



Maladaptation, migration and extirpation fuel climate change risk in a forest tree species

Andrew V. Gougherty^{1,3}✉, Stephen R. Keller² and Matthew C. Fitzpatrick¹

Accounting for population-level adaptation and migration remains a central challenge to predicting climate change effects on biodiversity. Assessing how climate change could disrupt local climate adaptation, resulting in maladaptation and possibly extirpation, can inform where climate change poses the greatest risks across species ranges. For the forest tree species balsam poplar (*Populus balsamifera*), we used climate-associated genetic loci to predict population maladaptation with and without migration, the distance to sites that minimize maladaptation, and the emergence of novel genotype–climate associations. We show that the greatest disruptions to contemporary genotype–climate associations occur along the longitudinal edges of the range, where populations are predicted to be maladapted to all future North American climates, rescue via migration is most limited and novel genotype–climate associations emerge. Our work advances beyond species-level range modelling towards the long-held goal of simultaneously estimating the contributions of maladaptation and migration to understanding the risks that populations may face from shifting climates.

To avoid extirpation (local extinction) in rapidly changing environments, species must either migrate to other suitable locations or adapt or respond through phenotypic plasticity to new conditions in current habitats^{1,2}. However, most efforts to forecast biological responses to climate change do not fully consider the contributions of migration and adaptation at the population level and instead assume that all populations within a species respond similarly to environmental change. While some studies have integrated intraspecific variation with predictions of species range shifts (for instance, using species distribution models^{3–5}), these approaches rarely capture the continuous nature of population-level adaptive variation present across species ranges (but see refs. ^{6,7} for examples of other approaches of integrating intraspecific variation with climate change predictions). Assessing the need for populations to migrate and/or adapt in rapidly changing environments has the potential to provide richer insights into climate change vulnerability that may help inform where conservation efforts may be most effective.

Rapid climate change will pose numerous threats to climatically adapted populations. While spatial shifts in climate could require populations to migrate rapidly to track suitable habitats, the emergence of novel climates and the loss of existing climates may require some level of adaptation for populations to persist. For instance, populations adapted to disappearing climates (that is, contemporary combinations of temperature and precipitation with no future analogue^{8,9}) can avoid extirpation through plasticity or genetic rescue by gene flow or allele frequency change in situ, potentially resulting in novel genotype–climate associations (that is, associations with no near equivalent with standing patterns of genetic variation). Similarly, novel climates (that is, future combinations of temperature, precipitation and so on with no contemporary analogue⁹), in which existing genotypes may have had no recent selection history, could favour the evolution of novel genotype–climate associations through the recombination of existing climate-adaptive alleles. High climate novelty could also cause range contractions if future

climates are far outside the climatic tolerances of existing populations. While both disappearing and novel climates represent risks to climatically adapted populations in situ, migration has the potential to dampen these effects. If propagules or alleles can disperse (or be moved) to climates to which they are preadapted, the risk of extirpation and range contraction could be lessened. In the absence of natural migration, moving individuals long distances (for example, through assisted gene flow or migration) has been proposed as a way to conserve vulnerable populations by minimizing future maladaptation^{10–12}. Quantifying the need for populations to adapt and migrate (naturally or otherwise) can therefore provide a more comprehensive and biologically meaningful assessment of the risks that populations may face in future climates.

Here, we integrate concepts of novel and disappearing climates with adaptive genetic variation to map where populations may be most preadapted or maladapted to future climates (Fig. 1). As a case study, we used 75 single nucleotide polymorphisms (SNPs) in the *Populus* flowering-time gene network in balsam poplar (*P. balsamifera* L., Supplementary Table 1)^{13,14} to model future maladaptation to climate. Flowering-time genes are involved in phenology, which will be crucial to the maintenance of species ranges, tree growth, reproductive success and ultimately population viability in future climates^{15,16}. Using generalized dissimilarity models (GDMs (ref. ¹⁷), Supplementary Fig. 1) and gradient forests (GFs (ref. ¹⁸), Supplementary Fig. 2), we modelled the turnover of adaptive genetic variation as a function of climatic differences between 81 populations (Extended Data Fig. 1) and derived three metrics of potential maladaptation that we term local, forward and reverse genetic offsets. The results from the GDMs and GFs were similar, so we focus on GDMs in the main text unless specifically noted. The GF results are shown in the Supplementary Information.

Genetic offset, following ref. ¹⁹, represents the disruption of current genotype–climate relationships due to rapid shifts in climate. Local offset assumes that populations do not migrate in response to climate change and represents vulnerability in the context of a

¹Appalachian Lab, University of Maryland Center for Environmental Science, Frostburg, MD, USA. ²Department of Plant Biology, University of Vermont, Burlington, VT, USA. ³Present address: Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada. ✉e-mail: andy.gougherty@botany.ubc.ca

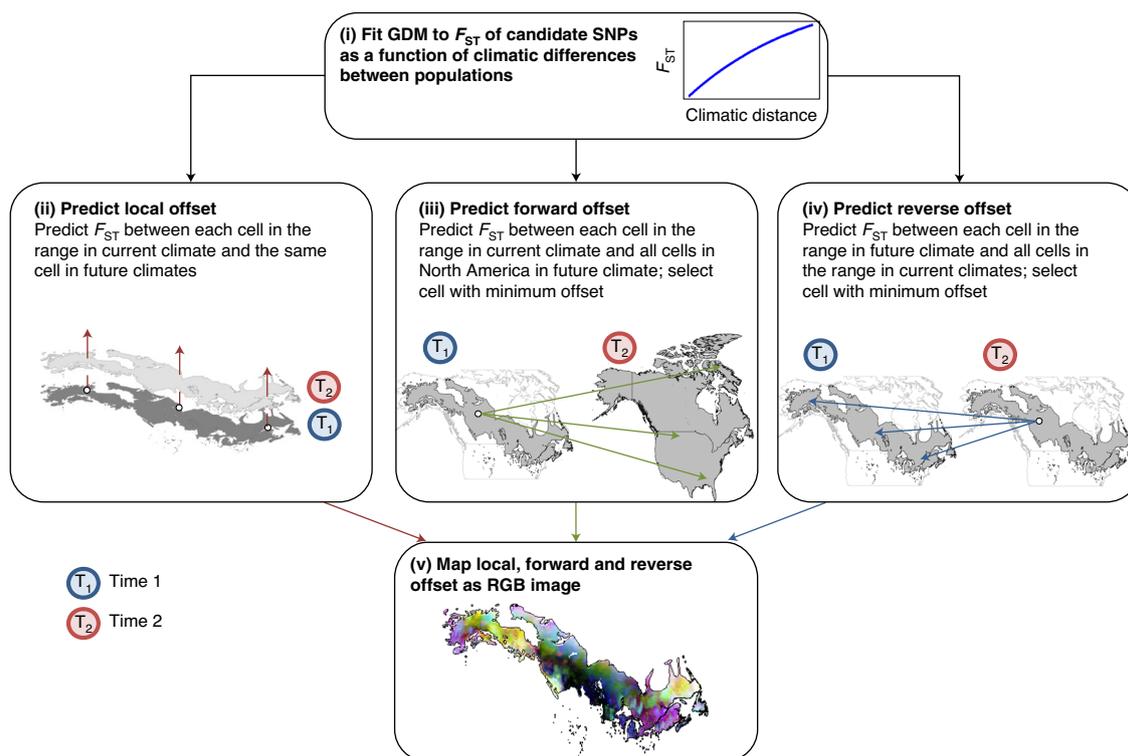


Fig. 1 | Schematic of how local, forward and reverse offsets were calculated and mapped for GDMs. After fitting a GDM to the pairwise fixation index (F_{ST}) of climatically adaptive SNPs (i), the model is used to predict local (ii), forward (iii) and reverse (iv) offsets. Local offset is calculated following Fitzpatrick and Keller¹⁹. Forward offset is calculated by predicting F_{ST} between each cell in the range in the current climate and all cells in North America in the future climate and selecting the minimum value. Reverse offset is calculated by predicting F_{ST} between each cell in the range in the future climate and all cells in the range in the current climate and selecting the minimum value. The maps are simultaneously plotted as a red–green–blue (RGB) image (v) (Fig. 2). The grey polygons are the balsam poplar’s contemporary range³⁵, with the exception of T_2 in iii (North America exclusive of Mexico).

population tolerating climate change in situ (Fig. 1(ii)). Local offset was calculated by estimating a temporal measure of offset for a resident population on the basis of the predicted in situ change in allele frequencies at climate-adaptive loci necessary to respond to local changes in climate over the next 50 years (2070). For long-lived sessile organisms, with long generation times, local offset may be the most relevant metric of vulnerability. Forward offset, in contrast, assumes that populations have unlimited migration ability and was calculated by identifying the minimum predicted offset if propagules or alleles, via gene flow, could move (or be moved) anywhere in North America (Fig. 1(iii)). Forward offset can be interpreted as the relative possibility of contemporary genotype–climate associations disappearing from the landscape assuming that populations were unconstrained by migration. Reverse offset is similar to the concepts of novel climate and novel species assemblages^{8,9}, but applied to the adaptive genetic composition of populations, representing the possibility that any population in the current range will be preadapted to a particular location in the future. Reverse offset was calculated by identifying the minimum offset between hypothetical populations within the current range in the future climate and populations in the current range in the current climate (Fig. 1(iv)). We also mapped the distance and direction populations would need to migrate to minimize forward offset, and we tested the effect of limiting migration to five distance bins (50, 100, 250, 500 and 1,000 km).

While the predicted patterns of local, forward and reverse offsets varied throughout the range of balsam poplar, some generalities did emerge (Fig. 2, Extended Data Figs. 2–4 and Supplementary Figs. 3 and 4). Local, forward and reverse offsets tended to be lowest in the centre of the range, indicating that these populations were predicted

to (1) experience minimal future disruption to the genotype–climate association at their current location (that is, low local offset, assuming no migration), (2) have current genotype–climate associations that are similar to those in other regions of North America under future climate (that is, low forward offset, assuming migration) and (3) have future genotype–climate associations that are similar to those in the current climate existing elsewhere in the range (that is, low reverse offset). The GDM and GF results were similar for local, forward and reverse offsets (spatially corrected r , 0.60, 0.67 and 0.45; P , 0.06, <0.01 and <0.01, respectively). Together, low local, forward and reverse offsets suggest that populations in the centre of the range may be the most preadapted to future climate for the phenology-associated genes we assessed. In contrast, the easternmost and northernmost parts of the range were predicted to have relatively high local, forward and reverse offsets (Fig. 2), indicating that there are no existing populations, either locally or elsewhere in the range, that are predicted to be preadapted to future climates in the eastern and northern parts of the range. High local and forward offsets suggest that eastern and northern populations are likely to be particularly vulnerable to climate change, as the impacts of local climate shifts cannot be mitigated by migration or movement to more suitable climate. As such, for this set of phenology-associated genes, populations in the eastern and northern parts of the range are likely to be the most at risk of extirpation due to climate change, potentially resulting in range contraction near the contemporary longitudinal range edges.

Local and forward offsets were most strongly associated with shifts in winter and/or summer precipitation (Supplementary Fig. 5), but they were also correlated with the underlying pattern of genetic

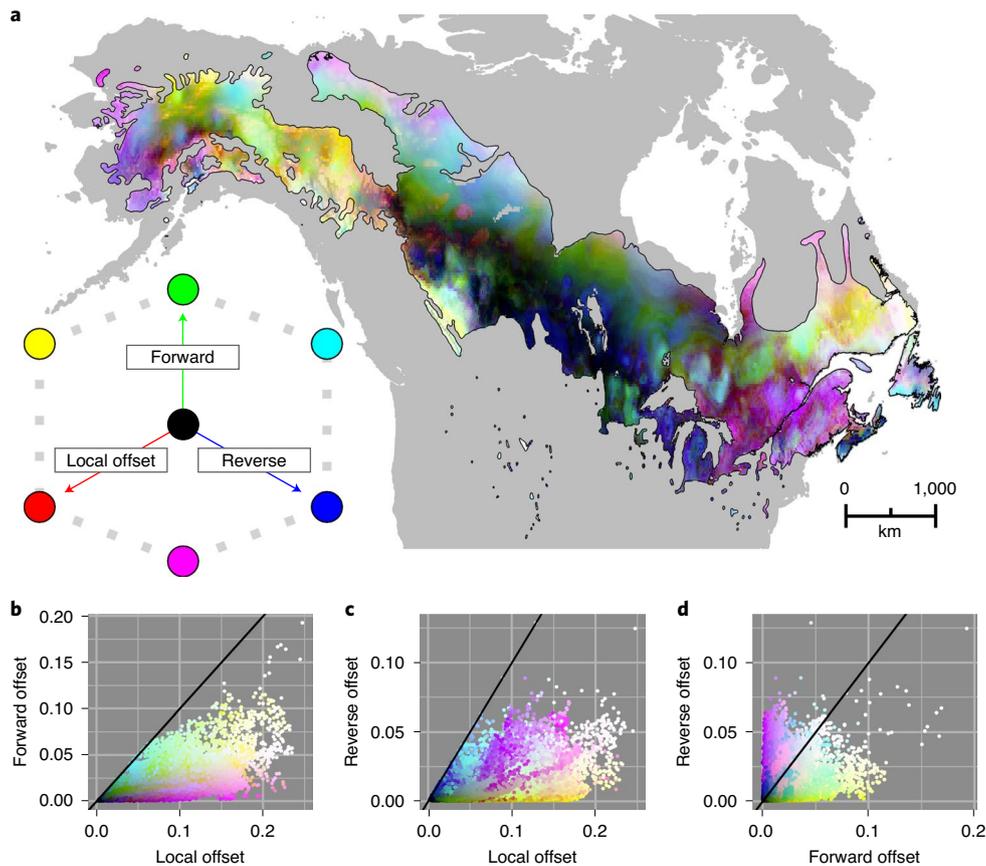


Fig. 2 | RGB map of local (red), forward (green) and reverse (blue) offsets. a, RGB map as shown in Fig. 1. **b–d**, Bivariate scattergrams of **a**, with 1:1 lines. The offset values were calculated from GDMs throughout the range of balsam poplar for 2070 and RCP 8.5. Brighter cells (closer to white) have relatively high values along each of the three axes, while darker cells (closer to black) have relatively lower values. The individual maps used in **a** are shown in Extended Data Fig. 3.

variability (that is, population structure). The strengths of the correlations between local/forward offsets and climatic shifts largely match the most important variables used to parameterize the GDMs (Supplementary Table 2) and GFs (Supplementary Fig. 2), in particular winter precipitation. The potential confounding influence of population structure on local/forward offsets is further complicated in that population structure itself is correlated with climate shifts, especially winter/summer precipitation and mean diurnal range. Additional work is needed to parse the influences of local climate shifts and population structure on the offset metrics.

Migration distances and direction

The geographic distances to locations that minimized forward offset were only weakly correlated with forward offset ($n_{\text{cell}}=29,942$, $r=0.12$, $P=0.35$), suggesting that, across populations, the need to migrate longer distances was not necessarily associated with a higher adaptive offset. The distances to locations that minimized future maladaptation for the phenology-associated genes we assessed (D_{min}) were largest in the eastern portion of the range, where they exceeded 5,000 km (Fig. 3, Extended Data Fig. 5 and Supplementary Figs. 6 and 7). The locations corresponding to D_{min} for much of the northeastern portion of the range were in mountainous regions in the western half of North America, indicating that populations in the easternmost portion of the range would need to migrate (or be moved) across nearly the entire North American continent to minimize future disruption of existing genotype–climate associations. These distances probably exceed what could

realistically be expected naturally given seed (and even pollen) dispersal limitations and the predominant west–east wind patterns in North America. The shortest D_{min} estimates, in contrast, occurred along the southern/southeastern range edge (near the northeastern United States and Ontario, Canada) and in the northern portion of the range in Alaska, often near mountainous areas. These short distances could be reachable by 2070, given that poplars are wind dispersed with short generation times. Few (<1%) locations had a D_{min} of zero, suggesting that populations in nearly all parts of the range would need to disperse propagules some distance to reach the locations that they are most preadapted to in the future, barring allele frequency change in situ.

Restricting the maximum allowable migration distance to one of multiple distance bins (that is, 50, 100, 250, 500 and 1,000 km) necessarily increased forward offset (Fig. 4 and Supplementary Figs. 8–12) but also revealed a declining benefit of lowered forward offset as migration distances increased. For instance, forward offset across all cells decreased by 52.8% ($n_{\text{cell}}=29,942$) when the maximum migration distance was expanded between 0 and 500 km, whereas the average decline in forward offset between 500 and 1,000 km was only 14.3%. This pattern suggests that, while there is a benefit to migration, forward offset can also be considerably reduced even if the global minimum offset cannot be reached.

Our analyses further indicate considerable variability in the direction populations would migrate to minimize future maladaptation of phenology-associated genes. While the GFs predicted mostly north–northwestward shifts to minimize maladaptation, the

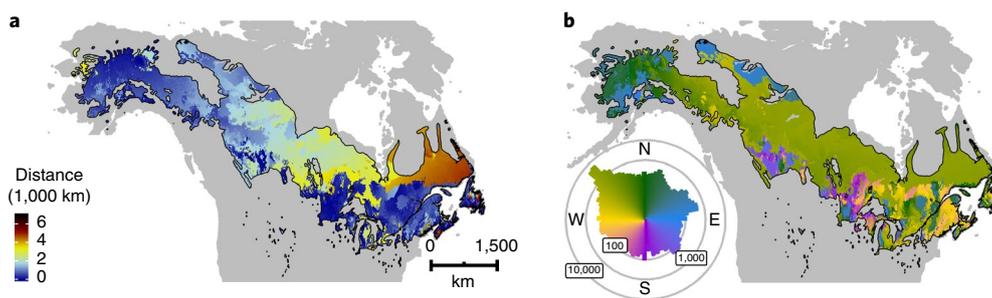


Fig. 3 | Distance and initial bearing to locations that minimize forward offset. a,b, Distance (**a**) and initial bearing (**b**) were calculated from the focal cell to the location in the future North American climate (2070, RCP 8.5) that minimizes the predicted F_{ST} from the GDMs. The polar histogram in **b** shows the \log_{10} number of cells in each bearing bin.

predictions from the GDMs were more complex. For example, while most (82.8%) locations within the range had an overall northward trajectory (that is, the location that minimized forward offset was at a higher latitude than the source cell), the GDMs predicted considerable variability along the southernmost range edge (Fig. 3b), where populations exhibited westward, eastward or even southward trajectories. The variability in population-level trajectories was especially apparent in the upper Great Lakes region, where migration direction varied over short distances. Recent observational studies have shown similar variability when populations are aggregated over entire species^{20,21}. Fei et al.²⁰, for instance, showed that over the past 30 years, eastern North American tree species have more often shown a westward shift in abundance than a poleward shift. The authors propose that this is due to shifting precipitation regimes and moisture availability increasing climatic suitability for eastern tree species in the centre of North America. Precipitation variables were similarly among the most important in our GDMs (Supplementary Table 2) and GFs (Supplementary Fig. 2), and changes in precipitation patterns may explain some of the non-northward trajectories in our predictions.

Advances and applications

The impacts of climate change across species ranges will be mediated by the collective ability of climatically adapted populations to adapt and/or migrate in response. Our study estimates the contributions of migration and adaptation by developing a spatially explicit understanding of the influence of local maladaptation, minimum migration distances and genotype–climate novelty on climatically adapted populations. In doing so, we attempt to shed new light on several major unresolved questions concerning how climate change may affect biodiversity, specifically: (1) where will climate change cause the greatest mismatch between climatically adapted populations and climate, and (2) which existing populations may be the most preadapted or maladapted to future climates? By simultaneously calculating multiple metrics of maladaptation and considering migration distances, our approach provides insight into the magnitude of in situ adaptation versus migration and helps elucidate where migration and adaptation may be most effective at reducing the genetic offset under future climates within the range. It is important to emphasize that our approach makes no attempt to predict the actual responses of populations over multiple generations to adapt or migrate²²; rather, we shed light on where these evolutionary processes may be most needed to maintain the status quo of current climate adaptation with regard to a set of phenology-associated genes. Nevertheless, our approach provides informative metrics of population-level vulnerability to climate change that may serve as a useful baseline for understanding where populations may be most at risk from climate change.

Contrary to some theoretical work (for example, see ref. ²³), we found that the vulnerability of balsam poplar to climate change

was not exclusively greatest along the trailing (southern) edge of the range, but rather at the longitudinal extremes of the range. The higher vulnerability of populations at the longitudinal margins of the range partly reflects the fact that the effects of temperature were secondary to the effects of winter precipitation in driving the differentiation of the SNPs we investigated. The effect of winter precipitation is probably partially due to several of the SNPs in the flowering-time gene network being involved in phenological processes that occur during the cool season (for example, vernalization and dormancy) and/or are related to water use (for example, the biosynthesis of abscisic acid, which contributes to the regulation of bud dormancy and drought response). These findings are consistent with previous work showing pre-season precipitation to be an important driver of spring phenology^{24,25}. Change in winter precipitation, the most important variable in our model, is predicted to be the greatest in the eastern and northernmost parts of the range of balsam poplar. Greater winter precipitation combined with warmer winter temperatures probably results in warmer/wetter winter conditions that populations have not experienced in the recent past, resulting in high adaptive offset in the eastern and northern parts of the range. These findings are consistent with recent work suggesting that accounting for local adaptation when predicting range shifts could yield results contrary to the leading/trailing-edge paradigm of range shifts (that is, as ranges shift poleward, trailing edge populations are most vulnerable to climate change as they will be the first to experience temperatures outside species' climatic niches^{26,27}). Similarly, numerous empirical studies have reported that recent climate-driven range shifts are rarely uniformly poleward^{20,21,28,29} and may be in multiple directions. Together, our results suggest that explicitly considering genotype–climate associations of multiple loci across multiple climatic gradients could yield a considerably more complex view of climate change responses than the poleward range shifts typically assumed in response to increased temperatures.

The metrics of population vulnerability to climate change that our approach provides may have applications in conservation and restoration efforts. For instance, assisted migration or gene flow, which rely on moving individuals to areas with suitable future climates outside the range (assisted migration) or within the range (assisted gene flow), could benefit from being able to identify which populations are most preadapted to future climates and how far these populations would have to disperse propagules. Assisted migration and gene flow are likely to provide the greatest benefit to dispersal-limited species with long generation times and may be the only option for some populations to avoid extirpation. Conversely, small, isolated populations with poor pollen and seed dispersal ability that are predicted to be maladapted to all future climates could be candidates for targeted germplasm sampling to ensure that rare adaptive alleles remain available for future breeding or restoration programmes. In instances when climatically adapted genes can be identified and ideally something about their function is known,

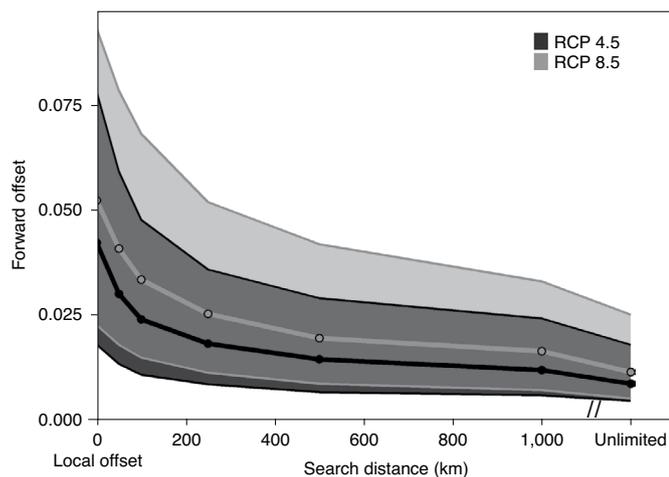


Fig. 4 | Effect of search distance on forward offset. Forward offset for 2070 was calculated from a GDM. The bands extend between the 25th and 75th percentiles, and the points are the median values.

genetic offset metrics could provide biologically meaningful guidance for the conservation of climatically adapted populations, beyond that offered by climate matching approaches or species distribution models.

Limitations and future work

While our approach offers unique insights into the magnitude of climate change disruption of climate adaptation, it also makes numerous simplifying assumptions that warrant further investigation and lend caution to their use in applied studies. First, the offset metrics implicitly assume that the current pattern of genetic differentiation across space can be used to estimate the disruption of genotype–climate associations. Because space-for-time substitutions are correlative, lack the mechanisms that will shape future patterns of genetic variation (for example, selection and gene flow) and do not account for potential plastic responses to climate change, they must be used with caution. However, using SNPs with an a priori relationship with climate¹⁴, related to a temporally and spatially varying trait (that is, phenology³⁰), helps ensure that we are modelling a robust, reliable genotype–climate signal. Furthermore, recent experimental work using information from common gardens suggests that offset metrics can be reliable metrics of climate maladaptation³¹. Second, because offset metrics are based on metrics of genetic differentiation (especially GDMs), using our approach does not identify the specific alleles that may need to shift frequencies to minimize maladaptation. We therefore cannot predict the actual novel genotypes in future populations—only that future populations may be genetically similar or dissimilar from current populations at the set of loci we assessed. Finally, our models emphasize only adaptation to climate and have been parameterized for a small set of SNPs involved in one aspect of climate adaptation. While focusing on climate adaptation is appropriate for SNPs related to phenology, many additional genes and gene pathways probably contribute in a polygenic fashion to climate adaptation, especially those related to abiotic stress responses. Additionally, populations are likely to experience unique biotic interactions in future climates³². It is therefore important to emphasize that the offset metrics calculated here are only relevant to the specific SNPs used in our study and are not meant to portray a comprehensive genome-wide view of climate maladaptation, nor are they generalizable to other portions of the genome responsive to non-climatic factors, such as biotic interactions or edaphic factors.

Our analyses were conducted on a small set of loci in balsam poplar associated with phenology, which have a well-studied physiological and phenotypic relationship with climate and are climatically adaptive in balsam poplar³³. However, our approach is generalizable to any number of loci having a robust association with climate. Further insights could be gained by assessing the pattern of forward and reverse offsets of loci associated with other climatically adaptive traits (for example, heat and drought tolerance and growth rates). Assessing loci associated with other traits could help elucidate the variable impacts that climate change may have on different parts of the genome, and could inform whether populations are preadapted to a single location on the landscape or, more likely, whether genomic regions underlying different functional traits will be preadapted to different locations on the future landscape. Such information will be crucial to understanding and mitigating the effects of climate change on adaptation in forest trees.

Conclusion

Populations, not species, respond to climate change, and these responses are likely to include a complex interplay of migration and adaptation to avoid extirpation^{1,2,34}. Here, we attempt a step towards accommodating this complexity in assessments of climate change vulnerability by estimating the contributions of population-level climate maladaptation, minimum migration distances and genotype–climate novelty. We found a rich assortment of risks from climate change across the range of balsam poplar, including high local, forward and reverse offsets and migration distances in the eastern portion of the range and low offsets and migration distances in the centre of the range. This suggests that eastern populations may face the greatest relative vulnerability to climate change and the greatest relative risk of future extirpation. More broadly, our work shows that, just as some climates and biological communities may disappear from the future landscape and novel ones may emerge in their place⁹, the same concepts apply to the genetic composition of climatically adapted tree populations. The concepts of forward and reverse genetic offsets provide a new way to consider population-level risk of future climate change that accounts for climate adaptation and goes beyond the constraints of species-level predictions.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41558-020-00968-6>.

Received: 12 May 2020; Accepted: 20 November 2020;

Published online: 18 January 2021

References

- Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T. & Curtis-McLane, S. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evol. Appl.* **1**, 95–111 (2008).
- Davis, M. B. & Shaw, R. G. Range shifts and adaptive responses to Quaternary climate change. *Science* **292**, 673–679 (2001).
- Ikeda, D. H. et al. Genetically informed ecological niche models improve climate change predictions. *Glob. Change Biol.* **23**, 164–176 (2017).
- Maguire, K. C., Shinneman, D. J., Potter, K. M. & Hipkins, V. D. Intraspecific niche models for ponderosa pine (*Pinus ponderosa*) suggest potential variability in population-level response to climate change. *Syst. Biol.* <https://doi.org/10.1093/sysbio/syy017> (2018).
- Smith, A. B., Godsoe, W., Rodríguez-Sánchez, F., Wang, H.-H. & Warren, D. Niche estimation above and below the species level. *Trends Ecol. Evol.* **34**, 260–273 (2019).
- 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).

7. Wang, T., O'Neill, G. A. & Aitken, S. N. Integrating environmental and genetic effects to predict responses of tree populations to climate. *Ecol. Appl.* **20**, 153–163 (2010).
8. Radeloff, V. C. et al. The rise of novelty in ecosystems. *Ecol. Appl.* **25**, 2051–2068 (2015).
9. Williams, J. W. & Jackson, S. T. Novel climates, no-analog communities, and ecological surprises. *Front. Ecol. Environ.* **5**, 475–482 (2007).
10. Aitken, S. N. & Whitlock, M. C. Assisted gene flow to facilitate local adaptation to climate change. *Annu. Rev. Ecol. Evol. Syst.* **44**, 367–388 (2013).
11. Vitt, P., Havens, K., Kramer, A. T., Sollenberger, D. & Yates, E. Assisted migration of plants: changes in latitudes, changes in attitudes. *Biol. Conserv.* **143**, 18–27 (2010).
12. Williams, M. I. & Dumroese, R. K. Preparing for climate change: forestry and assisted migration. *J. For.* **111**, 287–297 (2013).
13. Keller, S. R., Levensen, N., Olson, M. S. & Tiffin, P. Local adaptation in the flowering-time gene network of balsam poplar, *Populus balsamifera* L. *Mol. Biol. Evol.* **29**, 3143–3152 (2012).
14. Keller, S. R., Chhatre, V. E. & Fitzpatrick, M. C. Influence of range position on locally adaptive gene–environment associations in *Populus* flowering time genes. *J. Hered.* **109**, 47–58 (2018).
15. Chuine, I. Why does phenology drive species distribution? *Phil. Trans. R. Soc. B* **365**, 3149–3160 (2010).
16. Morin, X., Viner, D. & Chuine, I. Tree species range shifts at a continental scale: new predictive insights from a process-based model. *J. Ecol.* **96**, 784–794 (2008).
17. Ferrier, S., Manion, G., Elith, J. & Richardson, K. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib.* **13**, 252–264 (2007).
18. Ellis, N., Smith, S. J. & Pitcher, C. R. Gradient forests: calculating importance gradients on physical predictors. *Ecology* **93**, 156–168 (2012).
19. 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).
20. Fei, S. et al. Divergence of species responses to climate change. *Sci. Adv.* **3**, e1603055 (2017).
21. VanDerWal, J. et al. Focus on poleward shifts in species' distribution underestimates the fingerprint of climate change. *Nat. Clim. Change* **3**, 239–243 (2013).
22. Shaw, R. G. From the past to the future: considering the value and limits of evolutionary prediction. *Am. Nat.* **193**, 1–10 (2018).
23. Hampe, A. & Petit, R. J. Conserving biodiversity under climate change: the rear edge matters. *Ecol. Lett.* **8**, 461–467 (2005).
24. Yun, J. et al. Influence of winter precipitation on spring phenology in boreal forests. *Glob. Change Biol.* **24**, 5176–5187 (2018).
25. Fu, Y. H. et al. Unexpected role of winter precipitation in determining heat requirement for spring vegetation green-up at northern middle and high latitudes. *Glob. Change Biol.* **20**, 3743–3755 (2014).
26. Peterson, M. L., Doak, D. F. & Morris, W. F. Incorporating local adaptation into forecasts of species' distribution and abundance under climate change. *Glob. Change Biol.* **25**, 775–793 (2019).
27. Atkins, K. E. & Travis, J. M. J. Local adaptation and the evolution of species' ranges under climate change. *J. Theor. Biol.* **266**, 449–457 (2010).
28. Chen, I.-C., Hill, J. K., Ohlemuller, R., Roy, D. B. & Thomas, C. D. Rapid range shifts of species associated with high levels of climate warming. *Science* **333**, 1024–1026 (2011).
29. Groom, Q. J. Some poleward movement of British native vascular plants is occurring, but the fingerprint of climate change is not evident. *PeerJ* **1**, e77 (2013).
30. Olson, M. S. et al. The adaptive potential of *Populus balsamifera* L. to phenology requirements in a warmer global climate. *Mol. Ecol.* **22**, 1214–1230 (2013).
31. Fitzpatrick, M., Chhatre, V., Soolanayakanahally, R. & Keller, S. Experimental support for genomic prediction of climate maladaptation using the machine learning approach Gradient Forests. Preprint at <https://doi.org/10.22541/au.159863198.86187354> (2020).
32. Blois, J. L., Zarnetske, P. L., Fitzpatrick, M. C. & Finnegan, S. Climate change and the past, present, and future of biotic interactions. *Science* **341**, 499–504 (2013).
33. Keller, S. R. et al. Climate-driven local adaptation of ecophysiology and phenology in balsam poplar, *Populus balsamifera* L. (Salicaceae). *Am. J. Bot.* **98**, 99–108 (2011).
34. Alberto, F. J. et al. Potential for evolutionary responses to climate change—evidence from tree populations. *Glob. Change Biol.* **19**, 1645–1661 (2013).
35. Little, E. L. *Atlas of United States Trees* (US Dept of Agriculture, Forest Service, 1971).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Methods

Balsam poplar. Balsam poplar is a northern broad-leaved forest tree species that occurs over a large portion of the boreal region in North America spanning more than 30 degrees of latitude and multiple broad climatic gradients³⁵. The centre of the range is in central Canada, which is expected to see among the highest levels of future warming in North America³⁶. Trees in the *Populus* genus have emerged as a model system for landscape genomic studies of local adaptation to climate³⁷, and studies of balsam poplar, in particular, have shown it to be climatically adapted for numerous functional traits^{33,38} and its adaptive variation to be climatically structured³⁹.

GDMs and GFs. We used GDMs and GFs to map predictions of local adaptation to climate (Supplementary Figs. 1 and 2). GDMs (ref. ¹⁷) use a type of nonlinear matrix regression that accounts for the curvilinear relationship between climatic (and optionally, geographic) distance and genetic differentiation among populations separated along environmental gradients. A GDM fits a linear combination of monotonic 1-spline functions to each predictor that characterizes its unique contribution to the magnitude (height) and rate (shape) of genetic differentiation along that gradient. GFs (ref. ¹⁸), in contrast, are an extension of the random forest approach, which models allele frequencies as nonlinear functions along climate gradients¹⁹. A GF fits an ensemble of regression trees using random forest and then constructs ‘cumulative importance’ turnover functions from these models by determining how well partitions distributed at numerous ‘split values’ along each gradient explain genetic patterns on either side of a split. Locations of high split importance represent thresholds where genetic change is rapid, and important gradients will have greater overall cumulative importance.

We fit GDMs and GFs to genetic differentiation (F_{ST}) and minor allele frequencies, respectively, of 75 SNPs in the *Populus* flowering-time gene network genotyped in 995 individuals from 81 populations from across the range of balsam poplar¹⁴. Genes in the flowering-time network are associated with both reproductive and vegetative plant phenology by regulating the timing of seasonal growth, dormancy and reproduction within the permissive growing season. We selected SNPs that had a relationship with environment, identified using Bayenv2 (ref. ⁴⁰) and latent factor mixed models⁴¹. We used SNPs and population data described in ref. ¹⁴ and used a similar approach to identifying outlier SNPs but substituted principal components for summer and winter mean temperature and precipitation, isothermality and mean diurnal range⁴²—identical to the climate variables used in our GDMs and GFs (see below). Outliers from the latent factor mixed models were identified on the basis of a significant relationship with any of the six bioclimatic variables, assuming a false discovery rate of 0.01. We transformed the P values before calculating the false discovery rate using the genomic-control method, implemented in the `lfmm.pvalues` function in the LEA package⁴³ in R. Outliers from Bayenv2 were identified on the basis of having both a significant Spearman's ρ and a Bayes factor in the top 5% of loci for each bioclimatic variable. This resulted in 75 SNPs identified as outliers across 24 genes by either Bayenv2 or latent factor mixed models. On the basis of these 75 climate-associated candidate SNPs, for the GDMs, we calculated a multilocus pairwise F_{ST} among the 81 populations with at least 5 individuals using the `genet.dist` function in the `hierfstat` package⁴⁴ in R. Any pairwise F_{ST} values less than zero were assigned a value of zero. For the GFs, we used minor allele frequencies from ref. ¹⁴.

The GDMs and GFs were parameterized with six bioclimatic variables that lacked strong correlation ($|r| < 0.75$). These included summer and winter mean temperature (`bio10` and `bio11`) and precipitation (`bio18` and `bio19`), isothermality (`bio3`), and mean diurnal range (`bio2`) downloaded from the WorldClim dataset⁴², at a resolution of ten arc minutes. We considered including geographic distance in the GDMs but found a negligible improvement in model performance with its inclusion (cross-validation R^2 with geographic distance, 68.38%; without geographic distance, 68.04%), suggesting that the climate variables in the models are accounting for much of the explained spatial variation in genetic differentiation. The models were parameterized in the current climate (centred on ~1975) and predicted to the future climate (centred on 2070) of a composite average of five global circulation models (UCAR Community Climate System Model, NOAA Geophysical Fluid Dynamics Laboratory Coupled Physical Model, MET Office Hadley Center Earth System Model, NASA Goddard Institute for Space Studies-E2-R and Norwegian Earth System Model). We performed all analyses using two different emission scenarios (RCP 4.5 and 8.5) for 2070. The results and discussion refer to the composite mean of the five RCP 8.5 projections for 2070 unless specifically noted. The GDMs were fit using the `gdm` package⁴⁵ and the GFs using the `gradientForest` package⁴⁶ in R.

Genetic offset metrics. The GDMs used to predict genetic offset metrics explained 70.8% of the deviance in F_{ST} , and tenfold cross-validation revealed an R^2 of 68.0%. Variable permutation revealed that winter precipitation was the most important variable in the model. Isothermality (mean diurnal range divided by annual temperature range) and summer temperature were of secondary importance, while summer precipitation, mean diurnal range and winter temperature were least important (Supplementary Table 2). The GFs had an average R^2 of 38.5% across all 75 SNPs (range, 1–71%) and identified winter precipitation as the most important

variable, followed by mean diurnal range, summer precipitation, isothermality, winter temperature and summer temperature (Supplementary Fig. 2).

We used GDMs (Fig. 1(i)) and GFs to quantify the disruption of adaptive genotype–climate associations expected under climate change using three different formulations of genetic offset: local (Fig. 1(ii)), forward (Fig. 1(iii)) and reverse (Fig. 1(iv)). It is important to note that the units of the offset metrics differ between GDMs and GFs. While GDMs directly model F_{ST} , GFs transform climate variables to genetic importance values. Genetic distances are then calculated as the Euclidean distance between genetic importance values¹⁹. Below, we refer to these predicted F_{ST} values and Euclidean distances as ‘genetic differentiation’ unless otherwise specified.

Following Fitzpatrick and Keller¹⁹, local offset (a metric of maladaptation with regard to a population tolerating climate change *in situ*) was calculated by predicting differentiation for climatically adaptive SNPs between the present and future climate at the same location, assuming no migration or gene flow. Forward offset is the minimum expected disruption in the genotype–climate association assuming that populations have unlimited dispersal capacity. Forward offset was quantified by first using the fitted GDMs and GFs to predict genetic differentiation between the current climate at each focal grid cell within the current range of balsam poplar and all future climate grid cells in North America (exclusive of Mexico). In other words, for each grid cell in the current range, we obtained a predicted genetic offset from that grid cell under the current climate to all other grid cells in North America under the future climate. Next, for each current climate grid cell, we identified the minimum predicted differentiation value (F_{ST} or Euclidean distance) from the pool of differentiation values across all future grid cells. We term the minimum value ‘forward offset’. The grid cell that minimizes the predicted genetic differentiation represents the location that minimizes forward offset, and the distance to that location represents the migration distance required to minimize maladaptation (D_{min}). If multiple cells had an identical minimum offset, the one nearest to the focal pixel was chosen. High values of forward offset indicate that the population in that location is maladapted to all future North American climates and that contemporary genotype–climate associations in that population may disappear from the landscape. To assess the sensitivity of forward offset to dispersal constraints, we tested how forward offset varied when migration was limited to five distance classes: 50, 10, 250, 500 and 1,000 km. In addition to geographic distances, we calculated the initial bearing that populations would follow if they were to migrate to the location that minimized forward offset. Distance and bearing were calculated with the `distGeo` and `bearing` functions, respectively, in the `geosphere` package⁴⁶ in R.

Reverse offset follows the same idea as forward offset but is calculated from the future climate to the current climate. In this case, we first used the fitted GDMs and GFs to predict genetic differentiation between each future climate grid cell within the current range of balsam poplar and all current climate grid cells within the current range of balsam poplar. We thus obtained a predicted genetic offset from each grid cell under the future climate to all other grid cells under the current climate, both within the current range of balsam poplar. From these predictions, we then identified the current climate grid cell with the minimum predicted offset from the pool of offset predictions across all current grid cells. We term the minimum value ‘reverse offset’. Reverse offset provides a metric of how novel the future genotype–climate association is predicted to be at a given site, relative to existing genotype–climate associations present at any location throughout the range under the current climate. As such, high values of reverse offset indicate genetic novelty, as there is no analogous genotype–climate association found anywhere in the current landscape. Note that for reverse offset, we only considered pixels within the current range of balsam poplar in both periods to ensure that future novelty in genotype–climate associations was quantified only with respect to locations where balsam poplar currently occurs (that is, within the current range) as opposed to the entirety of North America, as was done for forward offset. In other words, source populations for the reverse offset estimates can come only from existing populations, whereas existing populations could in principle occur at any future location, given suitable climate and sufficient dispersal.

While other approaches could be used to identify forward and reverse offsets (such as binning, rounding or thresholding offset values), our metrics were based on minimum values, as this allowed us to avoid selecting an offset threshold and allowed the minimum to occur anywhere in North America (in the case of forward offset) or anywhere within the range (in the case of reverse offset). Other well-studied systems, however, could benefit from applying thresholds, especially if a particular F_{ST} was known to be linked to a reduction in fitness or climate maladaptation (as could be quantified from provenance trials).

Note that while the offset metrics can inform the magnitude of expected genetic change required to minimize maladaptation, they tell us less about the nature of that change. Because offset metrics represent an aggregate metric of genetic change (F_{ST} and multidimensional Euclidean distance for GDMs and GFs, respectively), a more mechanistic understanding of particular loci responsive to shifts in climate would require a disaggregated, per-locus estimate.

To better understand the relationship between local climate change, population structure and local and forward offsets, we quantified the correlations (Spearman's) between offset metrics, population structure and the magnitude of local climate change across our six climate variables. Reverse offset was not included in these analyses because this metric is not necessarily relevant to populations in their

current locations. Because we lacked a common set of neutral markers across all populations, we used a standardized set of ancestry coefficients from ref. ⁴⁷ for 85 balsam poplar populations to approximate population structure. The magnitude of local climate change was estimated as the difference between the future and current climates for each variable in our analysis.

To simultaneously visualize local, forward and reverse offsets, we mapped these three metrics as the red, green and blue bands of an RGB image, respectively. Because the values of local offset were systematically higher than those of forward and reverse offsets, we rescaled the values within each band to their quantiles before plotting. This ensured that the full range of each colour was possible in the RGB images, and is analogous to a histogram equalization performed on each band. We also tested for correlations between local, forward and reverse offsets and distances. We used a spatially corrected Pearson correlation coefficient to quantify these relationships. The spatial correlations were implemented with the SpatialPack package⁴⁸ in R after projecting latitude/longitude coordinates to an equidistant projection (Azimuthal equidistant). The R code, population occurrences and genetic data to calculate local, forward and reverse offsets are available at github.com/agougher/poplarAdaptiveOffset.

Because the results from the GDMs and GFs were similar, we focus the main text on GDMs as the output from the GDM prediction (F_{ST}) has direct biological meaning, unlike that from GFs (Euclidean distance of genetic importance values).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data are publicly available. The allele frequencies are available in ref. ¹⁴, and the climate data are available at <https://www.worldclim.org>.

Code availability

The R code and genetic data to calculate local, forward and reverse offsets are available at github.com/agougher/poplarAdaptiveOffset.

References

36. Romero-Lankao, P. et al. in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects* (eds Barros, V. R. et al.) 1439–1498 (Cambridge Univ. Press, 2014).
37. Fetter, K. C., Gugger, P. F. & Keller, S. R. in *Comparative and Evolutionary Genomics of Angiosperm Trees* (eds Groover, A. & Cronk, Q.) 303–333 (Springer International, 2017); https://doi.org/10.1007/7397_2016_19
38. Soolanayakanahally, R. Y., Guy, R. D., Silim, S. N., Drewes, E. C. & Schroeder, W. R. Enhanced assimilation rate and water use efficiency with latitude through increased photosynthetic capacity and internal conductance in balsam poplar (*Populus balsamifera* L.). *Plant Cell Environ.* **32**, 1821–1832 (2009).
39. Chhatre, V. E. et al. Climatic niche predicts the landscape structure of locally adaptive standing genetic variation. Preprint at <https://doi.org/10.1101/817411> (2019).
40. Günther, T. & Coop, G. Robust identification of local adaptation from allele frequencies. *Genetics* **195**, 205–220 (2013).
41. Frichot, E., Schoville, S. D., Bouchard, G. & François, O. Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol. Biol. Evol.* **30**, 1687–1699 (2013).
42. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
43. Frichot, E. & François, O. LEA: an R package for landscape and ecological association studies. *Methods Ecol. Evol.* **6**, 925–929 (2015).
44. Goudet, J. & Jombart, T. hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-22 (2015).
45. Manion, G., Lisk, M., Nieto-Lugilde, D., Mokany, K. & Fitzpatrick, M. gdm: Generalized dissimilarity modeling. R package version 1.3.11 (2018).
46. Hijmans, R. J. geosphere: Spherical trigonometry. R package version 1.5-10 (2019).
47. Gougherty, A. V., Chhatre, V. E., Keller, S. R. & Fitzpatrick, M. C. Contemporary range position predicts the range-wide pattern of genetic diversity in balsam poplar (*Populus balsamifera* L.). *J. Biogeogr.* **47**, 1246–1257 (2020).
48. Vallejos, R., Osorio, F. & Bevilacqua, M. *Spatial Relationships Between Two Georeferenced Variables: With Applications in R* (Springer, 2018).

Acknowledgements

We thank V. E. Chhatre for providing comments on an earlier draft of this manuscript. This work was supported by National Science Foundation Plant Genome Research award no. 1461868 to S.R.K. and M.C.F., and an UMCES PhD fellowship to A.V.G.

Author contributions

A.V.G. and M.C.F. conceived the study. S.R.K. processed and provided the genetic information. A.V.G. analysed the data and led the writing, with contributions and discussion from all authors.

Competing interests

The authors declare no competing interests.

Additional information

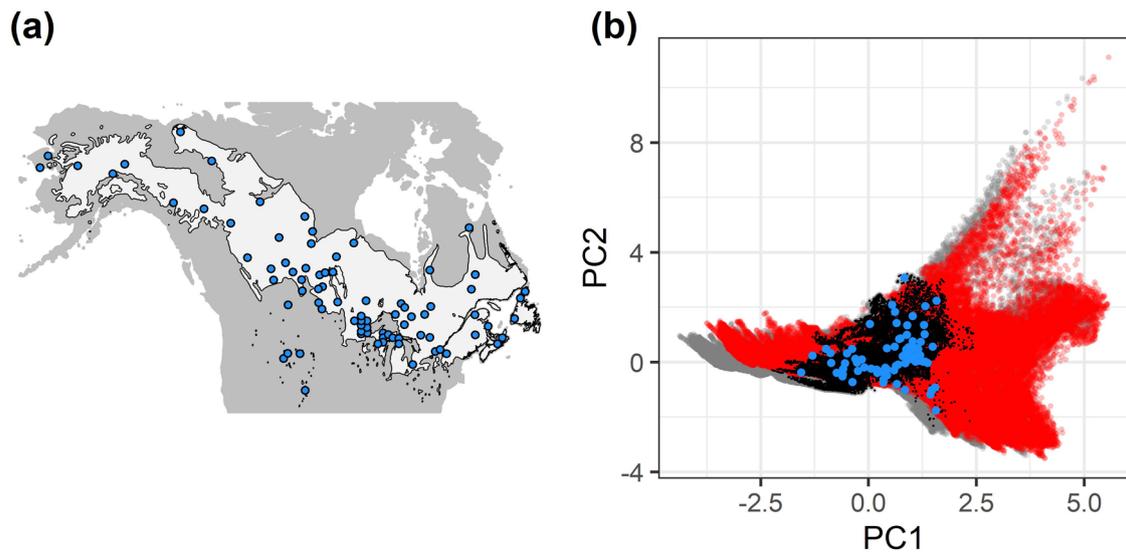
Extended data is available for this paper at <https://doi.org/10.1038/s41558-020-00968-6>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41558-020-00968-6>.

Correspondence and requests for materials should be addressed to A.V.G.

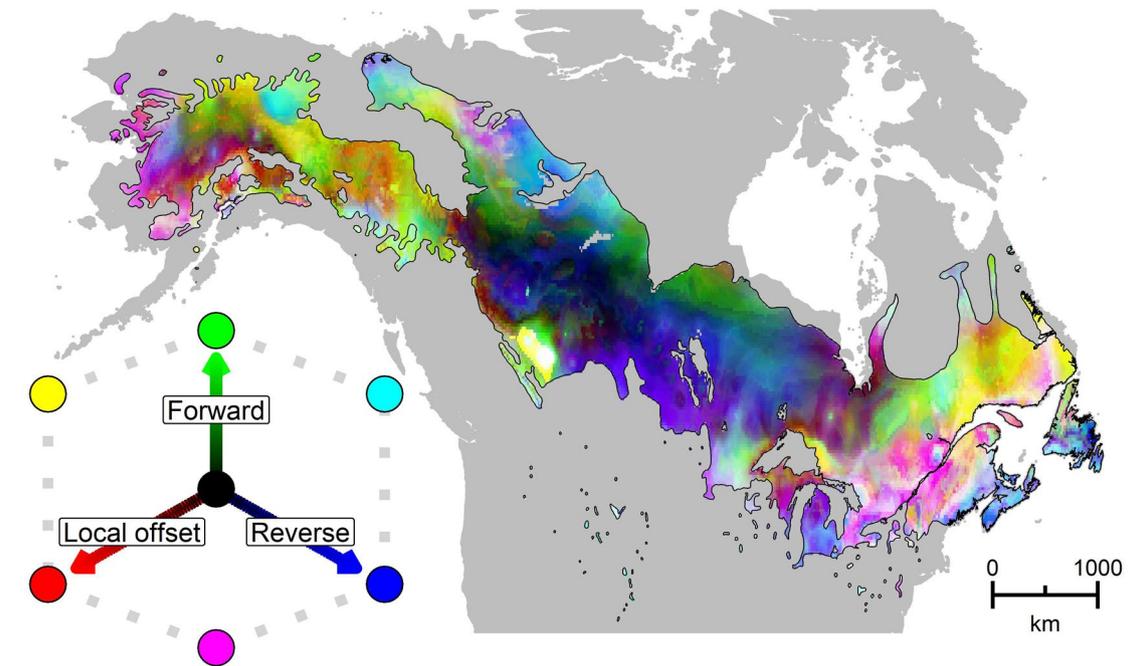
Peer review information *Nature Climate Change* thanks Erin Landguth, Christian Rellstab and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

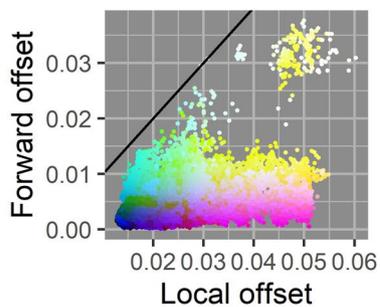


Extended Data Fig. 1 | Sampled population locations in geographic and climatic space. a, Geographic locations of populations used in GDM and GF models (blue circles), within balsam poplar's range (white polygon). **b**, Position of populations (blue circles), and cells in balsam poplar's range (black dots) in current North American climate space (gray circles). Red circles show the composite future climate (RCP 8.5) of North America used in predictions. Climate space is shown as the first two principal components (PCs) of current North American climate (mean diurnal range, isothermality, mean summer temperature, mean winter temperature, summer precipitation, winter precipitation), with future climates predicted into the same PCA space. **b**, is shown only for visualization purposes.

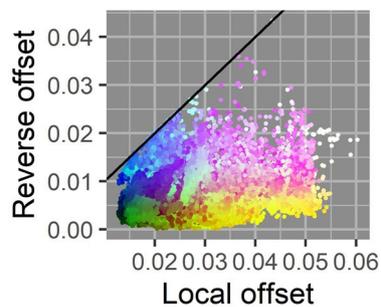
(a)



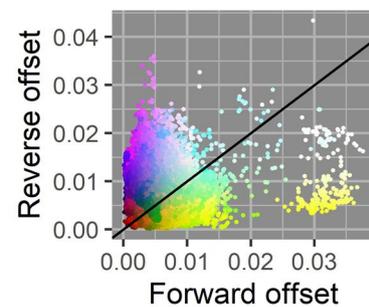
(b)



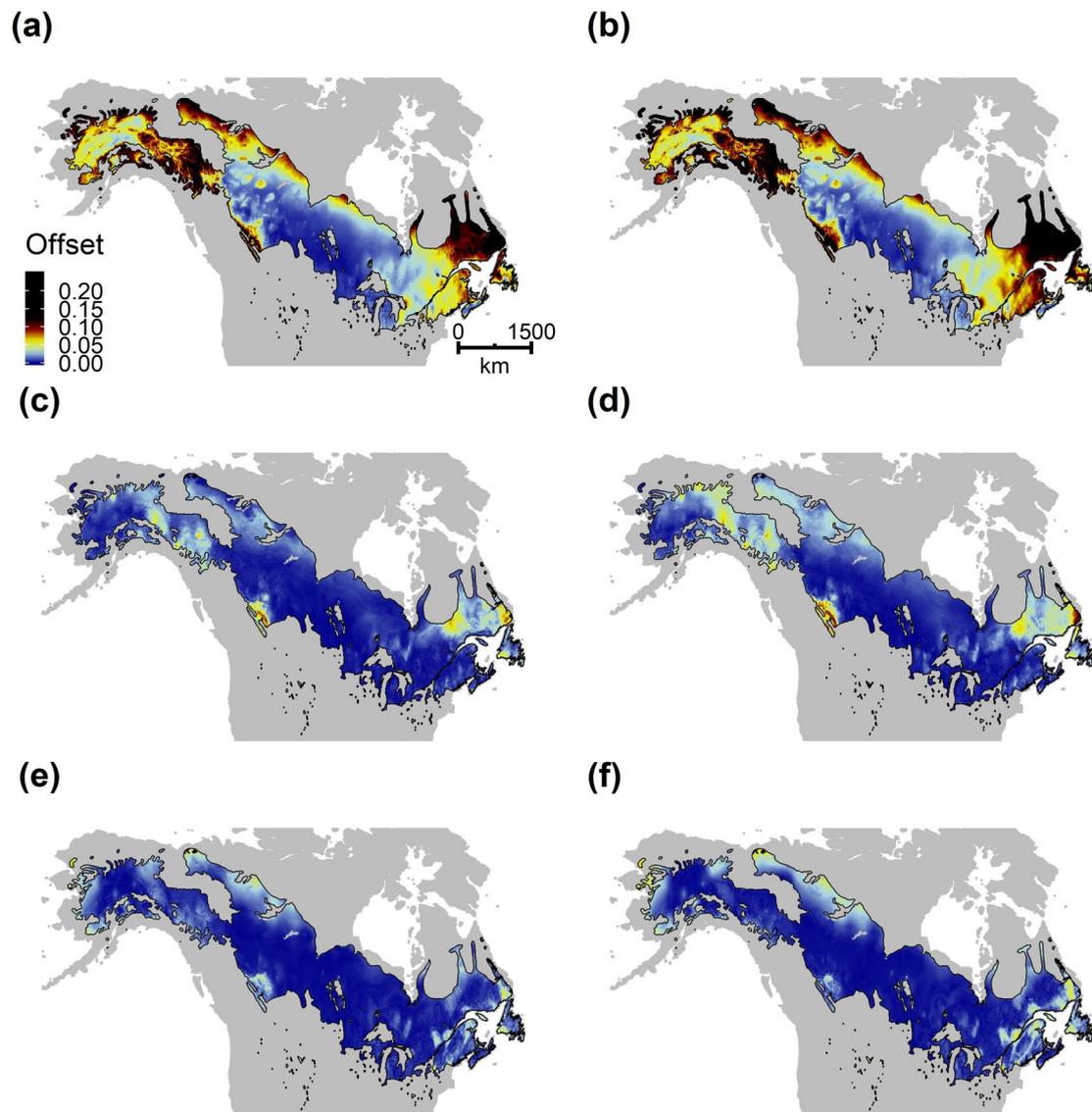
(c)



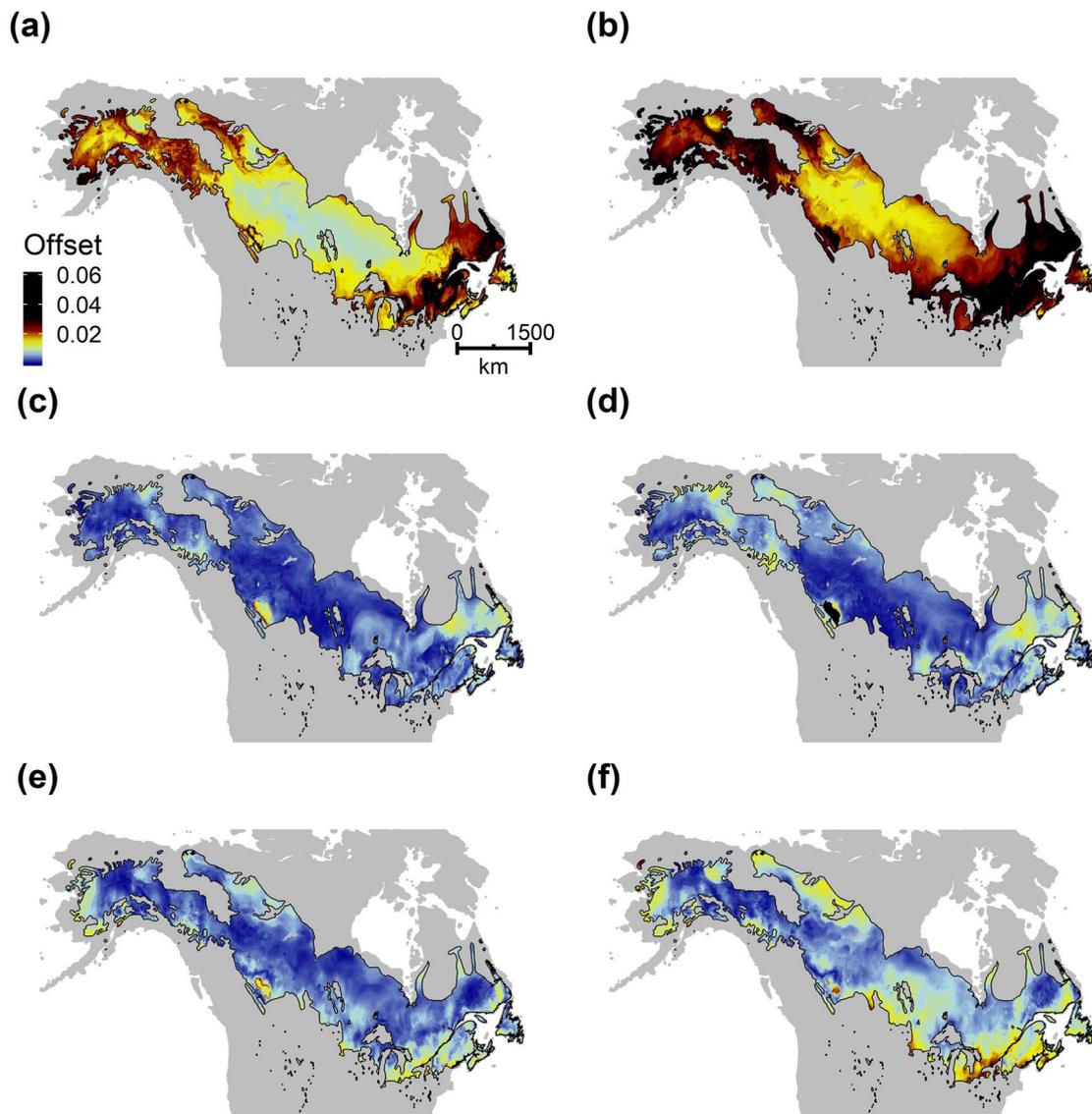
(d)



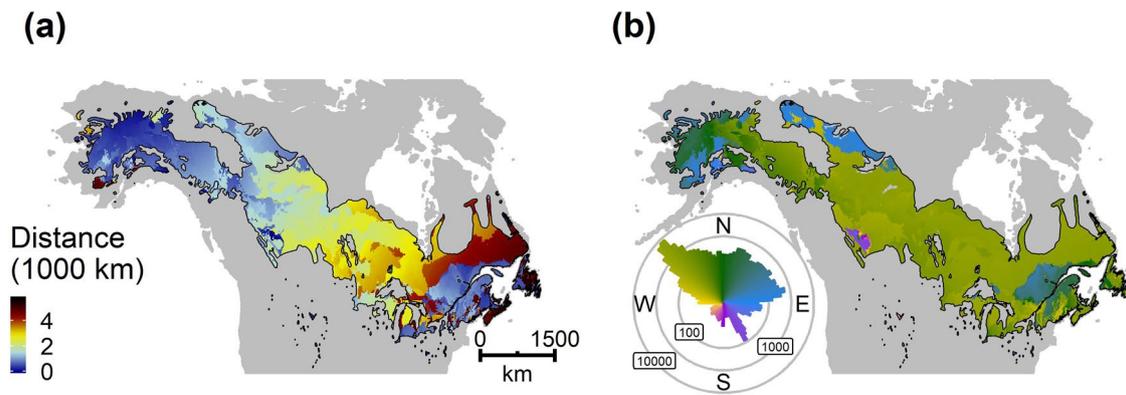
Extended Data Fig. 2 | Red-green-blue map of local (red), forward (green), and reverse (blue) offsets. Offset values were calculated from Gradient Forest throughout the range of balsam poplar for 2070 and RCP 8.5. Brighter cells, closer to white, have relatively high values along each of the three axes while darker cells, closer to black, have relatively lower values. **b-d**, Bivariate scattergrams of (a), with 1:1 lines. Individual maps used in (a) are shown in Extended Data Fig. 4.



Extended Data Fig. 3 | Local, forward, and reverse offsets from generalized dissimilarity models for balsam poplar. **a & b**, Local genetic offset, **(c & d)** forward offset, and **(e & f)** reverse offset from a generalized dissimilarity model for RCP 4.5 (first column; a, c, e) and RCP 8.5 (second column; b, d, f) for 2070. Note the non-linear color scale.



Extended Data Fig. 4 | Local, forward, and reverse offsets from Gradient Forest for balsam poplar. a & b, Local genetic offset, (**c & d**) forward offset, and (**e & f**) reverse offset from a Gradient Forest model for RCP 4.5 (first column; a, c, e) and RCP 8.5 (second column; b, d, f) for 2070.



Extended Data Fig. 5 | Distance and initial bearing to locations that minimize forward offset. Distance and (b) initial bearing were calculated from the focal cell to the location in future North American climate (2070, RCP 8.5) that minimizes predicted offset from a Gradient Forest model. Polar histogram in (b) shows the log₁₀ number of cells in each bearing bin.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Genetic differentiation of phenology-associated, climatically adaptive loci in balsam poplar was predicted spatially and temporally, in order to assess the effect of climate change on local climate adaptation.
Research sample	Previously published genetic data was collect for 995 individuals in 81 populations. We used 33 loci in the flowering time gene network that have previously shown an associated with climate. Genetic data is previously published in Keller, Chhatre, & Fitzpatrick (2018),
Sampling strategy	The 81 populations are dispersed throughout the range of balsam poplar, and capture the major climatic gradients within the range.
Data collection	Data were collected from published sources, including the literature and public websites.
Timing and spatial scale	Models of genetic differentiation were parameterized in current climate (centered on ~1975) and predicted to climates in 2070. Spatial scale of models were 10 arc-minutes.
Data exclusions	No data were excluded from analyses.
Reproducibility	All attempts to replicate the results of our model were successful.
Randomization	Randomization was not relevant to this study as genotyped individuals were assigned the population described in the source publication.
Blinding	Blinding was not relevant to this study as all data came from previously published sources.
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging