



HAL
open science

Evolving spatial conservation prioritization with intraspecific genetic data

Marco Andrello, Cassidy D'aloia, Alicia Dalongeville, Marco A. Escalante,
Jimena Guerrero, Charles Perrier, Juan Pablo Torres-Florez, Amanda Xuereb,
Stéphanie Manel

► **To cite this version:**

Marco Andrello, Cassidy D'aloia, Alicia Dalongeville, Marco A. Escalante, Jimena Guerrero, et al..
Evolving spatial conservation prioritization with intraspecific genetic data. *Trends in Ecology &
Evolution*, 2022, 37 (6), pp.553-564. 10.1016/j.tree.2022.03.003 . hal-03823288

HAL Id: hal-03823288

<https://hal.umontpellier.fr/hal-03823288>

Submitted on 16 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Evolving spatial conservation prioritization with intraspecific genetic data

Andreollo Marco ^{1,*}, D'aloia Cassidy ², Dalongeville Alicia ³, Escalante Marco A. ⁴, Guerrero Jimena ⁵, Perrier Charles ⁶, Torres-Florez Juan Pablo ⁷, Xuereb Amanda ⁸, Manel Stéphanie ⁹

¹ Institute for the study of Anthropic impacts and Sustainability in the marine environment, National Research Council, CNR-IAS, Rome, Italy

² Department of Biology, University of Toronto Mississauga, Mississauga, ON, Canada

³ MARBEC, University of Montpellier, CNRS, Ifremer, IRD, Montpellier, France

⁴ Laboratory of Molecular Ecology, Institute of Animal Physiology and Genetics of the Czech Academy of Sciences, Liběchov, Czech Republic

⁵ Sociedad Científica de Investigación Transdisciplinaria y Especialización (SCITE), Calimaya, México

⁶ CBGP, INRAe, CIRAD, IRD, Montpellier SupAgro, University of Montpellier, Montpellier, France

⁷ Instituto Chico Mendes de Conservação da Biodiversidade, Centro Nacional de Pesquisa e Conservação de Mamíferos Aquáticos, Santos, Brazil

⁸ Département de Biologie, Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC, Canada

⁹ CEFE, Univ Montpellier, CNRS, EPHE-PSL University, IRD, Montpellier, France

* Corresponding author : Marco Andreollo, email address : marco.andreollo@cnr.it

Abstract :

Spatial conservation prioritization (SCP) is a planning framework used to identify new conservation areas on the basis of the spatial distribution of species, ecosystems, and their services to human societies. The ongoing accumulation of intraspecific genetic data on a variety of species offers a way to gain knowledge of intraspecific genetic diversity and to estimate several population characteristics useful in conservation, such as dispersal and population size. Here, we review how intraspecific genetic data have been integrated into SCP and highlight their potential for identifying conservation area networks that represent intraspecific genetic diversity comprehensively and that ensure the long-term persistence of biodiversity in the face of global change.

Highlights

► Conservation area networks on land and sea need to be expanded to meet the objectives of the post-2020 global biodiversity framework. ► Spatial conservation prioritization (SCP) is a rigorous framework to identify suitable areas for protection on the basis of scientific data. ► Integrating intraspecific genetic data in SCP can help identify networks of conservation areas that are more representative of biological diversity and likely better at ensuring its long-term persistence.

Keywords : adaptive genetic diversity ; biodiversity features ; evolutionarily significant units ; reserve design ; systematic conservation planning

44 Main

45 **The benefits and challenges of intraspecific genetic data for spatial conservation** 46 **prioritization**

47 Facing worldwide declines in biodiversity and nature’s contributions to people [1], the post-2020 global
48 biodiversity framework under discussion by the UN will prescribe to create “ecologically representative and
49 well-connected” networks of conservation areas (CAs) that cover 30% of marine, aquatic and terrestrial
50 habitats, and to ensure that 90% of within species genetic diversity is maintained by 2030 (see
51 www.cbd.int/conferences/post2020). To achieve these goals, spatial conservation prioritization (SCP) is an
52 effective framework to identify new CAs on the basis of the spatial distribution of conservation costs and
53 **biodiversity features** (see **Glossary**) such as species and ecosystems [2].

54 Over time, SCP has evolved to integrate increasingly complex aspects of biodiversity, such as connectivity,
55 ecosystem services and functional diversity [3–5]. Recently, attempts have been made to use intraspecific
56 genetic data to gain knowledge on several aspects of species’ biology that are critical for their conservation
57 (see also **Online Supplemental Information Table S1**). In particular, genetic data can provide information
58 on intraspecific genetic diversity, dispersal and population size [6,7]. The published studies listed in **Table**
59 **S1** show that such information can increase the **comprehensiveness** of CA networks and the long-term
60 persistence of biodiversity. However, the successful integration of intraspecific genetic data with other
61 types of data in SCP presents challenges. Here, we briefly review the available techniques to estimate
62 intraspecific genetic diversity, dispersal and population size from intraspecific genetic data and we discuss
63 how to best integrate them in SCP.

64 **Obtaining unbiased information from intraspecific genetic data**

65 *Intraspecific genetic diversity*

66 Species are not static in time and show phenotypic variation throughout their range. This intraspecific
67 diversity, which arises through the interplay of environmental and genetic variation, has consequences for
68 population viability, community and ecosystem functioning, and nature's contributions to people [8,9].

69 Intraspecific variation is an important asset that can allow species to persist in the face of rapid
70 environmental change, such as those expected from the outcomes of global climate change [10]. There is
71 evidence that intraspecific genetic diversity has declined in many wild species [11,12]; therefore, the post-
72 2020 biodiversity framework will commit to protecting intraspecific genetic diversity and CAs can be a
73 valuable tool to reach this objective [13].

74 Intraspecific genetic diversity can be partitioned into within-population diversity and between-population
75 diversity, analogous to partitioning species diversity into alpha and beta components [14]. Within-
76 population genetic diversity can be measured using metrics such as allelic richness and observed and
77 expected heterozygosity, while between-population genetic diversity can be represented by metrics of
78 genetic differentiation [15]. Genetic differentiation can be used to identify conservation units below the
79 species level, such as management units (MUs) and evolutionarily significant units (ESUs) [16]. The
80 maintenance of genetic differentiation between MUs implies significant demographic isolation or selection
81 against immigrants, which justifies considering them as distinct conservation units [16]. Conversely,
82 genetically homogenous sets of individuals cannot be considered as MUs given that the level of migration
83 that is sufficient for genetic homogeneity might not be sufficiently high to ensure **demographic**
84 **connectivity** [17,18].

85 ESUs are populations or groups of populations that have evolved independently and can be identified by
86 reconstructing phylogenetic trees within species [16,19]. ESUs are important conservation units because a

87 comprehensive view of biodiversity includes the full set of nested clades representing phylogenetic
88 relationships among organisms [20]. In addition, regions with maximum phylogenetic diversity for a given
89 taxon will also have the greatest trait diversity and thus potential to respond evolutionarily to future
90 environmental change [21]. Finally, focusing on ESUs instead of species can help conserve biodiversity
91 when taxonomy is uncertain [22]. Although the identification of ESUs has frequently relied on finding
92 monophyletic clades [19], the general agreement is that ESUs should not be designated solely on the basis
93 of genetic distinctiveness: ecological exchangeability and existence of genetic adaptations are among the
94 proposed criteria to define ESUs [16,23,24]. Furthermore, the steps and choice of methods involved in
95 reconstructing phylogenies can influence the inferred relationships among population units [25,26].

96 A further distinction can be made between **neutral genetic diversity** and **adaptive genetic diversity**
97 according to the effects of genetic variation on individual and population fitness [27]. However, the effect
98 of different alleles on the fitness of individuals and the viability of populations is seldom known, especially
99 for non-model organisms. Genotype-phenotype association studies aim to identify genes responsible for
100 phenotypic variation through correlative tests between variation in phenotypic traits and genetic variation
101 [28]. Such genes can be considered important for the viability of populations when the phenotypic traits
102 studied are of key importance for the persistence of populations and the identified genes have sufficiently
103 large phenotypic effects for their variation to significantly affect phenotypic variation [29]. A second set of
104 methods (**outlier tests** and **environmental association analyses**) investigate the signatures of selection to
105 detect candidate loci underlying local adaptation [30,31]. However, it is always difficult to distinguish the
106 signatures of positive selection from those of genetic drift [30] and, even when adaptive loci have been
107 identified with high confidence, the effects of their genetic diversity on population persistence usually
108 remain unknown [29]. Faced with these challenges, it is often difficult to partition neutral from adaptive
109 genetic diversity. One possibility is using genome-wide genetic variation as a proxy for the viability of
110 populations [32]. However, for some cases where genetic variation in phenotypic traits has been

111 quantified, neutral genetic variation has proven to be a poor predictor of adaptive genetic variance [33].
112 Furthermore, genomic techniques allow typing thousands of loci and if all these loci were included as
113 biodiversity features, they could lead to computationally prohibitive problems and redundant information.
114 As large genomic data sets accumulate [34], there is a need to consider how measures of intraspecific
115 genetic diversity can be used in SCP.

116 *Dispersal*

117 The post-2020 global framework emphasizes that biodiversity should be protected through “well-
118 connected systems” of CAs. The functioning of systems of CAs as well-connected networks depends
119 critically on the dispersal of organisms, which facilitates recolonization after catastrophic disturbances
120 (**demographic rescue**) and allows the spread of adaptive variants that increase the viability of local
121 populations facing environmental change (**genetic rescue**). In some species, dispersal can be studied using
122 telemetry methods, but these techniques are not practical for many animal and plant species that disperse
123 during life stages (such as larvae or seeds) when they are too small to be equipped with emitters. In these
124 cases, genetic techniques can be a useful alternative to estimate dispersal at the temporal scale of a few
125 generations in the past (**Box 1**). Four types of methods have been identified to estimate dispersal from
126 genetic data: assignment tests [35], parentage analysis [35–37], analysis of the pattern of isolation-by-
127 distance [37] and clinal analysis [38]. The results are estimates of dispersal probabilities between sites
128 (summarized in a **dispersal matrix**) and dispersal distances (summarized in a **dispersal kernel**). Each of
129 these methods has strengths and weaknesses (reviewed in [7,35]): for example, the accuracy of assignment
130 tests depends on the degree of genetic differentiation between populations, while parentage and clinal
131 analyses require intensive sampling or sequencing efforts [35] and cannot realistically be applied to a large
132 number of species occupying an area being considered for SCP. However, gaining direct dispersal data for a
133 small number of representative taxa could be useful to complement other, more feasible genetic
134 approaches, such as the analysis of isolation by distance [37].

135 *Census and effective population size*

136 Various statistical frameworks are available to estimate **census population size** N_c from samples of
137 individuals from natural populations typed with molecular markers [39]. These methods offer a valuable
138 alternative to direct observation for obtaining estimates of population density in species that are difficult to
139 observe and count, such as aquatic animals. For example, close-kin mark-recapture is an extension of
140 traditional mark-recapture approaches where each juvenile carries the “marks” of its parents within its
141 DNA [40]; using this method with a panel of 8,961 SNPs, Hillary *et al.* [41] estimated that N_c in the white
142 shark *Carcharodon carcharias* population in eastern Australia and New Zealand ranges between 2,500–
143 6,750 individuals. Intraspecific genetic data are also useful to estimate **effective population size** N_e , which
144 is related to the risk of **inbreeding depression** and loss of genetic diversity [42], through several statistical
145 frameworks applicable to a variety of life-histories [43]. Although uncertainty increases when the real N_e is
146 large, with appropriate sampling designs and sufficient numbers of genetic markers, genetic data can
147 provide precise and unbiased estimates of N_c and N_e , in some cases using the same dataset [44].
148 Temporally repeated sampling can provide estimates of population trends in time and thus help identify
149 declining populations [45].

150 **Integrating information obtained from intraspecific genetic data in spatial** 151 **conservation prioritization**

152 SCP can be treated as a mathematical problem using equations linking the spatial distribution of
153 biodiversity features and conservation costs [46] (**Box 2**). While there are different formulations of SCP
154 problems [46], almost all of them involve four parameters: the representation level r_{ij} of biodiversity
155 feature j in site i , the cost c_i of protecting site i , the spatial target T_j for biodiversity feature j , and the
156 adjacency cost cv_{ih} between site i and h . The general principle to integrating the estimates from
157 intraspecific genetic data is to link them explicitly to the parameters of SCP (**Figure 1**).

158 *Intraspecific genetic diversity*

159 There are various ways to integrate information on intraspecific genetic diversity into SCP. The simplest
160 approach is to use alleles as biodiversity features instead of (or in combination with) species (“AL” method
161 in **Table 1**), but it can be difficult to decide which and how many genetic markers and alleles to consider as
162 biodiversity features. Estimates of within-site diversity, such as allelic richness and heterozygosity, can be
163 used as biodiversity features (“GM”); however, defining a target of representation T_j for them is not
164 meaningful since T_j considers the total sum of a biodiversity feature across the planning area and such
165 genetic metrics are not additive across space. These metrics could be better integrated as cost layers, for
166 example by setting costs proportional to the inverse of allelic richness to select sites with high local genetic
167 diversity (“CS” method). Another option is to use site-specific metrics to rank sites according to the metric
168 of interest (e.g. sites with low and high allelic richness) and split the taxon (species or conservation unit)
169 occurrence layer into several distinct layers with specific representation targets (“ST” method).

170 Conservation units (MUs and ESUs), when present, can also be used directly as biodiversity features (“CU”
171 method). As an alternative to using ESUs, the branches of the phylogenetic tree can be used directly as
172 biodiversity features to assign higher priorities to older genetic lineages [22,47]. This approach may be
173 useful because conserving lineages separated by longer branches results in protecting larger amounts of
174 genetic diversity, compared to conserving more closely related lineages. In addition, using branches
175 ensures cost-effective protection as deeper branches representing shared evolutionary histories are only
176 accounted for once in the prioritization [48].

177 Some species do not have a discrete spatial genetic structure that permits researchers to unambiguously
178 identify conservation units. A solution to this problem is to use continuous measures of genetic distance
179 [49] in the ‘environmental diversity’ formulation of the SCP problem (“ED” method), used to identify a set
180 of conservation priority sites on the basis of continuous intraspecific variation (genetic or environmental
181 [50]).

182 Importantly, as genetic sampling is usually sparse, there will not be enough observations to measure the
183 spatial occurrence r_{ij} of alleles, conservation units or genetic metrics nor to measure costs c_i in all sites. This
184 requires a spatialization step to go from sampled points to values for all sites in the regular grid (planning
185 units) used in SCP. This can be done using several methods relying on sampled genetic data only (e.g.
186 inverse distance weighting) or making use of environmental variables (e.g. **ecological niche models**).
187 **Supplementary Table 1** indicates the methods used for each published paper that incorporates genetic
188 data in SCP. There is currently no comparison of the various methods to infer genetic data to cover
189 unsampled sites (see **Outstanding questions**).

190 *Dispersal*

191 Several methods are available to constrain the sites chosen for protection to be spatially contiguous
192 [46,51,52], such as introducing a boundary cost cv_{ih} for not protecting pairs of bordering sites (**Box 2**). This
193 formulation can easily accommodate the information of dispersal contained in a dispersal matrix, whose
194 elements d_{ij} give the probabilities of dispersal from site j to site i . Whether it is estimated from genetic data
195 or obtained through other methods, the dispersal matrix can be used to define the cv_{ih} parameter in the
196 SCP problem, which becomes a connectivity penalty cost paid when site i is chosen for protection and site h
197 is not [53]. Depending on the goals of SCP, researchers can choose the extent to which connectivity should
198 be prioritized by changing parameter b , which becomes the connectivity strength modifier (**Equation 2** in
199 **Box 2**)[53]. In other formulations of the SCP problem, the dispersal matrix can be used to maximize metrics
200 of metapopulation performance, such as the expected time to extinction [54,55]. Alternatively, dispersal
201 distances can be used to set the maximal size of CAs and distances between different CAs in a network to
202 ensure that propagules and juveniles generated in one CA can disperse to and recruit in nearby CAs [56,57].
203 The dispersal matrix can also be used to define site-specific metrics measuring the importance of each site
204 for population persistence using graph theory [58] or matrix analysis [55]. When used as biodiversity

205 features [58] or costs [59], such site-specific metrics lead to the selection of sites that are well-connected,
206 and this connectivity may enhance persistence within the CA network [58].

207 *Population size*

208 Estimates of N_c and N_e are useful to refine the targets T_j 's of species representation that constrain the
209 solution of the SCP problem (**equation 2** in **Box 2**)[60]. These targets define the minimum proportions of
210 the geographical ranges of species that need to be included in the sets of CAs to consider those species
211 adequately covered. Species with smaller geographical ranges are usually given higher proportional targets
212 of representation because they might face a higher risk of extinction than species with larger ranges [61].
213 Despite being easy to implement, this approach is an approximation for the complexity of demographic,
214 genetic and ecological factors affecting the long-term persistence of species. Estimates of N_c and N_e could
215 help set more appropriate targets, for example by increasing T_j for species that have low numbers of
216 individuals even if their geographical range is large or for species showing a negative temporal trend in
217 abundance. The information provided by population abundance complements that of occurrence in setting
218 conservation priorities [62] and many species are showing signs of declining abundance despite keeping
219 stable geographical ranges [63]. The approach used by the IUCN to classify species into threat categories is
220 also based on criteria of geographical ranges and population abundance [64].

221 When estimates of N_c and N_e are available per site, they can be used to define SCP problems in terms of
222 abundance: the representation levels (r_{ij}) are the site-specific population numbers and the target T_j is the
223 total species abundance required for long-term persistence, which can be found using population viability
224 analysis or set following the general 50/500 rule [65–67]. This approach requires a comprehensive sampling
225 across the range of the species, or a method to spatialize the estimates of N_c and N_e . While there are
226 several abundance-based species distribution models that predict N_c [68], similar approaches for N_e have
227 yet to be developed.

228 **Building adaptive conservation area networks**

229 A primary goal of well-connected and genetically representative CA networks is to support the persistence
230 of species in the face of anthropogenic disturbance, such as land use and climate change [69]. When loci of
231 large effect on fitness can be identified, there are two alternative conservation strategies that can be
232 adopted to account for future adaptation. First, when the direction of environmental change can be
233 predicted and the relationship between alleles and environmental variables is known, a decision can be
234 made to conserve the alleles that confer stronger adaptation to future environmental conditions, or the
235 sites that show the smallest genetic offset with future predicted conditions [70]. However, focusing on the
236 winners of environmental change relies on many strong assumptions, among which that the populations
237 are optimally adapted to current environmental conditions and that the relationships between alleles and
238 environmental variables are correctly characterized. In addition, when the direction of environmental
239 change is unclear, it is even more difficult to predict biological responses accurately.

240 A safer strategy is to conserve a portfolio of alleles at adaptive loci (i.e. adaptive genetic diversity) as
241 opposed to conserving only some alleles, as this confers higher adaptation capacity when future
242 environmental conditions are uncertain [71] and buffers the risk of incorrectly characterized gene-
243 environment associations. Depending on the genetic structure of the species, intraspecific genetic diversity
244 can be conserved either by prioritizing sites with the highest within-site diversity (alpha diversity) or by
245 protecting sets of sites with complementary genetic variants to maximize adaptation capacity at the
246 landscape scale (beta diversity; **Box 3**).

247 **Concluding remarks: getting the best (out of) genetic data**

248 Despite the potential for improving CA planning, there are still numerous challenges that should be tackled
249 by future research (see **Outstanding questions**). First, information from intraspecific genetic data is
250 affected by various types of uncertainty [72]. Some estimated variables, such as dispersal distance and

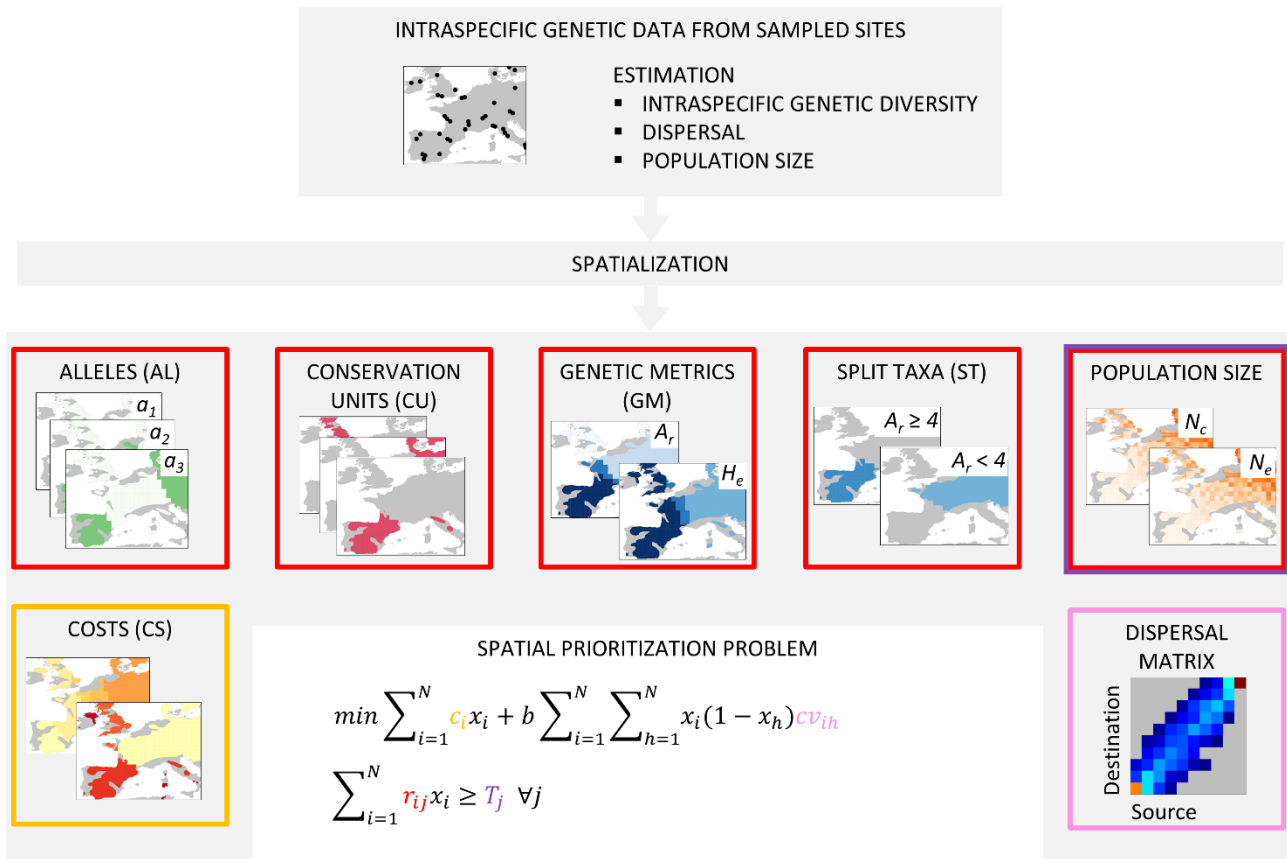
251 population size, can have wide confidence intervals [37,41] and the identity of conservation units and
252 adaptive genetic markers often depends on the methods used [24,30]. There is also limited knowledge
253 about the real effects of intraspecific genetic diversity on the adaptive potential of populations [73]. An
254 important area for future research is to evaluate the impact of these types of uncertainties on the selection
255 of CAs [74] and develop standardized, efficient workflows to integrate the uncertainty of inputs into multi-
256 species SCP [72].

257 Secondly, characterizing intraspecific genetic diversity requires multiple samples distributed throughout the
258 entire geographic range of species, and possibly replicated in time to estimate population abundance
259 trends. To estimate dispersal and population size, sampling must be carefully planned [35,43]. Multi-
260 species genetic studies are becoming more common [75,76] and efforts are made to bring together genetic
261 data sets for multiple species in free databases [34]. However, intraspecific genetic data are still lacking for
262 many species and attempts to replace them with surrogate variables (e.g. environmental variables) have
263 yielded mixed results [77–79]. Obtaining spatially and temporally replicated genetic samples for multiple
264 species, in line with local conservation priorities and involving all stakeholders [80], remains a main goal for
265 future research.

266 The scientific community has set ambitious goals to obtain genomic information for wild species: for
267 example, the Earth Biogenome Project aims to sequence all known eukaryotic species in a ten-year
268 timeframe [81]. The availability of genome sequences will undoubtedly help develop genetic markers for
269 wild species, but it will be necessary to understand how to best use the knowledge obtained from
270 reference genomes [82], for example the identification of deleterious mutations and the quantification of
271 mutation load [83], to plan networks of CAs. When genomic data are used to identify putatively adaptive
272 genetic markers, SCP solutions might be similar [85] or substantially different [74,84] to those found using
273 putatively neutral loci or traditional markers such as microsatellites.

274 Similarly, conserving intraspecific genetic diversity, dispersal and population size might require specific sets
275 of sites that increase conservation costs relative to the surface area needed to conserve species, possibly
276 making it more difficult to reach other conservation objectives. This is likely to happen each time new
277 objectives and constraints are added to the conservation problem. For example, the sites needed to
278 maintain ecosystem services and functional diversity are often different from those needed to conserve
279 species [86,87]. These conflicts in CA siting are eased when the connections between seemingly different
280 objectives are recognized: for example, ensuring that marine reserves ensure population persistence within
281 their borders (biodiversity conservation objective) and fishery supply beyond their borders (ecosystem
282 service objective) can be reconciled by siting them according to the dispersal capacity of the targeted
283 species [88]. This also shows that information obtained from intraspecific genetic data has an added benefit
284 [89] and may justify the extra money and time required to obtain them.

285 Systematic approaches to biodiversity conservation will be increasingly needed in the near future to reach
286 the targets of the post-2020 global biodiversity framework. Intraspecific genetic data are a wealthy source
287 of information not only for characterizing intraspecific genetic diversity, but also for estimating important
288 demographic parameters such as dispersal and population size. In addition to the framework briefly
289 illustrated here, there might be other ways, which will be important to assess, to expand SCP towards these
290 data. Early examples show that information from intraspecific genetic data is likely to improve the planning
291 of CAs to reach multiple ecological objectives.



292

293 **Figure 1. Integration of intraspecific genetic data into the ‘minimum set coverage’ spatial conservation prioritization (SCP)**
 294 **problem.** Intraspecific genetic data can enter the minimum set SCP problem in various ways. Estimates of intraspecific genetic
 295 diversity obtained from sampled sites can be converted into spatial layers through a spatialization step. Layers of alleles (AL),
 296 conservation units (CU), genetic metrics (GM; such as allelic richness A_r and expected heterozygosity H_e) and split taxa occurrences
 297 (ST) can be used as biodiversity features and enter the SCP problem via variable r_{ij} , the representation level of biodiversity feature j
 298 in site i . The ST example shows the distribution range of a taxon split into two layers on the basis of the A_r value, with a threshold of
 299 4 alleles. Information on dispersal, arranged in a dispersal matrix, can be used to define the connectivity penalty costs cv_{ih} .
 300 Estimates of population size at the species level can be used to refine the specific spatial representation targets T_j and estimates at
 301 the site level can be used as a layer to define SCP problems in terms of abundance. All three types of intraspecific genetic data can
 302 also be used to define layers of conservation costs c_i (CS). See **Table 1** and main text for detailed explanation of each method and
 303 **Box 2** for notation of the SCP problem.

304 **Table 1. Methods to integrate information from intraspecific genetic data in spatial conservation**
 305 **prioritization.**

| METHOD | DESCRIPTION | EXAMPLES |
|--------------------------------|--|------------------|
| ALLELES (AL) | Alleles are the biodiversity features. Allele presence or frequencies are mapped on the landscape and spatial layers are used as inputs in the prioritization. | [75,90–94] |
| CONSERVATION UNITS (CU) | Conservation units (management units, evolutionarily significant units or the branches of the phylogenetic tree) are treated as biodiversity features. As intraspecific genetic data are usually spatially sparse, the spatial distribution of individual conservation units is usually not known from observations, but can be predicted using spatial interpolation techniques or ecological niche models . In this latter case, each conservation unit is treated as a distinct entity in a model using environmental variables as predictors of its occurrence, with the possibility to include future environmental projections to forecast the response of each conservation unit under different climate change scenarios. | [22,47,84,95–97] |
| GENETIC METRICS (GM) | Genetic metrics, calculated for each species or conservation unit in each site, are the biodiversity features. Values in unsampled sites are predicted using spatialization techniques. A conceptual and practical difficulty with this method is the need to set representation targets for genetic metrics. | [74,93,98] |
| SPLIT TAXA (ST) | Taxa (species or conservation units) are the biodiversity features. Each taxon is represented by several spatial layers grouping sites sharing similar genetic characteristics. For example, distinct layers are used to represent sites with low, medium and high allelic richness or areas of low, medium and high genetic differentiation. Each layer has a spatial representation target. A limitation of this approach is that the number of distinct spatial layers and the limits among them are usually arbitrary. | [84,85,99–102] |
| COSTS (CS) | Costs are calculated as a function of site-specific or between-site genetic metrics. For example, sites with lower allelic richness are given higher protection costs to favor the | [99] |

selection of sites with higher genetic diversity. Pairwise genetic metrics can be integrated through boundary costs c_{ij} : pairs of sites with lower genetic differentiation are given lower pairwise costs to favor the selection of genetically connected sets of sites. One drawback of this approach is the need to combine information that may be incommensurable, e.g. genetic-based and monetary costs, or when costs are used to define layers of unsuitable habitats.

ENVIRONMENTAL DIVERSITY (ED) The ED formulation finds the subset of sites that contain the most representative set of environmental conditions among all candidate sites, subject to a limit on the number of sites that can be selected [50,103]. It uses a dissimilarity matrix to characterize the differences between each pair of sites: thus, it can be adapted to generate prioritizations that ensure a representative sample of genetic diversity among sites, using a genetic distance matrix instead of environmental dissimilarity. [50,77,78,104]

307 **Box 1. Dispersal estimates and their potential usefulness in spatial conservation**

308 **prioritization**

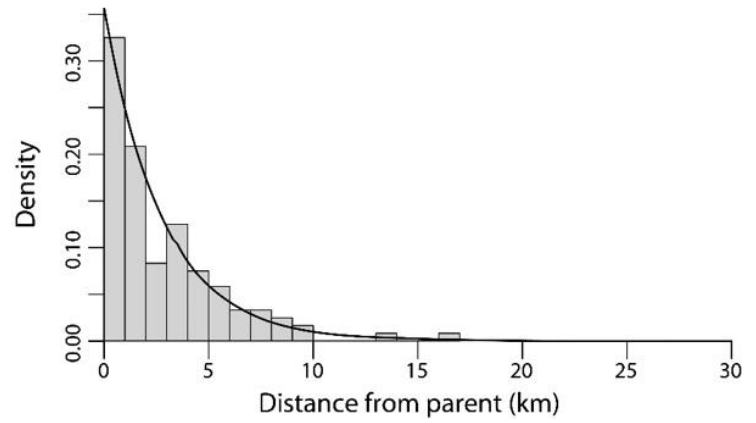
309 Intraspecific genetic data offer various ways to estimate dispersal in organisms and habitats that are
310 otherwise difficult to study using direct observations, such as tiny fish and invertebrate larvae that have the
311 potential to disperse widely on ocean currents.

312 Parentage analysis - whereby offspring are assigned to parents based on their DNA - can be used to directly
313 detect dispersal events [78]. In one example, D'Aloia *et al.* [36] genotyped over 7,000 individuals of the
314 neon goby *Elacatinus lori* and used parent-offspring matches to estimate the species' dispersal kernel on
315 the Belize barrier reef (Fig. I-A-B). They found that most larvae dispersed less than 2 km from their parents,
316 despite larvae spending nearly one month dispersing. Like most parentage studies, this was constrained to
317 a relatively small spatial area and required a large amount of sampling that will not be feasible to undertake
318 for all species of interest in SCP. However, follow-up studies have corroborated this strongly limited
319 dispersal pattern. For example, genetic sibship reconstruction revealed that full siblings are spatially
320 arranged as predicted by the parentage dispersal kernel [105] and genetic assignment tests at the scale of
321 the species' range revealed a low frequency of long-distance dispersal events [106]. The congruence
322 between multiple genetic estimates of dispersal in *E. lori* is promising for the application of more feasible
323 genetic-based estimates of dispersal in other species.

A



B



324

325 **Figure 1. Using genetic-based dispersal estimates to inform spatial conservation prioritization.** (A) A larva of the neon goby

326 *Elacatinus lori* (photo: J. Majoris); (B) The species' estimated dispersal kernel overlaid on a histogram of dispersal events detected

327 by parentage analysis. Fig. 1b drawn using data from [36].

328

(end of Box 1)

**Box 2. Spatial conservation prioritization as a framework to place new
conservation areas**

Spatial conservation prioritization (SCP) can be treated as a mathematical problem involving the spatial distribution of biodiversity features (e.g. species, indexed by $j = 1, \dots, S$) and conservation costs in a set of sites indexed by $i = 1, \dots, N$. r_{ij} indicates the spatial occurrence (binary variable) or abundance (continuous variable) of biodiversity feature j in site i . In one of the several possible types of SCP problems, the minimum set coverage [71], the mathematical formulation involves two equations:

$$\min \left(\sum_{i=1}^N c_i x_i + b \sum_{i=1}^N \sum_{h=1}^N x_i (1 - x_h) cv_{ih} \right) \quad (1)$$

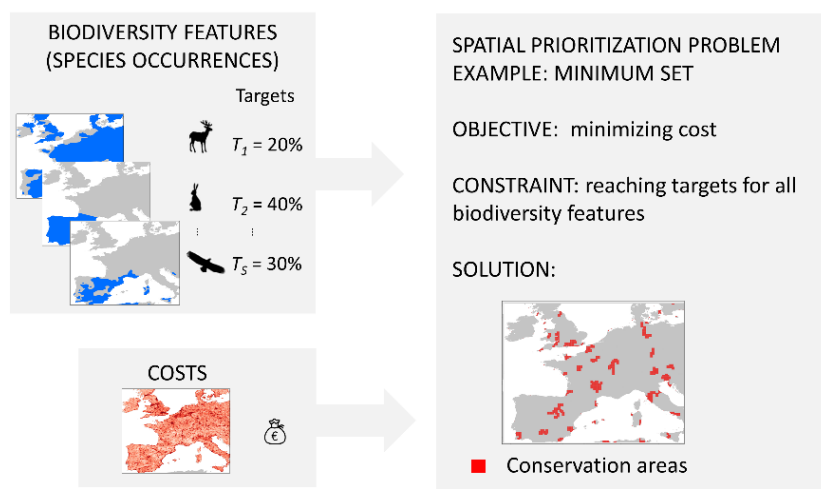
$$\sum_{i=1}^N r_{ij} x_i \geq T_j \quad \forall j \quad (2)$$

where x_i is the unknown variable indicating whether a site is selected for protection ($x_i = 1$) or not ($x_i = 0$). Solving the problem means finding the vector of x_i 's that satisfies the two equations.

Equation (1) states that the total cost of protection should be minimized. The total cost is the sum of two terms: the first term is the sum of the site-specific costs of protection, c_i 's, which can be defined as the monetary costs required to purchase the sites, as the opportunity costs of other excluded territorial uses or, in the absence of such information, simply as the surface area of the sites. The second term of equation (1) is used to limit the spatial fragmentation of the solution by introducing a boundary costs cv_{ih} , which is typically the length of the physical boundary between site i and h [5]; more simply, when sites have the same shape and size and are placed on a regular grid, $cv_{ih} = 1$ for adjacent sites and 0 otherwise. For two adjacent sites, the cost is paid when site i is protected but site h is not ($x_i = 1$ and $x_h = 0$). The "boundary length modifier" b is set according to the degree of fragmentation that is deemed acceptable (a lower b leads to a more fragmented solution).

348 Equation (2) constrains the solutions to sets of sites that include a minimum target proportion T_j of the
349 geographical range of each species. T_j is set according to ecological considerations: for example, species
350 with smaller ranges are given higher targets because they might be at higher risk of extinction than species
351 with larger ranges [61].

352 The SCP problem can be solved using exact or heuristic methods implemented in several software packages
353 [5,107,108]. The solution is a list of priority sites for the creation of new conservation areas (**Figure I**).



354

355 **Figure I**

356

(end of Box 2)

357

358 **Box 3. Retaining adaptive genetic diversity to foster persistence under uncertain**
359 **future conditions**

360 Prioritizing portfolios of genetic combinations increases the probability that “winning” combinations can
361 persist during periods of environmental change [71]. Depending on the genetic structure of the species,
362 targeting sites with high within-site adaptive genetic diversity or sites with populations adapted to different
363 local conditions will help build conservation area networks that retain the genetic diversity of species.

364 An example of prioritizing within-site diversity is given by Xuereb *et al.* [74]. They used environmental
365 association analysis to identify 51 SNPs associated with mean bottom temperature in the California sea
366 cucumber *Parastichopus californicus* living in the coastal seas of British Columbia (Canada). Then, they used
367 within-site heterozygosity at these putatively adaptive SNPs as a biodiversity feature in spatial conservation
368 prioritization (SCP), which led to the selection of sites in the northern region of the study area. In a second
369 prioritization exercise, they used the frequency of warm-temperature-associated alleles as a biodiversity
370 feature, which led to the selection of sites in the southern region. These results illustrate a trade-off
371 between prioritizing specific alleles versus prioritizing genetic diversity.

372 The second option is protecting a portfolio of sites with a diverse set of adaptations. Hanson *et al.* [84]
373 genotyped three amphibian species living in the Iberian peninsula at several thousand SNPs. Using outlier
374 detection and environmental association analyses with climatic and soil variables, they identified several
375 putatively adaptive loci in each species. They then identified sets of populations sharing similar adaptations
376 (adaptive units [24]) by applying genetic clustering techniques to these putatively adaptive loci and used
377 them as distinct biodiversity features in SCP. This allowed them to identify a set of complementary priority
378 areas for the conservation of adaptive genetic diversity at the species level.

379 When selecting different sites, it is important that the genetic variants that may be favorable under future
380 conditions will be able to spread to the other sites. For this reason, it is advisable to combine the

381 prioritization of genetically diverse sites with estimates of dispersal to build adaptation networks capable of
382 exchanging favorable genetic variants when needed [71]. Various approaches are available to integrate this
383 type of information in SCP (see main text). It should be noted, however, that prioritizing portfolios of
384 genetic combinations is still subject to the difficulties of correctly characterizing adaptive genetic diversity.

385 **(end of Box 3)**

386 **GLOSSARY**

387 **Adaptive genetic diversity.** The genetic diversity that is estimated at adaptive genes, i.e. those that have an
388 effect on fitness [27]

389 **Biodiversity feature.** A component of biodiversity (e.g. species, alleles, ecosystems) that can be mapped in
390 a landscape.

391 **Census population size.** The count of individuals in a population, often restricted to adult individuals.

392 **Comprehensiveness.** The degree to which a set of conservation areas includes all elements of biodiversity
393 features [51].

394 **Demographic connectivity.** The relative contribution of dispersal to population dynamics.

395 **Demographic rescue.** A decrease in population extinction probability owing to the simple addition of
396 immigrants.

397 **Dispersal kernel.** The statistical distribution of dispersal distances in a population.

398 **Dispersal matrix.** A dispersal matrix describes the probability of dispersal between a set of sites in the
399 landscape. Each element of the dispersal matrix is the dispersal probability from site j to site i , which may
400 be different from the dispersal probability from site i to site j (asymmetric dispersal).

401 **Ecological niche model.** A statistical model linking the spatial occurrence of a biodiversity feature to a set of
402 environmental variables. It is often used to predict species occurrences in places where no data are
403 available (spatial prediction) or in the future (forecasting).

404 **Effective population size.** The size of an ideal population experiencing the same rate of genetic drift or
405 inbreeding as the population under study. The ideal population is usually a closed population of constant
406 size with discrete generations and a Poisson variance in reproductive success between individuals.

407 **Environmental association analysis.** A statistical approach to identify genetic variants strongly associated
408 with specific environmental conditions.

409 **Genetic rescue.** A decrease in population extinction probability owing to gene flow.

410 **Inbreeding depression.** Reduced fitness of offspring with related parents, often due to deleterious
411 recessive alleles that become expressed in homozygous state.

412 **Outlier test.** A statistical approach to identify loci involved in local adaptation by screening for alleles that
413 show unusually high genetic differentiation among populations, i.e. outside of the distribution expected
414 under neutrality.

415 **Neutral genetic diversity.** The genetic diversity estimated at putatively neutral genes, i.e. those that do not
416 have any direct effect on fitness. This type of genetic diversity is selectively neutral and is useful to estimate
417 dispersal and population size [27].

418

419 **References**

- 420 1 Díaz, S. *et al.* (2019) Pervasive human-driven decline of life on Earth points to the need for
421 transformative change. *Science* 366, 1327
- 422 2 Moilanen, A. *et al.*, eds. (2009) *Spatial Conservation Prioritization: Quantitative Methods and*
423 *Computational Tools*, Oxford University Press.
- 424 3 Villarreal-Rosas, J. *et al.* (2020) Advancing Systematic Conservation Planning for Ecosystem Services.
425 *Trends Ecol. Evol.* 35, 1129–1139
- 426 4 Pollock, L.J. *et al.* (2020) Protecting Biodiversity (in All Its Complexity): New Models and Methods.
427 *Trends Ecol. Evol.* 35, 1119–1128
- 428 5 Daigle, R.M. *et al.* (2020) Operationalizing ecological connectivity in spatial conservation planning with
429 Marxan Connect. *Methods Ecol. Evol.* 11, 570–579
- 430 6 Hohenlohe, P.A. *et al.* (2021) Population genomics for wildlife conservation and management. *Mol.*
431 *Ecol.* 30, 62–82
- 432 7 Cayuela, H. *et al.* (2018) Demographic and genetic approaches to study dispersal in wild animal
433 populations: A methodological review. *Mol. Ecol.* 27, 3976–4010
- 434 8 Des Roches, S. *et al.* (2018) The ecological importance of intraspecific variation. *Nat. Ecol. Evol.* 2, 57–
435 64
- 436 9 Des Roches, S. *et al.* (2021) Conserving intraspecific variation for nature’s contributions to people. *Nat.*
437 *Ecol. Evol.* 5, 574–582
- 438 10 Razgour, O. *et al.* (2019) Considering adaptive genetic variation in climate change vulnerability
439 assessment reduces species range loss projections. *Proc. Natl. Acad. Sci.* 116, 10418–10423

- 440 11 Leigh, D.M. *et al.* (2019) Estimated six per cent loss of genetic variation in wild populations since the
441 industrial revolution. *Evol. Appl.* 12, 1505–1512
- 442 12 Exposito-Alonso, M. *et al.* (2021) Quantifying the scale of genetic diversity extinction in the
443 Anthropocene. *bioRxiv* DOI: 10.1101/2021.10.13.464000
- 444 13 Munguía-Vega, A. *et al.* (2015) Marine reserves help preserve genetic diversity after impacts derived
445 from climate variability: Lessons from the pink abalone in Baja California. *Glob. Ecol. Conserv.* 4, 264–
446 276
- 447 14 Gaggiotti, O.E. *et al.* (2018) Diversity from genes to ecosystems: A unifying framework to study
448 variation across biological metrics and scales. *Evol. Appl.* 11, 1176–1193
- 449 15 Jost, L. *et al.* (2018) Differentiation measures for conservation genetics. *Evol. Appl.* 11, 1139–1148
- 450 16 Funk, W.C. *et al.* (2012) Harnessing genomics for delineating conservation units. *Trends Ecol. Evol.* 27,
451 489–496
- 452 17 Lowe, W.H. and Allendorf, F.W. (2010) What can genetics tell us about population connectivity? *Mol.*
453 *Ecol.* 19, 3038–3051
- 454 18 Palsbøll, P.J. *et al.* (2007) Identification of management units using population genetic data. *Trends*
455 *Ecol. Evol.* 22, 11–16
- 456 19 Moritz, C. (1994) Defining ‘Evolutionarily Significant Units’ for conservation. *Trends Ecol. Evol.* 9, 373–
457 375
- 458 20 Mishler, B.D. *et al.* (2014) Phylogenetic measures of biodiversity and neo- and paleo-endemism in
459 Australian Acacia. *Nat. Commun.* 5, 4473
- 460 21 Davis, E.B. *et al.* (2008) The California Hotspots Project: identifying regions of rapid diversification of
461 mammals. *Mol. Ecol.* 17, 120–138

- 462 22 Rosauer, D.F. *et al.* (2018) Real-world conservation planning for evolutionary diversity in the Kimberley,
463 Australia, sidesteps uncertain taxonomy. *Conserv. Lett.* 11, e12438
- 464 23 Casacci, L.P. *et al.* (2014) The “Evolutionarily Significant Unit” concept and its applicability in biological
465 conservation. *Ital. J. Zool.* 81, 182–193
- 466 24 Barbosa, S. *et al.* (2018) Integrative approaches to guide conservation decisions: Using genomics to
467 define conservation units and functional corridors. *Mol. Ecol.* 27, 3452–3465
- 468 25 Yang, Z. and Rannala, B. (2012) Molecular phylogenetics: principles and practice. *Nat. Rev. Genet.* 13,
469 303–314
- 470 26 Kapli, P. *et al.* (2020) Phylogenetic tree building in the genomic age. *Nat. Rev. Genet.* 21, 428–444
- 471 27 Holderegger, R. *et al.* (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics.
472 *Landsc. Ecol.* 21, 797–807
- 473 28 Santure, A.W. and Garant, D. (2018) Wild GWAS—association mapping in natural populations. *Mol.*
474 *Ecol. Resour.* 18, 729–738
- 475 29 Kardos, M. and Shafer, A.B.A. (2018) The Peril of Gene-Targeted Conservation. *Trends Ecol. Evol.* 33,
476 827–839
- 477 30 Hoban, S. *et al.* (2016) Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and
478 Future Directions. *Am. Nat.* 188, 379–397
- 479 31 Manel, S. *et al.* (2016) Genomic resources and their influence on the detection of the signal of positive
480 selection in genome scans. *Mol. Ecol.* 25, 170–184
- 481 32 Kardos, M. *et al.* (2021) The crucial role of genome-wide genetic variation in conservation. *Proc. Natl.*
482 *Acad. Sci. U. S. A.* 118, e2104642118

- 483 33 Mittell, E.A. *et al.* (2015) Are molecular markers useful predictors of adaptive potential? *Ecol. Lett.* 18,
484 772–778
- 485 34 Leigh, D.M. *et al.* (2021) Opportunities and challenges of macrogenetic studies. *Nat. Rev. Genet.* 22,
486 791–807
- 487 35 Christie, M.R. *et al.* (2017) Disentangling the relative merits and disadvantages of parentage analysis
488 and assignment tests for inferring population connectivity. *ICES J. Mar. Sci.* 74, 1749–1762
- 489 36 D’Aloia, C.C. *et al.* (2015) Patterns, causes, and consequences of marine larval dispersal. *Proc. Natl.*
490 *Acad. Sci. U. S. A.* 112, 13940–13945
- 491 37 Pinsky, M.L. *et al.* (2017) Marine Dispersal Scales Are Congruent over Evolutionary and Ecological Time.
492 *Curr. Biol.* 27, 149–154
- 493 38 Gagnaire, P. *et al.* (2015) Using neutral, selected, and hitchhiker loci to assess connectivity of marine
494 populations in the genomic era. *Evol. Appl.* 8, 769–786
- 495 39 Luikart, G. *et al.* (2010) Estimation of census and effective population sizes: the increasing usefulness of
496 DNA-based approaches. *Conserv. Genet.* 11, 355–373
- 497 40 Bravington, M.V. *et al.* (2016) Close-Kin Mark-Recapture. *Stat. Sci.* 31, 259–274
- 498 41 Hillary, R.M. *et al.* (2018) Genetic relatedness reveals total population size of white sharks in eastern
499 Australia and New Zealand. *Sci. Rep.* 8, 2661
- 500 42 Willi, Y. *et al.* (2022) Conservation genetics as a management tool: The five best-supported paradigms
501 to assist the management of threatened species. *Proc. Natl. Acad. Sci.* 119, e2105076119
- 502 43 Wang, J. *et al.* (2016) Prediction and estimation of effective population size. *Heredity* 117, 193–206
- 503 44 Waples, R.S. and Feutry, P. (2022) Close-kin methods to estimate census size and effective population
504 size. *Fish Fish.* 23, 273–293

- 505 45 Luikart, G. *et al.* (2021) Detecting population declines via monitoring the effective number of breeders
506 (N_b). *Mol. Ecol. Resour.* 21, 379–393
- 507 46 Moilanen, A. *et al.* (2009) A mathematical classification of conservation prioritization problems. In
508 *Spatial conservation prioritization: Quantitative methods and computational tools* (Moilanen, A. *et al.*,
509 eds), pp. 28–42, Oxford University Press
- 510 47 Carvalho, S.B. *et al.* (2017) Spatial conservation prioritization of biodiversity spanning the evolutionary
511 continuum. *Nat. Ecol. Evol.* 1, 0151
- 512 48 Rodrigues, A.S.L. and Gaston, K.J. (2002) Maximising phylogenetic diversity in the selection of networks
513 of conservation areas. *Biol. Conserv.* 105, 103–111
- 514 49 Diniz-Filho, J.A.F. and Telles, M.P. de C. (2002) Spatial autocorrelation analysis and the identification of
515 operational units for conservation in continuous populations. *Conserv. Biol.* 16, 924–935
- 516 50 Hanson, J.O. *et al.* (2018) RAPTR: Representative and adequate prioritization toolkit in R. *Methods Ecol.*
517 *Evol.* 9, 320–330
- 518 51 Wilson, K.A. *et al.* (2009) Fundamental concepts of spatial conservation prioritization. In *Spatial*
519 *conservation prioritization: Quantitative methods and computational tools* (Moilanen, A. *et al.*, eds),
520 pp. 16–27, Oxford University Press
- 521 52 Haight, R.G. and Snyder, S.A. (2009) Integer programming methods for reserve selection and design. In
522 *Spatial conservation prioritization: Quantitative methods and computational tools* (Moilanen, A. *et al.*,
523 eds), pp. 43–57, Oxford University Press
- 524 53 Beger, M. *et al.* (2010) Incorporating asymmetric connectivity into spatial decision making for
525 conservation. *Conserv. Lett.* 3, 359–368

- 526 54 Nicholson, E. and Ovaskainen, O. (2009) Conservation prioritization using metapopulation models. In
527 *Spatial conservation prioritization: Quantitative methods and computational tools* (Moilanen, A. et al.,
528 eds), pp. 110–121, Oxford University Press
- 529 55 Nilsson Jacobi, M. and Jonsson, P.R. (2011) Optimal networks of nature reserves can be found through
530 eigenvalue perturbation theory of the connectivity matrix. *Ecol. Appl.* 21, 1861–1870
- 531 56 D’Aloia, C.C. et al. (2017) A multiple-species framework for integrating movement processes across life
532 stages into the design of marine protected areas. *Biol. Conserv.* 216, 93–100
- 533 57 Balbar, A.C. and Metaxas, A. (2019) The current application of ecological connectivity in the design of
534 marine protected areas. *Glob. Ecol. Conserv.* 17, e00569
- 535 58 Magris, R.A. et al. (2018) Biologically representative and well-connected marine reserves enhance
536 biodiversity persistence in conservation planning. *Conserv. Lett.* 11, e12439
- 537 59 Weeks, R. (2017) Incorporating seascape connectivity in conservation prioritisation. *PLOS ONE* 12,
538 e0182396
- 539 60 Burgman, M.A. et al. (2001) A Method for Setting the Size of Plant Conservation Target Areas. *Conserv.*
540 *Biol.* 15, 603–616
- 541 61 Rodrigues, A.S.L. et al. (2004) Global Gap Analysis: Priority Regions for Expanding the Global Protected-
542 Area Network. *BioScience* 54, 1092–1100
- 543 62 Burgess, M. et al. (2019) A new framework of spatial targeting for single-species conservation planning.
544 *Landsc. Ecol.* 34, 2765–2778
- 545 63 WWF (2020) *Living Planet Report 2020: Bending the Curve of Biodiversity Loss*, WWF.
- 546 64 Mace, G.M. et al. (2008) Quantification of extinction risk: IUCN’s system for classifying threatened
547 species. *Conserv. Biol.* 22, 1424–1442

- 548 65 Jamieson, I.G. and Allendorf, F.W. (2012) How does the 50/500 rule apply to MVPs? *Trends Ecol. Evol.*
549 27, 578–584
- 550 66 Jamieson, I.G. and Allendorf, F.W. (2013) A school of red herring: reply to Frankham et al. *Trends Ecol.*
551 *Evol.* 28, 188–189
- 552 67 Frankham, R. et al. (2013) 50/500 rule and minimum viable populations: response to Jamieson and
553 Allendorf. *Trends Ecol. Evol.* 28, 187–188
- 554 68 Waldock, C. et al. (2022) A quantitative review of abundance-based species distribution models.
555 *Ecography* 2022, e05694
- 556 69 Reside, A.E. et al. (2018) Adapting systematic conservation planning for climate change. *Biodivers.*
557 *Conserv.* 27, 1–29
- 558 70 Fitzpatrick, M.C. and Keller, S.R. (2015) Ecological genomics meets community-level modelling of
559 biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecol.*
560 *Lett.* 18, 1–16
- 561 71 Webster, M.S. et al. (2017) Who should pick the winners of climate change? *Trends Ecol. Evol.* 32, 167–
562 173
- 563 72 Regan, H.M. et al. (2009) Conservation prioritization and uncertainty in planning inputs. In *Spatial*
564 *conservation prioritization: Quantitative methods and computational tools* (Moilanen, A. et al., eds),
565 pp. 145–157, Oxford University Press
- 566 73 Hoffmann, A.A. et al. (2017) Revisiting Adaptive Potential, Population Size, and Conservation. *Trends*
567 *Ecol. Evol.* 32, 506–517
- 568 74 Xuereb, A. et al. (2021) Incorporating putatively neutral and adaptive genomic data into marine
569 conservation planning. *Conserv. Biol.* 35, 909–920

- 570 75 Taberlet, P. *et al.* (2012) Genetic diversity in widespread species is not congruent with species richness
571 in alpine plant communities. *Ecol. Lett.* 15, 1439–1448
- 572 76 Selkoe, K.A. *et al.* (2016) The DNA of coral reef biodiversity: predicting and protecting genetic diversity
573 of reef assemblages. *Proc. R. Soc. B Biol. Sci.* 283, 20160354
- 574 77 Hanson, J.O. *et al.* (2017) Environmental and geographic variables are effective surrogates for genetic
575 variation in conservation planning. *Proc. Natl. Acad. Sci. U. S. A.* 114, 12755–12760
- 576 78 Hanson, J.O. *et al.* (2021) Evaluating surrogates of genetic diversity for conservation planning. *Conserv.*
577 *Biol.* 35, 634–642
- 578 79 Ponce-Reyes, R. *et al.* (2014) Geographical surrogates of genetic variation for selecting island
579 populations for conservation. *Divers. Distrib.* 20, 640–651
- 580 80 Rossetto, M. *et al.* (2021) A conservation genomics workflow to guide practical management actions.
581 *Glob. Ecol. Conserv.* 26, e01492
- 582 81 Lewin, H.A. *et al.* (2022) The Earth BioGenome Project 2020: Starting the clock. *Proc. Natl. Acad. Sci. U.*
583 *S. A.* 119, e2115635118
- 584 82 Formenti, G. *et al.* (2022) The era of reference genomes in conservation genomics. *Trends Ecol. Evol.*
585 37, 197–202
- 586 83 van Oosterhout, C. (2020) Mutation load is the spectre of species conservation. *Nat. Ecol. Evol.* 4,
587 1004–1006
- 588 84 Hanson, J.O. *et al.* (2020) Conservation planning for adaptive and neutral evolutionary processes. *J.*
589 *Appl. Ecol.* 57, 2159–2169
- 590 85 Nielsen, E.S. *et al.* (2020) A comparison of genetic and genomic approaches to represent evolutionary
591 potential in conservation planning. *Biol. Conserv.* 251, 108770

- 592 86 O'Connor, L.M.J. *et al.* (2021) Balancing conservation priorities for nature and for people in Europe.
593 *Science* 372, 856–860
- 594 87 Jung, M. *et al.* (2021) Areas of global importance for conserving terrestrial biodiversity, carbon and
595 water. *Nat. Ecol. Evol.* 5, 1499–1509
- 596 88 Krueck, N.C. *et al.* (2017) Incorporating larval dispersal into MPA design for both conservation and
597 fisheries. *Ecol. Appl.* 27, 925–941
- 598 89 Costello, C. *et al.* (2010) The value of spatial information in MPA network design. *Proc. Natl. Acad. Sci.*
599 *U. S. A.* 107, 18294–18299
- 600 90 Diniz-Filho, J.A.F. *et al.* (2012) Planning for optimal conservation of geographical genetic variability
601 within species. *Conserv. Genet.* 13, 1085–1093
- 602 91 Diniz-Filho, J.A.F. *et al.* (2016) Exhaustive search for conservation networks of populations representing
603 genetic diversity. *Genet. Mol. Res.* 15, gmr.15017525
- 604 92 Diniz-Filho, J.A.F. *et al.* (2020) Overcoming the worst of both worlds: integrating climate change and
605 habitat loss into spatial conservation planning of genetic diversity in the Brazilian Cerrado. *Biodivers.*
606 *Conserv.* 29, 1555–1570
- 607 93 Schlottfeldt, S. *et al.* (2015) Multi-objective optimization in systematic conservation planning and the
608 representation of genetic variability among populations. *Genet. Mol. Res.* 14, 6744–6761
- 609 94 von Takach, B. *et al.* (2021) Population genomics and conservation management of a declining tropical
610 rodent. *Heredity* 126, 763–775
- 611 95 Vasconcelos, R. *et al.* (2012) Identifying priority areas for island endemics using genetic versus specific
612 diversity – The case of terrestrial reptiles of the Cape Verde Islands. *Biol. Conserv.* 153, 276–286
- 613 96 Vasconcelos, R. *et al.* (2018) Combining molecular and landscape tools for targeting evolutionary
614 processes in reserve design: An approach for islands. *PLOS ONE* 13, e0200830

- 615 97 Hermoso, V. *et al.* (2016) Species distributions represent intraspecific genetic diversity of freshwater
616 fish in conservation assessments. *Freshw. Biol.* 61, 1707–1719
- 617 98 Bonin, A. *et al.* (2007) Population adaptive index: a new method to help measure intraspecific genetic
618 diversity and prioritize populations for conservation. *Conserv. Biol.* 21, 697–708
- 619 99 Beger, M. *et al.* (2014) Evolving coral reef conservation with genetic information. *Bull. Mar. Sci.* 90,
620 159–185
- 621 100 Nielsen, E.S. *et al.* (2017) Multispecies genetic objectives in spatial conservation planning. *Conserv.*
622 *Biol.* 31, 872–882
- 623 101 Thomassen, H.A. *et al.* (2011) Mapping evolutionary process: a multi-taxa approach to conservation
624 prioritization. *Evol. Appl.* 4, 397–413
- 625 102 Phair, N.L. *et al.* (2021) Applying genomic data to seagrass conservation. *Biodivers. Conserv.* 30,
626 2079–2096
- 627 103 Faith, D.P. and Walker, P.A. (1996) Environmental diversity: on the best-possible use of surrogate
628 data for assessing the relative biodiversity of sets of areas. *Biodivers. Conserv.* 5, 399–415
- 629 104 Moritz, C. (2002) Strategies to protect biological diversity and the evolutionary processes that
630 sustain it. *Syst. Biol.* 51, 238–254
- 631 105 D’Aloia, C.C. *et al.* (2018) Limited dispersal explains the spatial distribution of siblings in a reef fish
632 population. *Mar. Ecol. Prog. Ser.* 607, 143–154
- 633 106 D’Aloia, C.C. *et al.* (2022) Population assignment tests uncover rare long-distance marine larval
634 dispersal events. *Ecology* 103, e03559
- 635 107 Lehtomäki, J. and Moilanen, A. (2013) Methods and workflow for spatial conservation prioritization
636 using Zonation. *Environ. Model. Softw.* 47, 128–137

637 108 Hanson, J.O. *et al.* (2021) *prioritizr: Systematic Conservation Prioritization in R. R package version*

638 7.0.1. <https://prioritizr.net/>,

639

Outstanding questions

What is the risk of disregarding intraspecific genetic data in spatial conservation prioritization (SCP) for biodiversity persistence? Conservation decisions have been and will be made in the absence of intraspecific genetic data, especially when they are too demanding to be collected. In what cases is it worth spending more time and money to collect genetic data?

As intraspecific genetic data are still lacking for many species, to what extent can they be replaced by surrogate information such as environmental variables in setting spatial conservation priorities?

What is the best method to spatialize genetic data to obtain information for unsampled sites?

How much more land and sea surface area will have to be protected to represent intraspecific genetic diversity? Previous studies showed that moderate extension of the current global system of conservation areas (CAs) would be sufficient to represent phylogenetic and functional diversity, but would this be true for intraspecific genetic diversity?

What is the impact of uncertainty in genetic data on the outcome of SCP? When uncertain genetic data are used in conjunction with other types of information to represent additional constraints to prioritization, the results risk being economically inefficient or unfavorable for conservation.

What is the risk of integrating intraspecific genetic data for some species only? Maximizing genetic diversity of one species can lower diversity of others. How would one select the species to collect genetic data on?

Table S1. List of published studies integrating intraspecific genetic data in spatial conservation prioritization

| Ref | Citation | Number of species | Molecular markers ^a | Integration of genetic data ^b | SCP method ^c | Genetic metrics ^d | Inference of information in unsampled sites |
|-----|----------------------------|-------------------|--------------------------------|--|------------------------------|--|--|
| [1] | Moritz (2002) | 10 | mtDNA | ED | Environmental diversity | Nei's [26] genetic distance averaged across species | Not performed. Only sampled sites were included in the prioritization. |
| [2] | Bonin et al. (2007) | 2 | AFLP | GM | Exhaustive search | Proportion of polymorphic loci Population adaptive index [2] | Not performed. Only sampled sites were included in the prioritization. |
| [3] | Thomassen et al. (2011) | 7 | AFLP, msat, nuDNA | ST | RESNET | Nei's [26] genetic distance F_{ST} ϕ_{ST} | Generalized dissimilarity modelling [27] |
| [4] | Diniz-Filho et al. (2012) | 1 | msat | AL | Simulated annealing | | Not performed. Only sampled sites were included in the prioritization. |
| [5] | Taberlet et al. (2012) | 39 | AFLP | AL | ZONATION | | Not needed. Sampling was performed using a regular grid |
| [6] | Vasconcelos et al. (2012) | 30 | None ^e | CU | ZONATION | | Ecological niche model with MAXENT [28] |
| [7] | Beger et al. (2014) | 1 | msat | ST, CM | MARXAN | Genetic clusters identified with STRUCTURE [29] Allelic richness Local F_{ST} estimated with GESTE [30] Asymmetric recent migration rates estimated with BAYESASS+ [31] | Allelic richness and local F_{ST} were interpolated in ARCGIS. Asymmetric migration rates were applied to proximate neighborhood identified using Thiessen polygons |
| [8] | Schlottfeldt et al. (2015) | 1 | msat | AL, GM | Multi-objective Evolutionary | Expected heterozygosity | Not performed. Only sampled sites were included in the prioritization. |

| | | | | | Algorithms (MOEA, [36]) | p-value of χ^2 test for Hardy-Weinberg equilibrium | |
|------|---------------------------|----|-------------------|----|----------------------------|--|--|
| [9] | Diniz-Filho et al. (2016) | 1 | msat | AL | Exhaustive search | | Not performed. Only sampled sites were included in the prioritization. |
| [10] | Hermoso et al. (2016) | 4 | Msat, mtDNA | CU | MARXAN | | Generalized dissimilarity modelling [27] |
| [11] | Carvalho et al. (2017) | 33 | mtDNA | CU | ZONATION, MARXAN | | Phylogeographical interpolation with PHYLIN [32] |
| [12] | Hanson et al. (2017) | 27 | AFLP | ED | RAPTR | Gower's [33] distance | Not needed. Sampling was performed using a regular grid |
| [13] | Nielsen et al. (2017) | 5 | mtDNA | ST | MARXAN | Haplotype diversity Nucleotide diversity Number of private haplotypes Local genetic differentiation | Inverse distance-weighting |
| [14] | Hanson et al. (2018) | 1 | AFLP | ED | RAPTR | Gower's [33] distance | Not needed. Sampling was performed using a regular grid |
| [15] | Vasconcelos et al. (2018) | 23 | mtDNA | CU | ZONATION | | Ecological niche model with MAXENT [28] |
| [16] | Paz-Vinas et al. (2018) | 6 | msat | AL | MARXAN | Allelic richness Private allelic richness Jost's [34] differentiation | Generalized linear models for spatial stream networks [35,36] |
| [17] | Rosauer et al. (2018) | 11 | None ^e | CU | MARXAN | | Lineage distribution model [37] |
| [18] | Hanson et al. (2019) | 9 | AFLP | CM | PRIORITIZR | Landscape resistance estimated from Nei's [26] genetic distance between sites | Not needed. Sampling was performed using a regular grid |

| | | | | | | | |
|------|---------------------------|----|------------|--------|-------------------------|---|--|
| [19] | Diniz-Filho et al. (2020) | 1 | msat | AL | Exhaustive search | | Not performed. Only sampled sites were included in the prioritization. |
| [20] | Hanson et al. (2020) | 3 | SNP | CU, ST | PRIORITIZR | Mean individual heterozygosity | Thin plate splines; phylogenetic interpolation with PHYLIN [32] |
| [21] | Nielsen et al. (2020) | 5 | mtDNA, SNP | ST | MARXAN | Nucleotide diversity Percentage of private alleles Percent of outlier SNPs | Inverse distance-weighting |
| [22] | Hanson et al. (2021) | 10 | msat | ED | Environmental diversity | Jost's [34] genetic differentiation | Not performed. Only sampled sites were included in the prioritization. |
| [23] | Phair et al. (2021) | 1 | SNP | ST | MARXAN | Nucleotide diversity Expected heterozygosity Allelic richness Number of shared SNPs and private SNPs Proportion of outlier SNPs | Inverse distance-weighting |
| [24] | von Takach et al. (2021) | 1 | SNP | AL | PRIORITIZR | | Not performed. Only sampled sites were included in the prioritization. |
| [25] | Xuereb et al. (2021) | 1 | SNP | GM | PRIORITIZR | Expected heterozygosity Local F_{ST} Adaptive score [38] Population adaptive index [39]. | Inverse distance-weighting |

Notes

The table includes only papers using intraspecific genetic data to obtain information that is used to define the input of a spatial conservation prioritization problem. The table was prepared starting from papers known to the authors and searching within the literature cited in them.

^a Type of molecular marker used: mitochondrial DNA (mtDNA), nuclear DNA (nuDNA), amplified fragment length polymorphisms (AFLP), microsatellites (msat), single nucleotide polymorphisms (SNP)

^b Methods used to integrate intraspecific genetic data in spatial conservation prioritization: alleles (AL), conservation units (CU), genetic metrics (GM), split taxa (ST), environmental diversity (ED). See **Table 1** in the main text for description of the methods

^c Method or software package used to perform spatial conservation prioritization (SCP): environmental diversity [40], MARXAN [41], ZONATION [42], PRIORITIZR [43], RAPTR [14], RESNET [44], MOEA (multi-objective evolutionary algorithms, [45]). Exhaustive search means that all the possible combinations of sites were considered.

^d Genetic metrics used in the prioritization either as biodiversity features (GM method), to split taxa layers (ST methods), as a distance or dissimilarity metric in the environmental diversity method (ED) or to integrate information on connectivity.

^e ESUs had been identified in other studies

References

- 1 Moritz, C. (2002) Strategies to Protect Biological Diversity and the Evolutionary Processes That Sustain It. *Syst. Biol.* 51, 238–254
- 2 Bonin, A. *et al.* (2007) Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conserv. Biol. J. Soc. Conserv. Biol.* 21, 697–708
- 3 Thomassen, H.A. *et al.* (2011) Mapping evolutionary process: a multi-taxa approach to conservation prioritization. *Evol. Appl.* 4, 397–413
- 4 Diniz-Filho, J.A.F. *et al.* (2012) Planning for optimal conservation of geographical genetic variability within species. *Conserv. Genet.* 13, 1085–1093
- 5 Taberlet, P. *et al.* (2012) Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecol. Lett.* 15, 1439–1448
- 6 Vasconcelos, R. *et al.* (2012) Identifying priority areas for island endemics using genetic versus specific diversity – The case of terrestrial reptiles of the Cape Verde Islands. *Biol. Conserv.* 153, 276–286
- 7 Beger, M. *et al.* (2014) Evolving coral reef conservation with genetic information. *Bull. Mar. Sci.* 90, 159–185
- 8 Schlottfeldt, S. *et al.* (2015) Multi-objective optimization in systematic conservation planning and the representation of genetic variability among populations. *Genet. Mol. Res.* 14, 6744–6761
- 9 Diniz-Filho, J.A.F. *et al.* (2016) Exhaustive search for conservation networks of populations representing genetic diversity. *Genet. Mol. Res.* 15,
- 10 Hermoso, V. *et al.* (2016) Species distributions represent intraspecific genetic diversity of freshwater fish in conservation assessments. *Freshw. Biol.* 61, 1707–1719
- 11 Carvalho, S.B. *et al.* (2017) Spatial conservation prioritization of biodiversity spanning the evolutionary continuum. *Nat. Ecol. Evol.* 1, 0151
- 12 Hanson, J.O. *et al.* (2017) Environmental and geographic variables are effective surrogates for genetic variation in conservation planning. *Proc. Natl. Acad. Sci.* 114, 12755–12760
- 13 Nielsen, E.S. *et al.* (2017) Multispecies genetic objectives in spatial conservation planning. *Conserv. Biol.* 31, 872–882
- 14 Hanson, J.O. *et al.* (2018) RAPTR: Representative and adequate prioritization toolkit in R. *Methods Ecol. Evol.* 9, 320–330

- 15 Vasconcelos, R. *et al.* (2018) Combining molecular and landscape tools for targeting evolutionary processes in reserve design: An approach for islands. *PLOS ONE* 13, e0200830
- 16 Paz-Vinas, I. *et al.* (2018) Systematic conservation planning for intraspecific genetic diversity. *Proc. R. Soc. B Biol. Sci.* 285, 20172746
- 17 Rosauer, D.F. *et al.* (2018) Real-world conservation planning for evolutionary diversity in the Kimberley, Australia, sidesteps uncertain taxonomy. *Conserv. Lett.* 11,
- 18 Hanson, J.O. *et al.* (2019) Conventional methods for enhancing connectivity in conservation planning do not always maintain gene flow. *J. Appl. Ecol.* 56, 913–922
- 19 Diniz-Filho, J.A.F. *et al.* (2020) Overcoming the worst of both worlds: integrating climate change and habitat loss into spatial conservation planning of genetic diversity in the Brazilian Cerrado. *Biodivers. Conserv.* 29, 1555–1570
- 20 Hanson, J.O. *et al.* (2020) Conservation planning for adaptive and neutral evolutionary processes. *J. Appl. Ecol.* 57, 2159–2169
- 21 Nielsen, E.S. *et al.* (2020) A comparison of genetic and genomic approaches to represent evolutionary potential in conservation planning. *Biol. Conserv.* 251, 108770
- 22 Hanson, J.O. *et al.* (2021) Evaluating surrogates of genetic diversity for conservation planning. *Conserv. Biol.* 35, 634–642
- 23 Phair, N.L. *et al.* (2021) Applying genomic data to seagrass conservation. *Biodivers. Conserv.* 30, 2079–2096
- 24 von Takach, B. *et al.* (2021) Population genomics and conservation management of a declining tropical rodent. *Heredity* 126, 763–775
- 25 Xuereb, A. *et al.* (2021) Incorporating putatively neutral and adaptive genomic data into marine conservation planning. *Conserv. Biol.* 35, 909–920
- 26 Nei, M. (1987) *Molecular Evolutionary Genetics*, Columbia University Press.
- 27 Ferrier, S. *et al.* (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib.* 13, 252–264
- 28 Phillips, S.J. *et al.* (2006) Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259
- 29 Pritchard, J.K. *et al.* (2000) Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959
- 30 Foll, M. and Gaggiotti, O. (2006) Identifying the Environmental Factors That Determine the Genetic Structure of Populations. *Genetics* 174, 875–891
- 31 Wilson, G.A. and Rannala, B. (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163, 1177–1191
- 32 Tarroso, P. *et al.* (2019) Phylin 2.0: Extending the phylogeographical interpolation method to include uncertainty and user-defined distance metrics. *Mol. Ecol. Resour.* 19, 1081–1094
- 33 Gower, J.C. (1971) A General Coefficient of Similarity and Some of Its Properties. *Biometrics* 27, 857–871
- 34 Jost, L. (2008) G_{ST} and its relatives do not measure differentiation. *Mol. Ecol.* 17, 4015–4026
- 35 Hoef, J.M.V. *et al.* (2014) SSN: An R Package for Spatial Statistical Modeling on Stream Networks. *J. Stat. Softw.* 56,
- 36 Peterson, E.E. and Hoef, J.M.V. (2014) STARS: An ArcGIS Toolset Used to Calculate the Spatial Information Needed to Fit Spatial Statistical Models to Stream Network Data. *J. Stat. Softw.* 56,
- 37 Rosauer, D.F. *et al.* (2015) Lineage Range Estimation Method Reveals Fine-Scale Endemism Linked to Pleistocene Stability in Australian Rainforest Herpetofauna. *PLOS ONE* 10, e0126274
- 38 Manel, S. *et al.* (2018) Predicting genotype environmental range from genome-environment associations. *Mol. Ecol.* 27, 2823–2833
- 39 Bonin, A. and Bernatchez, L. (2009) Challenges in assessing adaptive genetic diversity: Overview of methods and empirical illustrations. In *Population Genetics for Animal Conservation* (Bertorelle, G. *et al.*, eds), pp. 123–147, Cambridge University Press

- 40 Faith, D.P. and Walker, P.A. (1996) Environmental diversity: on the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodivers. Conserv.* 5, 399–415
- 41 Ball, I.R. *et al.* (2009) Marxan and relatives: Software for spatial conservation prioritisation. In *Spatial conservation prioritization: Quantitative methods and computational tools* (Moilanen, A. *et al.*, eds), pp. 185–195, Oxford University Press
- 42 Moilanen, A. *et al.* (2009) The Zonation framework and software for conservation prioritization. In *Spatial conservation prioritization: Quantitative methods and computational tools* (Moilanen, A. *et al.*, eds), pp. 196–210, Oxford University Press
- 43 Hanson, J.O. *et al.* (2021) *prioritizr: Systematic Conservation Prioritization in R. R package version 7.0.1.* <https://prioritizr.net/>,
- 44 Sarkar, S. *et al.* (2009) The ConsNet software platform for systematic conservation planning. In *Spatial conservation prioritization: Quantitative methods and computational tools* (Moilanen, A. *et al.*, eds), pp. 235–248, Oxford University Press
- 45 Coello-Coello, C. *et al.* (2007) *Evolutionary Algorithms for Solving Multi-Objective Problems*, (2nd edn) Springer US.