Received: 10 June 2022 | Revised: 25 October 2023 | Accepted: 30 October 2023

DOI: 10.1111/1755-0998.13893

INVITED TECHNICAL REVIEW

MOLECULAR ECOLOGY RESOURCES WILEY

Genomics and conservation: Guidance from training to analyses and applications

Lauren M. Schiebelhut[1] | Annie S. Guillaume[2] | Arianna Kuhn[3,4] | Rena M. Schweizer[5] | Ellie E. Armstrong[6] | Mark A. Beaumont[7] | Margaret Byrne[8] | Ted Cosart[9] | Brian K. Hand[10] | Leif Howard[9] | Steven M. Mussmann[11] | Shawn R. Narum[12] | Rita Rasteiro[13] | Angel G. Rivera-Colón[14] | Norah Saarman[15] | Arun Sethuraman[16] | Helen R. Taylor[17] | Gregg W. C. Thomas[18] | Maren Wellenreuther[19,20] | Gordon Luikart[5,9]

Correspondence
Lauren M. Schiebelhut, Life and Environmental Sciences, University of California, Merced, CA, USA.
Email: lschiebelhut@ucmerced.edu

Funding information
American Genetics Association; Dovetail Genomics; National Aeronautics and Space Administration, Grant/Award Number: 80NSSC19K0185; National Science Foundation, Grant/Award Number: DEB-1655809 and DoB-1639014; Springer Publishing; Wiley Publishing

Handling Editor: Andrew P. Kinziger

Abstract

Environmental change is intensifying the biodiversity crisis and threatening species across the tree of life. Conservation genomics can help inform conservation actions and slow biodiversity loss. However, more training, appropriate use of novel genomic methods and communication with managers are needed. Here, we review practical guidance to improve applied conservation genomics. We share insights aimed at ensuring effectiveness of conservation actions around three themes: (1) improving pedagogy and training in conservation genomics including for online global audiences, (2) conducting rigorous population genomic analyses properly considering theory, marker types and data interpretation and (3) facilitating communication and collaboration between managers and researchers. We aim to update students and professionals and expand their conservation toolkit with genomic principles and recent approaches for conserving and managing biodiversity. The biodiversity crisis is a global problem and, as such, requires international involvement, training, collaboration and frequent reviews of the literature and workshops as we do here.

KEYWORDS
bioinformatics, climate adaptation, conservation management, data analysis, genome assembly, molecular ecology, online courses, population genomics, sequence quality control

Annie S. Guillaume, Arianna Kuhn and Rena M. Schweizer are equal authorship contribution.

For Affiliation refer page on 13

1 | INTRODUCTION

Anthropogenic activities have increased the frequency and intensity of environmental change (AghaKouchak et al., 2020) and impose serious threats to biodiversity, with steep socio-economic costs (Bradbury et al., 2021). The past 5 years have been the hottest period on record (IPCC, 2021), with no signs of relenting. Often there is no single threat contributing to species' declines, but rather multiple stressors, including habitat loss, predation or competition from invasive species, disease and overharvesting (Lovejoy & Hannah, 2019). In many cases, swift conservation decisions, and actions that are well informed, must be taken to mitigate biodiversity loss. Population genomics can help inform these conservation actions in a number of ways. As the field expands and as genomic data become more easily generated, there is a growing need to train students, early-career scientists and practitioners in genomic data analysis and interpretation to help address diverse conservation issues (Schweizer et al., 2021; Taylor et al., 2017).

In this review, we discuss three major themes from the online ConGen 2021 course, usually held in Montana, USA that focused on training students and professionals to improve their use of genomics for conservation. First, we address the growing need for more extensive training to improve use of genomics principles and data (genotyping and sequencing) by scientists, managers and policy-makers worldwide; this includes recommendations to improve virtual (online) workshops. Second, we explain that rigorous science, rooted in theory, is necessary to generate actionable information (Allendorf, 2017). Improvements and rigour should be more thoroughly considered at all steps of genetics research and management workflows – from considering the appropriate molecular marker(s) and sample sizes for each research question to avoiding pitfalls associated with data analysis, interpretation and practical applications. Third, we discuss how population genomic knowledge and data can be used more effectively for conservation by building collaborations and communication between researchers and stakeholders (Taylor et al., 2017). Because the three themes of this review are distinct, readers might focus on those that interest them most.

1.1 | Improving training: advice for online workshops and teaching hands-on data analysis exercises

1.1.1 | Reaching global audiences via virtual training

The transition from in-person to virtual training in 2020, 2021 and 2022 allowed us to consider how to best engage people from different time zones and increase participation by students and instructors from more countries outside of the United States – especially countries and participants with limited funding to travel. Importantly, the online format fundamentally improved our ability to include and promote interactions among international and local audiences – including asynchronous learning by students from distant time zones who watched recorded lectures several hours after they occurred. This is an important difference between this online ConGen and a number of population genomic workshops and bioinformatic courses, which typically offer long-term or in-person courses (e.g. Stephen O'Brien's ConGen course; https://conservationgenetics.org/). For instance, a semester-long interdisciplinary, distributed graduate course between institutes in the United States and Europe combines local teaching with distant learning to teach landscape genetics (Wagner et al., 2012), while the annual 2-week Evomics Workshop on Population and Speciation Genomics held in Czech Republic brings up-to-date training in bioinformatics to students and early-career researchers across Europe (Barth et al., 2023).

We harnessed cloud-computing technology to maximize online engagement and learning. Providing hands-on exercises through a single virtual (cloud) server accessible to all instructors and students had multiple benefits. First, the single server reduced complications arising from each student preparing their own computers across different operating systems and downloading exercise data files. Second, it provided equitable access to computational power and speed throughout the course. Further, cloud computing through the RStudio Server (Rstudio Team, 2020) interface leverages a ready-made, offsite, flexible and temporary network and server infrastructure, with a versatile user connection to the server through a web browser (Langmead & Nellore, 2018).

Multiple software installation procedures, each typically dependent on other software and operating systems, can be daunting and sometimes fail completely on some students' computers. Time required for troubleshooting is obviated when using a single server with pre-installed software for hands-on exercises, available over the Internet. A virtual server hosted in the cloud also avoids burdening the host university's computing resources, including IT tasks associated with distributing temporary credentials to allow ConGen attendees access to a private university network.

By offering ephemeral, single-purpose, virtual machines (VMs), cloud computing services allow a single Linux image to move from one machine to another, with each VM configured for changing performance and storage needs. For example, the week before the first ConGen class, a modest VM was instantiated to allow instructors to test their hands-on software and procedures. When classes started, the (now tested and refined) Linux image was moved to a single 'production' VM with sufficient processors, RAM and storage to provide plentiful performance for combined use by instructors and students through remote login. By monitoring system resource use, we determined that our initial production VM was over-provisioned, and moved the image to a less costly machine, with only a very brief interruption in access. Finally, after terminating the ConGen final VM, we saved the Linux image in the cloud service to ease and speed-up server reconstruction for the following year.

To give organizers some idea of computer resources needed for a course similar to ConGen, with 12 of our 21 instructors offering computer-based exercises, and 41 (online) students, we met computer demands using a server with 72 virtual CPUs, 196 Gigabytes of RAM, 2 Terabytes of storage space for instructors and students

and network performance rated at 100 Gigabits. Varying numbers of students and instructors, as well as computational and storage demands of instructor-designed exercises and associated scheduling, likely render these specifications either too generous or insufficient for a given course. For VM management, it is important to test resource use before a course/workshop and to monitor it during server use.

The RStudio Server paired conveniently with cloud computing for hands-on exercises. It presents an interface accessible by web browsers that is secured using the underlying Linux user password system and properly limits user access with a Linux directory and file permission configuration. RStudio's client/server design efficiently accommodates many users with a single R installation (R Core Team, 2019), whose base system and extensible packaging system were used by many instructors. Along with an R terminal and plotting window, the RStudio interface also provides file transfers between local computers and the VM, a full-featured scripting editor, a graphical file browser and a Bash shell terminal that gives access to non-R executables. These resources encapsulate nearly the complete set of computing interactions students and instructors needed to perform the hands-on exercises for ConGen.

Only three programs required an operating-system-based graphical interface – EasyPop (Balloux, 2001), NeEstimator (Do et al., 2014) and Tracer (Rambaut et al., 2018). The solution, though inelegant, was to have students install these few programs on their laptops. We tried previously to provide these three programs on the VM using a virtual desktop server but found its performance too unpredictable to be useful. For future course iterations, we recommend making all software available virtually and will achieve this by modifying programs (e.g. EasyPop) to run with an R terminal (William Hemstrom & Jones, 2023).

In addition to streamlining the course with a virtual server, we used three software programs to facilitate communication and the transition to a remote offering of ConGen: Zoom (https://zoom.us) for video conferencing, Box (www.box.com) for document sharing and Slack (https://slack.com) for messaging. These programs allowed instructors and participants to share course content, participate in lectures and hands-on sessions and maintain communication, either live or asynchronously for those attending in distant time zones. Some instructors who did not want to learn or use Slack instead shared their emails with students to communicate directly.

Zoom recordings were converted to an MP3 format immediately following each lecture for uploading to a shared Box drive. This allowed asynchronous participants across different time zones to watch lectures a few hours later, while also archiving videos for future access by all participants (e.g. re-watching or teaching). Zoom breakout rooms for hands-on sessions created opportunities for students to problem solve in smaller groups with peers and instructors; students self-reported these activities as being useful to their learning. Based on participant feedback, we encourage future workshops to incorporate planned, unrecorded virtual 'coffee breaks' and casual question sessions with instructors. We recommend using a scheduler application that automatically adjusts for time zones when arranging one-on-one meetings with instructors to avoid confusion and scheduling conflicts.

The online format provides an opportunity to involve researchers globally who may not have joined in person due to geographic, financial, time or environmental constraints (López-Uribe et al., 2022). We saw an average increase from 8 international students at the in-person workshops (12–29% attendees in 2016–2019) to 20 at the online workshops (46–56% attendees in 2020–2021). Additionally, the online format allows for a diverse set of international instructors who bring in new skills and expertise. This increased diversity of participants and expert instructors worldwide shows that the online workshop format is cost-effective to provide population genomics training in countries where biodiversity is most threatened but least funded and where the lessons from a conservation course could have high impact.

1.1.2 | Designing effective hands-on learning activities: Examples testing for Hardy–Weinberg proportions

With the rapid increase in genomic resources, conservation geneticists are faced with a need to not only understand the theory behind population genomics tests, but also to efficiently analyse big datasets and interpret the results in applied conservation settings. Considering this, our ConGen course focused specifically on integrating the principles or assumptions of population genomics tests along with hands-on applications of tests and interpretations of results in practical scenarios (e.g. Table S1).

Among the most fundamental and widely used yet inconsistently applied tests in population genetics are tests for Hardy–Weinberg (HW) proportions (Sethuraman et al., 2019; Waples, 2014). Anyone publishing population genetic data or reading or using the population genetics literature must be able to conduct or understand these relatively basic tests (e.g. Allendorf, 2017; Allendorf et al., 2022). With the large number of loci in high-throughput sequencing (HTS) datasets, the number of HW tests per dataset has increased, which adds additional challenges of correcting for multiple testing (Graffelman & Weir, 2018; Sethuraman et al., 2019; Section 1.2.9 below). To improve student understanding of HW test assumptions and applications, the ConGen 2021 course had students test for deviations from HW proportions that were caused by different phenomena (e.g. cryptic population structure, small N_e or locus-specific selection).

In a somewhat novel teaching approach, student groups were given simulated datasets and asked to test for deviations from HW proportions, without information on the cause of the deviation (see the hands-on exercise in Table S1 and supplementary material). Student groups were asked to interpret test results and determine if their dataset showed signatures of a Wahlund effect (cryptic substructure) versus effects of small N_e, or selection on individual loci. Students learned to calculate and plot basic statistics, including F_{IS}, using the hierfstat R package; they also performed PCA using the

ADEGENET R package (Jombart, 2008) in R to distinguish population structure from other causes of HW proportion deviations. By using simulated datasets, students can analyse datasets as if they were empirical studies, having a chance to interpret results and discuss putative drivers of the recovered signal before receiving the correct answer from instructors (Figure 1).

We recommend giving students time and space (e.g. through Zoom breakout rooms) to work together and discuss the problem, with guidance from instructors (visiting the breakout rooms), focusing on population parameters (in our examples, negative F_{IS} at one or most loci), and assess genome-wide and locus-specific patterns. For example, differences in allele frequencies between the sexes caused by drift and small N_e will cause a genome-wide negative F_{IS} (i.e. an excess of heterozygotes). General discussion can include considering how to test for an excess of heterozygotes (negative F_{IS}) and allele frequency differences between sexes, which can be done separately or simultaneously/jointly (e.g. Graffelman & Weir, 2018), and how to estimate N_e to follow-up this hypothesis that N_e is small (e.g. Do et al., 2014; Waples & Do, 2010). This hands-on teaching strategy using simulated data can be applied to teaching other statistical tests, metrics and concepts (Schweizer et al., 2021).

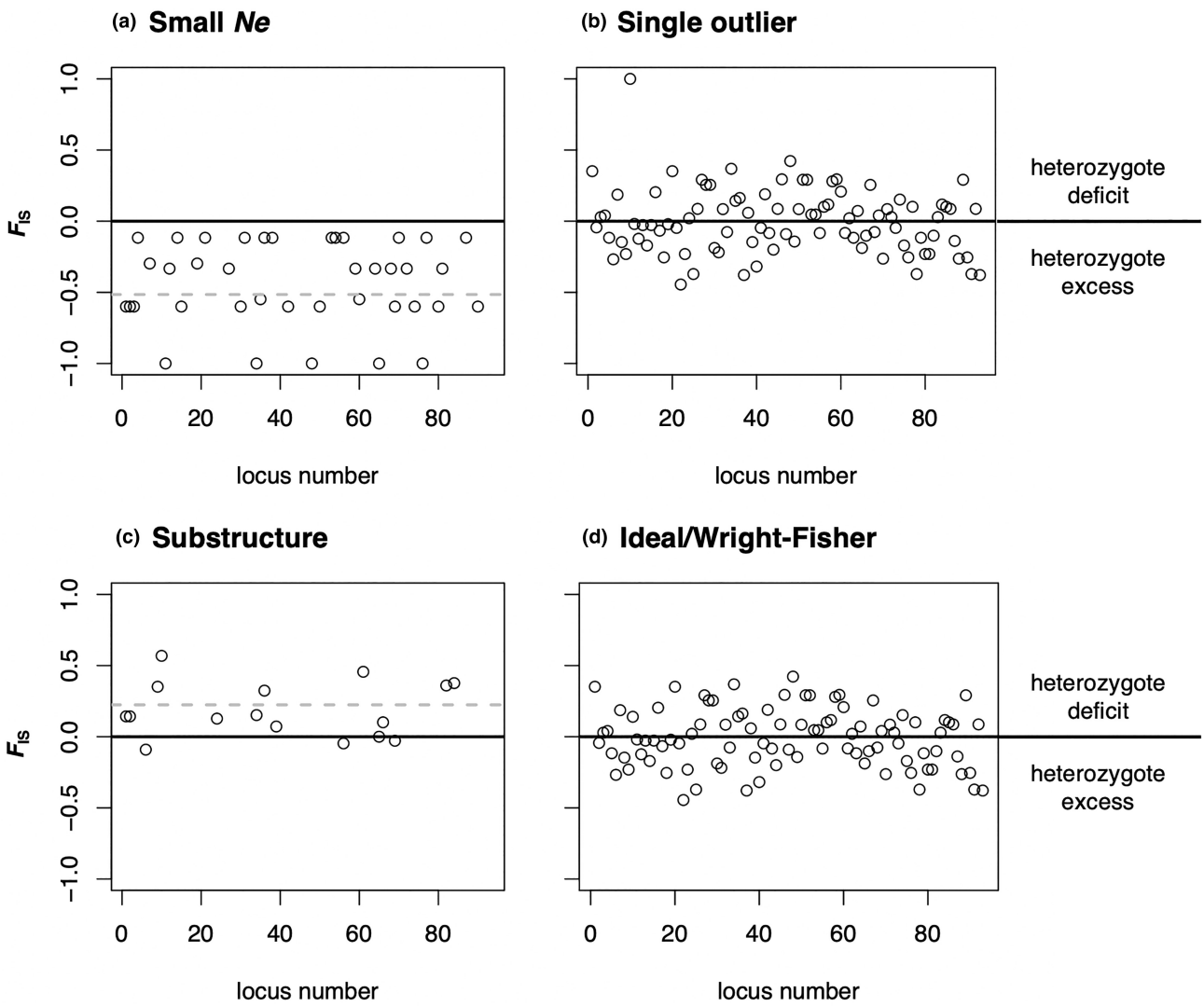

FIGURE 1 Representation of data simulated to represent four realistic biological scenarios (a–d; grey dashed lines represent mean F_{IS}) for the HW proportions hands-on learning activity from ConGen 2021. Each group of students was blindly assigned one of four simulated datasets: (a) a population with small N_e, (b) a dataset from a genome-wide scan with one outlier locus experiencing very strong selection against heterozygotes, (c) substructure, illustrating the Wahlund effect or (d) an ideal/Wright–Fisher population (i.e. random mating with large N_e and no substructure). Simulations were completed in EASYPOP (Balloux, 2001) under a Wright–Fisher model for a diploid organism with separate sexes and random mating (except b and c had selection and substructure added), details available in online materials (Section 1.1.2 in Table S1). All loci that went to fixation during the simulations (e.g. panels a and c) are not shown because F_{IS} cannot be estimated when there is zero heterozygosity. The students calculated F_{IS} and other summary statistics to connect genetic patterns to real-life processes and identify which dataset they received.

1.2 | Novel considerations at each step of the data analysis pipeline

This section discusses novel developments or timely issues associated with eight important topics covered at ConGen 2021 and 2022. We first present a novel decision tree to help choose the best molecular marker(s) for a given research question or management problem (2.1), which is particularly important when working with rare and endangered species and a limited budget or sample sizes; and (2.2) the importance of identifying structural variants in addition to SNPs because they are abundant across genomes and often influence phenotypes and fitness. We then discuss how to appropriately process molecular genetic data, including filtering and analysing RADseq data from raw reads to genotypes using the recently revised version of STACKS – v2 (2.3). We next discuss key factors to consider when assembling reference genomes, such as genome size, ploidy, heterozygosity, sample quality and quantity (2.4) and alignment (2.5). We then discuss detection of selection signatures including flexible approaches not requiring population identification, as implemented in PCADAPT (2.6). We then describe novel approaches for estimating dispersal rates using BAYESASS (2.7). Finally, we illustrate the importance of correcting for multiple testing in large genomic datasets (2.8).

1.2.1 | Choosing the right marker for your study

Recent advances in sequencing technology, targeted sequencing and associated cost reductions have increased the use of reduced representation sequencing and low-coverage whole genome sequencing approaches (Ali et al., 2016; Baird et al., 2008; Campbell et al., 2015; Lou et al., 2021). These advances provide researchers with improved tools to measure and conserve genome-wide variation (e.g. Kardos et al., 2021). Hundreds to thousands of SNPs and HTS provide increased power to address conservation questions and are less expensive than genotyping microsatellites (e.g. Allendorf et al., 2022).

This variety of available sequencing (and genotyping) approaches can lead to uncertainty and confusion when selecting the most appropriate marker type or sampling method. Indeed, while there is broad understanding that specific study questions and the biology of target organisms must guide the selection of molecular methods, many nuanced factors involved in these decisions remain. It is important for students to understand how to approach these decisions. Unfortunately, no publications provide a comprehensive review of or advice for choosing among marker types for different questions (e.g. Hohenlohe et al., 2019; Meek & Larson, 2019).

Population genetics studies often address multiple questions relating to variation at putatively neutral and adaptive regions of the genome (Willi et al., 2022). Such studies have traditionally addressed questions relating to population structure, diversity, connectivity, pedigrees, relatedness and phylogenetic relationships with a modest set of putatively neutral markers (Hohenlohe et al., 2019; Schweizer et al., 2021). However, complex patterns of relatedness or deep inference of pedigree relationships may require a relatively large number of neutral markers (e.g. Delomas & Campbell, 2022; Galla et al., 2022). Research questions that aim to address selective processes within or between species, populations, ecotypes or phenotypes typically require markers from adaptive regions of the genome.

While there remain cases where microsatellites are fully adequate, there are many scenarios where larger panels of genome-wide markers are less costly (Campbell et al., 2015) or allow genome-wide investigation of both neutral and adaptive variation (e.g. Lou et al., 2021; Meek & Larson, 2019). In cases of reduced representation or whole genome sequencing, there are circumstances when individual genotypes must be determined with high sequencing depth and quality (e.g. when testing for HW proportions, LD, parentage, forensics cases, GWAS or when individual fitness or phenotypes are compared to individual loci) in contrast to low-depth approaches to estimate allele frequencies within a group or population (Hendricks et al., 2018; Lou et al., 2021; Prince et al., 2017).

Based on these concepts, the flow chart in Figure 2 will help researchers choose appropriate marker methods for their study question. The chart begins with a question regarding the need of neutral versus adaptive markers since many studies in conservation genetics may only require neutral markers (e.g. population structure, estimates of *Ne*, parentage analyses) as opposed to those that require more extensive genomic coverage to identify adaptive variation (e.g. local adaptation; genomic basis of trait variation). This chart is not comprehensive of all scenarios, and largely targets sequencing or GBS-based approaches for conservation genomics.

Other tools are recommended for specialized applications (reviewed in Hohenlohe et al., 2019; Schweizer et al., 2021), such as epigenetics and gene expression (Anastasiadi et al., 2021), genome assembly (Whibley et al., 2021) or environmental DNA (Rodríguez-Ezpeleta et al., 2021). Furthermore, decisions regarding the discovery and genotyping of markers throughout the genome will typically involve factors of both scientific importance (e.g. sample size, genome size, statistical power, evolutionary context) and practical nature (e.g. sample quality, funding, time, resources, career goals). Ultimately, there are many perspectives regarding preferred methods, such that guidance should come from a variety of resources and people with expertise in marker selection for conservation genomics applications (Benestan et al., 2016; Schweizer et al., 2021).

1.2.2 | Conducting quality control in RADseq analysis

Among the most important breakthroughs in the 2010s for conservation genetics was the development of restriction site-associated DNA sequencing (RADseq; Anonymous, 2010). This sequencing approach has been reviewed elsewhere but is mentioned here because RADseq is increasingly applied to questions and systems – particularly those lacking a reference genome or genomic resources (Bootsma et al., 2021; Garrison et al., 2021; Nash et al., 2018; Stuart et al., 2014).

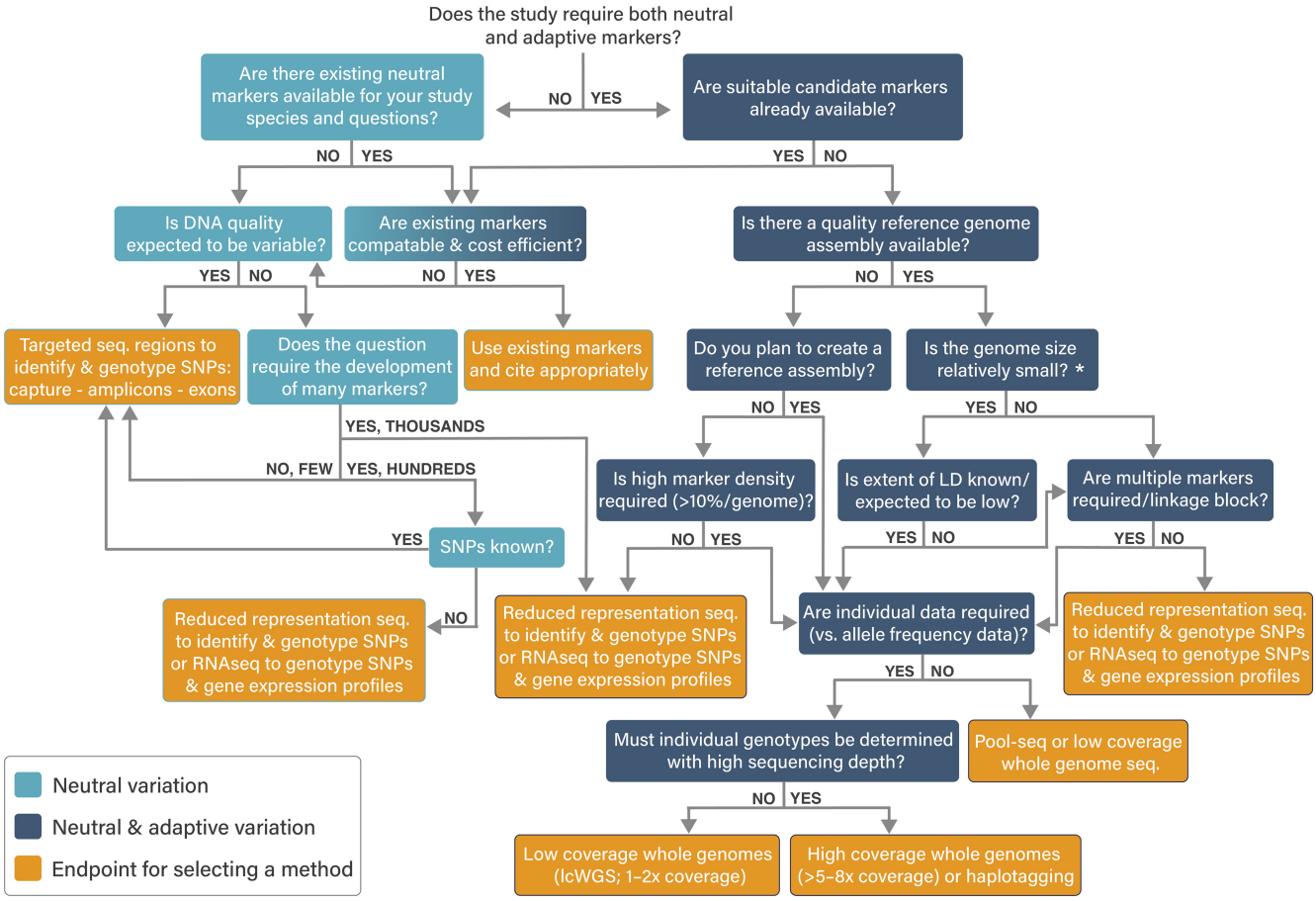

FIGURE 2 A decision tree for guiding the choice of a molecular marker method for conservation and ecological genomics applications. The chart begins with a question regarding the need of neutral versus adaptive markers since many studies in conservation genetics may only require neutral markers (e.g. population structure, estimates of N_e, gene flow, parentage analyses) as opposed to those that require more extensive genomic coverage to identify adaptive variation (e.g. local adaptation; genomic basis of trait variation; Allendorf et al., 2010). However, this flow chart is not comprehensive and decisions should be guided by specific project needs, for example, *a large genome size (e.g. > a few Gb) can limit genome-wide marker approaches because they require large computational resources or sequencing costs for a specific project.

Many countries, laboratories and non-model taxa are only now transitioning to use RADseq or still use microsatellites for conservation work; whole genome sequencing is often too expensive or unnecessary (Clugston et al., 2019; Gargiulo et al., 2021; fig. 4.1 in Allendorf et al., 2022). Efficient implementation of RADseq approaches often requires understanding of issues such as linkage disequilibrium, marker density, sequencing coverage and PCR duplicates to ensure sufficient power to detect the biological signal of interest (Davey et al., 2013; reviewed in Andrews et al., 2016; see also McKinney et al., 2017). Factors such as low sequencing coverage and an excess of PCR duplicate reads can reduce power and accuracy of genotyping by missing loci or alleles (allele dropout caused by heterozygotes scored as homozygotes; Hendricks et al., 2018). Technical sources of allele dropout in RADseq experiments (e.g. PCR, sequencing or bioinformatics problems) far eclipse the effects of dropout caused by biological polymorphism and so are critical for students to consider in analysis (Rivera-Colón et al., 2021).

There are several analytical pipelines available for the analysis of RADseq data, including Stacks 2 (Rochette et al., 2019), dDocent (Puritz et al., 2014), iPyrad (Eaton & Overcast, 2020), among others. While the underlying algorithmic approaches might be different in each software, their results can all be negatively impacted by technical artefacts (e.g. low coverage and high duplicates). Therefore, monitoring the data at each step of the pipeline is crucial for any RADseq analysis, regardless of the pipeline used. Although problems with coverage and PCR duplicates originate from the construction and sequencing of the library itself (Rochette et al., 2023), steps can be taken post-sequencing to assess, and potentially mitigate, their effects. For example, monitoring the distribution of reads across samples when de-multiplexing, verifying coverage before and after duplicate removal and filtering the resulting genotypes (see supplementary Box S1 for advice on using Stacks v2 software; Rochette et al., 2019). These checks provide a robust and reliable implementation of the library protocol, enabling the experimenter to trace the data (and problem solve) at each major step of the analysis – from

raw read to final genotype. Different pipelines yield different results; thus, different pipelines (and filtering choices) can be used to improve certainty of variant identification, genotyping quality and downstream inferences (Graham et al., 2020; Hemstrom et al., n.d. in review; Mona et al., 2023; Shafer et al., 2017).

1.2.3 | Structural variants are important markers and components of genomic variation

Recent work has revealed that structural variants – variation in the presence, abundance, position and/or direction of a nucleotide sequence – are ubiquitous in nature, often influence fitness and can impact larger sections of the genome relative to the more commonly studied SNP variants (Catanach et al., 2019; Feuk et al., 2006). Structural variants can influence population dynamics, adaptation and eco-evolutionary dynamics of threatened species or invasive species of management concern (e.g. Cohen et al., 2023). Structural changes can also impact the three-dimensional structure of the genome, which may influence gene expression and regulation (Mérot et al., 2020). Students and conservationists should understand this form of variation.

Structural variants are best detected with high-coverage whole genome data (Box 1). The role of structural variants in conservation and complementing SNP-based approaches to enhance species recovery is increasingly realized (Wold et al., 2021). These variants often detect significant genetic–environment associations where SNPs do not (e.g. sea surface temperature and American lobsters, *Homarus americanus*; Dorant et al., 2020). Structural variants can be species diagnostic and detected by PCR tests to inform management decisions (e.g. distinguishing a rare subspecies of white-cheeked gibbon, *Nomascus leucogenys leucongenys*; Carbone et al., 2009). As structural variants can also reduce reproductive success, they are of interest to programmes conserving small populations (Deakin et al., 2019).

Structural variants can lead to genomic incompatibilities that influence hybridization and introgression, and variants like inversions can impact the genome-wide recombination landscape (Sturtevant, 1913; Sturtevant & Mather, 1938; reviewed in Wellenreuther & Bernatchez, 2018). Recent studies show structural variants often have a significant influence on the phenotype (Sanchis-Juan et al., 2018; reviewed in Wellenreuther & Bernatchez, 2018), as seen in morphotypes of rainbow trout (*Oncorhynchus mykiss*: anadromous and resident; Pearse, 2016) and the ruff bird (*Calidris pugnax*: satellites, faeders and independents; Lamichhaney et al., 2016).

Most work has focused on the adaptive significance of gene duplications and large-effect inversions, with much less attention on other structural variant types, e.g. fusions or translocations (Wellenreuther et al., 2019). Future studies should focus on reproducible and well-documented frameworks for structural variant detection to uncover the full spectrum of types and sizes, and to relate their presence to their function and evolutionary significance (e.g. as outlined in Mérot et al., 2020), which will also facilitate comparative meta-analyses. Finally, databases will become important tools for detecting candidate adaptive variants in related species, something that is so far only available for human disease (e.g. DBVar: https://www.ncbi.nlm.nih.gov/dbvar/).

BOX 1 Discovery and genotyping of structural variants.

Structural variants are best assessed with high-coverage whole genome data (e.g. NovaSeq), and long variants, specifically, should be targeted with long-read data [e.g. Pacific Bioscience (PacBio)] that can span variant boundaries (Mahmoud et al., 2019). Other types of sequencing data can also be used, e.g. paired end data from reduced representation libraries or data derived from standard shotgun libraries (i.e. with short insert size, generally <1 kb), but their suitability in detecting a diverse set and size range of structural variants is more limited.

Detection of structural variants occurs via the application of both indirect and direct methods. Indirect approaches apply unsupervised methods such as principal component analysis (PCA) combined with information about linkage disequilibrium (LD) to detect genome regions that co-segregate and show limited recombination, i.e. haploblocks (e.g. Mérot et al., 2021). Direct methods use information based on the read depth or from split reads and read orientation to identify variants (e.g. duplications or deletions). Advancements, such as optical mapping, Hi-C data and linked reads using 10x Genomics, all provide information to directly detect structural variants de novo (Mérot et al., 2020). Software for the analysis of structural variants is rapidly developing and continuously updated and should be selected based on the type of available data and structural variants of interest [e.g. LUMPY (Layer et al., 2014), MANTA (Chen et al., 2016), DELLY (Rausch et al., 2012)].

1.2.4 | Recent advances in genome sequencing and assembly

Several technologies, such as Pacific Biosciences and Oxford Nanopore, have greatly increased their capabilities of sequencing long stretches of DNA (long reads), which has allowed for a greater ease of assembly. Longest reads up to 15–30 kb can now be achieved fairly reliably (Marx, 2023), and with a reduction in error rates (e.g. 99.8% accuracy using circular consensus long-read sequencing on the PacBio HiFi Sequel System; Wenger et al., 2019). In addition to long reads improving our ability to construct high-quality references, other technologies such as Hi-C or optical mapping can be incorporated to provide additional information on how various sequences fit together within chromosomes (Rhie et al., 2021; Whibley

et al., 2021). This, and the accompanying declining costs, make de novo assembly a more attractive and viable option for non-model organisms (for review, see Whibley et al., 2021).

Although it can be time-consuming and computationally intensive, building a reference genome for a species of interest is preferable to the alternate approach of mapping reads against the reference genome of a closely related species. This is because the differences in the sequence order and content between species can ultimately affect downstream analyses, a consequence broadly termed 'reference bias' (Günther & Nettelblad, 2019; Prasad et al., 2022). In fact, reference bias can even occur within species (Thorburn et al., 2023). To accommodate within-species reference bias, reference genomes from several individuals within a species can also be used to construct pangenomes, which further improves our ability to represent diversity within a species (e.g. The Human Pangenome Project; Wang et al., 2022).

In addition to reducing bias, high-quality, contiguous assemblies allow us to better infer variation in the context of recombination since they enable the use of haplotype-level information, thus facilitating our understanding of demography and gene flow (Leitwein et al., 2020). High-quality assemblies also enable more accurate detection of runs of homozygosity (ROH), which are important for the inference of recent or historic inbreeding events. For example, Prasad et al. showed that when mapping to non-conspecific assemblies, ROH was undetectable in both the rowi kiwi and the beluga whale, and more fragmented assemblies impacted the detection of ROH in unpredictable ways (Prasad et al., 2022).

Once annotated, comparisons of genome assemblies from different individuals, populations or species can capture the full range of genome variation, including SNPs, repetitive elements, copy-number variants and structural variation (Dominguez Del Angel et al., 2018). For example, annotation of both repetitive elements and genic regions in a threatened and endemic land snail, *Oreohelix idahoensis*, revealed that expansions of long terminal repeats are responsible for differences in both genome size and gene composition, specifically expanding gene families related to stress and biomineralization compared to other *Oreohelix* species (Linscott et al., 2022). Structural variation that would go undetected in more fragmented assemblies can also provide critical information about the maintenance of polymorphisms within species. For example, a new assembly for rainbow trout (*On. mykiss*) revealed a 55 Mb supergene that is responsible for maintaining sex-specific migratory tendencies (Pearse et al., 2019).

While many initiatives are producing high-quality reference genomes for species of conservation concern – e.g. the Vertebrate Genomes Project (Rhie et al., 2021), the Earth BioGenome Project (Lawniczak et al., 2022), the Darwin Tree of Life Project (The Darwin Tree of Life Project Consortium et al., 2022), the California Conservation Genomics Project (Shaffer et al., 2022) – it will be of critical importance for conservationists to contextualize these data to support our understanding of the diversity between and within species. Building reference genomes represents an important and sometimes overlooked first step in the conservation genomics pipeline, but greatly improves the accuracy of downstream analyses and subsequent inferences.

1.2.5 | An alternate to genome assembly: Reference-based read mapping

As genome assembly is both a time-consuming and computationally intensive endeavour, an often-used alternative strategy involves aligning sequence reads to a previously assembled genome of the same, or closely related, species. This approach, commonly called 'read mapping', is much faster than de novo assembly, but has the drawback that structural variation between the sequenced sample and the reference genome (being mapped to) will be difficult to discern (Theissinger et al., 2023). Yet, for many biological questions, read mapping offers a quick and sufficient alternative to de novo assembly. For instance, this method was recently employed to identify genes responsible for seasonal coat colour changes in white-tailed jackrabbits (Ferreira et al., 2023). Several pipelines have been developed to facilitate mapping-based population genetics of non-model organisms with many individuals at high throughput (Czech & Exposito-Alonso, 2022; Mirchandani et al., 2023). In some cases, it may also be a more efficient strategy to sequence many samples to low depth, pool reads from samples by population, and map the reads to a closely related, high-quality reference genome (Lou et al., 2021).

Read mapping has some drawbacks, including reference bias, which occurs when highly diverged reads fail to map to the reference genome (Armstrong et al., 2020; Sarver et al., 2017). When assessing the quality of mapped reads and variants called from them, many of the programs used to map reads [e.g. BWA (Li & Durbin, 2009), Bowtie2 (Langmead & Salzberg, 2012), minimap2 (Li, 2018)] also provide a convenient mapping score that can be used for filtering in downstream analyses. Beyond that, the most informative quality metric of mapped reads may be coverage or read depth. One can compute expected coverage and compare this to actual coverage output from the read mapper to identify potentially unsuccessful read mapping.

1.2.6 | Quality control of genome assemblies

Whether one is performing de novo assembly or read mapping, as with all bioinformatic tasks, quality control at every step is important. First, tools that assess the quality of raw reads like FastQC (Andrews, 2010) and MultiQC (Ewels et al., 2016) can help determine if the trimming or filtering of reads is necessary. For de novo assembly, the ideal outcome is to have one assembled contig or scaffold per chromosome for the sequenced species. However, due to the biological complexity of the genome, this is often not easily achieved. Nonetheless, non-chromosome-level assemblies are useful for improving applications of genomics to conservation (e.g. RADseq or related genotyping; Hendricks et al., 2018). Importantly, statistics have been developed to assess how well a genome has been assembled.

Tools are available to calculate basic assembly statistics such as the number, size and distributions of scaffold and contig lengths. These tools include but are not limited to BBTools (Bushnell, 2014), QUAST (Gurevich et al., 2013) or tools from the Assemblathon 2 (Bradnam et al., 2013). A commonly used assembly statistic is the

scaffold N50, which is computed as the length of the shortest scaffold of all scaffolds that sum to at least half of the genome assembly (Lander et al., 2001). In contrast, the scaffold L50 describes the number of scaffolds (when considered in order from largest to smallest) that comprise 50% of the genome.

While N50 and L50 are useful for estimating the contiguity of an assembly, these statistics are calculated independent of expected genome size (instead of using total *assembly* size as the expected sequence length), which makes comparisons of N50 statistics between genomes difficult. One solution is to use the NG50 statistic, which is similar to N50 but is computed using expected genome size rather than assembly size (e.g. NG50 of 100 means that half of the *genome* is assembled in scaffolds 100 bp or larger) and is now the preferred metric for comparing assemblies between individuals or species (Earl et al., 2011; Whibley et al., 2021) as we look forward to a near future where comparative genomics of non-model organisms in conservation settings becomes commonplace. Additional methods for evaluating assembly accuracy and polishing are provided in supplementary materials (Box S2).

1.2.7 | Detecting selection in the context of adaptive divergence

Detecting selection is important both to facilitate the study of adaptive variation and to remove adaptive loci prior to computing neutral population parameters (Luikart et al., 2003). Methods for detecting evidence of natural selection from gene frequency differences have been developed since the early 1970s (Lewontin & Krakauer, 1973), primarily to account for biases arising from population structure (Beaumont & Balding, 2004; Beaumont & Nichols, 1996; Coop et al., 2010; Excoffier et al., 2009; Fariello et al., 2013; Foll & Gaggiotti, 2008; Privé et al., 2020). Many approaches have been developed and implemented to detect selection using genome data (Lin et al., 2022; Nielsen, 2005; Rajawat et al., 2022; Smith et al., 2022). In the ConGen course, we discussed recent approaches that compare populations that are potentially experiencing divergent selection in order to illustrate potential confounding factors such as demographic history, with a focus on PCADAPT (Privé et al., 2020).

Many of the methods have been based on estimates of F_{ST}. Tight linkage and strong selection leads to higher F_{ST} for genetic markers linked to a selected variant. It is important to note that these methods identify linked regions of the genome and not necessarily the causal variant itself. Recently, Liu et al. (2019) studying horses at different elevations found high F_{ST} values at a number of markers, especially those linked to the EPAS1 gene, and the authors concluded that this provided evidence of adaptation to hypoxia at high altitude.

The classic literature on modelling hybrid zones (Barton & Bengtsson, 1986) makes use of the approach described earlier, which implicitly underpins genomic cline analysis, a statistical method for detecting natural selection in hybridizing populations (Gompert & Buerkle, 2011, 2012) and implemented in BGC (Gompert & Buerkle, 2012). Genomic cline analysis is more appropriate for situations in which the hybrid status of individuals can be identified (McFarlane et al., 2021), as it presupposes that pure parental types can be identified and gene flow follows a particular demographic model (Gompert & Buerkle, 2011). However, demographic scenarios are often not known, or can be complex, particularly in cases of extensive introgressive hybridization.

An alternative and flexible approach, which does not require pre-identification of parental populations, is to use the method of Duforet-Frebourg et al. (2016), implemented in PCADAPT (Privé et al., 2020), which is based on the distribution of PCA loadings and identifies genes that contribute an unusually large weight to the PCA score. A recent example of local adaptation using PCADAPT is from an analysis of spinner dolphin ecotypes which provides evidence of local adaptation in social behaviour (Andrews et al., 2021). This study compared several different approaches to identifying outliers, and also illustrated the use of enrichment analysis (also discussed below) for identifying networks of genes related to particular functions.

A further advance in methods for detecting the effects of natural selection in the genome has arisen from recognition that individual gene effects on a trait, and associated selection coefficients, may be very small (Boyle et al., 2017). This has led to methods focusing on groups of genes in sub-networks, for which statistical inference can be combined, leading to greater power (Berg & Coop, 2014; Daub et al., 2013). In the case of introgressing genomes, it is also possible to test whether groups of genes in the same sub-network are significantly enriched in recipient genomes as a consequence of introgression (Gouy & Excoffier, 2020).

Along with improvements in the detection of selection at the genomic level, there have also been advances in methods for simulating sequence evolution at a genomic level (Haller & Messer, 2019; Kelleher et al., 2016). These methods have a variety of uses: calibration of population genetic tools, such as those described earlier (Howard-McCombe et al., 2021; McFarlane et al., 2021), and the development of models that jointly infer demography and selection (Johri et al., 2020). These simulation tools are particularly valuable for calibrating methods for detecting selection (Howard-McCombe et al., 2021; McFarlane et al., 2021), because, though there have been major advances in their accuracy, it is not possible to develop an all-encompassing method that is accurate under all possible circumstances, and they tend to be sensitive to details of the demographic history of populations (Hoban et al., 2016).

Finally, methods to detect selection can also be used to remove putative adaptive loci from the analyses (Mamoozadeh et al., 2020; Reynes et al., 2021). This potentially allows researchers to focus on neutral drivers of evolution and to infer the demographic history of populations with less potential bias from loci under selection (Luikart et al., 2003).

1.2.8 | Contemporary migration estimation and a need for simulation studies

Assessment of contemporary migration (gene flow) aids in delineation of conservation units (Lowe & Allendorf, 2010) and is a common application that conservation geneticists encounter

(Burgess et al., 2022; Forester et al., 2022). Despite this conservation importance, methods of migration assessment often focus on coalescent-based approaches to determine historical migration rates and population dynamics (Al-Asadi et al., 2019; Beerli & Felsenstein, 2001; Hey & Nielsen, 2007; Nielsen & Wakeley, 2001). In comparison, algorithms for contemporary migration assessment have been less frequently published [BayesAss (Wilson & Rannala, 2003); divMigrate (Sundqvist et al., 2016)].

As software ages, it often requires modification to remain analytically relevant as datasets become larger. The modification of older, popular programs can be advantageous due to existing awareness of shortcomings or situations in which these methods are inaccurate (e.g. BA3-SNPs; Mussmann et al., 2019). However, the vetting process must continue when new genotyping methods are developed, with previous studies serving as a baseline for additional validation of analytical methods. For example, BayesAss is less suited for scenarios in which large numbers of populations are evaluated (e.g. >5) and where genetic differentiation is low among population pairs (e.g. $F_{ST} < 0.1$) because these scenarios can cause Markov chain Monte Carlo (MCMC) convergence problems or underestimation of true migration rates (Faubet et al., 2007; Meirmans, 2014).

A little known approach for evaluating large numbers of populations in networks (e.g. streams, riparian zones or mountain ridgelines) is to analyse subsets or populations. These environments can be deconstructed into discrete groups of neighbouring populations along the linear path of a stream between which organisms may plausibly move (Neuenschwander, 2006). This geographic sliding window approach, which assumes a stepping stone model of migration (Kimura & Weiss, 1964), has been proposed as one method for evaluating large numbers of discrete populations in a stream network (Mussmann, 2018).

Analytical methods must continue to evolve as new sequencing and SNP genotyping technologies increase in popularity. SNPs and new sequencing datasets undergo many filtering regimes including heterozygosity, minor allele frequency and missing data (Andrews & Luikart, 2014; O'Leary et al., 2018; Schweizer et al., 2021). Different SNP filters have been demonstrated to impact various analytical methods (Ahrens, Jordan, et al., 2021; Martin et al., 2021), but the explicit impacts of these filters have not yet been investigated for BayesAss. SNPs may represent either neutral or adaptive loci; the inclusion of adaptive loci in studies of migration estimation needs to be evaluated.

Researchers must also consider the benefits of genotyping more individuals compared to additional loci. Some population genomic statistics (e.g. heterozygosity; pairwise F_{ST}) can often be reliably determined from low sample size ($n \leq 10$ per population) when thousands of genomic loci are available (Nazareno et al., 2017; Schmidt et al., 2021). However, other analyses require larger samples per population ($n \geq 20$–30) to be accurate or useful – e.g. estimating N_e via linkage disequilibrium (Luikart et al., 2021; Nunziata & Weisrock, 2018). Testing for loci associated with fitness or a phenotype (GWAS) often require samples of hundreds to thousands of individuals to achieve high power for natural populations (Santure & Garant, 2018). The benefit from genotyping more loci can quickly plateau (Waples et al., 2022), and it remains unknown how the interactions between genome-scale samples of loci and individuals impact the ability to infer migratory rates. Therefore, precious computation time could, in some cases, be better spent including more individuals (rather than more loci) in analyses. As availability of genomic resources increases for many organisms, contemporary migration estimation could also benefit from development of algorithms focusing specifically on low-coverage genome sequencing (Lou et al., 2021).

The aforementioned knowledge gaps indicate areas of research in which simulation would be vital for evaluating SNP filter impacts on contemporary migration estimation (Hoban, 2014; Hoban et al., 2012; Yuan et al., 2012). Simulation studies can address realistic scenarios as conservation efforts come ever closer to the analytical and real-life situations in which small errors can cause management problems and population declines (Schweizer et al., 2021).

1.2.9 | Correcting for multiple testing

Recent meta-studies in conservation genomics have described the persistent underuse and inconsistent use of methods correcting for multiple testing, which is extremely problematic with datasets of thousands of loci (Hauser et al., 2019; Sethuraman et al., 2019) – an issue students need to be aware of in data analysis. Both studies reported conscious or unconscious bias by researchers applying conservative and/or liberal corrections within publications without explaining why [e.g. Benjamini–Yuktieli, Benjamini–Hochberg based false discovery rate (FDR) correction methods] to aid the discovery of significantly differentiated outlier genomic loci that are then correlated with functional adaptive trait variation (see Lotterhos & Whitlock, 2014). Hauser et al. (2019) and Sethuraman et al. (2019) also report that researchers are using more conservative corrections (e.g. Bonferroni, sequential Bonferroni correction methods) to identify genomic loci that are not in HW proportions, or in strong linkage disequilibrium so as to filter those loci for non-conformance to independence or neutrality (see Waples, 2014). While these are not inherently 'biases', most studies also lack sufficient explanation or analysis of FDRs chosen and possible effects on subsequent results and conclusions.

In this context, it is informative to consider the methods described in White et al. (2019) and Narum (2006) to estimate the family-wise error rate (FWER) under the complete null model for any statistical test (i.e. there are no correct discoveries, and therefore here equivalent to the FDR). Consider performing 1000 independent null chi-square tests with two degrees of freedom (similar to performing chi-square tests of HW proportions at 1000 bi-allelic SNP loci). FWER is estimated as 1-(1-alpha)^k, where *alpha* is the critical value threshold and *k* is the number of independent tests performed. Without adjusting for multiple

testing, the FWER quickly reaches 1.0, implying that the inference from these tests is always incorrect. However, all other correction methods – (in order of the degree of conservativeness) Bonferroni, Benjamini–Yekutieli (BY) and Benjamini–Hochberg (BH) – effectively control for FWER (and FDR), maintaining it the rate of false discovery across all *k* tests at a threshold of 0.2 and below (Figure 3). Importantly, however, controlling for FWER using the BY and BH methods also correspondingly inflates the false negative rate (see White et al. (2019) for a detailed treatment of this issue). Therefore, researchers must consider false negative and positive rates based on their question, and explain or justify their choice of correction method – including effects of sample sizes, number of tests and the desired degree of conservativeness (White et al., 2019).

Recent studies have detailed approaches for multiple testing correction (Austin et al., 2022; Flores-Manzanero et al., 2022; Marques et al., 2022). Correction procedures are straightforward to apply, and are often included with statistical analysis packages such as *p.adjust* command in R (https://www.r-project.org/), FDRESTIMATION package in R (Murray & Blume, 2020) and *multipletests* function in the Python STATSMODELS module (Seabold & Perktold, 2010).

1.3 | Applications and being an effective conservation genomicist

1.3.1 | Applied conservation genomics

The most exciting applications of genomics in conservation include questions that could not be addressed with previous molecular and computational approaches. We encouraged ConGen students to consider novel applications of genetic markers and discussed recent practical examples. A few recent examples are provided here, recognizing many other examples are reviewed elsewhere (e.g. Allendorf et al., 2022).

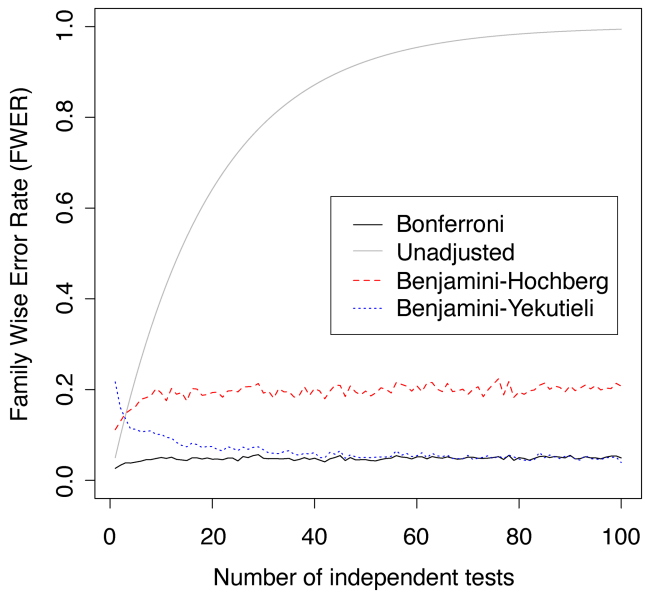

FIGURE 3 Family-wise error rates (FWER) plotted as a function of the number of independent tests, under different correction schemes and the unadjusted FWER in comparison. All correction methods effectively control for FWER, with varying degrees of conservativeness.

Adaptive genomic variation research is guiding assisted gene flow management (Aitken & Whitlock, 2013) and climate-assisted provenancing (Prober et al., 2015) to facilitate adaptation to changing climates. For example, genomic analysis was used to understand adaptive variation in *Corymbia calophylla*, a key species of the forests of south-western Australia. Association of SNPs with precipitation or temperature variables (and mapped in *Corymbia* and *Eucalyptus* reference genomes) was used to assess function of climate-associated variants (Ahrens et al., 2019). This genomic information is being combined with common garden experiments (Ahrens et al., 2020) and ecophysiological analysis (Ahrens, Challis, et al., 2021) to predict species' responses to changing climate, and inform adaptation strategies for forest conservation.

Landscape genomics allows genome-wide associations with environmental variables [also known as genotype–environment association (GEA) studies]; enabling inferences about local adaptation (Rellstab et al., 2015). For example, genome-wide SNP frequencies of coral in New Caledonia were associated with remote-sensed environmental heat data to investigate the local adaptation of reefs to thermal stress. Candidate-adaptive SNPs experiencing selection were used to characterize the functional role of nearby genes and make inferences about adaptive traits that could allow coral to persist in the face of warming oceans (Selmoni et al., 2021). Similarly, genomics is being used to fast-track thermal adaptation in coral across the Great Barrier Reef, Australia through identification of putatively adaptive loci in warm-adapted coral using genomic association methods, which will be used to target individuals for selective breeding, and implement assisted gene flow (Quigley et al., 2019; Suggett & van Oppen, 2022).

In another study, RADseq data for the greenlip abalone (*Haliotis laevigata*) revealed signs of local adaptation in candidate SNPs associated with genes involved in tolerance to high temperature and low oxygen (Lampert, 2018; Sandoval-Castillo et al., 2018). These results could influence management of this commercially important species in the face of climate change (Lopez et al., 2019).

Genomics provides numerous and informative markers often needed to identify the origins and track spread of invasive species, and thereby reduce future spread. For example, *Passiflora foetida*, a climbing vine that is native to Central and South America, has spread to South-east Asia and the Pacific. Genome skimming was used to produce whole chloroplast and ITS sequences from samples collected across the invasive range, which directed collection of herbarium specimens from across the native range. The analysis suggested three independent origins of the species into Australia from Ecuador and Peru, Brazil and the Caribbean, with only one of these lineages becoming widespread (Hopley et al., 2021). This has guided the search for a biological control agent for this species by identifying where in the native range to search for agents specific to these lineages.

1.3.2 | How to be an effective conservation geneticist

Genetic considerations are often missing from conservation management planning and policy (Cook & Sgrò, 2017; Laikre et al., 2020; Pierson et al., 2016). If this gap between genetics research and application persists, the tools and techniques described here will have limited influence in practice. Thus, we need to understand the factors that contribute to this gap and identify the tools and techniques to bridge it.

Many factors may contribute to the lack of integration of genetic data and principals into conservation action plans, where the shift to genome-scale data could exacerbate the gap between population genomics and its application in conservation. Hypothesized drivers of this gap include a lack of awareness of the benefits of integrating genetics and genomics into conservation (Pierson et al., 2016), heavy use of jargon in published communications (Allendorf et al., 2022, chap. 24; Cook & Sgrò, 2019), workflow bottlenecks from the generation or analysis of massive amounts of data and financial constraints. Funding is a problem, including lack of access to pay-to-view peer-reviewed publications by practitioners (Cook & Sgrò, 2017), and genetics may be perceived as costly to implement (Vernesi & Bruford, 2009). There also can be lack of funding and other incentives to motivate geneticists to bridge the gap (Haig et al., 2016; Shafer et al., 2015; reviewed in Allendorf et al., 2022, chap. 24).

It is often still under-appreciated that genetic factors such as inbreeding or fixation of harmful alleles can cause immediate or short-term population problems as suggested by genome sequencing in orcas (Kardos et al., 2023; see also Kardos & Luikart, 2021). Additionally, some practitioners are concerned about the increased handling time of vulnerable species needed to collect the high-quality samples required for some genomic work, while others are confused about the differences between genetics, genomics and genetically modified organisms (Kadykalo et al., 2020; Luikart et al., 2019).

Genomics could widen the gap as the questions that practitioners need answered may be misaligned with what researchers can publish in high-impact journals. For example, a management question might be answered using 10–20 microsatellite markers, but such a dataset can be difficult to publish in a high-profile journal. Fortunately, it is often possible to achieve both aims through effective communication and well-targeted research (see examples in Section 1.3.1).

Geneticists and practitioners worldwide are keen to close the gap between knowledge and application, where practitioners typically want to see genetic data better integrated into management (Cook & Sgrò, 2018; Taft et al., 2020). In the United States and some other places, an effective approach is to have in-house geneticists embedded in agencies conducting conservation activities such as the National Oceanic and Atmospheric Administration (NOAA), Forest Service, U.S. Geological Survey (USGS) and U.S. Fish & Wildlife Service (USFWS). Interestingly, practitioners surveyed in New Zealand do not favour this model due to fears that resources could become monopolized by certain teams and instead prefer external collaborations (Taylor et al., 2017).

An alternative is to form boundary organizations – typically a collection of researchers and practitioners from different organizations working together to bridge the gap via communication and training (Cook et al., 2013). Zoos can also bridge conservation genetic research and practice, where zoos globally have in-house genetics laboratories and well-established functional partnerships with practitioners: the Antwerp Zoo, Chicago Zoological Society, Copenhagen Zoo, Omaha's Henry Doorly Zoo, Royal Zoological Society of Scotland, San Diego Zoo Global, Smithsonian's National Zoo, Taronga Zoo in Sydney Australia and South African National Biodiversity Institute. Nations, governments, tribes and agencies can benefit from understanding advantages or disadvantages of the different strategies used by different countries and organizations mentioned earlier.

Incorporating genetic metrics into the International Union for Conservation of Nature (IUCN) RedList of threatened species would also be an effective motivator to gather and analyse more genetic data for threatened species (Garner et al., 2020). Although integrating genetic data into threat status can be challenging, it is not insurmountable (Funk et al., 2019). Publishing in languages other than English (Holderegger et al., 2019) and producing different versions of textbooks that cater to the needs of geneticists and practitioners (e.g. Frankham et al., 2019) are also practical solutions for broadening access to and adoption of conservation genomics approaches.

After completing academic training in population genetics (including theory), a conservation geneticist can follow numerous career paths, each requiring different skills (see Allendorf et al., 2022, chap. 24). Although under-taught in academic science programmes, skills in communication and building and facilitating collaborations can be honed by seeking out real-world experiences. For example, emergent conservation geneticists can interact with advisors and colleagues who sit on threatened species management steering groups, and ask to sit in on meetings. Courses, such as the facilitation training offered by the IUCN Conservation Planning Specialist Group (http://www.cpsg.org/our-approach/training) or the structured decision-making training from the IUCN Conservation Translocation Specialist Group (https://iucn-ctsg.org/), can help researchers practice using soft skills to engage with conservation managers. While effort is required up front to obtain and hone these skills, the investment will reap benefits, making early-career conservation geneticists more employable in a competitive job market and, more importantly, making them more effective conservationists.

2 | CONCLUSIONS

This review highlights key and emerging topics in conservation genomics research, including the importance of bioinformatics training, novel computational analyses and the understanding of population genetics theory to ensure appropriate data collection and analysis. We provide a novel flow chart (Figure 2) to help researchers choose appropriate marker methods for a range of study questions including the use of

neutral and adaptive loci. We also provide advice on cultivating relationships with practitioners through effective communication to ensure research projects will address conservation concerns. We hope this review helps readers worldwide stay updated on recent advances in population genomics and the growing number of applications of genomics that can help slow the global extinction crisis.

AUTHOR CONTRIBUTIONS

All authors wrote, edited and approved the manuscript.

AFFILIATIONS

[1]Life and Environmental Sciences, University of California, Merced, California, USA
[2]Geospatial Molecular Epidemiology group (GEOME), Laboratory for Biological Geochemistry (LGB), École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[3]Department of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, Canada
[4]Virginia Museum of Natural History, Martinsville, Virginia, USA
[5]Division of Biological Sciences, University of Montana, Missoula, Montana, USA
[6]Washington State University, Pullman, Washington, USA
[7]School of Biological Sciences, University of Bristol, Bristol, UK
[8]Department of Biodiversity, Conservation and Attractions, Biodiversity and Conservation Science, Perth, Western Australia, Australia
[9]Flathead Lake Biology Station, University of Montana, Missoula, Montana, USA
[10]Flathead Lake Biological Station, University of Montana, Polson, Montana, USA
[11]Southwestern Native Aquatic Resources and Recovery Center, U.S. Fish & Wildlife Service, Dexter, New Mexico, USA
[12]Hagerman Genetics Lab, University of Idaho, Hagerman, Idaho, USA
[13]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[14]Department of Evolution, Ecology, and Behavior, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA
[15]Department of Biology and Ecology Center, Utah State University, Logan, Utah, USA
[16]Department of Biology, San Diego State University, San Diego, California, USA
[17]Royal Zoological Society of Scotland, Edinburgh, Scotland
[18]Informatics Group, Harvard University, Cambridge, Massachusetts, USA
[19]Plant and Food Research, Nelson, New Zealand
[20]University of Auckland, Auckland, New Zealand

ACKNOWLEDGEMENTS

We thank Eric Anderson (and other instructors) for providing open access to resources for computing and bioinformatics in conservation genomics and various other ConGen tutorials. We also thank the participants of ConGen 2021 for their feedback on the course. Sponsors supporting ConGen include the National Science Foundation of the United States (grant DoB-1639014), the National Aeronautics and Space Administration (NASA grant 80NSSC19K0185), the American Genetics Association (AGA), Dovetail Genomics, Springer Publishing and Wiley Publishing. The course virtual server was also supported through Paul Hohenlohe (NSF grant DEB-1655809). The use of trade, product, industry or firm names is for informative purposes only and does not constitute an endorsement by the U.S. Government or the U.S. Fish and Wildlife Service. Links to non-Service websites do not imply any official U.S. Fish and Wildlife Service endorsement of the opinions or ideas expressed therein or guarantee the validity of the information provided. The findings, conclusions and opinions expressed in this article represent those of the authors, and do not necessarily represent the views of the U.S. Fish & Wildlife Service.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data used to generate the plots in the figures are available through the hands-on activities presented in this manuscript.

ORCID

Lauren M. Schiebelhut https://orcid.org/0000-0002-5417-5426
Annie S. Guillaume https://orcid.org/0000-0002-2188-2861
Arianna Kuhn https://orcid.org/0000-0002-6891-4672
Rena M. Schweizer https://orcid.org/0000-0002-8812-8177
Ellie E. Armstrong https://orcid.org/0000-0001-7107-6318
Mark A. Beaumont https://orcid.org/0000-0002-8773-2743
Margaret Byrne https://orcid.org/0000-0002-7197-5409
Brian K. Hand https://orcid.org/0000-0003-1152-665X
Leif Howard https://orcid.org/0000-0002-7535-0350
Steven M. Mussmann https://orcid.org/0000-0002-5237-5088
Shawn R. Narum https://orcid.org/0000-0002-7470-7485
Rita Rasteiro https://orcid.org/0000-0002-4217-3060
Angel G. Rivera-Colón https://orcid.org/0000-0001-9097-3241
Norah Saarman https://orcid.org/0000-0001-8974-0301
Arun Sethuraman https://orcid.org/0000-0002-8201-8292
Helen R. Taylor https://orcid.org/0000-0001-7951-0772
Gregg W. C. Thomas https://orcid.org/0000-0001-7621-2980
Maren Wellenreuther https://orcid.org/0000-0002-2764-8291
Gordon Luikart https://orcid.org/0000-0001-8697-0582

REFERENCES

AghaKouchak, A., Chiang, F., Huning, L. S., Love, C. A., Mallakpour, I., Mazdiyasni, O., Moftakhari, H., Papalexiou, S. M., Ragno, E., & Sadegh, M. (2020). Climate extremes and compound hazards in a warming world. *Annual Review of Earth and Planetary Sciences*, *48*(1), 519–548.

Ahrens, C. W., Andrew, M. E., Mazanec, R. A., Ruthrof, K. X., Challis, A., Hardy, G., Byrne, M., Tissue, D. T., & Rymer, P. D. (2020). Plant functional traits differ in adaptability and are predicted to be differentially affected by climate change. *Ecology and Evolution*, *10*(1), 232–248.

Ahrens, C. W., Byrne, M., & Rymer, P. D. (2019). Standing genomic variation within coding and regulatory regions contributes to the adaptive capacity to climate in a foundation tree species. *Molecular Ecology*, *28*(10), 2502–2516.

Ahrens, C. W., Challis, A., Byrne, M., Leigh, A., Nicotra, A. B., Tissue, D., & Rymer, P. (2021). Repeated extreme heatwaves result in higher leaf thermal tolerances and greater safety margins. *The New Phytologist*, *232*(3), 1212–1225.

Ahrens, C. W., Jordan, R., Bragg, J., Harrison, P. A., Hopley, T., Bothwell, H., Murray, K., Steane, D. A., Whale, J. W., Byrne, M., Andrew, R., & Rymer, P. D. (2021). Regarding the F-word: The effects of data filtering on inferred genotype–environment associations. *Molecular Ecology Resources*, *21*(5), 1460–1474.

Aitken, S. N., & Whitlock, M. C. (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 367–388. https://doi.org/10.1146/annurev-ecolsys-110512-135747

Al-Asadi, H., Petkova, D., Stephens, M., & Novembre, J. (2019). Estimating recent migration and population-size surfaces. *PLoS Genetics*, *15*(1), e1007908.

Ali, O. A., O'Rourke, S. M., Amish, S. J., Meek, M. H., Luikart, G., Jeffres, C., & Miller, M. R. (2016). RAD capture (rapture): Flexible and efficient sequence-based genotyping. *Genetics*, *202*(2), 389–400.

Allendorf, F. W. (2017). Genetics and the conservation of natural populations: Allozymes to genomes. *Molecular Ecology*, *26*(2), 420–430.

Allendorf, F. W., Chris Funk, W., Aitken, S. N., Byrne, M., & Luikart, G. (2022). *Conservation and the genomics of populations*. Oxford University Press.

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, *11*(10), 697–709.

Anastasiadi, D., Venney, C. J., Bernatchez, L., & Wellenreuther, M. (2021). Epigenetic inheritance and reproductive mode in plants and animals. *Trends in Ecology & Evolution*, *36*(12), 1124–1140.

Andrews, K. R., Epstein, B., Leslie, M. S., Fiedler, P., Morin, P. A., & Rus Hoelzel, A. (2021). Genomic signatures of divergent selection are associated with social behaviour for spinner dolphin ecotypes. *Molecular Ecology*, *30*(9), 1993–2008.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews. Genetics*, *17*(2), 81–92.

Andrews, K. R., & Luikart, G. (2014). Recent novel approaches for population genomics data analysis. *Molecular Ecology*, *23*(7), 1661–1667.

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics, Babraham Institute. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Anonymous. (2010). Breakthrough of the year. Areas to watch. *Science*, *330*(6011), 1608–1609.

Armstrong, E. E., Taylor, R. W., Miller, D. E., Kaelin, C. B., Barsh, G. S., Hadly, E. A., & Petrov, D. (2020). Long live the King: Chromosome-level assembly of the lion (*Panthera leo*) using linked-read, Hi-C, and long-read data. *BMC Biology*, *18*(1), 3.

Austin, J. D., Gore, J. A., Hargrove, J. S., Elizabeth, C., de Torrez, B., Carneiro, C. M., Ridgley, F. N., & Wisely, S. M. (2022). Strong population genetic structure and cryptic diversity in the Florida bonneted bat (*Eumops floridanus*). *Conservation Genetics*, *23*, 495–512. https://doi.org/10.1007/s10592-022-01432-y

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, *3*(10), e3376.

Balloux, F. (2001). EASYPOP (version 1.7): A computer program for population genetics simulations. *Journal of Heredity*, *92*, 301–302. https://doi.org/10.1093/jhered/92.3.301

Barth, J. M. I., Handley, S. A., Kintzl, D., Leonard, G., Malinsky, M., Matschiner, M., Meyer, B. S., Salzburger, W., Stefka, J., & Trucchi, E. (2023). The history and organization of the workshop on population and speciation genomics. *Evolution*, *16*(1), 2.

Barton, N., & Bengtsson, B. O. (1986). The barrier to genetic exchange between Hybridising populations. *Heredity*, *57*(Pt 3), 357–376.

Beaumont, M. A., & Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, *13*(4), 969–980.

Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *263*(1377), 1619–1626.

Beerli, P., & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in N subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(8), 4563–4568.

Benestan, L. M., Ferchaud, A.-L., Hohenlohe, P. A., Garner, B. A., Naylor, G. J. P., Baums, I. B., Schwartz, M. K., Kelley, J. L., & Luikart, G. (2016). Conservation genomics of natural and managed populations: Building a conceptual and practical framework. *Molecular Ecology*, *25*(13), 2967–2977.

Berg, J. J., & Coop, G. (2014). A population genetic signal of polygenic adaptation. *PLoS Genetics*, *10*(8), e1004412.

Bootsma, M. L., Miller, L., Sass, G. G., Euclide, P. T., & Larson, W. A. (2021). The ghosts of propagation past: Haplotype information clarifies the relative influence of stocking history and Phylogeographic processes on contemporary population structure of walleye (*Sander vitreus*). *Evolutionary Applications*, *14*(4), 1124–1144.

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: From polygenic to omnigenic. *Cell*, *169*(7), 1177–1186.

Bradbury, R. B., Butchart, S. H. M., Fisher, B., Hughes, F. M. R., Ingwall-King, L., MacDonald, M. A., Merriman, J. C., Peh, K. S.-H., Pellier, A.-S., Thomas, D. H. L., Trevelyan, R., & Balmford, A. (2021). The economic consequences of conserving or restoring sites for nature. *Nature Sustainability*, *4*(7), 602–608.

Bradnam, K. R., Fass, J. N., Alexandrov, A., Baranay, P., Bechner, M., Birol, I., Boisvert, S., Chapman, J. A., Chapuis, G., Chikhi, R., Chitsaz, H., Chou, W.-C., Corbeil, J., del Fabbro, C., Docking, T. R., Durbin, R., Earl, D., Emrich, S., Fedotov, P., … Korf, I. F. (2013). Assemblathon 2: Evaluating de Novo methods of genome assembly in three vertebrate species. *GigaScience*, *2*(1), 10.

Burgess, B. T., Irvine, R. L., & Russello, M. A. (2022). Population genomics of Sitka black-tailed deer supports invasive species management and ecological restoration on islands. *Communications Biology*, *5*(1), 223.

Bushnell, B. (2014). BBTools Software Package. http://sourceforge.Net/projects/bbmap 578: 579.

Campbell, N. R., Harmon, S. A., & Narum, S. R. (2015). Genotyping-in-thousands by sequencing (GT-Seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Molecular Ecology Resources*, *15*(4), 855–867.

Carbone, L., Mootnick, A. R., Nadler, T., Moisson, P., Ryder, O., Roos, C., & de Jong, P. J. (2009). A chromosomal inversion unique to the northern white-cheeked gibbon. *PLoS One*, *4*(3), e4999.

Catanach, A., Crowhurst, R., Deng, C., David, C., Bernatchez, L., & Wellenreuther, M. (2019). The genomic pool of standing structural variation outnumbers single nucleotide polymorphism by threefold in the marine teleost *Chrysophrys auratus*. *Molecular Ecology*, *28*(6), 1210–1223.

Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Källberg, M., Cox, A. J., Kruglyak, S., & Saunders, C. T. (2016). Manta: Rapid detection of structural variants and Indels for germline and cancer sequencing applications. *Bioinformatics*, *32*(8), 1220–1222.

Clugston, J. A. R., Kenicer, G. J., Milne, R., Overcast, I., Wilson, T. C., & Nagalingum, N. S. (2019). RADseq as a valuable tool for plants with large genomes – a case study in cycads. *Molecular Ecology Resources*, *19*(6), 1610–1622.

Cohen, Z. P., Schoville, S. D., & Hawthorne, D. J. (2023). The role of structural variants in pest adaptation and genome evolution of the Colorado potato beetle, *Leptinotarsa decemlineata* (say). *Molecular Ecology*, *32*(6), 1425–1440.

Cook, C. N., Mascia, M. B., Schwartz, M. W., Possingham, H. P., & Fuller, R. A. (2013). Achieving conservation science that bridges the knowledge-action boundary. *Conservation Biology: The Journal of the Society for Conservation Biology*, *27*(4), 669–678.

Cook, C. N., & Sgrò, C. M. (2017). Aligning science and policy to achieve evolutionarily enlightened conservation. *Conservation Biology: The Journal of the Society for Conservation Biology*, *31*(3), 501–512.

Cook, C. N., & Sgrò, C. M. (2018). Understanding managers' and scientists' perspectives on opportunities to achieve more evolutionarily enlightened management in conservation. *Evolutionary Applications*, *11*(8), 1371–1388.

Cook, C. N., & Sgrò, C. M. (2019). Conservation practitioners' understanding of how to manage evolutionary processes. *Conservation Biology: The Journal of the Society for Conservation Biology*, *33*(5), 993–1001.

Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, *185*(4), 1411–1423.

Czech, L., & Exposito-Alonso, M. (2022). Grenepipe: A flexible, scalable and reproducible pipeline to automate variant calling from sequence reads. *Bioinformatics*, *38*(20), 4809–4811.

Daub, J. T., Hofer, T., Cutivet, E., Dupanloup, I., Quintana-Murci, L., Robinson-Rechavi, M., & Excoffier, L. (2013). Evidence for polygenic adaptation to pathogens in the human genome. *Molecular Biology and Evolution*, *30*(7), 1544–1558.

Davey, J. W., Cezard, T., Fuentes-Utrilla, P., Eland, C., Gharbi, K., & Blaxter, M. L. (2013). Special features of RAD sequencing data: Implications for genotyping. *Molecular Ecology*, *22*(11), 3151–3164.

Deakin, J. E., Potter, S., O'Neill, R., Ruiz-Herrera, A., Cioffi, M. B., Eldridge, M. D. B., Fukui, K., Marshall Graves, J. A., Griffin, D., Grutzner, F., Kratochvíl, L., Miura, I., Rovatsos, M., Srikulnath, K., Wapstra, E., & Ezaz, T. (2019). Chromosomics: Bridging the gap between genomes and chromosomes. *Genes*, *10*(8), 627. https://doi.org/10.3390/genes10080627

Del Angel, D., Victoria, E. H., Sterck, L., Capella-Gutierrez, S., Notredame, C., Pettersson, O. V., Amselem, J., Bouri, L., Bocs, S., Klopp, C., Gibrat, J. F., Vlasova, A., Leskosek, B. L., Soler, L., Binzer-Panchal, M., & Lantz, H. (2018). Ten steps to get started in genome assembly and annotation. *F1000Research*, *7*, 148. https://doi.org/10.12688/f1000research.13598.1

Delomas, T. A., & Campbell, M. R. (2022). Grandparent inference from genetic data: The potential for parentage-based tagging programs to identify offspring of hatchery strays. *North American Journal of Fisheries Management*, *42*(1), 85–95.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimatorv2: Re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, *14*, 209–214. https://doi.org/10.1111/1755-0998.12157

Dorant, Y., Cayuela, H., Wellband, K., Laporte, M., Rougemont, Q., Mérot, C., Normandeau, E., Rochette, R., & Bernatchez, L. (2020). Copy number variants outperform SNPs to reveal genotype–temperature association in a marine species. *Molecular Ecology*, *29*, 4765–4782. https://doi.org/10.1111/mec.15565

Duforet-Frebourg, N., Luu, K., Laval, G., Bazin, E., & Blum, M. G. B. (2016). Detecting genomic signatures of natural selection with principal component analysis: Application to the 1000 genomes data. *Molecular Biology and Evolution*, *33*(4), 1082–1093.

Earl, D., Bradnam, K., John, J. S., Darling, A., Lin, D., Fass, J., Yu, H. O., Buffalo, V., Zerbino, D. R., Diekhans, M., Nguyen, N., Ariyaratne, P. N., Sung, W. K., Ning, Z., Haimel, M., Simpson, J. T., Fonseca, N. A., Birol, İ., Docking, T. R., … Paten, B. (2011). Assemblathon 1: A competitive assessment of de Novo short read assembly methods. *Genome Research*, *21*(12), 2224–2241.

Eaton, D. A. R., & Overcast, I. (2020). Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*(8), 2592–2594.

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, *32*(19), 3047–3048.

Excoffier, L., Hofer, T., & Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*(4), 285–298.

Fariello, M. I., Boitard, S., Naya, H., SanCristobal, M., & Servin, B. (2013). Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics*, *193*(3), 929–941.

Faubet, P., Waples, R. S., & Gaggiotti, O. E. (2007). Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Molecular Ecology*, *16*(6), 1149–1166.

Ferreira, M. S., Thurman, T. J., Jones, M. R., Farelo, L., Kumar, A. V., Mortimer, S. M. E., Demboski, J. R., Mills, L. S., Alves, P. C., Melo-Ferreira, J., & Good, J. M. (2023). The evolution of white-tailed jackrabbit camouflage in response to past and future seasonal climates. *Science*, *379*(6638), 1238–1242.

Feuk, L., Carson, A. R., & Scherer, S. W. (2006). Structural variation in the human genome. *Nature Reviews. Genetics*, *7*(2), 85–97.

Flores-Manzanero, A., Valenzuela-Galván, D., Cuarón, A. D., & Vázquez-Domínguez, E. (2022). Conservation genetics of two critically endangered island dwarf carnivores. *Conservation Genetics*, *23*(1), 35–49.

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*(2), 977–993.

Forester, B. R., Murphy, M., Mellison, C., Petersen, J., Pilliod, D. S., Van Horne, R., Harvey, J., & Chris Funk, W. (2022). Genomics-informed delineation of conservation units in a desert amphibian. *Molecular Ecology*, *31*(20), 5249–5269.

Frankham, R., Ballou, J. D., Ralls, K., Eldridge, M., Dudash, M. R., Fenster, C. B., Lacy, R. C., & Sunnucks, P. (2019). *A practical guide for genetic management of fragmented animal and plant populations*. Oxford Academic. https://doi.org/10.1093/oso/9780198783411.001.0001

Funk, W. C., Forester, B. R., Converse, S. J., Darst, C., & Morey, S. (2019). Improving conservation policy with genomics: A guide to integrating adaptive potential into U.S. endangered species act decisions for conservation practitioners and geneticists. *Conservation Genetics*, *20*(1), 115–134.

Galla, S. J., Brown, L., Couch-Lewis Ngāi Tahu Te Hapū O Ngāti Wheke Ngāti Waewae, Y., Cubrinovska, I., Eason, D., Gooley, R. M., Hamilton, J. A., Heath, J. A., Hauser, S. S., Latch, E. K., Matocq, M. D., Richardson, A., Wold, J. R., Hogg, C. J., Santure, A. W., & Steeves, T. E. (2022). The relevance of pedigrees in the conservation genomics era. *Molecular Ecology*, *31*(1), 41–54.

Gargiulo, R., Kull, T., & Fay, M. F. (2021). Effective double-digest RAD sequencing and genotyping despite large genome size. *Molecular Ecology Resources*, *21*(4), 1037–1055.

Garner, B. A., Hoban, S., & Luikart, G. (2020). IUCN red list and the value of integrating genetics. *Conservation Genetics*, *21*(5), 795–801.

Garrison, N. L., Johnson, P. D., & Whelan, N. V. (2021). Conservation genomics reveals low genetic diversity and multiple parentage in the threatened freshwater mussel, *Margaritifera hembeli*. *Conservation Genetics*, *22*(2), 217–231.

Gompert, Z., & Buerkle, C. A. (2011). Bayesian estimation of genomic clines. *Molecular Ecology*, *20*(10), 2111–2127.

Gompert, Z., & Buerkle, C. A. (2012). Bgc: Software for Bayesian estimation of genomic clines. *Molecular Ecology Resources*, *12*(6), 1168–1176.

Gouy, A., & Excoffier, L. (2020). Polygenic patterns of adaptive introgression in modern humans are mainly shaped by response to pathogens. *Molecular Biology and Evolution*, *37*(5), 1420–1433.

Graffelman, J., & Weir, B. S. (2018). On the testing of Hardy–Weinberg proportions and equality of allele frequencies in males and females at biallelic genetic markers. *Genetic Epidemiology*, *42*(1), 34–48.

Graham, C. F., Boreham, D. R., Manzon, R. G., Stott, W., Wilson, J. Y., & Somers, C. M. (2020). How 'simple' methodological decisions affect interpretation of population structure based on reduced representation library DNA sequencing: A case study using the lake whitefish. *PLoS One*, *15*(1), e0226608.

Günther, T., & Nettelblad, C. (2019). The presence and impact of reference bias on population genomic studies of prehistoric human populations. *PLoS Genetics*, *15*(7), e1008302.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072–1075.

Haig, S. M., Miller, M. P., Bellinger, R., Draheim, H. M., Mercer, D. M., & Mullins, T. D. (2016). The conservation genetics juggling act: Integrating genetics and ecology, science and policy. *Evolutionary Applications*, *9*(1), 181–195.

Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward genetic simulations beyond the Wright–Fisher model. *Molecular Biology and Evolution*, *36*(3), 632–637.

Hauser, S., Wakeland, K., & Leberg, P. (2019). Inconsistent use of multiple comparison corrections in studies of population genetic structure: Are some type I errors more tolerable than others? *Molecular Ecology Resources*, *19*(1), 144–148.

Hemstrom, W., Grummer, J., Luikart, G., & Christie, M. (n.d.). The 'F-word': Next-generation data filtering in the genomics era.

Hemstrom, W., & Jones, M. (2023). snpR: User friendly population genomics for SNP data sets with categorical metadata. *Molecular Ecology Resources*, *23*(4), 962–973.

Hendricks, S., Anderson, E. C., Antao, T., Bernatchez, L., Forester, B. R., Garner, B., Hand, B. K., Hohenlohe, P. A., Kardos, M., Koop, B., Sethuraman, A., Waples, R. S., & Luikart, G. (2018). Recent advances in conservation and population genomics data analysis. *Evolutionary Applications*, *11*(8), 1197–1211.

Hey, J., & Nielsen, R. (2007). Integration within the Felsenstein equation for improved Markov Chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(8), 2785–2790.

Hoban, S. (2014). An overview of the utility of population simulation software in molecular ecology. *Molecular Ecology*, *23*(10), 2383–2401.

Hoban, S., Bertorelle, G., & Gaggiotti, O. E. (2012). Computer simulations: Tools for population and evolutionary genetics. *Nature Reviews. Genetics*, *13*(2), 110–122.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A., & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, *188*(4), 379–397.

Hohenlohe, P. A., Hand, B. K., Andrews, K. R., & Luikart, G. (2019). Population genomics provides key insights in ecology and evolution. In O. P. Rajora (Ed.), *Population genomics: Concepts, approaches and applications* (pp. 483–510). Springer International Publishing.

Holderegger, R., Balkenhol, N., Bolliger, J., Engler, J. O., Gugerli, F., Hochkirch, A., Nowak, C., Segelbacher, G., Widmer, A., & Zachos, F. E. (2019). Conservation genetics: Linking science with practice. *Molecular Ecology*, *28*(17), 3848–3856.

Hopley, T., Webber, B. L., Raghu, S., Morin, L., & Byrne, M. (2021). Revealing the introduction history and phylogenetic relationships of *Passiflora foetida* sensu Lato in Australia. *Frontiers in Plant Science*, *12*(July), 651805.

Howard-McCombe, J., Ward, D., Kitchener, A. C., Lawson, D., Senn, H. V., & Beaumont, M. (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*(15), 3688–3702.

IPCC. (2021). Climate change 2021: The physical science basis. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*. Cambridge University Press.

Johri, P., Charlesworth, B., & Jensen, J. D. (2020). Toward an evolutionarily appropriate null model: Jointly inferring demography and purifying selection. *Genetics*, *215*(1), 173–192.

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405.

Kadykalo, A. N., Cooke, S. J., & Young, N. (2020). Conservation genomics from a practitioner lens: Evaluating the research-implementation gap in a managed freshwater fishery. *Biological Conservation*, *241*, 108350.

Kardos, M., Armstrong, E. E., Fitzpatrick, S. W., Hauser, S., Hedrick, P. W., Miller, J. M., Tallmon, D. A., & Chris Funk, W. (2021). The crucial role of genome-wide genetic variation in conservation. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(48), e2104642118. https://doi.org/10.1073/pnas.2104642118

Kardos, M., & Luikart, G. (2021). The genetic architecture of fitness drives population viability during rapid environmental change. *The American Naturalist*, *197*(5), 511–525.

Kardos, M., Zhang, Y., Parsons, K. M., Yunga, A., Kang, H., Xun, X., Liu, X., Matkin, C. O., Zhang, P., Ward, E. J., Hanson, M. B., Emmons, C., Ford, M. J., Fan, G., & Li, S. (2023). Inbreeding depression explains killer Whale population dynamics. *Nature Ecology & Evolution*, *7*(5), 675–686.

Kelleher, J., Etheridge, A. M., & McVean, G. (2016). Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLOS Computational Biology*, *12*, e1004842. https://doi.org/10.1371/journal.pcbi.1004842

Kimura, M., & Weiss, G. H. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, *49*(4), 561–576.

Laikre, L., Hoban, S., Bruford, M. W., Segelbacher, G., Allendorf, F. W., Gajardo, G., Rodríguez, A. G., Hedrick, P. W., Heuertz, M., Hohenlohe, P. A., Jaffé, R., Johannesson, K., Liggins, L., MacDonald, A. J., OrozcoterWengel, P., Reusch, T. B. H., Rodríguez-Correa, H., Russo, I.-R. M., Ryman, N., & Vernesi, C. (2020). Post-2020 goals overlook genetic diversity. *Science*, *367*(6482), 1083–1085.

Lamichhaney, S., Fan, G., Widemo, F., Gunnarsson, U., Thalmann, D. S., Hoeppner, M. P., Kerje, S., Gustafson, U., Shi, C., Zhang, H., Chen, W., Liang, X., Huang, L., Wang, J., Liang, E., Wu, Q., Lee, S. M.-Y., Xu, X., Höglund, J., … Andersson, L. (2016). Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nature Genetics*, *48*(1), 84–88.

Lampert, K. P. (2018). Oxygen concentration drives local adaptation in the Greenlip abalone (*Haliotis laevigata*). *Molecular Ecology*, *27*, 1521–1523.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., … Wyman, D. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921.

Langmead, B., & Nellore, A. (2018). Cloud computing for genomic data analysis and collaboration. *Nature Reviews. Genetics*, *19*(4), 208–219.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nature Methods*, *9*(4), 357–359.

Lawniczak, M. K. N., Durbin, R., Flicek, P., Lindblad-Toh, K., Wei, X., Archibald, J. M., Baker, W. J., Belov, K., Blaxter, M. L., Bonet, T. M., Childers, A. K., Coddington, J. A., Crandall, K. A., Crawford, A. J., Davey, R. P., Di Palma, F., Fang, Q., Haerty, W., Hall, N., … Richards, S. (2022). Standards recommendations for the earth BioGenome project. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(4), e2115639118. https://doi.org/10.1073/pnas.2115639118

Layer, R. M., Chiang, C., Quinlan, A. R., & Hall, I. M. (2014). LUMPY: A probabilistic framework for structural variant discovery. *Genome Biology*, *15*(6), R84.

Leitwein, M., Duranton, M., Rougemont, Q., Gagnaire, P.-A., & Bernatchez, L. (2020). Using haplotype information for conservation genomics. *Trends in Ecology & Evolution*, *35*(3), 245–258.

Lewontin, R. C., & Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, *74*(1), 175–195.

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Lin, J., Zhang, W., Zhang, X., Ma, X., Zhang, S., Chen, S., Wang, Y., Jia, H., Liao, Z., Lin, J., Zhu, M., Xu, X., Cai, M., Zeng, H., Wan, J., Yang, W., Matsumoto, T., Hardner, C., Nock, C. J., & Ming, R. (2022). Signatures of selection in recently domesticated macadamia. *Nature Communications*, *13*(1), 242.

Linscott, T. M., González-González, A., Hirano, T., & Parent, C. E. (2022). De Novo genome assembly and genome skims reveal LTRs dominate the genome of a limestone endemic Mountainsnail (*Oreohelix idahoensis*). *BMC Genomics*, *23*(1), 796.

Liu, X., Zhang, Y., Li, Y., Pan, J., Wang, D., Chen, W., Zheng, Z., He, X., Zhao, Q., Pu, Y., Guan, W., Han, J., Orlando, L., Ma, Y., & Jiang, L. (2019). EPAS1 gain-of-function mutation contributes to high-altitude adaptation in Tibetan horses. *Molecular Biology and Evolution*, *36*(11), 2591–2603.

Lopez, J. V., Kamel, B., Medina, M., Collins, T., & Baums, I. B. (2019). Multiple facets of marine invertebrate conservation genomics. *Annual Review of Animal Biosciences*, *7*, 473–497.

López-Uribe, M. M., Urbina, J. V., Mejía, A. I., Duque, L. O., Riaño-Jiménez, D., Cure, J. R., Ramos, V., Martel, C., Fuentes, J. D., & González, V. H. (2022). Creating a virtual international research experience. *American Entomologist*, *68*(1), 24–27.

Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, *23*(9), 2178–2192.

Lou, R. N., Jacobs, A., Wilder, A. P., & Therkildsen, N. O. (2021). A Beginner's guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*, *30*(23), 5966–5993.

Lovejoy, T. E., & Hannah, L. J. (2019). *Biodiversity and climate change: Transforming the biosphere*. Yale University Press.

Lowe, W. H., & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology*, *19*(15), 3038–3051.

Luikart, G., Antao, T., Hand, B. K., Muhlfeld, C. C., Boyer, M. C., Cosart, T., Trethewey, B., Al-Chokhachy, R., & Waples, R. S. (2021). Detecting population declines via monitoring the effective number of breeders (Nb). *Molecular Ecology Resources*, *21*(2), 379–393.

Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews. Genetics*, *4*(12), 981–994.

Luikart, G., Kardos, M., Hand, B. K., Rajora, O. P., Aitken, S. N., & Hohenlohe, P. A. (2019). Population genomics: Advancing understanding of nature. In O. P. Rajora (Ed.), *Population genomics: Concepts, approaches and applications* (pp. 3–79). Springer International Publishing.

Mahmoud, M., Gobet, N., Cruz-Dávalos, D. I., Mounier, N., Dessimoz, C., & Sedlazeck, F. J. (2019). Structural variant calling: The long and the short of it. *Genome Biology*, *20*(1), 246.

Mamoozadeh, N. R., Graves, J. E., & McDowell, J. R. (2020). Genome-wide SNPs resolve spatiotemporal patterns of connectivity within striped Marlin (*Kajikia audax*), a broadly distributed and highly migratory pelagic species. *Evolutionary Applications*, *13*(4), 677–698.

Marques, A. J., Doncheff, J. O., Hanson, M. C.-S., Martínez-Solano, I., Moritz, C., Tarroso, P., Velo-Antón, G., Veríssimo, A., & Carvalho, S. B. (2022). Range-wide genomic scans and tests for selection identify non-neutral spatial patterns of genetic variation in a non-model amphibian species (*Pelobates cultripes*). *Conservation Genetics*, *23*, 387–400. https://doi.org/10.1007/s10592-021-01425-3

Martin, B. T., Chafin, T. K., Douglas, M. R., Placyk Jr, J. S., Birkhead, R. D., Phillips, C. A., & Douglas, M. E. (2021). The choices we make and the impacts they have: Machine learning and species delimitation in north American box turtles (*Terrapene* spp.). *Molecular Ecology Resources*, *21*(8), 2801–2817.

Marx, V. (2023). Method of the year: Long-read sequencing. *Nature Methods*, *20*(1), 6–11.

McFarlane, S. E., Senn, H. V., Smith, S. L., & Pemberton, J. M. (2021). Locus-specific introgression in Young hybrid swarms: Drift may dominate selection. *Molecular Ecology*, *30*(9), 2104–2115.

McKinney, G. J., Larson, W. A., Seeb, L. W., & Seeb, J. E. (2017). RADseq provides unprecedented insights into molecular ecology and evolutionary genetics: Comment on breaking RAD by Lowry et al. (2016). *Molecular Ecology Resources*, *17*, 356–361.

Meek, M. H., & Larson, W. A. (2019). The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. *Molecular Ecology Resources*, *19*(4), 795–803.

Meirmans, P. G. (2014). Nonconvergence in Bayesian estimation of migration rates. *Molecular Ecology Resources*, *14*(4), 726–733.

Mérot, C., Berdan, E. L., Cayuela, H., Djambazian, H., Ferchaud, A.-L., Laporte, M., Normandeau, E., Ragoussis, J., Wellenreuther, M., & Bernatchez, L. (2021). Locally adaptive inversions modulate genetic variation at different geographic scales in a seaweed fly. *Molecular Biology and Evolution*, *38*(9), 3953–3971.

Mérot, C., Oomen, R. A., Tigano, A., & Wellenreuther, M. (2020). A roadmap for understanding the evolutionary significance of structural genomic variation. *Trends in Ecology & Evolution*, *35*(7), 561–572.

Mirchandani, C. D., Shultz, A. J., Thomas, G. W. C., Smith, S. J., Baylis, M., Arnold, B., Corbett-Detig, R., Enbody, E., & Sackton, T. B. (2023). A fast, reproducible, high-throughput variant calling workflow for evolutionary, ecological, and conservation genomics. *bioRxiv*. https://doi.org/10.1101/2023.06.22.546168

Mona, S., Benazzo, A., Delrieu-Trottin, E., & Lesturgie, P. (2023). Population genetics using low coverage RADseq data in non-model organisms: Biases and solutions. *Authorea Preprints*. https://doi.org/10.22541/au.168252801.19878064/v1

Murray, M. H., & Blume, J. D. (2020). False discovery rate computation: Illustrations and modifications. *arXiv [stat.ME]*. arXiv. http://arxiv.org/abs/2010.04680

Mussmann, S. M. (2018). *Diversification across a dynamic landscape: Phylogeography and Riverscape genetics of speckled dace (Rhinichthys osculus) in Western North America*. University of Arkansas. https://scholarworks.uark.edu/etd/2969

Mussmann, S. M., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2019). BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data. *Methods in Ecology and Evolution/British Ecological Society*, *10*(10), 1808–1813.

Narum, S. R. (2006). Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics*, *7*(5), 783–787.

Nash, H., Wirdateti, W., Low, G., Choo, S. W., Chong, J. L., Semiadi, G., Hari, R., Sulaiman, M. H., Turvey, S., Evans, T. A., & Rheindt, F. (2018). Conservation genomics reveals possible illegal trade routes and admixture across pangolin lineages in Southeast Asia. *Conservation Genetics*, *19*(5), 1083–1095.

Nazareno, A. G., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: An empirical study from an Amazonian plant species. *Molecular Ecology Resources*, *17*(6), 1136–1147.

Neuenschwander, S. (2006). AQUASPLATCHE: A program to simulate genetic diversity in populations living in linear habitats. *Molecular Ecology Notes*, *6*(3), 583–585.

Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, *39*, 197–218.

Nielsen, R., & Wakeley, J. (2001). Distinguishing migration from isolation: A Markov Chain Monte Carlo approach. *Genetics*, *158*(2), 885–896.

Nunziata, S. O., & Weisrock, D. W. (2018). Estimation of contemporary effective population size and population declines using RAD sequence data. *Heredity*, *120*(3), 196–207.

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci You're looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, *27*, 3193–3206. https://doi.org/10.1111/mec.14792

Pearse, D. E. (2016). Saving the spandrels? Adaptive genomic variation in conservation and fisheries management. *Journal of Fish Biology*, *89*(6), 2697–2716.

Pearse, D. E., Barson, N. J., Nome, T., Gao, G., Campbell, M. A., Abadía-Cardoso, A., Anderson, E. C., Rundio, D. E., Williams, T. H., Naish, K. A., Moen, T., Liu, S., Kent, M., Moser, M., Minkley, D. R., Rondeau, E. B., Brieuc, M. S. O., Sandve, S. R., Miller, M. R., … Lien, S. (2019).

Sex-dependent dominance maintains migration supergene in rainbow trout. *Nature Ecology & Evolution*, *3*(12), 1731–1742.

Pierson, J. C., Coates, D. J., Gerard, J., Oostermeijer, B., Beissinger, S. R., Bragg, J. G., Sunnucks, P., Schumaker, N. H., & Young, A. G. (2016). Genetic factors in threatened species recovery plans on three continents. *Frontiers in Ecology and the Environment*, *14*(8), 433–440.

Prasad, A., Lorenzen, E. D., & Westbury, M. V. (2022). Evaluating the role of reference-genome phylogenetic distance on evolutionary inference. *Molecular Ecology Resources*, *22*(1), 45–55.

Prince, D. J., O'Rourke, S. M., Thompson, T. Q., Ali, O. A., Lyman, H. S., Saglam, I. K., Hotaling, T. J., Spidle, A. P., & Miller, M. R. (2017). The evolutionary basis of premature migration in Pacific Salmon highlights the utility of genomics for informing conservation. *Science Advances*, *3*(8), e1603198.

Privé, F., Luu, K., Vilhjálmsson, B. J., & Blum, M. G. B. (2020). Performing highly efficient genome scans for local adaptation with R package Pcadapt version 4. *Molecular Biology and Evolution*, *37*(7), 2153–2154.

Prober, S., Byrne, M., McLean, E., Steane, D., Potts, B., Vaillancourt, R., & Stock, W. (2015). Climate-adjusted provenancing: A strategy for climate-resilient ecological restoration. *Frontiers in Ecology and Evolution*, *3*, 65. https://doi.org/10.3389/fevo.2015.00065

Puritz, J. B., Hollenbeck, C. M., & Gold, J. R. (2014). dDocent: A RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, *2*, e431.

Quigley, K. M., Bay, L. K., & van Oppen, M. J. H. (2019). The active spread of adaptive variation for reef resilience. *Ecology and Evolution*, *9*(19), 11122–11135.

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rajawat, D., Panigrahi, M., Kumar, H., Nayak, S. S., Parida, S., Bharat Bhushan, G. K., Gaur, T. D., & Mishra, B. P. (2022). Identification of important genomic footprints using eight different selection signature statistics in domestic cattle breeds. *Gene*, *816*, 146165.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Systematic Biology*, *67*(5), 901–904.

Rausch, T., Zichner, T., Schlattl, A., Stütz, A. M., Benes, V., & Korbel, J. O. (2012). DELLY: Structural variant discovery by integrated paired-end and Split-read analysis. *Bioinformatics*, *28*(18), i333–i339.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370.

Reynes, L., Aurelle, D., Chevalier, C., Pinazo, C., Valero, M., Mauger, S., Sartoretto, S., Blanfuné, A., Ruitton, S., Boudouresque, C.-F., Verlaque, M., & Thibaut, T. (2021). Population genomics and Lagrangian modeling shed light on dispersal events in the Mediterranean endemic *Ericaria zosteroides* (=*Cystoseira zosteroides*) (Fucales). *Frontiers in Marine Science*, *8*, 683528. https://doi.org/10.3389/fmars.2021.683528

Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., Lee, C., Ko, B. J., Chaisson, M., Gedman, G. L., Cantin, L. J., Thibaud-Nissen, F., Haggerty, L., Bista, I., Smith, M., … Jarvis, E. D. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, *592*(7856), 737–746.

Rivera-Colón, A. G., Rochette, N. C., & Catchen, J. M. (2021). Simulation with RADinitio improves RADseq experimental design and sheds light on sources of missing data. *Molecular Ecology Resources*, *21*(2), 363–378.

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737–4754.

Rochette, N. C., Rivera-Colón, A. G., Walsh, J., Sanger, T. J., Campbell-Staton, S. C., & Catchen, J. M. (2023). On the causes, consequences, and avoidance of PCR duplicates: Towards a theory of library complexity. *Molecular Ecology Resources*, *23*(6), 1299–1318.

Rodríguez-Ezpeleta, N., Zinger, L., Kinziger, A., Bik, H. M., Bonin, A., Coissac, E., Emerson, B. C., Lopes, C. M., Pelletier, T. A., Taberlet, P., & Narum, S. (2021). Biodiversity monitoring using environmental DNA. *Molecular Ecology Resources*, *21*(5), 1405–1409.

Rstudio Team. (2020). *RStudio: Integrated development for R*. RStudio, PBC. http://www.rstudio.com/

Sanchis-Juan, A., Stephens, J., French, C. E., Gleadall, N., Mégy, K., Penkett, C., Shamardina, O., Stirrups, K., Delon, I., Dewhurst, E., Dolling, H., Erwood, M., Grozeva, D., Stefanucci, L., Arno, G., Webster, A. R., Cole, T., Austin, T., Branco, R. G., … Carss, K. J. (2018). Complex structural variants in Mendelian disorders: Identification and breakpoint resolution using short- and long-read genome sequencing. *Genome Medicine*, *10*(1), 95.

Sandoval-Castillo, J., Robinson, N. A., Hart, A. M., Strain, L. W. S., & Beheregaray, L. B. (2018). Seascape genomics reveals adaptive divergence in a connected and commercially important Mollusc, the Greenlip abalone (*Haliotis laevigata*), along a longitudinal environmental gradient. *Molecular Ecology*, *27*(7), 1603–1620.

Santure, A. W., & Garant, D. (2018). Wild GWAS-association mapping in natural populations. *Molecular Ecology Resources*, *18*(4), 729–738.

Sarver, B. A. J., Keeble, S., Cosart, T., Tucker, P. K., Dean, M. D., & Good, J. M. (2017). Phylogenomic insights into mouse evolution using a pseudoreference approach. *Genome Biology and Evolution*, *9*(3), 726–739.

Schmidt, T. L., Jasper, M.-E., Weeks, A. R., & Hoffmann, A. A. (2021). Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods in Ecology and Evolution/British Ecological Society*, *12*(10), 1888–1898.

Schweizer, R. M., Saarman, N., Ramstad, K. M., Forester, B. R., Kelley, J. L., Hand, B. K., Malison, R. L., Ackiss, A. S., Watsa, M., Nelson, T. C., Beja-Pereira, A., Waples, R. S., Funk, W. C., & Luikart, G. (2021). Big data in conservation genomics: Boosting skills, hedging bets, and staying current in the field. *The Journal of Heredity*, *112*(4), 313–327.

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. In *Proceedings of the 9th python in science conference* (Vol. *57*, p. 61).

Selmoni, O., Lecellier, G., Magalon, H., Vigliola, L., Oury, N., Benzoni, F., Peignon, C., Joost, S., & Berteaux-Lecellier, V. (2021). Seascape genomics reveals candidate molecular targets of heat stress adaptation in three coral species. *Molecular Ecology*, *30*(8), 1892–1906.

Sethuraman, A., Gonzalez, N. M., Grenier, C. E., Kansagra, K. S., Mey, K. K., Nunez-Zavala, S. B., Summerhays, B. E. W., & Wulf, G. K. (2019). Continued misuse of multiple testing correction methods in population genetics – a wake-up call? *Molecular Ecology Resources*, *19*(1), 23–26.

Shafer, A. B. A., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B. W. (2017). Bioinformatic processing of RAD-Seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, *8*(8), 907–917.

Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., Colling, G., Dalén, L., de Meester, L., Ekblom, R., Fawcett, K. D., Fior, S., Hajibabaei, M., Hill, J. A., Hoezel, A. R., Höglund, J., Jensen, E. L., Krause, J., Kristensen, T. N., … Zieliński, P. (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution*, *30*(2), 78–87.

Shaffer, H. B., Toffelmier, E., Corbett-Detig, R. B., Escalona, M., Erickson, B., Fiedler, P., Gold, M., Harrigan, R. J., Hodges, S., Luckau, T. K., Miller, C., Oliveira, D. R., Shaffer, K. E., Shapiro, B., Sork, V. L., & Wang, I. J. (2022). Landscape genomics to enable conservation actions: The California conservation genomics project. *The Journal of Heredity*, *113*(6), 577–588.

Smith, E. G., Hazzouri, K. M., Choi, J. Y., Delaney, P., Al-Kharafi, M., Howells, E. J., Aranda, M., & Burt, J. A. (2022). Signatures of

selection underpinning rapid coral adaptation to the world's warmest reefs. *Science Advances*, *8*(2), eabl7287.
Stuart, Y. E., Campbell, T. S., Hohenlohe, P. A., Reynolds, R. G., Revell, L. J., & Losos, J. B. (2014). Rapid evolution of a native species following invasion by a congener. *Science*, *346*(6208), 463–466.
Sturtevant, A. H. (1913). The linear arrangement of six sex-linked factors in drosophila, as shown by their mode of association. *The Journal of Experimental Zoology*, *14*(1), 43–59.
Sturtevant, A. H., & Mather, K. (1938). The interrelations of inversions, heterosis and recombination. *The American Naturalist*, *72*(742), 447–452.
Suggett, D. J., & van Oppen, M. J. H. (2022). Horizon scan of rapidly advancing coral restoration approaches for 21st century reef management. *Emerging Topics in Life Sciences*, *6*(1), 125–136.
Sundqvist, L., Keenan, K., Zackrisson, M., Prodöhl, P., & Kleinhans, D. (2016). Directional genetic differentiation and relative migration. *Ecology and Evolution*, *6*(11), 3461–3475.
Taft, H. R., McCoskey, D. N., Miller, J. M., Pearson, S. K., Coleman, M. A., Fletcher, N. K., Mittan, C. S., Meek, M. H., & Barbosa, S. (2020). Research–management partnerships: An opportunity to integrate genetics in conservation actions. *Conservation Science and Practice*, *2*(9), e218. https://doi.org/10.1111/csp2.218
Taylor, H. R., Dussex, N., & van Heezik, Y. (2017). Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Global Ecology and Conservation*, *10*, 231–242.
The Darwin Tree of Life Project Consortium, Blaxter, M., Mieszkowska, N., Di Palma, F., Holland, P., Durbin, R., Richards, T., Berriman, M., Kersey, P., Hollingsworth, P., Wilson, W., Twyford, A., Gaya, E., Lawniczak, M., Lewis, O., Broad, G., Howe, K., Hart, M., Flicek, P., & Barnes, I. (2022). Sequence locally, think globally: The Darwin tree of life project. *Proceedings of the National Academy of Sciences*, *119*(4), e2115642118.
Theissinger, K., Fernandes, C., Formenti, G., Bista, I., Berg, P. R., Bleidorn, C., Bombarely, A., Crottini, A., Gallo, G. R., Godoy, J. A., Jentoft, S., Malukiewicz, J., Mouton, A., Oomen, R. A., Paez, S., Palsbøll, P. J., Pampoulie, C., Ruiz-López, M. J., Secomandi, S., … Zammit, G. (2023). How genomics can help biodiversity conservation. *Trends in Genetics: TIG*, *39*(7), 545–559.
Thorburn, D.-M. J., Sagonas, K., Binzer-Panchal, M., Chain, F. J. J., Feulner, P. G. D., Bornberg-Bauer, E., Reusch, T. B. H., Samonte-Padilla, I. E., Milinski, M., Lenz, T. L., & Eizaguirre, C. (2023). Origin matters: Using a local reference genome improves measures in population genomics. *Molecular Ecology Resources*, *23*, 1706–1723. https://doi.org/10.1111/1755-0998.13838
Vernesi, C., & Bruford, M. W. (2009). Recent developments in molecular tools for conservation. In G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, & C. Vernesi (Eds.), *Population genetics for animal conservation* (pp. 321–344). Cambridge University Press.
Wagner, H. H., Murphy, M. A., Holderegger, R., & Waits, L. (2012). Developing an interdisciplinary, distributed graduate course for twenty-first century scientists. *Bioscience*, *62*(2), 182–188.
Wang, T., Antonacci-Fulton, L., Howe, K., Lawson, H. A., Lucas, J. K., Phillippy, A. M., Popejoy, A. B., Asri, M., Carson, C., Chaisson, M. J. P., Chang, X., Cook-Deegan, R., Felsenfeld, A. L., Fulton, R. S., Garrison, E. P., Garrison, N.'. A., Graves-Lindsay, T. A., Ji, H., Kenny, E. E., … the Human Pangenome Reference Consortium. (2022). The human Pangenome project: A global resource to map genomic diversity. *Nature*, *604*(7906), 437–446.
Waples, R. S. (2014). Testing for Hardy–Weinberg proportions: Have we lost the plot? *The Journal of Heredity*, *106*(1), 1–19.
Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary N E using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, *3*(3), 244–262.
Waples, R. S., Waples, R. K., & Ward, E. J. (2022). Pseudoreplication in genomic-scale data sets. *Molecular Ecology Resources*, *22*(2), 503–518.
Wellenreuther, M., & Bernatchez, L. (2018). Eco-evolutionary genomics of chromosomal inversions. *Trends in Ecology & Evolution*, *33*(6), 427–440.
Wellenreuther, M., Mérot, C., Berdan, E., & Bernatchez, L. (2019). Going beyond SNPs: The role of structural genomic variants in adaptive evolution and species diversification. *Molecular Ecology*, *28*(6), 1203–1209.
Wenger, A. M., Peluso, P., Rowell, W. J., Chang, P.-C., Hall, R. J., Concepcion, G. T., Ebler, J., Fungtammasan, A., Kolesnikov, A., Olson, N. D., Töpfer, A., Alonge, M., Mahmoud, M., Qian, Y., Chin, C.-S., Phillippy, A. M., Schatz, M. C., Myers, G., DePristo, M. A., … Hunkapiller, M. W. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nature Biotechnology*, *37*(10), 1155–1162.
Whibley, A., Kelley, J. L., & Narum, S. R. (2021). The changing face of genome assemblies: Guidance on achieving high-quality reference genomes. *Molecular Ecology Resources*, *21*(3), 641–652.
White, T., van der Ende, J., & Nichols, T. E. (2019). Beyond Bonferroni revisited: Concerns over inflated false positive research findings in the fields of conservation genetics, biology, and medicine. *Conservation Genetics*, *20*(4), 927–937.
Willi, Y., Kristensen, T. N., Sgrò, C. M., Weeks, A. R., Ørsted, M., & Hoffmann, A. A. (2022). Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(1), e2105076119. https://doi.org/10.1073/pnas.2105076119
Wilson, G. A., & Rannala, B. (2003). Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, *163*(3), 1177–1191.
Wold, J., Koepfli, K.-P., Galla, S. J., Eccles, D., Hogg, C. J., Le Lec, M. F., Guhlin, J., Santure, A. W., & Steeves, T. E. (2021). Expanding the conservation genomics toolbox: Incorporating structural variants to enhance genomic studies for species of conservation concern. *Molecular Ecology*, *30*(23), 5949–5965.
Yuan, X., Miller, D. J., Zhang, J., Herrington, D., & Wang, Y. (2012). An overview of population genetic data simulation. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, *19*(1), 42–54.

SUPPORTING INFORMATION

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

How to cite this article: Schiebelhut, L. M., Guillaume, A. S., Kuhn, A., Schweizer, R. M., Armstrong, E. E., Beaumont, M. A., Byrne, M., Cosart, T., Hand, B. K., Howard, L., Mussmann, S. M., Narum, S. R., Rasteiro, R., Rivera-Colón, A. G., Saarman, N., Sethuraman, A., Taylor, H. R., Thomas, G. W. C., Wellenreuther, M., & Luikart, G. (2024). Genomics and conservation: Guidance from training to analyses and applications. *Molecular Ecology Resources*, *24*, e13893. https://doi.org/10.1111/1755-0998.13893