PERSPECTIVE
https://doi.org/10.1038/s41559-021-01526-9

nature ecology & evolution

The evolutionary genomics of species' responses to climate change

Jonás A. Aguirre-Liguori[1], Santiago Ramírez-Barahona[2] and Brandon S. Gaut[1]✉

Climate change is a threat to biodiversity. One way that this threat manifests is through pronounced shifts in the geographical range of species over time. To predict these shifts, researchers have primarily used species distribution models. However, these models are based on assumptions of niche conservatism and do not consider evolutionary processes, potentially limiting their accuracy and value. To incorporate evolution into the prediction of species' responses to climate change, researchers have turned to landscape genomic data and examined information about local genetic adaptation using climate models. Although this is an important advancement, this approach currently does not include other evolutionary processes—such as gene flow, population dispersal and genomic load—that are critical for predicting the fate of species across the landscape. Here, we briefly review the current practices for the use of species distribution models and for incorporating local adaptation. We next discuss the rationale and theory for considering additional processes, reviewing how they can be incorporated into studies of species' responses to climate change. We summarize with a conceptual framework of how manifold layers of information can be combined to predict the potential response of specific populations to climate change. We illustrate all of the topics using an exemplar dataset and provide the source code as potential tutorials. This Perspective is intended to be a step towards a more comprehensive integration of population genomics with climate change science.

Climate change contributes to the loss of biodiversity[1–6], highlighting a pressing need to identify how populations respond to climate change and to use this information to conserve species. To date, these needs have been addressed primarily using species distribution models (SDMs)[7–9]. The principal goals of SDMs are to identify climatic factors that define the geographical distribution of a species and then to predict shifts in the distributions of species over time and space as a function of climate change[8–10]. However, at their most basic, SDMs assume that all individuals and populations within a species are equivalent, and they ignore the potential for evolution. The forces that affect evolution are complex, and SDMs are not able to incorporate this complexity[1,11,12].

Recent research, prompted by a series of landmark studies[13–15], has integrated landscape genomic data with global circulation models that project future climatic conditions. This work recognizes that the response of populations to climate change can vary across a species' range and also emphasizes the potential importance of local adaptation to specific environments. However, the evolutionary fate of species does not rely on adaptation alone, because evolution is a function of many interacting factors[16]. The integration of evolutionary forces beyond adaptation may help to predict the fate of species, provide additional insights into evolutionary dynamics, and ultimately guide conservation and management[17–20].

In this Perspective, we have two goals: (1) to argue that additional evolutionary forces should be incorporated into predicting the response of species to climate change, and (2) to make concrete suggestions to achieve that end. To do so, we begin by briefly reviewing SDMs and the emerging synthesis of landscape genomics, local adaptation and climate projections. These topics have been reviewed recently[18,21–24], but we cover them here to lay a foundation for our proposed extensions. Some of these reviews have mentioned the need to consider additional evolutionary processes such as gene flow in predictions of species' responses to climate change[18,23,24], but they have not suggested potential solutions. Accordingly, in the second part of this Perspective we discuss three evolutionary processes—gene flow, population dispersal and genetic load—and suggest ways to build them into a framework for assessing the climate change vulnerability of populations. We illustrate our suggestions using an exemplar dataset from the wild plant species teosinte (*Zea mays* ssp. *mexicana*; hereafter *mexicana*) and provide code to perform all analyses. By providing background information, examples and code, we intend for this work to be accessible for researchers who seek to consolidate landscape genomics into their work or, conversely, for population genomicists who want to merge their data with climate projections. We conclude by proposing a conceptual framework to evaluate the responses of species to climate change.

Genetic structure and species distribution models

Correlative SDMs have been widely used to estimate the distribution of species across space and time. In essence, the climatic conditions in which a species occurs are used to construct a model of the species' ecological niche that is then projected into geographical space. The use and misuse of SDMs has been reviewed previously[1,8,10,16]; here we provide a brief overview to highlight pertinent information. The theory behind SDMs can be summarized with the BAM diagram[8,25], which depicts the geographical distribution of a species in three components: suitable areas in terms of biotic factors (B), abiotic factors (A) and areas that are accessible through migration (M) (Fig. 1a). The intersection among these components (G_O) characterizes regions in geographical space that are suitable for positive population growth. In the context of climate change, SDMs forecast how G_O will change under future environmental conditions by comparing present-day and future models. To achieve this, SDMs rely on several simplifying assumptions[8,26,27], primarily that (1) all the abiotically suitable areas also represent a suitable biotic environment, including biotic interactions and interactors ($A \approx B$);

---

[1]Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA. [2]Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México (UNAM), Mexico City, Mexico. ✉e-mail: bgaut@uci.edu

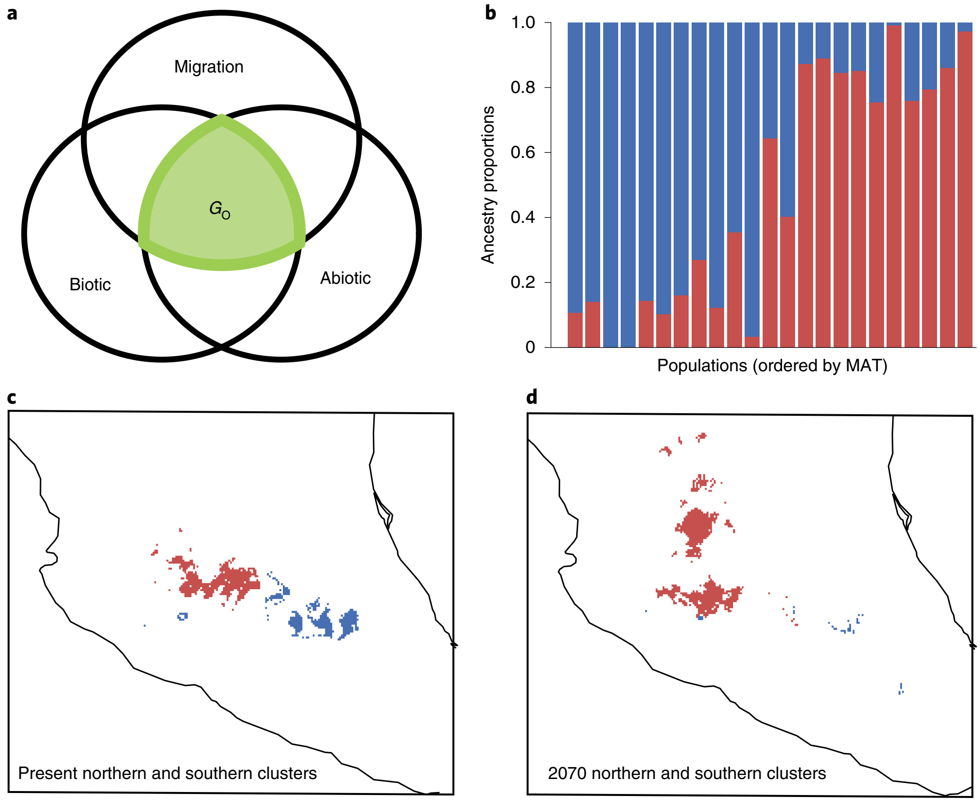

Fig. 1 | Genetic structure and species distribution models. a, The BAM model—a theoretical construction that integrates the biotic, abiotic and migration potential of species. G_O indicates the intersection in geographical space where population growth is positive. **b**, Genetic clustering of *mexicana* identifies two genetic clusters ($K = 2$), the southern cluster (blue) and the northern cluster (red). The bars are ordered by the mean annual temperature (MAT) in which populations grow. **c**,**d**, SDMs project the geographical range of southern (blue) and northern (red) genetic clusters in the present (**c**) and future (2070) (**d**).

(2) species' distributions are in equilibrium to environmental conditions; and (3) all the accessible space will be fully (and equally) occupied by a species.

The BAM model also assumes that all populations respond to climate change similarly. In other words, the projection of SDMs into different time periods assumes both niche conservation and spatial homogeneity of populations relative to climate and other biotic factors[8,27]. The problem with the assumption of homogeneity is it ignores how genetic factors affect the way a species (or its populations) interacts with *B*, *A* and *M* across both space and time. Ultimately, the probability of survival and dispersal across space and time is a function of environmental suitability and of the adaptive landscape. SDMs omit the latter and implicitly assume that genetic heterogeneity acts only to generate random and spatially uncorrelated noise between species' occurrence and the environment.

One advance has been to incorporate spatial genetic structure into SDMs[28,29] by (1) identifying genetic clusters using standard population genetics methods[30–32]; (2) building SDMs for individual clusters; and (3) comparing the per-cluster SDMs between the present and future to predict species' responses to climate change[22,33–36]. We illustrate this approach using the *mexicana* dataset, which consists of 33,454 SNPs from 348 individuals representing 23 populations from regions of Central Mexico[37,38] (Markdown 1, available in Supplementary Information). Genetic clustering (Markdown 2 Supplementary Information) indicates that populations can be divided roughly into two groups (Fig. 1b): a northern cluster (N) from warmer lowland regions and a southern cluster (S) in cooler highland regions. However, note that few populations consist of genetic components that are solely attributable to one cluster, emphasizing that clustering is inherently approximate. We next constructed SDMs for each cluster separately using data from the WorldClim database[39] that averages the climate over the recent past (1970–2010; Markdowns 1 and 2, Supplementary Information). The results show that the present-day SDMs for the two genetic clusters predict different areas of potential distribution (Fig. 1c), suggesting that the two clusters occupy non-overlapping areas of the available climatic space. Importantly, differences between the clusters are magnified when SDMs are projected to the year 2070 (Fig. 1d) because the future projected distribution is expected to increase slightly in area for the warm-adapted northern cluster but decline markedly for the cold-adapted southern cluster.

Similar to previous studies[12,14,22,33,35], the *mexicana* example supports the idea that different genetic clusters (or populations) vary in their relationships with environmental conditions (but see ref. [28]). By doing so, this example illustrates one of the many limitations of SDMs and the BAM model, specifically, the assumption that all populations are identical and will respond similarly to climate change. We nonetheless embrace the intellectual construct of the BAM model, especially its emphasis on overlapping criteria to consider the fate of a species. Below we will introduce a similar conceptual model that incorporates population-level evolutionary processes.

Incorporating local adaptation

The identification of genetic clusters provides additional resolution to SDMs, but they treat landscape genomic data as binary, static predictors of inclusion in a cluster. This approach ignores the fact that genomic data are a rich source of information about evolutionary processes. More specifically, the frequencies of alleles across populations contain clues about the genetics of adaptation to climate[40,41]. The basic adaptive premise is simple: if a single SNP (for example, an A to G variant) is at a low frequency in one population but at a high frequency in a second population, the SNP may contribute to adaptation in the local environment of one or both populations. Evidence for local adaptation is strengthened when the observed allele frequency differences between populations exceed the expectation on the basis of genetic drift and when allele frequencies are correlated with bioclimatic variables such as temperature, precipitation or other factors. However, it is important to accentuate that neutral and selective forces often covary in space, so that variant frequencies are shaped by evolutionary processes other than selection, including population history, drift and gene flow. For example, another potential explanation for the A to G variant is that allelic frequencies track an historical split between populations, rather than adaptation, and therefore reflect the combined effects of population history and genetic drift.

If the goal is to incorporate adaptation into predictions about the fate of populations under climate change, one approach is to filter genomic data to identify a subset of SNPs that is particularly likely to contribute to local adaptation. This filtering step can be completed using standard population genetic tools[30,42–44] that identify SNPs with allele frequency patterns that are not easily attributable to population structure and genetic drift[45–48]. For example, one method (bayescenv[42]) calculates genetic divergence among populations and then partitions individual SNPs into one of three classes: (1) neutral SNPs that reflect population history; (2) SNPs that do not fit the inferred pattern of population history (and are therefore likely to have been affected by forces beyond demography and drift); and (3) SNPs that also associate with environmental data, suggesting they are particularly strong candidates for contributing to local adaptation. Such methods can be applied either on their own or in combination. For example, Aguirre-Liguori et al.[38] applied two separate tools[42,43] to the *mexicana* dataset and identified a set of 32 putatively adaptive (or candidate) SNPs from both methods (Markdown 3.1, Supplementary Information). It is worth mentioning that this filtering step may not be able to discriminate between causative SNPs that contribute to adaptive change and variants that are linked to adaptive SNPs. Moreover, the filtering step contains an inherent trade-off between retaining a high number of SNPs, including numerous false positives, or tolerating a high rate of false negatives to focus on a particularly robust set of candidates. We suspect that the use of two filtering methods, as applied to *mexicana*, leans towards the latter but, more generally, procedures can be altered depending on project objectives (reviewed in ref. [23]).

Once identified, candidate SNPs are used to predict how population frequencies will shift under climate change.

This is done by building a model that identifies nonlinear associations between allele frequencies and environmental variables and then using those associations to project allele frequencies into the future, on the basis of predicted changes in bioclimatic variables. To address these steps, while efficiently handling the potential for massive amounts of data, researchers have turned to machine learning algorithms, such as gradient forest (GF)[13] and random forest[14]. As explained in recent publications[13,23,24,49], the goal of GF is to take SNP allele frequencies (as response variables) and climatic data (as predictors) to identify environmental gradients that associate with genetic variation and also to determine how allele frequencies turnover along that gradient[13,18,23]. To illustrate the approach, we applied GF to the 32 candidate *mexicana* SNPs and to a separate set of 500 neutral reference SNPs for comparison (Markdown 3.2, Supplementary Information) using contemporary bioclimatic data as predictors. On the basis of these analyses, the allele frequencies of candidate SNPs across populations are most strongly correlated with one climatic variable (BIO9, the mean temperature of the driest quarter) but also with others (Fig. 2a). A graph of BIO9 against projected allelic frequencies (Fig. 2b) demonstrates that: (1) candidate SNPs have a stronger association with BIO9 than the reference SNPs, providing some reassurance that the former reflect some feature of local adaptation to climate; (2) allelic frequencies for candidate SNPs turnover in regions in which BIO9 ranges between ~14 and 16 °C; and (3) candidate SNPs differ substantially in frequency between northern and southern populations (Fig. 2b).

The turnover functions make an implicit suggestion—if southern populations are located where BIO9 temperatures increase above ~14 °C in the future, then the frequencies of the adaptive alleles are expected to increase over time. In fact, if these SNPs do underlie local adaptation, then the adaptive alleles need to increase in frequency for southern populations to survive. The expected difference between allele frequencies in the present and in the future, on the basis of all relevant bioclimatic variables, is the local genetic offset[13,50]. Populations with higher genetic offsets are expected to be more vulnerable to climate change because they must undergo more dramatic changes in adaptive allele frequencies to adapt.

Genetic offsets are calculated across geographical space on the basis of the output of GF models. GF reports the cumulative importance of each bioclimatic variable across the landscape. For a given location, the local genetic offset is calculated between the present observed climate and a predicted future climate, which is a forecast from a global circulation model, by taking the Euclidean distance between the vectors of cumulative importance (Markdown 3.3, Supplementary Information). The end result is the identification of geographical regions in which the genetic space is predicted to be disrupted as a function of climate change. To illustrate this concept, we projected the genetic offset for *mexicana* populations using 2070 climate predictions (Fig. 2c), showing that the southern populations have higher average genetic offsets compared with the northern populations (Fig. 2d). These analyses suggest that northern populations will generally be less vulnerable to climate change than southern populations, an observation that could be crucial for conservation efforts.

Although recommended[23], it may be not necessary to identify candidate SNPs before GF analyses because one can infer turnover functions from an entire SNP dataset. However, if GF is applied to all SNPs (rather than to a subset of candidate SNPs), the inferred correlations may have little bearing on the process of local adaptation[51,52], because many SNPs are likely to covary in frequency among populations due to population structure. In this situation, genetic offsets will reflect expected shifts in neutral variation, confounding climate adaptation with the demographic history of populations. An alternative solution is to focus on a subset of SNPs based on an independent analysis[53] or to perform a correction for population structure before the GF analyses[49]. Whichever the case, it drives home a key point—some correction for population history is likely to be crucial.

All of these approaches are likely to have low statistical power to identify adaptive SNPs that covary with population structure[23]. In fact, one study of poplars has found that genetic offsets based on randomly selected SNPs predicted performance as well (or even slightly better) as sets of candidate SNPs[49], reflecting either limitations in the approach or that many adaptive SNPs covary with population structure in this system. Genetic offsets also omit biotic interactions, thereby probably underestimating the effect of climate change on the persistence of species[54]. Ultimately, more theoretical and experimental studies are needed to understand the utility and applicability of genetic offsets[24]. Nonetheless, incorporating adaptation into the framework of predicting species' fate under climate change is likely to provide important insights into causal relationships between genotypes and climate-related phenotypes and also

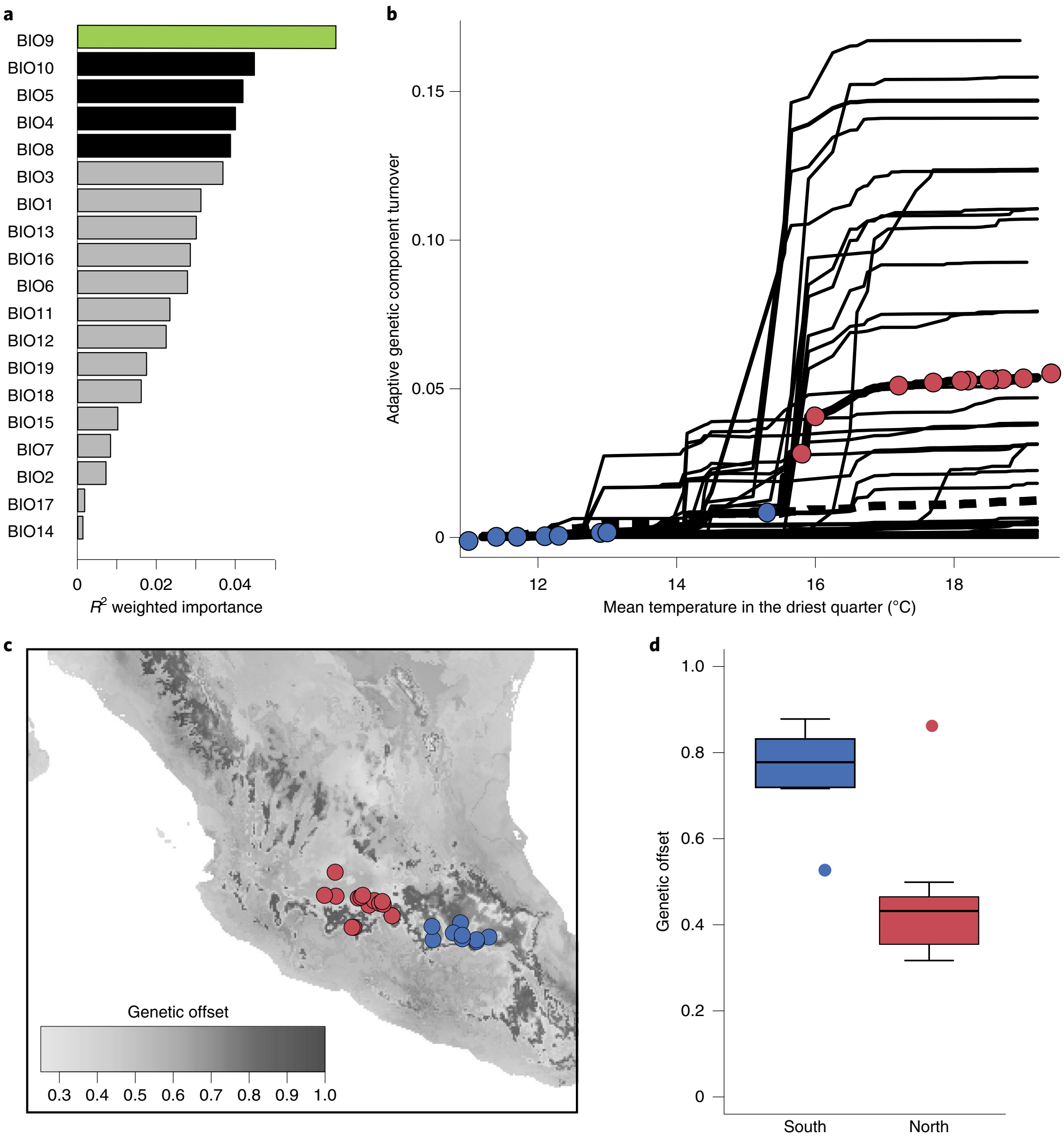

Fig. 2 | Turnover functions and genetic offsets. a, The inferred correlation between SNPs and bioclimatic variables on the basis of a GF analysis of 32 candidate SNPs. The bioclimatic variable in green (BIO9, the mean temperature in the driest quarter) has the strongest inferred correlation. The next four bioclimatic variables (BIO10, BIO5, BIO4 and BIO8), in rank of strongest contribution to the model, are represented in black. **b**, Allelic turnover functions relative to BIO9 (*x* axis), which is reported in °C. The *y* axis reflects the cumulative importance, which refers to the importance of SNPs in the GF models; it reflects the total amount of turnover in allele frequency across the temperature gradient[13]. The cumulative turnover for individual candidate SNPs is shown as thin grey lines. The solid black line indicates the turnover across all candidate SNPs; circles along this line represent populations arranged according to BIO9, with colours indicating their inclusion in the warm northern cluster (red) or the cold southern cluster (blue). The dashed line represents the average turnover across 500 reference SNPs that are not considered to be putatively adaptive. The superior performance of candidate SNPs is evidenced by the higher turnover values of the solid bold line (candidate SNPs) compared with the dotted line (reference SNPs). **c**, Genetic offset estimated across the landscape, based on the 32 candidate SNPs. Darker hues indicate higher genetic offset (higher expected vulnerability to climate change). The *y* and *x* axes represent latitude and longitude, respectively. The circles represent individual populations, with colours indicating their inclusion in the cold southern cluster (blue) or the warmer northern cluster (red). **d**, Estimated genetic offsets combined across all populations within the colder southern cluster (blue; $n=12$) and warmer northern cluster (red; $n=11$). The predicted genetic offset is higher, on average (two-sided *t*-test, $t=9.87$; d.f. = 17.53, $P<0.001$), for populations from the south.

to aid management decisions based on putatively adaptive variants within populations.

Additional evolutionary processes

To date, local adaptation has been integrated into ~20 climate vulnerability studies across a broad taxonomic array of organisms[23]. These studies generally omit other evolutionary processes that will impact the fate of populations under climate change. The importance of including more evolutionary processes such as gene flow, population movement and genetic load has been mentioned in various reviews and papers[18,21,23], but they have not been yet integrated into a comprehensive strategy. These processes are likely to affect the pace of species' adaptation and colonization under climate change[18], but not always positively. Below we discuss the possible impacts of these processes and review ways to integrate them into the framework of landscape genomics and predicting species' responses to climate change.

Gene flow. Gene flow refers to the genetic exchange between existing populations. A fascinating feature of this exchange is that it can either speed the process of adaptation, by introducing favourable alleles from one population to another[55–63], or slow adaptation by contributing non-adaptive alleles from one population to another[64–68]. Adaptive gene flow can have major effects on the geographical range of a species. One interesting example of adaptive gene flow is from *mexicana* into maize (*Zea mays* ssp. *mays*), which helped to expand high-altitude cultivation of the crop[57]. Remarkably, gene flow in the reverse direction has also been adaptive, because gene flow from maize into *mexicana* has enabled the latter to expand its geographical range as an introduced weed in Europe[62].

Maladaptation has been studied less than adaptation, but maladaptive alleles can also be introduced into a population. This occurs when there is gene flow between populations that are adapted to different conditions; an allele from a population in one environment will lead to hybrid and introgressed individuals with reduced fitness in the second environment[65,66,68–71]. Another way to state the point is to say that maladaptive gene flow reverses the divergence of populations caused by local adaptation[72,73]. Maladaptive gene flow may be particularly important when there are populations adapted to the edges of a species' environmental niche. If there are larger, older populations in the centre of the niche, then gene flow may predominate from the centre, potentially swamping locally adapted alleles in the 'edge' populations[65,66,74]. Although it has not been studied extensively, there are interesting examples of maladaptive gene flow. One comes from a study of two morphs of walking stick insects (*Timema cristinae*), where gene flow between the two morphs increased rates of predation on both[75].

The important point is that gene flow should be considered when trying to predict the fate of populations because it may either accelerate or decelerate the pace of adaptation. But how should one proceed? Fortunately, population genetics has several existing useful tools to evaluate gene flow. We illustrate one analysis here based on coalescent simulations[76,77] (Markdown 4, Supplementary Information). The simulations are fit to features of the observed data and output estimates of effective population size (N_e), the migration rate per generation (m) between populations and their product N_em (Fig. 3a). Under equilibrium conditions, N_em reflects the average number of individuals that move between populations each generation; a value of $N_em > 1$ is sufficient to homogenize populations over the long term assuming no countervailing evolutionary forces[78]. A key feature of N_em estimates based on this method is that they are directional, that is, gene flow to and from a population can be estimated separately. Other common methods to estimate N_em—such as traditional F_{ST} statistics[79,80]—do not provide directional insights.

For the purposes of illustration, we implemented this approach on two *mexicana* populations—one southern (population South 2) and one northern (population North 4). The goal was to assess whether these focal populations receive an influx of warmer-adapted alleles from northern populations or are inundated with maladaptive alleles from the south. We estimated migration into South 2 and North 4 from every other population in the dataset. The N_em estimates indicate that most populations contribute alleles to North 4 through gene flow (Fig. 3b). However, northern populations contribute to North 4 at much higher rates than southern populations, suggesting that North 4 is unlikely to be swamped by maladaptive gene flow from the south.

By contrast, the South 2 population receives more gene flow from populations that are adapted to colder climates (Fig. 3c), but it also receives substantive levels of gene flow from several northern populations at a N_em of ~1. The difficult question is whether putatively adaptive gene flow from the north will overcome more frequent gene flow from southern populations. Here the selection coefficient (s) of adaptive alleles is key, because maladaptive gene flow is expected to swamp the population if the benefit of adaptive alleles is less than roughly half the rate of maladaptive migration (that is, $s < \sim 0.5m$)[81,82]. Because the s of adaptive alleles is not known, except in rare exceptions where they have been determined experimentally[15], it is difficult to predict the outcome of competing gene flow into a population like South 2. It is also difficult to know whether additional, unsampled populations could contribute to South 2 and whether N_em estimates, which integrate over genetic history, will be accurate in the future. Finally, at least one study has shown that maladaptive and adaptive gene flow can be 'self-cancelling' and may therefore not impede divergence among populations due to local adaptation[83]. We nonetheless argue that gene flow is important for interpreting population vulnerability because it can have implications both for a population's response to climate change and for potential management strategies—for example, introducing pre-adaptive genetic diversity into vulnerable populations[55,84].

Population dispersal. Gene flow refers to the exchange of genes between existing populations, but some populations may need to move (or disperse) to a new location to survive[2,3,85]. As concrete examples, 102 montane moth species have increased their average altitude by 67 m along a transect on Mount Kinabalu in Borneo over a span of 42 years[86]. Similarly, more than 2,000 species of animals and plants have moved to either higher elevations or more extreme latitudes in response to climate change[85].

Dispersal is a potential mechanism for the persistence of species, but few studies have used landscape genomic data to examine this phenomenon in the context of climate change predictions. One example focuses on pearl millet in Africa; Rhoné et al.[53] identified landraces with high predicted genomic offsets as potential risks for crop failure. They then searched for landraces that have low predicted genetic offsets in the location of predicted failure. The most likely replacement landraces were geographically distant from the location of crop failure—that is, >1,000 km away on average. This study highlights the need for human-mediated, long-distance dispersal of landraces to avoid climate-related failures of the pearl millet crop. Gougherty et al.[50] also used genetic offsets to study the dispersal potential of wild poplar populations. They predicted that populations at the centre of the species' distribution will need to move shorter distances to maintain low genetic offsets in the future, as compared with populations at the edge of the distribution.

Both of these studies identified putative locations for movement on the basis of genetic offsets, but they did not estimate the cost of dispersal across the landscape. In the field of landscape genomics, the calculation of costs has usually relied on circuit theory to calculate the 'resistance' or 'current' of potential movement and to estimate least-cost paths between locations[87]. For example, Razgour et al.[35] measured the migration potential of two bat species between two distinct locations. By focusing on features that correlate with the current geographical distribution of species—specifically forest cover and slope—they assessed whether an endangered population could be replaced by the migration of an existing, less-endangered population. They concluded that landscape connectivity was not sufficient for an existing warm-adapted population to replace an endangered population. Similarly, Aguirre-Liguori et al.[88] used landscape resistance to predict potential areas of dispersion for wild *Zea* populations, based on present and future SDMs.

Circuit theory has been used in the field of landscape genetics to estimate the cost of migration across specific distances and routes. However, cost has been based on outputs from SDMs or from environmental data given physiological knowledge (for example,

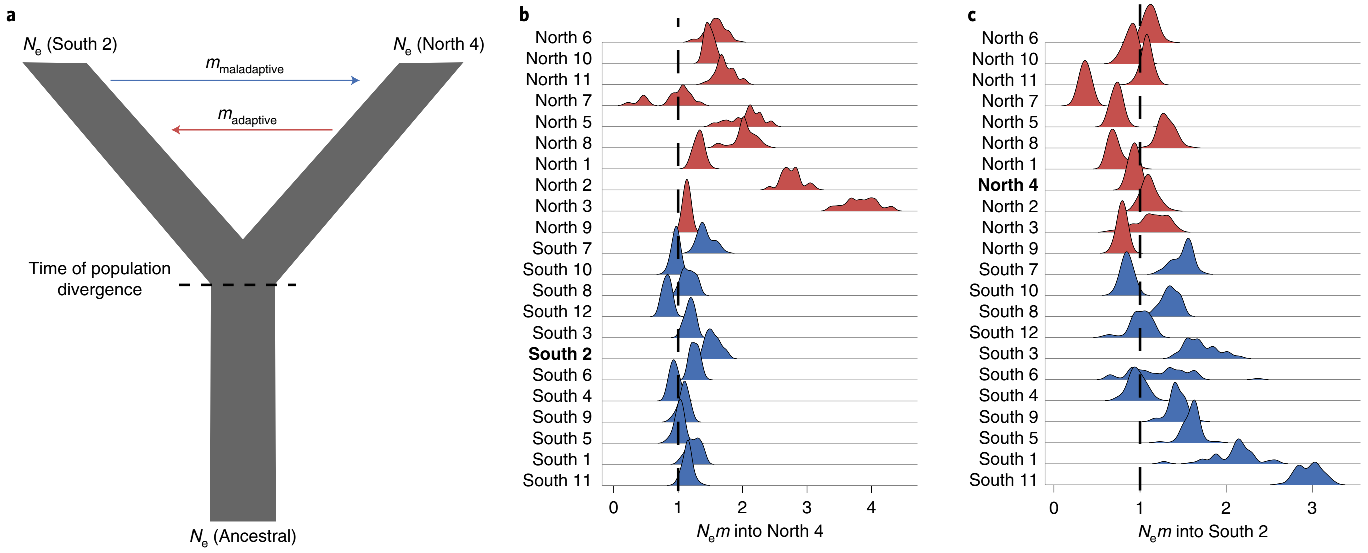

Fig. 3 | Gene flow into focal populations. a, The model used to compare gene flow between a pair of populations, using the two focal populations (South 2 and North 4) as an example. The model includes parameters for the population size (N) of each population, directional migration rates (m) between populations and a divergence time. We have labelled migration as adaptive and maladaptive because generally gene flow from populations in the cold southern cluster with high genetic offsets will be maladaptive relative to gene flow from populations in the warm northern cluster. **b**, Estimates of gene flow (N_em) into population North 4 from other sampled populations. Red distributions represent estimates from populations in the warm northern cluster, whereas blue distributions represent estimates from populations in the cold southern cluster. **c**, Summary of gene flow into population South 2, as described in **b**. In **b** and **c**, the dashed line represents $N_em = 1$, above which gene flow can homogenize allele frequencies between populations.

thermal niche thresholds), without utilizing landscape genomic data. One approach is to identify potentially habitable future locations for an existing population, based on genetic offsets[50,53], and then to evaluate the feasibility of potential migration routes to those habitable locations. The latter step requires a matrix that estimates the cost of movement from an existing location to all other potentially habitable locations in the future. Here we illustrate this approach by using turnover functions to determine whether movement from one location to another entails high 'landscape resistance', as reflected by high genetic offsets, or low resistance, which represents a corridor for potential population movement. In this framework, the total landscape resistance between a current location and a future settlement site is therefore a summation of the offsets across the landscape paths between the locations. A nice feature of this approach is that GF is used to weight bioclimatic variables to genetic patterns and is then employed to estimate the dispersal costs. Depending on the goals of the analysis and the weights, resistance between geographical sites can be based on all bioclimatic variables, on only uncorrelated variables or on a single variable (for example, BIO9). In Fig. 4a, we show the estimation of landscape resistance for the South 2 population of *mexicana* (Markdown 5, Supplementary Information). The resulting map suggests, as expected, that distant dispersal is less probable than nearby movement. It also suggests possible dispersal routes to the east and west but not to the north and south.

This approach also permits comparisons among populations. For example, we compared the number of habitable regions within 20 km of each sampled *mexicana* population, suggesting that some populations (for example, North 9 and South 7) have few available dispersal routes (Fig. 4b). Nonetheless, this approach based on genetic offsets cannot define the limit to dispersal—that is, when are offsets, resistances and geographical distances simply too much for a population to disperse? This question cannot be answered without detailed information about the physiology, phenotypic plasticity, environmental tolerances and movement capabilities of a specific study organism[16]. This approach also does not include dispersal by wind, water or biological vectors. Their effects can be substantive; for example, a recent meta-analysis found that stronger wind patterns between populations are negatively correlated with genetic differentiation[89]. One potential solution to the problem of wind and water dispersal is to estimate predicted settlement areas on the basis of genetic offsets and then use wind and water current models to identify the most probable routes of dispersal.

Genomic load. Finally, we discuss genomic load, which refers to the presence of genetic variants that reduce the fitness of a population relative to a local fitness optimum. These deleterious genetic variants are targeted by purifying selection. However, forces such as genetic drift and linked selection limit the efficacy of purifying selection, so that deleterious mutations can increase in frequency. Drift is higher in small populations, and—although the relationship among load, population size and history can be complex[90,91]—the reduced efficacy of selection can accelerate rates of population extinction[92]. It is therefore reasonable to surmise, as a first approximation, that the potential for a population to respond to climate change is inversely related to its load[21].

Load is difficulty to study, because its estimation requires knowledge about the fitness effects of specific variants. Lacking such information, empirical studies have focused on measuring genetic variants that are predicted to have deleterious effects, including variants that alter amino acids (particularly with non-conserved biochemical properties or that alter protein structure), introduce premature stop mutations or modify sequence motifs that have been otherwise highly conserved over evolutionary time[93–95]. Studies have shown that these measures can correlate with phenotype[96] and also that deleterious variants may increase in specific populations—for example, some inbreeding populations[92] and species with histories of genetic bottlenecks[90,91]. There has also been some focus on populations that are on the edges of a species' geographical distribution[97–99] or that expand a species' range[100,101]. In both cases, load tends to be higher than for populations near to the centre of the species' range, with evidence for associated fitness decreases in edge populations[102].

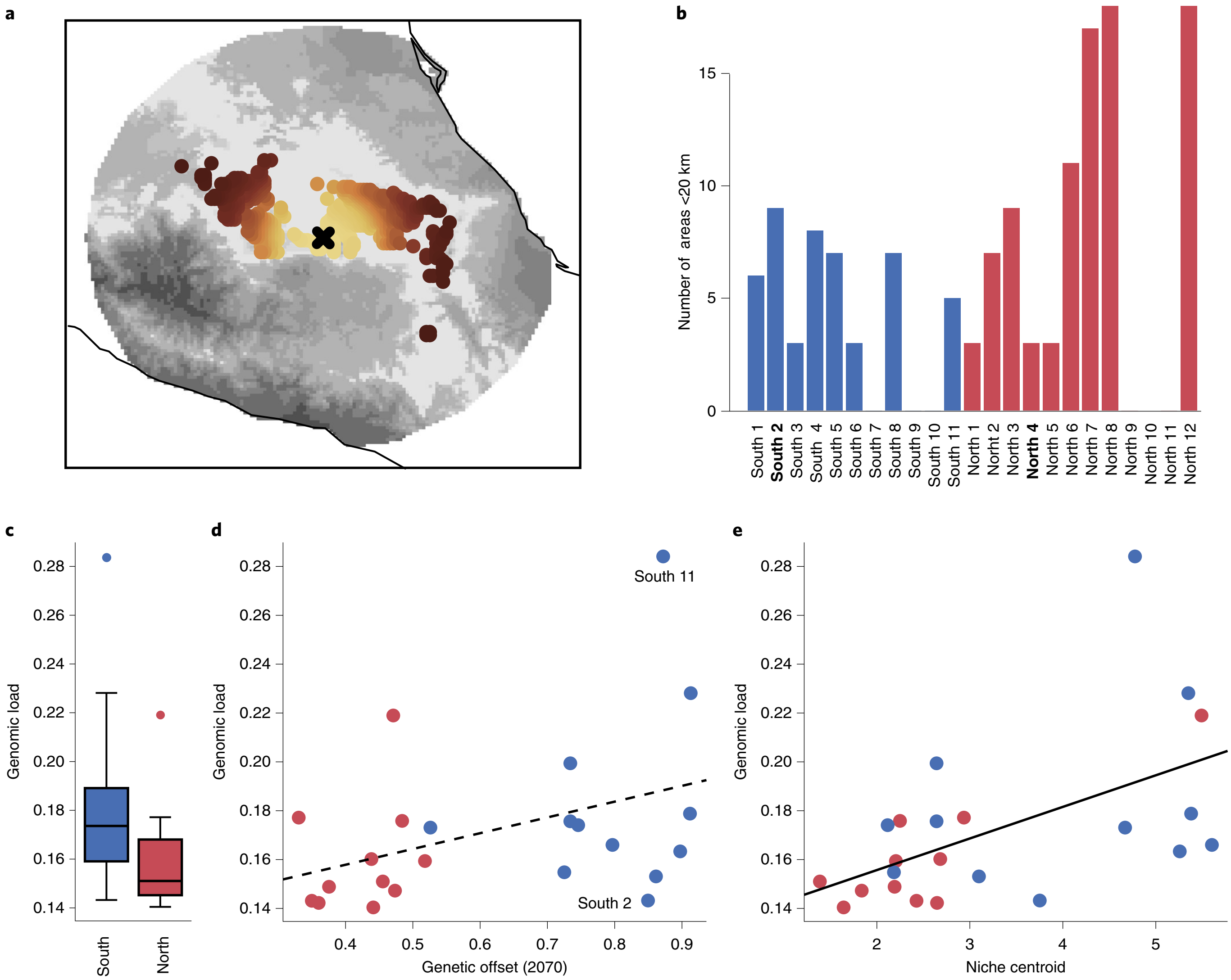

Fig. 4 | Predicting potential areas of future dispersal. a, The map shows the focal population (South 2) as a cross (×). Possible dispersal paths to new locations are coloured, with lower costs in yellow and higher costs in darker and warmer colours. The grey background indicates the predicted genetic offsets for future climate forecasts, with darker colours indicating higher offsets. **b**, The number of potential areas (grid cells in the map) that can be colonized within a 20 km radius for each population. Populations with fewer predicted settlement areas have fewer dispersal options. The two focal populations from Fig. 3 (South 2 and North 4) are indicated in bold. **c**, The estimated genomic load, as measured for populations in the cold southern cluster (blue) and warm northern cluster (red). The difference between clusters is not significant (two-sided t-test, $t=1.69$; d.f. = 18.15, $P=0.1$). **d**, The relationship between the estimated genomic load (y axis) and genetic offset (x axis) measured between the present and future climate forecasts ($r^2=0.121$; $P=0.057$; $F=4.03$ with 1 and 21 d.f.). The individual circles represent populations in the cold southern cluster (blue) and warm northern cluster (red); the South 2 and South 11 populations have similar offsets but very different genomic loads. **e**, The estimated genomic load plotted for each population as a function of the distance from the niche centroid of either the cold southern or the warm northern clusters. The correlation is significantly positive ($r^2=0.295$; $P=0.007$; $F=8.79$ with 1 and 21 d.f.).

Given the relevance of load and the fact that it can be approximated from genomic data, it is puzzling that it has not been incorporated into genetic predictions of the effects of climate change[67]. However, the concept has not been fully ignored as genetic offsets have been discussed as a measure of maladaptation[23,50]—that is, populations with high offsets will be maladapted in the future if they do not have time to adapt. This interpretation of maladaptation touches on genomic load but is incomplete. In theory, two populations can have identical genetic offsets based on adaptive SNPs but markedly different loads depending on the distribution of deleterious SNPs (and structural variants[103]) throughout the remainder of the genome. To illustrate this concept empirically, we estimated the load for *mexicana* populations (Markdown 6.1, Supplementary Information) using a measure that compares the proportion and population frequency of non-synonymous SNPs to synonymous SNPs[104]. We found that the southern populations have higher estimated loads, on average, compared with northern populations (Fig. 4c). Moreover, two southern populations (Fig. 4d) have similar genetic offsets but markedly different estimated loads. Assuming that load is correlated inversely with fitness, one expects that the population with higher load is vulnerable to climate change.

This approach considers only genomic load in the present, without reference to how it may change in relation to future climate. The challenge is that load is not deterministic and is therefore difficult to predict. However, if load is associated with another variable that is predictable, then it may be possible to roughly assess future trends.

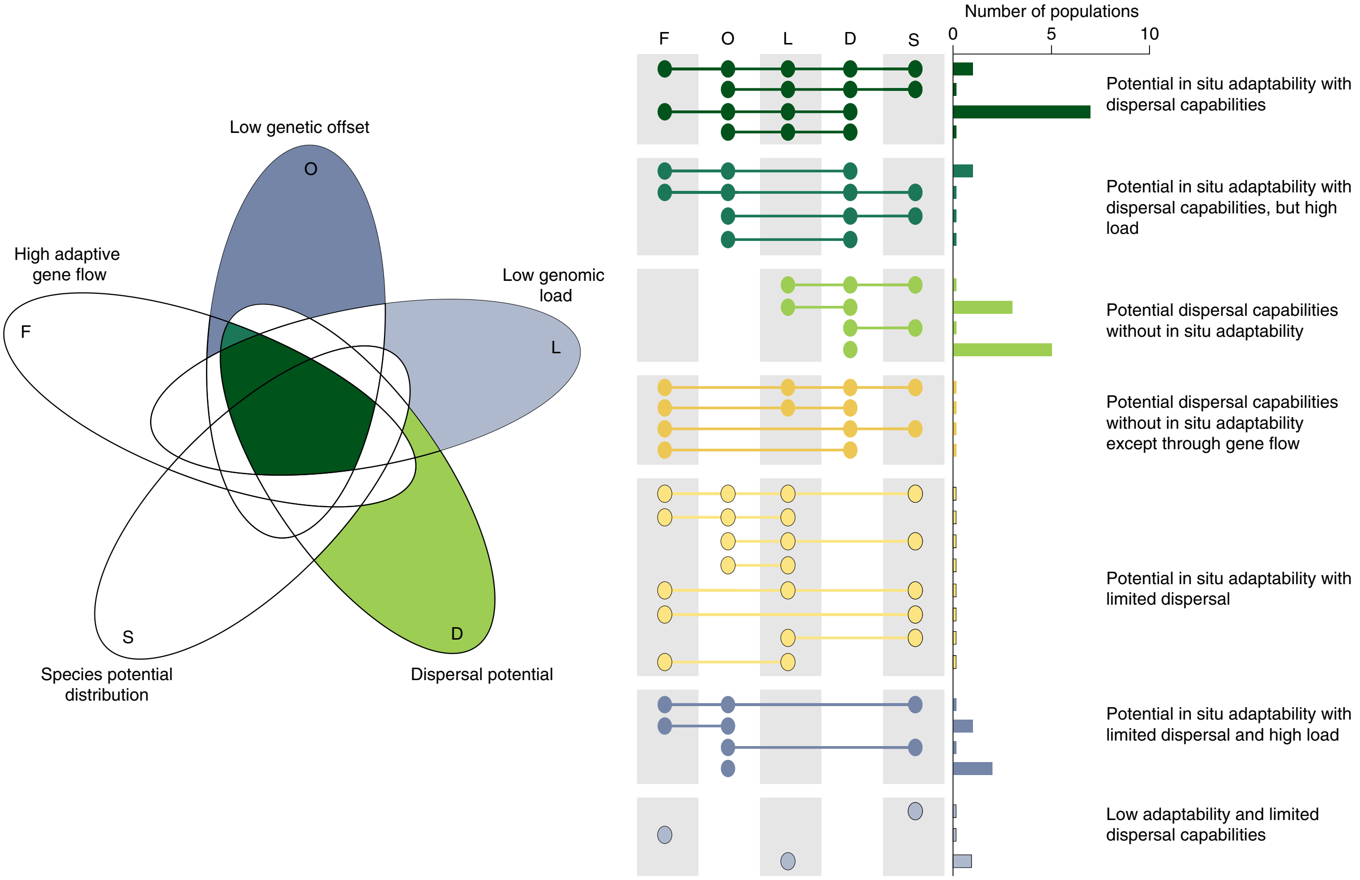

Fig. 5 | The FOLDS integrated framework. The Venn diagram represents an extension of the BAM model in Fig. 1a as a conceptual framework. The components of the Venn diagram represent the five types of analysis discussed in the text. The purpose of the diagram is to integrate information and to identify populations that are at risk due to specific factors—for example, populations that lack dispersal opportunities but have favourable gene flow, offsets, load, dispersal and SDMs. The colours in the Venn diagram map the 23 *mexicana* populations based on the following thresholds: more estimated gene flow from warm-adapted versus cold-adapted populations, genetic offsets of <50%, load in the lower half of populations, dispersal potential within 20 km and presence in future SDMs. In addition to the Venn diagram, we visualized the intersection of sets and their size using an UpSetR plot[109], for which the connected coloured dots indicate multiple overlapping features. The histogram shows the corresponding number of *mexicana* populations that have specific combinations of features. The text on the right provides verbal interpretation of sets of feature combinations.

The central abundance hypothesis (CAH)[74] is useful in this context because it predicts that environmental conditions are less ideal at the edges of the geographical distribution of a species and also that populations at edges have lower effective sizes and higher load than populations in the centre[65]. Although the CAH does not apply universally[70], it does hold in some cases. For example, genomic load is higher at the geographical limits of *Arabidopsis lyrata*[97], suggesting that founder events at the geographical limits, along with gene surfing[105], have caused deleterious mutations to accumulate on the expanding range[21]. To illustrate this concept, we have analysed how load for individual *mexicana* populations is correlated with the distance from the niche centroid[106], focusing on the northern and southern groups separately. For both groups, there is a positive and significant trend between load and the distance from the niche centroid (Fig. 4e).

When the CAH holds, as it seems to do for the northern and southern *mexicana* groups, the question of load can be simplified to the following: if climate changes in the future, is a population expected to be farther from the projected centroid? If so, it is expected to have increased genetic load. For example, most *mexicana* populations are expected to be farther from the niche centroid in the future than they are now. However, the effect is more pronounced for northern populations, suggesting that they are particularly likely to be subjected to evolutionary forces that produce genetic load (Markdown 6.2, Supplementary Information).

We recognize that this is, at best, a gross approximation of how load will change under future climates, but this discussion brings up one of the paradoxical features of populations and their predicted responses to climate change: the fate of small populations at the edge of species' geographical ranges[21,97]. These populations are expected to experience higher genetic drift (and therefore potential for higher genetic load). They may also be swamped by maladaptive gene flow from populations in the niche centroid and may be in locations nearer to extreme environments that limit their migration potential[65,106]. Yet, it is precisely these populations that are also evolving adaptively to new climatic factors, because they are often climatic outliers[38,55]. It is this give and take among adaptation and potentially contravening forces such as gene flow, dispersal and load that is not yet adequately integrated into the framework of landscape genomics and predictions about population vulnerability under climate change.

An integrated framework

The synthesis of genetic adaptation with climate change models is a remarkable advance, because it explicitly recognizes that populations are not homogenous across the landscape[13,14,107,108]. However, current approaches do not yet capture most of the complexities of the evolutionary process, some of which may be as important as adaptation[18]. We argue that additional processes can and should be considered to predict the fate of populations and species in the face of a changing environment. To support that argument, we reviewed three such processes—gene flow, dispersal and genetic load—in the context of climate change.

Of course, it is important to bring separate analyses together into an integrated whole. An integrated framework should synthesize levels of risk based on factors such as SDMs, genetic variation and landscape resistance[35]. But how can one perform this integration? One simple approach is a conceptual extension of the BAM model, in which the components of the Venn diagram represent manifold outcomes from different analyses. To demonstrate, we introduce the FOLDS model, which is named after the five features discussed in this Perspective: gene flow, genetic offsets, genetic load, dispersal and SDMs (Fig. 5). By applying empirical thresholds to all five features (Fig. 5), we have assessed the empirical distribution of populations in the FOLDS model. Of 23 populations, only one passed all five empirical thresholds, suggesting that it is the most likely to survive climate change without human intervention (Fig. 5). However, 7 out of 23 populations had low genetic offsets and high dispersal capabilities, despite being located in regions that are not predicted as habitable in the future by SDMs (Fig. 5; Markdown 7, Supplementary Information). We predict that these populations are likely to survive climate change, based on the weight of evolutionary evidence.

A useful feature of the FOLDS model is its flexibility; theoretically, it can synthesize an unlimited number of components, each component can be weighted on the basis of project objectives and thresholds can be varied to assess effects on the robustness of conclusions. Although imperfect, FOLDS may help to fuel this rapidly growing field, which needs to continue to expand in at least three directions. The first is experimental validations of predicted genetic offsets[23,24,49,51], which are beginning to appear[49,53] but still nascent. A second is further consideration of evolutionary processes, along with more complete measurement of genetic variants. An example of the latter is structural variants, which have not been integrated systematically into landscape genomics but are common, often unlinked to SNPs and clearly affect phenotypes[91,103]. Finally, we are intrigued about including additional features into FOLDS, including human impacts such as urbanization, susceptibility to disturbance (for example, fire or flood) and biotic interactions. It will be especially exciting to integrate landscape genomic data among interacting set of species—for example, pathogens, predators, vectors and hosts—or perhaps complete community assemblages.

Data availability

The exemplar *mexicana* data used in all analyses are available at Zenodo (https://doi.org/10.5281/zenodo.4746517).

Code availability

The Markdown file is available as Supplementary Information. All R code is also available at Zenodo (https://doi.org/10.5281/zenodo.4746517).

Received: 5 February 2021; Accepted: 5 July 2021;

References

1. Foden, W. B. et al. Climate change vulnerability assessment of species. *Wiley Interdiscip. Rev. Clim. Change* **10**, e551 (2019).
2. Parmesan, C. & Yohe, G. A globally coherent fingerprint of climate change. *Nature* **421**, 37–42 (2003).
3. Parmesan, C. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* **37**, 637–669 (2006).
4. Walther, G.-R. et al. Ecological responses to recent climate change. *Nature* **416**, 389–395 (2002).
5. Feeley, K. J., Bravo-Avila, C., Fadrique, B., Perez, T. M. & Zuleta, D. Climate-driven changes in the composition of New World plant communities. *Nat. Clim. Change* **10**, 965–970 (2020).
6. Román-Palacios, C. & Wiens, J. J. Recent responses to climate change reveal the drivers of species extinction and survival. *Proc. Natl Acad. Sci. USA* **117**, 4211–4217 (2020).
7. Dyderski, M. K., Paź, S., Frelich, L. E. & Jagodziński, A. M. How much does climate change threaten European forest tree species distributions? *Glob. Change Biol.* **24**, 1150–1163 (2018).
8. Peterson, A. T. et al. *Ecological Niches and Geographic Distributions* (Princeton Univ. Press, 2011).
9. Thuiller, W., Guéguen, M., Renaud, J., Karger, D. N. & Zimmermann, N. E. Uncertainty in ensembles of global biodiversity scenarios. *Nat. Commun.* **10**, 1446 (2019).
10. Fourcade, Y., Besnard, A. G. & Secondi, J. Evaluating interspecific niche overlaps in environmental and geographic spaces to assess the value of umbrella species. *J. Avian Biol.* **48**, 1563–1574 (2017).
11. Feeley, K. J., Rehm, E. M. & Machovina, B. perspective: The responses of tropical forest species to global climate change: acclimate, adapt, migrate, or go extinct? *Front. Biogeogr.* **4**, 69–84 (2012).
12. Razgour, O. et al. An integrated framework to identify wildlife populations under threat from climate change. *Mol. Ecol. Resour.* **18**, 18–31 (2018).
13. Fitzpatrick, M. C. & Keller, S. R. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecol. Lett.* **18**, 1–16 (2015).
14. Exposito-Alonso, M. et al. Genomic basis and evolutionary potential for extreme drought adaptation in *Arabidopsis thaliana*. *Nat. Ecol. Evol.* **2**, 352–358 (2018).
15. Exposito-Alonso, M., Burbano, H. A., Bossdorf, O., Nielsen, R. & Weigel, D. Natural selection on the *Arabidopsis thaliana* genome in present and future climates. *Nature* **573**, 126–129 (2019).
16. Reside, A. E., Butt, N. & Adams, V. M. Adapting systematic conservation planning for climate change. *Biodivers. Conserv.* **27**, 1–29 (2018).
17. Brown, J. L. et al. Predicting the genetic consequences of future climate change: the power of coupling spatial demography, the coalescent, and historical landscape changes. *Am. J. Bot.* **103**, 153–163 (2016).
18. Waldvogel, A. et al. Evolutionary genomics can improve prediction of species' responses to climate change. *Evoution Lett.* **4**, 4–18 (2019).
19. Allendorf, F. W., Hohenlohe, P. A. & Luikart, G. Genomics and the future of conservation genetics. *Nat. Rev. Genet.* **11**, 697–709 (2010).
20. Barbosa, S. et al. Integrative approaches to guide conservation decisions: using genomics to define conservation units and functional corridors. *Mol. Ecol.* **27**, 3452–3465 (2018).
21. Nadeau, C. P. & Urban, M. C. Eco-evolution on the edge during climate change. *Ecography* **42**, 1280–1297 (2019).
22. Hällfors, M. H. et al. Addressing potential local adaptation in species distribution models: Implications for conservation under climate change. *Ecol. Appl.* **26**, 1154–1169 (2016).
23. Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M. & Keller, S. R. Genomic prediction of (mal)adaptation across current and future climatic landscapes. *Annu. Rev. Ecol. Evol. Syst.* **51**, 245–269 (2020).
24. Rellstab, C., Dauphin, B. & Exposito-Alonso, M. Prospects and limitations of genomic offset in conservation management. *Evol. Appl.* **14**, 1202–1212 (2021).
25. Soberón, J. M. Niche and area of distribution modeling: a population ecology perspective. *Ecography* **33**, 159–167 (2010).
26. Araújo, M. B. & Peterson, A. T. Uses and misuses of bioclimatic envelope modeling. *Ecology* **93**, 1527–1539 (2012).
27. Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A. & Snyder, M. A. Niches, models, and climate change: assessing the assumptions and uncertainties. *Proc. Natl Acad. Sci. USA* **106**, 19729–19736 (2009).
28. Collart, F., Hedenäs, L., Broennimann, O., Guisan, A. & Vanderpoorten, A. Intraspecific differentiation: Implications for niche and distribution modelling. *J. Biogeogr.* **48**, 415–426 (2020).
29. Benito Garzón, M., Robson, T. M. & Hampe, A. ΔTraitSDMs: species distribution models that account for local adaptation and phenotypic plasticity. *N. Phytol.* **222**, 1757–1765 (2019).
30. Frichot, E., Schoville, S. D., Bouchard, G. & François, O. Testing for association between loci and environmental gradients using latent factor mixed models. *Mol. Biol. Evol.* **30**, 1687–1699 (2013).
31. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
32. Alexander, D. H. & Lange, K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinform.* **12**, 246 (2011).
33. Ikeda, D. H. et al. Genetically informed ecological niche models improve climate change predictions. *Glob. Change Biol.* **23**, 164–176 (2017).
34. Jay, F. et al. Forecasting changes in population genetic structure of alpine plants in response to global warming. *Mol. Ecol.* **21**, 2354–2368 (2012).

35. Razgour, O. et al. Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proc. Natl Acad. Sci. USA* **116**, 10418–10423 (2019).
36. Gotelli, N. J. & Stanton-Geddes, J. Climate change, genetic markers and species distribution modelling. *J. Biogeogr.* **42**, 1577–1585 (2015).
37. Pyhäjärvi, T., Hufford, M. B., Mezmouk, S. & Ross-Ibarra, J. Complex patterns of local adaptation in teosinte. *Genome Biol. Evol.* **5**, 1594–1609 (2013).
38. Aguirre-Liguori, J. A. et al. Connecting genomic patterns of local adaptation and niche suitability in teosintes. *Mol. Ecol.* **26**, 4226–4240 (2017).
39. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
40. Kawecki, T. J. & Ebert, D. Conceptual issues in local adaptation. *Ecol. Lett.* **7**, 1225–1241 (2004).
41. Leimu, R. & Fischer, M. A meta-analysis of local adaptation in plants. *PLoS ONE* **3**, e4010 (2008).
42. de Villemereuil, P. & Gaggiotti, O. E. A new FST-based method to uncover local adaptation using environmental variables. *Methods Ecol. Evol.* **6**, 1248–1258 (2015).
43. Coop, G. M., Witonsky, D., Di Rienzo, A. & Pritchard, J. K. Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185**, 1411–1423 (2010).
44. Gautier, M. Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics* **201**, 1555–1579 (2015).
45. De Mita, S. et al. Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Mol. Ecol.* **22**, 1383–1399 (2013).
46. Schoville, S. D. et al. Adaptive genetic variation on the landscape: methods and Cases. *Annu. Rev. Ecol. Evol. Syst.* **43**, 23–43 (2012).
47. Tiffin, P. & Ross-Ibarra, J. Advances and limits of using population genetics to understand local adaptation. *Trends Ecol. Evol.* **29**, 673–680 (2014).
48. Forester, B. R., Lasky, J. R., Wagner, H. H. & Urban, D. L. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Mol. Ecol.* **27**, 2215–2233 (2018).
49. Fitzpatrick, M. C., Chhatre, V. E., Soolanayakanahally, R. Y. & Keller, S. R. Experimental support for genomic prediction of climate maladaptation using the machine learning approach Gradient Forests. *Mol. Ecol. Resour.* https://doi.org/10.1111/1755-0998.13374 (2021).
50. Gougherty, A. V., Keller, S. R. & Fitzpatrick, M. C. Maladaptation, migration and extirpation fuel climate change risk in a forest tree species. *Nat. Clim. Change* **11**, 166–171 (2021).
51. Fitzpatrick, M. C., Keller, S. R. & Lotterhos, K. E. Comment on 'Genomic signals of selection predict climate-driven population declines in a migratory bird'. *Science* **361**, eaat7279 (2018).
52. Booker, T. R., Yeaman, S. & Whitlock, M. C. Variation in recombination rate affects detection of outliers in genome scans under neutrality. *Mol. Ecol.* **29**, 4274–4279 (2020).
53. Rhoné, B. et al. Pearl millet genomic vulnerability to climate change in West Africa highlights the need for regional collaboration. *Nat. Commun.* **11**, 5274 (2020).
54. Bascompte, J., García, M. B., Ortega, R., Rezende, E. L. & Pironon, S. Mutualistic interactions reshuffle the effects of climate change on plants across the tree of life. *Sci. Adv.* **5**, eaav2539 (2019).
55. Hampe, A. & Petit, R. J. Conserving biodiversity under climate change: the rear edge matters. *Ecol. Lett.* **8**, 461–467 (2005).
56. Sexton, J. P., Strauss, S. Y. & Rice, K. J. Gene flow increases fitness at the warm edge of a species' range. *Proc. Natl Acad. Sci. USA* **108**, 11704–11709 (2011).
57. Hufford, M. B. et al. The genomic signature of crop-wild introgression in maize. *PLoS Genet.* **9**, e1003477 (2013).
58. Pease, J. B., Haak, D. C., Hahn, M. W. & Moyle, L. C. Phylogenomics reveals three sources of adaptive variation during a rapid radiation. *PLoS Biol.* **14**, e1002379 (2016).
59. Figueiró, H. V. et al. Genome-wide signatures of complex introgression and adaptive evolution in the big cats. *Sci. Adv.* **3**, e1700299 (2017).
60. Bontrager, M. & Angert, A. L. Gene flow improves fitness at a range edge under climate change. *Evol. Lett.* **3**, 55–68 (2019).
61. Todesco, M. et al. Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature* **584**, 602–607 (2020).
62. Le Corre, V., Siol, M., Vigouroux, Y., Tenaillon, M. I. & Délye, C. Adaptive introgression from maize has facilitated the establishment of teosinte as a noxious weed in Europe. *Proc. Natl Acad. Sci. USA* **117**, 25618–25627 (2020).
63. Oziolor, E. et al. Adaptive introgression enables evolutionary rescue from extreme environmental pollution. *Science* **364**, 455–457 (2019).
64. Bolnick, D. I. & Nosil, P. Natural selection in populations subject to a migration load. *Evolution* **61**, 2229–2243 (2007).
65. Eckert, C. G., Samis, K. E. & Lougheed, S. C. Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Mol. Ecol.* **17**, 1170–1188 (2008).
66. Sexton, J. P., McInyre, P. J., Angert, A. L. & Rice, K. J. Evolution and ecology of species range limits. *Annu. Rev. Ecol. Evol. Syst.* **40**, 415–436 (2009).
67. Brady, S. P. et al. Causes of maladaptation. *Evol. Appl.* **12**, 1229–1242 (2019).
68. Micheletti, S. J. & Storfer, A. Mixed support for gene flow as a constraint to local adaptation and contributor to the limited geographic range of an endemic salamander. *Mol. Ecol.* **29**, 4091–4101 (2020).
69. Sagarin, R. D. & Gaines, S. D. The 'abundant centre' distribution: to what extent is it a biogeographical rule? *Ecol. Lett.* **5**, 137–147 (2002).
70. Sagarin, R. D., Gaines, S. D. & Gaylord, B. Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends Ecol. Evol.* **21**, 524–530 (2006).
71. Fedorka, K. M., Winterhalter, W. E., Shaw, K. L., Brogan, W. R. & Mousseau, T. A. The role of gene flow asymmetry along an environmental gradient in constraining local adaptation and range expansion. *J. Evol. Biol.* **25**, 1676–1685 (2012).
72. Nosil, P., Harmon, L. J. & Seehausen, O. Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* **24**, 145–156 (2009).
73. Cenzer, M. L. Adaptation to an invasive host is driving the loss of a native ecotype. *Evolution* **70**, 2296–2307 (2016).
74. Hengeveld, R. & Haeck, J. The distribution of abundance. I. Measurements. *J. Biogeogr.* **9**, 303–316 (1982).
75. Farkas, T. E., Mononen, T., Comeault, A. A., Hanski, I. & Nosil, P. Evolution of camouflage drives rapid ecological change in an insect community. *Curr. Biol.* **23**, 1835–1843 (2013).
76. Excoffier, L. & Foll, M. fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics* **27**, 1332–1334 (2011).
77. Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C. & Foll, M. Robust demographic inference from genomic and SNP data. *PLoS Genet.* **9**, e1003905 (2013).
78. Wright, S. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In *Proc. Sixth International Congress of Genetics* Vol. 1, 356–366 (1932).
79. Weir, B. S. & Cockerham, C. C. Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
80. Nei, M. & Chesser, R. K. Estimation of fixation indices and gene diversities. *Ann. Hum. Genet.* **47**, 253–259 (1983).
81. Yeaman, S. & Otto, S. P. Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution* **65**, 2123–2129 (2011).
82. Feder, J. L., Flaxman, S. M., Egan, S. P., Comeault, A. A. & Nosil, P. Geographic mode of speciation and genomic divergence. *Annu. Rev. Ecol. Evol. Syst.* **44**, 73–97 (2013).
83. Endler, J. Gene Flow and population differentiation. *Science* **179**, 243–250 (1973).
84. Rehm, E. M., Olivas, P., Stroud, J. & Feeley, K. J. Losing your edge: climate change and the conservation value of range-edge populations. *Ecol. Evol.* **5**, 4315–4326 (2015).
85. Chen, I. C., Hill, J. K., Ohlemüller, R., Roy, D. B. & Thomas, C. D. Rapid range shifts of species associated with high levels of climate warming. *Science* **333**, 1024–1026 (2011).
86. Chen, I. C. et al. Elevation increases in moth assemblages over 42 years on a tropical mountain. *Proc. Natl Acad. Sci. USA* **106**, 1479–1483 (2009).
87. McRae, B. H., Dickson, B. G., Keitt, T. H. & Shah, V. B. Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology* **89**, 2712–2724 (2008).
88. Aguirre-Liguori, J. A., Ramírez-Barahona, S., Tiffin, P. & Eguiarte, L. E. Climate change is predicted to disrupt patterns of local adaptation in wild and cultivated maize. *Proc. R. Soc. B* **286**, 20190486 (2019).
89. Kling, M. M. & Ackerly, D. D. Global wind patterns shape genetic differentiation, asymmetric gene flow, and genetic diversity in trees. *Proc. Natl Acad. Sci. USA* **118**, e2017317118 (2021).
90. Henn, B. M. et al. Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. *Proc. Natl Acad. Sci. USA* **113**, 440–449 (2016).
91. Gaut, B. S., Seymour, D. K., Liu, Q. & Zhou, Y. Demography and its effects on genomic variation in crop domestication. *Nat. Plants* **4**, 512–520 (2018).
92. Frankham, R. Genetics and extinction. *Biol. Conserv.* **126**, 131–140 (2005).
93. Choi, Y., Sims, G., Murphy, S., Miller, J. & Chan, A. Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE* **7**, e46688 (2012).
94. Ng, P. C. & Henikoff, S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).
95. Davydov, E. et al. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput. Biol.* **6**, e1001025 (2010).
96. Yang, J. et al. Incomplete dominance of deleterious alleles contributes substantially to trait variation and heterosis in maize. *PLoS Genet.* **13**, e1007019 (2017).

97. Willi, Y., Fracassetti, M., Zoller, S. & Van Buskirk, J. Accumulation of mutational load at the edges of a species range. *Mol. Biol. Evol.* **35**, 781–791 (2018).
98. Koski, M. H., Layman, N. C., Prior, C. J., Busch, J. W. & Galloway, L. F. Selfing ability and drift load evolve with range expansion. *Evol. Lett.* **3**, 500–512 (2019).
99. Micheletti, S. J. & Storfer, A. A test of the central-marginal hypothesis using population genetics and ecological niche modelling in an endemic salamander (*Ambystoma barbouri*). *Mol. Ecol.* **24**, 967–979 (2015).
100. Peischl, S. & Excoffier, L. Expansion load: recessive mutations and the role of standing genetic variation. *Mol. Ecol.* **24**, 2084–2094 (2015).
101. Braasch, J. & Barker, B. S. Expansion history and environmental suitability shape effective population size in a plant invasion. *Mol. Ecol.* **28**, 2546–2558 (2019).
102. Perrier, A., Sánchez-Castro, D. & Willi, Y. Expressed mutational load increases toward the edge of a species' geographic range. *Evolution* **74**, 1711–1723 (2020).
103. Zhou, Y. et al. The population genetics of structural variants in grapevine domestication. *Nat. Plants* **5**, 965–979 (2019).
104. Peischl, S., Kirkpatrick, M. & Excoffier, L. Expansion load and the evolutionary dynamics of a species range. *Am. Nat.* **185**, 81–93 (2015).
105. Excoffier, L., Foll, M. & Petit, R. J. Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.* **40**, 481–501 (2009).
106. Lira-Noriega, A. & Manthey, J. D. Relationship of genetic diversity and niche centrality: a survey and analysis. *Evolution* **68**, 1082–1093 (2014).
107. Bay, R. A. et al. Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* **359**, 83–86 (2018).
108. Ruegg, K. et al. Ecological genomics predicts climate vulnerability in an endangered southwestern songbird. *Ecol. Lett.* **21**, 1085–1096 (2018).
109. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**, 2938–2940 (2017).

Acknowledgements

The study was supported by a UC-Mexus postdoctoral fellowship to J.A.A.-L., National Science Foundation grant no. 1741627 to B.S.G. and CONACyT Ciencia de Frontera 2019 grant no. 263962 to S.R.-B.

Author contributions

J.A.A.-L., S.R.-B. and B.S.G shaped ideas and content, discussed the results and wrote the manuscript. J.A.A.-L. wrote the code, and S.R.-B. and J.A.A.-L. constructed the Markdown file.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-021-01526-9.

Correspondence should be addressed to B.S.G.

Peer review information *Nature Ecology & Evolution* thanks Matthew Fitzpatrick, Ann-Marie Waldvogel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.