Received: 5 December 2022 | Revised: 1 June 2023 | Accepted: 19 June 2023

DOI: 10.1111/1755-0998.13828

MOLECULAR ECOLOGY RESOURCES WILEY

RESOURCE ARTICLE

Sampling effect in predicting the evolutionary response of populations to climate change

Jonás A. Aguirre-Liguori[1,2] | Abraham Morales-Cruz[3] | Brandon S. Gaut[3] | Santiago Ramírez-Barahona[1]

[1]Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México (UNAM), Ciudad de México, Mexico

[2]Departamento de Ecología Tropical, Campus de Ciencias Biológicas y Agropecuarias, Universidad Autónoma de Yucatán, Mérida, Mexico

[3]Department of Ecology and Evolutionary Biology, University of California, Irvine, Irvine, California, USA

Correspondence
Jonás A. Aguirre-Liguori and Santiago Ramírez-Barahona, Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México (UNAM), Circuito Exterior s/n, Ciudad de México, 04510, Mexico.
Email: jonas.aguirre@hotmail.com and santiago.ramirez@ib.unam.mx

Funding information
Consejo Nacional de Ciencia y Tecnología, Grant/Award Number: FORDECYT-PRONACES/263962/2020

Handling Editor: Catherine Grueber

Abstract

Genomic data and machine learning approaches have gained interest due to their potential to identify adaptive genetic variation across populations and to assess species vulnerability to climate change. By identifying gene–environment associations for putatively adaptive loci, these approaches project changes to adaptive genetic composition as a function of future climate change (genetic offsets), which are interpreted as measuring the future maladaptation of populations due to climate change. In principle, higher genetic offsets relate to increased population vulnerability and therefore can be used to set priorities for conservation and management. However, it is not clear how sensitive these metrics are to the intensity of population and individual sampling. Here, we use five genomic datasets with varying numbers of SNPs (N_{SNPs} = 7006–1,398,773), sampled populations (N_{pop} = 23–47) and individuals (N_{ind} = 185–595) to evaluate the estimation sensitivity of genetic offsets to varying degrees of sampling intensity. We found that genetic offsets are sensitive to the number of populations being sampled, especially with less than 10 populations and when genetic structure is high. We also found that the number of individuals sampled per population had small effects on the estimation of genetic offsets, with more robust results when five or more individuals are sampled. Finally, uncertainty associated with the use of different future climate scenarios slightly increased estimation uncertainty in the genetic offsets. Our results suggest that sampling efforts should focus on increasing the number of populations, rather than the number of individuals per populations, and that multiple future climate scenarios should be evaluated to ascertain estimation sensitivity.

KEYWORDS
climate change, conservation genomics, genetic offsets, genetic structure, genomics of population, sampling design

1 | INTRODUCTION

It is unlikely that greenhouse gas emissions will decrease by the year 2100 (IPPC, 2022) and thus these will continue to generate changes in mean global climatic conditions, including an increased frequency and severity of more-extreme seasonal temperatures, drought and other natural disasters. Overall, these changes will continue to negatively impact the persistence of species (Parmesan, 2006; Parmesan & Yohe, 2003) and affect their interactions (Zamora-Gutiérrez et al., 2021), leading to increased risk of population decimations, local extinctions and losses of genetic diversity (Ceballos et al., 2017; Wuest et al., 2021).

Climate adaptation strategies are necessary to conserve and manage wild populations and to facilitate an adequate and timely response of populations to climate change (LeDee et al., 2021; Thompson et al., 2023). Unfortunately, reducing the potential impacts of climate change is due to the difficulties of obtaining adequate biological information and designing successful conservation strategies. Conservation and mitigation strategies should favour identifying or promoting the evolutionary potential or adaptive capacity of species and populations, based on the identified or hypothesized pathway(s) and mechanism(s) by which climate acts on the focal taxa (Forester et al., 2022; LeDee et al., 2021; Meek et al., 2023; Thompson et al., 2023; Thurman et al., 2020). To date, the evaluation of a population's risk has been most commonly based on future persistence probabilities estimated through the use of species distribution models (SDMs) (Peterson et al., 2011). However, SDMs do not incorporate the demographic and the evolutionary potential of species, which limits their potential use (Forester et al., 2022). The adaptive capacity of a species in the future can also be assessed using phenotypic, ecological, geographic, experimental, life-history and genetic data, in addition to addressing pathways affecting the exposure, sensitivity and adaptive capacity of focal taxa (Thompson et al., 2023; Thurman et al., 2020).

When genetic data are available, the future persistence of populations can also be assessed based on patterns of genetic diversity, inbreeding metrics and population structure across species' ranges (Frankham, 2005; Thompson et al., 2023; Tobón-Niedfeldt et al., 2022). Many studies have used neutral or adaptive genetic markers to assess genetic diversity and population connectivity as indirect indicators of the potential for persistence. However, considering adaptive capacity is necessary for predicting the possible response of populations to future climate change (Thurman et al., 2020, 2022). The analysis of patterns of adaptive genetic diversity and the evolutionary history of populations has been bolstered by the development of genomic tools, which allow the production of genome-wide data on a massive scale even for nonmodel species (Aguirre-Liguori et al., 2021; Capblancq, Fitzpatrick, et al., 2020; Schoville et al., 2012; Stapley et al., 2010; Tiffin & Ross-Ibarra, 2014; Waldvogel et al., 2019). The use of population genomic data to assess risks and evolutionary capacities across populations can help to identify more (or less) vulnerable populations and to recognize priorities for population conservation and management (Allendorf et al., 2010; Thurman et al., 2020; Waldvogel et al., 2019). Indeed, genome-wide data have been used to characterize genetic-environment associations, identify genetic variants involved in climate adaption and assess how populations will cope under new conditions brought on by a rapidly changing climate (Capblancq, Fitzpatrick, et al., 2020).

The rationale for the use of population genomic data is straightforward: some populations can either be locally adapted to environmental and biotic conditions expected under future climate change scenarios (Meek et al., 2023), or they may contain standing genetic variation that could become adaptive in the future (Barrett & Schluter, 2007; Hampe & Petit, 2005). The challenge is connect-ing genetic variants, and the projected change of genetic variants over time, to climatic change. In recent years, machine learning approaches have gained popularity for assessing population vulnerability and the capacity for adaptation to new climate conditions. One of these approaches, Gradient Forest (GF), has been used to identify gene–environment associations, to model genetic turnover across the landscape and to project changes in genetic composition under future climate change scenarios (Fitzpatrick & Keller, 2015). The estimated change in genetic composition over time is called the genetic offset (Fitzpatrick & Keller, 2015), and it is interpreted as measuring the maladaptation of populations in the face of predicted climate change (Capblancq, Fitzpatrick, et al., 2020). In theory, populations with a high genetic offset are expected to be more vulnerable to climate change either because these will need to respond substantively through adaptation or migration or because their current genetic composition is projected to be out-of-sync with future conditions. This is clearly an oversimplification of the process, because genetic offsets omit relevant mechanisms of populations' adaptation, such as climatic tolerances driven by phenotypic plasticity (Thurman et al., 2020), adaptation through new mutations or standing genetic variation (Barrett & Schluter, 2007), the introgression of adaptive alleles by gene flow and maladaptation associated with the accumulation of deleterious mutations (Aguirre-Liguori et al., 2021). However, the genetic offset of a given population can also be projected across the landscape (Gougherty et al., 2021; Rhoné et al., 2020) to potentially identify areas where the population may remain locally adapted if actively translocated (Gougherty et al., 2021).

The temporal and spatial projections of genetic offsets across a landscape can be used to assess populations' response to climate change either through in situ local adaptation or migration. Few experimental validations of the projected genomic offsets have been performed (Rellstab et al., 2021), but these show that genomic offsets have the potential to successfully predict the response of populations to climate change (Exposito-Alonso et al., 2019; Fitzpatrick et al., 2021; Rhoné et al., 2020). However, it is still not clear how powerful or informative these methods are, and how useful they are for the conservation and management of nonmodel species (Rellstab et al., 2021). For instance, a recent simulation study revealed the sensitivity of genetic offsets to species' genetic structure and demographic history (Láruson et al., 2022). In addition, genetic offsets appear to be inversely correlated with deme size, suggesting that

genetic drift could impact the estimation of genetic offsets (Láruson et al., 2022).

In addition to these limitations, there is a practical aspect to genetic offsets that remains unexplored: their sensitivity to nonrandom and limited population (and genomic) sampling. To construct a GF model and estimate genetic offsets, an ideal sampling likely consists of populations distributed across the entire geographic and environmental distribution of the species, with large number of loci across the genome and enough individuals per population to provide reasonable estimates of allele frequencies (Aguirre-Liguori et al., 2020). However, there is always a tradeoff between the number of sampled individuals and sequencing resources (Meirmans, 2015) and many highly vulnerable nonmodel species have few genetic resources or known populations. Therefore, it is critical to empirically evaluate the sensitivity of genomic offset projections under limited, but realistic, sampling of natural populations.

Here, we perform an assessment of the sensitivity of genetic offsets under both varying degrees of sampling effort and different climate change scenarios. To answer this question, we use published genomic datasets (Table 1) to evaluate how the estimation of genetic offsets is impacted by sampling designs by varying the number of populations, ecological variation and geographic spread. We also evaluate the uncertainty in the estimation of genetic offset associated with the use of multiple climate projections, based on different socioeconomic pathways models of greenhouse gas effects and global circulation models. Finally, we use our results to generate a list of recommendations for adequate sampling designs and sensitivity tests to consider when estimating genetic offsets.

2 | MATERIALS AND METHODS

2.1 | Bioclimate and genomic datasets

We downloaded from WorldClim (Fick & Hijmans, 2017), 19 bioclimatic variables for the present-day (1970–2000) at a 2.5 arc-minutes resolution. We also obtained future models from the Coupled Model Intercomparison Project Phase 6 (CMIP6, publicly available at https://esgf-node.llnl.gov/search/cmip6/, Eyring et al., 2016). Since future climates are modelled using different shared socioeconomic pathways and different global circulation models, it is important to include this uncertainty to have a range of possible future scenarios (Foden et al., 2019). To have a good representation of future climatic scenarios (Sanderson et al., 2015), we obtained data derived from five circulation models (GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0, UKESM1-0-LL), four shared socioeconomic pathways (i.e. SSPs 126, 245, 370, 585) and three time periods (2041–2060, 2061–2080 and 2081–2100).

We downloaded published genomic datasets (Table 1) belonging to different taxonomic groups (plants and animals), with different life histories (i.e. annuals, perennials) that inhabit several regions across the world (Figure S1). Because we aimed to explore sampling properties, we focussed on large datasets that were sufficient for subsampling. We selected datasets with more than 20 sampled populations and with geographic information available for every population. For each species, we ran the bioinformatic analyses described in Supplementary Methods (Supporting Text S1) to process the data, obtain SNPs and climatic data for each population and generate the input datasets used for further analyses. For each species, we generated: (1) a table with populations' allelic counts for each SNP; and (2) a table with the geographic and environmental data of each population. For species with individual-level data, we also generated a *genind* object (package in R). Since we used different numbers of individuals per subsampling design with different missing data (see below) we needed to impute these datasets to obtain individual genotypes for most loci. We used the imputation procedure of the package in R (Frichot & François, 2015) to estimate the number of ancestral populations (K) using sparse non-negative matrix factorization algorithms (*snmf* function, with $k = 1{:}10$) (Frichot & François, 2015) and impute missing genotypes across individuals (*impute* function in , using the K selected in the *snmf* analysis).

We used the resulting SNPs as the reference datasets to estimate genetic offsets and to assess estimation sensitivity under different

TABLE 1 Description of the datasets and summary statistics of the species analysed.

Species	Life-history trait	Distribution	Dataset	N_{Pop}	N_{Ind} (range)	N_{SNPs}	H_S	F_{ST}	Reference
Empidonax traillii	Bird	North America	Radseq	25	185 (1–19)	10,790	0.14	0.04	Ruegg et al. (2018)
Fagus sylvatica	Perennial/ Plant	Europe	Radseq	36	595 (9–63)	7006	0.22	0.05	Capblancq, Fitzpatrick, et al. (2020), Capblancq, Morin, et al. (2020)
Zea mays mexicana	Annual/plant	Mexico	Chip	23	333 (11–15)	33,456	0.22	0.16	Aguirre-Liguori et al. (2017)
Zea mays parviglumis	Annual/plant	Mexico	Chip	24	313 (10–16)	33,456	0.22	0.24	Aguirre-Liguori et al. (2017)
Arabidopsis lyrata	Annual/plant	North America	WGS	47	poolSeq	1,398,773	0.14	0.48	Willi et al. (2018)

sampling schemes. This approach assumes that the reference datasets are representative of the true patterns and are themselves without bias, which is unlikely because these datasets are impacted by their original sampling (e.g. not sampling an adequate representation of individuals, populations and SNPs) (see *Caveats* section in the Section 4). Nonetheless, the reference datasets provide a convenient point of comparison to assess sampling sensitivity.

2.2 | Summary statistics and outlier SNPs

The estimation of genetic offsets depends on a robust estimation of the demographic history and genetic structure of species. Genetic structure and genetic drift make genome-environment association (GEA) prone to identify false positives—that is, if genetic structure is co-aligned with a selective environment, false-positive SNPs can be identified due to population history rather than genetic adaptation. To contextualize our estimates of genetic offsets in each dataset described in Table 1, we estimated two summary statistics that correlate with levels of genetic drift and genetic structure. We first estimated the genetic diversity (*Hs*) using the *basic.stats* function in the *Hierfstat* (Goudet, 2005) package in R. The *basic.stats* function estimates the genetic diversity of each locus in each population and the overall mean *Hs* across all populations. We also used the *apply* function in R to estimate the mean *Hs* for each population. To account for varying patterns and strength of genetic structure, we used the *calculate.pairwise.Fst* function in the package in R (Bradburd et al., 2013) to estimate pairwise F_{ST} between each pair of populations, the mean F_{ST} across all populations and the mean F_{ST} between each population against the remaining populations. Finally, we assessed environmental and geographic variation across populations, which frequently correlate with genetic structure, by estimating the distances to the species' geographic and environmental centroids for every population (Lira-Noriega & Manthey, 2014); these distances describe how far populations are from the species' optimal environmental conditions (Eckert et al., 2008). The geographic centroid was estimated as the median longitude and latitude of species' occurrences. For the niche centroid, first, we used the *prcomp* function in R to obtain a principal component analysis (PCA) of the 19 bioclimatic variables (explained >95% of variance in each dataset). Then, we estimated the niche centroid as the mean score across the first six principal components (PCs). The distances to the two centroids were estimated as the Euclidian distance between the populations' position (geographic coordinates, PC scores) and the respective centroids (Lira-Noriega & Manthey, 2014).

The application of GF often relies on outlier loci (Fitzpatrick et al., 2018; Fitzpatrick & Keller, 2015), which we identified using latent factor mixed model 2 (LFMM2) (Caye et al., 2019). Briefly, LFMM2 identifies SNPs whose allelic frequencies are significantly associated with continuous variables (that is, climate), while controlling for genetic structure using latent factors. For simplicity, we only used the scores from the first PC obtained from the bioclimatic PCA described above. To determine the number of Latent Factors (*K*), we used the *prcomp* function in R to perform a genetic PCA using all SNPs, calculated the percentage of variance explained by each PC and estimated the rate of change of the variance explained between consecutive PCs. The optimal *K* was defined where the rate of change of the variance explained between two consecutive PCs was negative (meaning that there was a 'knee' in the screeplot). Next, we ran the LFMM2 model using *K* latent factors, a ridge penalty (*lfmm_ridge* function), and controlling for the genomic inflation (*lfmm_test* function, with calibrate='gif'). To make sampling consistent across datasets, for each dataset we defined the 100 SNPs with the lowest *p*-value as outliers. Applying strict corrections (Bonferroni or FDR) for outlier detection would decrease the number of outlier SNPs and hinder our capacity to compare results between replicates.

2.3 | Reference genetic offsets

We used GF (Ellis et al., 2012) to model the turnover in the genetic composition of species across the landscape as a function of continuous variables, usually climate (Fitzpatrick & Keller, 2015). The GF models estimate differences in genetic composition among populations, assess which response variables have the strongest contribution to model construction and identify particular SNPs with significant nonlinear associations with response variables (Capblancq, Fitzpatrick, et al., 2020; Fitzpatrick & Keller, 2015; Waldvogel et al., 2019).

For each genomic dataset (that is species), we built a GF model using the *gradientForest* function from the R package , based on 500 trees per run and *corr.threshold*=0.5. The initial models used all sampled populations and estimated the corresponding genetic offset using a reference climate scenario (see below). We built GF models with the 100 outlier SNPs described above using the 19 bioclimatic variables. For every model replicate, we ranked-ordered the bioclimatic variables based on their weighted importance for model construction and also ranked-ordered SNPs with significant associations with climate ($R^2>0$). To calculate the genetic offsets of populations, we used the GF models to predict the genetic composition of populations across the landscape using present-day and future climate layers. Keeping with the original methodology of Fitzpatrick and Keller (2015), we used the multidimensional Euclidian distances between the present and future genetic compositions to calculate populations' genetic offsets.

To incorporate uncertainty in future climate projections, we estimated the genetic offsets of populations under each of the 54 different climate change projections (see above), which allowed us to assess the impact of climate model uncertainty on the estimation of genetic offsets. Finally, we estimated the rank correlation (*cor.test* function in R) between the reference genetic offset of populations against offsets estimated under each climatic scenario. In this case, a rank correlation approaching 1.0 indicates that two climate scenarios project similar genetic offsets. For tractability, we defined a *reference climate scenario* as the one with the strongest rank correlations of genetic offsets relative to the remaining 53 climate scenarios. To

further assess uncertainty among climate change scenarios, for each bioclimatic variable we estimated the correlation between present values and their future projections under the 54 models; these correlations were estimated for each species separately.

2.4 | Subsampling designs

Given the costs of genome-wide sequencing, there is a trade-off between the number of individuals and populations that can be sequenced. Previous studies have shown that sampling a limited number of populations (relative to the 'true' number of populations) can inflate estimates of F_{ST} between populations (Aguirre-Liguori et al., 2020; Landguth & Schwartz, 2014; Schwartz & McKelvey, 2008; Willing et al., 2012) and increase the number of false positives outlier SNPs (De Mita et al., 2013). Likewise, Gradient Forest models rely on estimates of allelic frequencies across populations and therefore are, in principle, sensitive to varying levels of sampling intensity both in terms of populations and individuals per population. The reference genetic offset datasets described above were used as a baseline to assess the sensitivity of the inference of genetic offsets under different subsampling designs. To approximate real-world situations, we designed our subsampling to vary sampling intensities across populations and individuals within those populations. For each genomic dataset (Table 1), we generated random subsamples of N populations without replacement using varying levels of sampling intensity: $N_{pop}=5$, $N_{pop}=10$, $N_{pop}=15$, $N_{pop}=20$, and when possible, $N_{pop}=25$ and $N_{pop}=30$. Each random subsampling (i.e. $N_{pop}=5$) was replicated 1000 times. To describe the geographic and environmental structure of the sampled populations in each replicate, we evaluated how well each replicate represented the climatic and geographic distribution of the entire reference set of populations. For each replicate, we estimated the absolute sum and the standard deviation of the distances to the niche (Σ_{niche}, sd_{niche}) and geographic (Σ_{geo}, sd_{geo}) centroid of sampled populations, divided by N_{pop}. High values of Σ_{niche} or Σ_{geo} indicate that some of the sampled populations were near the niche edge and that there was wide coverage of the underlying environmental or geographic space.

For genomic datasets with individual-level data (Table 1), we also evaluated the effect of sampling different numbers of individuals per population. The number of individuals within a population is in principle an important parameter because it impacts the estimation of allele frequencies, a key input parameter for estimating genomic offsets. For each dataset, we chose the 20 populations that had the highest number of individuals and generated random subsamples of N_{ind} individuals per population using varying sampling intensities from $N_{ind}=1$ to $N_{ind}=10$ (i.e. $N_{ind}=1$ replicates consisted of 20 populations each with one sampled individual). For each N_{ind} we performed 1000 replicates. In the few instances where N_{ind} was higher than the number of individuals in a given population, we sampled all individuals present in the population. We also tested the effect of sampling all individuals for 20 populations and adding one to five extra populations (20+1 to 20+5) but with only one individual in each added population (singleton populations). The rationale behind this test is to assess the effect of including populations known by single individuals (for instance, museum collections or accessions).

2.5 | Estimation sensitivity of genetic offsets

For each subsampled replicate (using subsamples of either populations or individuals), we performed the same analyses described for the reference datasets: (1) identify 100 outlier SNPs; (2) estimate genetic offsets; (3) rank-order bioclimatic variables; and (4) identify SNPs significantly associated with climate.

Since we performed random subsamplings, an individual population was not selected for all replicates. Thus, each time a population was sampled in each replicate, we recorded its estimated genetic offset. We then estimated the mean and range of estimated genetic offsets across replicates for each population under different sampling intensities (See Figure S2). To assess inference sensitivity associated with different sampling intensities, we estimated the Spearman rank correlation (*cor.test* function in R) between populations' reference genetic offsets and their corresponding replicate genetic offsets. A rank correlation approaching 1.0 indicates that, irrespective of the sampling context, populations always tend to rank in the same order of genetic offsets. Likewise, a rank correlation approaching 0 (or becoming negative) indicates there the estimation of genetic offsets deviates significantly from the reference under limited sampling.

An important application of GF models is the ability to extrapolate the turnover functions across the landscape and predict genetic offsets at unsampled locations (Fitzpatrick & Keller, 2015). Extrapolating genetic offsets may help identify areas more or less vulnerable to future climate change (Fitzpatrick & Keller, 2015) or identify areas suitable for active migration or translocation in the future (Gougherty et al., 2021; Rhoné et al., 2020). In real-life scenarios, extrapolation to unsampled locations may correspond to instances in which populations are known but remain unsampled due to budget or logistic limitations. Thus, we also tested the capacity to predict genetic offsets in nonsampled populations under different levels of sampling intensity. To do this, for each sampling replicate we divided the original populations into sampled or unsampled. We then used the GF models estimated with the sampled populations (e.g. $N_{pop}=10$; see Section 2.4 above) to extrapolate the turnover functions and estimate genetic offset for the unsampled populations. We called these genetic offsets for unsampled populations the 'extrapolated genetic offset'. We estimated the rank correlation between the reference genetic offsets and the extrapolated genetic offsets of the unsampled populations. A rank correlation approaching 1.0 indicates that, irrespective of sampling, the extrapolation of genetic offsets is not sensitive to sampling.

The five genomic datasets we used were generated using different sequencing platforms with varying depth and coverage. Thus, the number of SNPs varied among species. These differences may result in varying power to detect the full range of genomic patterns of local adaptation across species, especially when considering only 100

outlier SNPs. To assess possible bias associated with the selection of SNPs, we also performed replicate sampling with varying number of population (N_{pop} = 5, N_{pop} = 10, N_{pop} = 15, N_{pop} = 20, and when possible, N_{pop} = 25 and N_{pop} = 30; see *Subsampling designs* above) using the complete datasets with all SNPs. Our rational was that using all SNPs relaxed the constraints and biases associated with the use of only 100 outlier SNPs. Due to computational limitations, we performed 100 replicates per subsampling and excluded the *Arabidopsis lyrata* dataset from this analysis, which had more than 1 million SNPs in the dataset (Table 1).

2.6 | Identifying relevant bioclimatic variables

GF models identify the environmental variables that most strongly contribute to the construction of the models, measured by their 'weight'. Weights have been used to determine which variables have the strongest effects on populations in the future (Aguirre-Liguori et al., 2021; Capblancq, Fitzpatrick, et al., 2020). For each dataset, from the GF models we extracted the estimated R^2 weighted importance of the bioclimatic variables that have the strongest contribution to the reference GF model. To compare among species, we renamed these bioclimatic variables as pos1 to pos19 based on their order of importance. For each replicate we rank-ordered the bioclimatic variables based on their contribution to model construction; we then compared the variability of ranking across replicates compared with that from the reference model. We visualized the similarity between replicates and the reference model using a heat map in which an aggregation of warmer colours across the diagonal indicates that the true contribution of bioclimate variables is retrieved irrespective of sampling.

3 | RESULTS

3.1 | Focal datasets

To estimate both the sensitivity of genetic offsets to sampling intensity and the robustness of inference based on climate models, we analysed complete and subsampled datasets from five species (Table 1). We chose the five datasets based largely on sampling intensity; we required that each dataset represents at least 20 populations, with the sufficient sample within populations to have reasonable estimates of allele frequencies. However, the datasets varied in myriad ways. For example, one dataset from the annual plant *Arabidopsis lyrata* was based on pooled sequencing of multiple individuals within each population. Other datasets—such as from the annual plants *Zea mays* ssp. *mexicana* and *Zea mays* ssp. *parviglumis*—were based on chip-based assays rather than resequencing data. The species also varied dramatically in genetic diversity (H_S) and population differentiation (F_{ST}). Finally, the species themselves varied quite dramatically, in that one represented a bird (*Empidonas traillii*), another represents a perennial plant (*Fagus sylvatica*), while the remainder are annual plants. Overall, our goal was to find datasets with sufficient sampling to be able to assess the effect of subsample, but we also sought to choose datasets that represent diverse ecological and organismal histories.

3.2 | Population sampling

3.2.1 | Estimation sensitivity

We tested whether the number of populations sampled impacted the estimation of genetic offsets. To assess this, we estimated genetic offsets for replicate subsamples consisting of 5, 10, 15 and 20 randomly selected populations—and 25 and 30 populations when possible—for each species. We compared the genetic offsets across replicates to reference estimates based on the entire dataset and the *reference climate scenario* (i.e. the scenario that had the highest correlations to other scenarios; see Section 2). Overall, we found a consistent trend across species, where the rank correlation between the reference and the replicate genetic offset estimates increased with increasing sampling intensity (Figure 1a, Table S1). Decreasing the number of populations resulted in genetic offsets that departed from the reference metrics across species and, interestingly, this sensitivity varied as a function of the species' genetic structure. More pronounced genetic structure (higher F_{ST}) resulted in more sensitive estimates, such that a higher intensity of population sampling was needed to converge on the reference genetic offsets. For example, in *Arabidopsis lyrata* (F_{ST} = 0.48) the median rank correlation decreased from 0.8 to −0.1 when decreasing sampling from 30 to 5 populations. By contrast, the median rank correlation in *Empidonax traillii* (F_{ST} = 0.04) decreased from 1 to 0.6 with an equivalent decrease in population sampling (Table S1).

Moreover, in the two species with the highest F_{ST} (*A. lyrata* and *Z. mays parviglumis*), we found that more than 15 populations were needed to retrieve a median rank correlation above 0.5 (Figure 1a, Table S1). For *A. lyrata* (F_{ST} = 0.48), at least 20 populations were needed to retrieve a median rank correlation >0.5 (the median rank correlations for N_{pop} = 15 and N_{pop} = 20 were 0.39 and 0.59, respectively). By contrast, we found that sampling 5 populations was enough to obtain rank correlations >0.5 for the remaining three species: *Z. mays mexicana* (median rank correlation of 0.5), *E. traillii* (median rank correlation of 0.6) and *F. sylvatica* (median rank correlation of 0.58) (Table S1).

A large proportion of replicates with fewer than 15 populations showed rank correlations <0.5 between the reference and replicate genetic offsets (Figure 1a, Table S1). The proportion of sampling replicates with rank correlations <0.5 when sampling 15 populations ranged from 15.4% in *F. sylvatica* to 62.3% in *A. lyrata* (Table S1). We wondered whether maximizing the environmental or geographic distribution of populations (Σ_{env}, Σ_{geo}), while minimizing population clustering (sd_{env}, sd_{geo}), could lead to better performance. To evaluate this idea, we performed replicates with low sampling intensity (N_{pop} = 5 and N_{pop} = 10) that maximized $\Sigma_{env/geo}$ but that minimized

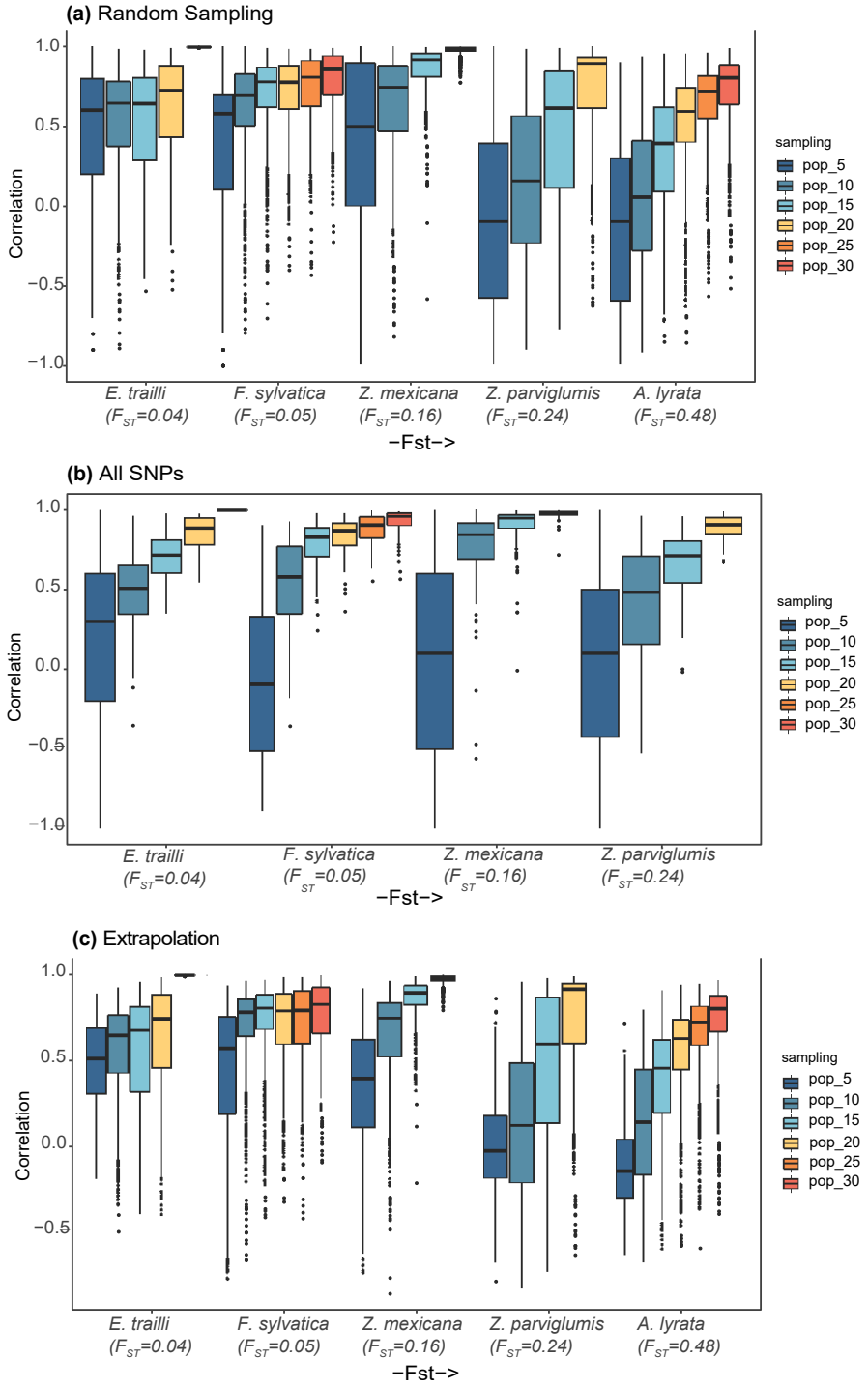

FIGURE 1 Estimation sensitivity of genetic offsets to varying population sampling intensities. In each panel, the boxplots depict the distribution of the rank correlation of genetic offsets between replicates and the reference datasets for each of five species. (a) reports correlations based on 100 outlier SNPs and (b) reports on all available SNPs. The rank correlations in (c) report correlations from extrapolations of unsampled populations. Each boxplot is based on 1000 replicates. Rank correlations closer to 1.0, with less variability across replicates, indicate both lower estimation sensitivity of genetic offsets and greater agreement between replicates and the reference datasets. In each panel, the species are ordered from left to right according to increasing genetic structure (F_{ST}). The boxes show the first and third quartiles. The upper and lower whisker show values that are at the limit of 1.5 times the interquartile range. The remaining dots show outliers.

$sd_{env/geo}$ (See Section 2). Counterintuitively, we found that maximizing the geographic or environmental distribution of sampled populations did not increase the correlation between the reference and subsampled genetic offsets (Figure S3).

It is possible that our results reflect properties of the methods used to identify SNP outliers rather than properties of genomic offset estimation. To assess this possibility, we also performed replicate sampling on the complete datasets (all SNPs) and found that increasing the number SNPs had little to no impact on the estimation of genetic offsets; the rank correlations between replicates and the reference estimates resulting from considering all SNPs were similar to those obtained using only 100 outlier SNPs. Overall, we found that increasing the number of populations resulted in higher rank correlations, irrespective of the number of SNPs used in the analyses (Figure 1b). This suggests that the estimation of genetic offsets is likely to be robust to the restrictions imposed by selecting few SNPs with the strongest signals of local adaptation (Aguirre-Liguori et al., 2019; Fitzpatrick et al., 2021; Rhoné et al., 2020).

3.2.2 | Extrapolation to unsampled locations

We assessed the effects of population samples to extrapolate GF models and genetic offsets to unsampled localities. For every replicate and sampling intensity, we estimated the genetic offsets of populations that were not sampled and estimated the rank correlation between extrapolated and reference genetic offsets. As expected, we found that increasing sampling intensity led to a more robust extrapolation of genetic offsets (Figure 1C, Table S2), but this depended on the underlying genetic structure of populations. For example, the median rank correlations were high in species with low genetic structure (*E. trailli* ranged from 0.51 in $N_{pop}=5$ to 1 in $N_{pop}=25$), whereas these were low or even negative in species with high genetic structure (*A. lyrata* ranged from −0.15 in $N_{pop}=5$ to 0.8 in $N_{pop}=30$). Extrapolated genetic offsets deviated more from the reference estimates as sampling intensity decreased ($N<15$) and in species with high F_{ST}; for the two species with more pronounced genetic structures, the capacity to extrapolate remained poor (as indicated by low-rank correlations) even when sampling >20 populations. For example, sampling 15 populations in *Z. mays parviglumis* and *A. lyrata* resulted in 43.3% and 55.5% of the replicates with rank correlations <0.5, respectively. By contrast, the same sampling in *F. sylvatica*, *E. trailli* and *Z. mays mexicana* resulted in 11.6%, 36% and 23% of replicates with rank correlations <0.5, respectively. Thus, sampling high numbers of populations appear to be especially important if the goal is to estimate offsets for localities that have not been explicitly sampled.

3.2.3 | Relevant bioclimatic variables

We evaluated whether sampling populations with varying intensity changed the estimated weight (contribution) of bioclimatic variables to the underlying GF models. Overall, we found that sampling intensity had a substantial impact on the identification of relevant bioclimatic variables (Figure 2). For N_{pop} <10, we found highly inconsistent results across replicates compared with the reference dataset, indicating a poor capacity to identify bioclimatic variables with the highest contribution to model building. Our results suggest that with N_{pop} >10, the contribution of different variables can be approximated more consistently but still with substantial sensitivity to sampling (Figure 2). For example, with $N>10$ the top three most important variables were consistently identified for most replicates (Figure 2c,d).

3.3 | Individual sampling

3.3.1 | Estimation sensitivity

To gauge the impact of sampling different numbers of individuals (N_{ind}) on the estimation of genetic offsets, we used a subsampling approach to test whether genetic offsets were sensitive to individual sampling intensity ($N_{ind}=1$–10 individuals per population sampled). We could not perform these analyses for *A. lyrata* because the SNP data were based on pooled samples within each population (Table 1). For the remaining four species, we found that sampling more than five individuals per population resulted in only a marginal increase in the median rank correlations, albeit with a reduced range of estimates (Figure 3). For example, the median rank correlations were >0.9 for *E. trailli*, *Z. mays parviglumis* and *Z. mays mexicana* when sampling more than five individuals per population (Table S3). However, for the two species with the lowest number of SNPs (*E. trailli* and *F. sylvatica*), estimates of genetic offsets were sensitive to sampling fewer than four individuals (Figure 3a). In *E. trailli*, for example, the mean rank correlations increased from 0.26 to 0.88 when increasing sampling from one to four individuals per population, and 30% of replicates had rank correlations <0.5 when sampling less than two individuals per population. Interestingly, in *F. sylvatica* we observed highly consistent rank correlations across sampling intensities, but the median rank correlation always remained <0.7 (Table S3). In this case, the 20 populations selected for these analyses had a low-rank correlation relative to the reference genetic offsets such that the observed rank correlations were lower than 0.7.

Occasionally it may be possible to add single individuals to represent additional sampling locations, which will provide frequency estimates based on bi-allelic SNPs. We evaluated whether it is helpful to increase the number of populations by adding single individuals from more locations. To do this, we sampled all individuals across 20 populations and then added single individuals from randomly chosen populations (singleton populations), ranging from one to five extra populations. Overall, across our five species, we did not find a consistent effect of increasing the number of singleton populations (Figure 3b), but for some species, it appeared that incorporating singleton populations increased estimation sensitivity. For example, in *Z. mays parviglumis* and *E. trailli* we found that adding one extra population increased the range of rank correlations observed across replicates. Interestingly, for *F. sylvatica* we observed that increasing the number of singleton populations produced genetic offsets that increasingly deviated from the reference estimates (Figure 3b, Table S4). Overall, these results suggest that sampling few individuals ($N_{ind}=4$–5) across many populations, while keeping N_{ind} homogenous across populations, is likely to result in genetic offsets that converge on the reference estimates.

3.4 | Climate uncertainty

Climate models use different circulation models and shared socioeconomic pathways to project climate conditions in the future. In studies of climate change, it is important to incorporate this uncertainty to have a range of future potential scenarios (Foden et al., 2019). For this reason, we assessed the variance in the estimation of genetic offsets across climate change scenarios. For each species, we estimated genetic offsets of populations under 54 future climate scenarios and calculated pairwise rank correlations of genetic offsets

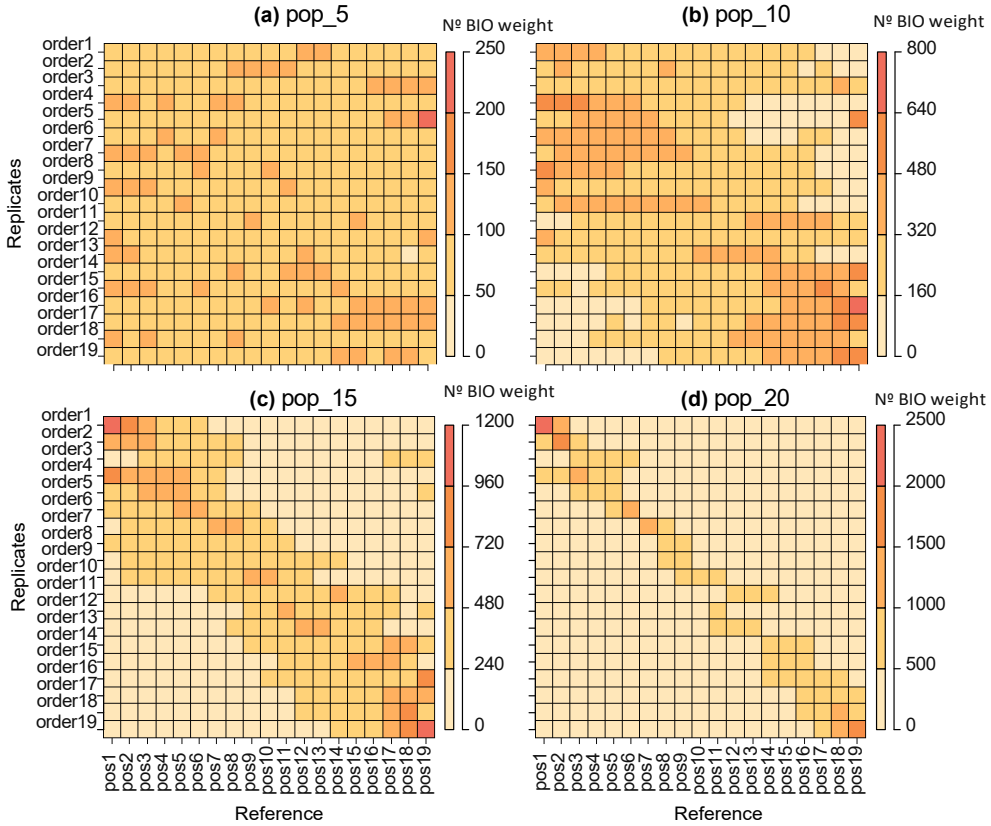

FIGURE 2 Detection sensitivity of the most important bioclimatic variables to varying sampling intensity. The heat maps depict the number of times a bioclimatic variable was selected to have a specific weight (importance) in the construction of the GF model. These maps interrogate across the five species, with separate maps for subsampled (a) 5, (b) 10, (c) 15 and (d) 20 populations. The *x* axis shows the rank order of variables obtained with the reference datasets and the *y*-axis the rank order obtained with the replicates. Warmer colours indicate that a variable was identified more frequently in a given rank order across replicates. Warmer colours across the diagonal indicate a similar rank order of variables in the replicates and the reference. To compare among species, the 19 bioclimatic variables were anonymized and renamed (e.g. order1, order19) according to their weights (importance) in the reference analysis (see Section 2).

among scenarios (Figure 4). Overall, our results indicated that using data from different future climate scenarios did not have a substantial effect on the estimation of genetic offsets. For most species, the use of different circulation models had a marginal impact on the estimation of genetic offsets across populations. However, these impacts also appeared to be species- or region-specific. More specifically, we found weak deviations of genetic offsets from the reference under different models: only 0%–8.73% of replicates (for all species except *F. sylvatica*) had median rank correlations <0.5 under distinct future scenarios. For *F. sylvatica*, the rank correlations among climatic scenarios were highly variable, with 51.9% less than 0.5. This suggests that, at least in some cases, the uncertainty associated with future climate scenarios propagated into the estimation of genetic offsets.

Finally, we explored variations in the projected bioclimatic variables under the 54 climate models for each species. Overall, we found that present-day conditions were strongly correlated to future conditions across all models (Figure S4A–E) but with some variation among variables and species; this pattern could partly explain the high uncertainty in genetic offsets observed across climate models for *F. sylvatica*. To further explore this idea, we estimated the standard deviation (sd) in correlation statistics for each variable and summed the sd across variables (Σsd) for each species. In principle, higher Σsd reflected an overall greater uncertainty among future climate models. We found *F. sylvatica* had the highest Σsd, which explained, to some extent, the high estimation sensitivity of genetic offsets to future climatic models for this species (Figure S4F). The higher variance across climate models for this species did not appear to be explained by either elevation or latitude (Figures S4G and S5).

4 | DISCUSSION

Genetic offsets have become an important metric to evaluate how vulnerable or maladapted populations might respond to ongo-ing climate change (Aguirre-Liguori et al., 2019; Bay et al., 2018; Capblancq, Fitzpatrick, et al., 2020; Ruegg et al., 2018). In this respect, they have become useful to assess priorities for conservation, to design schemes for assisted migrations (Meek et al., 2023) and to implement genetic crosses that may help populations adapt to climate change (Rhoné et al., 2020). However, genetic offsets have not been sufficiently validated, and it remains unclear whether these

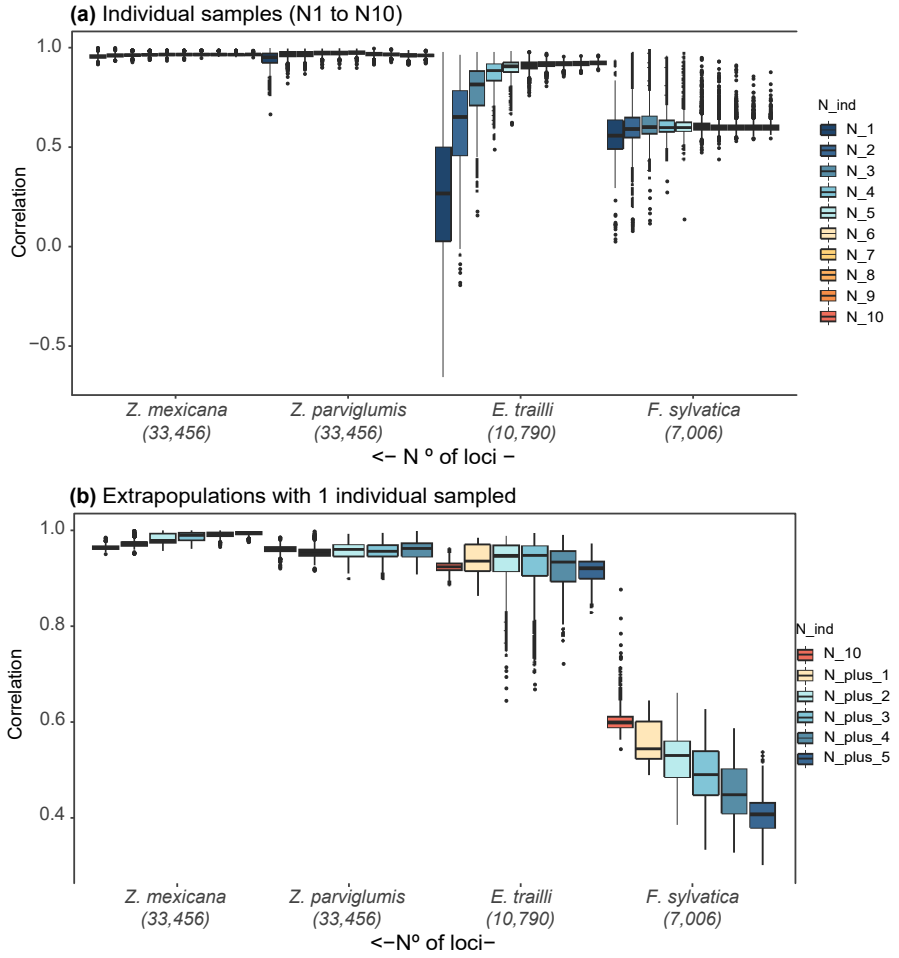

FIGURE 3 Estimation sensitivity of genetic offsets to varying sampling intensity of individuals. In each panel, the boxplots depict the distribution of the rank correlation of genetic offsets between replicates and the reference datasets for each of the four species. Boxplots depict the distribution of the rank correlation of genetic offsets between replicates and the reference datasets for five species. Rank correlations estimated for the sampling replicates using (a) varying the number of sampled individuals per population, based on 20 populations, and (b) varying the number of added singletons—that is, populations with a single sampled individual. One thousand replicates were performed for each test. Rank correlations closer to 1.0 indicate that populations' ranks are similar to those estimated for the reference and also that genetic offsets were less sensitive to inputs. Species are ordered from left to right according to decreasing number of loci (loci). The boxes show the first and third quartiles. The upper and lower whisker show values that are at the limit of 1.5 times the interquartile range. The rest of the dots show the outliers.

metrics are robust and work as intended (Rellstab et al., 2021). In this study we used multiple genomic datasets of species collected across different regions (Figure S1) and with different life-history traits (Table 1), to test how sampling intensity (i.e. the number of populations and individuals) impacts the estimation of genetic offsets and, ultimately, our capacity to set conservation priorities and management strategies.

4.1 | Sampling recommendations

Our results allow us to outline several sampling recommendations and suggest possible avenues to evaluate the estimation sensitivity of genetic offsets in real-life situations. Below we discuss these recommendations and present a markdown file with code (See Supplementary Text S2) that can be used to perform evaluations of sampling adequacy.

4.1.1 | Number of populations

We first tested the effect of sampling different numbers of populations on the estimation of genetic offsets. We have found that sampling more populations increases the rank correlation between the reference and replicate genetic offsets. The effect of increased population sampling is especially prominent for species with complex genetic structures, typified in this study by high F_{ST} (Table 1). These results are not surprising, since many simulation and sampling studies have shown that sampling a higher number of populations decreases sensitivity in the estimation of populations genetic parameters

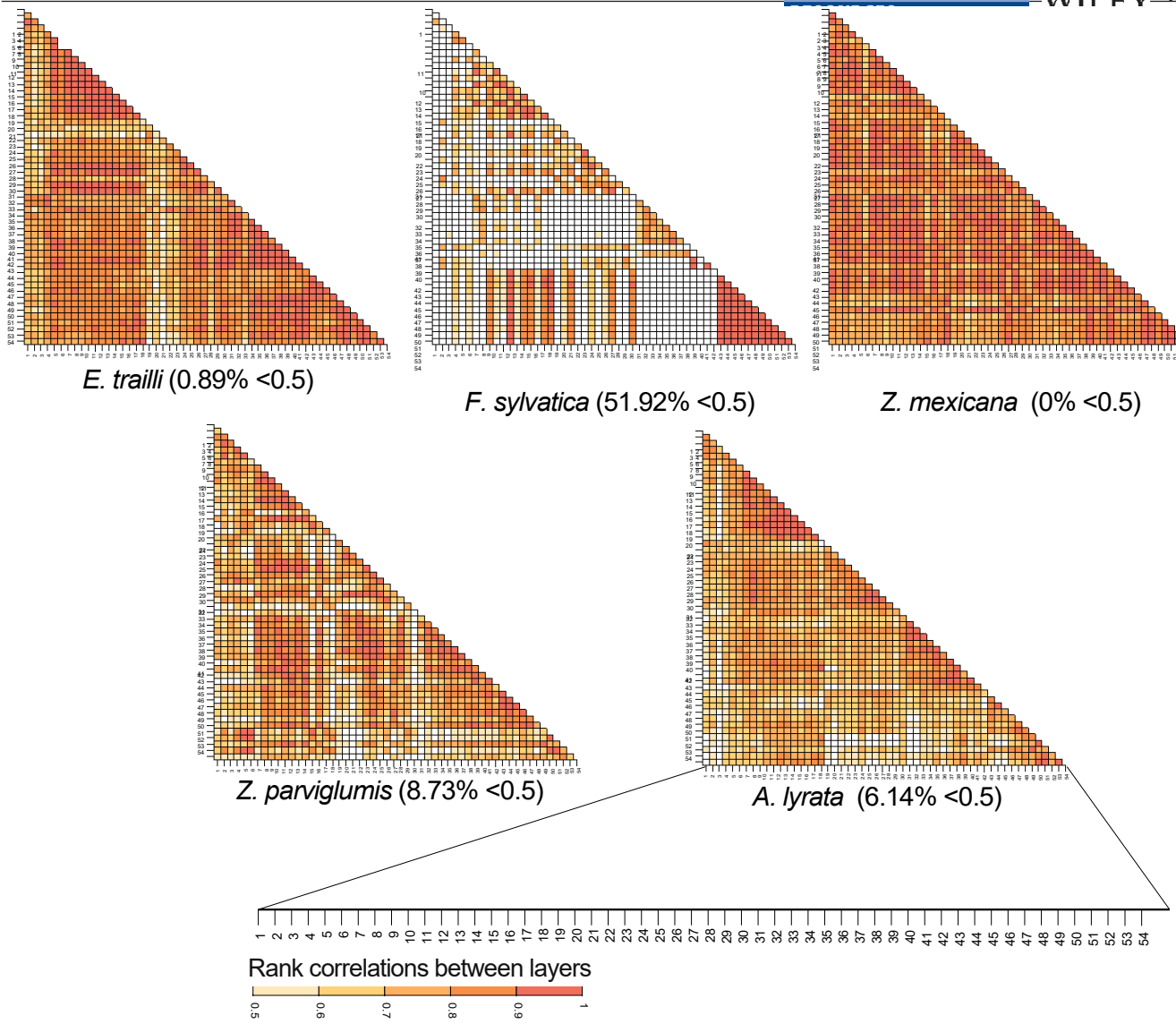

FIGURE 4 Uncertainty in the estimation of genetic offset under future climate scenarios. Heat maps depict the rank correlation of genetic offsets across populations, compared between pairs of 54 future climate scenarios. Warmer colours reflect higher rank correlations between genetic offsets estimated under different models. White squares depict rank correlations <0.5. Climate models in the *y* and *x* axis are coded according to Table S5.

(Aguirre-Liguori et al., 2020; Landguth & Schwartz, 2014; Schwartz & McKelvey, 2008; Willing et al., 2012). It has also been shown that sampling more populations is necessary to identify outlier SNPs and to discard potential false positives (De Mita et al., 2013). Importantly, for all species, we have found that sampling more than 15 populations was enough to retrieve median rank correlations above 0.5 between the reference and replicate datasets (Table 2 shows a summary of the results found and our sampling recommendations). Interestingly, for species with low F_{ST} sampling fewer than 15 populations was enough to converge on the reference patterns of genetic offsets across populations (Figure 1).

Our results also indicate that sampling fewer than 10 populations generated negative or close to zero rank correlations between replicates and the reference dataset, especially in species with high genetic structure. We recognize that gathering such extensive samples can be prohibitive, especially for nonmodel or endangered species or for research groups with limited funding. We therefore have also examined whether an explicit sampling scheme encompassing high geographic or environmental variation, but few populations reduce the bias in genetic offset estimations. We were surprised that this more systematic sampling of populations did not have a major effect on the estimation of genetic offsets (Figure S3). In other words, based on our analyses, maximizing geographic or environmental sampling is not a convenient shortcut that obviates the need for extensive population sampling. We nonetheless recommend that sampling should be designed to maximize the ecological and geographic distributions of populations, because these do impact the ability to detect local adaptation and estimate genetic structure. When only a few populations can be (or are) sampled, we recommend that the limitations of genetic offsets be considered,

TABLE 2 Main observations and recommendations for sampling design.

Test	Main observation	Recommendation
Number of populations	Sampling more populations: • Reduces the sensitivity of genetic offsets, making them more repeatable • Increases the power to extrapolate genetic offsets • Improves detection of relevant bioclimatic variables	At least 15 populations should be sampled to: (1) attain robust estimates of genetic offsets, (2) extrapolate genetic offsets to unsampled locations and (3) identify the most important bioclimatic variables. If genetic structure is high, more populations need to be sampled. A subsampling procedure can be used to evaluate if a plateau of genomic offset is reached; if not, results should recognize this caveat
Number of individuals	• Sampling more individuals has a marginal effect on genetic offset estimates • Adding singleton populations did not decrease the sensitivity in the estimation of genetic offsets	At least 4–5 individuals per population should be sampled, with the number of individuals similar across populations. If new singleton populations are added, it should be tested whether this increase/decrease the sensitivity in the estimation of genetic offsets
Climatic models	For some species, different climatic scenarios predict different rank ordination of genetic offsets among populations	Evaluate the uncertainty associated with different climatic models. Uncertainty should be reported and discussed

especially when implementing conservation and management strategies. In such cases, it may prove more informative to incorporate additional layers of information to inform conservation priorities of populations (e.g. factors associated with life-history traits, demography, niche) (Thurman et al., 2020). Also, for critically endangered species, it might not even be necessary to estimate the genetic offsets of populations, because the goal may be to protect all the known populations of the species.

GF models can project the present and future genetic turnover as a function of bioclimatic variables and thus genetic offsets can be estimated for any geographic area, even when populations have not been sampled. This permits extrapolating to unsampled geographic locations, with the potential to identify areas that are more or less vulnerability to future climate change (Aguirre-Liguori et al., 2021; Capblancq, Fitzpatrick, et al., 2020; Fitzpatrick & Keller, 2015) or that may serve as areas for potential future re-settlement (Aguirre-Liguori et al., 2021; Gougherty et al., 2021; Rhoné et al., 2020). This feature of genetic offsets is particularly attractive because it can be used to design assisted migration programme while avoiding outbreeding depression and maladaptation during the migration process (Aguirre-Liguori et al., 2021; Aitken & Whitlock, 2013; Allendorf et al., 2010). However, we have found that the extrapolation of genetic offsets to unsampled populations is highly sensitive to sampling intensity; more than 15 populations are needed to produce robust estimations of extrapolated genetic offsets with our datasets. This conclusion is especially important in species with complex genetic structures (Figure 1c).

Finally, GF models have the potential to identify bioclimatic variables that are important for the structure of adaptive variation across species ranges (Fitzpatrick & Keller, 2015). Identifying these variables can be valuable to assess the sensitivity of populations to climate change. Therefore, we also tested the capacity of GF models to identify the most important bioclimatic variables under varying levels of sampling intensity. Again, we found that at least 15 populations were needed to be able to accurately identify the most important bioclimatic variables driving allelic turnover across species ranges (Figure 2), but the identification of these variables was consistent with >15 population samples.

Overall, we found that for most species, and particularly for species with complex genetic structures (as defined by increasingly higher F_{ST} values), sampling more than 15 populations serves three purposes: (1) decreases the sensitivity in the estimation of genetic offsets; (2) increases the capacity to extrapolate genetic offsets to unsampled locations; and (3) consistently identifies the most important bioclimatic variables for climate adaptation.

The markdown and code provided in Supporting Text S2 show how to implement our sampling strategy both to evaluate whether the number of sampled populations is sufficient to make a robust estimation of genetic offsets and to identify important bioclimatic variables. Briefly, when the estimation sensitivity of genetic offsets is high (i.e. rank correlations do not converge to 1.0) and the rank order of bioclimatic variables does not maximize the number of values across the diagonal (as in Figure 2), this might be an indication that the number of sampled populations is likely not sufficient, and that caution is warranted for interpreting genetic offsets.

4.1.2 | Number of individuals

With the development of landscape genomics, it has become evident that sampling a greater number of populations is more important than sampling more individuals to adequately identify outlier SNPs (Nazareno et al., 2017; Willing et al., 2012). There is a clear tradeoff between the number of individuals and populations that can be sampled at finite sequencing costs. We have tested if sampling different numbers of individuals affected the estimation of genetic offsets and, as previously suggested by Aguirre-Liguori et al. (2022), found that increasing the number of individuals does not have an important

effect on the estimation of genetic offsets (Figure 3). However, for the two species (*E. trailli* and *F. sylvatica*) with the lowest number of SNPs, we found that sampling more than four individuals per population is needed for high correlations to reference genetic offsets.

Moreover, we have examined whether sampling additional populations, but with only single individuals (singleton populations), increases the estimation sensitivity of genetic offsets. Interestingly, increasing the number of singleton populations does not result in a substantial decrease in estimation sensitivity in our datasets. By contrast, for *F. sylvatica* we have found that increasing the number of singleton populations not only did not decrease the sensitivity but even increased it, with estimated genetic offsets deviating more substantially from the reference. Assuming a tradeoff between the number of sampled populations and the number of sampled individuals per population, our recommendation is to increase the number of populations even if fewer individuals are sampled, as long as at least 4–5 individuals are sampled per population. In addition, we recommend that when additional singleton populations are obtained (i.e. museum specimens or accessions), one should evaluate whether adding samples decreases the sensitivity in the estimation of genetic offsets. The markdown in Supporting Text S2 describes a way to evaluate whether estimation sensitivities decrease with more individuals per population or additional singleton populations. Again, when the estimation sensitivity of genetic offsets remains high or even increases, this would indicate that either more individuals per population are needed or that adding singletons is not advisable.

4.1.3 | Climatic models

Finally, a common problem in studies evaluating the genetic offsets of populations is that different climatic models are rarely incorporated into analyses. This is potentially problematic because future climates can be modelled with different global circulation models and different socioeconomic pathways (Sanderson et al., 2015). This variation leads to intermodel uncertainty that should be explicitly considered, especially in conservation studies (Foden et al., 2019).

For most of our species, we found that using different climatic models to estimate genetic offsets had a negligible effect on the estimation uncertainty of genetic offsets. For most species, we found high rank correlations between genetic offsets estimated using different climatic models (Figure 4). However, for *F. sylvatica*, we observed that more than 50% of the models had pairwise rank correlations below 0.5, indicating that climatic models identify different populations with high climate vulnerability. This uncertainty cannot be explained by the geographic and orographic distribution of *F. sylvatica*, which grows at high latitudes and altitudes (Capblancq, Morin, et al., 2020) (Figures S4G and S5). Instead, we found that correlations between present and future bioclimatic variables were particularly variable for this species, partially explaining the sensitivity of genetic offsets to climate models for *F. sylvatica* (Figure S4F). Even when genetic offsets of most species were not substantially affected using different circulation models, the odd behaviour in *F. sylvatica* shows that this source of uncertainty should be tested and discussed (Table 2). A high genetic offset under a particular climate model may suggest future maladaptation, even if other models yield low genetic offsets. Altogether, the use of multiple climate models to estimate genetic offsets may help identify populations or localities that are consistently projected to be highly vulnerable.

The markdown in Supplementary Text S2 describes a way to estimate the rank correlation between different climatic models and to identify the climate model with consistently high correlations to other models. The model with high correlations can then be used as a reference climatic model.

4.2 | Caveats of our study

Here we were not interested in identifying biological signals, but rather identifying how sampling intensity affects the estimation of genetic offsets. Therefore, the results herein presented should only be interpreted methodologically. Below, we list five caveats that should be considered while interpreting our results. Although it is likely that our results are affected by some of the issues outlined below, in general, we found consistent patterns among datasets.

First, we used the complete datasets to represent reference (or 'true') values of genetic offsets, and thus, our reference parameters are impacted by 'unknown' sampling biases imposed by the original sampling design and sequencing specifications. It is highly likely that the reference datasets do not include all the individuals, genotypes and populations that best describe the genetic composition, geographic range and climatic tolerance of the species (e.g. the *F. sylvatica* dataset only contains a small proportion of populations in the South of France). To reduce misinterpretations and to be able to compare both among species and among realistic sampling scenarios, we decided to perform our replicate samples of individuals and populations based on absolute numbers (N_{pop} = 5, 10, 15, 20, 25 and 30) instead of the percentage of sampled populations. This allowed us to compare realistic sampling sizes and directly compare observed patterns among datasets. We found that, under the same sampling intensity, the estimation sensitivity of genetic offsets was higher for species with more populations originally sampled (e.g. *A. lyrata* with 47 populations vs. *Z. mays mexicana* with 23 populations). This result likely reflects the varying percentage of populations sampled in the replicates (e.g. N_{pop} = 20 corresponds to 48% of *A. lyrata* populations but corresponds to 87% of *Z. mays mexicana* populations). However, the observed impact of lowering sampling intensity was consistent across datasets, supporting previous simulation results suggesting the impacts of sampling design on the estimation of genetic offsets (Láruson et al., 2022).

Second, the design of our study has included thousands of replicates for the estimation of outlier loci and GF models. For each species, we performed more than 20,000 replicates, therefore it was impossible to perform detailed analyses to identify the most supported outlier SNPs while avoiding false positives (De Mita

et al., 2013; Tiffin & Ross-Ibarra, 2014). Instead, we selected the 100 SNPs with the lowest *p*-value according to LFMM2 analyses. However, we believe that the misidentification of outlier SNPs will not affect our conclusions due to the low sensitivity to the selection of SNPs (Figure 1b). Moreover, while studies are needed to show that randomly selected SNPs are representative of the genomic landscape, previous studies have shown that randomly selected SNPs can be sufficient to estimate genetic offsets and identify patterns associated with changes in fitness (Aguirre-Liguori et al., 2019; Fitzpatrick et al., 2021; Rhoné et al., 2020). Along these lines, we did not attempt to perform robust SNP callings and instead employed the widely used *stacks* pipeline to identify SNPs (Catchen et al., 2013). Our approach to identify SNPs has the advantage that it can be used to obtain polymorphic SNPs for nonmodel species without a reference genome (Catchen et al., 2013).

A third caveat is that we were limited by the type of data available to perform the analyses. We downloaded data obtained with different genomic tools and with varying ascertainment biases. Ascertainment bias has been found to affect the estimation of demographic parameters estimated from the site frequency spectrum, yet it appears that F_{ST} is not strongly affected by ascertainment bias (Albrechtsen et al., 2010). It is interesting to note that both *E. trailli* and *A. lyrata* had few low-frequency SNPs (Figure S6). These might explain why estimation of genetic offsets in *E. trailli* and *A. lyrata* was highly sensitive to individual and population sampling, respectively (Figures 1 and 3a). However, we currently have no information to make inferences about the sensitivity of genetic offsets to ascertainment bias.

Fourth, the genetic structure and genetic diversity of populations can depend on the life-history traits of species (Aguinagalde et al., 2005; Nybom, 2004). While the data that we analysed did not allow us to conclude life-history traits could be important in the estimation of genetic offsets, it will be interesting to test whether generation time, range sizes, dispersal, reproductive system and other life-history traits affect the sensitivity in the estimation of genetic offsets.

Finally, genetic offsets have a more profound limitation. These are measures of the predicted level of the future maladaptation of populations (Capblanc et al., 2021), but assume that maladaptation occurs because of deviations from the optimum fitness values. However, local adaptation to climate change is a very complex process in which fitness is affected by other nonselective processes, such as inbreeding, genetic drift and genomic load, all of which are more impactful on small populations (Aguirre-Liguori et al., 2021; Willi et al., 2022). In general, these small populations are the focus of conservation priorities but potentially have low adaptive capac-ity that limits the utility of genomic offsets (Meek et al., 2023). Moreover, it is increasingly being reported that species-climate relationships vary across space or time (Schultz et al., 2022; Smith et al., 2019), and therefore, different populations could be affected by different processes. Finally, populations can tolerate changes in climate before suffering reductions in fitness through phenotypic plasticity or standing genetic variation, which are not considered by the approach to estimate genetic offsets.

AUTHOR CONTRIBUTIONS

JAAL and SRB conceived the idea and designed the experiments. JAAL and AMC wrote the code and performed the analyses. JAAL wrote the first draft of the manuscript with contributions from SRB. All authors interpreted the results and contributed to the writing of the final draft of the manuscript.

ACKNOWLEDGEMENTS

This manuscript was supported by the Consejo Nacional de Ciencia y Tecnología (CONACYT) Project FORDECYT-PRONACES/263962/2020 granted to SRB. JAAL received funding from CONACYT in the form of a postdoctoral fellowship (Project: FORDECYT-PRONACES/263962/2020 register number 28416). We also thank the editor and two anonymous reviewers that helped improve the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All analyses were performed using published datasets. Sequences from (1) the *Arabidopsis lyrata* dataset were downloaded from the European National Archive under Bioproject ID: PRJEB19338 and (2) the *Empidonax traillii* dataset were downloaded from the NCBI under Bioproject: PRJNA453612. The genotypes of *Fagus sylvatica* were downloaded from the dryad link: https://datadryad.org/stash/dataset/doi:10.5061/dryad.pc866t1k5. The genotypes from the *Zea mays* ssp. *mexicana* and *Zea mays* ssp. *parviglumis* were downloaded from the dryad link: https://datadryad.org/stash/dataset/doi:10.5061/dryad.tf556. All codes used to perform the analyses are available at https://github.com/spiritu-santi/GradientForest_Sampling.

BENEFIT-SHARING STATEMENT

The data analysed were downloaded from published data, so the benefit-sharing statement does not apply.

ORCID

Jonás A. Aguirre-Liguori https://orcid.org/0000-0003-1763-044X
Brandon S. Gaut https://orcid.org/0000-0002-1334-5556
Santiago Ramírez-Barahona https://orcid.org/0000-0003-3999-6952

REFERENCES

Aguinagalde, I., Hampe, A., Mohanty, A., Martin, J. P., Duminil, J., & Petit, R. J. (2005). Effects of life-history traits and species distribution on genetic structure at maternally inherited markers in European trees and shrubs. *Journal of Biogeography*, *32*(2), 329–339. https://doi.org/10.1111/j.1365-2699.2004.01178.x

Aguirre-Liguori, J. A., Luna-Sánchez, J. A., Gasca-Pineda, J., & Eguiarte, L. E. (2020). Evaluation of the minimum sampling design for population genomic and microsatellite studies: An analysis based on wild maize. *Frontiers in Genetics*, *11*, 870. https://doi.org/10.3389/fgene.2020.00870

Aguirre-Liguori, J. A., Morales-Cruz, A., & Gaut, B. S. (2022). Evaluating the persistence and utility of five wild *Vitis* species in the context of climate change. *Molecular Ecology*, *31*, 6457–6472. https://doi.org/10.1111/mec.16715

Aguirre-Liguori, J. A., Ramírez-Barahona, S., & Gaut, B. S. (2021). The evolutionary genomics of species' responses to climate change. *Nature Ecology & Evolution*, *10*(5), 1350–1360. https://doi.org/10.1038/s41559-021-01526-9

Aguirre-Liguori, J. A., Tenaillon, M. I., Vázquez-Lobo, A., Gaut, B. S., Jaramillo-Correa, J. P., Montes-Hernandez, S., Souza, V., & Eguiarte, L. E. (2017). Connecting genomic patterns of local adaptation and niche suitability in teosintes. *Molecular Ecology*, *16*(26), 4226–4240.

Aguirre-Liguori, J. A., Ramírez-Barahona, S., Tiffin, P., & Eguiarte, L. E. (2019). Climate change is predicted to disrupt patterns of local adaptation in wild and cultivated maize. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1906), 20190486. https://doi.org/10.1098/rspb.2019.0486

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

Albrechtsen, A., Nielsen, F. C., & Nielsen, R. (2010). Ascertainment biases in SNP chips affect measures of population divergence. *Molecular Biology and Evolution*, *27*(11), 2534–2547. https://doi.org/10.1093/molbev/msq148

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews. Genetics*, *11*(10), 697–709. https://doi.org/10.1038/nrg2844

Barrett, R. D. H., & Schluter, D. (2007). Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, *23*(1), 38–44. https://doi.org/10.1016/j.tree.2007.09.008

Bay, R. A., Harrigan, R. J., Le Underwood, V., Gibbs, H. L., Smith, T. B., & Ruegg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science*, *359*(6401), 83–86. https://doi.org/10.1126/science.aat7956

Bradburd, G. S., Ralph, P. L., & Coop, G. M. (2013). Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution*, *67*(11), 3258–3273. https://doi.org/10.1111/evo.12193

Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M., & Keller, S. R. (2020). Genomic prediction of (mal)adaptation across current and future climatic landscapes. *Annual Review of Ecology, Evolution, and Systematics*, *51*, 245–269. https://doi.org/10.1146/annurev-ecolsys-020720-042553

Capblancq, T., Morin, X., Gueguen, M., Renaud, J., Lobreaux, S., & Bazin, E. (2020). Climate-associated genetic variation in *Fagus sylvatica* and potential responses to climate change in the French Alps. *Journal of Evolutionary Biology*, *33*(6), 783–796. https://doi.org/10.1111/jeb.13610

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Caye, K., Jumentier, B., Lepeule, J., & François, O. (2019). LFMM 2: Fast and accurate inference of gene-environment associations in genome-wide studies. *Molecular Biology and Evolution*, *36*(4), 852–860. https://doi.org/10.1093/molbev/msz008

Ceballos, G., Ehrlich, P. R., & Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), E6089–E6096. https://doi.org/10.1073/pnas.1704949114

De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, *22*(5), 1383–1399. https://doi.org/10.1111/mec.12182

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. *Molecular Ecology*, *17*(5), 1170–1188. https://doi.org/10.1111/j.1365-294X.2007.03659.x

Ellis, N., Smith, S. J., & Roland Pitcher, C. (2012). Gradient forests: Calculating importance gradients on physical predictors. *Ecology*, *93*(1), 156–168. https://doi.org/10.1890/11-0252.1

Exposito-Alonso, M., Burbano, H. A., Bossdorf, O., Nielsen, R., & Weigel, D. (2019). Natural selection on the *Arabidopsis thaliana* genome in present and future climates. *Nature*, *573*(7772), 126–129. https://doi.org/10.1038/s41586-019-1520-9

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model Intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fitzpatrick, M. C., Chhatre, V. E., Soolanayakanahally, R. Y., & Keller, S. R. (2021). Experimental support for genomic prediction of climate maladaptation using the machine learning approach gradient forests. *Molecular Ecology Resources*, *21*(8), 2749–2765. https://doi.org/10.1111/1755-0998.13374

Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, *18*(1), 1–16. https://doi.org/10.1111/ele.12376

Fitzpatrick, M. C., Keller, S. R., & Lotterhos, K. E. (2018). Comment on "Genomic signals of selection predict climate-driven population declines in a migratory bird.". *Science*, *361*(6401), eaat7956. https://doi.org/10.1126/science.aat7956

Foden, W. B., Young, B. E., Akçakaya, H. R., Garcia, R. A., Hoffmann, A. A., Stein, B. A., Thomas, C. D., Wheatley, C. J., Bickford, D., Carr, J. A., Hole, D. G., Martin, T. G., Pacifici, M., Pearce-Higgins, J. W., Platts, P. J., Visconti, P., Watson, J. E. M., & Huntley, B. (2019). Climate change vulnerability assessment of species. *Wiley Interdisciplinary Reviews: Climate Change*, *10*(1), e551. https://doi.org/10.1002/wcc.551

Forester, B. R., Beever, E. A., Darst, C., Szymanski, J., & Funk, W. C. (2022). Linking evolutionary potential to extinction risk: Applications and future directions. *Frontiers in Ecology and the Environment*, *20*(9), 507–515. https://doi.org/10.1002/fee.2552

Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, *126*(2), 131–140. https://doi.org/10.1016/j.biocon.2005.05.002

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*(8), 925–929. https://doi.org/10.1111/2041-210X.12382

Gougherty, A. V., Keller, S. R., & Fitzpatrick, M. C. (2021). Maladaptation, migration and extirpation fuel climate change risk in a forest tree species. *Nature Climate Change*, *11*(2), 166–171. https://doi.org/10.1038/s41558-020-00968-6

Hampe, A., & Petit, R. J. (2005). Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, *8*(5), 461–467. https://doi.org/10.1111/j.1461-0248.2005.00739.x

IPCC. (2022). Climate change 2022: Impacts, adaptation and vulnerability. In H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama (Eds.), *Contribution of working group II*

to the sixth assessment report of the intergovernmental panel on climate change (p. 3056). Cambridge University Press. https://doi.org/10.1017/9781009325844

Landguth, E. L., & Schwartz, M. K. (2014). Evaluating sample allocation and effort in detecting population differentiation for discrete and continuously distributed individuals. *Conservation Genetics*, *15*(4), 981–992. https://doi.org/10.1007/s10592-014-0593-0

Láruson, Á. J., Fitzpatrick, M. C., Keller, S. R., Haller, B. C., & Lotterhos, K. E. (2022). Seeing the forest for the trees: Assessing genetic offset predictions from gradient forest. *Evolutionary Applications*, *15*(3), 403–416. https://doi.org/10.1111/eva.13354

LeDee, O. E., Handler, S. D., Hoving, C. L., Swanston, C. W., & Zuckerberg, B. (2021). Preparing wildlife for climate change: How far have we come? *Journal of Wildlife Management*, *85*(1), 7–16. https://doi.org/10.1002/jwmg.21969

Lira-Noriega, A., & Manthey, J. D. (2014). Relationship of genetic diversity and niche centrality: A survey and analysis. *Evolution*, *68*(4), 1082–1093. https://doi.org/10.1111/evo.12343

Meek, M. H., Beever, E. A., Barbosa, S., Fitzpatrick, S. W., Fletcher, N. K., Mittan-Moreau, C. S., Reid, B. N., Campbell–Staton, S. C., Green, N. F., & Hellmann, J. J. (2023). Understanding local adaptation to prepare populations for climate change. *Bioscience*, *73*(1), 36–47. https://doi.org/10.1093/biosci/biac101

Meirmans, P. G. (2015). Seven common mistakes in population genet-ics and how to avoid them. *Molecular Ecology*, *24*(13), 3223–3231. https://doi.org/10.1111/mec.13243

Nazareno, A. G., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: An empirical study from an Amazonian plant species. *Molecular Ecology Resources*, *17*(6), 1136–1147. https://doi.org/10.1111/1755-0998.12654

Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, *13*(5), 1143–1155. https://doi.org/10.1111/j.1365-294X.2004.02141.x

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, *37*(1), 637–669. https://doi.org/10.1146/annurev.ecolsys.37.091305.110100

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change. *Nature*, *421*, 37–42. https://doi.org/10.1038/nature01286

Peterson, A. T., Soberón, J., Richard G. P., Robert P. A., Martínez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). *Ecological niches and geographic distributions*. Princeton University Press.

Rellstab, C., Dauphin, B., & Exposito-Alonso, M. (2021). Prospects and limitations of genomic offset in conservation management. *Evolutionary Applications*, *14*(5), 1–11. https://doi.org/10.1111/eva.13205

Rhoné, B., Defrance, D., Berthouly-Salazar, C., Mariac, C., Cubry, P., Couderc, M., Dequincey, A., Assoumanne, A., Kane, N. A., Sultan, B., Barnaud, A., & Vigouroux, Y. (2020). Pearl millet genomic vulnerability to climate change in West Africa highlights the need for regional collaboration. *Nature Communications*, *11*(1), 5274. https://doi.org/10.1038/s41467-020-19066-4

Ruegg, K., Bay, R. A., Anderson, E. C., Saracco, J. F., Harrigan, R. J., Whitfield, M., Paxton, E. H., & Smith, T. B. (2018). Ecological genomics predicts climate vulnerability in an endangered southwestern songbird. *Ecology Letters*, *21*(7), 1085–1096. https://doi.org/10.1111/ele.12977

Sanderson, B. M., Knutti, R., & Caldwell, P. (2015). A representative democracy to reduce interdependency in a multimodel ensemble. *Journal of Climate*, *28*(13), 5171–5194. https://doi.org/10.1175/JCLI-D-14-00362.1

Schoville, S. D., Bonin, A., Francois, O., Lobreaux, S., Melodelima, C., & Manel, S. (2012). Adaptive genetic variation on the landscape: Methods and cases. *Annual Review of Ecology, Evolution, and Systematics*, *43*(43), 23–43. https://doi.org/10.1146/annurev-ecolsys-110411-160248

Schultz, E. L., Hülsmann, L., Pillet, M. D., Hartig, F., Breshears, D. D., Record, S., Shaw, J. D., DeRose, R. J., Zuidema, P. A., & Evans, M. E. K. (2022). Climate-driven, but dynamic and complex? A reconciliation of competing hypotheses for species' distributions. *Ecology Letters*, *25*(1), 38–51. https://doi.org/10.1111/ele.13902

Schwartz, M. K., & McKelvey, K. S. (2008). Why sampling scheme matters: The effect of sampling scheme on landscape genetic results. *Conservation Genetics*, *10*(2), 441–452. https://doi.org/10.1007/s10592-008-9622-1

Smith, A. B., Beever, E. A., Kessler, A. E., Johnston, A. N., Ray, C., Epps, C. W., Lanier, H. C., Klinger, R. C., Rodhouse, T. J., Varner, J., Perrine, J. D., Seglund, A., Hall, L. E., Galbreath, K., MacGlover, C., Billman, P., Blatz, G., Brewer, J., Castillo Vardaro, J., & Yandow, L. (2019). Alternatives to genetic affinity as a context for within-species response to climate. *Nature Climate Change*, *9*(10), 787–794. https://doi.org/10.1038/s41558-019-0584-8

Stapley, J., Reger, J., Feulner, P. G., Smadja, C., Galindo, J., Ekblom, R., Bennison, C., Ball, A. D., Beckerman, A. P., & Slate, J. (2010). Adaptation genomics: The next generation. *Trends in Ecology and Evolution*, *25*(12), 705–712. https://doi.org/10.1016/j.tree.2010.09.002

Thompson, L. M., Thurman, L. L., Cook, C. N., Beever, E. A., Sgrò, C. M., Battles, A., Botero, C. A., Gross, J. E., Hall, K. R., Hendry, A. P., Hoffmann, A. A., Hoving, C., LeDee, O. E., Menglet, C., Nicotra, A. B., Niver, R. A., Pérez-Jvostov, F., Quiñones, R. M., Schuurman, G. W., … Whiteley, A. (2023). Connecting research and practice to enhance the evolutionary potential of species under climate change. Conservation science and practice. *Conservation Science and Practice*, *5*, e12855. https://doi.org/10.1111/csp2.12855

Thurman, L. L., Gross, J. E., Mengelt, C., Beever, E. A., Thompson, L. M., Schuurman, G. W., Hoving, C. L., & Olden, J. D. (2022). Applying assessments of adaptive capacity to inform natural-resource management in a changing climate. *Conservation Biology*, *36*(2), e13838. https://doi.org/10.1111/cobi.13838

Thurman, L. L., Stein, B. A., Beever, E. A., Foden, W., Geange, S. R., Green, N., Gross, J. E., Lawrence, D. J., LeDee, O., Olden, J. D., Thompson, L. M., & Young, B. E. (2020). Persist in place or shift in space? Evaluating the adaptive capacity of species to climate change. *Frontiers in Ecology and the Environment*, *18*(9), 520–528. https://doi.org/10.1002/fee.2253

Tiffin, P., & Ross-Ibarra, J. (2014). Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology and Evolution*, *29*(12), 673–680. https://doi.org/10.1016/j.tree.2014.10.004

Tobón-Niedfeldt, W., Mastretta-Yanes, A., Urquiza-Haas, T., Goettsch, B., Cuervo-Robayo, A. P., & Urquiza-Haas, E. (2022). Incorporating evolutionary and threat pro- cesses into crop wild relatives conservation. *Nature Communications*, *13*, 6254. https://doi.org/10.1038/s41467-022-33703-0

Waldvogel, A., Rellstab, C., Bataillon, T., Feldmeyer, B., Rolshausen, G., Exposito-alonso, M., Rellstab, C., Kofler, R., Mock, T., Schmid, K., Schmitt, I., Bataillon, T., Savolainen, O., Bergland, A., Flatt, T., Guillaume, F., & Pfenninger, M. (2019). Evolutionary genomics can improve prediction of species' responses to climate change. *Evoution Letters*, *4*(1), 4–18. https://doi.org/10.1002/evl3.154

Willi, Y., Fracassetti, M., Zoller, S., & Van Buskirk, J. (2018). Accumulation of mutational load at the edges of a species range. *Molecular Biology and Evolution*, *35*(4), 781–791. https://doi.org/10.1093/molbev/msy003

Willi, Y., Kristensen, T. N., Sgro, C. M., Weeks, A. R., Ørsted, M., & Hoffmann, A. A. (2022). Conservation genetics as a management tool: The five best-supported paradigms to assist the management of threatened species. *Proceedings of the National Academy*

of Sciences of the United States of America, *119*(1), 1–10. https://doi.org/10.1073/pnas.2105076119

Willing, E.-M., Dreyer, C., & van Oosterhout, C. (2012). Estimates of genetic differentiation measured by fst do not necessarily require large sample sizes when using many snp markers. *PLoS One*, *7*(8), e42649. https://doi.org/10.1371/journal.pone.0042649

Wuest, S. E., Peter, R., & Niklaus, P. A. (2021). Ecological and evolutionary approaches to improving crop variety mixtures. *Nature Ecology and Evolution*, *5*(8), 1068–1077. https://doi.org/10.1038/s41559-021-01497-x

Zamora-Gutiérrez, V., Rivera-Villanueva, A. N., Martínez Balvanera, S., Castro-Castro, A., & Aguirre-Gutiérrez, J. (2021). Vulnerability of bat-plant pollination interactions due to environmental change. *Global Change Biology*, *27*(14), 3367–3382. https://doi.org/10.1111/gcb.15611

SUPPORTING INFORMATION

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

How to cite this article: Aguirre-Liguori, J. A., Morales-Cruz, A., Gaut, B. S., & Ramírez-Barahona, S. (2023). Sampling effect in predicting the evolutionary response of populations to climate change. *Molecular Ecology Resources*, *00*, 1–17. https://doi.org/10.1111/1755-0998.13828