943). IBD is usually assessed through the regression of pairwise genetic distances (e.g. F_{ST}) against geographic distances, which allows estimating absolute or relative average long-term dispersal distances, provided that populations are at migration-drift equilibrium (see Glossary) and that both population sizes and gene flow are spatially homogeneous. Assuming that local densities were properly estimated, the slope of the regression of $F_{ST}/(1-F_{ST})$ values against geographic distances may for instance allow inferring the absolute mean axial parent–offspring dispersal distance (Rousset 1997; Rousset and Billiard 2000). Alternatively, the distance thresholds at which pairwise correlations become null in spatial correlograms (Sokal and Oden 1978) may be compared across distinct genetic data sets and hence under certain conditions interpreted as estimates of the relative -rather than absolute- spatial extents of gene flow (Vekemans and Hardy 2004). Studying IBD is an excellent alternative to direct approaches in systems where capture-mark-recapture or telemetry is not an option. For instance, Filipović *et al.* (2020) successfully determined the mean dispersal distance of the mosquito *Aedes aegypti*, a major arboviral vector, as 45m, with a 10% probability of dispersal higher than 100m. Other examples can be found for instance in fishes (Puebla *et al.* 2009, De Keyzer *et al.* 2019) and reptiles (Wen *et al.* 2021).

Migration rate, m

An early genetic approach to estimating migration rate (*m* in the population genetic literature) relies on the use of Wright's F-statistics to measure correlation of allele frequencies within and among populations. These statistics were developed under the *island model*, a theoretical model in which each population is made of the same effective number of individuals N_e (or *effective population size*, see Glossary) and receives and provides the same number of immigrants at a rate *m* per generation (symmetric migration). Under this migration-drift equilibrium scenario, there is no selection or mutation and $F_{ST} \approx 1/(4N_e m+1)$ (Wright 1943). This relationship can be used to estimate $N_e m = 1/4(1/F_{ST}-1)$ the product of the migration rate and the effective population size. This approach has been fundamental to study the movement of lineages with geographic distances under the coalescent theory (section 'Coalescent and forward approaches') and to understand patterns of IBD (section 'Isolation-By-Distance'). Nevertheless, the number of assumptions in the above equation makes its interpretation far from straightforward (Whitlock and McCauley 1999; Marko and Hart 2011). Moreover, low F_{ST} can result from either large *Ne* or small *m* (Waples and Gaggiotti 2006) making it necessary to distinguish effective population size from migration. From an applied perspective and given all the body of research related to the above-mentioned limits, we caution against a literal interpretation of F_{ST} in terms of dispersal and against its use for making a direct link with any dispersal process. Fortunately, improved methods have been developed for simultaneously estimating both *Ne* and *m* as presented below.

Coalescent and forward approaches

Coalescent theory describes the movement of lineage as one moves backward in time. Development of the coalescent theory by Kingman (1982) has enabled the emergence of methods to infer population genetic parameters including the migration rate, *m*. When applied to a small number of loci, the coalescent theory, as well as F_{ST}-based methods, provide information about gene genealogy in the deep past. It is therefore informative about patterns of historical migration on long evolutionary time scales (in the order of N_e generations). This approach has been developed initially under the isolation with migration framework (IM) (Wakeley 1996). Recent improvement now allows inferring historical migration under more realistic demographic histories (reviewed in Cayuela *et al.* 2018). Furthermore, developments in Approximate Bayesian Computation (ABC) combined with the flexibility of coalescent simulations enable the implementation of models estimating migration rates among multiple populations connected by gene flow (Beaumont 2019). For instance, recent anthropogenic hybridisation, between domesticated and wild cats could be inferred with ABC with a surprising precision with ABC applied on RAD-sequencing data (Howard-McCombe *et al.* 2021). However, it is worth mentioning that ABC approaches generally rely on the use of summary statistics (e.g. F_{ST}, π, θ) and therefore remain limited to characterising historical gene flow events. Machine

learning methods exploit the raw genotype data, bypassing the compressing step incurred by the use of summary statistics. For instance, Battey *et al.* (2020) developed a neural network to infer the location of individuals based on unphased diploid genotypes and applied their methods to *Plasmodium* parasites, *Anopheles* mosquitoes, and *homo sapiens*. Their work takes advantage of the correlation between the genealogy and geography to infer the location of 'spatial pedigrees', which has been described in great details by Bradburd and Ralph (2019, see Box 22.1). Similarly, Osmond and Coop (2021) developed a method to infer recent dispersal and locate genetic ancestors. Applying their method to the plant *Arabidopsis thaliana*, they were able to infer recent long-distance dispersal. This important progress is made possible by the availability of whole genome sequences data and progress in forward simulation of whole genome across geographic landscapes (Haller and Messer 2019). They make it possible to move beyond characterising dispersal between discrete populations connected by discrete migration events to continuously distributed samples of individuals characterised by a distribution of dispersal distances across recent times. Combining IBD theory and coalescent simulation, Aguillon *et al.* (2017) were able to infer a signal of IBD with dispersal distances ranging from 500 meters to 10 km in Florida Scrub-Jays. While appealing, the use of these methods remained relatively limited due to ongoing progress in methodological development and scarce sequencing data available outside of a few well studied systems. Therefore, while these methods may be practically useful in the future, they are not yet fully relevant for managers with limited budget and sequencing data.

Box 22.1: Illustration of the concept of spatial pedigree by Bradburd and Ralph 2019 and its potential use for wildlife management

The spatial pedigree can be used to track the spatial distribution of genetic ancestors across the geographic landscape at different time points (Bradburd and Ralph 2019). In theory, and with dense enough sampling and whole genome data, it should be possible to infer whether the position of a genetic ancestor has greatly changed in recent time. This information could then be used to draw correlation with human alterations of the connectivity landscape or infer change in the distribution of natural barriers to dispersal. The spatial pedigree has recently been used to infer long range dispersal in *Arabidopsis thaliana* (Osmond and Coop 2021) or to identify ancestor and infer dispersal in three different species (Battey *et al.* 2020).

Figures were taken from Bradburd and Ralph (2019) based on simulation of a spatial pedigree. Figure (a) provides an example of a pedigree with one modern day sample with its ancestors (maternal in red hues and paternal in blue hues, dashed line = mating; solid line = parentage). Figure (b) shows the geographic distribution of individuals and the spatial pedigree. The plane in the spatial pedigree corresponds to a sampled region in non-overlapping generations. Each dot displays the birth location of each individual. The pedigree of the focal individual is highlighted back through time and across space. If there is a strong change in connectivity (for instance due to landscape artificialisation) associated with an alteration of the population demography, variation in pedigree through space and time is expected.

Bradburd GS, Ralph PL. 2019.
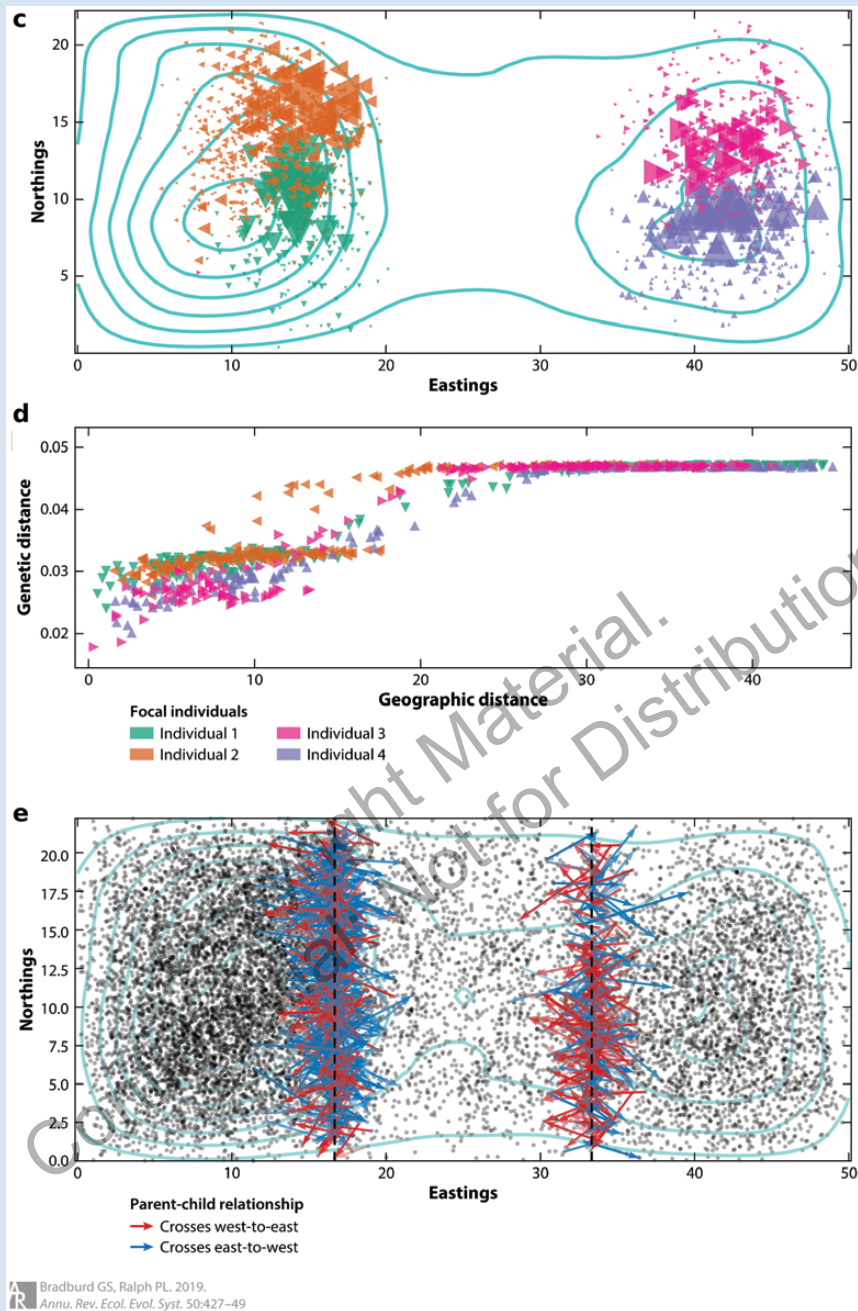*Annu. Rev. Ecol. Evol. Syst.* 50:427–49
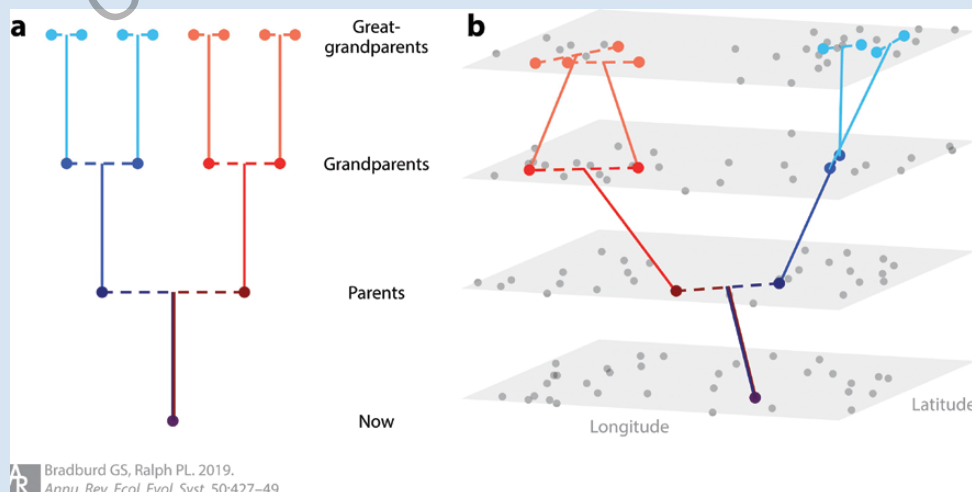
Figure (c) shows a map of four focal individuals along with their close genealogical relatives, each represented by a different colour with their geographical location. The triangle size is proportional to the expected proportion of genome shared. Figure (d) displays the positive relationship between genetic distance and geographic distance for the four focal individuals and 100 other individuals corresponding to panel a). Figure (e) shows the dispersal flow (i.e. dispersal events at each generation) across the two valleys separated by a given boundary on the same continuous map; all individuals from Figure (a) are shown as small dots at approximately 8 generations and every parent-child relationship crossing a dotted line is represented by red arrow for west-to-east crosses and blue arrow for east-to-west crosses.

Overall, the spatial pedigree is an area of active research and a promising approach to study the influence of landscape modifications (e.g. habitat fragmentation and loss) and conservation management (restoration of ecological corridors) on fine-scale dispersal movement across multiple generations.

Identity-by-descent blocks

The decreasing cost of whole genome sequencing enables the analysis of long continuous tracks of the genome known as haplotype blocks (reviewed in Browning and Browning 2012; Leitwein et al. 2020). These tracks, when unbroken by recombination, form large blocks whose size is gradually reduced at each generation of reproduction (Leitwein et al. 2020). Therefore, it is possible to study the distribution of track length between hybridising populations to evaluate the timescale and intensity of recent gene flow (Pool and Nielsen 2009; Gravel 2012). Similarly, blocks that are identical-by-descent (i.e. haplotypes inherited t generations ago from a common ancestor by a pair of individuals) can be used to infer migration rates (m) on recent time scales (Browning and Browning 2012; Palamara et al. 2012). The length of a segment decreases approximately at a rate $1/t$ and can therefore provide information on dispersal from a few tens of generations ago. A potentially useful method was developed recently to infer recent dispersal distance using these IBD blocks (Ringbauer et al. 2017). Similarly, Al-Asadi et al. (2019) developed maps of population effective size and migration rate at different time periods. The accuracy of the inference from these methods is still coarse and will be affected by the population demographic history, suggesting that substantial methodological improvement could be made in the future. While promising, to the best of our knowledge, the methods of Ringbauer et al. (2017) and Al-Asadi et al. (2019) have not been deployed beyond human, due to the lack of genomic resources in managed species or species with conservation issues. Nonetheless, an interesting approach based on admixture tracts has been developed and used in an exploited species, the Atlantic and Mediterranean seabass, as exemplified in Box 22.2. Overall, approaches based on Identity-by-descent blocks could be used to measure contemporary dispersal and recent changes in genetic connectivity (for instance due to anthropogenic factors) among populations without suffering the limitations associated to F_{st}-based measures. These methods could therefore be of high relevance to best manage populations and improve opportunity for gene flow between isolated patches having undergone strong human-driven fragmentation.

INVESTIGATING THE INFLUENCE OF ENVIRONMENT ON TRANSIENCE AND SUBSEQUENT GENE FLOW USING NGS APPROACHES

The transience phase of dispersal may be affected by various between-patch processes such as habitat fragmentation or predation (e.g. Fahrig 2007; Bonnot et al. 2013), in turn influencing gene flow and spatial genetic structures. Two main kinds of tools may be considered to indirectly investigate these processes, namely boundary-based and link-based methods (Wagner and Fortin 2013). Because of their genome-wide distribution, SNPs may provide a better spatiotemporal resolution than markers with lower genome coverage (e.g. microsatellites) in the detection of weak and/or recent landscape-driven changes in dispersal patterns (McCartney-Melstad et al. 2018; Maigret et al. 2020), provided only neutral markers are considered (Storfer et al. 2018; but see Landguth and Balkenhol 2012).

Boundary-based approaches

Boundary-based methods are primarily used to spatially delineate populations through the detection of significant edges between genetic entities (Jacquez et al. 2000) or through the identification of homogeneous genetic clusters, with and without assumptions about underlying population genetic models (i.e. markers are at both Hardy-Weinberg and linkage equilibrium, see Glossary; Jombart et al. 2008, 2009; François and Durand 2010). Once mapped, inferred genetic boundaries can be visually or statistically examined with respect to landscape features so as to identify potential barriers to gene flow (Balkenhol et al. 2009). For instance, LaCava et al. (2021) used both Principal Component Analysis (PCA) and Bayesian clustering to spatially delineate three admixed genetic groups in the mule deer Odocoileus hemionus. They then used a link-based approach to assess landscape functional connectivity both within each group and across groups. They revealed important regional variations in functional connectivity in that species, with gene flow in some groups being only limited by IBD, whereas gene flow in other groups were constrained by elevation, vegetation cover and/or anthropogenic infrastructures such as highways.

Link-based approaches

Link-based methods are used to assess the influence of landscape heterogeneity on genetic differentiation between habitat patches, here considered a proxy for extent of gene flow. Link-based methods generally rely on the statistical comparison of pairwise matrices and are designed to explicitly take the inherent non-independence of pairwise data into consideration. The primary frameworks are the Mantel test (Mantel 1967) and its various extensions and refinements (e.g. Smouse et al. 1986; Ferrier et al. 2007; Crabot et al. 2019), the maximum-likelihood population-effects model (MLPE; Clarke et al. 2002; Selkoe et al. 2010; Van Strien

et al. 2012), non-linear Bayesian or machine-learning algorithms (e.g. Bradburd *et al.* 2013; Peterman 2018; Box 22.3) and causal models (Fourtune *et al.* 2018). Pairwise genetic distances (response matrix) may be computed at the individual or at the population level from a stratified random sampling design (Anderson 2010; Storfer *et al.* 2018). When the sampling units are populations, a limitation of link-based methods is that they are only valid under the strict assumption of equal effective population sizes, or once the spatial heterogeneity in drift has been taken into account (e.g. Murphy *et al.* 2010; Prunier *et al.* 2017). The use of genome-wide SNPs may allow increasing the number of sampling locations and thus optimising the spatial sampling design (Prunier *et al.* 2013) since reliable allelic frequencies may be obtained from a reduced number of samples from each deme (Willing *et al.* 2012). Landscape distances (predictor matrices) represent the probability of transience between sampling locations under the hypotheses of IBD (e.g. Euclidean distances) and/or isolation-by-resistance (IBR; e.g. distances based on least-cost-paths or circuit theory; McRae and Beier 2007). In the latter case, different

approaches may be considered to assign resistance values to landscape features (Spear *et al.* 2010; Peterman *et al.* 2019) and to take spatial autocorrelation into account (e.g. Dray *et al.* 2006; Arjona *et al.* 2020). Linked-based methods using SNP markers have been deployed to assess the influence of landscape configuration on patterns of gene flow in a diversity of organisms, such as mammals (Kozakiewicz *et al.* 2019; LaCava *et al.* 2021), insects (Trense *et al.* 2021; Dudaniec *et al.* 2022), amphibians (Ferreira *et al.* 2020), and reptiles (Maigret *et al.* 2020) (see also Box 22.3). For instance, Creel *et al.* (2019) used RAD sequencing data and Mantel tests to compare the impact of anthropogenic activities as measured from a spatial Human Footprint Index on functional connectivity in two interacting large African carnivores, the lion (*Panthera leo*), a dominant competitor, and the African wild dog (*Lycaon pictus*), a subordinate competitor. They found a much stronger impact of anthropogenic activities on patterns of gene flow in lions than in African wild dogs, in line with the hypothesis that subordinate species are able (and adapted) to cross unfavourable landscape areas to reduce competition with dominant species.

Box 22.2: Using admixture tracts to infer dispersal distance: study case of Duranton *et al.*

Duranton *et al.* (2019) used the distribution of admixture tracts to infer dispersal distance in the European sea bass (*Dicentrarchus labrax*) within the Mediterranean Sea (the figure below, panel (a)). The authors considered individuals from the eastern Mediterranean Sea (orange circle, in panel (a)) and from the western Mediterranean Sea (red circle). Admixture between Atlantic (red chromosomes in panel (b)) and Mediterranean (yellow in panel (b)) sea bass lineages resulted in the introgression of Atlantic tracts in the Mediterranean Sea (panel (b), red arrows display the diffusion process). As shown in
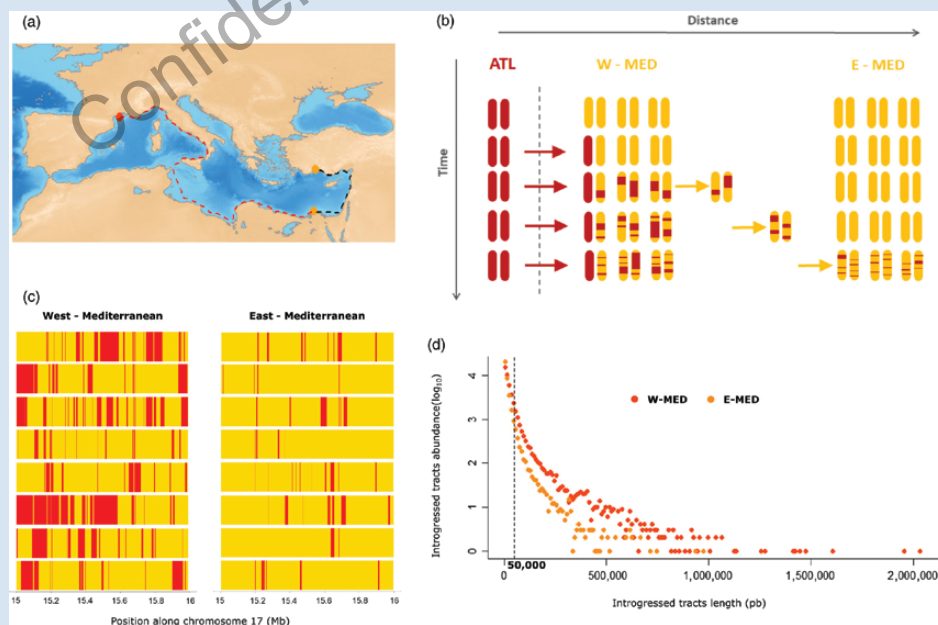
panel (b), the blocks of foreign ancestry are shortened by recombination during the diffusion process of Atlantic block from the west to the east (yellow arrow) of the Mediterranean Sea, enabling the estimation of effective dispersal. One individual (two chromosomes) is represented in the Atlantic and three-individuals (six chromosomes) are represented in western and eastern populations. As expected, due to recombination (see panel (c)), the distribution of ancestry tracts from Atlantic population (red) to the Mediterranean (yellow) populations along the chromosome 17 differs between the western Mediterranean (W-MED) and eastern Mediterranean (E-MED) populations. The abundance of tracts was negatively correlated with their length (panel (d)) in both the W-MED and E-MED populations, with shorter tracts expected to provide information on a longer time-scale. To estimate recent dispersal, Duranton *et al.* limited their analyses to tracks longer than 50kb (vertical dotted line in panel (d)). From an applied point of view, this study emphasises that the refined knowledge of dispersal and connectivity provided by the analysis of admixture tracts may help delineate Marine Protected Areas in a fish with high commercial value.

INVESTIGATING THE INFLUENCE OF ENVIRONMENT ON EMIGRATION-IMMIGRATION AND SUBSEQUENT GENE FLOW

Methods used to study the transience phase of dispersal usually rely on the assumption of random gene flow, all individuals being considered equally likely to disperse (Edelaar and Bolnick 2012). However, individuals may make informed choices, deciding to leave a patch (emigration) and/or to settle in a new one (immigration) depending on both their internal state and their perception of local environment (e.g. local carrying capacity or patch quality; Edelaar *et al.* 2008; Clobert *et al.* 2009). Furthermore, settlement may fail because of local environment or sexual selection acting upon some immigrants (Nosil *et al.* 2005). Understanding the whole dispersal process thus implies disentangling the influence of these at-site mechanisms on the spatial distribution of genome-wide allelic frequencies from the influence of among-site landscape processes impacting the transience phase. The isolation-by-environment (IBE) framework states that neutral pairwise genetic distances should increase with environmental differences between patches because of the non-randomness of immigration, whatever the processes operating during transience (Wang *et al.* 2013; Wang and Bradburd 2014). That is, organisms are most likely to move between similar environments (habitat matching choice; Edelaar *et al.* 2008) or may fail to settle in environments that differ from their original one (local adaptation). IBE can be assessed using linked-based statistical analyses, by measuring the effect of pairwise IBE metrics on pairwise genetic distances, in addition to classical IBD and/or IBR pairwise metrics (e.g. Sexton *et al.* 2014; Murray *et al.* 2019). For instance, Fenker *et al.* (2021) used SNP markers to explore whether dispersal was limited by geographic distance (IBD), geographic barriers (IBR) and/or environmental heterogeneity (IBE) in several co-distributed lizard taxa from the topographically and climatically complex monsoonal tropics of Australia. They found that IBE was the strongest predictor of genetic divergence in all taxa, with both temperature and precipitation acting as factors limiting dispersal, suggesting patterns of local adaptation. By considering both at-site (nodes) and among-site (edges) processes within a network-based analysis, gravity models (Fotheringham and O'Kelly 1989) allow identifying the respective drivers of emigration (e.g. patch productivity) and immigration (e.g. predation) and disentangling their influence and that of transience on spatial patterns of genetic differentiation (Murphy *et al.* 2010; Parsley *et al.* 2020). Gravity models might also be used to rule out the influence of spatial heterogeneity in effective population sizes, a prerequisite for the proper interpretation of genetic differentiation in terms of gene flow.

FUTURE DIRECTIONS

We have provided a synthetic snapshot of the most recently developed analytical methods to quantify dispersal, with a specific emphasis on wildlife management issues. Among the approaches that are currently under active development, long identical-by-descent blocks and migrants tracts appear highly promising as they can potentially provide information on gene flow (and population size) on recent timescale both within spatially structured populations and among diverging populations (reviewed in Leitwein *et al.* 2020). Recent timescales are likely the most relevant for conservation geneticists who may want to investigate how environmental factors constraint gene flow and effective sizes, identify population boundaries, and define management areas (see the example in Boxes 22.2 and 22.3). These approaches stand in contrast with methods relying on the site frequency spectrum (Gutenkunst *et al.* 2009), which provides information on long evolutionary timescale (in the order of N_e generations).

One of the main limitations of identity-by-descent methods is the need for large amount of genetic data, namely dense SNP data from whole genome sequencing or SNP array, and eventually a linkage map (e.g. Ringbauer *et al*. 2017). However, these methods can be used with a small number of spatially geolocalised individuals in the study area, which can be useful in the context conservation programs on rare/elusive species. Successful application of identity-by-descent methods will certainly necessitate extensive collaboration between practitioners and research laboratories. To date, to the best our knowledge, the study of Duranton *et al*. (2019) is the only study case using identity-by-descent tracts to infer dispersal and connectivity in a context of wildlife management (stock management of the Atlantic sea-bass in the Mediterranean basin), as detailed in Box 22.2. However, with decreasing cost of sequencing technologies, this approach could be applied in the near future in many species of conservation concerns. Indeed, recent technological progress currently makes long read sequencing available for conservation genomics studies at relatively low cost (Browning and Browning 2011; Meier *et al*. 2021).

Box 22.3: Using linked-based methods to quantify and predict population connectivity of an outbreaking forest insect pest: study case of Larroque *et al*.

The eastern spruce budworm (*Choristoneura fumiferana*) is a lepidopteran that periodically outbreaks (~ every 35 years) and defoliates millions of hectares of forests in North America, with severe socio-economic consequences. In Quebec, the last outbreak started in 2006. Larroque *et al*. (2022) investigated the spatial environmental determinants of dispersal in this species using 3562 SNP loci from 447 larvae covering most of the outbreak area. They optimised landscape resistance surfaces given pairwise measures of genetic differentiation using the machine-learning algorithm ResistanceGA (Peterman 2018), and found that landscape connectivity decreased drastically when mean daily July precipitation was less than 50mm (figure below, panel (a)), suggesting that water stress may hinder both dispersal and the probability of success of settlement through survival. Mapping the most likely paths of the spread of the outbreak given present environmental conditions, they identified one large high connectivity area surrounded by two low-connectivity areas to the East and West (panel (b)), suggesting that the western part of Quebec could potentially act as a barrier with Ontario. Similarly, mapping the most likely dispersal paths given expected environmental conditions in the next outbreak (~2040), they predicted an increase in connectivity in the western part of Quebec (panel (c)), paving the way to targeted pest control policies in well-connected but not yet attacked areas, through pre-emptive treatments such as spraying or harvest.
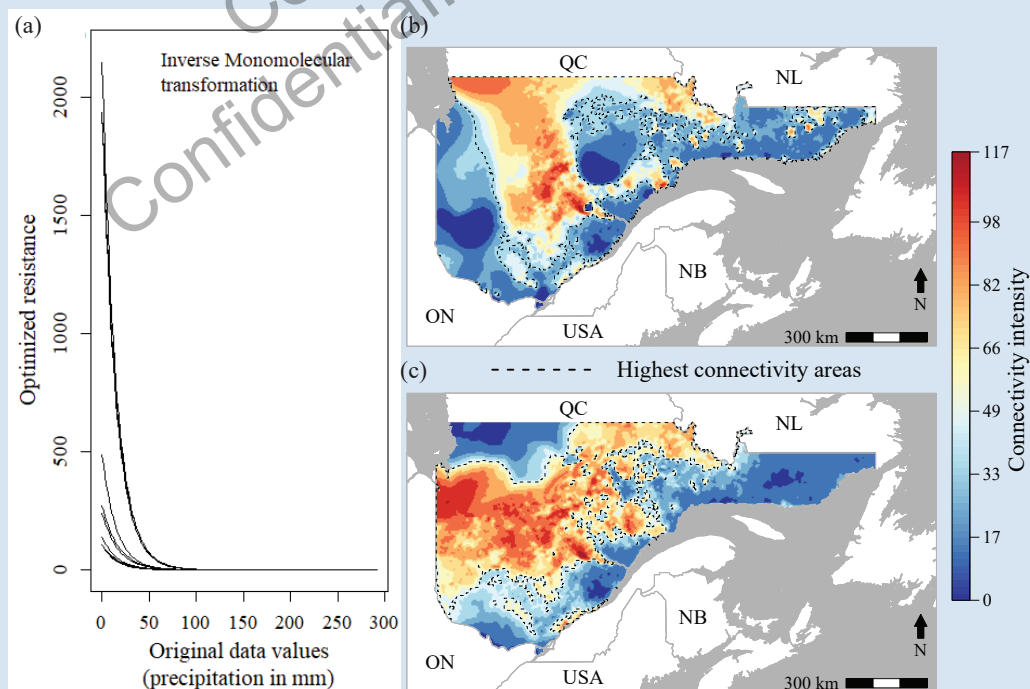
Lastly, we emphasise the critical importance of the genomic marker choice to investigate dispersal and gene flow. The vast majority of the methods described above rely on population genetic concepts that holds true under neutrality. However, the neutrality hypothesis is often neglected by genomic studies that quantify gene flow or assess landscape connectivity from a mix of neutral and adaptive markers. We argue that gene flow estimates (rate and distance) should preferentially be performed using markers that are virtually non-affected by selection to meet the neutrality assumption of population genetic theory. This step of marker selection is especially important in study systems (e.g. many marine organisms) with high gene flow and/or large effective population size since adaptive markers may have a strong contribution to population genetic structure in these systems (Xuereb *et al.* 2021). When a reference genome is available in the studied species, the filtering of non-neutral markers can be achieved by removing the SNPs that are located within the sequence of protein-coding genes and regulatory regions, and considering the extent of linkage disequilibrium around these genomic regions and keeping SNPs in the most highly recombining regions (Pouyet *et al.* 2018; see also Chapter 14). When no reference genome is available, putatively adaptive SNPs associated with environmental variables can be identified using usual genome scan approaches (e.g. F_{ST}-based analysis) such as BayPass (Gautier 2015) or mixed models (implemented in LFMM2, Caye *et al.* 2019). We recommend employing conservative thresholds when delimiting putatively neutral and adaptive markers since processes such as linked selection and biased gene conversion may affect neutrality along large stretches of the genome (Pouyet *et al.* 2018; see also Johri *et al.* 2022 for a review and recommendations).

CONCLUSION

Dispersal plays a central role in population dynamics and viability and is therefore of critical importance for wildlife management. In many organisms, quantifying dispersal using direct observations is a complicated task and genetic approaches have proven particularly useful for studying gene flow resulting from dispersal. For a long time, studies on gene flow suffered from a lack of power to assess the genetic structure of populations due to a limited number of markers. Twenty years ago, the advent of NGS approaches – generating thousands of SNP markers – revolutionised our ability to study dispersal and gene flow, especially in the context of conservation operations, natural resource management, and pest control programs. The ongoing decrease in sequencing costs allows assembling reference genomes and producing large amounts of genomic data that are particularly useful to analyse dispersal and its genetic consequences at different spatial and temporal scales. NGS approaches provide wildlife managers with robust tools that may improve the delimitation of conservation and harvesting areas, notably in species showing seemingly weak genetic structure. In addition, these methods allow estimating accurate dispersal parameters even when the number of available samples (i.e. individuals) is small, which it is very often the case in rare, endangered species.

REFERENCES

Aguillon SM, Fitzpatrick JW, Bowman R, Schoech SJ, Clark AG, *et al.* (2017) Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. *PLoS Genetics* **13**, e1006911.

Al-Asadi H, Petkova D, Stephens M, Novembre J (2019) Estimating recent migration and population-size surfaces. *PLOS Genetics* **15**, e1007908.

Alexander DH, NovembreJ, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* **19**, 1655–1664.

Almany GR, Planes S, Thorrold SR, Berumen ML, Bode M, *et al.* (2017) Larval fish dispersal in a coral-reef seascape. *Nature Ecology & Evolution* **1**, 1–7.

Anderson CD (2010) Effects of movement and mating patterns on gene flow among overwintering hibernacula of the Timber Rattlesnake (*Crotalus horridus*). *Copeia* **1**, 54–61.

Anderson EC, Waples RS, Kalinowski ST (2008) An improved method for predicting the accuracy of genetic stock identification. *Canadian Journal of Fisheries and Aquatic Sciences* **65**, 1475–1486.

Arjona Y, Fernández-López J, Navascués M, Alvarez N, Nogales M, *et al.* (2020) Linking seascape with landscape genetics: Oceanic currents favour colonization across the Galápagos Islands by a coastal plant. *Journal of Biogeography* **47**, 2622–2633.

Baguette M, Van Dyck H (2007) Landscape connectivity and animal behavior: Functional grain as a key determinant for dispersal. *Landscape Ecology* **22**, 1117–1129.

Baguette M, Blanchet S, Legrand D, Stevens VM, Turlure C (2013) Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews* **88**, 310–326.

Balkenhol N, Waits LP, Dezzani RJ (2009) Statistical approaches in landscape genetics: An evaluation of methods for linking landscape and genetic data. *Ecography* **32**, 818–830.

Battey CJ, Ralph PL, Kern AD (2020) Predicting geographic location from genetic variation with deep neural networks. *ELife* **9**, e54507.

Beacham TD, Jonsen K, Sutherland BJG, Lynch C, Rondeau EB (2022) Assessment of mixed-stock fisheries and hatchery broodstocks for coho salmon in British Columbia, Canada via parentage-based tagging and genetic stock identification. *Fisheries Research* **245**, 106136.

Beaumont MA (2019) Approximate bayesian computation. *Annual Review of Statistics and Its Application* **6**, 379–403.

Blouin MS (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology & Evolution* **18**, 503–511.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology* **74**, 21–45.

Bonnot N, Morellet N, Verheyden H, Cargnelutti B, Lourtet B, *et al.* (2013) Habitat use under predation risk: Hunting, roads and human dwellings influence the spatial behaviour of roe deer. *European Journal of Wildlife Research* **59**, 185–193.

Bonte D, Van Dyck H, Bullock JM, Coulon A, Delgado M, *et al.* (2012) Costs of dispersal. *Biological Reviews* **87**, 290–312.

Boulinier T, McCoy KD, Yoccoz NG, Gasparini J, Tveraa T (2008) Public information affects breeding dispersal in a colonial bird: kittiwakes cue on neighbours. *Biology Letters* **4**, 538–540.

Bradburd GS, Ralph PL, Coop GM (2013) Disentangling the effects of geographic and ecological isolation on genetic differentiation: Isolation by geographic and ecological distance. *Evolution* **67**, 3258–3273.

Bradburd GS, Coop GM, Ralph PL (2018) Inferring continuous and discrete population genetic structure across space. *Genetics* **210**, 33–52.

Bradburd GS, Ralph PL (2019) Spatial population genetics: It's about time. *Annual Review of Ecology, Evolution, and Systematics* **50**, 427–449.

Broquet T, Petit EJ (2009) Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics* **40**, 193–216.

Browning BL, Browning SR (2011) A fast, powerful method for detecting identity by descent. *The American Journal of Human Genetics* **88**, 173–182.

Browning SR, Browning BL (2012) Identity by descent between distant relatives: detection and applications. *Annual Review of Genetics* **46**, 617–633.

Caye K, Jay F, Michel O, François O (2018) Fast inference of individual admixture coefficients using geographic data. *The Annals of Applied Statistics* **12**, 586–608.

Caye K, Jumentier B, Lepeule J, François O (2019) LFMM 2: Fast and accurate inference of gene-environment associations in genome-wide studies. *Molecular Biology and Evolution* **36**, 852–860.

Cayuela H, Rougemont Q, Prunier JG, Moore JS, Clobert J, *et al.* (2018) Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Molecular Ecology* **27**, 3976–4010.

Cayuela H, Boualit L, Laporte M, Prunier JG, Preiss F, *et al.* (2019) Kin-dependent dispersal influences relatedness and genetic structuring in a lek system. *Oecologia* **191**, 97–112.

Cayuela H, Prunier JG, Laporte M, Gippet JM, Boualit L, *et al.* (2021) Demography, genetics, and decline of a spatially structured population of lekking bird. *Oecologia* **195**, 117–129.

Clarke RT, Rothery P, Raybould AF (2002) Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics* **7**, 361–372.

Clobert J, Ims RA, Rousset F (2004) Causes, mechanisms and consequences of dispersal. In *Ecology, Genetics and Evolution of Metapopulations.* (Eds I Hansk, OE Gaggiotti) pp. 307–335. Academic Press.

Clobert J, Le Galliard J-F, Cote J, Meylan S, Massot M (2009) Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology Letters* **12**, 197–209.

Corander J, Waldmann P, Sillanpää MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics* **163**, 367–374.

Cote J, Bestion E, Jacob S, Travis J, Legrand D, *et al.* (2017) Evolution of dispersal strategies and dispersal syndromes in fragmented landscapes. *Ecography* **40**, 56–73.

Coyne JA, Orr HA (2004) *Speciation.* Sinauer & Associates, Sunderland, MA, USA.

Crabot J, Clappe S, Dray S, Datry T (2019) Testing the Mantel statistic with a spatially-constrained permutation procedure. *Methods in Ecology and Evolution* **10**, 532–540.

Creel S, Spong G, Becker M, Simukonda C, Norman A, *et al.* (2019) Carnivores, competition and genetic connectivity in the Anthropocene. *Scientific Reports* **9**, 16339.

De Keyzer ELR, De Corte Z, Van Steenberge M, Raeymaekers JAM, Calboli FCF, *et al.* (2019) First genomic study on Lake Tanganyika sprat *Stolothrissa tanganicae*: A lack of population structure calls for integrated management of this important fisheries target species. *BMC Evolutionary Biology* **19**, 6.

Dobzhansky T (1943) Genetics of natural populations IX. Temporal changes in the composition of populations of Drosophila pseudoobscura. *Genetics* **28**, 162.

Dodds KG, McEwan JC, Brauning R, van Stijn TC, Rowe SJ, *et al.* (2019) Exclusion and genomic relatedness methods for assignment of parentage using genotyping-by-sequencing data. *G3: Genes, Genomes, Genetics* **9**, 3239–3247.

Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* **196**, 483–493.

Dudaniec RY, Carey AR, Svensson EI, Hansson B, Yong CJ, *et al.* (2022) Latitudinal clines in sexual selection, sexual size dimorphism and sex-specific genetic dispersal during a poleward range expansion. *Journal of Animal Ecology* **91**, 1104–1118.

Duranton M, Bonhomme F, Gagnaire PA (2019) The spatial scale of dispersal revealed by admixture tracts. *Evolutionary Applications* **12**, 1743–1756.

Edelaar P, Siepielski AM, Clobert J (2008) Matching habitat choice causes directed gene flow: A neglected dimension in evolution and ecology. *Evolution* **62**, 2462–2472.

Edelaar P, Bolnick DI (2012) Non-random gene flow: An under-appreciated force in evolution and ecology. *Trends in Ecology & Evolution* **27**, 659–665.

Fahrig L (2007) Non-optimal animal movement in human-altered landscapes. *Functional Ecology* **21**, 1003–1015.

Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Molecular Ecology* **16**, 1149–1166.

Fenker J, Tedeschi LG, Melville J, Moritz C (2021) Predictors of phylogeographic structure among codistributed taxa across the complex Australian monsoonal tropics. *Molecular Ecology* **30**, 4276–4291.

Ferreira AS, Lima AP, Jehle R, Ferrão M, Stow A (2020) The influence of environmental variation on the genetic structure of a poison frog distributed across continuous Amazonian rainforest. *Journal of Heredity* **111**, 457–470.

Ferrier S, Manion G, Elith J, Richardson K (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* **13**, 252–264.

Feutry P, Berry O, Kyne PM, Pillans RD, Hillary RM, *et al.* (2017) Inferring contemporary and historical genetic connectivity from juveniles. *Molecular Ecology* **26**, 444–456.

Filipović I, Hapuarachchi HC, Tien W-P, Razak MABA, Lee C, *et al.* (2020) Using spatial genetics to quantify mosquito dispersal for control programs. *BMC Biology* **18**, 104.

Flanagan SP, Forester BR, Latch EK, Aitken SN, Hoban S (2018) Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evolutionary Applications* **11**, 1035–1052.

Fotheringham AS, O'Kelly ME (1989) *Spatial Interaction Models: Formulations and Applications*. Kluwer Academic Publishers, Dordrecht, Netherlands.

Fountain T, Husby A, Nonaka E, DiLeo MF, Korhonen JH, *et al.* (2018) Inferring dispersal across a fragmented landscape using reconstructed families in the Glanville fritillary butterfly. *Evolutionary Applications* **11**, 287–297.

Fourtune L, Prunier JG, Paz-Vinas I, Loot G, Veyssière C, *et al.* (2018) Inferring causalities in landscape genetics: An extension of Wright's causal modeling to distance matrices. *The American Naturalist* **191**, 491–508.

François O, Durand E (2010) Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources* **10**, 773–784.

Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: Clusters or isolation by distance? *Journal of Applied Ecology* **46**, 493–505.

Frichot E, François O (2015) LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution* **6**, 925–929.

Gagnaire PA, Broquet T, Aurelle D, Viard F, Souissi A, *et al.* (2015) Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications* **8**, 769–786.

Galla SJ, Brown L, Couch-Lewis Y, Cubrinovska I, Eason D, *et al.* (2022) The relevance of pedigrees in the conservation genomics era. *Molecular Ecology* **31**, 41–54.

Galpern P, Peres-Neto PR, Polfus J, Manseau M (2014) MEMGENE: Spatial pattern detection in genetic distance data. *Methods in Ecology and Evolution* **5**, 1116–1120.

Gautier M (2015) Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics* **201**, 1555–1579.

Grashei KE, Ødegård J, Meuwissen TH (2018) Using genomic relationship likelihood for parentage assignment. *Genetics Selection Evolution* **50**, 1–11.

Gravel S (2012) Population genetics models of local ancestry. *Genetics* **191**, 607–619.

Guillot G, Renaud S, Ledevin R, Michaux J, Claude J (2012) A unifying model for the analysis of phenotypic, genetic, and geographic data. *Systematic Biology* **61**, 897–911.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics* **5**, e1000695.

Haller BC, Messer PW (2019) SLiM 3: Forward genetic simulations beyond the Wright–Fisher model. *Molecular Biology and Evolution* **36**, 632–637.

Hanski I (1998) Metapopulation dynamics. *Nature* **396**, 41–49.

Hanski IA, Gaggiotti OE (2004) *Ecology, Genetics and Evolution of Metapopulations*. Academic Press.

Hansson L (1991) Dispersal and connectivity in metapopulations. *Biological Journal of the Linnean Society* **42**, 89–103.

Howard-McCombe J, Ward D, Kitchener AC, Lawson D, Senn HV, *et al.* (2021) On the use of genome-wide data to model and date the time of anthropogenic hybridisation: An example from the Scottish wildcat. *Molecular Ecology* **30**, 3688–3702.

Huisman J (2017) Pedigree reconstruction from SNP data: Parentage assignment, sibship clustering and beyond. *Molecular Ecology Resources* **17**, 1009–1024.

Jacob S, Legrand D, Chaine AS, Bonte D, Schtickzelle N, *et al.* (2017) Gene flow favours local adaptation under habitat choice in ciliate microcosms. *Nature Ecology & Evolution* **1**, 1407–1410.

Jacquez GM, Maruca S, Fortin MJ (2000) From fields to objects: A review of geographic boundary analysis. *Journal of Geographical Systems* **2**, 221–241.

Jasper M, Schmidt TL, Ahmad NW, Sinkins SP, Hoffmann AA (2019) A genomic approach to inferring kinship reveals limited intergenerational dispersal in the yellow fever mosquito. *Molecular Ecology Resources* **19**, 1254–1264.

Johri P, Aquadro CF, Beaumont M, Charlesworth B, Excoffier L, *et al.* (2022) Recommendations for improving statistical inference in population genomics. *PLoS Biology* **20**, e3001669.

Jombart T, Devillard S, Dufour AB, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* **101**, 92–103.

Jombart T, Dray S, Dufour A-B (2009) Finding essential scales of spatial variation in ecological data: A multivariate approach. *Ecography* **32**, 161–168.

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. *Trends in Ecology & Evolution* **17**, 230–241.

Kess T, Galindo J, Boulding EG (2018) Genomic divergence between Spanish *Littorina saxatilis* ecotypes unravels limited admixture and extensive parallelism associated with population history. *Ecology and Evolution* **8**, 8311–8327.

Kingman JFC (1982) The coalescent. *Stochastic Processes and their Applications* **13**, 235–248.

Kozakiewicz CP, Burridge CP, Funk WC, Salerno PE, Trumbo DR, *et al.* (2019) Urbanization reduces genetic connectivity in bobcats (*Lynx rufus*) at both intra– and interpopulation spatial scales. *Molecular Ecology* **28**, 5068–5085.

LaCava MEF, Gagne RB, Gustafson KD, Oyler-McCance S, Monteith KL, *et al.* (2021) Functional connectivity in a continuously distributed, migratory species as revealed by landscape genomics. *Ecography* **44**, 987–999.

Landguth EL, Balkenhol N (2012) Relative sensitivity of neutral versus adaptive genetic data for assessing population differentiation. *Conservation Genetics* **13**, 1421–1426.

Larroque J, Wittische J, James PMA (2022) Quantifying and predicting population connectivity of an outbreaking forest insect pest. *Landscape Ecology* **37**, 763–778.

Lawson DJ, Van Dorp L, Falush D (2018) A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications* **9**, 1–11.

Leitwein M, Duranton M, Rougemont Q, Gagnaire PA, Bernatchez L (2020) Using haplotype information for conservation genomics. *Trends in Ecology & Evolution* **35**, 245–258.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution* **17**, 183–189.

Lepais O, Darvill BEN, O'connor S, Osborne JL, Sanderson RA, *et al.* (2010) Estimation of bumblebee queen dispersal distances using sibship reconstruction method. *Molecular Ecology* **19**, 819–831.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology* **19**, 3038–3051.

Maigret TA, Cox JJ, Weisrock DW (2020) A spatial genomic approach identifies time lags and historical barriers to gene flow in a rapidly fragmenting Appalachian landscape. *Molecular Ecology* **29**, 673–685.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* **18**, 189–197.

Manel S, Holderegger R (2013) Ten years of landscape genetics. *Trends in Ecology & Evolution* **28**(10), 614–621.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209–220.

Marko PB, Hart MW (2011) The complex analytical landscape of gene flow inference. *Trends in Ecology & Evolution* **26**, 448–456.

Matthysen E (2012) Multicausality of dispersal: A review. In *Dispersal Ecology and Evolution*. (Eds J Clobert, M Baguette, TG Benton, JM Bullock) pp. 3–18. Oxford University Press.

Mayr E (1963) *Animal Species and Evolution*. Harvard University Press.

McCartney-Melstad E, Vu JK, Shaffer HB (2018) Genomic data recover previously undetectable fragmentation effects in an endangered amphibian. *Molecular Ecology* **27**, 4430–4443.

McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences* **104**, 19885–19890.

Meier JI, Salazar PA, Kučka M, Davies RW, Dréau A, *et al.* (2021) Haplotype tagging reveals parallel formation of hybrid races in two butterfly species. *Proceedings of the National Academy of Sciences* **118**, e2015005118.

Meirmans PG (2012) The trouble with isolation by distance. *Molecular Ecology* **21**, 2839–2846.

Meirmans PG (2014) Nonconvergence in Bayesian estimation of migration rates. *Molecular Ecology Resources* **14**, 726–733.

Morin PA, Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution* **19**, 208–216.

Moore J-S, Bourret V, Dionne M, Bradbury I, O'Reilly P, *et al.* (2014) Conservation genomics of anadromous Atlantic salmon across its North American range: outlier loci identify

the same patterns of population structure as neutral loci. *Molecular Ecology* **23**, 5680–5697.

Murphy MA, Dezzani R, Pilliod DS, Storfer A (2010) Landscape genetics of high mountain frog metapopulations. *Molecular Ecology* **19**, 3634–3649.

Murray KD, Janes JK, Jones A, Bothwell HM, Andrew RL, *et al.* (2019) Landscape drivers of genomic diversity and divergence in woodland Eucalyptus. *Molecular Ecology* **28**, 5232–5247.

Nosil P, Vines TH, Funk DJ (2005) Reproductive isolation caused by natural selection against immigration from divergent habitats. *Evolution* **59**, 705–719.

Osmond MM, Coop G (2021) Estimating dispersal rates and locating genetic ancestors with genome-wide genealogies. *BioRxiv* 2021.07.13.452277; doi: 10.1101/2021.07.13.452277

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.

Palamara PF, Lencz T, Darvasi A, Pe'er I (2012) Length distributions of identity by descent reveal fine-scale demographic history. *The American Journal of Human Genetics* **91**, 809–822.

Parsley MB, Torres ML, Banerjee SM, Tobias ZJC, Goldberg CS, *et al.* (2020) Multiple lines of genetic inquiry reveal effects of local and landscape factors on an amphibian metapopulation. *Landscape Ecology* **35**, 319–335.

Peterman WE (2018) ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods in Ecology and Evolution* **9**, 1638–1647.

Peterman WE, Winiarski KJ, Moore CE, Carvalho CdaS, Gilbert AL, *et al.* (2019) A comparison of popular approaches to optimize landscape resistance surfaces. *Landscape Ecology* **34**, 2197–2208.

Pool JE, Nielsen R (2009) Inference of historical changes in migration rate from the lengths of migrant tracts. *Genetics* **181**, 711–719.

Pouyet F, Aeschbacher S, Thiéry A, Excoffier L (2018) Background selection and biased gene conversion affect more than 95% of the human genome and bias demographic inferences. *eLife* **7**, e36317.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Prunier JG, Kaufmann B, Fenet S, Picard D, Pompanon F, *et al.* (2013) Optimizing the trade-off between spatial and genetic sampling efforts in patchy populations: Towards a better assessment of functional connectivity using an individual-based sampling scheme. *Molecular Ecology* **22**, 5516–5530.

Prunier JG, Dubut V, Chikhi L, Blanchet S (2017) Contribution of spatial heterogeneity in effective population sizes to the variance in pairwise measures of genetic differentiation. *Methods in Ecology and Evolution* **8**, 1866–1877.

Prunier JG, Saint-Pé K, Blanchet S, Loot G, Rey O (2021) Molecular approaches reveal weak sibship aggregation and a high dispersal propensity in a non-native fish parasite. *Ecology and Evolution* **11**, 6080–6090.

Puebla O, Bermingham E, Guichard F (2009) Estimating dispersal from genetic isolation by distance in a coral reef fish (*Hypoplectrus puella*). *Ecology* **90**, 3087–3098.

Puechmaille SJ (2016) The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources* **16**, 608–627.

Ringbauer H, Coop G, Barton NH (2017) Inferring recent demography from isolation by distance of long shared sequence blocks. *Genetics* **205**, 1335–1351.

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.

Ronce O, Clobert J (2012) In *Dispersal Syndromes. Dispersal Ecology and Evolution*. (Eds J Clobert, M Baguette, TG Benton, JM Bullock) pp. 119–138. Oxford University Press.

Rougemont Q, Carrier A, Le Luyer J, Ferchaud AL, Farrell JM, *et al.* (2019) Combining population genomics and forward simulations to investigate stocking impacts: A case study of Muskellunge (*Esox masquinongy*) from the St. Lawrence River basin. *Evolutionary Applications* **12**, 902–922.

Rougemont Q, Perrier C, Besnard AL, Lebel I, Abdallah Y, *et al.* (2022) Population genetics reveals divergent lineages and ongoing hybridization in a declining migratory fish species complex. *Heredity* **129**, 137–151.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219–1228.

Rousset F, Billiard S (2000) A theoretical basis for measures of kin selection in subdivided populations: Finite populations and localized dispersal. *Journal of Evolutionary Biology* **13**, 814–825.

Schwartz MK, McKelvey KS, Cushman SA, Luikart G (2010) Landscape genomics: a brief perspective. In *Spatial Complexity, Informatics, and Wildlife Conservation*. pp. 165–174. Springer, Tokyo.

Selkoe KA, Watson JR, White C, Horin TB, Iacchei M, *et al.* (2010) Taking the chaos out of genetic patchiness: Seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology* **19**, 3708–3726.

Sexton JP, Hangartner SB, Hoffmann AA (2014) Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution* **68**, 1–15.

Slatkin M (1981) Estimating levels of gene flow in natural populations. *Genetics* **99**, 323–335.

Slatkin M (2017) 1. Gene Flow and Population Structure. In: *Ecological Genetics*. (Ed L Real) pp. 1–17. Princeton University Press, Princeton, NJ, USA.

Smouse PE, Long JC, Sokal RR (1986) Multiple-regression and correlation extensions of the mantel test of matrix correspondence. *Systematic Zoology* **35**, 627–632.

Snyder-Mackler N, Majoros WH, Yuan ML, Shaver AO, Gordon JB, *et al.* (2016) Efficient genome-wide sequencing and low-coverage pedigree analysis from noninvasively collected samples. *Genetics* **203**, 699–714.

Söderquist P, Elmberg J, Gunnarsson G, Thulin CG, Champagnon J, *et al.* (2017) Admixture between released and wild game birds: a changing genetic landscape in European mallards (*Anas platyrhynchos*). *European Journal of Wildlife Research* **63**(6), 1–13.

Sokal RR, Oden NL (1978) Spatial autocorrelation in Biology. 1. Methodology. *Biological Journal of the Linnean Society* **10**, 199–228.

Spear SF, Balkenhol N, Fortin M-J, Mcrae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: Considerations for parameterization and analysis. *Molecular Ecology* **19**, 3576–3591.

Städele V, Vigilant L (2016) Strategies for determining kinship in wild populations using genetic data. *Ecology and Evolution* **6**, 6107–6120.

Storfer A, Patton A, Fraik AK (2018) Navigating the interface between landscape genetics and landscape genomics. *Frontiers in Genetics* **9**, 68.

Sunde J, Yıldırım Y, Tibblin P, Forsman A (2020) Comparing the performance of microsatellites and RADseq in population genetic studies: Analysis of data for pike (*Esox lucius*) and a synthesis of previous studies. *Frontiers in Genetics* **11**, 218.

Taylor PD, Fahrig L, Henein K, Merriam G (1993) *Connectivity Is a Vital Element of Landscape Structure*. Oikos, 571–573.

Telfer S, Piertney SB, Dallas JF, Stewart WA, Marshall F, *et al.* (2003) Parentage assignment detects frequent and large-scale dispersal in water voles. *Molecular Ecology* **12**, 1939–1949.

Thomas CD, Kunin WE (1999) The spatial structure of populations. *Journal of Animal Ecology* **68**, 647–657.

Tigano A, Friesen VL (2016) Genomics of local adaptation with gene flow. *Molecular Ecology* **25**, 2144–2164.

Trense D, Schmidt TL, Yang Q, Chung J, Hoffmann AA, *et al.* (2021) Anthropogenic and natural barriers affect genetic connectivity in an Alpine butterfly. *Molecular Ecology* **30**, 114–130.

Unglaub B, Cayuela H, Schmidt BR, Preißler K, Glos J, *et al.* (2021) Context-dependent dispersal determines relatedness and genetic structure in a patchy amphibian population. *Molecular Ecology* **30**, 5009–5028.

van der Zee JP, Christianen MJ, Bérubé M, Nava M, Schut K, *et al.* (2021) The population genomic structure of green turtles (*Chelonia mydas*) suggests a warm-water corridor for tropical marine fauna between the Atlantic and Indian oceans during the last interglacial. *Heredity* **127**, 510–521.

Van Strien MJ, Keller D, Holderegger R (2012) A new analytical approach to landscape genetic modelling: Least-cost transect analysis and linear mixed models. *Molecular Ecology* **21**, 4010–4023.

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* **13**, 921–935.

Vu NT, Zenger KR, Silva CN, Guppy JL, Jerry DR (2021) Population structure, genetic connectivity, and signatures of local adaptation of the Giant Black tiger shrimp (*Penaeus monodon*) throughout the Indo-Pacific region. *Genome Biology and Evolution* **13**, evab214.

Wagner HH, Fortin M-J (2013) A conceptual framework for the spatial analysis of landscape genetic data. *Conservation Genetics* **14**, 253–261.

Wakeley J (1996) Distinguishing migration from isolation using the variance of pairwise differences. *Theoretical Population Biology* **49**, 369–386.

Wang IJ, Glor RE, Losos JB (2013) Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecology Letters* **16**, 175–182.

Wang IJ, Bradburd GS (2014) Isolation by environment. *Molecular Ecology* **23**, 5649–5662.

Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity* **89**, 438–450.

Waples RS, 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**, 1419–1439.

Wei S, Li Z, Momigliano P, Fu C, Wu H, *et al.* (2020) The roles of climate, geography and natural selection as drivers of genetic and phenotypic differentiation in a widespread amphibian *Hyla annectans* (Anura: Hylidae). *Molecular Ecology* **29**, 3667–3683.

Wen G, Jin L, Wu Y, Wang X, Fu J, *et al.* (2021) Low diversity, little genetic structure but no inbreeding in a high-density island endemic pit-viper *Gloydius shedaoensis. Current Zoology* **68**, 526–534.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: FST not equal to 1/(4Nm + 1). *Heredity* **82**, 117–125.

Whitlock MC, Ingvarsson PK, Hatfield T (2000) Local drift load and the heterosis of interconnected populations. *Heredity* **84**, 452–457.

Willing E-M, Dreyer C, van Oosterhout C (2012) Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. *PLoS ONE* **7**, e42649.

Woltmann S, Sherry TW, Kreiser BR (2012) A genetic approach to estimating natal dispersal distances and self-recruitment in resident rainforest birds. *Journal of Avian Biology* **43**, 33–42.

Wultsch C, Waits LP, Kelly MJ (2016) A comparative analysis of genetic diversity and structure in jaguars (*Panthera onca*), pumas (*Puma concolor*), and ocelots (*Leopardus pardalis*) in fragmented landscapes of a critical Mesoamerican linkage zone. *PLoS one* **11**, e0151043.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114.

Xuereb A, D'Aloia CC, Andrello M, Bernatchez L, Fortin MJ (2021) Incorporating putatively neutral and adaptive genomic data into marine conservation planning. *Conservation Biology* **35**, 909–920.

Yadav S, J Stow A, Dudaniec RY (2021) Microgeographical adaptation corresponds to elevational distributions of congeneric montane grasshoppers. *Molecular Ecology* **30**, 481–498.