

Systematic conservation planning for intraspecific genetic diversity

Ivan Paz-Vinas, Géraldine Loot, Virgilio Hermoso, Charlotte Veyssière, Nicolas Poulet, Gaêl Grenouillet, Simon Blanchet

▶ **To cite this version:**

Ivan Paz-Vinas, Géraldine Loot, Virgilio Hermoso, Charlotte Veyssière, Nicolas Poulet, et al.. Systematic conservation planning for intraspecific genetic diversity. Proceedings of the Royal Society B: Biological Sciences, 2018, 285 (1877), pp.2-10. 10.1098/rspb.2017.2746 . hal-01797434

HAL Id: hal-01797434
https://univ-lyon1.hal.science/hal-01797434

Submitted on 25 Apr 2019

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.



Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: http://oatao.univ-toulouse.fr/
Eprints ID: 23483

To link to this article: DOI: 10.1098/rspb.2017.2746
URL: http://dx.doi.org/10.1098/rspb.2017.2746

To cite this version: Paz-Vinas, Ivan and Loot, Geraldine and Hermoso, Virgilio and Veyssière, Charlotte and Poulet, Nicolas and Grenouillet, Gaël and Blanchet, Simon *Systematic conservation planning for intraspecific genetic diversity*. (2018) Proceedings of the Royal Society B: Biological Sciences, 285 (1877). 2-10. ISSN 0962-8452

Any correspondence concerning this service should be sent to the repository administrator:
staff-oatao@listes-diff.inp-toulouse.fr

Systematic conservation planning for intraspecific genetic diversity

Ivan Paz-Vinas[1,2,3], Géraldine Loot[1,4], Virgilio Hermoso[5], Charlotte Veyssière[1], Nicolas Poulet[6], Gaël Grenouillet[1,4] and Simon Blanchet[1,7]

[1]CNRS, UPS, IRD; UMR-5174 EDB, Université de Toulouse, 118 route de Narbonne, 31062 Toulouse cedex 4, France
[2]Aix-Marseille Université, CNRS, IRD, Avignon Université; UMR-7263 IMBE, 3 place Victor Hugo, 13331 Marseille cedex 3, France
[3]CNRS, ENTPE; UMR-5023 LEHNA, Université de Lyon, 6 rue Raphaël Dubois, 69622 Villeurbanne, France
[4]Institut Universitaire de France, Paris, France
[5]Centre Tecnologic Forestal de Catalunya, Crta. Sant Llorenc de Monunys, Km 2, 25280 Solsona, Lleida, Spain
[6]French Biodiversity Agency, pôle écohydraulique, Allée du professeur Camille Soula, 31400 Toulouse, France
[7]CNRS, Station d'Écologie Théorique et Expérimentale, UMR-5321, 09200 Moulis, France

Intraspecific diversity informs the demographic and evolutionary histories of populations, and should be a main conservation target. Although approaches exist for identifying relevant biological conservation units, attempts to identify priority conservation areas for intraspecific diversity are scarce, especially within a multi-specific framework. We used neutral molecular data on six European freshwater fish species (*Squalius cephalus*, *Phoxinus phoxinus*, *Barbatula barbatula*, *Gobio occitaniae*, *Leuciscus burdigalensis* and *Parachondrostoma toxostoma*) sampled at the riverscape scale (i.e. the Garonne-Dordogne river basin, France) to determine hot- and coldspots of genetic diversity, and to identify priority conservation areas using a systematic conservation planning approach. We demonstrate that systematic conservation planning is efficient for identifying priority areas representing a predefined part of the total genetic diversity of a whole landscape. With the exception of private allelic richness (PA), classical genetic diversity indices (allelic richness, genetic uniqueness) were poor predictors for identifying priority areas. Moreover, we identified weak surrogacies among conservation solutions found for each species, implying that conservation solutions are highly species-specific. Nonetheless, we showed that priority areas identified using intraspecific genetic data from multiple species provide more effective conservation solutions than areas identified for single species or on the basis of traditional taxonomic criteria.

1. Background

Biodiversity conservation is often addressed by identifying protected areas with high biodiversity and/or landscape values [1]. Conservation areas are generally identified as areas with high proportions of endemic, rare or iconic species [2]. Alternatively, conservation planning can be based on the concept of complementarity between conservation areas [3], and on cost-effectiveness analyses such as systematic conservation planning procedures (SCP [4]). SCP aims at identifying a number of complementary and irreplaceable sites best representing a predefined proportion of the biodiversity observed in a region and that should be managed for conservation in priority at a minimum cost.

There have been major conceptual and empirical leaps advocating the importance of considering species evolutionary history into conservation plans [5,6]. For instance, some authors included information on the phylogenetic history of species assemblages in SCP to preserve both species identities and their macro-evolutionary history [7–9]. However, genetic diversity observed at the population level (within species) has rarely been considered

in SCP (but see [10–13]). Intraspecific genetic diversity (IGD) is a fundamental biodiversity facet, as it informs the demographic and colonization history of populations (i.e. changes in effective population size over time and space), their level of biological connectivity, and their potential for adapting to environmental changes [14–16]. Conservation geneticists generally use neutral genetic diversity (i.e. genome portions unaffected by adaptive processes) to identify relevant evolutionary units within species [17], mainly by analysing genetic differentiation among populations [18,19]. For instance, the 'evolutionarily significant units' (ESUs) framework has been proposed for identifying different lineages for conservation prioritization [20,21]. By emphasizing complementarity among areas—rather than differentiation—SCP appears as a promising and complementary framework for identifying optimal sets of sites preserving predefined amounts of the total IGD observed within landscapes.

Most SCP studies yet to consider IGD aimed at identifying efficient surrogates for preserving genetic diversity [12,13]. For instance, two recent studies showed that protected areas identified using traditional species distribution data [13], and environmental and geographical descriptors [12] can be efficient surrogates in conservation planning for representing IGD. Although these studies shed light on how SCP can help preserve the IGD of multiple species, only a few studies have directly used multi-specific genotypic data for SCP (but see [11]). Datasets generated using neutral codominant genetic markers such as microsatellite markers are similar to datasets generally used for SCP (species presence/absence matrices), hence making possible applying SCP to genotypic surveys without needing intermediate modelling procedures or the calculation of summary genetic statistics (e.g. [10,12]) that may fail to capture the whole IGD of populations.

Here, we tested the potential of SCP to identify priority conservation areas accounting for IGD of a species assemblage at the landscape scale. We first considered a set of four common freshwater fish species (the chub *Squalius cephalus*, the gudgeon *Gobio occitaniae*, the stone loach *Barbatula barbatula* and the minnow *Phoxinus phoxinus*) to test the influence of conservation targets (proportion of the total amount of IGD to be secured by irreplaceable sites) and analytical strategies (analysing each species independently or all species pooled) on final conservation solutions (number and identity of irreplaceable sites). We then included two rare species of particular conservation interest (the beaked dace *Leuciscus burdigalensis* and the toxostome *Parachondrostoma toxostoma*) to test the relevance of the SCP approach in a 'real conservation-oriented study'. For all species, we ran SCP analyses considering a typical conservation target [22] to explore the spatial distribution of irreplaceable sites in the riverscape, and to test for surrogacy in irreplaceable sites among species. We further tested whether irreplaceable sites identified using SCP are more efficient for preserving IGD than randomly selected sites or than sites selected using traditional taxonomic criteria (e.g. sites displaying high species richness). We finally tested whether irreplaceable sites were correctly predicted by classical genetic diversity indices. We demonstrate that preserving the IGD of a species assemblage using systematic conservation planning is an efficient, yet complex approach necessitating appropriate analyses to assist the decision-making process.

2. Material and methods

(a) Biological models

We focused on six freshwater fish species: *Squalius cephalus*, *Phoxinus phoxinus* and *Barbatula barbatula*, which are widespread in Europe, and *Gobio occitaniae*, *Leuciscus burdigalensis* and *Parachondrostoma toxostoma*, which are endemic to southern France [23]. These species were all from the Cyprinidae family apart from *B. barbatula*, which is a Nemacheilidae. This species assemblage covers a large functional trait space representative of many freshwater fish communities. For instance, *S. cephalus* is a large-bodied species with a long lifespan (i.e. up to 60 cm long and 15 years old [23]), whereas *Ph. phoxinus* is a small-bodied species with a shorter lifespan (i.e. less than 12 cm long and 4–5 years old [23]). From an ecological perspective, *G. occitaniae*, *Pa. toxostoma* and *B. barbatula* are bottom feeders, whereas *S. cephalus* and *Ph. phoxinus* are water column feeders. *Leuciscus burdigalensis* is more opportunistic. Further, *B. barbatula* is mainly active during the night, while the other species are active during the day. Four of these species are abundant (*S. cephalus*, *G. occitaniae*, *Ph. phoxinus* and *B. barbatula*), although their ecological niche and spatial occupancy in the Garonne-Dordogne river basin varies (see electronic supplementary material, figure S1). We will hereafter refer to this set of species as the 'common' species. The two other species are rare (*L. burdigalensis*) to very rare (*Pa. toxostoma*; electronic supplementary material, figure S1) in the riverscape, and are of particular interest for conservation. *Leuciscus burdigalensis* is a recently described species that is experiencing both demographic and genetic bottlenecks in many populations [24,25]. *Parachondrostoma toxostoma* is a vulnerable species [26] listed in the IUCN Red List, in Annex II of the European Union Habitats Directive and in Appendix III of the Bern Convention [26].

(b) Sampling design

During Spring/Summer 2010–2011, we used electrofishing to sample 92 sites distributed across 35 rivers from a whole European river basin (the Garonne-Dordogne river basin, southwestern France; greater than 100 000 km^2; electronic supplementary material, figure S1). Sampling sites were chosen to cover the entire distribution of each species in the riverscape, and to allow characterizing their spatial patterns of IGD. Up to 25 individuals per species per site were sampled when possible. Not all species were present at all sampling sites (electronic supplementary material, figure S1 and table S1), and not all sites provided 25 individuals per species due to low densities. In these cases, we captured as many individuals as possible. We anaesthetized each individual and then collected and stored in 90% ethanol a fragment of their pelvic fin. All individuals were released alive at their sampling location.

(c) Genotyping

Genomic DNA was extracted using a salt-extraction protocol [27]. We used multiplexed polymerase chain reactions to co-amplify 8–15 microsatellite loci depending on the species (8 for *G. occitaniae*, 9 for *B. barbatula*, 10 for *S. cephalus* and *Ph. phoxinus*, 14 for *L. burdigalensis* and 15 for *Pa. toxostoma* (see electronic supplementary material, appendix A1 for details on recipes).

(d) Genetic diversity assessment

Given the good spatial resolution of the sampling obtained for the common species, descriptive genetic analyses were conducted for sampling sites displaying minimum sample sizes of $n = 10$ individuals for these species to maximize consistency on subsequent allelic frequency-based genetic analyses. For the two rare species, for which the sampling was more restricted,

genetic analyses were conducted for sampling sites displaying minimum sample sizes of $n = 6$ individuals to maximize the number of sampling sites included in SCP. Reducing the minimum sample size from 10 to 6 permitted increasing considerably the number of sites retained for genetic analyses (from 21 to 29 sites and from 8 to 14 sites for *L. burdigalensis* and *Pa. toxostoma*, respectively), hence improving the representation of their spatial distribution. We then determined for all six species the occurrence of null alleles, potential scoring errors, departures from Hardy–Weinberg equilibrium and linkage disequilibria among loci within sites (see electronic supplementary material, appendix A2 for details concerning these quality controls).

To assess within-site IGD (α-IGD), we applied the rarefaction procedures implemented in ADZE [28] to calculate both allelic richness (AR) and private allelic richness (PA) per sampling site (considering a minimum of $n = 10$ individuals for common species or $n = 6$ individuals for rare species). To assess among-site IGD (β-IGD), we used the R package 'mmod' [29] to calculate—for each species—a pairwise genetic differentiation index (D_{est} [30]). For each site (and species), we then derived the averaged value of all pairwise D_{est} values estimated between one given site and all the remaining sites to obtain a single value per site. For rare species, the impact of considering sites with $n \geq 6$ individuals instead of $n \geq 10$ on AR, PA and D_{est} was low (no clear impact neither on PA nor on D_{est}, and mean decreases in AR of 0.295 and 1.183 alleles per site for *Pa. toxostoma* and *L. burdigalensis*, respectively). This impact was however overcompensated by the increase in the number of sites retained for rare species when considering sites with $n \geq 6$ individuals.

(e) Testing the suitability of systematic conservation planning procedures for intraspecific genetic diversity

In this first step, we tested the influence of conservation targets and analytical strategies on conservation solutions. We focused on data from the four common species, as their large coverage of the sampling area is more suited for the demonstrative exercise done in this step.

(i) Spatial patterns of intraspecific genetic diversity

We first used geostatistical modelling tools to explore spatial patterns of α- and β-IGD for the four common species in the riverscape by predicting AR, PA and D_{est} distributions from the observed empirical values using generalized linear models for spatial stream networks (GLMSSN) [31,32]. These models were built by assuming three geographical descriptors (i.e. topological distance from the outlet, longitude and latitude) as explanatory variables (see electronic supplementary material, appendix A3 for further details). We used predictions from the best models (obtained through an Akaike information criteria model comparison) to produce kriged maps for each common species and each genetic index (electronic supplementary material, Appendix A3). This allowed us to represent the spatial distribution of IGD across the whole river drainage, and hence to visually highlight hot- and coldspots of IGD. Finally, we calculated Pearson's correlation coefficients between AR, PA and D_{est} values calculated at the site level for each pair of common species to test for spatial congruence in IGD patterns among species.

(ii) Identification of irreplaceable sites

We tested whether conservation targets (i.e. the percentage of the total number of conservation units present in conservation solutions) and analytical approaches (i.e. species-specific or species-pooled analyses) influence the identification of irreplaceable sites. SCP methods traditionally use species presence/absence data as input data to identify irreplaceable sites (i.e. sites that cannot be excluded from an optimal selection of sites for conservation [33]) for the conservation of taxonomic diversity at the community level. Here, we replaced species presence/absence data with allele presence/absence data as in [11] to identify irreplaceable sites for the conservation of IGD of each species at the riverscape scale. We used MARXAN v. 2.1 [33] and genotypic data from common species to identify, at the species level, an optimal set of sites best representing at least 50, 75, 90 or 100% of the total number of alleles present in the whole riverscape at a minimum 'cost' (see electronic supplementary material, appendix A4 for further details). We then tested and compared visually how the proportions of irreplaceable sites vary among species and conservation targets.

To test how pooling data from several species affects the identification of irreplaceable sites, we further performed a 'pooled' analysis, in which all alleles found for each common species at a site were pooled together in a single input dataset. We then selected random pulls of 30, 50, 75, 90 and 100% of the total number of alleles existing at the riverscape level (all common species confounded), and performed 100 MARXAN runs per conservation target to identify the minimum set of sites representing these particular selections of alleles.

(f) A real conservation-oriented study using systematic conservation planning procedures

In this second step, we (i) explored the spatial distribution of irreplaceable sites, (ii) tested whether conservation solutions found for each species are congruent among species, (iii) tested whether solutions found for one species can be used as a surrogate for other species, (iv) assessed the efficiency of the different conservation solutions we identified for preserving the overall IGD of all species, (v) compared the efficiency of conservation solutions identified through SCP to the efficiency of alternative conservation solutions based on traditional taxonomic criteria (e.g. in sites with high species richness) and finally (vi) tested whether irreplaceable sites are correctly predicted by classical genetic diversity indices. In this step, we included the two rare species, because surrogacy is particularly important to measure for rare species for which data are difficult to collect. We therefore focus more specifically on common versus rare species comparisons.

(i) Identification of irreplaceable sites for genetic conservation

We here focused only on irreplaceable sites identified for the 90% target and used the program MARXAN as described above to identify these sites for *L. burdigalensis* and *Pa. toxostoma* independently. We then mapped these irreplaceable sites on the river network to test (i) whether specific areas harboured more irreplaceable sites (e.g. upstream areas that are generally thought to be of high conservation priority [34]) and (ii) if irreplaceable sites are spatially congruent among species and, notably, among common and rare species. Additionally, we ran generalized linear models (GLMs) including whether a sampling site is irreplaceable at the 90% target as a binomial-dependent variable, and distance to the outlet of sites and per site betweenness centrality values [35] as explanatory variables. Betweenness centrality is an index quantifying the positional importance of each sampling site within the riverscape [36,37]. We assumed binomial error terms distributions and tested the significance of each term at the $\alpha = 0.05$ threshold.

(ii) Surrogacy in irreplaceable sites among species

We estimated surrogacy levels among species by calculating the percentage of the total number of alleles observed for a given species secured by irreplaceable sites identified for another species. Although surrogacy was calculated for all species pairs, we specifically focused on rare species by calculating (i) the percentage of the total number of alleles observed for rare species secured by irreplaceable sites identified for all common species and (ii) the percentage of the total number of alleles observed for common species secured by irreplaceable sites identified for all rare species.

(iii) Assessing the efficiency of systematic conservation planning procedures

The goals of this analysis were (i) to assess the efficiency of SCP for preserving the overall levels of IGD observed in the riverscape for the six species, and (ii) to compare the efficiency of SCP-based conservation solutions to the efficiency of alternative conservation solutions based on traditional taxonomic criteria. We first computed an average accumulation curve of 'protected genetic diversity' by estimating the average cumulative percentage of the total number of alleles (over all species) secured by a cumulative number of randomly selected sampling sites. Specifically, we increased the number of sampled sites from 1 to 92 with a continuous increment of one unit. At each incremental step, we randomly selected sites (without replacement) and calculated the percentage of the total number of alleles (over all species) secured by the random set of sites. We repeated this random selection 1000 times and averaged the percentages over these 1000 permutations to obtain the averaged accumulation curve of protected genetic diversity. We then tested to what extent, for an identical number of sites, the conservation solutions we identified through SCP (i.e. irreplaceable sites) secured higher percentages of IGD than those secured by a random selection of sites. Notably, we explored whether certain conservation solutions (i.e. by considering irreplaceable sites identified for single species or for the sets of common, rare and all six species) were more efficient than others to sustain a large percentage of IGD.

As in most conservation studies species diversity is the main conservation target, we finally checked the efficiency of conservation solutions based on taxonomic criteria for preserving the overall IGD of the six studied species. Specifically, we calculated the percentages of total number of alleles (over all species) secured by sites (i) in which a number equal or greater than 10, 15 or 20 freshwater fish species co-occur (the total number of fish species observed in the riverscape being 34, with a mean number of species observed at the site level of 12.77 ± 0.55), and (ii) in which the six studied species co-occur. We finally compared the efficiencies of taxonomic-data-based and SCP-based conservation solutions to assess if solutions based on taxonomic criteria are good surrogates of the overall IGD of all species. We derived data on fish species occurrence from [37] for 81 of our 92 sampling sites.

(iv) Relationships between irreplaceable sites and genetic diversity indices

To test the ability of classical genetic indices to predict the propensity of sites to be irreplaceable, we ran for each species GLMs (considering binomial error terms distributions) including whether sites were designated as irreplaceable at the 90% target as a binomial-dependent variable, and AR, PA and D_{est} as explanatory variables. We tested the significance of each term at the $\alpha = 0.05$ threshold. Explanatory variables were centred and scaled to compare the relative strength of the predictors among species.

3. Results

(a) Descriptive statistics

Overall, *Pa. toxostoma* showed the lowest within-site IGD estimates. Mean AR ranged from 2.114 for *Pa. toxostoma* to 5.821 for *Ph. phoxinus*, and mean PA ranged from 0.036 for *Pa. toxostoma* to 0.162 for *L. burdigalensis* (electronic supplementary material, table S2). *Parachondrostoma toxostoma* also showed the lowest mean D_{est} value (0.069), while *B. barbatula* displayed the highest mean value (0.383; electronic supplementary material, table S2).

(b) Testing the suitability of systematic conservation planning procedures for intraspecific genetic diversity

(i) Spatial patterns of intraspecific genetic diversity

Spatial patterns of IGD largely varied and were poorly congruent among common species (figure 1). For instance, hotspots of AR for *S. cephalus* were mainly found in the western part of the riverscape and on core streams, whereas these same areas were identified as coldspots of AR for *B. barbatula* (figure 1*a*(i)–*a*(iii)). Similarly, PA hotspots were inversely related between *G. occitaniae* and *Ph. phoxinus* (figure 1*b*(ii)–*b*(iv)). Similar conclusions were reached for D_{est} (figure 1*c*). For instance, D_{est} hotspots were observed in opposite riverscape areas for the species pair *B. barbatula*/*Ph. phoxinus* (figure 1*c*(iii)–*c*(iv)). Consequently, the sign, slope and significance of GLMSSNs explanatory variables strongly varied among species (figure 1). This incongruence in spatial patterns of IGD among species was also reflected by the low-to-moderate correlation coefficients measured among all possible species pairs and for each genetic diversity index (i.e. AR, PA and D_{est}; electronic supplementary material, table S3). Indeed, Pearson's correlation coefficients were lower than 0.6 for all comparisons but two (i.e. between *B. barbatula*/*G. occitaniae* and between *Ph. phoxinus*/*B. barbatula* for AR; electronic supplementary material, table S3).

(ii) The influence of conservation targets and analytical strategies to identify irreplaceable sites for genetic conservation

When species were analysed independently, we found that the number of irreplaceable sites increased as the conservation target increased, with a steep increase for conservation targets higher than 75% of the total number of alleles (figure 2). However, the percentage of irreplaceable sites strongly varied among species. For instance, for a 90% target, the proportion of irreplaceable sites ranged from 3.61% of the total number of sampled sites for *G. occitaniae* to 28.57% for *Ph. phoxinus* (electronic supplementary material, table S4; figure 2). For extreme conservation targets (100% of alleles to be secured), the proportion of irreplaceable sites varied from 25.30% for *G. occitaniae* to 68.26% for *Ph. phoxinus* (electronic supplementary material, table S4; figure 2).

When alleles from the four common species were analysed in a single pooled analysis, we similarly found that the proportion of irreplaceable sites increases as the conservation target increases (figure 2). Interestingly, no irreplaceable sites were identified for the 30% target, and only three irreplaceable sites were found for the 50% target (figure 2).

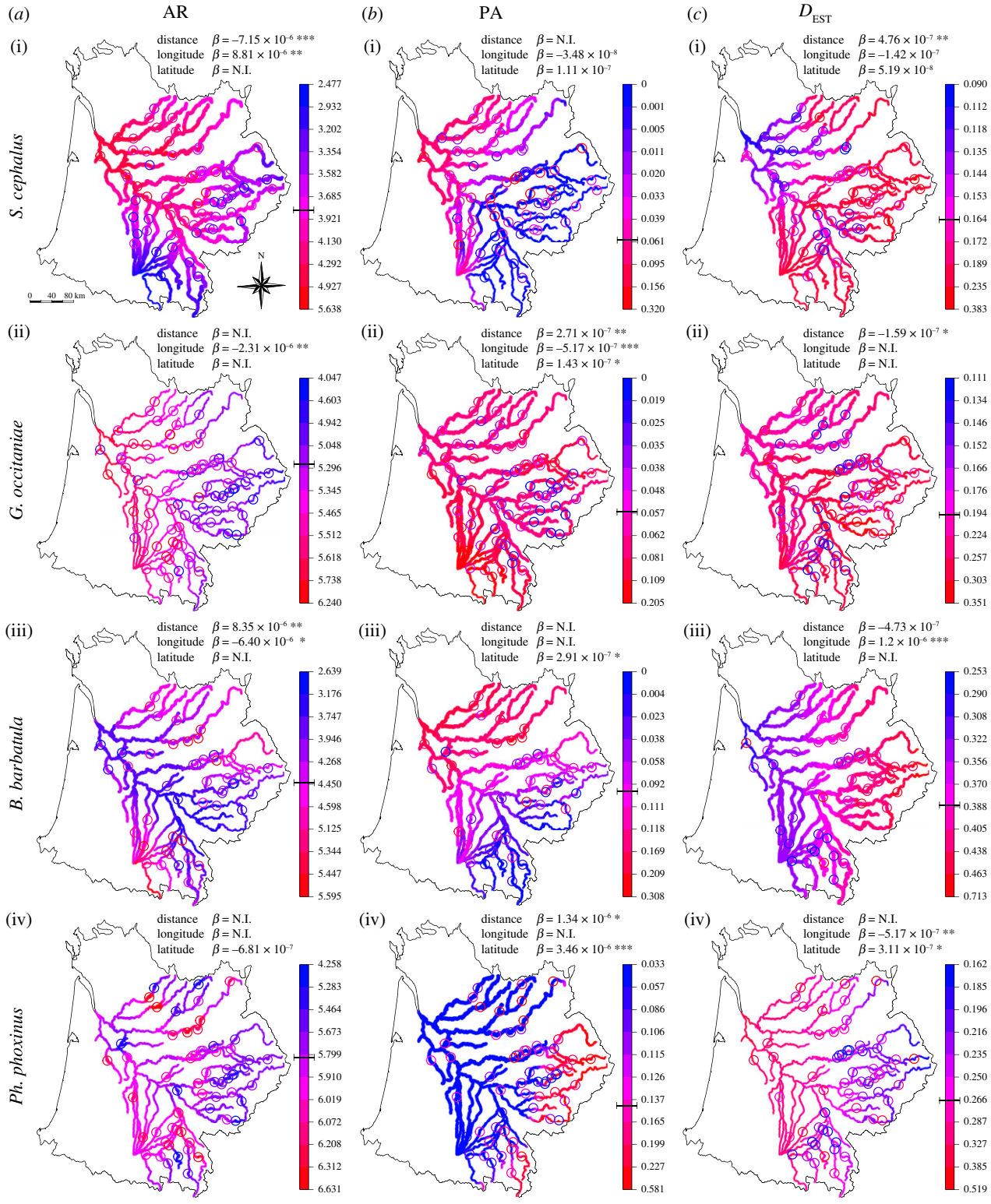


Figure 1. Spatial distribution of observed (coloured circles) and interpolated (coloured lines) values of (*a*) AR, (*b*) PA and (*c*) D_{est} for (i) *Squalius cephalus*, (ii) *Gobio occitaniae*, (iii) *Barbatula barbatula* and (iv) *Phoxinus phoxinus* obtained with GLMSSNs. Coloured lines width is inversely related to prediction standard error. The cursor on the vertical coloured scale indicates mean AR, PA and D_{est} values. The slope (β) of each explanatory variable (topological distance to the outlet, longitude and latitude) and its significance are also reported. N.I. indicates that the explanatory variable was not included in the model; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

The proportion of irreplaceable sites increased moderately to 17.39% for the 75% target, and then steeply increased for higher conservation targets to reach 55.70% for the 90% target and 76.08% for the 100% target (figure 2). This later result suggests that almost all the riverscape should be protected to reach high conservation targets when adopting a pooled strategy.

(c) A real conservation-oriented study using systematic conservation planning procedures

(i) Identification of irreplaceable sites for genetic conservation

We first visually explored the spatial distribution of irreplaceable sites. Overall, the localization of irreplaceable sites in the riverscape strongly varied among species (electronic

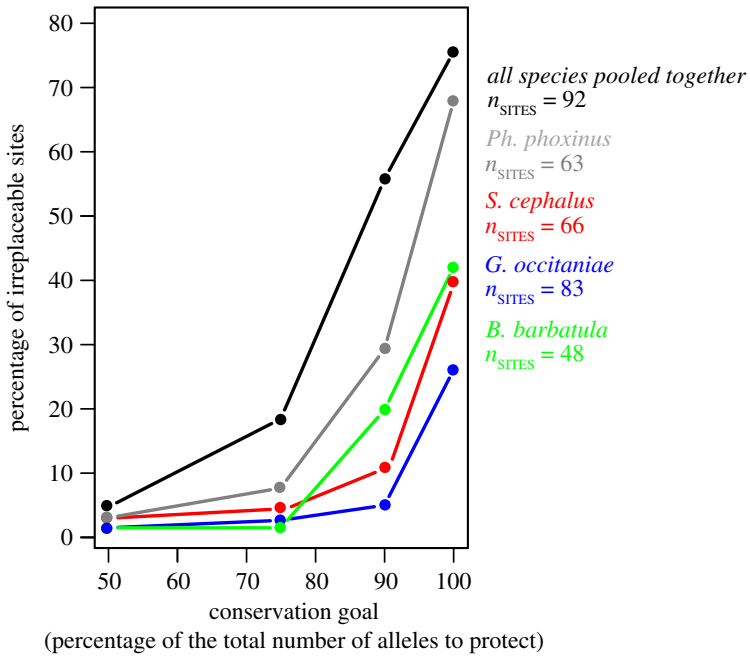


Figure 2. Percentage of irreplaceable sites identified by MARXAN for conservation targets of 50, 75, 90 and 100% of the total number of alleles present in the riverscape for each common species and for a pooled analysis in which all alleles from all common species were pooled together. n_{SITES} represents the number of sites included in MARXAN analyses.

supplementary material, figure S2). We failed to identify areas (e.g. upstream/downstream locations) clustering irreplaceable sites for any species (electronic supplementary material, figure S2A–F). This apparent lack of clustering was statistically confirmed by GLMs (electronic supplementary material, table S5). Indeed, neither the distance from the outlet of sampling sites nor their betweenness centralities were significant predictors of sites irreplaceability for all species, except for a significant effect of distance from the outlet for *Ph. phoxinus* (electronic supplementary material, table S5).

Second, we tested whether conservation solutions found for each species were spatially congruent among species. Over the six fish species, 42 sites (45.65%) were irreplaceable at the 90% conservation target for at least one species (figure 3). Thirty-two of these 42 sites were irreplaceable for at least one common species (figure 3), and 14 of the 42 sites were irreplaceable for at least one rare species (figure 3). Among the six species, only eight of these 42 sites were irreplaceable for at least two species, and only one site was irreplaceable for three species (figure 3).

(ii) Surrogacy in irreplaceable sites among species

Surrogacy levels were generally low to moderate, and strongly varied among species pairs (electronic supplementary material, table S6). For instance, irreplaceable sites identified for *G. occitaniae* failed to secure rare species IGD, and covered only 18–43% of the total number of alleles of the other common species. This indicates that irreplaceable sites identified for *G. occitaniae* (the most widespread species) are poor surrogates for preserving the IGD of other species (electronic supplementary material, table S6). Conversely, irreplaceable sites found for *Ph. phoxinus* and *L. burdigalensis* are better surrogates for *G. occitaniae*, as 79.58% and 70.27% of the total number of alleles of *G. occitaniae* was secured by irreplaceable sites identified for *Ph. phoxinus* and *L. burdigalensis*, respectively (electronic supplementary material, table S6). Overall, irreplaceable sites best sustaining other species IGD were those identified for *B. barbatula*, which secured in average 68.28% of the total number of alleles of other species (electronic supplementary material, table S6).

Interestingly, the 32 irreplaceable sites identified for common species secured 81.79% and 90% of the total number of alleles of *L. burdigalensis* and *Pa. toxostoma*, respectively, suggesting that irreplaceable sites identified for a set of common species can be good surrogates for rare species IGD. The 14 irreplaceable sites identified for the two rare species secured a total number of alleles ranging from 65.18% for *S. cephalus* to 79.40% for *B. barbatula*, which is not negligible given that conservation priorities are generally directed towards rare species.

(iii) Assessing the efficiency of systematic conservation planning procedures

Overall, irreplaceable sites identified for single species using SCP did not significantly secured higher percentages of IGD than randomly selected sites, excepting irreplaceable sites identified for *L. burdigalensis* and *B. barbatula* (figure 4), which were more efficient than randomly selected sites for preserving the overall IGD of all species. However, we found significantly higher percentages of secured alleles (compared to randomly selected sites) on irreplaceable sites identified through SCP for (i) the set of common species (ii) the set of rare species and (iii) all the six species (figure 4).

Conservation solutions based on the co-occurrence of high numbers of species (i.e. sites with numbers of co-occurring species equal or greater than 10, 15 or 20) did not secured more alleles than randomly selected sites, suggesting that raw interspecific diversity (i.e. species richness) may not be a good surrogate of IGD in our study (figure 4). Nevertheless, we found that a higher percentage of alleles would be secured by selecting sites on which the six species co-occur (compared to randomly selected sites; figure 4).

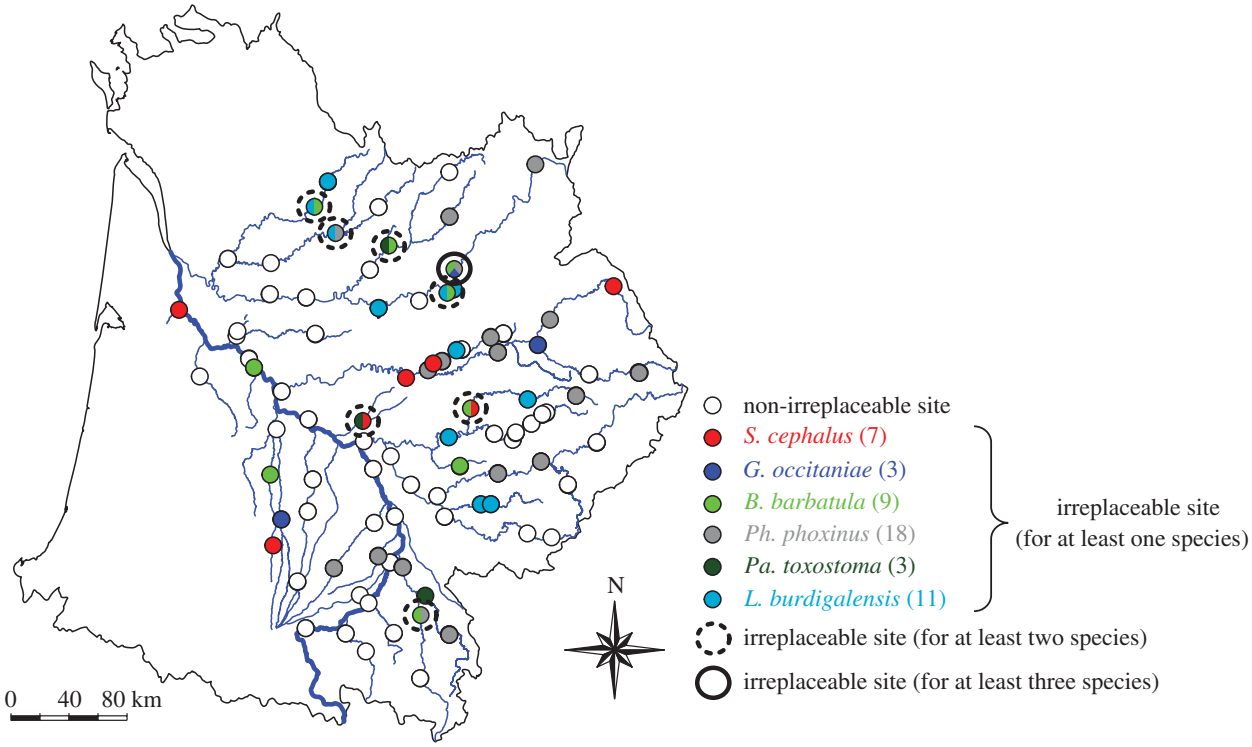


Figure 3. Irreplaceable sites identified by MARXAN for at least one (unicoloured-filled points), two (black dotted circles surrounding bicoloured points) or three (black bolded circle surrounding a tricoloured point) species, assuming a conservation target of 90% of the total number of alleles present in the riverscape.

(iv) Relationships between genetic indices and irreplaceable sites

For all species but *G. occitaniae* and *Pa. toxostoma* (for which no variables were significant predictors), PA was identified as the only variable significantly predicting the probability of sites to be irreplaceable (electronic supplementary material, table S7). This probability increased with the number of PA in a site.

4. Discussion

Intraspecific genetic diversity reflects the evolutionary and demographic history of populations, and hence sustains the capacity of species and populations to cope with environmental changes [14–16]. This biodiversity facet is the first that should respond to global change [15], and it is hence surprising that IGD has yet been rarely integrated in dedicated optimization planning tools. We here fill this gap by demonstrating that systematic conservation planning for IGD is an efficient, yet complex approach, necessitating careful considerations.

(a) From idiosyncratic intraspecific genetic diversity distributions . . .

Our results suggest that the IGD of the targeted species did not follow a common spatial pattern, but rather species-specific (idiosyncratic) spatial distributions. This conclusion holds true for all genetic diversity indices, and corroborates the few previous studies investigating simultaneously the spatial distribution of IGD at large spatial scales and for sympatric species [38,39]. This was, however, unexpected given that recent meta-analyses on freshwater organisms demonstrated that α-IGD is generally higher in downstream than in upstream areas [40], and that β-IGD tends to be higher in upstream than in downstream sections [34]. Overall, these general spatial patterns of IGD were verified in our datasets. For instance, a negative relationship between AR and distance from the outlet is expected in freshwater organisms [40], and was actually observed for three of the six studied species (electronic supplementary material, figure S3). However, when using a precise and novel approach to map genetic diversity across the network, we demonstrated that the distribution of cold- and hotspots of α- and β-IGD was subtler and idiosyncratic. This probably reflects interactions between recent colonization histories, species life-history traits and the network structure, which are expected to drive neutral IGD patterns in rivers [40–42].

(b) . . . to systematic conservation planning for intraspecific genetic diversity

The spatial mismatch in IGD among species probably explains why we found that the surrogacy level among irreplaceable sites identified for each species was extremely low. For instance, an extremely low proportion of irreplaceable sites was common to two or three species (and never more than three species; figure 4). In the same way, we detected no clear patterns in the spatial distribution of irreplaceable sites. Indeed, irreplaceable sites (for any of the six species) were not particularly situated in specific locations such as upstream areas and/or in areas of high connectivity (i.e. with high centrality values [36,37]), hence suggesting that priority areas for the conservation of multi-specific IGD should not be restricted to specific areas in the riverscape. It is, however, noteworthy that the greatly variable position of irreplaceable sites found for the studied species may reflect both their spatial occupancy in the riverscape and their own ecological niche. In this demonstrative study, we only

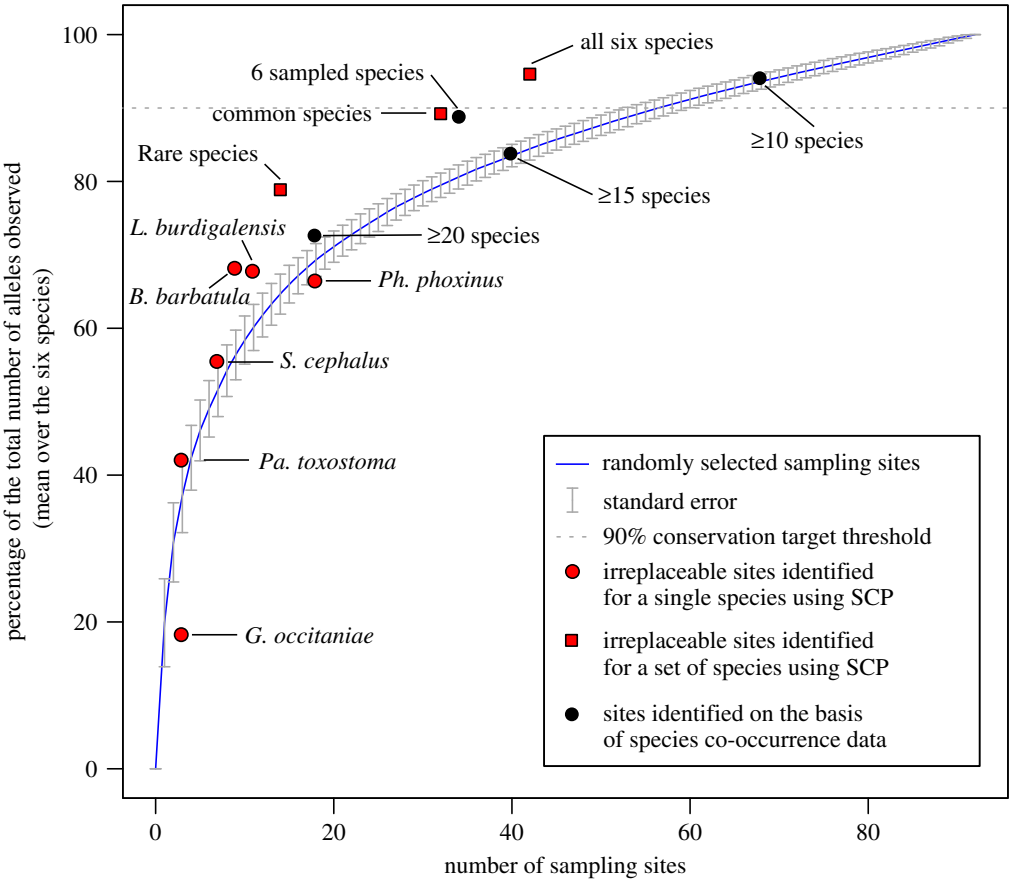


Figure 4. Mean percentage of the total number of alleles observed in the riverscape for the six species that are secured by (i) a cumulative number of randomly selected sampling sites (blue line; average accumulation curve obtained under 1000 permutations of the sampling sites; grey vertical lines: standard error), (ii) irreplaceable sites identified using SCP for each species independently (red dots), (iii) irreplaceable sites identified using SCP for different sets of species (red squares), and (iv) sites identified on the basis of freshwater fish species co-occurrence data (black dots).

focused on the influence of spatial variables (longitude, latitude, distance to the outlet, centrality) on conservation solutions. Considering other environmental variables that may directly or indirectly influence the distribution of IGD such as water temperature, oxygen concentration and/or water pollution (e.g. [43]) would certainly shed important light on the drivers of sites irreplaceability, and should hence be the focus of future SCP studies. Finally, the surrogacy level among irreplaceable sites was low to moderate, and never attained the 90% threshold we assumed when we considered all species independently. Combined with our finding that the number of irreplaceable sites can be high for reasonable conservation targets (up to 46% sites identified as irreplaceable for at least one species at the 90% conservation target), we conclude that our ability to identify priority areas for IGD may be highly species-specific and may depend on the capacity to tackle the trade-off between the amount of IGD to protect, and the extent of priority areas we can realistically protect.

However, when surrogacy between all irreplaceable sites identified for the entire set of common species and those identified for the rare species was tested, we reached reasonable proportions of the total number of alleles to protect (approx. 80–90%) for the two rare species (i.e. *Pa. toxostoma* and *L. burdigalensis*). This result suggests that, in some cases, genetic data obtained for a set of widely distributed, 'easier-to-sample' common species displaying varying life-history traits can be used for identifying protection areas for the IGD of other rare species that can be more problematic to sample. This would require, however, a good overlap between the spatial distributions of both the common species used as surrogates and the targeted rare species, and a wide sampling design allowing representing the whole spatial occupancy of species in the landscape. Failing to fulfil these two prerequisites may have a negative impact on conservation costs, as the number of irreplaceable sites necessary to preserve a predefined amount of IGD of rare species would be higher than expected. In our demonstrative study, for instance, 34 of the 42 (81%) irreplaceable sites were sites in which at least four study species co-occur, and 19 of these sites (54.3%) were sites in which the six species co-occur. It is worth mentioning that our study builds on a large-scale (but scattered) sampling design to test SCP for IGD at the riverscape scale. The efficiency of SCP for preserving IGD at finer sampling scales (e.g. by concentrating sampling sites within a restricted landscape area) remains to be tested. Another important finding we report here is that selecting conservation areas based on traditional taxonomic criteria (i.e. by selecting sites with high species richness) is not an efficient strategy for preserving the overall IGD of a set of species. This highlights the need for finding efficient ways of considering IGD in SCP, as we and others [9–13] proposed. Here, we have focused on six freshwater fish species, among which five were cyprinids. Future works should consider wider taxonomic diversities (e.g. by also considering IGD from vertebrate, invertebrate and/or plant species in SCP) to explore the distribution and drivers of site irreplaceability at a large taxonomic scale.

Overall, our results suggest that two different analytical strategies can be employed in real-case SCP studies aiming at preserving IGD of a species assemblage: (i) identifying conservation areas for a set of rare species or (ii) identifying conservation areas for a set of representative common species. Both strategies have their own advantages and inconveniences. The first strategy may optimally preserve IGD of rare species at competitive costs (e.g. only 14 irreplaceable sites to protect in our study), but at the expense of the IGD of other sympatric species. Conversely, the second strategy may optimally preserve IGD of a set of common species while maintaining high levels of IGD for rare species, but at a higher cost (32 irreplaceable sites to protect in our study). Whether to choose one of these two strategies will therefore depend on many factors such as how difficult is the sampling of rare species compared to common species, or the extent of the resources available for setting new protected areas. We recommend adopting the second strategy when possible, because it allows simultaneously maintaining IGD from rare and common species. Indeed, IGD of common species is vital for ensuring ecosystem stability, as it ultimately influences species interactions, population dynamics and ecosystem functioning [44]. It is noteworthy that further studies must be conducted to test if the two analytical strategies we propose can be applicable across taxa and landscapes.

Although it was not a main goal of the study, we identified an alternative strategy that may also be efficient for maintaining high levels of IGD for a species assemblage without requiring SCP analyses. Indeed, selecting sites where all the studied species co-occur as conservation solutions allowed maintaining high levels of IGD for both common and rare species in our study, although this strategy was less efficient than selecting irreplaceable sites found for common species through SCP. The efficiency of this strategy also needs to be thoroughly tested across taxa and landscapes, as it can be a useful surrogate strategy in some cases (e.g. when molecular data cannot be produced due to funding or technical limitations). It is, however, noteworthy that such strategy requires highly reliable taxonomic data gathered over the whole studied landscape.

We also demonstrate that microsatellite-based genotypic data can be easily used as direct input data in SCP to identify efficient conservation areas for IGD of a species assemblage at the landscape scale. Relatively small numbers of microsatellite markers such as those considered in our study (8–15) have been proved to be sufficient for capturing the confounding genetic effects of recent past demographic history, genetic drift and connectivity issues between populations [25,40,41], and their efficiency for informing recent evolutionary processes can be equal to or higher than low-to-moderate numbers of SNP markers for instance (e.g. fewer than 300 SNPs [18]). However, microsatellite markers have the disadvantage of being unsuitable for drawing conclusions on populations' adaptive potential, contrarily to other markers derived at the genome scale. Further, the sample size threshold we set (up to 25 individuals per site and species) was sufficient to properly estimate IGD of the species assemblage at the riverscape scale without implying an unreasonable sampling effort. We recall, however, that genetic diversity was probably slightly underestimated for sites where small numbers of individuals were captured (e.g. sites with $n = 6$ individuals for rare species; see the material and methods section for details). As our capacity to compile genetic datasets at large spatial, temporal and taxonomic scales greatly increased in recent decades, a next methodological step may be using high-throughput techniques to develop high numbers of markers informing both neutral and non-neutral processes to assess whether SCP solutions obtained using different markers allow efficiently preserving the whole genomic diversity of populations, and hence their whole evolutionary history.

5. Conclusion

Our study provides novel, insightful and promising knowledge on the setting of priority conservation areas for intraspecific diversity. It shows that systematic conservation planning methods are useful and efficient objective tools for conservation geneticists whose conservation solutions will strikingly depend on the species to be conserved and the quantity of genetic information that managers aim at preserving in a landscape. Given our results, we suggest that two strategies could be employed in real-case conservation programmes: (i) identifying priority conservation areas for a set of rare species when resources allocated to conservation are scarce or (ii) identifying priority conservation areas based on the analysis of a set of representative common species that may serve as 'umbrellas' for rare sympatric species when resources allocated to conservation are higher.

Our study also raises many additional questions that should be considered in the near future. We believe that the next steps will be to formally identify sound conservation targets for intraspecific diversity, to test whether neutral intraspecific diversity appropriately mirrors quantitative and adaptive diversity [16], and to quantify the influence of intraspecific diversity on ecosystem functioning and services to better evaluate the added value of preserving this biodiversity facet [16,44].

Ethics. Sampling was made in accordance with current French ethical laws (permit nos 2010-161-10, 2010-108, 2011-185-06 and 2011-318).

Data accessibility. Supplementary tables, figures and appendices are provided as electronic supplementary material. We also provide raw allele counts and genetic diversity estimates for all species, and raw outputs for Hardy–Weinberg equilibrium, linkage disequilibrium and null allele's tests.

Authors' contributions. S.B., G.L. and I.P.-V. designed the study. I.P.-V. and S.B. wrote the manuscript with the help of V.H., G.G., N.P. and G.L. I.P.-V., S.B., G.L. and C.V. collected the samples. G.L. and C.V. produced the genetic data. I.P.-V., V.H. and S.B. conducted statistical and population genetics analyses.

Competing interests. We declare we have no competing interests.

Funding. V.H. was funded by the National Environmental Research Pro-gram Northern Australia Hub and a 'Ramon y Cajal' contract (RYC-2013-13979) funded by the Spanish government. This study is part of the European project 'IMPACT'. It was carried with financial support from the Commission of the European Communities, specific RTD pro-gramme 'IWRMNET'. This work was supported by the 'Laboratoire d'Excellence' (LABEX) entitled TULIP (ANR-10-LABX-41).

Acknowledgements. We thank all our colleagues and ONEMA staff that contributed to the field sampling, especially Laurence Blanc. Lisa Fourtune is acknowledged for having calculated betweenness centrality values. We also thank Marie-Hélène Lizée, Camille Pagès and Je ´rôme Prunier for their comments, the 'Génopole Toulouse' for help with genotyping and Keoni Saint-Pée for correcting the English. We are grateful to Loeske Kruuk, Robert Wilson and five anonymous reviewers for their comments on earlier versions of the manuscript.

References

1. Watson JEM, Dudley N, Segan DB, Hockings M. 2014 The performance and potential of protected areas. *Nature* **515**, 67–73. (doi:10.1038/nature13947)
2. Filipe AF, Marques TA, Tiago P, Ribeiro F, Da Costa LM, Cowx IG, Collares-Pereira MJ. 2004 Selection of priority areas for fish conservation in Guadiana river basin, Iberian Peninsula. *Conserv. Biol.* **18**, 189–200. (doi:10.1111/j.1523-1739.2004.00620.x)
3. Kirkpatrick JB. 1983 An iterative method for establishing priorities for the selection of nature reserves: an example from Tasmania. *Biol. Conserv.* **25**, 127–134. (doi:10.1016/0006-3207(83)90056-3)
4. Margules CR, Pressey RL. 2000 Systematic conservation planning. *Nature* **405**, 243–253. (doi:10.1038/35012251)
5. Vane-Wright RI, Humphries CJ, Williams PH. 1991 What to protect? Systematics and the agony of choice. *Biol. Conserv.* **55**, 235–254. (doi:10.1016/0006-3207(91)90030-D)
6. Faith DP. 1992 Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **61**, 1–10. (doi:10.1016/0006-3207(92)91201-3)
7. Asmyhr MG, Linke S, Hose G, Nipperess DA. 2014 Systematic conservation planning for groundwater ecosystems using phylogenetic diversity. *PLoS ONE* **9**, e115132. (doi:10.1371/journal.pone.0115132)
8. Buerki S, Callmander MW, Bachman S, Moat J, Labat J-N, Forest F. 2015 Incorporating evolutionary history into conservation planning in biodiversity hotspots. *Phil. Trans. R. Soc. B* **370**, 20140014. (doi:10.1098/rstb.2014.0014)
9. Carvalho SB, Velo-Antón G, Tarroso P, Portela AP, Barata M, Carranza S, Moritz C, Possingham HP. 2017 Spatial conservation prioritization of biodiversity spanning the evolutionary continuum. *Nat. Ecol. Evol.* **1**, 0151. (doi:10.1038/s41559-017-0151)
10. Thomassen HA *et al.* 2011 Mapping evolutionary process: a multi-taxa approach to conservation prioritization. *Evol. Appl.* **4**, 397–413. (doi:10.1111/j.1752-4571.2010.00172.x)
11. Diniz-Filho JAF *et al.* 2012 Planning for optimal conservation of geographical genetic variability within species. *Conserv. Genet.* **13**, 1085–1093. (doi:10.1007/s10592-012-0356-8)
12. Hanson JO, Rhodes JR, Riginos C, Fuller RA. 2017 Environmental and geographic variables are effective surrogates for genetic variation in conservation planning. *Proc. Natl Acad. Sci. USA* **114**, 12 755–12 760. (doi:10.1073/pnas.1711009114)
13. Hermoso V, Kennard MJ, Schmidt DJ, Bond N, Huey JA, Mondol RK, Jamandre BW, Hughes JM. 2016 Species distributions represent intraspecific genetic diversity of freshwater fish in conservation assessments. *Freshw. Biol.* **61**, 1707–1719. (doi:10.1111/fwb.12810)
14. Carroll SP, Jorgensen PS, Kinnison MT, Bergstrom CT, Denison RF, Gluckman P, Smith TB, Strauss SY, Tabashnik BE. 2014 Applying evolutionary biology to address global challenges. *Science* **346**, 1245993. (doi:10.1126/science.1245993)
15. Rey O, Danchin E, Mirouze M, Loot C, Blanchet S. 2016 Adaptation to global change: a transposable element–epigenetics perspective. *Trends Ecol. Evol.* **31**, 514–526. (doi:10.1016/j.tree.2016.03.013)
16. Mittell EA, Nakagawa S, Hadfield JD. 2015 Are molecular markers useful predictors of adaptive potential? *Ecol. Lett.* **18**, 772–778. (doi:10.1111/ele.12454)
17. Pritchard JK, Stephens M, Donnelly P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.
18. Putman AI, Carbone I. 2014 Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol. Evol.* **4**, 4399–4428. (doi:10.1002/ece3.1305)
19. Blanchet S, Prunier JG, De Kort H. 2017 Time to go bigger: emerging patterns in macrogenetics. *Trends Genet.* **33**, 579–580. (doi:10.1016/j.tig.2017.06.007)
20. Moritz C. 1994 Defining 'Evolutionarily Significant Units' for conservation. *Trends Ecol. Evol.* **9**, 373–375. (doi:10.1016/0169-5347(94)90057-4)
21. Waples RS *et al.* 2001 Characterizing diversity in salmon from the Pacific Northwest. *J. Fish Biol.* **59**, 1–41. (doi:10.1111/j.1095-8649.2001.tb01376.x)
22. Frankham R, Briscoe DA, Ballou JD. 2002 *Introduction to conservation genetics*. Cambridge, UK: Cambridge University Press.
23. Kottelat M, Freyhof J. 2007 *Handbook of European freshwater fishes*. Cornol, Switzerland: Publications Kottelat.
24. Poulet N, Beaulaton L, Dembski S. 2011 Time trends in fish populations in metropolitan France: insights from national monitoring data. *J. Fish Biol.* **79**, 1436–1452. (doi:10.1111/j.1095-8649.2011.03084.x)
25. Paz-Vinas I, Quéméré E, Chikhi L, Loot G, Blanchet S. 2013 The demographic history of populations experiencing asymmetric gene flow: combining simulated and empirical data. *Mol. Ecol.* **22**, 3279–3291. (doi:10.1111/mec.12321)
26. Paz-Vinas I, Comte L, Chevalier M, Dubut V, Veyssiere C, Grenouillet G, Loot G, Blanchet S. 2013 Combining genetic and demographic data for prioritizing conservation actions: insights from a threatened fish species. *Ecol. Evol.* **3**, 2696–2710. (doi:10.1002/ece3.645)
27. Aljanabi SM, Martinez I. 1997 Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* **25**, 4692–4693. (doi:10.1093/nar/25.22.4692)
28. Szpiech ZA, Jakobsson M, Rosenberg NA. 2008 ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* **24**, 2498–2504. (doi:10.1093/bioinformatics/btn478)
29. Winter DJ. 2012 mmod: an R library for the calculation of population differentiation statistics. *Mol. Ecol. Resour.* **12**, 1158–1160. (doi:10.1111/j.1755-0998.2012.03174.x)
30. Jost L. 2008 G(st) and its relatives do not measure differentiation. *Mol. Ecol.* **17**, 4015–4026. (doi:10.1111/j.1365-294X.2008.03887.x)
31. Peterson E, Ver Hoef JMV. 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**, 1–17. (doi:10.18637/jss.v056.i02)
32. Ver Hoef J, Peterson E, Clifford D, Shah R. 2014 SSN: an R package for spatial statistical modeling on stream networks. *J. Stat. Softw.* **56**, 1–43.
33. Ball IR, Possingham HP, Watts M. 2009 Marxan and relatives: software for spatial conservation prioritisation. In *Spatial conservation prioritisation: quantitative methods and computational tools* (eds A Moilanen, KA Wilson, HP Possingham), pp. 185–195. Oxford, UK: Oxford University Press.
34. Finn DS, Bonada N, Múrria C, Hughes JM. 2011 Small but mighty: headwaters are vital to stream network biodiversity at two levels of organization. *J. North Am. Benthol. Soc.* **30**, 963–980. (doi:10.1899/11-012.1)
35. Estrada E, Bodin O. 2008 Using network centrality measures to manage landscape connectivity. *Ecol. Appl.* **18**, 1810–1825. (doi:10.1890/07-1419.1)
36. Altermatt F. 2013 Diversity in riverine metacommunities: a network perspective. *Aquat. Ecol.* **47**, 365–377. (doi:10.1007/s10452-013-9450-3)
37. Fourtune L, Paz-Vinas I, Loot G, Prunier JG, Blanchet S. 2016 Lessons from the fish: a multi-species analysis reveals common processes underlying similar species-genetic diversity correlations. *Freshw. Biol.* **61**, 1830–1845. (doi:10.1111/fwb.12826)
38. Fortuna MA, Albaladejo RG, Fernandez L, Aparicio A, Bascompte J. 2009 Networks of spatial genetic variation across species. *Proc. Natl Acad. Sci. USA* **106**, 19 044–19 049. (doi:10.1073/pnas.0907704106)
39. 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. (doi:10.1111/ele.12004)
40. Paz-Vinas I, Loot G, Stevens VM, Blanchet S. 2015 Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. *Mol. Ecol.* **24**, 4586–4604. (doi:10.1111/mec.13345)
41. Paz-Vinas I, Blanchet S. 2015 Dendritic connectivity shapes spatial patterns of genetic diversity: a simulation-based study. *J. Evol. Biol.* **28**, 986–994. (doi:10.1111/jeb.12626)
42. Thomaz AT, Christie MR, Knowles LL. 2016 The architecture of river networks can drive the evolutionary dynamics of aquatic populations. *Evolution* **70**, 731–739. (doi:10.1111/evo.12883)
43. Fourtune L, Prunier JG, Paz-Vinas I, Loot G, Veyssière C, Blanchet S. 2018 Inferring causalities in landscape genetics: an extension of Wright's causal modeling to distance matrices. *Am. Nat.* **191**, 491–508. (doi:10.1086/696233)
44. Mimura M *et al.* 2017 Understanding and monitoring the consequences of human impacts on intraspecific variation. *Evol. Appl.* **10**, 121–139. (doi:10.1111/eva.12436)