Genetic and environmental constraints causing species' range limits

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Antoine Perrier

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Yvonne Willi (Universität Basel)

Dr. Kay Lucek (Universität Basel)

PD Dr. Lukas Schärer (Universität Basel)

Dr. Joëlle Ronfort (Institut national de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, Montpellier, Frankreich)

Basel, den 15. September 2020

Prof. Dr. Martin Spiess

Dekan

Table of contents

Acknowledgements

I would like to express my sincere gratitude to my supervisor Prof. Dr. Yvonne Willi for providing me the opportunity to conduct research in her laboratory, and for her continued guidance and support throughout my work.

I would like to thank Dr. Lukas Schärer, Dr. Joëlle Ronfort and Dr Kay Lucek for being part of my thesis committee and final jury of my PhD defense. I also want to thank Kay Lucek for his numerous advices in statistical programing and writing, and the inspiring discussions on speciation, and more generally on Evolutionnary Biology.

I also thank Dr. Josh Van Buskirk, for his advices on statistical analysis and his helpful comments on my manuscripts.

I further want to thank my colleague Darío Sánchez-Castro, for sharing this PhD journey, and all the adventures and new experiences in our time spent in the USA.

A big thank you also to Olivier Bachmann, Susanna Riedl, Georg Armbruster and Markus Funk for their advice and help to carry out all my experiments, and for the countless hours spent counting seeds and measuring leaves; and to Franziska Grob and Maura Ellenberger for their resourcefulness in solving administrative and organizational questions.

Many thanks to all my friends and colleagues from the University of Basel, for the great discussions, advices and generally the great times spent together: Nora Walden, Selim Bouaouina, Alessio Maccagni, Janisse Deluigi, Hannah Augustijnen, Jessica Heblack, Judith Schepers, Jennifer Mark, Guillaume Wos, Markus Fracassetti, and all my colleagues from the Botanical Institute at Schönbeinstrasse and Hebelstrasse.

Finally, I would like to thank all the members of my family and my friends from all over the planet, for their moral support and encouragement during this thesis, and especially Wesley, I could never thank you enough for all the love and support you keep providing me.

Summary

Human-caused global change has led to shifts in the geographic distribution of many wild species. This has renewed the interest of understanding the factors that shape species' contemporary range limits from both an ecological and evolutionary perspective. Recent evolutionary theory particularly emphasized the role of past demographic processes and neutral evolution in contributing to range limits. The aim of my thesis was to study these factors and their interaction with the environment experienced at range edges in an empirical system, the North American plant *Arabidopsis lyrata*. By crossing populations of varying range position and demographic history, and raising their offspring in gardens distributed across and beyond the species range, I found that populations with a history of small size due to past range expansion or rear-edge isolation suffered from increased expression of mutational load driven by genetic drift. This latter effect was even stronger under environmental stress, particularly under a warmer climate. Furthermore, populations at range edges with heightened past exposure to genetic drift had a reduced signature of climate adaptation. Finally, I compared *A. lyrata* and a novel species it gave rise to, *A. arenicola*, with a more northern distribution, in a climate chamber experiment. This new taxon diverged from *A. lyrata* in coping with a cool climate and strong reproductive isolation, most likely allowing it to to colonize subarctic regions and escape maladaptive gene flow. Results generally support the newer evolutionary theory about a predominant role of neutral evolution in contributing to geographic range limits, via genetic drift opposing purifying and directional selection. The study of sister taxa however shows that these constraints to evolution at range limits are not absolute, and can be broken.

Introduction

What shapes a species' geographic distributions? This question is central to the fields of ecology and evolution (Sutherland et al., 2013) and has yet no clear answer. Past theoretical and empirical research has identified several intrinsic and extrinsic factors contributing to setting range limits, with potentially complex interactions (Gaston, 2009; Roy et al., 2009; Sexton et al., 2009; Louthan et al., 2015; Connallon and Sgrò, 2018; Willi and Van Buskirk, 2019). Especially genetic drift has been under particular scrutiny in evolutionary research on the causes of species' range limits. Recent evolutionary models have explored how drift accumulating in small populations at range limits can constrain range expansion by negatively impacting population fitness (Peischl et al., 2013; Henry et al., 2015; reviewed in Willi, 2019) and constrain adaptation along environmental gradients (Polechová and Barton, 2015; Polechová, 2018). While these studies provide a strong framework to explore why species are limited in their distributions, empirical evidences of these processes at range limits are still scarce.

The classic "center–periphery' hypothesis", based on the assumption that the range of a species is a representation of its ecological niche (Hutchinson, 1957), states that lower population abundance at range limits results from a decline in habitat suitability (Hengeveld and Haeck, 1982; Brown, 1984). This hypothesis is supported by meta-analytic studies, reporting a general decline in population occurrence and density of individuals toward range limits (Pironon et al., 2017), in line with a strong overlap between range limits and niche limits (Cahill et al., 2014; Hargreaves et al., 2014; Lee-Yaw et al., 2016). Another evolutionary theory states that decline in population size results from serial demographic bottlenecks during range expansion (Wade and McCauley, 1988; Peter and Slatkin, 2013) supported by the strong relation between decline in genetic diversity and expansion distance (Pironon et al., 2017).

Evolutionary theory states that population with a history of small size, or which have undergone demographic bottlenecks are exposed to strong genetic drift opposing purifying selection (Wright, 1931). As a result, (mostly) recessive deleterious mutations accumulate in

small populations, negatively affecting fitness, *i.e.* mutational load (Wright, 1931; Kimura et al., 1963; Kirkpatrick and Jarne, 2000; Peischl and Excoffier, 2015). Simulation studies predict that mutational load accumulated through serial bottlenecks during fast range expansion could be strong enough to slow down or even halt range expansion, if recombination is low (Peischl et al., 2013, 2015; Peischl and Excoffier, 2015). With the general decline in habitat suitability toward and beyond range limits (Brown, 1984; Cahill et al., 2014; Hargreaves et al., 2014; Lee-Yaw et al., 2016), the expression of mutational load may also increase in range-edge populations due to higher exposure to environmental stress, as has been suggested for inbreeding depression (Reed et al., 2012). Strong genetic drift also opposes selection on beneficial alleles and erodes selection in small populations (Wright, 1931). Simulation studies have identified increased genetic drift as a main constraint to adaptation along environmental gradients at range limits (Polechová and Barton, 2015; Polechová, 2018).

Increased genetic drift could be a predominant factor shaping the range limits of many temperate species, often characterized by a history of recent range expansion (Hewitt, 2000, 2004) since last glacial maximum (LGM), and a general decline in population sizes and genetic diversity toward species range limits (Eckert et al., 2008; Sexton et al., 2009; Pironon et al., 2017). However empirical evidences of the role of drift at range limits are still scarce, leaving several open questions. Past empirical studies support the accumulation of mutational load toward range limits: Signatures of mutational load have been shown to increase with further distance from expansion core in several plant species (González-Martínez et al., 2017; Willi et al., 2018; Koski et al., 2019), and phenotypic studies performed in laboratory, greenhouse or common gardens hint toward increased expression of mutational load toward range limits (Bosshard et al., 2017; Willi et al., 2018; Koski et al., 2019). Whether the patterns of increased expression of mutational load are uphold in populations exposed to their natural environment, and whether this effect is strong enough to reduce population demographic rates at range limits remain to be tested. The environmental dependency of the expression of mutational load has

also rarely been observed in natural populations (Fenster and Galloway, 2000; Prill et al., 2014; Li et al., 2018) and has never been tested in the context of limits to range expansion. Furthermore, local adaptation has been shown to decline in populations at range limits, linked to lower census size (Vergeer and Kunin, 2013) or genetic diversity (Halbritter et al., 2015, Hämälä et al., 2018), but these patterns are yet to be tested in the context of drift accumulated at range limits through a history of fast range expansion or long-term isolation. A final aspect to explore is the strength and permeability of the limits to range expansion exerted by drift. In fact, in the short evolutionary time since LGM, several taxa have been able to extend their range over large ecological gradients (*e.g.* Skrede et al., 2006; Koch et al., 2006; Smickl et al., 2010), hinting toward a leakiness on the constraints imposed by drift.

Study system

I empirically assessed these open questions in the North American *Arabidopsis lyrata* subsp. *lyrata* (L.). This species is ideal to explore the role of mutational load and its environmental dependency in shaping range limits: the current distribution of *A. lyrata* is characterized by a history of fast post-glacial range expansion from two separate refugia (Willi and Määttänen, 2010; Griffin and Willi, 2014; Willi et al., 2018). In addition, while most populations are outcrossing, selfing population occur predominantly at range limits of each cluster (Griffin and Willi, 2014), also expected to lead to increased genetic drift (Pollak, 1987; Nordborg and Donelli, 1997) and mutation accumulation (Lynch et al., 1995; Schultz and Lynch, 1997). In line, increased genomic signatures of mutational load have been linked with longer range expansion distance, or long-term isolation at the rear-edge, further increased in selfing populations (Willi et al., 2018). Furthermore, previous distribution modelling suggest that current northern and southern range limits of the species are well defined by steep decline in habitat suitability, excluding dispersal limitation (Lee-Yaw et al., 2018), but also potentially exposing populations to increased environmental stress toward and beyond range limits.

Finally, clines in adaptation across the range of this species have also been identified, along its main niche defining variables (Paccard et al., 2014; Wos and Willi, 2015; Walden et al., 2020), presenting an ideal setup to explore the role of drift in adaption at range limits.

Genetic and environmental constraints shaping the range limits of Arabidopsis lyrata

In a first project, I explored the variation in expression of mutational load and the variation in adaptation in 20 population representing the whole range of *A. lyrata* (Fig. 1). The expression of mutational load is classically inferred from heterosis, the difference in performance between hybrid populations, with expected reduced homozygosity on recessive deleterious mutations, and their parental populations (dominance model of heterosis, Crow, 1987). I raised offspring of laboratory-generated within- (WPC) and between-population crosses (BPC) in a transplant experiment along a latitudinal gradient at five sites in the USA, representing the conditions within the range, as well as conditions at and beyond the northern and southern edges.

In **Chapter 1**, I addressed the question whether the magnitude of the expression of mutational load is dependent on the history of range expansion or rear edge isolation, and whether this process results from range expansion alone or from the increased occurrence of selfing populations at range limits. In **Chapter 2**, I then tested whether the expression of mutational load is dependent on environmental stress resulting from the exposure to climates different from those experienced at the site of origin of each population. I also tested specifically in range-edge populations, whether the expression of mutational load increases when transplanted in unsuitable conditions beyond their respective edges. In **Chapter 3,** I assessed whether range limits actually correspond with niche limits in *A. lyrata*, by testing if WPC population performance estimates and demographic rates declined in sites at and beyond range limits. I then tested the levels of local adaptation across this species range, and whether adaptation was reduced in range-edge populations, with low levels of genetic diversity.

In a second project, I focused on the difference in adaptation to cold climates between *A. lyrata* and its selfing sister species *A. arenicola.* Despite sharing a common post-glacial ancestor with *A. lyrata*, *A. arenicola* successfully colonized subarctic and arctic regions of North America (Fig. 1), providing a striking counter-example to the limitations on range expansion tested in Chapter 1, 2 and 3. I raised populations of both species, and their previously laboratory generated hybrids in a climate chamber experiment simulating stress linked to growth in cold climates over a whole life cycle. In **Chapter 4**, I tested whether *A. lyrata* and *A. arenicola* differ in their adaptation to cold temperatures and frost events, to understand their divergent distribution. Finally, in **Chapter 5**, I assessed if both *A. lyrata* and *A. arenicola* are separated by other reproductive barrier contributing to adaptive differentiation in addition to their parapatric distribution and the selfing mating system of *A. arenicola*. I therefore tested the levels of intrinsic reproductive isolation by comparing performance of interspecific hybrids and their parental populations, and further tested if hybridization disrupted the adaptation strategy to cold climates of both species.

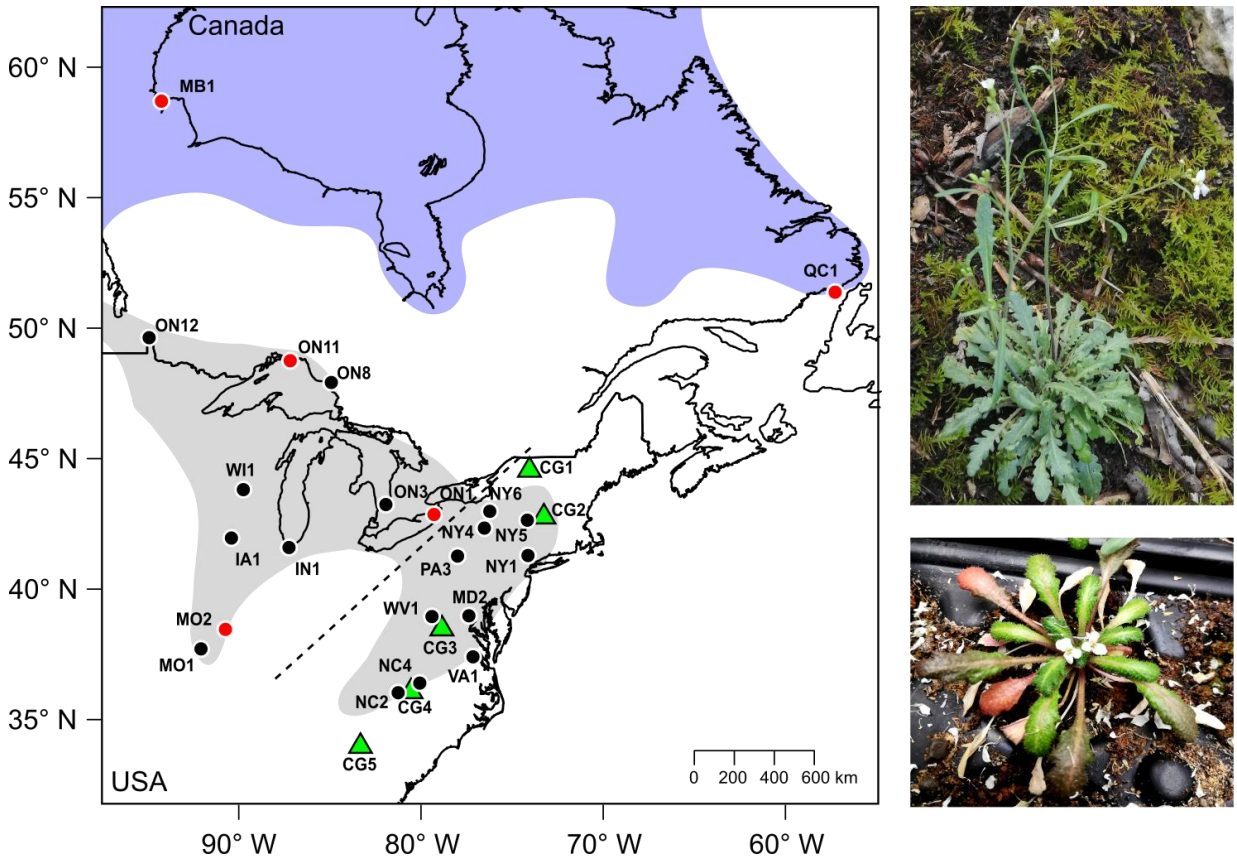

Figure 1: Distribution of *Arabidopsis lyrata* subsp. *lyrata* and *Arabidopsis arenicola* in eastern North America with the locations of the populations studied and the 5 common garden sites. Left: The grey shaded area represents the current North-American range of *A. lyrata*, and the blue shaded area the current range of *A. arenicola*. Circles filled in black or red represent outcrossing and selfing populations studied in this thesis. Population labels consist of the abbreviation for state (USA) or province (Canada) and a number (as in Willi et al., 2018). Green triangles represent the five common garden (CG) sites; numbers added to labels are in sequence of north to south. The dashed line is the split between eastern and western genetic clusters. **Right, top**: Flowering *A. lyrata* individual of the population NY6 in its natural localization; **bottom**: Flowering *A. arenicola* individual of the populations MB1 raised in a climate chamber in Basel. Switzerland (photographs taken by A. Perrier, 2017).

References

Bosshard, L., I. Dupanloup, O. Tenaillon, R. Bruggmann, M. Ackermann, S. Peischl, and L. Excoffier. 2017. Accumulation of deleterious mutations during bacterial range expansions. *Genetics* 207: 669–684.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. *The American Naturalist* 124: 255–279.

Cahill, A. E., M. E. Aiello-Lammens, M. C. Fisher-Reid, X. Hua, C. J. Karanewsky, H. Y. Ryu, et al. 2014. Causes of warm-edge range limits: systematic review, proximate factors and implications for climate change. *Journal of Biogeography* 41: 429–442.

Connallon T., and C. M. Sgrò. 2018. In search of a general theory of species' range evolution. *PLoS Biology* 16: e2006735.

Crow, J. F. 1987. Muller, Dobzhansky, and overdominance. *Journal of the History of Biology* 20: 351–380.

Eckert, C. G., K. E. Samis, and S. C. Lougheed. 2008. Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology* 17: 1170–1188.

Fenster, C. B., and L. F. Galloway. 2000. Population differentiation in an annual legume: genetic architecture. *Evolution* 54: 1157–1172.

Gaston, K. 2009. Geographic range limits of species. *Proceedings of the Royal Society B: Biological Sciences* 276: 1391–1393.

González-Martínez, S. C., K. Ridout, and J. R. Pannell. 2017. Range expansion compromises adaptive evolution in an outcrossing plant. *Current Biology* 27: 2544–2551.

Griffin, P. C., and Y. Willi. 2014. Evolutionary shifts to self-fertilization restricted to geographic range margins in North American *Arabidopsis lyrata*. *Ecology Letters* 17: 484–490.

Halbritter, A. H., S. Fior, I. Keller, R. Billeter, P. J. Edwards, R. Holderegger, et al. 2018. Trait differentiation and adaptation of plants along elevational gradients. *Journal of Evolutionary Biology* 31: 784–800.

Hämälä, T., T. M. Mattila, and O. Savolainen. 2018. Local adaptation and ecological differentiation under selection, migration, and drift in *Arabidopsis lyrata*. *Evolution* 72: 1373–1386.

Hargreaves, A. L., K. E. Samis, and C. G. Eckert. 2014. Are species' range limits simply niche limits writ large? A review of common garden experiments beyond the range. *The American Naturalist* 183: 157–173.

Hengeveld, R., and Haeck, J. 1982. The distribution of abundance. 1. Measurements. *Journal of Biogeography* 9: 303–316.

Henry, R. C., K. A. Barton, and J. M. J. Travis. 2015. Mutation accumulation and the formation of range limits. *Biology Letters* 11: 20140871–20140871.

Hewitt, G. M. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907–913.

Hewitt, G. M. 2004. Genetic consequences of climatic changes in the Quaternary. *Philosophical Transactions of the Royal Society B, Biological Sciences* 359: 183–195.

Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22: 415–427.

Kimura, M., T. Maruyama, and J. F. Crow. 1963. The mutation load in small populations. *Genetics* 48: 1303–1312.

Kirkpatrick, M., and P. Jarne. 2000. The effects of a bottleneck on inbreeding depression and the genetic load. *The American Naturalist* 155: 154–167.

Koch, M.A., C. Kiefer, D. Ehrich, J. Vogel, C. Brochmann, and K. Mummenhoff. 2006. Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). *Molecular Ecology* 15: 825–839.

Koski, M. H., N. C. Layman, C. J. Prior, J. W. Busch, and L. F. Galloway. 2019. Selfing ability and drift load evolve with range expansion. *Evolution Letters* 3: 500–512.

Lee-Yaw, J. A., H. M. Kharouba, M. Bontrager, C. Mahony, A. M. Csergő, A. M. Noreen, et al. 2016. A synthesis of common garden experiments and ecological niche models suggests that range limits are often niche limits. *Ecology Letters* 19: 710–722.

Lee-Yaw, J. A., M. Fracassetti, and Y. Willi. 2018. Environmental marginality and geographic range limits: a case study with *Arabidopsis lyrata* ssp. *lyrata*. *Ecography* 41: 622–634.

Li, Y., M. Stift, M. van Kleunen, and S. Bonser. 2018 Admixture increases performance of an invasive plant beyond first-generation heterosis. *Journal of Ecology* 106: 1595–1606.

Louthan, A. M., F. D. Daniel and A. L. Angert. 2015. Where and when do species interactions set range limits? *Trends in Ecology and Evolution* 30: 780–792.

Lynch, M., J. Conery, and R. Bürger. 1995. Mutational meltdowns in sexual populations. *Evolution* 49: 1067–1080.

Nordborg, M., and P. Donnelly. 1997. The coalescent process with selfing. *Genetics* 146: 1185–1195.

Paccard, A., A. Fruleux, and Y. Willi. 2014. Latitudinal trait variation and responses to drought in *Arabidopsis lyrata*. *Oecologia* 175: 577–587.

Peischl, S., and L. Excoffier. 2015. Expansion load: recessive mutations and the role of standing genetic variation. *Molecular Ecology* 24: 2084–2094.

Peischl, S., I. Dupanloup, M. Kirkpatrick, and L. Excoffier. 2013. On the accumulation of deleterious mutations during range expansions. *Molecular Ecology* 22: 5972–5982.

Peischl, S., M. Kirkpatrick, and L. Excoffier. 2015. Expansion load and the evolutionary dynamics of a species range. *The American Naturalist* 185: E81–E93.

Peter, B. M., and M. Slatkin. 2013. Detecting range expansions from genetic data. *Evolution* 67: 3274–3289.

Pironon, S., G. Papuga, J. Villellas, A. L. Angert, M. B. García, and J. D. Thompson. 2017. Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biological Reviews* 92: 1877–1909.

Polechová, J. 2018. Is the sky the limit? On the expansion threshold of a species range. *PLoS Biology* 16: e2005372.

Polechová, J., and N. H. Barton. 2015. Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences of the United States of America* 112: 6401–6406.

Pollak, E. 1987. On the theory of partially inbreeding finite populations. I. Partial selfing. *Genetics* 117:353–360.

Prill, N., J. M. Bullock, N. M. van Dam, and R. Leimu. 2014. Loss of heterosis and family-dependent inbreeding depressipost-on in plant performance and resistance against multiple herbivores under drought stress. *Journal of Ecology* 102: 1497–1505.

Reed, D. H., C. W. Fox, L. S. Enders, and T. N. Kristensen. 2012. Inbreeding–stress interactions: evolutionary and conservation consequences. *Annals of the New York Academy of Sciences* 1256: 33-48.

Roy, K., G. Hunt, D. Jablonski, A. Z. Krug, and J. W. Valentine. 2009. A macroevolutionary perspective on species range limits. *Proceedings of the Royal Society B: Biological Sciences* 276: 1485–1493.

Schmickl, R., M. H. Jørgensen, A. K. Brysting, and M. A. Koch. 2010. The evolutionary history of the *Arabidopsis lyrata* complex: a hybrid in the amphi-Beringian area closes a large distribution gap and builds up a genetic barrier. *BMC Evolutionary Biology* 10: 98.

Schultz, S. T., and M. Lynch. 1997. Mutation and extinction: the role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. *Evolution* 51: 1363–1371.

Sexton, J. P., P. J. McIntyre, A. L. Angert, and K. J. Rice. 2009. Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics* 40: 415–436.

Skrede, I., P. B. Eidesen, R. P. Portela, and C. Brochmann. 2006. Refugia, differentiation and postglacial migration in arctic-alpine Eurasia, exemplified by the mountain avens (*Dryas octopetala* L.). *Molecular Ecology* 15: 1827–1840.

Sutherland W. J., R. P. Freckleton, H. C. J. Godfray, S. R. Beissinger, T. Benton, D. D. Cameron, et al. 2013. Identification of 100 fundamental ecological questions. *Journal of Ecology* 101: 58–67.

Vergeer, P., and W. E. Kunin. 2013. Adaptation at range margins: common garden trials and the performance of *Arabidopsis lyrata* across its northwestern European range. *New Phytologyst* 197: 989–1001.

Wade, M. J., and D. E. McCauley. 1988. Extinction and recolonization: their effects on the genetic differentiation of local populations. *Evolution* 42: 995–1005.

Walden, N., K. Lucek, and Y. Willi. 2020. Lineage-specific adaptation to climate involves flowering time in North American Arabidopsis lyrata. *Molecular Ecology* 29: 1436–1451.

Willi, Y., and K. Määttänen. 2010. Evolutionary dynamics of mating system shifts in *Arabidopsis lyrata. Journal of Evolutionary Biology* 23: 2123–2131.

Willi Y., and J. Van Buskirk. 2019. A practical guide to the study of distribution limits. *The American Naturalist* 193: 773–785.

Willi, Y., M. Fracassetti, S. Zoller, and J. Van Buskirk. 2018. Accumulation of mutational load at the edges of a species range. *Molecular Biology and Evolution* 35: 781–791.

Willi, Y. 2019. The relevance of mutation load for species range limits. *American Journal of Botany* 106: 757–759.

Wos, G., and Y. Willi. 2015. Temperature-stress resistance and tolerance along a latitudinal cline in North American *Arabidopsis lyrata*. *PLoS ONE* 10: e0131808.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.

Chapter 1: Expressed mutational load increases toward the edge of a species' geographic range

Antoine Perrier*, Darío Sánchez-Castro, Yvonne Willi

Department of Environmental Sciences, University of Basel, Schönbeinstrasse 6, 4056 Basel, Switzerland.

*Corresponding author:

E-mail: antoine.perrier@unibas.ch

This article has been published in *Evolution* (15.06.2020).

Perrier, A., D. Sánchez-Castro, and Y. Willi. Y. 2020., Expressed mutational load increases toward the edge of a species' geographic range. *Evolution* 74: 1711–1723.

doi:10.1111/evo.14042

Abstract

There is no general explanation for why species have restricted geographic distributions. One hypothesis posits that range expansion or increasing scarcity of suitable habitat result in accumulation of mutational load due to enhanced genetic drift, which constrains population performance toward range limits and further expansion. We tested this hypothesis in the North American plant, *Arabidopsis lyrata*. We experimentally assessed mutational load by crossing plants of 20 populations from across the entire species range and by raising the offspring of within- and between-population crosses at five common garden sites within and beyond the range. Offspring performance was tracked over three growing seasons. The heterosis effect, depicting expressed mutational load, was increased in populations with heightened genomic estimates of load, longer expansion distance or long-term isolation, and a selfing mating system. The decline in performance of within-population crosses amounted to 80%. Mutation accumulation due to past range expansion and long-term isolation of populations in the area of range margins is therefore a strong determinant of population-mean performance, and the magnitude of effect may be sufficient to cause range limits.

Keywords: *Arabidopsis lyrata*, genetic drift, geographic species distribution, heterosis, mutational load, range expansion, range limit, small population size.

Introduction

What determines the limits of species' geographic distributions has been a long-standing question in biology, yet the more ultimate evolutionary causes are still not fully understood (Gaston 2009; Sutherland et al. 2013; Connallon and Sgrò 2018; Willi and Van Buskirk 2019). Ecological research has focused on limiting environmental factors and used the concept of the ecological niche of species to understand range limits (e.g., Hargreaves et al. 2014; Lee-Yaw et al. 2016). In contrast, evolutionary theory has focused on constrains in adapting to ecological gradients, for which few direct empirical tests exist to date (recent theory: e.g., Polechová and Barton 2015; Polechová 2018; older theory and empirical work reviewed in: Kawecki 2008; Gaston 2009; Sexton et al. 2009). Another evolutionary explanation for distribution limits is enhanced genetic drift and the accumulation of deleterious mutations toward the range edge, due to a history of small population size either produced by past range expansion or a scarcity of suitable habitat (Peischl et al. 2013; Peischl and Excoffier 2015; Henry et al. 2015; reviewed in Willi 2019). These theoretical studies have described conditions favoring mutational load in contributing to range limits, but few empirical estimates of mutational load across species distributions have been made and the fitness consequences of mutational load in nature are unknown. If the phenotypic effect of mutational load due to past expansion or habitat scarcity is considerable, it may constrain population persistence and establish a range limit by preventing further expansion (Peischl et al. 2015; Henry et al. 2015).

Populations at range edges may often have a history of small size, with the predicted consequence of heightened genetic drift that erodes genetic variation and opposes the effect of (mostly weak) selection (Wright 1931; Kimura et al. 1963). In long-term small populations, the consequence of drift opposing purifying selection is the accumulation of deleterious mutations, leading to a reduction in fitness called mutation(al) load (Kimura et al. 1963). Similar to stable small population size, demographic bottlenecks are also expected to enhance genetic drift, erode genetic variation (Nei et al. 1975), and heighten mutational load (Kirkpatrick and Jarne 2000). Recent theoretical work by Peischl and co-workers suggested that serial bottlenecks during rapid range

expansion lead to the accumulation of mutational load, in this context termed expansion load, which decreases population mean performance and slows down expansion or even halts expansion if recombination is low (Peischl et al. 2013, 2015; Peischl and Excoffier 2015). The increased frequency of recessive deleterious mutations contributes strongest to mutational load (Peischl and Excoffier 2015) and load can persist for thousands of generations (Peischl et al. 2013). Most notably, predictions of this general model apply in the absence of any environmental gradient. A different type of neutral model also predicted stable range margins due to mutation accumulation along a gradient of habitat quality. Henry et al. (2015) performed simulations along linear arrays of habitat patches of decreasing carrying capacity and found that the range limits retract to a stable point, before reaching the limit of habitat patches, due to mutation accumulation if both dispersal and population growth rate are small.

Empirical research suggests that a history of past range expansion is common in many taxa and that the habitat often deteriorates at range edges, both of which are associated with enhanced genetic drift. Quaternary ice ages caused retraction of the geographic distribution of many species into refugia, from which they have re-expanded, leaving many with distribution margins characterized by a history of recent range expansion and lowered effective population size (Hewitt 2000). Furthermore, several recent meta-studies confirmed the general trend for enhanced habitat deterioration and habitat isolation toward and around the geographic range limits and lower effective population sizes. A meta-study on transplant experiments with sites beyond the range edge revealed significant performance declines beyond range edges in about 80% of studies (Hargreaves et al. 2014), which was paralleled by a decline in habitat suitability deduced by niche modelling (Lee-Yaw et al. 2016). Furthermore, the density of individuals and populations of species were found to generally decline toward the range edge (Pironon et al. 2017). Other meta-level studies show that populations at range edges have reduced within-population genetic marker variation and are genetically more differentiated, documenting the enhanced action of genetic drift (Eckert et al. 2008; Sexton et al. 2009; Pironon et al. 2017). In the context of range margins, the evolution of mating

system shifts received additional attention. In hermaphroditic organisms, the incidence of self-fertilization increases toward range edges due to a history of mate limitation and the lowering of inbreeding load (Pujol et al. 2009; Griffin and Willi 2014; Matos et al. 2015). One consequence of a shift to selfing is increased genetic drift (Pollak 1987; Nordborg and Donelli 1997) and mutation accumulation (Lynch et al. 1995a; Schultz and Lynch 1997). Indeed, estimates of effective population sizes are typically lower in selfing compared to outcrossing taxa (Ingvarsson 2007; Hartfield et al. 2017).

The accumulation of deleterious mutations during range expansion has been studied best in humans. Populations with a longer history of expansion out-of-Africa, European Americans, had higher proportions of non-synonymous to all single-nucleotide polymorphisms (SNPs) compared to African Americans (Lohmueller et al. 2008). Similarly, an increased frequency of predicted deleterious mutations was observed in out-of-Africa populations compared to humans from southern Africa (Henn et al. 2016). In plants, increased genomic estimates of mutational load with range expansion have been described in at least three species (González-Martínez et al. 2017; Willi et al. 2018; Koski et al. 2019). However, empirical evidence of the link between expansion history and performance decline are scarce. In an experimental-evolution study with bacteria, lines with high mutation rates evolved to have reduced growth under range expansion over 1650 generations compared to their ancestral lines, suggesting accumulation of mutational load (Bosshard et al. 2017). Increased genomic estimates of mutational load toward the distribution edge were associated with reduced performance assessed in a common garden in the species *Arabidopsis lyrata* (Willi et al. 2018). In *Campanula americana*, populations further away from a putative glacial refugium in the southern Appalachians expressed increased mutational load in the greenhouse (Koski et al. 2019). However, these studies have not tested the contribution of mutational load to reducing population performance and demographic rates under natural conditions or across the distribution of a species. Moreover, the life stage at which mutational load is expressed is not known (Hansen and Price 1999).

Based on theory, we expect that the cumulative effects of numerous deleterious mutations each of small effect become most detectable at later life stages (Husband and Schemske 1996).

In this study, we estimated the expression of mutational load of natural populations of *A. lyrata* subsp. *lyrata* (L.) from across the species range in common gardens within and beyond the distribution range. The species is ideal for investigating genetic causes of range limits because niche modelling has shown that the species is not dispersal-limited in the south and north, indicating that range limits reflect niche limits (Lee-Yaw et al. 2018). Furthermore, previous population genomics studies demonstrate a history of fast post-glacial range expansion from two distinct refugia, resulting in two genetically distinct clusters with a small contact zone at Lake Erie (Willi and Määttänen 2010; Griffin and Willi 2014; Willi et al. 2018). Distance of expansion or rear-edge distance to the glacial refugia was positively associated with genomic estimates of mutational load, indicating that both past range expansion and long-term isolation at the south-western range edge left a signature of mutation accumulation. The highest genomic estimates of mutational load were found in selfing populations, which in this species are restricted to areas at or close to the edge of the range (Griffin and Willi 2014; Willi et al. 2018). To estimate expressed mutational load, we used the proxy of heterosis, *i.e.* the increase in fitness of between-population crosses compared to within-population crosses due to increased heterozygosity of recessive deleterious mutations (dominance model of heterosis, Crow 1987). We tested the following predictions: (i) Mutational load expressed in the field is tightly correlated with mutational load estimated on a genomic level. (ii) As with genomic estimates of load, expressed mutational load is correlated with post-glacial expansion distance or long-term isolation at the rear edge and with mating system. (iii) Expressed mutational load is based on weakly deleterious mutations, whose cumulative effect is greatest at late life stages.

Material and Methods

Plant material and the crossing of plants

Twenty populations of *A. lyrata* subsp. *lyrata* were selected to represent the whole range of distribution of the species (Fig. 1, Table S1). They represented: the two genetic clusters of the species in North America; different histories during and since the last glaciation cycle, either one of being close to glacial core distribution or one of expansion or rear-edge isolation; different mating systems, either being predominantly outcrossing or predominantly selfing (Griffin and Willi 2014). Seeds of different maternal plants per population were collected between 2007 and 2014 over an area of about 450 m^2 in each population. Seeds had been stored in separate bags per maternal plant at 4 °C under dry, dark conditions.

We raised 26 plants per population in growth chambers, one per field-collected maternal plant and that we assumed were unrelated, for the production of within- and between-population crosses. Three seeds per maternal plant were initially sown in individual pots filled with a 1:1 mixture of sand and peat. Pots were watered to saturation and seeds stratified for 12 days at 4 °C in the dark. Pots were then transferred to growth chambers (CLF Plant Climatics, Wertingen, Germany) with the following conditions to promote germination: 8h of light at 100 μmol m^{-2} s^{-1} and 20 °C, 16h of dark at 20 °C. Germinated plants were thinned to one per pot, 36 days after sowing. To promote growth and flowering, day length and light intensity were increased every three days by 1h and 20 μmol m^{-2} s^{-1}, respectively, over a period of 25 days, and day temperature was increased by 2 °C. The final conditions were kept until the end of the crossing experiment: 16h of light at 240 μmol m^{-2} s^{-1} and 22 °C, 8h of dark at 20 °C. After 25 days, when the first individuals started to bolt, all pots were transferred to a greenhouse with similar conditions as in the growth chambers to perform the crosses (all growth conditions detailed in Table S2).

Of the 20 populations, 18 were considered target populations, and two served as pollen donors for between-population crosses. The latter two populations were located in the center of distribution of the two ancestral clusters and had high genomic diversity (NY1 for the eastern cluster, IA1 for the

western cluster). For each of the 18 target populations, 12 of the 26 individuals were randomly chosen as being "mothers" (pollen recipients) and 12 individuals as "fathers" (pollen donors); the remaining plants were used as backups. The 12 mother plants of a target population were crossed with pollen from a randomly chosen father plant of the same population (WPC) and from a randomly chosen plant of the partner population (BPC); crosses were non-reciprocal. WPC crosses were also performed for the two partner populations (list of families and cross combinations in Table S3). We made hand-pollinations at the bud stage to exclude unwanted cross- and spontaneous self-pollination. Flower buds of the mother plant were opened with tweezers, the immature anthers were removed, and mature anthers of a father plant gently rubbed over the stigma. Pollen contamination was avoided by sterilizing the tweezers after each contact with a flower, and placing each pollinated plant into an insect-proof growth chamber until fruit elongation began (3-5 days). Each cross combination was repeated to obtain a sufficient number of seeds for the outdoor common gardens (at least six siliques or 60 healthy-looking seeds). Cross combinations were changed if no siliques or no viable seeds could be obtained. We collected mature siliques and left them to dry for two weeks at ambient temperature in the dark. Afterwards, they were stored at 4 °C, under dry and dark conditions.

Raising of plants in common gardens

Expressed mutational load, the heterosis effect in F1 individuals, was assessed at five common garden sites along a 1400 km latitudinal gradient in the eastern USA (Fig. 1). One site was in the center of the range of *A. lyrata*, in Harrisonburg, VA, two sites were close to the southern and northern borders of the range, in Winsten-Salem, NC, and Williamstown, MA, respectively, and two sites were beyond the southern and northern range edge, in Athens, GA, and the Adirondacks, NY, respectively (Table S4). In the analyses presented here, sites were treated as a level of replication for estimating mutational load. Our main goals were to analyze the relationships between expressed mutational load and a genomic estimate of mutational load (prediction i) and between expressed mutational load and past range dynamics or mating system (prediction ii). The common garden study started in fall 2017

and used the same protocol for each garden, with slight deviations due to local facilities. We sowed seeds from all successful cross combinations that had more than 15 healthy seeds in each garden. If a cross combination failed to produce enough viable seeds, we added an additional cross combination from the same population with a sufficient number of seeds. In total, 401 cross combinations contributed to the field experiment (Table S3). Per cross combination and common garden, three pots were filled each with two seeds (in some cases only one seed was available). Pots were randomly positioned across thirteen 38-cell propagation trays within each of three blocks per common garden. Across the five gardens, a total of 12,933 seeds were sown. In all common gardens, we used the same substrate mixture of washed river sand and peat (1:1.5 sand:peat). Sowing was done in early fall to early winter and started at the northernmost site. To prevent seeds from being washed away by heavy rainfall, germination was carried out under a ventilated greenhouse or temporary tent for 17-19 days until the peak of germination was reached. The trays were then exposed to natural conditions for the rest of the experiment. During fall 2017, the trays were regularly watered during periods of no rain, to ensure a constant moisture of the substrate, until snow fell or the first night frosts occurred. We weeded the pots manually, and seedlings were thinned starting 11 weeks after sowing to keep only one individual per pot. Herbivory by grazing was prevented by a fence, and organic slug repellent was used in the beginning of spring, after snowmelt. No further interventions were made until the end of the experiment in summer 2019 (2018 for Harrisonburg because the garden was needed for another experiment).

We measured performance on the level of the individual pot/plant. Day of germination, when a seedling had two fully open cotyledons, was checked three times a week until the peak of germination was over (4-5 weeks after sowing) and then once a week until the first thinning. Germination was again checked in spring 2018. Death of seedlings was recorded at the same time as germination was checked, and later, mortality was checked once a week unless there was a snow cover. We scored the day of first flower opening three days a week, starting when bolting was observed in 2018. Day of germination, death, and flowering were corrected by the mid-time between

previous checking and actual observation. Reproductive output was estimated in 2018 and 2019 by counting the number of fruits, pedicels (flowers that did not develop into a fruit), open flowers, and flower buds on all inflorescences. Female reproductive output of each individual was the total number of fruits and potential additional fruits that could have formed from buds and open flowers: fruits + ((flowers + buds) × (fruits / (fruits + pedicels))). We assessed reproductive output several weeks after peak flowering: in 2018 ~ 9 weeks after opening of the first flowers within each common garden, and in 2019 ~ 5 weeks after first flowering, estimated from flowering dates of the previous year.

To assess the contribution of the seed bank to population growth, we carried out a seed survival experiment over the winter of 2018/19. One hundred healthy seeds of five to twelve mother plants from each WPC and BPC cross combination were pooled on the level of the population and cross type, and packed in groups of 10 seeds in 10 separate bags made out of micro-perforated fabric (nonwoven polypropylene-felt, 40 g/m^2) that allowed the penetration of air and moisture. Two bags of each pool were placed in each of the five common gardens in October 2018 on freshly weeded and homogenized soil next to the pots to expose them to natural conditions, and they were retrieved in late spring 2019. Each pool was then visually screened to discriminate between seedlings and seeds. We then judged survival by first stratifying seeds on filter paper disks soaked with 1.5 ml of 0.05% gibberellic acid in petri-dishes (10 days, 4 °C, no light). Germination was assessed under similar conditions as detailed for the crossing experiment and scored over 20 days. Seed survival over winter was then estimated for each bag as: (germinated seedlings + germinated seedlings with gibberellic acid)/10.

Statistical analysis

We analyzed two measures of performance, using pot as the level of replication. *Multiplicative performance I* was the fraction of seeds that germinated multiplied by the total reproductive output in year 2 plus in year 3, and *multiplicative performance II* was germination multiplied by the number of fruits only. Components of these overall performance estimates were analyzed separately and are

described in Table S5; these analyses were used to identify the life stages most impacted by mutational load (prediction iii). Finally, to assess how mutational load affected demography, we estimated population growth rates for all WPC (20) and BPC combinations (18) in each common garden by constructing stage-classified matrices (Caswell 2001), based on population mean data of each common garden. The matrices were composed of three stages: 1–healthy seeds, 2–individuals capable of reproducing in spring of year 2 (2018), 3–individuals capable of reproducing in spring of year 3 (2019), with a projection interval set to one year for each stage. The exact parametrization of the matrices is described in Fig. S6. For each combination of population, cross type and common garden, we calculated λ, the finite rate of increase in one time-step (Caswell 2001).

Preliminary analyses on the level of the pot/plant (described below) revealed that the effect of cross-type was highly significant for *multiplicative performance I* and *II*, and therefore we present the analyses and results on heterosis first. Population-level heterosis was calculated as the increase in performance due to between-population crossing relative to within-population crossing, as follows: $(W_{BPC} - W_{WPC})/W_{WPC}$. W_{WPC} and W_{BPC} were calculated for each population in each common garden based on family means. In the case of W_{WPC}, the final value was an average of the two types of WPC, of the target and the pollen-donor population. In case either W_{BPC} or W_{WPC} was equal to zero for a specific cross combination in a specific common garden, we chose to replace this value by the smallest non-zero value observed within cross type (12 cases for *survival summer year 2*, three cases for *survival winter year 2*). Heterosis estimates were $\log_{10}$-transformed (after making all values positive by adding +1), and tested by hierarchical mixed-effects models using restricted maximum likelihood with the packages lme4 (Bates et al. 2015) and LmerTest (Kuznetsova et al. 2017; model parametrization in Appendix S7A) in *R* (R Core Team 2019). Fixed effects were either the genomic estimate of mutational load, or the recent range-dynamics history of a population and mating system. The genomic estimate of mutational load (hereafter genomic load) was the ratio of non-synonymous polymorphic sites to synonymous polymorphic sites, adjusted for their mean derived allele frequency relative to *A. thaliana*, $P_n f_n / P_s f_s$ (Willi et al, 2018). The range-dynamics history of a population was

its $\log_{10}$-transformed distance to distribution cores. Cores were glacial refugia that gave rise to range expansion, identified by means of the map-projection of a population phylogeny. More precisely, cores were defined as the location of the ancestral node from which a first ancestral population appeared that was located in an area covered by ice during the last glacial maximum (Willi et al. 2018). For younger populations, distance to core was calculated as the sum of great circle distances [km] from the location of the extant population back along the map-projected phylogeny to the core and reflected the expansion distance. Populations that had diverged earlier were considered rear-edge relative to the core sites. For these, the direct great-circle distance to the ancestral core population was calculated. The two Missouri populations, although part of a separate third cluster, were considered as being part of the western cluster due to proximity and a closer shared history of admixture (Willi et al. 2018). As a proxy for mating system we used the population inbreeding coefficient, F_{IS} (Griffin and Willi 2014). Continuous fixed effects were mean-centered before running each analysis. The random part of models included the crossed effects of maternal population and common garden.

Further analyses validated the use of heterosis as a proxy of expressed mutational load. First, we verified consistency in results between population-level analyses and pot/plant-level analyses. Dependent variables were the two measures of *multiplicative performance* and the separate performance components. Fixed effects were cross type, genomic load, and their interaction. Preliminary analyses showed that the best random structure was: maternal plant nested within maternal population and maternal population, for which intercepts and slopes of cross type were estimated, and block nested within common garden, and common garden. The two *multiplicative performance* variables were 0 inflated, which suggested the modelling of two processes, a Gaussian process (for $\log_{10}$-transformed performance values > 0), and a logistic process (modelling the probability of 1, assigned to performance values > 0). Analyses were performed in a Bayesian framework, with the package MCMCglmm (Hadfield 2010, 2019) on 10 parallel chains (model and prior parametrization detailed in Appendix S7B). Analyses on variables depicting life stage

components made use of restricted maximum likelihood (model parametrization detailed in Appendix S7C). Next, analyses were repeated on (log_{10}-transformed) population means for each cross type and each common garden, by use of restricted maximum likelihood. Fixed effects were cross type, mean-centred genomic load and the interaction between the two. Crossed random effects were maternal population and common garden. To validate if heterosis is the result mainly of dominance and load due to fully recessive deleterious mutations, we tested for a relationship between W_{WPC} and genomic load, and W_{BPC} and genomic load similar to above.

Results

Overall, 64.2% of all seeds germinated (Table S8). Plants had high survival rates at each life stage (61.8-99.6%), except for *survival summer year 2*, which was the most critical life stage (29.8%), with most deaths happening after reproducing. Surprisingly, despite high *survival to flowering year 2* (99.6%), only 60.2% initiated flowering, while 95% of plants that survived to year 3 initiated flowering (data not shown). Finally, individuals that flowered in year 2 produced on average 135 flowers, with values ranging from 1 to 2607 flowers, with an average fertilization rate of 67.2%. Heterosis in *multiplicative performance I* and *II* up to year 3, assessed per population and common garden, ranged from -0.96 to 23.50 (mean: 1.88) and from -0.92 to 30.23 (mean: 2.65), respectively (Table S8). Finally, heterosis estimated on λ was between -0.53 and 7.29 (mean: 0.73; Table S8).

Expressed mutational load, here estimated by heterosis in *multiplicative performance I* and *II* up to year 3, was positively related with the genomic estimate of mutational load (Table 1, Fig. 2; results on *MP I* and *II* to year 2 reported for comparison). The model-predicted increase between the population with the lowest and that with the highest genomic load was up to 5.6-fold (Table S9). Also, heterosis in *multiplicative performance I* and *II* up to year 3 significantly increased with the distance between the site of origin of a population and the glacial core distribution (Table 1, Figs. 2, 3). The predicted maximal increase in heterosis between the closest and farthest population from the

glacial cores was up to 3.4-fold (Table S9). Analyses on outcrossing populations confirmed the positive effect of distance to core on heterosis (Table S10). Heterosis was higher in selfing populations for *multiplicative performance I* and *II* up to year 2 but not to year 3 (Table 1). The predicted maximal increase in heterosis between the most outcrossed and the most inbred population was 3.3-fold for *multiplicative performance II* to year 2 (Table S9). The intercept of the linear models was significant for heterosis in *multiplicative performance I* and *II* to year 3, indicating that the average population suffered from mutational load (Table 1).

Heterosis associated with genomic load was significant relatively early in life (Table 1). *Survival fall year 1* was the second variable after germination in the life stage analyses and for this variable a significant positive relationship between heterosis and genomic load was found. Further variables with a significant positive relationship between heterosis and genomic load were: *bolting*, *reproductive output* and *number of fruits* produced, all in year 2. *Germination* was the first life stage for which the relationship between heterosis and distance to core was significant (Table 1). Further variables with a significant positive relationship between heterosis and distance to core were *reproductive output* and *number of fruits* produced in year 2. Results were similar when analysis was restricted to outcrossing populations (Table S10). Heterosis in *survival fall year 1, survival winter year 2*, and *bolting* were significantly positively related with F_{IS} (Table 1). Finally, heterosis for λ was positive and significant for genomic load, and as a trend for distance to core, and for F_{IS} (Table 1, Fig. 2). The model-predicted increase between the population with the lowest genomic load to the population with the highest genomic load was 1.3-fold (Table S9). For F_{IS}, the model-predicted increase between the most outbred to the most inbred population was 1.7-fold (Table S9).

Analyses similar to those presented above were performed on the level of individual pots/plants, with cross-type in the fixed effects part of the model (Table S11). Hierarchical mixed-effects model analyses revealed a significant effect of cross type on *multiplicative performance I* and *II* to year 3 in both the log-normal process and the logistic process (Table S11A; results on *MP I* and *II* to year 2 reported for comparison). Between-population crosses (BPC) had higher performance

than within-population crosses, supporting a general heterosis effect. No direct effect of genomic load on *multiplicative performances* was observed. However, the cross type-by-genomic load interaction was significantly positive in the log-normal aspect of both *multiplicative performance* estimates; the performance of BPC declined less with genomic load than the performance of WPC. Similarly, when averaging both *multiplicative performance* estimates on the level of population for each cross type and common garden (Table S12), BPC performed significantly better than WPC. Furthermore, both *multiplicative performance* estimates were negatively related with genomic load, while again the cross type (BPC)-by-genomic load interaction had a significant positive effect on *multiplicative performance I* (marginally significant for *multiplicative performance II*). These results indicated that the relationship between performance and genomic load was more negative for WPC than BPC. Also analyzing both cross types separately confirmed the negative relationship between *multiplicative performance I* and *II* of WPC and genomic load, while no significant relationship was found for BPC (Table S13, Fig. 4). The predicted decline of WPC performance had a maximum value of 80.3% for *multiplicative performance I* to year 3(Table S9).

Discussion

Recent evolutionary theory proposes that the neutral process of genetic drift can contribute to slowing further range expansion in a species or cause stable range edges due to the accumulation of mutational load (reviewed in Willi 2019). Here we showed experimentally that both leading and rear edge populations suffered from the increased expression of mutational load – estimated by heterosis based on life-time performance, demographic rates, and performance at individual life stages. The expression of mutational load was also higher in selfing populations predominantly located at the distribution edge, aggravating the negative effect of load at range edges. Overall, this study provides empirical support for an important role of mutational load in range limits.

The expression of mutational load increased by a factor of 3.4 with distance along the expansion route or distance from the historic core of distribution toward the distribution edges of *A. lyrata*. The decline in population mean multiplicative performance of within-population crosses due to increasing genomic load was up to 80%. These results constitute some of the first *in-situ* evidence on the expression of mutational load toward range limits, and support predictions from simulation studies (Peischl et al. 2013; Peischl and Excoffier 2015), genomic data (Willi et al. 2018), or similar phenotypic data from the greenhouse or garden (Willi et al. 2018; Koski et al. 2019). The strong link between expressed mutational load, mutational load estimated with sequence data, and range position observed in our study system sheds light on the processes shaping range limits (reviewed in Willi 2019). Further colonization by leading edge populations, already suffering from high levels of load, may be impeded by additional accumulation of mutational load, reducing performance below critical thresholds necessary to maintain persisting populations. Similarly, at the rear edge, population isolation and low effective population sizes may lead to mutational melt-down (Lynch et al. 1995b), such that rear-edges are unstable over the long term and in a state of gradual retraction.

We found that the most inbred populations, the three predominantly selfing populations located at the northern, eastern, and southern edges of the western cluster, expressed even higher levels of load than outcrossing populations, with a predicted 3.3-fold maximal increase in heterosis based on multiplicative performance (up to the second year), and 1.3-fold increase in heterosis based on demographic rates. A similar result was found earlier on a different set of selfing populations of *A. lyrata* (Willi 2013). Higher levels of mutational load in selfing populations is expected due to their generally lower effective population size combined with increased exposure to genetic drift (Lynch et al. 1995a). Indeed, genomic signatures of mutational load are increased in several other selfing taxa (reviewed in Wright et al. 2013; Laenen et al. 2018). Theoretical and empirical studies predict higher rates of selfing toward range limits (reviewed in Pannell 2015), as observed in *A. lyrata* (Griffin and Willi 2014). This overrepresentation of selfing populations at range edges could lead to a biased estimation of the effect of expansion on mutational load, but our conclusions are not affected by this

because the statistical models accounted for mating system. This was also confirmed by analysis of outcrossing populations only, which produced similar effect sizes for distance to core on heterosis. The most important insight, however, is that selfing populations may often bear a double load, one from the long expansion history and one from selfing. Both are likely to increase extinction risk (Goldberg et al. 2010) and be effective in causing range limits (Peischl et al. 2015).

As predicted, our results generally supported the expectation that the correlation between load and either a genomic estimate of load or distance to core strengthened over the life cycle of the plants. In an early phase of the life cycle, survival shortly after germination showed heightened heterosis with genomic load, and germination showed heightened heterosis with distance to core. But effect sizes were weaker than those found for later life stages (Table 1). Heterosis linked to genomic load or distance to core was found consistently for several performance variables of the first reproductive period (bolting, reproductive output, and number of fruits). Finally, the strongest associations between heterosis and either genomic load or distance to core occurred in the multiplicative performance estimates. These results agree with the prediction that expression of load is due to deleterious mutations with cumulative effect over an organism's life (Husband and Schemske 1996). More and more genes contribute to performance over the course of the life of an organism, so the number of genes potentially experiencing load also increases, and this should produce a cumulative effect. Other empirical support for this model comes from studies assessing inbreeding depression in long lived perennials (e.g., Koelewijn et al. 1999; Griffin et al. 2019). Another prediction, according to theory, is that the magnitude of genetic drift determines the effect sizes of mutations that become targets of neutral evolution and are freed from purifying selection (Kimura et al. 1963). Here our results suggest that drift associated with past range dynamics must have been very strong, allowing mutations with significant phenotypic effect to accumulate in the presumably fewer genes relevant early in life (e.g., already at the time of germination).

For selfing populations, the pattern of expression of mutational load through the life cycle is probably similar to that in outcrossing populations. Survival shortly after germination showed

heightened heterosis in populations with a selfing mating system. Later life stages with significantly increased heterosis were bolting during the first reproductive season and the survival during the second winter. In a previous study including five other selfing or mixed-mating populations of *A. lyrata*, Willi (2013) reported heightened heterosis associated with selfing only in reproductive output in the third year, but not in earlier life stages. However, another study focusing only on early life stages reported lower performance of within-population crosses for germination in selfing *A. lyrata* (Joschinski et al. 2015). Overall, it seems that also in selfing populations, the magnitude of the expression of load increases over the lifetime of a plant, and that early life phases can already be affected.

Our results suggest that heterosis accurately reflects the fitness effect of mutational load. Just as the phenotypic comparison between in- and outbred individuals can accurately estimate inbreeding depression (Keller and Waller 2002), heterosis can indicate the expression of mutational load *in vivo* and, with an appropriate rearing design, *in situ*. One advantage of this approach is that other confounding effects can be excluded. For example, by using between-population crosses as the reference for performance, we control for the potential influence of population-specific local adaptation of within-population crosses. However, this method depends on two important assumptions: that heterosis is affected only by dominance (and not overdominance) and that load is primarily due to fully recessive deleterious mutations (Oakley et al. 2015; Peischl and Excoffier 2015). We verified both assumptions. The fact that performance of between-population crosses did not increase with genomic estimates of load indicates that overdominance was not important. Likewise, the fact that performance of BPC did not decline with genomic load suggests that partially recessive deleterious mutations do not contribute appreciably to mutational load. A previous common garden study with *A. lyrata* found that part of the load was caused by partially recessive mutations (Willi et al. 2018). A further challenge is that heterosis can be affected by a performance decay due to outcrossing with distantly-related individuals, due to hybrid breakdown or disruption of coadapted gene complexes. In fact, the negative estimates of heterosis in our study, which occurred in nearly a

fourth of the BPC, may reflect genetic incompatibilities such as the Dobzhansky-Muller type (Lynch 1991; reviewed in Oakley et al. 2015). Peripheral populations were generally genetically more isolated, so we assume that outbreeding depression was stronger for these populations. If this is true, our estimates of mutational load for range-edge populations would be slight underestimates, and the increasing load with post-glacial expansion distance would be even greater than reported here.

Our findings clearly show that populations with the longest expansion history suffer most from the expression of mutational load. Populations with the highest genomic signatures of load, located at both leading and rear-edges of the distribution, suffer from the expression of load to the extent that it impairs their demographic rates. The accumulation of mutational load is therefore likely to be involved in shaping range limits by impeding further expansion at the leading edge and causing retraction at the rear edge. The discovery that population history impacts population persistence at range edges argues for the integration of evolutionary history into biodiversity conservation management (Hoffmann et al. 2015). These processes are also important in the context of climate change: strong mutational load at range edges could impair expansion into newly available habitats while rear-edge populations would suffer from increasing isolation due to habitat fragmentation, mutation accumulation, and eventual extinction. Genetic drift at range margins is predicted to limit adaptation and expansion into empty habitat (Polechová and Barton 2015; Polechová 2018). Our results imply that models of range limits along environmental gradients should integrate increasing drift and mutation accumulation toward range edges. This will produce deeper insights in the relative importance of factors contributing to maladaptation, range limits, and responses to climate change.

Acknowledgements: This work was supported by the Swiss National Science Foundation (31003 A_166322). We are grateful to Celia Evans (Paul Smith's College, Paul Smith, NY), Joan Edwards (Williams College, Williamstown, MA), Heather Peckham Griscom (James Madison University, Harrisonburg, VA), William K. Smith (Wake Forest University, Winston-Salem, NC) and Rodney Mauricio (University of Georgia, Athens, GA) for logistical support in the US. For field assistance we thank Mary Anderson, Michael Boyd, Bennet Coe, Scott Cory, Rachel Hillyer, Andrew Jones, Deidre Keating, Larry Kummer, David Lampman, Anastasia Levie-Sprick, Blake Macko, Shannon Malisson, Kathryn McGee, Althea Neighbors, Debra Rogers-Gillig, Caleb Rose, Amber Scarabaggio, Anna Shutley, Caroline Vath and Audrey Werner. For assistance with seed counts we thank Olivier Bachmann, Markus Funk and Susanna Riedl. Collection permits were provided by the Clinton County Conservation Board, Cornell University, Fort Leonard Wood Army Base, Iowa Department of Natural Resources, Missouri Department of Conservation, New York State Office of Parks, Ontario Parks, Palisades Interstate Park Commission, Rock Island Lodge, United States National Park Service, Virginia Department of Conservation and Recreation and the Wisconsin Department of Natural Resources.

Data accessibility statement: All data is stored in Dryad (https://doi.org/10.5061/dryad.cc2fqz642)

References

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67:1–48.

Bosshard, L., I. Dupanloup, O. Tenaillon, R. Bruggmann, M. Ackermann, S. Peischl, and L. Excoffier. 2017. Accumulation of deleterious mutations during bacterial range expansions. Genetics 207:669–684.

Caswell, H. 2001. Matrix population models. 2nd ed. Sinauer Associates, Sunderland, MA, USA.

Connallon T., and C. M. Sgrò. 2018. In search of a general theory of species' range evolution. PLoS. Biol. 16:e2006735.

Crow, J. F. 1987. Muller, Dobzhansky, and overdominance. J. Hist. Biol. 20: 351–380.

Eckert, C. G., K. E. Samis, and S. C. Lougheed. 2008. Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. Mol. Ecol. 17:1170–1188.

Gaston, K. 2009. Geographic range limits of species. Proc. R. Soc. B Biol. Sci. 276:1391–1393.

Goldberg, E. E., J. R. Kohn, R. Lande, K. A. Robertson, S. A. Smith, and B. Igić. 2010. Species selection maintains self-incompatibility. Science 330:493–495.

González-Martínez, S. C., K. Ridout, and J. R. Pannell. 2017. Range expansion compromises adaptive evolution in an outcrossing plant. Curr. Biol. 27:2544–2551.

Griffin, P. C., and Y. Willi. 2014. Evolutionary shifts to self-fertilisation restricted to geographic range margins in North American *Arabidopsis lyrata*. Ecol. Lett. 17:484–490.

Griffin, R. A., B. M. Potts, R. E. Vaillancourt, and J. C. Bell. 2019. Life cycle expression of inbreeding depression in *Eucalyptus regnans* and inter-generational stability of its mixed mating system, Ann. Bot. 124:179–187

Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J. Stat. Softw. 33:1–22.

Hadfield, J. D. 2019. MCMCglmm course notes. Available at: https://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf

Hansen, T. F., and D. K. Price. 1999. Age- and sex-distribution of the mutation load. Genetica 106:251

Hargreaves, A. L., K. E. Samis, and C. G. Eckert. 2014. Are species' range limits simply niche limits writ large? A review of common garden experiments beyond the range. Am. Nat. 183:157–173.

Hartfield, M., T. Bataillon, and S. Glémin. 2017. The evolutionary interplay between adaptation and self-fertilization. Trends Genet. 33:420–431.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

Henn, B. M., L. R. Botigué, S. Peischl, I. Dupanloup, M. Lipatov, B. K. Maples, A. R. Martin, S. Musharoff, H. Cann, and M. P. Snyder. 2016. Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. Proc. Natl. Acad. Sci. USA. 113:E440–E449.

Henry, R. C., K. A. Barton, and J. M. J. Travis. 2015. Mutation accumulation and the formation of range limits. Biol. Lett. 11:20140871–20140871.

Hoffmann, A., P. Griffin, S. Dillon, R. Catullo, R. Rane, M. Byrne, R. Jordan, J. Oakeshott, A. Weeks, L. Joseph, P. Lockhart, J. Borevitz, and C. Sgrò. 2015. A framework for incorporating

evolutionary genomics into biodiversity conservation and management. Clim. Change Responses. 2:1–23.

Husband, B. C., and D. W. Schemske. 1996. Evolution of the magnitude and timing of inbreeding depression in plants. Evolution 50:54–70.

Ingvarsson, P. 2007. A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. Evolution 56:2368–2373.

Joschinski, J., M. van Kleunen, and M. Stift. 2015. Costs associated with the evolution of selfing in North American populations of *Arabidopsis lyrata*? Evol. Ecol. 29: 749–764.

Kawecki, T. J. 2008. Adaptation to marginal habitats. Annu. Rev. Ecol. Evol. Syst. 39:321–342.

Keller, L. F., and D. M. Waller. 2002. Inbreeding effects in wild populations. Trends Ecol. Evol. 17:230–241.

Kimura, M., T. Maruyama, and J.F. Crow. 1963. The mutation load in small populations. Genetics 48:1303–1312.

Kirkpatrick, M., and P. Jarne. 2000. The effects of a bottleneck on inbreeding depression and the genetic load. Am. Nat. 155:154–167.

Koelewijn, H. P., V. Koski, and O. Savolainen. 1999. Magnitude and timing of inbreeding depression in Scots pine (*Pinus sylvestris* L.). Evolution 53:758–768.

Koski, M. H., N. C. Layman, C. J. Prior, J. W. Busch, and L. F. Galloway. 2019. Selfing ability and drift load evolve with range expansion. Evol. Lett. 3:500–512.

Kuznetsova A., P. B. Brockhoff, and R. H. B. Christensen. 2017. lmerTest package: tests in linear mixed-effects models. J. Stat. Softw. 82. 1-26.

Laenen, B., A. Tedder, M. D. Nowak, P. Toräng, J. Wunder, S. Wötzel, K. A. Steige, Y. Kourmpetis, T. Odong, A. D. Drouzas, M. C. A. M. Bink, J. Ågren, G. Coupland, and T. Slotte. 2018. Demography and mating system shape the genome-wide impact of purifying selection in *Arabis alpina*. Proc. R. Soc. B. 117:201707492–201707496.

Lee-Yaw, J. A., H. M. Kharouba, M. Bontrager, C. Mahony, A. M. Csergő, A. M. Noreen, Q. Li, R. Schuster, and A. L. Angert. 2016. A synthesis of common garden experiments and ecological niche models suggests that range limits are often niche limits. Ecol. Lett. 19:710–722.

Lee-Yaw, J. A., M. Fracassetti, and Y. Willi. 2018. Environmental marginality and geographic range limits: a case study with *Arabidopsis lyrata* ssp. *lyrata*. Ecography 41:622–634.

Lohmueller, K. E., A. R. Indap, S. Schmidt, A. R. Boyko, R. D. Hernandez, M. J. Hubisz, J. J. Sninsky, T. J. White, S. R. Sunyaev, R. Nielsen, A. G. Clark, and C. D. Bustamante. 2008. Proportionally more deleterious genetic variation in European than in African populations. Nature 451:994–998.

Lynch, M. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. Evolution 45:622–629.

Lynch, M., J. Conery, and R. Bürger. 1995a. Mutational meltdowns in sexual populations. Evolution 49:1067–1080.

Lynch, M., J. Conery, and R. Bürger. 1995b. Mutation accumulation and the extinction of small populations. Am. Nat. 146:489–518.

Matos, G., C. Palma-Silva, M. H. Bodanese-Zanettini, C. Lexer, and F. Bered. 2015. Limited pollen flow and high selfing rates toward geographic range limit in an Atlantic forest bromeliad. Flora Morphol. Distrib. Funct. Ecol. Plants 211:1–10.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution 29:1–10.

Nordborg, M., and P. Donnelly. 1997. The coalescent process with selfing. Genetics 146:1185–1195.

Oakley, C. G., J. Ågren, and D. W. Schemske. 2015. Heterosis and outbreeding depression in crosses between natural populations of *Arabidopsis thaliana*. Heredity 115:73–82.

Pannell, J. R. 2015. Evolution of the mating system in colonizing plants. Mol. Ecol. 24:2018–2037.

Peischl, S., and L. Excoffier. 2015. Expansion load: recessive mutations and the role of standing genetic variation. Mol. Ecol. 24:2084–2094.

Peischl, S., I. Dupanloup, M. Kirkpatrick, and L. Excoffier. 2013. On the accumulation of deleterious mutations during range expansions. Mol. Ecol. 22:5972–5982.

Peischl, S., M. Kirkpatrick, and L. Excoffier. 2015. Expansion load and the evolutionary dynamics of a species range. Am. Nat. 185:E81–E93.

Pironon, S., G. Papuga, J. Villellas, A. L. Angert, M. B. García, and J. D. Thompson. 2017. Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. Biol. Rev. 92: 1877–1909.

Polechová, J. 2018. Is the sky the limit? On the expansion threshold of a species range.

Polechová, J., and N. H. Barton. 2015. Limits to adaptation along environmental gradients. P. Natl. Acad. Sci. USA 112:6401–6406.

Pollak E. 1987. On the theory of partially inbreeding finite populations. I. Partial selfing. Genetics 117:353–360.

Pujol., B., S. R. Zhou, J. Sanchez Vilas, and J. R. Pannell. 2009. Reduced inbreeding depression after species range expansion. P. Natl. Acad. Sci. USA 106:15379–15383.

R Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/

Schultz, S. T., and M. Lynch. 1997. Mutation and extinction: the role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. Evolution 51:1363–1371.

Sexton, J. P., P. J. McIntyre, A. L. Angert, and K. J. Rice. 2009. Evolution and ecology of species range limits. Annu. Rev. Ecol. Evol. Syst. 40:415–436.

Sutherland W. J., R. P. Freckleton, H. C. J. Godfray, S. R. Beissinger, T. Benton, D. D. Cameron, et al. 2013. Identification of 100 fundamental ecological questions. J. Ecol. 101:58–67.

Willi, Y. 2013. Mutational meltdown in selfing *Arabidopsis lyrata*. Evolution 67:806–815.

Willi, Y. 2019. The relevance of mutation load for species range limits. Am. J. Bot. 106:757–759.

Willi, Y., and K. Määttänen. 2010. Evolutionary dynamics of mating system shifts in *Arabidopsis lyrata*. J. Evol. Biol. 23:2123–2131.

Willi Y., and J. Van Buskirk. 2019. A practical guide to the study of distribution limits. Am. Nat. 193:773–785.

Willi, Y., M. Fracassetti, S. Zoller, and J. Van Buskirk. 2018. Accumulation of mutational load at the edges of a species range. Mol. Biol. Evol. 35:781–791.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16: 97–159.

Wright, S. I., S. Kalisz, and T. Slotte. 2013. Evolutionary consequences of self-fertilization in plants. Proc. R. Soc. B. 280: 20130133.

Table 1: Summary of models testing for the effect of genomic estimates of mutational load or geographic distance to core and mating system (F_{IS}) on population-level heterosis at five common garden sites

		Genomic load						Distance to core [km]			F_{IS}					
Dependent variable	***N***	***Estimate***	***χ²***		***R²m***	***R²c***		***Estimate***	***χ²***		***Estimate***	***χ²***		***R²m***	***R²c***	
Multiplicative performance (MP)																
MP I to year 3	89	**1.90**	**6.99**	**	0.117	0.342	†,‡,§	**0.75**	**6.53**	*	0.32	2.05		0.175	0.350	†,‡,§
MP II to year 3	89	**2.11**	**9.30**	**	0.138	0.431	†,‡	**0.75**	**6.94**	**	0.38	3.12	(*)	0.188	0.438	†,‡
MP I to year 2	89	**2.47**	**14.31**	***	0.179	0.405	‡,§	**0.68**	**5.85**	*	**0.50**	**5.35**	*	0.200	0.414	‡,§
MP II to year 2	89	**2.52**	**14.65**	***	0.168	0.423		**0.72**	**6.52**	*	**0.51**	**5.66**	*	0.192	0.431	
Life stage components																
Germination	89	-0.10	0.19		0.005	0.465	§	**0.19**	**4.61**	*	-0.10	2.18		0.105	0.470	
Survival fall year 1	89	**0.24**	**7.38**	**	0.056	0.334		0.00	0.01		**0.10**	**10.43**	**	0.087	0.361	§
Survival winter year 1	89	0.02	0.18		0.002	0.077	‡	0.03	1.19		0.00	0.04		0.013	0.088	‡
Survival summer year 2	89	0.07	0.04		0.000	0.160		0.13	0.69		-0.03	0.06		0.007	0.161	
Survival winter year 2	44	0.27	0.45		0.010	0.010	†	-0.25	2.19		**0.29**	**4.78**	*****	0.113	0.113	†
Reproduction year 2																
Survival to flowering year 2	89	0.06	2.81	(*)	0.026	0.178	§	-0.01	1.02		0.02	3.57	(*)	0.035	0.186	
Bolting	89	**0.64**	**9.90**	**	0.067	0.408		0.05	0.3		**0.16**	**4.83**	*	0.050	0.399	
Reproductive output	82	**1.24**	**7.51**	**	0.127	0.281	‡	**0.45**	**5.4**	*	0.16	1.07		0.151	0.301	‡,§
Number of fruits	82	**1.09**	**6.26**	*	0.104	0.244	‡	**0.45**	**6.39**	*	0.13	0.87		0.150	0.263	‡
Demographic rate																
λ	89	**0.31**	**6.16**	*	0.108	0.470		0.17	2.71	(*)	**0.23**	**4.51**	*	0.160	0.476	

Population heterosis estimates ($\log_{10}$-transformed) were assumed to follow Gaussian distributions. The effect of distance to core and F_{IS} were assessed in the same model. Test statistics include regression coefficient (*estimate*), χ^2-value and the marginal and conditional R^2 of the model. Genomic load, distance to core and F_{IS} were standardized prior to analyses (mean = 0). Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

†For one of the five common gardens (CG3), the experiment stopped early and variables consider performance to year 2 only.

‡Model fits with significant (positive) intercept.

§The *bobyqa* optimizer was used when models initially failed to converge.

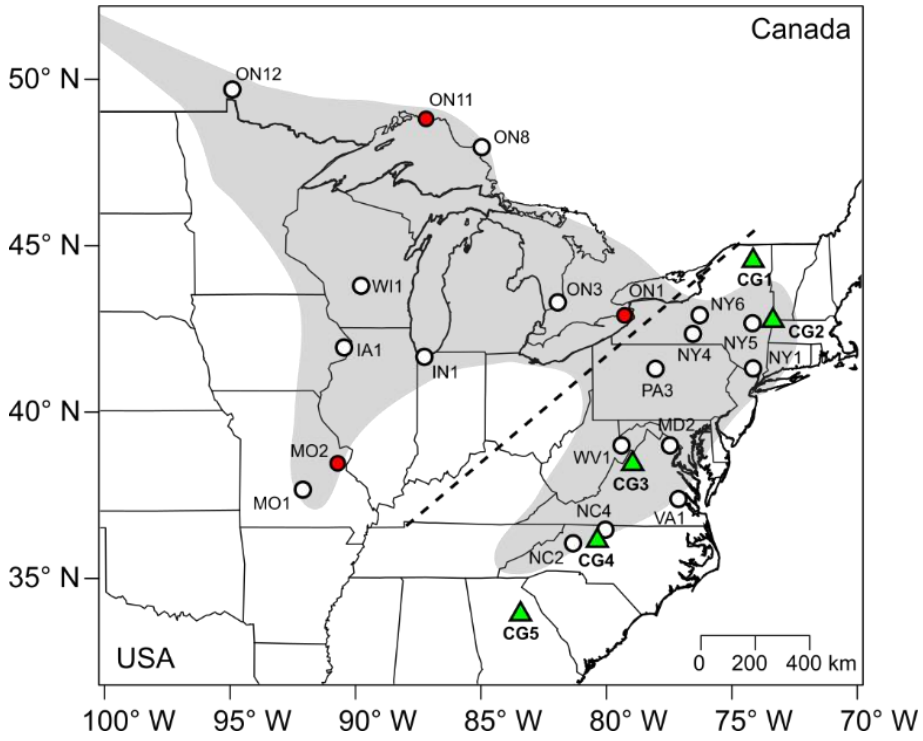

Figure 1: Distribution map of *Arabidopsis lyrata* in eastern North America with the locations of the 20 populations studied and the 5 common garden sites. Circles filled in white or red represent outcrossing and selfing populations, respectively. Population labels consist of the abbreviation for state (USA) or province (Canada) and a number (as in Willi *et al*. 2018). Green triangles represent the five common garden (CG) sites; numbers added to labels are in sequence of north to south. The dashed line is the split between eastern and western genetic clusters. Of the 20 populations, two were used as partner-populations for between-population crosses, NY1 for crosses with eastern populations, and IA1 for crosses with western populations.

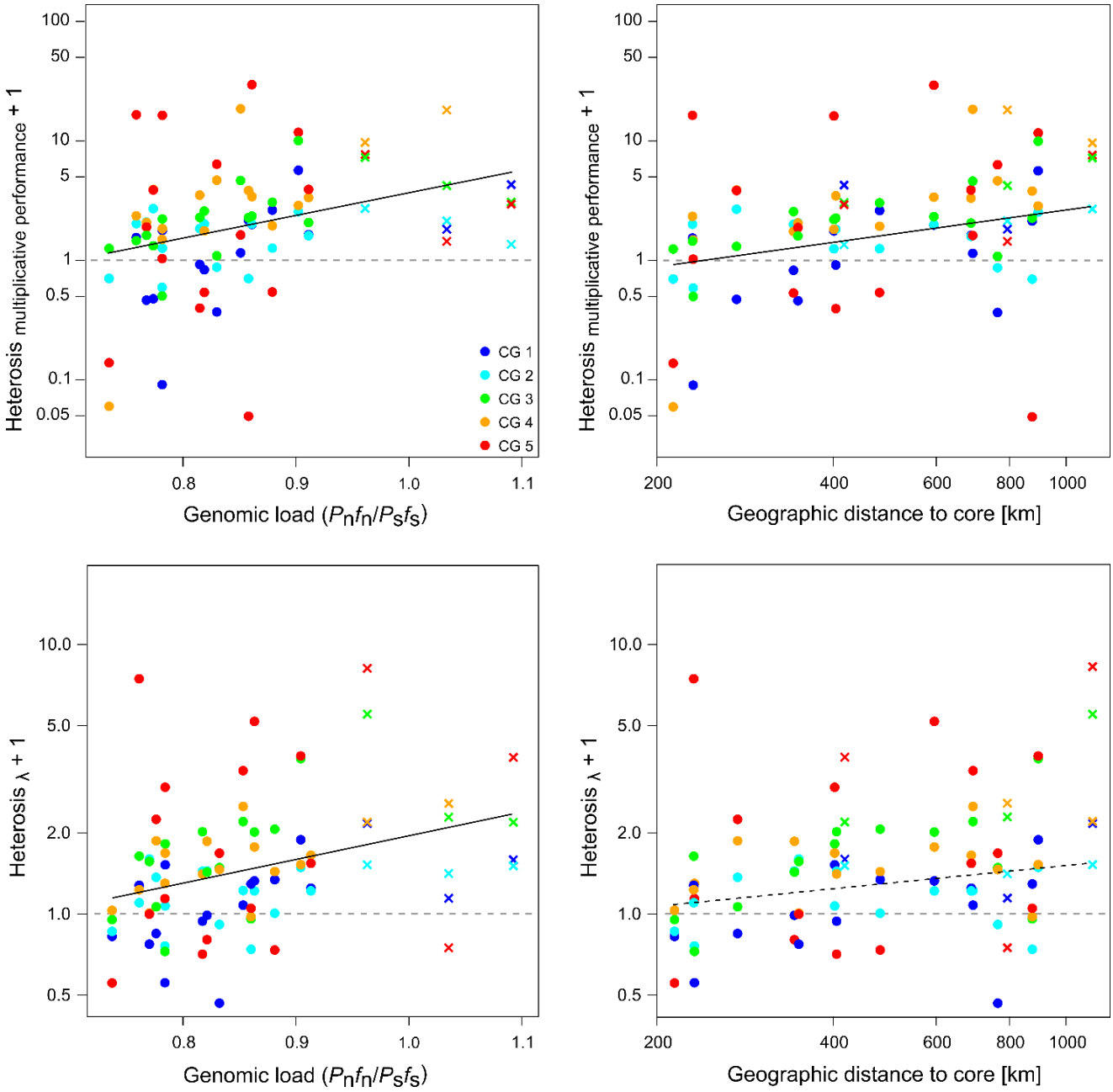

Figure 2: Relationship between heterosis in multiplicative performance or demographic rate and a genomic estimate of mutational load or geographic distance to core. Heterosis was estimated based on *multiplicative performance I* up to year 3 (**top**) or on λ (**bottom**) at the population level within each common garden site (CG1-5). Outcrossing populations are indicated by dots, selfing populations are indicated by crosses. Black lines represent the significant (full) or marginal (dashed) model-predicted slopes for heterosis (from test statistics in Table 1). The gray dashed line represents the value at which heterosis drops below 0, indicating outbreeding depression. Genomic estimate of mutational load (genomic load, **left**) was the ratio of genome-wide non-synonymous polymorphic sites multiplied by their derived mean frequency to synonymous polymorphic sites multiplied by their

derived mean frequency. Geographic distance to core (**right**) was the distance of a population back to the glacial refugium along the map-projected population phylogeny, or the direct great-circle distance to the glacial refugium for older populations. Test statistics are reported in Table 1 and Table S10.

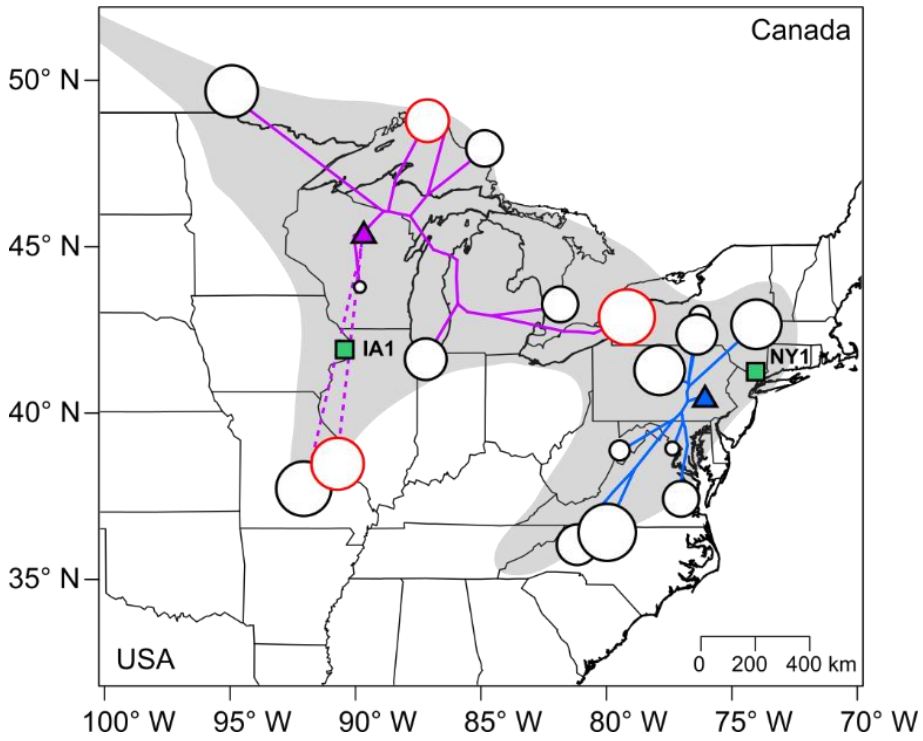

Figure 3: Expressed mutational load estimated by heterosis is increased at range edges of *Arabidopsis lyrata*. The 18 populations studied are represented by dots of varying diameter, porportional to their $\log_{10}$-transformed mean heterosis across common garden sites, calculated based on multiplicative performance including flower production during two reproductive seasons ($\log_{10}(\text{heterosis}_{\text{multiplicative performance I to Y3}} + 1)$ ranging from -1.4 to 1.4). Solid lines in purple and blue indicate the map-projected phylogeny from the western and eastern cores (presumed glacial refuge areas indicated by trianges) or connections to the core for older populations in the southwest (dashed purple lines). Mating system of populations is indicated by circle color: black for outcrossing and red for selfing. The two populations used as pollon donors in between-population crosses are represented by green squares. The approximate range of the species is shown by the gray-shaded area.

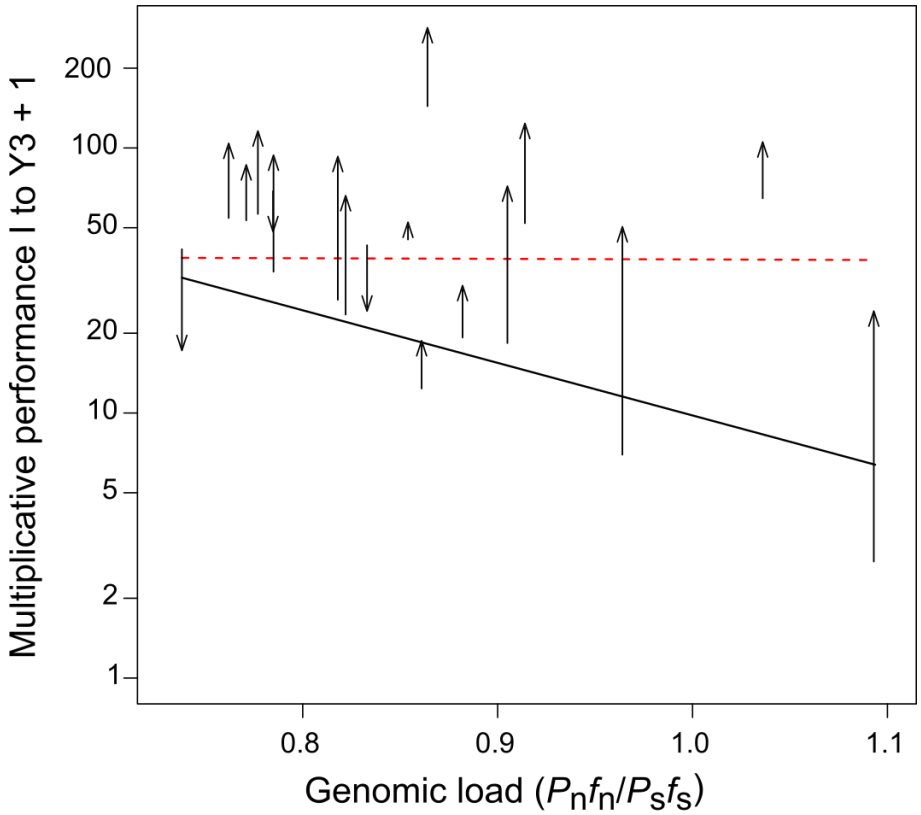

Figure 4: Relationshisp between population mean performance of within- (WPC) or between-population crosses (BPC) and THE genomic estimate of mutational load. Lines represent model-predicted slopes for population mean *multiplicative performance I* up to year 3 of WPC (solid black) and BPC (dashed red) (from test statistics in Table S13). Genomic estimate of mutational load was the ratio of the genome-wide number of non-synonymous polymorphic sites multiplied by their mean derived frequency to the number of synonymous polymorphic sites multiplied by their mean derived frequency ($P_n f_n / P_s f_s$). Arrows represent the direction of change in mean performance across common garden sites from WPC (tail of the arrow) to BPC (head of the arrow). Fifteen out of 18 populations have arrows pointing upward, indicating heterosis.

Supporting information

Table S1: Information on the *Arabidopsis lyrata* populations studied

Population code	Latitude [° N]	Longitude [° W]	Ecological variables		Variables on population history				
			Min. temp. early spring [°C] †	Mean prec. summer [mm] †	Cluster‡	Mating system §	Genomic load ($P_n f_n / P_s f_s$) ‡	Distance to core [km]‡	F_{IS} §
IA1	41.97	90.37	-6.65	106.3	West	outcrossing	0.80690	402.03	0.065
IN1	41.61	87.19	-4.60	97.3	West	outcrossing	0.83283	761.43	-0.031
MD2	38.99	77.25	-1.75	95.3	East	outcrossing	0.78476	230.85	0.017
MO1	37.72	92.06	-2.40	94.3	West	outcrossing	0.90542	893.01	0.090
MO2	38.47	90.71	-2.65	88.0	West	selfing	1.03637	791.00	0.677
NC2	36.04	81.16	-1.65	118.3	East	outcrossing	0.91391	686.11	0.043
NC4	36.41	79.96	0.20	104.0	East	outcrossing	0.86410	593.27	0.021
NY1	41.30	73.98	-3.90	100.0	East	outcrossing	0.77297	193.11	0.051
NY4	42.35	76.39	-7.25	94.0	East	outcrossing	0.77737	273.62	0.011
NY5	42.66	74.02	-8.45	97.3	East	outcrossing	0.78503	400.86	0.094
NY6	43.00	76.09	-7.25	93.3	East	outcrossing	0.77053	348.12	-0.047
ON1	42.87	79.18	-6.10	82.3	West	selfing	0.96393	1105.49	0.700
ON11	48.77	87.13	-15.60	81.7	West	selfing	1.09270	417.24	0.950
ON12	49.65	94.92	-16.75	84.7	West	outcrossing	0.85352	690.89	0.025
ON3	43.26	81.84	-6.80	79.7	West	outcrossing	0.86054	872.34	-0.041
ON8	47.93	84.85	-14.70	84.0	West	outcrossing	0.88166	479.45	-0.038
PA3	41.28	77.87	-5.60	101.0	East	outcrossing	0.76180	230.24	0.020
VA1	37.42	77.02	0.35	108.0	East	outcrossing	0.81838	403.90	0.026
WI1	43.83	89.72	-10.00	98.7	West	outcrossing	0.73834	213.46	0.033
WV1	38.96	79.29	-3.70	96.0	East	outcrossing	0.82191	342.15	-0.052

† Data extracted from WorldClim database version 1.4 (Hijmans 2005). ‡ Willi et al. 2018. § Griffin and Willi, 2014

Table S2: Growth conditions during the crossing experiment

Growth Phase	Duration [days]	Temperature daytime [°C]	Temperature nighttime [°C]	Day length [h]	Light intensity [µmol m^{-2} s^{-1}]
Stratification	12	4	4	0	0
Germination	22	20	20	8	100
Growth †	21	22	20	10	140
Flowering initiation	10	22	20	16	240
Flowering and crossing	205	22	20	16	240

† Day length and light intensity were gradually increased every three days by 1h and 20 µmol m^{-2} s^{-1}, respectively.

Table S3: Summary of the crossing experiment and number of seeds sown in each common garden

Mother population	Father population	No. of cross families	No. of seeds sown in each common garden					Cross type
			CG1	CG2	CG3	CG4	CG5	
NY1	NY1	9	70	69	63	69	66	WPC
NY5	NY5	12	72	72	72	72	72	WPC
IN1	IN1	11	72	72	72	72	72	WPC
MO1	MO1	12	72	72	69	72	72	WPC
ON11	ON11	12	72	69	66	72	66	WPC
ON12	ON12	12	69	69	72	72	72	WPC
MO2	MO2	12	72	72	72	72	72	WPC
NC2	NC2	12	69	70	72	72	72	WPC
NC4	NC4	11	72	72	72	72	72	WPC
IA1	IA1	9	70	72	57	72	66	WPC
VA1	VA1	9	72	72	72	72	72	WPC
MD2	MD2	10	69	69	69	72	69	WPC
WV1	WV1	12	69	69	66	72	60	WPC
PA3	PA3	11	72	72	72	72	72	WPC
NY6	NY6	12	72	72	66	72	63	WPC
NY4	NY4	10	70	69	69	72	72	WPC
WI1	WI1	10	66	72	72	72	72	WPC
ON8	ON8	6	27	30	27	37	27	WPC
ON3	ON3	8	69	69	69	72	66	WPC
ON1	ON1	12	72	72	72	72	69	WPC
NY5	NY1	12	72	72	72	72	72	BPC
IN1	IA1	11	72	69	69	72	72	BPC
MO1	IA1	12	72	72	72	72	72	BPC
ON11	IA1	9	55	54	54	*0*	72	BPC
ON12	IA1	10	72	72	72	72	72	BPC
MO2	IA1	12	72	72	72	72	72	BPC
NC2	NY1	12	69	70	69	69	72	BPC
NC4	NY1	10	72	72	72	72	72	BPC
VA1	NY1	10	72	72	72	78	72	BPC
MD2	NY1	9	72	72	72	72	72	BPC
WV1	NY1	11	72	72	66	72	66	BPC
PA3	NY1	11	72	72	72	72	72	BPC
NY6	NY1	11	72	72	72	72	72	BPC
NY4	NY1	11	72	72	72	72	72	BPC
WI1	IA1	11	78	72	72	72	72	BPC
ON8	IA1	5	34	35	33	39	33	BPC
ON3	IA1	10	63	63	63	72	63	BPC
ON1	IA1	12	72	72	72	72	69	BPC

Total number of seeds sown: 12,933

Total number of cross combinations: 401

Table S4: Information on common garden sites

Transplant site	Location	Latitude [° N]	Longitude [° W]	Min. temperatures early spring (°C) †
CG1 (NY)	Beyond northern edge	44.51	74.02	-5.65
CG2 (MA)	Northern edge	42.72	73.22	-2.40
CG3 (VA)	Center	38.43	78.86	1.60
CG4 (NC)	Southern edge	36.13	80.28	4.90
CG5 (GA)	Beyond southern edge	33.93	83.36	6.95
Mean northern pop. ‡	North	-	-	-2.13
Mean center pop. ‡	Center	-	-	1.01
Mean southern pop. ‡	South	-	-	3.98

† Data extracted from WorldClim database version 1.4 (Hijmans 2005); ‡ Data measured for mean population of the eastern cluster. Northern populations: NY4, NY5, NY6; center populations: PA3, MD2, WV1; southern populations: VA1, NC2, NC4.

Table S5 Description of performance estimates

Performance estimate	Type	Level	Description			Analysis
Multiplicative performance (MP)						
MP I to year 3	continuous	Pot	Germination rate * reproductive output 2018 + 2019		†	MCMC
MP II to year 3	continuous	Pot	Germination rate * number of fruits 2018 + 2019		†	MCMC
MP I to year 2	continuous	Pot	Germination rate * reproductive output 2018		†	MCMC
MP II to year 2	continuous	Pot	Germination rate * number of fruits 2018		†	MCMC
Life stage components			*From*	*To*		
Germination	binary	Seed	Day 0	31 days after sowing	‡	REML
Survival fall year 1	binary	Seed	31 days after sowing	Soil temp. <5 °C (fall 2017)	*§	REML
Survival winter year 1	binary	Pot	Soil temp. <5 °C (fall 2017)	Soil temp. >10 °C (spring 2018)	*§	REML
Survival summer year 2	binary	Pot	Soil temp. >10 °C (spring 2018)	Soil temp. <5 °C (fall 2018)	*	REML
Survival winter year 2	binary	Pot	Soil temp. <5 °C (Fall 2018)	Count of reproductive output (spring 2019)	*†	REML
Reproduction year 2						
Survival to flowering year 2	binary	Pot	Plants that survived from end of winter 2017/18 to flowering 2018			REML
Bolting	binary	Pot	Plants that survived to flowering and produced inflorescences or not			REML
Reproductive output	continuous	Pot	Sum of all flower organs for plants that bolted			REML
Number of fruits	continuous	Pot	Potential total number of fruits from plants that produced at least one flower			REML

* Soil temperature was monitored every hour over the whole length of the experiment by 5 iButton® (Maxim Integrated, San Jose, CA, USA) per common garden site, buried under 5 cm of substrate in an empty pot.

† except for CG3: no survival after assessment of reproductive output 2018

‡ except for CG1: second cohort in spring 2018

§ except for CG4: experiment restarted in December 2017

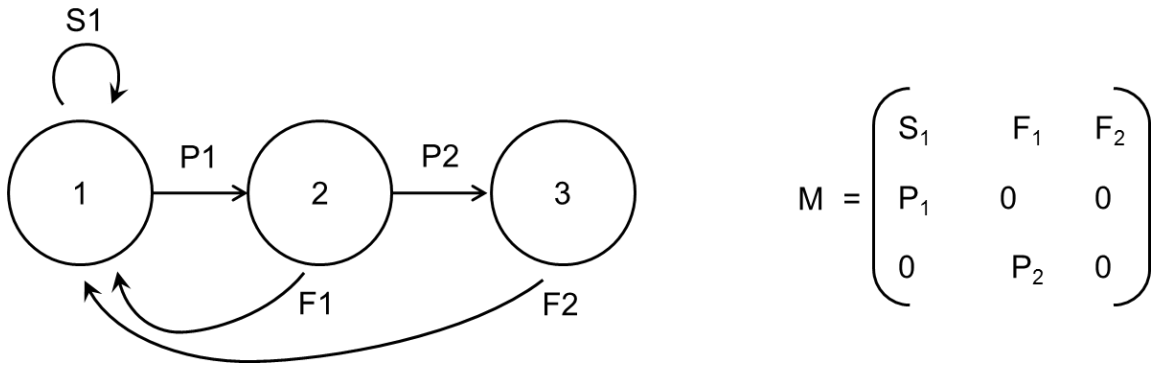

Figure S6: Estimation of population growth rate. For each combination of population, cross type and common garden site, a stage-classified matrix (**right**) was constructed based on assumptions about the life cycle (**left**). The life cycle was composed of three stages: 1–healthy seeds, 2–individuals capable of reproducing in the second year, 3–individuals capable of reproducing in the third year. The projection interval was set to one year for each stage. Survival between stage 1 and 2 (*P1*) was estimated as: germination rate in 2017 x survival from the seedling stage until the date of first flowering in the first reproductive period (year 2) at each site. Survival between stage 2 and 3 (*P2*) was estimated as the survival from the date of first flowering in the first reproductive period to the date of recording of reproductive output in the second reproductive period (year 3). We assumed that seeds that did not germinate in the first year (graduation from step 1 to 2) could survive over winter and contribute to the seed pool of the next years. We defined the probability to remain at the same stage (*S1*) as the survival of seeds over one winter. This estimate was calculated based on the seed burial experiment over one winter. The probability to remain at the same stage was set to 0 for both stage 2 and 3, assuming that no plants survived after the third year. Preliminary analysis showed that allowing individuals to remain in stage 3 indefinitely with the same probability of surviving each year than between stage 2 and 3 did not significantly affect the population growth rates (data not shown). Fecundity of stage 2 and 3 (*F1* and *F2* respectively) were estimated for each stage separately as: probability to reproduce * number of fruits * number of healthy seeds per fruit. While an estimate of the latter could have come from the crossing experiment, we assumed that these values were not reflective of the natural conditions and could introduce too much bias. Furthermore, fecundity of both stages would need to be adjusted for the chance of landing in a suitable environment for germination (including environmental effects, inter- and intraspecific competition, predation, etc.). We therefore decided to assign to all populations a standard value representing both number of healthy seeds per fruit and the probability to land in an environment suitable for germination, estimated as the value that yielded an average λ of 1 across all WPC over all sites.

Appendix S7: Parametrization of the hierarchical mixed-effects models

S7A: Hierarchical mixed-effects model with population heterosis as dependent variable

Model = lmer($\log_{10}$(heterosis + 1) ~ genomic load + (1 | maternal population) + (1 | common garden),
data = data)

S7B: Priors, and hierarchical mixed-effects model analyzed in a Bayesian (MCMC) framework, with individual multiplicative performance as dependent variable

Priors

Priors were set to be weak, using parameter expansion to improve convergence. *R* specifies the priors for the fixed effects, *G* specifies the priors for the random effects.

```
priors.model=list(
  R=list(V=diag(2), n=1, fix = 2),
  G=list(G1=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2),
       G2=list(V=diag(4), n=4, alpha.mu = rep(0,4),alpha.V = diag(4)*25^2),
       G3=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2),
       G4=list(V=diag(4), n=4, alpha.mu = rep(0,4),alpha.V = diag(4)*25^2),
       G5=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2),
       G6=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2)))
```

Parametrization of hierarchical mixed-effects models analyzed in a Bayesian (MCMC) framework

Multiplicative performance estimates were split into two datasets: the *zero_part*, a binary transformation of performance estimates with *zero_part* = 1 if performance > 0, else *zero_part* = 0; and the *norm_part* containing only the performance measures if *zero_part* = 1.

```
model = MCMCglmm(cbind(norm_part, zero_part) ~ trait -1 + trait:cross type * trait:genomic load,
            random = ~ us(trait):maternal population
            + us(trait:cross type):maternal population
            + us(trait): maternal population: maternal family
            + us(trait: cross type):maternal population:maternal family
            + us(trait):common garden + us(trait):common garden:block,
            prior = priors.model,
            rcov = ~idh(trait):units,
            family=c('gaussian', 'categorical'),
            burnin = 5000, thin = 100, nitt = 50000,
            data=data)
```

S7C: Hierarchical mixed-effects model with individual performance estimate as dependent variable

Dependent variable with binary distribution

Model = glmer(performance ~ cross type * genomic load
+ (1 + cross type | maternal population / maternal family)
+ (1 | common garden / block),
family = "binomial",
data = data)

Dependent variable with log-normal distribution

Model = lmer($\log_{10}$(performance + 1) ~ cross type * genomic load
+ (1 + cross type | maternal population / maternal family)
+ (1 | common garden / block),
data = data)

Table S8: Summary of individual performance, and heterosis based on population mean performance (in % for binary variables) up to three years at five common garden sites

	Individual performance		Population heterosis			
Dependent variable	***N***	**Mean**	***N***	**Min.**	**Mean**	**Max.**
Multiplicative performance (MP)						
MP I to year 3	6703	58.50	89	-0.96	1.88	23.50
MP II to year 3	6703	36.67	89	-0.92	2.65	30.23
MP I to year 2	6703	37.54	89	-0.95	2.96	41.21
MP II to year 2	6703	26.27	89	-0.97	3.89	71.45
Life stage components						
Germination	12933	64.2 %	89	-0.68	0.01	0.75
Survival fall year 1	7214	77.6 %	89	-0.40	0.10	1.75
Survival winter year 1	5072	85.2 %	89	-0.28	0.07	0.48
Survival summer year 2	4319	29.8 %	89	-0.83	0.64	12.42
Survival winter year 2	608	61.8 %	44	-0.68	0.11	5.00
Reproduction year 2						
Survival to flowering year 2	3731	99.6 %	89	-0.25	0.17	2.46
Bolting	3719	60.2 %	89	-0.53	0.57	5.42
Reproductive output †	2240	134.72	89	-0.82	0.69	8.00
Number of fruits	2218	96.09	82	-0.84	0.74	8.98
Fertilization rate ‡	2240	67.2 %	82	-0.47	0.09	0.76
Demographic rate						
λ	189	1.18	89	-0.53	0.73	7.29

† Sum of buds, flowers, fruits and pedicels produced by one individual

‡ Not analyzed

Table S9: Magnitude of effect of the genomic estimate of mutational load or geographic distance to core and mating system (F_{IS}) on (log_{10}-transformed) heterosis, and magnitude of effect of the genomic estimate of mutational load on (log_{10}-transformed) multiplicative performance (MP) of within-population crosses (WPC)

	Increase in heterosis + 1 (x-fold)				***Decrease in WPC performance (%)***	
Dependent variable	***N***	**Genomic load**	**Distance to core**	F_{IS}	***N***	**Genomic load**
MP I to year 3	89	4.7 *(1.0; 4.6)*	3.4 *(0.9; 3.0)*	*NS*	189	-80.3 *(32.6; 6.4)*
MP II to year 3	89	5.6 *(1.1; 6.1)*	3.4 *(1.0; 3.6)*	*NS*	189	-73.3 *(19.8; 5.3)*
MP I to year 2	89	7.5 *(1.0; 7.2)*	3.1 *(1.1; 3.3)*	3.2 *(1.5; 4.8)*	189	-75.7 *(21.0; 5.1)*
MP II to year 2	89	7.9 *(1.0; 8.0)*	3.2 *(1.1; 3.6)*	3.3 *(1.6; 5.3)*	189	-69.5 (*14.4; 4.4)*
λ	89	1.3 *(1.06; 1.36)*	*NS*	1.7 *(1.0; 1.7)*	-	-

For the general models of Table 1 analyzing heterosis, the magnitude of effect of a predictor variable was calculated as: ratio between the back-transformed predicted heterosis corresponding to the maximal value of a predictor variable, compared to the back-transformed predicted heterosis corresponding to the minimal value of a predictor variable. As in some cases outbreeding depression (negative heterosis values) was observed, 1 was added to predicted heterosis as outbreeding depression can take the maximal value of -1. For the general model of Table S13 analyzing WPC, the magnitude of effect of genomic load was calculated as: percentage difference between the back-transformed predicted performance corresponding to the maximal value of the predictor variable (in parenthesis, right) in our sampling and the back-transformed predicted performance corresponding to the minimal value of the predictor variable (in parenthesis, left).

Table S10: Summary of models testing for the effect of geographic distance to core on heterosis of outcrossing populations only, estimated on population mean performance estimates up to three years at five common garden sites

		Distance to core					
Dependent variable	***N***	***Estimate***	***χ²***		***R²m***	***R²c***	
Multiplicative performance (MP)							
MP I to year 3	75	**0.75**	**4.69**	*	0.096	0.289	†,§
MP II to year 3	75	**0.76**	**5.43**	*	0.101	0.373	†
MP I to year 2	75	0.65	3.75	(*)	0.080	0.331	§
MP II to year 2	75	**0.68**	**4.26**	*	0.081	0.341	
Life stage components							
Germination	75	0.10	1.90		0.033	0.455	
Survival fall year 1	75	0.02	0.19		0.002	0.243	§
Survival winter year 1	75	0.01	0.14		0.002	0.040	‡
Survival summer year 2	75	0.10	0.42		0.005	0.132	
Survival winter year 2	38	-0.23	1.63		0.042	0.042	†
Reproduction year 2							
Survival to flowering year 2	75	-0.01	0.13		0.001	0.249	
Bolting	75	0.08	0.70		0.007	0.310	
Reproductive output	70	**0.45**	**4.08**	*	0.097	0.286	‡,§
Number of fruits	70	**0.43**	**4.57**	*	0.096	0.235	‡

Population heterosis estimates (log_{10}-transformed) were assumed to follow Gaussian distributions. Test statistics include the regression coefficient (*estimate*), χ^2-value, and the marginal and conditional R^2 of the model. Distance to core was standardized prior to analyses (mean = 0). Model fits with significant (positive) intercept are indicated by ‡. Estimates with P-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. The *bobyqa* optimizer was used when models initially failed to converge (§). Results for random effects are not shown. For one of the five common gardens (CG3), the experiment stopped early and variables consider performance to year 2 only (indicated by †).

Table S11A: Summary of models performed on the level of individual pots, testing for the effect of cross type (between- compared to within-population crosses [0]), genomic estimate of mutational load and their interaction on multiplicative performance (MP) based either on flower production (MP I) or fruit production (MP II) up to one or two reproductive seasons, at five common garden sites

		Log-normal process, fixed effects									
		Cross type			**Genomic load**			**Cross type * genomic load**			
Dependent variable	***N***	***Mean***	***HPD interval***		***Mean***	***HPD interval***		***Mean***	***HPD interval***		
MP I to year 3	6703	**0.134**	(0.021,0.229)	*	-0.708	(-1.465,0.113)	(*)	**0.995**	(0.252,1.774)	***	†,‡
MP II to year 3	6703	**0.143**	(0.049,0.243)	**	-0.517	(-1.308,0.241)		**0.844**	(0.148,1.574)	***	†,‡
MP I to year 2	6703	**0.121**	(0.026,0.222)	*	-0.635	(-1.415,0.137)		**0.87**	(0.081,1.573)	***	‡
MP II to year 2	6703	**0.135**	(0.04,0.233)	**	-0.494	(-1.202,0.283)		**0.82**	(0.044,1.533)	***	‡
		Logistic process, fixed effects									
		Cross type			**Genomic load**			**Cross type * genomic load**			
Dependent variable	***N***	***Mean***	***HPD interval***		***Mean***	***HPD interval***		***Mean***	***HPD interval***		
MP I to year 3	6703	**0.491**	(0.226,0.754)	***	-0.972	(-3.438,1.803)		1.568	(-0.446,3.644)		†
MP II to year 3	6703	**0.497**	(0.241,0.742)	***	-1.15	(-3.713,1.447)		1.756	(-0.103,3.610)	(*)	†
MP I to year 2	6703	**0.471**	(0.223,0.761)	***	-1.134	(-3.788,1.713)		1.594	(-0.492,3.641)		
MP II to year 2	6703	**0.462**	(0.221,0.703)	***	-1.172	(-3.835,1.321)		1.632	(-0.262,3.432)	(*)	

Multiplicative performance estimates (log_{10}-transformed if >0) were assumed to follow Gaussian distributions with 0-inflation. Therefore, models assessed all fixed and random effects for their importance in both the Gaussian process (total number of flowers or fruits produced during one or two reproductive seasons) and the logistic process (binary variable depicting germination combined with survival and the capacity to initiate flowering). Estimates of coefficients are modes of an MCMC sample from the posterior distribution of parameters (mean and higher posterior density, *HPD* interval). The logistic part of the model predicts non-zeros in the distribution on the logit scale. Genomic load was standardized prior to analyses (mean = 0). Model fits with significant (positive) intercept are indicated by ‡. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown. For one of the five common gardens (CG3), the experiment stopped early and variables consider performance to year 2 only (indicated by †).

Table S11B: Summary of mixed-effects models performed on the level of individual pots, testing for the effect of cross type (between- compared to within-population crosses [0]), genomic estimate of mutational load and their interaction on performance based on germination and survival up to three years at five common garden sites, and reproduction in the second year

		Cross type (CT)			Genomic load			CT * genomic load				
Dependent variable	***N***	***Estimate***	χ^2		***Estimate***	χ^2		***Estimate***	χ^2		***R^2m***	***R^2c***
Life stage components												
Germination	12933	0.07	0.32		0.42	0.08		-1.15	0.71		0.001	0.260
Survival fall year 1	7214	**0.22**	**11.78**	***	-1.10	1.34		0.23	0.10		0.003	0.237 ‡
Survival winter year 1	5072	**0.47**	**22.40**	***	-0.53	0.34		0.92	0.66		0.006	0.297 ‡,§
Survival summer year 2	4319	**0.51**	**17.40**	***	-0.96	0.40		1.57	1.29		0.008	0.552 §
Survival winter year 2	608	-0.16	0.60		1.60	0.63		-0.11	0.00		0.003	0.227 †,‡,§
Reproduction year 2												
Survival to flowering year 2	3731	0.18	2.96	(*)	0.22	0.04		1.30	1.10		0.002	0.404 ‡
Bolting	3719	**0.51**	**14.62**	***	**-4.49**	**8.25**	**	**3.00**	**3.97**	*	0.019	0.453
Reproductive output	2240	**0.12**	**5.75**	*	-1.01	3.81	(*)	0.99	3.04	(*)	0.018	0.314 ‡,§
Number of fruits	2218	**0.14**	**7.88**	**	-0.73	2.04		0.85	2.24		0.014	0.306 ‡,§

Germination, survival and bolting were binary variables; the respective models predict non-zeros on the logit scale. All other performance estimates were $\log_{10}$-transformed and assumed to follow Gaussian distributions. Test statistics include regression coefficients of each fixed effect (*estimate*), χ^2-values, and the marginal and conditional R^2 of the model. Genomic load was standardized prior to analyses (mean = 0). Model fits with significant (positive) intercept are indicated by ‡. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. The *bobyqa* optimizer was used when models initially failed to converge (§). Results for random effects are not shown. For one of the five common gardens (CG3), the experiment stopped early and variables consider performance to year 2 only (indicated by †).

Table S12: Summary of models performed on the level of population means, testing for the effect of cross type (between- compared to within-population crosses [0]), genomic estimate of mutational load and their interaction on mean performance up to three years (per cross type) at five common garden sites

		Cross type (CT)			Genomic load			CT * genomic load					
Dependent variable	***N***	***Estimate***	χ^2		***Estimate***	χ^2		***Estimate***	χ^2		***R^2m***	***R^2c***	
Multiplicative performance (MP)													
MP I to year 3	189	**0.29**	**12.75**	***	**-1.99**	**5.20**	*	**1.91**	**4.57**	*	0.097	0.571	†,‡
MP II to year 3	189	**0.31**	**16.03**	***	**-1.62**	**3.92**	*	1.62	3.79	(*)	0.096	0.606	†,†
MP I to year 2	189	**0.28**	**15.35**	***	**-1.73**	**5.76**	*	**1.71**	**4.70**	*	0.086	0.685	‡
MP II to year 2	189	**0.28**	**17.03**	***	**-1.45**	**4.56**	*	**1.54**	**4.31**	*	0.083	0.692	‡
Life stage components													
Germination	189	0.01	0.75		0.01	0.02		-0.10	1.61		0.007	0.806	
Survival fall year 1	189	**0.01**	**8.66**	**	-0.08	3.19	(*)	0.05	1.53		0.015	0.829	‡
Survival winter year 1	189	**0.01**	**13.59**	***	0.02	0.60		0.00	0.02		0.029	0.642	‡,§
Survival summer year 2	189	**0.02**	**6.93**	**	0.02	0.11		0.04	0.28		0.006	0.909	‡,§
Survival winter year 2	189	-0.01	0.31		-0.10	0.62		0.15	0.74		0.008	0.317	†,‡,§
Reproduction year 2													
Survival to flowering year 2	189	0.00	0.67		0.01	0.22		0.03	1.68		0.007	0.807	‡
Bolting	189	**0.02**	**11.31**	***	**-0.19**	**10.12**	**	**0.15**	**5.44**	*	0.021	0.862	‡
Reproductive output	189	**0.11**	**4.00**	*	-1.00	3.79	(*)	**1.32**	**4.69**	*	0.041	0.651	‡,§
Number of fruits	189	**0.13**	**6.03**	*	-0.75	1.84		1.15	3.69	(*)	0.038	0.672	‡,§

Population mean performance estimates for each cross type ($\log_{10}$-transformed) were assumed to follow Gaussian distributions. Test statistics include regression coefficients of each fixed effect (*estimate*), χ^2-values, and the marginal and conditional R^2 of the model. Genomic load was standardized prior to analyses (mean = 0). Model fits with significant (positive) intercept are indicated by ‡. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. The *bobyqa* optimizer was used when models initially failed to converge (§). Results for random effects are not shown. For one of the five common gardens (CG3), the experiment stopped early and variables consider performance to year 2 only (indicated by †).

Table S13: Summary of models testing for the effect of genomic estimate of mutational load on population mean performance up to three years of within- (WPC) and between-population crosses (BPC) separately, at five common garden sites

		Genomic load						Genomic load				
Dependent variable	***N* (WPC)**	***Estimate***	χ^2		***R²m***	***R²c***		***N* (BPC)**	***Estimate***	χ^2	***R²m***	***R²c***	
Multiplicative performance (MP)													
MP I to year 3	100	**-1.99**	**5.20**	*	0.082	0.575	†,‡	89	-0.02	0.00	0.000	0.552	†,‡
MP II to year 3	100	**-1.62**	**3.92**	*	0.060	0.622	†,‡	89	0.04	0.00	0.000	0.568	†,‡
MP I to year 2	100	**-1.73**	**5.76**	*	0.060	0.704	‡	89	-0.01	0.00	0.000	0.652	‡
MP II to year 2	100	**-1.45**	**4.56**	*	0.047	0.718	‡	89	0.09	0.01	0.000	0.649	‡
Life stage components													
Germination	100	0.01	0.02		0.785	0.784		89	-0.08	0.78	0.817	0.827	
Survival fall year 1	100	-0.08	3.38	(*)	0.825	0.836	‡	89	-0.02	0.33	0.824	0.824	‡
Survival winter year 1	100	0.02	0.47		0.603	0.621	‡,§	89	0.02	1.15	0.699	0.715	‡,§
Survival summer year 2	100	0.02	0.11		0.889	0.865	‡,§	89	0.06	2.68	0.923	0.920	‡,§
Survival winter year 2	100	-0.25	0.27		0.232	0.232	†,‡,§	89	0.05	0.22	0.431	0.472	†‡,§
Reproduction year 2													
Survival to flowering year 2	100	0.01	0.26		0.782	0.782	‡	89	0.04	3.67	(*)	0.799	0.807	‡
Bolting	100	**-0.19**	**10.45**	**	0.870	0.875	‡	89	-0.02	0.20	0.856	0.866	‡
Reproductive output	100	-1.00	3.72	(*)	0.045	0.625	‡,§	89	0.27	0.26	0.004	0.647	‡,§
Number of fruits	100	-0.75	1.84		0.023	0.670	‡,§	89	0.37	0.55	0.007	0.646	‡,§

Population mean performance estimates for each cross type ($\log_{10}$-transformed) were assumed to follow Gaussian distributions. Test statistics include regression coefficient (*estimate*), χ^2-value, and the marginal and conditional R^2 of the model. Genomic load was standardized prior to analyses (mean = 0). Model fits with significant (positive) intercept are indicated by ‡. Estimates with P-values < 0.05 are written in bold; significance is indicated: (*) P<0.1, * P<0.05, ** P<0.01, *** P<0.001. The *bobyqa* optimizer was used when models initially failed to converge (§). Results for random effects are not shown. For one of the five common gardens (CG3), the experiment stopped early and variables consider performance to year 2 only (indicated by †).

Chapter 2: Environment dependence of the expression of mutational load, and species range limits

Antoine Perrier*, Darío Sánchez-Castro, Yvonne Willi

Department of Environmental Sciences, University of Basel, Schönbeinstrasse 6, 4056 Basel, Switzerland.

*Corresponding author:

E-mail: antoine.perrier@unibas.ch

This manuscript is in preparation for submission in *New Phytologist*.

Summary:

- Theoretical and empirical research on the causes of species' range limits suggests the contribution of several intrinsic and extrinsic factors, with potentially complex interactions. Recent theory proposes that populations at geographic range limits suffer from mutational load, due to increased genetic drift resulting from past range expansion or long-term population isolation. High mutational load at range edges may coincide with eroding environmental quality, and both may interact such that the expression of load becomes enhanced under the stressful conditions.
- Here we tested the hypothesis of environment dependence of the expression of mutational load associated with range limits in the North American plant *Arabidopsis lyrata*. For twenty populations from across the species range, within- and between-population crosses were performed and offspring raised at sites across the distribution and beyond.
- Heterosis, reflecting the expression of load, increased with heightened estimates of genomic load and with increasing environmental stress.
- We conclude that range-edge populations suffer from a two-fold genetic Allee effect, one by increased mutational load, and one by adverse environmental conditions that increase the expression of load, independent of its magnitude. Estimates of decline in demographic rates suggest that the two Allee effects may strongly contribute to range limits.

Keywords: *Arabidopsis lyrata*, environmental stress, genetic drift, geographic distribution limits, heterosis, mutational load, range dynamics, small population size.

Introduction

Geographic species distributions can be constrained by multiple factors, which may interact in complex manners (Gaston, 2009; Roy *et al.,* 2009; Sexton *et al.,* 2009; Louthan *et al.,* 2015; Connallon & Sgrò, 2018; Willi & Van Buskirk, 2019). One body of theory proposes that range expansion and long-term isolation of populations away from the distribution core are associated with enhanced genetic drift. The resulting accumulation of deleterious mutations contributes to range limits or may be a sole explanation under some circumstances (Peischl *et al.,* 2013; Henry *et al.,* 2015; reviewed in Willi, 2019). Environmental conditions may also degrade toward and beyond a species' range limits (Brown, 1984; Cahill *et al.,* 2014; Hargreaves *et al.,* 2014; Lee-Yaw *et al.,* 2016). In addition to their respective effects on population performance, mutational load and environmental stress may interact synergistically, as has been suggested for inbreeding depression (Roff, 1997, pages 285-338; Reed *et al.,* 2012). If this applies, the contribution of mutational load to range limits may be much stronger (Liao & Reed, 2009), and likely be relevant in shaping the distribution of many species with a recent history of dynamic ranges. Here we studied the environment dependence of mutational load at species range limits.

In populations of small effective size (N_e), the accumulation of deleterious mutations results from genetic drift opposing purifying selection (Kimura *et al.,* 1963), with as consequence reduced mean performance and persistence of populations (Lynch *et al.,* 1995). Mutational load was predicted to accumulate under range expansion accompanied by serial bottlenecks (Peischl *et al.,* 2013, 2015; Peischl & Excoffier, 2015). The accumulation of mutational load at the expansion front slows down further expansion or may even halt the expansion when genome-wide recombination is low (Peischl *et al.,* 2015). Theoretical work also showed that this so-called expansion load depends mostly on recessive deleterious mutations (Peischl & Excoffier, 2015), and can persist in populations for thousands of generations (Peischl *et al.,* 2013). Mutation accumulation may also set range limits without expansion but when the carrying capacity of suitable habitats declines along a line of habitat

patches, particularly under low dispersal and population growth rates (Henry *et al.,* 2015). In nature, mutational load may play an important role in explaining species distribution, as many temperate taxa are characterized by a history of post-glacial expansion (Hewitt, 2000, 2004) and population isolation generally increases toward range limits (Pironon *et al.,* 2017). In line there is a general trend of reduced genetic diversity and effective population size toward range limits (Eckert *et al.,* 2008; Pironon *et al.,* 2017). Genomic evidence for the accumulation of mutational load following range expansion was found for humans (Lohmueller *et al.,* 2008; Henn *et al.,* 2016; Peischl *et al.,* 2018) and in plants for expansion and rear-edge isolation (Zhang *et al.,* 2016; González-Martínez *et al.,* 2017; Willi *et al.,* 2018; Koski *et al.,* 2019). Phenotypic studies, though scarce, also provided evidence for the increased expression of mutational load toward the edges of species distribution in natural systems (Willi *et al.,* 2018; Koski *et al.,* 2019; Perrier *et al.,* 2020), and in experimental evolution with bacteria (Bosshard *et al.,* 2017). In the former studies, the expression of mutational load was estimated by crossing experiments and the estimation of the heterosis effect of between- compared to within-population crosses. Heterosis should depict load adequately if it is due to recessive deleterious mutations that are then paired with healthy, wild type alleles in the between-population crosses (dominance model of heterosis, Crow, 1987).

Toward range limits, apart from the accumulation of mutational load, the environment may exhibit a pattern of increasing stressfulness. Empirical studies indeed documented a general decline in habitat suitability toward range limits in many taxa (Sexton *et al.,* 2009; Pironon *et al.,* 2017). A more consistent and pronounced effect was found when species were experimentally transplanted beyond their range edges (Sexton *et al.,* 2009; Hargreaves *et al.,* 2014). For the same set of transplant experiments, reduced fitness beyond range edges coincided with a decline in predicted habitat suitability revealed by niche modelling (Lee-Yaw *et al.,* 2016). These studies suggest that the decay in habitat suitability acts additionally to mutational load. The two together may cause a double load in range-edge populations and be the cause of range limits, by impeding colonization beyond the current range limits at the leading edge and reducing persistence at the rear edge.

Furthermore, the expression of mutational load may become even stronger under environmental stress. Mechanisms of an interaction between mutational load and environmental stress are still not completely understood (Fujimoto *et al.,* 2018). The interaction has been discussed and tested extensively in the context of inbreeding depression (Armbruster & Reed, 2005; Willi *et al.,* 2007; Fox & Reed, 2010). Three mechanisms have been proposed (Reed *et al.,* 2012). A first and empirically well-supported mechanism is that stress induces the expression of deleterious mutations normally silent under benign conditions (Kondrashov & Houle, 1994; Elena & de Visser, 2003; Reed *et al.,* 2012). Such mutations were found for example in *Drosophila*, where recessive alleles were linked to increased mortality under temperature stress in inbred lines (Vermeulen & Bijlsma, 2004). Extreme stress and inbreeding impede the function of heat shock proteins such as Hsp90, which are essential for the buffering of the expression of deleterious mutations (Rutherford & Lindquist, 1998; Queitsch *et al.,* 2002; Bergman & Siegal, 2003). A second mechanism is that mutational load lowers stress resistance or tolerance, by impairing the maintenance of cellular homeostasis under stress, or affecting tissue and genome repair after stress exposure (reviewed in Agrawal & Whitelock, 2010; Reed *et al.,* 2012). For example, hybrids of *Arabidopsis thaliana* expressed more metabolites from central pathways linked to higher freezing tolerance compared to inbred lines (Korn *et al.*, 2010), and the regulation of stress-response pathways led to higher recovery after stress exposure (Miller *et al.,* 2015). The third mechanism is that stress increases phenotypic variance, increasing the opportunity for relative fitness measures to be lower in inbred individuals compared to outbred ones (Waller *et al.,* 2008). Reed *et al.,* (2012) found that variation in inbreeding depression in nine studies was predicted by stress itself and – to a weaker extent – also by phenotypic variance. In natural populations, the effect of stress on heterosis caused by crossing divergent populations, as a way of estimating the expression of mutational load, has rarely been investigated (Fenster & Galloway, 2000; Prill *et al.,* 2014; Li *et al.,* 2018), and never in the context of range limits.

In this study, we tested for environment dependence of the expression of mutational load at the range limits of the North American species *Arabidopsis lyrata* subsp. *lyrata* (L.). Distribution

modelling showed that the current range limits of the species reflect niche limits well in the south and the north (Lee-Yaw *et al.*, 2018). Furthermore, the current distribution has been strongly impacted by fast post-glacial range expansion from two separate refugia (Willi & Määttänen, 2010; Griffin & Willi, 2014; Willi *et al.*, 2018). Expansion was associated with an increase in mutational load, with the highest load nowadays found at range edges (Willi *et al.*, 2018; Perrier *et al.*, 2020). Here we estimated the expression of mutational load via a heterosis effect and tested whether it was increased under stress at range limits. Plants of within- and between-population crosses were raised in a latitudinal transplant experiment with five common garden sites within and beyond the distribution range. Environmental stress was estimated by the difference in conditions experienced by populations in each common garden compared to their site of origin (Hoffmann & Hercus, 2000), assuming local adaptation to the site of origin, and by relative performance of each population at a site compared to the best site. We tested three main hypotheses: (*i*) Heterosis, the expression of mutational load, is higher under more environmentally stressful conditions (significant effect of environmental stress on heterosis). (*ii*) Heterosis is highest for populations with high genomic estimates of load when raised under stressful conditions (significant interaction between a genomic estimate of load and environmental stress). (*iii*) Heterosis is highest when range-edge populations with highest load are transplanted beyond range limits.

Materials and Methods

Plant material and crossing design

We selected 20 populations of *A. lyrata* subsp. *lyrata* of a species-wide seed collection for a crossing experiment within and between populations (Fig. 1, Supporting Information Table S1). The populations represented the two ancestral clusters existing in the species (Griffin & Willi, 2014; Willi *et al.*, 2018), different positions within the range, the two mating systems of predominant outcrossing and predominant selfing (MO1, ON1, ON11), and the range of genomic signatures of mutational load

(Willi *et al.,* 2018). Selfing populations were shown to occur predominantly at the range edges (Griffin & Willi, 2014). Eighteen of the populations (9 per ancestral cluster) were considered target populations in the assessment of the expression of mutational load by heterosis. The remaining two were partner populations for producing between-population crosses (NY1 for the eastern cluster, IA1 for the western cluster). These partner populations were selected because they occurred in the core of each ancestral cluster and had high genomic diversity, indicating a history of little genetic drift. Seeds of all selected populations had been collected homogeneously over an area of about 450 m^2 in each population, between 2007 and 2014. Seeds of individual mother plants (seed families) had been stored separately in bags, at 4 °C under dry and dark conditions.

We raised individuals of 26 randomly selected seed families per population in growth chambers, assuming low relatedness between families. Three haphazardly chosen seeds per family were initially sown in separate pots filled with a 1:1 mixture of sand and peat. Pots were saturated with water and kept 12 days at 4 °C in the dark to stratify the seeds. To promote germination, pots were then transferred to growth chambers (CLF Plant Climatics, Wertingen, Germany) with the following conditions: 8h of light at 100 µmol m^{-2} s^{-1} and 20 °C, 16h of dark at 20 °C. Seedlings were thinned to one per pot 36 days after sowing. To promote growth and flowering, day length, light intensity and temperatures were increased over 25 days to reach the final conditions kept until the end of the crossing experiment: 16h of light at 240 µmol m^{-2} s^{-1} and 22 °C, 8h of dark at 20 °C. Then, when the first plants started to bolt, all pots were transferred to a greenhouse with similar conditions, to perform the crosses.

Plants of each population were randomly appointed to “mothers”/pollen recipients (12), “fathers”/pollen donors (12), and backups (2). Each mother plant was crossed using pollen from two “fathers”: one of the same population (WPC), and one of the partner population (BPC). The WPC crosses were also performed for the two partner populations (families and cross combinations listed in Table S2). Crosses were non-reciprocal. Pollination was performed at the bud stage to prevent spontaneous cross- or self-pollination. Flower buds were opened with sterile tweezers and immature

anthers removed. Mature anthers of the father plant were gently rubbed over the stigma. Unwanted pollination by insects was avoided by placing pollinated plants into insect-proof growth chambers until fruit formation (3-5 days). Mature siliques were collected and pooled in seed bags per family-cross-combination, and left to dry for two weeks at ambient temperature in the dark. We repeated each cross combination per mother plant to obtain enough seeds to perform the outdoor experiment (at least 60 healthy-looking seeds). In case of systematic cross failure, we replaced the father plant with a backup plant. Seeds were stored at 4 °C, under dry and dark conditions. In total, 401 family-cross type combinations where kept for the common garden experiment (Table S2).

Transplant experiment

Five common garden sites (CG) were selected along a *ca.* 1400 km latitudinal gradient crossing the species distribution of *A. lyrata* in the eastern USA (Fig. 1). The five sites represented the centre of the range of *A. lyrata* (CG3, Harrisonburg, VA), the southern and northern range edges (CG4, Winsten-Salem, NC, and CG2, Williamstown, MA, respectively), and areas beyond the edges (CG5, Athens, GA, and CG1, in the Adirondacks, NY, respectively). Sites were chosen based on extrapolations of their minimum temperature in early spring (WorldClim database version 2.0, Fick & Hijmans, 2017; Table S3), the most constraining environmental variable of *A. lyrata* at the southern and northern range edges (Lee-Yaw *et al.*, 2018). In each common garden, seeds of all crosses were sown in pots randomly distributed within each of three spatial blocks. In each pot, two seeds of a cross were sown, later thinned to one per pot (at least five weeks after sowing). Sowing was performed in early fall to early winter 2017, starting in CG1, timed about 6 weeks before the long-term daily average temperature fell below 10°C. Initially we protected the seeds from heavy rainfall by a temporary tent during the first 17 to 19 days after sowing. At CG4 we had to restart the experiment because of issues with water quality; we started the experiment again in December 2017 and kept the plants initially in a greenhouse but moved pots outdoors after an acclimation period and before the main winter frosts occurred. In all common gardens, the sowing substrate was regularly kept moist

until the first day of snow or night frost. Herbivory was prevented with organic slug repellent in an early phase, and by setting up a permanent fence to keep out larger herbivores. No further interventions were made until the end of the experiment in summer 2019. (At CG3, the experiment had to be stopped in fall 2018 because the garden was needed for another experiment). Five data loggers (iButton®, Maxim Integrated Products, Inc) were used to monitor air temperature every hour (1.5 m above ground, in the shadow) at each common garden site for the duration of the experiment.

From the day of sowing until the first thinning, individual performance was tracked on the level of the seedling and later on the level of the pot/plant. Germination, defined as when a seedling had two fully opened cotyledons, was recorded three days a week for four to five weeks until the peak of germination was over, then once a week until thinning. Death was recorded at the same time as germination was checked, later once a week unless there was a snow cover. Once bolting was observed at a common garden in 2018, day of first flower opening was recorded on three days a week. Germination, death and day of flowering were corrected by the mid-time between recording. Reproductive output of a plant was assessed several weeks after peak flowering: ~ 9 weeks after opening of the first flowers within each common garden in the second year (2018), and ~ 5 weeks after first flowering, in the third year (2019), estimated from flowering dates of the previous year. Total reproductive output was the sum of fruits, pedicels (flowers that did not develop into a fruit), flowers and flower buds. Female reproductive output was calculated as the total number of fruits and potential additional fruits that could have formed from buds and open flowers: fruits + ((flowers + buds) × (fruits / (fruits + pedicels))).

During the second winter, we also carried out a seed survival experiment to assess the potential contribution of the seed bank to population performance. For each combination of cross type and population, we pooled 100 seeds gathered from five to twelve mother plants. Seeds were then randomly dispatched in 10 groups of 10 seeds, placed in separate bags made out of micro-perforated fabric allowing moisture and air exchange (nonwoven polypropylene-felt, 40 g/m^2). Two replicate bags per cross combination were randomly attributed to each common garden site. Bags

were placed in common gardens in October 2018 on bare ground next to pots, exposed to natural conditions. Bags were retrieved in late spring 2019 and visually screened to count germinated seedlings and seeds. Survival of non-germinated seeds was tested by first stratifying them on filter paper disks soaked with 1.5 ml of 0.05% gibberellic acid in petri-dishes (10 days, 4 °C, no light), then keeping them under similar conditions as for the crossing experiment. Germination was scored four times over 20 days. Seed survival was then estimated for each replicate bag as: (germinated seedlings + germinated seedlings with gibberellic acid)/10.

Statistical analysis

Main analyses were based on lifetime performance calculated on the level of the pot up to year 3 (or year 2 for CG 3). *Multiplicative performance* (*MP*) *I* was the product of the fraction of seeds that germinated and the sum of total reproductive output recorded in year 2 and 3. *MP II* was germination multiplied by the sum of total number of fruit recorded up to year 3 (or year 2). Initial analyses checked for a general effect of cross type, genomic load and environmental stress, and were performed on the level of the pot. As cross type and interactions were significant, main analyses were done on the level of population heterosis. Population-level heterosis was calculated as the increase in performance due to between-population outcrossing relative to within-population crossing: $(W_{BPC} - W_{WPC})/W_{WPC}$. W_{WPC} and W_{BPC} were estimated as the performance within each common garden averaged on the level of population-cross type combination (population means of family means of individual performance, within cross-type). For heterosis, the final W_{WPC} was the average of the population-level performance of each target population and its partner population within a common garden. Heterosis was also assessed for the finite rate of increase, λ, per year. We constructed stage-classified matrices (Caswell, 2001) for each combination of population and cross type in each common garden, based on population mean performance over the three years. These matrices comprised three stages: (1) healthy seeds in year 1, (2) reproducing individuals of year 2, (3)

reproducing individuals of year 3, with a projection interval of one year between each stage. The parametrization of the matrices is detailed in Fig. S1.

Pre-analysis on *MP* on the level of the pot were performed in a Bayesian framework on 10 parallel chains, with the package *MCMCglmm* (Hadfield, 2010, 2019), because *MP* was zero-inflated (parametrization of models and details on analyses given in Method S1). Main analyses focused on the effect of genomic load, environmental stress and their interaction on $\log_{10}$-transformed heterosis by hierarchical mixed-effects models based on restricted maximum likelihood, using the packages *lme4* (Bates *et al.*, 2015) and *LmerTest* (Kuznetzova *et al.*, 2017) in *R* (R Core Team, 2020). Crossed random effects were population and common garden, while fixed effects were environmental stress, the genomic estimate of mutational load of the mother population, as well as their interaction (model parametrization detailed in Method S2A). The genomic estimate of mutational load (hereafter genomic load) was the ratio of non-synonymous polymorphic sites to synonymous polymorphic sites, adjusted for their mean derived allele frequency (relative to *A. thaliana*), $P_n f_n / P_s f_s$ (calculated in Willi *et al.*, 2018). Environmental stress was the difference in the average minimum temperature in April and May between common garden and site of origin of populations: $\Delta\ T_{min} = T_{min\ CG} - T_{min\ origin}$. $T_{min\ CG}$ was calculated based on records taken by the temperature loggers (Table S4). $T_{min\ origin}$ was calculated based on the WorldClim 2.0 database and averaged over target and partner population (spatial resolutions of 30 seconds, Fick & Hijmans, 2017; Table S1). Positive values of $\Delta\ T_{min}$ indicated transplanting toward warmer temperatures, and negative values a transplant toward colder temperatures. As we had no expectation about the sign of stressfulness, whether it was relevant or not, and whether its relation with heterosis was linear or quadratic, we first performed model selection based on AICc values. The fixed effects were either signed or absolute $\Delta\ T_{min}$, included an orthogonal quadratic term or not, and included no interaction or interaction terms. Covariates were mean-centered and square terms were calculated based on the mean-centered values. For the best model, we tested the significance of the effect of stress, genomic load, and their interaction by calculating the likelihood-ratio χ^2 using the *Anova* function (type III) of the package *car* (Fox & Weisberg, 2019).

To test whether heterosis is highest when range-edge populations with highest load are transplanted beyond range limits, we considered only the performance of six northern edge populations (ON8, ON11, ON12, NY4, NY5, NY6; in blue in Fig. 1) raised at the northern edge site (CG2) and the site beyond-northern edge (CG1), and the six southern edge populations (MO1, MO2, IN1, NC2, NC4 and VA1; in red in Fig. 1) raised at the southern edge site (CG4) and beyond the southern edge (CG5). Log_{10}-transformed heterosis was tested in hierarchical mixed-effect models based on restricted maximum likelihood. Crossed random effects were population and common garden, while fixed effects were the binary effect of transplant (edge = 0), the position of the edge (north = 0), as well as their interaction (model parametrization detailed in Method S2B).

In further analyses we used another estimate of stress, the performance decline at a site relative to the most favourable site (Reed *et al.,* 2012). For each target population in each common garden site, we calculated the relative decline in population-level performance as 1 - W_{WPC}/W_{max}., where W_{WPC} was the performance in each site averaged on both target and partner populations, and W_{max} was the highest W_{WPC} for the population pair within our experiment. Stress values close to 0 represent conditions of maximal performance, and values close to 1 represent maximal stress (*i.e.* $W_{WPC} = 0$). We performed model selection based on AICc values, comparing models including an orthogonal quadratic term of stress or not, with or without interaction with genomic load. On the best model, we then tested the effect of genomic load, stress and their interaction in similar analyses as described before. A further set of similar analyses as described above were performed on mean population WPC estimates of all 20 populations instead of heterosis. Here $T_{min\ origin}$ was calculated for each of the 20 populations.

Results

Pre-analysis on *multiplicative performance* on the level of the pot supported a role of cross type (WPC versus BPC) and its interaction with genomic load and environmental stress (Table S5), which motivated our focus on analyses based on population-level heterosis, on *MP I* and *II* up to year 3. Model selection on heterosis revealed that the best model was the one with signed stress, meaning that the difference in minimum temperature in early spring at the common garden site and the site of population origin ($\Delta\ T_{min} = T_{min\ CG} - T_{min\ origin}$) was allowed to be positive (warmer common garden) or negative (colder common garden), that considered the quadratic term for stress and included interactions (Table 1). The estimation of coefficients for the best model with *MP I* and *II* up to year 3 and λ revealed that genomic load and the linear term of stress had both a significant positive effect on heterosis (Table 2, Fig. 2a; results on *MP I* and *II* to year 2 reported in Table S6 for comparison). Heterosis increased with genomic load. Furthermore, heterosis was higher when the common garden was warmer compared to the site of origin of populations, *i.e.* with heat stress, but not when the common garden was colder. In contrast, the quadratic term of stress was not significant and interactions were generally not significant. Population-level heterosis based on *MP I* and *II* up to year 3 in the five common gardens ranged from -0.96 to 23.50 (mean: 1.88) and from -0.92 to 30.23 (mean: 2.65), respectively (Table S7). Heterosis based on λ ranged from -0.53 to 7.29 (mean: 0.73; Table S7).

For range-edge populations, heterosis based on *MP I* and *II* up to year 3, and λ, showed no significant variation between common gardens simulating respective range-edge conditions and common gardens simulating respective conditions beyond the edge (Table 3, Fig. 2c; results on *MP I* and *II* to year 2 reported in Table S8 for comparison). The position of the edge had a significant positive effect on heterosis based on *MP I* and *II* up to year 3, indicating that southern populations expressed generally higher heterosis. No significant interaction was observed

When stress was estimated based on relative performance decline compared to difference in temperature, the best model predicting heterosis was again one considering the quadratic term for

stress and including interactions (Table S9). Genomic load showed a significant positive effect on heterosis based on *MP I* and *II* up to year 3 *(*Table S10, results on *MP I* and *II* to year 2 reported for comparison). Heterosis significantly increased with stress (except *MP I* to year 3) and the quadratic term of stress (marginally significant for λ). The latter relationship pointed to an increase in heterosis both toward minimal and maximal stress. Genomic load and the quadratic term of stress further showed a significant positive interaction on heterosis based on λ, indicating that heterosis increased with increasing stress in populations with positive centered genomic load (*i.e.* genomic load > 0.86), but decreased in populations with negative centered genomic load.

Similar analyses to those testing the effect of environmental stress and genomic load on heterosis were performed on population-level WPC performance. Genomic load showed no significant effect on population-level WPC performance based on *multiplicative performance I* and *II* up to year 3, and λ, despite a consistent negative trend (Table S6, results on *MP I* and *II* to year 2 reported for comparison). Stress and the quadratic term of stress had both a significant negative effect on *MP I* and *II* up to year 3, and λ (Table S6, Fig. 2b). WPC performance decreases when plants were exposed to both warmer and colder temperatures, with a stronger effect toward warmer temperatures. The model-predicted decrease in WPC performance for *MP I* and *II* up to year 3 and λ was of 87%, 93% and 90% respectively when exposed to warmer temperatures (Table S11). When exposed to colder temperatures, *MP I* and *II* up to year 3 decreased of 14% and 7% respectively, while λ increased of 24% (Table S11). No significant interaction was observed

For range-edge populations, transplanting from common gardens simulating these populations respective range-edge conditions to common gardens simulating their respective beyond range-edge conditions significantly reduced only λ (Table S8, Fig. 2d; results on *MP I* and *II* to year 2 reported for comparison). The position of the edge interactions were generally not significant.

Discussion

Recent evolutionary theory proposes that species' geographic range limits result from the accumulation of mutational load in range-edge populations as a consequence of heightened levels of genetic drift (reviewed in Willi, 2019). Here we tested whether the expression of such mutational load is enhanced under stress (Reed *et al.,* 2012). In an experimental study on natural populations we found that heightened levels of genomic load and increased exposure to stress independently increased the expression of mutational load. Furthermore, while the exposure to warmer climates than the one experienced at the site of origin of populations led to heightened expression of mutational load, cooler climates led to lower heterosis. Overall, our study provides empirical evidences that edge populations, already suffering from high levels of genomic load, could suffer from additional load particularly under warmth stress, raising concern about the effects of rapidly warming climate on the persistence of small and isolated populations.

The expression of mutational load was generally stronger in populations with the highest levels of genomic load. Latter was found to accumulate in small populations with a history of long range expansion or isolation at the range limits of *A. lyrata* due to heightened exposure to drift (Willi *et al.,* 2018), resulting in increasing expression of mutational load toward range limits (Perrier *et al.,* 2020). The expression of mutational load was also higher in the populations most exposed to environmental stress in the transplant experiment, independently of the populations' levels of genomic load. The measure of environmental stress showed more consistent results on the expression of mutational load than stress based on relative performance. The latter also revealed increased expression of mutational load under stress, however following either or both linear and quadratic relationships, depending on which variable was assessed. Both calculations of stress are complementary: differences in environmental variables specifically assess the contribution of different types of stresses, while stress based on performance integrates other stressors that were not measured in our study, such as herbivory or soil properties. Most studies in the field of inbreeding depression do not test a range of stressors, but rather test benign conditions against stressful

conditions (Fox & Reed, 2010). Relative performance measures allow to characterize the strength of the stress applied. Here, our experimental design was based on the knowledge that our study organism was mostly constrained in its distribution by temperature (Lee-Yaw *et al.,* 2018), justifying using this variable as a main quantitative expression of stress. The results presented in this study represent some of the few empirical evidences that the expression of mutational load in natural populations is conditioned by environmental factors in addition of their evolutionary history. The environmental dependency of the expression of mutational load has often been suggested in previous studies due to similarity between heterosis and inbreeding depression (*e.g.* Oakley *et al.,* 2015), but has rarely been observed. Heterosis varied depending on the transplant environment in crosses between natural populations of *Chamaecrista fasciculate* with varying degrees of isolation (Fenster & Galloway, 2000), however the effect of the growth environment and resulting stress was not tested. Exposure to drought stress reduced heterosis in crosses between randomly paired natural populations of *Brassica nigra* in a recent greenhouse experiment (Prill *et al.,* 2014), linked to outbreeding depression counteracting heterosis. In contrast, drought stress increased heterosis based on survival and propensity to flower in crosses between native and invasive populations of *Mimulus guttatus* (Li *et al*., 2018), latter populations having experienced a mild bottleneck and likely to have accumulated mutational load. Agricultural research has also provided examples of heterosis variation under stress (reviewed in Blum, 2013, Fujimoto *et al.,* 2018), but with little context on population's demographic history. The strong effect of environmental stress on the expression of mutational load further highlights potential mechanisms for drift and environmental conditions to shape range limits.

The increased expression of mutational load under environmental stress was especially strong when populations were exposed to warmer climates than at their site of origin. This effect was supported by a strong decline of population mean multiplicative performance in within-population crosses, and a drop of 90% in demographic rates of within-population crosses when exposed to warmer climates. In contrast, the exposure to colder climates reduced the expression of mutational load, while leading to a decline in within-population crosses mean multiplicative performance, but a

25% increase in within-population cross demographic rates. These results suggest that populations exposed to warmer climates than at their site of origin were more stressed than when exposed to colder climates, latter potentially even increasing population persistence. Similar patterns were reported in a parallel study based on this transplant experiment (Sánchez-Castro *et al., in prep.*): the conditions at the northern edge and beyond were found to be not stressful, as performance and growth rates did not differ between sites at the centre of distribution, the northern edge and beyond, while performance and growth rates strongly declined at the southern edge and beyond. Furthermore, in *A. lyrata*, populations at lower latitudes are exposed to higher temperatures in summer but also to increased frost in winter, while populations at higher latitudes benefit from the snow cover to escape most frost events, therefore less frost tolerant than populations at lower latitudes (Wos & Willi, 2015). Transplanting frost sensitive northern populations in the southern sites most likely exposed them to increased frost, explaining the high stressfulness of transplanting toward warmer climates. On the contrary, southern populations transplanted in northern sites likely experienced lower stress than at their site of origin due to the longer snow cover, explaining the increased population growth rates when transplanted toward colder climates.

In our main analysis, genomic load and environmental stress both affected the expression of mutational load assessed over the three years, but did not interact. Previous studies in our study system showed that range-edge populations with the highest levels of genomic load suffer from increased levels of expressed load, linked to decrease in lifetime performance and demographic rates (Willi *et al.,* 2018; Perrier *et al.,* 2020), in line with predictions from simulation and genetic studies (reviewed in Willi, 2019). In Perrier *et al.* (2020), constraint on further colonization have been suggested to result from a genetic Allee effect (Luque *et al.,* 2016), due to additional accumulation of mutational load above tolerable levels to maintain population demographic rates. Both increase in expression of mutational load and decline in demographic rates under environmental stress suggest that range-edge populations could suffer from a second, stress-dependent genetic Allee effect, independent of the magnitude of mutational load. Small populations are more sensitive to random

environmental fluctuations, resulting in ecological Allee effects (Fauvergue *et al.,* 2012). Beyond range limits, where habitats are often unsuitable (Heargraves *et al.,* 2014; Lee-Yaw *et al.,* 2016), both genetic Allee effects could interact with ecological Allee effects in small populations (Wittmann *et al.,* 2016), precipitating their extinction faster than the sole magnitude of mutational load.

In populations originating from the northern and southern distribution edge, the expression of mutational load was however not affected when transplanting these populations from common gardens sites simulating their respective edge condition to sites simulating conditions beyond their respective edge. Demographic rates declined in both beyond-edge sites compared to their respective edge sites, implying that environmental conditions beyond the edge were stressful for range-edge populations, as suggested by the strong decline in habitat suitability at both northern and southern edge in our system (Lee-Yaw *et al.,* 2018). Several factors could explain why this specific stress did not trigger the stress-dependent genetic Allee effect described before. The extent to which environmental stress affects the expression of mutational load is dependent on the strength and the type of stress (Fox & Reed, 2010; Sandner & Matthies, 2016). The strength of the stress caused by transplanting beyond the edge might have been reduced by pre-adaptation in edge populations, if this stress does not differ in magnitude from conditions experienced in unfavorable years at range limits. While adaptive potential is predicted to be low in populations of small *Ne* (Weber, 1990; Willi *et al.,* 2006; Markert *et al.,* 2010; Polechová & Barton, 2015), especially under strong environmental gradient (Polechová, 2018), locally adapted populations could persist at range limits along environmental gradients despite expansion load (Gilbert *et al.,* 2017). In addition, frequent pre-exposure to stress caused by transplanting beyond the edge in the evolutionary history of edge populations could increase selection against deleterious alleles expressed under this particular type of stress (Hedrick, 1994; Bijlsma *et al.,* 1999; Plough, 2012; Enders & Nunney, 2016). Alternatively, the stress-dependent genetic Allee effect may have been masked by outbreeding depression resulting from outcrossing (Lynch, 1991), counteracting heterosis. In fact, outbreeding depression has also been reported to be sensitive to environmental conditions, reducing heterosis under stress (*e.g.* Prill

et al., 2014). In our experiment, nearly a fourth of the heterosis observation were negative in our study, suggesting the common occurrence of genetic incompatibilities such as the Dobzhansky-Muller type (Lynch, 1991; *e.g.* in Oakley *et al.,* 2015). Especially range-edge populations of *A. lyrata* exposed to higher genetic drift (Willi *et al.,* 2018) could have accumulated genetic incompatibilities leading to partial reproductive isolation, hence outbreeding depression (Orr & Turelli, 2001). Overall, while our results do not support an increased expression of mutational load due to stress in sites just beyond the range limits, the stress-dependent genetic Allee effect could be triggered by extreme environmental events occurring more frequently beyond range limits, and could impair long range dispersal in habitats less suitable than the immediate conditions beyond range limits. Future simulation studies based on range dynamic models (*e.g.* Gilbert *et al.,* 2017; Polechová, 2018) should therefore also consider the consequences of the additional stress-dependent genetic Allee effect on range expansion.

Our results contradict at the first glance the expectation that the magnitude of effect of stress on the expression of mutational load is dependent on the levels of genomic load. However, if the fitness effect of stress and genomic load were indeed independent *i. e.* if the proportion of fitness decline due to stress was similar for between- and within-population crosses, heterosis should not vary under stress. In our secondary analysis, genomic load and stress based on relative performance decline showed an interaction on heterosis based on demographic rates, suggesting that the interaction with environmental stress could be masked by other processes impacting heterosis. Stronger outbreeding depression in edge populations could lower heterosis measures as discussed above. In addition, populations with the lowest estimates of load may show higher heterosis under stress as expected. In *A. lyrata*, populations with the lowest levels of genomic load occur predominantly in suitable habitats (Lee-Yaw *et al.,* 2018). This could lead to the accumulation of mutation on genes not under selection in benign conditions, but expressed under stress (Hoffmann & Merilä, 1999). The increase in outbreeding depression in populations bearing high levels of genomic load and the increased heterosis in populations with lower levels of genomic load could artificially mask the

interaction between environmental stress and genomic load. In addition, genomic load and environmental stress interacted when tested on heterosis based on multiplicative performances up to the second year, suggesting that the processes masking the interaction gain in magnitude later in life. This supports the involvement of either outbreeding depression or heterosis as described above, the expression of different types of load being generally stronger in late life stages (Husband & Schemske, 1996). Already outlined by previous studies (Oakley *et al.*, 2015; Fujimoto *et al.*, 2018), more knowledge on the genetic basis of the environmental dependency of heterosis and outbreeding depression is required to fully understand their interplay in natural populations.

Our study is one of the first to empirically highlight how the expression of mutational load is dependent on the exposure to environmental stress, independently of population's levels of genomic load. The magnitude of this second genetic Allee effect might very well contribute in constraining range expansion in unfavorable habitats, in addition to additional accumulation of mutational load. The effect of environmental stress on the expression of mutational load was especially strong when populations were exposed to warmer conditions than at their location of origin, raising concerns for the persistence of small isolated populations under climate change. The rapidly declining habitat suitability at the warmer range limits could precipitate the extinction of already fragile rear-edge populations (Hampe & Petit, 2005; Lenoir & Svenning, 2013). At the leading edge, while new suitable habitats should be available, colonization could still be impaired by genomic load. The conjoint effect of both genetic Allee effects could lead to drastic range contraction rather than shift predicted from range dynamics simulations.

Acknowledgements: This work was supported by the Swiss National Science Foundation (31003 A_166322). We are grateful to Celia Evans (Paul Smith's College, Paul Smith, NY), Joan Edwards (Williams College, Williamstown, MA), Heather Peckham Griscom (James Madison University, Harrisonburg, VA), William K. Smith (Wake Forest University, Winston-Salem, NC) and Rodney Mauricio (University of Georgia, Athens, GA) for logistical support in the US. For field assistance we thank Mary Anderson, Michael Boyd, Bennet Coe, Scott Cory, Rachel Hillyer, Andrew Jones, Deidre Keating, Larry Kummer, David Lampman, Anastasia Levie-Sprick, Blake Macko, Shannon Malisson, Kathryn McGee, Althea Neighbors, Debra Rogers-Gillig, Caleb Rose, Amber Scarabaggio, Anna Shutley, Caroline Vath and Audrey Werner. For assistance with seed counts we thank Olivier Bachmann, Markus Funk and Susanna Riedl. Collection permits were provided by the Clinton County Conservation Board, Cornell University, Fort Leonard Wood Army Base, Iowa Department of Natural Resources, Missouri Department of Conservation, New York State Office of Parks, Ontario Parks, Palisades Interstate Park Commission, Rock Island Lodge, United States National Park Service, Virginia Department of Conservation and Recreation and the Wisconsin Department of Natural Resources.

Data accessibility statement: All data is stored in Dryad (https://doi.org/10.5061/dryad.cc2fqz642)

References

Agrawal AF, Whitlock MC. 2010. Environmental duress and epistasis: how does stress affect the strength of selection on new mutations. *Trends in Ecology and Evolution* **25**: 450–458.

Armbruster P, Reed DH. 2005. Inbreeding depression in benign and stressful environments. *Heredity* **95**: 235–242.

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**: 1–48.

Bijlsma R, Bundgaard J, Van Putten WF. 1999. Environmental dependence of inbreeding depression and purging in *Drosophila melanogaster*. *Journal of Evolutionary Biology* **12**: 1125–1137.

Bergman A, Siegal ML. 2003. Evolutionary capacitance as a general feature of complex gene networks. *Nature* **424**: 549–52.

Blum A. 2013. Heterosis, stress, and the environment: a possible road map towards the general improvement of crop yield. *Journal of Experimental Botany* **64**: 4829–4837.

Bosshard L, Dupanloup I, Tenaillon O, Bruggmann R, Ackermann M, Peischl S, Excoffier L. 2017. Accumulation of deleterious mutations during bacterial range expansions. *Genetics* **217**: 669–684.

Brown JH. 1984. On the relationship between abundance and distribution of species. *The American Naturalist* **124**: 255–279.

Cahill AE, Aiello-Lammens ME, Fisher-Reid MC, Hua X, Karanewsky CJ, Ryu HY, Sbeglia CG, Spagnolo F, Waldron JB, Wien JJ. 2014. Causes of warm-edge range limits: systematic review, proximate factors and implications for climate change. *Journal of Biogeography* **41**: 429–442.

Caswell H. 2001. *Matrix population models, 2nd edn*. Sunderland, MA, USA: Sinauer Associates

Connallon T, Sgrò CM. 2018. In search of a general theory of species' range evolution. *PLoS Biology* **16**: e2006735.

Crow JF. 1987. Muller, Dobzhansky, and overdominance. *Journal of the History of Biology* **20**: 351–380.

Eckert CG, Samis KE, Lougheed SC. 2008. Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology* **17**: 1170–1188.

Elena SF, De Visser JAGM. 2003. Environmental stress and the effects of mutation. *Journal of Biology* **2**: 12.

Enders LS, Nunney L. 2016. Reduction in the cumulative effect of stress-induced inbreeding depression due to intragenerational purging in *Drosophila melanogaster*. *Heredity* **116**: 304–313.

Fauvergue X, Vercken E, Malausa T, Hufbauer RA. 2012. The biology of small, introduced populations, with special reference to biological control. *Evolutionary Applications* **5**: 424–443.

Fenster CB, Galloway LF. 2000. Population differentiation in an annual legume: genetic architecture. *Evolution* **54**: 1157–1172.

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**: 4302–4315.

Fox CW, Reed DH. 2011. Inbreeding depression increases with environmental stress: an experimental study and meta-analysis. *Evolution* **65**: 246–258.

Fox J, Weisberg S. 2019. *An R Companion to Applied Regression, 3rd edn*. Thousand Oaks, CA, USA: Sage.

Fujimoto R, Uezono K, Ishikura S, Osabe K, Peacock WJ, Dennis ES. 2018. Recent research on the mechanism of heterosis is important for crop and vegetable breeding systems. *Breeding Science* **68**: 145–158.

Gaston K. 2009. Geographic range limits of species. *Proceedings of the Royal Society B: Biological Sciences* **276**: 1391–1393.

Gilbert KJ, Sharp NP, Angert AL, Conte GL, Draghi JA, Guillaime F, Hargreaves AL, Matthey Doret R, Whitlock MC. 2017. Local adaptation interacts with expansion load during range expansion: maladaptation reduces expansion load. *The American Naturalist* **189**: 368–380.

González-Martínez SC, Ridout K, Pannell JR. 2017. Range expansion compromises adaptive evolution in an outcrossing plant. *Current Biology* **27**: 2544–2551.

Griffin PC, Willi Y. 2014. Evolutionary shifts to self-fertilisation restricted to geographic range margins in North American *Arabidopsis lyrata*. *Ecology Letters* **17**: 484–490.

Hadfield JD. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Softwares* **33**: 1–22.

Hadfield JD. 2019. MCMCglmm course notes. [WWW document] URL https://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf [accessed 24 June 2020].

Hampe A, Petit RJ. 2005. Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters* **8**: 461–467.

Hargreaves AL, Samis KE, Eckert CG. 2014. Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. *The American Naturalist* **183**: 157–173.

Hedrick PW. 1994. Purging inbreeding depression and the probability of extinction: full-sib mating. *Heredity* **73**: 363–372.

Hewitt GM. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405:** 907–913.

Hewitt GM. 2004. Genetic consequences of climatic changes in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences* **359**: 183–195.

Henn BM, Botigué LR, Peischl S, Dupanloup I, Lipatov M, Maples BK, Martin AR, Musharoff S, Cann H, Snyder MP. 2016. Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. *Proceedings of the National Academy of Sciences of the United States of America* **113**: E440–E449.

Henry RC, Barton KA, Travis JMJ. 2015. Mutation accumulation and the formation of range limits. *Biology Letters* **11**: 20140871–20140871.

Hoffmann AA, Merilä J. 1999. Heritable variation and evolution under favourable and unfavourable conditions. *Trends in Ecology and Evolution* **14**: 96–101.

Hoffmann AA, Hercus MJ. 2000. Environmental stress as an evolutionary force. *BioScience* **50**: 217–226.

Husband BC, Schemske DW. 1996. Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution* **50**: 54–70.

Kimura M, Maruyama T, Crow JF. 1963. The mutation load in small populations. *Genetics* **48**: 1303–1312.

Kondrashov AS, Houle D. 1994. Genotype-environment Interactions and the estimation of the genomic mutation rate in Drosophila melanogaster. *Proceedings of the Royal Society B: Biological Sciences* **258**: 221–227.

Korn M, Gärtner T, Erban A, Kopka J, Selbig J, Hincha DK. 2010. Predicting *Arabidopsis* freezing tolerance and heterosis in freezing tolerance from metabolite composition. *Molecular Plant* **3**: 224–235.

Koski MH, Layman NC, Prior CJ, Busch JW, Galloway LF. 2019. Selfing ability and drift load evolve with range expansion. Evolution Letters **3**: 500–512.

Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: Tests in linear mixed-effects models. *Journal of Statistical Software* **82**: 1–26.

Lee-Yaw JA, Kharouba HM, Bontrager M, Mahony C, Csergő AM, Noreen AM, Li Q, Schuster R Angert AL. 2016. A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. *Ecology Letters* **19**: 710–722.

Lee-Yaw JA, Fracassetti M, Willi Y. 2018. Environmental marginality and geographic range limits: a case study with *Arabidopsis lyrata* ssp. *lyrata*. *Ecography* **41**: 622–634.

Lenoir J, Svenning JC. 2013. Latitudinal and elevational range shifts under contemporary climate change. In S. A. Levin, eds. *Encyclopedia of Biodiversity, second edition.* Academic Press. Encyclopedia of Biodiversity, 599–611.

Li Y, Stift M, van Kleunen M, Bonser S. 2018 Admixture increases performance of an invasive plant beyond first-generation heterosis. *Journal of Ecology* **106**: 1595–1606.

Liao W, Reed DH. 2009. Inbreeding-environment interactions increase extinction risk. *Animal Conservation.* **12**: 54–61.

Lohmueller KE, Indap AR, Schmidt S, Boyko AR, Hernandez RD, Hubisz MJ, Sninsky JJ, White TJ, Sunyaev SR, Nielsen R, et al. 2008. Proportionally more deleterious genetic variation in European than in African populations. *Nature* **451**: 994–998.

Louthan AM, Daniel FD, Angert AL. 2015. Where and when do species interactions set range limits? *Trends in Ecology and Evolution* **30**: 780–792.

Luque GM, Vayssade C, Facon B, Guillemaud T, Courchamp F, Fauvergue X. 2016. The genetic Allee effect: a unified framework for the genetics and demography of small populations. *Ecosphere* **7**: e01413.

Lynch M. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution* **45**: 622–629.

Lynch M, Conery J, Bürger R. 1995. Mutational meltdowns in sexual populations. *Evolution* **49**: 1067–1080.

Markert, JA, Champlin DM, Gutjahr-Gobell R, Grear JS, Kuhn A, McGreevy TJ Jr, Roth A, Bagley M, Nacci DE. 2010. Population genetic diversity and fitness in multiple environments. *BMC Evolutionary Biology* **10**: 205.

Miller M, Song Q, Shi X, Juenger TE, Chen ZJ. 2015. Natural variation in timing of stress-responsive gene expression predicts heterosis in intraspecific hybrids of *Arabidopsis*. *Nature Communications* **6**: 7453.

Oakley CG, Ågren J, Schemske DW. 2015. Heterosis and outbreeding depression in crosses between natural populations of *Arabidopsis thaliana*. *Heredity* **115**: 73–82.

Orr HA, Turelli M. 2001. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution* **55**: 1085–1094.

Peischl S, Excoffier L. 2015. Expansion load: Recessive mutations and the role of standing genetic variation. *Molecular Ecology* **24**: 2084–2094.

Peischl S, Dupanloup I, Kirkpatrick M, Excoffier L. 2013. On the accumulation of deleterious mutations during range expansions. *Molecular Ecology* **22**: 5972–5982.

Peischl S, Kirkpatrick M, Excoffier L. 2015. Expansion load and the evolutionary dynamics of a species range. *The American Naturalist* **185**: E81–E93.

Peischl S, Dupanloup I, Foucal A, Jomphe M, Bruat V, Grenier JC, Gouy A, Gilbert KJ, Gbeha E, Bosshard L, et al. 2018. Relaxed selection during a recent human expansion. *Genetics* **208**: 763–777.

Perrier A, Sánchez-Castro D, Willi Y. 2020. Expressed mutational load increases toward the edge of a species' geographic range. *Evolution* **74**: 1711–1723.

Pironon S, Papuga G, Villellas J, Angert AL, García MB, Thompson JD. 2017. Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biological Reviews* **92**: 1877–1909.

Plough LV. 2012. Environmental stress increases selection against and dominance of deleterious mutations in inbred families of the Pacific oyster *Crassostrea gigas*. *Molecular Ecology* **21**: 3974–3987.

Polechová J. 2018. Is the sky the limit? On the expansion threshold of a species range. *PLoS Biology* **16**: e2006735.

Polechová J, Barton NH. 2015. Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences of the USA* **112**: 6401–6406.

Prill N, Bullock JM, van Dam NM, Leimu R. 2014. Loss of heterosis and family-dependent inbreeding depression in plant performance and resistance against multiple herbivores under drought stress. *Journal of Ecology* **102**: 1497–1505.

Queitsch C, Sangster TA, Lindquist S. 2002. Hsp90 as a capacitor of phenotypic variation. *Nature* **417**: 618–624.

R Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reed DH, Fox CW, Enders LS, Kristensen TN. 2012. Inbreeding–stress interactions: evolutionary and conservation consequences. *Annals of the New York Academy of Sciences* **1256**: 33-48.

Roff DA. 1997. *Evolutionary Quantitative Genetics*. New York, NY, USA: Chapman and Hall.

Roy K, Hunt G, Jablonski D, Krug AZ, Valentine JW. 2009. A macroevolutionary perspective on species range limits. *Proceedings of the Royal Society B: Biological Sciences* **276**: 1485–1493.

Rutherford SL, Lindquist S. 1998. Hsp90 as a capacitor for morphological evolution. *Nature* **396**: 336–342.

Sandner TM, Matthies D. 2016. The effects of stress intensity and stress type on inbreeding depression in *Silene vulgaris*. *Evolution* **70**: 1225–1238.

Sexton JP, McIntyre PJ, Angert AL, Rice KJ. 2009. Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics* **40**: 415–436.

Vermeulen CJ, Bijlsma R. 2004. Characterization of conditionally expressed mutants affecting age-specific survival in inbred lines of *Drosophila melanogaster*: lethal conditions and temperature-sensitive periods. *Genetics* **167**: 1241–1248.

Waller DM, Dole J, Bersch AJ. 2008. Effects of stress and phenotypic variation on inbreeding depression in *Brassica rapa*. *Evolution* **62**: 917–931.

Weber KE. 1990. Increased selection response in larger populations. I Selection for wing-tip height in *Drosophila melanogaster* at three population sizes. *Genetic* **125**: 579–84.

Willi Y. 2019. The relevance of mutation load for species range limits. *American Journal of Botany* **106**: 757–759.

Willi Y, Määttänen K. 2010. Evolutionary dynamics of mating system shifts in *Arabidopsis lyrata*. Journal of Evolutionary Biology **23**: 2123–2131.

Willi Y, Van Buskirk J. 2019. A practical guide to the study of distribution limits. *The American Naturalist* **193**: 773–785.

Willi Y, Van Buskirk J, Hoffmann AA. 2006. Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics* **37**: 433–78.

Willi Y, Dietrich S, Kleunen MV, Fischer M. 2007. Inter-specific competitive stress does not affect the magnitude of inbreeding depression. *Evolutionary Ecology Research* **9**: 959–974.

Willi, Y., M. Fracassetti, S. Zoller, and J. Van Buskirk. 2018. Accumulation of mutational load at the edges of a species range. *Molecular Biology and Evolution* **35**: 781–791.

Wittmann MJ, Stuis H, Metzler D, Kuparinen A. 2016. Genetic Allee effects and their interaction with ecological Allee effects. *Journal of Animal Ecology* **87**: 11–23.

Wos G, Willi Y. 2015. Temperature-stress resistance and tolerance along a latitudinal cline in North American *Arabidopsis lyrata*. *PLoS ONE* **10**: e0131808.

Zhang M, Zhou L, Bawa R, Suren H, Holliday JA. 2016. Recombination rate variation, hitchhiking, and demographic history shape deleterious load in poplar. *Molecular Biology and Evolution* **33**: 2899–2910.

Tables:

Table 1: Model selection on potential predictors of heterosis

Model	***AICc***	***Δi***	***wi***
Genomic load + stress + stress^2 + (genomic load * stress) + (genomic load * stress^2)	115.2	0.0	0.93
Genomic load + stress + stress^2	121.0	5.8	0.05
Genomic load	123.3	8.2	0.02
Genomic load + stress	129.1	14.0	0.00
Genomic load + |stress|	131.2	16.0	0.00
Genomic load + stress + (genomic load * stress)	134.0	18.8	0.00
Genomic load + |stress| + (genomic load * |stress|)	135.2	20.0	0.00

The dependent variable was heterosis based on *multiplicative performance I* to year 3. Genomic load was the ratio of genome-wide non-synonymous to synonymous polymorphic sites adjusted by their mean derived ferquency (Willi *et al.*, 2018), and stress was the difference in minimum temperature in spring between common garden and site of origin ($\Delta\ T_{min} = T_{min\ CG} - T_{min\ origin}$). Models are ranked from best to worst fit, with lower *AICc* values indicating a better fit. The difference between the best fit model and the others is indicated as *Δi*, while the weight of each model is indicated by *wi*.

Table 2: Summary of mixed-effects models testing for the effect of the genomic estimate of mutational load, environmental stress ($\Delta T_{min} = T_{min\ CG} - T_{min\ origin}$) and their interaction on population-level heterosis based on *multiplicative performance* (*MP*) up to 3 years or the finite rate of increase, λ, at five common garden sites

		Genomic load (GL)			Stress (S)			Stress^2 (S^2)			GL * S			GL * S^2				
Dependent variable	***N***	***β***	***χ²***		***β***	***χ²***		***β***	***χ²***		***β***	***χ²***		***β***	***χ²***		***R^2m***	***R^2c***	
MP I to year 3	89	**1.89**	**6.63**	*	**1.03**	**5.77**	*	-0.73	2.98	(*)	-3.38	0.48		3.78	0.70		0.19	0.38	†
MP II to year 3	89	**1.96**	**7.46**	**	**1.76**	**19.07**	***	-0.26	0.45		-1.48	0.10		6.89	2.62		0.28	0.47	†
λ (Heterosis)	89	**0.82**	**5.30**	*	**0.79**	**15.45**	***	-0.08	0.16		0.14	0.00		1.72	0.75		0.24	0.47	†

Population heterosis estimates were $\log_{10}$-transformed prior to analysis. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include regression coefficient (β), χ^2-value of each fixed effect and the marginal and conditional R^2 of the model. Model fits with significant (positive) intercept are indicated by †. Regression coefficient with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table 3: Summary of mixed-effects models testing for the effect of transplant (beyond the edge compared to edge [0]), edge position (south compared to north [0]) and their interaction on population-level heterosis based on *multiplicative performance* (*MP*) up to 3 years or the finite rate of increase, λ, of range-edge populations transplanted in edge and beyond edge common garden sites

		Transplant		Edge position			Transplant * edge position			
Dependent variable	***N***	β	χ^2	β	χ^2		β	χ^2	R^2m	R^2c
MP I to year 3	24	0.10	0.16	**0.55**	**5.09**	*	-0.15	0.18	0.25	0.25
MP II to year 3	24	0.14	0.34	**0.57**	**5.28**	*	-0.12	0.12	0.28	0.28
λ (Heterosis)	24	0.05	0.04	0.17	0.53		-0.04	0.01	0.09	0.40

Population heterosis estimates were $\log_{10}$-transformed prior to analysis. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include regression coefficient (β), χ^2-value of each fixed effect and the marginal and conditional R^2 of the model. Regression coefficient with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Figures

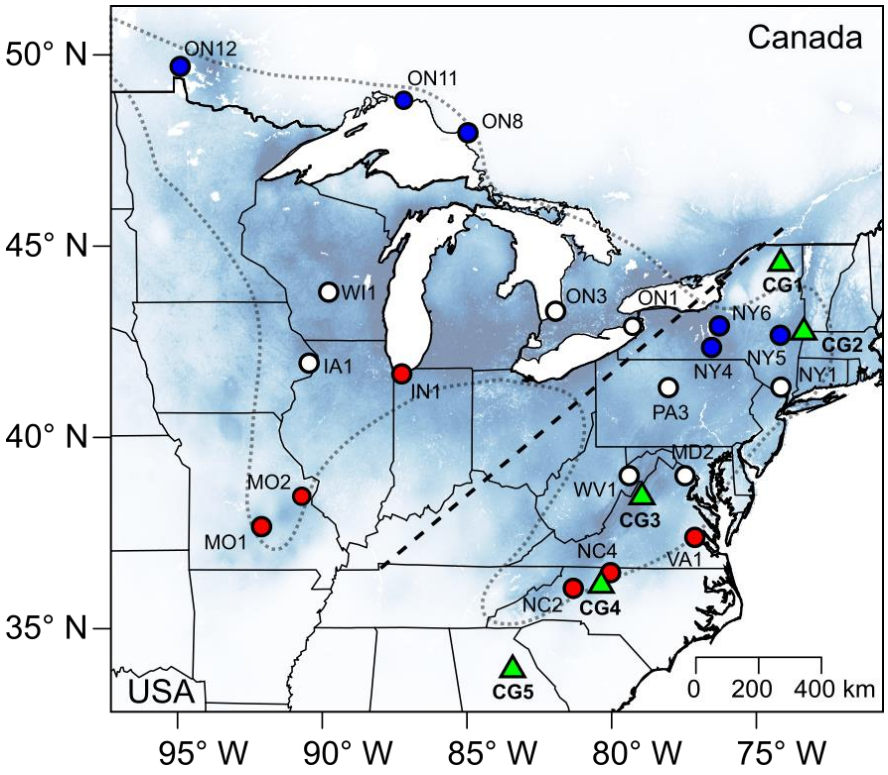

Figure 1: Map of the distribution of *Arabidopsis lyrata* in eastern North America, with information on habitat suitability and the location of the 20 populations studied and the 5 common garden sites. The range of *A. lyrata* is represented by a dotted line, habitat suitability by shades of blue, with darker blue indicating higher suitability (as in Lee-Yaw *et al.*, 2018). Populations are shown by circles, with abbreviations for state (USA) or province (Canada) and a number (as in Willi *et al.*, 2018). Blue and red circles represent northern and southern edge populations in our analysis. Green triangles represent the five common garden (CG) sites; numbers added to labels are in sequence of north to south. State outlines for the USA are shown, and the split between eastern and western genetic cluster is represented by the dashed line. Of the 20 populations, two were used as partner-populations for between-population crosses, NY1 for crosses with eastern populations, and IA1 for crosses with western populations.

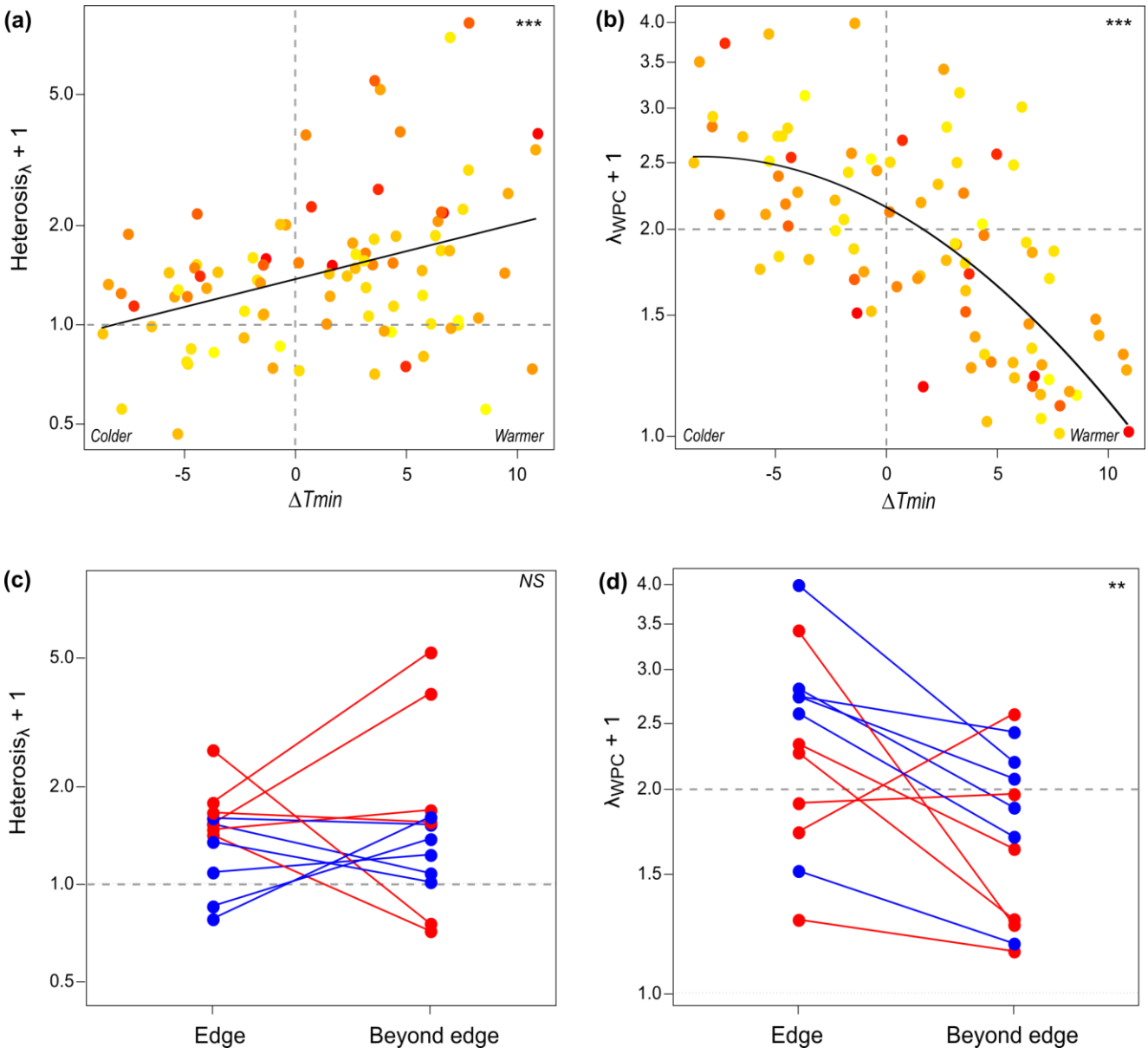

Figure 2: Environment dependence of the expression of mutational load. (a) and (c): Population heterosis was estimated based on the intrinic rate of increase, λ. The horizontal grey dashed line represents the value at which heterosis drops below 0, indicating outbreeding depression. **(b) and (d):** λ was calculated based on performance of within-population crosses (WPC). The horizontal grey dashed line represents the value at which λ drops below 1, indicating a decline in population size. **(a) and (b):** Environmental stress ($\Delta\ T_{min}$) was the difference in minimum temperature in early spring between common garden minus that at the site of origin. The black lines represent model-predicted slopes (from test statistics Table 2, Table S6). The vertical grey dashed line represents the transition between a negative and a positive $\Delta\ T_{min}$, indicating respectively a transplant toward colder and warmer sites. Genomic estimate of mutational load ($P_n f_n / P_s f_s$) is represented in shades of yellow (low) to red (high). Genomic load was the ratio of genome-wide non-synonymous to synonymous polymorphic sites adjusted by their mean derived frequency (as in Willi *et al.*, 2018). **(c) and (d):**

The effect of transplanting beyond the range edge was tested for both northern (blue) and southern (red) populations at their respective northern and southern edge (CG2, CG5) and beyond the edge (CG1, CG5). Lines indicate the change in heterosis or WPC performance of each populaton. Test statistics are reported in Table 3 and Table S8. Significance of the effect of $\Delta\ T_{min}$ or transplant is indicated as: *NS* $P>0.1$, (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Supporting information

Method S1: Parametrization of priors, and hierarchical mixed-effects model analyzed in a Bayesian (MCMC) framework, with individual *multiplicative performance I* up to year 3 as dependent variable

Method S2: Parametrization of the hierarchical mixed-effects models, with heterosis as dependent variable

Figure S1: Estimation of population growth rate

Table S1: Information on the *Arabidopsis lyrata* populations studied

Table S2: Summary of the crossing experiment and of the seeds sown in each common garden

Table S3: Information on the common garden sites

Table S4: Environmental stress per site

Table S5: Results of analysis performed on the level of individual pots, with a model testing for the effect of cross type (between- compared to within-population crosses [0]), genomic estimate of mutational load, environmental stress ($\Delta\ T_{min} = T_{min\ CG} - T_{min\ origin}$) and their interactions on *multiplicative performance I* up to year 3, at five common garden sites

Table S6: Summary of mixed-effects models testing for the effect of the genomic estimate of mutational load, environmental stress ($\Delta\ T_{min} = T_{min\ CG} - T_{min\ origin}$) and their interaction on population-level heterosis or within-population cross performance based on *multiplicative performance* (*MP*) up to year 2 or 3, and the finite rate of increase (λ), at five common garden sites

Table S7: Summary of population-level heterosis or within-population cross performance based on *multiplicative performance* (*MP*) up to year 3 or 2, and the finite rate of increase (λ), at five common garden sites

Table S8: Summary of mixed-effects models testing for the effect of transplant (beyond the edge compared to edge [0]), edge position (south compared to north [0]), and their interaction on population-level heterosis or within-population cross performance based on *multiplicative performance* (*MP*) up to year 2 or 3, and the finite rate of increase (λ), at five common garden sites

Table S9: Model selection on potential predictors of heterosis

Table S10: Summary of mixed-effects models testing for the effect of the genomic estimate of mutational load, stress based on relative performance, and their interaction on population-level heterosis based on multiplicative performance (MP) up to year 3 or 2 and the finite rate of increase (λ), at five common garden sites

Table S11: Magnitude of effect of the genomic estimate of mutational load and environmental stress ($\Delta\, T_{\text{min}} = T_{\text{min CG}} - T_{\text{min origin}}$) on ($\log_{10}$-transformed) *multiplicative performance (*MP) up to year 3 or the finite rate of increase, λ, of within-population crosses (WPC), at five common garden sites

Method S1: Parametrization of priors, and hierarchical mixed-effects model analyzed in a Bayesian (MCMC) framework, with individual *multiplicative performance I* up to year 3 as dependent variable

S1A: Priors

Priors were set to be weak, using parameter expansion to improve convergence. *R* specifies the priors for the fixed effects, *G* specifies the priors for the random effects.

```
priors.model=list(
  R=list(V=diag(2), n=1, fix = 2),
  G=list(G1=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2),
        G2=list(V=diag(4), n=4, alpha.mu = rep(0,4),alpha.V = diag(4)*25^2),
        G3=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2),
        G4=list(V=diag(4), n=4, alpha.mu = rep(0,4),alpha.V = diag(4)*25^2),
        G5=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2),
        G6=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2)))
```

S1 B: Parametrization of hierarchical mixed-effects models analyzed in a Bayesian framework

Multiplicative performance estimates were split into two parts: the *zero_part*, a binary transformation of performance estimates with *zero_part* = 1 if performance > 0, or else *zero_part* = 0; and the *norm_part* containing only the $\log_{10}$ transformed performance measures if *zero_part* = 1.

```
model = MCMCglmm(cbind(norm_part, zero_part) ~ trait -1 + trait:cross type * trait:genomic load
* trait:poly(stress, degree = 2),
            random = ~ us(trait):maternal population
            + us(trait:cross type):maternal population
            + us(trait): maternal population: maternal family
            + us(trait: cross type):maternal population:maternal family
            + us(trait):common garden + us(trait):common garden:block,
            prior = priors.model, rcov = ~idh(trait):units,
            family=c('gaussian', 'categorical'),
            burnin = 5000, thin = 100, nitt = 50000,
            data=data)
```

Method S2: Parametrization of the hierarchical mixed-effects models, with heterosis as dependent variable

S2A: Hierarchical mixed-effects model testing for the effect of stress and genomic load

Model = lmer($\log_{10}$(heterosis + 1) ~ genomic load * poly(stress, degree = 2)
+ (1 | maternal population)
+ (1 | common garden),
data = data)

S2B: Hierarchical mixed-effects model testing for the effect of transplant and genomic load

Model = lmer($\log_{10}$(heterosis + 1) ~ transplant * edge posisition
+ (1 | maternal population)
+ (1 | common garden),
data = data)

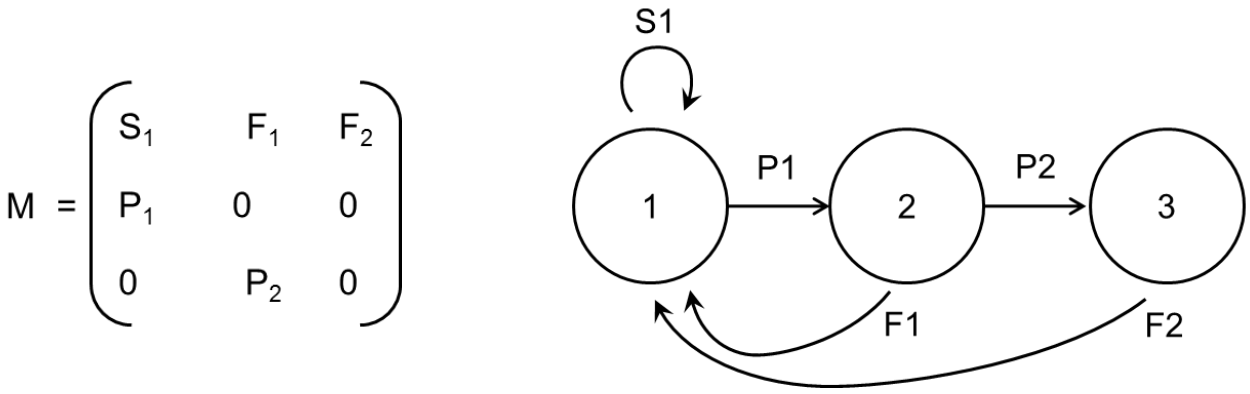

Figure S1: Estimation of population growth rate. We estimated population growth rate for all 20 WPC and 18 BPC combinations by constructing stage-classified matrices (**left**) at the level of the cross type of each population in each common garden. Each matrix was constructed to reflect the life cycle of individuals assessed in the common garden over three years. The life cycle (**right**) was composed of three stages: 1–healthy seeds (sown in 2017), 2–individuals capable of reproducing in spring of year 2 (2018), 3–individuals capable of reproducing in spring of year 3 (2019), each with a projection interval of one year. The transition from stage 1 to 2 (*P1*) was estimated as: proportion of germinated seedlings (2017) x survival of germinated seedlings until the date of first flowering in the first reproductive period (year 2) at each site. Transition from stage 2 to 3 (*P2*) was estimated as the survival of individuals alive at stage 2 to the date of recording of reproductive output in the second reproductive period (year 3). Seeds that did not germinate in the first year (transition from step 1 to 2) were assumed to remain at the same stage (*S1*) and to contribute to the seed pool of stage 1 in the next years. *S1* was defined as the probability of non-germinated seeds to survive over one winter. This estimate was calculated based on the seed survival experiment over one winter (2018 – 2019). We assumed no plants survived after the third year, and set the probability to remain at stage 2 or 3 at 0. Allowing individuals to remain in stage 3 indefinitely with a probability to survive each year equalling *P2* did not significantly affect the population growth rates (data not shown). Fecundity of stage 2 (*F1*) and 3 (*F2*) were estimated separately as: probability to reproduce * number of fruits * number of healthy seeds per fruit. The latter should reflect the effect of natural conditions on seed formation, as well as the chance of seeds landing in suitable environments to germinate. As these factors could not be estimated in the field experiment, number of fruits was a standard value representing both number of healthy seeds per fruit and the probability to land in an environment suitable for germination, calculated to yield an average λ of 1 across all WPC over all sites.

Table S1: Information on the *Arabidopsis lyrata* populations studied

Population	Latitude [° N]	Longitude [° W]	Ecological variable	Variables of population history	
			Min. temp. early spring [°C] †	Cluster	Genomic load ($P_n f_n / P_s f_s$) ‡
IA1	41.97	90.37	-0.9	West	0.8069
IN1	41.61	87.19	0.0	West	0.83283
MD2	38.99	77.25	3.8	East	0.78476
MO1	37.72	92.06	4.5	West	0.90542
MO2	38.47	90.71	4.0	West	1.03637
NC2	36.04	81.16	3.8	East	0.91391
NC4	36.41	79.96	5.0	East	0.8641
NY1	41.3	73.98	0.4	East	0.77297
NY4	42.35	76.39	-2.5	East	0.77737
NY5	42.66	74.02	-3.0	East	0.78503
NY6	43	76.09	-2.1	East	0.77053
ON1	42.87	79.18	-1.7	West	0.96393
ON11	48.77	87.13	-7.9	West	1.0927
ON12	49.65	94.92	-7.8	West	0.85352
ON3	43.26	81.84	-2.6	West	0.86054
ON8	47.93	84.85	-7.5	West	0.88166
PA3	41.28	77.87	-1.4	East	0.7618
VA1	37.42	77.02	5.5	East	0.81838
WI1	43.83	89.72	-3.3	West	0.73834
WV1	38.96	79.29	1.1	East	0.82191

† Data extracted from WorldClim database version 2.0 (Fick and Hijmans, 2017). ‡ Willi *et al.,* 2018.

Table S2: Summary of the crossing experiment and of the seeds sown in each common garden

Mother population	Father population	No. of cross families	No. of seeds sown in each common garden					Cross type
			CG1	CG2	CG3	CG4	CG5	
NY1	NY1	9	70	69	63	69	66	WPC
NY5	NY5	12	72	72	72	72	72	WPC
IN1	IN1	11	72	72	72	72	72	WPC
MO1	MO1	12	72	72	69	72	72	WPC
ON11	ON11	12	72	69	66	72	66	WPC
ON12	ON12	12	69	69	72	72	72	WPC
MO2	MO2	12	72	72	72	72	72	WPC
NC2	NC2	12	69	70	72	72	72	WPC
NC4	NC4	11	72	72	72	72	72	WPC
IA1	IA1	9	70	72	57	72	66	WPC
VA1	VA1	9	72	72	72	72	72	WPC
MD2	MD2	10	69	69	69	72	69	WPC
WV1	WV1	12	69	69	66	72	60	WPC
PA3	PA3	11	72	72	72	72	72	WPC
NY6	NY6	12	72	72	66	72	63	WPC
NY4	NY4	10	70	69	69	72	72	WPC
WI1	WI1	10	66	72	72	72	72	WPC
ON8	ON8	6	27	30	27	37	27	WPC
ON3	ON3	8	69	69	69	72	66	WPC
ON1	ON1	12	72	72	72	72	69	WPC
NY5	NY1	12	72	72	72	72	72	BPC
IN1	IA1	11	72	69	69	72	72	BPC
MO1	IA1	12	72	72	72	72	72	BPC
ON11	IA1	9	55	54	54	*0*	72	BPC †
ON12	IA1	10	72	72	72	72	72	BPC
MO2	IA1	12	72	72	72	72	72	BPC
NC2	NY1	12	69	70	69	69	72	BPC
NC4	NY1	10	72	72	72	72	72	BPC
VA1	NY1	10	72	72	72	78	72	BPC
MD2	NY1	9	72	72	72	72	72	BPC
WV1	NY1	11	72	72	66	72	66	BPC
PA3	NY1	11	72	72	72	72	72	BPC
NY6	NY1	11	72	72	72	72	72	BPC
NY4	NY1	11	72	72	72	72	72	BPC
WI1	IA1	11	78	72	72	72	72	BPC
ON8	IA1	5	34	35	33	39	33	BPC
ON3	IA1	10	63	63	63	72	63	BPC
ON1	IA1	12	72	72	72	72	69	BPC

Total number of seeds sown: 12,933; total number of cross combinations: 401; total number of population-level WPC: 100; total number of population-level BPC: 89. † ON11 x IA1 missing in CG4 due to too low numbers of seeds after restarting the experiment in December. Heterosis was not calculated for this cross-site-combination

Table S3: Information on the common garden sites

Transplant site	Location	Latitude [° N]	Longitude [° W]	Min. temperatures early spring (°C) †
CG1 (NY)	Beyond northern edge	44.51	74.02	-6.5
CG2 (MA)	Northern edge	42.72	73.22	-3.3
CG3 (VA)	Center	38.43	78.86	1.8
CG4 (NC)	Southern edge	36.13	80.28	5.6
CG5 (GA)	Beyond southern edge	33.93	83.36	8.0
Mean northern pop. ‡	North	-	-	-2.5
Mean center pop. ‡	Center	-	-	1.1
Mean southern pop. ‡	South	-	-	4.7

† Data extracted from WorldClim database version 2.0 (Fick and Hijmans, 2017); ‡ Data measured for mean population of the eastern cluster. Northern populations: NY4, NY5, NY6; center populations: PA3, MD2, WV1; southern populations: VA1, NC2, NC4.

Table S4: Environmental stress per site

Site	T_{min} year 2 [° C]	T_{min} year 2 + 3 [° C]
CG1 (NY)	-5.8	-6.0
CG2 (MA)	-2.8	-1.7
CG3 (VA)	2.2	†
CG4 (NC)	5.2	7.0
CG5 (GA)	6.4	7.0

Temperature recordings from data loggers (iButton®, Maxim Integrated Products, Inc) placed in each common garden site (1.5 m above ground, in the shadow).

† In the analysis, replaced by the value of year 2

Table S5: Results of analysis performed on the level of individual pots, with a model testing for the effect of cross type (between- compared to within-population crosses [0]), genomic estimate of mutational load, environmental stress ($\Delta\ T_{min} = T_{min\ CG} - T_{min\ origin}$) and their interactions on *multiplicative performance I* up to year 3, at five common garden sites

	Cross type (CT)	***Genomic load (GL)***	***Stress (S)***	S^2	***CT * GL***	***CT * S***	***CT * S^2***	***GL * S***	***GL * S^2***	***CT * GL * S***	***CT * GL * S^2***	
Process	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	
Log-normal	**0.15** *	**-1.35** *	**16.91** *	**-8.14** **	**1.53** *	-4.95	-2.26	69.36	-14.58	-24.26	37.86	†
Logistic	**0.63** ***	-2.65	**79.58** **	**-58.54** ***	2.79	**-65.53** ***	15.24	**384.49** ***	61.25	163.79	-30.87	

Multiplicative performance (log_{10}-transformed if >0) was assumed to follow a Gaussian distribution with 0-inflation. Therefore, the model assessed all fixed and random effects for their importance in both the Gaussian process (total number of flowers during one or two reproductive seasons) and the logistic process (binary variable depicting germination combined with survival and the capacity to initiate flowering). Estimates of coefficients are modes of an MCMC sample from the posterior distribution of parameters (*mean*). The logistic part of the model predicts non-zeros in the distribution on the logit scale. Genomic load and stress were standardized prior to analyses (mean = 0). Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table S6: Summary of mixed-effects models testing for the effect of the genomic estimate of mutational load, environmental stress ($\Delta T_{min} = T_{min\ CG} - T_{min\ origin}$) and their interaction on population-level heterosis or within-population cross performance based on *multiplicative performance* (*MP*) up to year 2 or 3, and the finite rate of increase (λ), at five common garden sites

		Genomic load (GL)		Stress (S)		Stress^2 (S^2)		GL * S		GL * S^2			
Dependent variable	***N***	***β***	***χ²***	***β***	***χ²***	***β***	***χ²***	***β***	***χ²***	***β***	***χ²***	***R^2m***	***R^2c***
Heterosis													
MP I to year 2	89	**2.21**	**9.48 *** **	**1.61**	**15.10 *** **	-0.24	0.34	3.06	0.42	**8.90**	**4.25 ***	0.31	0.48 †
MP II to year 2	89	**2.21**	**9.49 ** **	**2.00**	**21.66 *** **	-0.04	0.01	2.94	0.36	**9.92**	**4.93 ***	0.35	0.50 †
Performance (WPC)													
MP I to year 3	100	-1.42	3.38 (*)	**-2.39**	**7.19 ** **	**-1.50**	**9.93** **	-6.99	1.76	6.33	1.91	0.28	0.61 †, ‡
MP II to year 3	100	-1.17	2.61	**-2.58**	**9.11 ** **	**-1.41**	**10.77** **	-4.65	0.95	6.46	2.56	0.30	0.65 †, ‡
MP I to year 2	100	**-1.39**	**4.66 ***	**2.45**	**8.90 ** **	**-1.25**	**9.73** **	6.78	2.34	**10.43**	**7.60 ** **	0.28	0.71 †
MP II to year 2	100	-1.19	3.71 (*)	**2.37**	**9.13 ** **	**-1.14**	**9.30** **	5.90	2.04	**9.71**	**7.64 ** **	0.27	0.72 †
λ	100	-0.22	1.41	**-0.84**	**15.63 ** **	**-0.24**	**5.38 ***	-0.01	0.00	1.91	3.72 (*)	0.35	0.68 †

Population heterosis estimates and within-population cross (WPC) performance estimates were $\log_{10}$-transformed prior to analysis. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include regression coefficient (β), χ^2-value of each fixed effect, and the marginal and conditional R^2 of the model. Model fits with significant (positive) intercept are indicated by †. Regression coefficients with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown. For one of the five common gardens (CG3), the experiment stopped early and variables consider performance to year 2 only (indicated by ‡).

Table S7: Summary of population-level heterosis or within-population cross performance based on *multiplicative performance* (*MP*) up to year 3 or 2, and the finite rate of increase (λ), at five common garden sites

	Population heterosis				Population performance (WPC)	
Dependent variable	***N***	***Min.***	***Mean***	***Max.***	***N***	***Mean***
MP I to year 3	89	-0.96	1.88	23.50	100	41.31
MP II to year 3	89	-0.92	2.65	30.23	100	25.66
MP I to year 2	89	-0.95	2.96	41.21	100	28.75
MP II to year 2	89	-0.97	3.89	71.45	100	20.00
λ	89	-0.53	0.73	7.29	100	1.00

Table S8: Summary of mixed-effects models testing for the effect of transplant (beyond the edge compared to edge [0]), edge position (south compared to north [0]), and their interaction on population-level heterosis or within-population cross performance based on *multiplicative performance* (*MP*) up to year 2 or 3, and the finite rate of increase (λ), of range-edge populations transplanted in edge and beyond edge common garden sites

		Transplant (T)			Edge position (E)			T * E				
Dependent variable	***N***	***β***	***χ²***		***β***	***χ²***		***β***	***χ²***	***R²m***	***R²c***	
Heterosis												
MP I to year 2	24	0.04	0.00		0.34	0.11		0.30	0.04	0.10	0.74	
MP II to year 2	24	0.21	0.34		1.91	1.95		-1.23	0.40	0.09	0.38	
Performance (WPC)												
MP I to year 3	24	-0.45	0.73		-0.12	0.05		0.20	0.07	0.08	0.32	‡
MP II to year 3	24	-0.56	1.17		-0.29	0.30		0.38	0.27	0.12	0.41	‡
MP I to year 2	24	**-0.42**	**5.08**	*	-0.45	3.16	(*)	0.27	1.02	0.21	0.56	
MP II to year 2	24	-0.48	0.28		-0.54	0.35		0.42	0.10	0.10	0.85	
λ	24	**-0.15**	**5.34**	*	-0.11	1.94		0.03	0.14	0.28	0.48	

Population heterosis and within-population cross (WPC) performance were $\log_{10}$-transformed prior to analysis. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include regression coefficient (β), χ^2-value of each fixed effect and the marginal and conditional R^2 of the model. Regression coefficient with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown. For one of the five common gardens (CG3), the experiment stopped early and variables consider performance to year 2 only (indicated by ‡).

Table S9: Model selection on potential predictors of heterosis

Model	***AICc***	***Δi***	***wi***
Genomic load + stress + stress^2 + (genomic load * stress) + (genomic load * stress^2)	111.1	0.0	0.93
Genomic load + stress + stress^2	116.3	5.2	0.07
Genomic load	123.3	12.2	0.00
Genomic load + stress + (genomic load * stress)	127.0	16.0	0.00
Genomic load + stress	127.4	16.4	0.00

The dependent variable was heterosis based on *multiplicative performance I* to year 3. Genomic load was the ratio of genome-wide non-synonymous to synonymous polymorphic sites adjusted by their mean derived ferquency (Willi *et al.*, 2018), and stress was the relative difference of within-population cross performance of each population at a site compared to the best site. Models are ranked from best to worst fit, with lower *AICc* values indicating a better fit. The difference between the best fit model and the others is indicated as *Δi*, while the weight of each model is indicated by *wi*.

Table S10: Summary of mixed-effects models testing for the effect of the genomic estimate of mutational load, stress based on relative performance, and their interaction on population-level heterosis based on multiplicative performance (MP) up to year 3 or 2 and the finite rate of increase (λ), at five common garden sites

		Genomic load (GL)			Stress (S)			Stress^2 (S^2)			GL * S			GL * S^2					
Dependent variable	***N***	***β***	***χ²***		***β***	***χ²***		***β***	***χ²***		***β***	***χ²***		***β***	***χ²***		***R^2m***	***R^2c***	
MP I to year 3	89	**1.90**	**6.77**	**	0.36	0.60		**1.42**	**11.29**	***	-0.15	0.00		-0.91	0.03		0.20	0.44	†, ‡
MP II to year 3	89	**1.80**	**6.13**	*	**1.07**	**4.05**	*	**1.07**	**6.73**	**	0.32	0.01		6.66	2.20		0.23	0.51	†, ‡
MP I to year 2	89	**2.40**	**17.70**	***	**0.94**	**5.01**	*	**1.77**	**19.06**	***	1.99	0.23		5.07	1.85		0.37	0.49	†
MP II to year 2	89	**2.26**	**11.19**	***	**1.43**	**7.24**	**	**1.84**	**17.09**	***	1.38	0.11		7.14	2.95	(*)	0.38	0.55	†
λ (heterosis)	89	0.63	2.70		**0.92**	**16.98**	***	0.38	3.21	(*)	0.89	0.24		**4.16**	**4.82**	*	0.29	0.59	†

Population heterosis estimates were $\log_{10}$-transformed prior to analysis. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include regression coefficient (β), χ^2-value of each fixed effect the marginal and conditional R^2 of the model. Model fits with significant (positive) intercept are indicated by †. Regression coefficient with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. For one of the five common gardens (CG3), the experiment stopped early and variables consider performance to year 2 only (indicated by ‡).

Table S11: Magnitude of effect of the genomic estimate of mutational load and environmental stress ($\Delta T_{min} = T_{min\ CG} - T_{min\ origin}$) on ($log_{10}$-transformed) *multiplicative performance (*MP) up to year 3 or the finite rate of increase, λ, of within-population crosses (WPC), at five common garden sites

			Stress	
Dependent variable	***N***	***Genomic load***	***Warmer CG***	***Colder CG***
MP I to year 3	100	-56.1 *(0.74; 1.09)*	-86.9 *(0; 14.7)*	-14.2 *(0; -11.4)*
MP II to year 3	100	-54.9 *(0.74; 1.09)*	-92.6 *(0; 14.7)*	-6.6 *(0; -11.4)*
λ	100	-41.1 *(0.74; 1.09)*	-90.4 *(0; 14.7)*	24.1 *(0; -11.4)*

For the general model of Table S6 analyzing WPC performance, the magnitude of effect of genomic load and environmental stress was calculated as: percentage difference between the back-transformed predicted performance corresponding to the maximal value of the predictor variable (in parenthesis, right) in our sampling and the back-transformed predicted performance corresponding to the minimal value of the predictor variable (in parenthesis, left). The effect of genomic load was not significant in the analysis, but is reported for comparison. The magnitude of effect of genomic load was estimated considering temperatures close to those of the site of origin (stress = 0). The magnitude of effect of environmental stress was estimated considering mean genomic load values, and calculated separately if the common gardens (CG) were warmer or colder than the site of origin of the population.

Chapter 3: Reduced climate adaptation at range edges in North American *Arabidopsis lyrata*

Darío Sánchez-Castro*, Antoine Perrier, Yvonne Willi

Department of Environmental Sciences, University of Basel, Schönbeinstrasse 6, 4056 Basel, Switzerland.

*Corresponding author:

E-mail: dario.sanchezcastro@unibas.ch

This manuscript has been submitted to *Evolution Letter* (27.07.2020).

Abstract

Species range limits, when not caused by dispersal limitation, reflect constraints in the evolution of the ecological niche. Here we tested whether a history of small size and enhanced genetic drift of range-edge populations was linked with reduced adaptation. We performed a transplant experiment with sites across and beyond the species distribution of North American *Arabidopsis lyrata*, with plants from the centre of distribution, and the periphery with a history of range expansion or long-term isolation. Performance declined toward the southern range limit and beyond, suggesting that southern range limits – but not northern ones – reflected niche limits. Furthermore, we found adaptation to two important niche- and range-determining environmental variables, temperature in spring and precipitation during the wettest quarter. However, the signature of adaptation to precipitation was reduced in populations with a history of small population size. Therefore we conclude that reduced adaptation is a contributor to range limits.

Keywords: genetic diversity, genetic drift, local adaptation, marginal population, niche limits, range edge, small population size, transplant experiment.

Introduction

Several hypotheses have been put forward for why species are limited in their geographic distribution, but so far, it is unclear what the main constraining processes are (Hoffmann & Blows 1994; Kawecki 2008; Sexton *et al.* 2009; Willi & Van Buskirk 2019). If dispersal limitation is found not to be relevant at species range limits, geographic distributions reflect niche limits (Chown & Gaston 1999; Hargreaves *et al.* 2014). The evolutionary explanation for range limits then is that populations at the edges fail to adapt and expand their ecological niche. Evolutionary models have suggested the conditions under which adaptation at range edges fails and which therefore cause range limits (reviewed in Sexton *et al.* 2009). Here we focused on the fact that many species show enhanced signatures of genetic drift toward range edges (Pironon *et al.* 2017), which may be linked with the reduced adaptive potential. The lack of adaptation to range-edge conditions due to genetic drift may prevent the further spread of the species into more extreme environments and be one of the causes of range limits. Here, we tested the hypothesis that adaptation at current range edges is reduced and that this is connected with a history of long-term small size.

The role of increasing genetic drift toward range limits on adaptation has not been explored conclusively by evolutionary theory on species ranges. One set of models tracks adaptation and range expansion by assuming a linear gradient of environmental change and a polygenic trait under selection. Adaptation is predicted to lead to the expansion of the range unless dispersal is long and the environmental gradient steep, which leads to maladaptation, the gradual decline in population mean fitness, and eventually to range limits (Kirkpatrick & Barton 1997). When the action of both selection and genetic drift are considered, the same sort of model predicts that range limits establish by two contributors: steep environmental gradients, and either genetic drift opposing selection or genetic drift eroding genetic variation (Polechová & Barton 2015; Polechová 2018). Two aspects are noteworthy. Dispersal has a mixed effect; it increases dispersal load but lowers the magnitude of genetic drift. Furthermore, population sizes and genetic drift are fairly constant across the range. A second set of models works with source-sink dynamics and addresses whether a sink site that differs

in ecological conditions can be occupied and adapted to (reviewed in Kawecki 2008). Here the prediction is that adaptation and persistence in the sink is more likely if gene flow is not too restricted because it brings in recruits and genetic variation important for local adaptation (Holt & Gomulkiewicz 1997; Holt *et al*. 2003). While some of the source-sink models included the action of genetic drift, its role in the source, which could stand for the outermost edge-population, was not explored.

There are several reasons why range edges may commonly have a history of small population size that then affects the potential to adapt via genetic drift opposing selection or eroding genetic variation. First, based on empirical observations, a purely ecological hypothesis was formulated, namely that species have high abundance in the range centre and lower abundance at the range periphery because of a decline in habitat quality or habitat availability (abundant-centre hypothesis; Hengeveld & Haeck 1982; Brown 1984). A recent meta-study provided strong support for this hypothesis as 81% of studies were found to report a significant decline in population occurrence from the centre to the periphery (Pironon *et al.* 2017). In principle, this should lead to the enhanced exposure to genetic drift, as was suggested by the population-genetic extension of the abundant-centre hypothesis (Eckert *et al*. 2008). A completely different hypothesis advocates that during range expansion, serial demographic bottlenecks accompanied by genetic drift leave a pattern of declining genetic diversity from the area of expansion start toward the expansion end (Excoffier *et al*. 2009). Indeed, many species underwent relatively recent range expansion due to Pleistocene glaciation cycles, which left an imprint of small population size toward range edges (Hewitt 1996; Hewitt 2000). In support of both hypotheses outlined above, Pironon *et al*. (2017) found an overall significant decline in genetic marker diversity from the centre to the periphery across species ranges, with 47% of studies showing a significant decline. Therefore, we can conclude that many edges of species ranges, have a history of small size and heightened exposure to genetic drift, either due to less available habitat or past range expansion.

This motivates the question of whether range edge populations with a history of increased genetic drift are less adapted to local environmental conditions. The testing for local adaptation is best done with a reciprocal transplant experiment (Kawecki & Ebert 2004). By performing a general transplant experiment across a species distribution and beyond, we asked whether range limits coincide with niche limits (I), whether populations were adapted to the local climate (II), and whether range edge populations with low genetic diversity were less adapted (III). Study organism was the short-lived perennial plant *Arabidopsis lyrata* subsp. *lyrata* from North America. A previous niche-modelling study on *A. lyrata* indicated that northern and southern range limits coincided with niche limits, with minimum temperature in early spring and precipitation during the wettest quarter being the variables that predicted species occurrence best (Lee-Yaw *et al.* 2018). Furthermore, populations toward range edges were shown to have a history of small population size due to post-glacial range expansion and rear-edge isolation, and mating system shifts to selfing associated with range edges (Griffin & Willi 2014; Willi *et al.* 2018). In line, higher genetic diversity and therefore larger effective population sizes were found in areas from which colonization started, which nowadays are near to the geographic centre of distribution (Willi *et al.* 2018). The twenty populations involved in the study represented the geographic centre of distribution as well as the peripheries.

Material and methods

Study organism and within-population crosses

North American *A. lyrata* subsp. *lyrata* (from now on abbreviated as *A. lyrata*) is distributed along the eastern US, from North Carolina to Upstate New York, and in the Midwest, from Missouri to south-western Ontario, forming two distinct ancestral clusters (Willi & Määttänen 2010; Griffin & Willi 2014; Willi *et al.* 2018). It is a mostly self-incompatible, insect-pollinated plant that produces basal rosettes. However, a fraction of self-compatible and selfing populations were found at the edges

of species distribution (Griffin & Willi 2014). The species is found on sand dunes and rocky outcrops, as well as on sandy or rocky riverbanks and shorelines.

Twenty populations of *A. lyrata* were selected because they represented the total distribution in North America from south to north, two ancestral genetic clusters and both mating systems (Fig. 1, Table S1). For each natural population, in 2007, 2011, and 2014, mature fruits of 30-50 plants were collected over a surface area of about 450 m^2. To reduce the effects of the site of origin and to get a high number of seeds of known genotypic composition, we raised plants indoors to perform within-population crosses. For each population, two seeds of 26 seed families were sown in pots and later thinned to one plant per pot (see Table S2 for raising conditions). Plants of each population were randomly assigned to be either mother plants/dams receiving pollen (12), father plants/sires being pollen donors (12), and backup plants (2). Each dam was randomly assigned a sire from the same population. Hand pollinations were performed on emasculated flowers at the bud stage. The crossing was repeated until 6-7 fruits or about 60 seeds were available per cross combination. The experiment resulted in 224 crosses with seeds for sowing in the transplant experiment.

Transplant experiment

Five transplant sites were established along a latitudinal gradient in the eastern US (Fig. 1). Sites were selected based on the position relative to the species range: beyond the northern edge, in the Adirondacks, NY (CG1); near the northern range edge, in Williamstown, MA (CG2); in the centre of distribution, in Harrisonburg, VA (CG3); near the southern range edge, in Winston-Salem, NC (CG4); and, beyond the southern range edge, in Athens, GA (CG5), (Fig. 1, Table S3). The start of transplanting at the sites was adjusted to the local climate, about 6 weeks before the long-term daily average temperature fell to 10 °C. The setup started in August 2017 for the site beyond the northern edge (CG1) and ended two months later at the southernmost site (CG5). At the southern range edge (CG4), sowing had to be repeated in December of the same year because of chloride in the water.

At each transplant site, three replicate pots with two seeds each were prepared per cross combination. The three pots per cross were then split into three spatial blocks, and within the block, they were randomly assigned to 13 multi-pot trays with 38 pots each (note that not all pots were filled with seeds analyzed here; others contained between-population crosses; see Perrier *et al.* 2020). Pots had a diameter of 7 cm, a depth of 6 cm, were perforated at the bottom, and filled with a 1.5:1 mix of unfertilized peat moss and washed sand. The same protocol was followed at all sites. As some crosses had produced only a few seeds (cross combinations with less than 30 seeds), we replaced them with another maternal line of the same population, or only one seed was sown per pot. A total of 7,098 seeds were sown (5 transplant sites x 20 populations x 12 maternal lines x 3 blocks x 2 seeds per pot – 102 missing seeds, Table S4).

Pots were immediately placed outdoors, into a meadow, under a portable walk-in greenhouse to keep conditions favorable for germination for the first 10-12 days; an exception was the transplant site at the southern range edge where the second round of germination occurred inside the university glasshouse. When the portable greenhouse was removed, a white mesh cloth protected seedlings from being washed away for another week. Plants were watered as needed, keeping the soil surface moist during the first month to promote germination. Later on, plants were exposed to the natural local conditions at each site. However, competitive interactions were avoided by removing other plant species and covering the surrounding area with a black foil. Herbivore pressure was partially controlled: with a fence around the blocks, ant traps against seed predation in the first fall, and organic slug repellent in the first spring. When in the same pot two seedlings germinated, one was haphazardly removed.

Plant performance was tracked weekly or more regularly, starting with the sowing of seeds in late summer 2017 until the end of the reproductive season in June 2019 (for a list and description of traits see Table S5). *Reproductive output* was assessed in each of the two reproductive seasons, 9 weeks after the first few plants flowered at a site in year 2 (2018), and 5 weeks after the start of flowering in year 3 (2019). It was the total number of fruits, pedicels (flowers that did not produce a

fruit but contributed with pollen), flowers, and buds. Finally, *multiplicative performance* (*MP*) was calculated as g*ermination rate* observed in a pot times *reproductive output* up to year 3. At the end of the experiment, *root length* was measured as the distance from the centre of the rosette to the end of the longest root. Unfortunately, the transplant in the centre (CG3) had to be removed in fall 2018 because the site was needed for another experiment; to compare across sites, we therefore performed also analyses on *MP* up to year 2. Between fall 2018 and spring 2019, we additionally performed a seed-burial experiment to study seed survival. Seeds of multiple maternal plants of a population were pooled, packed in bags, and left on the soil surface in every common garden (Fig. S1).

Climate data

Analysis of climate adaptation focused on the two most niche- and range-determining climatic variables, minimum temperature in early spring and precipitation during the wettest quarter (Lee-Yaw *et al.* 2018). This data of the sites of origin of populations, and – for precipitation – for the transplant sites, was extracted from WorldClim database version 2.0 (Fick & Hijmans 2017). Temperature data at the transplant sites was collected by loggers in each garden. Five of them per site were installed 1.5 m above the ground, close to the pots and in shade, and recorded at an interval of 1 h. The difference between WorldClim-based minimum temperature in early spring, in March and April, at the site of origin of a population and the corresponding temperature measured with loggers at a transplant site was calculated and abbreviated with Δ_{Temp}. The difference between WorldClim-based precipitation during the wettest quarter at the site of origin of a population and a transplant site was abbreviated with Δ_{Prec}. The testing for local adaptation was based on absolute values, $|\Delta_{Temp}|$ and $|\Delta_{Prec}|$, with estimates close to zero indicating little difference in conditions between those populations had experienced at their site of origin and those experienced at a transplant site.

Statistical analysis

All main analyses were performed on *multiplicative performance* as the dependent variable. A first generalised linear mixed model (GLMM) tested whether range limits coincide with niche limits (research question I). The analysis was performed in a Bayesian framework (MCMCglmm in R; Hadfield 2009; R Core Team 2019) because *MP* was 0-inflated and required the analysis of both the logistic part with the 0s and the Gaussian part of the distribution (values $\log_{10}$-transformed if >0). Fixed effect was common garden, with the reference garden in the centre of the range (CG3). Random effects were block nested within transplant site, population, and family nested within population. MCMCglmm analysis was run on 10 parallel chains, with a burnin of 5000, thining of 100, and a nitt-value of 200,000. To assess whether the species had self-persistent populations within the range but not beyond the range limits, we estimated the growth rate of each population at each transplant site by creating stage-classified matrices (Caswell 2001; see Fig. S1). Furthermore, life stages were tested for their contribution to performance in the common gardens. *Germination* and *survival* were estimated as binary variables (0, 1). *Survival* $_{\text{year 1}}$ took into account the germination state (i.e. NA if the seed did not germinate; 0 if the plant died before the end of winter in 2017/18; 1 if it survived); and *survival* $_{\text{year 2}}$ was based on *survival* in year 1 (i.e. NA if the plant had died before). *Damage* on the rosette or on the inflorescence was also treated as binary. *Time to flowering*, the *severity of the damage* (1: 0-25%; 2: 26-50%; 3: 51-75%; 4: 76-100%), *reproductive output*, and *root length* were continuous variables. All these variables were analysed individually with restricted maximum likelihood, with the R package lme4 (Bates *et al.* 2015) and lmerTest (Kuznetsova *et al.* 2017). Fixed and random effects were the same as specified above.

In the second part of the analyses, tests addressed whether populations showed a signature of climate adaptation (research question II) and whether that signature depended on the history of genetic drift (III). The main dependent variable was again *multiplicative performance*, analysed by a GLMM and Bayesian statistics. Fixed effects were $|\Delta_{\text{Temp}}|$, $|\Delta_{\text{Prec}}|$, genomic diversity, and the interaction between the former two variables and genomic diversity. Genomic diversity was assumed

to reflect long-term population size, and species-wide estimates were shown to be well explained by expansion distance or rear-edge isolation, mating system and ancestral cluster (74% of variation explained; Willi *et al.* 2018). The estimate of genomic diversity was Tajima's π of intergenic regions revealed by pool-sequence analyses of population samples (Willi *et al.* 2018, Table S1). Random effects were common garden, block nested within common garden, population, and family nested within population. Secondary analyses focused on the role of the same fixed effects on life stage variables.

Results

Environmental conditions

Climate differed strongly between the five transplant sites (Table S2). Minimum temperature in early spring increased gradually from north to south, while precipitation during the wettest quarter increased from the centre toward beyond the range edges. Mean annual temperature at each transplant site was slightly warmer than expected based on longer-term averages depicted by the WorldClim data set (Table S2). The trend was strongest for the central and northern common garden sites.

I. Do range limits reflect niche limits?

A first main model tested for the effect of the position of common gardens across the *A.lyrata* distribution on *multiplicative performance* up to year 3 (Table 1; mean values of common gardens in Table S6). The effect of the common garden was significant in the logistic part of the model, depicting whether plants made it to flowering, for the site in the north, south, and beyond the southern edge. Values were significantly higher at the northern edge, and lower at the southern edge and beyond the southern edge compared to the common garden in the centre of distribution, CG3. For the log-normal part of the model, depicting the number of flowers produced, only the common garden at the northern edge differed, with lower values compared to CG3. Figure 2A combines results of the two parts of

distribution, illustrating low overall *multiplicative performance* at the gardens at the southern edge and beyond the southern edge of distribution, but little difference between common gardens in the north. In line, the population growth rate, *r*, was much reduced and median values across populations around 0 at the southern edge and beyond the southern edge, indicating that populations are mostly non-persistent when transplanted beyond the southern edge (Fig. 2B, Tables S6, S7). Growth rates at the northern sites were not significantly different from those in CG3 and overall on the positive side. Results indicate that southern range limits reflect niche limits, while northern range limits do not seem to represent the species niche limits.

Further analyses focused on the effect of the common garden on life stage components as summarized in Table S7. *Germination* was significantly lower in the north and beyond the northern edge, but significantly increased at the southern edge (there seeds were raised in a greenhouse). In the first year, *survival* was significantly higher in the northern sites and at the southern edge, compared to CG3, while *survival* $_{\text{year 2}}$ was significantly lower in all common gardens compared to CG3, and strongest in the south and at the northern range edge. *Time to flowering* in year 2 was not significantly different at the northern edge and beyond the northern edge but significantly longer at the southern edge and beyond the southern edge, indicating that plants in the south flowered later relative to when soil temperature increased above 5°C (or after snowmelt). The *reproductive output* to year 2 was significantly lower at both southern sites, but also in the north, and here in year 2 and when the output of year 2 and year 3 were added (Table S7). Roots were longer at the southern sites (not measured at CG3 and therefore comparison made with CG1). Finally, *damage to rosettes* was more common in the north and beyond the northern edge, but *damage severity* was lower compared to CG3. *Damage severity* on rosettes was also reduced at southern sites. Overall, these results supported the much reduced performance and population growth rate in the southern-most transplant site mainly due to reduced overall longevity of plants.

II. Are populations adapted to the climate conditions of their site of origin?

III. Is the effect of adaptation reduced in populations of range edges with low genetic diversity and a history of stronger genetic drift?

The main model on *multiplicative performance* revealed adaptation to both temperature and precipitation (Table 2). The absolute difference in minimum temperature in early spring between population origin and transplant site, $|\Delta_{Temp}|$, had a negative effect on the logistic process of performance, with fewer non-zero values the larger the difference was (Fig. 3A). In other words, plants had a higher chances to succeed from the seed stage to reproduction under more similar temperature conditions between home and transplant site, indicating temperature adaptation. A trend was already observed up to year 2, and the correlation became significant up to year 3. The absolute difference in precipitation during the wettest quarter between the site of population origin and transplant site, $|\Delta_{Prec}|$, had also a negative effect but this time on the log-normal part of the distribution of *multiplicative performance*. Once plants reproduced successfully, a greater performance was observed under similar precipitation as at the site of origin, indicating adaptation to the precipitation regime. Tajima's π did not affect the plant performance, however, there was an interaction with the ecological predictor of $|\Delta_{Prec}|$, with an effect on the normal part of the distribution for *multiplicative performance*. Once plants succeeded with germination and achieved reproduction, they revealed a signature of stronger climate adaptation the higher genomic diversity was and the weaker the history of genetic drift characteristic of marginal populations was (Fig. 3B).

Analyses on life-stage variables showed when patterns of adaptation emerged (Table S8). Adaptation to temperature was already expressed at the life stage of germination. *Germination* was reduced the more different the temperature regime was between the site of origin and site of assessment. Other life stages that contributed as a trend were *reproductive output* to year 3 and *root length*. *Reproductive output* tended to be lower and the roots shorter the more different the temperature regime was between site of origin and site of assessment. Adaptation to precipitation was expressed in later stages, by a decrease in *survival* $_{\text{year 2}}$, while the damage of inflorescences decreased

under more different precipitation conditions population experienced. However, there was no particular life stage when the interaction between Tajima's π and $|\Delta_{Prec}|$ was significant.

Discussion

I. Do range limits reflect niche limits?

Our study attempted to first answer the question of whether range limits are the spatial representation of niche limits by combining results from species distribution modelling and transplant experiments. For *A. lyrata*, Lee-Yaw *et al.* (2018) had found that southern and even more so northern range limits were predicted well by habitat suitability, mainly defined by average minimum temperature in early spring, and precipitation of the wettest quarter. Results of the transplant experiment across the distribution and beyond confirmed that range limits reflect niche limits in the south. At the southern edge and beyond, plant performance was significantly lower compared to the centre of distribution and growth rates were around 0 at the southern edge and beyond the southern edge (Fig. 2). However, the northern range edge may not be a reflection of niche limits anymore. At the northern edge and beyond, plant performance seemed comparable with that at the centre of distribution and growth rates were not significantly different. Overall, for the southern range edge of *A. lyrata*, there is good agreement with meta-analyses showing that range limits often equal niche limits (Hargreaves *et al.* 2014; Lee-Yaw *et al.* 2016).

However, secondary analysis performed on life stage variables showed that conditions in the north were not systematically better for *A. lyrata*. Germination was lower at northern sites. Then subsequent survival ($survival_{\text{year 1}}$) was initially higher at the northern sites but then changed to lower during the second year. Furthermore, the total reproductive output to year 2 was greater at the centre and was lower at least at the northern edge. Plants in the north seemed to be more affected by herbivores, while the severity of the damage was higher at the centre site. These results provide partial support that also northern conditions may be constraining for the species and to some extent limiting.

We also found that the climatic conditions during the study were on average warmer than the long-term average based on WorldClim data, with strongest increase in the centre and at the northern two sites, which may have made especially northern sites more benign for *A. lyrata* than they used to be. While the prevailing conditions may have been unusually warm during our study, the trend is also the one observed with climate warming. Therefore, we hypothesize that northern range limits may not reflect niche limits any more for *A. lyrata* due to global warming, but that the species is nowadays dispersal limited at the cold end of distribution. Such a result was found for example for 38% of non-forest plant species of the European Alps (Rumpf *et al.* 2019).

II. Are populations adapted to the climate conditions of their home site?

Adaptation to local conditions is a common finding of transplant studies (Leimu & Fischer 2008; Hereford 2009). Furthermore, local adaptation may most often be due to climate or other abiotic factors but to a lesser extent due to biotic interactions (Hargreaves et al. 2020). Here we found evidence for adaptation to minimum temperature in early spring, the most niche- and range-determining variable formerly revealed by distribution modelling, and precipitation during the wettest quarter; plant performance was better when conditions at transplant sites were similar to those at the site of origin. These signatures of adaptation measured on plant multiplicative performance were detected in the logistic part of the distribution for temperature, with more zeros when conditions were different than those at the sites of origin. Adaptation to precipitation was detected in the normal part of the distribution of multiplicative performance, determined by the total number of reproductive organs produced once plants reached the flowering stage. Similar results were found in the analysis of life stages. Adaptation to temperature was expressed early in life, mainly during germination (with additional trends for reproductive organs and root length). In contrast adaptation to precipitation was revealed later in life, in survival of the second year. This difference in timing of expression suggests that adaptation to temperature may involve fewer selection targets already expressed during germination, while adaptation to precipitation may be due to more genetic variants of small

phenotypic contribution with cumulative effect only visible later in life. Evidence for local adaptation to temperature has been numerous in plants, e.g., shown in transplant experiments performed across latitude (e.g., Ågren & Schemske 2012), or across elevation (Halbritter *et al.* 2018). Furthermore, it may not be uncommon that local adaptation to temperature is expressed at an early life phase, as also in *A. thaliana* adaptation to southern versus northern conditions strongly involved the seedling establishment phase (Postma & Ågren 2016).

III. Is the effect of adaptation reduced in populations of range edges with low genetic diversity and a history of stronger genetic drift?

A main result of this study is that adaptation to the second important niche- and range-determining variable was dependent on the history of genetic drift experienced by populations. Results pointed to long-term small populations having a relatively low fitness peak when the precipitation regime was similar between the site of origin and site of assessment, but that this fitness peak was higher and wider for populations with long-term large size, coming from the centre of the distribution. In this sense, populations of large size may be more tolerant of a wider range of precipitation regimes (Fig. 3B). In *A. lyrata*, populations with low genomic diversity and therefore a history of small size and genetic drift generally have a history of either post-glacial range expansion or long-term rear-edge isolation. The pattern was found for outcrossing populations, and the handful of selfing populations detected at range edges were confirmed to have an even more pronounced pattern of reduced genetic variation (Willi *et al.* 2018). The result of a reduced signature of adaptation in populations with a history of stronger genetic drift are in line with a result from European *A. lyrata*. A transplant experiment over an elevational gradient revealed local adaptation in one pair of low-/high-elevation populations with higher genetic diversity, but no signature of such adaptation in another pair of populations with lower genetic diversity (Hämälä *et al.* 2018). Higher tolerance to extreme environmental conditions in populations of range cores compared to peripheral populations was found in a *Triticum* species. Plants of the core of distribution and range edges were exposed to experimental

conditions found beyond the range edge, and core populations coped better with those (Volis *et al.* 2014). In our study, life-stage analyses did not reveal a particular timing of when genetic drift inferred with local adaptation or the evolution of tolerance, suggesting that its expression is due to many genetic variants each of small effect, which are more prone to be affected by genetic drift.

Range limits have recently been suggested to be a result of a failure of local adaptation due to genetic drift opposing selection or genetic drift eroding genetic variation (Polechová & Barton 2015; Polechová 2018). Our study supports that local adaptation is constrained by genetic drift associated with range-edge position. In line, Vergeer & Kunin (2013) found in a transplant experiment with sites and populations from the core and periphery of European *A. lyrata* that plant performance was generally higher in the core of distribution. Furthermore, populations from the southern range edge with the smallest census sizes were the least locally adapted. In *Plantago major*, a transplant experiment including a site in the core of distribution and two toward the northern edge revealed local adaptation of both core and edge populations, but the extent of local adaptation in edge populations with lower genetic diversity tended to be lower (Halbritter *et al.* 2015). In *A. lyrata*, the likely mechanism for this reduced adaptation is that genetic drift opposes selection. The alternative, that genetic drift impedes adaptation via a loss of genetic variation seems less likely. A quantitative genetics experiment involving populations from the centre and edges of distribution of *A. lyrata* showed that genetic variation for ecologically relevant traits was not much reduced in range-edge populations, and genetic correlations among them were weaker, which overall produced a pattern of similar adaptive potential (Paccard *et al.* 2016).

Our study reinforces the idea that populations at range margins with a history of strong genetic drift, caused by past range expansion, rear-edge isolation, or a selfing mating system, have a reduced signature of adaptation and lower tolerance of atypical environmental conditions. This puts them in a position of a lower population mean performance due to maladaptation on the one hand and makes them weak colonizers on the other hand. Apart from their higher mutational load by genetic drift

opposing purifying selection (Perrier *et al*. 2020), narrow and low adaptation to climate may together be main causes of geographic species distribution limits.

Acknowledgments

This research was supported by the Swiss National Science Foundation (31003A_166322). We are thankful to Celia Evans (Paul Smith's College, Paul Smith, NY), Joan Edwards (Williams College, Williamstown, MA), Heather Peckham Griscom (James Madison University, Harrisonburg, VA), William K. Smith (Wake Forest University, Winston-Salem, NC) and Rodney Mauricio (University of Georgia, Athens, GA) for logistical support in the USA. For field assistance we thank Mary Anderson, Michael Boyd, Bennet Coe, Scott Cory, Rachel Hillyer, Andrew Jones, Deidre Keating, Larry Kummer, David Lampman, Anastasia Levie-Sprick, Blake Macko, Shannon Malisson, Kathryn McGee, Althea Neighbors, Debra Rogers-Gillig, Caleb Rose, Amber Scarabaggio, Anna Shutley, Caroline Vath and Audrey Werner. For assistance with seed counts we thank Olivier Bachmann, Markus Funk, and Susanna Riedl. Collection permits were provided by the Clinton County Conservation Board, Cornell University, Fort Leonard Wood Army Base, Iowa Department of Natural Resources, Missouri Department of Conservation, New York State Office of Parks, Ontario Parks, Palisades Interstate Park Commission, Rock Island Lodge, United States National Park Service, Virginia Department of Conservation and Recreation and the Wisconsin Department of Natural Resources. The authors declare no conflicts of interest.

Data accessibility statement: All data is stored in Dryad (https://doi.org/10.5061/dryad.cc2fqz642)

References

Ågren, J. & Schemske, D.W. (2012). Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. *New Phytol.* 194:1112–1122.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67:1–48.

Brown, J.H. (1984). On the relationship between abundance and distribution of species. *Am. Nat.* 124:255–279.

Caswell, H. (2001). *Matrix Population Models: Construction, Analysis, and Interpretation*, 2nd edition. Sinauer Associates, Inc., Sunderland, MA.

Chown, S.L. & Gaston, K.J. (1999). Exploring links between physiology and ecology at macro-scales: the role of respiratory metabolism in insects. *Biol. Rev.* 74:87–120.

Eckert, C.G., Samis, K.E. & Lougheed S.C. (2008). Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Mol. Ecol.* 17:1170–1188.

Excoffier, L., Foll, M. & Petit, R.J. (2009). Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.* 40:481–501.

Fick, S.E. & Hijmans, R.J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37: 4302–4315.

Griffin, P.C. & Willi, Y. (2014). Evolutionary shifts to self-fertilisation restricted to geographic range margins in North American *Arabidopsis lyrata*. *Ecol. Lett.* 17:484–490.

Hadfield, J.D. (2009). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. https://cran.r-project.org/.

Halbritter, A.H., Billeter, R., Edwards, P.J. & Alexander, J.M. (2015). Local adaptation at range edges: comparing elevation and latitudinal gradients. *J. Evol. Biol.* 28:1849–1860.

Halbritter, A.H., Fior, S., Keller, I., Billeter, R., Edwards, P.J., Holderegger, R. *et al.* (2018). Trait

differentiation and adaptation of plants along elevational gradients. *J. Evol. Biol.* 31:784–800.

Hämälä, T., Mattila, T.M. & Savolainen, O. (2018). Local adaptation and ecological differentiation under selection, migration, and drift in *Arabidopsis lyrata*. *Evolution* 72:1373–1386.

Hargreaves, A.L., Samis, K.E. & Eckert, C.G. (2014). Are species' range limits simply niche limits writ large ? A review of transplant experiments beyond the range. *Am. Nat.* 183:157–173.

Hargreaves, A.L., Germain, R.M., Bontrager, M., Persi, J. & Angert, A.L. (2020). Local adaptation to biotic interactions: a meta-analyses across latitudes. *Am. Nat.* 195:395–411.

Hengeveld, R. & Haeck, J. (1982). The distribution of abundance. 1. Measurements. *J. Biogeogr.* 9:303–316.

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* 173:579–588.

Hewitt, G.M. (1996). Some genetic consequences of ice ages , and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58:247–276.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913.

Hoffmann, A.A. & Blows, M.W. (1994). Species borders: ecological and evolutionary perspectives. *Trends Ecol. Evol.* 9:223–227.

Holt, R.D. & Gomulkiewicz, R. (1997). How does immigration influence local adaptation? A reexamination of a familiar paradigm. *Am. Nat.* 149:563–572.

Holt, R.D., Gomulkiewicz, R. & Barfield, M. (2003). The phenomenology of niche evolution via quantitive traits in a ‘black-hole’ sink. *P. Roy. Soc. B.* 270:215–224.

Kawecki, T.J. & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecol. Lett.* 7:1225–1241.

Kawecki, T.J. (2008). Adaptation to marginal habitats. *Annu. Rev. Ecol. Evol. Syst.* 39:321–342.

Kirkpatrick, M. & Barton, N.H. (1997). Evolution of a species’ range. *Am. Nat.* 150:1–23.

Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82:1–26.

Lee-Yaw, J.A., Kharouba, H.M., Bontrager, M., Mahony, C., Csergő, A.M., Noreen, A.M.E. *et al.*

(2016). A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. *Ecol. Lett* 19:710–722.

Lee-Yaw, J.A., Fracassetti, M. & Willi, Y. (2018). Environmental marginality and geographic range limits: a case study with *Arabidopsis lyrata* ssp. *lyrata*. *Ecography* 41:622-634.

Leimu, R. & Fischer, M. (2008). A meta-analysis of local adaptation in plants. *PLoS ONE*, 3, e4010.

Paccard, A., Van Buskirk, J. & Willi, Y. (2016). Quantitative genetic architecture at latitudinal range boundaries: reduced variation but higher trait independence. *Am. Nat.* 187:667– 677.

Perrier, A., Sánchez-Castro, D. & Willi, Y. (2020). Expressed mutational load increases toward the edge of a species' geographic range. *Evolution* 74: 1711–1723.

Pironon, S., Papuga, G., Villellas, J., Angert, A.L., García, M.B. & Thompson, J.D. (2017). Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biol. Rev*. 92:1877–1909.

Polechová, J. & Barton, N.H. (2015). Limits to adaptation along environmental gradients. *PNAS* 112:6401–6406.

Polechová, J. (2018). Is the sky the limit? On the expansion threshold of a species' range. P*LoS Biol.*, 16, e2005372.

Postma, F.M. & Ågren, J. (2016). Early life stages contribute strongly to local adaptation in *Arabidopsis thaliana*. *PNAS* 113:7590–7595.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Rumpf, S.B., Hülber, K., Wessely, J., Willner, W., Moser, D., Gattringer, A., *et al.* (2019). Extinction debts and colonization credits of non-forest plants in the European Alps. *Nat. Commun.* 10:4293.

Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009). Evolution and ecology of species range limits. *Annu. Rev. Ecol. Evol. Syst*. 40:415–436.

Vergeer, P. & Kunin, W.E. (2013). Adaptation at range margins: common garden trials and the

performance of *Arabidopsis lyrata* across its northwestern European range. *New Phytol.* 197:989–1001.

Volis, S., Ormanbekova, D., Yermekbayev, K., Song, M. & Shulgina, I. (2014). Introduction beyond a species range: a relationship between population origin, adaptive potential and plant performance. *Heredity* 113:268–276.

Willi, Y. & Määttänen, K. (2010). Evolutionary dynamics of mating system shifts in *Arabidopsis lyrata*. *J. Evol. Biol.* 23:2123–2131.

Willi, Y., Fracassetti, M., Zoller, S. & Van Buskirk, J. (2018). Accumulation of mutational load at the edges of a species range. *Mol. Biol. Evol.* 35:781–791.

Willi, Y. & Van Buskirk, J. (2019). A practical guide to the study of distribution limits. *Am. Nat.* 193:773–785.

Table 1. Summary of model testing for the effect of common garden on multiplicative performance (*MP*) in a transplant experiment of *Arabidopsis lyrata*

		Fixed effects, logistic process										
		CG1 (Beyond north)		**CG2 (North, edge)**			**CG4 (South, edge)**			**CG5 (Beyond south)**		
Dependent variable	N	Mean	HPD	Mean	HPD		Mean	HPD		Mean	HPD	
MP $_{\text{to year 3}}$	1950	-0.154	(-0.821,0.365)	**0.599**	(-0.000,1.165)	*	**-0.978**	(-1.578,-0.428)	**	**-1.499**	(-2.136,-0.941)	***
MP $_{\text{to year 2}}$	1950	-0.215	(-0.838,0.358)	0.545	(-0.035,0.137)	(*)	**-1.055**	(-1.637,-0.497)	**	**-1.556**	(-2.149,-0.951)	***
		Fixed effects, log-normal process										
		CG1 (Beyond north)		**CG2 (North, edge)**			**CG4 (South, edge)**			**CG5 (Beyond south)**		
Dependent variable	N	Mean	HPD	Mean	HPD		Mean	HPD		Mean	HPD	
MP $_{\text{to year 3}}$	1950	0.061	(-0.126,0.278)	**-0.522**	(-0.730,-0.283)	***	0.176	(-0.069,0.412)		0.103	(-0.241,0.0.392)	
MP $_{\text{to year 2}}$	1950	-0.034	(-0.237,0.164)	**-0.532**	(-0.746,-0.317)	***	-0.256	(-0.498,0.019)	(*)	-0.085	(-0.450,0.218)	

The effect of each common garden is compared with the one in the centre of distribution (CG3). Multiplicative performance (log_{10}-transformed if >0) followed a Gaussian distribution with 0-inflation. Therefore, models assessed all fixed and random effects for their importance in both the logistic process (binary variable depicting germination, survival and capacity to initiate flowering; prediction of non-zeros) and the Gaussian process (total number of reproductive organs). Estimates of coefficients are modes of an MCMC sample from the posterior distribution of parameters (*mean* and higher posterior density, *HPD*, interval). Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table 2. Summary of model testing for the effect of the absolute difference in temperature between site of origin of populations and transplant site, $|\Delta_{Temp}|$, the absolute difference in precipitation, $|\Delta_{Prec}|$, genomic diversity depicted by Tajima's π and interactions on multiplicative performance (*MP*) in a transplant experiment of *Arabidopsis lyrata*

		Fixed effects, logistic process										
		$\|\Delta_{Temp}\|$			$\|\Delta_{Prec}\|$		**Tajima's π**		$\|\Delta_{Temp}\|*\pi$		$\|\Delta_{Prec}\|*\pi$	
Dependent variable	***N***	Mean	HPD		Mean	HPD	Mean	HPD	Mean	HPD	Mean	HPD
MP to year 3	1950	**-0.042**	**(-0.076,-0.004)**	*	-0.002	(-0.007,0.004)	0.064	(-0.063,0.193)	-0.004	(-0.013,0.007)	0.0001	(-0.0014,0.0016)
MP to year 2	1950	-0.035	(-0.075,-0.003)	(*)	-0.001	(-0.007,0.005)	0.081	(-0.047,0.216)	-0.006	(-0.017,0.004)	0.0002	(-0.0013,0.0018)
		Fixed effects, log-normal process										
		$\|\Delta_{Temp}\|$			$\|\Delta_{Prec}\|$		**Tajima's π**		$\|\Delta_{Temp}\|*\pi$		$\|\Delta_{Prec}\|*\pi$	
Dependent variable	***N***	Mean	HPD		Mean	HPD	Mean	HPD	Mean	HPD	Mean	HPD
MP to year 3	1950	-0.020	(-0.047,0.008)		**-0.005**	**(-0.010,-0.001)**	*	0.016	(-0.065,0.099)	0.002	(-0.005,0.010)	**0.001**	**(0.0002,0.0025)**	*
MP to year 2	1950	0.007	(-0.019,0.035)		**-0.005**	**(-0.009,-0.000)**	*	0.038	(-0.047,0.113)	-0.004	(-0.011,0.003)	**0.001**	**(0.0002,0.0026)**	*

Multiplicative performance (log_{10}-transformed if >0) followed a Gaussian distribution with 0-inflation. Therefore, models assessed all fixed and random effects for their importance in both the logistic process (binary variable depicting germination, survival and capacity to initiate flowering; prediction of non-zeros) and the Gaussian process (total number of reproductive organs). Estimates of coefficients are modes of an MCMC sample from the posterior distribution of parameters (*mean* and higher posterior density, *HPD*, interval). Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$. Results for random effects are not shown.

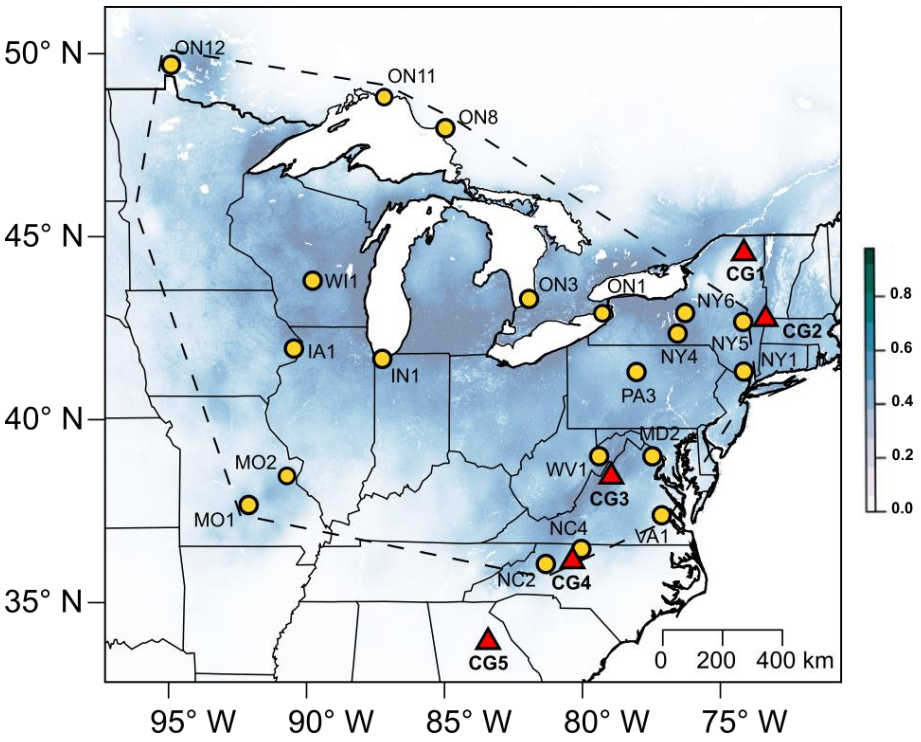

Figure 1. **Distribution of the 20 selected *Arabidopsis lyrata* populations and the location of the five common gardens (CG) transplant sites in North America**. Orange dots accompanied by a three-digit abbreviation represent the sites of origin of populations (Table S1, the two letters stand for the state in the US or the province in Canada, the number for latitudinal position within state, or longitudinal position within province as in Willi *et al.* 2018). Red triangles represent the location of each transplant site, followed by a number in sequence of initial sowing. The dashed line is the minimum convex polygon connecting the outermost populations of west and east. Shades of blue indicate habitat suitability revealed by niche-modelling (Lee-Yaw *et al.* 2018).

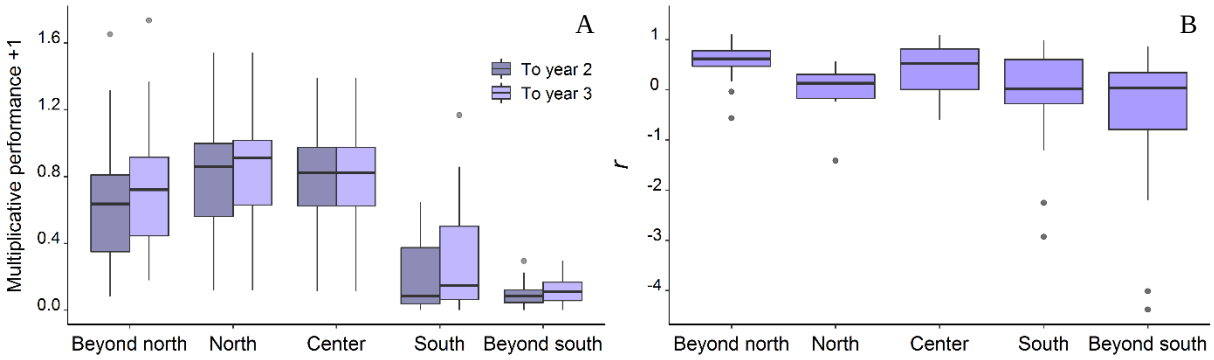

Figure 2. Multiplicative performance (A) and population growth rate (B) of *Arabidopsis lyrata* differing between transplant sites, sorted from north (left of the x-axis) to south (right). Panel A shows box plots based on population mean multiplicative performance up to year 2 or year 3. Population means were based on family means of pot-level multiplicative performance that was first log-transformed. Panel B shows box plots of population growth rate, *r*, of *Arabidopsis lyrata* populations at the five transplant sites, again sorted from north to south. The thick line of each box plot represents the median, the coloured box represents the interquartile range, the whiskers represent the variability outside the upper and lower quartiles, and individual dots represent the outliers.

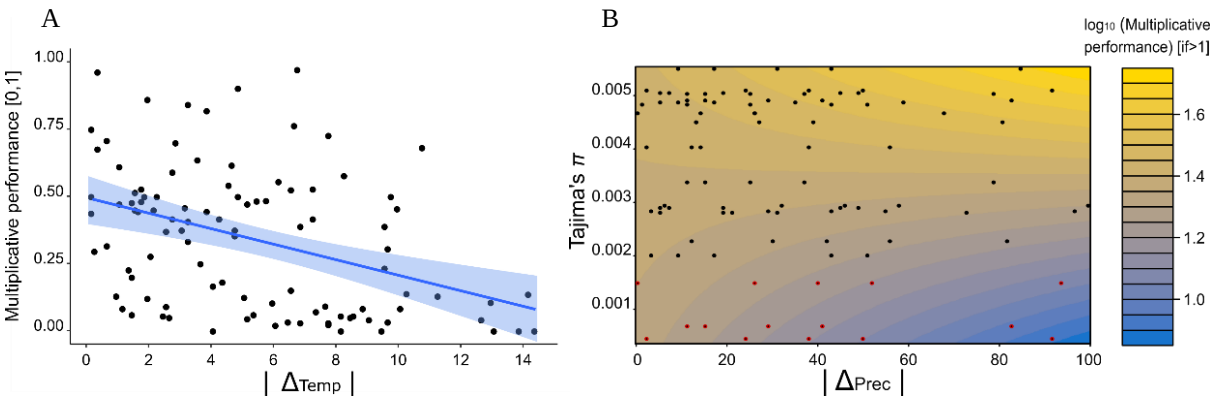

Figure 3. Relationship between multiplicative performance up to year 3 and absolute difference in temperature (A) and in precipitation, and genomic diversity, in *Arabidopsis lyrata* (B). In panel A, the population mean logistic response of multiplicative performance (0 or 1) is plotted against the absolute difference in minimum temperature in early spring between site of origin of populations and transplant site, $|\Delta_{Temp}|$ in °C. The model-predicted regression line is shown in blue, with the lower and upper 95% confidence interval. Panel B is a contour plot representing the predicted relationship between multiplicative performance up to year 3 (if values >0, in shades from blue to yellow) and both, the absolute difference in precipitation of the wettest quarter between site of origin of populations and transplant site, $|\Delta_{Prec}|$ in mm, and genomic diversity, Tajima's π. In both panels, dots are population means based on family means, for each common garden. Red dots represent selfing populations located at range edges, black dots the outcrossing populations.

Supporting Information

Table S1. Ecological and genomic information on the 20 *Arabidopsis lyrata* populations studied

Population	State/Province	Latitude [° N]	Longitude [° W]	Position	T_{min}espr [°C] †	PrecWQ [mm] †	Growing season †	Tajima's π ‡	Mating system ‡
IA1	Iowa	41.97	90.37	Center	-0.9	319	7	0.0040	Outcrossing
IN1	Indiana	41.61	87.19	South	0.0	296	7	0.0034	Outcrossing
MD2	Maryland	38.99	77.25	Center	3.8	290	9	0.0055	Outcrossing
MO1	Missouri	37.72	92.06	South	4.5	324	9	0.0020	Outcrossing
MO2	Missouri	38.47	90.71	South	4.0	292	9	0.0006	Selfing
NC2	North Carolina	36.04	81.16	South	3.8	363	9	0.0022	Outcrossing
NC4	North Carolina	36.41	79.96	South	5.0	326	9	0.0029	Outcrossing
NY1	New York	41.30	73.98	Center	0.4	326	8	0.0051	Outcrossing
NY4	New York	42.35	76.39	North	-2.5	283	7	0.0051	Outcrossing
NY5	New York	42.66	74.02	North	-3.0	296	7	0.0051	Outcrossing
NY6	New York	42.99	76.09	North	-2.1	292	7	0.0049	Outcrossing
ON1	Ontario	42.87	79.18	Center	-1.7	281	7	0.0014	Selfing
ON11	Ontario	48.77	87.13	North	-7.9	283	6	0.0004	Selfing
ON12	Ontario	49.65	94.92	North	-7.8	275	6	0.0029	Outcrossing
ON3	Ontario	43.26	81.84	Center	-2.6	278	7	0.0028	Outcrossing
ON8	Ontario	47.93	84.85	North	-7.5	302	6	0.0028	Outcrossing
PA3	Pensylvania	41.28	77.87	Center	-1.4	316	7	0.0049	Outcrossing
VA1	Virginia	37.42	77.02	South	5.5	332	10	0.0049	Outcrossing
WI1	Wisconsin	43.83	89.72	Center	-3.3	307	7	0.0047	Outcrossing
WV1	West Virginia	38.96	79.29	Center	1.1	294	9	0.0045	Outcrossing

The list reports: the name of the populations used in the experiment; the state (US) or province (CAN); coordinates; the position within the distribution area of *A. lyrata*; the average minimum temperature in early spring (T_{min}espr); precipitation of the wettest quarter (PrecWQ); length of the growing season, defined as the number of months with a mean temperature higher than 5 °C; Tajima's π of intergenic regions; the mating system. Data extracted from WorldClim (†). Genomic data from Willi et al. (2018) (‡).

Table S2. Conditions at each stage of the crossing experiment of *A. lyrata* for seed propagation

Process	Location	Temp. daytime [°C]	Temp. nighttime [°C]	Day length [h]	Light intensity [µmol m^{-2} s^{-1}]	Relative humidity [%]	Duration
Stratification	Cold room	-	4	0	0	0	12 days
Germination	Growth chambers *	20	20	8	150	50	22 days
Plant growth †	Growth chambers *	22	20	16	240	50	46 days
Crossing	University glasshouse	22	20	16	240	50	6 moths
Storage of siliques	Cold room	-	4	0	0	0	1-3 months

* CLF Plant Climatics, Wertingen, Germany

† Day length and light intensity were gradually increased every three days by 1 h and 20 µmol m^{-2} s^{-1}, respectively.

Table S3. Ecological information on transplant sites

Transplant site	Location	Position	Lat. [° N] Long. [° W]	Dist. to pop [km]	Starting date	T_{min}espr [°C]	T_{mean}annual [°C]	T_{mean}annual [°C] †	PrecWQ [mm]	Snow cover [days]	Growing season
CG1	Adirondacks (NY)	Beyond northern edge	43.96 74.22	120	11.08.2017	-5.8	6.3	4.4	333	173	6
CG2	Williamstown (MA)	Northern edge	42.71 73.20	8	28.08.2017	-2.9	9.0	7.4	307	116	7
CG3	Harrisonburg (VA)	Center	38.42 78.86	10	29.09.2017	2.2	13.6	11.8	281	10	10
CG4	Winston-Salem (NC)	Southern edge	36.12 80.28	41	07.12.2017	5.2	15.3	14.3	321	0	11
CG5	Athens (GA)	Beyond southern edge	33.90 83.38	152	19.10.2017	6.5	16.6	16.6	375	0	12

The list reports: the abbreviation of the transplant site (CG, common garden); the location; the position within the distribution area of *A. lyrata*; the coordinates; distance to the closest known natural population; the date of sowing; minimum temperature in early spring (T_{min}espr); mean annual temperature measured with loggers at each site, and extracted from WorldClim (†); precipitation of the wettest quarter (PrecWQ); the number of days with snow cover; and the growing season as the number of months when average monthly temperature was higher than 5°C.

Table S4. Summary of the total number of cross families and seeds sown per population in each common garden (CG)

Population	N° of cross families	Seeds sown in each transplant site				
		CG1	CG2	CG3	CG4	CG5
IA1	9	70	72	57	72	66
IN1	11	72	72	72	72	72
MD2	10	69	69	69	72	69
MO1	12	72	72	69	72	72
MO2	12	72	72	72	72	72
NC2	12	69	70	72	72	72
NC4	11	72	72	72	72	72
NY1	9	70	69	63	69	66
NY4	10	70	69	69	72	72
NY5	12	72	72	72	72	72
NY5	12	72	72	72	72	72
NY6	12	72	72	66	72	63
ON1	12	72	72	72	72	69
ON11	12	72	69	66	72	66
ON12	12	69	69	72	72	72
ON3	8	69	69	69	72	66
ON8	6	27	30	27	37	27
PA3	11	72	72	72	72	72
VA1	9	72	72	72	72	72
WI1	10	66	72	72	72	72
WV1	12	69	69	66	72	60

Table S5. Description of the traits measured and their analyses

Trait	Category	Level	Explanation	Measured			Analysis
Multiplicative performance				From	To		
MP $_{\text{to year 2}}$	Continuous	Pot	Germination rate * repro. output year 2	Sowing	Summer year 2		MCMC
MP $_{\text{to year 3}}$	Continuous	Pot	Germination rate * (repro. output year 2 + 3)	Sowing	Summer year 3	†	MCMC
Demographic rate							
λ	Continuous	Pop	Finite rate of increase				REML
r	Continuous	Pop	Growth rate, log-e transformation of λ				REML
Germination & survival							
Germination	Binary	Seed	Plants germinated (0/1)	Sowing day (day 0)	Spring year 2	‡	REML
Survival $_{\text{year1}}$	Binary	Seed	Survival until end winter year 1 (0/1)	Germination	Snowmelt or soil T°>5		REML
Survival $_{\text{year 2}}$	Binary	Seed	Survival spring year 2 to spring year 3 (0/1)	End of winter year 1	Snowmelt or soil T°>5	†	REML
Reproduction							
Time to flowering $_{\text{year 2}}$	Continuous	Pot	Number of days to flower	Snowmelt or soil T°>5	Spring/summer year 2		REML
Reproductive output $_{\text{year 2}}$	Continuous	Pot	Sum of flowers and buds in year 2	Once, 9 weeks after flowering			REML
Reproductive output $_{\text{to year 3}}$	Continuous	Pot	Sum of flowers and buds in year 2 + year 3	Spring/summer year 2	Spring/summer year 3	†	REML
Root length $_{\text{year 3}}$	Continuous	Pot	Length of the longest root, in mm	Once, 5 weeks after flowering		†	REML
Damage in year 2							
Damage to rosettes	Binary	Pot	Damaged rosettes (0/1)	Once, 9 weeks after flowering			REML
Damage to inflorescences	Binary	Pot	Damaged inflorescences (0/1)	Once, 9 weeks after flowering			REML
Damage severity $_{\text{rosettes}}$	Categorical	Pot	Severity of damage categorized from 1 to 4	Once, 9 weeks after flowering			REML
Damage severity $_{\text{inflorescences}}$	Categorical	Pot	Severity of damage categorized from 1 to 4	Once, 9 weeks after flowering			REML

† except for CG3: no data after summer 2018

‡ except for CG1: second cohort in spring 2018

Table S6. Summary of means with standard error (SE) of traits measured in each of the five common gardens (CG)

	CG1 (Beyond north)		CG2 (North, edge)		CG3 (Centre)		CG4 (South, edge)		CG5 (Beyond south)	
Trait	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Multiplicative performance										
MP to year 2	52	8	22	3	69	12	54	23	13	3
MP to year 3	44	8	20	3	69	12	8	2	7	2
Demographic rate										
λ	1.89	0.14	1.15	0.09	1.70	0.17	1.20	0.17	0.98	0.15
r	0.57	0.09	0.07	0.10	0.42	0.12	-0.14	0.23	-0.53	0.32
Germination & survival										
Germination	0.33	0.03	0.65	0.04	0.71	0.03	0.82	0.03	0.75	0.03
Survival year1	0.75	0.02	0.61	0.02	0.37	0.02	0.70	0.01	0.27	0.03
Survival year 2	0.34	0.03	0.06	0.01	0.24	0.03	0.04	0.01	0.005	0.002
Reproduction										
Time to flowering year 2	26	1	27	1	25	1	54	2	56	4
Reproductive output year 2	137	16	33	4	154	22	35	7	95	30
Reproductive output to year 3	172	18	35	4	154	22	198	65	175	47
Root length year 3	131	5	59	6	NA		193	18	285	52
Damage in year 2										
Damage to rosettes	0.51	0.05	0.82	0.04	0.35	0.04	0.32	0.06	0.33	0.08
Damage to inflorescences	0.59	0.05	0.79	0.04	0.94	0.06	0.38	0.05	0.44	0.05
Damage severity rosettes	0.62	0.04	0.20	0.04	0.74	0.03	0.38	0.09	0.65	0.09
Damage severity inflorescences	0.27	0.02	0.44	0.07	0.26	0.01	0.27	0.04	0.29	0.04

Means were calculated based on population means of family means.

Table S7. Summary of models testing for the effect of common garden on population growth rate and several plant traits of different life stages in a transplant experiment of *Arabidopsis lyrata*

		CG1 (Beyond north)			CG2 (North, edge)			CG4 (South, edge)			CG5 (Beyond south)		
Dependent variable	*N*	Estimate		SE	Estimate		SE	Estimate		SE	Estimate		SE
Demographic rate													
r	100	0.153		0.219	-0.35		0.219	**-0.562**	*	0.219	**-0.952**	***	0.235
Germination & survival													
Germination	7,098	**-1.722**	***	0.239	**-0.513**	*	0.237	**0.795**	***	0.241	0.167		0.238
Survival $_{\text{year 1}}$	3,844	**2.424**	***	0.480	**1.602**	***	0.463	**1.736**	***	0.460	-0.626		0.458
Survival $_{\text{year 2}}$	2,205	**-1.062**	*	0.521	**-3.355**	***	0.546	**-3.614**	***	0.532	**-5.045**	***	0.685
Reproduction													
Time to flowering $_{\text{year 2}}$	1,073	0.276		1.633	1.021		1.601	**24.805**	***	1.886	**32.620**	***	2.317
Reproductive output $_{\text{year 2}}$	1,256	-19.747		20.019	**-137.534**	***	19.398	**-135.530**	***	21.955	**-88.346**	***	26.539
Reproductive output $_{\text{to year 3}}$	1,256	5.713		39.087	**-133.515**	***	38.168	62.822		42.167	6.139		49.748
Root length $_{\text{year 3}}$	226				**-66.875**	***	12.666	**82.772**	***	11.658	**150.498**	***	28.645
Damage in year 2													
Damage to rosettes	1,255	**0.305**	***	0.104	**0.585**	***	0.104	0.073		0.107	0.007		0.114
Damage to inflorescences	1,079	0.013		0.154	**-0.386**	*	0.153	-0.184		0.158	0.027		0.166
Damage severity $_{\text{rosettes}}$	676	**-0.303**	***	0.059	**-0.118**	*	0.056	**-0.540**	***	0.071	**-0.477**	***	0.091
Damage severity $_{\text{inflorescences}}$	527	0.021		0.069	**0.185**	***	0.072	0.010		0.075	0.049		0.078

The effect of each common garden was compared with the one in the centre of the distribution (CG3; except for root length – CG1). Germination, survival, damage to rosettes or inflorescences were binary variables, all other variables were continuous. Test statistics include regression coefficients of each fixed effect (*estimate*) and standard error (*SE*). Coefficients are written in bold when $P < 0.05$. Significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. The *bobyqa* optimizer was used to improve model converge.

Table S8. Summary of models testing for the effect of the absolute difference in minimum temperature in early spring between site of origin of populations and transplant site, $|\Delta_{Temp}|$, the absolute difference in precipitation of the wettest quarter, $|\Delta_{Prec}|$, genomic diversity depicted by Tajima's π and interactions on several plant traits in a transplant experiment of *Arabidopsis lyrata*

		$\|\Delta_{Temp}\|$		$\|\Delta_{Prec}\|$		Tajima's π		$\|\Delta_{Temp}\|*\pi$		$\|\Delta_{Prec}\|*\pi$	
Dependent variable	*N*	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Germination & survival											
Germination	7,098	**-0.063** ***	0.019	0.001	0.003	-0.057	0.091	-2E-04	0.005	-1E-05	0.001
Survival year 1	3,844	-0.041	0.030	-0.004	0.003	0.078	0.068	-0.007	0.009	-5E-04	0.001
Survival year 2	2,205	-0.037	0.048	**-0.019** *	0.009	-0.023	0.130	-0.004	0.014	0.002	0.003
Reproduction											
Time to flowering year 2	1,073	0.476	0.293	0.079	0.049	-0.704	0.683	0.017	0.080	-0.007	0.013
Reproductive output year 2	1,256	2.884	3.928	0.155	0.617	16.988 (*)	9.084	-1.332	1.107	0.058	0.164
Reproductive output to year 3	1,256	-11.789 (*)	6.830	-1.380	1.068	1.903	15.072	1.104	1.926	0.267	0.285
Root length year 3	226	-5.610 (*)	3.321	-0.047	0.759	-4.948	8.921	1.067	0.925	0.196	0.198
Damage in year 2											
Damage to rosettes	1,255	0.004	0.010	-0.001	0.002	0.003	0.022	-0.003	0.003	2E-05	-4E-04
Damage to inflorescences	1,079	0.011	0.011	**-0.004** *	0.002	0.022	0.022	-0.005	0.003	0.001	0.001
Damage severity rosettes	676	0.017	0.011	0.002	0.002	0.027	0.028	-0.005 (*)	0.003	-0.001	0.001
Damage severity inflorescences	527	-0.005	0.007	0.001	0.001	-0.012	0.014	0.003 (*)	0.002	-4E-04	-3E-04

Germination, survival, damage to rosettes or inflorescences were binary variables, all other variables were continuous. Test statistics include regression coefficients of each fixed effect (*estimate*) and standard error values (*SE*). Coefficients are written in bold when $P< 0.05$. Significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. The *bobyqa* optimizer was used to help models to converge.

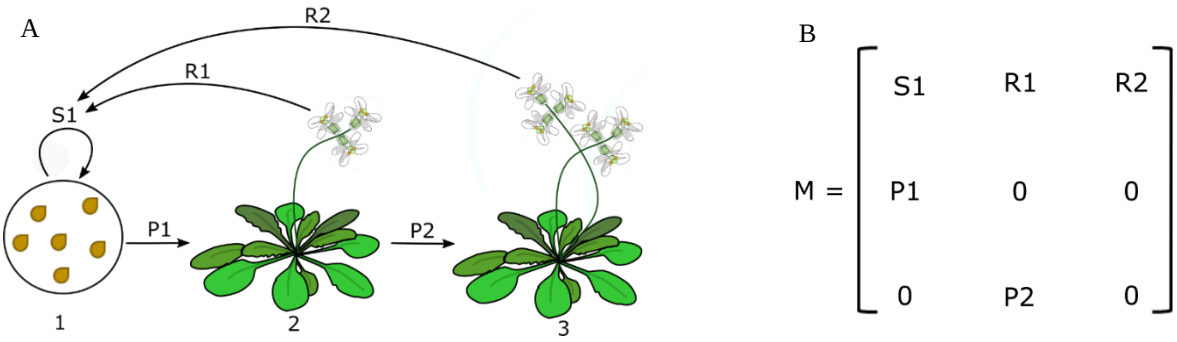

Figure S1: Estimation of population growth. For each population and common garden site, a stage-classified matrix (**panel B**) was constructed assuming three plant stages (**panel A**). The three stages were: 1– healthy seeds, 2– plants in spring of year 2 (2018), 3– plants in spring of year 3 (2019), with a projection interval for each stage set to one year. Survival between stage 1 and 2 (*P1*) was estimated as: germination rate x survival rate from the seedling stage to the reproduction period in year 2. Survival between stage 2 and 3 (*P2*) was calculated as: the survival rate from the first reproductive season to the second reproductive season (year 3). Seeds that did not germinate in the first year could have survived over winter and contributed to the seed pool of the next year, defined as the probability to remain at the same stage (*S1*). This was calculated based on a **seed-burial experiment*** over one winter. The probability to remain at the same stage was set to 0 for both stage 2 and 3, assuming that no plants survived after the third year. Reproduction in stage 2 and 3 (*R1* and *R2* respectively) were estimated as: probability to reproduce x number of fruits (plus fruits that were expected from flowers and buds) x number of healthy seeds per fruit that end in an environment suitable for germination. As we did not have any information on the latter term, we assigned to all populations a standard value leading to an average finite rate of increase in one time-step, λ, of 1 across common gardens. Estimates were $\log_e$-transformed, revealing the population growth rate, *r*.

***Seed-burial experiment**

One hundred healthy seeds per population coming from five to twelve different family lines of a population were pooled and then packed in 10 bags (nonwoven polypropylene-felt, 40 g/m^2), with 10 seeds each. Bags were brought to the transplant sites in fall 2018. In each of the five common gardens, the two bags of each population were split between two spatial blocks. Bags were placed on the ground and covered with a thin mixture of sand and peat. Seeds experienced the same ecological conditions as the plants in the transplant site until early summer 2019. Then bags were collected and examined. We distinguished between germinated and non-germinated seeds. Those seeds that had not germinated were stratified on paper disks saturated with 1.5 ml of 0.05% gibberellic acid in Petri-dishes for 10 days at 4 °C, and no light. Then, germination was assessed once every two days over a period of 20 days. *Seed survival over winter* was calculated for each replicate bag as: (germinated seedlings + germinated seedlings with gibberellic acid)/10.

Chapter 4: Divergent adaptive strategies to cold and frost condition the success of latitudinal range expansion in two *Arabidopsis* sister species

Antoine Perrier*, Yvonne Willi

Department of Environmental Sciences, University of Basel, Schönbeinstrasse 6, 4056 Basel, Switzerland.

*Corresponding author:

E-mail: antoine.perrier@unibas.ch

This manuscript is in preparation for submission in *Journal of Ecology*.

Abstract:

1. Past and present climate warming have forced species to colonize higher latitudes and altitudes. Such range expansion may reflect the occupation of newly available space tracking the species' niche shift, but particularly successful expansion may occur and be linked with niche evolution.
2. Here, we tested if the differential success of northern range expansion after last glacial maximum was linked to an evolutionary shift in coping with cold and frost in the two North American sister species *Arabidopsis lyrata* and *A. arenicola.* The former remained constrained to the northern shore of the Great Lakes, while the latter emerged there from *A. lyrata* and colonized subarctic regions. We tested differences in adaptation to cold and frost by tracking the performance of plants of replicate populations per species, raised under experimentally manipulated temperatures.
3. Subarctic *A. arenicola* was more tolerant to cold growth temperatures and more frost tolerant under mild conditions. In contrast, *A. lyrata* was more tolerant and resistant to frost under cold growth temperatures. The successful colonization of high latitudes by *A. arenicola* was most likely favored by higher tolerance to the generally colder growing season, and by higher frost tolerance during mid-summer, while *A. lyrata* persisted at mid-latitudes by evolving higher protection against frost in the cold periods early and late in the growing season, however constraining northward colonization.
4. *Synthesis:* Two closely related North American *Arabidopsis* species with a common post-glacial origin show differences in adaptation strategies to cold climates assessed in a climate chamber experiment: *Arabidopsis arenicola* was more adapted to cold growth temperatures and frost occurring in the warmer season at high latitudes, favoring the colonization of subarctic regions. *Arabidopsis lyrata,* was more adapted to frost occurring in the colder season at mid-latitudes, but less adapted to climates at higher latitudes, constraining northward post-glacial colonization.

Keywords: Adaptation, *Arabidopsis arenicola*, *Arabidopsis lyrata,* cold tolerance, frost resistance, frost tolerance, range expansion.

Introduction

Species distribution ranges are dynamic. Their fluctuations have been mainly linked to changes in climate, as observed under contemporary climate change (Parmesan 2006; Chen, Hill, Ohlemüller, Roy, & Thomas, 2011; Lenoir & Svenning, 2013; Bellard, Bertelsmeier, Leadley, Thuiller, & Courchamp, 2012), renewing the interest in understanding factors shaping species' range limits (Sutherland et al., 2013). The contemporary distributions of many temperate and arctic species result from past range shifts, expansions or contractions during the Pleistocene glacial cycles (Schluter, 2001; Hewitt, 2000, 2004), with a general trend of colonization toward high latitudes and altitudes after last glacial maximum (LGM, c. 23 000–19 000 years ago; Hughes, Gibbard & Ehlers, 2013). Most subarctic and arctic plant species are thought to have closely tracked the shift in their ecological niche during the last warming period, migrating northwards following the retreat of the ice sheet (Brochmann, Gabrielsen, Nordal, Landvik, & Elven, 2003; Schmitt, 2007; Birks, 2008). In the short evolutionary time since LGM, several taxa have also been able to extend their range while still persisting in areas close to former glacial refugia, resulting in ranges spanning from now temperate climate up to subarctic or arctic climates (*e.g.* Skrede, Eidesen, Portela, & Brochmann, 2006; Koch et al., 2006; Schmickl, Jørgensen, Brysting, & Koch, 2010). The evolutionary shifts required to successfully colonize over broad ecological gradients toward colder climates are not yet fully understood.

In plants, survival under cold climates at high latitudes and altitudes requires adaptation to long periods of non-lethal cold temperatures, and to extreme events of negative temperature (reviewed in Körner, 2016). Cold temperatures constrain development by reducing cellular respiration (Ap Rees et al., 1988) and cell division (Francis & Barlow, 1988). Frost leads to ice formation within tissues, resulting in cell membrane ruptures (Loehle, 1998). Ice forms first in extracellular compartments (Pearce & Ashworth, 1992), drawing water from cells (Thomashow, 1999), resulting in membrane damaging cellular dehydration (Pearce, 2001). Adaptation to cold temperatures has mostly been linked to increased cellular respiration (Wright et al., 2006), *e.g.* by

increasing the numbers of mitochondria (Miroslavov & Kravkina, 1991). Adaptation to frost relies on two mechanisms: tolerance, to mitigate the negative fitness impact of frost damage; and resistance, to reduce damage itself (Agrawal, Conner & Stinchcombe, 2004). Tolerance can be based on increased vegetative and reproductive regrowth after frost events, achieved by increased storage of resources (Baptist & Aranjuelo, 2011) and dormant meristems (Klimešomá & Klimeš, 2007). Frost resistance in plants is well documented (Sakai & Larcher, 1987; Pearce, 2001; Heidarvand & Amiri, 2010; Baxter, 2014; Körner, 2016), and mainly relies on preventing intracellular damage by favoring extracellular ice formation (Körner, 2016). This process is assisted by the production of metabolites such as membrane stabilizing sugars (Heidarvand & Amiri, 2010) or dehydration protective enzymes (Kosova et al., 2008).

Plants occurring in cold climates also adapted to avoid stress for most of their life cycle. Perennial plants survive the most stressful periods by entering a frost-resistant dormancy state over winter (Havranek & Tranquillini, 1995), and only re-initiate the production of sensitive tissue after the strongest frost events (reviewed in Neuner, 2014). Escaping stress is also achieved by small stature, to benefit from warmer conditions close to the ground (reviewed in Körner, 2012). Dormancy and small size also allow to exploit the insulating properties of snow cover in winter (constant 0°C to -5 °C; Larcher, Kainmuller & Wagner, 2010). Deep snow layers take time to melt, preventing premature exposition to stress and favoring delayed phenology until milder conditions (Inouye & Wielgolaski, 2003). Adapting phenology to exploit mild conditions comes at the cost of a short reproductive season, often requiring fast development and fast transition to flowering in the mild season (Prevéy et al., 2017).

While avoidance strategies reduce requirements for frost tolerance and resistance, plants growing in cold climates are still exposed to sharp frost events at the transition between cold and warm seasons (Inouye, Saavedra & Lee, 2003; Inouye, 2008), and even in summer (Taschler & Neuner, 2004; Sierra-Almeida & Cavieres, 2012). Maintaining stress tolerance or resistance is costly (Penning de Vries, 1974; Agrawal et al., 2004; Laureano et al., 2008), leading to reduced growth

(Loehle, 1998; Koehler, Center & Cavender-Barres, 2012; Wos & Willi, 2015; Sebastian-Azcona, Hamann, Hacke, & Rweyongeza, 2019). Frost resistance could also trade-off with reproduction in plants, suggested by the lower resistance of reproductive tissues in frost resistant taxa (Neuner, Erler, Ladinig, Hacker, & Wagner, 2013; Neuner, 2014). Adaptation to cold climates hence requires precise modulations of frost resistance (Sklenář, 2017) or tolerance, often initiated by cold acclimation (Browse & Xin, 2001; Knight & Knight, 2012). Plants surviving over winter with active vegetative tissue, such as short lived perennials or winter annuals, also show reduced developmental senescence in winter (Wingler, Juvany, Cuthbert, & Munné-Bosch, 2015), to maintain the protective role of sugars and to store nutrients. Under warmer temperature, higher recycling of nutrients through increased senescence allow greater investment in growth and reproduction (Davies & Gan, 2012), to exploit the warmer but short growing season.

Despite the extensive literature on adaptive strategies to cold climate, few studies have assessed them in the context of post-glacial range expansion. Adaptive strategies have mostly been investigated in arctic and alpine taxa (*e.g.* Prevéy et al., 2017; Sklenář, 2017) which have closely tracked the retreating ice front, benefiting from pre-adaptation to cold climates (Billings & Mooney, 1968; Birks, 2008). Empirical studies testing adaptive clines toward colder climates also focused either on single species (*e.g.* Colautti & Barrett, 2013; Wos & Willi, 2015), or on pairs of related lowland and alpine species (*e.g.* Kenta, Yamada & Onda, 2011; Ometto, Li, Bresadola, & Varotto, 2012) with little information on the history of colonization. Comparing adaptation strategies between taxa sharing a common post-glacial origin, but with diverging current distributions could allow to characterize the evolutionary shift required for successful post-glacial colonization toward colder climates.

Here we assess the differences in adaptation to cold climates in the North American *Arabidopsis lyrata* subsp. *lyrata* (L.; later referred as *A. lyrata*) and its northern selfing parapatric sister species *A. arenicola* (Fig 1). The contemporary distribution of *A. lyrata* results from post-glacial range expansion, originating from two distinct refugia (Griffin & Willi, 2014; Willi, Fracassetti,

Zoller, & Van Buskirk, 2018). Its northern range limits coincide with its niche limits, defined by minimum temperature in spring (Lee-Yaw, Fracassetti & Willi, 2018), excluding dispersal limitations. Previous studies suggest that northern populations rely on the snow cover to escape recurrent frost in winter, and have adapted to the shorter vegetation period and sharp frost events by evolving toward fast development, higher propensity to flower in the first year, and higher frost resistance (Paccard, Fruleux & Willi, 2014; Wos & Willi, 2015). Adaptation at the northern edge is also supported by genomic and phenotypic signatures of adaptation to temperature in early spring (Walden, Lucek & Willi, 2019; Sánchez-Castro, D., Perrier, A., & Willi, Y., *in prep.*). *Arabidopsis arenicola* occurs North of *A. lyrata,* and its range extends up to the subarctic regions of North America and Greenland (Hopkins, 1937; Mulligan, 1996; in Warwick, Al-Shehbaz & Sauder, 2006). Previous studies suggest *A. arenicola* diverged from *A. lyrata* during latter's post-glacial range expansion, from range-edge selfing populations of the northern shores of Lake Superior (Schmickl et al., 2010; Hohmann et al., 2014; Novikova et al., 2016; Walden, N., & Willi, Y., *in prep.*). *Arabidopsis arenicola* then colonized northwards, while *A. lyrata* expanded eastwards along the Great Lakes (Willi et al., 2018). Which adaptive strategies allowed one species but not the other to colonize further north are unknown, but could involve higher adaptation to cold growth temperatures or frost events to extend the shorter growth season, or on the contrary stronger delay in phenology and fast development to only exploit the mild and frost free season. Here we tested the divergence in adaptive strategy between *A. arenicola* and *A. lyrata* in a climate chamber experiment following three main axes: do both species diverge (*i*) in their tolerance to cold growth temperatures, (*ii*) in their tolerance or resistance to frost events, or (*iii*) in their modulation of growth, phenology or senescence under cold growth temperatures and frost events?

Material and methods

Plant material

We selected four populations of *A. lyrata* from both post-glacial genetic lineages (Willi et al., 2018) and two populations of *A. arenicola* (Fig. 1, Table S1). Seeds were collected from unrelated individuals (seed families) in each population between 2007 and 2017, and stored in separate bags per maternal plant at 4 °C, under dark and dry conditions. To limit maternal effects and effect of storage time, the climate chamber experiment was performed on laboratory generated F1 offspring of the seeds collected in the natural populations. Between 2016 and 2017, we raised one individual of 26 field collected families per population in growth chambers, later transferred in a greenhouse for crossing (see Table S2 for raising conditions). Per population, 12 randomly chosen "mother" plants (pollen recipients) were randomly paired with a remaining "father" (pollen donor), forming a cross combination. Hand pollination was performed on emasculated buds to exclude cross- and spontaneous self-pollination. Each cross combination was repeated to obtain enough healthy seeds for the climate chamber experiment (at least five siliques). If cross combinations failed, the father was replaced by a backup plant. Mature siliques were collected and dried two weeks at ambient temperature in the dark. Seeds were stored at 4 °C under dry and dark conditions. In total, we obtained 65 cross combinations (10 to 12 per population, Table S3) for the climate chamber experiment.

Climate chamber experiment

F1 individuals were raised in climate chambers (Climecab 1400, Kälte 3000 AG, Landquart, Switzerland) simulating to some extent the natural growth cycle of both species. To test adaptation to cold, two temperatures were setup for the growth phases: *mild* (20 °C) and *cold* (12 to 14 °C), simulating conditions close to the home climate of *A. lyrata* and *A. arenicola*, respectively. These temperatures were derived from extrapolations of the average temperatures in late summer (July and August) of the selected populations (WorldClim database version 2.0, Fick & Hijmans, 2017; Table S1, Fig. 1). To test for adaptation to frost (tolerance and resistance) as it may occur in fall and spring,

we simulated recurrent frost of low intensity. Each growth temperature setup was split in two treatments: *frost* or *control* (= no frost), resulting in a two-by-two factorial design, for a total of four growth temperature-treatment combinations (condition), each setup in one of four climate chambers. Three spatial blocks were assigned to each climate chamber resulting in 12 blocks, weekly re-positioned within each climate chamber, or between climate chambers simulating the same growth temperature when frost was not applied.

Per cross combination, two seeds (one if less than 24 seeds) were sown in twelve pots, filled with a standard substrate mixture of washed river sand and peat (1:1.5 sand:peat). Pots were randomly assigned to one of twelve replicate blocks across the four conditions, and randomly distributed across two 54-cell propagation trays within block (BK Qualipot, Otelfingen, Switzerland). In total, 1428 seeds were sown over 864 pots (Table S3, 6 populations * 12 seed families * 4 conditions * 3 blocks). Pots were watered to saturation and placed in the climate chambers for 20 days at 4 °C in the dark to stratify seeds. Seedlings were randomly thinned to one individual 28 days after germination.

The experiment started with a fall phase simulating environmental conditions typical for fall during germination and initial vegetative growth, followed by a winter phase with the effect of a snow cover simulated by vernalization (constant 4 °C, low light). A third spring phase simulated spring and summer with further vegetative growth and reproduction. Duration, day length, light intensity and temperatures of each phases are detailed in Table S4. Frost was applied for two weeks before (*frost I*) and for six weeks after (*frost II*) the simulated winter. The temperature was decreased to 4 °C during the night for three consecutive nights to acclimate the plants, followed by four nights at -4 °C. One hour after the start of the night phase, the temperature declined gradually to reach the target temperature at the centre of the night phase. This temperature was kept for one hour, then gradually increased to reach the night temperature of each condition one hour before the end of the night phase. This cycle was repeated twice for *frost I*, for a total of 14 days, and six times for *frost II* for a total of 42 days followed by an additional 9 nights at 4 °C. The experiment ended 213 days after initiation of germination. Trays were regularly watered ensuring constant substrate moisture. Fertilizer was added

every two weeks starting 87 days after initiation of germination (2% v/v Wuxal universal fertilizer, Hauert Manna Düngerwerke GmbH, Nürnberg, Germany).

Individual performance estimates

Individual plant performance was recorded at the level of the seedling until thinning, later at the level of the individual for the whole length of the experiment (recording rates detailed in Table S4). Day of germination, defined as when a seedling had two fully open cotyledons, and death of seedlings were checked five days a week until the peak of germination was over (four weeks after sowing), then one to three times a week. Rosette damage was recorded as visual estimation of the proportion of the rosette affected by discoloration, desiccation or necrosis, split into five classes (0 = 0 %, 1 = < 25 %, 2 = < 50 %, 3 = < 75 %, 4 = > 75 %). Damage was recorded starting on the last day before initiation of the first frost treatment, five days a week in the first frost treatment, then only once a week for the rest of the experiment. After vernalization, we scored the day of first flower opening four days a week for six weeks, later one to three days a week. Day of germination, flowering or death were corrected by the mid-time between previous checking and actual observation. Female reproductive output was estimated for each individual four weeks after the first flower opening, by counting the number of fruits (populations MB1, ON11 and QC1 were autonomously selfing), pedicels (flowers that did not develop into a fruit), open flowers, and flower buds on all inflorescences. On the same day, the length of the two longest leafs of the rosette were measured as estimation of the rosette radius.

Growth

Growth in simulated fall was estimated from photographs of every propagation tray, with pictures being taken 14 days after germination until 53 days after germination (second week of vernalization) three times a week, and then once a week. For each photograph, we estimated the radius of each rosette by averaging the length of the two longest leaves measured with ImageJ v1.53c (Rasband,

W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/, 1997-2020). For each individual, we fitted a three-parameter logistic function modeling the relationship between rosette radius and date of measure since germination (supported by preliminary model selection fitting seven alternative models: linear, exponential, power, two-, three-parameter logistic, Gompertz and Bertalanaffy, data not shown) using the package *drc* (Ritz, Baty, Streibig, & Gerhard, 2015). The parameters generated by this function were the asymptotic size, the scale parameter (1/growth rate *r*), and X_{mid} as the time until 50% of the asymptotic size was reached.

Electrolyte leakage assay

Resistance of leaves to frost damage was assessed by inferring the temperature at which 50% of tissue death occurred (LT_{50}) from measures of electrolyte leakage (EL) due to cell membrane damage by ice formation in tissues, adapting the protocol of Armstrong, Takebayashi, Sformo and Wolf (2015). Electrolyte leakage is quantified by measuring the conductivity of distilled water in which damaged tissues are incubated. We performed this assay on leaves from plants of the *mild control* condition without frost two weeks after vernalization. We collected six leaf discs per individual at the tip of fully expanded leaves of similar size, on the upper layer of the rosette. Each disc was placed in individual 15 mL centrifuge tubes and stabilized two hours at 5°C, dispatched in six programmable freezers. Each freezer was equipped with built-in temperature sensors, and two additional temperature loggers (iButton®, Maxim Integrated Products, Inc) setup to record temperature every 10 minutes. Tubes were then cooled at a rate of 4°/h to reach one of the six target temperature: 5 °C, -5 °C, -10 °C, -15 °C, -20 °C and -25 °C, maintained for one hour. The temperature was then increased to 5° C at a rate of 4°/h and stabilized at 5° C for 10h to allow for complete thawing of the leaf discs. To measure the initial electrolyte leakage (EL_I), we added 3ml of ddH_2O to each tube and allowed electrolytes to diffuse for two hours before transferring leaf discs into new tubes containing 3ml of ddH_2O. These were boiled in a water bath (Julabo TW20, HuberLab, Aesch, Switzerland) at 95°C for

30 minutes to allow for complete tissue destruction, then incubated at room temperature overnight to ensure maximal electrolyte diffusion for the measure of total electrolyte content (EL_T). EL_I and EL_T were measured using a conductivity meter (Fe30/EL30, Mettler Toledo, Columbus, USA). Each measure was corrected by the conductivity measured in control tubes without leaf discs, subjected to the same treatments as detailed above. The relative electrolyte leakage (EL_R) was calculated for each individual as $EL_I/(EL_T + EL_I)$. LT_{50} of each individual was then estimated as the temperatures at which 50% of EL_R is reached, by fitting a four-parameter logistic function modeling the relationship between EL_R and the measured temperature within each freezer.

Statistical analyses

The main dependent variable was *multiplicative performance*, calculated in each condition at the pot level, as the product between the germination rate within a pot, and the reproductive output (set to 0 if individuals died or did not flower). To assess the specific effect of cold and frost on the life cycle of both species we tested the variation in 13 additional traits, summarized in Table 1: We tested the individual components of multiplicative performance *germination*, as the binary success to germinate over the whole length of the experiment, and *reproductive output* assessed on plants that produced inflorescences. We assessed the effect of the two successive *frost* treatments on three variables, duplicated for both *frost I* and *II*: *Survival under frost I* and *II* were assessed as binary success of plants alive at the beginning of the respective *frost* treatment until 14 days after the end of each treatment, to allow for individuals that were completely damaged to recover and be recorded as alive. *Frost damage I* and *II* were assessed as binary increase in damage (0 = no variation, 1 = increase in damage), recorded during each *frost* treatment. In addition, the frost resistance of leaves in *mild* growth temperatures was inferred from the LT_{50} estimated by the electrolyte leakage assay. We assessed variation in phenology based on *time to germination* as the number of days from the end of stratification to germination, and on *time to flowering* as the number of days from the end of vernalization to first flower opening. We assessed growth in simulated fall based on the *growth rate*

parameter described above. Growth rate was only considered for the *mild* growth temperature, as individuals in *cold* germinated too late to fit growth curves. We additionally assessed growth in all conditions based on *size at flowering*, the average rosette radius measured while recording reproductive output. Finally, we assessed *senescence I* during the simulated winter phase and *senescence II* during the simulated spring phase as the binary increase in damage recorded 14 days after the end of each treatment, until the start of *frost II* for *senescence I*, and until the end of the experiment for *senescence II.* G*ermination*, *survival to frost I* and *II*, *frost damage I* and *II* and *senescence I* and *II* were assumed to follow binomial distributions, LT_{50} a normal distribution, and *reproductive output*, *time to germination*, *time to flowering*, *growth rate* and *size at flowering* were assumed to follow log-normal distributions.

Adaptation to non-lethal cold growth temperatures

We considered two aspects of adaptation to cold growth temperatures: tolerance, *i.e.* the capacity to maintain performance (similarly to frost tolerance, Agrawal et al., 2004), and the modulation of growth, phenology and senescence. This analysis considered only individuals from the *control* treatment of both *mild* and *cold* growth temperatures. In our main analysis, we assessed variation in tolerance to cold growth temperatures by testing the variation in *multiplicative performance* estimated on the level of the pot in hierarchical mixed-effects models. Fixed effects were the categorical variables *species*, estimated on three levels: *A. arenicola* (populations QC1, MB1), northern *A. lyrata* (*Lyr N*: ON11, NY5) and central *A. lyrata* (*Lyr C*: WI1, MD2), as well as *growth temperature* (*mild* = 0, *cold* = 1) and their interactions. Random effects were maternal family nested within population, and maternal population, crossed with the effect of block. *Multiplicative performance* was 0 inflated, suggesting the modelling of a Gaussian process (log_{10}-transformed for values > 0), and a logistic process (modelling the probability of 1, assigned to values > 0). Analyses were performed in a Bayesian framework, with the package *MCMCglmm* (Hadfield, 2010, 2019) in *R* (R Core Team 2019) on 10 parallel chains (model and prior parametrization detailed in Appendix S5). The contribution of

each fixed effect was assessed by comparing DIC values of the full model with three alternative models: one excluding *species*, one excluding *growth temperature*, and one with both fixed effect without interaction. For the fixed effects which removal led to a lower model fit, Tukey's tests were performed using the package *emmeans* (Lenth, 2019) to test the significant difference between each level of the fixed effect. Comparisons among *species* included the difference between *A. arenicola* and *Lyr N* as well as between *Lyr N* and *Lyr C*, within the *mild* or the *cold* growth environment. Comparisons between *growth temperature* included the difference between *cold* and *mild* growth temperatures within each *species* level. For the interaction between *species* and *growth temperature*, the contrast targeted the comparison of slopes on *growth temperature* between *A. arenicola* and *Lyr N* to test for differences in adaptation between species, as well as between *Lyr N* and *Lyr C* to test for differences in adaptation within *A. lyrata*. Differences between *Are* and *Lyr C* were out of the scope of this study, and are therefore not reported.

In secondary analysis, we assessed differences in tolerance to cold growth temperatures by testing the variation in the individual components of *multiplicative performance*: *germination* assessed on the level of the seed, and *reproductive output*, on the level of the pot, with the same hierarchical mixed-effects models structure as above, using restricted maximum likelihood with the packages *lme4* (Bates, Mächler, Bolker, & Walker, 2015) and *LmerTest* (Kuznetsova, Brockhoff & Christensen, 2017; model parametrization in Appendix S6). An ANOVA was performed on each model to test for the significance of each fixed effect. Tukey's tests were applied only on fixed effects significant in the ANOVA. We further assessed the effect of *species* and *growth temperatures* on the capacity to modulate growth, phenology and senescence by testing *size at flowering*, *time to germination*, *time to flowering* and *senescence I* and *II* following the exact same model structure and analysis based on maximum likelihood detailed above.

Adaptation to frost

We considered three aspects of adaptation to frost: tolerance, *i.e.* the capacity to maintain performance under frost (Agrawal et al., 2004); resistance, *i.e.* the capacity to limit frost damage or survive to frost, as well as the modulation of growth, phenology and senescence. This analysis was performed in parallel on two subsets, considering only individuals from both *control* and *frost* treatment of the *mild* growth temperatures in the first subset, and of *cold* temperatures in the second subset. In our main analysis, frost tolerance was assessed by testing the variation in *multiplicative performance* in similar hierarchical mixed-effects models performed in a Bayesian framework as used to test adaptation to cold temperatures, with the effect of *treatment* (*control* = 0, *frost* = 1) replacing *growth temperature*.

In secondary analysis, we assessed frost tolerance by testing the variation in *reproductive output* with the same hierarchical mixed-effects models structure as detailed above, using restricted maximum likelihood. *Germination* was not considered as most seedlings germinated before frost events. Further analyses tested three components of frost resistance: survival and rosette damage to assess resistance to recurrent frost events at the level of the individual, and LT_{50} to assess the frost resistance of leaf tissue to single frost events. We tested for variation in *survival to frost I* and *II* and *frost damage I* and *II* with the exact same hierarchical mixed-effects models structure as detailed above. LT_{50} was only assessed on individuals of the *mild* growth temperatures with no frost, and was therefore tested in hierarchical mixed-effects models with as fixed effects *species*, and as random effects maternal family nested within population, and maternal population, crossed with the effect of block. Similarly as for the adaptation to cold, we assessed the effects of *species* and *treatment* on the modulation of growth, phenology and senescence by testing *size at flowering, time to germination, time to flowering* and *senescence I* and *II* following the same analysis as detailed above.

Finally, we tested whether requirement of cold acclimation to increase frost resistance varied between species, considering only individuals from the *frost* treatments of both growth temperatures. We tested *survival to frost I* and *II* and *frost damage I* and *II* using the same hierarchical mixed-effects models structure as above, with as fixed effect *species, growth temperatures* and their interaction.

Results

Adaptation to non-lethal cold growth temperatures

In our main analysis assessing variation in tolerance to cold growth temperatures, the model comparison on the hierarchical mixed-effects model analyses testing the effects of *species*, *growth temperatures* (*cold* vs *mild* [0]) and their interactions on *multiplicative performance* revealed that only the interaction between fixed effects contributed to the fit of the full model (Table S7, mean values in Table S8), suggesting that the effect of *growth temperature* differed between *species*. The interaction between *species* and *growth temperature* was significantly positive comparing *A. arenicola* to northern *A. lyrata* in both log-normal and logistic processes (Table 2A, Fig. 2), indicating a more positive effect of *growth temperatures* on *A. arenicola.* On the contrary, *northern A. lyrata* were significantly more negatively affected by *growth temperatures* in both log-normal and logistic process than central *A. lyrata*.

In secondary analyses testing the effects of *species*, *growth temperatures* and their interactions on individual performance estimates, the effect of *species* was significant only for *senescence I,* significantly higher in *A. arenicola* compared to northern and A. *lyrata* populations, under *mild* and *cold* growth temperatures (Table 2B, Fig. 3, ANOVA reported in Table S9), and significantly higher in northern compared to central *A. lyrata* populations under *mild* growth temperatures. For all species, the effect of *growth temperatures* significantly reduced *germination.* The effect of *growth temperatures* significantly increased *reproductive output* in *A. arenicola*, but significantly decreased *reproductive output* in both northern and central *A. lyrata*. The effect of *growth temperatures* further significantly increased *time to germination* and *time to flowering* in all species. When comparing *A. arenicola* to northern *A. lyrata,* the interaction comparing the effect of *growth temperatures* between *species* was significantly positive for *germination, reproductive output, size at flowering* and *senescence II,* suggesting a stronger positive effect of *growth temperatures* on *A. arenicola*. This interaction was negative for *time to flowering*. When comparing northern to central *A. lyrata*

populations, this interaction was significantly negative for *germination* and significantly positive for *time to germination.*

Adaptation to frost events

In our main analysis assessing the variation in tolerance to frost, the model comparison testing the effects of *species*, *treatment* (*frost* vs *control* [0]) and their interactions on *multiplicative performance* revealed that only the interaction between fixed effects contributed to the fit of the full model in both mild and cold growth temperatures subsets (Table S10, Table S11A), suggesting that the effect of *treatment* differed between *species*. Under mild growth temperatures, *A. arenicola* was significantly more positively affected by the effect of *treatment* than northern *A. lyrata* populations in the log-normal process (Table 3A, Fig. 2, marginally significant in the logistic process), while northern *A. lyrata* were more negatively affected than central *A. lyrata.* Under cold growth temperatures, *A. arenicola* was more negatively affected than northern *A. lyrata* in the log-normal process (Table S11B, Fig. 2), and no differences were observed between northern and central *A. lyrata* populations.

Under mild growth temperatures, secondary analyses testing the effects of *species*, *treatment* and their interactions on individual performance estimates revealed that the effect of *species* was significant for *frost damage II*, significantly lower in northern compared to central *A. lyrata* in the *control* treatment (Table 3B, Fig. 3, ANOVA reported in Table S12). Furthermore, *senescence I* was significantly higher in *A. arenicola* compared to northern *A. lyrata* populations in both frost treatments, and significantly higher in northern compared to central *A. lyrata* in the *control* treatment. The effect of *treatment* significantly increased *reproductive output* in *A. arenicola*, *frost damage II* for *A. arenicola* and northern *A. lyrata*, and *time to flowering* in all species, but significantly reduced all species' *size at flowering* and *senescence II* (except central *A. lyrata*). The interaction comparing the effect of *treatment* between *species* was significantly positive for *reproductive output* when comparing *A. arenicola* to northern *A. lyrata*, and significantly negative for *senescence I* when comparing northern to central *A. lyrata*.

Under cold growth temperatures, secondary analyses testing the effects of *species*, *treatment* and their interactions on individual performance estimates revealed that the effect of *species* was significant for *frost damage II*, significantly lower in in *A. arenicola* compared to northern *A. lyrata* populations in the *control* treatment, and significant for *senescence I*, significantly higher in *A. arenicola* compared to northern *A. lyrata* populations in the *control* treatment (Table S13A, Fig. 3, ANOVA reported in Table S13B). The effect of *treatment* significantly reduced *reproductive output* and *senescence I* but increased *frost damage II* in *A. arenicola*, and increased *time to flowering* in all species. The interaction comparing the effect of *treatment* between *species* was significantly negative for *senescence I* but significantly positive for *frost damage II* when comparing *A. arenicola* to northern *A. lyrata.*

Finally, the hierarchical mixed-effects model analysis testing the effect of *species*, *growth temperatures* (*cold* vs *mild* [0]) and their interaction on dependent variables linked to frost resistance revealed no significant differences between *species* in both *mild* and *cold* growth temperatures (Table S14A, ANOVA reported in Table S14B). The effect of *growth temperatures* was significantly negative for *frost damage I* and *II* in all species. Interaction were generally not significant.

Discussion

We found clear differences in thermal tolerance in two North-American *Arabidopsis* species which can explain well their divergent post-glacial colonization patterns of northern areas. *Arabidopsis lyrata* originally expanded its range from two glacial refugia, but remained constrained to the northern shores of the Great Lakes (Willi et al., 2018). In contrast, *A. arenicola* split from *A. lyrata* at the northern shores of Lake Superior and expanded in the subarctic regions of North America and Greenland (Hopkins, 1937; Mulligan 1996; in Warwick et al., 2006). The species differed mainly in their tolerance to cold growing temperature, and their frost tolerance and resistance. *Arabidopsis arenicola* shows higher tolerance to cold growing temperatures, to compensate the shorter growing

season, and increased frost tolerance under mild growth temperatures, to compensate the occasional frost events occurring during mid-growing season. In contrast, *A. lyrata* shows lower tolerance to cold growing temperature, escaping the coldest month to exploit the warmer conditions in summer, and higher frost resistance and tolerance to frost events occurring in early spring, but reduced tolerance in summer.

Similarities in adaptive strategies to cold climates across species.

Our results outline similarities in the response of all populations from both species to environmental constraint linked to cold climate. Cold growth temperatures generally delayed time to germination and time to flowering, latter further delayed by frost under mild and cold growth temperatures. Frost also generally decreased size at flowering under mild growing temperatures, potentially due to reduced metabolism and cell division under the recurrent lower night temperatures (Ap Rees et al., 1988; Francis & Barlow, 1988). Both species were generally frost resistant, with high survival under frost (on average between 84 % and 96 %), and low average LT_{50} across populations (*ca.* -11 °C), with no differences between populations in these traits. Both species showed only variation in resistance to frost based on rosette damage. Latter was generally reduced by cold growth temperatures, indicating that both species rely on cold acclimation to increase frost resistance (Browse & Xin, 2001; Knight & Knight, 2012), in line with previous observation in *A. lyrata* (Wos & Willi, 2015). Overall, these results suggest both species have generally evolved toward escaping cold growth temperatures and frost events and exploiting the warmest frost-free season, while maintaining a general resistance to frost.

Differences in adaptation between A. arenicola and northern A. lyrata

Both species however also show clear differences in their response to cold growing temperatures or frost events. *Arabidopsis arenicola* showed higher tolerance to cold growth temperatures compared to northern *A. lyrata*, suggested by a better capacity to maintain multiplicative performance, both in

the zero and normal part of distribution. This difference was also reflected by the analysis of life-stage components: germination rates were reduced in both species under cold compared to mild growth temperatures, but significantly less in *A. arenicola*. Reduced seed dormancy is generally observed at higher latitudes (Wagmann et al., 2012; Debieu et al., 2013). In *A. thaliana*, higher dormancy at lower latitudes has been linked to dryer summers requiring to delay germination to fall (Kronholm, Pico, Alonso-Blanco, Goudet, & de Meaux, 2012; Postma & Ågren, 2015). Lower dormancy in *A. arenicola* could results from adaptation to generally lower temperatures during the vegetation period. The difference in tolerance to cold growth temperature was also reflected by the significant contrast in slope of effect of growth temperatures on reproductive output, increasing in *A. arenicola* but decreasing in northern *A. lyrata*. The significant increase in reproductive output suggests that *A. arenicola* is able to invest more resources in reproduction in response to cold growth temperatures.

Both species also showed differences in their modulation of phenology, growth and senescence when exposed to cold compared to mild growth temperatures: while flowering was delayed in all species, *A. arenicola* was less affected than northern *A. lyrata*, in opposition with the previous pattern of delayed flowering toward higher latitudes and altitudes reported in several related *Arabidopsis* taxa (Riihimaki & Savolainen, 2004; Stinchcombe et al., 2004; Kuittinen, Niittyvuopio, Rinne, & Savolainen, 2008; Montesinos-Navarro, Wig, Pico & Tonsor, 2010; Kenta et al., 2011). The lower effect of cold growth temperatures on flowering time in *A. arenicola* could allow this species to extend the short growth season at high latitudes. *Arabidopsis arenicola* seemed also better at maintaining size at flowering than northern *A. lyrata*, despite earlier flowering, suggesting higher investment in growth. Senescence in simulated winter was higher in *A. arenicola* compared to northern *A. lyrata*, but not affected by the growth temperatures. Senescence in simulated spring was generally high in both species (observed in ca. 92% of the individuals), but was less reduced by cold growth temperatures in *A. arenicola* than northern *A. lyrata*. Reduced senescence in simulated winter of high elevation populations of *Arabis alpina* has been linked to a strategy of maintaining of high

concentration of metabolites such as sugars to protect against frost damage and store resources in leaves (Wingler et al., 2015), followed by rapid recycling of these sugars through increased senescence, to invest in reproduction (Davies & Gan, 2012; Wingler et al., 2015). Here, the difference in senescence rates between both species suggest *A. arenicola* could have adapted to a shorter growth period than in northern *A. lyrata* by maintaining higher metabolic rates and resource recycling in winter, therefore benefiting from a head start in fast vegetative and reproductive development in spring. Furthermore, the higher maintenance of senescence rate in simulated spring in *A. arenicola* suggest that its higher tolerance to cold growth temperatures could result from a better capacity to maintain developmental senescence, maximally reinvesting resources in reproduction.

Under mild growth temperatures, *A. arenicola* also showed higher frost tolerance than northern *A. lyrata*, suggested by a better capacity at maintaining multiplicative performance in the log-normal process, and an increase in reproductive output compared to *A. lyrata*. This pattern is in opposition with previous observations of lower frost tolerance in northern compared to central and southern populations of *A. lyrata* (Wos & Willi, 2015). In latter study, lower tolerance was linked to the cost of the general adaptation toward fast growth and reproduction to exploit the warmer and frost free growing season. Higher frost tolerance in *A. arenicola* could be required to cope with the occasional frost events which can occur in the warmest months at high latitudes (Billings, 1974). Senescence in simulated winter was again higher in *A. arenicola* than northern *A. lyrata,* which could indicate that *A. arenicola* recycles faster resources dedicated to frost resistance into reproduction (Wingler et al., 2015), allowing for higher frost tolerance.

The opposite pattern was observed under cold growth temperatures: *A. arenicola* was less frost tolerant than northern *A. lyrata,* suggested by reduced capacity to maintain multiplicative performance under frost. This result is partially supported by a reduction in reproductive output in *A. arenicola* when exposed to frost, however not differing from northern *A. lyrata*. *Arabidopsis arenicola* also showed reduced frost resistance in simulated spring under cold growth temperatures compared to northern *A. lyrata,* as suggested by the increase in frost damage in *A. arenicola*,

significantly different from northern *A. lyrata*. Reduced frost resistance in high- compared to mid-latitude populations contradict the general patterns of increasing frost resistance toward higher latitudes in plants (Hurme, Repo, Savolainen, & Paakkonen, 1997; Pagter, Kristoffersen, Bronnum, & Jensen, 2010), also in *Arabidopsis* and related taxa (Hannah et al., 2006; Zhen & Ungerer, 2008; Ometto et al., 2012; Zuther, Schulz, Childs, & Hincha, 2012; Horton, Willems, Sasaki, Koornneef, & Nordborg, 2016; reviewed in Zuther, Lee, Erban, Kopka, & Hincha, 2018). Similar pattern as in our study was only reported in the closely related *A. kamchatica* (Armstrong et al., 2015), linked to the longer snow cover at high latitudes protecting against frost in the coldest months, requiring less frost resistance than populations at mid-latitudes exposed to frost in the cold season due to earlier snow melt. *Arabidopsis arenicola* showed again higher senescence in simulated winter than northern *A. lyrata* in the control treatment, but not under frost treatment. frost reduced senescence in simulated winter in *A. arenicola*, with a significant contrast in slope of effect compared to *A. lyrata*. These results suggest that *A. arenicola* modulates senescence in winter after frost exposure down to rates similar than *A. lyrata*, potentially to maintain resistance against frost in the coldest months (Wingler et al., 2015), however not sufficient to prevent frost damage in simulated spring. Seasonal change in frost defense mechanisms are common in taxa growing at high latitude and altitude (Inouye, 2008; Sierra-Almeida, Cavieres & Bravo, 2009; Preston & Sandve, 2013; Sklenář, 2017), but resistance generally increases at the transition phases between the cold and mild season. The downregulation of senescence in *A. arenicola* could also be linked to the reduced tolerance to frost in cold growth temperatures as outlined above, due to lower investment in reproduction to maintain higher resistance.

Differences in adaptation between northern and central A. lyrata

Populations of *A. lyrata* also showed differences in their response to cold growth temperatures and frost events: Northern *A. lyrata* populations showed reduced tolerance to cold growth temperatures compared to central *A. lyrata*, suggested by a lower capacity at maintaining multiplicative performance, both in the zero and normal part of the distribution. This difference was

also reflected by the analysis of individual life stage components: germination rates were significantly declined in both northern and central *A. lyrata* populations under cold compared to mild growth temperatures, but significantly more in northern populations. Reproductive output significantly declined in both northern and central populations, however with no significant difference between populations. Northern and central populations also showed differences in their modulation of germination and flowering phenology: northern populations showed a higher delay in germination and initiation of flowering than central populations when exposed to cold compared to mild growth temperatures, in lines with patterns of delayed flowering toward higher latitudes observed in related *Arabidopsis* taxa (Riihimaki & Savolainen, 2004; Stinchcombe et al., 2004; Kuittinen et al., 2008). Senescence in simulated winter was also higher in northern populations compared to central populations under mild growth temperatures, suggesting higher resource recycling (Davies & Gan, 2012; Wingler et al., 2015). Higher metabolic rates in northern population are in line with the finding of faster vegetative and reproductive development in northern populations to cope with the shorter growing season (Paccard et al., 2014; Wos & Willi, 2015).

Northern *A. lyrata* showed reduced frost tolerance under mild growth temperatures than central populations, suggested by a lower capacity at maintaining multiplicative performance in the logistic part, in line with patterns of reduced frost tolerance previously reported in this species (Wos & Willi, 2015). Other traits showed no variation between northern and central populations except senescence in simulated winter, which was again higher under control treatment, but significantly reduced under frost, indicating a modulation from high recycling of metabolites in the control treatment to reduced recycling under frost, potentially to maintain higher frost resistance (Davies & Gan, 2012; Wingler et al., 2015), while central populations could maintain a baseline of frost resistance. Northern populations showed also less damage in the control treatment during the frost period in simulated spring, which could result from a reduction in senescence rates after vernalization, necessary to maintaining frost resistance in early spring.

Divergence in adaptive strategies between both species

Both species have evolved divergent adaptive strategies to cold climate, allowing *A. arenicola* to successfully colonize high latitudes but constraining *A. lyrata* to mid-latitudes. *Arabidopsis arenicola* is able to exploit the generally colder temperatures at high latitudes, allowing it to extend the shorter growth season to complete its reproductive cycle, as suggested by its higher tolerance to cold growth temperatures, and lower reduction in time to flowering. *Arabidopsis arenicola* most likely remains completely protected by snow during the cold period, explaining its low frost resistance and tolerance under cold growth temperatures. Under the milder growth temperatures, *A. arenicola* then increases reproduction to tolerate the negative effects of frost, which can occur also in the warmest months at high latitudes (Billings, 1974). In contrast, northern *A. lyrata* seem to have evolved a strategy oriented toward escaping cold temperature and exploiting mostly the milder months in summer, as suggested by the lower tolerance to cold growth temperatures, and stronger increase in time to germination and flowering. This strategy reduces its growing season, requiring faster growth and reproduction compared to more southern populations (Wos & Willi., 2015), but potentially to the cost of reduced frost tolerance in summer. Northern populations of *A. lyrata* seem to have also evolved toward higher frost tolerance and resistance under cold growth temperatures compared to *A. arenicola* potentially linked to higher frost exposure due to earlier snow melt in winter at mid-latitudes, similar to patterns observed in *A. kamchatica* (Armstrong et al., 2015). While this adaptive strategy might have allowed *A. lyrata* to colonize mid-latitudes, persistence at higher latitudes could be constrained by its lower tolerance to frost under milder temperatures, and its delayed phenology could affect its capacity to complete its reproductive cycle.

Maximal post-glacial expansion of subarctic species

These divergent adaptive strategies despite a recent common ancestor (Novikova et al., 2016; Walden, N. & Willi, Y., *in prep*) raises the question whether maximal post-glacial range expansion is most successful under niche conservationism or niche expansion. Most subarctic and arctic plant

species are thought to have migrated northwards after LGM benefiting from the release of new habitats (Brochmann et al., 2003; Schmitt, 2007; Birks, 2008). Higher adaptation to climate at high latitudes but lower adaptation to climate at mid-latitudes indicate that *A. arenicola* may have followed a similar scenario: this species could have benefited from initial pre-adaptation to cold climates and tracked the shift in its ecological niches rather than experiencing new evolutionary shifts. The lower frost resistance and tolerance of *A. arenicola* to frost in cold growth temperatures could then have led to local extinctions at lower latitudes during the warming period after LGM, due to earlier snowmelt and increased frost exposure during the colder season, explaining the gap in distribution between both species. Phylogenetic studies of the *Arabidopsis* genus support a subarctic origin of the North American *A. lyrata* subsp. *lyrata* which likely split from the European lineage *A. lyrata* subsp. *petrea* and colonized North America over Beringia (Shmickl et al., 2010). The North American *A. lyrata* lineage is also connected to the European lineage by two arctic taxa: *A. petraea* subsp. *umbrosa* and *A. petraea* subsp. *septentrionalis* (Hohmann et al., 2014; Novikova et al., 2016). On the contrary, divergence in adaptation in northern *A. lyrata* could result from niche-expansion, allowing persistence at mid-latitude. However, the evolution of a strategy of fast vegetative and reproductive development to exploit only the mildest conditions could trade-off with adaptation to frost in the mid-growing season (Wos & Willi, 2015), necessary to colonize higher latitudes. In addition, a strategy oriented to exploit only the mildest conditions could reduce its capacity to complete its life cycle in the generally colder environments at higher latitudes. The evolution to persist at mid-latitudes could therefore represent an evolutionary dead-end constraining future northward colonization, suggesting post-glacial range expansion toward cold climates is more successful under niche conservationism.

The successful colonization of higher latitudes by *A. arenicola* could also have been favoured by its entirely selfing mating system: Reduced recombination rates in selfing populations have been predicted to increase the response to selection on polygenic traits by converting dominance and epistasis in additive variance (Lande & Porcher, 2015), and allowing better storage of cryptic genetic variation, which can be released after an environmental shift (Clo, Ronfort & Abu Awad, 2020),

leading to successful expansion over wide geographic ranges (Eriksson & Rafajlovi, 2020). A recent niche modeling study also found that selfing species generally show larger niche breadth toward more stressful environments compared to their outcrossing sister species (Grant & Kalitz, 2020). In contrast, the divergence in adaptation in northern *A. lyrata* could also result from increased genetic drift, recently identified as a main factor constraining adaptation along environmental gradients (Polechová & Barton, 2015; Polechová, 2018), as it opposes selection and erodes genetic variation (Wright, 1931). In line, northern range limits of *A. lyrata* are characterized by steep environmental gradients (Lee-Yaw et al., 2018), and range-edge populations suffer of increased genetic drift (Willi et al., 2018; Perrier, Sánchez-Castro & Willi, 2020), linked to reduced signatures of adaptation (Sánchez-Castro, D. et al., *in prep.*). The ecological niche and the levels of genetic drift in *A. arenicola* are however yet to be tested.

Conclusions

Our study is one of the first to empirically highlight how differences in adaptation could explain the difference is success of post-glacial colonization toward subarctic climate between two recently diverged sister species. The increased tolerance to cold growing temperatures and to frost in summer allowed *A. arenicola* to colonize subarctic climate, while the opposite adaptive strategy constrained *A. lyrata* to more temperate climates. In the wake of global warming, and the higher risk of exposure to early frost at higher latitudes (Inouye, 2008; Lancaster & Humpfreys, 2020), both species could be suitable model organisms to study the impacts of climate change on recently evolved subarctic and arctic plants.

Acknowledgements: This work was supported by the Swiss National Science Foundation (31003 A_166322). For assistance with crosses, seed counts, sowing, leaf measurements and electrolyte leakage assays, we thank Georg Armbruster, Olivier Bachmann, Markus Funk, Maja Pfund, Susanna Riedl and Darío Sánchez-Castro. Collection permits were provided by the New York State Office of Parks and the United States National Park Service.

References

Agrawal, A. A., Conner, J. K., & Stinchcombe, J. R. (2004). Evolution of plant resistance and tolerance to frost damage. Ecology Letters, 7(12), 1199-1208. doi:10.1111/j.1461-0248.2004.00680.x

Ap Rees, T., Burrell, M. M., Entwistle, T. G., Hammond, J. B., Kirk, D., & Kruger, N. J. (1988). Effects of low temperature on the respiratory metabolism of carbohydrates by plants. Symposia of the Society for Experimental Biology, (42), 377–393. doi: 10.1007/BF00193009

Armstrong, J. J., Takebayashi, N., Sformo, T., & Wolf, D. E. (2015). Cold tolerance in *Arabidopsis kamchatica*. American Journal of Botany, 102(3), 439–448. doi: 10.3732/ajb.1400373

Baptist, F., & Aranjuelo, I. (2011). Interaction of carbon and nitrogen metabolisms in alpine plants. In C. Lütz (Ed.), *Plants in alpine regions* (pp. 121–134). Springer, Vienna, Austria.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48. doi: 10.18637/jss.v067.i01

Baxter, R. (2014). Plant acclimation and adaptation to cold environments. In K. Franklin & P. Wigge (Eds.), *Temperature and plant development* (pp. 19–48). John Wiley & Sons, Inc, Oxford, UK.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. Ecology Letters, 15(4), 365–377. doi: 10.1111/j.1461-0248.2011.01736.x

Billings, W. D. (1974). Arctic and alpine vegetation: plant adaptations to cold summer climates. In J. D. Ives and R. G. Barry (Eds.), *Arctic and alpine environments* (pp. 403–443). Methuen, London, UK.

Billings, W. D., & Mooney, H. A. (1968). The ecology of arctic and alpine plants. Biological Reviews, 43(4), 481–529. doi: 10.1111/j.1469-185X.1968.tb00968.x

Birks, H. H. (2008). The Late-Quaternary history of arctic and alpine plants. Plant Ecology and Diversity, 1(2), 135–146. doi: 10.1080/17550870802328652

Brochmann, C., Gabrielsen, T. M., Nordal, I., Landvik, J. Y., & Elven, R. (2003). Glacial survival or *tabula rasa*? The history of North Atlantic biota revisited. *Taxon* 52: 417–450. doi: 10.2307/3647444

Browse, J., & Xin, Z. G. (2001). Temperature sensing and cold acclimation. Current Opinions in Plant Biology, 4(3), 241–246. doi: 10.1016/S1369-5266(00)00167-9

Chen, I. C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. Science, 333(6045), 1204–1026. doi: 10.1126/science.1206432

Clo, J., Ronfort, J., & Abu Awad, D. (2020). Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations. Journal of Evolutionary Biology, in press. doi: 10.1111/jeb.13660

Colautti, R. I., & Barrett, S. C. H. (2013). Rapid adaptation to climate facilitates range expansion of an invasive plant. Science, 342(6156), 364–366. doi: 10.1126/science.1242121

Davies, P. J., & Gan, S. (2012). Towards an integrated view of monocarpic plant senescence. Russian Journal of Plant Physiology, 59(4), 467–478. doi: 10.1134/S102144371204005X

Debieu, M., Tang, C., Stich, B., Sikosek, T., Effgen, S., Josephs, E., … de Maux, J. (2013). Co-variation between seed dormancy, growth rate and flowering time changes with latitude in *Arabidopsis thaliana*. PLoS ONE, 8(5), e61075. doi: 10.1371/journal.pone.0061075

Eriksson, M., & Rafajlović, M. (2020). The effect of genetic architecture and selfing on the capacity of a population to expand its range. bioRxiv. doi: 10.1101/2020.01.22.915454

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

Francis, D., & Barlow, P. W. (1988). Temperature and the cell cycle. In S. P. Long & F. I. Woodward (Eds.), *Plants and temperature, 42nd Symposium of the Society for Experimental Biology* (pp. 181–202). Company of Biologists, Cambridge, UK.

Grant, A. G., & Kalisz, S. (2020). Do selfing species have greater niche breadth? Support from ecological niche modeling. Evolution, 74(1), 73–88. doi: 10.1111/evo.13870

Griffin, P. C., & Willi, Y. (2014). Evolutionary shifts to self-fertilisation restricted to geographic range margins in North American *Arabidopsis lyrata*. Ecology Letters, 17(4), 484–490. doi: 10.1111/ele.12248

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. Journal of Statistical Softwares, 33(2), 1–22. doi: 10.18637/jss.v033.i02

Hadfield, J. D. (2019). MCMCglmm course notes. Website https://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf [accessed 18 August 2020].

Hannah, M. A., Wiese, D., Freund, S., Fiehn, O., Heyer, A. G., & Hincha, D. K. (2006). Natural genetic variation of freezing tolerance in *Arabidopsis*. Plant Physiology, 142(1), 98–112. doi: 10.1104/pp.106.081141

Havranek, W. M., & Tranquillini, W. (1995). Physiological processes during winter dormancy and their ecological significance. In W. K. Smith & T. M. Hinckley (Eds.), *Ecophysiology of Coniferous Forests* (pp. 95–124). Academic Press, San Diego, California, USA.

Heidarvand, L., & Amiri, R. M. (2010). What happens in plant molecular responses to cold stress? Acta Physiologiae Plantarum, 32(3), 419–431. doi: 10.1007/s11738-009-0451-8

Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. Nature, 405(6789), 907–913. doi: 10.1038/35016000

Hewitt, G. M. (2004). Genetic consequences of climatic changes in the Quaternary. Philosophical Transactions of the Royal Society B, Biological Sciences, 359(1442), 183–195. doi: 10.1098/rstb.2003.1388

Hohmann, N., Schmickl, R., Chiang, T. Y., Lučanová, M., Kolář, F., Marhold, K., & Koch, M. A. (2014). Taming the wild: resolving the gene pools of non-model *Arabidopsis* lineages. BMC Evolutionary Biology, 14(1), 224. doi: 10.1186/s12862-014-0224-x

Hopkins, M. (1937). *Arabis* in eastern and central North America. Rhodora, 39(460), 63–186.

Horton, M. W., Willems, G., Sasaki, E., Koornneef, M., & Nordborg, M. (2016). The genetic architecture of freezing tolerance varies across the range of *Arabidopsis thaliana*. Plant, Cell & Environment, 39(11), 2570–2579. doi: 10.1111/pce.12812

Hughes, P. D., Gibbard, P. L., & Ehlers, J. (2013). Timing of glaciation during the last glacial cycle: Evaluating the concept of a global ‘last glacial maximum’ (LGM). Earth-Science Reviews, 125, 171–198. doi: 10.1016/j.earscirev.2013.07.003

Hurme, P., Repo, T., Savolainen, O., & Paakkonen, T. (1997). Climatic adaptation of bud set and frost hardiness in Scots pine (*Pinus sylvestris*). Canadian Journal of Forest Research, 27(5), 716–723. doi: 10.1139/cjfr-27-5-716

Inouye, D.W. (2008). Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. Ecology, 89(2), 353–362. doi: 10.1890/06-2128.1

Inouye, D. W., & Wielgolaski, F. E. (2003). High altitude climates. In M. D. Schwartz (Ed.), *Phenology: an integrative environmental science, Tasks for Vegetation Science* (vol. 39, pp. 195–214). Springer, Dordrecht, Netherlands.

Inouye, D. W., Saavedra, F., & Lee, W. (2003). Environmental influences on the phenology and abundance of flowering by *Androsace septentrionalis* (Primulaceae). American Journal of Botany, 90(6), 905–910. doi: 10.3732/ajb.90.6.905

Kenta, T., Yamada, A., & Onda, Y. (2011). Clinal variation in flowering time and vernalisation requirement across a 3000-m altitudinal range in perennial *Arabidopsis kamchatica* ssp. *kamchatica* and annual lowland subspecies *kawasakiana*. Journal of Ecosystem and Ecography, S6, 001. doi: 10.4172/2157-7625.S6-001

Klimešomá, J., & Klimeš, L. (2007). Bud banks and their role in vegetative regeneration – a literature review and proposal for simple classification and assessment. Perspectives in Plant Ecology, Evolution and Systematics, 8(3), 115–129. doi: 10.1016/j.ppees.2006.10.002

Knight, M. R., & Knight, H. (2012). Low-temperature perception leading to gene expression and cold tolerance in higher plants. New Phytologist, 195(4), 737–751. doi: 10.1111/j.1469-8137.2012.04239.x

Koch, M.A., Kiefer, C., Ehrich, D., Vogel, J., Brochmann, C., & Mummenhoff, K. (2006). Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). Molecular Ecology, 15(3), 825–839. doi: 10.1111/j.1365-294X.2005.02848.x

Koehler, K., Center, A., & Cavender-Bares, J. (2012). Evidence for a freezing tolerance–growth rate trade-off in the live oaks (*Quercus series Virentes*) across the tropical–temperate divide. New Phytologist, 193(3), 730–744. doi: 10.1111/j.1469-8137.2011.03992.x

Körner, C. (2012) Treelines will be understood once the functional difference between a tree and a shrub is. Ambio, 41(S3), 197–206. doi: 10.1007/s13280-012-0313-2

Körner, C. (2016). Plant adaptation to cold climates. F1000Research, 5(F1000 Faculty Rev), 2769. doi: 10.12688/f1000research.9107.1

Kosova, K., Holkova, L., Prasil, I. T., Prasilova, P., Bradacova, M., Vitamvas, P., & Capkova, V. (2008). Expression of dehydrin 5 during the development of frost tolerance in barley (*Hordeum vulgare*). Journal of Plant Physiology, 165(11), 1142–1151. doi: 10.1016/j.jplph.2007.10.009

Kronholm, I., Pico, F. X., Alonso-Blanco, C., Goudet, J., & de Meaux, J. (2012). Genetic basis of adaptation in *Arabidopsis thaliana*: local adaptation at the seed dormancy QTL DOG1. Evolution, 66(7), 2287–2302. doi: 10.1111/j.1558-5646.2012.01590.x

Kuittinen, H., Niittyvuopio, A., Rinne, P., & Savolainen, O. (2008). Natural variation in *Arabidopsis lyrata* vernalization requirement conferred by a FRIGIDA indel polymorphism. Molecular Biology and Evolution, 25(2), 319–329. doi: 10.1093/molbev/msm257

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed-effects models. Journal of Statistical Software, 82(13), 1–26. doi: 10.18637/jss.v082.i13

Lancaster, L. T., & Humphreys, A. M. (2020). Global variation in the thermal tolerances of plants. Proceedings of the National Academy of Sciences, 117(24), 13580–13587. doi: 10.1073/pnas.1918162117

Larcher, W., Kainmueller, C., & Wagner, J. (2010). Survival types of high mountain plants under extreme temperatures. Flora, 205(1), 3–18. doi: 10.1016/j.flora.2008.12.005

Lande, R., & Porcher, E. (2015). Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. Genetics, 200(3), 891–906. doi: 10.1534/genetics.115.176693

Laureano, R. G., Lazo, Y. O., Linares, J. C., Luque, A., Martínez, F., Seco, J. I., & Merino, J. (2008). The cost of stress resistance: construction and maintenance costs of leaves and roots in two populations of *Quercus ilex*. Tree Physiology, 28(11), 1721–1728. doi: 10.1093/treephys/28.11.1721

Lee-Yaw, J. A., Fracassetti, M., & Willi, Y. (2018). Environmental marginality and geographic range limits: a case study with *Arabidopsis lyrata* ssp. *lyrata*. Ecography, 41(4), 622–634. doi: 10.1111/ecog.02869

Lenoir, J., & Svenning, J. C. (2013). Latitudinal and elevational range shifts under contemporary climate change. In S. A. Levin (Ed.), *Encyclopedia of Biodiversity* (2nd ed., pp. 599–611). Academic Press. Encyclopedia of Biodiversity.

Lenth, R. 2019. emmeans: estimated marginal means, aka least-squares means. R package version 1.4. Website https://CRAN.R-project.org/package=emmeans [accessed 18 August 2020]

Loehle, C. (1998). Height growth rate tradeoffs determine northern and southern range limits for trees. Journal of Biogeography, 25(4), 735–742. doi: 10.1046/j.1365-2699.1998.2540735.x

Miroslavov, E. A., & Kravkina, I. M. (1991). Comparative analysis of chloroplasts and mitochondria in leaf chlorenchyma from mountain plants grown at different altitudes. Annals of Botany, 68(3), 195–200. doi: 10.1093/oxfordjournals.aob.a088244

Montesinos-Navarro, A., Wig, J., Pico, F. X., & Tonsor, S. J. (2010). *Arabidopsis thaliana* populations show clinal variation in a climatic gradient associated with altitude. New Phytologist, 189(1), 282–294. doi: 10.1111/j.1469-8137.2010.03479.x

Mulligan, G. A. (1996). Synopsis of the genus *Arabis* (Brassicaceae) in Canada, Alaska, and Greenland. Rhodora, 97(890), 109–163.

Neuner, G. (2014). Frost resistance in alpine woody plants. Frontiers in Plant Science, 5, 654. doi: 10.3389/fpls.2014.00654

Neuner, G., Erler, A., Ladinig, U., Hacker, J., & Wagner, J. (2013). Frost resistance of reproductive tissues during various stages of development in high mountain plants. Physiologia Plantarum, 147(1), 88–100. doi: 10.1111/j.1399-3054.2012.01616.x

Novikova, P. Y., Hohmann, N., Nizhynska, V., Tsuchimatsu, T., Ali, J., Muir, G., … Nordborg, M. (2016). Sequencing of the genus *Arabidopsis* identifies a complex history of non-bifurcating speciation and abundant trans-specific polymorphism. Nature Genetics, 48(9), 1077–1082. doi: 10.1038/ng.3617

Ometto, L., Li, M., Bresadola, L., & Varotto, C. (2012). Rates of evolution in stress-related genes are associated with habitat preference in two *Cardamine* lineages. BMC Evolutionary Biology, 12(1), 7. doi: 10.1186/1471-2148-12-7

Paccard, A., Fruleux, A., & Willi, Y. (2014). Latitudinal trait variation and responses to drought in *Arabidopsis lyrata*. Oecologia, 175(2), 577–587. doi: 10.1007/s00442-014-2932-8

Pagter, M., Kristoffersen, A., Bronnum, P., & Jensen, M. (2010). Phenotypic differences in development of cold hardiness in three latitudinal populations of *Acer platanoides* L. Scandinavian Journal of Forest Research, 25(5), 412–420. doi: 10.1080/02827581.2010.512872

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

Pearce, R. S. (2001) Plant freezing and damage. Annals of Botany, 87(4), 417–424. doi: 10.1006/anbo.2000.1352

Pearce, R. S., & Ashworth, E. N. (1992). Cell-shape and localization of ice in leaves of overwintering wheat during frost stress in the field. Planta, 188(3), 324–331. doi: 10.1007/bf00192798

Penning de Vries, F.W.T. (1975). The cost of maintenance processes in plant cells. Annals of Botany, 39(1), 7–92. doi: 10.1093/oxfordjournals.aob.a084919

Perrier, A., Sánchez-Castro, D., & Willi, Y. (2020). Expressed mutational load increases toward the edge of a species' geographic range. Evolution, 74(8), 1711–1723. doi: 10.1111/evo.14042

Polechová, J. (2018). Is the sky the limit? On the expansion threshold of a species range. PLoS Biology, 16(6), e2006735. doi: 10.1371/journal.pbio.2005372

Polechová, J., & Barton, N. H. (2015). Limits to adaptation along environmental gradients. Proceedings of the National Academy of Sciences, USA, 112(20), 6401–6406. doi: 10.1073/pnas.1421515112

Postma, F. M., & Ågren, J. (2015). Maternal environment affects the genetic basis of seed dormancy in *Arabidopsis thaliana*. Molecular Ecology, 24(4), 785-797. doi: 10.1111/mec.13061

Preston, J., & Sandve, S. (2013). Adaptation to seasonality and the winter freeze. Frontiers in Plant Science, 4, 167. doi: 10.3389/fpls.2013.00167

Prevéy, J., Vellend, M. Rüger, N. Hollister, R. D. Bjorkman, A. D. Myers-Smith, I. H. … Oberbauer, S. F. (2017). Greater temperature sensitivity of plant phenology at colder sites: implications for convergence across northern latitudes. Global Change Biology, 23(7), 2660–2671. doi: 10.1111/gcb.13619

R Core Team. (2019). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Riihimaki, M., & Savolainen, O. (2004). Environmental and genetic effects on flowering differences between northern and southern populations of *Arabidopsis lyrata* (Brassicaceae). American Journal of Botany, 91(7), 1036–1045. doi: 10.3732/ajb.91.7.1036

Ritz, C., Baty, F., Streibig, J. C., & Gerhard, D. (2015). Dose-response analysis using R. PLoS ONE, 10(12), e0146021. doi: 10.1371/journal.pone.0146021

Sakai, A., & Larcher, W. (1987). Frost survival of plants: Responses and adaptation to freezing stress. Springer-Verlag, Berlin, Germany.

Schluter, D. (2001). Ecology and the origin of species. Trends in Ecology and Evolution, 16(7), 372–380. doi: 10.1016/S0169-5347(01)02198-X

Schmickl, R., Jørgensen, M. H., Brysting, A. K., & Koch, M. A. (2010). The evolutionary history of the *Arabidopsis lyrata* complex: a hybrid in the amphi-Beringian area closes a large distribution gap and builds up a genetic barrier. BMC Evolutionary Biology, 10(1), 98. doi: 10.1186/1471-2148-10-98

Schmitt T. (2007). Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. Frontiers in Zoology, 4(1), 11. doi: 10.1186/1742-9994-4-11

Sebastian-Azcona, J., Hamann, A., Hacke, U. G., & Rweyongeza, D. (2019) Survival, growth and cold hardiness tradeoffs in white spruce populations: Implications for assisted migration. Forest Ecology and Management, 433, 544–552. doi: 10.1016/j.foreco.2018.10.046

Sierra-Almeida, A., & Cavieres, L. A. (2012). Summer freezing resistance of high-elevation plant species changes with ontogeny. Environmental and Experimental Botany, 80, 10–15. doi: 10.1016/j.envexpbot.2012.01.009

Sierra-Almeida, A., Cavieres, L. A., & Bravo, L. A. (2009). Freezing resistance varies within the growing season and with elevation in high-Andean species of central Chiles. New Phytologist, 182(2), 461–469. doi: 10.1111/j.1469-8137.2008.02756.x

Sklenář, P. (2017). Seasonal variation of freezing resistance mechanisms in north-temperate alpine plants. Alpine Botany, 127(1), 31–39. doi: 10.1007/s00035-016-0174-6

Skrede, I., P. Eidesen, B., Portela, R. P., & Brochmann, C. 2006. Refugia, differentiation and postglacial migration in arctic-alpine Eurasia, exemplified by the mountain avens (*Dryas octopetala* L.). Molecular Ecology, 15(7), 1827–1840. doi: 10.1111/j.1365-294X.2006.02908.x

Stinchcombe, J. R., Weinig, C., Ungerer, M., Olsen, K. M., Mays, C., Halldorsdottir, S. S., … Schmitt, J. (2004). A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene FRIGIDA. Proceedings of the National Academy of Sciences, USA, 101(13), 4712–4717. doi: 10.1073/pnas.0306401101

Sutherland W. J., Freckleton, R. P., Godfray, H. C. J., Beissinger, S. R., Benton, T., Cameron, D. D., … Wiegand, T. (2013). Identification of 100 fundamental ecological questions. Journal of Ecology, 101(1), 58–67. doi: 10.1111/1365-2745.12025

Taschler, D., & Neuner, G. (2004). Summer frost resistance and freezing patterns measured *in situ* in leaves of major alpine growth forms in relation to their upper distribution boundary. Plant, Cell & Environment, 27(6), 737–746. doi: 10.1111/j.1365-3040.2004.01176.x

Thomashow, M. F. (1999). Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. Annual Review of Plant Physiology and Plant Molecular Biology, 50, 571–599. doi: 10.1146/annurev.arplant.50.1.571

Wagmann, K., Hautekèete, N. C., Piquot, Y., Meunier, C., Schmitt, S. E., & Van Dijk, H. (2012). Seed dormancy distribution: explanatory ecological factors. Annals of Botany, 110(6), 1205–1219. doi: 10.1093/aob/mcs194

Walden, N., Lucek, K., & Willi, Y. (2020). Lineage-specific adaptation to climate involves flowering time in North American *Arabidopsis lyrata*. Molecular Ecology, 29(8), 1436–1451. doi: 10.1111/mec.15338

Warwick, S. I., Al-Shehbaz, A., & Sauder, A. (2006). Phylogenetic position of *Arabis arenicola* and generic limits of *Aphragmus* and *Eutrema* (Brassicaceae) based on sequences of nuclear ribosomal DNA. Canadian Journal of Botany, 84(2), 269–281. doi: 10.1139/b05-161

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

Wingler, A., Juvany, M., Cuthbert, C., & Munné-Bosch, S. (2015). Adaptation to altitude affects the senescence response to chilling in the perennial plant *Arabis alpina*. Journal of Experimental Botany, 66(1), 355–367. doi: 10.1093/jxb/eru426

Wos, G., & Willi, Y. (2015). Temperature-stress resistance and tolerance along a latitudinal cline in North American *Arabidopsis lyrata*. PLoS ONE, 10(6), e0131808. doi: 10.1371/journal.pone.0131808

Wright, I. J., Reich, P. B., Atkin, O. K., Lusk, C. H., Tjoelker, M. G., & Westoby, M. (2006). Irradiance, temperature and rainfall influence leaf dark respiration in woody plants: evidence from comparisons across 20 sites. New Phytologist, 169(2), 309–319. doi: 10.1111/j.1469-8137.2005.01590.x

Wright, S. (1931). Evolution in Mendelian populations. Genetics, 16(2), 97–159.

Zhen, Y., & Ungerer, M. C. (2008). Clinal variation in freezing tolerance among natural accessions of *Arabidopsis thaliana*. New Phytologist, 177(2), 419–427. doi: 10.1111/j.1469-8137.2007.02262.x

Zuther, E., Schulz, E., Childs, L. H., & Hincha, D. K. (2012). Clinal variation in the non-acclimated and cold-acclimated freezing tolerance of *Arabidopsis thaliana* accessions. Plant, Cell & Environment, 35(10), 1860–1878. doi: 10.1111/j.1365-3040.2012.02522.x

Zuther E., Lee, Y. P., Erban, A., Kopka, J., & Hincha, D. K. (2018). Natural variation in freezing tolerance and cold acclimation response in *Arabidopsis thaliana* and related species. Advances in Experimental Medicine and Biology, 1081, 81–98. doi: 10.1007/978-981-13-1244-1_5

Tables

Table 1: Description of performance estimates

Performance estimate	Level	Description
Tolerance to cold growth temperatures and frost		
Multiplicative performance	Pot	Germination * reproductive output
Germination	Seed	Binary success of germination, from day 0 to the end of the experiment
Reproductive output	Pot	Sum of all flower organs counted four weeks after the first flower opening
Resistance to frost		
Survival to frost I *	Seed	Binary success of survival from the beginning of *frost I* to 14 days after the end of *frost I*
Survival to frost II *	Seed	Binary success of survival from the beginning of *frost II* to 14 days after the end of *frost II*
Frost damage I *	Pot	Binary success of increase in damage during *frost I*
Frost damage II *	Pot	Binary success of increase in damage during *frost II*
LT_{50}	Pot	Temperature at which 50 % of electrolyte leakage occurs
Phenology and growth		
Time to germination	Seed	Time [days] from day 0 until germination
Time to flowering	Pot	Time [days] from the end of vernalization to the first flower opening
Growth rate	Pot	Initial growth rate from a 3-parameter logistic growth model
Size at flowering	Pot	Average size [mm] of the two longest leaves four weeks after the first flower opening
Senescence		
Senescence I *	Pot	Binary success of increase in damage from 14 days after *frost I* to the beginning *frost II*
Senescence II *	Pot	Binary success of increase in damage from 14 days after *frost II* to the end of the experiment

* Frost was applied for two weeks before (*I*) and for six weeks after (*II*) the simulated winter phase

Table 2A: Summary of multiple comparisons testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), growth temperature (cold compared to mild [0]) and their interaction on multiplicative performance

		Species								
		Mild		*Cold*		*Growth temperature (cold vs mild)*			*Species * growth temperature*	
		A vs Lyr N	*Lyr N vs C*	*A vs Lyr N*	*Lyr N vs C*	*A*	*Lyr N*	*Lyr C*	*A vs Lyr N*	*Lyr N vs C*
Process	*N*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*
Log-normal	372	-0.67	-0.39	0.05	-0.59	0.17	-0.55	-0.35	**0.72** ***	**-0.20** *** †
Logistic	372	-4.78	2.62	3.17	-2.94	-1.31	-9.26	-3.70	**7.84** **	**-5.46** *

Multiplicative performance estimates (log_{10}-transformed if >0) were assumed to follow Gaussian distributions with 0-inflation. Therefore, models assessed all fixed and random effects for their importance in both the Gaussian process (total reproductive output) and the logistic process (binary variable depicting germination combined with survival and the capacity to initiate flowering). The logistic part of the model predicts non-zeros in the distribution on the logit scale. Coefficients (*mean*) depict estimated pairwise difference in performance between species within growth temperature, growth temperature within species, and estimated linear contrast in magnitude of effect of growth temperature between species. Estimates of coefficients are modes of an MCMC sample from the posterior distribution of parameters obtained from a Tukey's test. Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table 2B: Summary of multiple comparisons testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), growth temperature (cold compared to mild [0]) and their interaction on individual performance estimates and phenology estimates

		Species									
		Mild		*Cold*		*Growth temperature (cold vs mild)*			*Species* growth temperature*		
		A vs Lyr N	*Lyr N vs C*	*A vs Lyr N*	*Lyr N vs C*	*A*	*Lyr N*	*Lyr C*	*A vs Lyr N*	*Lyr N vs C*	
Dependent variable	*N*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	
Cold tolerance											
Germination	714	0.34	0.77	2.70	-0.72	**-0.87 ***	**-3.24 *****	**-1.74 *****	**2.36 *****	**-1.49**	** †, ‡
Reproductive output	215	-0.63	-0.48	-0.06	-0.61	**0.19 ***	**-0.38 ****	**-0.26 ***	**0.57 *****	-0.13	†
Phenology, growth and senescence											
Time to germination	418	-0.16	-0.03	-0.05	0.17	**0.70 *****	**0.59 *****	**0.39 *****	0.11	**0.20 ***	†
Time to flowering	215	0.27	0.01	0.06	0.16	**0.23 *****	**0.44 *****	**0.28 *****	**-0.21 *****	**0.16 *****	†
Size at flowering	215	0.00	-0.06	0.14	-0.00	0.01	-0.13	-0.19	**0.14 *****	0.06	†
Senescence I	252	**3.27 ****	**3.13 *****	**3.37 *****	2.27	-2.16	-2.26	-1.41	0.10	-0.85	†, ‡
Senescence II	243	-1.32	-0.52	2.33	-19.36	1.25	-2.40	16.45	**3.65 ***	-18.85	†, ‡

Binary variables (‡) were analyzed by models predicting non-zeros on the logit scale. All other performance estimates were $\log_{10}$-transformed and assumed to follow Gaussian distributions. Each model was optimized with the *bobyqa* optimizer to improve convergence. Coefficients (*β*) depict estimated pairwise difference in performance between species within growth temperature, growth temperature within species, and estimated linear contrast in magnitude of effect of growth temperature between species, obtained from a Tukey's test. Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) *P*<0.1, * *P*<0.05, ** *P*<0.01, *** *P*<0.001. Results for random effects are not shown.

Table 3A: Summary of multiple comparisons testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations]), treatment (frost compared to control [0]) and their interaction on multiplicative performance, in the mild growth temperature

		Species								
		Control		*Frost*		*Treatment (frost vs control)*			*Species* treatment*	
		A vs Lyr N	*Lyr N vs C*	*A vs Lyr N*	*Lyr N vs C*	*A*	*Lyr N*	*Lyr C*	*A vs Lyr N*	*Lyr N vs C*
Process	*N*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*
Log-normal	372	-0.72	-0.41	-0.46	-0.38	0.24	-0.02	-0.05	**0.25 ****	0.03
Logistic	372	-2.91	1.44	-1.18	-0.89	-0.57	-2.31	-0.03	1.73 (*)	**-2.35 ***

Multiplicative performance estimates (log_{10}-transformed if >0) were assumed to follow Gaussian distributions with 0-inflation. Therefore, models assessed all fixed and random effects for their importance in both the Gaussian process (total reproductive output) and the logistic process (binary variable depicting germination combined with survival and the capacity to initiate flowering). The logistic part of the model predicts non-zeros in the distribution on the logit scale. Coefficients (*mean*) depict estimated pairwise difference in performance between species within treatment, treatment within species, and estimated linear contrast in magnitude of effect of treatment between species. Estimates of coefficients are modes of an MCMC sample from the posterior distribution of parameters obtained from a Tukey's test. Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table 3B: Summary of multiple comparisons testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations]), treatment (frost compared to control [0]) and their interaction on individual performance estimates and phenology estimates, in the mild growth temperature

		Species									
		Control		*Frost*		*Treatment (frost vs control)*			*Species* treatment*		
		A vs Lyr N	*Lyr N vs C*	*A vs Lyr N*	*Lyr N vs C*	*A*	*Lyr N*	*Lyr C*	*A vs Lyr N*	*Lyr N vs C*	
Dependent variable	*N*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	
Frost tolerance											
Reproductive output	264	-0.69	-0.48	-0.45	-0.44	**0.20 ****	-0.04	-0.08	**0.24 ****	0.04	†
Frost resistance											
Survival to frost I	348	17.13	0.20	0.63	-0.44	-19.40	-2.90	-2.26	-16.50	0.64	‡
Survival to frost II	315	-0.00	-0.08	-0.75	-20.40	-21.16	-20.46	-0.10	-0.74	-20.31	†, ‡
Frost damage I	346	0.10	-0.04	-0.64	0.84	-21.20	-21.90	-21.00	-0.75	0.88	‡
Frost damage II	315	-0.18	**-2.07 ****	0.10	-1.41	**1.35 ****	**1.07 ***	0.43	0.28	0.66	‡
LT_{50}	128	0.06	-0.15	-	-	-	-	-	-	-	†
Phenology, growth and senescence											
Time to flowering	264	0.28	0.00	0.31	0.04	**0.17 *****	**0.14 *****	**0.10 ****	0.03	0.04	†
Growth rate	285	-0.04	-0.00	-0.07	0.05	0.05	0.09	0.04	-0.03	0.05	†
Size at flowering	264	0.01	-0.06	0.01	-0.01	**-0.05 ***	**-0.06 ****	**-0.11 *****	0.00	0.05	†
Senescence I	346	**3.14 ***	**3.06 *****	**1.17 *****	0.92	**-3.23 ****	**-1.26 ****	0.89	-1.97	**-2.14 ****	†, ‡
Senescence II	307	-1.37	-0.49	0.64	-0.70	-0.11	-2.12	-1.91	2.01	-0.21	†, ‡

Binary variables (‡) were analyzed by models predicting non-zeros on the logit scale. All other performance estimates were log_{10}-transformed (except LT_{50}) and assumed to follow Gaussian distributions. Each model was optimized with the *bobyqa* optimizer to improve convergence. Coefficients (*β*) depict estimated pairwise difference in performance between species within treatment, treatment within species, and estimated linear contrast of effect of treatment between species, obtained from a Tukey's test. Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) *P*<0.1, * *P*<0.05, ** *P*<0.01, *** *P*<0.001. Results for random effects are not shown.

Figures

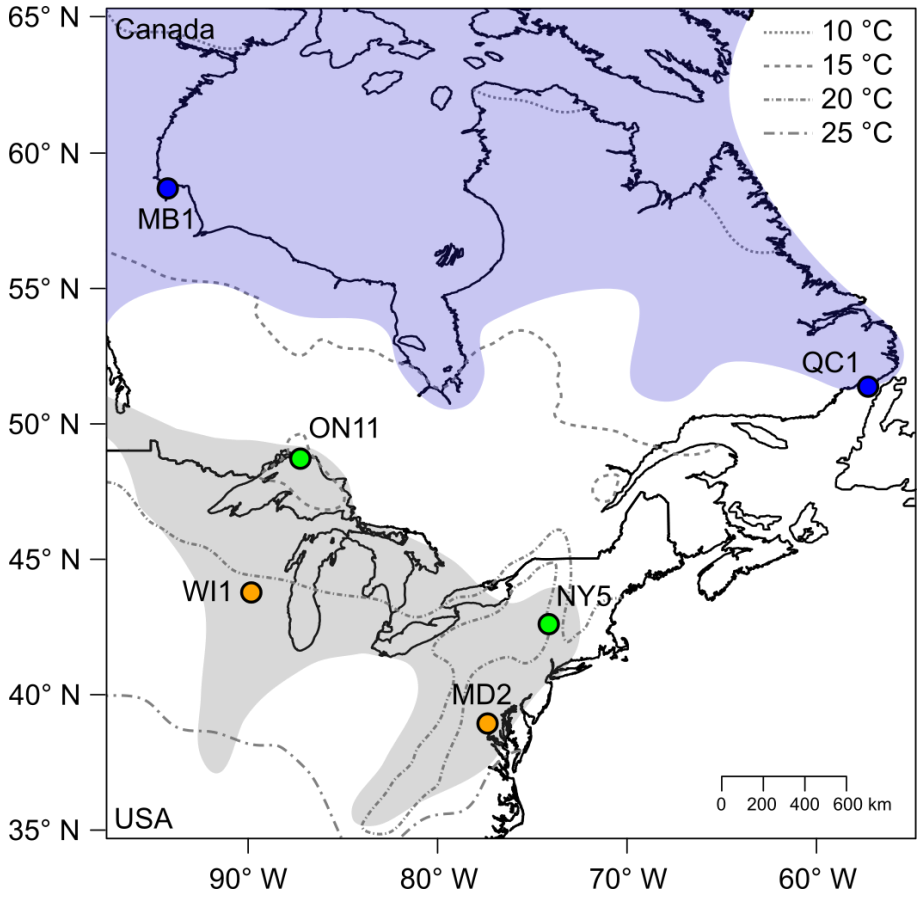

Figure 1: Distribution map of *Arabidopsis arenicola* and *A. lyrata* with the 6 populations studied. The blue shaded area represents the current range of *A. arenicola*, and the grey shaded area represents the North-American range of *A. lyrata*. Circles filled in blue, green and orange represent respectively populations of *A. arenicola* and northern and central populations of *A. lyrata*. Population labels consist of the abbreviation for state (USA) or province (Canada) and a number. Grey dashed lines represent isoclines of mean summer temperature (July and August, WorldClim database version 2.0, Fick & Hijmans, 2017).

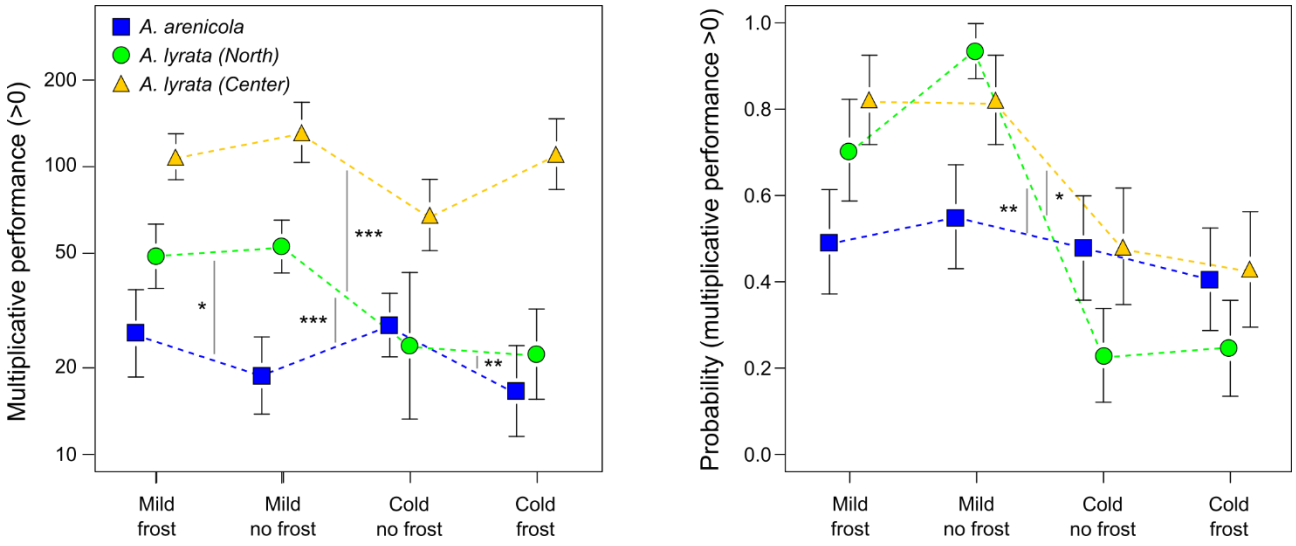

Figure 2: Variation in multiplicative performance of *A. arenicola* and *A. lyrata* under different temperature and frost regimes. Individuals of both species were grown under two growth temperatures: mild (20 °C) and cold (12 °C – 14 °C), both subdivided into two treatments: no frost (control) and frost. Symbols depict multiplicative performance estimates averaged for each species level: *A. arenicola* populations (blue squares), northern *A. lyrata* populations (green circles) and central *A. lyrata* populations (orange triangles). Multiplicative performance was analyzed on the level of the individual modeling two processes: a Gaussian process (values > 0, **left**) and a logistic process (modelling the probability of values > 0, **right**). Tukey's tests were performed to test for differences between species (not reported here), growth temperature or treatment, and the interaction between species and growth temperature or treatment. Horizontal bars represent the 95% confidence interval of the mean of species within each growth temperature and treatments. The significance of the effect of growth temperature or treatment within species is represented by full ($P < 0.05$) or dashed ($P > 0.05$) colored lines connecting each species. Significant differences between slopes of effect of growth temperature or treatments between *A. arenicola* and both northern populations of *A. lyrata*, and between northern and central populations of *A. lyrata*, are indicated by vertical grey lines, with significance indicated as: $*$ $P < 0.05$, $**$ $P < 0.01$, $***$ $P < 0.001$ Test statistics are reported in Table 2A, Table 3A and Table S11B.

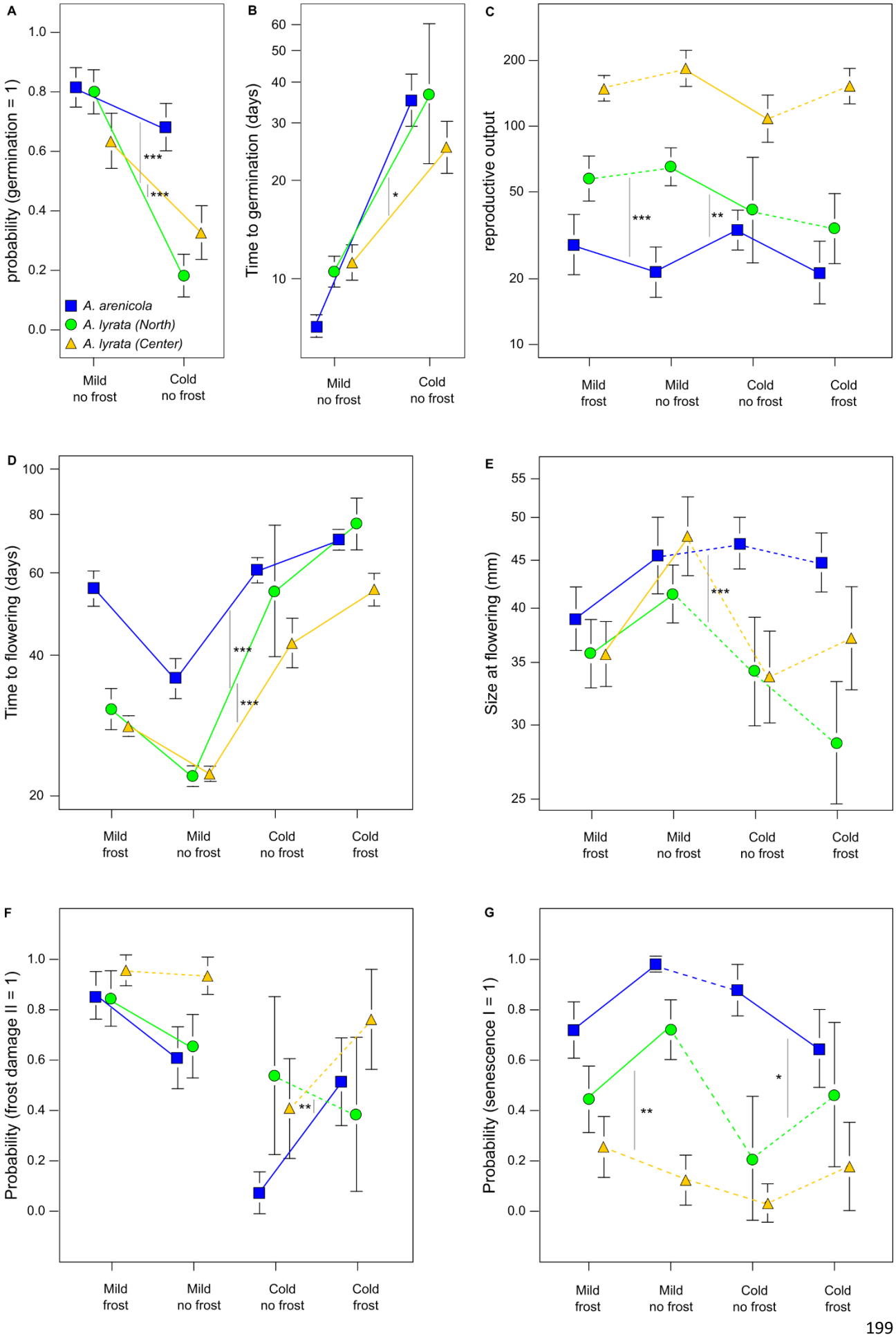
A
probability (germination = 1)
1.0
0.8
0.6
0.4
0.2
0.0
***
***
A. arenicola
A. lyrata (North)
A. lyrata (Center)
Mild no frost
Cold no frost
B
Time to germination (days)
60
50
40
30
20
10
*
Mild no frost
Cold no frost
C
reproductive output
200
100
50
20
10
***
**
Mild frost
Mild no frost
Cold no frost
Cold frost
D
Time to flowering (days)
100
80
60
40
20
***
***
Mild frost
Mild no frost
Cold no frost
Cold frost
E
Size at flowering (mm)
55
50
45
40
35
30
25
***
Mild frost
Mild no frost
Cold no frost
Cold frost
F
Probability (frost damage II = 1)
1.0
0.8
0.6
0.4
0.2
0.0
**
Mild frost
Mild no frost
Cold no frost
Cold frost
G
Probability (senescence I = 1)
1.0
0.8
0.6
0.4
0.2
0.0
**
*
Mild frost
Mild no frost
Cold no frost
Cold frost

Figure 3: Variation in individual performance estimates and phenology of *A. arenicola* and *A. lyrata* under different temperature and frost regimes. Individuals of both species were grown under two growth temperatures: mild (20 °C) and cold (12 °C – 14 °C), both subdivided into two treatments: no frost (control) and frost. Symbols depict individual performance or phenology estimates averaged for each species level: *A. arenicola* populations (blue squares), northern *A. lyrata* populations (green circles) and central *A. lyrata* populations (orange triangles). For each of the estimates of probability of germination (**A**), time to germination (**B**), reproductive output (**C**), time to flowering (**D**), size at flowering (**E**), frost damage II recorded in the simulated spring phase (**F**) and senescence I recorded during the simulated winter phase (**G**), Tukey's tests were performed to test for differences between species (not reported here), growth temperature or treatment, and the interaction between species and growth temperature or treatment. Horizontal bars represent the 95% confidence interval of the mean of species within each growth temperature and treatments. The significance of the effect of growth temperature or treatment within species is represented by full ($P < 0.05$) or dashed ($P > 0.05$) colored lines connecting each species. Significant differences between slopes of effect of growth temperature or treatments between *A. arenicola* and both northern populations of *A. lyrata*, and between northern and central populations of *A. lyrata*, are indicated by vertical grey lines, with significance indicated as: $*$ $P < 0.05$, $**$ $P < 0.01$, $***$ $P < 0.001$ Test statistics are reported in Table 2B, Table 3B and Table S13A.

Supplementary information

Table S1: Population estimates

Population code	Species	Species label	Latitude [° N]	Longitude [° W]	Mean temperature late summer [° C] †	Cluster	Mating system
QC1	*A. arenicola*	Are	51.43	-57.16	13.45	-	selfing
MB1	*A. arenicola*	Are	58.78	-94.20	12.00	-	selfing
NY5	*A. lyrata*	Lyr N	42.66	-74.02	20.20	E	outcrossing
ON11	*A. lyrata*	Lyr N	48.77	-87.13	14.65	W	selfing
MD2	*A. lyrata*	Lyr C	38.99	-77.25	24.55	E	outcrossing
WI1	*A. lyrata*	Lyr C	43.83	-89.72	20.75	W	outcrossing

† July and August, WorldClim database version 2.0, Fick & Hijmans, 2017

Table S2: Growth conditions of the crossing experiment

Growth Phase	Duration [days]	Temperature daytime [°C]	Temperature nighttime [°C]	Day length [h]	Light intensity [µmol m^{-2} s^{-1}]
Stratification	12	4	4	0	0
Germination	22	20	20	8	100
Growth †	21	22	20	10	140
Flowering initiation	10	22	20	16	240
Flowering and crossing	205	22	20	16	240

† Day length and light intensity were gradually increased every three days by 1h and 20 µmol m-2 s-1, respectively.

Table S3: Cross combinations generated in the crossing experiment and used in the climate chambers

Population code	No. of cross families	No. of seeds sown			
		Mild		Cold	
		Control	Frost	Control	Frost
QC1	11	63	63	63	63
MB1	12	72	72	72	72
NY5	12	64	64	64	64
ON11	10	51	51	51	51
MD2	10	52	52	52	52
WI1	10	55	55	55	55

Table S4: Growth conditions and performance tracking of the climate chamber experiment

Simulated season	Growth phase	Duration [days]	Day length [h] *	Light intensity [µmol m^{-2} s^{-1}] *	Mild temperatures [°C]		Cold temperatures [°C]		Weekly record rate		
					day	night	day	night	Germination and survival	Damage	Bolting and Flowering
Fall	Stratification	20	0	0	4	4	4	4	-	-	-
	Germination	14	8	120	20	18	12	10	5	-	-
	Growth †	26	10	140	20	18	12	10	5	5	-
Winter	Acclimation ‡	7	10	180	20	18	12	10	3	5	-
	Vernalization	46	8	140	4	4	4	4	1	1	-
	Acclimation ‡	4	8	140	4	4	4	4	1	1	-
Spring &	Growth I §	18	10	160	20	18	14	12	2	1	4
Summer	Growth II §	58	16	220	20	18	14	12	1	1	3
	Flowering ȿ	46	16	220	20	18	16	14	1	1	1

* Values at the beginning of each phase

† Light intensity was gradually increased every six days by 20 µmol m-2 s-1

‡ Gradual decrease / increase of day length, light intensity, and temperatures

§ Light intensity was gradually increased every three days by 20 µmol m-2 s-1

ȿ The whole experiment was first performed in climate chambers (Climecab 1400, Kälte 3000 AG, Landquart, Switzerland), and was transferd for the last 46 days into growth chambers (MobyLux GroBanks, CLF Plant Climatics, plantclimatics.de, Wertingen, Deutschland) for logistical reasons. Temperatures of the Cold growth temperature were increased due to technical limitations of this growth system

Appendix S5: Parametrization of priors and hierarchical mixed-effects model analyzis in a Bayesian (MCMC) framework, with individual multiplicative performance as dependent variable

Priors

Priors were set to be weak, convergence was improved using parameter expansion. *R* is the priors for the fixed effects, *G* is the priors for the random effects.

```
priors.model=list(
 R=list(V=diag(2), n=1, fix = 2),
 G=list(G1=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2),
    G2=list(V=diag(2), n=4, alpha.mu = rep(0,4),alpha.V = diag(2)*25^2),
    G3=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2)))
```

Hierarchical mixed-effects models analyzed in a Bayesian (MCMC) framework

Multiplicative performance was split into two datasets: the *zero_part*, a binary transformation of performance estimates with *zero_part* = 1 if performance > 0, else *zero_part* = 0; and the *norm_part* containing only the $\log_{10}$ transformed performance measures if *zero_part* = 1.

```
model = MCMCglmm(cbind(norm_part, zero_part) ~ trait -1 + trait:species * trait: growth
temperature,
        random = ~ us(trait):maternal population
        + us(trait): maternal population: maternal family
        + us(trait):block,
        prior = priors.model,
        rcov = ~idh(trait):units,
        family=c('gaussian', 'categorical'),
        burnin = 5000, thin = 100, nitt = 50000,
        data=data)
```

Appendix S6: Parametrization of restricted maximum likelihood based hierarchical mixed-effect models

Binary dependent variable

Model = glmer(performance ~ species * growth temperature
+ (1 | maternal population / maternal family)
+ (1 | block),
family = “binomial”,
control = glmerControl(optimizer = "bobyqa"),
data = data)

Dependent variable with log-normal distribution

Model = lmer($\log_{10}$(performance) ~ species * growth temperature
+ (1 | maternal population / maternal family)
+ (1 | block),
control = lmerControl(optimizer = "bobyqa"),
data = data)

Table S7: Summary of model comparison on models testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), growth temperature (cold compared to mild [0]) and their interaction on multiplicative performance

Model	Tested fixed effect	*DIC*	Δi
Species + temperature + (species* temperature)	-	342.46	-
Temperature + (species* temperature)	Species	342.23	0.23
Species + (species* temperature)	Temperature	343.07	-0.67
Species + temperature	Species * temperature	413.25	-70.79

The dependent variable was *multiplicative performance*. To test the contribution of each fixed effect to the fit of the full model, alternative models excluding each fixed effects were compared to the full model, with lower *DIC* values indicating a better fit. The difference between the full model and the others is indicated as Δi.

Table S8: Summary of performance in each treatment of each growth temperature

	Mild				Cold			
	Control		Frost		Control		Frost	
Dependent variable	*N*	*Mean*	*N*	*Mean*	*N*	*Mean*	*N*	*Mean*
Tolerance to cold growth temperatures and frost events								
Multiplicative performance	186	71.94	186	60.65	186	23.46	186	25.56
Germination †, ‡	357	74.06	357	80.79	357	42.41	357	40.34
Reproductive output	141	102.72	123	90.02	74	58.83	67	79.38
Resistance to frost								
Survival to frost I †, ‡	170	97.31	178	84.22	148	95.43	142	87.39
Survival to frost II ‡	168	100.00	147	96.15	102	95.81	89	87.95
Frost damage I ‡	169	0.00	177	71.77	83	0.00	77	31.29
Frost damage II ‡	168	73.88	147	89.18	81	28.88	69	44.80
LT_{50} [°C]	128	-10.84	-	-	-	-	-	-
Phenology, growth and senescence								
Time to germination †	270	11.05	286	11.39	148	48.45	142	45.69
Time to flowering	141	26.5	123	37.91	74	55.80	67	67.40
growth rate	143	0.04	142	0.06	18	0.02	15	0.10
Size at flowering [mm]	141	43.48	123	37.79	74	37.29	67	37.31
Senescence I ‡	169	62.49	177	48.88	83	34.18	77	41.49
Senescence II ‡	167	92.70	140	84.51	76	91.77	63	89.07

† Assessed at the seedling level

‡ Binary variable, expressed in %

Table S9: Summary of ANOVA on models testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), growth temperature (cold compared to mild [0]) and their interaction on individual performance estimates and phenology estimates

		Species		*Growth temperature*		*Species * growth temperature*				
Dependent variable	*N*	χ^2		χ^2		χ^2		R^2m	R^2c	
Cold tolerance										
Germination	714	1.45		**6.25**	*	**23.90**	***	0.28	0.46	†, ‡
Reproductive output	215	**14.19**	***	**5.21**	*	**40.26**	***	0.51	0.80	†
Phenology, growth and senescence										
Time to germination	418	1.98		**401.05**	***	**24.51**	***	0.49	0.64	†
Time to flowering	215	**9.39**	**	**98.39**	***	**34.89**	***	0.67	0.87	†
Size at flowering	215	0.35		0.24		**38.70**	***	0.15	0.64	†
Senescence I	252	**33.22**	***	3.70	(*)	0.41		0.66	0.70	†, ‡
Senescence II	243	2.25		2.17		**7.10**	*	0.88	0.90	†, ‡

Binary variables (‡) were analyzed by models predicting non-zeros on the logit scale. All other performance estimates were $\log_{10}$-transformed and assumed to follow Gaussian distributions. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include the χ^2-value, and the marginal and conditional R^2 of the model. Model fits with significant (positive) intercept are indicated by †. χ^2-value with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table S10: Summary of model comparison on models testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), treatment (frost compared to control [0]) and their interaction on multiplicative performance, in the mild growth temperature

Model	Tested fixed effect	*DIC*	Δi
Species + treatment + (species*treatment)	-	417.16	-
Treatment + (species*treatment)	Species	417.32	-0.16
Species + (species*treatment)	Treatment	417.16	0.00
Species + treatment	Species * treatment	428.32	-11.16

The dependent variable was *multiplicative performance*. To test the contribution of each fixed effect to the fit of the full model, alternative models excluding each fixed effects were compared to the full model, with lower *DIC* values indicating a better fit. The difference between the full model and the others is indicated as Δi.

Table S11A: Summary of model comparison on models testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), treatment (frost compared to control [0]) and their interaction on multiplicative performance, in the cold growth temperature

Model	Tested fixed effect	*DIC*	Δi
Species + treatment + (species*treatment)	-	408.93	-
Treatment + (species*treatment)	Species	408.61	0.32
Species + (species*treatment)	Treatment	409.41	-0.48
Species + treatment	Species * treatment	418.31	-9.38

The dependent variable was *multiplicative performance*. To test the contribution of each fixed effect to the fit of the full model, alternative models excluding each fixed effects were compared to the full model, with lower *DIC* values indicating a better fit. The difference between the full model and the others is indicated as Δi.

Table S11B: Summary of multiple comparisons testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), treatment (frost compared to control [0]) and their interaction on multiplicative performance, in the cold growth temperature

		Species									
		Control		*Frost*		*Treatment (frost vs control)*			*Species* treatment*		
		A vs Lyr N	*Lyr N vs C*	*A vs Lyr N*	*Lyr N vs C*	*A*	*Lyr N*	*Lyr C*	*A vs Lyr N*	*Lyr N vs C*	
Process	*N*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	
Log-normal	372	-0.01	-0.51	0.34	-0.93	0.01	-0.25	0.17	**-0.35 ***	-0.07	†
Logistic	372	-1.77	0.01	-0.79	-0.21	0.17	-0.81	-0.58	-0.98	0.75	

Multiplicative performance estimates (log_{10}-transformed if >0) were assumed to follow Gaussian distributions with 0-inflation. Therefore, models assessed all fixed and random effects for their importance in both the Gaussian process (total reproductive output) and the logistic process (binary variable depicting germination combined with survival and the capacity to initiate flowering). The logistic part of the model predicts non-zeros in the distribution on the logit scale. Coefficients (means) depict estimated pairwise difference in performance between species within treatment, treatment within species, and estimated linear contrast in magnitude of effect of treatment between species. Estimates of coefficients are modes of an MCMC sample from the posterior distribution of parameters obtained from a Tukey's test. Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table S12: Summary of ANOVA on models testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), treatment (frost compared to control [0]) and their interaction on individual performance estimates and phenology estimates, in the mild growth temperature

		Species		*Treatment*		*Species * treatment*				
Dependent variable	*N*	χ^2		χ^2		χ^2		R^2m	R^2c	
Frost tolerance										
Reproductive output	264	**12.19**	**	**8.42**	**	**14.37**	***	0.50	0.82	†
Frost resistance										
Survival to frost I	348	0.01		0.00		0.14		0.93	0.94	‡
Survival to frost II	315	0.00		0.00		0.00		0.97	0.97	†, ‡
Frost damage I	346	0.00		0.00		0.00		0.97	0.97	‡
Frost damage II	315	**11.50**	**	**7.35**	**	0.83		0.22	0.24	‡
LT_{50}	128	0.05		-		-		0.00	0.53	†
Phenology, growth and senescence										
Time to flowering	264	**8.91**	*	**52.46**	***	**9.47**	**	0.67	0.87	†
Growth rate	285	0.64		0.97		0.52		0.03	0.09	†
Size at flowering	264	0.40		**5.63**	*	4.20		0.08	0.57	†
Senescence I	346	**37.74**	***	**9.21**	**	**16.03**	***	0.53	0.56	†, ‡
Senescence II	307	4.99	(*)	0.04		5.17	(*)	0.19	0.19	†, ‡

Binary variables (‡) were analyzed by models predicting non-zeros on the logit scale. All other performance estimates were $\log_{10}$-transformed (except LT_{50}) and assumed to follow Gaussian distributions. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include the χ^2-value, and the marginal and conditional R^2 of the model. Model fits with significant (positive) intercept are indicated by †. χ^2-value with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table S13A: Summary of multiple comparisons testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), treatment within species (frost compared to control [0]) and their interaction on individual performance estimates and phenology estimates, in the cold growth temperature

		Species									
		Control		*Frost*		*Treatment (frost vs control)*			*Species* treatment*		
		A vs Lyr N	*Lyr N vs C*	*A vs Lyr N*	*Lyr N vs C*	*A*	*Lyr N*	*Lyr C*	*A vs Lyr N*	*Lyr N vs C*	
Dependent variable	*N*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	
Frost tolerance											
Reproductive output	141	-0.10	-0.51	-0.33	-0.63	**-0.21 ***	0.01	0.13	-0.22	-0.12	†
Frost resistance											
Survival to frost I	290	-0.14	-0.53	-0.40	-1.67	**-1.26 ***	-1.00	0.14	-0.26	-1.14	†, ‡
Survival to frost II	191	1.78	-19.30	0.74	-1.34	-1.01	0.03	-17.64	-1.04	17.67	†, ‡
Frost damage I	160	-0.04	0.13	0.08	0.11	20.00	19.90	19.90	0.12	-0.02	‡
Frost damage II	150	**-2.70 ** **	0.53	0.56	-1.64 (*)	**2.60 *** **	-0.63	1.54	**3.23 ** **	-2.17 (*)	†, ‡
Phenology, growth and senescence											
Time to flowering	141	0.03	0.11	-0.03	0.15	**0.08 ** **	**0.14 *** ***	**0.11 *** ***	-0.07	0.03	†
Size at flowering	141	0.12	0.05	0.17	-0.05	-0.03	-0.08	0.02	0.05	-0.10	†
Senescence I	160	**3.31 *** ***	1.96	0.75	1.38	**-1.38 ***	1.18	1.76	**-2.56 ***	-0.58	†, ‡
Senescence II	139	1.66	-19.34	-0.52	-0.79	-1.18	1.00	-17.54	-2.18	18.54	†, ‡

Binary variables (‡) were analyzed by models predicting non-zeros on the logit scale. All other performance estimates were $\log_{10}$-transformed and assumed to follow Gaussian distributions. Each model was optimized with the *bobyqa* optimizer to improve convergence. Coefficients (*β*) depict estimated pairwise difference in performance between species within treatment, treatment within species, and estimated linear contrast in magnitude of effect of treatment between species, obtained from a Tukey's test. Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table S13B: Summary of ANOVA on models testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), treatment within species (frost compared to control [0]) and their interaction on individual performance estimates and phenology estimates, in the cold growth temperature

		Species		*Treatment*		*Species * treatment*				
Dependent variable	*N*	χ^2		χ^2		χ^2		R^2m	R^2c	
Frost tolerance										
Reproductive output	141	**9.02**	*	**6.23**	*	**11.06**	**	0.51	0.72	†
Frost resistance										
Survival to frost I	290	0.36		**5.24**	*	0.83		0.15	0.15	†, ‡
Survival to frost II	191	2.48		1.21		0.58		0.90	0.93	†, ‡
Frost damage I	160	0.00		0.00		0.00		0.97	0.97	‡
Frost damage II	150	**12.43**	**	**13.93**	***	**9.41**	**	0.33	0.33	†, ‡
Phenology, growth and senescence										
Time to flowering	141	**12.0**	**	**9.43**	*	2.78		0.32	0.67	†
Size at flowering	141	**6.39**	*	0.69		4.39		0.30	0.54	†
Senescence I	160	**30.59**	***	**5.53**	*	**9.57**	**	0.51	0.51	†, ‡
Senescence II	139	2.81		1.79		1.97		0.94	0.94	†, ‡

Binary variables (‡) were analyzed by models predicting non-zeros on the logit scale. All other performance estimates were $\log_{10}$-transformed and assumed to follow Gaussian distributions. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include the χ^2-value, and the marginal and conditional R^2 of the model. Model fits with significant (positive) intercept are indicated by †. χ^2-value with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Table S14A: Summary of multiple comparisons testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), growth temperature within species (cold compared to mild [0]) and their interaction on individual performance estimates under frost treatment

		Species									
		Mild		*Cold*		*Growth temperature (cold vs mild)*			*Species * Growth temperature*		
		A vs Lyr N	*Lyr N vs C*	*A vs Lyr N*	*Lyr N vs C*	*A*	*Lyr N*	*Lyr C*	*A vs Lyr N*	*Lyr N vs C*	
Dependent variable	*N*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	
Survival to frost I	320	0.58	-0.40	-0.48	-1.63	-0.26	0.80	2.03	-1.06	-1.23	†
Survival to frost II	236	-0.75	-18.38	0.69	-1.65	-0.02	-1.46	-18.20	1.44	16.74	†
Frost damage I	254	-0.64	0.84	0.03	0.16	**-1.20 ****	**-1.86 ****	**-1.18 ***	0.66	-0.68	†
Frost damage II	216	0.10	-1.40	0.55	-1.65	**-1.74 ****	**-2.18 ****	**-1.94 ***	0.44	-0.25	†

All performance estimates were binary variables, and were analyzed by models predicting non-zeros on the logit scale. Each model was optimized with the *bobyqa* optimizer to improve convergence. Coefficients (*β*) depict estimated pairwise difference in performance between species within growth temperature, growth temperature within species, and estimated linear contrast in magnitude of effect of growth temperature between species, obtained from a Tukey's test. Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) *P*<0.1, * *P*<0.05, ** *P*<0.01, *** *P*<0.001. Results for random effects are not shown.

Table S14B: Summary of ANOVA on model testing for the effect of species (*A. arenicola* [A] compared to northern [Lyr N] populations, and northern compared to central [Lyr C] *A. lyrata* populations), growth temperature within species (cold compared to mild [0]) and their interaction on individual performance estimates under frost treatment

		Species	*Growth temperature*	*Species * growth temperature*			
Dependent variable	*N*	χ^2	χ^2	χ^2	R^2m	R^2c	
Survival to frost I	320	1.49	0.18	4.49	0.13	0.19	†
Survival to frost II	236	0.75	0.00	1.46	0.94	0.94	†
Frost damage I	254	3.62	**7.45** **	0.89	0.13	0.14	†
Frost damage II	216	3.03	**10.33** **	0.25	0.26	0.26	†

All performance estimates were binary variables, and were analyzed by models predicting non-zeros on the logit scale. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include the χ^2-value, and the marginal and conditional R^2 of the model. Model fits with significant (positive) intercept are indicated by †. χ^2-value with *P*-values < 0.05 are written in bold; significance is indicated: (*) *P*<0.1, * *P*<0.05, ** *P*<0.01, *** *P*<0.001. Results for random effects are not shown.

Chapter 5: Intrinsic and extrinsic postmating barriers contribute to reproductive isolation between two recently diverged *Arabidopsis* species

Antoine Perrier*, Yvonne Willi

Department of Environmental Sciences, University of Basel, Schönbeinstrasse 6, 4056 Basel, Switzerland.

*Corresponding author:

E-mail: antoine.perrier@unibas.ch

This manuscript is in preparation for submission in *American Journal of Botany*.

Abstract:

Recent speciation events in temperate and arctic species have mostly been linked to isolation in separate glacial refugia during Quaternary glaciation cycles. Rapid speciation during post-glacial range expansion has however received less attention. Here, we assessed whether postmating barriers contribute to reproductive isolation between the North American *Arabidopsis* and its northern selfing sister species *A. arenicola.* Despite sharing a common post-glacial origin, both species show different adaptive strategies to cold climates in line with their parapatric distribution, resulting in premating isolation due to ecological specialization and geographic distance. We assessed the strength of successive intrinsic reproductive barriers by performing within- and between-species crosses on populations of both species, and tracking the performance of their offspring raised in climate chambers. We also tested for selection against hybrids by tracking offspring performance under experimentally manipulated temperature in the climate chambers. All population pairs showed strong intrinsic reproductive isolation, with partial support for higher isolation between the most genetically distinct populations. Only early acting barriers were stronger between the most related populations, hinting toward reinforcing selection. Hybrids also showed lower adaptation on key traits necessary to survive in both parental ranges, suggesting selection against hybrids. We conclude that despite their recent divergence, both species have rapidly accumulated genetic incompatibilities potentially linked to the selfing mating system of *A. arenicola* and the strong drift in *A. lyrata.* Furthermore, strong reproductive isolation in early stages of divergence could have limited maladaptive geneflow from *A. lyrta* to *A. arenicola*, allowing it to colonize higher latitudes.

Keywords: Adaptation, *Arabidopsis arenicola*, *Arabidopsis lyrata*, extrinsic postmating barriers, intrinsic postmating barriers, reproductive isolation, speciation.

Introduction

Contemporary distribution and differentiation patterns of temperate and arctic species' have been strongly impacted by Pleistocene glacial fluctuations (Schluter, 2001; Hewitt, 2000, 2004), especially the last glacial cycle (c. 115 000–11 700 years ago) and last glacial maximum (LGM, c. 23 000–19 000 years ago; Hughes et al., 2013). After LGM, most northern hemisphere plant species migrated northwards following the retreat of the ice sheets (Brochmann et al., 2003; Schmitt, 2007), benefiting from pre-adaptation (Birks, 2008). Several taxa have also been able to expand their range from now temperate climates close to former glacial refugia, up to arctic climates (*e.g.* Skrede et al., 2006; Koch et al., 2006; Smickl et al., 2010). Recent divergence between post-glacial lineages has mostly been linked to the isolation in distinct glacial refugia (Hewitt, 2000, 2004; *e.g.* Escudero et al., 2010; Smickl et al., 2010; Jia et al., 2011; Chen et al., 2012). However, speciation events also occurred in the short evolutionary time since LGM within post-glacial lineages, during their migration or range expansion (*e.g.* Nies and Reusch, 2005; Escudero et al., 2019). Identifying the different reproductive barriers reducing gene flow between populations, ultimately leading to speciation (Dobzhansky 1937; Mayr 1942; Coyne and Orr 2004), is crucial to gain insight in the role of reproductive isolation (RI) in long-range post-glacial expansion.

Intrinsic reproductive isolation builds up between populations in long term geographical isolation or isolation-by-distance, due to genetic differences accumulated by drift (Wright, 1943, 1946). This mechanism is particularly strong in small populations exposed to high drift (*e.g.*, Carson, 1975; Templeton, 1981; Uyeda, 2009), and could be strengthened by the repeated founder events during range expansion (Barton, 1984; Gottlieb, 2004). Along ecological gradient, speciation could also result from ecological specialization, where RI emerges due to adaptation to different habitats (Schluter, 2000; Doebeli and Dieckmann, 2003; Coyne and Orr, 2004; Rundle and Nosil 2005; Lowry et al. 2008a; Nosil, 2012), even under moderate gene flow (Egan et al., 2015). Isolation due to divergence in adaptation can in turn lead to intrinsic barriers to gene flow, by differentiation through

drift or further divergent adaptation (Funk et al., 2011; Shafer and Wolf, 2013; Wang and Bradburd, 2014).

Along environmental gradients, the evolution of strong reproductive barriers could also favor adaptation necessary for successful range expansion: gene flow from more central populations, less adapted to the environmental conditions experienced by range-edge populations, could be maladaptive and constrain range expansion (Haldane 1956; García-Ramos and Kirkpatrick, 1997; Kirkpatrick and Barton 1997, reviewed in Lenormand, 2002; Bridle and Vines, 2007; recent theory in Polechová and Barton 2015; Polechová 2018). Fitness loss due to maladaptive hybridization can act as selection to locally drive the evolution of strong premating barriers, in a process called reinforcement (Dobzhansky 1940; Servedio and Noor 2003; Garner et al., 2018), typically observed in secondary contact.

To understand which isolation process acted between two recently diverged species, it is crucial to identifying the timing and strength of each reproductive barrier. These are usually classified in three categories, depending if they act before or after mating, and if they rely on extrinsic or intrinsic factors (Nosil et al., 2005). In plants, premating barriers can include spatial, temporal or ecological isolation (reviewed in Lowry et al. 2008a), pollinator or pollen discrimination (Moore and Pannell 2011; *e.g.* Hopkins and Rausher, 2011), or mating system incompatibilities such as selfing (Wright, 2013; Hu, 2015). Post mating, intrinsic prezygotic barriers prevent successful fertilization through pollen-pistil interactions such as divergences in pollen attraction (Swanson et al., 2004), or through gametophytic incompatibilities (Rieseberg and Willis, 2007), while intrinsic postzygotic barriers prevent the development or fertility of hybrids (Rice and Hostert, 1933), linked to genetic incompatibilities or the breackdown of favorable genetic interactions in both parents (Orr and Turelli, 2001). Extrinsic postmating barriers prevent the survival or reproduction of hybrids maladapted to the different parental environments (Rundle and Nosil, 2005)

Here, we investigate the timing and type of reproductive isolation between two recently diverged North American plant species: *Arabidopsis lyrata* subsp. *lyrata* (L.) and its northern selfing

parapatric sister species *A. arenicola* (Fig 1). *Arabidopsis lyrata* occurs in the eastern USA and southeastern Canada, with a well-defined northern distribution limit in North America, represented by the northern shores of the Great Lakes. The contemporary distribution of this species is characterized by a fast post-glacial range expansion, originating from two distinct refugia (Griffin and Willi, 2014; Willi et al., 2018). *Arabidopsis arenicola* occurs at higher latitudes, up to the subarctic regions of North America and Greenland (Hopkins, 1937; Mulligan, 1996; in Warwick et al., 2006). Previous phylogeographic studies suggest *A. arenicola* diverged during the post-glacial range expansion of *A. lyrata*, from range-edge selfing populations of the northern shores of Lake Superior (Schmickl et al., 2010; Hohmann et al., 2014; Novikova et al., 2016; Walden, N., and Willi, Y., *in prep.*). *Arabidopsis arenicola* then migrated toward higher latitudes, while *A. lyrata* remained constrained around the Great Lakes, expanding eastwards along its northern ecological niche limits (Lee-Yaw et al. 2018; Willi et al., 2018). A recent climate chamber study concluded that differences in adaptation strategies to cold climates between both species potentially allowed *A. arenicola* to colonize higher latitudes while *A. lyrata* remained constrained at lower latitudes (Perrier and Willi, *in prep*). This study also suggested that these difference in adaptation could prevent each species from colonizing its sister's current range, hinting toward reproductive isolation by selection against immigrants. In addition, both species are likely isolated by premating reproductive barrier due to their low geographical overlap and the selfing mating system of *A. arenicola.* Here we tested if additional reproductive barriers evolved between both species, contributing to their adaptive divergence. We addressed the following two questions: (*i*) Are both species separated by intrinsic reproductive barriers limiting hybrid formation, survival or fertility? (*ii*) Is hybridization maladaptive compared to one or both species?

Material and methods

Plant material and crossing experiment

We selected two populations of *A. arenicola*: MB1 and QC1, and four populations of both genetic lineages of *A. lyrata* (Table S1, Fig. 1; Willi et al. 2018): From the western genetic lineage that presumably gave rise to *A. arenicola,* we chose the selfing population ON11, also presumed the most related population to *A. arenicola* in our sampling (Walden, N., and Willi, Y., *in prep.*), and WI1. From the eastern genetic lineage, we chose NY5 and MD2, latter presumed to be the most genetically distinct from *A. arenicola* in our sampling due to its position at the leading edge of this lineage (Willi et al., 2018). Seeds were collected on supposedly unrelated individuals (seed families) in each population between 2007 and 2017, and stored in separate bags for each sampled maternal plant, at 4 °C under dark and dry conditions.

To assess the strength of reproductive isolation between the two species, we generated F1 offspring of each population (within-species crosses, WSC) as well as interspecific F1 hybrids (between-species crosses, BSC). In 2016 and 2017, we raised 26 individuals of each seed family per population in growth chambers, later transferred in a greenhouse for crossing (see Table S2 for raising conditions). We performed non-reciprocal WSC by crossing 12 randomly chosen "mother" plants (pollen recipients) with 12 "father" plants (pollen donor) randomly chosen within the remaining individuals of the same population, forming 12 cross combinations per population. To generate F1 hybrids between species, we paired each *A. arenicola* population with two of the *A. lyrata* populations (detailed in Table S3, Fig. 1). Within population pairs involving QC1, reciprocal BSC were performed between randomly chosen pairs of WSC "mother" individuals of both populations. Both individuals were used as pollen recipient and pollen donor. MB1 was collected in late summer 2017, after most individuals raised for the crossing experiment stopped flowering. Therefore, for population pairs involving MB1, reciprocal BSC were performed similarly as for QC1 by using the 12 WSC "mother" individuals of MB1, but using 12 F1 WSC offspring of each *A. lyrata* partner population considered as equivalent to F0 individuals, raised in similar conditions as in Table S2. In total we performed

crosses on 24 cross combinations per population pair, 12 for each cross direction. Hand pollination was performed on emasculated buds to prevent cross- and spontaneous self-pollination, with particular attention to not select buds with early opening anthers for individuals of selfing populations. Each cross combination was attempted at least 3 times. If WSC cross combinations failed, the father was replaced by one of the backup plants (to discard incompatible cross combinations due to this species' self-incompatibility system), while unsuccessful BPC cross combinations were not replaced. Each successful WSC and BSC cross combination was repeated to obtain at least five siliques with enough healthy seeds for the climate chamber experiment. Siliques were collected when ripe and dried two weeks at ambient temperature in the dark. Seeds were then stored at 4 °C under dry and dark conditions.

Climate chamber experiment

To assess selection against hybrids, we raised BSC and WSC offspring in a climate chamber experiment originally designed to study the difference in adaptation to cold climates between both species (Perrier and Willi, *in prep*). The climate chamber experiment followed a two-by-two factorial design over four climate chambers: we setup each climate chambers to one of the two different growth temperatures: *mild* (20 °C) and *cold* (12 to 14 °C), simulating conditions close to those experienced in nature by *A. lyrata* and *A. arenicola,* respectively. These temperatures were derived from extrapolations of the average temperatures in late summer (July and August) of the selected populations in each species (WorldClim database version 2.0, Fick and Hijmans, 2017; Table S1). and one the two frost treatments: *frost,* simulating recurrent frost events of low intensity or *control* (= no frost), resulting in four distinct temperature-treatment combinations (condition). Three spatial blocks were assigned to each condition, and were weekly re-positioned within each climate chamber, and between climate chambers simulating the same growth temperature when frost was not applied.

In total, we obtained 65 WSC cross combinations (10 to 12 per population, Table S3) usable for the climate chamber experiment. We only included 36 BSC combinations (6 per reciprocal

population pair, randomly chosen within the available combinations) in the climate chamber experiment due to high cross and seed failure. Two BSC cross directions of the population pairs were not included in the climate chamber experiment as they systematically failed to produce fruits or healthy seeds: QC1(mother)xWI1 and ON11(mother)xQC1. Two seeds per cross combination (one if less than 24 seeds) were sown in 12 individual pots randomly assigned to one of the twelve replicate blocks across conditions, filled with a standard substrate mixture of washed river sand and peat (1:1.5 sand:peat). Within blocks pots were randomly positionned across two 54-cell propagation trays (BK Qualipot, Otelfingen, Switzerland). Across the twelve blocks, a total number of 2088 seeds were sown in 1296 pots (Table S3, ((6 WSC populations * 12 seed families) + (6 BSC population pairs * 6 seed families)) * 4 conditions * 3 blocks). Pots were saturated with water, then placed for 20 days at 4 °C in the dark in the four climate chambers for seed stratification. Seedlings were randomly thinned to one individual 28 days after germination.

The climate chamber experiment first started with a phase simulating environmental conditions representative of fall, allowing germination and vegetative growth. Seedlings were then vernalized (constant 4 °C, low light) to simulate the presence of a snow cover in a second winter phase. A third spring phase simulated conditions representative of spring and summer for a second period of vegetative growth and reproduction. Nightly frost was simulated during two weeks before winter and during six weeks after winter. The length of each phase, the duration, temperatures and light conditions of day and night cycles, as well as the setup of frost cycles are summarized in Table S4. The experiment lasted for a total of 213 days. Trays were regularly watered for optimal substrate moisture. Fertilizer was added starting 87 days after initiation of germination, once every two weeks (2% v/v Wuxal universal fertilizer, Hauert Manna Düngerwerke GmbH, Nürnberg, Germany).

Individual performance estimates

During the crossing experiment, the success of each cross combination was recorded at the level of the bud. For at least two fruits per cross combination, the number of healthy looking seeds and the

total number of embryos were manually assessed under a binocular. During the climate chamber experiment, individual plant performance was tracked at the level of the seedling until thinning, then at the level of the individual for the whole length of the experiment (recording rates detailed in Table S4). We recorded the day of germination, defined as when a seedling had two fully open cotyledons, as well as the day of death over the whole length of the experiment. Variation in rosette damage was recorded starting at the first day of the first frost treatment, as visual estimation of the proportion of the rosette affected by discoloration, desiccation or necrosis, split into five classes (0 = 0 %, 1 = < 25 %, 2 = < 50 %, 3 = < 75 %, 4 = > 75 %). After vernalization we estimated reproductive output four weeks after the first flower opening of each individual, by counting for each individual the number of fruits (populations MB1, ON11 and QC1 were autonomously selfing), pedicels (flowers that did not develop into a fruit), open flowers, and flower buds on all inflorescences.

Statistical analyses

Intrinsic reproductive isolation: To estimate intrinsic postmating reproductive isolation (RI) between species (research question *i*), the main dependent variable was RI based on *total multiplicative performance (*MP_{total}*)*, integrating performance from the crossing experiment and from the individuals of the climate chamber experiment raised under mild growth temperatures in the control treatment, to express total RI. This variable was calculated for each cross combination as the product between the ratio of *crossing success* of each cross combination, the average *number of healthy seeds* produced per fruit, as well as the average F1 *germination* rate and *reproductive output*, both averaged at the level of the pot, then at the level of the cross combination. *Crossing success* represents an intrinsic prezygotic but postmating barrier, while the other three traits represent successive intrinsic postzygotic barriers. All traits are summarized in Table S5. For BSC cross-combinations that successfully produced healthy seeds but were not included in the climate chamber experiment, *MP* was calculated using the *germination* and *reproductive output* averaged on the level of the population pair, with respect to the crossing direction. For each BSC cross combination, RI was calculated using

the formula derived from Sobel and Chen (2014): RI = 1 – 2 (H / (H + C)), where RI varies between 1 (complete isolation), 0 (no RI) and -1 (facilitation of heteorspecific mating). Heterospecific mating (H) was based on BSC performance estimates, and conspecific mating (C) on the average performance of both parental WSC. To estimate at which life stage the reproductive barriers were acting, RI was further calculated on each of the four individual components of *MP*, averaged at the level of the cross combination. All RI estimates were analyzed assuming a normal distribution. In our main analysis, the strength of RI was first assessed by testing for each BSC population pair if the average RI based on *multiplicative performance* ($RI_{MPtotal}$) was significantly different from 0 by performing a simple one-sample t-test. We further tested differences in $RI_{MPtotal}$ between each BSC population pair by performing a one-way ANOVA, with *population pair* as fixed effect, followed by a Tukey's multiple comparison test. In this analysis, the direction of crossing was not taken into account: reciprocal cross combinations were considered as distinct combinations within a population pair. In secondary analyses, similar analytic steps were performed on RI estimations based on each individual component of *MP*. For RI based on *reproductive output*, the crossing pair MB1xMD2 (both cross directions) was removed from the analysis due to too low replication number, as few hybrids reached this life stage.

Extrinsic reproductive isolation: To estimate if hybridization was maladaptive (research question *ii*), we focused on the two most defining variables of the different adaptive strategies of both species. A previous study analyzing only WSC performance of the same climate chamber experiment (Perrier and Willi, *in prep*) revealed that *A. arenicola* was more tolerant to cold growing temperatures, potentially to extend the growing season in the coldest months, and also more frost tolerant after winter under warm growth temperatures, to compensate for the occasional frost events occurring in the warm season at high latitudes (Billings, 1974). *Arabidopsis arenicola* was however less frost resistant and tolerant after winter under cold growth temperatures, potentially constraining its survival at mid latitudes as early snow melt in the cold season would expose it to frost (Inouye et al., 2003; Inouye, 2008). On the contrary, northern populations of *A. lyrata* were more frost tolerant and

resistant to frost after winter under cold growth temperatures, but were less frost tolerant after winter under mild growth temperatures, and less tolerant to cold growing temperatures, constraining further northwards colonization. Similar to Perrier and Willi (*in prep*), we assessed tolerance to cold growing temperatures and to frost as the capacity to mitigate the negative fitness impact of both stresses (Agrawal et al., 2014), and resistance to frost as the capacity to reduce the damage and lethality of frost itself (Agrawal et al., 2014). Tolerance was based on *multiplicative performance* calculated at the level of the pot only on the climate chamber experiment (MP_{CC}), as the product between the F1 germination rate within a pot, and the F1 reproductive output (set to 0 if individuals died or did not flower). Frost resistance was based on *frost damage*, as the binary increase in damage (0 = no variation, 1 = increase in damage) recorded during the second frost treatment (all traits are summarized in Table S5).

We assessed variation in tolerance to cold growth temperatures by testing the variation in MP_{CC} estimated on the level of the pot in hierarchical mixed-effects models. This analysis considered only WSC individuals of *A. arenicola* and northern *A. lyrata*, as well as their hybrids. Fixed effects were the categorical variables *cross type*, estimated on three levels: WSC of *A. arenicola* (*A*, populations MB1 and QC1), or northern *A. lyrata* populations (*L*, populations ON11, NY5) and their hybrids (H, reciprocal cross-combinations considered as distinct combinations within a population pair), as well as *growth temperature* estimated on two levels: *mild* [0] and *cold*, and their interactions. Random effects were the effect of block, crossed with the combination of both parental families nested within the combination of both parental populations, and the combination of both parental populations. MP_{CC} was 0 inflated, which suggested the modelling of two processes, a Gaussian process (for performance values > 0, $\log_{10}$-transformed), and a logistic process (modelling the probability of 1, assigned to performance values > 0). We performed these analyses in a Bayesian framework, with the package *MCMCglmm* (Hadfield, 2010, 2019) in *R* (R Core Team 2019) on 10 parallel chains (model and prior parametrization detailed in Appendix S6A). The respective contribution of each fixed effect was tested by comparing DIC values of the full model against three

alternative models, one excluding *cross type*, one excluding *growth temperature*, and one with the two fixed effect but excluding the interaction. For the fixed effects which removal led to a lower model fit, Tukey's tests were performed using the package *emmeans* (Lenth, 2019) to test the significant difference between each level of the fixed effect. Comparisons among *cross-type* included the difference between each level of *cross type* within each level of *growth temperature.* Comparisons of *growth temperature* included the difference between *frost* and *control* within each *cross-type* level. For the interaction between *cross-type* and *growth temperature*, the contrast targeted the comparison of slopes of *growth temperature* between each level of *cross type.*

To assess variation in tolerance to frost, we tested MP_{CC} estimated on the level of the pot in similar hierarchical mixed-effects models, with the effect of *treatment* (*control* [0] and *frost*) replacing *growth conditions*. This analysis was performed in parallel in two subsets considering individuals grown in mild or cold growth temperatures. Similarly, to assess variation in resistance to frost, we tested *frost damage* in hierarchical mixed-effects models with the same structure as for frost tolerance, using restricted maximum likelihood with the packages *lme4* (Bates et al., 2015) and *LmerTest* (Kuznetsova et al., 2017; model parametrization in Appendix S6B). *Frost damage* was analyzed on the level of the pot assuming binomial distribution.

Results

Our main analysis testing for the strength of intrinsic reproductive barriers between population pairs (not distinguishing cross direction) revealed that all interspecific population pairs displayed significantly positive RI based on *total multiplicative performance* estimates (Table 1), ranging from 0.51 to 0.88 (0: random mating, 1: total isolation). $RI_{MPtotal}$ showed few significant differences between population pairs (results of the ANOVA in Table S7, multiple comparison in Table S8): RI of the population pair MB1xMD2 was significantly higher than for the two pairs QC1xON11 and MB1xNY5. The secondary analysis testing the strength of RI on each component of *MP* revealed

significant differences between population pairs (Table 1, Fig. 2, Table S8): RI based on *crossing success* was significant in all pairs except MB1xNY5, with the highest RI observed for QC1xON11, and RI based on the number of *healthy seeds* was significant in all pairs except QC1xON11, with the highest RI observed for QC1xWI1. Only MB1xMD2 showed significant positive RI based on *germination* rates, with however no significant differences with other population pairs. QC1xON11showed significant negative RI on *reproductive output*, indicating facilitation of hybrids, but only significantly lower than QC1xWI1.

In our analysis assessing variation in tolerance to cold growth temperatures between both species and their hybrids, the model comparison on the hierarchical mixed-effects model analyses testing the variation in *multiplicative performance* based on the climate chamber experiment *(MP_{CC})* between *cross type*, *growth temperatures* (*cold* vs *mild* [0]) and their interactions, revealed that only the interaction between fixed effects contributed to the fit of the full model (Table S9A), suggesting that the effect of *growth temperature* differed between *cross types*. The interaction between *cross type* and *growth temperature* was significantly positive when comparing *A. arenicola* to northern *A. lyrata* populations in both log-normal and logistic process, and significantly negative when comparing hybrids to *A. arenicola* in the log-normal process, indicating as a generally more positive effect of *growth temperatures* on *A. arenicola* (Table 2, Fig. 3). This interaction was also significantly positive when comparing hybrids to northern *A. lyrata* population in the logistic process.

In our analysis assessing variation in frost tolerance under mild and cold growth temperatures between both species and their hybrids, the model comparison on hierarchical mixed-effects model analyses testing the variation in MP_{CC} between *cross type*, *treatment* (*frost* vs *control* [0]) and their interactions, revealed that under both growing temperatures, only the interaction between fixed effects contributed to the fit of the full model (Table S9B and C). Under mild growth temperatures, the interaction between *cross type* and *treatment* was significantly positive when comparing *A. arenicola* to northern *A. lyrata* populations in the log-normal process, and significantly negative when comparing hybrids to *A. arenicola* in the log-normal process, indicating as a generally more positive

effect of *treatment* on *A. arenicola* under mild growth temperatures (Table 3A, Fig. 3). The opposite pattern was observed under cold growing temperatures, indicating as a generally more negative effect of *treatment* on *A. arenicola* under cold growth temperatures. This interaction was also significantly negative when comparing hybrids to northern *A. lyrata* population in the logistic process.

In our analysis assessing variation in frost resistance based on *frost damage* between both species and their hybrids under mild growth temperatures, *cross types* did not differ within *treatments* (Table 3B, results of the ANOVA in Table S9D). However, the effect of *treatment* significantly increased *frost damage* for all *cross types* (Table 3B, Fig. 3). Interactions between fixed effect were not significant. Under cold growth temperatures, *A. arenicola* had lower damage than northern *A. lyrata* in the *control* treatment (Table 3B) The effect of treatment significantly increased *frost damage* in A. *arenicola* and in the hybrids (Table 3B, Fig. 3). The interaction between fixed effects revealed a significant positive difference when comparing *frost damage* of *A. arenicola* and WSC of norther *A. lyrata*, and when comparing hybrids to *A. lyrata*, suggesting that *A. arenicola* and the hybrids were more positively affected by the effect of *treatment* under cold growth temperatures than *A. lyrata* (Table 3B, Fig. 3).

Discussion

We found clear patterns of reproductive isolation between the two North American plant species *Arabidopsis arenicola* and *A. lyrata* subsp. *lyrata*, in line with their divergent colonization patterns and divergent adaptation strategies despite a recent post-glacial common ancestor. In our study, both species are mainly separated by reproductive barriers acting early after pollination, reducing the rate of successful crosses and the production of healthy seeds. Past these two stages, reproductive isolation mainly acted by selection against hybrids, maladapted to both parental habitats.

Total intrinsic reproductive isolation (RI) between *A. arenicola* and *A. lyrata*, estimated on multiplicative performance, was significant for all four population pairs, with high levels of RI

varying between 0.51 and 0.88. These values are higher than average RI between closely related species pairs found in other studies (Lowry et al., 2008a; Baak et al., 2015), but reached similar levels than RI between lineages within *Campanula americana* (Barnard-Kubow et al., 2016; Barnard-Kubow and Galloway, 2017), a North American species with similar post-glacial range expansion history as *A. lyrata*. In *A. lyrata*, a previous study involving crosses between natural populations within each post-glacial genetic lineage further revealed outbreeding depression in a fourth of the tested cross combination (Perrier et al., 2020), linked to genetic incompatibilities such as the Dobzhansky–Muller type (Lynch, 1991; Oakley et al., 2015). The occurrence of reproductive barriers already within post-glacial lineages imply that RI could rapidly accumulate in our system, supporting the strong RI observed between both species in our study.

Intrinsic reproductive isolation between both species was also significant for the first three successive life stages composing multiplicative performance, but not for reproductive output. RI was especially strong for crossing success (on average ca. 0.28, maximum = 0.71) and number of healthy seeds (on average ca. 0.64, maximum = 0.75), significant for three out of four population pairs at each life stage. Failure at crossing success could result from differentially exclusion of pollen phenotypes (Moore and Pannell, 2011), or from gametophytic incompatibilities (Rieseberg and Willis, 2007). These incompatibilities are complex and affect the multiple stages between the moment pollen attaches to the stigma and the fusion of both gametes, but rely mostly on divergence in reproductive signaling such as the guiding of pollen tubes to the ovules (Swanson et al., 2004; Müller, 2014). In crosses between *A. thaliana* and the European *A. lyrata* subsp. *petrea*, 50% of crosses resulted in pollen tube overgrowth, linked to divergence in pollen tube attractants (Escobar-Restrepo et al., 2007). Furthermore, large degree of variation in pollen tube reception was found between crosses between both species (Müller, 2014), linked to the evolution of species-specific glycosylation patterns. Reduction in seed set was also observed in crosses between the closely related *Arabidopsis* species *A. lyrata* subsp. *petrea* and *A. arenosa*, with however higher levels of RI than observed in our study: crosses between parents of equal ploidy resulted in only 2% to 10% live seed (Josefsson et al.,

2006), and crosses between diploid individuals of both species only produced between 0 – 30% of normal seeds compared to parents (Burkart-Waco et al., 2012). Lower RI is to be expected in our system, as *A. thaliana* and *A. arenosa* diverged much earlier (3.8–5.8 million years ago), and additionally differ in their chromosome numbers (*A. thaliana* = 5, *A. arenosa* = 8). Finally, the levels of RI based on germination (ca. 0.14, maximum = 0.35) were on average higher in our study than in crosses between the North American and European *A. lyrata* (Hämälä et al., 2017; recalculated RI = ca. 0.02). Crosses between *A. lyrata* subsp. *petrea* and *A. arenosa* also showed reduced germination (Muir et al., 2015). The levels of RI based on germination in our study were however lower than observed between between post-glacial clades of *C. americana* (87% reduction in germination, Barnard-Kubow et al., 2016). Overall, intrinsic RI was expressed at several life stages, and varied strongly between population pairs, suggesting that recent speciation can result from multiple genes and incompatibilities (e.g., Levy 1991; Skrede et al. 2008; Barnard-Kubow and Galloway, 2017).

The comparison of RI based on multiplicative performance partially suggest that RI was stronger in crosses between *A. arenicola* and the south-eastern population of *A. lyrata* MD2. This pattern is less clear when analyzing individual life stage: Despite showing only significant expression in crosses involving MD2, RI based on germination showed no significant differences between population pairs. However, crosses involving MD2 consistently showed RI in the first three successive barriers, effectively reducing the production of viable hybrids, with only two seed families reaching the flowering stage. MD2 could be the most genetically distinct *A. lyrata* population from *A. arenicola* within our sampling, due to its position at the leading edge of the eastern cluster (Willi et al., 2018), separated before LGM from the western cluster which gave rise to *A. arenicola*. Stronger RI based on multiplicative performance in crosses involving MD2, and the expression of RI at three out of 4 life stages hints toward a clocklike evolution of multiple genetic incompatibilities in our study system, gradually increasing with divergence time (Moyle et al., 2004; Scopece et al., 2007; Barnard-Kubow et al., 2016; Barnard-Kubow and Galloway, 2017).

On the contrary, crosses involving ON11, the presumed closest relative to A. *arenicola* in our sampling (Walden, N., and Willi, Y., *in prep.*), showed the strongest RI based on crossing success, but showed no RI in subsequent life stages, and even heterosis for reproductive isolation. Stronger RI based on pollen-stigma incompatibilities in sympatric populations compared to allopatric populations of recently diverged tropical plants was previously interpreted as reinforcement (Kay and Schemske, 2008). Strong RI acting early after pollination with a population located close to the ancestral divergence zone could indicate an initial step of reinforcing selection to prevent maladaptive hybridization, while reduced RI in subsequent life stages are concordant with the recent divergence time between both populations. While the northward migration of *A. arenicola* resulted in a parapatric distribution with A. lyrata, the patterns of reinforcing selection could have been partially maintained due to the selfing mating system of ON11 and *A. arenicola*, allowing for better storage of kryptic genetic variation (Clo et al., 2020). Alternatively, selfing in ON11 could lead to stronger RI by facilitating the accumulation of hybrid incompatibilities by genetic drift (Wright et al., 2013). Previous results have indeed shown that selfing populations of *A. lyrata*, ON11 included, are subjected to strong genetic drift (Willi et al., 2018). The expression of heterosis in this population pair could potentially result from the accumulation of recessive deleterious mutations during range expansion (Willi et al., 2018; Perrier et al., 2020), which have been recently shown to result in heterosis in secondary contact hybrid zones despite RI (MacPherson et al., 2020).

The analysis of extrinsic reproductive barriers revealed that hybrids were generally maladaptive compared to both adaptive strategies to cold climates of *A. arenicola* and northern *A. lyrata* populations. The adaptive strategy of *A. arenicola* has been recently linked to higher tolerance to cold growth temperature, and higher frost tolerance based in mild growth temperatures (Perrier and Willi, *in prep*), both based on multiplicative performance. Multiplicative performance was reduced by cold growth temperature in the hybrids compared to *A. arenicola*, but still higher than northern *A. lyrata* populations. Multiplicative performance was also reduced by frost under mild growth temperatures in the hybrids to similar levels than northern *A. lyrata* populations. This results indicate

that hybrids between both species could be constrained by the shorter reproductive season and the higher frequency of frost events happening in summer at high latitudes (Billings, 1974). The adaptive strategy of northern *A. lyrata* populations was linked to higher frost tolerance under cold growth temperatures based on multiplicative performance, and higher resistance under cold growth temperatures based on frost damage (Perrier and Willi., *in prep*). In our study, multiplicative performance was also reduced by frost under cold growth temperatures in the hybrids compared to northern *A. lyrata,* but still higher than *A. arenicola*, while hybrids showed increased frost damage under cold growth temperatures than northern *A. lyrata* populations, to similar levels than *A. arenicola.* These results suggest that hybrids could also be more sensitive to frost event occurring in cold conditions after winter, in early spring right after snow melt (Inouye et al., 2003; Inouye, 2008). Differences in adaptation has long been suggested as a major contribution to reproductive isolation between parapatric species (Rundle and Nosil 2005; Martin and Willis, 2006; Lowry et al. 2008a,b; Melo et al. 2014; Cahenzli et al., 2018). However, most studies tend to find that premating extrinsic barriers linked to adaptation *i.e.* selection against immigrants, could play a stronger role than postmating extrinsic barriers, *i.e.* selection against hybrids, latter showing mostly intermediate levels of adaptation compared to their parents (Nosil et al., 2005; Martin and Willis, 2006, Lowry et al., 2008b). Previous results in *Arabidopsis* tend to show the same patterns: strong local adaptation was found between European and north American *A. lyrata* populations (Leinonen et al., 2011), but hybrids of both populations showed intermediate performance in the warmer north American transplant site, and even increased performance in the colder European transplant site, latter linked to heterosis. Similarly, interspecific hybrids between alpine and lowland European *A. lyrata* populations show mostly intermediate gene expression levels relative to both parents (Videvall et al 2015). Our results however suggest that some of the key adaptations in both species, *i.e.* tolerance to frost under mild growth temperatures in *A. arenicola* and frost resistance under cold growth temperatures in *A. lyrata,* are not additive but could be based on dominance. For these traits the complete breakdown of adaptation in hybrids to levels similar to the non-adapted parent could result in strong selection

against hybrids. Divergence in adaptation strategies between *A. lyrata* and *A. arenicola* could therefore contribute both to premating and postmating extrinsic barriers, similarly to results found in closely related *Dianthus* species (Cahenzli et al., 2018), reporting strong premating and postmating extrinsic RI linked to specialization to dry rocky habitats.

Overall, both intrinsic and extrinsic postmating barriers were strong in our study system, in contrast with magnitudes of RI found between *A. lyrata* and other species (Leinonen, 2011; Muir et al., 2015; Hämälä et al., 2017), but magnitudes closer to RI between the selfing *A. thaliana* and *A. arenosa* (Josefsson et al., 2006; Burkart-Waco et al., 2012). The strong RI observed in our study despite recent divergence could be in part linked the fast range expansion reported in *A. lyrata* (Willi et al., 2018) and presumed in *A. arenicola* (Schmickl et al., 2010; Walden, N., and Willi, Y., *in prep.*), as the accumulation of genetic differences through drift is strengthened by the repeated founder events during range expansion (Barton, 1984; Gottlieb, 2004). Heightened drift with a longer history of range expansion has indeed been reported in *A. lyrata* (Willi et al., 2018), but the levels of drift in *A. arenicola* are yet unknown. Strong RI could also be linked to the selfing mating system of *A. arenicola*. A selfing mating system is indeed expected to lead to increased genetic drift (Pollak, 1987; Nordborg and Donelli, 1997) and mutation accumulation (Lynch et al., 1995; Schultz and Lynch, 1997). Selfing further provides strong premating isolation (Hu et al., 2015), thereby facilitating the accumulation of hybrid incompatibilities by genetic drift, strengthening intrinsic reproductive barriers (Wright et al., 2013), as observed between selfing arctic *Draba* species (Grundt et al., 2006). The reduction of gene flow due to selfing has also been suggested to lead to increased local adaptation, contributing to budding speciation (Wright et al., 2013; Hodgins and Yeaman, 2019). Selfing could have therefore facilitated adaptive divergence *in A. arenicola* by increasing the strength of extrinsic reproductive barriers.

Conclusions

Our study empirically highlights how reproductive isolation can rapidly accumulate during a post-glacial speciation event, and potentially contributing to successful post-glacial colonization toward subarctic climate. Substantial reproductive isolation was found both in intrinsic and extrinsic barriers, in addition of selfing, biogeographical separation and diverging adaptation. The levels of intrinsic RI were especially high despite the recent divergence of both species, acting mostly in early stages after pollination by reducing the rate of successful crosses and the production of healthy seeds. Hybrids generated despite these barriers were however maladapted to both parental habitats, suggesting postmating extrinsic RI. Rapidly accumulated RI early in the divergence process between both species, potentially facilitated by selfing in *A. arenicola* and strong drift could have prevented maladaptive gene flow, allowing *A. arenicola* to expand toward higher latitudes.

Acknowledgements: This work was supported by the Swiss National Science Foundation (31003 A_166322). For assistance with crosses, seed counts, sowing, leaf measurements and electrolyte leakage assays, we thank Georg Armbruster, Olivier Bachmann, Markus Funk, Maja Pfund, Susanna Riedl and Darío Sánchez-Castro. For his advices on reproductive isolation and assistance in experimental design, we thank Kay Lucek. Collection permits were provided by the New York State Office of Parks and the United States National Park Service.

References

Agrawal, A. A., J. K. Conner, and J. R. Stinchcombe. 2004. Evolution of plant resistance and tolerance to frost damage. *Ecology Letters* 7: 1199–1208.

Baack, E., M. C. Melo, L. H. Rieseberg, and D. Ortiz-Barrientos. The origins of reproductive isolation in plants. *New Phytologist* 207: 968–984.

Barnard-Kubow, K. B., and L. F. Galloway. 2017. Variation in reproductive isolation across a species range. *Ecology and Evolution* 7: 9347– 9357.

Barnard-Kubow, K. B., N. So, and L. F. Galloway. 2016. Cytonuclear incompatibility contributes to the early stages of speciation. *Evolution* 70: 2752–2766.

Barton, N. H. 1984. Genetic revolutions, founder effects, and speciation, *Annual Review of Ecology and Systematics* 15:133–164.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.

Billings, W. D. 1974. Arctic and alpine vegetation: plant adaptations to cold summer climates. *In* J. D. Ives and R. G. Barry [eds.], Arctic and alpine environments, 403–443. Methuen, London, UK.

Birks, H. H. 2008. The Late-Quaternary history of arctic and alpine plants, *Plant Ecology and Diversity* 1: 135–146.

Bridle, J. R., and T. H. Vines. 2007. Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology and Evolution* 22: 140–147.

Brochmann, C., T. M. Gabrielsen, I. Nordal, J. Y. Landvik, and R. Elven. 2003. Glacial survival or *tabula rasa*? The history of North Atlantic biota revisited. *Taxon* 52: 417–450.

Burkart-waco, D., C. Josefsson, B. Dilkes, N. Kozloff, O. Torjek, R. Meyer, et al. 2012. Hybrid incompatibility in *Arabidopsis* is determined by a multiple-locus genetic network. *Plant Physiology* 158: 801–812.

Cahenzli, F., C. Bonetti, and A. Erhardt. 2018. Divergent strategies in pre- and postzygotic reproductive isolation between two closely related *Dianthus* species. *Evolution* 72: 1851–1862.

Carson, H. L. 1975. The genetics of speciation at the diploid level. *The American Naturalist* 109: 83–92.

Chen, D., X. Zhang, H. Kang, X. Sun, S. Yin, H. Du, et al. 2012. Phylogeography of *Quercus variabilis* based on chloroplast DNA sequence in East Asia: multiple glacial refugia and mainland-migrated island populations. *PLoS ONE* 7: e47268.

Clo, J., J. Ronfort, and D. Abu Awad. 2020. Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations. *Journal of Evolutionary Biology*, in press.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York, USA.

Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. *The American Naturalist* 74: 312–321.

Doebeli, M., and U. Dieckmann. 2003. Speciation along environmental gradients. *Nature* 421: 259–264.

Egan, S. P., G. J. Ragland, L. Assour, T. H. Q. Powell, G. R. Hood, S. Emrich, et al. 2015. Experimental evidence of genome-wide impact of ecological selection during early stages of speciation-with-gene-flow. *Ecology Letters* 18: 817–825.

Escobar-Restrepo, J. M., N. Huck, S. Kessler, V. Gagliardini, J. Gheyselinck, W. C. Ya, et al. 2007 The FERONIA receptor-like kinase mediates male-female interactions during pollen tube reception. *Science* 317: 656–660.

Escudero, M., M. Lovit, B. H. Brown, and A. L. Hipp. 2019. Rapid plant speciation associated with the last glacial period: reproductive isolation and genetic drift in sedges. *Botanical Journal of the Linnean Society* 190: 303–314.

Escudero, M., P. Vargas, P. Arens, N. J. Ouborg, and M. Luceño. 2010. The east-west-north colonization history of the Mediterranean and Europe by the coastal plant *Carex extensa* (Cyperaceae). *Molecular Ecology* 19: 352–370.

Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37: 4302–4315.

Funk, D. J., S. P. Egan, and P. Nosil. 2011. Isolation by adaptation in *Neochlamisus* leaf beetles: Host-related selection promotes neutral genomic divergence. *Molecular Ecology* 20: 4671–4682.

García-Ramos, G., and M. Kirkpatrick. 1997. Genetic models of rapid evolutionary divergence in peripheral populations. *Evolution* 51: 1–23

Garner, A. G., B. E. Goulet, M. C. Farnitano, Y. F. Molina-Henao, and R. Hopkins. 2018. Genomic Signatures of Reinforcement. *Genes* 9: 191.

Gottlieb LD. 2004. Rethinking classic examples of recent speciation in plants. *New Phytologist* 161: 71–82.

Griffin, P. C., and Y. Willi. 2014. Evolutionary shifts to self-fertilisation restricted to geographic range margins in North American *Arabidopsis lyrata*. *Ecology Letters* 17: 484–490.

Grundt, H. H., S. Kjolner, L. Borgen, L. H. Rieseberg, and C. Brochmann. 2006. High biological species diversity in the arctic flora. *Proceedings of the National Academy of Sciences, USA* 103: 972–975.

Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Softwares* 33: 1–22.

Hadfield, J. D. 2019. MCMCglmm course notes. Website https://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf [accessed 14 July 2020].

Haldane, J. B. S. 1956. The relation between density regulation and natural selection. *Proceedings of the Royal Society B, Biological Sciences* 145: 306–308.

Hämälä, T., T. M. Mattila, P. H. Leinonen, H. Kuittinen, and O. Savolainen. 2017. Role of seed germination in adaptation and reproductive isolation in *Arabidopsis lyrata*. *Molecular Ecology* 26: 3484–3496.

Hewitt, G. M. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907–913.

Hewitt, G. M. 2004. Genetic consequences of climatic changes in the Quaternary. *Philosophical Transactions of the Royal Society B, Biological Sciences* 359: 183–195.

Hodgins, K. A. and S. Yeaman. 2019. Mating system impacts the genetic architecture of adaptation to heterogeneous environments. *New Phytologist* 224: 1201–1214.

Hohmann, N., R. Schmickl, T. Y. Chiang, M. Lučanová, F. Kolář, K. Marhold, and M. A. Koch. 2014. Taming the wild: resolving the gene pools of non-model *Arabidopsis* lineages. *BMC Evolutionary Biology* 14: 224.

Hopkins, M. 1937. *Arabis* in eastern and central North America. *Rhodora* 39: 63–186.

Hopkins, R., and M. D. Rausher. 2011. Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature* 469: 411–414.

Hu, X. S. 2015. Mating system as a barrier to gene flow. *Evolution* 69: 1158–1177.

Hughes, P. D., P. L. Gibbard, and J. Ehlers. 2013. Timing of glaciation during the last glacial cycle: Evaluating the concept of a global 'last glacial maximum' (LGM). *Earth-Science Reviews* 125: 171–198.

Inouye, D.W. 2008. Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology* 89: 353–362.

Inouye, D. W., F. Saavedra, and W. Lee. 2003. Environmental influences on the phenology and abundance of flowering by *Androsace septentrionalis* (Primulaceae). *American Journal of Botany* 90: 905–910.

Jia, D., T. Liu, L. Wang, D. Zhou, and J. Liu. 2011. Evolutionary history of an alpine shrub *Hippophae tibetana* (Elaeagnaceae): allopatric divergence and regional expansion. *Biological Journal of the Linnean Society* 102: 37–50.

Josefsson, C., B. Dilkes, and L. Comai. 2006. Parent-dependent loss of gene silencing during interspecies hybridization. *Current Biology* 16: 1322–1328.

Kay, K. M., and D. W. Schemske. 2008. Natural selection reinforces speciation in a radiation of neotropical rainforest plants. *Evolution* 62: 2628–2642.

Kirkpatrick, M. and N. H. Barton. 1997. Evolution of a species' range. *The American Naturalist* 150: 1–23.

Koch, M. A., C. Kiefer, D. Ehrich, J. Vogel, C. Brochmann, and K. Mummenhoff. 2006. Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). *Molecular Ecology* 15: 825–839.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. lmerTest Package: Tests in linear mixed-effects models. *Journal of Statistical Software* 82: 1–26.

Lee-Yaw, J. A., M. Fracassetti, and Y. Willi. 2018. Environmental marginality and geographic range limits: a case study with *Arabidopsis lyrata* ssp. *lyrata*. *Ecography* 41: 622–634.

Leinonen, P.H., D. L. Remington, and O. Savolainen. 2011. Local adaptation, phenotypic differentiation, and hybrid fitness in diverged natural populations of *Arabidopsis lyrata*. *Evolution* 65: 90–107.

Lenormand, T. Gene flow and the limits to natural selection. *Trends in Ecology and Evolution* 17: 183–189.

Lenth, R. 2019. emmeans: estimated marginal means, aka least-squares means. R package version 1.4. Website https://CRAN.R-project.org/package=emmeans [accessed 14 July 2020]

Levy, F. 1991. A genetic analysis of reproductive barriers in *Phacelia dubia*. *Heredity* 67: 331–345.

Lowry, D. B., J. L. Modliszewski, K. M. Wright, C. A. Wu, and J. H. Willis. 2008a. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society B, Biological Sciences* 363: 3009–3021.

Lowry, D. B., R. C. Rockwood, and J. H. Willis. 2008b. Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution* 62: 2196–2214.

Lynch, M. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution* 45: 622–629.

Lynch, M., J. Conery, and R. Bürger. 1995. Mutational meltdowns in sexual populations. *Evolution* 49: 1067–1080.

MacPherson, A., S. Wang, R. Yamaguchi, L. H. Riesesberg, and S. P. Otto. 2020. Parental population range expansion before secondary contact promotes heterosis. *bioRxiv*. doi: https://doi.org/10.1101/2020.04.28.066308.

Martin, N. H., and J. H. Willis. 2007. Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution* 61: 68–82.

Melo, M. C., A. Grealy, B. Brittain, G. M. Walter, and D. Ortiz-Barrientos. 2014. Strong extrinsic reproductive isolation between parapatric populations of an Australian groundsel. *New Phytologist* 203: 323–334.

Moore, J. C., and J. R. Pannell. 2011. Sexual selection in plants. *Current Biology* 21: R176–R182.

Moyle, L. C., M. S. Olson, P. Tiffin, and D. Baum. 2004. Patterns of reproductive isolation in three angiosperm genera. *Evolution* 58:1195–1208.

Muir, G., P. Ruiz-Duarte, N. Hohmann, B. K. Mable, P. Novikova, R. Schmickl, et al. 2015. Exogenous selection rather than cytonuclear incompatibilities shapes asymmetrical fitness of reciprocal *Arabidopsis* hybrids. *Ecology and Evolution* 5: 1734–1745.

Müller, L. M. The role of pollen tube reception in reproductive isolation. Ph.D. dissertation, University of Zürich, Zürich, Switzerland.

Mulligan, G. A. 1996. Synopsis of the genus *Arabis* (Brassicaceae) in Canada, Alaska, and Greenland. *Rhodora* 97: 109–163.

Nies, G., and T. B. Reusch. 2005. Evolutionary divergence and possible incipient speciation in post-glacial populations of a cosmopolitan aquatic plant. *Journal of Evolutionary Biology* 18: 19–26.

Nordborg, M., and P. Donnelly. 1997. The coalescent process with selfing. *Genetics* 146: 1185–1195.

Nosil, P. 2012. Ecological speciation, 1st ed. Oxford University Press, Oxford, UK.

Nosil, P., T. H. Vines, and D. J. Funk. 2005. Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* 59: 705–719.

Novikova, P. Y., N. Hohmann, V. Nizhynska, T. Tsuchimatsu, J. Ali, G. Muir, A. Guggisberg, et al. 2016. Sequencing of the genus *Arabidopsis* identifies a complex history of non-bifurcating speciation and abundant trans-specific polymorphism. *Nature Genetics* 48: 1077–1082.

Oakley, C. G., J. Ågren, and D. W. Schemske. 2015. Heterosis and outbreeding depression in crosses between natural populations of *Arabidopsis thaliana*. *Heredity* 115: 73–82

Orr, H. A. and M. Turelli. 2001. The evolution of postzygotic isolation: Accumulating Dobzhansky-Muller incompatibilities. *Evolution* 55: 1085–1094.

Perrier, A., D. Sánchez-Castro, and Y. Willi. 2020. Expressed mutational load increases toward the edge of a species' geographic range. *Evolution* 74: 1711–1723.

Pollak, E. 1987. On the theory of partially inbreeding finite populations. I. Partial selfing. *Genetics* 117:353–360.

Polechová, J. 2018. Is the sky the limit? On the expansion threshold of a species range. *PLoS Biology* 16: e2006735

Polechová, J., and N. H. Barton. 2015. Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences, USA* 112: 6401–6406.

R Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rice,W. R., and E. E. Hostert. 1993. Laboratory experiments on speciation: what have we learned in 40 years. *Evolution* 47: 1637–1653.

Rieseberg, L. H., and J. H. Willis. 2007. Plant Speciation. *Science* 317: 910–914.

Riihimaki, M., and O. Savolainen. 2004. Environmental and genetic effects on flowering differences between northern and southern populations of *Arabidopsis lyrata* (Brassicaceae). *American Journal of Botany* 91: 1036–1045.

Rundle, H. D., and P. Nosil. 2005. Ecological speciation. *Ecology Letters* 8: 336–352.

Schluter, D. 2001. Ecology and the origin of species. *Trends in Ecology and Evolution* 16: 372-380.

Schmickl, R., M. H. Jørgensen, A. K. Brysting, and M. A. Koch. 2010. The evolutionary history of the *Arabidopsis lyrata* complex: a hybrid in the amphi-Beringian area closes a large distribution gap and builds up a genetic barrier. *BMC Evolutionary Biology* 10: 98.

Schmitt, T. 2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology* 4: 11.

Scopece, G., A. Musacchio, A. Widmer, and S. Cozzolino. 2007. Patterns of reproductive isolation in mediterranean deceptive orchids. *Evolution* 61: 2623–2642

Servedio, M. R., and M. A. Noor. 2003. The role of reinforcement in speciation: theory and data. *Annual Review of Ecology and Systematics* 34: 339–364.

Shafer, A. B. A., and J. B. W. Wolf. 2013. Widespread evidence for incipient ecological speciation: A meta-analysis of isolation-by-ecology. *Ecology Letters* 16: 940–950.

Skrede, I., C. Brochmann, L. Borgen, and L. H. Rieseberg. 2008. Genetics of intrinsic postzygotic isolation in a circumpolar plant species, *Draba nivalis* (Brassicaceae). *Evolution* 62: 1840–1851.

Sobel, J. M., and G. F. Chen. 2014. Unifi cation of methods for estimating the strength of reproductive isolation. *Evolution* 68: 1511–1522.

Schultz, S. T., and M. Lynch. 1997. Mutation and extinction: the role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. *Evolution* 51: 1363–1371.

Swanson, R., A. F. Edlund, and D. Preuss. 2004. Species specificity in pollen-pistil interactions. *Annual Review of Genetics* 38: 793– 818.

Templeton, A. R. 1981. Mechanisms of speciation-a population genetic approach. *Annual Review of Ecology and Systematics* 12: 23–48.

Uyeda, J. C., S. J. Arnold, P. A. Hohenlohe, and L. S. Mead. 2009. Drift promotes speciation by sexual selection. *Evolution* 63: 583–594.

Videvall, E., N. Sletvold, J. Hagenblad, J. Ågren, and B. Hansson. 2016. Strong maternal effects on gene expression in *Arabidopsis lyrata* hybrids. *Molecular Biology and Evolution* 33: 984–994.

Wang, I. J., and G. S. Bradburd. 2014. Isolation by environment. *Molecular Ecology* 23: 5649–5662.

Warwick, S. I., A. Al-Shehbaz, and A. Sauder. 2006. Phylogenetic position of *Arabis arenicola* and generic limits of *Aphragmus* and *Eutrema* (Brassicaceae) based on sequences of nuclear ribosomal DNA. *Canadian Journal of Botany* 84: 269–281.

Willi, Y., M. Fracassetti, S. Zoller, and J. Van Buskirk. 2018. Accumulation of mutational load at the edges of a species range. *Molecular Biology and Evolution* 35: 781–791.

Wright, S. 1943. Isolation by distance. *Genetics* 28: 114–138.

Wright, S. 1946. Isolation by distance under diverse systems of mating. *Genetics* 31: 39–59.

Wright., S. I., S. Kalisz, and T. Slotte. 2013. Evolutionary consequences of self-fertilization in plants. *Proceedings of the Royal Society B, Biological Sciences* 280: 20130133.

Tables

Table 1: Summary of total and individual contributions to reproductive isolation (RI) per population pair

	QC1xON11			*QC1xWI1*			*MB1xNY5*			*MB1xMD2*		
RI barriers	*N*	*Mean*		*N*	*Mean*		*N*	*Mean*		*N*	*Mean*	
Total RI												
Multiplicative performance	22	**0.52**	**	18	**0.79**	***	24	**0.51**	***	24	**0.88**	***
Individual RI barriers												
Crossing success	23	**0.71**	***	19	**0.25**	*	24	-0.02		24	**0.18**	**
Number of healthy seeds	10	-0.01		16	**0.75**	***	24	**0.28**	***	24	**0.42**	***
Germination	6	-0.05		6	-0.02		12	0.26	(*)	12	**0.35**	*
Reproductive output	6	**-0.45**	**	5	0.10		4	-0.07		2	0.10	‡

Mean RI was estimated for RI barrier within each population pair (not distinguishing cross direction), its significant difference from 0 was tested with a one-sample t-test. For RI based on reproductive output, the crossing pair MB1xMD2 (both directions) was not tested due to too low replication number, as few hybrids reached this life stage (‡). RI varied from 1 (complete isolation), to 0 (random mating), to −1 (outcrossing is favored). Means with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Table 2: Summary of multiple comparisons testing for the effect of cross type (within –species crosses (WSC) of *A. arenicola* [A] compared to WSC of northern *A. lyrata* [L], and hybrids between species [H] compared *A. arenicola* WSC or northern *A. lyrata* WSC), growth temperature (cold compared to mild [0]) and their interaction on multiplicative performance of individuals raised in the climate chamber experiment

		Cross type											
		Control			*Frost*			*Growth temperature (cold vs mild)*			*Cross type * growth temperature*		
		A vs L	*H vs A*	*H vs L*	*A vs L*	*H vs A*	*H vs L*	*A*	*L*	*H*	*A vs L*	*H vs A*	*H vs L*
Dependent variable	*N*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*
Log-normal	204	-0.65	0.71	0.06	0.09	0.13	0.23	0.18	-0.56	-0.40	**0.74** ***	**-0.5** ***	0.16
Logistic	368	-4.81	0.89	-3.93	3.07	-0.01	3.08	-1.34	-9.23	-2.24	**7.81** ***	-0.90	**7.00** **

Multiplicative performance estimates (log_{10}-transformed if >0) based on the climate chamber experiment were assumed to follow Gaussian distributions with 0-inflation. Therefore, models assessed all fixed and random effects for their importance in both the Gaussian process (total reproductive output) and the logistic process (binary variable depicting germination combined with survival and the capacity to initiate flowering). The logistic part of the model predicts non-zeros in the distribution on the logit scale. Coefficients (M = *mean*) depict estimated pairwise difference in performance between cross type within growth temperature, growth temperature within cross type, and estimated linear contrast in magnitude of effect of growth temperature between cross type. Estimates of coefficients are modes of an MCMC sample from the posterior distribution of parameters obtained from a Tukey's test. Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) *P*<0.1, * *P*<0.05, ** *P*<0.01, *** *P*<0.001. Results for random effects are not shown.

Table 3A: Summary of multiple comparisons testing for the effect of cross type (within –species crosses (WSC) of *A. arenicola* [A] compared to WSC of northern *A. lyrata* [L], and hybrids between species [H] compared *A. arenicola* WSC or northern *A. lyrata* WSC), treatment (*frost* compared to *control* [0]) and their interaction on multiplicative performance of individuals raised in the climate chamber experiment under mild or cold growth temperatures

		Cross type											
		Control			*Frost*			*Treatment (frost vs control)*			*Cross type * treatment*		
		A vs L	*H vs A*	*H vs L*	*A vs L*	*H vs A*	*H vs L*	*A*	*L*	*H*	*A vs L*	*H vs A*	*H vs L*
Dependent variable	*N*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*	*M*
Mild growth temperatures													
Log-normal	237	-0.71	0.76	0.05	-0.45	0.42	-0.03	0.23	-0.02	-0.10	**0.25 ****	**-0.34 *****	-0.08
Logistic	368	-3.14	0.55	-2.58	-1.18	0.03	-1.15	-0.62	-2.58	-1.15	1.94 (*****)	-0.53	1.43
Cold growth temperatures													
Log-normal	134	0.04	0.11	0.15	-0.32	0.37	0.06	-0.25	0.10	0.00	**-0.26 ***	**0.24 ***	-0.10
Logistic	368	1.78	-3.52	-0.74	0.79	-4.40	-3.60	-0.82	0.17	-1.70	0.89	-0.89	**-1.90 ***

Multiplicative performance estimates (log_{10}-transformed if >0) based on the climate chamber experiment were assumed to follow Gaussian distributions with 0-inflation. Therefore, models assessed all fixed and random effects for their importance in both the Gaussian process (total reproductive output) and the logistic process (binary variable depicting germination combined with survival and the capacity to initiate flowering). The logistic part of the model predicts non-zeros in the distribution on the logit scale. Models were performed on two subsets, considering individuals raised either under mild or cold growth temperatures. Coefficients (M = *mean*) depict estimated pairwise difference in performance between cross type within treatment, treatment within cross type, and estimated linear contrast in magnitude of effect of treatment between cross type. Estimates of coefficients are modes of an MCMC sample from the posterior distribution of parameters obtained from a Tukey's test. Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) *P*<0.1, * *P*<0.05, ** *P*<0.01, *** *P*<0.001. Results for random effects are not shown.

Table 3B: Summary of multiple comparisons testing for the effect of cross type (within –species crosses (WSC) of *A. arenicola* [A] compared to WSC of northern *A. lyrata* [L], and hybrids between species [H] compared *A. arenicola* WSC or northern *A. lyrata* WSC), treatment (*frost* compared to *control* [0]) and their interaction on frost damage of individuals raised in the climate chamber experiment grown under mild or cold growth temperatures

		Cross type											
		Control			*Frost*			*Treatment (frost vs control)*			*Cross type * treatment*		
		A vs L	*H vs A*	*H vs L*	*A vs L*	*H vs A*	*H vs L*	*A*	*L*	*H*	*A vs L*	*H vs A*	*H vs L*
Dependent variable	*N*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*	*β*
Frost damage (mild)	291	-0.20	-0.40	-0.60	0.09	0.12	0.21	**1.36 ****	**1.06 ***	**1.87 ****	0.29	0.52	0.81
Frost damage (cold)	155	**-2.69 ****	0.34	-2.35	0.53	0.21	0.73	**2.60 *****	-0.62	**2.46 *****	**3.22 ****	-0.14	**3.08 *** †

Frost damage was a binary variable, therefore analyzed by a model predicting non-zeros on the logit scale. Models were performed on two subsets, considering individuals raised either under mild or cold growth temperatures. Each model was optimized with the *bobyqa* optimizer to improve convergence. Coefficients (β) depict estimated pairwise difference in performance between cross types within each treatment, the effect treatment within cross type, and estimated linear contrast in magnitude of the effect of treatment between cross types, obtained from a Tukey's test. Test statistics include the χ^2-value, and the marginal and conditional R^2 of the model. Model fits with significant (positive) intercept are indicated by †. χ^2-value with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Figures

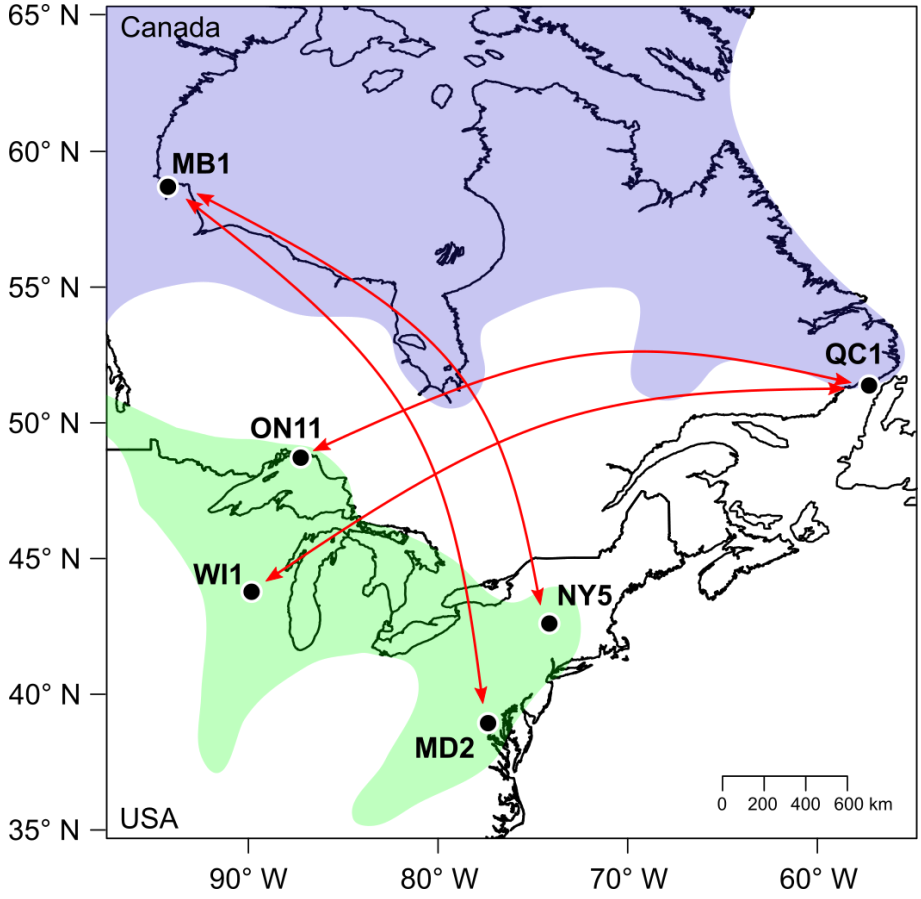

Figure 1: Distribution map of *Arabidopsis arenicola* and *A. lyrata* with the 6 populations studied. The blue shaded area represents the current range of *A. arenicola*, and the green shaded area represents the North-American range of *A. lyrata.* Black dots represent the position of each species. Population labels consist of the abbreviation for state (USA) or province (Canada) and a number. Red arrows represent the crosses performed between species.

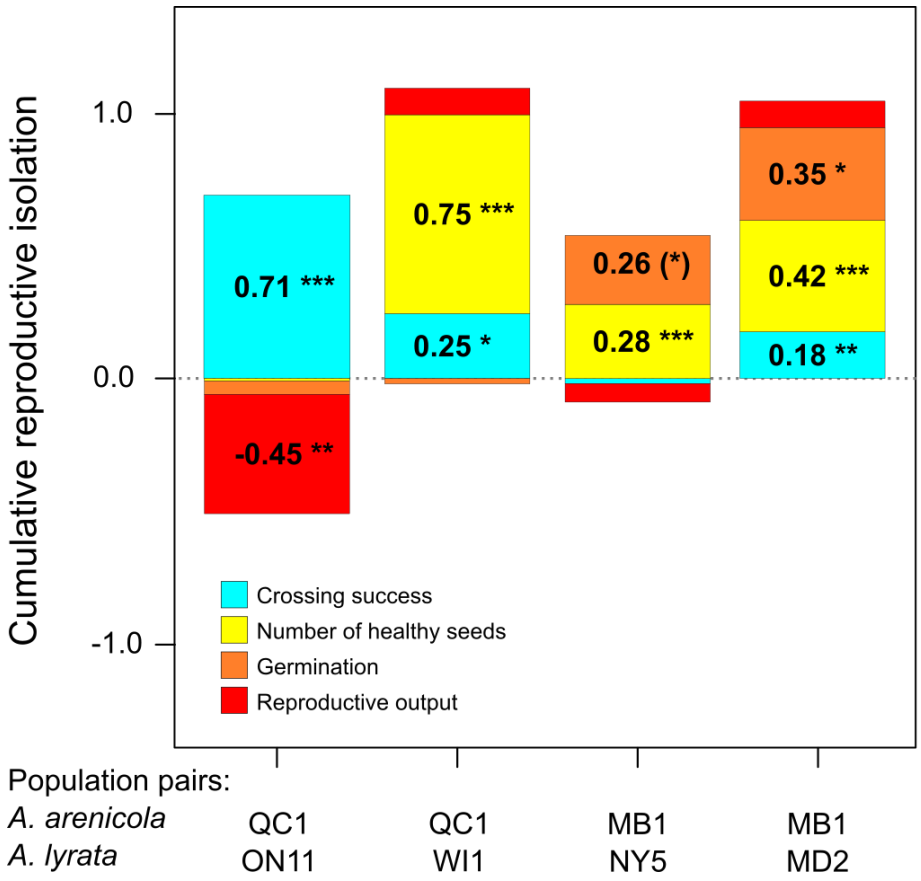

Figure 2: Cumulative reproductive isolation (RI) of each crossing pair between *A. arenicola* and *A. lyrata*. RI between pairs of populations (not distinguishing cross direction) of *A. arenicola* (QC1, MB1) and *A. lyrata* (ON11, WI1, NY5, MD2) was calculated based on four successive stages of within- and between-species crosses: crossing success (cyan) of each population pair, number of healthy seeds (yellow) produced by each successfully formed fruit, the germination (orange) rate of healthy seeds and the reproductive output (red) of successfully germinated individuals. The size of each bar represents the average RI of each stage, staked between positive RI (1 = complete isolation) and negative RI (-1 = outcrossing is favored). Numbers indicate average RI values significantly different from 0 (random mating, horizontal grey dashed line), tested in a one tailed Tuckey's test. Significance is indicated as: (∗) $P > 0.05$, ∗ $P < 0.05$, ∗∗ $P < 0.01$, ∗∗∗ $P < 0.001$. Test statistics are reported in Table 1.

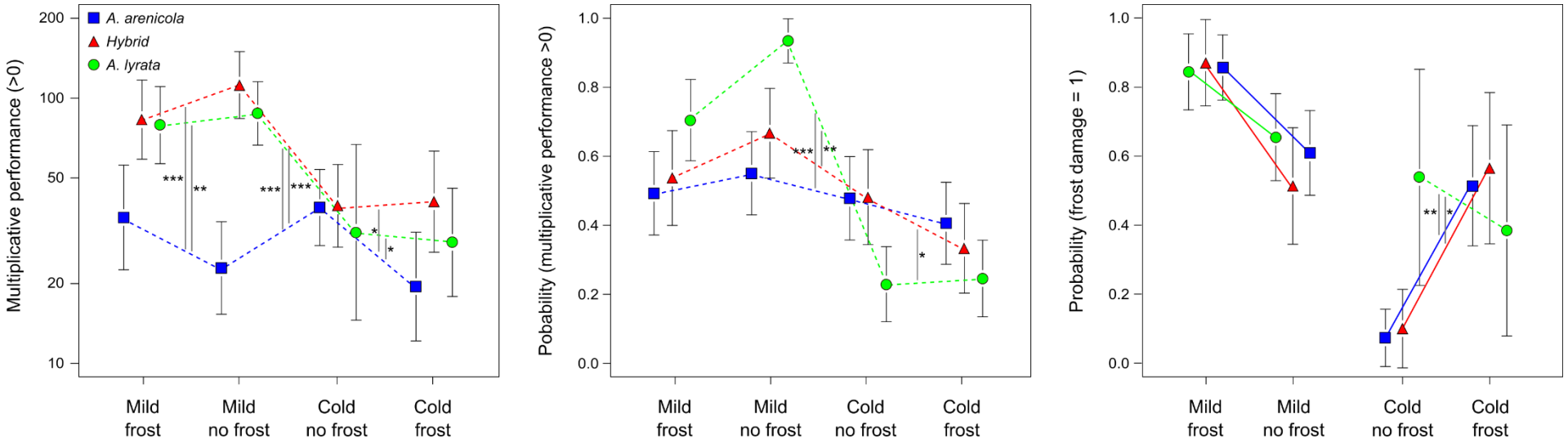

Figure 3: Variation in multiplicative performance and frost damage between *A. arenicola, A. lyrata* and their hybrids under different growth temperatures and frost treatments. Individuals of both species and their hybrids were raised in a climate chamber experiment under two growth temperatures: mild (20 °C) and cold (12 °C – 14 °C), and subjected to two treatments: no frost (control) and frost. Symbols depict multiplicative performance (**left** and **center**) or the probability of frost damage (**right**) averaged at the cross-type level, of within-species crosses of *A. arenicola* (blue squares) and *A. lyrata* (green circles), as well as their between-species hybrids (red triangle). Multiplicative performance was analyzed on the level of the individual modeling two processes: a Gaussian process (values > 0, **left**) and a logistic process (modelling the probability of values > 0, **center**). Tukey's tests were performed to test for differences between cross-type (not reported here) growth temperature or treatment, and the interaction between cross type and growth temperature or treatment. Horizontal bars represent the 95% confidence interval of the mean of each cross-type within each growth temperature and treatment combination. The significance of the effect of growth temperature or treatment within cross-type is represented by full ($P < 0.05$) or dashed ($P > 0.05$) colored lines connecting each species. Significant differences between slopes of effect of growth temperature or treatments between each cross-type are indicated by vertical grey lines, with significance indicated as: $*$ $P < 0.05$, $**$ $P < 0.01$, $***$ $P < 0.001$ Test statistics are reported in Table 2 and Table 3.

Supplementary information

Table S1: Population estimates

Population code	Species	Latitude [° N]	Longitude [° W]	Mean temperature late summer [° C] †	Cluster	Mating system
QC1	*A. arenicola*	51.43	-57.16	13.45	-	selfing
MB1	*A. arenicola*	58.78	-94.20	12.00	-	selfing
NY5	*A. lyrata*	42.66	-74.02	20.20	East	outcrossing
ON11	*A. lyrata*	48.77	-87.13	14.65	West	selfing
MD2	*A. lyrata*	38.99	-77.25	24.55	East	outcrossing
WI1	*A. lyrata*	43.83	-89.72	20.75	West	outcrossing

† July and August, WorldClim database version 2.0, Fick and Hijmans, 2017

Table S2: Growth conditions of the crossing experiment

Growth Phase	Duration [days]	Temperature daytime [°C]	Temperature nighttime [°C]	Day length [h]	Light intensity [µmol m^{-2} s^{-1}]
Stratification	12	4	4	0	0
Germination	22	20	20	8	100
Growth †	21	22	20	10	140
Flowering initiation	10	22	20	16	240
Flowering and crossing	205	22	20	16	240

† Day length and light intensity were gradually increased every three days by 1h and 20 µmol m-2 s-1, respectively.

Table S3: Cross combinations used in the climate chambers

Mother population	Father population	Cross Type	No. of cross combinations	No. of seeds sown			
				Mild		Cold	
				Control	Frost	Control	Frost
QC1	QC1	WSC	11	63	63	63	63
MB1	MB1	WSC	12	72	72	72	72
NY5	NY5	WSC	12	64	64	64	64
ON11	ON11	WSC	10	51	51	51	51
MD2	MD2	WSC	10	52	52	52	52
WI1	WI1	WSC	10	55	55	55	55
QC1	ON11	BSC	6	33	33	33	33
MB1	NY5	BSC	6	24	24	24	24
MB1	MD2	BSC	6	24	24	24	24
WI1	QC1	BSC	6	24	24	24	24
NY5	MB1	BSC	6	30	30	30	30
MD2	MB1	BSC	6	30	30	30	30

QC1(mother)xWI1 and ON11(mother)xQC1 did not produce viable seeds and where therefore not used in the climate chamber experiment

Table S4: Growth conditions and performance tracking of the climate chamber experiment

Simulated season	Growth phase	Duration [days]	Day length [h] *	Light intensity [µmol m^{-2} s^{-1}] *	Mild temperatures [°C]		Cold temperatures [°C]		Weekly record rate		
					day	night	day	night	Germination and survival	Damage	Bolting and Flowering
Fall	Stratification	20	0	0	4	4	4	4	-	-	-
	Germination	14	8	120	20	18	12	10	5	-	-
	Growth †	26	10	140	20	18 #	12	10 #	5	5	-
Winter	Acclimation ‡	7	10	180	20	18	12	10	3	3	-
	Vernalization	46	8	140	4	4	4	4	1	1	-
	Acclimation ‡	4	8	140	4	4	4	4	1	1	-
Spring and	Growth I §	18	10	160	20	18 #	14	12 #	2	1	4
Summer	Growth II §	58	16	220	20	18	14	12	1	1	3
	Flowering ȿ	46	16	220	20	18	16	14	1	1	1

* Values at the beginning of each phase

† Light intensity was gradually increased every six days by 20 µmol m-2 s-1

‡ Gradual decrease / increase of day length, light intensity, and temperatures

§ Light intensity was gradually increased every three days by 20 µmol m-2 s-1

ȿ The whole experiment was first performed in climate chambers (Climecab 1400, Kälte 3000 AG, Landquart, Switzerland), and was transferd for the last 46 days into growth chambers (MobyLux GroBanks, CLF Plant Climatics, plantclimatics.de, Wertingen, Deutschland) for logistical reasons. Temperatures of the Cold growth temperatures were increased due to technical limitations of this growth system

Nightly *frost* treatment was applied during two weeks before winter and during six weeks after winter. The temperature of night cycles was lowered to 4 °C for three nights to allow plant acclimation, then to -4°C for four nights to expose plants to frost. The cooling cycle started one hour after beginning of each night phase, and ended one hour before the beginning of the next day phase. Temperature declined gradually to reach the target temperature at the centre of the night phase. The target temperature was maintained for one hour, then temperature gradually increased back to the night temperature of each condition. This cycle was repeated two times for a total of 14 days before winter, and six times for a total of 42 days after winter, followed by an additional 9 nights at 4 °C.

Table S5: Description of performance estimates

Performance estimate	Level	Description
Multiplicative performance		
MP_{total}	Seed family	Crossing success * number of healthy seeds * germination * reproductive output
$MP_{climate\ chamber\ (CC)}$	Pot	Germination * reproductive output
Crossing experiment		
Crossing success	Seed family	Rate of crossing success (fruit elongation) over total number of crossing attempts
Number of healthy seed	Fruit	Number of healthy looking seeds per successful fruit
Climate chamber experiment		
Germination	Seed	Binary success of germination, from day 0 to the end of the experiment
Frost damage	Pot	Binary success of increase in damage under frost applied six weeks after vernalization
Reproductive output	Pot	Sum of all flower organs counted four weeks after the first flower opening

Appendix S6A: Parametrization of priors and hierarchical mixed-effects model analyzis in a Bayesian (MCMC) framework, with individual multiplicative performance as dependent variable

Priors

Priors were set to be weak, convergence was improved using parameter expansion. *R* is the priors for the fixed effects, *G* is the priors for the random effects.

```
priors.model=list(
 R=list(V=diag(2), n=1, fix = 2),
 G=list(G1=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2),
    G2=list(V=diag(2), n=4, alpha.mu = rep(0,4),alpha.V = diag(2)*25^2),
    G3=list(V=diag(2), n=2, alpha.mu = rep(0,2),alpha.V = diag(2)*25^2)))
```

Hierarchical mixed-effects models analyzed in a Bayesian (MCMC) framework

Multiplicative performance was split into two datasets: the *zero_part*, a binary transformation of performance estimates with *zero_part* = 1 if performance > 0, else *zero_part* = 0; and the *norm_part* containing only the $\log_{10}$ transformed performance measures if *zero_part* = 1.

```
model = MCMCglmm(cbind(norm_part, zero_part)
        ~ trait -1 + trait:cross type * trait: growth temperature,
        random = ~ us(trait):maternal population
        + us(trait): maternal population: maternal family
        + us(trait):block,
        prior = priors.model,
        rcov = ~idh(trait):units,
        family=c('gaussian', 'categorical'),
        burnin = 5000, thin = 100, nitt = 50000,
        data=data)
```

Appendix S6B: Parametrization of restricted maximum likelihood based hierarchical mixed-effect models

Binary dependent variable

Model = glmer(performance ~ cross type * treatment
+ (1 | maternal population / maternal family)
+ (1 | block),
family = “binomial”,
control = glmerControl(optimizer = "bobyqa"),
data = data)

Dependent variable with log-normal distribution

Model = lmer($\log_{10}$(performance) ~ cross type * growth temperature
+ (1 | maternal population / maternal family)
+ (1 | block),
control = lmerControl(optimizer = "bobyqa"),
data = data)

Table S7: Summary of ANOVA on model testing for variation in total and individual contributions to reproductive isolation (RI) between population pairs

		Population pair			
RI barriers	*N*	*F*		R^2	
Total RI					
Multiplicative performance	88	**4.11**	**	0.10	†
Individual RI barriers					
Cross success	90	**21.03**	***	0.40	†
Healthy seeds	73	**13.35**	***	0.34	
Germination	36	2.10		0.09	
Reproductive output	17	**6.76**	*	0.45	†

All dependent variables were assumed to follow Gaussian distributions. Test statistics include the *F*-value, and the adjusted R^2 of the model. Model fits with significant (positive) intercept are indicated by †. *F*-value with *P*-values < 0.05 are written in bold; significance is indicated: (*) *P*<0.1, * *P*<0.05, ** *P*<0.01, *** *P*<0.001.

Table S8: Summary of multiple comparisons testing for variation in total and individual contributions to reproductive isolation (RI) between population pairs

RI barriers	*N*	*QC1xON11 vs QC1xWI1* β	*QC1xON11 vs MB1xNY5* β	*QC1xON11 vs MB1xMD2* β	*QC1xWI1 vs MB1xNY5* β	*QC1xWI1 vs MB1xMD2* β	*MB1xNY5 vs MB1xMD2* β
Total RI							
Multiplicative performance	88	-0.27	0.00	**-0.36 ***	0.28	-0.09	**-0.36 *** †
Individual RI barriers							
Crossing success	90	**0.46 *****	**0.73 *****	**0.53 *****	**0.27 ***	0.07	-0.20 †
Number of healthy seeds	73	**-0.76 *****	-0.29 (*)	**-0.43 ****	**0.47 *****	**0.33 ***	-0.14
Germination	36	-0.03	-0.31	-0.41	-0.28	-0.38	-0.10
Reproductive output	17	**-0.55 ****	-0.38 (*)	‡	0.17	‡	‡ †

All dependent variables were assumed to follow Gaussian distributions. Coefficients (β) depict estimated pairwise difference in reproductive isolation between population pairs (not distinguishing cross direction), obtained from a Tukey's test. For RI based on reproductive output, the crossing pair MB1xMD2 (both cross directions) was not included in the analysis due to too low replication number, as few hybrids reached this life stage (‡). Model fits with significant (positive) intercept are indicated by †. Estimates with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Table S9A: Summary of model comparison on models testing for the effect of cross type (within –species crosses of *A. arenicola* populations [*A*] and northern *A. lyrata* populations [*L*], and their between-species hybrids [H]), growth temperature (cold compared to mild [0]) and their interaction on multiplicative performance of individuals raised in the climate chamber experiment

Model	Tested fixed effect	*DIC*	Δi
Cross type + temperature + (species* temperature)	-	315.3	-
Temperature + (cross type * temperature)	Cross type	315.4	-0.1
Cross type + (cross type * temperature)	Temperature	315.7	-0.4
Cross type + temperature	Cross type * temperature	397.3	-82.0

The dependent variable was *multiplicative performance* based on the climate chamber experiment. To test the contribution of each fixed effect to the fit of the full model, alternative models excluding each fixed effects were compared to the full model, with lower *DIC* values indicating a better fit. The difference between the full model and the others is indicated as Δi.

Table S9B: Summary of model comparison on models testing for the effect of cross type (within –species crosses of *A. arenicola* populations [*A*] and northern *A. lyrata* populations [*L*], and their between-species hybrids [H]), treatment (frost compared to control [0]) and their interaction on multiplicative performance of individuals raised in the climate chamber experiment, in the mild growth temperature

Model	Tested fixed effect	*DIC*	Δi
Cross type + treatment + (species* treatment)	-	394.8	-
Treatment + (cross type * treatment)	Cross type	394.5	0.3
Cross type + (cross type * treatment)	Treatment	394.7	0.1
Cross type + treatment	Cross type * treatment	406.0	-11.2

The dependent variable was *multiplicative performance* based on the climate chamber experiment. To test the contribution of each fixed effect to the fit of the full model, alternative models excluding each fixed effects were compared to the full model, with lower *DIC* values indicating a better fit. The difference between the full model and the others is indicated as Δi.

Table S9C: Summary of model comparison on models testing for the effect of cross type (within –species crosses of *A. arenicola* populations [*A*] and northern *A. lyrata* populations [*L*], and their between-species hybrids [H]), treatment (frost compared to control [0]) and their interaction on multiplicative performance of individuals raised in the climate chamber experiment, in the cold growth temperature

Model	Tested fixed effect	*DIC*	Δi
Cross type + treatment + (species* treatment)	-	389.4	-
Treatment + (cross type * treatment)	Cross type	389.4	0.0
Cross type + (cross type * treatment)	Treatment	389.6	-0.2
Cross type + treatment	Cross type * treatment	395.3	-5.9

The dependent variable was *multiplicative performance* based on the climate chamber experiment. To test the contribution of each fixed effect to the fit of the full model, alternative models excluding each fixed effects were compared to the full model, with lower *DIC* values indicating a better fit. The difference between the full model and the others is indicated as Δi.

Table S9D: Summary of ANOVA on models testing for the effect of cross type (within –species crosses of *A. arenicola* populations [*A*] and northern *A. lyrata* populations [*L*], and their between-species hybrids [H]), treatment (*frost* compared to *control* [0]) and their interaction on frost damage of individuals grown under mild or cold growth temperatures

		Cross type (CT)	*Treatment (T, frost vs control)*	*CT * T*			
Dependent variable	*N*	χ^2	χ^2	χ^2	R^2m	R^2c	
Frost damage (mild)	291	1.82	**8.45 ****	1.01	0.13	0.14	
Frost damage (cold)	155	**13.12 ****	**14.22 *****	**11.24 ****	0.32	0.32	†

Frost damage was a binary variable, therefore analyzed by a model predicting non-zeros on the logit scale. Models were performed on two subsets, considering individuals raised either under mild or cold growth temperatures. Each model was optimized with the *bobyqa* optimizer to improve convergence. Test statistics include the χ^2-value, and the marginal and conditional R^2 of the model. Model fits with significant (positive) intercept are indicated by †. χ^2-value with *P*-values < 0.05 are written in bold; significance is indicated: (*) $P<0.1$, * $P<0.05$, ** $P<0.01$, *** $P<0.001$. Results for random effects are not shown.

Conclusions

In this thesis, I empirically assessed the role of demographic, evolutionary and ecological factors in shaping range limits of North American *Arabidopsis lyrata* subsp *lyrata*. By performing a crossing experiment followed by a transplant experiment across and beyond the distribution of this species, I tested the dynamics of the expression of mutational load and local adaptation in small populations at range limits. I further assessed the factors conditioning successful post-glacial range expansion by testing the differences in adaptive strategy to cold climates and the reproductive isolation between *A. lyrata* and its sister species *A. arenicola,* latter colonizing much higher latitudes despite both species sharing a common post-glacial origin.

In the first subproject of this thesis, I showed that the expression of mutational load was higher in populations with increased genomic estimates of mutational load, with a longer history of range expansion or isolation, and a selfing mating system (**Chapter 1)**. Overall, the accumulation of mutational load led to a decline of 80% in performance. These findings suggest that mutation accumulation as a result of post-glacial range expansion, long-term isolation and a switch to a selfing mating system affect performance to magnitudes sufficient to cause range limits, supporting prediction from previous simulation studies (Peischl et al., 2013; Peischl and Excoffier, 2015). These results are also in line with patterns of mutational load reported on genomic data (Willi et al., 2018), and similar phenotypic data from greenhouse or common garden studies (Willi et al., 2018; Koski et al., 2019). I further showed that the expression of mutational load increases under environmental stress (**Chapter 2**), as has been suggested for inbreeding depression (Reed et al., 2012). This effect was especially strong in populations exposed to warmer climates than at their site of origin. This effect was independent of populations' genomic estimates of mutational load, indicating that small populations suffer from two genetic Allee effects: one directly linked to the increase in mutational load, and one linked to the increase in the expression of mutational load under adverse environmental conditions, independent of the magnitude of load. Climatic conditions immediately beyond this species range limits did not trigger this second genetic Allee effect, suggesting this effect was

dependent on high stress intensity and could be triggered by extreme climatic events, more frequent in unsuitable habitat beyond range limits. Finally, I showed that population performance declines in the common garden site beyond the southern range limits of this species, suggesting that its southern range limits reflect its ecological niche (**Chapter 3**). I found a general signature of local adaptation to the most niche defining variables in this species, however this signature was reduced in population of small size and reduced genomic diversity, suggesting drift contributes in shaping this species distribution limits by reducing adaptation in population of small size, mostly located at range limits. These findings are in line with previous simulation studies identifying drift as a main factor constraining adaptation along environmental gradients (Polechová and Barton, 2015; Polechová, 2018)

In the second subproject of this thesis I found that *A. lyrata* and its northern sister species *A. arenicola* diverged in their adaptive strategies to cold climates (**Chapter 4**): *A. lyrata* was more tolerant and resistant to frost events in early snow melt in late winter, but less tolerant to cold growth temperatures and less tolerant to late frost events in summer. On the contrary, *A. arenicola* showed increased tolerance to cold growth temperatures and higher tolerance to late frost events in summer, but reduced frost tolerance and resistance in winter, suggesting that the colonization of *A. lyrata* toward higher latitudes could have been constrained by reduced frost tolerance in summer. I also showed that both species were isolated by intrinsic reproductive barriers (**Chapter 5**), mainly preventing successful crossing and the production of viable seeds. In addition, hybridization disrupted the adaptation strategies of both species, indicating that intraspecific gene flow could be maladaptive in both species' natural habitats.

My thesis revealed three factors at play in shaping range limits of *A. lyrata*. First, the expression of mutational load increased toward range limits, leading to a strong decline in performance. This expression was also increased under strong environmental stress, which could be more frequent beyond range limits. The magnitude of both effects suggests that these might be strong enough to constrain further range expansion or isolation at both range edges, with a stronger effect in

unsuitable habitats. Second, adaptation to precipitation seemed impaired by drift, especially at the southern edge. At the northern edge, populations seemed less affected by conditions immediately outside of the niche limits, however further range expansion could have been halted due the adaptive strategy to cold climates evolved in northern populations: while this strategy is well suited for mid-latitudes, it would expose these populations to high frost stress at higher latitudes. Overall, southern range limits could be shaped by the accumulation of mutational load, its increased expression under degrading (warming) climate, while northern range limits could have been shaped by the accumulation of mutational load and maladaptation to climates at higher latitudes. Besides evolving a suitable adaptation strategy to subarctic and arctic climate, it remains an open question how *A. arenicola* could escape the negative effects of drift during range expansion. Previous simulation studies have predicted that only locally adapted populations can persist under strong expansion load (Gilbert et al., 2017), opening new perspective to explore the factors allowing or constraining fast post-glacial range expansion toward high latitudes.

In a broader context, my work reflects the complexity of factors which can play a role in shaping species distributions, and which could be acting simultaneously one species. The demographic and evolutionary factors shaping the range limits of *A. lyrata* could apply on many species with a history of post-glacial range expansion (Hewitt, 2000, 2004) and a history of small population size at range limits (Pironon et al., 2017). However, my results also show that the effect of these factors are not absolute, and that under certain conditions species having undergone post-glacial range expansion can escape the constraints exerted by drift. The results of my thesis therefore point toward a greater need of considering the effects of mutational load and the environmental dependency of the expression of mutational load in in simulation studies of colonization in empty habitats (Polechová and Barton, 2015; Polechová, 2018), and a general integration of demographic history and the role of drift in empirical studies in the context of range limits (Willi and Van Buskirk, 2019).

The findings of my thesis also support the call for greater implementation of demographic and evolutionary factors in biodiversity conservation and management (Hampe and Petit, 2005; Hoffmann et al., 2015): the negative effects of environmental stress generated by the rarefaction of suitable habitats through anthropogenic activities, and at large climate change, could be exacerbated by the increased expression of mutational load in small and isolated populations exposed to high drift. Furthermore, my results raise concern on the overall fate of species under climate change. On the short term, the already sensitive small populations at the warmer edge of a species' distribution (Hampe and Petit, 2005) might face the strongest changes in environmental conditions (Lenoir and Svenning, 2013), and could go extinct much faster due to increased expression of mutational load and lower adaptation. In the long term, further isolation due to habitat degradation could lead to extinction at the rear edge, and further colonization could be impeded at the leading edge. If increased genetic drift is a common factor shaping range limits, many species could face strong range contractions rather than range shifts (Chen et al., 2011; Lenoir and Svenning, 2013) in the next decades. Evolutionary and demographic factors are increasingly incorporated in simulation studies predicting species distribution dynamics under climate change (Bush et al., 2016; Wang et al., 2018), but the environmental dependency of mutational load has yet to be integrated in such studies

References

Bush, A., K. Mokany, R. Catullo, A. Hoffmann, V. Kellermann, C. Sgrò, et al. 2016. Incorporating evolutionary adaptation in species distribution modelling reduces projected vulnerability to climate change. *Ecology Letters* 19: 1468–1478.

Chen, I. C., J. K. Hill, R. Ohlemüller, D. B. Roy, and C. D. Thomas. 2011. Rapid range shifts of species associated with high levels of climate warming. *Science* 333: 1024–1026.

Gilbert, K. J., N. P. Sharp, A. L. Angert, G. L. Conte, J. A. Draghi, F. Guillaime, et al. 2017. Local adaptation interacts with expansion load during range expansion: maladaptation reduces expansion load. *The American Naturalist* 189: 368–380.

Hampe, A. and R. J. Petit. 2005. Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters* 8: 461–467.

Hewitt, G. M. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907–913.

Hewitt, G. M. 2004. Genetic consequences of climatic changes in the Quaternary. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 359: 183–195.

Hoffmann, A., P. Griffin, S. Dillon, R. Catullo, R. Rane, M. Byrne, et al. 2015 A framework for incorporating evolutionary genomics into biodiversity conservation and management. *Climate Change Responses* 2: 1–23.

Koski, M. H., N. C. Layman, C. J. Prior, J. W. Busch, and L. F. Galloway. 2019. Selfing ability and drift load evolve with range expansion. *Evolution Letters* 3: 500–512.

Lenoir, J., and J. C. Svenning. 2013. Latitudinal and elevational range shifts under contemporary climate change. *In* S. A. Levin, eds. *Encyclopedia of Biodiversity, second edition.* Academic Press. Encyclopedia of Biodiversity, 599–611.

Peischl, S., I. Dupanloup, M. Kirkpatrick, and L. Excoffier. 2013. On the accumulation of deleterious mutations during range expansions. *Molecular Ecology* 22: 5972–5982.

Peischl, S., M. Kirkpatrick, and L. Excoffier. 2015. Expansion load and the evolutionary dynamics of a species range. *The American Naturalist* 185: E81–E93.

Pironon, S., G. Papuga, J. Villellas, A. L. Angert, M. B. García, and J. D. Thompson. 2017. Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biological Reviews* 92: 1877–1909.

Polechová, J. 2018. Is the sky the limit? On the expansion threshold of a species range. *PLoS Biology* 16: e2005372.

Polechová, J., and N. H. Barton. 2015. Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences of the United States of America* 112: 6401–6406.

Reed, D. H., C. W. Fox, L. S. Enders, and T. N. Kristensen. 2012. Inbreeding–stress interactions: evolutionary and conservation consequences. *Annals of the New York Academy of Sciences* 1256: 33-48.

Wang, W. J., F. R. Thompson, H. S. He, J. S. Fraser, W. D. Dijak, and M. A. Spetich. 2018. Population dynamics has greater effects than climate change on tree species distribution in a temperate forest region. *Journal of Biogeography*, 45: 2766– 2778.

Willi Y., and J. Van Buskirk. 2019. A practical guide to the study of distribution limits. *The American Naturalist* 193: 773–785.

Willi, Y., M. Fracassetti, S. Zoller, and J. Van Buskirk. 2018. Accumulation of mutational load at the edges of a species range. *Molecular Biology and Evolution* 35: 781–791.