Title

Embracing the power of genomics to inform evolutionary significant units

Authors

Jayna C. Bergman[1*], Rebecca S. Taylor[1], Micheline Manseau[1]

[1]*Landscape Science and Technology, Environment and Climate Change Canada, Colonel By Drive, Ottawa, ON K1S 5B6, Canada*

[*]Corresponding author. Email: jaynabergman@gmail.com

Keywords

conservation units, conservation genomics, genetics, discreteness, adaptive variation, phylogeographic history

Acknowledgements

This research was supported by Environment and Climate Change Canada (ECCC).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors have nothing to report.

Abstract

Appropriate identification of evolutionary significant units (ESUs) is essential for effective conservation planning. Genomic data has emerged as a key tool to inform ESU decisions due to the increased resolution of information, yet little is known about how genomic data is being used in practice to inform these decisions. To address this, we conducted a systematic literature review and found that genomic data is increasingly being used to inform ESUs globally and across plant and animal taxa. We identified a lack of consistency in the methods used to analyze the genomic data, although with reduced representation sequencing and whole genome sequencing improving the ability to infer evolutionary significance. Adaptive loci, a key advantage of genomic data, need to be interpreted with caution and simply identifying these loci may lead to the inflation of ESUs. Our review further revealed inconsistencies in how ESUs are defined and that conclusions do not always meet the criteria outlined in the definition provided. For instance, the amount of gene flow determined rarely influenced the number of ESUs recommended. Overall, 67% of the studies included in this review suggested an increase in the number of ESUs regardless of vast differences across studies. A unifying definition of what constitutes an ESU and a detailed workflow that can provide reliable support to assess the evolutionary significance of these units is still needed. The unparalleled depth of information obtained from genomic data, and the application of a proven workflow has the potential to improve the identification of ESUs globally.

Introduction

Identifying intraspecific units for conservation is a critical strategy for protecting diversity within species (Funk et al., 2012; Coates et al., 2018; Hohenlohe et al., 2020; Allendorf et al., 2022; Forester et al., 2022; Turbek et al., 2023; Schmidt et al., 2024). One widely used concept is Evolutionary Significant Units (ESU), which are biologically informed and often legally recognized groups of individuals below the species level. ESUs are broadly recognized as being (1) geographically or genetically discrete and (2) representative of an important component of a species evolutionary history (see reviews by Crandall et al., 2000; Fraser and Bernatchez, 2001; Funk et al., 2012; Turbek et al., 2023). The identification of ESUs can have direct implications for the legal status and management of populations (Coates et al., 2018), and inappropriate identification can waste limited conservation resources (Forester et al., 2022). Therefore, carefully considering how ESUs are described is essential for effective conservation planning.

However, no single definition of ESUs is universally used and the general concept has evolved over time due to ongoing debate about which criteria should be emphasized (e.g. Ryder, 1986; Waples, 1991; Moritz, 1994; Fraser and Bernatchez, 2001; Funk et al., 2012). Ryder (1986) first introduced the term ESU for conservation purposes and described them as populations that “possess genetic attributes significant for the present and future generations of the species in question”. Other proposed definitions have highlighted different criteria. For example, Waples (1991) emphasized the importance of reproductive isolation, while Fraser and Bernatchez (2001) focused on requiring highly restrictive gene flow. Moritz (1994) argued the use of specific genetic markers and determined that ESUs should be “reciprocally monophyletic for mitochondrial DNA alleles and show significant divergence of allele frequencies at nuclear loci”. More recently, Funk et al., (2012) proposed that ESUs should be “historically isolated from

each other and likely have important adaptive differences among them", integrating both neutral and adaptive variation for this process. In contrast, Crandall et al., (2000) questioned the utility of the ESU concept altogether, and suggested "that the terminology [ESU] is abandoned and replaced with a more holistic concept of species". In addition to these biological definitions, there are also legally defined units designated by governing bodies to implement policy for legal protection and conservation status that overlaps conceptually with ESUs (Coates et al., 2018; Lehnert et al., 2023). For example, Designatable Units (DUs) are delineated by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC, 2023), and Distinct Population Segments (DPS) are delineated by the U.S. Endangered Species Act (US Fish and Wildlife Services, 1996). These legal units tend to use separate definitions with different criteria to inform the designations. Despite these varying perspectives, ESUs are widely used in conservation, underscoring the need for greater clarity and standardization in how they are identified.

Genomic data are emerging as a key tool in the complex decision-making process of identifying ESUs (Funk et al., 2012; Barbosa et al., 2018; Waples and Lindsey, 2018; Funk et al., 2019; Turbek et al., 2023). Traditional genetic markers, such as mitochondrial DNA (mtDNA) and microsatellites, have been commonly used to identify ESUs (e.g. Moritz, 1994; Kanthaswamy et al., 2006; Jensen et al., 2014). However, these markers typically target only a small number of loci (e.g. <25), providing limited genetic information that largely reflects neutral evolutionary processes (Funk et al., 2019; Barbosa et al., 2020; Ghildiyal et al., 2023; Chinna et al., 2025). In contrast, genomic data generated through high throughput sequencing can sample thousands to millions of loci across the entire genome, offering much higher resolution on both neutral

89 and adaptive variation (Coates et al., 2018; Ghildiyal et al., 2023; see Box 1 for an overview of
90 genomic sequencing approaches). This increased analytical power facilitates a wide range of
91 downstream analyses (Schmidt et al., 2024) that can be useful for identifying ESUs.

92 The unparalleled volume of information provided by genomic data has significantly
93 advanced the ability to accurately identify ESUs (Funk et al., 2012; Coates et al., 2018). Genome-
94 wide patterns of variation can reveal phylogenetic relationships with high precision (Barbosa et
95 al., 2020), demographic models can infer past evolutionary processes (Schmidt et al., 2024), and
96 putatively adaptive loci showing pronounced signatures of selection can be identified (Funk et
97 al., 2012; Barbosa et al., 2018). However, some types of genomic data may miss highly localized
98 signatures of adaptation (e.g. RADseq; Lou et al., 2021), that could lead to overlooking genetic
99 variation essential to a population's persistence. A more comprehensive approach is whole-
100 genome sequencing (WGS), which provides deep insights into a species' evolutionary history
101 (e.g. Ellegren, 2014; Jensen et al., 2021; Taylor et al., 2024b). When high coverage data are
102 available, WGS enables accurate genotype calling and the detection of regions under selection
103 (Fuentes-Pardo and Ruzzante, 2017; Lou et al., 2021). Despite the clear advantages of genomic
104 data in identifying ESUs (Funk et al., 2012; Ghildiyal et al., 2023; Schmidt et al., 2024), and the
105 availability of several published workflows to help guide analyses (e.g. Funk et al., 2012; Turbek
106 et al., 2023; Lehnert et al., 2023), it remains unclear how genomic data are being used in
107 practice to support these decisions.

108

109 **Aims of the study**

The aim of this study was to determine how genomic data is being used when directly applied to identify ESUs across taxa. To achieve this, we reviewed published literature that uses genomic data to inform ESU identification. Specifically, we aimed to identify the variation in data reporting among conservation genomic studies, the methods used to analyze the genomic data, and the definitions and criteria used to define what an ESU is. We further aimed to uncover how these decisions directly impacted the conclusions drawn on the number of ESUs suggested. Subsequently, we examined how WGS are used in ESU identification to understand if the increased amount of information available with WGS, when compared to other genomics methods, is being utilized to uncover additional variation. Finally, we offer recommendations for considerations to make when using genomic information to identify ESUs.

Literature Review

We conducted a literature search in Web of Science for studies that use genomic data to help inform the identification of ESUs. We included studies that were published until January 30, 2024, by searching the title, abstract, or keywords with the search string (genom* OR (WGS OR "whole genome" OR *seq OR GBS OR "genotyping-by-sequencing" OR SNP* OR "single nucleotide polymorphism*" OR "target capture" OR "target DNA" OR ddRAD OR "reduced representation sequencing" OR "next-generation sequencing" OR NGS OR transcript* OR *omics OR "adaptive loci" OR "adaptive locus") AND (“conservation unit*” OR “evolutionary significant unit*” OR "evolutionarily significant unit*" OR “adaptive unit*” OR “distinct population segment*” OR "designatable unit*"). We supplemented our search with articles from Scopus

using the same set of search terms. Resulting studies from both databases were combined and duplicate studies were removed. The titles and abstracts of the retrieved articles were first manually screened with the assistance of ASReview (ASReview LAB developers, 2023), followed by full-text review for those that met the initial criteria. We excluded studies that did not use genomic data in their analyses (e.g., irrelevant methods such as using only mtDNA, microsatellites, or cpDNA) and those that did not make direct conclusions about the number of ESUs (e.g., irrelevant topics such as generating a reference genome, assigning individuals to existing ESUs, or species delineation without a direct link to ESUs). For synthesis papers that did not include analysis of original data, we reviewed their cited references to identify additional relevant studies. We included studies on all vertebrate, invertebrate, and plant species and did not apply any geographic restrictions. In total, our search terms identified 415 articles of potential relevance (Table S1). Of these, 154 met our inclusion criteria and were evaluated in the literature review (Table S2).

For each study we recorded whether a written definition of what is considered an ESU was provided, and which reference was used to support that definition. We then extracted information on sampling details and the genomic datasets used from both the main text and supplementary materials as needed. If a study included separate datasets for more than one species, each species was treated independently and assigned a separate row in the data table. This resulted in 161 data entries in the final dataset used (Table S2).

The methods used to analyze the genomic data were recorded as either done or not done, with additional columns noting the specific programs used. All analytical methods were manually classified into two broad groups: methods assessing “discreteness” and those

153 assessing “evolutionary significance”. Methods classified under “discreteness” included
154 clustering and ancestry-based approaches, F_{st}/AMOVA, assessments of gene flow, tests for
155 isolation-by-distance, and network analyses. Methods under “evolutionary significance”
156 included phylogenetic analyses to uncover deeper evolutionary relationships, assessments of
157 demographic history, tests for separate glacial refugia, and estimating divergence times.
158 Because we were particularly interested in identifying trends related to adaptive loci, we
159 separately categorized studies that identified them. For these studies, we noted whether
160 adaptive loci were detected using F_{ST} outlier analyses or genotype-environment association
161 (GEA) approaches. We also recorded whether these loci were used as a dataset to explore
162 population structure or were simply identified as being present. In addition, for studies that
163 assessed gene flow, we determined if gene flow was directly calculated (e.g. migration rates, D-
164 statistic, f-statistics) or inferred (e.g. Admixture plots, presence of hybrids, amount of genetic
165 differentiation). Based on the study’s conclusions, we categorized the amount of gene flow
166 determined into three groups: lots of gene flow, limited gene flow, and no gene flow.

167 Overall, we aimed to understand the impact of genomic data on the suggested number
168 of ESUs. To do this, we extracted each study’s final conclusions regarding the number of ESUs
169 proposed and whether this represented an increase or decrease compared to the previously
170 identified number of ESUs. Additional details are provided in the supplementary material on
171 how data for each column was extracted from the studies (Table S3). All plots were generated in
172 R (R Core Team, 2024) using *ggplot2* (Wickham, 2016) and *patchwork* (Pedersen, 2024).

173

Results and Discussion

Our review showed that genomic data is used to inform the identification of ESUs across taxonomic groups around the globe and this use is steadily increasing. All classes of animals were represented in the collected studies (Fig. 1A), with fish having the most representation (n=42) and amphibians having the least (n=9). We also found that all continents, excluding Antarctica, were present in the dataset (Fig. 1B). However, there was a geographical bias in the studies with 42% done on species in North America, and of those studies 61% focused on populations within the United States (Fig. S1). Comparatively, some continents were severely underrepresented. For example, only 2% of the studies were done on species in Africa and 5% on species in South America (Fig. 1B). Plants were well represented in the collected studies (n=28), but a geographic bias was identified with 64% of these studies done on species in Asia (Table S2). Additionally, we identified a rapid increase in the use of genomic data to identify ESUs starting in 2020 (Fig. 1C). This increased use of genomic data across the world and across taxa highlights the importance of understanding how this data is produced, analyzed and interpreted to inform ESUs globally.

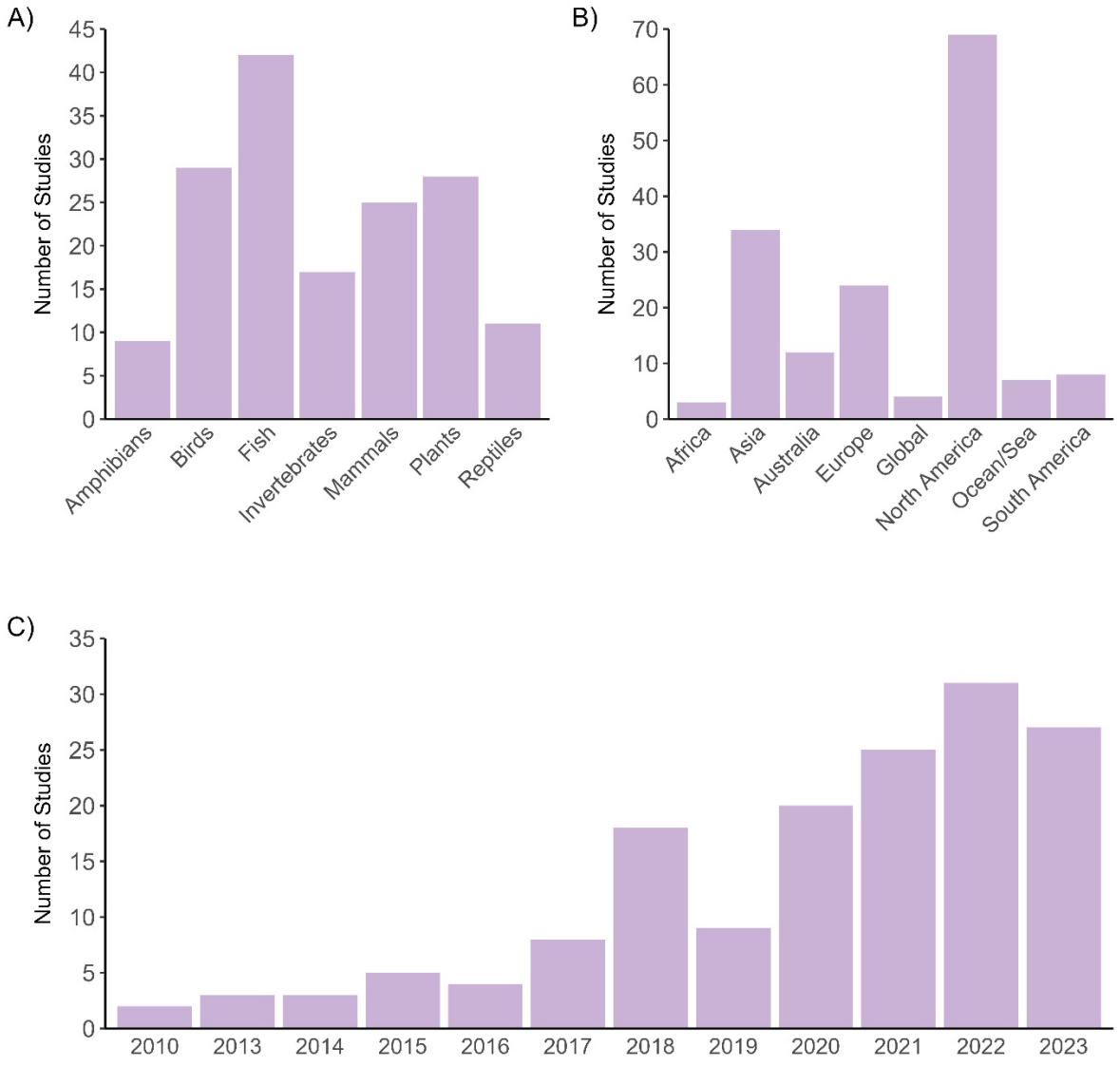


Figure 1. (A) The taxonomic groups represented by the 161 data entries included in the review, (B) the continents that the included studies had collected data from, and (C) the year that the study was published in. Our literature search ended in January 2024, and six studies are included in our literature review from that year. These studies are used in all other analyses and conclusions, but the year 2024 is excluded from panel C because it is not representative of the number of studies published in the entire year.

196

197 ***Data quality and Sequencing methods***

198 We found a lack of consistent reporting on key details that are needed to assess the
199 quality of genomic data used for identifying ESU. In particular, information on filtering decisions
200 and their effects on the dataset is critical for evaluating the reliability of data (O'Leary et al.,
201 2018). For example, read depth (also referred to as depth of coverage) is influenced by filtering
202 decisions and can affect the accuracy of genotype calls, which in turn impacts downstream
203 analyses (Hemstrom et al., 2024). Strikingly, only 40% of the studies in this review reported
204 either the average depth of coverage or the range of coverage values across samples (Table S2).
205 Of the remaining studies, 32% reported the minimum coverage filter that was used to call
206 genotypes, while 28% did not report any information to indicate coverage (Table S2). We
207 strongly advocate for all studies to report depth of coverage, in addition to the settings used to
208 filter genomic datasets.

209 We also observed inconsistencies in how geographic sampling information was reported.
210 Some studies did not clearly indicate the number of geographic sampling locations, or the
211 number of individuals sampled per location. These details are important, as the number of
212 geographic sampling locations is strongly associated with the ability to detect patterns of
213 isolation by distance (Hillis and Chambers, 2020), calculate reliable summary statics (e.g. F_{st}
214 values; Holsinger and Weir, 2009), and identify candidate loci for adaptation (Aguirre-Liguori et
215 al., 2020), all of which can be used to inform ESU decisions. Sampling information was not
216 reported in nine of the studies included in our review (Table S2). While full sampling tables were

often provided in the supplementary material, several studies did not summarize this information in the main text, reducing transparency and making it harder to assess potential biases or sources of error in the dataset.

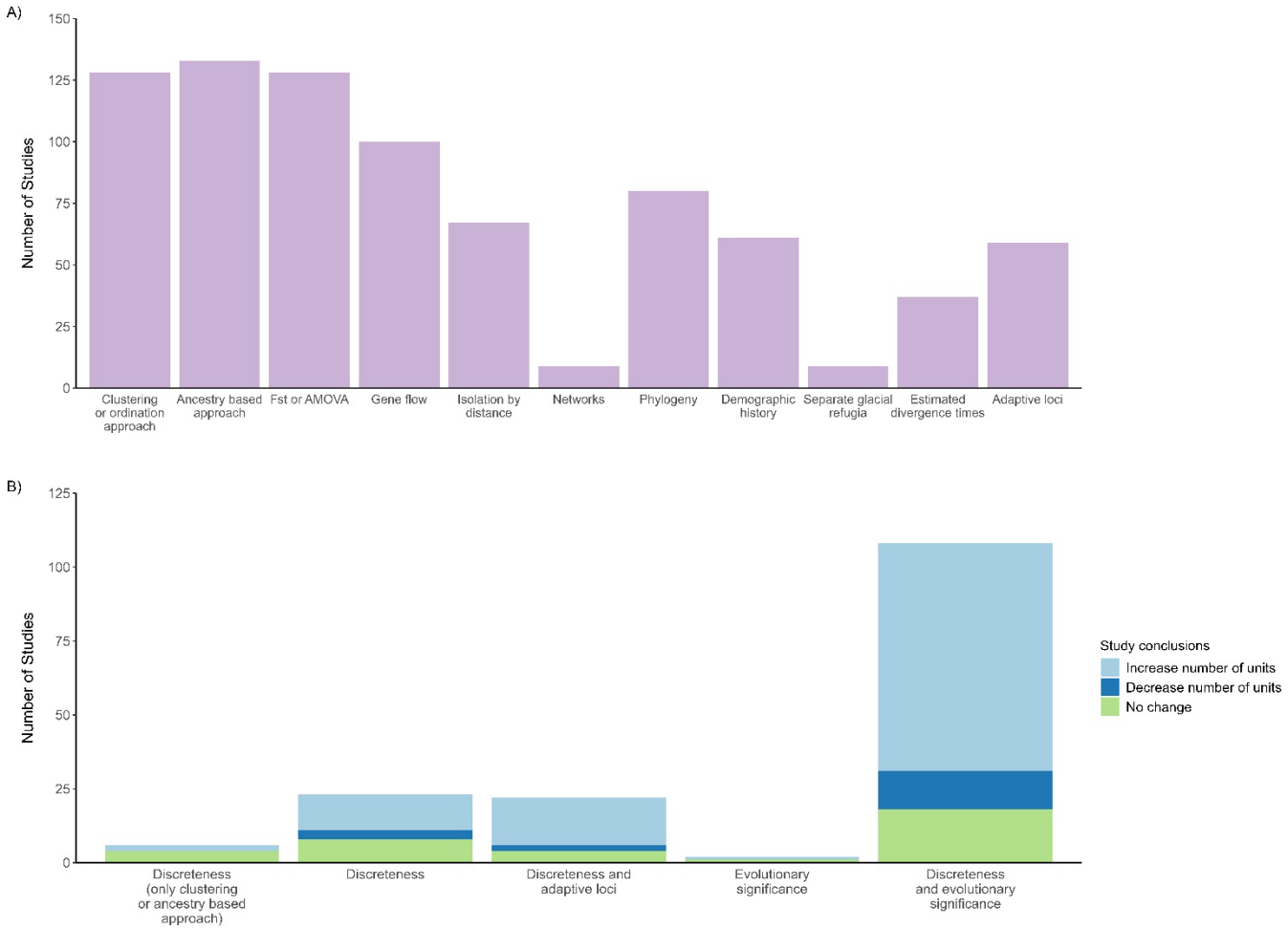
We grouped the sequencing methods used in the studies included in our review into six broad categories (see Box 1; Fig. S2). Reduced representation sequencing was the most commonly used approach with 103 studies (RRS; inclusive of restriction-site associated sequencing: RAD-seq, double digest restriction-site associated sequencing: ddRAD-seq, genotype by sequencing: GBS, and multiplexed inter-simple sequence repeats: MIG-seq; Fig. S2). The next most common methods were single nucleotide polymorphism (SNP)-based genotyping (e.g. SNP arrays, SNP assays, SNP chips, and SNP panels) and whole genome sequencing (WGS), both used in 22 studies (Fig. S2). The remaining sequencing approaches included targeted gene sequencing (e.g. major histocompatibility complex [MHC] genes and exon sequencing; n=10), transcriptomics (n=3), and pooled sequencing (pool-seq; n=1). Notably, most of these less commonly used methods focus on functional genomic regions and are therefore less suited for analyses to identify population structure or evolutionary history. Furthermore, pool-seq is often confounded with sequencing errors and individual genotypes are lost during the pooling process (Fuentes-Pardo and Ruzzante, 2017). Overall, RRS was used nearly five times more often than any other sequencing approach, but SNP-based genotyping and WGS were also common methods for generating genomic data to inform ESU identification.

Methods used to analyze genomic data

238 Among the various analyses available for genomic data, certain methods were more
239 common across studies (Fig. 2A). For example, 128 studies used a clustering or ordination
240 approach (e.g. PCA, DAPC) and 133 studies used an ancestry-based approach (e.g. Admixture,
241 Structure; Fig. 2A). Both these types of analyses are useful to explore population structure and
242 are recommended as starting points by recent workflows to determine ESUs (see Lehnert et al.,
243 2023; Turbek et al., 2023). However, not as many studies assessed the demographic history
244 (n=61), estimated divergence times (n=37), or identified if groups were in separate glacial
245 refugia (n=9; Fig. 2A). This is surprising, as assessments of demographic history, particularly
246 those using estimates of effective population size (Ne), have recently been highlighted as
247 important metrics for tracking genetic diversity (e.g., Hoban et al., 2022; Waples, 2024; Waples,
248 2025). Although tracking genetic diversity through Ne is not directly linked to the identification
249 of ESUs, and we do not recommend using Ne alone to define ESUs (Taylor et al., 2024a), it can
250 serve as an informative metric to help understand the evolutionary history of populations
251 (Chhina et al., 2025).

252 When we categorized methods into assessing for either “discreteness” or “evolutionary
253 significance” (see methods) to inform ESU identification, we found the majority of studies used
254 methods that could assess both (n=108; Fig. 2B). In contrast, 29 studies assessed only
255 discreteness and of these studies, six relied solely on clustering analyses and/or ancestry-based
256 methods (Fig. 2B). Interestingly, 9% of the studies recommended increasing the number of ESUs
257 even though they assessed only discreteness (Fig. 2B). This raises concerns, as determining
258 population structure using PCA and admixture-based analyses alone may produce unreliable
259 results and inaccurately reflect a population’s evolutionary history (Lawson et al., 2018; Elhaik,

2022), leading to the misidentification of ESUs. For example, inbred individuals can appear as distinct clusters on a PCA, which may be misinterpreted as evidence of discrete genetic groups (Taylor et al., 2024b). We recommend using multiple analyses to thoroughly explore population structure in ways that assess both discreteness and evolutionary significance when identifying ESUs. Clear and detailed examples of how to use several methods to provide strong support for ESU identification, including analyses addressing both discreteness and evolutionary significance, are provided in Box 2.



268 **Figure 2**. Out of the 161 data entries included in our literature review: (A) The number of
269 studies using different genomic analysis methods to determine how many ESUs should
270 be suggested. (B) The number of studies that applied methods testing for genetic
271 discreteness and/or evolutionary significance to inform ESU identification. Methods
272 classified under "discreteness" included clustering and ancestry-based approaches,
273 F_{st}/AMOVA, assessments of gene flow, tests for isolation-by-distance, and network
274 analyses. Methods under "evolutionary significance" included phylogenetic analyses to
275 uncover deeper evolutionary relationships, assessments of demographic history, tests
276 for separate glacial refugia, and divergence time estimation. Bar colors indicate the
277 proportion of studies in each category that recommended an increase, decrease, or no
278 change in the number of ESUs.

279

280 To understand whether the sequencing type influenced the choice of analyses, we
281 compared the methods used to analyze the genomic data across studies employing SNP-based
282 genotyping, RRS, and WGS. SNP-based studies accounted for all the studies that solely used
283 clustering and/or ancestry-based approaches, and they had the lowest proportion of studies to
284 assess both discreteness and evolutionary significance compared to RRS and WGS (Fig. 3). There
285 was a large range in the number of SNPs included in these datasets (25 to 95,043 SNPs; Table
286 S2), but 55% of these studies had fewer than 400 SNPs possibly resulting in a lack of resolution
287 required to carry out analyses needed to uncover deeper divergence across the genome. SNP-
288 based genotyping often leave large portions of the genome unsampled (Dokan et al., 2021).
289 They rely on targeted approaches that require prior knowledge of where variation is occurring

290 (Chhina et al., 2025), limiting the ability to detect cryptic population structure or assess
291 evolutionary significance, both of which are critical for informing ESU delineation. Despite these
292 limitations, the largest proportion of studies using SNP-based genotyping still recommended
293 increasing the number of ESUs.

294 In contrast, studies using RRS and WGS tended to use similar and more comprehensive
295 analyses to identify ESUs (Fig. 3; Fig. S3). In both cases, most studies assessed both discreteness
296 and evolutionary significance and likewise the majority suggested increasing the number of
297 ESUs (Fig. 3). However, WGS studies consistently included assessments of evolutionary
298 significance or identified adaptive loci, whereas a subset of RRS studies assessed only
299 discreteness (Fig. 3). This may be due to an advantage of WGS, which provides a more
300 comprehensive view of a species' evolutionary history by sequencing the entire genome rather
301 than sampling portions of it (Fuentes-Pardo and Ruzzante, 2017). This advantage may be further
302 highlighted, as a higher proportion of WGS studies estimated demographic histories and
303 divergence times compared to those using RRS or SNP-based genotyping (Fig. S3). We observed
304 a rising trend in the use of WGS, which is expected to continue in the coming years (Fig. S4). Our
305 review included studies published up to January 2024, and of the six studies from one month in
306 2024, half used WGS (Fig. S4). Altogether this suggests that WGS is becoming increasingly
307 common, and that the greater genomic resolution may facilitate the use of analyses that
308 provide stronger support for ESU identification in future studies.

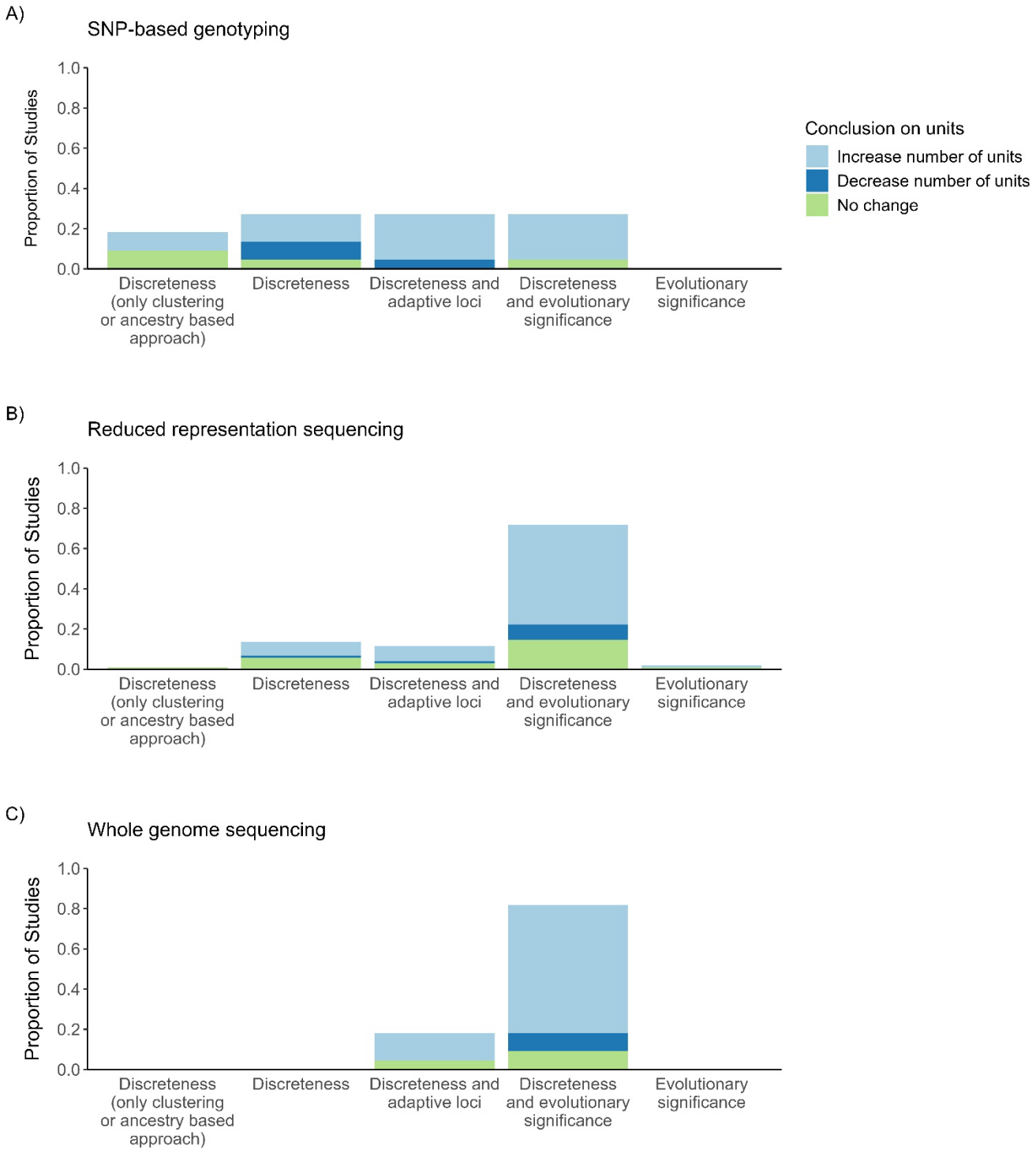


Figure 3. The proportion of studies that analyzed the genomic data with methods that tested for genetic discreteness and/or evolutionary significance, separated into different

sequencing methods of: (A) SNP-based genotyping (inclusive of SNP arrays, SNP assays, and SNP chips; n=22) (B) Reduced representation sequencing (inclusive of restriction-site associated sequencing: RAD-seq, double digest restriction-site associated sequencing: ddRAD-seq, Genotype by sequencing: GBS, and Multiplexed Inter-simple sequence repeat genotype by sequencing: MIG-seq; n = 103), and (C) Whole genome sequencing (n=22). Methods classified under "discreteness" included clustering and ancestry-based approaches, F_{st}/AMOVA, assessments of gene flow, tests for isolation-by-distance, and network analyses. Methods under "evolutionary significance" included phylogenetic analyses to uncover deeper evolutionary relationships, assessments of demographic history, tests for separate glacial refugia, and estimated divergence times. The filled color of the bars show the proportion of studies in each of these categories that suggested either increasing, decreasing, or no change to occur to the number of ESUs.

Gene flow

Gene flow is widely recognized as a critical evolutionary force that shapes genetic structure within and among species (Tigano and Friesen, 2016; Aitken et al., 2024; Chhina et al., 2025). It is typically defined as the movement of genetic information from one genetic group to another that often follows the migration of individuals but is not limited to just migration events (Funk et al., 2012; Tigano and Friesen, 2016). Several definitions of ESUs suggest that there should be a lack of gene flow or reproductive isolation between groups (i.e. no gene flow) as criteria for defining distinct evolutionary units (e.g. Waples, 1991; Fraser and Bernatchez, 2001;

333 COSEWIC, 2023). These perspectives may be based on the concern that gene flow could erode
334 genetic differences, potentially resulting in the loss of local adaptation (Lenormand, 2002). This,
335 in turn, could reduce the number of ESUs by diminishing the unique genetic diversity that
336 defines them. Alternatively, it is increasingly recognized that gene flow can occur between
337 populations adapted to different local conditions, and that local adaptation can be maintained
338 despite high levels of gene flow (Tigano and Friesen, 2016; Chhina et al., 2025). However,
339 whether gene flow increases or decreases local adaptation is largely population dependent
340 (Aitken et al., 2024). Other perspectives on ESUs have pointed out that it may be important to
341 balance the maintenance of gene flow to preserve adaptive differences (Crandall et al., 2000).
342 Thus, these different views offer no unified approach and leave substantial room for subjective
343 interpretation on how to consider gene flow when identifying ESUs.

344 We assessed the uncertainty surrounding gene flow by analyzing the methods that were
345 used to determine gene flow by the studies included in our review. We categorized assessments
346 into directly calculating gene flow (e.g. migration rates, D-statistics, f-statistics) and inferring
347 gene flow (e.g. admixture plots, amount of genetic differentiation, presence of hybrids). Results
348 showed that 15% of studies directly calculated gene flow, while 47% inferred it, and 35% did not
349 assess gene flow (Fig. 4A). Notably, the majority of studies recommended increasing the
350 number of ESUs, regardless of how gene flow was determined or if it was assessed at all (Fig.
351 4A). Studies that calculated gene flow never recommended decreasing the number of ESUs,
352 whereas those inferring gene flow or not assessing it had a portion of studies suggesting a
353 decrease (14% and 11%, respectively; Fig. 4A).

We then examined whether the amount of gene flow found influenced the suggested number of ESUs. Based on the language used in the conclusions of the reviewed studies, we categorized gene flow into three levels: lots of gene flow, limited gene flow, and no gene flow. These categories were somewhat arbitrary, as there was no consistent threshold that could be applied across the various methods that were used to assess gene flow. Surprisingly, the amount of gene flow reported did not appear to influence whether a study suggested increasing or decreasing the number of ESUs (Fig. 4B). Indeed, even among studies that described “lots of gene flow,” the largest proportion still recommended increasing the number of ESUs (Fig. 4B). We also found that whether gene flow was calculated or inferred did not affect the level of gene flow reported (Fig. S5). Overall, these results were unexpected, due to the several definitions of ESUs that emphasize the importance of limited gene flow or reproductive isolation between groups (e.g. Waples, 1991; Fraser and Bernatchez, 2001; COSEWIC, 2023). However, we suggest that some level of gene flow should be allowed between two ESUs, especially if this is the natural state of the populations. Gene flow among populations can maintain and even increase genetic diversity (Aitken and Whitlock, 2013; Aitken et al., 2024; Chhina et al., 2025). In such cases, conservation management should aim to maintain this gene flow, even if populations are considered as separate ESUs (Chhina et al., 2025). Our results suggest that this perspective may be shared by several other individuals.

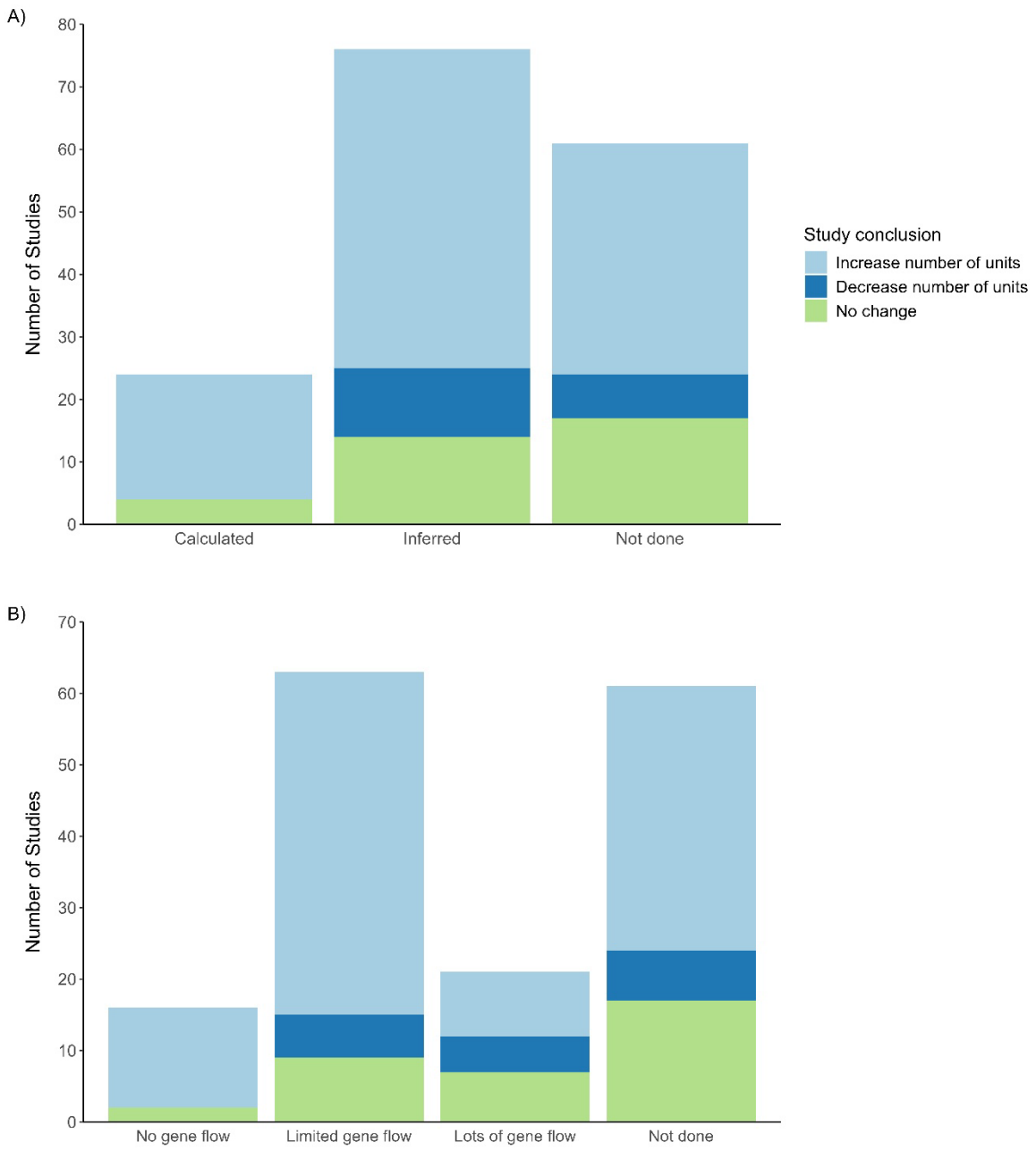


Figure 4. (A) The number of studies that assessed gene flow when determining how many ESUs to recognize, either by directly calculating gene flow (e.g. migration rates, D-statistics, f-statistics) or inferring gene flow (e.g. admixture plots, amount of genetic differentiation, presence of hybrids), as well as the number of studies that did not assess

377 gene flow. (B) Among studies that assessed gene flow, the author's conclusion regarding
378 the level of gene flow observed between the proposed ESUs. In both panels, bar colors
379 indicate the proportion of studies in each category that suggested increasing,
380 decreasing, or no change in the number of ESUs.

381

382 ***Adaptive loci***

383 A significant advantage of using genomic data over mtDNA or microsatellites to identify
384 ESUs is the ability to determine adaptive regions of the genome through the detection of
385 adaptive (or putatively adaptive) loci (Funk et al., 2012; Barbosa et al., 2018; Waples and
386 Lindsey, 2018). Two common approaches that are often used to identify adaptive loci are F_{st}
387 outlier analyses, and genotype-environment association (GEA) approaches. F_{st} outlier analyses
388 (e.g. BayScan, Lositan) detect loci that fall outside the range of the expected distribution of F_{st}
389 values (Funk et al., 2012). These loci are often assumed to be under divergent selection and
390 used as proxies for local adaptation (Waples and Lindley, 2018) and therefore thought to be
391 representative of evolutionary significance. Alternatively, GEA approaches (e.g. RDA, LFMM)
392 identify associations between allele frequencies and environmental variables, to reveal loci that
393 are associated with specific environmental conditions (Hoban et al., 2016; Forester et al., 2018).
394 Workflows to help guide ESU designation often incorporate the identification of adaptive loci
395 (e.g. Funk et al., 2012; Lehnert et al., 2023; Turbek et al., 2023). However, interpreting adaptive
396 loci needs to be handled cautiously (Hoban et al., 2016; Allendorf, 2017). F_{st} outliers can result
397 from processes other than local adaptation (Bierne et al., 2013), such as a species demographic

398 history (Whitlock and Lotterhos, 2015) or genetic drift (Lotterhos and Whitlock, 2014). Similarly,
399 loci identified through GEA methods often include false positives and may not reliably indicate
400 local adaptation (Nadeau et al., 2016). In some cases, GEA analyses can be paired with common
401 garden experiments or genome-wide association studies (GWAS) to help verify which loci are
402 associated with climate adaptation (e.g. Faske et al., 2021; Aitken et al., 2024).

403 We determined the number of studies in our review that identified adaptive loci and
404 examined whether they used F_{st} outlier analyses or GEA analyses, as these were the two
405 primary approaches used. In total, 59 studies identified adaptive loci, and of those studies, 22
406 used GEA approaches and 37 used F_{st} outlier analyses (Fig. 5A). Across both methods, a greater
407 proportion of studies recommended increasing the number of ESUs (GEA: 77%, F_{st} outlier: 65%;
408 Fig. 5A). However, even among studies that did not identify adaptive loci, the largest proportion
409 still suggested an increase in ESUs (66%; Fig. 5A). Interestingly, studies that used GEA
410 approaches only recommended either increasing the number of ESUs or maintaining the
411 current number, none suggested a decrease. These results may reflect greater confidence in the
412 adaptive potential of loci identified through GEA methods, as associations with environmental
413 conditions can strengthen the inference that these loci are truly adaptive (Hoban et al., 2016).

414 We also examined whether the way in which adaptive loci were used in a study influenced
415 conclusions about the number of ESUs suggested. We categorized studies with adaptive loci as
416 those that simply identified the presence of adaptive loci, and those that incorporated these
417 loci into downstream analyses that assessed population structure based on adaptive variation.
418 Among studies using GEA approaches, there was roughly an even split between those that only
419 identified loci and those that used them as a dataset (n=10 and n=12, respectively; Fig. 5B).

However, studies that additionally used adaptive loci as a dataset were more likely to suggest no change in the number of ESUs (Fig. 5B). In contrast, for F_{st} outliers there were more studies that only identified adaptive loci (n=25) than used them in downstream analyses (n=10) and these studies were more likely to suggest an increase in the number of ESUs (Fig. 5C). This pattern may suggest that identifying adaptive loci alone, without additionally examining their population structure, can inflate the number of ESUs identified, possibly due to the high rate of false positives associated with these methods (Bierne et al., 2013; Nadeau et al., 2016). However, further analyzing adaptive loci for population structure may mitigate this effect. Ultimately, while loci identified through F_{st} outliers and GEA approaches are a useful proxy for local adaptation, their interpretation should be approached with caution. Additional analyses may be useful to confirm a group's evolutionary significance beyond identifying adaptive loci, as these loci do not always provide reliable evidence for adaptive potential (Bierne et al., 2013; Chhina et al., 2024) and thus for identifying ESUs.

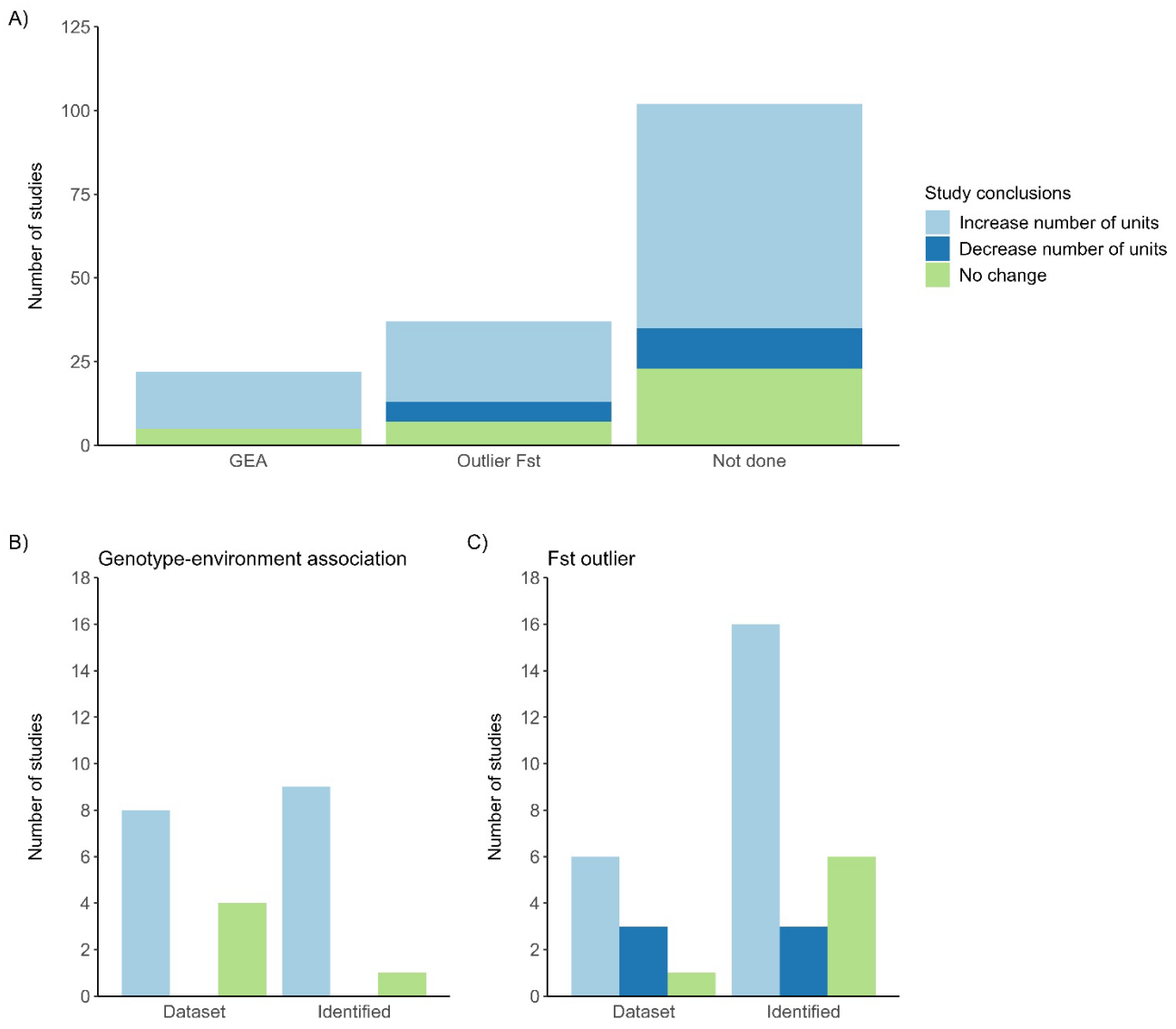


Figure 5. (A) The number of studies that determined adaptive loci using genotype-environment association (GEA; n=22) or F_{ST} outlier analyses (n=37), as well as those that did not identify adaptive loci (n=102). Bar colors indicate the number of studies in each category that recommended an increase, decrease, or no change in the number of ESUs. Among studies that identified adaptive loci, they were further categorized as "dataset" if the loci were used in downstream analyses, or as "identified" if the loci were only reported but not

used in further analyses. (B) Results for studies using GEA approaches; (C) Results for studies using F_{ST} outlier analyses.

Definitions and criteria used to define ESUs

One significant challenge we found in our review was the inconsistency in how ESUs are defined. A few definitions were more commonly cited in the empirical studies we reviewed, and we highlight these definitions in Table 1. Although ESUs are broadly recognized as being both genetically discrete and evolutionary significant, multiple definitions are available, each offering a different perspective on how to identify them (Table 1; also see reviews by Coates et al., 2018; Turbek et al., 2023). However, studies often reference multiple definitions simultaneously, resulting in merging the separate criteria of these definitions. In addition, ESU definitions often have subjective language, and this has been suggested to limit their practical application to real world situations (Fraser and Bernatchez, 2001). This ambiguity appears to be a persistent issue in the current studies we reviewed.

Table 1. Definitions of Evolutionary Significant Units (ESUs) from previous publications. This is not an exhaustive list but rather shows the reference and the corresponding definitions that were cited most frequently by the 154 publications included in this review.

Reference	Definition	Number of times cited*
Ryder, 1986	"a subset of the more inclusive entity species, which possess genetic attributes significant for the present and future generations of the species in question"	16
Waples, 1991	"A population must satisfy two criteria to be considered an ESU: (1) it must be substantially reproductively isolated from other conspecific	8

	population units, and (2) it must represent an important component in the evolutionary legacy of the species."	
Moritz, 1994	"ESUs should be (1) reciprocally monophyletic for mtDNA alleles and (2) show significant divergence of allele frequencies at nuclear loci."	29
Crandall et al., 2000	"…suggest that the terminology [ESU] is abandoned and replaced with a more holistic concept of species, consisting of populations with varying levels of gene flow evolving through drift and natural selection"	8
Fraser and Bernatchez, 2001	"A lineage demonstrating highly restricted gene flow from other such lineages within the higher organizational level (lineage) of the species"	12
Funk et al., 2012	"Major, intraspecific units that (1) have been historically isolated from each other and (2) that likely have important adaptive differences among them"	11
Committee on the Status of Endangered Wildlife in Canada (Designatable Units: COSEWIC, 2023)	"A unit below the level of a taxonomic species if it has attributes that make it both (1) discrete, and (2) evolutionary significant, where discrete means that there is currently very little transmission of heritable information from other such units and evolutionarily significant means that the unit harbors heritable adaptive traits or an evolutionary history not found elsewhere in Canada."	7
Other citations	NA	20
Not reported	NA	103

*This column sums to more than the total number of studies due to several studies citing multiple references

We found that only 36% of studies explicitly stated how they defined an ESU, including providing at least one reference for the given definition (Fig. 6A; Table 1). This raises concerns because of the multiple, and sometimes conflicting definitions of what constitutes an ESU. Interestingly, we found that studies that provided a definition were just as likely as those that did not to suggest an increase in the number of recognized ESUs (69% and 66%, respectively), and this outcome was more common than suggesting a decrease in the number of ESUs or no

462 change (Fig. 6A). Additionally, in some cases, the definition of ESU cited appeared to influence
463 whether a study recommended increasing or decreasing the number of ESUs (Fig. 6B). For
464 example, studies citing Waples (1991), more often suggested no change in the number of ESUs,
465 a pattern not observed with any other reference (Fig. 6B). In contrast, studies citing Moritz
466 (1994) and Funk et al. (2012) were proportionally more likely than those citing any other
467 definition to recommend an increase in the number of ESUs, rather than a decrease or no
468 change. Providing a clear definition for how an ESU is being considered is essential not only for
469 transparency, but also for having consistent and meaningful comparisons across studies.
470 Without providing a definition, it is unclear what evidence is being used to separate populations
471 into distinct ESUs, or alternatively, justify grouping them together.

472 Among the studies that provided a definition, the few key sources that were commonly
473 referenced (Table 1; Fig. 6B) differed in the main criteria required to determine separate ESUs.
474 We found that conclusions about increasing or decreasing the number of ESUs were often made
475 independently of the criteria outlined in the cited definition. For instance, 38% of studies citing
476 Waples (1991) and 33% citing Fraser and Bernatchez (2001) did not assess gene flow (Fig. S6),
477 despite the role of gene flow being specifically stated in those definitions. Similarly, 66% of
478 studies referencing Moritz (1994) did not include analyses with mtDNA (Table S2), even though
479 the definition clearly states the requirement of mtDNA to show reciprocal monophyletic groups.
480 This suggests that conclusions are often made independently of the criteria applied in an ESU
481 definition, raising concerns about the consistency and comparability of ESU identification across
482 studies, taxa, and regions.

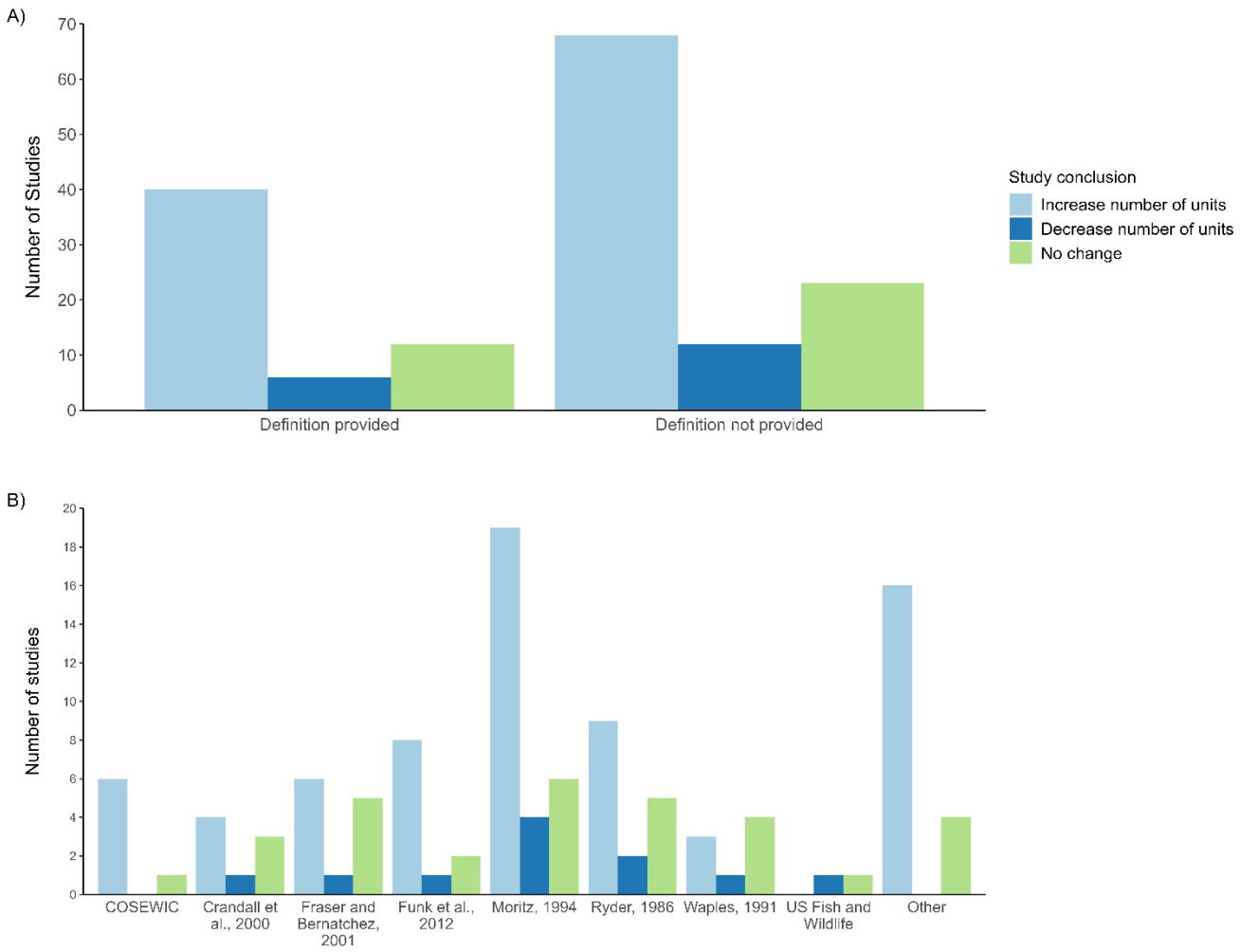


Figure 6. (A) The number of studies that provided a written definition of how they recognized an evolutionary significant unit, accompanied by a cited reference for that definition (n=58), and the number of studies that did not provide a definition (n=103). Bar colors indicate the study conclusion on whether they recommended an increase, decrease, or no change in the number of evolutionary significant units. (B) The most frequently cited references to define ESUs among the 161 data entries in our review, and the number of studies that suggested an increase, decrease, or no change in the number of ESUs.

Changes in the number and delineation of ESUs

The studies in our review more frequently suggested increasing the number of ESUs than maintaining the current number or decreasing them (Fig. 7). Specifically, 67% of the studies suggested an increase in ESUs, 22% suggested no change, and only 11% proposed a decrease (Fig. 7). We found this trend to be consistent regardless of sequencing type, the downstream methods used to analyze the genomic data, or the way ESU were defined. Thus, these findings indicate that genomic data is generally leading to the identification of more ESUs within species, despite the many differences across studies.

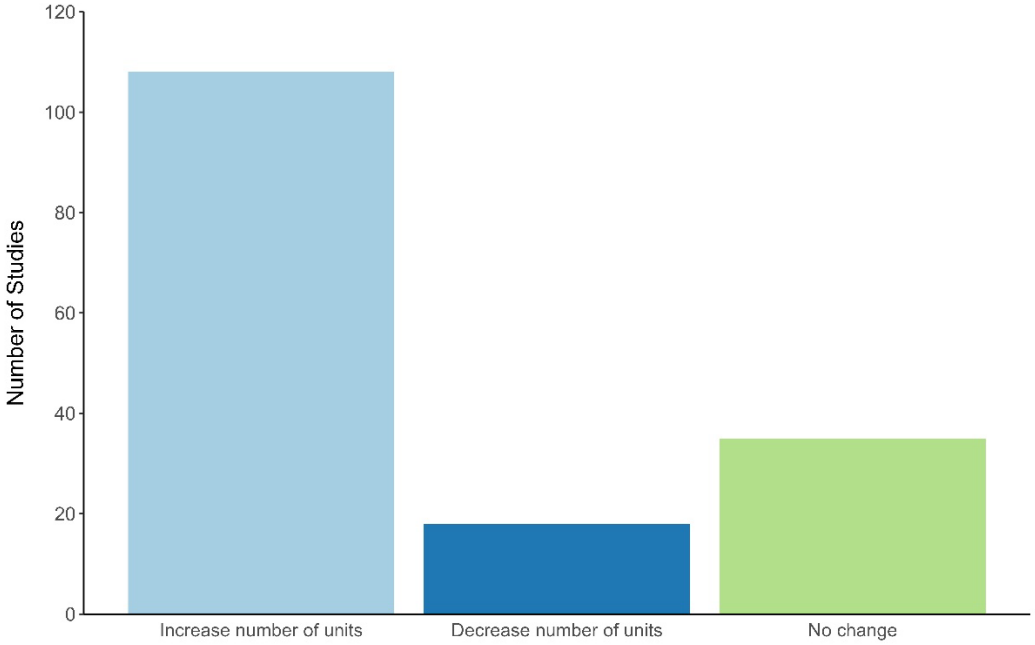


Figure 7. The overall conclusions on the suggested number of evolutionary significant units (increase, decrease, or no change the number of conservation units) across the 161 data entries included in this literature review.

An important finding is the inconsistency in how studies determined whether to increase the number of ESUs. One major source of this inconsistency may be due to the lack of a common threshold to determine if groups are different enough to be considered as separate ESUs. For example, F_{ST} values are one of the most widely used statistics in population genetics, but this statistic may give highly different estimates of genetic differences based on the number of individuals analyzed, the number of geographic sampling locations, and the type of genetic marker assessed (Holsinger and Weir, 2009). Implementing one common threshold may be extremely difficult, and a discussion on how to do this is long overdue (e.g. also recommended by Funk et al., 2012). Second, there is no unifying workflow available to help make these decisions in a reliable and consistent way. Notably, there are a couple of recent workflows that have been published to help guide the analyses for identifying separate conservation units. For example, Turbek et al., (2023) has developed a decision tree for species exhibiting patterns of Isolation-by-distance, and Lehnert et al., (2023) aimed to delineate Designatable Units in Canada (legal units often considered to have similar concepts to ESUs). Most likely due to their recent publication dates at the time of this literature review, these decision trees have yet to be widely used (but see Miller et al., 2024 which follows the approach of Lehnert et al., 2023). However, these workflows do not fully address the broader concept of evolutionary significance past identifying adaptive loci. Thus, a more detailed workflow to help guide studies to assess genetic discreteness and fully evaluate evolutionary significance is still needed to provide much-needed consistency for identifying ESUs.

526 **Conclusion and Recommendations**

527 The vast amount of information that genomic data provides greatly improves our ability
528 to inform the identification of ESUs (Funk et al., 2012; Coates et al., 2018; Funk et al., 2019;
529 Turbek et al., 2023). Genomic data facilitates a deeper understanding of the genetic differences
530 within a species and can reliably uncover their evolutionary history, both of which are essential
531 to accurately identify intraspecific ESUs. In this study, we conducted a comprehensive literature
532 review to explore how genomic data has been used to inform ESUs globally. We found that key
533 details about the genomic data being used, such as sequencing depth, are frequently not
534 reported. Additionally, we found that there is a lack of consistency in the methods used to
535 analyze genomic data or how ESUs are defined. Despite these inconsistencies, the majority of
536 studies recommended increasing the number of ESUs. Here we identify some gaps in how
537 genomic data is currently being used to identify ESUs and highlight some recommendations that
538 could be used to improve future studies.

539 We strongly recommend clear and consistent reporting of genomic data, study details, and
540 how ESUs are being defined. We found that an overarching problem when using genomic data
541 to inform ESUs is the lack of a unifying definition and an accompanying unifying framework, and
542 that future work is needed to agree upon a single definition of what constitutes an ESU which
543 can be used across all taxa. This lack of unification is causing inconsistencies in the analyses
544 being used, making it challenging to have comparable conclusions across different studies.
545 While some studies relied solely on measures of genetic discreteness (e.g. PCA), we recommend
546 that additional lines of evidence, such as demographic reconstructions, be incorporated to
547 assess evolutionary significance (see Box 2 for examples). Genetic structure identified through

548 analyses based on clustering and ancestry-based frequencies can be highly influenced by recent
549 isolation due to habitat fragmentation, genetic drift, or recent inbreeding, all of which can be
550 due to anthropogenic causes (Schlaepfer et al., 2018), and not a result of significant differences
551 evolved over time. Similarly, analyses used to detect adaptive loci, such as F_{ST} outliers, must be
552 interpreted cautiously, as they can be influenced by patterns other than adaptation (Bierne et
553 al., 2013; Lotterhos and Whitlock, 2014; Nadeau et al., 2016). Importantly, we suggest that
554 identifying adaptive loci is not always necessary if divergent evolutionary histories can be
555 demonstrated through other analyses.

556 We advocate for identifying ESUs in a manner which preserves the evolutionary legacy of a
557 species. We therefore recommend the use of phylogenomic and demographic history
558 reconstruction as important analyses to use when identifying ESUs. Ideally, phylogeny
559 reconstruction using genome-wide data should be used to provide a powerful and reliable
560 inference of evolutionary history. With the increase of accessibility to WGS (Fuentes-Pardo and
561 Ruzzante, 2017; Lou et al., 2021), this is a practical solution and could provide the much-needed
562 consistency for identifying ESUs across different taxa. Beyond improving phylogenetic
563 resolution, by using whole genomes we can improve the ability to test for local adaptation
564 beyond F_{st} outliers and GEA approaches. More sophisticated analyses for this need to be
565 developed and incorporated into workflows for ESUs enabling studies to fully utilize the large
566 amount of information that can be gained when using whole genomes. However, we recognize
567 that WGS is still not feasible for all systems, particularly those with large or complex genomes.
568 In such cases, RRS has been shown to provide reliable results to infer the demographic history
569 of a species (Elleouet and Aitken, 2018) and could be used as an alternative.

570 Despite these suggestions for improvement, significant progress has been made over the
571 past decade to better utilize the power of genomics to inform the identification of ESUs
572 (Flanagan, et al., 2018; Funk et al., 2019; Allendorf et al., 2022). Genomic data, including WGS,
573 is more accessible than ever before (Lou et al., 2021; Schmidt et al., 2024) and can provide
574 valuable insights into genome-wide variation inclusive of neutral, adaptive, and deleterious
575 variation that may be critical for current and future conservation management (Flanagan et al.,
576 2018; Chhina et al., 2024; Schmidt et al., 2024). Moving forward, it is essential that both
577 researchers and practitioners not only understand how to assess genomic data but also how to
578 apply it effectively for identifying ESUs. By fully embracing the power of genomics with more
579 standardization and thoughtful interpretation, we can improve species conservation and ensure
580 that decisions regarding the identification of ESUs are made accurately and informatively for
581 evolutionary relevance.

582 **Literature cited**

583 Aguirre-Liguori JA, Luna-Sánchez JA, Gasca-Pineda J, Eguiarte LE (2020). Evaluation of the
584 Minimum Sampling Design for Population Genomic and Microsatellite Studies: An Analysis
585 Based on Wild Maize. *Front Genet* **11**: 870.

586 Aitken SN, Jordan R, Tumas HR (2024). Conserving Evolutionary Potential: Combining Landscape
587 Genomics with Established Methods to Inform Plant Conservation. *Annual Review of Plant*
588 *Biology* **75**: 707–736.

589 Aitken SN, Whitlock MC (2013). Assisted Gene Flow to Facilitate Local Adaptation to Climate
590 Change. *Annu Rev Ecol Evol Syst* **44**: 367–388.

591 Allendorf FW (2017). Genetics and the conservation of natural populations: allozymes to
592 genomes. *Mol Ecol* **26**: 420–430.

593 Allendorf FW, Funk C, Aitken SN, Byrne M, Luikart G (2022). Conservation and the
594 genomics of populations. 3rd ed. Oxford (UK): Oxford University Press.

595 Barbosa S, Hendricks SA, Funk WC, Rajora OP, Hohenlohe PA (2020). Wildlife Population
596 Genomics: Applications and Approaches. In: Hohenlohe PA, Rajora OP (eds) *Population*
597 *Genomics: Wildlife*, Population Genomics. Springer International Publishing: Cham, pp 3–
598 59.

599 Barbosa S, Mestre F, White TA, Paupério J, Alves PC, Searle JB (2018). Integrative approaches to
600 guide conservation decisions: Using genomics to define conservation units and functional
601 corridors. *Mol Ecol* **27**: 3452–3465.

602 Bierne N, Roze D, Welch JJ (2013). Pervasive selection or is it…? why are F_{ST} outliers sometimes
603 so frequent? *Molecular Ecology* **22**: 2061–2064.

604 Chambers EA, Hillis DM (2020). The Multispecies Coalescent Over-Splits Species in the Case of
605 Geographically Widespread Taxa (R Glor, Ed.). *Systematic Biology* **69**: 184–193.

606 Chhina AK, Abhari N, Mooers A, Lewthwaite JMM (2024). Linking the spatial and genomic
607 structure of adaptive potential for conservation management: a review. *Genome* **67**: 403–
608 423.

609 Coates DJ, Byrne M, Moritz C (2018). Genetic Diversity and Conservation Units: Dealing With
610 the Species-Population Continuum in the Age of Genomics. *Front Ecol Evol* **6**: 165.

611 Committee on the Status of Endangered Wildlife in Canada, November (2023). COSEWIC
612 guidelines for recognizing designatable units. COSEWIC.
613 https://cosewic.ca/index.php/en/reports/preparing-status-reports/guidelines-recognizing-
614 designatable-units.html

615 Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000). Considering evolutionary
616 processes in conservation biology. *Trends in Ecology & Evolution* **15**: 290–295.

617 Dokan K, Kawamura S, Teshima KM (2021). Effects of single nucleotide polymorphism
618 ascertainment on population structure inferences (Y Kim, Ed.). *G3*
619 *Genes|Genomes|Genetics* **11**: jkab128.

620 Elhaik E (2022). Principal Component Analyses (PCA)-based findings in population genetic
621 studies are highly biased and must be reevaluated. *Sci Rep* **12**: 14683.

622 Ellegren H (2014). Genome sequencing and population genomics in non-model organisms.
623 *Trends in Ecology & Evolution* **29**: 51–63.

624 Elleouet JS, Aitken SN (2018). Exploring Approximate Bayesian Computation for inferring recent
625 demographic history with genomic markers in nonmodel species. *Molecular Ecology*
626 *Resources* **18**: 525–540.

627 Faske TM, Agneray AC, Jahner JP, Sheta LM, Leger EA, Parchman TL (2021). Genomic and
628 common garden approaches yield complementary results for quantifying environmental
629 drivers of local adaptation in rubber rabbitbrush, a foundational Great Basin shrub.
630 *Evolutionary Applications* **14**: 2881–2900.

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

634 Forester BR, Lasky JR, Wagner HH, Urban DL (2018). Comparing methods for detecting
635 multilocus adaptation with multivariate genotype–environment associations. *Molecular*
636 *Ecology* **27**: 2215–2233.

637 Forester BR, Murphy MA, Mellison C, Petersen J, Pilliod DS, Van Horne R, *et al.* (2022).
638 Genomics-informed delineation of conservation units in a desert amphibian. *Molecular*
639 *Ecology*: 1–21.

640 Fraser DJ, Bernatchez L (2001). Adaptive evolutionary conservation: towards a unified concept
641 for defining conservation units. *Mol Ecol* **10**: 2741–2752.

642 Fuentes-Pardo A, Ruzzante D (2017). Whole-genome sequencing approaches for conservation
643 biology: Advantages, limitations and practical recommendations. *Molecular Ecology* **26**:
644 5369–5406.

645 Funk WC, Forester BR, Converse SJ, Darst C, Morey S (2019). Improving conservation policy with
646 genomics: a guide to integrating adaptive potential into U.S. Endangered Species Act
647 decisions for conservation practitioners and geneticists. *Conserv Genet* **20**: 115–134.

648 Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012). Harnessing genomics for delineating
649 conservation units. *Trends in Ecology & Evolution* **27**: 489–496.

650 Gallego-García N, Vargas-Ramírez M, Shaffer HB (2023). The importance of cryptic diversity in
651 the conservation of wide-ranging species: The red-footed tortoise Chelonoidis carbonarius
652 in Colombia. *Molecular Ecology* **32**: 4531–4545.

Garrigan D, Hedrick PW (2003). Perspective: Detecting adaptive molecular polymorphism: Lessions from the MHC. *Evolution* **57**: 1707–1722.

Ghildiyal K, Nayak SS, Rajawat D, Sharma A, Chhotaray S, Bhushan B, *et al.* (2023). Genomic insights into the conservation of wild and domestic animal diversity: A review. *Gene* **886**: 147719.

Hemstrom W, Grummer JA, Luikart G, Christie MR (2024). Next-generation data filtering in the genomics era. *Nat Rev Genet*.

Hoban S, Archer FI, Bertola LD, Bragg JG, Breed MF, Bruford MW, *et al.* (2022). Global genetic diversity status and trends: towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition. *Biological Reviews* **97**: 1511–1538.

Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, *et al.* (2016). Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *The American Naturalist* **188**: 379–397.

Hohenlohe PA, Funk WC, Rajora OP (2021). Population genomics for wildlife conservation and management. *Mol Ecol* **30**: 62–82.

Holsinger KE, Weir BS (2009). Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nat Rev Genet* **10**: 639–650.

Jensen EL, Gaughran SJ, Garrick RC, Russello MA, Caccone A (2021). Demographic history and patterns of molecular evolution from whole genome sequencing in the radiation of Galapagos giant tortoises. *Molecular Ecology* **30**: 6325–6339.

Jensen EL, Govindarajulu P, Russello MA (2014). When the shoe doesn't fit: applying conservation unit concepts to western painted turtles at their northern periphery. *Conserv Genet* **15**: 261–274.

Jorge RF, Lima AP, Stow AJ (2022). Selection and localised genetic structure in the threatened Manauense Harlequin Frog (Bufonidae: Atelopus manauensis). *Conservation Genetics* **23**: 559–574.

Kanthaswamy S, Kurushima JD, Smith DG (2006). Inferring Pongo conservation units: a perspective based on microsatellite and mitochondrial DNA analyses. *Primates* **47**: 310–321.

Lawson DJ, Van Dorp L, Falush D (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat Commun* **9**: 3258.

Lehnert SJ, Bradbury IR, Wringe BF, Van Wyngaarden M, Bentzen P (2023). Multifaceted framework for defining conservation units: An example from Atlantic salmon (*Salmo salar*) in Canada. *Evolutionary Applications* **16**: 1568–1585.

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

Lotterhos KE, Whitlock MC (2014). Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology* **23**: 2178–2192.

Lou RN, Jacobs A, Wilder AP, Therkildsen NO (2021). A beginners guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology* **30**: 5966–5993.

Miller CV, Bossu CM, Sarraco JF, Toews DPL, Rushing CS, Roberto-Charron A, *et al.* (2024). Genomics-informed conservation units reveal spatial variation in climate vulnerability in a migratory bird. *Molecular Ecology* **33**.

Moritz C (1994). Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology & Evolution* **9**: 373–375.

Nadeau S, Godbout J, Lamothe M, Gros-Louis M-C, Isabel N, Ritland K (2015). Contrasting patterns of genetic diversity across the ranges of *Pinus monticola* and *P. strobus* : A comparison between eastern and western North American postglacial colonization histories. *American Journal of Botany* **102**: 1342–1355.

O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS (2018). These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology* **27**: 3193–3206.

Pederson T (2024). _patchwork: The Composer of Plots_. R package version 1.3.0, <https://CRAN.R-project.org/package=patchwork>.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species (L Orlando, Ed.). *PLoS ONE* **7**: e37135.

R Core Team (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ryder OA (1986). Species conservation and systematics: the dilemma of subspecies. Trends in Ecology and Evolution **1**: 9-10.

Schlaepfer DR, Braschler B, Rusterholz H, Baur B (2018). Genetic effects of anthropogenic habitat fragmentation on remnant animal and plant populations: a meta-analysis. *Ecosphere* **9**: e02488.

Schmidt TL, Thia JA, Hoffmann AA (2024). How Can Genomics Help or Hinder Wildlife Conservation? *Annu Rev Anim Biosci* **12**: annurev-animal-021022-051810.

Suyama Y, Matsuki Y (2015). MIG-seq: an effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. *Sci Rep* **5**: 16963.

Taylor RS, Manseau M, Keobouasone S, Liu P, Mastromonaco G, Solmundson K, *et al.* (2024b). High genetic load without purging in caribou, a diverse species at risk. *Current Biology* **34**: 1234-1246.e7.

Taylor RS, Manseau M, Wilson PJ (2024a). Delineating conservation units should be independent of effective population size. *Trends in Ecology & Evolution*: S0169534723003270.

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

Turbek SP, Funk WC, Ruegg KC (2023). Where to draw the line? Expanding the delineation of conservation units to highly mobile taxa (W Murphy, Ed.). *Journal of Heredity* **114**: 300–311.

U.S. Fish and Wildlife Service & National Marine Fisheries Service (1996). Policy regarding the recognition of distinct vertebrate population segments under the Endangered Species Act. *Federal Register*, **61**: 4722–4725. https://www.federalregister.gov/documents/1996/02/07/96-2639/policy-regarding-the-recognition-of-distinct-vertebrate-population-segments-under-the-endangered

Van de Schoot R, de Bruin J, Schram R, Zahedi P, de Boer J, Weijdema F et al. (2021). ASReview: open-source machine learning software for systematic reviews. *Nat Mach Intell* 3: 125–126.

Waples RS (1991). Pacific Salmon, Oncorhynchus spp., and the Definition of 'Species' Under the Endangered Species Act. *Marine Fisheries Review*.

Waples RS (2024). The N_e / N ratio in applied conservation. *Evolutionary Applications* **17**: e13695.

Waples RS (2025). The Idiot's Guide to Effective Population Size. *Molecular Ecology*: e17670.

Waples RS, Lindley ST (2018). Genomics and conservation units: The genetic basis of adult migration timing in Pacific salmonids. *Evolutionary Applications* **11**: 1518–1526.

Wickham (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

Whitlock MC, Lotterhos KE (2015). Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of F_{ST}. *The American Naturalist* **186**: S24–S36.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors have nothing to report.

759 **Box 1**

760 **Genomic data** are DNA sequences or molecular markers that encompass a large and
761 representative set of regions across the genome to provide high resolution information
762 (Schmidt et al., 2024). Genomic data can allow for the identification of the structure and/or
763 function of genetic variation across the genome.

764 Genomic data can be generated by the following sequencing methods which are grouped into
765 several categories:

766 1. **SNP-based genotyping** – Refers to techniques that have a set of predefined and
767 targeted probes with specific position of single nucleotide polymorphisms (SNPs). These
768 SNPs target variants that are based on relevance to specific individuals/populations,
769 adaptive loci, or particular traits. We have grouped SNP chips, SNP array, SNP assays and
770 SNP panels together because these are all types of targeted SNP-based genotyping.
771 2. **Reduced representation sequencing (RRS)** – Is an approach to generate genome-wide
772 sequencing data by sequencing only digested fragments of the genome. Fragments are
773 from across the entire genome, but only a fraction of the genome is sequenced. The
774 following sequencing types are RRS methods:
775 - ***Double digest restriction-site associated DNA sequencing (ddRAD)*** – Uses two
776 restriction enzymes to ensure more uniform fragments of the whole genome by
777 having two cut sites (Peterson et al., 2012). We have considered **DArTseq** methods as
778 ddRAD because they both deploy double enzyme digestion across the genome.
779 DArTseq uses a different system for commercial services.

780 - ***Genotype-by-sequencing (GBS)*** – Single enzyme digestion of the genome that uses
781 PCR for fragment size selection.
782 - ***Restriction-site associated DNA sequencing (RADseq)*** – uses one enzyme to perform
783 a single digestion of the whole genome and then randomly interrupts the digested
784 fragments.
785 - ***Multiplexed Inter-simple sequence repeat genotype by sequencing (MIG-seq***) – This
786 method does not use restriction-enzyme digestion but rather PCR-based markers
787 that amplify a large number of genome-wide regions without prior genetic
788 information (Suyama and Matsuki, 2015).
789 3. **Whole genome sequencing** – sequencing the entire genome of an organism to get a
790 comprehensive view of the genome. There are two types of whole genome sequencing
791 that differ by the depth of the sequencing:
792 - ***High coverage whole genome sequencing*** – Whole genome sequencing that
793 typically has coverage high enough to allow genotypes to be accurately called.
794 - ***Low coverage whole genome sequencing*** – Whole genome sequencing resulting in a
795 small number of reads covering the genomic loci (e.g. <10X coverage). Genotype
796 likelihood-based methods are used with low coverage data.
797 4. **Targeted gene sequencing** – This sequencing approach focuses on specific genes or
798 genomic regions of interest, rather than sequencing the entire genome.
799 - ***Major histocompatibility complex genes (MHC)*** – These functional genes play a vital
800 role in an organism's immune system and are highly polymorphic. MHC diversity can

801 be used as an indicator of a population's adaptive ability (Garrigan and Hedrick,
802 2003).

803 - ***Exome sequencing*** – Exons are protein-coding regions of the genome. Exome
804 sequencing is a targeted sequencing approach to capture variation within these
805 protein-coding regions.

806 5. **Pooled sequencing (pool-seq)** – Multiple individual DNA samples are pooled together
807 and then sequenced in a single run. Pooled sequencing typically has a high depth of
808 coverage (e.g. >50X) and can be cost effective (Fuentes-Pardo and Ruzzante, 2021).
809 However individual genotypes are undistinguishable due to the process of pooling the
810 DNA (Fuentes-Pardo and Ruzzante, 2021).

811 6. **Transcriptomics** – Is the characterization of RNA transcripts through RNA sequencing
812 that is typically used to understand gene expression and regulation.

813 **Box 2**

814 **Case study 1:** The importance of cryptic diversity in the conservation of wide-ranging species:
815 The red-footed tortoise (*Chelonoidis carbonarius*) in Columbia
816 *Gallego-Garcia et al., 2023*

817

818 The red-footed tortoise (*Chelonoidis carbonarius*) is a widespread, vulnerable species found in
819 Central America, South America, and several Caribbean islands. Gallego-Garcia et al. (2023)
820 sequenced 233 individuals from 25 sites across the species' range in Colombia using triple
821 enzyme restriction-site associated DNA sequencing (3RAD), generating a dataset of 30,327
822 SNPs. This summary focuses on the analyses conducted across all individuals to identify
823 Evolutionarily Significant Units (ESUs), although additional analyses were performed within each
824 identified ESU to determine additional smaller management units. To assess genetic
825 discreteness, the authors used Principal Component Analysis (PCA), STRUCTURE, and F_{ST} values.
826 All three methods supported the presence of distinct eastern and western lineages (i.e. two
827 ESUs). The lack of admixture in the STRUCTURE plot and high F_{ST} values (~0.42) were used to
828 infer no gene flow between the two groups. To assess evolutionary significance, a phylogenetic
829 tree was generated and revealed a basal split between the lineages, supporting deep
830 divergence and a lack of ancestral gene flow. Additionally, effective population sizes and
831 divergence times were estimated for both lineages to further support their evolutionary
832 significance. Finally, putatively adaptive loci were identified through genotype-environment
833 associations (GEA) analysis using a redundancy analysis and an F_{ST} outlier analysis using

834 PCAdapt. Overall, Gallego-Garcia et al., (2023) suggest that the eastern and western lineages
835 should be recognized as two separate ESUs based on analyses that support both genetic
836 discreteness and evolutionary significance, thereby increasing the number of ESUs for the
837 species.

838

839 **Case study 2:** Selection and localised genetic structure in the threatened Manauense Harlequin
840 frog (Bufonidae: *Atelopus manauensis*)
841 *Jorge et al., 2022*

842

843 The Manauense Harlequin frog (*Atelopus manauensis*) is a threaten, stream-breeding
844 amphibian with a narrow distribution in the Amazon. Jorge et al., (2022) sequenced 94
845 individual males from 21 streams using a modified protocol of double digest restriction-site
846 associated DNA sequencing (ddRAD), generating a dataset of 3,859 SNPs. To identify putatively
847 adaptive loci, the authors used F_{ST} outlier analyses (PCAdapt and ARLEQUIN) and a genotype-
848 environment association (GEA) approach using Latent Factor Mixed Models (LFMM). The GEA
849 analysis detected significant associations between genetic variation and environmental
850 variables, suggesting local adaptation. A neutral dataset of 3,121 SNPs was then created by
851 removing the adaptive loci identified through the three analyses, as well as loci not in Hardy-
852 Weinberg equilibrium. Genetic discreteness was assessed using the neutral dataset through
853 Discriminant Analysis of Principal Components (DAPC), sparse non-negative matrix factorization
854 (SNMf), STRUCTURE, and F_{ST} values. All four analyses identified six discrete genetic clusters, with

855 high F_{ST} values (0.3-0.5) between them. To assess evolutionary significance, in addition to
856 identifying the adaptive loci, the authors constructed a phylogenetic tree, and estimated
857 divergence times, which ranged from approximately 40,000 to 2,000 years ago depending on
858 the cluster. Finally, a stairway plot indicated population declines around 100,000 years ago, and
859 current estimated effective population sizes per group ranged from 20 to 100 individuals.
860 Overall, Jorge et al., (2022) suggest an increase in the number of ESUs, recognizing a total of six
861 separate ESUs that show both genetic discreteness and evolutionary significance.