



**HAL**  
open science

# Développement de nouveaux modèles éco-évolutifs pour évaluer les impacts démographiques des flux de gènes assistés dans le contexte des changements climatiques

Adèle Erlichman

## ► To cite this version:

Adèle Erlichman. Développement de nouveaux modèles éco-évolutifs pour évaluer les impacts démographiques des flux de gènes assistés dans le contexte des changements climatiques. *Ecologie, Environnement*. Université de Montpellier, 2024. Français. NNT : 2024UMONG005 . tel-04697452

**HAL Id: tel-04697452**

**<https://theses.hal.science/tel-04697452>**

Submitted on 13 Sep 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Sciences de l'Évolution et de la Biodiversité

École doctorale GAIA

Unité de recherche : Institut des Sciences de l'Évolution de Montpellier (ISEM, UMR 5554)

Développement de nouveaux modèles éco-évolutifs pour  
évaluer les impacts démographiques des flux de gènes  
assistés dans le contexte des changements climatiques

Présentée par Adèle ERLICHMAN  
Le 22 avril 2024

Sous la direction d'Ophélie RONCE  
et co-encadrée par Sarah P. OTTO

Devant le jury composé de

Sylvie ODDOU-MURATORIO, Directrice de recherches, INRAE, ECOBIOP, Saint-Pée sur Nivelle

Bruno COLAS, Professeur, AgroParisTech, ESE, Gif-sur-Yvette

Janneke HILLE RIS LAMBERS, Professeure, ETH, Institut de Biologie Intégrative, Zurich

Sébastien LION, Directeur de recherches, CNRS, CEFE, Montpellier

Jean-Baptiste MIHOUB, Maître de conférences, Sorbonne Université, Paris

Ophélie RONCE, Directrice de recherches, CNRS, ISEM, Montpellier

Sarah P. OTTO, Professeure, Dép. de Zoologie, Université de Colombie-Britannique, Vancouver

Rapportrice

Rapporteur

Examinatrice

Président du Jury

Examineur

Directrice de thèse

Encadrante de thèse



UNIVERSITÉ  
DE MONTPELLIER

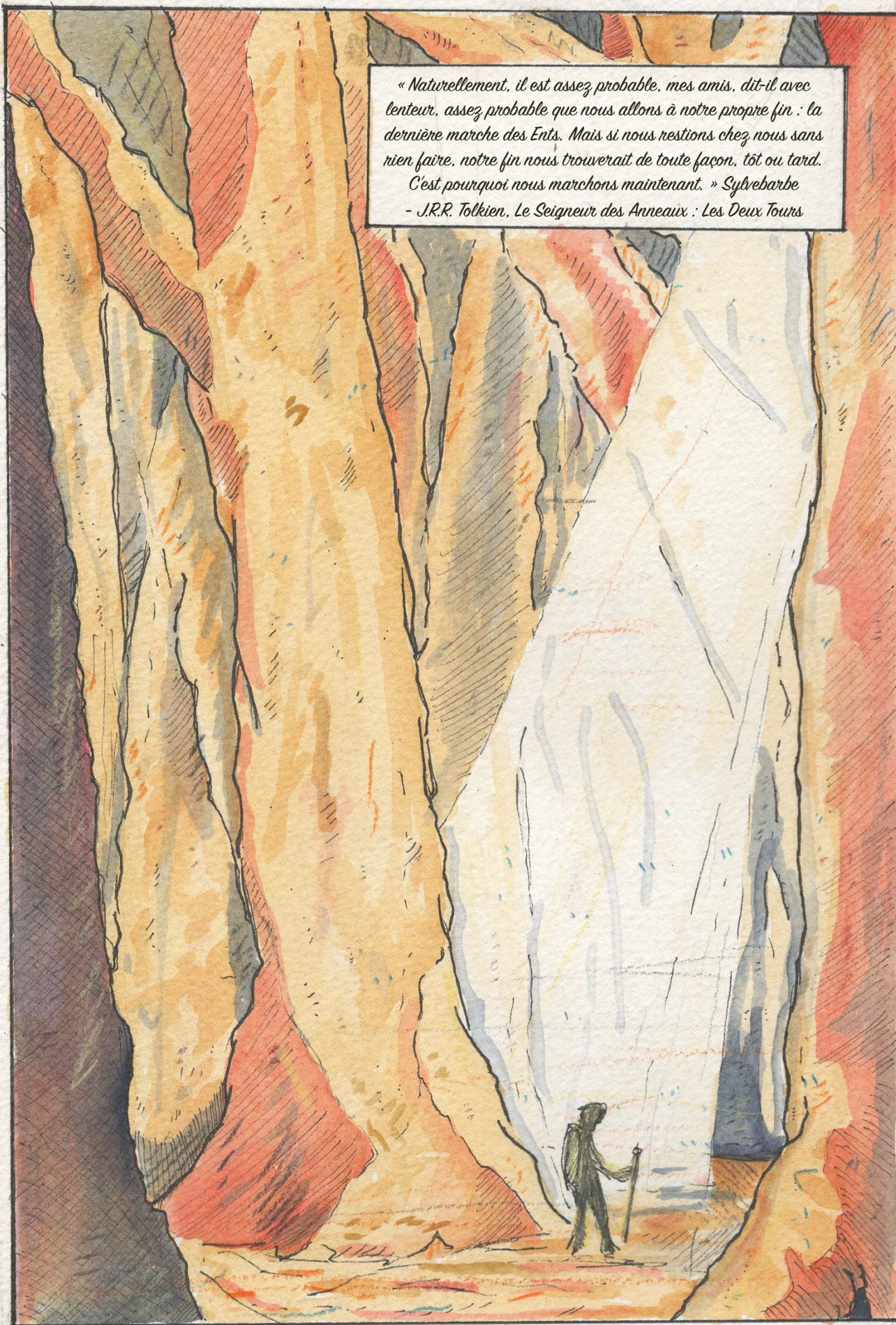
# Abstract

Biodiversity is facing an unprecedented extinction crisis due to anthropogenic pressures, such as climate change. There is growing concern about the slow spontaneous pace of evolutionary processes relative to the high speed of environmental changes. Assisted gene flow has been introduced as the managed translocation of individuals within the historical range of a species. Its purpose is to introduce, or increase, the frequency of genotypes expected to confer an advantage under new or future climatic conditions. It can also increase genetic variation in targeted populations in the hope that it can stimulate demography through heterosis and, by increasing the genetic variation on which selection can act on, the ability to adapt to a changing environment. Risks associated with assisted gene flow include pathogen introduction, outbreeding depression, genomic swamping, and increased maladaptation. Guidelines and information on optimal strategies and the uncertainties associated with these practices are therefore urgently needed. Using both analytical predictions and simulations, we have developed demo-evolutionary models to help identify critical parameters and sources of uncertainty when implementing these strategies. We first show the importance of taking into account the life cycle of targeted species, as the optimal choice of individuals to translocate in long-lived species facing a changing climate is subject to a trade-off between being well adapted at the beginning and end of life. We also propose a new method for integrating evolution into an integral projection model (IPM) and show how only a small range of introduced genetic diversity for adaptive traits would allow a rare and endangered monocarpic perennial plant, *Centaurea corymbosa*, to escape extinction in a warming climate. Lastly, we show with an analytical model that if beneficial alleles are introgressed through managed gene flow between genetically distinct populations or closely related species, small and early introductions of non local genetic material minimize swamping of the local population's genome and maximize the probability of rescue in small declining populations. Overall, these new models offer new perspectives for guiding assisted gene flow strategies and help inform debates about the relevance of their use and optimal implementation.

# Résumé

La biodiversité est confrontée à une crise d'extinction sans précédent en raison de pressions d'origine anthropique, telles que le changement climatique. La lenteur des processus évolutifs spontanés, par rapport à la rapidité des changements environnementaux, suscite de grandes inquiétudes quant à la persistance de nombreuses espèces. En réponse à ces menaces, les flux de gènes assistés sont présentés comme la translocation gérée d'individus au sein de leur aire de répartition historique. Leur objectif est d'introduire ou d'augmenter la fréquence des génotypes susceptibles de conférer un avantage dans des conditions climatiques nouvelles ou futures. Ils peuvent également accroître la variation génétique dans les populations ciblées, dans l'espoir de stimuler la démographie par l'hétérosis et, en augmentant la variation génétique sur laquelle la sélection peut agir, la capacité à s'adapter à un environnement changeant. Les risques associés aux flux de gènes assistés peuvent inclure l'introduction de pathogènes, la dépression hybride, l'envahissement génétique ou encore l'augmentation de la maladaptation. Il existe donc un besoin urgent de lignes directrices et d'études des stratégies optimales et des incertitudes associées à ces pratiques. Nous développons des modèles démo-évolutifs pour aider à identifier les paramètres critiques et les sources d'incertitude lors de la mise en œuvre de ces stratégies. Nous montrons tout d'abord, à l'aide d'un modèle structuré en stade, l'importance de la prise en compte du cycle de vie des espèces ciblées, car le choix optimal des individus à transférer chez les espèces à longue durée de vie confrontées à un climat changeant dépend d'un compromis entre une bonne adaptation au début et à la fin de la vie. Nous proposons également une nouvelle méthode pour intégrer l'évolution dans un modèle de projection intégrale (IPM) et montrons comment seule une petite gamme de diversité génétique introduite pour des traits adaptatifs permettrait à une plante rare et menacée, *Centaurea corymbosa*, d'échapper à l'extinction dans un climat qui se réchauffe. Enfin, nous montrons à l'aide d'un modèle analytique que si des allèles bénéfiques sont introduits entre des populations génétiquement distinctes ou des espèces étroitement apparentées, de petites introductions précoces de matériel génétique non local minimisent l'envahissement du génome de la population locale et maximisent la probabilité de sauvetage de petites populations en déclin. Ces nouveaux modèles offrent de nouvelles perspectives pour guider les stratégies de flux de gènes assistés et permettent d'éclairer les débats sur la pertinence de leur utilisation et leur mise en œuvre optimale.

*« Naturellement, il est assez probable, mes amis, dit-il avec lenteur, assez probable que nous allons à notre propre fin : la dernière marche des Ents. Mais si nous restions chez nous sans rien faire, notre fin nous trouverait de toute façon, tôt ou tard. C'est pourquoi nous marchons maintenant. » Sylvebarbe*  
- J.R.R. Tolkien, *Le Seigneur des Anneaux : Les Deux Tours*



# Acknowledgments / Remerciements

Ces années de doctorat sont passées à toute vitesse, et pourtant, la route a été longue. J'aimerais remercier tous ceux qui m'ont aidée à mener à bien cette thèse, en me soutenant au cours des trois dernières années (et demie !).

Ces remerciements vont être un mélange de français et d'anglais, et je m'en excuse auprès de ceux qui ne peuvent pas tout comprendre, même si finalement, je trouve que c'est une belle représentation de cet échange entre Montpellier et Vancouver. Cela m'amène à remercier le Fond France-Canada pour la Recherche (FCRF) qui a financé cette thèse et ces échanges.

I have been lucky enough to be mentored by a lot of incredible women during this PhD, and I think this is important enough to emphasize. It has been an honor to learn alongside them.

J'aimerais commencer par remercier Ophélie. J'avais tant à apprendre, et Ophélie m'a beaucoup appris, avec passion et patience. Je lui suis infiniment reconnaissante de m'avoir guidée et soutenue pendant ces années. Merci d'avoir été si disponible, malgré certaines circonstances (essayez de commencer votre thèse pendant un confinement !). Je n'aurais pas pu souhaiter une meilleure directrice de thèse !

Je voudrais également remercier Sally. Thank you, Sally, for your kindness and thoughtfulness. Thank you for making me feel so welcome in Vancouver, and for teaching me so much from our time together.

I only met Linnea Sandell once in person, but I'm sure we spent hours zoom-ing together. I learned so much from our exchanges, starting my PhD in Montpellier while she was finishing hers in Vancouver. Working with her has always been fun and joyful, and I have always looked forward to our meetings.

I would also like to thank Sally Aitken and the UBC Forestry Lab for their warm welcome to Vancouver. Sally, thank you for your valuable feedback on my first chapter.

Finally, I would like to thank the people I was lucky enough to meet in Vancouver. They made me feel at home, even though I was on the other side of the world. I will never forget the beers at sunset on Wreck Beach.

Merci également à toute l'équipe EvoDemo pour sa gentillesse au fil des années, et une

mention spéciale à Juliette, car je pense que nous avons essayé de nous soutenir mutuellement au cours de nos thèses.

Je tiens également à remercier les membres de mon comité de thèse: Olivier Gimenez, Luis-Miguel Chevin, Hélène Fréville, Alexandre Robert, Eric Imbert et Sandrine Maurice, qui m'ont soutenue et aidée par leurs remarques constructives tout au long de ce doctorat, toujours avec bienveillance. J'étends également ces remerciements aux membres du jury, qui ont accepté de lire et d'évaluer ce travail. Merci pour les échanges stimulants et enrichissants que nous avons eus lors de ma soutenance.

Enfin, je tiens à remercier ma famille et mes amis, pour leur soutien et leur affection, qui ont grandement contribué à la réussite de cette thèse. Mes parents, en particulier, qui m'ont toujours encouragée à poursuivre ce que j'aimais et m'ont donné les moyens de le faire.

Je souhaiterais aussi remercier mes amies : Marie, Sabrina, Justine, Mathilde et Violette qui m'ont soutenue tout au long de ces années. Merci aussi à Oscar, qui m'a fait sortir à Montpellier de temps en temps pour me changer les idées, et que je remercie également pour l'illustration de cette thèse.

Finalement, je souhaite remercier tout particulièrement Nicolas, qui a été mon soutien le plus important pendant ce doctorat. Il a été avec moi dans tous les moments de joie et de difficulté, même à distance, m'encourageant toujours à aller de l'avant et célébrant même les plus petits pas...

# Table of contents

<b>Résumé de la thèse en français</b>	<b>4</b>
<b>General introduction</b>	<b>11</b>
Loss of biodiversity . . . . .	11
Population responses to environmental pressures . . . . .	11
Human-assisted gene flow . . . . .	12
Examples of assisted gene flow . . . . .	13
Different challenges faced by different species . . . . .	13
Contributions from previous models . . . . .	14
Aims of this thesis . . . . .	15
<b>1 Planting long-lived trees in a warming climate: Theory shows the importance of stage-dependent climatic tolerance</b>	<b>17</b>
Supplementary Material 1 . . . . .	46
Appendix A: Estimation of the thermal tolerance of trees . . . . .	57
<b>2 Evolutionary rescue in a rare endemic Mediterranean plant: <i>Centaurea corymbosa</i></b>	<b>60</b>
Supplementary Material 2 . . . . .	79
Appendix B: R code for the evolutionary IPM . . . . .	83
<b>3 Evolutionary rescue by means of introgression: balancing genomic rescue and swamping</b>	<b>92</b>
Supplementary Material 3 . . . . .	107
<b>General discussion</b>	<b>117</b>
Relevance of assisted gene flow practices . . . . .	117
The importance of models . . . . .	118
Many uncertainties still surround these practices . . . . .	120
<b>Appendix C: publication from Master’s thesis</b>	<b>123</b>
<b>Bibliography</b>	<b>136</b>



# Résumé de la thèse en français

## Contexte

La biodiversité est confrontée à une crise d’extinction caractérisée par des taux d’extinction comparables à ceux des cinq grandes crises d’extinction massive (Barnosky et al. 2011; Dirzo et al. 2014). La crise actuelle est en grande partie le résultat de pressions anthropiques, notamment la surexploitation des écosystèmes, les activités agricoles, le développement urbain et le changement climatique (Maxwell et al. 2016a). Ces pressions entraînent une diminution de la taille des populations et de la diversité génétique disponible, ce qui accroît leur vulnérabilité aux facteurs de stochasticité et d’extinction associés aux populations de petite taille (Pimm et al. 2006b; Shaffer 1981). Le changement climatique (IPCC 2021) expose les populations à des événements extrêmes plus fréquents et plus intenses (*e.g.*, canicules, incendies, fortes pluies ; Mitchell et al. 2006) et perturbe progressivement, lorsqu’elle existe, l’adaptation locale des populations à leurs conditions environnementales historiques (Aitken and Bemmels 2016; Aitken and Whitlock 2013; Brady et al. 2019; Crespi 2000). À moins que les populations ne puissent suivre activement leur niche climatique, la maladaptation liée au changement climatique peut déclencher ou aggraver le déclin des populations (Duffy et al. 2022; Román-Palacios and Wiens 2020). Les conséquences de la diminution du nombre d’espèces, et de leur extinction, sont critiques pour les écosystèmes, en particulier lorsque les espèces touchées sont des espèces dites “clés de voûte”, telles que les arbres forestiers (Aitken and Whitlock 2013; Schmid et al. 2009).

## Réponses des populations aux pressions environnementales

Face au changement climatique, les populations doivent répondre pour éviter l’extinction. *In situ*, la plasticité phénotypique pourrait atténuer les effets de la maladaptation jusqu’à ce qu’une réponse adaptative soit possible (Anderson et al. 2012; Chevin 2015; Chevin et al. 2010). Toutefois, il n’est pas encore clair que la plasticité permettrait d’éviter l’extinction, quelle forme

prendrait la réponse plastique et quels seraient ses effets sur les réponses évolutives (Diamond and Martin 2021a; Gienapp et al. 2008; Parmesan 2006). La sélection naturelle pourrait conduire à une adaptation à condition qu'un génotype/phénotype adaptatif augmente en fréquence dans la population à la suite du changement environnemental. Si cette évolution est suffisamment rapide pour permettre à la population de conserver une taille importante et des perspectives de persistance élevées, ce processus est appelé sauvetage évolutif (Carlson et al. 2014; Gomulkiewicz and Holt 1995). Des études suggèrent toutefois que de nombreuses espèces n'évoluent pas assez rapidement pour suivre le rythme et l'ampleur actuels des changements environnementaux (*e.g.*, Kremer et al. 2012; Radchuk et al. 2019).

La dispersion est une autre forme de réponse possible face aux contraintes environnementales. Bien qu'il soit prouvé que les populations déplacent leurs aires de répartition vers les pôles et/ou les altitudes plus élevées pour suivre leur optimum climatique (Chen et al. 2011; Freeman et al. 2018) et que cela réduit les risques d'extinction (Román-Palacios and Wiens 2020), cela peut toutefois être limité par des barrières environnementales comme la fragmentation des habitats (Leimu et al. 2010). Les caractéristiques du cycle de vie de certaines espèces, telles qu'une nature sessile (*e.g.*, plantes, coraux, etc.) ou un long temps de génération, empêcheront très probablement ces espèces d'atteindre la vitesse de migration nécessaire pour persister (Ettinger and HilleRisLambers 2013; Shaw and Etterson 2012).

### **Flux de gènes assistés par l'Homme**

La crise actuelle conduit à des appels à des actions de conservation pour améliorer la condition des populations en déclin et accélérer leur adaptation à un climat changeant. En conséquence, les translocations assistées par l'homme ont été proposées pour aider les populations à se maintenir, à se rétablir et à s'adapter au changement climatique (IUCN/CSE, 2013). Il existe un continuum entre les différents types d'intervention en ce qui concerne l'endroit où les individus sont déplacés géographiquement (décrit dans la figure 1 de Aitken and Bemmels 2016, Seddon 2010), allant de (i) mesures de renforcement des populations *in situ*, visant à stimuler un sauvetage génétique des populations, en augmentant le nombre d'individus et la diversité génétique (Carlson et al. 2014) à (ii) des stratégies de migration assistée *i.e.*, la translocation de génotypes pré-adaptés en prévision ou en réponse à des changements environnementaux (Aitken and Bemmels 2016; Aitken and Whitlock 2013). Ces dernières peuvent prendre la forme d'une migration assistée en dehors de l'aire de répartition historique de l'espèce, parfois appelée "colonisation assistée", ou

d'une migration assistée au sein de l'aire de répartition actuelle d'une espèce également souvent appelée "flux de gènes assistés", bien que cette terminologie prête quelque peu à confusion car toutes les pratiques décrites ci-dessus décrivent des flux de gènes, assistés par l'Homme. La migration assistée en dehors de l'aire de répartition des espèces est très controversée en raison des risques qui y sont associés, tels que la maladaptation, l'hybridation, l'introduction d'agents pathogènes et les invasions biologiques (voir par exemple Hunter 2007; Ricciardi and Simberloff 2009). Le grand public est aussi généralement plus ouvert à des pratiques de migration assistée au sein de l'aire de répartition des espèces, qu'en dehors (Peterson St-Laurent et al. 2018). Les risques, cités ci-dessus, sont considérés comme plus faibles dans le cas de flux de gènes assistés entre des populations existantes, à l'exception de certains risques génétiques tels que la dépression hybride, l'envahissement génomique et les incompatibilités génétiques (Aitken and Whitlock 2013; Frankham et al. 2011; Muller 1942; Sáenz-Romero et al. 2021).

### **Exemples de flux de gènes assistés**

Le flux de gènes assistés peuvent être mis en œuvre chez un large éventail de taxons, animaux ou végétaux, mais ils sont principalement envisagés pour des espèces fondatrices (*e.g.*, les coraux), les espèces gérées (*e.g.*, les arbres) et les petites populations menacées (Aitken and Whitlock 2013). Les règles de transfert, qui définissent les meilleures pratiques lors de la sélection des sources de matériel biologique, mentionnent de plus en plus la ressemblance avec le climat futur comme un critère de sélection, par exemple, pour la replantation des arbres après récolte, la restauration ou le renforcement des populations, plutôt que l'utilisation exclusive de sources locales (Breed et al. 2018; O'Neill et al. 2008, 2017). Des appels clairs sont également lancés dans cette direction chez plusieurs taxons, comme dans les coraux (*e.g.*, Hagedorn et al. 2021), certains poissons (*e.g.*, Pavlova et al. 2017; Pregler et al. 2023), des mammifères (*e.g.*, Seddon and Schultz 2020) et les arbres (*e.g.*, Borrell et al. 2020; Browne et al. 2019; Girardin et al. 2021; Milesi et al. 2019; Young et al. 2020). De plus, bien que cet effet ne semble pas résulter d'une politique active d'adaptation au changement climatique, une étude récente a montré que les translocations de plantes étaient plus souvent effectuées vers des sites légèrement plus froids (Diallo et al. 2021), ce qui montre que ces interventions sont déjà en cours.

## Différentes difficultés rencontrées par différentes espèces

En fonction des défis auxquels sont confrontées les différentes espèces ciblées par les flux de gènes assistés, un large éventail de pratiques peut être envisagé. À l'origine, les flux de gènes assistés sont apparus, comme une stratégie visant à aider les populations à s'adapter au changement climatique, mais de nombreuses translocations ne sont pas liées au climat (Diallo et al. 2021; Kelly and Phillips 2016). Les "flux de gènes ciblés", par exemple, se réfère au mouvement de variants d'intérêt, d'une manière similaire à un sauvetage génétique, sauf que la variance introduite est délibérément choisie pour remédier à un problème particulier, par exemple lié à des agents pathogènes (Christie and Searle 2018; Gignoux-Wolfsohn et al. 2021), espèces envahissantes (Kelly and Phillips 2016; Van Dyken 2020) ou même au cancer (Hohenlohe et al. 2019).

L'aire de répartition des espèces détermine également en grande partie la forme de flux de gènes assistés qui peut être envisagée. En effet, des espèces à large aire de répartition sont susceptibles de présenter une grande diversité d'adaptations climatiques au sein de leur aire de répartition, qui peuvent être utilisées comme des sources intéressantes de pré-adaptation aux climats futurs (*e.g.*, Mimura and Aitken 2010). Les tests de provenances et les jardins communs fournissent des informations précieuses sur les adaptations des différentes sources aux conditions environnementales (*e.g.*, Risk et al. 2021). Pour les espèces dont l'aire de répartition est limitée, si une telle variation existe, les sources de mélange utilisées pour les renforcements et les translocations peuvent introduire une nouvelle variation génétique dans l'espoir de fournir des allèles bénéfiques, de réduire la consanguinité et d'avoir un effet positif global sur la fitness, même si des variants pré-adaptés ne peuvent pas être identifiés. Dans le cas de populations très menacées, l'hybridation avec des populations éloignées de la même espèce, ou d'espèces étroitement apparentées peut constituer un réservoir d'allèles adaptatifs et de nouvelles variations génétiques qui ne peuvent plus être trouvées au sein de l'espèce ou des populations locales.

Bon nombre des espèces ciblées par les flux de gènes assistés ont une longue durée de vie et/ou des cycles de vie complexes, *i.e.*, différents stades de vie ayant des écologies et des sensibilités différentes au climat, ce qui affecte leur potentiel évolutif dans le cadre du changement climatique (Marshall et al. 2016). Pour les espèces à très longue durée de vie, telles que les arbres, un compromis vient compliquer le choix des sources de semences. Étant donné que les arbres sont confrontés à des changements climatiques rapides au cours de leur longue vie, les allèles qui leur confèrent des performances optimales peuvent varier entre le début et la fin de leur vie ;

par conséquent, devons-nous introduire des gènes qui confèrent une bonne adaptation au stade juvénile, au risque d'être maladapté aux stades plus âgés, ou bien une bonne adaptation aux climats futurs lorsque ces juvéniles seront devenus des adultes ?

### **Contributions des modèles existants**

Bien qu'ils soient de plus en plus étudiés, les flux de gènes assistés laissent encore de nombreuses questions sans réponse. Pour répondre à certaines d'entre elles, les modèles peuvent être utiles. Tout d'abord, nous comprenons encore mal quand, et si, les stratégies de flux de gènes assistés constituent la meilleure option pour atténuer les conséquences du changement climatique. Certains modèles ont commencé à explorer la manière dont différentes stratégies de flux de gènes assistés peuvent ou non accélérer l'adaptation, notamment en modélisant l'introgession d'allèles adaptatifs dans les populations à risque, comme les coraux (Quigley et al. 2019). Certaines inquiétudes concernant les flux de gènes assistés portent sur la perte de toute ou d'une partie du génome de l'espèce cible. Certaines études concluent que la partie du génome de la population réceptrice perdue à la suite d'un flux de gènes est rarement supérieure à la fraction des individus introduits (Aitken and Whitlock 2013; Grummer et al. 2022), d'autres constatent, à l'aide de simulations, qu'elle est proportionnelle à l'avantage en termes de fitness conféré par les mutations bénéfiques introduites (Harris et al. 2019). En pratique, un autre paramètre important est le nombre d'individus qui devraient être introduits et la fréquence à laquelle les flux de gènes devrait être mis en œuvre. Des simulations ont révélé un compromis entre la préservation du génome de la population cible et la réduction du risque d'extinction de la population (Kelly and Phillips 2019a). Un autre paramètre qui s'avère être également important est la source des individus transférés. Les cycles de vie complexes peuvent conduire à une adaptation plus lente en raison des compromis qui se produisent au cours de l'histoire de vie des individus (Cotto et al. 2019; Marshall et al. 2016), il est donc urgent de comprendre comment les cycles de vie affectent les stratégies d'approvisionnement pour les flux de gènes assistés. Finalement, une autre question est de savoir comment déterminer la probabilité de sauvetage d'une population et comment en quantifier ces avantages. L'intégration des processus évolutifs aux analyses de viabilité des populations (PVA) ouvre la voie à des estimations plus réalistes de la persistance des populations dans le cadre du changement climatique (Kelly and Phillips 2019b; Pierson et al. 2015).

## Objectifs et principaux résultats

Dans cette thèse, nous avons étudié des questions autour de différents modèles biologiques qui sont des cibles pour les flux de gènes assistés et qui font face à différents défis. Nous résumons les questions, méthodes et les principaux résultats ci-dessous.

### Chapitre 1

Le changement climatique menace les arbres à longue durée de vie, qui risquent de ne pas s'adapter ou de ne pas migrer assez rapidement pour suivre la hausse des températures. Les flux de gènes assistés pourraient faciliter l'adaptation des populations aux climats futurs grâce à la translocation gérée de graines provenant d'un endroit plus chaud ("provenance") à l'intérieur de l'aire de répartition actuelle d'une espèce. Trouver la provenance qui aura les meilleures performances en termes de survie ou de croissance est compliqué par un compromis : étant donné que les arbres sont confrontés à un climat qui change rapidement au cours de leur longue vie, les allèles qui leur confèrent des performances optimales peuvent varier tout au long de leur vie. Par exemple, les arbres provenant de provenances plus chaudes seront bien adaptés à l'âge adulte mais pourraient souffrir de températures encore froides alors qu'ils ne sont encore qu'à un stade juvénile peu tolérant. Nous utilisons ici un modèle structuré en stade pour déterminer quelle provenance maximiserait la survie d'une cohorte d'arbres à longue durée de vie dans un climat en réchauffement. Notre modèle prédit que la meilleure provenance dépend de la vitesse à laquelle le climat changera, mais aussi en grande partie de la façon dont la tolérance climatique des arbres varie entre les différents stades de vie et du temps passé au stade le moins tolérant. Nous paramétrons le modèle avec des matrices de transition démographique estimées empiriquement pour vingt espèces d'arbres à longue durée de vie et nous montrons que les conclusions précédentes sont robustes aux variations du cycle de vie des arbres. Nous appelons donc à un effort accru pour mesurer comment la tolérance climatique change au cours de la vie, chez les espèces à longue durée de vie, ce qui reste inconnu pour la plupart des espèces d'arbres.

### Chapitre 2

Le sauvetage évolutif est un processus démographique et évolutif par lequel une population échappe à l'extinction en s'adaptant aux pressions environnementales. La Centaurée de la Clape (*Centaurea corymbosa*) est une plante vivace monocarpique, rare et menacée, endémique du Massif de la Clape dans le sud de la France. Elle fait l'objet d'un suivi démographique à long

terme depuis la fin des années 1990. Les petites populations restantes de cette espèce sont en déclin depuis plusieurs années et son extinction est prédite à moyen terme, accélérée par le changement climatique. L'objectif de cette étude est de comprendre si l'évolution d'un trait déterminant pour la fitness, le taux de croissance, pourrait sauver la population dans un climat qui se réchauffe. Nous proposons une nouvelle méthode pour intégrer de l'évolution dans un modèle de projection intégrale (IPM) et l'utilisons pour prédire le niveau de variance génétique critique pour ce trait adaptatif qui serait nécessaire pour éviter l'extinction. Nous étudions dans un premier temps, un scénario simple qui suppose qu'une seule cohorte est plantée, l'objectif étant de modéliser des scénarios plus complexes de renforcements à terme. Nous montrons que la tendance de déclin de la population ne peut être inversée qu'avec une petite gamme de diversité génétique introduite dans le taux de croissance. De plus, même dans le meilleur des cas, la population ne sera pas ne se remet pas à croître avant vingt ans, horizon auquel le taux d'extinction de certaines populations naturelles est déjà très élevé.

### Chapitre 3

La théorie suggère que la probabilité d'un sauvetage évolutif, lorsque l'évolution empêche une extinction autrement inévitable, augmente avec la taille de la population et la disponibilité de la variation génétique. L'introggression, c'est-à-dire le rétrocroisement d'hybrides avec une ou les deux lignées parentales, peut favoriser le sauvetage évolutif en introduisant de nouvelles variations génétiques dans les populations menacées. Bien que des cas d'introggression adaptative soient connus, ils aboutissent le plus souvent à un envahissement démographique ou génétique. Nous avons analysé un modèle analytique déterministe, soutenu par des simulations individu-centrées, afin de quantifier les effets de l'envahissement génétique après un seul événement d'introduction d'individus d'une autre population. Nos résultats analytiques identifient un nombre optimal d'individus introduits qui maximise la portion restante du génome de la population cible. En supposant que la population cible a décliné et n'exerce qu'une faible compétition, des introductions petites et précoces minimisent l'envahissement, car la population résidante reste à un niveau suffisant, tandis que des introductions plus importantes maximisent la probabilité d'un sauvetage. Non seulement nos résultats analytiques permettent de mieux comprendre le compromis entre l'envahissement génétique et le sauvetage, mais ils fournissent également des outils pour guider les gestionnaires de la conservation visant à sauver des populations de petite taille et/ou consanguines.

# General introduction

## Loss of biodiversity

Biodiversity has been facing an extinction crisis characterized by extinction rates comparable to those of the five great mass extinction crises (Barnosky et al. 2011; Dirzo et al. 2014). The current crisis is largely the result of anthropogenic pressures, including over-exploitation, agricultural activities, urban development and climate change (Maxwell et al. 2016a). These pressures lead to decreases in population sizes and available genetic diversity, which in turn lead to an increased vulnerability to stochasticity and extinction factors associated with small population sizes (Pimm et al. 2006b; Shaffer 1981). Climate change (IPCC 2021) exposes populations to more frequent and more intense extreme events (*e.g.*, heatwaves, fires, heavy rain; Mitchell et al. 2006) and gradually disrupts the existing local adaptation of populations to their historical environmental conditions (Aitken and Bemmels 2016; Aitken and Whitlock 2013; Brady et al. 2019; Crespi 2000). Unless populations can actively track their climatic niche, climate change-related maladaptation can initiate or worsen population declines (Duffy et al. 2022; Román-Palacios and Wiens 2020). The consequences of declining species numbers, and ultimately their extinction, are critical for ecosystems, particularly when the species affected are keystone species such as forest trees (Aitken and Whitlock 2013; Schmid et al. 2009).

## Population responses to environmental pressures

Faced with climate change, populations must respond to avoid extinction. *In situ*, phenotypic plasticity could mitigate the effects of maladaptation until an adaptive response is possible (Anderson et al. 2012; Chevin 2015; Chevin et al. 2010). It is still unclear to date, however, whether plasticity would allow the avoidance of extinction, what form the plastic response would take and its effects on evolutionary responses (Diamond and Martin 2021a; Gienapp et al. 2008;



Parmesan 2006). Natural selection could lead to such an adaptive response provided that an adaptive genotype increases in frequency in the population as a result of environmental change. If such evolution is sufficiently fast to allow the population to maintain a large population size and high prospects of persistence, this process is called evolutionary rescue (Carlson et al. 2014; Gomulkiewicz and Holt 1995). Studies suggest, however, that many species are not evolving fast enough to keep pace with the current rate and magnitude of environmental change (*e.g.*, Kremer et al. 2012; Radchuk et al. 2019).

Dispersal is another possible response to environmental constraints. Although there is evidence that populations are shifting their ranges towards the poles and/or higher altitudes to follow their climatic optimum (Chen et al. 2011; Freeman et al. 2018) and that this reduces extinction risks (Román-Palacios and Wiens 2020), it can however be limited by environmental barriers (*e.g.*, habitat fragmentation; Leimu et al. 2010). The life cycle characteristics of some species such as a sessile nature (plants, corals, etc.) or a long generation time will also most likely prevent species from achieving the migration speed required to persist (Ettinger and HilleRisLambers 2013; Shaw and Etterson 2012).

## Human-assisted gene flow

The current crisis has led to calls for conservation actions to improve the fitness of declining populations and accelerate their adaptation to a changing climate. As a result, human-assisted translocations have been proposed to help populations maintain, recover and adapt to climate change (IUCN/SSC, 2013). These practices can either aim at increasing the number of individuals and genetic diversity in targeted populations, thereby stimulating demographic and genetic rescue of populations (Carlson et al. 2014), or increase the frequency of pre-adapted genotypes in anticipation of or in response to environmental change, to stimulate evolutionary rescue (Aitken and Bemmels 2016; Aitken and Whitlock 2013).

There is a continuum between the different types of intervention in terms of where individuals are moved geographically (described in Figure 1 from Aitken and Bemmels 2016, Seddon 2010), ranging from (i) reinforcement measures of the populations *in situ* to (ii) assisted migration strategies. The latter can either take the form of assisted migration outside of the historical range of the species, sometimes referred to as “assisted colonization”, or assisted migration within the current range of a species also often referred to as “assisted gene flow”. This terminology is,

nevertheless, somewhat confusing as all practices describe variations of human-assisted flow of genes. Assisted migration beyond the species range is highly debated because of the associated risks, such as maladaptation, hybridization, pathogen introduction, and invasive spread (see for instance Hunter 2007; Ricciardi and Simberloff 2009). These risks are considered lower in the case of assisted gene flow among existing populations, with the exception of some genetic risks such as outbreeding depression, genomic swamping, genetic incompatibilities (Aitken and Whitlock 2013; Frankham et al. 2011; Muller 1942; Sáenz-Romero et al. 2021). The general public is also generally more open to assisted migration practices within than beyond the species range (Peterson St-Laurent et al. 2018).

### **Examples of assisted gene flow**

Assisted gene flow can be implemented in a wide range of species, animal or plant taxa, but is mainly considered for foundation species (*e.g.*, corals), managed species (*e.g.*, trees) and small, endangered populations (Aitken and Whitlock 2013). Transfer rules, defining best practices when selecting sources of biological material, increasingly mention climatic resemblance with future climate as a critical selection criterion, for replanting after harvest, restoration, or population reinforcement, rather than the exclusive use of local sources. This is notably the case of many tree species (Breed et al. 2018; O'Neill et al. 2008, 2017). Across taxa, clear calls are also being made in this direction, as in the case of corals (*e.g.*, Hagedorn et al. 2021), fishes (*e.g.*, Pavlova et al. 2017; Pregler et al. 2023), mammals (*e.g.*, Seddon and Schultz 2020) and trees (*e.g.*, Borrell et al. 2020; Browne et al. 2019; Girardin et al. 2021; Milesi et al. 2019; Young et al. 2020). Although this effect does not seem to be the result of an active policy of adaptation to climate change, a recent study showed that plant translocations were more often carried out towards slightly colder sites (Diallo et al. 2021), further showing these processes are already underway.

### **Different challenges faced by different species**

Based on the challenges faced by the different species targeted for assisted gene flow, a wide range of practices can be considered. Assisted gene flow originally emerged as a strategy to help populations adaptation to climate change, yet many translocations are not climate-related (Diallo et al. 2021; Kelly and Phillips 2016). “Targeted gene flow”, for instance, refers to the movement of variants of interest, in a similar way to genetic rescue, except the variance intro-

duced is deliberately chosen to remedy a particular concern, for instance related to pathogens (Christie and Searle 2018; Gignoux-Wolfsohn et al. 2021), invasive species (Kelly and Phillips 2019b; Van Dyken 2020) or even cancer (Hohenlohe et al. 2019).

Species range also largely determines the form of assisted gene flow that can be considered. Indeed, species with a wide range are likely to hold a large diversity of climatic adaptations within their range that can be used as interesting sources of pre-adaptation to future climates (*e.g.*, Mimura and Aitken 2010). Provenance tests and common gardens provide valuable information on the adaptations of different provenances to environmental conditions (*e.g.*, Risk et al. 2021). For species with a limited range, if such variation exists, mixing sources used for reinforcements and translocations can introduce new genetic variation in the hope of providing beneficial alleles, reducing inbreeding, and having an overall positive effect on fitness, even though pre-adapted variants cannot be identified. In the case of highly endangered populations, hybridization with distant populations of the same species or closely related species may provide a reservoir of adaptive alleles and new genetic variation that cannot be found anymore within the species or the local populations.

Many of the species targeted by assisted gene flow have long life spans and/or complex life cycles, with different stages having different ecologies and sensitivities to climate change, affecting their evolutionary potential under climate change (Marshall et al. 2016). For very long-lived species such as trees, a trade-off complicates the choice of seed sources. Because trees face a rapidly changing climates throughout their long lifetime, the alleles that confer optimal performance may vary over the course of their lives; thus, should we introduce genes that confer good adaptation now, at the juvenile stage, or adaptation to future climates when those juveniles have become adults?

## Contributions from previous models

Although increasingly studied, assisted gene flow still leave many questions unanswered. To answer some of them, models can be useful. First, we still have little understanding of when, and if, assisted gene flow strategies are the best option for mitigating the consequences of climate change. Some models have begun to explore how different assisted gene flow strategies may or may not accelerate adaptation, in particular by modeling the introgression of adaptive alleles into at-risk populations, *e.g.*, conferring tolerance to higher temperatures in coral populations

(Quigley et al. 2019). Concerns about assisted gene flow include the loss of the genome uniqueness of the target species. Some studies conclude that the portion of the recipient population's genome lost as a result of gene flow rarely constitutes more than the fraction of individuals introduced (Aitken and Whitlock 2013; Grummer et al. 2022) and others find, using simulations, that it is proportional to the fitness advantage conferred by the beneficial mutations introduced (Harris et al. 2019). An important practical parameter is the number of individuals that should be introduced, and the frequency at which gene flow should be implemented. Simulations revealed a trade-off between preserving the genome of the target population and reducing the risk of population extinction (Kelly and Phillips 2019a). In practice, another important parameter is the source of the translocated individuals. Complex life cycles may lead to slower adaptation due to trade-offs that occur during the life history of individuals (Cotto et al. 2019; Marshall et al. 2016; Marshall and Connallon 2023). There is, therefore, an urgent need to understand how species life cycles affect sourcing strategies for assisted gene flow. Another question regards how the probability of population rescue can be determined, and how these benefits can be quantified. The integration of evolutionary processes into Population Viability Analysis (PVA) paves the way to more realistic estimations of population persistence under climate change (Kelly and Phillips 2019b; Pierson et al. 2015).

## **Aims of this thesis**

In this thesis, we are investigating questions around different biological models that are targets for assisted gene flow and face different challenges.

First, we urgently need to understand the challenges faced by long-lived species growing in a changing climate to understand how this affects sourcing strategies for assisted gene flow. We investigate this question using staged-structured models where individuals have different tolerances to climate across their lifespan. We use trees, as an example since they are very long-lived organisms often targeted by assisted gene flow practices by parameterizing the model with empirically estimated demographic transition matrices for twenty long-lived tree species (Chapter 1).

Rather than predicting which genetic variants will be adaptive in the future, Chapter 2 focuses on increasing diversity and adaptive potential in a more general way. Our goal is to understand whether the evolution of a fitness-determining trait in a rare and endangered species

through gene flow could save the population in a warming climate. We predict the critical level of genetic diversity that would be required for this species to avoid extinction using an evolutionary explicit integral projection model (IPM). The methodology for integrating evolution into an IPM has been the subject of debates. Here, we propose a new approach, building on some previous work by Rees and Ellner 2019. We focus on the case of *Centaurea corymbosa*, a rare and endangered monocarpic perennial that has been the subject of a long-term demographic study and because a demographic IPM has already been fitted to this large data-set (Chapter 2).

Finally, Chapter 3 considers the costs of assisted gene flow to the target species. We focus on the trade-off between preserving the genome of the target population and reducing the risk of population extinction, generalizing previous simulation work with analytical predictions. We provide quantitative guidance on the optimal number of migrants to introduce, to maximize both the probability of evolutionary rescue through introgression of a beneficial allele, and the proportion of the target species genome preserved (Chapter 3).

## Chapter 1

# Planting long-lived trees in a warming climate: Theory shows the importance of stage-dependent climatic tolerance

## Planting long-lived trees in a warming climate: Theory shows the importance of stage-dependent climatic tolerance

Adèle Erlichman<sup>1,2\*</sup>, Linnea Sandell<sup>2,3,4\*</sup>, Sarah P. Otto<sup>2</sup>, Sally N. Aitken<sup>4</sup>, Ophélie Ronce<sup>1,2</sup>

<sup>1</sup> ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

<sup>2</sup> Department of Zoology, University of British Columbia, Vancouver, Canada

<sup>3</sup> Department of Organismal Biology, Uppsala University, Sweden

<sup>3</sup> Department of Urban and Rural Development, Swedish University of Agriculture, Sweden

<sup>4</sup> Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, Canada

Corresponding author email: [adeleerlichman@gmail.com](mailto:adeleerlichman@gmail.com)

\* These authors contributed equally to this work.

### Abstract

Climate change poses a particular threat to long-lived trees, which may not adapt or migrate fast enough to keep up with rising temperatures. Assisted gene flow could facilitate adaptation of populations to future climates by using managed translocation of seeds from a warmer location (provenance) within the current range of a species. Finding the provenance that will perform best in terms of survival or growth is complicated by a trade-off. Because trees face a rapidly changing climate during their long lives, the alleles that confer optimal performance may vary across their lifespan. For instance, trees from warmer provenances could be well-adapted as adults but suffer from colder temperatures while juvenile. Here we use a stage-structured model, using both analytical predictions and numerical simulations, to determine which provenance would maximize the survival of a cohort of long-lived trees in a changing climate. We parameterize our simulations using empirically estimated demographic transition matrices for twenty long-lived tree species. Unable to find reliable quantitative estimates of how climatic tolerance changes across stages in these same species, we vary this parameter to study its effect. Both our mathematical model and simulations predict that the best provenance depends strongly on how fast the climate changes and also how climatic tolerance varies across the life span of a tree. We thus call for increased empirical efforts to measure how climate tolerance changes over life in long-lived species, as our model suggests that it should strongly influence the best provenance for assisted gene flow.

*Keywords:* Assisted gene flow; Climate change; Local adaptation; Forestry; Complex life cycles; Seed sourcing

## Introduction

There is strong evidence for the local adaptation of many populations to their historical climate (Alberto et al. 2013; Lortie and Hierro 2022; Savolainen et al. 2007; Wadgymar et al. 2022). Current anthropogenic global warming is however disrupting climatic conditions (IPCC 2021), decoupling populations from the climate to which they were historically adapted (Benito-Garzón et al. 2018; Bontrager and Angert 2019; Gauli et al. 2022; Hanewinkel et al. 2013; Thomas et al. 2004). Climate change-related maladaptation could thus lead to, or exacerbate, population declines (Frank et al. 2017), especially when combined with other pressures faced by biodiversity (Maxwell et al. 2016b). While local adaptation to new climatic conditions may be restored with time due to natural evolutionary and ecological processes, such as adaptive phenotypic changes (Bell 2017; Yeh and Price 2004) and range shifts (Chen et al. 2011; Freeman et al. 2018; Lenoir and Svenning 2015), these natural processes may be insufficient in species with low evolutionary potential (Shaw and Etterson 2012), long generation time (*e.g.*, Cotto et al. 2017), maladaptive phenotypic plasticity (Diamond and Martin 2021b; Fox et al. 2019), or poor dispersal ability (Leimu et al. 2010).

In circumstances where life cycle attributes severely limit natural dispersal, human-assisted migration has been proposed to accelerate population adaptation to climate change, either as a conservation strategy or to maintain productivity for economically important species (Isabel et al. 2020). Assisted migration, also called assisted gene flow when focusing on movements within the historical range of populations (Aitken and Whitlock 2013), uses translocation of pre-adapted genotypes in anticipation of or in response to environmental change (Aitken and Bemmels 2016). Risks associated with these practices (Aitken and Whitlock 2013; Ricciardi and Simberloff 2009; Sáenz-Romero et al. 2021; Weeks et al. 2011) include maladaptation, outbreeding depression, gene swamping, hybridization, pathogen introduction, and invasive spread. We still have few tools to predict the efficiency of such interventions in mitigating the negative consequences of climate change, and their optimal design. Our aim here is to develop simple theoretical arguments contributing to a better understanding of trade-offs affecting assisted migration in long-lived species. Because risks are thought to be smaller in the case of assisted gene flow (Aitken and Whitlock 2013), we will focus on scenarios mimicking translocations within a species range. Our model however also helps understanding trade-offs affecting translocations beyond the range limits.



Assisted gene flow is particularly discussed in the context of forestry practices (Aitken and Whitlock 2013). Indeed, the long generation time and limited dispersal capabilities resulting from their prolonged sessile phase make long-lived trees particularly vulnerable to the effects of climate change (*e.g.*, Brodribb et al. 2020; Gougherty et al. 2021). Most temperate and boreal forest tree species are locally adapted to climate (Leites and Benito Garzón 2023) and already suffer negative consequences of climate change (*e.g.*, Abram et al. 2021; Davis et al. 2019; Forzieri et al. 2022; Hartmann et al. 2022; Hellmann et al. 2008; Sáenz-Romero et al. 2020; Stevens-Rumann et al. 2018; Sturrock et al. 2011). The vulnerability and decline of tree populations is of interest because trees provide numerous ecosystem services (Brockerhoff et al. 2017), with important ecological (*e.g.*, providing food resources and habitat for many other species (Yu et al. 2023), carbon storage, microclimate regulation), economic (Fraga et al. 2020), societal (Esperon Rodriguez et al. 2022), and cultural roles (Agnoletti and Santoro 2015; Lewis and Sheppard 2005; Trigger and Mulcock 2005).

To implement assisted gene flow, one needs to identify seed sources that maximise the survival and productivity of the transplanted individuals in the future, based on projected climate change. In forestry, the population used for seed sourcing is referred to as a "provenance", specifically in reference to the geographical location (Aitken and Bemmels 2016). Traditional seed-sourcing strategies use transfer functions and provenance trials to identify provenances best adapted to a plantation site (*e.g.*, O'Neill et al. 2017; Risk et al. 2021; Thomson et al. 2009). Historically, the use of local provenances was widely recommended for reforestation (O'Neill and Gómez-Pineda 2021). Managers however are increasingly being advised to anticipate near-future climate change when selecting seed sources for plantations (*e.g.*, Aitken and Bemmels 2016; Benito-Garzón et al. 2018; Girard et al. 2022; Sáenz-Romero et al. 2021; St. Clair et al. 2022; Whittet et al. 2016). For instance, in British Columbia, Canada, a recently adopted policy advises that planted populations be adapted to the expected future climate in the plantation site (O'Neill et al. 2017). Several assisted gene flow experiments are already underway (*e.g.*, McDonald et al. 2021; Prieto-Bemétez et al. 2021; Young et al. 2020).

Only a few recent models have however begun to investigate how different assisted gene flow strategies can accelerate adaptation (DeFilippo et al. 2022; Grummer et al. 2022; Quigley et al. 2019), and even fewer have attempted to quantify the demographic consequences of assisted gene flow and its impact on population persistence and function (*e.g.*, Bay et al. 2017; Kelly and Phillips 2019a; Kuparinen and Uusi-Heikkilä 2020). Many organisms targeted for assisted

gene flow, such as trees or corals, are long-lived and have complex life cycles: life-cycle stages (*e.g.*, seed, seedling, sapling, pole, mature tree, and senescent stages) differ in their tolerance to climatic variables (Du et al. 2019; Mašek et al. 2021; Pompa-Garciéa and Hadad 2016), and different individuals may spend different amounts of time in each stage before growing to the next (Jackson et al. 2009). Yet, we have little insight into how complex life cycles may affect optimal assisted gene flow strategies. In very long-lived organisms, a trade-off occurs when climate changes during an individual's lifetime: individuals will on average experience colder temperatures early in their lives and warmer temperatures later. The alleles conferring good performance may then be different at the beginning and end of life. In particular, there is concern that warm-adapted provenances that will perform well in the future could still suffer from damage due to colder temperatures while juvenile (Sebastian-Azcona et al. 2019). Previous theoretical studies have suggested that it will be challenging for long-lived species with complex life cycles to adapt to a changing climate, because their long generation time slows down evolution, but also because of trade-offs in adaptation to climate change between different life stages (Cotto and Chevin 2020; Cotto et al. 2019; Marshall et al. 2016; Marshall and Connallon 2022).

In organisms with a complex life cycle, different stages typically make different contributions to the growth of the population (Caswell 2000). In long-lived trees, for instance, individuals in young stages with high mortality typically have a low contribution to population dynamics compared to older individuals. In reintroduction or reinforcement programs, releasing individuals with high reproductive value and contribution to population growth can increase success of the program (Sarrazin and Legendre 2000). In assisted gene flow programs, it is however unknown whether similar reasoning applies: is it optimal to select seed sources that will be best adapted at the stage with the highest contribution to population growth, *i.e.*, when fully grown? Or is it optimal to source seeds that provide the most protection during life stages that are least tolerant to climate change? Some studies suggest that young stages tend to be less tolerant to stress than older stages (*e.g.*, Black and Bliss 1980; Kueppers et al. 2017; Munier et al. 2010). For example, in British Columbia, the current recommendation is to plant populations adapted to the expected future climate at one quarter of the rotation time (*i.e.*, time between planting and harvest), when trees are still relatively young, to account for the lower tolerance of seedlings to maladaptation (O'Neill et al. 2017). These rules of thumb however lack a strong theoretical foundation and we do not know how the choice of the best provenance depends on the exact life cycle of the tree, the speed of climate change and the impact it has on different life stages.

The aim of this paper is, therefore, to fill this gap by theoretically exploring the effect of (1) changes in tolerance to climate over the lifetime of trees and (2) variation in life history (*i.e.*, variation in stage-specific vital rates) on the choice of the best provenance for assisted gene flow. To this end, we use a simple scenario that describes the persistence of a single cohort of long-lived trees in a warming environment. This scenario approximates even-aged forest management, where seedlings of the same age are planted after a clearcut and harvested a number of decades later (Puettmann et al. 2015). It will serve as a baseline for future exploration of more complex forestry scenarios. To understand what determines the best provenance for maximizing survival of a cohort of trees under climate change, we extend the analytical work of Cotto and Chevin (2020), who used an age-structured model to examine the evolution of a quantitative trait in organisms that undergo multiple episodes of selection in a changing environment. We instead consider a stage-structured model that better represents the life history of species like trees, for which size more than age determines the survival prospects. We derive analytical predictions for the best provenance to maximize cohort survival under climate change. To develop a better understanding, we first illustrate these mathematical predictions in the simplest case where there are only two stages in the life cycle (juveniles and adults), exploring how the time spent in each stage and the difference in climatic tolerance across stages affect the choice of the best provenance. We then parameterize our model with empirically-measured life histories for 20 long-lived species. We check that qualitative conclusions derived in the two-stages case hold when the number of stages increases and for various tree life histories. In addition, we explore the effect of changes in various assumptions to assess robustness of our results to: (i) the change in thermal tolerance across life, (ii) the rate of environmental warming, (iii) random interannual variation in temperature and (iv) variation in thermal tolerance among individuals from the same cohort. We compare the potential benefits of implementing assisted gene flow predicted by our model for different choices for seed sourcing: planting only seeds from the local provenance, planting the best provenance predicted by our model, or using a simple rule of thumb to match seed provenance to the future climate in the plantation site, such as used in British Columbia (O'Neill et al. 2017).

## Methods

### Biological scenario

Given a single cohort of trees, planted at  $t = 0$  and harvested after  $H$  years under environmental warming, we aim to optimize the seed source ("provenance") considering the trade-off between current climate adaptation for seedlings and future climate adaptation for older trees. Trees in our model have a complex life cycle spanning  $n$  discrete stages, each with different tolerance to climate. The seed source that yields the highest number of surviving trees in the last stage  $n$  after  $H$  years will be denoted as the best provenance. Note that throughout this study, we use the term "harvest" to quantify the final size of the cohort. Maximizing the size of a cohort (i.e. the number of individuals reaching peak reproductive stage) also achieves conservation goals when individuals are not harvested.

For the sake of illustration, we characterize each provenances by the mean annual temperature that would maximize the survival of trees from that provenance in a constant climate (their thermal optimum,  $\theta$ ). This choice is motivated by the fact that mean annual temperature is frequently used to characterize the climatic niche of different provenances in experimental plantations (*e.g.*, Pedlar et al. 2021a,b; Wang et al. 2006). Our model could however be similarly applied to any climatic variable or combination of climatic variables affecting life history traits and characterizing the climatic niche of the different provenances ( $\theta$  would then be a vector with all relevant niche dimensions, *e.g.*, showing the optimal position along precipitation and temperature gradients). Each cohort is made up of trees from a single provenance, all of which share a constant thermal optimum throughout their life. The thermal niche width of this cohort is however allowed to vary across life stages, reflecting different sensitivities to temperature change. To evaluate the robustness of our predictions, we later relax the assumption that all individuals in the cohort have the same thermal optimum. Individuals from the same provenance may indeed vary in their response to temperature and seed sourcing strategies may involve mixing different provenances (Breed et al. 2018; Whittet et al. 2016). These numerical results are presented in supplementary material S1.

We assume that the annual survival rate declines as individuals are exposed to temperatures deviating from their thermal optimum. For mathematical convenience and building upon previous theoretical quantitative genetic models of adaptation to variable environment (Cotto and Chevin 2020), we more precisely define the survival rate in stage  $i$  ( $s_i$ ) as a Gaussian function

of the current temperature  $T_t$  in year  $t$ , with  $s_{i,\max}$  being the maximal survival rate in stage  $i$ :

$$s_i(\theta, T_t) = s_{i,\max} \exp\left(-\frac{(\theta - T_t)^2}{2\omega_i^2}\right) \quad (1.1)$$

An individual has the highest probability to survive when the yearly temperature matches its thermal optimum  $T_t = \theta$ . The parameter  $\omega_i$  scales the thermal tolerance of an individual in stage  $i$ : according to the above expression, when the temperature deviates from its optimal temperature ( $\theta$ ) by  $\omega_i$  degrees, the survival rate is reduced by a factor  $1 - e^{-\frac{1}{2}} \approx 40\%$ .

### Stage structure dynamics

We track the number of individuals in a stage-structured population under this scenario (Fig. 1). The transition probabilities between stages, from year to year, are described by an  $n$ -by- $n$  matrix  $\mathbf{M}$ , and the number of individuals in each stage is stored in a vector  $\mathbf{N}_t$ . Each year, individuals can either remain in their current stage if they survive and do not grow sufficiently to transition (with probability  $s_i(\theta, T_t) \times (1 - g_i)$ ), die (with probability  $1 - s_i(\theta, T_t)$ ), or transition to the next stage if they survive and grow (with probability  $s_i(\theta, T_t) \times g_i$ ), where  $g_i$  is the probability of growing to stage  $i + 1$  conditional on survival. We only allow transitions to occur to the same or the following stage, *i.e.*, we assume trees neither shrink nor skip stages. Hence, at the next time step  $t + 1$ , the number of individuals in each stage is:

$$\mathbf{N}_{t+1} = \mathbf{M}(\theta, T_t) \cdot \mathbf{N}_t \quad (1.2)$$

We do not explicitly consider the effects of density-dependent competition on survival, but assume that empirical estimates of survival and growth rates in natural populations that we use to parameterize the model already integrate the effect of competition and were measured at densities relevant to our forestry scenarios. Individuals can remain in a stage for several years, and different individuals can stochastically grow to the next stage at different ages. We denote by  $a_i$  the age at which an individual makes the transition from stage  $i$  to stage  $i + 1$ , where  $i \in [1, \dots, n - 1]$ . We define  $\vec{a} = [a_1, \dots, a_{n-1}]^T$  as the individual's life history, containing the age at which an individual transitioned to each stage, given that it reached the last stage before harvest, and  $p(\vec{a})$  as the fraction of surviving individuals at harvest with this realized life history.

We also investigate a scenario where both growth ( $g_i$ ) and survival ( $s_i$ ) rates depend on temperature, which shows very similar results to the case discussed in the main text (Supplementary

material S2; Suppl. Fig. 3).

## Analytical predictions

We derive an expression for the expected number of mature trees at harvest time,  $N_H$ , given an initial cohort of  $N_0$  seedlings from a provenance with thermal optimum  $\theta$ . We begin by assuming that all trees share a particular life history  $\vec{a}$ . Indeed, when the time spent in each stage is fixed, we can apply the age-structured results of Cotto and Chevin (2020) for multiple rounds of Gaussian selection on a single cohort. At harvest, the pool of remaining individuals is formed by the individuals that have survived each stage and reached the last:

$$E[N_H|\vec{a}] = N_0 \prod_{t=1}^{a_1} s_1(\theta, T_t) \times \cdots \times \prod_{t=a_{i-1}+1}^{a_i} s_i(\theta, T_t) \times \cdots \times \prod_{t=a_{n-1}+1}^H s_n(\theta, T_t) \quad (1.3)$$

Because equation 1.3 is a product of Gaussian functions, we can use the result of Cotto and Chevin (2020) to write the expected number of remaining individuals as:

$$E[N_H|\vec{a}] = N_0 s_{\text{tot,max}}(\vec{a}) \exp\left(-\frac{(\theta - T_{\text{tot}}(\vec{a}))^2}{2 \omega_{\text{tot}}^2(\vec{a})}\right) \quad (1.4)$$

with,

$$\omega_{\text{tot}}^2(\vec{a}) = \left(\frac{a_1}{\omega_1^2} + \cdots + \frac{a_i - a_{i-1}}{\omega_i^2} + \cdots + \frac{H - a_{n-1}}{\omega_n^2}\right)^{-1} \quad (1.5)$$

$$T_{\text{tot}}(\vec{a}) = \sum_{t=1}^{a_1} \frac{\omega_{\text{tot}}^2(\vec{a})}{\omega_1^2} T_t + \cdots + \sum_{t=a_{i-1}+1}^{a_i} \frac{\omega_{\text{tot}}^2(\vec{a})}{\omega_i^2} T_t + \cdots + \sum_{t=a_{n-1}+1}^H \frac{\omega_{\text{tot}}^2(\vec{a})}{\omega_n^2} T_t \quad (1.6)$$

$$s_{\text{tot,max}}(\vec{a}) = \left(s_{1,\text{max}}^{a_1} \times \cdots \times s_{i,\text{max}}^{a_i - a_{i-1}} \times \cdots \times s_{n,\text{max}}^{H - a_{n-1}}\right) \exp\left(\frac{T_{\text{tot}}(\vec{a})^2}{2\omega_{\text{tot}}^2(\vec{a})} - \sum_{t=1}^{a_1} \frac{T_t^2}{2\omega_1^2} - \cdots - \sum_{t=a_{i-1}+1}^{a_i} \frac{T_t^2}{2\omega_i^2} - \cdots - \sum_{t=a_{n-1}+1}^H \frac{T_t^2}{2\omega_n^2}\right) \quad (1.7)$$

According to equation 1.4, the probability that an individual with life history  $\vec{a}$  survives until harvest is maximal and equal to  $s_{\text{tot,max}}(\vec{a})$ , if its thermal optimum  $\theta$  matches the realized temperature experienced across its life span  $T_{\text{tot}}(\vec{a})$ , weighting the temperature at each age by the thermal tolerance of the stage corresponding to that age as in equation 6. The "cumulative thermal tolerance" experienced by an individual with a growth trajectory  $\vec{a}$  across its life span

is measured by  $\omega_{\text{tot}}(\vec{a})$ , which can be computed as  $\omega_{\text{tot}}^2(\vec{a}) = \frac{\tilde{\omega}^2}{H}$ , where  $\tilde{\omega}^2$  is the harmonic mean of the squared thermal tolerances experienced by an individual throughout its life, as given by equation 1.5. If temperature tolerance is constant throughout life (*i.e.*,  $\omega_i^2$  is constant at  $\omega^2$ ), then  $\omega_{\text{tot}}^2(\vec{a}) = \frac{\omega^2}{H}$ . In this case the best provenance has a thermal optimum equal to the midpoint temperature experienced during the life span of a tree (*i.e.*, when  $\theta = T_{\text{tot}}(\vec{a}) = \frac{1}{H} \sum_{t=1}^H T_t$ ; Cotto and Chevin 2020).

The above result applies only to stage-structured populations when there is a fixed amount of time between stages, in which case the stage-structured model is equivalent to an age-structured model. More generally, because individuals of the same cohort will vary stochastically in the time they spend in each stage, we must consider all possible life histories ( $\vec{a}$ ) to estimate the number of remaining individuals at harvest.

Accounting for all possible life histories, the pool of individuals that survive and grow to harvestable size by time  $H$  is expected to be:

$$E[N_H] = \sum_{\vec{a}} p(\vec{a}) E[N_H | \vec{a}] \quad (1.8)$$

where  $p(\vec{a})$  is the probability of the life history  $\vec{a}$  conditional on survival until harvest, given by  $p(\vec{a}) = \prod_{i=1}^{n-1} (1 - g_i)^{a_i - a_{i-1} - 1} g_i$  with  $a_0 = 0$ .

For the sake of illustration, we explore the prediction of this analytical model in the simple case where there are only two stages (*e.g.*, adults and juveniles). In this simple case, the life history of a surviving individual can be uniquely characterized by the age at transition between the two stages (*e.g.*, age at maturity). We also assess whether predictions about the best provenance, using equation (1.4) and the average age at transition in the cohort, provide a good approximation for what happens in the more complex stage-structured case where individuals may differ in their life history. We numerically look for the best provenance maximizing  $E[N_H]$  when integrating across all possible growth trajectories, using equation (1.8). We also compute the mean age of maturation  $a$  among survivors at harvest, integrating across all possible life histories for a given life cycle, and use that value to compute  $T_{\text{tot}}(a)$  as in equation (1.6). We compare the thermal optimum of the best provenance predicted by the two methods. To generate variation in the mean age of maturation among survivors at harvest, we repeat these calculations, drawing 500 values of  $g$ ,  $s_{1,\text{max}}$  and  $s_{2,\text{max}}$  from uniform distributions between 0.01 and 0.99.

## Simulations with vital rates for twenty long-lived tree species

We now turn to more complex life cycles, with more than two stages, and explore how the best provenance varies as a function of life history parameters. We extract empirically measured transition matrices ( $\mathbf{M}$ ) for long-lived tree species from the COMPADRE database (Salguero-Gómez et al. 2015) and perform simulations using equation (1.2) for these examples. It should be noted that the transition matrices gathered are quite variable in the way that they categorize life stages, both in the definition and number of stages (see details in Supplementary material S4). We assume that these trees are adapted to the climatic conditions in which they are currently growing. This allows us to estimate  $s_{i,\max}$  and  $g_i$  from the transition rates between stages in the empirical matrices by assuming that the vital rates are at their optima.

In our model, we assume that different provenances of the same species have different climatic optimum, but they have the same vital rates as long as they are each at their optimal temperature. For each species, we investigate how different provenances perform in the face of climate change. We vary the thermal optimum of the provenance  $\theta$  between the current temperature at the planting site  $T_0$  to  $T_0 + 4^\circ\text{C}$ , in increments of 0.01 degrees. The thermal optima of the provenances span the range of temperatures predicted over the life span of the trees for all warming scenarios considered (see below and Table 1.1). Each simulation iterates equation (1.2) for  $H$  years and computes the number of individuals that have reached the last stage ( $n$ ) at harvest,  $N_H$ . We then numerically look for the provenance whose thermal optimum ( $\theta$ ) results in the largest number of survivors, called the best provenance. Stand rotation periods in forestry vary widely among species, climates, and management practices, from less than 12 years for some fast-growing tropical species to more than 150 years for some oak stands (Bauhus et al. 2009). In the simulations, we consider a mid-range rotation period, with harvest occurring after  $H = 60$  years (some simulations were run with longer rotation time in Supplementary Material S6).

Simulations are run in Wolfram *Mathematica* (Version 12.0.0.0) while numerical searches for the best provenance are conducted in R (Version 4.2.3). Both scripts can be found in Supporting Information (Code S1 and Code S2). We checked that simulations, described above, and the analytical predictions from equation 1.8 yield identical results for a two-stage life cycle (Supplementary material S3; Suppl. Fig. 4).



**Species selection** We filter the COMPADRE database by “OrganismType” to obtain demographic parameters for tree species. As data on local adaptation for tropical tree species are scarcer and the impact of climate change is less straightforward (Aitken and Bemmels 2016; IPCC 2021), we also filter by “Ecoregion” to keep only non-tropical trees. The following ecoregions are retained in our data set: Temperate Broadleaf and Mixed forests (TBM), Temperate Coniferous Forests (TCF), Boreal forests and taiga (BOR), Montane grasslands and shrublands (MON), Temperate grasslands, savannas, and shrublands (TGS), Mediterranean forests, woodlands and scrubs (MED). We obtain 47 species in total. From these, we select 19 species which had complete matrices and for which individuals had reached the last stage before the time of harvest (set to 60 years). When several matrices are available for a species, we keep those that pooled all years, locations, or treatments to consider typical growth trajectories. We also add a matrix for American beech (*Fagus grandifolia*) originally described in Harcombe (1987). To test that our qualitative conclusions hold for trees with a slower growth and thus longer harvest time, we use a harvest time of 100 years and three matrices of species that reach the last stage after 60 years and before 100 years (see Supplementary Material S6). Details about the matrices used can be found in Supporting Information (Data S1).

**Modelling the change in tolerance across the life span of trees** The breadth of thermal tolerance in trees is typically estimated through translocation and provenance tests in common gardens (Aitken and Bemmels 2016; Aitken et al. 2008; Mátyás 1996; Schmidting 1994; Sork et al. 2013; Wang et al. 2010). We unfortunately failed to find such data on survival as a function of climatic distance for any of the species that we have selected in the COMPADRE database. Consequently, we cannot estimate thermal tolerance and how it changes across life history stages for any of the species with demographic data (see Appendix A for details on the data collected and the attempt to estimate the change in thermal tolerance with tree age). In the absence of such data, we explore the theoretical impact of changes in thermal tolerance on harvest using a simple hypothetical linear model for variation in thermal tolerance with stage:  $\omega_i^2 = \omega_1^2 (1 + b(i - 1))$ . A positive or negative value of  $b$  corresponds to a scenario where older or younger trees are more tolerant to temperature changes, respectively.

We investigate the effect of both scenarios of increasing or decreasing tolerance with age with our analytical model for a life cycle with two stages. Because this analysis shows the results to be quite symmetrical, and because younger trees are expected to be less tolerant

than fully grown trees, we focus only on cases with positive  $b$  in the simulations. We hold the cumulative thermal tolerance  $\omega_{\text{tot}}(\vec{a})$  constant across species and also within species when we contrast different scenarios for change in thermal tolerance across stages: we then vary the value of  $b$ , which, for a constant  $\omega_{\text{tot}}(\vec{a})$ , determines the value of the thermal tolerance of the first stage ( $\omega_1$ ).

To calculate stage-specific tolerances for each species, while holding the cumulative thermal tolerance  $\omega_{\text{tot}}(\vec{a})$  constant, we approximate the typical life history  $\vec{a}$  of each species by calculating the expected time spent in each stage  $i$  from the fundamental matrix  $\mathbf{F}$ , such that  $\mathbf{F} = (\mathbf{I} - \mathbf{M})^{-1}$  (see Caswell et al. 2018). The expected age of maturation from stage  $i$  to stage  $i + 1$  is  $a_i = \sum_{j=1}^i \mathbf{F}_{1,j}$ . These values are rounded to the next integer to obtain a typical life history for that species. We then use such values of  $a_i$  and replace the stage-specific tolerances by  $\omega_i^2 = \omega_1^2 (1 + b(i - 1))$  in equation 1.5, to obtain the value of  $\omega_1$  as a function of  $b$  and  $\omega_{\text{tot}}(\vec{a})$  for each species. This method to compute the average ages at transition between stages is not conditional on the survival of individuals to the last stage, so the life history used to estimate  $\omega_{\text{tot}}(\vec{a})$  may differ in some cases from the realized life history of harvested individuals. We find that, as a result, the effective cumulative tolerance of species is on average  $0.1^\circ\text{C}$  higher than the theoretical cumulative tolerance, with 95% of effective cumulative tolerance values less than  $0.5^\circ\text{C}$  above the theoretical cumulative tolerance initially set.

In the different figures, we represent the extent of change in thermal tolerance during ontogeny by the ratio of thermal tolerance in the last and first stage  $\omega_n/\omega_1$ . Note that varying this ratio within each species helps evaluating how much the choice of the best provenance depends on the (unknown) change in thermal tolerance across life. This ratio is actually likely to vary between species, in particular because of the variable definition of the first stage in the different tree life cycles from the COMPADRE database.

We replicate these simulations for 3 different values of the cumulative thermal tolerance  $\omega_{\text{tot}}(\vec{a})$ : 2, 3.5 and  $5^\circ\text{C}$ . The choice of order of magnitude for these values is inspired by the results of Rehfeldt et al. (1999). In that study, one-year-old seedlings of *Pinus contorta* were planted and survival was estimated at age 20 for two different provenances, yielding an estimate of the cumulative thermal tolerance of  $\approx 2.75^\circ\text{C}$  over 20 years. If thermal tolerance did not vary with age, this would predict a cumulative tolerance over 60 years of less than  $1^\circ\text{C}$  according to equation 1.5. If tolerance increases with age as we suspect, we however expect the cumulative tolerance over 60 years to exceed the latter estimation. The chosen values contrast situations

where the climatic niche of trees is relatively narrow, intermediate or wide, while being not widely unrealistic given extant evidence on survival tolerance to temperature in forest trees.

**Scenarios of environmental change** To better understand the interactions between the rate of warming and the species life cycle affecting the choice of the best provenance, we simulate scenarios where each species is submitted to the same average warming trend. We first consider a linear increase in mean annual temperature, at speed  $k$ , such that  $T_t = T_0 + kt$  (average warming in 60 years in table 1.1), where  $k = \frac{T_H - T_0}{H}$ .

We parameterize the rate of change  $k$  in mean annual near-surface air temperature using the latest climate change projections available, the CMIP6 climate models, available in the Intergovernmental Panel on Climate Change (IPCC) WGI Interactive Atlas. The projected mean increase in annual mean air temperature for the WGI reference regions can be downloaded as GeoTIFF maps from the compilation of 34 different climate models at <https://interactive-atlas.ipcc.ch>. We consider three temperature warming scenarios estimated in the IPCC Sixth Assessment Report (Table SPM.1; IPCC 2021): SSP2-4.5, SSP3-7.0 and SSP5-8.5. We choose to use the period 1995-2014 as a baseline and the projected temperatures in 2081-2100. To approximate the temperature increase in the regions where the species occur, instead of using the global average, we retrieve the predicted increase in mean annual air temperature at the geographic coordinates given for each demographic transition matrix that we downloaded from COMPADRE (see Supplementary Material S5; Suppl. Fig. 5). From this, we compute the predicted rate of increase in mean annual air temperature at each species location per year and average this value over all locations to obtain  $k$ . We also use a scenario with no warming as a control.

Table 1.1: Temperature warming scenarios

Scenario	IPCC scenario	Average warming in 60 years
Control	-	+0°C
Intermediate	SSP2-4.5	+1.7°C
High	SSP3-7.0	+2.6°C
Very high	SSP5-8.5	+3.3°C

To test the robustness of the results to random variation in mean annual temperature around that trend, we also run simulations where each year the temperature fluctuates around the

intermediate warming trend (+1.7°C). Each year, the value of the annual mean temperature is randomly drawn from a normal distribution centered on the value of  $T_t$ , which increases linearly over time.

We consider three temperature fluctuation scenarios based on the estimates of Olonscheck et al. (2021). Inspired by their reported historical global average of observed and simulated local standard deviations of 0.44 and 0.47°C, we explore a scenario, which we call moderate fluctuations, with the standard deviation around the mean annual temperature of 0.5°C, assuming no major change in the future in the extent of year-to-year fluctuations in temperature. Different climate models examined in Olonscheck et al. (2021) consistently predict decreased year-to-year fluctuations in mean temperature at high latitudes and increased fluctuations in the tropics by the end of the century, but the predictions do not converge in many parts of the world. Given this high uncertainty, we also explore two scenarios with 30% weaker or 30% stronger fluctuations (corresponding to standard deviations of 0.35 °C and 0.65 °C, respectively).

For the sake of illustration, we run the simulations with fluctuations around the mean annual temperature for two species with contrasting life histories (*C. decurrens* and *A. concolor*, both with  $n = 5$  stages). We replicate each scenario 100 times to explore different temperature trajectories and determine the best provenance for each replicate. We here simulate cohort survival for a broader range of provenances thermal optimum (ranging from -5°C below to 5°C above the current local temperature) to account for instances of low temperatures that may occur stochastically in some runs. We consider two cases: constant thermal tolerance over the trees lifespan ( $\omega_5/\omega_1 = 1$ ) and lower temperature tolerance in young trees compared to old ones ( $\omega_5/\omega_1 = 9$ ).

## Results

### Analytical results in a two-stage case

To better understand the impact of changing thermal tolerance across the lifespan, we first explore a simplified two-stage case with juvenile and mature trees, for which general analytical predictions are illustrated. While a two-stage model is a gross simplification of the life cycle of trees, this simple case has the advantage that life history impacts the best provenance through a single parameter,  $a$ , the age at which trees transition to the adult stage. This toy model will help us better understand the results for more complex life cycles. Fig. 1.1 shows how

the number of surviving trees in a cohort with a given life history varies depending on the thermal optimum of the planted provenance. Our equations predict that, quite intuitively, the best provenance to plant has a higher thermal optimum ( $T_{\text{tot}}(\vec{a})$  as in equation 1.6, *i.e.*, is more warm-adapted, when climate change is faster or when rotation time is longer (as illustrated in Fig. 1.1). The best provenance, however, does not vary with the cumulative thermal tolerance ( $\omega_{\text{tot}}(\vec{a})$ ; compare panels A and B in Fig. 1.1), depending only on the relative change in thermal tolerance across the life span as shown by equation 1.5. Interestingly, even when planting the best provenance, the maximal number of surviving trees in the cohort (*i.e.*,  $s_{\text{tot,max}}(\vec{a})$  as in equation 1.7) is less in scenarios where the climate changes within the course of trees life than in the absence of climate change (compare coloured curves in Fig. 1.1). Our model predicts that assisted gene flow can therefore, at best, only partly mitigate the negative consequences of a warming climate. The part of unavoidable loss in harvest increases with more rapid climate warming, longer time before harvest (compare full to dot-dashed line in Fig. 1.1), and when trees have narrower climatic niche (*i.e.*, smaller cumulative tolerance over their lifetime, compare panels A and B in Fig. 1.1).

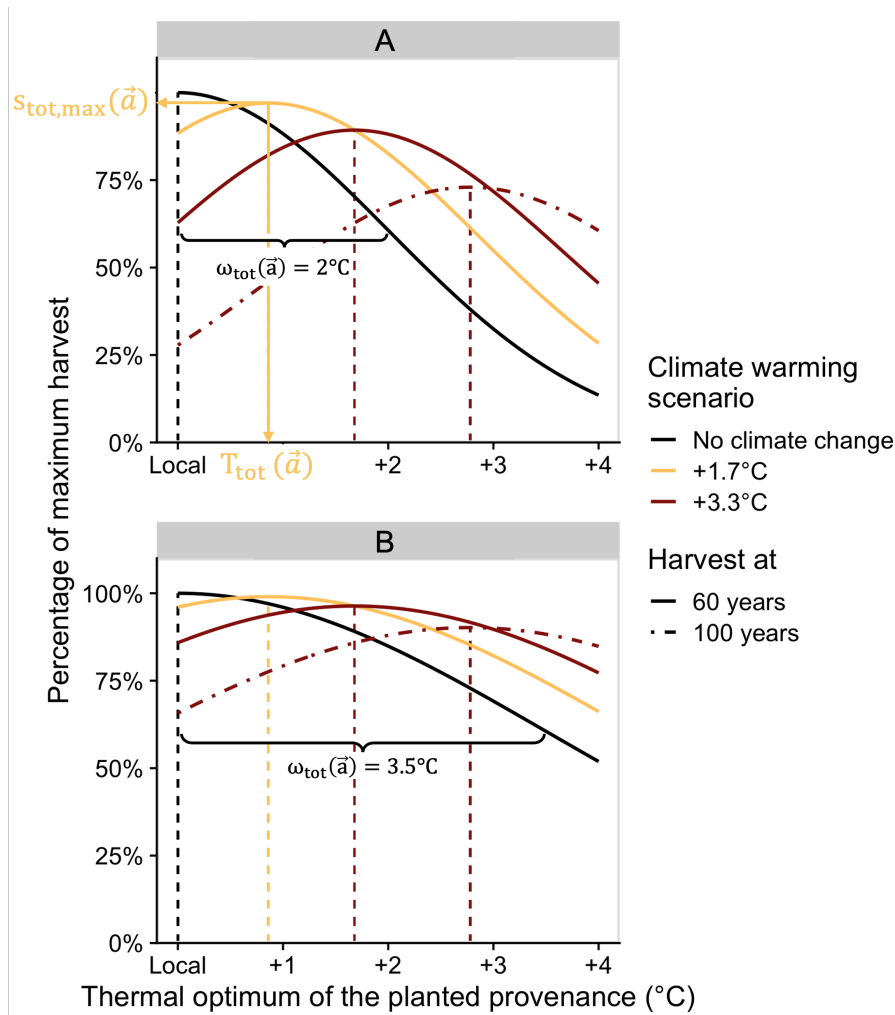


Figure 1.1: The potential loss in harvest from planting sub-optimal provenances is greater if trees have low cumulative tolerance (panel A versus panel B), especially under more severe climate warming scenarios in 60 years (colored curves) and if time before harvest is longer (dot-dashed curves versus solid curves). For each scenario, the dashed vertical lines represent the thermal optimum of the best provenance. For the climate warming scenario of  $+1.7^\circ\text{C}$  in panel A is indicated the best provenance to plant ( $T_{\text{tot}}(\vec{a})$ ), the maximal number of surviving trees in the cohort ( $s_{\text{tot,max}}(\vec{a})$ ) and the cumulative thermal tolerance ( $\omega_{\text{tot}}(\vec{a})$ ). The latter is also indicated on the same curve in panel B, for comparison. For the scenario of harvest after 100 years, the very high warming scenario (of  $3.3^\circ\text{C}$  in 60 years) gives a warming of  $5.5^\circ\text{C}$  in 100 years. The curves were produced with the two-stage analytical result with the following parameters:  $s_1 = 0.61$ ,  $s_2 = 0.98$  and  $g = 0.012$ .

Fig. 1.2 shows how life history influences the best provenance to plant. If thermal tolerance is equal in both stages ( $\omega_2/\omega_1 = 1$ ), the thermal optimum of the best provenance does not depend on the age of maturation from juvenile to mature tree ( $a$ ). When tolerance changes across the trees life, the thermal optimum of the best provenance is however pulled towards the temperatures experienced in the least tolerant stage, with the extent of this pull determined by the duration of this stage (Fig. 1.2). Time spent in the least tolerant stage furthermore has a

non-monotonic effect on the choice of the best provenance: when juveniles are less tolerant to temperature changes and the juvenile period is short, increasing the length of the juvenile period makes the choice of cooler-adapted provenance more advantageous, but when this juvenile period is very long it is best to select warmer-adapted provenances. If time spent in either phase is very long (such that the trees experience most years in that stage), the thermal optimum of the best provenance is close to the average temperature experienced during the period. Differences in thermal tolerance across life stages matter most when the age of the transition between stages is intermediate, such that a significant part of life is spent in each stage.

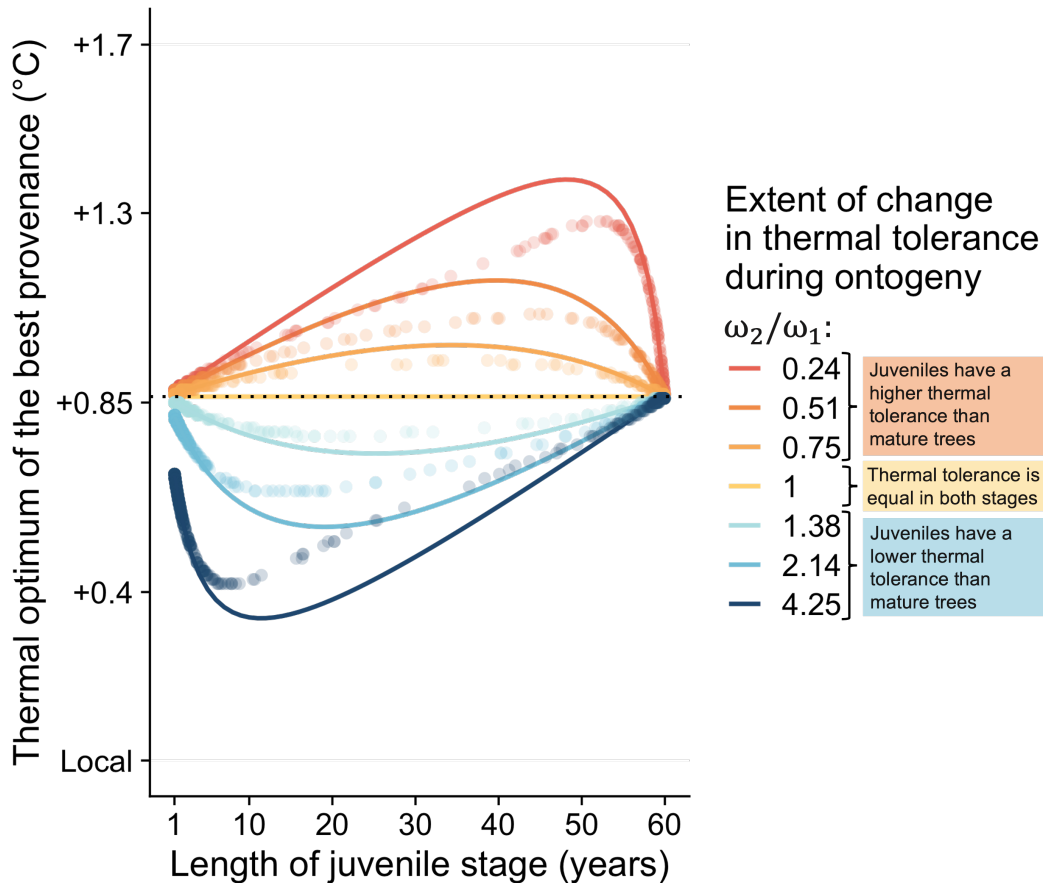


Figure 1.2: The best provenance is determined by variation in thermal tolerance across life stages (colored curves), and the duration of the juvenile phase. The analytical results from a two-stage model are illustrated with a climate warming scenario of  $+1.7^\circ\text{C}$  in 60 years. The extent of the change in thermal tolerance during ontogeny is represented by the ratio between the thermal tolerance of mature trees and juvenile trees ( $\omega_2/\omega_1$ ) shown with colors. Solid lines represent the thermal optimum of the best provenance using the age-structured analysis, *i.e.*, assuming the whole cohort grows from juveniles to mature trees at age  $a$  (predicted using equation 3). Dots show the thermal optimal provenance in a stage-structured framework, *i.e.*, when individuals stochastically can grow at different ages (predicted using equation 8), for 500 sets of random values for  $g$ ,  $s_{1,\max}$  and  $s_{2,\max}$  drawn from uniform distributions between 0.01 and 0.99, conditional upon survival. The x-coordinate gives the average maturation age among individuals surviving to harvest for each set of life history parameters. The dotted horizontal line indicates the average temperature between planting and harvest.

The above line of reasoning applies when all individuals have the same life history and should work well to predict the thermal optimum of the best provenance in a stage-structured population if there is little variation in life history among those that survive until harvest. In Fig. 1.2, we compare the thermal optimum of the best provenance as a function of the mean age of maturation among survivors at harvest (dots) in the stage-structured analytical predictions for a fixed age of maturation between stages (curves). All previous qualitative conclusions about the thermal optimum of the best provenance, assuming a fixed life history, hold in the stage-



structured model, and the quantitative predictions assuming a constant life history lack precision only when the age of transition between stages is intermediate. In those cases, the mean age at transition hides a great deal of variation in life histories with some individuals maturing early and others late. Stochastic realizations of each individual's life history then tend to lengthen one stage or the other, causing the optimum in the stage-structured model to be a bit less sensitive to differences in thermal tolerance than the strictly age-structured model (dots are closer to the case where  $\omega_1/\omega_2 = 1$  than the lines in Fig. 1.2).

We find that relaxing our assumption of a homogeneous cohort (*i.e.*, increasing the phenotypic variance at the plantation site) is unlikely to change our predictions about the best provenance much (Eq. S1.1b and Suppl. Fig. 1.1). In the scenario of a variable cohort, we observe that increasing the initial phenotypic variance decreases the expected number of survivors when planting the best provenance, but that higher phenotypic variance can minimize the loss of survivors when planting a subpar provenance, for example if climate deviates strongly from our expectation (Eq. S1.1b, Suppl. Fig. 1.2).

## Simulation results

**Optimal provenance for 20 non-tropical tree species in a warming climate** In accordance with the analytical predictions, when the thermal tolerance is constant over the lifetime of the trees ( $\omega_i/\omega_1 = 1$ ), the thermal optimum of the best provenance matches the average temperature over the period and does not vary across species (but is warmer if temperatures rise faster) (Fig. 1.3). When young trees are less tolerant to temperature changes than old trees ( $\omega_i/\omega_1 > 1$ ), the thermal optimum of the best provenance decreases, as predicted analytically, and becomes increasingly different among species depending on their life history (see Supplementary Material S6 for similar conclusions for species with longer rotation times). The optimal provenance for species that spend many years in the first and least tolerant stage is cooler and much nearer current temperatures (darker curves in Fig. 1.3) than for species that mature earlier (lighter curves). The ranking of species can reverse if this first stage is both poorly tolerant to climate variation and prolonged, as young individuals will then be exposed to warmer temperatures. This pattern is consistent with the non-monotonic effect of changing the age at transition between stages shown in Fig. 1.2. We however caution that stage-specific temperature sensitivities are likely to differ between species and that empirical data on  $\omega_i$  across the lifespan of trees is needed for species-specific predictions.

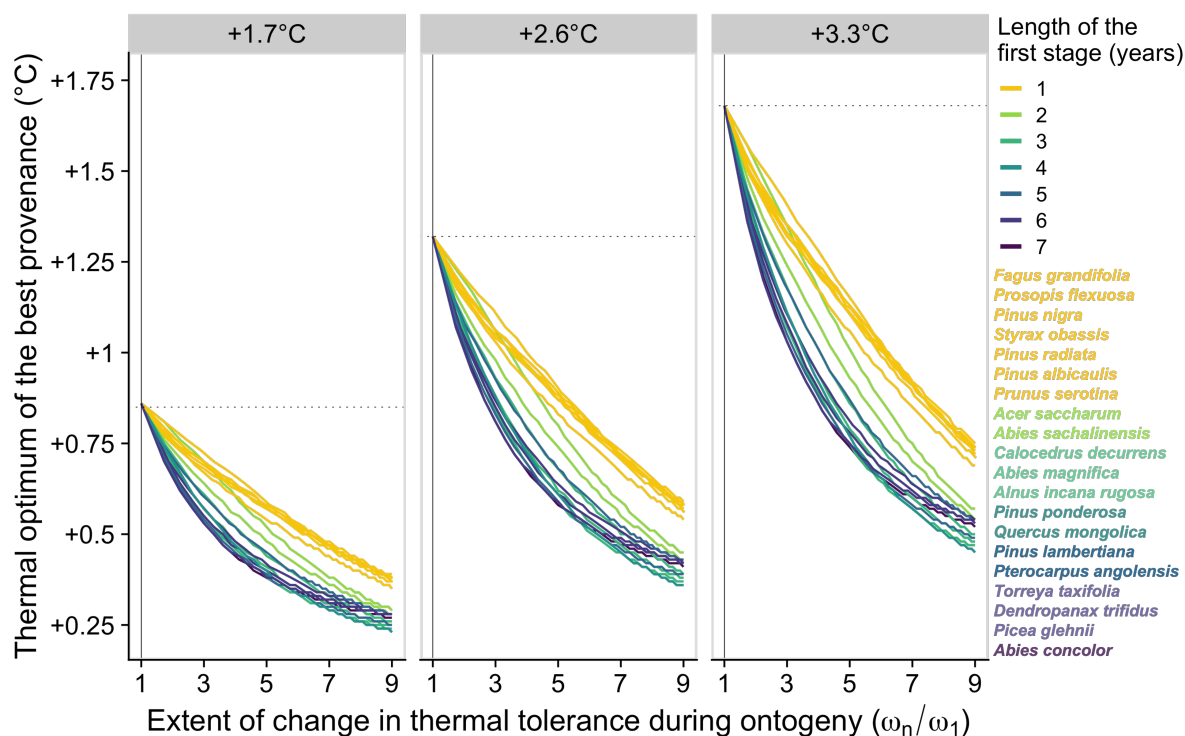


Figure 1.3: Variation in the best provenance to plant is illustrated using the life history trajectories for twenty long-lived tree species. The less tolerant younger trees are to changes in temperature compared to older trees (the larger  $\omega_n/\omega_1$  along the x-axis), the cooler the best provenance. The effect is shown for 3 warming scenarios (panels). Each curve represents a species and is coloured by the length of their first life stage (transition matrices that considered seeds or seedlings spend only one year in this stage). The vertical line indicates the case where tolerance is constant across all stages. For each warming scenario, the dotted horizontal lines correspond to the average temperature over the period.

**Comparing the performance of different seed sources** We seek to assess the accuracy of using simple rules for provenance selection, considering that critical parameters, such as how thermal tolerance changes over a species lifespan, are often unknown. In our simulations, we consider three seed sources: (1) the “local” source, which we here assume to be adapted to the current climate with optimal temperature matching the expected temperature when the cohort is planted (2) the model-predicted best provenance, and (3) a provenance adapted to the predicted temperature at one quarter of the rotation time (*i.e.*, in the present scenario in 15 years) (O’Neill et al. 2017). We compare the predicted harvest loss when planting each seed source (Fig. 1.4 and Supplementary Material 1 - S7; Suppl. Fig. 1.7). For all species examined, our model predicts that the benefit of planting the best provenance (yellow) compared to the local provenance (black) is significant when all stages are assumed to be equally tolerant to temperature changes: using the best provenance instead of the local provenance then allows

reducing loss from 4% to less than 1% in Fig. 1.4 and these gains are considerably greater if climate change is faster and/or cumulative thermal tolerance is set to be smaller (Supplementary material 1 - S7; Suppl. Fig. 1.7). The benefits of assisted gene flow are however predicted to diminish when younger trees are much less tolerant than older trees ( $\omega_n \gg \omega_1$ ), with the best provenance performing only slightly better than the local provenance (Fig. 5 and Supplementary material S7; Suppl. Fig. 1.7). This happens because the best provenance is then cooler-adapted and less different from the local provenance. Our model predicts that the recommendation from O'Neill et al. 2017 (blue) performs nearly as well as the best provenance (yellow) when juveniles are less tolerant to temperature changes than older trees, but not quite as well if all stages are equally tolerant (at most the loss of harvest is 2% as opposed to the 1% for the best provenance in Fig. 1.4).

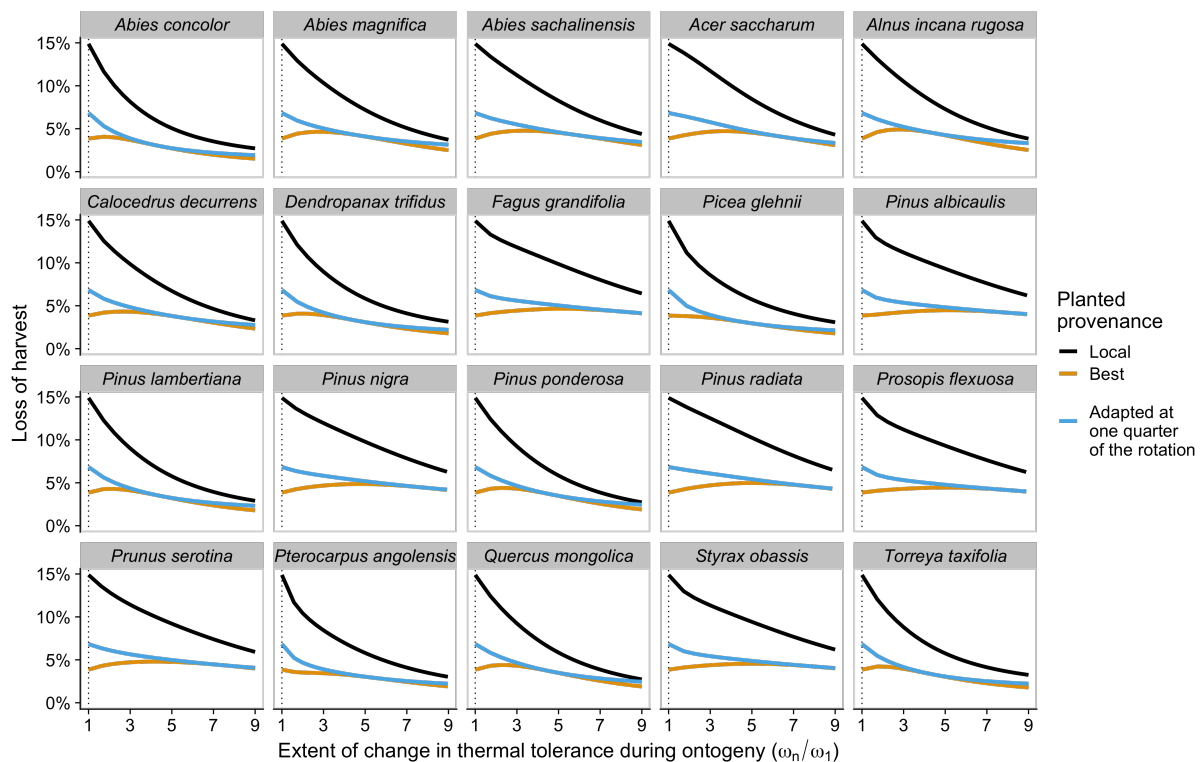


Figure 1.4: Large losses in harvest could be avoided for all species by planting the best provenance (yellow) or the provenance adapted to the temperature at one quarter of the rotation time ( $+0.43^\circ\text{C}$ ; blue) as opposed to planting the local provenance (black), but the benefit is reduced if young trees are less tolerant to changes in temperature as compared to older trees ( $\omega_n/\omega_1$  increases) in an intermediate climate warming scenario of  $+1.7^\circ\text{C}$ . The cumulative thermal tolerance is held fixed across the life span at  $\omega_{\text{tot}}(\vec{a}) = 3.5^\circ\text{C}$ . Harvest loss is measured against the performance of the local provenance in a scenario with no climate change. The vertical dotted line indicates the case where tolerance is constant across all stages.

**Climate fluctuations** We assess the robustness of the previous predictions by letting the annual temperature fluctuate from year to year around a warming trend. Fluctuations in the average annual temperature result in a reduction in harvest on average (Fig. 1.5). We however find that the provenance that performs best on average is not different from the one estimated without fluctuations (Fig. 1.5). There is nonetheless a lot of variability among runs (and thus uncertainty) in the thermal optimum of the best provenance with temperature fluctuations, indicating that the realized climate trajectory may randomly have been warmer or colder than expected, which would favour a warmer or cooler provenance, respectively. This variance in best provenance across realized climate trajectories is much larger when stages have different thermal tolerances (Fig. 1.5B and Fig. 1.5D) and when the least tolerant stage lasts only for a few years (compare Fig. 1.5B and Fig. 1.5D): when the first stage is poorly tolerant to changes in temperature, the best provenance depends primarily on the temperatures experienced randomly during the first few years of life: individuals must be well adapted to the temperatures in their first years to survive. In particular, if those early years are unusually cold, the thermal optimum of the best provenance to plant may even be cooler than the local provenance. The variability in harvest associated with interannual fluctuations is however smaller if diversified provenances are planted (Suppl. Fig. 2).

We assess the robustness of the previous predictions by letting the annual temperature fluctuate from year to year around a warming trend. Fluctuations in the average annual temperature result in a reduction in harvest on average (Fig. 1.5). We find, however, that the provenance that performs best on average is not significantly different from the one estimated without fluctuations (Fig. 1.5). There is nonetheless significant variance among runs (and thus uncertainty) in the thermal optimum of the best provenance with temperature fluctuations, indicating that the realized climate trajectory may randomly have been warmer or colder than expected, which would favour a warmer or cooler provenance, respectively. This variance in best provenance across realized climate trajectories is much larger when stages have different thermal tolerances (Fig. 1.5B and Fig. 1.5D) and when the least tolerant stage lasts only for a few years (compare Fig. 1.5B and Fig. 1.5D): when the first stage is poorly tolerant to changes in temperature, the best provenance depends primarily on the temperatures experienced randomly during the first few years of life: individuals must be well adapted to the temperatures in their first years to survive. In particular, if those early years are unusually cold, the thermal optimum of the best provenance to plant may even be cooler than the local provenance.

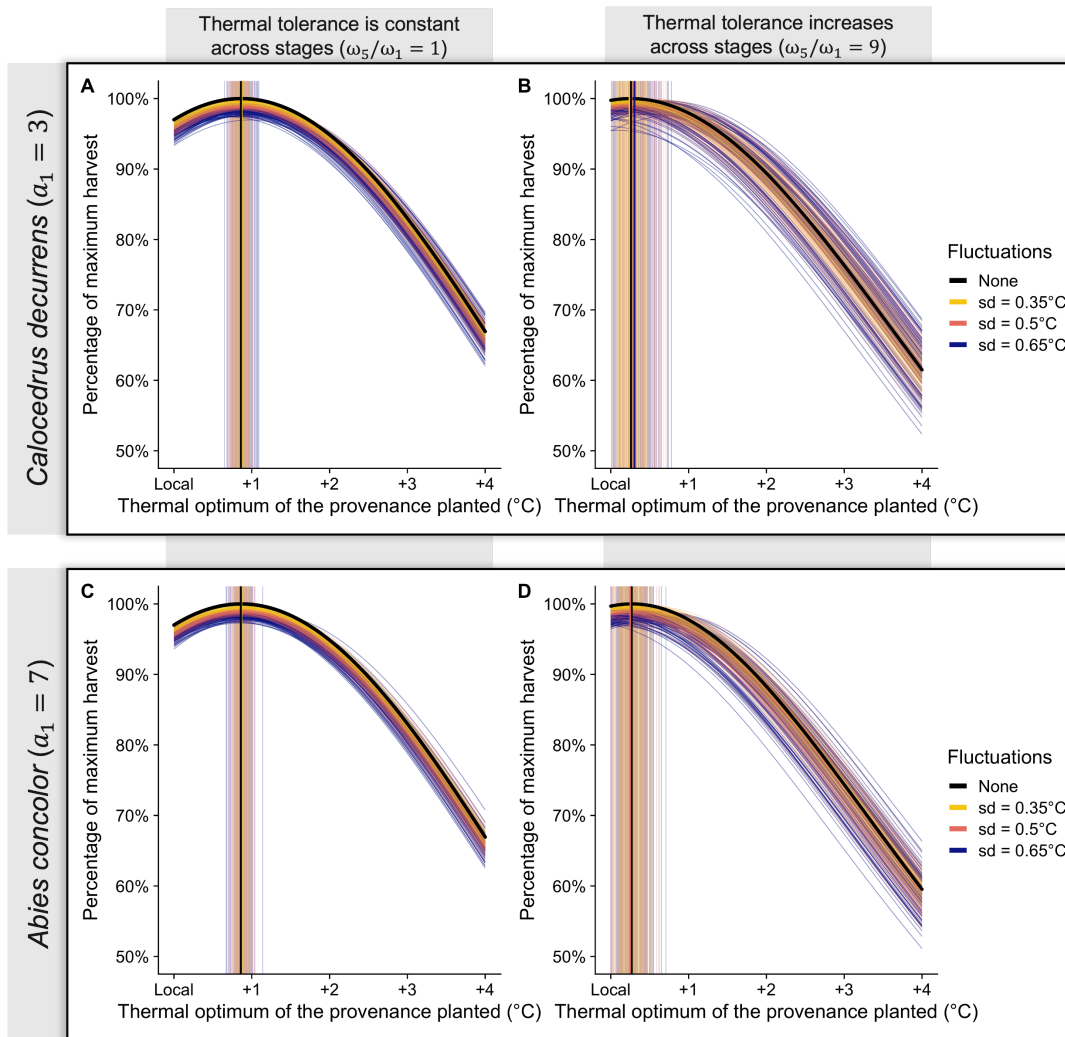


Figure 1.5: Random interannual fluctuations can lead to a reduced harvest for *Calocedrus decurrens* (panels A and B) and *Abies concolor* (panels C and D), but with increased variability when young trees are less tolerant to changes in temperature as opposed to older trees ( $\omega_5/\omega_1 = 9$  for panels B and D versus  $\omega_5/\omega_1 = 1$  for panel A and C), and for *Calocedrus decurrens*, a species that remains in the first stage for a shorter time. Environmental temperature increases under the intermediate climate warming scenario of  $+1.7\text{ }^\circ\text{C}$  either linearly (black), with a variance given by the global average of locally observed values for interannual fluctuations ( $\text{sd} = 0.5\text{ }^\circ\text{C}$ ; pink) or, with lower or higher variance ( $\text{sd} = 0.35\text{ }^\circ\text{C}$ ; yellow, or  $0.65\text{ }^\circ\text{C}$ ; blue). Each scenario of climate warming with fluctuations is replicated 100 times with random draws of temperature around the warming trend. Each curve represents the harvest for each provenance relative to the harvest of the local provenance without climate change. The light vertical lines represent the thermal optimum of the best provenance for each run with fluctuations. The darker vertical lines represents the average of such thermal optima. The cumulative thermal tolerance across the life span is held fixed at  $\omega_{\text{tot}}(\vec{a}) = 3.5\text{ }^\circ\text{C}$ .

## Discussion

### **The best seed source for assisted gene flow in long-lived species depends on how tolerance to climate changes along life**

Several authors have suggested that assisted gene flow could mitigate the negative effects of climate change (*e.g.*, Browne et al. 2019; Milesi et al. 2019). Recent modelling however suggests that the short-term conservation benefits of assisted gene flow may often be modest in large populations, unless pre-adapted genetic variants of large effects are introduced at large frequencies (Grummer et al. 2022). This makes the identification of pre-adapted sources of seeds particularly critical for the success of assisted gene flow in forest trees. The choice is indeed complicated by a trade-off in long-lived species, where adaptation to a warmer climate at the adult stage may come at the cost of poor adaptation to cooler climates while still juvenile. We have modeled the survival of a single cohort of trees in a changing climate and showed that, to maximize the number of surviving trees, the choice of the best provenance critically depends on differences between life stages in their tolerance to climate. When thermal tolerance is constant over the lifespan of a tree, the best provenance is the one whose thermal optimum corresponds to the average temperature observed over this period, which increases with the rate of climate warming before harvest. When there is stage-specific thermal tolerance, the best provenance is pulled towards the temperature experienced in the least tolerant stage, weighted by how many years are spent in that stage. The thermal optimum of the best provenance is cooler than the average temperature across the period when young trees are less tolerant to changes in temperature and warmer when older trees are less tolerant. This effect of time spent in the least tolerant stage is complicated, because, the longer individuals stay in this stage, the more variable temperatures they experience while in that stage. We showed the predictions of this simple model to be quite robust to variation in the trees life cycle and the number of stages, the presence of variation in climatic niche within the provenance and the stochastic fluctuations of climate around the warming trend. While we have focused on temperature for the sake of illustration, the same conclusion would apply for any climatic variable or combination of climatic variables affecting tree growth and survival (*e.g.*, precipitation).

## **We lack information on critical parameters to determine the best seed source**

Our model shows that changes in tolerance during the life of the species are particularly relevant in determining the best provenance for assisted gene flow. Yet, quantitatively estimating these changes in different species throughout the life of a tree is not trivial. There is a need for further empirical estimates of how tolerance to climatic variables varies with tree ontogeny and affects the growth and survival of trees at different life stages. Some studies have attempted to answer these questions by looking at changes in the stage distribution of trees across space along climatic gradients (*e.g.*, Bell et al. 2014; Lenoir et al. 2010; McLaughlin and Zavaleta 2012), but lags in responses to climate change confound the interpretation of these patterns (Heiland et al. 2022). Long-term provenance trials, where different seed sources are planted on multiple sites, are invaluable for assessing the risks and benefits of assisted gene flow. While there are older trials for some common species (*e.g.*, Carter 1996; Rehfeldt et al. 1999; Schmidting 1994; St. Clair et al. 2022), establishing new common gardens, with responses to climate changes in mind (such as Du et al. 2019; Kueppers et al. 2017; Munier et al. 2010), should be a priority. Another approach is the analysis of tree rings in regards to past climate change. Several such studies reveal that trees of different age and size vary in their growth response to past climatic variation (see Au et al. 2022; Carrer and Urbinati 2004; Depardieu et al. 2020; Housset et al. 2018; Latreille et al. 2017; Marquis et al. 2020; Mašek et al. 2021; Pompa-Garciéa and Hadad 2016). Unless the entire historical population can be sampled (including dead trees), studying only those individuals who survived past climatic fluctuations may however lead to an overestimation of tolerance in older age classes due to survivor bias (Duchesne et al. 2019, similar to the slow-grower survivorship bias reported in Brienen et al. 2012).

While younger life stages are generally thought to be less tolerant to climatic stress than larger trees (*e.g.*, Black and Bliss 1980; Kueppers et al. 2017; Munier et al. 2010; Pompa-Garciéa and Hadad 2016, though see Du et al. 2019 for an exception), several considerations may make our predictions about the best provenance more complex. First, younger and older stages may be sensitive to different climatic factors, rather than overall less or more tolerant to any stress (Mašek et al. 2021). Competition within and between species is a strong determinant of survival, especially in younger trees (Kunstler et al. 2021). Climate-associated mortality in young stages may be balanced by reduced density-dependent mortality. The effects of competition may therefore make seedling survival less sensitive to climatic conditions, favouring the choice of warmer-adapted provenances for plantation. The exact practices to implement assisted gene

flow could also affect our predictions: for instance, greenhouse or nursery cultivation of seedlings before transplantation makes their survival less affected by local temperature than trees regenerating from seed in colder natural environments.

Finally, even if our model correctly predicts the thermal optimum of the best provenance, information about thermal optimum is not available for many potential seed sources. Given that climate change has already increased temperatures by an average of around one degree C globally (IPCC 2021), current temperatures may not match thermal adaptation of local sources that are adapted to pre-Anthropocene temperatures (as for instance for *Quercus lobata*, see Browne et al. 2019). Predictions of climate adaptation based on genomic data and genetic-environment associations is increasingly considered as an alternative to the assumption of systematic local adaptation, but these methods are also in need of further validation (Capblancq et al. 2020). Our model also assumes that the temperature maximizing the survival of an individual is the same for all its life, which is not necessarily the case if juveniles and adults have different climatic preferences. While it is straightforward to modify our model to include this possibility, we critically lack information about differences in thermal optimum between adults and juveniles in most species.

### **Uncertainties remain high about the costs/benefits ratio of assisted gene flow**

Many uncertainties must be considered when discussing the balance of costs and benefits associated with shifting genotypes in response to or in anticipation of climate change (Srivastava et al. 2021).

The first level of uncertainty relates to the identification of the provenance that will perform best in a given climate scenario. Our model predicts how much tree mortality could be avoided by planting sources better adapted to future climate than the local source. Interestingly, we found that these losses can be reduced even when information to identify the best provenance is not available, by using a simple rule of thumb, such as planting provenances well adapted to the expected climate at one quarter of the rotation period. Using the best provenance also does not avoid all loss in a changing climate, because of the trade-off between early and late adaptation, described as the cost of within generation selection in Cotto and Chevin (2020). Our model predicts that planting warmer provenances in a warming environment can reduce tree mortality in a changing climate, but the benefits are only modest if climate change is not very fast, climatic tolerance is wide and young trees are much less tolerant to climate than old



ones. Costs of translocation could therefore outweigh the advantages when early life stages are less climate-tolerant than later stages. Indeed, populations are not only climate-adapted but also locally adapted to other factors such as photoperiod, pollinators, and pests (see Giencke et al. 2018 for an example of phenological mismatch and Wadgymar et al. 2022 for a review on local adaptation in plants). Translocated warmer provenances might, therefore, under-perform compared to local provenance even if they are better matched to the future climate. The benefits of assisted gene flow however increase if the local provenance is already maladapted to the current climate because of past climate warming: a warmer-adapted seed source may then increase the survival of both early and late stages.

Our model also reveals a second level of uncertainty, the fundamental uncertainty associated with interannual variation in climate and the prediction of future climate. Differences in climatic tolerance between stages make predicting the best provenance of seeds with interannual variation more difficult than when differences between stages are ignored. On average, the best provenance that maximizes survival is not affected by these fluctuations, but in any realization of a sequence of climatic years, the best provenance may differ greatly from this average. High fluctuations around the expected performance of a provenance will be especially pronounced when there are strong differences in climatic tolerance between stages and when sensitive stages are short and so do not average across many years of climate variation.

## Conclusions

This theoretical study helps answer questions about assisted gene flow in long-lived species. Our model suggests that planting seeds from warmer sites in a warming environment has the potential to mitigate mortality losses of a tree cohort, even under an optimistic climate change scenario. Importantly, the length of time spent in each stage and the tolerance of that stage to the expected climate can also be used to identify which stages are expected to experience high mortality, which may help managers allocate resources across the life cycle (*e.g.*, by thinning to reduce competition or spraying for pests during that stage). A great deal of uncertainty however remains about the optimal seed sourcing, particularly because variation in tolerance to different climatic variables as a function of tree age or stage are not well known. This deserves to be investigated further empirically, alongside more extensive modeling of assisted gene flow, simulating mixing seed sources and alternative regeneration approaches, to better quantify the

risks and benefits associated with assisted gene flow practices.

### **Acknowledgements**

This research project was initiated while O.R. was an invited research fellow at the Peter Wall Institute of Advanced Studies, and we gratefully acknowledge their support. O.R acknowledges support from CNRS for her secondment at the University of British Columbia and from the MUSE mobility program 2018-2019. The France-Canada Research Funds provided financial support for collaboration and travel between Vancouver and Montpellier and for A. E.'s PhD scholarship. O.R. and A. E. acknowledge support from the ANR project FloRes (ANR 22-CE02-0010-01). A CC-BY public copyright license has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission, in accordance with the grant's open access conditions.

### **Conflict of interest statement**

The authors declare no conflicts of interest.

### **Data availability statement**

The data that support the findings of this study is openly available in public databases (COM-PADRE - Plant Matrix Database, IPCC WGI Interactive Atlas). The scripts (Mathematica and R) are available in Supporting Information (Code S1, S2 and S3). *Note: These materials are not part of this thesis due to their length, but can be provided upon request.*

## Supplementary Material 1

### S1. Increasing phenotypic variance at the plantation site

We extend the two-stage framework developed in the main text, assuming here that phenotypes are normally distributed with mean phenotype  $\bar{\theta}$  and phenotypic variance  $P_0$ . We apply the age-structured result of Cotto and Chevin (2020) for multiple rounds of Gaussian selection on a single cohort (their equation 9) to write the expected number of mature trees at harvest time if all surviving trees had grown from juveniles to mature trees at age  $a$ , as:

$$E[N_H(\bar{\theta})|a] = N_0 s_{\text{tot,max}}(a) \sqrt{\frac{\omega_{\text{tot}}^2(a)}{P_0 + \omega_{\text{tot}}^2(a)}} \exp\left(-\frac{(\bar{\theta} - T_{\text{tot}}(a))^2}{\omega_{\text{tot}}^2(a) + P_0}\right) \quad (\text{S1.1a})$$

Considering all possible life histories until harvest, the remaining pool of individuals is formed by the sum of the cohorts of individuals making a transition at the same age  $a$ :

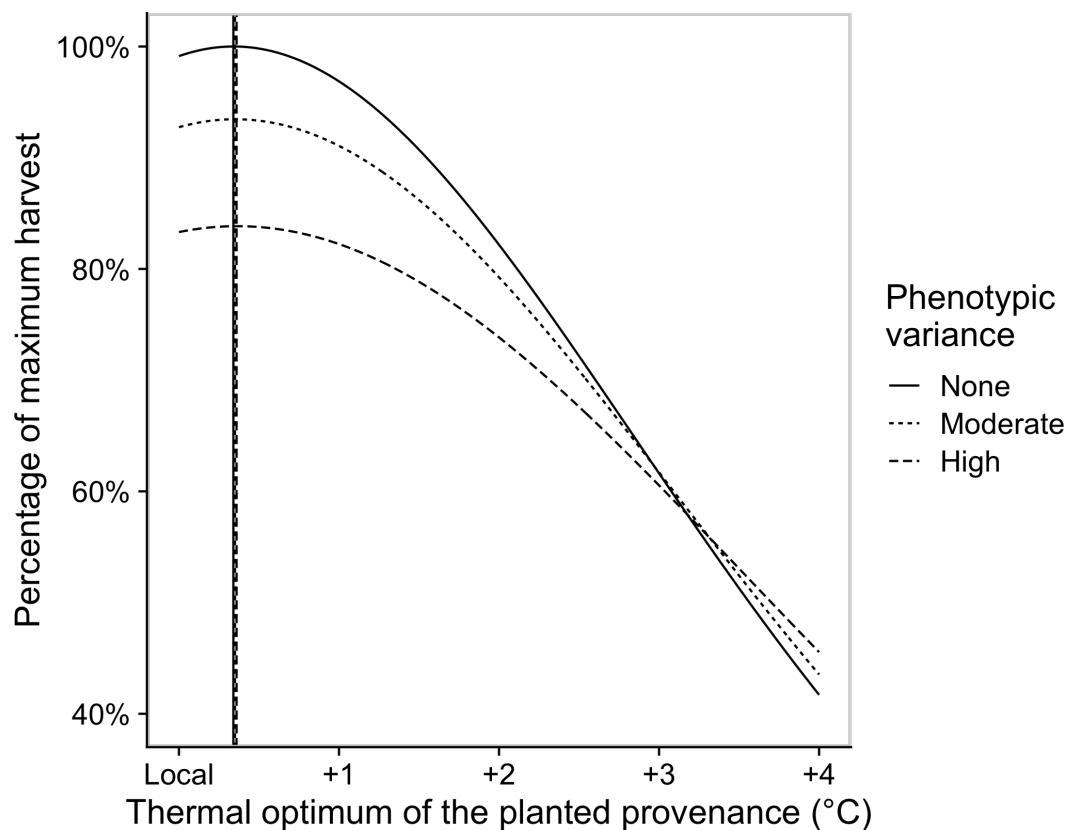
$$E[N_H] = g N_0 \sum_{a=1}^H (1-g)^{a-1} s_{\text{tot,max}}(a) \sqrt{\frac{\omega_{\text{tot}}^2(a)}{P_0 + \omega_{\text{tot}}^2(a)}} \exp\left(-\frac{(\bar{\theta} - T_{\text{tot}}(a))^2}{\omega_{\text{tot}}^2(a) + P_0}\right) \quad (\text{S1.1b})$$

The range of optimal annual temperatures maximizing survival over 20 years among different provenances of *Pinus contorta* was found to span over 7.6°C (Rehfeldt et al. 1999). We set the phenotypic variance ( $P_0$ ) in the plantation to 2°C<sup>2</sup> for a “moderate” variance scenario and to 6°C<sup>2</sup> for a “high” variance scenario, mimicking a situation where different provenances would be planted together in the cohort.

We find that increasing phenotypic variance at the planting site does not affect our prediction for the best provenance, but it does reduce the expected number of survivors when the best provenance is planted (Suppl. Fig. 1.1). This makes sense as fewer individuals will have the optimal phenotype in the initial cohort. When the planted provenance is very different from the best provenance, phenotypic variance among the planted trees conversely increases the harvest, because the cohort is more likely to contain some untypical individuals with a phenotype closer to the optimum.

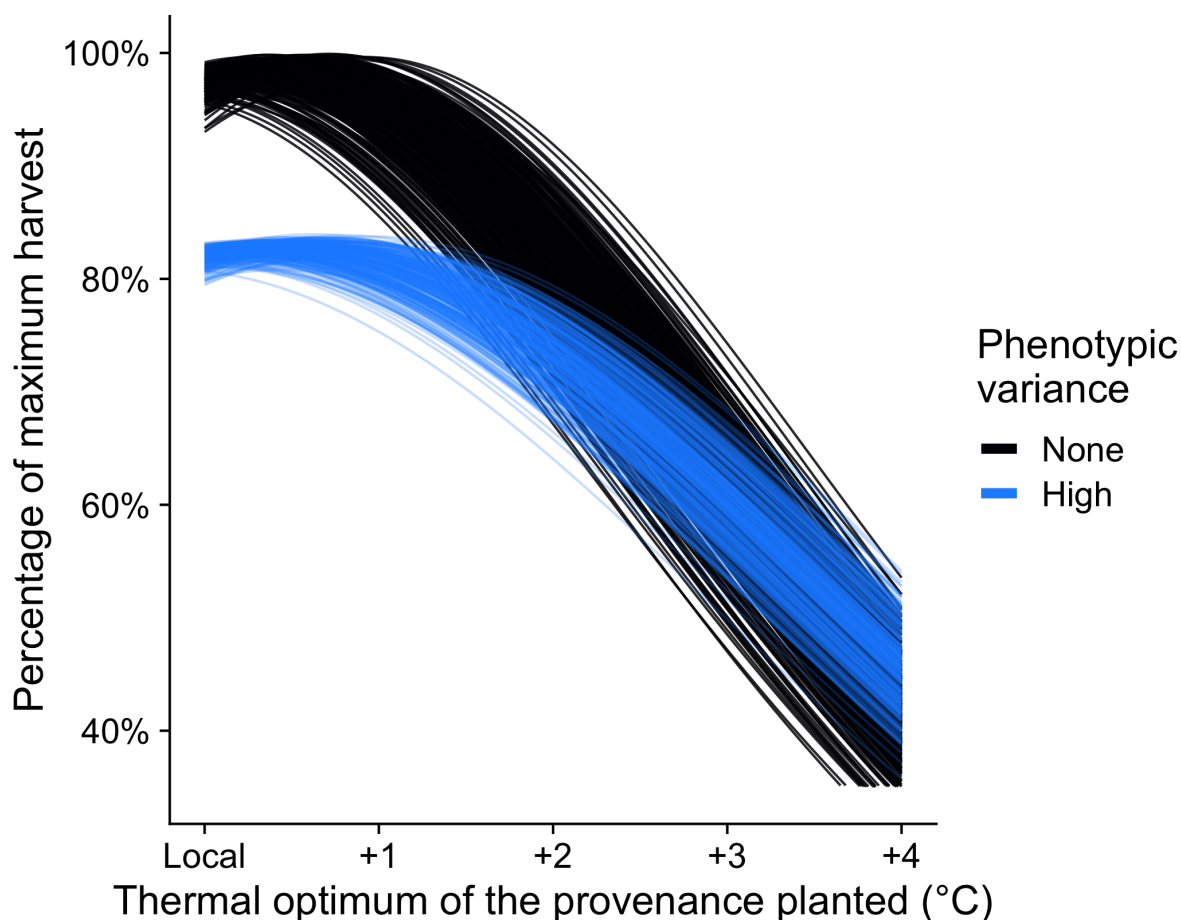
We next investigate if these conclusions hold in less predictable environments when the temperature fluctuates around a warming trend. In agreement with Suppl. Fig. 1.1, we find that increasing phenotypic variance tends to decrease the average harvest (Suppl. Fig. 1.2), but also decreases variance in harvest between runs and avoids very large loss when maladapted provenances are planted.

## Effect of increasing phenotypic variance



Suppl. Fig. 1.1: Increasing phenotypic variance at the planting site has no effect on the prediction of the thermal optimum of the best provenance (vertical lines), but it does reduce the harvest when the best provenance is planted. When the thermal optimum of the planted provenance is very far from that of the best provenance, however, phenotypic variance increases the expected harvest. This result is depicted for a climate warming scenario of  $+1.7\text{C}$  in 60 years and using parameters:  $s_1 = 0.61$ ,  $s_2 = 0.98$  and  $g = 0.012$ . Three scenarios of phenotypic variance are considered ("None",  $P_0 = 0\text{C}^2$ , solid line), moderate phenotypic variance ("Moderate",  $P_0 = 2\text{C}^2$ , short dash) or high phenotypic variance ("High",  $P_0 = 6\text{C}^2$ , long dash). Each curve represents the percentage of maximum harvest produced by each provenance in each scenario of phenotypic variance. The curves were produced with the two-stage stage-structured model (equation S1b). The cumulative thermal tolerance across the life span was held fixed at  $\omega_{\text{tot}}(\vec{a}) = 3.5\text{C}$  and the extent of change in thermal tolerance across the lifespan was set to  $\omega_2/\omega_1 = 46$ .

## Effect of increasing phenotypic variance when the climate fluctuates



Suppl. Fig. 1.2: Increasing phenotypic variance (blue lines,  $P_0 = 6^\circ\text{C}$ ) under random inter-annual fluctuations in mean annual temperature generally results in lower harvest than in the case without phenotypic variance (black lines,  $P_0 = 0^\circ\text{C}$ ) but phenotypic variance limits the loss in the case where the provenance planted is very different from the best provenance for a climate warming scenario of  $+1.7^\circ\text{C}$  in 60 years and using parameters:  $s_1 = 0.61$ ,  $s_2 = 0.98$  and  $g = 0.012$ . Environmental temperature increases with standard deviation around the trend of  $\text{sd} = 0.65^\circ\text{C}$ . Each scenario was replicated 300 times. Each line represents the percentage of maximum harvest produced by each provenance in each scenario of phenotypic variance and for each random climatic trajectory. Cumulative thermal tolerance across the life span was held fixed at  $\omega_{\text{tot}}(\bar{a}) = 3.5^\circ\text{C}$ . Thermal tolerance across the lifespan was set to  $\omega_2/\omega_1 = 46$ . The curves were produced with the two-stage stage-structured model (equation S1.1b).

## S2. Both survival and growth are sensitive to temperature

If both the survival and growth rates depend on temperature, we define, as before, the survival rate in stage  $i$  ( $s_i$ ), as a Gaussian function of the thermal optimum of the provenance  $\theta_s$  and the current temperature  $T_t$  in year  $t$ , with  $s_{i,\text{max}}$  the maximal survival rate in stage  $i$  and  $\omega_{s,i}$

the thermal tolerance for survival in stage  $i$ :

$$s_i(\theta_s, T_t) = s_{i,\max} \exp\left(-\frac{(\theta_s - T_t)^2}{2\omega_{s,i}^2}\right) \quad (\text{S1.2a})$$

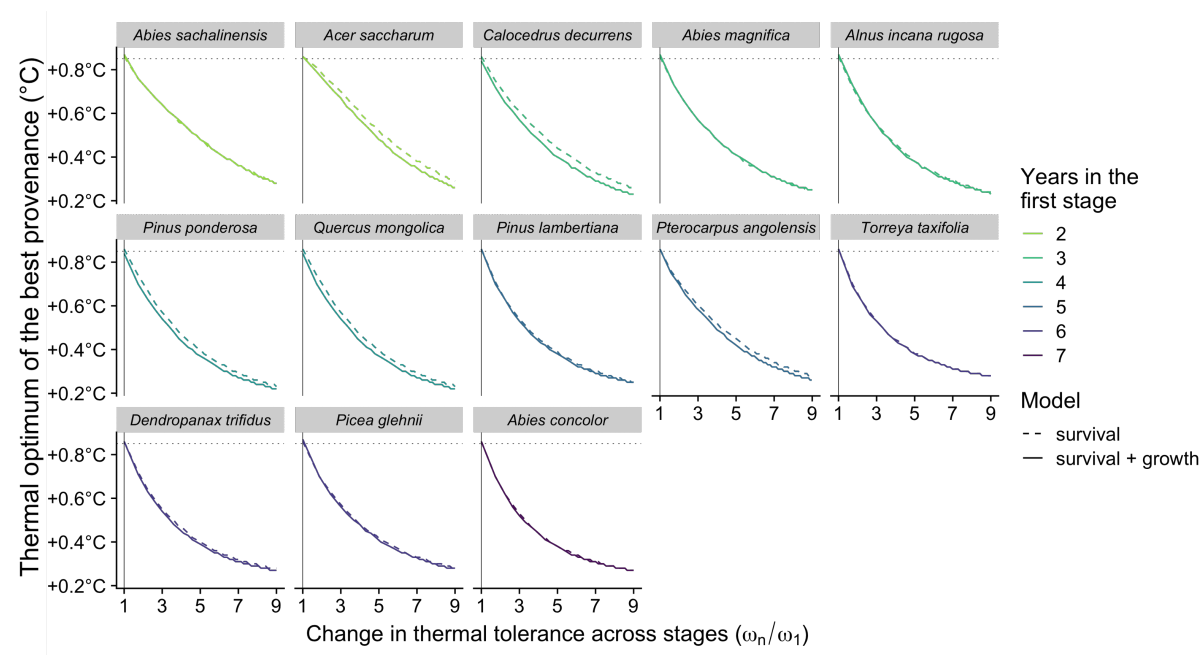
In this case, the growth rate in stage  $i$  ( $g_i$ ) is also expressed as a Gaussian function of the thermal optimum of the provenance  $\theta_g$  and the current temperature  $T_t$  in year  $t$ , with  $g_{i,\max}$  the maximal growth rate in stage  $i$  and  $\omega_{g,i}$  the thermal tolerance for growth in stage  $i$ :

$$g_i(\theta_g, T_t) = g_{i,\max} \exp\left(-\frac{(\theta_g - T_t)^2}{2\omega_{g,i}^2}\right) \quad (\text{S1.2b})$$

We then apply the same framework as the case where only survival is dependent on temperature: each year, individuals can either remain in their current stage if they survive and do not grow sufficiently to transition (with probability  $s_i(\theta_s, T_t) \times (1 - g_i(\theta_g, T_t))$ ), die (with probability  $1 - s_i(\theta_s, T_t)$ ), or transition to the next stage if they survive and grow (with probability  $s_i(\theta_s, T_t) \times g_i(\theta_g, T_t)$ ).

We focus on the effect of having both survival and growth dependent on temperature on the optimal provenance, and for simplicity assume that survival and growth have the same thermal optimum ( $\theta_s = \theta_g$ ) and the same thermal tolerance ( $\omega_{s,i} = \omega_{g,i}$ ). Note that, in this case, the effect on individuals of distance from the optimum is even more important, as it reduces both survival and growth. We therefore expect a lower number of survivors at harvest time.

For this analysis, we exclude species that spend only one year in the first stage. The definition of the first stage in these matrices is "seed" (see Supplementary material S4), with a probability of moving to the next stage set to 1. Making this transition probability dependent on temperature would make the duration of this seed stage longer than a year, which would be biologically implausible for these species.

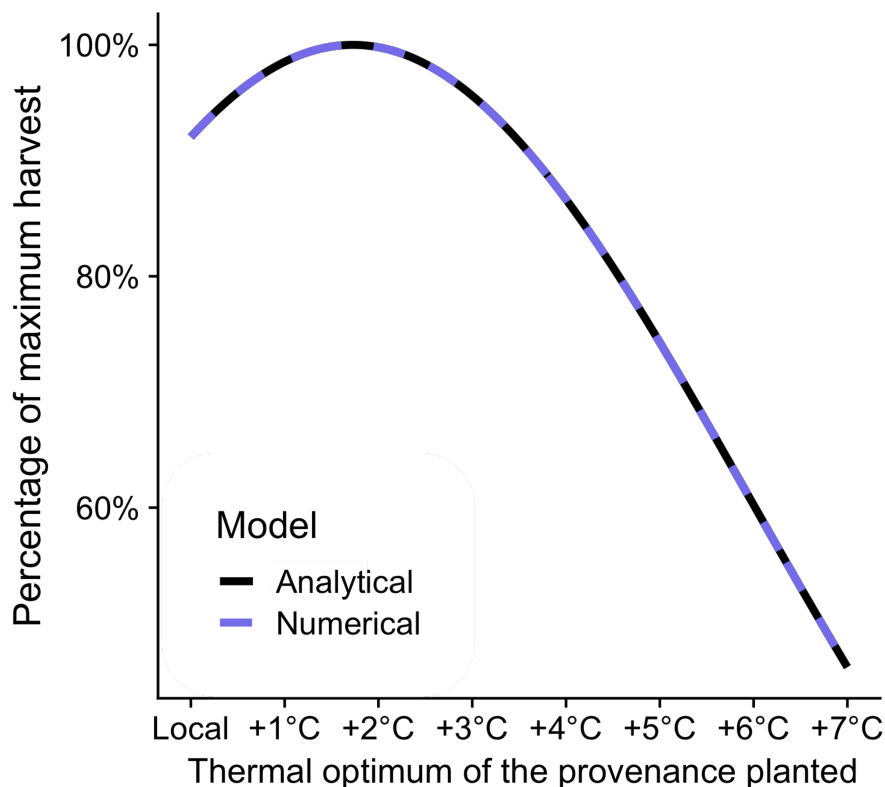


Suppl. Fig. 1.3: Comparison of model predictions for the thermal optimum of the best provenance as a function of changes in thermal tolerance across stages ( $\omega_n/\omega_1$ ) when only survival is dependent on temperature (dashed lines) and when survival and growth are both dependent on temperature (solid lines) for thirteen tree species (panels) in the intermediate climate warming scenario of  $+1.7^\circ\text{C}$ . Curves are coloured according to the length the first life stage of each species. In the case where survival and growth are both dependent on temperature, both share the same thermal optimum and thermal tolerance. The vertical line indicates the case where tolerance is constant across all stages. The dotted horizontal line corresponds to the mean temperature over the period. The total thermal tolerance across the life span is held fixed at  $\omega_{\text{tot}}(\bar{a}) = 3.5^\circ\text{C}$ .

We find when both survival and growth depend on temperature, the thermal optimum of the best provenance decreases when young trees are less tolerant to temperature changes than old trees, matching the relationship described when only survival depends on temperature ( $\omega_n/\omega_1 > 1$ ; Suppl. Fig. 1.3). In these examples, our predictions about the best provenance are also very close quantitatively to predictions when climate change only affects survival.

### S3. Comparing analytical predictions with simulations

We compared analytical prediction given by equation 1.8 and the numerical simulations for a two-stage life cycle for which we varied the optimal temperature of the provenance  $\theta$  between the local temperature in the planting site to the local temperature  $+7^\circ\text{C}$ , in increments of 0.01 degrees. Each simulation iterated the model for  $H = 60$  years and computed the number of individuals that have reached the last stage at harvest. We found both model outputs to be identical (Suppl. Fig. 1.4).



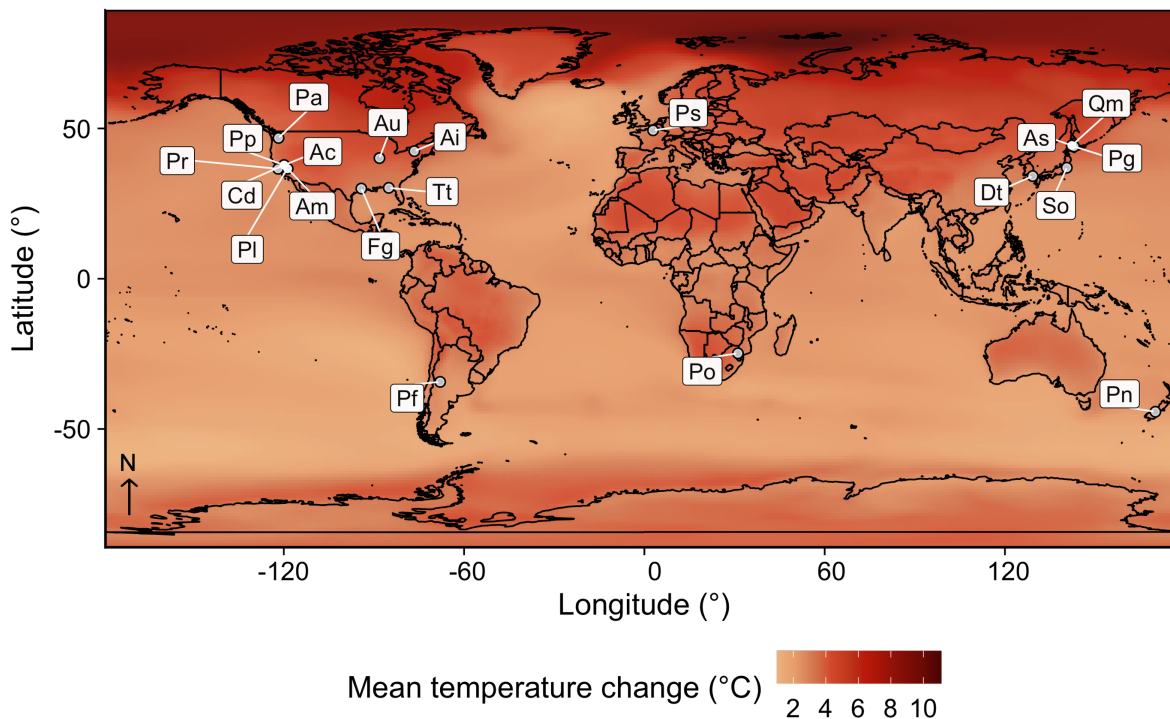
Suppl. Fig. 1.4: Comparison between analytical predictions (black) and numerical simulations (purple). The black curve was produced by the stage-structured analytical prediction (equation 1.8) and the purple curve by numerical simulations. Harvest includes all surviving individuals, regardless of their stage. The transition probabilities between stages were calculated with parameters:  $s_1 = 0.61$ ,  $s_2 = 0.98$  and  $g = 0.012$ . Simulations were carried out under the intermediate climate warming scenario of  $+1.7^\circ\text{C}$ , and the harvest period was set at  $H = 60$ . Thermal tolerance was held constant across the lifespan ( $\omega_2/\omega_1 = 1$ ) with the total thermal tolerance fixed at  $\omega_{\text{tot}}(\bar{a}) = 3.5^\circ\text{C}$ .

#### S4. Details about the demographic matrices gathered

The transition matrices gathered are quite variable in the way that they categorize life stages, both in the definition and number of stages. For example, the first stage for the matrix of white fir (*Abies concolor*) contains individuals below five cm in diameter at breast height and the second stage individuals between five to ten cm DBH, while the matrix for tortuous mesquite (*Prosopis flexuosa*) defines the first stage as “seedling”, and the second stage as saplings with a diameter at ground level below 1 cm. Some of the variability stems from real biological differences (*e.g.*, with a maximum height ranging from 10 meters for tortuous mesquite to 75 meters for white fir), but some differences may be more arbitrary. In particular, some transition matrices include a transient seed stage before germination, while others start with seedlings.



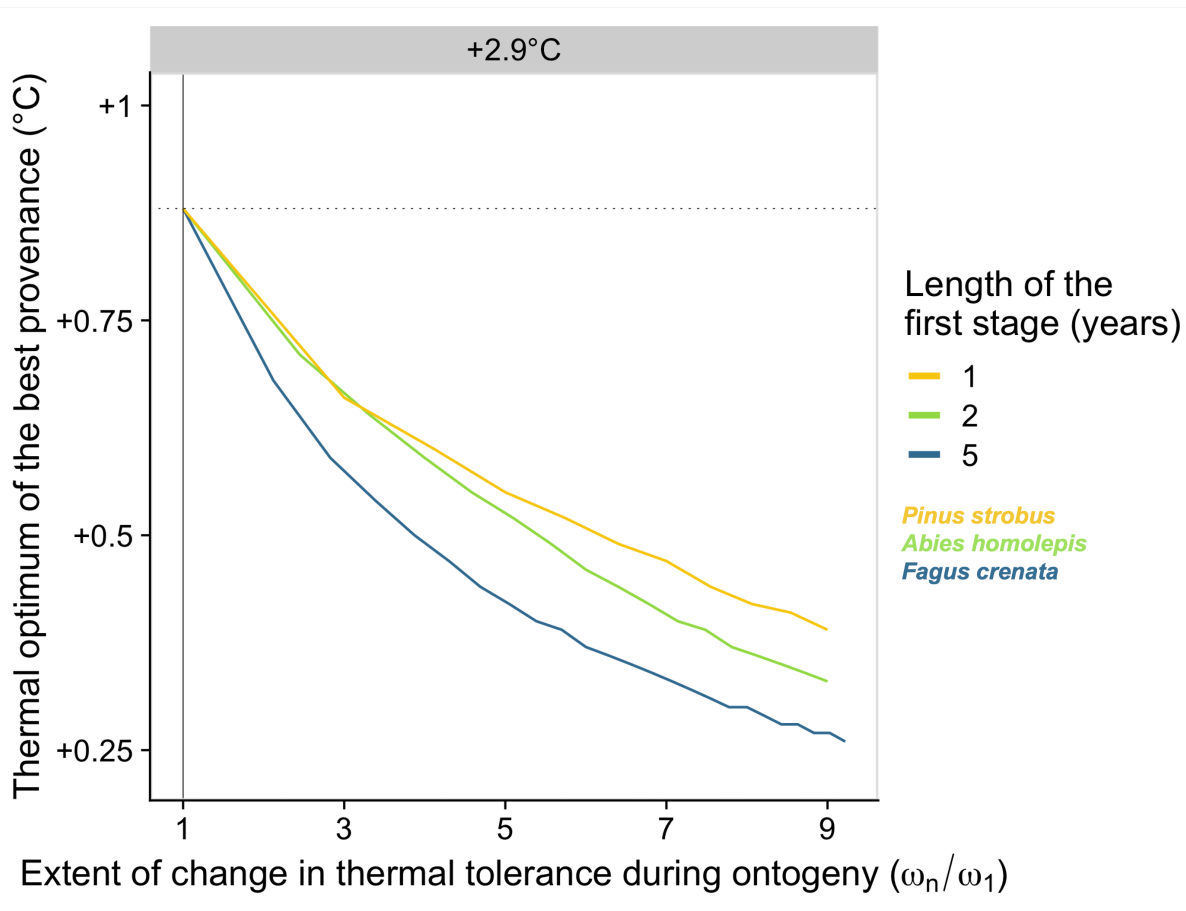
## S5. Expected warming at the geographic coordinates of the studied tree species



Suppl. Fig. 1.5: Geographic location of investigated tree species in our projected intermediate global climate warming of +1.7°C. The dots represent the geographic coordinates at the location of the tree population matrix estimation. Tree species are *Abies concolor* (Ac), *Abies magnifica* (Am), *Abies sachalinensis* (As), *Acer saccharum* (Au), *Alnus incana* (Ai), *Calocedrus decurrens* (Cd), *Dendropanax trifidus* (Dt), *Fagus grandifolia* (Fg), *Picea glehnii* (Pg), *Pinus albicaulis* (Pa), *Pinus lambertiana* (Pl), *Pinus nigra* (Pn), *Pinus ponderosa* (Pp), *Pinus radiata* (Pr), *Prosopis flexuosa* (Pf), *Prunus serotina* (Ps), *Pterocarpus angolensis* (Po), *Quercus mongolica* (Qm), *Styrax obassis* (So) and *Torreya taxifolia* (Tt). The estimated mean temperature change is from the IPCC CMIP6 - SSP2-4.5 climate warming scenario projected for a long term period 2081-2100 (rel. to 1995-2014) (data can be downloaded at <https://interactive-atlas.ipcc.ch>).

## S6. Longer rotation time

To test the effect of a longer harvest time ( $H = 100$  years), we use three matrices of species that reach the last stage after 60 years and before 100 years.

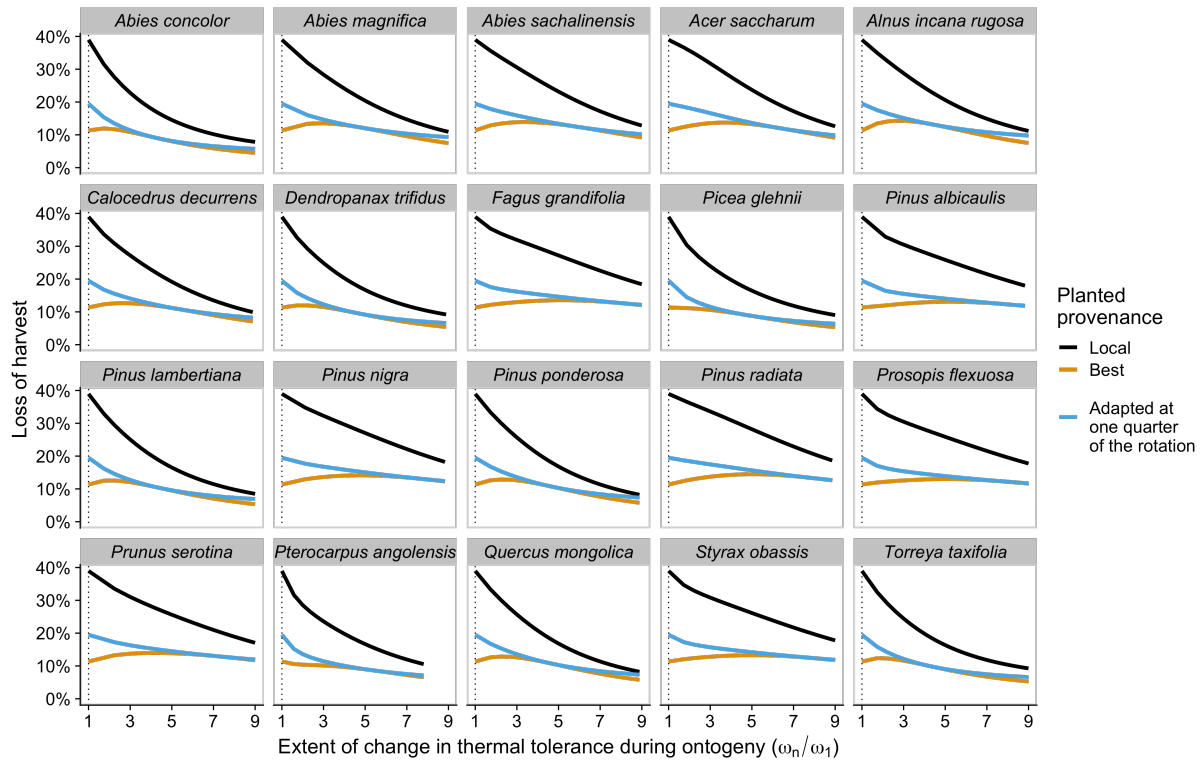


Suppl. Fig. 1.6: Variation in the best provenance to plant illustrated using the life history trajectories for three species that mature slower. The less tolerant younger trees are to changes in temperature compared to older trees (the larger  $\omega_n/\omega_1$  along the x-axis), the cooler the best provenance. The effect is shown for the intermediate warming scenario. For a rotation time of 100 years, the intermediate warming scenario (of 1.7°C in 60 years) gives a warming of 2.9°C in 100 years. Each curve represents a species and is coloured by the length of their first life stage. The vertical line indicates the case where tolerance is constant across all stages. For each warming scenario, the dotted horizontal lines correspond to the average temperature over the period.

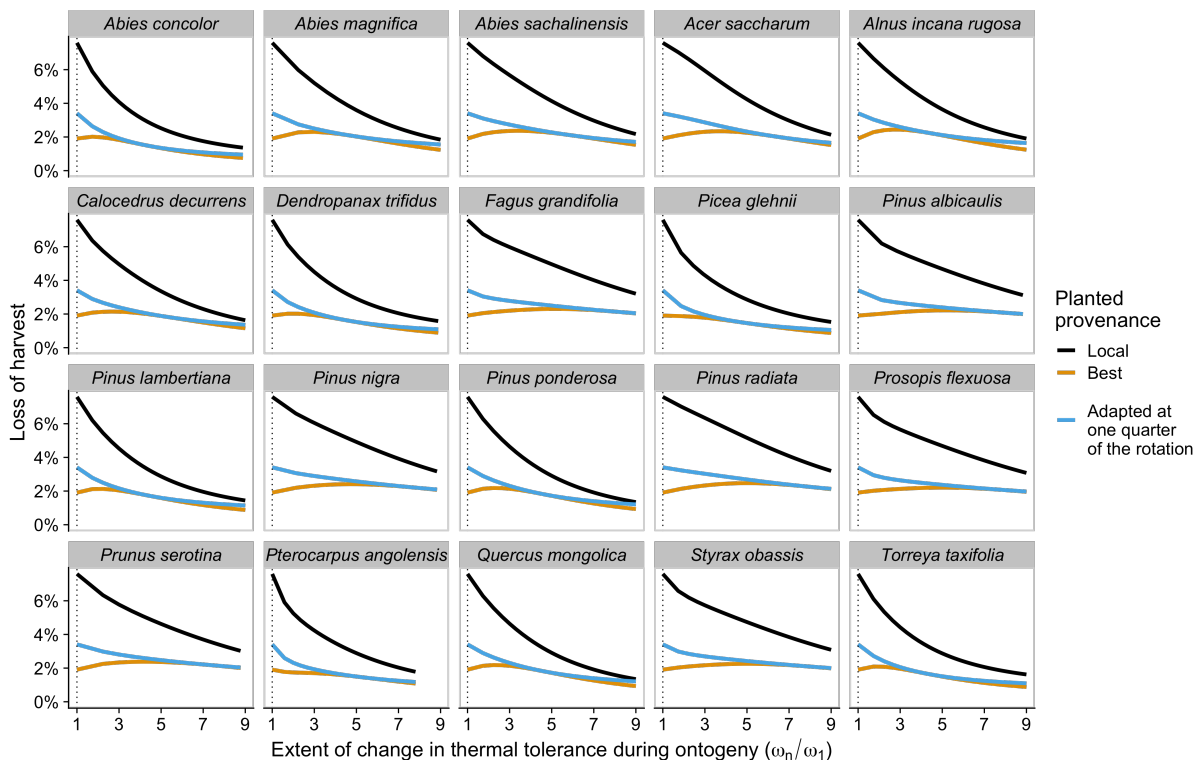
We find that our main qualitative predictions also hold for species that mature slower with a longer rotation time ( $H = 100$ ). When the thermal tolerance is constant over the lifetime of the trees ( $\omega_i/\omega_1 = 1$ ), the thermal optimum of the best provenance matches the average temperature over the period and does not vary across species (Suppl. Fig. 1.6). When young trees are less tolerant to temperature changes than old trees ( $\omega_i/\omega_1 > 1$ ), the thermal optimum of the best provenance decreases and becomes increasingly different among species depending on their life history. The optimal provenance for species that spend many years in the first and least tolerant stage is cooler and much nearer current temperatures (darker curves) than for species that mature earlier (lighter curves).

**S7. Comparing the performance of different seed sources - under multiple climate warming scenarios and for different values of the cumulative thermal tolerance ( $\omega_{\text{tot}}(\vec{a})$ )**

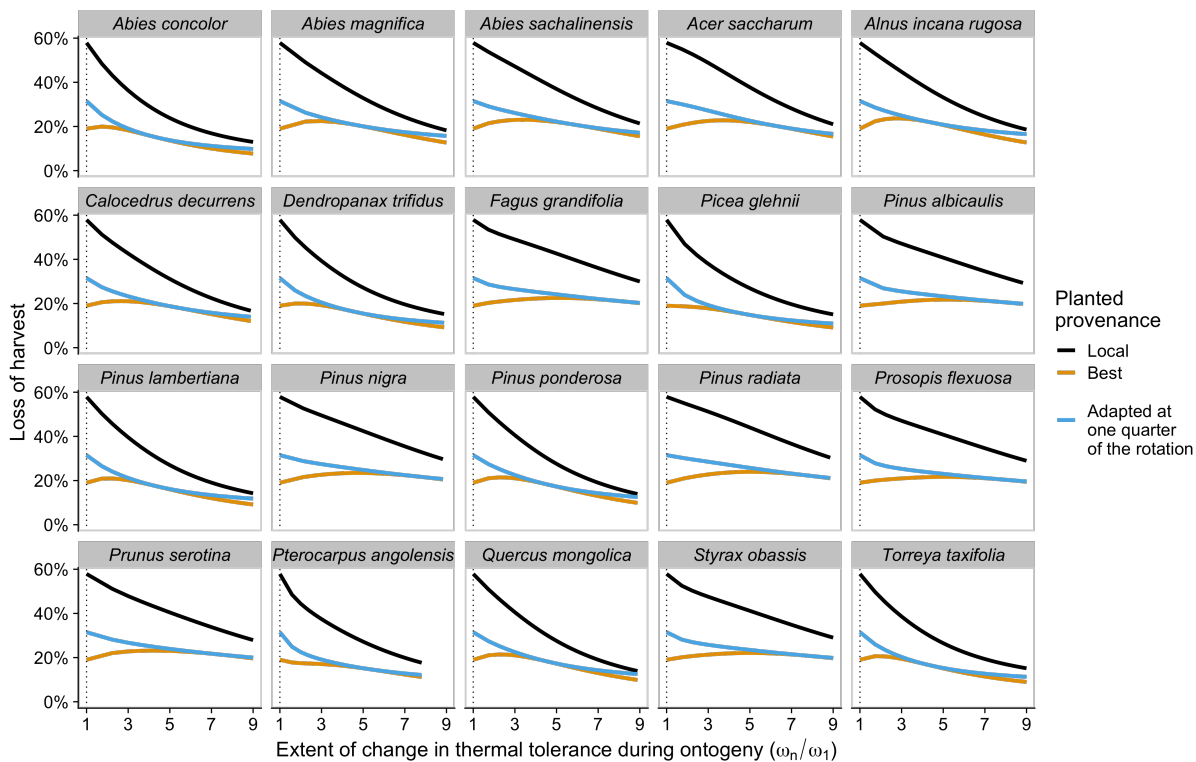
**A: Climate warming = +1.7°C and  $\omega_{\text{tot}}(\vec{a}) = 2^\circ\text{C}$**

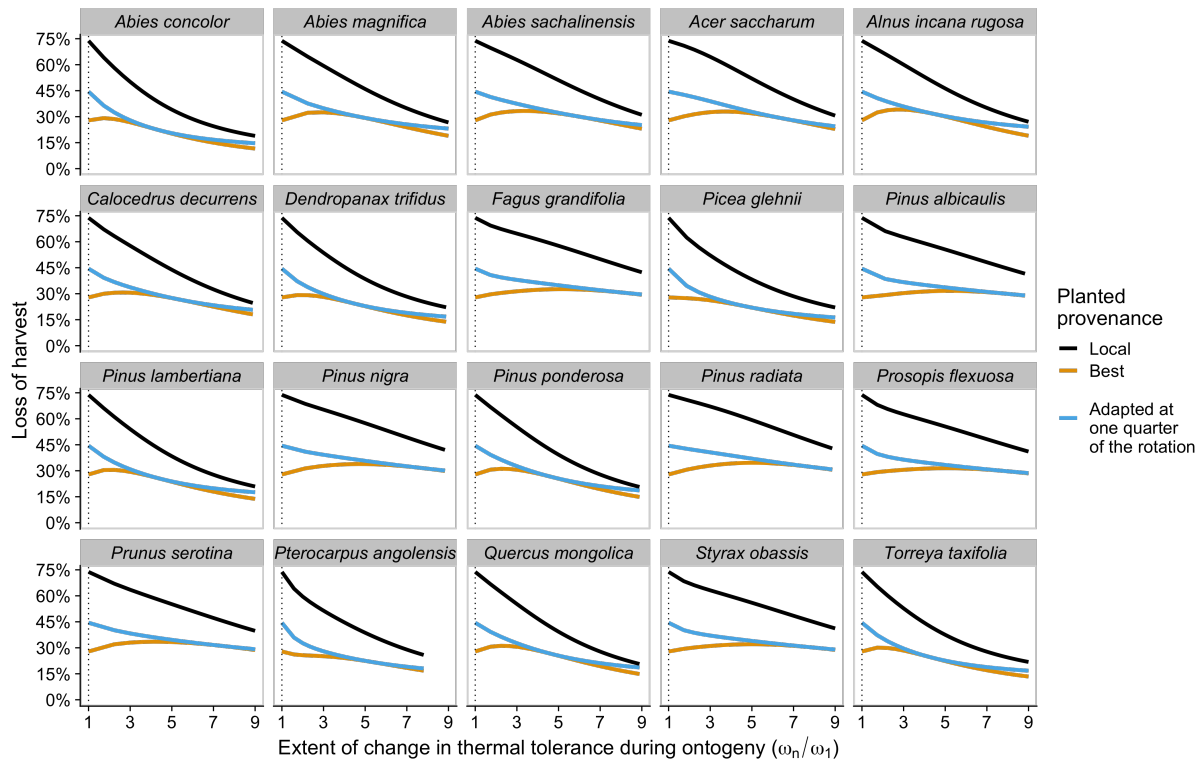


**B: Climate warming = +1.7°C and  $\omega_{tot}(\vec{a}) = 5^\circ\text{C}$**



**C: Climate warming = +2.6°C and  $\omega_{tot}(\vec{a}) = 2^\circ\text{C}$**



D: Climate warming =  $+3.3^{\circ}\text{C}$  and  $\omega_{\text{tot}}(\vec{a}) = 2^{\circ}\text{C}$ 

Suppl. Fig. 1.7: Modifying both the cumulative thermal tolerance ( $\omega_{\text{tot}}(\vec{a})$ ) and the strength of climate warming has a major impact on the percentage of loss of harvest - increasing with worse climate change and lower cumulative thermal tolerance. The shape of the response comparing the seed sources (local, best and adapted to temperature at one quarter of the rotation time), however, remains unchanged in all cases and for all species. For all panels, harvest loss was measured against the performance of the local provenance in a scenario with no climate change. The vertical dotted line indicates the case where tolerance is constant across all stages.

## Appendix A: Estimation of the thermal tolerance of trees

*This appendix is not included in the publication in Evolutionary Applications.*

The width of the thermal niche or thermal tolerance at each stage is a critical parameter of our model. In an attempt to parameterize our simulations, we conducted a literature review to estimate how much tree yearly performance declines when temperature deviates from the optimal temperature for a given provenance (*i.e.*, tree thermal tolerance,  $\omega_i^2$ , the width of the Gaussian response functions in equation 1.1). We reviewed studies that measured the cumulative growth or survival of trees from different provenances, either planted in the same common garden, or transplanted to multiple locations (provenance trials) to estimate how much performance declines with the climatic distance between the planting site and the site of origin of the provenance. Studies differ both in the tree species studied and the age at which they are measured. We identified 13 papers measuring temperature response curves for a total of 17 tree species, yielding 113 observations: 103 for cumulative height responses and 10 for cumulative survival responses to temperature. Tree ages when measured ranged from 2 to 97 years old (2 to 39 for height and 2 to 97 for survival). In these studies, climate was characterized either by the annual temperature at the planting site ( $T$ ), when trees from a source were transplanted to multiple sites, or by the temperature transfer distance ( $\Delta T$ ) between the planting site and the location of seed collection (*e.g.*, average annual, minimum, maximum, or mean summer temperature difference). For simplicity, we refer to the climate variable as  $T$  in the following expression.

In the papers retrieved, population responses were fitted by quadratic regressions of performance against temperature of the form:  $\alpha T^2 + \beta T + \gamma$ . In our age-structured model, the number of survivors after  $c$  years can also be approximated by a quadratic response to temperature if we assume that temperature varied little during the study and is not very far from the thermal optimum of the provenance. The cumulative number of survivors at census can then be approximated by:

$$E[N_c] \approx N_0 s_{\text{tot,max}} \left( 1 - \frac{(\theta - T)^2}{2 \omega_{\text{tot}}^2} \right), \quad (1.9)$$

with  $\omega_{\text{tot}}^2$ , the cumulative thermal tolerance. By identifying the coefficients of the quadratic, we can deduce that:  $\omega_{\text{tot}}^2 = \frac{\beta^2 - 4\alpha\gamma}{8\alpha^2}$ .

The different studies looked at cumulative survival or growth after a variable number of years. Such estimates of thermal tolerance are therefore not comparable among studies as performance

is expected to decrease with the number of cumulative years under a non-optimal temperature. To find a comparable measure across studies, we sought the average thermal tolerance per year that would result in the same cumulative tolerance ( $\omega_{\text{tot}}^2$ ) if all ages had the same tolerance, that is the harmonic mean of the squared thermal tolerances experienced each year by trees, from planting to census. Considering the number of years before census ( $c$ ), the average squared tolerance per year is then estimated as  $\tilde{\omega}^2 = c\omega_{\text{tot}}^2$ . If the thermal tolerance of trees increases with age in most species, we expect the harmonic mean of squared tolerances to increase when trees were measured at a later age (*i.e.*, in studies with a larger number of years before census).

Few papers estimated thermal tolerance from survival data (10 observations) and many more from growth responses (103 observations). We pooled both types of data together to estimate the average thermal tolerance per year and tested if it varies with age at census and differs between the considered trait (survival or growth).

### Statistical analysis

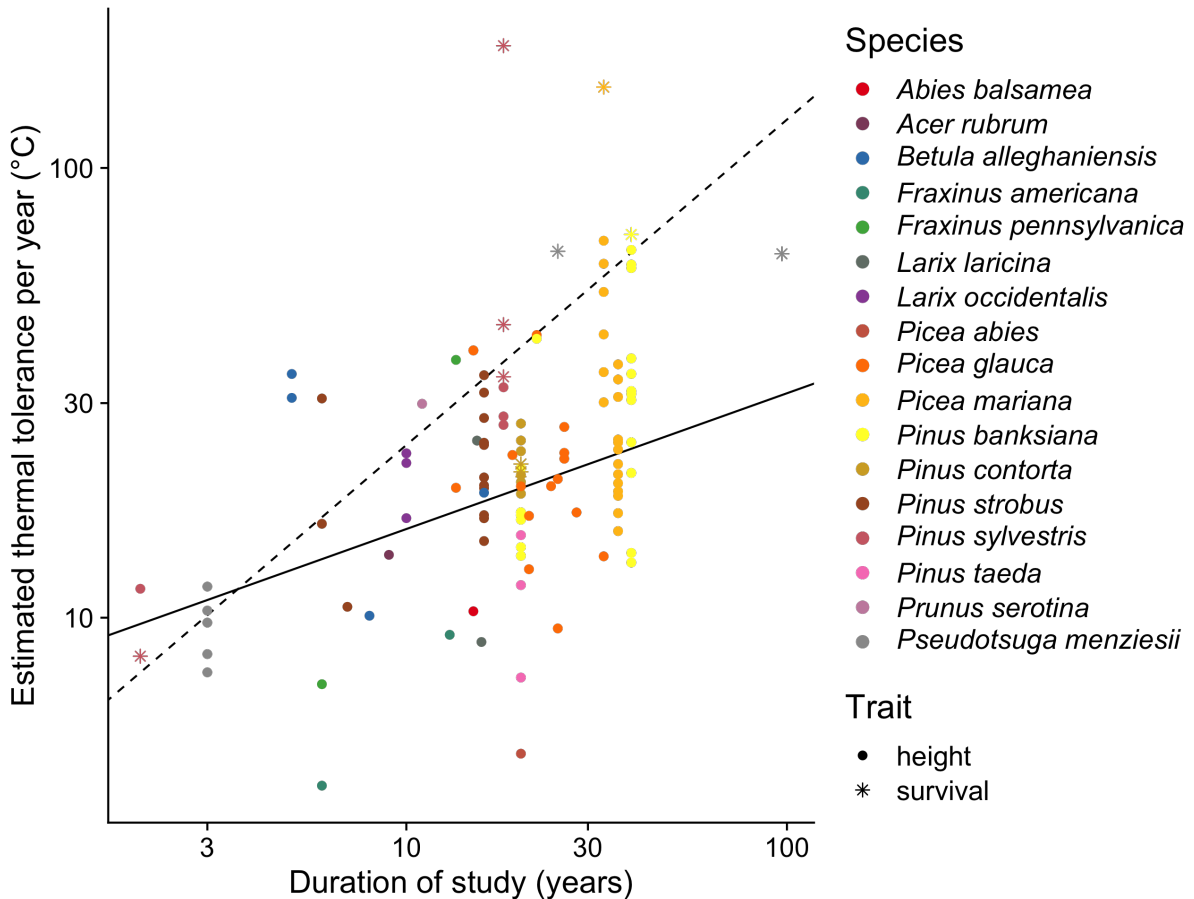
We used the following mixed effects model that takes into account different sources of variation to statistically test for a change in the estimated thermal tolerance with study duration

$$\log_{10}(\tilde{\omega}_{\kappa}^2) = \beta_0 + \beta_1 \log_{10}(c_{\kappa}) + \beta_2 X_{\kappa} + \beta_3 X_{\kappa} \log_{10}(c_{\kappa}) + \sigma_{s_{\kappa}} + \epsilon_{\kappa} \quad (1.10)$$

where  $\tilde{\omega}_{\kappa}^2$  is the estimated average squared yearly thermal tolerance for the study  $\kappa$ , where census occurred after  $c_{\kappa}$  years,  $X_{\kappa}$  indicates whether survival or height was measured,  $s_{\kappa}$  is the species measured in that study and  $\sigma_{s_{\kappa}}$  is the random effect associated with that species,  $\epsilon_{\kappa}$  is the residual effect and the  $\beta$  are the fitted parameters. This complete model allows the harmonic mean of thermal tolerance to change with age differently for survival and growth. We used log-likelihood ratio tests to compare nested models, which contained subsets of the explanatory variables.

We find a significant positive relationship between estimated average tolerance per year and study duration ( $\chi^2 = 23.415$ ,  $p < 0.001$ , Fig. 3), consistently with our expectation that tolerance increases with the age of the tree. We also find that the average tolerance is significantly different depending on the trait considered ( $\chi^2 = 16.577$ ,  $p < 0.001$ ), and there is a significant interaction between study duration and the trait considered ( $\chi^2 = 4.461$ ,  $p < 0.01$ ). There is no significant effect of the species considered ( $\chi^2 = 1.3163$ ,  $p = 0.2513$ ), which could be due to the very few

replicated studies per species in our data set. We however kept the random effect in the model to avoid pseudo-replication when multiple studies used the same species. Focusing on growth, for which there are a greater number of observations, after 60 years, the predicted harmonic mean of yearly tolerances is  $\tilde{\omega} = 27^\circ\text{C}$ . The cumulative thermal tolerance ( $\omega_{\text{tot}}$ ) over the period for a harvest at  $H = 60$  years would thus be  $3.5^\circ\text{C}$ .



Additional. Fig. 1.1: The estimated thermal tolerance per year of trees ( $\tilde{\omega}$ ) as a function of the duration of the study for different tree species (colors), measured for survival (circles) or height (stars). The plain black line is the regression line for height data ( $\log_{10}(Y) = 0.302 \log_{10}(D) + 0.895$ ) and the dashed black line is the regression line for survival data ( $\log_{10}(Y) = 0.724 \log_{10}(D) + 0.660$ ).

The limited empirical data for the same species at different ages, however, did not allow us to parameterize our model with confidence using estimates of thermal tolerance.



## Chapter 2

# Evolutionary rescue in a rare endemic Mediterranean plant: *Centaurea corymbosa*

Adèle Erlichman<sup>1</sup> and Ophélie Ronce<sup>1</sup>

<sup>1</sup> ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

### Abstract

Evolutionary rescue is a demographic and evolutionary process by which a population escapes extinction by adapting to environmental pressures. *Centaurea corymbosa* is a rare and threatened monocarpic perennial endemic to the Massif de la Clape in Southern France. It has been the subject of long-term demographic monitoring since the late 1990s. The small remaining populations of this species have been in decline for several years, and its extinction is predicted in the medium term, accelerated by climate change. The aim of this study is to understand whether the evolution of a fitness-determining trait, the growth rate, could save the population in a warming climate. We use a new method for integrating evolution into an integral projection model (IPM) and use it to predict the critical level of genetic variance for the intrinsic individual growth rate that would be necessary to avoid extinction. We begin by examining a simple scenario in which a single cohort is planted, with the goal of modeling more complex scenarios of reinforcements in the future. We show how the declining trend of the population can only be reversed with a small range of introduced genetic diversity in the growth rate. Even in the best-case scenario, the population will not start expanding again for another twenty years, by which time the extinction rate of some natural populations is already very high.

*Keywords:* integral projection model, population dynamics, evolution, conservation biology, climate change

## Introduction

Modeling population dynamics is key to understanding demographic trends and can help decide whether conservation measures should be taken. In a changing world, demographic and evolutionary processes are intertwined (Pelletier et al. 2009). This is particularly captured by evolutionary rescue processes, where a population escapes extinction through adaptation and natural selection (Carlson et al. 2014; Gomulkiewicz and Holt 1995). Demographic models that integrate eco-evolutionary processes are therefore useful to describe population dynamics in the context of climate change (*e.g.*, Cotto et al. 2019; Simmonds et al. 2020).

Matrix Projection Models (MPMs) have long been used to describe the dynamics of stage-structured populations (MPMs; Caswell 2000). More recently, Integral Projection Models (IPMs; Easterling et al. 2000; Ellner et al. 2016; Ellner and Rees 2006) were introduced as an extension of MPMs. IPMs use a continuous relationship between vital rates (survival, growth and fecundity) and quantitative traits, generally size, rather than dividing individuals into the more or less arbitrary discrete classes of an MPM. These relationships are given by regression models fitted to collected data. IPMs are particularly useful tools for describing the demography of continuously growing organisms (*e.g.*, trees) that do not have clearly defined life stages, such as some insects, for instance, with clear developmental stages (Zuidema et al. 2010). IPMs also reduce the number of parameters that need to be estimated from the data and have been shown to work better on small data sets (Ellner and Rees 2006; Ramula et al. 2009). Individual heterogeneity adds to the complexity of the predictions about how populations will respond to environmental changes (Forsythe et al. 2021; Vindenes and Langangen 2015). IPMs explicitly consider some variation between individuals, instead of treating all individuals from the same stage as identical (Easterling et al. 2000; Ellner and Rees 2006).

IPMs functions are, however, often too complex to be analytically integrated. The continuous variable of the IPM is therefore classically discretized during the numerical integration of the demographic functions, generally using the “mid-point rule” to compute the kernel (Ellner and Rees 2006). This kernel is a matrix containing all the transition probabilities for each age and size class. Conveniently, this allows similar predictions to be made from IPMs and MPMs (*e.g.*, long-term growth rate, sensitivities, elasticities). Comparison of the two types of model has shown that they give similar results, provided that a sufficient number of classes are used to avoid the errors introduced by the discretization method (Doak et al. 2021; Zuidema et al. 2010).

IPMs have been used to predict phenotypic optimums (*e.g.*, age or size at flowering) that maximize the population growth rate (Kuss et al. 2008; Rees and Ellner 2016), but these models did not describe the evolutionary dynamics towards such equilibria. Over the past decade, IPMs have been extended to the prediction of eco-evolutionary dynamics (*e.g.*, Childs et al. 2011; Coulson et al. 2011; Ozgul et al. 2010; Traill et al. 2014). Evolutionary IPMs have, however, been subject to debates. Concerns raised included the fact that some evolutionary explicit IPMs were developed by regressing a so-called heritability function to describe the transmission of trait values from parents to offspring. As the intercept of this function was held constant over generations, this was incompatible with current understanding of evolution (Chevin 2015; Janeiro et al. 2017). Some age-structured evolutionary IPMs (Ozgul et al. 2010; Traill et al. 2014) also considered size as the evolving trait in the population. These models were criticized because their use of growth functions ignored the cumulative effect of genetic difference in growth and the build-up of phenotypic and genetic variation with age, underestimating the contribution of evolution to phenotypic change (Chevin 2015). The recommendation was to, therefore, assume genetic variation for the growth function rather than the size of individuals (Chevin 2015). Other evolutionary explicit models have been developed since, with attempts to make these models more compatible with a classical quantitative genetics framework (Coulson et al. 2017, 2021). Examples of implementation consider the evolution of other continuous quantitative traits, such as the laying date in birds, but ignore the demographic heterogeneities due to size variation (Simmonds et al. 2020). The latter model separately tracks the distribution of breeding values and phenotypic values for the laying date in the population and age classes and assume that the distribution of breeding values in new-born is Gaussian with a fixed variance, corresponding to the additive genetic variance for the trait, and a mean equal to the mean breeding values in the selected mothers. Rees and Ellner (2019) proposed an evolutionary IPM where a single heritable trait is followed in a size-structured population. Their approach is more realistic because they do not assume that the distribution of breeding values is Gaussian throughout the newborn/seedling class, but they use the infinitesimal model which assumes that the distribution of breeding values is Gaussian and of fixed variance in each family, conditional on the breeding values of the two parents. Their framework allows for both the evolutionary dynamics of the trait interacting with size and individual size to be tracked, essentially as in a stage-structured case (Barfield et al. 2011).

In this study, we build upon this previous work to track the joint evolution of a trait and the

demography of an age- and size-structured population described by an IPM. We propose to allow the growth function of an IPM to evolve, such that the breeding value of a trait (the intrinsic individual growth rate) affects the probability of growing to a certain size at the next time step. We first wrote expressions for a size- and age-structured IPM (described in Supplementary Material S1) that explicitly track the distribution of individuals of a certain age, size and breeding value for the intrinsic individual growth rate value assuming an explicit infinitesimal model of inheritance (Fisher 1918), but making no assumption about the distribution of breeding values in the population or in particular cohorts, as in Rees and Ellner 2019. Integrating these equations across continuous variation of size and breeding values for growth rate is however cumbersome and our evolutionary model involves convolutions of parental breeding values distributions that make it non linear, contrary to the classic MPMs and IPMs. We therefore propose to reduce the dimensionality of the evolutionary model by (i) discretizing the demographic model in a large number of small size classes as is typically done when integrating IPMs, (ii) assuming a Gaussian distribution of breeding values for the evolving trait within each of these size classes. The latter assumption has the practical advantage that we now need to track only the mean breeding value in each class, rather than changes in its entire distribution.

We argue that numerous quantitative genetics models have been developed to explore the evolution of continuous traits in stage-structured models with discrete stages, building upon the work of Barfield et al. (2011) (Barfield et al. 2011; Cotto et al. 2019; de Vries and Caswell 2019; Orive et al. 2017; Schmid et al. 2022). As mentioned above, IPM functions are commonly discretized when implemented, hence, we propose to combine the IPM framework, which uses data-informed vital rate functions, with the stage-structured framework described by Barfield et al. (2011) for tracking the evolution of a trait in a population. Some version of the framework proposed by Barfield et al. (2011) requires an assumption about the distribution of phenotypic values within each class, which is assumed to be Gaussian. Considering very small class sizes, we can track the average phenotypic value in each class at each time step.

To illustrate this approach, we use the concrete example of an IPM fitted by Hadjou-Belaid et al. (*in prep*) on the long term demographic data of *Centaurea corymbosa*. *C. corymbosa* is a monocarpic perennial endemic to a small zone of approximately 3km<sup>2</sup> in the Massif de la Clape, in Southern France. Six populations of *C. corymbosa* are known to this day (Colas et al. 1997) and previous demographic studies showed that these populations have been declining (Fréville et al. 2004; Hadjou Belaid et al. 2018) and are especially threatened by climate change with a

high probability of extinction before as soon as 20 years (the asymptotic growth rate for pooled populations is  $\lambda = 0.880$ , Hadjou Belaid et al. 2018). Introductions have been attempted, with the creation of two new populations in 1994 (Colas et al. 2008), one has shortly gone extinct and the other still persists as a small isolated patch (Eric Imbert, personal communication). Reinforcements were also initiated, with the plantation of around 40.000 seeds, over 4 years, in the 6 extant natural populations (Ducrettet and Imbert 2023). Hadjou-Belaid (2018) shows that evolution of age and size at flowering in *C. corymbosa* is unlikely to rescue the populations from extinction. They predict the optimal age and size at flowering of this monocarpic perennial under current and future climate and find that even if the plants evolved to the optimal value for this life history trait, the population would still decline if other aspects of their life history remained unchanged. Conversely, the long-term population growth rate is much more sensitive to variation in rosette survival, which critically depends on their size. Size is also a strong determinant of plant fecundity (Hadjou Belaid et al. 2018). This suggests that rapid evolution of individual growth rates *i.e.*, the average annual size increment, by affecting size distribution in the population, could have a large impact on population demography and prospects of survival.

We are interested in the possibility that some level of genetic variability of the individual growth rate within the population would ensure long-term increase of the *C. corymbosa* population *i.e.*, reverse the declining population dynamics, before it goes extinct (Gomulkiewicz and Houle 2009). We propose a modeling framework that allows a life-history trait to evolve in an age- and size-structured population. We vary the initial level of genetic variability of the individual size growth rate within the population, and track both the demography and the evolution of this life-history trait over time. We begin by examining the effect of an initial addition of variance to individual growth on the size distribution of individuals over time. We then examine different scenarios aimed at stimulating evolutionary rescue of *C. corymbosa*.

## Methods

### General scenario

We start by simulating the introduction of a new population, tracking a single cohort of translocated individuals in their first year and are interested in finding for which level of genetic variance in the cohort, the population is expected to recover before 20 years, which is the extinction horizon for certain populations. This paves the way for more complex scenarios, simulating

reinforcements of extant populations for instance.

## Description of the Integral Projection Model

We use the IPM developed in Asma Hadjou-Belaid's thesis (Hadjou-Belaid 2018) and Hadjou-Belaid et al. (in prep) to describe the demography of *C. corymbosa* in the absence of genetic variation in the population. Like for many plants, size and age are very important determinants of variation in life history traits between individuals for *C. corymbosa*, and both are used to describe the state of an individual. Individuals are, thus, characterized by their age with the discrete variable  $a \in \{1, \dots, a_{max}\}$ , and size with the continuous variable  $x \in [x_{min}(a), x_{max}(a)]$  whose range varies with age. Based on the life cycle of *C. corymbosa*, and conditionally to their survival until the next flowering season, each individual can either reproduce and then die, or stay vegetative and change in size (see Fig. 2.1). We ignore the rare case where some plant survive and reproduce a second time (less than 0.3%, Hadjou Belaid et al. 2018).

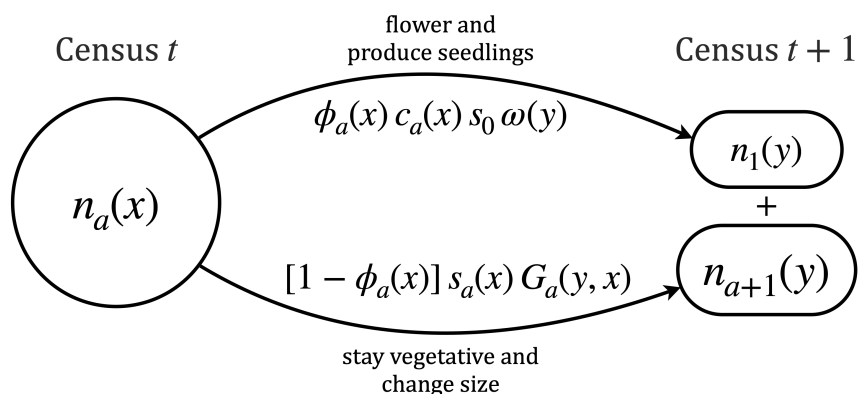


Fig. 2.1: Life cycle diagram for the IPM model representing the life history of *Centaurea corymbosa* that considers a prebreeding census. At each time step  $t$ , plants can either: flower ( $\phi_a(x)$ ), bear capitulas ( $c_a(x)$ ) that produce seeds that survive and germinate ( $s_0$ ) with size distribution ( $\omega(y)$ ) in March, or, stay vegetative ( $1 - \phi_a(x)$ ) and change size ( $G_a(y, x)$ ), conditional on their survival ( $s_a(x)$ ).

Conditional on its survival, an individual of size  $x$  at time  $t$  can produce new individuals, or change size to be, of size  $y$  the following year ( $t + 1$ ). At each time step, given the number of individuals in each age and size classes  $n_a(x)$  at time  $t$ , we can predict both the number of one-year-old seedlings with size  $y$ ,  $n_1(y)$ , and the number of individuals in the other age classes,  $n_{a+1}(y)$ , at the next time step  $t + 1$ , by summing the contributions of each individual through

reproduction, change in size and aging

$$\begin{cases} n_1(y, t + 1) = \sum_{a=1}^{a_{max}} \int_{x_{min}(a)}^{x_{max}(a)} n_a(x, t) \phi_a(x) c_a(x) s_0 \omega(y) dx \\ n_{a+1}(y, t + 1) = \int_{x_{min}(a)}^{x_{max}(a)} n_a(x, t) (1 - \phi_a(x)) s_a(x) G_a(y, x) dx \end{cases} \quad (2.1)$$

with  $\phi_a(x)$  the flowering probability of individuals of age  $a$  and size  $x$ ,  $c_a(x)$  the number of capitula produced by an individual of age  $a$  and size  $x$ ,  $s_0$  the establishment rate of seeds (*i.e.*, number of seedlings surviving to March  $t$ , relative to the number of capitula at  $t-1$ ),  $\omega(y)$  the probability density that a given seedling has size  $y$  in March,  $s_a(x)$  the survival rate of an individual of age  $a$  and size  $x$ , and  $G_a(y, x)$  the growth rate of an individual of age  $a$  from size  $x$  at  $t$  to size  $y$  at  $t + 1$ . Note that we assume that the number of capitula determines both the contribution of a mother plant to the total number of seeds that germinate, and the contribution of a father plant to the pollen pool that fertilized these seeds.

These demographic functions have been fitted with the long-term demographic data for *C. corymbosa*. Individual size is defined by the diameter of the plants, which has been measured every three months (in September, December, March and June) since 1994 in delimited plots of the six natural populations. Further details regarding the data collection can be found in Hadjou-Belaid et al. (*in prep*). Details regarding the fit of the statistical models are summarized in Tab. 2.1 and Fig. 2.2 and can be found in Hadjou-Belaid et al. (*in prep*).

Demographic function	Fitted model
Survival	$\text{logit}(s_a(x)) = -2.41 + 0.77x - 0.03x^2 + 0.10a$
Growth	$G_a(y, x) = \frac{1}{\sqrt{2\pi} \sigma(x)^2} \exp\left(-\frac{(y - \mu(a, x))^2}{2\sigma(x)^2}\right)$
	$\mu_a(x) = 2.01 + 1.341x - 0.03x^2 - 0.33a$
	$\sigma(x) = 0.93 + 0.32x$
Flowering probability	$\text{logit}(\phi_a(x)) = -8.96 + 0.72x - 0.02x^2 + 0.43a$
Number of capitula	$c_a(x) = 0.45 + 1.95x + 1.43a$
Seedling size distribution	$\omega(y) = \exp(-0.58 - 0.82y)$
Establishment rate of seeds	0.560

Tab. 2.1: Summary of the demographic functions fitted and parameters estimated by Hadjou-Belaid et al. *in prep*.

The statistical models fitted in Hadjou-Belaid et al. *in prep*. included a random population effect and a random year effect for survival and growth predictors and a random year effect for

flowering and fecundity predictors. For the sake of simplicity, we have here used predictions for the fixed part of the models, simulating the species demography for an hypothetical population and year and ignoring between years fluctuations. Our predictions of demography are therefore overly optimistic in this regard.

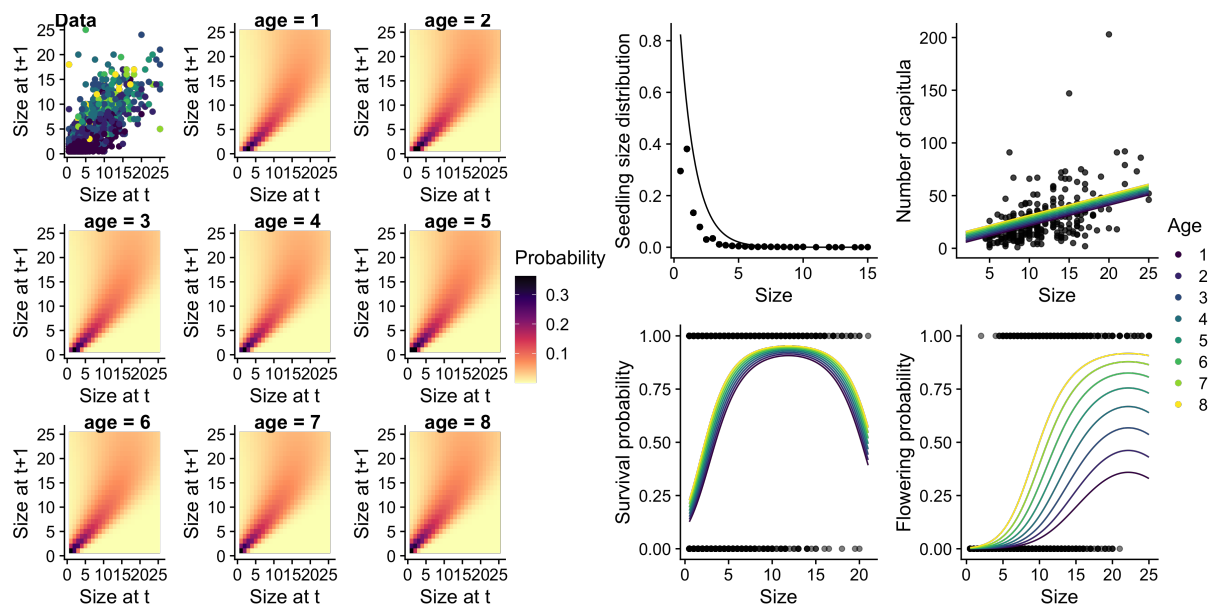


Fig. 2.2: Graphical representation of the available demographic data for *C. corymbosa* and the demographic transition functions fitted by Hadjou-Belaid et al. *in prep*. The growth function is shown on the left with the raw data on the top left corner and the growth predictions by age on the other panels. Number of life-histories: 7243.

## Adding evolution to the IPM

The framework described hereafter aims at tracking evolution of a genetically variable trait for which genetic variation is initially added, to see if this stimulates evolutionary rescue. We choose to introduce genetic variability into the individual growth function of the IPM model of *C. corymbosa* as this will affect the overall transition probabilities, survival-growth and reproduction. We, therefore, assumed that the individual size growth rate is genetically variable in the population. The Gaussian properties of the growth function also facilitate the derivation of the evolutionary expressions.

## Discretization of the IPM demographic functions

We propose a framework to combine the interesting properties of the IPM, using continuous functions adjusted on rich data sets to parameterize life-history traits variation, with the existing demographic and evolutionary framework developed in stage-structured models.



Given that the size range is not constant with age (see Suppl. Tab. 2 in Supplementary Material S2 for details), we define  $l$  size classes within each age class in the population with the width of size intervals at age  $a$ :

$$h(a) = \frac{x_{max}(a) - x_{min}(a)}{l}. \quad (2.2)$$

The choice to maintain constant the number of size classes between age categories, rather than the width of size classes between age categories, is motivated by the idea that small difference in size in younger age classes may have more consequences on their life history traits than in older age classes.

Let  $j$  be the size class within age  $a$  characterized by its midpoint size  $m_a(j)$  that contains all individuals of age  $a$  and size  $x$  such that  $x$  is contained in  $[m_a(j) - \frac{h(a)}{2}, m_a(j) + \frac{h(a)}{2}]$ . When  $h$  is small, the number of individuals of age  $a$  within that interval at each time step  $t$ ,  $N_a(j, t)$ , can be approximated by

$$N_a(j, t) = \int_{m_a(j) - \frac{h(a)}{2}}^{m_a(j) + \frac{h(a)}{2}} n_a(x, t) dx \approx h(a) n_a(m_a(j), t). \quad (2.3)$$

Let  $f_a(i, j)$  be the fecundity function that gives the number of established seedlings, produced by parents of age  $a$  and size contained in interval  $j$ , that fall in the size interval  $i = [m_1(i) - \frac{h(1)}{2}; m_1(i) + \frac{h(1)}{2}]$ . We next define  $p_a(i, j)$ , the transition function that gives the probability that a plant, whose size is contained in interval  $j$ , survives, stays vegetative, and falls in the size interval  $i$ , such that its size  $y$  is in  $[m_{a+1}(i) - \frac{h(a+1)}{2}; m_{a+1}(i) + \frac{h(a+1)}{2}]$  at the next time step.

We can express the size distribution of one-year-old seedlings,  $N_1(i, t + 1)$ , and the size distribution of individuals in the other age classes  $N_{a+1}(i, t + 1)$  at the next time step as

$$\begin{cases} N_1(i, t + 1) = \sum_{a=1}^{a_{max}} \sum_{j=1}^l f_a(i, j) N_a(j, t) \\ N_{a+1}(i, t + 1) = \sum_{j=1}^l p_a(i, j) N_a(j, t), \end{cases} \quad (2.4)$$

with

$$f_a(i, j) \approx h(1) \phi_a(m_a(j)) c_a(m_a(j)) s_0 \omega(m_a(i)) \quad (2.5)$$

and

$$p_a(i, j) \approx h(a + 1) s_a(m_a(j)) \left(1 - \phi_a(m_a(j))\right) G_a(m_a(i), m_a(j)). \quad (2.6)$$

### Evolution of the growth function

Let  $z$  be the value of some underlying phenotypic trait (*e.g.*, the concentration of a growth hormone) such that  $z$  affects the probability that an individual of age  $a$ , size  $x$  will have size  $y$  at  $t + 1$ . It modifies the intercept of the growth function such that the expected size at  $t + 1$  given phenotype  $z$ , size  $x$  and age  $a$  at time  $t$  is:

$$\mu_a(x) + z = (2.01 + z) + 1.341x - 0.03x^2 - 0.33a. \quad (2.7)$$

As a result, individuals of all age and size will tend to grow faster if  $z$  is larger. The growth function of an individual with trait  $z$  can then be written as

$$G_{a,z}(y, x) = \frac{1}{\sqrt{2\pi}\sigma(x)^2} \exp\left(\frac{-\left(z - (y - \mu_a(x))\right)^2}{2\sigma(x)^2}\right). \quad (2.8)$$

### Demographic transitions

We assume that phenotype  $z$  is distributed as a Gaussian in each class with  $\bar{z}_a(j, t)$ , the average phenotype of individuals of age  $a$  in the size interval  $j$  at time  $t$ , and  $\sigma_z^2$ , its phenotypic variance, which we assume constant across age and size classes and time. Let  $\tau_{a,j}(z, t)$  be the probability that an individual of age  $a$  and contained in interval  $j$  has phenotype  $z$  at time  $t$  (here assumed to be the probability density function of a Gaussian distribution). The average transition probability for an individual of age  $a$ , size contained in interval  $j$  and integrating over all possible phenotypic values  $z$  can then be written as

$$\begin{aligned} \bar{p}_a(i, j, t) &= \int \tau_{a,j}(z, t) p_{a,z}(i, j) dz \\ &= h(a+1) s_a(m_a(j)) \left(1 - \phi_a(m_a(j))\right) \int \tau_{a,j}(z, t) G_{a,z}(m_a(i), m_a(j)) dz \\ &= h(a+1) s_a(m_a(j)) \left(1 - \phi_a(m_a(j))\right) \\ &\quad \frac{1}{\sqrt{2\pi}(\sigma(m_a(j))^2 + \sigma_z^2)} \exp\left(\frac{-\left(\bar{z}_a(j, t) - (m_a(i) - \mu_a(m_a(j)))\right)^2}{2(\sigma(m_a(j))^2 + \sigma_z^2)}\right). \end{aligned} \quad (2.9)$$

The number of individuals in each age and size class at the next time step is then

$$\begin{cases} N_1(i, t+1) = \sum_{a=1}^{a_{max}} \sum_{j=1}^l f_a(i, j) N_a(j, t) \\ N_{a+1}(i, t+1) = \sum_{j=1}^l \bar{p}_a(i, j, t) N_a(j, t). \end{cases} \quad (2.10)$$

### Evolution of the trait

The quantitative trait  $z$  is determined by a large number of additive loci with small effects. Each new individual that is born has a phenotype  $z$  given by the sum of its breeding value,  $g$ , and a micro-environmental effect,  $e$ , drawn from a Gaussian distribution of variance  $\sigma_e^2$ . Its breeding value is also drawn from a Gaussian distribution, centered on the mean of the breeding values of its two parents, with variance  $V_0 = V_{LE}/2$ , which we assume constant across age and size classes, following the infinitesimal model of inheritance. We first describe this modeling framework that makes no assumptions about the distribution of breeding values in the population or in particular cohorts in Supplementary Material 2 - S1.

We use expressions from Barfield et al. 2011, which provide a framework for tracking changes in the phenotypic ( $z$ ) and breeding ( $g$ ) values in a stage-structured population, assuming Gaussian distributions of phenotypic and breeding values in every age and size classes. We describe the expected changes in phenotypic and breeding values in this age and size-structured population. We now assume that the variance of breeding values in every stage is  $\sigma_g^2$ . The expected mean breeding value for one-year-old seedlings,  $\bar{g}_1$ , and for the other age classes,  $\bar{g}_{a+1}$ , at the next time step, can be expressed with the following recursion (equation (5) from Barfield et al. 2011)

$$\begin{cases} \bar{g}_1(i, t+1) = \sum_{a=1}^{a_{max}} \sum_{j=1}^l c_a^f(i, j) \bar{g}_a(j, t), \\ \bar{g}_{a+1}(i, t+1) = \sum_{j=1}^l c_a^p(i, j) (\bar{g}_a(j, t) + \sigma_g^2 \frac{\partial \ln \bar{p}_a(i, j, t)}{\partial \bar{z}_a(j, t)}), \end{cases} \quad (2.11)$$

with  $c_a^f(i, j) = \frac{f_a(i, j) N_a(j, t)}{N_1(i, t+1)}$  and  $c_a^p(i, j) = \frac{\bar{p}_a(i, j, t) N_a(j, t)}{N_{a+1}(i, t+1)}$  the relative contributions to state  $i$  by stage  $j$  through direct reproduction and transition, respectively. Note that

$$\frac{\partial \ln \bar{p}_a(i, j)}{\partial \bar{z}_j} = \frac{(m_a(i) - \mu_a(m_a(j))) - \bar{z}_a(j, t)}{\sigma(m_a(j))^2 + \sigma_z^2}. \quad (2.12)$$

Similarly, changes in the expected mean phenotypic value for one-year-old seedlings,  $\bar{z}_1$ , and for the other age classes,  $\bar{z}_{a+1}$ , at the next time step, can be expressed with the following recursion

(equation (6) from Barfield et al. 2011)

$$\begin{cases} \bar{z}_1(i, t+1) = \sum_{a=1}^{a_{max}} \sum_{j=1}^l c_a^f(i, j) \bar{g}_a(j, t), \\ \bar{z}_{a+1}(i, t+1) = \sum_{j=1}^l c_a^p(i, j) (\bar{z}_a(j, t) + \sigma_z^2 \frac{\partial \ln \bar{p}_a(i, j, t)}{\partial \bar{z}_a(j, t)}). \end{cases} \quad (2.13)$$

## Simulations

The model described above allows us to construct projection kernels and calculate transition probabilities for each age and size classes under R (R Core Team 2022; see details about the code in Appendix B). Given the high probability of extinction of *C. corymbosa* under various climate scenarios between 20 and 50 years (see Hadjou Belaid et al. 2018), we are interested in finding out whether an increase in heritable phenotypic variance could allow the natural populations to reach a positive growth rate before 20 years.

### Introduction of a single cohort

In this first scenario, a single cohort is introduced with a given initial phenotypic variation among individuals (affecting the probability that their change in size is bigger or smaller at the next time step). We set the initial average phenotype  $\bar{z}_1(0)$  to 0, *i.e.*, on average all individuals in a given class grow neither faster nor slower than predicted for that species, and explore different scenarios for the initial phenotypic variance in the planted cohort,  $\sigma_z = 0$  being the control scenario where the probability of their change in size at the next time step remains unchanged, against higher values of variance. We consider heritability to be  $h^2 = 0.5$ , such that  $h^2 = \sigma_g^2 / \sigma_z^2$ .

### Parameters

Each simulation iterates the model for  $t = 50$  years and computes the number of individuals, as well as the average breeding and phenotypic values, in each age and size class at each time step. The reinforcements carried out in this population provided four times 10,000 individuals, over the course of several years. Table 5 from Ducrettet and Imbert 2023 summarizes the outcome of the reinforcement operations, for each operation between 410 and 730 seeds germinated into seedlings and survived until the following year. We chose to follow one of these seedlings cohorts, as if it was an introduction (no interaction with the natural populations), such that  $n_1(t = 0) = 500$ . We consider individuals to be initially distributed in each size class of the first age class following the predicted seedling size distribution,  $\omega(y)$ . Following Hadjou-Belaid et al. (in prep), each age

class is divided into 50 size classes such that  $l = 50$ . We calculate the asymptotic population growth rate  $\lambda$ , the dominant eigenvalue of the kernel, at each time step using the “popbio” package in R (Stubben and Milligan 2007). We verified the code for the evolutionary IPM by comparing the prediction given by the demographic IPM fitted by Hadjou-Belaid et al. (in prep), the sum of expressions (4), and the evolutionary explicit IPM, we checked that we obtain similar results as Hadjou-Belaid et al. *in prep* when we set the phenotypic variance such that  $\sigma_z = 0$  (see Supplementary Material S3 for details).

## Results

### Adding genetic variance on growth

We first show how increasing the average intrinsic growth rate value ( $\bar{z}$ ) does increase the probability that individual of size  $x$  at  $t$  grow to a bigger size at time  $t+1$  as a result of the intercept of the individual growth function being increased (Fig. 2.3A versus Fig. 2.3B). Although increasing phenotypic variance introduces a greater probability of faster size growth for some individuals, it also introduces a greater probability of slower growth for others (the red color is clearly more spread out on Fig. 2.3C and Fig. 2.3D).

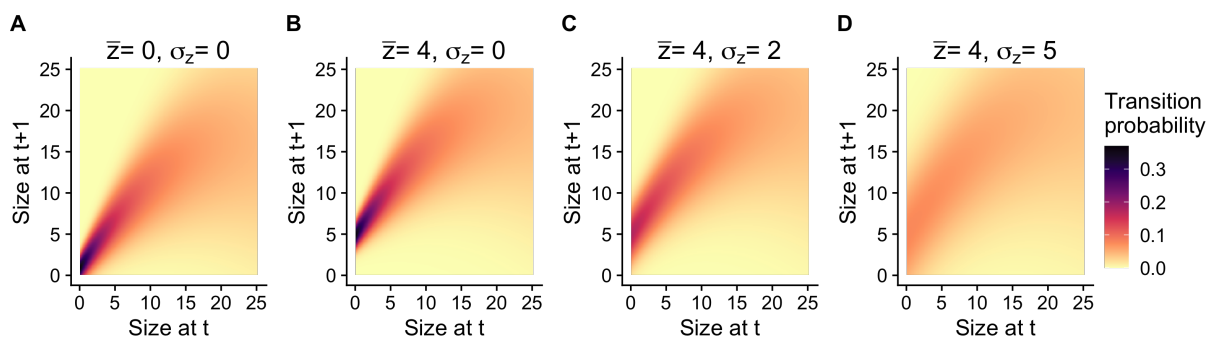


Fig. 2.3: Increasing the average individual intrinsic growth rate ( $\bar{z}$ ) increases the probability to grow bigger from size  $x$  at  $t$  to size  $y$  at  $t+1$ , although this distribution is more spread out if the phenotypic variance ( $\sigma_z$ ) is large. These growth probabilities have been computed for individuals of age  $a = 4$ .

We find that an initial input of phenotypic variance increases the average individual size intrinsic growth rate value in the cohort over time (Fig. 2.4). In each age class, the individual growth rate increases in time as selection acts on the phenotypic variance. This value eventually reaches a plateau, as becoming too big also reduces both survival and fecundity (Fig. 2.2). Note that this plateau is lower for older age classes than younger age classes. Selection thus increases

more the growth rate of the young plants than the older plants. This is due to the monocarpic life cycle of *Centaurea corymbosa*: as individuals die after flowering and flowering happens once a critical size has been reached; old individuals are those that grow fast enough to reach size allowing a good survival but not fast enough to reach the critical size for flowering. Increasing the amount of genetic variance allows faster evolution of the individual growth rate in young age classes, but the effect of increasing the phenotypic variance is more complex in older age classes because of the faster selective disappearance of fast growers, which flower and die early, from the old age classes.

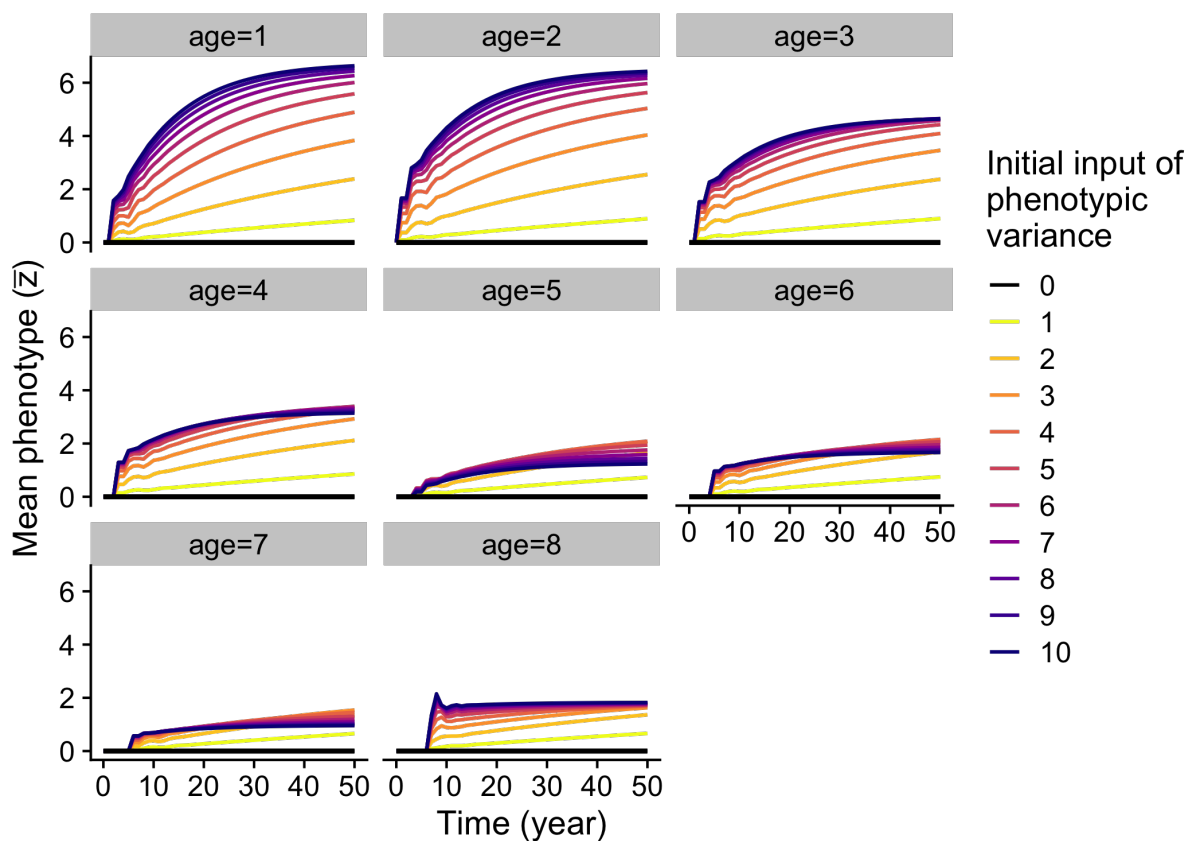


Fig. 2.4: The mean phenotypic value for the *C. corymbosa* cohorts is shown as a function of time for the scenarios with no phenotypic variance added (black curve) or with phenotypic variance added (colored curves).

As expected, the size distribution of seedlings (age=1) remains unchanged over time, even when there has been an initial input of phenotypic variation for growth rate as seedling size does not depend on maternal size in our model (coloured curves, Fig. 2.5). Within the other age classes, individuals do reach a larger size (the average size of individuals increases) with a greater initial input of phenotypic variance as the growth rate evolves to higher values. The increase in phenotypic variation in the initial cohort can also lead to an increase in the number of

individuals (yellow and pink curves), as opposed to the black curve without phenotypic variation. Note that it also lead to a decrease in the number of individuals (blue curve) if the initial input of phenotypic variance is too high.

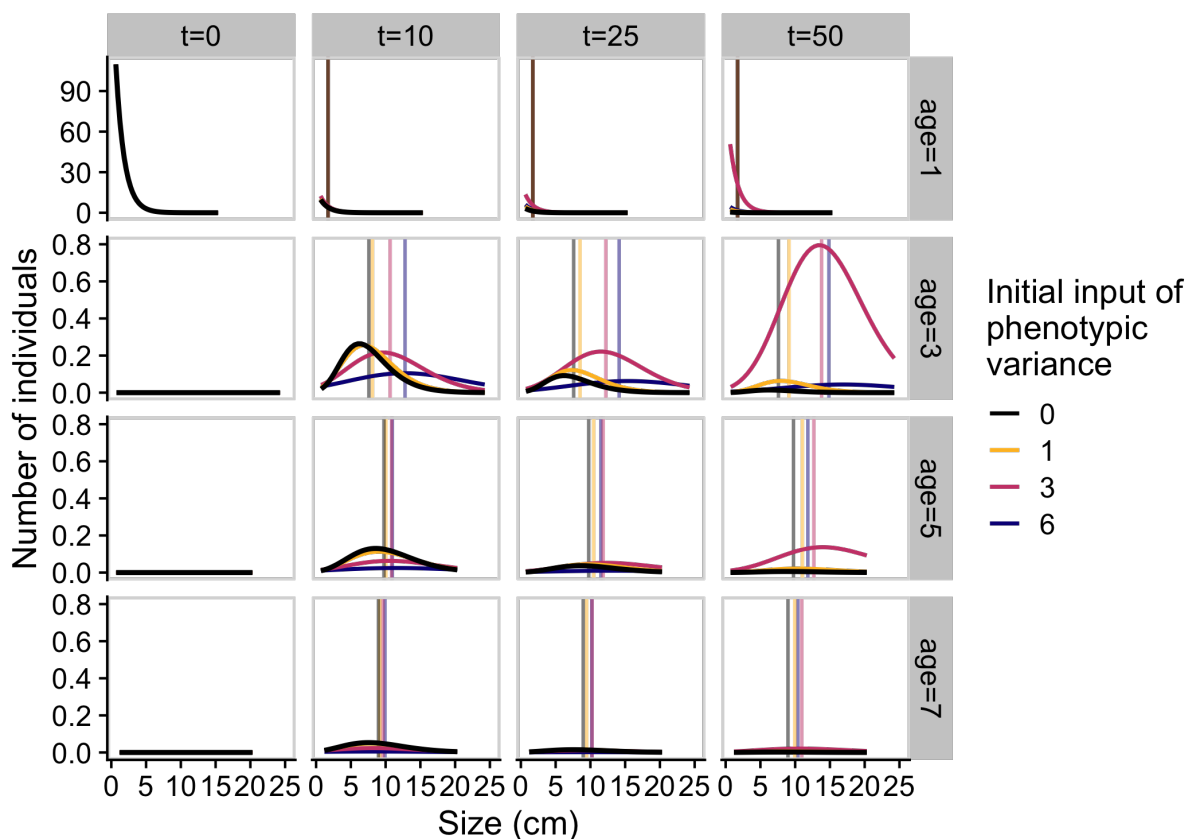


Fig. 2.5: The number of individuals and their size in the age classes 1, 3, 5 and 7 are shown at  $t = 0$ ,  $t = 10$ ,  $t = 25$  and  $t = 50$  years. The black curves represent the scenario with no initial input of phenotypic variance, while the colored curves represent different levels of initial phenotypic variance within the cohort.

### Evolutionary rescue of *C. corymbosa*

Without intervention, the size of the cohort of *C. corymbosa* decreases and heads towards extinction (black curve, Fig. 2.6). We show here that an input of phenotypic variance can lead the planted cohort to demographically expand. If the added phenotypic variance is too low ( $\sigma_z < 2$ ), the population fails to grow before 50 years, whereas if the added phenotypic variance is too high ( $\sigma_z > 6$ ), the population becomes extinct sooner than without any added phenotypic variance.

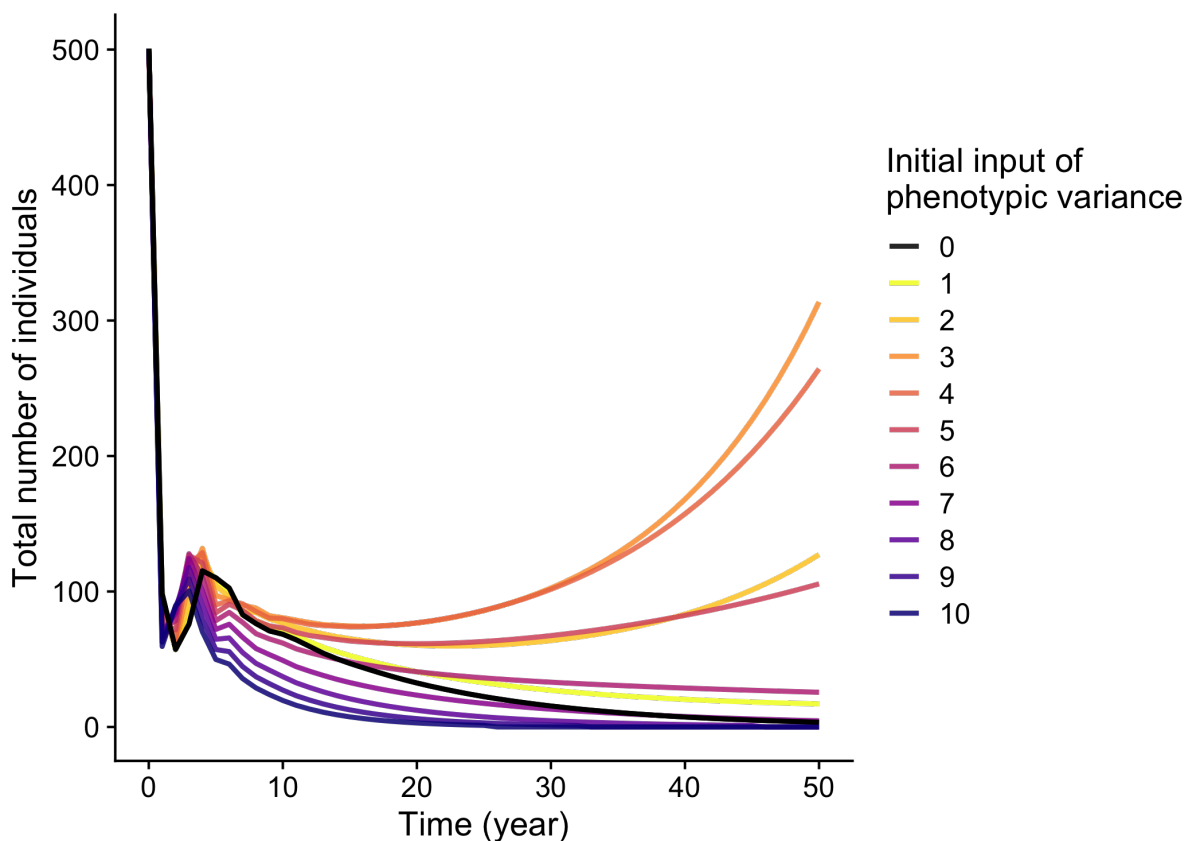


Fig. 2.6: The total number of individuals in the *C. corymbosa* cohort decreases with time, before eventually increasing again under certain scenarios of phenotypic variance input (colored curves), following a characteristic evolutionary rescue curve. This scenario involves a single introduced cohort, with no interaction with the natural populations.

Note that the disequilibrium in the age and size structure (single cohort planted) cause the initial fluctuations in the number of individuals (and the big drop in the first year where there are only fragile seedlings in the population).

The long-term population growth rate ( $\lambda$ ) only reaches the threshold of 1 during the 50 year simulation for scenarios of an input of phenotypic variation of  $2 < \sigma_z < 5$  (Fig. 2.7).



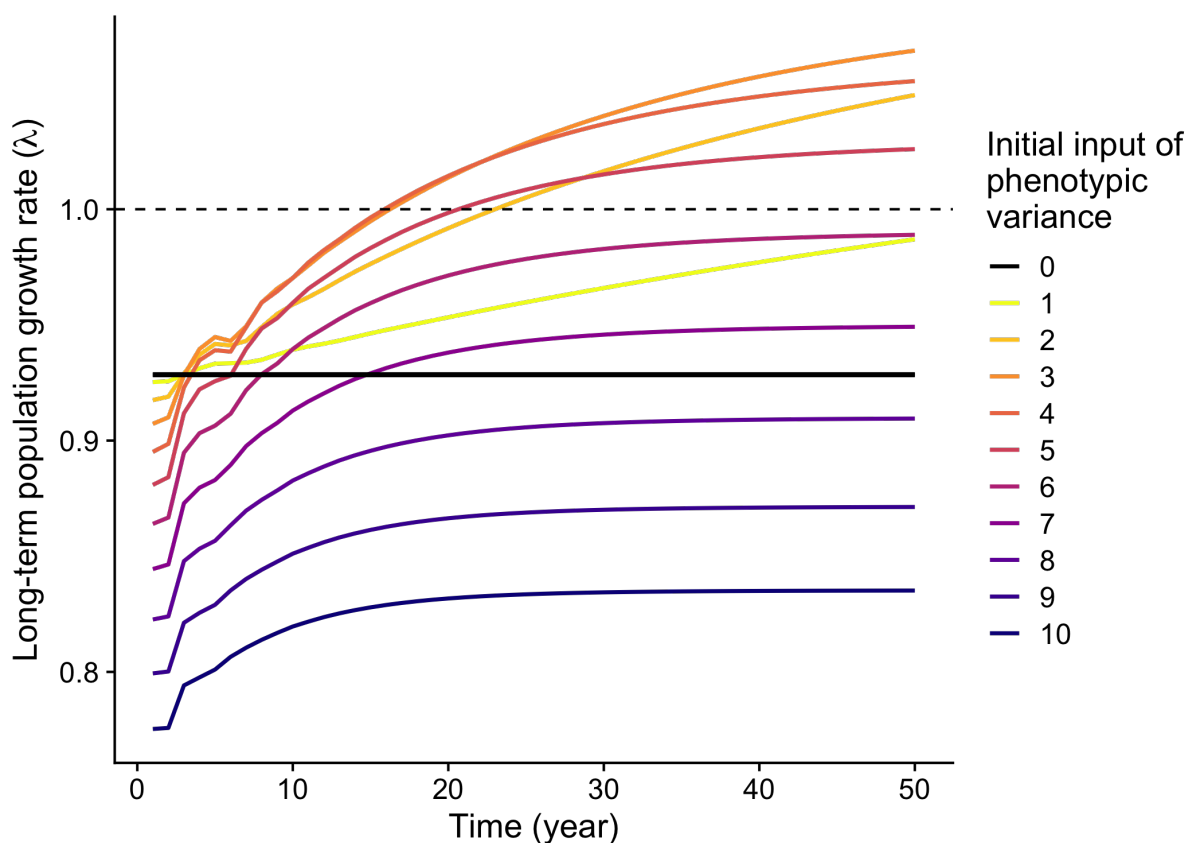


Fig. 2.7: The long-term growth rate of the cohort of *C. corymbosa* is shown as a function of time, for various scenarios of phenotypic variance input (colored curves) or without phenotypic variance (black curve). The threshold for population replacement ( $\lambda = 1$ ) is indicated by a dashed black line.

## Discussion

Reinforcements of declining populations can increase genetic variability and help reverse demographic trends (Gomulkiewicz and Holt 1995), particularly if certain genotypes have an advantage (*e.g.*, faster growth). We propose a new method to integrate evolutionary dynamics in an IPM model describing the demography of *C. corymbosa*. We predict the critical level of genetic diversity for a fitness-determining trait (Gomulkiewicz and Houle 2009), the individual size growth rate, that would be necessary for *C. corymbosa* to escape extinction in the medium term under the current climate.

## Development of a new evolutionary IPM

Our model assumes that intrinsic growth rate values are normally distributed in each class. We thus only followed the average value of each age and size class. This assumption may

be problematic, if the actual size distribution deviates from normality, especially after several generations. Individual-based simulations should enable us to check that we are not deviating too much from the predictions made by our model under this assumption. If it proves to be fairly correct, this assumption has the advantage of making the calculation more efficient when used for large matrices or a large number of years. The simulations presented here ran in 24 min.

### **Little hope for conservation efforts to rescue *C. corymbosa* from extinction**

We show that a certain initial variance in the intrinsic growth rate can reverse the declining trend of the population, because larger individuals have a better fitness overall. Yet, even in the best-case scenario, the size of the population does not increase for another 20 years, by which time the rate of extinction is already very high for some of the natural populations (Hadjou Belaid et al. 2018). We also find that too much initial variance induces variance load (Lande and Shannon 1996), which accelerates population extinction. Nevertheless, holding the phenotypic variance constant, probably leads to overly pessimistic expectations about the extent of this variance load: the initial mortality of individuals with a low growth rate will lead to an erosion of the variance initially introduced between stages and over time, and a lower variance load than that simulated here.

The average value of the trait for the intrinsic growth trait is constrained by the fact that individuals that grow very big have a drastic decrease in fitness. This could be due to the statistic models fitted to the data, and questions the confidence in the shape of this vital rate/size relationship fitted.

This model is simplistic in many ways, which makes our predictions optimistic overall. First, the fitted statistical models ignore variance between populations and fluctuations between years. We plan on testing the effect of environmental stochasticity by randomly drawing years to affect the vital rate estimations. This should not change the shape of the response, but population are likely to extinct sooner and take more time to recover. We also did not include environmental change, which is know to have a negative impact of population persistence (Hadjou Belaid et al. 2018). The model also ignores some processes that might further accelerate extinction of *C. corymbosa*. Notably, the population most probably suffers from Allee effects, which may result in a further decrease of the growth rate at low densities (Kirchner et al. 2006). Finally, we assume heritability is  $h^2 = 0.5$ , which is a parameter we could vary later.

Experiments are underway to estimate the amount of genetic variance on the growth rate in the reinforcements. Very preliminary results suggest there is very little variance in the growth rate between maternal families in the plants at the start of the growing season (Master's report of Alexander Fisher, 2023).

### **Acknowledgements**

We gratefully acknowledge the contributions of Asma Hadjou Belaid, Sandrine Maurice and Eric Imbert, who have kindly shared their data, R codes and input for this chapter. In particular, we would like to thank Eric for numerous discussions about *Centaurea corymbosa* and IPMs. We would also like to extend our gratitude to the numerous individuals engaged in the long-term monitoring of *Centaurea corymbosa*, who conduct fieldwork on multiple occasions each year.

## Supplementary Material 2

### S1 - Modeling framework of an evolutionary explicit IPM

We describe below a framework that allows the evolution of growth of *C. corymbosa* in an age- and size-structured model. Life cycle functions were previously fitted by Hadjou-Belaid et al. *in prep* with demographic data for *C. corymbosa* as described in the main text (Hadjou-Belaid et al. (in prep)).

#### Model assumptions

This model assumes

- growth is genetically variable in the population,
- fecundity depends only on the size  $x$  and the age  $a$  of the parents (*i.e.* not on their breeding value  $g$ ),
- parents mate at random and there is no pollen limitation,
- an infinitely large number of genes influence the traits (infinitesimal model),
- for the sake of simplicity we here present the case where the phenotype is fully heritable ( $z=g$ ) and we follow the breeding value,  $g$ , for the trait,
- the phenotype is not considered to be normally distributed in each class by default.

**Variable glossary**

Suppl. Tab. 2.1: Summary of variables used.

Notation	Parameter
$t$	time step
$a$	age
$x$	size at $t$
$y$	size at $t + 1$
$g$	breeding value
$\Omega$	range of sizes
$\mathbb{G}$	range of breeding values
$n_{a,g}(x, t)$	number of individuals of age $a$ , breeding value $g$ and size $x$ at $t$
$f_a(y, x)$	fecundity function ( $f'_a$ for females, $f''_a$ for males)
$p_{a,g}(y, x)$	survival-growth function
$\phi_a(x)$	flowering function
$c_a(x)$	number of capitula produced
$s_0$	establishment rate
$s_a(x)$	survival function
$G_{a,g}(y, x)$	growth function
$\omega(y)$	Probability Density Function (PDF) of seedling size
$\gamma(g)$	PDF of seedling breeding value
$\gamma_f$	PDF of breeding values among all mothers
$\gamma_m$	PDF of breeding values among all fathers

**Implementation of the IPM**

We describe the change of a continuous age- and size-structured population over discrete time with an Integral Projection Model (IPM; Easterling et al. 2000; Childs et al. 2003, Rees et al. 2019). Individuals are characterized by their age  $a$ , breeding value  $g$ , and size  $x$  and the IPM

tracks their distribution  $n_{a,g}(x, t)$  at time each time step  $t$ .

**Breeding value of descendants** The Probability Density Function (PDF) representing breeding value distribution among all mothers is  $\gamma_f$ . The probability that a mother is of breeding value  $g$  is

$$\gamma_f(g) = \frac{\sum_{a=1}^{a_{max}} \int_{\Omega} f'_a(\bullet, x) n_{a,g}(x, t) dx}{n_{1,\bullet}(\bullet, t_{+1})}, \quad (\text{S2.1})$$

the total number of descendants produced by a mother of breeding value  $g$ , with  $f'_a(\bullet, x) = \int_{\Omega} f'_a(y, x) dy$  the number of offspring of all sizes born to a parent of size  $x$  and age  $a$ , divided by the total number of descendants produced,  $n_{1,\bullet}(\bullet, t_{+1}) = \int_{\Omega} \int_{\mathbb{G}} n_{a,g}(x, t) dy dg$ .

The PDF representing breeding value distribution among fathers is  $\gamma_m$ . The probability that a father is of breeding value  $g$  is

$$\gamma_m(g) = \frac{\sum_{a=1}^{a_{max}} \int_{\Omega} f''_a(\bullet, x) n_{a,g}(x, t) dx}{\int_{\mathbb{G}} \sum_{a=1}^{a_{max}} \int_{\Omega} f''_a(\bullet, x) n_{a,g'}(x, t) dx dg'} \quad (\text{S2.2})$$

the total number of descendants produced by a father of breeding value  $g$  divided by the total quantity of pollen available.  $g'$  is a notation used to represent all possible breeding values of fathers.

The probability that a seedling has the breeding value  $g$  is then

$$\gamma(g) = \int_{\mathbb{G}_f} \int_{\mathbb{G}_m} R(g|g', g'') \gamma_m(g') \gamma_f(g'') dg' dg'', \quad (\text{S2.3})$$

with  $g''$  the notation used to represent all possible breeding values of mothers,  $R(g|g', g'')$  the probability for an offspring to be of breeding value  $g$  given that its mother is of breeding value  $g''$  and its father is of breeding value  $g'$ . In the infinitesimal model of inheritance,

$$R(g|g', g'') = \frac{1}{\sqrt{2\pi\text{VLE}}} \exp \frac{-(g - \frac{g'+g''}{2})^2}{2\text{VLE}}, \quad (\text{S2.4})$$

with variance VLE.

**Individual distribution across demographic categories** The number of individuals in their 1<sup>st</sup> year ( $a = 1$ ) of breeding value  $g$  and size  $y$  produced by parents of breeding value  $g$  at

the next time step  $t + 1$  can be written as

$$\begin{aligned} n_{1,g}(y, t + 1) &= \omega(y) \gamma(g) n_{1,\bullet}(\bullet, t + 1) \\ &= \omega(y) \int_{\mathbb{G}_f} \int_{\mathbb{G}_m} R(g|g', g'') \gamma_f(g) \gamma_m(g) dg dg' \sum_{a=1}^{a_{max}} \int_{\Omega} f_a(\bullet, x) n_{a,\bullet}(x, t) dx, \end{aligned} \quad (\text{S2.5})$$

where,  $\omega(y)$  is the PDF of offspring size,  $f_a(\bullet, x)$  is the total number of offspring produced by females of size  $x$  and age  $a$ ,  $n_{a,\bullet}(x, t)$  is the total number of parents of size  $x$ , age  $a$  at  $t$  integrated over the pool of all breeding value such that  $n_{a,\bullet}(x, t) = \int_{\mathbb{G}} n_{a,g}(x, t) dg$ .

The number of vegetative individuals older than 1 year old can be written as

$$n_{a+1,g}(y, t + 1) = \int_{\Omega} p_{a,g}(y, x) n_{a,g}(x, t) dx. \quad (\text{S2.6})$$

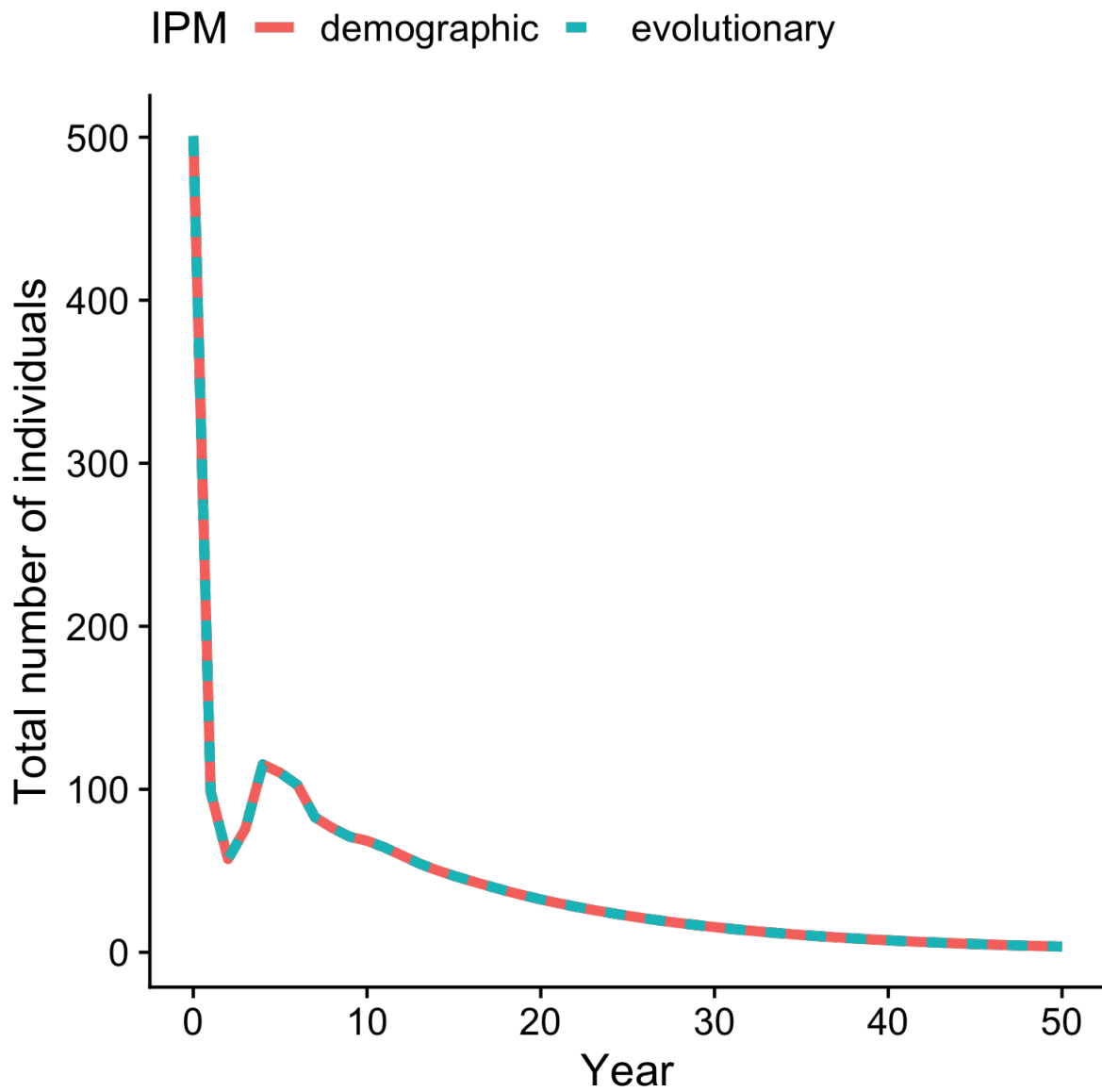
## S2 - Size range of *C. corymbosa* at each age

Age	Minimum size (cm)	Maximum size (cm)
1	0.5	15.5
2	0.5	20.5
3	0.5	24.5
4	0.5	25.5
5	0.5	20.5
6	0.5	25.5
7	1	20.5
8	0.5	25.5

Suppl. Tab. 2.2: Extreme size values for each age group of *C. corymbosa*.

## S3 - Comparison between the demographic and the evolutionary IPM

We compared the prediction given by the demographic IPM adjusted by Hadjou-Belaid et al. (in prep) (as described in ‘‘Description of the Integral Projection Model’’ of the main text) and the evolutionary explicit IPM with the initial phenotypic variance set to  $\sigma_z = 0$ , such that the probability of their change in size at the next time step remains unchanged, to check that they yielded identical predictions, which they do (Suppl. Fig. 1).



Suppl. Fig. 2.1: Comparing the predictions of demographic trends between the demographic-only IPM (pink) and the evolutionary IPM with the initial phenotypic variance set to  $\sigma_z = 0$  (blue) yields identical results.



## Appendix B: R code for the evolutionary IPM

Adèle Erlichman & Ophélie Ronce

### Define vital rate functions

#### Survival probability

```
s.x <- function(a, x) {  
  u = exp(- 2.411952 + 0.773796 * x - 0.032601 * x^2 + 0.102227 * a)  
  return(u / (1 + u)) # logit  
}
```

#### Flowering probability

```
f.x <- function(a, x) {  
  u = exp(-8.95979 + 0.717488 * x - 0.016180 * x^2 + 0.426482 * a)  
  return(u / (1 + u))  
}
```

#### Number of capitula

```
c.x <- function(a, x) {  
  value = 0.4536 + 1.9540 * x + 1.4276 * a  
  return(value)  
}
```

#### Seedling size distribution

```
ssd.y <- function(y) {  
  value = exp(-0.583055 - 0.824603 * y)  
  return(value)  
}
```

#### Establishment rate

```
s0 <- 0.560
```

#### Evolution of the growth function

```
muG <- function(a, x) {  
  value = 2.005839 + 1.340332 * x - 0.028022 * x^2 - 0.332909 * a  
  return(value)  
}
```

```

sigmaG <- function(x) {
  value = 0.9292 + 0.3153 * x
  return(value)
}

G.yxz <- function(a, x, y, z, sigmaz) {
  value = 1 / (sqrt(2 * pi * (sigmaG(x)^2 + sigmaz^2))) *
    exp(-(z - (y - muG(a, x)))^2 / (2 * (sigmaG(x)^2 + sigmaz^2)))
  return(value)
}

```

## Functions that describe life history

```

# fecundity function
fmean <- function(a, y, x) {
  value = s0 * f.x(a, x) * c.x(a, x) * ssd.y(y)
  valueFec = ifelse(value >= 0, value , 0)
  return(valueFec)
}

# survival growth function (mean transition rate from j to i)
tmean <- function(a, y, x, z, sigmaz) {
  value = s.x(a, x) * (1 - f.x(a, x)) * G.yxz(a, x, y, z, sigmaz)
  return(value)
}

# derivative of the log transition rate from j to i
# with respect to the mean phenotypic value (z) at j
dlntmean <- function(a, y, x, z, sigmaz) {
  value = (theta.xy(a, x, y) - z) / (sigmaG(x)^2 + sigmaz^2)
  return(value)
}

```

## Transition matrix generator

```

# Builds the transition matrix using the z vector
TransitionMatrixGenerator <-
  function(a, zvec, sigmazvec, n, min.size, max.size) {
    # individuals of age a
    b = min.size[a] + c(0:n) * (max.size[a] - min.size[a]) / n # class boundaries
    y.a = 0.5 * (b[1:n] + b[2:(n + 1)]) # midpoint of each class

    # individuals of age a+1 (a1)
    if (a != age.max) {
      b = min.size[a + 1] + c(0:n) * (max.size[a + 1] - min.size[a + 1]) / n
      # class boundaries
      y.a1 = 0.5 * (b[1:n] + b[2:(n + 1)]) # midpoint of each class
      h.a1 = y.a[2] - y.a[1] # step size for individuals of age a+1 (a1)
    } else {
      b = min.size[age.max] + c(0:n) * (max.size[age.max] - min.size[age.max]) / n
      # class boundaries
      y.a1 = 0.5 * (b[1:n] + b[2:(n + 1)]) # midpoint of each class
      h.a1 = y.a[2] - y.a[1] # step size for individuals of age max
    }
  }

```

```

}

# individuals of age 1
b = min.size[1] + c(0:n) * (max.size[1] - min.size[1]) / n # class boundaries
y.1 = 0.5 * (b[1:n] + b[2:(n + 1)]) # midpoint of each class
h.1 = y.1[2] - y.1[1] # step size for individuals of age 1

# Define zvec and sigmaz for this age class a
zvec_a = zvec[(a*n-(n-1)):(a*n)]
sigmazvec_a = sigmazvec[(a*n-(n-1)):(a*n)]

# Survival - Growth matrix
# Define empty matrix
SG = matrix(c(NA), ncol = length(y.a), nrow = length(y.a))
for (i in c(1:length(y.a))) {
  for (j in c(1:length(y.a))) {
    SG[i, j] <-
      h.a1 * tmean(a, y.a1[i], y.a[j], zvec_a[j], sigmazvec_a[j])
  }
}

# Fecundity matrix
F = matrix(c(NA), ncol = length(y.a), nrow = length(y.a))
for (i in c(1:length(y.a))) {
  for (j in c(1:length(y.a))) {
    F[i, j] <-
      h.1 * fmean(a, y.1[i], y.a[j])
  }
}

# Full matrix
kernel = SG + F

# dlntij
dlnSG = matrix(c(NA), ncol = length(y.a), nrow = length(y.a))
for (i in c(1:length(y.a))) {
  for (j in c(1:length(y.a))) {
    dlnSG[i, j] <-
      h.a1 * dlntmean(a, y.a1[i], y.a[j], zvec_a[j], sigmazvec_a[j])
  }
}

return(list(
  K = kernel, # full kernel (TM+FM)
  TM = SG, # transition matrix
  FM = F, # Fecundity matrix
  dlnTM = dlnSG # derivative of the log of TM
))
}

```

## Full matrix generator

```
# Builds the age + size structured Kernel
MatrixGenerator = function(age.max, zvec, sigmazvec, MatrixDim, min.size, max.size) {
  TM.k = matrix(data = NA,
                nrow = MatrixDim,
                ncol = MatrixDim)
  FM.k = matrix(data = NA,
                nrow = MatrixDim,
                ncol = MatrixDim)
  K.k = matrix(data = NA,
                nrow = MatrixDim,
                ncol = MatrixDim)
  Kernals = array(data = 0,
                  dim = c(age.max, 3, MatrixDim, MatrixDim))
  superMatrix = array(0, dim = c((age.max * MatrixDim),
                                (age.max * MatrixDim)))
  for (a in (1:(age.max - 1))) {
    Kls = TransitionMatrixGenerator(a, zvec, sigmazvec, MatrixDim,
                                    min.size, max.size)

    Kernals[a, 1, , ] = Kls$K
    Kernals[a, 2, , ] = Kls$TM
    Kernals[a, 3, , ] = Kls$FM
    debut = (a - 1) * MatrixDim + 1
    fin = debut + (MatrixDim - 1)
    superMatrix[1:MatrixDim, (debut:fin)] = Kls$FM
    superMatrix[((fin + 1):(fin + MatrixDim)), (debut:fin)] = Kls$TM
  }

  Kls = TransitionMatrixGenerator(age.max, zvec, sigmazvec, MatrixDim,
                                  min.size, max.size)

  Kernals[age.max, 1, , ] = Kls$K
  Kernals[age.max, 2, , ] = Kls$TM
  Kernals[age.max, 3, , ] = Kls$FM
  debut = (age.max - 1) * MatrixDim + 1
  fin = debut + (MatrixDim - 1)
  superMatrix[1:MatrixDim, (debut:fin)] = Kls$FM
  superMatrix[(debut:fin), (debut:fin)] = Kls$TM

  return(superMatrix)
}
```

## Full dln matrix generator

```
# Builds the dlnTT matrix
dlnMatrixGenerator = function(age.max, zvec, sigmazvec, MatrixDim, min.size, max.size) {
  TM.k = matrix(data = NA,
                nrow = MatrixDim,
                ncol = MatrixDim)
  FM.k = matrix(data = NA,
                nrow = MatrixDim,
                ncol = MatrixDim)
  K.k = matrix(data = NA,
```

```

        nrow = MatrixDim,
        ncol = MatrixDim)
Kernels = array(data = 0,
                dim = c(age.max, 3, MatrixDim, MatrixDim))
superMatrix = array(0, dim = c((age.max * MatrixDim),
                               (age.max * MatrixDim)))
for (a in (1:(age.max - 1))) {
  Kls = TransitionMatrixGenerator(a, zvec, sigmazvec, MatrixDim,
                                min.size, max.size)

  debut = (a - 1) * MatrixDim + 1
  fin = debut + (MatrixDim - 1)
  superMatrix[1:MatrixDim, (debut:fin)] = 0
  superMatrix[((fin + 1):(fin + MatrixDim)), (debut:fin)] = Kls$dlnTM
}

Kls = TransitionMatrixGenerator(age.max, zvec, sigmazvec, MatrixDim,
                                min.size, max.size)

Kernels[age.max, 1, , ] = Kls$K
Kernels[age.max, 2, , ] = Kls$TM
Kernels[age.max, 3, , ] = Kls$FM
debut = (age.max - 1) * MatrixDim + 1
fin = debut + (MatrixDim - 1)
superMatrix[1:MatrixDim, (debut:fin)] = 0
superMatrix[(debut:fin), (debut:fin)] = Kls$dlnTM

return(superMatrix)
}

```

## General recursion

```

# The general recursion defines the phenotypic and breeding values
# based on the transition matrices

GeneralRecursion <-
function(g0, z0, NO, sigmag, sigmaz, age.max, MatrixDim, min.size, max.size) {
  # define empty vectors for phenotype and breeding value
  zt1 = rep(NA, length(NO))
  gt1 = rep(NA, length(NO))

  # full kernel
  K = MatrixGenerator(age.max, z0, sigmaz, MatrixDim, min.size, max.size)

  # compute number of individuals at t+1
  Nt1 = K %*% NO

  # fecundity matrix
  FF = K
  FF[(MatrixDim + 1):(MatrixDim * age.max),] = 0

  # transition matrix
  TT = K
  TT[1:MatrixDim,] = 0
}

```

```

# derivative of the log transition matrix
dlnTT = dlnMatrixGenerator(age.max, z0, sigmaz, MatrixDim, min.size, max.size)

# recursion for g values at t
for (i in 1:(MatrixDim * age.max)) {
  if (Nt1[i] == 0) {
    gt1[i] = 0
  } else {
    j = 1:(MatrixDim * age.max)
    gt1[i] = sum((NO[j]*TT[i, j]/Nt1[i]) * (g0[j] + sigmag[j]^2 *dlnTT[i, j]) +
                ((NO[j]*FF[i, j]/Nt1[i])*(g0[j])))
  }
}

# recursion for z values at t
for (i in 1:(MatrixDim * age.max)) {
  if (Nt1[i] == 0) {
    zt1[i] = 0
  } else {
    j = 1:(MatrixDim * age.max)
    zt1[i] = sum((NO[j]*TT[i, j]/Nt1[i]) * (z0[j] + sigmaz[j]^2 *dlnTT[i, j]) +
                ((NO[j]*FF[i, j]/Nt1[i])*(g0[j])))
  }
}

return(list(gt = gt1, zt = zt1))
}

```

## Running recursion

```

RunRecursion <- function(g0, z0, NO, sigmag, sigmaz, age.max,
                        MatrixDim, min.size, max.size, nyears) {
  # Define empty lists
  zt = list()
  gt = list()
  Nt = list()
  TT = list()

  # Initialize for t=0
  zt[[1]] <- z0
  gt[[1]] <- g0
  Nt[[1]] <- NO

  # Recursion
  for (t in 1:nyears) {
    TT[[t]] = MatrixGenerator(age.max, zt[[t]], sigmaz, MatrixDim,
                              min.size, max.size)
    zt[[t + 1]] = GeneralRecursion(gt[[t]],
                                   zt[[t]],
                                   Nt[[t]],
                                   sigmag,
                                   sigmaz,
                                   age.max,

```

```

MatrixDim,
min.size,
max.size)$zt
gt[[t + 1]] = GeneralRecursion(gt[[t]],
                               zt[[t]],
                               Nt[[t]],
                               sigmag,
                               sigmaz,
                               age.max,
                               MatrixDim,
                               min.size,
                               max.size)$gt
Nt[[t + 1]] = TT[[t]] %*% Nt[[t]]
print(t)
}

# Data wrangling for nice outputs (number of individuals in each class)
TabNt = plyr::ldply(Nt, rbind)
year = rep(0:nyears, each = MatrixDim * age.max)
age = rep(rep(1:age.max, each = MatrixDim), nyears + 1)
size.age = rep(rep(1:MatrixDim), age.max*(nyears+1))
size = vector()
compteur=1
for(a in 1:age.max){
  b = min.size[a] + c(0:MatrixDim) * (max.size[a] - min.size[a]) / MatrixDim
  midpoints = 0.5 * (b[1:MatrixDim] + b[2:(MatrixDim + 1)])
  size[compteur:(compteur+(MatrixDim-1))] = midpoints
  compteur=compteur+MatrixDim
}
sizevec = rep(size, nyears + 1)
zini = rep(z0[1], each = MatrixDim * age.max)
gini = rep(g0[1], each = MatrixDim * age.max)
sigmazini = rep(sigmaz[1], each = MatrixDim * age.max)
sigmagini = rep(sigmag[1], each = MatrixDim * age.max)

NTT <-
  data.table(year, age, size.age, sizevec, zini, gini, sigmazini, sigmagini, TabNt)
colnames(NTT) <- c("year", "age", "size.age", "size", "zini", "gini",
                  "sigmazini", "sigmagini", "Nind")

# Data wrangling for nice outputs (breeding values)
GT = plyr::ldply(gt, rbind)
year = rep(0:nyears)
GT <- data.table(year, GT)
GT = data.table::melt(GT, id.vars = "year")
GT$size = rep(rep(1:MatrixDim, each = nyears + 1), age.max)
GT$age = rep(1:age.max, each = (nyears + 1) * MatrixDim)
colnames(GT) <- c("year", "age.size", "gt", "size", "age")

# Data wrangling for nice outputs (phenotypic values)
ZT = plyr::ldply(zt, rbind)
year = rep(0:nyears)
ZT <- data.table(year, ZT)

```

```

ZT = data.table::melt(ZT, id.vars = "year")
ZT$size = rep(rep(1:MatrixDim, each = nyears + 1), age.max)
ZT$age = rep(1:age.max, each = (nyears + 1) * MatrixDim)
colnames(ZT) <- c("year", "age.size", "zt", "size", "age")

return(list(
  NT = NTT,
  gT = GT,
  zT = ZT,
  TM = TT
))
}

```

## Simple run exemple

```

# Define matrix dimension
MatrixDim = 50
# Define number of years for simulation
nyears = 50

# Define oldest age class
age.max = 8
# Define max size in each age class
max.size = c(15.5, 20.5, 24.5, 25.5, 20.5, 25.5, 20.5, 25.5)
# Define min size in each age class
min.size = c(0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 1, 0.5)

# To define the distribution of seedlings in the first cohort
boundaries = min.size[1] + c(0:MatrixDim) * (max.size[1] - min.size[1]) / MatrixDim
midpoints = 0.5 * (boundaries[1:MatrixDim] + boundaries[2:(MatrixDim + 1)])
# Consider 500 seedlings
N0 <- matrix(c(500*ssd.y(midpoints), rep(0, (MatrixDim * age.max)-MatrixDim)), ncol = 1)

# Initial phenotypic values: 0 in each class
z0 = matrix(rep(0, MatrixDim * age.max), nrow = 1)
# Initial breeding values: 0 in each class
g0 = matrix(rep(0, MatrixDim * age.max), nrow = 1)

# Define input of variance
z=2
sigmaz = matrix(rep(z, MatrixDim * age.max), ncol = 1)
# consider heritability is 0.5
sigmag = matrix(rep(sqrt(0.5) * z, MatrixDim * age.max), ncol = 1)

# Run
output = RunRecursion(g0, z0, N0, sigmag, sigmaz, age.max,
  MatrixDim, min.size, max.size, nyears)

```



## Chapter 3

# Evolutionary rescue by means of introgression: balancing genomic rescue and swamping

Adèle Erlichman<sup>1,2\*</sup>, Freek J.H. de Haas<sup>2</sup>, Sarah P. Otto<sup>2</sup>

<sup>1</sup> ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France

<sup>2</sup> Department of Zoology, University of British Columbia, Vancouver, Canada

### Abstract

Theory suggests that the likelihood of evolutionary rescue, where evolution prevents otherwise inevitable extinction, increases with population size and availability of genetic variation. Introgression, which refers to the back-crossing of hybrids with one or both parental lineages, can boost evolutionary rescue by introducing new genetic variation into endangered populations. Although cases of adaptive introgression are known, more often it results in demographic or genetic swamping. We have analyzed a deterministic model, supported by individual-based simulations, to quantify the effects of genetic swamping after a single introduction event of individuals from another population. Our analytical results identify an optimal number of introduced individuals that maximizes the surviving portion of the focal population's genome. Assuming the focal population has declined and exerts little competition, small and early introductions minimize swamping, because the resident population remains large, while larger introductions maximize the probability of rescue. Not only do our analytical results provide insights into the trade-off between genetic swamping and rescue, they also provide tools to guide conservation managers aiming to rescue small and/or inbred populations.

*Keywords:* conservation genetics, translocation, hybridization, genetic swamping, genetic rescue.

## Introduction

Long-term environmental changes pose an additional challenge to the persistence of populations that may already be in decline, due to other anthropogenic threats (IPCC 2021; Parmesan 2006; Román-Palacios and Wiens 2020). The ability of populations to respond to these changes through genetic adaptation depends on the strength of selection relative to their adaptive capacity (Gomulkiewicz and Holt 1995; Lynch and Lande 1993; Orr and Unckless 2008). Scenarios where evolution can reverse demographic threats and so prevent otherwise inevitable extinction have been termed evolutionary rescue (Gomulkiewicz and Holt 1995). Theory predicts that the probability of evolutionary rescue decreases with stress intensity and increases with initial population size and with the abundance of genetic variation available to fuel adaptation to new conditions (for a review see Bell and Collins 2008). The potential role and importance of hybridization (*i.e.*, the crossing of different species or populations) as a conservation management tool for evolutionary rescue is currently being debated across taxa (Bell et al. 2019; Chan et al. 2019; Hamilton and Miller 2016), including crops (Burgarella et al. 2019), forest trees (Cortés et al. 2020) or animals (Hedrick 2013).

Hybrid back-crossing with one or both parental lineages (*i.e.*, introgression) can increase adaptive potential by introducing genetic variation through gene flow between populations. Introgression that results in adaptive evolution is referred to “adaptive introgression”. However, introgression does not always result in evolutionary rescue. Too much hybridization can lead to (1) demographic or (2) genetic swamping of the target population. Demographic swamping occurs when reproductive effort of the target population is diverted toward the production of low fitness hybrid offspring (outbreeding depression) (Wolf et al. 2001), decreasing population growth rate and hastening extinction. Genetic swamping occurs when hybrids are so fit (hybrid vigor) or so frequent that there is substantial gene flow into the target species. The resulting gene flow can lead to genetic loss of all or part of the endangered population’s genome (Allendorf et al. 2001; Woodruff and Gould 1987). High levels of introgression can thus pose a threat to preservation of the parental lineage (Rhymer and Simberloff 1996).

Studies to date are inconclusive about the most common outcome of introgressive hybridization. A literature review by Todesco *et al.* (2016) indicates that genetic and demographic swamping are more frequent than genetic rescue, with genetic swamping being much more frequent than demographic swamping (Todesco et al. 2016). Although examples of swamping have

been documented *e.g.*, Isle Royale wolves (*Canis lupus*; Adams et al. 2011; Hedrick et al. 2014), wildcats (*Felis silvestris*; Howard-McCombe et al. 2023), there are numerous examples of adaptive introgression for species rescue *e.g.*, Florida panthers (*Puma concolor coryi*; Johnson et al. 2010; Pimm et al. 2006a), Scandinavian wolves (*Canis lupus*; Åkesson et al. 2016; Vila et al. 2003), Iberian lynx (*Lynx pardinus*; Lucena-Perez et al. 2024) and for adaptation to novel environmental conditions *e.g.*, domestic goats (*Capra hircus*; Zheng et al. 2020), Tibetan mastiff (*Canis lupus familiaris*; Miao et al. 2017). In some instances, the parental genome has even remained relatively intact despite long histories of association and potential interbreeding *e.g.*, adders (*Vipera berus*; Madsen et al. 2004), along with cases of stable hybrid zone formation *e.g.*, darters (between *Etheostoma osburni* and *Etheostoma variatum*; Gibson et al. 2019).

One possible reason for genetic rescue appearing to be an uncommon outcome of hybridization in the studies summarized by Todesco *et al.* (2016) is the authors' choice of search terms, which included "extinction" but not "rescue". Furthermore, many studies of hybridization involve communities with on-going hybrid formation, making swamping more likely. Unfortunately, there is still little guidance as to when and how much hybridization might be beneficial in a conservation context (Bell et al. 2019). Lack of clear guidelines and practical tools contributes to the reluctance to engage in genetic rescue via hybridization out of concerns about losing the genetic distinctiveness of local populations and potentially contributing to extinction risk (Liddell et al. 2021; Weeks et al. 2016).

Due to its unpredictable and potentially damaging effects, genetic rescue by means of introgression remains a controversial conservation management tool (Bohling 2016). For species that can be bred in captivity, controlled back-crosses between distinct population can limit the harmful effects of both demographic and genetic swamping. For species that cannot be easily bred, however, it is challenging to predict the outcome of introgression. For conservation purposes the long-term effects of hybridization matter, but for many species little is known about hybrid fitness beyond the first few generations of crosses. Models can help to evaluate the relative risks of genetic swamping and to guide conservation approaches for evolutionary rescue of wild populations (see for instance Aitken and Whitlock 2013; Baskett and Gomulkiewicz 2011; Maroso et al. 2023; Vedder et al. 2022; Zilko et al. 2021). Simulations of Human/Neanderthal introgression has shown that genetic rescue by introgression leads to the loss of a portion of the native genome proportional to the fitness increase in the population, assuming the focal population has a stable effective size (Harris et al. 2019). Previous work explored the trade-off between maintaining the

recipient population’s genome and maximizing survival, with simulations, seeking to maximize both genetic diversity and population survival under continuous step-wise environmental change of different magnitude affecting the optimal value for a polygenic quantitative trait (Kelly and Phillips 2019a; Smart and Phillips 2023). These simulations showed that the timing of the introduction and the size of the introduced cohort are the most important parameters for maximizing the success of the targeted gene flow.

Here, we present an analytical framework that compare the long-term risks of genomic swamping relative to genetic rescue, allowing exploration beyond particular parameters. We focus on cases of introgression between genetically distinct populations or sub-species where hybrid fitness is moderate to high. Our model can be interpreted as a reference model for genome rescue by means of introgression and can be used to assess the optimal number of introduced individuals to minimize genetic swamping while allowing evolutionary rescue. We also provide individual-based simulations coded in SLiM (Haller and Messer 2017), which track changes across the genome and can be modified to explore parameters (see Perspectives) most appropriate for a species of interest.

## Methods

### Model description

We begin by describing the simplest scenario where one adaptive allele found in a related population can boost fitness of a declining population of concern. We refer to these populations as the *non-resident* and *resident* population, respectively. We track the fate of distinctive neutral alleles that differentiate the resident population, focusing on a particular locus **A**, following the introduction of an adaptive allele *B* at locus **B** from the non-resident population. The *B* allele could represent a gene of large effect that underlies adaptation to the current biotic or abiotic environment. We consider a randomly mating diploid population with crossovers occurring between loci **A** and **B** at rate  $c$ . At generation  $n = 0$ , non-residents of haplotype “aB” are introduced at a frequency of  $p_0$  into a population of residents, consisting of “Ab” haplotypes. Individuals bearing the *B* allele have a fitness advantage  $s$  ( $W_B = 1 + s$ ) over the resident individuals ( $W_b = 1$ ). We call the hybrid “AB” the *rescue* haplotype, as it combines the fitness advantage of allele *B* with the locally distinctive allele *A*.

We follow the notation of Maynard Smith and Haigh (1974), who investigated genetic hitch-

hiking of neutral alleles alongside the spread of an adaptive allele. The extent of hitchhiking was made analytically tractable by changing variables from haplotype frequencies to the frequency of the adaptive allele  $p_n$  across generations  $n$ , alongside the frequency of the neutral allele  $A$  among carriers of the adaptive  $B$  allele (called  $Q_n$ ) and among carriers of the non-adaptive  $b$  allele (called  $R_n$ ). The notation is summarized in Table 3.1. For the problem at hand, we seek to maximize  $Q_n$ , the frequency of the non-resident haplotype, which combines both the adaptive allele  $B$  and the resident allele at the neutral locus  $A$ .

Haplotype	$AB$	$aB$	$Ab$	$ab$
Frequency	$p_n Q_n$	$p_n(1 - Q_n)$	$(1 - p_n)R_n$	$(1 - p_n)(1 - R_n)$
Fitness	$1 + s$	$1 + s$	1	1

Tab. 3.1: Summary table of haplotypes, their frequencies at generation  $n$ , and their fitnesses

As locus **A** is neutral, the dynamics of allele  $B$  with fitness advantage  $s$  is given by:

$$p_n = \frac{p_0(1 + s)^n}{1 - p_0 + p_0(1 + s)^n}, \quad (3.1)$$

The initial frequency  $p_0$  is determined by when the introduction occurs and how large it is. If the resident population size currently consists of  $N_R$  haplotypes, then an introduction of  $N_I$  haplotypes gives  $p_0 = \frac{N_I}{N_R}$ . To be consistent with Maynard Smith & Haigh's haploid model, we count haplotypes throughout, with diploid populations consisting of  $N/2$  individuals. In a diploid population, we assume that diploid fitness is determined by the product of the fitness effects of each component allele ( $W_{BB} = (1 + s)^2$ ,  $W_{Bb} = (1 + s)$ ,  $W_{bb} = 1$ ). We expect that dominance should not matter too much, because most of the hitchhiking effect and fixation probability are determined early, before  $BB$  arises, but this reasoning may break down at very low recombination rates, we will do simulations to check how much it affects the predictions at a later point.

As noted below, the smaller  $p_0$ , the less genetic swamping is expected in the resident population, which makes it a relevant parameter to explore. From a conservation standpoint, introducing fewer individuals (smaller  $N_I$ ) or introducing them earlier (while the resident population  $N_R$  remains large) both reduce  $p_0$ .

We start by quantifying the fixation probability of  $B$  and then incorporate the degree of introgression in a later section.

### The probability of evolutionary rescue

We consider the probability that the adaptive allele establishes within the resident population, rescuing the population. Haldane (1927) investigated the probability of establishment of the rare advantageous allele  $B$  using the theory of branching processes. While he considered a constant population size, this framework is readily adapted to consider changing population sizes (as in Otto and Whitlock 1997). Let  $1 - P_n[N_I]$  be the probability that none of the  $N_I$  copies the adaptive allele  $B$ , present at generation  $n$ , will leave descendants in the next generation. The probability of  $B$  establishing in the population is then  $P_n[N_I]$ . If we assume a Poisson distribution for the number of offspring  $\lambda_n$  per individual bearing the adaptive allele  $B$  at generation  $n$ , then  $P_n[N_I]$  satisfies:

$$1 - P_n[N_I] = \sum_{j=0}^{\infty} e^{-N_I \lambda_n} \frac{N_I \lambda_n^j}{j!} (1 - P_{n+1}[1])^j, \quad (3.2)$$

where  $\lambda$  stands for the average number of offspring (Haldane 1927). We assume that the resident population size has declined to the point that it exerts little competition on other individuals. The population size is then expected to decline geometrically over time by a factor  $(1 + r)$  each generation among individuals bearing the non-adaptive allele  $b$  and to rise by a factor  $(1 + r)(1 + s)$  among individuals bearing the adaptive allele (see Otto and Whitlock 1995 for models allowing competition). For the evolutionary rescue scenario of interest,  $(1 + r) < 1$  while  $(1 + r)(1 + s) > 1$ .

As long as the adaptive allele establishes while competition remains negligible, the number of offspring per individual bearing the adaptive allele  $B$  remains constant at  $\lambda = (1 + r)(1 + s)$ , and the probability of establishment is the same from one generation to the next ( $P_n[N_I] = P_{n+1}[N_I] = P[N_I]$ ). The probability that the adaptive allele establishes when introduced in a single copy,  $P[1]$ , then equals the valid solution to:

$$P[1] = 1 - e^{-(1+r)(1+s)P[1]}, \quad (3.3)$$

where  $0 < P[1] < 1$  when the adaptive allele is capable of rescue ( $(1 + r)(1 + s) > 1$ ). If multiple individuals bearing  $B$  are initially introduced, then  $P[N_I] = 1 - e^{-N_I(1+r)(1+s)P[1]} = 1 - (1 - P[1])^{N_I}$ .

Equation 3.3 may be solved numerically for any value of  $r$  and  $s$ , but if we assume that the

number of offspring per  $B$  parent is near one ( $\lambda = (1+r)(1+s) \approx 1$ ) then  $P[1] \approx 1 - e^{-2(r+s+rs)} \approx 2(r+s+rs)$ . Introducing  $N_I$  individuals from the non-resident population then rescues the population with probability:

$$P[N_I] = 1 - e^{-2N_I(r+s+rs)}. \quad (3.4)$$

### Introgression

Here, we assume that the population sizes are large enough to ignore drift (we relax this assumption below). While the resident population is declining and will eventually go extinct if rescue does not occur in time, the frequency of each haplotype can still be described deterministically following the methods of Maynard Smith and Haigh (1974). Their Equation (7) gives the frequency of the non-resident haplotype ‘‘AB’’ among individuals bearing the adaptive allele  $B$  at each generation ( $Q_n$ ), which can be rearranged as the sum of terms

$$Q_n = R_0 \sum_{i=1}^n c(1-c)^{i-1} \frac{1-p_0}{1-p_0+p_0(1+s)^i}. \quad (3.5)$$

Similarly,  $1 - Q_n$  measures introgression of the non-resident allele  $a$  at time  $n$  among haplotypes bearing  $B$ . Within this deterministic framework, the frequency  $Q_n$  of the non-resident haplotype among individuals bearing the adaptive allele  $B$  can be explored over time, either assuming free recombination (see Supplementary Material 3 - S1) or allowing linkage (see Supplementary Material 3 - S2).

Of particular interest is the fraction of the resident genome ultimately preserved, long after the introduction event, given by  $Q_\infty$ . In the next section, we will use this quantity to determine the extent of introgression across the whole genome. To do so, however, we must approximate  $Q_\infty$ . Previous approximations have assumed large populations (Durrett and Schweinsberg 2004) and/or weak selection (Barton 2000). We consider an alternate approximation that allows for strong selection in populations of moderate size (see Supplementary Material - S3):

$$Q_\infty \approx R_0(1-p_0) \left( 1 - (1-c)^{\frac{\log\left(\frac{1-p_0}{p_0}\right)}{\log(1+s)}} \right). \quad (3.6)$$

Below, we assume that the resident population initially lacks the adaptive allele  $B$  ( $R_0 = 1$ ).

In Suppl. Tab. 3.1, we compare our approximation with the original expression of Maynard Smith and Haigh (1974), as well as the approximations of Barton (2000) and Durrett and

Schweinsberg (2004) (see Supplementary Material 3 - S4). We find that equation 3.6 slightly overestimates the final frequency of the rescue haplotype ( $Q_\infty$ ) but is more accurate, on average, across the range of parameters considered in Table 1 of Maynard Smith and Haigh (1974). Importantly, the approximation holds well even for stronger selection ( $s = 0.5$  considered in Suppl. Tab. 3.2), as might be experienced for a major effect allele capable of rescuing the resident population from extinction.

### Fraction of the resident genome preserved

Instead of considering a single neutral locus **A**, we now look at the impact of introgression of locus **B** on the whole genome, given a genetic map. We approximate the fraction of the resident genome preserved after the sweep is over ( $G$ ) using Haldane's mapping function, *i.e.*, for sites located  $\delta$  Morgans from locus **B** (Haldane 1919). Using  $c = \frac{1}{2}(1 - \exp^{-2\delta})$  and averaging  $Q_\infty$  across the genome, we have:

$$G = \frac{\int_{\delta=0}^M Q_\infty(\delta) d\delta}{M} \quad (3.7)$$

for a genetic map of length  $M$ . As detailed in the supporting *Mathematica* notebook, this expression can be simplified to

$$G = (1 - p_0) - \frac{1}{M}(1 - p_0) \left( 2^{-1-x} \left( \text{Beta}\left[\frac{e^{2M}}{1 + e^{2M}}, -x, 0\right] - \text{Beta}\left[\frac{1}{2}, -x, 0\right] \right) \right) \quad (3.8)$$

with  $x = \frac{\log\left(\frac{1-p_0}{p_0}\right)}{\log(1+s)}$  and  $\text{Beta}[a, b, c]$  is the incomplete beta function (see Abramowitz et al. 1988 p. 944). Note that for very long maps, *i.e.*,  $M \rightarrow \infty$ ,  $G$  approaches the following limit

$$G_{M \rightarrow \infty} = (1 - p_0) \left( 1 - 2^{-\frac{\log\left(\frac{1}{p_0} - 1\right)}{\log(1+s)}} \right). \quad (3.9)$$

While the above assumes a single chromosome, the average amount of introgression across a genome with multiple chromosomes is given by  $fG + (1 - f)G_{M \rightarrow \infty}$ , where  $f$  is the fraction of the genome on the chromosome bearing the adaptive allele  $B$  and  $M$  is now the length of that chromosome.



## Combining evolutionary rescue and genomic preservation

To determine the optimal balance between the probability of rescue and genomic preservation of the resident genome, we combine the above metrics. The expected weighted fraction of genome rescued ( $E[\mathcal{R}]$ ) is then

$$E[\mathcal{R}] = \hat{G}^v P[N_I], \quad (3.10)$$

where  $v$  is a weighting factor that determines how critical avoiding swamping is ( $v \gg 1$ : very critical,  $v \ll 1$ : not at all important). When  $v = 1$ ,  $E[\mathcal{R}]$  gives the the expected surviving portion of the resident population's genome that remains in the long-term.

## Individual-based simulations

We perform individual-based simulations using SLiM (Haller and Messer 2017). We model a declining population of diploid individuals where an individual possessing a beneficial mutation is introduced. To be able to compare with the haploid framework previously described, we use a diploid model of selection that matches haploid selection such that  $h = (-1 + \sqrt{1 + s})/s$ . Each generation, offsprings are stochastically drawn in a Poisson distribution according to population size and parent fitness. Each simulation stops once the beneficial mutation has fixed, and we compute the proportion of the genome at each loci that descends from the non-resident individual initially introduced. We run simulations for parameters detailed in 3.2 (but see Perspectives).

We use Haldane's mapping function to compute a recombination rate that matches the analytical framework. We consider a map with a length of  $M = 2$  Morgans and  $L = 100$  loci, which are distributed equidistantly along this map. The recombination rate between each pair of loci is then  $c = \frac{1}{2}(1 - \exp^{-2 \times (M/L)}) \approx 0.02$ . We repeat the simulations 400 times for each set of parameters.

We first use these individual based simulations simply to check the expectation of our expressions (see Supplementary material 3 - S5 and Perspectives). We average the fraction of the non-resident genome introgressed at each loci and compare it to equation (6) and across the whole genome and compare it to the prediction from our approximation, equation (8).

Parameter	Value(s)
Carrying capacity	7000
Resident population size	100, 1000
Relative selection coefficient	0.7
Number of runs	400

Tab. 3.2: Summary table of simulation parameters

## Results

### Degree of introgression

As illustrated in Fig. 3.1, a higher fraction of the resident genome is retained when proportionately fewer genomes are introduced (lower  $p_0$ ), when the species recombines at a higher rate across its genome (higher length of genetic map,  $M$ ), and when selection favouring the adaptive allele is weaker (lower  $s$ ). For species with large map lengths (large  $M$ ), the maximum amount of the resident genome preserved is expected to equal  $1 - p_0$  when selection is weak and  $1 - 2p_0$  when the adaptive mutation doubles fitness ( $s = 1$ ). Thus, the most important consideration for preserving local genetic diversity is to reduce  $p_0$  by maximizing the number of residents and minimizing the number of introduced individuals (compare panels for different  $p_0$  in Fig. 3.1). While waiting to introduce genetic variation is important for conservation managers to be confident that hybrid rescue is needed, waiting increases the extent of genetic swamping, simply because fewer resident genomes are present at the time of introduction.

As expected, stronger selection ( $s$ ) on the adaptive allele  $B$  leads to less preservation of the resident genome due to a faster sweep. By contrast, higher rates of recombination (longer genetic map lengths,  $M$ ) facilitate the maintenance of the resident genome, decreasing the extent of introgression (Fig. 3.1). Note that the fraction of the resident genome retained is independent of population dynamics ( $r$ ), and is only affected by the initial ratio of non-resident relative to the size of the resident population.

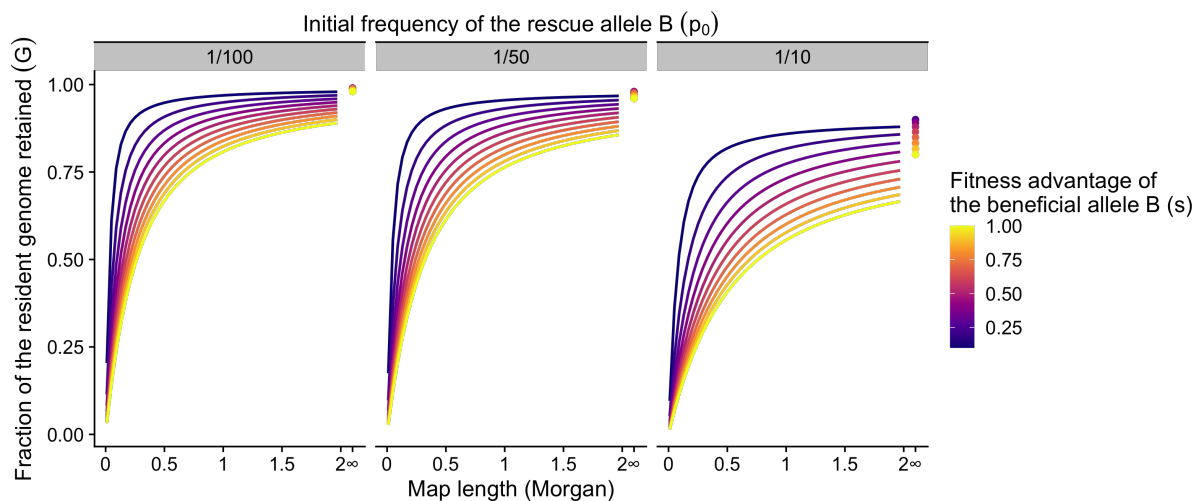


Fig. 3.1: The fraction of the resident genome retained after the sweep ( $G$ ) rises with the map length of the species ( $M$ ) but declines with the selective advantage of  $B$  (lighter shading) and its initial frequency  $p_0$  (from left to right).

### Balancing rescue and swamping

While minimizing introduction frequencies ( $p_0$ ) always reduces swamping, introducing too few non-resident individuals risks losing the adaptive allele while it remains rare. Thus, there is a trade-off in the optimal number of non-resident to introduce into the resident population to maximize the probability of rescue and the preservation of the resident genome (Equation 3.10). This trade-off is explored in Fig. 3.2), assuming that the goal is to maximize the surviving portion of the focal species' genome ( $v = 1$ ), accounting for the possibility of extinction (nothing remains) and persistence ( $\hat{G}$  remains). The larger the population size of the resident population at the time of introduction, the higher the optimal number of introduced individuals (from yellow to purple vertical lines), because swamping is expected to be weaker and rescue is the most important factor determining the expected amount of the resident genome that will remain in the population.

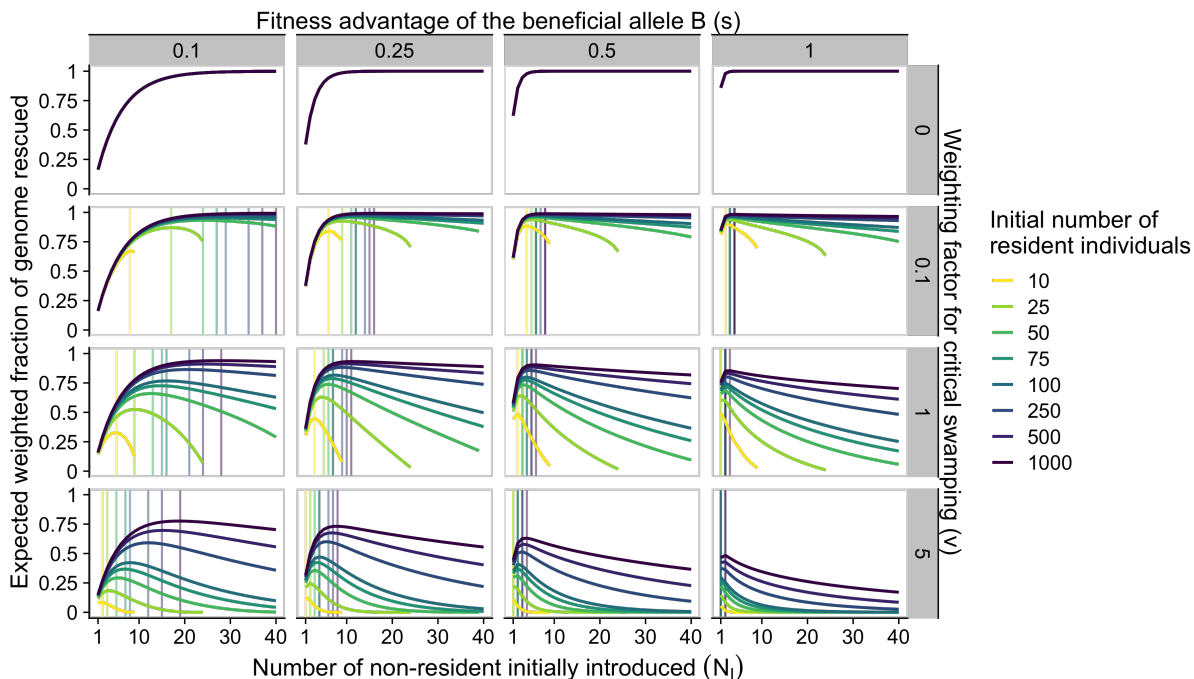


Fig. 3.2: The optimal number of non-resident individuals to introduce to maximize the probability of rescue (vertical lines) depends mainly on the size of the resident population (colored curves, affecting  $p_0$ ) and the fitness advantage of the adaptive allele ( $s$ , columns). The rows consider different conservation objectives, from considering the fixation probability alone ( $v = 0$ ), rescue mattering ten times more than swamping ( $v = 0.1$ ), both being of equal importance ( $v = 1$ ), or swamping being five times more important ( $v = 5$ ). Other parameters:  $r = -0.01$  (decline of the resident population), map length  $M = 1$ .

The optimal number of individuals to introduce also depends greatly on the fitness advantage of the rescue allele  $B$  ( $s$ ), the stronger the advantage the smaller the optimal introduction size is (from left to right panels, Fig. 3.2 and Fig. 3.3), both because genetic swamping is more severe and the probability of rescue is higher when selection is strong. In some cases, it may be preferable to ensure survival of the population, even at the risk of more genetic swamping ( $v < 1$ ), for example, when the population plays an important ecological role. In this case ( $v = 0.1$ ; second-row panel), the best compromise is to introduce slightly more non-resident to ensure maximal rescue. In other cases, it may be preferable to avoid swamping of the resident genome as much as possible, as when the distinctive traits of the resident population are culturally valuable. This can be investigated by increasing  $v$  above 1, which places more weight on keeping the resident genome intact and thereby overall decreases the optimal number of introduced copies of the  $B$  allele. Higher rates of decline in the resident population ( $r$  decreases) increase the optimal number of non-resident individuals to introduce to maximize the probability of rescue (Fig. 3.3), but lower the maximum fraction of the resident genome that can be saved (see

Supp. Fig 3.7 and 3.8 in Supplementary material 3 - S6).

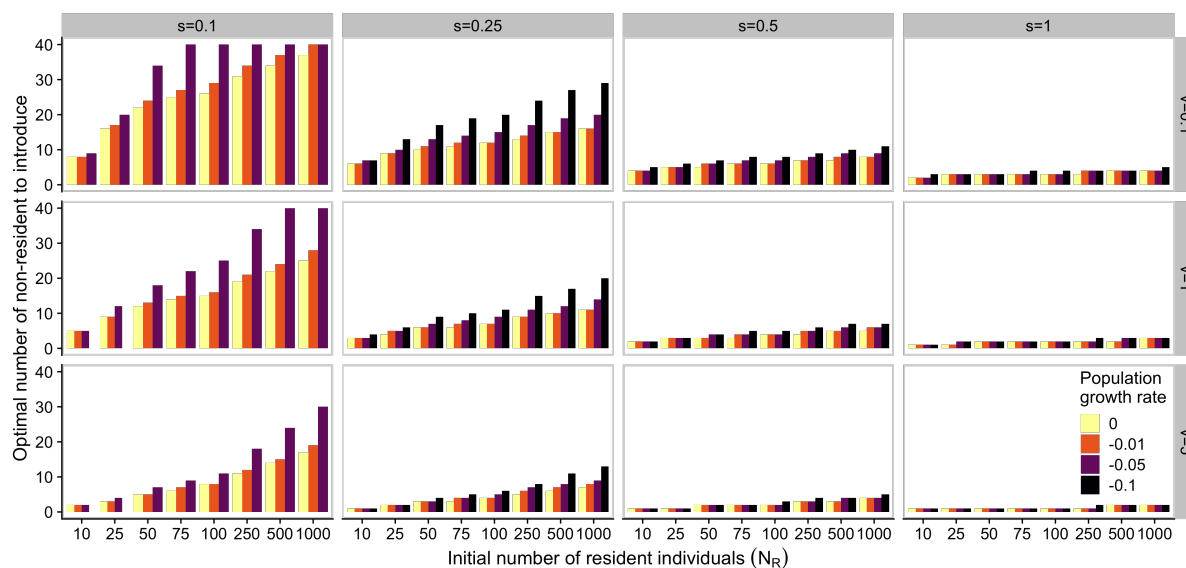


Fig. 3.3: Higher rates of decline in the resident population (colors) increase the optimal number of non-resident individuals to introduce to maximize the probability of rescue, particularly the fitness advantage of the adaptive allele is low ( $s$ , columns) and when rescue matters ten times more than swamping ( $v = 0.1$ ). There is no solution to equation (4) when  $r = -0.1$  and  $s = 0.1$  (darker color). Other parameters: map length  $M = 1$ .

## Discussion

Genetic swamping poses a risk for conservation management approaches that involve introgressive hybridization. To provide more insight into the balance between gene swamping and evolutionary rescue of endangered populations, we built a model and determined the optimal number of introduced individuals given the conservation value of preserving ( $v > 1$ ) or rescuing ( $v < 1$ ) the local population ( $v$ ). Genetic swamping is here defined as the extent of the non-resident genome that makes it into the resident,  $1 - \hat{G}$ , while rescue is defined as the probability that the rescue allele ( $B$ ) fixes in the resident population,  $P[N_I]$ , allowing the population to grow.

In this paper, we highlight critical parameters that determine the surviving portion of a focal population's, providing guidance for conservation managers considering this type of practice. As it is the case for other conservation strategies, such as conservation breeding programs, the size of the resident population to be saved is one the most critical parameter to assess the risks of such practice (Rabier et al. 2022, Appendix C). A larger resident population helps slow down hitchhiking of linked alleles along with the adaptive allele, encouraging planning them as early

as possible.

The model we present here investigates the long-term dynamics of hybridization and is simplistic in many aspects. For example, our model assumes uniform recombination across the genome. It is well known, however, that recombination rates vary greatly across the genome (Stapley et al. 2017). Adaptive alleles that occur within regions of lower (higher) recombination will lead to more (less) genetic swamping than we have calculated, although these effects should be localized within that region of the genome. Indeed, recent work has shown that patterns of introgression vary across the genome (Baack and Rieseberg 2007). Large genome-scale analyses will be required to understand whether rampant genetic exchange is the rule for hybridizing species or whether much of the genome is resistant to introgression.

As expected, there is a trade-off between genetic swamping and evolutionary rescue, which can be optimized by introducing an intermediate number of hybrids. Our results are in line with those of Kelly and Phillips 2019a, which showed that the timing (and therefore the size of the resident population) and the size of the introduction were the key parameters determining introgression success. Our analytical model also allows determination of that optimal number of invaders that would lead to the highest expected evolutionary rescue. Whether the recent theoretical results can be extended to inform when natural populations will undergo evolutionary rescue if faced with novel environmental conditions remains to be seen. To facilitate comparing these theoretical predictions with empirical data, we provide a simulation framework in SLiM which allows users to change parameters of the model and visualize the output instantly (see Perspectives).

## Perspectives

Before we consider submitting this work, further modifications and analyses remain to be carried out.

1. The deterministic calculations of Maynard Smith and Haigh (1974) are known to underestimate the effect of hitchhiking because the frequency of  $B$  rises faster among those cases that fix. Future work will involve changing  $p_0$  to  $p_0/(2s)$  as mentioned by Barton (2000).
2. We will add deleterious mutations, similarly to Hartfield and Otto 2011 and Kelly and Phillips 2019a, as they can decrease the fitness of individuals with too much swamping, and facilitate the persistence of recombinant lineages. We will also explore adding barriers

to gene flow with incompatibilities (BDMIs) into the individual-based simulations.

3. Ultimately, we would like to develop a Shiny app to allow users to enter parameters for a species of interest. Alternatively, we could explore writing an R wrapper using `slimr` (Dinnage et al. 2023) to implement our simulation framework directly onto R.

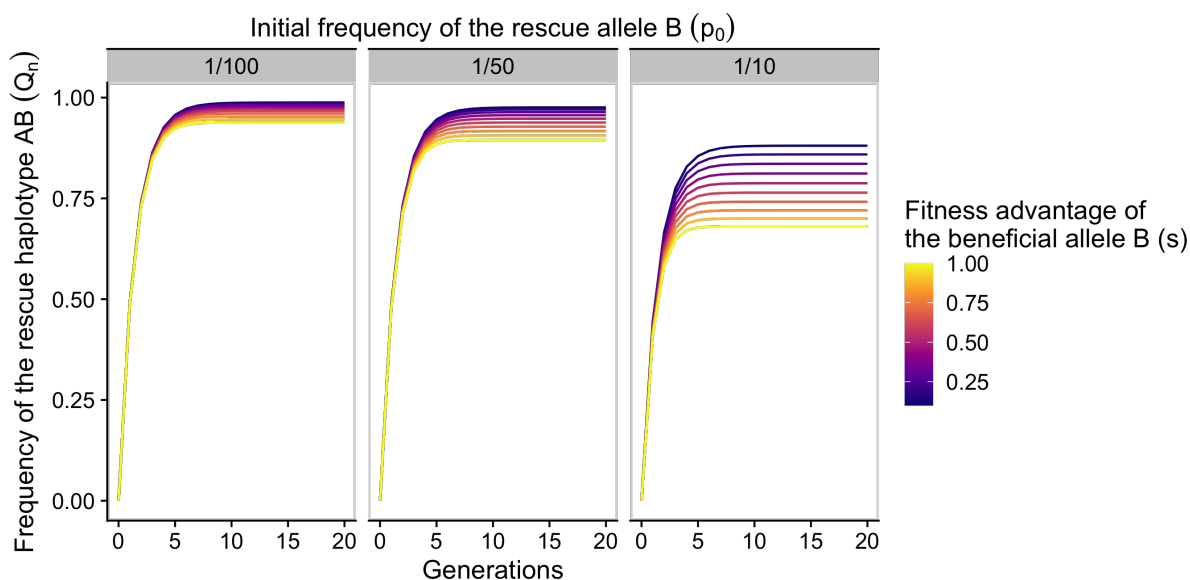
## Supplementary material 3

### S1. Free recombination

Consider a single non-resident individual is introduced into the resident population, at generation  $n = 0$ . In the notation of Maynard Smith and Haigh 1974, this means that  $R_0 = 1$  and  $Q_0 = 0$ , with  $p_0 = \frac{1}{N_R}$ . Assuming free recombination ( $c = 0.5$ ), the amount of rescue each generation can be expressed as

$$Q_n = \sum_{i=1}^n \frac{2^{-i} (1 - p_0)}{1 - p_0 + p_0(1 + s)^i}. \quad (\text{S1})$$

We explore the expected amount of rescue ( $Q_n$ ) after 20 generations for relatively small resident populations of 10, 50 or 100 individuals (Suppl. Fig. 3.1). We find that for the rescue haplotype to predominate in the population, the initial introduction must be at a low enough frequency relative to the size of the resident population. For a population of 10 individuals, the maximum frequency of the *rescue* haplotype barely exceeds 0.75. For populations of size 50 or 100, however, the final frequency of the rescue individuals is close to 1, if the fitness advantage of the beneficial allele is not too important. Otherwise, the proportion is slightly lower, notably for populations of 50 resident individuals, and falls to around 0.6 for populations of 10 resident individuals.

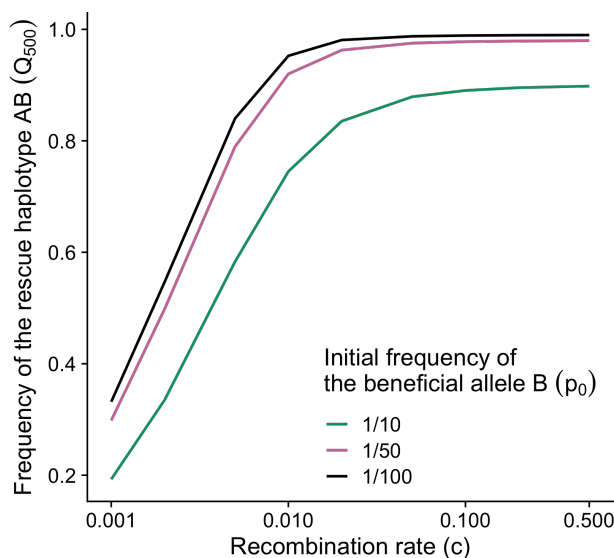


Suppl. Fig. 3.1: The rescue haplotype predominates in the population, at the end, if the initial introduction is at low enough frequency relative to the resident population size and the sweep not too fast (fitness advantage of the B allele not too high).



## S2. Allowing linkage

We next relax the assumption that the loci are unlinked and evaluate the fraction of the rescue haplotype AB that ultimately fixes. Due to imprecision when recombination rates are low, we use  $Q_{500}$  generations rather than  $Q_{\infty}$ .



Suppl. Fig. 3.2: Recombination increases the frequency of the rescue haplotype retained after 500 generations similarly with both initial frequency at introduction.

## S3 - Approximation for the extent of hitchhiking

The summation in Equation 3.5 has the form of a geometric probability distribution  $(c(1-c)^{i-1})$ , weighted by the frequency of the non-resident allele  $b$  at time  $i$ . Thus, the summation describes the waiting time until a successful recombination event moves the selected allele  $B$  from a haplotype carrying the non-resident allele  $a$  to one carrying the resident allele  $A$ , given that the probability that the homologous chromosome carries  $b$  declines over time according to  $1 - p_n$  (Equation 3.1). The  $R_0$  multiplying the summation in Equation 3.5 represents the probability that this homologous chromosome also carries  $A$ .

Because the adaptive allele frequency is expected to change logistically over time, the weighting is near one initially when  $1 - p_0 \approx 1$  and near zero once allele  $b$  becomes rare. The approximation we pursue replaces the logistic weighting term (given by the the fraction in Equation 3.5) with a rectangular function, starting with a height of  $1 - p_0$  and declining to zero at the mid-point of the logistic. The midpoint is found by setting  $p_n$  in Equation 3.1 to  $1/2$  and solving for the generation at which this occurs, giving  $n_{mid} = \log((1 - p_0)/p_0)/\log(1 + s)$ . The

cumulative distribution of the geometric distribution up until time  $n_{mid}$  gives Equation 3.6.

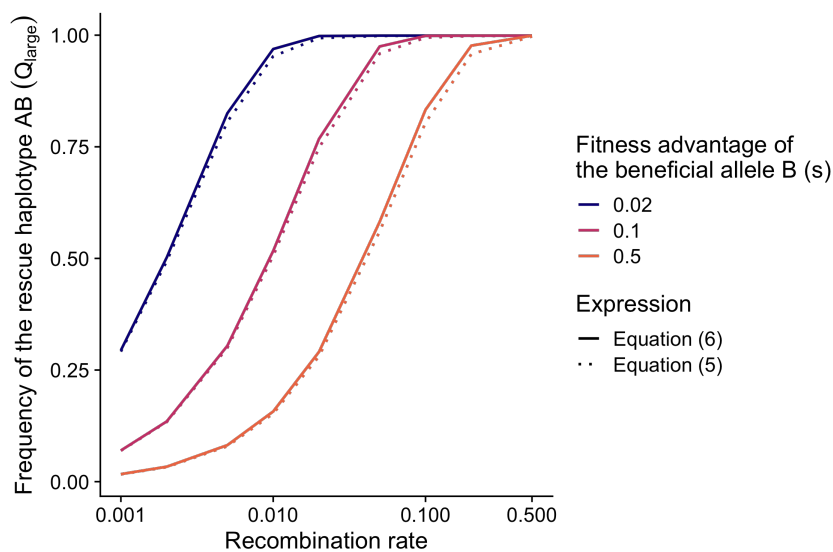
Note that for tight recombination (*i.e.*,  $c \rightarrow 0$ ), this approximation goes to

$$Q_{max} \approx \frac{c(1-p_0) \log\left(\frac{1-p_0}{p_0}\right)}{\log(1+s)}. \quad (\text{S1})$$

This approximation is expected to be most precise when selection is strong  $s \gg 0$ , so that the logistic change in allele frequency has a steep slope at the mid-point. As shown in Suppl. Tab. 3.1 and 3.2, however, the approximation works well across the range of parameters considered (see Supplementary Material S4 for calculations of these tables and illustrative figures).

## S4. Comparisons

### Comparison with the exact sum for $Q_n$ from Maynard Smith and Haigh 1974

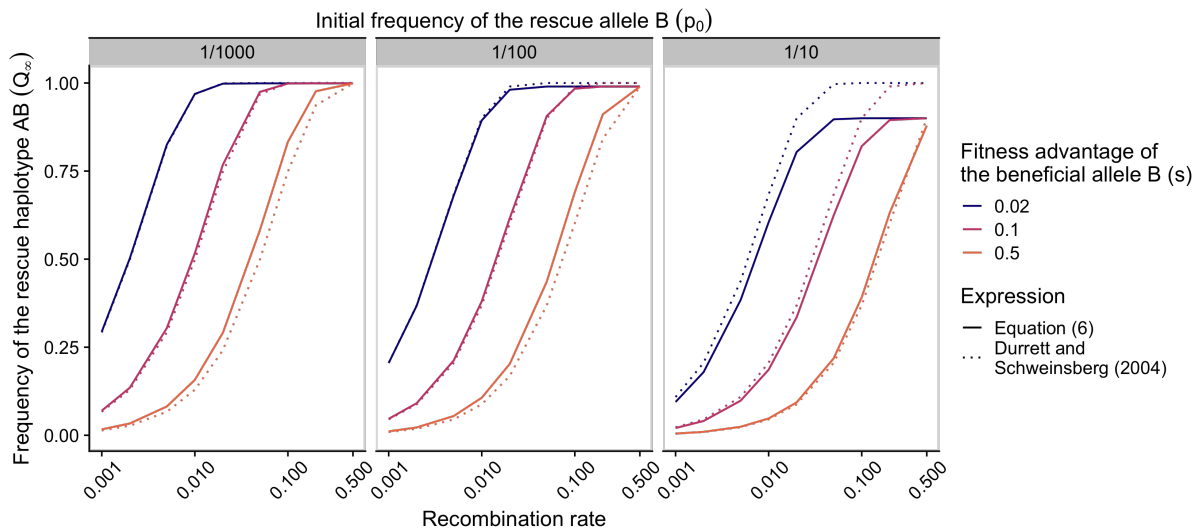


Suppl. Fig. 3.3: The approximation (equation (6); plain line) slightly overestimates the amount of rescue after a large number of generations against the exact value for  $Q_n$  (equation (5); dotted line) but performs well overall whether the fitness advantage is  $s = 0.02$  (in blue) after 700 generations,  $s = 0.1$  (in pink) after 150 generations, or  $s = 0.5$  after 50 generations.

### Comparison with previous approximations

#### Comparing to Durrett and Schweinsberg 2004

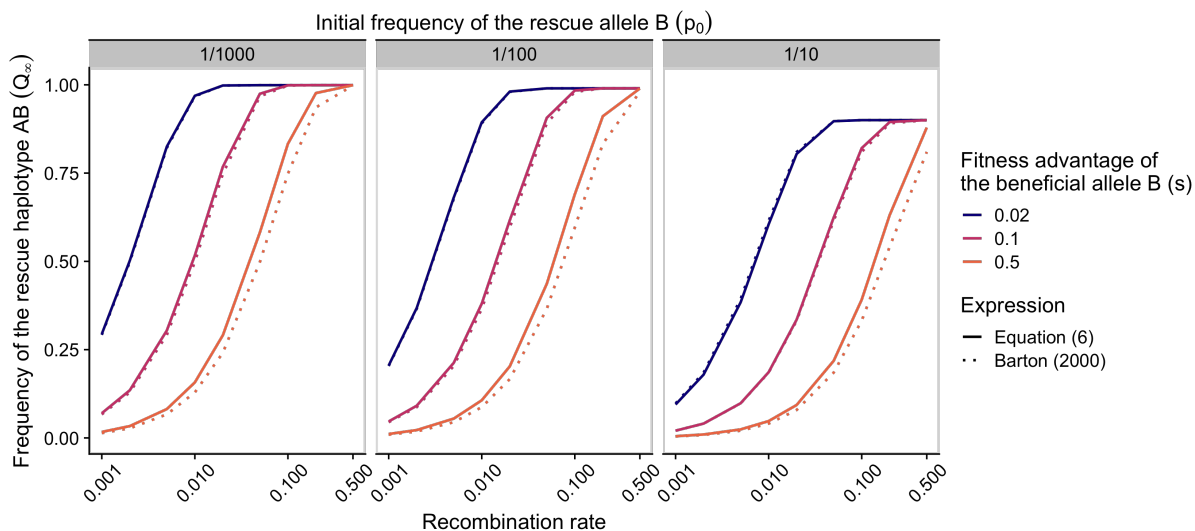
Durrett and Schweinsberg 2004 have an approximation when the population size is very large. We find the approximations to diverge for strong fitness advantage (in orange) or for higher starting frequencies ( $p_0 = 1/100$  or  $p_0 = 1/10$ ).



Suppl. Fig. 3.4: The approximation for  $Q_n$  (equation (6); plain line) works well for a lower initial frequency of the rescue allele B but underestimates the amount of rescue after a large number of generations compared to the approximation by Durrett and Schweinsberg 2004 (dotted line) especially for higher starting frequencies ( $p_0 = 1/100$  or  $p_0 = 1/10$ ) or a strong fitness advantage  $s = 0.5$  (orange).

### Comparing to Barton 2000

Barton 2000 approximates discrete time recursions with continuous time functions, assuming that  $(1 + s) \approx \exp(s)$  and  $(1 - r) \approx \exp(-r)$ . His approximation works well when selection is weak but breaks down with strong selection (when  $s$  is 0.5 or larger).



Suppl. Fig. 3.5: The approximation for  $Q_n$  (equation (6); plain line) works well when the fitness advantage of the rescue allele B is weak but breaks down with a stronger advantage (when  $s$  is 0.5, in orange) compared to the approximation by Barton 2000 (dotted line).

**Comparing all existing approximations to Maynard Smith and Haigh 1974**

Starting with  $p_0 = 1/N$  copies of the allele, we compare the approximation for  $Q_n$  (equation 3) to the approximation  $\frac{c}{s} \log(N)$  from Maynard Smith and Haigh 1974, to Durrett and Schweinsberg 2004 and Barton 2000.

We find when  $N$  is large,  $s$  small, and  $c$  small, there are substantial discrepancies unless enough time has passed but after 5000 generations we are able to recreate the Maynard Smith and Haigh 1974 table.

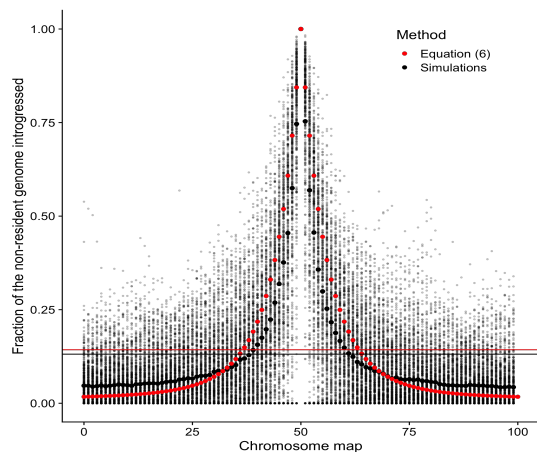
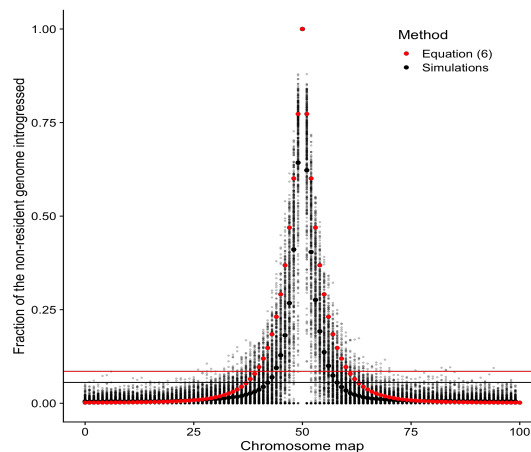
$N$	$s$	$c$	$Q_{5000}$	MS&H (1974)	D&S (2004)	Barton (2000)	Equation 3.6
100	0.01	0.00003	0.0137599	0.4%	-0.3%	-1.3%	-1.0%
100	0.1	0.0003	0.014231	-2.9	-3.6	-4.6	-0.1
1000000	0.01	0.00001	0.013782	0.2	-0.4	-0.4	0.0
1000000	0.1	0.0001	0.0143404	-3.7	-4.3	-4.3	0.4
100	0.01	0.00006	0.0273027	1.2	-0.2	-1.2	-0.9
100	0.1	0.0006	0.0282332	-2.1	-3.5	-4.4	0.0
1000000	0.01	0.00002	0.027371	0.9	-0.4	-0.4	0.1
1000000	0.1	0.0002	0.0284731	-3.0	-4.3	-4.3	0.4
100	0.01	0.00015	0.0666666	3.6	0.1	-0.9	-0.6
100	0.1	0.0015	0.068906	0.2	-3.1	-4.1	0.3
1000000	0.01	0.00005	0.067007	3.1	-0.4	-0.4	0.1
1000000	0.1	0.0005	0.0696546	-0.8	-4.2	-4.2	0.4
100	0.01	0.0003	0.128271	7.7	0.6	-0.4	-0.1
100	0.1	0.003	0.132476	4.3	-2.6	-3.6	0.8
1000000	0.01	0.0001	0.129455	6.7	-0.3	-0.3	0.1
1000000	0.1	0.001	0.13441	2.8	-4.0	-4.0	0.4
100	0.01	0.0006	0.237936	16.1	1.5	0.5	0.7
100	0.1	0.006	0.245355	12.6	-1.6	-2.6	1.6
1000000	0.01	0.0002	0.241909	14.2	-0.2	-0.2	0.2
1000000	0.1	0.002	0.25059	10.3	-3.7	-3.7	0.5
100	0.01	0.0015	0.482226	43.2	3.4	2.4	2.7
100	0.1	0.015	0.495058	39.5	0.8	-0.2	3.5
1000000	0.01	0.0005	0.498418	38.6	0.1	0.1	0.4
1000000	0.1	0.005	0.513015	34.7	-2.8	-2.8	0.7
<b>Mean error</b>				<b>9.5%</b>	<b>-1.4%</b>	<b>-1.9%</b>	<b>0.4%</b>
<b>Mean absolute error</b>				<b>10.5%</b>	<b>1.9%</b>	<b>2.1%</b>	<b>0.7%</b>

Suppl. Tab. 3.1: Comparing the approximations for  $Q_n$  across the parameters considered by Maynard Smith and Haigh (1974): the number of haplotypes ( $N$ ), the selection coefficient ( $s$ ), and the crossover frequency ( $c$ ). Starting with a frequency of  $p_0 = 1/N$  of the adaptive allele  $B$ , Equation (3.5) for  $Q_n$  was evaluated at generation 5000, which was sufficient time to see no further change and to recreate their Table 1 (4th column). The % error from  $Q_{5000}$  is reported in the following columns, with the mean % error and mean absolute % error reported at the bottom of the table. Approximations:  $c/s \log(1/p_0)$  from Maynard Smith and Haigh,  $1 - \exp(-c \log(1/p_0)/s)$  from Durrett and Schweinsberg (2004),  $(1 - p_0)(1 - p_0^{c/s})$  from Barton (2000), and Equation 3.6.

$N$	$s$	$c$	$Q_{5000}$	MS&H (1974)	D&S (2004)	Barton (2000)	Equation 3.6
100	0.5	0.0015	0.0161532	-14.5%	-15.1%	-15.9%	3.4%
1000000	0.5	0.0005	0.0166482	-17.0	-17.6	-17.6	1.5
100	0.5	0.003	0.0320271	-13.7	-14.9	-15.8	3.5
1000000	0.5	0.001	0.0330224	-16.3	-17.5	-17.5	1.5
100	0.5	0.0075	0.0780226	-11.5	-14.5	-15.3	3.8
1000000	0.5	0.0025	0.080546	-14.2	-17.1	-17.1	1.5
100	0.5	0.015	0.149547	-7.6	-13.7	-14.6	4.2
1000000	0.5	0.005	0.154676	-10.7	-16.6	-16.6	1.5
100	0.5	0.03	0.275298	0.4	-12.3	-13.2	5.0
1000000	0.5	0.01	0.285668	-3.3	-15.5	-15.5	1.5
100	0.5	0.075	0.54576	26.6	-8.6	-9.5	6.4
1000000	0.5	0.025	0.56982	21.2	-12.5	-12.5	1.4
<b>Mean error</b>				<b>-5.1%</b>	<b>-14.7%</b>	<b>-15.1%</b>	<b>2.9%</b>
<b>Mean absolute error</b>				<b>13.1%</b>	<b>14.7%</b>	<b>15.1%</b>	<b>2.9%</b>

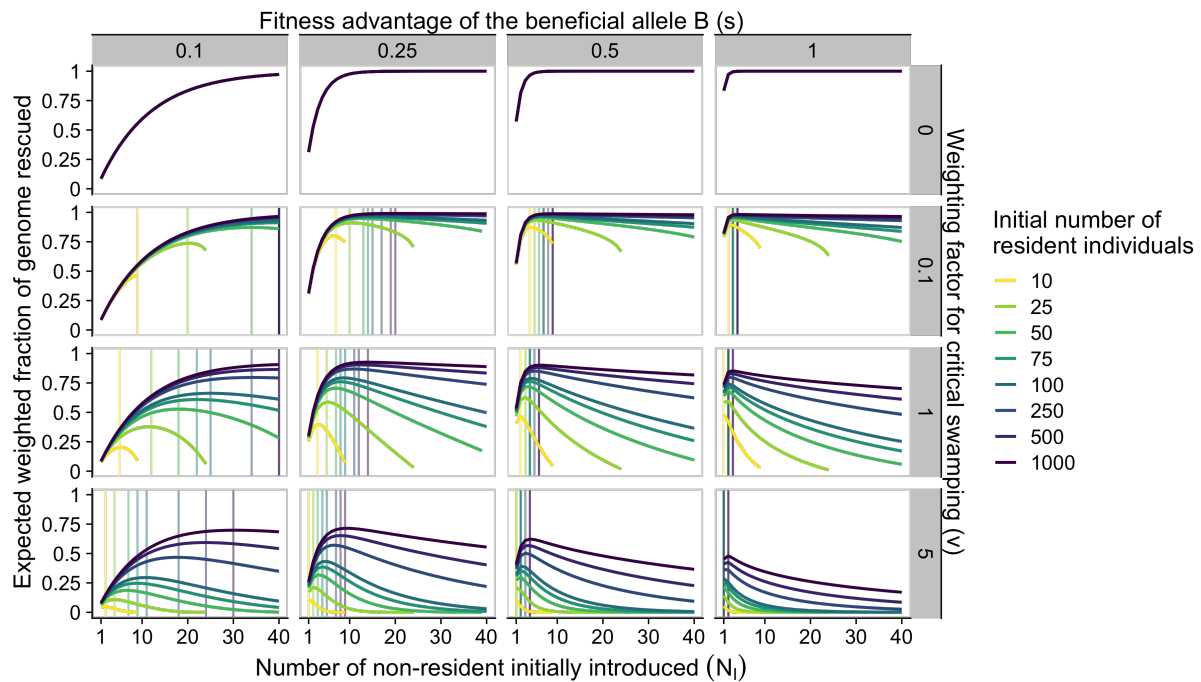
Suppl. Tab. 3.2: Comparing the approximations for  $Q_n$  with stronger selection. Equivalent to Table 3.1 but with  $s = 0.5$ .

## S5. Individual-based simulations

(a) Parameters:  $p_0 = 1/100$ ,  $s = 0.7$ .(b) Parameters:  $p_0 = 1/1000$ ,  $s = 0.7$ .

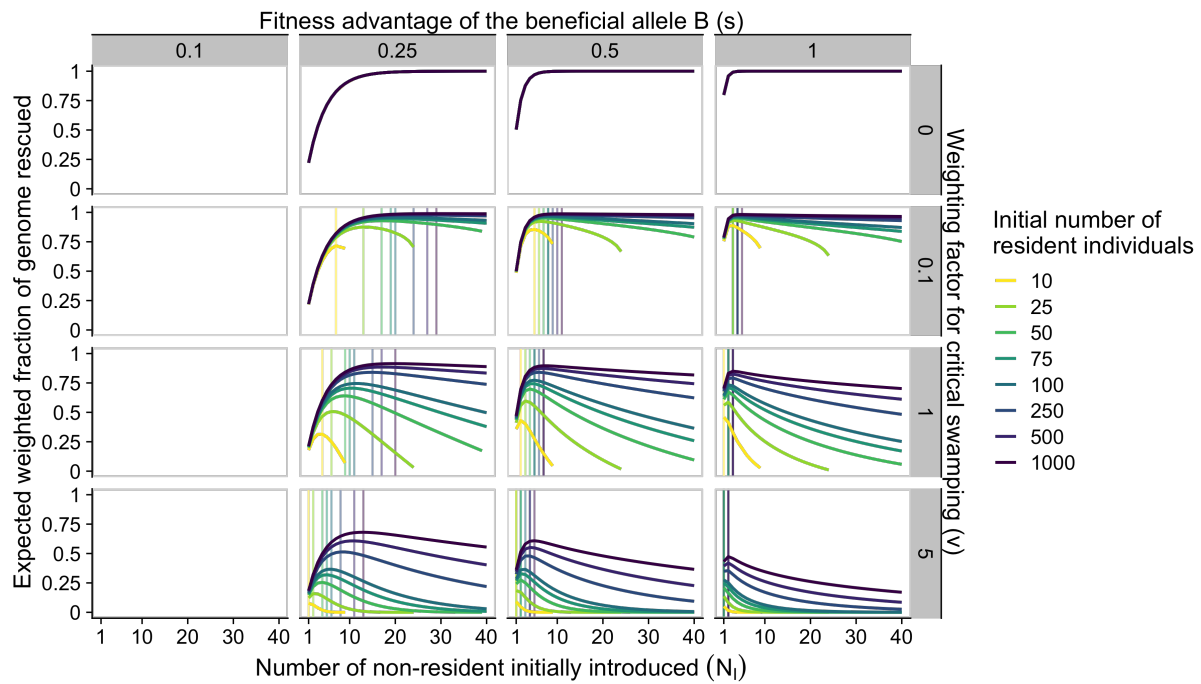
Suppl. Fig. 3.6: The fraction of non-resident genome that is introgressed increases toward the locus where the beneficial allele is introgressed. Equation (6) slightly overestimates the average fraction of non-resident genome introgressed at each locus when compared to the simulation results, especially in regions close to the introgressed allele. The total fraction of genome swamped is also slightly underestimated by equation (8) (14%) compared to the simulations (13%), although this difference is more important for a smaller initial resident population with equation (8) (8%) compared to the simulations (5%).

## S6. Higher rates of decline of the resident population



Suppl. Fig. 3.7: The optimal number of non-resident individuals to introduce to maximize the probability of rescue (vertical lines) increases with a higher rate of decline ( $r = -0.05$ ) and the maximum portion of the resident genome that can be saved decreases. Other parameter: map length  $M = 1$ .





Suppl. Fig. 3.8: The optimal number of non-resident individuals to introduce to maximize the probability of rescue (vertical lines) increases with a higher rate of decline ( $r = -0.1$ ) and the maximum portion of the resident genome that can be saved decreases. There is no solution to equation (4) when  $r = -0.1$  and  $s = 0.1$ . Other parameter: map length  $M = 1$ .

# General discussion

The work presented in this thesis focuses on assisted gene flow practices and their relevance as a conservation strategy.

## Relevance of assisted gene flow practices

In Chapter 1, we used a stage-structured framework, where individuals of each stage present different tolerances to a climatic variable (using temperature as an example). We show how assisted gene flow can effectively mitigate the effects of climate change by reducing the expected loss of individuals after climate change when planting trees adapted to warmer temperatures. We find, however, that the expected benefits of implementing assisted gene flow strategies are sometimes limited. We also highlight a high degree of uncertainty surrounding the outcome of these practices, mainly because an important parameter is poorly understood: the change in climatic tolerance of trees across life stages. The outcome of these practices was also strongly influenced by the stochasticity of the climate. We conducted a preliminary review of the literature looking for quantitative estimates of variation in thermal tolerance in trees, which revealed that few studies had a clear answer and none had quantitative measures of this parameter (Master's project of Freya Eriksen, 2021). We attempted to estimate the thermal tolerance of trees at different life stages by measuring climatic response curves such as reported in provenances tests and transplantation in many tree species (see Appendix A), which proved difficult due to (i) the lack of data on the effect of temperature on survival, although there are more data on the effect on growth, and (ii) the lack of data for the same species at different ages.

In Chapter 2, we developed a new framework for integral projection models to incorporate evolutionary dynamics, which is a step forward in overcoming some of the problems associated with existing models (see Appendix B). We use this model to investigate the possibility that an increase in genetic variance for growth in a declining population, as a result of assisted gene flow,

could stimulate evolutionary rescue. Our first, simple scenario assumes that a single cohort of the rare and endangered species *C. corymbosa* is planted, but the aim is to model more complex scenarios of reinforcements at a later stage. In this particular case, the species seems to have little chance of escaping extinction, even with an increase in genetic variation for individual growth rate.

In Chapter 3, we presented an analytical framework that introduces a new approximation for equation (7) of Maynard Smith and Haigh (1974), which gives the fraction of the genome that is conserved in a population where a beneficial allele is introduced (Maynard Smith and Haigh 1974). We show how this approximation produces more accurate results than previous approximations. We found that managed gene flow between genetically distinct populations or closely related species can lead to the introgression of beneficial alleles while still preserving a significant portion of the focal species' genome, although highly dependent on the number of individuals introduced relative to the size of the focal population. We draw attention to the fact that this type of managed gene flow, often considered a last resort out of fear and cost, carries greater risks when undertaken when natural populations are already in a very poor state. This echoes the message of a study I contributed to writing and published during my PhD (Rabier et al. 2022; Appendix C). Using simulations, we found that *ex situ* conservation breeding programs aimed at providing individuals for reinforcements are often implemented at a late stage, when populations have already suffered significant losses (notably due to the costs associated with these programs), and that this has a major impact on the inbreeding and residual genetic diversity of the founders of these programs, which in turn has a major impact on the natural population reinforced, for a large number of generations.

## The importance of models

We have used different types of theoretical models (demographic, quantitative genetics, population genetics), which are sometimes simplistic in the assumptions they make, but which can nevertheless allow us to make some predictions for species of interest for which parameters are known. Moreover, even models that do not explicitly make precise predictions are of interest for planning, as they can help to identify parameters and sources of uncertainty in the implementation of assisted gene flow practices. For example, the model developed in Chapter 1 identifies a trade-off that may arise during the lifetime of long-lived organisms such as trees when making

predictions about optimal provenances for planting and shows that uncertainties about change in climatic tolerance in the life of trees is critical to the implementation of assisted gene flow. There are many ways to improve this model and make its predictions more realistic. First, we could add costs for non-local provenances to incorporate the advantages that local provenance might have over other provenances by being adapted to aspects of the environment other than climate. Second, we use a stage-structured model, but an integral projection model may be more appropriate given that trees are slow-growing species that do not have very distinct biological stages. At the beginning of my PhD, we looked for published IPM models of temperate forest trees that could allow us to parameterize an IPM, but the one we found used forest inventory data (Kunstler et al. 2021) and started with individuals of size  $>10$  cm, *i.e.*, individuals that had already lived several years, which would have meant missing the important early years for climate tolerance.

In Chapter 1, we consider a single age- and stage-structured cohort growing in a warming climate with no regeneration. Although we do not model climate warming in Chapter 2, we complexify the previous framework by using an age- and size-structured IPM model that accounts for both regeneration and evolving growth trajectories.

In Chapters 1 and 2, we were interested in the potential benefits of gene flow. Although we modeled some costs in Chapter 2, with genetic load leading populations to extinction if there were too much genetic variance on the trait, in Chapter 3 we wanted to investigate the direct costs assisted gene flow can have on targeted populations, focusing on the swamping risks of a portion of the genome following an introgression event. We were able to generalize conclusions from previous models that used simulations (Harris et al. 2019; Kelly and Phillips 2019a; Smart and Phillips 2023) and allow exploration beyond specific parameters (Chapter 3).

All models (Chapters 1, 2, and 3) help us to quantify thresholds at which excessive gene flow could lead to deleterious consequences, such as introducing a genetic load (Lande and Shannon 1996), *i.e.*, fewer well-adapted individuals than poorly-adapted individuals will result in fewer individuals remaining in the cohort (Chapter 1), slow evolution or make rescue more difficult or impossible (Chapter 2), or lead to genomic swamping of the target population (Chapter 3).

## Many uncertainties still surround these practices

Biodiversity management and conservation practices are already integrating climate change into their transfer rules (*e.g.*, O'Neill et al. 2017), although most translocations are carried out in response to existing climate change impacts rather than in anticipation of its effects (Diallo et al. 2021). Many uncertainties remain regarding the outcomes and impacts of assisted gene flow practices.

We find that there is a first layer of uncertainty about the outcomes of assisted gene flow practices, which is related to the lack of knowledge of certain parameters, such as the change in climatic tolerance over the lifespan (Chapter 1), the selection coefficient of the beneficial allele, the size of the genome map (Chapter 3), genetic variance and covariance of adaptive traits (Chapter 2). Then there is a second layer of uncertainty related to the intrinsic stochasticity of the processes we study, with (i) the stochasticity of environmental change and (ii) the stochasticity of evolutionary processes *e.g.*, the beneficial allele may be randomly lost. We find that this uncertainty is amplified by trade-offs in the life cycle (Chapter 1), but can be reduced in some extent by diversity, as exemplified by the planting of diversified cohorts in Chapter 1.

Although our results suggest that substantial benefits are possible from the implementation of assisted gene flow, these benefits are sometimes very small (Chapter 1). Empirical studies show that increasing gene flow in small inbred populations can stimulate genetic rescue (Johnson et al. 2010; Madsen et al. 2004; Vila et al. 2003) but that some negative effects of outbreeding can appear after a few generations (*e.g.*, Hedrick et al. 2014). This supports the need for long-term studies to evaluate the success of these practices. Retrospective evidence that assisted gene flow can be an effective mean to accelerate population adaptation to climate change is, however, scarce (Breed et al. 2018; Young et al. 2020). Tools like the TransLoc database are, therefore, valuable to keep track of animal and plant translocations (Colas and Sarrazin 2017).

Our models in Chapters 1 and 2 suppose that we have identified source populations or genotypes that have high fitness in future climates. In reality, however, we lack information about the best sources of genetic material for AGF. Similarity between historical and future climates is often proposed as a selection criterion, but this strategy does not work if climate change has already decoupled populations from the climate to which they were historically adapted (Browne et al. 2019). Genomic offset is a method for predicting adaptation from genomic data, but information on population dynamics and fitness to validate this new approach is still very

scarce (Chen et al. 2022; Rellstab et al. 2021). In addition, for endemic species with a very small range and very small populations, there may not be enough variance to reach the critical variance that would save the population.

We only modeled a changing climate in Chapter 1, while in Chapters 2 and 3 we considered adaptation to an already deteriorated but constant environment, as typical in some models of evolutionary rescue (*e.g.*, Orr and Unckless 2014). Overall, however, recommendations for optimal assisted gene flow strategies are likely to be affected by climate change. First, it may affect the timing and number of individuals to be introduced. Favoring earlier introductions while populations are still large (Smart and Phillips 2023), and fewer migrants, which may not be possible, if the optimal number is already to introduce only one individual. A trade-off could also arise if climate changes during the lifetime of individuals, as the alleles that confer good adaptation may vary between the beginning and end of an individual's lifetime.

Questions also remain about the optimal frequency of migration events. On the one hand, to maintain the long-term viability of populations, it may be necessary to repeat gene flow on a regular basis to prevent inbreeding from increasing and stimulate population growth (Åkesson et al. 2022; Harrison et al. 2016). On the other hand, too much gene flow can be deleterious (genomic swamping, genetic load), so this balance must be taken into account when planning these strategies. In Chapter 1, we simulate the persistence of a single cohort of trees of the same species, with no local regeneration. In reality, regeneration from local trees does occur spontaneously, even in the case of clear-cut plantations such as those found in Canada (Harvey and Bergeron 1989).

As we have seen, assisted gene flow can have beneficial effects (evolutionary rescue, adaptation to climate change) as well as detrimental effects (swamping, invasion, introduction of deleterious alleles or diseases), all of which are largely unpredictable due to the inherent uncertainty surrounding both the climate change trajectory that will occur and the evolutionary response of populations to gene flow. Uncertainty is therefore a critical consideration for the design of practical interventions in natural environments when using predictive models (Maris et al. 2018). There is also no clear answer to the question of what state conservation practices should restore or adapt populations and ecosystems to (Robert 2024), nor to the much debated question of if species should be prioritized (Bottrill et al. 2008) or not (Pimm 2000).

Underlying, however, is an ethical, even philosophical question (*e.g.*, Aubin et al. 2011; Filbee-Dexter and Smajdor 2019): Should we try to manage natural populations or let them

adapt themselves, at the risk of extinction?

## **Appendix C: The necessity of considering founder kinships in conservation breeding programs (publication from Master's thesis)**

During my Masters I contributed to the conceptualisation and development of the methodology of the following study. In particular, I coded and interpreted the theoretical analyses and wrote the Master's thesis. I also participated in the process of writing and reviewing the manuscript.

This study combines a simulation framework with empirical pedigrees to investigate the consequences of the “founders assumption” for the genetic management of conservation breeding programs. This assumption consist in assuming that the founders of these programs are neither inbred nor related. In reality, these genetic relationships are unknown, and as these programs are usually established for highly endangered populations, this assumption is potentially erroneous. We used computer simulations of theoretical wild and captive populations to investigate the magnitude of the bias associated with this assumption, as well as various factors influencing this bias. In a second step, we replaced theoretical captive population pedigrees with real conservation breeding pedigrees for six species of conservation concern. Finally, we estimated molecular-based kinship among founders of a large captive population in one of these species to refine the pedigree analysis by including true founder kinship as an alternative to the founder assumption.



# The necessity of considering founder kinships in conservation breeding programs

R. Rabier<sup>1,2,3</sup> , A. Erlichman<sup>1,2</sup> , L. Lesobre<sup>1,3</sup>  & A. Robert<sup>2</sup>

<sup>1</sup> Reneco International Wildlife Consultants LLC, Abu Dhabi, United Arab Emirates

<sup>2</sup> Centre d'Ecologie et des Sciences de la Conservation (CESCO), Muséum national d'Histoire naturelle, Centre National de la Recherche Scientifique, Sorbonne Université, Paris, France

<sup>3</sup> Emirates Center for Wildlife Propagation, Missour, Morocco

## Keywords

ex situ; relatedness; simulation; gene drop; pedigree; microsatellites; conservation breeding; genetic diversity.

## Correspondence

Robin Rabier, Centre d'Ecologie et des Sciences de la Conservation (CESCO), Muséum national d'Histoire naturelle, CP 135, 57 rue Cuvier, 75005 Paris, France.  
Tel: +33 6 79 51 09 83  
Email: robin\_rabier@hotmail.fr

Editor: Jeff Johnson

Associate Editor: Catherine Grueber

Received 18 June 2021; accepted 03 March 2022

doi:10.1111/acv.12779

## Abstract

Conservation breeding programs pursue the double objective of preserving genetic diversity and producing individuals to support in situ conservation measures. Genetic management and monitoring are commonly based on pedigree analyses and are therefore dependent on the pedigree quality and the underlying assumption that founders of the captive population are neither inbred nor related to each other or to the captive population. Should founders be related, this assumption may lead to an overestimation of genetic diversity and an underestimation of inbreeding in captivity. In this study, we examined the effects of the founder assumption on estimates of genetic diversity by combining three approaches: (1) computer simulations of pedigrees of theoretical captive populations and their associated in situ source populations, (2) analysis of pedigrees of real conservation breeding programs and (3) microsatellite-based computation of kinship among founders from a captive population of houbara bustard. The theoretical approach revealed that the captive population's average mean kinship  $M_k$  increases linearly with founder average  $M_k$ . The bias in  $M_k$  due to the founder assumption was strongly related to the effective size of the source population and to pedigree quality while remaining critical after 25 generations of captivity. Results based on real populations were consistent with theoretical ones and confirmed the overestimation of genetic diversity in captive populations. Overall, our results indicate that under situations classically encountered in conservation breeding programs (e.g. small and isolated source population, incomplete pedigrees), estimates of genetic diversity are potentially highly overestimated, challenging the genetic management of captive populations of species of conservation concern.

## Introduction

Following the ongoing biodiversity crisis (Ceballos, Ehrlich & Dirzo, 2017), the combination of both in and ex situ conservation actions are necessary to save many species from extinction (IUCN SSC, 2013). In 2011, about 15% of threatened species (i.e. 691 species) were subjected to conservation breeding programs (Conde *et al.*, 2011). In animals, such programs are often under the guidance of regional or global organizations and managed through studbooks to optimize the conservation of the species. The objective of many conservation breeding programs is to produce sufficient individuals to support in situ conservation measures to save species from extinction in the wild (IUCN, 2014), with an emphasis on reintroductions (Fischer & Lindenmayer, 2000). Simultaneously, conservation breeding programs have to

maximize the maintenance of genetic diversity in captivity (Ballou, 1992; Ralls & Ballou, 2013). These aims are not independent since potential genetic changes associated with captivity can have deleterious effects on the dynamics and viability of restored populations (Lynch & O'Hely, 2001; Ford, 2002; Robert, 2009) and thus on the very success of programs (Robert *et al.*, 2015). Therefore, strict genetic management is required to minimize risks of genetic changes in captivity such as inbreeding, genetic drift or adaptation to captivity (Lacy, 1987; Frankham, 2008). Genetic management is usually based on the concept of kinship calculated from pedigree analyses and defined as the probability that two alleles of the same gene, randomly drawn from two individuals are identical by descent (Ballou & Lacy, 1995). The average kinship between an individual and the whole population including itself defines its mean kinship.

Although mean kinship and genetic diversity are not always negatively correlated, both empirical (Montgomery *et al.*, 1997; Wisely, McDonald & Buskirk, 2003; Rabier *et al.*, 2020) and simulation studies (Ballou & Lacy, 1995; Ivy & Lacy, 2012) have shown that minimizing mean kinship by mating individuals with the lowest value (i.e. individuals least related to the whole captive population) is an effective strategy to maximize the maintenance of genetic diversity. This approach focuses on minimizing the loss of genetic variation relative to the founding population's genetic diversity, which depends on the number of founders but also on their kinship and individual inbreeding. Inference of pedigree-based mean kinship is based on defining founders as neither inbred nor related to each other or to the captive population, except to their own descent (Lacy, 1989). In absence of information on relationships among individuals, it is assumed that these fit the definition of founders (hereafter, the founder assumption; Ruiz-López *et al.*, 2009; Hogg *et al.*, 2019).

This assumption is probably erroneous, as in many cases, conservation breeding programs are implemented to answer a conservation need to avoid global extinction of a threatened species or boosting the dynamics of small declining populations through relocations (McGowan, Traylor-Holzer & Leus, 2017). As these programs are typically established from small, relict or fragmented in situ populations (Jones *et al.*, 2002), they are likely to already exhibit a loss of genetic diversity and increased inbreeding, as a consequence of isolation and temporary or long-term reduced population size (Spielman, Brook & Frankham, 2004; Ruiz-López *et al.*, 2009; Hammerly, Morrow & Johnson, 2013). Furthermore, there are two situations where the founder assumption is likely to be erroneous. The first concerns programs where founders are regularly collected from the same in situ populations; a practice that is encouraged to minimize both the risk of adaptation to captivity and genetic drift (Lacy, 1987; Conway, 1995). Here, founders newly collected are likely to be related to those initially collected. In addition, when the in situ population is reinforced by captive-bred individuals, a genetic dependence between captive and in situ populations is expected through the two-way gene flow and newly collected founders are likely to be genetically related to the captive population. The second situation is that of programs involving captive-bred individuals of unknown ancestry (i.e. individuals for which parents are unknown due to pedigree uncertainty). In most analytical frameworks, such individuals are assumed to be outbred and unrelated to the captive population even if they are not founders, which brings the same potentially erroneous assumption (Russello & Amato, 2004).

Loss of genetic variation and inbreeding are cumulative processes (Ewing *et al.*, 2008; Biebach & Keller, 2010). While limiting the loss of genetic variation relative to the founding population is crucial for conservation breeding programs, it is necessary to investigate the actual genetic state of the captive population and its potential consequences on population viability, dynamics and evolutionary resilience. Ignoring true founder relationships in genetic management

may lead to an overestimation of genetic diversity and an underestimation of inbreeding (Rudnick & Lacy, 2008). As an alternative, it is possible to estimate founder kinships through molecular genetics and integrate the results into pedigree analyses (Russello & Amato, 2004; Ivy *et al.*, 2009).

Here, we aimed to assess and quantify the effects and implications of the founder assumption on pedigree analyses and the genetic management of conservation breeding programs. Our analyses rely on a combination of theoretical pedigree simulations, the use of real pedigrees of six species subject to conservation breeding programs and molecular estimation of founder kinship in a large captive population in one of these species. Within the theoretical approach, we simulated the pedigree of an ideal in situ source population from which founders were randomly drawn to establish a focal captive population. Pedigree analysis of this theoretical captive population was then conducted by considering either the true genetic relationships among founders or by considering that they were unrelated and outbred (i.e. the founder assumption). The magnitude of the error induced by the founder assumption was examined as a function of various genetic and demographic parameters (i.e. effective size of the source population, number of founders of the captive population, effective size of the captive population and percentage of pedigree known). Secondly, we replaced theoretical captive populations with real populations of six species subject to a conservation breeding program and performed the same analyses. Finally, we used 19 microsatellite loci to produce a molecular-based kinship matrix between founders of a captive population of North-African houbara bustard (*Chlamydotis undulata undulata*, Jacquin 1784) to compare pedigree analyses with and without its inclusion.

From a theoretical point of view, our first main general expectation was that the average mean kinship of the captive population was positively related to the founder average kinship, even after several generations of captive breeding. Thus, we expected that the founder assumption will lead to an underestimation of the true mean kinship of the captive population when founders were related. Because founders collected within a small source population are likely to be more related than founders from a large population, we also expected that the magnitude of the bias associated with the founder assumption (i.e. the difference between the true mean kinship of the captive population and mean kinship estimated under the founder assumption) will increase when the effective size of the source population decreases.

Our second general prediction concerned the effect of pedigree quality (i.e. the proportion of individuals with known parents). Given that individuals with unknown parents are considered outbred and unrelated to the captive population in pedigree-based genetic assessments, our prediction was that the captive population's mean kinship would be more underestimated in low-quality pedigrees. Thus, in both theoretical and real captive populations, we predicted that the biases related to pedigree quality and founder assumption added up.

## Material and methods

### Simulation model and theoretical parameters.

#### Model structure

The principle of the model was first to simulate an ideal source population (i.e. panmictic, closed, constant size and non-overlapping generations). All founders of the source populations were created as neither related nor inbred and of generation 0. After 25 generations, individuals were randomly drawn to create a captive population (hereafter, these founders of the captive population are referred to as the founders). Captive populations were simulated the same way as the source one over 25 generations. For both populations, we used the function *simul.pedigree* of the package *synbreed* 0.12–9 (Wimmer *et al.*, 2012) in R 3.5.1 (R Core Team, 2019) to perform random mating at each generation. Since there was no overlapping generation, a generation corresponded to a reproduction event (i.e. the generation of an individual was the average generation of its parents plus one). We used the following input to parameterize the models: effective size of the source population ( $N_{e_s}$ ), number of founders of the captive population ( $N_f$ ), effective size of the captive population ( $N_{e_c}$ ) and percentage of pedigree known ( $P_k$ ), which was computed as the proportion of individuals with known ancestry in the pedigree (see below). In addition, a gene drop was simulated whereby two unique alleles of one locus were assigned to each founder of the source population and offspring genotypes were generated following the principles of Mendelian segregation, assuming the absence of mutation and equal fertility capacity of alleles (MacCluer *et al.*, 1986). Gene drop function allowed generating the genotype of the founders of the captive population using two methods: (1) the genotype of each founder corresponded to its true genotype obtained from the gene drop applied on the source population or (2) two new unique alleles were attributed to each founder. Method 1 allowed considering true founder relationships while method 2 allowed simulating a gene drop under the founder assumption. The diagram of the simulation model is presented in Supporting Information Appendix S1. R script used for the simulations of the source and captive populations can be found at <https://github.com/R-script/The-necessity-of-considering-founders-kinship-in-conservation-breeding-programs/blob/059db918039d64e33fda6fdb1f39812194a4f051/simulped.R>

#### Pedigree analysis

We used the function *pedIBD* of the R package *optiSel* 2.0.2 (Wellmann, 2018) to perform pedigree analyses and compute the kinship. Kinship computations among captive individuals were either based on true founder kinships or under the founder assumption. Kinships were averaged to compute individual mean kinship ( $M_k$ ). Individuals with unknown parents were considered as individuals of generation 0,

outbred and not related to other captive individuals ( $M_k = 0$ ). Homozygosity was based on gene drop and computed as the proportion of homozygotes. Pedigree analyses were performed at generations 5 and 25. We computed the captive population's average mean kinship (average  $M_k$ ) as the average mean kinship of captive individuals at a given generation, and we computed the founder average mean kinship (founder average  $M_k$ ) as the average mean kinship of the founders of the captive population. To quantify the bias associated with the founder assumption, we computed the difference between average  $M_k$  under the assumption that all founders are unrelated (i.e. founder assumption) and average  $M_k$  when including the matrix of founder kinships within pedigree analyses ( $\Delta M_k$ ). To capture the effect of pedigree quality on the estimation of average  $M_k$ , we computed two types of bias in incomplete theoretical pedigrees. The first one ( $\Delta M_{k1}$ ) was the bias in average  $M_k$  due to the founder assumption for a specific percentage of pedigree known and the second one ( $\Delta M_{k2}$ ) was the bias in average  $M_k$  due to both the founder assumption and pedigree quality (see Supporting Information Appendix S2 for detailed computation methods of  $\Delta M_{k1}$  and  $\Delta M_{k2}$ ).

#### Model parameters

We used this simulation model to examine the influence of several parameters on the magnitude of  $\Delta M_k$ , by varying the input parameters described above:  $N_{e_s}$ ,  $N_f$ ,  $N_{e_c}$  and  $P_k$ . We tested 25 scenarios corresponding to combinations of these parameters (Table 1). For each scenario, the whole process (i.e. simulations of source and captive populations) was repeated 500 times and averaged results were computed.

#### Model validation

Model validation (Supporting Information Appendix S3) was based on the comparison of three inbreeding estimators: the pedigree-based individual inbreeding coefficient (*pedInbreeding* function of the package *optiSel*), the gene drop-based homozygosity and the population inbreeding coefficient computed according to the Wright's theoretical formula (Wright, 1969) which predicts that under genetic drift, the increase in inbreeding level per generation is equal to  $1/2N_e$ , where  $N_e$  is the effective population size.

**Table 1** Input parameters of the model used to simulate theoretical populations

Parameters	Values
Source population	
Number of generations	25
Effective population size $N_{e_s}$	30, 100, 1,000
Captive population	
Number of founders $N_f$	5, 30, 100
Number of generations	25
Effective population size $N_{e_c}$	20, 50, 250, 1,000

## Real populations

Alternatively to theoretical pedigrees of captive populations, we used real pedigrees of six populations of species that are subject to conservation breeding programs: the binturong (*Arctictis binturong*, Raffles 1821), the gaur (*Bos gaurus*, Hamilton Smith 1827), the European otter (*Lutra lutra*, Linnaeus 1758), the Arabian oryx (*Oryx leucorix*, Pallas 1777), the African houbara bustard (*C. u. undulata*, Jacquin 1784; hereafter, African houbara) and the Asian houbara bustard (*Chlamydotis macqueenii*, Gray 1832; hereafter, Asian houbara). Due to incomplete knowledge of the global pedigree of the oryx, along with discrepancies in pedigree records quality between both Middle East and European pedigrees, we performed analyses at the European level (oryx EU) and at the global level (oryx WORLD). Pedigrees are described in Supporting Information Appendix S4 and species are described in Supporting Information Appendix S5. Since the founder kinships of these captive populations were not available, we used our simulation model to simulate their source populations (with the same model and protocol as for the theoretical analysis). A single founding event was considered, and founders were drawn in the simulated source population with a number  $N_f$  derived from real pedigree records. We compared theoretical and real pedigrees at generation 5 since the maximum number of generations of the real programs ranged from 5 to 7 (Supporting Information Appendix S4). To study the effect of  $P_k$  in real pedigrees, we could only compute the bias associated with the founder assumption ( $\Delta Mk_1$ ) since genetic relationships of individuals of unknown parents with the rest of the captive population remained unknown.

## Molecular analysis in African houbara

### Microsatellite genotyping

We used 19 microsatellite markers to produce a kinship matrix between founders of a captive population of African houbara. The genotyping of 19 microsatellite loci, that is A10, A18, A205, A106, A113a, A120, A208, A21, A210, A29, A2, CAM-04, CAM-07, CAM-19, D117, D118, D119, Otmic27, Otmic38 (Chbel *et al.*, 2002; Pitra *et al.*, 2004), was performed by Genoscreen (Lille, France) following conditions described in (Lesobre *et al.*, 2010a).

### Molecular-based founder kinship

CERVUS 3.0.7 (Kalinowski, Taper & Marshall, 2007) was used to check that the null allele rate of microsatellites was below 0.1, the lack of linkage disequilibrium as well as the Hardy–Weinberg equilibrium of each locus. Microsatellite data were used to compute the founder kinships of the 2018 captive population of African houbara (Supporting Information Appendix S6). We also computed kinships among 99 wild individuals sampled within the in situ population of African houbara. To select the most appropriate kinship estimator, we compared 7 estimators of relatedness and 1

estimator of kinship following the protocol described in (Lesobre *et al.*, 2010b). We selected the kinship estimator of Loiselle (Loiselle *et al.*, 1995) that was computed with SPAGEDI 1.5 (Hardy & Vekemans, 2002) and whose negative values were set to 0. Details of the selection of the estimators can be found in Supporting Information Appendix S7.

## Pedigree analyses

Analysis of the pedigree of the 2018 captive population of African houbara was performed as described above (function *pedIBD* of the R package *optiSel* 2.0.2) but under four scenarios defining the founder kinships matrix:

- 1 *Classic*: all founders were assumed to be unrelated (founder assumption) and no founder kinships matrix was included in pedigree analyses.
- 2 *Pseudo-founders*: most founders were considered unrelated (founder assumption), except those known to have been collected within the same nest or laid by the same female, which was considered full sibs. A kinship matrix including a kinship of 0.25 between these related founders was included in pedigree analyses.
- 3 *Molecular-kinship*: a founder kinships matrix was computed using molecular estimates of Loiselle's kinship. This kinship matrix was included in pedigree analyses.
- 4 *Simulated-kinship*: founders were drawn from a simulated source population of relatively large effective size ( $N_{e_s} = 1000$ ), based on estimates of in situ population size at the beginning of the program ( $N \approx 10\,000$  individuals; Goriup, 1997) and the ratio  $N_e/N$  compiled by several extensive studies (Frankham, 1995; Palstra & Ruzzante, 2008). The pedigree-based founder kinships matrix was included in pedigree analyses following 500 iterative computations and the captive population average  $M_k$  was estimated over these iterations.

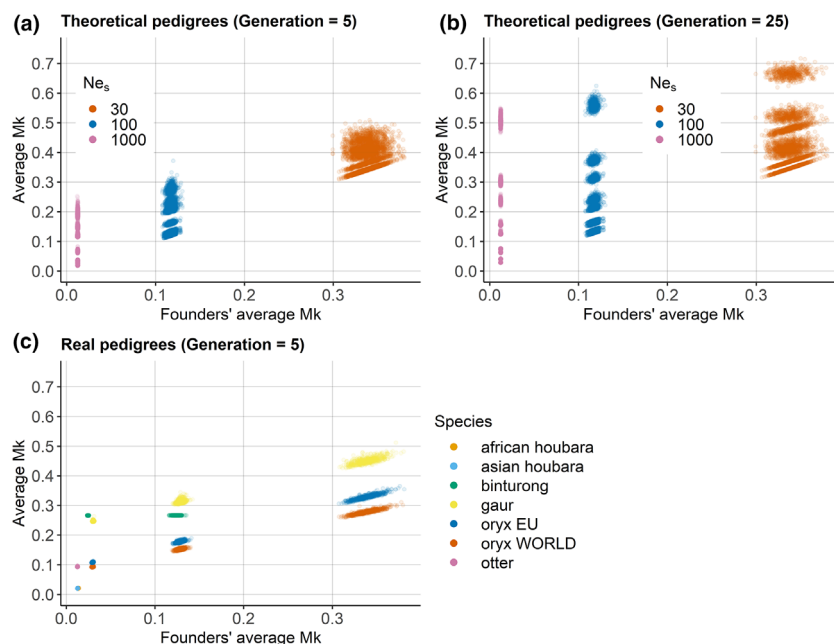
Significances of average  $M_k$  computed according to each scenario were tested with a Student's t-test while differences between scenarios were tested with a one-factor analysis of variance (ANOVA) including the scenario as factor and a following post hoc test of Tukey. All statistics were performed in R 3.5.1 (R Core Team, 2019).

## Results

### Theoretical and real pedigrees

Considering that results computed from gene drop and those computed from pedigree analyses were consistent, we only present the results of pedigree analyses (see Supporting Information Appendix S8 for gene drop results).

We performed the first set of simulations assuming that the percentage of pedigree known was 100% (i.e. in this case,  $\Delta Mk = \Delta Mk_1 = \Delta Mk_2$ ). The captive population's average mean kinship increased linearly with founder average mean kinships in both theoretical (Fig. 1a and b) and real pedigrees (Fig. 1c), regardless of the scenario and the generation number. After five generations, results computed



**Figure 1** Increase in captive population average mean kinship Mk according to average founder mean kinships Mk in theoretical pedigrees at generation 5 (a) and 25 (b), and in real pedigrees (c). Each point represents an iteration.  $N_{e_s}$ : effective size of the source population.

on theoretical and real pedigrees (Fig. 1a and c, respectively) exhibited similar ranges of variation and slopes. By generation 25, average Mk reached higher values and variation between iterations increased (Fig. 1a and b). Note that the vertical grouping of points in Fig. 1a and b was due to variations in other parameters (i.e.  $N_f$  and  $N_{e_c}$ ; Supporting Information Appendix S9). Bias in the estimations of average Mk due to the founder assumption ( $\Delta Mk$ ) decreased with the effective size of the source population ( $N_{e_s}$ ) in theoretical pedigrees at generations 5 and 25. The magnitude of this bias reached substantial values when the captive population was founded from small source one ( $0.12 < \Delta Mk < 0.37$  when  $N_{e_s} = 30$ ; Fig. 2a). Qualitatively, similar results were obtained with real pedigrees, which also showed a predominant effect of the effective size of the source population on the magnitude of the bias (Fig. 2b).

In theoretical pedigrees,  $\Delta Mk$  was weakly affected by the number of founders (Fig. 2c) or the effective size of the captive population (Fig. 2e). In contrast, in real pedigrees (Fig. 2d and 2f),  $\Delta Mk$  seemed to be affected by the number of founders (e.g. average  $\Delta Mk \approx 0.09$  for  $18 < N_f \leq 28$  while  $\Delta Mk \approx 0.01$  for  $39 \geq N_f$ ) and effective size of the captive population.

Secondarily, we considered various proportions of pedigree known. We computed both the bias in average Mk due to the founder assumption ( $\Delta Mk_1$ ) for a certain level of pedigree known and the overall bias in average Mk due to both the founder assumption and pedigree uncertainty ( $\Delta Mk_2$ ). In these simulations,  $\Delta Mk_1$  increased linearly with the proportion of pedigree known (Fig. 3a and b) while  $\Delta Mk_2$  decreased (Fig. 3c and d). This was especially true when the captive population was founded from a small

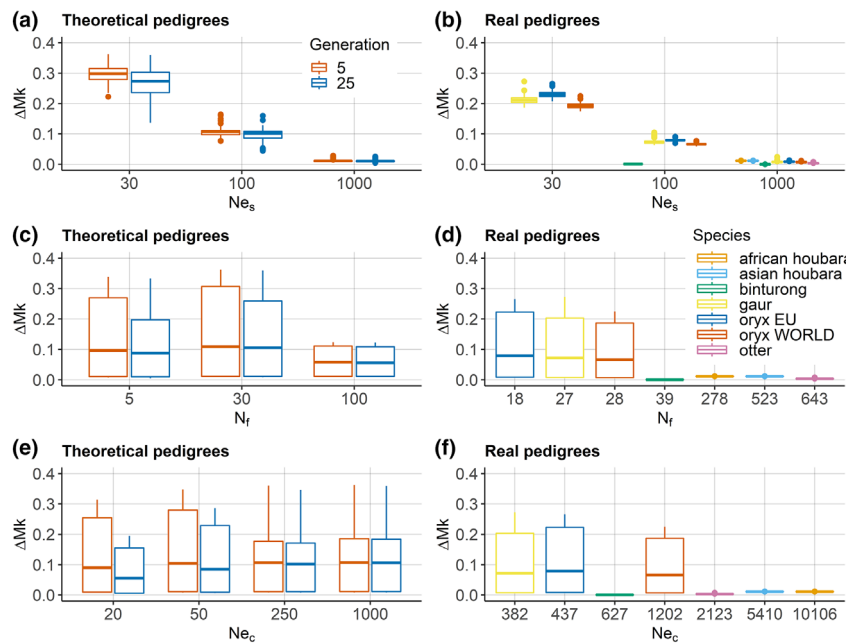
source population. Shifts in  $\Delta Mk_1$  were weaker at generation 25 than at generation 5 (Fig. 3a and b). The range of variation in  $\Delta Mk_2$  was higher at generation 25 than at generation 5 (Fig. 3c and d). For  $\Delta Mk_1$ , this trend was also found in real pedigrees, although the slope was smaller (i.e. increase from 0 to about 0.012; Fig. 4).

### Molecular analysis in African houbara

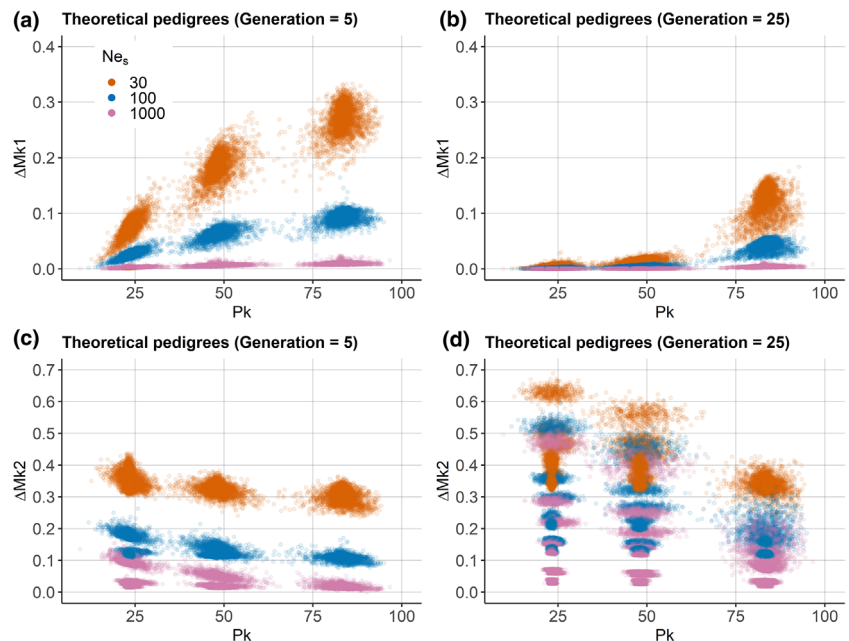
With the *Classic* scenario, average Mk was 0.007 ( $t = 205.4$ , d.f. = 8630,  $P$ value  $< 0.001$ ; 95% CI: 0.006–0.007), while it was 0.035 ( $t = 684.3$ , d.f. = 8630,  $P$ value  $< 0.001$ ; 95% CI: 0.035–0.035) with the scenario *Molecular-kinship*, 0.007 ( $t = 244.7$ , d.f. = 8630,  $P$ value  $< 0.001$ ; 95% CI: 0.007–0.007) with the scenario *Pseudo-founders* and 0.017 ( $t = 531.6$ , d.f. = 8630,  $P$ value  $< 0.001$ ; 95% CI: 0.017–0.017) with the scenario *Simulated-kinship* (Fig. 5). Differences in average Mk between all pairs of scenarios were all significant (see Supporting Information Appendix S10). The molecular-based founder average Mk was 0.033 ( $t = 109.05$ , d.f. = 257,  $P$ value  $< 0.001$ ; 95% CI: 0.033–0.034). The molecular-based average Mk among the 99 wild houbara was 0.042 ( $t = 27.91$ , d.f. = 97,  $P$ value  $< 0.001$ ; 95% CI: 0.039–0.045).

### Discussion

In agreement with previous works (Ballou & Lacy, 1995; Jiménez-Mena *et al.*, 2016), our study indicates that the widespread founder assumption, according to which founders harvested to constitute the initial captive population are unrelated and outbred, generates the risk of biasing the genetic



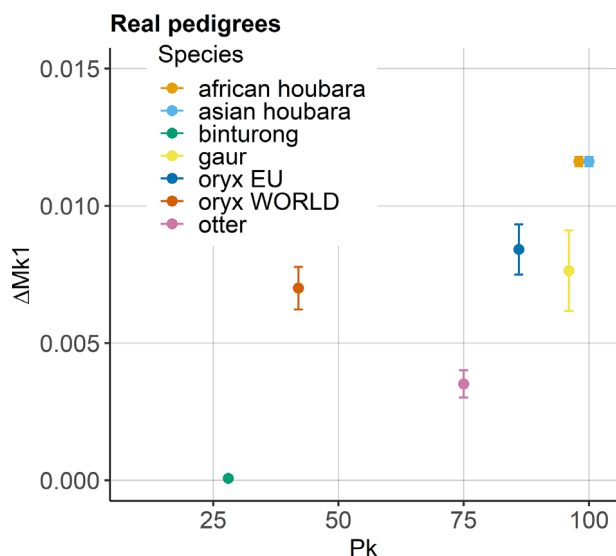
**Figure 2** Bias in the estimation of captive population average mean kinship due to the founder assumption ( $\Delta M_k$ ) in theoretical (a, c, e) and real pedigrees (b, d, f), according to the effective size of the source population  $N_{e_s}$  (a, b), the number of founders  $N_f$  (c, d) and the effective size of the captive population  $N_{e_c}$  (e, f).  $\Delta M_k$  was not computed for each value of  $N_{e_s}$  for the seven real pedigrees since,  $N_{e_s}$  was constrained by true  $N_f$  (e.g. the pedigree of the binturong included  $N_f = 39$  founders so the effective size of the source population could not be  $N_{e_s} = 30$  but only  $N_{e_s} = 100$  or  $N_{e_s} = 1000$ ).



**Figure 3** Bias in the estimation of captive population average mean kinship in theoretical pedigrees at generations 5 and 25, (a, b) due to the founder assumption for a certain percentage of pedigree known (excluding the bias due to pedigree uncertainty;  $\Delta M_k1$ ), and (c, d) due to both the founder assumption and the pedigree uncertainty ( $\Delta M_k2$ ). Pk, percentage of pedigree known.

monitoring of the captive population. Our theoretical analysis confirmed that the average mean kinship of the captive population is positively related to the average mean kinship of

the founder, even after 25 generations of captive breeding, which leads to an underestimation of the captive population's true mean kinship, especially when founders come from



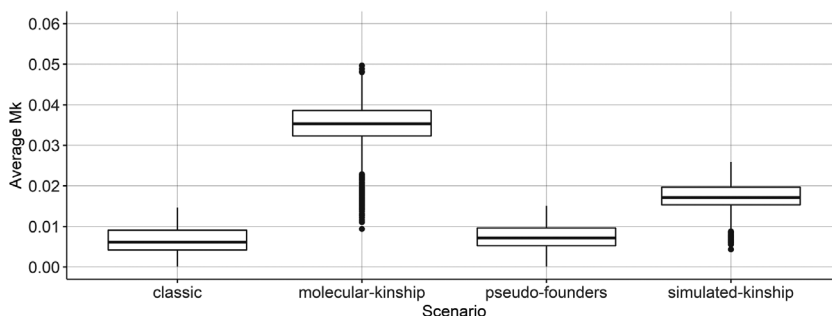
**Figure 4** Bias in the estimation of captive population average mean kinship in real pedigrees, due to the founder assumption for a certain percentage of pedigree known (excluding the bias due to pedigree uncertainty;  $\Delta Mk1$ ). Pk, percentage of pedigree known.

small source populations. Our analysis of real pedigrees was consistent with theoretical results and confirmed a potential overestimation of genetic diversity in captive populations, although results were strongly affected by the quality of these pedigrees. As expected, we found that a high proportion of individuals with unknown parents in the pedigree leads to an overestimation of genetic diversity (as the founder assumption does) and that the biases related to the pedigree quality and founder assumption are approximately additive.

### Short- and long-term consequences of the founder assumption on genetic monitoring

Our analyses were conducted after 5 generations in captivity, to fit with widespread situations of conservation breeding

programs, where the number of generations is kept low to minimize the risks of adaptation to captivity (Williams & Hoffman, 2009; Schäfer & Reiners, 2017), and after 25 generations, to approach the long-term goal of maintaining at least 90% of the initial genetic diversity (Soulé *et al.*, 1986). Our results indicated that, regardless of the number of generations (i.e. 5 or 25), the average mean kinship of the captive populations increases linearly with the founder kinships and that the bias due to the founder assumption is crucially related to the effective size of the source population. The impact of the effective size of the source population is explained by the characteristics of the source population that dictates founder kinships since a decrease in its size, associated with the closed nature of the theoretical source population, will lead to an increased probability of drawing two related founders. Increases in the average mean kinship of captive populations when accounting for founder kinships is expected to be larger when these are founded from small in situ populations (Rudnick & Lacy, 2008). Short-term results (at generation 5) echo a simulation study conducted by Rudnick & Lacy (2008) and an empirical one conducted by Schäfer & Reiners (2017) in a captive population of Nepalese red panda *Ailurus f. fulgens*. Both found that under realistic scenarios defining founder kinships, the founder assumption had an early strong effect on inbreeding level and loss of genetic diversity. However, both studies highlighted that the effects of considering true founder kinships rapidly decreased and reached stability after 5, for Rudnick & Lacy (2008), or 4 generations for Schäfer & Reiners (2017). In contrast, our theoretical results did not highlight a decrease in the effects of the founder assumption with generations. This discrepancy might partly be due to differences in simulated and actual captive population management. Both the simulated populations created by Rudnick & Lacy (2008) and the captive one of red panda studied by Schäfer & Reiners (2017) were managed to minimize the mean kinship in the population. Here, we simulated captive populations with random mating, therefore without any kind of genetic management. This suggests that strategy aiming at minimizing mean kinship can limit the medium- and long-term negative bias associated with the founder assumption.



**Figure 5** Captive population average mean kinship Mk of the 2018 captive population of African houbara computed according to four scenarios defining how founders were considered (i.e. unrelated, molecular-based kinship, founders of the same nest considered full sibs and or simulation-based kinship).

### **Pedigree uncertainty increases bias in estimates of the captive populations' genetic diversity**

With incomplete pedigrees, we measured biases associated with the founder assumption in two ways. The first one was only due to the founder assumption without accounting for the true genetic relationships between individuals of unknown parents and the rest of the captive population ( $\Delta\text{Mk1}$ ), while the second one was due to both the founder assumption and the genetic relationships between individuals of unknown parents and the rest of the captive population ( $\Delta\text{Mk2}$ ), which provides a measure of the overall bias in the estimation of the average mean kinship of captive populations. Our results indicate that pedigree quality (i.e. the proportion of captive individuals with known parents) has opposite effects on the indicators  $\Delta\text{Mk1}$  and  $\Delta\text{Mk2}$ . On one hand,  $\Delta\text{Mk1}$  increases with the proportion of pedigree known (Fig. 3a and b). Indeed, in bad quality pedigrees, most of the bias in the captive population  $\text{Mk}$  is caused by the presence of a large proportion of individuals with unknown parents, which are assumed to be outbred and unrelated to the rest of the captive population (Russello & Amato, 2004). In these cases, if the bias is measured considering only the kinships of founders ( $\Delta\text{Mk1}$ ), it is underestimated in comparison to a good quality pedigree. On the other hand,  $\Delta\text{Mk2}$  decreases with the proportion of pedigree known (Fig. 3c and d), which reflects that the overall bias not only increases with the initial kinship of the founders but also when pedigree uncertainty rises, and that these two effects add up. Thus, pedigrees of poor quality and where founders are likely to be strongly related present a higher risk of strong underestimation of the average mean kinship of captive populations.

### **Founder assumption is likely to be erroneous**

Conservation breeding programs are a response to the global loss of biodiversity, whose agendas are constrained by the adverse conservation status of species (McGowan, Traylor-Holzer & Leus, 2017). More than 15 years ago, a study found that, in 170 threatened taxa, average heterozygosity was lower than in related non-threatened taxa in 77% of the comparisons (Spielman, Brook & Frankham, 2004). Thus, many in situ relict populations of threatened species already suffer from genetic problems impairing their fitness and evolutionary resilience (Ruiz-López *et al.*, 2009; Hammerly, Morrow & Johnson, 2013). This is consistent with a population size threshold value of  $N = 1000$  or less defined by the International Union for Conservation of Nature for the classification of a species as vulnerable in the Red List (IUCN, 2012). This size corresponds to effective population sizes ranging from  $N_e = 100$  to  $N_e = 200$  according to the ratio  $N_e/N$  compiled by several studies (Frankham, 1995; Palstra & Ruzzante, 2008). This also corresponds to intermediate values of the effective size of the source population analyzed

in our simulation model (Table 1). Examples of captive populations founded from small in situ populations are numerous, such as the whooping crane *Grus americana* founded with 88 individuals collected within an in situ population that had suffered a sharp bottleneck (Jones *et al.*, 2002) or the black-lion-tamarin *Leontopithecus chrysopygus* founded with 35 individuals within an in situ population of about 200 individuals (Ayala-Burbano *et al.*, 2020). Founders collected in such populations are unlikely to be unrelated and outbred.

To investigate relationships among founders of conservation breeding programs, the use of molecular-based metrics has been proposed, for example Haig, Ballou & Casna (1994) in the Guam rail *Rallus owstoni*, Gautschi *et al.* (2003) in the bearded vulture *Gypaetus barbatus* and Jones *et al.* (2002) in the whooping crane *Grus americana*. Among the 50 founders of a captive population of Iberian lynx *Lynx pardinus*, an average relatedness of 0.1 (i.e. kinship of 0.05) was found (Kleinman-Ruiz *et al.*, 2019). While molecular-based metrics can deeply improve the understanding of the population-relatedness structure (Russello & Amato, 2004; Kleinman-Ruiz *et al.*, 2019), there is no consensus when it comes to incorporate them in pedigree analyses (Ivy *et al.*, 2009). Indeed, molecular- and pedigree-based estimates of genetic diversity rely on different concepts (Ballou, Amos & Coulson, 2004; Slate *et al.*, 2004; Rabier *et al.*, 2020). Like pedigrees, molecular estimates are associated with uncertainties and biases related to the variation among kinship estimators or characteristics of markers (e.g. type, number, variability; Gautschi *et al.*, 2003).

In the captive population of African houbara, molecular analyses yielded an average molecular mean kinship of 0.033 among the 258 founders. This kinship level is consistent with average value computed among 99 wild individuals sampled within in situ population of African houbara (i.e. 0.042). The inclusion of molecular founder kinships increased the captive population's average mean kinship from 0.007 (*Classic scenario*, i.e. founder assumption scenario) to 0.035 after 5.05 generations of captive breeding, confirming the underestimation of the average mean kinship of the captive population when ignoring founder kinships.

### **Founder assumption can have severe consequences on captive and in situ populations**

Genetic monitoring and management of conservation breeding programs rely on recommendations urging for the preservation of the initial genetic diversity present in the founding population (Soulé *et al.*, 1986; Leus, Traylor-Holzer & Lacy, 2011; Ralls & Ballou, 2013). This comes from the fact that (i) the actual genetic diversity of the founding population is generally unknown; (ii) it is impossible to manipulate or modify the initial genetic diversity, so the only purpose that can be implemented in conservation breeding programs focuses on preserving this initial genetic diversity. Our study



suggests that when founders are genetically related, estimates of the maintenance of the initial genetic diversity are not strongly affected by the founder assumption, that is when averaged on all scenarios, genetic diversity decreased by 0.17 between generation 0 and 25 under the founder assumption and by 0.14 when including true founder kinships. However, relying on the founder assumption does not allow to correctly estimate the level of genetic diversity in cases where founders are strongly related. For example, results showed that under the founder assumption, in intermediate cases with a captive population of 250 individuals founded with 30 founders from a source population of 100, the estimated average  $M_k$  was 0 at generation 0 (founding population), 0.02 at generation 5 and 0.06 at generation 25. When the founder assumption was relaxed and true founder kinships included, these values were respectively 0.12, 0.13 and 0.17. Thus, levels of estimated mean kinship were radically different depending on whether the founder assumption was considered or not. This underestimation of average mean kinship is associated with an underestimation of individual inbreeding (Ivy *et al.*, 2009; Hogg *et al.*, 2019). For example, in a captive population of Iberian lynx, inbreeding coefficient highly increased from  $F < 10^{-4}$  to  $F = 0.2$  when founder kinships were included (Kleinman-Ruiz *et al.*, 2019).

The ultimate goal of conservation breeding programs is often to release individuals in situ through conservation relocations (IUCN SSC, 2013). The viability of restored populations is a cornerstone for the success of such restoration programs (Robert *et al.*, 2015) and it has been suggested that reintroduction programs based on the release of captive-bred individuals are generally less successful than those based on the translocation of wild-bred individuals (Griffith *et al.*, 1989; Wolf *et al.*, 1996; Fischer & Lindenmayer, 2000; Jule, Leaver & Lea, 2008). A significant portion of these differences arises from divergent experiences between individuals born and raised in captivity compared to those born and raised in the wild (e.g. exploratory behavior, predator avoidance, resources usage). Additionally, while, such reduction of fitness in restored populations may be partly related to selective processes such as adaptation to captivity (Frankham, 2008), it may also be explained by reduced evolutionary resilience associated with accumulated genetic drift and inbreeding loads during the captive phase (Robert, 2009) or prior to the captive phase in cases of founders already genetically related. In relation to the inbreeding load, a previous meta-analysis showed that realistic levels of inbreeding depression (i.e.  $B = 12$  diploid lethal equivalents) significantly reduced the projected viability of populations over a broad taxonomic range (i.e.  $-37\%$  in median time to extinction), compared with a model without inbreeding depression (i.e. assuming  $B = 0$ ; O'Grady *et al.*, 2006). However, to correctly estimate the effect of this inbreeding depression in restored populations, it is necessary to accurately estimate the level of inbreeding of released individuals that combines the accumulation of inbreeding in captivity but also the initial inbreeding of the captive population.

## Conclusions

The success of conservation breeding programs is highly dependent on the genetic state of individuals released to create or supplement restored in situ populations. Our study suggests that in certain conditions classically encountered in conservation breeding programs founded from relict in situ populations, ignoring true founder relationships can lead to significant overestimation of genetic diversity and underestimation of inbreeding, especially when founders are highly related and collected within a small source population. One option when founders are known to be related is to analyze the costs and benefits associated with their inclusion or exclusion in the breeding management (Willis & Wiese, 1993). For instance, keeping related founders is associated with little loss of genetic diversity but increased risks of inbreeding when there are many founders. On the contrary, when there are few founders, the risk of excluding one founder is to lose unique genetic material (Wilcken & Lees, 1998). Our simulation approach can also be used to estimate the risks associated with a conservation breeding program before starting it. This approach would provide vital insights on the expected kinship level among founders and help managers in providing recommendations on the size and composition of the required founding population. Furthermore, since the effects of the founder assumption are crucially related to the source population size, we argue that conservation breeding programs should be initiated as early as possible when an in situ population decline is identified (Pritchard *et al.*, 2012). It will allow minimizing the risks of collecting related founders while sampling a large proportion of wild genetic diversity. Furthermore, this approach will grant further founder collections therefore limiting the risks of genetic loss or adaptation to captivity (Rabier *et al.*, 2020). An early start of conservation breeding programs permits a combined use of both in situ and ex situ conservation measures which represents a more effective strategy than the use of only one of these approaches (Redford, Jensen & Breheny, 2012).

## Acknowledgments

We are grateful to HH Sheikh Mohamed bin Zayed Al Nahyan, Crown Prince of Abu Dhabi and Founder of the IFHC, HH Sheikh Theyab Bin Mohamed Al Nahyan, Chairman of the IFHC and HE Mohammed Ahmed Al Bowardi, Deputy Chairman, for their support. This study was conducted under the guidance of Reneco International Wildlife Consultants LLC, a consulting company that manages the IFHC's conservation programs such as the Emirates Center for Wildlife Propagation and the Sheikh Khalifa Houbara Breeding Center. These programs provided pedigree records for the African houbara and the Asian houbara, respectively. We are grateful to M. Saint Jalme, A. Bourgeois and E. Rey who provided access to the pedigrees of the binturong, the gaur, the European otter and the Arabian oryx. We thank Dr. Frédéric Lacroix, Managing Director of Reneco, for his supervision. Two anonymous reviewers provided very helpful and constructive comments on a previous draft.

## Funding information

This work was supported by the Emirates Center for Wildlife Propagation and the International Fund for Houbara Conservation.

## Conflict of interest

The authors declare that they have no known conflict of interest that could have appeared to influence the work reported in this paper.

## References

- Ayala-Burbano, P.A., Galetti Junior, P.M., Wormell, D., Pissinatti, A., Marques, M.C. & de Freitas, P.D. (2020). Studbook and molecular analyses for the endangered black-lion-tamarin; an integrative approach for assessing genetic diversity and driving management in captivity. *Sci. Rep.* **10**, 6781.
- Ballou, J.D. (1992). Genetic and demographic considerations in endangered species captive breeding and reintroduction programs. In *Wildlife 2001: populations: 262–275*. McCullough, D.R. & Barrett, R.H. (Eds). Dordrecht: Springer Netherlands.
- Ballou, J.D. & Lacy, R.C. (1995). Identifying genetically important individuals for management of genetic diversity in pedigreed populations. In *Population management for survival and recovery: analytical methods and strategies in small population conservation: 76–111*. Ballou, J.D., Gilpin, M. & Foose, T.J. (Eds). New York: Columbia University Press.
- Balloux, F., Amos, W. & Coulson, T. (2004). Does heterozygosity estimate inbreeding in real populations? *Mol. Ecol.* **13**, 3021–3031.
- Biebach, I. & Keller, L.F. (2010). Inbreeding in reintroduced populations: the effects of early reintroduction history and contemporary processes. *Conserv. Genet.* **11**, 527–538.
- Ceballos, G., Ehrlich, P.R. & Dirzo, R. (2017). Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *PNAS* **114**, E6089–E6096.
- Chbel, F., Broderick, D., Idaghdour, Y., Korrida, A. & McCormick, P. (2002). Characterization of 22 microsatellites loci from the endangered houbara bustard (*Chlamydotis undulata undulata*). *Mol. Ecol. Notes* **2**, 484–487.
- Conde, D.A., Flesness, N., Colchero, F., Jones, O.R. & Scheuerlein, A. (2011). An emerging role of zoos to conserve biodiversity. *Science* **331**, 1390–1391.
- Conway, W. (1995). Wild and zoo animal interactive management and habitat conservation. *Biodivers. Conserv.* **4**, 573–594.
- Ewing, S.R., Nager, R.G., Nicoll, M.A.C., Aumjaud, A., Jones, C.G. & Keller, L.F. (2008). Inbreeding and loss of genetic variation in a reintroduced population of Mauritius kestrel. *Conserv. Biol.* **22**, 395–404.
- Fischer, J. & Lindenmayer, D.B. (2000). An assessment of the published results of animal relocations. *Biol. Conserv.* **96**, 1–11.
- Ford, M.J. (2002). Selection in captivity during supportive breeding may reduce fitness in the wild. *Conserv. Biol.* **16**, 815–825.
- Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: a review. *Genet. Res.* **66**, 95–107.
- Frankham, R. (2008). Genetic adaptation to captivity in species conservation programs. *Mol. Ecol.* **17**, 325–333.
- Gautschi, B., Müller, J.P., Schmid, B. & Shykoff, J.A. (2003). Effective number of breeders and maintenance of genetic diversity in the captive bearded vulture population. *Heredity* **91**, 9–16.
- Goriup, P.D. (1997). The world status of the houbara bustard *Chlamydotis undulata*. *Bird Conserv. Int.* **7**, 373–397.
- Griffith, B., Scott, J.M., Carpenter, J.W. & Reed, C. (1989). Translocation as a species conservation tool: status and strategy. *Science* **245**, 477–480.
- Haig, S.M., Ballou, J.D. & Casna, N.J. (1994). Identification of kin structure among Guam rail founders: a comparison of pedigrees and DNA profiles. *Mol. Ecol.* **3**, 109–119.
- Hammerly, S.C., Morrow, M.E. & Johnson, J.A. (2013). A comparison of pedigree- and DNA-based measures for identifying inbreeding depression in the critically endangered Attwater's prairie-chicken. *Mol. Ecol.* **22**, 5313–5328.
- Hardy, O.J. & Vekemans, X. (2002). Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* **2**, 618–620.
- Hogg, C.J., Wright, B., Morris, K.M., Lee, A.V., Ivy, J.A., Grueber, C.E. & Belov, K. (2019). Founder relationships and conservation management: empirical kinships reveal the effect on breeding programmes when founders are assumed to be unrelated. *Anim. Conserv.* **22**, 348–361.
- IUCN. (2012). *IUCN Red List categories and criteria*. 2nd edn. Gland, Switzerland: International Union for Conservation of Nature. <https://portals.iucn.org/library/node/10315>.
- IUCN. (2014). *Guidelines on the use of ex situ management for species conservation*. 2nd edn. Gland, Switzerland: International Union for Conservation of Nature. [http://www.iucn.org/about/work/programmes/species/publications/iucn\\_guidelines\\_and\\_policy\\_statements/](http://www.iucn.org/about/work/programmes/species/publications/iucn_guidelines_and_policy_statements/).
- IUCN SSC. (2013). *Guidelines for reintroductions and other conservation translocations*, 1st edn. Gland, Switzerland: International Union for Conservation of Nature.
- Ivy, J.A. & Lacy, R.C. (2012). A comparison of strategies for selecting breeding pairs to maximize genetic diversity retention in managed populations. *J. Hered.* **103**, 186–196.
- Ivy, J.A., Miller, A., Lacy, R.C. & DeWoody, J.A. (2009). Methods and prospects for using molecular data in captive breeding programs: an empirical example using Parma wallabies (*Macropus parma*). *J. Hered.* **100**, 441–454.

- Jiménez-Mena, B., Schad, K., Hanna, N. & Lacy, R.C. (2016). Pedigree analysis for the genetic management of group-living species. *Ecol. Evol.* **6**, 3067–3078.
- Jones, K.L., Glenn, T.C., Lacy, R.C., Pierce, J.R., Unruh, N., Mirande, C.M. & Chavez-Ramirez, F. (2002). Refining the whooping crane studbook by incorporating microsatellite DNA and leg-banding analyses. *Conserv. Biol.* **16**, 789–799.
- Jule, K.R., Leaver, L.A. & Lea, S.E.G. (2008). The effects of captive experience on reintroduction survival in carnivores: a review and analysis. *Biol. Conserv.* **141**, 355–363.
- Kalinowski, S.T., Taper, M.L. & Marshall, T.C. (2007). Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment: CERVUS LIKELIHOOD MODEL. *Mol. Ecol.* **16**, 1099–1106.
- Kleinman-Ruiz, D., Soriano, L., Casas-Marce, M., Szychta, C., Sánchez, I., Fernández, J. & Godoy, J.A. (2019). Genetic evaluation of the Iberian lynx ex situ conservation programme. *Heredity* **123**, 647–661.
- Lacy, R.C. (1987). Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. *Conserv. Biol.* **1**, 143–158.
- Lacy, R.C. (1989). Analysis of founder representation in pedigrees: founder equivalent and founder genome equivalents. *Zoo Biol.* **8**, 111–123.
- Lesobre, L., Lacroix, F., Caizergues, A., Hingrat, Y., Chalah, T. & Saint Jalme, M. (2010a). Conservation genetics of houbara bustard (*Chlamydotis undulata undulata*): population structure and its implications for the reinforcement of wild populations. *Conserv. Genet.* **11**, 1489–1497.
- Lesobre, L., Lacroix, F., Le Nuz, E., Hingrat, Y., Chalah, T. & Saint Jalme, M. (2010b). Absence of male reproductive skew, along with high frequency of polyandry and conspecific brood parasitism in the lekking houbara bustard *Chlamydotis undulata undulata*. *J. Avian Biol.* **41**, 117–127.
- Leus, K., Traylor-Holzer, K. & Lacy, R.C. (2011). Genetic and demographic population management in zoos and aquariums: recent developments, future challenges and opportunities for scientific research. *Int. Zoo Yearb.* **45**, 213–225.
- Loiselle, B.A., Sork, V.L., Nason, J. & Graham, C. (1995). Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* Lieckfeldt (Rubiaceae). *Am. J. Bot.* **82**, 1420–1425.
- Lynch, M. & O'Hely, M. (2001). Captive breeding and genetic fitness of natural populations. *Conserv. Genet.* **2**, 363–378.
- MacCluer, J.W., VandeBerg, J.L., Read, B. & Ryder, O.A. (1986). Pedigree analysis by computer simulation. *Zoo Biol.* **5**, 147–160.
- McGowan, P.J.K., Traylor-Holzer, K. & Leus, K. (2017). IUCN guidelines for determining when and how ex situ management should be used in species conservation. *Conserv. Lett.* **10**, 361–366.
- Montgomery, M.E., Ballou, J.D., Nurthen, R.K., England, P.R., Briscoe, D.A. & Frankham, R. (1997). Minimizing kinship in captive breeding programs. *Zoo Biol.* **16**, 377–389.
- O'Grady, J.J., Brook, B.W., Reed, D.H., Ballou, J.D., Tonkyn, D.W. & Frankham, R. (2006). Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biol. Conserv.* **133**, 42–51.
- Palstra, F.P. & Ruzzante, D.E. (2008). Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Mol. Ecol.* **17**, 3428–3447.
- Pitra, C., D'Aloia, M.-A., Lieckfeldt, D. & Combreau, O. (2004). Genetic variation across the current range of the Asian houbara bustard (*Chlamydotis undulata macqueenii*). *Conserv. Genet.* **5**, 205–215.
- Pritchard, D.J., Fa, J.E., Oldfield, S. & Harrop, S.R. (2012). Bring the captive closer to the wild: redefining the role of ex situ conservation. *Oryx* **46**, 18–23.
- R Core Team. (2019). *R: a language and environment for statistical computing (3.6.1)* [Computer software].
- Rabier, R., Robert, A., Lacroix, F. & Lesobre, L. (2020). Genetic assessment of a conservation breeding program of the houbara bustard (*Chlamydotis undulata undulata*) in Morocco, based on pedigree and molecular analyses. *Zoo Biol.* **39**, 365–447.
- Ralls, K. & Ballou, J.D. (2013). Captive breeding and reintroduction. In *Encyclopedia of biodiversity*: 662–667. Levin, S.A. (Ed.). Waltham: Elsevier Academic Press.
- Redford, K.H., Jensen, D.B. & Breheny, J.J. (2012). Integrating the captive and the wild. *Science* **338**, 1157–1158.
- Robert, A. (2009). Captive breeding genetics and reintroduction success. *Biol. Conserv.* **142**, 2915–2922.
- Robert, A., Colas, B., Guigon, I., Kerbiriou, C., Mihoub, J.-B., Saint-Jalme, M. & Sarrazin, F. (2015). Defining reintroduction success using IUCN criteria for threatened species: a demographic assessment. *Anim. Conserv.* **18**, 397–406.
- Rudnick, J.A. & Lacy, R.C. (2008). The impact of assumptions about founder relationships on the effectiveness of captive breeding strategies. *Conserv. Genet.* **9**, 1439–1450.
- Ruiz-López, M.J., Roldán, E.R.S., Espeso, G. & Gomendio, M. (2009). Pedigrees and microsatellites among endangered ungulates: what do they tell us? *Mol. Ecol.* **18**, 1352–1364.
- Russello, M.A. & Amato, G. (2004). Ex situ population management in the absence of pedigree information. *Mol. Ecol.* **13**, 2829–2840.
- Schäfer, F. & Reiners, T.E. (2017). Long term vs. short term impact of founder relatedness on gene diversity and inbreeding within the European endangered species

- Programme (EEP) of the red panda (*Ailurus f. Fulgens*). *J. Zoo Aquar. Res.* **5**, 86–91.
- Slate, J., David, P., Dodds, K.G., Veenvliet, B.A., Glass, B.C., Broad, T.E. & McEwan, J.C. (2004). Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity* **93**, 255–265.
- Soulé, M., Gilpin, M., Conway, W. & Foose, T. (1986). The millenium ark: how long a voyage, how many staterooms, how many passengers? *Zoo Biol.* **5**, 101–113.
- Spielman, D., Brook, B.W. & Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. *PNAS* **101**, 15261–15264.
- Wellmann, R. (2018). optiSel: Optimum Contribution Selection and Population Genetics (2.0.2) [R]. <https://CRAN.R-project.org/package=optiSel>
- Wilcken, J. & Lees, C. (1998). *Managing Zoo Populations: compiling and analysing studbook data*. Sydney, Australia: Australasian Regional Association of Zoological Parks and Aquaria.
- Williams, S.E. & Hoffman, E.A. (2009). Minimizing genetic adaptation in captive breeding programs: a review. *Biol. Conserv.* **142**, 2388–2400.
- Willis, K. & Wiese, R.J. (1993). Effect of new founders on retention of gene diversity in captive populations: a formalization of the nucleus population concept. *Zoo Biol.* **12**, 535–548.
- Wimmer, V., Albrecht, T., Auinger, H.-J. & Schön, C.-C. (2012). Synbreed: a framework for the analysis of genomic prediction data using R. *Bioinformatics* **28**, 2086–2087.
- Wisely, S.M., McDonald, D.B. & Buskirk, S.W. (2003). Evaluation of the genetic management of the endangered black-footed ferret (*Mustela nigripes*). *Zoo Biol.* **22**, 287–298.
- Wolf, C.M., Griffith, B., Reed, C. & Temple, S.A. (1996). Avian and mammalian translocations: update and reanalysis of 1987 survey data. *Conserv. Biol.* **10**, 1142–1154.
- Wright, S. (1969). *Evolution and the genetics of populations. Volume 2. The theory of gene frequencies*. London: University of Chicago Press.

## Supporting information

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

**Appendix S1.** Diagram of the simulation model.

**Appendix S2.** Computation of  $\Delta M_k1$  and  $\Delta M_k2$ .

**Appendix S3.** Validation of the simulation model.

**Appendix S4.** Description of the real pedigrees.  $N_f$ : number of founders;  $P_k$ : percentage of pedigree known;  $VU$ : vulnerable;  $NT$ : near threatened.

**Appendix S5.** Description of the six species subjected to ex situ conservation actions.

**Appendix S6.** Number of genotyped founders per number of loci within ECWP's 2018 captive population of African houbara.

**Appendix S7.** Diagnosis of the estimators of kinship tested.

**Appendix S8.** Results of gene-drop analysis.

**Appendix S9.** Captive population average mean kinship  $M_k$  according to average founder mean kinships  $M_k$  in theoretical pedigrees at generation 5 (A, C) and 25 (B, D).  $N_f$ : number of founders of the captive populations;  $N_{e_c}$ : effective size of the captive population.

**Appendix S10.** Results of the post-hoc Tukey's test allowing to compare captive population average mean kinship computed according each scenario.

# Bibliography

- Abram, N. J., Henley, B. J., Gupta, A. S., Lippmann, T. J., Clarke, H., Dowdy, A. J., Sharples, J. J., Nolan, R. H., Zhang, T., Wooster, M. J., Wurtzel, J. B., Meissner, K. J., Pitman, A. J., Ukkola, A. M., Murphy, B. P., Tapper, N. J., & Boer, M. M. (2021). Connections of climate change and variability to large and extreme forest fires in southeast Australia. *Communications Earth and Environment*, 2, 8. doi:10.1038/s43247-020-00065-8
- Abramowitz, M., Stegun, I. A., & Romer, R. H. (1988). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. American Association of Physics Teachers.
- Adams, J. R., Vucetich, L. M., Hedrick, P. W., Peterson, R. O., & Vucetich, J. A. (2011). Genomic sweep and potential genetic rescue during limiting environmental conditions in an isolated wolf population. *Proceedings of the Royal Society B: Biological Sciences*, 278(1723), 3336–3344.
- Agnoletti, M. & Santoro, A. (2015). Cultural values and sustainable forest management: the case of europe. *Journal of Forest Research*, 20, 438–444.
- Aitken, S. N. & Bemmels, J. B. (2016). Time to get moving: assisted gene flow of forest trees. *Evolutionary applications*, 9(1), 271–290.
- Aitken, S. N. & Whitlock, M. C. (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics*, 44(1), 367–388. doi:10.1146/annurev-ecolsys-110512-135747
- Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, 1, 95–111. doi:10.1111/j.1752-4571.2007.00013.x
- Åkesson, M., Flagstad, Ø., Aspi, J., Kojola, I., Liberg, O., Wabakken, P., & Sand, H. (2022). Genetic signature of immigrants and their effect on genetic diversity in the recently established scandinavian wolf population. *Conservation Genetics*, 1–15.

- Åkesson, M., Liberg, O., Sand, H., Wabakken, P., Bensch, S., & Flagstad, Ø. (2016). Genetic rescue in a severely inbred wolf population. *Molecular ecology*, *25*(19), 4745–4756.
- Alberto, F. J., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A., Lefèvre, F., Lenormand, T., Yeaman, S., Whetten, R. et al. (2013). Potential for evolutionary responses to climate change—evidence from tree populations. *Global change biology*, *19*(6), 1645–1661.
- Allendorf, F. W., Leary, R. F., Spruell, P., & Wenburg, J. K. (2001). The problems with hybrids: setting conservation guidelines. *Trends in ecology & evolution*, *16*(11), 613–622.
- Anderson, J. T., Inouye, D. W., McKinney, A. M., Colautti, R. I., & Mitchell-Olds, T. (2012). Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1743), 3843–3852.
- Au, T. F., Maxwell, J. T., Robeson, S. M., Li, J., Siani, S. M., Novick, K. A., Dannenberg, M. P., Phillips, R. P., Li, T., Chen, Z., & Lenoir, J. (2022). Younger trees in the upper canopy are more sensitive but also more resilient to drought. *Nature climate change*, *12*(12), 1168–1174. doi:10.1038/s41558-022-01528-w
- Aubin, I., Garbe, C., Colombo, S., Drever, C., McKenney, D., Messier, C., Pedlar, J., Saner, M., Venier, L., Wellstead, A. et al. (2011). Why we disagree about assisted migration: ethical implications of a key debate regarding the future of Canada's forests. *The Forestry Chronicle*, *87*(6), 755–765.
- Baack, E. J. & Rieseberg, L. H. (2007). A genomic view of introgression and hybrid speciation. *Current opinion in genetics & development*, *17*(6), 513–518.
- Barfield, M., Holt, R. D., & Gomulkiewicz, R. (2011). Evolution in Stage-Structured Populations. *The American Naturalist*, *177*(4), 397–409. doi:10.1086/658903
- Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C. et al. (2011). Has the earth's sixth mass extinction already arrived? *Nature*, *471*(7336), 51–57.
- Barton, N. H. (2000). Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1403), 1553–1562.
- Baskett, M. L. & Gomulkiewicz, R. (2011). Introgressive hybridization as a mechanism for species rescue. *Theoretical Ecology*, *4*(2), 223–239.

- Bauhus, J., Puettmann, K., & Messier, C. (2009). Silviculture for old-growth attributes. *Forest Ecology and Management*, *258*, 525–537. doi:10.1016/J.FORECO.2009.01.053
- Bay, R. A., Rose, N. H., Logan, C. A., & Palumbi, S. R. (2017). Genomic models predict successful coral adaptation if future ocean warming rates are reduced. *Science Advances*, *3*(11), e1701413. doi:10.1126/sciadv.1701413
- Bell, D. M., Bradford, J. B., & Lauenroth, W. K. (2014). Early indicators of change: divergent climate envelopes between tree life stages imply range shifts in the western United States. *Global Ecology and Biogeography*, *23*(2), 168–180. doi:10.1111/geb.12109
- Bell, D. A., Robinson, Z. L., Funk, W. C., Fitzpatrick, S. W., Allendorf, F. W., Tallmon, D. A., & Whiteley, A. R. (2019). The exciting potential and remaining uncertainties of genetic rescue. *Trends in Ecology & Evolution*, *34*(12), 1070–1079.
- Bell, G. (2017). Evolutionary rescue. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 605–627. doi:10.1146/annurev-ecolsys-110316-023011
- Bell, G. & Collins, S. (2008). Adaptation, extinction and global change. *Evolutionary Applications*, *1*(1), 3–16.
- Benito-Garzón, M., Fady, B., Davi, H., Vizcaíno-Palomar, N., & Fernández-Manjarrés, J. (2018). Trees on the move: using decision theory to compensate for climate change at the regional scale in forest social-ecological systems. *Regional Environmental Change*, *18*, 1427–1437.
- Black, R. A. & Bliss, L. (1980). Reproductive ecology of picea mariana (mill.) bsp., at tree line near inuvik, northwest territories, canada. *Ecological Monographs*, *50*(3), 331–354.
- Bohling, J. H. (2016). Strategies to address the conservation threats posed by hybridization and genetic introgression. *Biological Conservation*, *203*, 321–327.
- Bontrager, M. & Angert, A. L. (2019). Gene flow improves fitness at a range edge under climate change. *Evolution Letters*, *3*(1), 55–68. doi:10.1002/evl3.91
- Borrell, J. S., Zohren, J., Nichols, R. A., & Buggs, R. J. (2020). Genomic assessment of local adaptation in dwarf birch to inform assisted gene flow. *Evolutionary Applications*, *13*(1), 161–175.
- Bottrill, M. C., Joseph, L. N., Carwardine, J., Bode, M., Cook, C., Game, E. T., Grantham, H., Kark, S., Linke, S., McDonald-Madden, E. et al. (2008). Is conservation triage just smart decision making? *Trends in ecology & evolution*, *23*(12), 649–654.

- Brady, S. P., Bolnick, D. I., Angert, A. L., Gonzalez, A., Barrett, R. D., Crispo, E., Derry, A. M., Eckert, C. G., Fraser, D. J., Fussmann, G. F. et al. (2019). Causes of maladaptation. *Evolutionary Applications*, *12*(7), 1229–1242.
- Breed, M. F., Harrison, P. A., Bischoff, A., Durruty, P., Gellie, N. J., Gonzales, E. K., Havens, K., Karmann, M., Kilkenny, F. F., Krauss, S. L. et al. (2018). Priority actions to improve provenance decision-making. *BioScience*, *68*, 510–516. doi:10.1093/biosci/biy050
- Brienen, R. J. W., Gloor, E., & Zuidema, P. A. (2012). Detecting evidence for CO<sub>2</sub> fertilization from tree ring studies: the potential role of sampling biases. *Global Biogeochemical Cycles*, *26*(1). doi:10.1029/2011GB004143. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011GB004143>
- Brockhoff, E. G., Barbaro, L., Castagnyrol, B., Forrester, D. I., Gardiner, B., González-Olabarria, J. R., Lyver, P. O., Meurisse, N., Oxbrough, A., Taki, H., Thompson, I. D., van der Plas, F., & Jactel, H. (2017). Forest biodiversity, ecosystem functioning and the provision of ecosystem services. *Biodiversity and Conservation*, *26*, 3005–3035. doi:10.1007/s10531-017-1453-2
- Brodribb, T. J., Powers, J., Cochard, H., & Choat, B. (2020). Hanging by a thread? Forests and drought. *Science*, *368*(6488), 261–266.
- Browne, L., Wright, J. W., Fitz-Gibbon, S., Gugger, P. F., & Sork, V. L. (2019). Adaptational lag to temperature in valley oak (*Quercus lobata*) can be mitigated by genome-informed assisted gene flow. *Proceedings of the National Academy of Sciences*, *116*(50), 25179–25185. doi:10.1073/pnas.1908771116
- Burgarella, C., Barnaud, A., Kane, N. A., Jankowski, F., Scarcelli, N., Billot, C., Vigouroux, Y., & Berthouly-Salazar, C. (2019). Adaptive introgression: an untapped evolutionary mechanism for crop adaptation. *Frontiers in Plant Science*, *10*, 4.
- Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M., & Keller, S. R. (2020). Genomic prediction of (mal)adaptation across current and future climatic landscapes. *Annual Review of Ecology, Evolution, and Systematics*, *51*(1), 245–269. doi:10.1146/annurev-ecolsys-020720-042553
- Carlson, S. M., Cunningham, C. J., & Westley, P. A. (2014). Evolutionary rescue in a changing world. *Trends in Ecology & Evolution*, *29*(9), 521–530.
- Carrer, M. & Urbinati, C. (2004). Age-dependent tree-ring growth responses to climate in larch decidua and pinus cembra. *Ecology*, *85*(3), 730–740. doi:10.1890/02-0478



- Carter, K. (1996). Provenance tests as indicators of growth response to climate change in 10 north temperate tree species. *Canadian Journal of Forest Research*, *26*, 1089–1095.
- Caswell, H. (2000). *Matrix population models*. Sinauer Sunderland, MA.
- Caswell, H., de Vries, C., Hartemink, N., Roth, G., & van Daalen, S. F. (2018). Age  $\times$  stage-classified demographic analysis: a comprehensive approach. *Ecological monographs*, *88*(4), 560–584. doi:<https://doi.org/10.1002/ecm.1306>
- Chan, W. Y., Hoffmann, A. A., & van Oppen, M. J. (2019). Hybridization as a conservation management tool. *Conservation Letters*, *12*(5), e12652.
- Chen, I. C., Hill, J. K., Ohlemuller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate warming. *Science*, *333*(6045), 1024–1026. doi:10.1126/science.1206432
- Chen, Z., Grossfurthner, L., Loxterman, J. L., Masingale, J., Richardson, B. A., Seaborn, T., Smith, B., Waits, L. P., & Narum, S. R. (2022). Applying genomics in assisted migration under climate change: framework, empirical applications, and case studies. *Evolutionary Applications*, *15*(1), 3–21.
- Chevin, L. (2015). Evolution of adult size depends on genetic variance in growth trajectories: a comment on analyses of evolutionary dynamics using integral projection models. *Methods in Ecology and Evolution*, *6*, 981–986. doi:10.1111/2041-210X.12389
- Chevin, L., Lande, R., & Mace, G. M. (2010). Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS biology*, *8*(4), e1000357.
- Childs, D., Coulson, T., Pemberton, J., Clutton-Brock, T., & Rees, M. (2011). Predicting trait values and measuring selection in complex life histories: reproductive allocation decisions in soay sheep. *Ecology Letters*, *14*(10), 985–992.
- Christie, M. R. & Searle, C. L. (2018). Evolutionary rescue in a host–pathogen system results in coexistence not clearance. *Evolutionary Applications*, *11*(5), 681–693.
- Colas, B., Kirchner, F., Riba, M., Olivieri, I., Mignot, A., Imbert, E., Beltrame, C., Carbonell, D., & Fréville, H. (2008). Restoration demography: a 10-year demographic comparison between introduced and natural populations of endemic *centaurea corymbosa* (asteraceae). *Journal of Applied Ecology*, *45*(5), 1468–1476.
- Colas, B., Olivieri, I., & Riba, M. (1997). *Centaurea corymbosa*, a cliff-dwelling species tottering on the brink of extinction: a demographic and genetic study. *Proceedings of the National Academy of Sciences*, *94*(7), 3471–3476.

- Colas, B. & Sarrazin, F. (2017). Transloc database. *BBEES*. Retrieved from <http://translocations.in2p3.fr/>
- Cortés, A. J., Restrepo-Montoya, M., & Bedoya-Canas, L. E. (2020). Modern strategies to assess and breed forest tree adaptation to changing climate. *Frontiers in Plant Science*, *11*, 583323.
- Cotto, O. & Chevin, L.-M. (2020). Fluctuations in lifetime selection in an autocorrelated environment. *Theoretical Population Biology*, *134*, 119–128. doi:10.1016/j.tpb.2020.03.002
- Cotto, O., Sandell, L., Chevin, L., & Ronce, O. (2019). Maladaptive shifts in life history in a changing environment. *The American Naturalist*, *194*(4), 558–573.
- Cotto, O., Wessely, J., Georges, D., Klonner, G., Schmid, M., Dullinger, S., Thuiller, W., & Guillaume, F. (2017). A dynamic eco-evolutionary model predicts slow response of alpine plants to climate warming. *Nature Communications*, *8*. doi:10.1038/ncomms15399
- Coulson, T., Kendall, B. E., Barthold, J., Plard, F., Schindler, S., Ozgul, A., & Gaillard, J.-M. (2017). Modeling adaptive and nonadaptive responses of populations to environmental change. *The American Naturalist*, *190*(3), 313–336.
- Coulson, T., MacNulty, D. R., Stahler, D. R., VonHoldt, B., Wayne, R. K., & Smith, D. W. (2011). Modeling effects of environmental change on wolf population dynamics, trait evolution, and life history. *Science*, *334*(6060), 1275–1278.
- Coulson, T., Potter, T., & Felmy, A. (2021). Predicting evolution over multiple generations in deteriorating environments using evolutionarily explicit integral projection models. *Evolutionary Applications*, *14*(10), 2490–2501.
- Crespi, B. J. (2000). The evolution of maladaptation. *Heredity*, *84*, 623–629. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2540.2000.00746.x/full>
- Davis, K. T., Dobrowski, S. Z., Higuera, P. E., Holden, Z. A., Veblen, T. T., Rother, M. T., Parks, S. A., Sala, A., & Maneta, M. P. (2019). Wildfires and climate change push low-elevation forests across a critical climate threshold for tree regeneration. *Proceedings of the National Academy of Sciences*, *116*(13), 6193–6198.
- de Vries, C. & Caswell, H. (2019). Stage-structured evolutionary demography: linking life histories, population genetics, and ecological dynamics. *The American Naturalist*, *193*(4), 545–559.
- DeFilippo, L. B., McManus, L. C., Schindler, D. E., Pinsky, M. L., Colton, M. A., Fox, H. E., Tekwa, E., Palumbi, S. R., Essington, T. E., & Webster, M. M. (2022). Assessing the

- potential for demographic restoration and assisted evolution to build climate resilience in coral reefs. *Ecological applications*, 32(7), e2650. doi:10.1002/eap.2650
- Depardieu, C., Girardin, M. P., Nadeau, S., Lenz, P., Bousquet, J., & Isabel, N. (2020). Adaptive genetic variation to drought in a widely distributed conifer suggests a potential for increasing forest resilience in a drying climate. *New Phytologist*, 227(2), 427–439. doi:10.1111/nph.16551
- Diallo, M., Ollier, S., Mayeur, A., Fernández-Manjarrés, J., García-Fernández, A., Iriondo, J. M., Vaissière, A. C., & Colas, B. (2021). Plant translocations in Europe and the Mediterranean: geographical and climatic directions and distances from source to host sites. *Journal of Ecology*, 109, 2296–2308. doi:10.1111/1365-2745.13609
- Diamond, S. E. & Martin, R. A. (2021a). Buying time: plasticity and population persistence. In *Phenotypic plasticity & evolution* (pp. 185–209). CRC Press.
- Diamond, S. E. & Martin, R. A. (2021b). Buying time: plasticity and population persistence. In D. W. Pfennig (Ed.), *Phenotypic Plasticity & Evolution: Causes, Consequences, Controversies* (pp. 185–209). CRC Press.
- Dinnage, R., Sarre, S. D., Duncan, R. P., Dickman, C. R., Edwards, S. V., Greenville, A. C., Wardle, G. M., & Gruber, B. (2023). Slimr: an r package for tailor-made integrations of data in population genomic simulations over space and time. *Molecular Ecology Resources*, e13916.
- Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., & Collen, B. (2014). Defaunation in the anthropocene. *Science*, 345(6195), 401–406.
- Doak, D. F., Waddle, E., Langendorf, R. E., Louthan, A. M., Isabelle Chardon, N., Dibner, R. R., Keinath, D. A., Lombardi, E., Steenbock, C., Shriver, R. K. et al. (2021). A critical comparison of integral projection and matrix projection models for demographic analysis. *Ecological Monographs*, 91(2), e01447.
- Du, J., Li, K., He, Z., Chen, L., Zhu, X., & Lin, P. (2019). Age-mediation of tree-growth responses to experimental warming in the northeastern Tibetan Plateau. *Ecology and Evolution*, 9(4), 2242–2254. doi:10.1002/ece3.4920. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ece3.4920>
- Duchesne, L., Houle, D., Ouimet, R., Caldwell, L., Gloor, M., & Brienens, R. (2019). Large apparent growth increases in boreal forests inferred from tree-rings are an artefact of sampling biases. *Scientific Reports*, 9(1), 6832. doi:10.1038/s41598-019-43243-1

- Ducrettet, J. & Imbert, E. (2023). *Rapport final projet renpop. appel à projet fonctionnalité des milieux terrestres et efficacité des pratiques de gestion*. OFB-labex Cemeb.
- Duffy, K., Gouhier, T. C., & Ganguly, A. R. (2022). Climate-mediated shifts in temperature fluctuations promote extinction risk. *Nature Climate Change*, *12*, 1037–1044. doi:10.1038/s41558-022-01490-7
- Durrett, R. & Schweinsberg, J. (2004). Approximating selective sweeps. *Theoretical population biology*, *66*(2), 129–138.
- Easterling, M. R., Ellner, S. P., & Dixon, P. M. (2000). Size-specific sensitivity: applying a new structured population model. *Ecology*, *81*(3), 694–708.
- Ellner, S. P., Childs, D. Z., Rees, M. et al. (2016). Data-driven modelling of structured populations. *A practical guide to the Integral Projection Model*. Cham: Springer.
- Ellner, S. P. & Rees, M. (2006). Integral projection models for species with complex demography. *The American Naturalist*, *167*(3), 410–428.
- Eriksen, F. D., Erlichman, A., & Ronce, O. (2021). Investigating sensitivity of different life stages to climate change in terrestrial plant populations. *Master 1 Report*.
- Esperon Rodriguez, M., Tjoelker, M. G., Lenoir, J., Baumgartner, J. B., Beaumont, L. J., Nipperess, D. A., Power, S. A., Richard, B., Rymer, P. D., & Gallagher, R. V. (2022). Climate change increases global risk to urban forests. *Nature Climate Change*, *12*, 950–955. doi:10.1038/S41558-022-01465-8
- Ettinger, A. K. & HilleRisLambers, J. (2013). Climate isn't everything: competitive interactions and variation by life stage will also affect range shifts in a warming world. *American Journal of Botany*, *100*, 1344–1355. doi:10.3732/AJB.1200489
- Filbee-Dexter, K. & Smajdor, A. (2019). Ethics of assisted evolution in marine conservation. *Frontiers in Marine Science*, *6*, 20.
- Fisher, A., Maurice, S., & Imbert, E. (2023). Genotypic variation of drought resistance traits in the rare endemic plant species *Centaurea corymbosa*. *Master 1 Report*.
- Fisher, R. A. (1918). The correlation between relatives on the supposition of mendelian inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, *52*(2), 399–433.
- Forsythe, A. B., Day, T., & Nelson, W. A. (2021). Demystifying individual heterogeneity. *Ecology Letters*, *24*(10), 2282–2297.

- Forzieri, G., Dakos, V., McDowell, N. G., Ramdane, A., & Cescatti, A. (2022). Emerging signals of declining forest resilience under climate change. *Nature*, *608*, 534–539. doi:10.1038/s41586-022-04959-9
- Fox, R. J., Donelson, J. M., Schunter, C., Ravasi, T., & Gaitán-Espitia, J. D. (2019). Beyond buying time: the role of plasticity in phenotypic adaptation to rapid environmental change. *Philosophical Transactions of the Royal Society B*, *374*. doi:10.1098/RSTB.2018.0174
- Fraga, H., Moriondo, M., Leolini, L., & Santos, J. A. (2020). Mediterranean olive orchards under climate change: a review of future impacts and adaptation strategies. *Agronomy*, *11*(1), 56.
- Frank, A., Howe, G. T., Sperisen, C., Brang, P., Clair, J. B. S., Schmatz, D. R., & Heiri, C. (2017). Risk of genetic maladaptation due to climate change in three major European tree species. *Global Change Biology*, *23*(12), 5358–5371. doi:10.1111/gcb.13802. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.13802>
- Frankham, R., Ballou, J. D., Eldridge, M. D., Lacy, R. C., Ralls, K., Dudash, M. R., & Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology*, *25*(3), 465–475.
- Freeman, B. G., Scholer, M. N., Ruiz-Gutierrez, V., & Fitzpatrick, J. W. (2018). Climate change causes upslope shifts and mountaintop extirpations in a tropical bird community. *Proceedings of the National Academy of Sciences*, *115*(47), 11982–11987. doi:10.1073/pnas.1804224115
- Fréville, H., Colas, B., Riba, M., Caswell, H., Mignot, A., Imbert, E., & Olivieri, I. (2004). Spatial and temporal demographic variability in the endemic plant species *Centaurea corymbosa* (Asteraceae). *Ecology*, *85*(3), 694–703. doi:10.1890/03-0119
- Gauli, A., Neupane, P. R., Mundhenk, P., & Köhl, M. (2022). Effect of climate change on the growth of tree species: dendroclimatological analysis. *Forests*, *13*(4), 496. doi:10.3390/f13040496
- Gibson, I., Welsh, A. B., Welsh, S. A., & Cincotta, D. A. (2019). Genetic swamping and possible species collapse: tracking introgression between the native candy darter and introduced variegated darter. *Conservation Genetics*, *20*, 287–298.
- Gienapp, P., Teplitsky, C., Alho, J., Mills, J., & Merilä, J. (2008). Climate change and evolution: disentangling environmental and genetic responses. *Molecular ecology*, *17*(1), 167–178.

- Giencke, L. M., Denhof, R. C., Kirkman, L. K., Stuber, O. S., & Brantley, S. T. (2018). Seed sourcing for longleaf pine ground cover restoration: using plant performance to assess seed transfer zones and home-site advantage. *Restoration Ecology*, *26*, 1127–1136. doi:10.1111/rec.12673
- Gignoux-Wolfsohn, S. A., Pinsky, M. L., Kerwin, K., Herzog, C., Hall, M., Bennett, A. B., Fefferman, N. H., & Maslo, B. (2021). Genomic signatures of selection in bats surviving white-nose syndrome. *Molecular Ecology*, *30*(22), 5643–5657.
- Girard, Q., Ducouso, A., de Gramont, C. B., Louvet, J. M., Reynet, P., Musch, B., & Kremer, A. (2022). Provenance variation and seed sourcing for sessile oak (*quercus petraea* (matt.) Liebl.) in France. *Annals of Forest Science*, *79*(1), 1–16.
- Girardin, M. P., Isabel, N., Guo, X. J., Lamothe, M., Duchesne, I., & Lenz, P. (2021). Annual aboveground carbon uptake enhancements from assisted gene flow in boreal black spruce forests are not long-lasting. *Nature Communications*, *12*(1), 1169.
- Gomulkiewicz, R. & Holt, R. D. (1995). When does evolution by natural selection prevent extinction? *Evolution*, *49*(1), 201–207.
- Gomulkiewicz, R. & Houle, D. (2009). Demographic and genetic constraints on evolution. *The American Naturalist*, *174*(6), E218–E229.
- Gougherty, A. V., Keller, S. R., & Fitzpatrick, M. C. (2021). Maladaptation, migration and extirpation fuel climate change risk in a forest tree species. *Nature Climate Change*, *11*, 166–171. doi:10.1038/s41558-020-00968-6
- Grummer, J. A., Booker, T. R., Matthey-Doret, R., Nietlisbach, P., Thomaz, A. T., & Whitlock, M. C. (2022). The immediate costs and long-term benefits of assisted gene flow in large populations. *Conservation Biology*, *36*(4), e13911. doi:10.1111/cobi.13911
- Hadjou Belaid, A., Maurice, S., Fréville, H., Carbonell, D., & Imbert, E. (2018). Predicting population viability of the narrow endemic Mediterranean plant *Centaurea corymbosa* under climate change. *Biological Conservation*, *223*, 19–33. doi:10.1016/j.biocon.2018.04.019
- Hadjou-Belaid, A. (2018). *Démographie et réponses adaptatives des populations végétales aux changements environnementaux* (Doctoral dissertation, Montpellier).
- Hadjou-Belaid, A., Maurice, S., Fréville, H., & Imbert, E. ((in prep)). Variation in size at flowering in response to climate change is not enough for the monocarpic plant *centaurea corymbosa* to escape extinction.

- Hagedorn, M., Page, C. A., O'Neil, K. L., Flores, D. M., Tichy, L., Conn, T., Chamberland, V. F., Lager, C., Zuchowicz, N., Lohr, K. et al. (2021). Assisted gene flow using cryopreserved sperm in critically endangered coral. *Proceedings of the National Academy of Sciences*, *118*(38), e2110559118.
- Haldane, J. B. S. (1919). The combination of linkage values, and the calculation of distances between the loci of linked factors. *Journal of Genetics*, *8*(29), 299–309.
- Haldane, J. B. S. (1927). A mathematical theory of natural and artificial selection, part v: selection and mutation. *Mathematical Proceedings of the Cambridge Philosophical Society*, *23*(7), 838–844. doi:10.1017/S0305004100015644
- Haller, B. C. & Messer, P. W. (2017). Slim 2: flexible, interactive forward genetic simulations. *Molecular biology and evolution*, *34*(1), 230–240.
- Hamilton, J. A. & Miller, J. M. (2016). Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conservation Biology*, *30*(1), 33–41.
- Hanewinkel, M., Cullmann, D. A., Schelhaas, M. J., Nabuurs, G. J., & Zimmermann, N. E. (2013). Climate change may cause severe loss in the economic value of European forest land. *Nature Climate Change*, *3*. doi:10.1038/nclimate1687
- Harcombe, P. (1987). Tree life tables. *Bioscience*, *37*(8), 557–568.
- Harris, K., Zhang, Y., & Nielsen, R. (2019). Genetic rescue and the maintenance of native ancestry. *Conservation Genetics*, *20*, 59–64.
- Harrisson, K. A., Pavlova, A., Gonçalves da Silva, A., Rose, R., Bull, J. K., Lancaster, M. L., Murray, N., Quin, B., Menkhorst, P., Magrath, M. J. et al. (2016). Scope for genetic rescue of an endangered subspecies through re-establishing natural gene flow with another subspecies. *Molecular Ecology*, *25*(6), 1242–1258.
- Hartfield, M. & Otto, S. P. (2011). Recombination and hitchhiking of deleterious alleles. *Evolution*, *65*(9), 2421–2434.
- Hartmann, H., Bastos, A., Das, A. J., Esquivel-Muelbert, A., Hammond, W. M., Martínez-Vilalta, J., McDowell, N. G., Powers, J. S., Pugh, T. A., Ruthrof, K. X., & Allen, C. D. (2022). Climate change risks to global forest health: Emergence of unexpected events of elevated tree mortality worldwide. *Annual Review of Plant Biology*, *73*(1), 673–702. doi:10.1146/annurev-arplant-102820-012804
- Harvey, B. D. & Bergeron, Y. (1989). Site patterns of natural regeneration following clear-cutting in northwestern quebec. *Canadian Journal of Forest Research*, *19*(11), 1458–1469.

- Hedrick, P. W. (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular ecology*, *22*(18), 4606–4618.
- Hedrick, P. W., Peterson, R. O., Vucetich, L. M., Adams, J. R., & Vucetich, J. A. (2014). Genetic rescue in isle royale wolves: genetic analysis and the collapse of the population. *Conservation genetics*, *15*, 1111–1121.
- Heiland, L., Kunstler, G., Ruiz-Benito, P., Buras, A., Dahlgren, J., & Hülsmann, L. (2022). Divergent occurrences of juvenile and adult trees are explained by both environmental change and ontogenetic effects. *Ecography*, *2022*, e06042. doi:10.1111/ECOG.06042
- Hellmann, J. J., Byers, J. E., Bierwagen, B. G., & Dukes, J. S. (2008). Five potential consequences of climate change for invasive species. *Conservation Biology*, *22*, 534–543. doi:10.1111/J.1523-1739.2008.00951.X
- Hohenlohe, P. A., McCallum, H. I., Jones, M. E., Lawrance, M. F., Hamede, R. K., & Storfer, A. (2019). Conserving adaptive potential: lessons from tasmanian devils and their transmissible cancer. *Conservation Genetics*, *20*, 81–87.
- Housset, J. M., Nadeau, S., Isabel, N., Depardieu, C., Duchesne, I., Lenz, P., & Girardin, M. P. (2018). Tree rings provide a new class of phenotypes for genetic associations that foster insights into adaptation of conifers to climate change. *New Phytologist*, *218*(2), 630–645. doi:10.1111/nph.14968
- Howard-McCombe, J., Jamieson, A., Carmagnini, A., Russo, I.-R. M., Ghazali, M., Campbell, R., Driscoll, C., Murphy, W. J., Nowak, C., O'Connor, T. et al. (2023). Genetic swamping of the critically endangered scottish wildcat was recent and accelerated by disease. *Current Biology*, *33*(21), 4761–4769.
- Hunter, M. L. (2007). Climate change and moving species: furthering the debate on assisted colonization. *Conservation Biology*, *21*(5), 1356–1358. doi:10.1111/j.1523-1739.2007.00780.x
- IPCC. (2021). Summary for Policymaker. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. doi:10.1017/9781009157896.001



- Isabel, N., Holliday, J. A., & Aitken, S. N. (2020). Forest genomics: advancing climate adaptation, forest health, productivity, and conservation. *Evolutionary Applications*, *13*, 3–10. doi:10.1111/EVA.12902
- Jackson, S. T., Betancourt, J. L., Booth, R. K., & Gray, S. T. (2009). Ecology and the ratchet of events: climate variability, niche dimensions, and species distributions. *Proceedings of the National Academy of Sciences*, *106*(supplement\_2), 19685–19692.
- Janeiro, M. J., Coltman, D. W., Festa-Bianchet, M., Pelletier, F., & Morrissey, M. B. (2017). Towards robust evolutionary inference with integral projection models. *Journal of Evolutionary Biology*, *30*(2), 270–288.
- Johnson, W. E., Onorato, D. P., Roelke, M. E., Land, E. D., Cunningham, M., Belden, R. C., McBride, R., Jansen, D., Lotz, M., Shindle, D. et al. (2010). Genetic restoration of the florida panther. *Science*, *329*(5999), 1641–1645.
- Kelly, E. & Phillips, B. L. (2016). Targeted gene flow for conservation. *Conservation biology*, *30*(2), 259–267.
- Kelly, E. & Phillips, B. L. (2019a). How many and when? optimising targeted gene flow for a step change in the environment. *Ecology Letters*, *22*(3), 447–457.
- Kelly, E. & Phillips, B. L. (2019b). Targeted gene flow and rapid adaptation in an endangered marsupial. *Conservation Biology*, *33*(1), 112–121.
- Kirchner, F., Robert, A., & Colas, B. (2006). Modelling the dynamics of introduced populations in the narrow-endemic *Centaurea corymbosa*: A demo-genetic integration. *Journal of Applied Ecology*, *43*(5), 1011–1021. doi:10.1111/j.1365-2664.2006.01179.x
- Kremer, A., Ronce, O., Robledo-Arnuncio, J. J., Guillaume, F., Bohrer, G., Nathan, R., Bridle, J. R., Gomulkiewicz, R., Klein, E. K., Ritland, K., Kuparinen, A., Gerber, S., & Schueler, S. (2012). Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*, *15*, 378–392. doi:10.1111/j.1461-0248.2012.01746.x
- Kueppers, L. M., Conlisk, E., Castanha, C., Moyes, A. B., Germino, M. J., de Valpine, P., Torn, M. S., & Mitton, J. B. (2017). Warming and provenance limit tree recruitment across and beyond the elevation range of subalpine forest. *Global Change Biology*, *23*, 2383–2395. doi:10.1111/gcb.13561
- Kunstler, G., Guyennon, A., Ratcliffe, S., Rüger, N., Ruiz-Benito, P., Childs, D. Z., Dahlgren, J., Lehtonen, A., Thuiller, W., Wirth, C. et al. (2021). Demographic performance of European

- tree species at their hot and cold climatic edges. *Journal of Ecology*, *109*(2), 1041–1054. doi:10.1111/1365-2745.13533
- Kuparinen, A. & Uusi-Heikkilä, S. (2020). Atlantic cod recovery from the Allee effect zone: contrasting ecological and evolutionary rescue. *Fish and Fisheries*, *21*(5), 916–926. doi:10.1111/faf.12470
- Kuss, P., Rees, M., Ægisdóttir, H. H., Ellner, S. P., & Stöcklin, J. (2008). Evolutionary demography of long-lived monocarpic perennials: a time-lagged integral projection model. *Journal of Ecology*, *96*(4), 821–832.
- Lande, R. & Shannon, S. (1996). The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, 434–437.
- Latreille, A., Davi, H., Huard, F., & Pichot, C. (2017). Variability of the climate-radial growth relationship among *Abies alba* trees and populations along altitudinal gradients. *Forest Ecology and Management*, *396*, 150–159. doi:https://doi.org/10.1016/j.foreco.2017.04.012
- Leimu, R., Vergeer, P., Angeloni, F., & Ouborg, N. J. (2010). Habitat fragmentation, climate change, and inbreeding in plants. *Annals of the New York Academy of Sciences*, *1195*, 84–98. doi:10.1111/J.1749-6632.2010.05450.X
- Leites, L. & Benito Garzón, M. (2023). Forest tree species adaptation to climate across biomes: Building on the legacy of ecological genetics to anticipate responses to climate change. *Global Change Biology*, *29*(17), 4711–4730. doi:https://doi.org/10.1111/gcb.16711
- Lenoir, J., Gégout, J. C., Guisan, A., Vittoz, P., Wohlgemuth, T., Zimmermann, N. E., Dullinger, S., Pauli, H., Willner, W., & Svenning, J. C. (2010). Going against the flow: potential mechanisms for unexpected downslope range shifts in a warming climate. *Ecography*, *33*(2), 295–303. doi:10.1111/J.1600-0587.2010.06279.X
- Lenoir, J. & Svenning, J. C. (2015). Climate-related range shifts—a global multidimensional synthesis and new research directions. *Ecography*, *38*(1), 15–28. doi:10.1111/ECOG.00967
- Lewis, J. L. & Sheppard, S. R. (2005). Ancient values, new challenges: indigenous spiritual perceptions of landscapes and forest management. *Society and Natural Resources*, *18*(10), 907–920.
- Liddell, E., Sunnucks, P., & Cook, C. N. (2021). To mix or not to mix gene pools for threatened species management? few studies use genetic data to examine the risks of both actions, but failing to do so leads disproportionately to recommendations for separate management. *Biological Conservation*, *256*, 109072.

- Lortie, C. J. & Hierro, J. L. (2022). A synthesis of local adaptation to climate through reciprocal common gardens. *Journal of Ecology*, *110*, 1015–1021. doi:10.1111/1365-2745.13664
- Lucena-Perez, M., Paijmans, J. L., Nocete, F., Nadal, J., Detry, C., Dalén, L., Hofreiter, M., Barlow, A., & Godoy, J. A. (2024). Recent increase in species-wide diversity after inter-species introgression in the highly endangered iberian lynx. *Nature Ecology & Evolution*, 1–11.
- Lynch, M. & Lande, R. (1993). Evolution and extinction in response to environmental change. *Biotic interactions and global change*, 234–250.
- Madsen, T., Ujvari, B., & Olsson, M. (2004). Novel genes continue to enhance population growth in adders (*Vipera berus*). *Biological Conservation*, *120*(1), 145–147.
- Maris, V., Huneman, P., Coreau, A., Kéfi, S., Pradel, R., & Devictor, V. (2018). Prediction in ecology: promises, obstacles and clarifications. *Oikos*, *127*(2), 171–183.
- Maroso, F., Padovani, G., Muñoz Mora, V. H., Giannelli, F., Trucchi, E., & Bertorelle, G. (2023). Fitness consequences and ancestry loss in the apennine brown bear after a simulated genetic rescue intervention. *Conservation Biology*, *37*(6), e14133.
- Marquis, B., Bergeron, Y., Simard, M., & Tremblay, F. (2020). Growing-season frost is a better predictor of tree growth than mean annual temperature in boreal mixedwood forest plantations. *Global Change Biology*, *26*(11), 6537–6554. doi:10.1111/gcb.15327
- Marshall, D. J., Burgess, S. C., & Connallon, T. (2016). Global change, life-history complexity and the potential for evolutionary rescue. *Evolutionary Applications*, *9*(9), 1189–1201. doi:10.1111/eva.12396
- Marshall, D. J. & Connallon, T. (2022). Carry-over effects and fitness trade-offs in marine life histories: The costs of complexity for adaptation. *Evolutionary Applications*, *16*(2), 474–485. doi:10.1111/eva.13477
- Marshall, D. J. & Connallon, T. (2023). Carry-over effects and fitness trade-offs in marine life histories: the costs of complexity for adaptation. *Evolutionary Applications*, *16*(2), 474–485.
- Mašek, J., Tumajer, J., Rydval, M., Lange, J., & Treml, V. (2021). Age and size outperform topographic effects on growth-climate responses of trees in two central european coniferous forest types. *Dendrochronologia*, *68*, 125845. doi:10.1016/j.dendro.2021.125845
- Mátyás, C. (1996). Climatic adaptation of trees : rediscovering provenance tests. *Euphytica*, *92*, 45–54. doi:10.1007/BF00022827

- Maxwell, S. L., Fuller, R. A., Brooks, T. M., & Watson, J. E. (2016a). Biodiversity: the ravages of guns, nets and bulldozers. *Nature*, *536*(7615), 143–145.
- Maxwell, S. L., Fuller, R. A., Brooks, T. M., & Watson, J. E. (2016b). Biodiversity: The ravages of guns, nets and bulldozers. *Nature*, *536*, 143–145.
- Maynard Smith, J. & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetical research*, *23*(01), 23–35.
- McDonald, G., Appleby, M. W., Sime, H., Radford, J., & Hoffmann, A. A. (2021). Establishing a climate-ready revegetation trial in central victoria—a case study. *Ecological Management & Restoration*, *22*(3), 256–265.
- McLaughlin, B. C. & Zavaleta, E. S. (2012). Predicting species responses to climate change: demography and climate microrefugia in California valley oak (*Quercus lobata*). *Global Change Biology*, *18*(7), 2301–2312.
- Miao, B., Wang, Z., & Li, Y. (2017). Genomic analysis reveals hypoxia adaptation in the tibetan mastiff by introgression of the gray wolf from the tibetan plateau. *Molecular Biology and Evolution*, *34*(3), 734–743.
- Milesi, P., Berlin, M., Chen, J., Orsucci, M., Li, L., Jansson, G., Karlsson, B., & Lascoux, M. (2019). Assessing the potential for assisted gene flow using past introduction of Norway spruce in southern Sweden: Local adaptation and genetic basis of quantitative traits in trees. *Evolutionary Applications*, *12*, 1946–1959. doi:10.1111/eva.12855
- Mimura, M. & Aitken, S. N. (2010). Local adaptation at the range peripheries of sitka spruce. *Journal of evolutionary biology*, *23*(2), 249–258.
- Mitchell, J. F., Lowe, J., Wood, R. A., & Vellinga, M. (2006). Extreme events due to human-induced climate change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *364*(1845), 2117–2133.
- Muller, H. J. (1942). Isolating mechanisms, evolution, and temperature. In *Biol. symp.* (Vol. 6, p. 71).
- Munier, A., Hermanutz, L., Jacobs, J. D., & Lewis, K. (2010). The interacting effects of temperature, ground disturbance, and herbivory on seedling establishment: implications for treeline advance with climate warming. *Plant Ecology*, *210*, 19–30. doi:10.1007/s11258-010-9724-y
- Olonscheck, D., Schurer, A. P., Lücke, L., & Hegerl, G. C. (2021). Large-scale emergence of regional changes in year-to-year temperature variability by the end of the 21st century. *Nature Communications*, *12*(1), 7237.

- O'Neill, G. A. & Gómez-Pineda, E. (2021). Local was best: sourcing tree seed for future climates. *Canadian Journal of Forest Research*, 51(10), 1432–1439.
- O'Neill, G., Ukrainetz, N., Carlson, M., Cartwright, C., Jaquish, B., King, J., Krakowski, J., Russell, J., Stoehr, M., Xie, C., & a Yanchuk. (2008). *Assisted migration to address climate change in British Columbia: recommendations for interim seed transfer standards* (tech. rep. No. 048). Ministry of Forests and Range, Forest Science Program, Research Branch, British Columbia. Victoria, B.C. Retrieved from [www.for.gov.bc.ca/hfd/pubs/Docs/Tr/Tr048.htm](http://www.for.gov.bc.ca/hfd/pubs/Docs/Tr/Tr048.htm)
- O'Neill, G., Wang, T., Ukrainetz, N., Charleson, L., McAuley, L., Yanchuk, A., Zedel, S. et al. (2017). *A proposed climate-based seed transfer system for British Columbia* (tech. rep. No. 099). Ministry of Forests, Lands and Natural Resource Operations, British Columbia. Victoria, B.C.
- Orive, M. E., Barfield, M., Fernandez, C., & Holt, R. D. (2017). Effects of clonal reproduction on evolutionary lag and evolutionary rescue. *The American Naturalist*, 190(4), 469–490.
- Orr, H. A. & Unckless, R. L. (2008). Population extinction and the genetics of adaptation. *The American Naturalist*, 172(2), 160–169.
- Orr, H. A. & Unckless, R. L. (2014). The population genetics of evolutionary rescue. *PLoS genetics*, 10(8), e1004551.
- Otto, S. P. & Whitlock, M. C. (1997). The probability of fixation in populations of changing size. *Genetics*, 146(2), 723–733.
- Ozgul, A., Childs, D. Z., Oli, M. K., Armitage, K. B., Blumstein, D. T., Olson, L. E., Tuljapurkar, S., & Coulson, T. (2010). Coupled dynamics of body mass and population growth in response to environmental change. *Nature*, 466(7305), 482–485.
- Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, 637–669.
- Pavlova, A., Beheregaray, L. B., Coleman, R., Gilligan, D., Harrison, K. A., Ingram, B. A., Kearns, J., Lamb, A. M., Lintermans, M., Lyon, J. et al. (2017). Severe consequences of habitat fragmentation on genetic diversity of an endangered Australian freshwater fish: a call for assisted gene flow. *Evolutionary Applications*, 10(6), 531–550.
- Pedlar, J. H., McKenney, D. W., & Lu, P. (2021a). Critical seed transfer distances for selected tree species in eastern North America. *Journal of Ecology*, 109(6), 2271–2283. doi:10.1111/1365-2745.13605

- Pedlar, J. H., McKenney, D. W., Lu, P., & Thomson, A. (2021b). Response of northern populations of black spruce and jack pine to southward seed transfers: implications for climate change. *Atmosphere*, *12*(10), 1363. doi:10.3390/atmos12101363
- Pelletier, F., Garant, D., & Hendry, A. P. (2009). *Eco-evolutionary dynamics*. The Royal Society London.
- Peterson St-Laurent, G., Hagerman, S., & Kozak, R. (2018). What risks matter? public views about assisted migration and other climate-adaptive reforestation strategies. *Climatic change*, *151*, 573–587. doi:10.1007/s10584-018-2310-3
- Pierson, J. C., Beissinger, S. R., Bragg, J. G., Coates, D. J., Oostermeijer, J. G. B., Sunnucks, P., Schumaker, N. H., Trotter, M. V., & Young, A. G. (2015). Incorporating evolutionary processes into population viability models. *Conservation Biology*, *29*(3), 755–764.
- Pimm, S. L. (2000). Against triage. *Science*, *289*(5488), 2289–2289.
- Pimm, S. L., Dollar, L., & Bass Jr, O. L. (2006a). The genetic rescue of the florida panther. *Animal Conservation*, *9*(2), 115–122.
- Pimm, S. L., Raven, P., Peterson, A., Şekercioğlu, Ç. H., & Ehrlich, P. R. (2006b). Human impacts on the rates of recent, present, and future bird extinctions. *Proceedings of the National Academy of Sciences*, *103*(29), 10941–10946.
- Pompa-García, M. & Hadad, M. A. (2016). Sensitivity of pines in Mexico to temperature varies with age. *Atmósfera*, *29*(3), 209–219. doi:10.20937/ATM.2016.29.03.03
- Pregler, K. C., Obedzinski, M., Gilbert-Horvath, E. A., White, B., Carlson, S. M., & Garza, J. C. (2023). Assisted gene flow from outcrossing shows the potential for genetic rescue in an endangered salmon population. *Conservation Letters*, *16*(2), e12934.
- Prieto-Benítez, S., Morente-López, J., Rubio Teso, M. L., Lara-Romero, C., García-Fernández, A., Torres, E., & Iriondo, J. M. (2021). Evaluating assisted gene flow in marginal populations of a high mountain species. *Frontiers in Ecology and Evolution*, *9*, 638837.
- Puettmann, K. J., Wilson, S. M., Baker, S. C., Donoso, P. J., Drössler, L., Amente, G., Harvey, B. D., Knoke, T., Lu, Y., Nocentini, S., Putz, F. E., Yoshida, T., & Bauhus, J. (2015). Silvicultural alternatives to conventional even-aged forest management - what limits global adoption? *Forest Ecosystems*, *2*, 8. doi:10.1186/s40663-015-0031-x
- Quigley, K. M., Bay, L. K., & van Oppen, M. J. (2019). The active spread of adaptive variation for reef resilience. *Ecology and Evolution*, *9*, 11122–11135. doi:10.1002/ece3.5616

- R Core Team. (2022). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Rabier, R., Erlichman, A., Lesobre, L., & Robert, A. (2022). The necessity of considering founder kinships in conservation breeding programs. *Animal Conservation*, *25*(6), 759–770.
- Radchuk, V., Reed, T., Teplitsky, C., van de Pol, M., Charmantier, A., Hassall, C., Adamík, P., Adriaensen, F., Ahola, M. P., Arcese, P., Avilés, J. M., Balbontin, J., Berg, K. S., Borrás, A., Burthe, S., Clobert, J., Dehnhard, N., de Lope, F., Dhondt, A. A., Dingemanse, N. J., Doi, H., Eeva, T., Fickel, J., Filella, I., Fossøy, F., Goodenough, A. E., Hall, S. J., Hansson, B., Harris, M., Hasselquist, D., Hickler, T., Joshi, J., Kharouba, H., Martínez, J. G., Mihoub, J. B., Mills, J. A., Molina-Morales, M., Moksnes, A., Ozgul, A., Parejo, D., Pilard, P., Poisbleau, M., Rousset, F., Rödel, M. O., Scott, D., Senar, J. C., Stefanescu, C., Stokke, B. G., Kusano, T., Tarka, M., Tarwater, C. E., Thonicke, K., Thorley, J., Wilting, A., Tryjanowski, P., Merila, J., Sheldon, B. C., Møller, A. P., Matthysen, E., Janzen, F., Dobson, F. S., Visser, M. E., Beissinger, S. R., Courtiol, A., & Kramer-Schadt, S. (2019). Adaptive responses of animals to climate change are most likely insufficient. *Nature Communications*, *10*, 1–14. doi:10.1038/s41467-019-10924-4
- Ramula, S., Rees, M., & Buckley, Y. M. (2009). Integral projection models perform better for small demographic data sets than matrix population models: a case study of two perennial herbs. *Journal of Applied Ecology*, *46*(5), 1048–1053.
- Rees, M. & Ellner, S. P. (2016). Evolving integral projection models: evolutionary demography meets eco-evolutionary dynamics. *Methods in Ecology and Evolution*, *7*(2), 157–170.
- Rees, M. & Ellner, S. P. (2019). Why so variable: can genetic variance in flowering thresholds be maintained by fluctuating selection? *The American Naturalist*, *194*(1), E13–E29.
- Rehfeldt, G. E., Ying, C. C., Spittlehouse, D. L., & Hamilton Jr, D. A. (1999). Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecological monographs*, *69*(3), 375–407. doi:10.1890/0012-9615(1999)069[0375:GRTCIP]2.0.CO;2
- Rellstab, C., Dauphin, B., & Exposito-Alonso, M. (2021). Prospects and limitations of genomic offset in conservation management. *Evolutionary Applications*, *14*(5), 1202–1212.
- Rhymer, J. M. & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual review of ecology and systematics*, *27*(1), 83–109.
- Ricciardi, A. & Simberloff, D. (2009). Assisted colonization is not a viable conservation strategy. *Trends in Ecology and Evolution*, *24*, 248–253. doi:10.1016/j.tree.2008.12.006

- Risk, C., McKenney, D. W., Pedlar, J., & Lu, P. (2021). A compilation of North American tree provenance trials and relevant historical climate data for seven species. *Scientific Data*, 8, 29. doi:10.1038/s41597-021-00820-2
- Robert, A. (2024). Building references for nature conservation. *Conservation Biology*, e14202.
- Román-Palacios, C. & Wiens, J. J. (2020). Recent responses to climate change reveal the drivers of species extinction and survival. *Proceedings of the National Academy of Sciences*, 117(8), 4211–4217.
- Sáenz-Romero, C., Mendoza-Maya, E., Gómez-Pineda, E., Blanco-García, A., Endara-Agramont, A. R., Lindig-Cisneros, R., López-Upton, J., Trejo-Ramírez, O., Wehenkel, C., Cibrián-Tovar, D., Flores-López, C., Plascencia-González, A., Vargas-Hernández, J. J., Mendoza-Maya, E., Wehenkel, C., López-Upton, J., & Vargas-Hernández, J. (2020). Recent evidence of Mexican temperate forest decline and the need for ex situ conservation, assisted migration, and translocation of species ensembles as adaptive management to face projected climatic change impacts in a megadiverse country. *Canadian Journal of Forest Research*, 50(9), 843–854. doi:10.1139/cjfr-2019-0329
- Sáenz-Romero, C., O'Neill, G., Aitken, S. N., & Lindig-Cisneros, R. (2021). Assisted migration field tests in Canada and Mexico: lessons, limitations, and challenges. *Forests*, 12(1), 1–19. doi:10.3390/f12010009
- Salguero-Gómez, R., Jones, O. R., Archer, C. R., Buckley, Y. M., Che-Castaldo, J., Caswell, H., Hodgson, D., Scheuerlein, A., Conde, D. A., Brinks, E. et al. (2015). The compadre plant matrix database: an open online repository for plant demography. *Journal of Ecology*, 103(1), 202–218. doi:10.1111/1365-2745.12334
- Sarrazin, F. & Legendre, S. (2000). Demographic approach to releasing adults versus young in reintroductions. *Conservation Biology*, 14, 488–500. doi:10.1046/J.1523-1739.2000.97305.X
- Savolainen, O., Pyhajarvi, T., & Knürr, T. (2007). Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics*, 38, 595–619. doi:10.1146/annurev.ecolsys.38.091206.095646
- Schmid, B., Balvanera, P., Cardinale, B. J., Godbold, J., Pfisterer, A. B., Raffaelli, D., Solan, M., & Srivastava, D. S. (2009). Consequences of species loss for ecosystem functioning: meta-analyses of data from biodiversity experiments. *Biodiversity, ecosystem functioning, and human wellbeing: An ecological and economic perspective*, 14–29.



- Schmid, M., Paniw, M., Postuma, M., Ozgul, A., & Guillaume, F. (2022). A trade-off between robustness to environmental fluctuations and speed of evolution. *The American Naturalist*, *200*(1), E16–E35.
- Schmidtling, R. C. (1994). Use of provenance tests to predict response to climatic change: loblolly pine and Norway spruce. *Tree Physiology*, *14*, 805–817. Retrieved from <http://treephys.oxfordjournals.org/>
- 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:<https://doi.org/10.1016/j.foreco.2018.10.046>
- Seddon, J. & Schultz, B. (2020). Koala conservation in queensland, australia: a role for assisted gene flow for genetic rescue? *Conservation genetics in mammals: Integrative research using novel approaches*, 331–349.
- Seddon, P. J. (2010). From reintroduction to assisted colonization: moving along the conservation translocation spectrum. *Restoration Ecology*, *18*(6), 796–802.
- Shaffer, M. L. (1981). Minimum population sizes for species conservation. *BioScience*, *31*(2), 131–134.
- Shaw, R. G. & Etterson, J. R. (2012). Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. *New Phytologist*, *195*, 752–765. doi:10.1111/J.1469-8137.2012.04230.X
- Simmonds, E. G., Cole, E. F., Sheldon, B. C., & Coulson, T. (2020). Testing the effect of quantitative genetic inheritance in structured models on projections of population dynamics. *Oikos*, *129*(4), 559–571.
- Smart, A. S. & Phillips, B. L. (2023). Optimizing targeted gene flow to maximize local genetic diversity: when and how to act under various scenarios of environmental change. *Conservation Genetics*, *24*(6), 869–882.
- Sork, V., Aitken, S., Dyer, R., Eckert, A., Legendre, P., & Neale, D. (2013). Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*, *9*, 901–911.
- Srivastava, D. S., Coristine, L., Angert, A. L., Bontrager, M., Amundrud, S. L., Williams, J. L., Yeung, A. C., de Zwaan, D. R., Thompson, P. L., Aitken, S. N., Sunday, J. M., O'Connor, M. I., Whitton, J., Brown, N. E., MacLeod, C. D., Parfrey, L. W., Bernhardt, J. R.,

- Carrillo, J., Harley, C. D., Martone, P. T., Freeman, B. G., Tseng, M., & Donner, S. D. (2021). Wildcards in climate change biology. *Ecological Monographs*, *91*(4), e01471. doi:10.1002/ECM.1471
- St. Clair, J. B., Richardson, B. A., Stevenson-Molnar, N., Howe, G. T., Bower, A. D., Erickson, V. J., Ward, B., Bachelet, D., Kilkenny, F. F., & Wang, T. (2022). Seedlot selection tool and climate-smart restoration tool: Web-based tools for sourcing seed adapted to future climates. *Ecosphere*, *13*, e4089. doi:10.1002/ECS2.4089
- Stapley, J., Feulner, P. G., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160455.
- Stevens-Rumann, C. S., Kemp, K. B., Higuera, P. E., Harvey, B. J., Rother, M. T., Donato, D. C., Morgan, P., & Veblen, T. T. (2018). Evidence for declining forest resilience to wildfires under climate change. *Ecology Letters*, *21*, 243–252. doi:10.1111/ELE.12889
- Stubben, C. J. & Milligan, B. G. (2007). Estimating and analyzing demographic models using the popbio package in r. *Journal of Statistical Software*, *22*(11).
- Sturrock, R. N., Frankel, S. J., Brown, A. V., Hennon, P. E., Kliejunas, J. T., Lewis, K. J., Worrall, J. J., & Woods, A. J. (2011). Climate change and forest diseases. *Plant Pathology*, *60*, 133–149. doi:10.1111/J.1365-3059.2010.02406.X
- Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F. N., de Siqueira, M. F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, M. A., Townsend Peterson, A., Phillips, O. L., & Williams, S. E. (2004). Extinction risk from climate change. *Nature*, *427*(6970), 145–148. doi:10.1038/nature02121
- Thomson, A. M., Riddell, C. L., & Parker, W. H. (2009). Boreal forest provenance tests used to predict optimal growth and response to climate change: 2. Black spruce. *Canadian Journal of Forest Research*, *39*(1), 143–153. doi:10.1139/X08-167
- Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G. et al. (2016). Hybridization and extinction. *Evolutionary Applications*.
- Trails, L. W., Schindler, S., & Coulson, T. (2014). Demography, not inheritance, drives phenotypic change in hunted bighorn sheep. *Proceedings of the National Academy of Sciences*, *111*(36), 13223–13228.

- Trigger, D. & Mulcock, J. (2005). Forests as spiritually significant places: nature, culture and belonging 'in australia. *The Australian journal of anthropology*, 16(3), 306–320.
- Van Dyken, J. D. (2020). Evolutionary rescue from a wave of biological invasion. *The American Naturalist*, 195(1), 115–128.
- Vedder, D., Lens, L., Martin, C. A., Pellikka, P., Adhikari, H., Heiskanen, J., Engler, J. O., & Sarmiento Cabral, J. (2022). Hybridization may aid evolutionary rescue of an endangered east african passerine. *Evolutionary applications*, 15(7), 1177–1188.
- Vila, C., Sundqvist, A.-.-K., Flagstad, Ø., Seddon, J., Kojola, I., Casulli, A., Sand, H., Wabakken, P., Ellegren, H. et al. (2003). Rescue of a severely bottlenecked wolf (*canis lupus*) population by a single immigrant. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1510), 91–97.
- Vindenes, Y. & Langangen, Ø. (2015). Individual heterogeneity in life histories and eco-evolutionary dynamics. *Ecology letters*, 18(5), 417–432.
- Wadgymar, S. M., DeMarche, M. L., Josephs, E. B., Sheth, S. N., & Anderson, J. T. (2022). Local adaptation: causal agents of selection and adaptive trait divergence. *Annual Review of Ecology, Evolution, and Systematics*, 53, 87–111. doi:10.1146/annurev-ecolsys-012722-035231
- Wang, T., Hamann, A., Yanchuk, A., O'neill, G., & Aitken, S. (2006). Use of response functions in selecting lodgepole pine populations for future climates. *Global Change Biology*, 12(12), 2404–2416. doi:10.1111/j.1365-2486.2006.01271.x
- Wang, T., O'Neill, G. A., & Aitken, S. N. (2010). Integrating environmental and genetic effects to predict responses of tree populations to climate. *Ecological applications*, 20(1), 153–163.
- Weeks, A. R., Sgro, C. M., Young, A. G., Frankham, R., Mitchell, N. J., Miller, K. A., Byrne, M., Coates, D. J., Eldridge, M. D., Sunnucks, P., Breed, M. F., James, E. A., & Hoffmann, A. A. (2011). Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications*, 4(6), 709–725. doi:10.1111/j.1752-4571.2011.00192.x
- Weeks, A. R., Stoklosa, J., & Hoffmann, A. A. (2016). Conservation of genetic uniqueness of populations may increase extinction likelihood of endangered species: the case of australian mammals. *Frontiers in Zoology*, 13, 1–9.

- Whittet, R., Cavers, S., Cottrell, J., & Ennos, R. (2016). Seed sourcing for woodland creation in an era of uncertainty: an analysis of the options for great britain. *Forestry: An International Journal of Forest Research*, *90*(2), 163–173.
- Wolf, D. E., Takebayashi, N., & Rieseberg, L. H. (2001). Predicting the risk of extinction through hybridization. *Conservation Biology*, *15*(4), 1039–1053.
- Woodruff, D. S. & Gould, S. J. (1987). Fifty years of interspecific hybridization: genetics and morphometrics of a controlled experiment on the land snail cerion in the florida keys. *Evolution*, 1022–1045.
- Yeh, P. J. & Price, T. D. (2004). Adaptive phenotypic plasticity and the successful colonization of a novel environment. *The American Naturalist*, *164*(4), 531–542.
- Young, D. J., Blush, T. D., Landram, M., Wright, J. W., Latimer, A. M., & Safford, H. D. (2020). Assisted gene flow in the context of large-scale forest management in California, USA. *Ecosphere*, *11*(1), e03001. doi:<https://doi.org/10.1002/ecs2.3001>
- Yu, H., Wang, T., Skidmore, A., Heurich, M., & Bässler, C. (2023). How future climate and tree distribution changes shape the biodiversity of macrofungi across europe. *Diversity and Distributions*, *29*(5), 666–682.
- Zheng, Z., Wang, X., Li, M., Li, Y., Yang, Z., Wang, X., Pan, X., Gong, M., Zhang, Y., Guo, Y. et al. (2020). The origin of domestication genes in goats. *Science Advances*, *6*(21), eaaz5216.
- Zilko, J. P., Harley, D., Pavlova, A., & Sunnucks, P. (2021). Applying population viability analysis to inform genetic rescue that preserves locally unique genetic variation in a critically endangered mammal. *Diversity*, *13*(8), 382.
- Zuidema, P. A., Jongejans, E., Chien, P. D., During, H. J., & Schieving, F. (2010). Integral projection models for trees: a new parameterization method and a validation of model output. *Journal of Ecology*, *98*(2), 345–355.